University of California
Santa Barbara

Search for Higgs boson pair production in final states with photons and leptons

A dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy
in
Physics

by

Francesco Setti

Committee in charge:

Professor Claudio Campagnari, Chair
Professor Nathaniel Craig
Professor Hugh Lippincott

September 2023

The Dissertation of Francesco Setti is approved:

Professor Nathaniel Craig

Professor Hugh Lippincott

Professor Claudio Campagnari, Committee Chair

June 2023

Search for Higgs boson pair production in final states with photons and leptons

by

Francesco Setti

Acknowledgements

My whole PhD has been a fantastic adventure, which was made even more incredible by the people that I encountered along the way. Therefore, I feel the need to express my gratitude to many of them. Proceeding in a somewhat chronological order, I am obliged to thank Professor Claudio Campagnari for his animated and charismatic guidance and, more importantly, for giving me this unique opportunity. From Claudio I learned the importance of discerning primary and fundamental problems to focus on, as opposed to non-essential ones. This may not seem immensely valuable, but the ability to direct one's efforts and energies toward more fundamental problems is essential for success in an environment where "rabbit holes" emerge every day, with the risk of getting lost in them without a clear resolution. This is something that I feel has helped me grow not just as a scientist but also as a humanoid. Following Claudio, I must express my gratitude to another central figure in my PhD, Professor Avi Yagil. Avi inspired me with his innate curiosity and unfiltered conversations about the most various topics in particle physics. The harmonious duets between Avi and Claudio always kept me engaged in our meetings and demonstrated the most pure, real, and spontaneous form of discussion I have ever witnessed. The straightforward and simple manner in which they dissected and analysed many topics will always serve as a great example of effective communication to me. While bidding farewell to this remarkable group of individuals brings a touch of sadness, I also consider myself fortunate to have gained invaluable experience from working with them and to have been part of an unparalleled working environment that will remain close to my heart.

Next in line, I was fortunate to work with Dr. Hualin Mei, who is, simply stated, a genius physicist and an indefatigable machine. He is always (and I mean always) ready to help, responsive on Skype, possesses extensive knowledge about everything, and yet remains

extremely humble. His dedication to work and clarity of thought have been tremendously inspiring to the point that, I must admit, he has been the most influential person I have ever worked with. An unsurpassable example of a colleague and friend whom I doubt I will be able to find in the future, sadly. I also want to thank Samuel May, who, along with Hualin, guided me through the initial stages of this analysis and patiently taught me about Higgs searches. Sam constantly motivated me to expand my knowledge about my research by setting an exemplary standard.

When I first joined Claudio's group at UCSB, graduate students (at that time) Nick Amin and Bennett Marsh warmly welcomed me and made me feel at home from day one. Their dedication to research and work ethic were truly inspirational, and I will always be grateful to them for the time they spent explaining various topics to me, from computer science to physics. I hope I have made a similar impression on my current colleagues Joe, Keegan, and Jason. They have been fantastic friends, and I owe them the little sanity of mind that, I believe, I still possess. I could never forget the long walks on the beach after our lunchtime feasts and the countless adventures along the Santa Barbara coast.

This whole experience would have not been possible without the support of my mother and my sister Valeria. They have always been and will always be by my side advising me on the best decisions to make. I really wouldn't be the person and physicist I am today without them, for good and for bad. I also want to thank you Emily for putting up with me in this very stressful year of my life and for (almost) never losing patience with me. Your support has meant a lot to me.

As I notice that I am rapidly approaching the second page of acknowledgments in this double-spaced text, I will try to be more concise (since I haven't even started the main text yet) by expressing my gratitude to all of my friends and colleagues who supported me during this extraordinary journey, such as Estefania, Andre, Malik, Chris, Sam and Mel (and our brunches), Camilo, Ulascan, and many, many others!

Curriculum Vitæ

Francesco Setti

Education

2023	Ph.D. in Physics, University of California, Santa Barbara.
2020	M.A. in Physics, University of California, Santa Barbara.
2018	MSci in Physics with Theoretical Physics, Imperial College London.

Abstract

Search for Higgs boson pair production in final states with photons and leptons

by

Francesco Setti

A search for Higgs boson pair production via gluon-gluon fusion in final states with two photons and one or more leptons is presented. The analysis is based on data from proton-proton collisions at a center-of-mass energy of 13 TeV collected with the CMS detector at the LHC, corresponding to an integrated luminosity of 138 fb^{-1}. No significant excess above the background-only prediction is observed. An upper limit is set on the gluon-gluon fusion Higgs boson pair production cross section. The observed (expected) 95% confidence level (CL) upper limits are found to be 25 (23) times the Standard Model (SM) prediction. Assuming all other couplings of the Higgs boson have values as predicted by the SM, the observed (expected) constraints on the modifier of the Higgs boson trilinear self-coupling κ_λ are found to be $-11.1 < \kappa_\lambda < 16.6$ ($-8.7 < \kappa_\lambda < 14.7$) at 95% CL.

Table of Contents

Part I

Theoretical Framework

1 Arisal of the Standard Model

In this chapter we will present the theoretical background that is necessary to fully appreciate and understand the reasons behind our search for the di-Higgs boson production at the Large Hadron Collider (LHC). We will proceed with an overview of the structure of the Standard Model (SM), followed by a detailed discussion of the Higgs mechanism and the phenomenology associated with di-Higgs production at hadron colliders. Lastly, we will give an overview of the current experimental status of di-Higgs searches at the LHC.

The Standard Model (SM) of particle physics is a collection of quantum field theories that aims at concisely and elegantly summarising the fundamental forces and interactions of nature. It is an enormous effort that thousands of scientists have worked on and improved over the central decades of the XX century. However, it is nowhere near a complete or final theory. In fact, the SM includes only three of the four fundamental forces: electromagnetism or Quantum Electro Dynamic (QED), strong force or Quantum Chromo Dynamic (QCD) and weak force; whereas there is no current formulation of a quantum field theory of gravity. The weak, strong and electromagnetic interactions can be unified and described by a single theory that presents a local gauge invariance under the symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$. According to the SM, the building blocks of the universe are fermions, which constitute the matter that we observe around us, while

the interactions between fermions are mediated by bosons.

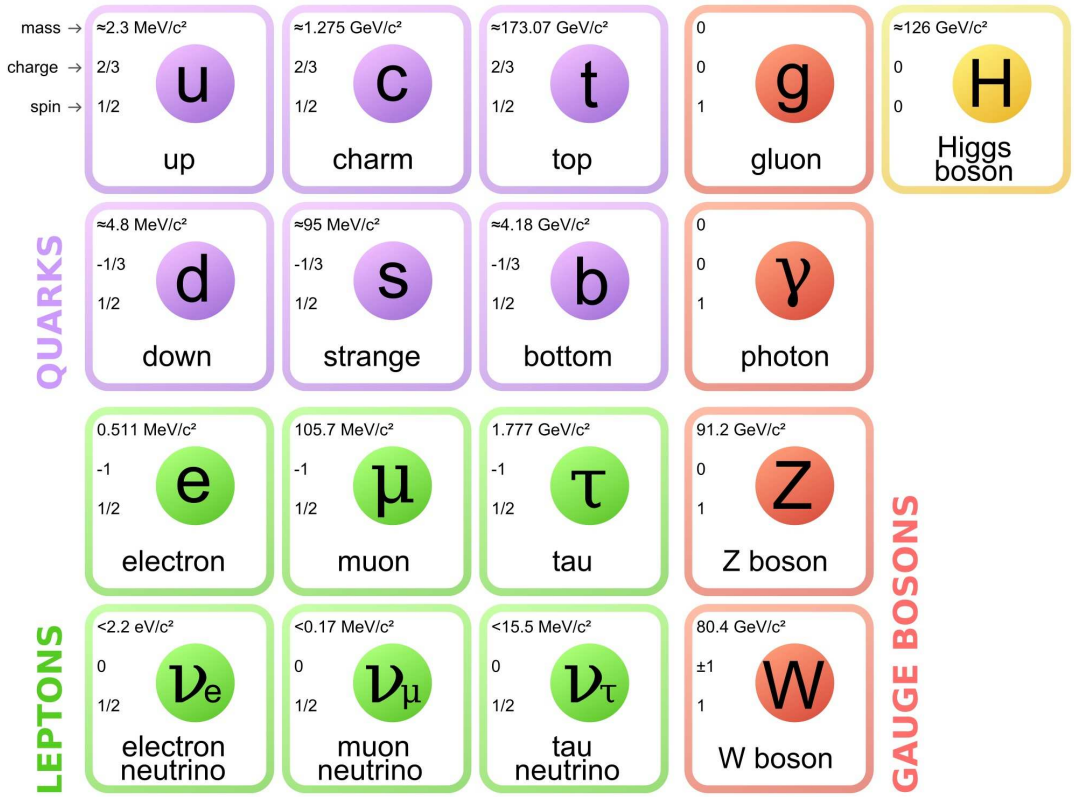


Figure 1: Schematic diagram of the particles currently described by the Standard Model. In the white boxes are shown fermions, while the colored boxes display the five elementary bosons. Picture from [1]

1.1 Fermions

Fermions are the building blocks of matter in our universe and, in the SM, they correspond to quantum fields with non-integer spin $S = 1/2$ which are subject to the Fermi-Dirac statistics. Fermions are divided into leptons and quarks: leptons couple to the weak force and electromagnetism (with the exception of neutrinos) but not to the strong interaction, whereas quarks couple with both the electro-weak and strong force. Each lepton and quark family is divided into three generations

with two fermions each , as shown in Fig. 1, where successive generation maintain the same physical properties as the previous ones but exhibit heavier masses. Along with fermionic particles (matter), there also exist fermionic anti-particles (anti-matter) that exactly match their particle counterparts in masses and spins but with opposite charge and magnetic moment.

1.1.1 Leptons

Charged leptons can be divided into three families: the electron (e with a mass of 0.511MeV/c^2), the muon (μ, mass 105.7 MeV/c^2) and the tau (τ, the heaviest with as mass of 1.77 GeV/c^2). Each lepton family includes a neutrino ν, therefore we have an electron neutrino ν_e, a muon neutrino ν_μ and a tau neutrino ν_τ. Neutrinos are not electrically charged hence they only interact via the weak force. Several experiments are dedicated to measure the properties of neutrinos like their masses, their flavor mixing, but to this date we only know that their masses are non-zero (although very close to).

1.1.2 Quarks

Quarks are also divided into three generations or families: up quark u and down quark d belong to the first and lightest family, strange quark s and charm quark c belong to the second family, and top t and bottom b quarks belong to the third and heaviest generation. Quarks are subject to all three fundamental forces (e.g. weak, strong and electromagnetic) and are the only fermions that couple to the strong force via the color charge. As an interesting consequence of the QCD color confinement properties, quarks do not exist in free states but are only observable in bound states. These are referred to as mesons if they contain a quark and an anti-quark

or baryons if instead they contain a triplet of quarks or anti-quarks. Generally, we refer to mesons and baryons collectively as hadrons, which also gives the name to the fantastical machine that allowed the discovery the Higgs Boson in 2012 and that we used to collect the data analised in this thesis.

1.2 Bosons

Bosons are quantum fields that obey the Bose-Einstein statistics, which means that they are not affected by spatial density and can overlap in space. They are the carriers of all the fundamental forces and have integer spins. Photons γ, $W^{\pm}$ and Z bosons are the mediators of the electro-weak force and correspond to the gauge bosons of the electro-weak group $SU(2)_L \times U(1)_Y$, whereas eight gluons mediate the strong interaction and correspond to the generators of the strong Lie group $SU(3)_C$. The photon has no electric charge and no mass; the two $W^{\pm}$ bosons instead have an electrical charge of $Q = \pm 1$ and mass $m_W = 80.3\ \mathrm{GeV/c^2}$, whereas the Z boson has a mass of $m_Z = 91.19\ \mathrm{GeV/c^2}$ and neutral electric charge.

Gluons belong to a different symmetry group and exhibit different characteristics: they carry color quantum numbers traditionally referred to as red, green and blue. There are eight color and anti-color charges in QCD, which are equivalent to the electric charges in QED. Gluons interact with any other colored particles, therefore they can self-interact with other gluons or with colored quarks and antiquarks. Neither mass nor electrical charge is associated with the mediators of the strong force. A more detailed discussion of their properties and their role in the SM is covered in the remaining paragraphs of this chapter.

1.3 Gauge Invariance

Since we want our mathematical description of our universe to be as precise and "physical" as possible, we need to require our SM Lagrangian to be symmetric under a local gauge transformation of a field. This mathematical constraints is equivalent to saying that the laws of physics should be the same over any point in spacetime. If we attempt to write a mathematical model / description of a free fermionic field ψ with no interactions and mass m, we would end up with a Lagrangian of the form:

$$\mathcal{L} = i\overline{\psi}\gamma_\mu\partial^\mu\psi - m\overline{\psi}\psi, \tag{1}$$

where the first term relates to the kinetic energy of the field and the second term to the mass of the fermionic field. If we then require such structure to be invariant under a local gauge transformation, this implies that the Lagrangian 1 is left unchanged by a transformation $U(x)$ of a field $\psi(x)$ as:

$$\psi(x) \to U(x)\,\psi(x) = e^{i\alpha_a(x)T_a}\,\psi(x), \tag{2}$$

where $\psi(x)$ is the fermionic field, U the unitary operator of the corresponding symmetry group of the transformation, a the index for each generator of the symmetry group, $\alpha_a(x)$ is the gauge parameter, which is a function of the spacetime coordinates x, and T_a the generators of the symmetry group associated with the interaction. If we expand the unitary operator by performing an infinitesimal transformation of the field, we get:

$$\psi(x) \rightarrow (1 + i\alpha_a T_a)\, \psi(x) \tag{3}$$

and consequently

$$\partial_\mu \psi(x) = (1 + i\alpha_a T_a)\, \partial_\mu \psi(x) + i T_a\, \psi(x)\, \partial_\mu \alpha_a \tag{4}$$

with ∂_μ being shorthand for the spacetime derivative:

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = (+\frac{1}{c}\frac{\partial}{\partial t}\, ,\, \nabla\,) \tag{5}$$

Now, if it wasn't for the last term in equation 4, the infinitesimal transformation would preserve local gauge invariance, but this is not the case. In order to account for this extra term we have to introduce a new vector field G^a_μ which needs to obey the transformation relation:

$$G^a_\mu \rightarrow G^a_\mu - \frac{1}{g}\, \partial_\mu\, \alpha_a(x). \tag{6}$$

However, in the case that the symmetry group of the interaction that we are dealing with is non-Abelian, the vector field G undergoes a slightly different transformation relation, given that the group generators do not necessarily commute ($[T_a, T_b] \neq 0$), so we have instead:

$$G^a_\mu \rightarrow G^a_\mu - \frac{1}{g}\, \partial_\mu\, \alpha_a(x) - f_{abc}\, \alpha_b\, G^c_\mu, \tag{7}$$

where f_{abc} is the tensor notation for the structure constants of the symmetry group. In other terms, the vector field transformation in Eq. 7 is equivalent to defining a new derivative D_μ, commonly referred to as the covariant derivate, that allows to preserve local gauge invariance of the Lagrangian under an infinitesimal gauge transformation of the field:

$$D_\mu = \partial_\mu + i\, g\, T_a\, G^a_\mu. \tag{8}$$

If we then replace $\partial_\mu \rightarrow D_\mu$ in Eq. 1 and add a new kinetic term for the new field to maintain the Lagrangian gauge invariant, the equation becomes:

$$\mathcal{L} = i\,\overline{\psi}\,\gamma_\mu\,\partial^\mu\,\psi \;-\; m\,\overline{\psi}\,\psi \;-\; g\,(\,\overline{\psi}\,\gamma^\mu\,T_a\,\psi\,)\,G^a_\mu \;-\; \frac{1}{4}\,F^a_{\mu\nu}\,F^{\mu\nu}_a, \tag{9}$$

in which we introduced a number a of new vector fields G^a_μ that couple to ψ with a coupling strength g and whose kinetic terms are summarised by the field strength tensor $F^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu$. We shall noticed that in order to preserve the local gauge invariance of the Lagrangian, there cannot be present a mass term related to the gauge fields G (e.g. no terms involving the interaction $G^a_\mu G^\mu_a$), and therefore such gauge fields must be massless.

1.4 The Strong Interaction

The strong force, or QCD, is a gauge theory that describes the interaction between gluons and quarks. It is based on the non-abelian Lie group $SU(3)_C$, where C is short for color, the charge associated with this interaction. The correct formulation of QCD is obtained by requiring local gauge invariance of the strong Lagrangian:

$$\mathcal{L} = \overline{\psi}(x)\,(\,i\,\not{\partial} - m\,)\,\psi(x) \tag{10}$$

under the $SU(3)$ group symmetry. In the above equation, $\not{\partial}$ is a shortcut for $\gamma_\mu \partial^\mu$ where we contract the partial four derivative ∂^μ with the gamma matrices γ_μ, and ψ is a fermionic field of mass m. Next, we perform an infinitesimal transformation under the $SU(3)$ Lie group according to Eq. 2 as:

$$\psi(x) \rightarrow e^{i\,g\,\frac{\lambda^a}{2}\,\theta_a(x)}\,\psi(x), \tag{11}$$

where $\frac{\lambda^a}{2}$ represent the eight Gell-Mann matrices and generators of the non-Abelian symmetry group, and g the strong coupling. In order to maintain this theory gauge invariant under a local transformation, it is necessary to introduce the gauge vector fields G^a_μ such that they transform under spacetime dependent rotations as:

$$G^a_\mu = G^a_\mu + \partial_\mu\,\theta^a(x) + g\,f^{abc}\,G^c_\mu\,\theta^c \tag{12}$$

which will be integrated into the covariant derivate to give:

$$D_\mu = \partial_\mu - i\,g\,G^a_\mu(x)\,\frac{\lambda^a}{2}. \tag{13}$$

It is important to notice that in this case the structure constants f^{abc} can be obtained via the relation:

$$\left[\frac{\lambda^a}{2}, \frac{\lambda^b}{2}\right] = i\, f^{abc}\, \frac{\lambda^c}{2}. \tag{14}$$

The missing ingredient to complete the strong force Lagrangian is the kinetic term associated with the gluon fields $-\frac{1}{4} F_a^{\nu\mu} F^a_{\mu\nu}$, where:

$$F^a_{\mu\nu} = \partial_\mu\, G^a_\nu - \partial_\nu\, G^a_\mu + g\, f^{abc}\, G^b_\mu\, G^c_\nu \tag{15}$$

and the last term arises due to the fact the $SU(3)$ is a non-Abelian group. To conclude, the full QCD Lagrangian density reads:

$$\mathcal{L}_{QCD} = \overline{\psi}(x)(\, i\, \partial\!\!\!/ - m\,)\, \psi(x) - g\, \overline{\psi}(x)\, \gamma^\mu\, \frac{\lambda^a}{2}\, \psi\, G^a_\mu - \frac{1}{4}\, F_a^{\nu\mu}\, F^a_{\mu\nu}, \tag{16}$$

where the first term is unchanged from the original Lagrangian and represents the free field quark propagation, the second term arises from the introduction of the covariant derivative in Eq. 13 and describes the interaction between two quark fields ψ and one gluon field G^a_μ. In Eq. 16, we implicitly sum over the index a which represent the eight gluon fields or generators of the $SU(3)_C$ symmetry group. It can be convenient to define the coupling strength of the strong interaction in terms of the strong coupling constant α_S as:

$$\alpha_S = \frac{g^2}{4\pi}. \tag{17}$$

An interesting consequence of the fact that the strong force can be represented by a non-Abelian group is the fact that, as per Eq. 14, the f^{abc} constants are not nec-

essarily zero and therefore a cubic and a quadratic interaction term for the gluon field arise from the expansion of their kinetic term (last one in Eq. 16). Such terms represent the self-interactive nature of gluons, a direct consequence of the structure of the $SU(3)$ group.

1.5 Electro-weak Interaction

Similarly to the arisal of the strong interaction due to the $SU(3)_C$ local gauge invariance of the SM Lagrangian, the Electro-Weak force is a direct consequence of the gauge invariance associated with the $SU(2)_L \times U(1)_Y$ group symmetry. The beautiful simplicity (or complexity) of the SM is that there is no standalone formulation of the weak interaction and the electromagnetic force, but instead they are intrinsically connected to one another and their manifestation is made possible by the symmetry breaking pattern of the $SU(2)_L \times U(1)_Y$ gauge group. $SU(2)_L$ is a non-Abelian gauge group and its associated charge is commonly referred to as the weak isospin quantum number I_3, the L subscript refers to the "left" chirality nature of this interaction, since the three vector fields of this group $W^i_\mu (i = 1, 2, 3)$ only couple to fermionic fields of left chirality also referred to as $SU(2)_L$ doublets; whereas fermions of right chirality are $SU(2)_L$ singlets and do not interact with the weak W^i_μ vector fields. The $U(1)_Y$ gauge group is instead Abelian and its associated field is denoted as B_μ, while its charge is the weak hypercharge Y. Both left hand (ψ_L) and right hand (ψ_R) fermionic fields interact with the $U(1)_Y$ associated gauge field. The concept of chirality can be explained by introducing the γ^5 matrix, defined as:

$$\gamma^5 = i\,\gamma^0\,\gamma^1\,\gamma^2\,\gamma^3, \tag{18}$$

where the γ^is are the usual gamma (or Dirac) matrices. The γ^5 matrix allow us to decompose any Dirac spinor into left- and right- handed chiral components using the projection operators P_L and P_R, which are defined as:

$$P_L = \frac{1-\gamma^5}{2}, \tag{19}$$

$$P_R = \frac{1+\gamma^5}{2}. \tag{20}$$

In the context of the electro-weak interaction it is useful to define the left and right chiral projections of a fermion field:

$$\Psi_L = \frac{1-\gamma^5}{2}\begin{pmatrix}\psi\\ \psi'\end{pmatrix} = \begin{pmatrix}\psi_L\\ \psi'_L\end{pmatrix} \tag{21}$$

$$\psi_R = \frac{1+\gamma^5}{2}\,\psi \tag{22}$$

$$\psi'_R = \frac{1+\gamma^5}{2}\,\psi' \tag{23}$$

in which ψ and ψ' represent the fermion fields. For instance, to describe the electron - neutrino doublet we use the formalism:

$$\psi_L(x) = \begin{pmatrix}\nu_e\\ e\end{pmatrix}_L; \qquad \psi_R(x) = \nu_{eR}; \qquad \psi'_R(x) = e_R \tag{24}$$

and similarly we can use the same formula to express the up and down quarks.

Since we are mentioning the coupling of the weak vector fields to fermions of left chirality, it is important to not confuse chirality with helicity. Helicity eigenstates are defined by the normalised projection of the spin of a particle onto its direction of motion. Chirality states, instead, are the eigenstates of the γ^5 matrix. If we look at the decomposition of a right-handed helicity spinor ($u_\uparrow$) into its left- and right-handed chiral components, u_L and u_R respectively, we obtain:

$$u_\uparrow \propto \frac{(1-\kappa)}{2} u_L + \frac{(1+\kappa)}{2} u_R \tag{25}$$

with $\kappa = \frac{p}{E+m}$, p the momentum, E the energy and m the mass of the particle. We notice that in the limit of a relativistic particle ($E \gg m$), then $\kappa \to 1$ and the u_L factor approaches 0, then only in this case we have an overlap between helicity and chirality states.

Going back to the electro-weak interaction, Kazuhiko Nishijima first and Murray Gell-Mann after were able to find a correlation between the conserved charges I_3 and Y with the more physical concept of electric charge Q. The formula takes the name from both physicists, since they came up with the idea independently and only a few years apart (1953 versus 1956):

$$Q = I_3 + \frac{Y}{2}. \tag{26}$$

Now let's take a look at the mathematics. A Lagrangian of this type would take the form:

$$\mathcal{L}_{EW} = i\,\overline{\Psi}_L \not{D}\,\Psi_L + i\,\overline{\psi}_R \not{D}\,\psi_R + i\,\overline{\psi}^{'}_R \not{D}\,\psi^{'}_R \tag{27}$$

in this case, the contracted covariant derivative $\not{D} = \gamma_\mu D^\mu$ is defined as:

$$D_\mu = \partial_\mu - i\,g\,W^i_\mu\,T_i + i\,g^{'}\,\frac{Y}{2}\,B_\mu \tag{28}$$

where the $SU(2)_L$ generators T_i take the form of Pauli matrices σ_i as $T_i = \sigma_i/2$ for the left handed chiral fermions and instead they become zero ($T_i = 0$) for the right handed chiral fermion fields ψ_R and $\psi^{'}_R$. In addition, g is the coupling strength associated with the gauge vectors of the $SU(2)_L$ group and $g^{'}$ the coupling strength related to B_μ. Per se, this formulation is not particularly enlightening in describing the electro-weak interaction; for this reason it is useful to re-write Eq. 27 in these terms:

$$\mathcal{L}_{EW} = \mathcal{L}_{kin} + \mathcal{L}_{CC} + \mathcal{L}_{NC} \tag{29}$$

where we have rearranged the terms as:

$$\mathcal{L}_{kin} = i\,\overline{\Psi}_L\,\partial\!\!\!/\,\Psi_L + i\,\overline{\psi}_R\,\partial\!\!\!/\,\psi_R + i\,\overline{\psi}^{'}_R\,\partial\!\!\!/\,\psi^{'}_R$$

$$\mathcal{L}_{CC} = \frac{g}{\sqrt{2}}\,W^{+}_{\mu}\,\overline{\psi}_L\,\gamma^{\mu}\,\psi^{'}_L + \frac{g}{\sqrt{2}}\,W^{-}_{\mu}\,\overline{\psi}^{'}_L\,\gamma^{\mu}\,\psi_L \tag{30}$$

$$\begin{aligned}\mathcal{L}_{NC} = {} & \frac{g}{\sqrt{2}}\,W^{3}_{\mu}(\,\overline{\psi}_L\,\gamma^{\mu}\,\psi_L - \overline{\psi}^{'}_L\,\gamma^{\mu}\,\psi^{'}_L) \\ & + \frac{g^{'}}{\sqrt{2}}B_{\mu}\,\left[Y_{\Psi_L}\,\left(\overline{\psi}_L\,\gamma^{\mu}\,\psi_L + \overline{\psi}^{'}_L\,\gamma^{\mu}\,\psi^{'}_L\right) + Y_{\psi_R}\,\overline{\psi}_R\,\gamma^{\mu}\,\psi_R + Y_{\psi^{'}_R}\,\overline{\psi}^{'}_R\,\gamma^{\mu}\,\psi^{'}_R\right]\end{aligned}$$

and defined

$$W^{\pm}_{\mu} = \frac{1}{\sqrt{2}}(\,W^{1}_{\mu} \mp i\,W^{2}_{\mu}\,) \tag{31}$$

to be the electrically charged W weak bosons W^{+} and W^{-} (hence CC, shortcut for Charged Current), which couples exclusively to the left chiral states ψ_L and $\psi^{'}_L$. Now, the two remaining fields W^{3}_{μ} and B_{μ} do not necessarily resemble the electromagnetic interaction that we are used to. In order to gain a deeper insight into the neutral current part of the Lagrangian ($\mathcal{L}_{NC}$), it is useful to redefine these fields via a linear superposition which takes as a parameter the Weinberg mixing angle θ_W:

$$\begin{pmatrix} W^{3}_{\mu} \\ B_{\mu} \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} Z_{\mu} \\ A_{\mu} \end{pmatrix} \tag{32}$$

Where we can now substitute the expressions for W_μ^3 and B_μ to formulate the neutral current in terms of the neutrally charged Z boson (Z_μ gauge field) and the photon field (A_μ). According to this redefinition of the fields, we must set the coupling strengths g and $g^{'}$ to match the electromagnetic charge e as:

$$g \sin\theta_W = g^{'} \cos\theta_W = e. \tag{33}$$

After constraining the remaining free parameters in the Lagrangian, the full electroweak Lagrangian formulation reads:

$$\mathcal{L}_{EW} = i\,\overline{\Psi}_L \not{D}\, \Psi_L + i\,\overline{\psi}_R \not{D}\, \psi_R + i\,\overline{\psi}^{'}_R \not{D}\, \psi^{'}_R - \frac{1}{4}\, B^{\mu\nu}\, B_{\nu\mu} - \frac{1}{4}\, W_i^{\mu\nu}\, W^i_{\nu\mu} \tag{34}$$

in which we have defined the $W_i^{\mu\nu}$ and $B^{\mu\nu}$ field strength tensors to be:

$$B^{\mu\nu} = \partial^\mu B^\nu - \partial^\nu B^\mu \tag{35}$$

$$W_i^{\mu\nu} = \partial^\mu W_i^\nu - \partial^\nu W_i^\mu + g\,\varepsilon^{ijk}\, W_j^\mu\, W_k^\nu \tag{36}$$

If we analyse Eq. 34, we observe that the first three terms are the free fermionic field kinetic terms for Ψ_L, ψ_R and $\psi^{'}_R$ and the last two terms contribute in the form of the neutral and charged current interactions. It is interesting to note that by expanding the $W_i^{\mu\nu}$ term, a number of new interactions appear in the Lagrangian; these are both trilinear couplings of the W, Z and γ like γWW or ZWW, but also quadrilinear terms such as $\gamma\gamma WW, \gamma ZWW, ZZWW$ and $WWWW$. All such verteces and

amplitudes have been measured experimentally and were found to be in excellent agreement with the predictions from the SM! A final remark should go the absence of mass terms for the fermion fields and also for the W_i^μ and B^μ gauge fields. As previously discussed, a mass term for the gauge fields would break the local gauge invariance. In addition, fermionic masses of the form $m\,\psi\,\overline{\psi}^{'} = m(\,\overline{\psi}_R\,\psi_L + \overline{\psi}_L\,\psi_R\,)$ would also break the symmetry, since it incorrectly mixes right handed singlets with left handed doublets that transform differently under the $SU(2)_L \times U(1)_Y$ group. Both of these issues will be discussed in the next section.

2 The Higgs Mechanism

Since no mass terms are allowed for the vector fields in the QCD or Electro-Weak Lagrangian, there has to be a way in which we can assign a mass to all of these particles. This method is called the Higgs Mechanism, and it generates masses for the gauge bosons along with fermion masses. Given the fact that a massive vector field has three degrees of polarisation and a massless one has only two, the minimum number of degrees of freedom for a theory that generates masses for both massive and massless gauge vector fields has to be at least three. The simplest choice consists of a complex scalar doublet of fields $\Phi(x)$ such as:

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix} \tag{37}$$

where ϕ^+ represents a charged scalar field and ϕ^0 a neutral scalar field, which in total add up to four degrees of freedom.

Let us study what happens If we add the following terms to the SM Lagrangian:

$$\mathcal{L}_{Higgs} = (D_\mu \Phi)^\dagger (D^\mu \Phi) + V(\Phi^\dagger \Phi) \tag{38}$$

where the potential $V(\Phi^\dagger \Phi)$ is defined as:

$$V(\Phi^\dagger \Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \qquad \text{with } \mu^2, \lambda > 0. \tag{39}$$

There are infinitely many minima with same energy (degenerate) that satisfy the conditions to be at the minimum:

$$|\Phi|^2 = \frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2} \tag{40}$$

with $v^2 = \mu^2/\lambda$ being the vacuum expectation value (vev) of the potential $V(\Phi^\dagger \Phi)$. Now, without loss of generalities, we can pick any state that we prefer, according to Eq. 40, breaking the initial symmetry of the potential. This is well illustrated in Fig. 2.

However, the Lagrangian is still invariant under the global $U(1)$ transformation, and therefore we can expand the newly introduced field around its minimum by a small excitation $H(x)$ according to:

$$\Phi(x) = \frac{1}{\sqrt{2}} \left[\frac{i\,\sigma_i\,\theta^i(x)}{v} \right] \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{41}$$

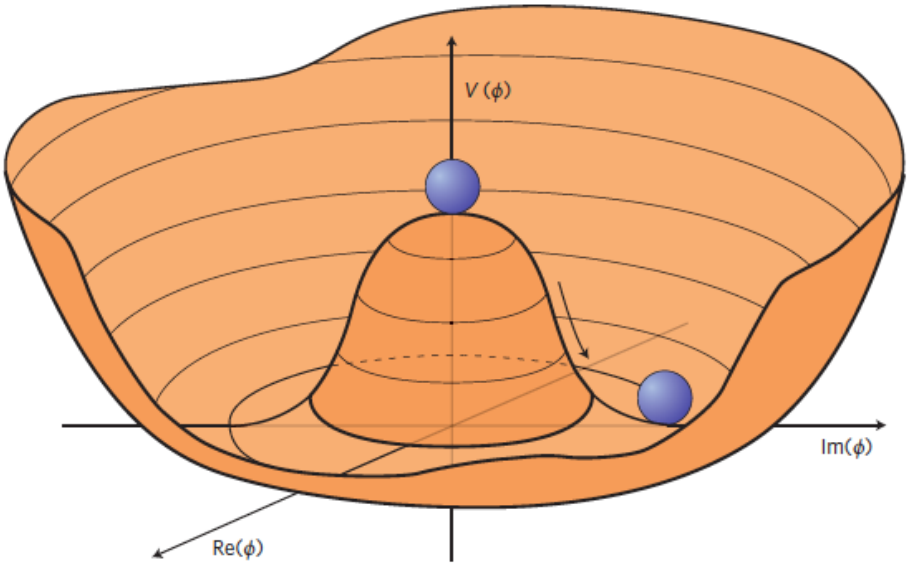


Figure 2: Higgs potential as described by Eq. 39 when $\mu^2 > 0$ and the minimum of the scalar field $|\phi_0| = v/\sqrt{2}$. By choosing any of the points along the minimum of the potential $V(\phi_0) = -\lambda v^4/4$ the $U(1)$ symmetry is broken. Figure from [2].

where $\theta^i(x)$ correspond to the three massless Goldstone bosons generated by the spontaneous symmetry breaking process. These massless fields are not physical, and can be removed by choosing an appropriate gauge, called the unitary gauge, via a $SU(2)_L$ transformation:

$$\Phi(x) \rightarrow \Phi^{'}(x) = \exp\left(-\frac{i\,\sigma_i\,\theta^i(x)}{v}\right)\,\Phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{42}$$

which leaves us with the real scalar field $H(x)$, corresponding to a physical massive field called the Higgs field, and the three Goldstone bosons are "eaten" by the W and Z fields in order to become massive.

If we write the Lagrangian with the inclusion of the Higgs field $H(x)$, we then have an interaction of the form:

$$\begin{aligned}\mathcal{L}_{Higgs} = & \frac{1}{2}\,\partial^{\mu}H\,\partial_{\mu}H - \frac{1}{2}\,(\,2\,\lambda\,v^{2})\,H^{2} \\ & + \left[\left(\frac{g\,v}{2}\right)^{2} W^{+}_{\mu}\,W^{\mu-} + \frac{1}{2}\frac{(g^{2}+g'^{2})\,v^{2}}{4} Z_{\mu}\,Z^{\mu}\right]\left(1+\frac{H}{v}\right)^{2} \\ & + \lambda\,v\,H^{3} + \frac{\lambda}{4}\,H^{4} + \frac{\lambda}{4}\,v^{4}\end{aligned} \tag{43}$$

The first two terms of Eq. 43 describe the free field of the Higgs boson, a massive particle with mass $m_H^2 = 2\lambda v^2 = 2\mu^2$. The other terms on the second line represent the interactions between the Higgs field and the W and Z vector bosons, along with the self-interaction of the latter two gauge fields. These terms also describe the masses for the W and Z as:

$$\begin{aligned} m_W^2 &= \frac{g^2\,v^2}{4} \\ m_Z^2 &= \frac{(g^2+g'^2)\,v^2}{4} = \frac{m_W^2}{cos^2\theta_W} \end{aligned} \tag{44}$$

Let us note that this theory is consistent with our model of physics, since the absence of an interaction term between the Higgs field and A_μ ensure that the photon remains massless.

Lastly, the third row of Eq. 43 describes the interacting terms between the Higgs field and itself, also referred to as "Higgs self-coupling" λ_{HHH}, and we are presented with a trilinear term HHH and quadrilinear term HHHH. It is useful to note that the Higgs trilinear and quadrilinear self-couplings, λ_{HHH} and λ_{HHHH} respectively, are usually expressed inside the Higgs potential as:

$$V_{Higgs}(\phi) = \frac{1}{2} m_H^2 \, H^2 + \lambda_{HHH} \, v \, H^3 + \frac{1}{4} \, \lambda_{HHHH} \, H^4 \tag{45}$$

where we defined:

$$\lambda_{HHH} = \lambda_{HHHH} = \frac{m_H^2}{2 \, v^2} \tag{46}$$

which suggests the interesting nature of this interaction, where the Higgs self-coupling is related to its own mass and the shape of the Higgs potential. In the following chapters, we will describe how our analysis is particularly relevant in constraining λ_{HHH} via a direct measurement of the Higgs pair production, and therefore providing a test for the Higgs potential as it is predicted by the SM.

One of the missing items, before we conclude our discussion of the Higgs mechanism, is the emergence of fermionic masses. So far, we only covered the way in which vector bosons become massive, but what about fermions? The process by which fermions gain mass is also mediated by the Higgs boson. First, we need to introduce a new type of interaction term into our Lagrangian. This type of interaction is called Yukawa interaction and, by using our notation of lepton fields as in Eq. 24, is expressed as:

$$\mathcal{L}_{Yukawa} = - \, y_e \left(\overline{\psi}_R \, \phi^\dagger \, \psi_L + \overline{\psi}_L \, \phi \, \psi_R \right), \tag{47}$$

where y_e represents the coupling constant between the Higgs field and an electron. If we now use the scalar doublet formalism to express the Higgs field (Eq. 42), we

end up with the expression:

$$\mathcal{L}_{Yukawa} = -y_e \frac{v+H}{\sqrt{2}} \left(\bar{e}_R e_L + \bar{e}_L e_R\right), \tag{48}$$

which can be re-arranged into electron fields $e = (\, e_R \,,\, e_L)$ and $\bar{e} = (\, \bar{e}_R \,,\, \bar{e}_L)$ as:

$$\mathcal{L}_{Yukawa} = -y_e \frac{v}{\sqrt{2}} \bar{e}\, e - y_e \frac{H}{\sqrt{2}} \bar{e}\, e \tag{49}$$

in which we recognise a mass term for the electron m_e in $m_e = y_e \cdot v \,/\, \sqrt{2}$, along with an interaction term between the Higgs field and the electron. A similar procedure is used to generate masses for all fermions. An interesting consideration is the fact that the coupling strength y_e between the Higgs boson and the electron (or any other fermion) is directly proportional to the mass of the electron (or fermion) itself! Because of this peculiar characteristic, the Higgs boson tends to decay to the heaviest kinematically accessible fermions. As we previously saw, the interaction strength between the Higgs and vector bosons depends quadratically on the vector boson masses, so even in this case our Higgs field is more inclined to decay or couple to heavier vector fields, all kinematic constraints considered.

2.1 Phenomenology

The cross section for producing a Higgs boson in proton-proton collisions at $\sqrt{s} =$ 13TeV is about 55 pb and is mediated by four main production mechanisms. The dominant one is via gluon-gluon fusion (ggH), where two gluons interact via a triangle diagram to produce a Higgs, the internal loop is generally mediated by a heavy quark loop (mostly via a top quark loop), since the coupling strength of the

Higgs to any particle is directly proportional to the mass of the particle the Higgs is interacting with. ggH can be used to set indirect constraints on the Yukawa couplings strength between the Higgs and the top quark y_t and has a cross section of about 49 pb at $\sqrt{s} = 13$ TeV. The second largest production mechanism is via Vector Boson Fusion (VBF) with a cross section of 3.7 pb. VBF occurs when the colliding quarks radiate two vector bosons (either Z or W) and these merge to produce a Higgs boson, hence the name. VBF is a particularly interesting production mode because of its final states, which include a Higgs boson on top of two jets with large pseudorapidity η separation and invariant mass. The jet characteristics of VBF are a distinctive signature of VBF production and are very useful in reducing QCD backgrounds. An additional production mechanism for the Higgs is by radiation of a Z or W vector boson, hence shortened to VH, where the radiating vector boson is still present in the final states, and its leptonic decays can be used to suppress background contamination from QCD processes. Lastly, the Higgs field can be produced in conjunction with a $t\bar{t}$ pair ($t\bar{t}H$) or just with a single top qualk (tH). These family of decays can be used to directly measure the y_t Yukawa coupling. All of the above processes are summarised in terms of their dominant Feynman diagrams in Fig. 3. In addition, a summary of the main Higgs production modes in terms of cross sections as a function of the center of mass energy $\sqrt{s}$ is reported in Fig. 4.

2.1.1 Higgs Decay Modes

In terms of Higgs decay, the scalar particle can decay to a large number of final states. These are summarised Table 1 along with Fig. 5 for the low mass region, and in Fig. 6 for the high mass region, where we report the Branching Ratios $\mathcal{B}$ of the

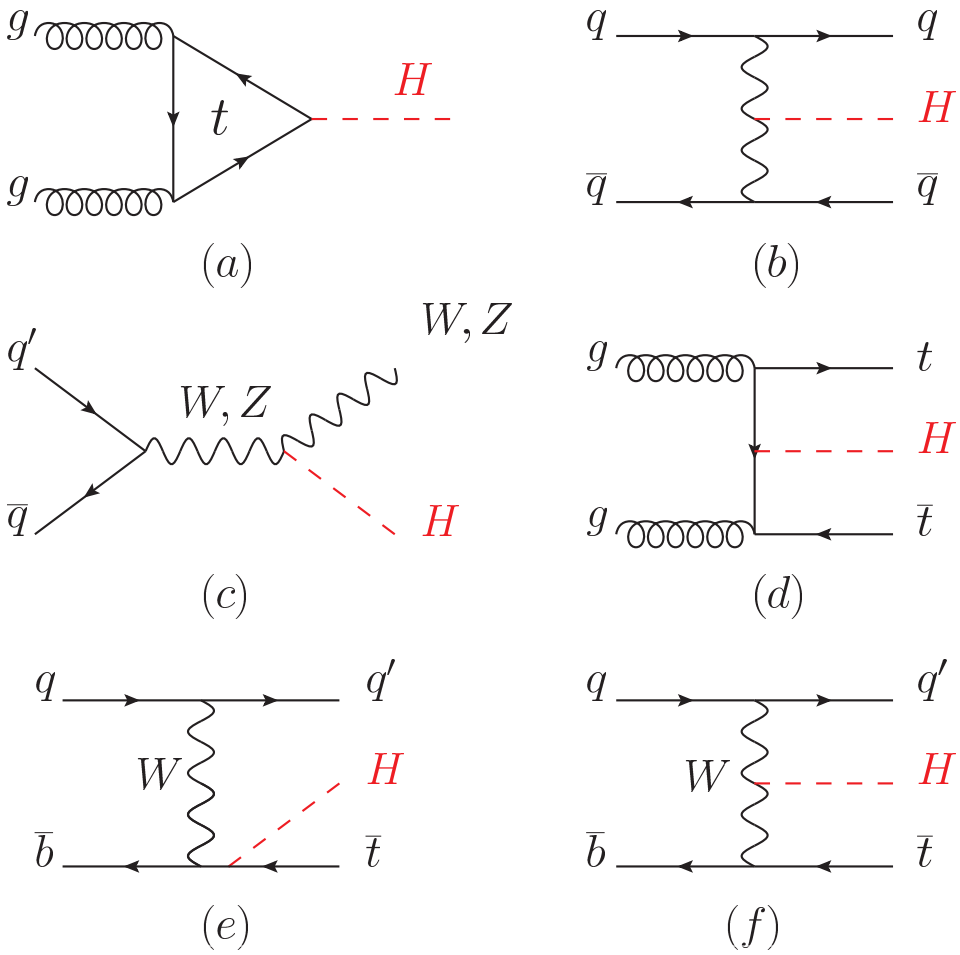


Figure 3: The main leading order Feynman diagrams involved in the production of a single Higgs boson are shown here. We can distinguish a few different production mechanisms such as: (a) gluon gluon fusion, (b) Vector-boson fusion, (c) Higgsstrahlung or associated production of a Higgs along with a gauge boson at tree level from a quark-quark interaction, (d) associated production with a gauge boson (at loop level from a gluon-gluon interaction), (e) associated production with a pair of top quarks (there is a similar diagram for the associated production with a pair of bottom quarks),(f) production in association with a single top quark. Figure from [3].

Higgs decays as a function of the Higgs boson mass. As we expected from our previous discussion, we can notice that the Higgs prefers to decay to heavier particles, when kinematically allowed. So why was it measured instead via some of its rarest

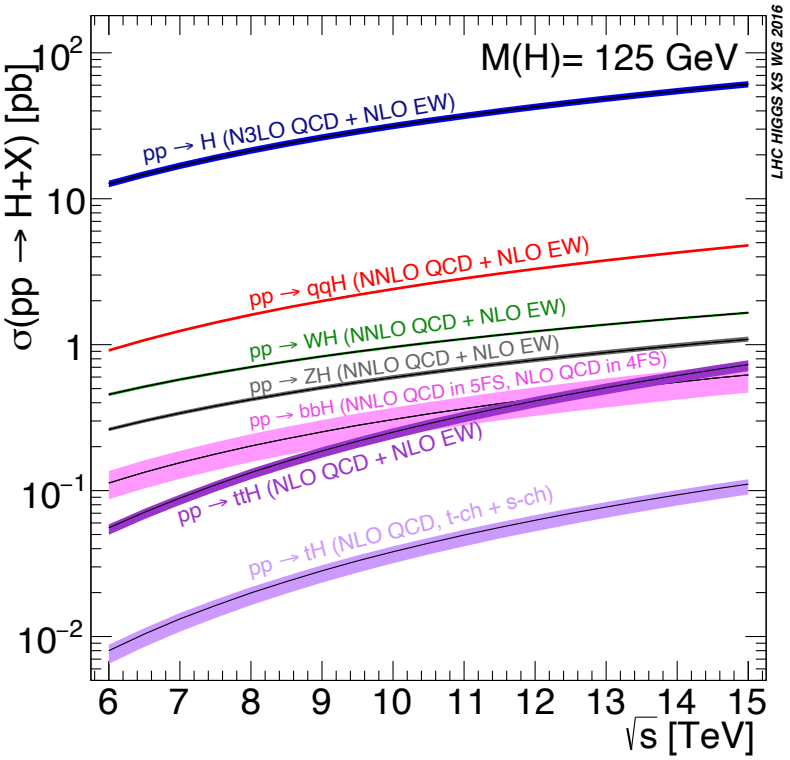


Figure 4: SM Higgs bosons production cross section as a function of the center of mass energies $\sqrt{s}$ and divided by production mechanisms. Figure from [4].

decay modes? There is no easy answer to this question, but we can provide an explanation to leading order. As we bump up the center of mass energy to produce more Higgs bosons, the relative production of other background processes also increases. In order to account for this problem, we tend to reconstruct the Higgs from final states that are rather rare in other background processes and can be precisely measured and reconstructed. This is the case for a pair of photons, since both the CMS and ATLAS detectors have excellent electromagnetic calorimeters (ECAL) that allow good reconstruction efficiencies for photons and electrons; but the reconstruction efficiency is excellent also for muons, where these weakly interacting particles are able to escape most of the detecting elements and are therefore well isolated in the outskirts of the detector, where other particles cannot reach because they have

been previously stopped by heavy layers of brass or iron. For these reasons, at CMS we tend to reconstruct particles from photon, electron or muons candidates over b quarks, a signature that can easily be contaminated by high levels of QCD backgrounds and rather poor jet measurements. In conclusion, the branching ratio of a particular final state is often not the most important factor in determining the analysis strategy, which is influenced by factors such as reconstruction quality of the final state particles and chances of background processes faking the signal.

Main Branching Ratios of the SM Higgs

Decay Channel	Branching ratio
$H \to \gamma\gamma$	2.27×10^{-3}
$H \to ZZ$	2.62×10^{-2}
$H \to W^+W^-$	2.14×10^{-1}
$H \to \tau^+\tau^-$	6.27×10^{-2}
$H \to b\bar{b}$	5.82×10^{-1}
$H \to c\bar{c}$	2.89×10^{-2}
$H \to Z\gamma$	1.53×10^{-3}
$H \to \mu^+\mu^-$	2.18×10^{-4}

Table 1: Decay branching fractions for a SM Higgs of mass m_H = 125.

2.2 Higgs Searches

The Higgs boson was measured for the first time in the history of mankind in 2012 at the Large Hadron Collider (LHC) facilities, by two distinct and independent experiments: CMS [5, 6] and ATLAS [7]. One of the free parameters in the SM Lagrangian is the mass of the Higgs boson, which was measured to be 125 GeV. The above experiments measured the production of a single Higgs from proton-proton collisions at a center of mass energy $\sqrt{s} = 7$ and 8 TeV and using a combination of the following channels: $H \to ZZ^* \to \ell^+\ell^-\ell^+\ell^-$ with $\ell = e, \mu$ and $H \to \gamma\gamma$. The

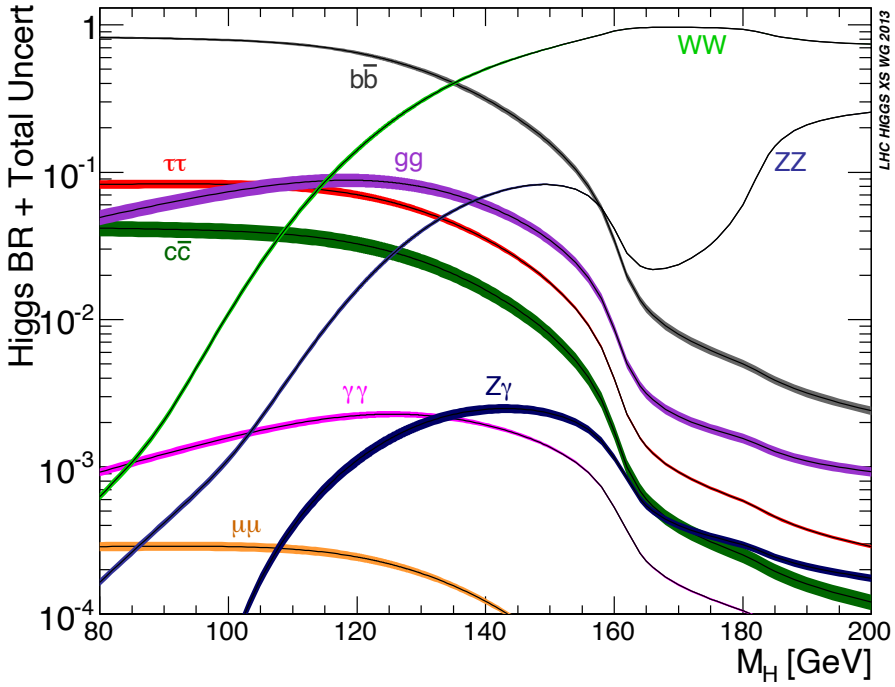


Figure 5: Decay branching fractions of the Standard Model Higgs boson as a function of its mass in the low mass region. Figure from [4].

experimental measurement of the Higgs boson marks one of the biggest successes of the SM, since it completed the spectrum of particles predicted by the SM. The existence of the Higgs boson was later confirmed with additional data during Run 2 (2016 through 2018) at $\sqrt{s} = 13$ TeV. At the current stage of the LHC operations, the Higgs boson has been extensively studied. Its mass precisely measured [8], its spin and parity found to be compatible with the $J^P = 0^+$ SM hypothesis [9], and its interactions with the electro weak gauge bosons observed to be consistent with the SM predictions [10, 11] .

However, there are still a number of precision measurements that can be carried out in order to confirm whether the Higgs that has been observed at the LHC in 2012 is exactly what the SM predicted, or there could be hints of physics beyond the SM (BSM). In particular, one of the key ingredients in identifying the scalar field is by

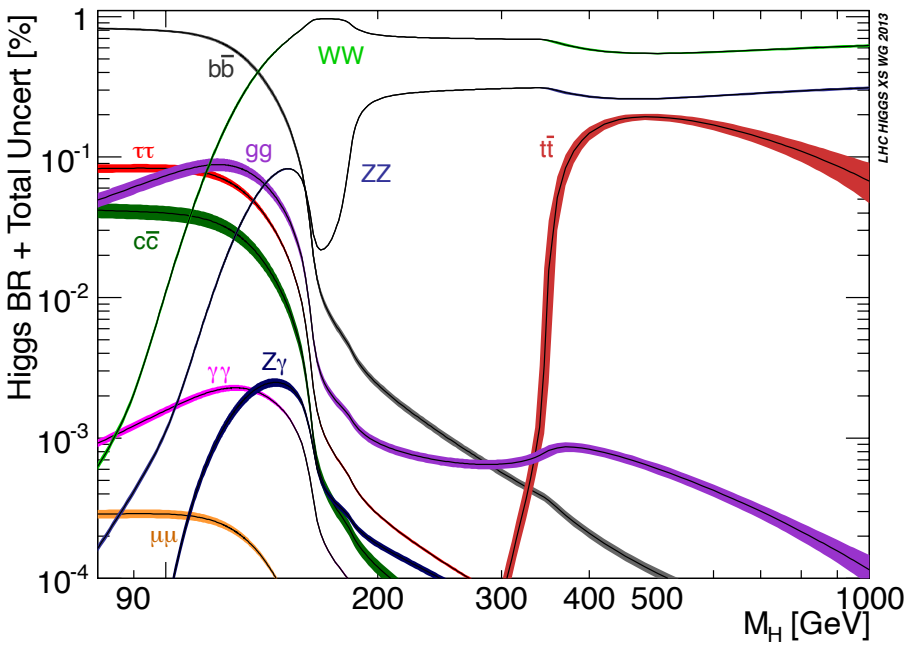


Figure 6: Decay branching fractions of the Standard Model Higgs boson as a function of its mass in the high mass region. Figure from [4].

performing a direct measurement of its self-coupling λ_{HHH}. So far, we can use a number of SM calculations to indirectly constrain λ_{HHH} to be ~ 0.13, but there is no direct measurement of it. A direct measurement of this quantity is of primary importance since $\lambda_{HHH} = m_H^2/2v^2$ directly relates to the shape of the Higgs potential in the SM. This measurement is accessible via a direct measurement of the Higgs pair production, which would therefore provide a fundamental test of the SM description of the Higgs potential.

3 Higgs Boson Pair Production

Double Higgs production is a rare process in the SM, with a cross section of about 30 fb at $\sqrt{s} = 13$ TeV for its dominant production mechanism, roughly 1000 times

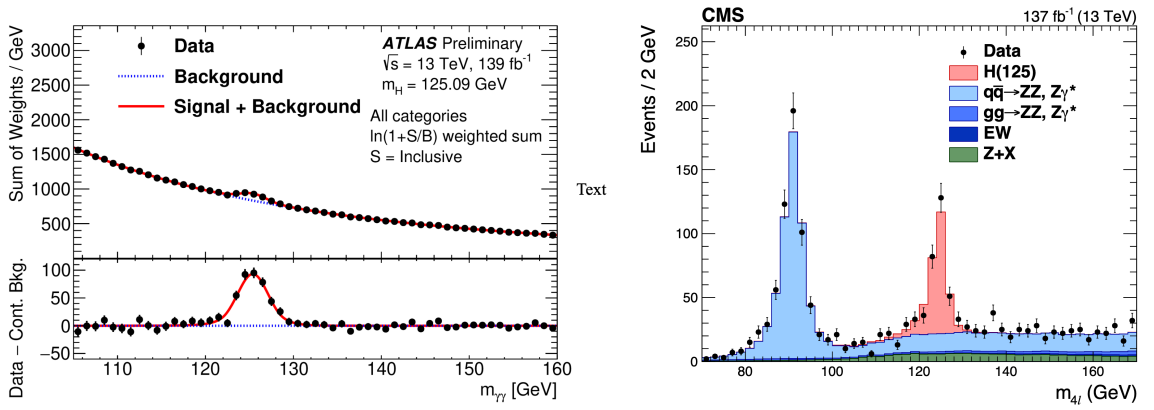


Figure 7: Current status of the Higgs measurements as performed by the ATLAS experiment (left) [12] and the CMS experiment (right) [13]. ATLAS reconstructed the scalar boson from a pair of photon candidates. CMS, instead, sought after the Higgs by focusing on its decays with four leptons. These plots refer to the full Run 2 data.

smaller than for single Higgs production. The sensitivity of the LHC is predicted to be just enough to probe the di-Higgs production after the High Luminosity runs, while it falls short to generate enough statistical evidence for a measurement of triple Higgs production, which has a cross section around 80 ab at $\sqrt{s} = 13$ TeV and it is even a rarer process than di-Higgs production.

Gluon Gluon Fusion, or ggF for short, is the main di-Higgs production mechanism at the LHC. Its cross section is $\sigma_{ggF} = 31.05$ fb and, to leading order calculations (LO), there are two diagrams that contribute to this process: a triangle diagram and a box diagram, as shown in Fig. 9.

The triangle diagram (on the left) is perhaps more interesting since it contains the Higgs trilinear self coupling λ_{HHH}, which is often referred to in terms of its units of the SM value as $\kappa_\lambda = \lambda_{HHH}/\lambda^{SM}_{HHH}$. So for $\kappa_\lambda = 1$ we retain the Standard Model coupling, and for other values of κ_λ we dive into BSM couplings. The box diagrams instead has a large contribution to the cross section, see Fig. 10, but it does not entail

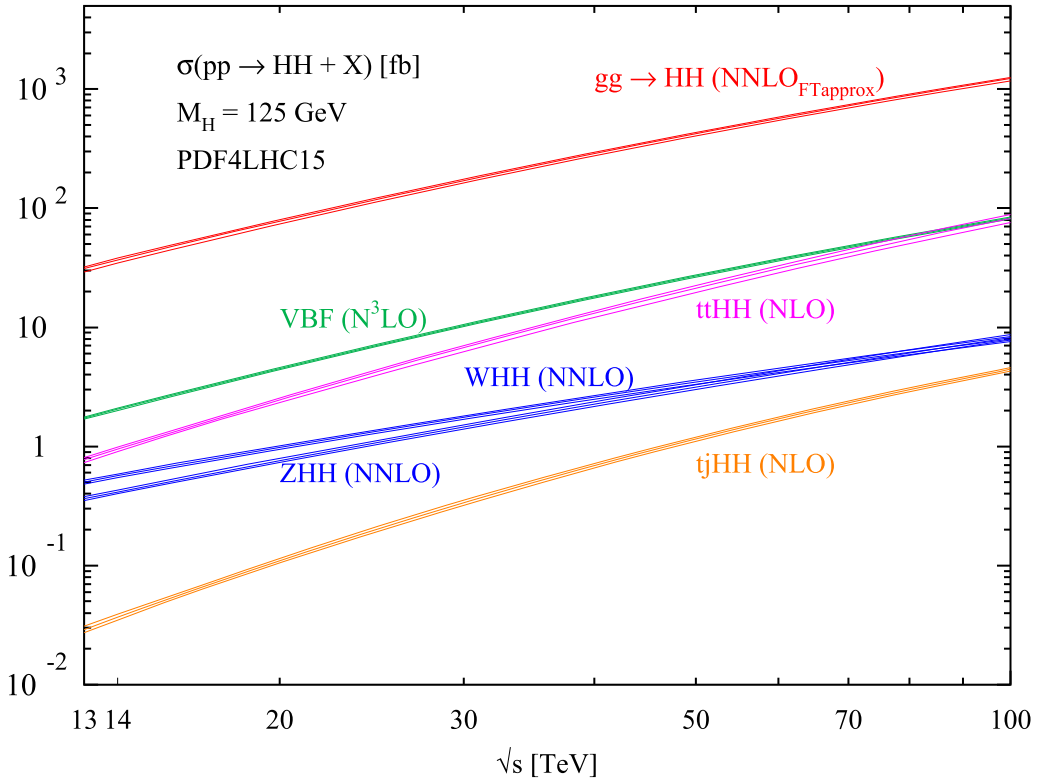


Figure 8: Total cross sections for HH production using NLO corrections in QCD for the six dominant production mechanisms in proton proton colliders. The uncertainties on the parton distribution functions are represented by the thickness of each line [4].

as much interesting physics. In this case, the loop is mediated by top or other heavy quarks and we have to associate production of two Higgs bosons. This diagram therefore is influenced by the Yukawa coupling between the Higgs and the top quark, y_t. For these reasons, measurements of the ggF di-Higgs cross section can be used to constrain both the κ_λ and y_t couplings.

I find interesting to note Fig. 10, which shows the individual contribution to the cross section from the triangle and box diagrams, along with their destructive interference as a function of the di-Higgs mass m_{HH}. As one may expect, there is a turn on effect around $m_{HH} \sim 380$ GeV, which is exactly twice the mass of the top

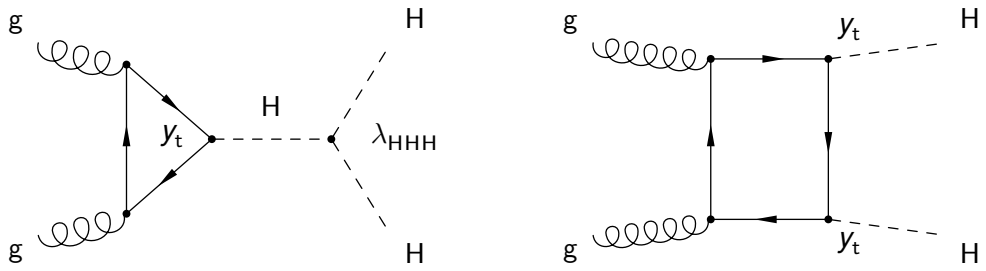


Figure 9: Representative Feynman diagrams for HH production via gluon-gluon fusion: involving the Higgs boson self-coupling (λ_{HHH}) and the top quark Yukawa coupling (κ_t) (left) and the box diagram involving κ_t (right). Figures taken from Ref. [14].

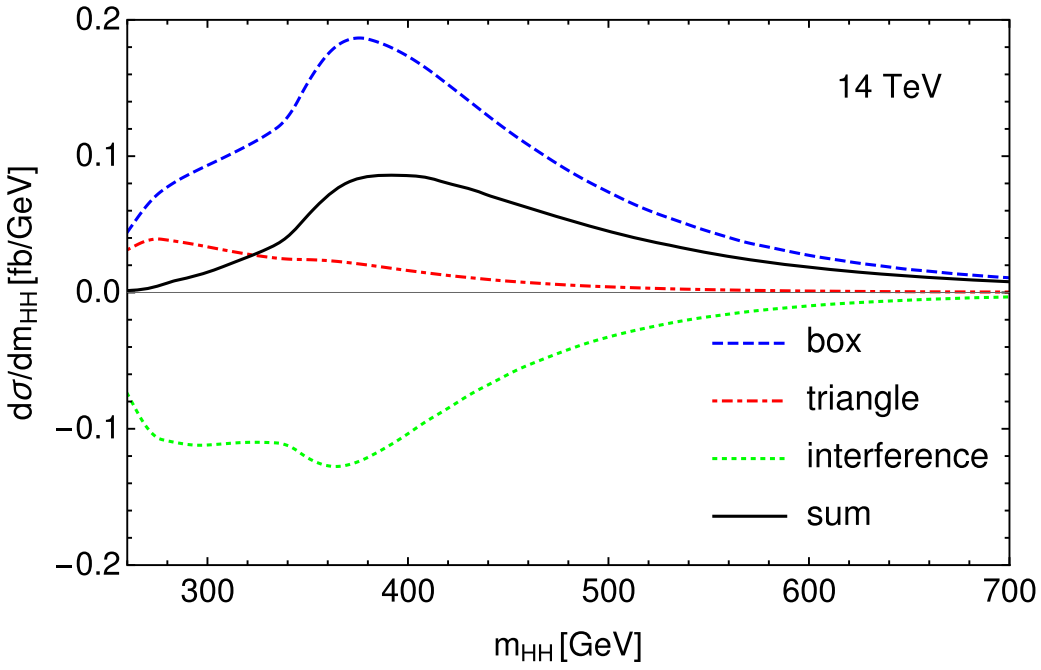


Figure 10: Differential cross section and destructive interference between the triangle and box HH diagrams as a function of the di-Higgs invariant mass distribution, LO accuracy [15].

quark. This enhances the contribution from the box diagram as can be seen from the graph.

Vector Boson Fusion, or VBF, is the second largest production mechanism. Its cross section is around 20 times smaller than ggF with $\sigma_{VBF} = 1.7$ fb also at $\sqrt{s} = 13$ TeV. The LO diagrams for VBF are shown in Fig. 11:

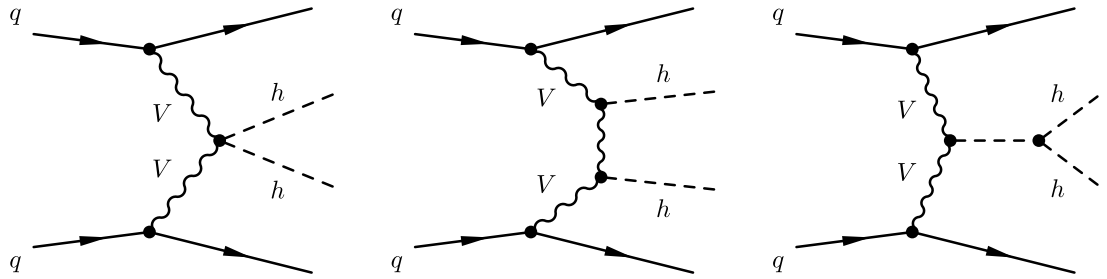


Figure 11: Feynman diagrams dominating the HH production in the VBF mode.

This production mode is of particular interest because of the appearance of the HHVV vertex with coupling strength c_{2V}, which is only accessible via this mode, and the HVV vertex c_V, for which VBF is the dominant production mechanism. The measurements of these couplings against their SM predicted values will provide a fundamental test for the SM and may lead to a new structure of the Electro-Weak Symmetry breaking if found not to be in agreement with the SM prediction. One last remark about VBF and why it has gained popularity in several di-Higgs searches, is the fact that, despite its lower cross section, the two final state jets with large pseudorapidity gap $\Delta\eta$ and invariant mass m_{jj} are a distinctive signature of this process and provide optimal discrimination power versus other types of background.

Vector boson associated production, or VHH, entails the production of a pair of Higgs bosons along with a vector boson V, that can be either a W or a Z. The Higgs is produced via *Higgstrahlung*, the radiation of a Higgs boson from a vector boson V. The LO Feynman diagrams are shown in Fig. 12.The cross section for this process

at $\sqrt{s} = 13$ TeV is rather small, only about $\sigma_{VHH} = 0.87$ fb.

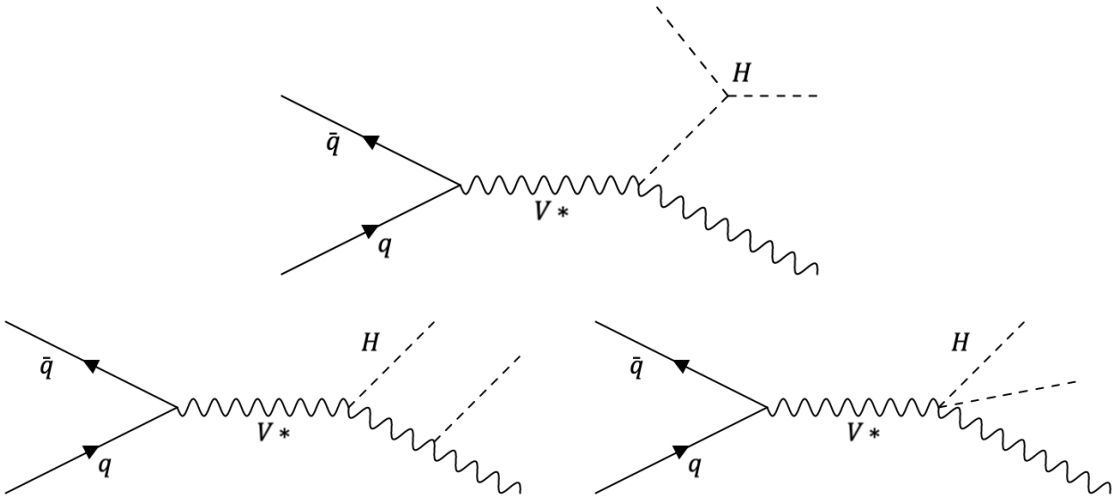


Figure 12: Feynman diagrams dominating the HH production in the VHH mode.

Top quark and anti-quark associated production, or ttHH, generates a top quark anti-quark pair along with two Higgs bosons. Main diagrams for this production mechanism are shown in Fig. 13, with the ttHH cross section being $\sigma_{ttHH} = 0.78$ fb at $\sqrt{s} = 13$ TeV.

Because ggF has the largest contribution to the overall HH cross section, our analysis will focus on this production mode in order to extract constraints about the Higgs potential. It is worth noting that, given the rather small SM cross section for HH processes, this type of search is particularly sensitive to BSM models, which could enhance the di-Higgs cross section and change the kinematics of the properties of the final state products. It is therefore in our interest to investigate more this production mechanism in order to compare the observed results versus the SM predictions and other potential BSM scenarios.

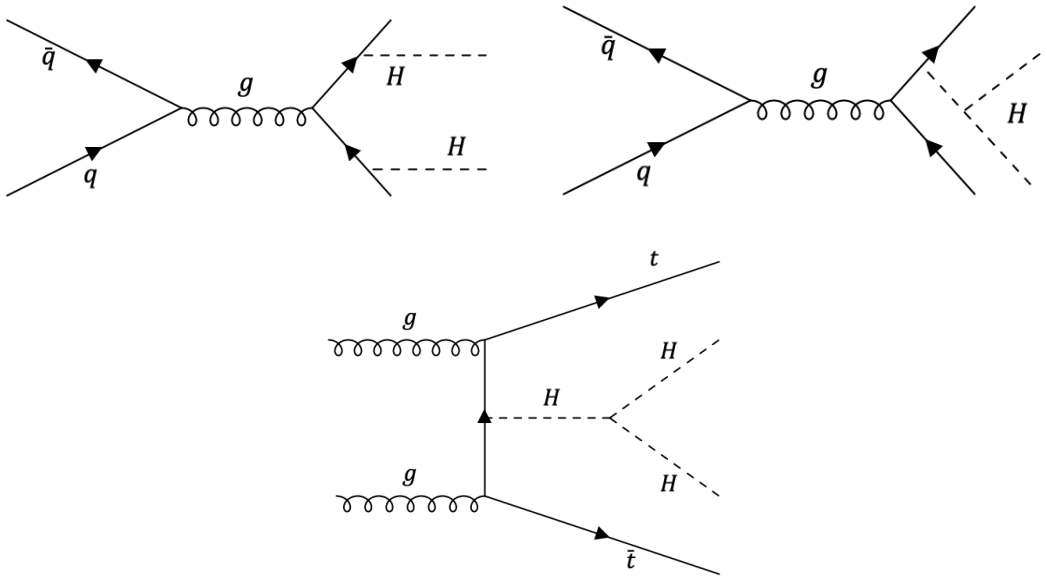


Figure 13: Feynman diagrams dominating the HH production in the ttHH mode.

3.1 BSM

The Standard Model of particle physics has been proven to be extremely accurate in its description of fundamental particles over the years. Its latest triumph being the discovery of the Higgs boson in 2012. However, there are still a number of phenomena that are not accounted for by the SM, such as the nature of dark matter, or the dominance of matter over anti-matter in our universe or, perhaps most importantly, there is no field description of gravity. On a more subtle level, there is no clear explanation as to why the masses of fermions should span such a large range in the SM, going from massless (or close to) neutrinos and electrons with a mass of half of an MeV to the top quark with a mass of about 170 000 MeV. For these reasons, it may seem that the SM is just the low energy approximation of a more comprehensive theory of the universe which could in principle answer all of the questions that the SM cannot.

The cross section and kinematics of HH production can be significantly affected in many theories of physics beyond the SM, as shown in Fig. 14 and Fig. 15.

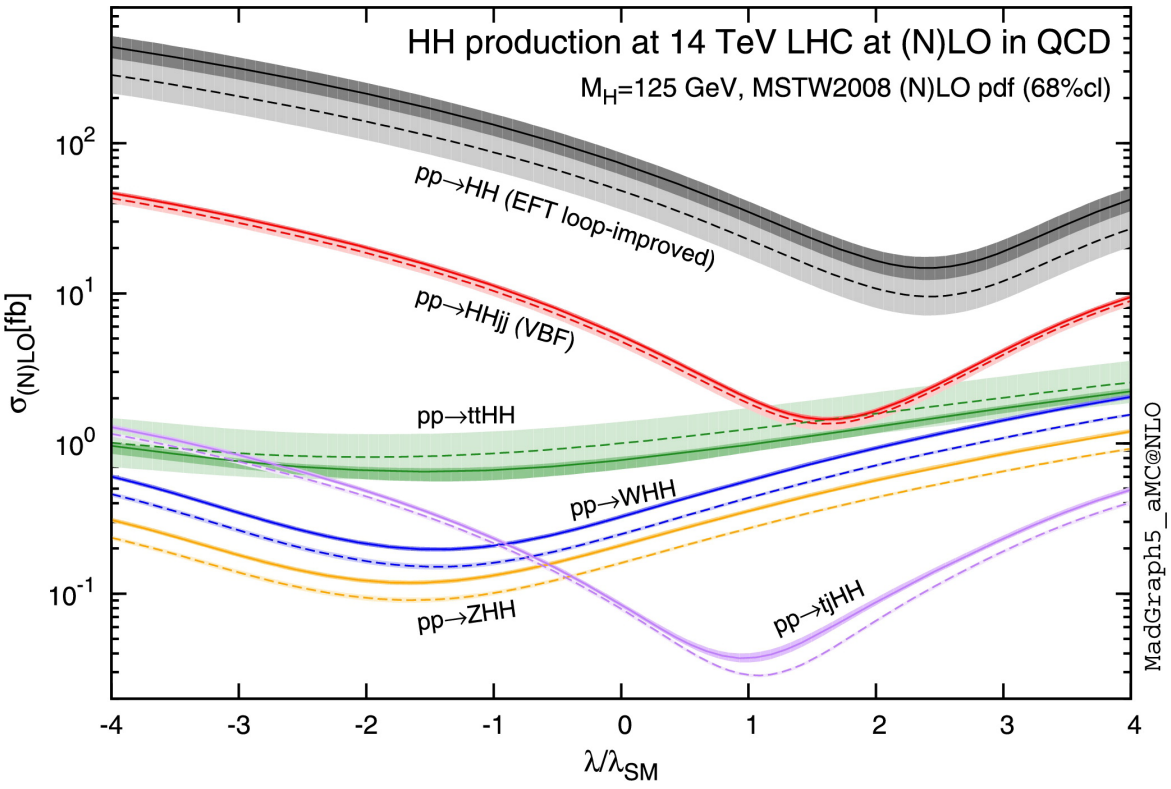


Figure 14: The cross sections for HH production channels [4] at the s=14 TeV LHC can be expressed as a function of the self-interaction coupling λ_{HHH}, both at leading order (LO) and next-to-leading order (NLO) in QCD. The results are depicted by dashed (solid) lines and light-(dark-)colour bands, representing the LO (NLO) values with scale and PDF uncertainties added linearly. The standard model (SM) values for the cross sections can be obtained when $\kappa_\lambda = \lambda/\lambda_{SM} = 1$.

The lack of discovery of BSM particles at the LHC may suggest that these particles, if they exist, have masses large enough to avoid direct production at current LHC energies and motivates the use of an effective field theory (EFT) framework to parametrize the contributions of new physics in terms of higher-dimensional operators suppressed by a large mass scale Λ. In this analysis, BSM contributions

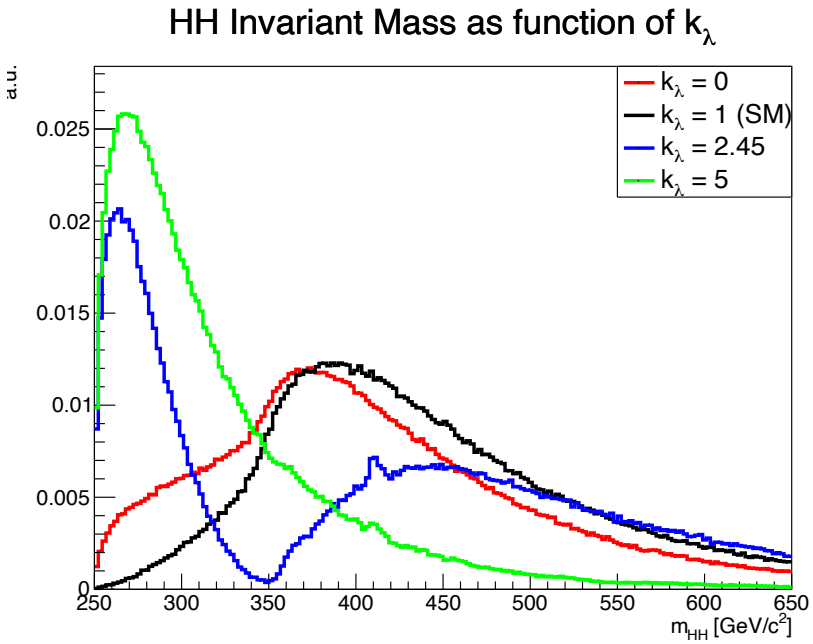


Figure 15: Normalised m_{HH} distribution for the different BSM scenarios. Depending on the value of the κ_λ coupling, the m_{HH} spectrum varies considerably both in shape (as shown in figure) and normalisation.

are parametrized through an effective Lagrangian that extends that of the SM with dimension-6 operators [16, 17], in the form:

$$\mathcal{L} = \mathcal{L}_{SM} + \sum_i \frac{c_i}{\Lambda^2} O_i^6 + ... \tag{50}$$

where c_i are the Wilson coefficients. In this way, any BSM model can be matched to the EFT order by order so to derive an expression for the Wilson coefficients as a function of the BSM model parameters, and BSM effects beyond the experimental reach of the LHC can be probed by deriving constraints that are independent of the model selected.

This effective Lagrangian parametrizes BSM effects in terms of five couplings involving the Higgs boson. Two are present in the SM and expressed as ratios to their

SM values: the self-coupling (κ_λ) and the top quark Yukawa coupling (κ_t) as shown in Fig. 9; while three are contact interactions not present in the SM: between two Higgs bosons and two gluons (c_{2g}), between one Higgs boson and two gluons (c_g) and between two Higgs bosons and two top quarks (c_2). Diagrams involving the non-SM couplings are shown in Fig. 16.

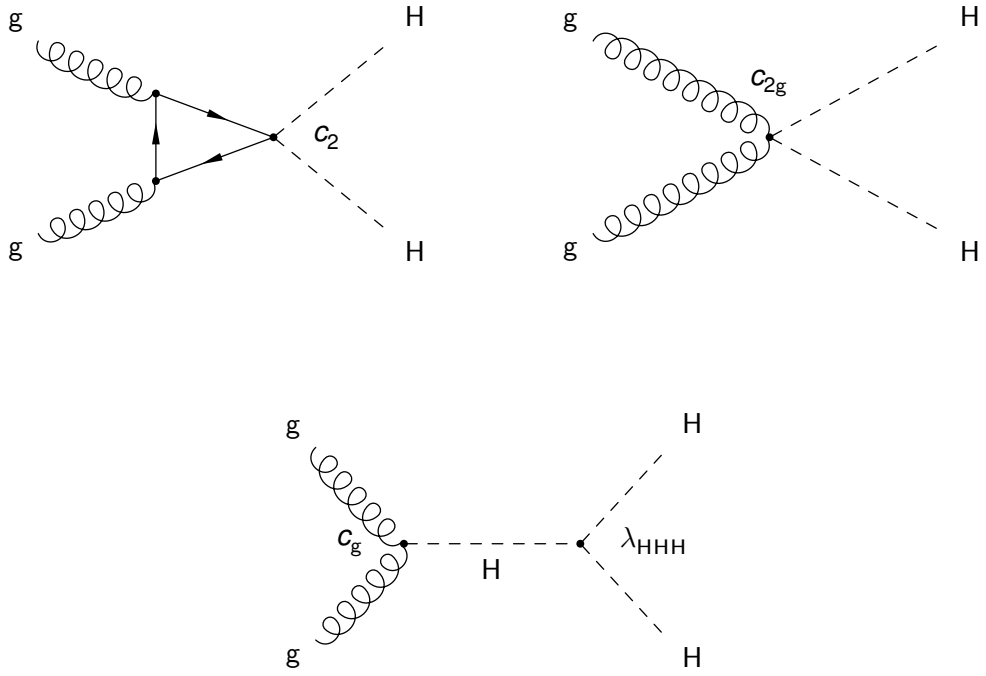


Figure 16: Feynman diagrams contributing to HH production at LO in the EFT framework [17], involving the contact interactions between: two Higgs bosons and two top quarks (c_2) (left), two Higgs bosons and two gluons (c_{2g}) (middle), and one Higgs boson and two gluons (c_g) (right). Figures taken from Ref. [14].

3.1.1 Monte Carlo Generation of BSM Samples

There have been many CMS recommendations on how to produce BMS di-Higgs samples, the latest ones that I am aware of use a re-weighting procedure that I will now go on to describe. Let us start by reviewing the ggF production of HH, this

is mediated by the two leading diagrams in Fig. 9. The cross section of ggF is proportional to the amplitude squared ($|A|^2$) so we have an equation of the form:

$$\begin{aligned} A &= \kappa_t \kappa_\lambda T + {\kappa_t}^2 B \\ \sigma(\kappa_t, \kappa_\lambda) \sim |A|^2 &= {\kappa_t}^2 {\kappa_\lambda}^2 T^2 + {\kappa_t}^4 B^2 + {\kappa_t}^3 \kappa_\lambda |T^* B + B^* T| \\ &= {\kappa_t}^2 {\kappa_\lambda}^2 t + {\kappa_t}^4 b + {\kappa_t}^3 \kappa_\lambda i \end{aligned}$$

where T stands for the triangle diagram amplitude, B for the box diagram one, t (b) for T^2 (B^2) and i for the interference between the two diagrams. At this point we should note that this formula is valid at any order of the calculation, as we have terms of $|A|^2$ that scale with the same powers of the couplings, and this still holds differentially $\partial\sigma/\partial x$ when using $\partial t/\partial x$, $\partial b/\partial x$ and $\partial i/\partial x$ as inputs. If we now introduce vectors to generalise this discussion we have:

$$\begin{aligned} \mathbf{c} &= \left(\kappa_t^2 \, \kappa_\lambda^2 \, , \, \kappa_t^4 \, , \, \kappa_\lambda \, \kappa_t^3 \right) \\ \mathbf{v} &= \left(t \, , \, b \, , \, i \right) \\ \sigma\left(\kappa_\lambda \, , \, \kappa_t \right) &= \mathbf{c} \cdot \mathbf{v} \end{aligned}$$

with $\mathbf{c}$ vector of couplings and $\mathbf{v}$ vector of components. We then consider three different κ_λ and κ_t scenarios for which we know their cross section, so that we can write:

$$\begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_2 \end{pmatrix} = \begin{pmatrix} c_1^1 & c_1^2 & c_1^3 \\ c_2^1 & c_2^2 & c_2^3 \\ c_3^1 & c_3^2 & c_3^3 \end{pmatrix} \begin{pmatrix} t \\ b \\ i \end{pmatrix} \tag{51}$$

or in short

$$\boldsymbol{\sigma}\left(\kappa_\lambda, \kappa_t\right) = \mathbf{C}\,\mathbf{v} \tag{52}$$

and by solving the three equations we find the values for t, b and i. Once we have these values, we can invert the matrix $\mathbf{C}$ and express the cross section for any arbitrary values of κ_λ and κ_t using the formula:

$$\sigma(\kappa_\lambda, \kappa_t) = \mathbf{c}^T(\kappa_\lambda, \kappa_t)\,\mathbf{C}^{-1}\,\boldsymbol{\sigma} \tag{53}$$

where $\mathbf{c}^T(\kappa_\lambda, \kappa_t)$ describes the couplings to model, $\mathbf{C}^{-1}$ is known from the couplings of the inputs samples and $\boldsymbol{\sigma}$ is the cross section vector of the samples that are scaled and summed. The formula in Eq. 53 is particularly useful since it allows us to generate any κ_λ and κ_t sample we desire just by reweighting any three existing samples. Practically, we prefer to use the triplet of $\kappa_\lambda = 1\,, 2.45\,,$ and 5 samples, since other κ_λ points may lead to large re-weighting factors which can result in numerical instabilities and errors if the factors do not counterbalance each other exactly.

This analysis places direct constraints on κ_λ, under the assumption that all other couplings of the Higgs boson have values as predicted by the SM, and places limits

on the cross sections for twelve scenarios of non-SM combinations of all five Higgs boson couplings (κ_λ, κ_t, c_{2g}, c_g, c_2), following the scenarios proposed by Ref. [18]. The twelve effective field theory (EFT) scenarios are generated using a kinematic reweighting of the available κ_λ samples (since these are at NLO accuracy) based on the di-Higgs invariant mass and opening angle. The couplings for each of the twelve EFT benchmarks are summarised in Table 2.

Table 2: Parameter values of nonresonant BSM benchmark hypotheses. The first column correspond to the SM sample, while the next 12 correspond to the benchmark hypotheses identified using the method from Ref. [18].

	SM	1	2	3	4	5	6	7	8	9	10	11	12
κ_λ	1.0	7.5	1.0	1.0	-3.5	1.0	2.4	5.0	15.0	1.0	10.0	2.4	15.0
κ_t	1.0	1.0	1.0	1.0	1.5	1.0	1.0	1.0	1.0	1.0	1.5	1.0	1.0
c_2	0.0	-1.0	0.5	-1.5	-3.0	0.0	0.0	0.0	0.0	1.0	-1.0	0.0	1.0
c_g	0.0	0.0	-0.8	0.0	0.0	0.8	0.2	0.2	-1.0	-0.6	0.0	1.0	0.0
c_{2g}	0.0	0.0	0.6	-0.8	0.0	-1.0	-0.2	-0.2	1.0	0.6	0.0	-1.0	0.0

Such points have been selected to represent the main kinematic observables of the HH processes over the full phase space.

3.2 HH Measurements at the LHC

Because of their connection to BSM theories, HH searches have been the object of several investigations at the Large Hadron Collider, both by the CMS and the ATLAS experiment. As the di-Higgs ggF cross section is rather small, the HH final states with the largest branching ratios, see Fig. 17, have been examined first, such as H$\rightarrow b\bar{b}$. There is however a tradeoff between large branching fractions and contamination from backgrounds presenting similar final states as those targeted. Every decay mode has its own experimental challenges that will be briefly described in the next paragraphs.

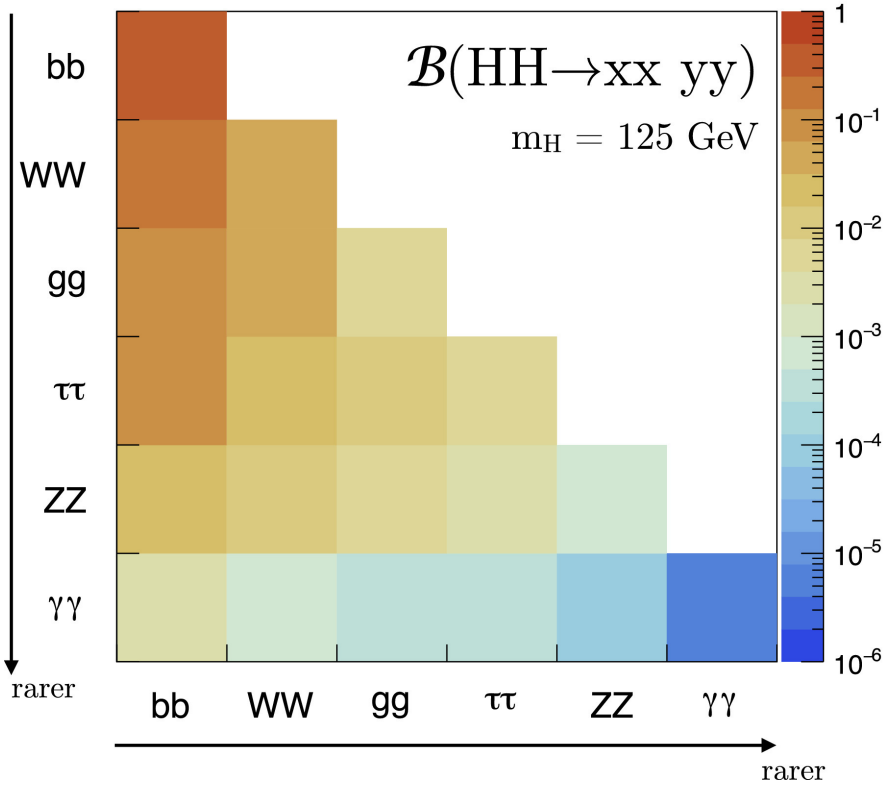


Figure 17: Summary of the main HH branching ratios as predicted by the SM.

To this day, the main HH searches have targeted:

HH→ $b\bar{b}b\bar{b}$ presents the largest branching fraction ($\mathcal{B} = 33.7\%$) but it also suffers from high levels of background contamination, especially those coming from QCD activity, which are notoriously difficult to model using Monte Carlo methods, because of the infrared divergence of the QCD theory. Let us also note that b-tagging of jets is fundamental in this type of analysis, but this is too slow of a process for the trigger logic, so that the trigger efficiency for this channel is lower than for other final states. The best upper limit on the ggF cross section has been set by the CMS experiment as $\sigma_{ggF}(HH) < 3.5\times$ SM prediction at the 95% confidence level (CL)

[19].

HH$\rightarrow$ $b\bar{b}\tau^+\tau^-$ retains a large branching ratio ($\mathcal{B} \sim 7\%$) but the presence of taus, hence neutrinos, does not allow the full reconstruction of the final states. This channel also suffers from QCD contamination where a $t\bar{t}$ pair is produced and each top quark decays into a W plus a b-jet, which would mimic the final state signature. The best limit on the ggF cross section so far is an expected limit, as the analysis hasn't been unblinded yet, and it is $\sigma_{ggF}(HH) < 4.8\times$ SM prediction at the 95% CL [20].

HH$\rightarrow$ $b\bar{b}VV$, where V can be either a W or a Z. Despite having a reasonable branching ratio (2.7%), this channel is swamped by $t\bar{t}$ irreducible background, hence hard to isolate [21].

HH$\rightarrow$ $b\bar{b}\gamma\gamma$ is a very interesting final state. Its branching ratio is not huge, $\mathcal{B} \sim$ 0.26%, however, the $H \rightarrow \gamma\gamma$ offers a clean signature with very high resolution that allows to suppress a large fraction of background. In this analysis a 2D fit is performed to extract the signal strength, where both the di-photon invariant mass and the di-jet one are fitted, for a joint limit on the ggF and VBF cross section as $\sigma_{ggF+VBF}(HH) < 7.7$ times the SM prediction [22].

This bring us to the end of the introductory section to the Standard Model of particle physics and we will soon dive into the more experimental aspects of particle physics, where we reconstruct particles and perform cross section measurements of different processes with the magical and unique machinery that is the CMS detector at the LHC.

Part II

Experimental Apparatus

4 The Large Hadron Collider

The Large Hadron Collider (LHC) is the most powerful particle accelerator ever built, and is operated by CERN (European Organization for Nuclear Research). It is located at the border between Switzerland and France, near Geneva, and it consists of a 26.7 km long tunnel about 100 m underground in which beams of protons are accelerated to the speed of light and eventually collide. The main goal of the LHC is to produce new heavy particles and observe rare processes by providing interactions at a high center-of-mass energy and a high number of collisions. The design of the machine and beam structure are driven by these goals. The LHC has four interaction points along its circumference where the proton beams collide and where the four detectors are placed, in order to analyse and examine each particle collision that takes place. One of the four detector is the Compact Muon Solenoid (CMS), which is the machinery that has been utilised to collect the data presented in this thesis.

The Large Hadron Collider (LHC) was built to investigate the Standard Model and the phenomenon of spontaneous symmetry breaking, with particular emphasis on the search for the Higgs boson existence, along with the discovery of BSM phenomena. Initially, the LHC was designed to only perform proton-proton collisions, but later heavy ion collisions were also included for the study of quarks and gluons'

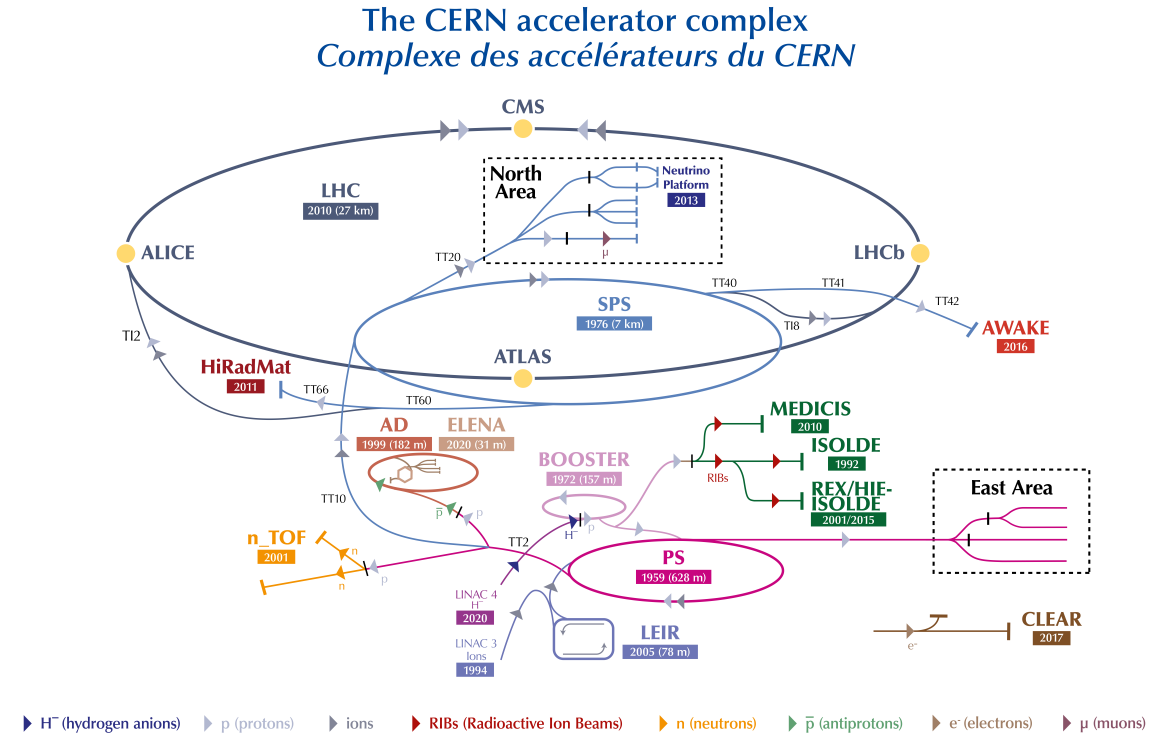


Figure 18: Schematic diagrams of the LHC complex [23]. The proton journey starts in the LINAC, PSB, PS and the SPS where they are gradually accelerated. Subsequently, they are injected and collimated into the LHC main ring where they form two bunches traveling in opposite directions and colliding at the four interactions points where CERN main experiments are located (ATLAS, CMS, LHCb and ALICE).

collective behavior (not covered in this thesis though). The LHC's scientific program is expected to span several decades; the first stable beam running conditions were attained in 2009 (Run 1) while the current data-taking phase is set to end in 2024 (Run 3) and will be followed by a major upgrade with the High Luminosity (HL) LHC phase. The LHC is designed to provide proton-proton collisions at a center-of-mass energy $\sqrt{s}$ of up to 14 TeV, with plans to deliver approximately 300

fb-1 by 2024, and 3000 fb-1 by the end of the HL-LHC period.

The CERN accelerator complex accelerates protons in multiple stages as shown in Fig. 18. Protons are first accelerated to an energy of 50 MeV by a linear accelerator, then passed to the proton synchrotron booster (PSB), proton synchrotron (PS), and super proton synchrotron (SPS) where they are accelerated to energies of 1.4 GeV, 25 GeV, and 450 GeV respectively. The protons are then injected into the LHC in bunches of approximately 10^{11} protons, with a nominal spacing of 25 ns between bunches, corresponding to a collision frequency of approximately 40 MHz, and around 2800 bunches per proton beam. At the LHC, each particle beam is accelerated to an energy of 6.5 TeV. The two proton beams travel in opposite directions in the LHC tunnel and collide at the four interaction points with a center-of-mass energy of $\sqrt{s} = 13$ TeV. To bend the two proton beams in opposite directions, magnets with opposite magnetic fields are installed around the beam pipe. The LHC consists of two concentric counter-rotating rings in a common cryogenic structure, where superconducting magnets are placed. Dipole magnets are used to bend the beam, quadrupole magnets are used to collimate the beams near the interaction points, and magnets of higher multipole orders control the LHC optics. The field strength required from the magnets is dictated by the target center of mass energy of 14 TeV, along with the fixed circumference of the previous LEP tunnel, which results to an enormous magnetic field of $B = 8.3$ T. This field is achieved using superconducting magnets that need to be cooled down to an operating temperature of 1.9 K, that can be achieved by using about 96 tonnes of superfluid helium He-4.

The number of events produced during proton-proton collisions at the LHC inter-

action points is related to the cross section of a given process. However, the total number of produced events N_{evt} is proportional not only to the cross section σ of a process but also to the instantaneous luminosity L of the accelerator, as described by the following equation:

$$N_{evt} = \sigma \int L \, dt \tag{54}$$

The integrated luminosity, which is the sum of the instantaneous luminosity over time, is crucial for observing rare events, as it results in more events being produced for a given cross section. Assuming that the proton beams at the LHC are identical, the machine luminosity can be computed using the formula:

$$L = \frac{N_b^2 n_b f_{LHC} \gamma}{4\pi \epsilon_n \beta^*} F \tag{55}$$

with N_b being the number of protons in a bunch, n_b the number of collimated bunches in the proton beam, f_{LHC} the number of revolutions per second, γ the relativistic Lorentz factor and F is a geometric factor that accounts for the luminosity reduction due to the cross angle between the two proton beams θ_c at the interaction point and is computed according to:

$$F = \left[1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*} \right)^2 \right]^{-1/2} \tag{56}$$

where σ_z is the bunch length and σ^* is transverse bunch extension. There are two additional parameters, ϵ_n and β^* which are related to the transverse size of the proton beams. Because proton-proton collisions and other small losses deplete the

beams, the instantaneous luminosity of the LHC decreases over time, or until new proton beams are injected into the accelerator. On average, a proton beam lasts for about 8 hours before it is dumped and replaced with a higher occupancy one. The average number of simultaneous interactions per bunch crossing, or pile-up (PU), can be computed as:

$$< PU >= \frac{L \cdot \sigma_{pp}^{inel}}{n_b \cdot f_{LHC}}, \tag{57}$$

where σ_{pp}^{inel} refers to the inelastic cross section for proton-proton collisions and is about 69 mb for a center of mass energy of 13 TeV. As we can see from the equation above, a large number of bunches is desirable to keep the PU low. This is an important factor to keep in mind when designing the running conditions of the LHC, which can reaches a maximum instantaneous luminosity of $L = 10^{34}$ cm^{-2} s^{-1} with an average pile-up of 22.

The LHC generates enormous amounts of proton-proton collision data, which is recorded by various detectors. The main experiments at the LHC are located at the four collision points and are known as ATLAS, CMS, LHCb, and ALICE. The CMS and ATLAS experiments use detectors that have a wide range of applications in physics research. These detectors are designed to be versatile so that any new discoveries made at the LHC can be independently verified. In contrast, the ALICE and LHCb experiments have specialized detectors that are specifically tailored to study specific physics phenomena.

4.1 The Compact Muon Solenoid

The CMS detector is located at one of the four collision points in the LHC and is designed to detect particles resulting from LHC collisions and measure their properties with high precision. As shown in Fig. 19, the CMS detector [24] has a cylindrical shape with an overall length of 29 meters, a diameter of 15 meters, and a weight of about 14,000 tonnes. The detector is divided into a central barrel and two forward endcaps; each component is composed of several sub-systems placed concentrically around the interaction point, covering most of the solid angle around it. Each sub-system, or sub-detector, is specialized in detecting and reconstructing different types of particles.

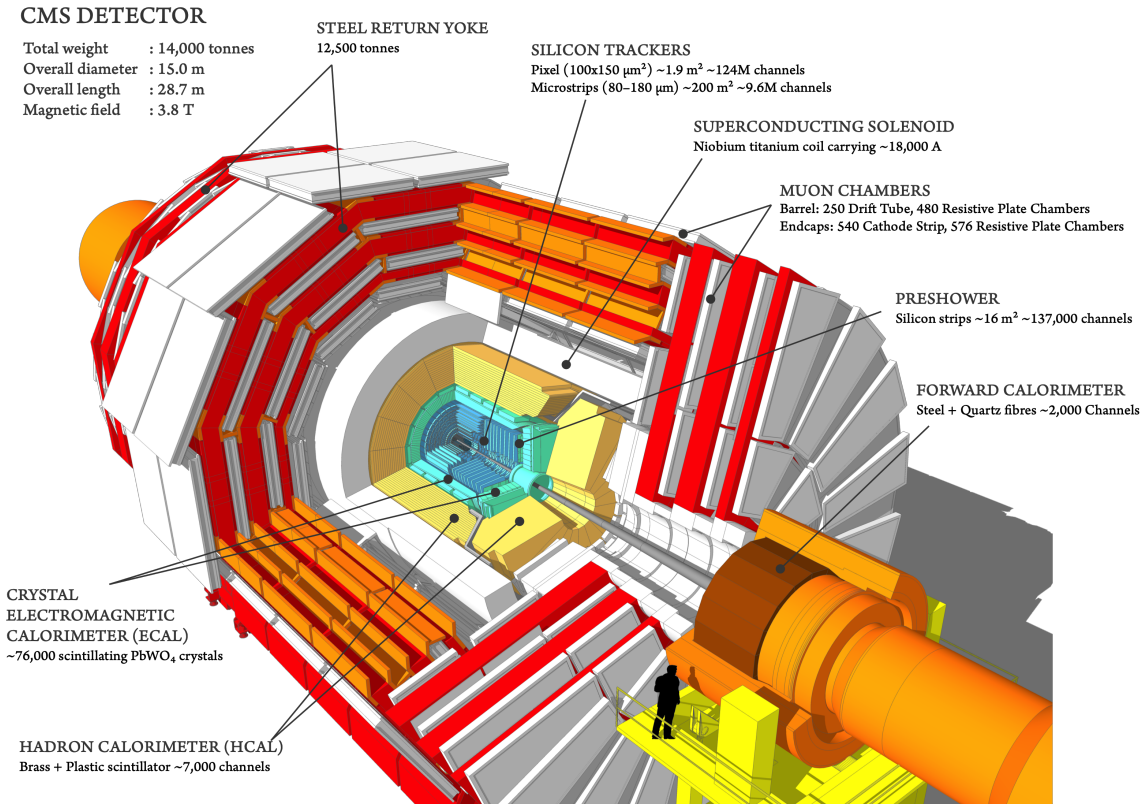


Figure 19: Schematic diagrams of the CMS detector [24].

The CMS detector is built around a superconducting solenoidal magnet that generates a magnetic field of 3.8 T. This strong magnetic field allows for accurate measurement of the momentum of charged particles based on their curved trajectories. The detector includes a pixel and tracker system located inside the magnet, surrounded by a lead tungstate crystal electromagnetic calorimeter (ECAL) and a brass-scintillator hadronic calorimeter (HCAL). Outside of the magnet is the steel return yoke with an embedded muon detection system that includes Drift Tubes (DT), Cathode Strip Chambers (CSC), and Resistive Plate Chambers (RPC). More detailed information about the CMS detector and its components can be found in Ref. [25].

4.2 Coordinate System

The CMS experiment utilises a coordinate system that is centered around the center of the detector and is right-handed. The x-axis is directed towards the center of the LHC ring, the y-axis is perpendicular to the LHC plane and points upwards, and the z-axis points in the anticlockwise direction of the beam. The position of particles is measured using both longitudinal and transverse coordinates along the z-axis and in the x-y plane, respectively. The polar angle, measured from the z-axis, is used to calculate the pseudorapidity which is denoted as $\eta = -\ln[\tan(\theta/2)]$. This is preferred over the polar angle as it is roughly constant for the production of low-momentum hadrons in proton-proton collisions and it is also Lorentz-invariant for longitudinal boosts if the masses of particles are much lower than their energies. The LHC collides protons which are not elementary particles, therefore, the primary focus of observation is the momentum of particles in the transverse x-y plane, referred to as transverse momentum p_T.

An illustration of the CMS Cartesian and polar coordinate systems is shown in Fig. 20.

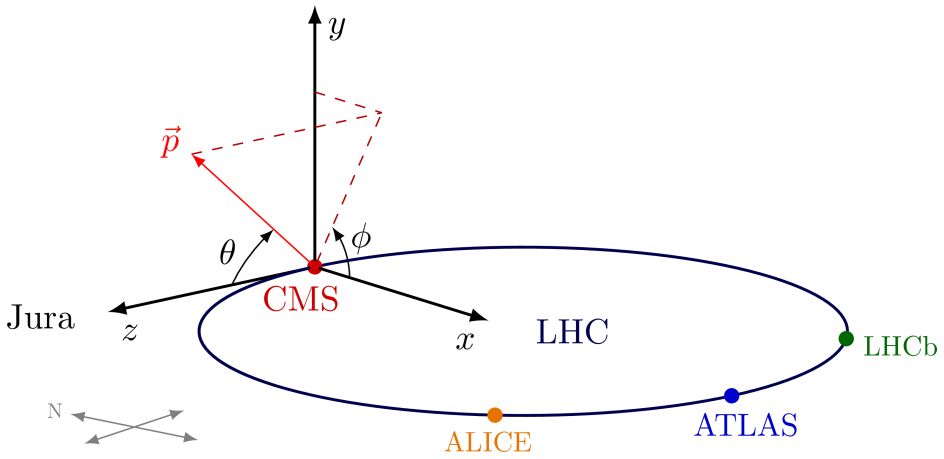


Figure 20: Coordinate system of the CMS detector.

4.3 Solenoid Magnet

The superconducting solenoidal magnet that gracefully names the CMS experiment measures 12 meters in length and 6 meters in diameter and it is used to generate the appropriate magnetic field needed for particle reconstruction [26]. The niobium-titanium solenoid is kept at an extremely low temperature of 4.5 K in order to maintain its superconductivity. The magnet provides a nearly uniform magnetic field of 3.8 T, which is necessary to accurately measure the momentum of high energy charged particles. The magnet is located outside of the tracker and calorimeter subsystems to prevent energy loss from particles showering in the coil of the magnet.

4.4 Tracking System

The tracking system is positioned directly around the beam interaction point and is used to measure the transverse momentum of charged particles from their reconstructed trajectory, as well as to identify the primary vertex (PV) of the proton

collision and secondary vertices (SV) associated with the decays of long-lived particles such as b hadrons or τ leptons [27]. It measures 5.6 meters in length and 2.4 meters in diameter. The design of this system was driven by the need to cope with the following constraints: high particle fluxes, association of signals to the correct bunch crossing, tolerance to high levels of radiation, and minimal amount of material to avoid particle energy losses and multiple scattering [28]. To meet all of these challenges, silicon sensors with diverse granularity were employed, and the tracking system was split into two sub-detectors: the pixel detector and the strip detector. A longitudinal view of the tracking system with the individual subsystems is shown in Fig. 21.

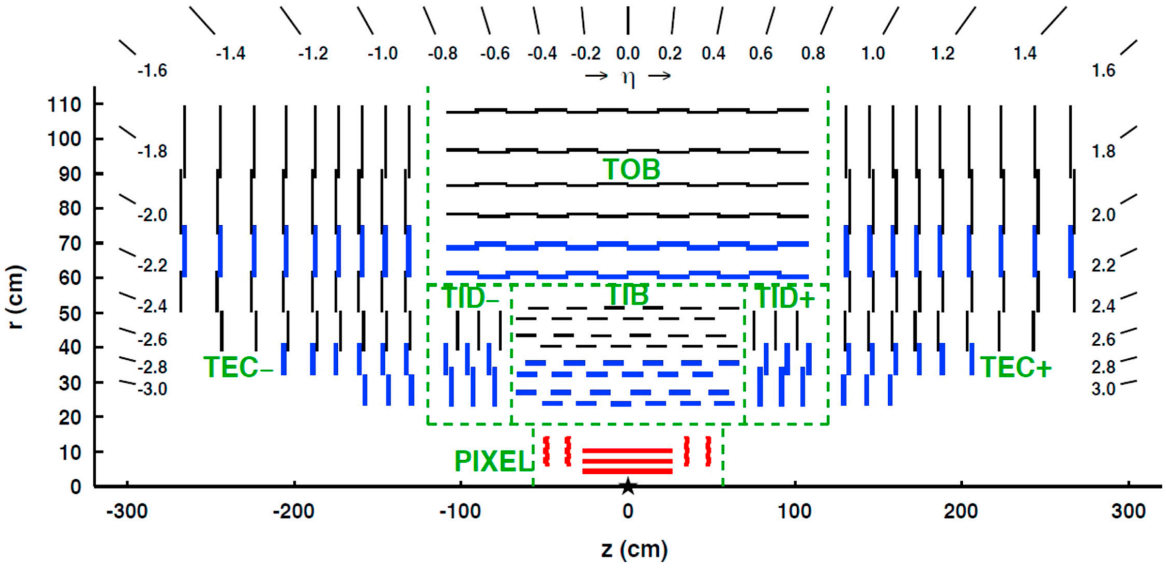


Figure 21: Longitudinal view of a slice of the CMS tracker system and sub-detectors in the *rz* plane. Red-colored modules indicate pixels, black thin lines represent single-sided strip modules, and blue thick lines represent strip stereo modules. In these modules, one is rotated at a stereo angle to enable the reconstruction of hit positions in 3D. The strip tracker detector comprises the tracker inner barrel (TIB) and the tracker inner disks (TID), as well as the outermost tracker outer barrel (TOB) and the tracker endcaps (TEC) [28].

The pixel detector is made up of silicon sensors and is located in the innermost region of the tracker, where the particle flux is highest. In order to handle the increased luminosity expected for LHC operations in 2017 and 2018, an upgrade was made to the pixel detector during the 2016-2017 year-end technical stop [29]. The new design is compared to the original in Fig. 22. The upgraded detector has pixel cells measuring 100 x 150 µm, placed in four layers in the barrel and three disks in the endcap, resulting in a spatial resolution of 15-20 µm, for vertex detection.

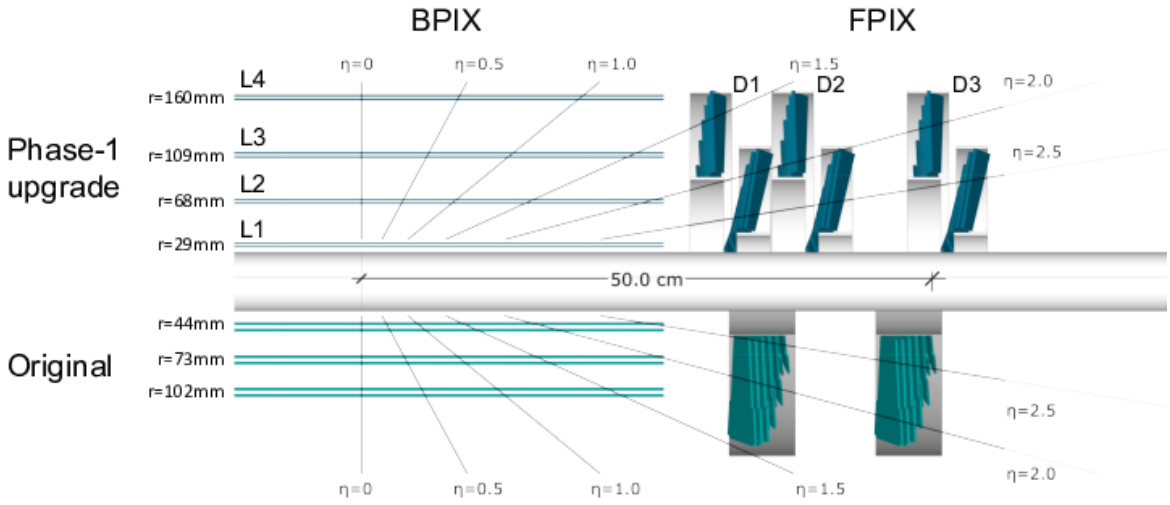


Figure 22: Comparison of the layout between the original CMS pixel detector (bottom) and after the Phase-1 upgrade (upper part) [29].

The intermediate region of the barrel, between 20 and 55 cm in radius, is covered by silicon microstrip detectors that are 10 cm wide and 80 µm thick. The outermost

region, between 55 and 120 cm in radius, is equipped with larger silicon strip detectors that measure 25 cm by 180 µm. These detectors have a resolution of 20 to 500 µm in the radial direction and 200 to 500 µm in the longitudinal direction for single point measurements.

The barrel tracking system is designed to minimize gaps in the acceptance by slightly shifting each module within a given layer with respect to its neighboring modules. The precise positioning of inactive elements surrounding the proton-proton collision point can be visualized in the x-y plane through a technique called hadrography [30], which uses reconstructed nuclear interactions. This can be seen in the Fig. 23, which shows the beam pipe position, the pixel detector, and the first layer of the inner barrel.

The tracking material and associated components, such as cables, support, and cooling system, take up a significant amount of space in front of the calorimeters, up to 1.6 radiation lengths [31]. The 2017 upgrade reduced the amount of pixel material by 40% in the endcaps and 10% in the barrel, see Fig. 24, improving the determination of the impact parameter and increasing the expected longitudinal IP resolution by up to 1.5 times. Additionally, the upgrade led to a 10% increase in b-tagging efficiency.

4.5 The Electromagnetic Calorimeter (ECAL)

The electromagnetic calorimeter (ECAL) [32] is built to accurately determine the energy of electrons and photons by measuring the scintillation light created via elec-

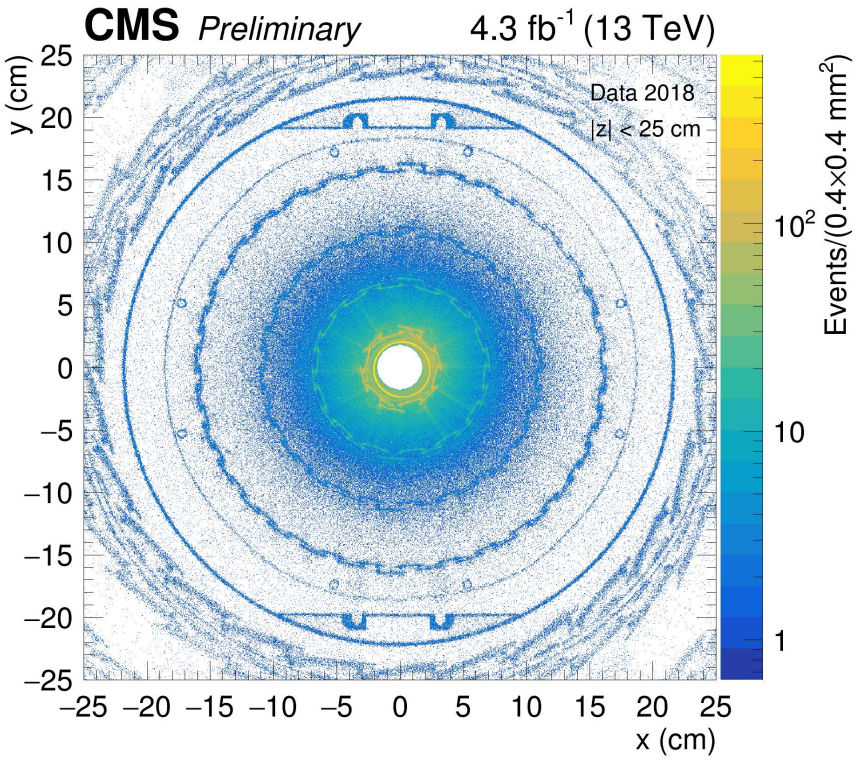


Figure 23: The hadrography $h(x, y)$ of the tracking system in the xy plane in the barrel region ($|z| < 25$ cm) is represented by a color scale indicating the density of nuclear interaction vertices. The beam pipe, the 4 layers of the barrel pixel detector with its support, and the first layer of the tracker inner barrel (TIB) detector are distinguishable above the background of mis-reconstructed nuclear interactions. [30].

tromagnetic showers caused by high-energy electrons or photons. It is composed of approximately 76000 scintillating lead tungstate crystals ($PbWO_4$). $PbWO_4$ was chosen for its high light yield, resistance to radiation, high density ($\rho = 8.28\,\mathrm{g/cm^3}$), small Moliere radius ($R_M = 2.2$ cm) and short radiation length ($X_0 = 0.89$ cm). Because of the fast light yield of the material chosen, more than 80% of the radiation photons are emitted before the next LHC bunch crossing occurs. The $PbWO_4$

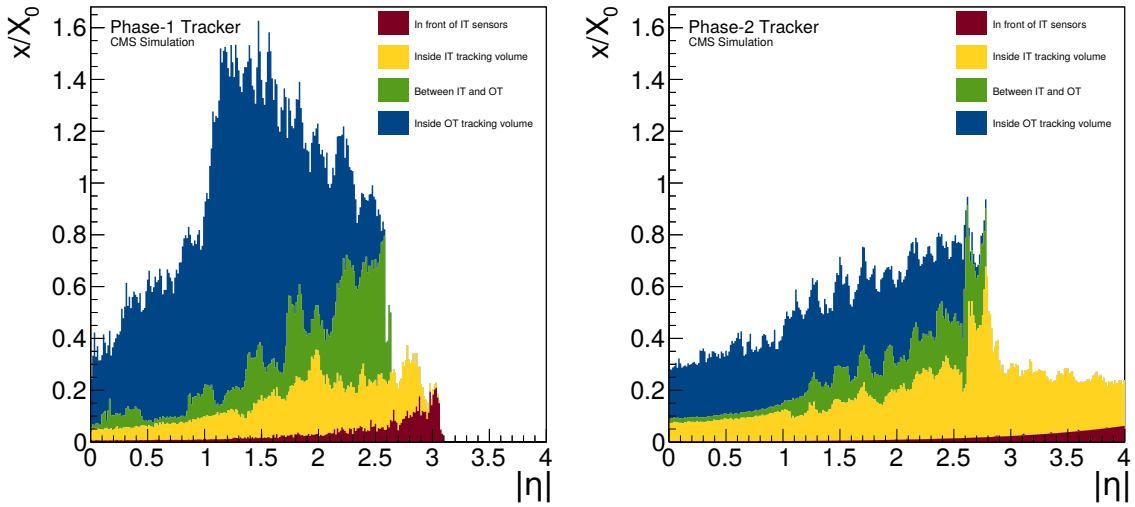


Figure 24: The material budget within the tracking volume is estimated in units of radiation lengths, and is compared between the Phase-1 (left) and Phase-2 (right) detectors. The stacked histograms display the amount of material in the following locations: front of the Inner Tracker (IT) sensors in brown, inside the IT tracking volume in yellow, in between IT and Outer Tracker (OT) sensors in green, and inside the OT tracking volume in blue. [31].

crystals have a short radiation length, allowing the electromagnetic shower to be mostly contained within a single crystal $\sim 25\,X_0$). The ECAL is made up of two main subdetectors, the barrel (EB) which has a inner radius of 129 cm and covers pseudorapidity range of $|\eta| < 1.479$, and the two endcaps (EE) which covers $1.653 < |\eta| < 3.0$. The EB is composed of around 60,000 crystals which are 23 cm in length and 22 mm x 22 mm in transverse size, while the EE are made up of 7,000 crystals, each measuring 22 cm in length and 28.62 mm x 28.62 mm in transverse size. The crystals in the EB are organized into 36 supermodules (18 per half barrel) that cover 20 degrees in ϕ, while crystals in the EE are grouped into two semicircular dees. Also, in both subdetectors the crystals are tilted by 3 degrees relative to the direction to the IP to ensure that particles do not escape detection by traversing the small gaps between crystals. A schematic view of the ECAL can be found in

Fig. 25 while its longitudinal layout is shown in Fig. 26.

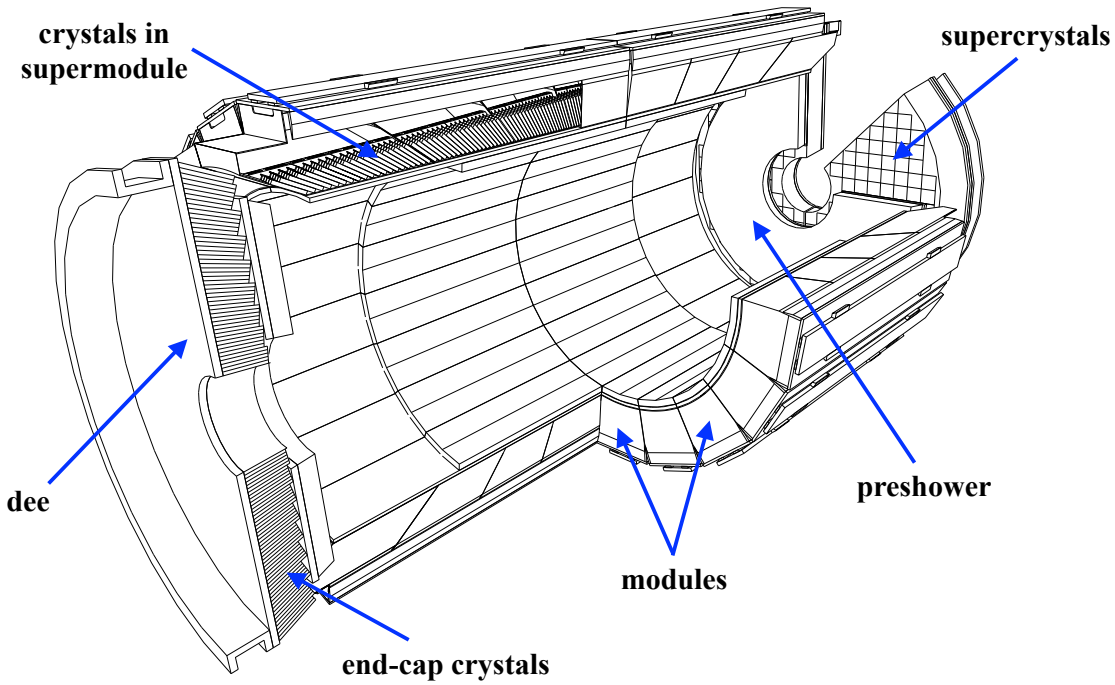


Figure 25: Illustration of the ECAL components. Picture adapted from [32].

The scintillation light yield of $PbWO_4$ is relatively low ($\sim 10\gamma$/MeV) and highly dependent on temperature, requiring amplification by high-gain photodetectors that can operate in high magnetic fields. The crystals' scintillation light is detected and amplified internally by Avalanche Photo Diodes (APDs) in the EB and Vacuum Photo-Triodes (VPTs) in the EE due to the differing levels of radiation and magnetic field configurations in these areas. A device called a pre-shower, consisting of two planes of silicon microstrip detectors and two lead absorbers, is placed before the EE to improve identification of neutral pions and covers a range of pseudorapidity $1.653 < |\eta| < 2.6$. The energy resolution of calorimeters can be described by the square sum of three factors:

$$\frac{\sigma_E}{E} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus C \tag{58}$$

where S represents the stochastic term for statistical fluctuations in light efficiency, N the noise term for electronic noise, and a constant term C for detector inhomogeneities and operational conditions. The ECAL energy resolution was measured in a dedicated test beam, with the resolution parameters found to be $S = 2.8\%$, $N = 12\%$, and $C = 0.30\%$. The energy resolution of the ECAL is outstanding, with a precision of around 1% for photons or electrons with energy above 10 GeV.

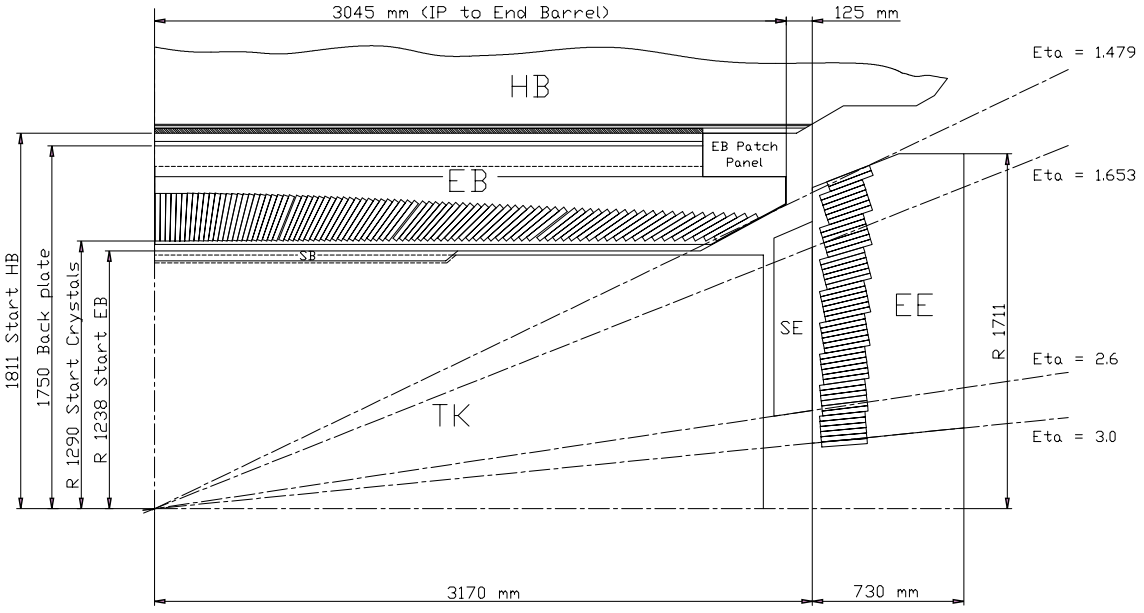


Figure 26: Longitudinal schematic layout of one quadrant of the ECAL subdetector. Picture adapted from [32].

4.6 The Hadronic Calorimeter (HCAL)

The Hadronic Calorimeter (HCAL), described in extensive details in [33], is meant to capture and quantify the energy of hadrons. Long-lived hadrons passing through

the ECAL cause showers in HCAL. Electromagnetic showers are generated by electromagnetic interactions and predominantly stem from photon conversions and bremsstrahlung radiation; hadronic showers, on the other hand, are the result of strong interactions that also create undetectable neutrinos, causing non-Gaussian tails in energy resolution. The proportion of hadronic and electromagnetic shower energies is dependent on the energy of incident particles.

The HCAL is situated immediately after the ECAL and before the magnet coil. It is a sampling calorimeter composed of alternating brass and plastic scintillator layers. The scintillation light is detected through optical fibers by Hybrid Photodiodes (HPD) and Silicon Photomultipliers (SiPM). Fig. 27 illustrates the schematic cross-section of a slice of HCAL, consisting of the central barrel part (HB) covering the region $|\eta| < 1.3$ and endcaps (HE) covering the area between $1.3 < |\eta| < 3.0$.

There are a couple of additional calorimeters to improve the energy resolution of hadrons, starting with a small outer hadronic calorimeter (HO) that measures and dampens hadronic showers from highly energetic hadrons that would have otherwise not deposited all of their energy in the inner calorimeter or the magnetic coil. The HO is situated just outside of the magnet solenoid. Lastly, there is a forward hadronic calorimeter (HF) that covers the range $3.0 < |\eta| < 5.0$ composed of steel to absorb the hadronic showers and quartz fibers to measure them.

In terms of performance, the HCAL obeys the same physics as the ECAL Eq. 58, just with worse energy resolution. Measurements yielded a stochastic term of $S = 84.7\%$, a constant $C = 7.4\%$, while the electronic noise term is negligible compared to the other contributions.

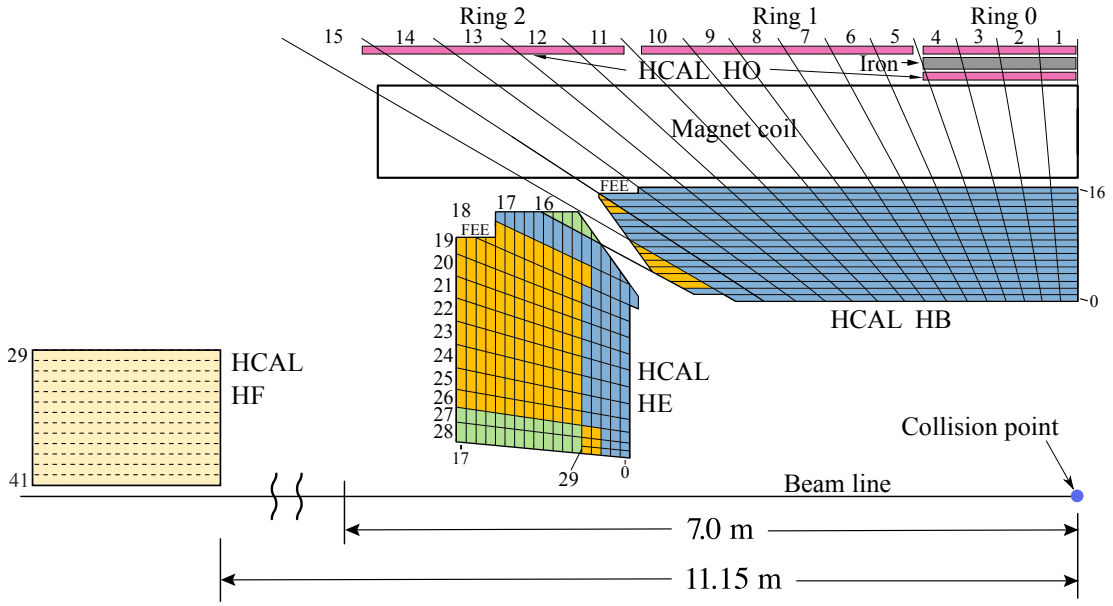


Figure 27: A schematic view of one quarter of the CMS HCAL during 2016 LHC operation. Here the positions of the four major components of the CMS HCAL are shown: the Hadron Barrel (HB), the Hadron Endcap (HE), the Hadron Outer (HO), and the Hadron Forward (HF) calorimeters. Picture adapted from [33].

4.7 The Muon System

Muons leave a trace in the tracking detectors and pass through ECAL and HCAL detectors without losing a significant portion of their energy. Thus, a separate muon system [34] is designed to identify muons and accurately measure their momentum. Such muon system is positioned within the iron return yoke of the solenoid magnet, which generates a strong magnetic field of 2T. Fig. 29 shows the schematic cross-section of the CMS muon system. The muon detection system utilises three types of gaseous detectors. In the barrel region, where the magnetic field is uniform and the muon rate low, drift tube (DT) detectors are used for $|\eta| < 1.2$. The DTs are organized into five wheels, each consisting of four concentric DT stations as seen in Fig. 30. Each DT cell is filled with a mixture of Ar and CO_2 gases and

Figure 28: Assembly of the 7th wedge of the HCAL during a hot day of August in 2000. Courtesy of L. Veillet.

has an anode wire in the middle and cathodes on the sides. DTs provide position measurements of traversing muons in the (r, ϕ) plane and the z position. In the endcap region, where radiation levels are high and the magnetic field is non-uniform ($0.9 < |\eta| < 2.4$), cathode strip chambers (CSC) are used. Each CSC is filled with a mixture of Ar, CO_2, and CF_4 gases and consists of six layers of anode wires interspersed between seven cathode plates. Both DTs and CSCs provide an excellent spatial resolution of about 40-150 µm, crucial for determining the muon momentum accurately. In addition, a complementary muon system made of resistive plate chambers (RPCs) is installed in both the barrel and endcap regions, covering $|\eta| < 1.6$. The RPCs consist of two resistive Bakelite layers separated by a volume filled with a mixture of $C_2H_2F_4$ and C_4H_{10}, SF_6 gases. The fast response and

good time resolution of RPCs make this system capable of determining the proton bunch crossing and can be used in the trigger system, as will be described in the next section.

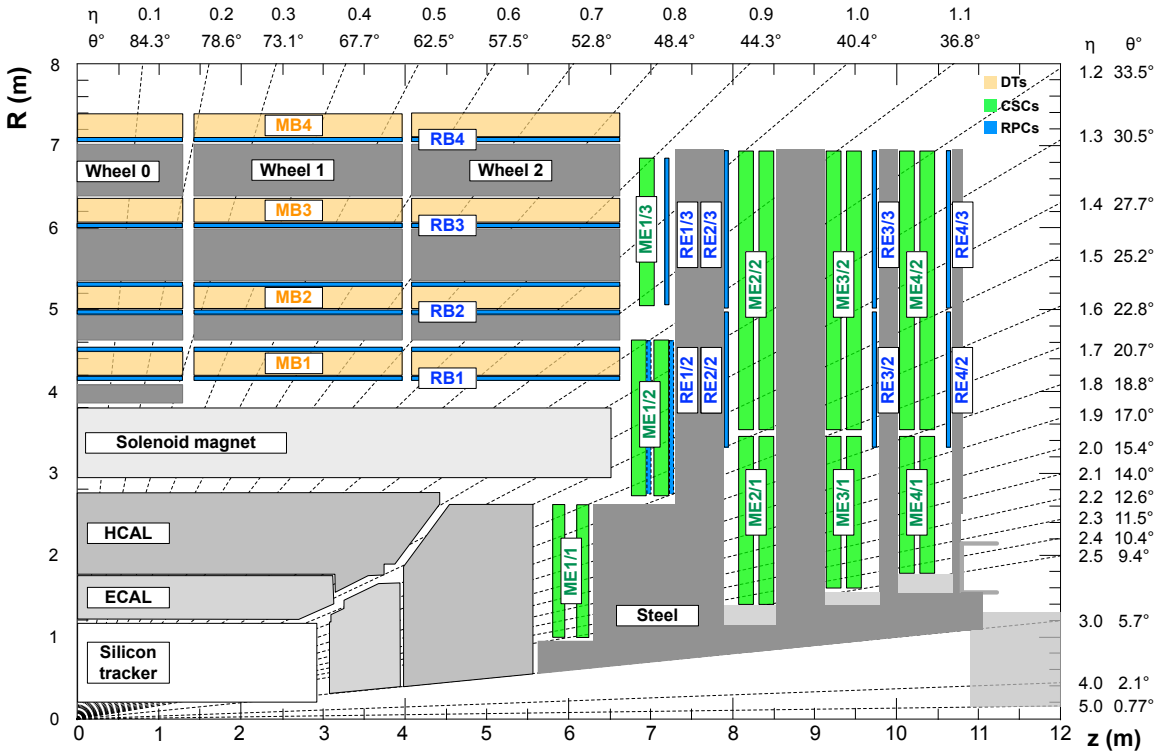


Figure 29: This figure depicts one quadrant of the CMS detector in its Run 2 configuration (from 2015), with the Muon detectors in colour.

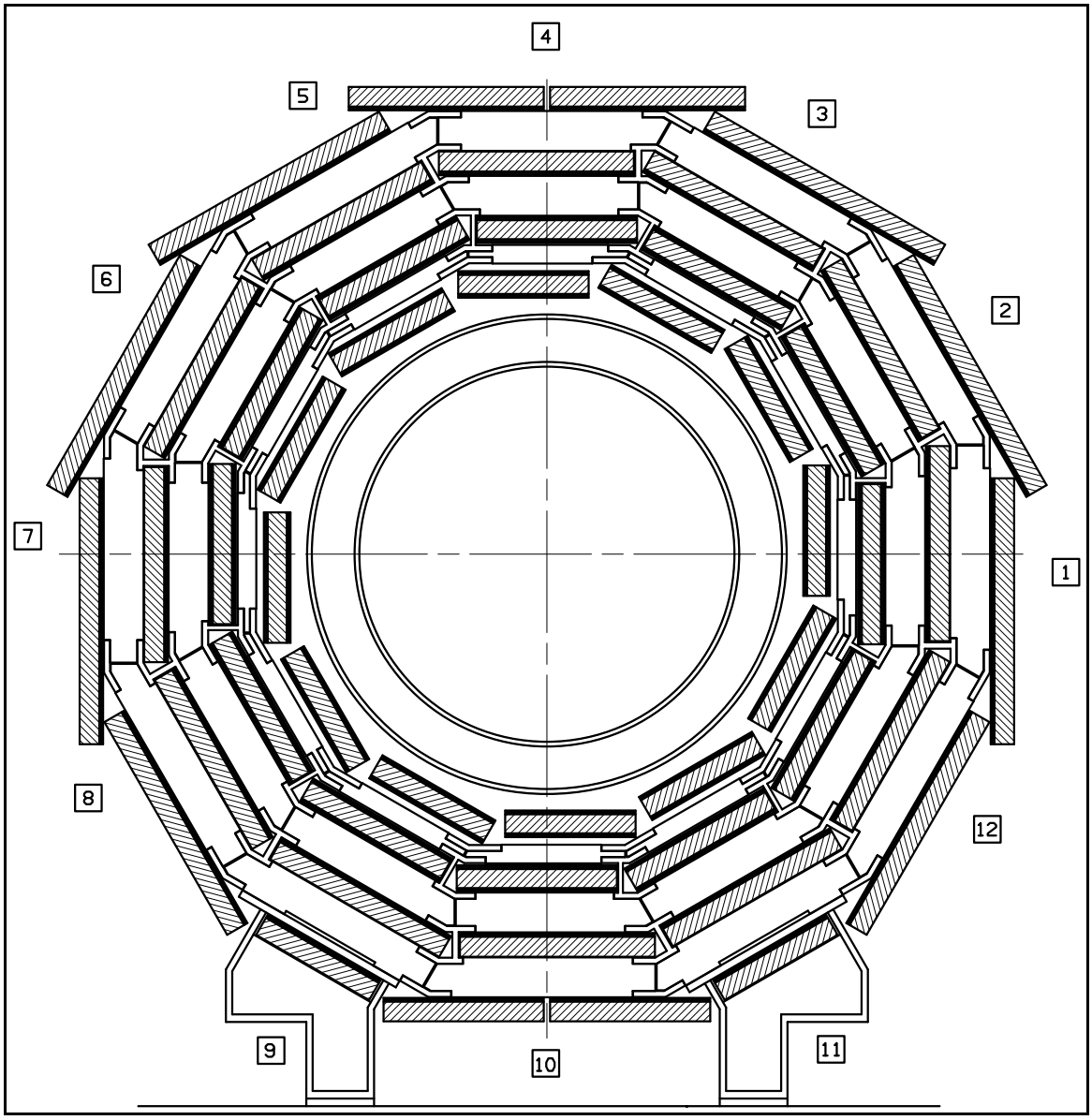


Figure 30: Layout of the CMS barrel muon DT chambers in one of the five wheels. All the chambers are identical, except for wheels -1 and +1 where the presence of the cryogenic chimneys requires the shortening of the chambers in two sectors. Additionally, in order to prevent wires longer than 4 meters, the MB4 chambers are cut in half in sectors 4 and 10. The RPC chambers, which are not shown to scale, are represented as bold lines and are composed of 2 layers for the MB1 and MB2 chambers and 1 layer for the MB3 and MB4 chambers.. Picture adapted from [34].

5 Event Reconstruction

The reconstruction of proton collisions, or as we call them in particle physics "events", starts with the CMS detector's trigger system, which records information from various subsystems, such as energy deposits from the calorimeters and particle hits in the muon system and tracker. These signals are then reconstructed offline to form individual physics objects, such as photons, leptons, jets, and missing transverse energy. The particle flow (PF) algorithm [35] is at the core of the CMS event reconstruction and aims to identify all stable particles in an event (PF candidates), using and correlating a combination of information from the sub-detectors. A diagram of the sub-detectors involved in the reconstruction of typical trajectories is provided in Fig. 31, which illustrates the different particle trajectories for muons, electrons, charged and neutral hadrons, and photons. The PF algorithm first reconstructs muons using the trajectories of hits in the silicon tracker and the muon system. Next, it identifies electrons using information from the calorimeters and tracker, and photons as ECAL energy deposits without tracks. Energy deposits in the ECAL and HCAL are associated with charged and neutral hadrons based on the tracker information. Finally, the algorithm computes the optimal four-momentum of the particles and uses this information to reconstruct hadronic jets and missing transverse momentum $\vec{p}_T^{miss}$. The missing transverse momentum vector, computed as the negative vector sum of the transverse momenta of all PF candidates in an event, has a magnitude denoted as p_T^{miss}. Given that the main interest of this analysis lies in the diphoton and di-tau final states, we will focus our attention on a detailed description of these objects and how they are reconstructed at CMS.

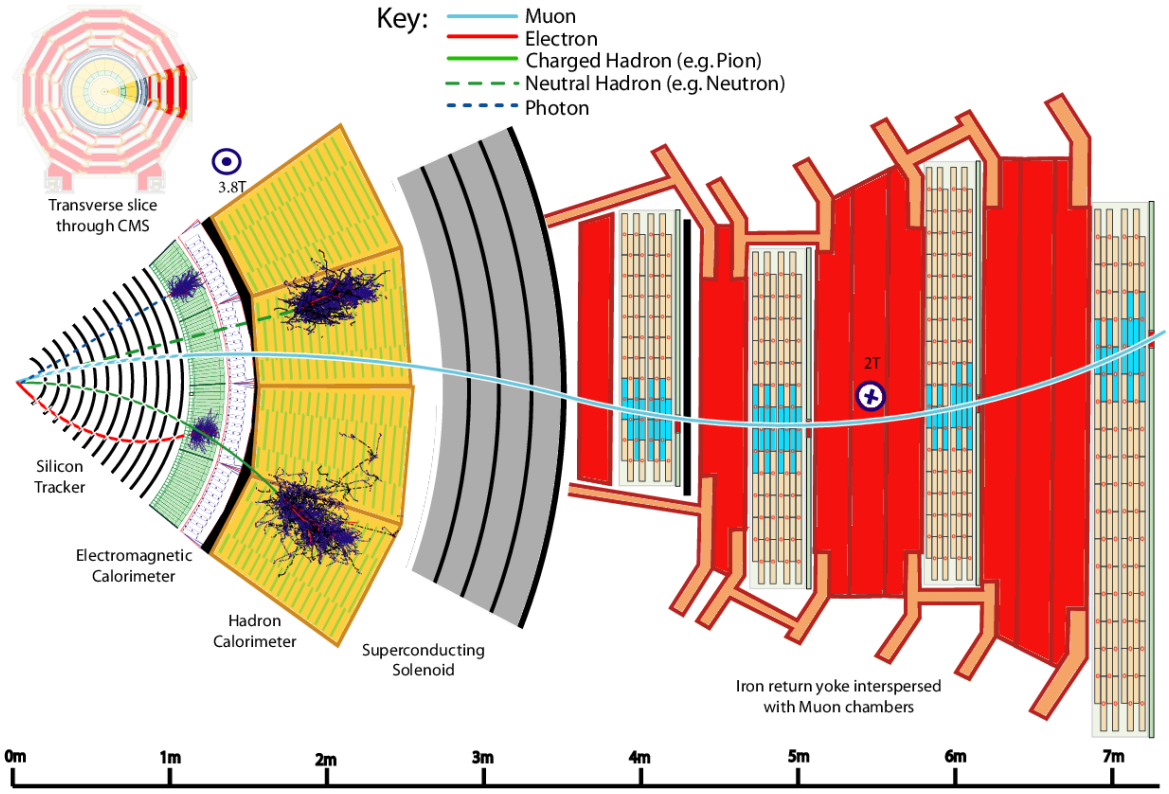


Figure 31: Transverse projection of the CMS detector [36] which emphasises the role that each sub-detector plays in identifying particles. Typical trajectories for muons, electrons, photons, charged and neutral hadrons are depicted in the picture along with the different sub-detectors they interact with.

5.1 Photon Reconstruction

The process of reconstructing photon candidates starts with the detection of energy clusters in the ECAL that are not linked to charged particle tracks, with the exception of converted photons. The electromagnetic shower from a photon in the ECAL usually tend to spread over multiple crystals, and the presence of the tracker upstream of the ECAL can lead to photon conversion into an electron-positron pair along with bremsstrahlung radiation from electrons and positrons. The strong magnetic field causes energy radiated by electrons and positrons to spread radially over the ϕ direction. On average, photons (electrons) deposit about 94% (97%) of their

total energy in a 3x3 (5x5) crystal matrix surrounding the crystal hit by the incident particle. To recover the radiated energy, different photon reconstruction algorithms group the ECAL energy deposits into clusters and superclusters (SC). The clustering technique begins with forming basic clusters from neighboring cells with deposited energy above a certain threshold, usually around a few hundred MeV; the exact value for the threshold depends on the detector conditions such as PU, noise, etc. along with the location of the cluster in the ECAL. These basic clusters are then merged to form a SC, where the clustering algorithms can recover around 95% of the photon energy, including converted photons. The photon candidates generated this way are then stored only if they fall within the ECAL and tracker fiducial region ($|\eta| < 2.5$), excluding the ECAL barrel-endcap transition region ($1.44 < |\eta| < 1.57$) where photon reconstruction is suboptimal.

5.2 Photon Energy Calculation

The energy of photon candidates is determined by summing the energy measured in the SC and the energy measured in the preshower subdetector, within its coverage. The reconstructed photon energy can be affected by various losses resulting from instrumental factors. For instance, the ECAL crystals can become less transparent over time due to exposure to ionizing radiation [37]. This issue can be partially resolved through thermal annealing of the crystals at room temperature; this practice can only take place when the crystals are not being irradiated like, for example, during the winter shutdown or the LHC interfill periods. To monitor the transparency of the crystals, a laser system is used to measure the amount of light that passes through each crystal and make adjustments accordingly. The degradation of the crystal transparency over time is shown in Fig. 32, where we notice how

the more forward regions in η suffers the most from the high radiation environment.

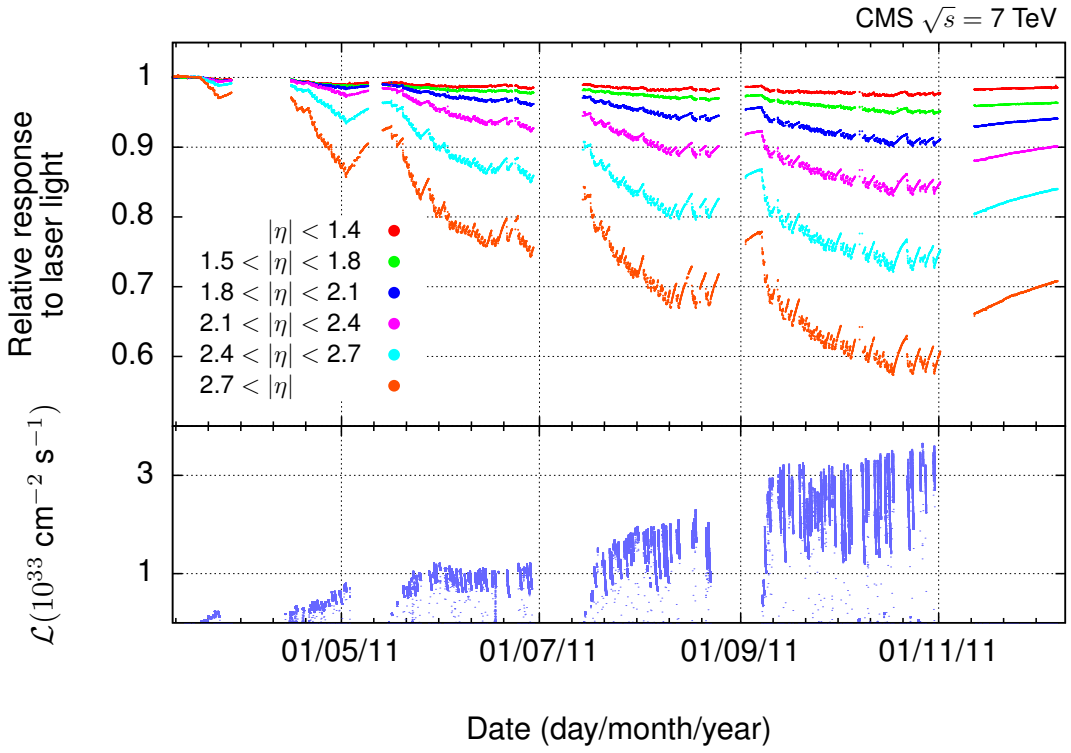


Figure 32: The relative response to laser light during 2011, normalized to data at the start of the year. The average response for each pseudorapidity range is presented. The bottom plot illustrates the corresponding instantaneous luminosity. Following the last LHC technical stop, a recuperation of crystal transparency is observed during the heavy-ion data-taking with low luminosity at the end of 2011. Picture from [37].

Additionally, a calibration process is performed to equalize the response of the ECAL channels. This is accomplished by exploiting the symmetry in ϕ of the energy flow, photon energy constraints from the reconstructed invariant mass of π^0 and η mesons, and constraints on the energy of electrons from vector boson decays. The results of these corrections are shown in Fig. 33, where we notice the excellent

reconstruction of the electron energy (E) and momentum (p) after applying the light monitoring corrections (green, middle plot).

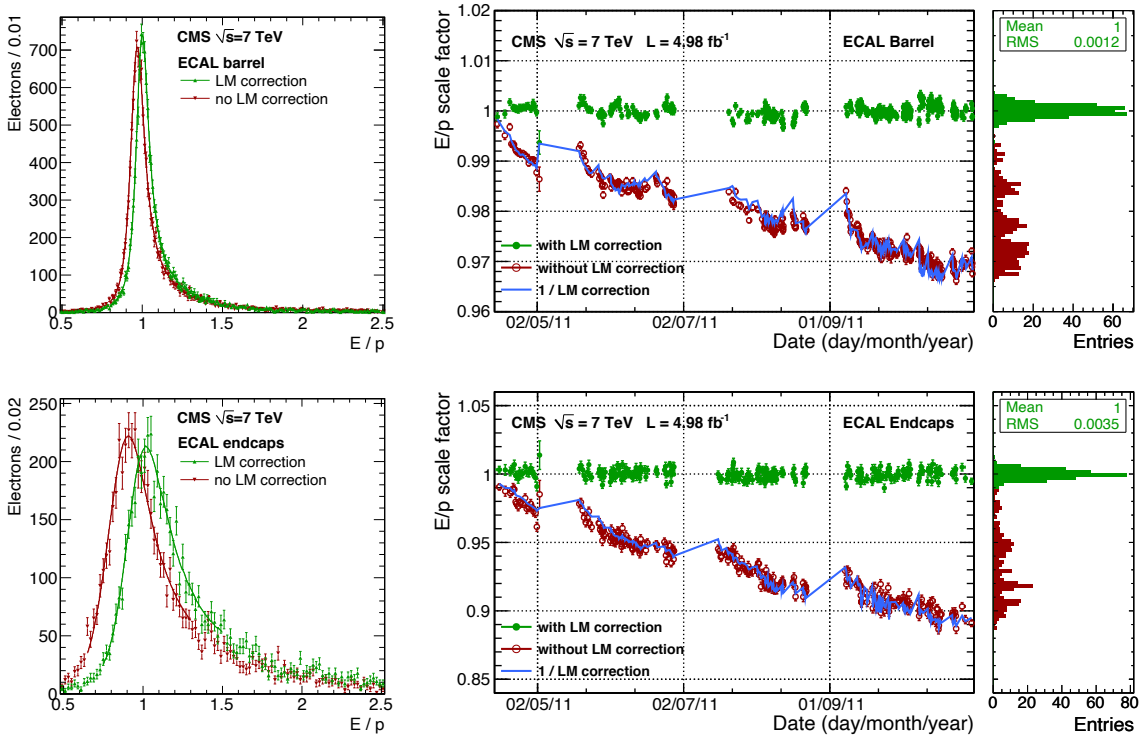


Figure 33: Relative variation in energy response for the Electromagnetic Calorimeter Barrel (EB) in the top panel and for the Electromagnetic Calorimeter Endcap (EE) in the bottom panel, as determined from the *energy over momentum* E/p analysis of electrons in W-boson decays. On the left, examples of fits to the E/p distributions before (red) and after (green) light monitoring corrections are applied. In the middle, the response stability during the 2011 pp data-taking period before (red open circles) and after (green points) response corrections is displayed. The blue line represents the inverse of the average light monitoring corrections. On the right, the distribution of the projected relative energy scales is shown. Picture from [37].

To correct for the radiation lost in front of the ECAL and suboptimal shower containment, a regression technique based on simulated events is employed [38]. This method, which uses a semi-parametric regression with a Gaussian core and two power law tails, provides estimates for the energy correction and the energy res-

olution σ_E / E. The regression is based on input variables that include information about the photon shower, the coordinates of clusters and superclusters, and the median energy density ρ in the event to account for pile-up effects.

Finally, a set of $Z \rightarrow ee$ events is used to correct residual discrepancies observed between data and Monte Carlo (MC) simulation, see Fig. 34. The ECAL energy scale in the data is adjusted using simulated $Z \rightarrow ee$ events, and the MC photon energy is smeared to match the resolution measured in the data, as shown in Fig. 35.

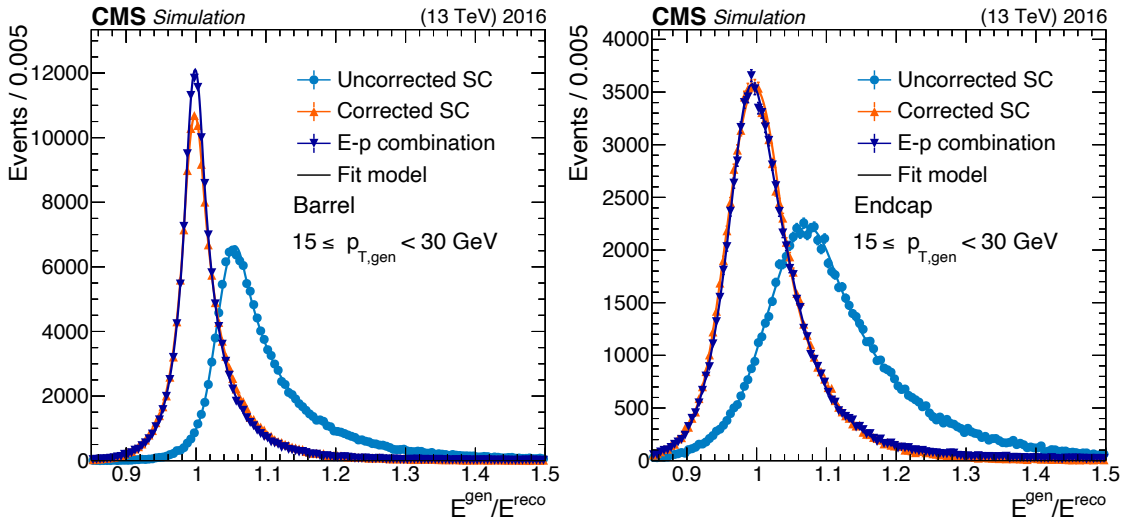


Figure 34: The figure displays the ratio of the true electron energy to the reconstructed electron energy in the pT range of 15-30 GeV, with and without regression corrections, for the barrel (left) and endcap (right) electrons in 2016 MC samples. A DSCB (Gaussian core and two power law tails) fit is overlaid on the plot. The statistical uncertainties of the MC samples are represented by the vertical bars on the markers. Picture from [38].

The energy resolution of photons varies based on the pseudorapidity region where the photon candidate is reconstructed. In the barrel section of the ECAL, the energy resolution of unconverted photons is about 1% for energies up to several TeV, while for all other barrel photons the energy resolution is roughly 1.3% for photons with

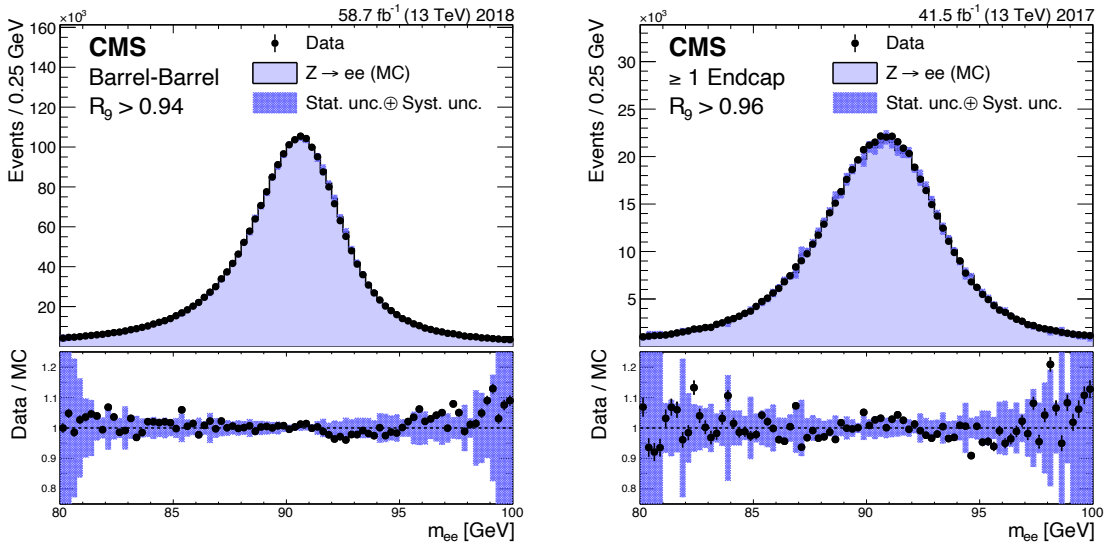


Figure 35: Invariant mass distribution of $Z \to ee$ events, after smearing the energy resolution in Monte Carlo and applying energy scale corrections to data. The results are presented for barrel (left) and endcap (right) electrons. Filled histograms represent MC events, while data are shown by the markers. The vertical bars on the markers represent the statistical uncertainties in data, while the combined statistical and systematic uncertainties in the simulation are depicted by the hatched regions. The lower panels display the ratio of the data to the simulation, with the uncertainties in the simulation represented by the bands. Picture from [38].

$|\eta| \sim 1.0$, and rises to 2.5% for photons with $|\eta| \sim 1.4$. The energy resolution of photons reconstructed in the endcap sections of the ECAL is worse, at around 2.5% for unconverted photons and 3-4% for converted ones.

5.3 Identification of Photon Candidates

A Boosted Decision Tree (BDT) multivariate classifier is trained to discern photons from jets, by assigning to each photon a score (photon ID). The photon identification is trained with variables that describe the electromagnetic shower shape and the isolation measures, which are defined as the sum of transverse momentum of photons and charged hadrons within a 0.3 radius of the cone $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$

around the photon candidate direction. To correct for the imperfections in the MC simulation modeling of the input variables, a correction method based on $Z \rightarrow ee$ events is used. This method, called a chained quantile regression, trains multiple BDTs to predict the cumulative distribution function using input variables such as the photon's transverse momentum p_T, pseudorapidity η, azimuthal angle ϕ, and global event energy density ρ. These correction models are then applied to the simulated photons to make sure their shower shape variables and photon ID accurately reproduces the distributions observed in data.

5.4 Tau Identification

Two main algorithms have been developed at CMS to identify hadronic tau decays (τ_H): the hadron plus strip (HPS) and the tau neutral classifier (TaNC) algorithms [39]. HPS is the common analysis tools, while TaNC has been used in the past as a cross-check and it won't be covered in this discussion. The HPS algorithm uses the PF algorithm to form particle candidates in the events; these particles are categorised in a mutually exclusive way into the following categories: photon, electrons, muons, charged hadrons and neutral hadrons. The dominant hadronic tau decays involve either one or three charged π mesons and up to two neutral π^0 mesons, the exact branching ratios are summarised in Table 3.

After the PF candidates have been reconstructed, and a PF jet has been selected as the candidate for the τ_H object, its four momentum is then reconstructed using the anti-k_T algorithm and adopting a distance parameter of $R = 0.5$. Note that this specific isolation requirement will be later propagated to the analysis taus, since the scale factors applied to our MC prediction would otherwise be invalid. After

Decay Mode	Resonance	Mass (MeV/c^2)	Branching fraction (%)
$\tau^- \to h^- \nu_\tau$			11.6
$\tau^- \to h^- \pi^0 \nu_\tau$	ρ^-	770	26.0
$\tau^- \to h^- \pi^0 \pi^0 \nu_\tau$	a_1^-	1200	9.5
$\tau^- \to h^- h^+ h^- \nu_\tau$	a_1^-	1200	9.8
$\tau^- \to h^- h^+ h^- \pi^0 \nu_\tau$			4.8

Table 3: Branching ratios of the dominant hadronic decays for the τ lepton. In this table, h represents both pions π and kaons K. We also included the symbol and the mass of any intermediate hadronic resonance.

we have chosen our PF candidate, the HPS algorithm first searches for the π^0 decay products of the τ_H, then combines the neutral hadrons found with the charged ones in order to compute the tau decay mode and evaluate the corrected four momentum along with the isolation measures. Particular attention is paid by the HPS algorithm to photon conversions in the CMS tracker, since the bending of electrons or positrons due to the high magnetic field tend to spread the calorimeter signature of π^0s along the ϕ direction. The HPS deals with this phenomenon by associating a "strip" (hence the name) to each of the converted photons, a strip is identified as an object that is built from electromagnetically charged PF candidates such as electrons or photons. First, the most energetic electromagnetic particle within the PF jet is identified and the strip centered around it, then the algorithms looks for other EM charged candidates in close proximity to the current strip center, within a window of $\Delta\eta < 0.05$ and $\Delta\phi < 0.20$. In the scenario where other EM candidates are found within these boundaries, the most energetic one out of those found is associated with the strip and its four momentum recalculated. This procedure is applied until no electromagnetic particles are found in proximity of the strip. Once a number of these strips are calculated, only those with $p_T^{strip} > 1$ GeV / c are combined into the charged hadrons that are used to reconstruct the τ_H decay modes. The allowed topologies of hadronic tau decays in the HPS algorithm are:

- Single hadron: $\tau \to h^- \nu_\tau$ and $\tau \to h^- \pi^0 \nu_\tau$ decays where the π^0s do not have enough energy to be reconstructed as a strip.
- One hadron + one strip: correspond to $\tau \to h^- \pi^0 \nu_\tau$ where the photons from the neutral meson decay are collimated on the calorimeter surface.
- One hadron + two strips: correspond to $\tau \to h^- \pi^0 \nu_\tau$ where the photons from the neutral meson decay are resolved as two separate physics objects on the calorimeter surface.
- Three hadrons: correspond to $\tau \to h^- h^+ h^-$ decays in which the three charged hadrons are required to originate from the same secondary vertex.

There are no additional final states topologies that account for $\tau \to h^- \pi^0 \pi^0$ and $\tau \to h^- h^+ h^- \pi^0 \nu_\tau$, as they are assimilated into the above categories. In addition, all strips and charged hadrons associated with the τ_H are required to be within a cone of size $\Delta R < (2.8 \text{ GeV/c}) \ / \ p_T^{\tau_H}$, with $p_T^{\tau_H}$ being the transverse momentum of the hadronic tau candidate, and the so-reconstructed tau momentum is required to be within a cone of radius $\Delta R < 0.1$ from the initial jet PF candidate that it originated from. Lastly, an isolation criteria is applied by requiring the absence of additional photons or charged hadrons in the proximity ($\Delta R < 0.5$) of the τ_H candidate. By tweaking the p_T threshold of what can be considered particles within the isolation cone, three working points are derived as ”Loose”, ”Medium” and ”Tight”. The ”Loose” working point (WP) is roughly equivalent to a probability of 1% for a jet to be misclassified as an hadronic tau. Successive WPs further reduce the misidentification rate by a factor of 2 with respect to the previous one. The performance of the HPS algorithm is shown in Fig. 36 and compared to the TaNC

algorithm for the three working points considered.

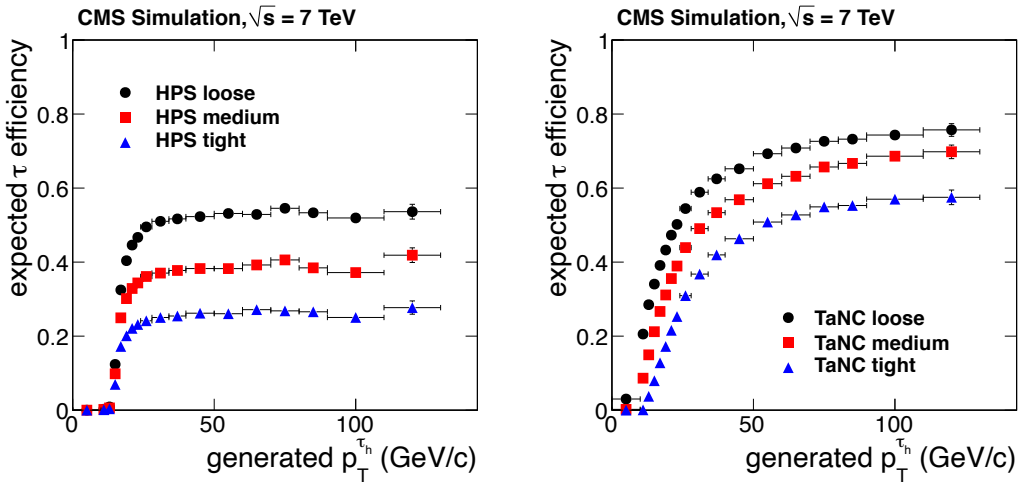


Figure 36: Expected reconstruction efficiency of the τ_H algorithms as a function of the generated $p_T^{\tau_H}$. The plot was generated using a simulated sample of $Z \rightarrow \tau\tau$ events which were reconstructed using the HPS (left) and the TaNC (right) algorithms for the "Loose", "Medium" and "Tight" working points. Picture from [39].

5.5 Muon Identification

Tau leptons decay to muons in about 17% of the cases, via $\tau \rightarrow \mu\bar{\nu}_\mu\nu_\tau$. Because of the presence of neutrinos, these tau candidates are measured with worse energy resolution. However, correctly identifying and measuring muon kinematics is rather straightforward in CMS and does not necessarily involve the PF algorithm, since muons have a distinctive signature in the dedicated muon trackers, which provide high efficiency across the entire detector acceptance. The design of CMS optimizes the absorption of all particles (except for muons and neutrinos) in the calorimeters, resulting in high muon purity in the outermost detector. Three reconstruction algorithms have been implemented to utilise information from the subdetectors:

- Standalone Muon: reconstructed using only hits in the muon detectors, obtained by fitting the data.

- Tracker Muon: reconstructed in the inner tracker, with the extrapolated track having to be consistent with at least one track segment in the muon detectors.
- Global Muon: if the parameters of the track reconstructed in the inner tracker and muon chambers, when propagated to a common surface, are compatible, they are combined into a single "global muon candidate".

Approximately 99% of muons produced within the geometrical acceptance of the muon system can be reconstructed as a standalone muon, tracker muon, or a combination of both (in which case, they are merged into a single candidate). Muons with low transverse momentum ($p_T^{\mu} < 5$ GeV) often do not meet the criteria for global-muon reconstruction due to increased multiple scattering in the yoke material and are typically reconstructed only as tracker muons. Additionally, charged hadrons can be erroneously reconstructed as muons if their showers extend beyond the HCAL volume, thus contaminating the low p_T tracker muons collection. Furthermore, the default identification criteria for muons are found to be too stringent at high p_T. Hence, in the PF algorithm, different criteria are applied to muon tracks to achieve an optimal balance between identification efficiency and purity across the whole kinematic spectrum.

5.6 Electron Identification

Tau leptons can decay to electrons in a similar way to what was described for muons. With a branching ratio of about 17%, the $\tau \to e\overline{\nu}_e\nu_\tau$ decay chain generates one electron and two neutrinos from the original tau. For what concerns the identification and reconstruction of electrons, it mostly involves linking a reconstructed

tracker track with a shower in the electromagnetic calorimeter. This distinct signature of the electron sets it apart from charged hadrons and muons, which typically behave as minimum ionizing particles in the electromagnetic calorimeter. However, the reconstruction of electrons is complicated by the substantial material in the tracker, which results in significant bremsstrahlung radiation. Bremsstrahlung radiation spreads the electron's energy over a cone in the electromagnetic calorimeter and more than half of the produced electrons will radiate away $\sim$ 50% of their production energy due to the significant amount of material, more than 0.75 radiation lengths, in between the interaction point (IP, where electrons are produced) and the detection location (ECAL). Additionally, radiative photons can become problematic if they undergo subsequent pair conversion, with the resulting electron-positron pairs further reducing the initial electron energy. To account for the effects of bremsstrahlung, electron reconstruction involves selecting a seed cluster that corresponds to the electron trajectory and a clustering of the nearby energy deposits to form a supercluster, similarly to what is done for photons. Then, the energy of the electron is measured using two distinct approaches: one involves the summation over the energies of the clusters, which also makes use of a calibration using the known linear response of the electromagnetic calorimeter, and the second method directly quantifies the transverse momentum of the reconstructed track. The final measurement of the electron energy is obtained by combining these two measurements using a weighted average, with weights depending on the shower profile of the electron [40].

5.7 Charged Hadrons and Jets

The energy of charged hadrons is determined from a combination of their momentum, as measured in the tracker, and the matching ECAL and HCAL energy deposits, after they have been corrected for zero-suppression effects and for the response function of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is obtained from the corresponding corrected ECAL and HCAL energies.

For each event, hadronic jets are clustered from the reconstructed particles using the infrared and collinear safe anti-k_{T} algorithm [41, 42] with a distance parameter of 0.4. Jet momentum is determined as the vectorial sum of all particle momenta in the jet, and is found from simulation to be, on average, within 5 to 10% of the true momentum over the whole transverse momentum (p_{T}) spectrum and detector acceptance. Additional pp interactions within the same or nearby bunch crossings can contribute with additional tracks and calorimetric energy depositions, increasing the apparent jet momentum. To mitigate this effect, tracks identified to be originating from pileup vertices are discarded, and an offset correction is applied to correct for remaining contributions. Jet energy corrections are derived from simulation studies so that the average measured energy of the jets becomes identical to that of particle-level jets. In situ measurements of the momentum balance in dijet, $\gamma +$ jets, Z+jet, and multijet events are used to determine any residual differences between the jet energy scale in data and simulation, and appropriate corrections are made [43]. Additional selection criteria are applied to each jet to remove jets that are potentially dominated by instrumental effects or reconstruction failures. The jet energy resolution amounts typically to 15–20% at 30 GeV, 10% at 100 GeV, and 5%

at 1 TeV [43].

The missing transverse momentum vector $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is computed as the negative vector sum of the transverse momenta of all the PF candidates in an event, and its magnitude is denoted as $p_{\mathrm{T}}^{\mathrm{miss}}$ [44]. The $\vec{p}_{\mathrm{T}}^{\,\mathrm{miss}}$ is modified to account for corrections to the energy scale of the reconstructed jets in the event.

5.8 The Trigger

The main idea behind the trigger system in a particle physics experiment is to reduce the storage data rate from the bunch crossing rate to a more manageable data acquisition rate. However, if we want to gain a deeper insight into why we need a trigger system and what it does, we must first understand the data flow at an experiment like CMS. As it was previously mentioned, the LHC generates proton-proton collisions every 25ns; these collisions then generate a number of secondary particles from quark and gluon interactions. Such particles travel through the detector leaving digital hits along their way. The digital hits are then used to reconstruct the trajectory of each particle and this information is stored for each individual event. Now, if we recoded every single event which occurs every 25ns, and stored it offline for later analysis, we would require around 40 Tb of storage space per second! This is an enormous number in terms of memory needed and processing power, that it is unfeasible for any computer cluster on earth. In order to mitigate this problem, CMS only stores those events and particles that it deems to contain any ”interesting” physics. What interesting physics means would require its own section and unfortunately we do not have enough human power to explore this topic in details, but let us just mention that for interesting physics is usually meant any type of interaction or generation of particles that we do not fully understand yet. If there is

an event in which all processes originate from a theory that is believed to be well understood, then there is no need to study this process further and we decide not to store the event. If, instead, the subsequent event contains some highly energetic photons that may originate from a Higgs decay or some other heavy particle then this corresponds to an interesting event that we may want to look at offline. In the latter scenario, a trigger is fired within the CMS detector that allows to store all the information regarding every single particle in the event for offline studies. In this way, we are able to greatly suppress the data storage rate needed to run the experiment to about $\sim$ 3 Gb/s. The trigger system is divided into two stages of data skimming: a Level 1 (L1) trigger [45] and a High Level Trigger (HLT) [46], as can be seen from Fig. 37.

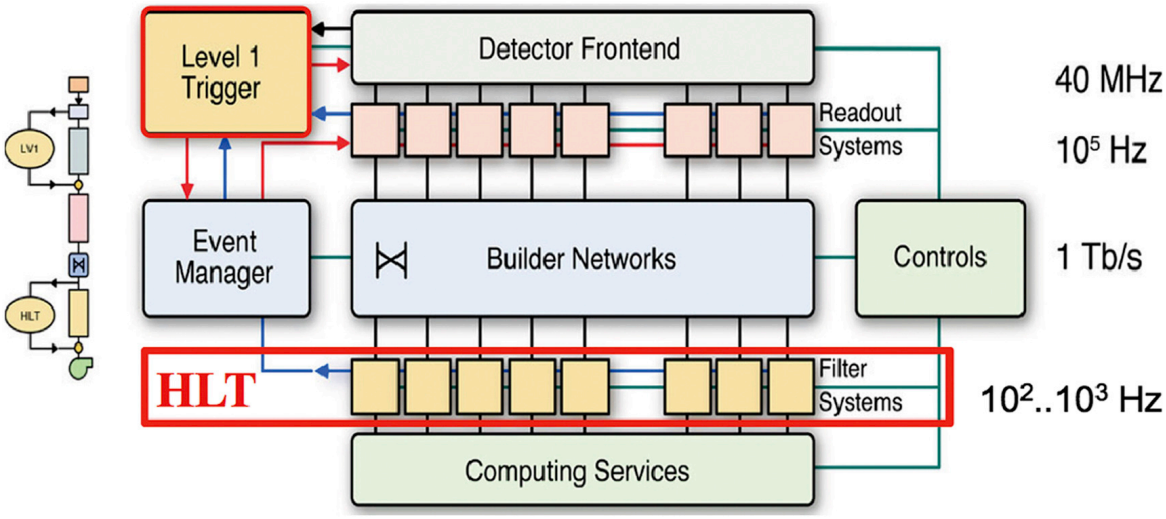


Figure 37: Schematic illustration of the full trigger system at CMS, which comprises of a Level 1 (L1) trigger to initially reduce the incoming data rate from 40MHz to about 100 kHz, and a High Level Trigger (HLT) which further suppresses the rate to about a few kHz. Picture from [47].

5.8.1 Level 1 Trigger

The L1 trigger operates at the hardware level, mostly making use of the beautiful piece of electrical engineering that Field Programmable Gate Arrays (FPGAs) and Application Specific Integrated Circuits (ASICs) are. L1 trigger reads and ”interprets” information from the calorimeters and the muon chambers in order to make a first rough decision whether to keep the event or not in 3.2μs [48], which allows the data rate to be reduced from over 40 MHz to about 100kHz. The L1 trigger information from the muon chambers and the calorimeters are merged into the L1 Global Trigger (GT) which ultimately makes the decision on whether to store the event or not. While the event is under scrutiny by the HLT, its data are temporarily stored in the pipelines of processing elements for the next step. A schematic illustration of the L1 trigger system logic is shown in Fig. 38.

5.8.2 High Level Trigger

After an event has passed the L1 GT, it is time for the HLT to make its own decisions. The HLT occurs at a later stage, when the data rate has already been decreased, therefore it can exploit the information collected by the whole detector in order to come up with a decision in around 100 ms. The HLT is an online software based system that makes use of around 20000 CPUs in order to handle the full L1 trigger information; this is the stage at which the PF candidates are reconstructed, usually by making use of the seeded tracks from the L1 objects around which the HLT candidates are built. In order for the HLT to store an event, the latter must satisfy the requirements of at least one ”HLT path”. Such path is a list of kinematic and particle quality cuts that aim at selecting interesting physics events; these can be events with two highly energetic photons (as needed for the Higgs discovery), or

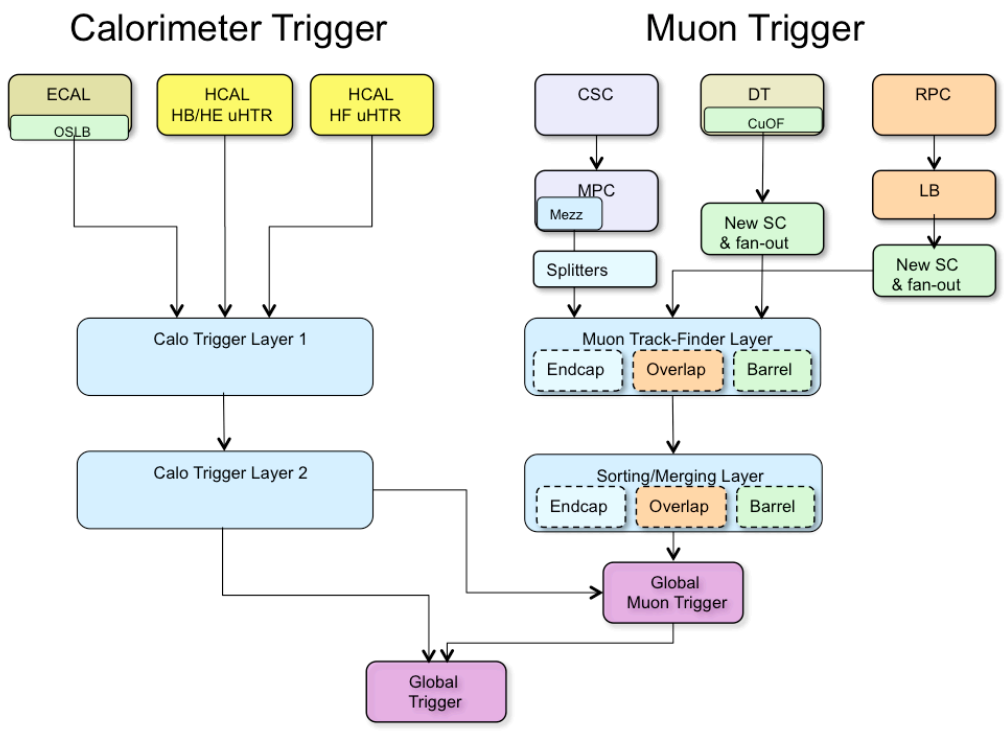


Figure 38: Illustration of the hierarchy within the L1 trigger [48].

events with highly energetic and isolated muons, or some other signature that has been deemed to be of interest when hardcoding the HLT paths into its software. Once the HLT accepts an event, the relevant data is permanently stored on the disk by the data acquisition system (DAQ).

6 Monte Carlo Events

In this chapter we will discuss the Monte Carlo (MC) algorithms that have been used to simulate our signal and background processes. The MC generation step is crucial in many analyses, since it has to take into account all of the physics effects and Feynman diagrams that may affect the production or suppression of a certain process in the phase space of interest. In order to understand what are the main backgrounds in our analysis, it is essential to understand the di-Higgs signature that we are after in this analysis. The final states we are interested in involve one Higgs decays to a pair of photons ($H \to \gamma\gamma$) and the other Higgs to a pair of tau leptons ($H \to \tau\tau$). On one hand, $H \to \gamma\gamma$ allows an excellent mass resolution in the reconstruction of the diphoton system along with a relatively low background contamination; its main drawback is the minimal branching fraction for this decay, $\mathcal{B}_{H\to\gamma\gamma} = 0.227\%$. On the other hand, $H \to \tau\tau$, retains a considerable branching ratio, $\mathcal{B}_{H\to\tau\tau} = 6.3\%$, although it suffers from worse resolution than $H \to \gamma\gamma$ due to the presence of neutrinos originating from the τs. This can be an important effect in the case where the di-tau invariant mass of the Higgs overlaps with the Z mass, which also can decay to a pair of taus. However, there are a few techniques discussed in Sec. 7 which allow an improved resolution of the di-tau invariant mass. In terms of $H \to \gamma\gamma$ contamination, the main backgrounds come from $\gamma\gamma + \text{jets}$ processes which can fake the $m_{\gamma\gamma}$ invariant distribution of the Higgs, although without a clear peak around the Higgs resonance. This important detail will be later exploited to greatly suppress $\gamma\gamma + \text{jets}$ backgrounds and limit its presence in the signal extraction window.

6.1 Data & Monte Carlo Propagation

Particle colliders are extremely complex systems and there are several steps that are required to simulate an event in going from its production via Feynman diagrams to the final reconstructed objects. When I started producing MC for my and other's analyses, I realised how little documentation there was about what every step in the production chain was doing and why, so here I decided to give a brief description of the workflow that goes into producing and validating the simulation at CMS. This is an essential and fundamental task, since if we did not trust our simulation we would not be able to make predictions about our models and what we understand or what we do not.

The first step in MC production involves declaring which types of Feynman diagrams we want to include in our simulation, the level of accuracy (LO or NLO or NNLO, etc.) and the various free parameters (such as masses of resonances, coupling strengths, etc.) of the models that we want to generate. This is typically done by softwares such as MG5_aMC@NLO , MADGRAPH or POWHEG, which output a gridpack or tarball with all of the needed information zipped in them. Subsequently, the following tasks takes place:

- **LHE-GEN**: at this step the gridpacks are unfolded and used to generate the specific processes that were requested. Here we are able to control the decay channels of the particles that are produced, so for example if we are only interested in the $H \rightarrow \gamma\gamma$ final states we can require the Higgs to only decay to photons, and so on. Also at this stage is where the QCD fragmentation takes place, using the PYTHIA package, which

models the hadronisation of quarks and gluons and we set the center of mass energy $\sqrt{s}$ that we want to run our simulation at, along with other parton shower details.

- **SIM** is where we propagate the particles generated at the LHE-GEN step through the detector, and we simulate the detector responses, such as the ECAL response to EM showers, or the muon detection. In order to do so we have to specify the running conditions and specifications of the detector, which are different for each year of data taking.
- **DIGIPremix** is the step that by far takes the longest to process! Here we inject the simulated events with a number of zero bias runs, which is real data that was collected with the detector triggering on any proton-proton collision, in order to account for the average level of QCD background to be expected in every event. This is also used to account for PU, which is not well modeled in simulation, and a number of other physical processes that degrade the performance of the detector and are extremely hard to simulate. Hence why we inject real data instead.
- **HLT**, as the name may suggest, this is the step where the PF candidates are built and the simulated hits throughout the detector are used to mimic the responses of the L1 and HLT triggers. Although this information is evaluated in simulation, in the analysis workflow we tend to not use any HLT triggers in our event selection but instead we scale the weight of each event by a trigger-associated scale factor (SF) that is derived from data, since the simulated HLT variables are not always in perfect agreement with data.

- **RECO/AOD** is the offline reconstruction of tracks, vertices, jets, electrons, muons, etc. along with hits and clusters. This step restructures the simulation into a more physics analysis friendly format, hence the AOD acronym: "Analysis Object Data".
- **miniAOD** is a more succinct format of AOD, which is able to store all of the event information in about $\sim$ 40 kb per event, AOD format uses around 400 kb so it is an impressive improvement! This size reduction is done by reducing the amount of information available per event, but still maintaining enough information for most analyses to be carried out.
- **nanoAOD** the last step of the MC production! This is a further data reduction from miniAOD, which allows to store all of the relevant information in just a few kb per event. The way in which nanoAOD achieves lower event-size is by further removing unnecessary information like, for instance, instead of storing all of the factors to apply jet energy corrections (JECs), as is done in miniAOD, nanoAOD only stores the Type-1 corrected jet so that, unless the JEC has to be undone and reapplied using a different correction scheme (which is very unlikely for most analysis), there is no need to apply JECs downstream. The main reason for moving away from miniAOD is that during Run 3, which just started last year around April 2022, the data flux is much larger than during Run 2 and it is not feasible to maintain all of the past and present datasets in miniAOD format. This issue will be even more relevant during the High Luminosity runs of the LHC (HL-LHC), which will further boost the instantaneous luminosity, hence a lighter data format is highly desirable.

All of our inputs have been processed in private nanoAOD production, where we started from centrally produced miniAOD for all of the backgrounds and from privately produced gridpacks for the signals considered. The private production of nanoAOD was necessary to include a number of photons variables needed for the standard $H \to \gamma\gamma$ diphoton preselection, which were missing in the centrally produced nanoAOD samples.

Data is directly handled by the CMS detector and its output is ready to be processed according to the RECO/AOD formats. Now, since the AOD reconstruction is performed offline, there have been several algorithms over the years that have been used to process the collected data. These algorithms are called reconstruction campaigns, and can differ from one another in terms of: thresholds cuts for certain objects to enter the electrons collection, Isolated Tracks collection, and so on, or also the offline reconstruction and identification of b-jets or tau leptons can be improved, etc. . The initial information available is always the same, but the offline reconstruction of physics objects can be always improved (both for Data and MC), especially given the monstrous advancement of machine learning algorithms in recent years. For this reason, in this analysis we use the latest reconstruction campaign available for Run 2, which corresponds to the `Summer20-UltraLegacy` campaign.

6.2 Signal Modelling

The ggF HH signal samples are generated with POWHEG 2.0 using the full top quark mass dependence [49] and with NLO accuracy [50–54]. Both SM ($\kappa_\lambda = 1$) and BSM ($\kappa_\lambda = 0\,,\, 2.45\,,\, 5$) scenarios are studied, all using $\kappa_t = 1$. As mentioned in Section 3.1 and in more details in Ref. [52], the HH cross section dependence on

κ_λ and κ_t can be expressed in terms of the three vertices / diagrams: κ_λ, κ_t and their interference. Hence, any linear superposition of three samples involving different κ_λ values can be used to obtain the signal sample corresponding to an arbitrary point in the $(\kappa_\lambda, \kappa_t)$ parameter space. Following the latest recommendations from the H group, we apply a kinematic re-weighting of the NLO κ_λ samples to produce 12 EFT benchmarks hypotheses [18], each corresponding to a different point in the 5D parameter space defined by the couplings: κ_λ, κ_t, c_2, c_g and c_{2g}. The analytical formula derived for the EFT re-weighting allows us to compute the cross section differentially as a function of the di-Higgs invariant mass and the angle between the beam direction in the di-Higgs center of mass frame and either of the Higgs bosons (Collins-Soper angle). The list of benchmark hypotheses generated is provided in Table 2.

6.2.1 Signal Processes

In this analysis, we focus on di-Higgs non resonant searches with final states involving $\gamma\gamma$ and $\tau\tau$; therefore, our main signal process is $HH \to \gamma\gamma\tau\tau$. However, we also include the process $HH \to \gamma\gamma WW$, where the WW pair can decay either dileptonically (both Ws decay as $W \to \ell\nu_\ell$, with $\ell = e, \mu, \tau$), or one W decay leptonically and the other W decays hadronically ($W \to q\bar{q}$). This decision is justified by the presence of two photons and two τs (or other charged leptons that overlap with our selection for leptonic τ decays) in the final states of all of the above states. Similarly to what was discussed in the previous paragraph, we also produce SM and BSM ($\kappa_\lambda = 0$, 2.45 , 5) for both the dileptonic and semileptonic decays of the $HH \to \gamma\gamma WW$, and we will refer to our signal HH as including both $HH \to \gamma\gamma\tau\tau$ and $HH \to \gamma\gamma WW$ for the remaining part of this analysis.

6.3 Backgrounds

The dominant backgrounds for this analysis are irreducible prompt diphoton production ($\gamma\gamma + \text{jets}$) and the reducible background from $\gamma + \text{jets}$ events, where jets can be misclassified as photons or tau leptons. The $\gamma\gamma + \text{jets}$ process is modeled at LO with SHERPA v.2.2.1 [55] and up to three additional jets as well as the box processes. To model $\gamma + \text{jets}$ we use instead PYTHIA 8 [56] at LO.

SM Higgs boson production, where the Higgs boson decays to a pair of photons, is considered as a resonant background. We include the following production modes: gluon-gluon fusion H (ggH), VBF H, vector boson associated production (VH) and $t\bar{t}H$. These processes are simulated using MADGRAPH5_aMC@NLO v2.4.2, with cross sections and decay branching ratios taken from Ref. [16]. No other processes from single H are considered, as their contributions would be negligible.

All simulated samples are interfaced using PYTHIA 8 with the tune CP5 [57, 58] for parton showering, fragmentation with the standard p_{T}-ordered parton shower (PS) scheme and the underlying event description. We use the GEANT4 [59] package to simulate the CMS detector response.

6.4 Data

The analysis uses data collected by CMS during the LHC pp collisions over a three year period from 2016 to 2018 for a total integrated luminosity of $138\,\mathrm{fb}^{-1}$. The events selected need to pass a diphoton trigger with asymmetric photon p_{T} thresholds of $p_{\mathrm{T}}^{\gamma 1} > 30\,\mathrm{GeV}$ and $p_{\mathrm{T}}^{\gamma 2} > 18(22)\,\mathrm{GeV}$ for the data collected in 2016 (2017 and 2018). The diphoton candidates are required to have an invariant mass exceeding 100 GeV, along with loose selections based on: electromagnetic calorimeter shower

shapes of the photon candidates, photons isolation with respect to the p_{T} of charged hadrons in their proximity and the ratio of energies deposited between the hadronic and electromagnetic calorimeters by photon showers [60].

Part III

Event Selection

7 Analysis Strategy

After producing the MC samples needed for the analysis, it is time to start our search! What we need to do first is to impose a loose selection based on kinematic cuts, particle identifications and other requirements in order to retrain most of our signal events while rejecting a large fraction of background. The first step in this process consists in defining the objects that we are after, photons and tau leptons, and require that these are well reconstructed in order to minimise the possibility of misclassified particles, e.g. particles that are erroneously classified as a particle that they are not. An overview of the analysis flow is depicted in Fig. 39, which involves a sequence of requirements that are imposed to all events in order to achieve higher signal over background purity in the region of interest.

7.1 $H \to \gamma\gamma$ candidate

The photon candidates selected for our analysis are reconstructed according to the procedure described in Sec. 5.1. After applying the photon energy correction, we require a further offline diphoton selection with criteria that are similar to, but more stringent than, those used in the trigger [60].

Photon candidates are also required to satisfy a loose identification (photon ID [60])

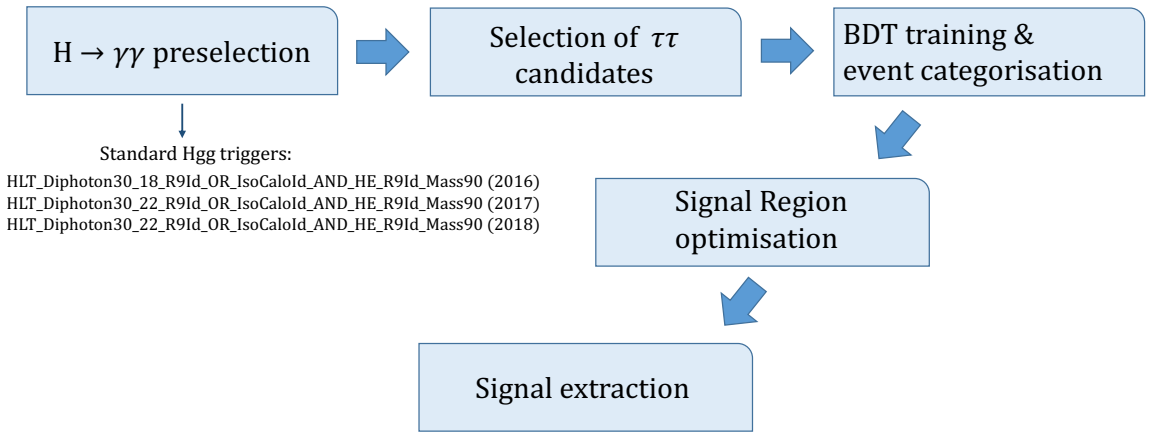


Figure 39: Flow of the $HH \to \gamma\gamma\tau\tau$ analysis requirements. The above selections are imposed to all events in order to achieve higher signal to background ratio before the extraction of the HH signal.

criterion based on a boosted decision tree (BDT) classifier trained to separate photons from jets. The photon ID BDT is trained using variables that describe the shape of the photon's electromagnetic shower and isolation. Photon candidates must also be within the fiducial regions of both the ECAL and the tracker, corresponding to a pseudorapidity requirement of $|\eta| < 2.5$, with the exception of those in the ECAL barrel-endcap transition region ($1.44 < |\eta| < 1.57$), which are excluded due to suboptimal reconstruction performance in this region.

Events are required to have at least two photon candidates and the Higgs boson candidates are then reconstructed from pairs of the selected photon. The diphoton invariant mass ($m_{\gamma\gamma}$) is required to be in the range 100–180 GeV for the searches targeting non resonant processes with a $H \to \gamma\gamma$ decay. We additionally impose mass-dependent photon p_{T} requirements of $p_{\mathrm{T}}/m_{\gamma\gamma} > 1/3$ and $p_{\mathrm{T}}/m_{\gamma\gamma} > 1/4$ for the highest p_{T} (leading) and second-highest p_{T} (subleading) photons, respectively.

7.2 $H \rightarrow \tau\tau$ candidate

When searching for opposite sign lepton candidates from H, we consider three physics objects: e, μ and τ_h, where τ_h is a hadronically decaying τ. If not enough of such candidates are found in the event, we then consider IsolatedTracks. Events with ditau candidates consistent with a $Z \rightarrow ll$ or $Z \rightarrow ll\gamma$ decay are rejected. Any event with an opposite-sign same flavor (OSSF) ee or $\mu\mu$ pair is rejected if either: $m_{ll} \in m_Z \pm 10$ GeV or $m_{ll\gamma} \in m_Z \pm 25$ GeV, where $m_{ll\gamma}$ is calculated with respect to both the leading and subleading photons. Events pass the preselection if they have a valid ditau candidate or if they have a single hadronic tau.

Electron candidates need to pass the requirements: $p_T > 10$ GeV, $|\eta| < 2.5$ and either $|\eta| < 1.4442$ or $|\eta| > 1.566$, $|d_{xy}| < 0.045$ and $|d_z| < 0.2$. We further require the electron to pass the working point `Electron_mvaFall17V2Iso_WP90` and an isolation cut with respect to the photon candidates of $\Delta R > 0.2$.

Muon candidates are selected after applying the cuts: $p_T > 15$ GeV, $|\eta| < 2.4$, $|d_{xy}| < 0.045$, $|d_z| < 0.2$ and `Muon_pfRelIso03_all` < 0.3. We then require that the muon is tagged as a GlobalMuon and passes medium cut-based ID, in addition to an isolation cut of $\Delta R > 0.2$ with respect to photon candidates.

The hadronic taus are selected according to: $p_T > 20$ GeV, $|\eta| < 2.3$, $|d_z| < 0.2$, and by requiring the `tauID('decayModeFindingNewDMs')` along with loose tau isolation IDs. We also make use of the deepTauID scores versus electrons, muons and jets with working points respectively at VVLoose, VLoose and Loose, in addition to an isolation requirement of $\Delta R > 0.5$ with respect to the selected photons and leptons.

If only a τ_h is reconstructed in the event, with no additional electrons or muons, we search for additional H candidates in the isolated track collection. We require these objects to be PF candidates, to originate from the PV, pass lepton veto and have $p_T > 5$ GeV, $|d_{xy}| < 0.2$ and $|d_z| < 0.1$. An isolation cut of $\Delta R > 0.2$ is applied with respect to the selected photons and leptons or taus.

In this analysis, jets are selected mostly to reject backgrounds like $t\bar{t}$ + jets. The selection implies: $p_T > 25$ GeV, $|\eta| < 2.4$ and pass Loose jet ID flag. Isolation between jets and photons, leptons or taus is attained by requiring $\Delta R > 0.4$. Additionally, b jets are identified from the selected jets using the Medium working point for the DeepJet b-tagger [61].

7.3 $\gamma\gamma\tau\tau$ events

In order for an event to pass the $\gamma\gamma\tau\tau$ selection, a photon pair must be present, as described in Sec. 7.1, along with a valid ditau candidate or a single hadronic tau (see Sec. 7.2). For events with more than one ditau candidate, the final candidate is chosen based on the composition of the objects forming the ditau candidate, with the following priority:

1. τ_h/τ_h
2. τ_h/μ
3. τ_h/e
4. μ/e
5. μ/μ

6. e/e

7. τ_h/isolated track

If an event still has more than one ditau candidate after selecting only those with the highest priority, the candidate with invariant mass closest to m_H is chosen.

The expected yields in the inclusive preselection (all categories) for signal, background, and data are shown in Table 4 below and for subsets of categories in Tables 5, 6, 7, 8, 9 below.

Preselection Yields: Inclusive of all Categories

Process	Yield	Stat. unc.	Syst. unc.	$\mathcal{F}$ of bkg
HHggWW-dileptonic	0.064	$\pm$0.000	$^{+0.012}_{-0.011}$	0.00
HHggWW-semileptonic	0.070	$\pm$0.000	$^{+0.013}_{-0.012}$	0.00
HHggTauTau	0.281	$\pm$0.000	$^{+0.053}_{-0.049}$	0.00
ttH-M125	7.7	$\pm$0.1	$^{+0.5}_{-1.0}$	0.00
VBFH-M125	10.4	$\pm$0.1	$^{+0.8}_{-0.8}$	0.00
VH-M125	12.9	$\pm$0.1	$^{+0.9}_{-0.9}$	0.00
ggH-M125	36.0	$\pm$0.4	$^{+1.8}_{-1.7}$	0.00
tt+X	200	$\pm$20	$^{+11}_{-12}$	0.01
WGamma	300	$\pm$10	$^{+15}_{-15}$	0.01
ZGamma	350	$\pm$8	$^{+30}_{-30}$	0.01
GJets	11800	$\pm$550	$^{+1600}_{-1400}$	0.46
DiPhoton	13200	$\pm$40	$^{+650}_{-650}$	0.51
Total MC bkg	25800	$\pm$550	$^{+1700}_{-1600}$	1.00
Data	24481	$\pm$156	$^{+0}_{-0}$	0.95

Table 4: Signal, resonant background, nonresonant background, and data yields in the preselection (inclusive of all categories).

From the inclusive Table 4, we see that for the full Run 2 we expect less than < 0.5 events from our HH signature and that the dominant backgrounds are $\gamma\gamma + \text{jets}$ and $\gamma + \text{jets}$ events. These two processes together make up for over 97% of our

Preselection Yields: 2τ / 0 ℓ category

Process	Yield	Stat. unc.	Syst. unc.	$\mathcal{F}$ of bkg
HHggWW-dileptonic	0.001	$\pm$0.000	$^{+0.000}_{-0.000}$	0.00
HHggWW-semileptonic	0.001	$\pm$0.000	$^{+0.000}_{-0.000}$	0.00
HHggTauTau	0.050	$\pm$0.000	$^{+0.009}_{-0.009}$	0.00
ttH-M125	0.139	$\pm$0.010	$^{+0.011}_{-0.020}$	0.00
VBFH-M125	0.037	$\pm$0.006	$^{+0.004}_{-0.004}$	0.00
VH-M125	0.251	$\pm$0.013	$^{+0.028}_{-0.028}$	0.00
ggH-M125	0.053	$\pm$0.020	$^{+0.003}_{-0.004}$	0.00
tt+X	0	$\pm$1	$^{+0}_{-0}$	0.00
WGamma	2	$\pm$0	$^{+0}_{-0}$	0.02
ZGamma	5	$\pm$1	$^{+0}_{-0}$	0.05
GJets	35	$\pm$16	$^{+7}_{-6}$	0.33
DiPhoton	63	$\pm$2	$^{+3}_{-3}$	0.60
Total MC bkg	106	$\pm$17	$^{+8}_{-7}$	1.00
Data	69	$\pm$8	$^{+0}_{-0}$	0.65

Table 5: Signal, resonant background, nonresonant background, and data yields in the 2τ / $0l$ category.

total background composition. In order to further characterise and differentiate the di-Higgs signature versus other processes, we calculate a number of additional variables that can improve the separation between signal and background. The first one that I want to mention is the corrected SVFit tau Lorentz vector. This is a procedure that makes use of the missing energy from neutrinos originating in τ decays to compute a corrected τ Lorentz vector with much improved mass resolution. This way we can reconstruct the invariant mass of the di-tau candidates from Higgs with a peak much closer to the nominal Higgs mass m_H, setting it apart from other backgrounds processes involving Z decays, such as $Z + \gamma$. Before we dive into the technicalities of the SVFit algorithm, I would prefer to give an overview of tau decays so that we can better understand the subtleties of the SVFit machinery.

Preselection Yields: 1τ / $1l$

Process	Yield	Stat. unc.	Syst. unc.	$\mathcal{F}$ of bkg
HHggWW-dileptonic	0.009	± 0.000	$^{+0.002}_{-0.002}$	0.00
HHggWW-semileptonic	0.008	± 0.000	$^{+0.002}_{-0.001}$	0.00
HHggTauTau	0.058	± 0.000	$^{+0.011}_{-0.010}$	0.00
ggH-M125	0.004	± 0.007	$^{+0.001}_{-0.001}$	0.00
VBFH-M125	0.005	± 0.002	$^{+0.001}_{-0.001}$	0.00
VH-M125	0.561	± 0.024	$^{+0.050}_{-0.052}$	0.01
ttH-M125	1.086	± 0.025	$^{+0.097}_{-0.161}$	0.02
GJets	2	± 1	$^{+0}_{-0}$	0.02
DiPhoton	9	± 0	$^{+0}_{-0}$	0.14
WGamma	9	± 2	$^{+0}_{-0}$	0.14
tt+X	11	± 3	$^{+1}_{-1}$	0.17
ZGamma	33	± 2	$^{+4}_{-3}$	0.51
Total MC bkg	64	± 5	$^{+4}_{-3}$	1.00
Data	54	± 7	$^{+0}_{-0}$	0.83

Table 6: Signal, resonant background, nonresonant background, and data yields in the in the 1τ / $1l$ category.

Preselection Yields: 0τ / $2l$

Process	Yield	Stat. unc.	Syst. unc.	$\mathcal{F}$ of bkg
HHggWW-dileptonic	0.042	$\pm$0.000	$^{+0.008}_{-0.007}$	0.00
HHggWW-semileptonic	0.000	$\pm$0.000	$^{+0.000}_{-0.000}$	0.00
HHggTauTau	0.016	$\pm$0.000	$^{+0.003}_{-0.003}$	0.00
ggH-M125	-0.002	$\pm$0.007	$^{+0.002}_{-0.005}$	-0.00
VBFH-M125	0.006	$\pm$0.002	$^{+0.001}_{-0.001}$	0.00
VH-M125	0.314	$\pm$0.015	$^{+0.026}_{-0.028}$	0.00
ttH-M125	1.763	$\pm$0.031	$^{+0.168}_{-0.267}$	0.02
GJets	0	$\pm$0	$^{+0}_{-0}$	0.00
WGamma	0	$\pm$0	$^{+0}_{-0}$	0.01
DiPhoton	13	$\pm$0	$^{+1}_{-0}$	0.16
tt+X	24	$\pm$4	$^{+2}_{-2}$	0.28
ZGamma	46	$\pm$2	$^{+4}_{-4}$	0.5
Total MC bkg	85	$\pm$5	$^{+4}_{-4}$	1.00
Data	62	$\pm$7	$^{+0}_{-0}$	0.72

Table 7: Signal, resonant background, nonresonant background, and data yields in the in the 0τ / $2l$ category.

Preselection Yields: 1τ / 1 IsoTrack

Process	Yield	Stat. unc.	Syst. unc.	$\mathcal{F}$ of bkg
HHggWW-dileptonic	0.001	$\pm$0.000	$^{+0.000}_{-0.000}$	0.00
HHggWW-semileptonic	0.006	$\pm$0.000	$^{+0.001}_{-0.001}$	0.00
HHggTauTau	0.026	$\pm$0.000	$^{+0.005}_{-0.005}$	0.00
VBFH-M125	0.289	$\pm$0.018	$^{+0.024}_{-0.025}$	0.00
VH-M125	0.545	$\pm$0.022	$^{+0.049}_{-0.051}$	0.00
ttH-M125	0.584	$\pm$0.019	$^{+0.049}_{-0.082}$	0.00
ggH-M125	0.960	$\pm$0.068	$^{+0.075}_{-0.074}$	0.00
WGamma	8	$\pm$1	$^{+0}_{-0}$	0.01
ZGamma	10	$\pm$1	$^{+1}_{-1}$	0.01
tt+X	14	$\pm$4	$^{+1}_{-2}$	0.02
GJets	350	$\pm$80	$^{+48}_{-51}$	0.45
DiPhoton	370	$\pm$5	$^{+20}_{-20}$	0.50
Total MC bkg	745	$\pm$80	$^{+52}_{-55}$	1.00
Data	1015	$\pm$31	$^{+0}_{-0}$	1.36

Table 8: Signal, resonant background, nonresonant background, and data yields in the in the 1τ / 1 iso-track category.

Preselection Yields: 1τ / 0 IsoTrack

Process	Yield	Stat. unc.	Syst. unc.	$\mathcal{F}$ of bkg
HHggWW-dileptonic	0.011	$\pm$0.000	$^{+0.002}_{-0.002}$	0.00
HHggWW-semileptonic	0.055	$\pm$0.000	$^{+0.010}_{-0.010}$	0.00
HHggTauTau	0.131	$\pm$0.000	$^{+0.025}_{-0.023}$	0.00
ttH-M125	4.161	$\pm$0.051	$^{+0.347}_{-0.558}$	0.00
VBFH-M125	10.095	$\pm$0.101	$^{+0.847}_{-0.838}$	0.00
VH-M125	11.268	$\pm$0.092	$^{+0.843}_{-0.851}$	0.00
ggH-M125	35.029	$\pm$0.402	$^{+1.769}_{-1.734}$	0.00
tt+X	152	$\pm$15	$^{+9}_{-9}$	0.01
ZGamma	250	$\pm$7	$^{+26}_{-23}$	0.01
WGamma	300	$\pm$10	$^{+14}_{-14}$	0.01
GJets	11500	$\pm$500	$^{+1508}_{-1377}$	0.46
DiPhoton	12700	$\pm$40	$^{+626}_{-629}$	0.51
Total MC bkg	24833	$\pm$550	$^{+1634}_{-1514}$	1.00
Data	23281	$\pm$152	$^{+0}_{-0}$	0.94

Table 9: Signal, resonant background, nonresonant background, and data yields in the in the $1\tau/0l$ category.

8 Tau Decays

τ leptons always decay to either leptons (e, μ) or a combination of charged and neutral hadrons. An illustration of a HH decay involving taus is depicted in Fig. 40.

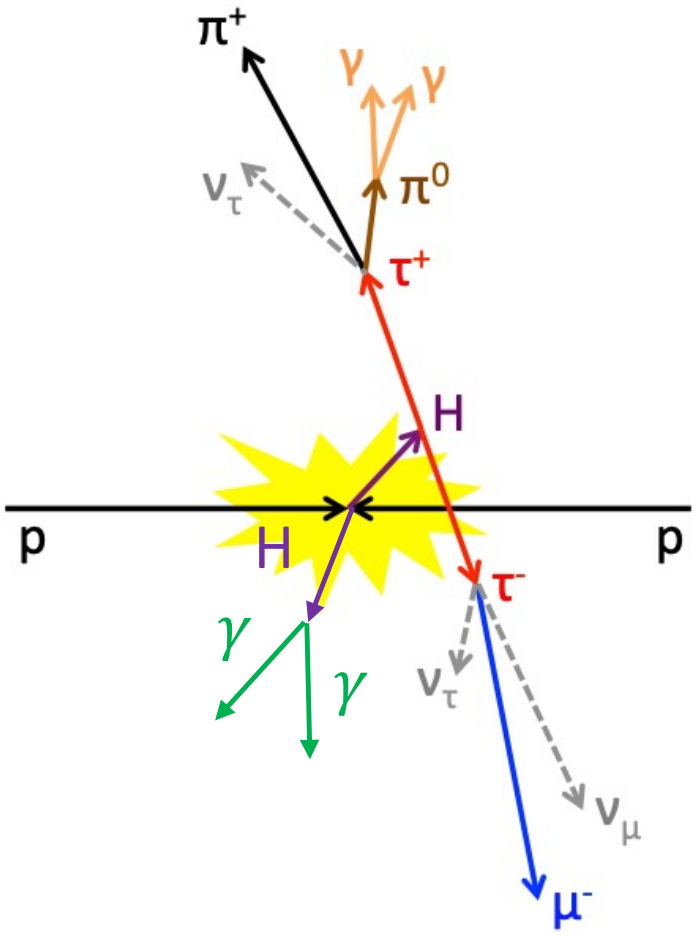


Figure 40: Illustration of a $\gamma\gamma\tau\tau$ event decay, where one Higgs decays to a pair of photons and the other Higgs to a pair of τ leptons, which further decay both leptonically and hadronically.

A summary of the τ branching ratios is given by:

- $\tau \to e\overline{\nu}_e\nu_\tau$, with $\mathcal{B} = 17.8\%$
- $\tau \to \mu\overline{\nu}_\mu\nu_\tau$, with $\mathcal{B} = 17.4\%$
- $\tau \to h^-\nu_\tau + \geq 0h^0$, with $\mathcal{B} = 49.5\%$
- $\tau \to h^-h^+h^-\nu_\tau + \geq 0h^0$, with $\mathcal{B} = 15.2\%$

where by $\geq 0h^0$ we mean any positive allowed number of neutral hadrons (mostly π^0). The average lifetime of tau leptons is around 290 fs, which corresponds to a distance of around $c\tau \times \gamma \sim 1$mm before it decays (we approximated $\gamma \sim p_T/m_\tau = 20/1.7 \sim 10$), hence failing to reach the first layer of the tracker and leaving us only with the τ decay products to reconstruct the mother particle. You may have noticed the presence of two νs in leptonic decays and only one ν in hadronic. This affects the resolution of the reconstructed tau candidates, which is much worse for leptonic decays since νs cannot be effectively measured in the CMS detector. The kinematic of hadronic or leptonic τ decays can be parametrised in terms of:

- $X = E_{vis}/E_\tau$: fraction of visible energy over the total τ energy
- ϕ : angle between the τ momentum and the momentum of its visible products
- $m_{\nu\nu}$: invariant mass of the neutrino system – only for leptonic τ decays

The variable X is also related to the angle in the laboratory frame between the τ momentum and its visible decays products, θ_{GJ} commonly known as the Gottfried–Jackson angle, by:

$$
\begin{aligned}
m_\tau^2 &= E_\tau^2 - \vec{p}_\tau^2 \\
&= (E_{vis} + E_{invis})^2 - \vec{p}_\tau^2 \\
&= \left(E_{vis} + \sqrt{|\vec{p}_\tau - \vec{p}_{vis}|^2 + m_{\nu\nu}^2} \right)^2 - \vec{p}_\tau^2 \\
&= \left(E_{vis} + \sqrt{\vec{p}_\tau^2 - 2|\vec{p}_\tau||\vec{p}_{vis}| \cos\theta_{GJ} + \vec{p}_{vis}^2 + m_{\nu\nu}^2} \right)^2 - \vec{p}_\tau^2
\end{aligned}
$$

we can then re-arrange the equation in terms of θ_{GJ} to get:

$$
\cos\theta_{GJ} = -\frac{m_\tau^2 + m_{vis}^2 - m_{\nu\nu}^2 \pm 2E_\tau E_{vis}}{2|\vec{p}_\tau||\vec{p}_{vis}|} \tag{59}
$$

but only one of the two solutions yields a value of $\cos\theta_{GJ}$ in the interval $[-1, 1]$. The $m_{\nu\nu}$ and X variables constrain the τ momentum to lie on the surface of a cone with a distance from the apex of $|\vec{p}_\tau| = \sqrt{(E_{vis}/X)^2 - m_\tau^2}$, whereas the axis of the cone is defined by the $\vec{p}$ of the visible decay products and θ_{GJ} represents the opening angle of the cone. The angular position of the τ momentum around the cone is instead defined by ϕ. Given the above quantities, the energy and momentum of the primary τ can be fully determined.

In our case, we are interested in system with two τ decays, hence the channels available become:

- $\tau_H\tau_H \rightarrow$ both taus decay hadronically
- $\tau_H\tau_l \rightarrow$ one tau decays hadronically one tau decays leptonically
- $\tau_l\tau_l \rightarrow$ both taus decay leptonically

These are graphically summarised in Fig. 41.

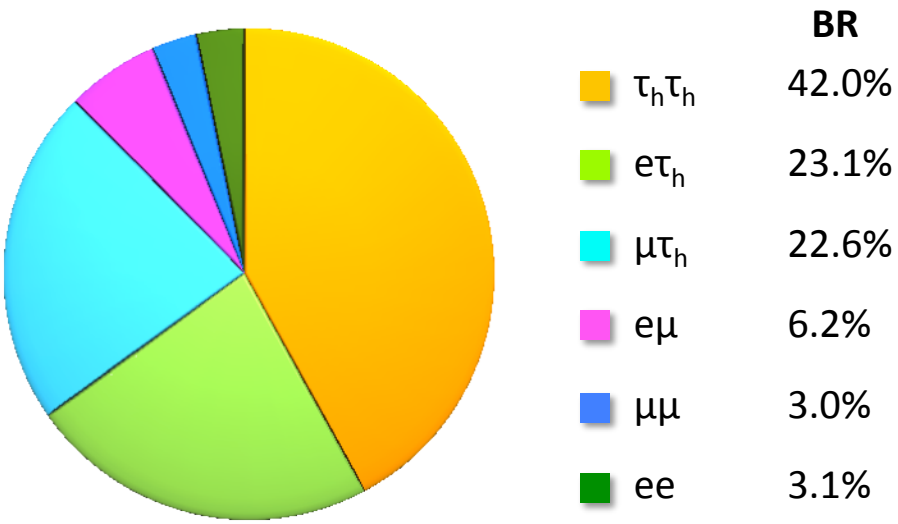


Figure 41: Branching ratios for the main $\tau\tau$ decay channels.

Therefore, the number of unknown parameters for each $\tau\tau$ decay mode is:

- $\tau_H\tau_H \rightarrow$ 4 parameters (X_1, ϕ_1, X_2, ϕ_2)
- $\tau_H\tau_l \rightarrow$ 5 parameters ($X_1, \phi_1, X_2, \phi_2, m^2_{\nu\nu}$)
- $\tau_l\tau_l \rightarrow$ 6 parameters ($X_1, \phi_1, m^1_{\nu\nu}, X_2, \phi_2, m^2_{\nu\nu}$)

while we only have two constraints / assumptions:

$$\sum p^{\nu}_x = E^{miss}_x \text{ and } E^{miss}_x = -\sum p^{PF}_x$$
$$\sum p^{\nu}_y = E^{miss}_y \text{ and } E^{miss}_y = -\sum p^{PF}_y$$

which equates to the assumption that the missing energy in the x (y) direction entirely comes from the neutrino candidates of the τ system and it is equal to the

negative of the sum of momenta along the x (y) axis of all of the PF candidates in the event. Now that we defined the objects and framework that we are working with we can introduce the SVFit algorithm.

8.1 The SVFit Algorithm

SVFit is a *Dynamical Likelihood Method* [62], which is based on a likelihood approach but it reconstructs kinematic quantities ($m_{\tau\tau}$) on an event-by-event basis, hence why *Dynamical*. In the estimation of $m_{\tau\tau}$, the measured observables E_x^{miss} and E_y^{miss} are combined with a probability model that accounts for missing E_T resolution and τ decay kinematics. A prediction for the probability density $p(\vec{x}|\vec{y},\vec{a})$ is generated by the model, where $p(\vec{x}|\vec{y},\vec{a})$ reflects the probability to observe the quantities $\vec{x} = (E_x^{miss}, E_y^{miss})$ measured in an event given a set of unknown parameters $\vec{a} = (\mathrm{X}_1, \phi_1, m_{\nu\nu}^1, X_2, \phi_2, m_{\nu\nu}^2)$ and assuming that the four momenta of the visible decay products equal the measured values $\vec{y} = (p_1^{vis}, p_2^{vis})$. The probability $P(m_{\tau\tau})$ of observing a particular mass is computed for a number of mass hypotheses $m_{\tau\tau}^i$ by *marginalisation* of the unknown parameters $\vec{a}$ as:

$$P(m_{\tau\tau}) = \int \delta\left(m_{\tau\tau} - m_{\tau\tau}(\vec{y},\vec{a})\right)\; p(\vec{x}|\vec{y},\vec{a})\, d\vec{a} \;\;, \tag{60}$$

where the right hand side of the equation can be interpreted as the weighted average over all hypothetic configurations that are compatible with the observed quantities $\vec{y}$. The numerical integration is carried out using the VEGAS algorithm [63] by using $\delta m_{\tau\tau}$ steps that evolve according to $\delta m_{\tau\tau}^{i+1} = 1.025 \cdot \delta m_{\tau\tau}^i$ with $\delta m_{\tau\tau}^0 = 2.5$ GeV and an overall $m_{\tau\tau}$ range of $m_{\tau\tau} \in [5, 2000]$ GeV. The optimal value $\hat{m}_{\tau\tau}$ is chosen so that it maximises $P(m_{\tau\tau}^i)$.

8.1.1 Choice of Matrix Elements

Matrix Elements (ME) for the leptonic and hadronic τ decays are taken from the literature [64, 65], where we assume the τs to be unpolarised in their decays to es and μs, which results in a probability density function:

$$\frac{d\Gamma}{dX\, dm_{\nu\nu}\, d\phi} \propto \frac{m_{\nu\nu}}{4m_\tau^2}\left[\left(m_\tau^2 + 2m_{\nu\nu}^2\right)\left(m_\tau^2 - m_{\nu\nu}^2\right)\right] \tag{61}$$

with the kinematic restrictions of $0 < X < 1$ and $0 \leq m_{\nu\nu} \leq m_\tau\sqrt{1-X}$. Otherwise the integrand of Eq.60 is taken to be zero.

The ME for the hadronic τ decays instead depends on the tau polarisation and its decay channel. In SVFit, an empirical model was found to be in good agreement with the predicted profiles and it can be expressed as:

$$\frac{d\Gamma}{dX\, d\phi} \propto \frac{1}{1 - m_{vis}^2/m_\tau^2} \tag{62}$$

where the hadronic system of decay products is assumed to be a single particle of mass m_{vis}. In this case, the above ME has validity in the range $m_{vis}^2/m_\tau^2 \leq X \leq 1$.

8.1.2 Missing Transverse Energy

SVFit also considers the likelihood of a τ decay hypothesis with the Missing Transverse Energy (MET) measured in an event. Potential differences between $\sum p_x^\nu$ and E_x^{miss} (similarly for y) are also accounted for in the model assuming a Gaussian profile resolution for the measurements, illustrated in Fig. 42. The likelihood term associated with the MET in the event and its resolution therefore takes the form:

$$\mathcal{L}_{MET} = \frac{1}{2\pi\sqrt{|V|}} \cdot \exp\left[-\frac{1}{2}\begin{pmatrix} E_x^{miss} - \sum p_x^{\nu} \\ E_y^{miss} - \sum p_y^{\nu} \end{pmatrix}^T \cdot V^{-1} \cdot \begin{pmatrix} E_x^{miss} - \sum p_x^{\nu} \\ E_y^{miss} - \sum p_y^{\nu} \end{pmatrix}\right] \quad (63)$$

where V is the MET covariance matrix (to account for the MET resolution in the event) and $|V|$ its determinant.

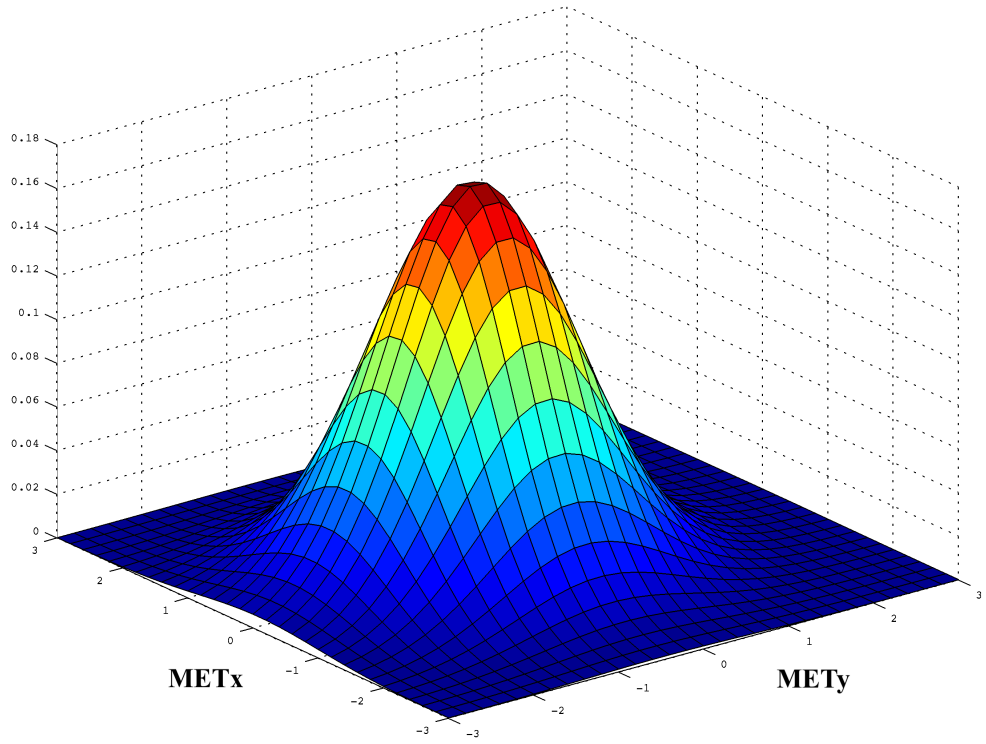


Figure 42: The MET variables in x and y are assumed to originate from a 2D Gaussian profile whose spread depends on the MET covariance matrix for that particular event, so to account for the MET resolution on an event basis.

8.1.3 SVFit performance

To compare the performance of the SVFit algorithm, the SVFit-reconstructed $\tau\tau$ invariant mass is shown next to the di-tau invariant mass from its visible products in Fig. 43; two distinct samples are considered: a signal samples from $H \to \tau\tau$ and a background sample from $Z \to \tau\tau$. Overall, we can notice an improvement in the $m_{\tau\tau}^{SVFit}$ separation between the two distinct processes when compared to the visible di-tau invariant mass.

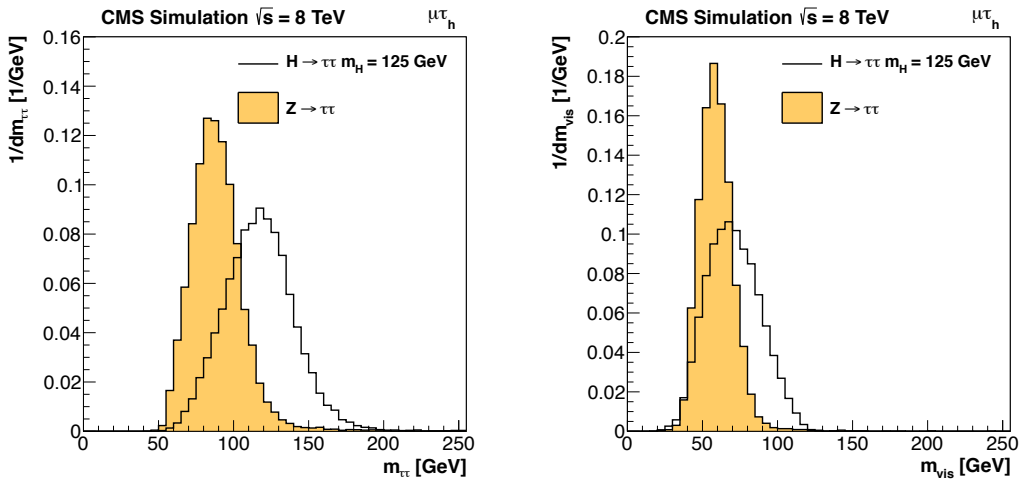


Figure 43: $m_{\tau\tau}$ distributions for $H \to \tau\tau$ (signal) and $Z/\gamma^* \to \tau\tau$ (background) events as reconstructed by the SVFit algorithm (left) and by the visible tau decay products. This plot only shows the channel $\tau\tau \to \mu\tau_H$. Picture from [62].

8.2 Other HH Signatures

On top of the SVFit di-τ invariant mass, there is a number of additional variables that can help us characterise the di-Higgs signature. One of these is the helicity of the $\gamma\gamma$ system, as shown in Fig. 44, where we see that our signal $H\ H$ has a rather flat profile, whereas $\gamma\gamma$ + jets and γ + jets tend to peak around 0.7 and are

highly suppressed at $\cos\theta_{\gamma\gamma} \sim 1$. Another angular variable of interest is the helicity of the $\tau\tau$ system, which, again, is pretty flat for $H\ H$ while it tends to cluster around $\cos\theta_{\tau\tau} \sim 1$ for $\gamma\gamma + \text{jets}$ and $\gamma + \text{jets}$. This variable, however, suffers from low statistics, especially in the background, since we require both tau candidates to be reconstructed. The Collins-Soper angle, $\cos\theta_{CS}$, also differentiates the HH signal with a falling profile towards values of $\cos\theta_{CS}$ close to unity from the other backgrounds. And finally, the helicity between the $\gamma\gamma$ and the $\tau\tau$ objects is also considered, with a falling distribution towards $|\cos\theta|_{\gamma\gamma\tau\tau} \to 1$ for the HH signal and an opposite profile for $\gamma + \text{jets}$ and $\gamma\gamma + \text{jets}$ (falling towards $|\cos\theta|_{\gamma\gamma\tau\tau} \to 0$). The angular properties of the di-Higgs system can be interpreted as a direct consequence of the tensor structure of the HH production mechanism [66, 67], which differs from any other background process.

We also considered the m_X variable defined as:

$$m_X = m_{\gamma\gamma\tau\tau} - \Big(m_{\tau\tau} - m_H\Big) - \Big(m_{\gamma\gamma} - m_H\Big) \tag{64}$$

which is less sensitive to the energy resolution of the $\gamma\gamma$ or $\tau\tau$ systems if the correct candidates from the Higgs decays are selected. This variable is also of particular interest for the BSM searches, since it varies considerably among the distinct EFT scenarios considered, see Fig. 45. For the SM scenario, m_X displays a rather broad spread, partially induced by the selection cuts applied to the reconstructed objects and partially due to the interference between the different Feynman diagrams that contribute to the overall ggf HH production. No requirement is imposed on any of the above variables, but they are instead fed to the BDT as a training input in order

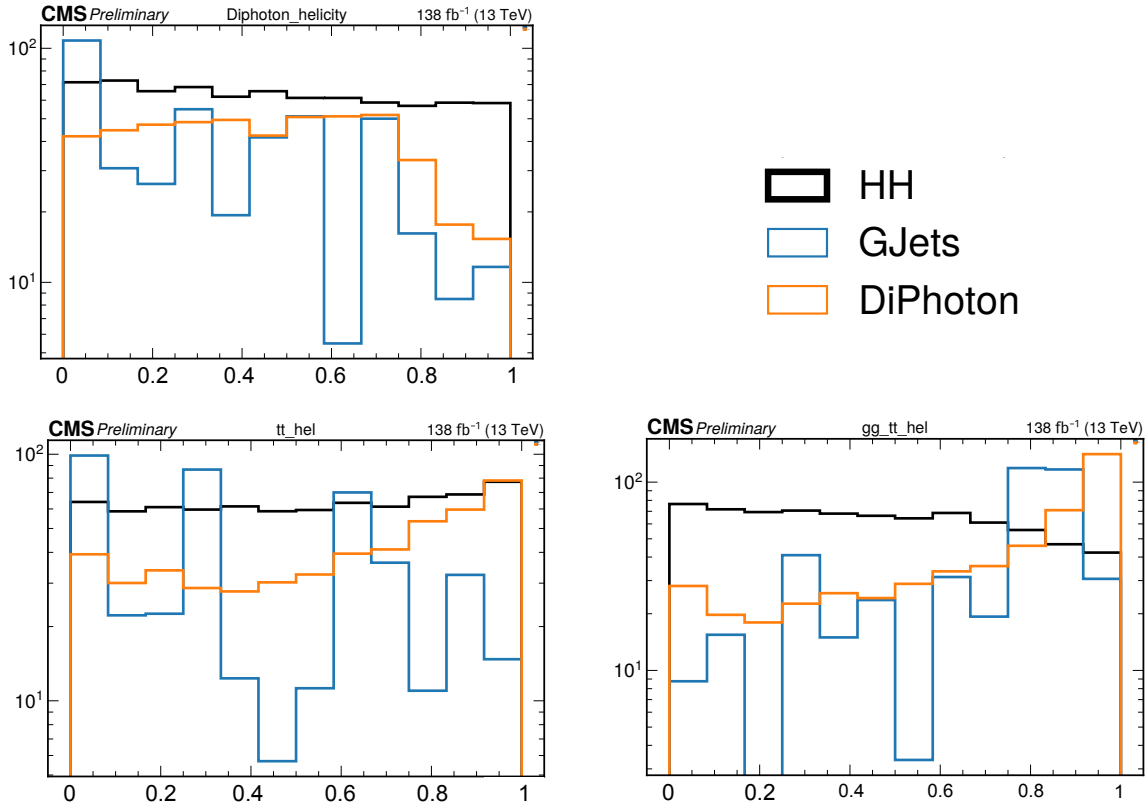


Figure 44: Absolute value of the cosine of the helicity angle for the $\gamma\gamma$ (top left), $\tau\tau$ (bottom left) and $\gamma\gamma - \tau\tau$ (bottom right) systems. The processes show the helicity profiles for our signal HH and backgrounds $\gamma +$ jets and $\gamma\gamma +$ jets, with shapes that are arbitrarily normalised and overlaid on top of each other.

to classify and discern signal from background events.

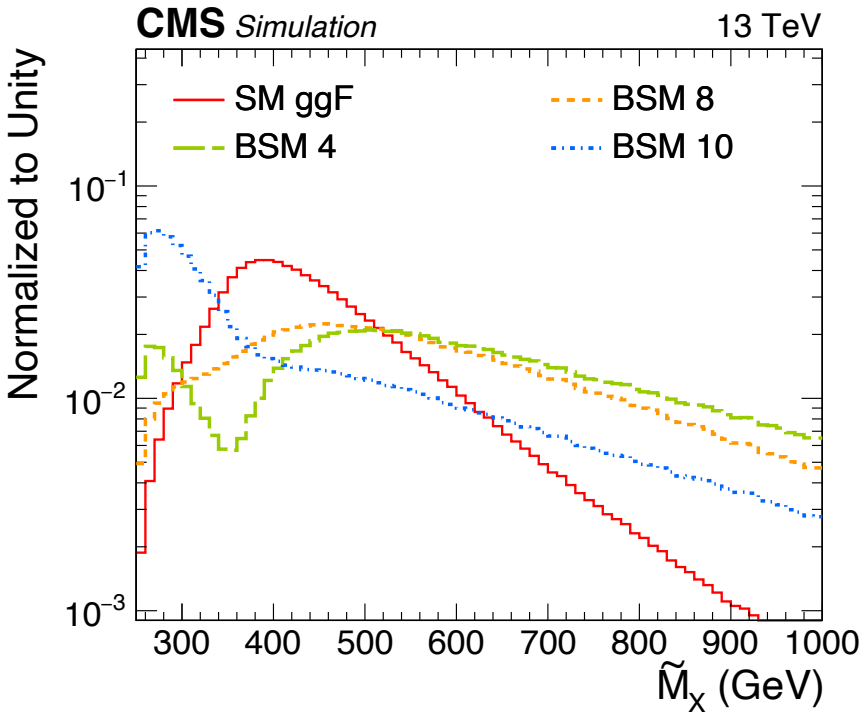


Figure 45: Normalised distribution of m_X for the SM and a number of BSM models from Table 2. The SM di-Higgs pair displays a broader spectrum than other BSM models, along with a very different profile and mean.

9 Multivariate Analysis

Events passing the preselection are then fed to a binary classification boosted decision tree (BDT) algorithm, trained to separate the HH signal from all backgrounds. The output of the BDT is then used to define multiple signal regions that are used for the statistical analysis and limit extraction procedure. The optimization of these signal regions is described in more details in Sec. 10. To avoid any chance of introducing bias in the analysis, we split the MC samples into independent datasets for each of the following steps: BDT training (16.7% of total MC dataset), BDT testing(16.7% of total MC dataset), signal region optimisation(16.7% of total MC

dataset), extraction of the expected upper limits (50% of total MC dataset).

I believe it is important to emphasise that, as the final statistical analysis estimates the nonresonant background through a fit to events in data in the $m_{\gamma\gamma}$ distribution, full knowledge of the nonresonant background composition is not strictly necessary. A simulation-based description is still studied, as other than being useful for developing an optimal analysis strategy, it is used for two purposes:

1. Training & optimizing BDT
2. Optimizing cuts on BDT to form signal regions

With this in mind, we develop a description of the background using MC simulation of the dominant resonant and nonresonant backgrounds.

9.1 Why Decision Forests?

In a time where machine learning (ML) algorithms get more and more complex, it becomes essential to acquire a basic understanding of the chosen MVA algorithm in order to make the most informed decision on which ML classifier best fits the task to solve. The main two choices for most particle physics searches are neural networks (NNs) and decision forests, which are based on a large assembly of Decision Trees (DTs) illustrated in Fig. 46. This thesis is probably not the best place to gather an in depth knowledge about either of these methods, but I do want to justify the reasons behind our choice of a BDT classifier rather than the more sophisticated NNs.

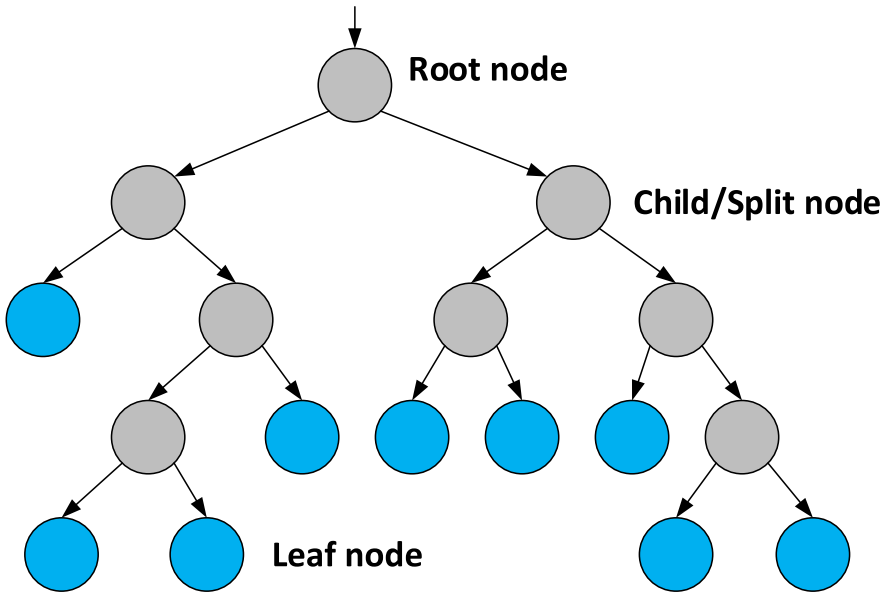


Figure 46: Illustration of a single Decision Tree (DT) [68].

9.1.1 Speed

In experimental particle physics, ML is just a tool to identify a signal region with higher signal purity than in a cut-based analysis. Because of this, we often need an adaptable and fast algorithm that can provide robust predictions even in the case of low statistics samples. Speed is an important factor as we often change the architecture and structure of the classifier before we finalise the analysis. Because of this, we tend to prefer faster training algorithms so that we can quickly check the effects of varying input features or parameters, and Neural Networks tend to have slower times compared to Boosted Decision Trees. The main reason is that NN are (more) complex models with several layers of interconnected nodes (neurons) which generally require more time to train. In addition, each neuron sees its weight and bias updated at each iteration, which is a computationally intensive task. There are a number of ways to obviate to this issue, by training NN on Graphical Processing

Units (GPUs) or FPGAs, but these tasks, although extremely interesting, quickly diverge from the more fundamental physics questions that we are trying to answer in this analysis. Also, NN typically require larger amounts of data, which can also contribute to slower training times. The optimization algorithm used in NNs, such as Gradient Descent, can also slow down the training process, especially for complex models. Finally, the backpropagation algorithm used to update the weights in NNs can also be computationally expensive, adding to the slower training times of NNs as compared to BDTs. To quantify the training time differences between BDTs and NNs, our analysis would require a training time of roughly tens of hours using NNs, instead of a few seconds using decision forests. Overall, BDTs are much simpler models, which are based on simple decision trees that are then combined into a decision forest in order to form a more complex model. The simplicity of the model makes this algorithm much faster to train. BDTs also perform well when trained on statistically limited samples and do not require specialised hardware or computational heavy resources as NN do.

9.1.2 Multicollinearity

Another reason for choosing BDT over NN is the handling of correlated training features. NN are notoriously ”delicate” to the amount, range and form of the input features provided. Because of this, there is usually a lot of pre-processing of the input features that is needed before one can input the training variables to a NN. These include re-normalising all of the inputs in a [0,1] (sometimes [-1,1]) range using sigmoid functions, this reduces the occurrancies of divergencies during the back-propagation step, decorrelation of input variables, since correlated variables can negatively affect the performance of the NN, and many other data processing

techniques. On the other hand, BDTs are much more robust and there is basically no pre-processing needed. By construction, decision trees are immune to multicollinearity. In the case of highly correlated variables, when deciding upon a split the BDT will choose only one of the correlated variables. And since the BDT forest is built from individual decision trees, the resulting classifier is unaffected by multicollinearity. That said, it is in general considered good practice to remove highly correlated variables, which can lead to a loss in performance if using other MVA techniques. However, in our case, there is no such risk.

9.2 Training Samples

During the training process, the BDT uses MC samples exclusively and we define our background processes to be: $\gamma\gamma$ + jets, $t\bar{t}+\gamma\gamma$, $t\bar{t}+\gamma$, $t\bar{t}$, $W+\gamma$, $Z+\gamma$, WW, WZ, ZZ, VH, ggH, $t\bar{t}H$ and VBF H; and our signal to be the SM ($\kappa_\lambda = 1$) samples for $HH \to \gamma\gamma\tau\tau$ and $HH \to \gamma\gamma WW$. We maintain the same configuration also when calculating the 12 EFT benchmark limits or performing the κ_λ scan. A natural question may rise, why do not include the BSM scenarios into the training of the binary classifier? Well the answer is that we did perform some studies investigating the gain in sensitivity by training a BDT to specific BSM points or a combination thereof, and the results were that the gain in sensitivity for the final exclusion limits in the κ_λ scan was minimal. The main reason for this involves the way in which the κ_λ scan is performed, which requires that all of the input samples (e.g. $\kappa_\lambda = 0$, 1, 2.45, 5) need to undergo the exact same selection in order for the re-weighting, described in Sec. 3.1, to be valid. This means that we cannot pick and choose which BDT cut to impose on the different κ_λ samples. For instance, if we trained the BDT on a specific κ_λ point but, given that the di-Higgs kinematics

vary considerably depending on the κ_λ value, the BDT may not perform as well on the other κ_λ samples that were not included in the training, and when performing the κ_λ scan one sample may retain a large efficiency, but the other κ_λ processes may not; so given that the cross-section at each arbitrary value of κ_λ is extracted from the linear combination of three κ_λ samples, this training approach does not necessarily results in an improved limit. Similarly, training on the full set of κ_λ values does not provide a tangible improvement in limits because of the large differences in HH kinematics, which cannot be fully captured by the BDT and at times may overlap with kinematic profiles of background processes. For these reasons, we chose to train the BDT only on the SM samples, and accept the tradeoff in performance for other BSM points. It would be interesting to run a more thorough study of this effect on the κ_λ limits, but that would require additional time and manpower than it is available at this stage of the analysis.

We train one single BDT to reject both the non-resonant and resonant background. Some studies were carried out to investigate whether it would be beneficial to train a dedicated MVA classifier to reject resonant backgrounds (VH, ggH, $t\bar{t}H$ and VBF H), however this investigation was not converging and the gain in sensitivity was minimal, so we decided to proceed with a single BDT also to maintain the analysis as simple as possible. When training the BDT, we assign equal cumulative weights to signal and background, this way the classifier pays equal attention to retaining a considerable fraction of signal while rejecting background. Within each class (signal or background), we assign weights to each event based on the relative cross section and branching ratio of the process considered, so that the relative contributions of the yields at preselection is reflected in the weights assigned in the training step.

9.3 Scaling of the Resonant Background

An additional study was carried out to examine how the BDT performance is affected when we scale the yields of the resonant backgrounds by various factors with respect to the non-resonant background. This concept can be better understood in the context of limit extraction. When we evaluate the significance of our search, we perform a fit to the $m_{\gamma\gamma}$ distribution around the Higgs peak. The Higgs window is defined to be roughly between 120 and 130 GeV. We have to bear in mind that, at the preselection level (when the BDT training is carried out), nonresonant backgrounds include all of the events in the $m_{\gamma\gamma}$ interval $100 < m_{\gamma\gamma} < 180$. If we assume the nonresonant background shape to be roughly a falling exponential in $m_{\gamma\gamma}$, we can then mathematically estimate that the nonresonant background contamination in the Higgs window (where the limit extraction is performed) is scaled down by a factor of 7. In other words, we compare the areas of the falling exponential in the Higgs window over the full $m_{\gamma\gamma}$ range, and this ratio is $\sim 1/7$. For this reason, we consider scaling the resonant backgrounds by different factors that range from 1 to 10. The BDT performance in terms of the upper limits on the HH cross section is summarised in Table 10 for the different scaling factors considered. From these results we see that a resonant scaling factor of 6 gives the lowest upper limit. This is also in good agreement with the expected factor obtained from the back-of-the-envelope calculation above. Among the training background processes, you may have notice the absence of $\gamma +$ jets. The issue here is that $\gamma +$ jets events are notoriously hard to model and their cross section is considerably larger than the other backgrounds. Because of how we normalise background events before training the BDT, the inclusion of $\gamma +$ jets events negatively affects the performance of the classifier as the BDT pays more attention to removing $\gamma +$ jets events (which are not

a dominant background in the signal region where the $m_{\gamma\gamma}$ fit is performed) due to their large weights rather than suppressing other types of background that are much harder to reject around the Higgs window, such as $Z + \gamma$ or resonant backgrounds.

Upper Limits vs of scaling of resonant bkg.

scaling factor	upper limit
1	23.1
2	23.1
3	21.3
4	22.3
5	22.1
6	20.4
7	21.8
8	22.0
9	22.6
10	21.7

Table 10: Upper limits on HH cross section as a function of the resonant background scaling applied during the training of the BDT. An optimal scaling factor of 6 is chosen, since it allows to further suppress resonant processes in the most sensitive signal region without allowing large fractions of non-resonant backgrounds in the least sensitive signal region.

9.4 Input Features

A large number of input variables has been studied, and the most discriminating ones have been selected based on their distinct profiles in the signal / background MC plots. The inputs to the BDT are a variety of high-level variables summarizing the kinematics of each event, and are listed in Table 11. A description of any non-trivial variables follows, and comparisons between data and simulation for each of the input features are shown below and also at the following link.

Other BDT variables not included in the table are:

Input Features to BDTs

Category	Features	
Photon Kinematics	$\gamma_1\ p_T/m_{\gamma\gamma}$	$\gamma_1\ \eta$
	γ_1 ID MVA	γ_1 Pixel Seed Veto
	$\gamma_2\ p_T/m_{\gamma\gamma}$	$\gamma_2\ \eta$
	γ_2 ID MVA	γ_2 Pixel Seed Veto
	min/max γ ID MVA	min/max $\gamma\ p_T/m_{\gamma\gamma}$
DiPhoton Kinematics	$p_T^{\gamma\gamma}/m_{\gamma\gamma}$	$\eta_{\gamma\gamma}$
	$\Delta R_{\gamma\gamma}$	$\lvert\cos(\Delta\phi)_{\gamma\gamma}\rvert$
	$\lvert\cos(\text{helicity angle}(\theta))\rvert$	
τ_H / e / μ / IsoTracks	multiplicity	p_T
	η	
Jets & b-Jets	multiplicity	p_T
	η	b-Tag
	max b-Tag	
Composite Objects	$\Delta\phi(MET,\gamma\gamma)$	$\Delta\phi(MET,\tau_{h/l}\tau_{h/l})$
	$\Delta\phi(\tau^1_{h/l},\tau^2_{h/l})$	$\Delta\phi(\gamma\gamma,\tau\tau^{SVFit})$
	$\Delta\phi(MET,\tau_{h/l})$	$\Delta R(\tau_1^{SVFit},\tau_2^{SVFit})$
	$\Delta R(\gamma\gamma,\tau\tau^{SVFit})$	$\Delta R(\tau^1_{h/l},\tau^2_{h/l})$
	$\Delta\eta(\tau^1_{h/l},\tau^2_{h/l})$	$\Delta\eta(\tau_1^{SVFit},\tau_2^{SVFit})$
	$p_{T\tau\tau}^{SVFit}$	$\eta_{\tau\tau}^{SVFit}$
	$m_{\tau\tau}^{SVFit}$	$p_{T\tau\tau}^{visible}$
	$\eta_{\tau\tau}^{visible}$	$m_{\tau\tau}^{visible}$
	$m_{\tau\tau\gamma^{lead}}$	$m_{\tau\tau\gamma^{sublead}}$
	$helicity(\tau\tau^{SVFit},\gamma\gamma)$	$cos\theta^{CS}(\tau\tau^{SVFit},\gamma\gamma)$
Event-level Kinematics	E_T^{miss}	

Table 11: Features used in training BDTs.

Jet & b-Jet variables:

- bTag: b jet tag of the selected jet
- max b-Tag: maximum b-Tag in the event

Composite objects variables:

- $\Delta\phi(MET,\gamma\gamma)$: $\Delta\phi$ between missing transverse energy (MET) and the diphoton system

- $\Delta\phi(MET, \tau_{h/l}\tau_{h/l})$: $\Delta\phi$ between MET and the $\tau\tau$ system as reconstructed by SVFit
- $\Delta\phi(\tau^1_{h/l}, \tau^2_{h/l})$: $\Delta\phi$ between the two τ SVFit candidates
- $\Delta\phi(\gamma\gamma, \tau\tau^{SVFit})$: $\Delta\phi$ between the $\gamma\gamma$ pair and the $\tau\tau$ system as reconstructed by the SVFit algorithm
- $\Delta\phi(MET, \tau_{h/l})$: angle between MET and τ candidate with leading p_T
- $\Delta R(\tau_1^{SVFit}, \tau_2^{SVFit})$: ΔR between the two τ candidates as reconstructed by the SVFit algorithm
- $\Delta R(\gamma\gamma, \tau\tau^{SVFit})$: ΔR between the $\gamma\gamma$ pair and the $\tau\tau$ system as reconstructed by SVFit
- $\Delta R(\tau^1_{h/l}, \tau^2_{h/l})$: ΔR between the visible τ candidates
- $\Delta\eta(\tau_1^{SVFit}, \tau_2^{SVFit})$: $\Delta\eta$ between the τ candidates as reconstructed by the SVFit algorithm
- $\Delta\eta(\tau^1_{h/l}, \tau^2_{h/l})$: $\Delta\eta$ between the visible τ candidates
- $p_{T\tau\tau}^{SVFit}, \eta_{\tau\tau}^{SVFit}, m_{\tau\tau}^{SVFit}$: quantities of the $\tau\tau$ pair as reconstructed by the SVFit algorithm
- $p_{T\tau\tau}^{visible}, \eta_{\tau\tau}^{visible}, m_{\tau\tau}^{visible}$: quantities of the visible $\tau\tau$ system
- $m_{\tau\tau\gamma^{lead}}, m_{\tau\tau\gamma^{sublead}}$: invariant mass of the di-tau pair (SVFit) and the leading (or subleading) γ
- $helicity(\tau\tau^{SVFit}, \gamma\gamma)$: absolute value of the cosine of the helicity angle between the $\gamma\gamma$ pair and the $\tau\tau$ system as reconstructed by the SVFit algorithm
- $cos\theta^{CS}(\tau\tau^{SVFit}, \gamma\gamma)$: absolute value of the cosine of the Collins-Soper an-

gle between the $\gamma\gamma$ pair and the $\tau\tau$ system as reconstructed by the SVFit algorithm

While the $m^{visible}$ is computed as the invariant mass of the visible decay products from τ decays, the $m_{\tau\tau}^{SVFit}$ quantity refers to the $\tau\tau$ invariant mass as reconstructed by the SVFit algorithm [62]. The Collins-Soper angle $\cos\theta^{\mathrm{CS}}(\tau\tau^{\mathrm{SVFit}}, \gamma\gamma)$ is discussed in more details in [69], but it can be approximately understood as the angle between the direction of the $H \to \gamma\gamma$ candidate to the Collins-Soper frame of reference, which assumes each incoming particle to have 6.5 TeV in the scattering (under Run 2 operating conditions).

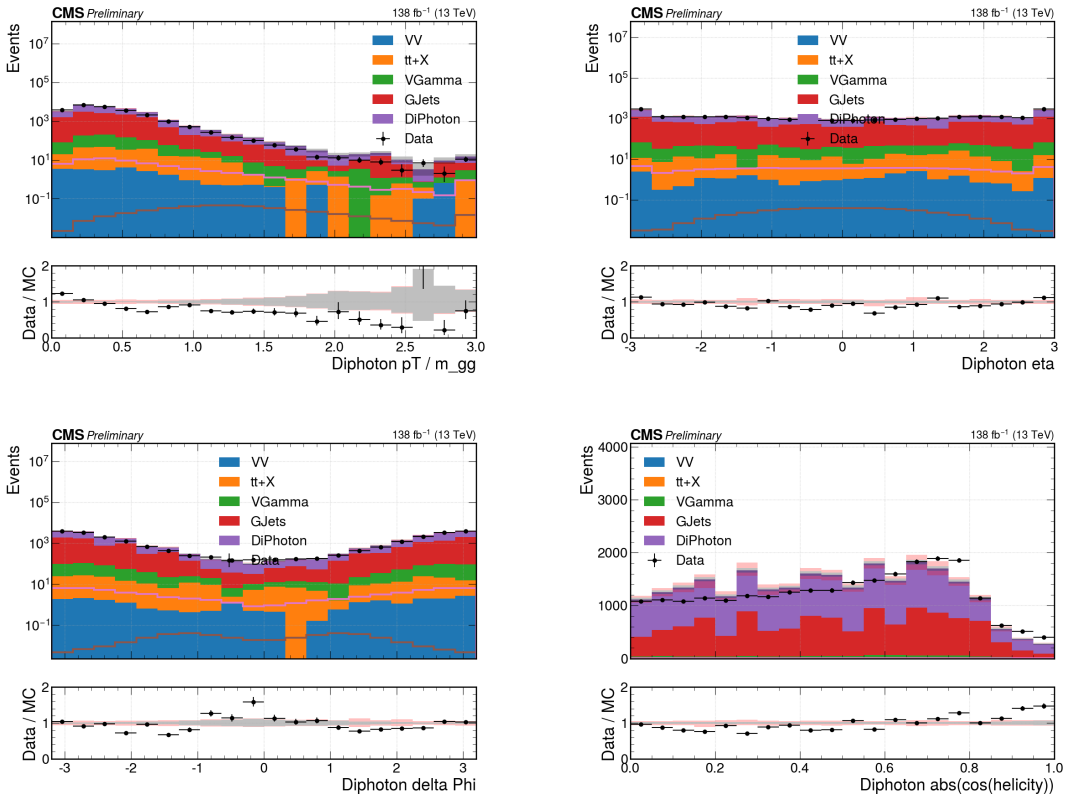


Figure 47: Data / MC comparison for the BDT input features.

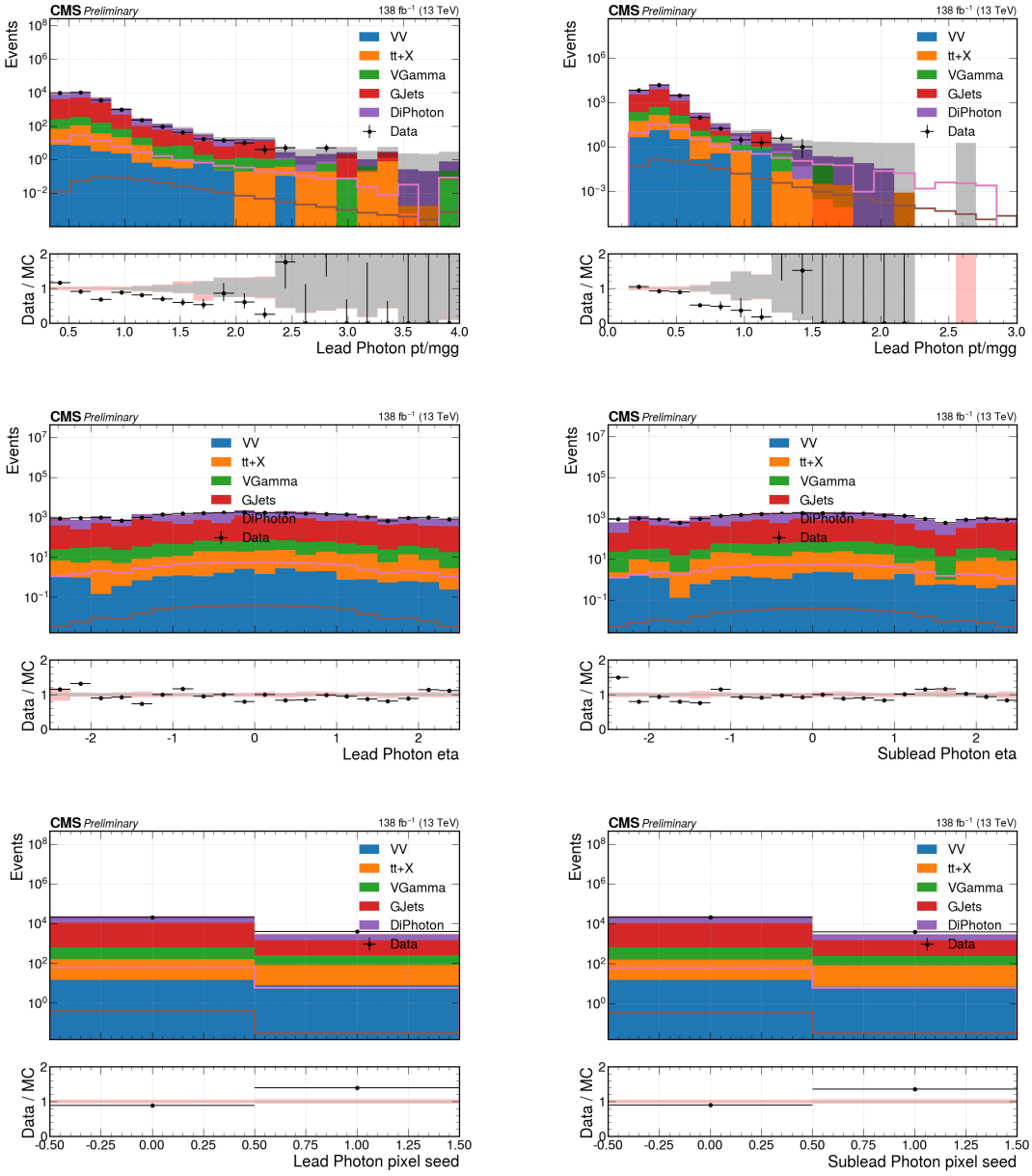


Figure 48: Data / MC comparison for the BDT input features.

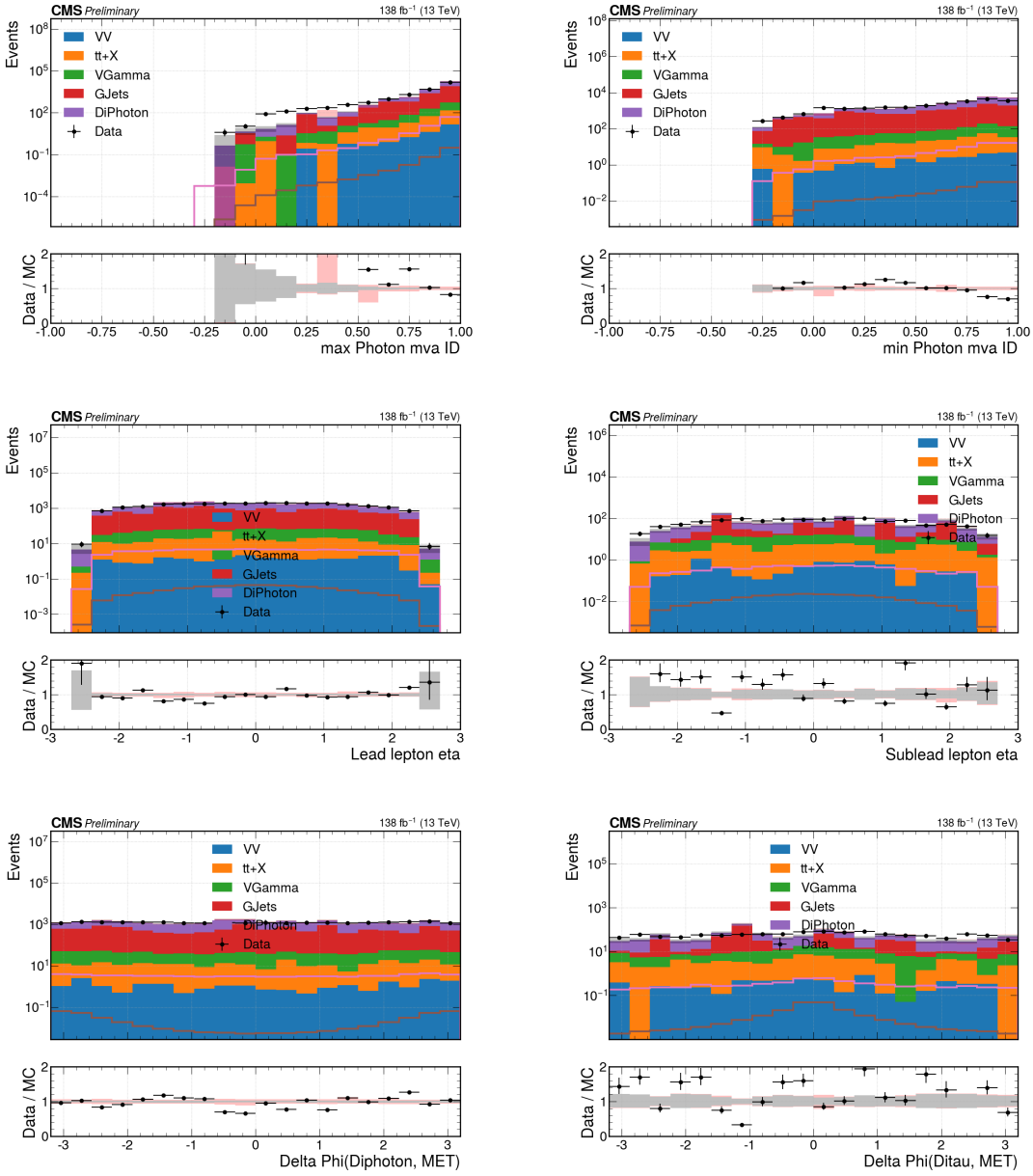


Figure 49: Data / MC comparison for the BDT input features.

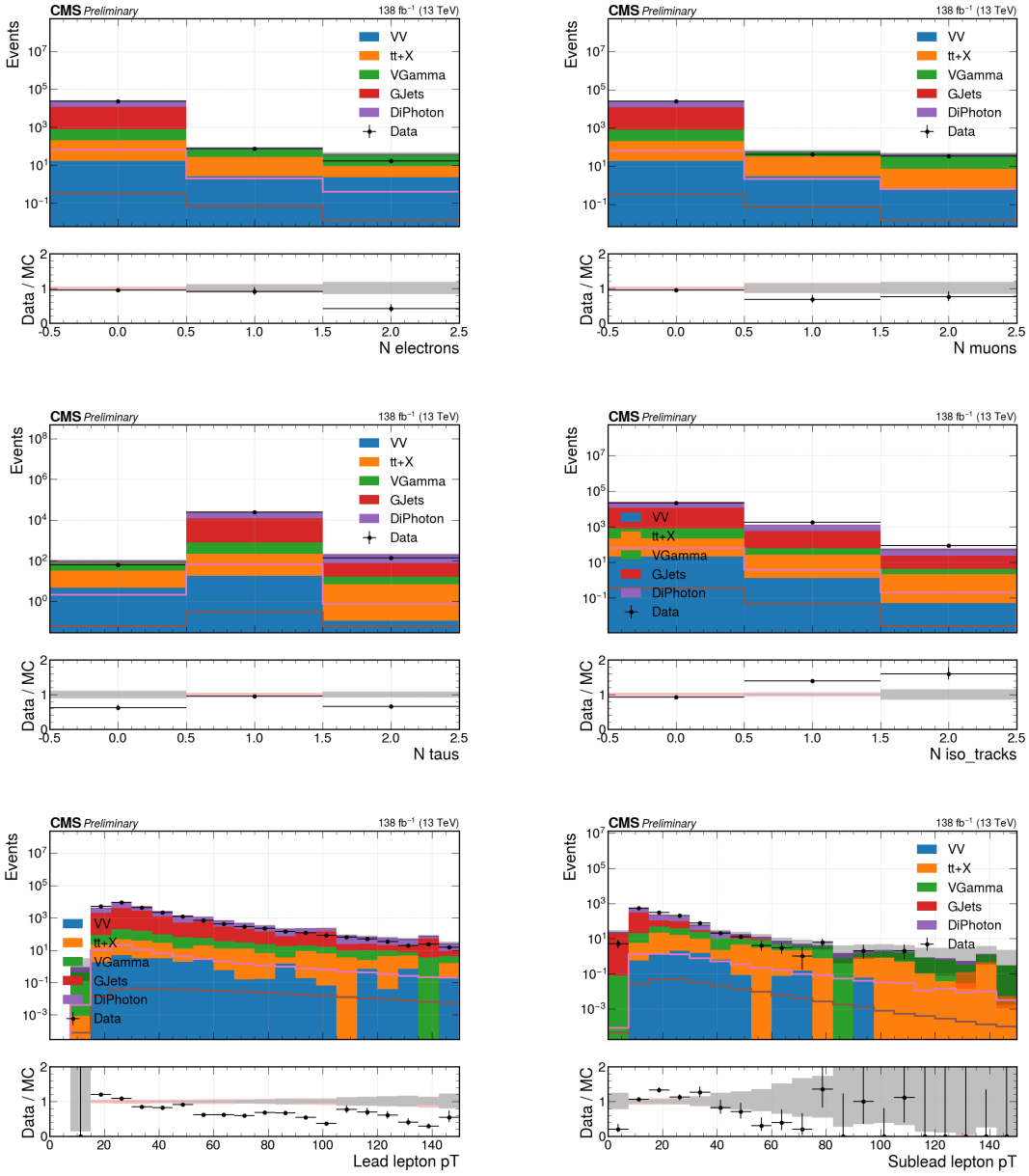


Figure 50: Data / MC comparison for the BDT input features.

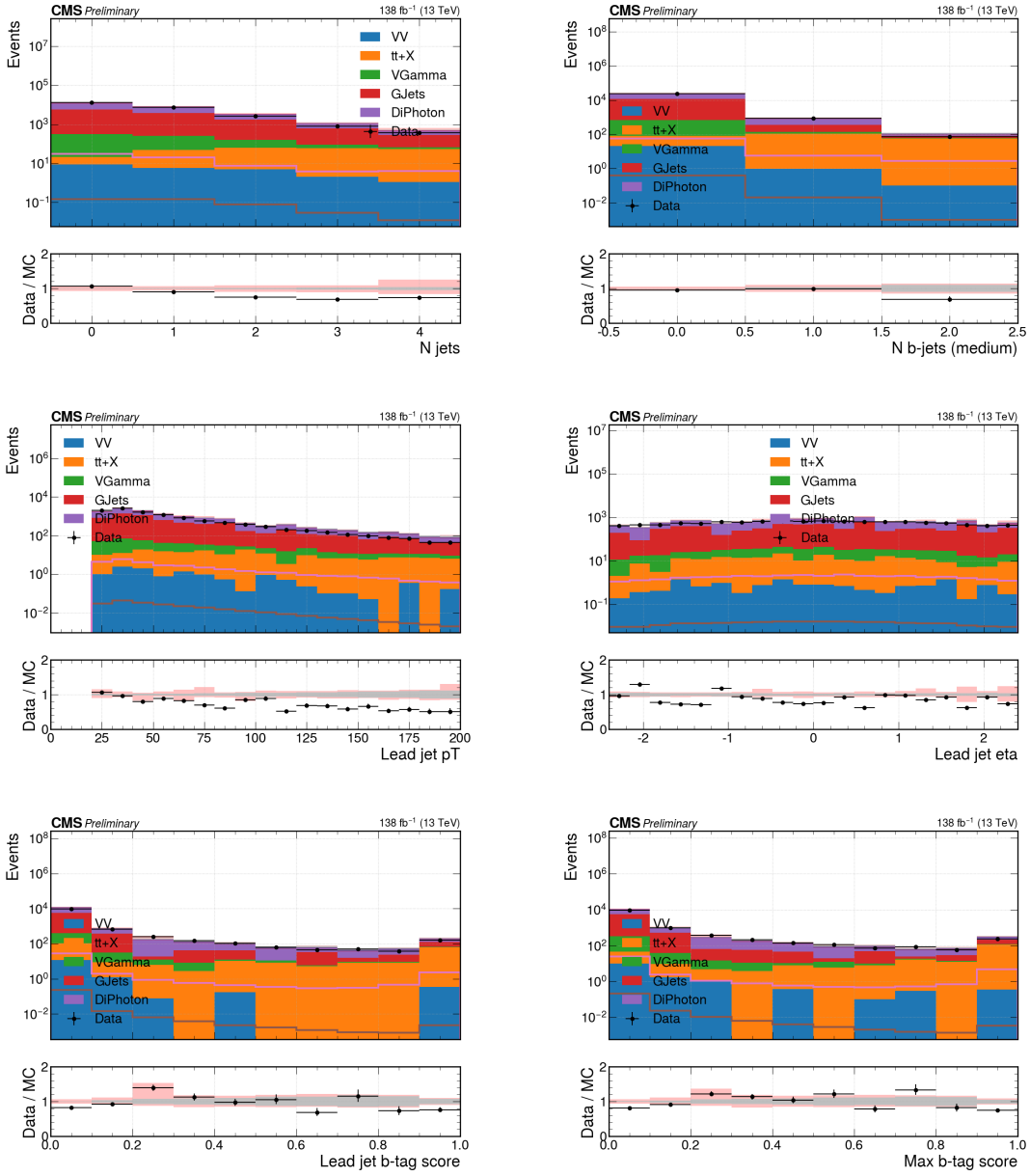


Figure 51: Data / MC comparison for the BDT input features.

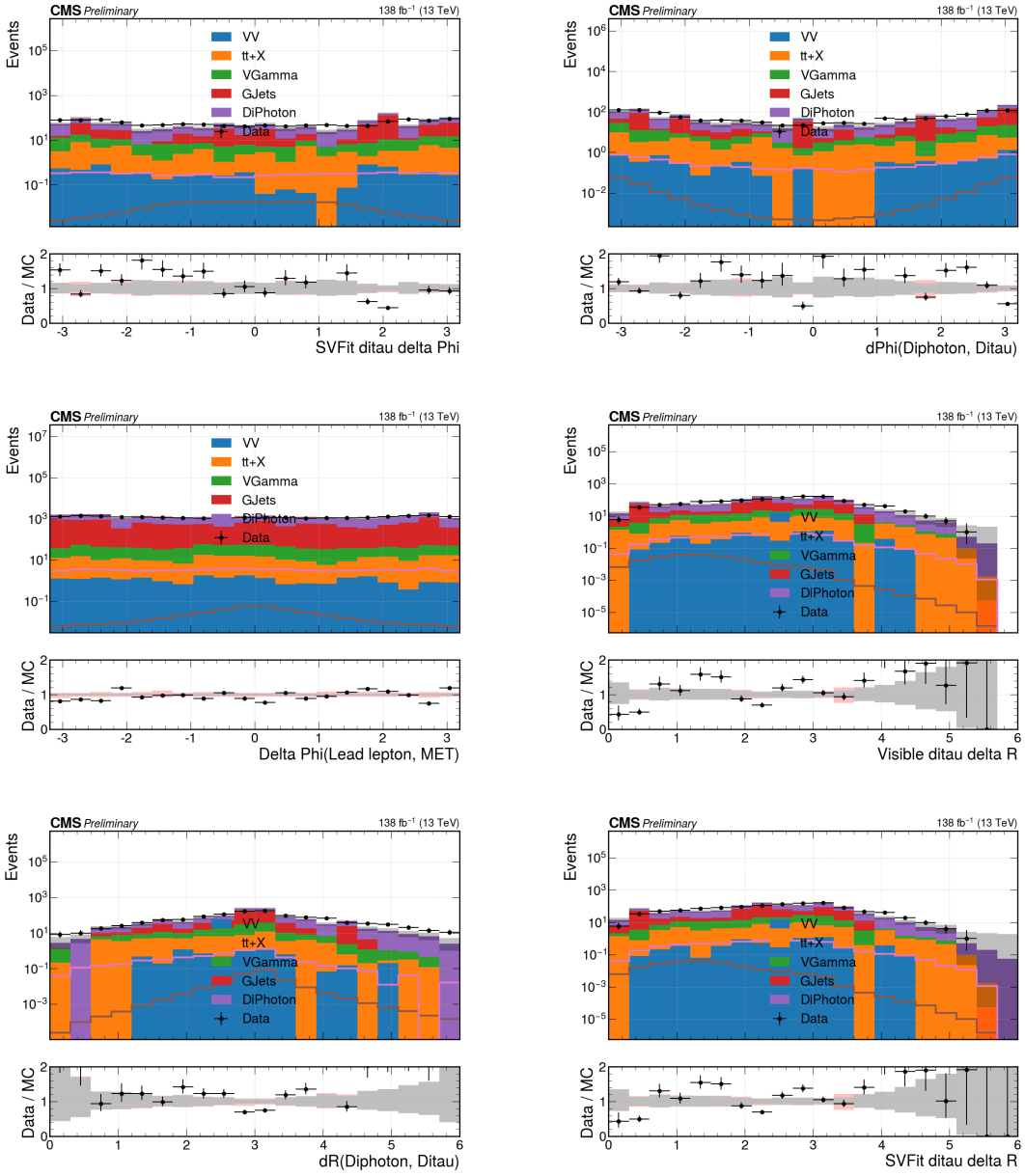


Figure 52: Data / MC comparison for the BDT input features.

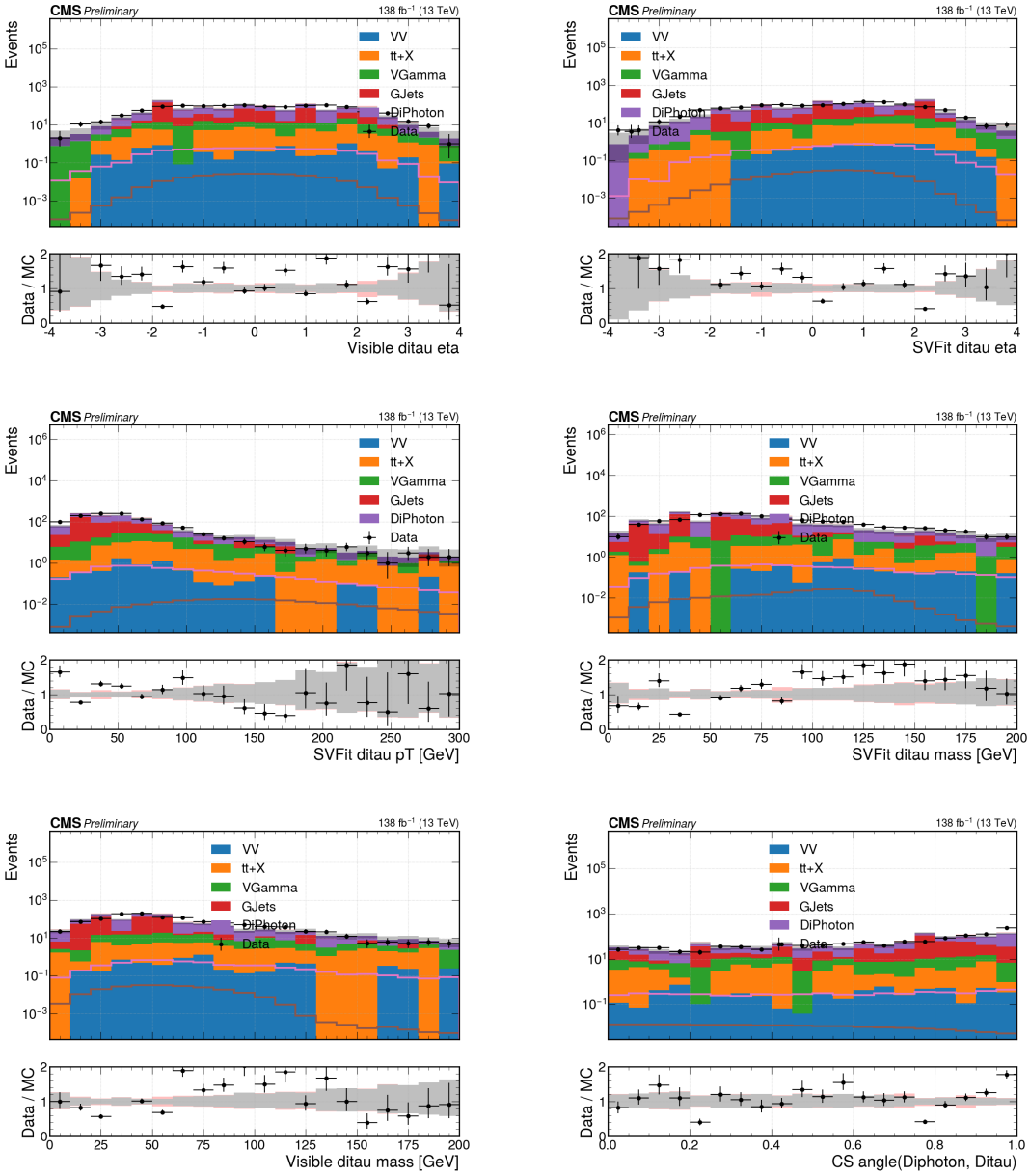


Figure 53: Data / MC comparison for the BDT input features.

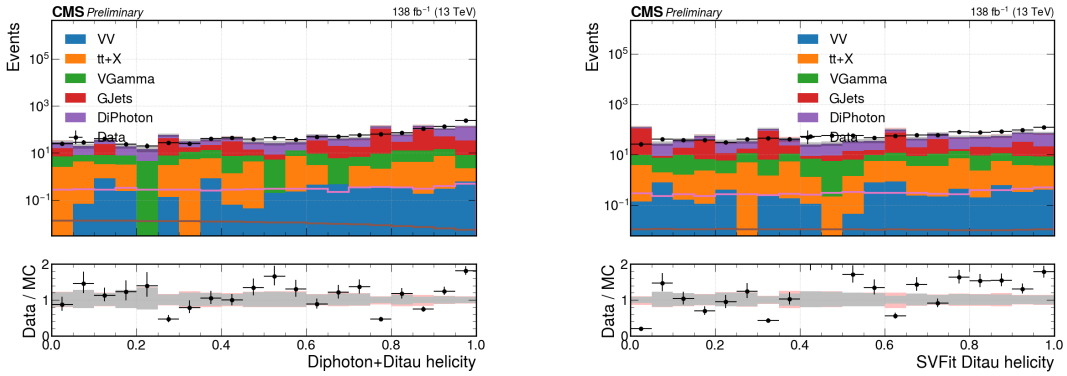


Figure 54: Data / MC comparison for the BDT input features.

You may have noticed in these plots the agreement between Data and Monte Carlo is not always excellent. This is a known fact and it does not necessarily represent an issue, see the discussion at the beginning of Sec. 9. The main reason behind the poor agreement is the intentionally "loose" selection for photons and tau candidates, which allows the presence of objects/events which are not particularly well reconstructed or modeled in MC in order to have enough statistics to train the BDT and optimise signal regions with statistically independent datasets. We must notice that we are not interested in a perfect modeling of such loose objects / events, because these do not enter our most sensitive signal regions anyway, and are easily rejected by our BDT. If we compare the above plots with the same distributions in the signal regions, see Fig. 62 and Fig. 63, we notice that in the latter the agreement is much improved. Additionally, since we model our background from Data in the extraction of limits, any discrepancy between Data and MC may only affect the optimisation of the analysis but it does not in any way affect the validity of the final results.

9.5 Training Details & Performance

The BDT is trained with the `xgboost` [70] framework and the loss function used to find the optimal structure of the decision forest is the binary cross-entropy function:

$$L(y, \hat{y}) = -y \log(\hat{y}) - (1 - y) \log(1 - \hat{y}), \tag{65}$$

where $\hat{y}$ is the BDT prediction of the label y, defined as 1 for signal events and 0 for background events.

The number of trees is determined through an early-stopping procedure in which

training is terminated when the loss on the testing set ceases to improve for 5 rounds. The technique prevents any form of overtraining to arise in the final model. The hyperparameters for the BDT are listed below:

- Maximum depth: 5
- Learning rate: 0.2

The distribution of the BDT output is shown for data and simulation in Fig. 55.

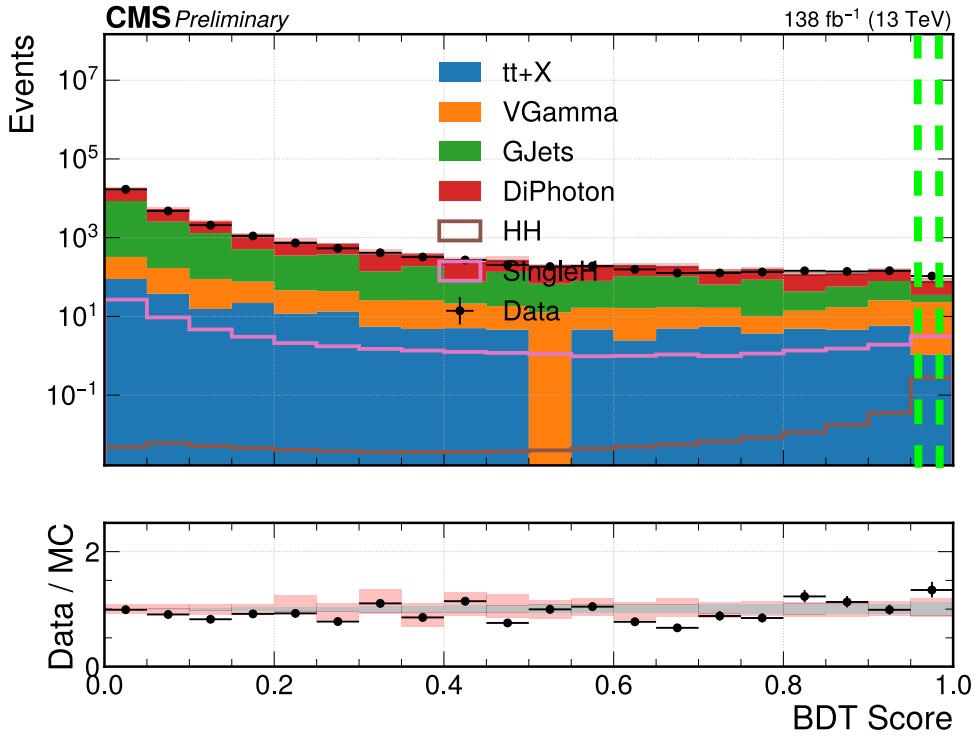


Figure 55: BDT distribution for events from data and simulation passing the preselection. Events from data (black points) and nonresonant backgrounds (stacked histograms) are blinded in the region $m_{\gamma\gamma} \in [120, 130]$ GeV. The dashed lines show the boundaries of each signal region.

The gain of the top 20 training features is showed in Fig. 56, where *gain* in XGBoost relates to the relative contribution of the corresponding feature to the model calcu-

lated by taking each feature's contribution for each tree in the model. A high value of this metric when compared to another feature implies it is more important in generating a prediction. The way in which variables are ranked is not always transparent when using MVA tools, and should be taken with a pinch of salt. For this reason, to gain a better understanding of which are the most important variables in separating signal from background, we ran a study in which we would remove a number of training variables and evaluate the BDT performance. So, one at the time, we remove the main sets of variables from Table 11 and compute the upper limits (UL) on the HH cross section after optimising the signal regions, where the latter procedure is described in detail in Sec. 10. The results from the study are summarised in Table 12, where we see that the γ and $\gamma\gamma$ variables drive the sensitivity of the classifiers, exploiting the $H \to \gamma\gamma$ signature and high p_T of the massless vector bosons, then follow the τ_H and lepton variables and eventually the jet-related variables have the least impact on the separation of signal from background (with the exception of the $t\bar{t}H$ rejection).

Input Variables Removed	Upper Limit Degradation
Photon & Diphoton	+ 21%
τ_H / Lepton	+ 11%
Jet	+ 7 %

Table 12: Degradation of the final upper limit on the HH cross section, as a function of the set of input variables removed during the BDT training. Higher upper limits result in less stringent phase space constraints.

The correlation matrix between BDT input features for signal (background) are

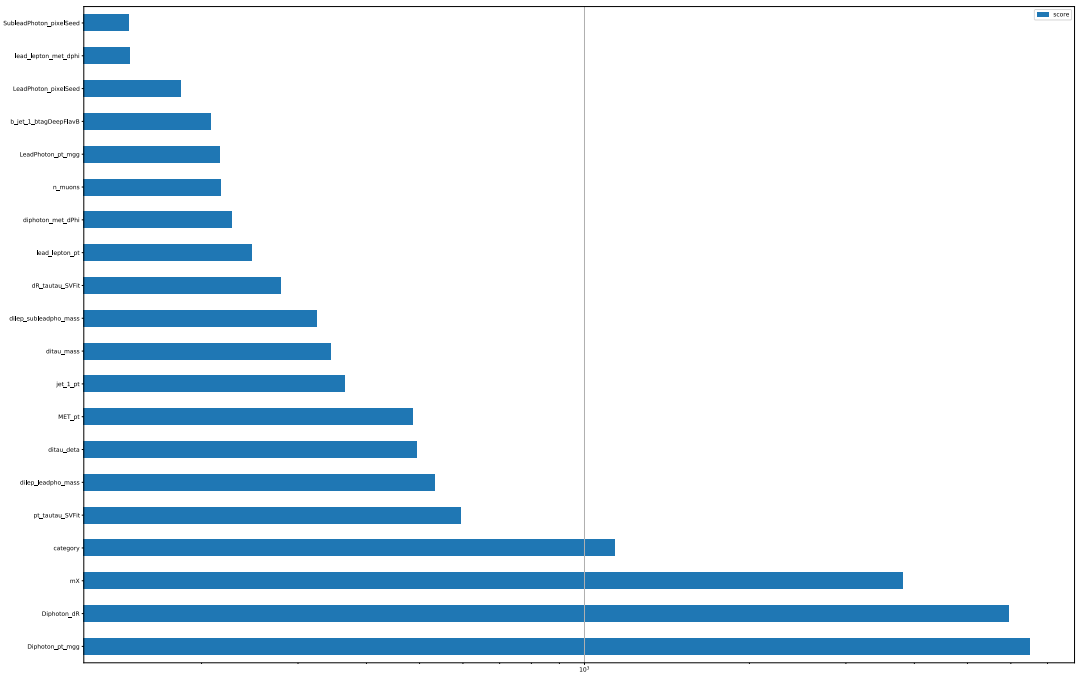


Figure 56: Gain for each of the top 20 most discriminating variables used to train the BDT.

shown in Fig. 57 (Fig. 58).

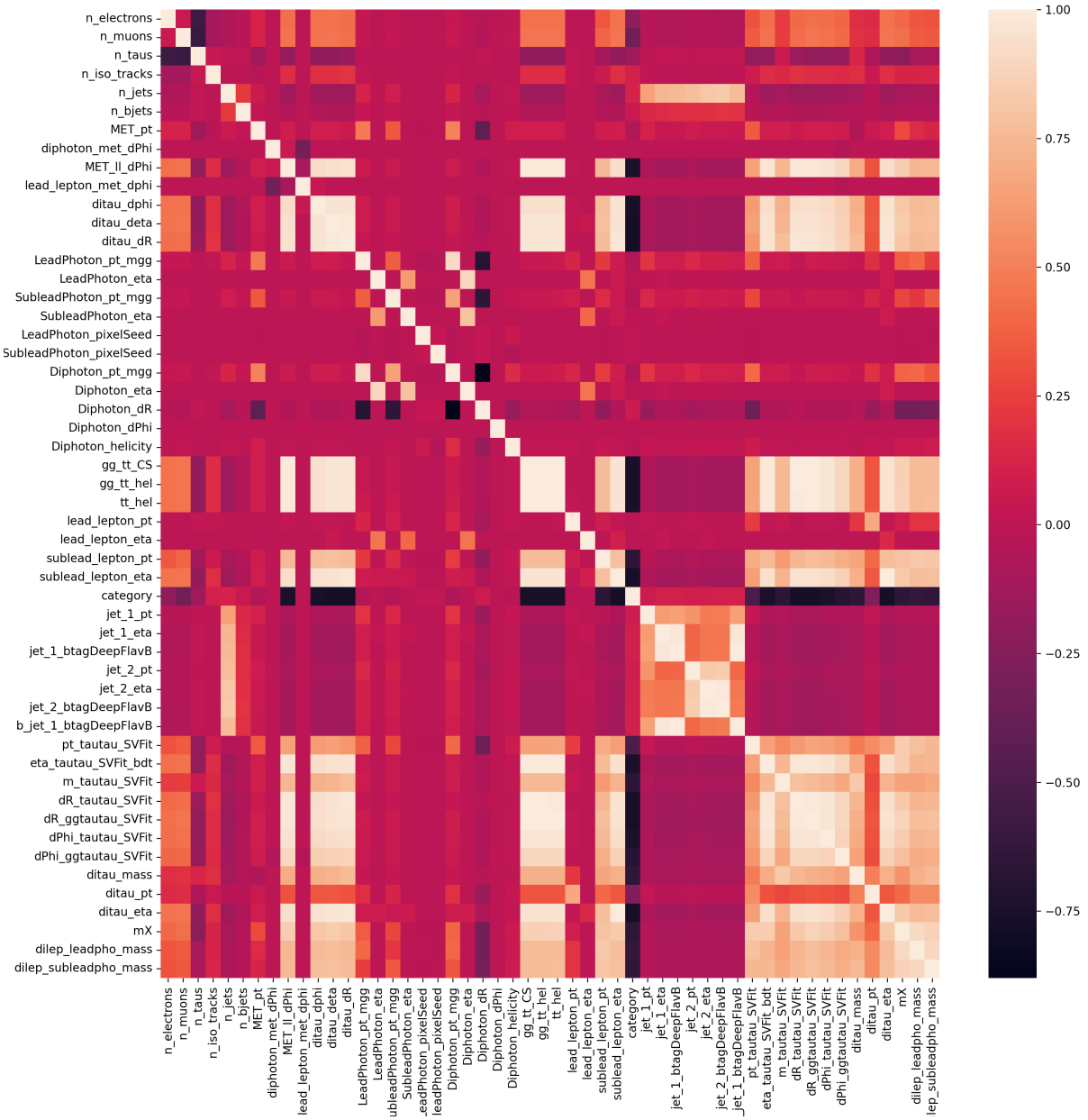


Figure 57: Correlation matrix for the BDT input features of the signal samples.

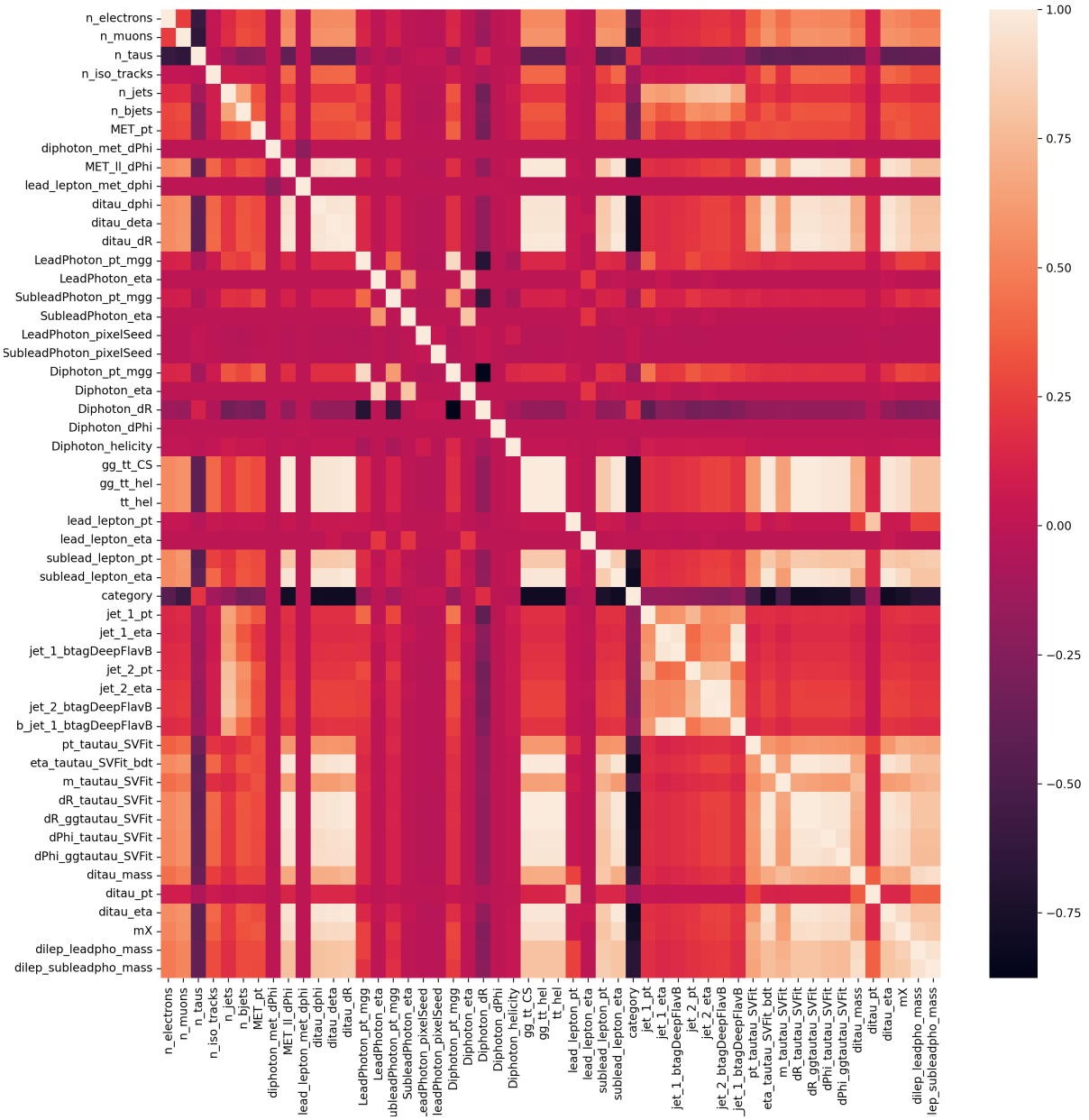


Figure 58: Correlation matrix for the BDT input features of the resonant & non-resonant background samples.

10 Signal Region Optimisation

Once the selection BDT has been trained, we use its score profile to define two signal regions (SRs). The SRs are chosen such that they minimise the expected upper limit on the HH cross section according to the following procedure.

10.1 Procedure

During this step both signal and background processes are exclusively modeled from MC, with the background yields being scaled to match the Data yields at pre-selection level, so as to improve the overall modeling of background. Then the entire BDT score spectrum is divided and scanned in two bins: signal region one (SR1) with a tighter BDT score and signal region two (SR2) with a looser BDT score. Once we have the SR boundaries to test, we perform a simplified fit of the $m_{\gamma\gamma}$ distribution. In performing the fit for signal and resonant backgrounds we use a Double Crystal Ball (DCB) function, instead of the sum of a DCB and Gaussian (as it will be done later), and all three years are combined into a single fit, since we do not take into account systematics variations whose effect may vary depending on the data taking period. Additionally, all resonant background processes are fit together with a single function rather than with separate fits for each production mode, and this is also done for all signal processes. Non-resonant backgrounds are fit with a single exponential function, rather than the variety of functional forms used in the discrete profiling method. During the optimising procedure, we blind the Higgs mass window in data and in the non-resonant background. The $m_{\gamma\gamma}$ fits are performed in the interval $120 < m_{\gamma\gamma} < 130$ GeV for signal and resonant backgrounds and in the interval between $100 < m_{\gamma\gamma} < 120$ GeV & $130 < m_{\gamma\gamma} < 180$

GeV for non-resonant backgrounds. The blinding of data is needed since we require a minimum number of events in the fitting region for the statistical modeling of the processes considered to be reliable. Plots of the simplified fits are shown in Fig. 59 for SR1 and in Fig. 60 for SR2.

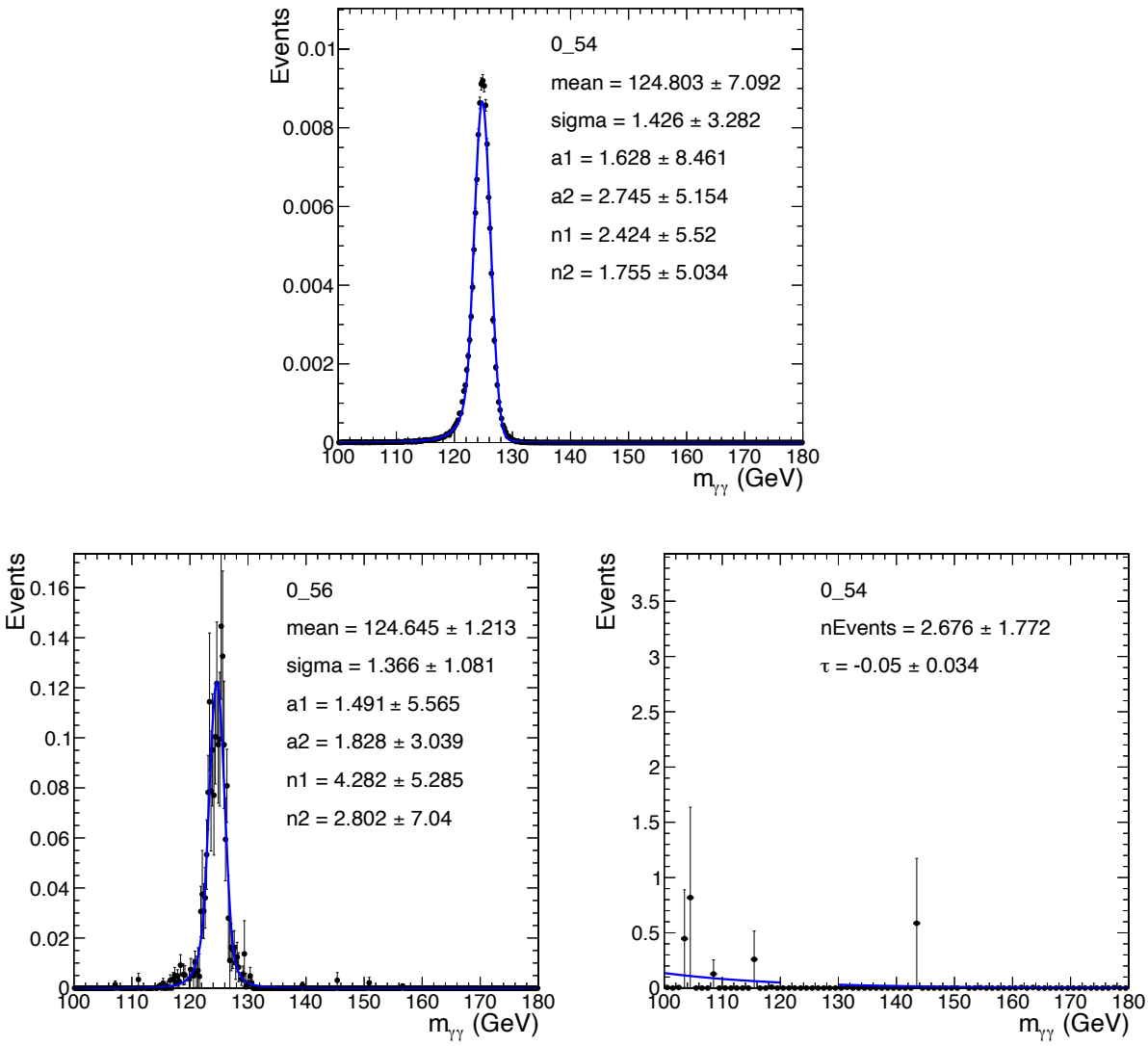


Figure 59: Simplified $m_{\gamma\gamma}$ fits combining all signal (top), resonant background (bottom left) and non-resonant background (bottom right) processes in the optimised SR1. These fits will be later used to extract the expected number of events in the Higgs mass window and run a counting experiment as a preliminary proxy of the sensitivity of the chosen SR boundaries.

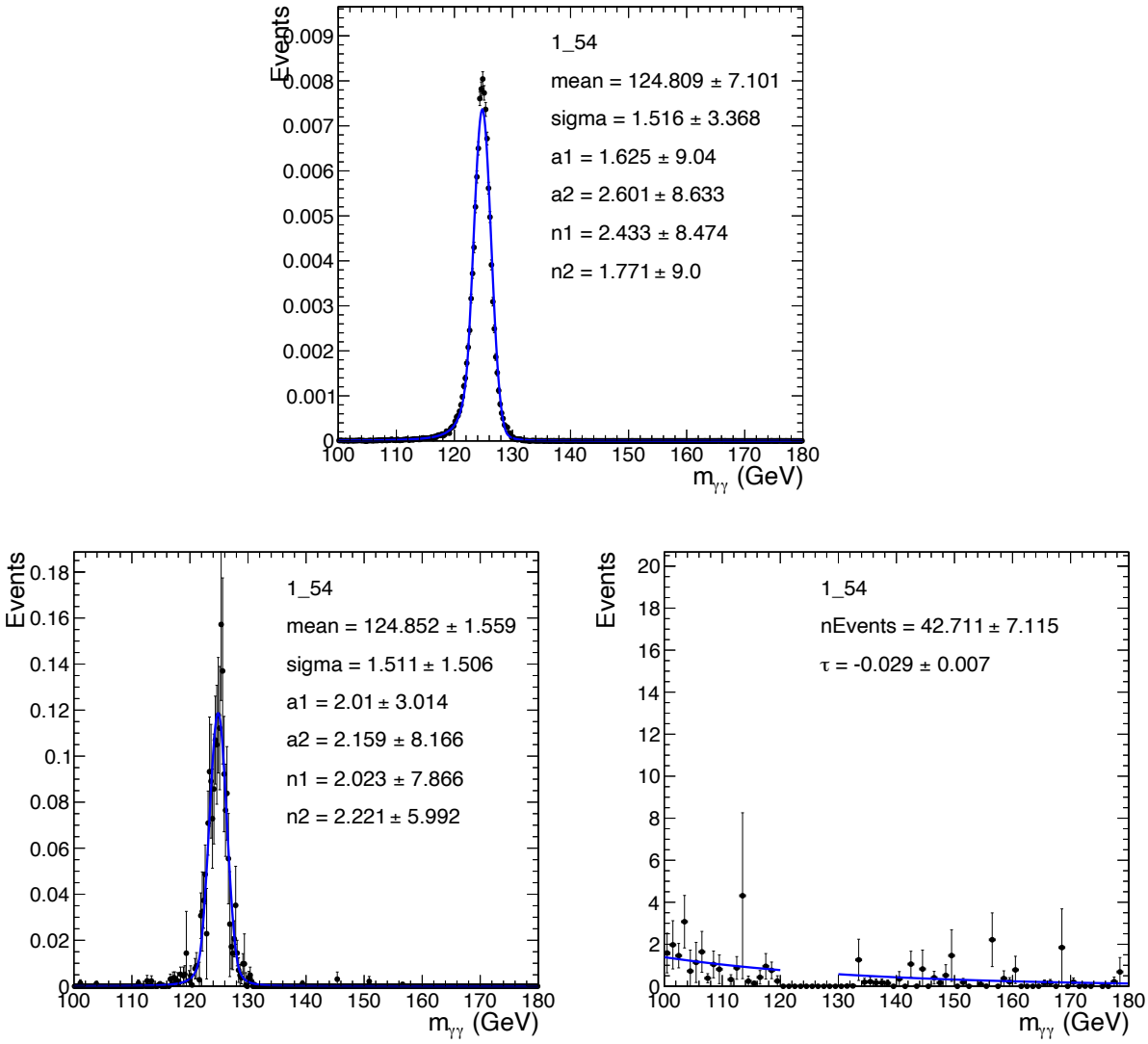


Figure 60: Simplified $m_{\gamma\gamma}$ fits combining all signal (top), resonant background (bottom left) and non-resonant background (bottom right) processes in the optimised SR2. These fits will be later used to extract the expected number of events in the Higgs mass window and run a counting experiment as a preliminary proxy of the sensitivity of the chosen SR boundaries.

Once the fits are performed for a given set of SR boundaries, we extract the expected upper limit (UL) on the HH cross section by calculating the number of events from signal, resonant backgrounds and non-resonant backgrounds. Using

these, we run a simplified counting experiment which adopts the confidence levels (CLs) criterion along with the profile likelihood ratio as the test statistic. Since the test statistics profiles are analytically computed using the asymptotic approximation, computing-heavy tasks such as toy throwing and fitting procedures are not needed in this approach. The expected limit on the HH cross section we obtain in this way is to be interpreted under the background only hypothesis. During this procedure, we require the presence of at least 8 data events in each SR for the tested boundaries to be accepted. This requirement prevents the algorithm to select signal regions where the fit to the background model is untrustworthy because of its limited statistics. Once the expected UL are computed for all scanned boundaries of SRs, we choose the SR definition that minimises the upper limit on the HH cross section.

10.2 Number of Signal Regions

As an effort to optimise the analysis strategy, we ran a study to evaluate the change in expected sensitivity as a function of the number of signal regions. The results of such study are summarised in Table 13, where we express the UL on the cross section as a multiplicative factor of the HH cross section as predicted in the SM:

Scan of signal sensitivity

Number of SRs	Upper Limit (x SM)
1	22.4
2	21.8
3	21.6
4	21.6

Table 13: Expected signal sensitivity vs number of signal regions considered.

Bearing in mind that the above numbers do not include a complete treatment of scale factors, hence do not necessarily reflect the final sensitivity, we can still use

the above table to identify a considerable improvement in performance between one and two SRs, while the sensitivity saturates past two SRs with marginal gain using three or four SRs. Because of this fact and the increase in computing time for the optimisation of more than two SRs, we choose to use two signal regions in this analysis.

10.3 Signal Region Characterisation

The distribution of the BDT output in the two signal regions is shown for data and simulation in Fig. 61, with the green lines indicating the cuts defining the signal regions.

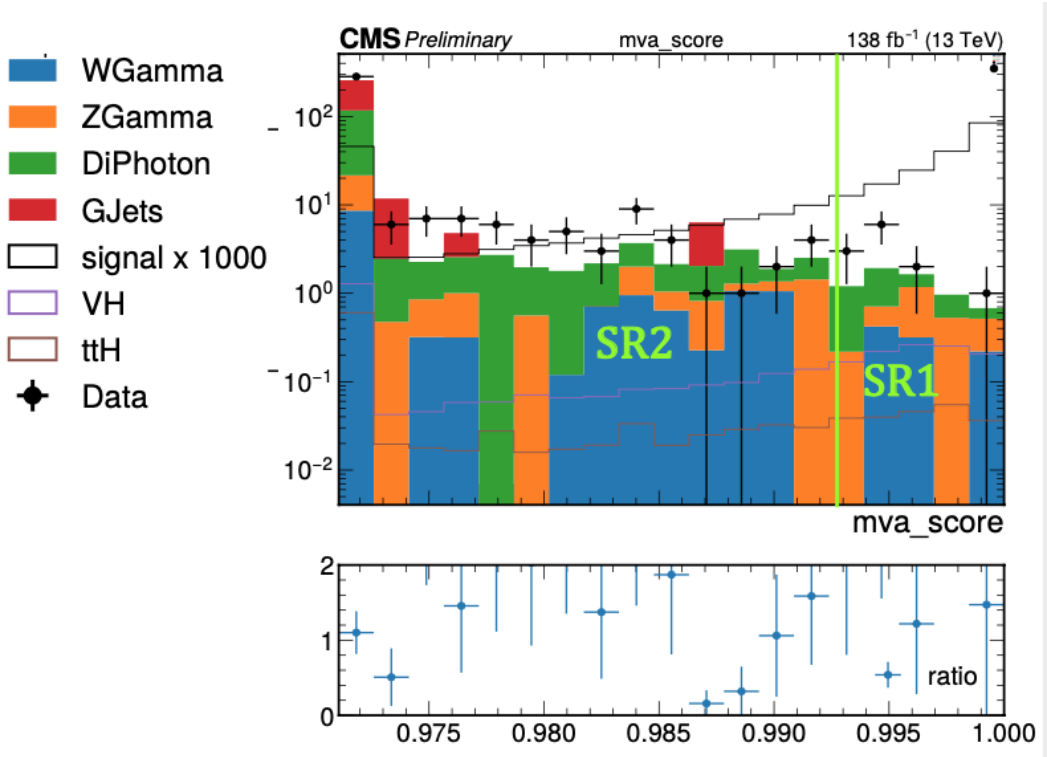


Figure 61: BDT distribution for events from data and simulation passing the signal region selection. Events from data (black points) and nonresonant backgrounds (stacked histograms) are blinded in the region $m_{\gamma\gamma} \in [120, 130]$ GeV. The dashed lines show the boundaries of each signal region.

The yields in the signal regions for signal, background, and data are shown in Tables 14, 15, 16 below. We notice that SR1 is our most sensitive region, with the highest signal over background ratio. Here the dominant backgrounds are $Z + \gamma$ and $\gamma\gamma +$ jets processes. SR2 instead has a lower sensitivity and much larger contamination from background processes. Combining both signal regions, we retain slightly less than 0.25 signal events, for an overall efficiency of about 50% from the $\gamma\gamma\tau\tau$ preselection yields. In the yields tables presented here, data and nonresonant backgrounds are blinded in the $m_{\gamma\gamma} \in [120, 130]$ GeV region.

Signal Region Yields: Inclusive

Process	Yield	Stat. unc.	Syst. unc.	$\mathcal{F}$ of bkg
HHggWW-semileptonic	0.023	$\pm$0.000	$^{+0.004}_{-0.004}$	0.00
HHggWW-dileptonic	0.051	$\pm$0.000	$^{+0.009}_{-0.009}$	0.00
HHggTauTau	0.162	$\pm$0.000	$^{+0.029}_{-0.027}$	0.00
VBFH-M125	0.07	$\pm$0.01	$^{+0.01}_{-0.01}$	0.00
ggH-M125	0.09	$\pm$0.03	$^{+0.02}_{-0.02}$	0.00
ttH-M125	0.35	$\pm$0.02	$^{+0.03}_{-0.05}$	0.01
VH-M125	1.53	$\pm$0.03	$^{+0.14}_{-0.14}$	0.06
tt+X	-1.4	$\pm$1.5	$^{+0.6}_{-0.2}$	-0.05
GJets	2.7	$\pm$2.1	$^{+1.8}_{-0.4}$	0.11
WGamma	2.7	$\pm$0.8	$^{+0.3}_{-0.4}$	0.11
ZGamma	7.2	$\pm$1.1	$^{+0.8}_{-0.7}$	0.28
DiPhoton	15.0	$\pm$0.9	$^{+1.0}_{-1.2}$	0.57
Total MC bkg	26.5	$\pm$3.2	$^{+2.4}_{-1.5}$	1.00
Data	33.0	$\pm$5.7	$^{+0.0}_{-0.0}$	1.25

Table 14: Signal, resonant background, nonresonant background, and data yields in the signal regions (both signal regions combined).

Compared to preselection, we can notice that the overall data/MC agreement is much improved in the SRs, see Fig. 62 and Fig. 63. This is an effect of the tighter cuts applied on the events the signal regions, which remove a large fraction of poorly reconstructed events from data. However, some discrepancies are still present mainly

Signal Region 1 Yields

Process	Yield	Stat. unc.	Syst. unc.	$\mathcal{F}$ of bkg
HHggWW-semileptonic	0.013	$\pm$0.000	$^{+0.002}_{-0.002}$	0.00
HHggWW-dileptonic	0.042	$\pm$0.000	$^{+0.008}_{-0.007}$	0.00
HHggTauTau	0.122	$\pm$0.000	$^{+0.021}_{-0.020}$	0.00
VBFH-M125	0.01	$\pm$0.00	$^{+0.00}_{-0.00}$	0.00
ggH-M125	0.02	$\pm$0.01	$^{+0.01}_{-0.01}$	0.00
ttH-M125	0.15	$\pm$0.01	$^{+0.01}_{-0.02}$	0.02
VH-M125	0.89	$\pm$0.03	$^{+0.08}_{-0.09}$	0.11
GJets	0.0	$\pm$0.0	$^{+0.0}_{-0.0}$	0.00
tt+X	0.8	$\pm$0.5	$^{+0.3}_{-0.2}$	0.10
WGamma	1.5	$\pm$0.6	$^{+0.2}_{-0.2}$	0.18
DiPhoton	2.8	$\pm$0.4	$^{+0.3}_{-0.3}$	0.34
ZGamma	3.0	$\pm$0.7	$^{+0.4}_{-0.4}$	0.37
Total MC bkg	8.2	$\pm$1.1	$^{+0.6}_{-0.6}$	1.00
Data	10.0	$\pm$3.2	$^{+0.0}_{-0.0}$	1.22

Table 15: Signal, resonant background, nonresonant background, and data yields in signal region 1 (SR 1).

due to the limited statistics of the γ + jets sample and its large event weights.

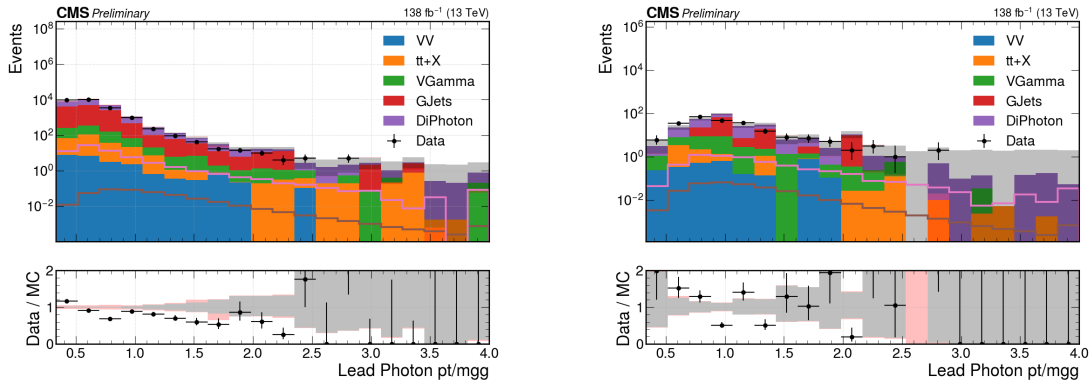


Figure 62: Comparison of the $p_T/m_{\gamma\gamma}$ for the leading photon at the preselection level (left) and in the signal regions (right). In the SR plot, we notice an improved modeling and overall better agreement between Data and Monte Carlo. The disagreement with simulation is mainly due to the limited statistics and high weights of the γ + jets sample.

Signal Region 2 Yields

Process	Yield	Stat. unc.	Syst. unc.	$\mathcal{F}$ of bkg
HHggWW-semileptonic	0.011	$\pm$0.000	$^{+0.002}_{-0.002}$	0.00
HHggWW-dileptonic	0.009	$\pm$0.000	$^{+0.002}_{-0.002}$	0.00
HHggTauTau	0.040	$\pm$0.000	$^{+0.008}_{-0.007}$	0.00
VBFH-M125	0.06	$\pm$0.01	$^{+0.01}_{-0.01}$	0.00
ggH-M125	0.07	$\pm$0.03	$^{+0.02}_{-0.02}$	0.00
ttH-M125	0.21	$\pm$0.01	$^{+0.02}_{-0.03}$	0.01
VH-M125	0.64	$\pm$0.02	$^{+0.06}_{-0.06}$	0.04
tt+X	-2.3	$\pm$1.4	$^{+0.7}_{-0.3}$	-0.12
WGamma	1.3	$\pm$0.6	$^{+0.4}_{-0.3}$	0.07
GJets	2.8	$\pm$2.2	$^{+1.8}_{-0.4}$	0.15
ZGamma	4.2	$\pm$0.9	$^{+0.8}_{-0.6}$	0.23
DiPhoton	12.3	$\pm$0.9	$^{+0.8}_{-0.9}$	0.67
Total MC bkg	18.3	$\pm$2.9	$^{+2.3}_{-1.3}$	1.00
Data	23.0	$\pm$4.8	$^{+0.00}_{-0.00}$	1.26

Table 16: Signal, resonant background, nonresonant background, and data yields in signal region 2 (SR 2).

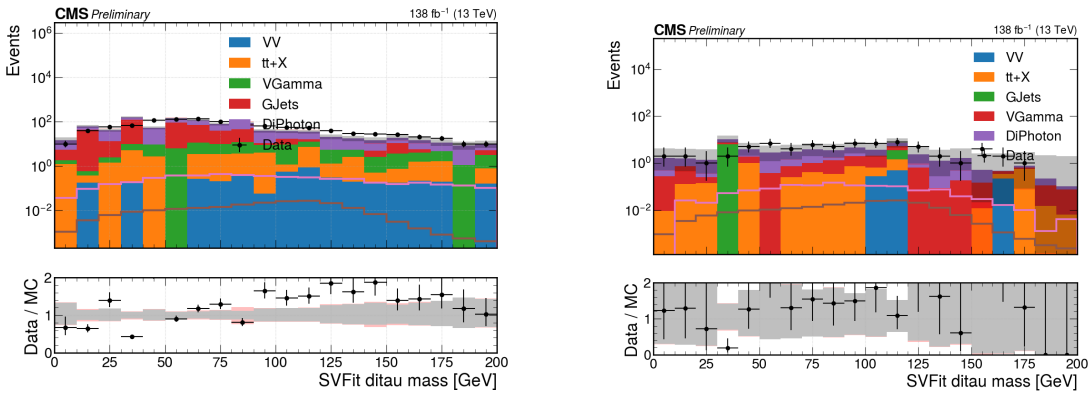


Figure 63: Comparison of the $m_{\tau\tau}^{SVFit}$ of the $\tau\tau$ candidates at the preselection level (left) and in the signal regions (right). In the SR plot, we notice an improved modeling and overall better agreement between Data and Monte Carlo. The only data-point in disagreement with simulation is due to the limited statistics and high weights of the γ + jets sample.

Part IV

Statistical Analysis

11 Signal & Background Modeling

The strategies used for modeling resonant (i.e. with a $H \to \gamma\gamma$ decay) and nonresonant processes follow those used in many CMS $H \to \gamma\gamma$ analyses, see for instance Ref. [71]. Resonant processes are modeled from simulation with a fit of the $m_{\gamma\gamma}$ distribution to a Double Crystal Ball summed to a Gaussian function. Nonresonant processes are modeled from data with a fit of the $m_{\gamma\gamma}$ distribution to a variety of functional forms. The choice of function used to model the nonresonant background is treated as a discrete nuisance parameter. This technique is commonly known as the discrete profiling method [72].

11.1 Signal and resonant background models

Both the HH signal and single H backgrounds are modeled from simulation with a fit of the $m_{\gamma\gamma}$ distribution to a Double Crystal Ball summed to a Gaussian function. In order to keep track of the year dependent systematic uncertainties and resolution effects, fits are performed separately for each production mode, signal region and year of data-taking. The signal and resonant background fits are shown in Fig. 64, Fig. 65 and Fig. 66 below.

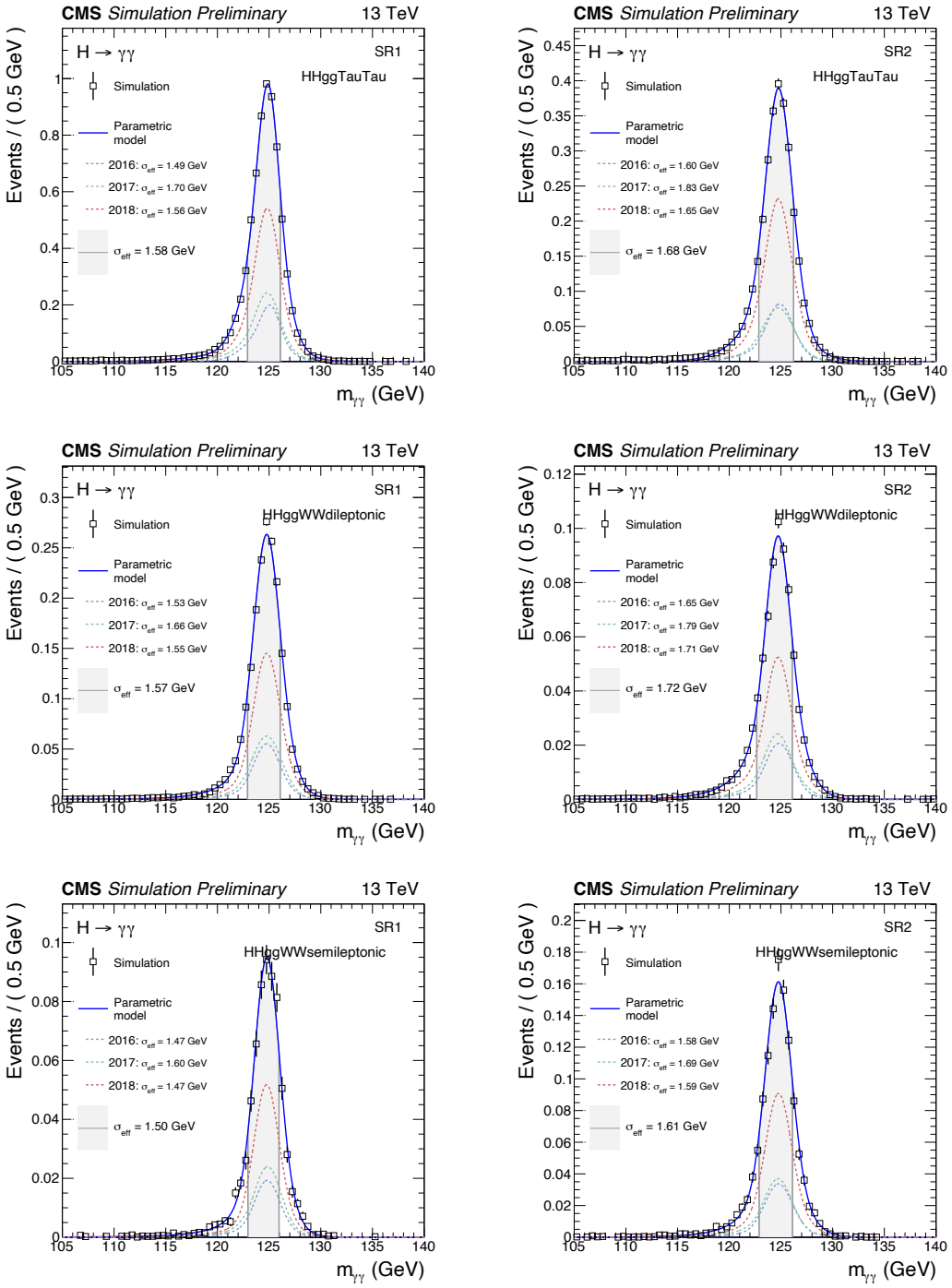


Figure 64: Parametric models of the signal process $HH \rightarrow \gamma\gamma\tau\tau$ and $HH \rightarrow \gamma\gamma WW$ in SR 1 (left) and SR 2 (right).

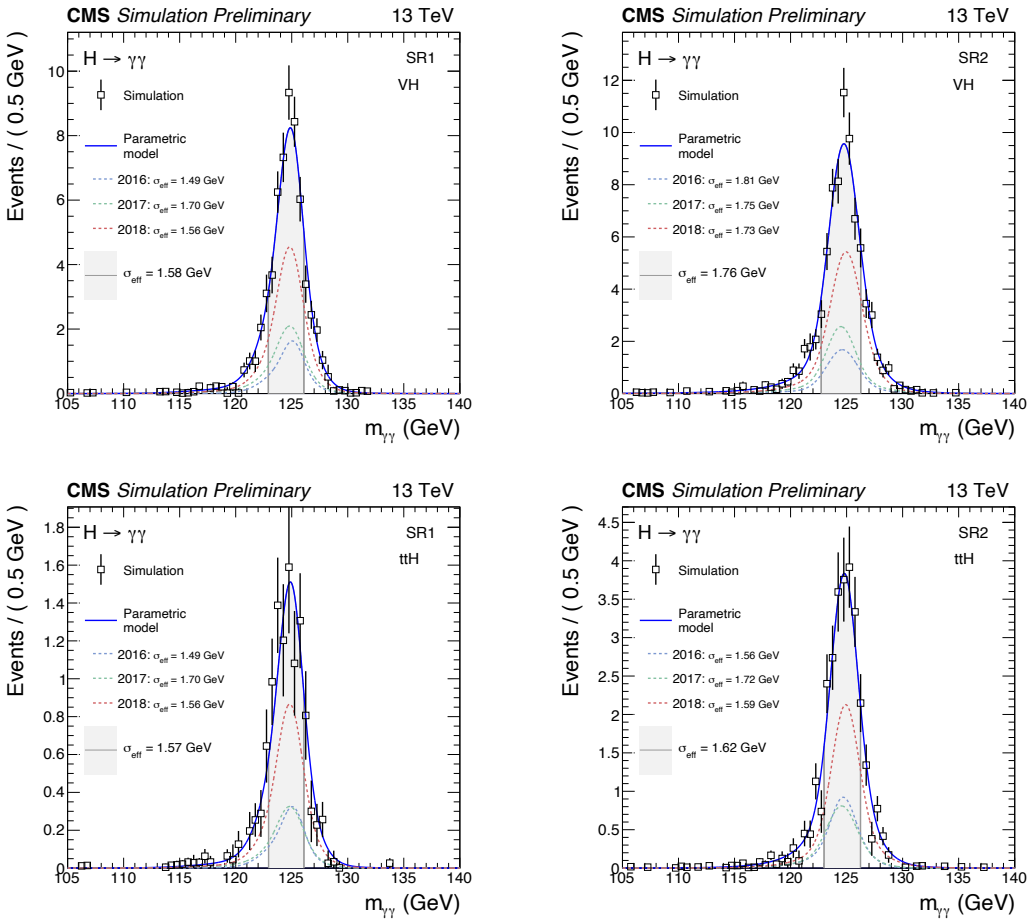


Figure 65: Parametric models of resonant background processes: VH and ttH in SR 1 (left) and in SR 2 (right).

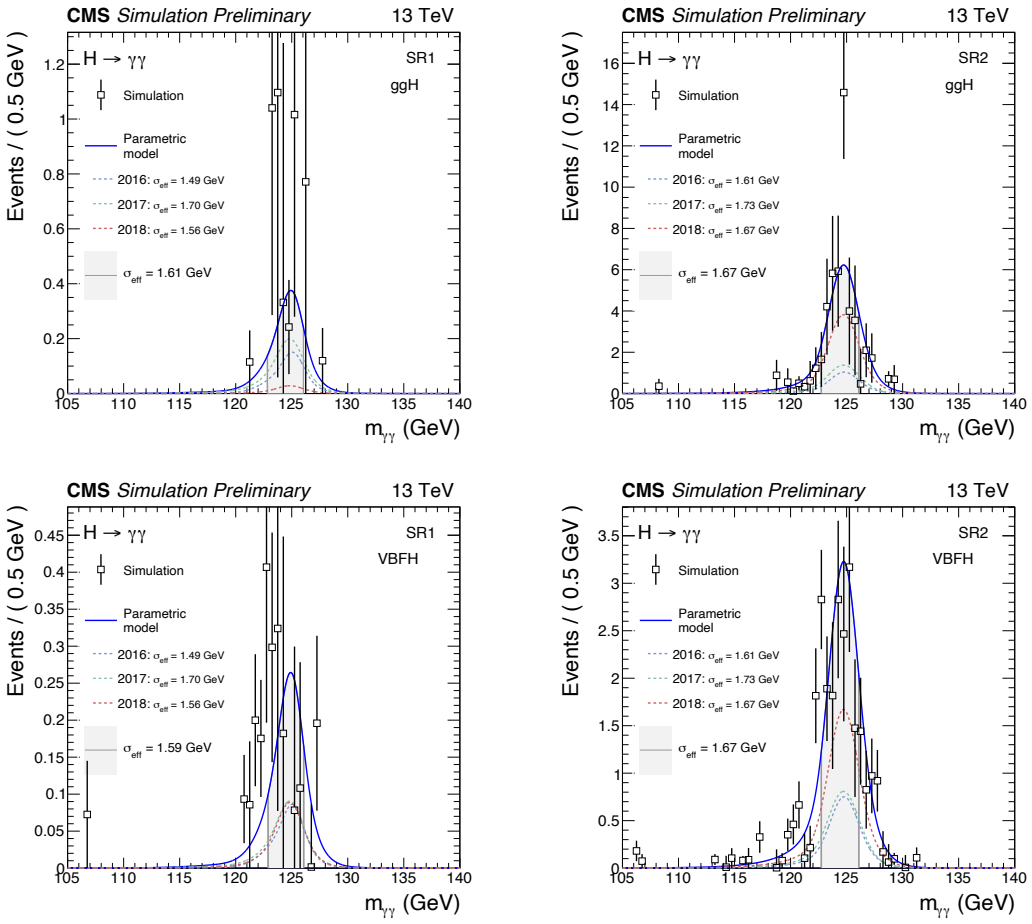


Figure 66: Parametric models of resonant background processes: ggH and VBFH in SR 1 (left) and in SR 2 (right).

In order to avoid fitting distributions with a limited number of points, we require each permutation of process, SR and year to have at least 1000 entries. If this is not the case, we instead "borrow" the DCB + Gaussian profile from the $HH \to \gamma\gamma\tau\tau$ signal model, which is known to not suffer from low statistics, and scale the shape according to the expected yields for the process, SR and year with limited statistics. This procedure allows to improve the agreement between the mean and standard deviation of the fits across years, although it has no impact on the final sensitivity. Lastly, this approach is possible due to the fact that the underlying physics behind the $H \to \gamma\gamma$ decay is the same, whether we are considering a resonant background, such as ggH, or the HH signals. All $m_{\gamma\gamma}$ fits are performed in the [120, 130] GeV mass interval and an independent fit is produced for each systematic variation that may affect the shape of the $m_{\gamma\gamma}$ distribution.

11.2 Nonresonant background models

All nonresonant backgrounds (i.e. any process without a $H \to \gamma\gamma$ decay) are modeled directly from data with fits of the $m_{\gamma\gamma}$ distribution in the [100, 180] GeV range. Nonresonant background processes have $m_{\gamma\gamma}$ distributions which are smoothly falling functions of $m_{\gamma\gamma}$ and it is ensured that any potential "turn-on" effects near the trigger threshold on $m_{\gamma\gamma}$ (around 90 GeV) are mitigated through the use of the scaling photon p_{T} cuts: $p_{\mathrm{T}}/m_{\gamma\gamma} > 0.33(0.25)$ for the leading (subleading) photon. However, the exact shape of the $m_{\gamma\gamma}$ spectrum is not known, and a particular choice of function may, in principle, bias the result. For this reason, a variety of functional forms are used and the choice of functional form is treated as a discrete nuisance parameter [72].

Four families of functions are considered for the nonresonant background fits:

1. Expontential

$$f_N(x) = \sum_{i=0}^{N} a_i \exp(-b_i x), \tag{66}$$

2. Power Law

$$f_N(x) = \sum_{i=0}^{N} a_i x^{-b_i}, \tag{67}$$

3. Bernstein polynomial

$$f_N(x) = \sum_{i=0}^{N} a_i \binom{N}{i} x^i (1-x)^{N-i}, \tag{68}$$

4. Laurent series

$$f_N(x) = \sum_{i=0}^{N} a_i x^{-4+\sum_{j=0}^{i}(-1)^i j}, \tag{69}$$

with a_i and b_i the parameters to be fitted and N the order of each function.

In general, a higher-order function will always give a better fit. To avoid having functions of arbitrarily higher-order being selected as the best-fit functions from each family, an F-test [73] is used to assess the goodness-of-fit while accounting for the number of degrees of freedom of the fit. In the final statistical analysis, the choice of function used to model the background is treated as a discrete nuisance parameter.

In contrast to the fits for signal and resonant backgrounds, the fits for nonresonant background are performed inclusively for all three years of data-taking. Although there may be year-dependent effects in the $m_{\gamma\gamma}$ distribution due to the evolving performance of the CMS electromagnetic calorimeter (ECAL), most of the sensitive signal regions have a modest number of data events entering the sidebands and aggregating the three years together improves the statistics for these fits.

The final set of functions and their respective orders for each signal region are shown in Fig. 67 and the background models with uncertainties are shown in Fig. 68.

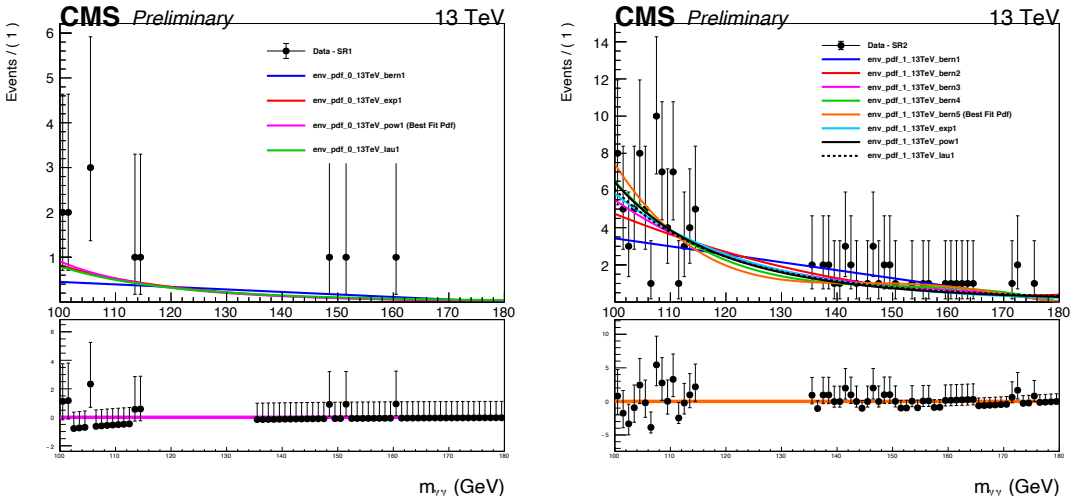


Figure 67: Set of functions considered for the background models in SR 1 (left) and SR 2 (right). The choice of functional form is considered as a discrete nuisance parameter in the final statistical analysis.

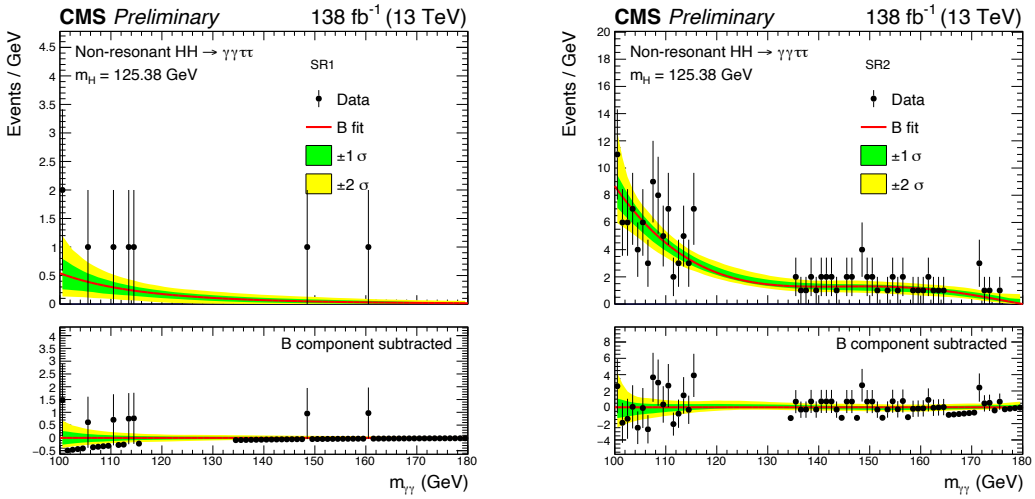


Figure 68: Nonresonant background model and uncertainty for SR 1 (left) and SR 2 (right).

11.3 Bias Study

To study any potential bias in the choice of the background model, we generate 1000 toy MC experiments for each background function considered in the envelope method. The number of background events in each toy is sampled from a Poisson distribution with mean equal to the number of background events in the signal region considered and an injected signal strength of one. We then fit the MC toys to extract the signal strength, plot its pull distribution and perform a Gaussian fit of it. Generally, we expect to see an approximately Gaussian shape with zero mean and unit width. As long as the absolute value of the mean is < 0.14, which implies that the coverage error and the uncertainty error are $< 1\%$, this is considered satisfactory. The pulls for this study are shown in Fig. 69 for SR1 and in Fig. 70 for SR2. No bias is observed in the choice of background models, with the exception of the Bernstein polynomial of degree 5 (bern5) in SR2. However, because of the large number of degrees of freedom, such function can easily pick up fake spikes and bumps in data (see right plot of Fig. 67) and yield the lowest χ^2. Therefore, we do not expect bern5 to accurately model our background and is in fact later removed in the extraction of final limits.

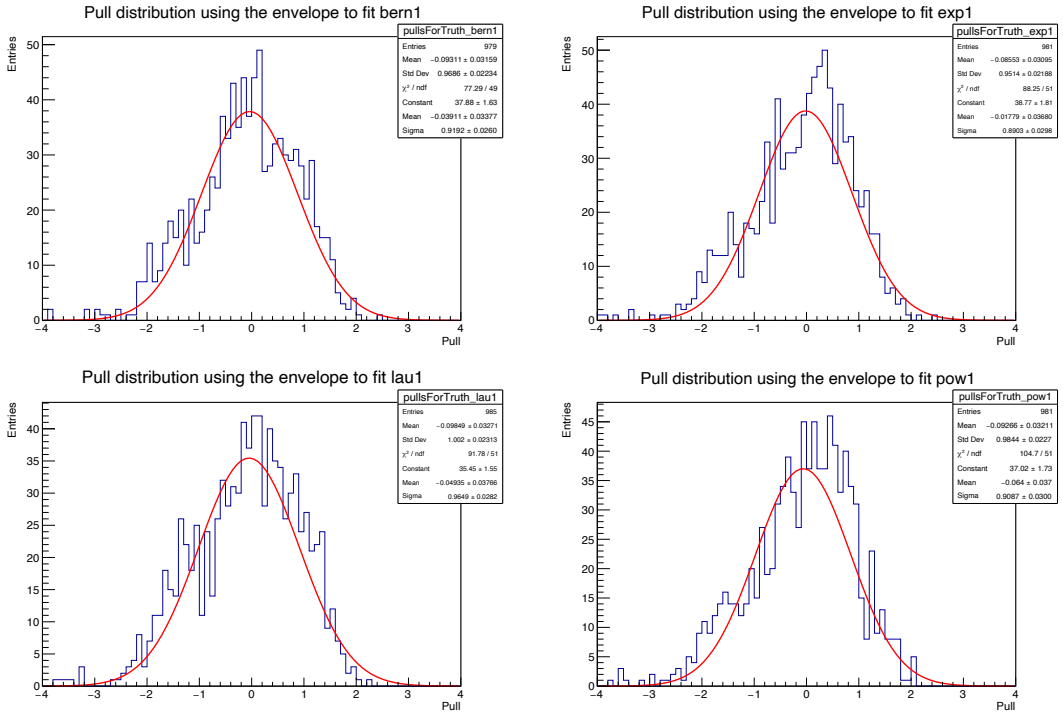


Figure 69: Pull distributions obtained by using the envelope method to fit the background functions in SR1. The numerical results of the Gaussian fit are displayed at the bottom of the statistics box.

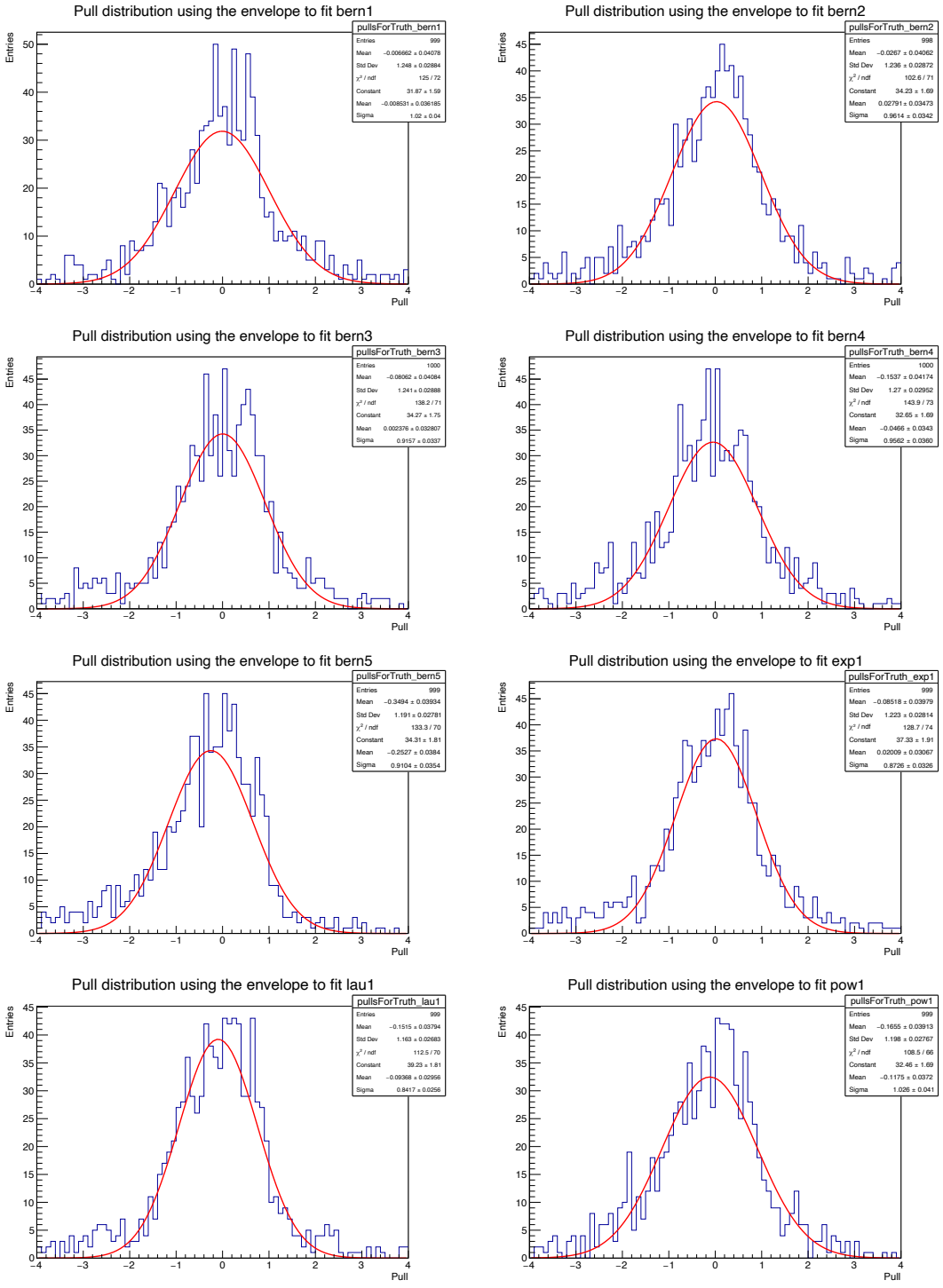


Figure 70: Pull distributions obtained by using the envelope method to fit the background functions in SR2. The numerical results of the Gaussian fit are displayed at the bottom of the statistics box.

12 Statistical Interpretation

After the analysis has been tuned and optimised for optimal sensitivity reach, we need to adopt a statistical modeling of the data in order to establish the presence of HH signal in the observed data or lack thereof. In doing so, we adopt a modified frequentist approach and utilise the same statistical framework that has been developed by the ATLAS and CMS experiments in their searches for the Higgs boson. First of all, the value that we want to constrain is the signal strength μ defined as

$$\mu = (\sigma_{HH} \cdot Br)_{obs} \,/\, (\sigma_{HH} \cdot Br)_{SM}$$

where $(\sigma_{HH} \cdot Br)_{obs}$ is the observed di-Higgs cross section times the $\gamma\gamma\tau\tau$ branching ratio, while $(\sigma_{HH} \cdot Br)_{SM}$ is the SM predicted one. In order to asses the presence, or absence, of signal in the observed data, we need to define a hypothesis testing procedure. The procedure that I am going to describe is pretty common to many high energy physics searches and it involves two hypothesis: a null hypothesis H_0, or background only hypothesis, which assumes no presence of signal and that all of the observed events come from background processes; and the alternative hypothesis H_1, also signal + background hypothesis, which assumes the existence of signal with a signal strength of μ on top of the background. We then use the LHC profile likelihood test statistics in order to quantify the compatibility of a particular hypothesis given the observed data. When we set a limit on the signal strength μ, that is equivalent to finding the value of μ for which the background only hypothesis H_0 can be excluded in favor of the signal + background hypothesis H_1.
Let us now define s and b to be the signal and background yields, respectively. The signal and background distributions, however, depend on the systematic uncer-

tainties (θs) that will be later discussed in Sec.13; for this reason we will refer to the signal and background yields distributions as $s(\theta)$ and $b(\theta)$. Given an observation of n data events, we can define the likelihood function $\mathcal{L}(n \mid \mu, \theta)$ as:

$$\mathcal{L}(n \mid \mu, \theta) = P\left[n \mid \mu \cdot s(\theta) + b(\theta)\right] \cdot p(\theta \mid \tilde{\theta}) \tag{70}$$

with P being the probability density function, $p(\theta \mid \tilde{\theta})$ the prior knowledge on the nuisance parameters θ given their best estimate $\tilde{\theta}$. Traditionally, the frequentist formulation does not include a term account for systematic uncertainties, which is an addition of the modified frequentist approach. To evaluate the likelihood we can either use the observed data, or a number of pseudo experiments in which we use the expected distributions for $\mu \cdot s$, b and θ to generate the events to be fitted. For the case of a binned distribution with total number of bins *nBins*, the probability density can be defined as the product of several Poisson probabilities as:

$$\mathcal{L}(n \mid \mu, \theta) = \prod_{i=1}^{nBins} \frac{(\mu \cdot s_i + b_i)^{n_i}}{n_i!} \exp\left[-(\mu \cdot s_i + b_i)\right] \cdot p(\theta \mid \tilde{\theta}) \tag{71}$$

where we run the multiplication over all of the bins i and s_i, b_i and n_i represent the per-bin signal, background and total yields, respectively.

Then we need to quantify the level of compatibility of the observed data with the background only hypothesis or the signal + background hypothesis, and to do so we make use of the LHC profile likelihood ratio [74, 75] as a test statistic $\tilde{q}_\mu$ as:

$$\tilde{q}_\mu = -2\ln\frac{\mathcal{L}(n,\tilde{\theta}\,|\,\mu,\hat{\theta}_\mu)}{\mathcal{L}(n,\tilde{\theta}\,|\,\hat{\mu},\hat{\theta})} = -2\ln\lambda(\mu)\,,\ \text{ with }\ 0 \leq \hat{\mu} \leq \mu \tag{72}$$

where we made use of $\hat{\mu}$ and $\hat{\theta}$ as best fit values for the signal strength and nuisance parameters, and $\hat{\theta}_\mu$ as best fit value for the nuisance parameters with a fixed μ, also known as the "conditional maximum likelihood estimator". The $\lambda(\mu)$ ratio is defined in the interval $0 \leq \lambda \leq 1$ where values of $\lambda(\mu)$ closer to 1 imply good compatibility between the signal + background hypothesis with signal strength μ and the observed data (or pseudo data). Now, the CMS convention is to require $\hat{\mu}$ to be positive, which I am not completely sure as to why. The value of $\hat{\mu}$ is just a post fit value from a statistical distribution, which doesn't have an intrinsic relation to nature; and especially if we consider regions where the background is dominating the signal yields, you may pretty easily convince yourself that statistical fluctuations in the events to be fitted can well give rise to negative signal strengths. But, as Prof. Campagnari would say, this is what the "statistics guru" at CMS recommend, so we will go along with it. The other constraint $\hat{\mu} \leq \mu$ is imposed to alleviate the potential statistical fluctuations of background in the downward direction, so to discourage evidence against an alternative hypothesis with small signal strength. This protective constraint, in fact, one of the main features that characterise the modified frequentist method that will be soon discussed.

12.1 Exclusion Limits

The modified frequentist approach $\mathrm{CL_s}$ allows us to compute exclusion limits by calculating the observed value of the $\tilde{q}_\mu^{obs}$ test statistics for each of the μ values that we are considering. We then define two p-values that correspond to the observation

of the background only hypothesis (p_b) and the signal + background hypothesis (p_μ) as:

$$p_\mu = \mathrm{CL}_{s+b} = P\left(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs} \mid H_1\right) = \int_{\tilde{q}_\mu^{obs}}^{\infty} f\left(\tilde{q}_\mu \mid \mu, \tilde{\theta}_\mu^{obs}\right) \mathrm{d}\tilde{q}_\mu \tag{73}$$

and

$$1 - p_b = \mathrm{CL}_b = P\left(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs} \mid H_0\right) = \int_{\tilde{q}_\mu^{obs}}^{\infty} f\left(\tilde{q}_\mu \mid 0, \tilde{\theta}_0^{obs}\right) \mathrm{d}\tilde{q}_\mu \tag{74}$$

where $f\left(\tilde{q}_\mu \mid \mu, \tilde{\theta}_\mu^{obs}\right)$ and $f\left(\tilde{q}_\mu \mid 0, \tilde{\theta}_0^{obs}\right)$ are probability density functions obtained via the production of pseudo data, in other terms events generated according to the Poisson probability distributions for the $s + b$ and for the b only hypotheses. In addition, we fix the nuisance parameters $\tilde{\theta}_\mu^{obs}$ and $\tilde{\theta}_0^{obs}$ to their optimum value obtained by performing a fit to the observed data.

We then defined the modified frequentist approach Confidence Level intervals ($\mathrm{CL_s}$) to be:

$$\mathrm{CL}_s(\mu) = \frac{\mathrm{CL}_{s+b}}{\mathrm{CL}_b} \tag{75}$$

and we treat it as the final figure of merit of the hypothesis testing procedure. A signal is considered to be excluded at a CL α if $\mathrm{CL_s}(\mu) < 1 - \alpha$; so that in order to exclude a signal with a given signal strength μ at the 95% CL we must impose $\mathrm{CL_s}(\mu) < 0.05$. In the computation of the exclusion limits, we need to estimate the

probability density functions of $\tilde{q}_\mu$, $f\left(\tilde{q}_\mu \,|\mu,\, \tilde{\theta}_\mu^{obs}\right)$ and $f\left(\tilde{q}_\mu \,|0,\, \tilde{\theta}_0^{obs}\right)$ from pseudo data. This step can be particularly demanding in terms of the computing resources needed. To obviate to such inconvenient, it is common practice to adopt the asymptotic limit approximation [76] of the probability density functions instead. The asymptotic limit approximation can be obtained by making use of the Wald theorem, which states that, in the limit of a sufficiently large data sample with N events, the test statistics $\tilde{q}_\mu$ can be approximated as:

$$\tilde{q}_\mu = -2 \ln \lambda(\mu) \rightarrow \frac{(\mu - \hat{\mu})^2}{\sigma^2} + \mathcal{O}(1/\sqrt{N}) \tag{76}$$

where σ^2 is defined as the standard deviation of $\hat{\mu}$ and derived from the covariance matrix of the likelihood function.

In this way, we obtain the upper limit on μ as:

$$\mu = \hat{\mu} + \sigma \, \Phi^{-1} \, (1 - \alpha) \tag{77}$$

with Φ being the cumulative distribution of the standard Gaussian. In the asymptotic limit approximation, the standard deviation σ can be evaluated using the representative *Asimov dataset*, which is distributed according to the likelihood function profile where the observed parameters are fixed to their expected values.

12.2 Signal Excess Quantification

In the case where an excess of events is observed, it comes handy to have a way to quantify the presence of signal at a given CL. To do this, we must compare the compatibility of the observed data with the background only using the test statistics:

$$\tilde{q}_0 = -2 \ln \frac{\mathcal{L}(n, \tilde{\theta} \,|\, 0, \hat{\theta}_0)}{\mathcal{L}(n, \tilde{\theta} \,|\, \hat{\mu}, \hat{\theta})} \tag{78}$$

where we are testing the $\mu = 0$ hypothesis. Given the above definition, when $\tilde{q}_0$ tends to larger values it implies a consistent incompatibility between the observed data and the background only hypothesis. We can quantify such incompatibility by introducing a p-value, defined as

$$p = P\left(\tilde{q}_0 \geq \tilde{q}_0^{obs} \,|\, H_0\right) ,$$

in other words, the probability of $\tilde{q}_0$ being larger (or equal) than $\tilde{q}_0^{obs}$ under the background only hypothesis (no signal). For easier interpretation, the p-value is then often expressed in terms of Z significance, where Z corresponds to the number of standard deviations σ of a one sided Gaussian integral as:

$$p = \int_Z^{+\infty} \frac{1}{2\pi} e^{-x^2/2} \mathrm{d}x \tag{79}$$

The high energy physics convention is to claim *evidence* of a new particle when the significance is 3σ or $Z = 3$ or $p = 1.3 \times 10^{-3}$, and *observation* of a new particle when the significance is 5σ or $Z = 5$ or $p = 2.8 \times 10^{-7}$.

13 Systematic Uncertainties

This section describes the various sources of systematic uncertainty which affect the measurement of $HH \to \gamma\gamma\tau\tau$ final states, along with their treatment.

The majority of systematic uncertainties are modeled with log-normal probability distribution functions that parametrize the event yield in each signal region. In other words, these affect the overall normalization of the signal and resonant background processes in each signal region (and in principle can vary by signal region) but do not affect the shape of the $m_{\gamma\gamma}$ distribution.

The second type of systematic uncertainty are those which affect the shape (and in principle the normalization) of the $m_{\gamma\gamma}$ distributions for signal and resonant background processes, as well as the non-resonant background modeled from data. These systematic uncertainties are treated as nuisance parameters in the signal models for each signal region. In particular, the uncertainty associated with the choice of function for modeling the non-resonant background is treated as a discrete nuisance parameter, following the prescription of the discrete profiling method [72].

13.1 Theoretical Uncertainties

The theoretical uncertainties can affect the normalization of both the signal HH processes as well as the background single H processes. The uncertainties considered for the signal HH processes are:

- *Uncertainty in HH cross section*: estimated by varying the factorization (μ_F) and renormalization (μ_R) scales, the parton distribution functions (PDFs), and the value of the strong force coupling constant (α_s), and are

found to be $^{+4.5\%}_{-6.4\%}$ with the calculations in Refs. [77–83].

The uncertainties for the resonant background (single H) processes are:

- *QCD renormalization (μ_R) and factorization (μ_F) scales*: Values are taken following the recommendations of Ref. [84] and are typically on the order of 5-10%.
- *Strong coupling constant (α_s)*: Values are taken following the recommendations of Ref. [85] and are up to 2% in magnitude.
- *Parton density function (PDF)*: Values are computed following the PDF4LHC prescription [85, 86] and are typically on the order of 1%.

Finally, the uncertainties considered for both the signal and resonant background processes are:

- $H \to \gamma\gamma$ *branching fraction*: Value taken from Ref. [84] and estimated to be around 2%.

13.2 Experimental uncertainties

Here we describe the sources of experimental uncertainty, which originate from imperfections or temporary deficiencies in the detection apparatus. Experimental uncertainties may either affect the shape and normalization of the $m_{\gamma\gamma}$ distribution or simply the normalization.

Those affecting the shape of the $m_{\gamma\gamma}$ distribution include:

- *Photon energy scale*: the uncertainty associated with the corrections derived to match the photon energy scale in simulation with that observed in data. The systematics variations are provided by the EGamma POG

and summarized in this document.

- *Photon energy resolution*: the uncertainty associated with the corrections derived to match the photon energy resolution in simulation with that observed in data. The sytematics are added to the v9 nanoAOD produced centrally. The branches used are the dESigma up and down variations - CMSSW reference.
- *Material uncertainty*: in general, electrons shower earlier than photons when passing through the tracker material. Dedicated samples with variations in the amount of material ahead of the ECAL are used to estimate the effect on the photon energy scale.
- *FNUF uncertainty*: differences in light collection efficiency (LCE) along the length of ECAL crystals result in different responses to electrons and photons. The uncertainty is estimated using the LCE model described in [87], derived from optical simulation [88].

Note that the scales and smearings applied by default in nanoAOD have been validated in the EGM and XPOG groups, as shown by the plots in Fig. 71 taken from the following presentations:

- XPOG 23 Feb 2022
- CMS Week 5 Apr 2022

Experimental uncertainties affecting only the normalization of a given process include:

- *Integrated luminosity*: An estimated uncertainty of 2.5% for 2018, 2.3% for 2017 and 2.5% for 2016 is used as recommended by the LUMI POG

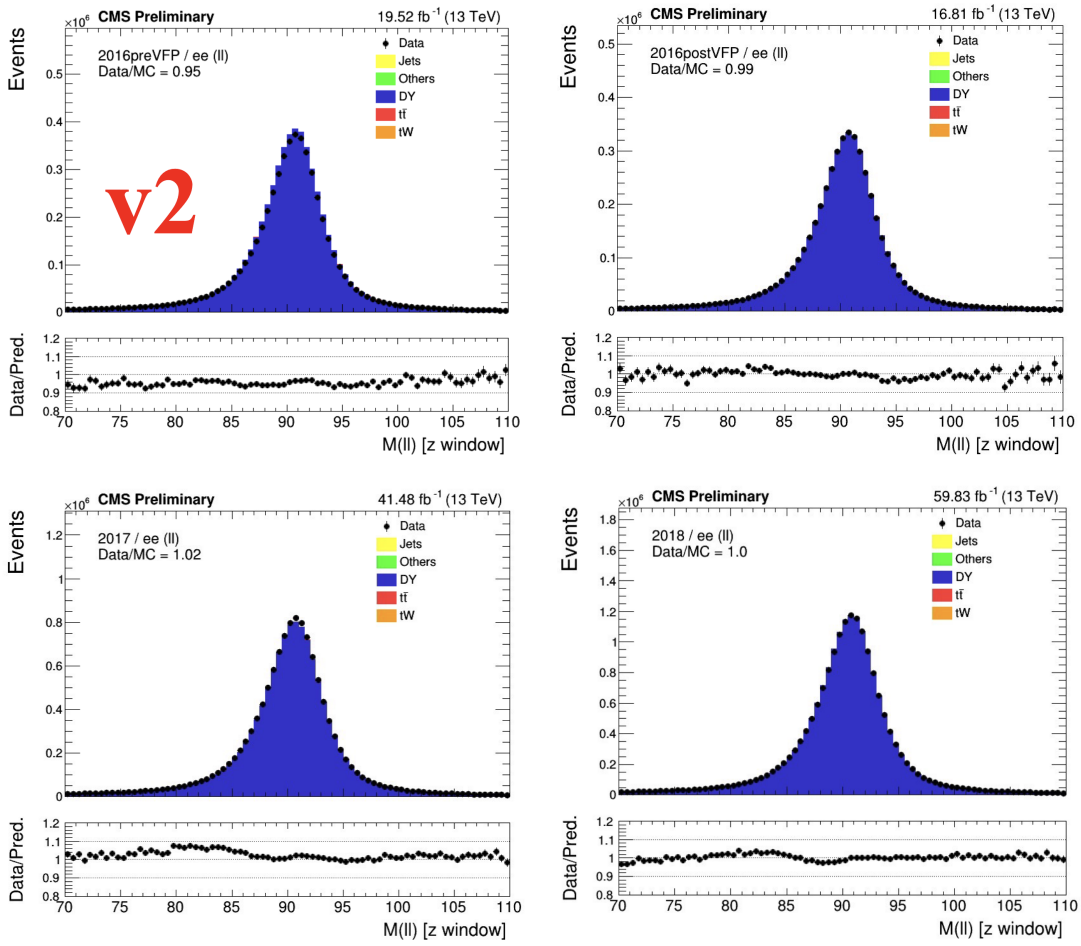


Figure 71: Comparison of data and simulation in $Z \rightarrow ee$ events: 2016 preVFP (upper left), 2016 postVFP (upper right), 2017 (lower left), and 2018 (lower right).

- *L1 prefiring*: the inefficiency due to the prefiring of the L1 trigger during 2016 and 2017 data-taking periods is estimated following the recommended treatment from the JetMET group [1]. The magnitude of the inefficiency is calculated as the product over all jets and photons in the event and the systematic uncertainty in the inefficiency is taken as the maximum of the per-object uncertainties (summed in quadrature) and 20%. The prefiring central weight and up/down variations are computed us-

[1] https://twiki.cern.ch/twiki/bin/viewauth/CMS/L1ECALPrefiringWeightRecipe

ing the recipe in nanoAOD tools.

- *PU weight*: the PU central reweighting and up/down variations are computed using the recipe in nanoAOD tools and are of the order of $\sim$ 5%.
- *Trigger scale factors*: the efficiency of the HLT triggers used in this analysis are calculated using the tag-and-probe method in events.
- *Diphoton preselection scale factors*: the uncertainty in the diphoton preselection efficiency is estimated from the ratio of the efficiency as calculated in data versus that calculated in simulation.
- *Electron Veto scale factors*: similar to the pixel seed veto uncertainty, the uncertainty in the scale factors applied to simulation to match the efficiency of the electron veto, derived using $Z \rightarrow \mu\mu + \gamma$ events.
- *Shape of the DeepJet b-tag discriminant*: the b-tagging discriminant is corrected in simulation by a continuous reshaping factor which is a function of p_T, η, and jet flavor. The uncertainty in the reshape factor is estimated as described in [89]. Nine individual sources of uncertainty are considered, including uncertainties relating to the jet energy scale, the tagging of light-flavor, charm-flavor, and heavy-flavor jets, and the limited statistics available in deriving the scale factors.
- *Jet and MET up/down variations*: the JES global up and down variations and the JER variations are used to compute the Jet momentum systematic uncertainties. The MET is affected by the same systematics. Additionally, an uncertainty source sue to the MET unclustered component is considered.
- *Muon pT corrections*: the latest available muon rochester corrections (v5,

see RochCorMuon TWiki) are used for the muon scale. The relative up and down uncertainties are applied.

- *Tau pT corrections*: the tau energy scale corrections are applied following the recipe in nanoAOD tools post-processing. The relative systematic uncertainties are also computed and used as recommended in TAU TWiki.
- *Muon ID and ISO scale factors*: the muon ID and isolation scale factors for the specific working points used in the selection are taken from MUO POG Gitlab. The computation is implemented in nanoAOD dedicated tools. Loose ID and ISO are used.
- *Tau ID scale factors*: the Deep tau ID scale factors and uncertainties are applied following the recipe in nanoAOD tools post-processing. The scale factors and uncertainties are used as recommended in used as recommended in TAU TWiki.

13.2.1 Treatment of known detector issues

Several issues with the CMS detector were identified during Run 2 data-taking, including:

1. *Prefiring of the L1 trigger*: described under experimental uncertainties
2. *Effect of EE noise on* E_T^{miss}: for 2017, in which the transparency loss of ECAL crystals in the endcap is very high and poorly modeled in simulation. We use the recommended recipe from the JetMET group, which excludes unclustered pf candidates and jets in the high $|\eta|$ region from the calculation of $E_{\mathrm{T}}^{\mathrm{miss}}$.
3. *Effect of hadronic endcap calorimeter (HEM) 15/16 failure on* E_T^{miss}: other $H \to \gamma\gamma$

analyses with similar final states including leptons and , e.g. VH ($H \to \gamma\gamma$) have found that the effect of the HEM issue is negligible, slides. Therefore, no dedicated treatment is used for this analysis.

13.3 Impacts

In order to estimate which nuisance parameters (NPs) have the largest effect on the uncertainty of the final limit, we plot the NP impacts $\Delta\hat{r}$ for the $m_{\gamma\gamma}$ fit in all categories. The impact $\Delta\hat{r}$ of a NP θ on a parameter of interest (POI) μ, in our case this being the signal strength, is defined to be the shift $\Delta\mu$ caused when we vary θ from its fixed nominal value to its $+1\sigma$ or -1σ post-fit value, while maintaining all other NPs profiled as normal. In addition, the central column in the plot shows the value of $(\theta - \theta_0)/\Delta_\theta$ where θ and θ_0 represent the post-fit and pre-fit values of the nuisance parameter respectively, and Δ_θ the pre-fit uncertainty. This method effectively determines the correlation between each individual NP and the POI. Nuisance parameter impacts and pulls are shown in Figs. 72- 74.

By looking at the impacts of the nuisance parameters, we notice that the main source of uncertainty is the background normalisation in the signal regions. This is expected since our analysis is statistically limited. The photon corrections and resolutions closely follow, as these systematic variations can impact the mean of the $m_{\gamma\gamma}$ distribution, hence affecting the underlaying background estimate and the final limit. Additional sources of experimental uncertainties come from the theoretical uncertainties of SM Higgs production, τ ID efficiencies and MET. However, since our analysis is statistically limited, the effect of systematic uncertainties is minor (a few %s) compared to the impact of the statistical ones.

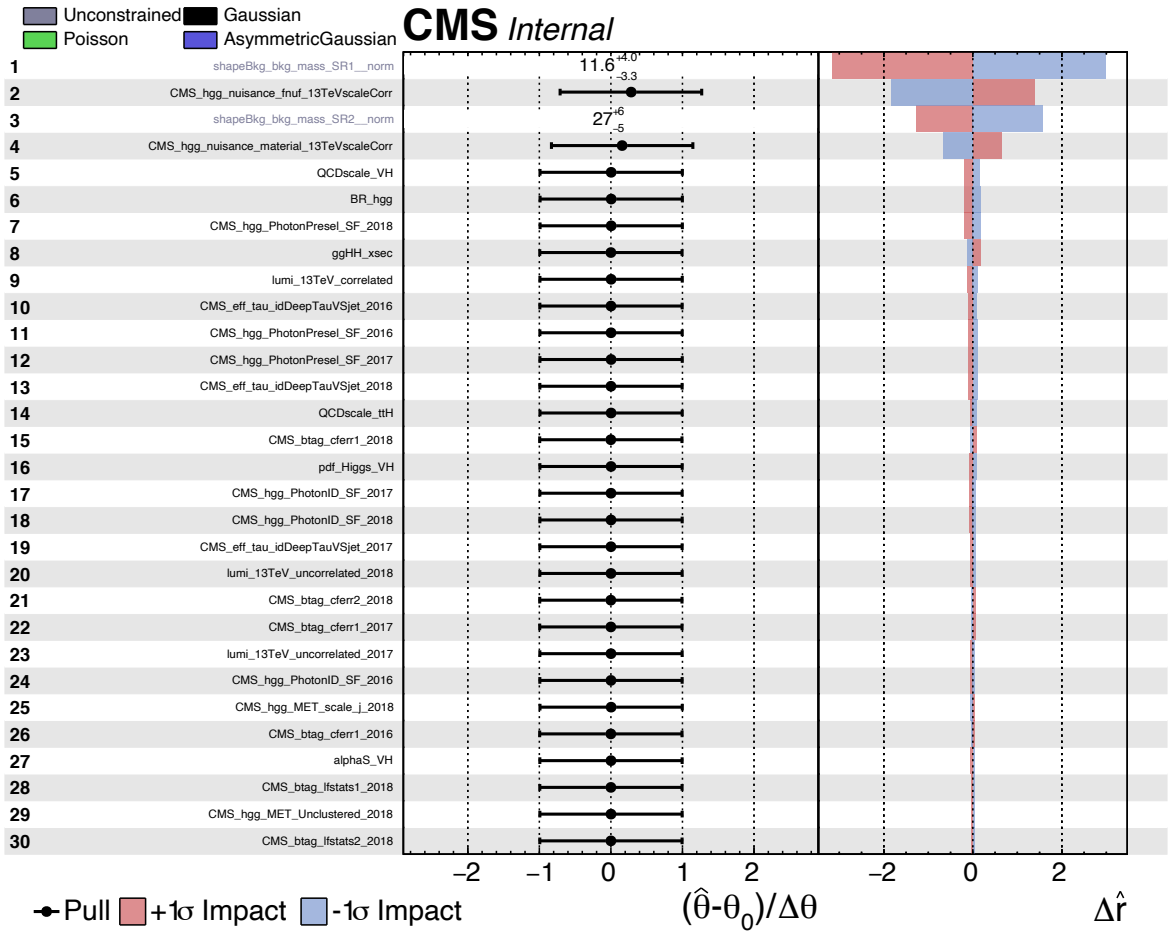


Figure 72: Pulls and impacts of the systematic uncertainties (1 of 3).

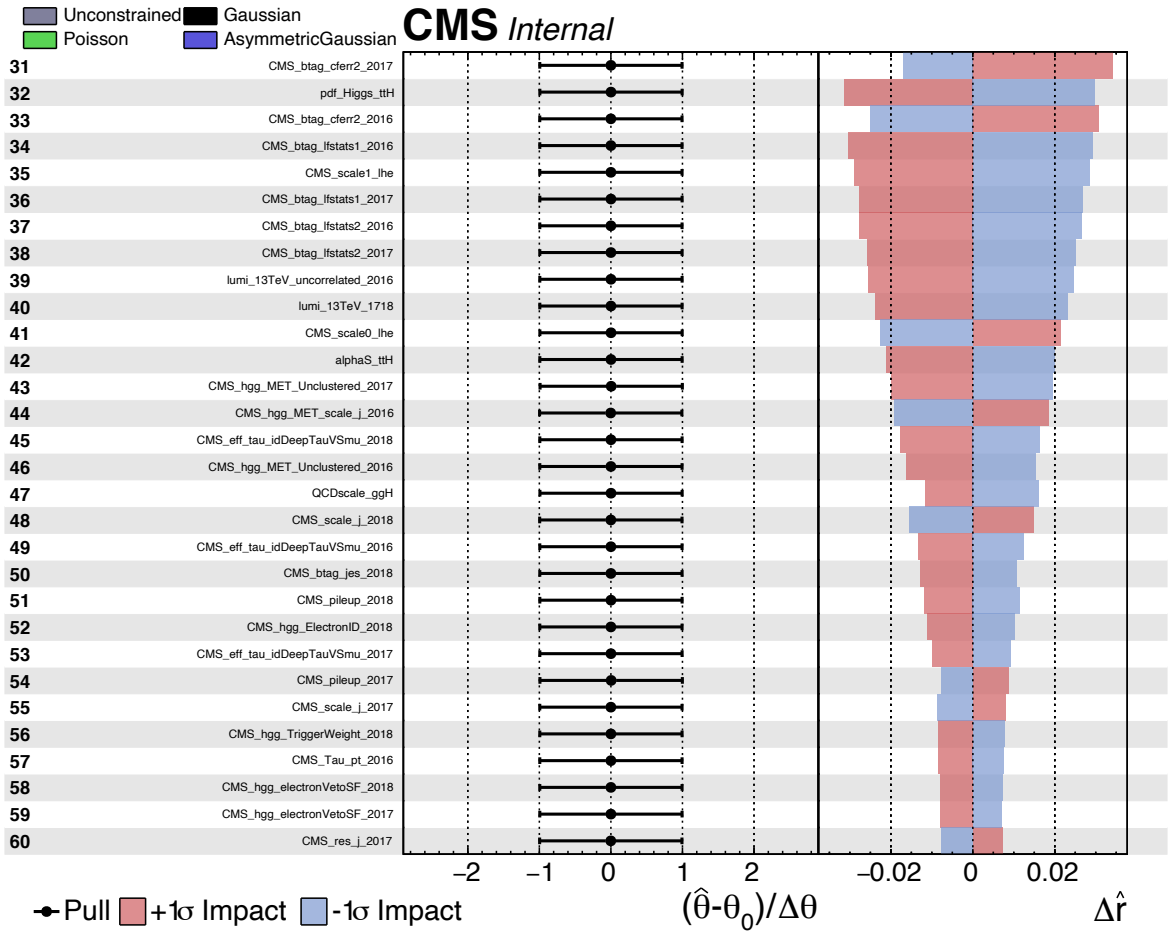


Figure 73: Pulls and impacts of the systematic uncertainties (2 of 3).

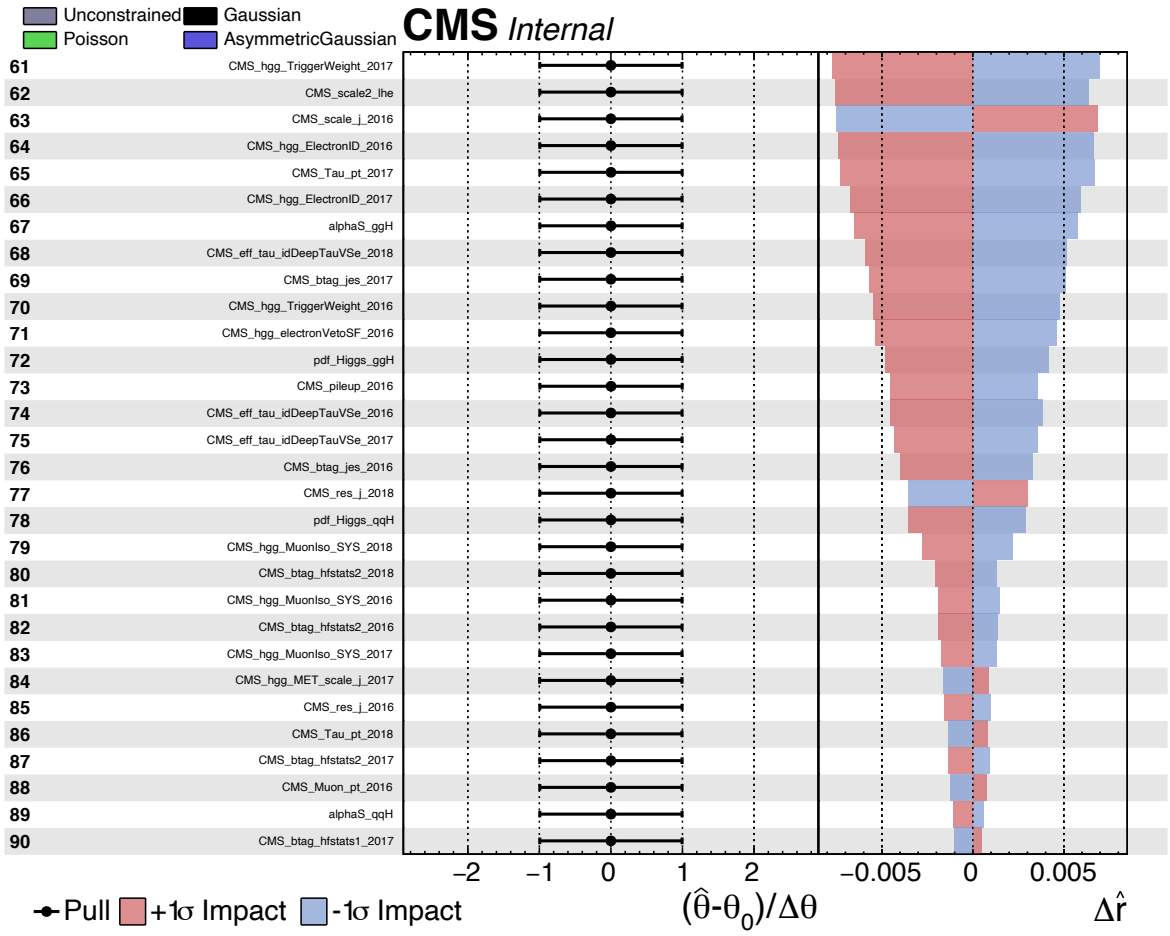


Figure 74: Pulls and impacts of the systematic uncertainties (3 of 3).

14 Results

The derivation of upper limits uses the modified frequentist approach for confidence levels (CL_s technique), with the LHC profile likelihood ratio as a test statistic [74, 75] in the asymptotic approximation [76], as detailed in Sec. 12. When extracting limits, we combine all of our simulated signals from the $HH \to \gamma\gamma\tau\tau$ channel along with $HH \to \gamma\gamma WW$, where the WW pair can decay either semileptonically or dileptonically; resonant backgrounds are also modeled from MC while the non-resonant background component is modeled from Data.

14.1 SM Results

Binned fits of the $m_{\gamma\gamma}$ distributions are performed simultaneously in all signal regions to extract central values and uncertainties on the parameters of interest, such as the signal strenght μ and the HH ggF cross section σ_{HH}. The bin width is set to 1 GeV. The observed (expected) 95% CL upper limit on the HH signal strength is found to be 24.7 (23.0) times the SM prediction, while the observed (expected) limit on the cross section is 764 (711) fb. Fig. 75 and Fig. 76 show the Signal + Background models for SR1 and SR2, respectively, where the green (yellow) band show the $\pm 1\sigma$ ($\pm 2\sigma$) uncertainty on the Background only model.

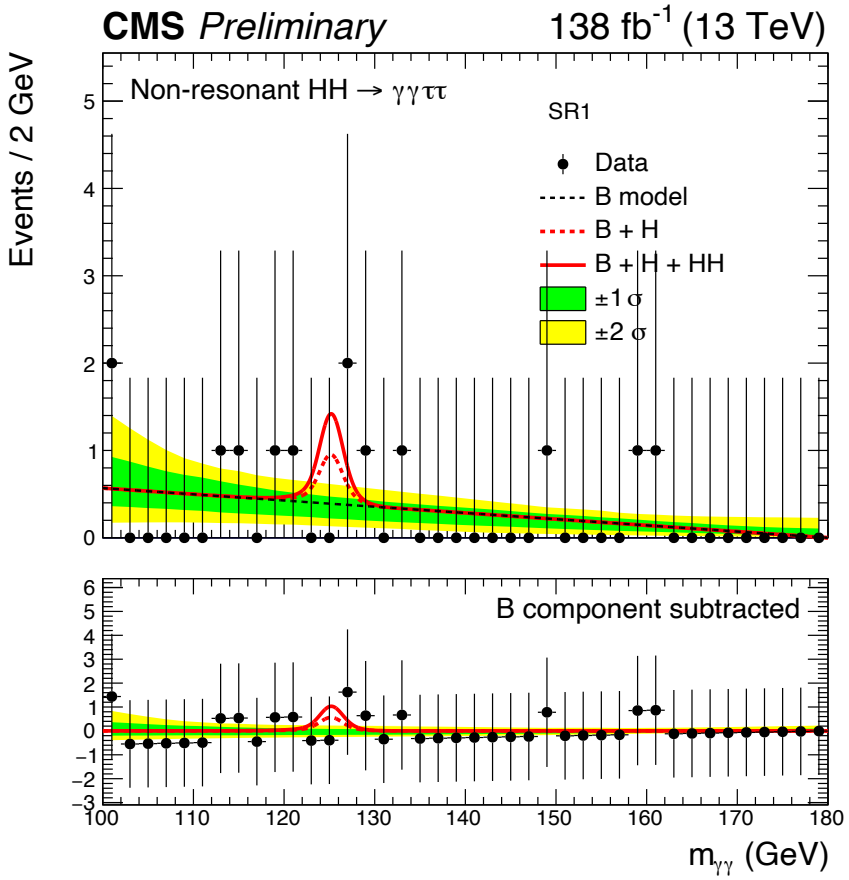


Figure 75: Unblinded $m_{\gamma\gamma}$ distribution in SR1. The plot shows the non-resonant background model (dashed black line), the addition of the SM Higgs background (dashed red line) and the addition of the di-Higgs signal model (solid red line). The green (yellow) bands display the one (two) standard deviation uncertainties for the non-resonant background.

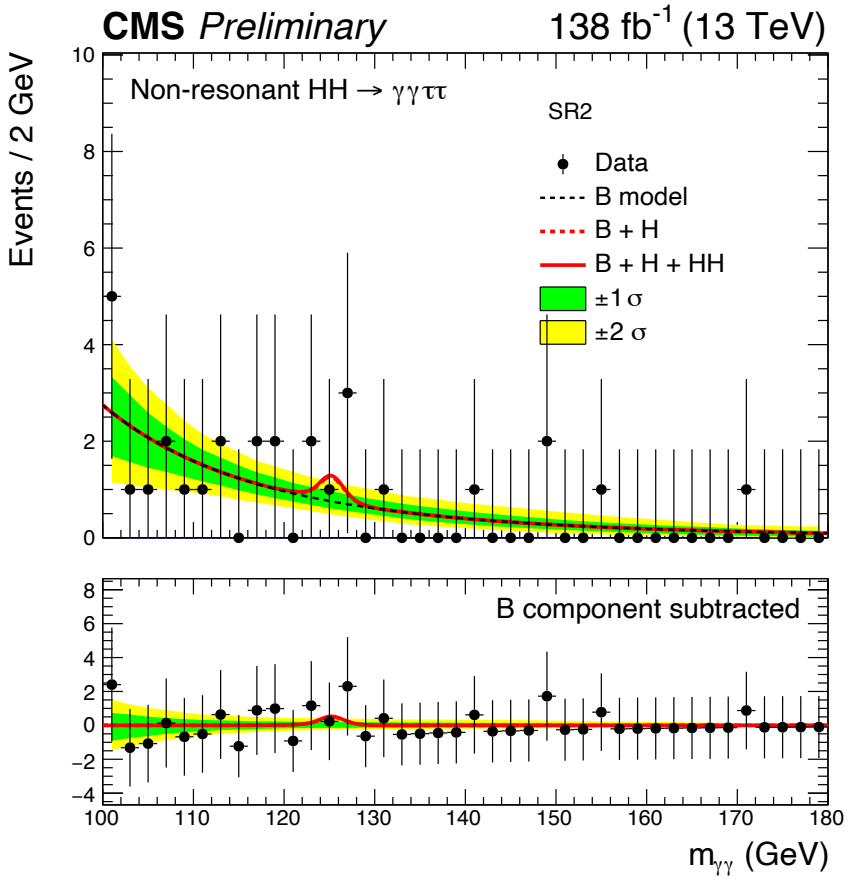


Figure 76: Unblinded $m_{\gamma\gamma}$ distribution in SR2. The plot shows the non-resonant background model (dashed black line), the addition of the SM Higgs background (dashed red line) and the addition of the di-Higgs signal model (solid red line). The green (yellow) bands display the one (two) standard deviation uncertainties for the non-resonant background.

14.2 κ_λ Scan

We also compute the σ_{HH} exclusion limits as a function of κ_λ, under the assumption that the top quark Yukawa coupling retains it SM value, $\kappa_t = 1$. The results are shown in Fig 77 for the expected 95% CL upper limits, where we normalise the input signals to $\sigma = 1 \times \mathcal{B}$ pb ($\mathcal{B}$ is the branching ratio of the decay channel of interest) while in the κ_λ scan the cross section varies as a function of the κ_λ/κ_t ratio. The green and yellow bands in the plot show the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands, respectively; while the theoretical prediction for the SM production ($\kappa_\lambda = \kappa_t = 1$) is depicted as a red parabola with its $\pm 1\sigma$ uncertainty bands. The expected limit is instead represented by a dashed line. The final limits in the κ_λ scan depend on various factors, including theoretical and experimental contributions. First of all, the predicted HH cross section has a minimum around $\kappa_\lambda \sim 2.45$, which is where the triangle and box diagrams maximise their destructive interference in the gluon gluon fusion production; hence the sensitivity dramatically varies around this point. For values of $\kappa_\lambda > 2.45$ and up to $\kappa_\lambda \sim 10$ the HH mass spectra generated via ggF are softer; therefore, their experimental acceptance in this regions is reduced. Lastly, the expected limits for $|\kappa_\lambda| > 10$ asymptotically tend to similar values.

We can then extract the observed (expected) upper limits on κ_λ by finding the intersection between the theoretical prediction (red) and the expected limits (dashed) so that κ_λ is constrained at the 95 % CL to be within the interval $-11.1 < \kappa_\lambda < 16.6$ ($-9.2 < \kappa_\lambda < 15.2$).

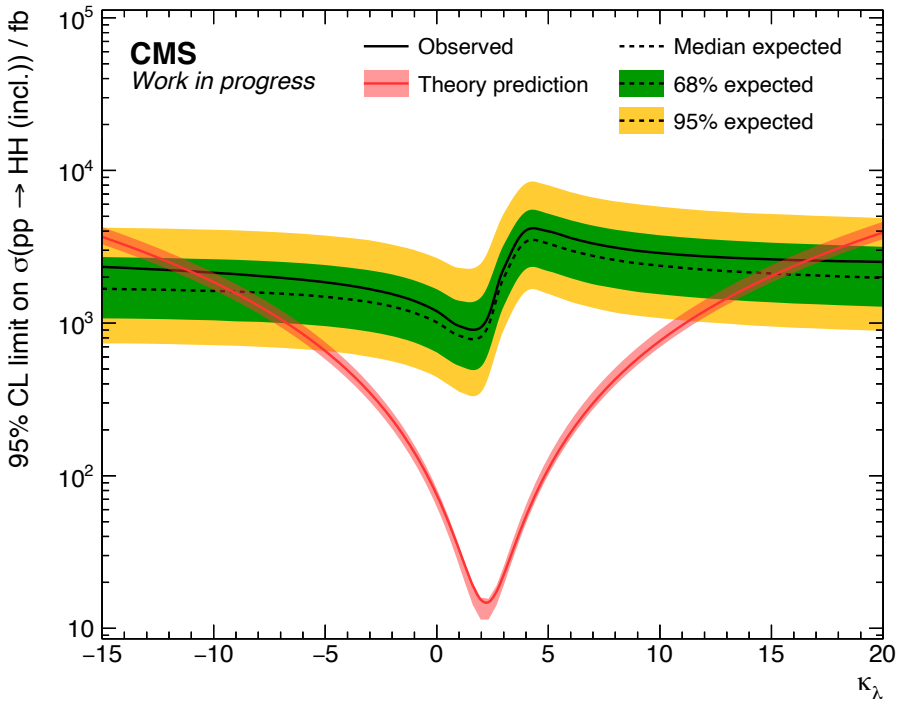


Figure 77: Expected 95% CL upper limits for different values of κ_λ. The green (yellow) bands display the one (two) standard deviation uncertainties in the expected limit. The red line shows the theoretical prediction for the HH cross section.

14.2.1 Closure Test

Given the procedure described in Sec. 3.1.1, we only need three out of the four κ_λ samples that are available to generate any arbitrary point in the (κ_λ, κ_t) space. Therefore, we can extract the limit with respect to the SM prediction for each of the four κ_λ samples in two scenarios: one is when the κ_λ sample in consideration is passed as an input for the production of the additional κ_λ samples, and second is when the considered sample is not used as an input and is instead inferred from the remaining three samples. Theoretically, both of the above methods should yield the same results. However, for certain choices of input samples the procedure may involve large cancellations, since the scaling factors can become much larger than

the κ_λ couplings under consideration, and therefore induce numerical instabilities. The results of the closure test are summarised in Table 17, where we show: the κ_λ sample under test, the expected signal strength (x SM) when the sample is used as an input, the expected signal strength when the sample is inferred from the other κ_λ points, $\pm 1\,\sigma$ variations on the signal strength (absolute values, not errors!) and the percentage difference between these two approaches.

Closure Test of the κ_λ Scan

κ_λ Point	Input Sample		Inferred Sample		Difference
	μ	$\pm\sigma$	μ	$\pm\sigma$	
0	12.6	$^{+7.7}_{-4.3}$	12.5	$^{+7.5}_{-4.4}$	$< 1\%$
1	23.1	$^{+14.8}_{-8.1}$	23.2	$^{+14.6}_{-8.2}$	$< 1\%$
2.45	62.0	$^{+41}_{-19}$	61.9	$^{+39}_{-19}$	$< 1\%$
5	25.8	$^{+14}_{-8.6}$	25.0	$^{+19}_{-8.9}$	3.2%

Table 17: Closure test for the κ_λ scan procedure. On the left most column we have the κ_λ sample under test; the second and third column from the left represent the extracted signal strength (μ) with respect to the SM prediction and its $\pm 1, \sigma$ variations in the case where the κ_λ sample is being passed as an input to the inference algorithm, following the discussion in Sec. 3.1.1; the third and penultimate columns from the right instead represent the signal strength and variations of the κ_λ point under consideration when the sample is being inferred from the remaining κ_λ values. The right-most column represents the percentage difference between these two approaches, and it shows perfect agreement for the first three points and slightly worse agreement for $\kappa_\lambda = 5$, but this is an expected feature given the arisal of numerical instabilities for this particular choice of input values.

Given these results, the official recommendation is to perform the κ_λ scan using the set of points: $\kappa_\lambda = 1\,,\, 2.45\,,\, 5\,$, hence omitting $\kappa_\lambda = 0$. This choice is motivated by the fact that the inclusion of $\kappa_\lambda = 5$ is strongly recommended given the potential arisal of numerical instabilities during the inference procedure; similarly $\kappa_\lambda = 2.45$ should also be included since it captures the variations of HH production around

its minimum. Because of this, we are left with one point to choose, which is going to be the SM $\kappa_\lambda = 1$ as it is a "special" point.

14.3 EFT Results

Upper limits on the HH cross section are also computed for the 12 EFT benchmark scenarios [18]. To allow for the interpretation in terms of the total inclusive HH production cross section, the samples are normalized to an inclusive (and hypothetical) cross section of 1 fb. Similarly to the SM procedure, the non-resonant background is extracted from a fit to Data while the resonant background is modeled from Monte Carlo. The expected EFT upper limits on the gluon gluon Fusion HH production are computed at the 95% CL and are summarised in Table 18 and shown in Fig. 78.

Upper Limits on HH cross section of EFT benchmarks

EFT BM	xsec [fb]
1	681
2	469
3	753
4	957
5	586
6	1140
7	2675
8	633
8a	700
9	606
10	2191
11	1209
12	1542

Table 18: Upper limits on HH cross section for the 12+1 EFT benchmarks.

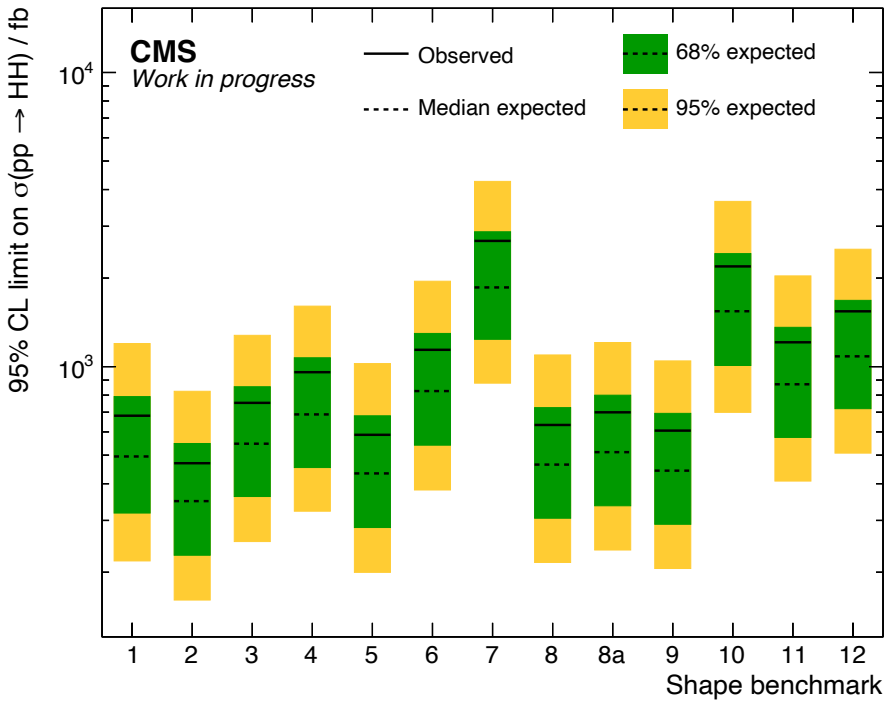


Figure 78: Expected 95% CL upper limits on the non-resonant HH cross section for the 12 EFT benchmarks considered.

15 Conclusion

A search for the production of two scalar bosons in the $\gamma\gamma\tau\tau$ final state is presented. The search uses data from proton-proton collisions collected by the CMS experiment at the LHC in 2016–2018 at a centre-of-mass energy of 13 TeV, corresponding to 138 fb^{-1} of total integrated luminosity. The channel analysed in this thesis targets the non-resonant HH production via gluon-fusion, where no significant deviation from the background-only hypothesis is observed. Upper limits at the 95% CL on the HH production cross section are extracted for production in the SM and in several BSM scenarios. The observed upper limit for the SM production is found to be 764 fb, corresponding to 24.7 times the SM prediction, whilst the expected upper limit is 711 fb, corresponding to 23 times the SM prediction. The limit is also derived as a function of the Higgs boson self-coupling modifier, κ_λ, assuming all other Higgs boson couplings are as predicted in the SM. The coupling modifier, κ_λ, is constrained within the range $-11.1 < \kappa_\lambda < 16.6$ at the 95% CL. In addition, the limit is extract for numerous BSM benchmark scenarios. The results are consistent with the SM predictions.

Part V

Appendices

A Input Features to BDT

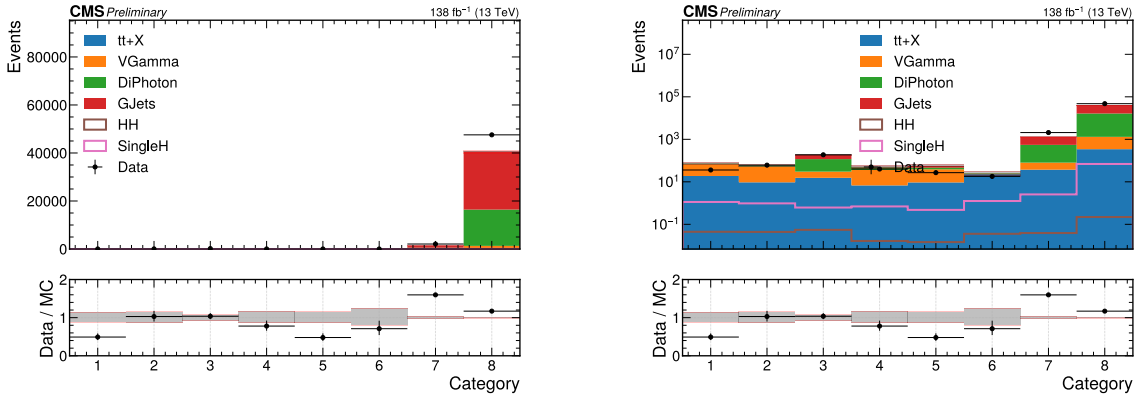


Figure 79: Comparison of data (black points) and simulation (stacked histograms) for the event categorization based on the flavor of ditau pair in linear scale (left) and logarithmic scale (right).

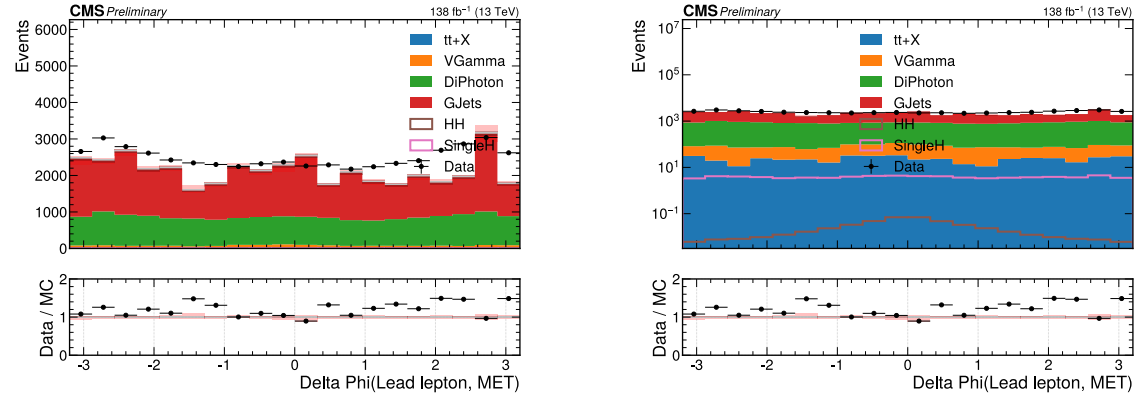


Figure 80: Comparison of data (black points) and simulation (stacked histograms) for the $\Delta\phi$ between the $p_{\mathrm{T}}^{\mathrm{miss}}$ and leading lepton in linear scale (left) and logarithmic scale (right).

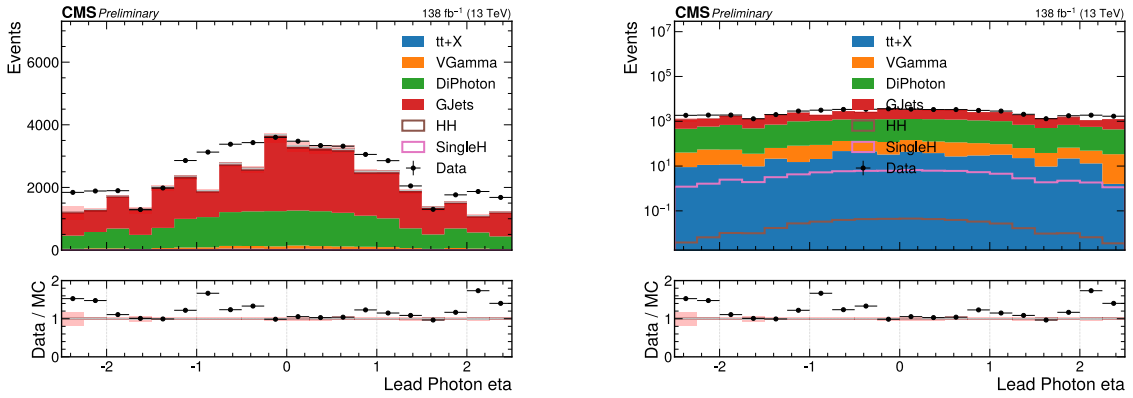


Figure 81: Comparison of data (black points) and simulation (stacked histograms) for the pseudorapidity of the leading photon in linear scale (left) and logarithmic scale (right).

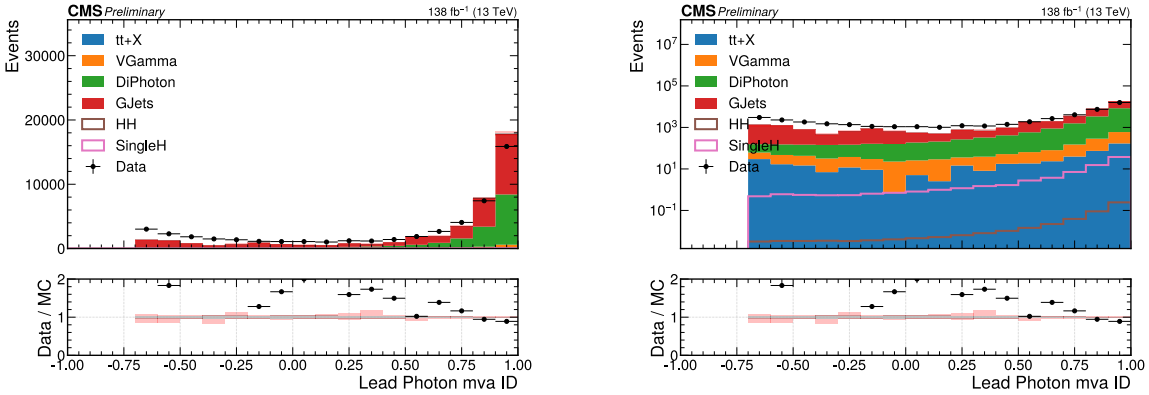


Figure 82: Comparison of data (black points) and simulation (stacked histograms) for the photon ID MVA of the leading photon in linear scale (left) and logarithmic scale (right).

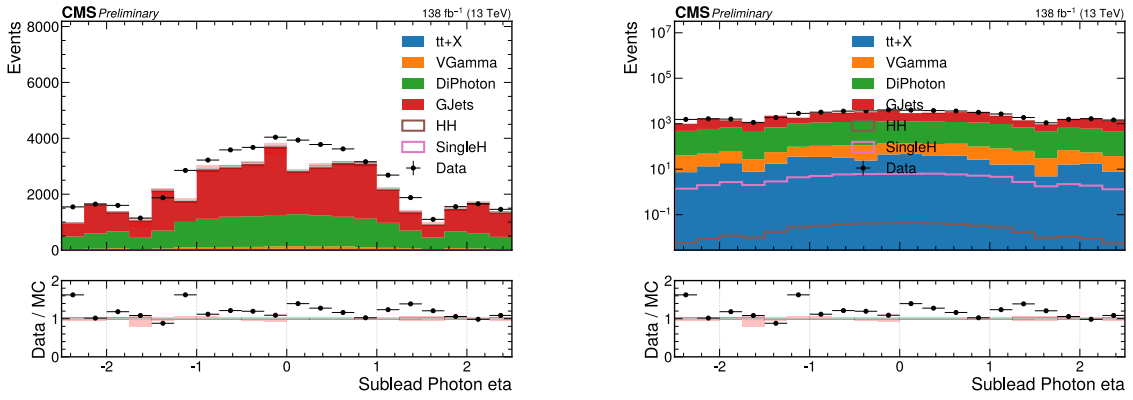


Figure 83: Comparison of data (black points) and simulation (stacked histograms) for the pseudorapidity of the subleading photon in linear scale (left) and logarithmic scale (right).

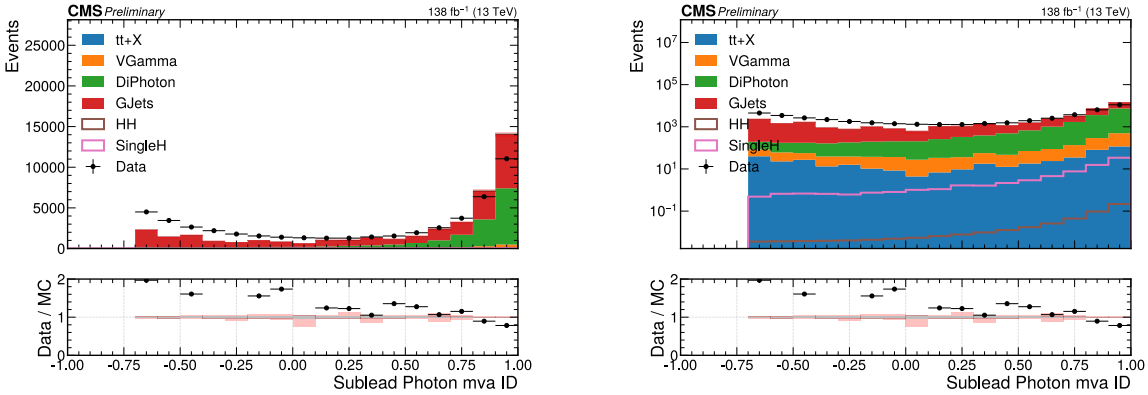


Figure 84: Comparison of data (black points) and simulation (stacked histograms) for the photon ID MVA of the subleading photon in linear scale (left) and logarithmic scale (right).

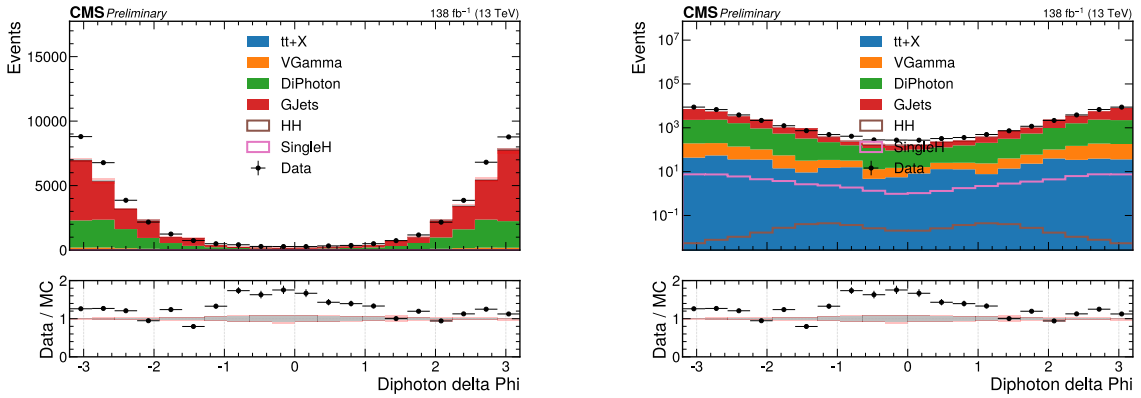


Figure 85: Comparison of data (black points) and simulation (stacked histograms) for the $\Delta\phi$ between the two photons in the diphoton candidate in linear scale (left) and logarithmic scale (right).

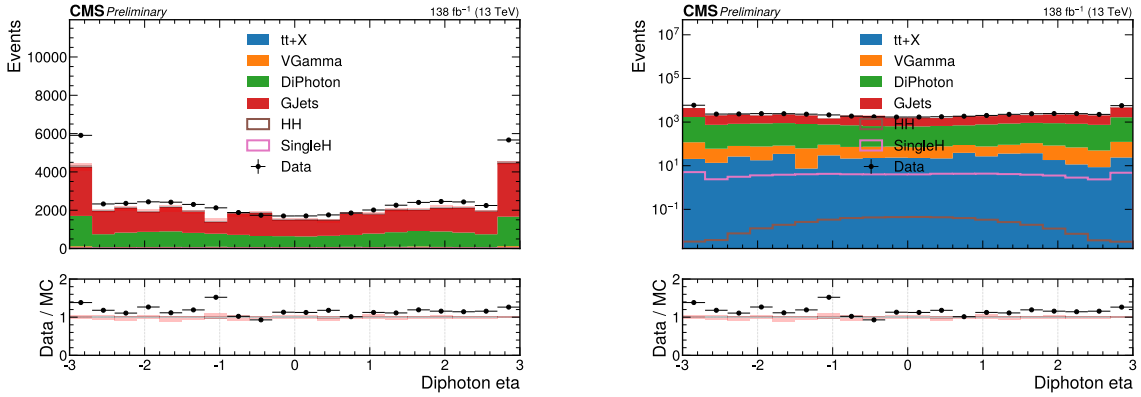


Figure 86: Comparison of data (black points) and simulation (stacked histograms) for the pseudorapidity of the diphoton candidate in linear scale (left) and logarithmic scale (right).

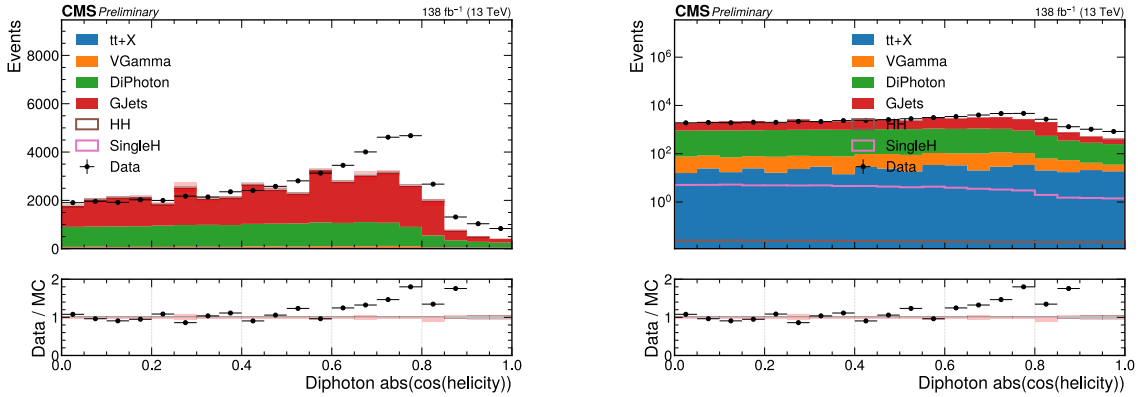


Figure 87: Comparison of data (black points) and simulation (stacked histograms) for the $|\cos(\text{helicity angle}(\theta))|$ of the two photons in the diphoton candidate in linear scale (left) and logarithmic scale (right).

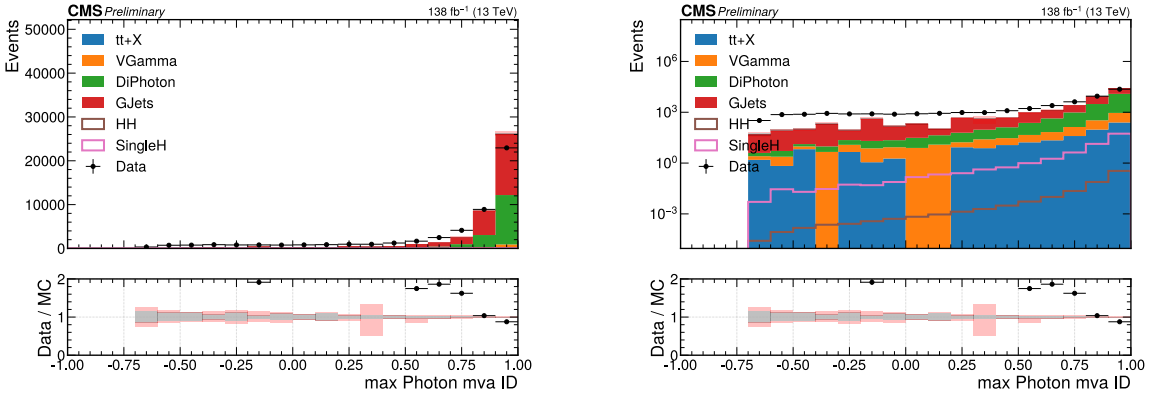


Figure 88: Comparison of data (black points) and simulation (stacked histograms) for the max photon ID MVA in linear scale (left) and logarithmic scale (right).

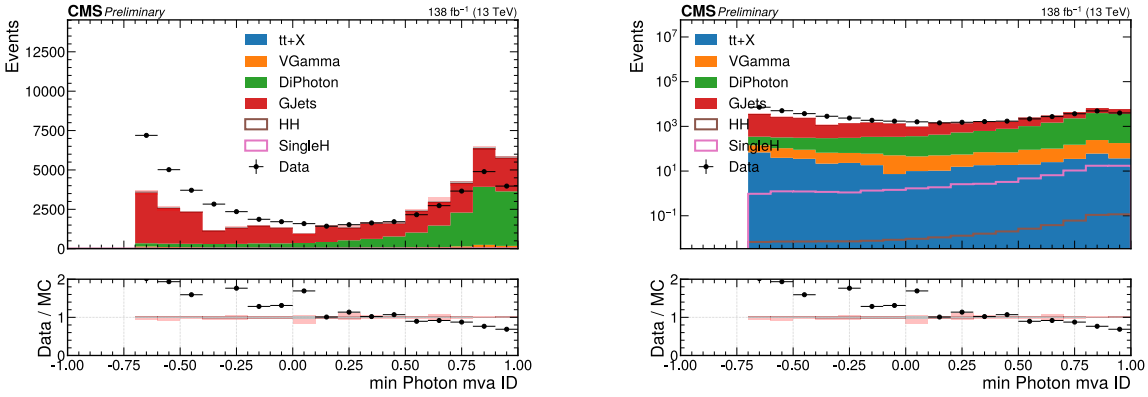


Figure 89: Comparison of data (black points) and simulation (stacked histograms) for the min photon ID MVA in linear scale (left) and logarithmic scale (right).

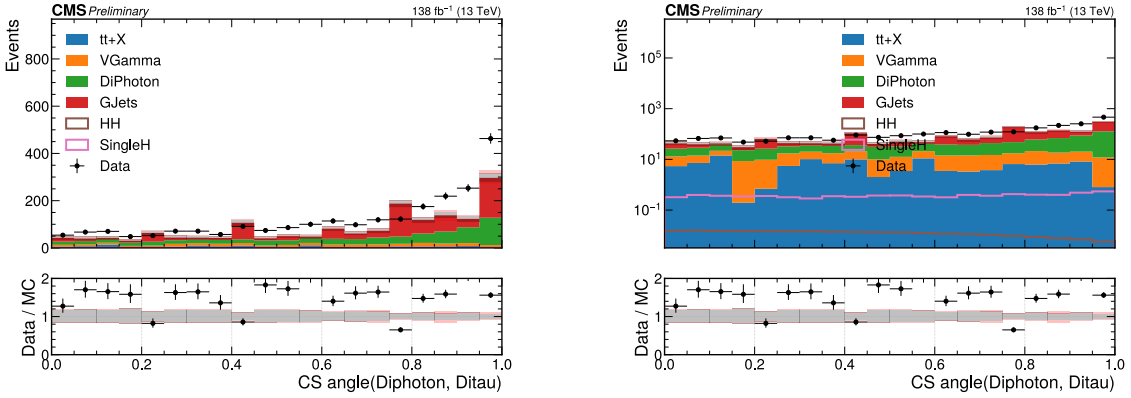


Figure 90: Comparison of data (black points) and simulation (stacked histograms) for the Collins-Soper [69] angle $\cos\theta^{\mathrm{CS}}(\tau\tau^{\mathrm{SVFit}}, \gamma\gamma)$ in linear scale (left) and logarithmic scale (right).

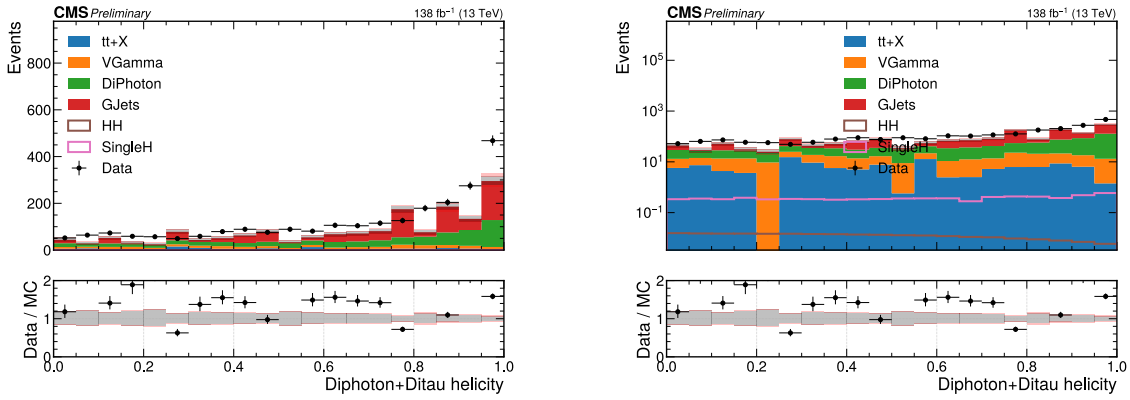


Figure 91: Comparison of data (black points) and simulation (stacked histograms) for the $|\cos(\text{helicity angle}(\theta))|$ of the ditau candidate and the diphoton candidate in linear scale (left) and logarithmic scale (right).

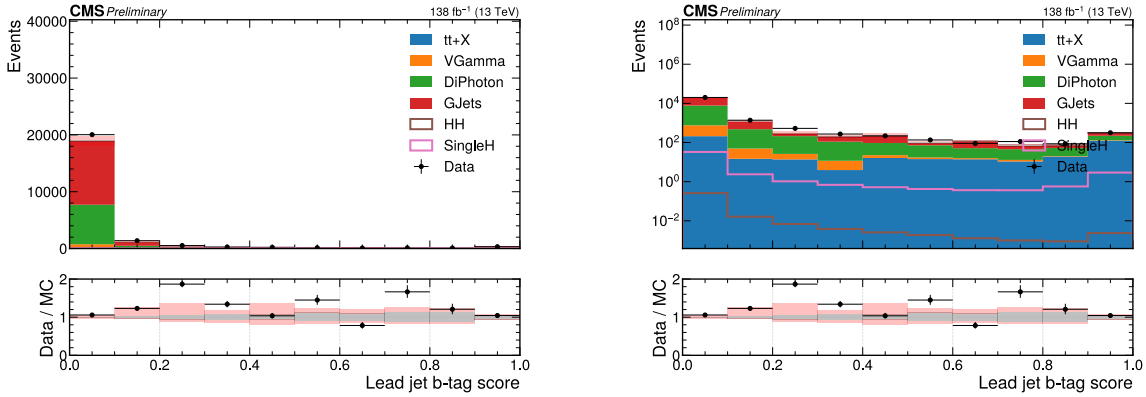


Figure 92: Comparison of data (black points) and simulation (stacked histograms) for the DeepJet b tagging score for the leading jet in linear scale (left) and logarithmic scale (right).

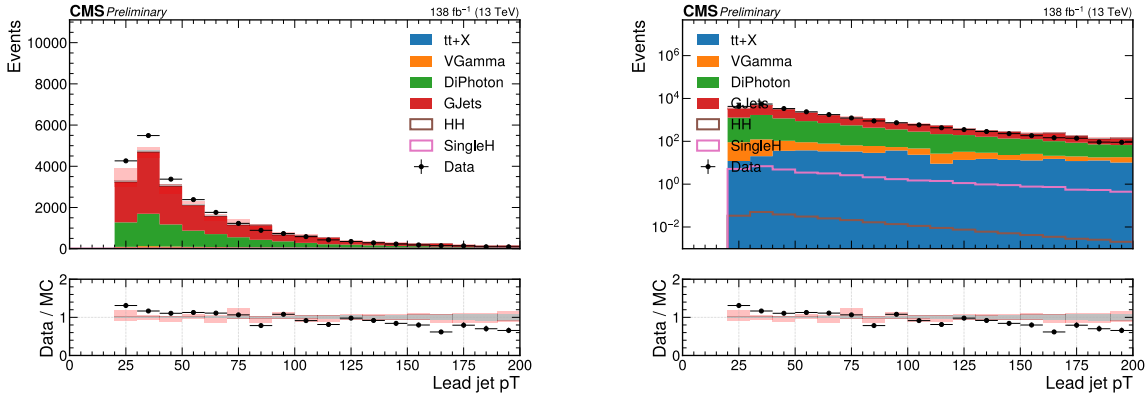


Figure 93: Comparison of data (black points) and simulation (stacked histograms) for the p_{T} of the leading jet in linear scale (left) and logarithmic scale (right).

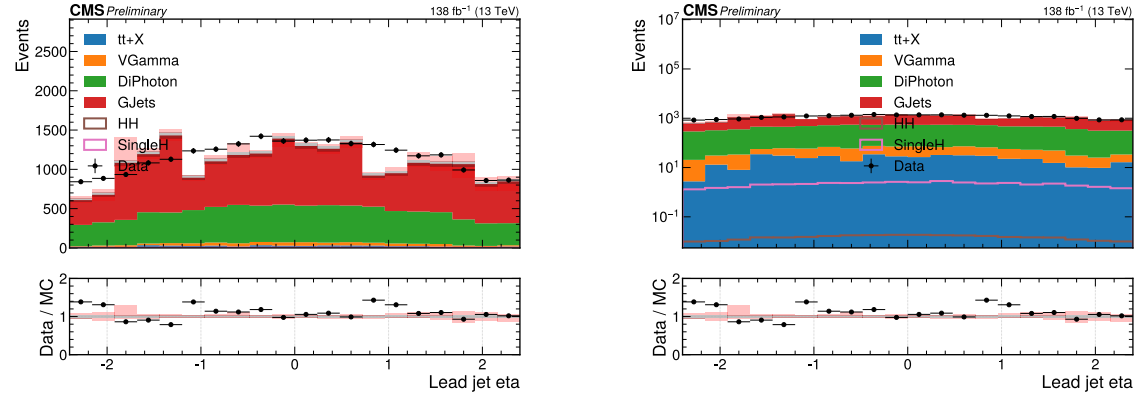


Figure 94: Comparison of data (black points) and simulation (stacked histograms) for the pseudorapidity of the leading jet in linear scale (left) and logarithmic scale (right).

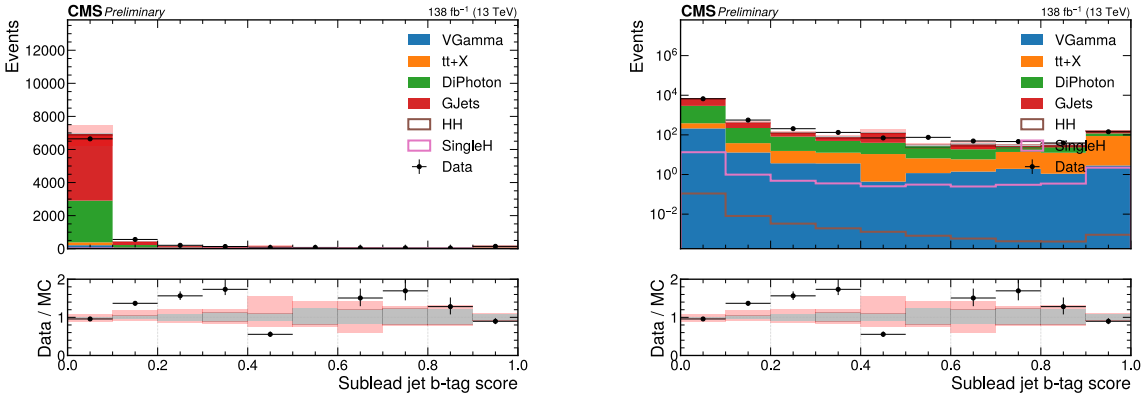


Figure 95: Comparison of data (black points) and simulation (stacked histograms) for the DeepJet b tagging score for the subleading jet in linear scale (left) and logarithmic scale (right).

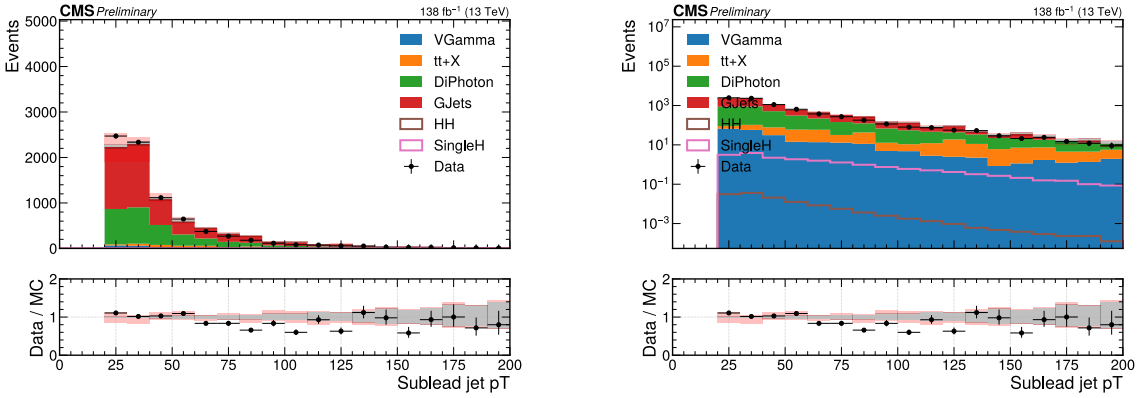


Figure 96: Comparison of data (black points) and simulation (stacked histograms) for the p_{T} of the subleading jet in linear scale (left) and logarithmic scale (right).

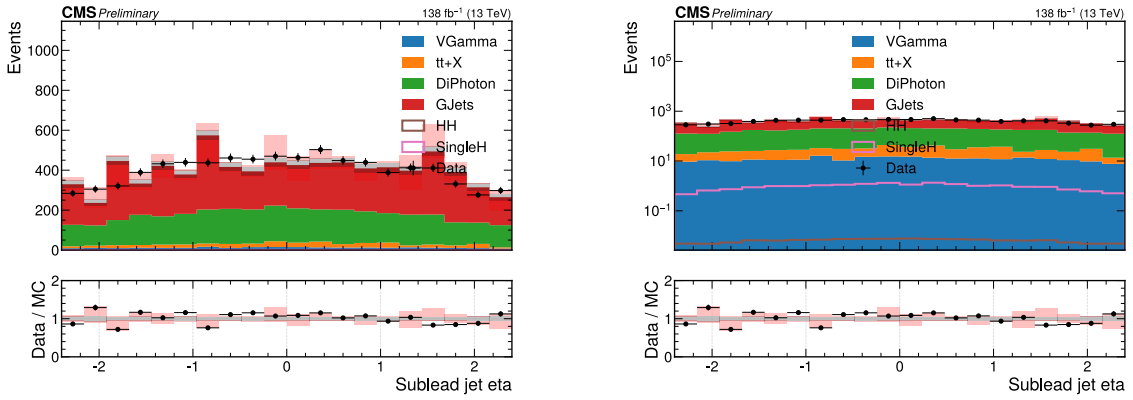


Figure 97: Comparison of data (black points) and simulation (stacked histograms) for the pseudorapidity of the subleading jet in linear scale (left) and logarithmic scale (right).

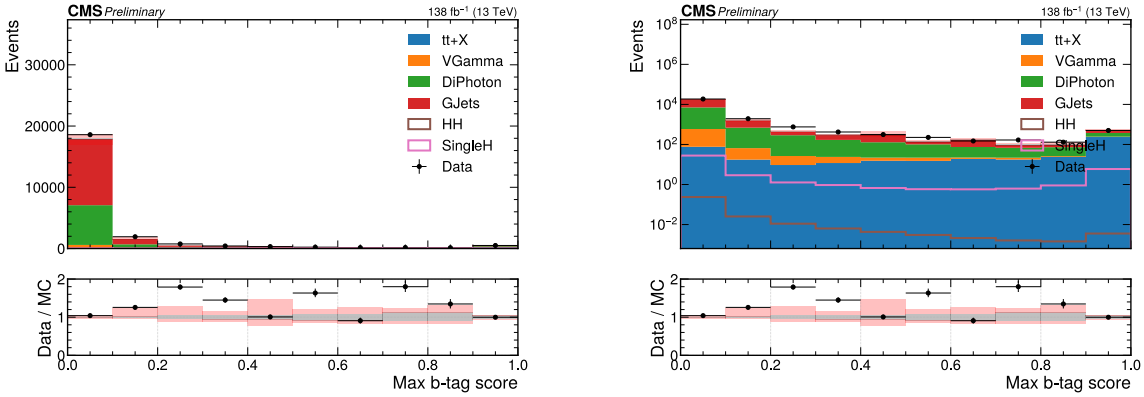


Figure 98: Comparison of data (black points) and simulation (stacked histograms) for the maximum DeepJet b tagging score of all jets in the event in linear scale (left) and logarithmic scale (right).

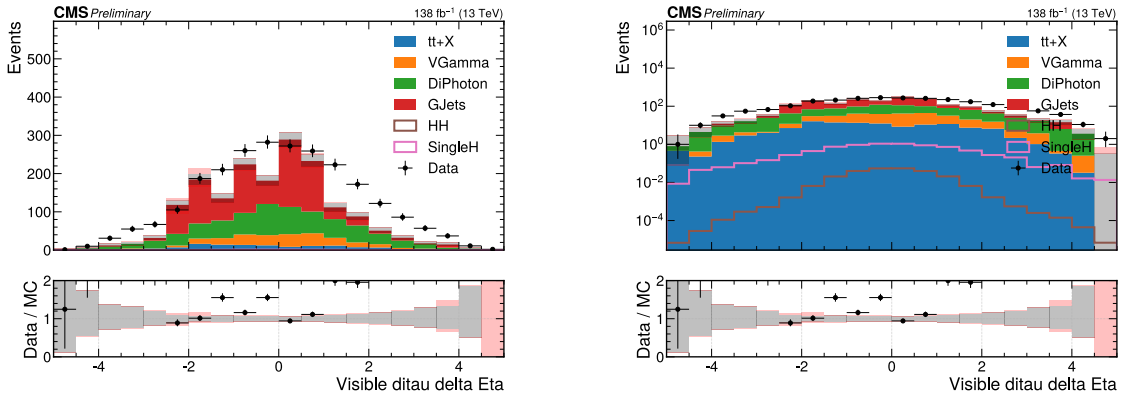


Figure 99: Comparison of data (black points) and simulation (stacked histograms) for the $\Delta\eta$ of the two τ_h/l in the ditau candidate in linear scale (left) and logarithmic scale (right).

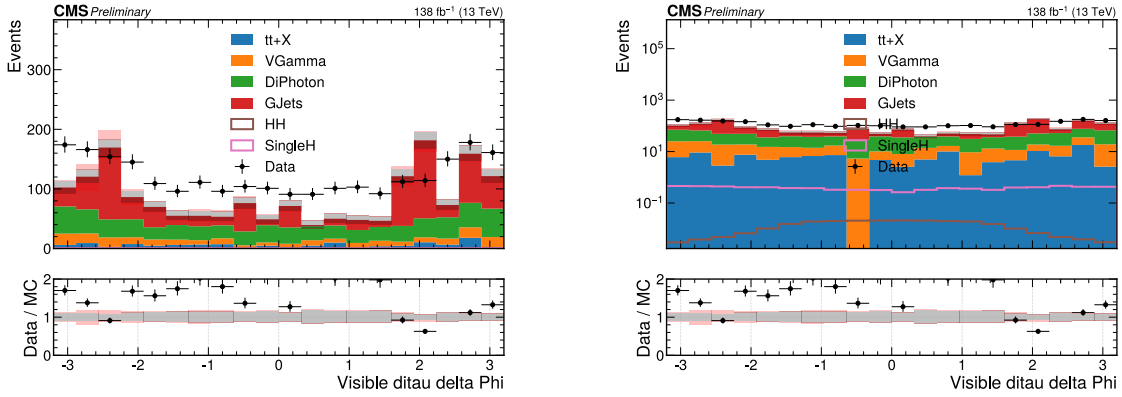


Figure 100: Comparison of data (black points) and simulation (stacked histograms) for the $\Delta\phi$ of the two τ_h/l in the ditau candidate in linear scale (left) and logarithmic scale (right).

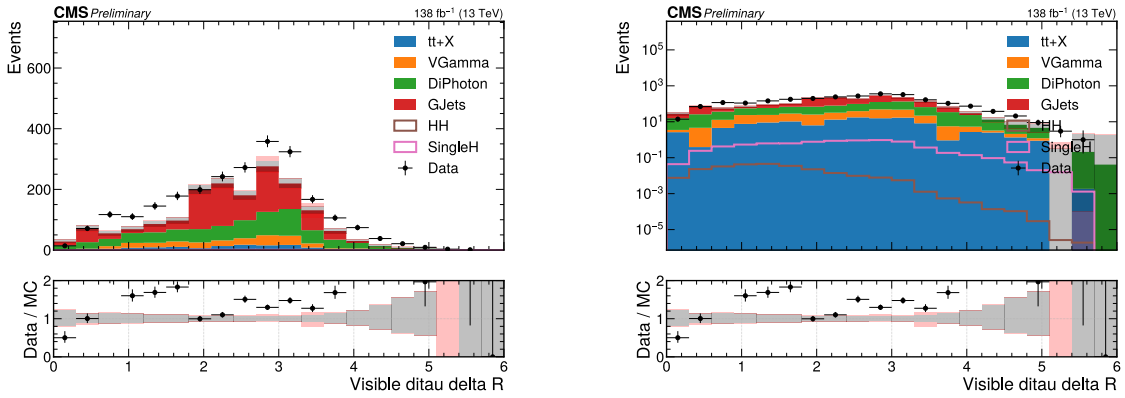


Figure 101: Comparison of data (black points) and simulation (stacked histograms) for the ΔR of the two τ_h/l in the ditau candidate in linear scale (left) and logarithmic scale (right).

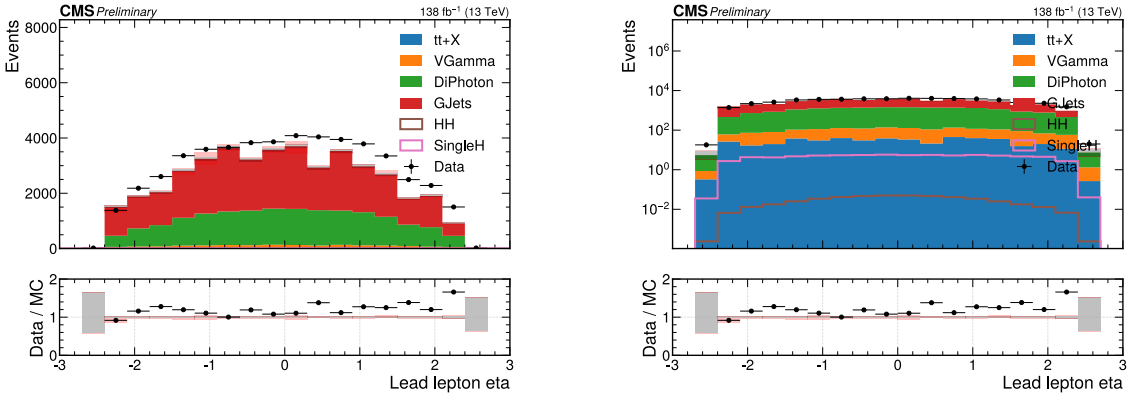


Figure 102: Comparison of data (black points) and simulation (stacked histograms) for the pseudorapidity of the leading τ_h/lepton in linear scale (left) and logarithmic scale (right).

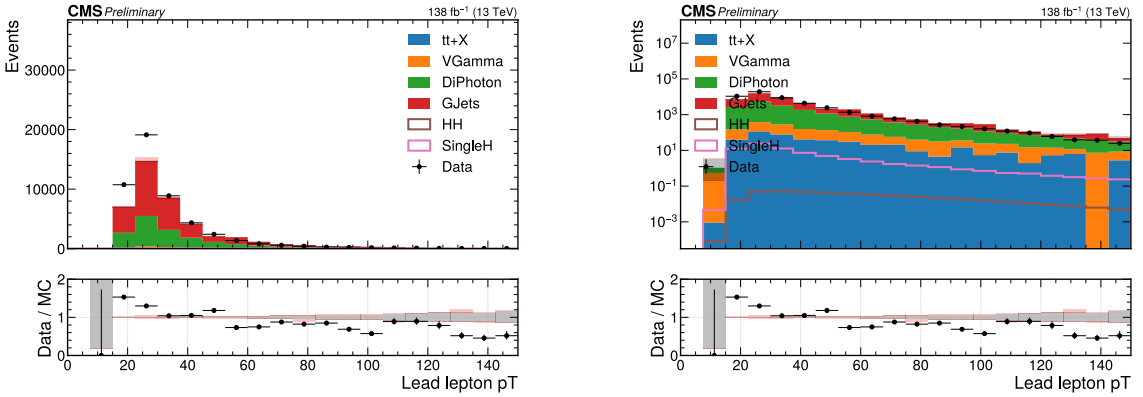


Figure 103: Comparison of data (black points) and simulation (stacked histograms) for the p_{T} of the leading τ_h/lepton in linear scale (left) and logarithmic scale (right).

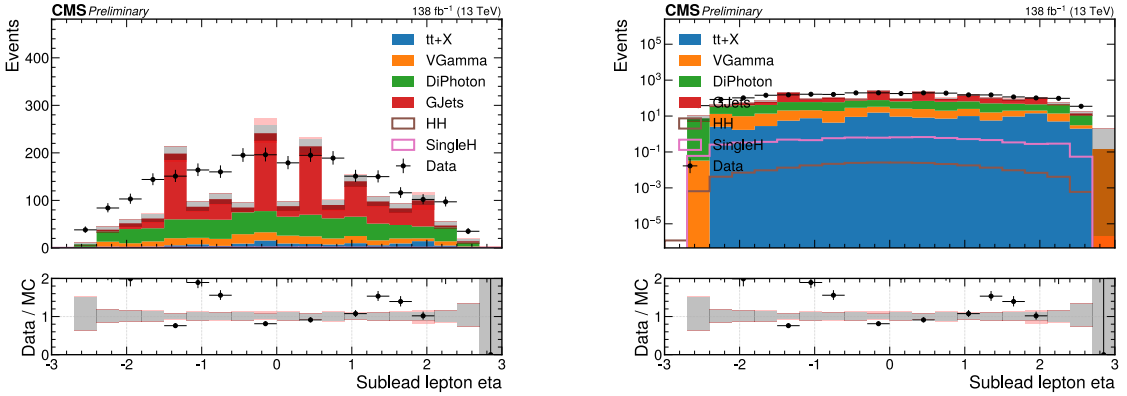


Figure 104: Comparison of data (black points) and simulation (stacked histograms) for the pseudorapidity of the subleading τ_h/lepton in linear scale (left) and logarithmic scale (right).

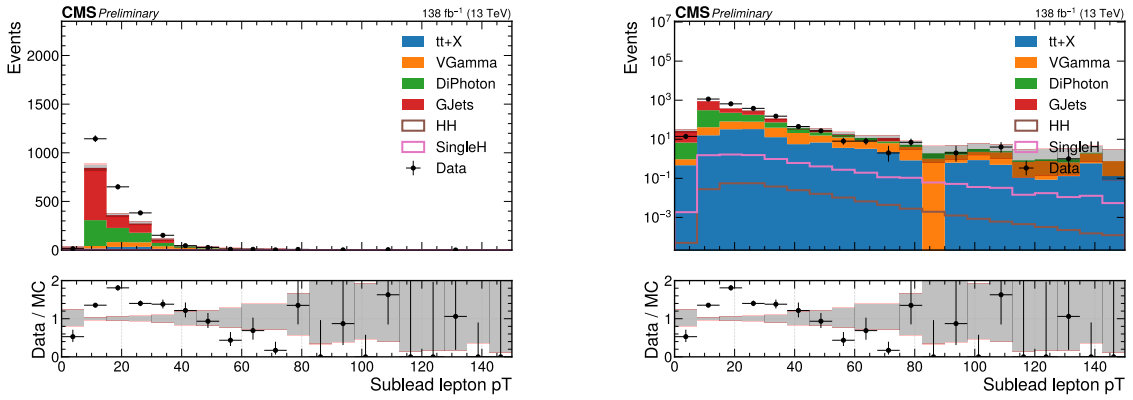


Figure 105: Comparison of data (black points) and simulation (stacked histograms) for the p_{T} of the subleading τ_h/lepton in linear scale (left) and logarithmic scale (right).

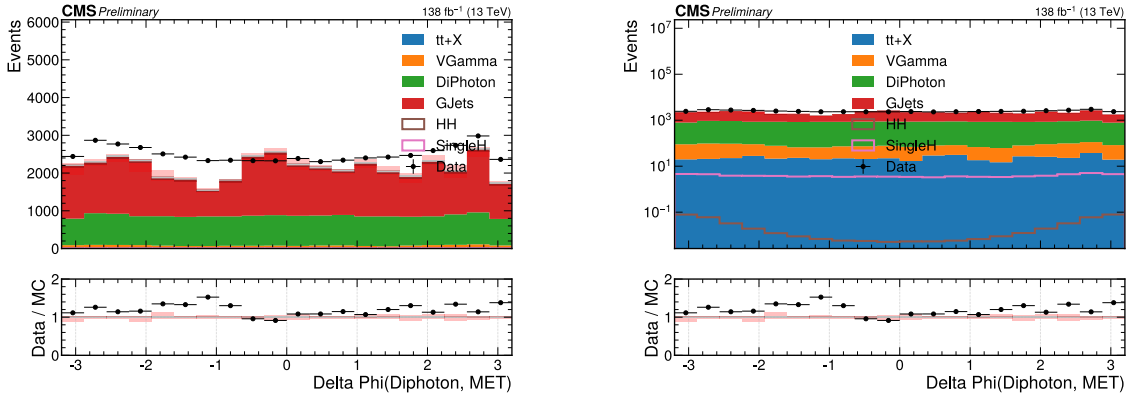


Figure 106: Comparison of data (black points) and simulation (stacked histograms) for the $\Delta\phi$ between the $p_{\mathrm{T}}^{\mathrm{miss}}$ and the diphoton candidate in linear scale (left) and logarithmic scale (right).

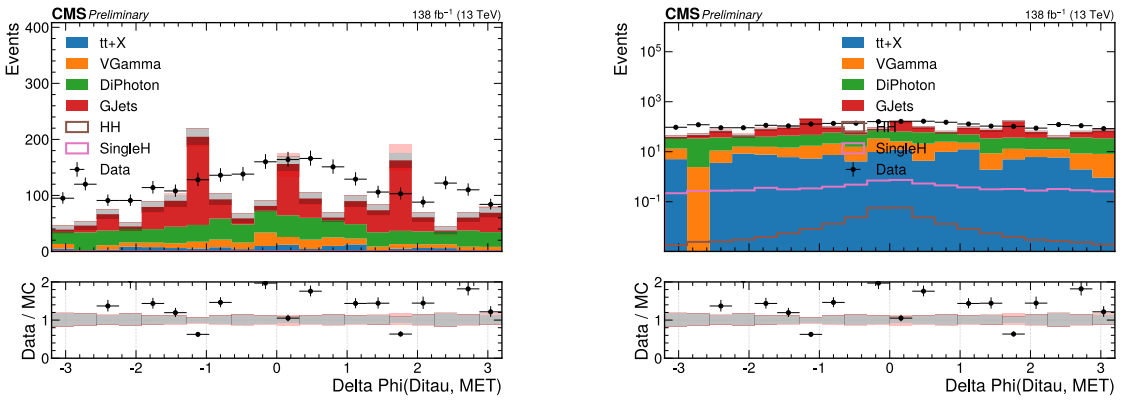


Figure 107: Comparison of data (black points) and simulation (stacked histograms) for the $\Delta\phi$ between the $p_{\mathrm{T}}^{\mathrm{miss}}$ and the ditau candidate in linear scale (left) and logarithmic scale (right).

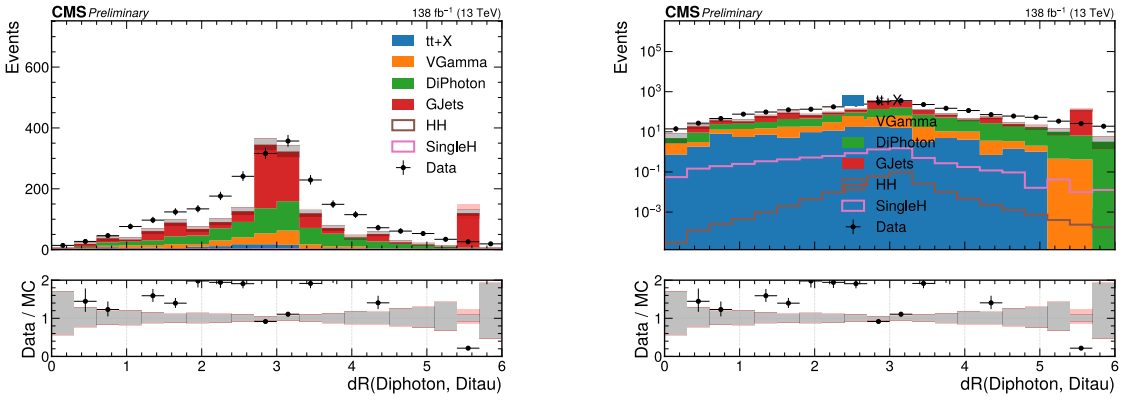


Figure 108: Comparison of data (black points) and simulation (stacked histograms) for the ΔR between the diphoton candidate and the ditau candidate (with the ditau candidate four-vector computed according to the SVFit algorithm) in linear scale (left) and logarithmic scale (right).

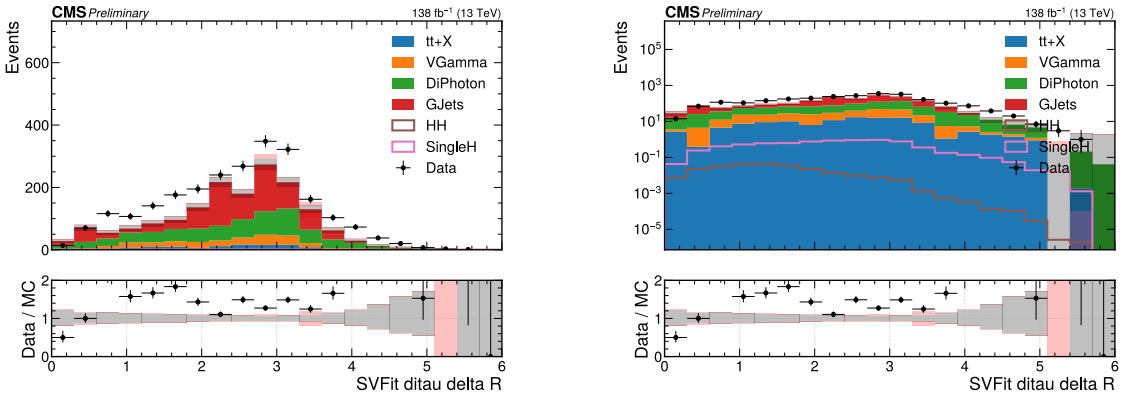


Figure 109: Comparison of data (black points) and simulation (stacked histograms) for the ΔR between the two τ candidates in the ditau candidate (with the four vectors computed according to the SVFit algorithm) in linear scale (left) and logarithmic scale (right).

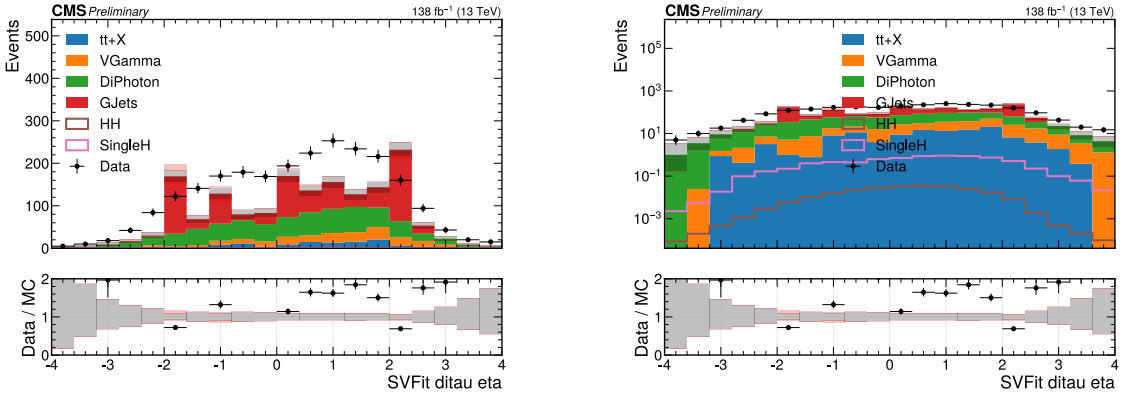


Figure 110: Comparison of data (black points) and simulation (stacked histograms) for the pseudorapidity of the ditau candidate (with the ditau candidate four-vector computed according to the SVFit algorithm) in linear scale (left) and logarithmic scale (right).

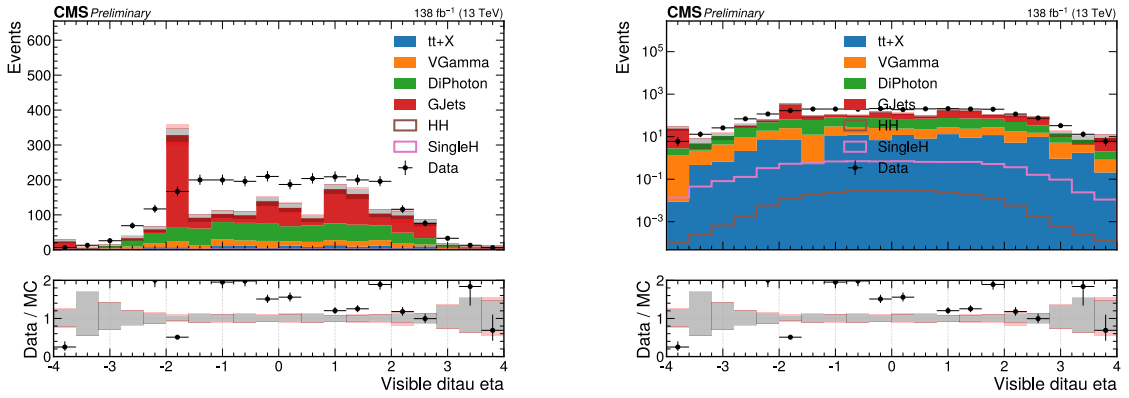


Figure 111: Comparison of data (black points) and simulation (stacked histograms) for the pseudorapidity of the ditau candidate (with the ditau candidate four-vector computed with the visible components of its τ_h/l candidates) in linear scale (left) and logarithmic scale (right).

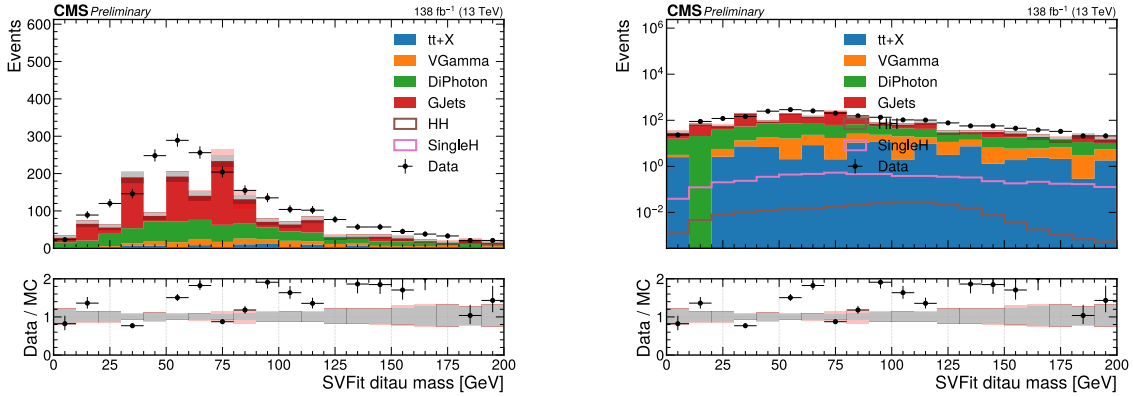


Figure 112: Comparison of data (black points) and simulation (stacked histograms) for the invariant mass of the ditau candidate (with the ditau candidate four-vector computed according to the SVFit algorithm) in linear scale (left) and logarithmic scale (right).

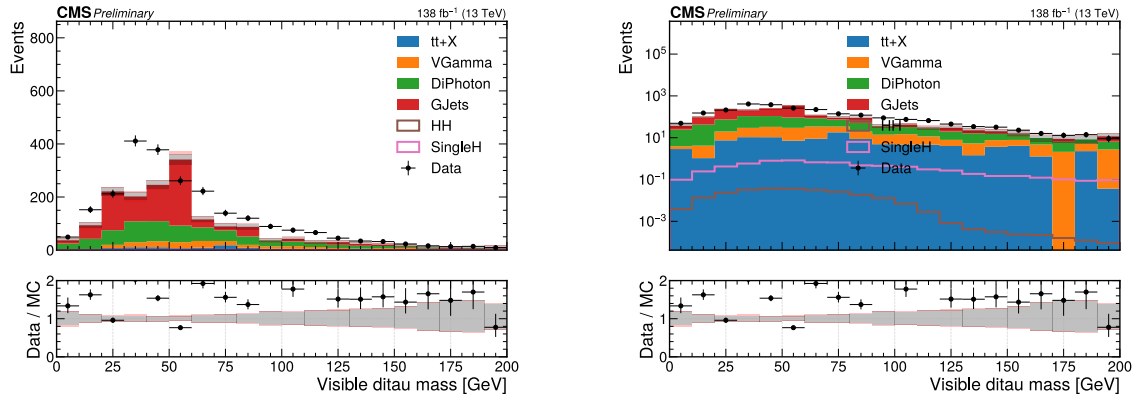


Figure 113: Comparison of data (black points) and simulation (stacked histograms) for the invariant mass of the visible components of the ditau candidate in linear scale (left) and logarithmic scale (right).

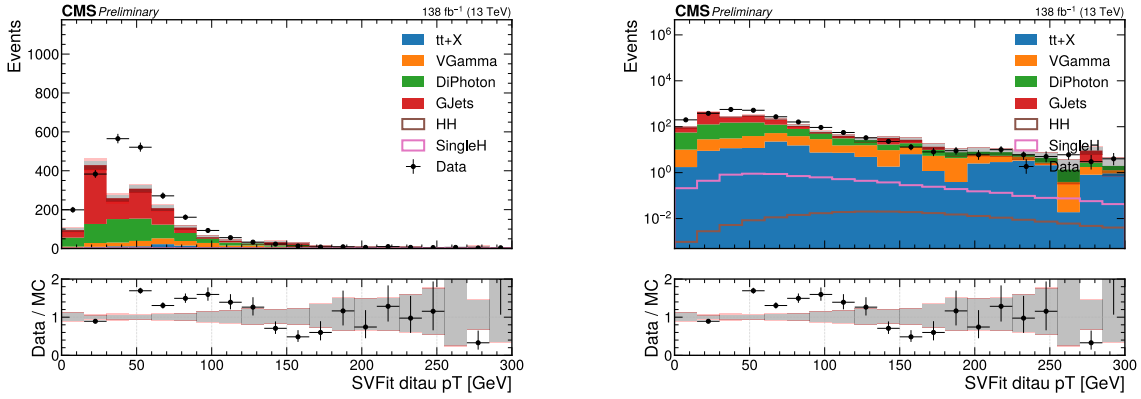


Figure 114: Comparison of data (black points) and simulation (stacked histograms) for the p_{T} of the ditau candidate (with the ditau candidate four-vector computed according to the SVFit algorithm) in linear scale (left) and logarithmic scale (right).

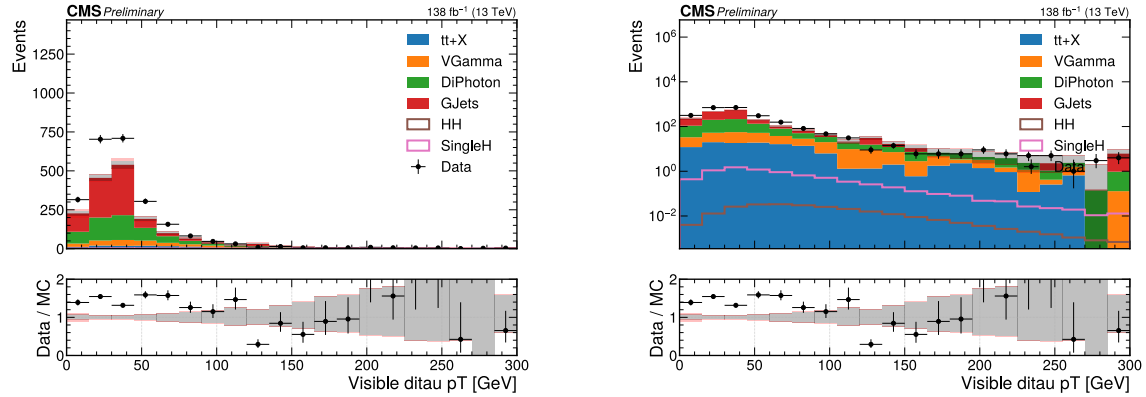


Figure 115: Comparison of data (black points) and simulation (stacked histograms) for the p_{T} of the visible components of the ditau candidate in linear scale (left) and logarithmic scale (right).

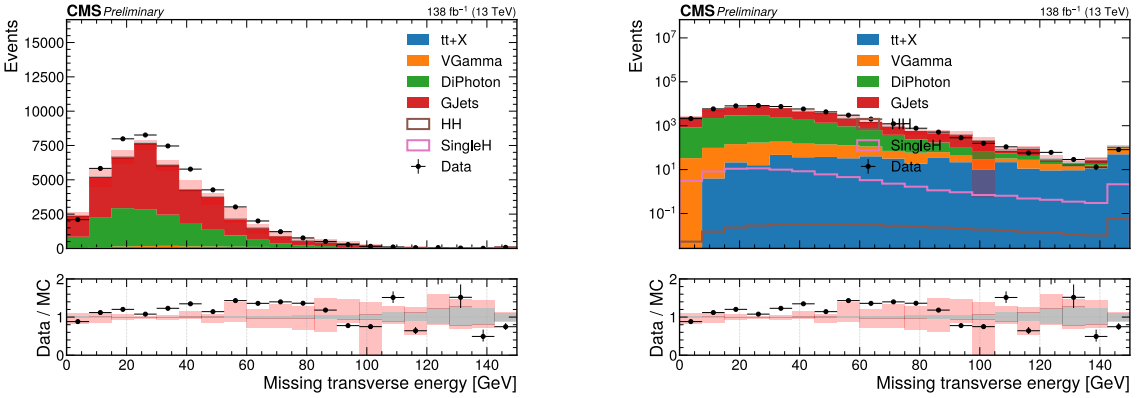


Figure 116: Comparison of data (black points) and simulation (stacked histograms) for the $p_{\mathrm{T}}^{\mathrm{miss}}$ in linear scale (left) and logarithmic scale (right).

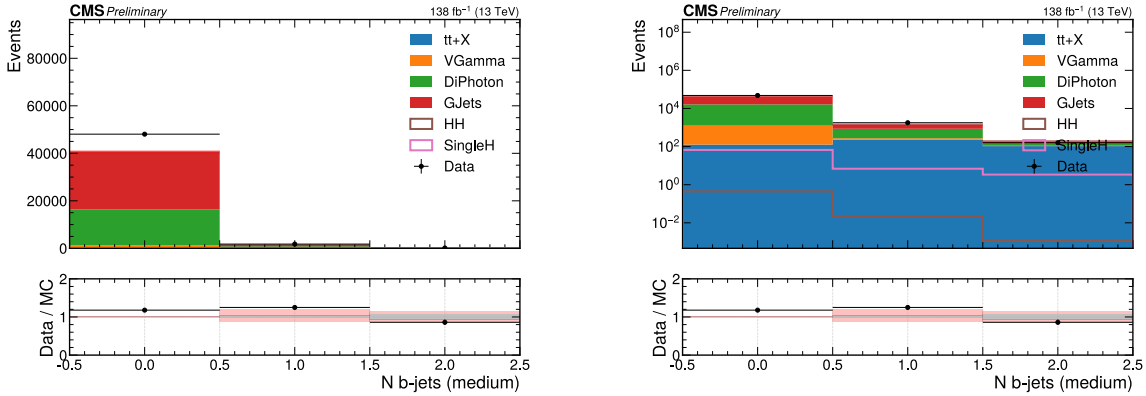


Figure 117: Comparison of data (black points) and simulation (stacked histograms) for the b-jet multiplicity in linear scale (left) and logarithmic scale (right).

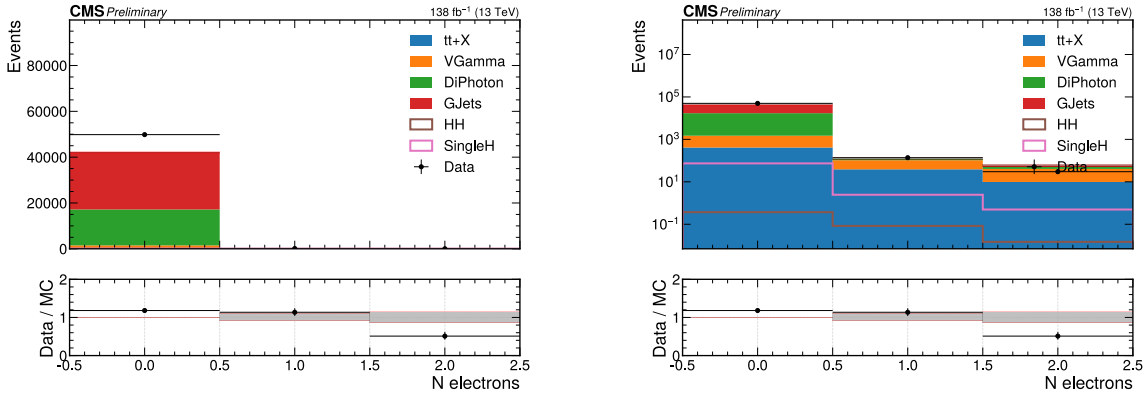


Figure 118: Comparison of data (black points) and simulation (stacked histograms) for the electron multiplicity in linear scale (left) and logarithmic scale (right).

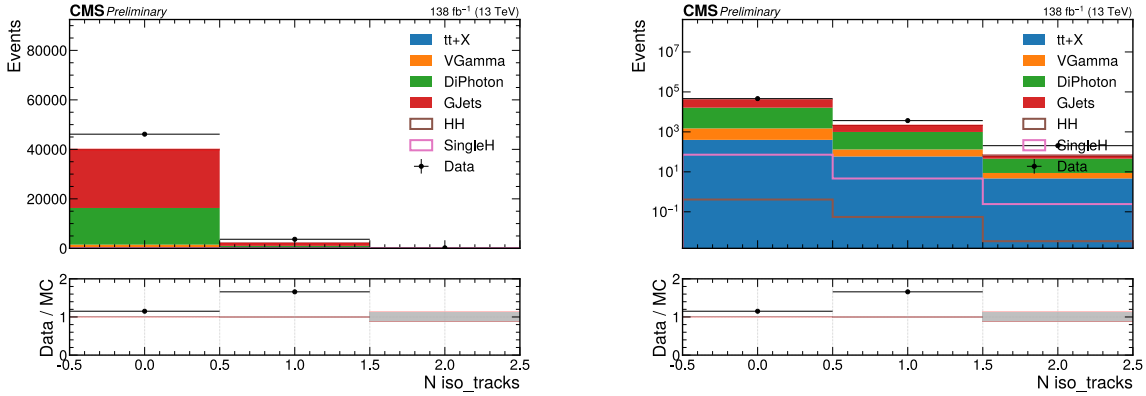


Figure 119: Comparison of data (black points) and simulation (stacked histograms) for the isolated track multiplicity in linear scale (left) and logarithmic scale (right).

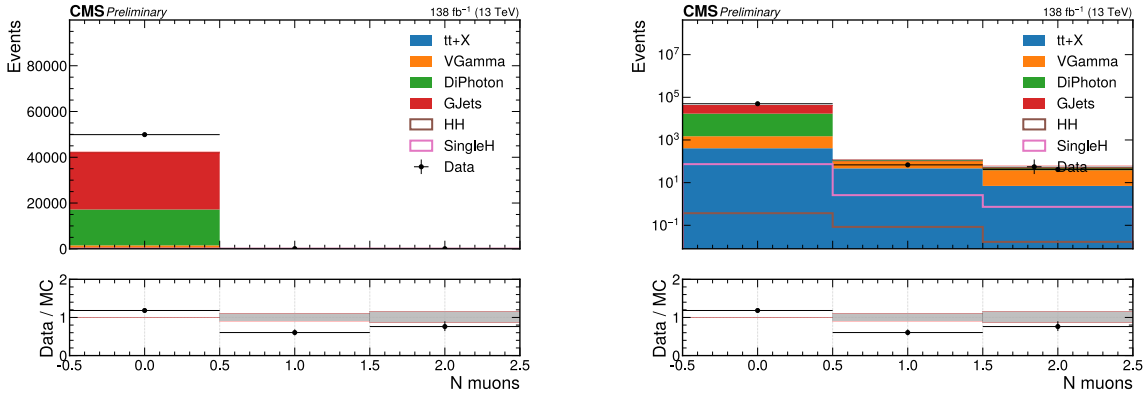


Figure 120: Comparison of data (black points) and simulation (stacked histograms) for the muon multiplicity in linear scale (left) and logarithmic scale (right).

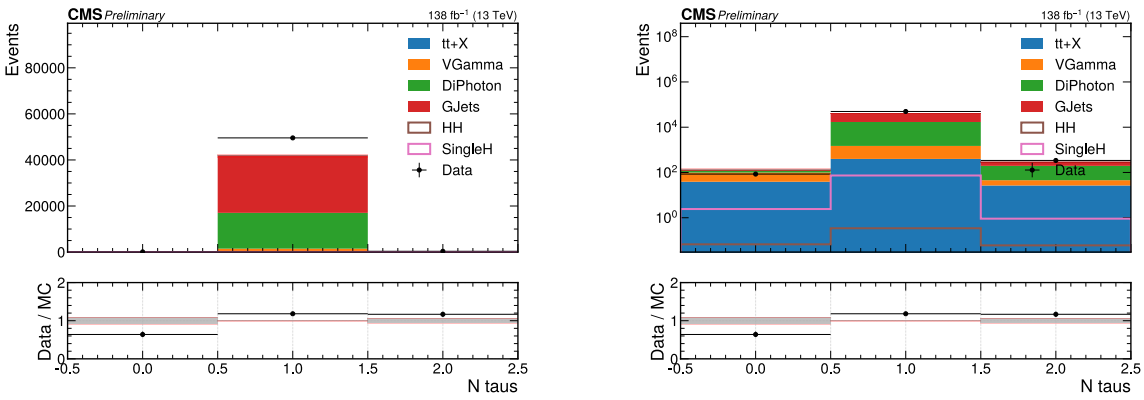


Figure 121: Comparison of data (black points) and simulation (stacked histograms) for the τ_h multiplicity in linear scale (left) and logarithmic scale (right).

B Photon preselection scale factor

The preselection scale factor for photons is estimated comparing Drell-Yan simulation to data, where electron pairs are selected as ”photon candidates” and used to reconstruct the Z mass. Events are required to have a ”tag” photon candidate, with a tight selection and a match to the trigger photon, and a ”probe” photon candidate, with a loose selection. The scale factor is computed comparing the efficiency of the probe to pass the photon preselection used in the analysis. The scale factors are derived by year in 2 bins of $|\eta|$ and R_9, with the 2016 data taking year is split into the pre-VFP and post-VFP periods.

The datasets used are the ones reported in table **??**. The simulated samples are listed below:

- `DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX_2016APV`
- `DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX_2016`
- `DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX_2017`
- `DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX_2018`

The trigger path `HLT_Ele32_eta2p1_WPTight_Gsf` is used for 2016 data and simulation, while the path `HLT_Ele32_WPTight_Gsf_L1DoubleEG` is used for 2017 and 2018.

Events are required to have two photons with $p_{\mathrm{T}} > 20$ GeV, $|\eta| < 2.5$ and $|\eta| < 1.4442$ OR $|\eta| > 1.566$. The photons are required to fail the electron veto requirement, with this being the main difference from the analysis selections, as we are

working with electrons as photon candidates instead of genuine photons. Only events with 2 photons are used to simplify the later steps. All the names of the variables used in this appendix are the same as in section 4.

The tag photon must fall into one of the categories defined in R_9bins (table 19), used in the analysis.

Table 19: Additional photon requirements for barrel and endcap photons at different ranges of R_9, intended to mimic the HLT requirements.

	R_9	$\sigma_{i\eta i\eta}$	$\mathcal{I}_{\text{ph}}$ (GeV)	$\mathcal{I}_{\text{tk}}$ (GeV)
Barrel	[0.50, 0.85]	<0.015	<4.0	<6.0
	>0.85	—	—	—
Endcaps	[0.80, 0.90]	<0.035	<4.0	<6.0
	>0.90	—	—	—

Additionally it is required to pass the following requirements:

- $p_{\text{T}} > 40$ GeV
- $\mathcal{I}_{\text{ch}} < 20$ GeV
- $\mathcal{I}_{\text{ch}} < 0.3$
- H/E < 0.08
- $R_9 > 0.8$
- EGM photon ID MVA > -0.7
- match to a trigger photon within ΔR = 0.3

The probe photon can pass or fail the following requirements:

- R_9category, as for the tag
- $p_{\text{T}} > 30(20)$ GeV for leading (subleading) photon candidate

- $\mathcal{I}_{\text{ch}} < 20$ GeV OR $\mathcal{I}_{\text{ch}} < 0.3$ OR $R_9 > 0.8$
- H/E < 0.08
- EGM photon ID MVA > -0.7

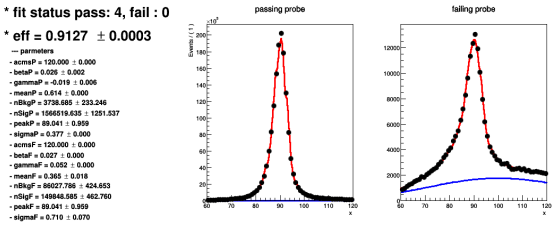
Diphoton candidates are built using the tag and the probe. The invariant mass of the Z peak from the candidates is fitted both in data and simulation in the pass and fail region of the probe. A floating background is allowed. The integral of the fitted signal function is used to derive the selection efficiency and scale factors. Statistical uncertainties and systematics due to the signal and background models are considered. The fits were done using the EGamma tag and probe framework (https://github.com/cms-egamma/egm_tnp_analysis).

The scale factor obtained are listed in table 20.

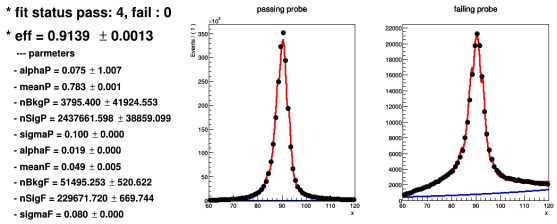
Table 20: scale factors and uncertainties.

	2016 APV	2016	2017	2018
EB high R9	1.002+-0.021	1.001+-0.021	1.003+-0.021	0.992+-0.023
EB low R9	1.008+-0.023	1.004+-0.023	1.013+-0.019	1.001+-0.027
EE high R9	1.010+-0.018	1.006+-0.018	1.013+-0.015	0.995+-0.019
EE low R9	1.013+-0.025	1.011+-0.027	1.021+-0.026	0.985+-0.033

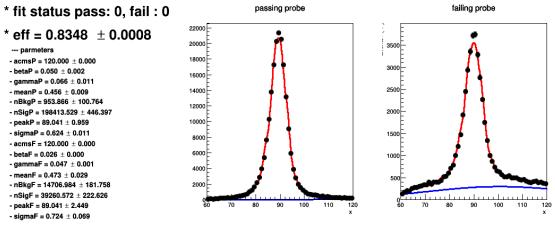
The figures below show the fits used to compute the scale factors and the uncertainties.



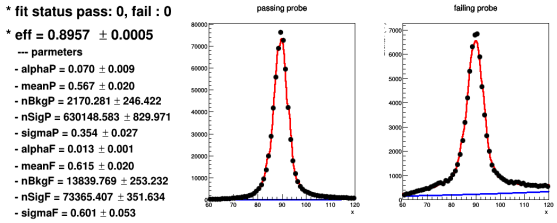
(a) low R_9category EB



(b) high R_9category EB

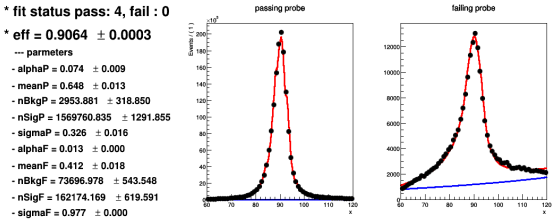


(c) low R_9category EE

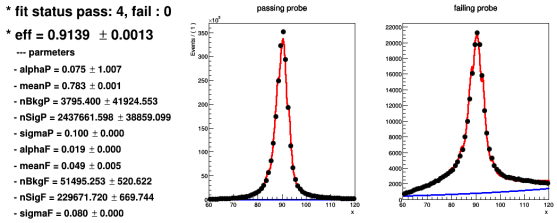


(d) high R_9category EE

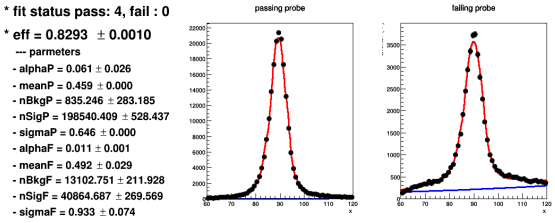
Figure 122: Data nominal fit 2016 APV



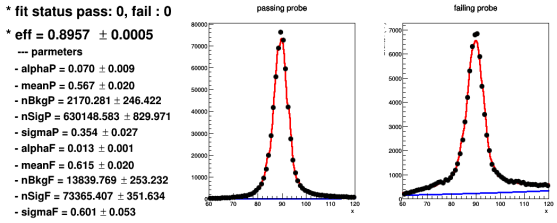
(a) low R_9category EB



(b) high R_9category EB

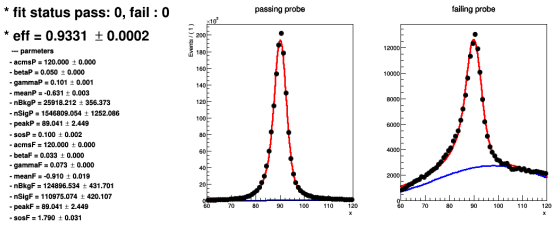


(c) low R_9category EE

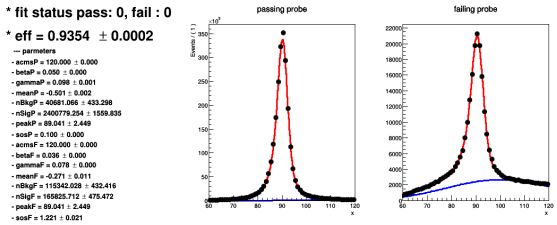


(d) high R_9category EE

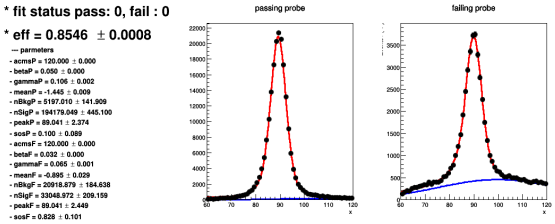
Figure 123: Data alternative background fit 2016 APV



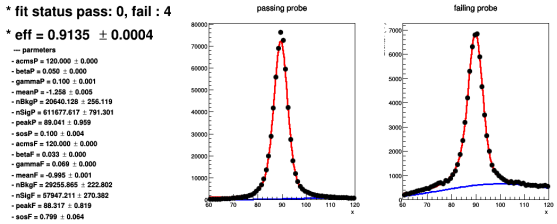
(a) low R_9category EB



(b) high R_9category EB

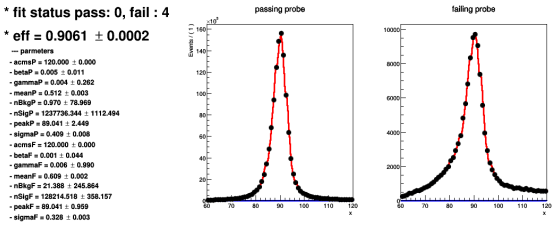


(c) low R_9category EE

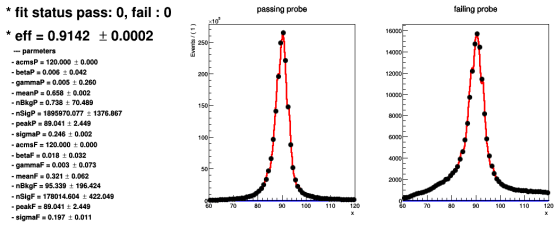


(d) high R_9category EE

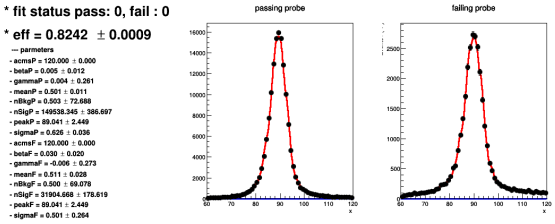
Figure 124: Data alternative signal fit 2016 APV



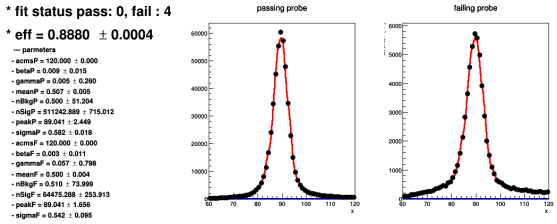
(a) low R_9category EB



(b) high R_9category EB

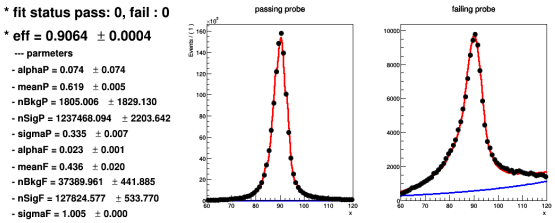


(c) low R_9category EE

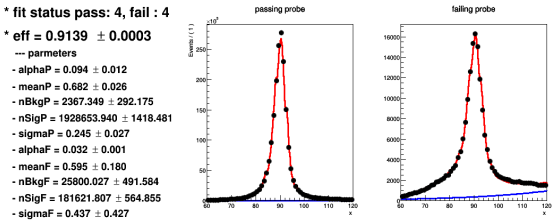


(d) high R_9category EE

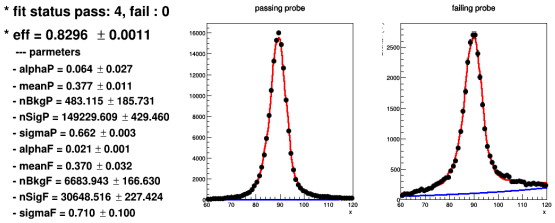
Figure 125: MC nominal fit 2016 APV



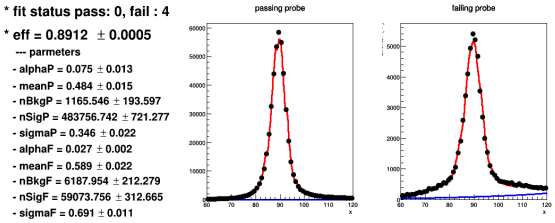
(a) low R_9category EB



(b) high R_9category EB

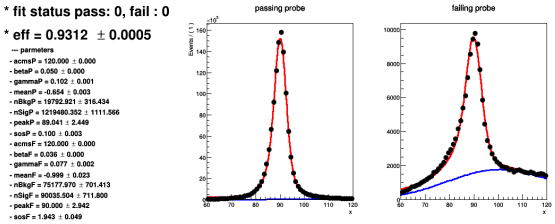


(c) low R_9category EE

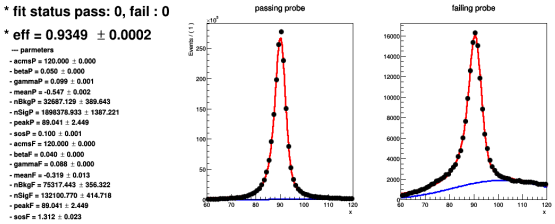


(d) high R_9category EE

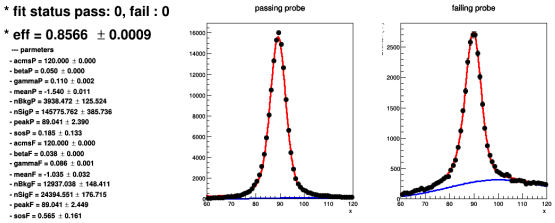
Figure 126: Data nominal fit 2016



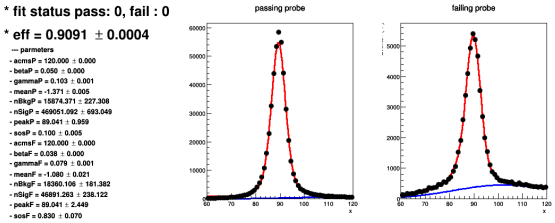
(a) low R_9category EB



(b) high R_9category EB

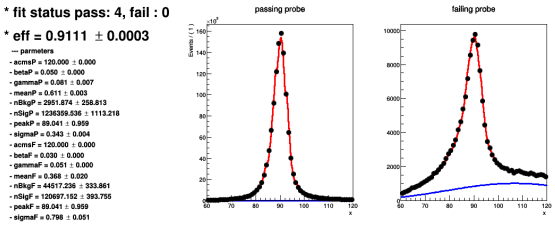


(c) low R_9category EE

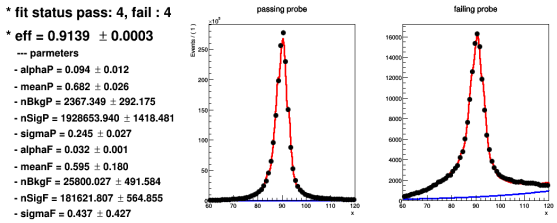


(d) high R_9category EE

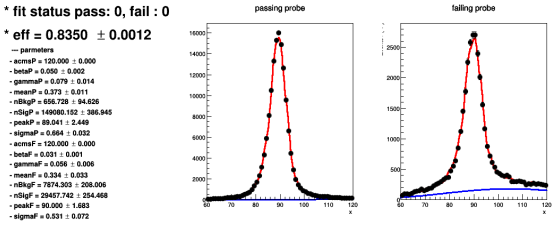
Figure 127: Data alternative background fit 2016



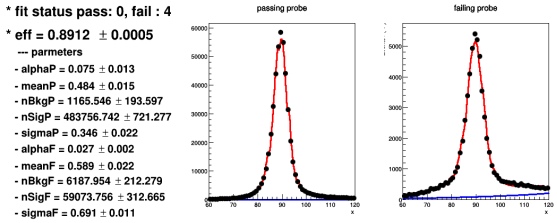
(a) low R_9category EB



(b) high R_9category EB

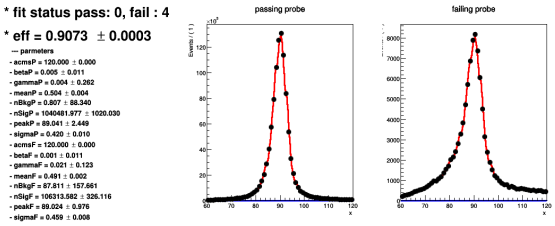


(c) low R_9category EE

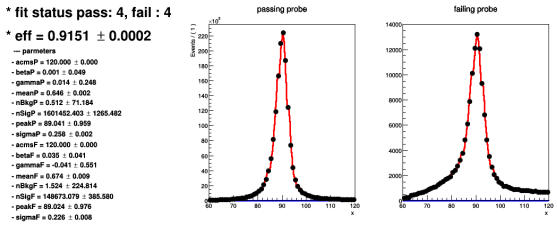


(d) high R_9category EE

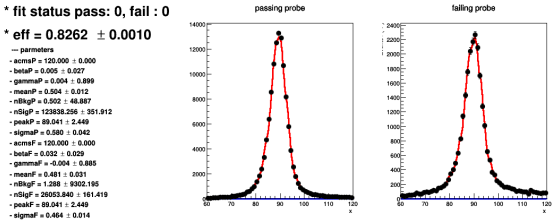
Figure 128: Data alternative signal fit 2016



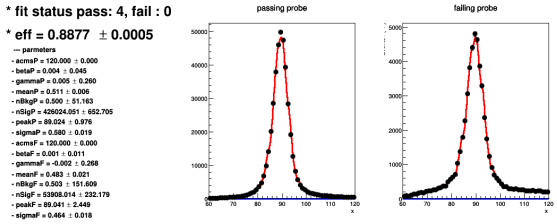
(a) low R_9category EB



(b) high R_9category EB

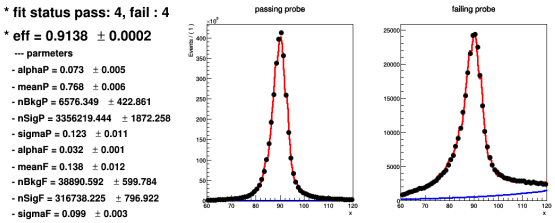


(c) low R_9category EE

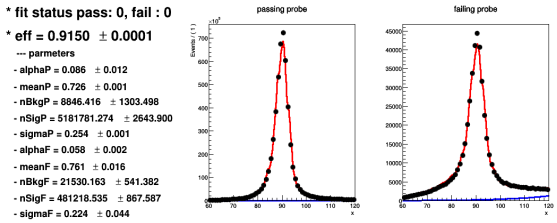


(d) high R_9category EE

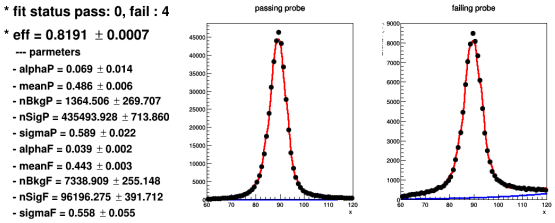
Figure 129: MC nominal fit 2016



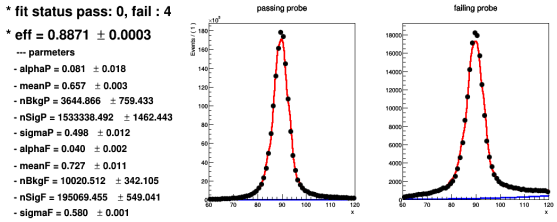
(a) low R_9category EB



(b) high R_9category EB

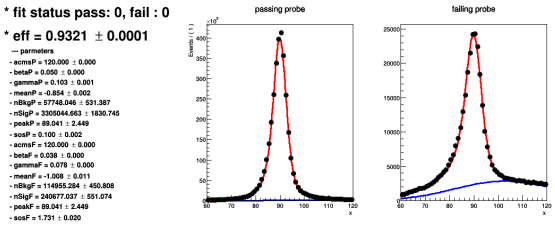


(c) low R_9category EE

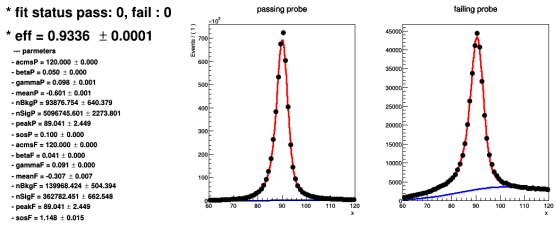


(d) high R_9category EE

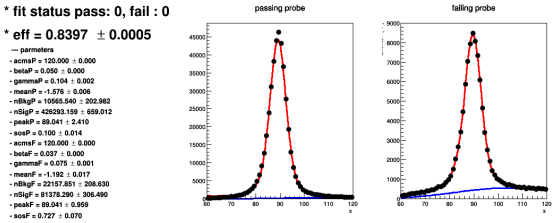
Figure 130: Data nominal fit 2017



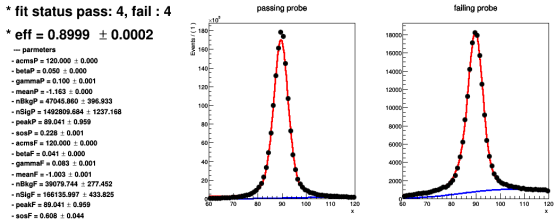
(a) low R_9category EB



(b) high R_9category EB

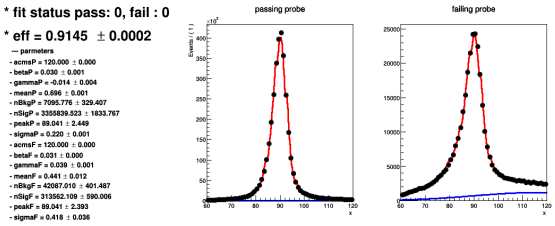


(c) low R_9category EE

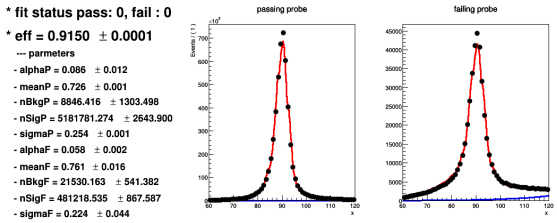


(d) high R_9category EE

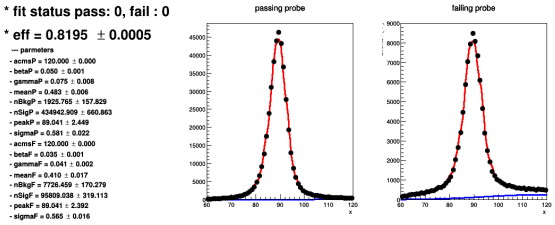
Figure 131: Data alternative background fit 2017



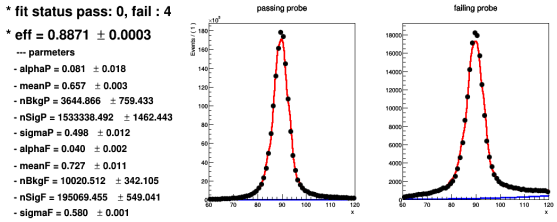
(a) low R_9category EB



(b) high R_9category EB

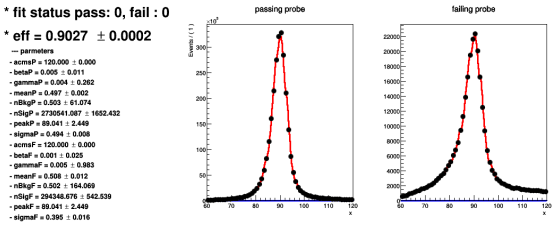


(c) low R_9category EE

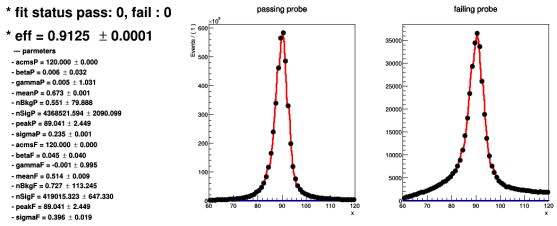


(d) high R_9category EE

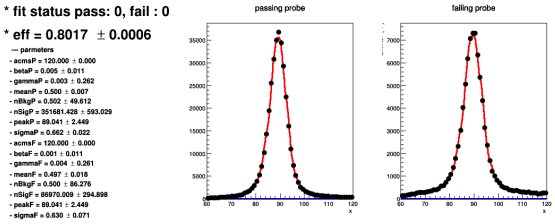
Figure 132: Data alternative signal fit 2017



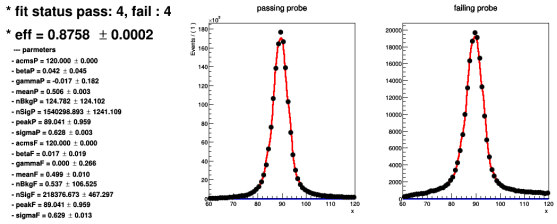
(a) low R_9category EB



(b) high R_9category EB

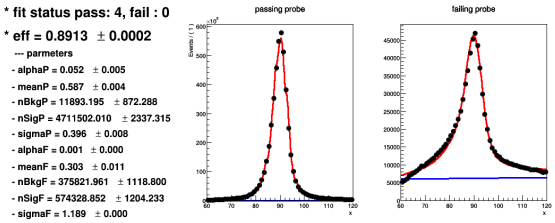


(c) low R_9category EE

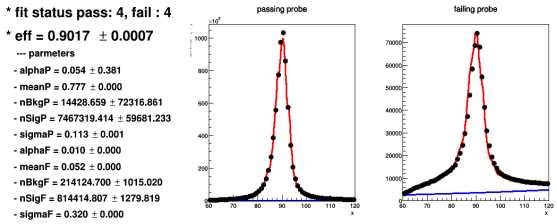


(d) high R_9category EE

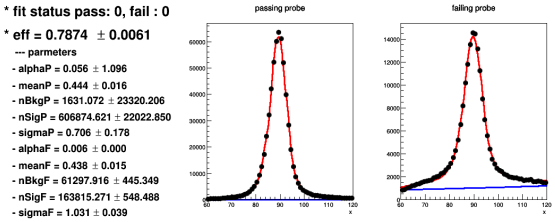
Figure 133: MC nominal fit 2017



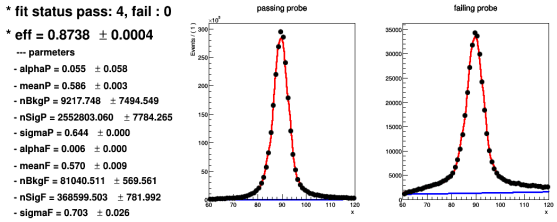
(a) low R_9category EB



(b) high R_9category EB

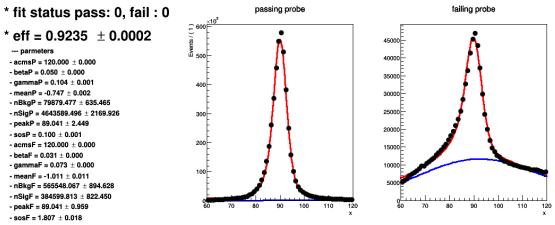


(c) low R_9category EE

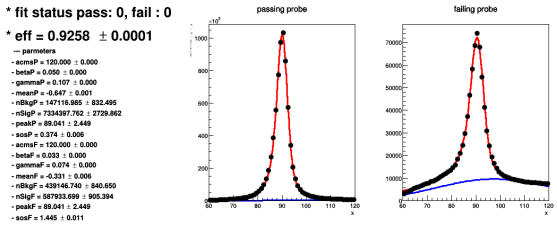


(d) high R_9category EE

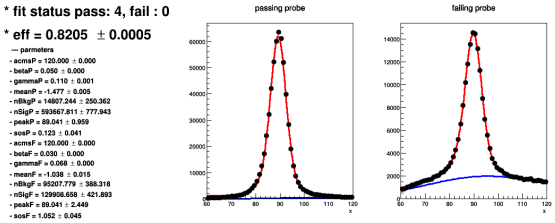
Figure 134: Data nominal fit 2018



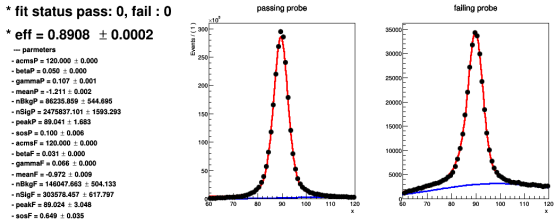
(a) low R_9category EB



(b) high R_9category EB

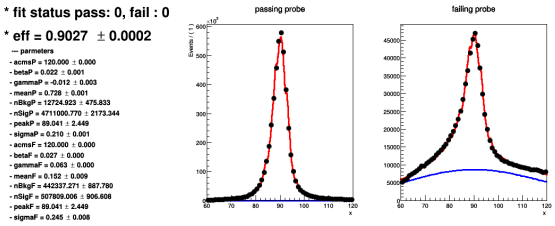


(c) low R_9category EE

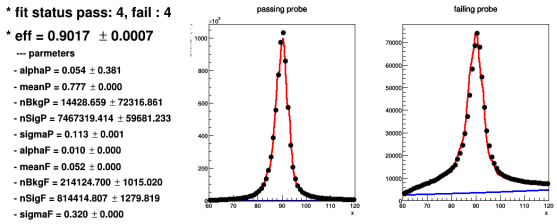


(d) high R_9category EE

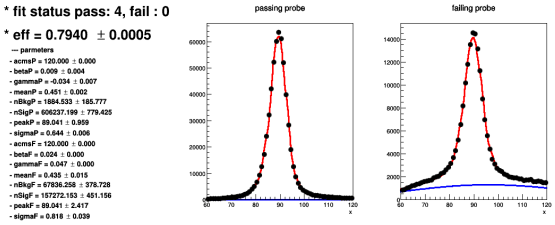
Figure 135: Data alternative background fit 2016



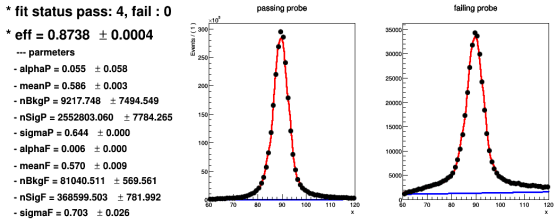
(a) low R_9category EB



(b) high R_9category EB

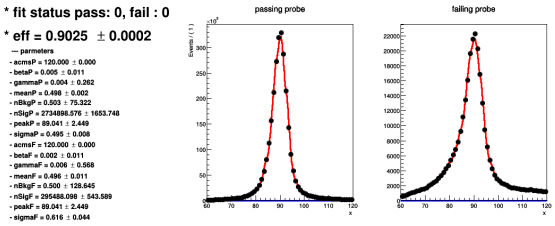


(c) low R_9category EE

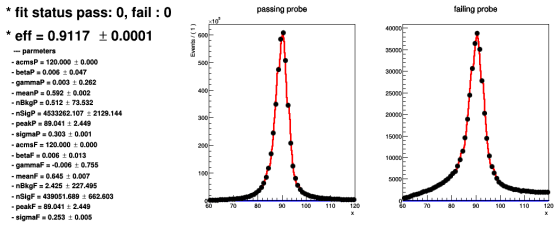


(d) high R_9category EE

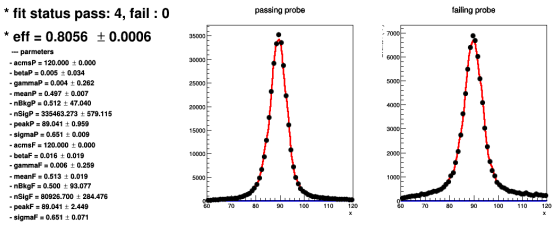
Figure 136: Data alternative signal fit 2018



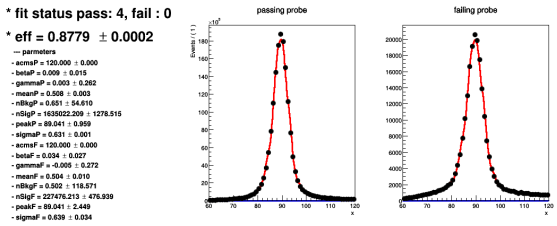
(a) low R_9category EB



(b) high R_9category EB



(c) low R_9category EE



(d) high R_9category EE

Figure 137: MC nominal fit 2018

References

[1] D. Galbrait, "UX: Standard Model of the Standard Model". Visited on March the 4th, 2023. https://davidgalbraith.org/portfolio/ux-standard-model-of-the-standard-model.

[2] J. Ellis, "Higgs Physics", doi:10.5170/CERN-2015-004.117, arXiv:1312.5672. 52 pages, 45 figures, Lectures presented at the ESHEP 2013 School of High-Energy Physics, to appear as part of the proceedings in a CERN Yellow Report.

[3] Particle Data Group Collaboration, "Review of particle physics", *Phys. Rev. D* **98** (Aug, 2018) 030001, doi:10.1103/PhysRevD.98.030001.

[4] C. Patrignani, "Review of particle physics", *Chinese Physics C* **40** (oct, 2016) 100001, doi:10.1088/1674-1137/40/10/100001.

[5] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[6] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s}$ = 7 and 8 TeV", *JHEP* **06** (2013) 081, doi:10.1007/JHEP06(2013)081, arXiv:1303.4571.

[7] ATLAS Collaboration, "Observation of a new particle in the search for the standard model Higgs boson with the detector at the LHC", *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[8] A. Sirunyan et al., "A measurement of the higgs boson mass in the diphoton decay channel", *Physics Letters B* **805** (jun, 2020) 135425, doi:10.1016/j.physletb.2020.135425.

[9] CMS Collaboration Collaboration, "Study of the mass and spin-parity of the higgs boson candidate via its decays to z boson pairs", *Phys. Rev. Lett.* **110** (Feb, 2013) 081803, doi:10.1103/PhysRevLett.110.081803.

[10] G. Aad et al., "Measurements of higgs boson production and couplings in the four-lepton channel in proton-proton collisions at center-of-mass energies of 7 and 8 TeV with the ATLAS detector", *Physical Review D* **91** (jan, 2015) doi:10.1103/physrevd.91.012006.

[11] V. Khachatryan et al., "Observation of the diphoton decay of the higgs boson and measurement of its properties", *The European Physical Journal C* **74** (oct, 2014) doi:10.1140/epjc/s10052-014-3076-z.

[12] A. Collaboration, "Atlas-conf-2020-026", 2020.

[13] C. Collaboration, "Measurements of production cross sections of the Higgs boson in the four-lepton final state in proton-proton collisions at sqrt(s) = 13 TeV",.

[14] CMS Collaboration, "Search for nonresonant Higgs boson pair production in final states with two bottom quarks and two photons in proton-proton collisions at $\sqrt{s}$ = 13 TeV", *JHEP* **03** (2021) 257, doi:10.1007/JHEP03(2021)257, arXiv:2011.12373.

[15] B. Di Micco et al., "Higgs boson pair production at colliders: status and perspectives. Higgs boson potential at colliders: Status and perspectives", *Rev. Phys.* **5** (2020) 100045, doi:10.1016/j.revip.2020.100045, arXiv:1910.00012.

[16] D. de Florian et al., "Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector", CERN Report CERN-2017-002-M, 2016. doi:10.23731/CYRM-2017-002, arXiv:1610.07922.

[17] F. Goertz, A. Papaefstathiou, L. L. Yang, and J. Zurita, "Higgs boson pair production in the D=6 extension of the SM", *JHEP* **04** (2015) 167, doi:10.1007/JHEP04(2015)167, arXiv:1410.3471.

[18] A. Carvalho et al., "Higgs pair production: Choosing benchmarks with cluster analysis", *JHEP* **04** (2016) 126, doi:10.1007/JHEP04(2016)126, arXiv:1507.02245.

[19] A. M. Sirunyan et al., "Search for nonresonant higgs boson pair production in the $$ \mathrm{b}\overline{\mathrm{b}}\mathrm{b}\overline{\mathrm{b}} $$ final state at $$ \sqrt{s} $$ = 13 TeV", *Journal of High Energy Physics* **2019** (apr, 2019) doi:10.1007/jhep04(2019)112.

[20] C. Collaboration, "Search for nonresonant higgs boson pair production in final state with two bottom quarks and two tau leptons in proton-proton collisions at $\sqrt{s}$ = 13 tev", 2022.

[21] A. Sirunyan et al., "Search for resonant pair production of higgs bosons in the bbzz channel in proton-proton collisions at 13 tev", *Physical Review D* **102** (aug, 2020) doi:10.1103/physrevd.102.032003.

[22] A. M. Sirunyan et al., "Search for nonresonant higgs boson pair production in final states with two bottom quarks and two photons in proton-proton

collisions at $$ \sqrt{s} $$ = 13 TeV", *Journal of High Energy Physics* **2021** (mar, 2021) doi:10.1007/jhep03(2021)257.

[23] R. Garoby, "Plans for upgrading the cern proton accelerator complex", *Journal of Physics: Conference Series* **110** (may, 2008) 112003, doi:10.1088/1742-6596/110/11/112003.

[24] T. Sakuma and T. McCauley, "Detector and event visualization with sketchup at the cms experiment", *Journal of Physics: Conference Series* **513** (jun, 2014) 022032, doi:10.1088/1742-6596/513/2/022032.

[25] T. C. Collaboration et al., "The cms experiment at the cern lhc", *Journal of Instrumentation* **3** (aug, 2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[26] CMS Collaboration, "The CMS magnet project: Technical Design Report". Technical design report. CMS. CERN, Geneva, 1997. doi:10.17181/CERN.6ZU0.V4T9.

[27] CMS CollaborationV. Karimaki, et al., "The CMS tracker system project: Technical Design Report". Technical design report. CMS. CERN, Geneva, 1997.

[28] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014) P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569. Comments: Replaced with published version. Added journal reference and DOI.

[29] CMSTrackerGroupoftheCMS Collaboration, "The CMS Phase-1 Pixel Detector Upgrade", *JINST* **16** (2021), no. 02, P02027, doi:10.1088/1748-0221/16/02/P02027, arXiv:2012.14304.

[30] CMS Collaboration, "Precision measurement of the structure of the CMS inner tracking system using nuclear interactions with data collected in 2018",.

[31] CMS Collaboration, "The Phase-2 Upgrade of the CMS Tracker", technical report, CERN, Geneva, 2017. doi:10.17181/CERN.QZ28.FLHW.

[32] CMS Collaboration, "The CMS electromagnetic calorimeter project: Technical Design Report". Technical design report. CMS. CERN, Geneva, 1997.

[33] CMS Collaboration, "The CMS hadron calorimeter project: Technical Design Report". Technical design report. CMS. CERN, Geneva, 1997.

[34] CMS Collaboration, J. G. Layter, "The CMS muon project: Technical Design Report". Technical design report. CMS. CERN, Geneva, 1997.

[35] A. Sirunyan et al., "Particle-flow reconstruction and global event description with the CMS detector", *Journal of Instrumentation* **12** (oct, 2017) P10003–P10003, doi:10.1088/1748-0221/12/10/p10003.

[36] T. C. Collaboration, S. Chatrchyan, G. Hmayakyan, and V. Khachatryan, "The cms experiment at the cern lhc", *Journal of Instrumentation* **3** (aug, 2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[37] "Energy calibration and resolution of the cms electromagnetic calorimeter in pp collisions at s = 7 tev", *Journal of Instrumentation* **8** (sep, 2013) P09009, doi:10.1088/1748-0221/8/09/P09009.

[38] A. Sirunyan and S. Trembath-Reichert, "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC", *Journal of Instrumentation* **16** (may, 2021) P05014, doi:10.1088/1748-0221/16/05/p05014.

[39] A. Sirunyan and F. Golf, "Performance of reconstruction and identification of τ leptons decaying to hadrons and iv/isubτ/sub in pp collisions at $\sqrt{}$is/i=13 TeV", *Journal of Instrumentation* **13** (oct, 2018) P10005–P10005, doi:10.1088/1748-0221/13/10/p10005.

[40] "Energy calibration and resolution of the cms electromagnetic calorimeter in pp collisions at s = 7 tev", *Journal of Instrumentation* **8** (sep, 2013) P09009, doi:10.1088/1748-0221/8/09/P09009.

[41] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_{T} jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.

[42] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[43] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

[44] CMS Collaboration, "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13\,\mathrm{TeV}$ using the CMS detector", *JINST* **14** (2019) P07004, doi:10.1088/1748-0221/14/07/P07004, arXiv:1903.06078.

[45] V. Khachatryan et al., "The CMS trigger system", *Journal of Instrumentation* **12** (jan, 2017) P01020–P01020, doi:10.1088/1748-0221/12/01/p01020.

[46] V. Khachatryan and N. Woods, "The cms trigger system", *Journal of Instrumentation* **12** (jan, 2017) P01020, doi:10.1088/1748-0221/12/01/P01020.

[47] M. Jeitler, "Upgrade of the trigger system of cms", *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **718** (2013) 11–15, doi:https://doi.org/10.1016/j.nima.2012.08.091. Proceedings of the 12th Pisa Meeting on Advanced Detectors.

[48] CMScollaboration Collaboration, "CMS Technical Design Report for the Level-1 Trigger Upgrade", technical report, 2013.

[49] G. Buchalla et al., "Higgs boson pair production in non-linear Effective Field Theory with full m_t-dependence at NLO QCD", *JHEP* **09** (2018) 057, doi:10.1007/JHEP09(2018)057, arXiv:1806.05162.

[50] E. Bagnaschi, G. Degrassi, P. Slavich, and A. Vicini, "Higgs production via gluon fusion in the POWHEG approach in the SM and in the MSSM", *JHEP* **02** (2012) 088, doi:10.1007/JHEP02(2012)088, arXiv:1111.2854.

[51] G. Heinrich et al., "NLO predictions for Higgs boson pair production with full top quark mass dependence matched to parton showers", *JHEP* **08** (2017) 088, doi:10.1007/JHEP08(2017)088, arXiv:1703.09252.

[52] G. Heinrich et al., "Probing the trilinear Higgs boson coupling in di-Higgs production at NLO QCD including parton shower effects", *JHEP* **06** (2019) 066, doi:10.1007/JHEP06(2019)066, arXiv:1903.08137.

[53] S. P. Jones and S. Kuttimalai, "Parton shower and NLO-matching uncertainties in Higgs boson pair production", *JHEP* **02** (2018) 176, doi:10.1007/JHEP02(2018)176, arXiv:1711.03319.

[54] G. Heinrich, S. P. Jones, M. Kerner, and L. Scyboz, "A non-linear EFT description of $gg \to HH$ at NLO interfaced to POWHEG", *JHEP* **10** (2020) 021, doi:10.1007/JHEP10(2020)021, arXiv:2006.16877.

[55] T. Gleisberg et al., "Event generation with SHERPA 1.1", *JHEP* **02** (2009) 007, doi:10.1088/1126-6708/2009/02/007, arXiv:0811.4622.

[56] T. Sjöstrand et al., "An introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.

[57] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J. C* **76** (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.

[58] CMS Collaboration, "Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements", *Eur. Phys. J. C* **80** (2020) 4, doi:10.1140/epjc/s10052-019-7499-4, arXiv:1903.12179.

[59] GEANT4 Collaboration, "GEANT4 — a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[60] CMS Collaboration, "Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **11** (2018) 185, doi:10.1007/JHEP11(2018)185, arXiv:1804.02716.

[61] CMS Collaboration, "Performance of the DeepJet b tagging algorithm using 41.9 fb^{-1} of data from proton-proton collisions at 13 TeV with Phase-1 CMS detector", CMS Detector Performance Note CMS-DP-2018-058, 2018.

[62] E. K. F. L. Bianchini, J. Conway and C. Veelken, "Reconstruction of the higgs mass in h to tautau events by dynamical likelihood techniques", *Journal of Physics: Conference Series* **513** (2014) doi:10.1088/1742-6596/513/2/022035.

[63] G. Peter Lepage, "A new algorithm for adaptive multidimensional integration", *Journal of Computational Physics* **27** (1978), no. 2, 192–203, doi:https://doi.org/10.1016/0021-9991(78)90004-9.

[64] B. Bullock, K. Hagiwara, and A. D. Martin, "Tau polarization and its correlations as a probe of new physics", *Nuclear Physics B* **395** (1993), no. 3, 499–533.

[65] S. Raychaudhuri and D. Roy, "Sharpening up the charged higgs boson signature using τ polarization at the cern lhc", *Physical Review D* **53** (1996), no. 9, 4902.

[66] M. V. Chizhov, V. A. Bednyakov, and J. A. Budagov, "On resonance search in dijet events at the lhc", 2011.

[67] A. Azatov, R. Contino, G. Panico, and M. Son, "Effective field theory analysis of double higgs boson production via gluon fusion", *Phys. Rev. D* **92** (Aug, 2015) 035001, doi:10.1103/PhysRevD.92.035001.

[68] M. R. Camana Acosta, S. Ahmed, C. E. Garcia, and I. Koo, "Extremely randomized trees-based scheme for stealthy cyber-attack detection in smart grid networks", *IEEE Access* **8** (2020) 19921–19933, doi:10.1109/ACCESS.2020.2968934.

[69] J. C. Collins and D. E. Soper, "Angular distribution of dileptons in high-energy hadron collisions", *Phys. Rev. D* **16** (Oct, 1977) 2219, doi:10.1103/PhysRevD.16.2219.

[70] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System", in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Association for Computing Machinery, 2016. doi:10.1145/2939672.2939785.

[71] CMS Collaboration, "Measurements of Higgs boson production cross sections and couplings in the diphoton decay channel at $\sqrt{s}$ = 13 TeV", *JHEP* **07** (2021) 027, doi:10.1007/JHEP07(2021)027, arXiv:2103.06956.

[72] P. D. Dauncey, M. Kenzie, N. Wardle, and G. J. Davies, "Handling uncertainties in background shapes", *JINST* **10** (2015) P04015, doi:10.1088/1748-0221/10/04/P04015, arXiv:1408.6865.

[73] R. Fisher, "On the interpretation of χ^2 from contingency tables, and the calculation of p", *Journal of the Royal Statistical Society* **85** (1922), no. 1, 87–94.

[74] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.

[75] A. L. Read, "Presentation of search results: the CL_s technique", *J. Phys. G* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.

[76] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727. [Erratum: doi:10.1140/epjc/s10052-013-2501-z].

[77] M. Grazzini et al., "Higgs boson pair production at NNLO with top quark mass effects", *JHEP* **05** (2018) 059, doi:10.1007/JHEP05(2018)059, arXiv:1803.02463.

[78] S. Dawson, S. Dittmaier, and M. Spira, "Neutral Higgs boson pair production at hadron colliders: QCD corrections", *Phys. Rev. D* **58** (1998) 115012, doi:10.1103/PhysRevD.58.115012, arXiv:hep-ph/9805244.

[79] S. Borowka et al., "Higgs boson pair production in gluon fusion at next-to-leading order with full top-quark mass dependence", *Phys. Rev. Lett.* **117** (2016) 012001, doi:10.1103/PhysRevLett.117.079901, arXiv:1604.06447.

[80] J. Baglio et al., "Gluon fusion into Higgs pairs at NLO QCD and the top mass scheme", *Eur. Phys. J. C* **79** (2019) 459, doi:10.1140/epjc/s10052-019-6973-3, arXiv:1811.05692.

[81] D. de Florian and J. Mazzitelli, "Higgs boson pair production at next-to-next-to-leading order in QCD", *Phys. Rev. Lett.* **111** (2013) 201801, doi:10.1103/PhysRevLett.111.201801, arXiv:1309.6594.

[82] D. Y. Shao, C. S. Li, H. T. Li, and J. Wang, "Threshold resummation effects in Higgs boson pair production at the LHC", *JHEP* **07** (2013) 169, doi:10.1007/JHEP07(2013)169, arXiv:1301.1245.

[83] D. de Florian and J. Mazzitelli, "Higgs pair production at next-to-next-to-leading logarithmic accuracy at the LHC", *JHEP* **09** (2015) 053, doi:10.1007/JHEP09(2015)053, arXiv:1505.07122.

[84] D. de Florian et al., "Handbook of lhc higgs cross sections: 4. deciphering the nature of the higgs sector", 2016.

[85] J. Butterworth et al., "PDF4LHC recommendations for LHC Run II", *J. Phys. G* **43** (2016) 023001, doi:10.1088/0954-3899/43/2/023001, arXiv:1510.03865.

[86] S. H. et al., "Handbook of lhc higgs cross sections: 3. higgs properties", doi:10.5170/CERN-2013-004, arXiv:arXiv:1307.1347.

[87] T. Adams et al., "Beam test evaluation of electromagnetic calorimeter modules made from proton-damaged $PbWO_4$ crystals", *JINST* **11** (2016), no. 04, P04012, doi:10.1088/1748-0221/11/04/P04012.

[88] F. Gentit, "Litrani: A General purpose Monte Carlo program simulating light propagation in isotropic or anisotropic media", *Nucl. Instrum. Meth. A* **486** (2002) 35–39, doi:10.1016/S0168-9002(02)00671-X.

[89] A. Sirunyan et al., "Identification of heavy-flavour jets with the cms detector in pp collisions at 13 tev", *Journal of Instrumentation* **13** (May, 2018) P05011P05011, doi:10.1088/1748-0221/13/05/p05011.