UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS FÍSICAS

TESIS DOCTORAL

Search for Higgs boson pair production at the CMS experiment of the LHC and studies of a new trigger algorithm for the muon chambers in HL-LHC

Búsqueda de producción de pares de bosones de Higgs en el experimento CMS del LHC y estudios de un nuevo sistema de disparo para las cámaras de muones durante HL-LHC

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Jaime León Holgado

DIRECTORA

Silvia Goy López

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS FÍSICAS

TESIS DOCTORAL

Search for Higgs boson pair production at the CMS experiment of the LHC and studies of a new trigger algorithm for the muon chambers in HL-LHC

Búsqueda de producción de pares de bosones de Higgs en el experimento CMS del LHC y estudios de un nuevo sistema de disparo para las cámaras de muones durante HL-LHC

Memoria para optar al grado de doctor

presentada por

Jaime León Holgado

Directora

Dra. Silvia Goy López

Contents

Abstract

This thesis reports a search for double Higgs boson production at the Large Hadron Collider (LHC). This process is predicted by the Standard Model of particle physics (SM), with a production cross section around 1000 times smaller than single Higgs boson production. Its study can provide information about the Higgs potential, the Higgs self-coupling, and its couplings to other SM particles such as the t quark or the vector bosons. Additionally, Beyond the Standard Model (BSM) effects can manifest as modifications on the production cross section or changes in event kinematics. Two production modes are studied, gluon fusion (ggF) and vector boson fusion (VBF). This search is performed using data from the Compact Muon Solenoid (CMS) experiment during the so-called Run 2 of the LHC, corresponding to proton-proton collision data collected in 2016, 2017, and 2018, and amounting to a total integrated luminosity of around 138 fb^{-1}. The analysis considers the final state where a Higgs boson decays into two b quarks and the other decays into two τ leptons. Since it has an intermediate branching ratio and relatively low background, it is considered as one of the golden channels in the search for double Higgs boson production.

The theoretical context of the double Higgs production is first presented, followed by a description of the experimental setup used in the analysis, including the techniques used within the CMS experiment for particle reconstruction and identification. The analysis strategy is then presented, starting by identifying the different background and signal samples to be considered and followed by a description of the selection applied to the different objects in the final state in order to reduce the background contamination. A further discrimination is then performed by splitting the data and simulated events into categories enriched in either ggF or VBF events. For the VBF categorisation a machine learning approach is considered, consisting in a multi-class deep neural network (DNN) that classifies events as signal or as one of the main background processes. Signal extraction is finally performed with another DNN. The main physics results of this thesis are: upper limits on the double Higgs boson production cross section via ggF and VBF, the production cross section for VBF production only (fixing ggF to its SM prediction), and exclusion ranges on the Higgs self-coupling and its couplings to vector bosons. The results obtained constitute an important improvement with respect to the previous $\text{HH} \rightarrow \text{bb}\tau\tau$ analysis, providing one of the best sensitivities among all the decay channels studied in the CMS Collaboration. For the new data taking period (Run 3, 2022-2025), a new trigger strategy has been studied, targeting the decay channel where both τ decay into hadrons and a jet is present in the final state. This new trigger strategy has been included in the trigger menus before the data-taking period starting in 2022, so its performance can be studied using real and simulated collision data.

In order to prepare for the next stage of LHC operations after 2029, the so-called High-Luminosity LHC (HL-LHC) era, the increase of luminosity, resulting in a larger occupancy and trigger rate, will require a CMS detector upgrade. A new trigger algorithm has been developed to perform the trigger primitive generation in the muon drift tube chambers during HL-LHC. This algorithm profits from the full time resolution provided by the new electronics at trigger level, computing the

trigger primitive's time, position, and direction with a resolution comparable to what is reachable with the present offline reconstruction system. The algorithm's performance is estimated using a software emulator with simulated and real data samples, and through hardware implementation tests.

Resumen

Esta tesis presenta una búsqueda de producción de pares de bosones de Higgs en el *Large Hadron Collider* (LHC). Este proceso está predicho por el Modelo Estándar de física de partículas (ME), con una sección eficaz de producción alrededor de 1000 veces más pequeña que la producción de un único bosón de Higgs. El estudio de este proceso aporta información acerca del potencial del Higgs, el autoacoplo del bosón de Higgs y su acoplo a otras partículas del ME, como el quark t or los bosones vectoriales. Adicionalmente, efectos de Física Más Allá del Modelo Estándar se pueden manifestar como modificaciones en la sección eficaz de producción o en la cinemática de los objetos del evento. Dos modos de producción son estudiados: fusión de gluones (*gluon fusion*, ggF) y de bosones vectoriales (*vector boson fusion*, VBF). Esta búsqueda se ha realizado con datos del experimento *Compact Muon Solenoid* (CMS) durante el *Run* 2 del LHC, correspondiente a los años 2016, 2017 y 2018, correspondiente a una luminosidad integrada de 138 fb^{-1}. Este análisis considera el estado final en el que un bosón de Higgs decae en dos quarks b y el otro en dos leptones τ. Dado que es este canal tiene una fracción de desintegración intermedia y un fondo relativamente bajo, es uno de los canales que provee mejores resultados dentro de la producción de pares de bosones de Higgs.

En primer lugar se presenta el contexto teórico de la producción de bosones de Higgs, seguida de una descripción de la configuración experimental usada en el análisis, incluyendo las técnicas usadas en el experimento CMS para la reconstrucción e identificación de partículas. A continuación se describe la estrategia seguida en el análisis, comenzando por una descripción de la selección aplicada a los diferentes objectos del estado final de cara a reducir la contaminación del fondo. Una mayor discriminación se produce separando los eventos de datos y simulación en categorías enriquecidas en eventos producidos por ggF o VBF. Para la categorización en eventos de VBF se consideran técnicas de aprendizaje automático, concretamente una red neuronal multiclase que clasifica los eventos como procedentes de señal o de alguno de los fondos. La extracción de la señal se realiza con otra red neuronal, la cual provee los principales resultados físicos de esta tesis: límites superiores en la sección eficaz de producción de pares de bosones de Higgs mediante ggF o VBF y en la sección eficaz de producción solo mediante VBF (asumiendo que ggF se comporta como su preddición en el ME), y rangos de exclusión para el autoacoplo del bosón de Higgs y su acoplo a los bosones vectoriales. Los resultados obtenidos constituyen una importante mejora con respecto al anterior análisis $\text{HH} \to \text{bb}\tau\tau$, obteniendo una de las mejores sensibilidades entre los canales de desintegración estudiados dentro de la colaboración CMS. De cara al próximo periodo de toma de datos (*Run* 3, 2022-2025), se ha estudiado una nueva estrategia de selección *online*, apuntando al canal en el que ambos τ se desintegran en hadrones y adicionalmente un *jet* se encuentra en el estado final. Esta nueva estrategia ha sido incluida en los menús de *trigger* antes del periodo de toma de datos que comenzó en 2022, de forma que su rendimiento puede ser estudiado usando datos de colisiones reales y simuladas.

De cara a preparar el siguiente periodo de operaciones del LHC que comienza en 2029, el llamado *High-Luminosity* LHC (HL-LHC), el experimento CMS está llevando a cabo una actualización del detector para que pueda trabajar con el aumento

de luminosidad previsto, el cual llevará asociado un aumento de ocupancia y *rate* de datos. Se ha desarrollado un nuevo algoritmo para el sistema de disparo para realizar la generación de *trigger primitives* en las cámaras de tubos de deriva (DT) durante HL-LHC. Este algoritmo se beneficia de la resolución temporal completa que proveerá la nueva electrónica directamente a nivel del sistema de disparo, calculando el tiempo, posición y dirección de la *trigger primitive* con una resolución comparable a la que obtiene actualmente el sistema de reconstrucción *offline*. El rendimiento de este algoritmo ha sido estimado con un emulador en *software* en muestras de datos reales y simulados y con tests para la implementación en *hardware*.

Introduction

The Standard Model of particle physics is a theory that describes the elementary particles and their interactions. The last piece of this model was included in 1964 with the introduction of the electroweak symmetry breaking and the Higgs boson. Since then, the Standard Model has provided many successful prediction experimentally verified thanks to, for instance, collider experiments such as the LHC. In fact, the ATLAS and CMS experiment using proton-proton collision data from the LHC were able to discover the Higgs boson in 2012 [1, 2], which strongly supports the validity of the Standard Model.

Nonetheless, some phenomena are not described by the Standard Model, such as the asymmetry between matter and antimatter, the existence of dark matter, gravity, or even the mass scale of the Higgs boson itself. Therefore, numerous theories that extend the Standard Model at higher energies have been proposed, so they can explain the so-called Beyond the Standard Model physics.

The production of Higgs boson pairs is a process that can help to further understand the Standard Model and test the validity of Beyond the Standard Model theories. It provides unique information about the Higgs self-coupling and its coupling to other particles. Deviations on their Standard Model predictions can be related to Beyond the Standard Model physics processes.

The observation of Higgs boson pair production (as predicted by the Standard Model) and the determination of the Higgs self-coupling are not possible with the data collected as of now by the LHC experiments. The study of some HH decay channels, however, can provide very stringent upper limits on its production cross section and exclusion ranges for the self-coupling already with LHC Run 2 data. Among these decay channels, the one were one Higgs boson decays into two b quarks and the other into two τ leptons gives some of the most significant results, thanks to its medium branching fraction and relatively low background.

Looking into the future, the High-Luminosity LHC era will continue the current LHC operations, providing much more instantaneous luminosity and collected data. This amount of new data will open the door to the study of rarer processes, or even provide more precise results for the already studied ones (such as the Higgs boson pair production). This accelerator upgrades, however, need to be addressed by the experiments, so they can cope with the increased data rate and radiation. Within the CMS experiment, various subdetectors will be replaced, upgraded or substituted, so the final trigger and reconstruction performance will be maintained or even improved during HL-LHC operation.

This thesis is structured as follows. Chapter 1 introduces the theoretical background needed in order to study the Higgs boson pair production at the LHC. Chapter 2 describes the experimental setup used in the analysis, including the LHC and the CMS experiment. Chapter 3 presents the particle reconstruction and identification methods used in CMS analyses. The HH $\rightarrow$ bb$\tau\tau$ analysis is discussed in Chapter 4, and its results are shown in Chapter 5. Chapter 6 discusses a new trigger strategy that could boost the sensitivity in the HH $\rightarrow$ bb$\tau\tau$ and H $\rightarrow \tau\tau$ analyses during Run 3. Finally, Chapter 7 describes a new algorithm implemented for the CMS Drift Tube chambers for HL-LHC and shows its performance results over simulated and collision data samples.

Chapter 1

Theoretical context

The Standard Model (SM) of particle physics is a theory that describes the elementary particles and the interactions between them. It provides a set of predictions very well experimentally tested. A brief description of it is given in Section 1.1. Section 1.2 introduces a description of the phenomenology of the double Higgs production. Finally, Section 1.3 describes the physics of a proton-proton collision and its modelling.

1.1 The Standard Model

The SM is a gauge invariant and renormalizable quantum field theory (QFT) based on quantum mechanics and special relativity. Elementary particles correspond to excited states of quantum fields. The propagation and interactions of fields in a QFT are fully determined by a Lagrangian density, $\mathcal{L}$, that depends on the fields, ψ_i, their derivatives, $\partial_\mu \psi_i$ and the space-time coordinates, x^μ.

The SM describes the effect of fundamental interactions as a QFT with an exchange of force-mediating particles. This theory shall be invariant under local gauge symmetries, which, by the Noether's theorem, enforces the conservation of charges associated to the interactions. Additionally, imposing the theory to be invariant under these transformations requires introducing additional vector fields, corresponding to the gauge bosons that mediate the interactions. Two interactions are present in the SM: the strong interaction, invariant under $SU(3)_C$, and the electroweak interaction, invariant under $SU(2)_L \times U(1)_Y$.

The Lagrangian $\mathcal{L}$ provides probability amplitudes for the involved physics processes. However, it cannot be solved analytically but requires perturbation theory. The effect of these perturbations can be approximated with expansions in leading-order (LO) or higher orders (next-to-leading-order or NLO, NNLO, and so on). These perturbation terms can be used to build *matrix elements*, which give information about the strength of the process (i.e. its *cross section*) and the kinematics of the final state particles. They are often illustrated as Feynman diagrams.

The particle content of the Standard Model is shown in Fig. 1.1. Two types

of particles are included in the SM: matter particles, fermions with spin = 1/2 associated to fields represented by spinors, and bosons with spin = 1, associated to vectorial fields, that mediate the interactions between fermions. Additionally, the Higgs boson, associated to a scalar field, is included as a mechanism to give masses to both fermions and bosons.

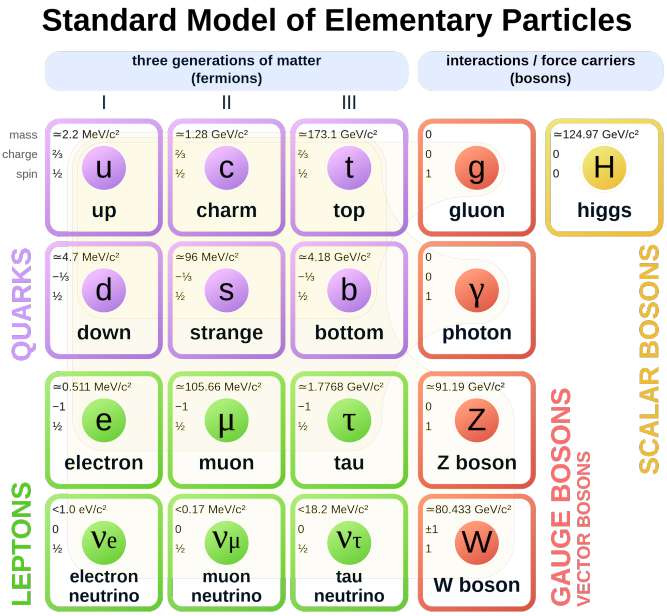


Figure 1.1: Particle content of the Standard Model of particle physics. Extracted from [3].

Even if the SM has been a very successful theory, explaining most of the experimental results with a very high level of accuracy, it does not provide a complete understanding of Nature. In particular, it is challenged by the existence of dark matter, the asymmetry in the Universe between matter and antimatter, or even the values of the fundamental parameters of the SM itself. The SM is clearly incomplete, and must be taken as an effective theory valid in a particular energy (or mass) range. The theories that extend the Standard Model in order to include explanations about these known problems are referred to as *Beyond the Standard Model* (BSM) theories. Some examples of them are theories based on supersymmetry, such as the Minimal Supersymmetric Standard Model (MSSM) and Next-to-Minimal Supersymmetric Standard Model (NMSSM), string theory, or extra dimensions.

1.1.1 Matter particles

Twelve fermions have been experimentally observed and are described in the SM, six *quarks* and six *leptons*. Both quarks and leptons are divided in three *generations*, each containing two quarks with electric charges $Q = +2/3$ and $-1/3$, and two leptons with $Q = -1$ and 0. Each fermion has its corresponding *antiparticle*,

with the same characteristics but opposite quantum numbers, including the electric charge. Quarks are the only particles in the SM affected by the three forces, the electromagnetic, the weak, and the strong forces, described in the next sections.

The first generation of quarks is composed by the *up* (u) and *down* (d) quarks, with masses of 2.2 MeV and 4.7 MeV [4], respectively. The second generation, the *charm* (c) and *strange* (s) quarks, have masses of 1.27 GeV and 93.4 MeV [4], respectively. The *top* (t) and *bottom* (b) quarks form the third generation, with masses of 172.7 GeV and 4.18 GeV. Each quark carries a *flavour* and a *colour*, which makes them subject to the electroweak and the strong forces, respectively. A property of the strong force is the *color confinement*, which implies that quarks cannot exist as free states, but are experimentally observed as bound states called *hadrons*. The top quark, however, cannot form hadrons, since its very high mass makes it very unstable, decaying right after being produced. At higher energies, the strong force gets increasingly weaker, so quarks behave like *asymptotically free* particles. This allows to study their interactions in proton colliders like the LHC.

Electrons (e), *muons* (μ) and *taus* (τ) are the three charged leptons included in the SM, with masses of 0.511 MeV, 105.66 MeV and 1.78 GeV respectively [4]. These particles are subject to the electromagnetic and the weak forces. The electron is the only stable charged lepton. Muons have a lifetime of 2.2 μs [4], nonetheless they are still detectable by collider experiments thanks to their high momentum and the experiment size. Tau leptons, however, have a lifetime of 290.3 fs, so they decay before being detected and can only be inferred from its decay products.

Each charged lepton is associated to a electrically neutral neutrino of the same flavour (ν_e, ν_μ and ν_τ). In the classical SM formulation, these neutrinos are massless. However, the observation of neutrino flavour oscillations [5] implies that neutrinos are massive, with upper limits on their masses of $m(\nu_e) < 1.1$ eV, $m(\nu_\mu) < 0.19$ MeV, and $m(\nu_\tau) < 18.2$ MeV [4]. Neutrinos only interact with matter via the weak force, so they can not be directly detected at collider experiments. Their presence is reconstructed from the energy imbalance in the event (see Section 3.2.5).

1.1.2 Interactions in the Standard Model

The strong interaction

The strong interaction is described by the Quantum Chromodynamics (QCD) theory. QCD is a non-abelian gauge theory based on the $SU(3)_C$ group. The conserved charge associated to this symmetry is the color charge. As only quarks have color charge, only quark fields have a fundamental representation under this group; the other fermions have the trivial representation. The $SU(3)_C$ group has eight generators, T^a, the Gell-Mann matrices, related to the eight *gluons*, the massless strong force mediators. The QCD Lagrangian can be written as

$$\mathcal{L}_{\text{QCD}} = \bar{\psi}\left(i\gamma^\mu\partial_\mu - g_s\gamma^\mu T_a G^a_\mu - m\right)\psi - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a, \tag{1.1}$$

where γ^μ are the Dirac matrices, ψ is the quark field, m is the quark mass, g_s is the coupling constant of the strong interaction (although $\alpha_s = \frac{g_s^2}{4\pi}$ is more commonly used) and the strength field tensor $G^a_{\mu\nu}$ is defined as

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f^{abc} G^b_\mu G^c_\nu, \tag{1.2}$$

where f^{abc} are the structure constants of the $SU(3)_C$ group. The last term of the strength field tensor is included since the $SU(3)_C$ is non-abelian. This term leads to 3- and 4-gluons self-interactions.

The electroweak interaction

The electroweak (EWK) interaction is a unified description of the electromagnetic and weak interactions. It's a theory locally gauge invariant under $SU(2)_L \times U(1)_Y$ transformations.

The weak interaction is a theory described with the non-abelian $SU(2)_L$ group. It has the peculiarity that it violates parity, i.e. left and right-handed particles couple differently. The concept of *chirality* is then introduced in the theory: the chirality is a Lorentz-invariant quantity corresponding to the eigenvalues of $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ (± 1), giving rise to the left (ψ_L) and right (ψ_R) chirality fields, obtained as

$$\psi_L = \frac{1}{2}(1 - \gamma^5)\ \psi, \tag{1.3}$$

$$\psi_R = \frac{1}{2}(1 + \gamma^5)\ \psi. \tag{1.4}$$

Every fermionic field of the SM is represented as one left chirality doublet (ψ_L) and two right chirality singlets (ψ_R, ψ'_R). For example, for the first family of quarks

$$\psi_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \psi_R = u_R, \psi'_R = d_R, \tag{1.5}$$

and, similarly, for the first family of leptons

$$\psi_L = \begin{pmatrix} \nu_{e,L} \\ e_L \end{pmatrix}, \psi_R = \nu_{e,R}, \psi'_R = e_R. \tag{1.6}$$

The conserved charge associated to the $SU(2)_L$ group is the *weak isospin* I_3. For left-handed fermions, the up-type takes a value of $I_3 = +1/2$ and the down-type $I_3 = -1/2$. Right-handed fermions, as they are singlets of $SU(2)_L$, take a value of $I_3 = 0$.

The weak Lagrangian can be written as

$$\mathcal{L}_{\text{WK}} = i\bar{\psi}_L \gamma_\mu D_\mu \psi_L + i\bar{\psi}_R \gamma^\mu \partial_\mu \psi_R + i\bar{\psi}'_R \gamma^\mu \partial_\mu \psi'_R - \frac{1}{4} W^i_{\mu\nu} W_i^{\mu\nu}, \tag{1.7}$$

where

$$D_\mu = \partial_\mu + i g_w T^i W^i_\mu \tag{1.8}$$

is the covariant derivative of the $SU(2)_L$ group with $i \in \{1, 2, 3\}$, g_w is the coupling constant of the weak force, W_μ^i are the three gauge fields, $T_\mu^i = \sigma_i/2$ ($\sigma_i \equiv$ Pauli matrices) are the generators of $SU(2)_L$ and the tensor field $W_{\mu\nu}^i$ is defined as

$$W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g_w \epsilon_{ijk} W_\mu^j W_\nu^k, \tag{1.9}$$

where ϵ_{ijk} are the structure constants of $SU(2)_L$. The last term of Eq. (1.9) adds interaction terms between three and four gauge fields. Note that a mass term $m\bar{\psi}\psi = m(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R)$ has not been added to the Lagragian, as it's not invariant under $SU(2)_L$.

The gauge fields W_μ^i do not correspond to the physical $W_\mu^\pm$ and Z_μ fields associated to the $W^\pm$ and Z bosons. The first ones can be obtained via a linear combination

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp i W_\mu^2). \tag{1.10}$$

To obtain the Z_μ boson field, an additional group $U(1)_Y$ inspired in the electromagnetic theory is introduced, with a Lagrangian

$$\mathcal{L}_Y = i\bar{\psi}\gamma^\mu D_\mu \psi - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \tag{1.11}$$

where

$$D_\mu = \partial_\mu + i g_Y Y B_\mu \tag{1.12}$$

is the covariant derivative of $U(1)_Y$, g_Y is the coupling constant and the tensor field $B_{\mu\nu}$ is defined as

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu. \tag{1.13}$$

The conserved quantity of this symmetry is the hypercharge, Y, related to the isospin and the electric charge by $Y = 2(Q - I_3)$. With this additional field, the Z boson and the photon (γ) fields Z_μ and A_μ can be obtained as

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}, \tag{1.14}$$

where θ_W is the Weinberg angle, defined as

$$\theta_W = \tan^{-1} \frac{g_W}{g_Y}. \tag{1.15}$$

Up to this point, the $W^\pm$, Z and γ bosons and the fermions have been assumed to be massless, since typical mass terms would break the gauge invariance. However, experimentally all those particles have been measured as massive particles (with the exception of the photon). The mechanism through which fermions and bosons acquire mass comes from the spontaneous symmetry breaking of the electroweak symmetry, described below.

1.1.3 The electroweak symmetry breaking

Since the EWK bosons are not massless, additional terms need to be introduced in the Lagragian by breaking the electroweak symmetry without breaking the gauge invariance. These mass terms arise from the so-called electroweak symmetry breaking (EWSB).

In the SM, the EWSB is achieved through the *Higgs mechanism*, proposed in 1964 independently by P. Higgs [6] and F. Englert and R. Brout [7], among others. This mechanism explains the masses of gauge bosons and fermions by proposing that the $SU(2)_L \times U(1)_Y$ symmetry from the EWK theory is broken at low energies and becomes $U(1)_{EM}$. This symmetry breaking is based on the introduction of two new complex scalar fields grouped in a weak isospin doublet ϕ known as the Higgs field:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \tag{1.16}$$

where ϕ_1, ϕ_2, ϕ_3 and ϕ_4 are real scalar fields. The Lagrangian associated to this field can be written as

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger (D_\mu \phi) - V(\phi), \tag{1.17}$$

where D_μ is the covariant derivative from the EWK theory

$$D_\mu = \partial_\mu + i g_w T^i W^i_\mu + i g_Y Y B_\mu \tag{1.18}$$

and $V(\phi)$ is a potential that has the form

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2, \tag{1.19}$$

where μ^2 and λ are constants. In the case where $\mu^2 > 0$ and $\lambda > 0$, the potential will have a parabolic shape with a minimum in $\phi = 0$, i.e. the corresponding to a scalar particle with mass μ and self-interaction with a coupling λ. However, if $\mu^2 < 0$, the shape of the potential is illustrated in Fig. 1.2, the so-called *mexican hat* shape. In this case, $\phi = 0$ becomes a local maximum, and the minima are found in the doublets that satisfy

$$\phi^\dagger \phi = \frac{-\mu^2}{2\lambda} \equiv \frac{v^2}{2}, \tag{1.20}$$

where v is the *vacuum expectation value*. Among all possible states, the ground state can be expressed, without loss of generality, as

$$\phi_{\text{ground}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{1.21}$$

By exciting around the ground state, $\phi(x)$ becomes

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \tag{1.22}$$

where $H(x)$ is the only field that remains, the one associated to a new massive particle, the Higgs boson (H).

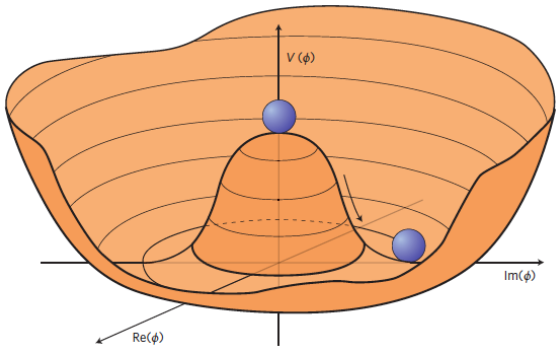


Figure 1.2: Illustration of the Higgs potential in the case that $\mu^2 < 0$. Extracted from [8].

By including Eq. 1.22 and the transformations shown in Eqs. 1.10 and 1.14 into the Lagrangian from Eq. 1.17, the Lagrangian can be rewritten as

$$\mathcal{L}_{\text{Higgs}} = \frac{1}{2}\partial_\mu H \partial^\mu H + \frac{g_w^2 v^2}{4} W_\mu^+ W^{-\mu} + \frac{1}{2}\frac{g_w^2 + g_Y^2}{4} v^2 Z_\mu Z^\mu + \frac{1}{2} 2\mu^2 H^2 \tag{1.23}$$

$$+ \frac{g_w^2}{2} v H W_\mu^+ W^{-\mu} + \frac{g_Y^2}{2} v H Z_\mu Z^\mu \tag{1.24}$$

$$+ \frac{g_w^4}{2} H^2 W_\mu^+ W^{-\mu} + \frac{g_Y^2}{4} H^2 Z_\mu Z^\mu \tag{1.25}$$

$$+ \frac{\mu^2}{v} H^3 + \frac{\mu^2}{4\nu^2} H^4. \tag{1.26}$$

The second, third, and fourth terms of Eq. (1.23) correspond to mass terms for the $W^\pm$, Z, and H bosons, so these bosons acquire masses given by

$$M_{\text{Z}} = \frac{\sqrt{g_w^2 + g_Y^2}}{2} v, \qquad M_{\text{W}^\pm} = \frac{g_w v}{2} = M_Z \cos\theta_W, \qquad M_{\text{H}} = \sqrt{2}\mu. \tag{1.27}$$

No mass term appears for the A_μ field, so the photon remains massless.

The rest of the terms in the Lagrangian are interaction terms. Eqs. (1.24) and (1.25) include trilinear HZZ and HW^+W^- and quadrilinear HHZZ and HHW^+W^- couplings. Eq. (1.26) includes the trilinear and quadrilinear Higgs boson self-interactions HHH and HHHH. With these terms, the Higgs *self-couplings* can be defined as

$$\lambda_{\text{HHH}} = \lambda_{\text{HHHH}} = \frac{M_H^2}{v^2} \equiv \lambda. \tag{1.28}$$

The presence of these interaction terms shows that the double (and triple) Higgs production is predicted by the SM at tree level. Studying this production is an important test of the electroweak symmetry breaking.

Finally, the mass of the fermions also raises from their interactions with the Higgs field. For a given fermionic field, using its chirality components (obtained

via Eqs. (1.3) and (1.4)), one can write its interaction with the Higgs field with a Yukawa Lagrangian:

$$\mathcal{L}_{\text{Yukawa}} = -y_F(\bar{\psi}_R\phi^\dagger\psi_L + \bar{\psi}_L\phi\psi_R), \tag{1.29}$$

where y_F is the coupling constant between the fermion and the Higgs boson. As an example, for the electron field shown in Eq. (1.6), by introducing ϕ parametrized as in Eq. (1.22), one obtains:

$$\mathcal{L}_{\text{Yukawa},e} = -y_e\frac{v+H}{\sqrt{2}}(\bar{e}_R e_L + \bar{e}_L e_R). \tag{1.30}$$

By re-grouping the electron fields as $\bar{e} = (\bar{e}_R, \bar{e}_L)$ and $e = (e_R, e_L)$, Eq. (1.30) becomes

$$\mathcal{L}_{\text{Yukawa},e} = -y_e\frac{v}{\sqrt{2}}\bar{e}e - y_e\frac{H}{\sqrt{2}}\bar{e}e. \tag{1.31}$$

The first term in Eq. (1.31) corresponds to the mass term of the electron field, so the mass of the electron can be obtained as $m_e = \frac{y_e v}{\sqrt{2}}$. The second term represents the interaction between the Higgs boson and the electron.

In conclusion, the EWSB allows both fermions and bosons to acquire mass via the Higgs field. All masses and interactions between the Higgs boson and the particles are related to v and the gauge couplings. In fact, the interactions between the Higgs boson and the particles are proportional to masses of the latter, so the Higgs boson will decay preferentially to the heaviest particles kinematically accesible.

1.1.4 Phenomenology of the Higgs boson

On July 4 2012, the CMS and ATLAS experiments at the LHC at CERN announced the observation of a new boson compatible with the SM prediction of the Higgs boson [1, 2]. The discovery was performed independently by the two experiments using data from the first two years of the LHC Run 1 (2011-2012, see Section 2.1).

Regarding the mass of the Higgs boson, the most recent and precise value was obtained by combining the results from the CMS experiment during 2011, 2012 and 2016 at centre-of-mass energies of $\sqrt{s}$ = 7, 8 and 13 TeV, respectively [9]. The resulting value was

$$m_{\text{H}} = 125.38 \pm 0.11\ (\text{stat}) \pm 0.08\ (\text{syst})\ \text{GeV}. \tag{1.32}$$

The SM predicts that the Higgs boson is a electrically neutral scalar particle (spin = 0) and even under charge conjugation parity (CP) transformations. Both properties have been determined. Due to the Landau-Yang theorem, the observation of the Higgs boson decaying into $\gamma\gamma$ forbids it to have a spin = 1 [10,11]. The spin-0 vs spin-2 and the CP hypotheses have been studied using the Higgs boson decay into two Z bosons [12], confirming the SM prediction.

Once the Higgs boson mass has been measured, its production modes and their cross sections can also be studied. Fig. 1.3 shows the Higgs boson main production

modes cross sections for a Higgs boson mass of 125 GeV as a function of the collision centre-of-mass energy. The dominant production mode is gluon fusion (ggF), where the Higgs boson is produced via loops of heavy quarks (mostly t quarks); its cross section is around 49 pb at $\sqrt{s} = 13$ TeV. The production mode with the second largest contribution is vector boson fusion (VBF), where the Higgs boson is produced in association with two high-momentum quarks very close to the direction of the incident particles. Its cross section is about 10 times smaller than ggF, 3.8 fb. Follows the associated production with a vector boson, with a cross section of 2.3 pb, and the associated production with heavy quarks (top pair, bottom pair or single top), with cross sections around one order of magnitude smaller than the previous. The Feynman diagrams of these processes are shown in Fig 1.4.

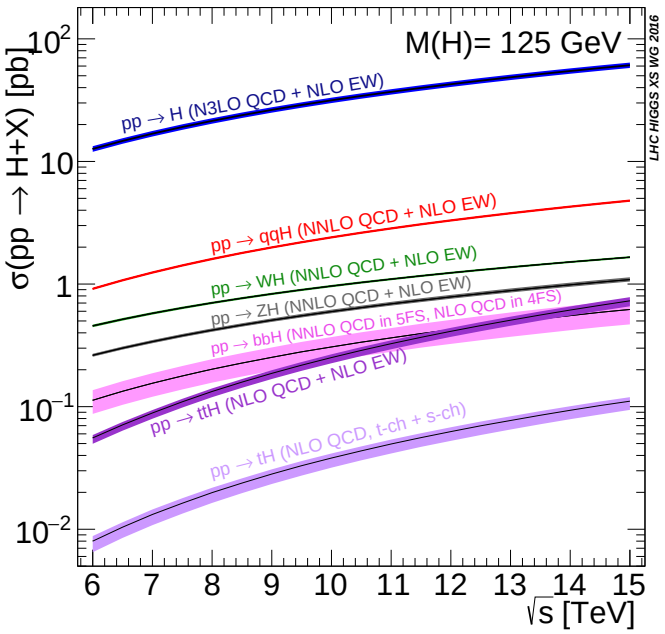


Figure 1.3: The SM Higgs boson production cross sections as a function of the LHC centre of mass energy for a Higgs boson mass of 125 GeV. The main production modes at the LHC are shown: gluon fusion (blue), vector boson fusion (red), vector boson associated production (green and black), top quark pair associated production (dark purple), bottom quark pair associated production (pink) and single top associated production (light purple). Extracted from [13].

The *branching fractions* or *branching ratios* (BR) of the Higgs boson (i.e. the probability of the Higgs boson decaying into different final states) for a mass of 125 GeV are shown in Table 1.1. The decay with the largest BR is the H → bb, with a BR of 58%. However, this channel is not the most appropriate to perform experimental studies at the LHC, as other processes with similar final states are very commonly produced from pp collisions. Instead, the most convenient final states are the H → $\gamma\gamma$ and H → ZZ → 4l ($l = e,\ \mu$), as the detector resolutions on the objects from these final states are very high.

After having experimentally established the Higgs boson production and its de-

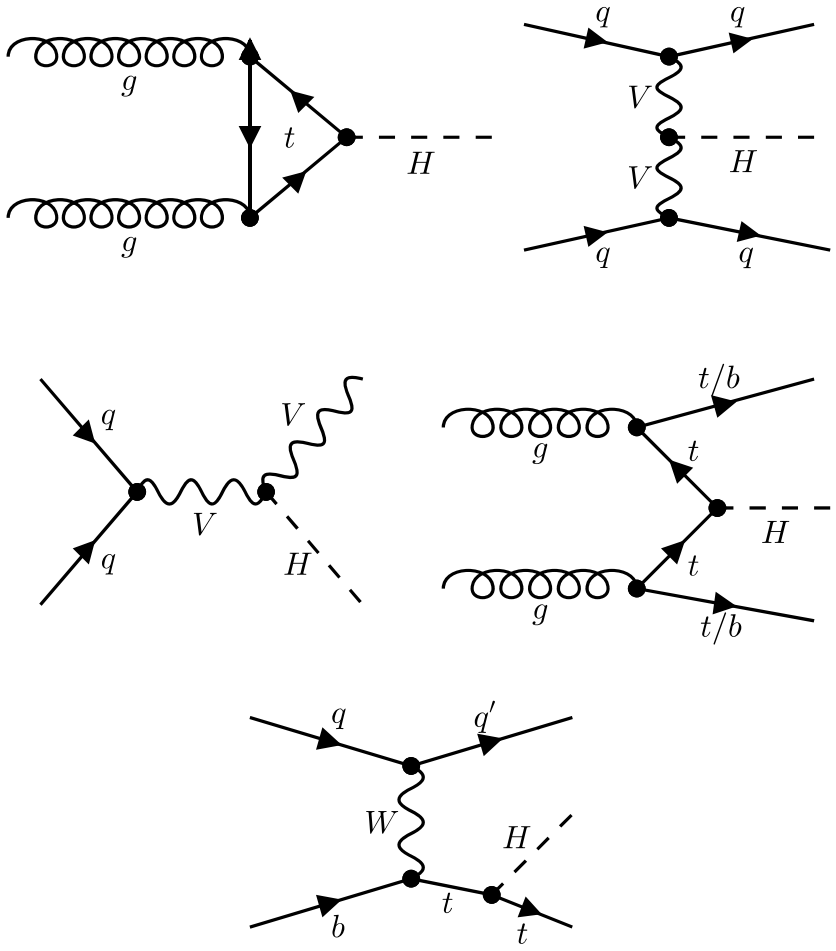


Figure 1.4: Leading-order Feynman diagrams of the main Higgs boson production modes at the LHC.

Decay channel	BR (%)
bb	58.24
WW	21.37
gg	8.187
$\tau\tau$	6.272
$c\bar{c}$	2.891
ZZ	2.619
$\gamma\gamma$	0.227

Table 1.1: Branching fractions (in %) for the H decay into the most probable final states for a Higgs boson mass of 125 GeV. Values extracted from [13].

cay modes, the linear proportionality between the Higgs boson couplings and the mass of the fermions, and the squared proportionality with respect to the mass of

the vector bosons can be verified. As shown in Fig. 1.5, the measured couplings with fermions and vector bosons follow the predicted dependency on the mass of the particles for a wide range of particle masses, which gives a strong evidence of the validity of the EWSB mechanism.

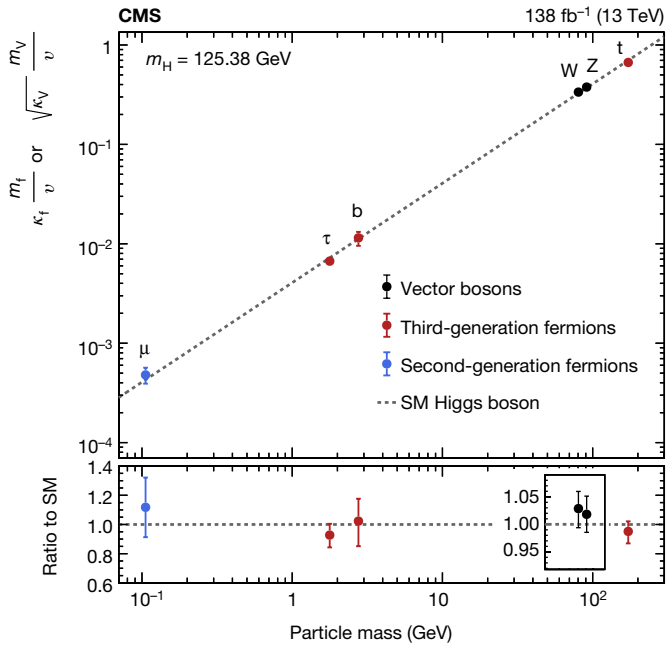


Figure 1.5: Measured normalised coupling constants of the Higgs boson to fermions and vector bosons, as functions of the fermion or vector boson mass. Extracted from [14].

1.2 Higgs boson pair production at the LHC

The production of Higgs boson pairs (HH) is a rare process (around 1000 times less probable than single Higgs boson production) predicted by the SM. Its study provides information not only about the Higgs self-coupling, but also about other couplings with the t quark or vector bosons. Precise measurements of these couplings can also probe Beyond the Standard Model (BSM) effects: small variations of their values can lead to big changes in the HH production cross section and the kinematics of the produced particles.

1.2.1 HH production modes

Fig. 1.6 shows the cross section for the six main HH production modes and their dependence with respect to the energy of the collision. At 13 TeV (collision energy during the Run 2 of the LHC, see Section 2.1), the two dominant production modes are gluon fusion and vector boson fusion. Both are described in the following.

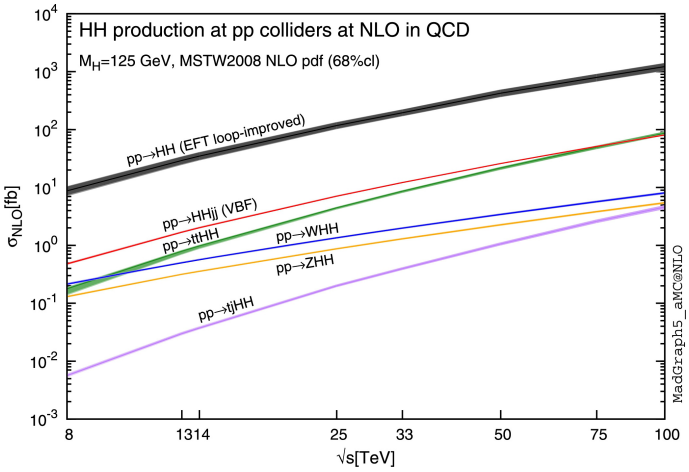


Figure 1.6: Total cross section at NLO for the six largest HH production modes at pp colliders: gluon fusion (black), vector boson fusion (red), top quark pair associated production (green), vector boson associated production (blue and yellow) and single top quark associated production (purple). The thickness of the lines corresponds to the scale and PDF uncertainties (see Section 1.3). Extracted from [15].

Gluon fusion (ggF) is the dominant production mode, with a cross section predicted by the Standard Model of around 31 fb. Since gluons are massless, they do not couple to the Higgs boson directly, but through heavy quark loops. This loop is dominated by the t quark ($\sim 99\%$), followed by the b quark with 1%. The LO diagrams of this production mode are shown in Fig. 1.7, denoted as *triangle* and *box* diagrams. The triangle contribution depends linearly on the Higgs boson self-coupling λ and the Yukawa coupling y_t, while the box diagram only depends quadratically on y_t.

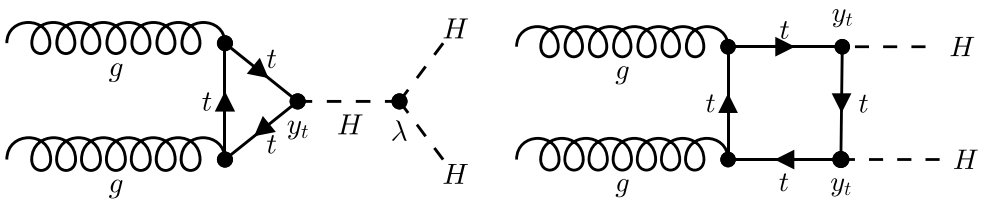


Figure 1.7: Feynman diagrams contributing to Higgs boson pair production via gluon fusion in the SM at leading order. The left diagram is often referred as the *triangle* diagram, whereas the right one, as the *box* diagram.

Vector boson fusion (VBF) is the subdominant production mode at the LHC energy ranges, with a cross section predicted in the Standard Model of ~ 1.7 fb. The most distinctive feature of this production mode is the presence of two additional quarks with high energy and very close to the initial direction of the colliding particles. Experimentally, this translates into a couple of particle jets with high invariant mass and large separation in pseudorapidity (see Section 2.1). This effect

can be seen in Fig. 1.8, showing the distributions of the invariant mass of the two jets with the largest invariant mass, and the difference in pseudorapidity between the two jets with the highest difference, for samples produced ggF and VBF. Three Feynman diagrams, shown in Fig. 1.9, contribute at LO to this production mode. The first diagram depends linearly on the coupling between two vector bosons and two Higgs bosons, λ_{2V}. The second depends quadratically on the coupling between two vector bosons and one Higgs boson (λ_V), while the third one linearly on both λ_V and the Higgs boson self-coupling.

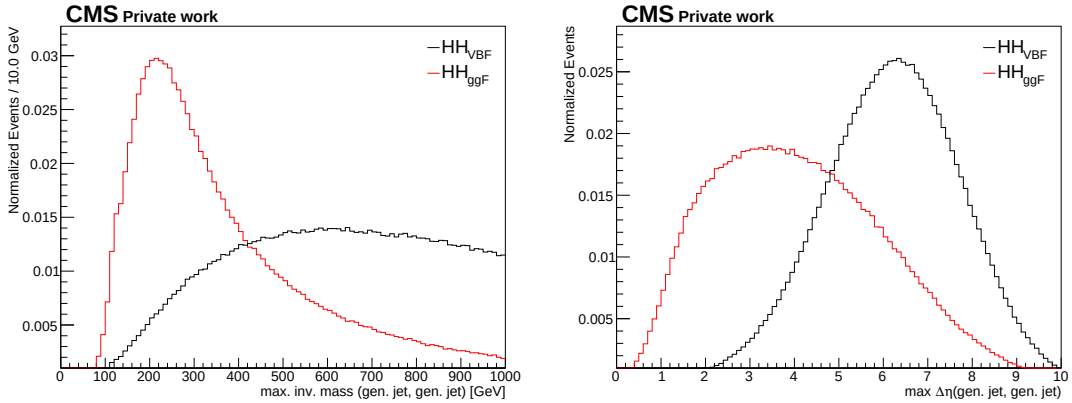


Figure 1.8: (Left) Invariant mass of the two jets at generator level with the highest invariant mass and (right) difference in pseudorapidity between the two jets at generator level with the highest difference in pseudorapidity, for HH samples produced via VBF (black) and ggF (red).

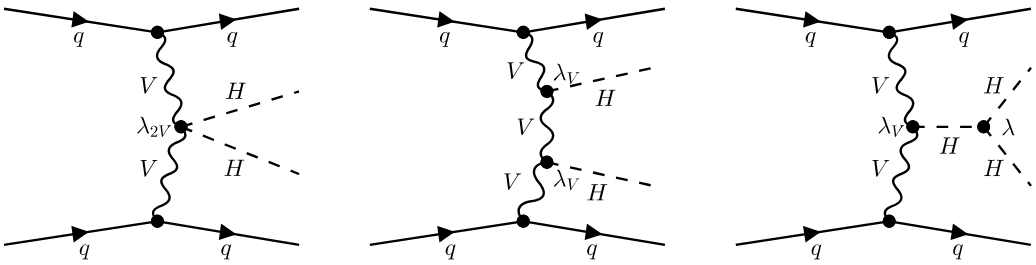


Figure 1.9: Feynman diagrams contributing to Higgs boson pair production via vector boson fusion in the SM at leading order.

Within each production mode, the final cross section is the result of the interference of its corresponding diagrams. For instance, the differential cross section corresponding to the diagrams related to the ggF production mode and their interference is represented in Fig. 1.10 as a function of the Higgs pair invariant mass. Its interference is shown to be destructive, which reduces the cross section of this process.

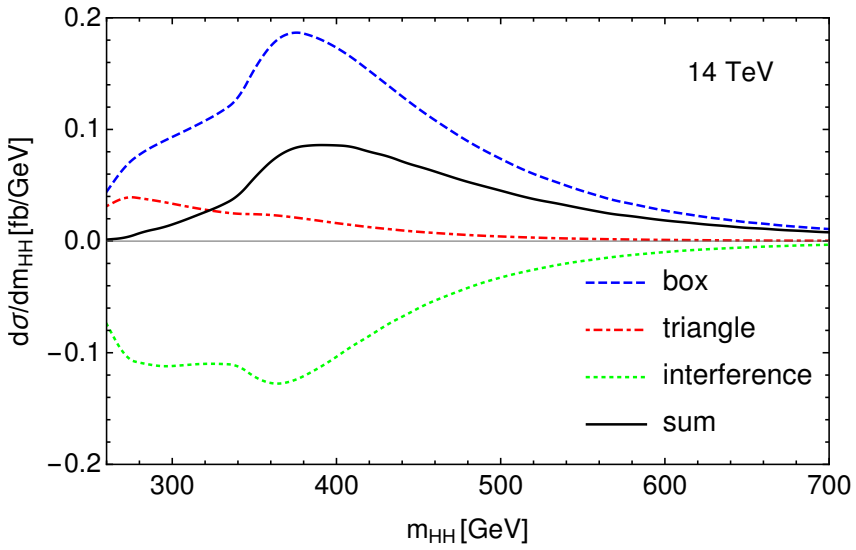


Figure 1.10: Differential cross section corresponding to the different contributions to the gluon fusion production mechanism and their interference. Extracted from [16].

The value of the λ coupling only depends in the SM on the values of v and m_H. However, some BSM theories predict a modification of the λ value, varying the properties of the Higgs boson pair production. To study possible BSM effects, a parametric approach is followed, so λ is left as a free parameter and deviations from the SM are measured by studying the Higgs self-coupling modifier, defined as

$$\kappa_\lambda = \lambda/\lambda_{\text{SM}}. \tag{1.33}$$

Changes on the value of κ_λ produce big modifications on the production cross section. Table 1.2 shows some values of the ggF production cross section for different values of κ_λ. By fitting those points, the production cross-section dependence on κ_λ can be obtained as

$$\sigma(\kappa_\lambda) = 70.3874 - 50.4111 \times \kappa_\lambda + 11.0595 \times \kappa_\lambda^2. \tag{1.34}$$

κ_λ	-1	0	1	2	2.4	3	5
σ [fb]	$131.9^{+2.5\%}_{-6.7\%}$	$70.38^{+2.4\%}_{-6.1\%}$	$31.05^{+2.2\%}_{-5.0\%}$	$13.81^{+2.1\%}_{-4.9\%}$	$13.1^{+2.3\%}_{-5.1\%}$	$18.67^{+2.7\%}_{-7.3\%}$	$94.82^{+4.9\%}_{-8.8\%}$

Table 1.2: HH ggF production cross section at $\sqrt{s} = 13$ TeV at NNLO precision for different values of the Higgs self-coupling modifier. Extracted from [17].

For the VBF production mode, a similar behaviour is obtained for different values of κ_λ, but also for the other couplings involved: κ_V (λ_V coupling modifier) and κ_{2V} (λ_{2V} coupling modifier). In fact, the dependence on κ_{2V} is even larger than the one on κ_λ, as shown in Fig. 1.11.

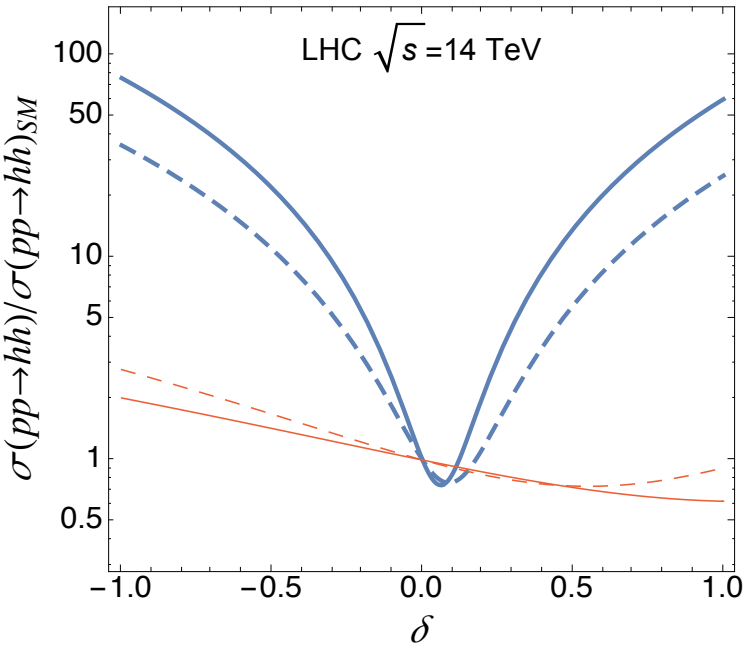


Figure 1.11: HH VBF production mode cross section at $\sqrt{s} = 14$ TeV, in units of the SM value, as a function of $\delta_{\kappa_{2V}} = \kappa_{2V} - 1$ (blue) and $\delta_{\kappa_\lambda} = \kappa_\lambda - 1$ (orange) after applying cuts reproducing the realistic acceptance of LHC detectors (solid) and more specific VBF analysis selections (dashed). Extracted from [18].

Variations on the value of the coupling strengths not only modify the production cross section, but also the kinematics of the events. As an example, Fig. 1.12 shows the distributions of the HH system mass and the difference in pseudorapidity between the two Higgs bosons at generator level produced via VBF with three different values of κ_{2V}. The one produced with the couplings set to their SM values populates much more the low mass region and the high η difference with respect to the other two samples, which vary in one unit the κ_{2V} coupling strength. This effect has important consequences for the experimental searches, since some variables that could be used to distinguish between signal and background could reduce its discriminating power when considering BSM couplings.

1.2.2 HH decay channels

Several final states can be produced by the decay of both Higgs bosons at the energy ranges of the LHC. Since HH production is a very rare process, targeting final states with sufficient branching ratio (for example, by requiring that one of the Higgs boson decays into two b quarks) is desirable. The different channels available and their corresponding branching fractions are shown in Fig. 1.13.

Each decay channel is affected by different experimental challenges, and not necessarily the channel with the largest BR will be the most sensitive. The three channels that provide the best sensitivity are summarised in the following.

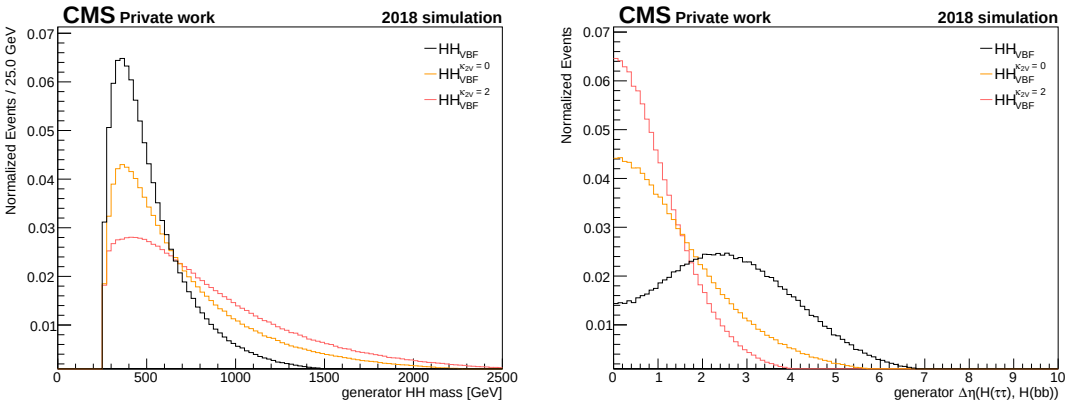


Figure 1.12: (Left) Mass of the HH system at generator level and (right) difference in pseudorapidity between the two Higgs bosons at generator level for HH VBF samples produced (black) with the couplings set to their SM values, (yellow) with $\kappa_{2V} = 0$ and (red) with $\kappa_{2V} = 2$.

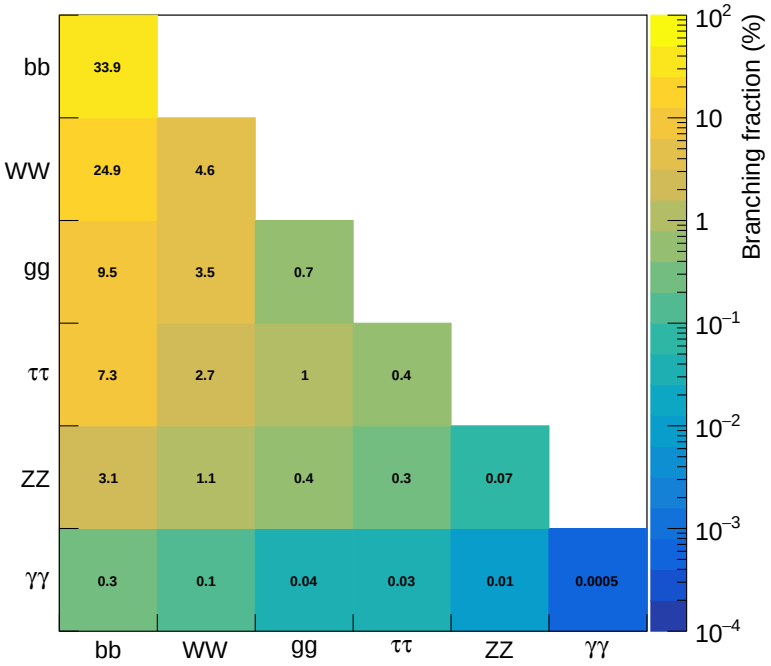


Figure 1.13: Branching fractions (in %) for the HH decay into some of the possible final states for a Higgs boson mass of 125 GeV. Values extracted from [13].

- HH → bbbb is the final state with the highest branching fraction (∼ 33.9%). However, this final state suffers from large background contamination coming from processes with jets in the final state (e.g. QCD multijet).

- HH $\rightarrow$ bb$\tau\tau$ has a smaller branching fraction ($\sim$ 7.3%) but higher purity thanks to the $\tau\tau$ pair. It suffers from contamination coming from QCD multijet, Z/γ^*+ jets and $t\bar{t}$ events.
- HH $\rightarrow$ bb$\gamma\gamma$ is a very pure final state thanks to the two photons, which provide a very clear signal; however, its branching fraction is very small ($\sim$ 0.3%).

The HH $\rightarrow$ bb$\tau\tau$ is the final state studied in this thesis. Its medium branching ratio and relatively low background provides an interesting compromise, even though efficient particle identification and background rejection algorithms need to be considered.

1.3 Proton-proton collision phenomenology

Several processes occur between the beginning of a proton-proton collision and the measurement of particles in the final state. This section introduces these different processes.

Protons are not elementary particles, but composed by *partons*, quarks and gluons, which are the ones that actually interact during the collision. The initial energy of each incident parton is described by probability density functions called *parton density functions* or PDF, which describe the amount of proton momentum carried by a certain parton. Two sets of PDFs are used in this thesis: NNPDF 3.0 [19] for the 2016 data sets, and NNPDF 3.1 [20] for 2017 and 2018 the data sets.

Among all parton interactions in the proton-proton beam collision, the *hard interaction* is the one having the largest momentum exchange between interacting partons. The matrix element that describes the hard interaction can be computed with different accuracy by including terms at different orders (represented as Feynman diagrams). The higher the order, the more accurate and computationally expensive. These computations are performed using Monte Carlo (MC) event generators [21]. Some of the most common generators, and the ones used in this thesis, are the POWHEG [22] and MADGRAPH5_AMC@NLO [23] generators.

Due to the strength of the strong force, partons can radiate gluons at any stage of the process, either as initial state radiation or final state radiation. The emission of gluons results in *parton showers*, where gluons split into quark-antiquark or gluon-gluon pairs and quarks emit more gluons. The energy of the produced particles decreases until the production of fundamental particles is no longer energetically favourable and hadron formation (*hadronization*) becomes the dominant process. Parton shower and hadronization effects are modelled in this thesis with PYTHIA [24]. To mediate between the event generation and the parton showering, different *merging schemes* are considered, so double-counting of jets is avoided and a better modelling is achieved. Some examples of merging schemes used in this thesis are the MLM [25] and FxFx [26] methods.

In parallel to the partons in the hard interaction, other partons may also interact with smaller transfer of momentum between them. The initial state radiation, the

final state radiation, and beam remnants constitute the *underlying event* (UE), i.e. everything that is not part of the hard interaction. To simulate the UE activity, generators rely on phenomenological models, whose parameters are then constrained using collision data. This calibration process is called *tuning* of the generator. Two different tunings are used for the samples considered in this thesis, CUETP8M1 [27] and CP5 [28].

Protons in the beams travel in bunches containing around 10^{11} of them, so multiple interactions can happen each time bunches cross. These additional interactions with respect to the hard interaction are called *pile-up* (PU). PU interactions generate much activity in the detectors that could complicate their reconstruction capabilities, so a correct modelling of PU is crucial. The expected PU profile is included in the MC simulations.

Finally, given the set of final-state particles produced through the event generation procedure, the detector response has to be simulated. The GEANT4 toolkit [29] models the propagation of particles through matter. Afterwards, the detector readout is then simulated so it resembles its actual response on real data, producing hits and energy deposits that can be used to reconstruct the event particles.

Chapter 2

The Large Hadron Collider and the CMS experiment

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the most powerful particle accelerator ever built [30]. It's located at CERN (*Organisation européenne pour la recherche nucléaire*, European Organization for Nuclear Research) in the border between France and Switzerland, near Geneva. The accelerator is a 27 km circular collider built around 100 m underground composed by two rings where two beams of protons or heavy ions circulate in opposite directions inside vacuum pipes. The beams are bended using 1232 superconducting dipoles, cooled down with superfluid helium to 1.9 K, which generate an 8.3 T magnetic field. To focus the beams and avoid spreading, 392 quadrupole magnets are used.

The LHC is the last step of the accelerating chain in which the particles reach the operating energies. The CERN accelerator complex is shown in Fig. 2.1. At the beginning of the chain, hydrogen atoms are ionized and injected into the LINAC2, where they are accelerated to an energy of around 50 MeV. Then, the beam is passed through a series of circular synchrotrons, reaching 1.4 GeV after the BOOSTER, 26 GeV after the PS and 450 GeV after the SPS, before they enter into the LHC and reach their target energy. The designed center-of-mass energy ($\sqrt{s}$) for proton-proton collisions is 14 TeV, although the target energy has changed over the different phases: 7 TeV during 2010 and 8 TeV during 2011 and 2012 (3-year period called Run 1), 13 TeV during Run 2 (2015-2018) and 13.6 TeV during Run 3 (which started in 2022).

The beams are made to intersect at four points in the LHC, where the four experiments are placed. ATLAS (*A Toroidal LHC ApparatuS*) and CMS (*Compact Muon Solenoid*) are two general purpose experiments with similar physics programs located in opposite sites of the LHC ring. ALICE (*A Large Ion Collider Experiment*) is focused mostly on heavy ion collision in order to study quark-gluon plasma, while LHCb (*LHC beauty*) focuses on CP-violation and the matter-antimatter asymmetry by studying b quark physics.

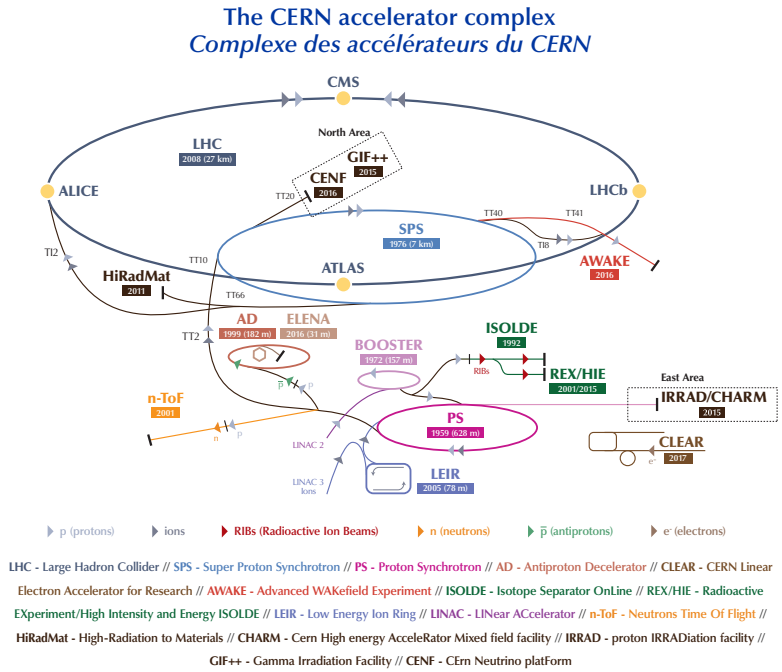


Figure 2.1: CERN accelerator complex. Extracted from [31].

2.1.1 Luminosity

In the beams, *bunches* are formed by grouping n protons within a 25 ns time spacing (Δt_0), so the probability of interaction between the beams when they collide increases. Sets of consecutive bunches are further structured in *trains*. A *fill* is defined as the period of data taking between the beam is injected in the accelerator until it is dumped.

The beam's main properties are the transverse emittance, ϵ, which reflects how close two protons in the same bunch are in the position-momentum phase space; and the amplitude function, β^*, which defines how focused the beam is.

The instantaneous luminosity, L_{ins}, determines the amount of collisions produced per unit of time and area. It can be obtained considering the previously described design and operation parameters and is given by,

$$L_{ins} = \frac{n_1 n_2 f}{4\Delta\epsilon\beta^*}, \tag{2.1}$$

where n_1 and n_2 are the number of protons in the two bunches that collide and f is the beam frequency. Typical values for these parameters during Run 2 are $n = 1.2\times10^{11}$ protons per bunch, $f = 40$ MHz, $\epsilon = 3.75$ mm$\times\mu$rad and $\beta^* = 0.55$ m, which give a nominal instantaneous luminosity in the order of 10^{34} cm^{-2}s^{-1}, with a maximum obtained during Run 2 of $\sim 2 \times 10^{34}$ cm^{-2}s^{-1}.

For a given process with cross section σ, the expected number of events is given

by

$$\frac{dN}{dt} = \sigma L_{ins}, \tag{2.2}$$

so, in order to study physical processes with low production cross section, a high luminosity is needed. The total number of pp events is proportional to the integrated luminosity:

$$L_{int} = \int L_{ins}\, dt. \tag{2.3}$$

Fig. 2.2 shows the total integrated luminosity delivered by the LHC and the recorded by the CMS experiment during the Run 2 data taking period.

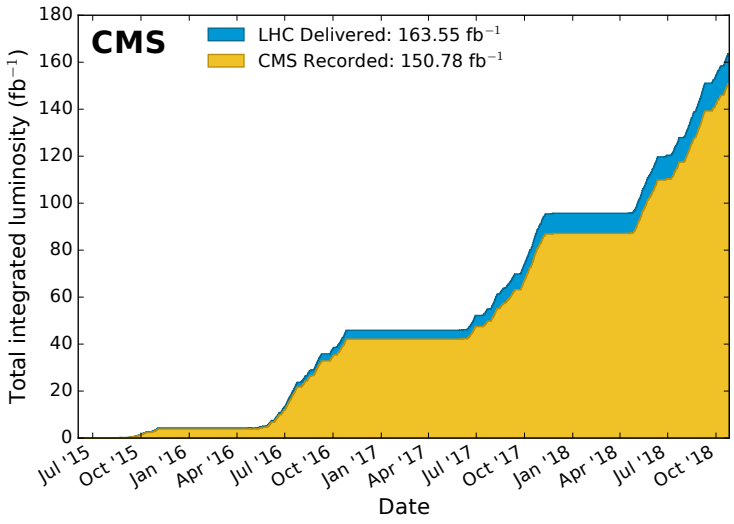


Figure 2.2: Integrated luminosity delivered by the LHC and recorded by the CMS experiment during the Run 2 data taking period. Extracted from [32].

The average number of simultaneous interactions per bunch crossing, i.e. the pile-up (PU), is obtained as

$$\langle PU \rangle = \frac{L_{ins}\ \sigma_{pp}^{inel}\ \Delta t_0}{n}, \tag{2.4}$$

where σ_{pp}^{inel} is the inelastic pp cross section. High PU values result in a very high detector occupancy, which degrades the efficiency and resolution of the particle reconstruction. Fig. 2.3 shows the PU distributions during the Run 2 data-taking period.

2.1.2 The High Luminosity LHC

Throughout the lifetime of the LHC, several improvements on the accelerator have been implemented during long shutdowns (LS) in order to extend its discovery potential, as displayed on the roadmap in Fig. 2.4. In LS1, the interconnections between

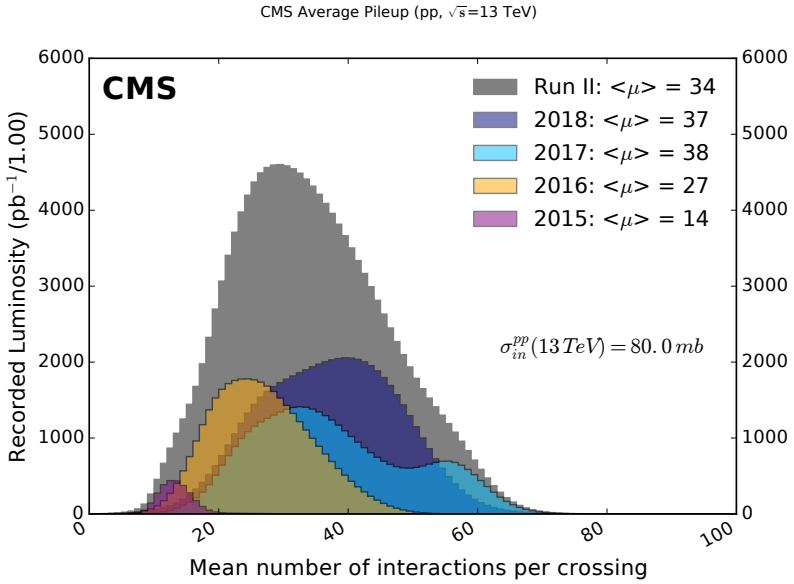


Figure 2.3: Distribution of the average number of interactions per crossing (pile-up) for pp collisions in 2015 (purple), 2016 (orange), 2017 (light blue), 2018 (navy blue), and the full Run 2 period (gray).

the LHC superconducting magnets were consolidated to increase the center-of-mass energy to 13 TeV and allowing the machine to reach instantaneous luminosities of up to 2.1×10^{34} cm^{-2}s^{-1}, equivalent to an average pile-up of 55. During LS2, an optimization of the LHC parameters was performed so the machine can increase its center-of-mass energy to 13.6 TeV and sustain a maximum instantaneous luminosity of 2×10^{34} cm^{-2}s^{-1} during longer periods of time during Run-3. The next long shutdown, LS3, will include a major upgrade to both the accelerator and the experiments, starting the Phase 2 operation of the LHC or High Luminosity LHC (HL-LHC) [33]. In its nominal operation, the accelerator will run at a centre-of-mass energy of up to 14 TeV, while the instantaneous luminosity will increase during Run 4 up to $5-7.5\times10^{34}$ cm^{-2}s^{-1}, corresponding to a PU of 140-200, potentially leading to a total integrated luminosity of up to 4000 fb^{-1} after 10 years of operations. This increase needs to be addressed by corresponding detector upgrades in order to be able to cope with the higher instantaneous and integrated luminosities and maintain the performance of the reconstruction system in this new PU scenario.

2.2 The CMS experiment

The CMS detector [34] is a general-purpose experiment located 100 m underground at the LHC point 5, near the french village of Cessy. The apparatus is a cylindrical detector with a length of 21.6 m and a diameter of 14.6 m. Its main feature is the presence of a superconducting solenoid of 6 m internal diameter and 12.5 m length, which creates an homogeneous magnetic field of 3.8 T. The intense magnetic

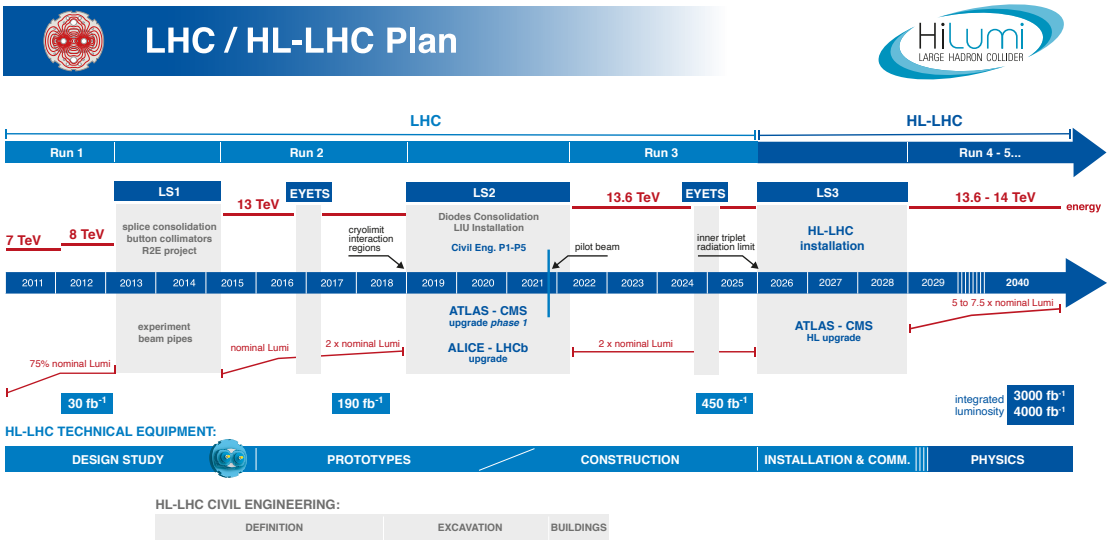


Figure 2.4: LHC baseline plan for the next decade and beyond showing the collision energy (upper line), the instantaneous luminosity (lower line) and the total integrated luminosity (lower boxes). Running periods and long shutdowns are also displayed.

field allows to measure the transverse momentum and the electric charge of charged particles from their curved trajectories.

2.2.1 Coordinate system

The CMS experiment uses a right-handed coordinate system centred in the nominal interaction point. The x-axis points towards the centre of the LHC ring, the y-axis vertically and the z-axis along the beam direction. Polar coordinates are usually used, so the azimuthal angle ϕ is defined in the xy plane by the angle with respect to the x-axis and the polar angle θ is defined in the rz plane as the angle formed with the z-axis. An schematic of this coordinate system is shown in Fig. 2.5.

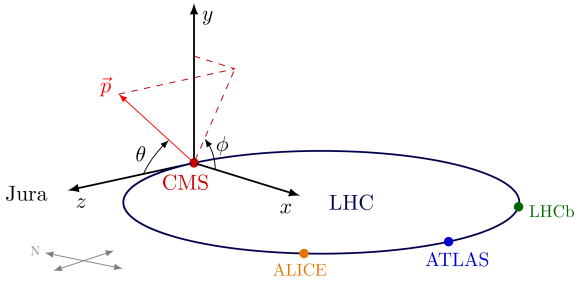


Figure 2.5: Illustration of the CMS coordinate system. Extracted from [35].

The *rapidity* y is defined as

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \tag{2.5}$$

where E and p_z are the energy and the momentum along the beam axis respectively for a given particle. Differences in rapidity are invariant under Lorentz boosts, so this variable is very useful in hadronic collision studies. Instead of this variable, for highly relativistic particles an approximated quantity called the *pseudorapidity* η can be used:

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right), \tag{2.6}$$

which does not depend on the particle energy or momentum. Pseudorapidity varies from 0 when $\theta = \pi/2$ to infinity when $\theta = 0$. Regions with large $|\eta|$ are often referred to as *forward*, while the ones with low $|\eta|$, as *central*.

Using the angle ϕ, the pseudorapidity η and the momentum in the plane orthogonal to the beam axis p_T (*transverse momentum*), the particle's momentum can be completely defined. They are related to the usual p_x, p_y and p_z by

$$p_x = p \sin\theta \cos\phi = p_T \cos\phi, \tag{2.7}$$

$$p_y = p \sin\theta \sin\phi = p_T \sin\phi, \tag{2.8}$$

$$p_z = p \cos\phi = p_T \sinh\eta. \tag{2.9}$$

Another variable that will be useful later is the angular distance ΔR between two particles with coordinates (η_1, ϕ_1) and (η_2, ϕ_2),

$$\Delta R = \sqrt{(\eta_1 - \eta_2)^2 + (\phi_1 - \phi_2)^2}. \tag{2.10}$$

This variable is usually considered to perform jet clustering or to evaluate the *isolation* of a particle with respect to its neighbouring particles, i.e. the amount of energy found in a cone around the particle. This metric helps to discriminate between *prompt* particles (i.e. coming from the hard interaction) from particles found inside jets or coming from soft scattering.

2.2.2 Subdetectors

The structure of the CMS detector is shown in Fig. 2.6. The central zone of the detector, or *barrel*, is divided intro cylindrical slices, called *wheels*. Each wheel is numbered from -2 to 2, being wheel 0 the one where the interaction point is located. The two forward regions, beyond wheels -2 and 2, are called the *endcaps*.

The detector is composed by several subdetectors nested around the interaction point, each with a different technology to detect different kinds of particles. Ordered by vicinity to the interaction point, they are located both before (silicon tracker, electromagnetic calorimeter and hadronic calorimeters) and after (muon system) the magnetic solenoid. A description of each subdetector (starting from the solenoid magnet) is included in the following.

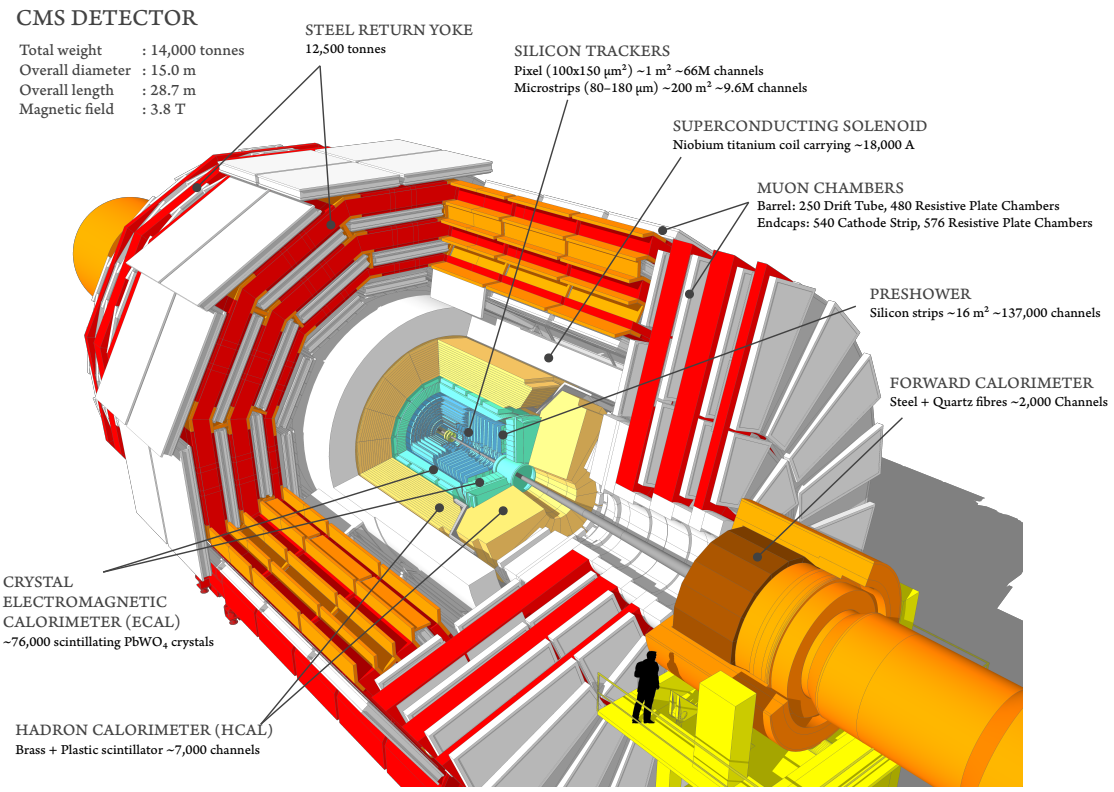


Figure 2.6: Schematic layout of the CMS subdetectors. Extracted from [36].

The solenoid magnet

The solenoid magnet is the central feature of the CMS detector. It is a niobium-titanium superconducting solenoid operating at 4.5 K. It provides a 3.8 T magnetic field inside the solenoid and up to 2 T outside it. The return yoke of the magnet is made of steel, and is located between the chambers of the muon system.

The magnetic field bends the trajectory of charged particles produced in the collision, and their transverse momentum can be inferred from the curvature of their trajectory. A larger magnetic field produces a larger bending, increasing the precision on the momentum determination.

The silicon tracker

The CMS tracker [37, 38] is the innermost part of the CMS detector. It has a diameter of 2.5 m and a length of 5.8 m, covering up to $|\eta| < 2.5$. It's built in silicon, so it can resist well the high radiation levels produced during the collisions while providing very good spatial resolution. When charged particles transverse the tracker volume, they deposit their energy by ionizing the silicon semiconductors, creating electron-hole pairs that induce currents that drift to the electrodes, where they are digitized and measured. These *hits* can be combined into *tracks* by using pattern recognition algorithms (see Section 3) to reconstruct both its trajectory

and origin. This origin could be the main interaction point, or *primary vertex*, or *secondary vertices* displaced a few millimetres from the primary vertex. These secondary vertices are produced from the decay of particles like heavy-flavour quarks or τ leptons.

A structural view of the tracker is shown in Fig. 2.7. It consists of two main detectors: a pixel detector surrounding the beam pipe and a strip detector surrounding the pixel detector.

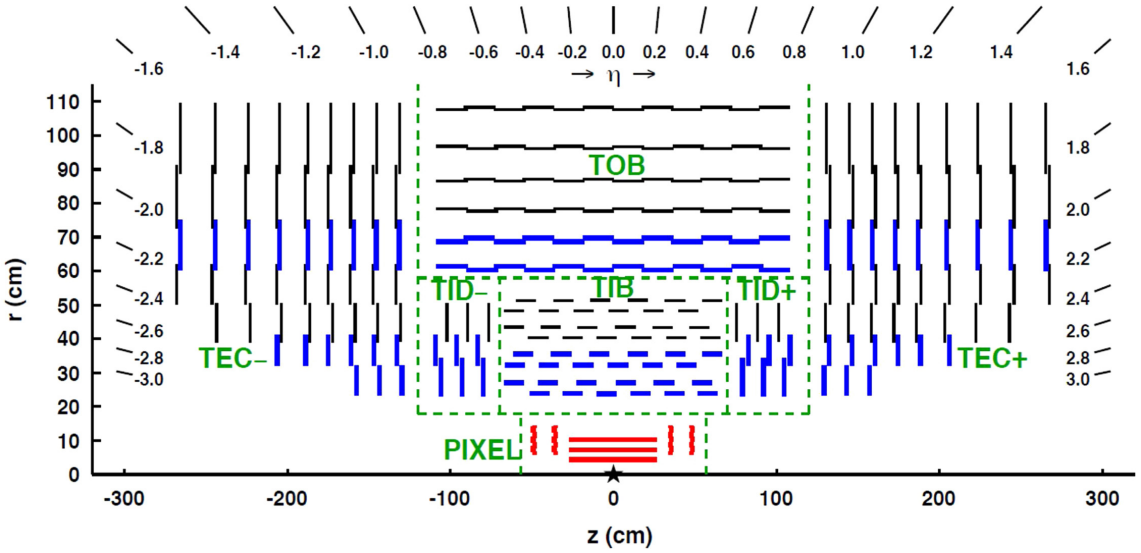


Figure 2.7: Schematic structure of a tracker slice in the r-z plane. Extracted from [39]

The pixel tracker covers the innermost region of the tracker, where the flux of particles coming from the collisions is higher. At these small distances to the interaction point, a very high granularity is then needed in order to resolve hits of neighbouring particles. The detector originally consisted of three detector layers in the barrel region with radii 4.4, 7.3, and 10.2 cm, and two layers on each side of the detector at distances of $\pm$34.5 cm and $\pm$46.5 cm. An upgrade on the detector was performed before the 2017 data-taking period [40], so a high tracking performance could be maintained when reaching luminosities up to 2×10^{34} cm^{-2}s^{-1} and |PU| exceeding 50. In this upgrade, the detector was replaced with four barrel and three endcap layers on each side. A comparison between before and after the upgrade is shown in Fig. 2.8. A total of 124 million pixels are included in the upgraded pixel detector, each with a 100$\times$150 μm^2 cross section. The detector produces 3-D measurements along the paths of the charged particles with single hit resolutions around 10 μm.

In the outer part of the tracking system, the strip detector consists of 9.3 million silicon micro-strip sensors. Given the lower occupancy in this region, the use of microstrips is enough to provide an unambiguous hit detection. The strip tracker is divided into four regions with different layouts. In the barrel, the Tracker Inner Barrel (TIB) covers the region of radii between 20 and 55 cm, while the Tracker

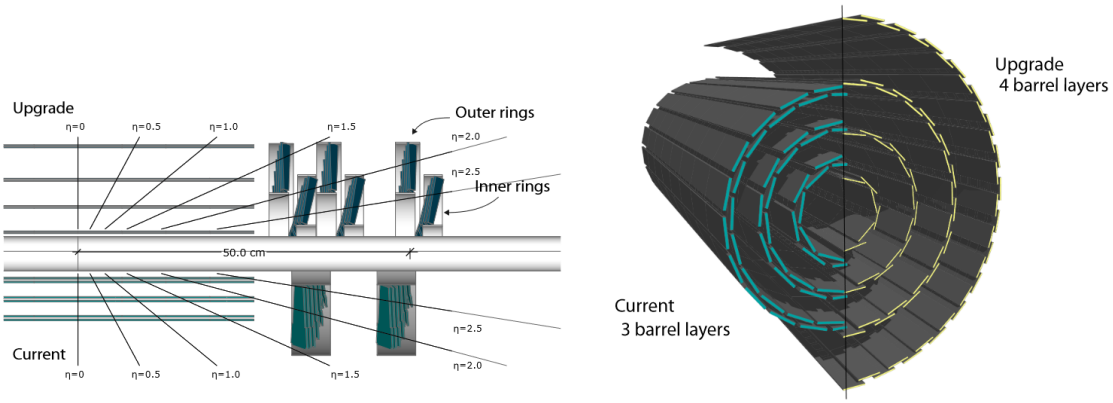


Figure 2.8: Comparison of the pixel detector before and after to 2017 upgrade (labeled as *Current* and *Upgrade* respectively). Extracted from [40].

Outer Barrel (TOB) reaches up to $r = 116$ cm. They contain four and six layers respectively. In the endcaps, the TIB is covered by the Tracker Inner Disk (TID), made out of three layers on each side of the detector, ant the Tracker Endcaps (TEC) in the forward regions, with nine layers. The TIB and TID provide single hit position resolutions of around 13-38 μm, while the TEC and TOB, approximately 18-47 μm.

The Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) [41] is a hermetic and homogeneous array of lead tungstate ($PbWO_4$) scintillating crystals. It is designed so an incident electron or photon generates an electromagnetic shower that deposits most of its energy inside the ECAL volume. With the location and magnitude of the energy deposits, the energy and the direction of the incident particle can be measured. The characteristics of the $PbWO_4$ crystals make them an appropriate choice for kind of operation at the LHC, thanks to its high density (8.28 g/cm^3), short radiation length (0.89 cm) and small Molière radius (2.2 cm), resulting in a fine granularity and a compact calorimeter.

The ECAL layout is shown in Fig. 2.9. It is divided in three different regions: the barrel ECAL (EB) ($|\eta| < 1.4.79$), two endcaps (EE), with a $1.57 < |\eta| < 3$ coverage, and a preshower detector (ES) in front of the endcaps ($1.65 < |\eta| < 2.6$).

The EB is instrumented with 61200 trapezoidal crystals, each one covering a 22×22 mm^2 surface in the front face, equivalent to 0.0174×0.0174 in the η-ϕ plane, and 26×26 mm^2 at the rear face. The crystal length is 230 mm, corresponding to $\sim$ 26 radiation lenghts. They are mounted in a way to avoid cracks between crystals aligned with the expected particle trajectories. In each EE, 7324 crystals are mounted, with a rear face cross section of 30×30 mm^2, a front face cross section of 28.62×28.62 mm^2 and a length of 220 mm ($\sim$ 25 radiation lenghts).

In front of each endcap, the ES is composed of two layers of lead absorber

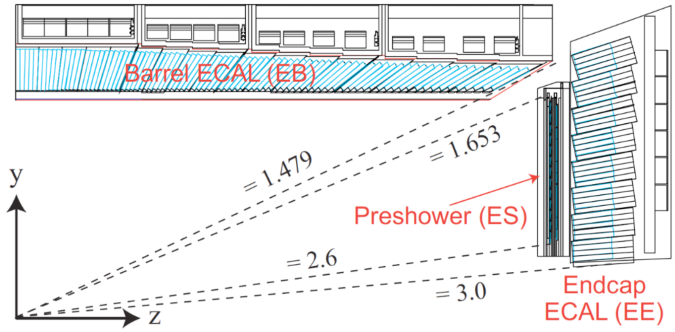


Figure 2.9: Schematic view of one quarter of the CMS ECAL. Extracted from [42].

instrumented with orthogonal layers of strip sensors in order to help with π^0/γ separation.

The energy resolution for the ECAL can be parametrized as follows [34]:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2, \tag{2.11}$$

where S is stochastic term (coming from fluctuations in the shower), N the noise term (electronics, digitization or pileup noise) and C as constant term (intercalibration errors or energy leakage from the back of the crystal). This estimation was confirmed in a test beam by summing 3×3 crystals, obtaining a resolution of:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{12\%}{E}\right)^2 + (0.3\%)^2. \tag{2.12}$$

The Hadron Calorimeter

The Hadron Calorimeter (HCAL) [43] completely surrounds the ECAL, covering pseudorapidities up to $|\eta| < 5.2$. It is a sampling calorimeter consisting of alternating layers of passive material (non-magnetic brass absorber) and plastic scintillators. In the dense layers of the passive material the hadrons interact, creating hadronic showers and inducing detectable light in the plastic scintillators. The total amount of light provides a measurement of the energy of the particle, and the calorimeter segmentation allows to measure the position of the shower.

The CMS HCAL consists of four different subsystems, as shown in Fig. 2.10. The HCAL Barrel (HB) covers a range of $|\eta| < 1.4$, while the HCAL Endcaps (HE) cover a range of $1.3 < |\eta| < 3$. The HCAL outer (HO), formed by additional layers of scintallators, is situated outside the solenoid and is used to measure energy leakages from the HB. It covers a range of $|\eta| < 1.26$ and matches the ϕ segmentation of the muon system. To make the HCAL as hermetic as possible, the HCAL Forward (HF) is placed on both sides of the detector beyond the muon system, covering the high η region, $3 < |\eta| < 5.2$. Unlike the other HCAL subsystems, the active material is quartz fibers embedded in the steel, which emit Cherenkov light when particles pass through.

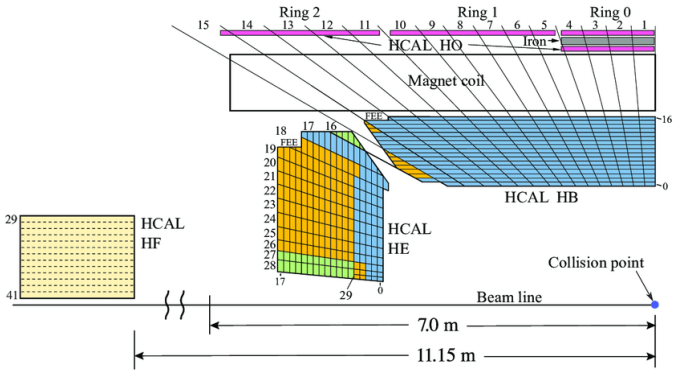


Figure 2.10: Schematic view of one quarter of the CMS HCAL. Extracted from [43].

The energy resolution of the HCAL is designed to be approximately

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{100\%}{\sqrt{E}}\right)^2 + (4.5\%)^2. \tag{2.13}$$

This was confirmed with cosmic and test beam data [44], as the hadronic energy resolution of the barrel HCAL and ECAL combination was measured to be

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{84.7\%}{\sqrt{E}}\right)^2 + (7.4\%)^2. \tag{2.14}$$

The Muon System

Muons are produced in several physical processes, so the precise measurement of their momenta is key for both Standard Model studies and new physics searches. However, muons interact poorly with the calorimeters and the inactive material, so they are able to travel outside the magnet. To detect them, the CMS Muon System [34,45] is located outside the solenoid, inserted between the layers of the iron return yoke of the magnet system. The charge and momentum of the muons can be measured using the magnetic field from the solenoid, providing a complementary measurement to the one performed in the tracker.

The muon chambers are instrumented with three different gas ionization detectors: the *Drift Tube chambers* (DT), the *Resistive Plate Chambers* (RPC) and the *Cathode Strip Chambers* (CSC). Their layout inside the CMS detector is shown in Fig. 2.11. In the three cases, when a muon transverses the subdetector, the gas inside the chambers is ionized, creating electric signals that are read out by dedicated electronics and are associated with several positions inside the chamber where the muon crossed it, known as *hits*. These are combined in each subsystem and then between subsystems to generate the muon candidates, as described in Section 3.2.1.

A brief description of the RPC and CSC systems is shown in the following. The DT system is further described in Section 2.3.

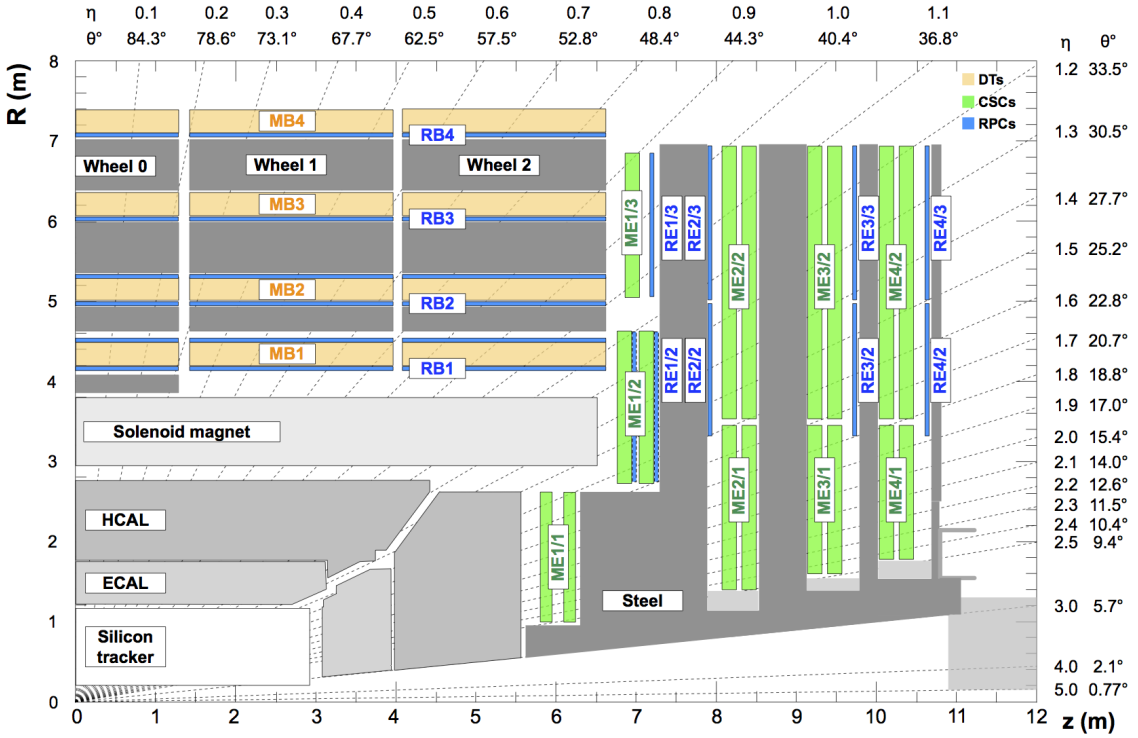


Figure 2.11: Schematic layout of a quadrant of the CMS detector. The location of the muon subsystems is highlighted. Extracted from [46].

Cathode Strip Chambers CSCs are located in the endcaps, covering η regions between $1.2 < |\eta| < 2.4$. They have shorter drift paths and faster time response, needed in order to cope with the higher muon and other particle rates and the stronger and non-uniform magnetic field found in the endcaps. They are also finely segmented, providing a three-dimensional spatial resolution. CSC detectors have a trapezoidal shape and contain alternate layers of anode wires and cathode stripes, filled with Ar, CO_2 and CF_4 (50%-40%-10%). Similarly to the DT chambers, the detection principle is the ionization of the gas when a muon crosses the chamber, inducing signals used to provide a precise measurement of the position, with an spatial resolution of 40-150 μm.

Resistive Plate Chambers RPCs are installed in both the barrel and endcaps, reaching up to $|\eta| < 1.9$. They are formed by two parallel plates, an anode and a cathode, separated by a mixture of gases. The detector works in avalanche mode: when the muon crosses a chamber, an electromagnetic cascade is produced by the high electric field, inducing a signal read out by the strips located in the outer surface. As the segmentation is very poor, so is the spatial resolution (0.8 - 1.2 cm), although they are faster detectors than both DTs and CSCs, with a time resolution better than 3 ns. Therefore, they are particularly useful in detecting the bunch crossing associated to the muon track, even in high PU scenarios.

2.2.3 CMS trigger system

Given the high luminosity provided by the LHC and the large amount of information coming from the detector cells, a huge data rate is produced by the experiment (∼TB/s). However, most of the collisions result in low energy pp collisions, not meant to be studied in the experiment, while the processes of interest for the LHC physics program have smaller cross sections by several orders of magnitude, as shown in Fig. 2.12. Therefore, a trigger system which reduces the event rate while keeping the events interesting for CMS physics analysis is needed.

The CMS trigger system [47] is implemented in two steps: the Level-1 (L1) Trigger, a hardware-based system, and the High Level Trigger (HLT), which is software-based.

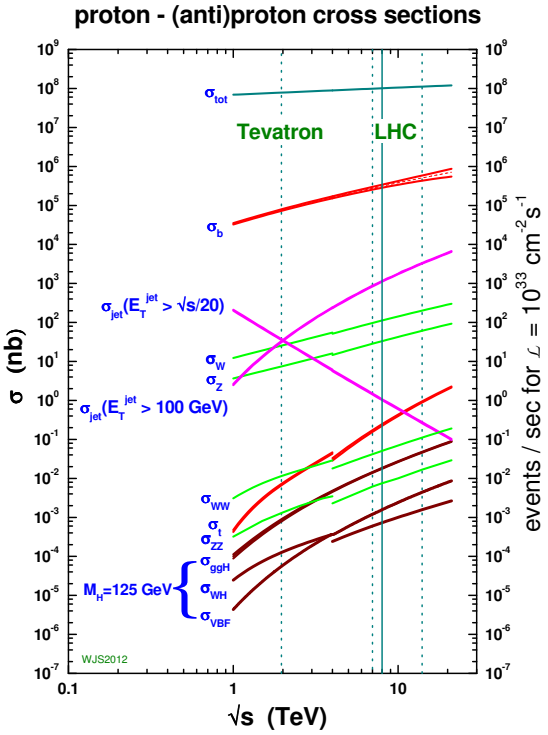


Figure 2.12: Summary of cross section calculations at NLO or NNLO at the LHC and Tevatron. Extracted from [48].

The Level-1 trigger

The L1 Trigger takes the data coming from the subdetectors and selects events at a rate of up to 100 kHz from the incoming 40 MHz, making decisions within a time interval of less than 4 μs after the collision. Due to the timing constrains, the L1 cannot run the whole reconstruction of the physics objects and their trajectories.

Instead, it produces the so-called L1 *candidates*, low resolution physics objects that rely only on the inputs coming from the calorimeters and the muon system.

Fig. 2.13 shows an schematic structure of the L1 trigger. The information of the energy deposits in the HCAL and ECAL subdetectors is processed by the *calorimeter trigger*. The transverse energy measured in the calorimeters is transmitted to the calorimeter trigger in form of *trigger primitives*, which are then combined into trigger object candidates: electrons, photons, hadronic τ (τ_h) leptons, jets and energy sums. As only the information of the calorimeter can be used, both electrons and photons are reconstructed together as L1 e/γ objects by clustering energy deposits. The L1 τ_h objects are identified from their decays into charged or neutral pions. These pions are reconstructed with a clustering algorithm adapted from the L1 e/γ algorithm. A key feature of the L1 τ_h objects is their isolation, which allows to discriminate them from QCD jets, keeping background under control and increasing the purity. L1 jets are reconstructed through energy deposits centred around a local maximum. The total sum of the transverse energy over all the energy deposits can provide an estimation of the missing transverse energy.

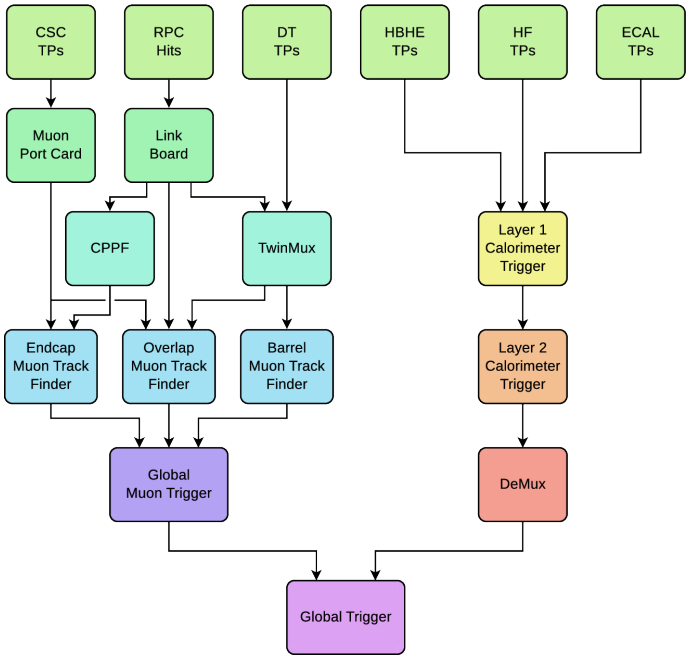


Figure 2.13: Diagram of the CMS L1 trigger system. Extracted from [49].

The DT and CSC subdetectors generate, at detector level, their trigger primitives by reconstructing segments on each chamber out of the received hits. These trigger primitives, together with RPC hits, are collected in the *muon trigger*. In the barrel, the Barrel Muon Track Finder (BMTF) uses the information from the DT and RPC subdetectors, whose information is merged into trigger primitives in the *TwinMux* system. The Endcap Muon Track Finder (EMTF) uses the data from the CSC and RPC detectors, while the Overlap Muon Track Finder (OMTF) makes use of the information coming from the three subdetectors. The three track finders create muon candidates with a p_T, η, ϕ and a quality criteria that are then sent to the

Global Muon Trigger (μGMT), which orders them by p_T and quality and removes the duplicates.

The candidates coming from the calorimeter and muon trigger are finally sent to the Global Trigger (μGT), which decides if the event is kept or discarded based on L1 selection algorithms commonly called *seeds*. Basic seeds can be built by applying thresholds on kinematic observables of one or more objects of one type (or *collection*), e.g. requiring one electron with p_T bigger than some value. More complicated seeds can also be included by applying thresholds on objects coming from different collections (often called *cross-triggers*) or on more elaborated observables, as $\Delta\eta$ or ΔR. A full set of seeds is known as the *L1 Menu*.

The L1 menu is configured to perform optimally under different luminosity and pile-up conditions, so the thresholds are constantly updated to keep the rate under 100 kHz. Different luminosity settings (*columns*) are available, each of them including different sets of seeds (or even the same seed with different thresholds), so the menu can be adapted to the LHC running conditions during each fill. Additionally, to reduce the rate coming from a particular seed, this seed could be *prescaled*, so only a predefined fraction of events that satisfy this seed's selection will actually trigger.

When designing a new seed to be included in the L1 menu, it's very important that the rate produced by it fits within the allocated rate budget defined by the different physics groups, trying to achieve an efficiency on the signal as high as possible, so the desired events are kept and move along the rest of the acquisition system.

After one of these selection algorithms is fulfilled, the μGT sends a L1 trigger signal to all subdetectors. The information collected by them is then packed into an *event*, which is assigned an *event number*. Then, the event continues through the next steps of the online processing before getting stored and analysed.

Even if the L1 trigger is a fully hardware-based system, every stage (from the trigger primitive generation to the Global Trigger) has its own software *emulator*, a C++ code that replicates the firmware implementation with a very high accuracy. These emulators are used to monitor the correct functioning of the trigger during data-taking and to evaluate the trigger algorithms over simulated samples.

The High Level Trigger

The High Level Trigger (HLT) is the second step of the CMS trigger. It is carried out by a processor farm able to reduce the 100 kHz coming out from the L1 trigger to 1-2 kHz, a bandwidth compatible with the data acquisition capabilities, with a processing time of hundreds of ms. The full detector data is taken as input, performing a reconstruction very close to offline. Some algorithms are even shared between HLT and offline reconstructions, such as Particle Flow (see Section 3); others are a simplified version of the offline algorithm, as is the case of the tracking.

HLT candidates are seeded by L1 objects and reconstructed around them. Their

reconstruction is performed in a sequence of steps, mainly reconstruction modules and selection filters, groups into *paths*. The full HLT selection, or HLT *menu*, consists of hundreds of paths. Some of them are used for physics; others, the *monitoring* paths, for efficiency computation or background studies. As in the L1 menu, the HLT can have several columns and its trigger paths can or cannot be prescaled.

2.2.4 Data Acquisition and Computing systems

The CMS Data Acquisition System (DAQ) is the system responsible for performing the read-out of all subdetectors after the event selection performed by the L1 trigger, building the events from the fragments delivered by each subdetector, and processing them in a computer farm called the *Event Filter Farm*, where the HLT selection code is run. A selection is done based on the HLT decision, such that any given event is either discarded or sent to storage [34].

The output rate of the HLT system is still in the order of 1 kHz. With the large granularity of the CMS detector, the event size is approximately 1 MB. The combination of both leads to petabytes of data produced per year, something not possible to store nor process in the traditional way (i.e. in the computing resources of the laboratory).

CMS is characterized by being a detector with a very high rate (order of hundreds of Hz) and producing big amounts of data (given that each event has an average size of 1 MB). The combination of both leads to petabytes of data produced per year, something not storable in the traditional way (i.e. in the computing resources of the laboratory). Instead, CMS makes use of the World-wide LHC Computing Grid (WLCG) [50, 51], a network created between more than 100 institutes around the world that helps spreading the work and resources necessary to store and analyse the data.

The CMS offline computing system [52] is arranged in four tiers, geographically distributed, in order to benefit from the resources and expertise from several institutes and achieve redundancy amongst multiple centres. First, a single Tier-0 centre at CERN accepts data from the CMS DAQ system, performing a prompt reconstruction of the raw data and creating *datasets* by grouping events with particular physics objects or conditions. The Tier-0 distributes raw and processed data to the seven Tier-1 centers, located in several CMS collaborating countries, such as the Tier-1 centre at PIC (Barcelona, Spain). These centres perform reconstruction, calibration and other data-intensive tasks. They provide the datasets to numerous Tier-2 centres, smaller but with substantial CPU resources, providing capacity for analysis, calibration activities and the production of simulated Monte Carlo (MC) events. CIEMAT (Madrid, Spain) is one of these Tier-2 centres. Finally, Tier-3 centres provide interactive resources for local groups.

The CMS Collaboration uses various data formats, depending on the level of complexity and information required. The RAW format, coming from the Tier-0, is the one that includes all the event information, requiring the most size per event. Then, Tier-1 and Tier-2 centres produce reconstructed (and re-reconstructed)

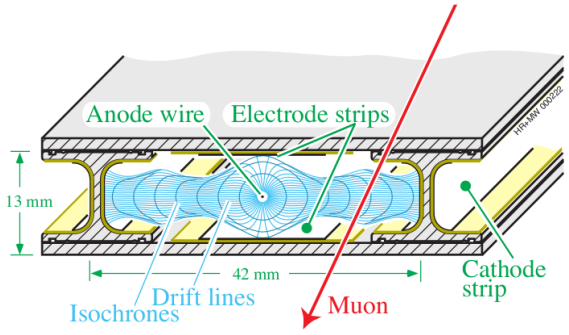


Figure 2.14: Sketch of a drift cell showing drift lines and isochrones. Extracted from [34].

formats from the RAW information. Finally, the data format that contains all the needed information for performing CMS analysis with all reconstructed objects but an smaller size is the Analysis Object Data (AOD). Similar formats reduced in size are miniAOD and nanoAOD, both used in the analyses presented in this thesis. In all cases, data is stored in ROOT [53] files, which contain all the event object information and kinematics. A ROOT file has a tree-like structure with the variables as branches.

A common software framework, the CMS Software (CMSSW), is used in the collaboration for the different analyses. It is an object-oriented structure based on C++ and python, constantly evolving during the different data-taking periods. It contains all services needed to perform simulation, calibration, alignment and reconstruction and the code run in the HLT.

2.3 The CMS Drift Tube system

The CMS Drift Tube (DT) system is the subsystem in charge of muon trigger, reconstruction, and identification in the barrel wheels, covering $|\eta| < 1.2$. DT chambers are characterized by having a very good single-point position resolution (80-120 μm) and an adequate timing resolution. The basic element of these chambers is the drift cell, depicted in Fig. 2.14. These cells have a size of 42×13 cm^2 and a length of approximately 2.4 m and are filled with a gas mixture of argon and carbon dioxide (85%/15%). Inside the DT cell there is a central wire at a high voltage of 3600 V, two electrodes (cathodes) on the sides at -1800 V, and two electrodes above and below the wires at +1200 V (strips). When a charged particle (a muon) crosses the cell, it ionizes the gas creating electrons that drift towards the wire with a constant velocity of about 54 μm/ns. Given the cell sizes, the maximum drift time (from one of the walls to the wire) is almost 400 ns. Once the drift time is known, the position can be obtained by its relation with the drift velocity,

$$x_{\text{drift}} = v_{\text{drift}} \cdot t_{\text{drift}}. \tag{2.15}$$

Fig. 2.15 shows an sketch of one of the five wheels of the CMS barrel. Each wheel

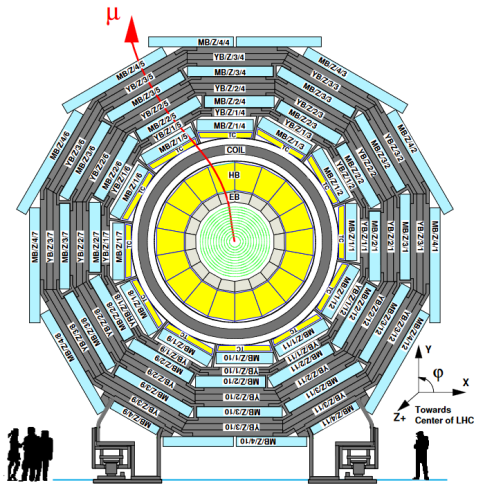


Figure 2.15: Layout of the CMS barrel muon DT chambers in one of the five wheels. Extracted from [34].

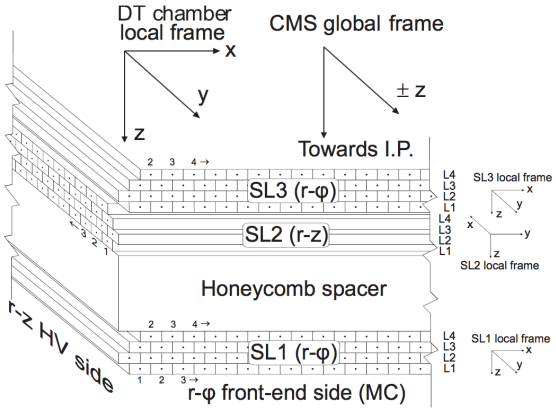


Figure 2.16: Schematic view of a DT chamber. Extracted from [54].

is divided in 12 *sectors* in azimuthal angle. Each sector consists of four *stations*, labelled from MB1 to MB4 (except sectors 4 and 10, which have two MB4 stations each). Each chamber includes three *superlayers* (SL), described in Fig. 2.16, from SL1 to SL3. SL1 and SL3 are oriented in the ϕ direction, while the second one is orthogonal to them and is oriented in the z direction. In the MB4 stations, only the two ϕ SL are available. Each SL includes four *layers* staggered by half a cell. This layout helps to resolve if a muon crossed a DT cell by the left or the right of the wire (also known as *laterality*). All this redundancy helps improving the efficiency and resolution of the system, making it able to cope with situations of dead zones in the detector, such as the space between wheels, electromagnetic showers produced in the chambers by very high p_T muons or the inefficiencies appearing in the DT cells after radiation damage (see Section 7.2).

2.3.1 The DT trigger system

The DT trigger system is a structured, robust, and redundant system devoted to the fast detection and reconstruction of the muons crossing the DT chambers. It is able to measure the position and trajectory or the crossing particles and identify the bunch crossing from which they were originated. It is implemented using dedicated electronics located both on-detector (in the so-called *minicrates*) and in the service cavern.

Inside the minicrates, the DT trigger chain starts at the superlayer level in the *Bunch and Track Identifier* (BTI), which combines the signals from the wires and generates a trigger at a fixed time after the passage of the muon. The association of hits is based on a meantimer technique [55], which uses the fact that there is a relation between the drift times of any three adjacent layers. From the associated hits, the BTI can obtain the BX, position (with a resolution of 1.4 mm), and direction (with a resolution better than 60 mrad) of the track segments [34]. After the BTI, the *Track Correlator* (TRACO) attempts to correlate the track segments measured in each r-ϕ superlayer, enhancing the parameter resolutions. Finally, the *Trigger Server* (TS) board is composed by two subsystems, one devoted to process the output of the TRACO for the r-ϕ segments, and another to process the segments coming from the BTI for the r-z superlayer. In both cases, the TS selects the segments with the highest quality (i.e. higher number of hits used) and transverse momentum. The TS obtains a position resolution of 0.8 mm and a direction resolution of 4 mrad, with an efficiency over 90% [56].

The output from the TS is transmitted via copper-to-optical-fiber translators (CuOF) to the service cavern, where the *TwinMux* systems, running in Xilinx Virtex 7 *Field Programmable Gate Arrays* (FPGA) combines the information from the DT and RPC systems, generating the final trigger primitives that will be sent to the Barrel Muon Track Finder.

2.4 The Phase-2 upgrade of the CMS experiment

During HL-LHC, the accelerator will potentially deliver up to 7.5×10^{34} cm^{-2}s^{-1} and up to 4000 fb^{-1} (see Section 2.1.2). In this scenario, the interaction rate, occupancy, PU levels, and radiation ageing will increase beyond the capabilities of the existing detector and trigger technologies. In the CMS *Phase-2* upgrade program [57], the goal is to maintain the excellent performance of the present *Phase-1* detector or even improve it wherever possible under these new challenging conditions throughout the whole operation of HL-LHC.

Two major concerns need to be addressed in the required detector upgrade. The first one is the radiation damage, that would reduce the performance of some subdetectors. The second, the higher PU and particle rate require a better precision on the detector to be able to discriminate particles coming from the interactions. A higher granularity is then needed, specifically in the endcaps.

Some details on the major detector, trigger, and DAQ elements needing upgrading are given in the following

Tracker The silicon tracker will suffer from significant radiation damage by LS3, so it will be completely replaced. To maintain the track reconstruction performance with the higher PU levels, the granularity of the pixel detector and the outer tracker will be increased by almost a factor 4. The new outer tracker will be lighter, providing a better resolution and new pixel disks will be added in the forward regions to extend the acceptance close to $|\eta| = 4$.

MIP Timing Detector The MIP Timing Detector (MTD) is a new detector planned for CMS during HL-LHC [58]. This detector will allow CMS to measure precisely the production time of minimum ionizing particles (MIP), which could help to assign charged tracks to the correct interaction vertices in the new PU 200 scenario. The detector will be placed as a thin layer between the Tracker and the calorimeters, divided into the barrel ($|\eta| < 1.5$) and two endcap sections covering up to $|\eta| = 3$. In the barrel, the technology used will be crystal scintillators read out by Silicon Photomultipliers (SiPM), while in the endcap, Low-Gain Avalanche Detectors (LGAD) will be used.

Calorimeter endcaps The electromagnetic and hadronic endcap calorimeters will also suffer significant radiation damage by LS3, so they must be replaced. The new replacement is called the *High-Granularity Calorimeter* (HGCAL) [59], a calorimeter with electromagnetic and hadronic sections and excellent longitudinal and transverse segmentation, providing three dimensional images of showers and a timing measurement.

Muon detectors The current barrel muon detectors will be kept for HL-LHC. Their electronics, however, will be upgraded in order to cope with the increased rate and luminosity. A more detailed description of the DT upgrade is shown in Section 2.4.1. In the endcap, the current muon system consists of four stations of CSCs and four stations of RPCs. To maintain the trigger acceptance in this region during Phase-2, additional chambers that make use of new detector technologies with high rate capability are going to be added [60], as shown in Fig. 2.17. Two new improved RPC (iRPC) stations will be installed along CSC stations 3 and 4, extending the RPC pseudorapidity coverage to $|\eta| < 2.4$. In the $1.6 < |\eta| < 2.8$, new *Gas Electron Multiplier* (GEM) chambers will be installed. These new chambers provide a precise measurement of the muon bending angle and help controlling the trigger rate while increasing the trigger efficiency and the operational resilience of the system. Three new GEM superchambers (that group two or six GEM chambers) will be installed, named as GE1/1, GE2/1 and ME0.

Trigger The latency of the present L1 trigger is limited to 3.4 μs by the tracker readout. During Phase-2, for the first time the tracker information will also be used

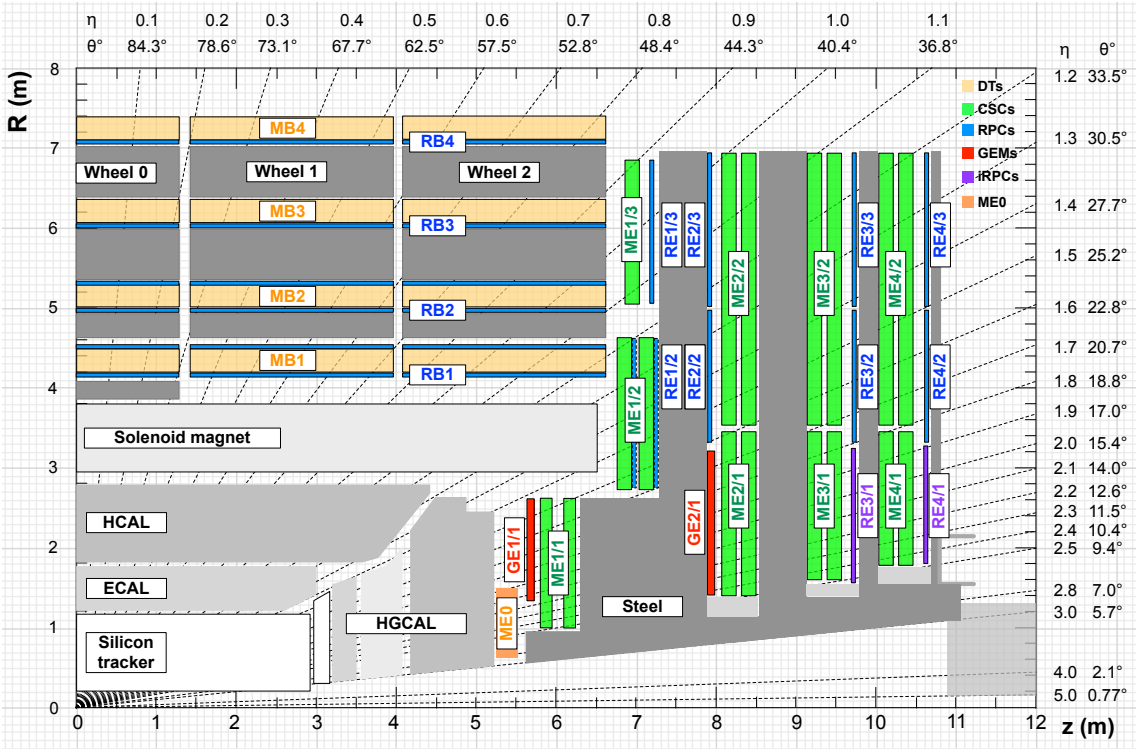


Figure 2.17: Longitudinal view of a quadrant of the CMS detector after the installation of the HGCal and the new muon chambers. Extracted from [61].

at trigger level. The tracker trigger will perform pattern recognition to reconstruct the tracks of primary particles with $p_T > 2$ GeV, discarding as many low energy tracks as possible. The reconstructed tracks will then be matched with the muon and calorimeter information, increasing latency to 12.5 μs. This will require upgrades of the readout electronics in some of the existing sub-detectors: barrel calorimeter, DTs (further described below) and CSCs. New detector electronics must be designed to sustain larger L1 rate, of 500(750) kHz for PU 140(200) scenario, keeping typical values of p_t and energy thresholds as in a Phase-1 trigger menu.

Data Acquisition and Trigger Control The DAQ system will be upgraded to implement the increasing of bandwidth and computing power needed to accommodate the larger event size and L1-trigger rate. Compared to Phase-1, bandwidth and computing power will increase by factors of about 10(15) and 15(30) for operation at PU 140(200).

Apart from these changes, some upgrades on the barrel calorimeters electronics, protections against beam background, the luminosity measurement system and software and computing for online and offline analysis are also planned.

2.4.1 Upgrade on the DT Trigger system for Phase-2

The DT chambers located in the barrel region of CMS are responsible of muon detection and precise momentum measurement over a wide range of energies. They provide identification, reconstruction and trigger on muons. During the Phase-2 upgrade [60,61], the goal is to maintain or improve the present system performance for trigger and offline reconstruction at the HL-LHC background rates and trigger conditions (500-750 kHz L1 trigger rate and 12.5 μs latency).

Present DT detectors will remain for HL-LHC operation. Even if the most exposed chambers will suffer from detector ageing at the end of HL-LHC due to radiation damage [60], the loss of performance is expected to be minor after this period (see Section 7.2). However, the DT on-detector electronics must be replaced since they are not prepared for the HL-LHC radiation conditions and can't cope with the expected increase of L1 rate foreseen during Phase-2 operation.

In the upgraded DT architecture, only the time digitization of the chamber signals will stay on-detector. This will be performed by the new *On-Board electronics for DT* boards (OBDT). These boards receive signals from each anode wire and assigns them a digital timestamp. The trigger primitive generation, presently performed also on-detector, will be moved from the experimental cavern to the service cavern. In the service cavern, new boards will be installed featuring the latest commercial FPGAs. A new trigger algorithm will be running in these boards, profiting from the full chamber information with the best DT time resolution (few ns), so it can generate DT trigger primitives with a more precise and robust bunch crossing identification, better spatial resolution (close to offline level), and better fakes rejection. This new algorithm is described in Chapter 7.

Once the DT trigger primitives have been generated, they are sent to the BMTF, where standalone muons are built. Standalone muons and L1 tracks (from the tracker) are sent to the Global Muon Trigger. The output from the Global Muon Trigger is combined with the information from the other trigger subsystems in the Global Trigger. Finally, the Global Trigger calculates a trigger decision based on the L1 menu.

Chapter 3

Particle identification and reconstruction

The CMS experiment was designed such that particles of different types, whether they are produced at the primary vertex or originate from subsequent particle decay or hadronization, leave a signature in the detector, as shown in Fig. 3.1. Muons lose an small fraction of its energy in the inner part of the detector, so, as they are charged particles, are detected in both the inner tracker and the muon chambers. Electrons and photons are both absorbed in the ECAL (where their energy can be measured), although only the electrons leave a track in the inner tracker. Hadrons may deposit a fraction of their energy in the ECAL, although they deposit a substantial amount in the HCAL. The trajectory of the charged hadrons can be measured in the tracker. Neutrinos hardly interact with the CMS detector and they escape it. However, they manifest their presence by the energy imbalance of the event in the transverse plane or missing transverse energy.

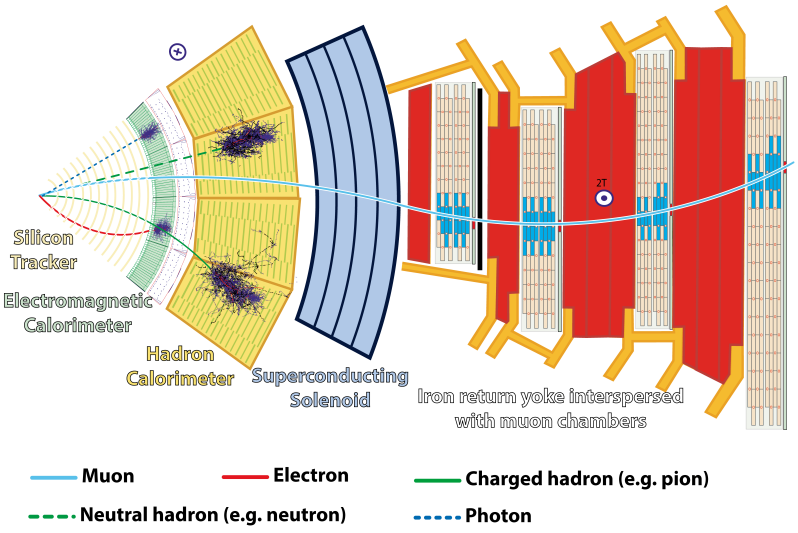


Figure 3.1: Slice of the CMS detector in the transverse view with the signals left by different particles. Extracted from [62].

In order to reconstruct the physics objects, the information coming from all subdetectors is combined by the Particle Flow (PF) [63] algorithm. A brief description of this algorithm is included in Section 3.1. The reconstruction of final state particles using this algorithm is described in Section 3.2.

3.1 The Particle Flow algorithm

The PF algorithm is the general reconstruction algorithm in CMS. This algorithm combines the information coming from the different CMS subdetectors to build, for each event, a set of stable particle candidates (e.g. muons, electrons, hadrons, photons, etc.), assigning them optimal estimates of their energy and direction. The main inputs to the algorithm are *tracks* and *clusters*.

Charged-particle tracks are reconstructed using a combinatorial track finder based on a Kalman Filtering approach [64] in three steps. First, an initial seed generation is performed with a few hits compatible with a charged-particle trajectory. Then, hits along the charged-particle trajectory are gathered before performing a final fitting to determined the charged-particle properties: origin, transverse momentum, and direction. This combinatorial track finder is applied in several successive iterations, keeping a moderate efficiency but low fake rate. At each step, a quality criteria is applied to the track seeds, adapted to the track's p_T, η, and number of hits. The hits associated with the selected tracks are masked so they are not considered at later stages, reducing the combinatorial complexity. From the reconstructed tracks, primary vertex corresponding to several interactions can be identified. Among them, the hard interaction vertex is selected as the one with the highest quadratic sum of the p_T of its corresponding tracks.

Clusters are constructed from energy depositions in the calorimeters, separately for each sub-detector (preshower layers, ECAL barrel and endcaps, and HCAL barrel and endcaps). Cluster seeds are identified as cells with an energy larger than a given seed threshold, and larger than the energy from their neighbours. Then, seeds are grown by aggregating cells with a corner in common with the seed cell that have an energy higher than a noise level.

Tracks and clusters are finally associated with a *link algorithm*, producing *PF blocks* that can be identified with the different particles.

3.2 Final state particle reconstruction

3.2.1 Muons

Muons are reconstructed independently by the silicon tracker and the muon chambers [65]. The former gives a precise measurement of the momentum of these muons, while the latter identifies muons with high efficiency over the full detector acceptance and provides an accurate measurement of the momentum of high p_T muons.

Muon objects are reconstructed as three different muon types:

- Hits coming from the DT and CSC systems are identified and clustered in ordered to build straight-line track *segments*. This is also known as *local* reconstruction. Then, local segments and hits coming from the DT, CSC and RPC can be combined into **standalone muons**.
- Standalone muon tracks can be matched to tracks in the inner tracker and, if they are compatible when propagated onto a common surface, a **global muon** is constructed.
- Tracks reconstructed in the inner tracker can be extrapolated to the muon system. If at least one muon segment matches the extrapolation, the track is identified as a **tracker muon** track.

Thanks to the high efficiency of the inner track and muon segment reconstruction, about 99% of the muons produced with the geometrical acceptance of the muon system are reconstructed as global or tracker muons, and very often as both. If this is the case, they are merged into a single candidate.

Charged hadrons may be misreconstructed as muons if, for example, some of the hadron shower remnants reach the muon system (*punch-through*). This is taken into account by the PF algorithm in order to build PF muons by selecting some properties from global and tracker muons. Different levels of requirements (e.g. track fit χ^2, number of hits per track or degree of matching between tracker and standalone tracks) result in three muon identification working points, *loose*, *medium* and *tight*, with increasing purity and decreasing efficiency. The information from the different subdetectors is used to define the PF muon isolation:

$$I^l_{rel} = \left(\sum p_T^{\text{charged}} + \max \left[0, \sum p_T^{\text{neutral-had}} + \sum p_T^{\gamma} - \frac{1}{2} \sum p_T^{\text{PU}} \right] \right) / p^l_T, \quad (3.1)$$

where $\sum p_T^{\text{charged}}$, $\sum p_T^{\text{neutral-had}}$, $\sum p_T^{\gamma}$, and $\sum p_T^{\text{PU}}$ are the scalar sums of the transverse momenta of charged hadrons originating from the primary vertex, neutral hadrons, photons, and charged hadrons not originating from the primary vertex respectively inside a $\Delta R < 0.4$ cone around the muon.

3.2.2 Electrons

Electrons are reconstructed using the information from the tracker and the ECAL [66]. Because of the material of the tracker, most electrons emit a sizeable fraction of their energy as bremsstrahlung photons before entering the ECAL, so several ECAL clusters are produced. The energy of the electron and all possible produced photons is collected by grouping compatible clusters in a small window in η and an extended window in ϕ (to account for the magnetic field) around the electron direction into *superclusters*. Electron tracks are then reconstructed from the superclusters and tracker tracks refitted using a Gaussian-Sum Filter algorithm [67].

Several strategies are used in order to discriminate prompt electrons from background sources. Electron PF isolation is computed using (3.1), although the cone around the electron is defined with a $\Delta R < 0.3$. Three isolation components are considered separately, the scalar sums of transverse momenta of charged hadrons coming from the primary vertex, neutral hadrons and photons.

Electron identification is performed with a multivariate approach (MVA) updated and improved for Run 2 analyses [68]. The discriminator used is based on a Boosted Decision Tree (BDT) that considers as input variables related to the shower shape, the track quality, the track-cluster matching or the fraction of momentum lost due to brehmsstrahlung. To increase the flexibility at analysis level, two different BDTs are trained, one with the previous input variables and another one with those variables and the three isolation components as input. Separate trainings are performed in bins of E_T and η. The split in E_T is at 10 GeV, so two regions where the background compositions are very different are defined. Two splits are defined in η: at $|\eta| < 0.8$, where the tracker material budget starts to steeply increase, and in the barrel-endcap transition. Fig. 3.2 shows the performance of the BDTs for electrons with $E_T > 20$ GeV in the barrel and the endcap. Three working points are defined for the two BDTs (with and without the isolation variables): *wp90* and *wp80*, associated to a 90 and 80% signal efficiency respectively, and a *loose* WP, with 98% signal efficiency.

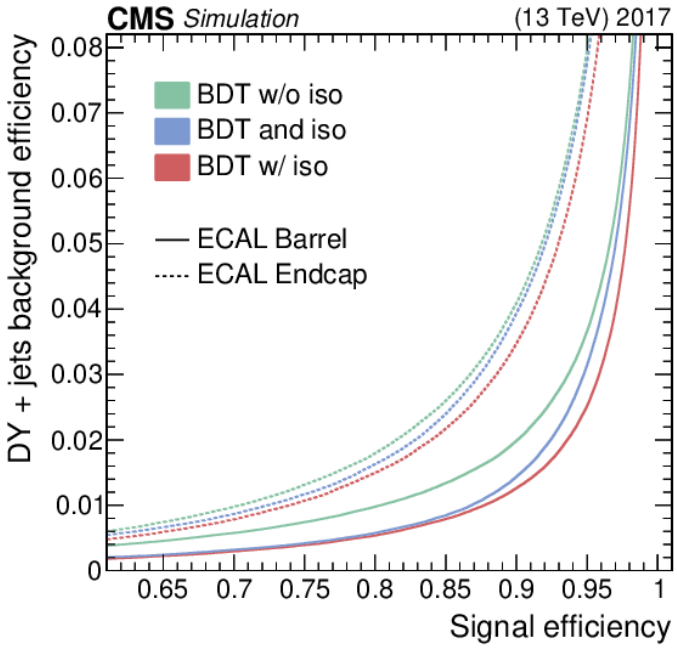


Figure 3.2: Performance of the electron MVA-based identification algorithm with (red) and without (green) the isolation variables, compared to an optimized sequential selection using the BDT without the isolations followed by a selection requirement on the combined isolation (blue). Extracted from [68].

3.2.3 Jets

Jets are originated by the hadronic showers produced from quarks and gluons. In the experiment, they can be reconstructed by clustering charged and neutral hadrons with the anti-k_T algorithm [69]. This clustering is done by grouping PF candidates close to each other in order to build a cone-shaped jet (*PF jet*) whose typical size is driven by the distance parameter R. Regular jets, or AK4 jets, are reconstructed with a distance parameter of 0.4. In boosted topologies, where the particle decaying into two jets has very high momentum, both AK4 jets can overlap. Therefore, a bigger object with distance parameter $R = 0.8$ can be built, corresponding to an AK8 jet.

To reduce jets poorly reconstructed or affected by instrumental noise in the calorimeters, some identifications requirements are applied to them [70]. Some of these requirements are the fraction of charged and neutral hadrons, electrons, photons and muons or the multiplicity of charged and neutral hadrons. Three working points are considered, *loose* and *tight* (designed to remove jets originating from calorimetric noise) and *tight lepton veto*, that rejects background coming from misreconstructed electron and muon candidates. The performance of the tight working point is shown in Fig. 3.3.

PU jets are jets not coming from the primary vertex, so they do not come from the interaction under study. To reduce the amount of them, several mitigation techniques exist to identify and reject such jets [71]. The most commonly used is the Charged Hadron Substraction (CHS), which removes all charged constituents associated to PU vertices before starting the jet clustering procedure. After applying the CHS mitigation technique, the remaining PU jets are further reduced by applying a multivariate approach, the PU jet ID. It consists of a BDT trained with 12-15 variables that characterize the jet and the event. The performance of this algorithm is shown in Fig. 3.4. Three working points are defined for this discriminant, the *tight* and medium working points, with 80 and 90% efficiency on quark jets respectively, and the loose working point, 99% efficient for quark jets in $|\eta| < 2.5$ and 95% efficient for quark jets in $|\eta| > 2.5$.

In order to identify jets coming from the hadronisation of b quarks (b-jets), dedicated algorithms have been developed. Among those we can find the DeepCSV [73] and DeepJet [74] algorithms. Both algorithms are DNN-based approaches, although the latter shows a better performance (shown in Fig. 3.5) thanks to including the full information of all jet constituents, neutral and charged particles, secondary vertices and global event variables. Three working points are defined, *loose*, *medium* and *tight*, corresponding to around 10, 1 and 0.1 % light jet misidentification probability respectively.

3.2.4 Tau leptons

The τ lepton is the heaviest lepton, with a mass of $m_\tau = 1.777$ GeV. This mass is large enough to make it decay almost immediately after its production (with an

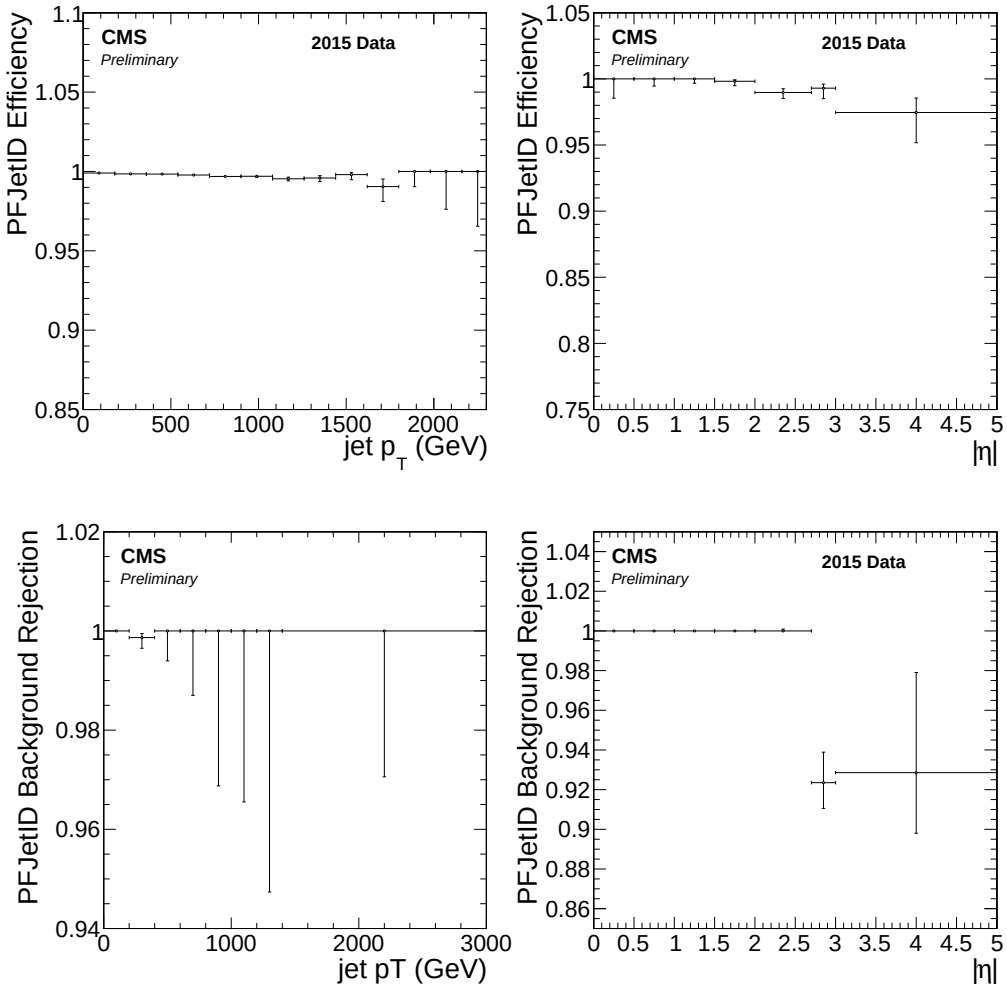


Figure 3.3: Performance of the CMS PF jet identification algorithm for the tight working point. Plots from the top row show the efficiency of correctly identifying a PF jet, while the ones in the bottom row show the noise jet background rejection rate. The two plots from the left are shown as a function of the jet p_T for central jets ($|\eta| < 0.5$), while right plots are shown as a function of the jet $|\eta|$ for jets with p_t between 30 and 100 GeV. Extracted from [70].

average lifetime of 3×10^{-13} s) into a lighter lepton and two neutrinos (leptonic decay) or into hadrons and a τ neutrino (hadronic decay). In the leptonic decays, the τ is reconstructed as the correspondent electron or muon. In the hadronic decays, however, the τ lepton usually decays to charged hadrons (mostly $\pi^{\pm}$), denoted as *prongs*, and neutral hadrons (π_0). Therefore, the hadronic τ (τ_h) has to be reconstructed as a composite object from its PF components. In both decays, neutrinos remain undetected, so it's not possible to measure exactly the momentum of the decaying τ. A summary of the existing τ decay modes and their branching fractions in shown in Table 3.1.

For the τ_h reconstruction, the *hadrons-plus-strips* (HPS) algorithm [76] is used.

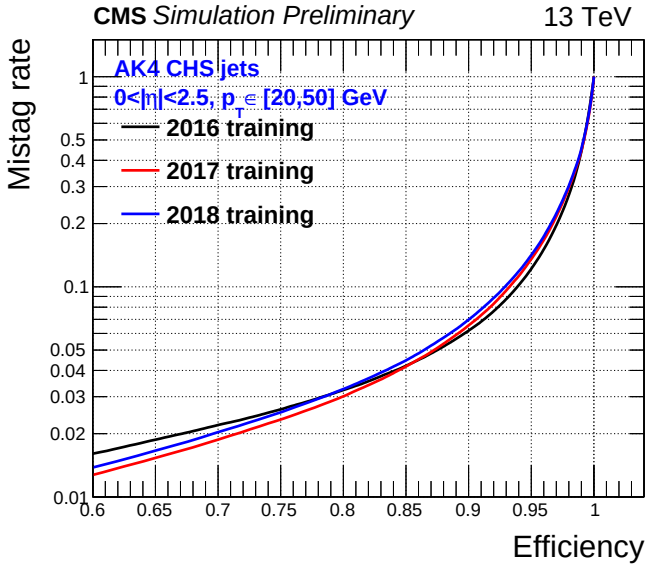


Figure 3.4: PU Jet ID mistag rate as a function of its efficiency for jets with $|\eta| < 2.5$ and $p_T \in [20, 50]$ GeV. Performance is measured by using simulated samples of Drell-Yan events. Extracted from [72]

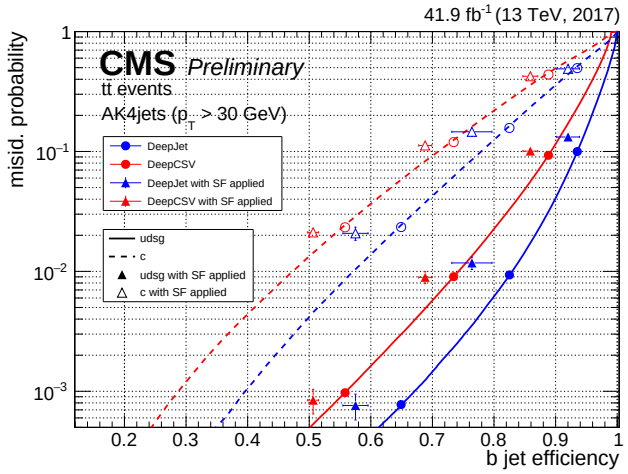


Figure 3.5: Performance of the CMS DeepJet (blue) and DeepCSV (red) algorithms. The plot represents the probability of misidentifying jets coming from light quarks or gluons (solid lines) or c jets (dashed lines) with respect to the efficiency of correctly identifying jets coming from b quarks. Results are obtained using jets with $|\eta| < 2.5$ and $p_T > 30$ GeV coming from the decay of simulated $t\bar{t}$ events. Loose, medium and tight working points for both algorithms are shown in the plot as circles when obtained directly from the simulation or as triangles after applying the scale factors obtained through data-simulation studies. Extracted from [75].

Decay Mode	BR (%)
Leptonic decays	
$\tau^- \to e^- \bar{\nu}_e \nu_\tau$	17.8
$\tau^- \to \mu^- \bar{\nu}_\mu \nu_\tau$	17.4
Hadronic decays	
$\tau^- \to \mathrm{h}^- \nu_\tau$	11.5
$\tau^- \to \mathrm{h}^- \pi^0 \nu_\tau$	25.9
$\tau^- \to \mathrm{h}^- \pi^0 \pi^0 \nu_\tau$	9.5
$\tau^- \to \mathrm{h}^- \mathrm{h}^+ \mathrm{h}^- \nu_\tau$	9.8
$\tau^- \to \mathrm{h}^- \mathrm{h}^+ \mathrm{h}^- \pi^0 \nu_\tau$	4.8
Other	3.3

Table 3.1: Decay modes of the τ lepton and their branching fractions (BR) in % [4]. Charged hadrons are denoted by $\mathrm{h}^\pm$.

This algorithm is seeded by the AK4 PF jets reconstructed as in Section 3.2.3 and its constituents. Regarding the π_0, they promptly decay into a pair of photons, which can then convert into e^+e^- pairs due to the amount of tracker material. Electrons and photon candidates in a certain region $\Delta\phi \times \Delta\eta$ are clustered together, building *strips*, so the energy of the π_0 is measured. Both charged hadrons and strips inside the PF jet are then combined by the HPS algorithm into all the possible combinations of hadrons for the following decay modes: $\mathrm{h}^\pm$, $\mathrm{h}^\pm\pi^0$, $\mathrm{h}^\pm\pi^0\pi^0$ and $\mathrm{h}^\pm\mathrm{h}^\mp\mathrm{h}^\pm$ (plus the additional ν_τ).

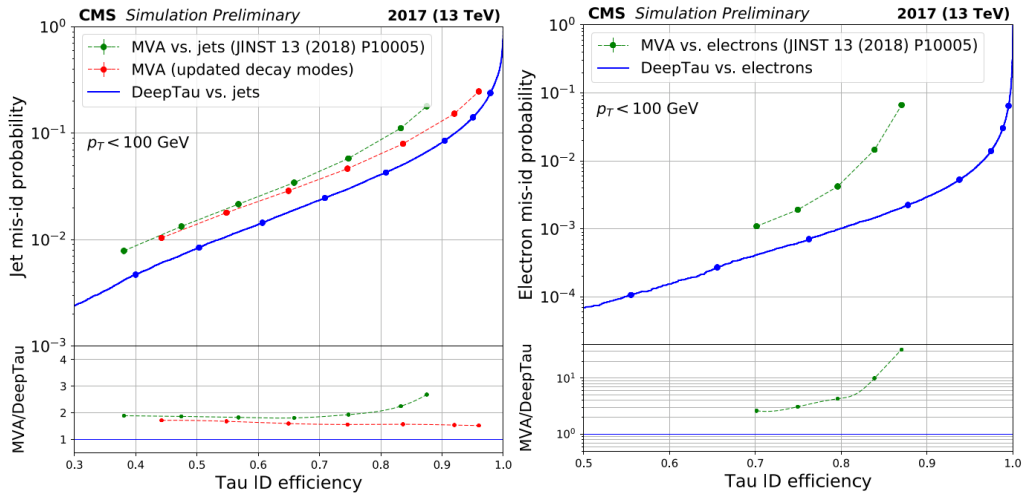
In order to increase the purity of the τ_h reconstruction, an identification method, DeepTau [77], has been implemented to discriminate real τ_h from quark and gluon jets, electrons and muons. This algorithm follows a multi-class DNN approach that combines information from high-level reconstructed τ features, the decay mode extracted from the HPS algorithm and low-level information used to build the PF candidates. Three different discriminators, built to identify τ_h against jets, electrons and muons, respectively, are defined within the DeepTau algorithm, each with its own associated working points:

- *DeepTauVSjet*, with working points VVVLoose (very-very-very-loose), VVLoose, VLoose, Loose, Medium, Tight, VTight and VVTight.

- *DeepTauVSe*, with working points VVVLoose, VVLoose, VLoose, Loose, Medium, Tight, VTight and VVTight.

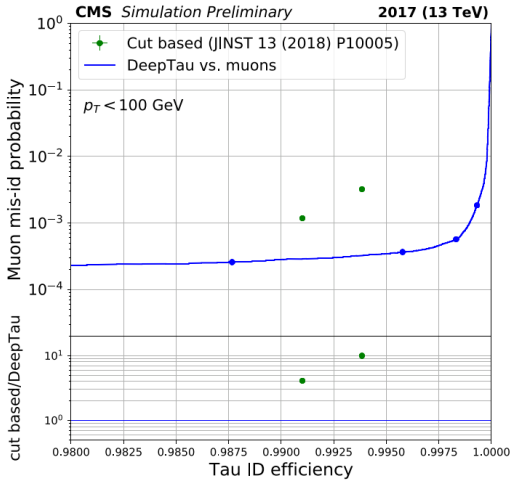
- *DeepTauVSmu*, with working points VLoose, Loose, Medium and Tight.

The performance of the three discriminators is shown in Fig. 3.6, where a substantial improvement with respect to older algorithms can be seen.



(a) DeepTauVSjet

(b) DeepTauVSe



(c) DeepTauVSmu

Figure 3.6: Performance of the CMS DeepTau algorithm. Plots represent the probability of misidentify jets (a), electrons (b) or muons (c) as τ_h with respect to the efficiency of correctly identifying the τ_h for the DeepTauVSjet (a), DeepTauVse (b) and DeepTauVSmu (c) discriminators. Working points of the discriminators are indicated as dots. The performance of older discriminators used before the appearance of the DeepTau algorithm are also shown for the sake of comparison. Extracted from [77].

3.2.5 Missing transverse energy

When two protons collide, the amount of momentum carried by their constituents is unknown. However, their momentum in the transverse plane can be considered as negligible. Therefore, by conservation of momentum, the total momentum in the transverse plane in the initial and in the final state should also be negligible. Then, particles that can not be measured in the detector (neutrinos, BSM particles,...) can

be observed by an energy imbalance in the transverse plane. The *missing transverse momentum* is then defined as

$$\vec{p}_T{}^{\text{miss}} = -\sum \vec{p}_T, \tag{3.2}$$

where the sum extends to all the PF particles available in each event. The modulus of $\vec{p}_T{}^{\text{miss}}$ can be denoted as p_T^{miss}.

To perform a precise measurement of this quantity, the CMS detector covers almost the whole solid angle, so most of the particles that can be measured go through the detector. But its determination is also affected by inefficiencies of tracking and clustering algorithms or from the actual detector (for example, calorimeter noise). Anomalous events that could have a large missing energy coming from detector issues are removed by using specific filters prior to any analysis [78].

Part I

Search for Higgs boson pair production at the CMS experiment of the LHC

Chapter 4

Analysis setup

This chapter describes the analysis strategy followed in the search for Higgs pair production via gluon-gluon fusion and vector boson fusion and decaying into two b quarks and two τ leptons. The aim of the analysis is to provide a value of the production cross section and study deviations from the Higgs boson couplings involved. The statistics collected during the LHC Run 2 are not enough to provide accurate measurements of these quantities. Therefore, upper limits on the cross section of the HH produced via ggF and VBF and exclusion ranges on the relevant couplings are established, which still are very relevant results. The statistical model considered and the results obtained are shown in Chapter 5.

Being the τ lepton an unstable particle, several final states can be reconstructed. As described in Section 3.2.4, the τ lepton mainly decays into hadrons, with a total branching fraction of 64.8%. Alternatively, it can decay leptonically into an electron, an electron neutrino and a τ neutrino or into a muon, a muon neutrino and a τ neutrino, with similar branching fractions of 17.8 and 17.4%, respectively. These three decays are denoted as τ_h, τ_e, and τ_μ. The combination of τ decays leads to six possible decay pairs. However, only three are considered in the analysis, the ones where at least one τ decays hadronically: the fully hadronic final state, $\tau_h\tau_h$, and the semi-leptonic ones, $\tau_e\tau_h$ and $\tau_\mu\tau_h$, covering 87.6% of the total branching fraction. Additionally, at least two jets are also present in the final state.

HH signal contribution is estimated using simulated samples produced via ggF and VBF. In order to study possible BSM scenarios, several samples with different values of the Higgs boson couplings involved are produced, so they can be reweighed to generate distributions for other coupling values. The modelling of the signal processes and the reweighing methods are described in Section 4.1.1.

Several processes have a final state that is the same or can mimic the final state of the HH $\rightarrow$ bb$\tau\tau$ signal, contributing as background sources in the analysis. Two types of backgrounds are usually distinguished: irreducible backgrounds, whose final state is the same as in the signal, and reducible backgrounds, whose final state mimics the one from the signal due to the misidentification of one or more objects. The modelling of these backgrounds is described in Section 4.1.2, although a minor description is given in the following.

Two processes constitute the major irreducible backgrounds. The first one is the production of pair of leptons through Drell-Yan (DY) processes ($Z/\gamma^* \to ll$) in association with jets. When the Z/γ^* decays into two τ, the two τ can be associated to the decay products of one of the Higgs bosons and a pair of jets can be selected as the ones that come from the other Higgs decay, so the final state mimics the one from the signal. If the Z/γ^* decays into two e or two μ, if at least one lepton is misidentified as a τ_h, the final selection could be satisfied as well.

The other major irreducible background source in the pair production of top quarks ($t\bar{t}$). The top quark decays almost exclusively to a b quark and a W boson, which in turn can decay into a lepton and a neutrino or to a pair of quarks. This results in three possible $t\bar{t}$ decay final states: the di-leptonic final state, where the two W bosons decay into a lepton and a neutrino; the semi-leptonic final state, where one W boson decays leptonically and the other one into two quarks; and the fully-hadronic final-state, where both W bosons decay into quarks. The first two final states are the ones that give the largest contribution to the background, since the di-leptonic decay has the same final state as the signal and the semi-leptonic decay a similar final state (if one jet is misidentified as a lepton), but a higher branching ratio than the fully-leptonic decay.

The largest reducible background comes from QCD multijet events, where two jets are misidentified as leptons. This process has a relatively large contribution in the $\tau_h\tau_h$ channel, but its contribution to the $\tau_e\tau_h$ and $\tau_\mu\tau_h$ is smaller due to the stringent requirements set at analysis level to select prompt high p_t isolated leptons from the signal process.

Finally, W+jets events, with one lepton coming from the W decay and one jet misidentified as a τ_h, have a minimal contribution thanks to the b jet requirements. There are other backgrounds whose contributions are even smaller, e.g. single Higgs boson events or di- and tri-vector boson production.

To increase the signal purity, several selections are applied to the objects that could belong to the signal final state. Section 4.2 describes these selections. Section 4.3 includes all the dedicated steps of the analysis: selecting the $H \to \tau\tau$ and $H \to bb$ candidates, the VBF jet selection, and the event categorization. Section 4.4 describes the methods used to estimate and normalize the main backgrounds and Section 4.5 includes the corrections applied to the simulated events used in the analysis. Finally, Section 4.7 describes the multi-class classification strategy devoted to categorizing VBF-like events and Section 4.8 shows the signal extraction strategy used in the analysis.

My personal contributions to this chapter include the study of the VBF topology and the development of the multi-class classification strategy and its use in the VBF signal extraction. I also performed a detailed study of the QCD background estimation and its related uncertainties (detailed in Chapter 5). Additionally, I developed a new analysis framework devoted to replicate the $HH \to bb\tau\tau$ analysis workflow using the latest available samples (produced in the nanoAOD data format). The results obtained with this new framework are in very good agreement with the ones obtained with the framework used for the Run 2 publication. It will continue

to be used during Run 3 in the HH $\rightarrow$ bb$\tau\tau$ analysis and other analyses in the CIEMAT-CMS group.

4.1 Data and simulated samples

This analysis uses the data collected at CMS at $\sqrt{s} = 13$ TeV during 2016, 2017 and 2018, corresponding to integrated luminosities of 36.3, 41.5, and 59.7 fb^{-1} respectively. MC simulated samples, produced with state-of-the-art generators, are used for both signal and background processes.

4.1.1 Signal modelling

The exploration of several BSM scenarios requires the modelling of the signals for several sets of couplings. As only a limited set of simulated samples can be produced, weighting techniques are implemented for both ggF and VBF HH signals, so additional BSM scenarios can be modelled starting only from a few BSM samples.

Regarding ggF, the reweighing is based on the sum of different samples [79]. At LO, the amplitude of the ggF production can be written as

$$\mathcal{A} = \kappa_\lambda \kappa_t T + \kappa_t^2 B, \tag{4.1}$$

where T and B can be associated to the triangle and box diagrams of Fig. 1.7. Therefore, the ggF production cross section can be computed from the square of the amplitude,

$$\sigma(\kappa_\lambda, \kappa_t) \sim |\mathcal{A}|^2 = \kappa_\lambda^2 \kappa_t^2 t + \kappa_t^4 b + \kappa_\lambda \kappa_t^3 i, \tag{4.2}$$

where $t = |T|^2$, $b = |B|^2$ and $i = |TB^* + B^*T|$. Eq. (4.2) can be rewritten in matrix notation form as

$$\sigma(\kappa_\lambda, \kappa_t) = \mathbf{c}(\kappa_\lambda, \kappa_t) \cdot \mathbf{v}, \tag{4.3}$$

where $\mathbf{c}(\kappa_\lambda, \kappa_t) = (\kappa_\lambda^2 \kappa_t^2, \kappa_t^4, \kappa_\lambda \kappa_t^3)$ is the vector of coupling constants and $\mathbf{v} = (t, b, i)$ is the vector of the values of the three components.

Given three particular samples with fixed values of κ_λ and κ_t and cross sections σ_1, σ_2 and σ_3, Eq. (4.3) can be expressed as a linear combination of these three samples:

$$\begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \sigma_3 \end{pmatrix} = \begin{pmatrix} c_1^1 & c_1^2 & c_1^3 \\ c_2^1 & c_2^2 & c_2^3 \\ c_3^1 & c_3^2 & c_3^3 \end{pmatrix} \begin{pmatrix} t \\ b \\ i \end{pmatrix}, \tag{4.4}$$

where (c_i^1, c_i^2, c_i^3) is the vector of coupling constants for each particular sample. In matrix notation form, Eq. (4.4) can be expressed as

$$\boldsymbol{\sigma} = \mathbf{C}\mathbf{v}. \tag{4.5}$$

Inverting this equation and substituting the obtained expression of $\mathbf{v}$, Eq. (4.3) becomes

$$\sigma(\kappa_\lambda, \kappa_t) = \mathbf{c}(\kappa_\lambda, \kappa_t)\ \mathbf{C}^{-1}\ \boldsymbol{\sigma}. \tag{4.6}$$

Eq. (4.6) shows that the ggF production cross section for generic values of κ_λ and κ_t can be obtained just as a linear combination of the cross sections of three given samples with fixed values of κ_λ and κ_t[1]. Similarly, the distribution of a particular variable of a generic ggF samples can be obtained just by combining the same distribution from the three produced ggF samples.

To perform these studies, three ggF samples at NLO were produced using POWHEG with a fixed value of $\kappa_t = 1$ and $\kappa_\lambda = 1$, $\kappa_\lambda = 2.45$ and $\kappa_\lambda = 5$ (whose cross sections are shown in Table 4.1)[2]. Additionally, another sample with $\kappa_\lambda = 0$ was produced to test the performance of the reweighing method. Fig. 4.1 shows a comparison between the HH system mass obtained directly from the sample with $\kappa_\lambda = 0$ and the same distribution obtained by reweighing the histograms from the first three samples. Good agreement between the original and the reweighed distribution is observed, thus validating the reweighing method.

κ_λ	σ (fb)
1	31.05
2.45	13.12
5	91.17

Table 4.1: κ_λ parameter values considered in the ggF HH $\to$ bb$\tau\tau$ simulated samples used and their corresponding cross sections. κ_t takes a value of 1 in all cases. Cross sections are taken from [80], although their values have to be scaled by $\sigma_{\text{NNLO+FTapprox}}/\sigma_{\text{NLO}}$.

These NLO samples are used for both signal extraction and training the machine learning approach described in Section 4.7. To train the neural network used for signal extraction (described in Section 4.8), additional LO samples with larger statistics were generated using MADGRAPH5_AMC@NLO.

A similar procedure as the one used for the ggF production mode is also used for VBF modelling. In this case, the HH production cross section can be written as

$$\begin{aligned}\sigma(\kappa_V, \kappa_{2V}, \kappa_\lambda) &\sim |\kappa_V\kappa_\lambda A + \kappa_V^2 B + \kappa_{2V} C|^2 \\ &= \kappa_V^2\kappa_\lambda^2 a + \kappa_V^4 b + \kappa_{2V}^2 c + \kappa_V^3\kappa_\lambda i_{ab} + \kappa_V\kappa_{2V}\kappa_\lambda i_{ac} + \kappa_V^2\kappa_{2V} i_{bc}, \quad (4.7)\end{aligned}$$

where A, B and C can be associated to the diagrams from Fig. 1.9, $a = |A|^2$, $b = |B|^2$, $c = |C|^2$ and i_{ij} are the inference terms. Thus, the VBF signal can be modelled by summing six components $\mathbf{V} = (a, b, c, i_{ab}, i_{ac}, i_{bc})$, scaled by $\mathbf{K} = (\kappa_V^2\kappa_\lambda^2, \kappa_V^4, \kappa_{2V}^2, \kappa_V^3\kappa_\lambda, \kappa_V\kappa_{2V}\kappa_\lambda, \kappa_V^2\kappa_{2V})$. By defining $\mathbf{M}$ as the 6×6 coefficient matrix and σ as the vector of the cross sections of six different samples, the cross section σ_{target} of a generic sample with given κ_V, κ_{2V} and κ_λ can be obtained as

$$\sigma_{\text{target}} = \mathbf{K}\mathbf{M}^{-1}\boldsymbol{\sigma}. \quad (4.8)$$

[1]This method can only be used if, for the three given samples, $\mathbf{C}$ is invertible.

[2]Note that the dependency on the κ_t coupling modifier has been dropped to increase the accuracy in the production of distributions with different values of κ_λ.

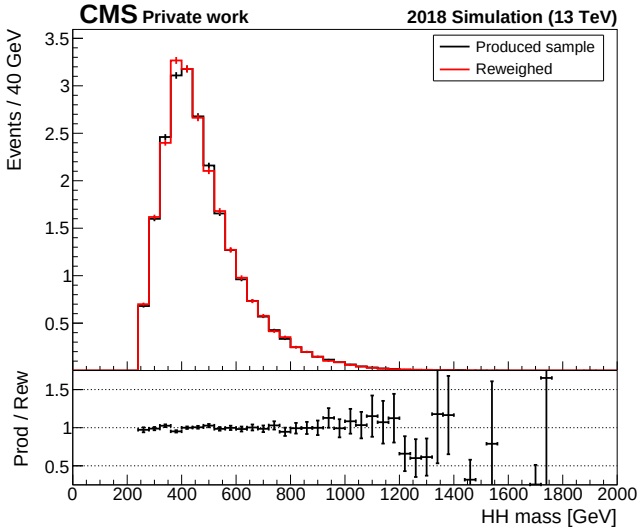


Figure 4.1: Distribution of the invariant mass of the HH system obtained (black) from a HH sample produced via gluon-gluon fusion with $\kappa_\lambda = 0$ (red) by reweighing the $\kappa_\lambda = 1$, $\kappa_\lambda = 2.45$ and $\kappa_\lambda = 5$ samples to produce the distribution expected from a $\kappa_\lambda = 0$ sample.

Similarly, the shape and yield of the distribution from a generic sample can be obtained by reweighing the same distribution obtained from six produced samples. The samples used in the analysis were generated at LO with MadGraph5_aMC@NLO with the κ_V, κ_{2V} and κ_λ values shown in Table 4.2. An additional LO sample with $\kappa_V = 0.5$, $\kappa_{2V} = 1$ and $\kappa_\lambda = 1$ was also produced to perform closure tests on the reweighing method. Fig. 4.2 shows a comparison between the HH system mass obtained directly from the sample with $\kappa_V = 0.5$ and the same distribution obtained by reweighing the histograms from the six given samples. Again, good agreement between the original and the reweighed distribution is observed, thus validating the reweighing method.

κ_V	κ_{2V}	κ_λ	σ (fb)
1	1	1	1.726
1.5	1	1	66.018
1	1	0	4.609
1	1	2	1.423
1	2	1	14.218
1	0	1	27.08

Table 4.2: Combinations of the (κ_V, κ_{2V}, κ_λ) parameter values considered in the VBF HH $\rightarrow$ bb$\tau\tau$ simulated samples used and their corresponding cross sections.

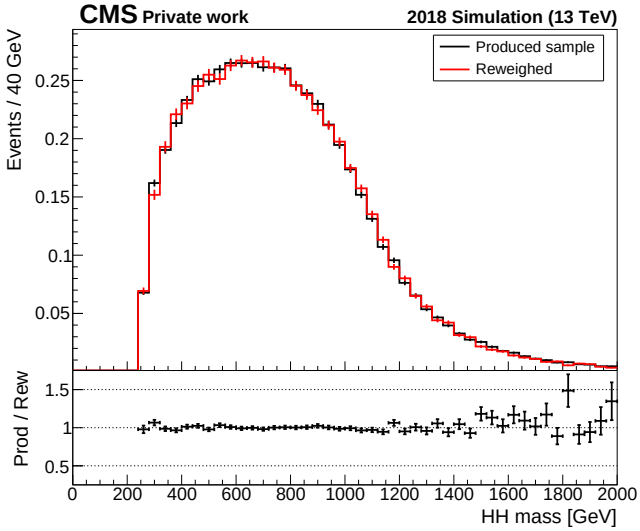


Figure 4.2: Distribution of the invariant mass of the HH system obtained from a HH sample produced via vector boson fusion with $\kappa_V = 0.5$ (and the other couplings set to their SM values) (black) and by reweighing the samples in Table 4.2 to produce the distribution expected for a sample with $\kappa_V = 0.5$ (red).

4.1.2 Background modelling

Simulated events are generally used to predict the background contributions in this analysis.

$Z/\gamma^* \to ll$ + jets is modelled with MadGraph5_aMC@NLO at LO accuracy and normalized to its cross section computed at NNLO. The agreement between data and the simulation is improved by applying weights derived from dedicated $Z \to \mu\mu$ control regions, as shown in Section 4.4.2.

$t\bar{t}$ production is modelled with POWHEG at NLO, and its cross section is extracted by fitting in a $t\bar{t}$ control region (see Section 4.4.3).

QCD multi-jet background is estimated fully from data, as shown in Section 4.4.1.

All background processes considered, their modelling approaches and the cross sections used are summarized in Table 4.3.

4.2 Physics objects preselection

To increase the purity of the reconstructed objects, a first selection on the muons, electrons, τ_h, and jets is performed.

Process	Modelling (accuracy)	$\sigma \times$ B (pb)
$Z/\gamma^* \to ll$ + jets	MadGraph5_aMC@NLO (LO)	6077.22
$t\bar{t}$	Powheg (NLO)	
$\to$ Di-leptonic		88.29
$\to$ Semi-leptonic		365.34
$\to$ Fully-hadronic		377.96
QCD multijet	Data-driven (Section 4.4.1)	
$W \to l\nu$ + jets	MadGraph5_aMC@NLO (LO)	61526.7
Electroweak	MadGraph5_aMC@NLO (LO)	
W^+ + jj		25.62
W^- + jj		20.25
Z + jj		3.987
Single-t (Single-$\bar{t}$)	Powheg (NLO)	
W-channel		35.85 (35.85)
t-channel		136.02 (80.95)
WW	Powheg (NLO)	
$\to 2l2\nu$		12.18
$\to 2l2q$		50.0
$\to 4q$		51.72
ZZ	Powheg (NLO)	
$\to 4l$		1.26
$\to 2l2\nu$		0.564
$\to 2l2q$		5.52
$\to 2q2\nu$		4.07
WZ		
$\to 3l\nu$	Powheg (NLO)	4.43
$\to l\nu 2q$	MadGraph5_aMC@NLO (NLO)	10.71
$\to l3\nu$	MadGraph5_aMC@NLO (NLO)	3.06
$\to 2l2q$	MadGraph5_aMC@NLO (NLO)	10.71
Triboson	MadGraph5_aMC@NLO (NLO)	
WWW		0.209
WWZ		0.168
WZZ		0.057
ZZZ		0.0147
Single Higgs	Powheg (NLO)	
ggH ($H \to \tau\tau$)		3.07
VBFH ($H \to \tau\tau$)		0.238
ttH		
$H \to bb$		0.29
$H \to \tau\tau$ (2017-18)		0.032
$H \not\to bb$		0.214 (2016), 0.182 (2017-18)
WH ($H \to \tau\tau$)		0.0858
ZH ($H \to bb$)		
$Z \to ll$		0.05
$Z \to qq$		0.35
ZH ($H \to \tau\tau$)		0.0556
$t\bar{t}W$ ($W \to l\nu$, qq)	MadGraph5_aMC@NLO (NLO)	0.61
$t\bar{t}Z$ ($Z \to 2l2\nu$)	MadGraph5_aMC@NLO (NLO)	0.2529
$t\bar{t}WW$	MadGraph5_aMC@NLO (LO)	0.000698
$t\bar{t}WZ$	MadGraph5_aMC@NLO (LO)	0.000244
$t\bar{t}ZZ$	MadGraph5_aMC@NLO (LO)	0.000139

Table 4.3: Background processes, modelling used and cross sections at 13 TeV.

4.2.1 Muons

Muons in the analysis are reconstructed as explained in Section 3.2.1 and are only selected if they pass the tight identification working point and the tight isolation working point (corresponding to $I^{\mu}_{rel} < 0.15$). To ensure the compatibility of the muon candidate with the primary vertex, the distance between the muon track and the primary vertex must be in the transverse plane $d_{xy} < 0.045$ cm and in the longitudinal direction $d_z < 0.2$ cm.

4.2.2 Electrons

Electron identification is performed using the MVA described in Section 3.2.2. The tight working point, characterized by an 80% efficiency, is used in the analysis. The same $d_{xy} < 0.045$ cm and $d_z < 0.2$ cm requirements as for muons are imposed.

4.2.3 Hadronic taus

Decays of τ leptons into hadrons and a neutrino are reconstructed by the HPS identification algorithm, as described in Section 3.2.4. In the analysis only the decays into $h^{\pm}$, $h^{\pm}\pi^0$, $h^{\pm}\pi^0\pi^0$, and $h^{\pm}h^{\mp}h^{\pm}$ are considered. In order to discriminate the correct hadronic τ decays from quark and gluon jets and from electrons and muons, the three DeepTau discriminants will be considered. Different working points (as described in Section 3.2.4) will be used depending on the τ pair decay channel. τ_h candidates also need to fulfil the $d_z < 0.2$ cm requirement.

4.2.4 Jets

Both AK4 and AK8 jets, described in Section 3.2.3 are considered. b and VBF jet candidates are reconstructed as AK4 jets with $p_T > 20$ GeV ($p_T > 30$ GeV for VBF jets). VBF jets are reconstructed within $\eta < |4.7|$. Since b-tagging relies on the information from the tracker and this is only available in the central region, b jet candidates are required to be with $|\eta| < 2.4$. In 2017, all jets with $20 < p_T < 50$ and $2.65 < |\eta| < 3.139$ are removed since they populate a very specific phase space region where it is known 2017 MC simulation does not match the data.

All jets are required to pass the tight working point of the particle-flow jet identification. Jets with $p_T < 50$ GeV are also required to pass the loose pile-up jet ID discriminator working point.

Depending on the separation ΔR_{jj} between the two b jets finally identified as coming from the Higgs decay, there are three possible regimes. If $\Delta R_{jj} > 0.8$, they are reconstructed as two different AK4 jets. If $0.4 < \Delta R_{jj} < 0.8$, the jets are partially in overlap and are reconstructed as two different AK4 jets and also as one merged AK8 jet. Finally, if $\Delta R_{jj} < 0.4$, the two jets will be reconstructed as one merged AK8 jet. The two final regimes correspond to the boosted production.

However, for non resonant HH production, the latest highly boosted regime is almost never reached. Therefore, events are classified into two categories, the *resolved* and the *boosted* categories, associated to the $\Delta R_{jj} > 0.8$ and $0.4 < \Delta R_{jj} < 0.8$ topologies, respectively.

4.3 Analysis flow

4.3.1 $H \to \tau\tau$ pair selection

As initial step, the decay products of the $H \to \tau\tau$ decay are identified, considering the three decay modes described ($\tau_\mu\tau_h$, $\tau_e\tau_h$ and $\tau_h\tau_h$).

After applying the preselection described in Section 4.2 on the muons, electrons and hadronic τ, the $\tau\tau$ decay mode is assessed. All signal events are required to have at least one hadronic τ. Events are classified as $\tau_\mu\tau_h$ if a muon satisfying the baseline selection is found, otherwise are classified as $\tau_e\tau_h$ if a baseline-electron is found and lastly as $\tau_h\tau_h$ if two hadronic τ are present. For each channel, if more than one pair is found, they are sorted according to the isolation of the muon or the electron in the $\tau_\mu\tau_h$ and $\tau_e\tau_h$ channels and the highest isolated τ_h (i.e. the one with the highest DeepTauVSjet score) in the $\tau_h\tau_h$ channel.

Additionally to the preselection described in Sections 4.2.1, 4.2.2, and 4.2.3, a kinematic selection is applied to the different objects depending on the trigger strategy followed, which depends on the channel considered. A summary of these triggers and the p_t thresholds applied to the trigger objects is included in Table 4.4. In the semileptonic channels, single-lepton and cross-lepton triggers are used. The single-lepton triggers require an isolated muon or electron, while the cross-triggers require an additional hadronic τ. The cross-lepton triggers allow to reduce the threshold on the lepton p_T, increasing the acceptance of the analysis. Online p_t thresholds vary between trigger strategies and years. In 2016, the single-μ trigger required a muon of 22 GeV, while the cross trigger $\mu\tau_h$ required a muon of 19 GeV and a τ_h of 20 GeV. These numbers increased in 2017 and 2018 to 24 GeV for the single-μ trigger and 20 and 27 GeV for the cross trigger. Similarly, during the 2016 (2017-2018) campaign single-e triggers required a 25 (32) GeV electron, while cross triggers $e\tau_h$ (only available in 2017 and 2018) required a 24 GeV electron and a 30 GeV τ_h. For the $\tau_h\tau_h$ channel, a double-τ_h trigger was used during the three years, requiring two 35 GeV τ_h. Additionally, during 2017 a new double-τ_h trigger targeting VBF production was introduced to increase the $\tau_h\tau_h$ channel phase-space, so on top of requiring two 20 GeV τ_h, one jet with $p_T > 115$ GeV and another one with $p_T > 45$ GeV are selected at trigger level (if their invariant mass is $m_{jj} > 620$ GeV). All electrons, muons, and hadronic τ_h considered for triggering are required to be within $|\eta| < 2.1$.

Each reconstructed offline object is required to geometrically match (within $\Delta R < 0.5$) their corresponding trigger object and also pass a p_T threshold depending

$\tau\tau$ channel	Year	Trigger p_T thresholds
$\tau_\mu\tau_h$	2016	Single-μ: $p_t^\mu > 22$ GeV Cross $\mu\tau_h$: $p_t^\mu > 20$ GeV, $p_t^{\tau_h} > 20$ GeV
	2017	Single-μ: $p_t^\mu > 24$ GeV Cross $\mu\tau_h$: $p_t^\mu > 20$ GeV, $p_t^{\tau_h} > 27$ GeV
	2018	Single-μ: $p_t^\mu > 24$ GeV Cross $\mu\tau_h$: $p_t^\mu > 20$ GeV, $p_t^{\tau_h} > 27$ GeV
$\tau_e\tau_h$	2016	Single-e: $p_t^e > 25$ GeV
	2017	Single-e: $p_t^e > 32$ GeV Cross $e\tau_h$: $p_t^e > 24$ GeV, $p_t^{\tau_h} > 30$ GeV
	2018	Single-e: $p_t^e > 32$ GeV Cross $e\tau_h$: $p_t^e > 24$ GeV, $p_t^{\tau_h} > 30$ GeV
$\tau_h\tau_h$	2016	Double-τ_h: $p_t^{\tau_h} > 35$ GeV
	2017	Double-τ_h: $p_t^{\tau_h} > 35$ GeV VBF+H $\rightarrow \tau_h\tau_h$ trigger: $p_t^{\tau_h} > 20$ GeV
	2018	Double-τ_h: $p_t^{\tau_h} > 35$ GeV VBF+H $\rightarrow \tau_h\tau_h$ trigger: $p_t^{\tau_h} > 20$ GeV

Table 4.4: Trigger selections applied in the analysis split by channel and year.

on the HLT trigger path fired by the event,

$$p_t^{\text{offline}} >= p_t^{\text{trigger}} + \text{offset}, \tag{4.9}$$

where p_t^{offline} and p_t^{trigger} are the p_t of the offline object and the p_t threshold of the trigger object respectively and offset is a quantity chosen to be conservative with respect to the trigger turn-on curves. This offset depends on the lepton type: 1 GeV for electrons and muons and 5 GeV for hadronic τ. Events passing the single-lepton triggers also require an hadronic τ_h of 20 GeV and $|\eta| < 2.3$.

All selections applied to offline objects are summarized in Tables 4.5, 4.6 and 4.7. Additional selections are applied to the lepton pair to increase the signal purity, such as both τ candidates having opposite charge or a spatial separation between them of $\Delta R > 0.5$.

Finally, a third lepton veto is applied on all events, discarding the ones where an additional muon or electron is found besides the ones selected in the $\tau\tau$ pair. This selection reduces the background coming from Z/γ^* + jets and ensures that the three $\tau_h\tau_h$ decay modes considered are mutually exclusive. These additional leptons are required looser selections than the signal leptons. Events are vetoed if they contain an additional electron with $|\eta| < 2.5$ and $p_T > 10$ GeV that passes the loose MVA ID and loose relative isolation working points. Similarly, they are also vetoed if a muon within $|\eta| < 2.4$, $p_T > 10$ GeV and that passes the tight or medium ID working points and the loose relative isolation working point is present.

Object	Year	Selections
μ	All	Tight ID, Tight Iso. $\|\eta\| < 2.1$ $d_{xy} < 0.045$ cm, $d_z < 0.2$ cm
	2016	single-μ: $p_t > 23$ GeV cross $\mu\tau_h$: $p_t > 20$ GeV
	2017	single-μ: $p_t > 25$ GeV cross $\mu\tau_h$: $p_t > 21$ GeV
	2018	single-μ: $p_t > 25$ GeV cross $\mu\tau_h$: $p_t > 21$ GeV
τ_h	All	Tight DeepTauVSmu working point VLoose DeepTauVSe working point Medium DeepTauVSjet working point $d_z < 0.2$ cm
	2016	single-μ: $p_t > 20$ GeV, $\|\eta\| < 2.3$ cross-$\mu\tau_h$: $p_t > 25$ GeV, $\|\eta\| < 2.1$
	2017	single-μ: $p_t > 20$ GeV, $\|\eta\| < 2.3$ cross $\mu\tau_h$: $p_t > 32$ GeV, $\|\eta\| < 2.1$
	2018	single-μ: $p_t > 20$ GeV, $\|\eta\| < 2.3$ cross $\mu\tau_h$: $p_t > 32$ GeV, $\|\eta\| < 2.1$
pair	All	$\Delta R(\mu, \tau_h) > 0.5$ Opposite charge

Table 4.5: Offline selection for the $\tau_\mu\tau_h$ final state.

4.3.2 H → bb pair and VBF jet selection

Once a $\tau\tau$ pair candidate is identified, a selection of the two b jets coming from the decay of the other Higgs boson is performed. All jets are required to pass a baseline selection (described in Section 4.2.4), before they are sorted with the newly developed *HH-btag* algorithm: the two jets with the highest scores are taken to be the two b jets coming from the Higgs boson decay. Additionally, the two forward jets characteristic from the VBF production are identified.

The HH-btag algorithm

Studies performed on the previous HH → bb$\tau\tau$ CMS result on 2016 data [81] showed that one of the main sources of inefficiency in the analysis was the selection of the H → bb candidate, that at that moment was using the combined secondary vertex [82] algorithm for identifying the b-quarks coming from the Higgs decay. A similar result was found after moving to the DeepJet algorithm. Therefore, a new algorithm was developed explicitly for the full Run 2 analysis, the HH-btag algorithm, based on a neural network architecture. The input features of this network are the kinematic variables of the b-jet candidates, of the H → $\tau\tau$ candidate and of the missing transverse energy. More complex variables are also included, such as the DeepJet

Object	Year	Selections
e	All	Tight MVA ID $\|\eta\| < 2.1$ $d_{xy} < 0.045$ cm, $d_z < 0.2$ cm
	2016	Single-e: $p_t > 26$ GeV
	2017	Single-e: $p_t > 33$ GeV Cross $e\tau_h$: $p_t > 25$ GeV
	2018	Single-e: $p_t > 33$ GeV Cross $e\tau_h$: $p_t > 25$ GeV
τ_h	All	Tight DeepTauVSmu working point VLoose DeepTauVSe working point Medium DeepTauVSjet working point $d_z < 0.2$ cm
	2016	Single-e: $p_t > 20$ GeV, $\|\eta\| < 2.3$
	2017	Single-e: $p_t > 20$ GeV, $\|\eta\| < 2.3$ Cross $e\tau_h$: $p_t > 35$ GeV, $\|\eta\| < 2.1$
	2018	Single-e: $p_t > 20$ GeV, $\|\eta\| < 2.3$ Cross $e\tau_h$: $p_t > 35$ GeV, $\|\eta\| < 2.1$
pair	All	$\Delta R(e, \tau_h) > 0.5$ Opposite charge

Table 4.6: Offline selection for the $\tau_e\tau_h$ final state.

Object	Year	Selections
τ_h	All	$\|\eta\| < 2.1$ VLoose DeepTauVSmu working point VVLoose DeepTauVSe working point Medium DeepTauVSjet working point $d_z < 0.2$ cm
	2016	Double-τ_h: $p_t > 40$ GeV
	2017	Double-τ_h trigger: $p_t > 40$ GeV VBF+H $\rightarrow \tau_h\tau_h$ trigger: $p_t > 25$ GeV
	2018	Double-τ_h trigger: $p_t > 40$ GeV VBF+H $\rightarrow \tau_h\tau_h$ trigger: $p_t > 25$ GeV
pair	All	$\Delta R(\tau_h, \tau_h) > 0.5$ Opposite charge

Table 4.7: Offline selection for the $\tau_h\tau_h$ final state.

score of the b jet candidates, the angular separation between the b jets and the H $\rightarrow \tau\tau$ candidate and some global event variables like the year of the sample of the H $\rightarrow \tau\tau$ decay channel.

To evaluate the performance of the network, a variable named purity is defined:

$$\text{purity} = \frac{N_{\text{true}}}{N}, \tag{4.10}$$

where N is the total number of events where a bb pair can be identified by a given algorithm and N_{true} the number of events where this bb pair matches with the bb pair produced from the Higgs boson candidate decay at generator level. With this definition, the HH-btag obtains a purity of 95% and 97% for ggF and VBF events, respectively, an improvement of around 5% with respect to the DeepJet discriminator.

VBF jet selection

Among all the jets satisfying the baseline selection described in Section 4.2.4, the ones that have not been identified as bb pair candidates, are now considered as possible VBF jets. The two VBF jet candidates are selected as the ones that, combined, give that highest invariant mass among all the possible combinations with the available jets. This criteria gives a selection purity (defined as in Eq. 4.10) when running on the SM VBF HH $\rightarrow$ bb$\tau\tau$ sample of around 85%.

4.3.3 Event categorization

Events where a $\tau\tau$ and a bb pair candidates are found (the so-called *baseline* categorization) are then split into eight mutually exclusive categories.

In order to select signal events produced via VBF, a first selection (denoted as *VBF inclusive*) is applied by requiring two VBF jet candidates as described in Section 4.3.2 that also satisfy two further requirements in order to reduce the background contamination: invariant mass of the jet-jet system larger than 500 GeV and a separation in pseudorapidity between both VBF jets of $\Delta\eta > 3$. Regarding the H $\rightarrow$ bb pair candidate, at least one of the b jets should pass the medium DeepJet b-tag working point. For events in the $\tau_h\tau_h$ channel that were included in the VBF+H $\rightarrow \tau_h\tau_h$ trigger phase space, i.e. the ones where both τ_h have a $p_T > 25$ GeV and at least one has a $p_T < 40$ GeV, the selection on the VBF jets is tightened to be at the trigger efficiency plateau. The VBF jet candidates in these events are required to have an invariant mass of $M_{jj} > 800$ GeV, the leading VBF jet candidate must have a $p_T > 140$ GeV and the subleading one a $p_T > 60$ GeV.

To further reduce the remaining contamination from backgrounds and from HH events produced through the ggF mechanism, a multi-class classification approach is introduced so events passing the VBF inclusive selection are separated into two signal categories, associated to the HH VBF and HH ggF processes, and three background-enriched ones, associated to the $t\bar{t}$, $t\bar{t}H$ and $Z/\gamma^* +$ jets processes. This multi-class classification strategy is described in Section 4.7.

Events where two VBF jets could not be identified or that do not pass the VBF inclusive selection are classified into two categories depending on the topology of

the $H \rightarrow$ bb decay, as shown in Section 4.2.4. Events that belong to the resolved category don't require any further selection on the b jet candidates. However, for the boosted category, the presence of an AK8 jet is needed. In order to extract the substructure of this fat jet, the *soft drop declustering* [83] algorithm is used. This algorithm not only reconstructs the two subjets but also mitigates the contributions from initial state radiation and pile-up. The invariant mass of the two jets can be also obtained by this algorithm, leading to a better data-MC agreement than computing the invariant mass from the two AK4 jets.

Events in the boosted category should have an AK8 jet whose softdrop mass is $m_{\text{AK8}} > 30$ GeV and its two subjets must geometrically match the previously selected AK4 jets with $\Delta R < 0.4$. Both jets should also pass the loose DeepJet b-tag working point. If they do not belong to this boosted category, events are categorized as resolved. In addition, resolved events are finally split into two categories depending on the b-tag multiplicity: the *Resolved 1 b-tag* category includes events where only one b-jet passes the medium DeepJet working point, while in the *Resolved 2 b-tag* category both b-jets pass the medium DeepJet working point.

To summarize, the final analysis categories are eight: Resolved 1 b-tag, Resolved 2 b-tag, Boosted, and the five VBF subcategories associated to five processes, HH VBF, HH ggF, $\mathrm{t\bar{t}}$, $\mathrm{t\bar{t}H}$ and $\mathrm{Z}/\gamma^* +$ jets. A diagram with the categorization workflow is shown in Fig. 4.3.

Even with the selections described in this Section and while performing the H$\rightarrow$ $\tau\tau$ and H$\rightarrow$ bb pair candidates determination, the events belonging to previously described analysis categories will be still mostly coming from background processes. An additional selection is included in the boosted and the two resolved categories (not to the VBF categories due to the low event yields) by including a requirement on the invariant mass of the $\tau\tau$ ($m_{\tau\tau}$) and the bb (m_{bb}) systems. The former is computed using the *SVFit* [84] algorithm, that is based on a likelihood function that quantifies the level of compatibility of a Higgs mass hypothesis (in the HH $\rightarrow$ bb$\tau\tau$ analysis, 125 GeV) with the momenta of the visible τ decay products (e, μ and τ_h) and the missing transverse energy reconstructed in the event. This algorithm provides a better invariant mass reconstruction so that is closer to 125 GeV when compared to the invariant mass of the τ visible decay products or even the visible mass plus the missing transverse energy, as shown in Fig. 4.4.

The main aim of the mass cut is to remove significantly outlying background events where no signal overlap is expected. Regarding the shape of the mass cut, it was chosen to be elliptical, since is still simple enough (only needs four parameters) and provides better results than a simpler rectangular cut. The parameters used are different for the two resolved and the boosted categories, as the kinematics in these two cases are different, and are obtained by minimising the background acceptance while keeping the signal acceptance over 90%. The final mass cut used for the resolved categories is

$$\frac{(m_{\tau\tau} - 129\ \text{GeV})^2}{(53\ \text{GeV})^2} + \frac{(m_{\text{bb}} - 169\ \text{GeV})^2}{(145\ \text{GeV})^2} < 1, \tag{4.11}$$

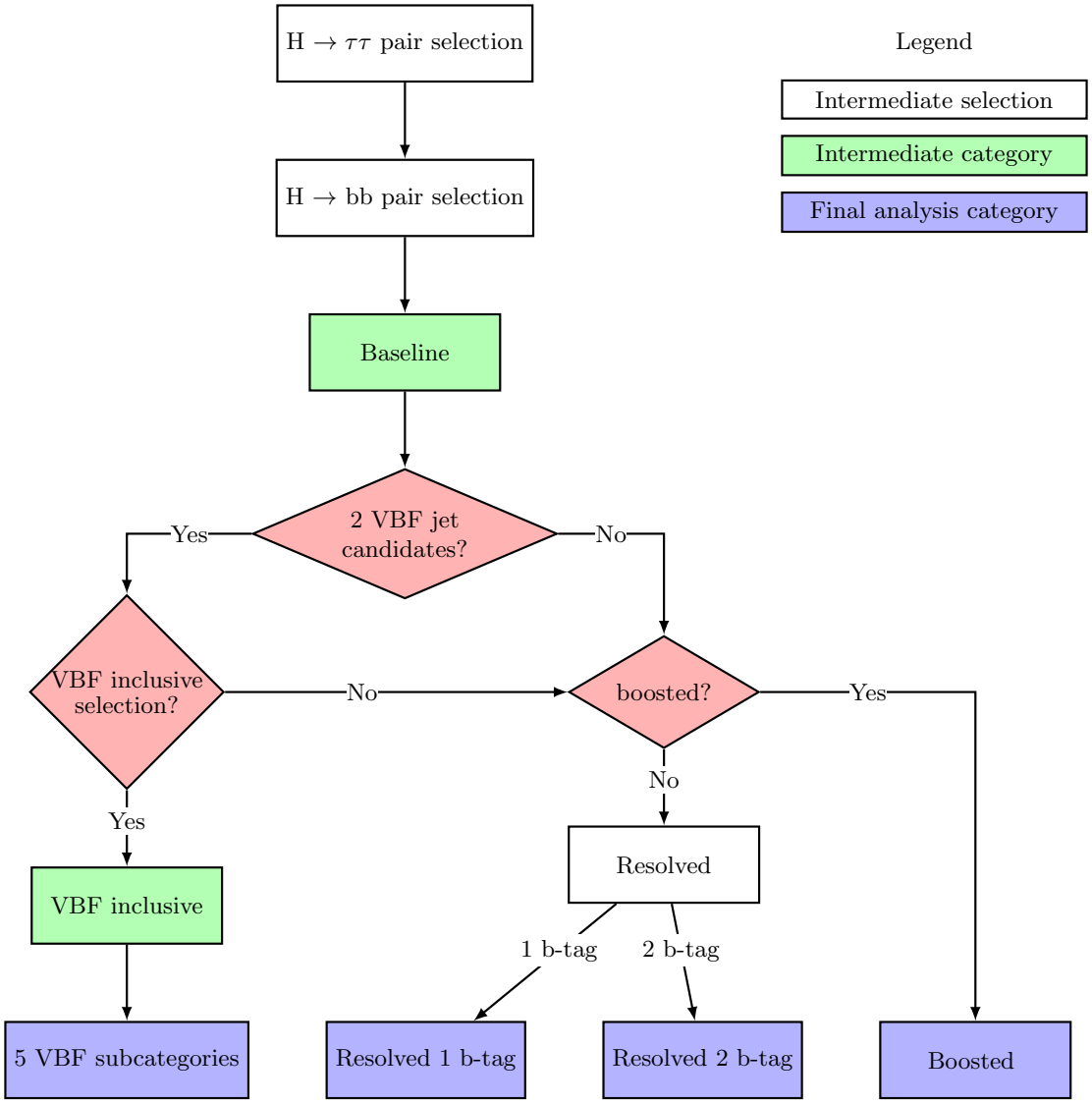


Figure 4.3: Description scheme of event categorization in the HH → bbττ analysis. The 5 VBF subcategories are shown in the Figure merged for the sake of visualization.

while, for the boosted category,

$$\frac{(m_{\tau\tau} - 128\ \text{GeV})^2}{(60\ \text{GeV})^2} + \frac{(m_{\text{bb}} - 159\ \text{GeV})^2}{(94\ \text{GeV})^2} < 1. \tag{4.12}$$

The values of the acceptance × efficiency for the different analysis categories (Baseline, Resolved 1 b-tag, Resolved 2 b-tag, Boosted, and VBF inclusive) in 2018 are shown in Tables 4.8, 4.9, 4.10, and 4.11 for the $t\bar{t}$ and Z/γ^* + jets backgrounds and the HH ggF and HH VBF signals, both produced as predicted by the SM.

	$\tau_\mu\tau_h$	$\tau_e\tau_h$	$\tau_h\tau_h$
Baseline	$(11002.6 \pm 6.9) \times 10^{-6}$	$(6932.5 \pm 5.3) \times 10^{-6}$	$(2211.9 \pm 9.2) \times 10^{-7}$
Resolved 1 b-tag	$(1604.6 \pm 2.6) \times 10^{-6}$	$(961.1 \pm 2.0) \times 10^{-6}$	$(222.4 \pm 2.9) \times 10^{-7}$
Resolved 2 b-tag	$(1401.4 \pm 2.3) \times 10^{-6}$	$(844.1 \pm 1.8) \times 10^{-6}$	$(248.8 \pm 3.0) \times 10^{-7}$
Boosted	$(93.8 \pm 2.3) \times 10^{-7}$	$(61.9 \pm 1.8) \times 10^{-7}$	$(21.8 \pm 3.3) \times 10^{-8}$
VBF inclusive	$(761.2 \pm 1.8) \times 10^{-6}$	$(4766.8 \pm 14.1) \times 10^{-7}$	$(206.1 \pm 2.8) \times 10^{-7}$

Table 4.8: Acceptance × efficiency values for the $t\bar{t}$ background in the different analysis categories in 2018.

	$\tau_\mu\tau_h$	$\tau_e\tau_h$	$\tau_h\tau_h$
Baseline	$(4709.4 \pm 13.9) \times 10^{-7}$	$(4465.1 \pm 12.9) \times 10^{-7}$	$(540.1 \pm 2.9) \times 10^{-7}$
Resolved 1 b-tag	$(285.0 \pm 3.7) \times 10^{-7}$	$(270.9 \pm 3.1) \times 10^{-7}$	$(318.3 \pm 7.9) \times 10^{-8}$
Resolved 2 b-tag	$(292.7 \pm 13.6) \times 10^{-8}$	$(265.8 \pm 10.4) \times 10^{-8}$	$(39.9 \pm 1.9) \times 10^{-8}$
Boosted	$(71.4 \pm 4.4) \times 10^{-9}$	$(65.6 \pm 4.1) \times 10^{-9}$	$(34.9 \pm 3.7) \times 10^{-9}$
VBF inclusive	$(166.0 \pm 6.1) \times 10^{-8}$	$(181.2 \pm 5.5) \times 10^{-8}$	$(375.0 \pm 13.6) \times 10^{-9}$

Table 4.9: Acceptance × efficiency values for the Z/γ^* + jets background in the different analysis categories in 2018.

	$\tau_\mu\tau_h$	$\tau_e\tau_h$	$\tau_h\tau_h$
Baseline	$(270.4 \pm 1.7) \times 10^{-5}$	$(1611.2 \pm 12.8) \times 10^{-6}$	$(1658.5 \pm 13.4) \times 10^{-6}$
Resolved 1 b-tag	$(767.2 \pm 9.3) \times 10^{-6}$	$(446.8 \pm 6.9) \times 10^{-6}$	$(443.1 \pm 7.1) \times 10^{-6}$
Resolved 2 b-tag	$(995.6 \pm 9.8) \times 10^{-6}$	$(585.8 \pm 7.2) \times 10^{-6}$	$(652.9 \pm 7.8) \times 10^{-6}$
Boosted	$(104.9 \pm 3.6) \times 10^{-6}$	$(73.2 \pm 2.8) \times 10^{-6}$	$(110.2 \pm 3.6) \times 10^{-6}$
VBF inclusive	$(150.0 \pm 4.1) \times 10^{-6}$	$(92.7 \pm 3.1) \times 10^{-6}$	$(109.9 \pm 3.4) \times 10^{-6}$

Table 4.10: Acceptance × efficiency values for the HH ggF signal process in the different analysis categories in 2018.

	$\tau_\mu\tau_h$	$\tau_e\tau_h$	$\tau_h\tau_h$
Baseline	$(1846.7 \pm 8.8) \times 10^{-6}$	$(1020.4 \pm 6.4) \times 10^{-6}$	$(1009.8 \pm 6.4) \times 10^{-6}$
Resolved 1 b-tag	$(257.8 \pm 3.3) \times 10^{-6}$	$(136.8 \pm 2.3) \times 10^{-6}$	$(115.9 \pm 2.2) \times 10^{-6}$
Resolved 2 b-tag	$(273.0 \pm 3.2) \times 10^{-6}$	$(150.8 \pm 2.3) \times 10^{-6}$	$(199.3 \pm 2.6) \times 10^{-6}$
Boosted	$(179.9 \pm 8.5) \times 10^{-7}$	$(129.5 \pm 7.4) \times 10^{-7}$	$(253.7 \pm 10.2) \times 10^{-7}$
VBF inclusive	$(717.7 \pm 5.5) \times 10^{-6}$	$(405.7 \pm 4.0) \times 10^{-6}$	$(434.1 \pm 4.1) \times 10^{-6}$

Table 4.11: Acceptance × efficiency values for the HH VBF signal process in the different analysis categories in 2018.

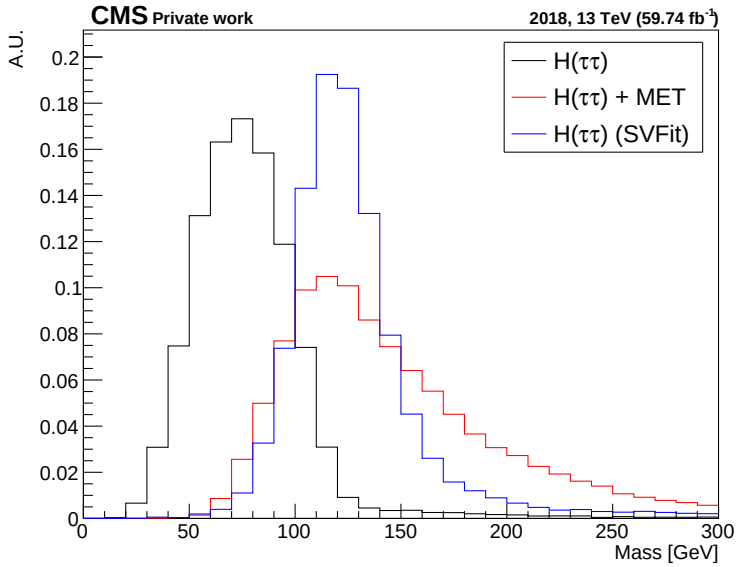


Figure 4.4: Distribution of the $\tau\tau$ pair mass for HH events produced via ggF computed from (black) the visible decay products of both τ, (red) the visible decay products of both τ and the missing transverse energy and (blue) with the SVFit algorithm.

4.4 Background estimation

4.4.1 QCD multi-jet

Generic QCD multi-jet events can enter the final selection if two jets are misidentified as the $\tau\tau$ pair. The process creating QCD multi-jets has the largest cross section among all the background processes so, event it gets reduced thanks to the analysis selections, its contribution ends up being notably important, particularly in the $\tau_h\tau_h$ channel. Estimating the QCD background contribution via simulated events would require a prohibitively large number of them, so the estimation is performed using a data-driven method. The strategy used is often called "ABCD" method, and is described as follows.

QCD background estimation using the ABCD method

The ABCD takes its name from the fact that four regions are used: the signal region, A, and three other regions, B, C and D, obtained by reverting two different selection criteria. In the HH $\rightarrow$ bb$\tau\tau$ analysis, region A (*opposite-sign, isolated*) requires that both leptons are isolated and have opposite sign. Region B (*same-sign, isolated*) is obtained by inverting the tau pair charge requirement. In region C (*opposite-sign, non-isolated*), the DeepTauVsJet selection is inverted by requiring that the selected τ_h (the one with the lowest p_t in the $\tau_h\tau_h$ channel) fails the medium working point but still passes the VVVLoose one. Region D (*same-sign, non-isolated*) combines

these two inverted selections. An sketch of the four regions is shown in Fig. 4.5.

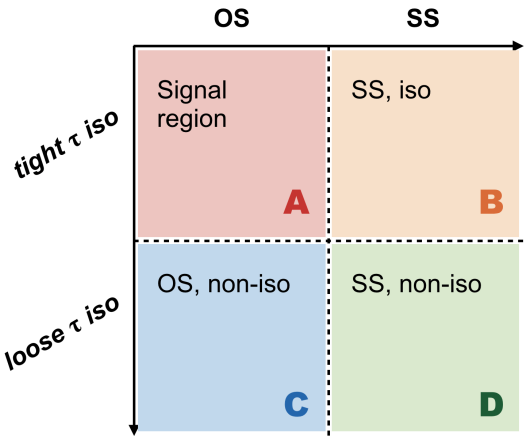


Figure 4.5: Sketch of the four regions used to estimate the QCD multi-jet background.

The distribution of the multi-jet QCD contribution in any given binned variable is estimated from the C region: yields from the other backgrounds are substracted from the data in this region and the remaining number of events in each bin is then multiplied by a factor $k^{\text{iso/non-iso}}$. This factor accounts for moving from the non-isolated selection (present in region C) to the isolated selection (used in the signal region), and is measured as the ratio of event yields in regions B and D after substracting the other background yields in each region. Therefore, the ABCD method gives a QCD estimation whose shape comes from region C and its yield from regions B, C and D as given by the formula C $\times$ B / D. However, in some analysis categories, due to the lack of statistics, a different selection is applied to regions B and D. In the boosted category, the B / D ratio is computed with all the selections defined in the category excluding the b-tagging requirement, while in the five VBF subcategories, the B / D correction factor is estimated from the VBF inclusive category, which includes the five subcategories.

Something to be noted is that decision of considering region C to obtain the shape from the QCD distributions instead of region B is quite arbitrary. This decision will be discussed in Section 5.3.

4.4.2 $Z/\gamma^* + \text{jets}$

The contribution of $Z/\gamma^* \to ll$ associated with jets is estimated from Monte-Carlo simulations. Samples can be produced with MadGraph5_aMC@NLO at LO or NLO precisions. Two sets of LO-precision samples are produced: inclusive samples, with up to four jets produced at matrix element with MLM merging, and different samples with different number of jets at matrix element (from one to four) or requiring two b jets. The combination of both inclusive and exclusive samples lead

to large statistics for this process. However, LO modelling of the jets is worse than in NLO, where a better modelling is observed but less statistics are available. The final solution is use the samples with LO precision but improving the modelling by correcting the simulation with a data-driven method. The correction factors used in the MC samples are derived from several $Z \to \mu\mu$ sideband regions.

For the correction factor computation, data and simulated events are required to pass a selection similar to the one applied in the analysis. At trigger level, events are selected by the single-muon trigger, with p_T thresholds of 22 GeV in 2016 and 24 GeV in 2017 and 2018. At offline level, events should have two muons with $p_T > 20$ GeV and $|\eta| < 2.4$ and one muon should match the muon used for triggering. A vertex constraint of $d_{xy} < 0.045$ mm and $d_z < 0.2$ mm is applied to both muons, and they should satisfy the tight muon identification and tight isolation working points. Each of the two selected muons should be separated from the other by $\Delta R > 0.1$, have opposite charge and an invariant mass of $m_{\mu\mu} > 50$ GeV. The third lepton veto and the jet section are the same as for the analysis final states.

To reduce the QCD and $t\bar{t}$ contributions in the sideband region, a cut on the missing transverse energy is applied by requiring $E_T^{\text{miss}} < 45$ GeV.

Scale factors are estimated using 18 control regions. First, events are split in 3 regions based on the number of jets that pass the medium DeepJet b-tag working point: 0, 1 or 2. The latter is the one with the largest contribution in the phase-space of interest, since it has exactly the same final state as the signal. Then, each control region is divided in six control regions based on the p_T of the reconstructed Z candidate. The six control regions are named as *Very Low* p_T, *Low* p_T, *Medium 1* p_T, *Medium 2* p_T, *High* p_T and *Very High* p_T. The p_T range of each control region is shown in Table 4.12.

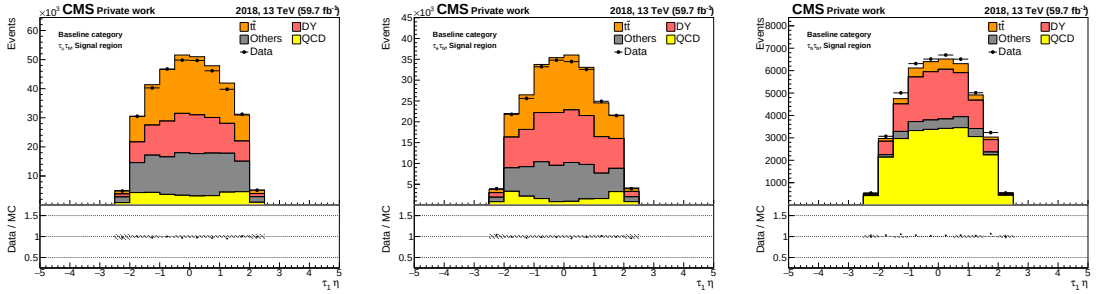
In each of the 18 regions, LO $Z/\gamma^* \to ll$ MC samples are subdivided in another 18 contributions, based on the number of b-partons at Matrix Element level (0, 1, 2) and the p_T of the Z boson at generator level (same p_T ranges as for the sideband regions). The data distribution of the invariant mass of the muon pair is fitted simultaneously in all categories, leaving the normalization of the 18×18 DY contributions floating.

p_T region	p_T range - 2016	p_T range - 2017	p_T range - 2018
Very Low p_T	$0 < p_T \leq 10$	$0 < p_T \leq 10$	$0 < p_T \leq 10$
Low p_T	$10 < p_T \leq 50$	$10 < p_T \leq 30$	$10 < p_T \leq 30$
Medium 1 p_T	$50 < p_T \leq 80$	$30 < p_T \leq 50$	$30 < p_T \leq 50$
Medium 2 p_T	$80 < p_T \leq 110$	$50 < p_T \leq 100$	$50 < p_T \leq 100$
High p_T	$110 < p_T \leq 190$	$100 < p_T \leq 200$	$100 < p_T \leq 200$
Very High p_T	$p_T > 190$	$p_T > 200$	$p_T > 200$

Table 4.12: Definition of the Z p_T regions for the $Z/\gamma^* \to ll$ + jets scale factor computation.

4.4.3 Top-antitop background

The contribution of the $t\bar{t}$ background is obtained relying on the MC simulation. However, while the shape of the process if well modelled by the simulation, the normalization of the background shows a disagreement with respect to the data, as can be seen in Fig. 4.6.



(a) (b) (c)

Figure 4.6: Distributions of the η of the first lepton (μ, e, or the most isolated τ_h depending on the $\tau\tau$ decay channel) in 2018 for the (a) $\tau_\mu\tau_h$ channel, (b) $\tau_e\tau_h$ channel, and (c) $\tau_h\tau_h$ channel. Note that the content of each bin is scaled by the bin width. The two channels with the biggest $t\bar{t}$ contribution ($\tau_\mu\tau_h$ and $\tau_e\tau_h$) show a data/background disagreement in the η region where more $t\bar{t}$ is expected, while good agreement can be seen where poor contamination of $t\bar{t}$ is present, which points to an issue in the $t\bar{t}$ normalization.

In order to solve this normalization problem, a specific $t\bar{t}$ control region has been defined and fitted to extract an scale factor to be applied to the $t\bar{t}$ background yield. This control region had to be as close as possible to the signal region but orthogonal to it, being as much enriched and pure on $t\bar{t}$ events as possible. The control region was finally defined by applying the same selections shown in Section 4.3.3 for the Resolved 2 b-tag category except for the mass cut, which will be the inverse of the mass cut used in that category:

$$\frac{(m_{\tau\tau} - 129\ \text{GeV})^2}{(53\ \text{GeV})^2} + \frac{(m_{\text{bb}} - 169\ \text{GeV})^2}{(145\ \text{GeV})^2} > 1. \tag{4.13}$$

A fit is performed in this region by including only the $t\bar{t}$ scale factor (that applies only to $t\bar{t}$ background) as a free parameter and the statistical uncertainties in each of the three years. To reduce these statistical uncertainties, the three channels are considered altogether. The results for these scale factors are found in Table 4.13.

4.5 Corrections to simulated events

In general, simulations do not perfectly reproduce the behaviour found in real data. To improve the agreement between data and simulated samples, some correction

2016	2017	2018
0.908 ± 0.006	0.988 ± 0.006	0.966 ± 0.009

Table 4.13: Results of the $t\bar{t}$ control region fits for the three years. The values of the $t\bar{t}$ scale factors and their uncertainties are reported.

factors or *scale factors* are included in the simulated events. These scale factors can modify the whole yield of particular samples (as the normalization factors obtained for $Z/\gamma^* \to ll$ + jets or $t\bar{t}$ in Section 4.4), weight each event independently according to some characteristics of the event, or even only modify particular objects in every event.

In the following, all the correction factors applied to simulated events (except the normalization factors described in Section 4.4)) are summarised.

Pile-up reweighting

The production of additional objects coming from PU vertices can lead to variations in the analysis results. Therefore, the effect of additional proton-proton interactions is also modelled in simulated events. The probability distribution of the number of PU interactions is modelled before the data taking period itself, so the PU distributions between data and simulated samples can be different (as shown in Fig. 4.7). Therefore, simulated events have to be weighted according to the ratio between the PU distributions of data and MC. For the data, the pile-up is given by

$$\mathrm{PU} = \frac{L_{ins} \cdot \sigma_{\mathrm{inel}}}{\mathrm{f_{rev}}}, \tag{4.14}$$

where σ_{inel} is the total pp inelastic cross section and $\mathrm{f_{rev}}$ is the LHC orbit frequency (11246 Hz). The cross section takes a value of 69.2 fb, as recommended by the Collaboration.

Pile-up jet identification scale factors

As described in Section 4.2.4, jets with $p_t < 50$ GeV only are considered if they satisfy the loose working point of the pile-up jet ID discriminator. As this discriminator is not 100%(0%) efficient on real(PU) jets and its results can differ between data and MC events, some scale factors are needed to weight each particular MC event. These scale factors are produced by studying the pile-up jet ID efficiency, i.e. probability of a real jet (a reconstructed jet matched to a MC generator-level jet) to pass the PU jet ID working point, and the mistag rate, i.e. the probability of a PU jet (a reconstructed jet not matched to a MC generator-level jet) to pass the PU jet ID working point.

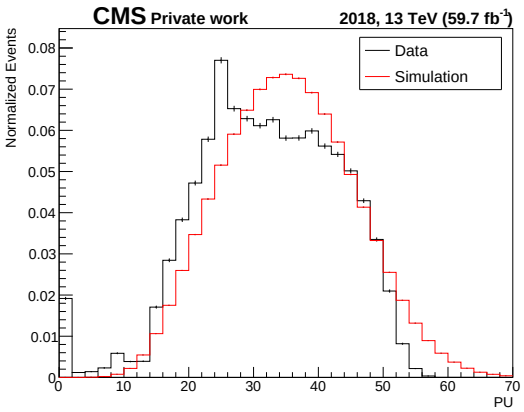


Figure 4.7: Distribution of the number of PU vertices in 2018 for both data and simulated samples.

Jet smearing

Measurements have shown that the jet energy resolution (JER) in data is worse than in simulation [85]. Then, the four-momentum of the jets in simulated events need to be further smeared to describe the data. The smearing procedure consists on a hybrid method used to scale the reconstructed four-momentum of a jet with an scaling factor f. First, a matching is performed between the jet and a generator-level jet if $\Delta R(\text{jet}, \text{genjet}) < 0.2$ and $|p_T^{\text{jet}} - p_T^{\text{genjet}}| < 3\sigma p_T^{\text{jet}}$, where σ is the relative jet momentum resolution measured in simulation. This resolution depends on the p_t and η of the jet and the measured PU. In case this matching is performed, the scaling factor f is obtained as

$$f = 1 + (s-1)\frac{p_T^{\text{jet}} - p_T^{\text{genjet}}}{p_T^{\text{jet}}}, \tag{4.15}$$

where s is the data-to-simulation core resolution scale factor. Otherwise, f is obtained in an stochastic approach as

$$f = 1 + N(0,\sigma)\sqrt{\max(s^2-1, 0)}, \tag{4.16}$$

where $N(0,\sigma)$ is a random number sampled from a normal distribution with zero mean and σ standard deviation.

The effect of the jet smearing correction over the p_T and energy of the highest p_T b jet candidate is shown in Fig. 4.8.

Level-1 ECAL prefiring

In 2016 and 2017 operation, a slowly developing shift in the shape of the ECAL pulses was observed [49]. This effect was manifested in an increasing offset in the timing calibration of the pulses related to the transparency loss of the ECAL crystals,

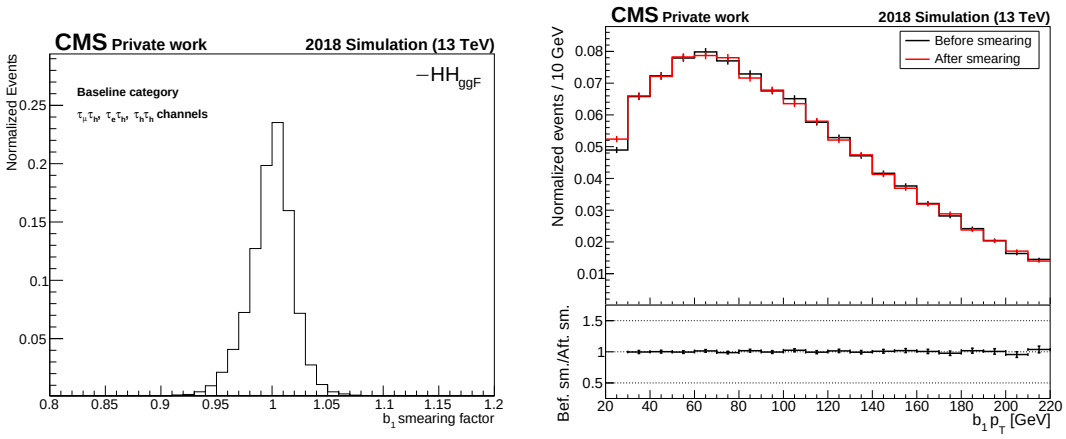


Figure 4.8: (Left) First b jet smearing factor distribution. (Right) First b jet p_T distribution before applying the jet smearing correction (black) and after applying it (red). In both cases, a 2018 HH ggF samples is considered.

mostly in the endcap. This offset was compensated offline by recalibrating the pulses, but was not corrected in the ECAL trigger primitives. With time, this effect ended up moving the endcap pulses to a time region where the bunch crossing assignment would be affected, leading to the Level-1 trigger system to *prefire*, i.e. accept the earlier collision in BX - 1, whereas the one in BX 0 is the one of interest. Since Level-1 trigger rules forbid two consecutive bunch crossings to fire, the actual event would not be triggered and therefore lost. This effect was not accounted for by the simulations, so additional weights are included to account for the probability of an event to prefire according to the p_T and η of forward jets and photons in the event. This weight is obtained as the product of the non prefiring probability of all jets and photons from the event:

$$w = \prod_{i \in \text{photons, jets}} (1 - \varepsilon_i^{\text{pref}}(p_T, \eta)). \tag{4.17}$$

The distribution of this weight for a 2016 HH ggF sample is shown in Fig. 4.9 (left).

In early 2018, the endcap timing delays in the ECAL front-end electronics were corrected, optimizing the pulse synchronization and solving these prefiring issues.

Trigger efficiency scale factors

As the simulations do not reproduce exactly the data, there are some discrepancies in the efficiencies obtained in the single-lepton, cross-lepton and VBF triggers used in the analysis. These discrepancies can be compensated by including some event weights.

For the $\tau_h\tau_h$ final state, trigger efficiencies and scale factors are measured using $Z \to \tau_\mu\tau_h$ events selected with a tag and probe technique, both in data and sim-

ulation. For the semi-leptonic final states, a similar technique is considered using $Z \to \mu\mu$ and $Z \to ee$ events. In these cases, the scale factors must take into account the logical OR of the single-lepton and cross-lepton triggers. Assuming the lepton and τ_h legs are independent, the efficiency for either simulation and data can be obtained as

$$\text{Eff} = \epsilon_L(1 - \epsilon_{\tau_h}) + \epsilon_l\epsilon_{\tau_h}, \tag{4.18}$$

where ϵ_L is the single lepton trigger efficiency, ϵ_l is the cross lepton trigger efficiency for the τ_e or τ_μ leg and ϵ_{τ_h} is the cross lepton trigger efficiency for the τ_h leg.

For the VBF trigger paths, efficiencies and scale factors are computed separately for the τ_h legs and the jet legs. For the τ_h legs, a tag and probe technique focused on $Z \to \tau_\mu\tau_h$ events is exploited. Both τ_h are considered independents, and their correspondent scale factors are obtained as a function of their p_T and decay mode. For the jet legs, events are required to satisfy the requirements of one of the $\tau_h\tau_h$ triggers considered and the additional requirements applied on the jets by the VBF trigger. The scale factors in this case are computed as a function of the transverse momenta of the two jets with the highest invariant mass and on the actual invariant mass. Note that these VBF trigger scale factors are only when at least one of the τ_h has a p_T between 25 and 40 GeV. If both τ_h have a $p_t > 40$ GeV, the $\tau_h\tau_h$ scale factors are considered instead.

The distribution of the combination of the trigger scale factors for a 2018 HH ggF sample is shown in Fig. 4.9.

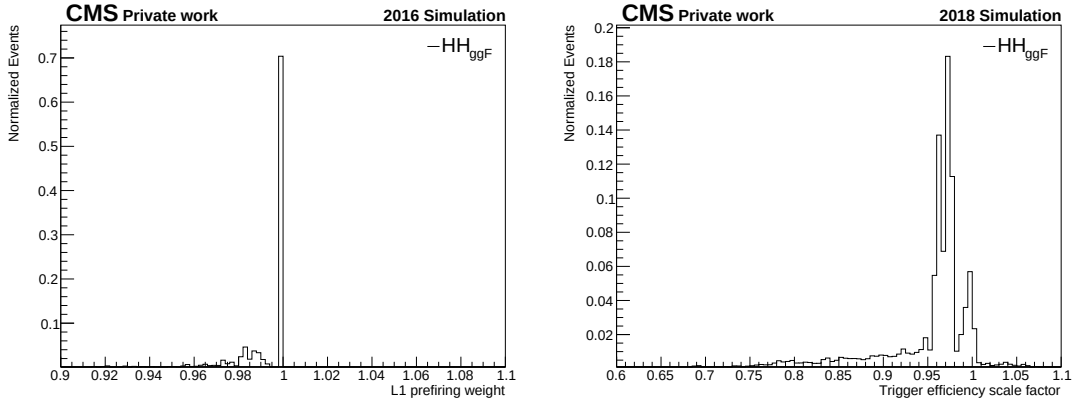


Figure 4.9: (Left) L1 prefiring weight for a 2016 HH ggF simulated sample. (Right) Trigger efficiency scale factor for a 2018 HH ggF simulated sample.

Trigger prescale weights

In order to account for the prescales used for data taking for some of the HLT triggers, the simulated events that fired only prescaled paths are corrected with weights obtained as the ratio between the integrated luminosity covered by the prescaled trigger over the total integrated luminosity collected during the year.

Only two triggers used in the analysis need these prescale weights: the VBF triggers used in both 2017 and 2018. In the first case, the VBF trigger was included and enabled once 14.37 fb^{-1} out of the total 41.5 fb^{-1} had already been collected. Then, simulated events that were selected as they satisfied this VBF trigger selection need to be weighted by around 0.65. Analogously, events that satisfied the 2018 VBF trigger selection are weighted by around 0.99.

b-tagging efficiency

To account for discrepancies in the b-tagging performance of the DeepJet algorithm between data and simulation, the discriminant distribution is corrected by applying weights to the simulated events. These weights are obtained as

$$w = \prod_{i}^{N} SF(D^i, p_T^i, \eta^i), \tag{4.19}$$

where D^i, p_T^i, and η^i stand for the discriminator score, the transverse momentum and the pseudorapidity of the jet and the product considers all jets satisfying the requirements described in Section 4.2.4. The scale factor for each jet are obtained as

$$SF(D^i, p_T^i, \eta^i) = \varepsilon^{\text{data}}(D^i, p_T^i, \eta^i)/\varepsilon^{\text{MC}}(D^i, p_T^i, \eta^i), \tag{4.20}$$

where $\varepsilon^{\text{data}}(D^i, p_T^i, \eta^i)$ and $\varepsilon^{\text{MC}}(D^i, p_T^i, \eta^i)$ are the b-tagging efficiencies in data and simulation, respectively [86]. In simulation, the b-tagging efficiency is defined as the number of identified b jets in a particular bin of the DeepJet discriminator, jet p_T and jet η with respect to the total number of b jets. The number of b jets is measured by matching the jets with the generated hadrons. In data, the tagging efficiency is measured using samples pure in b jets; e.g. a sample with high p_T jets with a muon inside them (a topology usually found in heavy-flavour jets), or a di-lepton plus jets sample, expected to include b jets coming from top quark decays. To account for possible misidentification of jets in real data, the misidentification probability obtained from simulation is subtracted from the tagging efficiency in data. This misidentification probability is obtained as the number of jets misidentified as b jets and matched to generated non-b jets out of the total jets matched to generated non-b jets.

Fig. 4.10 shows the distribution of the b-tag efficiency weight for a 2018 HH ggF sample. As stated before, these additional weights intend to only modify the shape of distribution, not the expected event yields. As only applying these weights could lead to modification on the total event yields, the final distributions are corrected by scaling with the ratio of the sum of event weights without considering the b-tag weight and considering it.

Lepton identification and reconstruction scale factors

Regarding the identification of τ_h, the DeepTau algorithm shows small differences when running it on data or simulated events. To account for these differences, some

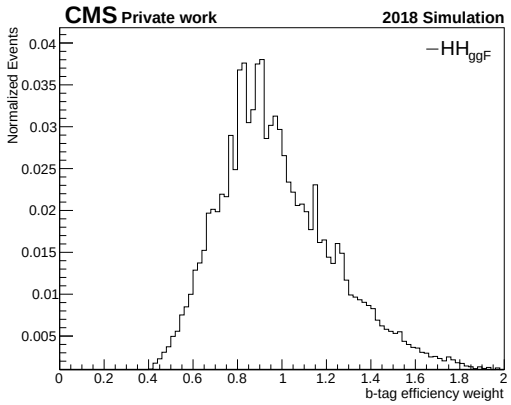


Figure 4.10: b-tag efficiency weight for a 2018 HH ggF simulated sample.

scale factors are included depending on if a genuine τ_h or a misidentified particle is selected as the τ_h [87].

For genuine τ_h, scale factors are computed with a tag and probe technique by selecting $Z \to \tau\tau$ events decaying into a $\tau_\mu\tau_h$ final state. Several quality cuts are applied to both μ and τ_h to increase the signal purity. Two sets of scale factors are then computed, either binned in the τ_h p_t or in its decay mode. In the bb$\tau\tau$ analysis, the former is used for all τ_h, modifying the event weight.

For genuine electrons and muons misidentified as τ_h, scale factors are computed with a tag and probe technique by selecting $Z \to ee$ and $Z \to \mu\mu$ events, respectively, where one of the leptons is correctly reconstructed and the other is reconstructed as a τ_h. For electrons, scale factors are provided split into barrel and endcap, while for muons, binned as a function of η.

To account for possible disagreements between data and simulation regarding the reconstruction and identification of muons and electrons in the $\tau_\mu\tau_h$ and $\tau_e\tau_h$ channels, specific scale factors binned in the lepton p_t and η are provided. Scale factors are obtained with tag and probe techniques, selecting $Z \to ee$ and $Z \to \mu\mu$ events.

Fig. 4.11 shows the lepton reconstruction and identification scale factors for the three $\tau\tau$ channels, i.e. the product of the scale factors corresponding to the final state particles of the decay channel (either DeepTau, electron, or muon scale factors).

Tau Energy Scale

The correction to the τ_h energy scale is defined by the deviation of the average reconstructed τ_h energy from the generator-level energy of the visible τ_h decay products [87, 88]. The data-to-simulation correction factor is obtained by fitting distributions of observables sensitive to the τ_h energy scale in $Z/\gamma^* \to \tau\tau \to \tau_e\tau_h,\ \tau_\mu\tau_h$, such as the m_{τ_h} and the mass of the $l\tau_h$ system. These corrections modify the energy and momentum on the τ_h, modifying also the shape of the final discriminants. Four

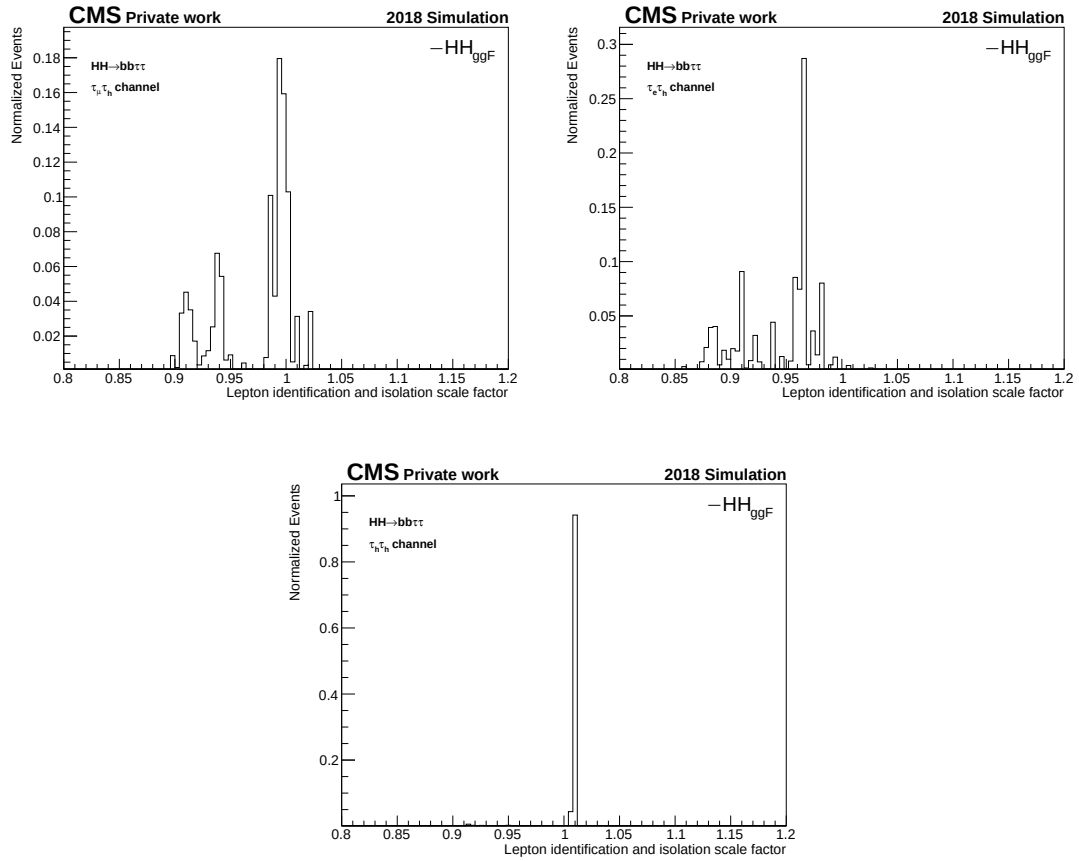


Figure 4.11: Lepton reconstruction and identification scale factors in the $\tau_\mu\tau_h$ channel (top left), $\tau_e\tau_h$ channel (top right), and $\tau_h\tau_h$ channel (bottom) for a 2018 HH ggF simulated sample.

independent corrections are provided, one for each decay mode used in the analysis, taking values of up to 2%. Fig. 4.12 shows the distributions of τ_h p_T and mass before and after applying the tau energy scale corrections. Very subtle differences are observed between with and without the corrections.

4.6 Distributions after the baseline selection

The distributions of some of the main variables after the baseline categorization are shown in Fig. 4.13, 4.14, 4.15, and 4.16, for the data and backgrounds in 2018. All corrections described in Section 4.5 are applied to simulated events. In all distributions there is a good agreement (in general within the statistical uncertainty range) between data and the expected backgrounds.

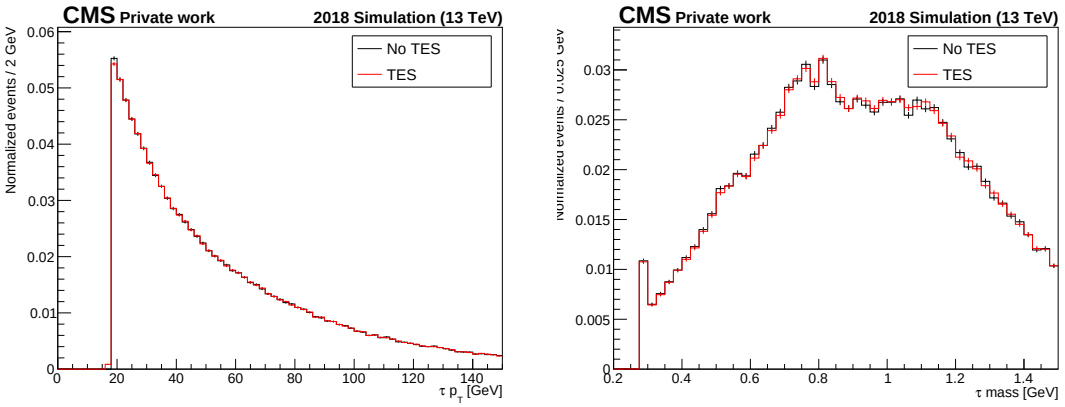


Figure 4.12: τ_h p_T (left) and mass (right) distributions for a 2018 HH ggF sample before applying the tau energy scale correction (black) and after applying it (red).

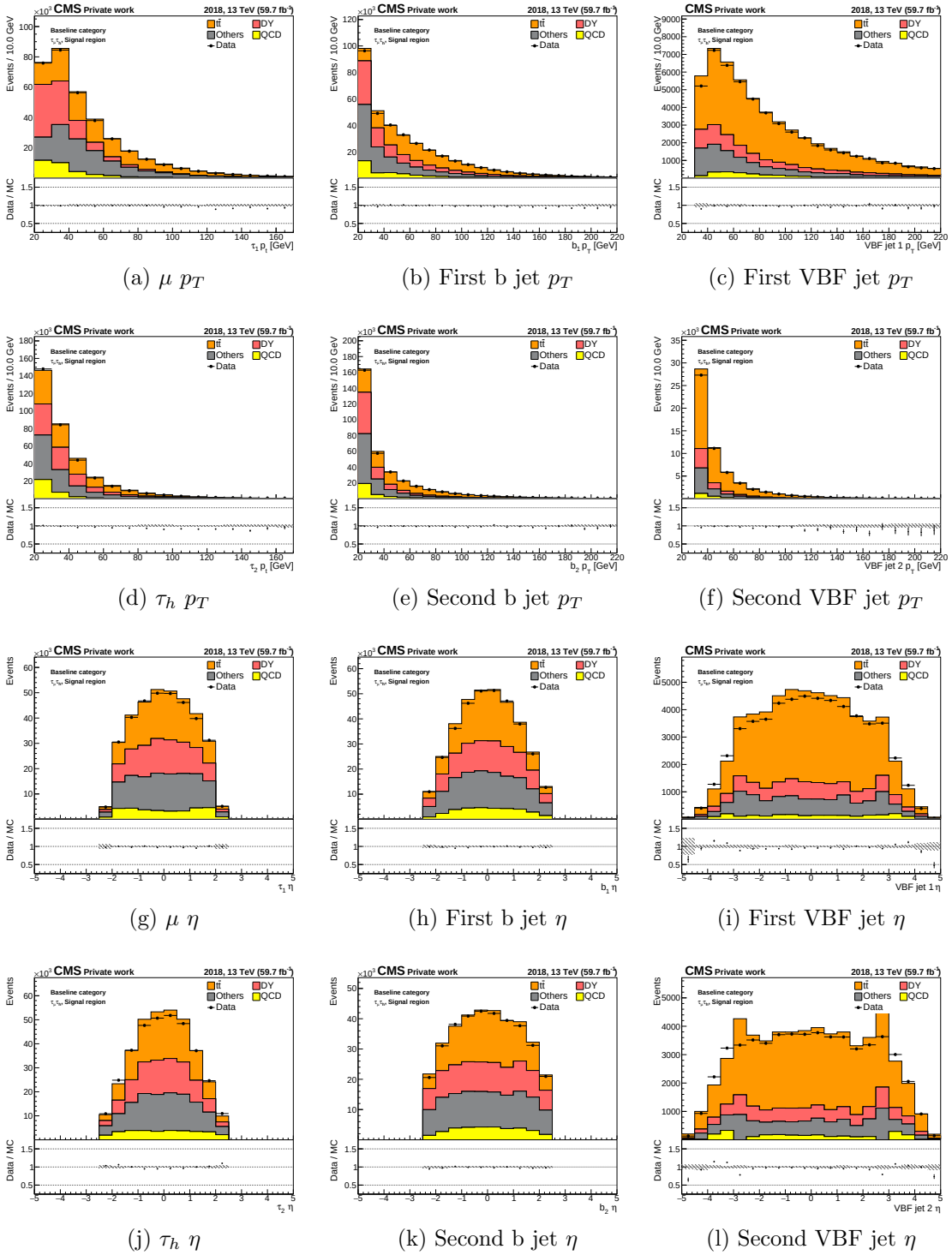
4.7 Multi-class classification for VBF categorization

In the multi-class classification strategy, machine learning techniques are used to assign probability estimates for an event to belong to categories associated to any of the relevant physics processes under consideration. The presence of several categories allows to study different signals with different topologies and to constrain nuisance parameters associated to backgrounds with large systematic uncertainties. By doing so, the analysis sensitivity can be enhanced.

In the HH $\rightarrow$ bb$\tau\tau$ search, both of these characteristics are present. On the one hand, distinguishing between ggF and VBF production helps to study the κ_{2V} coupling strength, which is only accessible in the VBF production. On the other hand, the analysis is dominated by backgrounds, among the ones we find QCD multijet, Drell-Yan processes and processes involving the t quark ($t\bar{t}$ and $t\bar{t}H$). The first two have a big contribution due to its huge cross section, while the latter have a similar final state to the expected signal signature. Therefore, a strategy that discriminates between the signals and the different backgrounds could help increasing the analysis sensitivity.

4.7.1 Classification strategy

Fig. 4.17 shows the concept of the event categorization using DNNs performing a multi-class classification, which can be described as follows. Per event, a predefined set of variables is used as input for a neural network whose architecture and weights are pre-trained and optimized. The network evaluates the variables and outputs for each event a vector of nine floating point numbers, each of them correspondent to the probability estimate for the event to be produced by the physical processes under consideration. In the HH $\rightarrow$ bb$\tau\tau$ analysis, nine processes are defined. Three of them



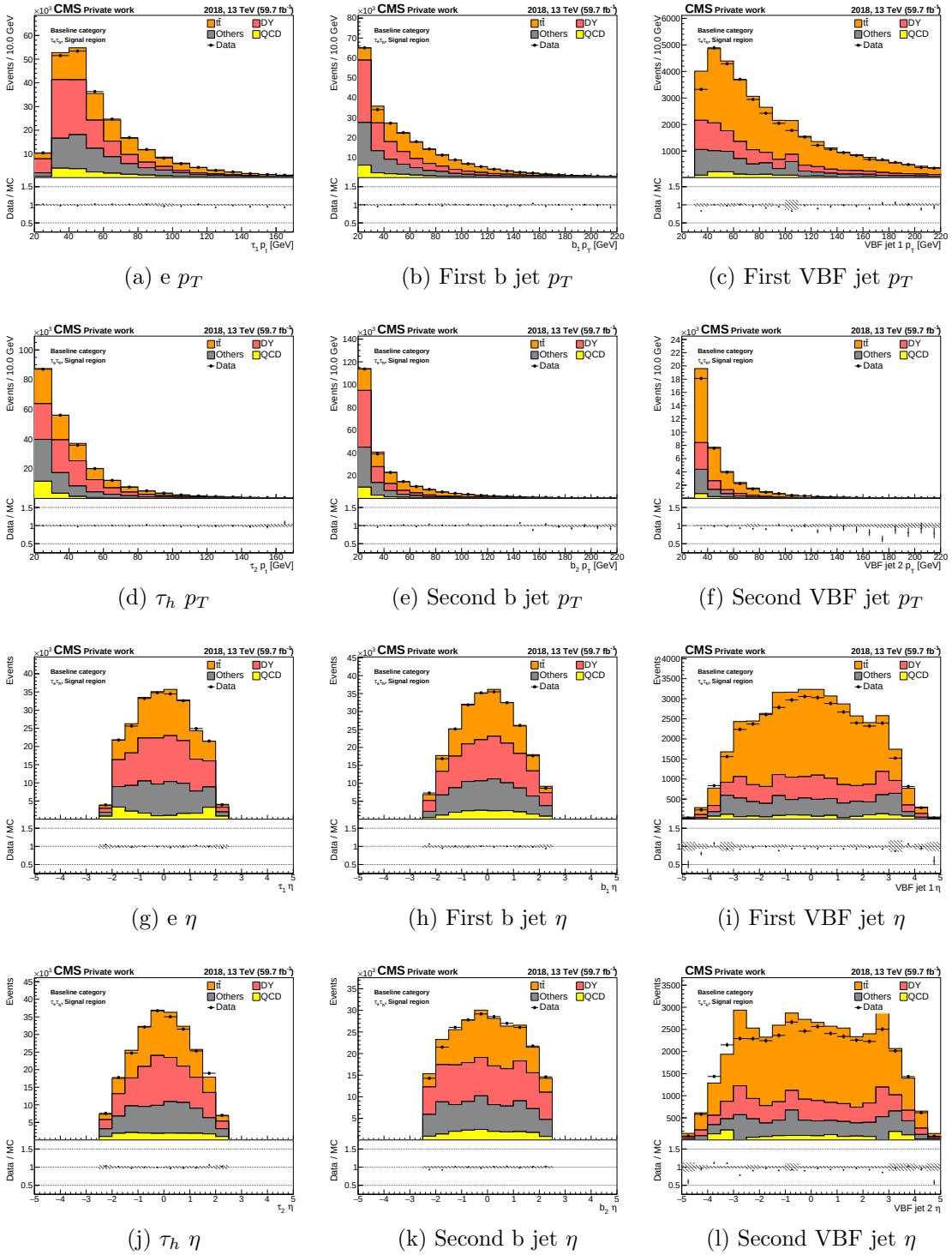
(a) μ p_T (b) First b jet p_T (c) First VBF jet p_T

(d) τ_h p_T (e) Second b jet p_T (f) Second VBF jet p_T

(g) μ η (h) First b jet η (i) First VBF jet η

(j) τ_h η (k) Second b jet η (l) Second VBF jet η

Figure 4.13: Leptons and jets variable distributions in the baseline category, $\tau_\mu\tau_h$ channel, and 2018. The first b and VBF jet candidates out of the corresponding jet pairs are the ones with the highest p_T.



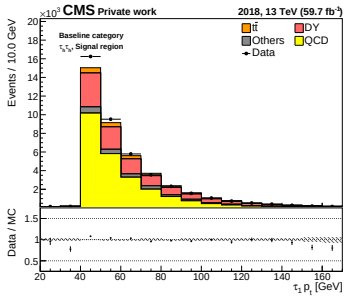
(a) e p_T (b) First b jet p_T (c) First VBF jet p_T

(d) τ_h p_T (e) Second b jet p_T (f) Second VBF jet p_T

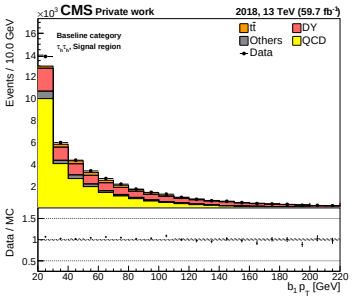
(g) e η (h) First b jet η (i) First VBF jet η

(j) τ_h η (k) Second b jet η (l) Second VBF jet η

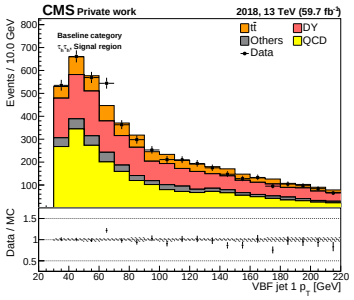
Figure 4.14: Leptons and jets variable distributions in the baseline category, $\tau_e\tau_h$ channel, and 2018. b and VBF jet candidates are ordered by p_T.



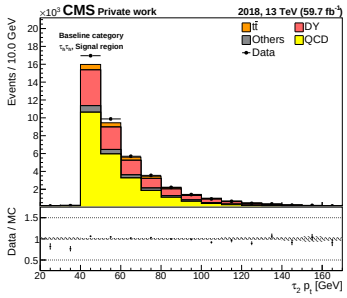
(a) First τ_h p_T



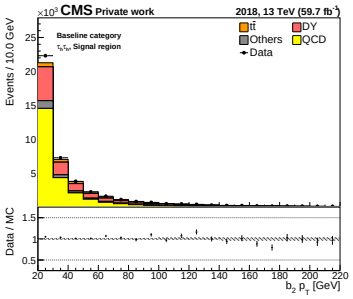
(b) First b jet p_T



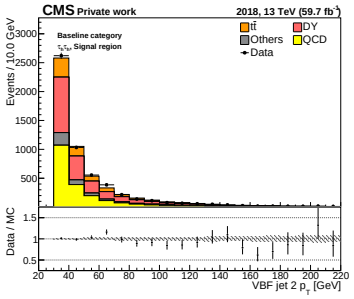
(c) First VBF jet p_T



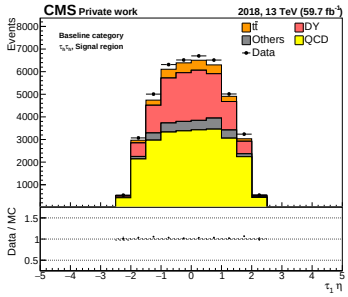
(d) Second τ_h p_T



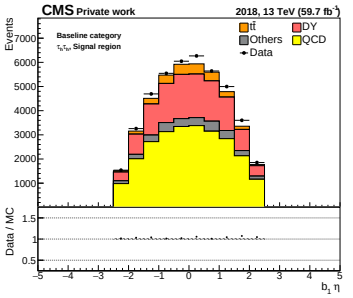
(e) Second b jet p_T



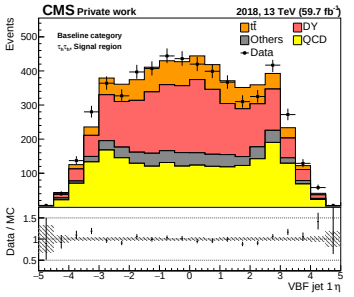
(f) Second VBF jet p_T



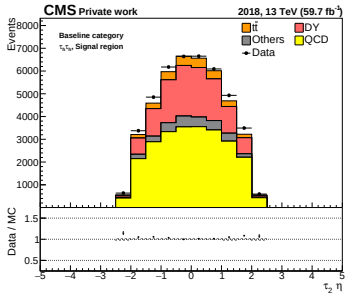
(g) First τ_h η



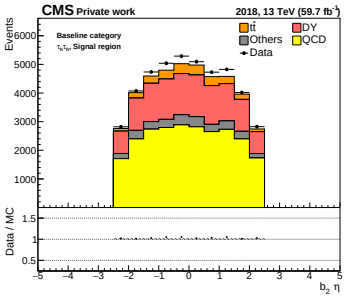
(h) First b jet η



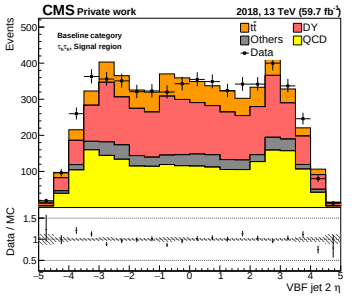
(i) First VBF jet η



(j) Second τ_h η

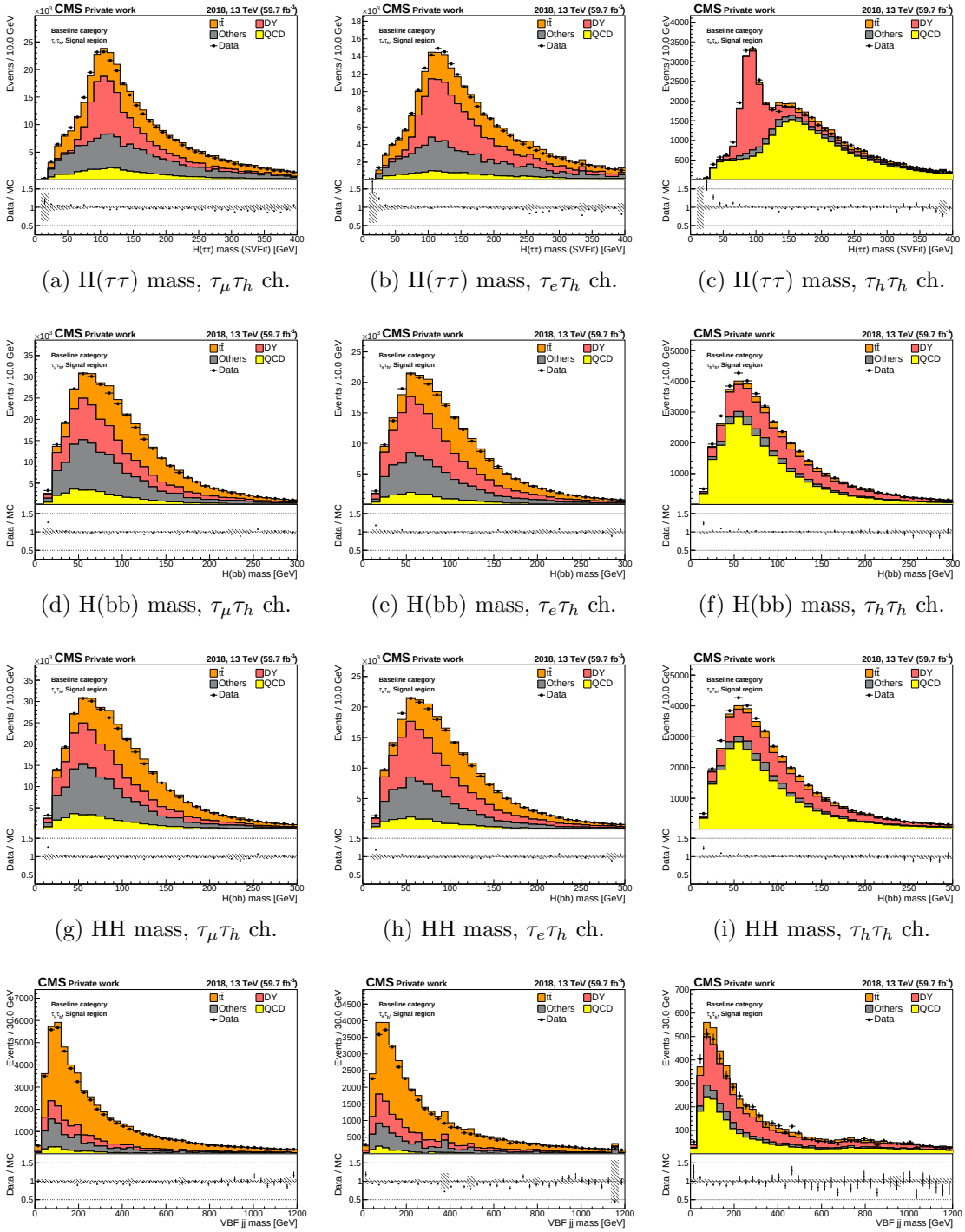


(k) Second b jet η



(l) Second VBF jet η

Figure 4.15: Leptons and jets variable distributions in the baseline category, $\tau_h\tau_h$ channel, and 2018. First τ_h corresponds to the most isolated τ_h of the $\tau\tau$ pair, while b and VBF jet candidates are ordered by p_T.



(a) H($\tau\tau$) mass, $\tau_\mu\tau_h$ ch. (b) H($\tau\tau$) mass, $\tau_e\tau_h$ ch. (c) H($\tau\tau$) mass, $\tau_h\tau_h$ ch.

(d) H(bb) mass, $\tau_\mu\tau_h$ ch. (e) H(bb) mass, $\tau_e\tau_h$ ch. (f) H(bb) mass, $\tau_h\tau_h$ ch.

(g) HH mass, $\tau_\mu\tau_h$ ch. (h) HH mass, $\tau_e\tau_h$ ch. (i) HH mass, $\tau_h\tau_h$ ch.

(j) VBF (j_1, j_2) mass, $\tau_\mu\tau_h$ ch. (k) VBF (j_1, j_2) mass, $\tau_e\tau_h$ ch. (l) VBF (j_1, j_2) mass, $\tau_h\tau_h$ ch.

Figure 4.16: H($\tau\tau$), H(bb), HH, and VBF (j_1, j_2) system mass distributions in the baseline category, $\tau_\mu\tau_h$, $\tau_e\tau_h$, and $\tau_h\tau_h$ channels, in 2018.

correspond to signal processes: HH ggF, HH VBF produced with $\kappa_{2V} = 1$, and a mixture of HH VBF produced with $\kappa_{2V} = 0$ and $\kappa_{2V} = 2$. Splitting the HH_{VBF} contribution into a SM- and BSM-like component comes from the fact that the topology of the event differs a lot for different values of κ_{2V}. Therefore, considering only the HH VBF produced as in the SM could lead to an artificial bias of the network towards the SM topology, reducing the sensitivity for excluding BSM scenarios. Regarding the background processes, the ones with the highest contribution (that were introduced before) are included. The contribution from $t\bar{t}$ is splitted according to the decays of the two W bosons into di-leptonic (DL), semi-leptonic (SL) and fully-hadronic (FH). $t\bar{t}H$ is also included in two different samples, with the H decaying into $b\bar{b}$ or $\tau\bar{\tau}$, while an inclusive Z+jets sample is used to model the DY process. The QCD background, as described in Section 4.4.1, is obtained in a data-driven way, so it can not be used as an output value by the DNN. Therefore, the six background processes previously described are the only ones we will consider.

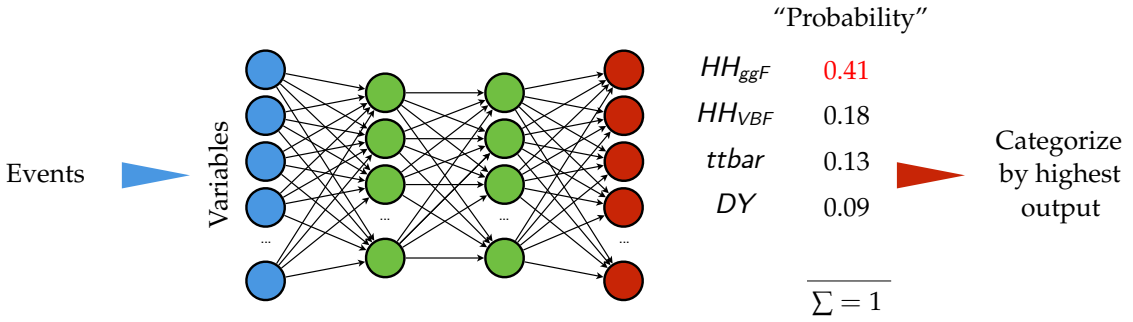


Figure 4.17: Concept of the event categorization approach using a multi-class DNN. Every event is given nine floating point values, each of them expressing the probability of the event of being originated by the associated process. An event is categorized by the process class associated to the output with the highest value (HH_{ggF} in the given example).

Each event, after obtaining its corresponding output values, can be categorised by selecting the process whose corresponding node obtained the highest value. Nonetheless, the final categorization used in the analysis for signal extraction is obtained by merging some of the outputs: the three $t\bar{t}$, the two $t\bar{t}H$ and the two HH VBF respectively. This way, only five distinct outputs remain: HH VBF, HH ggF, $t\bar{t}$, $t\bar{t}H$ and DY. These merging could have been done initially by training and validating with 5 samples instead of 9. However, splitting the samples provides the network additional information on the underlying process, which can lead to a better training performance.

4.7.2 Network architecture

A representation of the neural network architecture is shown in Fig. 4.18. First, input variables are splitted into four-vector components of particles and higher-level variables. Four-vector variables are passed to a Lorentz Boost Network (LBN) [89].

The purpose of this network is to extract high-level variables from four-vectors of particles and their combinations as part of the learning process. Together with the remaining input higher-level variables, they are forwarded to a standard, fully connected neural network that actually does the multi-class classification.

The final architecture used is the result of a hyperparameter optimization, and it will be described in Section 4.7.4.

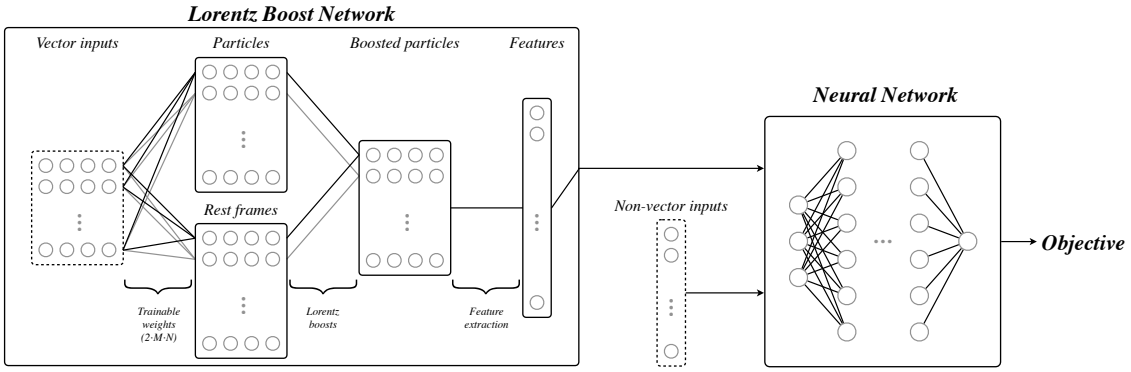


Figure 4.18: Architecture of the multi-class classification network. Four-vector input features are passed through a Lorentz Boost Network for extracting high-level feature representations. These, together with the rest of input features, are forwarded to the actual neural network in order to do the classification of events.

4.7.3 Input variables

To perform the training and validation of the DNN, events from 2016, 2017, and 2018 passing the VBF inclusive category selection (described in Section 4.3.3) are considered. Table 4.14 summarises all variables considered as input to the neural network. Most of them are four-vector variables from different objects selected during the analysis: lepton (e, μ or τ_h) candidates to be the decay product of one of the H, b jet candidates coming from the decays from the other H, VBF jets, additional central jets not selected as the b jet candidates, forward jets not selected as the VBF jet candidates, MET and the H $\rightarrow$ bb and H $\rightarrow \tau\tau$ candidates. Additionally, some higher level features are included, such as the deep flavour and HH-btag scores for the b, VBF and additional central jets, and categorical features to provide information about the year of the input sample (so differences in detector alignment, trigger configurations or tuning parameters of the simulation can be inferred by the network) or the decay channel of the selected $\tau\tau$ pair. The distributions of some of the most important variables (selected by permutation importance [90]) are shown in Fig. 4.19. The distributions from all years and channels are shown in Appendix A. A good agreement is observed between data and background in all distributions, validating its use as input for the neural network.

Variable name(s)	Description
is_201{6,7,8}	Flag denoting the campaign / year of input events.
is_{etau,mutau,tautau}	Flag denoting reconstructed lepton channel.
lep{1,2}_{e,pt,eta,phi}	Four-vector components of the two leptons.
bjet{1,2}_{e,pt,eta,phi}	Four-vector components of the two b-tagged jets.
bjet{1,2}_{deepflavor,hhbtag}	b-tag and HH-b-tag of the two b-tagged jets.
vbfjet{1,2}_{e,pt,eta,phi}	Four-vector components of the two VBF jets.
vbfjet{1,2}_{deepflavor,hhbtag}	b-tag and HH-b-tag of the two VBF jets.
ctjet{1,2,3}_{e,pt,eta,phi}	Four-vector components of the three additional central jets.
ctjet{1,2,3}_{deepflavor,hhbtag}	b-tag and HH-b-tag of the three additional central jets.
fwjet{1,2}_{e,pt,eta,phi}	Four-vector components of the two additional forward jets.
met_{pt,phi}	Missing transverse energy and azimuthal direction.
bh_{e,pt,eta,phi}	Four-vector components of the reconstructed $H_{b\bar{b}}$ candidate.
tauh_{e,pt,eta,phi}	Four-vector components of the reconstructed $H_{\tau\bar{\tau}}$ candidate.

Table 4.14: Names and descriptions of variables considered as inputs to the neural network.

4.7.4 Training procedure and network optimization

In order to exploit the full amount of events without introducing a potential bias on the DNN outputs due to overtraining, a two-fold cross validation approach is performed: the available simulated events are divided into two independent datasets, the ones with *even* event numbers and the ones with *odd* event numbers. This way, the network trained and validated with even event numbers will be used for infering the DNN outputs for the events with odd event numbers and vice versa. On each dataset, a dedicated training is performed using 70% for the actual training, while 30% is used to perform the immediate cross validation. To increase the statistical robustness, ensembles of ten networks with different random seeds are trained and merged with the same weight after each particular training finishes, so the final outputs of the DNN inference are the average of the 10 DNN outputs.

Training is performed with the TENSORFLOW [91] algorithm interfaced with KERAS [92] and independently for even and odd event numbers. In fact, a hyper-parameter optimization is performed for each network via a grid-based scanning procedure, so the networks for even and odd event numbers could end up being totally different. The results of these optimizations, however, lead to only slightly different networks. In both cases, the optimal network contains four hidden layers, each with 128 units and tanh as activation function, and an output layer with nine units subject to a softmax activation function (so the DNN scores take values between 0 and 1 and sum 1). The training with even (odd) event numbers was performed using the Adam optimizer with a learning rate of 0.001 (0.006), an L_2 normalization factor λ of $5(2) \times 10^{-5}$ and a common batch size of 1024, equally populated by the nine physics processes under investigation. During the training, the method of batch normalization is applied as well as a random unit dropout with a probability of 5%. The loss function in both cases is constructed as a sum of three

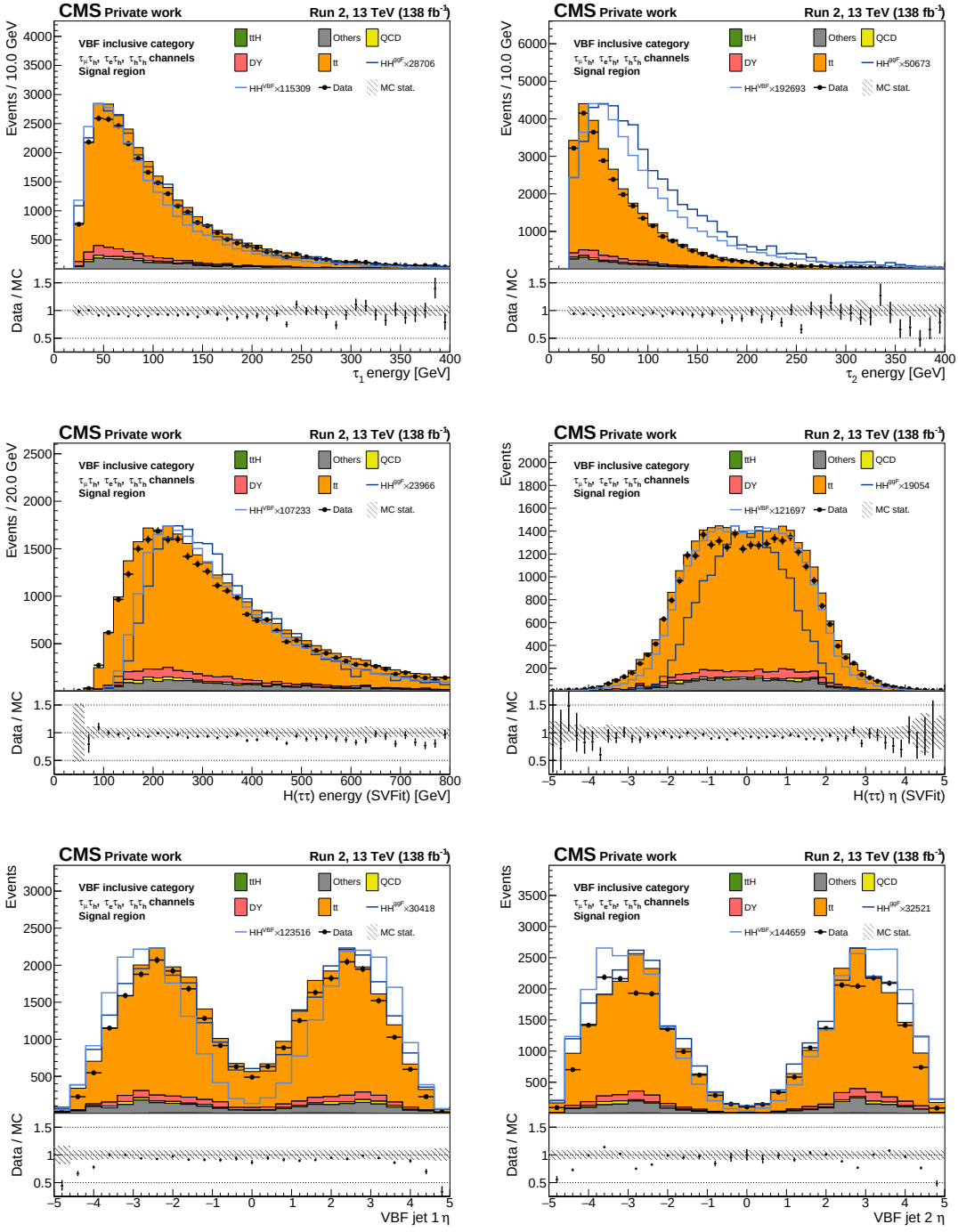


Figure 4.19: Distributions of the energies of the first lepton (μ, e, or the most isolated τ_h depending on the $\tau\tau$ decay channel) (top left) and the second lepton candidate (top right), the energy (medium left) and pseudorapidity (medium right) of the H$\rightarrow \tau\tau$ candidate, and the pseudorapidity of the first (bottom left) and second (bottom right) VBF jets for the VBF inclusive category. Each distribution contains the full Run 2 data sets and the three $\tau\tau$ decay channels.

individual loss terms,

$$L(y(x), y_G(x), \hat{y}, \hat{y}_G)_{\text{CE}} = \frac{1}{N}\sum_{n}^{N} -\hat{y}_n \log y_n(x) \;+\; \frac{1}{N}\sum_{n}^{N} -\hat{y}_{G,n} \log y_{G,n}(x) \;+\; L_2(\lambda), \tag{4.21}$$

where the first two terms represent categorical cross entropy (CE) functions and the third term a typical L_2 normalization (to keep the model weights as small as possible) factorized by λ. The first CE is applied to all nine output classes and compares the prediction $y_n(x)$ of an event n given the inputs x with the true value $\hat{y}_n$. For the second CE, outputs are combined into a signal class (all HH processes) and a background class. This increases the relative importance of signal events and balances the importance of an HH event of a given class being classified as an event from a different signal class (and analogously for background events). This second CE term is often called *grouped* cross entropy.

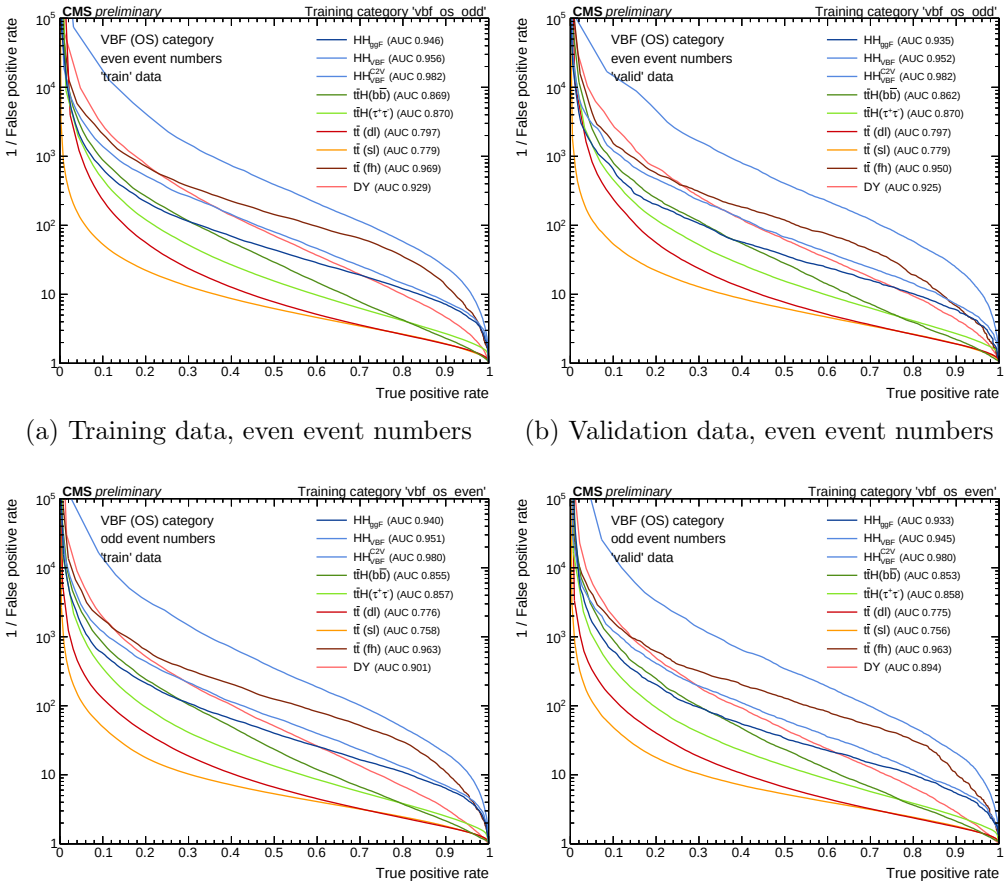
4.7.5 Training and validation results

To study the performance of a classifier, *Receiver Operating Characteristic* (ROC) curves are commonly used metrics. These curves plot the true positive rate (i.e. percentage of events correctly assigned to their corresponding processes) against the false positive rate (i.e. percentage of events wrongly assigned to a different process) at various thresholds settings. Fig. 4.20 shows the ROC curves of the trained networks with even and odd numbers separately for training and validation events. No significant differences between the training and validation datasets, then overtraining has been sufficiently suppressed. The separation performance of the training networks can be seen in the large area-under-curve (AUC) values, above 0.93 for HH signals and between 0.75-0.95 for backgrounds.

The efficiency of the categorization achieved by this approach can be seen in the so-called confusion matrices shown in Fig. 4.21. Numbers are normalized per row, and represent the probability of an event to have its process correctly predicted or not. Accuracies are denoted by the diagonal, while off-diagonal values correspond to confusion rates. Note how most of the signal events stay in the signal categories, while very little background contamination can be observed in these categories. A sixth row has been added in the matrices containing QCD events obtained by the data-driven method. Even if the training was not performed with the QCD process as input, QCD events get classified mostly in background categories, with little contamination affecting the HH_{VBF} category.

4.8 Signal extraction

To perform the signal extraction, a different DNN was developed in order to identify $\text{HH} \to \text{bb}\tau\tau$ events. The approach used to train this DNN [93] follows the one used in the CMS di-Higgs HL-LHC projection analysis for the $\text{bb}\tau\tau$ channel [94]. The goal of the DNN is to classify events as originating either from signal (SM $\text{HH} \to \text{bb}\tau\tau$



(a) Training data, even event numbers

(b) Validation data, even event numbers

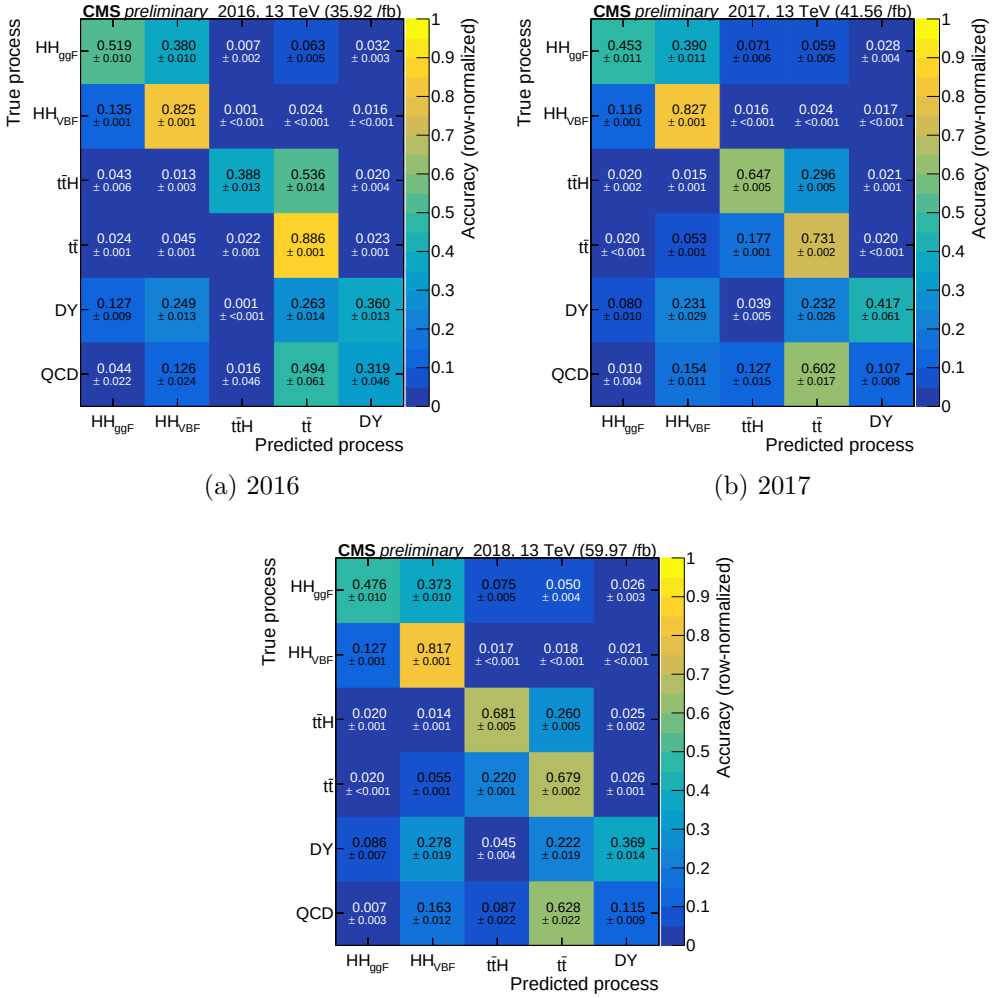
(c) Training data, odd event numbers

(d) Validation data, odd event numbers

Figure 4.20: ROC curves showing the relation between true and false positive rates, evaluated for a particular process against all others separately on the training (left) and validation (right) data and for even (top) and odd (bottom) event numbers. Corresponding area-under-curve values are shown in the legend.

produced via ggF) or background processes by assigning to them a single prediction score. Score values closer to one indicates "signal-like" events, while closer to zero, "background-like" events. The final DNN distributions obtained by inferring this network prediction in each category, $\tau\tau$ decay channel and year are used in the signal extraction fit.

Training is performed with the PyTorch [95] algorithm interfaced with LUMIN [96] using all MC-driven backgrounds as background processes and SM HH $\rightarrow$ bb$\tau\tau$ as signal. Simulated events used for training are split in two halves, the ones with even event number and the ones with odd event number. A pair of discriminators are then trained with each half of the data. At inference time, the discriminators are used to predict the classes of events in the complementary halves of the data, i.e. the one trained with even event numbers for the odd event numbers, and vice versa. This way, all MC data can be used for the training without adding a bias on the predictions that would arise if only one network was used and the same data



(a) 2016

(b) 2017

(c) 2018

Figure 4.21: Row-normalized process classification accuracies and confusion rates, evaluated for the three data-taking years. Note that the two HH VBF, the two $t\bar{t}H$ and the three $t\bar{t}$ processes are merged following the categorization strategy described in Section 4.7.1.

used for training was also used for inference. To further increase the statistical robustness, each discriminator consists of 10 neural networks, each trained with a different starting random seed.

For the training, a total of 26 features (selected among over a 100) are used as input to the neural network. Both categorical (year of data taking, the $\tau\tau$ decay mode or the number of VBF jet candidates) and object features are used. From the latter, the most important features (ranked by permutation importance) are the DeepJet scores of the b jets, the invariant masses of the bb, $\tau\tau$, and HH systems and kinematic variables of the reconstructed particles. The distributions of the masses of the bb, $\tau\tau$, and HH systems are shown in Fig. 4.22 and 4.23 in four analysis categories: Resolved, 1 b-tag and 2 b-tag, Boosted, and VBF inclusive. The three

years from Run 2 data taking and the three $\tau\tau$ decay channels are displayed together in each distribution. Each plot shows the data, the background contributions and two SM signals, HH ggF and VBF, scaled so their maximum is close to the one from the sum of backgrounds. All corrections described in Section 4.5 have been applied to the simulated events. In all distributions, there is a good agreement between the data and the backgrounds, in general within the statistical uncertainty ranges.

4.8.1 VBF signal extraction

After the development of the DNN devoted to signal extraction and the multi-class classification strategy for VBF analysis, a complete study was performed in order to determine the signal extraction strategy for the VBF signal. Four different strategies were proposed:

1. Consider the output score of the VBF node from the multi-class classification approach in the VBF inclusive category.
2. Consider the signal extraction DNN (from now onwards, DNN) output score in the VBF inclusive category, i.e. avoid using the multi-class classification approach.
3. Consider the output score of the VBF node from the multi-class classification approach but splitting the VBF inclusive category into five subcategories defined by the processes considered in the multi-class classification: HH ggF, HH VBF, $t\bar{t}$, $t\bar{t}H$, and $Z/\gamma^* + jets$. Events are categorised by selecting the process whose corresponding node obtained the highest value.
4. Consider the DNN output score in the five VBF subcategories.

The output score in the categories defined by the different strategies is shown in Fig. 4.24 (in strategies 3. and 4., only the output score in the HH VBF subcategory is displayed, as it should give the largest sensitivity in the VBF analysis). Some observations can be extracted from this Figure:

- The DNN output score from the HH VBF process is contained mostly in the last bin from the distributions, where the background event yield is the smallest. The output score of the VBF node is more uniformly distributed (although mostly contained in the last bins of the distributions).
- The categorisation strategy defined by the multi-class classification approach helps drastically reducing the amount of background in the category, which could provide more sensitivity.

In order to estimate the sensitivity that could be provided by the four strategies described, the binned significance is considered, obtained as the ratio between the signal yield (in this case, HH VBF) and the square root of the total background yield

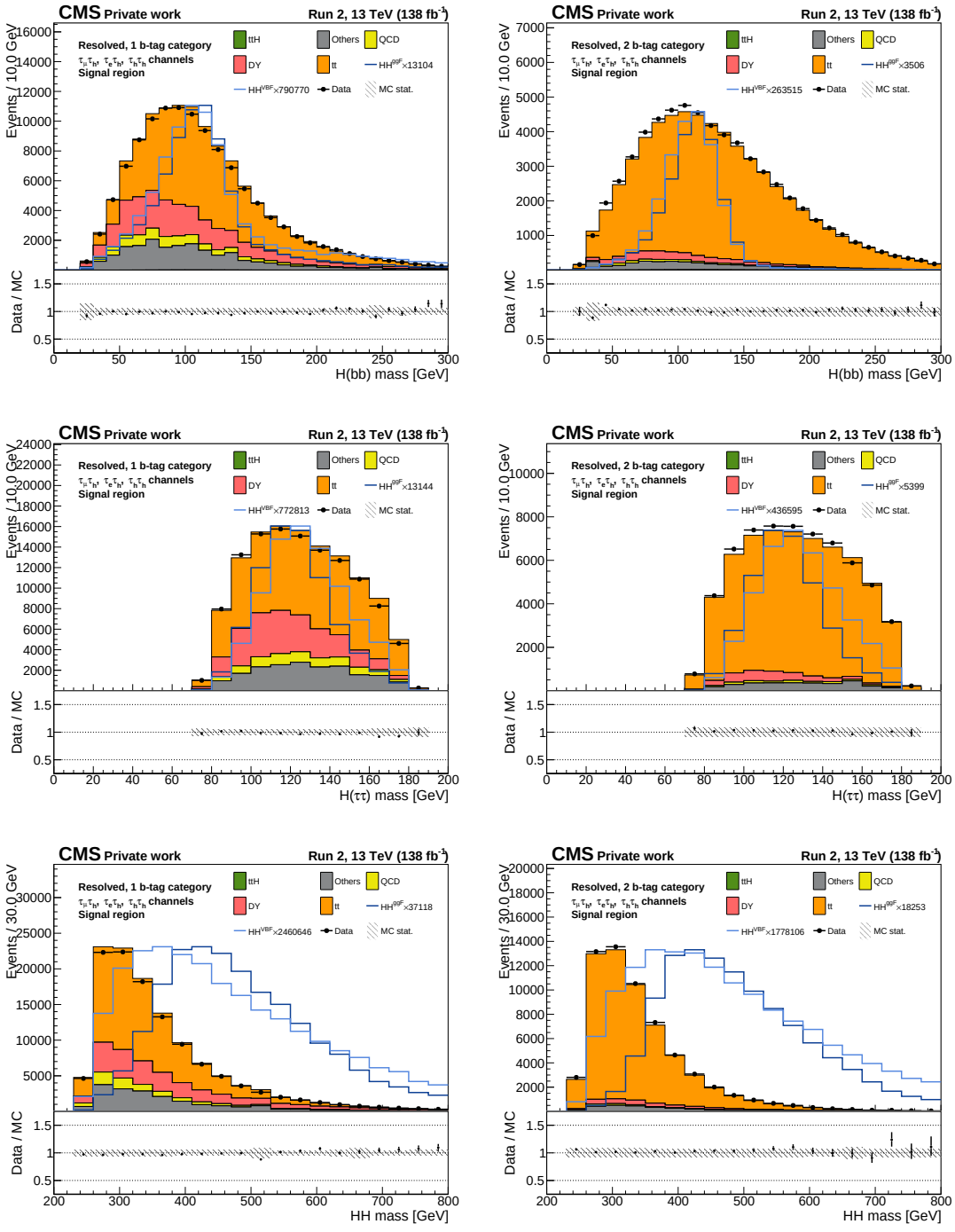


Figure 4.22: Distributions of the masses of the H → bb candidate (top), H → $\tau\tau$ candidate (middle), and HH system (bottom) for the Resolved, 1 b-tag (left), and Resolved, 2 b-tag (right) analysis categories. Each distribution contains the full Run 2 data sets and the three $\tau\tau$ decay channels.

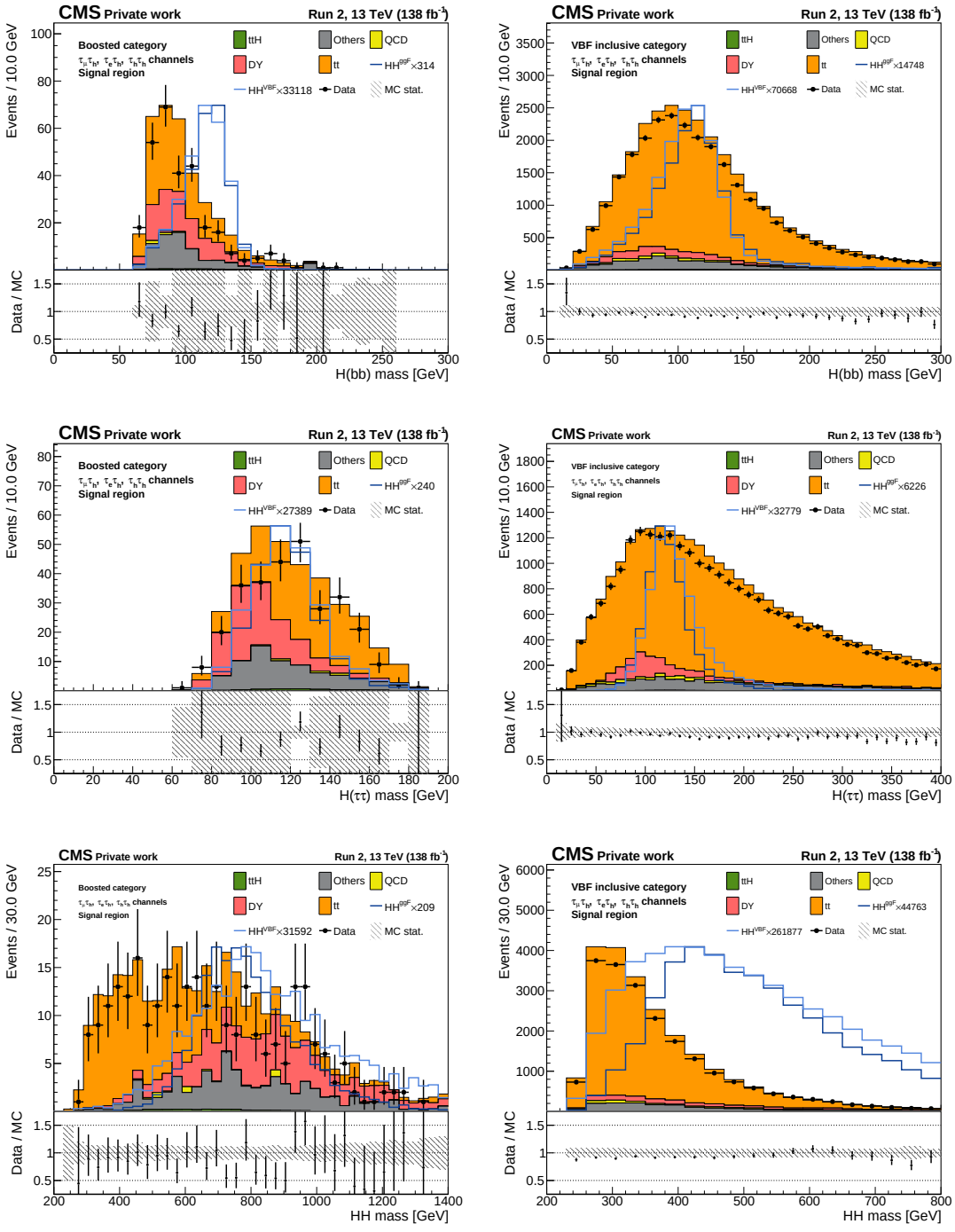


Figure 4.23: Distributions of the masses of the H → bb candidate (top), H → $\tau\tau$ candidate (middle), and HH system (bottom) for the Boosted (left), and VBF inclusive (right) analysis categories. Each distribution contains the full Run 2 data sets and the three $\tau\tau$ decay channels.

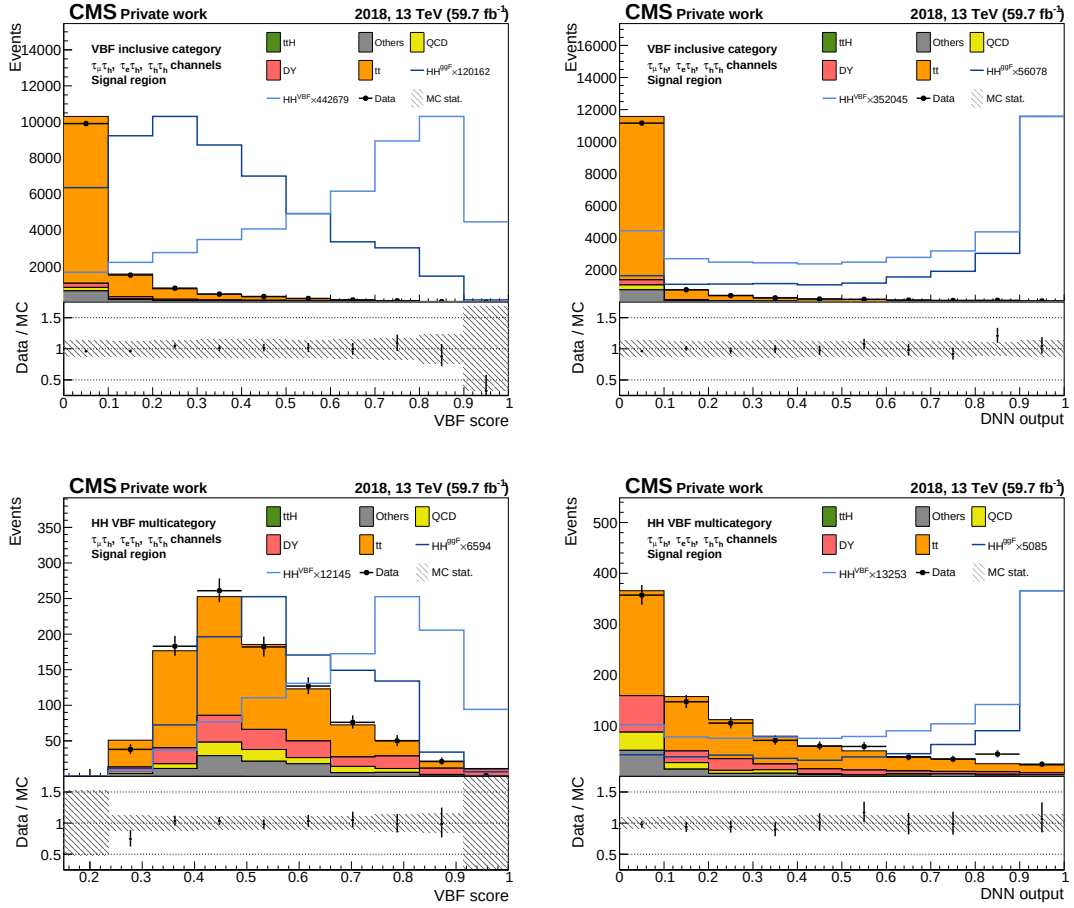


Figure 4.24: Distributions of the output score of the VBF node in the VBF inclusive category (top left) and the HH VBF subcategory (bottom left) and distributions of the DNN output score in the VBF inclusive category (top right) and the HH VBF subcategory (bottom right). Results are shown for 2018 and the three $\tau\tau$ decay channels together.

in a particular bin. With this binned significance, two quantities are used to measure the sensitivity: the maximum binned significance over all bins in the distribution and the sum of the binned significances from all bins. The highest their values, the more sensitivity the distribution should provide. The values obtained for 2018 split by $\tau\tau$ decay are shown in Table 4.15. In general, the largest values are obtained by the fourth strategy, i.e. the DNN output score obtained in the subcategories defined by the multi-class classification approach. In fact, only the HH VBF is considered here, but the other subcategories could also provide more sensitivity: some signal could be allocated in the HH ggF subcategory, and the background subcategories could help constraining the values of the systematic uncertainties.

In conclusion, the DNN developed for signal extraction will be used in the eight orthogonal categories (Resolved 1 b-tag, Resolved 2 b-tag, Boosted and the five VBF subcategories).

	$\tau_\mu\tau_h$ channel		$\tau_e\tau_h$ channel		$\tau_h\tau_h$ channel	
	$\max_i S_i/\sqrt{B_i}$	$\sum_i S_i/\sqrt{B_i}$	$\max_i S_i/\sqrt{B_i}$	$\sum_i S_i/\sqrt{B_i}$	$\max_i S_i/\sqrt{B_i}$	$\sum_i S_i/\sqrt{B_i}$
Strategy 1	0.0022	0.0072	0.0021	0.0067	0.0054	0.0157
Strategy 2	0.0025	0.0056	0.0017	0.0038	0.0061	0.0137
Strategy 3	0.0020	0.0073	0.0018	0.0068	0.0049	0.0164
Strategy 4	0.0036	0.0085	0.0025	0.0060	0.0139	0.0244

Table 4.15: Values obtained for the maximum binned significance and the sum of binned significances for the distributions defined by the four VBF signal extraction strategies. Results are obtained using 2018 data and the three $\tau\tau$ decay channels separately.

4.9 Summary

The analysis strategy followed in the HH $\to$ bb$\tau\tau$ search has been presented in the context of the non-resonant double Higgs production. This search targets the ggF and VBF production modes and the fully-hadronic ($\tau_h\tau_h$) and semi-leptonic ($\tau_e\tau_h$, $\tau_\mu\tau_h$) τ pair decay modes.

The analysis flow followed in this search has been described, including the identification of the processes that contribute as reducible or irreducible background sources and the signal reweighing procedures used to study different BSM scenarios. The different selections applied in the analysis have been described, profiting from the newly developed techniques at object identification and reconstruction level, in order to reduce the background contamination. A further discrimination is obtained by splitting the events into ggF- and VBF-enriched categories. For the VBF categorization, I developed a new machine learning approach, consisting on a multi-class DNN that classifies events as signal or as one of the main backgrounds. This approach achieves a very high signal accuracy (i.e. most signal events are classified in the signal subcategories) and low background contamination in the signal subcategories, which improves the analysis sensitivity. In the final analysis categories, a different DNN has been developed to identify HH $\to$ bb$\tau\tau$ events. The DNN output distributions in the analysis categories are used for signal extraction, obtaining the results shown in Chapter 5.

Chapter 5

Results

This chapter presents the results obtained in the CMS HH $\rightarrow$ bb$\tau\tau$ analysis using the full Run 2 data set. Two sets of results are produced: inclusive results, considering as signal both ggF and VBF altogether, and VBF-only results, where only the VBF production mode is considered as signal and ggF is fixed to its SM prediction.

The chapter is structured as follows. Section 5.1 shows an overview of the statistical methodology followed for performing the signal extraction. Section 5.3 describes the systematic uncertainties considered in the analysis. The final results are shown in Section 5.4 and a comparison with the results obtained from other HH analyses is performed in Section 5.5.

My main contributions to this chapter are the detailed study of the QCD background estimation, including several tests that try to determine the validity of the estimation method. These validity tests lead to the inclusion of several systematic uncertainties related to the QCD estimation. Additionally, I also studied the uncertainty related to the VBF dipole recoil.

5.1 Statistical procedure

The statistical methodology used to evaluate the presence or not of signals in data follows the method used in the 2011 combination of the CMS and ATLAS Higgs boson observation results [97]. This procedure, based on the modified frequentist method (often referred to as CL_s), aims to quantify the compatibility between the observed data and the background plus signal hypothesis or between the data and the background only hypothesis.

The procedure considers the DNN distributions in the categories (Resolved, 1 b-tag, Resolved 2 b-tag, Boosted and 5 VBF subcategories), three channels and three years (from here onwards, event categories) described during the analysis, for both the data, background estimate and signal processes. For each event category, the number of events n_i predicted by the modelling in the bin i of the distribution is defined as

$$n_i(\mu, \boldsymbol{\theta_i}) = \mu \cdot s_i(\boldsymbol{\theta_i}) + b_i(\boldsymbol{\theta_i}), \tag{5.1}$$

where $s_i(\boldsymbol{\theta_i})$ and $b_i(\boldsymbol{\theta_i})$ are the modelled signal and background yields in the given bin, μ is the *signal strength modifier*, defined as the ratio between the observed signal cross section and the one predicted by the SM, and $\boldsymbol{\theta_i}$ are the *nuisance parameters*, which parametrize the different uncertainties. These nuisance parameters are described in Section 5.1.2.

5.1.1 Observed and expected limits

In order to estimate the values of the μ and their associated uncertainties, a binned *likelihood function* is used to quantify the compatibility between the observed data and the prediction for given values of the statistical parameters (μ and nuisance parameters). This function can be written as the product of the Poisson probabilities to observe n_i^{obs} events in the bin i, scaled by the prior nuisance probability distribution (defined in Section 5.1.2):

$$L(\mu, \boldsymbol{\theta}|\text{data}) = \prod_i \frac{(\mu \cdot s_i(\boldsymbol{\theta_i}) + b_i(\boldsymbol{\theta_i}))^{n_i^{\text{obs}}}}{n_i^{\text{obs}}!} e^{(\mu \cdot s_i(\boldsymbol{\theta_i}) + b_i(\boldsymbol{\theta_i}))} \cdot \rho(\boldsymbol{\theta_i}|\tilde{\boldsymbol{\theta}}_{\boldsymbol{i}}) \tag{5.2}$$

To compare the compatibility of the data with with different signal plus background hypotheses against the background-only hypothesis a *test statistic* q_μ can be defined as:

$$q_\mu = -2 \ln \frac{L(\mu, \hat{\boldsymbol{\theta}}_\mu|\text{data})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}}|\text{data})}, \tag{5.3}$$

where $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are the parameters that maximise L, and $\hat{\boldsymbol{\theta}}_\mu$ the set of nuisances that corresponds to the maximum value of L for a given μ value. Higher values of q_μ correspond to increasing incompatibility of the data with the signal plus background hypothesis.

The probability density functions $f(q_\mu|\mu, \hat{\boldsymbol{\theta}}_\mu^{\text{obs}})$ (signal plus background hypothesis) and $f(q_0|0, \hat{\boldsymbol{\theta}}_0^{\text{obs}})$ (background-only) are obtained by generating pseudo-data, i.e. simulations that follow the same Poisson probability distribution. However, this approach is very CPU consuming, so instead an asymptotic approximation is considered [98]. Instead of using the generated pseudo-data, only the Asimov dataset (where the observations are equal to the predictions and the nuisance parameters are equal to their nominal values) is considered. Thus, the probabilities that the observed value $q_\mu^{n^{\text{obs}}}$ is compatible with the signal plus background ($\equiv \text{CL}_{s+b}(\mu)$) or with the background-only ($\equiv \text{CL}_b$) hypothesis are computed as

$$\begin{aligned} \text{CL}_{s+b}(\mu) = P(q_\mu \geq q_\mu^{n^{\text{obs}}}|\mu \cdot s + b) = \int_{q_\mu^{n^{\text{obs}}}}^{\infty} f(q_\mu|\mu, \hat{\boldsymbol{\theta}}_\mu^{\text{obs}})\ \text{d}q_\mu, \\ \text{CL}_b = P(q_0 \geq q_0^{n^{\text{obs}}}|b) = \int_{q_0^{n^{\text{obs}}}}^{\infty} f(q_0|\mu, \hat{\boldsymbol{\theta}}_0^{\text{obs}})\ \text{d}q_0. \end{aligned} \tag{5.4}$$

Their ratio is known as the $\text{CL}_s(\mu)$,

$$\text{CL}_s(\mu) = \text{CL}_{s+b}(\mu)/\text{CL}_b. \tag{5.5}$$

In the HH $\to$ bb$\tau\tau$ analysis, the confidence level (CL) is taken as the value 1 - CL_s. The 95% CL value ($\mathrm{CL}_s(\mu) = 0.05$) is usually chosen as a convention for presenting CMS results.

The observed upper exclusion limit for the signal strength modifier μ_{up} is obtained so that $\mathrm{CL}_s(\mu_{\mathrm{up}}) = 0.05$. This observed limit can be compared with the expected exclusion limit (i.e. obtained only using the Asimov dataset) and its $\pm 1\sigma$ or $\pm 2\sigma$ error bands (corresponding to 68% and 95% confidence intervals respectively), so the agreement between the observations and the model predictions can be studied. In the asymptotic approach used, the Nσ error bands $\mu^{\mathrm{exp}}_{\mathrm{up+N}}$ and the median $\mu^{\mathrm{exp}}_{\mathrm{up}}$ of the expected limit are obtained using the Asimov dataset and given by [98]

$$\mu^{\mathrm{exp}}_{\mathrm{up+N}} = \sigma_A \cdot \left(\Phi^{-1}(1 - \mathrm{CL}_s\Phi(N)) + \mathrm{N}\right), \tag{5.6}$$

where $\sigma^2_A = \mu^2/q_{\mu,A}$ is evaluated by using the Asimov Dataset and $\Phi(x)$ is the cumulative standard Gaussian distribution (zero mean, unit variance).

5.1.2 Nuisance parameters

Both background and signal event yields are affected by several sources of uncertainties, modelled by introducing nuisance parameters. These nuisance parameters have a probability density function $\rho(\theta|\tilde{\theta})$ associated to some estimate of the nominal value $\tilde{\theta}$ and other parameters regulating its shape. Due to the Bayes' theorem, the probability density can be re-interpreted as a posterior arising from dedicated measurements for the estimations of $\tilde{\theta}$:

$$\rho(\theta|\tilde{\theta}) \sim p(\theta|\tilde{\theta}) \cdot \pi_\theta(\theta), \tag{5.7}$$

where $p(\theta|\tilde{\theta})$ is the probability density function for the auxiliary measurement of $\tilde{\theta}$ and $\pi_\theta(\theta)$ the prior probability distribution.

In the HH $\to$ bb$\tau\tau$ analysis, three types or nuisances are considered.

- Normalization uncertainties: some uncertainties only modify the normalization of one or more processes, affecting either all event categories or a limited number of those. Up and down variations of these uncertainties, obtained by adding or subtracting to the bin content the value correspondent to one standard deviation to the central value, vary the number of events of the distribution equally in all bins. The prior probability distributions on these uncertainties involve Gaussian functions. However, for large values of the uncertainties, the normal distribution is not appropriate. Instead, the log-normal distribution is considered:

$$\pi_\theta(\theta) = \frac{1}{\theta}\frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{(\ln\theta)^2}{2\sigma^2}\right), \tag{5.8}$$

  where σ is the width of the distribution and corresponds to the relative uncertainty on θ, estimated using theoretical calculations or alternative measurements.

- Shape uncertainties: instead of only modifying the whole integral of the distributions, some uncertainties vary the content of individual bins in an event category distribution. These variations are modelled by recreating the distributions by modifying the quantity affected by the systematic according to the boundaries of its central 68% confidence interval. Using the three distributions (no variation and up and down variations), the number of events per bin is smoothly interpolated between the boundaries defined by the up and down distributions.

- Statistical uncertainties: the amount of simulated events that can be produced for a given process per bin and event category is limited. To model the statistical uncertainties, one possibility is to include one separate nuisance per process and bin following a Poisson distribution, the so-called Barlow-Beeston approach [99]. However, if enough events are available per bin, the sum of Poisson distributions can be approximated by a Gaussian distribution with reasonable accuracy. In that case, only one nuisance per bin is incorporated in the statistical model.

In addition to providing a value for the signal strength modifier, the likelihood maximization provides the best fit values for each nuisance parameter of the statistical model and an estimate of their a-posteriori widths. The nuisance values before and after the fit are referred to as *pre-fit* and *post-fit* values, respectively.

5.2 Validation tests for the QCD background estimation

In order to evaluate the validity of the followed procedure for the QCD background estimation, several tests were performed. Details and results of these tests are given here.

In the first test, the stability of the estimated QCD normalization has been evaluated by modifying the definition of the C and D regions using four different DeepTau working points. These alternative regions are defined as:

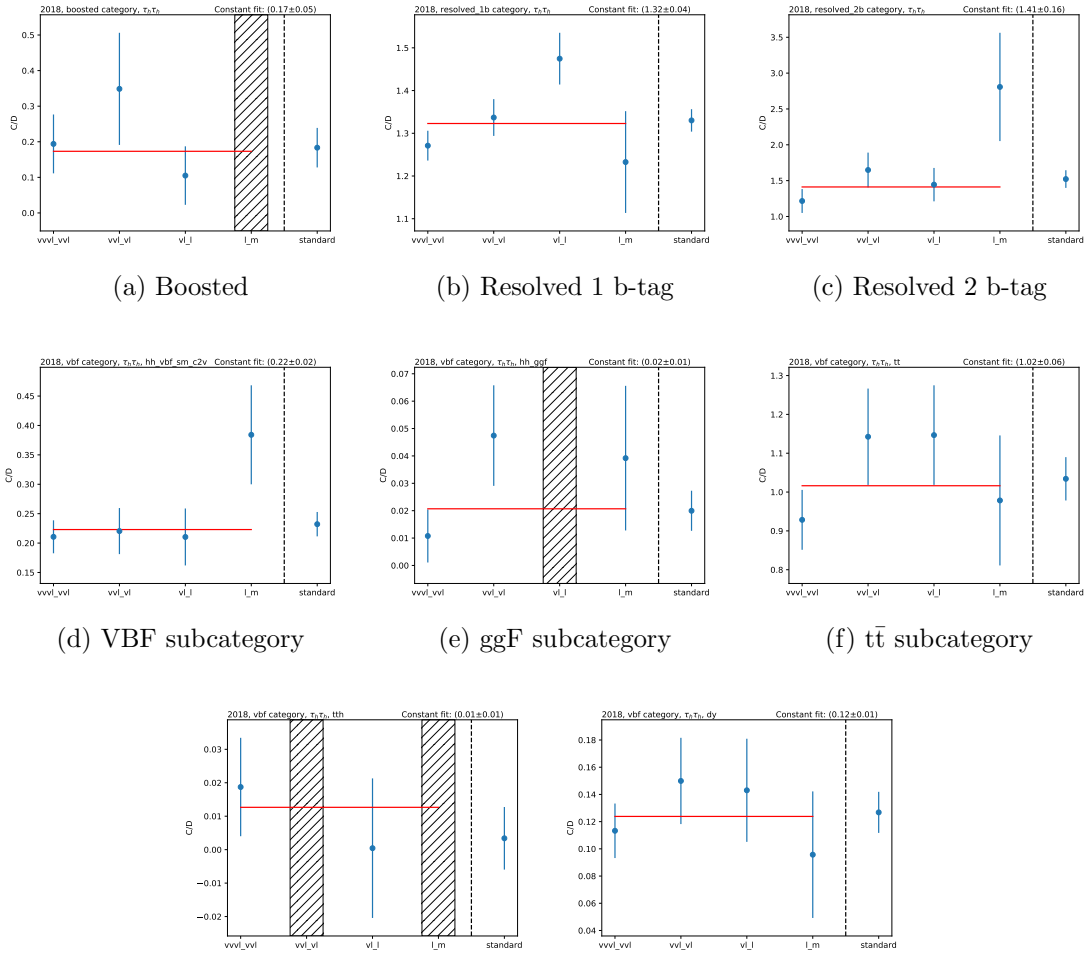
- Events where the τ_h passes the VVVLoose working point of the DeepTauVSjet discriminator, but not the VVLoose one.

- Events where the τ_h passes the VVLoose working point of the DeepTauVSjet discriminator, but not the VLoose one.

- Events where the τ_h passes the VLoose working point of the DeepTauVSjet discriminator, but not the Loose one.

- Events where the τ_h passes the Loose working point of the DeepTauVSjet discriminator, but not the Medium one.

In these four alternative regions the ratio between the C and D yields is computed and compared with the one obtained with the standard definition of the C and D regions. Results for the C/D yield estimation for the $\tau_h\tau_h$ channel (where the QCD contribution is really important) in 2018 are shown in Fig. 5.1. The other channels and years are shown in Appendix B. A straight line showing a fit to a constant of the C/D yield values for the four alternative working point definitions is also plotted. When a working point leads to a negative yield in either of regions B, C or D, a dashed band is shown and the point is not used to compute the fit nor the average. In general the C/D yield obtained with the standard definition of C and D regions is in agreement within uncertainties with both the result of this constant fit, and with the weighted average of these four alternative C/D yield points. In the few cases where this does not happen, an additional uncertainty is proposed, as defined in Section 5.3.



(a) Boosted (b) Resolved 1 b-tag (c) Resolved 2 b-tag

(d) VBF subcategory (e) ggF subcategory (f) $t\bar{t}$ subcategory

(g) $t\bar{t}H$ subcategory (h) DY subcategory

Figure 5.1: C/D yield estimations for the five working points for the $\tau_h\tau_h$ channel in 2018. A dashed band is plotted where the correspondent working point leads to a negative yield for regions B, C and/or D.

In the second test, the QCD estimation obtained from the ABCD method is compared with a direct data after background subtraction in a sideband region where signal presence is negligible. This sideband region has been defined by inverting the elliptic mass cut from the Resolved 1 b-tag category (defined in Section 4.3.3). Fig. 5.2 shows the DNN output distribution per channel and year in the sideband region considered. Again, the QCD estimation obtained with the ABCD method has been found to be in good agreement with the direct data-background subtraction in this sideband, validating the QCD estimations obtained with the ABCD method used in the analysis.

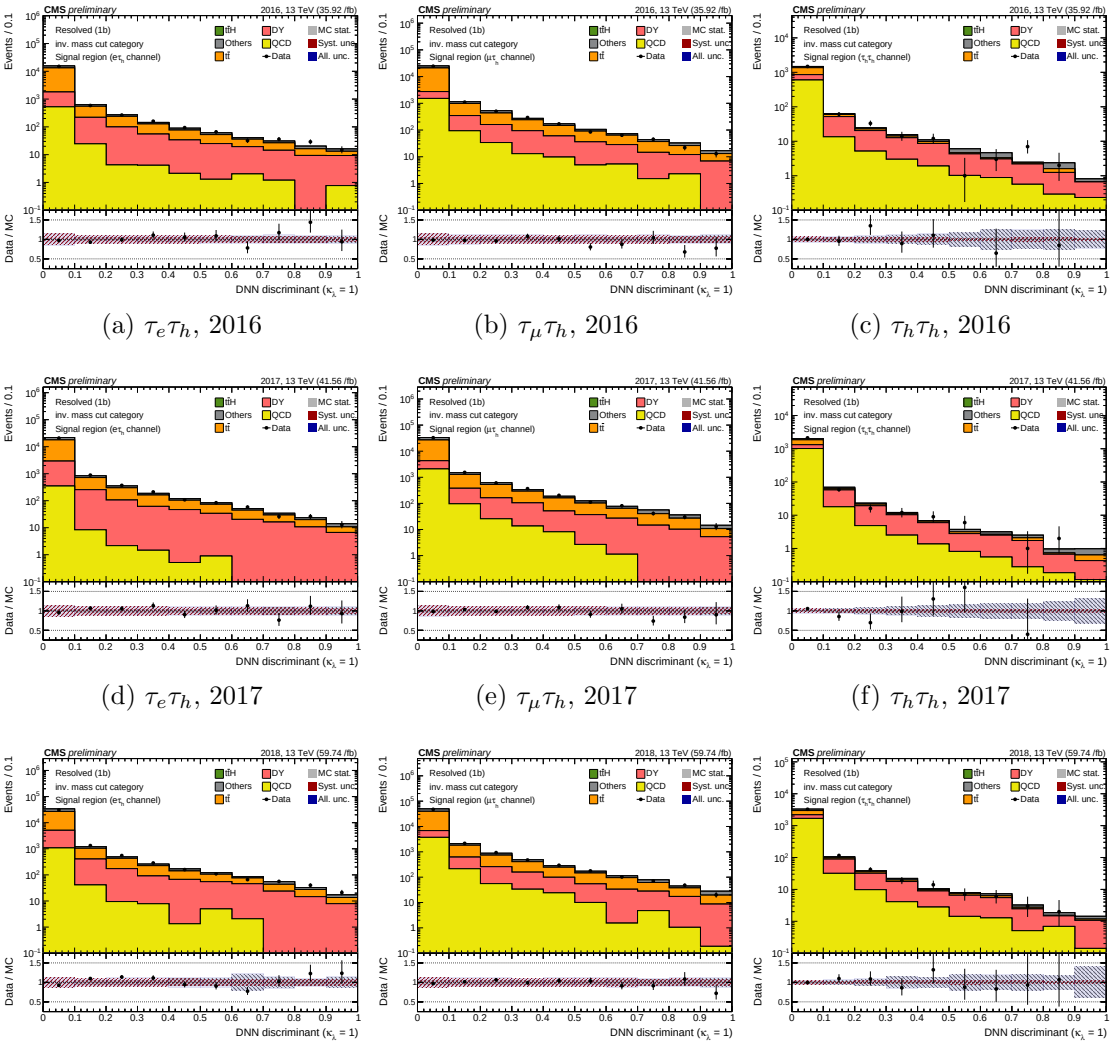
Finally, the last validity test consists of obtaining the QCD contribution using the ABCD method in the final analysis categories but redefining the usual regions using an alternative DeepTauVSjet working point that was already introduced in the first test. This way, the new regions A', B', C' and D' use the following selections:

- A': Both τ have opposite charge and the τ_h passes the VVLoose working point but not the VLoose one.
- B': Both τ have the same charge and the τ_h passes the VVLoose working point but not the VLoose one.
- C': Both τ have opposite charge and the τ_h passes the VVVLoose working point but not the VVLoose one.
- D': Both τ have the same charge and the τ_h passes the VVVLoose working point but not the VVLoose one.

This way, we can estimate the QCD contribution in regions close to the analysis signal regions but poorly populated in signal (thanks to considering a very loose working point) and enhanced in QCD contribution. Fig. 5.3 shows the DNN output distributions for the $\tau_h\tau_h$ channel in 2018 in this alternative signal region. Agreement within the statistical uncertainties can be seen between the prediction from all background sources, including the QCD estimation obtained with the ABCD method, and the data in this validation region, reassuring the validity of the ABCD method for QCD estimation. In particular, we see a very large QCD contribution in the $\tau_h\tau_h$ channel in the Resolved 1 b-tag category, where also a very good agreement within uncertainties between data and backgrounds can be observed.

5.3 Systematic uncertainties

The systematic uncertainties considered in the HH $\rightarrow$ bb$\tau\tau$ analysis are discussed in the following sections. Two types of uncertainties are considered: *experimental uncertainties*, coming from the imperfect knowledge of the collected data and the differences in object reconstruction between real data and simulation, and *theoretical uncertainties*, arising from the imperfect modelling of the MC simulations or the limited knowledge of the fundamental parameters.

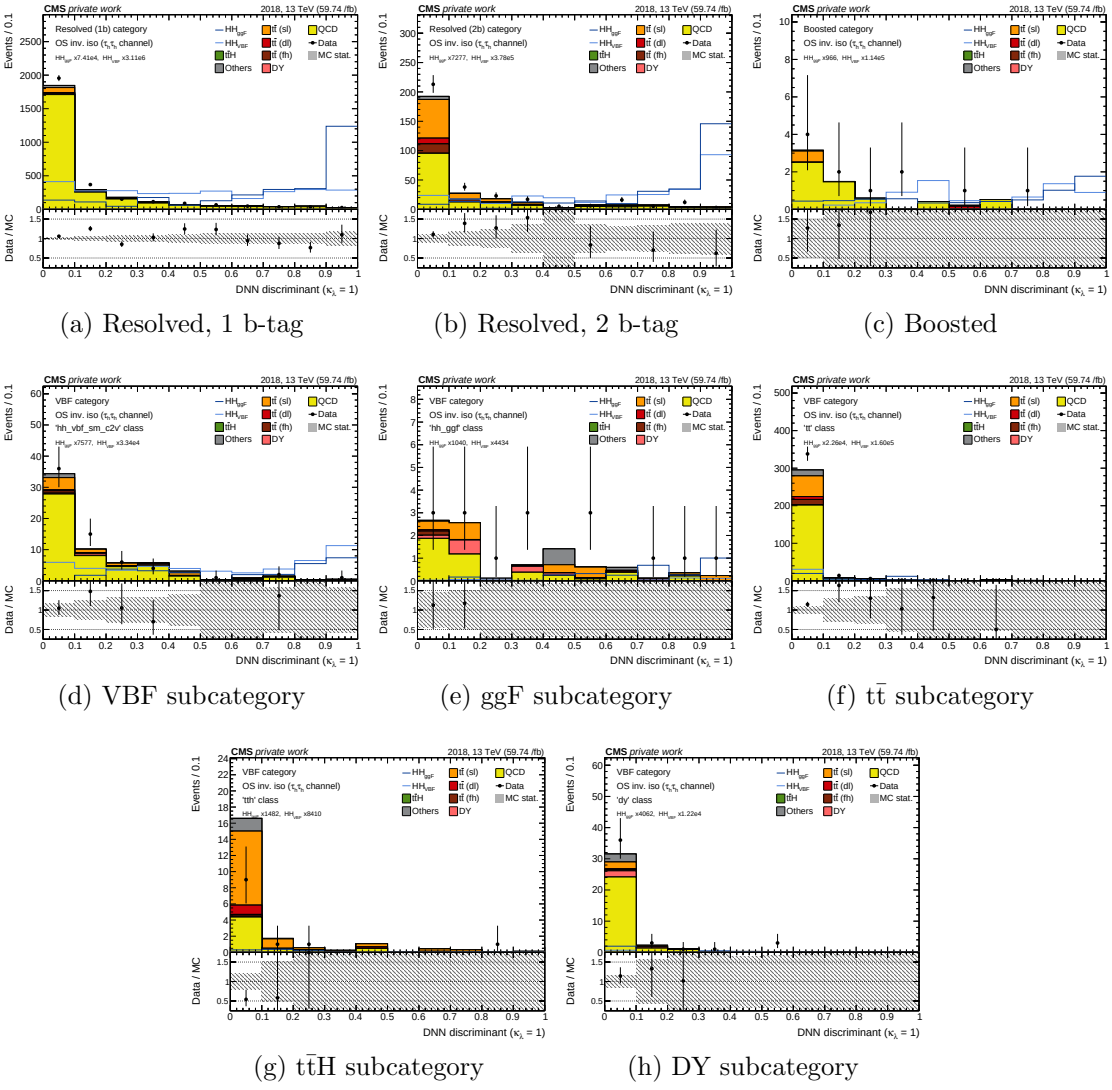


(a) $\tau_e\tau_h$, 2016 (b) $\tau_\mu\tau_h$, 2016 (c) $\tau_h\tau_h$, 2016

(d) $\tau_e\tau_h$, 2017 (e) $\tau_\mu\tau_h$, 2017 (f) $\tau_h\tau_h$, 2017

(g) $\tau_e\tau_h$, 2018 (h) $\tau_\mu\tau_h$, 2018 (i) $\tau_h\tau_h$, 2018

Figure 5.2: DNN output distribution per channel for the sideband region defined by inverting the elliptic mass cut in the Resolved 1 b-tag category. Statistical uncertainties for the data and all background contributions and the normalization systematic uncertainties for the main background processes are included.



(a) Resolved, 1 b-tag (b) Resolved, 2 b-tag (c) Boosted

(d) VBF subcategory (e) ggF subcategory (f) $t\bar{t}$ subcategory

(g) $t\bar{t}H$ subcategory (h) DY subcategory

Figure 5.3: DNN output distribution for the different analysis categories in the alternative signal region defined by a looser DeepTauVSjet working point for the $\tau_h\tau_h$ channel in 2018. Statistical uncertainties for the data and all background contributions are included.

Both normalization and shape uncertainties are considered in the analysis. Unless the opposite is stated, all uncertainties are uncorrelated among the three years, the analysis categories, and the $\tau\tau$ decay channels.

Experimental uncertainties

Luminosity

The uncertainty in the measured integrated luminosity is obtained via dedicated Van-der-Meer scans and the stability of the detector response during data-taking [100]. It takes values of 1.2%, 2.3% and 2.5% for the 2016 [101], 2017 [102] and

2018 [103] datasets respectively. These uncertainties are partially correlated among the three years. These uncertainties are applied to all signals and to all backgrounds that are estimated fully from MC simulation. Given that the normalizations of the $t\bar{t}$, Z/γ^*+ jets and QCD multi-jet processes are obtained from data, the uncertainty is not applied for these processes.

Trigger scale factors

Four uncertainties in the trigger scale factors are included to take into account the four different τ_h decay modes considered in the analysis. They are applied to the τ_h present in the final state (both in the $\tau_h\tau_h$ channel). The largest effect in the DNN discriminant is given for the decay mode 1 in the $\tau_h\tau_h$ channel, as shown in Fig. 5.4a. Two additional uncertainties are used for the τ decays into an electron or a muon, whose effects are shown in Fig. 5.4bc. In addition, in 2017 and 2018 in the $\tau_h\tau_h$ channel one further uncertainty source is considered arising from the jet scale factors of the VBF trigger. Its effect on the $\tau_h\tau_h$ channel and the VBF subcategory is shown in Fig. 5.4d.

Lepton reconstruction, isolation and identification

Electron and muon reconstruction, isolation, and identification uncertainties are determined from the simulation-to-data scale factors; a value of 1 % is obtained for both electrons and muons. For τ_h, the corresponding systematic uncertainty is 3 (15)% for $p_T < 100$ GeV ($p_T > 100$ GeV).

L1 ECAL trigger prefiring

The uncertainty on the correction factor used to account for the L1 prefiring is obtained from the combination of the statistical uncertainty associated to the p_T and η bin considered for each photon and jet, and the systematic uncertainty, taken as the 20% of the prefiring rate. This uncertainty only has a rate-changing effect and takes a value of 2%.

PU reweighting

The uncertainty on the PU reweighting technique comes from the uncertainty on the inelastic pp total cross section, which takes a value of 4.6%. This value is then propagated to the PU weight generating its up and down variations. The resulting systematic uncertainty is estimated to have a value of 1%.

PU jet identification

Uncertainties in he PU jet identification scale factors are included as functions of the jet p_T and η. These uncertainties modify the weight of each event, not the

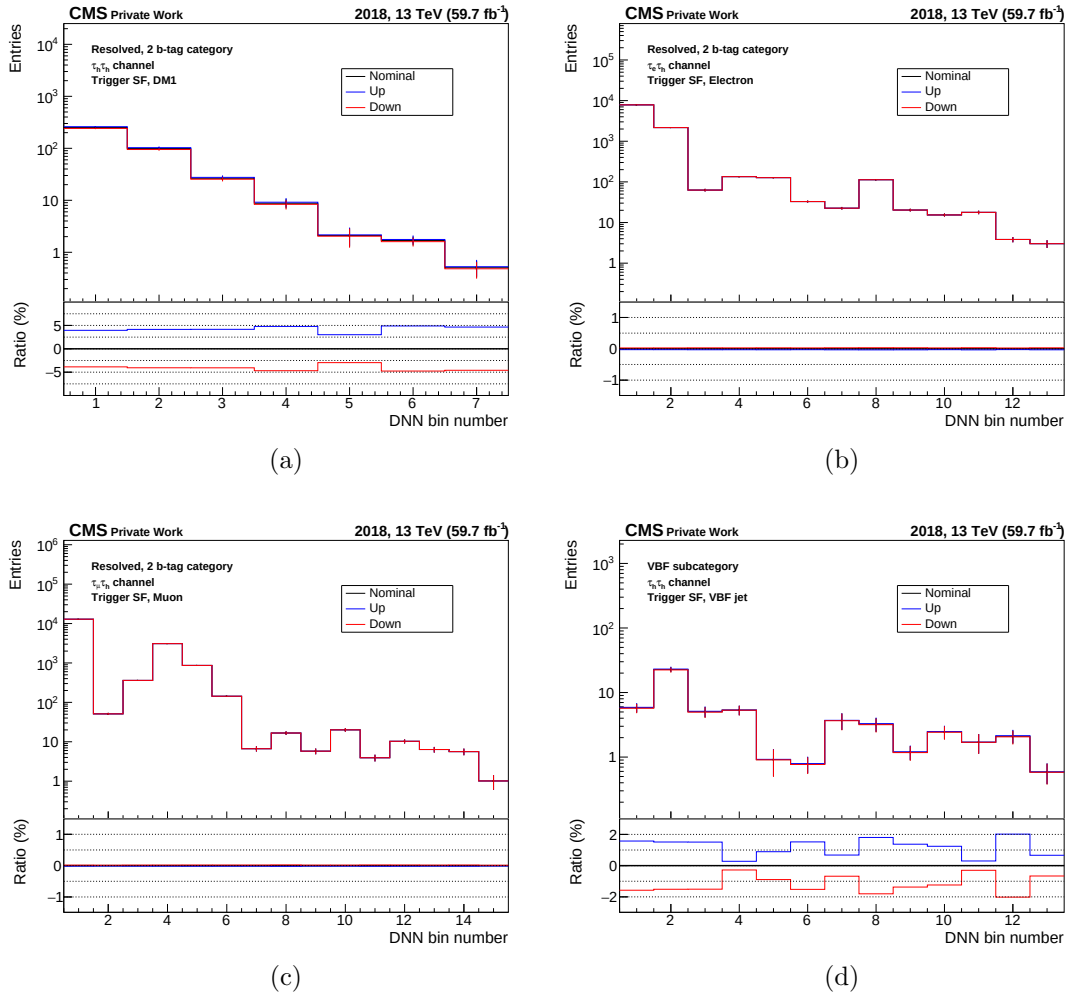


Figure 5.4: (a) Trigger scale factor uncertainty variations on the DNN discriminant in the Resolved, 2 b-tag category and the $\tau_h\tau_h$ channel in 2018 for the decay mode 1. (b) Trigger scale factor uncertainty variations on the DNN discriminant in the Resolved, 2 b-tag category and the $\tau_e\tau_h$ channel in 2018 for the τ decay into an electron. (c) Trigger scale factor uncertainty variations on the DNN discriminant in the Resolved, 2 b-tag category and the $\tau_\mu\tau_h$ channel in 2018 for the τ decay into a muon. (d) Trigger scale factor uncertainty variations on the DNN discriminant in the VBF subcategory and the $\tau_h\tau_h$ channel in 2018 for the trigger jets in the VBF trigger. Contributions of all expected backgrounds are added.

individual objects. The resulting up and down templates for the DNN discriminant in the Resolved, 2 b-tag category and the $\tau_h\tau_h$ channel are shown in Fig. 5.5.

$\mathrm{t\bar{t}}$ normalization

The normalization of the $\mathrm{t\bar{t}}$ background is taken from a fit to a $\mathrm{t\bar{t}}$ enriched control region for each year, as shown in Section 4.4.3. The resulting systematic uncertainty

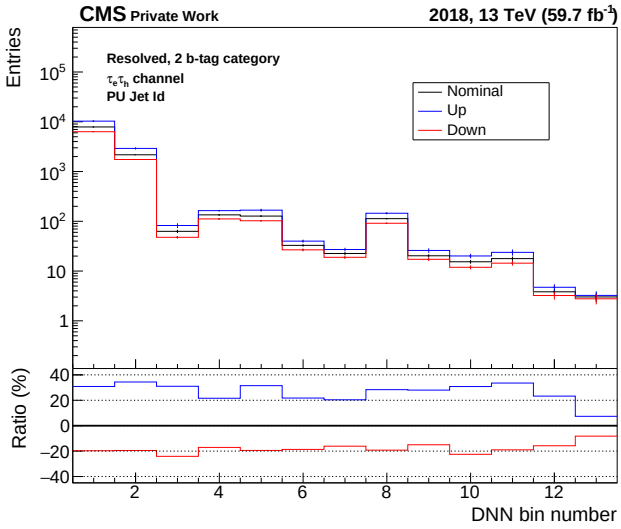


Figure 5.5: PU jet ID uncertainty variations on the DNN discriminant in the Resolved, 2 b-tag category and the $\tau_h\tau_h$ channel in 2018. Contributions of all expected backgrounds are added.

is below 1% for each year.

Z/γ^*+ jets normalization

The normalization of the Z/γ^*+ jets background is taken from a fit to 18 Z/γ^*+ jets enriched control regions per year, as described in Section 4.4.2. The resulting uncertainties range from 0.1 to 60% depending on the year and the control region considered.

QCD multi-jet

The contribution from QCD multi-jet background is determined using the ABCD method as described in Section 4.4.1. Assuming that the contribution of the QCD multi-jet background is constant in the four regions considered (A, B, C and D), three uncertainties can be defined, all uncorrelated across categories, $\tau\tau$ decay channels and years:

- **Shape uncertainty**: as shown in Section 4.4.1, the QCD multi-jet background shape is extracted from the selection included in region C. However, an additional estimation can be obtained if, instead of taking the shape from region C and computing the factor k as the ratio of the yields coming from regions B and D, the shape was taken from region B and the factor k is computed as the ratio between regions C and D. In fact, as shown in Figs. 5.6 and 5.7, the data/background agreement is very comparable when considering these two strategies. Therefore, the QCD background shape used in the

analysis will come from the average of the shapes from the two regions. The two alternative shapes (from C and B) are considered as the up and down templates for the QCD shape uncertainty.

- **Uncertainty on the yield correction factor**: the statistical uncertainty on the correction factor B/D is a normalization uncertainty defined as the sum in quadrature of the statistical uncertainties on the event yields in regions B and D. The value of this uncertainty ranges from 5% to 100% depending on the category, channel and year considered.
- **Additional uncertainty**: this normalization uncertainty is added only in the cases where the C/D factor obtained using the standard QCD definition of the C and D regions does not agree within uncertainties with the weighted average of the four alternative C/D yield points defined in the first validation test (see Section 5.2) for the QCD multi-jet background estimation. It follows the expression

$$\text{add\_unc} = \sqrt{\left(\frac{(\text{C/D})_{\text{standard}} - (\text{C/D})_{\text{average}}}{(\text{C/D})_{\text{standard}}}\right)^2 - \left(\frac{\Delta(\text{C/D})_{\text{standard}}}{(\text{C/D})_{\text{standard}}}\right)^2}. \quad (5.9)$$

Table 5.1 shows the QCD normalization percentage relative uncertainty for all years, channels and categories considered. For each cell, the first number is the statistical uncertainty on the yield correction factor, while the second (if it appears) is the additional uncertainty as defined above. The cell is empty if no QCD could be estimated (as either regions B, C or D have negative yield). Note that the statistical uncertainty for all the VBF subcategories in the same year and channel is the same, as it comes from the VBF inclusive category (described in Section 4.3.3).

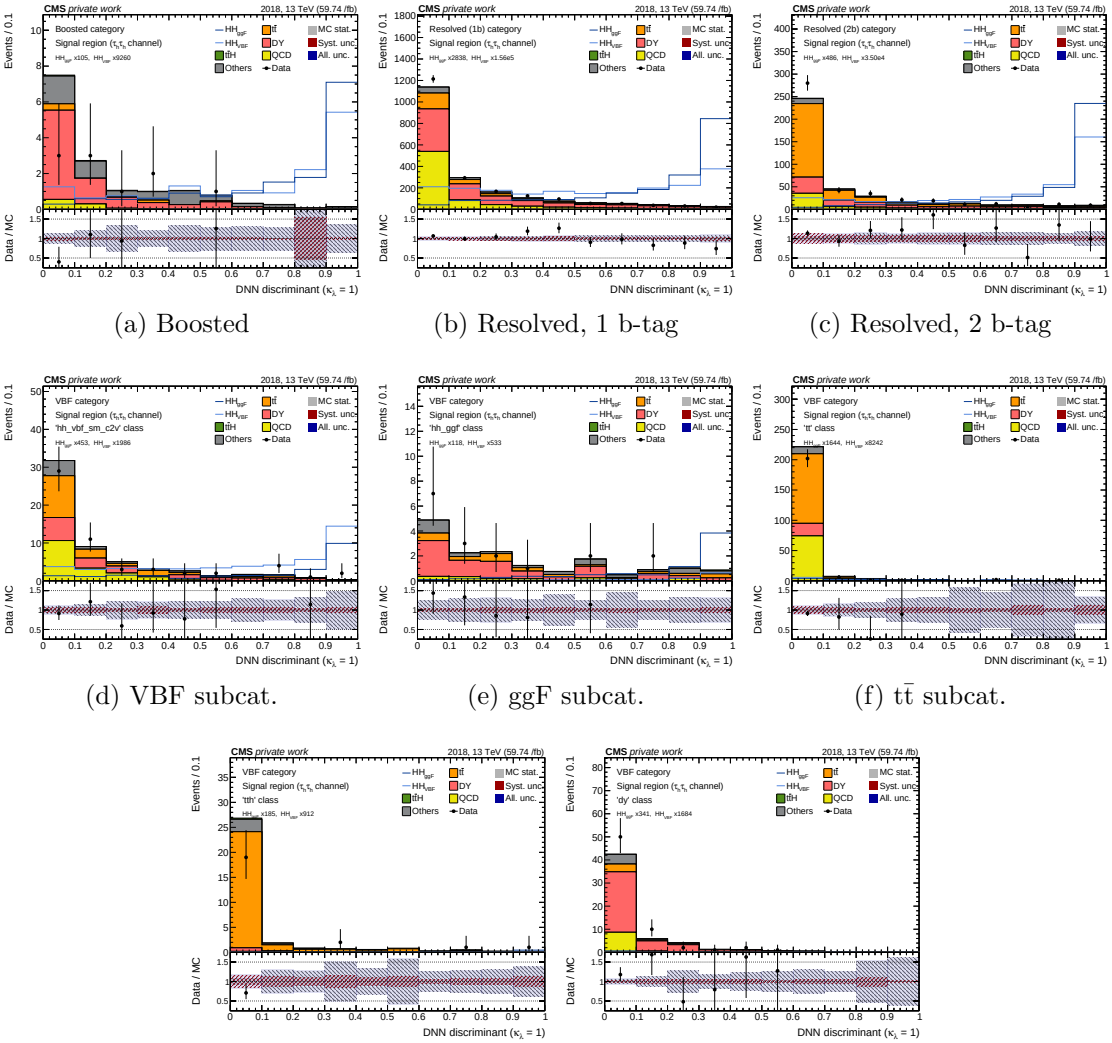
τ_h energy scale

The uncertainty on the measurement of the τ_h energy measurement can lead to a change in the distribution of the discriminant variables. An uncertainty on the energy scale of each τ_h is derived by combining low- and high-p_T measurements in $Z \rightarrow \tau\tau$ and $W^* \rightarrow \tau\nu$ events. Four different uncertainties are included, one for each τ_h decay mode considered in the analysis. Each uncertainty only applies to the τ_h reconstructed with the specific decay mode, the others remain unchanged. The largest effects are found in the $\tau_h\tau_h$ channel and the decay modes 1 and 10, as shown in Fig. 5.8ab.

Energy scale of electrons and muons misidentified as τ_h

Separate uncertainties in the energy scale of electrons misidentified as τ_h candidates are provided to take into account the $h^\pm$ and $h^\pm\pi^0$ decay modes. The largest effect from these uncertainties is seen in the $\tau_h\tau_h$ channel, shown in Fig. 5.8cd for the Resolved, 2 b-tag category.

For the muons misidentified as τ_h candidates, the uncertainty in the energy scale is 1%, uncorrelated across decay modes.



(a) Boosted (b) Resolved, 1 b-tag (c) Resolved, 2 b-tag

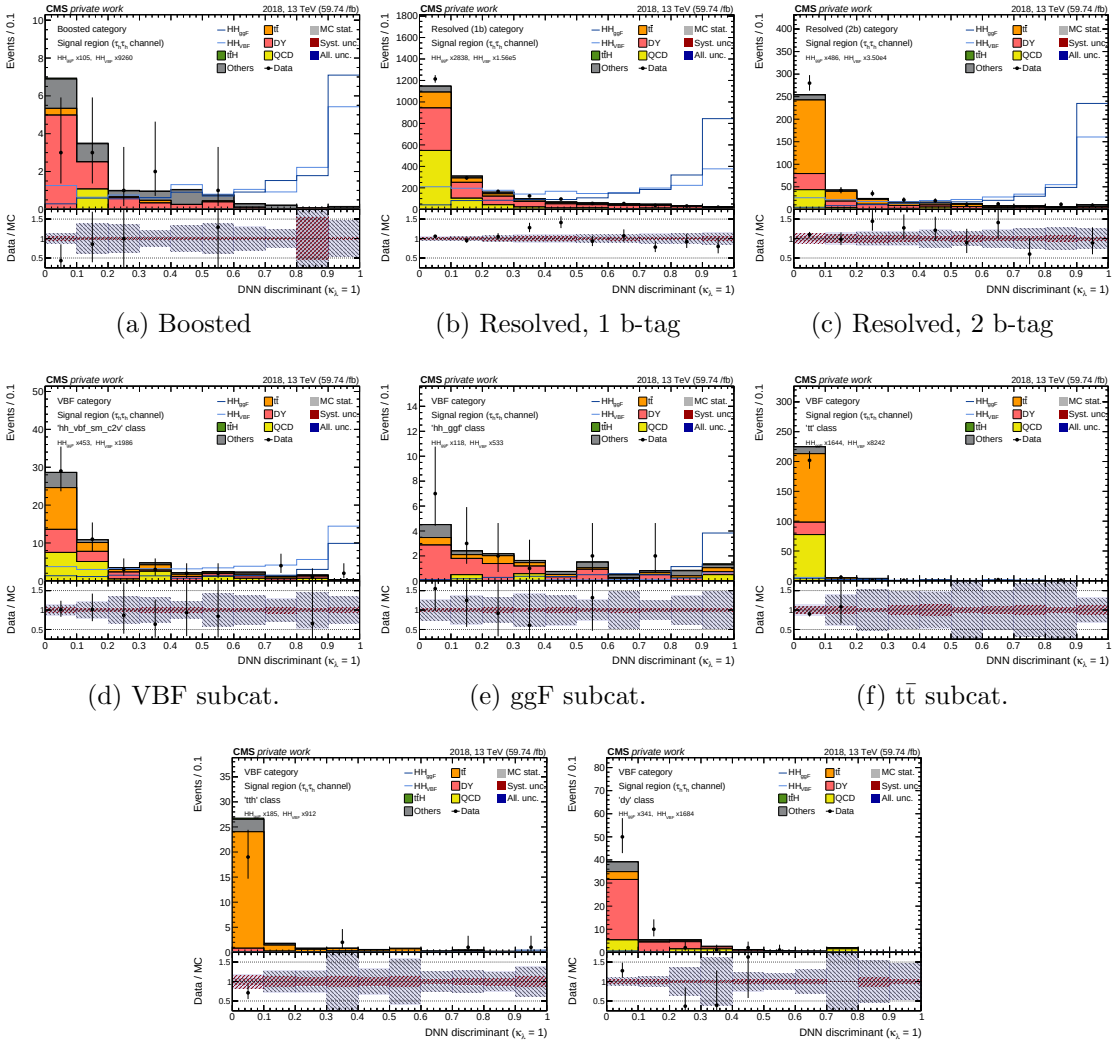
(d) VBF subcat. (e) ggF subcat. (f) $t\bar{t}$ subcat.

(g) $t\bar{t}H$ subcat. (h) DY subcat.

Figure 5.6: DNN output distributions for the analysis categories in the signal region, $\tau_h\tau_h$ channel and 2018. QCD background is estimated from the ABCD method taking the shape from region C and the k factor from B/D.

Jets faking τ_h

Uncertainties coming from the misidentification of jets as τ_h candidates are estimated with a control region in the $\tau_\mu\tau_h$ channel defined by inverting the charge requirement on the $\tau\tau$ pair and imposing that neither of the b jet candidates pass the medium DeepJet working point. The uncertainties are computed as the difference between the total yield from data and from background in this region divided by the total yield in data. Two uncorrelated uncertainties are derived per year, one for the barrel and another one for the endcap, taking values from 11% to 28% depending on the year and region.



(a) Boosted (b) Resolved, 1 b-tag (c) Resolved, 2 b-tag

(d) VBF subcat. (e) ggF subcat. (f) $t\bar{t}$ subcat.

(g) $t\bar{t}H$ subcat. (h) DY subcat.

Figure 5.7: DNN output distributions for the analysis categories in the signal region, $\tau_h\tau_h$ channel and 2018. QCD background is estimated from the ABCD method taking the shape from region B and the k factor from C/D.

Jet energy scale and resolution

Several uncertainties related to the calibration of the jet energy scale (JES) and resolution (JER) are included. For JES, 26 independent sources are identified [85]. However, for analysis mildly sensitive to jet energy corrections (as is the HH $\rightarrow$ bb$\tau\tau$ analysis), these uncertainties are grouped into 11 separate sources per year. The ones common between years are treated as fully correlated, while the ones appearing only in one year are left as uncorrelated. The effects of two of the sources with the largest impact are shown in Fig. 5.9ab. For JER, the energy of the simulated jets is smeared in order to match the observed energy resolution in data, as shown in Section 4.5. An uncertainty on the JER correction factor is provided in different

		2016	2017	2018
Boosted	$\tau_h\tau_h$	198.0	44.5	60.6
	$\tau_\mu\tau_h$	–	–	–
	$\tau_e\tau_h$	–	–	–
Resolved, 1 b-tag	$\tau_h\tau_h$	7.6	5.9	4.7
	$\tau_\mu\tau_h$	7.1	7.4	13.6
	$\tau_e\tau_h$	14.8	10.0	10.0
Resolved, 2 b-tag	$\tau_h\tau_h$	118.5	21.9	21.6
	$\tau_\mu\tau_h$	18.6	17.7	10.8
	$\tau_e\tau_h$	42.8	–	–
VBF subcategory	$\tau_h\tau_h$	27.2	23.2	14.1
	$\tau_\mu\tau_h$	81.7	26.7 , 36.2	11.6
	$\tau_e\tau_h$	122.0	33.1	–
ggF subcategory	$\tau_h\tau_h$	27.2	23.2	14.1
	$\tau_\mu\tau_h$	–	–	11.6
	$\tau_e\tau_h$	–	–	–
$t\bar{t}$ subcategory	$\tau_h\tau_h$	27.2	23.2	14.1
	$\tau_\mu\tau_h$	–	–	11.6
	$\tau_e\tau_h$	–	–	25.7
$t\bar{t}H$ subcategory	$\tau_h\tau_h$	–	23.2	14.1
	$\tau_\mu\tau_h$	81.7	–	–
	$\tau_e\tau_h$	–	–	–
DY subcategory	$\tau_h\tau_h$	27.2	23.2	14.1
	$\tau_\mu\tau_h$	81.7 , 81.2	26.7	11.6 , 7.7
	$\tau_e\tau_h$	–	–	25.7

Table 5.1: QCD normalization percentage relative uncertainty. For each cell, the first number is the statistical uncertainty on the yield correction factor, while the second (if it appears) is the additional uncertainty. The cell is empty if no QCD could be estimated (as either regions B, C or D have negative yield). Note that the statistical uncertainty for all the VBF subcategories in the same year and channel is the same, as it comes from the VBF inclusive category.

regions of η and p_T for each year, resulting in a shift in the final distributions. These shifted distributions are shown in Fig. 5.9cd for the $\tau_e\tau_h$ and $\tau_h\tau_h$ channels in the Resolved, 2 b-tag category in 2018.

DeepTau identification

Several uncertainties coming from the application of the different DeepTau identification scale factors is determined by using a tag-and-probe procedure as a function of the τ_h candidate p_T. For the identification of τ_h against jets, five uncertainties are provided, binned between 20 and 25 GeV, 25 and 30 GeV, 30 and 35 GeV, 35 and 40 GeV, and from 40 GeV onwards. The last bin produces the largest variations, shown in Fig. 5.10ab for the $\tau_\mu\tau_h$ and $\tau_h\tau_h$ channels. For the identification

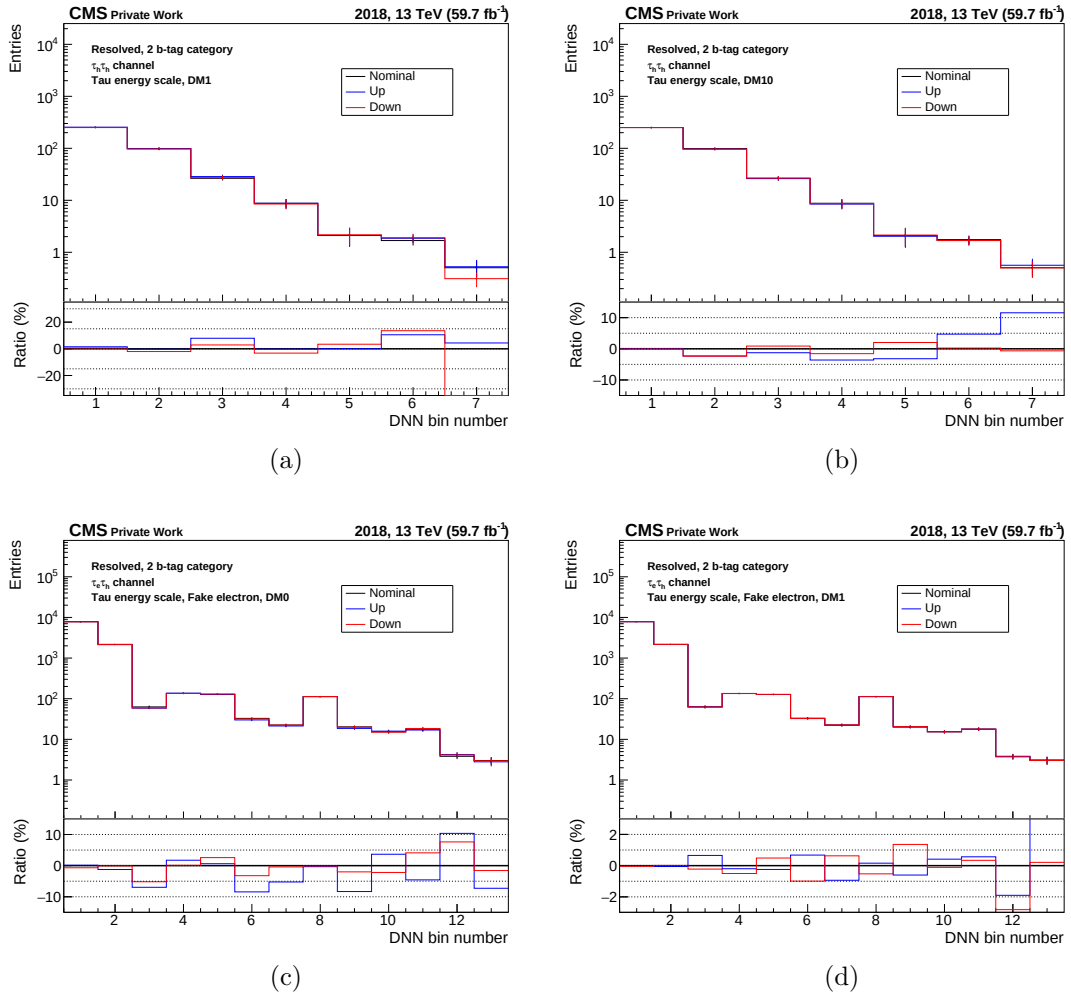


Figure 5.8: (Top) τ_h energy scale uncertainty variations on the DNN discriminant in the Resolved, 2 b-tag category and the $\tau_h\tau_h$ channel in 2018 for the decay modes 1 (a) and 10 (b). (Bottom) Electron faking τ_h energy scale uncertainty variations in the Resolved, 2 b-tag category and the $\tau_e\tau_h$ channel in 2018 for the decay modes 0 (c) and 1 (d). Contributions of all expected backgrounds are added.

of tau leptons against electrons, two additional uncertainties are included, one for the barrel and another one for the endcap, treated as uncorrelated between them. The shifted distributions produced by these systematics are shown for the $\tau_e\tau_h$ in Fig. 5.10cd.

b-tagging efficiency

The uncertainties included in the b-tagging efficiency cover both the contamination from udscg (cb) jets in heavy- (light-) flavor regions and the statistical fluctuations in the data and simulation samples used for the computation of the efficiency. The uncertainties coming from statistical fluctuations are uncorrelated between years,

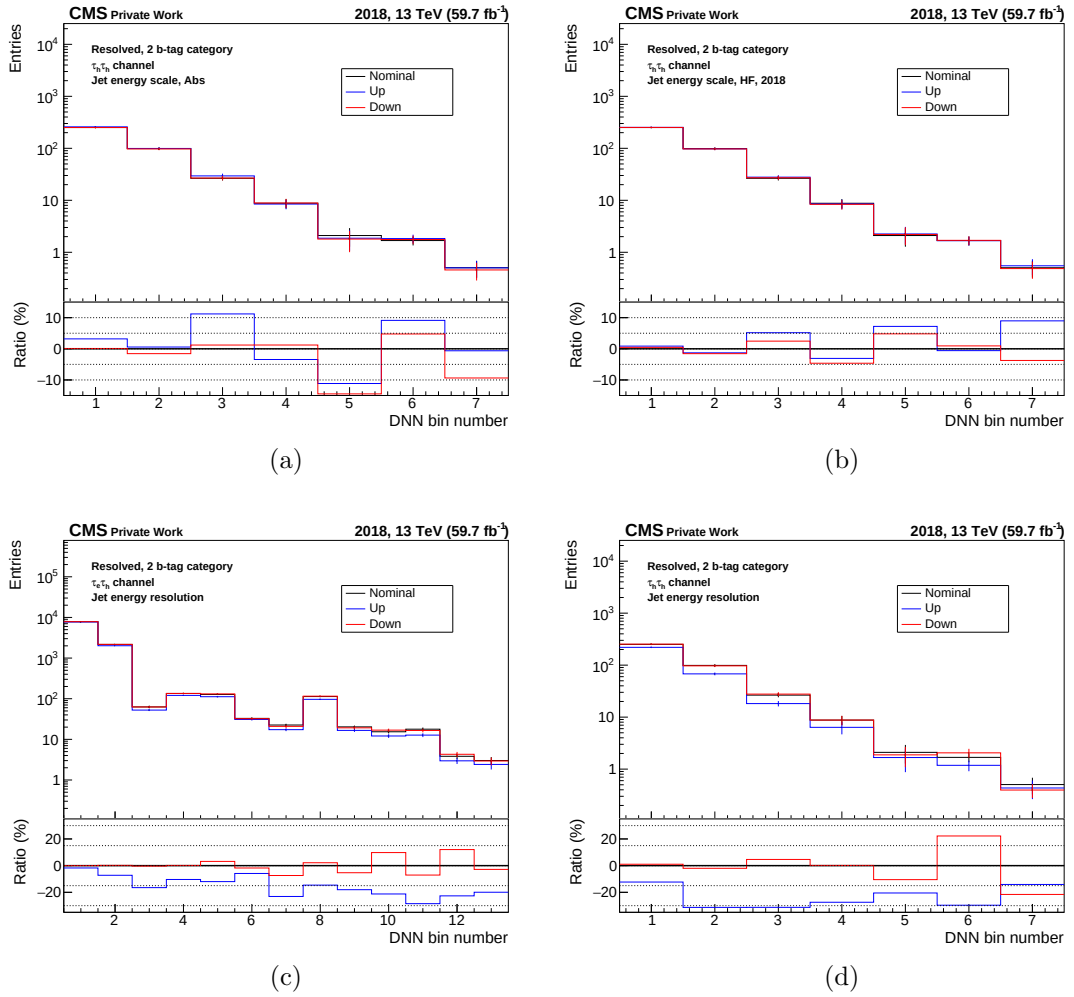


Figure 5.9: (Top) Jet energy scale uncertainty variations on the DNN discriminant in the Resolved, 2 b-tag category and the $\tau_\mu\tau_h$ (a) and $\tau_h\tau_h$ (b) channels in 2018. (Bottom) Jet energy resolution uncertainty variations in the Resolved, 2 b-tag category and the $\tau_e\tau_h$ (c) and $\tau_h\tau_h$ (d) channels in 2018. Contributions of all expected backgrounds are added.

while the ones arising from the jet contamination in light-flavor or heavy-flavor regions are correlated between years. The latter produce the largest variations on the DNN discriminator shape and rate; their corresponding shifted distributions on the $\tau_h\tau_h$ channel are shown in Fig. 5.11.

Theoretical and modelling uncertainties

HH production cross section

The theoretical uncertainty in the cross section HH production via ggF is $^{+6\%}_{-23\%}$ (scale + m_t) $\pm 3\%$ (PDF + α_s) fb [104] and via VBF $^{+0.03\%}_{-0.04\%}$ (scale) $\pm 2.1\%$ (PDF +

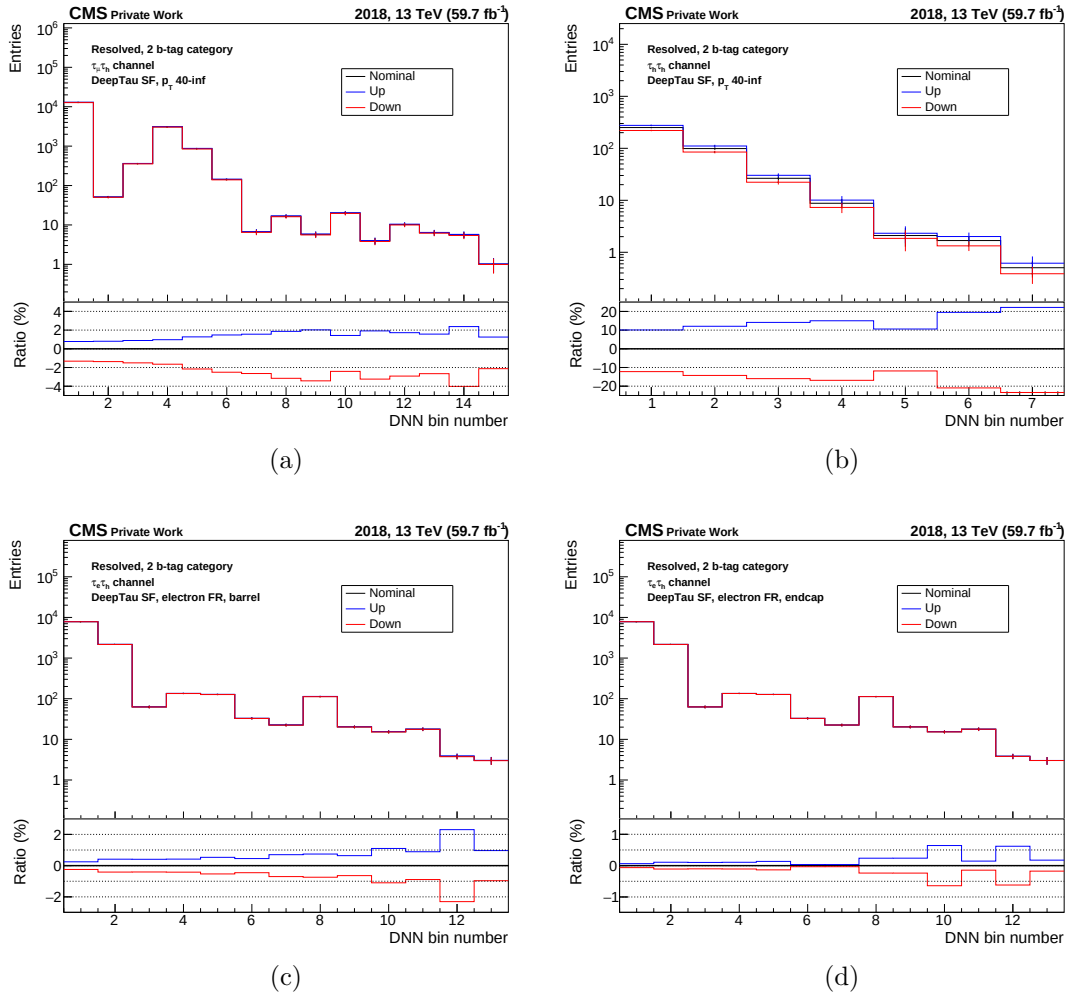
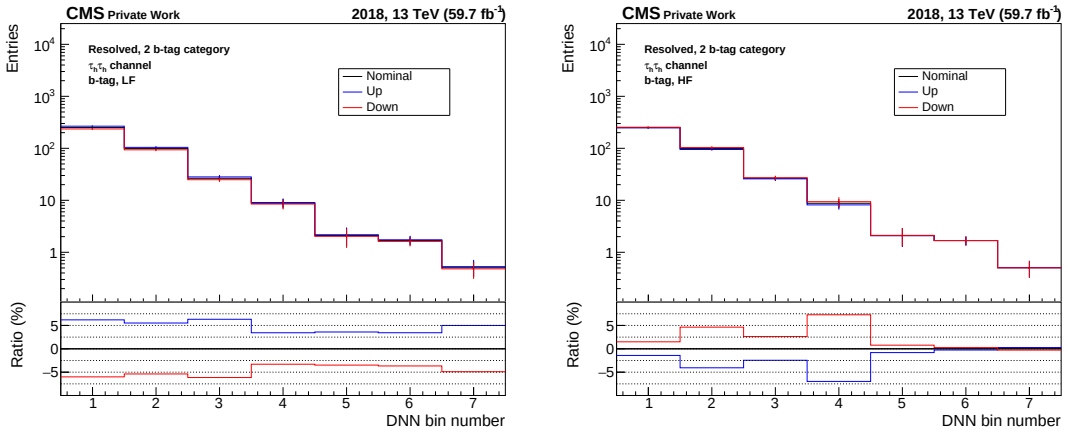


Figure 5.10: (Top) DeepTau identification scale factor ($p_T(\tau_h) > 40$ GeV) uncertainty variations on the DNN discriminant in the Resolved, 2 b-tag category and the $\tau_\mu\tau_h$ (a) and $\tau_h\tau_h$ (b) channels in 2018. (Bottom) Electron faking τ_h scale factor uncertainty variations in the Resolved, 2 b-tag category and the $\tau_e\tau_h$ channel in the barrel (c) and endcap (d). Contributions of all expected backgrounds are added.

α_s) [105]. These uncertainties are only considered when upper limits on the signal strength modifier are obtained, not in the upper limits on the cross section.

H branching fractions

Two uncertainties are included to account for the Higgs boson decay [106] into bb and $\tau\tau$: $\pm 0.65\%$ (theory) $^{+0.72\%}_{-0.74\%}$ (m_q) $^{+0.78\%}_{-0.80\%}$ (α_s) and $^{+1.16\%}_{-1.17\%}$ (theory) $^{-0.98\%}_{-0.99\%}$ (m_q) $\pm 0.62\%$ (α_s) respectively, where m_q stands for the quark mass. As in the HH production cross section uncertainties, these are only included when computing the limit with respect to the SM expectation.



(a) b-tagging efficiency uncertainties

Figure 5.11: b-tagging efficiency scale factor uncertainty variations on the DNN discriminant in the Resolved, 2 b-tag category and the $\tau_h\tau_h$ channel in 2018 in the light-flavor (left) and heavy-flavor (right) regions. Contributions of all expected backgrounds are added.

Background cross sections

For the processes modelled exclusively from simulated events, several normalization uncertainties are included due to the imperfect knowledge of the process cross section. Their values are reported in Tab. 5.2. These uncertainties come from the modelling of the PDFs (denoted in the table as pdf_*) and the determination of α_S (strong interaction coupling) and QCD scale (energy scale used to absorb infrared divergences), denoted as α_S_* and QCDscale_*, respectively.

VBF dipole recoil uncertainty

The default CMS showering Pythia8 configuration used in the standard MC production does not provide a good modelling of the third leading jet distribution for the VBF signal [107]. This can be improved by setting the dipole recoil option, available in recent Pythia8 versions, to ON. Additional VBF signal samples with κ_{2V}=1 and κ_{2V}=2 (with the other couplings set to their SM expectations) were produced for 2017 and 2018 using this alternative option, and have been used to compute a normalization systematic uncertainty to the standard VBF signal samples due to this effect.

The uncertainty is computed per event category taking the most conservative value between:

- The ratio of integrated yields in the dipole recoil ON/OFF samples augmented by its statistical uncertainty.
- The ratio in normalised per bin yields in the corresponding DNN distributions,

Uncertainty	Value (%)
QCDscale_ZH	+3.8/-3.0
pdf_ZH	±1.3
α_S_ZH	±0.9
QCDscale_WH	+0.5/-0.7
pdf_WH	±1.7
α_S_WH	±0.9
QCDscale_t$\bar{t}$H	+5.8/-9.2
pdf_t$\bar{t}$H	±3.0
α_S_t$\bar{t}$H	±2
QCDscale_ggH	±3.9
pdf_ggH	±1.9
α_S_ggH	±2.6
QCDscale_qqH	+0.4/-0.3
pdf_qqH	±2.1
α_S_qqH	±0.5
QCDscale_singleT	+4.2/-3.5
QCDscale_tW	±5.4
QCDscale_W	+0.8/-0.4
QCDscale_EWK	±2
QCDscale_VV	±10
QCDscale_VVV	±10

Table 5.2: Theoretical uncertainties on the background production cross sections.

where the yield per DNN bin has been normalized by the yield in the highest statistics bin in the standard VBF sample with dipole recoil OFF.

An example of both computations is shown in Fig. 5.12 for the 2018 samples with κ_{2V}=1 in the VBF subcategory.

Category / Channel	$\tau_e\tau_h$	$\tau_\mu\tau_h$	$\tau_h\tau_h$
Boosted	20.4	39.3	19.8
Resolved, 1 b-tag	14.5	11.8	21.9
Resolved, 2 b-tag	26.9	15.4	17.1
VBF subcategory	16.3	18.7	9.0
ggF subcategory	55.2	50.4	33.7
t$\bar{t}$ subcategory	42.0	41.7	56.7
t$\bar{t}$H subcategory	67.2	69.6	70.5
DY subcategory	46.7	48.3	52.3

Table 5.3: VBF dipole recoil uncertainty values in %.

The final uncertainty has been set to be the same for all six VBF processes used in the signal modelling in all years for each category and channel by taking the most

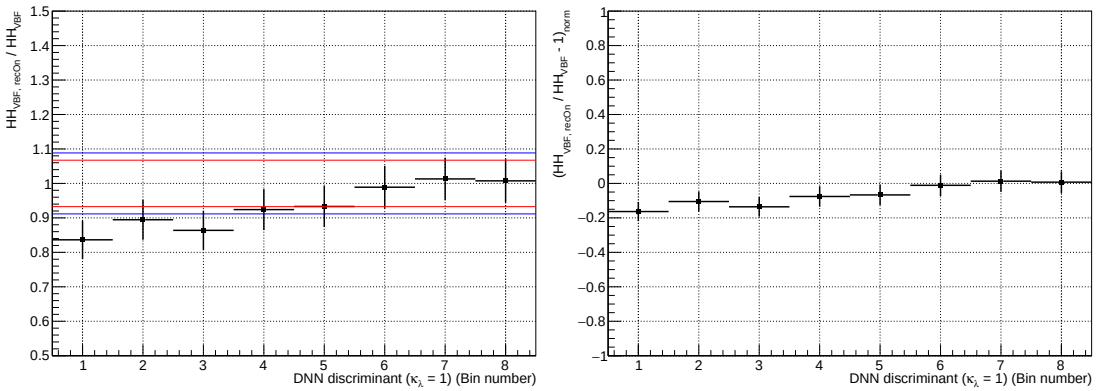


Figure 5.12: Dipole recoil ON/OFF uncertainty tests for the 2018 VBF samples with κ_{2V}=1 in the VBF subcategory. Left: Ratio between the DNN distributions. Red lines show the ratio of integrated yields and its symmetric with respect to 1 and blue lines after augmenting this ratio by the statistical uncertainty. The dipole recoil uncertainty (up and down templates) derived from this plot is taken by substracting 1 to the blue lines. Right: Ratio between the DNN distributions after normalizing by the yield in the bin with the highest statistics. The uncertainty (up and down templates) derived from this plot is taken from the bin with the largest deviation in absolute value and its opposite.

conservative value between 2017 and 2018 and both κ_{2V}values. The uncertainty can be seen in Table 5.3 and is considered as correlated between years and categories, but uncorrelated between the three channels.

5.4 Results

The results shown in this section are obtained by performing a binned maximum likelihood fit of the DNN prediction in eight categories per channel simultaneously for all three years, resulting in a total of 72 input distributions. The three types of nuisance parameters introduced in Section 5.1.2 are considered, being the normalization and systematic uncertainties further described in Section 5.3.

The DNN output distributions in 2018 for the $\tau_h\tau_h$ channel in the eight analysis categories is shown in Fig. 5.13. The binning schemes used in the DNN distributions are chosen to minimise the expected upper limit, while keeping as low as possible the number of bins per distribution and ensuring the stability of the fit. Each bin is required to contain at least one $t\bar{t}$ and one Z/γ^* + jets event and to have a weighted background yield larger than 0.03. The uncertainty bands shown in the distributions are the post-fit ones. The other channels and years can be found in Ref. [108].

The expected and observed limits for the ggF + VBF HH production cross section at the SM and those for the VBF only HH production cross section are listed in

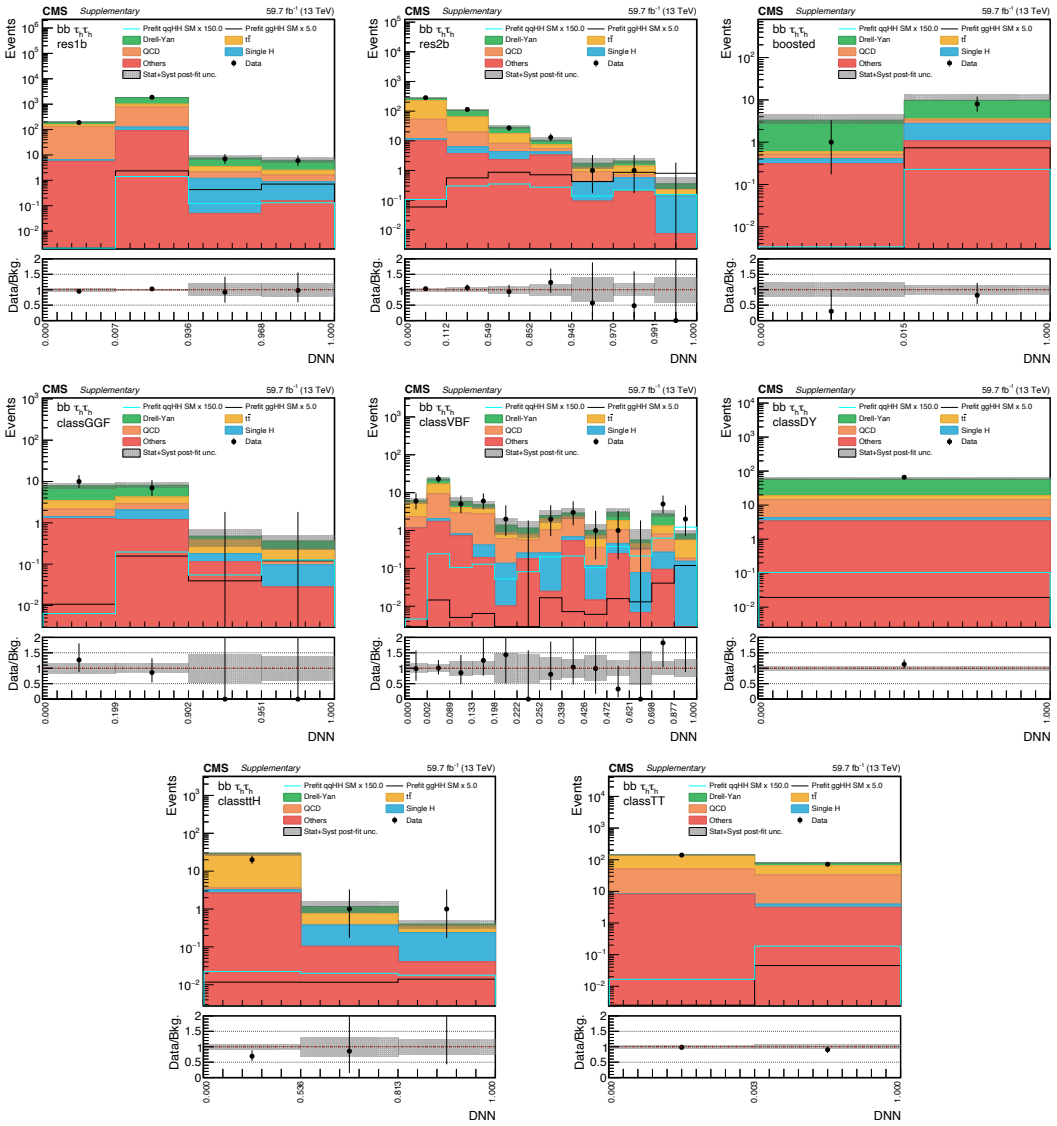


Figure 5.13: DNN output distributions in the $\tau_h\tau_h$ channel in 2018 for the eight analysis categories. The shaded band in the plots represents the statistical plus systematic uncertainty. Extracted from [108].

Tables 5.4 and 5.5, respectively. A visual representation of these limits is provided in Fig. 5.14.

	Obs. (Exp.) limit on $\sigma_{\mathrm{ggF+VBF}}(\mathrm{pp} \to \mathrm{HH})/\sigma^{\mathrm{theory}}_{\mathrm{ggF+VBF}}$	Obs. (Exp.) limit on $\sigma_{\mathrm{ggF+VBF}}(\mathrm{pp} \to \mathrm{HH})$ [fb]
2016	8.9 (10.6)	272 (324)
2017	9.5 (11.7)	291 (356)
2018	5.5 (8.2)	169 (249)
Combined	3.3 (5.2)	102 (159)

Table 5.4: Expected and observed upper limits on the ratio of the experimentally expected at 95% CL for the SM point, where $\sigma^{\mathrm{theory}}_{\mathrm{ggF+VBF}} = 32.776$ fb is the sum of the ggF and VBF HH production modes cross sections.

	Obs. (Exp.) limit on $\sigma_{\mathrm{VBF}}(\mathrm{pp} \to \mathrm{qqHH})/\sigma^{\mathrm{theory}}_{\mathrm{VBF}}$	Obs. (Exp.) limit on $\sigma_{\mathrm{VBF}}(\mathrm{pp} \to \mathrm{qqHH})$ [fb]
2016	283 (357)	487 (616)
2017	280 (392)	485 (676)
2018	241 (226)	414 (391)
Combined	124 (154)	212 (266)

Table 5.5: Expected and observed upper limits on the ratio of the experimentally expected at 95% CL for the SM point, where $\sigma^{\mathrm{theory}}_{\mathrm{VBF}} = 1.726$ fb is the VBF HH production mode cross section.

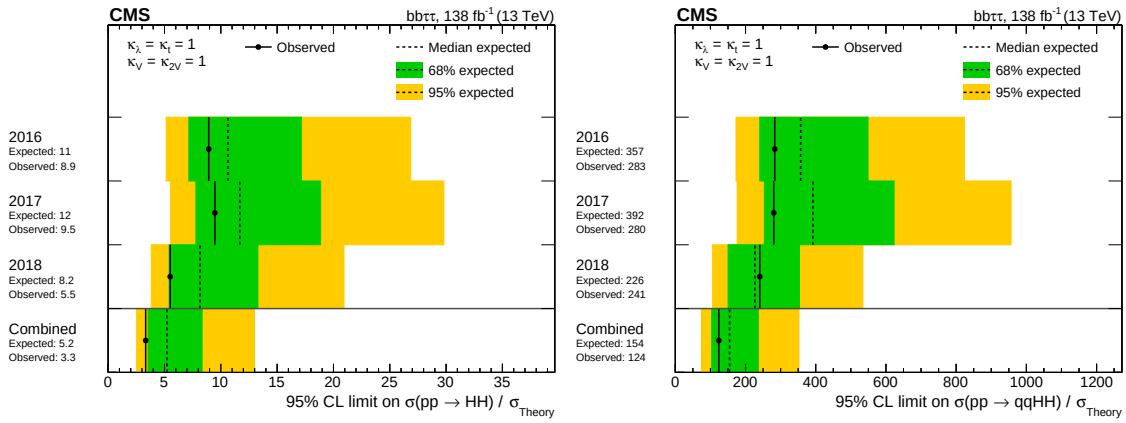


Figure 5.14: Expected and observed upper limits on the ratio of experimentally expected (left) ggF + VBF ($\sigma_{\mathrm{ggF+VBF}}$) and (right) VBF only ($\sigma_{\mathrm{VBF}}$) HH production cross section and the expectation from the SM (σ_{theory}) at 95%, split by the different years and combined for the full Run 2 dataset. Extracted from [109].

In order to study the Higgs couplings involved, the upper limits on the HH production cross section can also be computed for fixed values of the parameters.

Fig. 5.15 shows the scans of the expected and observed upper limits on the HH production cross sections as functions of κ_λ (ggF + VBF) and κ_{2V} (VBF only). With 95% CL, κ_λ gets constrained between $-1.7 < \kappa_\lambda < 8.7$ (observed) and $-2.9 < \kappa_\lambda < 9.8$ (expected) and κ_{2V} between $-0.4 < \kappa_{2V} < 2.6$ (observed) and $-0.6 < \kappa_{2V} < 2.8$ (expected). In both cases, the other couplings are set to 1 (i.e. SM value).

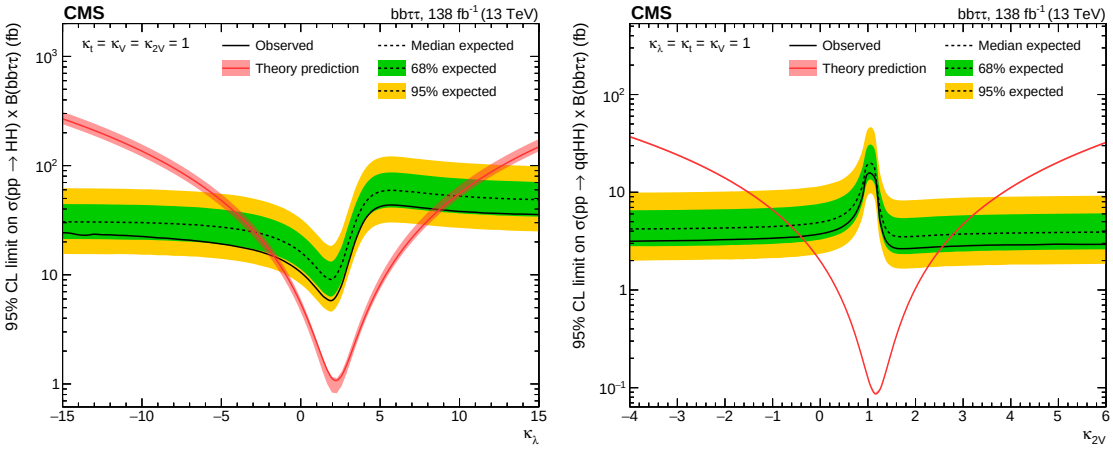


Figure 5.15: Expected and observed upper limits on the experimentally expected (left) ggF + VBF ($\sigma_{\text{ggF+VBF}}$) HH production cross section as functions of κ_λ and (right) VBF only (σ_{VBF}) HH production cross section as functions of κ_{2V}. In both cases, all the other couplings are set to their SM expectations. The red solid lines show the theoretical prediction for the HH production cross sections and its uncertainty (red shaded bands). Extracted from [109].

For both the ggF and VBF results reported, the largest sensitivity comes from the $\tau_h\tau_h$ channel, followed by the $\tau_\mu\tau_h$ and the $\tau_e\tau_h$ channels. The Resolved, 2 b-tag category provides the largest sensitivity when constraining the κ_λ value, while the VBF subcategory is the most dominant in the κ_{2V} measurement.

5.5 Comparison with other HH analyses

In this section a comparison between the results shown above and the ones provided by other analyses is performed.

First, within CMS the most direct comparison is with the HH $\rightarrow$ bb$\tau\tau$ analysis using 2016 data published in 2018 [81]. In this analysis, the signal extraction was performed by fitting the m_{T2} distribution [110] and a BDT was used in order to identify and reject the $t\bar{t}$ background. The observed (expected) limit on the measured HH production cross section at the SM point times the bb$\tau\tau$ branching fraction was found to be 75.4(61.0) fb, correspondent to 31.5(25.5) times the cross section predicted by the SM. These numbers can be directly compared with the upper limits obtained in latest analysis using only 2016 data, shown in Tab. 5.4 (observed (expected) limit of 8.9(11.0) times the SM prediction). Therefore, in the

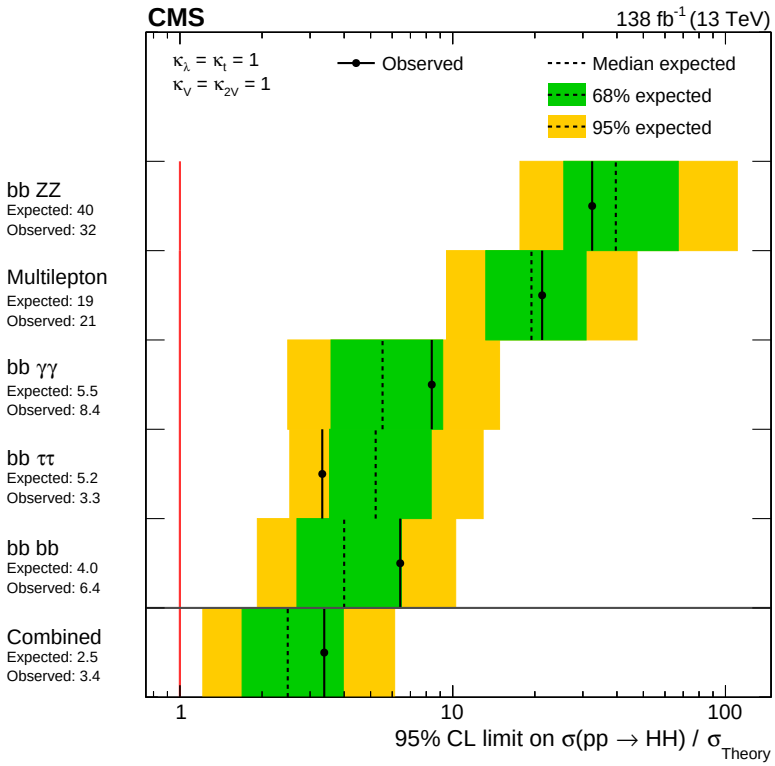


Figure 5.16: Expected and observed limits on the ratio of experimentally estimated production cross section and the expectation from the SM σ_{Theory} in searches using different final states and their combination. The search modes are ordered, from upper to lower, by their expected sensitivities from the least to the most sensitive. The overall combination of all searches is shown by the lowest entry. Extracted from [14].

full Run-2 analysis we can see an improvement of ∼58% in the expected limit and ∼70% in the observed limit. Regarding the κ_λ measurement, the 2016 analysis obtained observed (expected) constraints of $-18(-14) < k_\lambda < 26(22)$, a much wider range than the one obtained in the full Run-2 analysis. These new results benefit from the improved trigger strategy and the adoption of new techniques not only at analysis level, with the new neural network approaches used for categorization and signal extraction, but in the identification and reconstruction algorithms developed within the CMS Collaboration.

Considering the full Run-2 dataset, the ATLAS Collaboration also performed a HH → bb$\tau\tau$ search [111, 112]. Similarly to the CMS analysis, the signal extraction is performed by fitting the output score of a multivariate approach. Assuming a $\sigma^{\text{theory}}_{\text{ggF+VBF}} = 32.7$ fb, the observed (expected) upper limit on the HH production cross section obtained was 4.7(3.9) times $\sigma^{\text{theory}}_{\text{ggF+VBF}}$. The value of κ_λ is constrained between $-2.5 < \kappa_\lambda < 9.3$ (observed) and $-1.9 < \kappa_\lambda < 9.1$ (expected). These results are a bit more stringent than the CMS ones.

With respect to the VBF HH production, ATLAS did not perform a dedicated

analysis for this production mode, although some results on the production cross section and the κ_{2V} constraints were provided in the combination of ATLAS HH results [112]. The observed (expected) upper limit on the VBF HH production cross section was found to be 433(366) fb, while κ_{2V} is constrained between $-0.5 < \kappa_{2V} < 2.7$ (observed and expected). The value of the upper limit is much larger with respect to the one obtained in the CMS analysis. In fact, the CMS HH $\rightarrow$ bb$\tau\tau$ analysis provides all the decay channels studied in ATLAS. Regarding κ_{2V}, its constraints are quite similar to the ones obtained in the CMS analysis.

Regarding other CMS HH searches with Run-2 data, a combination of the CMS HH results was published to celebrate the tenth anniversary of the Higgs boson discovery [14]. This combination includes five final states, and the bb$\tau\tau$ analysis ranks among the first three with the highest sensitivity (see Fig. 5.16).

5.6 Future prospects

Given the importance of the Higgs pair production in the SM and to test BSM theories, searches for this production will be conducted with data taken during Run 3 (2022-2025) and HL-LHC (2029-).

During Run 3, CMS will collect around 220 fb^{-1}. With this increase in statistics, the expected upper limit on the inclusive HH production cross section obtained by the HH $\rightarrow$ bb$\tau\tau$ analysis could be lowered to $\sim$ 3.2 times the SM prediction. However, new techniques are being studied already for their inclusion in a future Run 3 analysis, such as the identification of τ_h and jets using particle clouds [113]. Additionally, a new trigger strategy has been studied for the $\tau_h\tau_h$ channel in Run 3, providing a sensitivity increase of up to 10%. This trigger strategy is described in Chapter 6.

The prospects for the study of the Higgs boson pair production after HL-LHC have been conducted by both CMS and ATLAS experiments [14, 94, 114]. These projection studies assume a total integrated luminosity of 3000 fb^{-1}, and take into account the upcoming detector upgrades and state-of-the-art analysis techniques.

The CMS HH $\rightarrow$ bb$\tau\tau$ projection [94] follows an analysis strategy similar to the Run 2 analysis presented in this thesis: two b-tagged jets and two leptons are selected, although only the fully-hadronic ($\tau_h\tau_h$) and semi-leptonic ($\tau_\mu\tau_h$, $\tau_e\tau_h$) decay channels are considered. Signal extraction is performed with a neural network that follows the same approach as the one described in this thesis. By fitting the expected DNN output distributions in the three final states considered, the expected upper limit on the measured HH production cross section at the SM point times the bb$\tau\tau$ branching fraction was found to be 1.4 times the SM prediction. Fig. 5.17 shows a comparison between the upper limits obtained with early Run 2 data, full Run 2, and the projection for HL-LHC for the bbbb, bb$\tau\tau$, and bb$\gamma\gamma$ CMS analyses, and the results for the three decay channels combined. Something to be noted is that, as the combined limit after HL-LHC is expected to be below unity, the sensitivity could be sufficient to establish the existence of the SM HH production. In fact,

when combining the results of CMS and ATLAS after HL-LHC, a significance of 4 standard deviations could be achieved, providing an evidence of the existence of SM HH production [114].

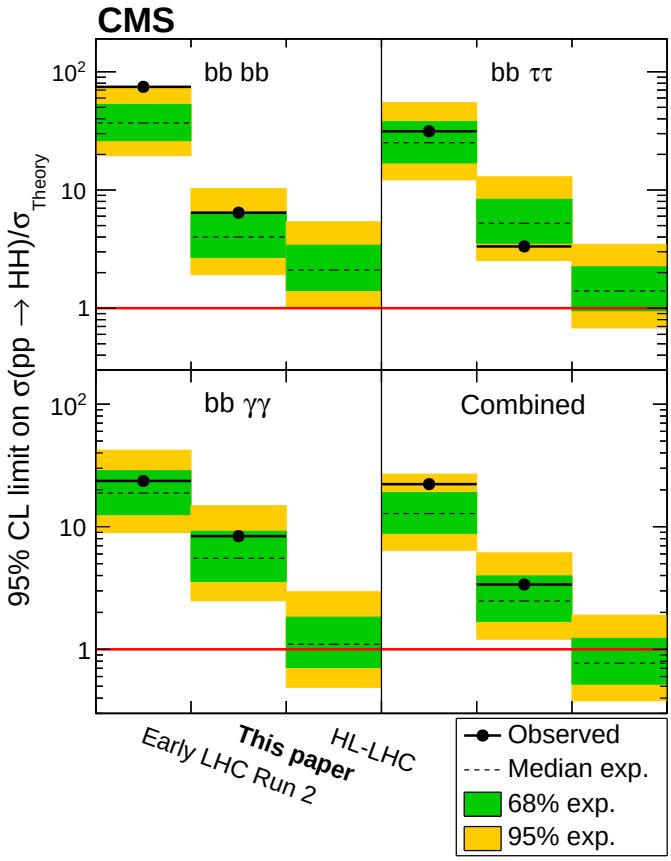


Figure 5.17: Expected and observed limits on HH production in different datasets: early LHC Run 2 data (35.9 fb^{-1}), present results using full LHC Run 2 data (138 fb^{-1}), and projections for HL-LHC (3000 $fb^{-}1$). Extracted from [14].

Even if these projections already assume an improvement on the theoretical predictions (scaling down the theoretical uncertainties) and an upgraded CMS detector (scaling down the experimental uncertainties), there is still room for improving these results. For instance, the τ_h and b jet identification algorithms used in the Run 2 analysis (DeepTau and HH-btag, respectively) are not considered in these projections. Moreover, all techniques that are being studied for Run 3 could boost the sensitivity in this future HL-LHC analyses.

5.7 Summary

This chapter includes the results for the search for Higgs boson pair production using Run 2 proton-proton collision data at 13 TeV, where one Higgs boson decays into two τ leptons and the other into two b quarks. Two sets of results have been

shown: inclusive results, where both ggF and VBF production modes are considered as signal, and VBF-only results, where only the VBF production mode is considered as signal and the ggF signal strength is fixed to its SM prediction.

The statistical procedure followed to obtain these results has been described. It is based on a binned maximum likelihood fit for the DNN discriminator described in Chapter 4. A complete description of the systematic uncertainties affecting the analysis has been included, coming from experimental and theoretical sources. Among the first ones, several uncertainties affecting the QCD contribution are included, coming from a set of validity tests I performed to study the QCD background estimation method. Additionally, among the theoretical uncertainties I have studied the uncertainty coming from the modelling of the dipole recoil in the HH VBF samples.

The observed (expected) upper limit with a 95% CL on inclusive HH production cross section corresponds to 3.3 (5.2) times the theoretical cross section predicted by the SM. For the VBF-only production, the limit is set at 124 (154) times the SM prediction. Additionally, 95% constraints have been established on the Higgs self-coupling (κ_λ) and its coupling to the vector bosons (κ_{2V}), as deviations of these couplings could be produced by BSM effects. For κ_λ, the observed (expected) constraint has been set to $-1.7 < \kappa_\lambda < 8.7$ ($-2.9 < \kappa_\lambda < 9.8$). For κ_{2V}, the obtained constraint was $-0.4 < \kappa_{2V} < 2.6$ ($-0.6 < \kappa_{2V} < 2.8$).

Regarding the inclusive results, the full Run-2 analysis shows a large improvement with respect to the previously published CMS HH $\rightarrow$ bb$\tau\tau$ results (using 2016 data). This improvement comes not only from the larger statistics but from the inclusion of a new trigger strategy and the adoption of new techniques at analysis level. Additionally, new particle identification and reconstruction algorithms developed within the CMS Collaboration have been considered. With respect to the ATLAS Collaboration HH $\rightarrow$ bb$\tau\tau$ search, the results are very similar. Within the CMS Collaboration, the HH $\rightarrow$ bb$\tau\tau$ result ranks among the most sensitive HH decay channels under study. Regarding the VBF-only results, the CMS HH $\rightarrow$ bb$\tau\tau$ analysis provides a better sensitivity than the ATLAS analysis.

Looking into the future, the CMS and ATLAS Collaborations have conducted projections for the HH analyses after HL-LHC. For the CMS HH $\rightarrow$ bb$\tau\tau$ analysis, an expected upper limit on the HH production cross section times the branching fraction of 1.4 times the SM prediction is obtained. A combination of the CMS and ATLAS results after HL-LHC is expected to provide the first evidence of HH production, with a significance of 4 standard deviations. However, there is room for improvement thanks to, among others, the inclusion of new particle identification techniques or a new trigger strategy.

Chapter 6

The double-τ_h + jet trigger

In view of Run 3, a lot of effort has been dedicated in order to maximize the collection efficiency from physics processes such as H $\rightarrow \tau\tau$ and HH $\rightarrow$ bb$\tau\tau$. In these two cases, in most events the produced $\tau\tau$ pair is accompanied by one or more jets coming from the hard-scattering processes or from QCD radiation (in addition to the 2 b-jets coming from the other H in the HH $\rightarrow$ bb$\tau\tau$ process), as seen in Fig. 6.1.

In both H $\rightarrow \tau\tau$ and HH $\rightarrow$ bb$\tau\tau$ analyses, the trigger strategy during Run-1 and Run-2 for the $\tau_h\tau_h$ channel consisted on a double-τ_h selection, with a p_T threshold of 32 and 35 GeV at L1 and HLT level, respectively, during 2018 data-taking. By including in the menu a new trigger that requires two τ_h and an additional jet (whose p_T spectrum at L1 is shown in Fig. 6.2), the p_T threshold on the τ_h could be lowered. By doing so, as more τ_h will be available for triggering (see Fig. 6.3), the acceptance of both analyses could be enlarged.

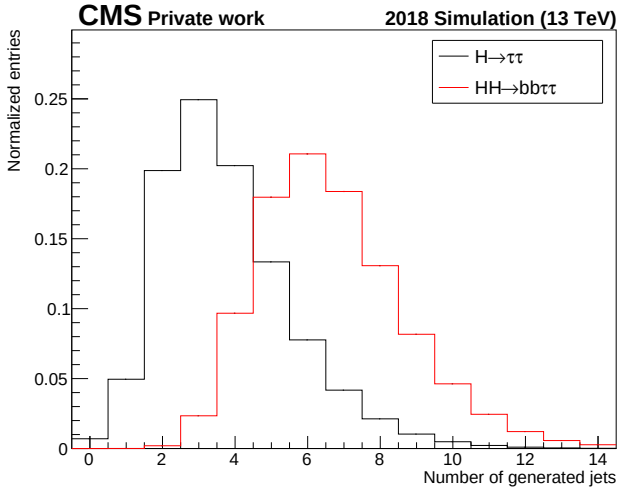


Figure 6.1: Total number of jets per event at generator level for H $\rightarrow \tau\tau$ (black) and HH $\rightarrow$ bb$\tau\tau$ (red) processes, both produced via ggF.

The following sections discuss the design and performance of a new L1 seed with

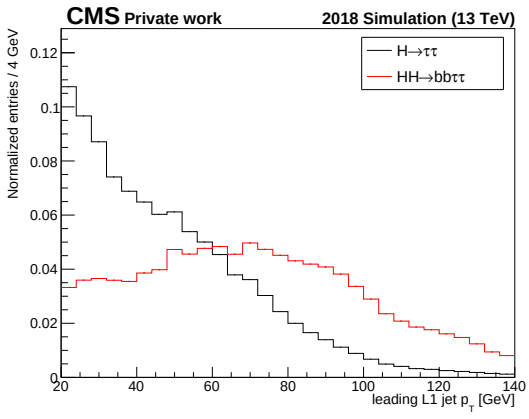


Figure 6.2: L1 leading jet p_T for the H $\rightarrow \tau\tau$ (black) and HH $\rightarrow$ bb$\tau\tau$ (red) processes, both produced via ggF.

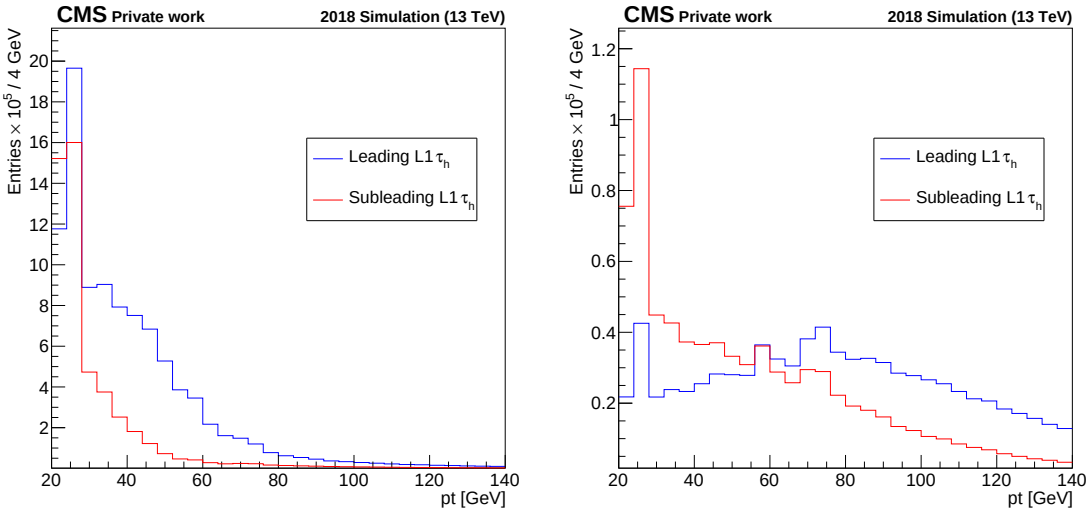


Figure 6.3: L1 leading and subleading τ_h p_T for the H $\rightarrow \tau\tau$ (left) and HH $\rightarrow$ bb$\tau\tau$ (right) processes, both produced via ggF. In both cases, below the double-τ_h trigger p_T threshold (32 GeV in 2018), many τ_h would be available for triggering.

two τ_h and one jet to be included in the L1 Menu and a new HLT path seeded by this new L1 seed. The goal is that the newly defined L1 seed and HLT path add as little rate as possible but increase the final acceptance on the H $\rightarrow \tau\tau$ and HH $\rightarrow$ bb$\tau\tau$ analyses. The acceptance increase at trigger level is enhanced by the possibility of loosening correspondingly the offline selection cuts. In the H $\rightarrow \tau\tau$ analysis [115], the two τ_h are selected if their $p_T > 40$ GeV, while the selection on the jets is different in the four categories defined. In this study, only two categories are considered, obtained as a simplification of the most sensitive analysis categories: the 1-jet, high p_T category, where we consider events with one jet with $p_T > 70$ GeV; and the 2-jet category, where we select events with at least two jets with $p_T > 30$ GeV.

In the HH $\rightarrow$ bb$\tau\tau$ analysis, as was described in Chapter 4, events with at least two τ_h with $p_T > 40$ GeV and at least two jet candidates with $p_T > 20$ GeV are selected.

The goal of this study was to include the new trigger seed and path in the menus devoted to collisions during 2022, where an average PU of 53 and 2748 bunches per beam were expected. The design and optimization of the triggers, however, will consider samples from 2018, which will be adapted to the expected 2022 conditions. To estimate the rate added by the new trigger, zero bias samples will be considered. These samples are characterised by being unbiased by the trigger decision, as they only require the crossing of two bunches.

Section 6.1 describes the optimization procedure followed to include the L1 double-τ_h+jet seed. Section 6.2 shows the performance results of the new L1 seed obtained from Run 3 data and simulated samples. Section 6.3 describes the HLT double-τ_h+jet path included in the menu, and its performance is evaluated in Section 6.4.

The development of this new trigger was a substantial part of my thesis, from the inclusion of the new L1 seed after a long optimization period, to the development of the corresponding HLT path. I am the main responsible for the performance studies from both L1 seed and HLT path and the first studies on the sensitivity gained thanks to this new trigger in a future Run 3 HH $\rightarrow$ bb$\tau\tau$ analysis.

6.1 The L1 double-τ_h + jet trigger

In the L1 menu used during the Run-2 data taking, the only seeds that consider hadronic τ decays were the `L1_DoubleIsoTauXer2p1` seeds, where two isolated τ_h with $p_T \geq X$ GeV and $|\eta| < 2.1$ are selected. The threshold X varied during the data-taking, with a value of $X = 32$ GeV during 2018. The strategy followed in this study is to include a new seed in the L1 Menu with two isolated τ_h objects and one jet object identified by the μGT. This new seed complements the already present double-τ_h seed with the same 32 GeV threshold (or with a larger p_T threshold in order to keep the rate as low as possible). A particular treatment has to be given to seeds involving τ_h and jet objects simultaneously, as both, τ_h and jets, are reconstructed as purely calorimetric objects. Therefore, all τ_h objects enter the L1 jet collection, and some of the L1 jets also appear in the L1 τ_h collection. This effect is illustrated in Fig. 6.4, where the minimal ΔR between all L1 τ_h's and jets is computed. At small values of ΔR a peak is observed, coming from the overlap between the L1 τ_h and jet collections. To reduce this effect, only L1 jets that are separated more than $\Delta R > 0.5$ with the selected τ_h are considered. This operation is called overlap removal [49]. It was already present in the μGT before the inclusion of the new seed, both in the firmware and its software emulator. However, while including the new seed, some discrepancies arose between both implementations, so the μGT software emulator had to be was modified accordingly.

As previously stated, this new double-τ_h+jet seed (encoded as `L1_DoubleIsoTauYer2p1_JetZ_RmOvlp_dR0p5`, where Y and Z are the p_T thresholds for the τ_h and jet,

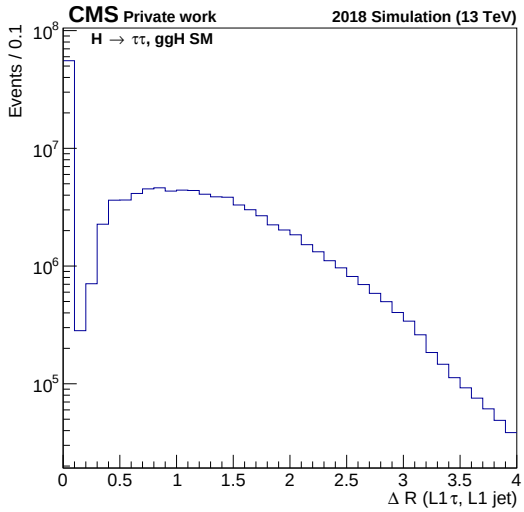


Figure 6.4: Minimal ΔR between all L1 τ_h's and jets per event in a H $\rightarrow \tau\tau$ ggF simulated sample. The peak at small ΔR values is due to the overlap between the L1 τ_h and jet collections.

respectively) does not aim to replace the present `L1_DoubleIsoTau32er2p1` seed, but to complement it. However, in order to make some rate available for the new seed, increasing the double-τ_h p_T threshold to a value X could be needed. The rates corresponding to different values of the double-τ_h threshold are shown in Fig. 6.5. After the inclusion of the new seed, the aim is to keep the total rate of both seeds as close as possible to the original rate of the `L1_DoubleIsoTau32er2p1` seed, so a target rate of $\sim$ 17 kHz is considered.

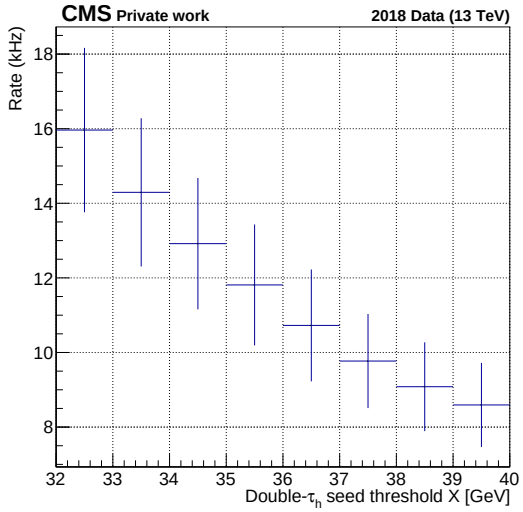


Figure 6.5: Rate of the `L1_DoubleIsoTauXer2p1` trigger seed under Run 3 conditions (PU 53, 2748 bunches) as a function of the X threshold.

Boolean	Offline p_T selections		
	$H \to \tau\tau$		$HH \to bb\tau\tau$
	1-jet, high p_T	2-jet	
`PassOfflDoubleTau32`	$p_T^{\tau_1} > 40$ $p_T^{\tau_2} > 40$ $p_T^{j_1} > 70$	$p_T^{\tau_1} > 40$ $p_T^{\tau_2} > 40$ $p_T^{j_1} > 30$ $p_T^{j_2} > 30$	$p_T^{\tau_1} > 40$ $p_T^{\tau_2} > 40$ $p_T^{j_1} > 20$ $p_T^{j_2} > 20$
`PassOfflDoubleTauX`	$p_T^{\tau_1} > X + 8$ $p_T^{\tau_2} > X + 8$ $p_T^{j_1} > 70$	$p_T^{\tau_1} > X + 8$ $p_T^{\tau_2} > X + 8$ $p_T^{j_1} > 30$ $p_T^{j_2} > 30$	$p_T^{\tau_1} > X + 8$ $p_T^{\tau_2} > X + 8$ $p_T^{j_1} > 20$ $p_T^{j_2} > 20$
`PassOfflDoubleTauYJetZ`	$p_T^{\tau_1} > Y + 8$ $p_T^{\tau_2} > Y + 8$ $p_T^{j_1} > Z + 10$ $p_T^{j_1} > 70$	$p_T^{\tau_1} > Y + 8$ $p_T^{\tau_2} > Y + 8$ $p_T^{j_1} > Z + 10$ $p_T^{j_1} > 30$ $p_T^{j_2} > 30$	$p_T^{\tau_1} > Y + 8$ $p_T^{\tau_2} > Y + 8$ $p_T^{j_1} > Z + 10$ $p_T^{j_1} > 20$ $p_T^{j_2} > 20$

Table 6.1: Offline p_T selections associated to the L1 selections of `L1_DoubleIsoTau32er2p1`, `L1_DoubleIsoTauXer2p1` and `L1_DoubleIsoTauYer2p1_JetZ_RmOvlp_dR0p5` for the different categories, where τ_1 and τ_2 are the leading and subleading τ_h and j_1 and j_2 are the leading and subleading jets. All thresholds are expressed in GeV.

As the aim to increase the signal acceptance, the acceptance gain when considering the two new seeds instead of the reference `L1_DoubleIsoTau32er2p1` seed is computed. Four signal samples are considered, two ggF and VBF $H \to \tau\tau$ and two ggF and VBF $HH \to bb\tau\tau$. These acceptance gains are computed not only from the L1 objects that pass the required selection cuts (grouped in the booleans `PassL1DoubleTau32`, `PassL1DoubleTauX` and `PassL1DoubleTauYJetZ`), but also applying selections to the objects obtained by the offline reconstruction that evolve accordingly to the L1 thresholds considered. In general, this evolution means that, for a given L1 p_T threshold T, the offline threshold is obtained as $T + \Delta T$. As in both target analyses the offline τ_h p_T threshold is set to 40 GeV when using a L1 threshold of 32 GeV, a $\Delta T = 8$ GeV for the τ_h is considered. For the jets, a $\Delta T = 10$ GeV is taken. All p_T selections for the different categories are summarised in Table 6.1, where one boolean associated to each seed has been defined in order to group the needed offline selections. On top of these p_T cuts, all offline jets must be within $|\eta| < 4.7$ and pass tight jet ID and loose PU jet ID and an overlap removal criteria with the offline τ_h ($\Delta R(\text{jet}, \tau_h) > 0.5$).

Thus, the acceptance gain when adding a double-τ_h+jet seed and modifying the

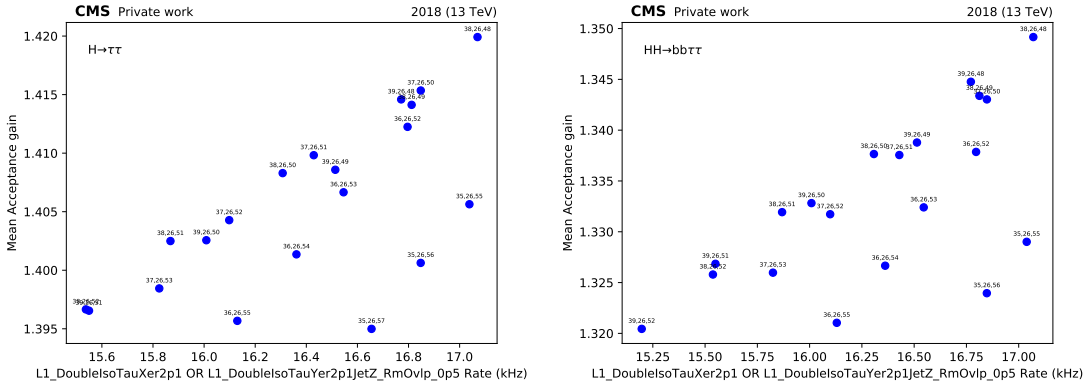


Figure 6.6: Mean acceptance gain for the H $\to \tau\tau$ (left) and HH $\to$ bb$\tau\tau$ (right) processes when considering the new trigger strategy for given (X, Y, Z) p_T thresholds.

threshold of the present double-τ_h seed can be defined as

$$g(X,Y,Z) = \frac{N[(\texttt{PassL1DoubleTauX \&\& PassOfflDoubleTauX}) \;||\; (\texttt{PassL1DoubleTauYJetZ \&\& PassOfflDoubleTauYJetZ})]}{N[\texttt{PassL1DoubleTau32 \&\& PassOfflDoubleTau32}]} \tag{6.1}$$

where N is the number of events that pass the conditions required for given values of the X, Y and Z p_T thresholds. The average acceptance gains obtained for different sets of thresholds (X, Y, Z) for the samples and categories considered is shown in Fig. 6.6. Only the 20 sets with the highest acceptance gain in a 15-17 kHz rate range are displayed. Given these results, the decision was to consider thresholds $Y = 26$ and $Z = 55$ GeV for the double-τ_h+jet seed, i.e. adding the seed encoded as `L1_DoubleIsoTau26er2p1_Jet55_RmOvlp_dR0p5`. In addition, another double-τ_h+jet seed with thresholds $Y = 26$ GeV and $Z = 70$ GeV was added to the L1 menu as a backup seed in case the rate coming from the main seed ends up being larger than expected. Rates computed with Run 3 conditions for these new seeds and their main acceptance gains on top of the `L1_DoubleIsoTau32er2p1` seed for the two processes considered are shown in Table 6.2. Note that the H $\to \tau\tau$ column groups both 1-jet, high p_T and 2-jet categories.

Trigger seed	Acceptance gain		Rate (kHz)
	H $\to \tau\tau$	HH $\to$ bb$\tau\tau$	
`L1_DoubleIsoTau26er2p1_Jet55_RmOvlp_dR0p5`	44%	36%	8.3 ± 2.8
`L1_DoubleIsoTau26er2p1_Jet70_RmOvlp_dR0p5`	34%	29%	5.3 ± 1.8

Table 6.2: Rates and mean acceptance gains for the two new double-τ_h+jet seeds.

6.2 Performance of the L1 double-τ_h + jet trigger

The LHC Run 3 started during summer 2022 with the first proton-proton collisions at $\sqrt{s} = 13.6$ TeV. In general, 2450 bunches were allocated per beam, resulting in an average PU of around 50 GeV at the start of the fill.

Before the start of the LHC Run 3 (summer 2022), the new L1 double-τ_h+jet seed (and its backup seed) were included in the L1 menu, so they can be used in physics analyses. Being included in the menu also allowed to study their rate and performance directly using 2022 collision data, where te seeds were enabled, and validating the expected results obtained by emulating the new seeds on 2018 data. Note that the expected beam conditions have changed between those results and the actual data taking, so the predictions need to be slightly modified.

In the following, performance results will be shown for the main L1 trigger seed.

Rate with 2022 data

Fig. 6.7 shows the rate obtained for the L1 double-τ_h+jet seed using a collision data sample taken during 2022 as a function of the average PU. Note how the rate scales almost linearly with respect to the PU. Regarding the 2018 estimations, by rescaling the 8.3 ± 2.8 kHz obtained with 2018 data for PU 53 and 2748 bunches to 2450 bunches, a value of 7.4 ± 2.5 kHz is obtained, which agrees with the values obtained with 2022 data.

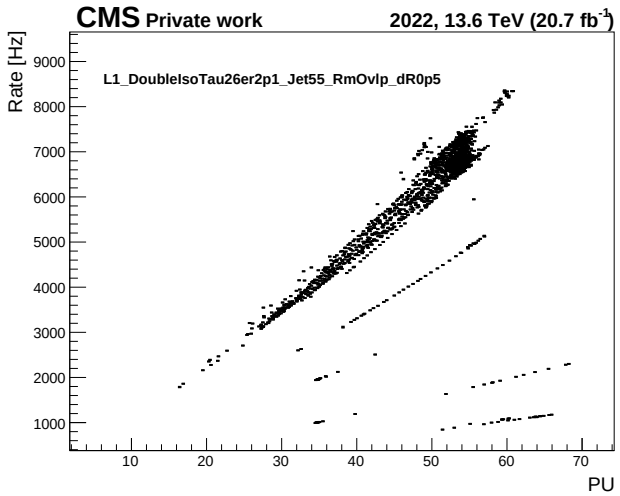


Figure 6.7: `L1_DoubleIsoTau26er2p1_Jet55_RmOvlp_dR0p5` seed rate as a function of the PU for a collision data sample collected during 2022. Outliers come from runs taken with smaller number of bunches.

Trigger efficiency

In order to measure the performance of a trigger selection, a good estimator is the trigger efficiency. This efficiency is computed as the fraction of events satisfying some requirements on the offline and trigger objects of the event over the ones satisfying the offline requirements. In general, this efficiency is computed with respect to the offline quantities related to the objects used in the trigger selection. For instance, in the double-τ_h+jet trigger, this efficiency will be measured with respect to some quantities related to a pair of τ_h and a jet. In a simple 1-D case, where the algorithm only depends on one variable from one object, the shape of the efficiency as a function of the corresponding offline variable is commonly known as a *turn-on* curve, which is a convolution of an step function (changing from 0 to 1 in the threshold used in the online selection) and a Gaussian resolution function, which represents the difference in resolution between the trigger and offline reconstruction. The more similar the trigger and offline reconstruction, the sharper would be the turn-on curve.

The efficiency for the double-τ_h+jet trigger can be obtained as the product of the efficiencies for each individual trigger object, i.e.

$$\text{efficiency}_{(\tau_1,\tau_2,\text{jet})} = \text{efficiency}_{\tau_1} \times \text{efficiency}_{\tau_2} \times \text{efficiency}_{\text{jet}}, \tag{6.2}$$

where the individual trigger objects should satisfy the same kinematic requirements as in the double-τ_h+jet trigger. A caveat from this method is that the efficiencies for each object are being considered independent from the other objects, while a correlation between them could appear. This particularity will be studied in future analysis.

In order to measure the efficiency for each single object, a tag and probe technique (already presented in previous chapters) is considered. In this method, we profit of the production of a known resonance that decays into two particles: a *tag* particle, which satisfy very tight identification and isolation requirements to ensure it comes from the decay of the resonance, and a *probe* particle, which is compatible with the resonance and is used to measure the efficiency.

In the double-τ_h analysis, the τ_h efficiency will be measured by considering the production of a Z boson that decays into two τ_h. One τ_h decays leptonically into a muon and two neutrinos; the other, hadronically. The muon from the first decay will be used for tagging, while the τ_h will be used for efficiency computation. The tau trigger efficiency will be measured using a newly developed monitoring seed coded as `L1_Mu18er2p1_Tau26er2p1`, a trigger that requires one muon with $p_T \geq 18$ GeV and $|\eta| < 2.1$, and a τ_h with $p_T \geq 26$ GeV and $|\eta| < 2.1$ (i.e. the same kinematic restrictions as the double-τ_h+jet seed under study with the exception of the isolation cut). Thus, the τ_h efficiency will be obtained as

$$\text{efficiency} = \frac{N_{\text{pass}}}{N_{\text{all}}}, \tag{6.3}$$

where N_{all} is the number of events firing a single muon L1 trigger (encoded as `L1_SingleMu22`), where one muon and at least one τ_h were reconstructed and the

muon is matched to the muon used for triggering. Several quality cuts are applied to the reconstructed muon ($p_T > 24$ GeV, $|\eta| < 2.1$), the reconstructed τ_h ($|\eta| < 2.1$, $d_z < 0.2$ cm, VVVLoose DeepTauVsmu, VVVLoose DeepTauVse and Medium DeepTauVsjet working points) and the $\mu\tau_h$ system ($\Delta R(\mu, \tau_h) > 0.5$, invariant mass between 30 and 80 GeV). N_{pass} is the number of events out of N_{all} where the offline τ_h was matched to a trigger-level τ_h associated to the monitoring seed `L1_Mu18er2p1_Tau26er2p1`. The trigger-level τ_h is required to be isolated in order to fully match the τ_h requirements in the double-τ_h+jet seed.

When computing the efficiency for a given trigger with a tag and probe method, the samples used for this calculation need to be unbiased at trigger level, i.e. the events in the sample must not be selected by any trigger that targets the probe object. CMS stores in a particular data set all events that fired a muon-related trigger. Therefore, without any bias, we can use this data set to obtain the τ_h and jet trigger efficiencies using real collision data, and then obtain the efficiency for the double-τ_h+jet trigger seed. The efficiency will also be computed with a Z + 2 jets simulated sample, so a comparison between the performance in real data and simulation can be obtained. Fig. 6.8 shows the τ_h efficiency with respect to the offline τ_h p_T and η for the data and simulated sample.

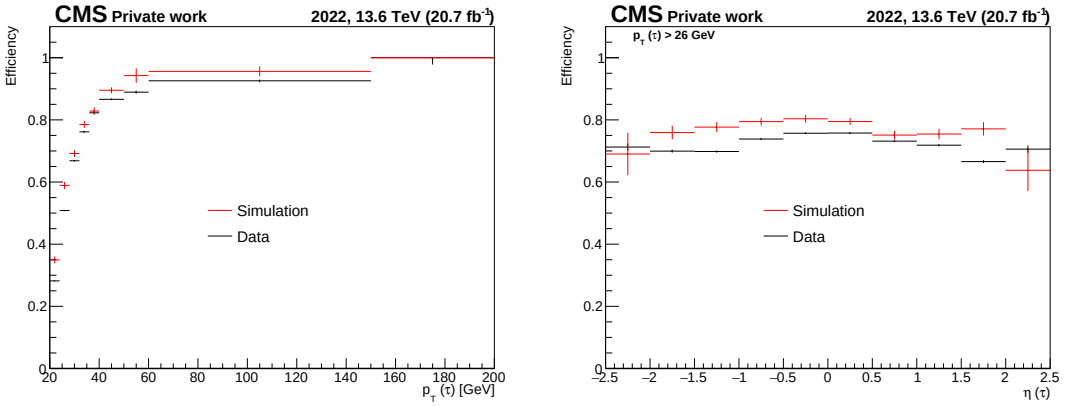


Figure 6.8: L1 τ_h trigger efficiency obtained with a tag and probe method with respect to the offline τ_h p_T (left) and η (right) for data events collected during 2022 (black) and simulated events from a Z + 2 jets sample (red). Simulated events are reweighed so their PU distribution matches the one from the data (see Fig. 6.10). Generator-level weights are applied to the simulated events. Note that a τ $p_T >$ 26 GeV cut is imposed in the efficiency with respect to τ η to match the τ p_T threshold at trigger level.

The trigger jet efficiency can be also computed in a similar way. In this case, N_{all} includes all events firing the `L1_Mu18er2p1_Tau26er2p1` trigger, where one muon, at least one τ_h, and at least one jet were reconstructed, and the muon and the τ_h are matched to the muon and the τ_h used for triggering. The previous quality cuts are applied for the muon and the τ_h, while the jets are required to satisfy the tight particle-flow jet identification working point, $\Delta R(\mu, \text{jet}) > 0.5$, and

$\Delta R(\tau_h, \text{jet}) > 0.5$. Out of these events, N_{pass} are the ones where the offline jet matched a trigger jet associated to a newly developed monitoring seed encoded as `L1_Mu18er2p1_Tau26er2p1_Jet55`. Fig. 6.9 shows the jet efficiency with respect to the offline jet p_T and η.

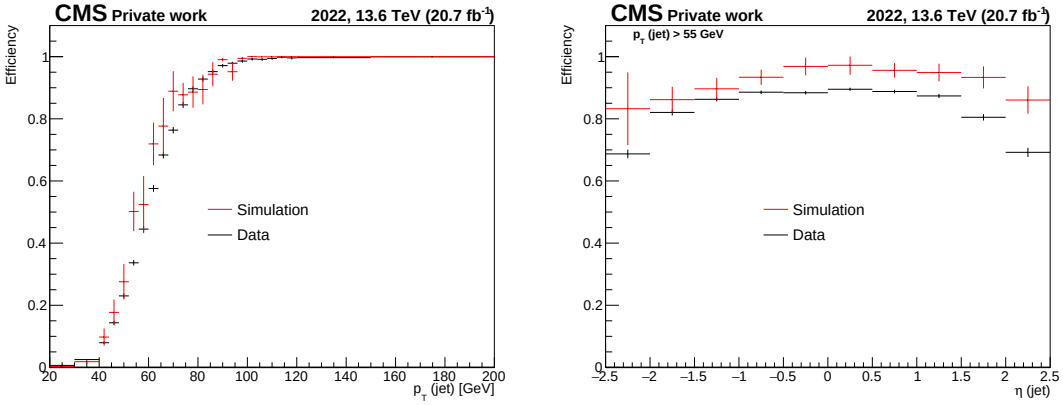


Figure 6.9: L1 jet trigger efficiency obtained with a tag and probe method with respect to the offline jet p_T (left) and η (right) for data events collected during 2022 (black) and simulated events from a Z + 2 jets sample (red). Simulated events are reweighed so their PU distribution matches the one from the data (see Fig. 6.10). Generator-level weights are applied to the simulated events. Note that a jet $p_T >$ 55 GeV cut is imposed in the efficiency with respect to jet η to match the jet p_T threshold at trigger level.

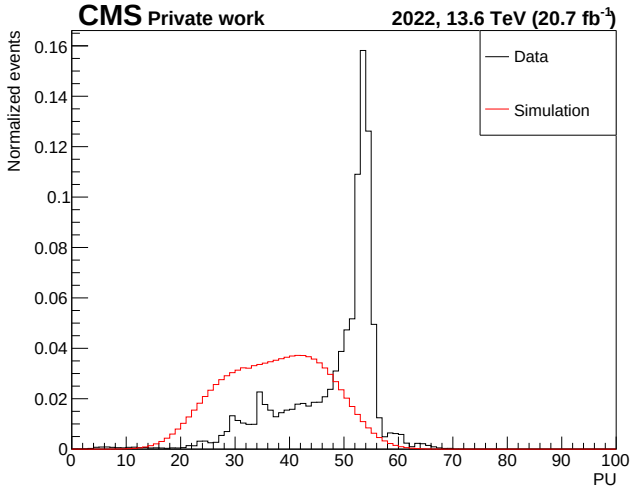


Figure 6.10: PU distribution for the data (black) and simulated (red) events used for the efficiency computation.

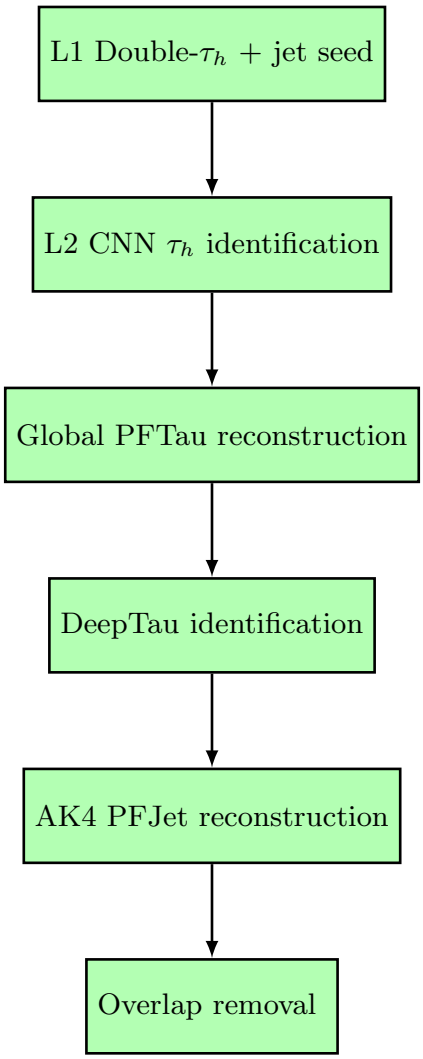


Figure 6.11: Sketch of the structure followed by the implemented HLT double-τ_h + jet path.

6.3 The HLT double-τ_h + jet path

In order to trigger on double-τ_h + jet events, an HLT path was designed and implemented, associated to the main L1 seed described in Section 6.1. Additionally, a backup HLT path was also implemented triggered by the backup L1 seed. The structure of these paths is shown in Fig. 6.11. This structure is very similar to the existing double-τ_h paths in Runs 1 and 2, but with some improvements developed for Run 3.

For an event to pass the HLT path, the first requirement is to pass the L1 seed requirements. On top of these L1 requirements, a typical HLT di-τ_h requirement is applied [47]: each L1 tau is used to build a Level-2 (L2) τ by associating the L1 tau all objects that fall into a ΔR cone of 0.5 with respect to it. Then, a

filter is added based on a Convolutional Neural Network (CNN), aimed to reduce the number of background objects entering the next steps of the path (more time consuming) while reducing the least possible the number of true τ filtered. This machine learning approach is one of the new additions to Run 3 τ paths, resulting in an increase of efficiency and a decrease of rate with respect to the approaches used in the previous data-taking periods.

After building the L2 τ, global PF τ are reconstructed, similarly as how it was shown in Section 3.2.4. PF τ are then filtered based on the same DeepTau algorithm as in the offline reconstruction level. This is another improvement introduced during Run 3, resulting in better signal purity and rate reduction.

At each step, at least a pair of τ leptons needs to survive the selection. Also, the two finally selected τ_h need to match the L1 τ used for triggering at L1 within a ΔR of 0.5, and their p_T must be higher than 30 GeV.

After selecting the two τ_h, the PF algorithm is used again to reconstruct the PF jets in the event. These PF jets will be selected if they satisfy an overlap removal criteria similar to the one used at L1: no matching (within a ΔR cone of 0.5) between the jet and a pair of PF τ that passed the whole HLT selection. On top of this criteria, the jet must match a L1 jet used for triggering and have a p_T higher than a given threshold. For the main HLT path (triggered with the L1 seed with the 55 GeV jet p_T threshold), this threshold is set to 60 GeV, while for the backup path, it is set to 75 GeV.

All these selections are grouped into two HLT paths, a main path and a backup path, encoded as `HLT_DoubleMediumDeepTauPFTau30_L2NN_eta2p1_PFJet60` and `HLT_DoubleMediumDeepTauPFTau30_L2NN_eta2p1_PFJet75`, respectively.

6.4 Performance of the double-τ_h + jet HLT path

Expected rate using 2018 data

To compute the expected rate from the new path, a similar approach to the one used in Section 6.1 is considered. The rate is estimated from a zero bias data sample taken during several 2018 runs corrected with the expected at that time Run-3 conditions (PU linearly scaled to 53 and 2748 colliding bunches). The rate obtained for this path is around 15.7 ± 3.8 Hz. Note that, as the expected rate by an HLT path is very small (around two to three orders of magnitude smaller than the correspondent L1 seed), a lot of zero bias events are needed to obtain a reasonably accurate result. In fact, even with all the 2018 runs that were used to obtain this value, only 15 events were able to pass the trigger, so the statistical uncertainty drives the result obtained.

For the sake of validating this result, the rate for the main L1 seed (`L1_DoubleIsoTau26er2p1_Jet55_RmOvlp_dR0p5`) described in Section 6.1 was also obtained in the same sample. The resulting value was 8.94 kHz, in agreement with the value obtained in the previous section.

Rate using 2022 data

In parallel to the inclusion of the new L1 double-τ_h+jet seed (and its backup seed) in the L1 menu, their corresponding HLT paths were also included in the HLT menu. Fig. 6.12 shows the rate distribution for the HLT double-τ_h+jet path obtained from the same collision data sample used for the L1 seed rate estimation. In this case, the rate scaling with respect to PU is not as linear as for the L1 seed. However, for PU 53 we obtain a rate of around 15 Hz, in agreement with our expectation using 2018 data (after rescaling it by the number of bunches).

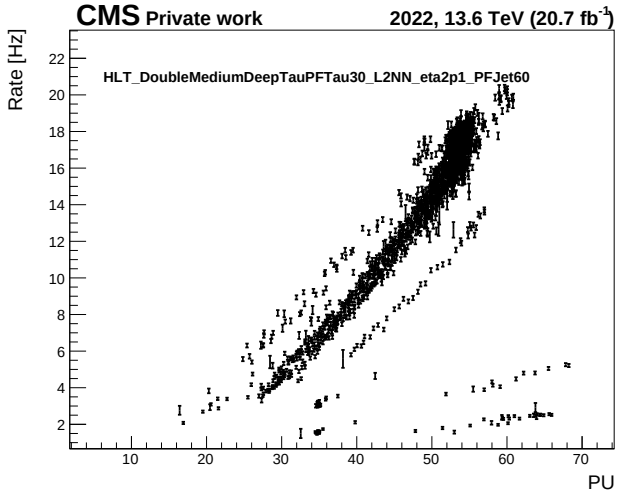


Figure 6.12: `HLT_DoubleMediumDeepTauPFTau30_L2NN_eta2p1_PFJet60` path rate as a function of the PU for a Run 3 collision data sample collected during 2022. Outliers come from runs taken with smaller number of bunches.

Total trigger efficiency

In order for a trigger path to be used in the different analyses, the total trigger efficiency (i.e. the efficiency of firing both the HLT path and its corresponding L1 seed) has to be computed for both data and simulated events, so the latter can be reweighed by applying adequate scale factors. The procedure to obtain this efficiency is similar to the one shown for the L1 seeds: assuming the τ_h and jet objects are uncorrelated, the total efficiency can be computed as the product of the efficiency of selecting the leading τ, the subleading τ and the jet.

The τ_h trigger efficiency can be measured through a tag-and-probe technique, as was done to compute the L1 efficiency, by selecting events that were triggered by a single-muon HLT path. Out of these events, the efficiency will be measured from the ones that also passed a HLT cross-trigger that requires a muon (used as tag) and a τ_h (used as probe). This path was implemented so that the selections applied on the muon are the same as in the single-muon path and the ones applied on the τ_h are the same requirements than in the double-τ_h+jet path. Fig. 6.13 shows the

efficiency as a function of the τ_h p_T and η for real data collected during Run 3 and simulated events from the previous Z + 2 jets sample. Some differences can be seen between both samples, but will be corrected by including an scale factor for the simulated events.

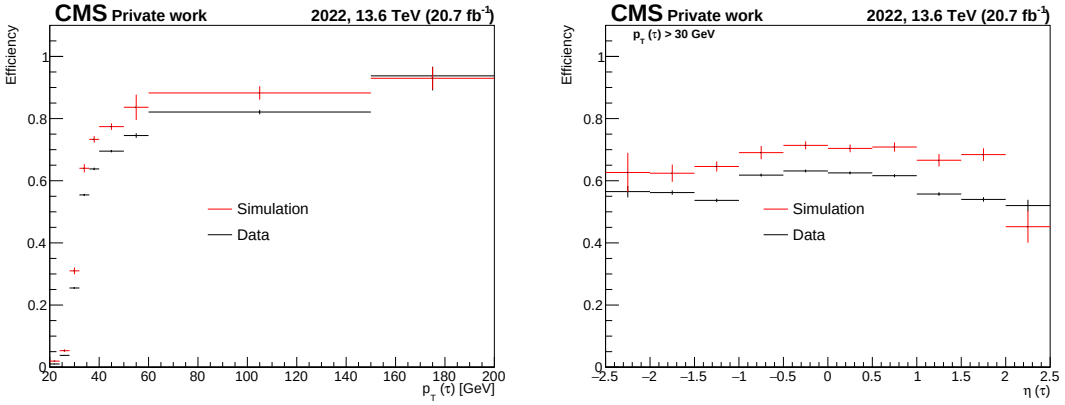


Figure 6.13: HLT τ_h trigger efficiency with respect to the offline τ_h p_T (left) and η (right) for data events collected during 2022 (black) and simulated events from a Z + 2 jets sample (red). Simulated events are reweighed so their PU distribution matches the one from the data (see Fig. 6.10). Generator-level weights are applied to the simulated events. Note that a τ $p_T > 30$ GeV cut is imposed in the efficiency with respect to τ η to match the τ p_T threshold at trigger level.

In a similar way, the jet trigger efficiency is measured by selecting events with one muon, at least one τ_h and at least one jet that were triggered by the muon-τ_h cross-trigger. Out of these events, the efficiency will be computed from the ones that were triggered by an HLT muon-τ_h-jet cross trigger, where the requirements on the trigger muon and τ_h are the same as in the muon-τ_h cross trigger and the requirements on the jet are the same as in the double-τ_h+jet trigger. Fig. 6.14 shows the efficiency as a function of the jet p_T and η for real data collected during Run 3 and simulated events from the previous Z + 2 jets sample. The agreement between data and simulated events is even better than for the τ_h, although some disagreement is appreciated mostly in the energy range where the turn-on is produced.

Expected sensitivity with Run 3 simulation

A different test that can be performed, more close to analysis level, is to try to estimate the sensitivity gain when considering both the double-τ_h and double-τ_h+jet triggers instead of only considering the double-τ_h trigger, as done in the Run 2 HH $\rightarrow$ bb$\tau\tau$ analysis. Four different samples generated for Run 3 analysis are considered, two signals (SM HH $\rightarrow$ bb$\tau\tau$ produced via ggF and VBF) and two of the main backgrounds (inclusive $t\bar{t}$ and the previous Z + 2 jets sample). Over these samples, we apply the selection described in Sec. 4.3.1 in order to select the two τ coming from one of the Higgs bosons in the HH $\rightarrow$ bb$\tau\tau$ analysis. In the $\tau_h\tau_h$ channel,

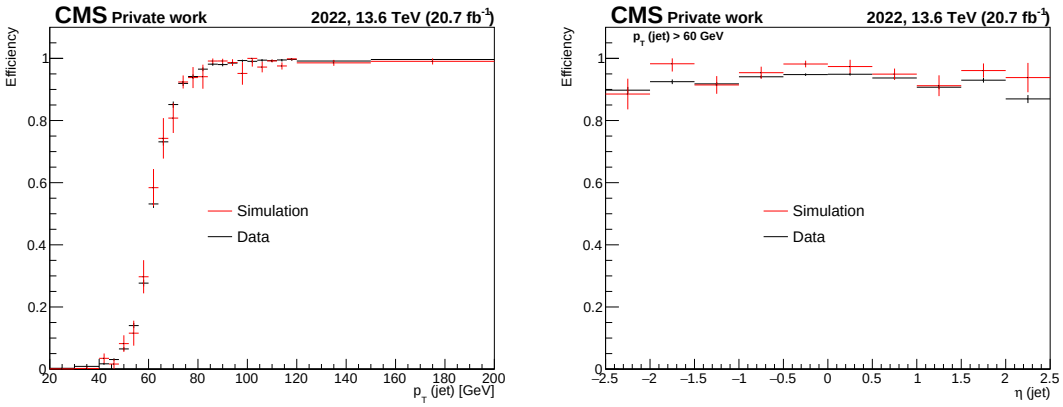


Figure 6.14: HLT jet trigger efficiency with respect to the offline jet p_T (left) and η (right) for data events collected during 2022 (black) and simulated events from a Z + 2 jets sample (red). Simulated events are reweighed so their PU distribution matches the one from the data (see Fig. 6.10). Generator-level weights are applied to the simulated events. Note that a jet $p_T > 60$ GeV cut is imposed in the efficiency with respect to jet η to match the jet p_T threshold at trigger level.

two trigger strategies will be considered. The first one consists of selecting only the events that fired the double-τ_h trigger with a 35 GeV p_T threshold. This trigger was updated before Run 3, so it includes the new features also added to the double-τ_h+jet trigger. In this case, the selected τ_h are required to have a $p_T > 40$ GeV, as done in the HH $\rightarrow$ bb$\tau\tau$ analysis. On the other hand, the second strategy consists of selecting events triggered by the double-τ_h or the new double-τ_h+jet triggers. As the double-τ_h+jet trigger requires two τ_h with $p_T > 30$ GeV, the offline requirement can be lowered to 35 GeV. This way, more events could satisfy the lepton pair selection and be used for the signal extraction. Additionally, for events trigged by the double-τ_h+jet path, an offline jet with $p_T > 65$ GeV is required.

Table 6.3 shows the event yields for the four samples and the two trigger strategies for a given luminosity of 38 fb^{-1} (approximately the collected by CMS during 2022). An increase in the amount of signal events collected with the two triggers of 15% and 19% can be seen for the HH ggF and HH VBF samples, respectively. With the obtained event yields for both signals and backgrounds, the relative increase of significance can be computed as

$$\Delta S = \frac{\left(\frac{N(S)}{\sqrt{\sum_i N(B_i)}}\right)_{2\tau_h \text{ OR } 2\tau_h+j}}{\left(\frac{N(S)}{\sqrt{\sum_i N(B_i)}}\right)_{2\tau_h}}, \tag{6.4}$$

where $N(S)$ is the signal event yield, $N(B_i)$ is the event yield for a background i, and OR is the logical function OR between the two triggers. The relative increase of significance for the ggF sample takes a value of 6%, while for the VBF sample, a 9% value is obtained.

	HH ggF	HH VBF	$t\bar{t}$	Z + 2 jets
Double-τ_h trigger	43.03	1.57	23440	94499
Double-τ_h OR double-τ_h+jets triggers	49.25	1.84	28530	108501
Double-τ_h OR double-τ_h+jets / Double-τ_h	1.14	1.18	1.22	1.15

Table 6.3: Expected event yields in the $\tau_h\tau_h$ channel after the HH $\rightarrow$ bb$\tau\tau$ $\tau\tau$ pair selection for HH ggF, HH VBF, $t\bar{t}$ and Z + 2 jets for two different trigger strategies, considering only the double-τ_h trigger or both the double-τ_h and the double-τ_h+jet triggers. The third row of the table is the event ratio between the two trigger strategies. An integrated luminosity of 38 fb^{-1} (approximately the collected during 2022) has been considered.

6.5 Summary

In view of the new LHC data taking period (Run 3, 2022-2025), a new trigger strategy for the H $\rightarrow \tau\tau$ and HH $\rightarrow$ bb$\tau\tau$ analyses has been studied, focusing on the fully-hadronic τ pair decay mode ($\tau_h\tau_h$). This new strategy consists of including in the L1 and HLT menus a trigger that requires two τ_h and one jet, where the p_T threshold on the τ_h is smaller than the one present in the double-τ_h trigger, already available in the Run 2 menus. By lowering this p_T threshold, more τ_h would be available for triggering and the acceptance on the H $\rightarrow \tau\tau$ and HH $\rightarrow$ bb$\tau\tau$ analyses would increase.

Regarding the L1 seed, a complete process was performed in order to obtain the optimal p_T thresholds to be used for both τ_h and jet in the seed. This optimization was based on having a compromise between the acceptance gained and the rate added. The optimization process lead to adding two new double-τ_h+jet seeds to the L1 menu: a main seed, with a τ_h p_T threshold of 26 GeV and a jet p_T threshold of 55 GeV, and its backup seed, with p_T thresholds of 26 GeV and 70 GeV for the τ_h and jet, respectively. By considering the main seed, an estimation of the acceptance increases in the H $\rightarrow \tau\tau$ and HH $\rightarrow$ bb$\tau\tau$ obtained values of 44% and 36%, respectively, while adding around 8 kHz total rate on a Run 2 sample adapted to the Run 3 data taking conditions.

The new L1 seeds were included in the L1 menu before the start of the LHC Run 3 (summer 2022), so they can be used in physics analyses. The rate and efficiency of the main L1 seed have been studied using 2022 collision data. The rate values obtained agree with the 2018 extrapolations. The trigger efficiency for both τ_h and jet objects has also been obtained with tag and probe techniques for both 2022 collision data and a Z + 2 jets simulated sample, achieving a good agreement between the two samples.

Once the L1 seed was implemented, its corresponding HLT path was also developed based on the existing double-τ_h path, already present in the Run 3 menu. This path includes new features developed for Run 3, such as a CNN for Level-2 τ filtering or the DeepTau algorithm for Particle Flow τ identification at trigger level. The rate of the Double-τ_h+jet path was estimated using 2018 data with Run-

3 conditions, obtaining around 16 Hz. The rate computed with 2022 collision data agrees with the expectations. The efficiency of both τ_h and jet objects has also been computed with 2022 collision data and the Z + 2 jets simulated sample. A very good agreement is achieved between the two samples.

Finally, a first test has been performed to compute an estimation of the sensitivity gain in the HH $\rightarrow$ bb$\tau\tau$ analysis when considering the new trigger strategy: instead of just considering the double-τ_h trigger (as was done in the Run 2 analysis), computing the logical OR of this trigger and the new double-τ_h+jet trigger. When applying the $\tau\tau$ pair selection (both at trigger and offline level) and a minimal jet requirement if the event satisfied the double-τ_h+jet trigger selection, 6% and 9% sensitivity gains were obtained for the HH ggF and HH VBF samples, respectively.

Part II

Studies of a new trigger algorithm for the muon chambers in HL-LHC

Chapter 7

The Analytical Method

For Phase-2, present DT on-detector electronics will be replaced by new boards, the so-called OBDT, which will perform the time digitisation of the chamber signals inside radiation-tolerant FPGAs with a time bin of 0.78 ns (25/32 ns) and a dead time of 100 ns as a result of the minimum DT front-end board deadtime. This time resolution is the same that is being used at present in the DT readout system and therefore, in the offline reconstruction. However, the legacy trigger system needed to work with lower resolution sampling of 12.5 ns and dead times of 400 ns. During Phase-2, all the incoming chamber signals will be sent asynchronously streamed via high-speed optical links to the back-end system, outside the experimental cavern, where the trigger primitives are generated in commercial FPGAs.

In order to implement a new algorithm for trigger primitive generation during HL-LHC, several concerns exist and need to be addressed.

- When a new LHC orbit starts, during the first bunch crossing (BX0), all counters used as a reference for the time of the hits in the DT cells are reset. Then, a hit timestamp (t_{TDC}) can be split in several contributions:

$$t_{\mathrm{TDC}} = t_0 + t_{\mathrm{TOF}} + t_e + t_d, \tag{7.1}$$

  where t_0 is the time difference between the BX0 and the BX of the collision that produced the muon, t_{TOF} is the time of flight of the muon, t_e is the sum of all processing times coming from the electronics involved, and t_d is the drift time from the ionized electrons, which depends on where the muon crossed the DT cell. t_{TOF} and t_e can be determined per cell by calibrating the detector, although they introduce an uncertainty on the value of t_{TDC}. Both t_d and t_{TDC} need to be determined when generating the trigger primitives, taking into account a collision is produced every 25 ns and the maximum drift time is ~ 400 ns, so a signal could be produced by the sixteen previous collisions.

- The timing information from the hits is not intrinsic to the time of the arrival of the hit timestamps, as it was in the Phase-1 system. Therefore, the new algorithm must be capable of sorting the hits and cope with a longer buffering time to ensure that all the relevant information is available before generating the trigger primitive.

- Since the dead time is reduced from 400 ns to approximately 100 ns, the number of hits in the same cell that could potentially belong to the same muon can increase by a factor 4, increasing the possible combinations of hits from different layers exponentially.

- The ageing of the chamber due to accumulated charge, as a consequence of the increased radiation, will reduce the efficiency of each DT cell. Accordingly, the DT trigger has to allow low quality combinations to account for this drop in efficiency in order to guarantee a fully efficient muon measurement.

- The increased PU during HL-LHC leads to a higher activity in the detector, increasing the hit rate in the DT chambers.

In summary, an algorithm with the capabilities of the offline reconstruction algorithm but in real time and handling hits with unknown arrival times had to be implemented. It is to be noted that the actual algorithm, without the HL-LHC added requirements, is already quite complex since it requires to perform a fit of the hits positions with the added unknown of the hits laterality (left or right to the wire) and the hits arrival time.

The *Analytical Method* (AM) [116] is the algorithm developed for performing muon trigger primitive (TP) generation for Phase-2 in the CMS barrel by using information from the DT system. It profits from the improved time digitization at trigger level in Phase-2 DT input hits, 25/32 ns ($\equiv 1$ *TDC count*), much finer than the 25 ns needed for bunch crossing (BX) identification. The input information to the AM algorithm is the wire numbers of the hit cells and the corresponding hit times with respect to the start of the LHC orbit. From these, using the known value of the drift velocity, and assuming a given laterality hypothesis (whether the signal was produced right or left of the cell wire), the hit position is reconstructed. For a given hypothesis of muon trajectory within a given chamber, by assuming it is a straight segment, the collision time, local direction and local position can be analytically determined, with a resolution comparable to what is reachable with the present offline reconstruction software [65].

A brief description of the algorithm and its performance results were recently published in Ref. [116]. In this chapter, the latest version of the algorithm will be presented. Section 7.1 includes a detailed description of the algorithm. Section 7.2 shows its performance results using simulated events, while Section 7.3 and 7.4 describe studies performed with the software emulator and the firmware implementation in the laboratory and in the CMS DT Slice Test.

I have been involved in the development of the new trigger algorithm during the whole length of my thesis. I started as one of the developers of the first version of the algorithm, being the responsible of all its performance tests (both in simulation and real data and its comparison with the firmware implementation) and its integration inside the CMSSW framework. Regarding the version of the algorithm described in this thesis, I am the only developer of the software emulator, and contributed to the firmware implementation with the laterality prediction module. I have also done all

the performance and firmware-emulator comparison tests described in the following sections.

7.1 Description of the algorithm

The AM algorithm has been implemented in firmware and in software as an emulator, being the latter as similar as possible to the former. This section describes the algorithm as included in both implementations. The algorithm can be logically separated in different parts, as shown in Fig. 7.1. The inputs to it are only the time and cell number of all signals collected in each SL. In the first part of the algorithm, both ϕ SLs are treated independently, starting by combining the available hits into groups and then fitting into SL TPs. Finally, a correlation between the TPs coming from both SLs is attempted.

Each logical step is further explained in the following.

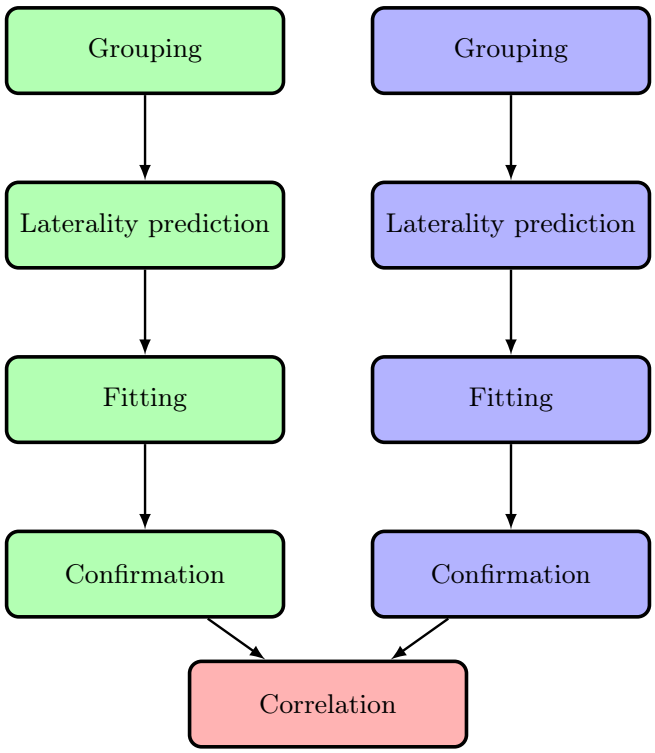


Figure 7.1: Sketch of the structure followed by the AM algorithm. The same four steps independently are followed for each ϕ SL, before a correlation between SLs is attempted.

Grouping

In this step, the aim is to build hit groups, sets of three or four signals in a given superlayer whose cells are compatible with a muon crossing the chamber in a straight trajectory. The cell distributions compatible with straight trajectories are called *cell layouts*, and are shown in Fig. 7.2. Hit groups are built synchronously whenever a hit is detected and introduced as input to the algorithm.

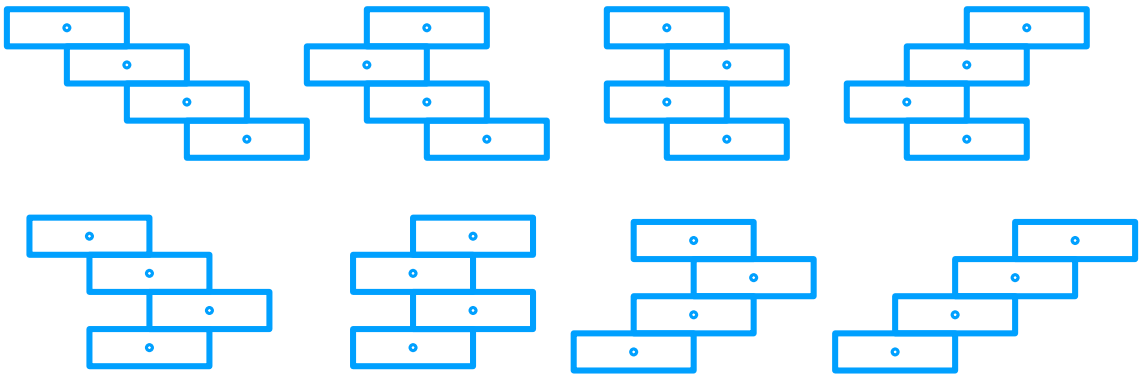


Figure 7.2: Sketch of the cell layouts compatible with straight muon trajectories in a DT superlayer.

Whenever a new hit is detected and introduced as input to the algorithm, it gets stored and compared with all the hits that were previously stored in a window of 400 ns and that could be used to build a hit group compatible with one of the eight available cell layouts. These is done by looking at ten cells at once, shown in Fig. 7.3, which depend on the layer of the new hit. Then, a loop is performed to build the hits compatible with the cell layouts of Fig. 7.2 from left to right. Each cell layout will be considered only if at least three hits in different layers are available. If more than one hit per layer is available, all the possible combinations are built.

Laterality prediction

As it was shown in Section 2.2.2, there is a left-right (laterality) ambiguity problem on the hit position. This step in the algorithm aims to provide all posible laterality combinations feasible for each hit group.

In the firmware implementation, the architecture is so that each output laterality combination is treated by a dedicated fitting module, i.e. if a hit group could allow at most N laterality combinations, N fitters would be instantiated. Since the amount of FPGA resources needed per fitter is very substantial, reducing the maximum number of allowed lateralities per group is desired.

Due to geometrical constraints, at most four laterality combinations are available per hit group. However, this number can be reduced by also considering the timing information from each hit in the group. Nonetheless, this reduction could potentially decrease the algorithm efficiency, so it needs to be studied. The strategy to follow in

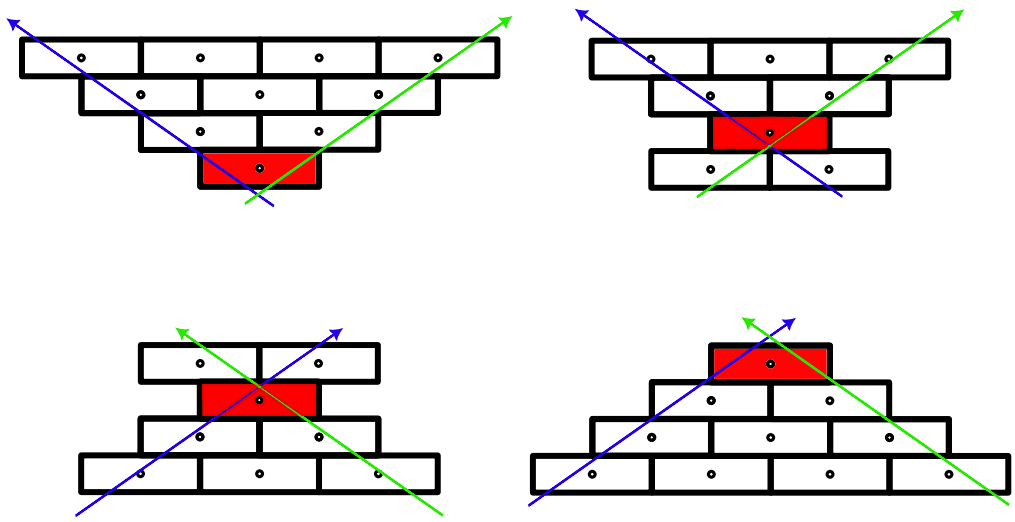


Figure 7.3: Structures of 10 cells where the combinations of three or four hits are looked for. The cell shown in red indicates the one that detected the hit under study. Blue arrows indicate the first combinations of cells that are studied, while the green ones are the last ones.

this laterality prediction step is, instead of considering the full timing information (which will be done later in the fitting), obtaining for each hit its *coarsified time difference*. For each hit, the process to follow is:

1. Obtain the time difference with respect to the hit in the group with the highest time value (i.e. the ones that was measured the latest).
2. Out of the time difference, take only the most significant part. To do so, the time difference is translated to binary code, the resulting code is cut to its N most significant bits (MSB) and then translated back to decimal code, obtaining its *reduced time difference*.
3. The reduced time difference is then compared to some fixed thresholds to obtain a number from zero to three: reduced time differences between 0 and the first threshold are assigned a value of 0; between the first threshold and the second, a 1, and so on.

With the three or four input hits, a *coarsed set* of three or four elements, respectively, is produced. The combination of this set and the cell layout is associated via a previous classification to up to three laterality combinations. An example of the coarsification process is shown in Table 7.1.

To obtain the thresholds needed for the coarsification, a dataset of straight tracks (simulating a muon crossing a DT superlayer) is produced. These tracks follow the equation

$$x = x_0 + \text{slope} \cdot y, \tag{7.2}$$

	Layer 0	Layer 1	Layer 2	Layer 3
Input times (ns)	541	313	531	560
Input times (TDC counts)	692	400	679	716
Time difference w.r.t. latest hit (abs. value)	24	316	37	0
Reduced time differences to their 6 MSB	3	39	4	0
Coarsified time differences	1	2	1	0

Table 7.1: Example of the time coarsification process. The coarsification has been performed using as thresholds 1, 31 and 40 (i.e. reduced time differences between 0 and 1 are assigned a value of 0; between 1 and 31, a 1; between 31 and 40, a 2; and between 40 and $2^6 - 1$ (the largest value that can be obtained using 6 bits), a 3).

where x is the position in the horizontal axis and y the position in the vertical axis. x_0 and slope take values between -21 mm and 21 mm and between -1.6 and 1.6, respectively (with granularities of 0.2 mm and 0.005). Each track gives a total of four hits, one for each layer, and from the position of each hit inside the cell both the drift time and the laterality can be extracted. To account for the detector resolution, each drift time is smeared by adding a random quantity sampled from a normal distribution with zero mean and $\sigma = 10$ ns.

Once the dataset is built, a dedicated algorithm is used in order to find the optimal thresholds for the coarsification. The algorithm tries to minimize number of coarsed sets for each cell layout that are associated to four or more laterality combinations by applying Bayesian optimization techniques [117, 118]. Two sets of thresholds are then selected, one for four-hit primitives and another one for three-hit primitives. With these sets of thresholds, if at most three laterality combinations are associated to each coarsed set, 99.98% of the four-hit primitives find their true laterality combination among the three allowed. For three-hit primitives, all of them find their true laterality combination. Thus, almost no efficiency is lost due to adding the laterality prediction module, and it fulfils the objective of reducing the maximum number of allowed lateralities per group from four to three. Moreover, the mean number of possible lateralities per group drops from 2.78 to 2.16. Given this reduction, another architecture change is being studied, where only two fitting modules are instantiated and hit groups with their corresponding lateralities are input to the fitters in a queue approach. This new approach would reduce even more the FPGA occupancy with little change (if any) in performance.

Fitting

After building the hit patterns and for each laterality combination obtained in the previous steps, the muon primitive's time t_0, the bunch crossing (BX) of the proton-proton interaction that produced the muon, the segment local position and the local direction with respect to the chamber perpendicular axis $\tan\psi$ are computed in a given superlayer using formulas extracted with the least squares minimization method. The χ^2 of the fit is also computed.

In general, 4-hit fits will provide much better results than 3-hit fits, as one additional degree of freedom is available. In fact, in the latter the χ^2 of the fit is meaningless, as three parameters are fitted with three degrees of freedom. Not only the resolution achieved with these primitives will be worse, they will also generate more TPs out of PU, background or electronic noise.

Confirmation

Previous to the correlation between SLs, another step aimed to select the best single-SL TPs is performed. Each TP is *confirmed* if, after extrapolating it to the other SL, the extrapolation matches to at least two hits. If this is the case, the segment gets tagged as a *confirmed* segment candidate.

Once this confirmation is done, a quality code is assigned to each SL TP, as shown in Table 7.2. By comparing this quality code between TPs, a filtering is done as follows:

- If two primitives have at least one hit in common, only the primitive with the highest quality is kept. In this way, a confirmed 4-hit TP prevails over a non-confirmed 4-hit, which prevails over confirmed and non-confirmed 3-hit TPs.
- If two same-quality TPs have at least one hit in common, the behaviour depends on the quality considered. For both 4-hit and 4+2 hits TPs, the TP with the smallest χ^2 obtained in the fitting prevails. For 3-hit and 3+2 hits TPs, as the χ^2 is meaningless, both TPs are kept.
- If no hit is shared between two TPs, both are kept, regardless of the quality of each of them.

Quality	Description	Type
1	3-hit TP	Uncorrelated
2	3+2 hits TP	Confirmed
3	4-hit TP	Uncorrelated
4	4+2 hits TP	Confirmed

Table 7.2: SL quality descriptions.

Correlation

If at least one fit was obtained in each of the two r-ϕ SLs, a combination of those fits can be performed if their corresponding times lay in a window of ± 25 ns and their position and slope are relatively similar. In that case, a fitting is performed again using the six to eight hits used in the SL-fits, obtaining new values for the t_0, BX, local position and local direction. The corresponding SL input segments are

then discarded. If no match is found or the fit was invalid, all SL candidates are then kept at this stage.

Depending on the number of hits used for the correlation, the quality code is updated with the new numbers described in Table 7.3.

Quality	Description	Type
6	3+3 hits TP	Correlated
7	4+3 hits TP	Correlated
8	4+4 hits TP	Correlated

Table 7.3: Chamber TP quality descriptions.

A final filter is performed at this stage similarly to the per-SL filter:

- If two TPs match (within given position, local direction and time ranges) and they have different quality, only the one with the highest quality remains. If they have the same quality, the one with the smallest re-fitted χ^2 is kept.
- If they do not match, both TPs prevail.

7.2 Algorithm performance

In order to estimate the AM performance, several sets of simulated and real data samples have been used. For the results presented in this chapter, we will consider a 3700 simulated event sample with four prompt muon pairs per event. Each pair consists of 2 back-to-back generated muons with flat p_T, ϕ and η distributions with a p_T between 2 and 200 GeV and within $|\eta| < 1.2$. Overimposed to the signal, additional proton-proton interactions (PU) are generated within a window of ± 16 BX around the central BX (where the signal muons are generated), fully covering the maximum drift time of $\sim$390 ns, with an average PU of 200 events per bunch crossing. Additionally, the GEANT4 simulation configuration takes into account backgrounds from long-lived particles originating from collisions, in particular low-energy neutrons, that can produce hits in the DT chambers evenly over the LHC orbit. The simulation assumes a perfect inter-chamber calibration.

During Phase-2, DT chambers will be exposed to high radiation levels, specially in some regions of the detector, potentially producing aging effects degrading the DT cell performance and lowering the DT chamber efficiency. Several DT ageing scenarios can be simulated by removing in a random way DT hits, according to predefined probabilities. In our results, a scenario equivalent to 3000 fb^{-1} has been considered, corresponding to extreme ageing effects in the DT detector at the end of Phase-2. This conservative scenario has been based on measurements of the DT chamber performance under high radiation conducted in the new CERN Gamma Irradiation Facility. Theses hit efficiencies have been estimating considering a safety factor of 2 from the expected HL-LHC instantaneous luminosity ($2\times5 \times 10^{34}$ cm^{-2}s^{-1}) and the same safety factor for the expected integrated luminosity ($2\times$3000 fb^{-1}). In this

scenario, the lowest DT chamber efficiencies are in the order of 70% in MB1 of the most external barrel wheels (Wh ±2), raising to 90% in some sectors from the MB4 and remaining significantly higher for the rest of the DT chambers, as can be seen in Fig. 7.4. Despite these values, thanks to the redundancy of the system and to the mitigation measures currently implemented, a good performance of the muon triggering and reconstruction in still expected during Phase-2.

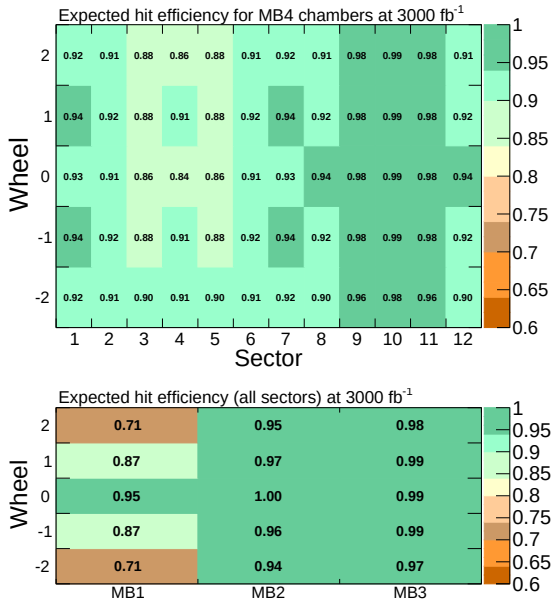


Figure 7.4: Expected hit efficiencies at the end of Phase-2 for all the DT chambers of the CMS muon system. The upper plot shows MB4 chambers, the lower shows MB1, MB2 and MB3. Extracted from [116].

7.2.1 Efficiencies to trigger on prompt muons

In order to verify the performance of the algorithm, it is important to measure certain metrics such as the efficiency to reconstruct a segment out of the available hits. To obtain these efficiencies, the denominator is obtained as the number of DT offline segments with at least 4 hits in the r-ϕ view and also 4 hits in the r-z view (if available) geometrically matched with a generated muon within a window of 0.15 in η and 0.1 rad in ϕ. In order to remove segments coming from PU events, a cut on the reconstructed segment time of ±15 ns is applied. The numerator of the efficiency is defined as the number of trigger primitives whose BX is the one from the collision and matching these segments within a ϕ window of 0.1 rad.

Fig. 7.5 summarizes the DT Phase-2 TP efficiency per station and wheel for two ageing scenarios: with no ageing applied and with the ageing scenario corresponding to 3000 fb^{-1}. In both cases, 4 different TP quality thresholds are considered. When no ageing is considered, TP efficiencies are higher than 98% for a quality threshold of 1 or 2, while the efficiency for correlated TPs is above 80% in the whole detec-

tor. When the 3000 fb^{-1} ageing scenario is applied, the efficiency drop is larger in the chambers more affected by these ageing effects (as expected), overall for high TP quality thresholds. When lower qualities are considered, the efficiency can be substantially recovered.

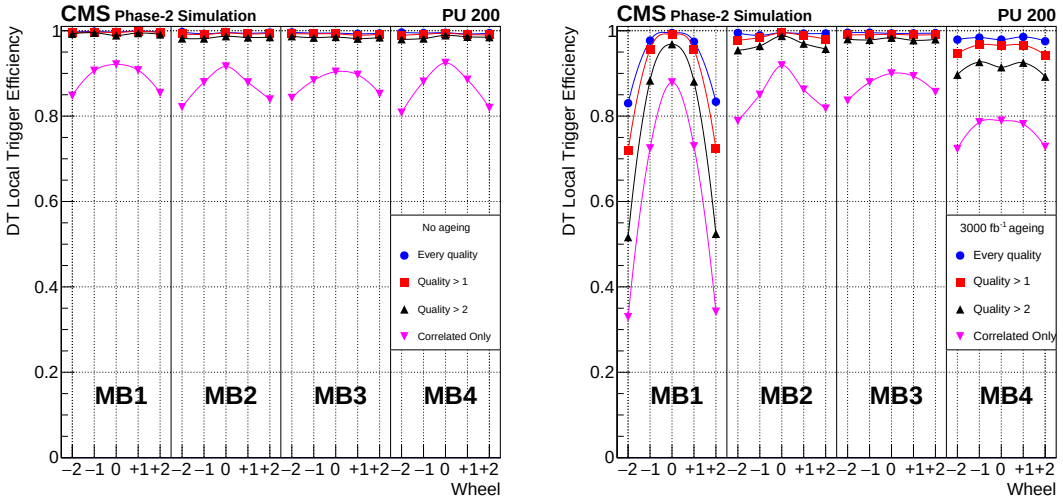


Figure 7.5: TP efficiency with respect to segments reconstructed with the offline system considering (left) no ageing scenario or (right) 3000 fb^{-1} end of HL-LHC ageing scenario, for four different quality groups: every quality, shown in blue; quality > 1 (i.e. confirmed three-hit TPs or better), in red; quality > 2 (i.e. four-hit TPs or better), in black; and correlated TPs, in magenta.

7.2.2 Comparison with offline segments

A further evaluation of the performance of the algorithm and the quality of the TP fit results can be done by comparing the values obtained from this fit and the corresponding values of offline reconstructed segments. For this purpose, only the offline segments that satisfy the same matching and quality criteria used in the efficiency studies are considered. In the following results, ageing effects have been applied to the hits before generating the TPs.

Regarding the performance of 3- and 4-hit fits, Fig. 7.6 shows a comparison between the time distribution of TPs obtained from fits with 3 or 4 hits only. Note how not only the standard deviation of the core gaussian distribution is bigger in the 3-hit case, but how a flat component is much more considerable.

Fig. 7.7 shows the time distribution from all TPs and only from correlated TPs. By comparing both distributions, it is to be noted how the core of the all-TP distribution is dominated by correlated TPs, which at the end drive most of the sensitivity. The σ of the Gaussian fit in the correlated-TP distribution is 3.2 TDC counts, equivalent to 2.5 ns. Note how the precision achieved at trigger level, even with the end

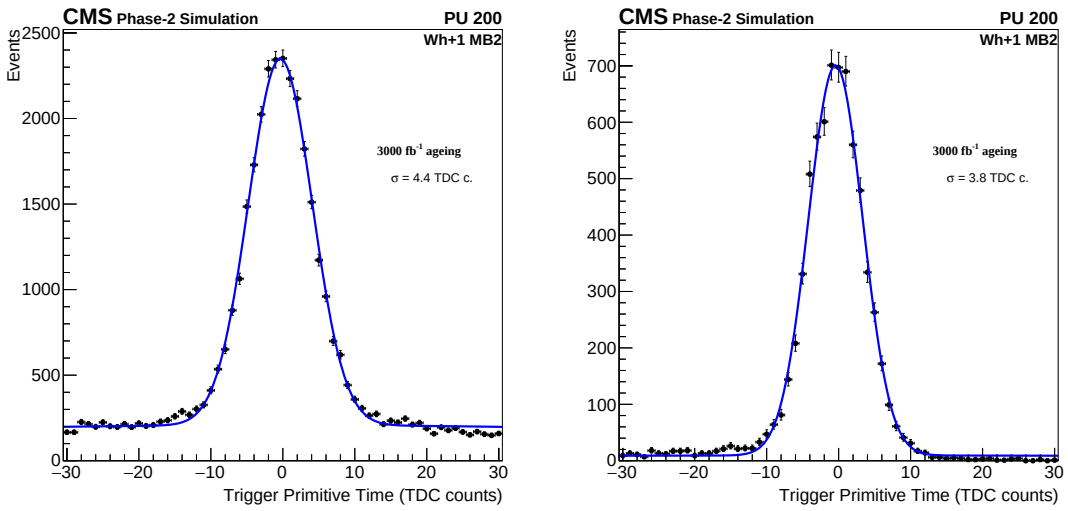


Figure 7.6: TP time distribution for (left) 3-hit and (right) 4-hit TPs in Wh+1 MB2. End of HL-LHC ageing is applied to hits before TP generation. Both distributions have been obtained by deactivating the correlation between SLs, so most primitives included in them won't be in the final algorithm's output.

of HL-LHC ageing scenario applied, is only a few ns away from the current offline reconstruction system.

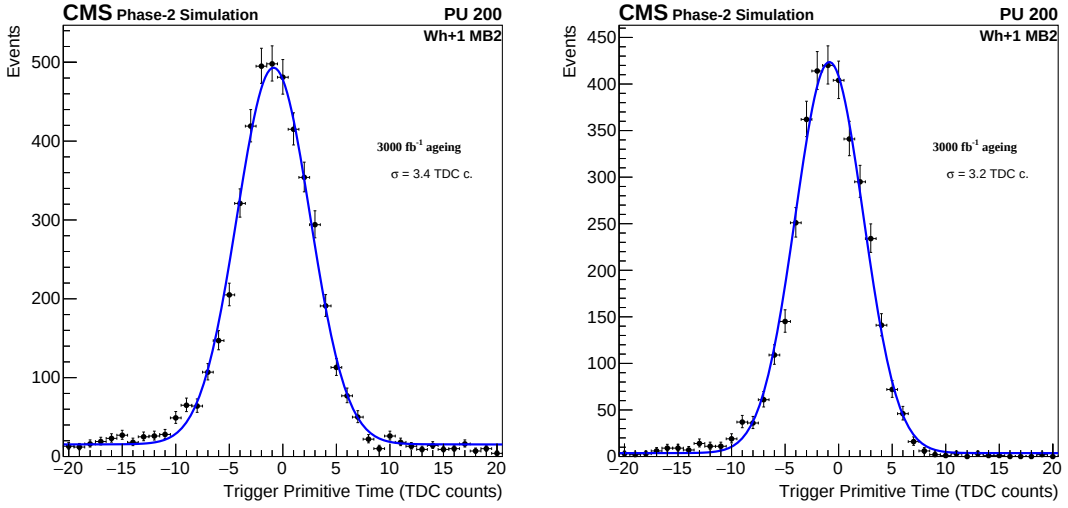


Figure 7.7: TP time distribution for (left) all TPs and (right) correlated TPs in Wh+1 MB2. End of HL-LHC ageing is applied to hits before TP generation. Larger tails at negative values are compatible with the known effect of delta rays and detector deadtime [46].

Fig. 7.8 shows the difference in local position and local direction for correlated

TPs in Wh+1 MB2. The standard deviation (σ) of the Gaussian fit to the local position (local direction) difference distribution is 59.8 μm (0.6 mrad).

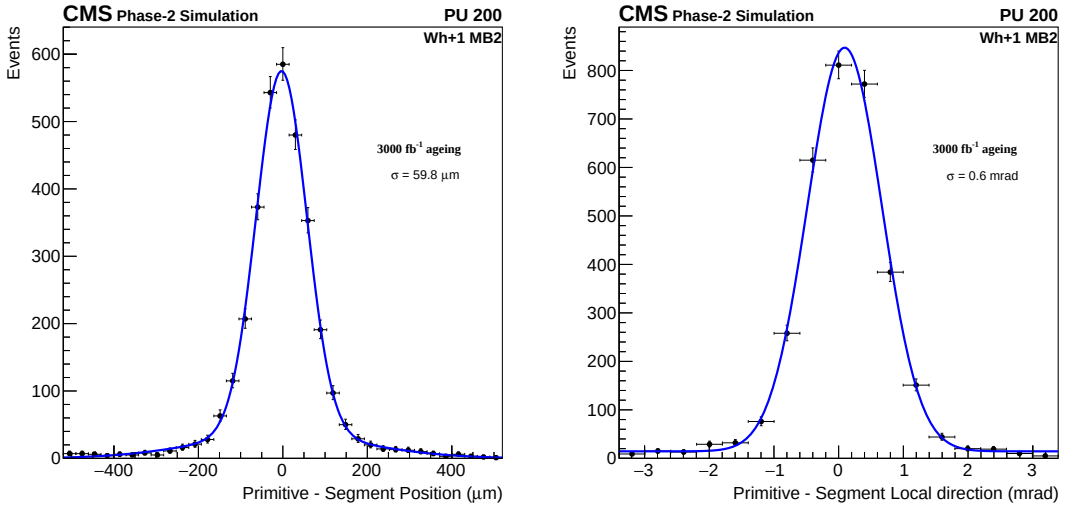


Figure 7.8: Difference with respect to offline reconstructed segments of (left) TP local position and (right) TP local direction, for correlated TPs in Wh+1 MB2. End of HL-LHC ageing is applied to hits before TP generation.

Fig. 7.9 shows the σ of the Gaussian fit to the local direction different distribution for chambers in the same wheel and station. Correlated TPs are shown in blue and uncorrelated TPs in red. The improvement seen in correlated TPs is not only due to the added degrees of freedom but to the effect of the larger level arm between SL1 and SL3 with respect to the measurement in a single superlayer.

It must be stressed that these parameter differences between offline segments and TP can't be used to measure the TP parameter resolution (as measured from generated quantities), since significant correlations between both TPs and segments are to be expected due to sharing of hits. However, dedicated tests were performed using the previous AM algorithm version comparing the trigger primitives directly to the simulated information, as shown in Ref. [116]. The resolutions obtained with respect to the simulated information were very close to the ones with respect to the offline reconstructed segments.

7.2.3 Rate studies

TP rates have been computed for the AM algorithm using the previously described simulated sample, with and without considering ageing end of HL-LHC effects. For each event, only the chambers not crossed by offline reconstructed muons were considered. Without ageing, rates for DT TPs with at least 4 hits are in fair agreement with the ones extrapolated from Phase-1 data, of the order $\sim$ 0.5 MHz at 10×10^{34} cm^{-2}s^{-1} in MB1 external wheels [60]. Rates elsewhere are at least a factor

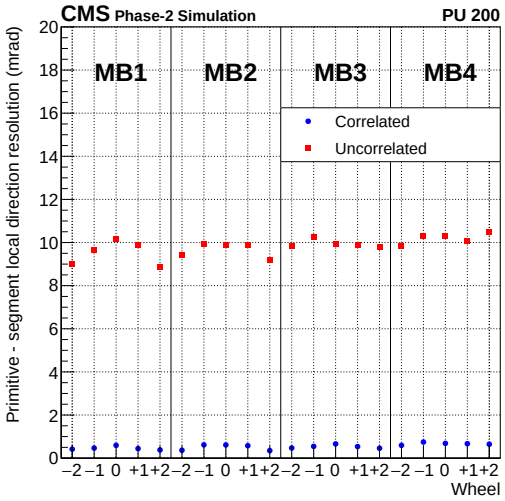


Figure 7.9: σ of the Gaussian fit of the difference in local direction between offline segments and TPs. Correlated TPs are shown in blue and uncorrelated TPs in red. End of HL-LHC ageing is applied to hits before TP generation.

of two smaller, with most chambers having one order of magnitude less rate than the MB1 external wheels. Ageing reduces the trigger primitive rates as expected by the hit efficiency loss. In the MB1 external wheels, the rate is around three times smaller. All these estimated rates fit within the specification for the new Phase-2 L1 trigger system.

7.3 Firmware implementation

The AM trigger algorithm, already designed in a firmware-oriented approach, has been ported to VHDL in order to estimate the FPGA resources required. This implementation has been validated with real data using emulator-to-firmware comparisons in prototype boards.

The present firmware implementation performs the TP generation in the r-ϕ view of the DT chamber and its consistent with the algorithm description included in Section 7.1. However, some small differences are present, mainly due to practical constraints. In particular, confirmed qualities are not implemented in the firmware version for the moment.

The present implementation of the algorithm has been performed in a Xilinx Virtex Ultrascale Plus (VU9P) FPGA. On top of the AM algorithm logic, several functionalities have been included to allow the control and operation of the system.

7.3.1 Firmware-Emulator comparison

In order to study the level of agreement between the current emulator and the firmware implementation, a series of studies have been done at a test stand at CIEMAT, shown in Fig. 7.10. In these tests, the input is given by the DT hits from the previously described simulated sample, coming from all chambers in the CMS detector. These hits are stored in a file and injected at the input buffers of the VCU118 board. These hits already have the Phase-2 data format and are injected at a predefined time, emulating the behaviour of the OBDT. Hits go through all the trigger chain inside the VU9P FPGA, and TPs are generated in the board and are read out. The emulator is also run on the same sample, so results from both emulator and firmware can be compared.

Figure 7.10: Test stands devoted to firmware-emulator comparison tests based on a Xilinx Virtex 7 FPGA (left), used with the previous AM algorithm version, and on a Xilinx Virtex Ultrascale Plus FPGA (right), used in the most recent tests.

The first test tries to quantify how many TPs are produced in both emulator and firmware and what is the level of additional TPs produced in each implementation. To do so, for any given primitive output by the firmware(emulator), a corresponding primitive is searched for in the emulator(firmware) output, sharing the same fitted hits with the same laterality assignments. This *matching efficiency* is quantified in Table 7.4, where the percentage of matched primitive pairs is shown per TP quality. Matching is in general greater than 97% for all qualities (except quality 6, which suffers from a high statistical uncertainty). This matching is much better than the achieved in the Phase-1 system [56], even considering that the sample used in our computation has much harsher conditions. Something to be highlighted is the matching efficiency for quality-8 TPs, which takes a value of 99.99% in both cases. At the end, quality-8 TPs drive the efficiency and the resolution of the algorithm; being able to reproduce them almost perfectly in both implementations is a very big success.

The second test studies the agreement in the parameter values computed by both firmware and emulator. This is done by comparing the obtained parameters between pairs of emulator-firmware TPs that fit the same hits with the same laterality assignment. In Fig. 7.11, the left panel shows the difference between the TP BX

Quality	Matching w.r.t. emulator (%)	Matching w.r.t. firmware (%)
1	98.17 ± 0.05	99.46 ± 0.03
3	99.70 ± 0.03	99.93 ± 0.02
6	94.22 ± 1.19	93.52 ± 1.24
7	97.63 ± 0.21	97.61 ± 0.21
8	99.99 ± 0.01	99.99 ± 0.01
Average	98.76 ± 0.03	99.56 ± 0.02

Table 7.4: Matching efficiency per quality between the firmware and the emulator implementations. The second (third) column shows the percentage of firmware (emulator) TPs that could also be found in the emulator (firmware). Note that, as the confirmation is not included in the firmware implementation, it is also skipped in the emulator version used for these tests, so no quality 2 and 4 TPs are produced.

obtained by the emulator (in blue) or the firmware (in red) and the event BX. As shown in the insert, agreement in time is at the level of least significant bit (1 TDC count). Similarly, the right panel compares the time fitted by the emulator (in blue) and the firmware (dashed red) once subtracted the BX from the event (multiplied by 32 in order to convert the BX number to units of TDC counts). Perfect agreement between both curves can be seen.

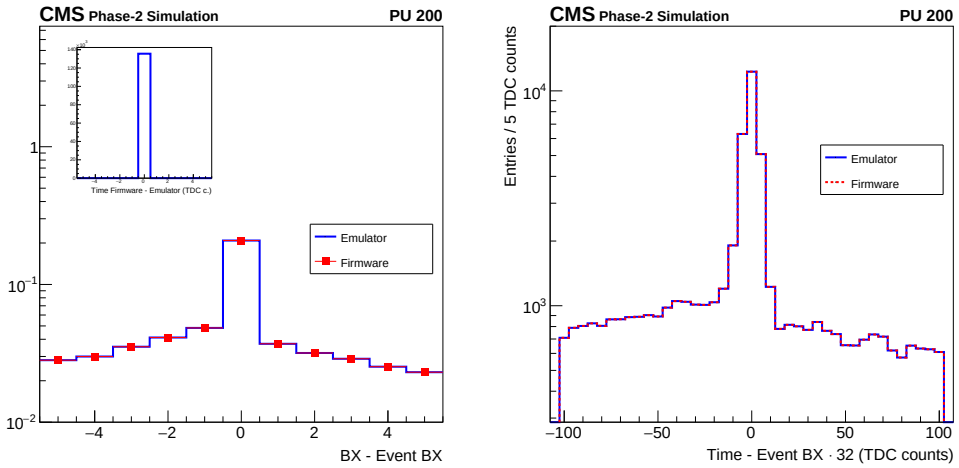


Figure 7.11: Left: Difference in BX assignment between emulator TPs and event BX (blue) and firmware TPs and event BX (red). Insert shoes agreement in the fitted time value at the level of Least Significant Bit (1 ns). Right: Difference between TP time and the event BX · 32, as obtained by the emulator (blue) and the firmware (dashed reds). In both plots, only pairs of primitives fitting the same hits with the same laterality assignment are considered.

In Fig. 7.12, the left panel shows a comparison between the trigger positions obtained by emulator and firmware. The right panel shows the distributions of local direction as obtained by emulator (blue) and the firmware (dashed red). In both

cases, the corresponding inserts show agreement at the level of Least Significant Bit on these variables.

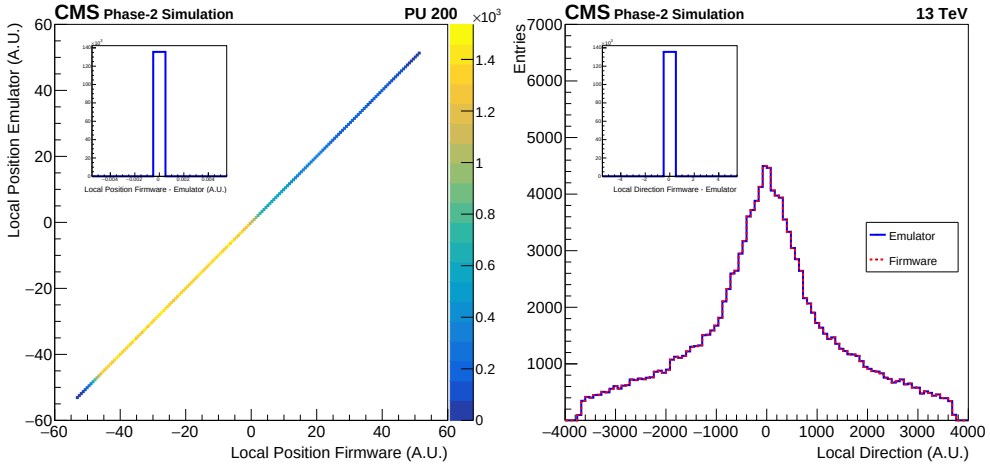


Figure 7.12: Left: TP local position as obtained by the software emulator versus TP local position obtained by the firmware. Right: TP local direction as obtained by the software emulator (blue) and obtained by the firmware (dashed red). In both variables, the insert shows an agreement at the level of Least Significant Bit. In both plots, only pairs of primitives fitting same hits with same laterality assumptions are considered.

7.4 AM performance in the CMS DT Slice Test

During LS2, a complete exercise has been made to instrument one sector (Wheel +2, Sector 12) of the CMS detector with the new Phase-2 DT electronics prototypes to setup a demonstrator of the upgraded system, the *DT Slice Test*. Chamber signals were split into the four DT stations, so they can be readout at the same time with the Phase-1 electronics and the first OBDT board prototypes. In the current setup, each OBDT covers one superlayer, except in the MB4, where two OBDTs per superlayer are required, and the MB3, which is not fully instrumented for the trigger.

In the back-end system, several prototypes of boards based on the TM7 board [119] are used either for Slow Control, the so-called MOCO (Monitoring and Control) board, or for readout and triggering purposes through AB7 (AM algorithm Beta on Virtex 7) boards. Five AB7 boards are used for the whole sector, one for each MB1, MB2 and MB3, and two for the MB4, each of them reading one OBDT. The AB7 board is in charge of performing the trigger primitive generation using the previous version of the Analytical Method algorithm, described in Ref. [116]. The latest firmware version (described in this thesis) is expected to be deployed inside the CMS infrastructure in the upcoming months.

This prototype system was integrated into the DAQ system of CMS, so both

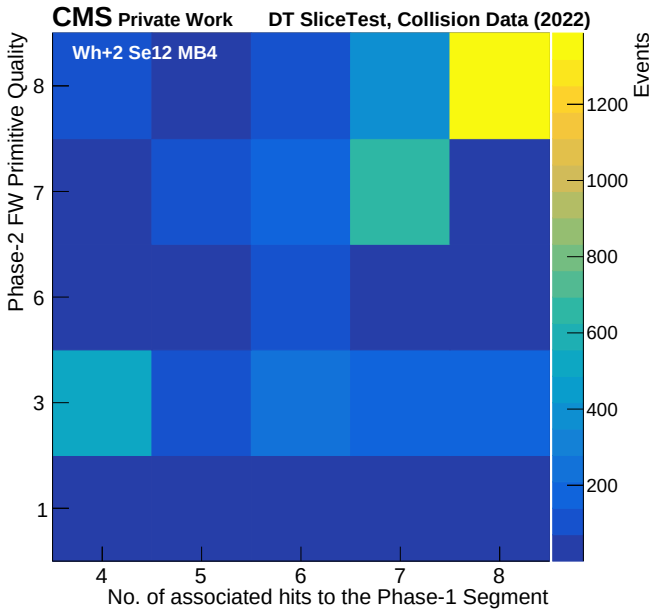


Figure 7.13: 2D distribution of the Phase-2 TP quality obtained by the AM firmware versus the number of hits associated with the offline reconstructed segment in the MB4 station for the DT Slice Test collision data taken during the start of Run 3 in 2022.

cosmic (during LS2) and collision data (during the start of Run 3 in 2022) could be taken. As the information from the chamber signals can be read out by both Phase-1 and Phase-2 systems, a comparison between the Phase-1 trigger and the Analytical Method algorithms could be performed directly on-detector. The following results are obtained using collision data taken during the start of Run 3 in 2022.

Fig. 7.13 shows the 2D distribution of the Phase-2 TP qualities, as obtained by the AM firmware in the MB4, as a function of the number of hits associated with the offline reconstructed segments. Offline segments are required to have at least four hits in the $r - \phi$ view. Only the segment with more associated hits, and the TP with the highest quality is considered. A good correlation between TP qualities and the number of hits in the offline segment is observed.

Fig. 7.14 shows the efficiency of finding a Phase-2 TP in any BX with respect to the local position of the offline reconstructed segment in the MB4. Three quality groups are considered: every quality, quality ≥ 6 (i.e. correlated TPs), and quality ≥ 3 (i.e. TPs built out of four or more hits). The values obtained are in general very high for all quality groups considered. Note that the efficiency drop found in position = 0 cm is due to edge effects appearing as two AB7 generates TPs in this chamber, one for each half. This effect will disappear in the final system as only one board will be used.

Fig. 7.15 shows the time difference between the AM TP and the offline reconstructed segment. As can be seen, the time resolution with respect to segments is

of the order of few ns, while in the current Phase-1 system, the TP output time was given in steps of bunch crossings (25 ns).

7.5 Summary

For Phase-2, the increase in the luminosity delivered by the LHC will require a replacement of the DT on-detector electronics. These new electronics will provide the time digitisation of the chamber signals with a time bin smaller than 1 ns and a dead time of 100 ns already to the DT trigger system. This DT trigger system in Phase-2 will run a new algorithm I helped to develop, the Analytical Method, to perform muon trigger primitive generation. The algorithm is already implemented in firmware and in software as an emulator, and uses as input only the time and cell number of all signals collected in each DT superlayer. These signals are first

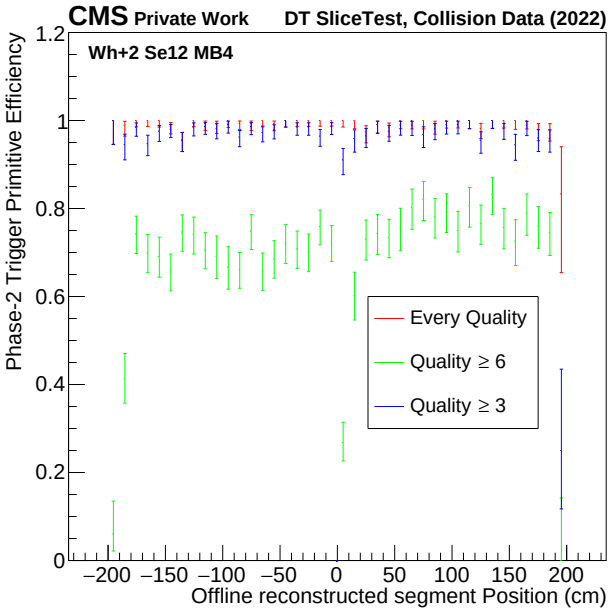


Figure 7.14: Efficiency of finding a Phase-2 TP in any BX with respect to the local position of the offline reconstructed segment considering every primitive (red), primitives built with more than four hits (blue), and primitives with more than six hits, i.e. three or more hits per ϕ superlayer (green) for DT Slice Test collision data taken in the MB4 during the start of Run 3 in 2022. Selected segments are built with at least four hits and have an inclination in the radial coordinate smaller than 30° with respect to the direction perpendicular to the chamber. No geometrical matching between the offline segment and the TP is required. Since one AB7 board only generates primitives from one half of the MB4 chamber, an efficiency drop is visible at the boundary (position = 0 cm) between the two regions due to edge effects. The final system will use only one board for the full chamber, so this effect will disappear.

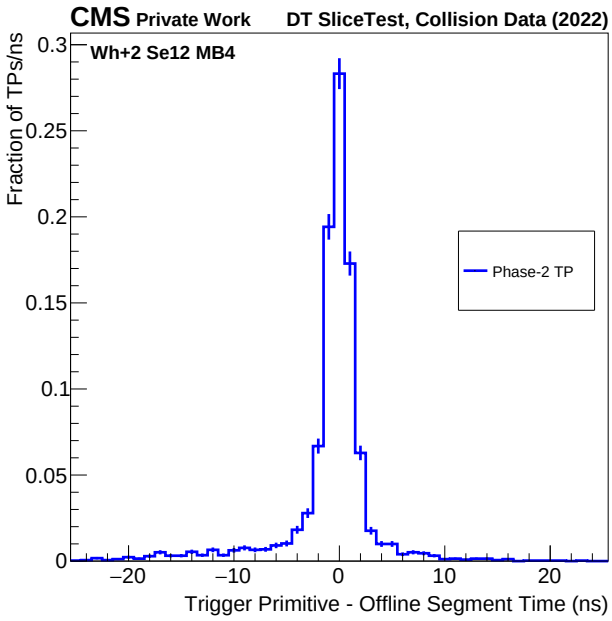


Figure 7.15: Difference between the AM TP fitted time and the offline reconstructed segment time for DT Slice Test collision data in the MB4 station during the start of Run 3 in 2022. Selected segments are built with at least four hits and have an inclination in the radial coordinate smaller than 30° with respect to the direction perpendicular to the chamber. No geometrical matching between the offline segment and the TP is required. Only the TP with the highest quality is considered.

combined into hit groups, which are then fitted into superlayer trigger primitives. Then, the information from both superlayers is combined to build the final trigger primitives. The algorithm is designed to adapt to the FPGA architecture, so its occupancy and performance are improved. An example of this design is shown in the laterality prediction module.

The performance of the algorithm has been studied with a simulated sample assuming the worst expected conditions with respect to the ageing of the DT chambers. The efficiencies to trigger on prompt muons are very high in the whole detector, even when considering only high quality trigger primitives. The resolutions computed with respect to offline segments show a very high precision achieved already at trigger level, with a time resolution only a few ns away from the current offline reconstruction system.

The AM algorithm has been ported to VHDL in order to estimate the FPGA resources required. This implementation has been validated with simulated data using emulator-to-firmware comparisons in prototype boards. These comparisons show a matching between both implementations above 98% when all TP qualities are considered, and almost 100% for very high quality trigger primitives. When performing a fit with the same hits and laterality combinations, the trigger primitive parameters (time, position, and direction) agree between both implementations to

the least significant bit.

During LS2, one sector of the CMS detector with the new Phase-2 electronics, the so-called DT Slice Test. In the back-end system, a prototype board runs the previous version of the Analytical Method algorithm, so it can be tested with cosmic and 2022 collision data. With 2022 collision data, the algorithm provides very good results in efficiency, time resolution, and quality matching with the offline reconstructed system.

Conclusions

This thesis presents a search for Higgs boson pair production via gluon fusion (ggF) and vector boson fusion (VBF) in the channel where one of the Higgs bosons decays into two τ leptons and the other Higgs decays into two b quarks ($HH \to bb\tau\tau$). This process is predicted by the Standard Model (SM), but could be affected by Beyond the Standard Model (BSM) effects. The search is performed using proton-proton collision data at $\sqrt{s} = 13$ TeV collected by the CMS experiment in the LHC during 2016, 2017, and 2018, equivalent to 138 fb^{-1}. The analysed final state is where two b quarks and two τ leptons are produced. The three decay modes of the $\tau\tau$ pair with the largest branching ratio have been studied, requiring one τ to decay into hadrons and the other leptonically or hadronically.

This analysis has been based on the previously published $HH \to bb\tau\tau$ search, performed only with 2016 data. Many improvements have been included since, both in object reconstruction and identification (e.g. DeepJet and DeepTau algorithms for the identification of b jets and τ leptons decaying into hadrons) and at analysis level. A new categorization scheme has been studied focused on the VBF production mode, in order to assign events to categories associated enriched in events belonging to the VBF, ggF, and the main background processes. This categorization is performed using a multi-class deep neural network (DNN), providing more sensitivity to the VBF analysis thanks to the additional discrimination and the ability to constrain nuisance parameters. A complete study of the main backgrounds has been performed, focusing on the QCD background estimation. Several validity tests have been done to study the method used for QCD estimation, with satisfactory results.

Signal extraction is performed with a binned maximum likelihood fit to another DNN that discriminates between signal- and background-like events. Upper limits at 95% confidence level (CL) are set considering as signal HH produced via ggF and VBF (inclusive results) or produced by only VBF (fixing the ggF cross section to its SM prediction). The observed (expected) 95% CL upper limit on inclusive HH production cross section corresponds to 3.3 (5.2) times the theoretical cross section predicted by the SM, the second most sensitive result among all CMS HH analyses. For VBF-only production, this limit is set at 124 (154) times the theoretical prediction, which is the best result among all CMS and ATLAS HH analyses. Additionally, 95% CL constraints can be set to the Higgs self-coupling and its coupling to the vector bosons. Deviations of these couplings from their SM predictions could be produced by BSM effects, so it is really important to try to constrain their values. For κ_λ, the Higgs self-coupling modifier, the observed (expected) constraint

has been set to $-1.7 < \kappa_\lambda < 8.7$ ($-2.9 < \kappa_\lambda < 9.8$). For κ_{2V}, the modifier for the coupling between two Higgs bosons and two vector bosons, the obtained constraint was $-0.4 < \kappa_{2V} < 2.6$ ($-0.6 < \kappa_{2V} < 2.8$). Looking into the future, projections for Run 3 provide an expected upper limit on the inclusive HH cross section of 3.2 times the SM prediction, while for HL-LHC this limit is expected to drop to 1.4 times the SM prediction.

For the new data taking period (Run 3, 2022-2025), a new trigger strategy has been studied, focusing on the $\tau\tau$ decay mode where both τ decay hadronically. This new strategy consists of including in the L1 and HLT menus a trigger that requires two τ_h and one jet, where the p_T threshold on the τ_h is smaller than the one from the double-τ_h triggers, already present in the Run 2 menus. The goal of this strategy is to increase the acceptance gain on the $H \to \tau\tau$ and $HH \to bb\tau\tau$ analyses by lowering the τ p_T threshold. A complete optimization was performed in order to obtain the online τ and jet p_T thresholds of a new L1 seed to be included in the menu, having a compromise between the acceptance gained and the rate added. Once the L1 seed was implemented, its corresponding HLT path was also developed. The L1 seed and HLT path were included prior to the 2022 data taking period, so its performance can be studied using real data and simulation. The rate added was in agreement with the expectations obtained with 2018 zero bias data. The efficiencies from both τ_h and jet objects have also been studied, with a good agreement between data and simulation. A first test has been performed to compute the sensitivity gain in the $HH \to bb\tau\tau$ analysis when considering the new trigger strategy (logical OR of the double-τ_h and the double-τ_h plus jet triggers instead of just the double-τ_h trigger). When only applying the $\tau\tau$ pair selection (both at trigger and offline level) and a minimal jet requirement if the event satisfied the double-τ_h plus jet trigger selection, 6% and 9% sensitivity gains were obtained for the HH ggF and HH VBF samples, respectively.

Finally, the Analytical Method algorithm for muon trigger primitive generation in Phase-2 has been presented. A thorough description of the algorithm has been included. The algorithm has been validated with a simulated sample assuming the worst expected conditions, showing excellent performance both in efficiency and resolution. The agreement between the software emulator and the firmware implementation has been studied using a dedicated setup in the laboratory, obtaining a matching between both implementations above 98% when all qualities are considered, and almost 100% for very high quality trigger primitives. When performing a fit with the same hits and laterality combinations, the trigger primitive parameters (time, position, and direction) match between both implementations to the least significant bit. The hardware implementation (from the previous algorithm version) has been tested with collision data at CMS, providing very good results in efficiency, time resolution, and matching with the offline reconstructed segments. It is expected that the latest firmware version will be deployed inside the CMS infrastructure in the upcoming months, showing the significant improvement in performance the CMS DT trigger will experience during HL-LHC.

Bibliography

[1] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30–61, 2012.

[2] The ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716:1–29, 09 2012.

[3] `https://upload.wikimedia.org/wikipedia/commons/thumb/0/00/Standard_Model_of_Elementary_Particles.svg/2000px-Standard_Model_of_Elementary_Particles.svg.png`.

[4] R. L. Workman et al. Review of Particle Physics. *PTEP*, 2022:083C01, 2022.

[5] Y. Ashie et al. Measurement of atmospheric neutrino oscillation parameters by Super-Kamiokande I. *Physical Review D*, 71(11), Jun 2005.

[6] P. W. Higgs. Broken symmetries and the masses of gauge bosons. *Physical review letters*, 13(16):508, 1964.

[7] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Physical review letters*, 13(9):321, 1964.

[8] J. Ellis. Higgs Physics. pages 117–168, 2015. 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.

[9] CMS Collaboration. A measurement of the Higgs boson mass in the diphoton decay channel. *Physics Letters B*, 805:135425, 2020.

[10] L. D. Landau. On the angular momentum of a system of two photons. *Dokl. Akad. Nauk SSSR*, 60(2):207–209, 1948.

[11] C. N. Yang. Selection Rules for the Dematerialization of a Particle into Two Photons. *Phys. Rev.*, 77:242–245, Jan 1950.

[12] CMS Collaboration. Study of the Mass and Spin-Parity of the Higgs Boson Candidate via Its Decays to Z Boson Pairs. *Phys. Rev. Lett.*, 110:081803, Feb 2013.

[13] D. de Florian et al. *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*. CERN Yellow Reports: Monographs. CERN, Geneva, 2017. 869 pages, 295 figures, 248 tables and 1645 citations. Working Group web page: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHXSWG.

[14] A portrait of the Higgs boson by the CMS experiment ten years after the discovery. *Nature*, 607(7917):60–68, Jul 2022.

[15] Frederix R. et al. Higgs pair production at the LHC with NLO and parton-shower effects. *Physics Letters B*, 732:142–149, 2014.

[16] B Di Micco et al. Higgs boson potential at colliders: status and perspectives. Technical report, CERN, Geneva, 2019.

[17] S. Amoroso et al. Les Houches 2019: Physics at TeV Colliders: Standard Model Working Group Report, 2020.

[18] F. Bishara, R. Contino, and J. Rojo. Higgs pair production in vector-boson fusion at the LHC and beyond. *The European Physical Journal C*, 77(7), Jul 2017.

[19] R. D. Ball et al. Parton distributions for the LHC Run II. *Journal of High Energy Physics*, 2015(4):1–148, 2015.

[20] R. D. Ball et al. Parton distributions from high-precision collider data. *The European Physical Journal C*, 77:1–75, 2017.

[21] P.A. Zyla et al. (particle data group), prog. theor. exp. phys. 2020, 083c01. 2020.

[22] S. Frixione, P. Nason, and C. Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *Journal of High Energy Physics*, 2007(11):070–070, Nov 2007.

[23] J. Alwall et al. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7), Jul 2014.

[24] T. Sjöstrand et al. An introduction to PYTHIA 8.2. *Computer Physics Communications*, 191:159–177, 2015.

[25] J. Alwall et al. Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *The European Physical Journal C*, 53:473–500, 2008.

[26] R. Frederix and S. Frixione. Merging meets matching in MC@NLO. *Journal of High Energy Physics*, 2012(12):1–41, 2012.

[27] V. Khachatryan et al. Event generator tunes obtained from underlying event and multiparton scattering measurements. *The European Physical Journal C*, 76:1–52, 2016.

[28] CMS Collaboration. Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements. *Eur. Phys. J. C*, 80:4, 2020.

[29] S. Agostinelli et al. Geant4 — a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.

[30] O. S. Brüning et al. *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004.

[31] E. Mobs. The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Août 2018. 2018. General Photo.

[32] Public CMS Luminosity Information. Available at `https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults`.

[33] G. Apollinari et al. High-Luminosity Large Hadron Collider (HL-LHC). Technical Design Report V. 0.1. Technical report, 2017.

[34] CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004–S08004, Aug 2008.

[35] I. Neutelings. CMS coordinate system. Available at `https://tikz.net/axis3d_cms/`.

[36] T. Sakuma. Cutaway diagrams of CMS detector. 2019.

[37] CMS Collaboration. *The CMS tracker system project: Technical Design Report*. Technical design report. CMS. CERN, Geneva, 1997.

[38] CMS Collaboration. *The CMS tracker: addendum to the Technical Design Report*. Technical design report. CMS. CERN, Geneva, 2000.

[39] CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *Journal of Instrumentation*, 9(10):P10009–P10009, Oct 2014.

[40] CMS Collaboration. CMS Technical Design Report for the Pixel Detector Upgrade. Technical report, 2012.

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

[42] A. Benaglia. The CMS ECAL performance with examples. Technical report, CERN, Geneva, 2014.

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

[44] CMS Collaboration. Performance of the CMS hadron calorimeter with cosmic ray muons and LHC beam data. *Journal of Instrumentation*, 5(03):T03012–T03012, mar 2010.

[45] CMS Collaboration. *The CMS muon project: Technical Design Report*. Technical design report. CMS. CERN, Geneva, 1997.

[46] CMS Collaboration. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV. *JINST*, 13(06):P06015, 2018.

[47] CMS Collaboration. The CMS trigger system. *Journal of Instrumentation*, 12(01):P01020–P01020, Jan 2017.

[48] W.J. Stirling. Parton luminosity and cross section plots. Available at `http://www.hep.ph.ic.ac.uk/~wstirlin/plots/plots.html`.

[49] CMS Collaboration. Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JINST*, 15:P10017. 63 p, Jun 2020.

[50] K. Bos et al. *LHC computing Grid: Technical Design Report*. Technical design report. LCG. CERN, Geneva, 2005.

[51] I. Bird et al. Update of the Computing Models of the WLCG and the LHC Experiments. Technical report, 2014.

[52] CMS Collaboration. *CMS computing: Technical Design Report*. Technical design report. CMS. CERN, Geneva, 2005. Submitted on 31 May 2005.

[53] R. Brun and F. Rademakers. ROOT — An object oriented data analysis framework. *Nucl. Inst. & Meth. in Phys. Res. A*, 389(1):81–86, 1997.

[54] CMS Collaboration. Calibration of the CMS drift tube chambers and measurement of the drift velocity with cosmic rays. *Journal of Instrumentation*, vol. 5:T03016, 03 2010.

[55] F. Gasparini et al. Bunch crossing identification at LHC using a mean-timer technique. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 336(1):91–97, 1993.

[56] CMS Collaboration. Performance of the CMS drift-tube chamber local trigger with cosmic rays. *JINST*, 5(03):T03003, 2010.

[57] CMS Collaboration. Technical Proposal for the Phase-II Upgrade of the CMS Detector. Technical report, Geneva, 2015.

[58] CMS Collaboration. A MIP Timing Detector for the CMS Phase-2 Upgrade. Technical report, CERN, Geneva, 2019.

[59] CMS Collaboration. The Phase-2 Upgrade of the CMS Endcap Calorimeter. Technical report, CERN, Geneva, 2017.

[60] CMS Collaboration. The Phase-2 Upgrade of the CMS Muon Detectors. Technical report, CERN, Geneva, 2017.

[61] CMS Collaboration. The Phase-2 Upgrade of the CMS Level-1 Trigger. Technical report, CERN, Geneva, 2020. Final version.

[62] D. Barney. CMS Detector Slice. Available at `https://cds.cern.ch/record/2120661`., 2016.

[63] CMS Collaboration. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12(10):P10003, 2017.

[64] W. Adam, B. Mangano, T. Speer, and T. Todorov. Track Reconstruction in the CMS tracker. Technical report, CERN, Geneva, 2006.

[65] CMS Collaboration. The performance of the cms muon detector in proton−proton collisions at $\sqrt{s} = 7$ tev at the lhc. *Journal of Instrumentation*, 8(11):P11002, Nov 2013.

[66] CMS Collaboration. Performance of electron reconstruction and selection with the CMS detector in proton−proton collisions at $\sqrt{s} = 8$ tev. *Journal of Instrumentation*, 10(06):P06005, Jun 2015.

[67] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov. Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC. *Journal of Physics G: Nuclear and Particle Physics*, 31(9):N9, Jul 2005.

[68] CMS Collaboration. Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC. *JINST*, 16(05):P05014, 2021. Replaced with the published version. Added the journal reference and the DOI. All the figures and tables can be found at http://cms-results.web.cern.ch/cms-results/public-results/publications/EGM-17-001 (CMS Public Pages).

[69] M. Cacciari, G. P. Salam, and G. Soyez. The anti-kt jet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063, Apr 2008.

[70] CMS Collaboration. Jet algorithms performance in 13 TeV data. Technical report, CERN, Geneva, 2017.

[71] CMS Collaboration. Pileup mitigation at CMS in 13 TeV data. *Journal of Instrumentation*, 15(09):P09018–P09018, Sep 2020.

[72] CMS Collaboration. Performance of the pile up jet identification in CMS for Run 2. 2020.

[73] CMS Collaboration. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. *Journal of Instrumentation*, 13(05):P05011–P05011, may 2018.

[74] E. Bols, J. Kieseler, M. Verzetti, M. Stoye, and A. Stakia. Jet flavour classification using DeepJet. *Journal of Instrumentation*, 15(12):P12012, Dec 2020.

[75] LHC Higgs Cross Section HH Sub-group. Performance of the CMS DeepJet b tagging algorithm using 41.9 /fb of data from proton-proton collisions at 13 TeV with Phase 1 CMS detector. Last visited on March 4th, 2023.

[76] CMS Collaboration. Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s} = 13$ tev. *Journal of Instrumentation*, 13(10):P10005, Oct 2018.

[77] CMS Collaboration. Performance of the DeepTau algorithm for the discrimination of taus against jets, electron, and muons. 2019.

[78] CMS Collaboration. Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector. *Journal of Instrumentation*, 14(07):P07004–P07004, Jul 2019.

[79] G. Heinrich et al. Probing the trilinear Higgs boson coupling in di-Higgs production at NLO QCD including parton shower effects. *Journal of High Energy Physics*, 2019(6):1–17, 2019.

[80] LHC Higgs Cross Section HH Sub-group. Current recommendations for HH cross-sections. Last visited on February 13th, 2023.

[81] CMS Collaboration. Search for higgs boson pair production in events with two bottom quarks and two tau leptons in proton–proton collisions at $\sqrt{s} = 13$ tev. *Physics Letters B*, 778:101–127, 2018.

[82] CMS Collaboration. Identification of b-quark jets with the CMS experiment. *Journal of Instrumentation*, 8(04):P04013, apr 2013.

[83] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler. Soft Drop. *JHEP*, 05:146, 2014.

[84] L. Bianchini, J. Conway, E. K. Friis, and C. Veelken. Reconstruction of the Higgs mass in H $\to \tau\tau$ Events by Dynamical Likelihood techniques. *Journal of Physics: Conference Series*, 513(2):022035, Jun 2014.

[85] CMS Collaboration. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *Journal of Instrumentation*, 12(02):P02014–P02014, Feb 2017.

[86] CMS Collaboration. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. *Journal of Instrumentation*, 13(05):P05011–P05011, may 2018.

[87] CMS Collaboration. Identification of hadronic tau lepton decays using a deep neural network. *Journal of Instrumentation*, 17(07):P07023, Jul 2022.

[88] CMS Collaboration. Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s}$=13 TeV. *Journal of Instrumentation*, 13(10):P10005, Oct 2018.

[89] M. Erdmann, E. Geiser, Y. Rath, and M. Rieger. Lorentz Boost Networks: autonomous physics-inspired feature engineering. *Journal of Instrumentation*, 14(06):P06006–P06006, Jun 2019.

[90] L. Breiman. Random Forests. *Machine Learning*, 45:5–32, 2001.

[91] Tensorflow documentation. Available at `https://www.tensorflow.org`.

[92] Keras documentation. Available at `https://keras.io/`.

[93] G. C. Strong. On the impact of selected modern deep-learning techniques to the performance and celerity of classification models in an experimental high-energy physics use case. *Machine Learning: Science and Technology*, 1(4):045006, Sep 2020.

[94] CMS Collaboration. Prospects for HH measurements at the HL-LHC. Technical report, CERN, Geneva, 2018.

[95] PyTorch documentation. Available at `https://pytorch.org/`.

[96] G. C. Strong. Lumin, Mar 2019.

[97] The ATLAS and CMS Collaborations and The LHC Higgs Combination Group. Procedure for the LHC Higgs boson search combination in Summer 2011. Technical report, CERN, Geneva, 2011.

[98] G. Cowan, K Cranmer, E Gross, and O Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2), Feb 2011.

[99] R. Barlow and C. Beeston. Fitting using finite Monte Carlo samples. *Computer Physics Communications*, 77(2):219–228, 1993.

[100] CMS Collaboration. Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS. *Eur. Phys. J. C*, 81(9):800, 2021. Replaced with the published version. Added the journal reference and the DOI. All the figures and tables can be found at http://cms-results.web.cern.ch/cms-results/public-results/publications/LUM-17-003 (CMS Public Pages).

[101] CMS Luminosity Measurements for the 2016 Data Taking Period. Technical report, CERN, Geneva, 2017.

[102] CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV. Technical report, CERN, Geneva, 2018.

[103] CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV. Technical report, CERN, Geneva, 2019.

[104] M. Grazzini et al. Higgs boson pair production at NNLO with top quark mass effects. *Journal of High Energy Physics*, 2018(5):1–21, 2018.

[105] F. A. Dreyer and A. Karlberg. Vector-boson fusion Higgs pair production at N^3LO. *Phys. Rev. D*, 98:114016, Dec 2018.

[106] D. de Florian et al. Handbook of lhc higgs cross sections: 4. deciphering the nature of the higgs sector. *arXiv.org*, 2016.

[107] B. Cabouat and T. Sjöstrand. Some dipole shower studies. *The European Physical Journal C*, 78(3):1–20, 2018.

[108] Search for nonresonant Higgs boson pair production in final states with two bottom quarks and two tau leptons in proton-proton collisions at $\sqrt{s}$ = 13 tev. Public results. Available at `http://cms-results.web.cern.ch/cms-results/public-results/superseded/HIG-20-010`.

[109] CMS 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. *Physics Letters B*, 842:137531, 2023.

[110] C.G. Lester and D.J. Summers. Measuring masses of semi-invisibly decaying particle pairs produced at hadron colliders. *Physics Letters B*, 463(1):99–103, Sep 1999.

[111] ATLAS Collaboration. Search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel using 13 TeV *pp* collision data from the ATLAS detector. 2022.

[112] ATLAS Collaboration. Constraining the Higgs boson self-coupling from single- and double-Higgs production with the ATLAS detector using *pp* collisions at $\sqrt{s} = 13$ TeV. 2022.

[113] H. Qu and L. Gouskos. Jet tagging via Particle Clouds. *Physical Review D*, 101(5), Mar 2020.

[114] M. Cepeda et al. Report from Working Group 2: Higgs Physics at the HL-LHC and HE-LHC. *CERN Yellow Rep. Monogr.*, 7:221–584, 2019.

[115] CMS Collaboration. Measurements of Higgs boson production in the decay channel with a pair of τ leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV. 2022.

[116] G. Abbiendi et al. The Analytical Method algorithm for trigger primitives generation at the LHC Drift Tubes detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1049:168103, 2023.

[117] J. Snoek, H. Larochelle, and R. P. Adams. Practical bayesian optimization of machine learning algorithms. *Advances in neural information processing systems*, 25, 2012.

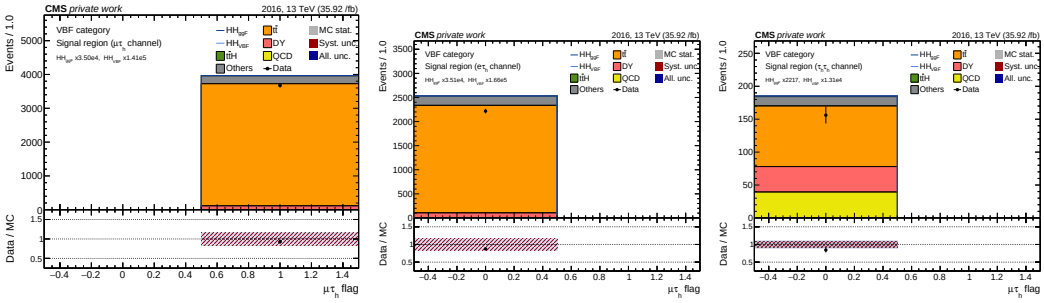
[118] F. Nogueira. Bayesian Optimization: Open source constrained global optimization tool for Python, 2014–.

[119] A. Triossi et al. The CMS Barrel Muon trigger upgrade. *Journal of Instrumentation*, 12:C01095–C01095, 01 2017.

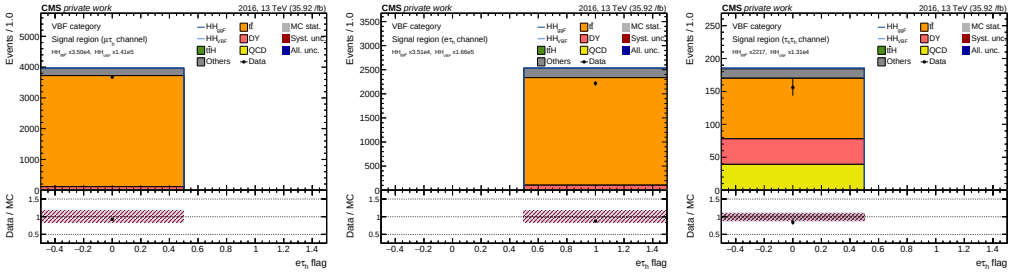
Appendix A

Multi-class classification input variable distributions

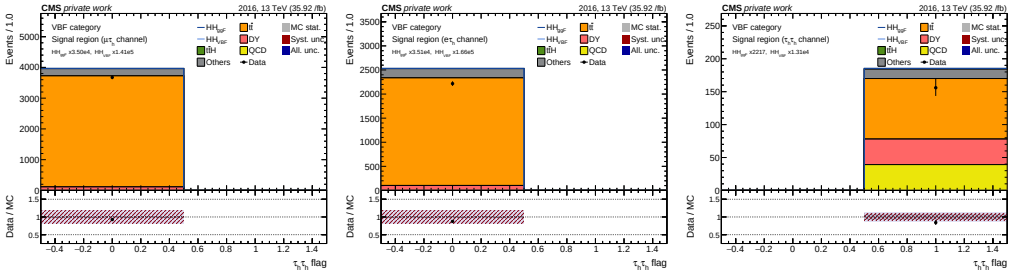
This appendix includes the data and simulated distributions from all the variables used as input for the multi-class classification approach. The three years (2016, 2017, and 2018) and $\tau\tau$ decay channels are considered. The definition from each variable can be found in Table 4.14.



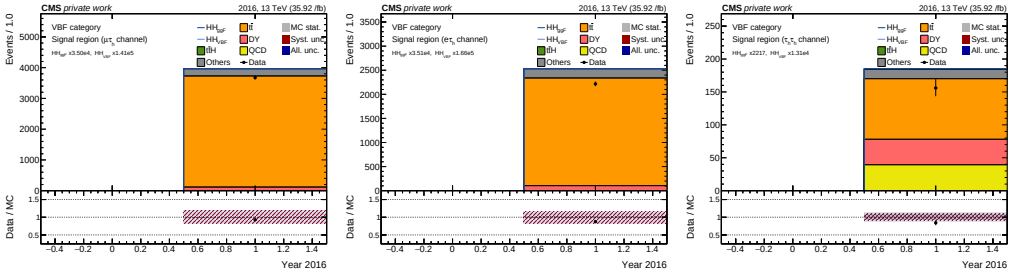
(b) `is_mutau`



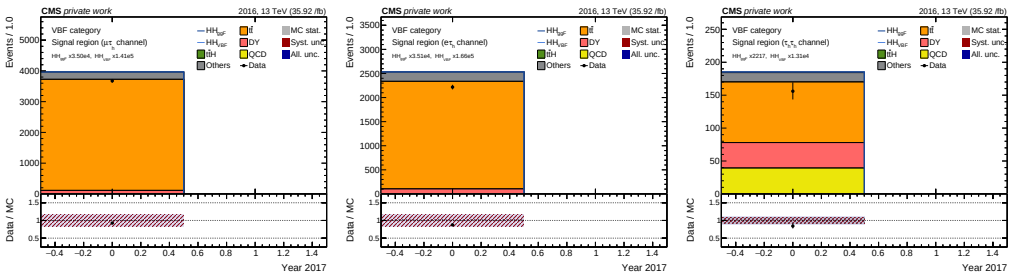
(e) `is_etau`



(h) `is_tautau`

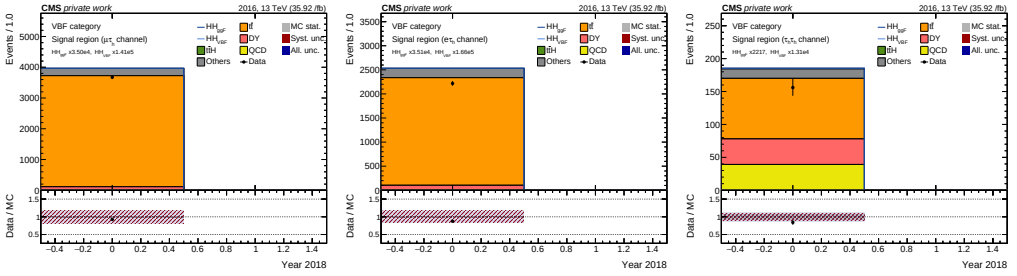


(k) `is_2016`

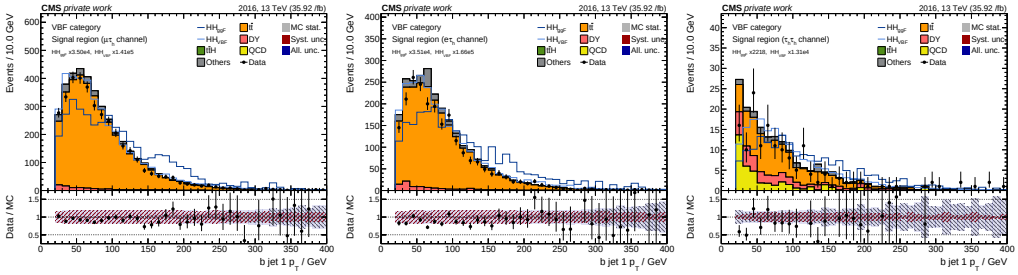


(n) `is_2017`

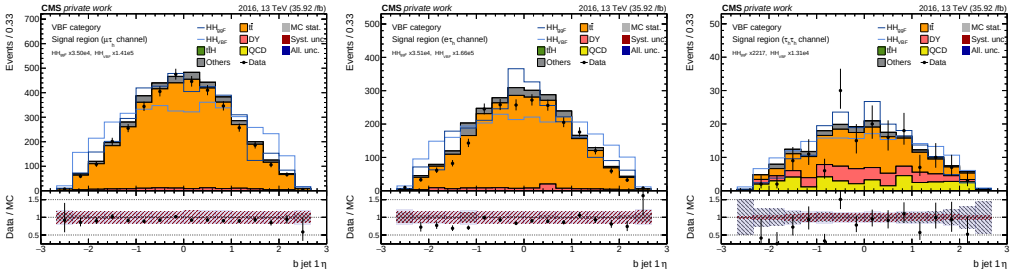
Figure A.1: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2016 data taking period.



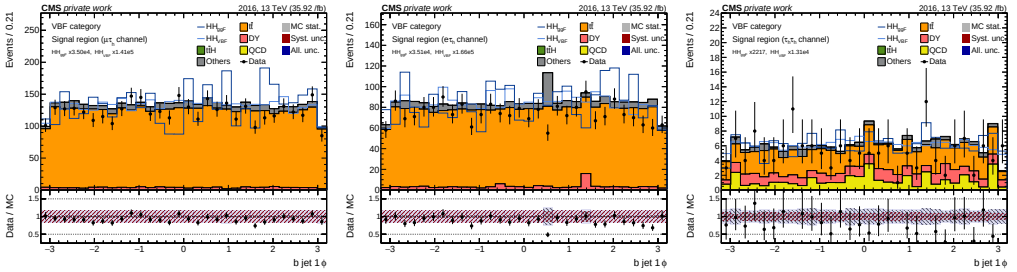
(b) `is_2018`



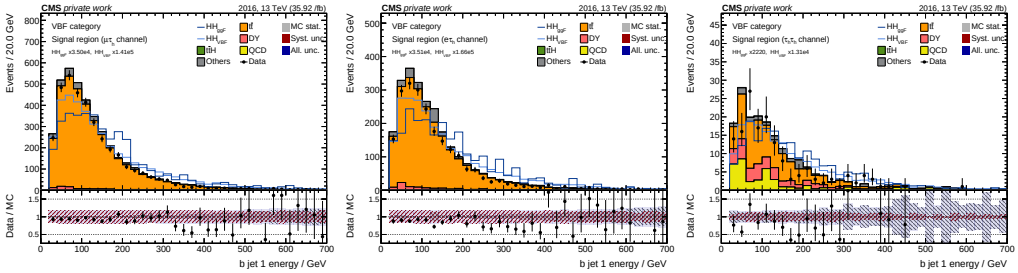
(e) `bjet1_pt`



(h) `bjet1_eta`

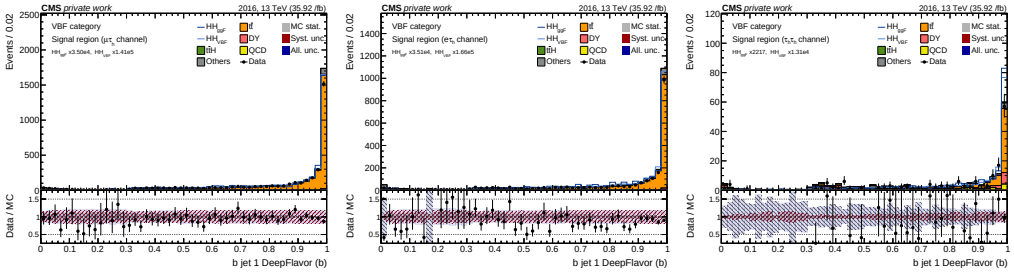


(k) `bjet1_phi`

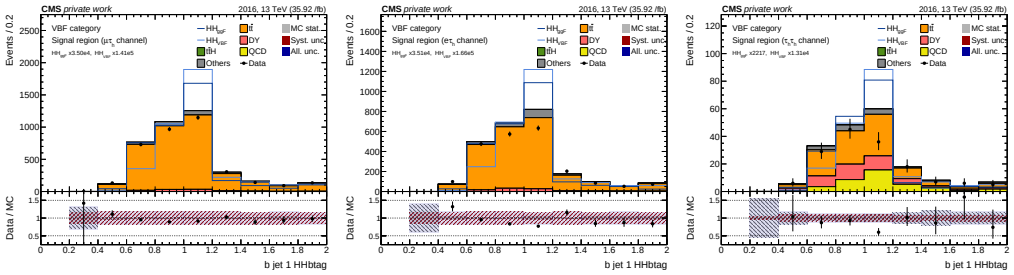


(n) `bjet1_e`

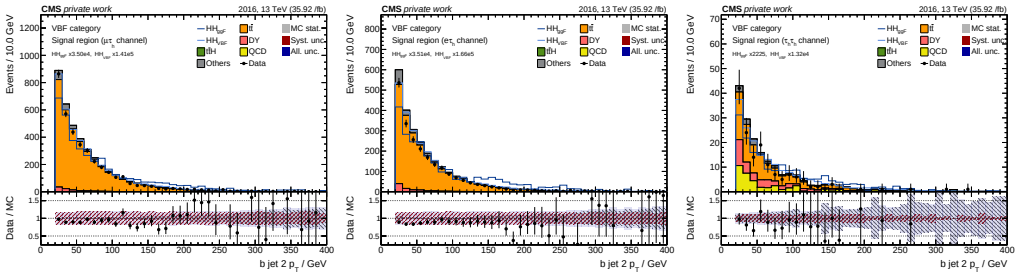
Figure A.2: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2016 data taking period.



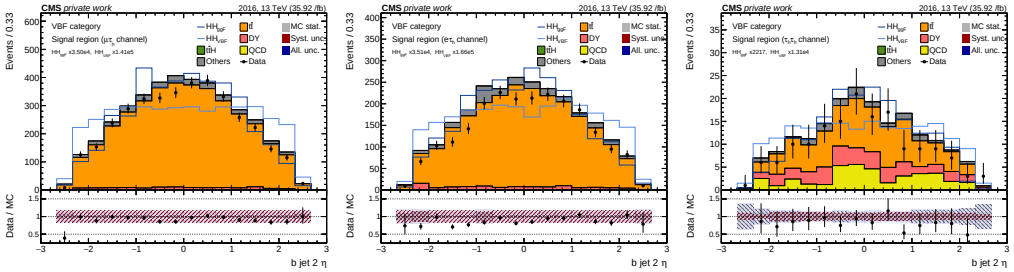
(b) `bjet1_deepflavor_b`



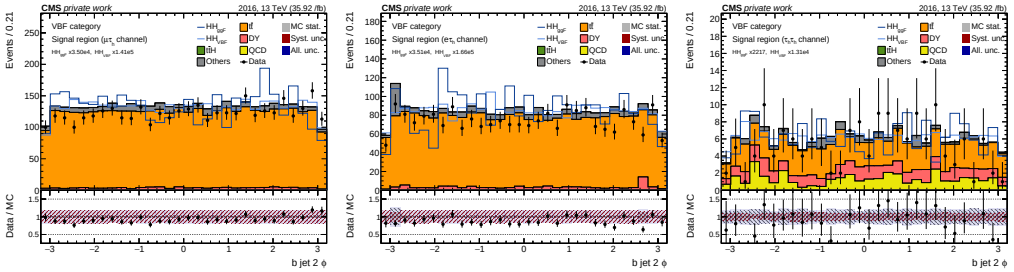
(e) `bjet1_hhbtag`



(h) `bjet2_pt`

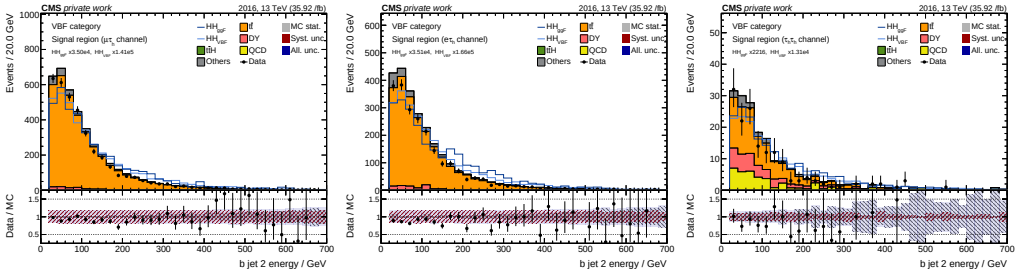


(k) `bjet2_eta`

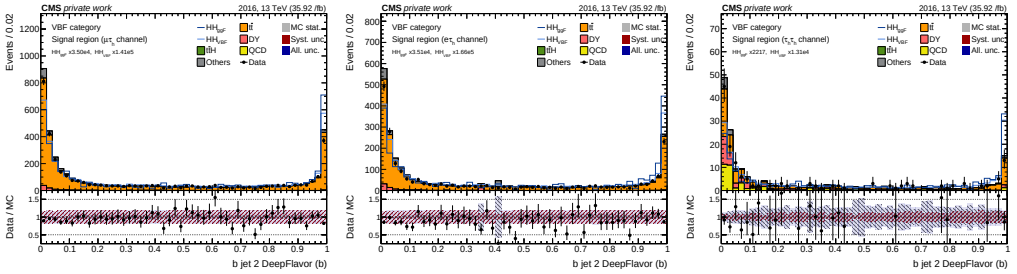


(n) `bjet2_phi`

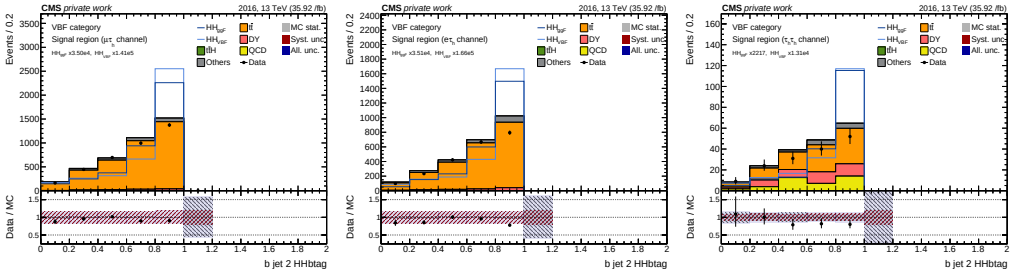
Figure A.3: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2016 data taking period.



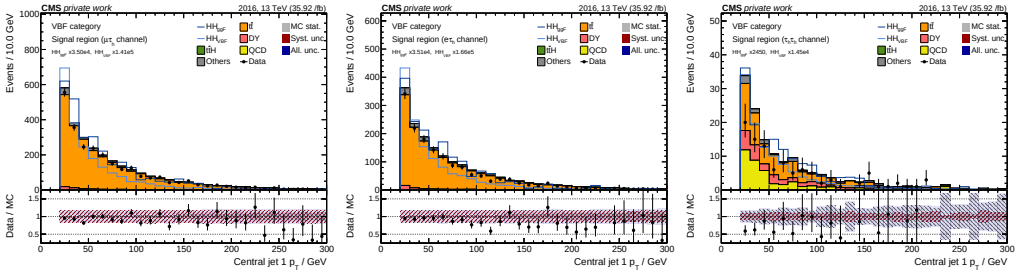
(b) `bjet2_e`



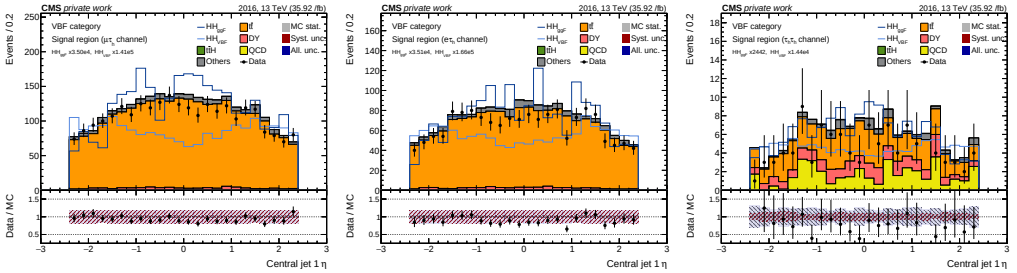
(e) `bjet2_deepflavor_b`



(h) `bjet2_hhbtag`

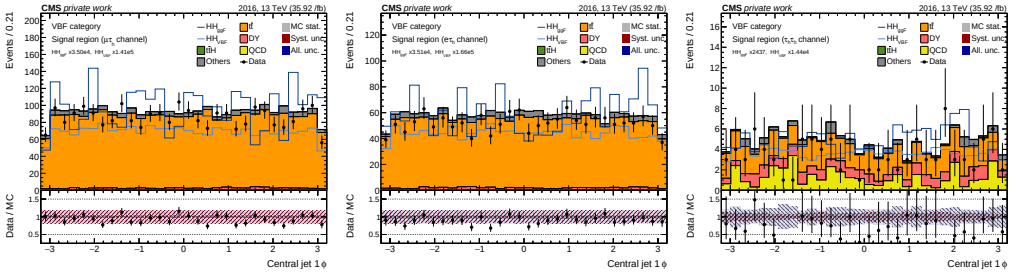


(k) `ctjet1_pt`

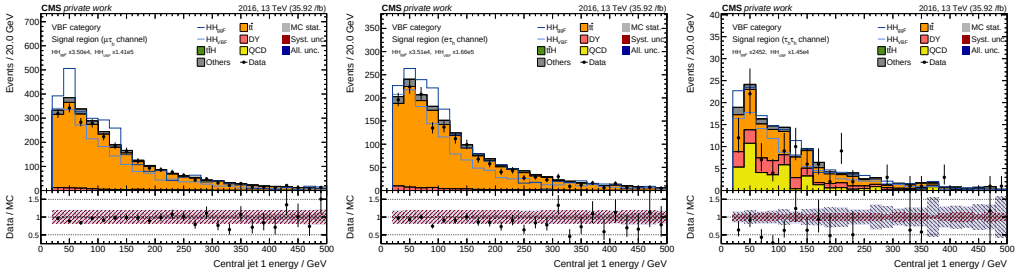


(n) `ctjet1_eta`

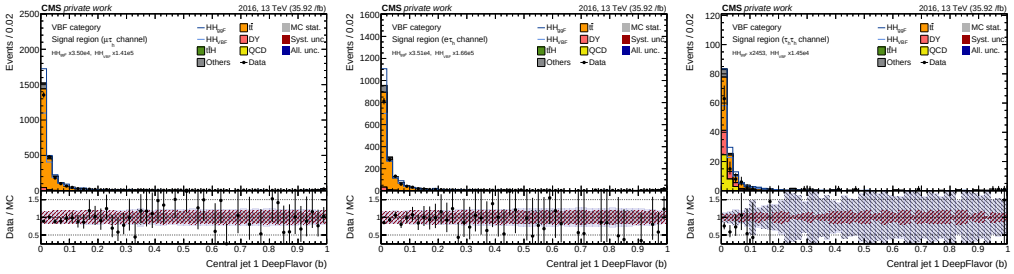
Figure A.4: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2016 data taking period.



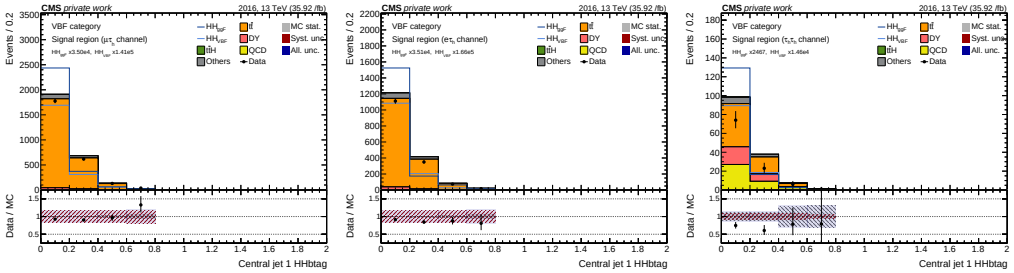
(b) `ctjet1_phi`



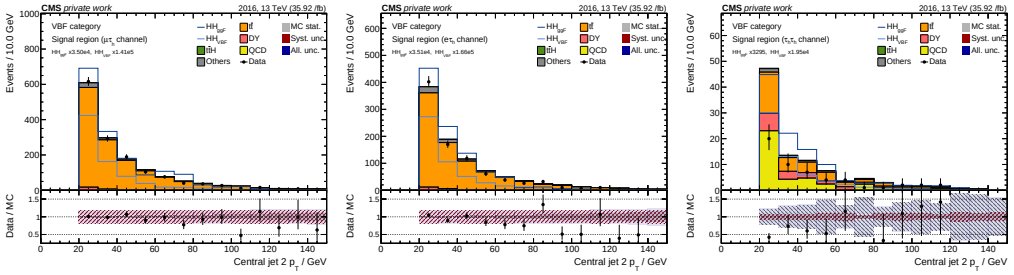
(e) `ctjet1_e`



(h) `ctjet1_deepflavor_b`

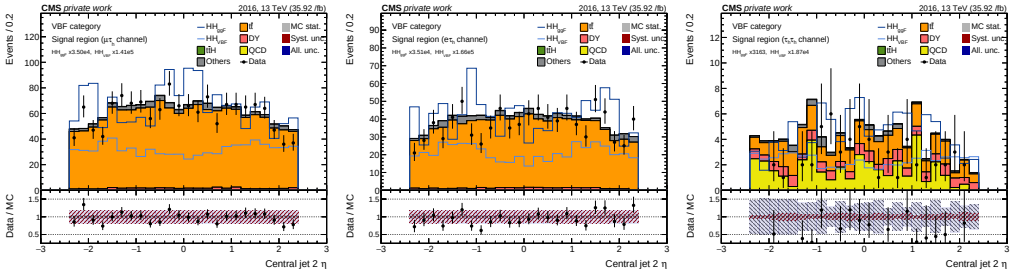


(k) `ctjet1_hhbtag`

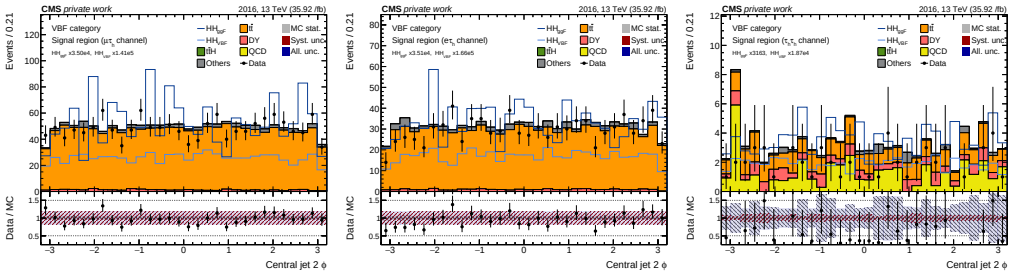


(n) `ctjet2_pt`

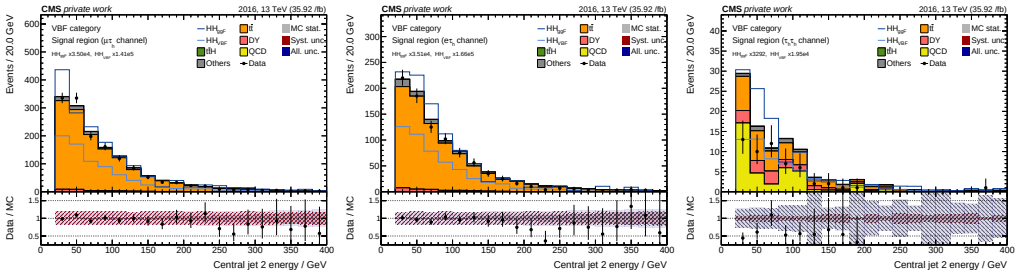
Figure A.5: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2016 data taking period.



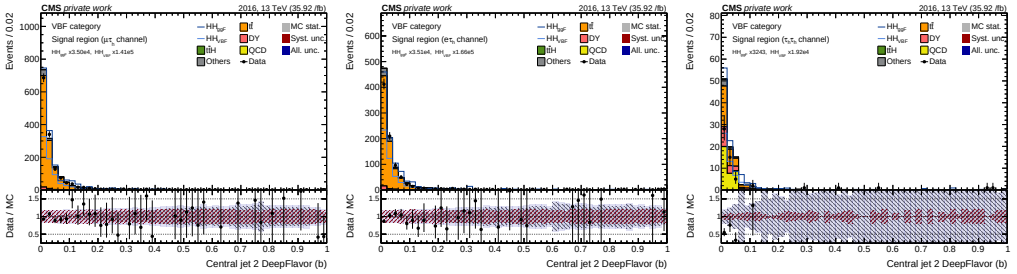
(b) `ctjet2_eta`



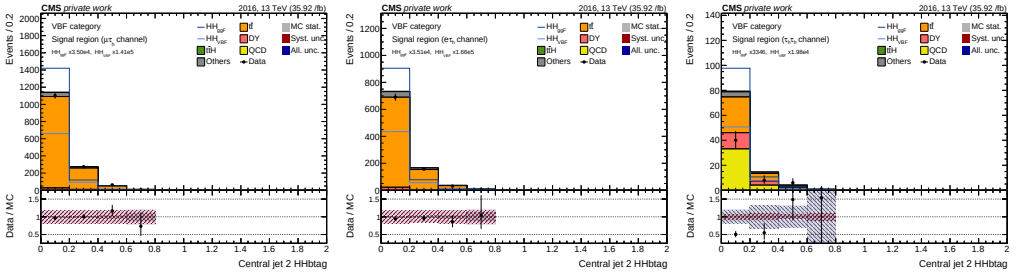
(e) `ctjet2_phi`



(h) `ctjet2_e`

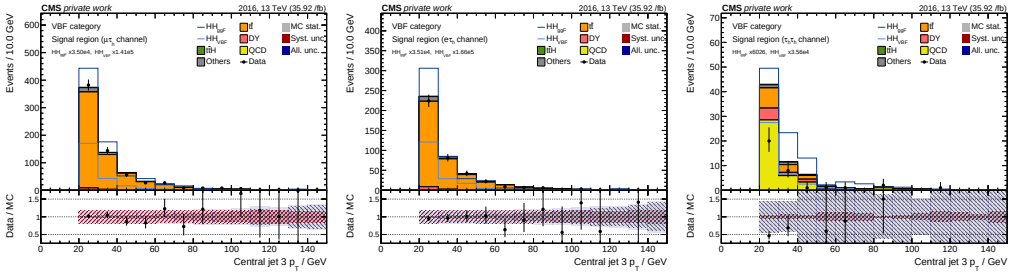


(k) `ctjet2_deepflavor_b`

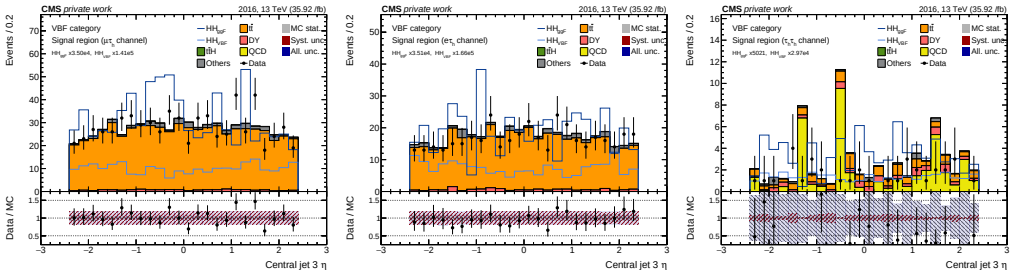


(n) `ctjet2_hhbtag`

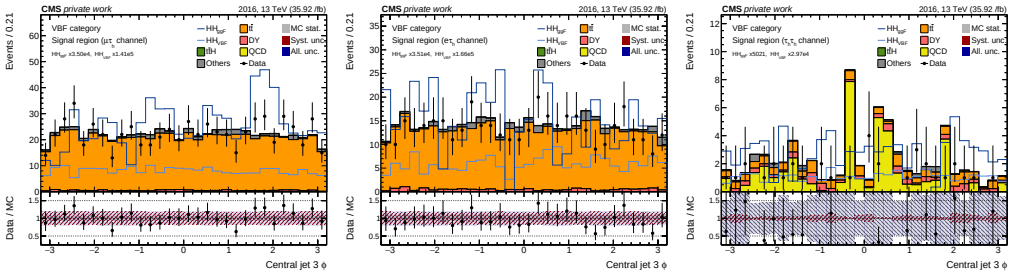
Figure A.6: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2016 data taking period.



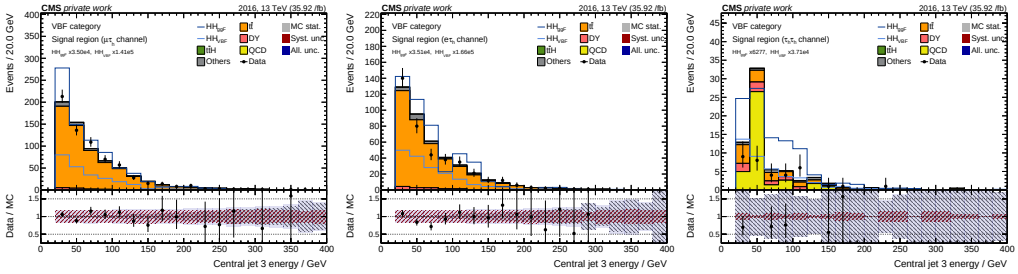
(b) `ctjet3_pt`



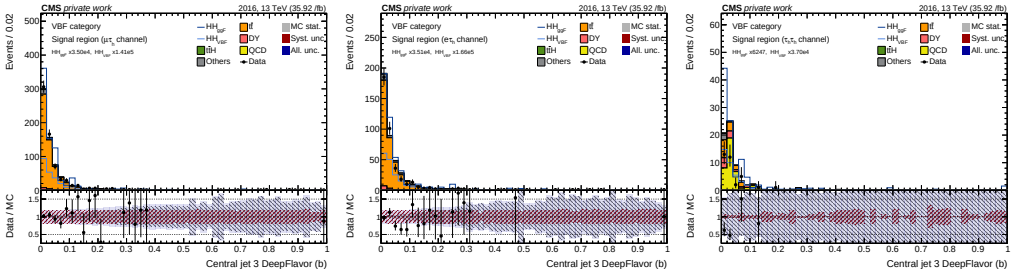
(e) `ctjet3_eta`



(h) `ctjet3_phi`

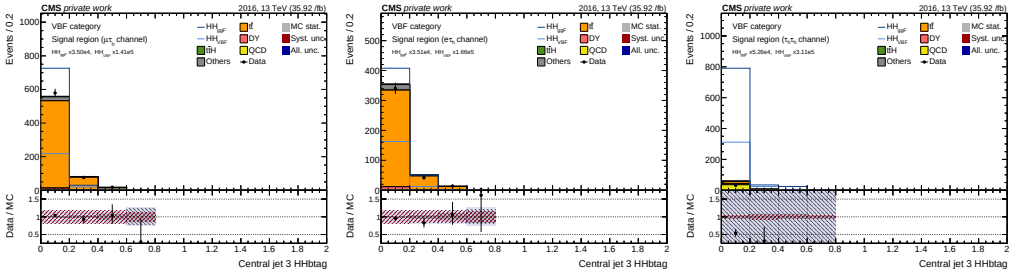


(k) `ctjet3_e`

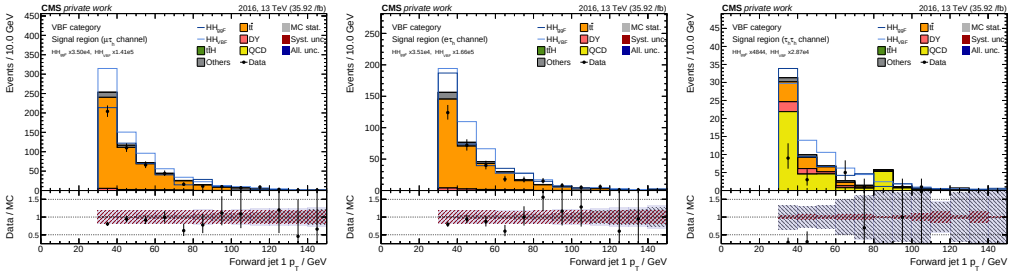


(n) `ctjet3_deepflavor_b`

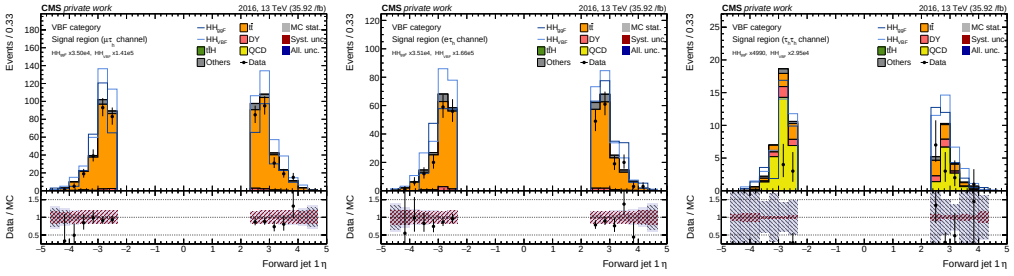
Figure A.7: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2016 data taking period.



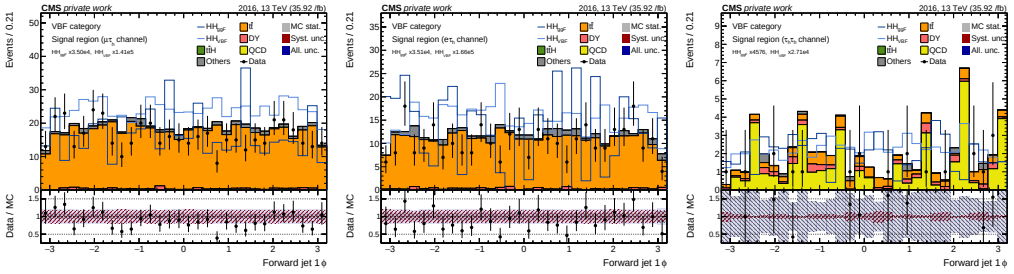
(b) `ctjet3_hhbtag`



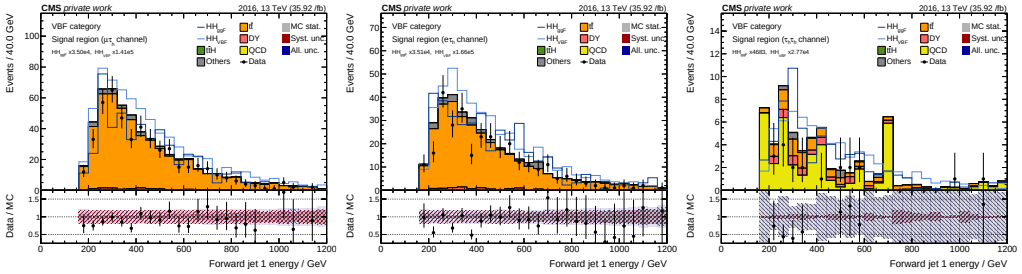
(e) `fwjet1_pt`



(h) `fwjet1_eta`

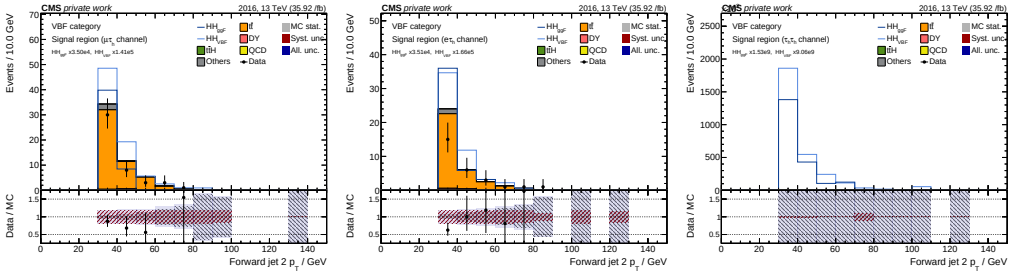


(k) `fwjet1_phi`

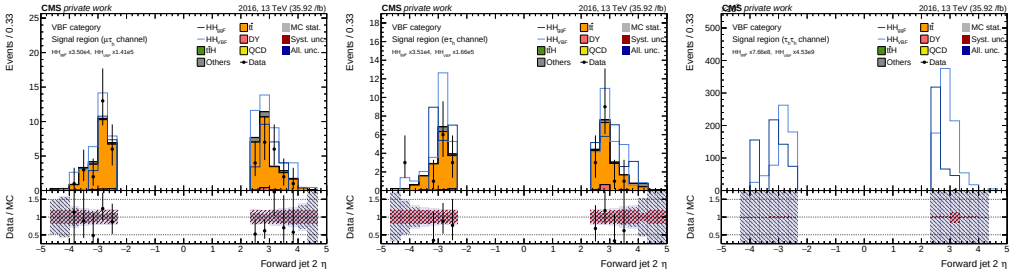


(n) `fwjet1_e`

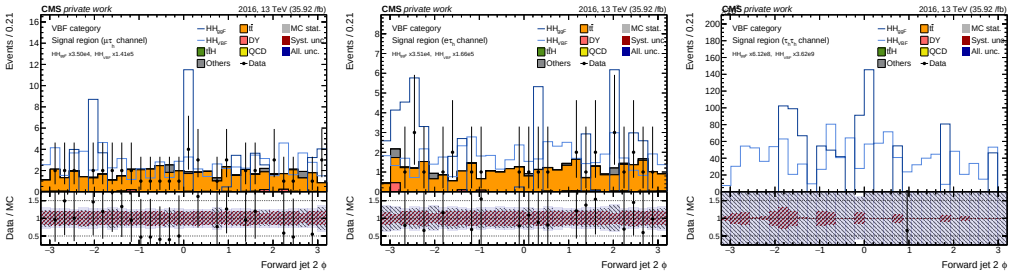
Figure A.8: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2016 data taking period.



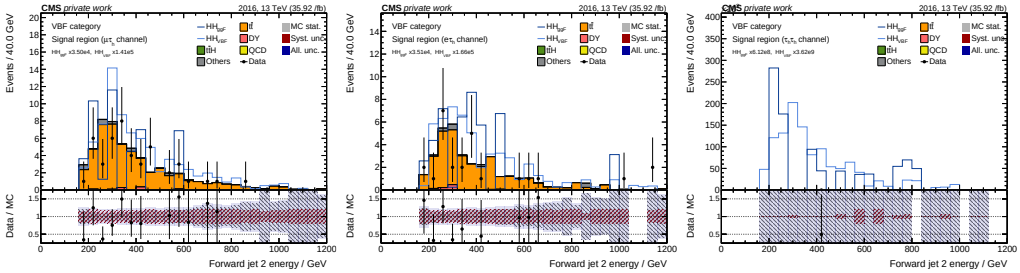
(b) `fwjet2_pt`



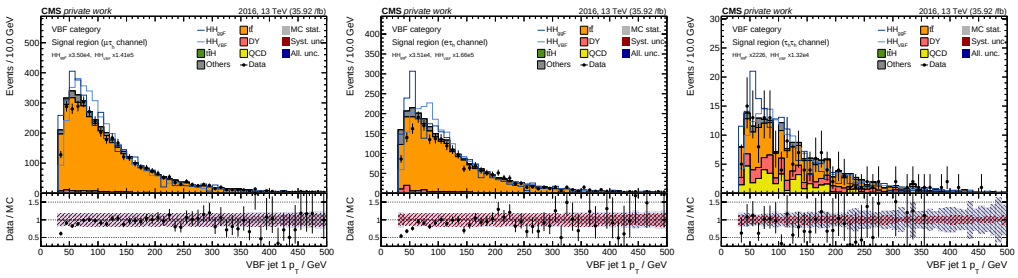
(e) `fwjet2_eta`



(h) `fwjet2_phi`

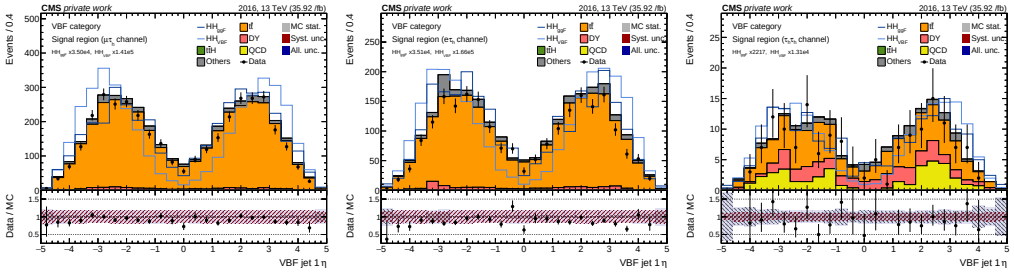


(k) `fwjet2_e`

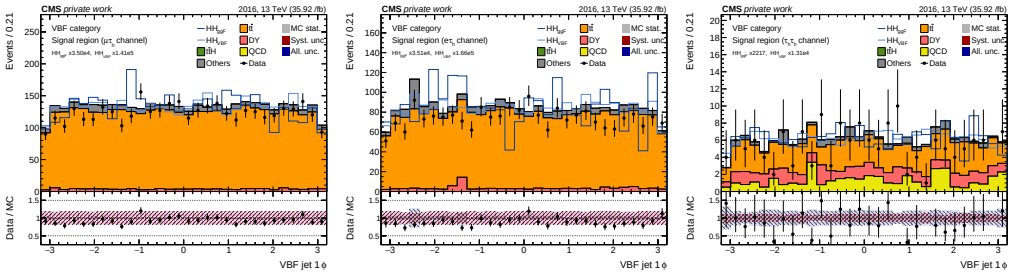


(n) `vbfjet1_pt`

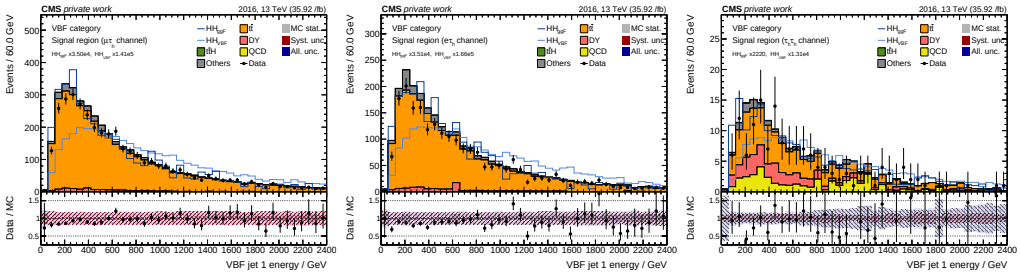
Figure A.9: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2016 data taking period.



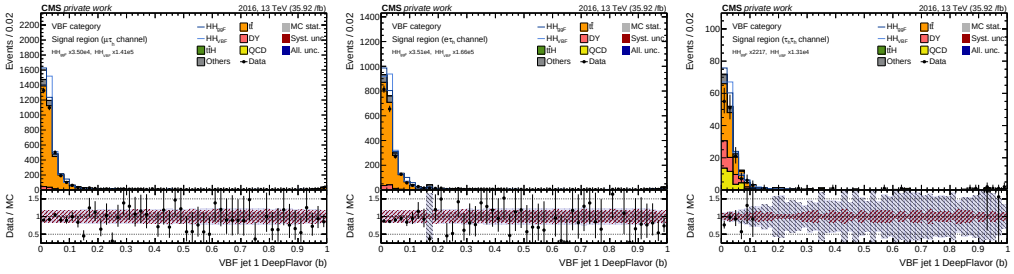
(b) `vbfjet1_eta`



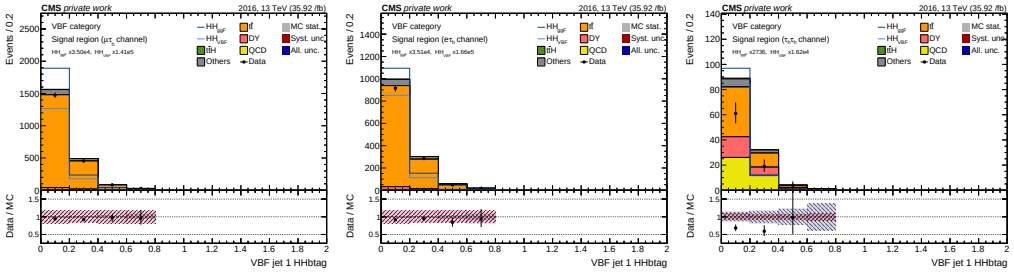
(e) `vbfjet1_phi`



(h) `vbfjet1_e`

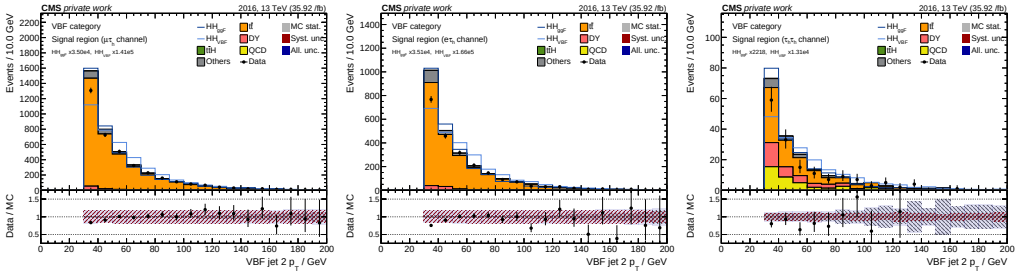


(k) `vbfjet1_deepflavor_b`

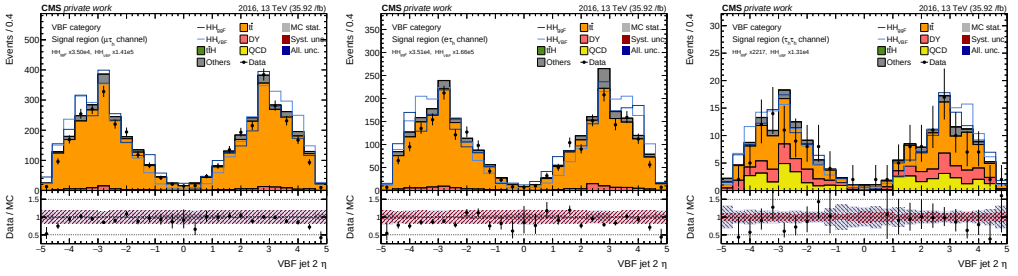


(n) `vbfjet1_hhbtag`

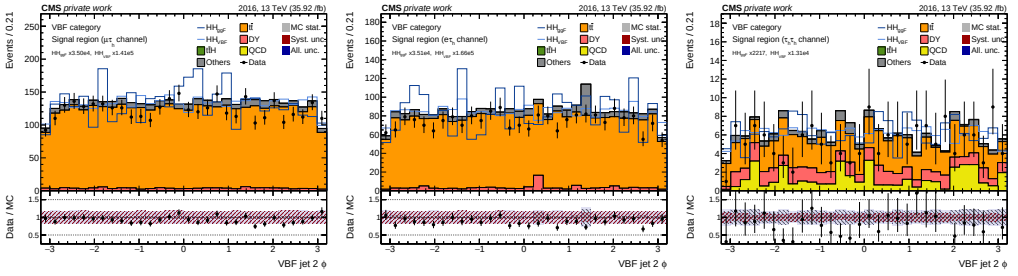
Figure A.10: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2016 data taking period.



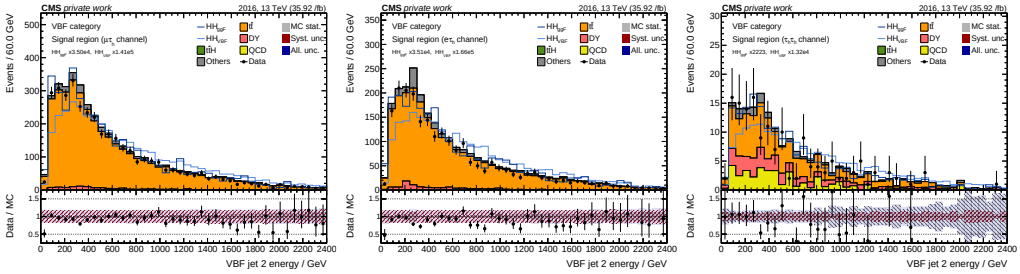
(b) `vbfjet2_pt`



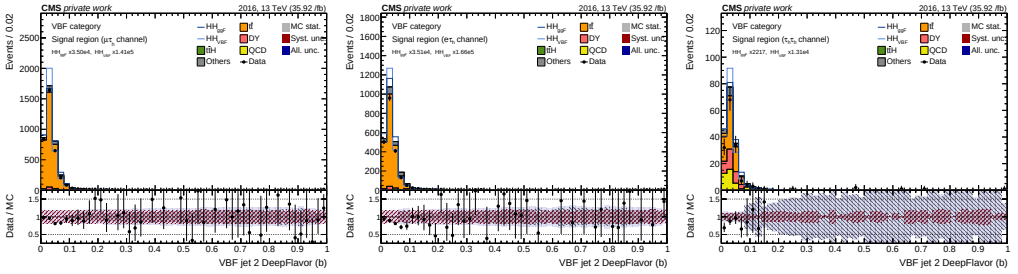
(e) `vbfjet2_eta`



(h) `vbfjet2_phi`

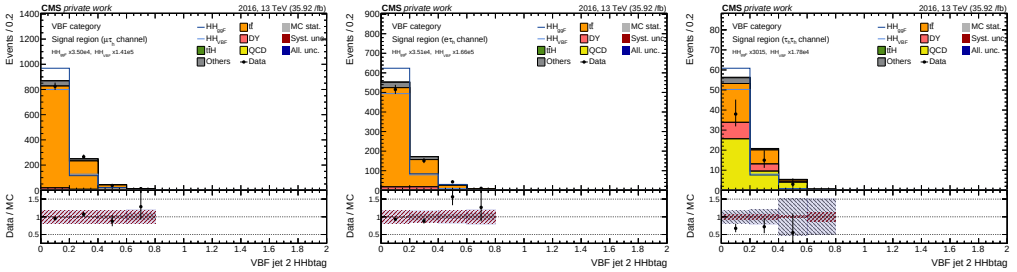


(k) `vbfjet2_e`

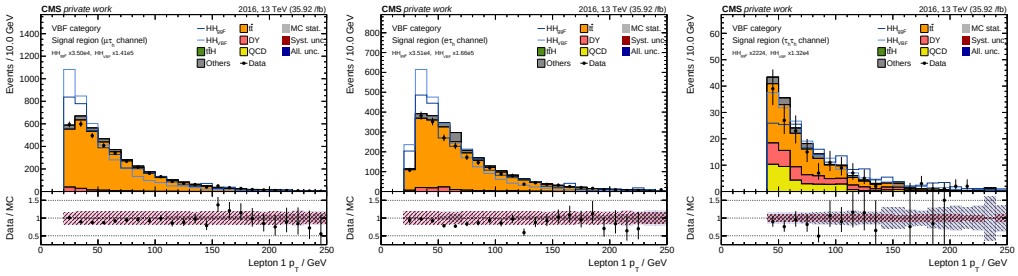


(n) `vbfjet2_deepflavor_b`

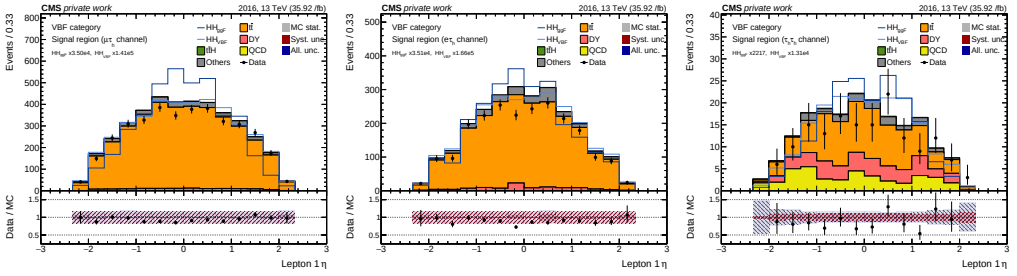
Figure A.11: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2016 data taking period.



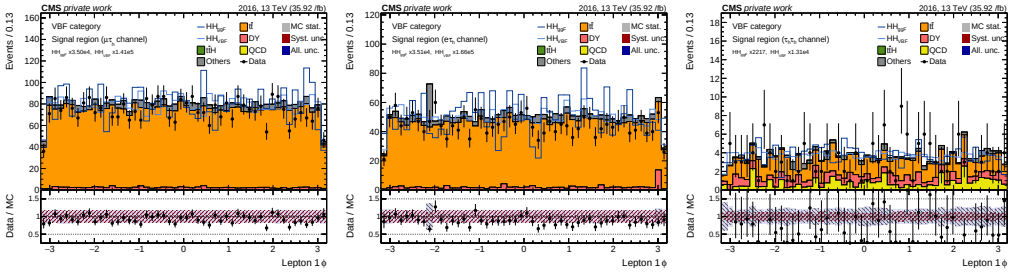
(b) `vbfjet2_hhbtag`



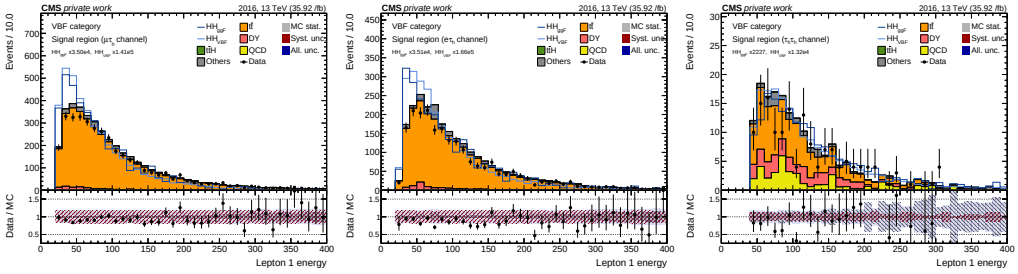
(e) `lep1_pt`



(h) `lep1_eta`

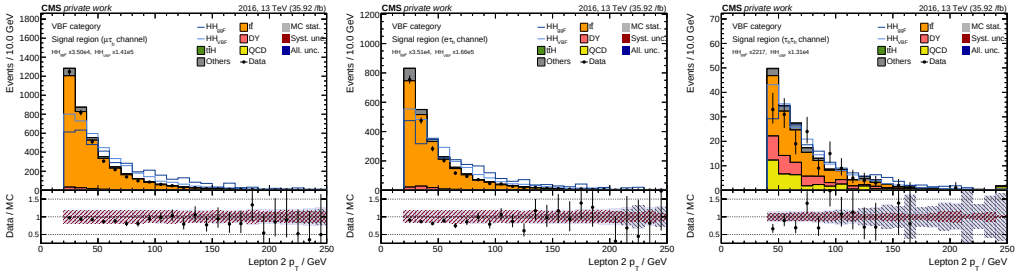


(k) `lep1_phi`

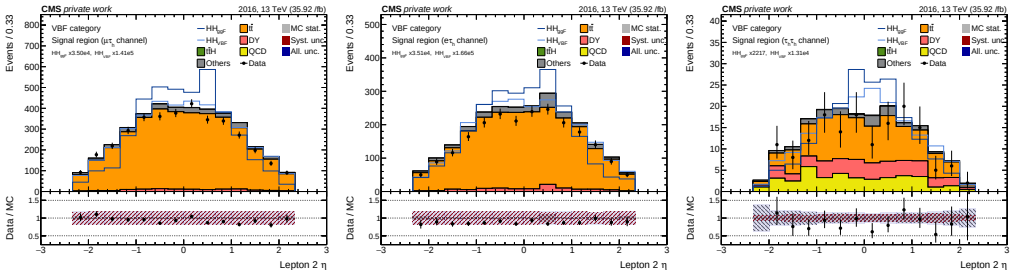


(n) `lep1_e`

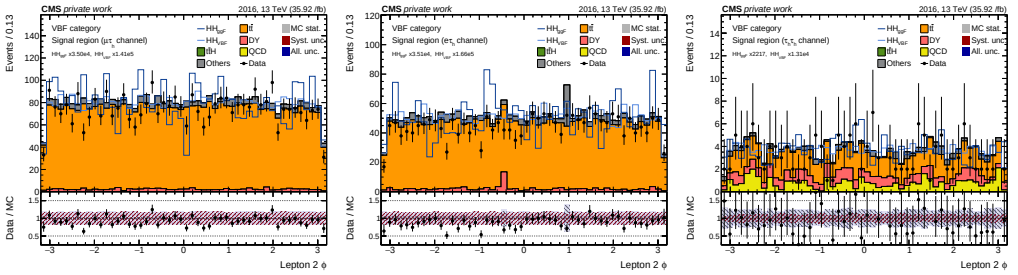
Figure A.12: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2016 data taking period.



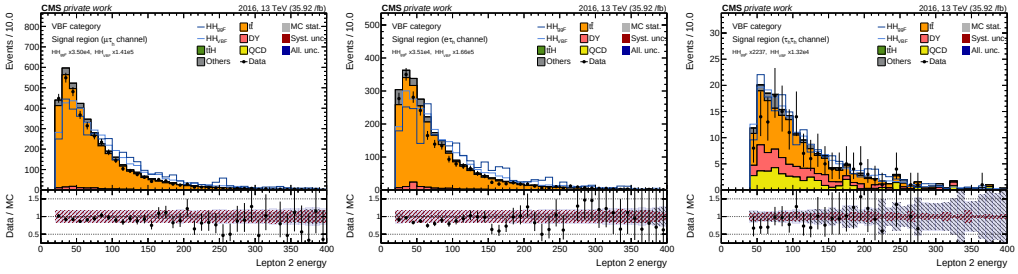
(b) `lep2_pt`



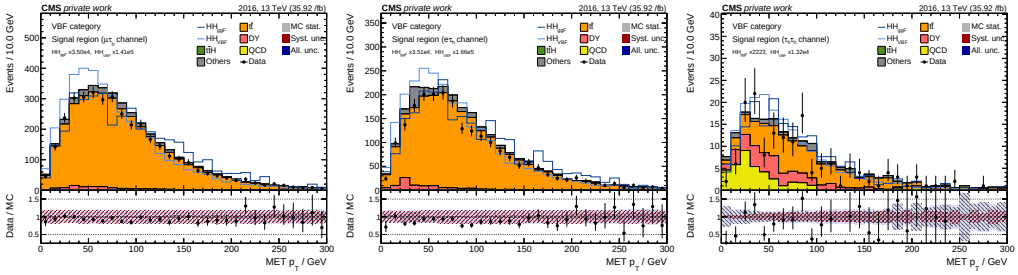
(e) `lep2_eta`



(h) `lep2_phi`

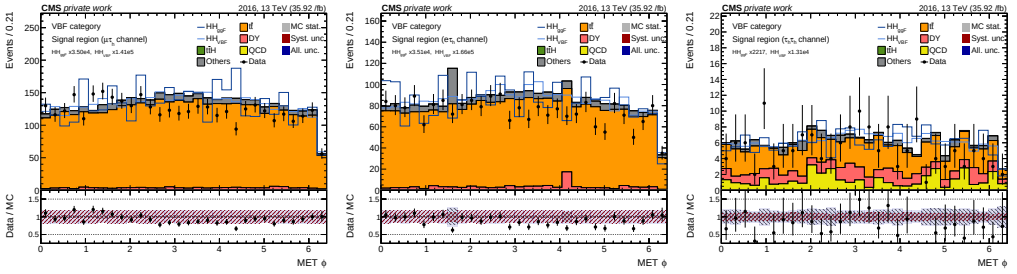


(k) `lep2_e`

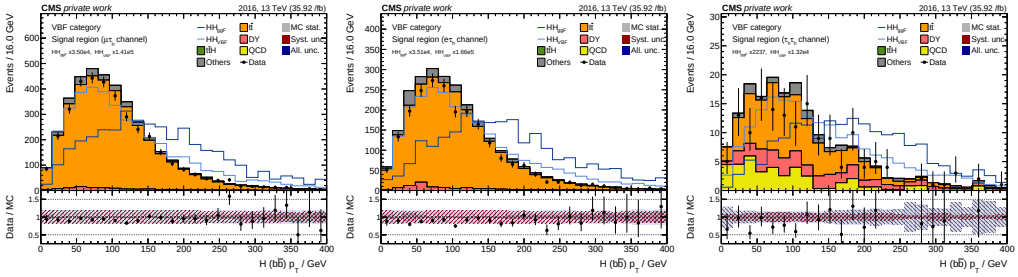


(n) `met_pt`

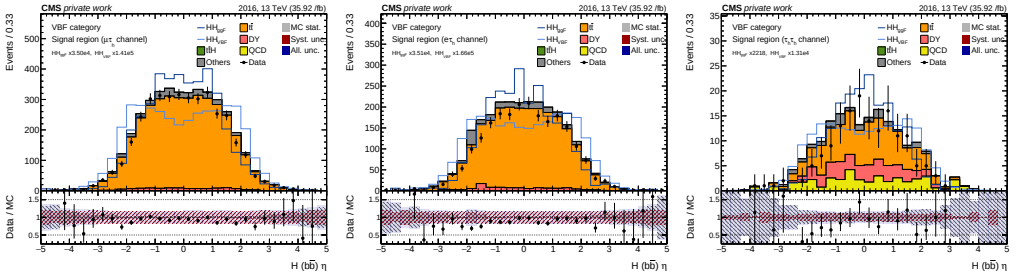
Figure A.13: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2016 data taking period.



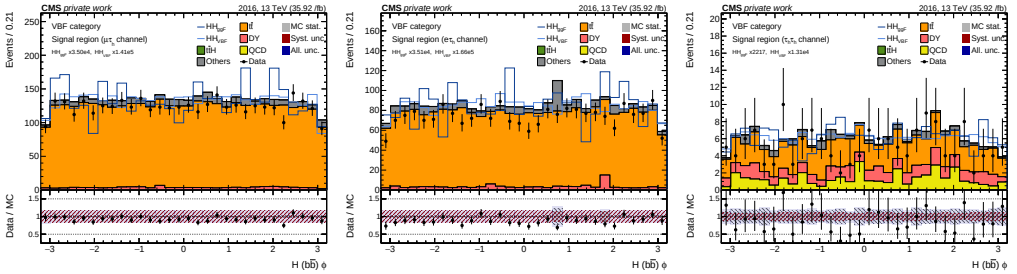
(b) `met_phi`



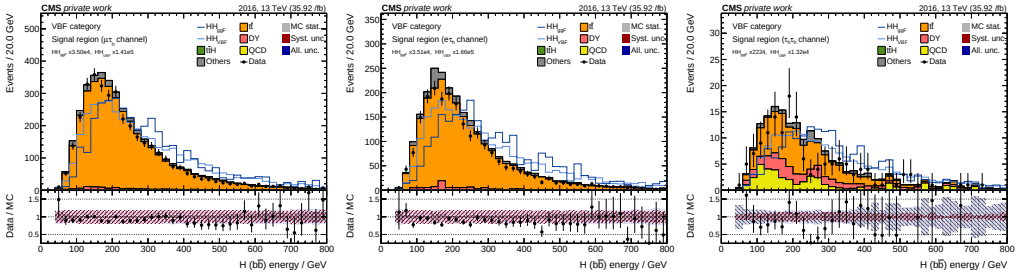
(e) `bh_pt`



(h) `bh_eta`

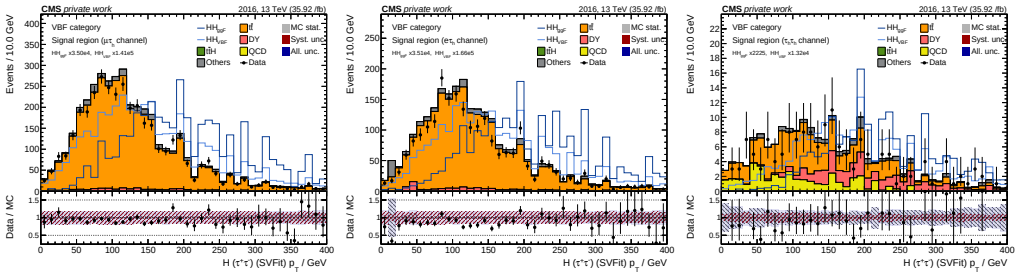


(k) `bh_phi`

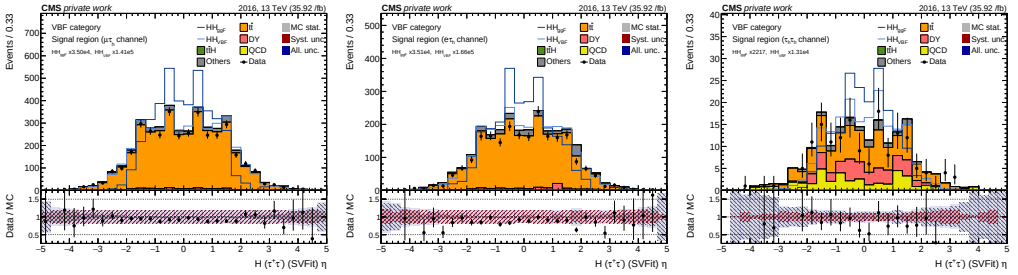


(n) `bh_e`

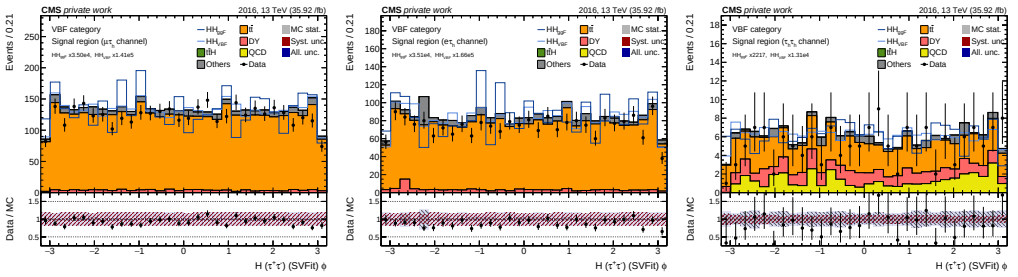
Figure A.14: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2016 data taking period.



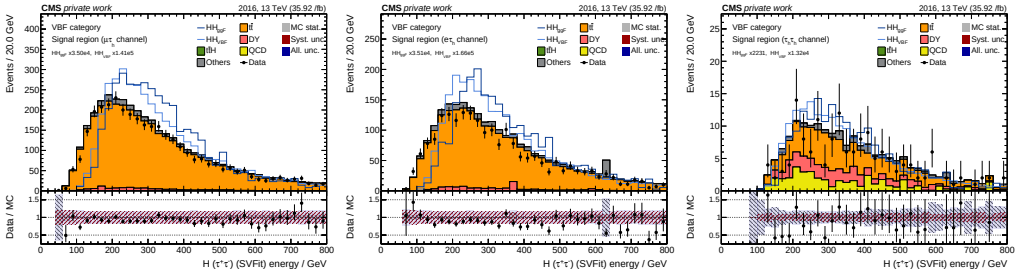
(b) `tauh_sv_pt`



(e) `tauh_sv_eta`

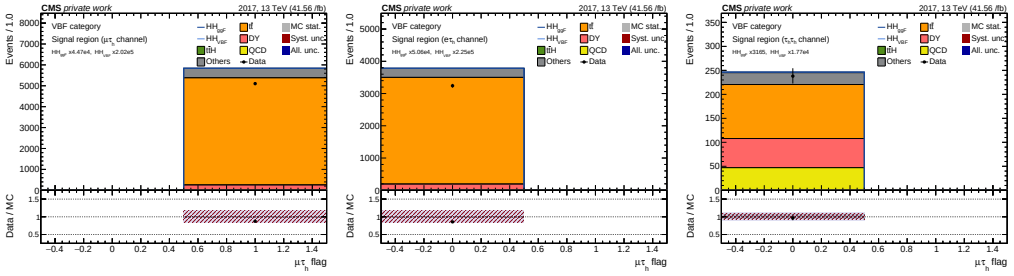


(h) `tauh_sv_phi`

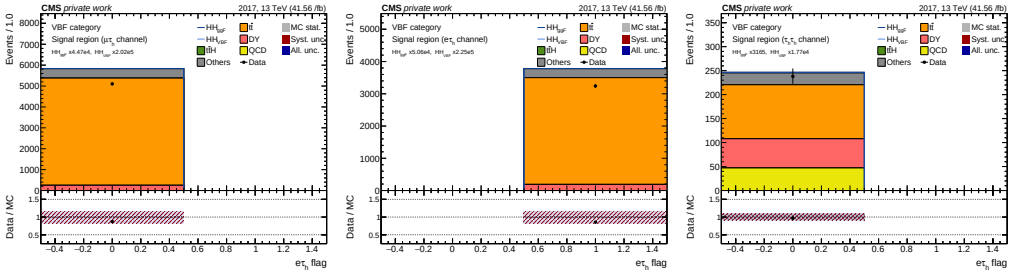


(k) `tauh_sv_e`

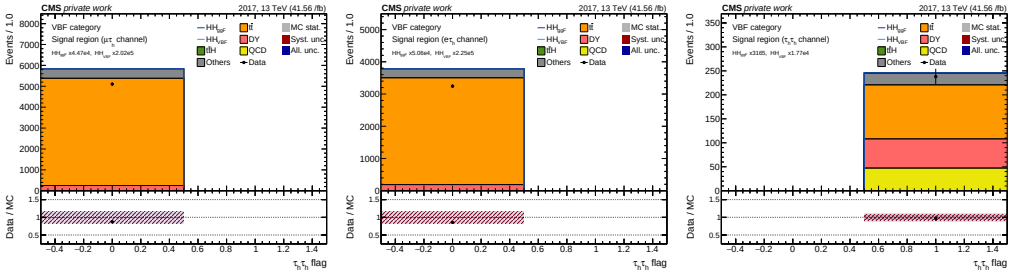
Figure A.15: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2016 data taking period.



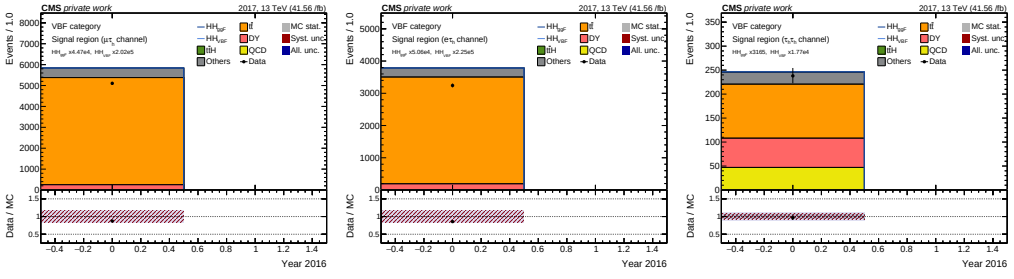
(b) `is_mutau`



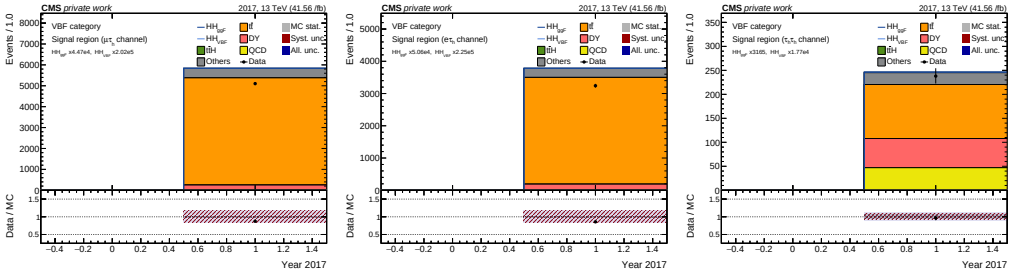
(e) `is_etau`



(h) `is_tautau`

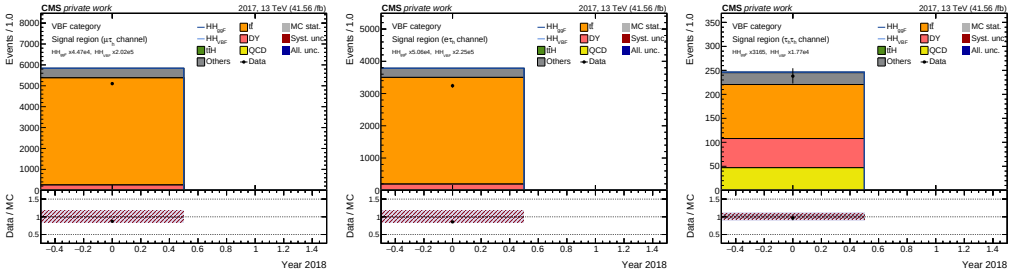


(k) `is_2016`

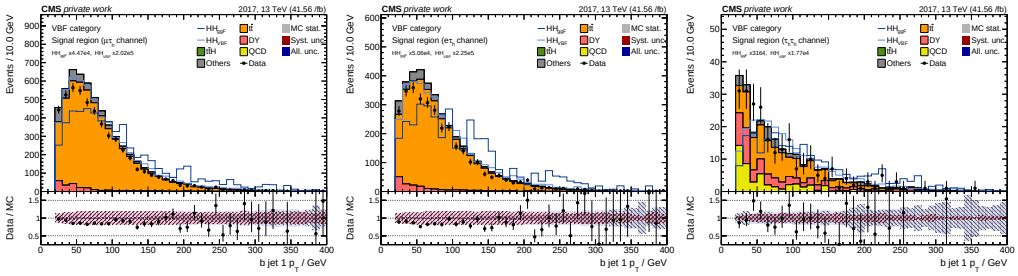


(n) `is_2017`

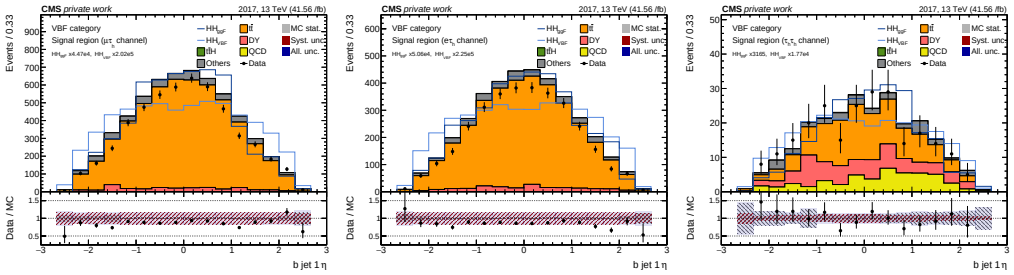
Figure A.16: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2017 data taking period.



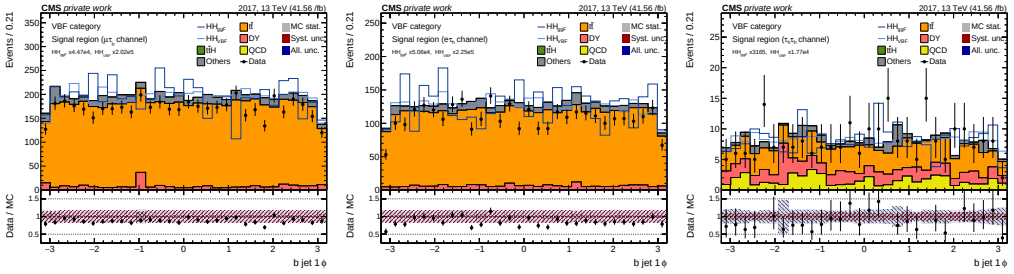
(b) `is_2018`



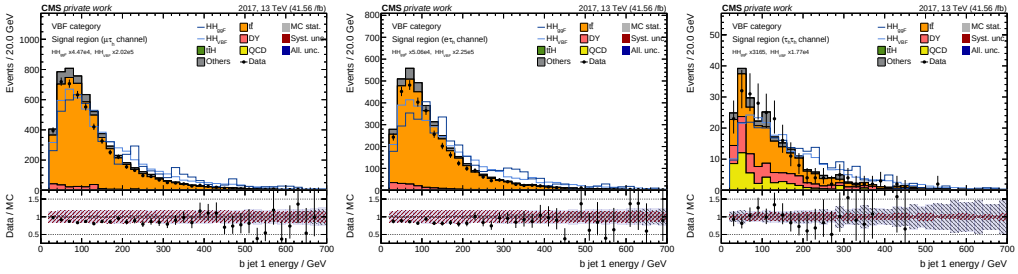
(e) `bjet1_pt`



(h) `bjet1_eta`

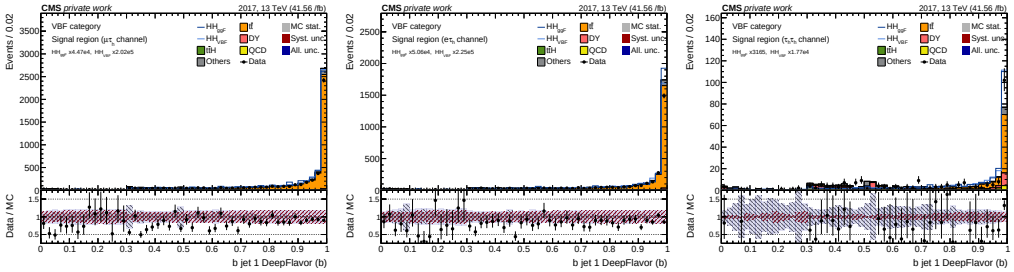


(k) `bjet1_phi`

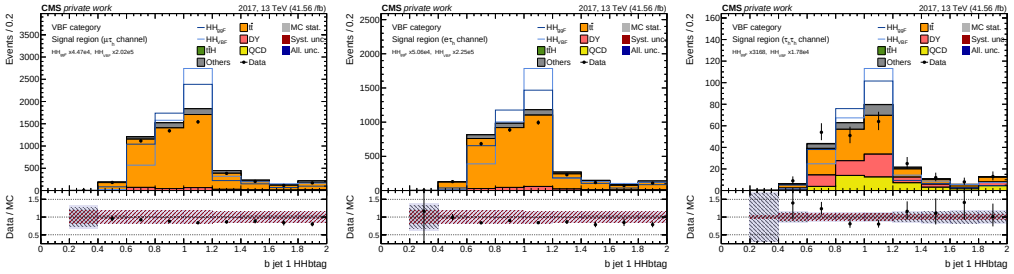


(n) `bjet1_e`

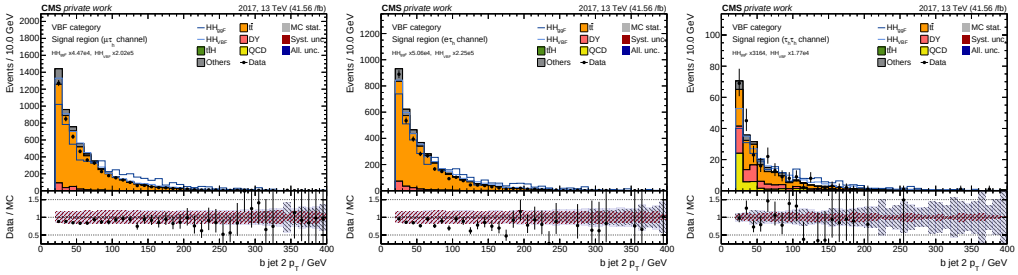
Figure A.17: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2017 data taking period.



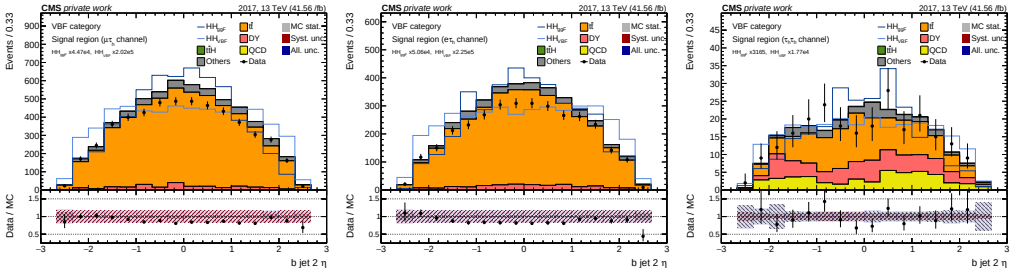
(b) `bjet1_deepflavor_b`



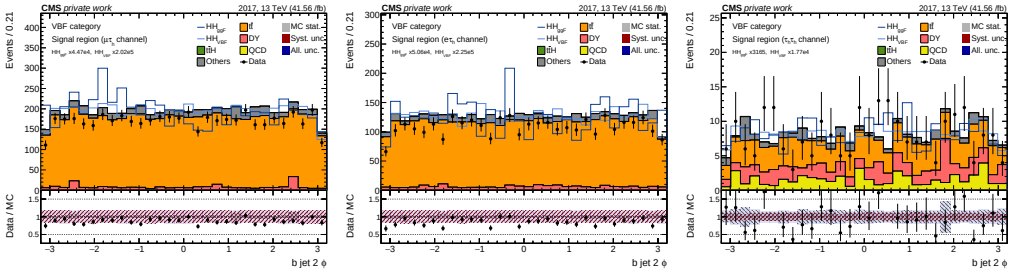
(e) `bjet1_hhbtag`



(h) `bjet2_pt`

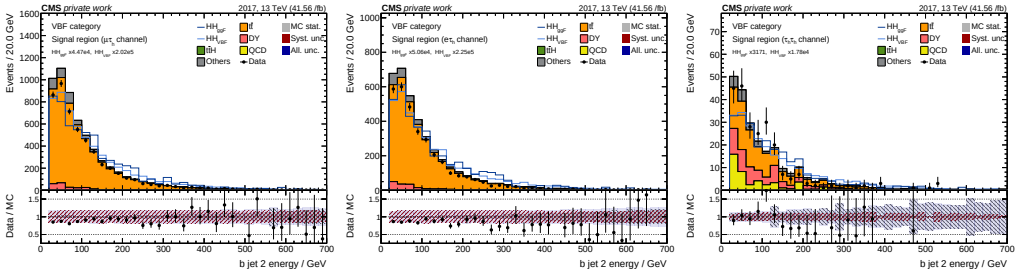


(k) `bjet2_eta`

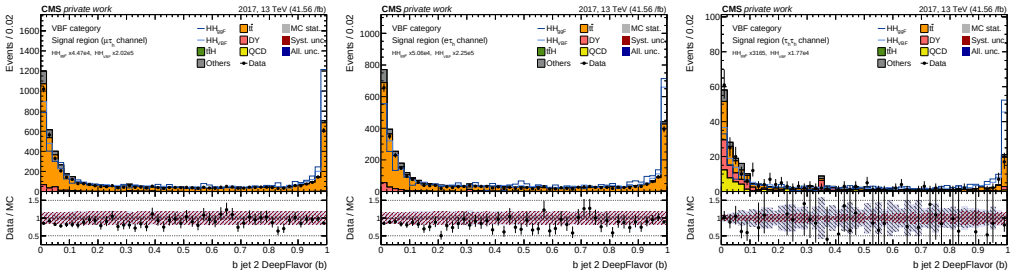


(n) `bjet2_phi`

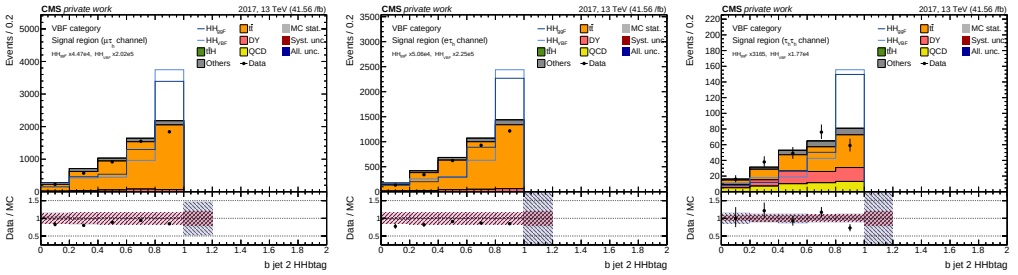
Figure A.18: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2017 data taking period.



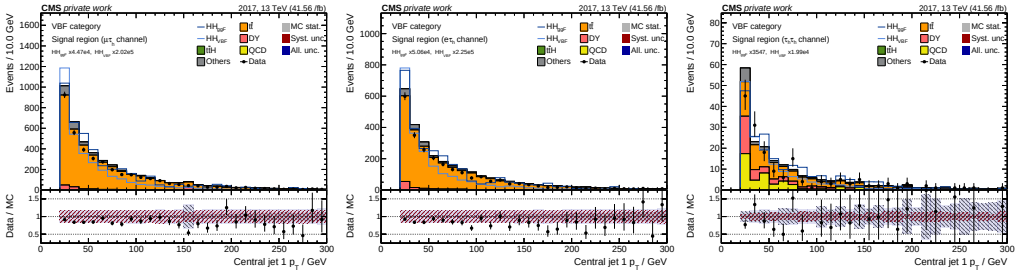
(b) `bjet2_e`



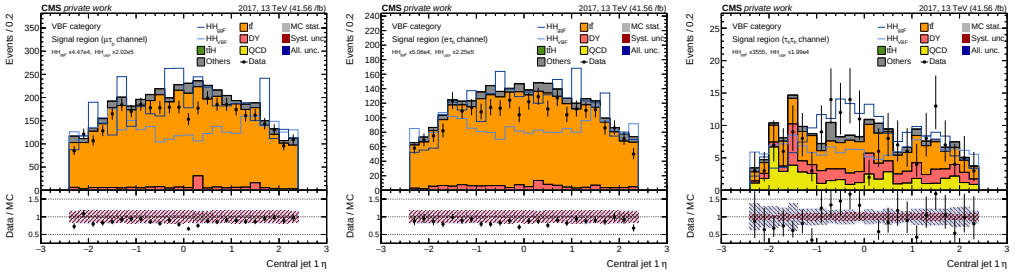
(e) `bjet2_deepflavor_b`



(h) `bjet2_hhbtag`

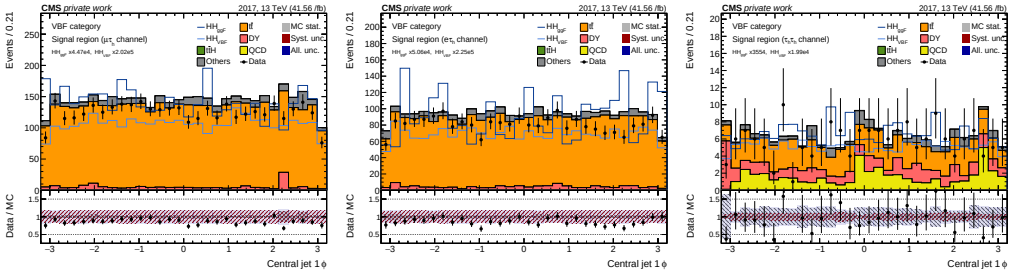


(k) `ctjet1_pt`

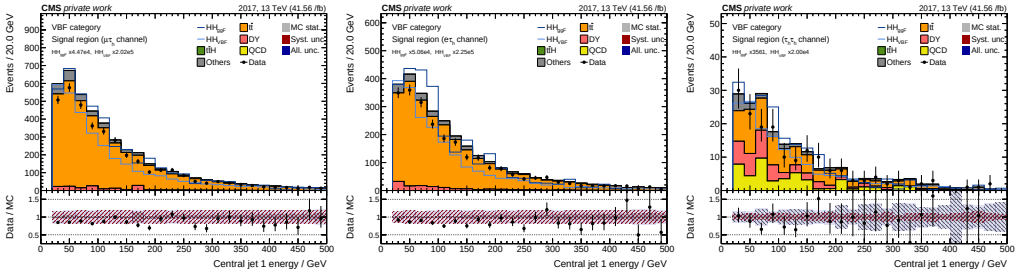


(n) `ctjet1_eta`

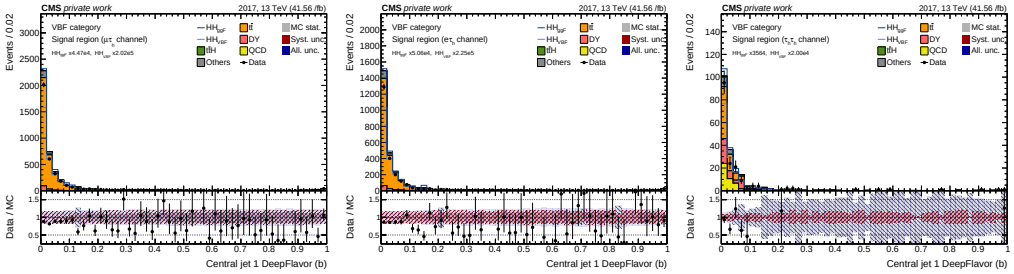
Figure A.19: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2017 data taking period.



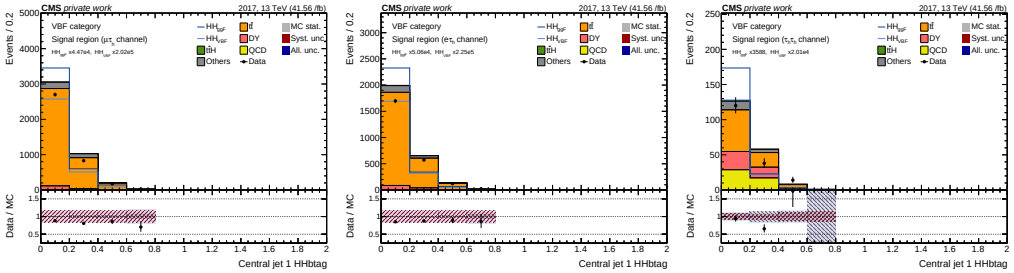
(b) `ctjet1_phi`



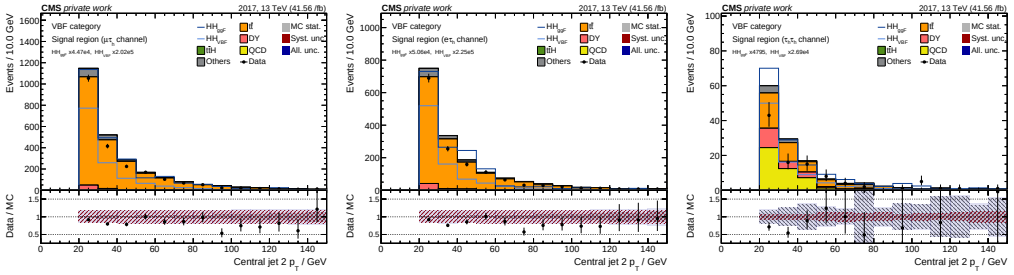
(e) `ctjet1_e`



(h) `ctjet1_deepflavor_b`

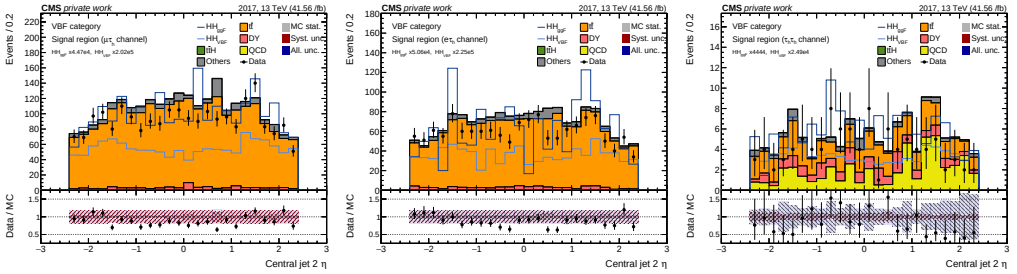


(k) `ctjet1_hhbtag`

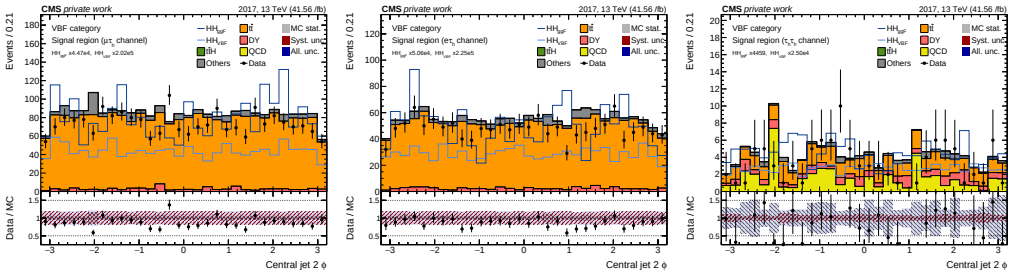


(n) `ctjet2_pt`

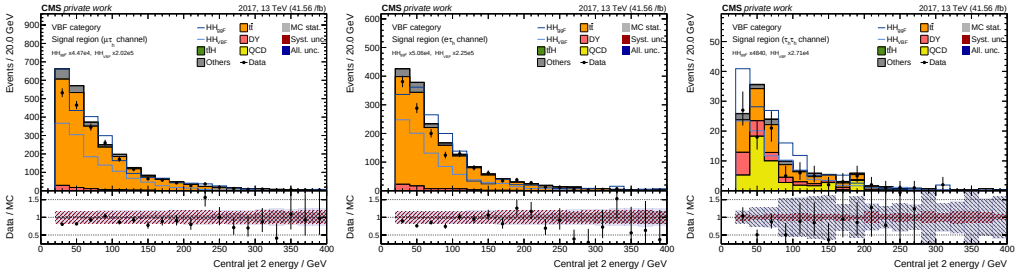
Figure A.20: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2017 data taking period.



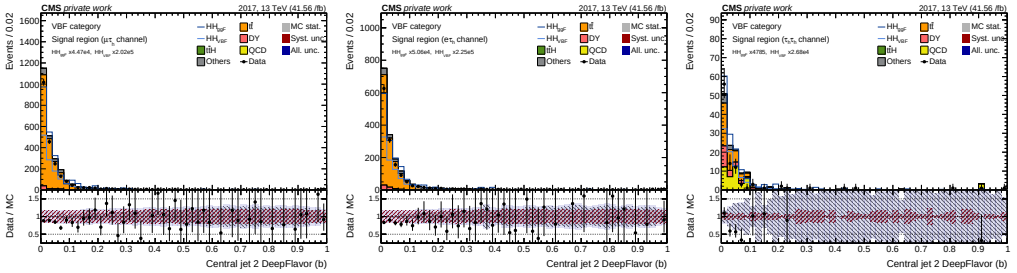
(b) `ctjet2_eta`



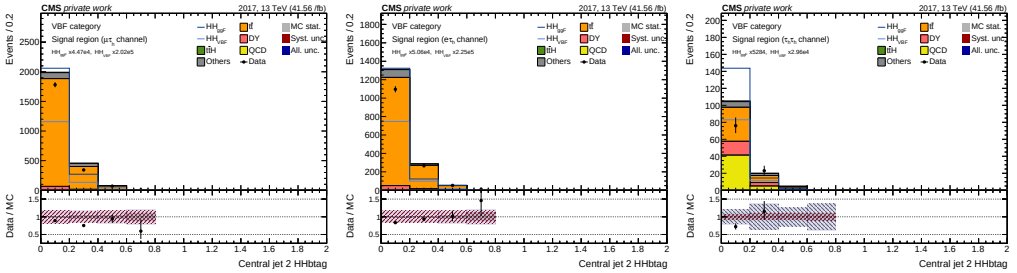
(e) `ctjet2_phi`



(h) `ctjet2_e`

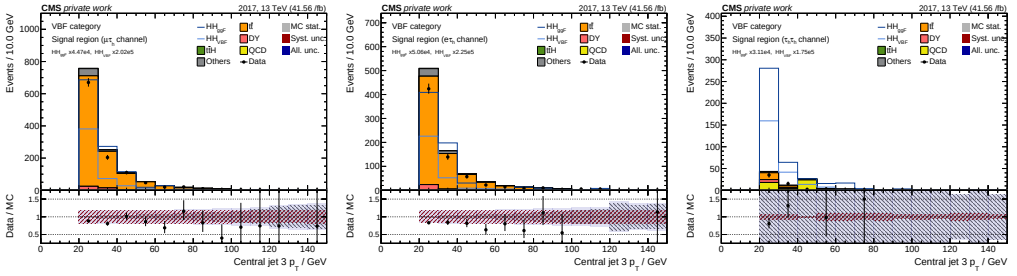


(k) `ctjet2_deepflavor_b`

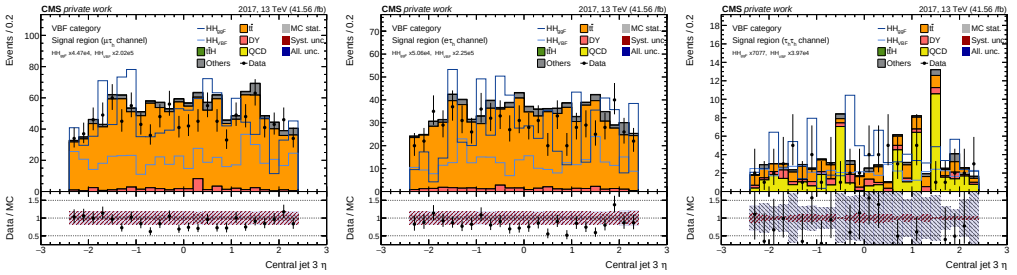


(n) `ctjet2_hhbtag`

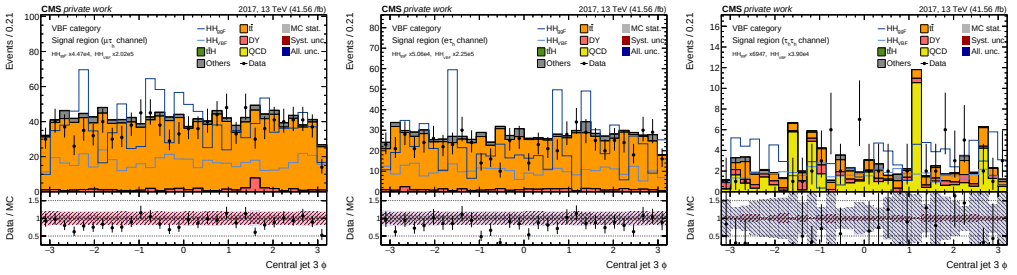
Figure A.21: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2017 data taking period.



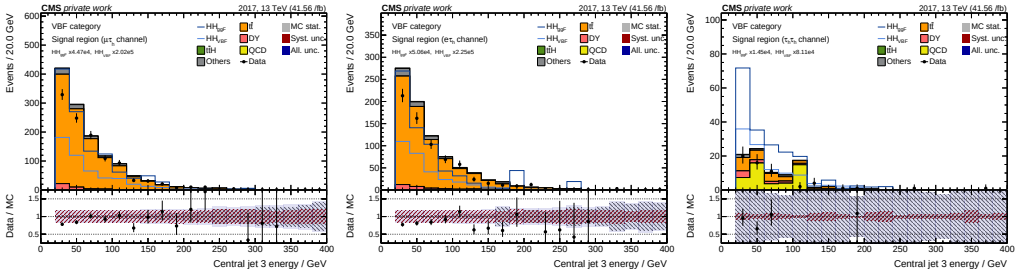
(b) `ctjet3_pt`



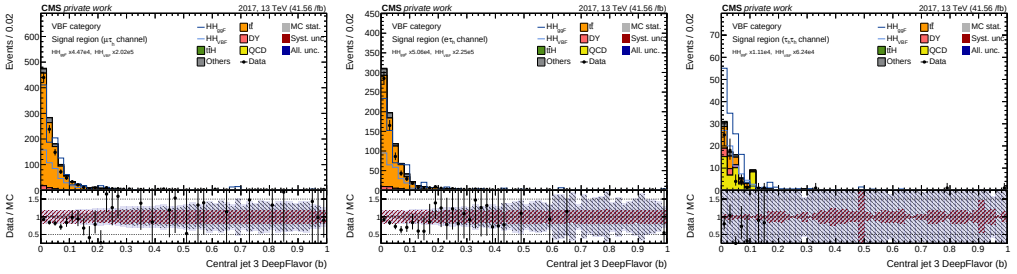
(e) `ctjet3_eta`



(h) `ctjet3_phi`

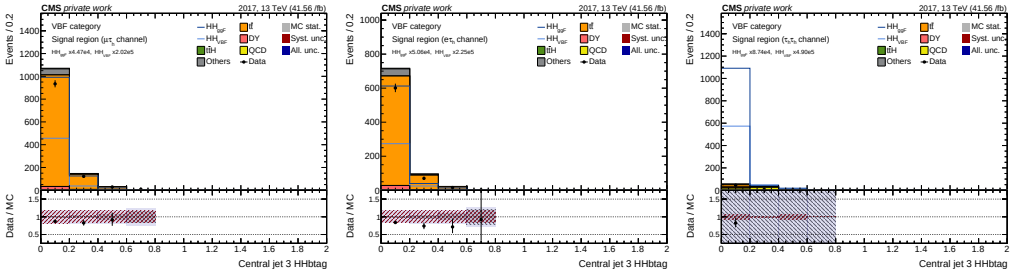


(k) `ctjet3_e`

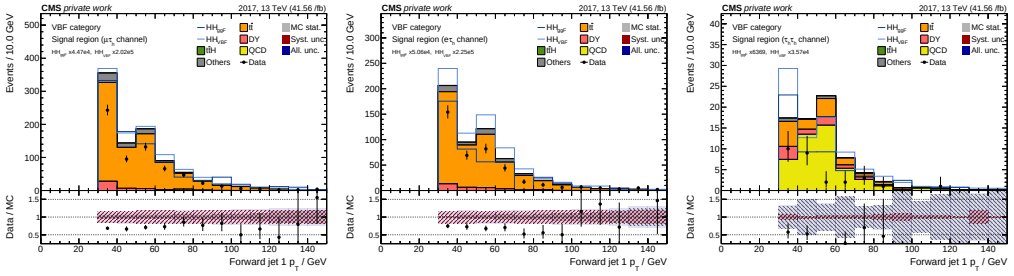


(n) `ctjet3_deepflavor_b`

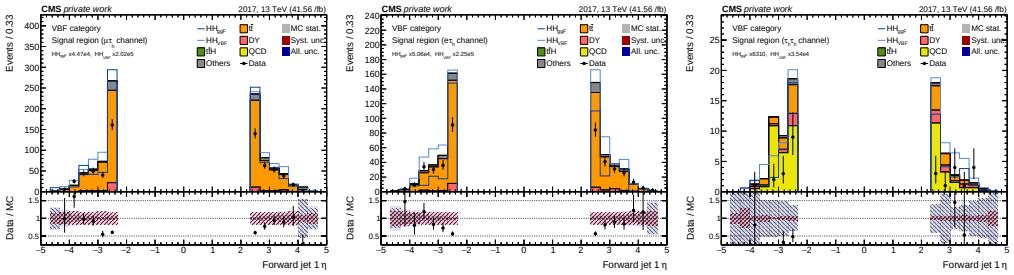
Figure A.22: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2017 data taking period.



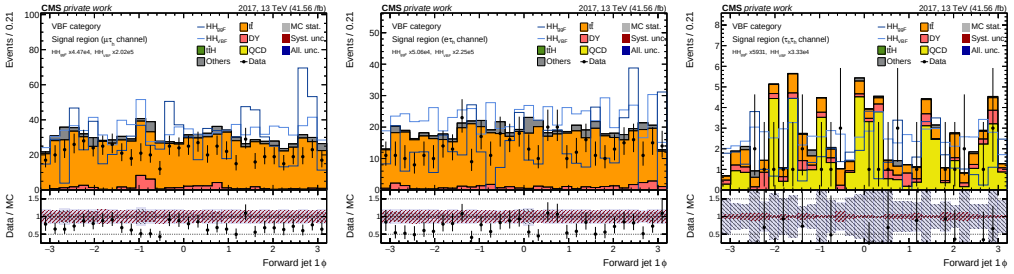
(b) `ctjet3_hhbtag`



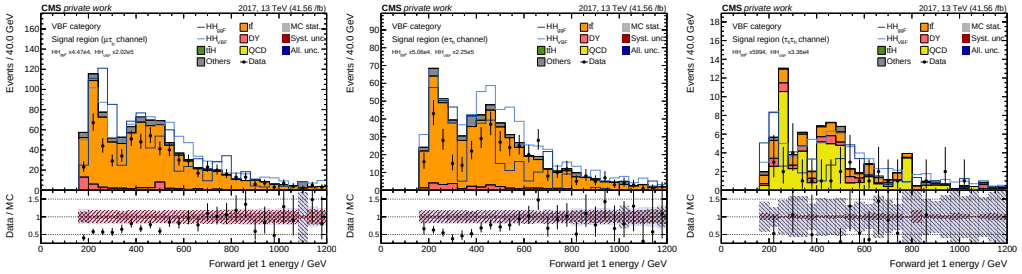
(e) `fwjet1_pt`



(h) `fwjet1_eta`

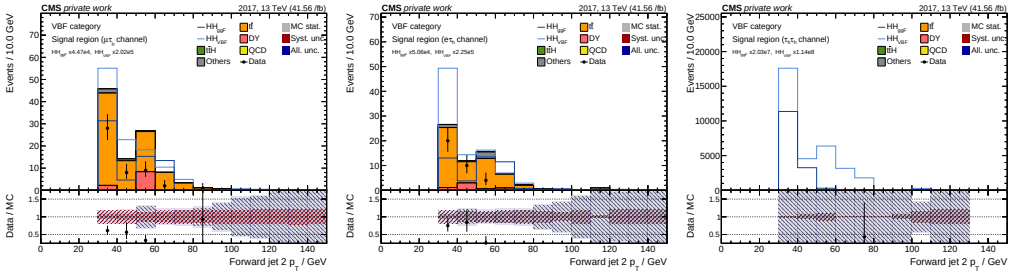


(k) `fwjet1_phi`

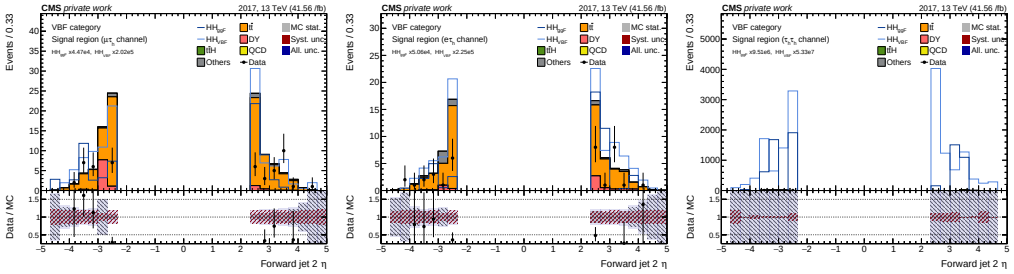


(n) `fwjet1_e`

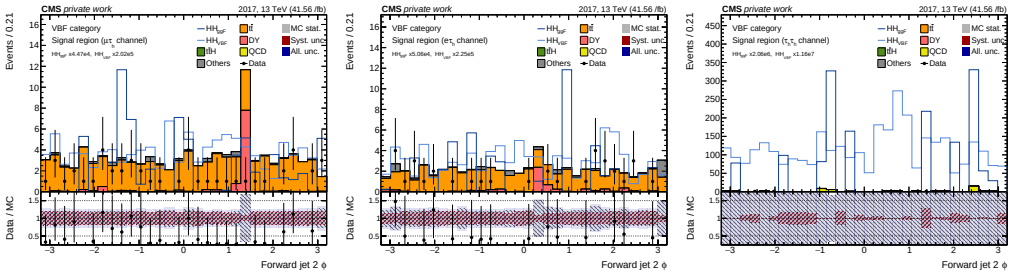
Figure A.23: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2017 data taking period.



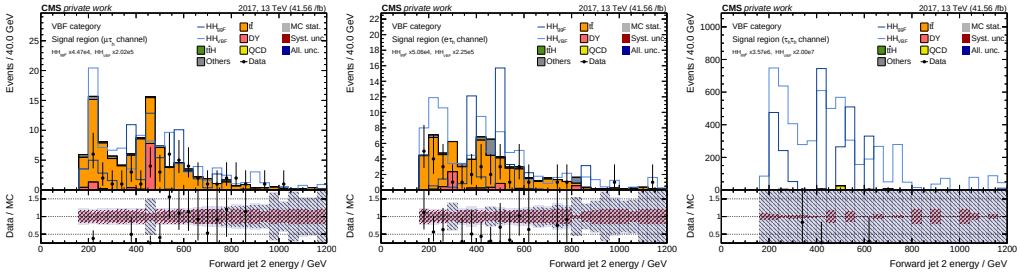
(b) `fwjet2_pt`



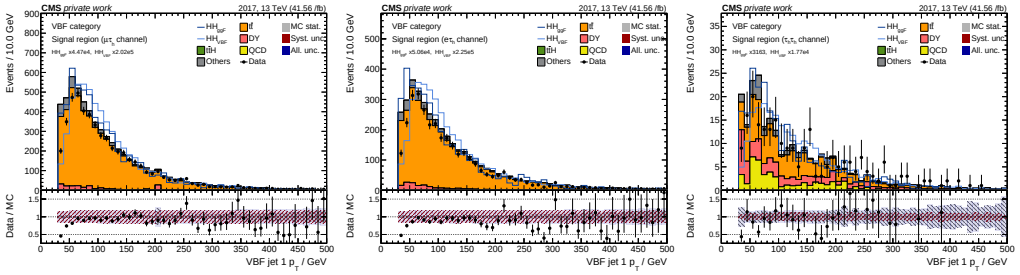
(e) `fwjet2_eta`



(h) `fwjet2_phi`

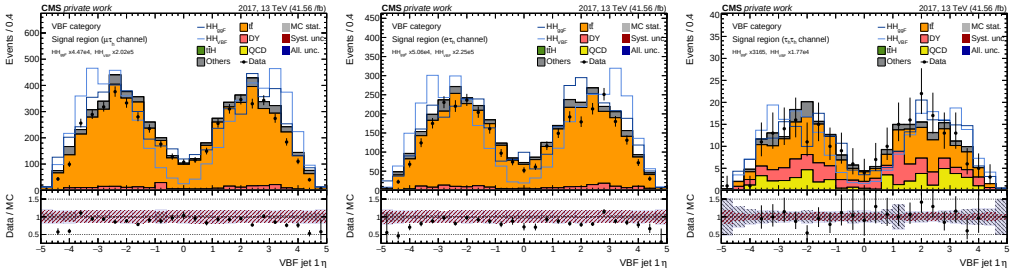


(k) `fwjet2_e`

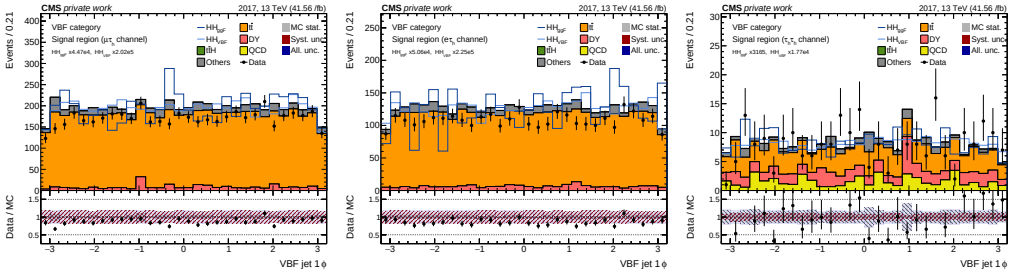


(n) `vbfjet1_pt`

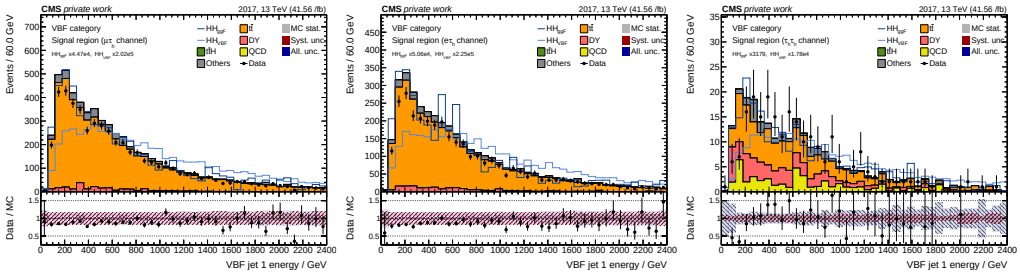
Figure A.24: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2017 data taking period.



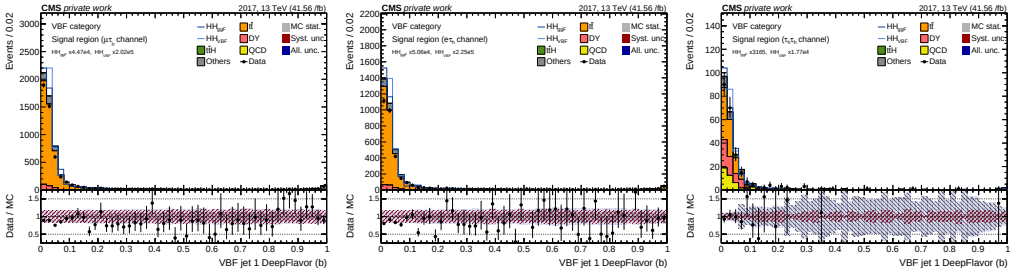
(b) `vbfjet1_eta`



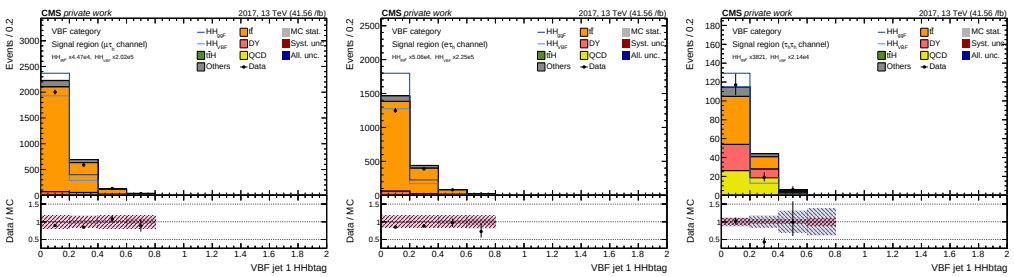
(e) `vbfjet1_phi`



(h) `vbfjet1_e`

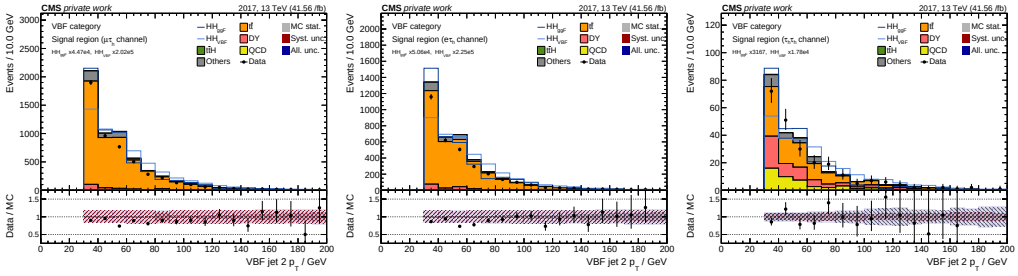


(k) `vbfjet1_deepflavor_b`

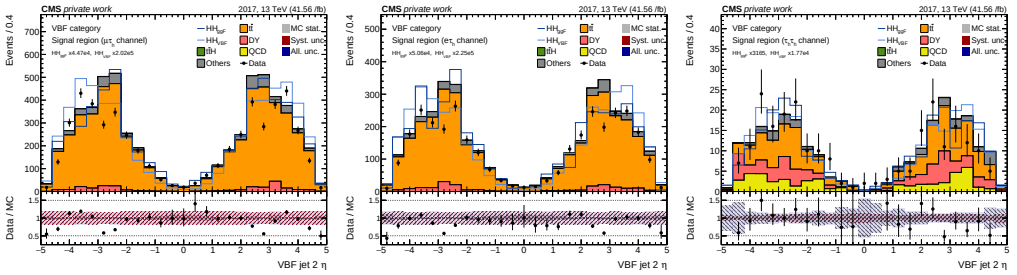


(n) `vbfjet1_hhbtag`

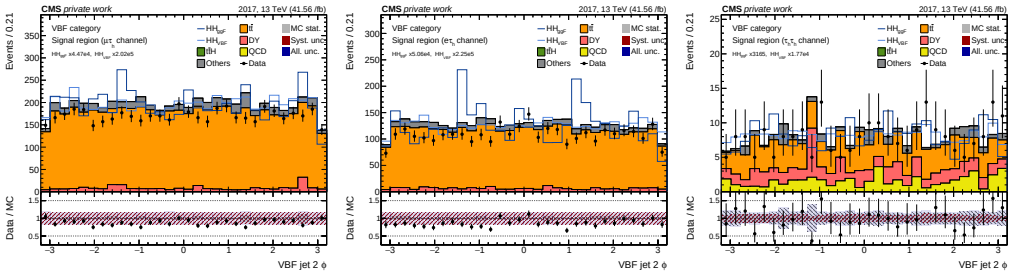
Figure A.25: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2017 data taking period.



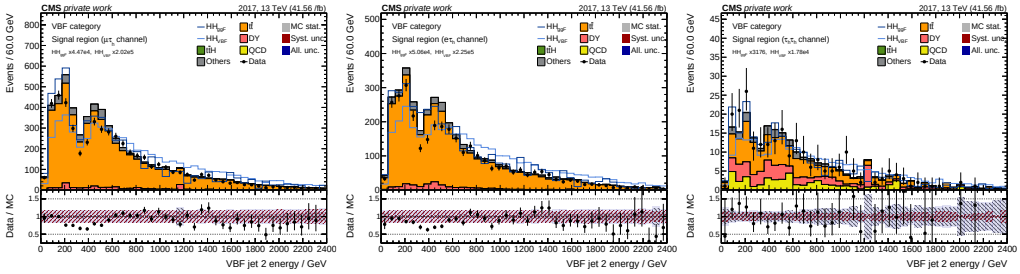
(b) `vbfjet2_pt`



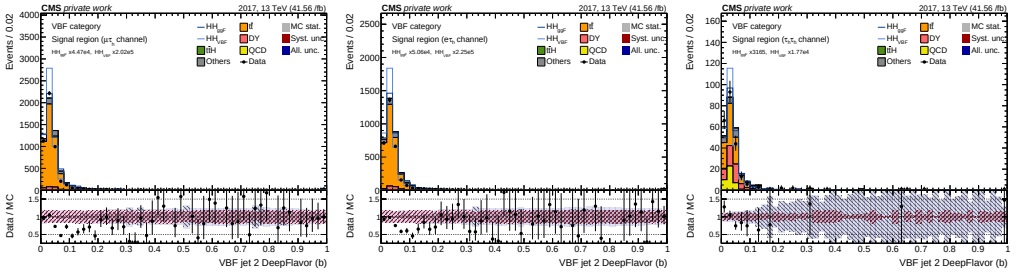
(e) `vbfjet2_eta`



(h) `vbfjet2_phi`

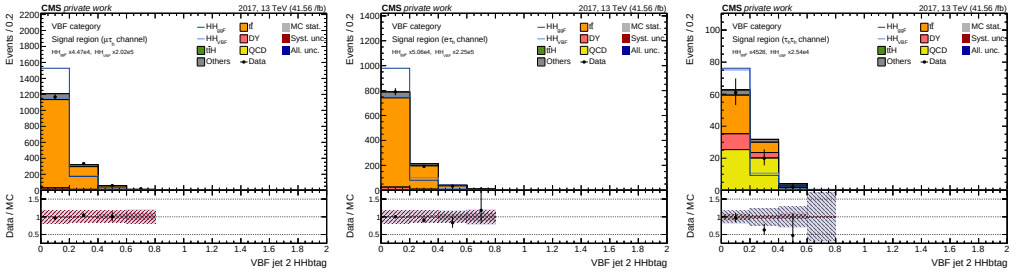


(k) `vbfjet2_e`

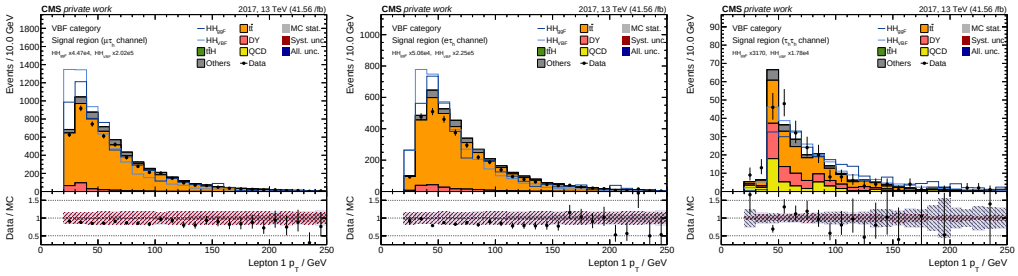


(n) `vbfjet2_deepflavor_b`

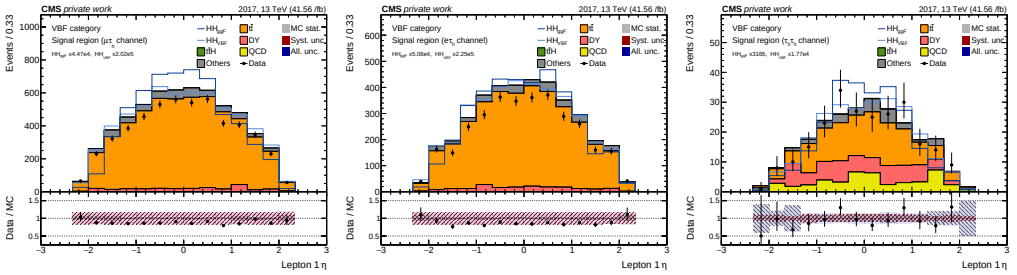
Figure A.26: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2017 data taking period.



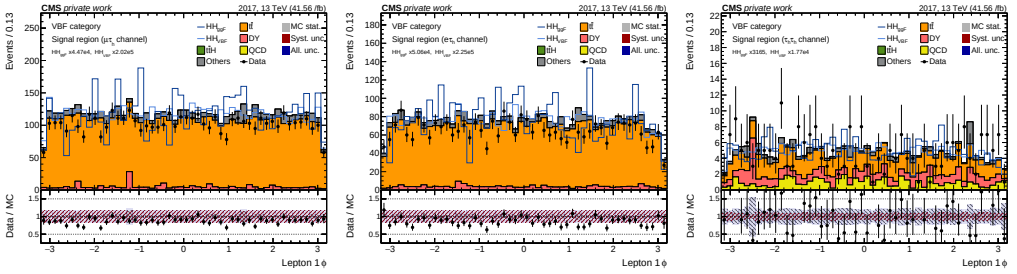
(b) `vbfjet2_hhbtag`



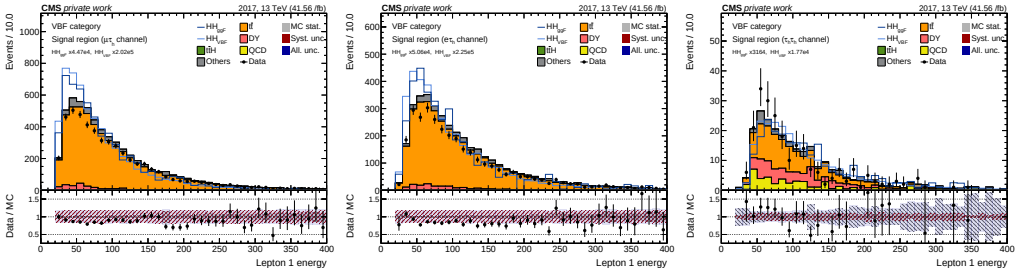
(e) `lep1_pt`



(h) `lep1_eta`

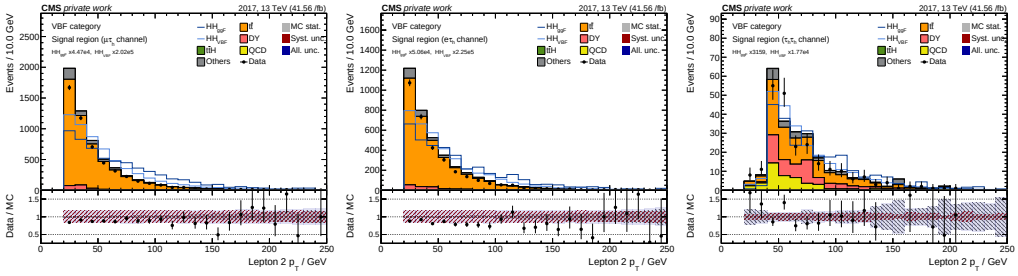


(k) `lep1_phi`

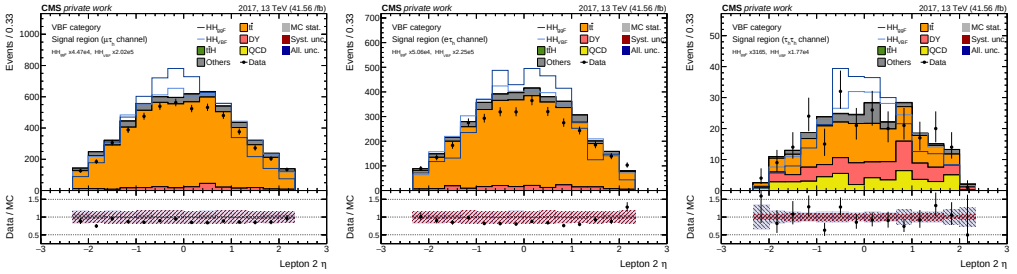


(n) `lep1_e`

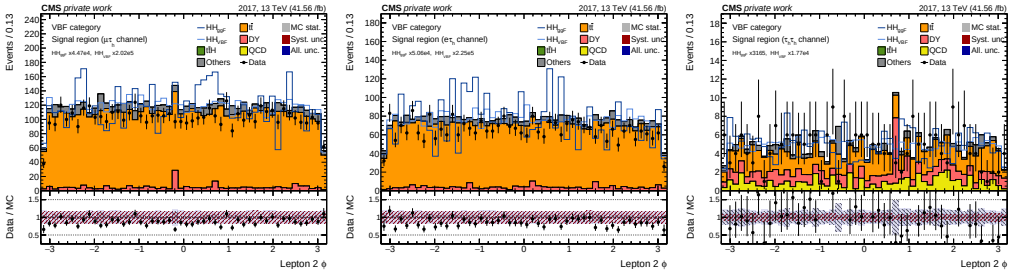
Figure A.27: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2017 data taking period.



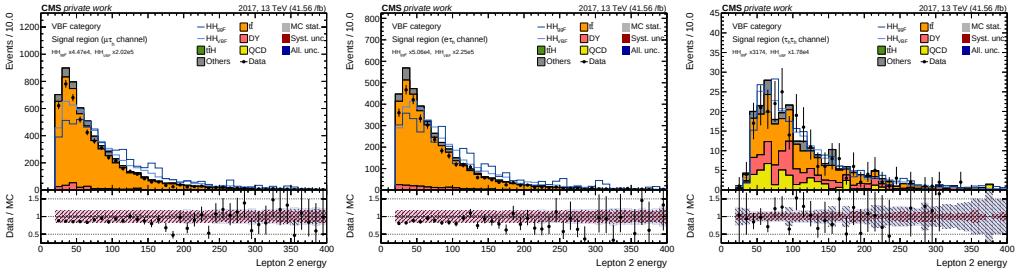
(b) `lep2_pt`



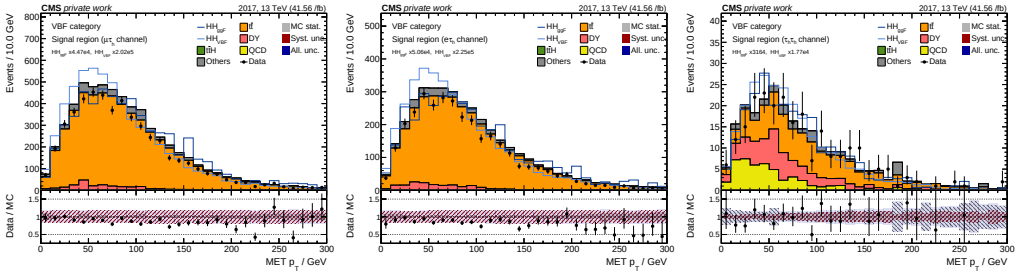
(e) `lep2_eta`



(h) `lep2_phi`

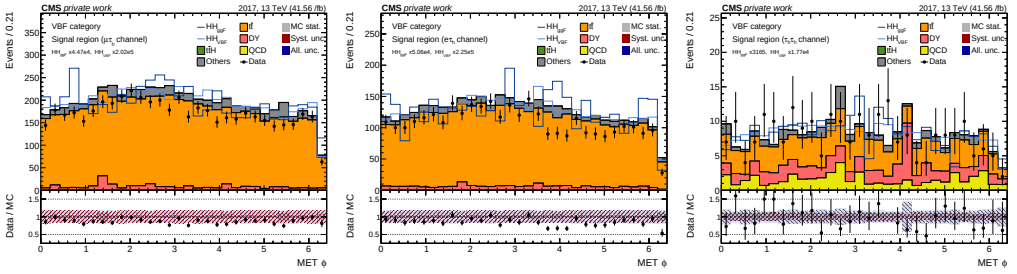


(k) `lep2_e`

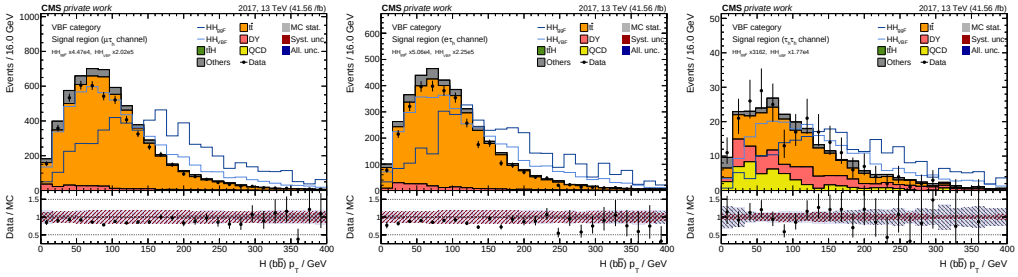


(n) `met_pt`

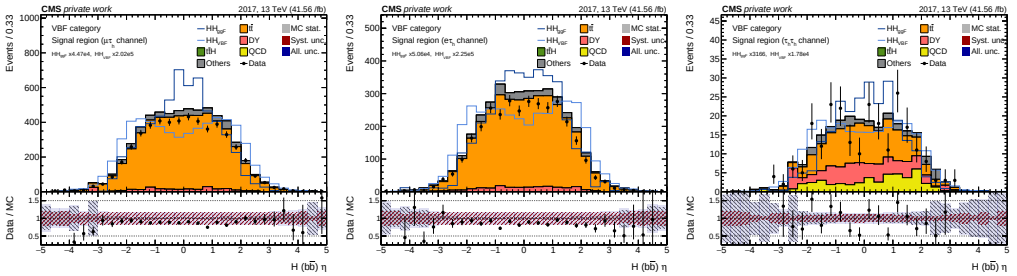
Figure A.28: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2017 data taking period.



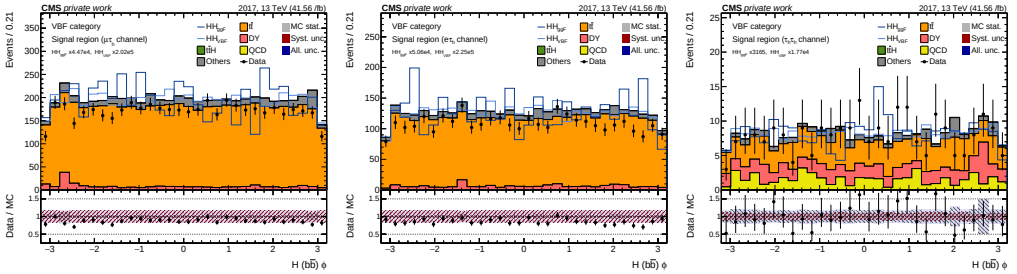
(b) `met_phi`



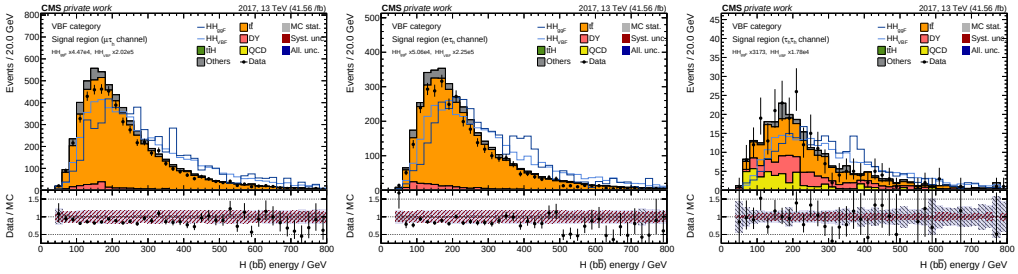
(e) `bh_pt`



(h) `bh_eta`

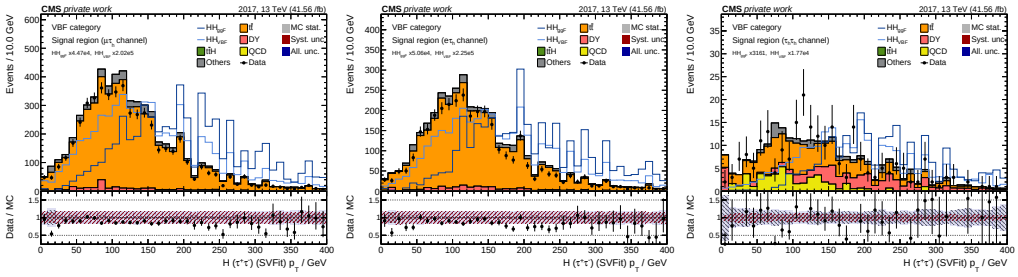


(k) `bh_phi`

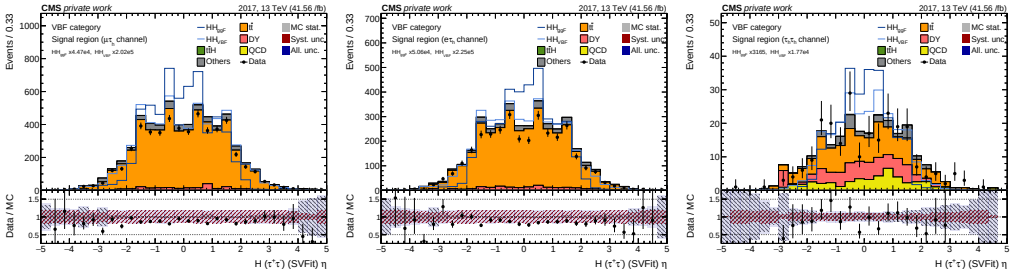


(n) `bh_e`

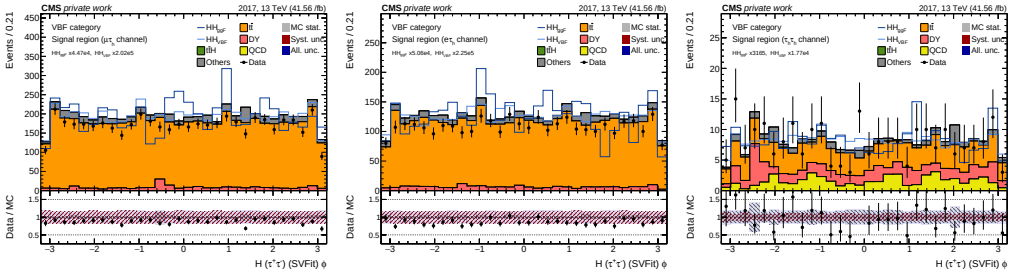
Figure A.29: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2017 data taking period.



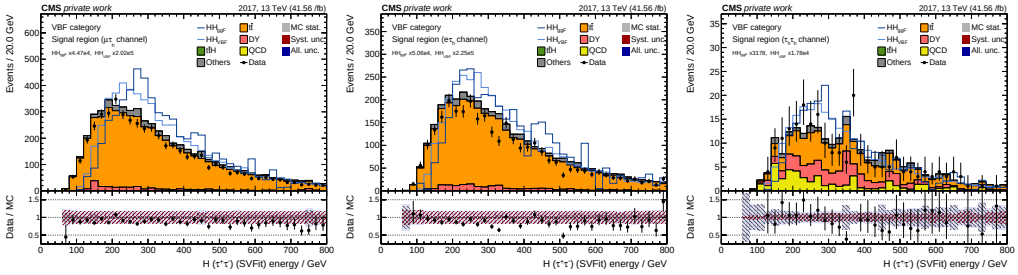
(b) `tauh_sv_pt`



(e) `tauh_sv_eta`

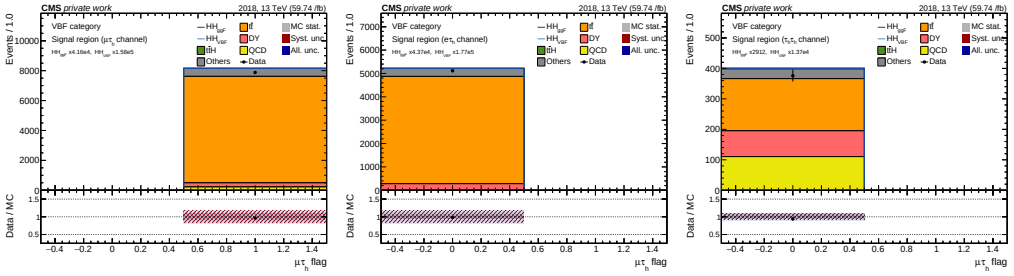


(h) `tauh_sv_phi`

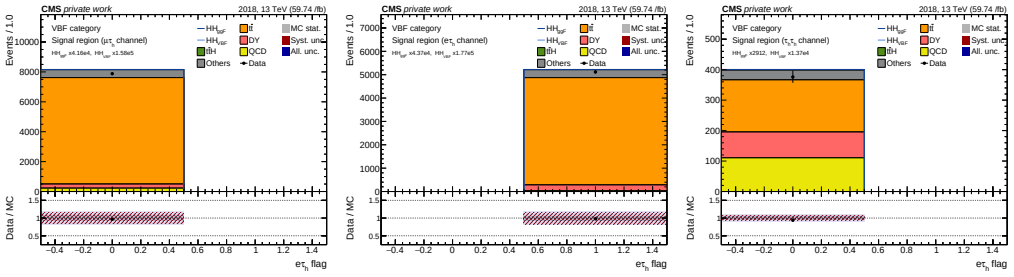


(k) `tauh_sv_e`

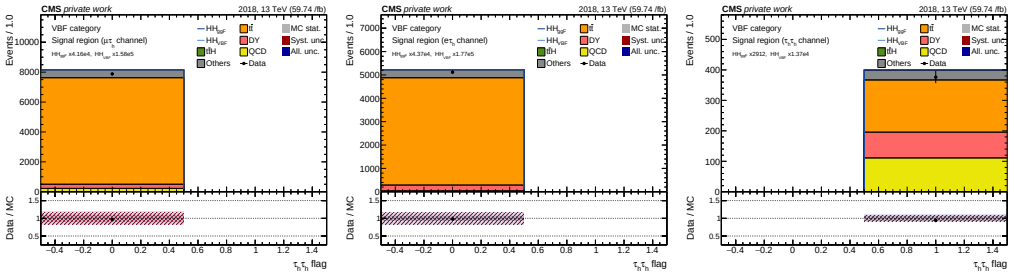
Figure A.30: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2017 data taking period.



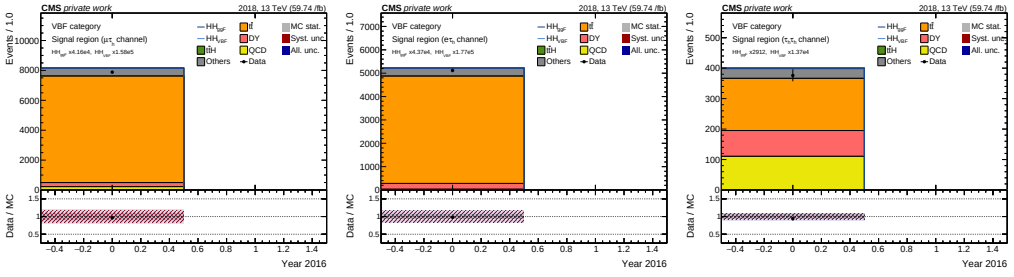
(b) `is_mutau`



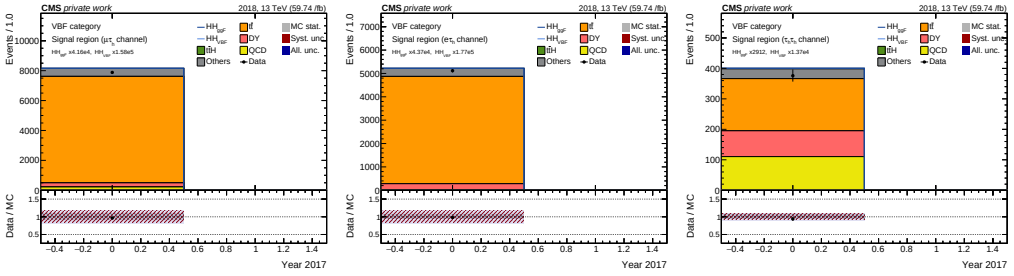
(e) `is_etau`



(h) `is_tautau`

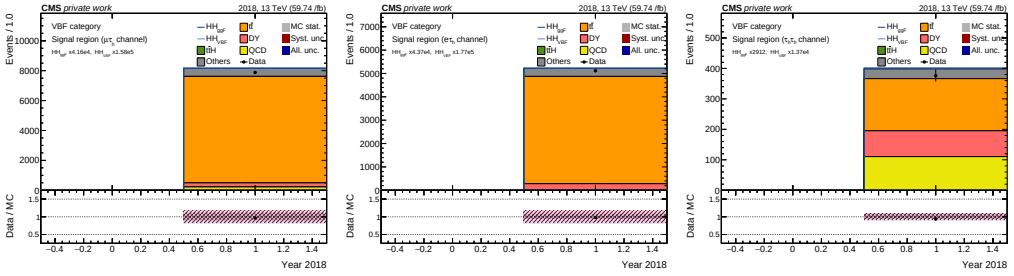


(k) `is_2016`

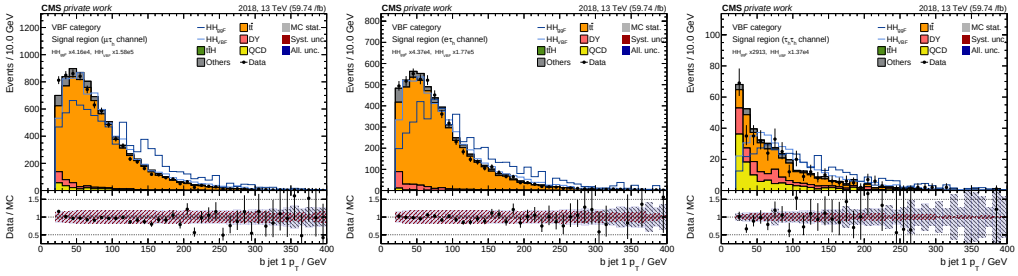


(n) `is_2017`

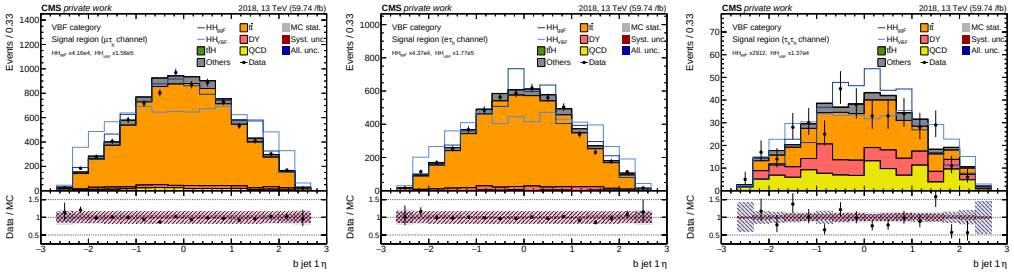
Figure A.31: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2018 data taking period.



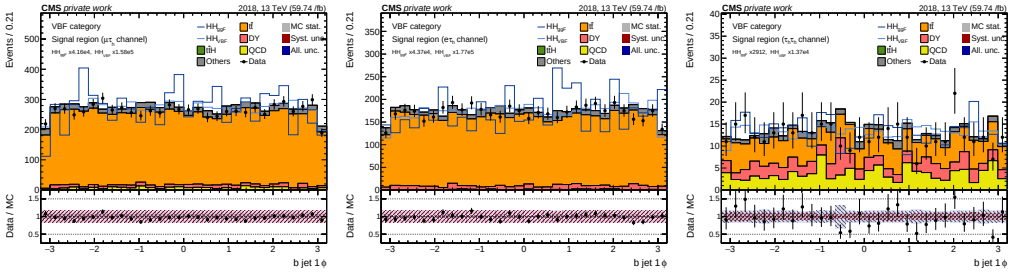
(b) `is_2018`



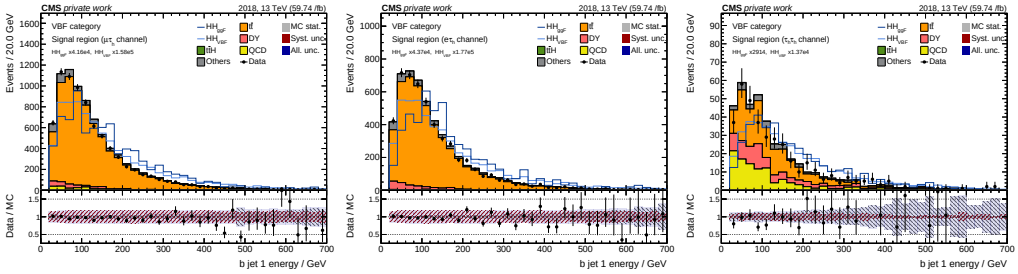
(e) `bjet1_pt`



(h) `bjet1_eta`

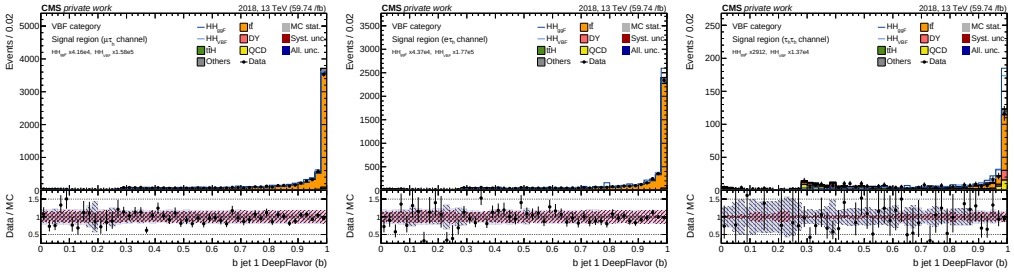


(k) `bjet1_phi`

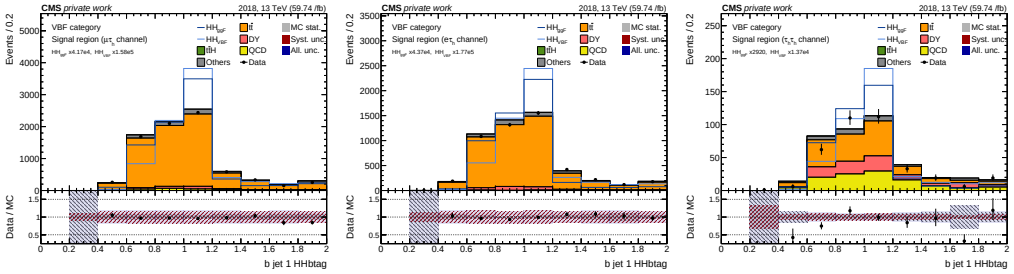


(n) `bjet1_e`

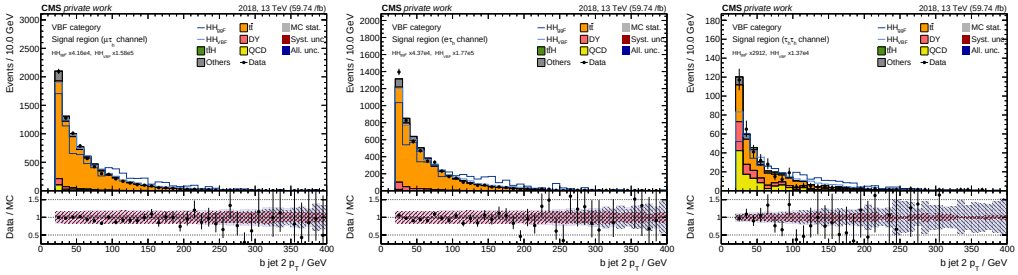
Figure A.32: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2018 data taking period.



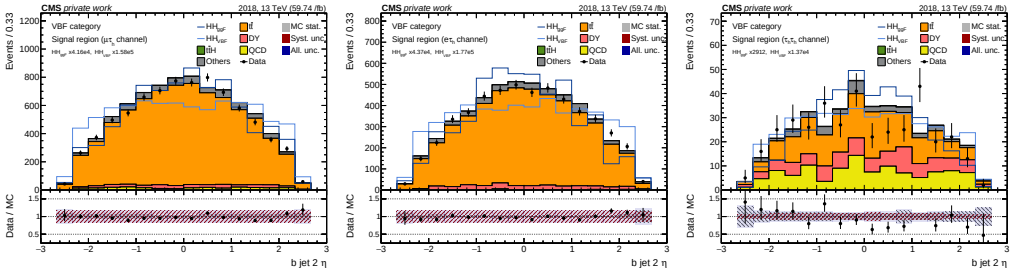
(b) `bjet1_deepflavor_b`



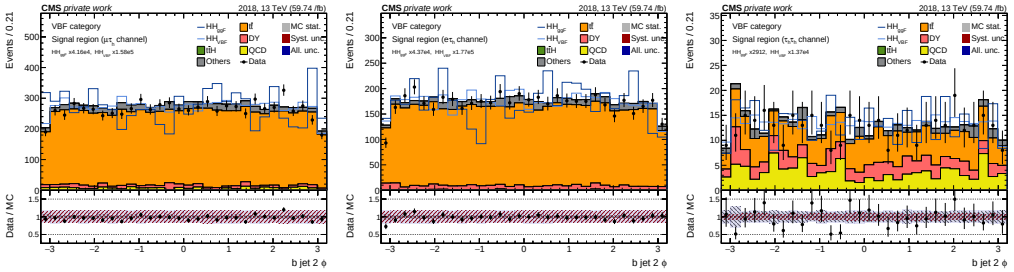
(e) `bjet1_hhbtag`



(h) `bjet2_pt`

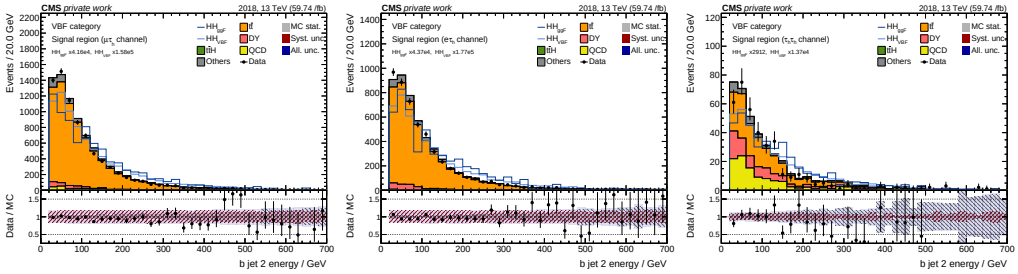


(k) `bjet2_eta`

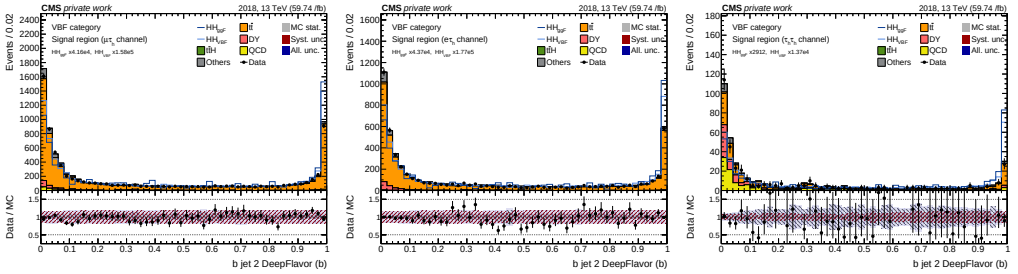


(n) `bjet2_phi`

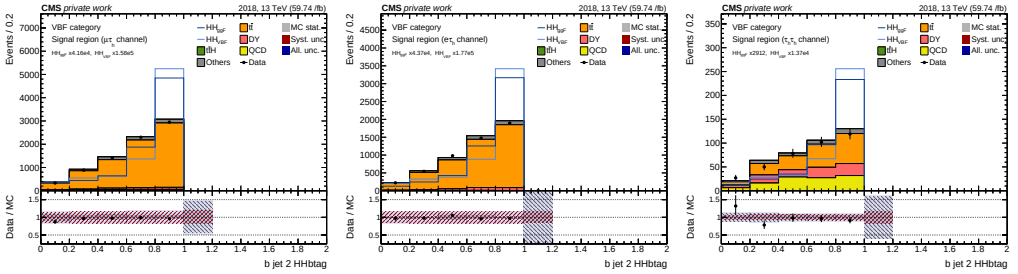
Figure A.33: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2018 data taking period.



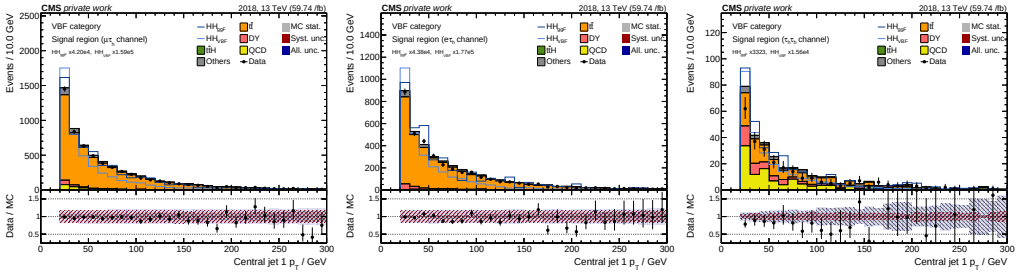
(b) `bjet2_e`



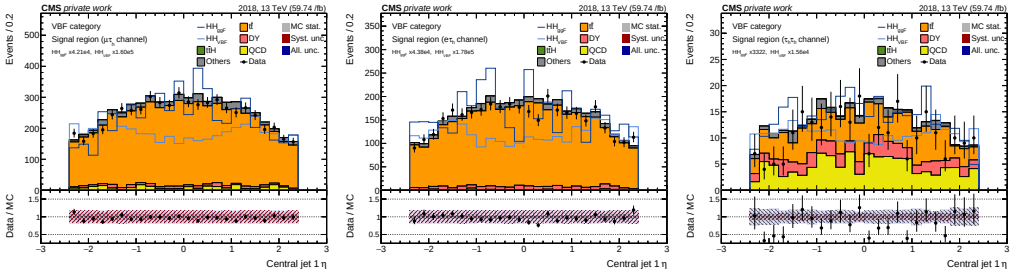
(e) `bjet2_deepflavor_b`



(h) `bjet2_hhbtag`

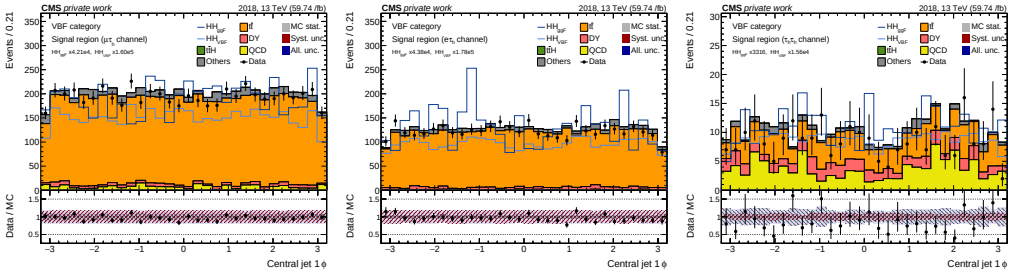


(k) `ctjet1_pt`

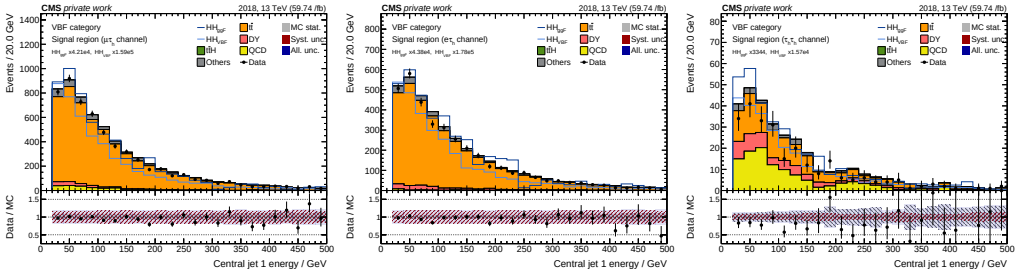


(n) `ctjet1_eta`

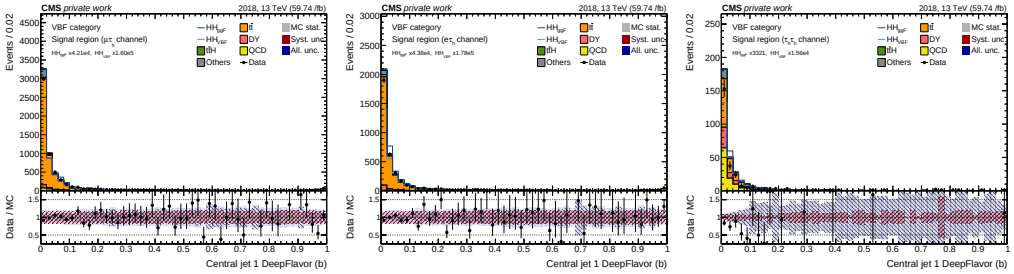
Figure A.34: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2018 data taking period.



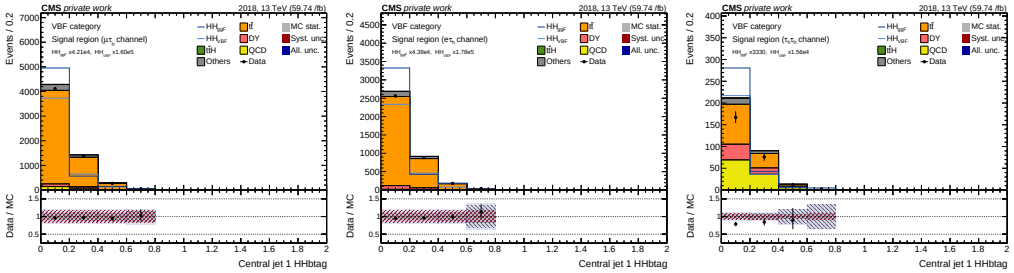
(b) `ctjet1_phi`



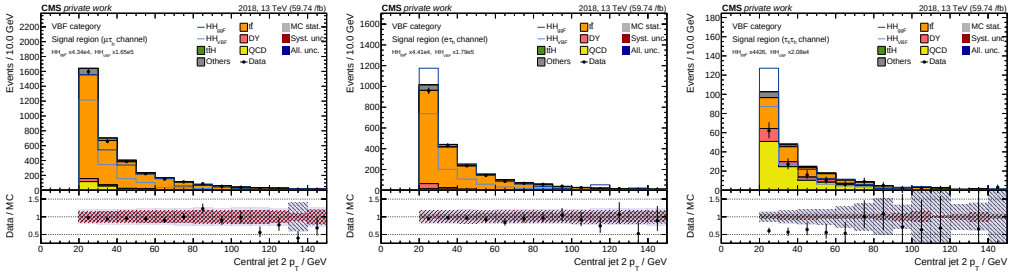
(e) `ctjet1_e`



(h) `ctjet1_deepflavor_b`

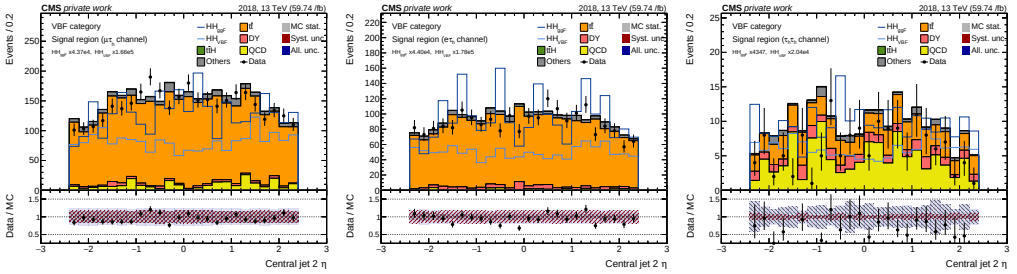


(k) `ctjet1_hhbtag`

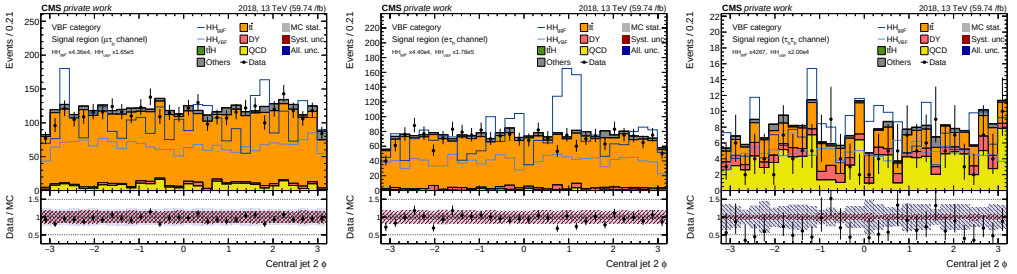


(n) `ctjet2_pt`

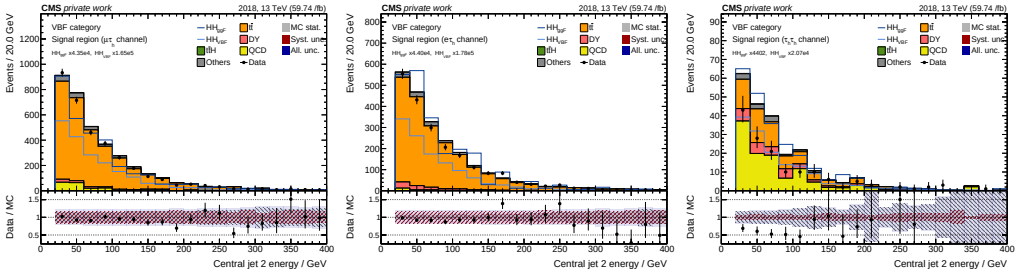
Figure A.35: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2018 data taking period.



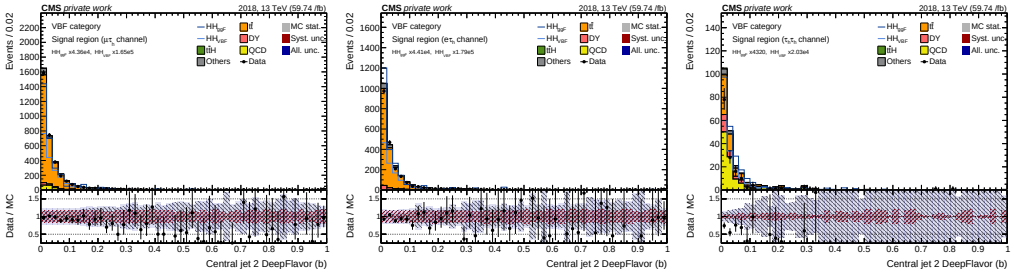
(b) `ctjet2_eta`



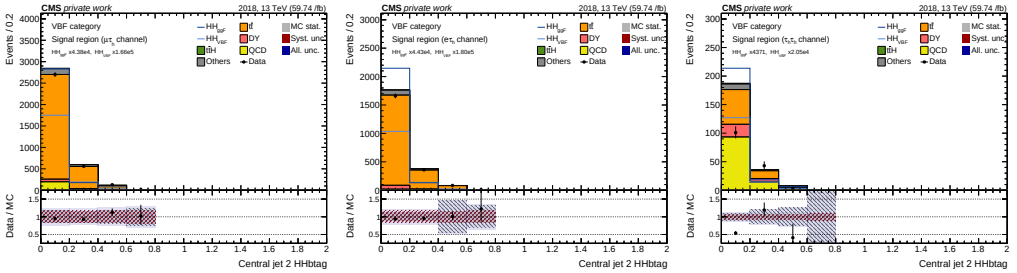
(e) `ctjet2_phi`



(h) `ctjet2_e`

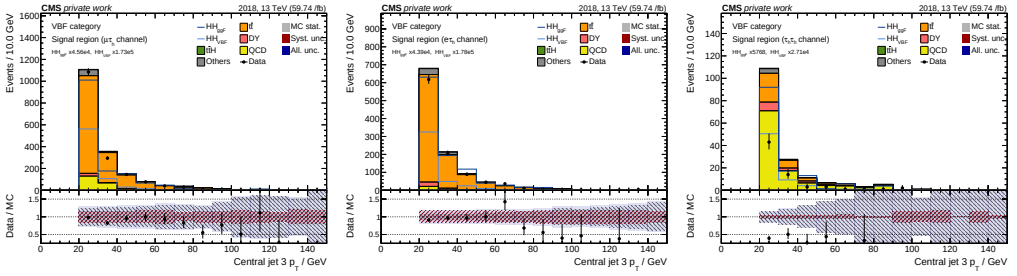


(k) `ctjet2_deepflavor_b`

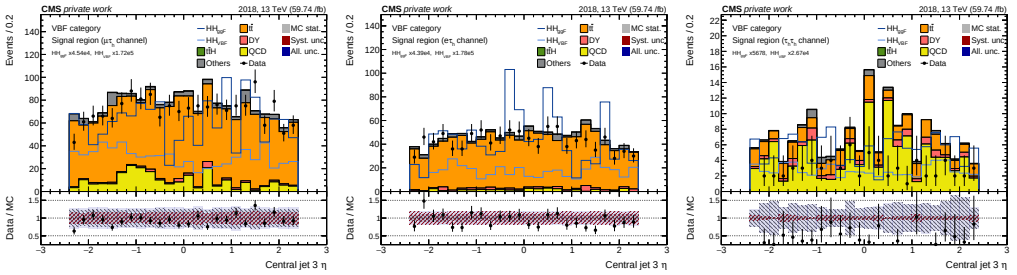


(n) `ctjet2_hhbtag`

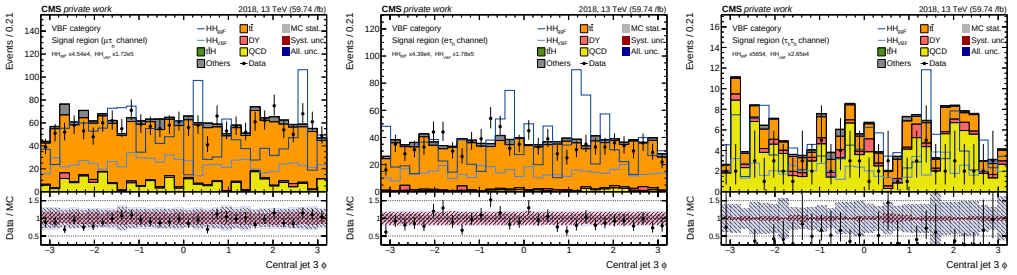
Figure A.36: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2018 data taking period.



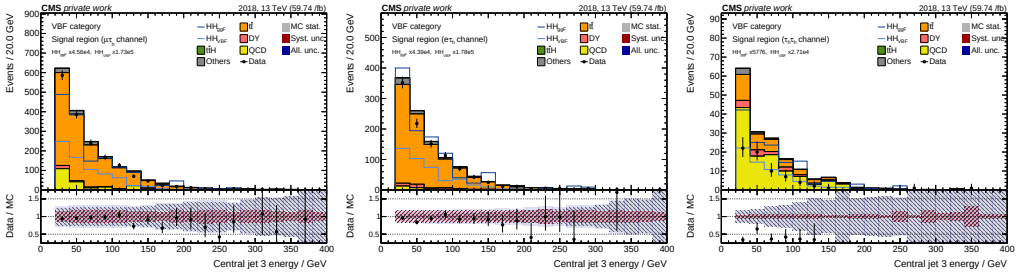
(b) `ctjet3_pt`



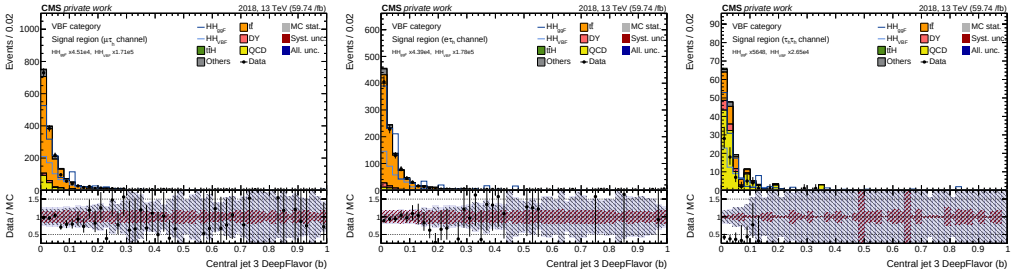
(e) `ctjet3_eta`



(h) `ctjet3_phi`

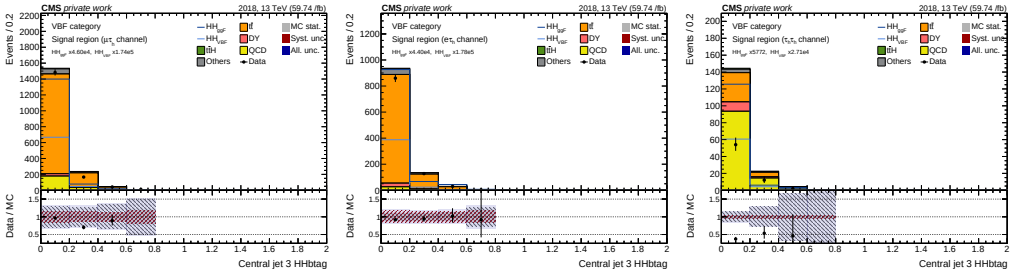


(k) `ctjet3_e`

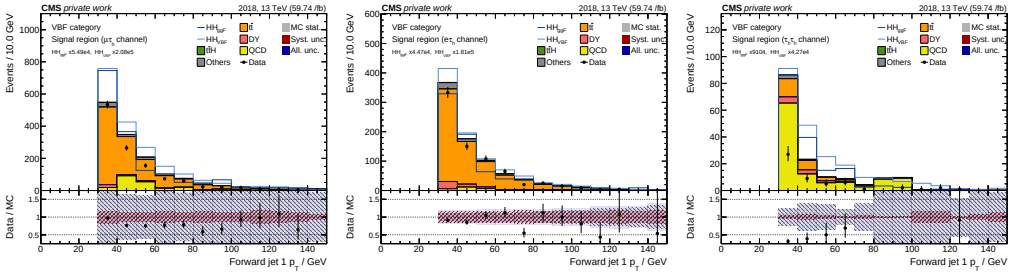


(n) `ctjet3_deepflavor_b`

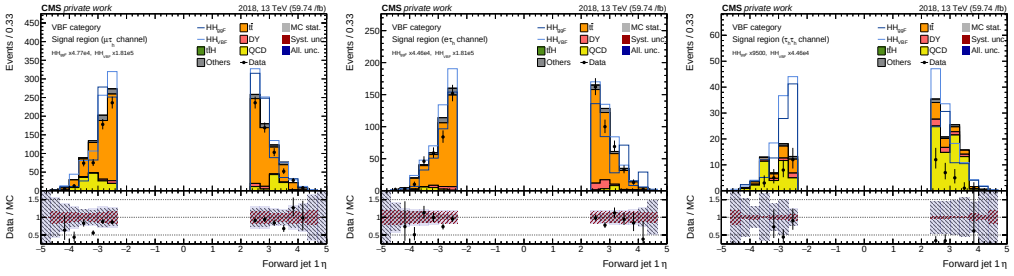
Figure A.37: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2018 data taking period.



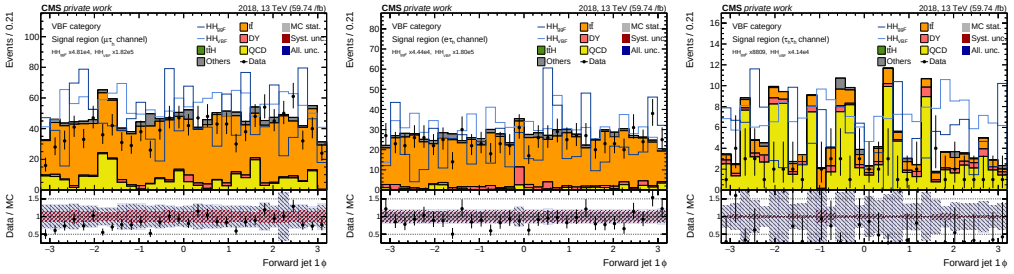
(b) `ctjet3_hhbtag`



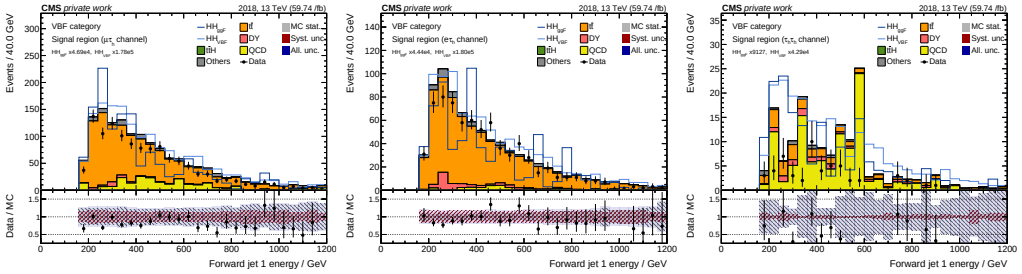
(e) `fwjet1_pt`



(h) `fwjet1_eta`

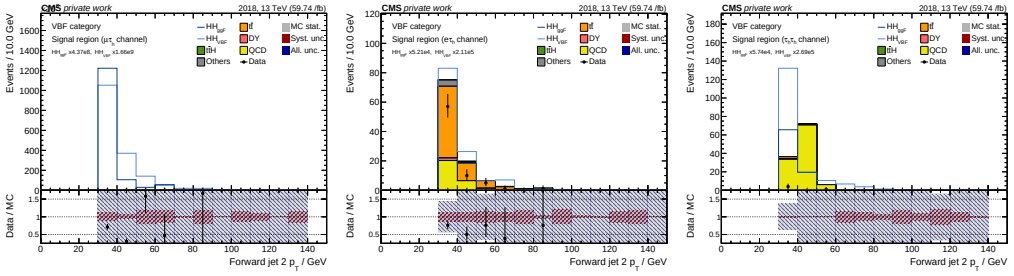


(k) `fwjet1_phi`

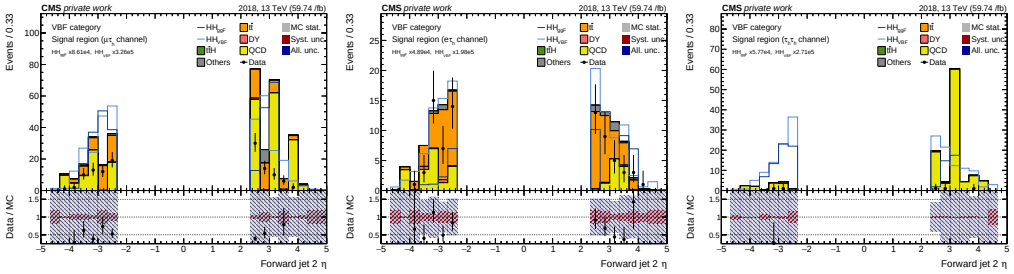


(n) `fwjet1_e`

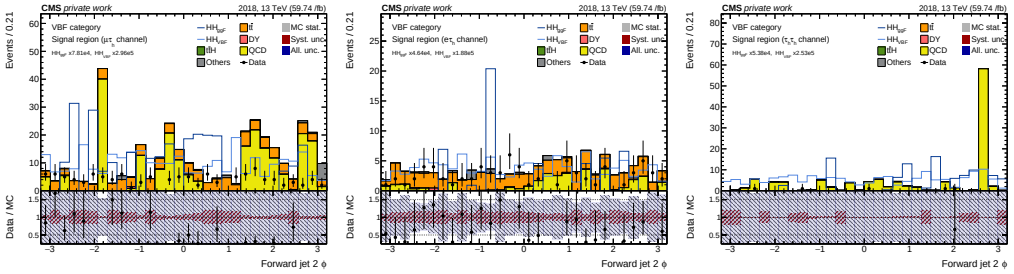
Figure A.38: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2018 data taking period.



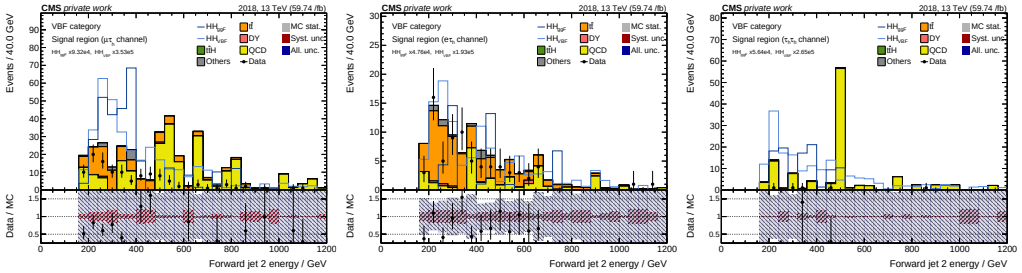
(b) `fwjet2_pt`



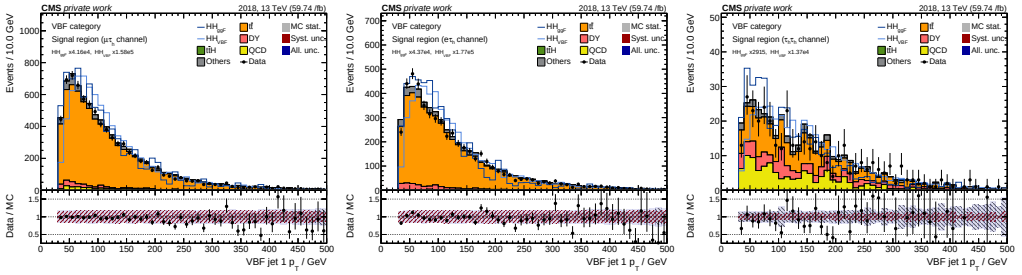
(e) `fwjet2_eta`



(h) `fwjet2_phi`

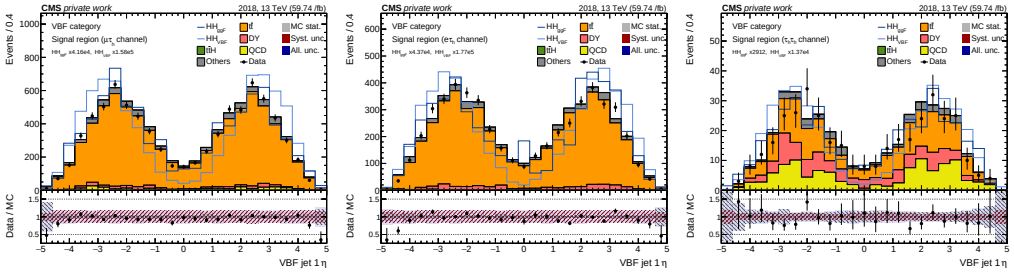


(k) `fwjet2_e`

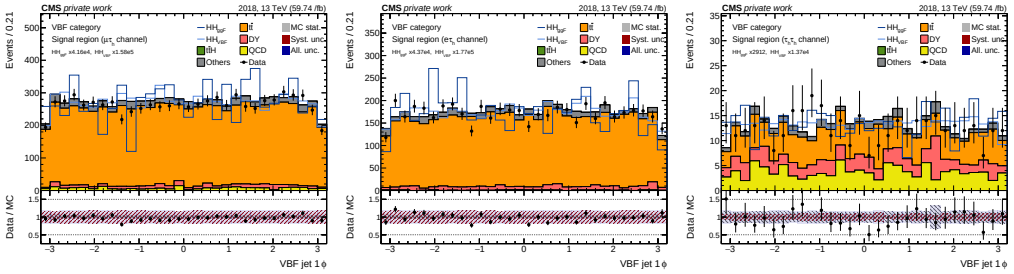


(n) `vbfjet1_pt`

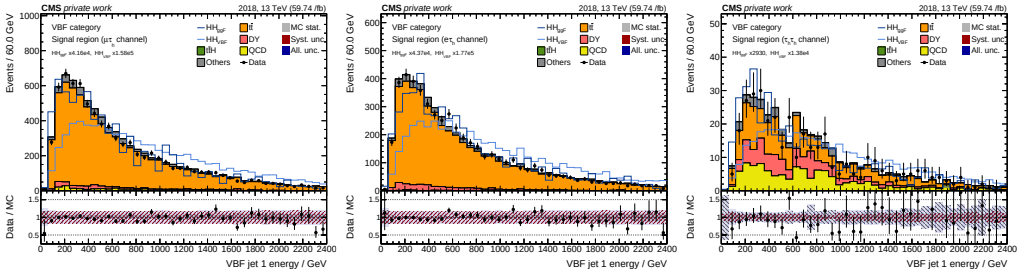
Figure A.39: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2018 data taking period.



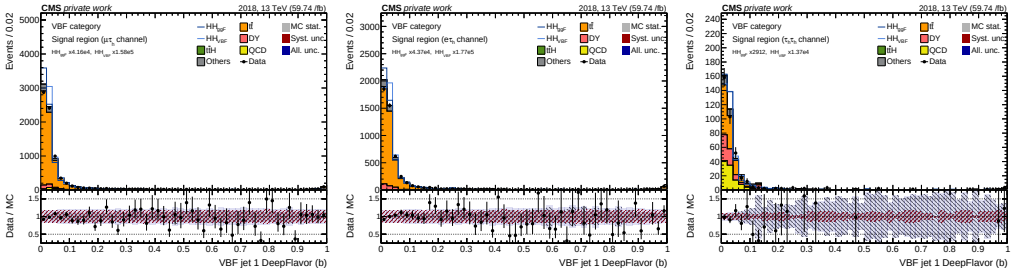
(b) `vbfjet1_eta`



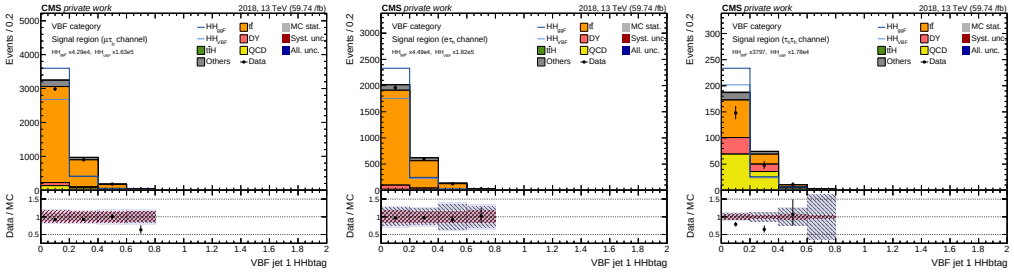
(e) `vbfjet1_phi`



(h) `vbfjet1_e`

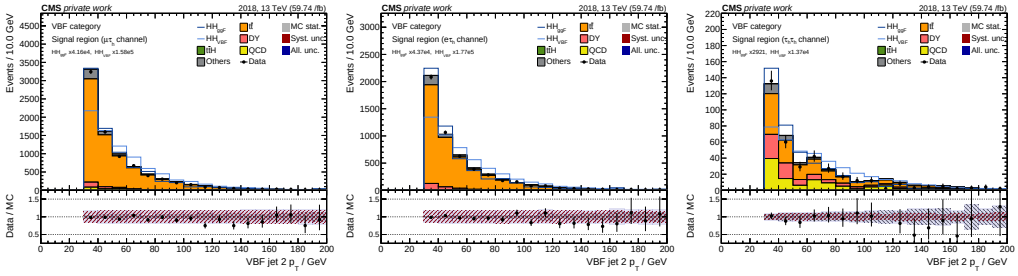


(k) `vbfjet1_deepflavor_b`

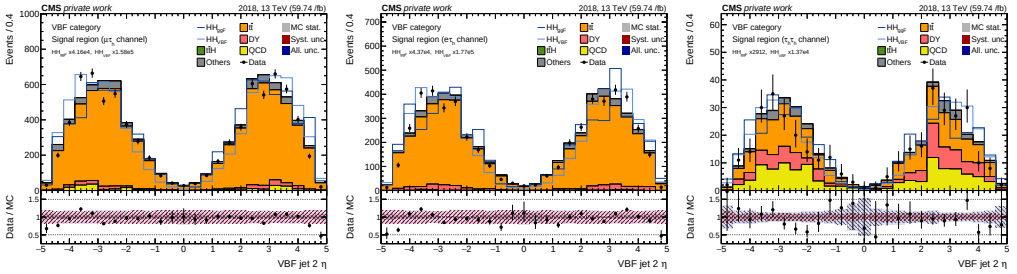


(n) `vbfjet1_hhbtag`

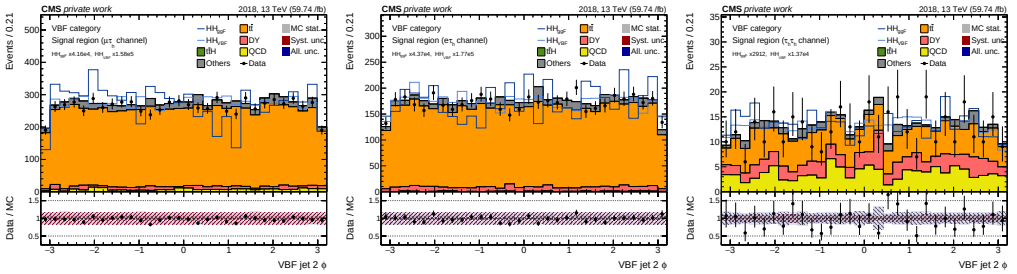
Figure A.40: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2018 data taking period.



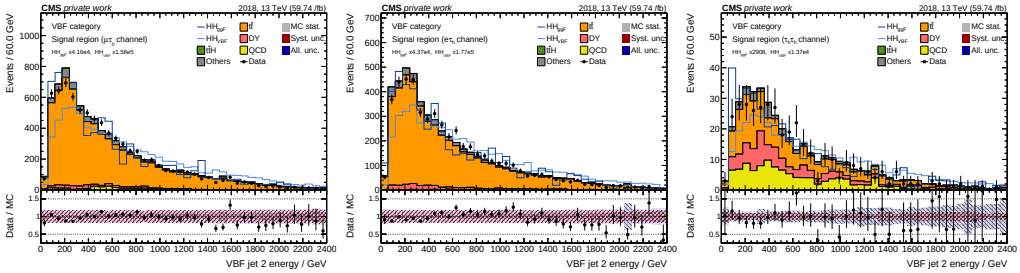
(b) `vbfjet2_pt`



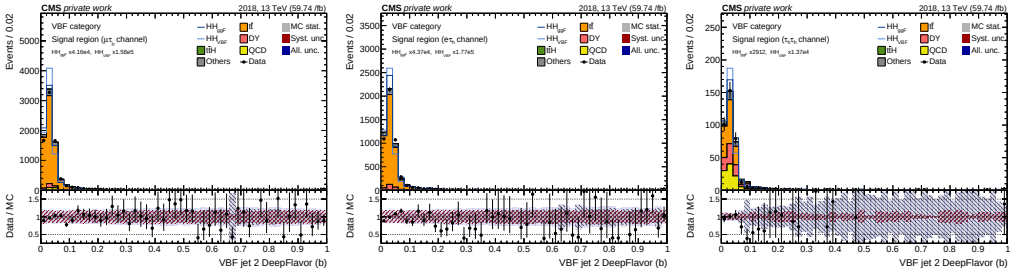
(e) `vbfjet2_eta`



(h) `vbfjet2_phi`

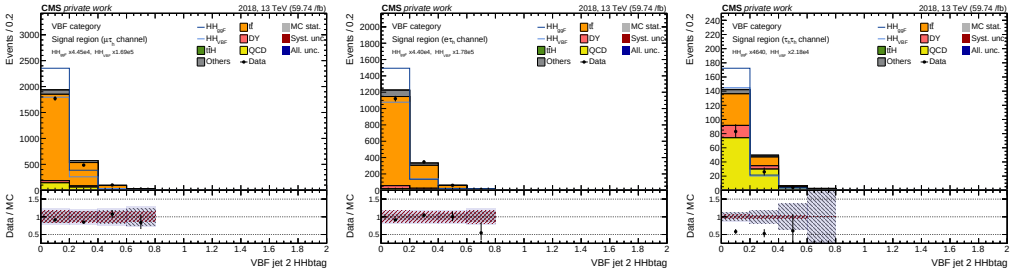


(k) `vbfjet2_e`

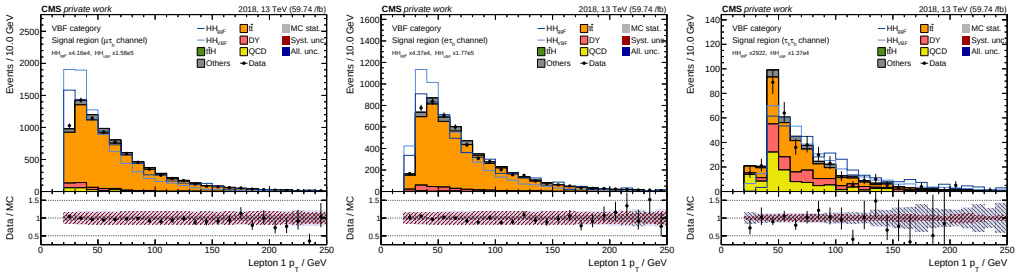


(n) `vbfjet2_deepflavor_b`

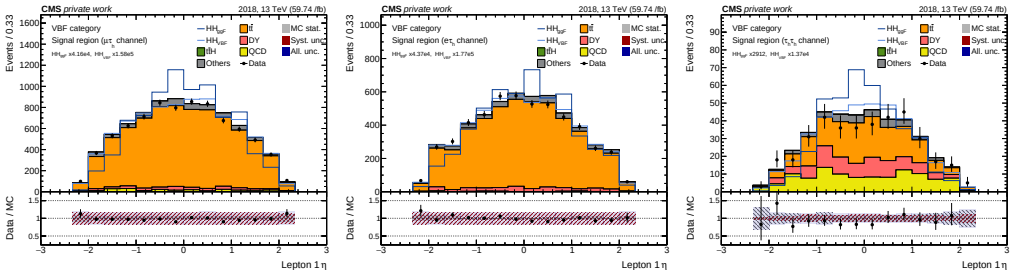
Figure A.41: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2018 data taking period.



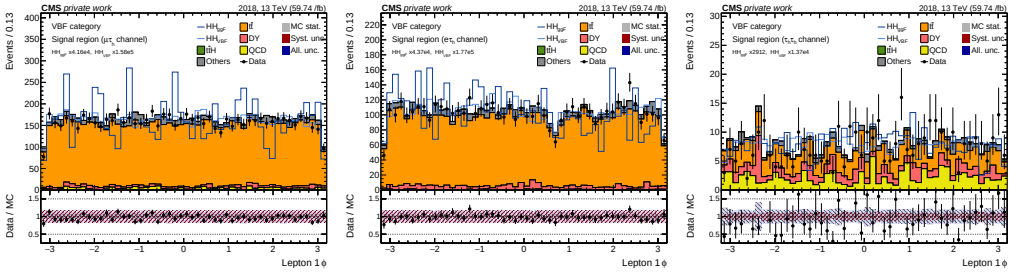
(b) `vbfjet2_hhbtag`



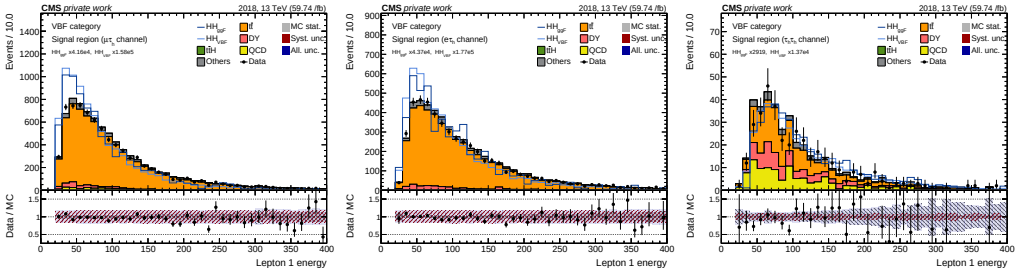
(e) `lep1_pt`



(h) `lep1_eta`

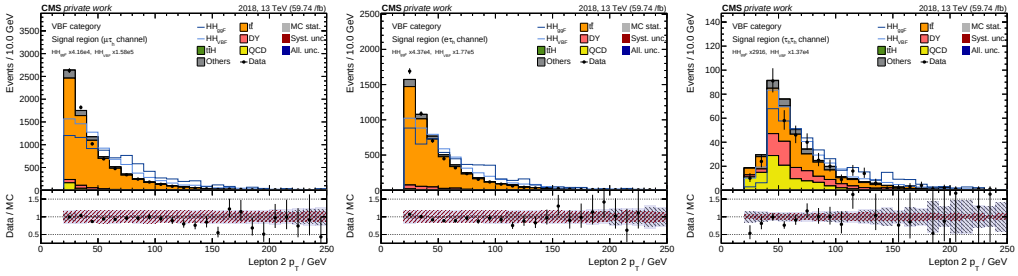


(k) `lep1_phi`

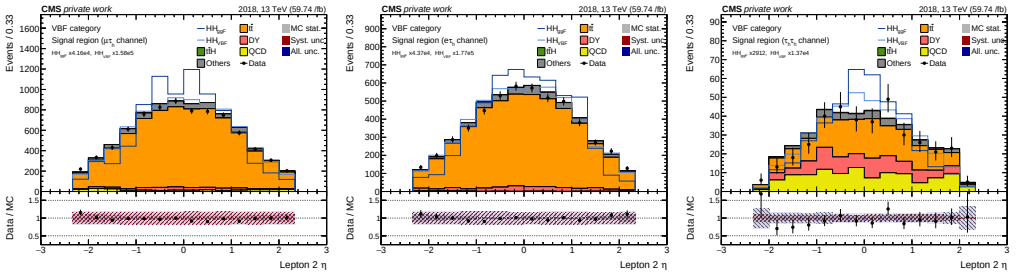


(n) `lep1_e`

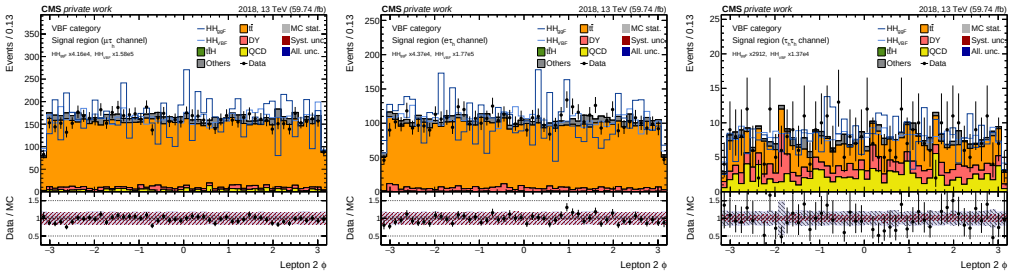
Figure A.42: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2018 data taking period.



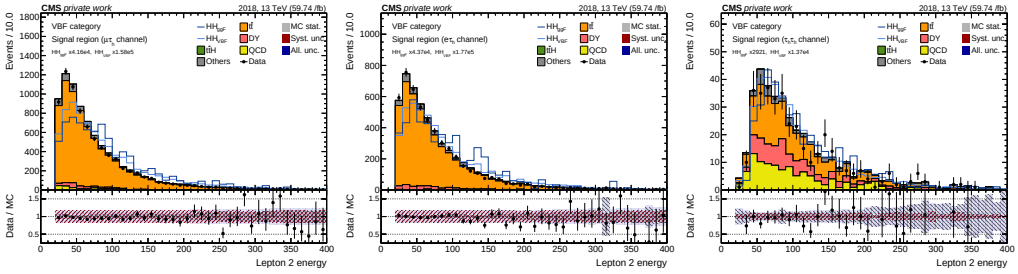
(b) `lep2_pt`



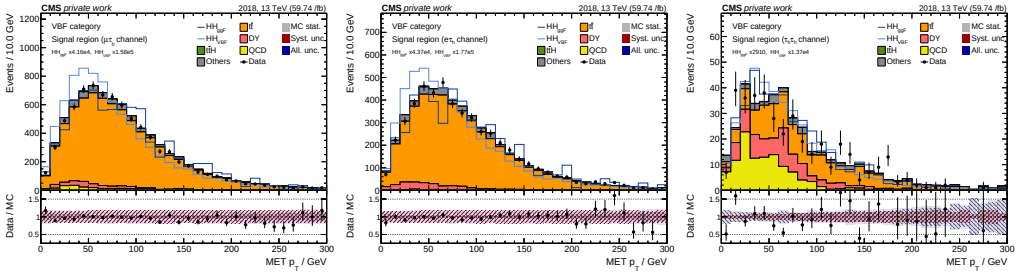
(e) `lep2_eta`



(h) `lep2_phi`

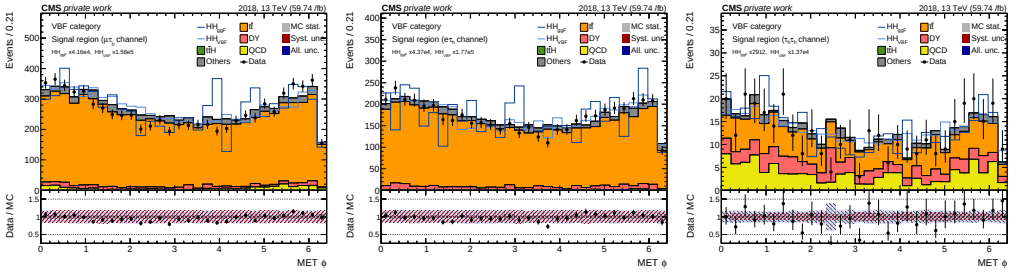


(k) `lep2_e`

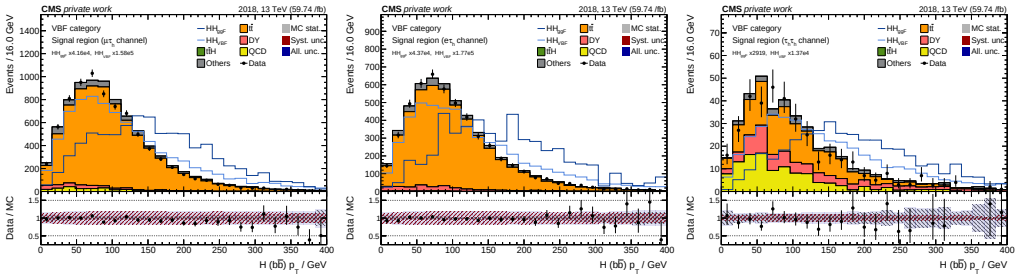


(n) `met_pt`

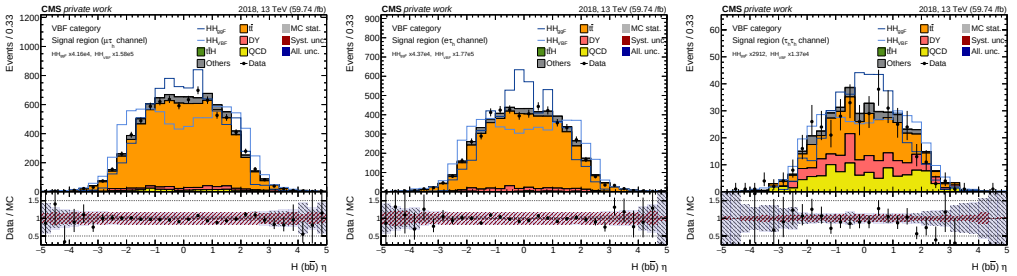
Figure A.43: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2018 data taking period.



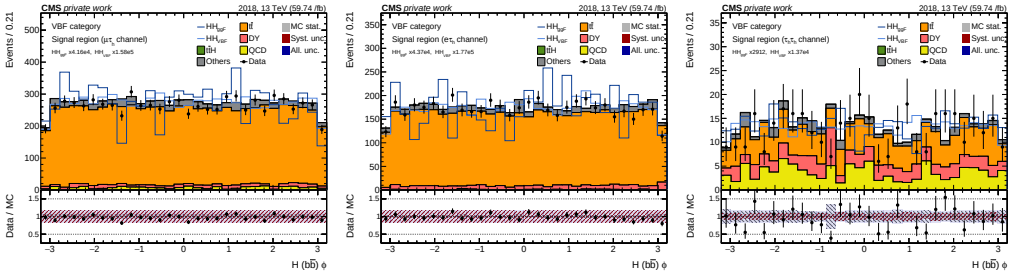
(b) `met_phi`



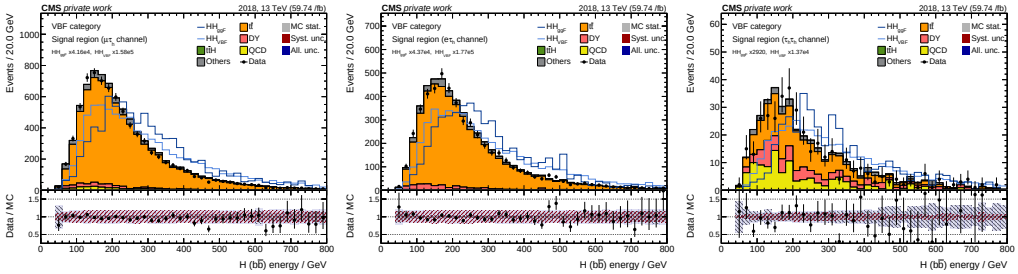
(e) `bh_pt`



(h) `bh_eta`

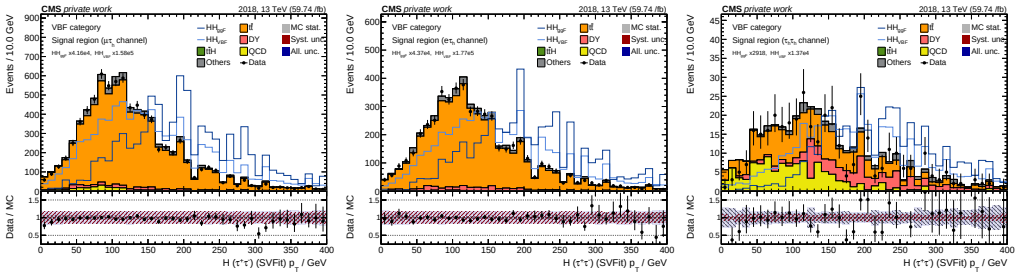


(k) `bh_phi`

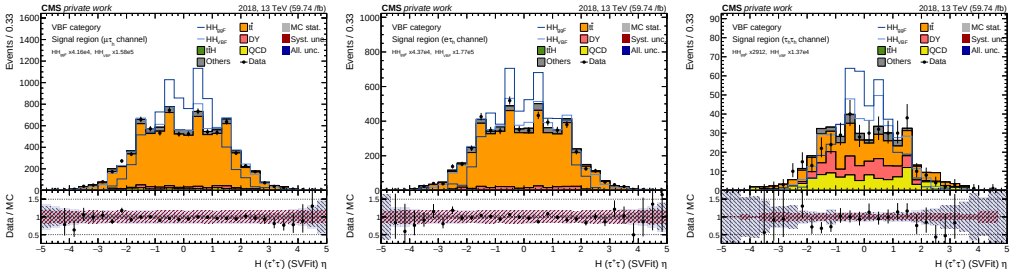


(n) `bh_e`

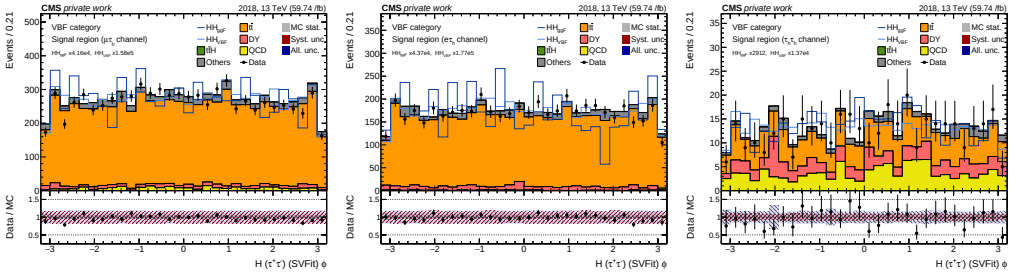
Figure A.44: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2018 data taking period.



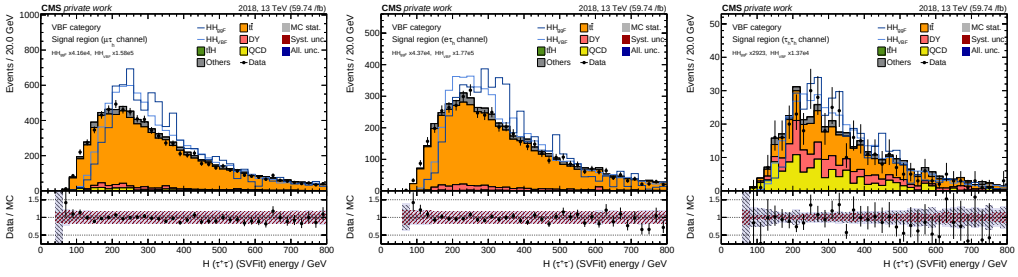
(b) `tauh_sv_pt`



(e) `tauh_sv_eta`



(h) `tauh_sv_phi`



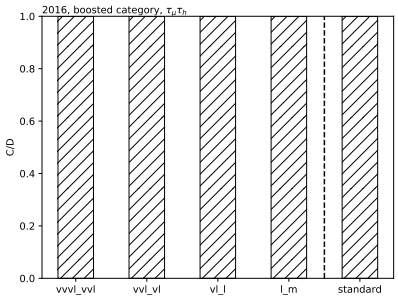
(k) `tauh_sv_e`

Figure A.45: Input feature distributions in the $\tau_\mu\tau_h$ (left), $\tau_e\tau_h$ (right), and $\tau_h\tau_h$ channels for the 2018 data taking period.

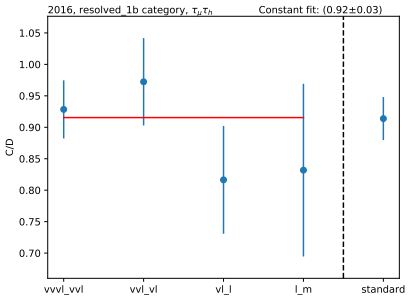
Appendix B

First validation test for QCD estimation

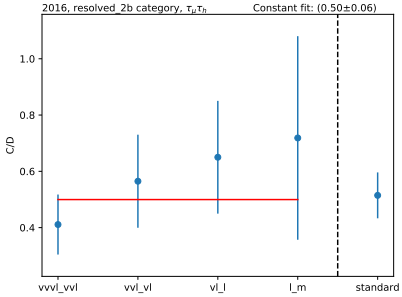
This appendix shows the results for the C/D yield estimation, corresponding to the first QCD validation test, in all years (2016, 2017, and 2018) and $\tau\tau$ decay channels. This test is fully described in Section 5.2.



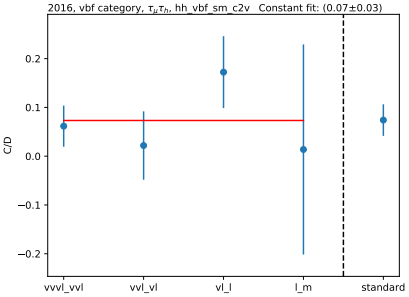
(a) Boosted



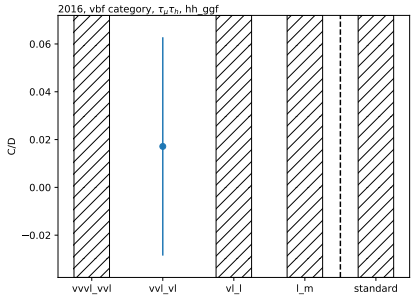
(b) Resolved 1 b-tag



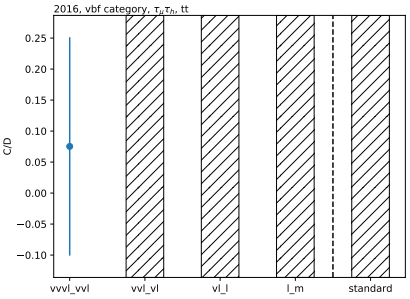
(c) Resolved 2 b-tag



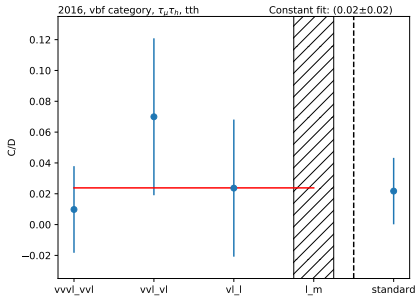
(d) VBF subcategory



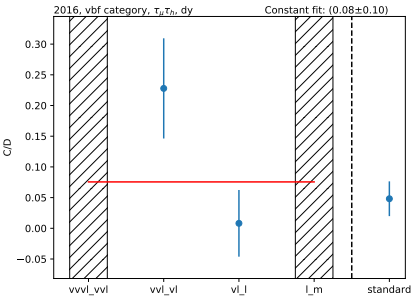
(e) ggF subcategory



(f) $t\bar{t}$ subcategory

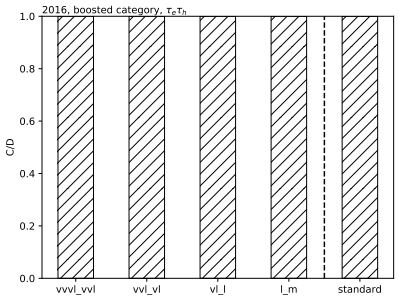


(g) $t\bar{t}H$ subcategory

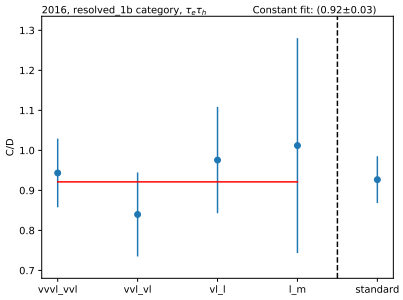


(h) DY subcategory

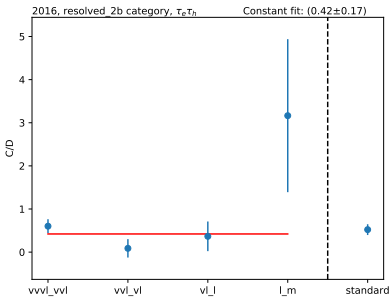
Figure B.1: C/D yield estimations for the five working points for the $\tau_\mu\tau_h$ channel in 2016. A dashed band is plotted where the correspondent working point leads to a negative yield for regions B, C and/or D.



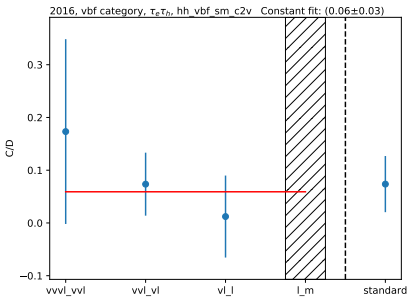
(a) Boosted



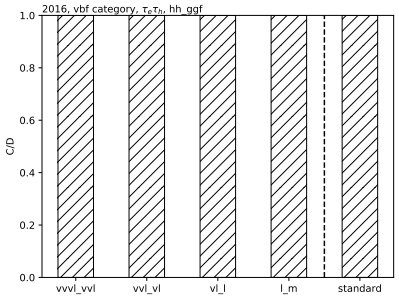
(b) Resolved 1 b-tag



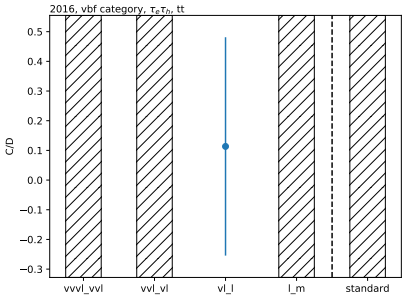
(c) Resolved 2 b-tag



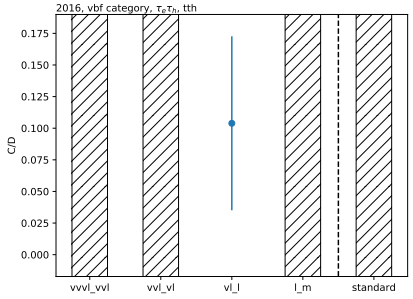
(d) VBF subcategory



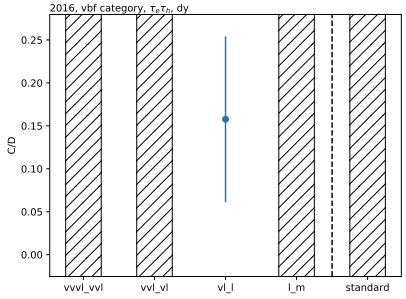
(e) ggF subcategory



(f) $t\bar{t}$ subcategory

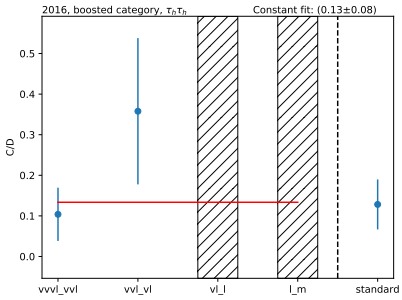


(g) $t\bar{t}H$ subcategory

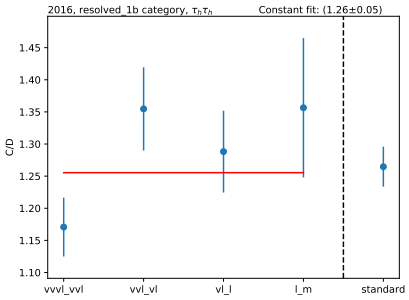


(h) DY subcategory

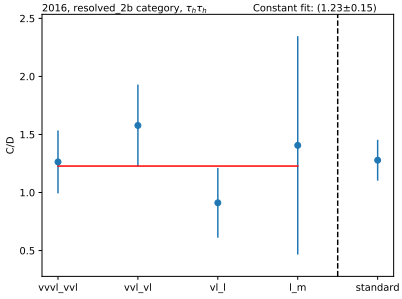
Figure B.2: C/D yield estimations for the five working points for the $\tau_e\tau_h$ channel in 2016. A dashed band is plotted where the correspondent working point leads to a negative yield for regions B, C and/or D.



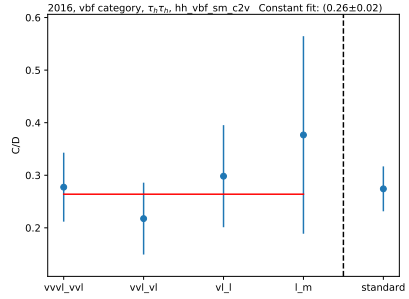
(a) Boosted



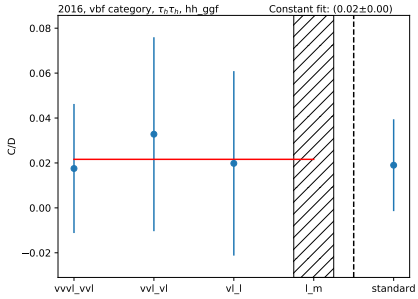
(b) Resolved 1 b-tag



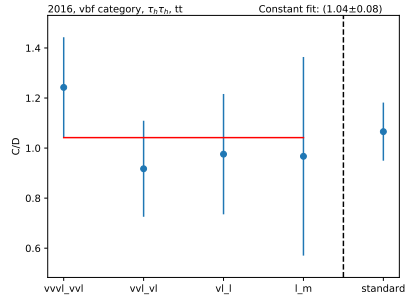
(c) Resolved 2 b-tag



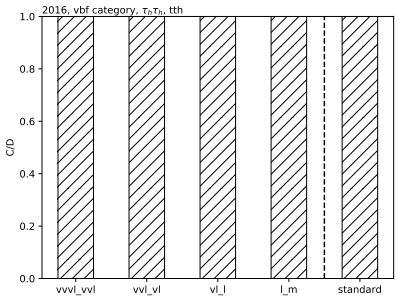
(d) VBF subcategory



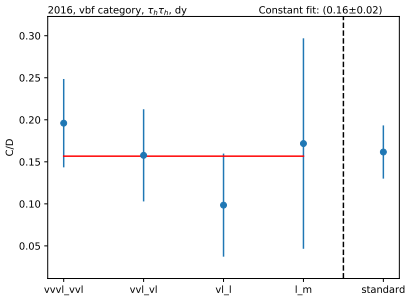
(e) ggF subcategory



(f) $t\bar{t}$ subcategory

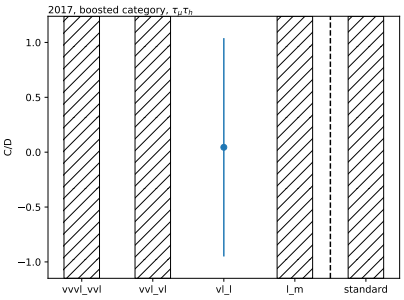


(g) $t\bar{t}H$ subcategory

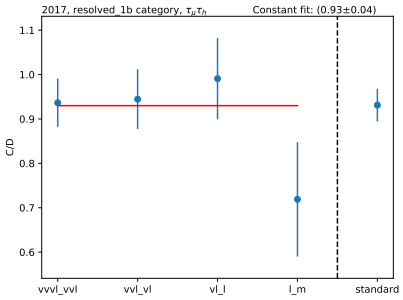


(h) DY subcategory

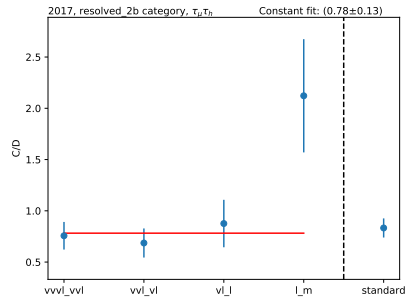
Figure B.3: C/D yield estimations for the five working points for the $\tau_h\tau_h$ channel in 2016. A dashed band is plotted where the correspondent working point leads to a negative yield for regions B, C and/or D.



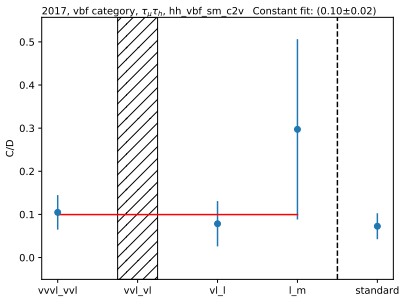
(a) Boosted



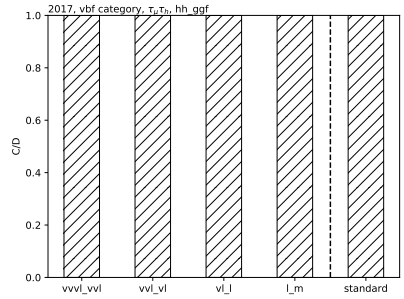
(b) Resolved 1 b-tag



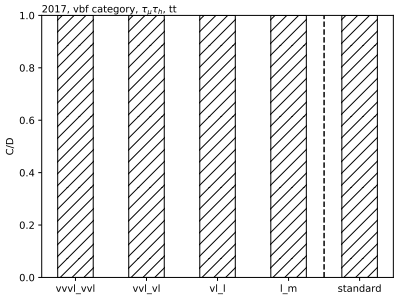
(c) Resolved 2 b-tag



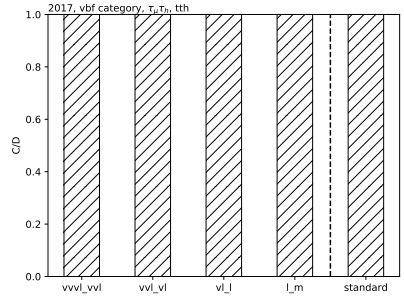
(d) VBF subcategory



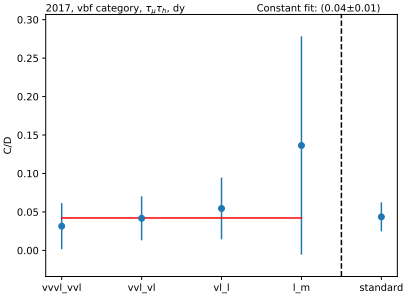
(e) ggF subcategory



(f) $t\bar{t}$ subcategory

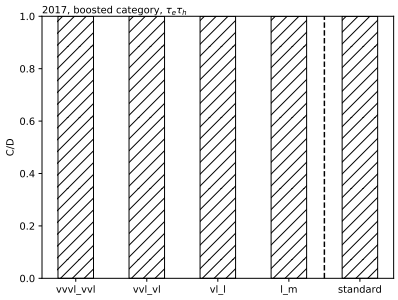


(g) $t\bar{t}H$ subcategory

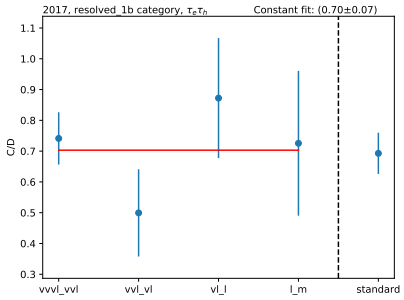


(h) DY subcategory

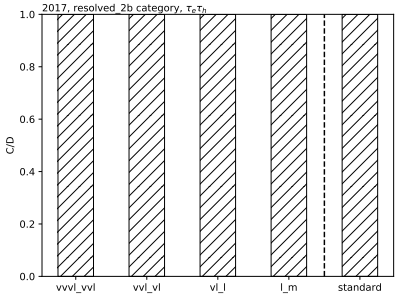
Figure B.4: C/D yield estimations for the five working points for the $\tau_\mu\tau_h$ channel in 2017. A dashed band is plotted where the correspondent working point leads to a negative yield for regions B, C and/or D.



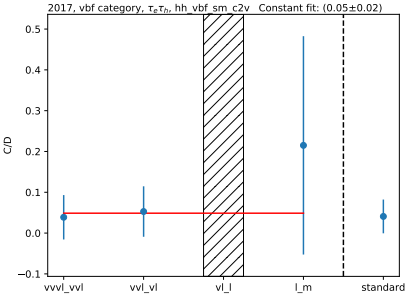
(a) Boosted



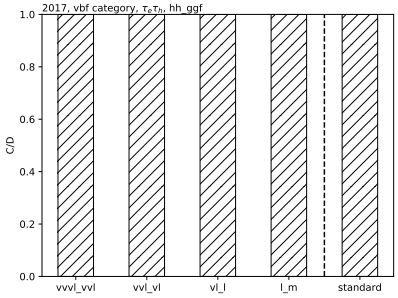
(b) Resolved 1 b-tag



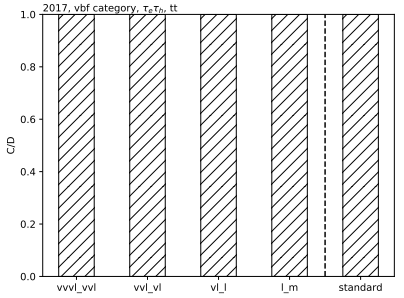
(c) Resolved 2 b-tag



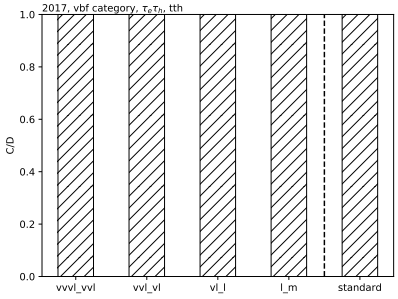
(d) VBF subcategory



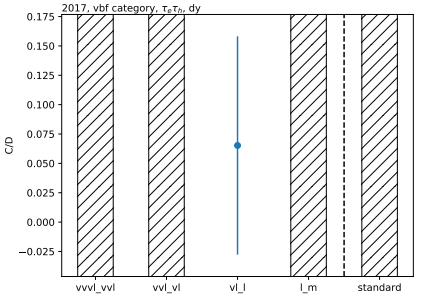
(e) ggF subcategory



(f) $t\bar{t}$ subcategory

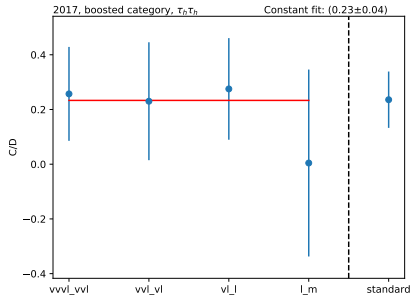


(g) $t\bar{t}H$ subcategory

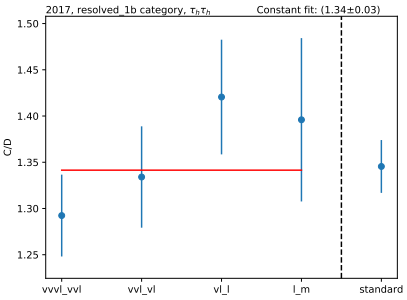


(h) DY subcategory

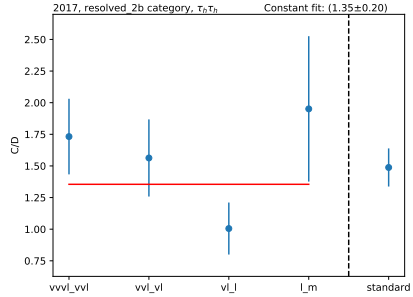
Figure B.5: C/D yield estimations for the five working points for the $\tau_e\tau_h$ channel in 2017. A dashed band is plotted where the correspondent working point leads to a negative yield for regions B, C and/or D.



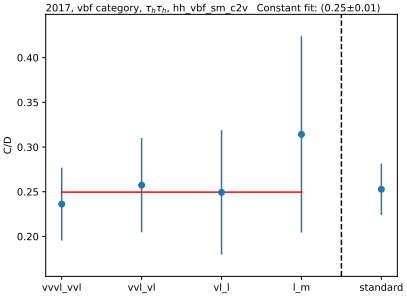
(a) Boosted



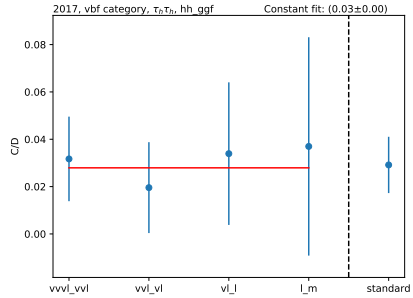
(b) Resolved 1 b-tag



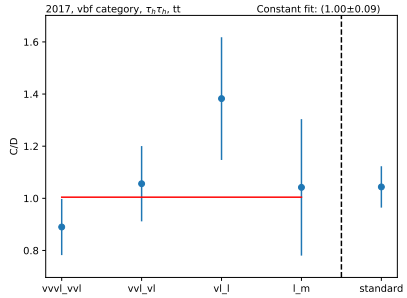
(c) Resolved 2 b-tag



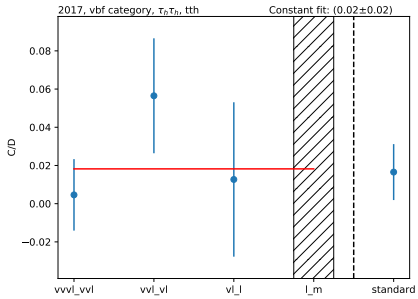
(d) VBF subcategory



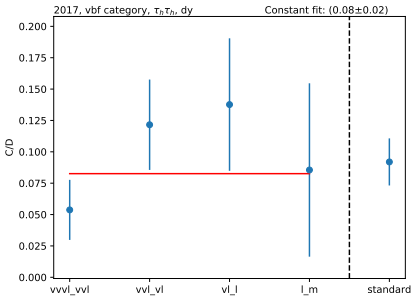
(e) ggF subcategory



(f) $t\bar{t}$ subcategory

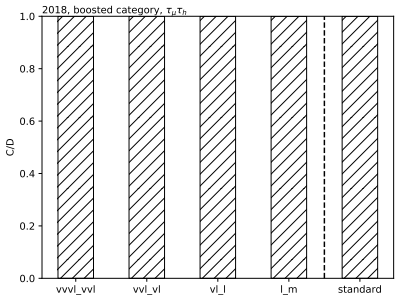


(g) $t\bar{t}H$ subcategory

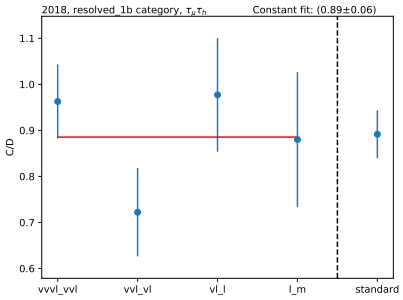


(h) DY subcategory

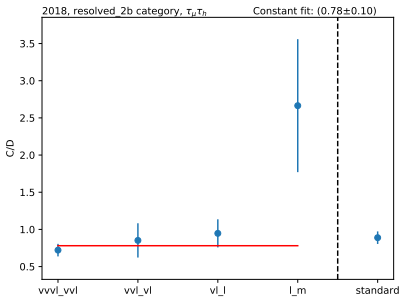
Figure B.6: C/D yield estimations for the five working points for the $\tau_h\tau_h$ channel in 2017. A dashed band is plotted where the correspondent working point leads to a negative yield for regions B, C and/or D.



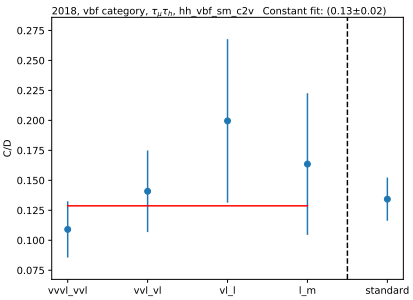
(a) Boosted



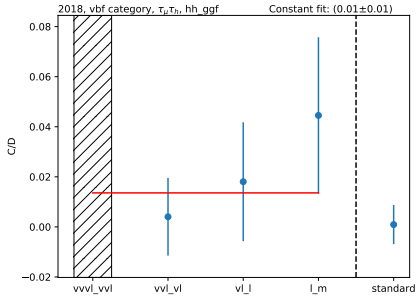
(b) Resolved 1 b-tag



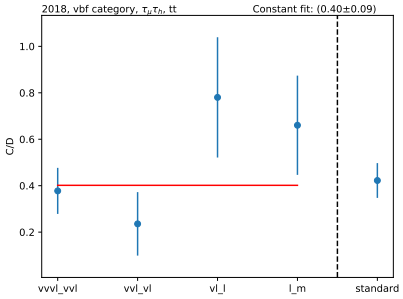
(c) Resolved 2 b-tag



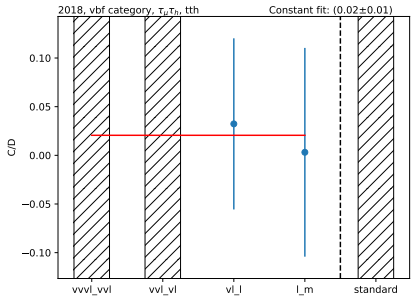
(d) VBF subcategory



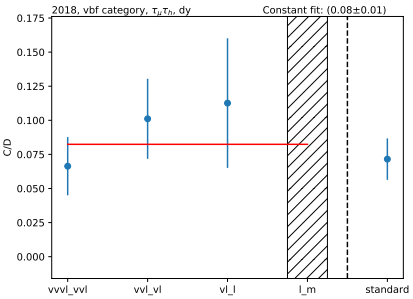
(e) ggF subcategory



(f) $t\bar{t}$ subcategory

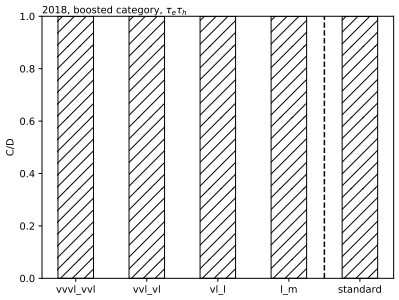


(g) $t\bar{t}H$ subcategory

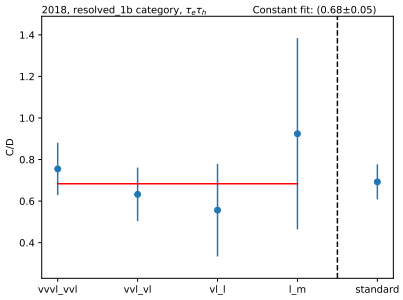


(h) DY subcategory

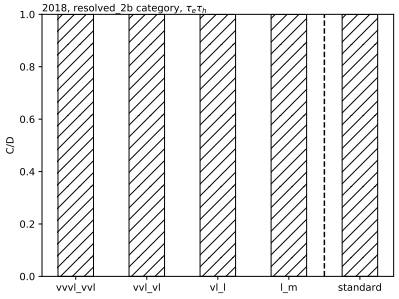
Figure B.7: C/D yield estimations for the five working points for the $\tau_\mu\tau_h$ channel in 2018. A dashed band is plotted where the correspondent working point leads to a negative yield for regions B, C and/or D.



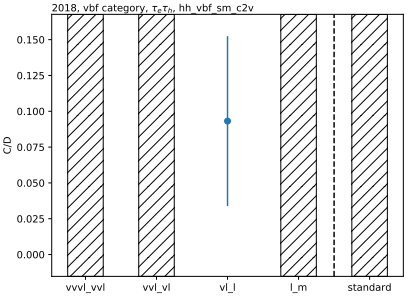
(a) Boosted



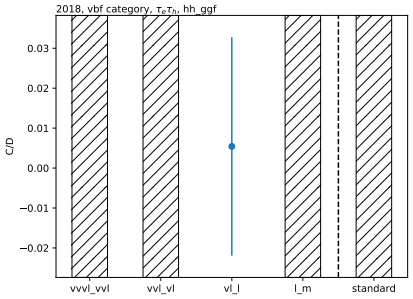
(b) Resolved 1 b-tag



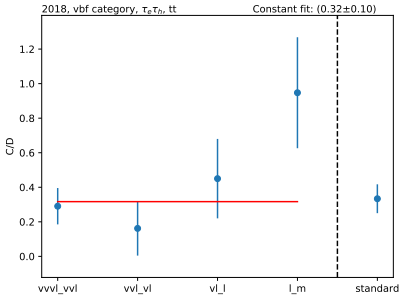
(c) Resolved 2 b-tag



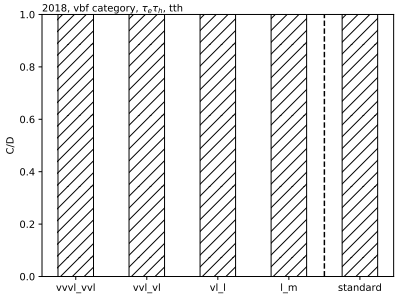
(d) VBF subcategory



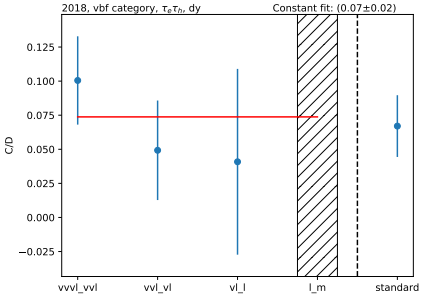
(e) ggF subcategory



(f) $t\bar{t}$ subcategory

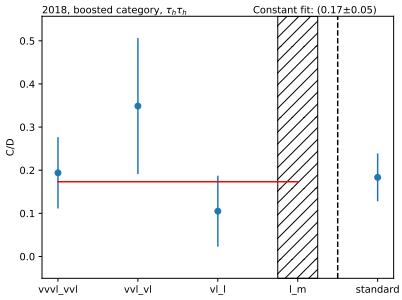


(g) $t\bar{t}H$ subcategory

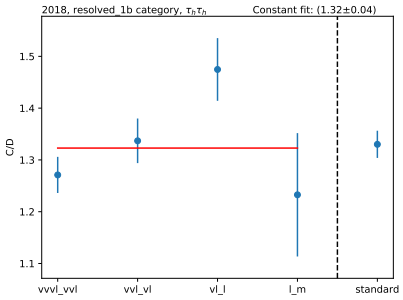


(h) DY subcategory

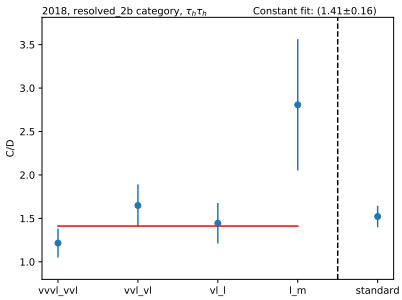
Figure B.8: C/D yield estimations for the five working points for the $\tau_e\tau_h$ channel in 2018. A dashed band is plotted where the correspondent working point leads to a negative yield for regions B, C and/or D.



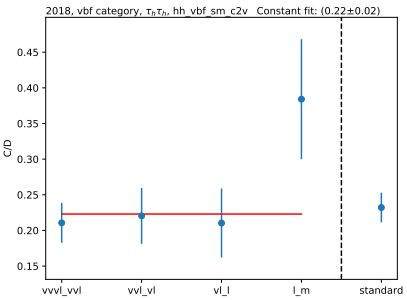
(a) Boosted



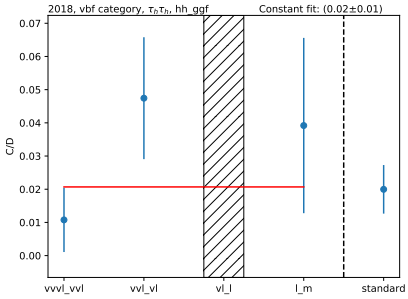
(b) Resolved 1 b-tag



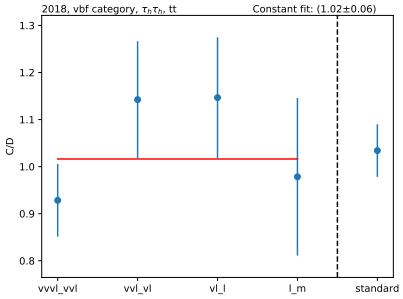
(c) Resolved 2 b-tag



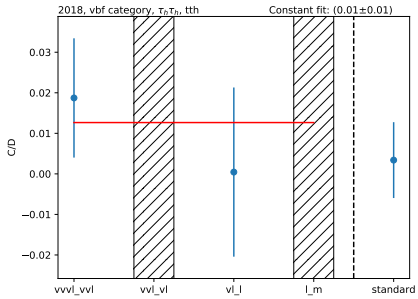
(d) VBF subcategory



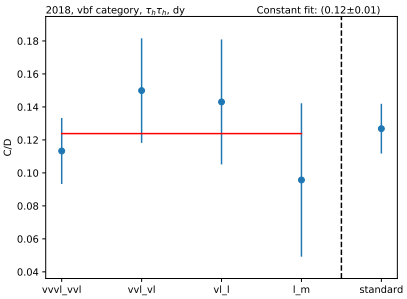
(e) ggF subcategory



(f) $t\bar{t}$ subcategory



(g) $t\bar{t}H$ subcategory



(h) DY subcategory

Figure B.9: C/D yield estimations for the five working points for the $\tau_h\tau_h$ channel in 2018. A dashed band is plotted where the correspondent working point leads to a negative yield for regions B, C and/or D.

List of Figures

List of Tables