SAPIENZA
UNIVERSITÀ DI ROMA

Search for new exotic particles decaying to photons at the LHC

Scuola di dottorato Vito Volterra

Dottorato di Ricerca in Fisica – XXVII Ciclo

Candidate

Livia Soffi

ID number 1154143

Thesis Advisors

Prof. Shahram Rahatlou

Dr. Daniele del Re

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

October 2014

Thesis defended on 27 January 2015
in front of a Board of Examiners composed by:

Prof. Anna Di Ciaccio (chairman)

Prof. Nadia Pastrone

Prof. Valeria Ferrari

Search for new exotic particles decaying to photons at the LHC
Ph.D. thesis. Sapienza – University of Rome

This thesis has been typeset by LaTeX and the Sapthesis class.

Author's email: liviasoffi@gmail.com

"...you can't connect the dots looking forward; you can only connect them looking backwards. So you have to trust that the dots will somehow connect in your future... This approach has never let me down, and it has made all the difference in my life."
Steve Jobs

To those who, in this life, connect the dots with me.

"...non e' possibile unire i puntini guardando avanti; potete solo unirli guardandovi all'indietro. Dovete quindi avere fiducia che, nel futuro, i puntini che ora vi paiono senza senso possano in qualche modo unirsi...Questo approccio non mi ha mai lasciato a terra, e ha fatto la differenza nella mia vita."
Steve Jobs

A quelli che, in questa vita, uniscono i puntini insieme a me.

Contents

Introduction

The Standard Model of elementary particles (SM) and their fundamental interactions provides a very elegant theoretical framework for the description of what matter is made of and how it holds together. The development of this fascinating theory is a long and involving tale which finds its origin at the beginning of the 19th century with the study of the atoms as the fundamental building blocks of matter and through the identification of the different elements present in nature.

During the last decades, a sequence of brilliant ideas and experiments have shown an excellent agreement between the theoretical expectation and experimental results for the Standard Model [1]. The success of such a model was recently crowned with the observation of the last elusive particle predicted but yet undiscovered: the Higgs boson. The hunt for the Higgs boson, started 40 years ago, found indeed its conclusion in July 2012, when both the ATLAS and the CMS collaborations at the CERN Large Hadron Collider (LHC) of Geneva announced the discovery of a new boson having a mass of 125 GeV [2], [3] with properties very similar to those of the Higgs boson predicted by the Standard Model. The new observed particle, the first truly elementary particle discovered since the Z-boson, is the first example of an elementary scalar field and a new form of matter in addition to fermions and vector gauge bosons.

Despite the new particle is looking more and more like the Standard Model Higgs boson, precise measurements of its interactions, decays and quantum properties are underway and will continue also in the future of the LHC. Indeed, an open question remains, whether is this the Higgs boson predicted by the Standard Model of particle physics, or possibly the lightest of several bosons predicted in some theories that go beyond the Standard Model itself.

The search for physics beyond the Standard Model (BSM) is nowadays mainly motivated by the fact that, despite its experimental success, the theory still has some drawbacks and unsolved problems that represent a strong conceptual indication for new physics in the high energy domain. One major problem of the Standard Model is that it does not include the gravity, one of the four fundamental forces. The model also fails to explain why the gravity is so much weaker than the electromagnetic or nuclear forces. In addition, the Standard Model only describes visible matter, that is the all matter we see around us on Earth as well as in stars and galaxies. But proofs tell us the universe contains about five times more "dark matter", a type of matter completely different from the ordinary matter. Among all the particles predicted by the Standard Model, none has the properties of dark matter [4]. Hence, it is clear that the Standard Model gives an incomplete picture of the content of the universe and new physics is needed to solve the problem.

There are several alternative theories to the Standard Model which try to solve such open issues. In these models, new physics, in terms of new particles and new interactions, is expected to be visible at the TeV energy scale and thus accessible at the LHC. In the Higgs sector, the presence of physics beyond the Standard Model can be performed studying unusual decays of the already observed scalar or looking for new heavy states in interesting final states.

Following the latter approach, this thesis describes a model independent search for new heavy scalars in the diphoton final state performed with the general-purpose detector of the Compact Muon Solenoid (CMS) experiment. The electromagnetic calorimeter (ECAL) of CMS allows for the identification of high energy photons, thanks respectively to its fine lateral segmentation and the optimal performance of measuring energy with a design resolution better than 1% for energies above 50 GeV [5]. The analysis is based on 19.7 fb^{-1} of data collected during 2012 at a centre-of-mass energy of 8 TeV. This search finds its motivations in Extended Higgs Sector theories [6] and in Models with Extra-Dimension [7], where, under specific assumptions, the production of heavy scalars which decay to photons is favoured. If present, the signal will appear as an enhancement in the number of event counts at a particular value of its reconstructed mass over a continuum diphoton background. In order to be sensitive to any possible signal, both the mass and natural width of the new particles are investigated over a wide range of energies.

The dissertation is organized as follows:

- A theoretical outline of the Standard Model together with the recent discovery of a new scalar and the unsolved problems of the theory are discussed in Chapter 1. A description of the present searches for physics beyond the Standard Model in the Extended Higgs sector is also provided here;
- The experimental signature of a heavy resonance decaying to two photons together with the background topologies are discussed in Chapter 2. Theoretical motivations for a search of new scalars in the diphoton channel are also given together with an overview of the existing diphoton searches at colliders;
- A brief overview of the most important features of the LHC and the CMS experiment is presented in Chapter 3;
- The photon reconstruction and the selection are discussed in Chapter 4;
- The analysis strategy, the event selection and categorization together with the estimation of the background and the signal parametrization are presented in Chapter 5. The fit technique and the method to establish exclusion limits, including the description of the systematic effects affecting the analysis are also described;
- The results of the analysis are presented in Chapter 6 together with their interpretation in different theoretical models. The results obtained extend significantly previous diphoton searches results in terms of mass and width ranges.

In addition to the high mass resonance search in the diphoton channel, which represents the main topic of this thesis and the work of my last two year of PhD,

another physics analysis is presented in Appendix A. This work, of which I have been one of the main authors, has been the main activity of my first year of PhD. This is a search for new long-lived particles which decay to photons in the CMS detector. The results of this analysis, interpreted in a specific Supersymmetry scenario and published in 2013 in [8], are the most stringent results on long-lived neutralinos to date.

Chapter 1

Beyond the Standard Model Physics at the LHC

This Chapter starts with a brief introduction to the Standard Model of particle physics in Section 1.1. The recent discovery of the new boson with properties similar to those predicted for the Higgs boson associated to the Higgs mechanism of the SM is described in Section 1.2. Although all measurements of the recently observed 125 GeV boson to date, indicate compatibility with the SM Higgs boson, the associated uncertainties are still large, and the possibility for non-SM properties remains. Any significant deviation from SM predictions for Higgs couplings would indeed provide an immediate indication of new physics. Moreover, although additional SM-like Higgs bosons have been excluded over a wide range of masses, the existence of additional scalars with exotic decay modes remains a possibility.

The need for physics beyond the Standard Model is motivated by the fact that despite the beauty and the simplicity of the theory and its success in describing data, there are several open questions which find a solution only if the SM is assumed to be just the low energy limit of a more fundamental theory. The onset of new physics is expected at scales of the order of, or just above, the electroweak scale. The open issues in the SM and the motivations for searches of physics beyond it at the Large Hadron Collider of Geneva are outlined in Section 1.3.

Section 1.4 gives finally a description of searches for new physics in the Higgs sector. The possible channels studied at the CMS and ATLAS collaborations, together with the most recent results from both the two experiments, are summarized.

1.1 Introduction to the Standard Model

The Standard Model of particle physics is a description of nature which best summarizes our understanding of the fundamental structure of matter and the fundamental forces which govern all known phenomena in this regime. Mathematically, the Standard Model is a theory of interacting quantum fields. Excitations in these fields correspond to particles, and each separate field corresponds to a different type (or flavor) of particle. Developed in the early 1970s, it has successfully explained almost all experimental results and precisely predicted a wide variety of phenomena. Over time and through many experiments, the Standard Model has become established

as a well-tested physics theory [9].

The main constituents of the SM are shown in Fig. 1.1. The particles involved

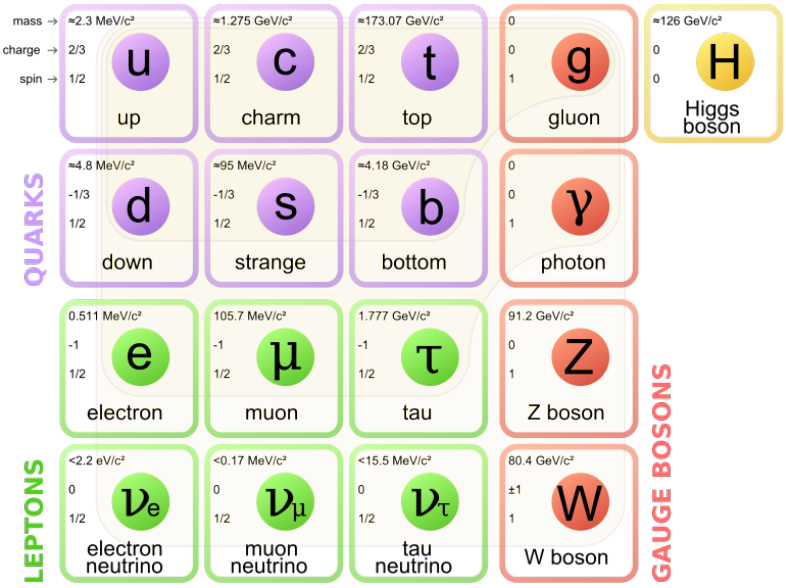


Figure 1.1. Overview of the Standard Model constituents.

are characterized by their spin, their mass, and the quantum numbers (charges) determining their interactions. Fermions are spin 1/2 particles and are organized in three different families according to their masses. Within each family they are distinguished by their charges under strong (quarks) and electromagnetic interactions (leptons). The neutrinos instead are neutral under both the strong and the electromagnetic interactions and are at least six order of magnitudes lighter than all the other SM fermions. The masses of the SM fermions span a range going from the sub-eV neutrino masses to the $1.7 \cdot 10^2$ GeV top quark [10].

Fermions interact in different ways via the exchange of spin-1 particles, the so called bosons. The SM provides a description of the strong, weak and electromagnetic interactions in terms of "gauge theories", i.e. theories which possess invariance under a set of space-time dependent transformations (called "local transformations"). According to the gauge principle, all the fundamental forces (with the exception of gravity) are mediated by the exchange of the gauge fields, corresponding to a specific symmetry group. The symmetry group which stands at the basis of the SM is:

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{1.1}$$

obtained as direct product of the symmetry group of strong interactions, $SU(3)_C$, and the symmetry group of the electroweak interactions, $SU(2)_L \otimes U(1)_Y$. The symmetry group of the electromagnetic interactions, $U(1)_{em}$, appears in the SM as a subgroup of $SU(2)_L \otimes U(1)_Y$ and it is in this sense that the weak and electromagnetic interactions are said to be unified. The gauge sector of SM is composed of 12 spin-1 gauge bosons: gluons, the $W^{\pm}$ bosons, the Z boson and the photon γ. The gluons are the mediators of the strong interactions. They are massless, electrically

neutral and carry color quantum number. The consequence of gluons being colorful is that they can interact with themselves. The $W^{\pm}$ and Z bosons are massive, self-interacting, particles and mediate the weak interactions. Finally the photon is massless, chargeless and non self-interacting and mediates the electromagnetic interactions.

The fact that the weak gauge bosons have a mass different from zero indicates that the electroweak group is not a symmetry of the vacuum. Also the fermion masses can not be included without violating gauge symmetry. Mass terms can be introduced exploiting the Spontaneous Symmetry Breaking Mechanism (EWSB) [11], [12], [13], which produces the following pattern in the SM:

$$SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \rightarrow SU(3)_C \otimes U(1)_{em} \tag{1.2}$$

with the purpose to give mass to the weak bosons and fermions and, at the same time, keep the photon massless.

The EWSB mechanism predicts the existence of a new scalar and electrically neutral particle: the Higgs boson. This particle was the last missing piece of the theory until July 2012 when the LHC experiments found a SM-like Higgs boson having mass around 125 GeV, in excellent agreement with the indirect predictions from electroweak precision data.

1.2 Discovery of a New Scalar Particle

At the end of July 2012 (after only two and a half years of data taking at the Large Hadron Collider in Geneva) both the ATLAS and the CMS collaborations [2], [3], published on Physics Letters B the observation of a new boson with properties very similar to those of the Standard Model Higgs boson. CMS observed an excess of events at a mass of approximately 125 GeV with a statistical significance of five standard deviations (5σ) above background expectations in more than one decay channel. This means that the probability of the background alone fluctuating up by this amount or more is about one in three million. The observation of the decay of this new scalar into two photons implies that the new particle is a boson, not a fermion, and that it cannot be a "spin 1" particle. Figure 1.2 shows the excess in the diphoton invariant mass, obtained by CMS analyzing 10.4 fb^{-1} of data collected during the 2011-2012 data-taking.

Subsequent studies of the production and decay rates [14] and of the spin-parity quantum numbers [15] of the new boson have been performed with the new data collected in 2012. New measurements show that the new boson properties are compatible with those expected for the SM Higgs boson. Nowadays CMS studies five main Higgs boson decay channels. Three channels result in pairs of bosonic particles ($\gamma\gamma$, ZZ or WW) and two channels result in pairs of fermionic particles (bb or $\tau\tau$). The bosonic channels are equally sensitive in the search for a Higgs boson and all are more sensitive than the bb and $\tau\tau$ channels.

The H$\rightarrow \gamma\gamma$ and $H \rightarrow ZZ \rightarrow 4l$ channels play a special role because of the excellent mass resolution of the reconstructed diphoton and four-lepton final states, respectively. The $H \rightarrow WW \rightarrow ll\nu\nu$ measurement has a high sensitivity but relatively poor mass resolution because of the presence of neutrinos in the final state.

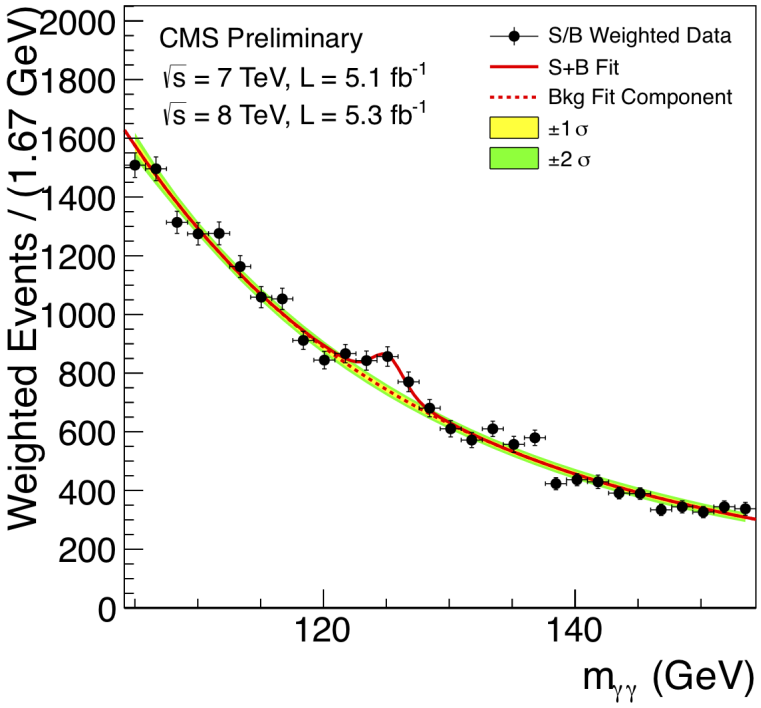


Figure 1.2. Diphoton invariant mass distribution as published in July 2012 by the CMS experiment [2]. The lines represent the fitted background and signal, and the colored bands represent the 1 and 2 standard deviation uncertainties in the background estimate.

The bb and $\tau\tau$ decay modes are afflicted by large background contributions and have relatively poor mass resolution, making it challenging to obtain a sensitivity comparable to that of the other channels. Figure 1.3 shows two-dimensional 68% Confidence Level (CL) regions for two parameters of interest in the search for the Higgs boson, the signal strength relative to the SM expectation, $\mu = \sigma/\sigma_{SM}$, and the mass, m_H, based on these channels. The value of σ_{SM} at a centre-of-mass energy of 8 TeV is equal to 19.9 pb [16].

From the high-resolution $\gamma\gamma$ and ZZ channels, the mass of this Higgs boson is measured by the CMS collaboration to be $m_H = 125.03^{+0.26}_{-0.27}(stat.)^{+0.13}_{-0.15}(syst.)$ GeV, with the precision dominated by the statistical uncertainty [17]. For this mass, the event yields obtained in the different analyses tagging specific decay modes and production mechanisms are consistent with those expected for the Standard Model Higgs boson, as shown in Figure 1.4. The combined best-fit signal strength, relative to the Standard Model expectation, is found to be $\mu = 1.00 + 0.09(stat.)^{+0.08}_{-0.07}(theo.) + 0.07(syst.)$ at the measured mass [17].

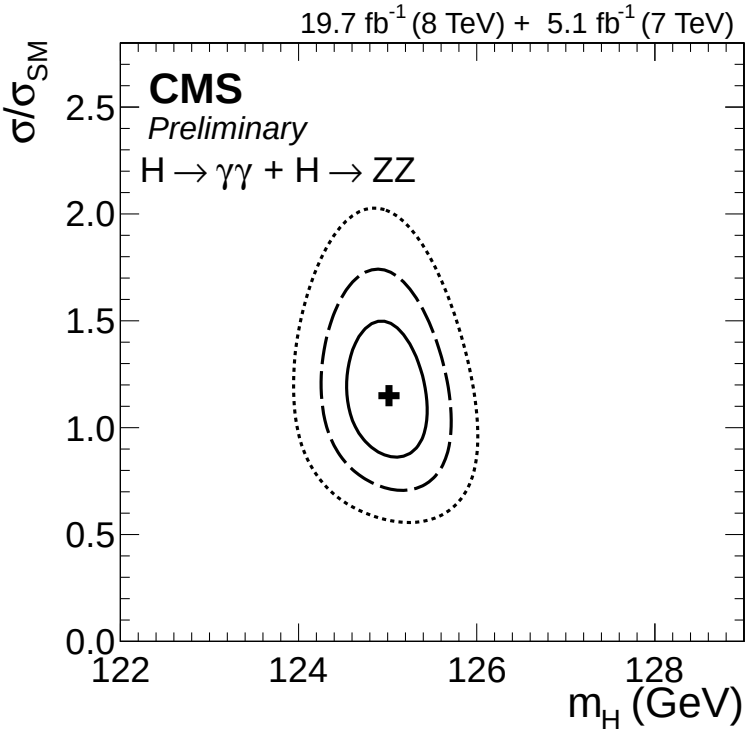


Figure 1.3. The 68% CL contours for the signal strength σ/σ_{SM} versus the boson mass m_H for the $\gamma\gamma$ and $4l$ combination [17].

1.3 Open Questions in the Standard Model

Even though the Standard Model is currently the best description of the subatomic world, and a new scalar with properties very close to those predicted for the Higgs boson has been recently observed, this theory does not explain the complete picture of the nature. In fact, there are strong conceptual indications for physics beyond the Standard Model (BSM), which is expected to be manifest at energies in the TeV domain, such as the fact that gravity is not described in the SM, that dark matter is there and we do not know what it is and that neutrinos mix and have mass. In addition there also other theoretical motivations like the unification and the naturalness problems which however can also be interpreted as a priori biases about how nature should be.

Dark Matter in the Universe

The distribution of matter in the universe has been observed to be dramatically different with respect to the one obtained considering visible objects only (stars, dusts). The observations of clusters of galaxies [18] and the studies on rotation curves of spiral galaxies strongly support the existence of some sort of matter that does not interact with photons and hence is not directly visible to astronomers. The studies by WMAP experiment on cosmic microwave background anisotropies [19]

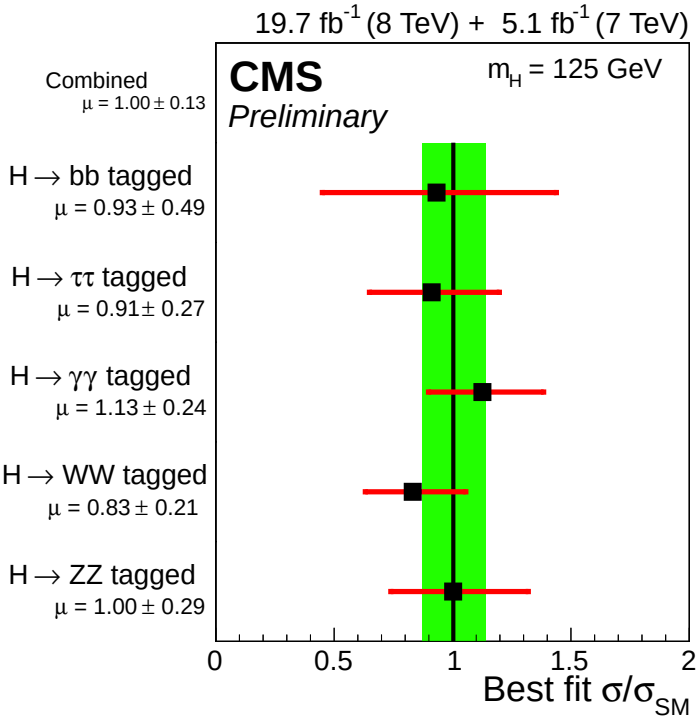


Figure 1.4. Sub-combinations by predominant decay mode of the new observed SM-like Higgs boson as presented in [17].

also confirm that 23% of the energy density of the universe is carried by dark matter.

All these observations suggest that this so-called dark matter is made of stable neutral particles produced in the big bang, permeating the universe. From the observations of the cosmic microwave background radiation we can in addition infer the fact that dark matter particles have to be non-relativistic (cold dark matter) [19]. Neither the density, nor the non-relativistic nature are consistent with the only candidates the Standard Model can provide, the neutrinos. So if the dark matter problem is to be solved with the existence of a new weakly interacting massive particle (a WIMP), it has to be part of a theory beyond the Standard Model.

Neutrino Masses

The simplest form of the Higgs mechanism for Spontaneous Symmetry Breaking provides masses for all the elementary fermions but fails to explain the oscillations observed between neutrinos, which imply non-zero masses of at least two neutrinos [20]. Recent report of atmospheric neutrino oscillations, as well as previous other indications from solar neutrino studies [21] and studies at accelerators [22], nowadays establish experimentally that neutrino masses have to be considered different from zero.

Gravity

There are four fundamental forces at work in the universe: the strong force, the weak force, the electromagnetic force, and the gravitational force. They work over different ranges and have different strengths. Looking at relative strengths of fundamental interactions, gravity is by far the most weak, and it is not included in the description of nature given by the SM. Near the Plank scale (10^{19}GeV) however the relative contribution of gravitational effects becomes relevant and including gravity in a consistent theoretical framework becomes mandatory.

The lack of a theoretical description or experimental clues of the bosons which mediate gravitational interactions within the Standard Model represents indeed an undeniable proof that a more general theory has to be introduced to replace the Standard Model by unifying all 4 fundamental forces in the so-called "Theory of Everything" [23].

Unification of the Gauge Couplings

In addition to a quantum mechanic description of gravity, many other unsolved questions related to the origin of the mass in the universe and to the nature and the intensity of the known interactions at any range of energy, find promising answers in the so called Grand Unified Theory (GUT) [24]. Among them the Gauge Problem related to the three independent symmetry groups of the theory, as well as the Fermion Problem [25] about the number and the origin of quarks and leptons and the Charge Quantization Problem [26] are just few of the issues that the GUT theory try to solve in an elegant way.

The basic principle of a GUT theory is that the known electromagnetic, weak and strong nuclear forces are combined into a single theory as different branches of a unique interaction, associated with a simple gauge group. Therefore, a unification scale should exist, at which the electromagnetic, weak and strong interactions have to be described by one unified coupling constant rather than three independent ones.

In Quantum Field Theory the coupling constants are only effective constants at a certain energy. They are energy or equivalently distance dependent through virtual corrections both in Quantum Electrodynamics (QED) and in Quantum Chromodynamics (QCD). However in QED the coupling constant increases as function of the energy scale Q while in QCD the coupling constant decreases. Specifically, they evolve, according with the renormalization group equation, as a function of Q [27] as following:

$$\alpha_i^{-1}(Q) = \alpha_i^{-1}(m_Z) + \frac{b_i}{2\pi} log \frac{Q}{m_Z} \tag{1.3}$$

where $\alpha_i(m_Z)$ is the coupling constant calculated at the electroweak scale, b_i is a coefficient different for each coupling, depending on the number of quark and lepton generations and the number of Higgs doublet fields of the theory, and m_Z is the mass of the Z boson.

Figure 1.5 shows the evolution of the inverse of the couplings as a function of the energy scale. Although the values of the coupling constants tend to converge for high Q values, they do not come to a common value at any scale.

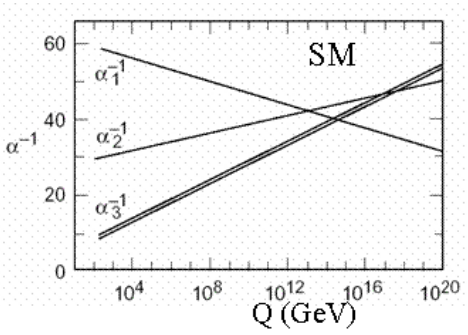


Figure 1.5. Evolution of the inverse of the three coupling constants in the Standard Model [28].

The failure in providing a unification of the fundamental interactions suggests the presence of new physics phenomena beyond the SM, occurring at a higher energy scale.

Naturalness of the Higgs Mass

Under the assumption that the Standard Model is an effective theory which manifests at low-energy corresponding to the electroweak scale $\langle v \rangle = 246$ GeV, another scale which became relevant is the cut-off Λ which represent the scale up to which the theory is valid. This means that above the cut-off scale the Standard Model ceases to be valid and the new physics takes over.

However the Higgs mass diverges quadratically with the cut-off value Λ: the quantum correction to the mass term is indeed proportional to the cut-off scale squared because of the quadratic divergence as follow [28]:

$$\left(\Delta M_h^2\right)_f \simeq \frac{\lambda_f^2}{4\pi^2}\left(\Lambda^2 + m_f^2\right) + ... \tag{1.4}$$

Therefore while the mass term of the Higgs field is of the order of the weak scale (125 GeV), the natural scale for the mass term is, however, the cut-off scale of the theory [29]. If the cut-off scale of the Standard Model is near the Planck scale (10^{19} GeV), one needs to fine-tune the bare mass term of the Higgs potential to roughly 17 orders of magnitude to keep the weak scale very tiny compared to the Planck scale. In order to solve this problem a theory more general and complete than the SM is needed. The most famous scenario indeed is Supersymmetry [30], where a new symmetry is introduced to eliminate these quadratic divergences.

1.4 Search for New Physics in the Higgs Sector

From a general point of view the search for new physics within the Higgs sector can be performed in three principal ways. The first is to make precision measurements of the couplings of the SM-like Higgs itself, as these couplings could be altered from their Standard Model values by mixing between scalars. The second possibility is to

search for new decay modes of the already observed Higgs which are not predicted by the SM. The third is to search for new heavy states in additional channels, including those in which heavier scalars decay to final states involving the SM-like Higgs itself [31]. There is extensive interplay between these three avenues, since the couplings of the SM-like Higgs are correlated with the variation in production and decay modes of additional scalars and vice versa [32]. Although both charged and neutral new particles are predicted, this thesis focuses only on the searches for neutral candidates in the extended Higgs sector.

1.4.1 Non Standard Model Properties of the Higgs Boson

The first possibility, in the search for BSM physics in the Higgs sector, is to study the properties of the already observed Higgs boson and look for something unexpected which may be an hint for new physics. Among all, particularly relevant are searches for unusual and rare decays of the 125 GeV boson. These observations may signify a Higgs boson with non-SM couplings to vector bosons and fermions.

Invisible decays

One could imagine that there are additional particles beyond the Standard Model, that escape our detectors, to which the Higgs can couple. This will enhance the Higgs invisible branching fraction significantly, compared to O(0.1%) as predicted in the Standard Model via $H \to ZZ \to 4\nu$ decay. At a mass of 125 GeV, the invisible branching fraction, $BR(inv) = BR(H \to invisible)$, of the Higgs is especially sensitive to new particles at electroweak scale. Supersymmetric theories (SUSY) tend to contain a stable neutral Lightest SUSY Particle (LSP), and this can open up decays of the Higgs boson into pairs of LSP's, for example neutralinos [30]. Some theories of large extra dimensions predict graviscalars that could mix with the Higgs boson [33]. The signature would be equivalent to an invisible decay of the Higgs boson.

Such searches were performed at LEP [34], using the ZH associated production mode. They excluded at 95% CL an invisible Higgs boson of mass smaller than 105 GeV and produced with a cross section higher than 0.2 times the Standard Model ZH cross section. Phenomenological studies of hadron collider searches for H(inv) have considered all production mechanisms. The ATLAS collaboration reported a search for invisible decays of a Higgs boson produced in association with a Z boson that decays to leptons [35], placing an upper limit on the invisible Higgs boson branching fraction of 0.75 at 95% CL for $m_H = 125.5$ GeV. The ATLAS collaboration also searched for an invisibly decaying Higgs boson in association with either a W or Z boson decaying hadronically [36]. CMS reports searches for H(inv) in the ZH mode, where the Z boson decays to leptons or a $b\bar{b}$ quark pair, and the first search for H(inv) in the vector boson fusion (VBF) production mode, where the Higgs boson is produced in association with two quarks. By assuming standard model production cross sections, and combining all channels, the upper limit on the invisible branching fraction of a Higgs boson for $m_H = 125$ GeV, is found to be 0.58, with an expected limit of 0.44, at 95% CL [37].

In general, interactions of the Higgs boson with the unknown dark matter (DM)

sector may also introduce invisible decay modes, and bounds on these decays can constrain DM models. In the so-called "Higgs-portal" models of DM interactions [38], [39], the Higgs boson takes the role of mediator between the SM particles and the DM particle. If the DM candidate has a mass below $m_H/2$, the Higgs boson can decay directly in a pair of DM candidates and the invisible decay width, Γ_{inv}, can be directly translated to the DM-nucleon elastic cross section. As shown in Figure 1.6, the results from searches for invisible decays of the observed boson (performed by CMS [37] but also ATLAS makes a similar interpretation) can be compared with results from direct detection experiments which studies the elastic interaction between DM and nuclei in terms of DM mass, M_χ, and DM-nucleon cross section. This kind of search at colliders became particularly relevant in the low DM candidate mass region where the direct searches loose sensitivity.

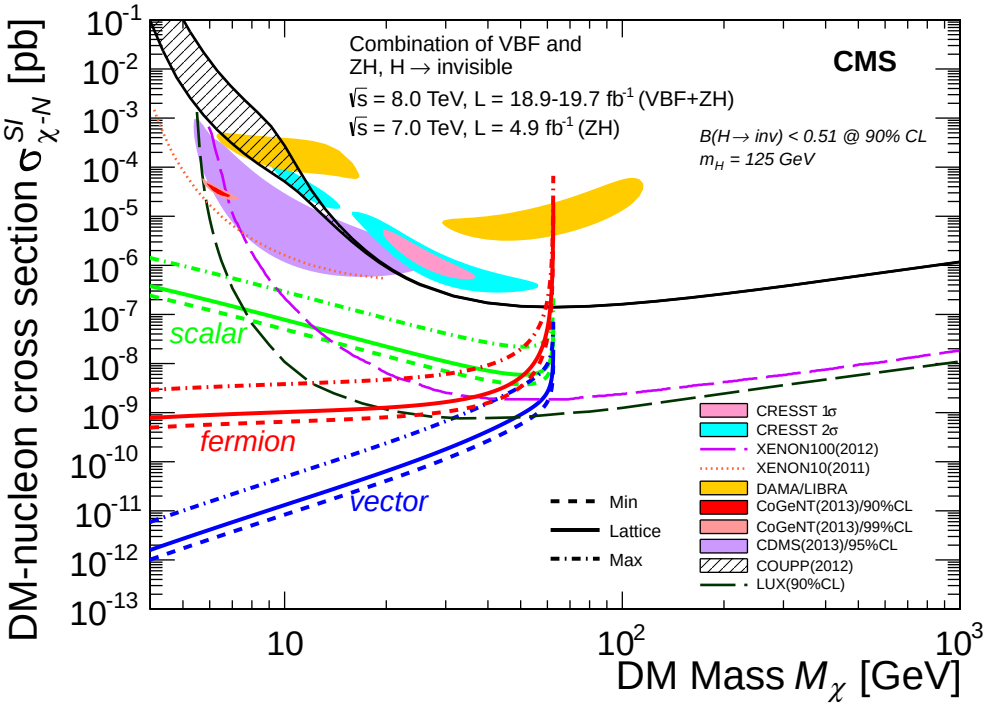


Figure 1.6. Upper limits on the spin-independent DM-nucleon cross section in Higgs-portal models, derived for $m_H = 125$ GeV and $BR(H \to inv) < 0.51$ at 90% CL, as a function of the DM mass by the CMS experiment [37]. Limits are shown separately for scalar, vector and fermion DM. The solid lines represent the central value of the Higgs-nucleon coupling, which enters as a parameter, and is taken from a lattice calculation, while the dashed and dot-dashed lines represent lower and upper bounds on this parameter. Other experimental results are shown for comparison, from the CRESST, XENON10, XENON100, DAMA/LIBRA, CoGeNT, CDMS II, COUPP, LUX Collaborations.

Recent theories proposing that the Higgs boson played a central role in the evolution of the early universe [40] provide further motivation to understand the relationship between the Higgs boson and DM.

Rare decays

Rare Higgs boson decays searches, such as $H \to \mu\mu$ or $H \to Z\gamma$ performed by ATLAS collaboration [41], [42], or the $H \to \gamma^*\gamma \to \mu\mu\gamma$ studied by CMS [43], are rich in information, that will not only increase our understanding of the basic properties of the EWSB, but can also probe novel couplings caused by possible extensions of the Standard Model [44].

Concerning for example the $H \to \mu\mu$ search presented by ATLAS [41], the observed di-muon spectrum in data has been studied and found to be consistent with the expected background as shown in Figure 1.7. Since no evidence for a signal

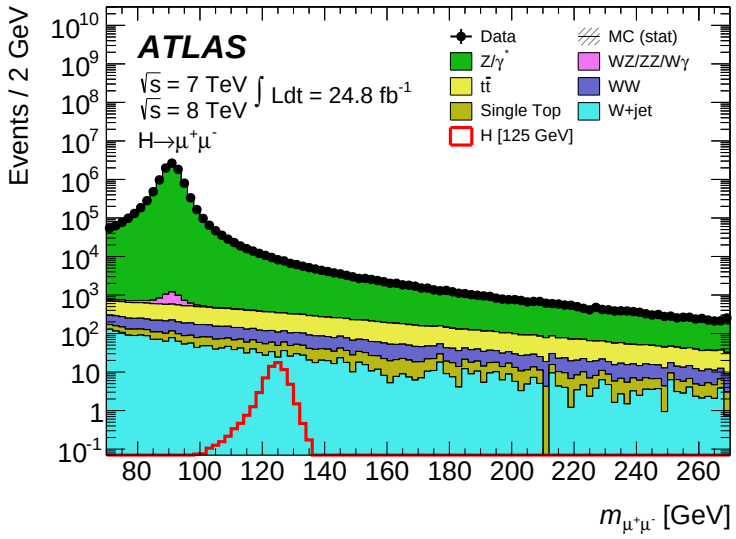


Figure 1.7. The distribution of the dimuon invariant mass with all the selection requirements applied as presented by the ATLAS experiment for the $H \to \mu\mu$ search. The expected signal is shown for $m_H = 125$ GeV.

is observed, upper limits are set on the signal strength of the $H \to \mu\mu$ process as a function of the Higgs boson mass. For a SM Higgs boson with a mass of 125.5 GeV, the observed (expected) limit on the signal strength at the 95% CL is 7.0 (7.2) and the quoted 95% CL upper limit on the $H \to \mu\mu$ branching ratio is $1.5 \cdot 10^{-3}$.

Conversely, in the search for the rare $H \to \gamma^*\gamma \to \mu\mu\gamma$ decay , which is performed by CMS, the various contributions shown in Figure 1.8 need be disentangled to a certain extent in order to identify the signal. Requirements on the di-muon invariant mass and transverse momentum of the photon are used to separate the loop-induced processes $H \to \gamma^*\gamma$ and $H \to Z\gamma$. Similarly, contributions from tree-level diagrams with final-state radiation such as $H \to \mu\mu$ and $H \to q\bar{q}$ (where $q\bar{q}$ hadronizes into mesons that decay to $\mu\mu$) are reduced by requiring the photon and the leptons to be isolated from each other. Also for this search upper limits are set on the signal strength as a function of the Higgs boson mass, since no significant excess has been observed over the full mass range, as shown in Figure 1.9.

The expected exclusion limits at 95% CL are between 8 and 13 times the SM cross section and the observed limit ranges between about 4 and 19 times the SM

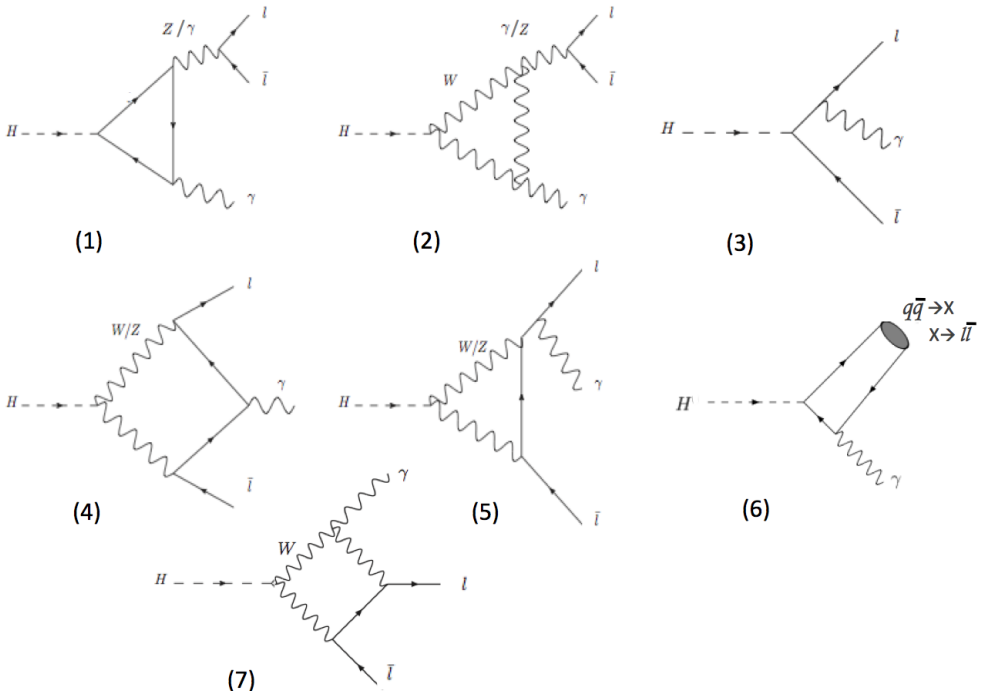


Figure 1.8. Diagrams contributing to $H \to \mu\mu\gamma$ [43]. The contributions from diagrams (1) and (2) dominate.

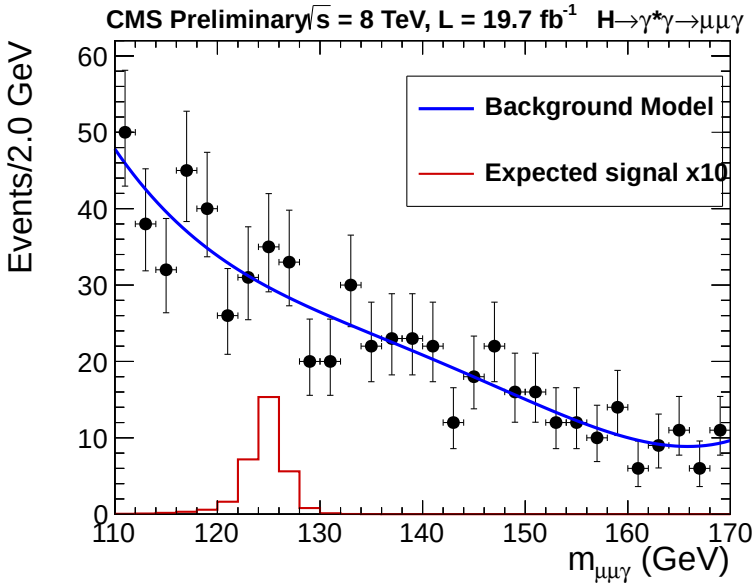


Figure 1.9. The $m_{\mu\mu\gamma}$ spectrum for 8 TeV data [43]. Also shown is the simulated SM Higgs boson signal events, $m = 125$ GeV, (histogram), scaled by a factor 10.

cross section. The observed and expected limits for $m_{\mu\mu\gamma}$ at 125 GeV and $m_{\mu\mu} < 20$ GeV are about ten times of the SM prediction. Due to the parity-violating nature of weak interactions and the interference between $H \to \gamma^*\gamma$ and $H \to Z\gamma$ decay modes, a forward-backward asymmetry of -0.025/+0.02 in $H \to \mu\mu\gamma$ events is also expected [44]. This property along with non-trivial angular distributions, germane to this three-body decay, can be used to understand the Higgs beyond what we learn from H$\to \gamma\gamma$.

Exotic decays

One example of exotic decay of the Higgs boson are decays which violate the lepton number conservation law (LFV) [45]. In the Standard Model LFV decays of the Higgs boson are not allowed if the theory is to be renormalizable. If however, the requirement that it is renormalizable is relaxed, so that it is a theory that is valid only to a finite mass scale, then LFV couplings may be introduced. LFV decays can also occur naturally in models with more than one Higgs doublet without giving up on renormalizability [46]. They also arise in Composite Higgs models [47], models with flavor symmetries [48] and Randall-Sundrum Models [49].

The presence of LFV Higgs couplings would allow LFV effects in decays mediated by virtual Higgs. There are three possibilities: $\mu \to e$, $\tau \to \mu$ and $\tau \to e$ transitions. The experimental constraints available on these transitions have been reviewed and translated into constraints on the following branching ratios: $B(H \to \tau\mu)$, $B(H \to e\tau)$ [50]. The present limits on these decays are: $B(H \to \mu\tau) < O(0.1)$ and $B(H \to e\tau) < O(0.1)$. The observation of the Higgs boson offers for the first time the possibility of sensitive direct searches for the LFV Higgs decays $H \to \tau\mu$ and $H \to e\tau$. In particular the constraint on $H \to \tau\mu$ can be interpreted in terms of LFV Higgs Yukawa couplings. As shown in Figure 1.10, results from a CMS search [51] in the $H \to \tau\mu$ channel can be compared to the constraints from previous indirect measurements on the $|Y_{\tau\mu}|$. It can be seen from this plot that the direct search improves the constraints by roughly an order of magnitude.

Another interesting search for exotic decays of the already discovered scalar, is the first ATLAS search for the Higgs decay, $h_0 \to \pi_\nu\pi_\nu$, to two identical neutral particles (π_ν) that have a displaced decay to fermion/anti-fermion pairs [52]. These long-lived particles occur in many models, including gauge-mediated extensions of the Minimal Supersymmetric Standard Model with R-parity violation [53], inelastic dark matter [54] and the Hidden Valley (HV) scenario [55]. In many of these BSM scenarios, the lifetime of the neutral states is not specified and can have a very large range extending to about 20 m by exploiting the size and layout of the ATLAS muon spectrometer. Also for this search, no events in the data sample pass the final selection requiring two isolated, back-to-back vertices in the muon spectrometer. Since no significant excess over the background prediction is found, exclusion limits for $\sigma_{h_0} \times BR(h_0 \to \pi_\nu\pi_\nu)$ are set by rejecting the signal hypothesis at the 95% CL. Figure 1.11 shows the 95% CL upper limit on $\sigma_{h_0} \times BR(h_0 \to \pi_\nu\pi_\nu)/\sigma_{SM}$ as a function of the π_ν proper decay length ($c\tau$) in multiples of the SM Higgs cross section, σ_{SM}. As expected the Higgs and π_ν mass combinations with the largest boosts leading to larger $\beta\gamma c\tau$ have the smallest exclusion limits.

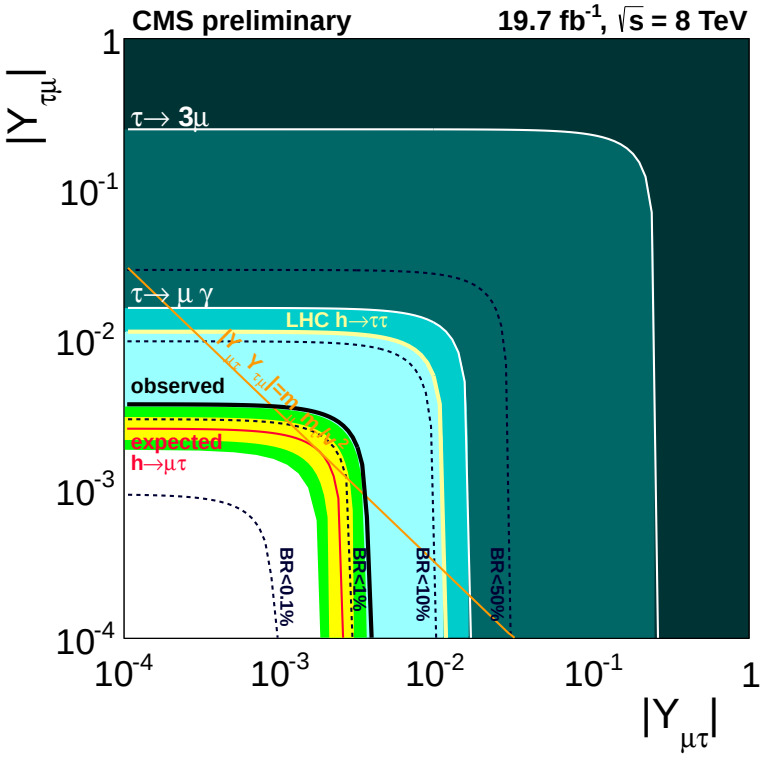


Figure 1.10. Constraints on the flavor violating Yukawa coupling $Y_{\mu\tau}$ from the CMS experiment [51]. The expected (red solid line) and observed (black solid line) limits are derived from the limit on $B(H \to \mu\tau)$ from the present analysis. The diagonal Yukawa couplings are approximated by their SM values.

1.4.2 Additional Neutral Scalars in an Extended Sector

In addition to looking for evidence of unexpected decay modes or deviations from SM predictions for the couplings of the observed SM-like Higgs, it is instrumental to search directly for additional scalars, both in standard channels and in cascade decays involving multiple scalars. These searches are motivated in several new physics models as SUSY [30], Little Higgs [56] and Gauge Higgs Unification [57] which often predict extended Higgs sectors as a low energy effective theory. The Higgs sector in the SM takes a minimal form, which is composed of only one isospin doublet Higgs field. However, there are in principle many possible extensions and deformations of the minimal Higgs sector. The simplest extension to the SM, which provides a framework addressing naturalness, gauge coupling unification, and the existence of dark matter, is the Minimal Supersymmetric Standard Model (MSSM) [58].

Within the MSSM, of great interest are theories with two Higgs fields transforming as doublets under $SU(2)_L$ with unit $U(1)_Y$ charge. Such models, which provide a general effective theory framework for extensions of the electroweak symmetry breaking sector are referred as Two Higgs Double Models (2HDMs) [6]. The phenomenology of 2HDMs is rich, as five physical Higgs sector particles remain after electroweak symmetry breaking: two neutral CP-even scalars, h, H; one neutral

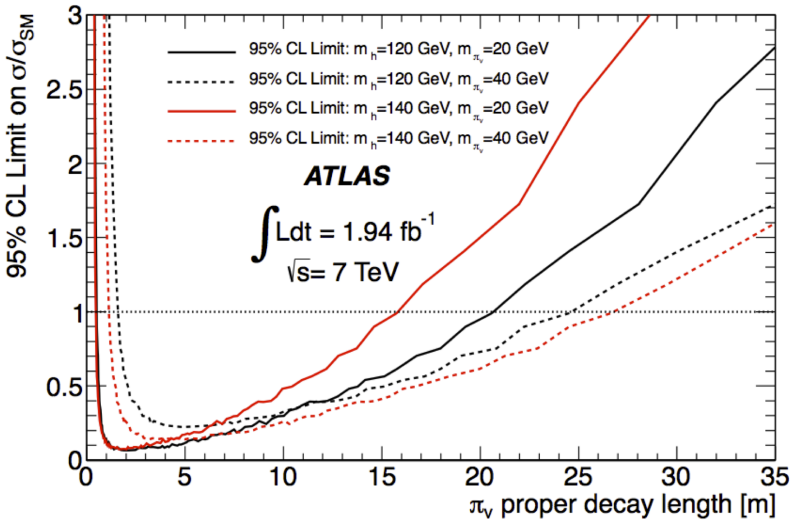


Figure 1.11. Observed 95% upper limits on the process $h_0 \to \pi_\nu\pi_\nu$, vs. the π_ν proper decay length, expressed as a multiple of the SM cross section for Higgs production [52]. Exclusion limits assume 100% branching ratio for the Higgs decaying to π_ν's.

CP-odd pseudoscalar, A; and two charged scalars, H^+ and H^- [59] . All of these states could have masses at or below the TeV scale, in a regime accessible to the LHC. The parameter space of these models can accommodate a wide range of variations in the production and decay modes of the SM-like Higgs boson, as well as discoverable rates for the production and decay of additional scalars [60]. A generic feature in the 2HDM Higgs sector is the prediction of a decoupling limit in which the heavy Higgs bosons have similar masses and the light CP-even Higgs boson (h) in practice becomes identical to a SM Higgs boson with the same mass.

Depending on the values of the masses of the heavy Higgs as well as on the values of the space parameters of the model, several interesting final states can be studied at LHC in order to probe the extended Higgs sector. The most recent results to date from both the ATLAS and CMS collaborations concerning the search for new neutral heavy scalars are summarized in the following sections. A more detailed description of the 2HDM focusing on the role and possible interpretation of a search for diphoton resonances within this framework is given in Chapter 2.

Neutral scalar decays to SM particles

A variety of searches in $\tau\tau$, bb and $\gamma\gamma$ decays of neutral H and A scalars find their motivations in an enhancement in the couplings with respect to the SM for defined regions of the MSSM space parameters, resulting in increased branching fractions to τ leptons, b quarks and photons. Among them the most recent results on $\tau\tau$ and $\gamma\gamma$ searches at the LHC are briefly described in this section.

- **$H/A \to \tau\tau$**

  In the study of this decay, five different $\tau\tau$ final state signatures are studied by both the CMS [61] and ATLAS [62] collaborations: $\tau_e\tau_h$, $\tau_\mu\tau_h$, $\tau_e\tau_\mu$, $\tau_\mu\tau_\mu$,

and $\tau_h\tau_h$, where τ_h denotes a hadronically decaying τ. These results are complementary to the searches at the Tevatron [63] and LEP [64]. The sensitivity is improved by performing a categorization based on expected Higgs boson mass and production mechanisms. Figures 1.12 show $m_{\tau\tau}$ distributions in one of the possible channels from both the CMS and ATLAS collaborations.

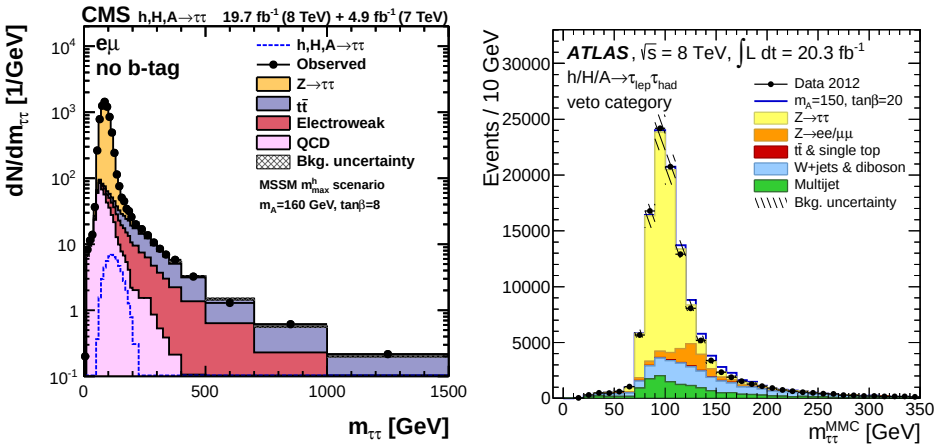


Figure 1.12. Left: Observed and predicted $m_{\tau\tau}$ mass distributions in the $e\mu$ channel at the CMS experiment [61]. Right: $m_{\tau\tau}$ mass distributions for the low-mass categories of the $h/H/A \to \tau_{lept}\tau_{had}$ channel at ATLAS [62]. The background uncertainty includes statistical and systematic uncertainties.

The searches find no indication of an excess over the expected background in the channels considered and 95% CL limits are set, which provide tight constraints in the MSSM parameter space for several modified scenarios. The results on the so called "m_h^{mod+}" MSSM model [65] are shown in Figures 1.13 comparing both CMS and ATLAS.

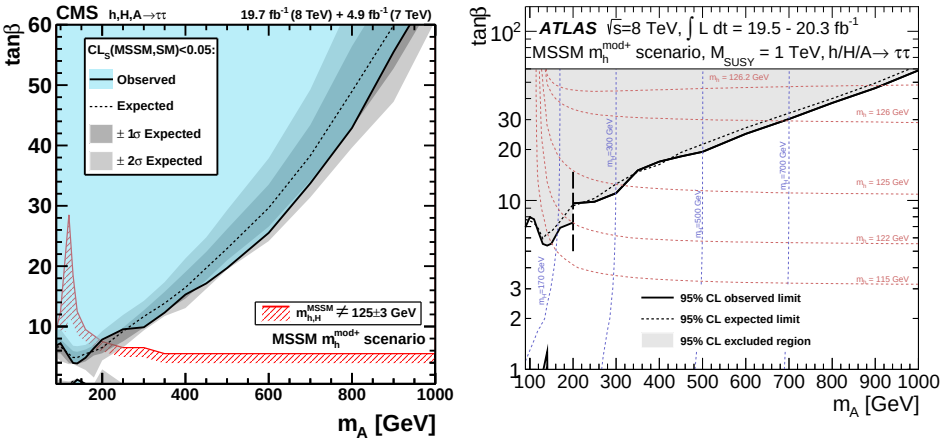


Figure 1.13. Left: MSSM vs SM limit in the MSSM m_h^{mod+} scenario in the m_A - [1]$tan\beta$ space parameters by CMS [61]. Right: Expected (dashed line) and observed (solid line) 95% CL exclusion limits on $tan\beta$ as a function of m_A by ATLAS [62]. The light grey area denotes the excluded region. The exclusion limits are shown for the MSSM m_h^{mod+} scenario. The vertical dashed line at 200 GeV indicates the transition point between low and high mass categories.

- **H/A → γγ**

  The $\gamma\gamma$ final state becomes relevant in the search for an extended Higgs sector when the h in the 2HDMs approaches the "alignment limit" (see Section 2.1.1) and its couplings become almost the same of the SM couplings of the Higgs boson. Recent results on an inclusive diphoton search have been published by both CMS and ATLAS. The two analyses are similar and different at the same time. ATLAS extends the search also in the low mass region, while CMS looks only at the high mass region above 150 GeV but extends the search to new resonances with natural widths up to 10% of the value of the mass. The search preformed by CMS, of which I am the main author, is the topic of this thesis, hence, a more detailed discussion will be given in Chapter 2.

Heavy neutral scalar decays to SM-like Higgs pairs

If the mass of the heavy H were close to twice the mass of SM-like observed Higgs, then apart from SM channels, the only extra way for them to decay would be via two SM-like Higgs [66]. Similarly for A decaying to Z boson together with a SM-like Higgs is the primary mode, when mass of A is between twice of SM Higgs mass and twice the top mass. Among the several final states the most interesting, and easiest to identify, ones are signatures with leptons and photons.

- **H/A → hh →multilepton final states**

  Decays of the heavy H/A to two SM-like Higgs lead to interesting distinct possibilities which provide multilepton final states such as WWWW, WWZZ,

[1] $tan\beta$ is the ratio of the vacuum expectation values of the two Higgs doublets

WW$\tau\tau$, ZZZZ, ZZ$\tau\tau$, ZZbb, and $\tau\tau\tau\tau$, which are easily identifiable at the LHC and represent promising search channels for discovery or exclusion of the new physics [67]. Observed multilepton events in CMS with or without diphoton candidates are organized into exclusive search channels based on event kinematics. The search channels are ordered by the amount of expected Standard Model background. Data-based estimation of the Standard Model backgrounds is emphasized, but data-validated simulations are also employed as appropriate. Observations are consistent with the Standard Model and lead to the exclusion of significant portions of Two Higgs Doublet Model parameter space presented in terms of the relevant parameters of the model, as shown in Figures 1.14.

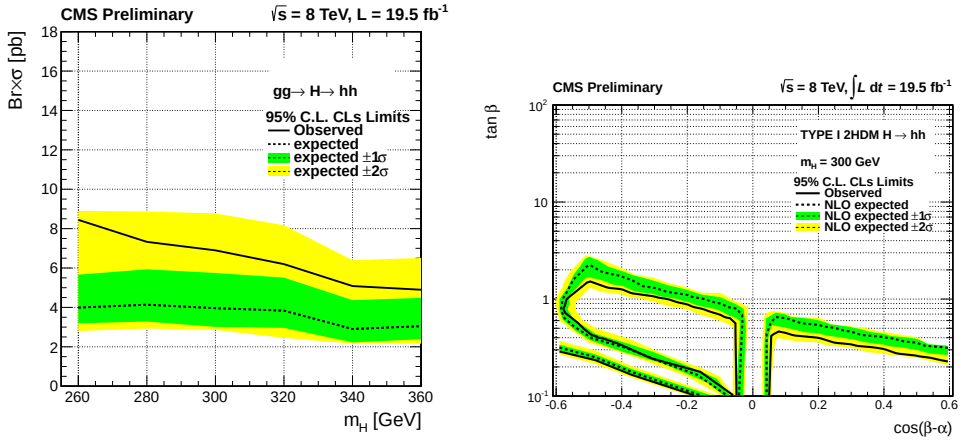


Figure 1.14. Left: Observed and expected limits with 1 and 2-σ bands for $H \to hh$ in terms of $\sigma \cdot BR$ by the CMS experiment [67]. These limits are based on multilepton channels. Branching ratios for h are assumed to have Standard Model values. Right: Observed and expected limits on Heavy higgs of mass 300 GeV in Type I 2HDM (see Section 2.1.1). The parameters α an β determine the cross section for H production, the branching ratio $BR(H \to hh)$ and the branching ratios $BR(h \to WW, ZZ, \tau\tau, \gamma\gamma)$. Region below the observed limit and within the observed limit loop is excluded.

- **$H/A \to hh \to \gamma\gamma b\bar{b}$**

  The $\gamma\gamma b\bar{b}$ channel is a fully reconstructed final state and a powerful final state in which to search for Higgs boson pair production [68] thanks to the large $h \to b\bar{b}$ branching ratio, clean diphoton trigger, excellent diphoton invariant mass resolution, and low backgrounds. The search is performed in the mass range 260-500 GeV by ATLAS [69] and in the range 260-1100 GeV by CMS [70].

  The $\gamma\gamma b\bar{b}$ search is interpreted by the two experiments in terms of Warped Extra Dimensions (WED) [71], models which predict the existence of spin-0 radion [72] or the first Kaluza-Klein (KK) excitation of the spin-2 graviton [73].

  Figures 1.15 shows simulated $\gamma\gamma b\bar{b}$ signal models at different masses and the observed and expected exclusion power for this $H \to hh$ searches, which are very similar among the ATLAS and CMS collaborations.

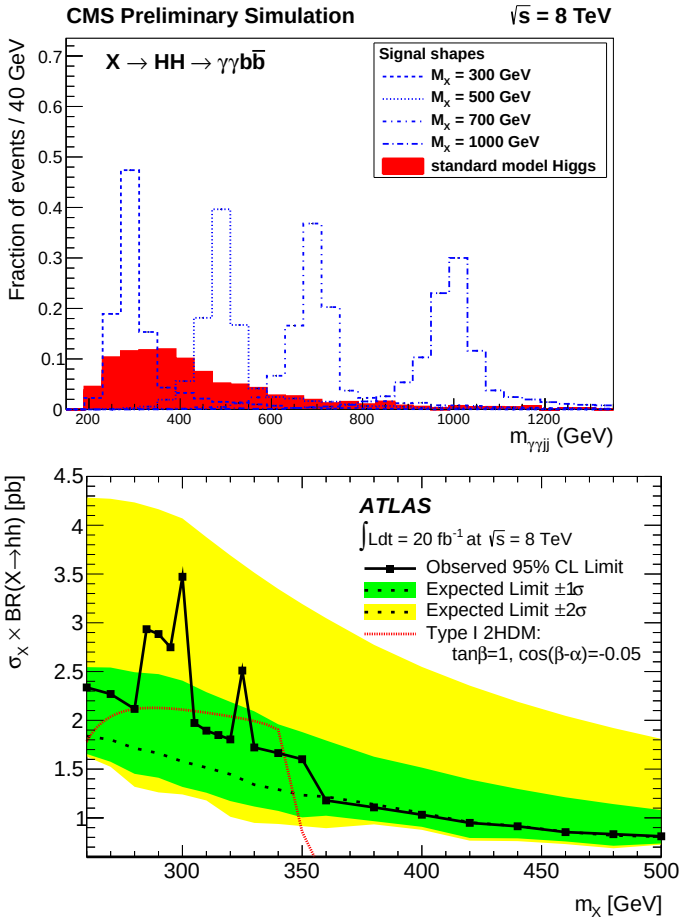


Figure 1.15. Top: Simulated mass spectra for the $\gamma\gamma b\bar{b}$ signal after basic selections on photons and requesting at least one loose b–tagged jet at CMS for different mass hypotheses [70]. Bottom: A 95% CL upper limit on the cross section times branching ratio of a narrow resonance decaying to pairs of Higgs bosons as a function of $m_{\gamma\gamma b\bar{b}}$ from ATLAS.

Another interesting channel, on which both the ATLAS and CMS collaboration are working at present, is also the $H/A \to hh \to b\bar{b}b\bar{b}$.

Chapter 2

Search for New Scalar Particles Decaying to Two Photons

The search presented in this thesis covers for the first time the unexplored intermediate energy range in the diphoton spectrum immediately above the energies investigated by the SM-like Higgs search [2] which stops around 160 GeV since already-small branching ratio of the SM-like Higgs falls rapidly with increasing mass [74]. This has been conceived as a model independent search for any possible signal of a new resonance which decays to two photons with a mass above 150 GeV. The analysis strategy and the choice of a cut-based event selection rather than a multivariate one (as described in Section 5.3) have been guided by the need to be sensitive to whichever new particle (produced via gluon fusion or via vector boson fusion) with different spin hypotheses. However new particles which decay to two photons are predicted in different scenarios of physics BSM as 2HDMs or models with extra-dimensions. An insight on these models is given in Section 2.1. Section 2.2 gives an overview on the analysis strategy exploited in this thesis for the identification of new resonant phenomena in the diphoton final state at CMS experiment. This chapter ends with a review of the existing searches for new resonances decaying to photons performed by the experiments at colliders, in Section 2.3.

2.1 Theoretical Motivations

Although a model independent approach is here considered, in order to be sensitive to any possible model which predicts the existence of a new state detectable via the diphoton decay mode, new scalars decaying to photons can be found in several BSM scenarios. The importance of the diphoton channel in the 2HDM and in models with extra-dimensions are described in this section.

2.1.1 2HDM Phenomenology

Even though the general parameter space of 2HDMs is large as outlined in Section 1.4.2, it can be significantly reduced using a set of reasonable assumptions which constrain the model, as for example the absence of three-level flavor-changing neutral currents processes. The mass spectrum can be categorized into two approximate

scales: a light SM-like Higgs h at $m_h = 125$ GeV, and the remaining physical Higgs scalars $H, A, H^{\pm}$ clustered together at an equal or higher scale with $m_H \sim m_A \sim m_H^{\pm}$. In this case the only available decay modes of the SM-like Higgs are those involving SM states. The absence of tree-level flavor-changing neutral currents in theories with multiple Higgs doublets is guaranteed by the Glashow-Weinberg condition [75] that all fermions of a given representation receive their masses through renormalizable Yukawa couplings to a single Higgs doublet, in which case the tree-level couplings of neutral Higgs bosons are diagonal in the mass eigenbasis [76]. In theories with only two Higgs doublets $\Phi_{1,2}$, the Yukawa couplings are [66]:

$$V_{yukawa} = -\sum_{i=1,2} (Q\tilde{\Phi}_i y_i^u \bar{u} + Q\Phi_i y_i^d \bar{d} + L\Phi_i y_i^e \bar{e} + h.c.) \tag{2.1}$$

and the Glashow-Weinberg condition is satisfied by four discrete assignments, where by convention up-type quarks are always taken to couple to Φ_2^0:

- Type 1, in which $y_1^{u,d,e} = 0$; all fermions couple to one doublet.
- Type 2, in which $y_1^u = y_2^d = y_2^e = 0$; the up-type quarks couple to one doublet and the down-type quarks and leptons couple to the other.
- Type 3, in which $y_1^u = y_1^d = y_2^e = 0$; quarks couple to one doublet and leptons to the other.
- Type 4, in which $y_1^u = y_2^d = y_1^e = 0$; up-type quarks and leptons couple to one doublet and down-type quarks couple to the other.

Often the theory is represented as a function of few free parameters:

- $tan\beta = |\langle\Phi_2^0\rangle / \langle\Phi_1^0\rangle|$; the ratio of vacuum expectation values of $\Phi_{1,2}^0$
- α; the mixing angle that diagonalizes the 2×2 neutral scalar $h - H$ mass squared matrix;
- the four physical masses $m_h; m_H; m_A; m_H^{\pm}$; and three more couplings.

The angles α and β fully determine the couplings between a single physical Higgs boson and two gauge bosons or two fermions, as well as the coupling between two Higgses and a single gauge boson. Under the assumption that the new observed Higgs boson is the light CP-even Higgs scalar of the 2HDM theory (h), the consistency of its couplings to the ones predicted by the SM push the model close to the alignment limit ($cos(\beta - \alpha) = 0$) where, as illustrated in Figure 2.1, certain decay modes of heavy scalars vanish, including $H \to VV$, $H \to hh$, and $A \to Zh$. Moreover, in proximity to the alignment limit, the associated production modes for H or A involving vector couplings (such as vector boson fusion or production in association with a W or Z boson) are likely to be suppressed.

Second Higgs Doublet Decay Topology	Alignment Limit
$H \rightarrow WW, ZZ$	–
$H, A \rightarrow \gamma\gamma$	✓
$H, A \rightarrow \tau\tau, \mu\mu$	✓
$H, A \rightarrow tt$	✓
$A \rightarrow Zh$	–
$H \rightarrow hh$	–
$t \rightarrow H^{\pm} b$	✓

Figure 2.1. Leading 2HDM decay topologies with unsuppressed production cross sections near the alignment limit for neutral scalars with mass spectra $m_H, m_A > 2m_h$, and top quark with $m_{H^{\pm}} < m_t + m_b$. A check-mark (dash) indicates that the partial decay width approaches a constant (vanishes) in the $cos(\beta - \alpha) = 0$ alignment limit.

Importance of the Diphoton Channel

In the search for the most promising final states for the discovery of additional Higgs scalars, one has to take into account that, although all the decay channels play complementary roles, when kinematically available the decay of $H \rightarrow hh$ dominates the decay products of H for $2m_h < m_H < 2m_t$. However, as the alignment limit is approached, the branching ratio for $H \rightarrow hh$ is diminished, and at the same time decays to $\gamma\gamma$ and $\tau\tau$ become increasingly important because they do not decouple. The same is true for the pseudoscalar Higgs A with respect to the decays to Zh and $\gamma\gamma$ and $\tau\tau$ respectively. This feature strongly motivates a search for diphoton production in the high energy range (above 150 GeV) at the LHC.

In order to have an estimation of the intensity of a possible H$\rightarrow \gamma\gamma$ ($A \rightarrow \gamma\gamma$) signal in the high mass region, one can look at the contours of the inclusive $BR(H \rightarrow \gamma\gamma)$ shown in Figure 2.2 for Type 1 and Type 2 2HDMs [66]. For Type 1 the rate falls with increasing $tan\beta$ due to the $tan\beta$ suppression of the production mode, while it peaks in the alignment limit because the total width drops precipitously as $H \rightarrow VV$ and $H \rightarrow hh$ decouple. In Type 2 2HDM, the rate falls more rapidly with increasing $tan\beta$ due to the suppressed production and the growing total width, though at large $tan\beta$ the contributions from $H \rightarrow b\bar{b}$ associated production cause the rate to increase once again. Inclusive production of A with $A \rightarrow \gamma\gamma$ is entirely analogous to H in its parametric scaling. Contours of the inclusive $BR(A \rightarrow \gamma\gamma)$ are shown in Figure 2.3 and scale as discussed above, with slightly different modulation as the alignment limit is approached due to the fact that there is no contribution from W loops in

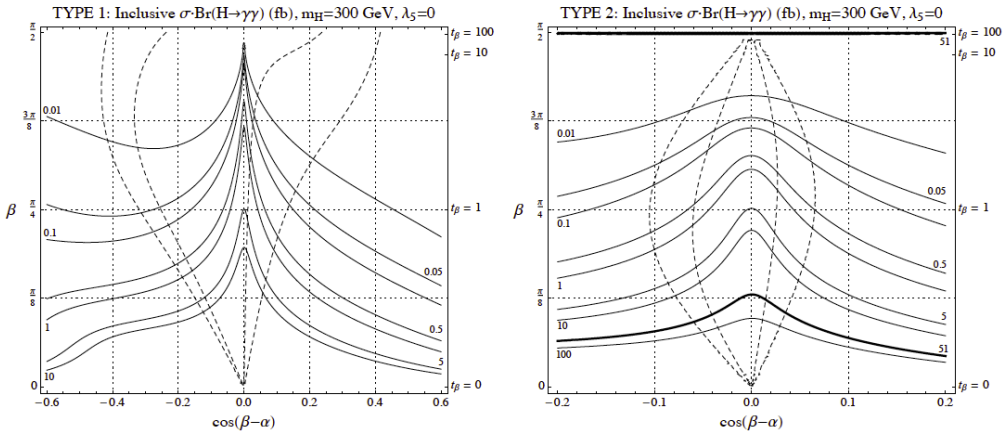


Figure 2.2. Contours of the inclusive $\sigma \cdot BR(H \to \gamma\gamma)$ in units of fb for 8 TeV pp collisions for the pseudoscalar Higgs boson with $m_H = 300$ GeV, shown as a function of $cos(\beta - \alpha)$ and β (left y-axis) and $tan\beta$ (right y-axis) for Type 1 (left) and Type 2 (right) 2HDM. The inner (outer) dashed contour denotes the 68% (95%) CL best fit to the signals of the SM-like Higgs [66].

the A effective coupling. However the values for the possible absolute cross sections time branching ratio to this kind of signals range from 0.01 up to 10 fb for both H and A decays. The model independent results obtained in this analysis are also

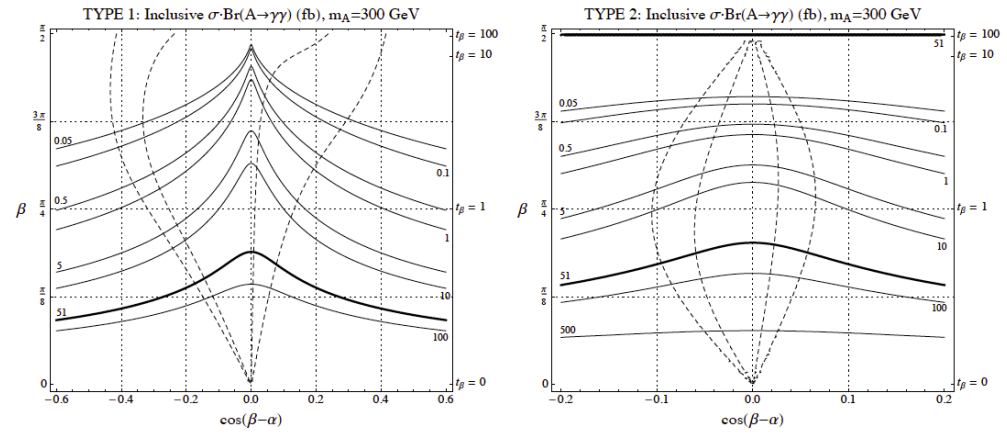


Figure 2.3. Contours of the inclusive $\sigma \cdot BR(A \to \gamma\gamma)$ in units of fb for 8 TeV pp collisions for the pseudoscalar Higgs boson with $m_A = 300$ GeV, shown as a function of $cos(\beta - \alpha)$ and β (left y-axis) and $tan\beta$ (right y-axis) for Type 1 (left) and Type 2 (right) 2HDM. The inner (outer) dashed contour denotes the 68% (95%) CL best fit to the signals of the SM-like Higgs [66].

interpreted in terms of the 2HDM space parameters and shown in Section 6.2.1.

2.1.2 Models with Extra Dimensions

Another set of models which are introduced to address the hierarchy problem are those which predict the existence of extra dimensions. These models have been proposed with two different approaches: the ADD model, named after its inventors Arkani-Hamed, Dimopoulos and Dvali [77] and the RS model, proposed by Lisa Randall and Raman Sundrum [78].

In the ADD model, the authors propose a scenario whereby the SM is constrained to the common 3+1 space-time dimensions (brane), while gravity is free to propagate throughout a larger multidimensional space (bulk). The gravitational flux on the brane is therefore diluted by virtue of Gauss Law in the bulk, which relates the fundamental Planck scale M_D to the apparent reduced scale $M_{Pl} \sim 2 \cdot 10^{18}$ GeV according to the formula $M_D^{n_{ED}+2} = \frac{M_{Pl}^2}{r^{n_{ED}}}$, where r and n_{ED} are the size and number of the extra dimensions (ED), respectively. Postulating a fundamental Planck scale to be on the order of the electroweak symmetry breaking scale (~ 1 TeV) results in ED with $r < 1$ mm for $n_{ED} \geq 2$.

Another model of ED that solves the hierarchy problem was proposed by Randall and Sundrum. In this model, the observed hierarchy is related instead to the warped anti-de Sitter (AdS) geometry of the ED. Specifically the RS1 model is considered, whereby only one finite ED exists separating two branes, one at each end of the ED. The geometry of the bulk is based on a slice of a five-dimensional AdS space with a length πr_c, where r_c is the compactification radius. The full metric is given by $ds^2 = e^{-kr_c y}\eta_{\mu\nu}dx^\mu dx^\nu - r_c^2 dy^2$, where Greek indices run over 4-dimensional space-time, $\eta_{\mu\nu}$ is the Minkowski metric tensor, and $0 \leq y \leq \pi$ is the coordinate along the single ED of radius r_c. The value of k specifies the curvature scale (or "warp factor") and relates the fundamental Planck scale on one brane to the apparent scale on the other by $\Lambda_\pi = M_{Pl}e^{-kr_c\pi}$. As a consequence, a value of $kr_c \sim 10$ would provide a natural solution to the hierarchy problem, yielding $\Lambda_\pi \sim 1$ TeV.

One observable effect of large EDs is due to a virtual graviton acting as a propagator in Drell-Yan like processes similar to the one shown in Figure 6.9, which result in production of a fermion or a boson pair in the final state. In this case, the

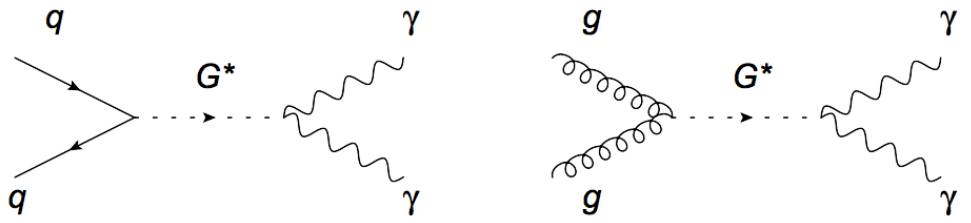


Figure 2.4. Feynman diagram for virtual graviton production through gluon fusion and decaying into two photons.

graviton-induced diagram interferes with the analogous SM diagrams and results in an enhancement of the invariant mass spectrum of the di-boson or di-fermion system, particularly at high masses, where the number of excitations contributing to this process is large [79].

Phenomenologically, the excited gravitons in both models preferentially decay into two gauge bosons, such as photons, rather than into two leptons, because the graviton has spin 2, and so fermions cannot be produced in an s wave function. In the RS scenario, gravitons appear as well separated Kaluza-Klein (KK) excitations with masses and widths determined by the parameters of the RS1 model. In the ADD model, the wave function of the KK gravitons must satisfy periodic boundary conditions, resulting in discrete energy levels with modal spacing of the order of the inverse ED size, from 1 MeV to 100 MeV, much smaller than the spacing in the RS1 model, which is expected to be of the order of 1 TeV.

2.2 Analysis Overview

The decision of investigating new physics by means of the diphoton channel lies on the fact that this final state, reconstructed with great precision using the CMS high resolution electromagnetic calorimeter, provides a clean topology, where the identification of the two high energy photons is performed with a very high efficiency. The signal events are characterized by the presence of two high transverse momentum (p_T) photons, from the decay of the new scalar. The diphoton system is fully reconstructed and the invariant mass is computed as follow:

$$m_{\gamma\gamma} = \sqrt{2p_{\gamma_1}p_{\gamma_2}(1 - cos\theta_{12})} \tag{2.2}$$

where p_{γ_1} (p_{γ_2}) is the transverse momentum of the leading(sub-leading) photon and $cos\theta_{12}$ represents the angle, in the transversal plane, between the directions of the two photons. The existence of the new particle is tested looking for local excesses over the continuum diphoton background in proximity of its predicted mass. Excellent performance in photon identification, as well as in vertex position measurement and energy resolution are the factors which determine the power of this bump search at CMS. The search is supposed to be as much inclusive as possible. This means that no requirements are applied on the presence of additional jets, leptons, or missing transverse energy in the event. Given the model-independent approach of this study, no assumption have been done on the theoretical model which predicts these diphoton resonances, and results can be interpreted in terms of different spin hypotheses (spin 0, spin 2) and different production modes (gluon fusion and vector boson fusion) as described in Section 6.2. The diphoton spectrum is studied for energies above 150 GeV in order to be far enough in energy from the already observed 125 GeV scalar thus simplifying the analysis technique. Not only the mass of the resonance is investigated, but also its natural width, looking for both narrow and wide resonances (i.e. up to 10% of the value of the mass) over the full diphoton spectrum.

The analysis is in principle simple and straightforward, since it consists in a classic bump hunt on a smooth background shape. The background contribution from SM processes is relatively low compared to hadronic searches. A brief discussion of the main SM processes which contribute to the diphoton spectrum in the energy range of interest, is provided in Section 2.2.1. However the technique used in this analysis to estimate the expected background contribution (as detailed in Section 5.4.1) is to take it directly fitting data assuming negligible signal. This approach is

inherited from the SM Higgs analysis performed in CMS [17]. Because of the use of a data driven technique, the different background contributions to the diphoton spectrum, as well as their cross sections measurement and any possible correction factor to be applied to the simulated events are not relevant in this analysis and will not be discussed.

The analysis is performed exploiting the expertise from the SM Higgs search [2] in its decay into two photons at CMS, for what concern the photon identification and the event selection and classification. A significant contribution has been given in order to improve the signal parametrization and the background estimation as described in Chapter 5. Unfortunately, due to the very low number of events in data above 1 TeV, the analysis strategy itself does not allow to extend the search to the very high mass region where a more complex strategy to estimate the misidentification rate for the QCD contribution, and thus extrapolate the background shape, is needed.

2.2.1 Sources of Background

Two types of background processes are considered as contribution to the diphoton production at the LHC. The irreducible component of the background has two real high p_T isolated photons. The reducible one has at least one non-isolated photon.

A two real photon signature can be produced by both quark-antiquark annihilation (namely Born $qq \to X$) and gluon-gluon fusion (namely Box $gg \to X$) as well as quark-gluon Compton scattering with isolated Bremsstrahlung processes (namely Brem $qg \to X$) as shown in Figure 2.5.

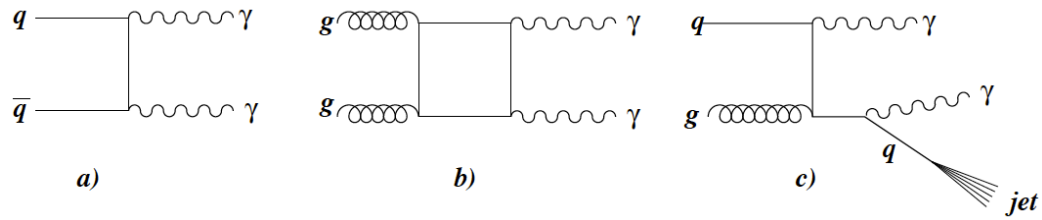


Figure 2.5. Spectator diagrams for irreducible background with two isolated photons: (a) Born $qq \to X$ process, (b) Box $gg \to X$, (c) Brem $qg \to X$

Backgrounds in which at least one final state jet is interpreted as a photon are considered as reducible. The dominant reducible backgrounds arise from QCD processes like γ+jet and dijet, which can lead to fake photons induced by neutral hadrons π^0 or produced in the jet fragmentation processes. A measurement of the isolated photon production cross section is reported in [80], while a measurement of the diphoton production cross section is published in [81].

2.3 Diphoton Searches at Colliders

The diphoton mass spectrum has been widely studied at experiments at hadron colliders to search for new particles beyond the SM. The resonant diphoton production has been studied at both the Fermilab Tevatron Collider, and the CERN Large Hadron Collider, with the progressively increasing energy collision and integrated

luminosity. Each machine in its time has therefore probed the highest masses of diphoton resonances accessible. In this section a review of the different searches performed by D0, CDF, ATLAS and CMS, their techniques, data, results, and limits on new particles decaying to photons are described in order to contextualize the search described in this thesis. The searches considered are listed in Table 2.1 in chronological order, along with a summary of the center-of-mass energy and luminosity of the dataset, and the energy range of the search. A summary of the analyses is given in this section.

Exp	Yr	$\sqrt{s}$[TeV]	$\int L[fb^{-1}]$	Mass range [GeV]	Ref.
CDF	2010	1.96	5.94	100-700	[82]
D0	2010	1.96	5.94	220-1050	[83]
ATLAS	2013	7	4.9	142-3000	[84]
CMS	2013	7	2.2	200-2000	[79]
ATLAS	2014	8	20	65-650	[85]
CMS	2014	8	19.7	150-850	[86]

Table 2.1. Summary of diphoton searches at hadron colliders.

A search for KK graviton has been performed for the first time at an hadron collider by both the CDF [82] and D0 [83] experiments. Having obtained the shapes of the invariant mass spectra of the various background sources, at D0 the background normalization is determined by fitting the invariant mass spectrum of the data to a superposition of the backgrounds in a low-mass control region (60 $< M_{\gamma\gamma} <$ 200 GeV), where KK gravitons have been excluded at the 95% C.L. by previous searches [87]. CDF uses a functional form which is a sum of the diphoton shape and the photon-like jets shape to the invariant mass spectrum of the data. The search is optimized for a narrow resonance, but still retains sensitivity to other signals which would produce an excess over SM predictions. The mass window is scanned over the mass region 100-700 GeV. In Figures 2.6 the diphoton spectra as observed by the CDF and the D0 experiments are shown.

Most recent results on graviton searches in the diphoton final state have been published by the ATLAS [84] and CMS [88] collaborations at the LHC. The techniques used for the background estimation and their associated uncertainties have been also significantly improved exploiting data-driven techniques for the estimation of the reducible γ +jets background. The quoted results show an overall improvement of a factor 2 to 3 in the excluded production cross section of diphoton events predicted by the Randall-Sundrum model, corresponding to a graviton mass limit which is 10% to 15% larger than limits set by CDF and D0 experiments. A comparison of the observed invariant mass spectrum of diphoton events and the background expectation from both CMS and ATLAS is shown in Figures 2.7. The results presented by the two collaborations are of the same order of magnitude. It is interesting to note that the two direct observables chosen for the interpretation of the results of these searches within the RS scenario, are the mass of the lightest graviton excitation, M_1, which is in the TeV range, and the dimensionless coupling to the SM fields, $\tilde{k} = k/M_{Pl}$ [89]. Figure 2.8 shows the excluded regions in the $M_1 - \tilde{k}$ plane as obtained from the CMS experiment. Also shown are bounds due to precision electroweak measurements and

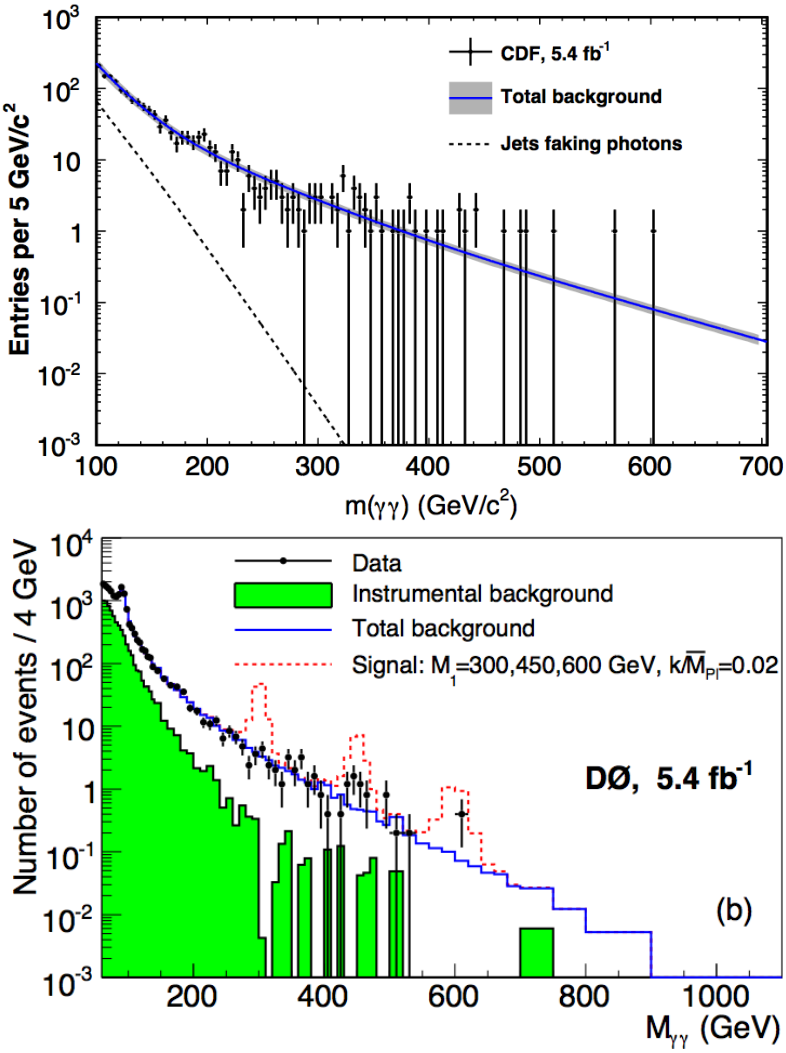


Figure 2.6. Top: The diphoton invariant mass spectrum with the fitted total background overlaid as presented by the CDF collaboration [82] . The points are the data. The dotted line shows the contribution from events where at least one selected photon is from a misidentified jet, and the solid line shows this background plus the diphoton SM distribution. Bottom: Invariant mass spectrum from $\gamma\gamma$ data (points) at D0 experiment [83]. Superimposed are the fitted total background shape from SM processes including instrumental background (open histogram) and the fitted contribution from events with misidentified photons (shaded histogram). The open histogram with dashed line shows the signals expected from KK gravitons with M_1 = 300, 450, 600 GeV (from left to right) and k/M_{Pl} = 0.02 on top of the total background. Invariant masses below 200 GeV are taken as the control region.

to naturalness arguments [90].

Recently model independent searches for diphoton resonances X have been finally performed on the yet unexplored "low" energy range by both the ATLAS [85] and the CMS [86] experiments on the 8 TeV collision data. The analyses extend the methods developed for the measurement of the Higgs couplings in the H$\rightarrow \gamma\gamma$ channel [2], [3]

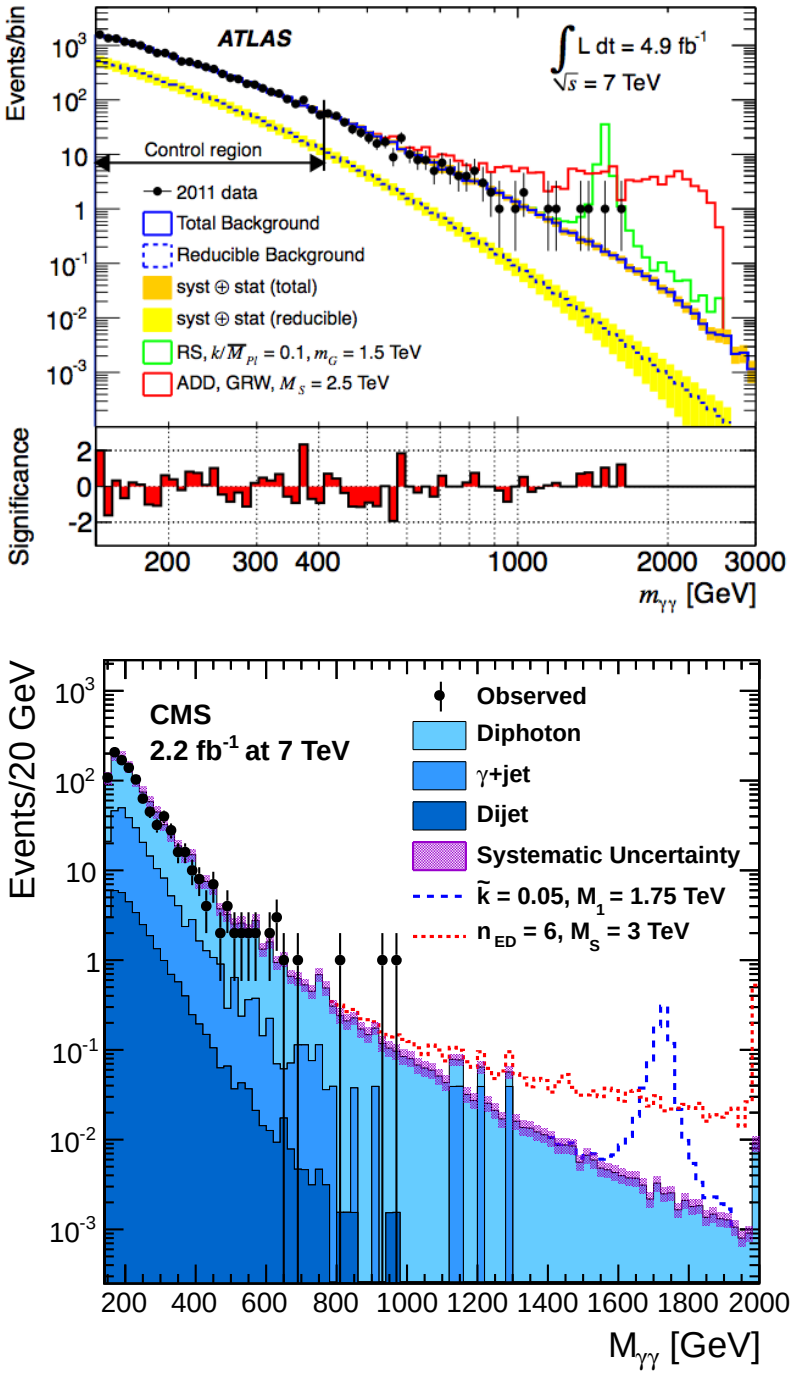


Figure 2.7. Observed invariant mass distribution of diphoton events from the 7 TeV ATLAS analysis [84] (top) and CMS analysis [88] (bottom). Superimposed are the SM background expectation and the expected signals for an example each for RS and ADD models.The bin-by-bin significance of the difference between data and predicted background is shown below the main plot.

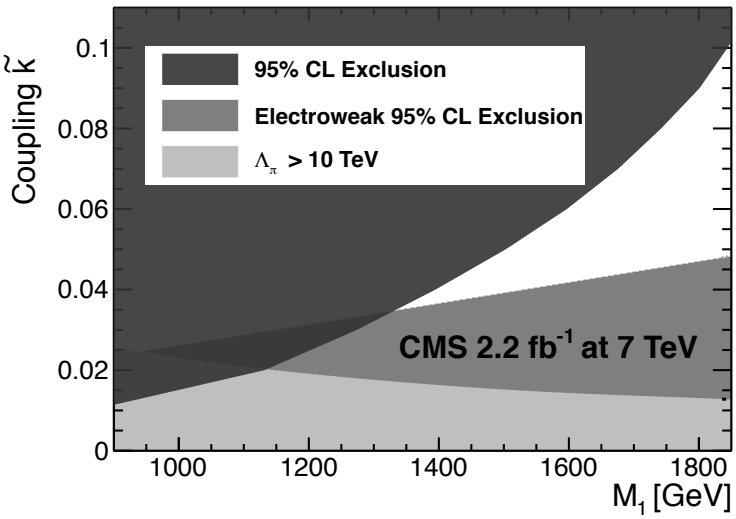


Figure 2.8. The 95% CL exclusion region for the RS1 graviton model in the $M_1 - \tilde{k}$ plane by the CMS experiment [88] . The expected limits coincide very closely with the measured limits and so are not shown in the figure. Also shown are bounds due to electroweak constraints and naturalness ($\Lambda_\pi > 10$ TeV).

to search for diphoton resonances in the wider mass range. The narrow resonances search from ATLAS is performed in the $65 \leq m_{\gamma\gamma} \leq 650$ GeV by fitting analytical descriptions of the signal and background distributions to the measured diphoton invariant mass spectrum. The background yield in the signal region is estimated by interpolation between the $m_{\gamma\gamma}$ sidebands. In the low-mass search, the most challenging background is the Drell-Yan production, dominated by the Z boson resonance where both electrons are misidentified as photons, mostly classified as converted photons. The loss of signal sensitivity is mitigated by separating the events into three categories, according to the number of converted photon candidates in each event. To determine the continuum background shape in the high mass range, an analytical function is fitted within a sliding $m_{\gamma\gamma}$ window. Figure 2.9 shows the results of several background-only fits to the data in the high-mass and low-mass analyses. The observed and the expected limits, shown in Figure 2.10, are consistent with the absence of a signal. The limits on $\sigma \cdot BR(X \rightarrow \gamma\gamma)$ for an additional scalar resonance range from 90 fb at the low end of the search interval, to 1 fb for $m_X = 600$ GeV.

The diphoton model independent search performed by CMS is the topic of this thesis and will be described in detail in the following chapters. It will be shown that the sensitivity of the CMS and the ATLAS analyses are similar in the high energies range. The improvement in the CMS search relies in extending the search for narrow resonances to resonances with widths up to 10% of the value of the mass of the hypothetical signal. The upper limits set on the absolute cross section times branching fraction of the X$\rightarrow \gamma\gamma$ process are given in the two-dimensional plane mass vs natural width of the signal, as will be described in Chapter 6.

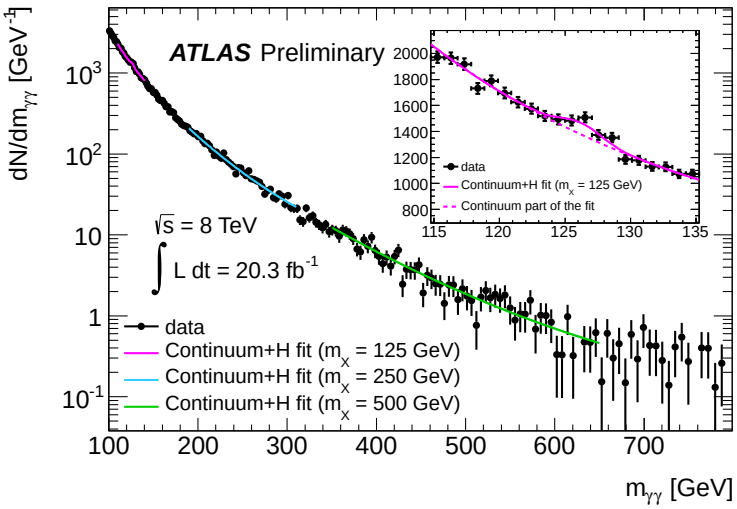


Figure 2.9. Background-only fits to the data (black dots) as functions of the diphoton invariant mass $m_{\gamma\gamma}$ for the inclusive high and low mass analyses at ATLAS [85]. The solid line shows the sum of the Higgs boson and the continuum background components, while the dashed line shows the continuum background component only.

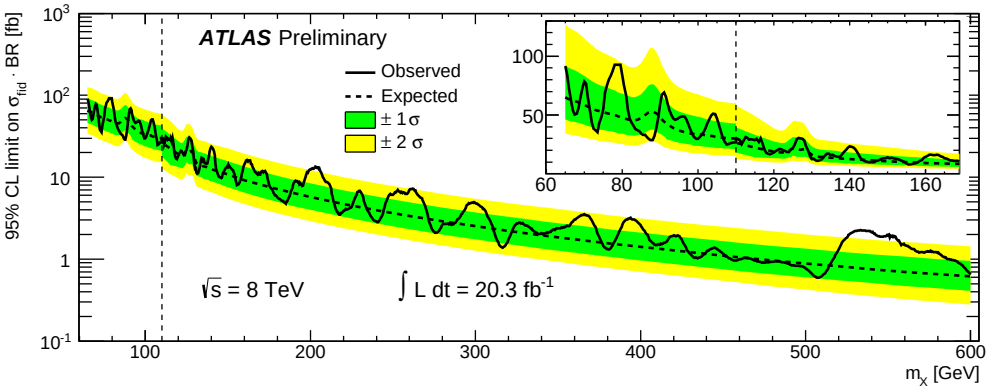


Figure 2.10. Observed and expected 95% CL limit on the fiducial cross-section times branching fraction as a function of m_X in the range $65 \leq m_{\gamma\gamma} \leq 650$ GeV at ATLAS [85]. The discontinuity in the limit at $m_X = 110$ GeV (vertical dashed line) is due to the transition between the low-mass and high-mass analyses. The dark and light shaded bands show the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainties on the expected limit. The inset shows a zoom of the limit in the $65 \leq m_{\gamma\gamma} \leq 170$ GeV range.

Chapter 3

The Large Hadron Collider and the CMS experiment

This chapter gives a brief overview of the most important features of the Large Hadron Collider and the Compact Muon Solenoid experiment. A concise description of the LHC is discussed in Section 3.1. The descriptions of the CMS experiment and its main sub-detectors are given in Section 3.2.

3.1 Large Hadron Collider

The Large Hadron Collider [91], at the CERN laboratories of Geneva, is a proton-proton (pp) collider built to work at the design center of mass energy of $\sqrt{s} = 14$ TeV, with a bunch crossing every 25 ns and a design luminosity of 10^{34} cm^{-2} s^{-1}. The LHC is installed in the same circular underground tunnel occupied until the year 2000 by the Large Electron Positron collider (LEP). The pp collision are used, instead of the e^+e^- one of LEP, to reduce the synchrotron radiation, in order to accelerate the particles up to a very large energy. It was preferred to a $p\bar{p}$ collider because it allows to reach higher rate of events. In fact the low anti-proton production efficiency (10^5 protons are needed to create an anti-proton) and larger time needed to accumulate them, would make almost impossible to reach the high design luminosity of the LHC.

The luminosity L is the parameter to quantify the performances of a collider, because the event rate R_i of a given process i, defined as the number of events occurring per unit of time, can be written as:

$$R_i = \frac{dN_i}{dt} = L \cdot \sigma_i \tag{3.1}$$

where σ_i is the cross section of the process i. The luminosity depends only on the machine parameters. Assuming a small crossing angle between the beams and Gaussian-shaped beam bunches, the luminosity L can be written as:

$$L = \frac{f n_b N^2}{4\pi\sigma^2} \tag{3.2}$$

where f is the revolution frequency of particle bunches, n_b is the number of bunches rotating in the accelerator, N is the number of protons in the two colliding bunches,

σ is the RMS of beam profile distributions in the plane orthogonal to the beam direction.

The LHC works as follows. A small linear accelerator (Linac2) will bring the proton beams up to an energy of 50 MeV and then the Proton Synchrotron Booster (PSB) will further accelerate protons up to 1.4 GeV. After passing through the Proton Synchrotron (PS) at 50 GeV, the beam will be finally injected in the Super Proton Synchrotron (SPS) where it will reach the energy of 450 GeV, the last stage before entering the LHC (see Figure 3.1).

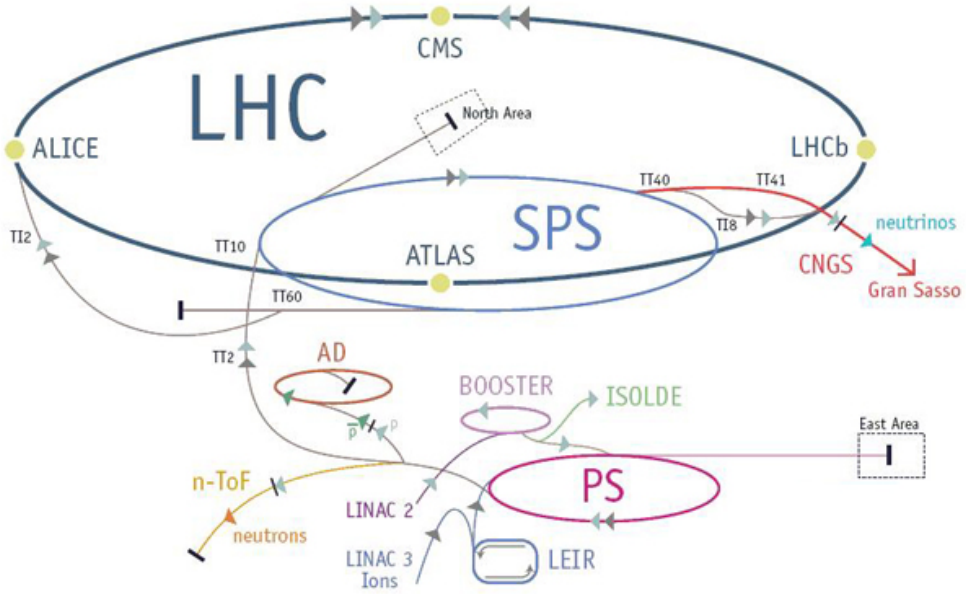


Figure 3.1. An overall view of particle accelerators at CERN [92].

In the LHC design, 1232 main dipole magnets (made of niobium-titanium superconductor chilled with superfluid Helium at 1.9 K) generating a magnetic field up to 8.3 T, will be used to steer the particles into curvilinear trajectories. The two beams will run in two contiguous pipes with vacuum inside, separated by 19.4 cm, that will be unified in proximity of the interactions points, where the experiments will be placed. Because of the high luminosity of the LHC, large thermal power will be generated near the pipes due to the synchrotron radiation, making necessary the presence of a suitable cooling system. For this reason also the pipes will be in contact with superfluid Helium at 1.9 K.

3.2 The Compact Muon Solenoid Overall Concept

The Compact Muon Solenoid (CMS) is one of the two general purpose experiments which takes data at the LHC. Its physics goals range from the search for the Higgs boson to the searches for physics beyond the Standard Model, to the precision measurements of already known particles and phenomena [93].

The overall layout of CMS is shown in 3.2. The inner tracker and the two

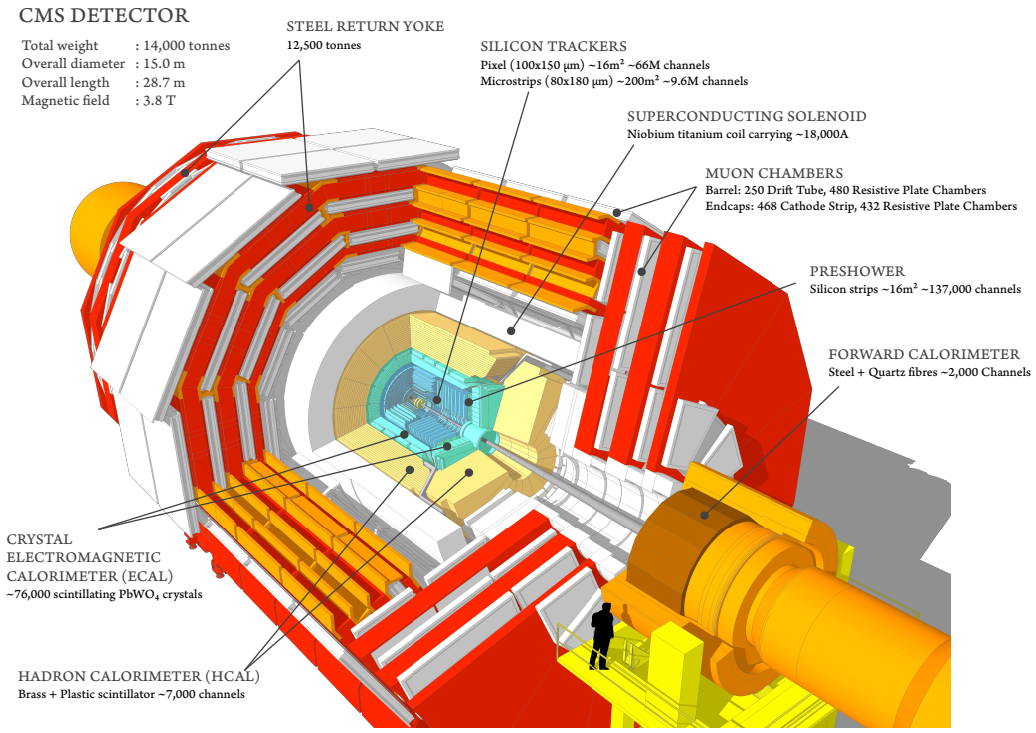


Figure 3.2. An exploded view of the CMS detector.

calorimeters of CMS are located inside a 13 m-long, 5.9 m inner diameter, 3.8 T superconducting solenoid. In order to achieve good momentum resolution within a compact spectrometer without making stringent demands on muon-chamber resolution and alignment, a high magnetic field was chosen. The return field is large enough to saturate 1.5 m of iron, allowing four muon stations to be integrated to ensure robustness and full geometric coverage. The central part of CMS is called *barrel* while the two edges of the detector are denoted as *endcaps*.

The tracking volume is given by a cylinder of length 5.8 m and diameter 2.6 m. In order to deal with high track multiplicities, CMS employs 10 layers of silicon microstrip detectors, which provide the required granularity and precision. In addition, 3 layers of silicon pixel detectors are placed close to the interaction region to improve the measurement of the impact parameter of charged-particle tracks, as well as the position of secondary vertexes. The electromagnetic calorimeter (ECAL) uses lead tungstate ($PbWO_4$) crystals with coverage in pseudorapidity up to $|\eta| < 3.0$. A preshower system is installed in front of the edges of ECAL for π^0 rejection.

In the following, the CMS sub-detectors are described from the innermost region (the closest to the interaction point) to the outermost region. The chapter ends with a description of the trigger and data acquisition systems.

3.3 Coordinate Conventions

The coordinate system adopted by CMS has the origin centered at the nominal collision point inside the experiment, the y-axis pointing vertically upward, and the x-axis pointing radially inward toward the center of the LHC. Figure 3.3 shows the coordinate system in CMS.

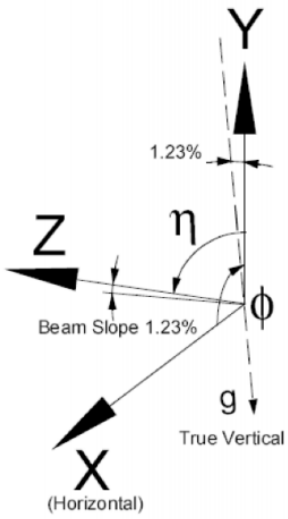


Figure 3.3. Coordinate system in CMS [92].

The azimuthal angle ϕ is measured from the x-axis in the x-y plane. The polar angle θ is measured from the z-axis. Pseudorapidity is defined as

$$\eta = -\ln\tan(\theta/2) \tag{3.3}$$

The value $\eta = 0$ corresponds to a direction perpendicular to the beamline, while the limit $\eta = \infty$ gives a direction parallel to the beamline.

The momentum and energy measured transverse to the beam direction, denoted by p_T and E_T, respectively, are computed as follow:

$$p_T = psin\theta \tag{3.4}$$

$$E_T = Esin\theta \tag{3.5}$$

Finally, particles which escape the detection leave an imbalance in the transverse plane which is quantified as missing transverse energy in the following way:

$$E_T^{miss} = -\sum_i p_T^i \tag{3.6}$$

as the negative vectorial sum of the transverse momentum of all the visible particles in the event.

3.4 Inner Tracking System

The tracker [94], placed within the magnetic field, is the subdetector which is closer to the interaction point. It is dedicated to track and vertex finding. The silicon (Si) technology has been chosen for the whole tracker in order to provide good radiation hardness, high granularity and large hit redundancy to perform a good pattern recognition. The layout of the CMS tracker is shown in Figure 3.4. Close to the interaction vertex, in the barrel region, are 3 layers of hybrid pixel detectors at a radius (r) of about 4, 7 and 10 cm. The size of the pixel detector is $100 - 150$ m^2. In the barrel part, the Si microstrip detectors are placed at r between 20 and 110 cm. The forward region has 2 pixel and 9 microstrip layers in each of the two endcaps. In order to avoid excessively shallow track crossing angles, the Inner Barrel is shorter than the Outer Barrel, and there are additional three Inner Disks in the transition region between barrel and endcaps, on each side of the Inner Barrel. The total area of the Si detectors is around 200 m^2, providing a coverage up to $\eta = 2.5$. The material budget inside the active volume of the tracker increases from 0.4 X_0 at $\eta = 0$ to around 1 X_0 at $|\eta| = 1.6$, before decreasing to 0.6 X_0 at $|\eta| = 2.5$.

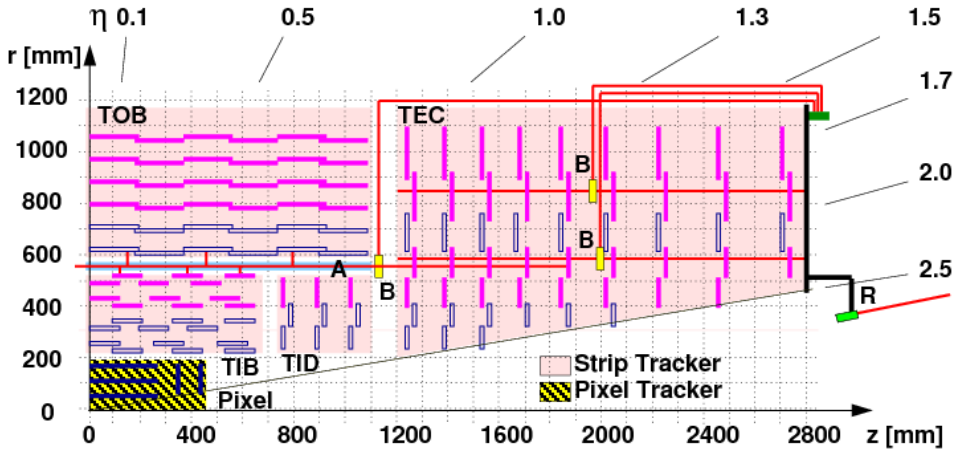


Figure 3.4. A global view of the tracker layout.

3.5 Electromagnetic Calorimeter

The electromagnetic calorimeter plays an essential role in the study of the physics of electroweak symmetry breaking, and in the exploration of beyond the Standard Model scenarios. ECAL is a homogeneous calorimeter of almost 76000 Lead Tungstate $PbWO_4$ scintillating crystals divided into a barrel and two endcaps. A 3D view of the barrel and endcap electromagnetic calorimeter is shown in Figure 3.5.

3.5.1 The Barrel Calorimeter

The barrel part of the ECAL covers the pseudorapidity range $|\eta| < 1.479$ (see Figure 3.6). The front face of the crystals is at a radius of 1.29 m and each crystal has a square cross-section of $22 \times 22\,\text{mm}^2$ and a length of 230 mm corresponding

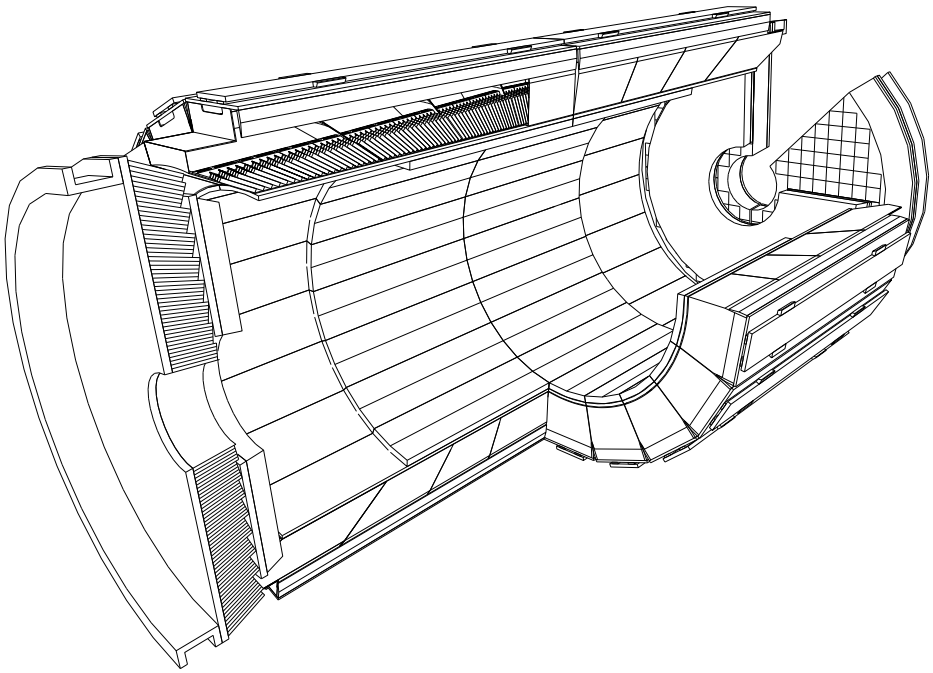


Figure 3.5. A 3D view of the electromagnetic calorimeter.

to $25.8\,X_0$. The truncated pyramid-shaped crystals are mounted in a geometry which is off-pointing with respect to the mean position of the primary interaction vertex, with a 3° tilt in both ϕ and in η. The crystal cross-section corresponds to $\Delta\eta \times \Delta\phi = 0.0175 \times 0.0175$ (1°). The barrel granularity is 360-fold in ϕ and

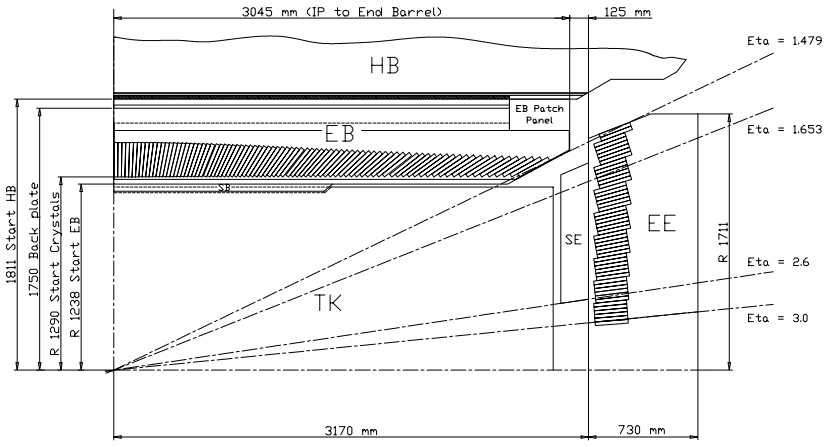


Figure 3.6. Longitudinal section of the electromagnetic calorimeter (one quadrant)

(2×85)-fold in η, resulting in a total number of 61 200 crystals. The crystal volume in the barrel amounts to $8.14\,m^3$ (67.4 t). Crystals for each half-barrel are grouped in 18 supermodules each subtending 20° in ϕ. Each supermodule comprises four modules with 500 crystals in the first module and 400 crystals in each of the remaining

three modules. For simplicity of construction and assembly, crystals have been grouped in arrays of 2×5 crystals which are contained in a very thin wall ($200\,\mu$m) alveolar structure and form a submodule. Thermal regulation is carried out by two active systems: 1) a specially regulated cooling circuit which keeps the operating temperature (ambient temperature) of the crystal array and of the APDs within a tight temperature spread of $\pm 0.05\,^{\circ}$C, ensuring adequate thermal stability; 2) the power cooling circuit evacuates the heat generated by all power sources in the supermodule (each supermodule is designed as a separate thermal entity).

3.5.2 The Endcap Calorimeter

The endcap part of the crystal calorimeter covers a pseudorapidity range from 1.48 to 3.0. The design of the endcaps provides precision energy measurement up to $|\eta| = 2.5$. Crystals are however installed up to $|\eta| = 3$ in order to augment the energy-flow measurement in the forward direction. The mechanical design of the endcap calorimeter is based on an off-pointing pseudo-projective geometry using tapered crystals of the same shape and dimensions ($24.7 \times 24.7 \times 220\,\mathrm{mm}^3$) grouped together into units of 36, referred to as supercrystals. A total of 268 identical supercrystals is used to cover each endcap with a further 64 sectioned supercrystals used to complete the inner and outer perimeter. Each endcap contains 14648 crystals, corresponding to a volume of $1.52\,\mathrm{m}^3$ (12.6 t). Both endcaps are identical. Each endcap detector is constructed using Dee-shaped sections as seen in Figure 3.7.

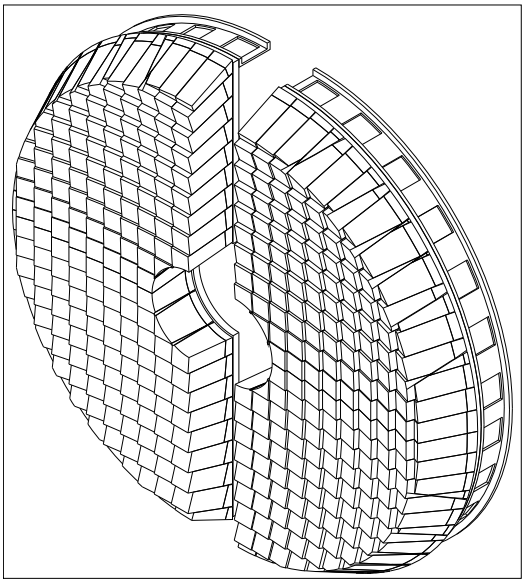


Figure 3.7. A single endcap with Dees apart.

Figure 3.8 shows the total thickness (in radiation lengths) of the ECAL as a function of pseudorapidity. The endcap part also includes the preshower detector. Because of the high radiation levels in the endcaps all materials used in this region must tolerate very large doses and neutron fluences.

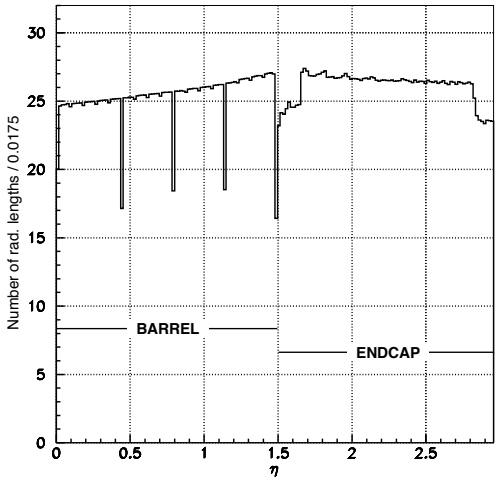


Figure 3.8. Total thickness in X_0 of the ECAL as a function of pseudorapidity, averaged over ϕ.

3.5.3 The Preshower Detector

The endcap preshower covers a pseudorapidity range from $|\eta| = 1.65$ to 2.61. Its main function is to provide π^0-γ separation. The preshower detector, placed in front of the crystals, contains two lead converters of a total thickness of $2\,X_0$ and $1\,X_0$ respectively, followed by detector planes of silicon strips with a pitch of $< 2\,\text{mm}$. The impact position of the electromagnetic shower is determined by the center-of-gravity of the deposited energy. The accuracy is typically $300\,\mu\text{m}$ at 50 GeV. In order to correct for the energy deposited in the lead converter, the energy measured in the silicon is used to apply corrections to the energy measurement in the crystal. The fraction of energy deposited in the preshower (typically 5% at 20 GeV) decreases with increasing incident energy. Figure 3.9 shows the layout of the preshower.

To maintain its performance during the lifetime of the experiment, the endcap silicon detector has to be operated at $-5\,^\circ\text{C}$. Heating films and insulating foam glued on the moderators guarantee that the external surfaces are kept at the ambient temperature of the neighboring detectors.

3.5.4 Lead Tungstate Crystals

The characteristics of the Lead Tungstate crystals ($PbWO_4$) make them an appropriate choice for operation at LHC [95]. The high density ($8.3\,\text{g/cm}^3$), short radiation length (0.89 cm) and small Molière radius (2.2 cm) results in a fine granularity and a compact calorimeter. The scintillation decay time is of the same order of magnitude as the LHC bunch crossing time: about 80% of the light is emitted in 25 ns. The light output is relatively low: about 4.5 photoelectrons per MeV are collected in both the avalanche photodiodes (APDs) and the vacuum phototriodes (VPTs), where the higher APD quantum efficiency is balanced by their smaller surface coverage on the back face of the crystal. The crystals emit blue-green scintillation light with a broad maximum at 420 nm [96].

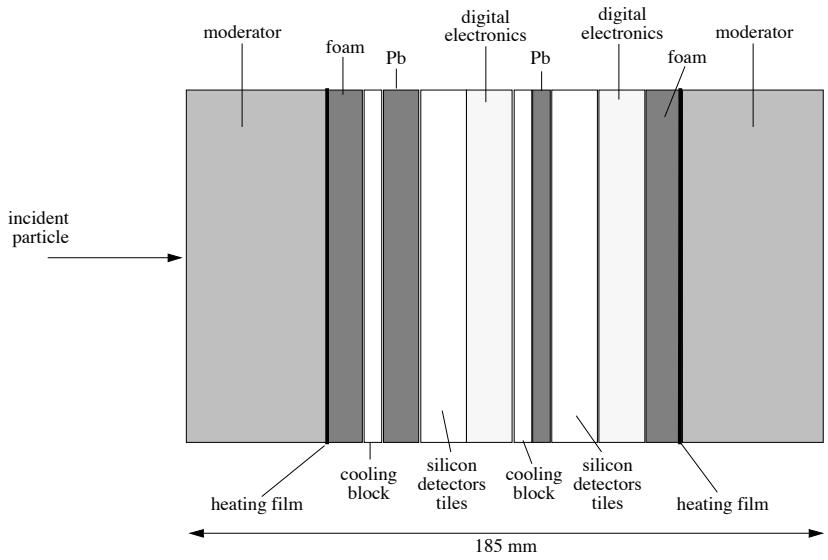


Figure 3.9. Schematic section through the endcap preshower.

3.5.5 Amplitude reconstruction

The raw data for a single channel consists of a series of consecutive digitization of the signal making up a time frame [97]. The number of samples is adjustable (2+4n) with a default of 10. The digitizations are made at the bunch crossing frequency of 40 MHz, i.e. one sample each 25 ns. In addition, the timing of the signal is adjusted in LHC running so that the signal pulse maximum corresponds to one of the samplings. Figure 3.10 shows an example of the time sampling for a signal pulse as a function of the time difference $(T - T_{max})$, where T and T_{max} indicate the time of the generic ADC sample and the time corresponding to the maximum of the pulse shape respectively. The simplest method of reconstructing the amplitude of the channel is to take the sampling on the maximum as the measurement of the signal. However, a larger number of samples is preferred since it allows more sophisticated digital processing of the signal to reduce noise contribution. The other reason is to enable the identification of pile-up events from other bunch crossing. The signal amplitude is computed as a linear combination of discrete time samples:

$$A = \sum_{i=0}^{N} S_i \times w_i \tag{3.7}$$

where w_i are the weights, S_i the time sample values in ADC counts and N is the number of samples used in the filtering. The weights are determined to minimize the noise contribution. Amplitude and time measurement are strongly correlated. After a signal has been amplified and shaped by the front-end electronics, the channel timing reconstruction consists in a precise measurement of the time the pulse reaches its maximum values A_{max}. Looking at Figure 3.10, the reconstructed time of a channel corresponds to the value of T_{max}.

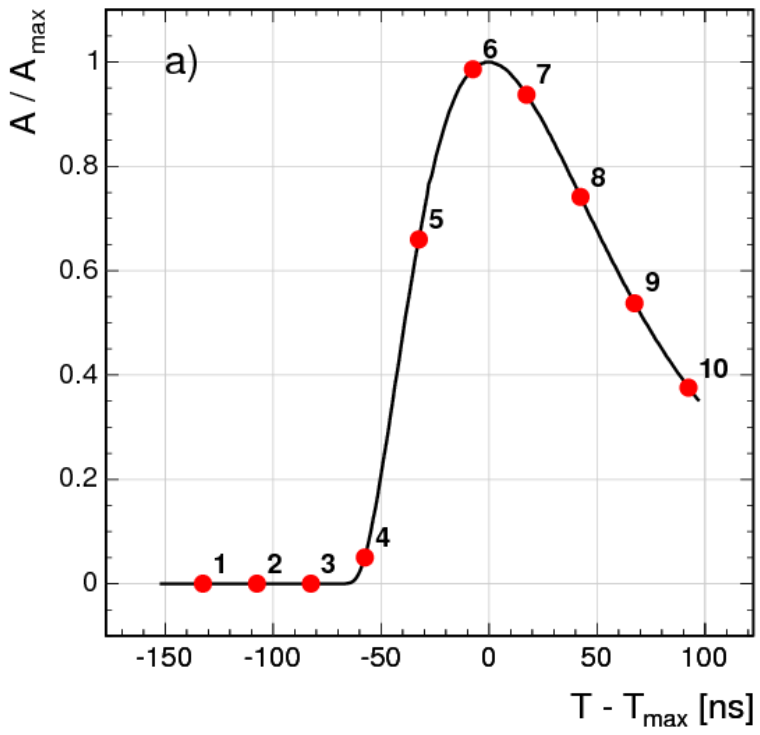


Figure 3.10. Pulse shape measured in the ECAL as a function of $(T - T_{max})$.

3.5.6 Energy Resolution

For the energy range of about 25 GeV to 500 GeV, the ECAL energy resolution has been parameterized as:

$$\frac{\sigma(E)}{E} = \frac{a}{\sqrt{E}} \oplus \frac{\sigma_n}{E} \oplus c \quad \text{(E in GeV)} \tag{3.8}$$

where a is the stochastic term, σ_n the noise, and c the constant term. Figure 3.11 summarizes the different contributions expected for the energy resolution. Terms representing the degradation of the energy resolution at extremely high energies have not been included. The stochastic term includes fluctuations in the shower containment as well as a contribution from photostatistics. The noise term contains the contributions from electronic noise and pile-up energy; the former is quite important at low energy, the latter is negligible at low luminosity. The curve labeled *intrinsic* includes the shower containment and a constant term of 0.55%. The constant term must be kept down to this level in order to profit from the excellent stochastic term of $PbWO_4$ in the energy range relevant for the search for new physics. To achieve this goal, in situ calibration/monitoring using isolated high p_T electrons is performed.

3.5.7 ECAL Time measurement and resolution

The algorithm used to extract T_{max} relies on an alternative representation of the pulse shape, provided by a a variable defined as the ratio between the amplitudes of

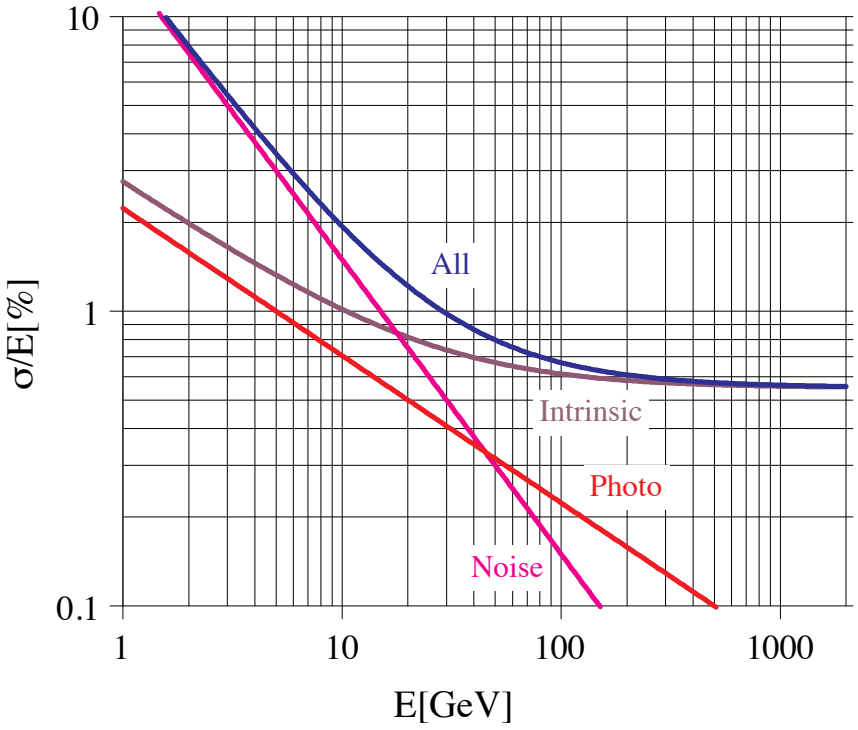


Figure 3.11. Different contributions to the energy resolution of the $PbWO_4$ calorimeter.

two consecutive samples [97]:

$$R(T) = \frac{A(T)}{A(T + 25ns)} \tag{3.9}$$

where $A(T)$ represents the pulse amplitude at time T . Figure 3.12 illustrates the parametrization of time difference $(T - T_{max})$ as a function of $R(T)$. In view of the universal character of the pulse shape, this representation is independent on the maximum amplitude A_{max} and can be described well with a simple polynomial parametrization.

Each pair of consecutive samples gives a measurement of the ratio:

$$R_i = \frac{A(T + [i] \cdot 25ns)}{A(T + [i + 1] \cdot 25ns)} \tag{3.10}$$

An estimate $T_{max,i}$ of the maximum time can be obtained from each R_i ratio as $T_{max,i} = T_i - T(R_i)$. A more precise determination of the maximum time and its uncertainty is then obtained from the weighted average of the estimated $T_{max,i}$:

$$T_{max} = \frac{\sum_i \frac{T_{max,i}}{\sigma_i^2}}{\sum_i \frac{1}{\sigma_i^2}} \tag{3.11}$$

$$\frac{1}{\sigma_T^2} = \sum_i \frac{1}{\sigma_i^2} \tag{3.12}$$

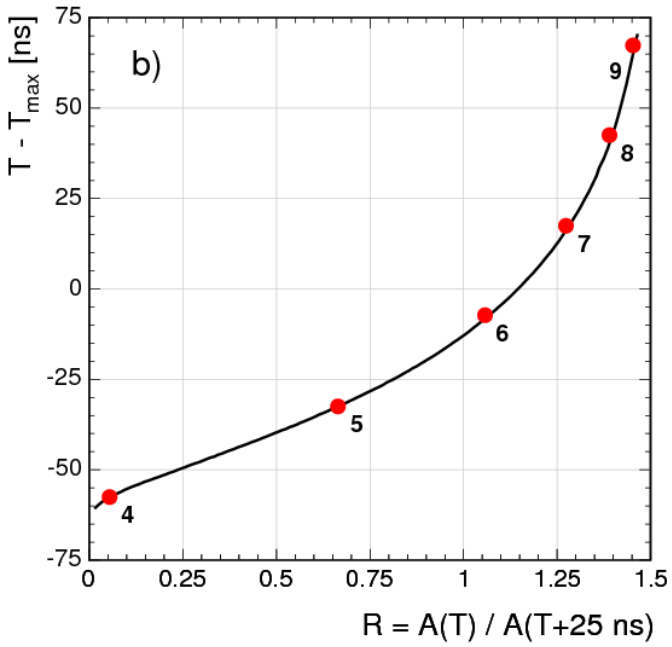


Figure 3.12. Time difference $(T - T_{max})$ as a function of the ratio of the amplitudes $R(T)$.

The typical number of available ratios R_i is five or six.

To determine the intrinsic time resolution of the ECAL, electrons from a test beam with energy between 15 and 300 GeV are used. The time resolution is extracted from the distribution of the time difference between adjacent crystals that share the same electromagnetic shower and measure similar energies. The distribution of the time difference is well described by a Gaussian function, whose width can be parametrized as [97]:

$$\sigma^2(t_1 - t_2) = \left(\frac{N\sigma_n}{A_{eff}}\right)^2 + C^2 \tag{3.13}$$

where $A_{eff} = E_1 E_2 / \sqrt{E_1^2 + E_2^2}$, with $t_{1,2}$ and $E_{1,2}$ corresponding to the times and energies measured in the two crystals, σ_n is a parameter related to the noise level, N and C represent the noise and constant term coefficients of time resolution. The extracted width is presented in Figure 3.13 as a function of the variable A_{eff}/σ_n. The energy scales for barrel and endcap are superimposed in the plot.

Precise ECAL time determination results to be crucial in many respects. The better the precision of time measurement and synchronization, the larger the rejection of backgrounds with a broad time distribution. Such backgrounds are cosmic rays, beam halo muons, electronic noise, and out-of- time proton-proton interactions. Precise time measurement also makes it possible to identify particles predicted by different models beyond the Standard Model. Slow heavy charged R-hadrons [98], which travel through the calorimeter and interact before decaying, and photons from the decay of long-lived new particles that reach the calorimeter out-of-time with respect to particles travelling at the speed of light from the interaction point. To achieve these goals the time measurement performance both at low energy (1 GeV or

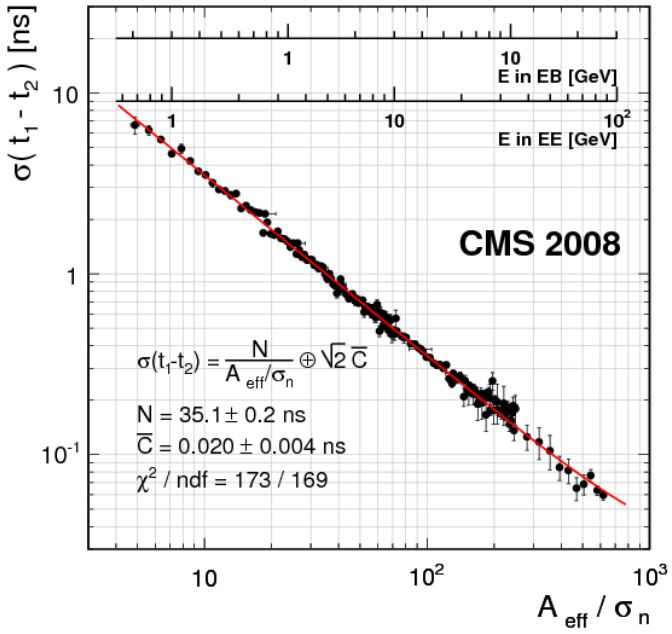


Figure 3.13. Gaussian width of the time difference between two neighboring crystals as a function of the variable A_{eff}/σ_n, for test beam electrons between 15 and 300 GeV. The equivalent single-crystal energy scales for barrel and endcaps are overlaid on the plot.

less) and high energy (several tens of GeV) for showering photons have been studied also with collision data in order to reproduce the results obtained at the test beam. Indeed during collisions several effects can worsen significantly the design resolution, as for example, run by run variations, inter-calibration, effects vs energy, radiation, the presence of the magnet field and of the tracker material in front of the ECAL. The time resolution estimated with 8 TeV data, looking at the time difference of electrons produced in the decay of a Z boson in CMS, is shown in Figure 3.14. The noise term results to be consistent with the one obtained at the test beam, while the constant term is about 150 ps. This value although far from the design performance, it is almost ok for every physics application. An example of the use of the ECAL time measurement in physics analysis is presented in appendix A. In this analysis, for the first time, the novel technique of exploiting the ECAL time measurement is used to identify off time photons produced in the decays of long-lived neutralinos, with decay lengths comparable to the ECAL radial size.

3.6 Magnet

The required performance of the muon system, and hence the bending power, is defined by the narrow states decaying into muons and by the unambiguous determination of the sign for muons with a momentum of 1 TeV/c. This requires a momentum resolution of $\Delta p/p \sim 10\%$ at $p = 1$ TeV.

To achieve this goal, CMS chose a large superconducting solenoid, the parameters of which are given in table 3.1.

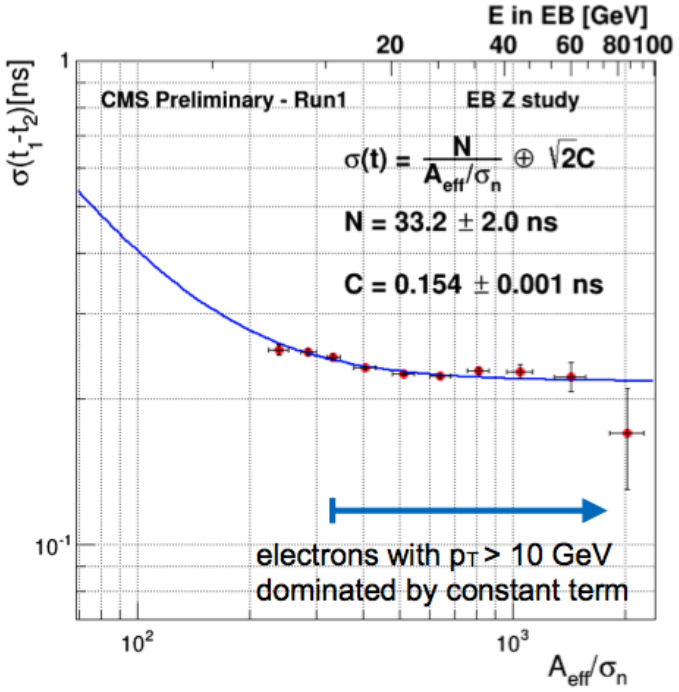


Figure 3.14. Gaussian width of the time difference between two electrons produced in the decay of a Z boson as a function of the variable A_{eff}/σ_n. The equivalent single-crystal energy scales for barrel and endcaps are overlaid on the plot.

Parameter	Value
Field	3.8 T
Inner bore	5.9 m
Length	12.9 m
Number of turns	2168
Current	19.5 kA
Stored energy	2.7 GJ

Table 3.1. Parameters of the CMS superconducting solenoid.

3.7 Hadron Calorimeter

The design of the hadron calorimeter (HCAL) [99] is strongly influenced by the choice of the magnet parameters since most of the CMS calorimetry is located inside the magnet coil and surrounds the ECAL system (see figure 3.15). An important requirement of HCAL is to minimize the non-Gaussian tails in the energy resolution and to provide good containment and hermeticity. Hence, the HCAL design maximizes material inside the magnet coil in terms of interaction lengths. This is complemented by an additional layer of scintillators, referred to as the hadron outer (HO) detector, lining the outside of the coil. Brass has been chosen as absorber material as it has a reasonably short interaction length, is easy to machine and is non-magnetic. Maximizing the amount of absorber before the magnet requires

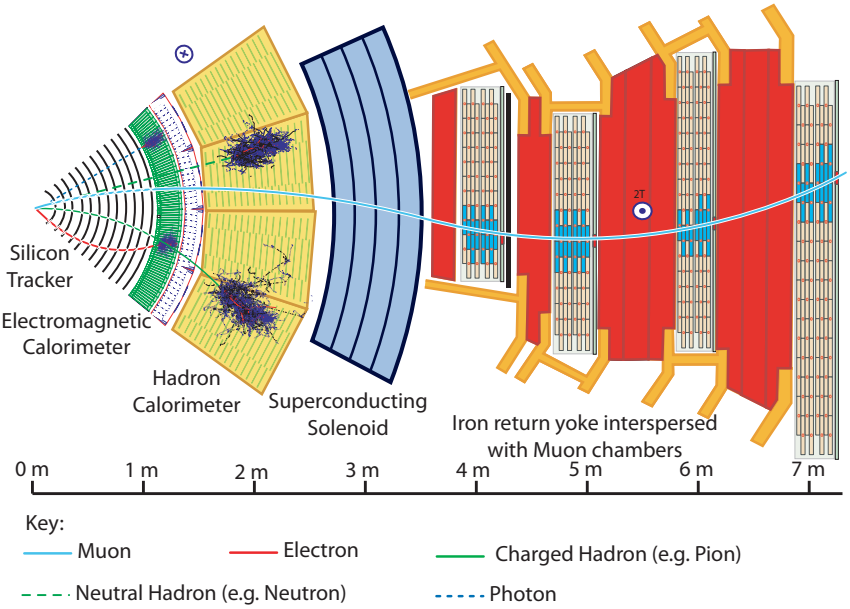


Figure 3.15. Schematic view of a transverse slice of the central part of the CMS detector.

minimizing the amount of space devoted to the active medium. The tile/fiber technology makes for an ideal choice. It consists of plastic scintillator tiles read out with embedded wavelength-shifting (WLS) fibers. The WLS fibers are spliced to high-attenuation-length clear fibers are just outside the scintillator carrying the light to the readout system. The photodetection readout is based on multi-channel hybrid photodiodes (HPDs). The absorber structure is assembled by bolting together precisely machined and overlapping brass plates so as to leave space to insert the scintillator plates, which have a thickness of 3.7 mm. The overall assembly enables the HCAL to be built with essentially no uninstrumented cracks or dead areas in ϕ.

3.8 Muon System

The muon system is the outermost of the CMS subdetectors. Its main goals are the identification of muons, thanks to their high penetrating power, and a precise measurement of their momentum, with the help of the information coming from the tracker. The muon system also works as trigger for events which involve muons and it provides a precise time measurement of the bunch crossing. The CMS muon system [100] relies on three kinds of gaseous detectors: drift tubes (DT), cathode strip chambers (CSC) and resistive plate chambers (RPC). The DT and the CSC provide an excellent spatial resolution for the measurement of charged particle momentum; the RPC are used for trigger issues because of the very good timing. The active parts of the muon system are hosted into stations which are interleaved by the iron layers of the return yoke of the magnet. The longitudinal view of a quarter of the muon system is given in Figure 3.16. The barrel extends up to $|\eta| < 1.4$, the endcaps up to $|\eta| < 2.4$.

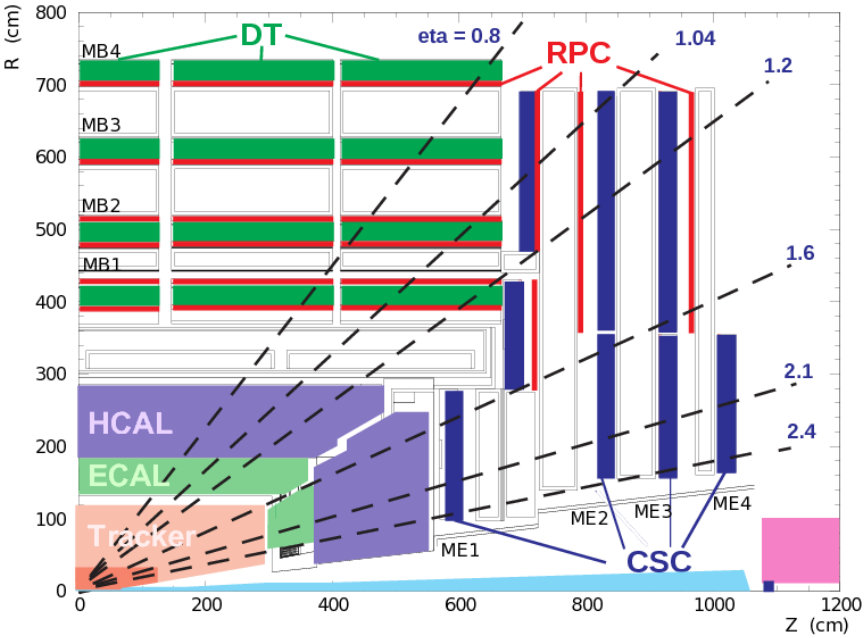


Figure 3.16. Longitudinal view of one quarter of the CMS muon system

3.9 Trigger and Data Acquisition

The trigger system in CMS is the start of the physics event selection process. A decision to retain an event for further consideration has to be made every 25 ns. This decision is based on the event's suitability for inclusion in one of the various datasets to be used for analysis. The datasets to be taken are determined by CMS physics priorities as a whole. These datasets include dilepton and multilepton datasets, diphoton datasets, lepton plus jet datasets for top, Higgs and BSM physics, and inclusive electron datasets for calorimeter calibrations. In addition, other samples are necessary for measuring efficiencies in event selection and studying backgrounds. The trigger has to select these samples in real time along with the main data samples.

For the nominal LHC design luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, an average of 17 events occurs at the beam crossing frequency of 25 ns. This input rate of 10^9 interactions every second must be reduced by a factor of at least 10^7 to 100 Hz, the maximum rate that can be archived by the on-line computer farm. CMS has chosen to reduce this rate in two steps. At the first level (L1) all data is stored for 3.2 μs, after which no more than 100 kHz of the stored events are forwarded to the High Level Triggers (HLT). The L1 system uses only coarsely segmented data from calorimeter and muon detectors, while holding all the high-resolution data in pipeline memories in the front-end electronics. The HLT is provided by a subset of the on-line processor farm which, in turn, passes a fraction of these events to the remainder of the on-line farm for more complete processing.

3.9.1 Level 1 Trigger

The design of the CMS Trigger and Data Acquisition system is illustrated in figure 3.17. At the first level all information about the event is preserved. The first

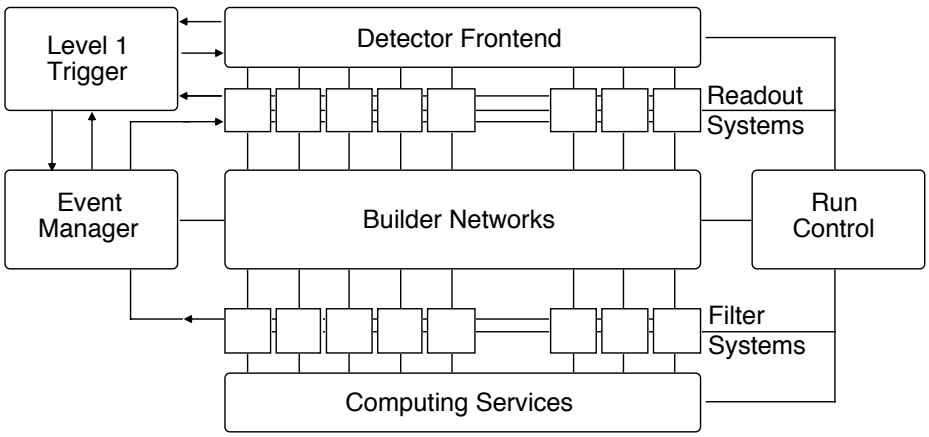


Figure 3.17. CMS Trigger and Data Acquisition System.

level decision is made, with negligible dead-time, on a subset of the total information available for the events. Since signal propagation delays are included in this pipeline time, the L1 trigger calculations must be done in many cases in less than 1 μs. If the first level trigger generates an accept, the event data are moved or assigned to a buffer for readout and processing by the High Level Triggers.

The L1 trigger involves the calorimetry and muon systems as well as some correlation of information from these systems. The L1 decision is based on the presence of local objects such as photons, electrons, muons, and jets, using information from calorimeters, and muon systems in a given element of η-ϕ space. It also employs global sums of E_T and missing E_T. Each of these items is tested against several p_T or E_T thresholds.

3.9.2 High Level Trigger

The CMS Level-1 Trigger System is required to reduce the input interaction rate of 1 GHz to a filtered event rate of 75 kHz. To match the capabilities of the mass storage and offline computing systems, the final output of the experiment should not exceed 100 events per second.

The High Level Triggers have access to all the information used in L1 since this is stored locally in the L1 trigger crates. Consequently, High Level Triggers can make further combinations and other topological calculations on the digital list of objects transmitted from L1. Eventually, the High Level Triggers use the full event data for the decision to keep an event.

Chapter 4

Photon Reconstruction and Selection

The following Chapter details the photon reconstruction and selection that is employed in the analysis. Section 4.1 gives the details of the photon reconstruction technique performed with the ECAL. Sections 4.2 and 4.3 list the preselection and selection requirements that signal photons are required to satisfy.

4.1 Photon Reconstruction

The use of photons in the observation and measurement of the decay of a new resonance is of particular interest. For this channel the energy resolution has a significant impact on the sensitivity of the search, and the steps used to optimize the performance of the energy reconstruction and its accurate simulation are described in this section.

4.1.1 Photon Super Cluster

Photon showers deposit their energy in several crystals in the ECAL. Approximately 94% of the incident energy of a photon is contained in 3x3 crystals, and 97% in 5x5 crystals [5]. The presence of material in front of the calorimeter results in bremsstrahlung and photon conversions. Because of the strong magnetic field the energy reaching the calorimeter is spread in ϕ. The energy is therefore clustered at the electromagnetic calorimeter level by building a cluster of clusters (Super Cluster or SC), which is extended in ϕ. The topological variable $R_9 = E_{3\times3}/E_{SC}$ (energy sum of 3x3 crystals centered on the most energetic, divided by the energy of the Super Cluster), is used to discriminate between unconverted ($R_9 > 0.94$) and converted ($R_9 < 0.94$) photons.

Two main algorithms are used to reconstruct the Super Cluster associated to a photon or electron in ECAL. The Hybrid algorithm [5] uses the $\eta - \phi$ geometry of the barrel crystals to exploit the knowledge of the lateral shower shape in the η direction (taking a fixed bar of 3 or 5 crystals in η), while searching dynamically for separated energy in the ϕ direction. The Island algorithm starts by a search for seeds which are defined as crystals with an energy above a certain threshold.

Starting from the seed position, adjacent crystals are examined, scanning first in ϕ and then in η. Along each scan line crystals are added to the cluster until a rise in energy or crystal that has not been read out is encountered. In much the same way as energy is clustered at the level of calorimeter cells or crystals, non-overlapping Island clusters can in turn be clustered into Super Clusters. The procedure is seeded by searching for the most energetic cluster and then collecting all the other nearby clusters in a very narrow η-window (1-2 crystals), and much wider ϕ-window(5-6 crystals).

The choice of the best clustering algorithm depends on the spread of the deposited energy. For compact energy deposits, originating mainly from the single showers of unconverted photons, the best energy measurement is achieved using a 5x5 crystal array, while Super Clustering algorithms provide better measurement for multiple showers originating from conversions where the bending of the electron and positron tracks, and their radiation in tracker material, spreads the energy over a larger area of the calorimeter [5]. A Super Cluster is promoted to a photon candidate if its reconstructed transverse energy is greater than 10 GeV.

Given the limited calorimeter depth, the longitudinal containment of high energy electron and photon showers depends critically on the initial particle energy, on shower length fluctuations and on the actual shower path in the ECAL.

4.1.2 Photon Conversions

Conversion track pairs are reconstructed from a combination of standard tracks reconstructed by using iterative tracking steps [101]. A mixture of these tracks is preselected with basic quality criteria (nHits> 4, $\chi^2 < 10$), hence opposite-charge pairs are considered which satisfy the photon conversion topology. The signature which distinguishes photon conversion candidates from massive V0 particles[1] or from vertices from mis-reconstructed tracks is the parallelism of the conversion electrons tracks at the conversion vertex. Conversions are also required to have a fitted $p_T > 1$ GeV and they are considered matched to a photon Super Cluster if the $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ measure between the direction of the Super Cluster and the conversion is less than 0.1. About one quarter of the events have at least one of the photon reconstructed and selected as a conversion.

4.1.3 Energy Reconstruction and Calibration

A good understanding of the expected signal shape is crucial as an input to statistical analysis procedure. The resolution of the reconstructed mass peak for the energies involved in this high mass search is driven by the detector energy resolution and by the vertex resolution.The mass resolution improves linearly with the vertex resolution down to a vertex resolution of 1 cm. With better accuracy than this, the mass resolution is totally dominated by the energy measurements in the calorimeter. It is thus necessary to optimize the agreement in energy measurement between MC and the observed data, such as the simulated events match the observed detector performance. In this section the different ingredients entering the photon energy

[1]The appearance of the decay of a neutral strange particle into two observed charged daughter particles gives rise to the terminology 'V0' to describe the decay topology

calibration and the SC energy calibration performed with a multivariate technique (MVA) are discussed. The procedure through which data is scaled and simulation is smeared to obtain a signal model that is as realistic as possible is documented in Section 4.1.4.

There are three main ingredients affecting the determination of the photon energy [102]:

- Channel intercalibration
- Transparency loss corrections
- Clustered energy corrections

Channel Intercalibration

Individual channel calibrations (IC) are derived in situ mainly by equalizing the response to low mass diphoton resonances (π^0, η) across the detector [103]. Supplementary information and cross checks are provided by studying the approximate ϕ-invariance of the energy flow in minimum bias data [103], and the ratio of the energy over the momentum (E/p) of isolated electrons [104]. In Fig. 4.1 one can appreciate the effect of the individual channel calibration in Z$\rightarrow ee$ events from 2011 data by comparing the blue (uncalibrated) and red curves (calibrated).

Transparencies Loss Correction

The response stability in time is achieved using per-channel corrections to compensate for variations of light transmission in the crystals due to radiation damage and recovery.

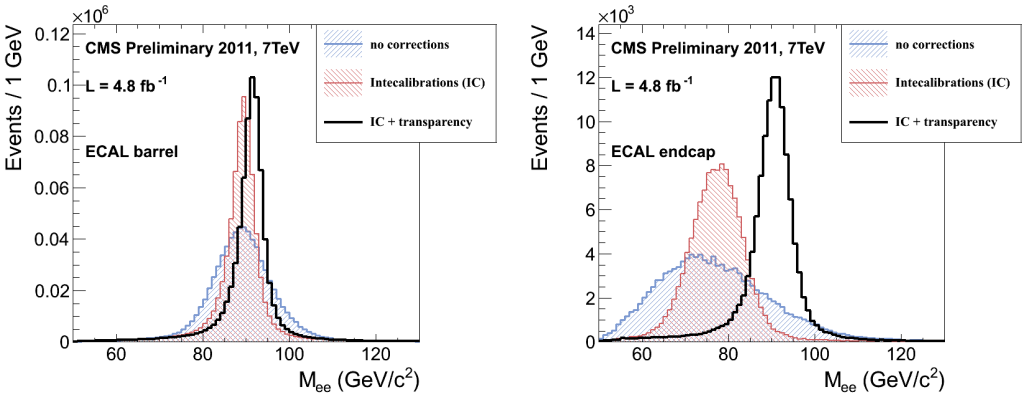
Laser light is injected in each crystal every 40 minutes and the signal measured by the photon detectors, relative to the signal at a reference time, is used to monitor the light transmission and derive corrections [105]. The time history of reference signals, such as the E/p of isolated electrons, or the invariant mass peak of Z$\rightarrow ee$ decays and of π^0 and η decays into photons, is used to monitor the quality of the correction.

In Fig. 4.1 one can appreciate the effect of the transparency loss corrections in Z$\rightarrow ee$ events from 2011 data by comparing the red (calibrated, uncorrected) and black curves (calibrated and corrected). As expected, the transparency losses being smaller in barrel, the effect is correspondingly smaller there due to the different dose of radiation to which the crystals are exposed.

Clustered Energy Corrections

The photon energy is computed starting from the raw Super Cluster energy (adding also the preshower energy in the endcap) as detailed in Section 4.1.1. In order to obtain the optimal resolution, the raw energy must be corrected for local containment of the shower in the calorimeter, as well as the global containment of the shower for photons which convert and shower in material upstream of the calorimeter.

These corrections are computed using a multivariate regression technique originally based on the TMVA Gradient Boosted Decision Tree implementation [106],



(a) barrel (b) EE

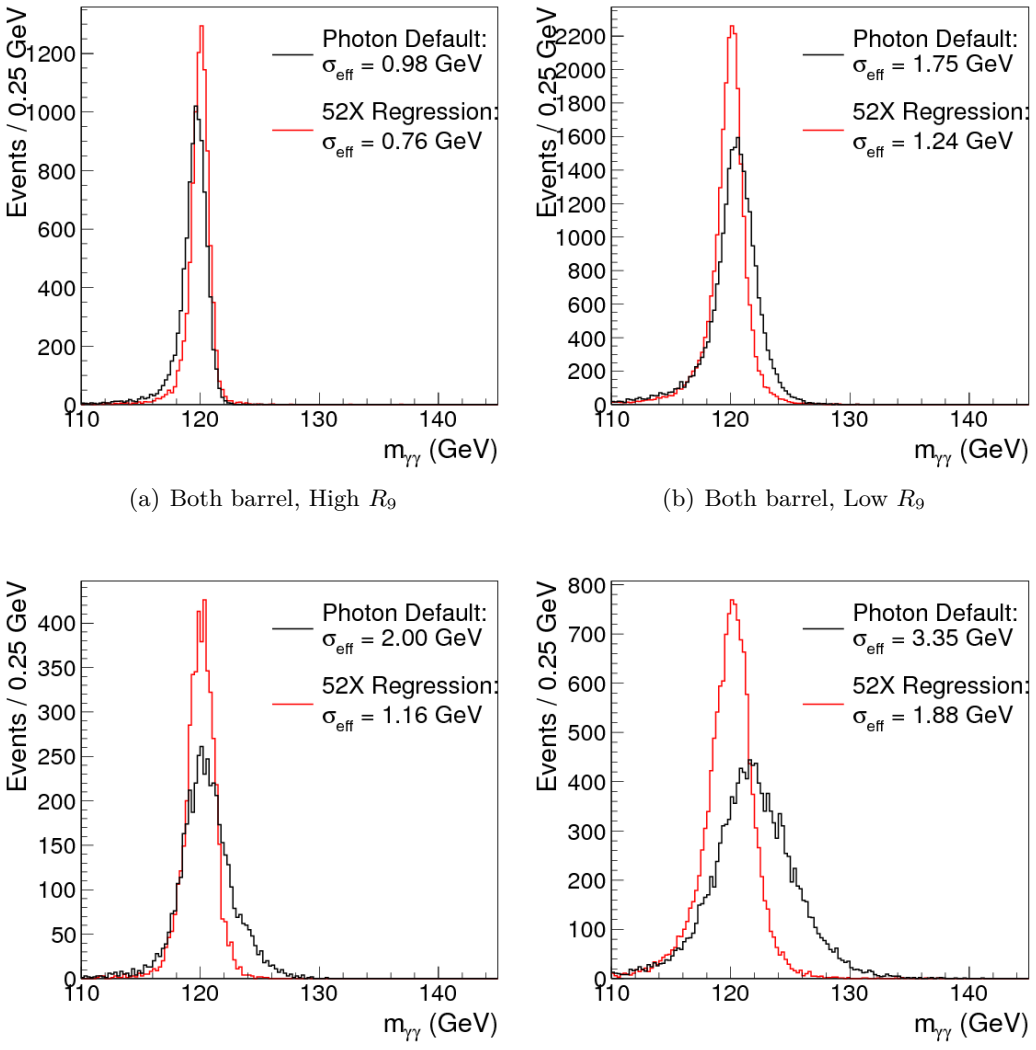
Figure 4.1. Using the best energy clustering and correction procedures, one can appreciate the effect of the individual channel calibration (IC) and transparency loss corrections (Laser) on 2011 Z→ *ee* data: uncalibrated and uncorrected (no IC, no Laser) data in blue, calibrated but uncorrected (IC, no Laser) data in red, and calibrated and corrected (IC, Laser) data in black.

though substantially optimized for the SM Higgs search in the diphoton channel [17]. The regression is trained on prompt photons in Monte Carlo (from the γ + Jets sample) using the ratio of generator level photon energy to the raw Super Cluster (+ preshower) energy as the target variable. The input variables are:

- The global η and ϕ coordinates of the Super Cluster
- The R_9 of the Super Cluster
- The ratio of the 5x5 crystal energy to the raw Super Cluster energy
- The energy weighted η-width and ϕ-width of the Super Cluster
- The number of basic clusters
- The ratio of hadronic energy behind the Super Cluster to the electromagnetic energy of the cluster.

In the endcap, the ratio of preshower energy to raw Super Cluster energy is additionally included. These variables provide information on the likelihood and location of a photon conversion and the degree of showering in the material, and together with their correlation with the global η and ϕ position of the Super Cluster, drive the degree of global containment correction predicted by the regression. Finally the number of primary vertices and median energy density ρ in the event are included in order to correct residual energy scale effects from pile-up. The ρ is evaluated for each event by taking the median value of the following distribution: $\rho = median\left[\frac{p_{T_j}}{A_j}\right]$, where the index j runs over all the jets reconstructed in the event, and A_j is the jet area.

The performance of the regression for photons in Monte Carlo is shown in Figure 4.2 in terms of the $H \to \gamma\gamma$ mass resolution in the same four event classes used for this analysis (see Section 5.3), comparing the signal shape obtained after the regression to the shape computed using the default photon energy in the reconstruction (using $E_{5\times5}$ for high R_9 photons and electron-tuned Super Cluster corrections for low R_9).



(a) Both barrel, High R_9

(b) Both barrel, Low R_9

(c) One EE, High R_9

(d) One EE, Low R_9

Figure 4.2. Comparison of the diphoton mass resolution in Higgs→ 120 GeV Monte Carlo. The default reconstructed photon energy is shown in Black, and the full regression-corrected energy is shown in red.

4.1.4 Resolving Data and Simulation Discrepancies

Having corrected the data as much as possible, some discrepancies with respect to simulation are still present. Based on Z→ *ee* data and simulated events, these discrepancies are dealt with by correcting the energy scale in data and by then determining the smearing needed to apply to the simulated samples so as to have the best match between data and simulation.

Energy Scale Correction

The Super Cluster energy scale is tuned and corrected varying the scale in the data to match the Monte Carlo in $Z \to ee$ events. An analytic fit to the Z invariant mass peak, build with Super Cluster energies, is performed using a convolution of a Breit-Wigner with a Crystal-Ball (CB). Data and MC distribution are fitted separately and fit results are compared to extract scale offset. In the fit the parameter of the Breit-Wigner are fixed to the PDG [10] values: $M_Z = 91.188$ GeV and $\Gamma_Z = 2.495$ GeV. On the contrary the parameters of the CB, which gives a reasonable description of the calorimeter resolution effects and bremsstrahlung losses in front of the calorimeter, are free parameters of the fit.

The data-MC difference is time dependent; moreover the time dependence is not the same in different pseudorapidity region while it is very similar for showering and non showering electrons. The relative scale difference is defined as:

$$\Delta P = \frac{\Delta m_{data} - \Delta m_{MC}}{m_Z} \tag{4.1}$$

where Δm is peak shift of the CB function in respectively data and simulation. The quantity $(1 - \Delta P)$ ranges up to 1.5% in the barrel, and up to 4.9% in the endcap, and it is applied as a multiplicative factor to data to correct the photon energy scale.

Monte Carlo Energy Smearing

Finally a direct smearing to the MC energies has been developed to estimate more efficiently the effective resolution of the electromagnetic calorimeter. The Super Cluster energy is modified by applying a Gaussian multiplicative factor centered in $1 + \Delta P$ and with a $\Delta\sigma$ resolution, where ΔP is the energy scale correction and $\Delta\sigma$ is the additional constant term in the energy resolution.

After defining n exclusive electron categories, the $[n(n+1)/2]$ $Z \to ee$ invariant mass distributions are build in data and simulation. The method is based on the maximization of the likelihood between the data and the smeared MC in the $[n(n+1)/2]$ $Z \to ee$ invariant mass distributions as a function of the $2n$ parameters: (ΔP, $\Delta\sigma$) for each electron category. The big advantage of this method is therefore to exploit the full $Z \to ee$ data sample in the search of the smearing parameters, including all the events with the two electrons lying in different categories.

In this analysis $n = 8$ categories are used[2], accounting for different η regions of barrel and endcap, for different R_9 intervals and for the distance between the cluster

[2]Four region in pseudorapidity($|\eta| < 1.0$, $1.0 < |\eta| < 1.4442$, $1.566 < |\eta| < 2.0$, $2.0 < |\eta| < 2.5$) and 2 categories in R_9 ($R_9 < 0.94$: high interaction with upstream material, $R_9 > 0.94$: low level of interaction with the upstream material)

and the ECAL module borders. Two examples of mass distributions are visible in Figure 4.3: the left plot shows the Z→ *ee* with both electrons belonging to the (barrel, $|\eta| < 1$, $R_9 > 0.94$) category, while the right plot shows one mixed category: one (barrel,$|\eta| < 1$, $R_9 > 0.94$) and one (endcap,$|\eta| > 2$, $R_9 > 0.94$) electron. The final smearing factors range from 1.1% to 2.4% in the barrel, and from 3.3% to 6.1% in the endcaps.

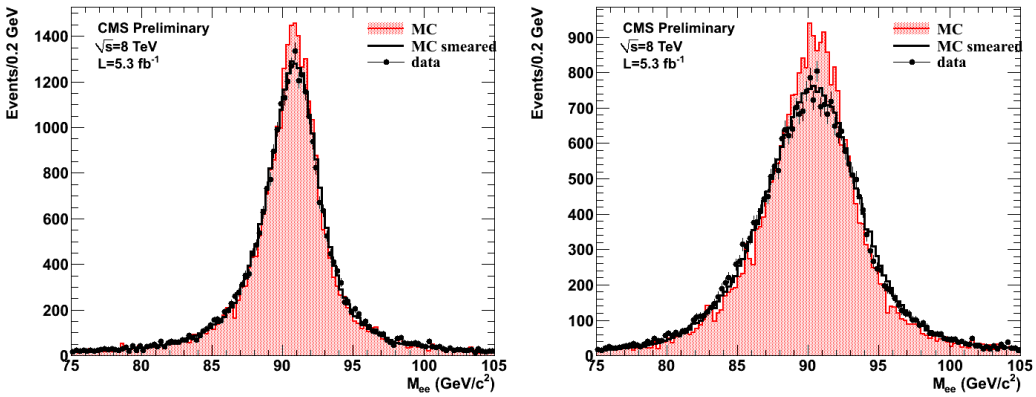


Figure 4.3. Invariant mass distribution of Z→ee events per category: (left) both (barrel, $|\eta| < 1$, R_9 >0.94) electrons, (right) one (barrel, $|\eta| < 1$, R_9 >0.94) and one (EE, $|\eta| < 2$, R_9 <0.94) electron. Red histogram is MC reweighted for the pile-up distribution while the black empty histogram is the optimal smearing of MC energies to match Data distributions.

4.2 Photon Preselection

A loose preselection is applied to all prompt and non-prompt photons within the fiducial region of ECAL. This preselection is also a crucial step for the photon identification that is explained later, in Section 4.3. First, to avoid misidentifying an electron as a photon, a conversion-safe electron veto is applied. Second, to make sure the photon identification is performed in a region where simulation can properly describe the behavior of data, a selection is applied to keep the common phase space between data passing the trigger and the MC where no trigger requirement is applied. The variables used for preselection are built to be similar to ones used in the trigger and in the electromagnetic filter applied to simulated QCD background at generation level. This specific filter requires the presence in the event of at least two particles that can produce an energy deposit in the ECAL sufficient to mimic a photon (i.e. "fake photons").

They are based on the Particle Flow (PF) reconstruction of the involved objects [107]. The preselection cuts for photons in the Barrel ($|SC_\eta| < 1.4442$) and in the Endcap ($1.566 < |SC_\eta| < 2.5$) are listed in Table 4.1 and the variables are defined as follows:

1. H/E (HoE): The ratio of hadronic energy in HCAL towers behind the Super Cluster to the ECAL energy in the Super Cluster.

2. $\sigma_{i\eta i\eta}$ (CovIEtaIEta): The energy weighted (single crystal energy over the Super Cluster energy) standard deviation of single crystal η within the 5x5 crystals centered at the crystal with maximum energy.

3. EtCorrHcalIso: The HCAL isolation within $\Delta R < 0.3$ cone minus 0.005 times the transverse energy of the photon.

4. EtCorrTrkIso: The track isolation within $\Delta R < 0.3$ hollow cone minus 0.002 times the transverse energy of the photon.

5. ChargedPFIso: The sum of p_T of the charged particle flow candidates within $0.02 < \Delta R < 0.2$.

	Barrel		Endcap		Both Barrel and Endcap		
R9	HoE	CovIEtaIEta	HoE	CovIEtaIEta	EtCorrHcalIso	EtCorrTrkIso	ChargedPFIso
≤ 0.9	< 0.075	< 0.014	< 0.075	< 0.034	< 4 GeV	< 4 GeV	< 4 GeV
> 0.9	< 0.082	< 0.014	< 0.075	< 0.034	< 50 GeV	< 50 GeV	< 4 GeV

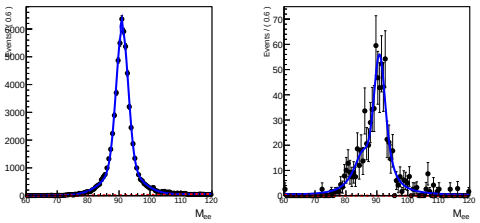
Table 4.1. Preselection cuts.

Table 4.2 shows preselection efficiencies measured using tag and probe with Z→ee events for data (TNP(data)), Monte Carlo (TNP(MC)) and the ratio TNP(data)/TNP(MC), for the four photon categories. The tag and probe technique is the same described in detail in Section 5.1. Figures 4.4 show the invariant mass distributions of tag and probe electrons used for measuring the data to Monte Carlo preselection efficiency scale factors in the four categories.

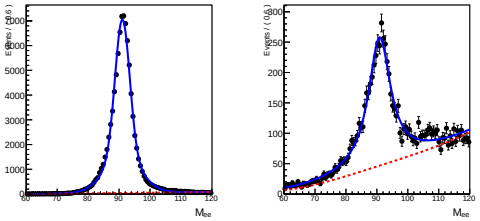
	DATA			MC		R	
	Eff.	Stat. Err.	Syst. Err.	Eff.	Stat. Err.	Eff.	Err.
Barrel; R_9 >0.9	0.9879	0.0002	0.0030	0.9864	0.0001	0.999	0.003
Barrel; R_9 <0.9	0.9566	0.0006	0.0055	0.9610	0.0002	0.995	0.006
Endcap; R_9 >0.9	0.9838	0.0003	0.0090	0.9789	0.0002	1.005	0.009
Endcap; R_9 <0.9	0.9545	0.0009	0.0170	0.9445	0.0005	1.011	0.018

Table 4.2. Photon identification efficiencies measured in the 4 photon categories using tag and probe with Z→ee events (for all cuts except electron rejection).

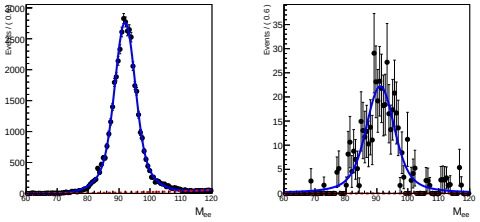
The effect on the signal photon efficiency of the electron veto requirement has been evaluated from data in a sample of Z→ $\mu\mu\gamma$ events compared with events selected in a simulated sample of Drell-Yan events. The photon from Z→ $\mu\mu\gamma$ in a mass window 70 to 110 GeV is additionally subjected to the preselection criteria described above, except the electron veto, and used as probe. The electron veto efficiency is measured as the ratio of the number of photons "passing" the electron veto divided by the total number of pre-selected photons. Table 4.3 lists the results in data and simulation and their ratios. Values are given for the four categories used in the analysis.



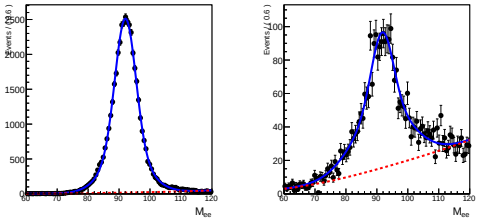
(a) Barrel; $R_9 > 0.94$



(b) Barrel; $R_9 < 0.94$



(c) Endcap; $R_9 > 0.94$



(d) Endcap; $R_9 < 0.94$

Figure 4.4. Invariant mass of the tag-electron plus probe-electron pair computed in the tag and probe procedure for measuring the data to Monte Carlo preselection efficiency scale factors for the four categories at 8 TeV. On the left the passing probe, on the right the failing ones.

	DATA		MC		R	
	Eff.	Stat. Err.	Eff.	Stat. Err.	Eff.	Err.
Barrel; R_9 >0.94	0.998	0.001	1.000	0.000	0.995	0.001
Barrel; R_9 <0.94	0.986	0.002	0.993	0.002	0.998	0.001
Endcap; R_9 >0.94	0.991	0.002	0.998	0.001	0.992	0.002
Endcap; R_9 <0.94	0.962	0.004	0.972	0.005	0.990	0.007

Table 4.3. Efficiency of the conversion-safe electron veto, measured in the four photon categories using tag and probe with Z$\rightarrow \mu\mu\gamma$ events. The data to simulation ratio is also shown with its error. The efficiency is built using for the denominator the number of photons passing all cuts except the electron veto, and for the numerator the number of photons passing all cuts including the electron veto.

4.3 Photon Identification

The photon identification is performed using a set of cuts on six discriminating variables. Cut values are optimized separately in four categories defined in terms of pseudorapidity and R_9. These categories have significantly different levels of background and mass resolution and their use provides increased sensitivity. The following variables are used to distinguish isolated photons originating from the primary interaction from the background due to low multiplicity jets with high electromagnetic content:

- Relative combined isolation using selected event vertex. To compute this variable first an isolation sum is calculated as follows:

$$\sum Iso = Iso^{track} + Iso^{ECAL} + Iso^{HCAL} \tag{4.2}$$

  where:

  - Iso^{track} is the scalar sum of the transverse momenta of tracks which are consistent with originating from the primary vertex (within $\pm$1 cm along the beam direction and within $\pm$0.1 cm transverse to the beam direction) selected as described in Section 5.2 and lying within a hollow cone of size $\Delta R < 0.3$ centered around a line joining the selected primary vertex to the ECAL Super Cluster, excluding an inner cone ($\Delta R_i = 0.02$) in order to avoid including the momenta of conversion tracks.
  - Iso^{ECAL} is computed as the transverse energy sum of ECAL energy deposits in crystals located within a cone of size $\Delta R < 0.3$ (approximately 1250 crystals), centered around the Super Cluster position, excluding an inner veto cone ($\Delta R_i = 3.5$ crystals) and eta-slice ($\Delta\eta = 2.5$ crystals) in order to exclude the footprint of the signal photon, which can be extended in the ϕ direction in the case of converted photons.
  - Iso^{HCAL} is the sum of the energies of HCAL towers whose centers lie within a ring-shaped region of outer radius $\Delta R = 0.4$ and inner radius $\Delta R = 0.15$, centered on the ECAL Super Cluster position.

For each of the isolation sums defined above, the energy deposited within the isolation cone is contaminated by energy from pile-up and from the underlying event. Since the contamination increases with the number of pile-up vertices n_{PV}, the efficiency of the isolation cut decreases with increasing pile-up. In order to maintain high efficiency under high pile-up conditions, the contribution to $\sum Iso$ from pile-up and the underlying event is estimated on an event-by-event basis as the product of the measured energy density ρ for the event determined using the FastJet algorithm [108], and an effective area A_{eff} corresponding to the isolation cone excluding veto regions. A_{eff} is determined empirically as the ratio of the slopes of linear fits to the mean value of $\sum Iso$ vs n_{PV} and to the value of ρ vs n_{PV} in Z events. The value of A_{eff} for the definition of isolation described here is found to be 0.17.

The pile-up corrected isolation sum is then given by:

$$\sum Iso^{PUcorr} = \sum Iso - \rho A_{eff} \tag{4.3}$$

The isolation sum is then scaled by p_T^{pho}/50GeV/c, where p_T^{pho} is the transverse energy of the photon determined using the selected primary vertex. The relative isolation is thus given by:

$$Iso^{rel} = \frac{\sum Iso^{PUcorr}}{p_T^{pho}/50GeV/c} \tag{4.4}$$

- Relative combined isolation using event vertex giving highest Iso^{Trk}. The relative combined isolation using event vertex is computed as above with the following differences:
  - Iso^{Trk} is computed for each reconstructed primary vertex and the largest value is used.
  - For all three sub-detector isolation the outer cone size ΔR to 0.4 is used. The value of the effective area A_{eff} of the isolation cone is here set to 0.52.

  This definition of isolation adds discrimination since in the previous definitions Iso^{Trk} gives no discrimination in the case that the wrong primary vertex is selected, although it is more powerful in the case that the correct vertex is selected. In this way the different definitions are complementary.

- Relative track isolation using selected event vertex. Since track isolation is the most discriminating of the three sub-detector isolation variables, a cut is additionally applied on relative isolation defined using track isolation only:

$$Iso^{rel,track} = \frac{\sum Iso^{track}}{p_T^{pho}/50GeV/c} \tag{4.5}$$

where Iso^{track} is the one computed using the selected primary vertex. No pileup subtraction is required since only tracks consistent with the selected primary vertex are included in the sum.

- H/E. The ratio of hadronic energy to electromagnetic energy is calculated as the ratio of the sum of HCAL tower energies within a cone of size $\Delta R < 0.15$ centered on the ECAL Super Cluster position, to the energy of the Super Cluster. Due to the 25 radiation length thickness of the ECAL crystals, isolated photons have a value close to or equal to zero.

- $\sigma_{\eta\eta}$. The transverse shape of the electromagnetic cluster is computed with logarithmic weights and is defined as

$$\sigma^2_{\eta\eta} = \frac{\sum_i^{5\times5} w_i(\eta_i - \bar{\eta}_{5\times5})^2}{\sum_i^{5\times5} w_i};\ w_i = \max\left(0, 4.7 + \ln\frac{E_i}{E_{5\times5}}\right) \tag{4.6}$$

  where E_i and η_i are the energy and pseudorapidity of the i^{th} crystal within the 5×5 electromagnetic cluster and $E_{5\times5}$ and $\eta_{5\times5}$ are the energy and η of the entire 5×5 cluster. The value of $\sigma^2_{\eta\eta}$ tends to be smaller for single isolated photons (including converted photons, since the cluster is spread in the ϕ direction only), than for the background which is dominated by jets consisting of π^0s decaying to two photons.

- Minimum threshold on R_9. A minimum threshold on R_9 is applied to photons in the ECAL endcaps in order to exclude very poorly reconstructed photons.

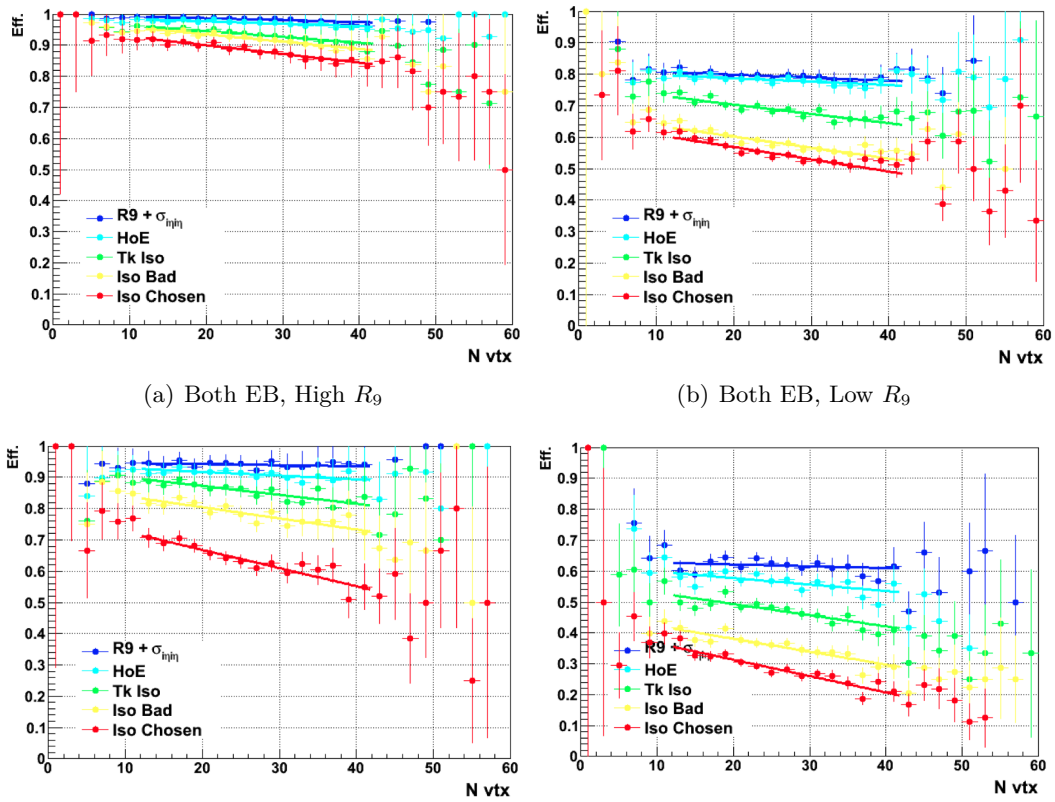
The cuts are set to get the best compromise between signal efficiency and fake rate. This means that the cuts are tighter in the low R_9 category than in the high R_9 category and tighter in the endcap than in the barrel. The chosen photon identification working point is applied to both legs of the diphoton pair and the cuts are listed in Table 4.4.

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

Table 4.4. Photon ID selection cut values. The cuts are applied to both the leading and sub leading photons.

Photon Identification performance

The efficiency of the photon identification variables as a function of the number of reconstructed vertexes in shown in Figure 4.5 in the four photon categories. The single-photon efficiency evaluated on a simulated signal samples ranges from 97% to 83% going from the first to last category. Figure 4.6 shows the signal efficiency of the



(a) Both EB, High R_9

(b) Both EB, Low R_9

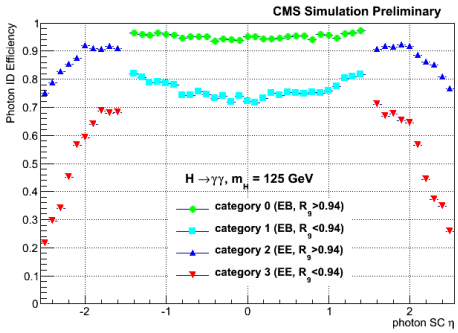
(c) One EE, High R_9

(d) One EE, Low R_9

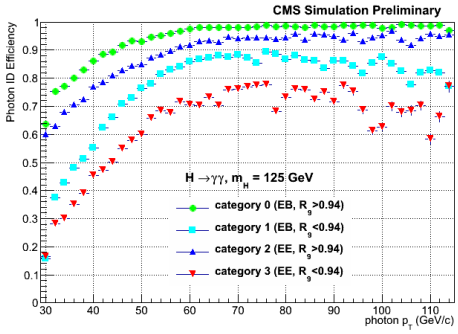
Figure 4.5. Photon Identification variables as a function of the number of vertexes in the four photon categories. Within the legend, "Iso Good" stands for Relative Combined Isolation with respect to the selected vertex; "Iso Bad" stands for Relative Combined Isolation with respect to the worst vertex; "HoE" stands for H/E

photon identification from the decay of a Higgs boson of mass 125 GeV as a function of the pseudorapidity and the p_T of the photon, for each of the four categories.

In case of a low mass resonance (200 GeV) decaying to $\gamma\gamma$, the kinematics is very similar to the one from the SM Higgs. For higher resonance masses the photons are more boosted, consequently easier to identify (see Figure 4.7). This property could be used to improve the photon identification efficiency in the high mass region, where anyway the background is much smaller. The drawback of the photons of the investigated signal to have sizeable boost is that the p_T of the photons can be out of the range for which the identification uncertainties were computed. The propagation of this to the final uncertainties is detailed in Section 5.6.



(a) η



(b) p_T

Figure 4.6. Photon Identification efficiency for signal photons from the decay of a Higgs boson of mass 125 GeV, as a function of pseudorapidity (a) and p_T (b), for each of the four photon categories.

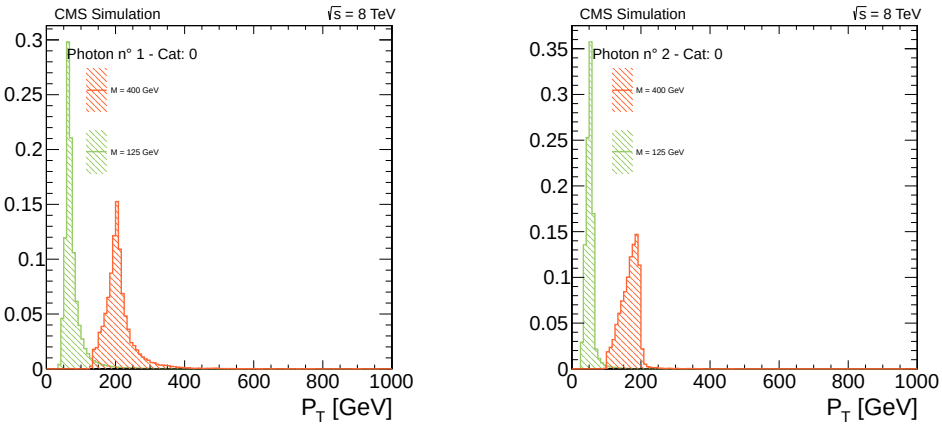


Figure 4.7. Transverse momentum of the two selected photons for the signal at high mass compared with the SM higgs production.

Chapter 5

Analysis Strategy

The search for a new resonance is performed by fitting analytical descriptions of the signal and background distributions to the measured diphoton invariant mass spectrum $M_{\gamma\gamma}$. Section 5.1 reports the analyzed datasets, both data and simulated events. Subsequent Section 5.2 reports the vertex determination technique employed in the analysis for the identification of the primary vertex in the event. The event classification and categorization is shown in Section 5.3.

In searching for a small signal over a huge continuum background in a wide energy range, the two main issues are the background modelling and the signal parametrization. The former topic is discussed in Section 5.4 while the latter is described in Section 5.5. A detailed description of all the systematic effects entering in the analysis is given in Section 5.6. Section 5.7 finally describes the fit technique exploited in this analysis and the limit setting procedure.

5.1 Data sample and Trigger

This analysis focuses on the search of a massive diphoton resonance with mass above 150 GeV. Most of the signal events, therefore, will fire the double-photon HLT paths, and be stored in double photon primary dataset called *DoublePhoton*. The results presented here make use of a total of 19.7 fb^{-1} of data collected by the CMS detector during the 2012 data taking at a centre-of-mass energy of 8 TeV. The analyzed running periods and their corresponding run ranges and integrated luminosities are reported in Table 5.1.

Run period	Int. lumi (fb^{-1})
RunA+B	5.3
RunC	7.1
RunD	7.3
RunA+B+C+D	19.7

Table 5.1. Subdivision of the 2012 dataset into running periods: each period is identified by its name and the corresponding integrated luminosity

An example of a real event with two high energy photons selected from data is shown in Figure 5.1. The two photons in this event have respectively $p_{T1} = 136.9$

GeV and $p_{T2} = 99.9$ GeV. They are produced almost back to back in the transverse plane and their invariant mass is equal to $m_{\gamma\gamma} = 368.1$ GeV. The number of multiple interactions in this event is equal to 12. The red towers represent the energy deposits in the ECAL. The height of each tower is proportional to the energy released there by the incident particle. The green lines are additional tracks in the same event, produced in secondary interactions.

5.1.1 Trigger

The trigger system at CMS is responsible of the evaluation of each LHC collision inside the detector. The system is conceived as a two-levels system which includes the Level-1 Trigger and the High Level Trigger as described in Section 3.9. The trigger criteria for identifying diphoton events have evolved with the increasing instantaneous luminosity delivered by the LHC. The diphoton triggers used by the analysis can be categorized into two types (two HLT paths [109]), one with E_T thresholds of 26/18 GeV and the other with E_T thresholds of 36/22 GeV, where the two thresholds are applied respectively to the leading and the sub-leading photon. Each 26/18 path is required to be initiated by at least two hardware L1 e/γ candidate, whereas each 36/22 path is required to be initiated by at least one hardware L1 e/γ candidate.

The HLT selection criteria can be grouped into general isolation plus calorimeter identification("Iso+CaloId"), and shape requirements("R_9"). The variables used include the following:

- General variables: E_T of both photon legs and the diphoton invariant mass $M_{\gamma\gamma}$
- Other variables used in CaloId + Iso paths: The ratio between the energy deposition in HCAL and ECAL for a solid cone of radius 0.15 (H/E), ECAL isolation, HCAL isolation, Track isolation
- Other variables used in R9 paths: The ratio between the energy of the 3x3 crystals around the most energetic crystal and the energy of the whole supercluster (R_9)

Any photon that passes the general cuts and either of the Iso+CaloId or the R9 cuts is considered a "good" photon. Therefore, the general trigger strategy is to keep all possible good photon pairs using the "OR" triggers [109].

To evaluate the trigger efficiencies, the efficiency of the L1-seeding (L1 efficiency) and the efficiency of the HLT filters once the L1 requirement has been satisfied are computed separately. In this analysis, a tag and probe method on Z$\rightarrow$ *ee* data for efficiency measurements is employed. This can be done assuming that photons will have similar efficiencies as electrons since at reconstruction level they are almost the same electromagnetic object in the calorimeter. The systematic bias on this assumption is included in the analysis when computing the effects of the scale and resolution on the energy measurement in ECAL (see Section 5.6). The procedure of the tag and probe strategy is as follow:

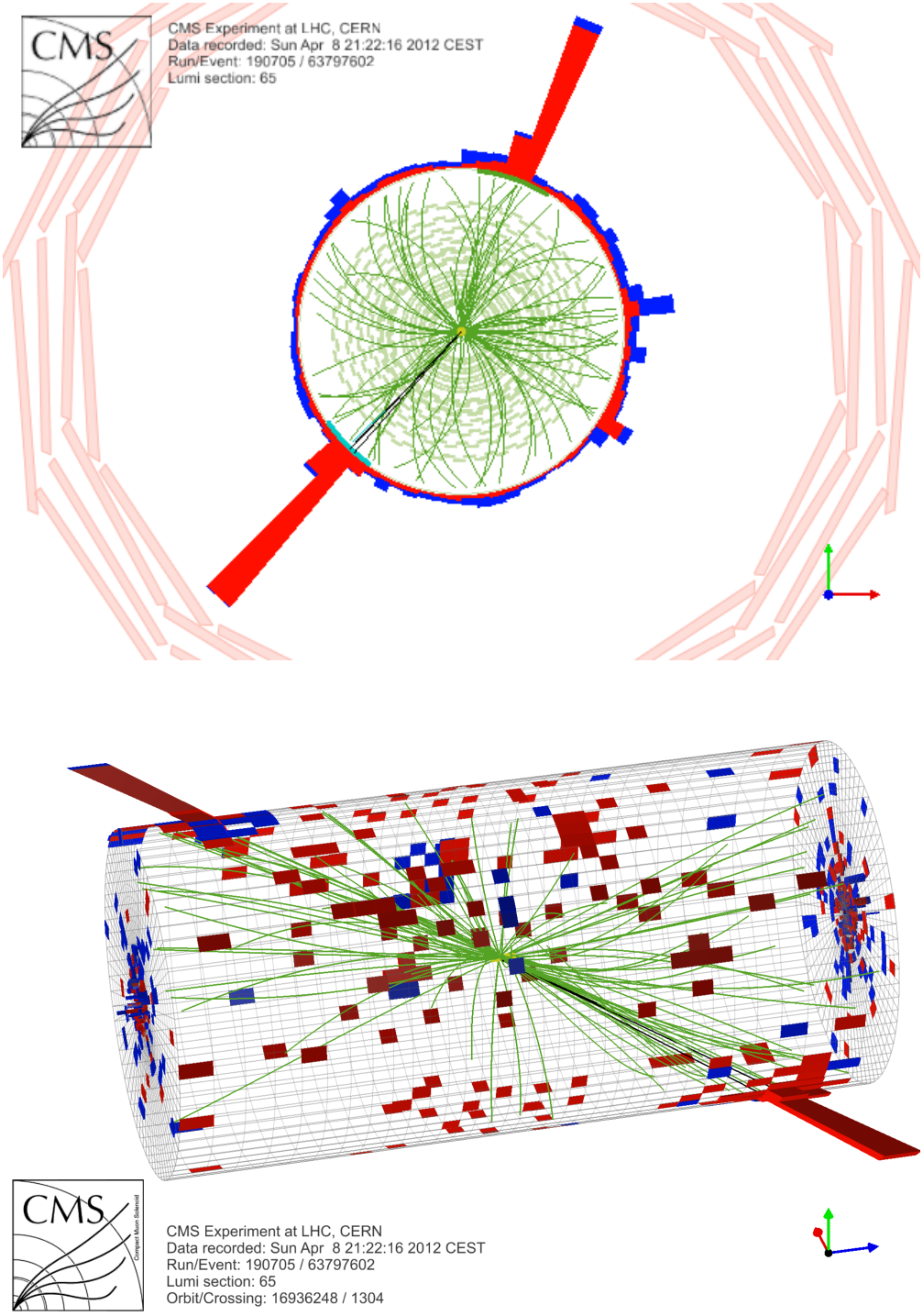


Figure 5.1. Event display of a diphoton production in CMS. The diphoton system has an invariant mass of $m_{\gamma\gamma} = 368.1$ GeV.

1. A sample of data is selected where one electron is required to pass a set of tight HLT requirements and the other electron is required to pass only a very

loose selection

2. At offline level, a reconstructed electron is required to match to the high quality HLT electron and at least two photons are required to match the two HLT electromagnetic objects and have an invariant mass compatible with the Z peak

3. The event is required to pass the preselection described in 4.2

4. The electron which match to the HLT electron leg is labelled as tag and the other one as probe

5. The number of "tag+passing probe" and "tag+failing probe" events are used to compute the efficiencies, where "passing probes" are defined according to whatever efficiency to measure

6. The procedure is repeated in bins of the probe variables (e.g. p_T, η ...) to compute efficiency as a function of those variables.

An example of the HLT 26/18 trigger efficiencies as a function of the p_T of the electron is shown in Figure 5.2 for the cluster shape requirements and for the isolation requirements only and for the OR of the two. For energies above 30 GeV the HLT OR trigger is 100% efficient.

5.1.2 Monte Carlo Simulation

The description of a new resonance decaying in two photons and of all the background processes is obtained from Monte Carlo (MC) simulations. The parton hadronization and jet fragmentation processes are generated with PYTHIA 6 [110] while the interaction of particles with the CMS detector is simulated with GEANT 4 [111]. Monte Carlo signal and background events are generated using of combination of samples. Simulated signal samples of $X \to \gamma\gamma$ events are generated with PYTHIA [110] for the gluon fusion (ggH) and vector boson fusion (VBF). Signal events are generated for resonance mass points corresponding to 150, 200, 250, 300, 400, 600, 850 GeV. Signals are generated assuming SM Higgs branching ratio for the diphoton final state. Resonances are generated under the narrow hypothesis with a fixed width of 0.1 GeV, i.e. smaller than the detector resolution which is of the order of 1% for energies above 60 GeV. Additional width smearing is applied later in the off-line analysis. A detailed study on the proper theoretical lineshape for a wide high mass resonance is discussed in detail in Section 5.5.2.

Simulated backgrounds include the diphoton continuum production via box and born processes ("DiPhotonJets" and "DiPhotonBox"), generated with MadGraph 5 [112] and SHERPA 1.4.2 [113], and processes where one of the photon candidates arises from misidentified jet fragments ("PhotonJet" and "QCD"), simulated with PYTHIA. All the SM background samples which are used in this analysis are summarized in Table 5.2. A filter was applied during the production of the QCD Dijet and γ+Jets samples in order to improve the production efficiency. This filter requires an electromagnetic activity compatible with the one expected from a photons, coming from electrons or neutral hadrons, with $E_T > 15$ GeV within a small region.

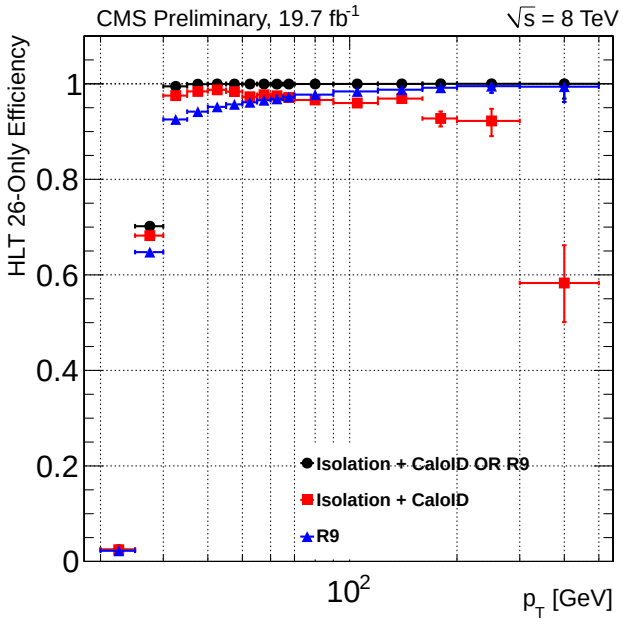


Figure 5.2. HLT trigger efficiency of the 26/28 GeV HLT trigger computed with the tag and probe procedure with Z→ *ee* events in 8 TeV data. In red the "Iso+CaloId" trigger efficiency while in blue the "R9" trigger efficiency. The "Iso+CaloId" efficiency dropping is due to the isolation cut which does not behave properly increasing the energy of the incident particle. The black points represent the OR of the two triggers. From this turn on curves we conclude that using the OR of the two available trigger requirements assure a 100% efficiency for energies above 60 GeV.

In addition it is required that this potential photon signal have no more than one charged particle in a cone of radius 0.2 to take into account converted photons, thus mimicking a tracker isolation.

Dataset	$\widehat{p_T}[GeV/c]$	σ (pb)	ϵ_{filter}
/DiPhotonJets_Madgraph	-	75.4	1.
/DiPhotonBox_Pt-10to25_Pythia	$10-25$	424.8	1.
/DiPhotonBox_Pt-25to250_Pythia	$25-250$	15.54	1.
/DiPhotonBox_Pt-250toInf_Pythia	>250	$1.18\cdot10^{-3}$	1.
/PhotonJet_DoubleEmenriched_Pt20to40	$20-40$	$8.19\cdot10^{4}$	$1.84\cdot10^{-3}$
/PhotonJet_DoubleEmenriched_Pt40	>40	$8.84\cdot10^{3}$	$5.39\cdot10^{-2}$
/QCD_DoubleEmenriched_Pt30to40	$30-40$	$5.20\cdot10^{7}$	$2.35\cdot10^{-4}$
/QCD_DoubleEmenriched_Pt40	>40	$2.37\cdot10^{7}$	$2.18\cdot10^{-3}$

Table 5.2. Background Monte Carlo samples, their production cross-section and filter efficiency for different $\hat{p_T}$-bins generated with PYTHIA program.

5.1.3 Pile-Up Re-weighting

The simultaneous collisions (pile-up) scenario accounts for both multiple pp collisions happening in the same bunch crossing (in-time), and for 50 ns out-of-time pile-up[1]. The number of in- and out-of-time interactions to be overlaid are selected individually from a Poisson distribution based on the chosen luminosity and the total inelastic cross section (71.3 mb). Out-of-time interactions are simulated for each beam crossing that is "in scope" for a given production run. Any arbitrary bunch configuration can be generated in 25 ns steps. Times of Geant 4 [111] simulated hits are shifted to match bunch assignment and the shifted times are considered in generating pulse shapes in the the digitization simulation. Typically, $\pm$125 ns worth of bunch crossings are simulated. Collection of Geant 4 simulated hits from all of the minimum bias events and the hard-scatter "signal" event are merged, and then processed by digitization or electronics simulation [114].

Both signal and background events have been generated with a default pile-up configuration which has constant probability for up to twenty additional interactions, and then a poissonian tail above this value. This distribution is obviously different from what is observed in the data. To fully reproduce the expected distribution of the number of interactions taking place in real data, the simulated events must be reweighted. The distribution of reconstructed primary vertices after the re-weighing is shown in Figure 5.3. The result of the re-weighting technique shows an excellent agreement between the observed and the simulated distribution of the number of multiple interactions in the events.

5.2 Vertex Identification

The mass resolution of a resonance decaying into two photons is driven by two factors: the photon energy resolution and the resolution in measuring the opening angle between the two photons. In this section the treatment of the latter is described.

The opening angle resolution strongly depends on determining the interaction point where the two photons were produced. The identification of the collision point is more difficult in the event of diphoton production since photons are neutral and do not leave a track in the inner tracker from which they can be readily assigned to a given interaction vertex. Therefore, in order to achieve the best possible mass resolution and signal over background discrimination, it is important to have an efficient determination of the correct diphoton vertex.

The luminous region of the LHC at P5 has an RMS spread of about 5 *cm* and the mean number of interactions per bunch crossing is about 19.9. The resolution on the photon opening angle makes a negligible contribution to the mass resolution, as compared to the ECAL energy resolution, only if the interaction point is known within about 10 mm. The mass resolution can be preserved by correctly assigning the reconstructed photons to one of the interaction vertices reconstructed from the charged tracks using the standard CMS algorithm [115].

The method used in this analysis to select the best primary vertex candidate is based on a multivariate approach (MVA) [116] exploiting the kinematic properties

[1]Out-of-time pileup refers to minimum bias collsions occurring in neighbouring bunch crossings (within approximately 10 BX)

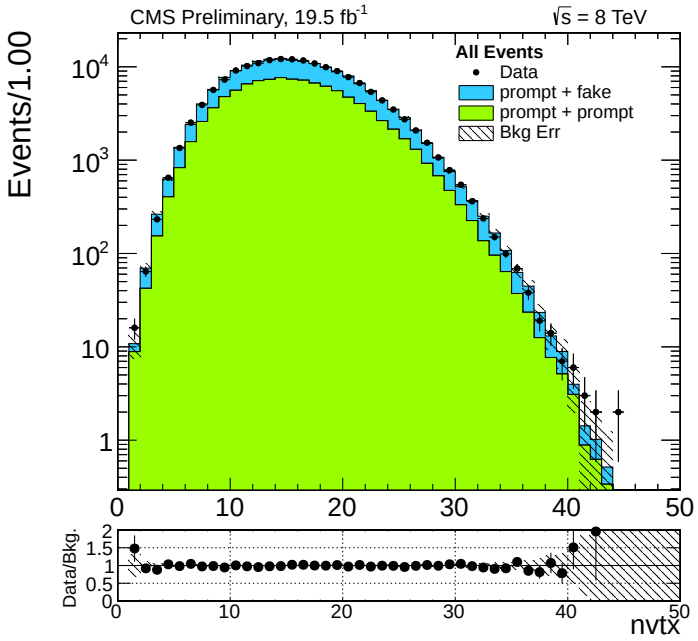


Figure 5.3. Number of reconstructed primary vertices in data and in the simulation after the pile-up reweighing procedure is applied to the simulation

of the vertex tracks in the event and their correlation with the diphoton kinematics, and adding the tracker information for converted photons [17].

The three variables used as inputs for the vertex identification are:

- *sumpt2*: $\sum_i |\vec{p}_T^{\,i}|^2$.
- *ptbal*: $-\sum_i (\vec{p}_T^{\,i} \cdot \frac{\vec{p}_T^{\,\gamma\gamma}}{|\vec{p}_T^{\,\gamma\gamma}|})$.
- *ptasym*: $(|\sum_i \vec{p}_T^{\,i}| - p_T^{\gamma\gamma})\ /\ (|\sum_i \vec{p}_T^{\,i}| + p_T^{\gamma\gamma})$.

where the sums run over all tracks associated to a given vertex.

The distribution of such variables is shown in Figure 5.4 for primary vertices and vertices coming from multiple interactions.

In events with at least one photon conversion, a fourth variable $pull_{conv} = |z_{vertex} - z_{conv}|/\sigma_{conv}$ (where z_{conv} is the estimated primary vertex position and σ_{conv} is the resolution measured in data) is added as input for the multivariate discriminant computation.

Performance of the identification algorithm described above is tested in Z$\to \mu\mu$ events including a comparison between data and simulated events. The lepton tracks are used to identify the hard interaction vertex and the vertices are refitted after removing the leptons tacks from collection of tracks used in the vertex reconstruction algorithm to mimic the topology of a new scalar decaying into two photons.

Figure 5.5 (a) shows the MVA discriminant output for Z$\to \mu\mu$ events in the MC simulation and in data, demonstrating a general good agreement. The efficiency

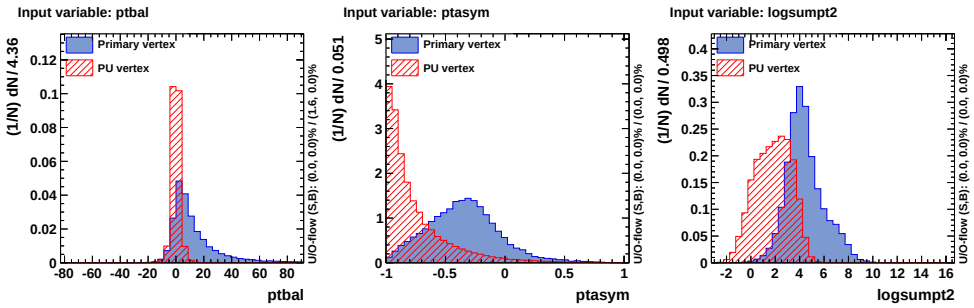
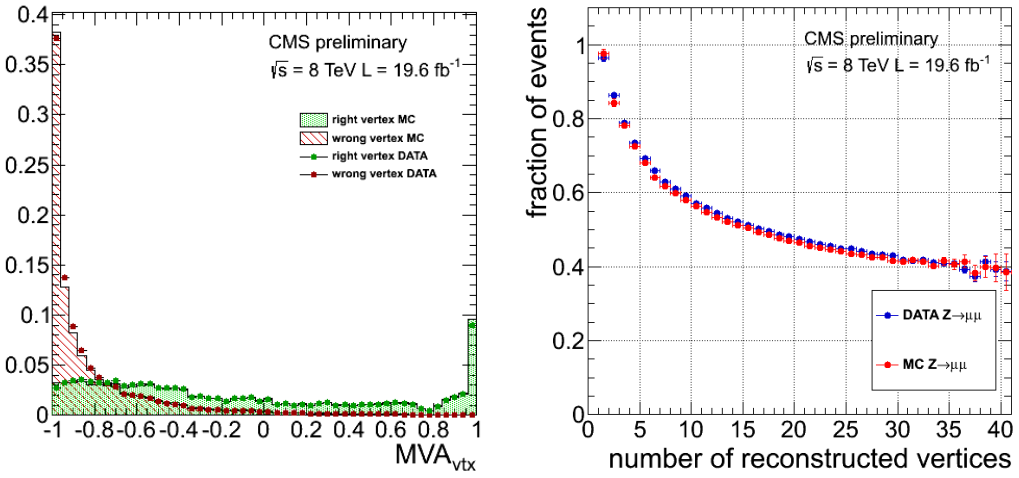


Figure 5.4. Vertex identification variables extracted from the general tracks collection.

measured in data on the Z→ $\mu\mu$ sample is reported in Figure 5.5 (b) as a function of the number of reconstructed vertices.



(a) (b)

Figure 5.5. (a) MVA output for signal and pile-up vertices in Z→ $\mu\mu$ events. (b) Comparison between data and simulation of the fraction of events in which the reconstructed vertex is found within 10 mm from the true vertex position as a function of the number of reconstructed primary vertices [17].

The fraction of X→ $\gamma\gamma$ events, from a sample of simulated events, where the vertex chosen using the MVA discriminant is found within 1 cm from the true vertex is shown as a function of the new X boson mass in Fig. 5.6. It can be seen that the efficiency is 85% in average and increases as a function of the resonance mass due to the harder boost of the diphoton system.

5.3 Selection Criteria and Event Categorization

Signal significance in the search for a new resonance decaying to two photons depends on the mass resolution and the signal-to-background ratio. Search sensitivity can

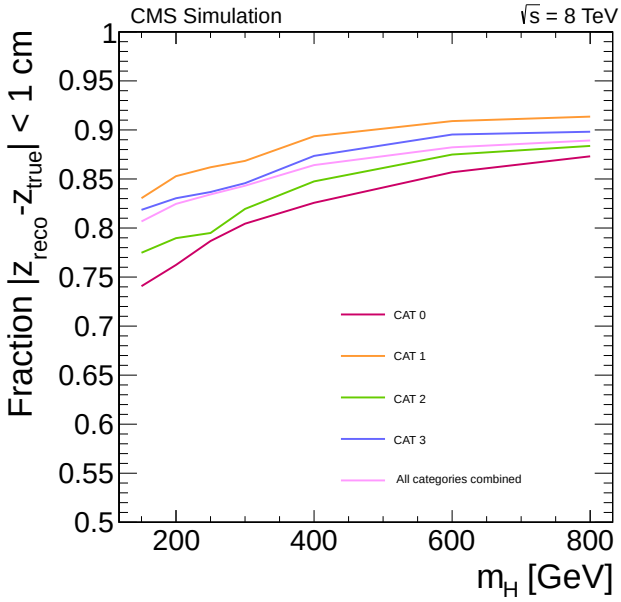


Figure 5.6. Performance of the MVA vertex identification on a X→ $\gamma\gamma$ MC sample for the 2012 pileup conditions as a function of the new resonance mass. Efficiencies are reported also for the four categories used in this analysis. Categories description can be found in Section 5.3.

be increased by subdividing the events into classes according to indicators of mass resolution and predicted signal-to-background ratio and combining the results of the searches from the different classes.

Two simple classifiers are used: the minimum R_9 of the two photons, and the maximum pseudorapidity of the two photons. Both classifiers are effective in separating diphotons with good mass resolution from those with less good resolution, and in separating events with a higher signal to background probability from those with a lower signal/background probability. The class boundary values are chosen to match those used in the photon categories for the photon identification cuts. The event class definitions are shown in Table 5.3.

cat 0	Both photons in barrel	min(R_9)>0.94
cat 1	Both photons in barrel	min(R_9)<0.94
cat 2	One or more in endcap	min(R_9)>0.94
cat 3	One or more in endcap	min(R_9)<0.94

Table 5.3. Inclusive event classes for the analysis

To obtain similar invariant mass shapes over the different categories (hence simplifying the background modelling), instead of cutting on the p_T of the photon it has been chosen to cut on $p_T/m_{\gamma\gamma}$, where $m_{\gamma\gamma}$ is the candidate diphoton invariant mass. The "sliding cut" has been fixed to be 1/3 and 1/4, on $p_T/m_{\gamma\gamma}$ variable, respectively on the leading and sub leading photon. If more than one photon pair passes the selection, the selected pair is that with the highest scalar sum of the

photon transverse momenta.

The diphoton invariant mass distribution for selected events in data and Monte Carlo is shown in Figure 5.7 for all event classes combined, and for each event class separately. It is clear from these distributions that the agreement between data and simulation is not perfect. This is visible from the mean value of the pull distribution in Figure 5.7 which is significantly lower than zero. This is one of the motivations which drive the analysis strategy of determining the background contribution from data instead of using simulation, as described in Section 5.4

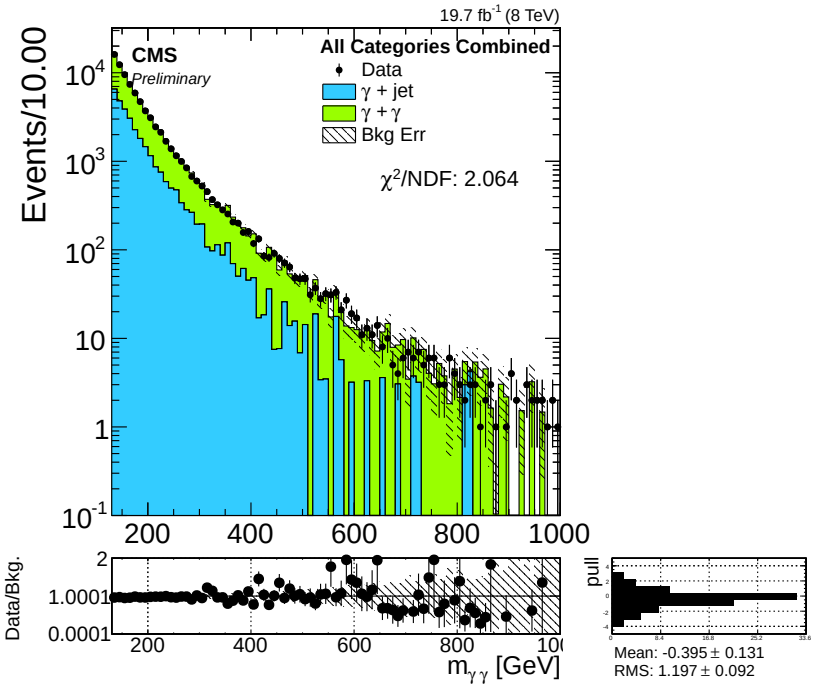
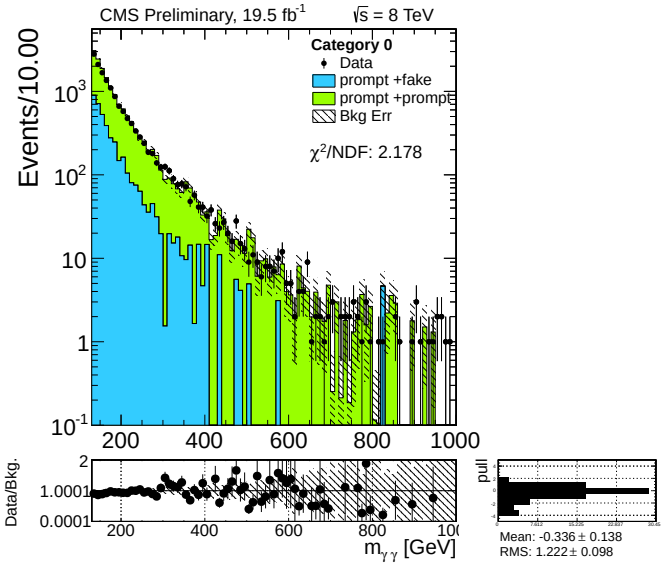
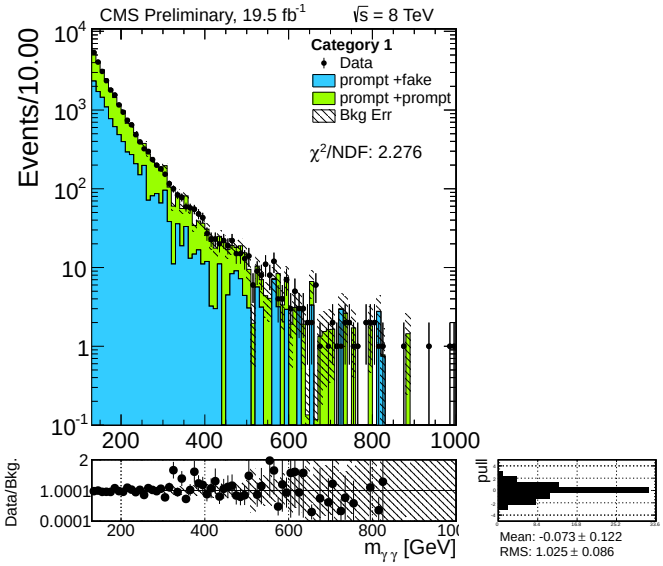


Figure 5.7. Diphoton invariant mass distributions for events selected for data and simulation. Backgrounds are represented by the filled histograms. The shaded band represents the Poisson uncertainty on the the MC prediction. The agreement between data and simulation is not so good. Also, the mean value of the pull distribution is significantly different to zero (30%). This observed disagreement is not problematic for the purpose of this analysis because since the estimation of the background will be obtained with a data driven technique, as explained in Section 5.4.1.

Figure 5.8 shows the signal efficiency × acceptance of the event selection as a function of mass hypothesis for gluon fusion produced resonance combining all the events (a) and for the different categories (b). The efficiency × acceptance is determined for each mass hypothesis computing the fraction of events which satisfy all the selection requirements on the vertex position and on the photon isolation, together with requirements on the $p_T/m_{\gamma\gamma}$ and on the pseudorapidity of the two photons. The shown systematic uncertainties are described in Section 5.6. The increasing of the efficiency times acceptance as a function of the mass is expected and reflects the improvement observed in the vertex determination efficiency as a function of the mass shown in Figure 5.6. Also one should think that the greater the energy



Category 0

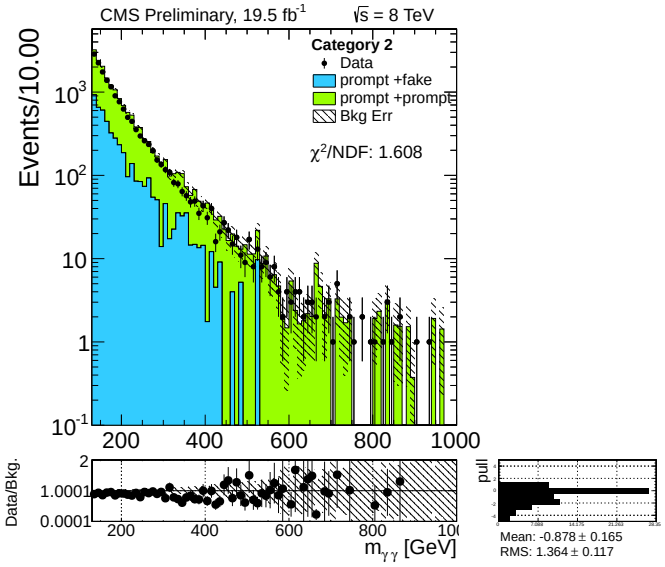


Category 1

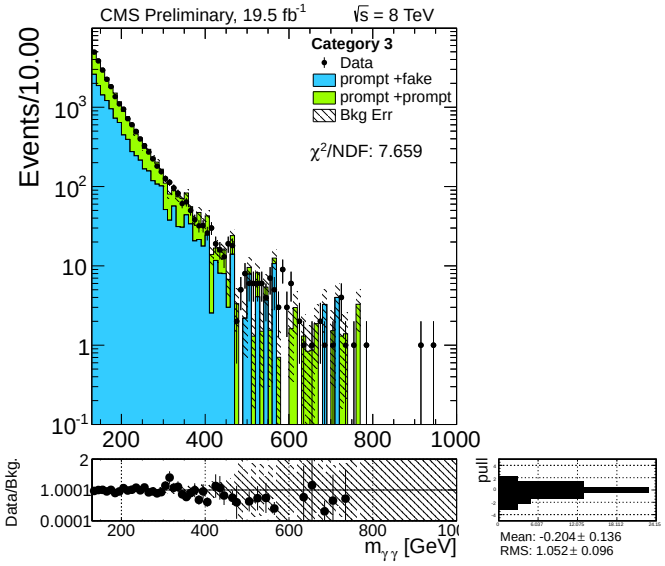
of the photon, the better the performance one gets in the photon identification.

5.4 Background Model

The distribution of background events for each channel described in Table 5.3 is estimated with a data driven technique since we do not intend to fully rely on the simulation to estimate the expected background yields after applying the event selection. This approach is the same exploited in the SM H→ $\gamma\gamma$ analysis in



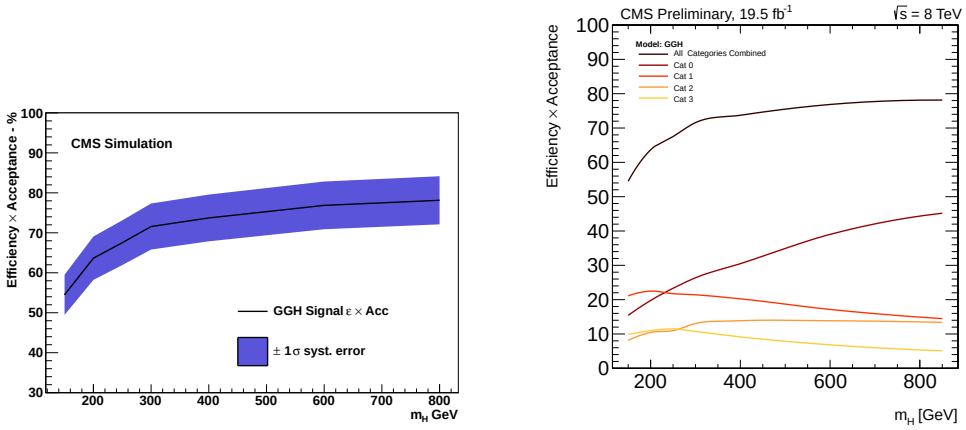
Category 2



Category 3

CMS [17].

The reason to describe the background using a functional form rather than measuring background shape in signal-free regions, or using Monte Carlo simulation predictions, is that inverting cuts to define a signal-free region (e.g. isolation) would lead inevitably to modified ratios of the different background components and this could affect the diphoton mass spectrum. Using simulated events would not only require very large amount of Monte Carlo events so that one of the possible models is clearly favored over the others, it would also require the Monte Carlo prediction



(a) (b)

Figure 5.8. Signal efficiency × acceptance as a function of mass hypothesis. (a) All categories combined, (b) Efficiencies for each category separated.

to be verified against data using, again, a signal-free region with the caveats already exposed. Due to the difficulties above we focus on describing the background with a functional form that describes the true underlying mass distribution and does not introduce undesirable long-range effects in the exclusion limits. The desired background model is also the one that gives an accurate signal strength with as few model parameters as possible.

To determine the continuum background shape over such a large mass range ($[150 - 850]$ GeV), a sliding $M_{\gamma\gamma}$ window fit range is exploited. Section 5.4.1 investigates different analytical functions for the background estimation together with the data fit performance of each model. Section 5.4.2 proceeds to the description of the bias study technique exploited to define the fit function and the fit parameters. The results of the bias study and the final fit to data are finally reported in Section 5.4.3.

5.4.1 Background-Only Bias Study

The first step in the determination of the background shape is defining a set of functions to test as the possible functional forms describing the true $M_{\gamma\gamma}$ background distribution (i.e. the distribution to which signal-free data would tend if the integrated luminosity were increased) in the signal expected region. Using m as a shorthand for $M_{\gamma\gamma}$, we define the following five analytic functions:

1. $$\mathrm{f}_0(m) := e^{-p_1 m}\, m^{-p_2}, \tag{5.1}$$
   with two free parameters p_1 and p_2.

2. $$\mathrm{f}_1(m) := (1-x)^{p_1}/x^{p_2+p_3\, log(x)}, \tag{5.2}$$
   with $x = m/\sqrt{s}$ and three free parameters p_1, p_2 and p_3.

3.
$$\mathrm{f}_2(m) := e^{-m/(p_1+p_2\,m)}, \tag{5.3}$$
with two free parameters p_1 and p_2.

4.
$$\mathrm{f}_3(m) := (1-f)\cdot m^{-4} + f\cdot m^{-5}, \tag{5.4}$$
with one free parameters f.

5.
$$\mathrm{f}_4(m) := (1-f)\cdot e^{-p_1 m} + f\cdot e^{-p_2 m}, \tag{5.5}$$
with three free parameters f, p_1, p_2.

The first three functional forms have been used in previous dijet searches [117] to describe both data and QCD predictions, while the latter two models are frequently used in diphoton searches to describe the background shape. The fits of each of these models in each event class to the data, are shown in Figure 5.9- 5.12 for the mass range $[300, 740]$ GeV. In the bottom part of the figures the ratio of each fit model with respect to the data is shown. All these models fit the data very well. This is also reflected in the χ^2 and χ^2 probability values for the different fits, shown in Table 5.4-5.7, which are very similar among the different models in each category. It is worth to note that the errors used in the χ^2 computation are simply the associated poissonian uncertainties to the bin content in the data histograms and that these fits have been performed blind, without looking at data points until the analysis was approved within the collaboration.

Model	$\chi^2/ndof$	Prob($\chi^2/ndof$)
f_0	0.58	0.99
f_1	0.59	0.99
f_2	0.58	0.99
f_3	0.56	0.99
f_4	0.59	0.99

Table 5.4. χ^2 and χ^2 probability of the ft to data with the different models in category 0

Model	$\chi^2/ndof$	Prob($\chi^2/ndof$)
f_0	0.63	0.99
f_1	0.64	0.98
f_2	0.65	0.58
f_3	0.61	0.99
f_4	0.63	0.98

Table 5.5. χ^2 and χ^2 probability of the ft to data with the different models in category 1

This set of analytical functions aims to cover the range of possible truth-models sufficiently to render the conclusions drawn reliable to a satisfactory level. We also tried to use the Bernstein polynomial as truth model, as done in the $H \to \gamma\gamma$ SM

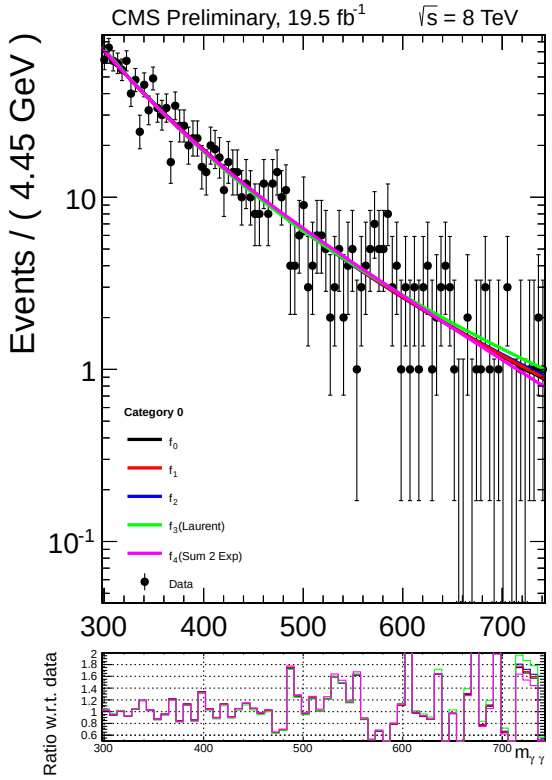


Figure 5.9. Cat 0 : Fit to data with several tested models in the region analyzed for a resonance with mass of 450 GeV. In the bottom plots for each category, the ratio of each model with respect to f_0 is shown.

Model	$\chi^2/ndof$	Prob($\chi^2/ndof$)
f_0	0.80	0.87
f_1	0.81	0.84
f_2	0.89	0.73
f_3	1.12	0.23
f_4	0.81	0.84

Table 5.6. χ^2 and χ^2 probability of the ft to data with the different models in category 2

Model	$\chi^2/ndof$	Prob($\chi^2/ndof$)
f_0	1.33	0.047
f_1	1.37	0.032
f_2	1.48	0.019
f_3	1.71	0.0036
f_4	1.34	0.042

Table 5.7. χ^2 and χ^2 probability of the ft to data with the different models in category 3

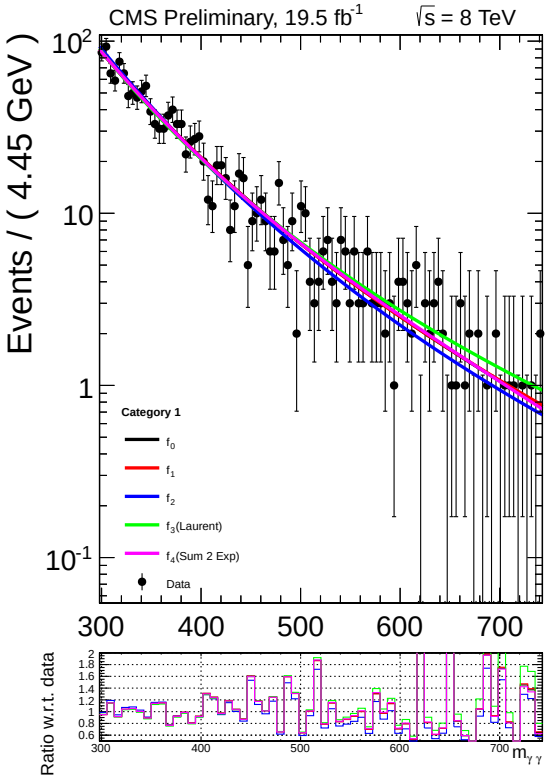


Figure 5.10. Cat 1: Fit to data with several tested models in the region analyzed for a resonance with mass of 450 GeV. In the bottom plots for each category, the ratio of each model with respect to f_0 is shown.

search [17]. Unfortunately the fits to data in the different ranges did not converge easily, also increasing the degree of the polynomial itself.

Due to no prior knowledge of the true background shape in data it is essential to chose, among this set of models, the one that is flexible enough to describe the background reliably without introducing biases on the measured signal yields/exclusion limits, but is also suitably constrained such that the fit is as sensitive as possible to a signal peak. The choice of the fit model has indeed a direct effect on the analysis sensitivity: a model that under-predicts (over-predicts) the number of events in the signal region will worsen (improve) the final exclusion limits in the absence of signal.

The first test we do in order to verify which is the most flexible model among these five functions, is to assume that each of them could somehow describe the unknown true mass distribution, and test each model against each other performing a fit. We will choose the model which will be less biased among all. To do that, pseudo-experiments describing the possible experimental outcomes of the $M_{\gamma\gamma}$ shape are randomly generated using f_0, f_1, f_2, f_3 and f_4 functions as template for the background. We test then each fit model against each other truth model for each signal mass point (150,250,350,450,550,650,750,850 GeV) and for each category (0,1,2,3). For each combination 1000 background-only simulated datasets

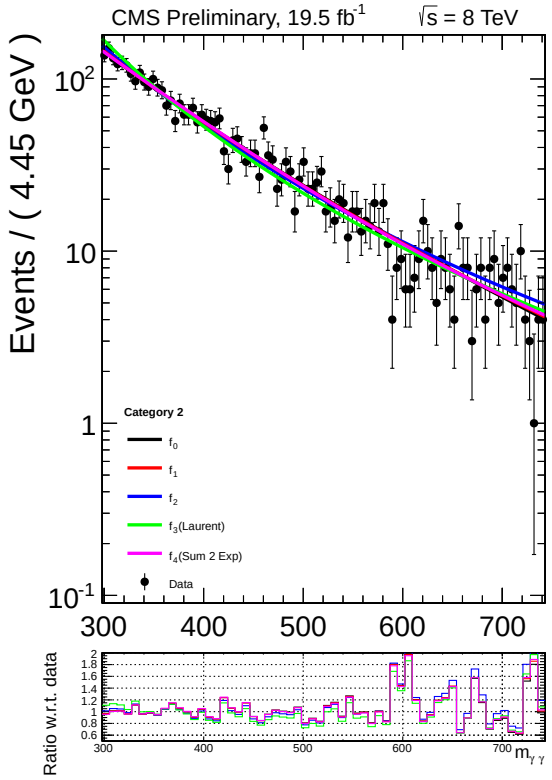


Figure 5.11. Cat 2 : Fit to data with several tested models in the region analyzed for a resonance with mass of 450 GeV. In the bottom plots for each category, the ratio of each model with respect to f_0 is shown.

are generated and fitted with the *signal + background* pdf. The number of background events generated is equal to the number of events expected in the analyzed fit range. The signal pdf is described in Signal 5.5. We then compute the pulls on the number of fitted events and we define *Bias* of the fit the ratio between the fitted mean number of signal events (whose expected value is zero) and the uncertainty on the fitted number of signal events:

$$\text{Bias}(m_{\text{H}}) := \frac{N_{\text{fit}}^{\text{sig}}}{\Delta N_{\text{fit}}^{\text{sig}}} \tag{5.6}$$

In order for the fit-model to be labeled adequate it is required that the maximal potential bias introduced by its choice against all the truth-models and over the entire mass range of interest is negligible (i.e. $< 15\%$). This assures that the deviation between the truth model and the candidate background function does not exceed five times the uncertainty of the candidate prediction. Among the available functions f_0 is the one that gives, at the end of the study, the smaller bias against all the other truth models. Therefore f_0 has been chosen as fit model to describe the background shape in this analysis.

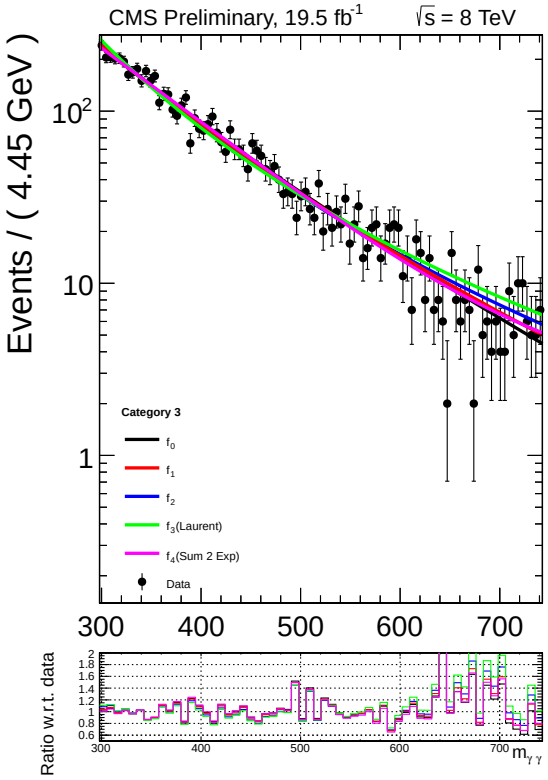


Figure 5.12. Cat 3 : Fit to data with several tested models in the region analyzed for a resonance with mass of 450 GeV. In the bottom plots for each category, the ratio of each model with respect to f_0 is shown.

The approach which exploits the fit technique to extract the number of signal and background events cannot be used for masses above 850 GeV due to the very low numbers of events in data. This is visible in Table 5.8 which shows the number of events in the very high mass region for this analysis.

Range	N° of events
$[700, 950]$ GeV	85
$[950, 1200]$ GeV	23

Table 5.8. Number of events in the very high mass region

850 GeV is therefore the highest value of the mass we consider in this search for new resonances. For masses above this value another analysis strategy, alternative to the fit technique, is needed for the estimation of the background. The lowest value of mass considered is fixed to 150 GeV and the entire fitting range is fixed to $[130, 1000]$ GeV. This lowest value has been chosen in order to have a fit range far enough from the recent observed scalar at 125 GeV, whose signal is not added in the fitting procedure. In order to extend the search to resonances with a natural width ranging from 0.1 to several GeV, the study of the bias is performed also varying the

width of the resonance over the full range of mass investigated. The maximum value of the width of the resonance has been fixed to 10% of the resonance mass itself. This value is limited by the width of the resonance mass points at the edges of the range (150 and 850 GeV) which have to be included in the fitting range ([130, 1000] GeV) within at least one sigma. The final fits of f_0 to data are shown in Figures 5.13 for the four categories for the mass range [240 , 640] GeV.

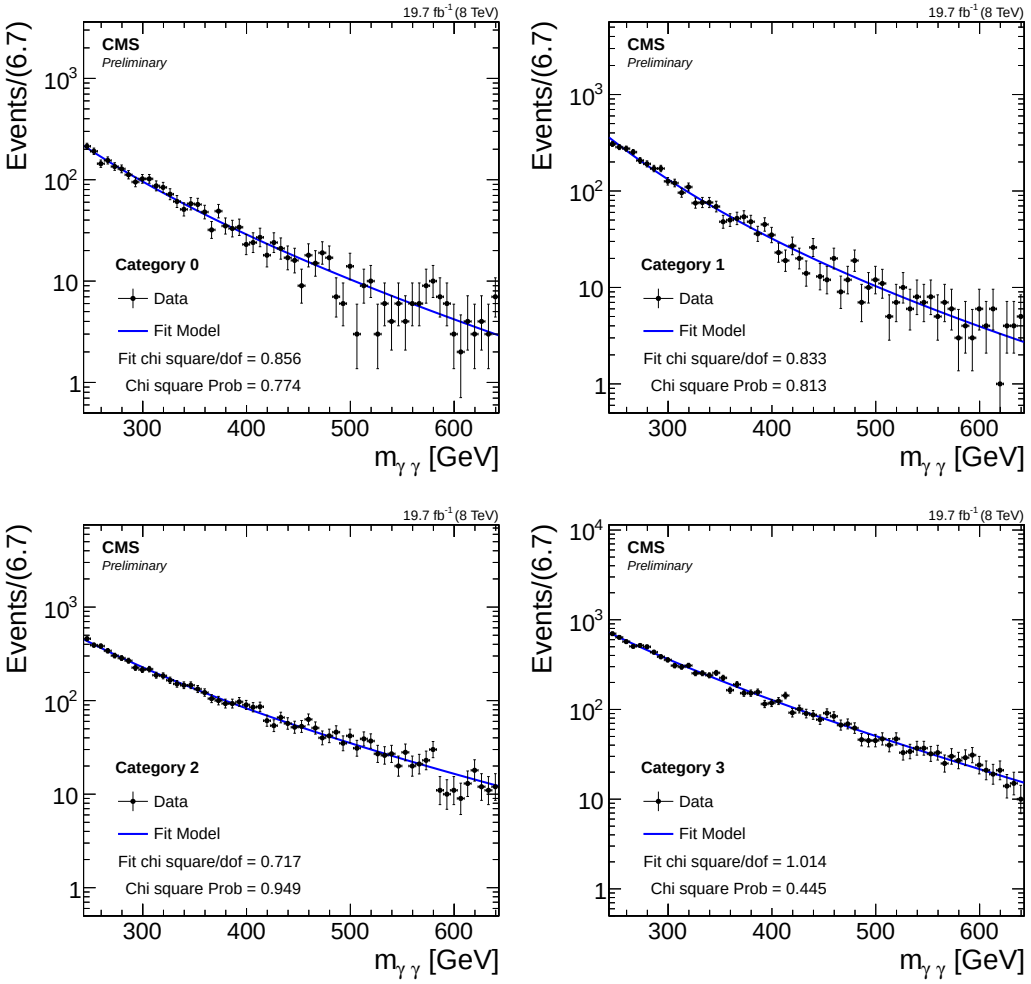


Figure 5.13. Fit to the f_0 model for the 4 categories of events in the invariant mass range $[240, 640]$ GeV.

5.4.2 Determination of the Sliding Window Fit Range

This section aims to explain how, the fit function being fixed, the edges of the sliding window fit range are chosen for each mass hypothesis. Fixed the mass of the resonance, the lower and the upper bound of the fit range are varied in order to find out the best fit range which minimizes the bias. For each combination of upper and lower bounds the bias study is performed and the bias as defined in Equation 5.6 is

computed. In order to be conservative, this study is performed considering a wide resonance with a width of 10% of the mass itself, since the bias increases increasing the width of the resonance. Fit boundaries are computed using the events in the first category (both photons with high R_9 and in the barrel). For each mass and for each truth model is then possible to complete a grid where the lower bound of the fit is indicated in the x axis and the upper bound of the fit in the y axis. On the z axis (identified with a color scale) the bias for each combination is reported. In Figure 5.14 an example of these maps is shown for a mass of 650 GeV.

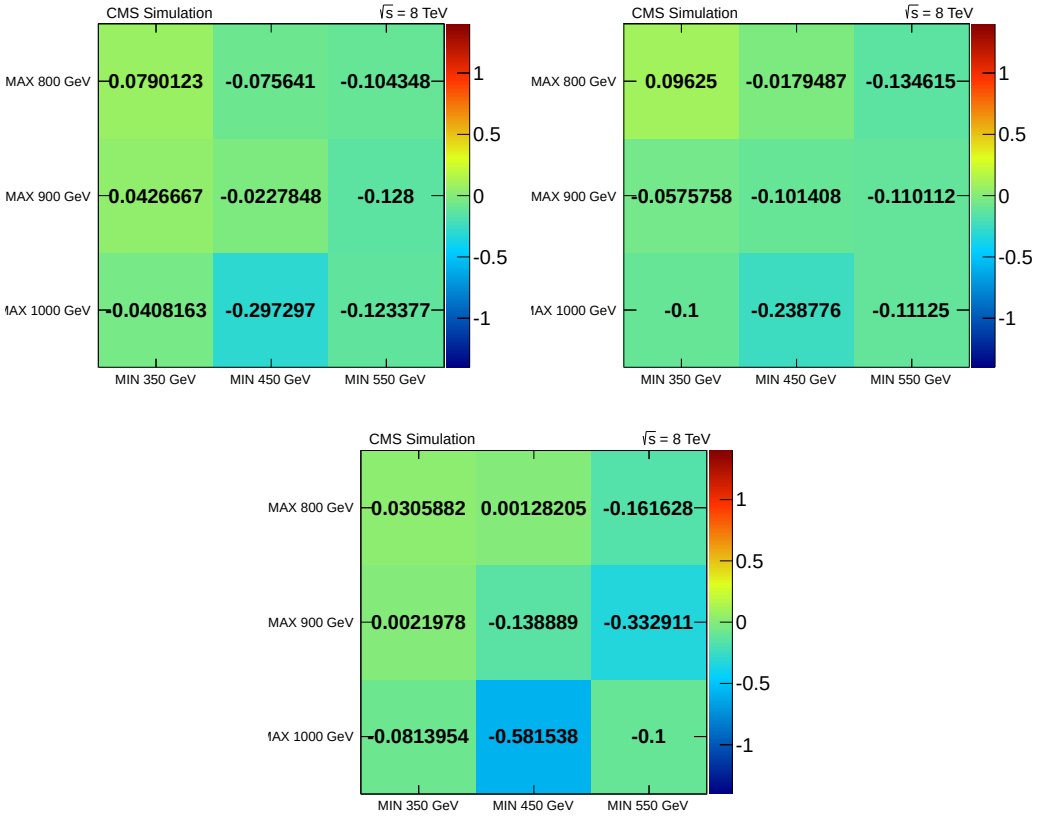


Figure 5.14. Maps with the bias study results for a mass of 650 GeV. In the top-left plot the truth model is the f_0 function itself, in the top-right the truth model is the f_1 and in the bottom plot the truth model is the f_2 function.

For this example, the lower bound is varied from 350 GeV to 550 GeV and the upper bound from 800 GeV to 1000 GeV . We find that for a mass of 650 GeV the best combination, which minimizes the bias for all the truth models hypotheses is the $[450, 800]$ GeV fit range. For the two masses at the edges of the range (150 GeV and 850 GeV) only one of the two boundaries (upper and lower respectively) is varied for this study. In Figure 5.15 the bias for a mass of 150 GeV scanning the upper bound of the fit range is shown for each truth model. The best combination in this case is $[130, 230]$ GeV. In Table 5.9 the upper and lower bounds for the fit range for each mass, obtained with the procedure explained above, are reported. The trend of

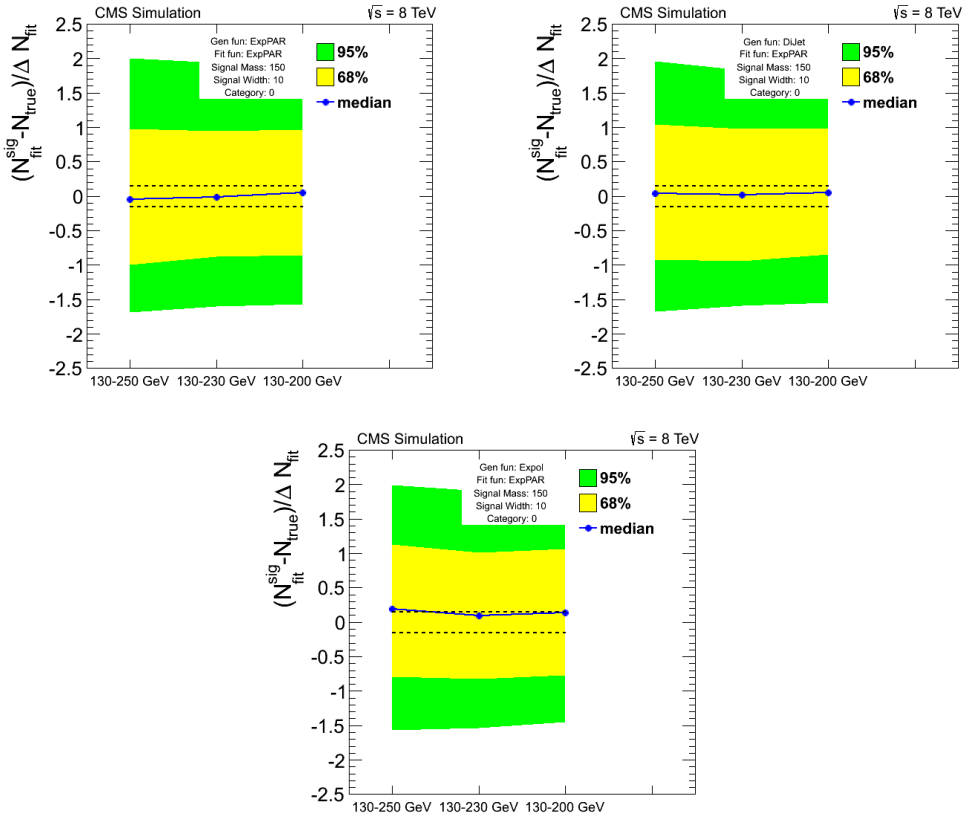


Figure 5.15. Scan of the results of the bias study for a mass of 150 GeV. In the right plot the truth model is the f_0 function itself, in the top-left the truth model is the f_1 and on the bottom the truth model is the f_2 function.

Mass [GeV]	Lower Bound [GeV]	Upper Bound [GeV]
150	130	230
250	200	550
350	250	650
450	300	700
550	300	800
650	450	800
750	450	1000
850	550	1000

Table 5.9. Upper and lower boundaries for the sliding window background fit.

the upper and the lower bounds as a function of the mass is shown in Figure 5.16. The points are fitted with polynomial functions to obtain a continuous trend of the bound vs the mass. The upper bounds are fitted with a 2 degree polynomial while the lower bounds are fitted with a 1 degree polynomial.

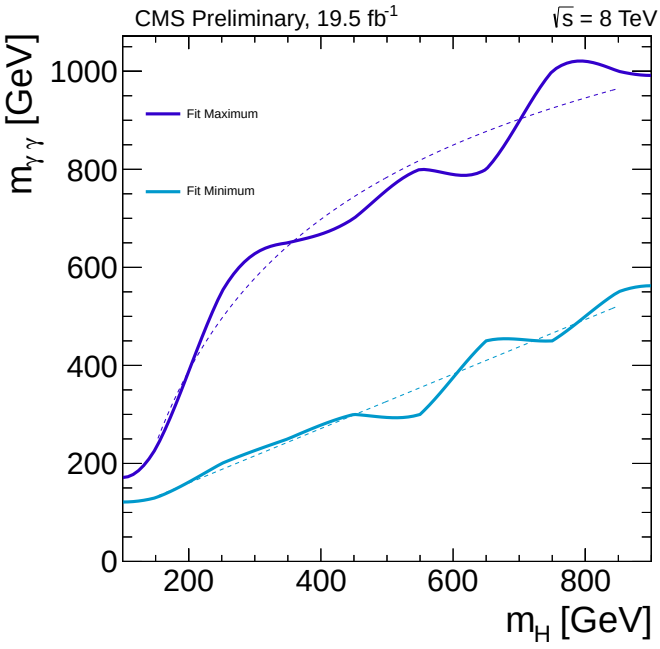


Figure 5.16. Trend of the upper and of the lower bounds of the fit range as a function of the mass.

5.4.3 Bias Study Results

As a closure test the bias as a function of the mass is computed again using for each mass the upper and lower bounds of the fit given by the two polynomial functions as shown in Figure 5.16. The trend of the bias for each truth model and for events in the first category is shown in Figure 5.17. As expected the bias is negligible over the full mass range.

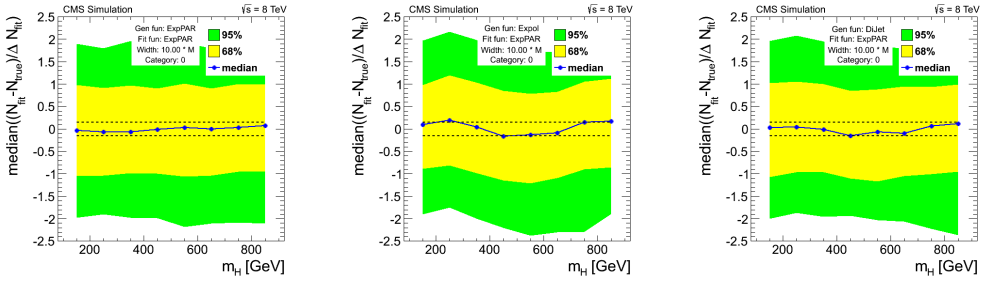


Figure 5.17. Trend of the bias as a function of the resonance mass for events in CAT 0. The width of the resonance is equal to 10% of the mass itself.

Several further tests have been performed to check the goodness of this study.

- The bias is negligible also for the other three categories as shown in Figures 5.18.

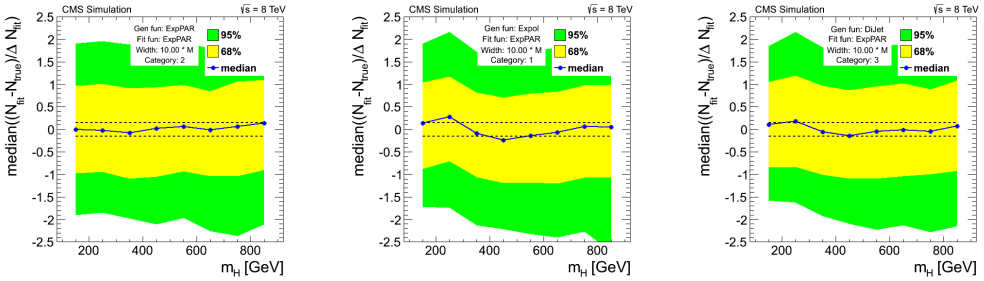


Figure 5.18. Trend of the bias as a function of the resonance mass for events in other categories. The width of the resonance is equal to 10% of the mass itself.

- The bias is negligible also if we consider narrow resonances with width $\Gamma_H = 0.1$ GeV as shown in Figure 5.19.

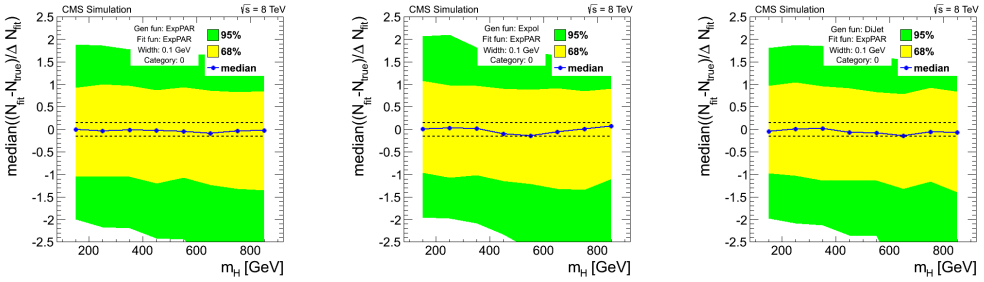


Figure 5.19. Trend of the bias as a function of the resonance mass for events in CAT 0. The width of the resonance is equal to 0.1 GeV.

- The bias is negligible also if we use as truth models two additional functions (f_3 and f_4) as shown in Figure 5.20.

- The bias is negligible also if we repeat the study injecting the signal in simulation toys. This is particularly interesting to verify that the chosen background model is suitably constrained. For each mass we generate a number of signal events equal to the number of events corresponding to the expected UL at 95% CL that we are able to exclude for that mass. If after the *signal+background* fit we get a number of signal events compatible with what we generated, this means that the background function is completely de-correlated from the signal and our technique could be sensitive to any possible peak, if present. The result of this study is shown in Figure 5.21.

All these tests demonstrate that f_0 is a good unbiased analytic model for the description of our background shape in any fit range and thus it will be used to model data in the statistical analysis described in Section 5.7.1.

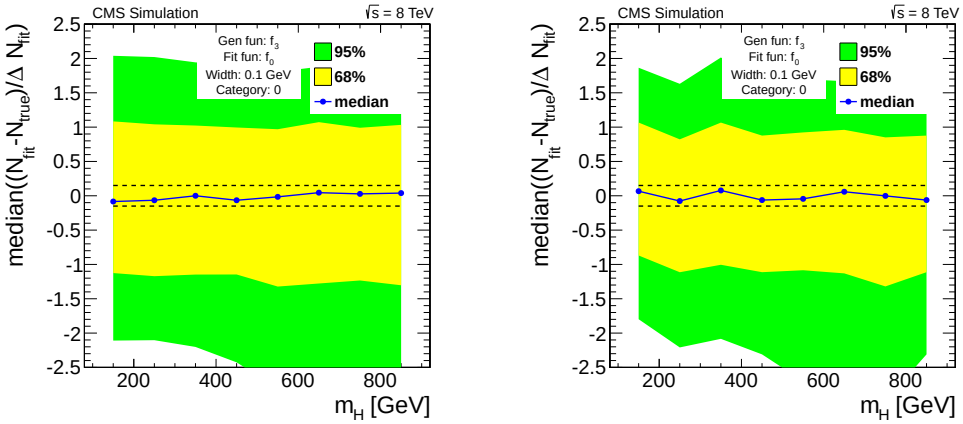


Figure 5.20. Trend of the bias as a function of the resonance mass for two additional truth models. The width of the resonance is equal to 0.1 GeV.

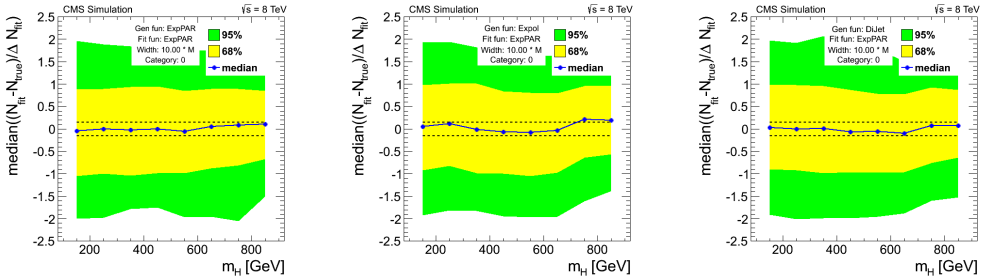


Figure 5.21. Trend of the bias as a function of the resonance mass for events in CAT 0 when a signal is injected. The width of the resonance is equal to 0.1 GeV.

5.5 Signal model

The signal shapes, which are mainly dominated by the detector and reconstruction response in the ECAL, need to be well described in each of the four categories of events. The Monte Carlo is used, after the smearing of the resolution, and the application of all efficiency corrections and scale factors, to build a parametrized model for the signal. The strategy for the shape model is to describe the signal Monte Carlo with an analytic function in which the hypothetical signal mass (m_X) represents a parameter which can vary continuously for any value in the range $[150 - 850]$ GeV. This can be extended in a simple fashion to allow for an additional free parameter, the natural width of the new resonance, Γ_X, by convoluting the signal shape with a Breit-Wigner function which has a mean of m_X and a width of Γ_X. Performing a 2D scan of m_X and Γ_X a search can be performed as a function of the mass and of the natural width of the new resonance.

5.5.1 ECAL Response

As first step the performance of the ECAL in the mass measurement (response function) need to be evaluated for the different signal hypotheses. The distribution of the reduced mass ($\frac{\Delta m}{m}$) is thus computed for each available mass point as the ratio between the reconstructed diphoton mass in the event (m_{reco}) and the true mass (m_{true}) of the event minus one (Figures 5.22- 5.23):

$$\frac{\Delta m}{m} = \frac{m_{reco}}{m_{true}} - 1 \tag{5.7}$$

In order to construct the parametric model, the response function is fitted with an analytic function, namely the sum of two single-sided Crystal-Ball functions (Equation 5.9) with common mean m_0 and sigma σ and different n and α. We choose to use the Crystal-Ball function (CB) $f_{CB}(\frac{\Delta m}{m})$ to describe our signal model because this function combines a Gaussian core and a power-law tail with an exponent n to account for photon energy loss due to pair production and bremsstrahlung. The parameter α defines the transition between the Gaussian and the power-law functions:

$$f_{CB}(x) = \begin{cases} \frac{N}{\sqrt{2\pi}\sigma} exp(-\frac{(x-x_0)^2}{2\sigma^2}), & \text{for } \frac{x-x_0}{\sigma} > \alpha; \\ \frac{N}{\sqrt{2\pi}\sigma} \left(\frac{n}{|\alpha|}\right)^n exp(-\frac{|\alpha|^2}{2}) \left(\frac{n}{|\alpha|} - |\alpha| - \frac{x-x_0}{\sigma}\right)^{-n} & \text{for } \frac{x-x_0}{\sigma} \leq \alpha \end{cases} \tag{5.8}$$

The resulting fit function:

$$\text{Response}\left(\frac{\Delta m}{m}\right) = (1-f)\cdot f_{CB_1}\left(\frac{\Delta m}{m}|m_0, \alpha_1, \sigma, n_1\right) + f \cdot f_{CB_2}\left(\frac{\Delta m}{m}|m_0, \alpha_2, \sigma, n_2\right) \tag{5.9}$$

has seven free parameters; in addition to the four parameters used in Equation 5.8, one parameter is the fraction of the second Crystal-Ball to the total signal yield (f) and the other two are α_2 and n_2 of the second CB.

In Figures 5.24- 5.25 the fit to the reduced mass is shown for the signal mass point $m_X = 150$ GeV in the four categories.

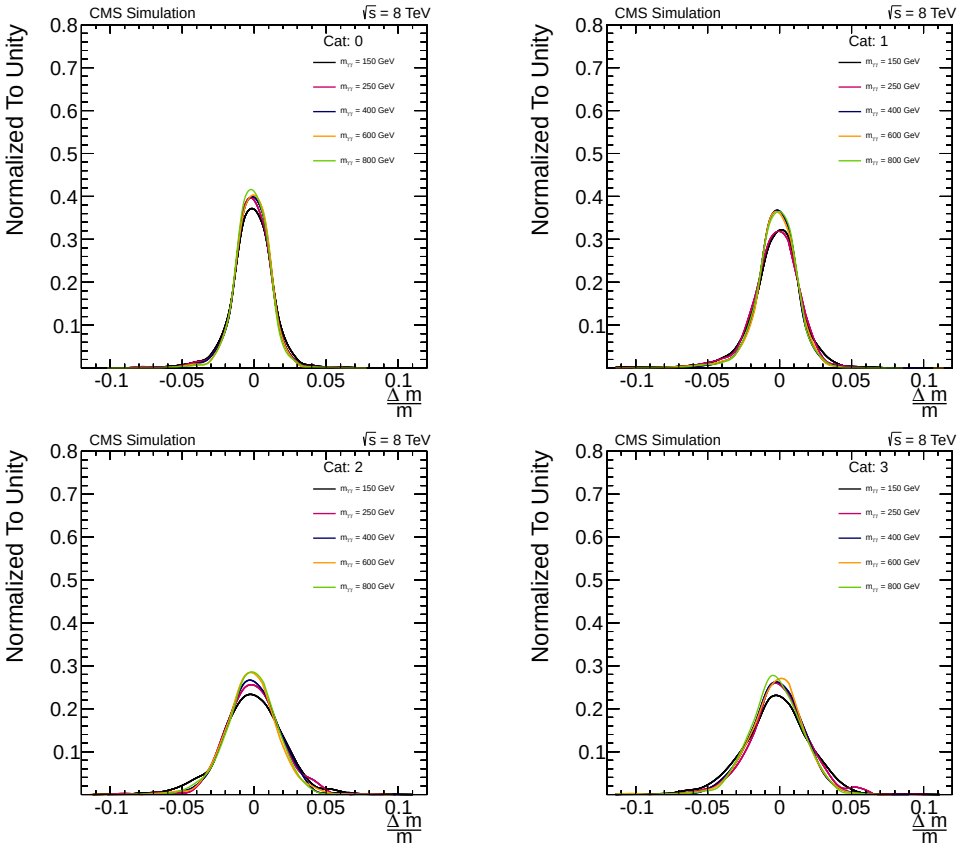


Figure 5.22. Signal shapes for different categories. Linear scale.

Figures 5.23 clearly demonstrate that while simulated signal shapes are almost centered at the same value, the resolution improves increasing the mass of the resonance. This is visible also in Figures 5.26 which show the trend of the mean and the sigma of the signal distributions as a function of the mass for the different categories as given from the fit to the response function together with their statistical uncertainties. All the fit parameters, except μ_{CB} and σ_{CB} have been fixed to the values of the fit at $m_X = 150$ GeV.

Thus, in order to assume the same shape for the response distribution for any mass hypothesis, m_X, the resolution improvement with respect to the 150 GeV case shown in Figure 5.27, is modelled with a 2-degree polynomial and the sigma of the signal model is parametrized as a function of m_X as detailed in Equation 5.10.

$$\sigma_{CB}(m_X) = (p_o + p_1 \cdot m_X + p_2 \cdot m_X^2) \cdot \sigma_{CB}(150) \tag{5.10}$$

5.5.2 Theoretical Signal Shape

The theoretical lineshape of an unstable particle is usually described at NLO in Monte Carlo generators with a Breit-Wigner (BW) distribution [74]. However in

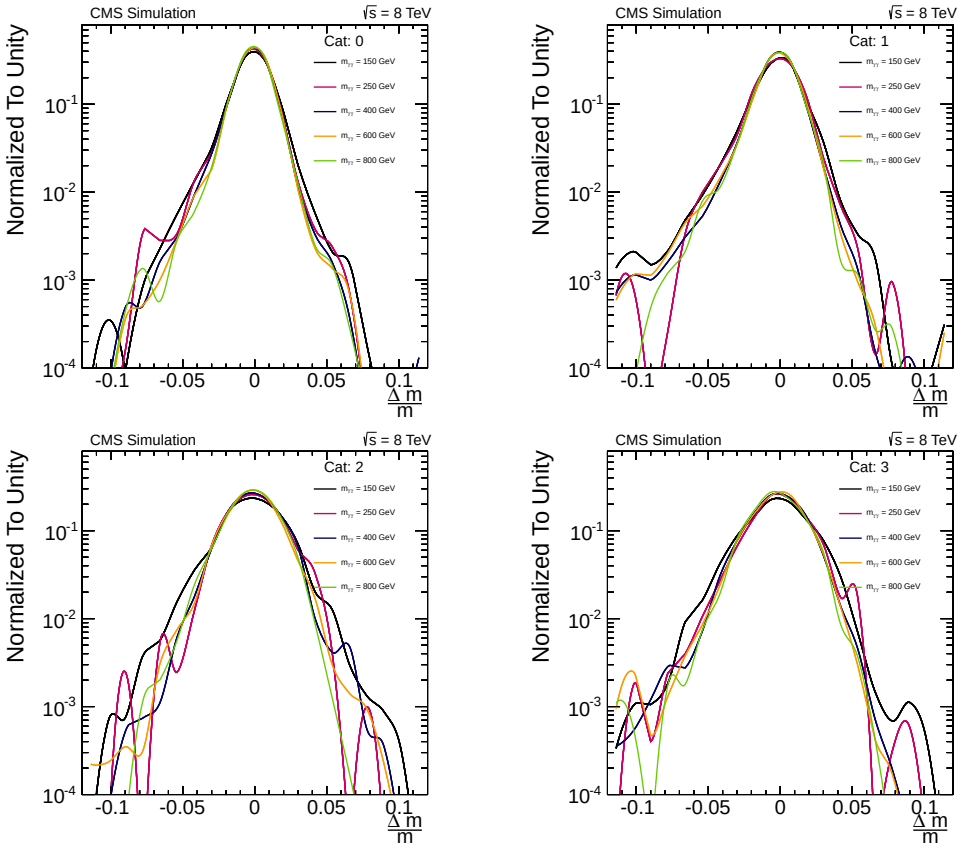


Figure 5.23. Signal shapes for different categories. Logarithmic scale.

the event of a high mass resonance there are several effects to be taken into account for the description of its theoretical lineshape:

- The signal shape is influenced not only by the natural width of the state, but also by the proton PDFs [118].
- A more correct approach to describe the lineshape has been proposed, known as Complex Pole Scheme (CPS) [118], which is valid for a SM high mass Higgs bosons, whose natural width increases rapidly above 400 GeV. In the event of whichever wide high mass resonance, as the case of our analysis, it may be not needed to use this parametrization if one wants to be as much model independent as possible. We study this kind of model just to understand the size of the problem.
- The interference between the signal and the non resonant background.

Effect of proton PDFs It has been observed [118] that in the event of a high mass, wide resonance the gluon fusion parton luminosity introduces a low energy tail, creating an artificial increase of the lineshape at low virtualities. This effect has been studied looking at the shape of a gluon fusion produced SM Higgs boson

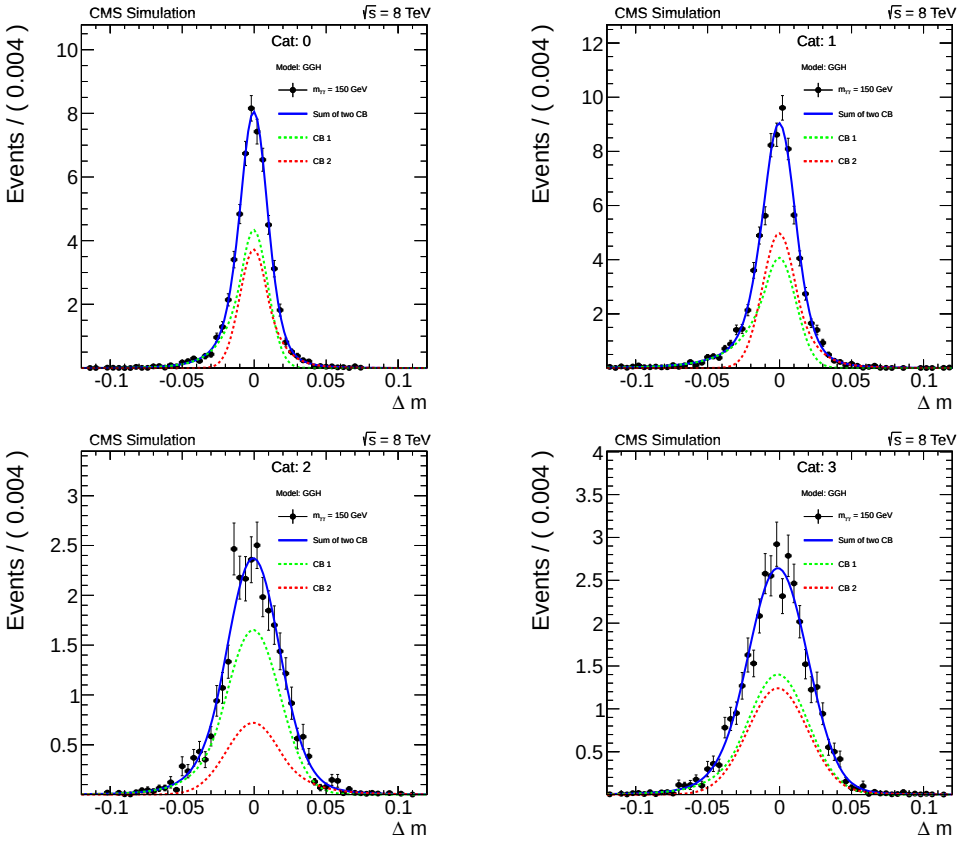


Figure 5.24. Two Crystal-Ball fit (blue line) to the response distributions for $m_X = 150$ GeV. The green dashed line is the first Crystal-Ball component while the red is the second Crystal-Ball component. Linear scale.

at high mass. We used the POWHEG MINLO [119] generator which includes the effect of the proton PDFs. We studied a SM Higgs bosons with masses: 400, 600 and 800 GeV and correspondent widths of 30, 120, and 300 GeV. Since these natural Higgs widths are greater than the widths considered in our analysis which ranges up to 10% of the resonance mass for each mass hypothesis (see Section 5.4.2) we assume that any possible estimation of the observed effect will be conservative for our analysis.

We compare the generated shapes with a modified Breit-Wigner model defined as:

$$BW(m_X, \Gamma_X) = \frac{1}{(m^2 - m_X^2)^2 + m_X^2 \Gamma_X^2} \tag{5.11}$$

in order to evaluate the effect of the distortion at different masses and widths in Figures 5.28

It is visible from Figures 5.28 that the effect of the distortion increases with the mass as expected. Nonetheless the global effect results negligible(i.e. $< 1\%$) and the BW model of Equation 5.11 can be safely used to describe our resonances.

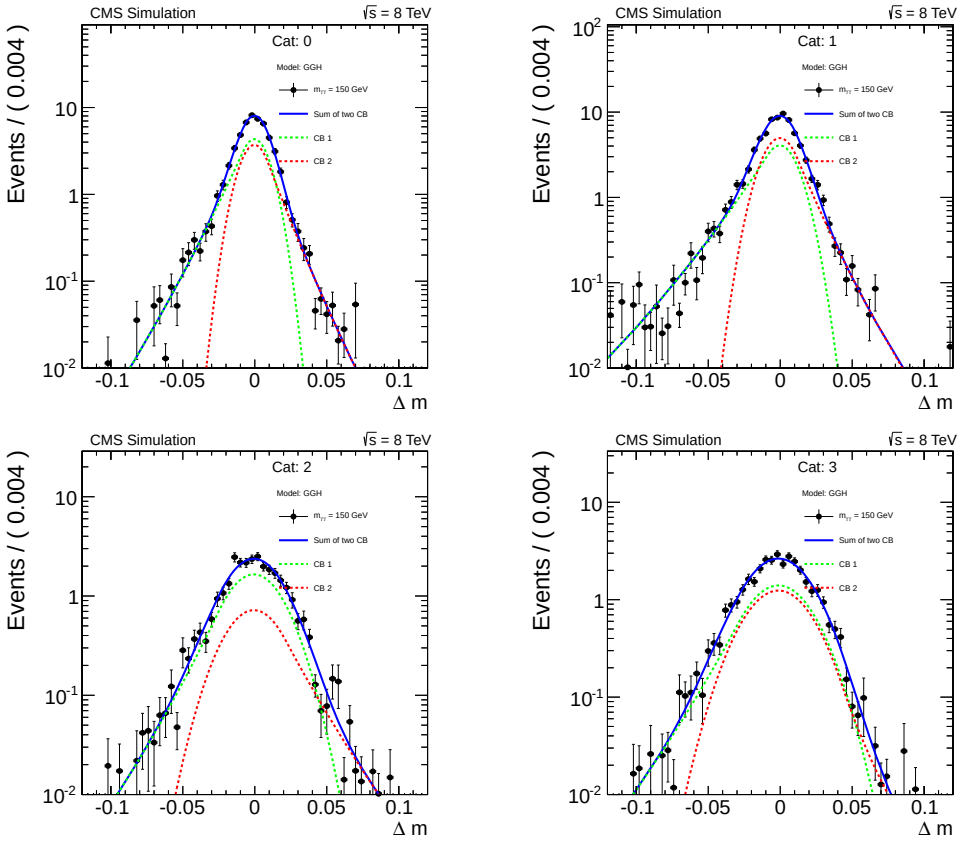


Figure 5.25. Two Crystal-Ball fit (blue line) to the response distributions for $m_X = 150$ GeV. The green dashed line is the first Crystal-Ball component while the red is the second Crystal-Ball component. Logarithmic scale.

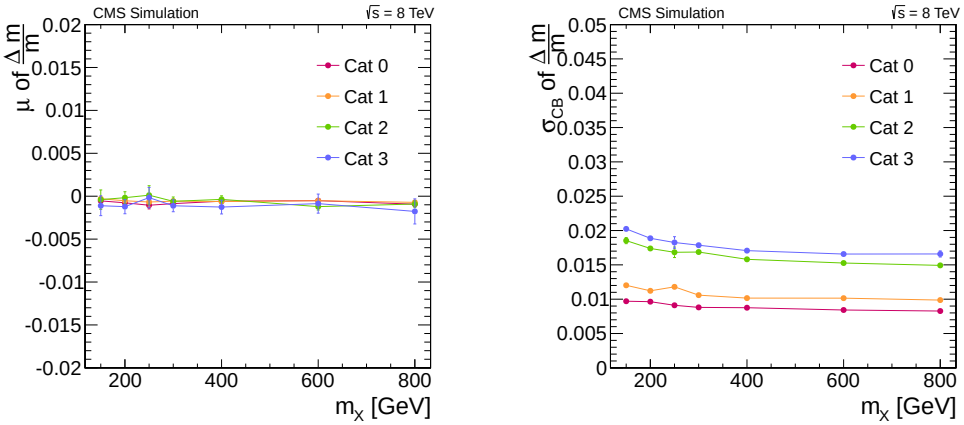


Figure 5.26. Trend of μ_{CB} e and σ_{CB} as a function of the mass for the different categories.

Complex Pole Scheme Approximation The description of a resonant theoretical lineshape via the zero-width approximation (ZWA) (narrow-width approximation)

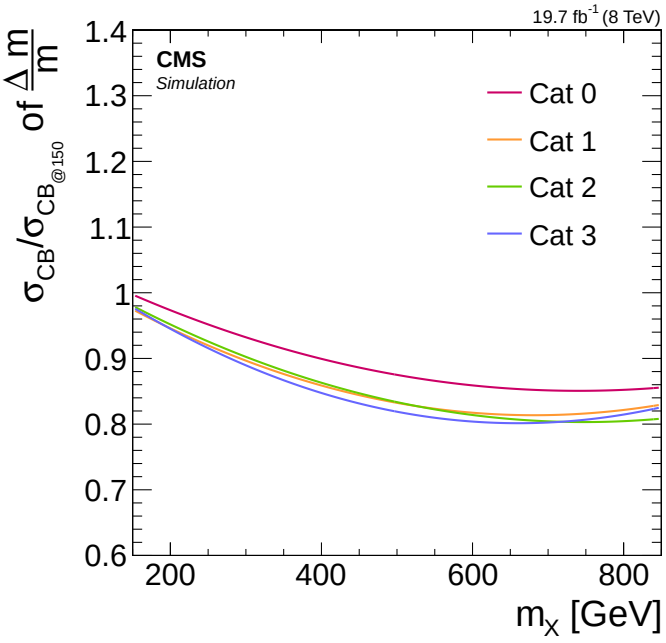


Figure 5.27. Signal resolution improvement as a function of the mass hypothesis with respect to the 150 GeV mass.

exploits the asymptotic equality of the squared modulus of the unstable (scalar) particle propagator with 4-momentum q to $\pi/(M\Gamma)\delta(q^2 - M^2)$ in the limit $\Gamma \to 0$ [120]. The Dirac delta function restricts the unstable particle to on-shell states without otherwise affecting the production and decay subprocesses. In general, an uncertainty of $O(\Gamma/M)$ is expected for the ZWA. For the particles involved in this search one finds that Γ/M is of $O(10\%)$, which implies a ZWA error that is similar to other theoretical and experimental errors in the analysis. However, even if the ZWA is in this case applicable, it may be interesting to evaluate the size of the effect. In the event of a heavy SM Higgs boson with $m_H = 300$ GeV, for instance, the expected ZWA uncertainty is of O(3%) and increases to O(20%) for $m_H = 600$ GeV. When ZWA is evidently not adequate for a heavy Higgs boson, a more detailed description of the lineshape, known as Complex Pole Scheme (CPS), is required and has been developed recently in Refs. [118], [121].

In order to evaluate the effect of using in this analysis a simple BW description (Equation 5.11) instead of a more complicated CPS lineshape, we look at the shape of a gluon fusion produced SM Higgs boson at high mass. We use the POWHEG 15 [122] generator which includes the effect of the proton PDFs and the CPS. We describe these distributions with [123]:

$$CPS(m_X, \Gamma_X, c', \delta) = f(m_{\gamma\gamma}) \frac{m_{\gamma\gamma}}{(m_{\gamma\gamma} - (m_X + \delta)^2)^2 + (m_{\gamma\gamma} c' \Gamma_X)^2} \tag{5.12}$$

where c' and δ are two parameters of the CPS theory and $f(m_{\gamma\gamma})$ is a universal factor, which has been found to improve the description of the signal tails around the pole mass. In Figures 5.29 the CPS shapes for three different Higgs mass hypotheses

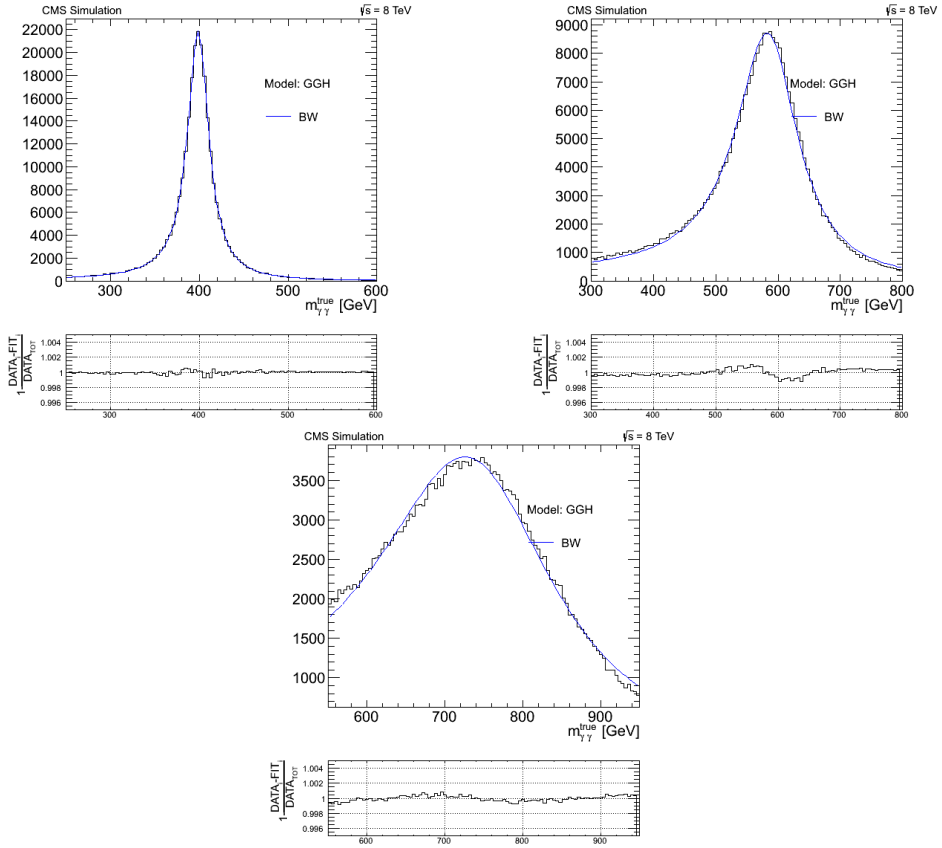


Figure 5.28. Top plots: POWHEG MINLO true mass distribution for a SM Higgs boson with masses 400, 600 and 800 GeV and natural widths of 30, 120, and 300 GeV. Superimposed is a Breit-Wigner fit as given in Equation 5.11. Bottom: Bin by bin difference between prediction and fit normalized to the data integral.

are shown. It is visible that in this model the effective width of the Higgs becomes narrower ($c'\Gamma_X$ with c' less than 1) and the shape distorted in the tails.

Fixed the width of the resonances (40,60, 80 GeV for masses of 400, 600 and 800 GeV) the expected and observed upper limit at 95% of CL as detailed in Section 5.7.2 are computed exploiting both the BW model in Equation 5.11 and the CPS model in Equation 5.12. The results obtained with the BW model are 2-10% more conservative w.r.t. those with the CPS model. This is mainly due to the fact that in the CPS the effective width of the resonance is narrower. We conclude that the simple BW model can be safely used in this analysis if a systematic uncertainty of order Γ/M is added in the statistical procedure as detailed in Section 5.6.

Interference effect The observation of a new state decaying in pairs of photons at the LHC turns out to be one of the main discovery channels and, therefore, requests precise theoretical calculations for the corresponding cross section. The cross-section of the main background, which is the non resonant diphoton production, has been recently computed up to NNLO [124], but the interference between signal

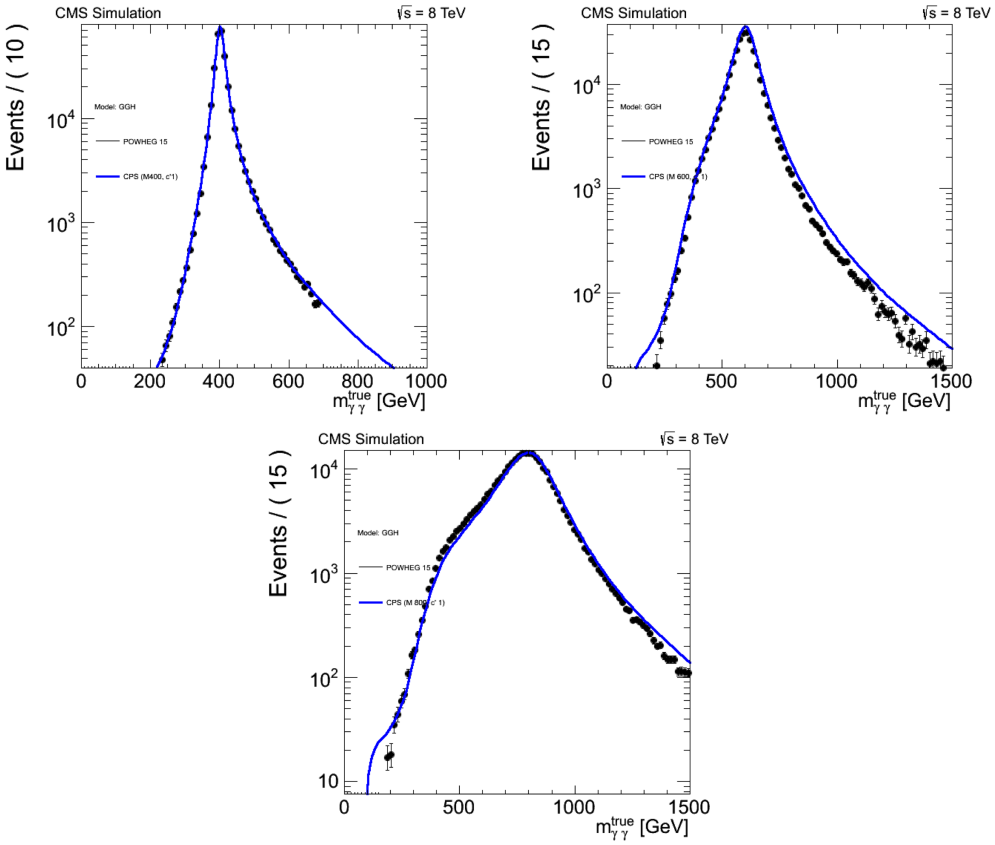


Figure 5.29. POWHEG 15 true mass distribution for a SM Higgs boson with CPS fit function as given in Equation 5.12.

and background has not been evaluated to such level of accuracy yet. In particular the interference of the resonant process $ij \to X \to \gamma\gamma$ with the continuum QCD background $ij \to \gamma\gamma$ induced by quark loops can have several contributions which are not quantified so far for the mass range analyzed in this search. However this interference effect should be negligible in the hypothesis of narrow resonances and could become relevant only in the event of a wide resonance. To determine the effect of the interference the amplitude and the complex phase of the signal should be known but this is not the case of this analysis, where the search is completely model independent. The only strategy that we can adopt in order to give just an estimate of the effect, is to assume a SM-like Higgs particle and evaluate the effect of the interference under this hypothesis. Interference computation for a generic spin 0 gluon fusion produced resonance which decays to two photons with arbitrary width has been implemented in Sherpa MC generator [125] by theorists from Standford specifically for this search. Signal plus interference processes have been generated for different mass hypotheses (i.e. 200, 400, 800 GeV). Since interference effects are expected to increase, increasing the width of the resonance, the conservative situation of widths equal to 10% of the mass is considered. The effect of including or not the interference in the analysis is evaluated comparing the results of the search

(the upper limits on the cross section times branching ratio of the X→ $\gamma\gamma$ process as described in Section 6.1) using either the simple Breit-Wigner function (Equation 5.11, described in Section 5.5.3) or the template given by Sherpa as theoretical model for the resonance. Figures 5.30 show the true mass distributions as given from Sherpa when interference effect is included. The differences in the observed and expected limits including or not the interference are of the order of 3% and can be neglected since smaller that the total systematic error (described in Section 5.6) considered in the analysis.

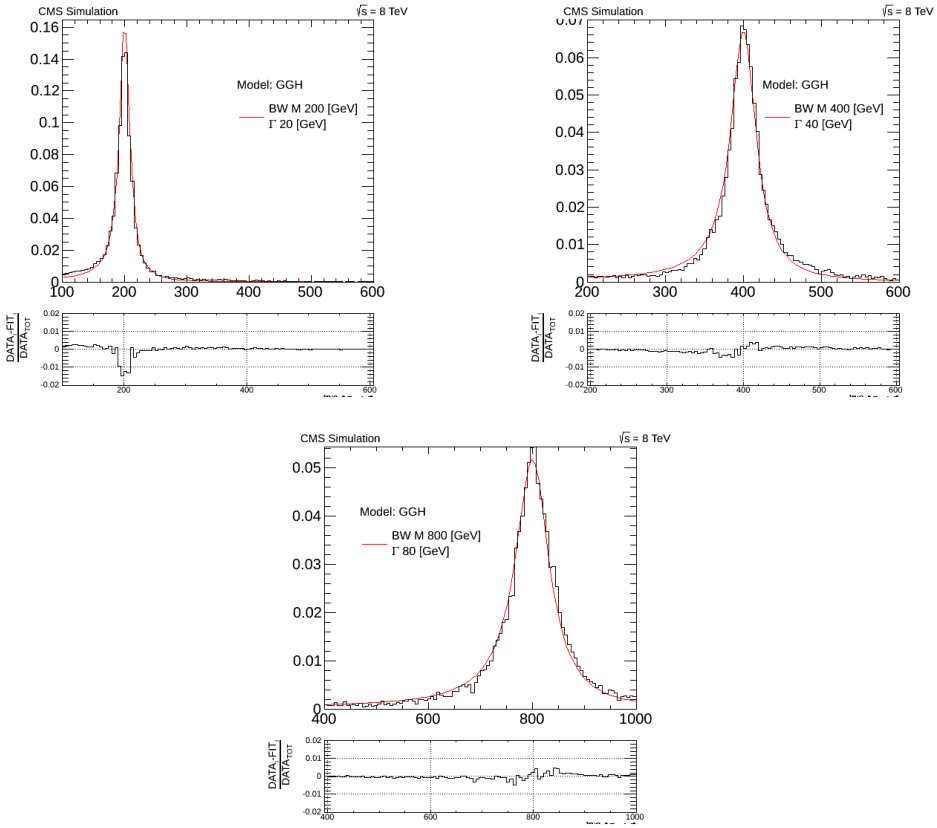


Figure 5.30. Signal plus interference shape for the true mass distribution for three different mass hypotheses as obtained with Sherpa MC generator. Overlaid is a Breit-Wigner fit. In the bottom of the plots the difference between the histogram (labelled as "data") and the fit is reported.

5.5.3 Parametric Signal Shape

The theoretical signal line shape described in Equation 5.11 is convoluted with the response function (Equation 5.9) to account for the experimental resolution, to

parametrize the reconstructed signal mass distribution:

$$\text{Signal Model}(m) = \text{Response}(m|m_0, \alpha_1, \alpha_2, n_1, n_2, \sigma(m_X)\cdot m_X, f) \otimes BW(m|m_X, \Gamma_X) \tag{5.13}$$

The signal model in Equation 5.13 depends continuously upon the parameter m_X through the sigma of the response function and the position of the peak in the Breit-Wigner and upon the parameter Γ_X, through the width of the Breit-Wigner. This signal model is tested fitting the signal distribution for the mass points available in Monte Carlo, as shown in Figures 5.31-5.32. The width of the Breit-Wigner for these fits is fixed to $\Gamma_X = 0.1$ GeV, since the signal has been generated with this nominal width. In Figure 5.33 different signal shapes simulating resonances with natural widths greater than zero are shown.

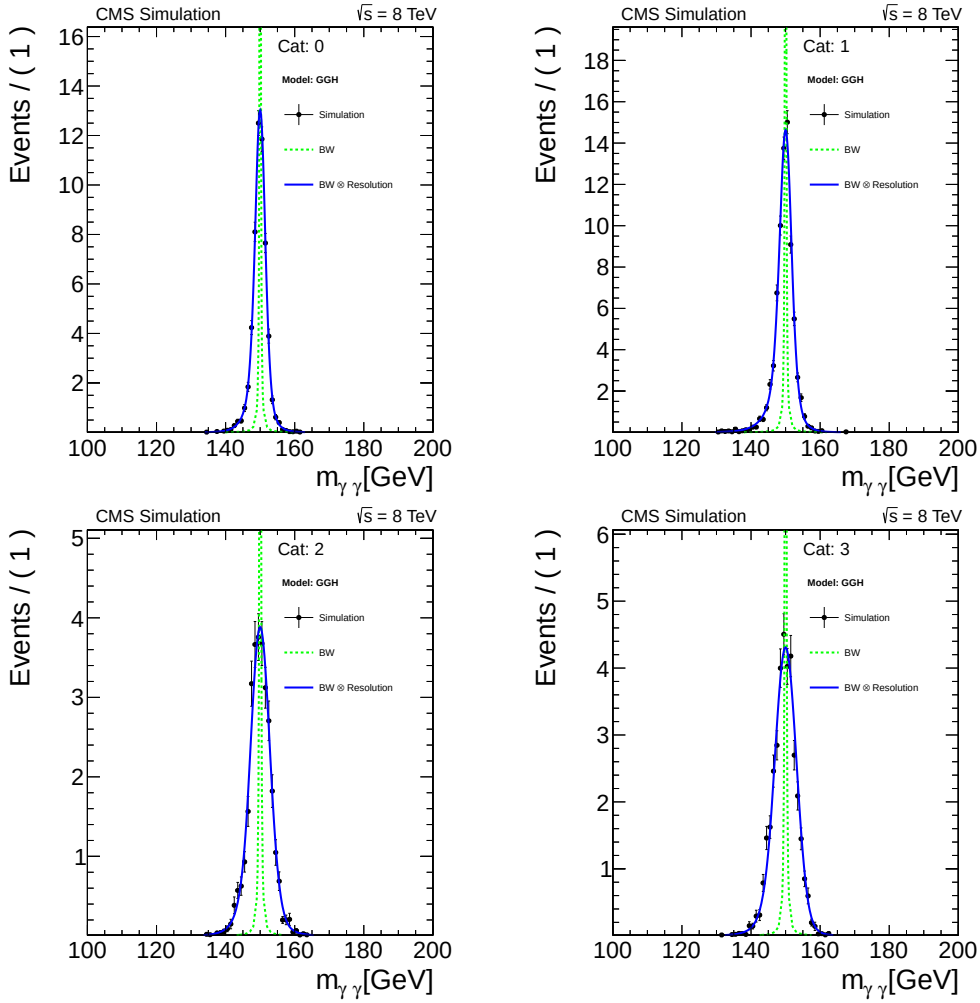


Figure 5.31. Full parametrized signal shape for the 4 classes for $m_X = 150$ GeV signal in linear scale. The black points are the weighted Monte Carlo events and the blue line is the corresponding model. The green line represents the Breit-Wigner component of the model.

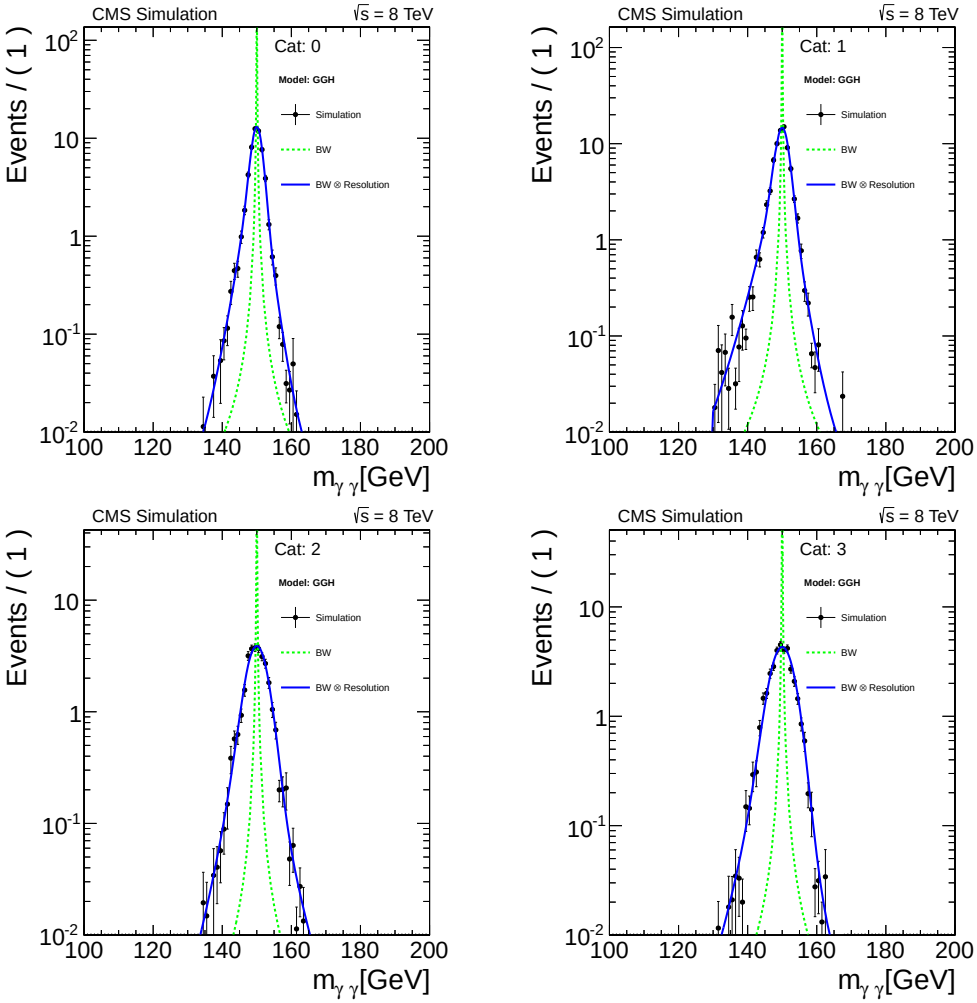


Figure 5.32. Full parametrized signal shape for the 4 classes for $m_X = 150$ GeV signal in logarithmic scale. The black points are the weighted Monte Carlo events and the blue line is the corresponding model. The green line represents the Breit-Wigner component of the model.

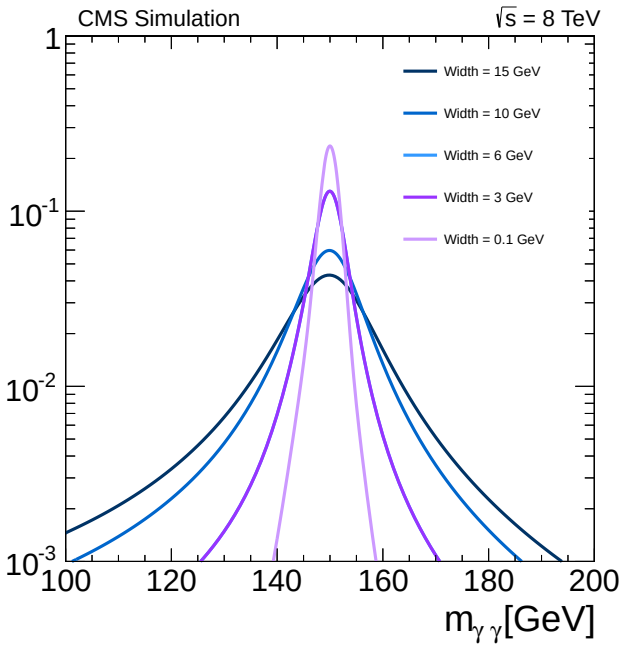


Figure 5.33. Simulated signals with different widths.

5.6 Systematics

In this section, the systematic uncertainties related to common features of all the event categories are described. Since the background is modelled in a fully data-driven manner, the result does not rely on the Monte Carlo prediction for the diphoton shape of the various background components. The signal modelling is driven instead by the Monte Carlo prediction with relevant Monte Carlo to data correction factors applied.

The experimental systematic uncertainties affecting the analysis can be separated between systematics affecting the yield and systematics affecting the shape of the signal itself. Systematics affecting the yield are incorporated as correlated log-normal uncertainties on the class yields while systematic uncertainties affecting the shape are incorporated as parametric variations of the model.

Single photon systematics account for the photon identification efficiency, in particular for the efficiency of the R_9 cut which reflects in a migration of the photons among the categories, and for the ECAL energy scale and resolution. Single event systematics account for the uncertainty on the measured integrated luminosity, for the vertex finding efficiency and for the efficiency of the diphoton trigger.

5.6.1 Photons Related Systematics

The photon related uncertainties can be classified in the following sources:

- Photons selection efficiency: normalization uncertainty on the efficiency of the selection of signal photons. It is defined as the conservative envelope shown on

the Figure 5.34 which shows the photon efficiency uncertainty as a function of the photon p_T applied to a high mass signal ($M = 400$ GeV).

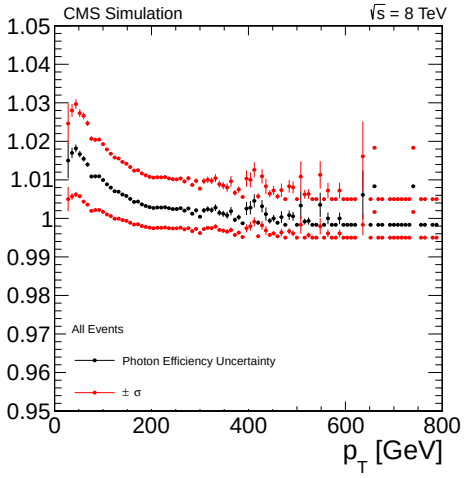


Figure 5.34. Photon efficiency uncertainty as coming from the SM H→ $\gamma\gamma$ analysis applied on a high mass resonance of 400 GeV [17].

- Diphoton trigger acceptance: normalization uncertainty on the signal HLT efficiency, 1%.
- Photon energy scale uncertainty: modifies the mean position of the signal model. We evaluate this source of uncertainty by systematically shifting all reconstructed photon transverse momenta by ±1 standard deviation of the measured photon energy scale uncertainty and then recomputing the invariant mass of the system and the resolution distribution used in the parametric signal model. For each category we evaluate the shift (positive and negative) of the mean of this distribution and we use this shift to propagate the systematics on the energy scale in the signal model and thus in the limit extraction. Figures 5.35 show the resolution shapes applying the energy scale systematics. Numeric results are shown in Tab 5.10

Category	Standard	Scale UP	Scale DOWN	Shift
0	-0.00028	0.00470	-0.00534	0.5%
1	-0.00115	0.00392	-0.00605	0.5%
2	-0.00110	0.00705	-0.00988	0.7%
3	-0.00064	0.00659	-0.00791	0.7%

Table 5.10. Mean position of the signal model applying the energy scale systematics to photons.

- Photon energy resolution uncertainty: is typically 200 MeV as given by the most recent CMS analyses [17].

  We apply an additional gaussian smearing to the photon energy to simulate a worse photon transverse momentum resolution. The invariant mass of the

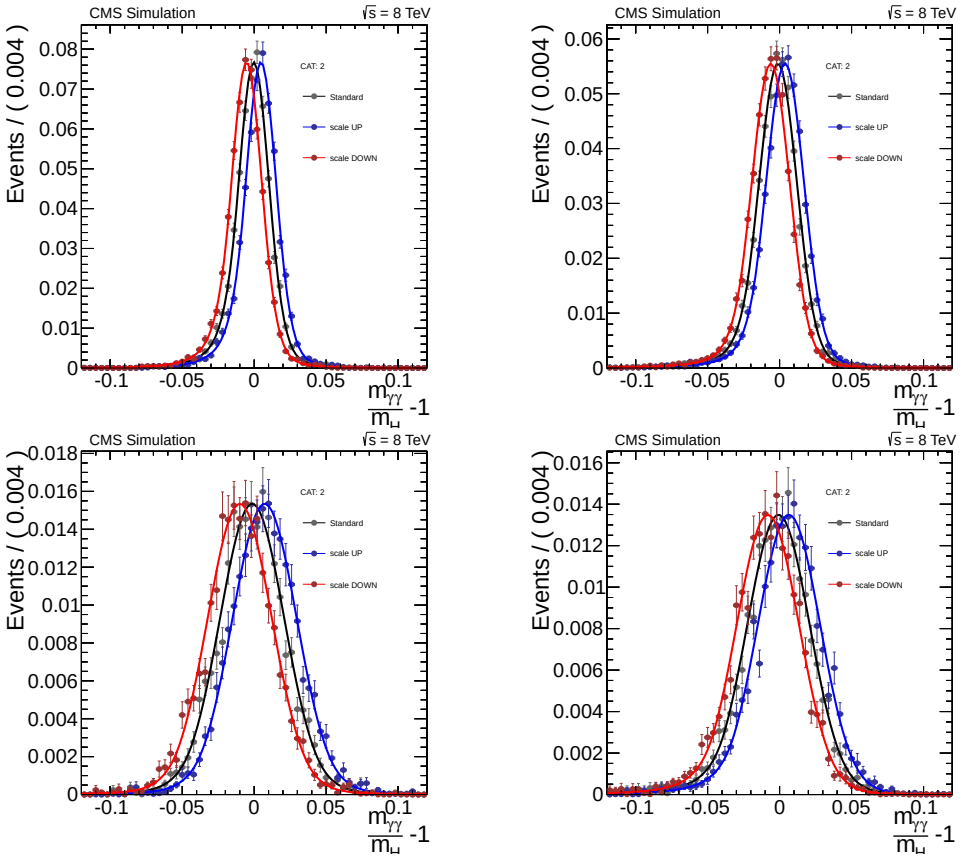


Figure 5.35. Shifts observed in the resolution function after applying the energy shift due to the systematic uncertainty on the photon energy scale. Results are shown separately for the four categories.

diphoton system in then recomputed together with the resolution distribution used in the parametric signal model. For each category we evaluate the smearing (positive and negative) of the σ of this distribution and we use this smearing to propagate the systematics on the energy resolution in the signal model and thus in the limit extraction. Figures 5.36 show the resolution shapes applying the energy resolution systematics. Numeric results are shown in Tab 5.11

Category	Standard	Smear UP	Smear DOWN	Smear
0	0.00624	0.00795	0.00501	0.5%
1	0.00872	0.01049	0.00706	0.58%
2	0.01538	0.01834	0.01289	1%
3	0.01606	0.01817	0.01435	1%

Table 5.11. Resolution values of the signal model applying the uncertainty related to the photon energy resolution.

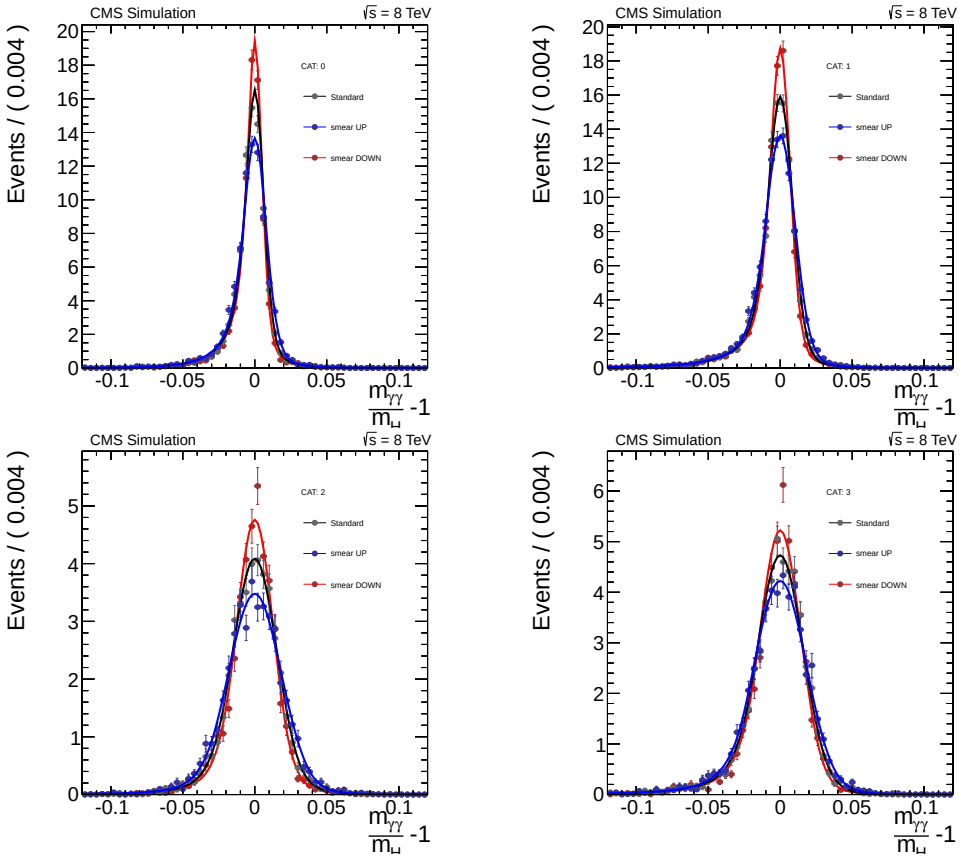


Figure 5.36. Smearings observed in the resolution function after applying the energy smear due to the systematic uncertainty on photon energy resolution. Results are shown separately for the four categories.

Since both the uncertainty on the energy scale and on the energy resolution are evaluated looking at $Z \to ee$ events in data and MC, the main source of systematic uncertainty in the determination of the energy shift and the additional energy smearing to be applied to the signal MC is the difference between electrons and photons . The final estimate of the systematic uncertainty is obtained comparing Z$\to$ ee events and $H \to \gamma\gamma$ events with the Higgs mass equal to 90 GeV. For this reason the systematics estimated for the SM Higgs search [17] can be safely used for this analysis as well for low mass resonances where the bulk of the photon p_T distribution is close to the one of the photons coming from the Standard Model Higgs (Figure 4.7). This is not the case for high mass resonances, therefore additional photon systematics have to be considered. To account for this, we assign an additional normalization systematics of 5% for diphoton pair which can be motivated in the following way:

- The ECAL energy resolution is improving with energy, therefore the extra smearings applied to get a better data to MC agreement will be smaller at high masses.

- Possible ECAL non linearities at high energy were studied in the contest of the Z'→ee CMS analysis for energies above 350 GeV and found to be <1% [126].

5.6.2 Other Sources of Systematics

The luminosity uncertainty is taken to be 2.6% [127].

As detailed in [17] we consider two further sources of systematics: the uncertainty related to the vertex finding efficiency and the class migration due to R9 cut efficiency.

Table 5.12 reports all the relevant systematic uncertainties accounted for the analysis.

Sources of systematic uncertainty	**Uncertainty**	
Per photon	**Barrel**	**Endcap**
Energy resolution $R_9 > 0.94$ (low η, high η)	0.05%, 0.10%	0.07%, 0.03%
Energy resolution $R_9 < 0.94$ (low η, high η)	0.05%, 0.09%	0.09%, 0.06%
Photon energy scale	0.5%	2%
Photon identification efficiency	1.0%	2.6%
Per event		
Integrated luminosity	2.2%	2.5%
Vertex finding efficiency	0.2%	0.2%
Trigger efficiency	1.0%	1.0%
R_9 category migration	2.3%	5.5%
Additional normalization systematic	5%	5%

Table 5.12. Sources of systematic uncertainty on the signal accounted for in the analysis, and applicable to events in all classes.

In general the analysis is observed to be statistics limited and for both resonances with 0.1 GeV width and wide resonances with width of 10% of the mass itself the systematic uncertainties worsen the results by at most 2%.

5.7 Likelihood Fit and Signal Extraction

The fit technique used to extract the signal yield in this analysis is based on the construction of the combined probability distribution of all measurements in the data sample. No excess has been observed over the SM background expectation and limits have been set on the product of the cross section times branch ration of the X→ $\gamma\gamma$ process. Section 5.7.1 starts with the description of the likelihood function and the fit method used to estimate the signal yield. Section 5.7.2 gives a summary of the procedure used to set upper limits on the production rate of the new scalar in the analyzed energy range. The outcomes of the limits setting procedure are shown in the chapter dedicated to the results and to their interpretation 6.

5.7.1 Maximum Likelihood Fit to Data

The Maximum-Likelihood (ML) [128] is the fit technique exploited in this analysis for estimating model parameters which best characterize the observed data. The first ingredient of the ML method is the model which describes the data. This model is given by the sum of the probability density function modelling the background and the signal. As detailed in Section 5.4 the estimation of the background $f_b(x)$ is taken directly from data while the signal parametrization $f_s(x)$ is obtained in Section 5.5 from simulation as a function of the mass and the natural width of the resonance. The model depends on input parameters which can be classified as either nuisance parameters or parameters of interest. A nuisance parameter(or systematic effect) is any parameter that is not under investigation but still has an impact on the predictions. Conversely, parameters of interest are the parameters that are being constrained in the respective analysis. In this analysis the variable of interest is r which is interpreted as the signal strength with respect to a reference cross-section for the production of an hypothetical resonance over the continuum background. Once obtained the signal strength r of the hypothetical signal it will be possible to make assumption on the absolute cross-section time branching ratio of the X$\rightarrow \gamma\gamma$ process. If any excess will not be observed upper limits will be placed as described in Section 5.7.2.

The likelihood function $L(data|r, \boldsymbol{\nu})$ is constructed as:

$$L(data|r, \boldsymbol{\nu}) = Poisson(data|r, \boldsymbol{\nu}) \cdot p(\boldsymbol{\nu}|\tilde{\boldsymbol{\nu}}) \tag{5.14}$$

Here *data* represents either the actual experimental observation or pseudo-data used to construct sampling distributions to be discussed further below. The parameter of interest r is the signal strength modifier and $\boldsymbol{\nu}$ represents the full suite of nuisance parameters described in detail in Section 5.6 as systematic sources affecting the measurement. The systematic error pdfs ($p(\boldsymbol{\nu}|\tilde{\boldsymbol{\nu}})$) reflect the degree of belief on what the true parameter value $\tilde{\boldsymbol{\nu}}$ might be.

For an unbinned likelihood over k events in the data sample, $Poisson(data|r, \boldsymbol{\nu})$ stands for:

$$k^{-1} \prod_i (r \cdot S(\boldsymbol{\nu}) \cdot f_s(x_i) + B(\boldsymbol{\nu}) \cdot f_b(x_i)) e^{-(r \cdot S(\boldsymbol{\nu}) + B(\boldsymbol{\nu}))} \tag{5.15}$$

where $f_s(x)$ and $f_b(x)$ are pdfs of signal and background of the observable x, while S and B are total event rates expected for signal and backgrounds.

As the name implies, ML proceeds to maximize the likelihood function with respect to the parameter of interest, which in turn maximizes the agreement between the model and the data. The value of the parameter r which maximizes the likelihood function is the estimate $\hat{r}$ of the strength for the signal under analysis.

5.7.2 Limit Setting Procedure

The statistical procedure that is used to exclude the existence of a new boson, if any excess is not observed over the continuum background is the CL_S method [129] which provides a means of setting upper limits on cross-sections derived from a model and constrains the possible range of fundamental parameters [130]. To compare

the compatibility of the data with the background-only and signal+background hypotheses, where the signal is allowed to be scaled by some factor r, we construct the test statistic $\tilde{q}_r$ [129] based on the profile likelihood ratio:

$$\tilde{q}_r = \frac{L(data|r, \hat{\boldsymbol{\nu}}_{\boldsymbol{r}})}{L(data|\hat{r}, \hat{\boldsymbol{\nu}})} \tag{5.16}$$

where $\hat{\boldsymbol{\nu}}_{\boldsymbol{r}}$ refers to the conditional maximum likelihood estimators of $\boldsymbol{\nu}$, given the signal strength parameter r and *data* that, as before, may refer to the actual experimental observation or pseudo-data (toys). The pair of parameter estimators $\hat{\boldsymbol{\nu}}_{\boldsymbol{r}}$ and $\hat{r}$ correspond to the global maximum of the likelihood.

Observed Limit

In order to quote the 95% Confidence Level Observed Upper Limit on r the strategy is the following:

- Compute the observed value of the test statistic $\tilde{q}_r^{obs}$ for the given signal strength modifier r under test.
- Find values of the nuisance parameters $\hat{\boldsymbol{\nu}}_{\mathbf{0}}$ and $\hat{\boldsymbol{\nu}}_{\boldsymbol{r}}$ best describing the experimentally observed data (i.e. maximizing the likelihood as given in Equation 5.14), for the background-only and signal+background hypotheses, respectively.
- Generate toy Monte Carlo pseudo-data to construct pdfs $f(\tilde{q}_r; r, \hat{\boldsymbol{\nu}}_{\boldsymbol{r}})$ and $f(\tilde{q}_r; 0, \hat{\boldsymbol{\nu}}_{\mathbf{0}})$ assuming a signal with strength r in the signal+background hypothesis and $r = 0$ for the background-only hypothesis.
- Having constructed $f(\tilde{q}_r; r, \hat{\boldsymbol{\nu}}_{\boldsymbol{r}})$ and $f(\tilde{q}_r; 0, \hat{\boldsymbol{\nu}}_{\mathbf{0}})$ distributions, we define two p-values to be associated with the actual observation for the signal+background and background-only hypothesis, p_r and p_b:

$$p_r = P(\tilde{q}_r \geq \tilde{q}_r^{obs}|signal + background) = \int_{\tilde{q}_r^{obs}}^{\infty} f(\tilde{q}_r; r, \hat{\boldsymbol{\nu}}_{\boldsymbol{r}}) d\tilde{q}_r \tag{5.17}$$

$$1 - p_b = P(\tilde{q}_r \geq \tilde{q}_r^{obs}|background - only) = \int_{\tilde{q}_0^{obs}}^{\infty} f(\tilde{q}_r; 0, \hat{\boldsymbol{\nu}}_{\mathbf{0}}) d\tilde{q}_r \tag{5.18}$$

and calculate $CL_s(r)$ as a ratio of these two probabilities:

$$CL_s(r) = \frac{p_r}{1 - p_b} \tag{5.19}$$

- Finally to quote 95% Confidence Level Upper Limits on r, denoted as $r^{95\%CL}$, we adjust r until we reach $CL_s = 0.05$.

Expected Limit

In order to define the expected median upper limit and $\pm 1\sigma$ and $\pm 2\sigma$ bands for the background-only hypothesis we generate a large number of pseudo-data according to the background-only pdf as given in Section 5.4. We then calculate for each of them the CL_s and $r^{95\%CL}$ as if they were real data. Integrating the distribution of $r^{95\%CL}$ obtained for all the pseudo-data, one can build the cumulative probability distribution of the results. The point at which the cumulative probability distribution crosses the quantile 50% is the median expected value. The $\pm 1\sigma$ (68%) band is defined by crossing the 16% and 84% quantiles. Crossing at 2.5% and 97.5% define the $\pm 2\sigma$ (95%) band. The results corresponding to the observed and expected upper limits set in this analysis are shown in details in Section 6.1.

Chapter 6

Results and Interpretation

This analysis searches for a new scalar not predicted by the Standard Model which decays to two photons across the $[150-850]$ GeV invariant mass range. The strategy adopted is to make no assumptions on the natural width of the resonances. Therefore the search is extended also to resonances with natural widths ranging from zero up to several GeV (i.e. 10% of the mass hypothesis). No excesses have been observed over the SM expectation (the measured significance as shown in Figure 6.5 is less than 2 σ over the full analyzed mass range with only one local excess that does not exceed the 3 σ) and limits have been placed on the production cross section of new heavy resonances which decay to photons, as outlined in Section 6.1.

6.1 Results

The $m_{\gamma\gamma}$ invariant mass spectrum is analyzed in each of the four categories for numerous hypothetical signals, varying the mass and the width of the resonance. The properties of interest to be measured in this analysis, such as the signal and background yields, the mass and width of the resonance, are determined performing a maximum likelihood fit to the diphoton invariant mass distribution in the four event classes in the range $[130-1000]$ GeV as described in Section 5.7.

Since signal samples have been generated only for a number of reference mass points ($[150-850]$ GeV with a step of 50 GeV), where the behaviour of the expected signal is studied, the results of the analysis need to be interpolated at every intermediate mass point. The shape of the signal is thus parametrized as a function of m_X and Γ_X as described in Section 5.5.3. The selection efficiency is also given as a function of m_X in Figure 5.8 under the hypothesis of a gluon fusion produced resonance.

Given the sliding window fit range technique adopted for the background estimation and described in Section 5.4.3, it may happen that, fixed the mass of the resonance, the tails of the signal model extend beyond the boundaries of the range. This effect results to be appreciable only in the event of a wide resonance and the efficiency loss observed increasing the width is of the order of at most [0.1-0.5] %. Figure 6.1 shows the signal efficiencies as a function of the mass and the width of the resonance under the hypothesis of a gluon fusion produced signal.

As we are searching for BSM signature the signal cross section is not fixed. The

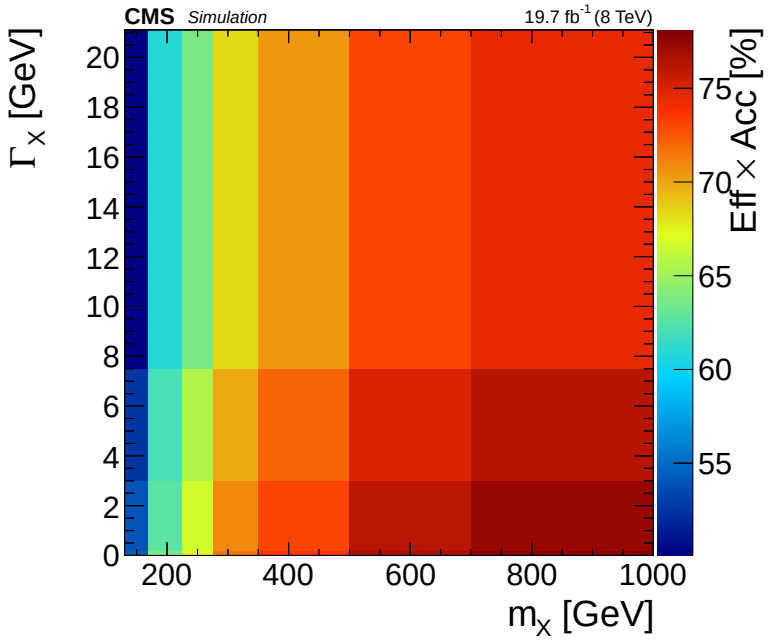


Figure 6.1. Signal efficiencies as a function of the mass and width for a gluon fusion produced resonance. Fixed the mass, the loss in efficiency observed as a function of the width is due to the possibility for the tails of the signal model to extend beyond the window of the fit range.

signal normalization is taken from the parametrization of the signal yield assuming the same production cross section for all the masses. The reference cross section is the $H \to \gamma\gamma$ production cross section at 150 GeV. In this way the signal yield parametrization takes into account only the efficiency of the selection. Results are then given in terms of absolute cross section times branching ratio. Since no excess has been observed over the background expectation, upper limits are set on the signal production. The upper limit here is defined as the cross section times branching ratio ($\sigma \times BR(X \to \gamma\gamma)$) that the analysis is able to falsify at 95% Confidence Level assuming no evidence for the signal.

Fixed the signal parameters (i.e. its mass and its natural width), a simultaneous likelihood fit [131] of the $m_{\gamma\gamma}$ distributions in the four event categories is performed as described in Section 5.7.2, exploiting the CLs modified frequentist technique [129]. The expected statistical power of the analysis is computed generating a large number of random pseudo-experiments with, as starting point, the expected normalization and shape of the $m_{\gamma\gamma}$ distributions and the corresponding systematic uncertainties. For each background-only toy a likelihood fit is performed using the signal + background model and the exclusion limit is computed at 95% of CL. The median of the results is taken as central value of the expected statistical power of the analysis, and the distribution is integrated to define 68% and 95% probability intervals about the median. These values are then compared to the observed limit, which is obtained by the fit to the analyzed data.

Observed (markers) and expected (dashed line) exclusion limits on the product

of the production cross section times the branching fraction of X→ $\gamma\gamma$ are presented in the following figures. The expected limit also shows the 68% and 95% probability ranges, respectively marked by a green and a yellow shade.

Limits are set for the four event categories as a function of the resonance mass as shown in Figure 6.2. Figures 6.3 show the combined limit for a narrow and a wide resonance in the analyzed mass range. Figures 6.4 show the limit worsening increasing the width of the resonance for two values of the resonance mass (150 and 850 GeV).

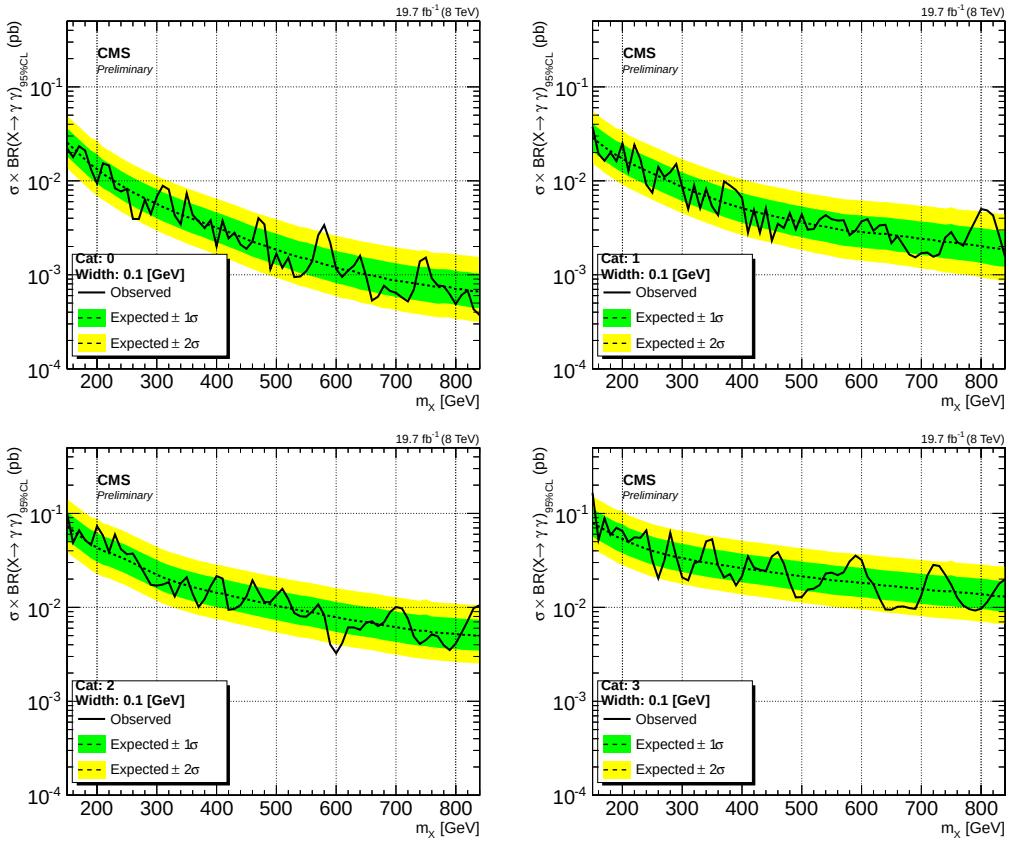


Figure 6.2. Expected and Observed exclusion limits at 95% CL on the cross section times the branching ratio of a new resonance decaying into two photons as a function of the resonance mass hypothesis for a resonance's width of 0.1 GeV.

It can be argued that a local excess over the background only expectation can be observed around 578 GeV, as visible in Figure 6.3. This excess is quantified by the background-only p-value, i.e. the probability for the background to fluctuate and give an excess of events as large or larger than the observed one. The p-value is then converted into a significance in terms of standard deviation as detailed in [131]. As reference, an excess of 5 σ corresponds to a p-value of $2.8 \cdot 10^{-7}$. The p-value in the mass range around 578 GeV is reported in Figure 6.5 for different hypotheses on the natural width of the resonance.

The local excess observed is quantified of 2.7 σ. This results even if appreciable

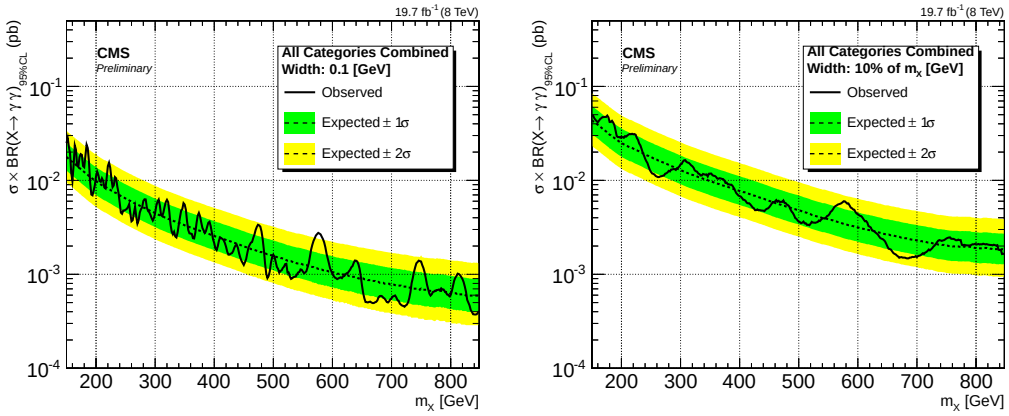


Figure 6.3. Expected and Observed exclusion limit at 95% CL on the cross section times the branching ratio of a new resonance decaying into two photons as a function of the resonance mass hypothesis combining the four classes of events. Narrow resonance hypothesis on the left and wide resonance hypothesis on the right.

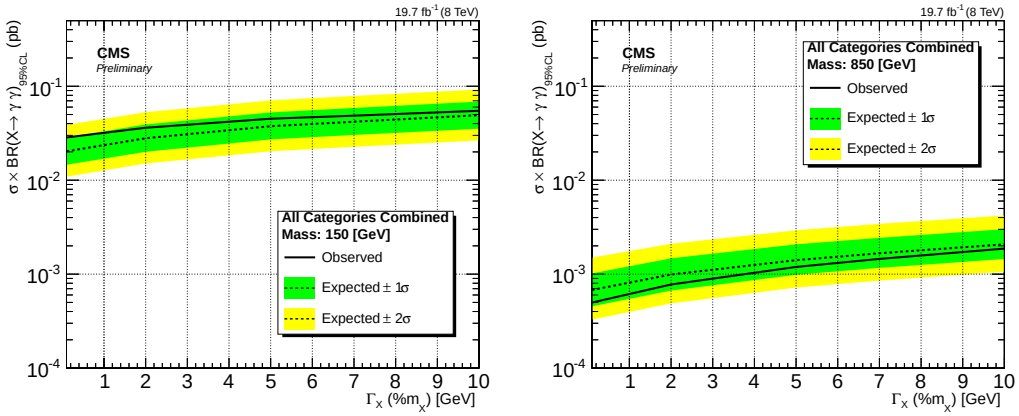


Figure 6.4. Expected and Observed exclusion limit at 95% CL on the cross section times the branching ratio of a new resonance decaying into two photons as a function of the resonance width hypothesis combining the four classes of events. Low mass resonance on the left and high mass resonance on the right.

is not relevant considering the "look elsewhere effect". The latter is introduced to take into account that such a local excess could in principle come out at any value of the mass in the analyzed range [128]. In this case a trial factor should be computed by which one need to "de-rate" the local p-value derived from the maximal value observed in the scan. In the case of this analysis this trial factor is fairly estimated to be greater than 100, reducing the significance of this excess to less than 2 σ.

As a summary of this analysis, results are presented in a two-dimensional plane as a function of the mass and the width of the hypothetical signal. Figure 6.6 shows the 95% CL exclusion limits on the production cross section times branching ratio dividing the events in the four categories. Expected limits are on the left

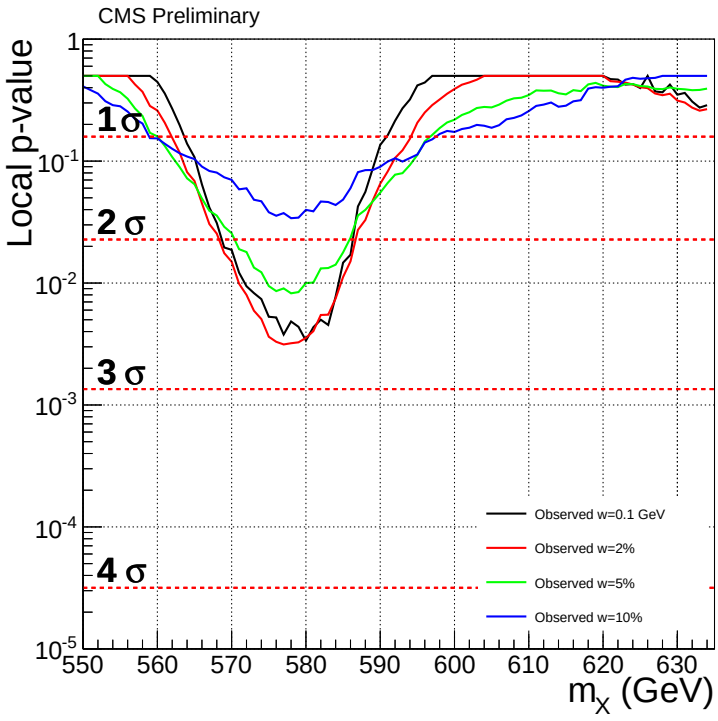


Figure 6.5. Expected p-value computed for a Signal Mass of 578 GeV (the minimum of the observed pval) and for a signal strength corresponding to the best fit at the same mass.

and observed limits are on the right. Contours are reported for different values of $\sigma \times BR(X \to \gamma\gamma)$. Figures 6.7 are the correspondent combined expected and observed limits in the 2D plane.

6.2 Interpretation in different models

Such a model-independent approach in this search for diphoton resonant processes over a continuum diphoton background allows us to interpret the results in different signal scenarios. So far results, including efficiencies and upper limits, have been shown for the hypothesis of a spin 0 gluon fusion produced resonance. In this section we discuss the interpretations of the results on two other narrow width possible models:

- A Vector Boson Fusion produced spin 0 resonance
- A Gluon Fusion produced spin 2 resonance

The figure of merit we use to compare the performance of the analysis on the different models are the efficiencies and the acceptances, including all the weights, smearings and scale factors. The trend of the efficiency as a function of the resonance mass for the three models is shown in Figure 6.8.

Given the results shown in Section 5.7.1 and these efficiencies, analysis results can be easily interpreted in any possible signal scenario. As an example, exclusion limits as a function of the mass are re-computed for the hypothesis of a narrow spin 2 resonance produced via gluon fusion, as shown in Figure 6.9.

6.2.1 Interpretation in 2HDM scenario

In this section the model independent limits obtained for an hypothetical heavy resonance which decays to two photons are interpreted in the context of the diphoton decay of the two heavy scalar Higgs, H and A, predicted in the 2HDM scenario [132]. In this SUSY scenario, the production cross sections for H and A, as well as the branching fractions for them to decay to two photons depend on two parameters: α and β. As explained in Section 2.1.1 the mixing angle between H and h (the lightest Higgs boson predicted by the theory) is given by α where as β, via $tan\beta$ gives the relative contribution of each higgs doublet to electroweak symmetry breaking. We use the SusHi [133] program to calculate the 2HDM cross sections and BRs are calculated using 2HDMC [134].

Exclusion regions in the plane $tan\beta$ vs $cos(\alpha-\beta)$ are shown only for the diphoton decay of the A scalar, since no region of the phase space can be excluded for the decay of the heavy H scalar.

Figures 6.10 show expected and observed limits for Heavy Higgs A of mass 200 and 300 GeV for Type I 2HDM along with the associated $\sigma \cdot BR$ contours for the respective model. The case where H and A are degenerated in mass is considered. These contour plots are similar to one from theory paper by Nathaniel Craig et al. [31]. In Figures 6.10 the regions below the curve is excluded. In case of a closed curve, the region enclosed by the curve is excluded.

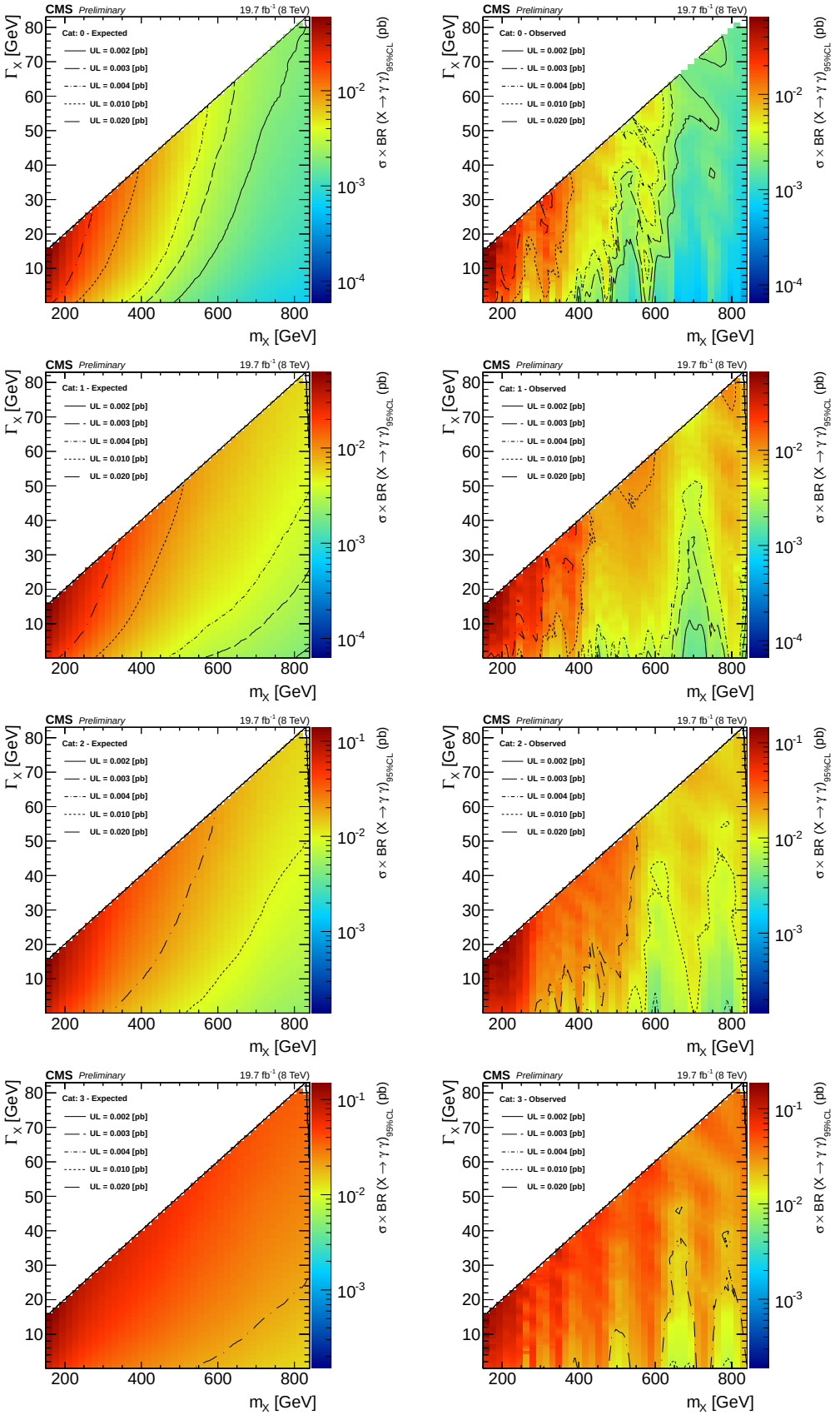


Figure 6.6. Exclusion limit at 95% CL on the cross section times the branching ratio of a new resonance decaying into two photons as a function of the resonance mass hypothesis and width in the four classes of events. Contours are reported for different values of $\sigma \times BR(X \to \gamma\gamma)$. Expected limits are on the left and observed limits are on the right.

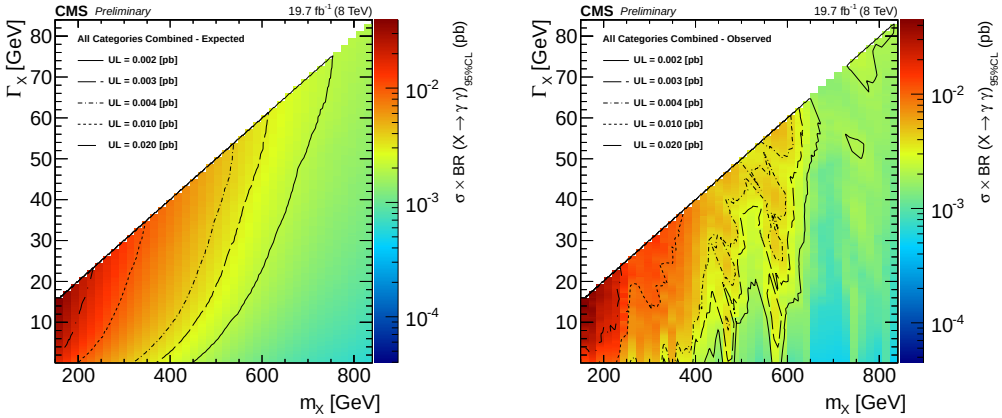


Figure 6.7. Exclusion limit at 95% CL on the cross section times the branching ratio of a new resonance decaying into two photons as a function of the resonance mass hypothesis and width combing the four classes of events. Contours are reported for different values of $\sigma \times BR(X \to \gamma\gamma)$. Expected limits are on the left and observed limits are on the right.

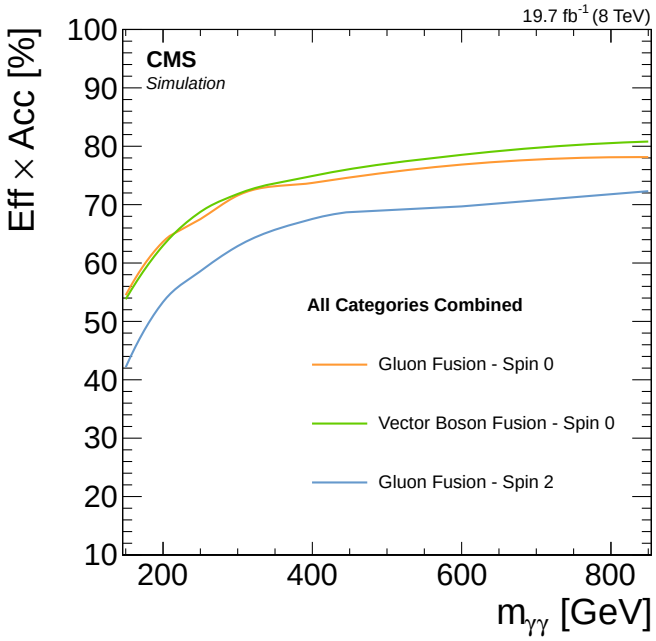


Figure 6.8. Selection efficiency times acceptance as a function of the resonance mass for different benchmark models

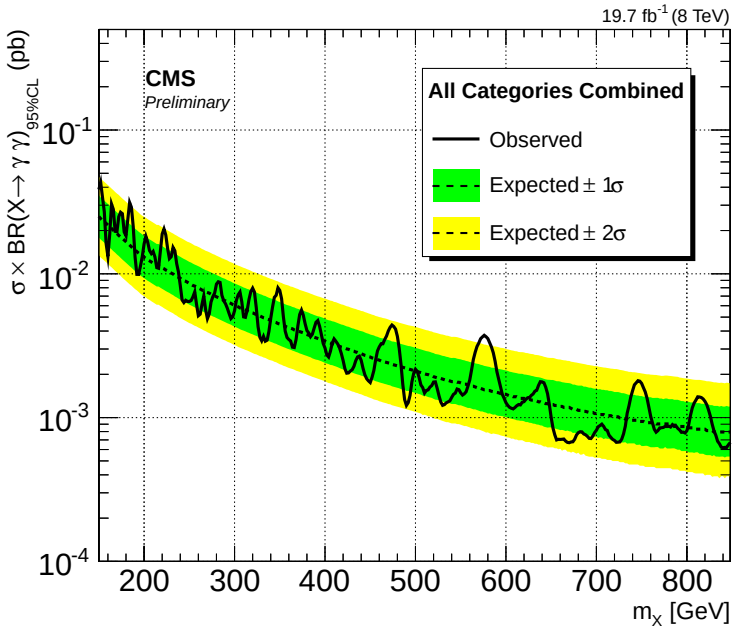


Figure 6.9. Exclusion limit at 95% CL on the cross section times the branching ratio of a new spin 2 resonance produced via gluon fusion and decaying into two photons as a function of the resonance mass hypothesis combining the four classes of events.

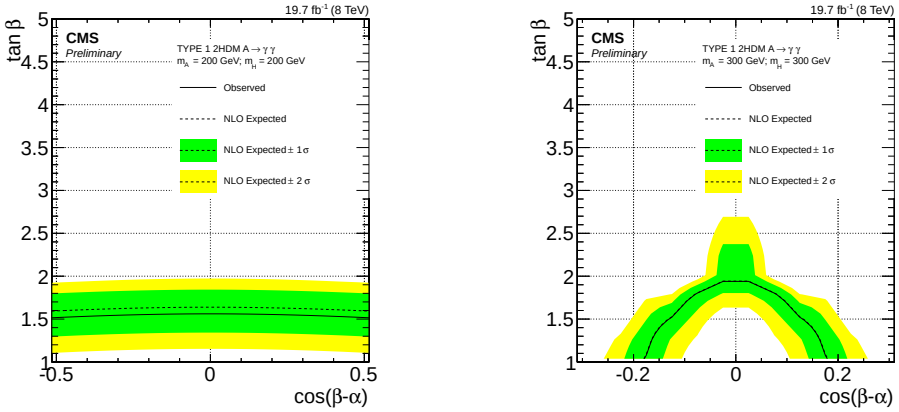


Figure 6.10. Observed and expected 95% CL upper limits for gluon fusion production of a heavy Higgs boson A of mass 200 (left) and 300 (right) GeV as a function of parameters $tan\beta$ and $cos(\alpha - \beta)$ of the Type I 2HDM assuming the H boson heavier than A (i.e. $m_H = 800$ GeV).

Conclusion

In this dissertation we have presented a model independent search for new heavy scalars decaying into two photons. The analyzed dataset comprises 19.7 fb^{-1} of proton-proton collisions delivered by the LHC at the center-of-mass energy of 8 TeV, and recorded by the CMS detector. All the events were evaluated by two levels of trigger (hardware and software), requiring two isolated photons to be reconstructed.

Such photons underwent a tight selection aimed at rejecting fake photons ascribable to mis-reconstructed jets, electrons with few tracker hits, and genuine photons coming from the decay of boosted mesons. The photon selection was based mainly on isolation variables and on the shape of the electromagnetic shower. A precise reconstruction of the photon energy mainly affected by the crystal intercalibration and by the crystal transparency loss due to irradiation was performed reaching a precision better than 1% in the barrel region. The clusterized photon energy was further corrected to account for the energy leaking outside the clusterization region, for the hardware gaps between the calorimeter modules, and for the average energy density in the event due to pile-up interactions

The vertex determination particularly challenging in events where the hard scattering collision produces only two neutral objects has been performed through a multivariate approach combining the transverse momentum of the diphoton system with the transverse momentum of all the other particles coming from a given vertex.

The search for new heavy states decaying to photons is performed by fitting analytical descriptions of the signal and background distributions to the measured diphoton invariant mass spectrum. The number of background events in the signal region was estimated from fitting the data with a sliding window technique according to the value of the mass and the natural width of the studied resonance. A detailed bias study has been performed to find out the analytical function which better estimates the background shape without introducing any bias in the measured signal yield. The reconstructed signal shape, mainly dominated by the resolution and reconstruction in ECAL, is obtained by convoluting a response function which takes into account the ECAL performance in the mass measurement with a Breit-Wigner describing theoretical lineshape for the resonance.

No assumptions have been made on the mass and on the natural width of the resonance, and the search is performed in the diphoton mass range between 150 GeV and 850 GeV. Heavy resonances are investigated with natural widths up to 10% of their mass. Since no excesses have been observed over the background expectation, upper limits have been set at 95% of CL on the product of the cross section times branching ratio. The final result of this analysis is a two-dimensional plot reporting the expected and observed upper-limits as a function of the mass and the natural

width of the heavy resonance. These results are the most recent to quote on a search for heavy particles in the diphoton final state and the model independent upper limits set extend over a considerably wider mass and width ranges than previous searches. We further interpret these limits in the context of 2HDMs, presenting exclusion contours in the $tan\beta$ versus $cos(\beta - \alpha)$ plane. This is the first search for heavy Higgs decaying to two photons carried out at the LHC.

Ringraziamenti

Ringraziare non e' mai una cosa semplice, direi pure un' usanza un po' dimenticata. Ma dopo essere sopravvissuti a tre anni di dottorato, non ci si puo' davvero fermare davanti all'ultima difficolta', quindi procedo. Quelli che adesso ringrazio sono quelle stesse persone, che come dicevo a pagina uno, condividono con me i "puntini" di questa vita. Questa tesi e' stata scritta con i vostri consigli e con le vostre critiche, con la vostra esperienza e con il vostro esempio, e ancora piu' importante, con la vostra presenza e il vostro sostegno. Percio' grazie.

Grazie prima di tutto ai miei compagni di vita romani che negli anni mi hanno insegnato a scoprire e ad amare incondizionatamente la citta' piu' bella del mondo. La devozione che i romani dimostrano per la loro citta' e' qualcosa di travolgente e contagioso che si puo' solo cercare di imparare col tempo se qualcuno e' disposto ad insegnartelo. Grazie dunque ad Andrea R., F.C., Matteo, Simone, Nick e Rob per avermi regalato un pezzetto della loro romanita', che invidio infinitamente. Un pensiero speciale poi va a Gu, Peter, Gizza e Giuliano che sono arrivati tardi nella mia vita rispetto agli altri, ma sono riusciti a farsi volere incredibilmente bene in pochi giorni. Un ringraziamento anche ai miei amici storici Gigi e Rick, che erano con me a studiare materia condensata sul tetto sotto il sole cocente di un giugno romano, o il 23 dicembre seduti sul pavimento del corridoio a fisica ad aspettare di fare l'orale di MQR. Quando saro' vecchia queste sono le cose che raccontero' ai mei nipoti, e vi pensero' con affetto. E grazie anche a Salvatore che in questi tre anni ogni tanto mi ha rapito dalla vita frenetica del dipartimento di fisica per portarmi in giro ad esplorare il mondo. Ai miei nipoti raccontero' anche questo. Una parola gentile infine anche per Guido, Michele e Peppe, trio indiscusso di simpatia e presenza constante nella mia vita dal lontano 2001. Loro stanno li' a ricordarmi che anche se diventiamo grandi certe cose non cambiano mai, per fortuna.

Ringrazio poi le donne della mia vita, Vale, Dudi, Claudia, Daria, Nau e Veronica. Le ringrazio perche' ciascuna di loro riesce a starmi vicina nonostante i chilometri, le montagne o gli oceani spesse volte ci separino. Loro sono la mia squadra, che festeggia con me le mie vittorie e che raccoglie i pezzi delle mie sconfitte. La fiducia che mi hanno sempre dimostrato non si puo' certo ripagare con quattro righe, ma spero che la vita mi dara' modo e tempo di restituire loro tutto quello che hanno dato a me.

Direi che poi c'e' da ringraziare i miei compagni di lavoro, quel controverso e curioso gruppo di persone che si firma CMS Roma. Ringrazio, in ordine di eta', Giuseppe, Vittorio e Giulia per avere condiviso con me lunghe e interminabili giornate nell'ufficio del 32 a litigare con CMSSW, e altrettanto interminabili serate a condividere pensieri e ansie per il futuro. Realizzare che esiste qualcuno, proprio

vicino a te, che ha le tue stesse paure, ti fa capire che "quella che al momento ci sembra incompletezza, a conti fatti si chiama solo gioventu' ". A Francesco M. spetta un ringraziamento speciale. E' forse l'unico di questa lista, che fitterebbe bene allo stesso tempo nelle categorie amici, colleghi e famiglia, e quello che ha condiviso con me piu' puntini che altri. Ho cominciato con lui questa avventura quando della fisica sapevamo a malapena spiegare F=ma, e nonostante tutto il tempo passato, mi ritrovo spesso a pensare "ah, bella questa analisi, magari l'anno prossimo io e Fra' potremmo farla insieme". Non so se questo significa qualcosa, ma a me fa sentire a casa. A Shervin, beh, non c'e' molto da dire. Semplicemente senza di lui non avrei mai concluso questa tesi. E' forse, tra tutti i membri del gruppo, il piu' strano e controverso, ma lo ringrazio per aver lasciato sempre una porta aperta per me, che fosse la porta del suo ufficio o la porta di casa. Specialmente nelle ultime settimane di dottorato e' stata davvero una presenza indispensabile, come collega ma ancora di piu' come amico, dunque grazie. Passando poi ai big del gruppo (come piace a loro essere chiamati) o ai piu' vecchi (come invece piace chiamarli a me) ringraziamo subito Daniele e Paolo, per i loro consigli e suggerimenti direi senza dubbio indispensabili per portare a termine questo lavoro, e tanti altri. Non credo che diventero' mai un fisico al loro livello, ma li ringrazio per tutto quello che mi hanno insegnato anche inconsapevolmente. Un grande grazie anche a Riccardo e Francesca che sono stati i primi ad accogliermi nel gruppo nel lontano 2010 e a Giovanni, che ha sopportato pazientemente per ben 5 lunghi anni prima che io riuscissi a dargli del "tu". Un ringraziamento speciale va poi a Emanuele e Chiara. Li ringrazio non tanto per l'aiuto che mi hanno dato a lavoro, o per i consigli o suggerimenti sulle scelte del post-doc che sono stati certamente preziosi. Li ringrazio piuttosto per l'esempio che mi danno ogni giorno come famiglia, e per avermi mostrato che nonostante le difficolta' che un lavoro affascinante e impegnativo come il nostro possa portare, costruire qualcosa e' ancora possibile, e che vale la pena provarci. Rimangono infine, in ordine di eta' non me ne vogliano, Marcella, indiscutibile "capa" ad honorem della nostra ciurma ed Egidio, che ringrazio soprattutto per avermi dato la possibilita' di insegnare. La fiducia che ha dimostrato nel mettermi nelle mani un'aula piena di studenti del terzo anno, e' senza dubbio una delle cose piu' belle che qualcuno avesse potuto fare per me.

Due parole poi le spenderei pure per ringraziare i miei colleghi e amici al CERN, Marco, AndreaCarlo e Andrea M. per tutte le volte che si sono dimostrati disponibili ad aiutarmi, ad orari improbabili del giorno e della notte. Ringrazio pure i vari convener che ho incontrato lungo il mio cammino, primi tra tutti Mauro, Pasquale e Federico, che ho imparato essere delle belle persone prima ancora che dei bravissimi fisici.

Restano mi sa' poche persone da ringraziare. I miei genitori e la mia famiglia prima di tutto. Loro li ringrazio perche' nonostante la strada che mi sono scelta di percorrere sia spesso impervia e rischiosa, riescono incredibilmente a seguirmi ovunque vado, e molte volte ad arrivare prima di me. Li ringrazio per esserci sempre e per mettere i puntini al posto mio quando vedono che in quel momento le forze per farlo mi mancano. Portano avanti con me un progetto da circa 27 anni con la convinzione che ad un certo punto tutti gli sforzi fatti verranno ripagati. Io posso solo assicurargli che faro' il modo che questo accada.

E per ultimo ringrazio il mio capo che in questi anni non e' stato solo un capo, ma

piuttosto una guida e dispensatore di consigli. Lo ringrazio per tutte le opportunita' che mi ha dato e per avermi sempre fatto intuire, secondo me anche inconsciamente, quale fosse la strada giusta da seguire. Ha avuto praticamente sempre ragione, ed e' forse per questo che la mia stima e considerazione di lui negli anni hanno potuto solo che aumentare. In una delle nostre ultime conversazioni gli ho detto che un giorno, in futuro, mi piacerebbe tornare a lavorare a Roma, e lui, al suo solito, ha sentenziato qualcosa del tipo "Beh se e' cosi' sei cretina", perche' la situazione in Italia e' disastrata e tra sei anni lo sara' ancora di piu'. Che dire.. spero che, una volta tanto, sia lui ad avere torto :)

Appendix A

Search for long-lived particles decaying to photons

In this appendix, a search for new long-lived particles decaying to photons, predicted in several scenarios of BSM physics is described. This analysis has been performed during the first year of my PhD. A paper describing the result of these studies is now public [8]. Results given by this analysis are the most stringent on long-lived neutralinos.

A.1 Introduction

New, heavy particles with long lifetimes are predicted in many models of physics beyond the SM, such as hidden valley scenarios [55] or supersymmetry with gauge-mediated supersymmetry breaking (GMSB) [135]. Under the assumption of R-parity conservation [136], strongly-interacting supersymmetric particles would be pair-produced at the Large Hadron Collider. The decay chain may include one or more quarks and gluons, as well as the lightest supersymmetric particle (LSP), which escapes detection, giving rise to a momentum imbalance in the transverse plane. A GMSB benchmark scenario, commonly described as Snowmass Points and Slopes 8 (SPS8) [137] is used as the reference in this search. In this scenario, the lightest neutralino ($\tilde{\chi}_1^0 \to \gamma\tilde{G}$) is the next-to-lightest supersymmetric particle, and can be long-lived. It decays to a photon (or a Z boson) and a gravitino ($\tilde{\chi}_1^0 \to \gamma\tilde{G}$), which is the LSP [138]. If $\tilde{\chi}_1^0$ consists predominantly of the bino, the superpartner of the $U(1)$ gauge field, its branching fraction to a photon and gravitino is expected to be large. If $\tilde{\chi}_1^0 \to \gamma\tilde{G}$ is wino-like, the superpartner of the $SU(2)$ gauge fields, its branching fraction to a photon and gravitino is reduced. Figure A.1 shows several diagrams of possible squark and gluino pair-production processes that result in a single-photon or diphoton final state.

The search criteria require only one identified photon in order to be sensitive to scenarios with a large branching fraction for the neutralino decay to a Z boson and a gravitino. For a long-lived neutralino, the photon from the $\tilde{\chi}_1^0 \to \gamma\tilde{G}$ decay is produced at the $\tilde{\chi}_1^0 \to \gamma\tilde{G}$ decay vertex, at some distance from the beam line, and reaches the detector at a later time than the prompt, relativistic particles produced at the interaction point. The search for $\tilde{\chi}_1^0 \to \gamma\tilde{G}$ is performed with a novel technique

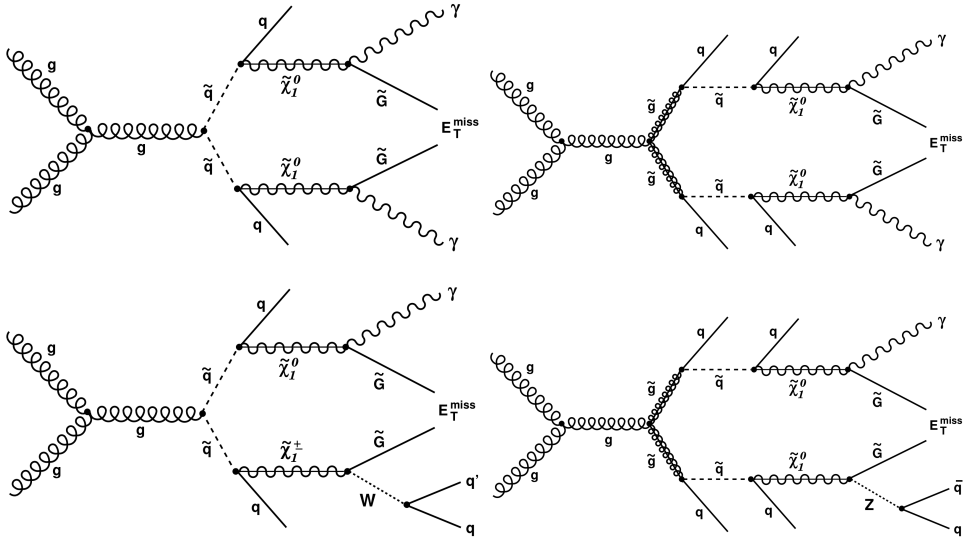


Figure A.1. Example diagrams for SUSY processes that result in a diphoton (top) and single-photon (bottom) final state through squark (left) and gluino (right) production at the LHC.

using the excellent time measurement with the electromagnetic calorimeter for the identification of the off-time photons..

Previous searches for long-lived neutralinos have been performed by the CMS Collaboration [139], using the impact parameter of converted photons relative to the beam collision point, and by the CDF [140] collaboration, using only the missing transverse energy in the event. Other searches with prompt photons, by the ATLAS [141] and D0 [142] collaborations, place lower limits on the mass of the $\tilde{\chi}_1^0 \to \gamma\tilde{G}$ at 280 GeV and 175 GeV, respectively, in the SPS8 scenario, assuming BR($\tilde{\chi}_1^0 \to \gamma\tilde{G}$)=100%.

A.2 Analysis Strategy

This analysis has been performed on the proton-proton collision data at a center-of-mass-energy of 7 TeV corresponding to an integrated luminosity of $4.9 fb^{-1}$.

Events are selected with a primary vertex with at least four associated tracks and a position less than 2 cm from the center of the CMS detector in the direction transverse to the beam and 24 cm in the direction along the beam. Events are also required to have at least three jets with $p_T > 35$ GeV and spatially separated from photons by at least $\Delta R = 0.5$. Jets are reconstructed from objects identified using the Particle-Flow (PF) algorithm [143] with anti-kt clustering [144] and a distance parameter of 0.5. In this analysis, the missing transverse energy ($\not\!\!E_T$) is defined as the magnitude of the vector sum of the transverse momentum of all particles identified in the PF algorithm in the event excluding muons.

Photon candidates are required to have $p_T \geq 100$ GeV and $|\eta| \leq 1.44$ and to be isolated in the HCAL, the ECAL, and the tracker. The photon candidates are reconstructed from clusters of energy in the ECAL as detailed in Chapter 4. The

identification criteria similar to those exploited in the high mass diphoton search, are described in detail in [8]. However, one additional distinctive feature is here used to discriminate between real and fake photons: the shape of the energy deposits they leave in the ECAL. This is detailed in Section A.2.2. The performance of the applied selection on different signal samples are described in Section A.2.3.

The excellent performance of the ECAL allows also for the identification of off-pointing and off-time photons thanks to the optimal resolution (< 1 ns for $E > 3$ GeV [97]) on time measurement. The definition of time of arrival of the photon in ECAL will be described in Section A.2.1. Figure A.2 shows a comparison between the distributions of ECAL Time for photons from SM events, and photons from long-living neutralino with different $\tilde{\chi}_1^0 \to \gamma\tilde{G}$ lifetimes. All distributions have been normalized to unity. The plot clearly demonstrates that GMSB distribution populates the region at large reconstructed time and that the measured time in ECAL is a very powerful variable in identifying out-of-time photons.

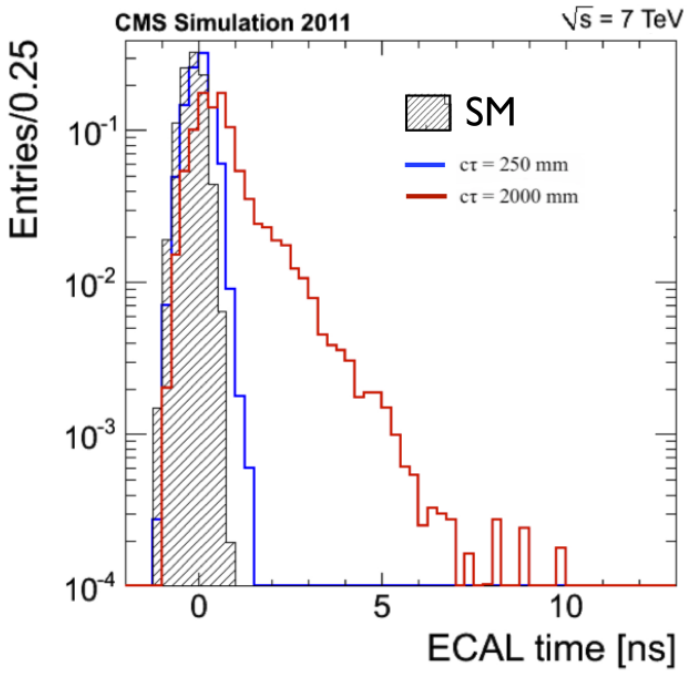


Figure A.2. ECAL time distribution for prompt and off-time photons. Two different GMSB scenario are shown with neutralino proper decay length equal to 250 mm and 2000 mm.

In this analysis signal is expected when both large $\not{E}_T$ and ECAL Time are measured therefore the time of arrival of the photon at the detector and the missing transverse energy are used to discriminate signal from background. The signal and background yields are finally determined with a binned maximum likelihood fit to the two-dimensional distribution in these variables.

Irreducible SM backgrounds, i.e. events with real missing transverse energy and high P_T photons in the final state, are negligible. The main backgrounds are instead reducible and arise from SM processes with misidentified photons and/or mismeasured missing transverse energy like QCD events and γ+jets. The contribution of these two main backgrounds are estimated to make up more than 99% of the total data sample and are estimated using data-driven control samples as detailed in Section A.3. The remaining 1% are other (non-QCD) processes taken into account since they can play a role in the tails of the $\not\!\!E_T$ distribution where signal is expected. Simulated events are used to estimate the contribution of these processes.

A.2.1 ECAL Time measurement

The time of impact, T_{raw}, for the photon on the surface of the ECAL is the weighted time of impact measured in the crystals within the cluster associated with a photon candidate. An event-by-event correction (T_{prompt}) is applied to T_{raw} to account for possible biases due to the jitter in the trigger system, and to the imperfect knowledge of the time of the interaction within the bunch crossing. This correction is computed using the time of impact of all crystals in the event, excluding those belonging to the two most energetic photon candidates, which are typically due to prompt jets, low-energy prompt photons, and photons from π^0 and η decays. The new calibrated ECAL time is defined as $T_{calib} = T_{\text{raw}} - T_{\text{prompt}}$. With this definition, a particle produced at the interaction point has a time of arrival of zero, whereas a delayed photon has a non-zero T_{calib}. The distributions in data for T_{raw} and T_{calib}, after the nominal selection, are shown in Figure A.3. The width of the main, Gaussian, component of T_{calib} is slightly smaller than that of T_{raw}, while there is some increase in the tails. For the dominant background processes, the tails are taken into account by using control samples in data, as described in Section A.3.

In the determination of the yield, the distribution of T_{calib} in simulated signal events is used as a template for the signal contribution. This distribution is narrower in simulation than in the data, because the uncertainties in the time inter-calibration constants are not emulated. A convolution with a Gaussian, whose parameters vary as a function of the photon energy, is performed to reproduce the T_{calib} resolution observed in data [145].

A.2.2 Photon Cluster Shape

Prompt photons have a roughly circular projected energy deposit on the ECAL surface, while the energy deposits from jets typically have a larger width along the η direction. Non-prompt photons are expected to have an elliptical shape along an arbitrary direction, as illustrated in Figure A.4, therefore the width of the energy deposit along the η direction is not optimal for the discrimination of jets. In this search, the shape of the energy deposit is characterized by the minor axis (S_{Minor}) of its projection on the internal ECAL surface. The axis S_{Minor} is computed using the geometrical properties of the distribution of the energy deposit, and is defined as

$$S_{\text{Minor}} = \frac{(S_{\phi\phi} + S_{\eta\eta}) - \sqrt{(S_{\phi\phi} - S_{\eta\eta})^2 + 4S_{\phi\eta}^2}}{2}, \tag{A.1}$$

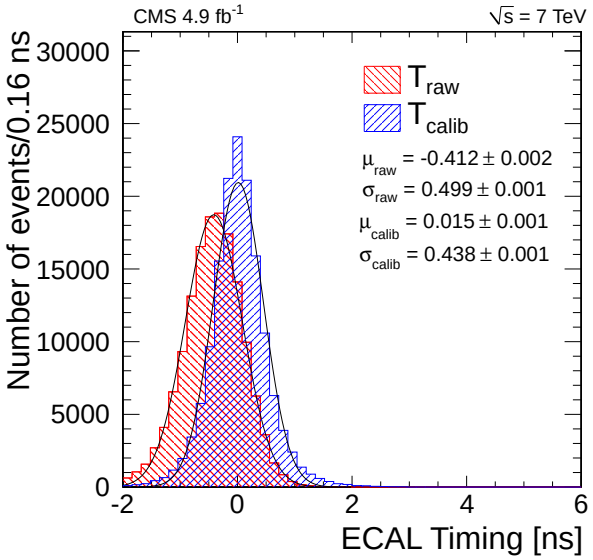


Figure A.3. The ECAL timing distribution for data, before and after calibration, overlaid with the results of the Gaussian fits.

where $S_{\phi\phi}$, $S_{\eta\eta}$, and $S_{\phi\eta}$ are the second moments of the spatial distribution of the energy deposit in the ECAL in η-ϕ coordinates. A large fraction of QCD multijet events can be rejected by applying requirements on S_{Minor} as illustrated in Figure A.5, where the normalized distributions of S_{Minor} for simulated signal and QCD multijet background events are shown.

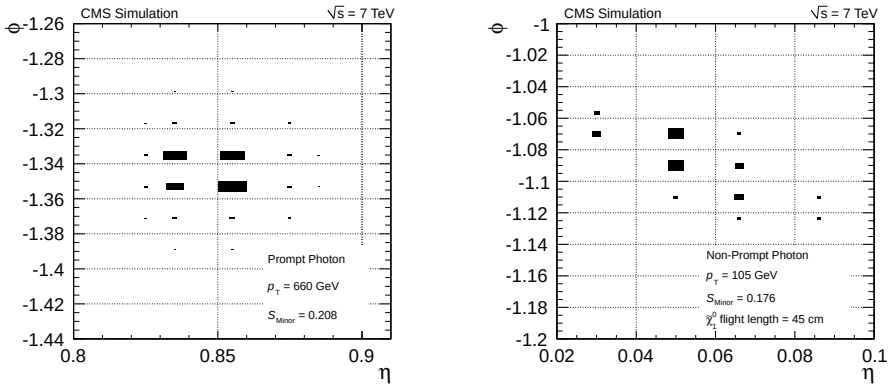


Figure A.4. The distribution of energy deposition in the ECAL crystals for a prompt (left) and a non-prompt (right) photon. Each rectangle represents an ECAL crystal and has a size that is proportional to the energy deposited in that crystal. The non-prompt illustration is for a $\tilde{\chi}_1^0 \to \gamma\tilde{G}$ flight length of 45 cm.

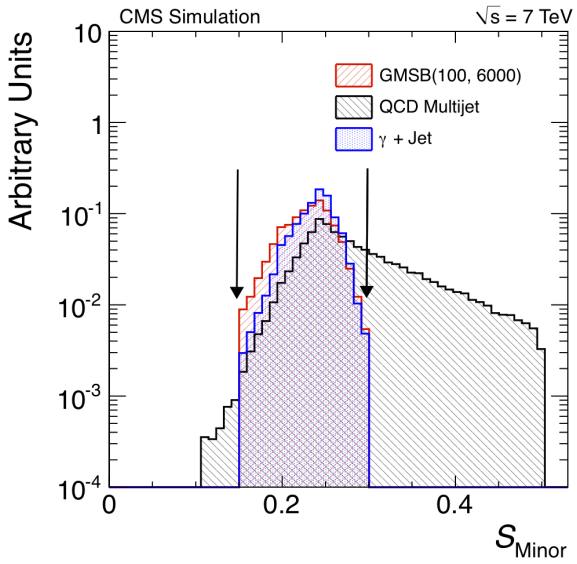


Figure A.5. Normalized distribution of S_{Minor} for simulated signal, γ + jets, and QCD multijet events. The arrows indicate the S_{Minor} selection interval.

A.2.3 Selection Efficiencies on Signal Samples

The signal efficiencies for selecting one photon and at least three jets are summarized in Table A.1 for proper decay lengths between 1 mm and 6000 mm and for Λ between 100TeV and 180TeV. The efficiency drops by a factor of two between $c\tau = 1$ mm and 6000 mm, since, with increasing decay time, the probability of the $\tilde{\chi}_1^0$ to decay outside the detector is enhanced.

Λ (TeVns)	$M_{\tilde{\chi}_1^0}$ (GeVns)	$c\tau = 1$ mm	$c\tau = 250$ mm	$c\tau = 2000$ mm	$c\tau = 6000$ mm
100	140	18.7 ± 0.3	18.4 ± 0.2	8.4 ± 0.2	3.3 ± 0.1
120	170	24.9 ± 0.4	24.6 ± 0.3	15.1 ± 0.4	6.6 ± 0.1
140	200	30.4 ± 0.3	31.3 ± 0.3	22.2 ± 0.4	11.4 ± 0.3
160	230	35.5 ± 0.3	36.1 ± 0.6	29.4 ± 0.4	17.0 ± 0.4
180	260	40.1 ± 0.7	38.0 ± 0.5	36.0 ± 0.5	22.2 ± 0.4

Table A.1. Selection efficiency in percent. The reported uncertainties include the contributions of systematic effects, for various signal samples.

A.3 Background estimation

The primary sources of background in the analysis are QCD multijet events and γ + jets events, which together make up 99% of the sample. Improper reconstruction of jets can give rise to fake missing transverse energy, while photons produced in the decays of hadrons (mostly energetic π_0 and η) can sometimes pass the isolation criteria.

A large fraction of γ + jets events, characterized by a smaller jet multiplicity compared to signal, are rejected by requiring at least three jets in the event. The residual contribution of these backgrounds is estimated from the data.

In addition, there are other (non-QCD) processes with genuine $\not\!\!E_T$, largely comprised of $W/Z + \gamma +$ jets and $t\bar{t}$ events, where the W boson decays into a lepton and a neutrino. There is also a small contribution from Drell-Yan processes. These events make up less than 1% of the total sample but are taken into account since they can play a role in the tails of the $\not\!\!E_T$ distribution where signal is expected. Simulated events are used to estimate the contribution of these processes.

Finally, additional backgrounds from events not originating from proton-proton collisions, including cosmic rays and beam-halo muons, are also expected. The contribution of these events is reduced to negligible levels by requiring T_{calib} of the most energetic photon candidate to be greater than -2 ns, and the event to have an identified primary vertex and at least three jets.

Because of the difficulty of accurately predicting cross sections and jet multiplicities for multijet and $\gamma +$ jets processes, their contribution is estimated with methods based on the data. The QCD multijet control sample is obtained by selecting events with at least three jets and a photon candidate passing a less stringent identification requirement but failing the nominal photon selection criteria. The $\gamma +$ jets control sample consists of events with one photon which satisfies the nominal selection. Events with the angle in the transverse plane between the highest-p_T jet (leading jet) and the photon smaller than $2/3\,\pi$ are rejected. The ratio of the transverse momenta of the leading jet to that of the photon is required to be between 0.6 and 1.4, while for the subleading jet the ratio is required to be less than 0.2. The contribution of non-QCD and signal events to these two control samples is estimated to be, respectively, 1% and less than 0.01%.

To estimate the number of background and signal events in data, a maximum likelihood fit is performed to the two-dimensional distribution of $\not\!\!E_T$ and T_{calib}. The correlation coefficient between $\not\!\!E_T$ and T_{calib} is 0.05 for events with $\not\!\!E_T > 100$ GeV and $T_{\text{calib}} > 0.5$ ns, and 0.001 when all events are considered. Binned shape templates are derived from simulated events for signal and non-QCD backgrounds. Templates for QCD multijet and $\gamma +$ jets are derived from data control samples as described earlier. The relative normalization of the QCD multijet and $\gamma +$ jets components is fixed to 67% and 33%, respectively, based on studies with simulated events. The normalization of the non-QCD templates are fixed in the fit according to the measured cross sections (statistical uncertainties in the cross sections are less than 3%) and the integrated luminosity of the data sample. Studies have been performed with pseudo-experiments to confirm the stability of the fit and to verify that the fit results are unbiased. The measured signal and background yields in data, obtained with the likelihood fit, are summarized in Table A.2. The one-dimensional projections of $\not\!\!E_T$ and T_{calib} for the data and expected backgrounds, as determined from the fit, are illustrated in Figure A.6. No excess of events is observed beyond the SM backgrounds and the fitted signal yield is compatible with zero. It should be noted that the discriminating power of individual variables is not apparent in these projections because the largest sensitivity to signal is in the region with both large $\not\!\!E_T$ and large T_{calib}. The improved background discrimination is visible in Figure A.7 where the one-dimensional projection of $\not\!\!E_T$ for events with $T_{\text{calib}} > 0.5$ ns is illustrated.

	Events
GMSB (100, 250)	6 ± 8
GMSB (100, 2000)	4 ± 4
QCD multijet and $\gamma +$ jets	$80\,900 \pm 300$
$t\bar{t}$ + jets (fixed)	73
$W \to e\nu +$ jets (fixed)	116
Drell–Yan + jets (fixed)	67
$W/Z +$ jets $+ \gamma$ (fixed)	215
Total background	$81\,400 \pm 300$
Data	81 382

Table A.2. The measured signal and background yields determined with the maximum likelihood fit to the data. The relative composition of QCD multijet and $\gamma +$ jets backgrounds have been normalized to 67% and 33% with respect to each other. The expected signal yields are 211 events for the GMSB(100,250) benchmark point and 96 for GMSB(100,2000). The GMSB(100,250) benchmark point corresponds to $\Lambda = 100$TeV, $c\tau = 250$ mm and the GMSB(100,2000) benchmark point corresponds to $\Lambda = 100$TeV, $c\tau = 2000$ mm. The reported uncertainties are statistical only and are determined in the fit.

A.4 Systematic uncertainties

Several sources of systematic uncertainty have been considered and their contributions are summarized in Table A.3. The largest single contribution to the systematic uncertainties derives from the uncertainty in the modeling of the background shape. A bin-by-bin variation of the background shape template according to the Poisson uncertainty is used to determine the contribution of each type of background. An additional uncertainty is assessed for the QCD multijet and $\gamma +$ jets processes using simulated events, by comparing the shapes of $\not\!\!E_T$ and T_{calib} for the control sample and for a sample obtained with the nominal selection criteria. The difference observed in simulation is used to re-weight the shapes obtained in data control samples. The dominant contribution is due to the difference in the $\not\!\!E_T$ distributions. The small tails in the distribution of T_{calib} are accounted for by using data control samples to derive the templates, rather than relying on simulation. The uncertainty in the relative fraction of QCD multijet and $\gamma +$ jets events is estimated to be 33%. The main contribution to this uncertainty is due to the next-to-leading correction for the $\gamma +$ jets cross section. Additional contributions are included to take into account the the observed difference between the number of events in the $\gamma +$ jets control sample in data and the expected number of events according to PYTHIA (10%), and to the QCD multijet events misidentified as $\gamma +$ jets events (10%).

The main contributions to the uncertainty in the signal shape modeling derive from the uncertainty in the $\not\!\!E_T$ resolution and the determination of T_{calib}. The contribution of the $\not\!\!E_T$ resolution uncertainty is estimated by smearing the $\not\!\!E_T$ distribution of simulated signal events. A systematic uncertainty of 0.1 ns is assigned to the measurement of the time of impact T_{calib}. This uncertainty is determined using a sample of $\gamma +$ jets events by measuring the difference between the average T_{calib} values in data and simulation, as a function of the photon p_T.

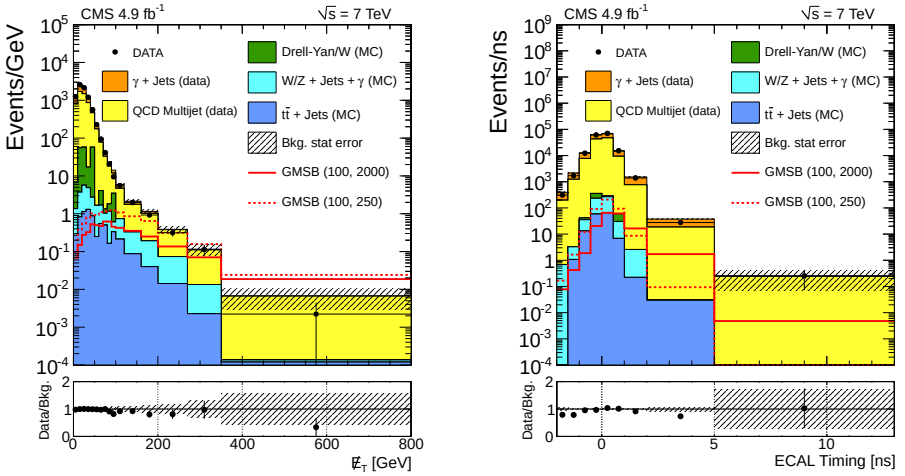


Figure A.6. The one-dimensional projection for $\not{E}_T$ (left) and for ECAL timing (right), after all selection requirements. The multijet and γ + jets backgrounds are normalized to the yields from the fit. The rest of the backgrounds are fixed according to the integrated luminosity of the data. The GMSB(100,2000) benchmark point corresponds to $\Lambda = 100$TeV, $c\tau = 2000$ mm and the GMSB(100,250) benchmark point corresponds to $\Lambda = 100$TeV, $c\tau = 250$ mm.

The uncertainty in the luminosity determination is 2.2% [146]. The remaining sources of systematic uncertainty affecting the signal acceptance are the following. The calorimeter response to different types of particles is not perfectly linear and hence corrections are made to properly map the measured jet energy deposition. The uncertainty on this correction is referred to as the uncertainty on the jet energy scale and varies as a function of position and transverse momentum of the jet. Similarly, the uncertainty on the photon energy scale in the barrel is estimated to be 1.0%, based on the final-state radiation measurement with Z bosons [147]. Following the recommendations of the PDF4LHC group [148], PDF and the strong coupling constant (α_s) variations of the MSTW2008 [149], CTEQ6.6 [150] and NNPDF2.0 [151] PDF sets are taken into account and their impact on the signal acceptance is estimated.

A.5 Results

The observed event yield in data is consistent with the SM background prediction, and upper limits are obtained on the production cross section of a long-lived neutralino in the context of the GMSB model, assuming BR($\tilde{\chi}_1^0 \to \gamma\tilde{G}$)=100%. Exclusion limits are computed with a modified frequentist CL_s method [152–154], using the asymptotic approximation for the test statistic as described in Ref. [155]. The background normalization and the corresponding uncertainty are taken from the fit to the data. The uncertainties in the shapes are taken into account by vertical interpolation of the templates. The shapes are interpolated quadratically for shifts

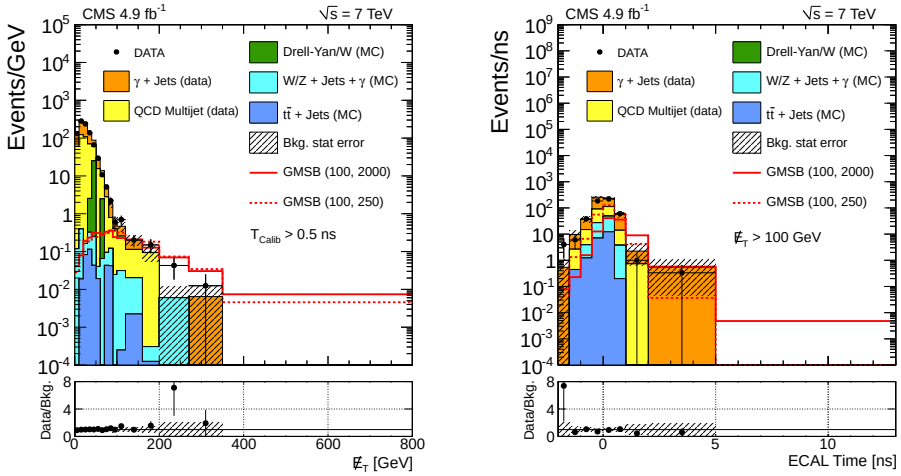


Figure A.7. The one-dimensional projection after all selection requirements for $\not{E}_T$ for events with $T_{\text{calib}} > 0.5ns$ (left) and for ECAL timing (right) for events with $\not{E}_T > 100GeV$. The multijet and γ + jets backgrounds are normalized to the yields from the fit. The rest of the backgrounds are fixed according to the integrated luminosity of the data. The GMSB(100,2000) benchmark point corresponds to $\Lambda = 100$ TeV, $c\tau = 2000$ mm and the GMSB(100,250) benchmark point corresponds to $\Lambda = 100$TeV, $c\tau = 250$ mm.

Source	Uncertainty (%)
Background	
Shape	10
Normalization	0.3
Multijet/γ + jets fraction	0.8
Signal shape	
$\not{E}_T$ resolution	0.2–2
ECAL timing uncertainty	1–5
Signal acceptance × efficiency	
Photon energy scale	0.5–3
Jet energy scale	0.02–0.05
Jet energy resolution	0.01–2
PDF uncertainties	0.1–2

Table A.3. Summary of the systematic uncertainties in the background and signal shapes, as well as in the signal acceptance × efficiency. The signal uncertainties are evaluated individually for every signal point, although only the maximum and minimum values associated with each source are quoted.

below one standard deviation and linearly beyond. Log-normal multiplicative corrections are used for the normalization, the signal acceptance, and the integrated luminosity. Figure A.8 shows the observed and expected 95% CL upper limits on the cross section for GMSB production in terms of$\tilde{\chi}_1^0 \to \gamma\tilde{G}$ mass (left), and proper decay

length (right). The signal cross section is computed at leading order precision and the theoretical uncertainty is evaluated by using the PDF4LHC recommendation for the PDF uncertainty. The one-dimensional limits are combined to provide exclusion limits in the mass and proper decay length plane of the long-lived$\tilde{\chi}_1^0 \to \gamma\tilde{G}$ in Figure A.9.

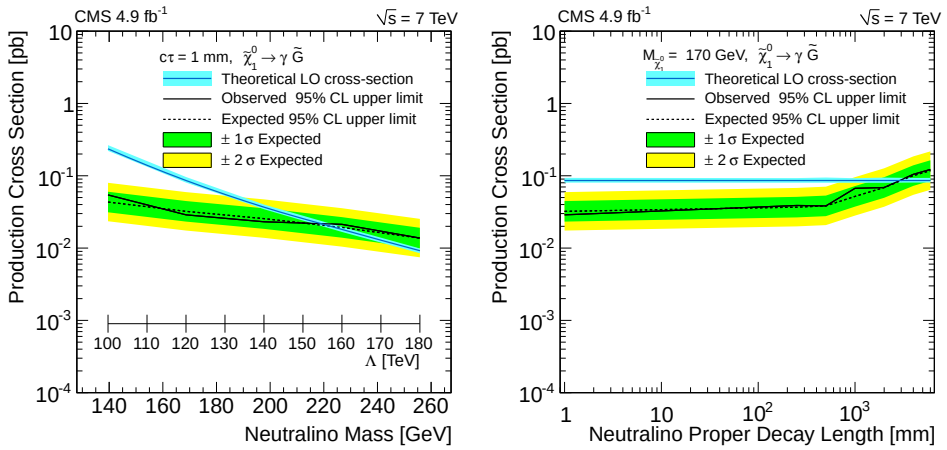


Figure A.8. Upper limits at the 95% CL on the cross section as a function of the $\tilde{\chi}_1^0 \to \gamma\tilde{G}$ mass for $c\tau = 1$ mm (left), and for the$\tilde{\chi}_1^0 \to \gamma\tilde{G}$ proper decay length for $M_{\tilde{\chi}_1^0} = 170$ GeV (right) in the SPS8 model of GMSB supersymmetry.

A.6 Summary

The CMS experiment has performed a search for long-lived particles produced in association with jets using LHC proton-proton collision data at a center-of-mass energy of 7TeV corresponding to an integrated luminosity of $4.9fb^{-1}$. A GMSB scenario with a long-lived neutralino decaying to a photon and a gravitino is used as the reference. The missing transverse energy and the timing information from the ECAL are used to search for an excess of events over the expected SM background prediction. A fit to the two-dimensional distribution in these variables yields no significant excess of events beyond the SM contributions, and upper limits at 95% CL are obtained on the GMSB production cross section in the SPS8 model of GMSB supersymmetry. In this scheme, we obtain an exclusion region as a function of both the neutralino mass and its proper decay length, assuming BR($\tilde{\chi}_1^0 \to \gamma\tilde{G}$)=100%. The mass of the lightest neutralino is then restricted to values $m(\tilde{\chi}_1^0) > 220$ GeV (for neutralino proper decay length $c\tau < 500$ mm) at 95% CL, and the neutralino decay length $c\tau$ must be greater than 6000 mm (for $m(\tilde{\chi}_1^0) < 150$ GeV). These limits are the most stringent for long-lived neutralinos.

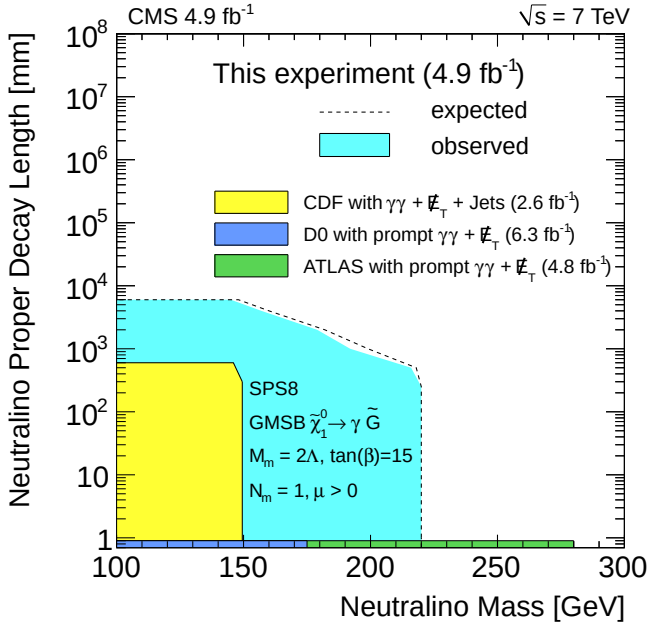


Figure A.9. The observed exclusion region for the mass and proper decay length of the long-lived $\tilde{\chi}_1^0$ in the SPS8 model of GMSB supersymmetry.

Bibliography

[1] Steven Weinberg. The Making of the Standard Model. *Eur.Phys.J.*, C34:5–13, 2004.

[2] S. Chatrchyan et al. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):30 – 61, 2012.

[3] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys.Lett.*, B716:1–29, 2012.

[4] B. Burch and R. Cowsik. Properties of Galactic Dark Matter: Constraints from Astronomical Observations. *The Astrophysical Journal*, 779:35, December 2013.

[5] The CMS Collaboration. Cms physics technical design report, volume ii: Physics performance. *Journal of Physics G: Nuclear and Particle Physics*, 34(6):995, 2007.

[6] A. Barroso, P.M. Ferreira, Rui Santos, Marc Sher, and Joao P. Silva. 2HDM at the LHC - the story so far. 2013.

[7] T.S. Berry. Extra Dimension Searches at Accelerators . Technical report, SLAC, 33rd SLAC Summer Institute on Particle Physics (SSI 2005), 25 July - 5 August 2005.

[8] Serguei Chatrchyan et al. Search for long-lived particles decaying to photons and missing energy in proton-proton collisions at $\sqrt{s} = 7$ TeV. *Phys.Lett.*, B722:273–294, 2013.

[9] M. Herrero. The Standard Model. *NATO Sci.Ser.C*, 534:1–59, 1999.

[10] J. Beringer et al. Review of Particle Physics (RPP). *Phys.Rev.*, D86:010001, 2012.

[11] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, Aug 1964.

[12] Peter W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[13] Peter W. Higgs. Spontaneous symmetry breakdown without massless bosons. *Phys. Rev.*, 145:1156–1163, May 1966.

[14] Precise determination of the mass of the Higgs boson and studies of the compatibility of its couplings with the Standard Model. Technical Report CMS-PAS-HIG-14-009, CERN, Geneva, 2014.

[15] Serguei Chatrchyan et al. Study of the Mass and Spin-Parity of the Higgs Boson Candidate Via Its Decays to Z Boson Pairs. *Phys.Rev.Lett.*, 110:081803, 2013.

[16] LHC Higgs Cross Section Working Group, S. Dittmaier, C. Mariotti, G. Passarino, and R. Tanaka (Eds.). Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables. *CERN-2011-002*, CERN, Geneva, 2011.

[17] Updated measurements of the Higgs boson at 125 GeV in the two photon decay channel. Technical Report CMS-PAS-HIG-13-001, CERN, Geneva, 2013.

[18] F. Zwicky. On the masses of nebulae and of clusters of nebulae. *Journal of Astrophysics*, 86(217), 1937.

[19] J. Dunkley et al. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Likelihoods and Parameters from the WMAP data. *Astrophys.J.Suppl.*, 180:306–329, 2009.

[20] Y. Fukuda et al. Evidence for oscillation of atmospheric neutrinos. *Phys.Rev.Lett.*, 81:1562–1567, 1998.

[21] R. Davis. A review of the homestake solar neutrino experiment. *Prog. Part. Nucl. Phys.*, 32:13–32, 1994.

[22] C. et al. Athanassopoulos. Candidate events in a search for $\overline{\nu}_\mu \to \overline{\nu}_e$ oscillations. *Phys. Rev. Lett.*, 75:2650–2653, Oct 1995.

[23] Ali H. Chamseddine, Alain Connes, and Matilde Marcolli. Gravity and the Standard Model with neutrino mixing. *Adv.Theor.Math.Phys.*, 11:991–1089, 2007.

[24] Stuart Raby. Grand Unified Theories. 2006.

[25] Shailesh Chandrasekharan. Fermion Bag Approach to Fermion Sign Problems. *Eur.Phys.J.*, A49:90, 2013.

[26] O.B. Abdinov, F.T. Khalil-zade, and S.S. Rzaeva. Electric Charge Quantization in Standard Model. 2008.

[27] C. GIUNTI, C. W. KIM, and U. W. LEE. Running coupling constants and grand unification models. *Modern Physics Letters A*, 06(19):1745–1755, 1991.

[28] Michael E. Peskin. Beyond the standard model. 1997.

[29] Gerard 't Hooft. Naturalness, chiral symmetry, and spontaneous chiral symmetry breaking. *NATO Sci.Ser.B*, 59:135, 1980.

[30] G. Belanger, F. Boudjema, A. Cottrant, R.M. Godbole, and A. Semenov. The MSSM invisible Higgs in the light of dark matter and g-2. *Phys.Lett.*, B519:93–102, 2001.

[31] Nathaniel Craig, Jamison Galloway, and Scott Thomas. Searching for Signs of the Second Higgs Doublet. 2013.

[32] Kei Yagyu. Studies on Extended Higgs Sectors as a Probe of New Physics Beyond the Standard Model. 2012.

[33] Gian F. Giudice, Riccardo Rattazzi, and James D. Wells. Graviscalars from higher dimensional metrics and curvature Higgs mixing. *Nucl.Phys.*, B595:250–276, 2001.

[34] Searches for invisible Higgs bosons: Preliminary combined results using LEP data collected at energies up to 209-GeV. 2001.

[35] Georges Aad et al. Search for Invisible Decays of a Higgs Boson Produced in Association with a Z Boson in ATLAS. *Phys.Rev.Lett.*, 112:201802, 2014.

[36] Georges Aad et al. Search for dark matter in events with a hadronically decaying W or Z boson and missing transverse momentum in *pp* collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *Phys.Rev.Lett.*, 112(4):041802, 2014.

[37] Serguei Chatrchyan et al. Search for invisible decays of Higgs bosons in the vector boson fusion and associated ZH production modes. *Eur.Phys.J.*, C74:2980, 2014.

[38] Abdelhak Djouadi, Oleg Lebedev, Yann Mambrini, and Jeremie Quevillon. Implications of LHC searches for Higgs–portal dark matter. *Phys.Lett.*, B709:65–69, 2012.

[39] Brian Patt and Frank Wilczek. Higgs-field portal into hidden sectors. 2006.

[40] Geraldine Servant and Sean Tulin. Higgsogenesis. *Phys. Rev. Lett.*, 111(arXiv:1304.3464. CERN-PH-TH-2013-068):151601, Apr 2013.

[41] Georges Aad et al. Search for the Standard Model Higgs boson decay to $\mu^+\mu^-$ with the ATLAS detector. *Phys.Lett.*, B738:68–86, 2014.

[42] Georges Aad et al. Search for Higgs boson decays to a photon and a Z boson in pp collisions at $\sqrt{s}$=7 and 8 TeV with the ATLAS detector. *Phys.Lett.*, B732:8–27, 2014.

[43] Search for a Higgs boson decaying into $\gamma * \gamma$ to $\mu\mu\gamma$ with dilepton mass below 20 GeV in pp collisions at sqrt-s = 8 TeV. Technical Report CMS-PAS-HIG-14-003, CERN, Geneva, 2014.

[44] Yi Sun, Hao-Ran Chang, and Dao-Neng Gao. Higgs decays to gamma l+ l- in the Standard Model. *JHEP*, 1305:061, 2013.

[45] Roni Harnik, Joachim Kopp, and Jure Zupan. Flavor Violating Higgs Decays. *JHEP*, 1303:026, 2013.

[46] James D. Bjorken and Steven Weinberg. Mechanism for nonconservation of muon number. *Phys. Rev. Lett.*, 38:622–625, Mar 1977.

[47] Aleksandr Azatov, Manuel Toharia, and Lijun Zhu. Higgs Mediated FCNC's in Warped Extra Dimensions. *Phys.Rev.*, D80:035016, 2009.

[48] Hajime Ishimori, Tatsuo Kobayashi, Hiroshi Ohki, Yusuke Shimizu, Hiroshi Okada, et al. Non-Abelian Discrete Symmetries in Particle Physics. *Prog.Theor.Phys.Suppl.*, 183:1–163, 2010.

[49] Gilad Perez and Lisa Randall. Natural Neutrino Masses and Mixings from Warped Geometry. *JHEP*, 0901:077, 2009.

[50] Gianluca Blankenburg, John Ellis, and Gino Isidori. Flavour-Changing Decays of a 125 GeV Higgs-like Particle. *Phys.Lett.*, B712:386–390, 2012.

[51] Search for Lepton Flavour Violating Decays of the Higgs Boson. Technical Report CMS-PAS-HIG-14-005, CERN, Geneva, 2014.

[52] Search for pair produced long-lived neutral particles decaying in the ATLAS hadronic calorimeter in pp collisions at $\sqrt{s} = 8$ TeV. Technical Report ATLAS-CONF-2014-041, CERN, Geneva, Jul 2014.

[53] Gregory Moreau. Phenomenological study of the interactions violating the R parity symmetry in the supersymmetric theories. 2000.

[54] David Smith and Neal Weiner. Inelastic dark matter. *Phys. Rev. D*, 64:043502, Jul 2001.

[55] Matthew J. Strassler and Kathryn M. Zurek. Echoes of a hidden valley at hadron colliders. *Phys.Lett.*, B651:374–379, 2007.

[56] Tao Han, Heather E. Logan, and Lian-Tao Wang. Smoking-gun signatures of little Higgs models. *JHEP*, 0601:099, 2006.

[57] Yutaka Hosotani. Gauge-Higgs Unification Approach. *AIP Conf.Proc.*, 1467:208–213, 2012.

[58] P. Fayet. Supersymmetry and weak, electromagnetic and strong interactions. *Physics Letters B*, 64(2):159 – 162, 1976.

[59] G.C. Branco, P.M. Ferreira, L. Lavoura, M.N. Rebelo, Marc Sher, et al. Theory and phenomenology of two-Higgs-doublet models. *Phys.Rept.*, 516:1–102, 2012.

[60] Howard E. Haber and Deva O'Neil. Basis-independent methods for the two-Higgs-doublet model. II. The Significance of tan beta. *Phys.Rev.*, D74:015018, 2006.

[61] Vardan Khachatryan et al. Search for neutral MSSM Higgs bosons decaying to a pair of tau leptons in pp collisions. 2014.

[62] Georges Aad et al. Search for neutral Higgs bosons of the minimal supersymmetric standard model in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. 2014.

[63] Doug Benjamin et al. Combined CDF and D0 Upper Limits on MSSM Higgs Boson Production in tau-tau Final States with up to 2.2 fb-1. 2010.

[64] S. Schael et al. Search for neutral MSSM Higgs bosons at LEP. *Eur.Phys.J.*, C47:547–587, 2006.

[65] M. Carena, S. Heinemeyer, O. Stål, C.E.M. Wagner, and G. Weiglein. MSSM Higgs Boson Searches at the LHC: Benchmark Scenarios after the Discovery of a Higgs-like Particle. *Eur.Phys.J.*, C73:2552, 2013.

[66] Nathaniel Craig, Jared A. Evans, Richard Gray, Can Kilic, Michael Park, et al. Multi-Lepton Signals of Multiple Higgs Bosons. *JHEP*, 1302:033, 2013.

[67] 2HDM scenario, H to hh and A to Zh. Technical Report CMS-PAS-HIG-13-025, CERN, Geneva, 2013.

[68] U. Baur, T. Plehn, and D. Rainwater. Probing the higgs self-coupling at hadron colliders using rare decays. *Phys. Rev. D*, 69:053004, Mar 2004.

[69] Georges Aad et al. Search for Higgs Boson Pair Production in the $\gamma\gamma b\bar{b}$ Final State using pp Collision Data at $\sqrt{s} = 8$TeV from the ATLAS Detector. 2014.

[70] Search for resonant HH production in 2gamma+2b channel. Technical Report CMS-PAS-HIG-13-032, CERN, Geneva, 2014.

[71] H. Davoudiasl, J.L. Hewett, and T.G. Rizzo. Phenomenology of the Randall-Sundrum Gauge Hierarchy Model. *Phys.Rev.Lett.*, 84:2080, 2000.

[72] O. DeWolfe, D.Z. Freedman, S.S. Gubser, and A. Karch. Modeling the fifth-dimension with scalars and gravity. *Phys.Rev.*, D62:046008, 2000.

[73] Lisa Randall and Raman Sundrum. A Large mass hierarchy from a small extra dimension. *Phys.Rev.Lett.*, 83:3370–3373, 1999.

[74] S. Heinemeyer et al. Handbook of LHC Higgs Cross Sections: 3. Higgs Properties. 2013.

[75] Graham Cree and Heather E. Logan. Yukawa alignment from natural flavor conservation. *Phys.Rev.*, D84:055021, 2011.

[76] Sheldon L. Glashow and Steven Weinberg. Natural conservation laws for neutral currents. *Phys. Rev. D*, 15:1958–1965, Apr 1977.

[77] Nima Arkani-Hamed, Savas Dimopoulos, and G.R. Dvali. The Hierarchy problem and new dimensions at a millimeter. *Phys.Lett.*, B429:263–272, 1998.

[78] Lisa Randall and Raman Sundrum. An alternative to compactification. *Phys. Rev. Lett.*, 83:4690–4693, Dec 1999.

[79] Serguei Chatrchyan et al. Search for signatures of extra dimensions in the diphoton mass spectrum at the Large Hadron Collider. *Phys.Rev.Lett.*, 108:111801, 2012.

[80] S. Chatrchyan and V. Khachatryan. Measurement of the differential cross section for isolated prompt photon production in *pp* collisions at 7 tev. *Phys. Rev. D*, 84:052011, Sep 2011.

[81] Serguei Chatrchyan et al. Measurement of the Production Cross Section for Pairs of Isolated Photons in *pp* collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 1201:133, 2012.

[82] T. Aaltonen et al. Search for Randall-Sundrum Gravitons in the Diphoton Channel at CDF. *Phys.Rev.*, D83:011102, 2011.

[83] Victor Mukhamedovich Abazov et al. Search for Randall-Sundrum gravitons in the dielectron and diphoton final states with 5.4 fb-1 of data from $p\bar{p}$ collisions at sqrt(s)=1.96 TeV. *Phys.Rev.Lett.*, 104:241802, 2010.

[84] Georges Aad et al. Search for Extra Dimensions in diphoton events using proton-proton collisions recorded at $\sqrt{s} = 7$ TeV with the ATLAS detector at the LHC. *New J.Phys.*, 15:043007, 2013.

[85] Search for scalar diphoton resonances in the mass range 65-600 GeV with the ATLAS detector in pp collision data at $\sqrt{s} = 8$ TeV. Technical Report ATLAS-CONF-2014-031, CERN, Geneva, Jun 2014.

[86] Search for an Higgs Like resonance in the diphoton mass spectra above 150 GeV with 8 TeV data. Technical Report CMS-PAS-HIG-14-006, CERN, Geneva, 2014.

[87] V. M. et al. Abazov. Search for randall-sundrum gravitons with 1 fb^{-1} of data from $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. Lett.*, 100:091802, Mar 2008.

[88] Search for High-Mass Diphoton Resonances in pp Collision at 8 TeV with the CMS Detector. Technical Report CMS-PAS-EXO-12-045, CERN, Geneva, 2014.

[89] H. Davoudiasl, J. L. Hewett, and T. G. Rizzo. Phenomenology of the randall-sundrum gauge hierarchy model. *Phys. Rev. Lett.*, 84:2080–2083, Mar 2000.

[90] H. Davoudiasl, J.L. Hewett, and T.G. Rizzo. Experimental probes of localized gravity: On and off the wall. *Phys.Rev.*, D63:075004, 2001.

[91] Thomas Sven Pettersson and P Lefèvre. The Large Hadron Collider: conceptual design. Technical Report CERN-AC-95-05 LHC, CERN, Geneva, Oct 1995.

[92] Fabienne Marcastel. CERN's Accelerator Complex. La chaîne des accélérateurs du CERN. Oct 2013. General Photo.

[93] The CMS Collaboration. Cms physics technical design report, volume ii: Physics performance. *Journal of Physics G: Nuclear and Particle Physics*, 34(6):995, 2007.

[94] Christian Piasecki. Development of the CMS tracker and reconstruction of secondary vertices of b and c hadrons. 2006.

[95] J. Beringer and J. Arguin. Review of particle physics. *Phys. Rev. D*, 86:010001, Jul 2012.

[96] A.A Annenkov, M.V Korzhik, and P Lecoq. Lead tungstate scintillation material. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 490(1–2):30 – 50, 2002.

[97] CMS Collaboration. Time reconstruction and performance of the cms electromagnetic calorimeter. *Journal of Instrumentation*, 5(03):T03011, 2010.

[98] Serguei Chatrchyan et al. Searches for long-lived charged particles in pp collisions at $\sqrt{s}$=7 and 8 TeV. *JHEP*, 1307:122, 2013.

[99] CMS: The hadron calorimeter technical design report. 1997.

[100] *The CMS muon project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.

[101] Studies of Tracker Material. Technical Report CMS-PAS-TRK-10-003, 2010.

[102] Paolo Meridiani, Chris Seez. Definition of calibrated ecal rechits and the ecal calibration and correction scheme. IN-2011/002, 2011.

[103] The CMS Collaboration. Electromagnetic calorimeter calibration with 7 tev data. *CMS Physics Analysis Summary*, CMS-PAS-EGM-2010-003, 2010.

[104] A. Benaglia, F. Colombo, F. De Guio, L. Di Matteo, R. Gerosa, A. Ghezzi, B. Marzocchi, T. Tabarelli de Fatis, M. Malberti. Analysis of the ecal response stability and uniformity with isolated electrons. *CMS Analysis Note*, 2011/500, 2011.

[105] M. Dejardin, J.-L. Faure, F. Ferri, P. Gras, G. Hamel de Montchenault, P. Jarry, J. Malcles, I. Tecker. Measurement of the crystal transparency losses in cms-ecal using the laser monitoring system. AN-2011/505, 2011.

[106] Andreas Hoecker, Peter Speckmayer, Joerg Stelzer, Jan Therhaag, Eckhard von Toerne, and Helge Voss. TMVA: Toolkit for Multivariate Data Analysis. *PoS*, ACAT:040, 2007.

[107] Florian Beaudette. The CMS Particle Flow Algorithm. pages

[108] Matteo Cacciari and Gavin P. Salam. Pileup subtraction using jet areas. *Phys.Lett.*, B659:119–126, 2008.

[109] W. Adam et al. The CMS high level trigger. *Eur.Phys.J.*, C46:605–667, 2006.

[110] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 0605:026, 2006.

[111] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl.Instrum.Meth.*, A506:250–303, 2003.

[112] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer. MadGraph 5: Going Beyond. *JHEP*, 06:128, 2011.

[113] T. Gleisberg, Stefan. Hoeche, F. Krauss, M. Schonherr, S. Schumann, et al. Event generation with SHERPA 1.1. *JHEP*, 0902:007, 2009.

[114] S Banerjee. Cms simulation software. *Journal of Physics: Conference Series*, 396(2):022003, 2012.

[115] E. Chabanat and N. Estre. Deterministic Annealing for Vertex Finding at CMS. 2005.

[116] Pushpalatha C. Bhat. Multivariate Analysis Methods in Particle Physics. *Ann.Rev.Nucl.Part.Sci.*, 61:281–309, 2011.

[117] The CMS collaboration. Search for Dijet Resonances in 7 TeV pp Collisions at CMS. *Phys.Rev.Lett.*, 105:211801, 2010.

[118] Stefano Goria, Giampiero Passarino, and Dario Rosco. The Higgs Boson Lineshape. *Nucl.Phys.*, B864:530–579, 2012.

[119] Emanuele Re. Higgs production at NNLOPS. 2014.

[120] Nikolas Kauer. Inadequacy of zero-width approximation for a light Higgs boson signal. *Mod.Phys.Lett.*, A28:1330015, 2013.

[121] Diogo Buarque Franzosi, Fabio Maltoni, and Cen Zhang. Effective field theory approach to the Higgs lineshape. *Phys.Rev.*, D87(5):053015, 2013.

[122] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 1006:043, 2010.

[123] Search for the Standard Model Higgs boson in the H to WW to lnujj decay channel in pp collisions at the LHC. Technical Report CMS-PAS-HIG-13-027, CERN, Geneva, 2012.

[124] Daniel de Florian, Nerina Fidanza, R.J. Hernández-Pinto, Javier Mazzitelli, Yamila Rotstein Habarnau, et al. A complete $O(\alpha_S^2)$ calculation of the signal-background interference for the Higgs diphoton decay channel. *Eur.Phys.J.*, C73:2387, 2013.

[125] Stefan Höche, Ye Li, and Stefan Prestel. Higgs-boson production through gluon fusion at NNLO QCD with parton showers. *Phys.Rev.*, D90(5):054011, 2014.

[126] Serguei Chatrchyan et al. Search for heavy narrow dilepton resonances in pp collisions at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV. Technical report, 2013.

[127] Cms luminosity based on pixel cluster counting - summer 2013 update. Technical Report CMS-PAS-LUM-13-001, CERN, Geneva, 2013.

[128] The CMS Collaboration. Procedure for the lhc higgs boson search combination in summer 2011. *CMS Analysis Note*, 2011/298, 2011.

[129] A L Read. Presentation of search results: the cl s technique. *Journal of Physics G: Nuclear and Particle Physics*, 28(10):2693, 2002.

[130] Bernhard Mistlberger and Falko Dulat. Limit setting procedures and theoretical uncertainties in Higgs boson searches. 2012.

[131] SM Higgs Combination. Technical Report CMS-PAS-HIG-11-011, CERN, Geneva, 2011.

[132] P. S. Bhupal Dev and Apostolos Pilaftsis. Maximally Symmetric Two Higgs Doublet Model with Natural Standard Model Alignment. 2014.

[133] Robert V. Harlander, Stefan Liebler, and Hendrik Mantler. SusHi: A program for the calculation of Higgs production in gluon fusion and bottom-quark annihilation in the Standard Model and the MSSM. *Computer Physics Communications*, 184:1605–1617, 2013.

[134] David Eriksson, Johan Rathsman, and Oscar Stal. 2HDMC: Two-Higgs-Doublet Model Calculator Physics and Manual. *Comput.Phys.Commun.*, 181:189–205, 2010.

[135] G. F. Giudice and R. Rattazzi. Theories with gauge-mediated supersymmetry breaking. *Phys. Rept.*, 322:419, 1999.

[136] Glennys R. Farrar and Pierre Fayet. Phenomenology of the production, decay, and detection of new hadronic states associated with supersymmetry. *Phys. Lett. B*, 76:575, 1978.

[137] B. C. Allanach et al. The Snowmass Points and Slopes: benchmarks for SUSY searches. *Eur. Phys. J. C*, 25:113, 2002.

[138] Savas Dimopoulos, Michael Dine, Stuart Raby, and Scott D. Thomas. Experimental signatures of low-energy gauge mediated supersymmetry breaking. *Phys. Rev. Lett.*, 76:3494, 1996.

[139] CMS Collaboration. Search for new physics with long-lived particles decaying to photons and missing energy in pp collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 11:172, 2012.

[140] T. Aaltonen et al. Search for Supersymmetry with Gauge-Mediated Breaking in Diphoton Events with Missing Transverse Energy at CDF II. *Phys. Rev. Lett.*, 104:011801, 2010.

[141] G. Aad et al. Search for diphoton events with large missing transverse momentum in 7 TeV proton-proton collision data with the ATLAS detector. *Phys. Lett. B*, 710:519, 2012.

[142] V. M. Abazov et al. Search for diphoton events with large missing transverse energy in 6.3 fb^{-1} of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Rev. Lett.*, 105:221802, 2010.

[143] CMS Collaboration. Commissioning of the particle-flow reconstruction in minimum-bias and jet events from *pp* collisions at 7 TeV. CMS Physics Analysis Summary CMS-PAS-PFT-10-002, 2010.

[144] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-k_t jet clustering algorithm. *JHEP*, 04:063, 2008.

[145] The CMS Collaboration. Study of timing reconstruction with photons. Technical Report 2012-486, 2011. CMS NOTE-2012/486.

[146] CMS Collaboration. Absolute calibration of the luminosity measurement at CMS: Winter 2012 update. CMS Physics Analysis Summary CMS-PAS-SMP-12-008, 2012.

[147] Serguei Chatrchyan et al. Measurement of Wγ and Zγ production in pp collisions at $\sqrt{s} = 7$ TeV. *Physics Letters B*, 701:535, 2011.

[148] Michiel Botje, Jon Butterworth, Amanda Cooper-Sarkar, Albert de Roeck, Joel Feltesse, Stefano Forte, Alexander Glazov, Joey Huston, Ronan McNulty, Torbjörn Sjöstrand, and Robert Thorne. The PDF4LHC Working Group Interim Recommendations. 2011.

[149] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt. Parton distributions for the LHC. *Eur. Phys. J. C*, 63:189, 2009.

[150] Pavel M. Nadolsky, Hung-Liang Lai, Qing-Hong Cao, Joey Huston, Jon Pumplin, Daniel Stump, Wu-Ki Tung, and C.-P. Yuan. Implications of CTEQ global analysis for collider observables. *Phys. Rev. D*, 78:013004, 2008.

[151] Richard D. Ball, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, José I. Latorre, Juan Rojo, and Maria Ubiali. A first unbiased global NLO determination of parton distributions and their uncertainties. *Nucl. Phys. B*, 838:136, 2010.

[152] ATLAS and CMS Collaborations. Procedure for the lhc higgs boson search combination in summer 2011. ATL-PHYS-PUB 2011-011, 2011. CMS NOTE-2011/005.

[153] Thomas Junk. Confidence level computation for combining searches with small statistics. *Nucl. Instrum. Meth. A*, 434:435, 1999.

[154] Alexander L. Read. Presentation of search results: The CL_s technique. *J. Phys. G*, 28:2693, 2002.

[155] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C*, 71:1554, 2011.