



UNIVERSITÀ
DEGLI STUDI
FIRENZE

PhD IN
PHYSICS AND ASTRONOMY

CYCLE XXXV

COORDINATOR Prof. Raffaello D'Alessandro

**Precision measurement of the Higgs boson
production cross sections and couplings in the
WW decay channel with the CMS experiment
at the LHC using the Run II data set**

Academic Discipline (SSD) FIS/04

Doctoral Candidate

Dr. Roberto Seidita

Supervisors

Prof. Vitaliano Ciulli

Dr. Lorenzo Viliani

Coordinator

Prof. Raffaello D'Alessandro

Years 2019/2022

*Ce qui est admirable, ce n'est pas que le champ des étoiles soit si vaste,
c'est que l'homme l'ait mesuré.*

Jacques Anatole François Thibault

Contents

2	1 The coupling structure of the Higgs field	1
3	1.1 The SM in a nutshell	1
4	1.1.1 Interactions from symmetries	2
5	1.1.2 The symmetry structure of the SM	4
6	1.2 The spontaneous breaking of electroweak symmetry	6
7	1.3 Phenomenology of the Higgs boson at the LHC	11
8	1.3.1 Gluon fusion production	14
9	1.3.2 Vector boson fusion production	14
10	1.3.3 Associated production with a vector boson	15
11	1.3.4 Decay of the Higgs boson	16
12	1.4 The strong interaction at hadron colliders	16
13	1.4.1 The running of the strong coupling constant	18
14	1.4.2 Hadronic jets	18
15	2 The CMS detector at the LHC	21
16	2.1 The large hadron collider	21
17	2.2 The Compact Muon Solenoid	27
18	2.2.1 The silicon tracker	27
19	2.2.2 The electromagnetic calorimeter	30
20	2.2.3 The hadronic calorimeter	32
21	2.2.4 The superconducting magnet	33
22	2.2.5 The muon system	33
23	2.3 The trigger system in CMS	35
24	2.3.1 The level-1 trigger	35
25	2.3.2 The high level trigger	35
26	2.4 Event reconstruction in CMS	36
27	2.4.1 Muon reconstruction	36
28	2.4.2 Electron and photon reconstruction	37
29	2.4.3 Hadron reconstruction	38
30	2.4.4 Jet reconstruction	38
31	2.4.5 Missing transverse momentum reconstruction	42

32	3	The $H \rightarrow WW$ channel	45
33	3.1	Signal signatures	45
34	3.1.1	The ggH signature	45
35	3.1.2	The VBF signature	48
36	3.1.3	The WH and ZH signatures	50
37	3.2	Sources of background	54
38	3.2.1	Non-resonant WW production	54
39	3.2.2	The Drell-Yan process	54
40	3.2.3	Top quark production	55
41	3.2.4	Non-prompt lepton background	57
42	3.2.5	Multiboson production	58
43	4	Analysis strategy	59
44	4.1	Data sets and simulated samples	59
45	4.1.1	Trigger setup	59
46	4.1.2	Background modelling	60
47	4.1.3	Signal modelling	61
48	4.2	Definition of analysis objects and corrections	62
49	4.2.1	Trigger efficiencies in MC samples	65
50	4.2.2	Muons	69
51	4.2.3	Electrons	71
52	4.2.4	Hadronic jets	71
53	4.2.5	b tagging	73
54	4.2.6	Non-prompt leptons	74
55	4.3	Event selection and categorization	78
56	4.3.1	Different flavor ggH category	80
57	4.3.2	Same flavor ggH category	84
58	4.3.3	Different flavor VBF category	91
59	4.3.4	Same flavor VBF category	96
60	4.3.5	VH2j categories	97
61	4.3.6	WHSS category	99
62	4.3.7	WH3l category	102
63	4.3.8	ZH3l category	102
64	4.3.9	ZH4l category	106
65	4.4	The simplified template cross section measurement	109
66	4.4.1	The STXS framework	109
67	4.4.2	The STXS analysis	112

68	5	Signal extraction and systematic uncertainties	117
69	5.1	Signal extraction fit	117
70	5.1.1	Maximum likelihood parameter estimation	117
71	5.1.2	Treatment of systematic uncertainties	120
72	5.1.3	Data-driven background normalization	122
73	5.1.4	Goodness of fit	123
74	5.2	Data-driven background estimation setup	124
75	5.3	Sources of systematic uncertainty	126
76	5.3.1	Sources of theoretical uncertainty	126
77	5.3.2	Sources of experimental uncertainty	127
78	5.3.3	Process-specific uncertainties	129
79	5.3.4	Theoretical uncertainties in the STXS measurement . .	130
80	6	Results	131
81	6.1	Signals strength modifiers	134
82	6.2	Couplings to vector bosons and fermions	137
83	6.3	STXS	141
84	6.4	Combination with other decay channels	143

1 Introduction

2 The discovery of the Higgs boson in 2012 by the ATLAS and CMS collabora-
3 tions at the CERN LHC accelerator marked the final experimental confirma-
4 tion of the Standard Model (SM) of particle physics, the theory that encodes
5 our best understanding of fundamental particles and their interactions. Ten
6 years after the discovery, Run 2 of the LHC has marked the transition to the
7 precision era in the characterization of the Higgs boson.

8 In this work I will present the measurement of the production cross sec-
9 tions and couplings of the Higgs boson in proton-proton collision events at
10 a center of mass energy of 13 TeV, produced by the LHC. The analysis is
11 performed on the data set collected by CMS during Run 2 of the LHC, cor-
12 responding to an integrated luminosity of 138 fb^{-1} . The focus is on events
13 in which the Higgs boson decays to a pair of electroweak W bosons with
14 opposite electric charge. This channel is characterized by a large branching
15 ratio and relatively low-background final state, making it a prime candidate
16 for precision characterization of the Higgs boson. The analysis is designed to
17 tackle the four most probable production modes of a Higgs boson at the LHC,
18 achieving 10% accuracy on the total Higgs boson production cross section.
19 A second set of results is presented in terms of couplings of the Higgs boson
20 to vector bosons and fermions, probing the structure of the Higgs sector of
21 the SM to 5% accuracy. Finally, a measurement of 14 cross sections in the
22 simplified template cross section framework is presented.

23 During its development, I was the person responsible for the analysis,
24 the main author of the resulting CMS publication, and performed the signal
25 extraction fits on the combination of all analysis categories described in this
26 work.

27 This thesis is organized as follows. Chapter 1 contains a brief overview
28 of the theoretical foundations of the SM, with focus on the Brout-Englert-
29 Higgs mechanism and phenomenology of the Higgs boson. Chapter 2 gives
30 an overview of the CMS detector at the CERN LHC. Chapter 3 details the
31 experimental signatures accessible through the $H \rightarrow WW$ channel. Chapter 4
32 outlines the analysis strategy, event selection, and discriminating observables
33 used to extract the results. Chapter 5 details the statistical procedure for
34 extracting the result and the treatment of systematic uncertainties. Finally,
35 results are given in Chapter 6.

Chapter 1

The coupling structure of the Higgs field

Our current understanding of the fundamental structure of matter is encoded in the aptly named standard model (SM) of particle physics. Developed in the latter half of the 1900s, the theory is able to accurately predict phenomena on a wide range of characteristic energies, starting from a relatively slim set of assumptions. Arguably the last outstanding issue in the theoretical framework of the SM was that any elementary particle carrying mass was in apparent contradiction with the symmetry assumptions on which the model is built. The issue was elegantly solved by Peter Higgs [1], Robert Brout, and François Englert [2] in 1964 by adding an additional particle to the model, nowadays commonly referred to as the Higgs boson (H). Experimental confirmation of the Brout-Englert-Higgs hypothesis arrived in 2012, when the ATLAS and CMS collaborations at the LHC announced the observation of a new scalar resonance [3, 4, 5], compatible with the H boson.

While a comprehensive description of the SM is well beyond the scope of this thesis, I will present a short summary of its most prominent features, with a higher level of detail reserved for the physics of the H boson, the subject of this work. Throughout this thesis, unless otherwise specified, natural units are used i.e., $c = \hbar = 1$.

1.1 The SM in a nutshell

The two fundamental assumptions underlying the SM are that:

- at its core, matter is made up of quantum fields with spin $1/2$ (fermions) that interact with each other by way of another set of quantum fields

with integer spin (bosons). Particles are the quanta of excitation of these fields.

- Nature shows a set of symmetries pertaining to both spacetime (translational and Lorentz invariance) as well as the structure of the quantum fields themselves (gauge symmetry).

The choice of symmetries and fields determines the dynamical behavior that the theory predicts. Over a few decades, the model has been tuned and expanded to the point where it successfully describes two of the three known fundamental interactions: the electroweak interaction and the strong interaction. The remaining known force in Nature, the gravitational force, is not included in the SM in any fashion. The reason is twofold: on the theoretical side, combining the quantum field formulation (on which the SM is built) with general relativity (the currently accepted description of gravity) is a challenge yet to be overcome; on the experimental side, the effects of gravity on the scale of elementary particles are so weak as to be made irrelevant by any of the other interactions, and are therefore not measurable with currently available technologies.

1.1.1 Interactions from symmetries

The most common formulation of the SM is in the lagrangian description. Very briefly, the dynamics predicted by the theory is the one that makes the action stationary (the stationary point is commonly a minimum). In natural units, the action is a dimensionless quantity, and in the typical framework of quantum field theory it is expressed as an integral over spacetime of a lagrangian density $\mathcal{L}$:

$$S = \int d^4x \mathcal{L}(\phi_i, \partial_\mu \phi_i, x). \quad (1.1)$$

Although the action S is the fundamental quantity in the quantum field theory framework, one usually works with the lagrangian density $\mathcal{L}$. Here ϕ_i represents the generic set of fields that make up the theory. The first requirement is for $\mathcal{L}$ to be invariant under a reference frame change i.e., under the set of spacetime translations, rotations, and Lorentz boost transformations, which make up the Poincaré group. This requirement ensures conservation of energy, momentum, and angular momentum. Additional symmetries can be enforced on the lagrangian density to achieve the desired predictions. In the SM, these are the so-called gauge symmetries i.e., local transformations in the space of internal degrees of freedom of the fields. Enforcing gauge

invariance on the lagrangian density has proven to be a powerful tool in developing the SM, as it naturally gives rise to interactions between the different fields, and thus to the dynamics the theory predicts. The simplest example is that of quantum electrodynamics (QED), which is contained in the SM. The non-interacting QED lagrangian density is:

$$\mathcal{L}_{QED}^{free} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \quad (1.2)$$

where ψ is a spin 1/2 fermion (e.g., an electron) and $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$ is the strength tensor of the electromagnetic potential A^μ . Eq. 1.2 is invariant under the following transformation:

$$\psi \longrightarrow e^{i\alpha}\psi, \quad A_\mu \longrightarrow A_\mu + i\partial_\mu\beta(x). \quad (1.3)$$

While in the above expression β can generally be a function of the spacetime coordinate x , the same does not hold for α i.e., the phase shift of the fermion field ψ has to be the same throughout spacetime for Eq. 1.2 to be invariant. This restriction can (and should) appear unphysical: it is established from basic quantum mechanics that a phase shift of the field should not impact observable quantities, and should thus leave the dynamics of the theory unchanged. In a quantum field theoretical framework, this is equivalent to stating that the lagrangian density should be invariant under a local phase shift of the fields i.e., the transformation parameter α in the example above should be allowed to depend on x . This apparent limitation can be elegantly solved by noticing the form of the extra term that is generated in Eq. 1.2 when ψ undergoes a local phase shift:

$$\bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi \longrightarrow \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi + i\bar{\psi}\gamma^\mu\psi\partial_\mu\alpha(x). \quad (1.4)$$

The extra term, negating invariance of the lagrangian density, is the product of $\partial_\mu\alpha(x)$ and $j^\mu = \bar{\psi}\gamma^\mu\psi$, the vector current associated to ψ . This term can be exactly cancelled by adding an interacting term to Eq. 1.2 between the current and A_μ :

$$\mathcal{L}_{QED} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - j^\mu A_\mu. \quad (1.5)$$

Under the condition that $\alpha = \beta$, a local phase shift of ψ is now exactly cancelled by a simultaneous gauge transformation of A_μ , thus leaving the lagrangian density invariant. The local phase transformation $\psi \longrightarrow e^{i\alpha(x)}\psi$ falls under the group of commutative and unitary transformations $U(1)$; invariance of the lagrangian density under this family of transformations is therefore commonly referred to as local $U(1)$ symmetry.

125 This very condensed (and minutia-deprived) example hopefully gives a
 126 sense of the power of symmetry assumptions in quantum field theory, as
 127 they naturally give rise to interactions between fermions and gauge fields.

128 1.1.2 The symmetry structure of the SM

129 The SM can be thought of as an expansion upon the idea laid out in the
 130 previous paragraph. Instead of a single fermion field, the SM contains six:
 131 three quarks and three leptons. The group of symmetries under which
 132 the lagrangian density is assumed to be invariant is likewise expanded to
 133 $SU(3)_C \times SU(2)_L \times U(1)_Y$. The symbols C, L and Y stand for color, left-
 134 handed chiral, and hypercharge, respectively. The $SU(3)_C$ symmetry de-
 135 scribes the strong interaction, while the $SU(2)_L \times U(1)_Y$ symmetry describes
 136 the electroweak one. To specify the structure of the SM, one needs to specify
 137 how each of the fermion fields behaves with respect to each symmetry group.

138 The three lepton fields (referred to as the three generations for historical
 139 reasons) are spin 1/2 fermions. Their left handed chiral components form
 140 $SU(2)$ doublets and carry a hypercharge of -1 :

$$\begin{pmatrix} \nu \\ \ell \end{pmatrix}_L = \left\{ \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L \right\}. \quad (1.6)$$

141 The $SU(2)_L$ degree of freedom is the third component of weak isospin T_3 ,
 142 conventionally equal to $+1/2$ for the up components and $-1/2$ for the down
 143 components. The down components of the lepton fields are the electron,
 144 muon, and tau, which also have right-handed equivalents carrying $Y = -2$
 145 and behaving as scalars under $SU(2)_L$:

$$\ell_R = e_R, \mu_R, \tau_R. \quad (1.7)$$

146 The up components are the electron neutrino, muon neutrino and tau neu-
 147 trino. These fields have no right-handed equivalents. The fact these fields
 148 transform non-trivially under the $SU(2)_L \times U(1)_Y$ group is equivalent to stat-
 149 ing that leptons are subject to the electroweak interaction. On the other
 150 hand, leptons behave as singlets (i.e., they transform trivially) under the
 151 $SU(3)_C$ group, meaning that they are not subject to the strong interaction.
 152 It can be shown that electric charge Q is related to weak isospin and hyper-
 153 charge through the relation:

$$Q = T_3 + \frac{Y}{2}, \quad (1.8)$$

also encoding the fact that neutrinos do not carry electric charge, making them the most weakly interacting particles in the SM.

The second set of fermions in the SM is that of quarks. These are once again spin 1/2 fields whose left-chiral components form $SU(2)_L$ doublets:

$$\begin{pmatrix} q_u \\ q_d \end{pmatrix}_L = \left\{ \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L \right\}, \quad (1.9)$$

and carry $Y = 1/3$. Contrary to leptons, both up and down components of quark fields have right-handed equivalents behaving as $SU(2)_L$ singlets:

$$\begin{aligned} q_{u,R} &= u_R, \quad c_R, \quad t_R, & Y &= 4/3 \\ q_{d,R} &= d_R, \quad s_R, \quad b_R, & Y &= -2/3. \end{aligned} \quad (1.10)$$

It follows from Eq. 1.8 that all quarks carry electric charge. The most important difference with respect to leptons however is that quarks also form triplets under $SU(3)_C$. This encodes the fact that they carry an additional form of charge, called the color charge, which makes them subject to the strong interaction.

Imposing the $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry on this set of fermion fields gives rise to the rich interaction structure of the SM. These interactions are mediated by a set of bosonic fields, similarly to what was described for the electromagnetic case in the simpler QED example given previously. From imposing $SU(2)_L \times U(1)_Y$ symmetry, three spin 1 gauge fields arise. Two of these, the Z boson and the photon (γ), are electrically neutral, while the $W^\pm$ boson carries one unit of electric charge. The Z and W bosons couple to every left-handed field in the SM, while the photon couples to every particle carrying electric charge (thus also including the $W^\pm$ bosons). The $SU(3)_C$ symmetry, on the other hand, gives rise to eight spin 1 bosons called gluons, which are electrically neutral but carry color charge. These couple to every colored particle, meaning all quarks, as well as themselves. The self-coupling of gluons is one of the peculiarities of quantum chromodynamics (QCD), the theory of strong interactions. The interaction scheme of the SM is summarized in Fig. 1.1. While the SM, as (briefly) summarized above, has proven effective in describing a wide range of natural phenomena observed in high energy physics (HEP) experiments, there is one important missing point. Under the assumption of $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetry, no particle in the SM is allowed to carry mass. For fermions, a mass term in the lagrangian density would be of the form:

$$m\bar{\psi}\psi = m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L), \quad (1.11)$$

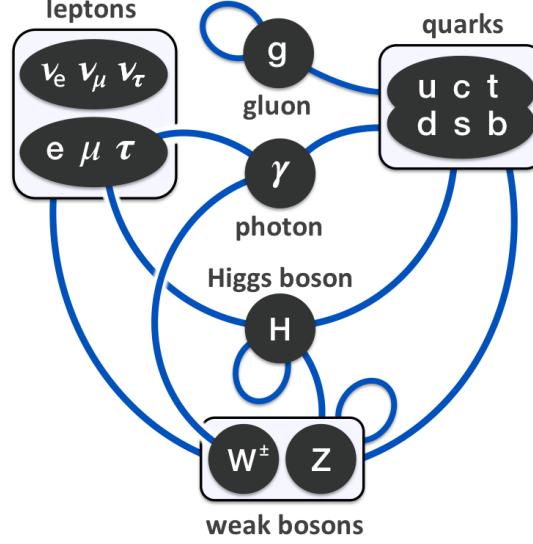


Figure 1.1: Interaction scheme in the SM. Particles connected through lines share an interaction vertex through at least one of the interactions in the SM.

185 coupling left- and right-handed parts of the field and thus breaking $SU(2)_L$
 186 symmetry. A similar argument can be made for a generic vector boson V ,
 187 whose mass term would take the form:

$$m^2 V^\mu V_\mu, \quad (1.12)$$

188 which is not invariant under the gauge transformation for a vector field, as
 189 given in Eq. 1.3 for A_μ . This is in contrast with the experimental evidence
 190 that all charged leptons, all quarks, and the W and Z bosons carry nonzero
 191 mass. In the case of the weak gauge bosons in particular, this aspect is crucial,
 192 since their large (compared to up and down type quarks or electrons) mass
 193 is needed to explain the short range of the weak force.

194 1.2 The spontaneous breaking of electroweak 195 symmetry

196 The solution to the issue of massive particles in the SM came through the
 197 introduction of an additional field to the model. The idea of spontaneous
 198 breaking of symmetry (SSB) is not unique to particle physics: every time
 199 there is a system whose equations of motion (equivalently, the lagrangian

density in the framework of HEP) are invariant under a given transformation, but the ground state of the theory is not, spontaneous breaking of symmetry occurs. While abstaining from any claim to generality (a more formally complete treatment can be found in [6] and [7]), I will give an overview of the specific case of the SM.

A simple, yet complete, example of spontaneous symmetry breaking in the framework of quantum field theory is that of a charged, massless, scalar field ϕ coupled to the electromagnetic field A_μ , driven by the free lagrangian density:

$$\mathcal{L}_{free} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + D_\mu\phi(D^\mu\phi)^\dagger, \quad (1.13)$$

which is invariant under phase shifts of ϕ or, in other terms, under U(1) gauge transformations. Here $D_\mu \equiv \partial_\mu + ieA_\mu$ is the covariant derivative, a compact notation for the gauge-invariant construction given in the previous section. The field ϕ takes values in the complex plane. The dynamics are made more interesting by adding a potential term to Eq. 1.13, in the form:

$$V(\phi) = -\mu^2\phi^\dagger\phi + \frac{\lambda}{2}(\phi^\dagger\phi)^2, \quad (1.14)$$

which is still invariant under U(1) transformations, thus leaving $\mathcal{L} = \mathcal{L}_{free} - V$ invariant as well. The parameter λ is bound to positive values in order for the potential to be lower bound and thus not give rise to unphysical dynamics. The peculiarity of this lagrangian density is that for $\mu^2 > 0$, while it itself is U(1) invariant, it produces a ground state that is not. This can be seen simply by deriving the potential with respect to the fields:

$$\frac{\partial V}{\partial \phi} = \phi^\dagger(-\mu^2 + \lambda\phi^\dagger\phi). \quad (1.15)$$

The minimum of the potential is thus reached for $|\phi|^2 = \mu^2/\lambda \equiv v$ (for $\mu^2 > 0$ the other extremum in $|\phi| = 0$ is a maximum). This forms an infinitely degenerate set of minima located on a circle of radius v in the complex plane. Crucially, a state located on this set is not invariant under U(1) transformations of the field, since it is not symmetrical with respect to the origin. This is a general feature, known as spontaneous breaking of symmetry. One important thing to note is that the theory can no longer be treated perturbatively around the origin, as it doesn't correspond to the ground state. This is easily circumvented by redefining the field. Decomposing it into its real components $\phi = \phi_1 + i\phi_2$, and accounting for the non-zero ground state, this can be done as:

$$\phi = v + \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2). \quad (1.16)$$

231 This is one of infinite possible choices (all points on the circumference of
 232 minima are equally valid, the real direction was chosen for convenience), but
 233 all are equivalent since they only differ by a rotation. This new field can be
 234 treated perturbatively, since it is zero at the ground state or, as it is more
 235 commonly said in the context of HEP, its vacuum expectation value (VEV) is
 236 0. The first physical consequence of this construction is that ϕ_1 has acquired
 237 mass. This can be seen by rewriting the potential in terms of ϕ_1 and ϕ_2 :

$$V(\phi) = -\frac{\mu^4}{2\lambda} + \frac{1}{2}2\mu^2\phi_1^2 + \mathcal{O}(\phi_i^3). \quad (1.17)$$

238 The first term is constant and thus doesn't impact the dynamics; the second
 239 term is indeed a mass term for ϕ_1 , with $m_{\phi_1} = \sqrt{2}\mu$. The ϕ_2 component of
 240 the field remains massless, as excitations of this field correspond to movement
 241 along the circle of minima, a direction in which the potential is flat. This
 242 component is called a Goldstone boson.

243 The second notable aspect is made evident by rewriting the kinetic term
 244 of ϕ in the same fashion:

$$\begin{aligned} D_\mu\phi(D^\mu\phi)^\dagger &= \frac{1}{2}(\partial_\mu\phi_1\partial^\mu\phi_1)^2 + \frac{1}{2}(\partial_\mu\phi_2\partial^\mu\phi_2)^2 \\ &+ \sqrt{2}evA_\mu\partial^\mu\phi_2 + \frac{1}{2}2e^2v^2A_\mu A^\mu + \mathcal{O}(\phi_i\phi_jA_\mu), \end{aligned} \quad (1.18)$$

245 where the last term contains everything of order 3 or higher in any combi-
 246 nation of the fields. The last explicitly written term is a mass term for the
 247 photon field A_μ , with $m_A = \sqrt{2}ev$. The crucial aspect of this construction
 248 is that this mass term for the photon does not violate gauge symmetry, as
 249 it arises from an interaction with the VEV of the scalar field ϕ . This mech-
 250 anism is the key to accommodate massive gauge fields in a gauge invariant
 251 quantum field theory. The final aspect to note is that the Goldstone boson
 252 ϕ_2 in this construction does not behave as an independent physical particle.
 253 This is due to the fact that with a particular choice of gauge transformation,
 254 called the unitary gauge, it is possible to make ϕ_2 vanish from the lagrangian
 255 density, leaving only the real field ϕ_1 and the (massive) photon:

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \partial_\mu\phi_1\partial^\mu\phi_1 + e^2v^2A_\mu A^\mu - V(\phi). \quad (1.19)$$

256 This can be thought of as the degree of freedom encoded in the Goldstone
 257 boson being recast as the longitudinal polarization of the photon, which it
 258 can only have if it is massive.

259 The application of SSB to the SM is done through the Brout-Englert-
 260 Higgs (BEH) mechanism. The basis is the same, although with an added

degree of complexity. The main difference is that the symmetry we wish to be spontaneously broken is the local $SU(2)_L \times U(1)_Y$ one, and to do so a set of 4 scalar fields is added to the theory. These can be conveniently grouped in an $SU(2)_L$ doublet:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi_0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \varphi - i\eta \end{pmatrix} \quad (1.20)$$

The $U(1)_Y$ charge of Φ can be derived by preemptively stating that we will assign non-vanishing VEV to its down component. This assignment is arbitrary, since it can in principle be assigned to any direction in the (ϕ^+, ϕ_0) plane and mapped to our choice through an $SU(2)_L$ rotation. Given ϕ_0 will be assigned a non-vanishing VEV, it follows that this component must be electrically neutral, since otherwise the vacuum itself would present a net electric charge. Following Eq. 1.8 the hypercharge of the doublet should therefore be +1, also fixing the electric charge of the ϕ^+ component to +1. The potential term added to the SM lagrangian density is also of the same form:

$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2. \quad (1.21)$$

This is essentially a 4-dimensional realization of Eq. 1.14. As anticipated, the non-vanishing VEV will be assigned to ϕ_0 , so the doublet will be shifted to:

$$\Phi' = \Phi - \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad (1.22)$$

where v is again the VEV of the field. In analogy with the previous example, ϕ_0 , ϕ_1 , and η behave as Goldstone bosons, while h is the direction along which the potential is not flat and will thus acquire mass. This is the Higgs field. As in the electromagnetic case, we have the freedom to operate a transformation to the unitary gauge, in which the Goldstone degrees of freedom are recast as longitudinal polarizations of the $W^\pm$ and Z bosons, which acquire mass:

$$\begin{aligned} \mathcal{L}_\Phi^{\text{kin.}} &= D_\mu \Phi (D^\mu \Phi)^\dagger \xrightarrow{SSB} \\ &\xrightarrow{SSB} \frac{1}{2} \partial_\mu h \partial^\mu h + \left(1 + \frac{h}{v}\right)^2 \left[m_W^2 W^{+\mu} W_\mu^- + \frac{1}{2} m_Z^2 Z^\mu Z_\mu \right], \end{aligned} \quad (1.23)$$

where:

$$m_W^2 = \frac{g^2 v^2}{4}, \quad m_Z^2 = \frac{g^2 v^2}{4 \cos^2 \theta_W}. \quad (1.24)$$

The additional parameter θ_W is the Weinberg mixing angle [8]. It enters the model since the physical Z and γ fields are a linear combination of the two

287 electrically neutral $SU(2)_L \times U(1)_Y$ gauge fields, which can be parameterized
 288 as a rotation by θ_W in the space of the two fields. In addition to mass terms
 289 for the W and Z bosons, Eq. 1.23 also contains interaction terms with the H
 290 boson:

$$\mathcal{L}_\Phi \supset \frac{2}{v} m_W^2 (h W^{+\mu} W_\mu^-) + \frac{1}{v} m_Z^2 (h Z^\mu Z_\mu), \quad (1.25)$$

291 with couplings proportional to the square masses of the vector bosons. The
 292 mass of the Higgs boson in this picture is found to be, in complete analogy
 293 with Eq. 1.17:

$$m_H = \sqrt{2}|\mu|. \quad (1.26)$$

294 Note that, while the VEV of the Higgs field can be calculated by knowing m_W
 295 and the electroweak coupling constant g , m_H is only proportional to μ , which
 296 is a free parameter of the model. Its mass therefore cannot be predicted and
 297 must be measured experimentally, a fact that made the experimental search
 298 for the H boson noticeably more challenging. The most precise measurement
 299 of m_H at the time of writing has been performed by the CMS collaboration,
 300 yielding a value of 125.38 ± 0.14 GeV [9].

301 Although mass terms for fermions do not arise directly from the BEH
 302 mechanism itself, the addition of the Higgs field allows to include them in the
 303 SM in a gauge invariant way. The addition of the Φ doublet with hypercharge
 304 $Y = 1$ permits the inclusion of an additional set of terms in the SM lagrangian
 305 density, the so-called Yukawa terms. For leptons, these are written as:

$$g_\ell (\bar{\nu}_{\ell,L}, \bar{\ell}_L) \Phi \ell_R + h.c., \quad (1.27)$$

306 where ℓ is any lepton field and L, R indicate the left- and right-handed
 307 components, respectively. Note that such a term is indeed neutral with
 308 respect to all charges and is thus gauge invariant, and this is only possible
 309 because the positive hypercharge of the Φ doublet balances the negative
 310 one of the lepton field. Following spontaneous breaking of the electroweak
 311 symmetry, Eq. 1.27 can be expanded in two parts:

$$\begin{aligned} \frac{g_\ell}{\sqrt{2}} (\bar{\nu}_{\ell,L}, \bar{\ell}_L) \begin{pmatrix} 0 \\ v+h \end{pmatrix} \ell_R + h.c. &= \frac{g_\ell}{\sqrt{2}} (v+h) (\bar{\ell}_L \ell_R + \bar{\ell}_R \ell_L) \\ &= \frac{g_\ell}{\sqrt{2}} (v+h) \bar{\ell} \ell = m_\ell \bar{\ell} \ell + \frac{m_\ell}{v} h \bar{\ell} \ell. \end{aligned} \quad (1.28)$$

312 The first term is a mass term for the charged lepton field ℓ , while the second
 313 is an interaction term between the massive lepton and the H field. Note that
 314 neutrinos remain massless in this construction, a feature whose discussion is
 315 beyond the scope of this work. The same procedure can be followed to give
 316 mass to quarks. The coupling constant of the resulting interaction between

Table 1.1: Masses of the fields contained in the SM. Precise values and their uncertainties can be found in [10].

Field	Mass [GeV]	Field	Mass [GeV]
u	$2.2 \cdot 10^{-3}$	d	$4.7 \cdot 10^{-3}$
c	1.27	s	0.096
t	173.2	b	4.18
e	$0.511 \cdot 10^{-3}$	ν_ℓ	~ 0
μ	$105.7 \cdot 10^{-3}$	τ	1.78
γ	0	g	0
$W^\pm$	80.4	Z	91.2

the H boson and a fermion (quark or charged lepton) is proportional to the mass of the fermion, meaning that the H boson will interact more strongly with more massive particles. Note that the mass of each fermion is a free parameter of the model (being proportional to the Yukawa coupling g_ℓ), and has to be determined by experiment. Currently accepted values are given in Table 1.1. Because this interaction is not the result of imposing a gauge symmetry on the lagrangian, it cannot be identified with neither the strong nor the electroweak interactions, and can in a sense be regarded as a new force of Nature, although with a peculiar origin. Recalling that the H boson also interacts with the massive vector bosons and itself, the only scalar in the SM interacts with every massive particle, making its phenomenology uniquely rich. Moreover, the nature of the H boson's couplings to other SM particles is unique in different ways, since the couplings with vector bosons stem from the symmetry breaking mechanism itself, and those to fermions constitute the only Yukawa-type interaction present in the SM. Probing these couplings is thus a stringent test of solidity of the SM, as well as a potential window on physics not yet encoded in the SM, since it is reasonable to expect additional, yet unseen, massive particles (e.g., a candidate for dark matter) to interact with the H boson in a similar way.

1.3 Phenomenology of the Higgs boson at the LHC

The unique way the Higgs field interacts with quarks, charged leptons, and weak gauge bosons, gives rise to a very rich phenomenology in the scalar sector of the SM. In a sense, the dynamics induced by the Higgs field can be regarded as an additional fundamental interaction, with the peculiarity

that it does not arise from a gauge structure. This means that the coupling strength is not quantized in the usual manner, as made evident by the fact that it is proportional to the (squared) mass of the (boson) fermion the H boson is interacting with, as discussed in the previous section. This fact imposes a clear hierarchy on the interactions in the scalar sector of the SM: the H boson interacts more strongly the larger the mass of the particle it is interacting with. This is reflected both in the way a H boson can be produced at a hadron collider such as the LHC, as well as in its decay branching ratios (BRs).

Production of a H boson in a proton-proton (p-p) collider such as the LHC is regulated by an interplay between the physics of the SM's scalar sector and the structure of the proton. At the energies achievable at the LHC, the internal structure of the proton is completely resolved, meaning that the scattering can effectively be thought of as occurring between two partons i.e., the constituents of the proton. These are quarks and the gluons through which they interact. Each carries a fraction of the 4-momentum of the proton, typically parametrized through the Bjorken variable x as:

$$p_{parton} = xp_{proton}. \quad (1.29)$$

By construction x is bound to the $(0, 1)$ interval. The first consequence of Eq. 1.29 is that the center-of-mass energy of the elementary scattering $\sqrt{s}$ is lower than the collision energy of the protons:

$$\sqrt{s}_{parton} = \sqrt{x_1 x_2 s}, \quad (1.30)$$

where x_1 and x_2 are the 4-momentum fractions of the scattering partons from the two colliding protons. These are stochastic variables whose distributions are commonly referred to as parton distribution functions (PDFs). Each type of parton has its associated PDF $f_i(x)$, where i runs over all flavors of quarks as well as gluons. Because these are fractions of the 4-momentum of the proton, they are bound to satisfy the condition:

$$\sum_i \int_0^1 x f_i(x) dx = 1. \quad (1.31)$$

In this picture no parton carries 4-momentum in any direction except the one the proton is moving in. In the case of p-p collisions at the LHC, where the proton momentum is aligned with the beam axis, this means that the same is true for the momenta of all partons it contains, a fact that will be relevant later. More formally, the notion that the p-p scattering reduces to a parton-parton scattering at high enough energies is enclosed in the QCD

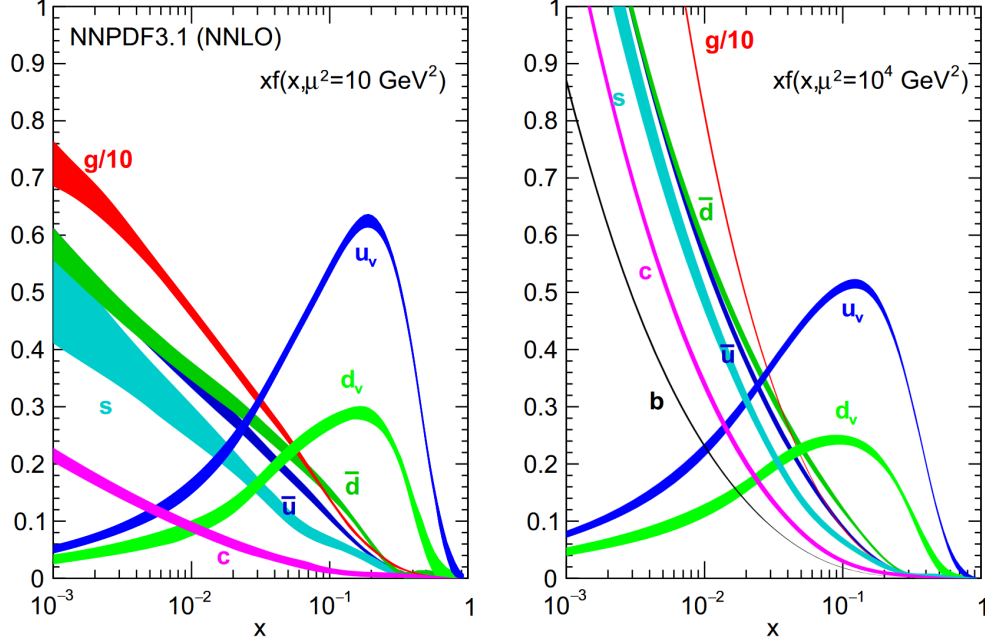


Figure 1.2: Parton distribution functions evaluated at factorization scales $\mu_F^2 = 10 \text{ GeV}^2$ (left) and $\mu_F^2 = 10^4 \text{ GeV}^2$ (right) [15]. The gluon contribution is scaled down by a factor 10 for better visualization.

factorization theorem [11], stating that the cross section for a p-p collision producing some final state F can be factorized as:

$$\sigma(pp \rightarrow F) = \sum_{a,b} \int dx_1 dx_2 f_a(x_1, \mu_F^2) f_b(x_2, \mu_F^2) \hat{\sigma}_{a,b \rightarrow F}. \quad (1.32)$$

Here the indices a and b run over all partons that can initiate the process of interest, μ_F is the factorization scale, and $\hat{\sigma}_{a,b \rightarrow F}$ is the parton level cross section for the process of interest. The evolution of the parton distribution functions between two factorization scales is governed by the DGLAP equations [12]. Eq. 1.32 encodes the idea that the dynamics are well separated between the low energy scale of the bound proton state, and the high energy of the parton-parton scattering, whose cross section can be calculated perturbatively. The set of parton distribution functions used for this analysis has been determined by precisely combining data from the LHC, HERA, and Tevatron colliders [13, 14, 15], and an example is shown in Fig. 1.2. Note that while the contributions of the u and d quarks (the so-called valence quarks) are dominating at high x , the gluon contribution is by far the most important one at lower values. This fact plays a crucial role in the phenomenology

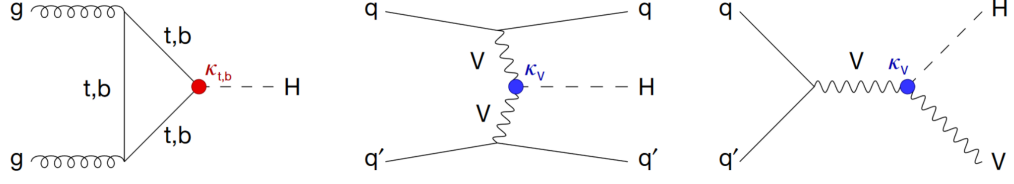


Figure 1.3: Leading order Feynman diagrams for the main production modes of a H boson at the LHC. From left to right: gluon fusion, vector boson fusion, and associated production with a vector boson. The κ_i labels indicate coupling modifiers, which will be explained in detail in Section 6.2.

of H boson production, whose leading diagrams are shown in Fig. 1.3. The three most probable (in terms of total cross section) means of producing a H boson at the LHC are gluon fusion (ggH), vector boson fusion (VBF) and associated production with a vector boson (VH). These are the processes of interest in this analysis.

1.3.1 Gluon fusion production

The most probable means of producing a H boson at the LHC is through gluon fusion. Because the gluon is massless, it has no direct coupling with the H boson. Instead, the process is mediated by a loop of quarks, with the t and b quarks giving the dominant contributions. This reflects the stronger coupling of the H boson to particles with larger mass. The ggH cross section at the LHC, with a p-p collision energy of 13 TeV, amounts to approximately 48.3 pb [16] and accounts for over 90% of the total cross section. At leading order there is no particle produced besides the H boson, although at higher orders in QCD additional gluons and quarks can be produced by initial state radiation. These manifest as hadronic jets in the detector, as will be discussed in Section 1.4.

1.3.2 Vector boson fusion production

The second production mode in terms of cross section is vector boson fusion. Two quarks from the colliding protons radiate a pair of W or Z bosons, which fuse to produce a H boson. The total cross section at the LHC at $\sqrt{s} = 13$ TeV is approximately 3.77 pb [17], one order of magnitude below the ggH equivalent. This fact alone makes measuring VBF independently a challenge, since its distinction from ggH (with additional jets from QCD radiation) is not trivial; it is however of great interest to disentangle the two processes, since ggH proceeds through the Yukawa interaction with fermions

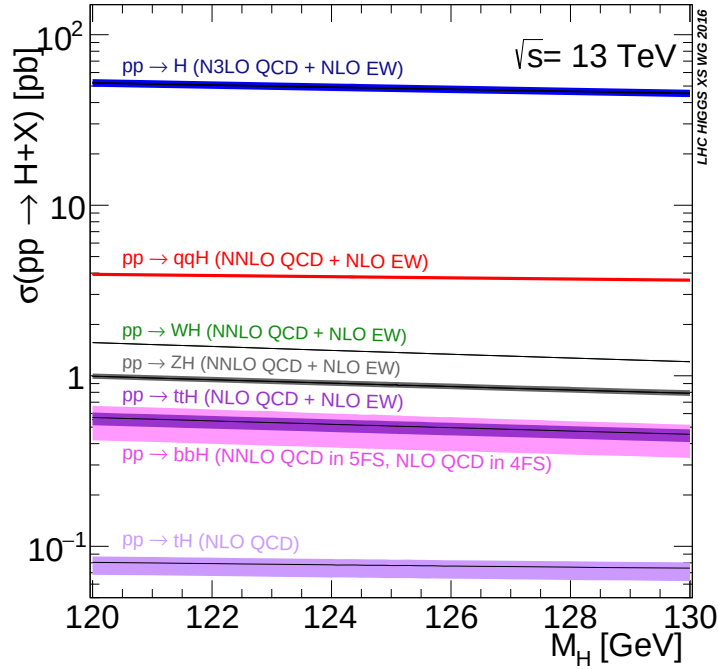


Figure 1.4: Summary of the total cross sections of different production modes of a H boson at the LHC, for $\sqrt{s} = 13$ TeV, as a function of the H boson's mass [17].

while VBF involves couplings to vector bosons. The two processes therefore probe different aspects of the scalar sector of the SM.

1.3.3 Associated production with a vector boson

The final production process of interest for this analysis is the associated production of a H boson with a vector boson V , standing for either W or Z . The total cross sections at the LHC for $\sqrt{s} = 13$ TeV are approximately 1.36 pb and 0.877 pb for WH and ZH production, respectively [17]. While the cross sections are both lower than VBF, this fact is balanced out to a certain degree by the presence of the additional vector boson in the final state, which reduces backgrounds significantly. Additional modes of production that are experimentally accessible at the LHC are associated production with a pair of t (ttH) or b (bbH) quarks. These are however not the direct target of this analysis, since dedicated measurements by the CMS collaboration exist [18]. The total cross sections for the different production modes at the LHC are summarized in Fig. 1.4

1.3.4 Decay of the Higgs boson

With its uniquely rich interaction landscape, the H boson has a very short lifetime of about 1.6×10^{-22} s [17], meaning that once produced it effectively decays at the point of interaction. The H boson is thus only detectable indirectly through its decay products, which can be both fermions as well as vector bosons. The decay width to a pair of fermions can be calculated to be:

$$\Gamma(H \rightarrow f\bar{f}) \propto m_H \frac{m_f^2}{m_W^2} \left(1 - \frac{4m_f^2}{m_H^2}\right)^{3/2} \cdot N_C(f), \quad (1.33)$$

where $N_C(f)$ is equal to 1 for leptons and 3 for quarks. Taken as a function of the fermion mass, this decay width is monotonically increasing as long as $m_f < m_H/2$. Excluding the t quark, since $m_t > m_H$ forbids the decay to a $t\bar{t}$ pair, this means that BRs follow the mass hierarchy of fermions, starting with b quarks. For decays to W or Z bosons the picture is complicated by the fact that Nature chose a value of m_H lower than $2m_W$, meaning that if the H boson is produced on shell, one of the vector bosons will be pushed off shell. Finally, it is worth noting that the H boson can also decay to photons and gluons: although there is no direct coupling with the corresponding fields (since the photon and gluon are massless), the decay can take place through a loop of quarks or vector bosons, in complete analogy with the ggH production mechanism. Diagrams for the various decay modes are given in Fig. 1.5. The precise values of the BRs of the H boson are shown in Fig. 1.6. The most probable decay is that in a pair of b quarks; experimentally however, this final state suffers from a large background from QCD scattering in p-p collision. Although the decay to b quarks has been observed [19], the most experimentally sensitive decay channels are those with either photons or charged leptons with high momentum in the final state. This includes the decay to a pair of W bosons, the target of this analysis, with a BR of 22.0% [20], and whose phenomenology will be discussed in more detail in Section 3.1.

1.4 The strong interaction at hadron colliders

Even though the electroweak and Higgs sectors of the SM are the main focus of this work, at a hadron collider the strong interaction is bound to play an important role. Once again, a detailed treatment of QCD is well beyond the scope of this thesis; I will present only the specific aspects relevant to this analysis.

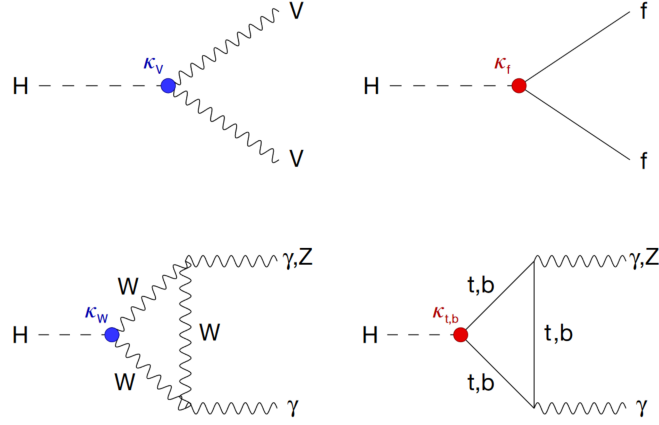


Figure 1.5: Leading order Feynman diagrams for the main decay modes of a H boson to a pair of massive vector bosons (upper left), a pair of fermions (upper right), and a pair of photons mediated by a W (lower left) or heavy quark loop (lower right).

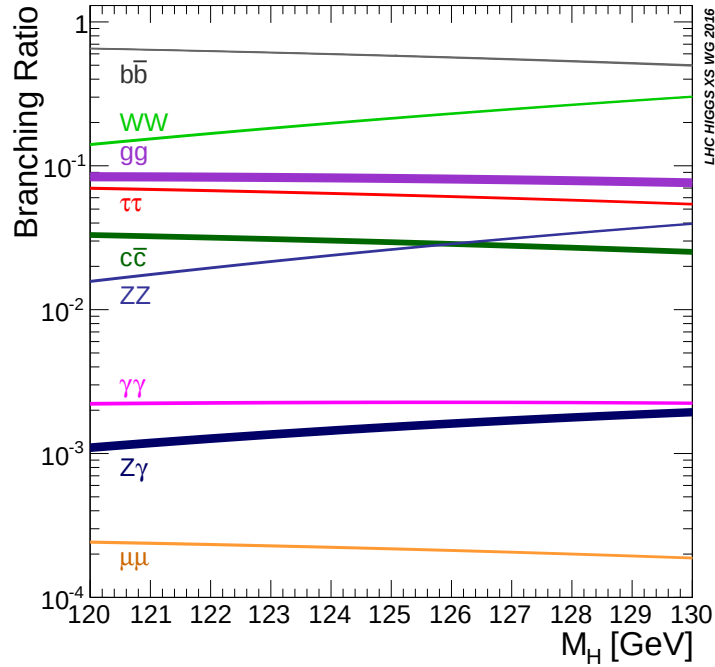


Figure 1.6: Branching ratios of the H boson as a function of its mass [17].

464 1.4.1 The running of the strong coupling constant

465 One of the defining factors of QCD when compared to the electroweak inter-
 466 action is how the coupling α_s changes with the energy scale of the process in
 467 question. This is common in quantum field theories, and is in fact true also
 468 of the electroweak coupling α_{EW} . The crucial difference is that while α_{EW}
 469 increases with the energy scale q of the interaction, α_s decreases:

$$\frac{d\alpha_s(q^2)}{dq^2} < 0. \quad (1.34)$$

470 This feature is known as asymptotic freedom [21], and has the net effect
 471 of reducing the coupling strength of the strong force when the interaction
 472 energy becomes large. In other words, quarks interact more weakly at high
 473 energies, making a perturbative approach feasible for calculating observables,
 474 and more strongly at low energies, where the theory enters a non-perturbative
 475 regime. Arguably the most important experimental consequence of this be-
 476 havior is the phenomenon commonly known as color confinement [22]. In
 477 a (very) simplified view, when a pair of quarks is moved apart, the self-
 478 interacting property of gluons (Fig. 1.7 left) causes the field between them
 479 to form a flux tube, whose cross section varies very little when the distance
 480 between the quarks grows (Fig. 1.7 center). This means that the energy
 481 density of the flux tube is roughly constant, in turn meaning that the total
 482 energy of the flux tube increases quasi-linearly with distance (as shown in
 483 Fig. 1.8), instead of dropping off in the way the static electromagnetic po-
 484 tential does when two charged particles are moved apart (Fig. 1.7 right). At
 485 a threshold then, it becomes energetically favorable to extract a $q\bar{q}$ pair from
 486 the vacuum and break the flux tube, forming two bound states of quarks.
 487 Experimentally, this has the major consequence that no free quark can
 488 be observed, since it will always be energetically favorable to form a color-
 489 less bound state, called a hadron. In a high energy environment moreover,
 490 partons are observed in the form of hadronic jets.

491 1.4.2 Hadronic jets

492 When a high energy colored particle is produced at the LHC, the effect of
 493 color confinement is the production of a so-called hadronic jet. Because of
 494 the confinement property of QCD, briefly mentioned in the preceding section,
 495 the emitted parton will extract other colored particles from the vacuum.
 496 This will produce a shower of particles, sharing the initial momentum of the
 497 parton produced in the hard scattering, and thus being collimated along the
 498 direction it was emitted in. Once the energy of individual partons in the

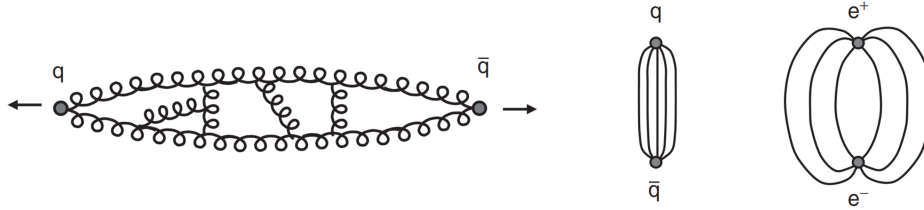


Figure 1.7: Schematic view of the formation of a flux tube of the gluon field between two quarks (left, center), in contrast to the behavior of the electromagnetic field between two electric charges (right) [23].

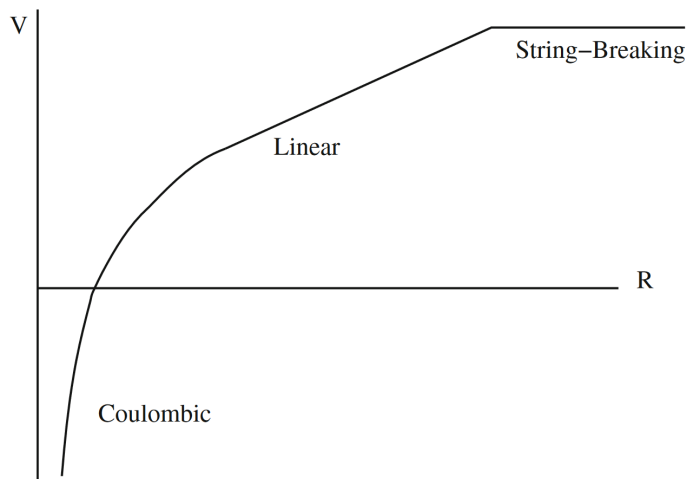


Figure 1.8: Qualitative view of the strong potential between two quarks as a function of their separation [22].

shower becomes low enough, the non-perturbative regime of QCD is reached and hadrons are formed. The net result is that the experimental signature of a high energy colored particle is a collimated spray of hadrons, called a jet. Jets may also contain leptons and photons, originating from the decay of unstable hadrons.

Special care has to be taken when deciding how to cluster the individual hadrons into a jet, both theoretically and experimentally. The prime reason is the infrared and collinear divergence of QCD. In short, when calculating a QCD diagram at fixed order, adding the emission of a gluon as either initial or final state radiation results in cross sections that diverge for either low energy (infrared divergence) or low emission angle (collinear divergence) of the gluon. Although this diverging behavior is a mathematical artefact, it is crucial that the operational definition of a jet be robust with respect to these divergences. A similar argument can be made on the experimental side, since two particles emitted too close to each other will be detected as a single particle because of the finite resolution of the detector. Arguably the simplest example of an observable that is not robust in this regard is the number of particles in a jet. Jets therefore have to be defined in such a way to be collinear and infrared safe. An observable $O(p_1, \dots, p_N)$, function of the momenta of N particles, is said to be infrared safe if:

$$O(p_1, \dots, p_N) = O(p_1, \dots, p_N, \varepsilon) \text{ for small } \varepsilon^2, \quad (1.35)$$

meaning that it is left unchanged by the addition of an extra particle with low momentum. Similarly, O is said to be collinear safe if:

$$O(p_1, \dots, p_N) = O(p_1, \dots, p_{i,1}, p_{i,2}, \dots, p_N) \text{ with } p_{i,1} + p_{i,2} = p_i, \quad (1.36)$$

meaning that O does not change if the i -th particle is replaced by two particles. In this analysis, the anti- k_t algorithm is used [24], and will be detailed in Section 2.4.

Chapter 2

The CMS detector at the LHC

2.1 The large hadron collider

The large hadron collider (LHC) is the largest and most powerful particle accelerator active in the world at the time of writing. Located at an average depth of around 100 meters underground at the CERN facilities in the Geneva area, it spans a 27 km circumference across the french-swiss border. The LHC is the final step of a multi-tiered accelerator complex, shown in Fig. 2.1. The machine is capable of accelerating beams of protons or heavy ions in opposite directions, producing ion-ion, ion-proton and proton-proton collisions. Collisions involving heavy ions are used mainly for the study of quark-gluon plasma dynamics, while p-p collisions are used to probe fundamental interactions and are the focus of this work. The reason p-p collisions are more suitable for studying elementary SM scattering processes is the much higher center-of-mass energy of the nucleon-nucleon interaction that can be reached in this configuration, with the peak energy allowed by the design of the accelerator of 14 TeV. The highest energy achieved during actual operation of the LHC is 13.6 TeV, although during the 2016-2018 time frame (during which the data set which this analysis exploits was gathered) the accelerator ran at an energy of 13 TeV. The beams are made to collide in four points along the circumference, where the four main experiments are located:

- ATLAS [25] (A Toroidal Lhc ApparatuS) and CMS [26] (Compact Muon Solenoid) are multi-purpose experiments, located at opposite points on the collider. They are the two detectors at the LHC capable of probing a wide gamut of phenomena. They are designed to cover the largest portion of solid angle possible around the interaction point, with focus on the region orthogonal to the beam axis. Although similar in

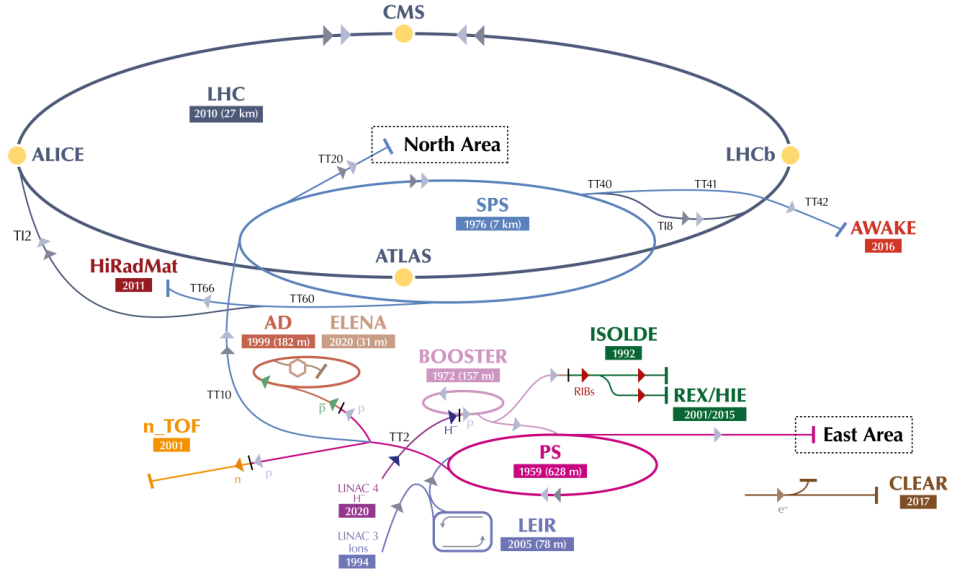


Figure 2.1: Schematic view of the CERN accelerator Complex, with the LHC as its final stage.

scope and goal, they are independently designed, so that any observation claimed by one of the experiments can be confirmed or disproved by the other. Both experiments explore a vast range of phenomena, including but not limited to physics of the H boson, measurement of properties of the t quark and vector bosons, and study of rare decays of heavy hadrons. Both experiments also pursue a comprehensive program to search for physics beyond the SM (BSM), both directly (via searches for yet unseen particles) as well as indirectly (through deviations from the SM in precision measurements).

- LHCb [27] (LHC beauty) is focused on the study of flavor physics and CP violation. To achieve this it is designed to precisely measure the decays of B-type hadrons. Its geometry is peculiar with respect to the other LHC experiments, as it focuses on the solid angle portion along the direction of one of the beams.
- ALICE [28] (A Large Ion Collider Experiments) focuses on ion-ion collisions, with the primary aim of studying the quark-gluon plasma state of matter.

Acceleration of the beams is achieved through the use of radio frequency cavities that generate an oscillating electric field. In order to keep the beams

on a circular trajectory and to focus them, a system of over 1200 dipole magnets and 8000 higher order magnets is employed. The LHC is not exactly circular; rather it is built by alternating straight and curved sectors. The straight sectors are where the beams are accelerated by cavities or focused by quadrupole and octupole magnets. The curved sectors are where the beams are bent using dipoles. A simple relation between a particle's momentum p and the radius of its curvature in a magnetic field is given by:

$$\rho[\text{m}] = \frac{p[\text{GeV}]}{0.3 \cdot B[\text{T}]} \quad (2.1)$$

This means that to keep beams within the LHC circumference at an energy of 6.5 TeV, very strong magnetic fields of the order of 8 T are required. This is only achievable by cooling the magnets to ~ 2 K, a temperature at which they become superconductive and therefore able to sustain the strong currents needed to generate these magnetic fields. One peculiarity of the LHC is that it aims to simultaneously accelerate beams of positive particles in two opposite directions, requiring opposite magnetic fields for the beams travelling in the two directions. This is achieved with a particular design of the dipole magnets, capable of inverting the 8 T magnetic field over the span of a few cm, as shown in Fig. 2.2. The beams are then made to cross and collide at the desired interaction points.

One of the most important performance metrics of a particle accelerator is its instantaneous luminosity $\mathcal{L}$. This quantity is defined by the relation:

$$R = \sigma \mathcal{L}, \quad (2.2)$$

where R is the rate of occurrence of a given process with cross section σ . This quantity effectively measures how efficient the collider is in producing collisions and thus directly determines the statistical power of the data set it is able to yield over a given running time. The instantaneous luminosity can be parameterized in terms of quantities related to the beam structure. Particles in the LHC travel in bunches, which are made to collide in the interaction points, as shown in Fig. 2.3. Given the transverse area a of the beams at the interaction point, the number of particles in the colliding bunches n_1 and n_2 , and the bunch collision frequency f , the instantaneous luminosity scales as:

$$\mathcal{L} \propto f n_1 n_2 / a. \quad (2.3)$$

The instantaneous luminosity is usually given in units of $\text{cm}^{-2}\text{s}^{-1}$. Integrating the instantaneous luminosity over a given data taking period gives the integrated luminosity:

$$\mathcal{L}_{int} = \int_{T_{data}} \mathcal{L}(t) dt, \quad (2.4)$$

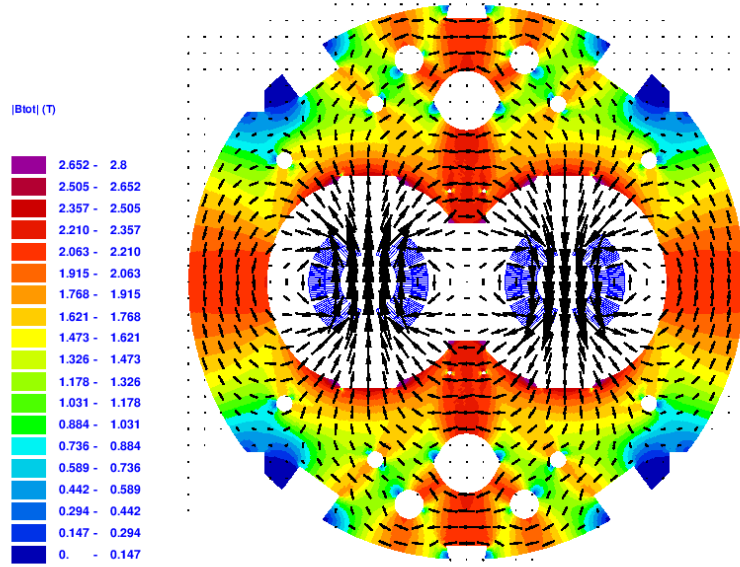


Figure 2.2: Section of an LHC dipole magnet. The design is such that the magnetic field is opposite in the two beam pipes, allowing for two proton beams to circulate in opposite directions.

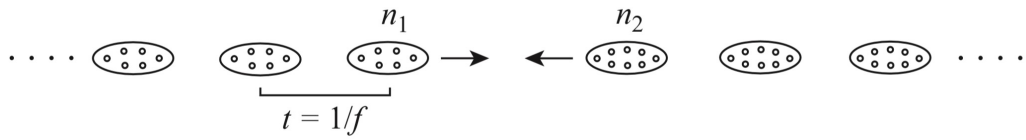


Figure 2.3: Schematic of a bunch collision in the LHC.

which is directly related to the expected number of occurrences of a process in the data set by the relation:

$$N = \sigma \cdot \mathcal{L}_{int}. \quad (2.5)$$

Given the typical values of cross sections in particle physics applications, units of cm^{-2} are not convenient. At the LHC, integrated luminosity is therefore usually given in fb^{-1} , where $1 \text{ fb}^{-1} = 10^{39} \text{ cm}^{-2}$. The peak luminosity reached by the LHC during the 2016–2018 data taking period, referred to as Run 2, is approximately $2 \cdot 10^{34} \text{ cm}^{-2}\text{s}^{-1}$. Precise measurement of $\mathcal{L}$ is a key factor in the precision of cross section measurements, as these are typically obtained by measuring the number of occurrences of the process and inverting Eq. 2.5. In CMS, luminosity is measured through an array of methods, including counting hits in the muon system, energy deposits in the forward calorimeter, energy clusters in the pixel tracker, as well as dedicated detectors [29]. Because the protons are grouped in bunches, at every bunch crossing multiple interactions occur. The number of interactions per bunch crossing is referred to as pileup (PU) and depends on the number of protons per bunch and conditions of the beam. High PU conditions can spoil the performance of the detector, since a large number of tracks can degrade the efficiency in disentangling nearby interaction points from each other. When a hard interaction (i.e., an interaction with high momentum transfer between partons) occurs, the colliding protons are dismembered. The remnants of the protons give rise to what is called the underlying event (UE).

Hard scattering processes only account for a part of the total pp cross section at the LHC. Other processes, called diffractive, can occur. Such interactions are mediated by the exchange of a colorless object: if both protons remain intact following the interaction, it is called elastic or central diffractive; if one or both of the protons are dismembered it is called single or double diffractive. The total pp cross section at the LHC is $\sim 100 \text{ mb}$ [30], with a sizable $\sim 30 \text{ mb}$ coming from diffractive processes.

The collider setup provides an intrinsic cylindrical symmetry to the collision. It is therefore natural to use a cylindrical-type coordinate system to describe events, with the z axis coinciding with the beam direction. One complication however arises in the case of hadronic colliders: because the interaction happens between partons, each carrying a random fraction of the proton's momentum, the center of mass of the collision event differs from the detector one by an unknown Lorentz boost along the beam axis. For this reason, collision events are described through a set of Lorentz invariant observables, effectively bypassing this effect. The most commonly utilized are:

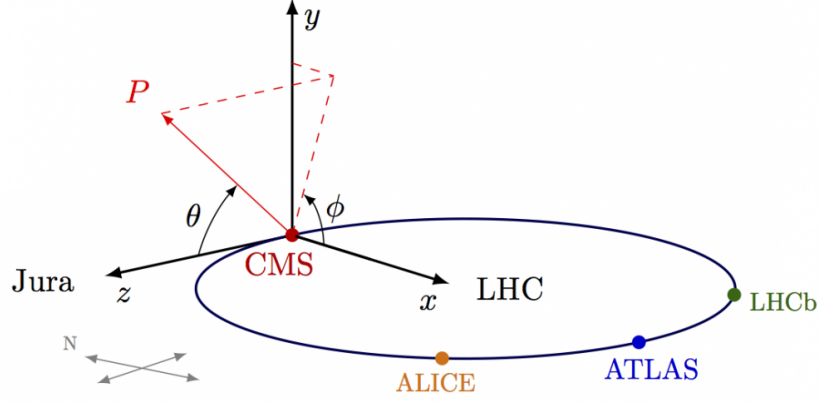


Figure 2.4: The coordinate system used in CMS.

- the transverse momentum $\vec{p}_T$, namely the projection of the spacial momentum $\vec{p}$ on the plane orthogonal to z (transverse plane);
- the azimuthal angle ϕ , also defined in the transverse plane;
- the rapidity interval between two objects Δy . Rapidity of a particle is defined as:

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right), \quad (2.6)$$

intervals of which are Lorentz invariant. In ultra-relativistic regimes, rapidity can be approximated by pseudo-rapidity:

$$y \xrightarrow{\beta \sim 1} \eta = -\ln \left(\tan \frac{\theta}{2} \right), \quad (2.7)$$

where θ is the angle between the particle momentum and the z axis. In this sense, η can be seen as a re-parametrization of the usual polar angle in spherical coordinates.

The coordinate system is visualized in Fig. 2.4. Another variable of crucial importance is the missing transverse momentum, defined as:

$$\vec{p}_T^{\text{miss}} = - \sum_i \vec{p}_{T,i}. \quad (2.8)$$

If all products of a collision were detected, this quantity would vanish since the protons carry no momentum on the transverse plane. This is however not the case: neutrinos for example are too weakly interactive to be directly detected, resulting in a non vanishing p_T^{miss} . This quantity is also of great importance for searches beyond the SM, as many theories predict particles (e.g. dark matter candidates) that would escape direct detection.

2.2 The Compact Muon Solenoid

The CMS is one of the four major experiments installed at the LHC. In 2012, together with ATLAS, the CMS collaboration reported the observation of a new scalar resonance, compatible with the SM H boson. Although this observation was the driving reason behind its construction, CMS is designed to pursue a vast range of physics studies. At the time of writing, Run 3 of the LHC has started, with a number of upgrades to various CMS systems designed to tackle the new data taking period. Given that this work is based solely on data gathered in Run 2, the detector will be described as it was set up during the 2016–2018 period, omitting recent modifications and additions. Built with a cylindrical geometry coaxial to the beam pipe, the detector measures approximately 22 m in length and 15 m in diameter, as shown in Fig. 2.5. It is comprised of a complex system of subdetectors, able to detect electrons, photons, muons and hadronic jets produced in collisions. The coordinate system in use is a hybrid reference frame (r, η, ϕ) , where r is the distance from the beam axis (which also defines the z axis) and η, ϕ have been previously defined. The subsystems of CMS can be divided in two main groups: the barrel covers the low η (central) region, and is the closest to the interaction point, while the two endcaps detect particles emitted at high η (low angle with respect to the LHC beam). In the following an overview of the main subdetectors is given, in an inside-out order starting from the center (Fig. 2.6).

2.2.1 The silicon tracker

The tracking system in CMS is designed to precisely reconstruct the trajectories of charged particles produced during collisions. The solenoid magnet (described in the following) produces a strong magnetic field parallel to the z axis, so that charged particles follow helicoidal trajectories with a curvature determined by Eq. 2.1. Reconstructing these trajectories thus allows to measure the transverse momenta of the particles. High precision in track measurement is also required to achieve good vertex reconstruction i.e., to minimize the probability of assigning a track to the wrong vertex. This is especially important to distinguish secondary vertexes, that may arise due to the decay of meta-stable particles, from the primary interaction vertex. One of the challenges in the design of the tracker is the requirement to have good performance over a wide range of momenta: on the one hand, leptons with $p_T \sim 1$ TeV are expected to be produced by new, high mass resonances; on the other hand, good reconstruction efficiency on tracks with p_T as low as 100 MeV is required in order to precisely measure the energy of hadronic

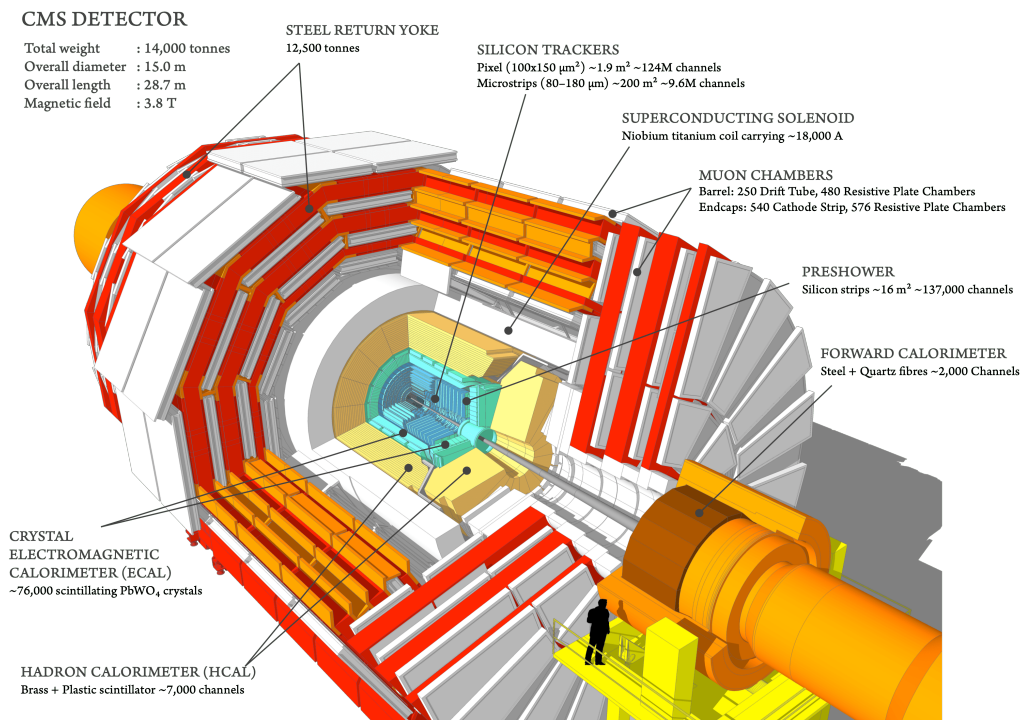


Figure 2.5: Schematic view of the CMS detector during Run 2 of the LHC.

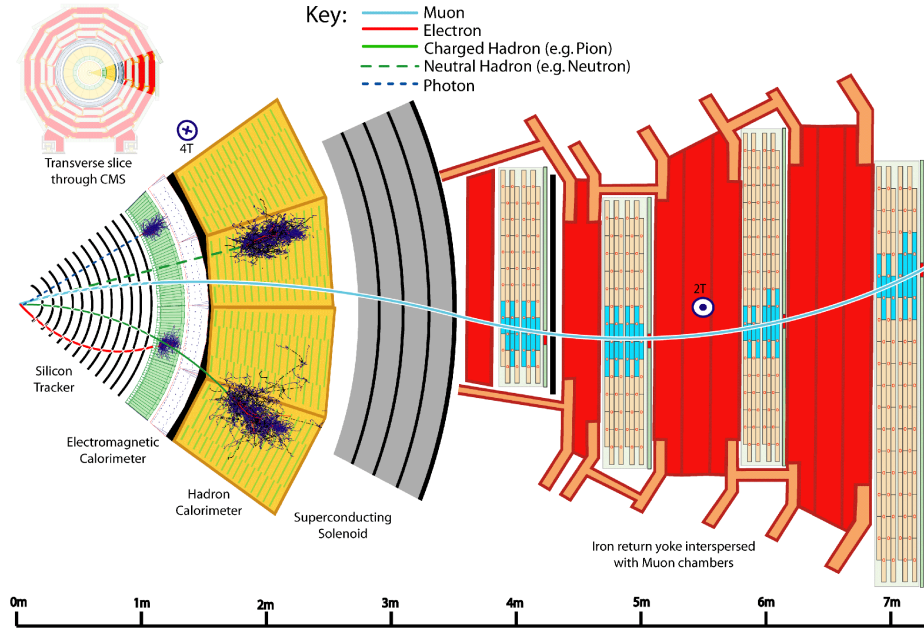


Figure 2.6: Schematic view of a section of CMS orthogonal to the beam direction in the barrel region. Example trajectories for different particles are shown.

jets. Moreover, because the tracker is the core of the CMS detector and thus every particle has to cross it, the amount of material has to be kept as low as possible in order to minimize multiple scattering and photon conversion probabilities. The former spoils momentum measurement in the tracker itself and in the muon system, while the second impacts the precision of energy measurements in the electromagnetic calorimeter.

Fig. 2.7 shows a section of the tracking system. The detector situated closest to the nominal interaction point (starting at $r = 3.0$ cm) is a pixel detector, comprised of four cylindrical layers coaxial to the beam, flanked on each side by three discs orthogonal to the beam. This setup allows for a 3-dimensional measurement of the position of a particle interacting with the tracker: the $r\phi$ and z coordinates are given by the pixel itself, while r is determined by the corresponding layer's position (this is true for the barrel; in the endcaps, z is given by the layer and $r, r\phi$ by the pixel). Each pixel has a pitch of $100 \times 150 \mu\text{m}^2$, for a total of 124 million pixels. Spatial resolution on a single hit varies between $10 \mu\text{m}$ and $40 \mu\text{m}$ depending on the region.

The strip detector system consists of a total of 10 cylindrical layers and three disks on each side in the barrel region, complemented by one set of nine disks in each endcap. Layers indicated in red in Fig. 2.7 house a single

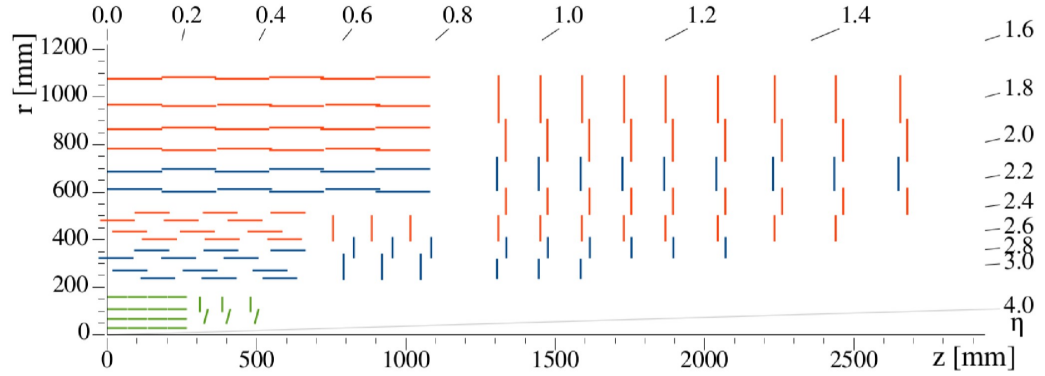


Figure 2.7: Schematic view of a quarter of the tracking system, along the (r, z) plane. Pixel detectors are shown in green, double strip detectors in blue and single strip detectors in red.

strip detector, while layers in blue feature a second set of strip detectors mounted on the rear of the first set, tilted by a 100 mrad angle to allow for a 3-dimensional reconstruction (similarly to the pixel detector). Single hit spacial resolution varies between 10 μm and 50 μm depending on the η of the track.

In order to reduce effects such as leakage currents due to radiation damage, the whole tracker is operated at a temperature of -15°C .

2.2.2 The electromagnetic calorimeter

The electromagnetic calorimeter (ECAL) in CMS is a hermetic, homogeneous detector with a cylindrical geometry. Its role is to identify and accurately measure the energy of electrons and photons. This is achieved using scintillating PbWO_4 crystals, in number of 61200 in the barrel and 7324 in each endcap (Fig. 2.8). The choice of PbWO_4 crystals is motivated by a number of factors: high density, short radiation length and small Molière radius allow for a detector both compact as well as highly granular; resilience to radiation damage grants a long lifetime to the detector; finally, fast scintillation time ($\tau = 10$ ns) allows to gather $\sim 80\%$ of the emitted light between two consecutive bunch crossings. The crystals are read out by avalanche photodiodes (APDs) in the barrel region and vacuum phototriodes (VPTs) in the endcaps. A preshower system is placed before the endcap sections of ECAL, with the goal of separating photons coming from the primary vertex from those coming from the in-flight decay of a π^0 , which could be misreconstructed as a single photon. The energy resolution of a homogeneous electromagnetic calorimeter

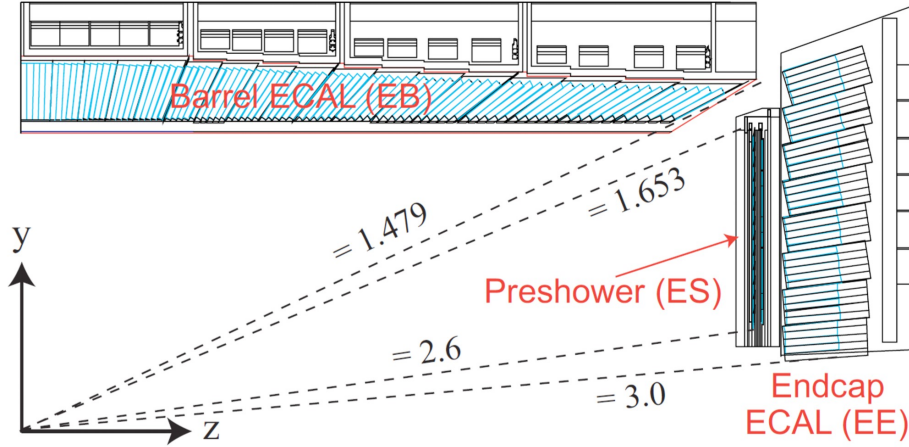


Figure 2.8: Schematic view of a quarter of the ECAL system in the (r, z) plane.

can be expressed as a sum in quadrature of three distinct sources:

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{a}{\sqrt{E}}\right)^2 + \left(\frac{b}{E}\right)^2 + c^2, \quad (2.9)$$

where:

- the first term is the stochastic contribution, accounting for fluctuations in the number of collected photoelectrons. These fluctuations include variations in the number of photons emitted by the crystal, number of photons collected and the quantum efficiency of the photodetector. The fact that this contribution is inversely proportional to $\sqrt{E}$ makes calorimeters very precise at high energies.
- The second term is due to electronic noise in the readout system. The contribution depends on the instantaneous luminosity and on η .
- The third term, which is dominant at high energies, accounts for a number of effects, e.g. longitudinal non-uniformities in light collection, energy leakage from the rear of the crystal and calibration.

The energy resolution for electrons in the barrel region is measured to be:

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{2.8\% \text{ GeV}^{1/2}}{\sqrt{E}}\right)^2 + \left(\frac{12\% \text{ GeV}}{E}\right)^2 + (0.3\%)^2. \quad (2.10)$$

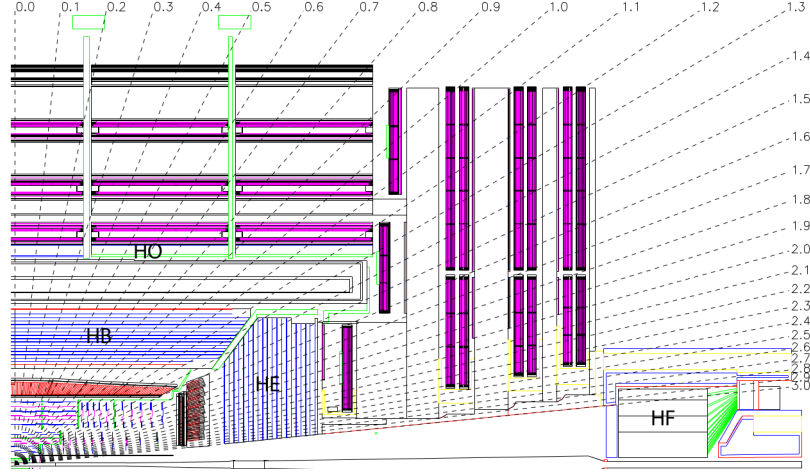


Figure 2.9: Schematic view of a quarter of the HCAL system in the (r, z) plane, showing the barrel (HB), outer (HO), endcap (HE), and forward (HF) regions of the system.

2.2.3 The hadronic calorimeter

The hadronic calorimeter (HCAL) in CMS is a hermetic sampling calorimeter, divided into three distinct regions: the barrel, the endcaps, and the forward region (Fig. 2.9). The HCAL complements the ECAL to build a complete calorimetric system, able to measure energy of electrons, photons and hadronic jets. Its hermetic design also allows indirect measurement of weakly interacting particles (e.g. neutrinos) by means of p_T^{miss} . Because of its placement inside the solenoid, a non magnetic material such as brass has to be used for the absorbing layers of the detector in the barrel and endcap regions. The active layers consist of plastic scintillators paired to wavelength shifters and hybrid photodiodes for readout. In order to maximize the hermeticity of the system, the fraction of absorber material is substantial, with a thickness of 5 cm for absorber layers compared to 3.7 mm for active layers. The forward part of the detector is placed outside the magnet. It consists of alternating layers of iron absorber and Cherenkov emitting quartz fibers. Energy resolution in the barrel and endcap regions can be expressed as:

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{115\% \text{ GeV}^{1/2}}{\sqrt{E}}\right)^2 + (5.5\%)^2. \quad (2.11)$$

2.2.4 The superconducting magnet

At the heart of CMS is the largest superconducting solenoid ever constructed with a length of 12.5 m and a diameter of 6 m. Built out of four Niobium-Titanium layers, at maximum capacity it is capable of generating a 3.8 T magnetic field in the whole volume, resulting in 2.7 GJ of energy stored in the field. To achieve this, the material is rendered superconductive by cooling it to about 4 K through a liquid Helium cooling system. Currents of ~ 18000 A can then be injected into the solenoid, producing the desired magnetic field. Outside the solenoid is the return yoke, made of three sections (barrel and two endcaps) each divided into four layers. The magnetic field in the yoke is about 1.8 T. The muon system is placed in between these layers.

2.2.5 The muon system

The muon system is situated outside the solenoid magnet, exploiting the fact that muons are among the only particles (besides neutrinos) that are able to consistently travel through the rest of the detector without being absorbed. Exploiting the residual field in the return yoke, it complements the tracker in the measurement of muon p_T , and is of crucial importance to the trigger, as will be explained in the following. The system consists of three types of gaseous detectors:

- The barrel region is covered by 4 layers of drift tube (DT) detectors. DTs are capable of precisely measuring the impact point of a charged particle, effectively serving as a complementary tracking system in the outer region of CMS.
- The two endcaps house 4 stations of cathode strip chambers (CSCs) each. These are multi-wire proportional chamber detectors, and their prime characteristic is their excellent spatial resolution on hits of around 100 μm .
- Both the barrel and the endcap regions of CMS are complemented by a total of 1056 resistive plate chambers (RPCs). Apart from providing redundancy to the muon system, they serve the primary purpose of giving very fast hit information that is vital to the trigger system (described in the following section), given their time resolution of the order of a few ns.

A section of the muon system is shown in Fig. 2.10.

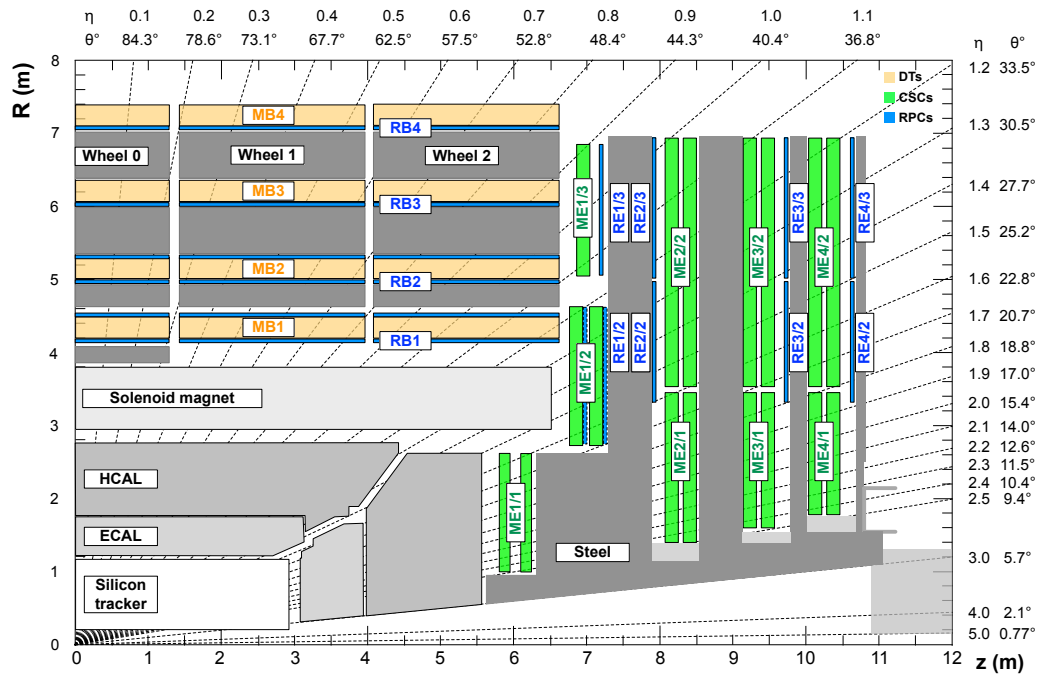


Figure 2.10: Schematic view of a quarter of the muon system. DT detectors form the muon barrel (MB) system; CSC detectors for the muon endcap (ME) system; RPC detectors are located in both the barrel (RB) and endcap (RE) regions.

2.3 The trigger system in CMS

The bunch crossing frequency in LHC is of 40 MHz, and at each crossing multiple interactions occur. This results in an amount of data that is impossible to process and store with currently available technologies. Moreover, only a fraction of the events contains a process of potential interest. The role of the trigger system is to perform an on-the-fly selection on collision events based on an initial set of information, choosing which events to store for analysis and which ones to discard. The trigger in CMS is built on two levels: the level-1 trigger (L1) and the high level trigger (HLT), which together are capable of reducing the rate of events from the 40 MHz crossing frequency to about 1 kHz, a data flux that can then be stored for analysis.

2.3.1 The level-1 trigger

The L1 trigger runs on custom processors, and is designed to make decisions within $3.2 \mu\text{s}$. To do this it has access to information coming from the calorimeters and the muon system:

- calorimeters provide information on so-called towers i.e., clusters of energy measured by both ECAL and HCAL. The readout electronics of the calorimeters is already able to identify electron, photon and jet candidates, providing this information to the L1 system.
- The muon system provides what are called track segments, together with a first estimate of their p_T . It is also able to assign a quality flag to each of these segments, sending the 4 highest quality ones to L1.

This information is combined by the L1 global trigger, which decides whether or not to keep the event based on a set of pre-programmed thresholds, called the L1 trigger table. In case the event is flagged as interesting, the full set of detector subsystems is read out and the event is sent to the HLT. This reduces the event rate from 40 MHz to 100 kHz.

2.3.2 The high level trigger

Contrary to L1, the HLT runs on commercial processors. It reconstructs the event output by the L1 system using information from all detection systems. The reconstruction algorithm in HLT is implemented completely via software, making both hardware as well as software upgrades easier. The HLT selection proceeds by steps: at first only information from calorimeters and muon chambers is used for a coarse selection; the subsequent step adds information

838 from the tracker to further refine the reconstruction; finally, the last step has
839 access to the full information of the detector to make the final decision.
840 This division is crucial, since computation time for an event processing is
841 limited, and this way more complex algorithms are only run on a subset of
842 the events pre-selected by L1. HLT is thus able to reduce the rate to ~ 1
843 kHz. These events are then permanently stored and made available for full
844 offline reconstruction and analysis.

845 2.4 Event reconstruction in CMS

846 After an event has been selected and saved, it has to be processed: the raw
847 information (i.e. tracker hits, calorimeter deposits etc.) have to be combined
848 to produce higher level objects on which data analysis can be more easily
849 performed. This is done by the Particle-Flow (PF) algorithm [31]. The
850 precise tracking system and highly segmented calorimeter are the key features
851 of CMS that allow the PF technique to function. The first step is to convert
852 raw data into higher level objects e.g., combining tracker hits into charged
853 particle tracks and merging calorimeter deposits into clusters. These can
854 then be combined to build the actual physics objects (electrons, muons, jets,
855 etc.), which will be used in analysis. Every track is assigned to a vertex. The
856 vertex with the highest total p_T associated to it is labeled as the primary
857 vertex (PV). The reconstruction happens in the order given below. Once an
858 object has been identified, the corresponding tracks and clusters are removed
859 from the subsequent searches.

860 2.4.1 Muon reconstruction

861 Muons are reconstructed by combining information from the tracker and the
862 muon system [32]. First, so-called standalone muon tracks are built by only
863 using hits in the muon system. These are then matched to corresponding
864 tracks in the central tracker. This matching can happen in two ways:

- 865 • Inside-out: each track in the central tracker is extrapolated to the muon
866 system. If a matching standalone muon track is found, the two tracks
867 qualify for what is called a tracker muon.
- 868 • Outside-in: each standalone muon track is extrapolated to the tracker,
869 and if a matching track is found, the object is registered as a so-called
870 global muon.

871 The two methods combined reach a muon reconstruction efficiency of about
872 99%, with the first being more efficient for low p_T muons emitted in the

central region of the detector, and the second being optimized for higher p_T muons that penetrate more than one muon station. For this analysis, muons are required to be reconstructed as both a tracker muon and a global muon, with hits in at least six layers of the tracker, of which at least one in the pixel tracker, and at least two muon stations. Requiring a hit in the pixel tracker greatly suppresses muons originating from in-flight decays of other particles. Furthermore, the impact parameter of the track with respect to the PV is required to be below 0.02 (0.01) cm for muons with p_T above (below) 20 GeV in the transverse plane, and below 0.1 cm in the longitudinal direction.

Muons are also required to be isolated. The isolation metric is calculated as the ratio of the muon p_T to the total p_T of all other particles found within a radius of 0.4 from the muon in the (η, ϕ) plane. This quantity is required to be below 0.15.

2.4.2 Electron and photon reconstruction

Electrons are reconstructed by combining energy deposits in ECAL with compatible charged tracks in the tracker. A peculiar feature of electrons when traversing the detector is that they are prone to interact with the material in the tracker by emitting bremsstrahlung radiation along their trajectory. To account for this phenomenon, so-called superclusters in ECAL are used for electron reconstruction. A window around the initial cluster is built, spanning a narrow range in η and a more extended one in ϕ , and clusters in the window are grouped into a supercluster. This asymmetric window is designed to gather bremsstrahlung photons emitted by the electron on its helicoidal path in the magnetic field. The ratio of the energy in the supercluster and the momentum of the corresponding track is also used to disentangle electrons from charged hadrons, which deposit less of their energy in the ECAL.

To reject reconstructed electrons from spurious sources, a dedicated identification algorithm is used, taking as inputs a variety of calorimetric, tracking, and isolation variables [33]. These include the ratio of energy deposited in geometrically correspondent ECAL and HCAL clusters, geometric shape of the supercluster, the ratio of the supercluster energy to the track momentum, the relative isolation of the electron (defined in a similar way as for muons), number of hits in the tracker, and various measures of the track reconstruction quality. When a high energy photon traverses the tracker material, it can be converted into an e^+e^- pair. These can be disentangled from prompt electrons (i.e., ones emitted at the interaction vertex) by rejecting candidates whose tracks are missing more than one hit in the innermost layers of the tracker.

912 Remaining ECAL clusters not associated with a track or a matching
 913 HCAL cluster are classified as photons. These are of marginal interest for
 914 this analysis, and more details on their reconstruction and identification can
 915 be found in [33].

916 2.4.3 Hadron reconstruction

917 Once tracks and calorimetric deposits identified as muons, electrons, or pho-
 918 tons are removed from the event, the remaining energy deposits are recon-
 919 structed as either charged or neutral hadrons. Energy clusters in ECAL and
 920 HCAL that match geometrically are assumed to originate from the same
 921 particle.

922 If one or more tracks that point to the calorimetric deposit are found,
 923 they are classified as charged hadrons. The energy of the calorimetric cluster
 924 is compared to the sum of the momenta of all compatible tracks. If they
 925 match within the expected resolution, it is assumed that the energy of the
 926 cluster is only due to charged hadrons, the energy of which is then extracted
 927 by simultaneously fitting the track momenta and the energy fractions in
 928 the calorimeter cluster. This combination ensures good energy resolution
 929 across the p_T spectrum, since at very high p_T the calorimeter will exceed
 930 the performance of the tracker (whose momentum measurement is degraded
 931 when the curvature of the track is too low). If the energy of the calorimetric
 932 cluster is larger than the sum of track momenta, the excess energy deposits in
 933 ECAL and HCAL are identified as photons and neutral hadrons, respectively.

934 2.4.4 Jet reconstruction

935 After every particle in the event has been identified, jets are reconstructed
 936 by clustering PF objects. As anticipated in section 1.4, clustering algorithms
 937 are required to provide jet definitions which are collinear and infrared safe.
 938 In this analysis jets are defined by the anti- k_T algorithm [24]. Given a set of
 939 N tracks, two kinds of distance are calculated:

$$d_{i,B} \equiv p_{T,i}^{-2}, \quad (2.12)$$

940 and

$$d_{ij} \equiv \min\{p_{T,i}^{-2}, p_{T,j}^{-2}\} \left(\frac{\Delta_{ij}^2}{R^2} \right). \quad (2.13)$$

941 Eq. 2.12 measures the distance between object i and the beam, while Eq. 2.13
 942 measures the distance between the pair of objects i and j . The quantity Δ_{ij}^2 is

the distance measured on the (y, ϕ) plane (y being the rapidity, as previously defined):

$$\Delta_{ij}^2 \equiv (y_i - y_j)^2 + (\phi_i - \phi_j)^2, \quad (2.14)$$

while R is a parameter to be chosen for optimization, and is set to 0.4 in this analysis. The algorithm then proceeds as follows:

- distances $d_{i,B}$ and d_{ij} are calculated for each particle and for each particle pair, respectively;
- the smallest distance is found. If it is of type d_{ij} , objects i and j are combined into a single object summing their 4-momenta and the previous step is repeated;
- as soon as the smallest distance is found to be of type $d_{i,B}$, object i is defined as a jet and removed from the list of input particles;
- the algorithm restarts, until no particles are left.

This algorithm is both infrared and collinear safe: if a soft hadron is added to the initial set of particles, the same jet is found, and the same is true if a particle is replaced by two collinear ones. Moreover, the anti- k_T algorithm starts the clustering procedure from the highest p_T objects, resulting in jets whose shape in the (y, ϕ) plane is much smoother when compared to other algorithms (Fig. 2.11).

The reconstructed energy of a jet does not, on average, match the energy of the parton that initiated it. This is primarily due to two factors: nonlinearities in the response of the detector, and PU. The presence of PU is a crucial factor in achieving the high instantaneous luminosity of the LHC: with the bunch crossing frequency fixed, having multiple interactions per crossing increases the possibility of an event of interest occurring in each crossing. The downside is a more experimentally challenging environment, since the number of particles interacting with the detector will be much larger, as shown in Fig. 2.12. This has a direct effect on the measured energy of jets, since particles produced in PU collisions can enter the clustering of a jet coming from the PV. PU therefore manifests as an upwards shift of the measured energy of jets, which needs to be corrected for. The PU correction used on jets for this analysis proceeds in three steps [34]:

- Charged particles whose tracks originate from a PU vertex are removed from the clustering sequence. This procedure is known as charged hadron subtraction, and corrects for around 50% of the energy offset from PU.

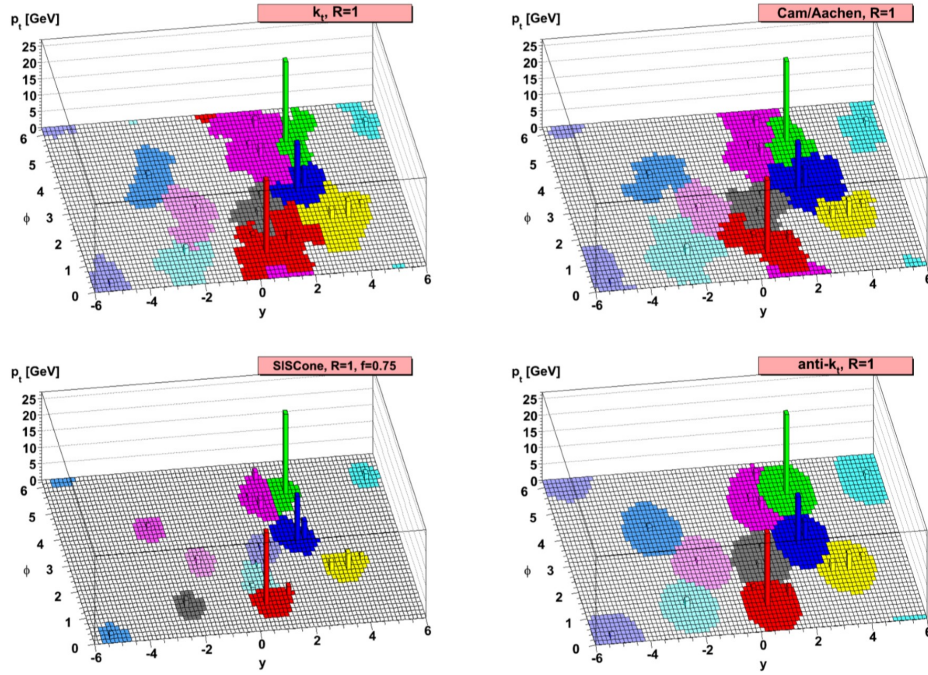


Figure 2.11: Comparison between different jet reconstruction algorithms in the (y, ϕ) plane. Notice how jets defined through the anti- k_T algorithm present a more regular i.e., cone-like shape.

- 977 • A simulated QCD sample is used to estimate the average energy density
 978 ρ due to PU interactions as a function of η and p_T of the jet. This is
 979 achieved by simulating the same events with and without additional
 980 PU vertices and comparing the reconstructed energies. This density is
 981 then used to estimate and subtract the the remaining PU offset from
 982 the jet based on its area in the (η, ϕ) plane. This second step corrects
 983 mostly for neutral particles from PU (which cannot be assigned to an
 984 interaction vertex), as well as remaining charged particles that could
 985 not be associated to an interaction vertex.

- 986 • Finally, an additional scale factor is applied to correct for residual dis-
 987 crepancies between data and MC. This is done on samples recorded
 988 with the only trigger requirement being that a bunch crossing has oc-
 989 curred (zero bias sample). The reason is that the high p_T jet activity
 990 in such a sample is negligible, allowing for a clean evaluation of the PU
 991 contribution. The average p_T distribution is evaluated in cells of the
 992 (η, ϕ) plane, and the ratio of this quantity calculated in data and MC
 993 is used as a correction factor to the jet 4-momentum.

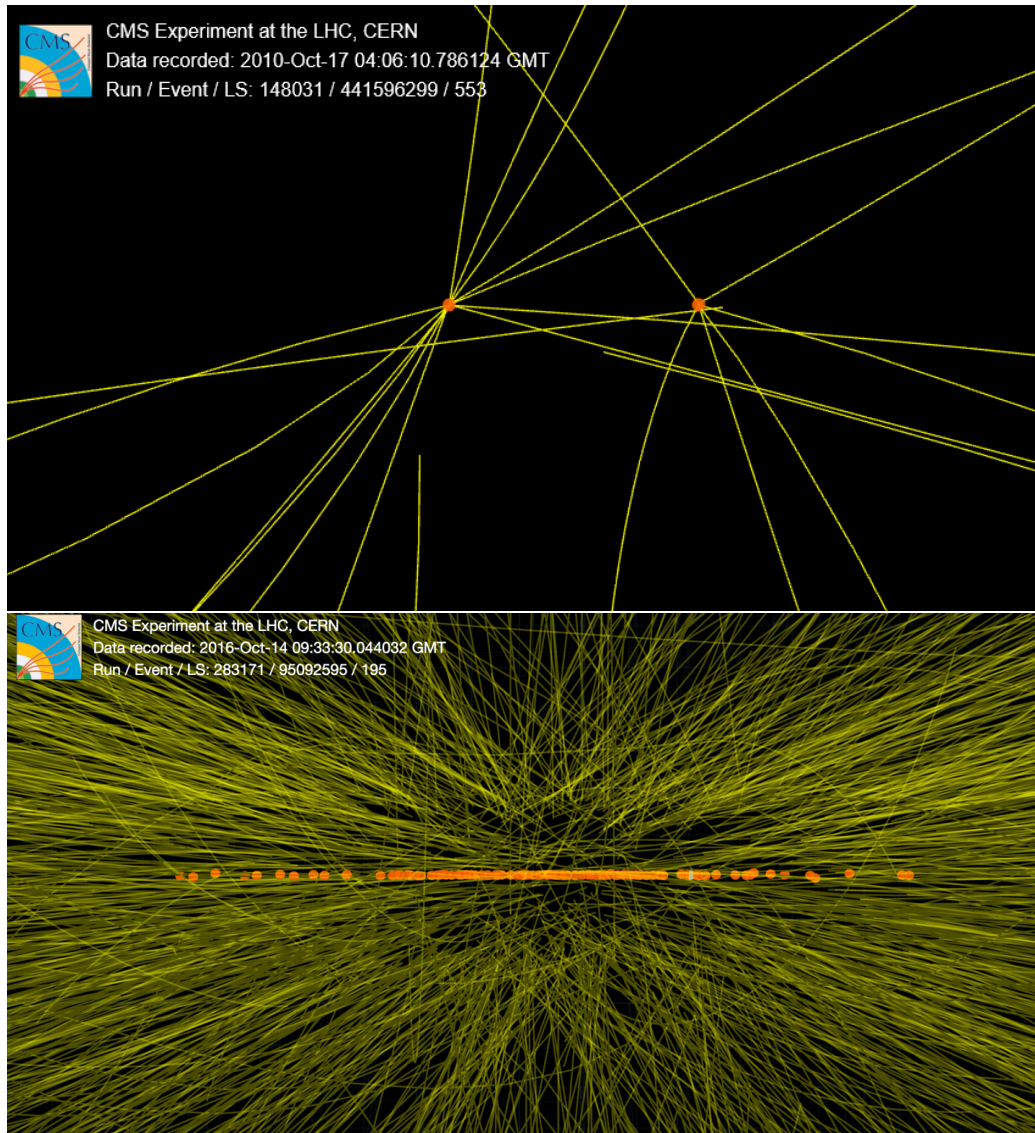


Figure 2.12: Comparison of a low PU event recorded in 2010 (upper) and a high PU event collected during a high PU run in 2016 (lower).

After accounting for the PU offset, the energy response of the detector is also corrected for. This is done on a simulated sample, where reconstructed jets are matched with the corresponding generator level partons (i.e., before simulating the hadronic shower and interaction with the detector). The jet response is then calculated as the ratio between the reconstructed jet momentum and the parton momentum.

Finally, a residual correction to the jet response is measured in data by exploiting transverse momentum imbalance. The idea is that in events with no undetected particles, the total momentum on the transverse plane should vanish. A reference object with well measured transverse momentum is identified: a jet in the central detector region (where the detector performs best) in dijet events, a pair of electrons or muons (the decay products of a Z boson) in Z +jet events, or a photon in γ +jet events. The response is then calculated as the ratio between the p_T of the jet to be calibrated and the one of the reference object.

After applying the full set of corrections, the jet response is found to be within 3% from unity in the whole detector volume, and within 1% from unity in the barrel region.

2.4.5 Missing transverse momentum reconstruction

Finally, the calculation of $\vec{p}_T^{\text{miss}}$ is performed considering the previously reconstructed PF objects [35]. As is the case for jets, PU can spoil the calculation of $\vec{p}_T^{\text{miss}}$ if done simply by applying Eq. 2.8. In this analysis the pileup per particle identification (PUPPI) method [36] is used to make the estimate of the missing transverse momentum more robust with respect to PU. The PUPPI algorithm calculates a local shape variable α_i for every PF particle (charged or neutral) in the event:

$$\alpha_i = \log \sum_{j \neq i} \left(\frac{p_{Tj}}{\Delta R_{ij}} \right)^2, \quad (2.15)$$

where j runs over all charged particles inside a cone of radius 0.4 in (η, ϕ) around the i -th particle, and ΔR_{ij} is the (η, ϕ) distance. This variable is then compared to its median PU value $\bar{\alpha}_{PU}$ with a χ^2 approximation:

$$\chi_i^2 = \frac{(\alpha_i - \bar{\alpha}_{PU})^2}{\text{RMS}_{PU}^2}. \quad (2.16)$$

The values of $\bar{\alpha}_{PU}$ and RMS_{PU}^2 are calculated using all charged particles associated to PU vertices. Finally, a weight is calculated for each particle as:

$$w_i = F_{\chi^2, \text{NDF}=1}(\chi_i^2), \quad (2.17)$$

1025 where F is the cumulative distribution function of a χ^2 with one degree of
 1026 freedom. The weights range from 0 for particles associated with a PU vertex,
 1027 to close to 1 for particles associated with the PV, with charged particles from
 1028 the PV being assigned $w_i = 1$. The $\vec{p}_T^{\text{miss}}$ can then be calculated following
 1029 Eq. 2.8, but with every particle's transverse momentum $p_{T,i}$ weighted by the
 1030 corresponding PUPPI weight w_i :

$$\vec{p}_T^{\text{miss}} = - \sum_i w_i \cdot \vec{p}_{T,i}. \quad (2.18)$$

1031 Finally, the jet energy corrections (JECs) detailed in the previous section are
 1032 also propagated to the calculation of $\vec{p}_T^{\text{miss}}$:

$$\vec{p}_T^{\text{miss,corr}} = \vec{p}_T^{\text{miss}} - \sum_{\text{jets}} (\vec{p}_{T,jet}^{\text{corr}} - \vec{p}_{T,jet}). \quad (2.19)$$

Chapter 3

The $H \rightarrow WW$ channel

3.1 Signal signatures

This analysis targets events in which the H boson decays to a pair of W bosons. Like the H boson, the W boson has a very short lifetime of $\sim 10^{-25}$ s [10], meaning that what is detected are the decay products of the W boson pair. The W boson can decay to a pair of quarks (hadronic decay) or a charged lepton and its associated neutrino (leptonic decay), with BRs of 67% and 11% (per lepton family), respectively. Although more probable, the hadronic final state suffers from a very large background in a QCD-dominated environment such as the LHC. Leptonic decays are therefore experimentally more sensitive overall, and drive the sensitivity of the analysis. Some semileptonic decays are also considered when targeting VH production with a leptonic decay of the V boson, as will be discussed later. Apart from the choice of the decay mode of the W boson pair, the final state of the process of interest also depends on how the H boson was produced.

3.1.1 The ggH signature

As anticipated in previous sections, ggH is the most probable production mode of a H boson at the LHC. As shown in Fig. 1.3, the leading order process does not present any particles in the final state other than the decay products of the H boson; at higher order in QCD however, additional hadronic jets may arise from initial state radiation. In this analysis, unless otherwise specified, both W bosons coming from the decay of the H boson are required to decay leptonically:

$$W \rightarrow \ell \nu_\ell. \quad (3.1)$$

1057 The neutrinos are undetected and give rise to p_T^{miss} . Of the three charged
 1058 lepton families, the electron and muon are detected directly, while the τ
 1059 lepton also decays in within the volume of the detector. This decay can once
 1060 again be leptonic or hadronic:

$$\tau \rightarrow e(\mu)\nu_e(\nu_\mu)\nu_\tau, \quad \tau \rightarrow qq'\nu_\tau. \quad (3.2)$$

1061 The former case is included in the definition of the signal, the latter is not.
 1062 The final state that drives the sensitivity of the analysis is therefore:

$$H \rightarrow WW \rightarrow 2\ell 2\nu. \quad (3.3)$$

1063 This final state benefits not only from the presence of leptons, distinguishing
 1064 it from much more abundant QCD processes, but also from a manifestation
 1065 of the scalar nature of the H boson in the form of a spin correlation effect.
 1066 Angular momentum conservation means that the two spin 1 W bosons orig-
 1067 inating from the spin 0 H boson will have opposite spin projections (these
 1068 being forced to sum to 0 in the rest frame of the H boson). On subse-
 1069 quent decay of the W bosons to an $\ell\nu_\ell$ pair, the (right)left-handedness of
 1070 the (anti)neutrino forces is to be emitted with a momentum pointing in the
 1071 (same) opposite direction with respect to its spin. As shown qualitatively in
 1072 Fig. 3.1, this results in the neutrino pair and the charged lepton pair being
 1073 emitted in opposite hemispheres. Experimentally, this means that the two
 1074 charged leptons in the $H \rightarrow WW$ decay will typically have a small separation
 1075 angle. Because this effect is intrinsically tied to the fact that the W bosons
 1076 originate from a spin 0 resonance, it is unique for this process and can serve
 1077 to distinguish it from sources of background. Another distinguishing feature
 1078 comes from the fact that, because $m_W > m_H/2$, one of the W bosons in the
 1079 decay will have high virtuality. This is reflected in one of the charged leptons
 1080 in the leptonic decay having, on average, a lower p_T than when emitted from
 1081 an on-shell W boson. These effects are captured by the two main discrim-
 1082 inating variables in this analysis, being the invariant mass of the dilepton
 1083 pair and the transverse mass of the H boson. The former is defined as:

$$m_{\ell\ell} = \sqrt{2p^{\ell_1}p^{\ell_2}(1 - \cos \Delta\alpha)}, \quad (3.4)$$

1084 where p^{ℓ_1} and p^{ℓ_2} are the momenta of the leptons with highest (leading)
 1085 and second highest (subleading) p_T in the event, and $\Delta\alpha$ is their separation
 1086 angle. This formulation assumes $m_e \simeq m_\mu \simeq 0$, a condition that holds well
 1087 at the energy scales under consideration ($p_T \sim 10 \text{ GeV}$ or higher). For the
 1088 latter, the transverse mass for a generic set of particles $\{P_i\}$ with transverse

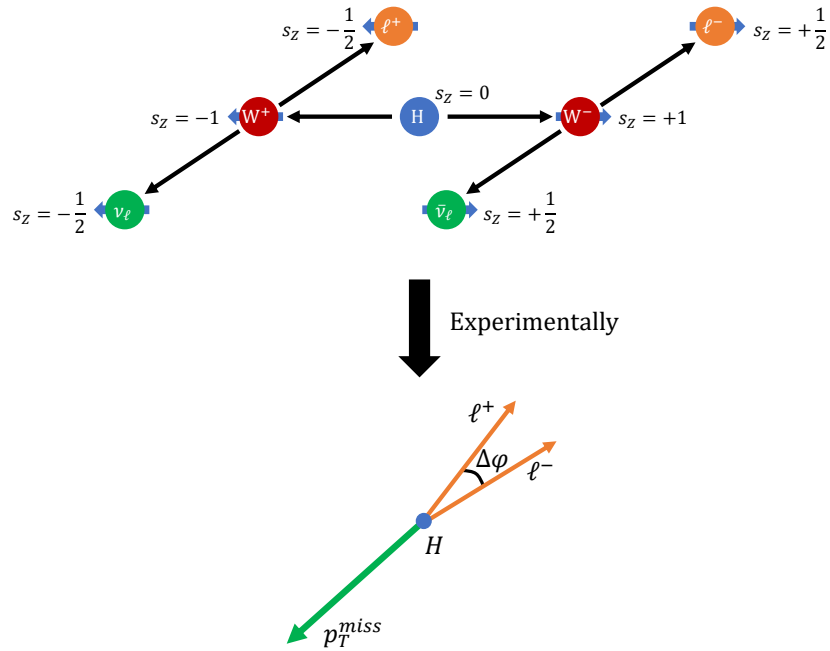


Figure 3.1: Qualitative depiction of the spin correlation effect in the $H \rightarrow WW$ decay.

1089 momenta $\vec{p}_{T,i}$ is defined as:

$$m_T(\{P_i\}) = \sqrt{\left(\sum |\vec{p}_{T,i}|\right)^2 - \left|\sum \vec{p}_{T,i}\right|^2}; \quad (3.5)$$

1090 the transverse mass of the H boson is then defined as $m_T^H = m_T(\vec{p}_T^{\ell\ell}, \vec{p}_T^{\text{miss}})$.
 1091 Note that this definition neglects the invariant mass of the dilepton system;
 1092 this is done to decorrelate m_T^H from $m_{\ell\ell}$, since the two are used together in
 1093 the fit, as described in Section 4.3. Because it measures the mass of the decay
 1094 system projected on the transverse plane, m_T^H is bound to be below m_H in
 1095 events in which a H boson is produced (modulo detector resolution effects).
 1096 Since no such requirement is present in background processes, this variable
 1097 also gives a high discrimination power against backgrounds. Altogether, the
 1098 experimental signature of the fully leptonic $H \rightarrow WW$ decay consists of two
 1099 oppositely charged leptons with high p_T , moderate p_T^{miss} coming from the
 1100 neutrinos, low $m_{\ell\ell}$, and $m_T^H \leq 125 \text{ GeV}$. An important distinction has to
 1101 be made concerning the flavor composition of the charged leptons in the
 1102 final state. Although decays to electrons and muons are roughly equally
 1103 probable, events in which the W bosons decay to leptons of the same flavor
 1104 (SF) are subject to a much larger background contamination coming from
 1105 Drell-Yan (DY) processes, as will be detailed in Section 3.2. For this reason,
 1106 so-called different flavor (DF) final states i.e., with an $e\mu$ pair in the final
 1107 state, are much more sensitive. Finally, because the ggH process is initiated
 1108 by gluons, a varying number of hadronic jets can be produced by initial state
 1109 radiation, although with probabilities suppressed by increasing powers of α_s .
 1110 The leading order ggH diagram with a leptonic $H \rightarrow WW$ decay is shown in
 1111 Fig. 3.2.

1112 3.1.2 The VBF signature

1113 Everything described so far applies regardless of production mode; what
 1114 distinguishes the VBF signature from ggH is the presence of the two jets
 1115 coming from the quarks that radiate the vector bosons (Fig. 3.3). These
 1116 jets (called VBF jets in the following) can be distinguished from e.g., QCD
 1117 radiation in the ggH process, through their peculiar kinematic features. Since
 1118 the two quarks do not interact directly, they typically retain much of their
 1119 momentum along the beam axis, resulting in a typically large η separation
 1120 $\Delta\eta_{jj}$, as well as a high invariant mass of the dijet pair m_{jj} . This is shown in
 1121 Fig. 3.4. For the same reason (the radiating quarks not being connected
 1122 by colored objects), emission of additional jets within the $\Delta\eta$ region between
 1123 the two VBF jets is suppressed.

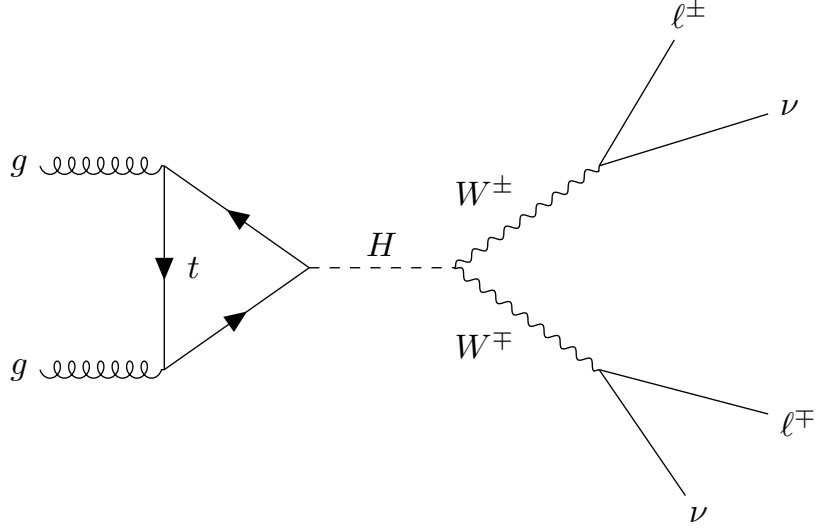


Figure 3.2: Leading order diagram of the ggH production mechanism with a fully leptonic $H \rightarrow WW$ decay.

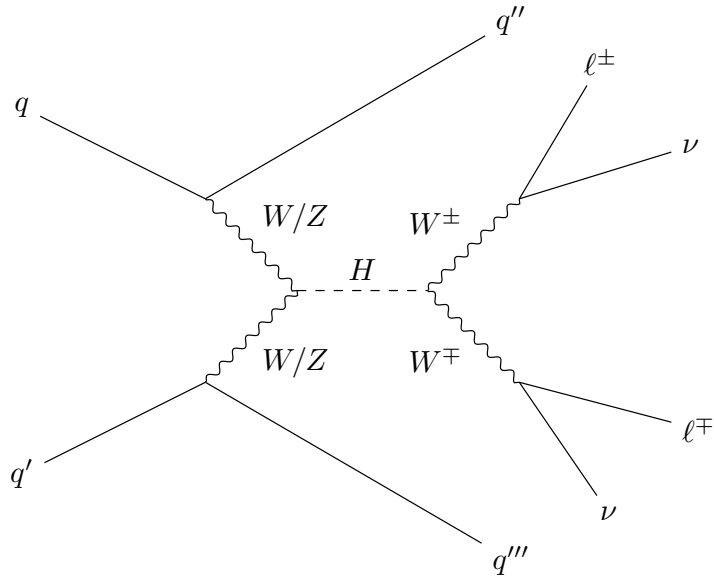


Figure 3.3: Leading order diagram of the VBF production mechanism with a fully leptonic $H \rightarrow WW$ decay.

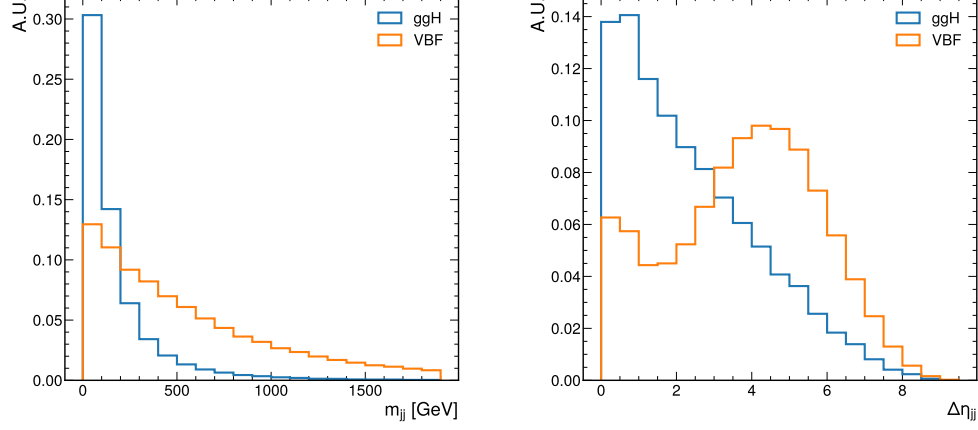


Figure 3.4: Comparison of the m_{jj} (left) and $\Delta\eta_{jj}$ (right) spectra in ggH and VBF processes.

3.1.3 The WH and ZH signatures

The WH and ZH production modes can be distinguished by exploiting the presence of the additional vector boson ($V = W, Z$). If the V boson decays hadronically, this will result in a pair of jets whose invariant mass peaks around the mass of the V boson, as shown in Fig 3.5. This final state will be referred to as VH2j in the following. In this analysis, only fully leptonic decays of the H boson are considered in this case, meaning that all considerations on $m_{\ell\ell}$ and m_T^H done previously still apply. In the case of leptonic decay of the associated V boson, four different final states are accounted for in the analysis, including also semileptonic decays of the H:

- In WH production with a semileptonic decay of the $H \rightarrow WW$ system (WHSS), the final state consists of two charged leptons, moderate p_T^{miss} , and two jets. Because the two leptons do not both arise from the H boson, there is no requirement for them to have opposite electric charge; this fact can be exploited by restricting the analysis to events with same-sign (SS) leptons in the final state, severely reducing background contamination.
- In WH production with a fully leptonic decay of the $H \rightarrow WW$ system (WH3l), the final state consists of three charged leptons and moderate p_T^{miss} .
- In ZH production with a semileptonic decay of the $H \rightarrow WW$ system

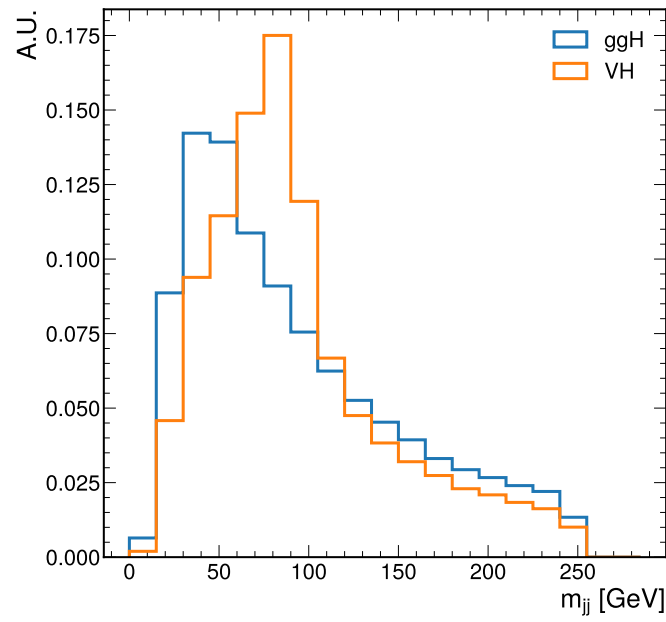


Figure 3.5: Comparison of the m_{jj} spectrum in ggH and VH processes.

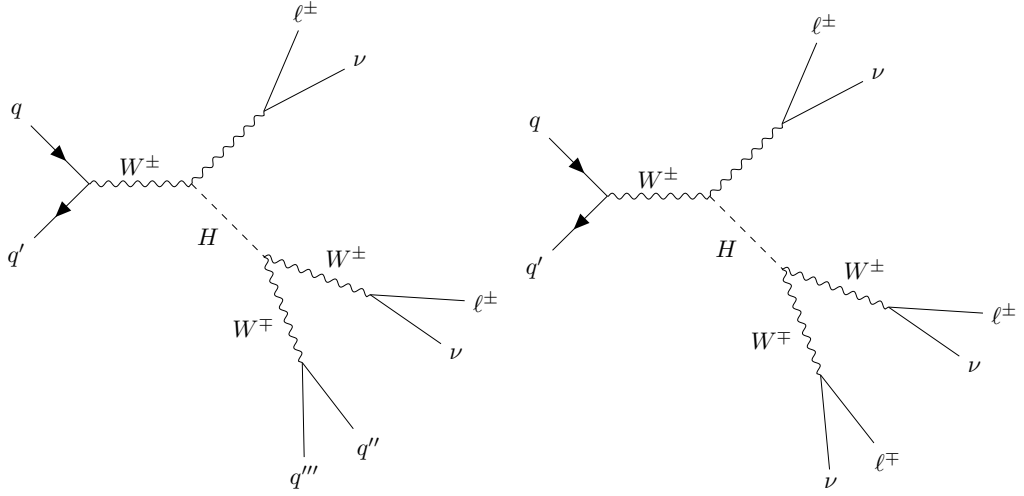


Figure 3.6: Leading order diagrams for WH production with leptonic decay of the associated W boson and semileptonic (WHSS, left) or fully leptonic (WH3l, right) decay of the $H \rightarrow WW$ system.

1145 (ZH3l), the final state consists again of three charged leptons and mod-
 1146 erate p_T^{miss} . The fact that out of the three leptons two originate from a
 1147 Z boson can be used to reduce background contamination.

- 1148 • Finally, in ZH production with a fully leptonic decay of the $H \rightarrow WW$
 1149 system (ZH4l), the final state consists of four leptons and p_T^{miss} . This
 1150 final state has very low background contamination, however the proba-
 1151 bility for both the H and the Z bosons to decay leptonically is relatively
 1152 low.

1153 Leading order diagrams for all of the above processes are shown in Fig. 3.6
 1154 and 3.7. Accounting for the split in DF and SF in the ggH, VBF, and VH2j
 1155 processes, this analysis targets a total of 10 final states, each with its own
 1156 background composition, detailed in Section 3.2.

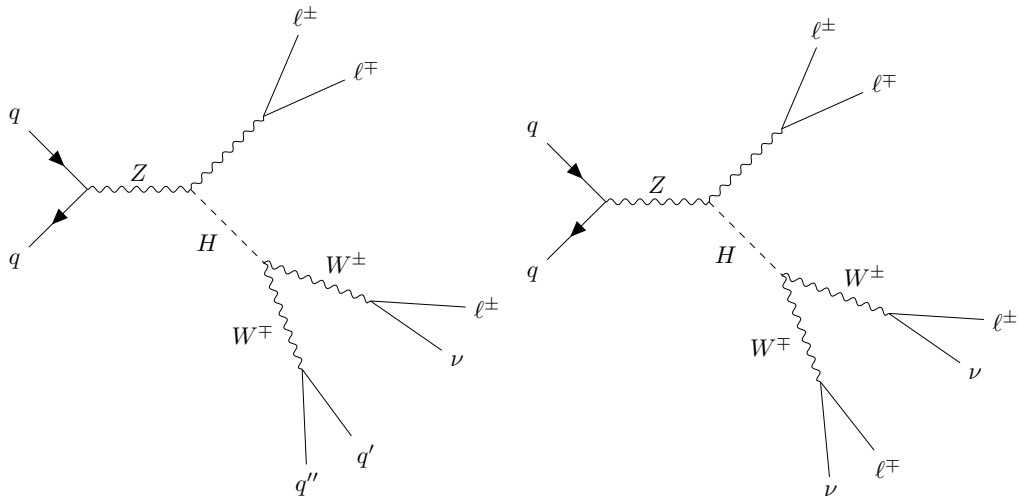


Figure 3.7: Leading order diagrams for ZH production with leptonic decay of the associated Z boson and semileptonic (ZH3l, left) or fully leptonic (ZH4l, right) decay of the $H \rightarrow WW$ system.

1157 3.2 Sources of background

1158 Other SM processes which can result in the same final state composition of
 1159 the signal give rise to background events. The approach in tackling them is
 1160 twofold: on the one hand, differences in the kinematic distributions of the
 1161 final state particles between signal and background processes can be exploited
 1162 to separate them, by applying kinematic selections and fitting the shape
 1163 of discriminating observables; on the other hand, the expected background
 1164 yields in the phase space of the analysis are estimated as precisely as possible
 1165 by a combination of MC simulation and data-driven methods and effectively
 1166 subtracted from data in the fit procedure, as will be detailed in the following.

1167 3.2.1 Non-resonant WW production

1168 Arguably the most immediately apparent source of background comes from
 1169 the production of pairs of W bosons through processes not involving a H
 1170 boson. At the LHC, production of WW pairs can be initiated by either
 1171 quarks or gluons (with the quark-initiated contribution being the dominant
 1172 one), as shown in Fig. 3.9. Since the particle content in the final state is
 1173 the same as in the $H \rightarrow WW$ decay, these processes are one of the main
 1174 sources of background in the analysis. As anticipated in Section 3.1 however,
 1175 the scalar H boson resonance compresses the $m_{\ell\ell}$ spectrum to lower values.
 1176 This is not the case for non-resonant WW production, resulting in a higher
 1177 average $m_{\ell\ell}$, as shown in Fig. 3.8, a fact that can be exploited to separate the
 1178 two processes. In the VBF phase space especially, a particular subset of the
 1179 WW production processes, commonly referred to as vector boson scattering
 1180 (VBS), becomes important. This is because, as shown in Fig. 3.9, VBS
 1181 presents a dijet structure very similar to VBF.

1182 3.2.2 The Drell-Yan process

1183 Another major source of background comes from the production of dilepton
 1184 pairs through Drell-Yan (DY) processes, an example of which is given in
 1185 Fig. 3.10. The key aspect of DY production is that the leptons in the final
 1186 state have the same flavor. This is the prime reason why DF channels are
 1187 typically more sensitive than SF channels in the analysis, since the DY con-
 1188 tamination of the DF final state is greatly reduced. Compared to $H \rightarrow WW$
 1189 processes, when pairs of electrons or muons are produced, there are no neu-
 1190 trinos in the final state; non-zero p_T^{miss} can however arise due to the finite
 1191 resolution of the detector. This, combined with its very large (compared to
 1192 the signal) cross section, makes DY the main source of background in SF

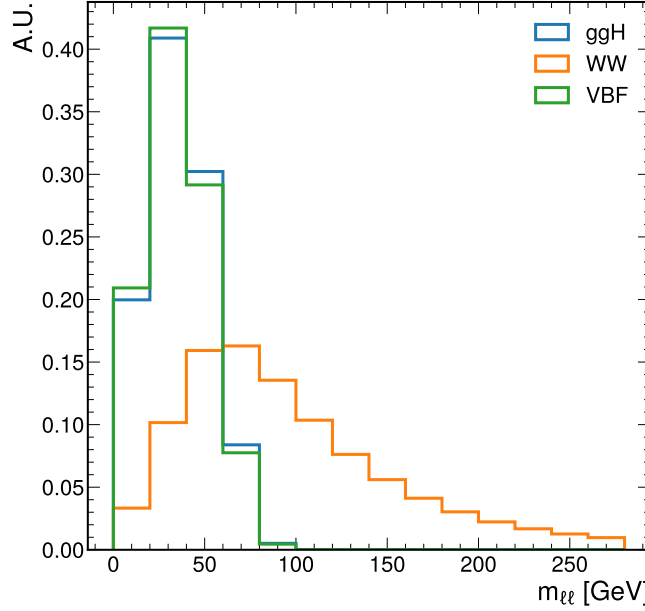


Figure 3.8: Comparison of the $m_{\ell\ell}$ spectrum in ggH, VBF, and in non-resonant production of a pair of W bosons, one of the main background processes in this analysis.

channels. Estimation of this background is made more complicated by the instrumental nature of the reconstructed p_T^{miss} , which is not well modeled by the simulation of the detector. A dedicated data-driven estimation technique is therefore used instead, and is described in Section 4.3. While in DF channels the DY contribution is much reduced, it can still enter the selection when a pair of tau leptons is produced:

$$qq \rightarrow Z/\gamma^* \rightarrow \tau\tau \rightarrow e\mu 4\nu, \quad (3.6)$$

which results in a signature analogous to $H \rightarrow WW$ with a pair of DF leptons and p_T^{miss} .

3.2.3 Top quark production

The production of t quarks is the dominant background contribution in DF channels. This is due to the fact that the t quark, due to its very large mass of 173 GeV [10], decays on a timescale shorter than the hadronization one to a b quark and a W boson, as shown in Fig. 3.11. The b quarks

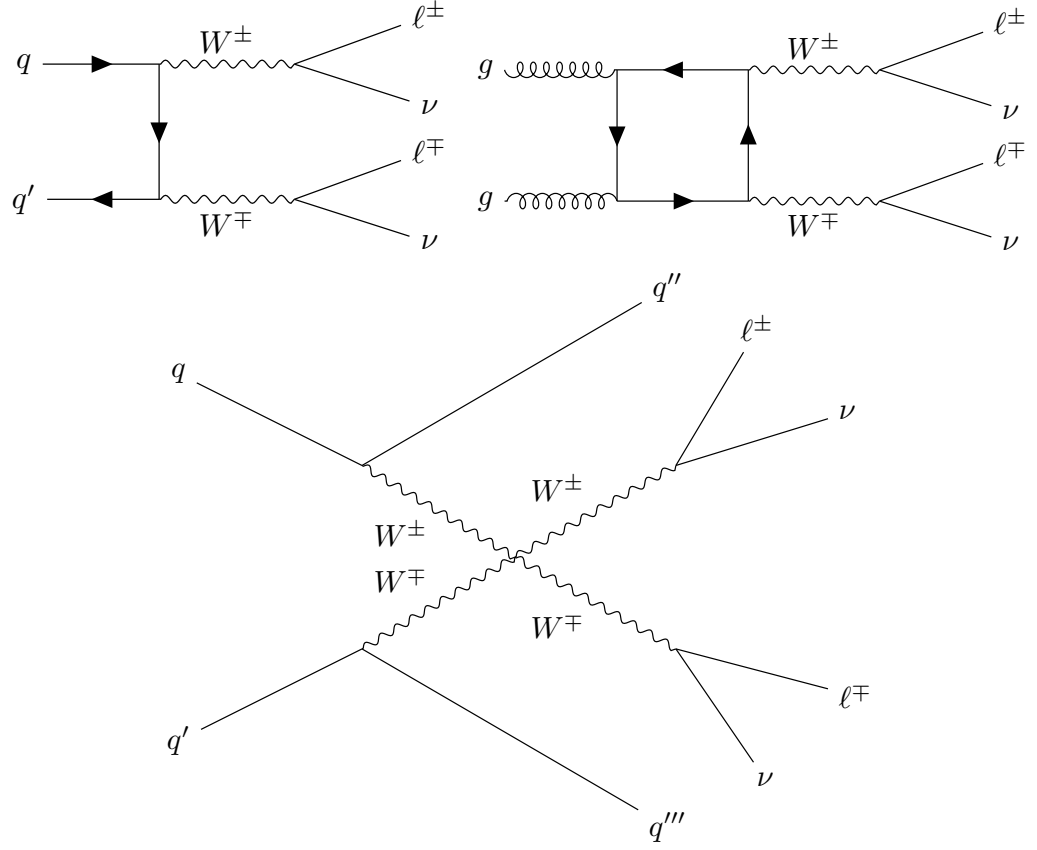


Figure 3.9: Example diagrams for production of WW pairs by a pair of quarks (upper left), a pair of gluons (upper right), and through the VBS process (lower).

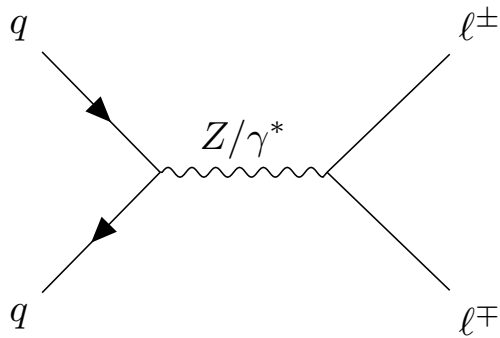


Figure 3.10: Leading order diagram for DY production of a dilepton pair.

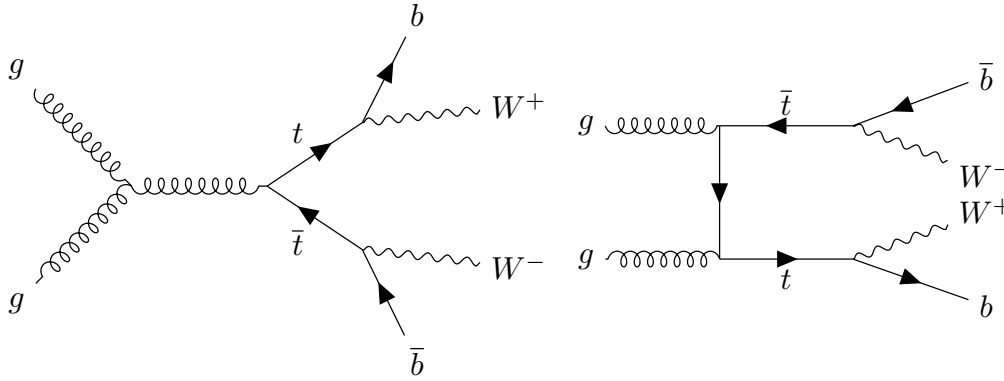


Figure 3.11: Example diagrams for production of pairs of t quarks.

give rise to hadronic jets with a peculiar structure, due primarily to the long lifetime and large mass of the bound states they form compared to other unstable hadrons. These jets can be distinguished from so-called light flavor jets i.e., ones coming from u , d , s , c quarks, or gluons, by means of dedicated algorithms, called b taggers. Although the performance of b taggers in recognizing b quark-induced jets (b jets) is excellent, the large cross section for t quark production at the LHC makes this one of the largest sources of background in most categories of the analysis.

3.2.4 Non-prompt lepton background

Another instrumentation-driven background arises from so-called non-prompt leptons. These are any reconstructed leptons which do not correspond to genuine leptons coming from the primary interaction i.e., prompt leptons. The sources can be leptonic decays of heavy-flavor hadrons (typically B and D hadrons), hadrons erroneously reconstructed as leptons, prompt photons that convert to electrons by interacting with the tracker, or jets misidentified as leptons. Because these misreconstructed events are instrumental in nature (i.e., they arise from the finite resolution of the detector), they are not well modeled by the simulation of the apparatus. The bulk of these events is suppressed by the identification and isolation requirements outlined in Section 2.4. The remaining contribution, mostly concentrated in the low p_T and, consequently, low $m_{\ell\ell}$ phase space, is estimated from data, as will be described in Section 4.1.

1228 **3.2.5 Multiboson production**

1229 Diboson and triboson events include production of WZ , $V\gamma^*$, $V\gamma$, and ZZ
1230 pairs, as well as various combinations of triple vector boson production. The
1231 importance of these backgrounds varies greatly depending on the final state
1232 of interest, as will be detailed in later sections.

Chapter 4

Analysis strategy

4.1 Data sets and simulated samples

4.1.1 Trigger setup

The analysis is performed on data gathered by the CMS detector in the 2016–2018 time frame, corresponding to Run 2 of the LHC, and on a corresponding set of Monte Carlo (MC) simulations that form the SM prediction. Because of variations in the operating setup of CMS and the LHC over the data taking period, the data are divided into three sets, corresponding to each year of data taking. For the same reason, MC samples are produced independently for each year as well, since they include the simulation of the particles interacting with the detector, as will be detailed later. The three data sets correspond to an integrated luminosity of 36.3, 41.5, and 59.7 fb⁻¹ for 2016, 2017, and 2018, respectively, for a total of 138 fb⁻¹ [29, 37, 38]. Events selected for analysis are required to satisfy HLT requirements on the number and reconstruction quality of leptons (electrons and muons). In particular, events are selected that contain at least one high p_T electron or muon, or a pair of charged leptons in any combination. Precise thresholds vary throughout the data taking period, and are specified in Table 4.1. Thresholds on electrons are generally higher since the L1 trigger does not have access to information from the tracker, and must therefore only rely on energy deposits in the ECAL to reconstruct electrons. At low energies, calorimetric measurements become sensitive to noise, spoiling the reconstruction. Also note that for this reason, at this stage, electrons and photons cannot be distinguished. For muons on the other hand, the L1 trigger can rely on the muon system, somewhat making up for the missing information from the silicon tracker at this stage and allowing for lower thresholds. The overall efficiency of the combination of triggers used is 62%, 66%, 72%, and 76% for ggH, VBF, WH,

Table 4.1: Trigger requirements on the data set used in the analysis.

Trigger	Year	Requirements
Single electron	2016	$p_T > 25 \text{ GeV}$, $ \eta < 2.1$ or $p_T > 27 \text{ GeV}$, $2.1 < \eta < 2.5$
	2017	$p_T > 35 \text{ GeV}$, $ \eta < 2.5$
	2018	$p_T > 32 \text{ GeV}$, $ \eta < 2.5$
Single muon	2016	$p_T > 24 \text{ GeV}$, $ \eta < 2.4$
	2017	$p_T > 27 \text{ GeV}$, $ \eta < 2.4$
	2018	$p_T > 24 \text{ GeV}$, $ \eta < 2.4$
Double electron	All years	$p_T^{\ell_1} > 23 \text{ GeV}$, $p_T^{\ell_2} > 12 \text{ GeV}$, $ \eta_{1,2} < 2.5$
Double muon	All years	$p_T^{\ell_1} > 17 \text{ GeV}$, $p_T^{\ell_2} > 8 \text{ GeV}$, $ \eta_{1,2} < 2.4$
Electron-muon	All years	$p_T^{\ell_1} > 23 \text{ GeV}$, $p_T^{\ell_2} > 12 \text{ GeV}$
		$p_T^{\mu_2} > 8 \text{ GeV}$ in first part of 2016 data taking $ \eta_e < 2.5$, $ \eta_\mu < 2.4$

and ZH events, respectively.

4.1.2 Background modelling

A full set of MC simulation samples is used to describe both the signal and background processes entering the analysis. All MC samples pertaining to a given year of data taking share the same set of PDFs, UE tune and parton shower (PS) configuration. The PS interfaces with MC generators and simulates the hadronic showering of colored particles. The PDF set is NNPDF 3.0 [13, 14] at next-to-leading order (NLO) for 2016 and NNPDF 3.1 [15] at next-to-next-to-leading order (NNLO) for 2017 and 2018. All event generators are interfaced with PYTHIA [39] for the simulation of the UE, PS, and hadronization. Versions 8.226 and 8.230 are used for 2016 and 2017–2018, respectively. The UE tune is set to CUETP8M1 [40] and CP5 [41] in 2016 and 2017–2018, respectively.

Several event generators are used to describe different physics processes. These are POWHEG v2 [42], possibly with the MINLO HVJ [43] or MINLO HJJ [44] extensions, MADGRAPH5_aMC@NLO [45] (MG5), and MCFM v7.0 [46]. The full list of background samples with the corresponding generators is given in Table 4.2. All events generated via MC are fed into a detailed simulation of the CMS apparatus built in the GEANT4 package [47] to accurately describe the interaction of final state particles with the detector material. The effect of PU is added to all MC samples by generating addi-

Table 4.2: Full set of MC simulation samples used to model background processes in the analysis.

Process	Generator	Accuracy	Notes
(qq)WW	POWHEG v2	NLO	-
(gg)WW	MCFM v7.0	LO	Normalized to NLO cross section
VBS	MG5 v2.4.2	LO	-
Top quark	POWHEG v2	NLO	-
DY	MG5 v2.4.1	NLO	-
$V\gamma$	"	NLO	-
$V\gamma^*$	"	NLO	$m_{\gamma^*} < 100 \text{ MeV}$
VV	POWHEG v2	NLO	Events with a Z or γ^* with mass below 100 MeV
VVV	MG5 v2.4.2	NLO	-

tional p-p scattering events via PYTHIA and adding the resulting particles to the simulated hard scattering processes. The number of additional interactions obtained this way is tuned to match the mean number of interactions per crossing observed in data, whose mean is 23 in 2016 and 32 in 2017 and 2018.

The quark-induced WW background in particular is generated at NLO in QCD and then reweighted, so that the p_T spectrum of the W boson pair matches the NNLO prediction with NNLL resummation.

A dedicated strategy is used for DY production of τ lepton pairs. This background is estimated with a hybrid technique called τ embedding [48], which exploits both MC simulation as well as data. A sample is selected in data consisting of events with well reconstructed $\mu^+\mu^-$ pairs. The energy deposits corresponding to these muons are then removed from the event, and replaced with a $\tau^+\tau^-$ pair with the same kinematics. The decay and interaction of the τ leptons with the detector is simulated. The advantage of this procedure is that it leverages data to get a better description of quantities that are less well modeled by MC simulation, such as the instrumental component of p_T^{miss} .

4.1.3 Signal modelling

Signal samples for the analysis are produced for the ggH, VBF, WH, and ZH processes, as well as for production of a H boson in association with a t (ttH) or b (bbH) quark pair; the latter processes are however not the target

of this analysis, since a dedicated measurement by CMS already exists [49], and their contributions are found to be negligible in all categories. The ZH process includes contributions initiated both by quarks and gluons. All signal samples are generated with the POWHEG v2 generator, providing a description at NLO accuracy in QCD, except for the bbH sample which is produced with MG5.

The ggH sample in particular is generated with POWHEG at NLO accuracy in QCD and reweighted to match the NNLO prediction. The reweighting is done as a function of the p_T of the Higgs boson and the number of hadronic jets produced in association with it [50, 51]. The NNLO accuracy drops to NLO when requiring at least one hadronic jet, and LO when requiring at least two. Since one of the key aspects of this analysis is the simultaneous measurement of ggH and VBF, with the latter having at least two jets in the final state, higher accuracy in this region of phase space is desirable. The MINLO HJJ generator is thus used for the phase space with at least two jets, ensuring NLO accuracy. The MINLO sample has been compared to the one obtained from POWHEG with NNLOPS reweighting in a 2-jet selection, as shown in Fig. 4.1. The comparison is focused on features of the dijet system, with the pseudorapidity separation between the two jets $\Delta\eta_{jj}$ in particular showing a noticeable difference in the more accurate MINLO HJJ sample.

The VBF sample, generated in POWHEG and interfaced with PYTHIA for the PS, presents a mismodelling in the emission of additional jets in the pseudorapidity gap between the two VBF ones. In order to evaluate the impact of this behavior, a corrected sample has been produced for the 2018 data set and compared to the nominal one used in the analysis. While the differences in shape of the relevant observables were found to be negligible, as shown in Fig. 4.2, a difference in event yields in the VBF selection (described in following sections) of about 7% was observed. This discrepancy is due to the requirements on events to contain exactly two jets with $p_T > 30$ GeV, which is affected by the mismodelling of the kinematics of the third jet in the event (if present). To correct for this discrepancy in event yields, an additional weight is applied to the VBF samples used for analysis. This weight is derived separately in each analysis category.

4.2 Definition of analysis objects and corrections

On top of the reconstruction outlined in Section 2.4, additional requirements and corrections on reconstructed objects are applied for this analysis, and

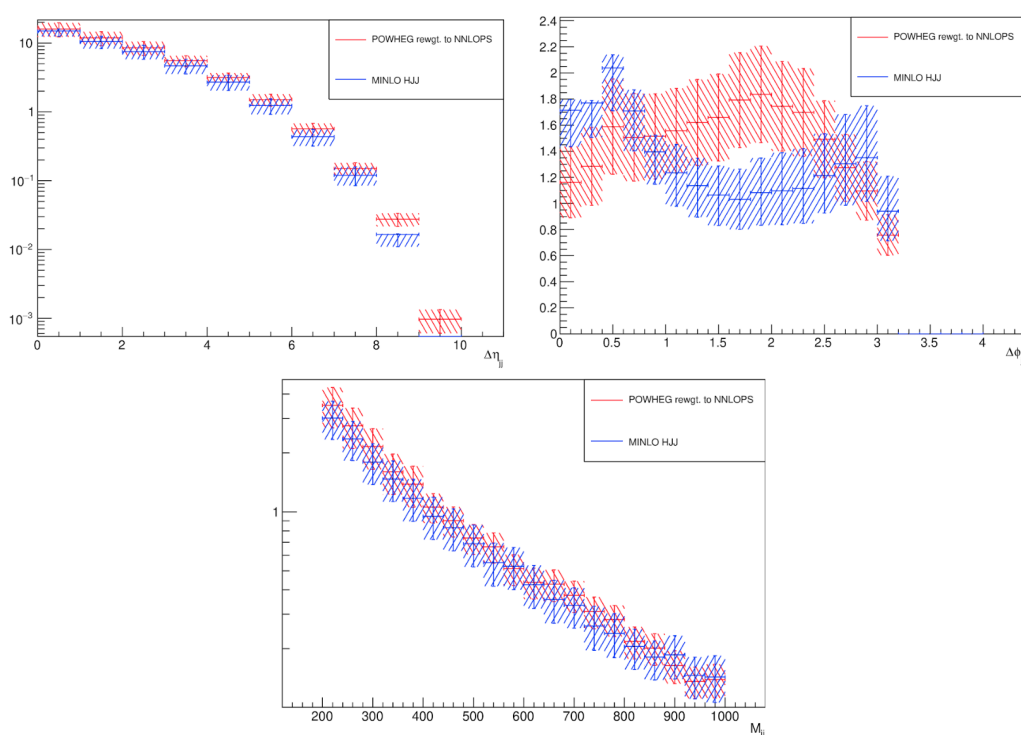


Figure 4.1: Comparison between the MINLO HJJ and POWHEG with NNLOPS reweighting samples in a 2-jet phase space.

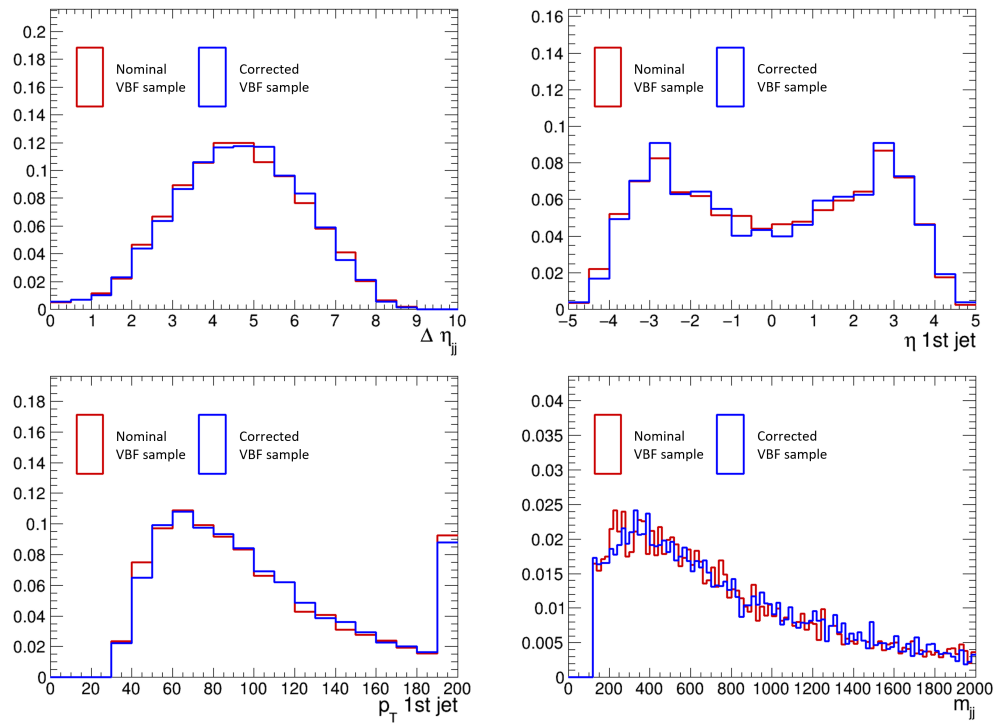


Figure 4.2: Shape comparison between the nominal POWHEG VBF samples used in the analysis (red) and the one with the corrected modelling of the third jet (blue).

are outlined in this section.

4.2.1 Trigger efficiencies in MC samples

To describe the trigger response in MC samples, a tag and probe data driven method is used. First, a sample of dilepton events with $60 < m_{\ell\ell} < 120$ GeV is selected in data. The lepton that fired the trigger is defined as the tag lepton, while the second is defined as the probe. The per-lepton efficiency of the trigger can then be estimated as the ratio of the number of probes that also satisfy the trigger conditions to the total number of probes. This efficiency is estimated as a function of p_T and η of the probe lepton, separately for electrons and muons.

While these numbers could in principle be applied directly to MC events as weights to account for the trigger efficiency, two orders of complication arise in the case of this analysis. The first is that the bulk of events of interest use the double muon ($\mu\mu$), double electron (ee), or electron-muon triggers ($e\mu$), as shown in Table 4.1, meaning that the single lepton trigger efficiencies need to be combined to estimate the efficiency of these dilepton triggers. This is done as follows. Taking an $e\mu$ event as an example, let $\varepsilon_{e\mu}^{\ell_i}$ be the trigger efficiency of the $e\mu$ trigger for the i -th lepton in the event, ordered by decreasing p_T . Let $t_{e\mu}^{\ell_i}$ be the p_T threshold of the $e\mu$ trigger on the i -th lepton. Note that since the experimental resolution on electrons and muons is different, the efficiencies have to be calculated separately for the $e\mu$ and μe cases, indicating trigger selections with a leading electron or muon, respectively. The probability $P(D)$ of an event passing the dilepton trigger can then be expressed as:

$$\begin{aligned} P(D)_{t\&p} &= P(e\mu) + P(\mu e) - P(e\mu \wedge \mu e) = \\ &= P(p_T^e > t_{e\mu}^{\ell_1} \wedge p_T^\mu > t_{e\mu}^{\ell_2}) \\ &\quad + P(p_T^e > t_{\mu e}^{\ell_2} \wedge p_T^\mu > t_{\mu e}^{\ell_1}) \\ &\quad - P(p_T^e > t_{e\mu}^{\ell_1} \wedge p_T^\mu > t_{\mu e}^{\ell_1}). \end{aligned} \tag{4.1}$$

The last term is the probability of the event passing both the $e\mu$ and μe trigger selections. Under the assumption that the probabilities of the two leptons passing the respective trigger thresholds are independent, the above expression can be rewritten in terms of the per-lepton trigger efficiencies:

$$P(D)_{t\&p} = \varepsilon_{e\mu}^{\ell_1} \varepsilon_{e\mu}^{\ell_2} + \varepsilon_{\mu e}^{\ell_1} \varepsilon_{\mu e}^{\ell_2} - \varepsilon_{e\mu}^{\ell_1} \varepsilon_{\mu e}^{\ell_1}. \tag{4.2}$$

The assumption of the per-lepton trigger efficiencies being independent holds unless the two leptons are emitted very close to each other, at which point

their respective isolation cones can overlap resulting in a non-negligible correlation between the two. To account for this effect, a correction factor is calculated as follows. The same tag and probe procedure as described above is replicated on a dilepton MC sample yielding $P(D)_{t\&p}^{MC}$. On the same MC sample, the response of the trigger is also simulated, yielding $P(D)_{sim}^{MC}$. The correction factor is then calculated as the ratio between the two, as a function of the ΔR separation between the leptons:

$$w(\Delta R) = \frac{P(D)_{sim}^{MC}}{P(D)_{t\&p}^{MC}}. \quad (4.3)$$

This correction factor is found to significantly deviate from 1 only for low values of ΔR . The electron-muon trigger efficiency is then calculated as:

$$P(D) = w(\Delta R) \cdot P(D)_{t\&p}. \quad (4.4)$$

Analogous expressions hold for the ee and $\mu\mu$ triggers.

The second order of complication is due to the addition of single lepton triggers to the analysis. This is done to recover efficiency on events in which the subleading lepton is below the dilepton trigger thresholds. Let $P(S)$ be the probability of an event passing the single lepton triggers, calculated in the same way as $P(D)$. The total trigger efficiency is:

$$\varepsilon = P(D \vee S) = P(D) + P(S) - P(D \wedge S). \quad (4.5)$$

Rewriting $P(D \wedge S)$ in terms of per-lepton trigger efficiencies is not straightforward, since $P(D)$ and $P(S)$ are intrinsically dependent probabilities. A first step is expressing it in terms of conditional probabilities:

$$\begin{aligned} P(D \wedge S) &= P(D)P(S|D) \\ &= P(D) [P(S_1|D) + P(S_2|D) - P(S_1 \wedge S_2|D)], \end{aligned} \quad (4.6)$$

where $P(S_i)$ is the probability of the i -th lepton passing the single lepton trigger. The advantage of this formulation is that the last term can be factorized as:

$$P(S_1 \wedge S_2|D) = P(S_1|D)P(S_2|D), \quad (4.7)$$

under the assumption of $P(S_1)$ and $P(S_2)$ being independent, which holds to good approximation. These conditional probabilities can be expanded through Bayes' theorem:

$$P(S_{1(2)}|D) = \frac{P(D|S_{1(2)})P(S_{1(2)})}{P(D)} \simeq \frac{P(D_{2(1)}^{low})P(S_{1(2)})}{P(D)}, \quad (4.8)$$

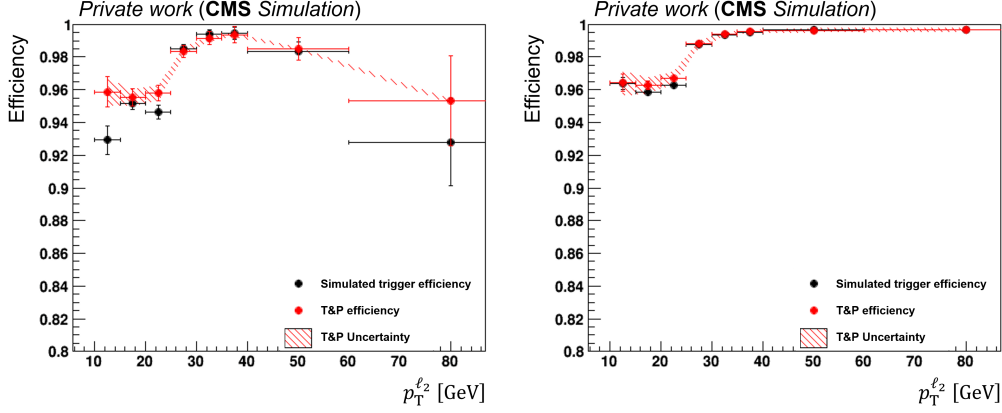


Figure 4.3: Comparison between the trigger efficiency as a function of $p_T^{\ell_2}$ obtained through simulation of the trigger (black curves) and calculated via the tag and probe method described above (red curves). The efficiencies are calculated on ggH (left) and WW (right) MC events containing a $e\mu$ pair with a leading electron.

1395 where $P(D_{2(1)}^{low})$ is the probability of the subleading (leading) lepton passing
 1396 the lower threshold of the dilepton trigger. The last step in Eq. 4.8 holds
 1397 because single lepton trigger thresholds are higher than the lower thresholds
 1398 in dilepton triggers, and the ID and isolation requirements match. Finally,
 1399 combining all of the above yields the combined trigger efficiency expressed
 1400 in terms of data driven single lepton efficiencies:

$$\begin{aligned} \varepsilon = & P(S) + P(D) - P(D_2^{low})P(S_1) - P(D_1^{low})P(S_2) + \\ & + \frac{P(D_1^{low})P(D_2^{low})P(S_1)P(S_2)}{P(D)}. \end{aligned} \quad (4.9)$$

1401 To validate this construction of the data driven trigger efficiencies, the whole
 1402 procedure has been repeated on a set of signal and background MC samples,
 1403 and compared to the full simulation of the trigger. A selection of the com-
 1404 parisons is shown in Fig. 4.3, 4.4, 4.5, and 4.6. A drop in efficiency can be
 1405 noticed, especially in di-electron signal events (Fig. 4.5), when the p_T of the
 1406 subleading lepton is high i.e., above ~ 50 GeV. This is due to the fact that
 1407 when the p_T of both leptons is high, the isolation cones of the two electrons
 1408 overlap, reducing the efficiency of the trigger.

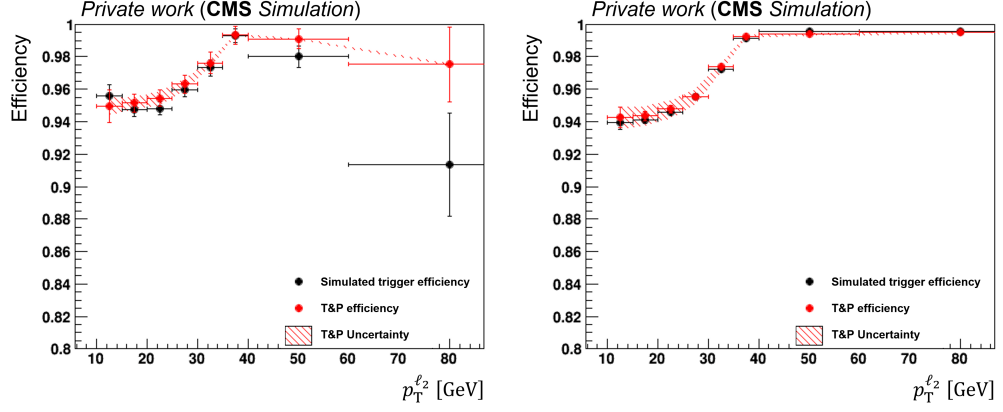


Figure 4.4: Comparison between the trigger efficiency as a function of $p_T^{\ell_2}$ obtained through simulation of the trigger (black curves) and calculated via the tag and probe method described above (red curves). The efficiencies are calculated on ggH (left) and WW (right) MC events containing a $e\mu$ pair with a leading muon.

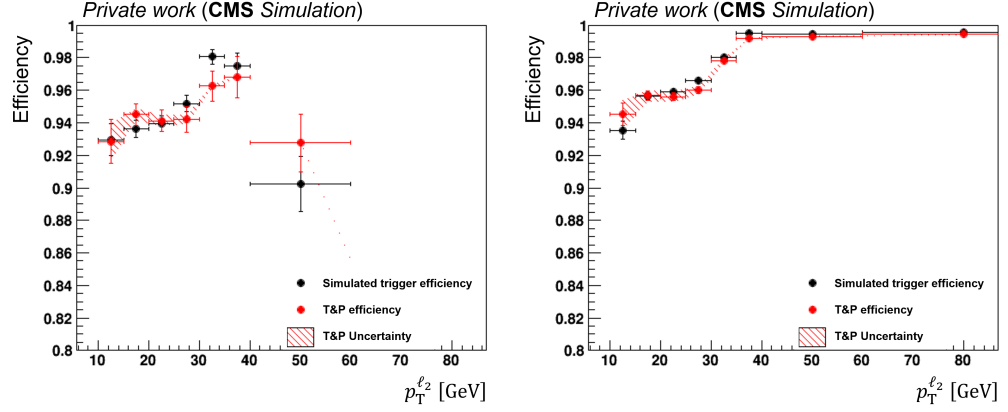


Figure 4.5: Comparison between the trigger efficiency as a function of $p_T^{\ell_2}$ obtained through simulation of the trigger (black curves) and calculated via the tag and probe method described above (red curves). The efficiencies are calculated on ggH (left) and WW (right) MC events containing a ee pair.

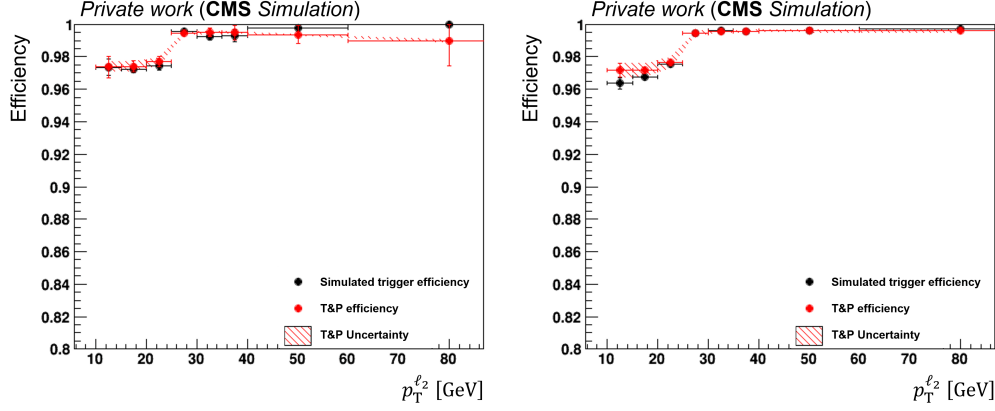


Figure 4.6: Comparison between the trigger efficiency as a function of $p_T^{\ell_2}$ obtained through simulation of the trigger (black curves) and calculated via the tag and probe method described above (red curves). The efficiencies are calculated on ggH (left) and WW (right) MC events containing a $\mu\mu$ pair.

4.2.2 Muons

On top of the identification and isolation criteria outlined in Section 2.4, muons entering the analysis are required to satisfy the following conditions:

- have $p_T > 10$ GeV and $|\eta| < 2.4$;
- have $|d_{xy}| < 0.01(0.02)$ cm for $p_T < 20$ GeV ($p_T > 20$ GeV), where d_{xy} is the impact parameter of the track with respect to the PV i.e., the distance from the point where the track intersects the transverse plane containing the PV to the PV itself;
- have $|d_z| < 0.1$ cm, where d_z is the longitudinal distance from the track to the PV.

In order to further reduce the contamination from non-prompt leptons, a selection on a BDT developed for the ttH analysis [18], called ttHMVA, is also applied. The BDT takes as input kinematics and isolation observables of the lepton, as well as information on the jet closest to the lepton. Muons used in the analysis are required to have a ttHMVA score higher than 0.8. The full set of selections given above will be referred to as the HWWID for muons in the following.

The efficiency of the HWWID requirement is evaluated in data and on a DY MC sample through the tag and probe method described in Section 4.2.1. On top of satisfying all the criteria outlined above, the tag is required to have

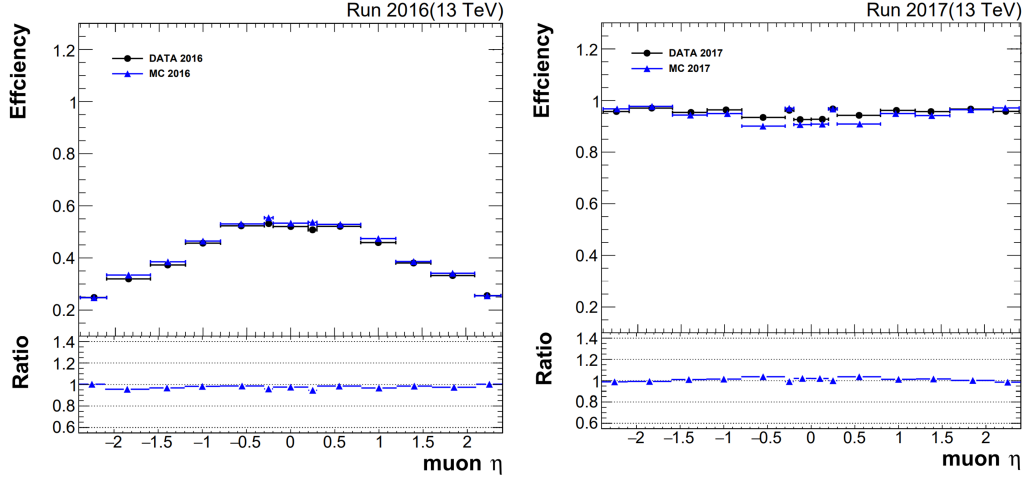


Figure 4.7: Examples of tag and probe efficiencies of the HWWID selection for muons as a function of η , for $10 < p_T < 15$ GeV in the 2016 data set (left) and $60 < p_T < 100$ GeV in 2017 data set (right). Efficiencies are shown for both data and MC events, and the ratio of the two is used as a correction factor.

1429 $p_T > 29$ GeV in order to reduce the contamination from non-prompt muons.
 1430 For data, the tag is also required to match the object that fired the single
 1431 muon HLT; this removes any trigger-induced bias in the procedure. The
 1432 probe is required to satisfy a much looser set of identification criteria, and
 1433 the tag and probe pair is required to have an invariant mass between 70 and
 1434 130 GeV, therefore selecting $Z \rightarrow \mu\mu$ events. A kinematic fit to the Z boson
 1435 mass peak is performed to eliminate the background contribution. Finally,
 1436 the efficiency can be evaluated by counting the number of probes that also
 1437 satisfy the full set of requirements for a muon in the analysis, divided by the
 1438 total number of selected probes. This procedure is carried out in bins of η
 1439 and ϕ of the probe, on both data and MC. The ratio of the data and MC
 1440 efficiencies, or scale factor, is then used to correct the HWWID efficiency in
 1441 MC, as a function of η and ϕ of the muon. Most scale factors are found to
 1442 be very close to unity, with rare outliers at $\sim 5\%$ from one. An example of
 1443 the resulting efficiencies as a function of η of the muon is shown in Fig. 4.7.
 1444 The systematic uncertainty on the the HWWID efficiency is evaluated by
 1445 repeating the procedure while varying the isolation requirement on the tag
 1446 muon, as well as changing the fit function, binning, and range of the fit to the
 1447 Z boson mass, and evaluating the corresponding change in the scale factors.
 1448 The largest change in each (η, ϕ) bin is taken as the systematic uncertainty,
 1449 which is found to be below 1% on average, and around 6% at most.

4.2.3 Electrons

On top of the identification and isolation criteria outlined in Section 2.4, and additional set of requirements is imposed on electrons entering the analysis in order to reduce the non-prompt lepton contamination. Instead of a selection based on a set of track quality criteria as is applied on muons, for electrons the HWWID is based on a multivariate approach. A BDT is built from a set of observables pertaining to the electron reconstruction in both the tracker and the calorimeter:

- calorimetric observables include the ratio of the energy deposits in the ECAL and the HCAL, difference in position between the track and calorimeter cluster, the shape and position of the cluster in (η, ϕ) , and the ratio of the momentum of the track to the energy of the corresponding cluster;
- track quality observables include the number of hits in the tracker along the trajectory of the track, the χ^2 of the track fit, as well as the number of pixel layers without a hit corresponding to the track. This last quantity is aimed at suppressing electrons originating from the conversion of a photon in the material of the tracker, which manifests as a track originating after the innermost layers of the pixel tracker.

A selection is then performed on the output of the BDT, tuned to have a 90% efficiency on genuine electrons. On top of the requirement on the BDT output, the isolation of the electron (defined in Section 2.4) is required to be below 0.06.

The same tag and probe method described for muons is applied to electrons to correct the HWWID efficiency in MC from data. Examples of the resulting efficiencies and scale factors are shown in Fig. 4.8. Compared to the muon case outlined before, the scale factors for electrons are found to be farther from unity on average, with deviations of up to 15%. The corresponding uncertainties are typically of the order of 2-5%, with the highest ones being around 10%.

4.2.4 Hadronic jets

On top of the corrections to energy response outlined in Section 2.4, hadronic jets in the analysis are required to satisfy a set of quality criteria aimed at suppressing jets originating from spurious sources (e.g., PU). These include requirements on the energy fraction assigned to neutral hadrons, charged hadrons, and photons in the jet, varying depending on the η region of the

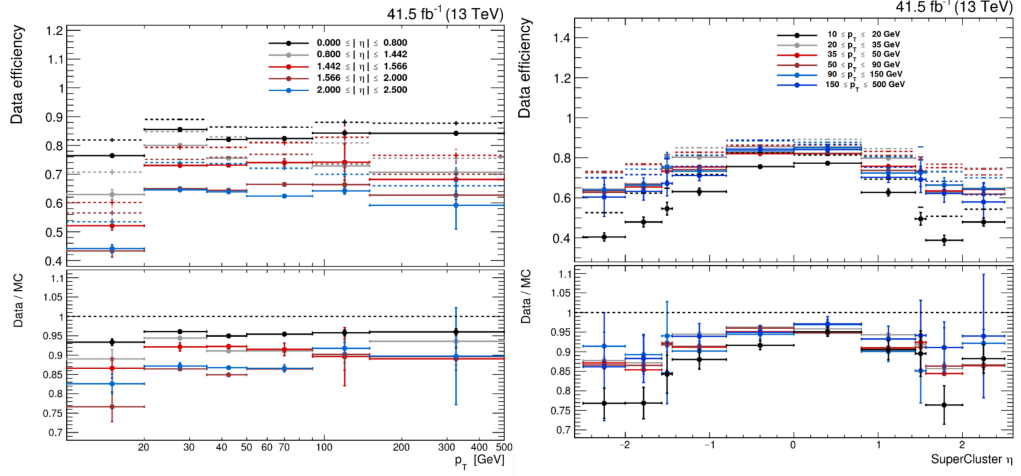


Figure 4.8: Examples of the tag and probe efficiencies of the HWWID selection for electrons as a function of p_T (left) and η (right) in the 2017 data set. In the upper panels, solid and dashed lines indicate efficiencies calculated in data and MC, respectively. The lower panels show the ratio of the two, which is used as a correction factor.

1486 detector. In MC samples corresponding to the 2017 and 2018 data-taking
 1487 periods, the jet energy resolution (JER) was found to be overestimated when
 1488 compared to data; to correct for this discrepancy, the energy of jets in MC
 1489 is smeared to match the one observed in data by multiplying their energy by
 1490 a smearing factor, which is calculated in two ways:

- 1491 • If for a reconstructed jet a matching generator level jet (i.e., a jet
 1492 clustered from generator-level particles) is found, the correction factor
 1493 c_{JER} is calculated as:

$$c_{\text{JER}} = 1 + (s_{\text{JER}} - 1) \frac{p_T - p_T^{\text{gen}}}{p_T}, \quad (4.10)$$

1494 where s_{JER} is the data to MC resolution scale factor.

- 1495 • If no matching generator level jet is found in the event, a stochastic
 1496 smearing is done instead, with the correction factor calculated as:

$$c_{\text{JER}} = 1 + \mathcal{N}(0, \sigma_{\text{JER}}) \sqrt{\max(s_{\text{JER}}^2 - 1, 0)}, \quad (4.11)$$

1497 where σ_{JER} is the p_T resolution in MC events. The downside of this ap-
 1498 proach is that the resolution can only be worsened, while the matching
 1499 method also allows for improvement.

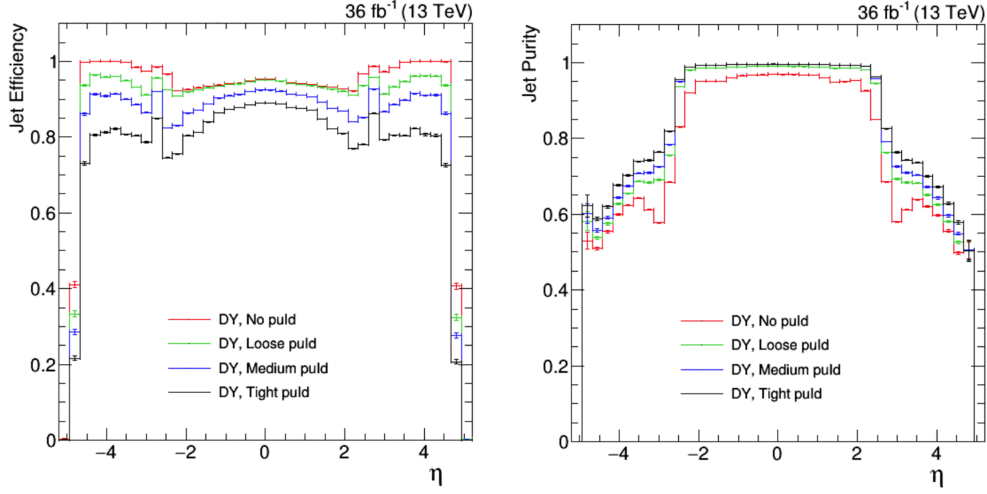


Figure 4.9: Efficiency (left) and purity (right) of the PUJetID requirements, evaluated on a MC sample of DY events corresponding to the 2016 data set. Different working points, labelled as tight, medium, and loose are shown, with the loose working point being the one used in the analysis.

The effect of applying the JER smearing has been checked in all three data sets of the 2016–2018 period. On the 2017 and 2018 data sets, it was found to significantly improve the data to MC agreement, especially in DY events. When applied to the 2016 data set however, a degrading effect on the data to MC agreement was observed, and JER smearing was therefore not applied.

Finally, to better reject jets arising from PU interactions, an additional requirement, labelled PUJetID, is imposed. The PUJetID is defined on the score of a BDT trained to discriminate genuine jets, arising from hard scattering, from PU jets, using information on the charged and neutral particle multiplicities in the jet, as well as its geometric shape. The threshold on the BDT output is tuned in bins of η and ϕ of the jet to achieve a 99% efficiency on non-PU jets. The PUJetID requirement is applied only on jets with $p_T < 50$ GeV. An example of the resulting efficiency and purity evaluated on a MC sample of DY events corresponding to the 2016 data set is shown in Fig. 4.9.

4.2.5 b tagging

As anticipated in Section 3.2, the peculiar structure of jets originating from a b quark enables their recognition through so-called b tagging algorithms, or b taggers. The b tagger used for all categories of the analysis is the DeepCSV

algorithm [52], with the exception of the VBF DF category in which the DEEPFLAVOR tagger [53] was found to perform better. The tagger assigns a score to each jet in the event, quantifying the likelihood that it originated from a b quark. Different thresholds on the score can be selected to consider a jet as b tagged. These are referred to as loose, medium, and tight working points (WPs), and correspond to a probability of mistakenly b tagging a light flavor jet of 10, 1, and 0.1%, respectively. Unless otherwise stated, the loose WP is used in this analysis, since it gives the highest b jet rejection power. The b tagging score is calculated on all jets with $p_T > 20$ GeV and $|\eta| < 2.4$.

To improve the agreement between data and MC simulation, the latter is weighted by a set of scale factors. These are calculated per-jet as a function of its η and p_T , flavor of the initiating quark, and b tagger score. This allows to also correct the shape of the b tagger score in simulation to better describe data. The weight that is applied to each MC event is calculated as a product of the corresponding per-jet scale factors, including all jets on which the b tagger is evaluated:

$$w_{\text{event}} = \prod_{i=0}^{N_{\text{jets}}} \text{SF}(\eta_i, p_{T,i}, D_i), \quad (4.12)$$

where D_i is the b tagger score for the i -th jet in the event. The score of the DeepCSV algorithm, after applying the correction, is shown in Fig. 4.10, 4.11, and 4.12 in control regions enriched in t quark and DY events, for the 2016, 2017, and 2018 data sets, respectively.

4.2.6 Non-prompt leptons

The contribution from non-prompt leptons in the analysis phase space is estimated directly from data. A control sample is defined by selecting events containing one lepton passing the identification and isolation requirements, and a second lepton that fails them but passes a looser set of criteria (pass-fail sample). This sample is dominated by non-prompt leptons. A second independent sample, enriched in dijet events, is used to estimate the probability of an object that satisfies the looser lepton identification criteria to also satisfy the full analysis requirements (non-prompt rate, R_{NP}). This probability is then used to re-weight events from the pass-fail sample:

$$w_{\text{NP}} = \frac{R_{\text{NP}}(\eta, p_T)}{1 - R_{\text{NP}}(\eta, p_T)}, \quad (4.13)$$

yielding the contribution from non-prompt leptons in the signal region. The non-prompt rate is measured as a function of the object's p_T and η . Sources of prompt leptons in the dijet sample are subtracted using MC simulation.

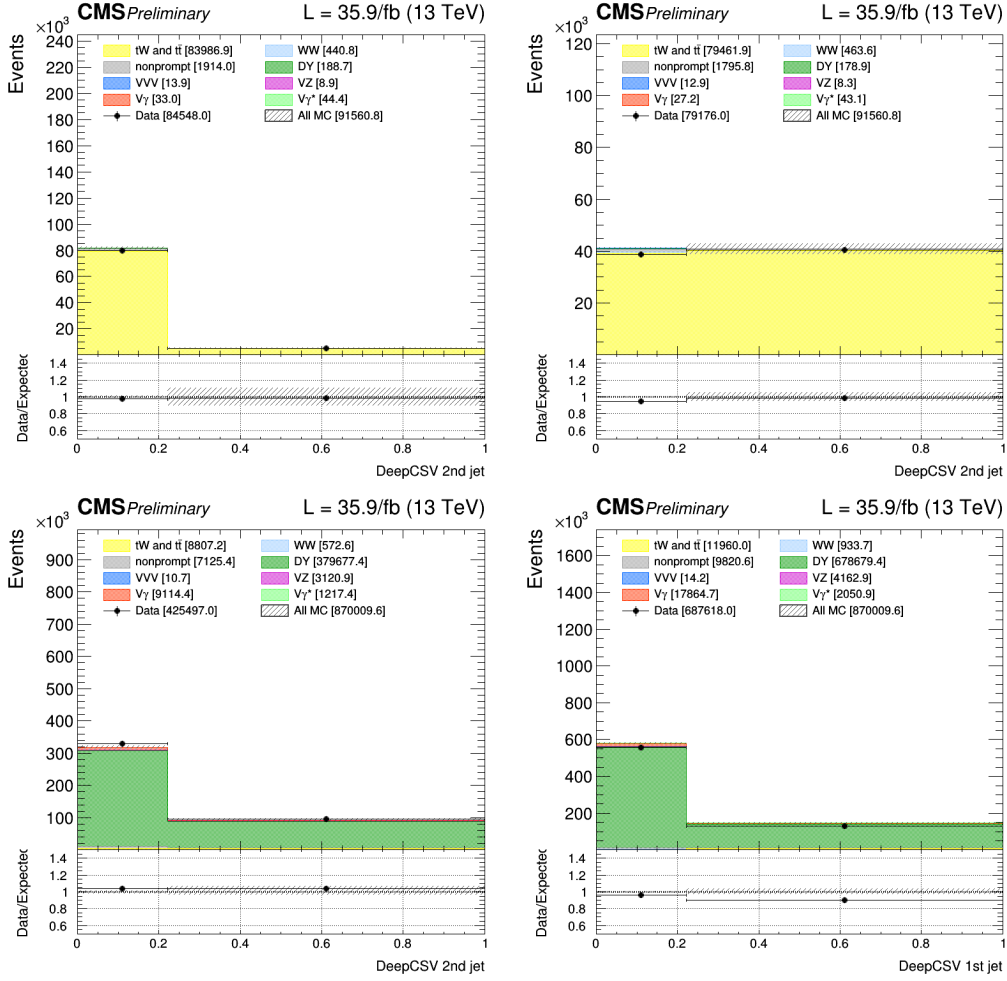


Figure 4.10: Data to MC comparison of the DeepCSV score after applying the correction outlined above, for the 2016 data set. Events are shown in control regions enriched in t quark (top) and DY (bottom) events, for jets with p_T below (left) and above (right) 30 GeV. The boundary between the two bins is the threshold on the score corresponding to the loose WP.

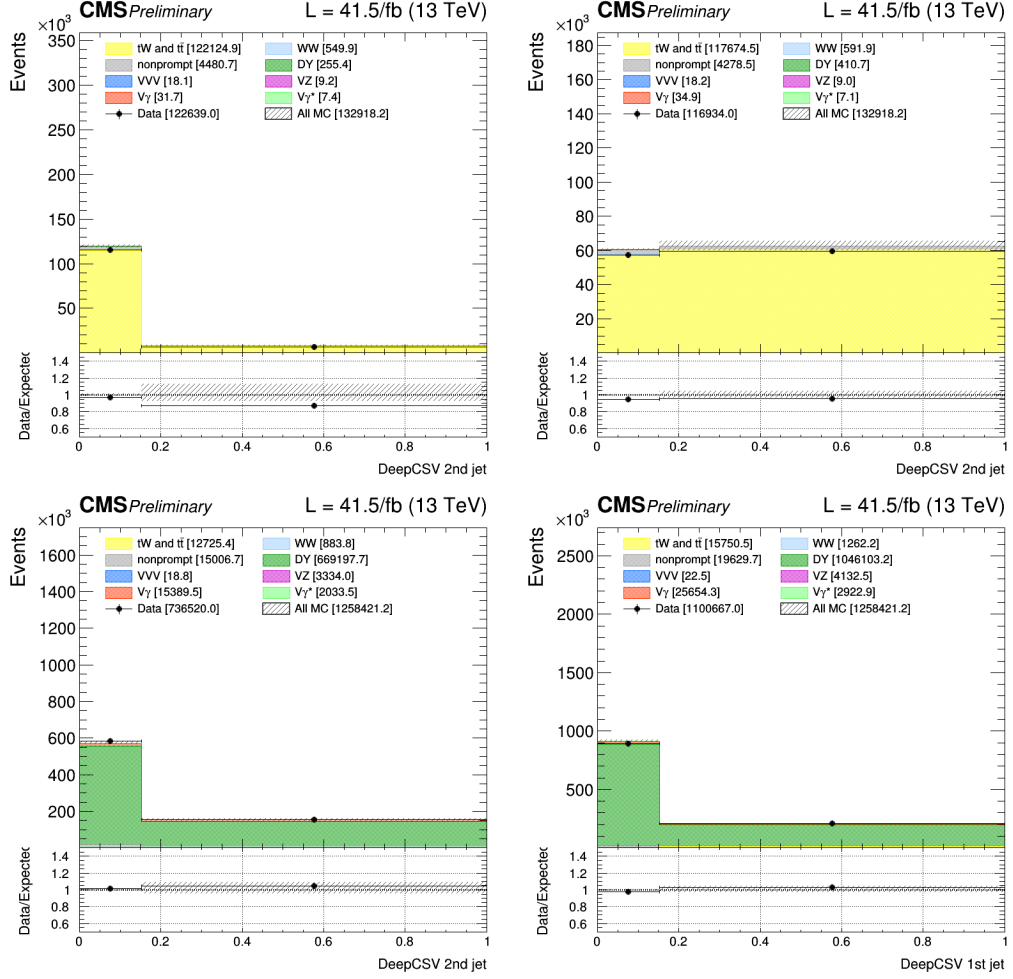


Figure 4.11: Data to MC comparison of the DeepCSV score after applying the correction outlined above, for the 2017 data set. Events are shown in control regions enriched in t quark (top) and DY (bottom) events, for jets with p_T below (left) and above (right) 30 GeV. The boundary between the two bins is the threshold on the score corresponding to the loose WP.

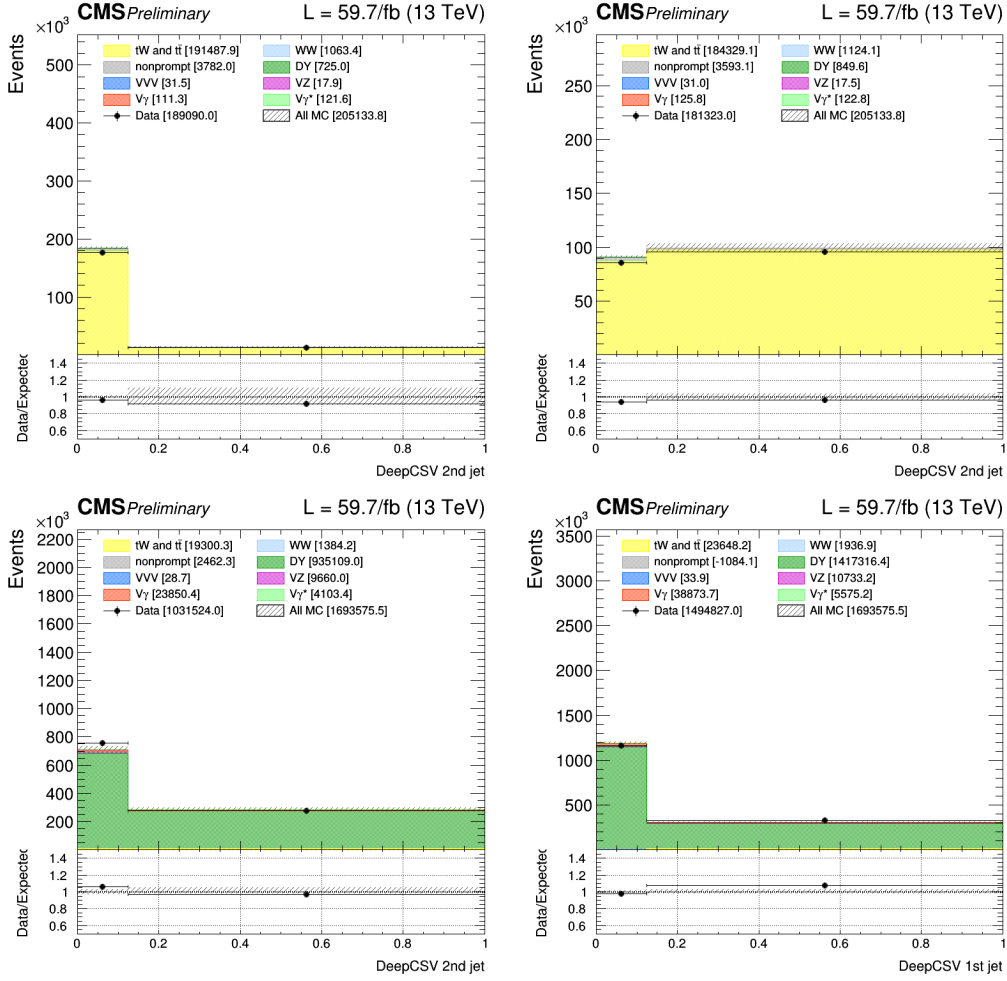


Figure 4.12: Data to MC comparison of the DeepCSV score after applying the correction outlined above, for the 2018 data set. Events are shown in control regions enriched in t quark (top) and DY (bottom) events, for jets with p_T below (left) and above (right) 30 GeV. The boundary between the two bins is the threshold on the score corresponding to the loose WP.

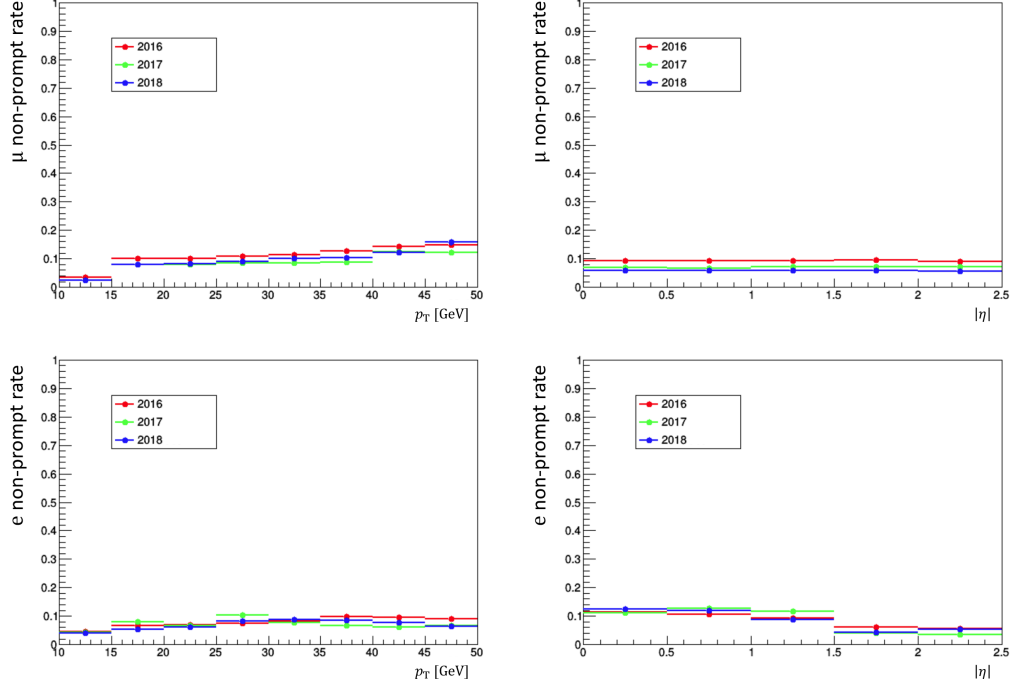


Figure 4.13: Non-prompt rates for muons (top) and electrons (bottom), as a function of p_T (left) and η (right).

1552 Since the non-prompt rate is estimated in data and is then applied to a
 1553 sample likewise selected in data, this method allows to extract an estimate of
 1554 the non-prompt lepton distribution with a negligible reliance on simulation.
 1555 Examples of the non-prompt rates measured as described above are shown
 1556 in Fig. 4.13. The method is validated in a CR in which dilepton pairs with
 1557 same electric charge are required. Since scattering processes that give rise
 1558 to same-sign lepton pairs are rare, such a sample is enriched in non-prompt
 1559 leptons. A selection of data to MC comparisons is shown in Fig. 4.14.

1560 4.3 Event selection and categorization

1561 In order to effectively measure all of the target H boson production mech-
 1562 anisms, events are divided into mutually exclusive categories designed to
 1563 exploit the peculiarities of each final state, as detailed in Section 3.1. The di-
 1564 vision is based on the number and flavor composition of charged leptons and
 1565 the kinematics and number of hadronic jets in the final state, and is summa-
 1566 rized in Table 4.3. In each of these categories, as anticipated in Section 3.2,
 1567 a kinematic selection is applied on events in order to suppress background

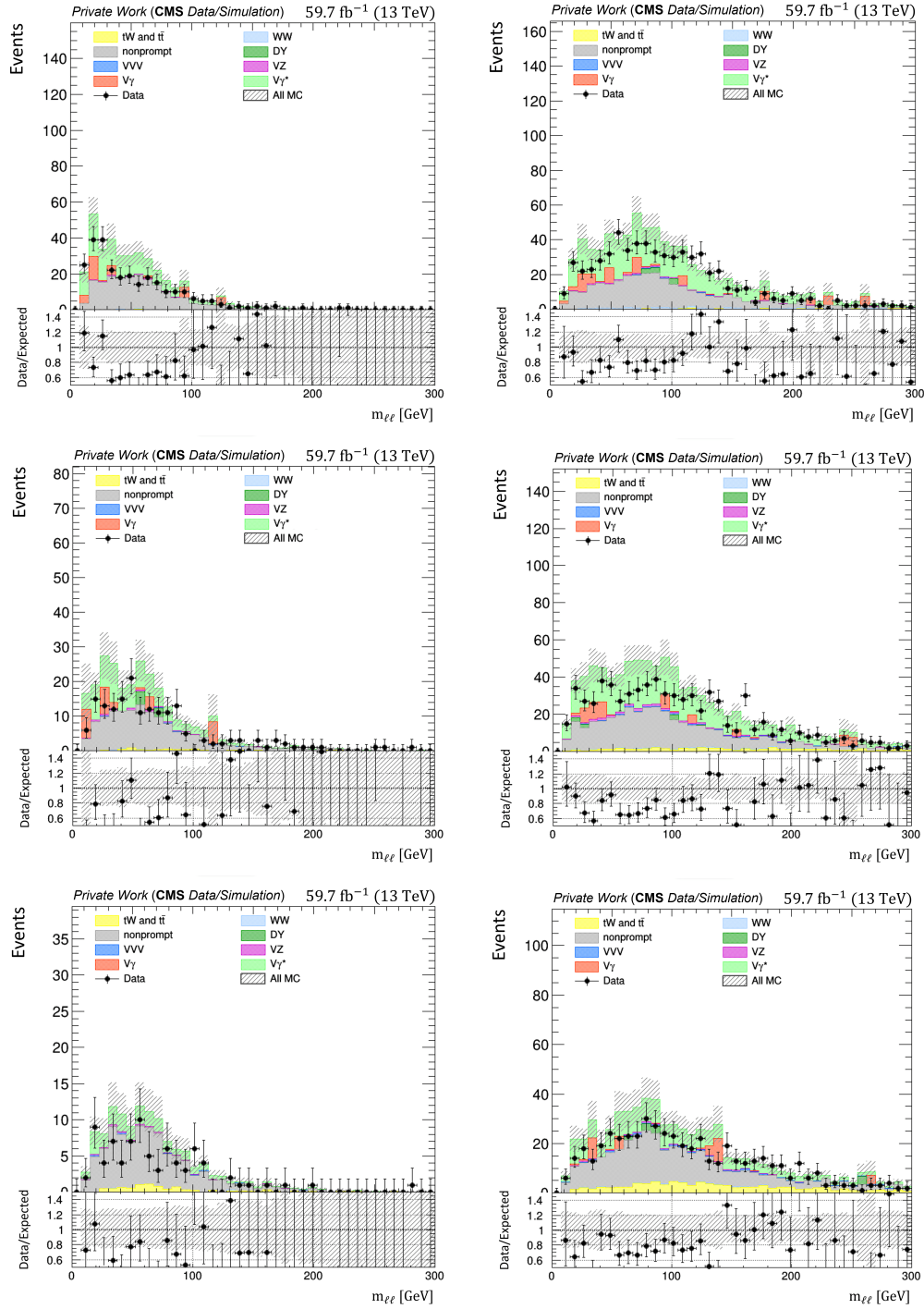


Figure 4.14: Validation of the non-prompt lepton estimation technique in a same-sign CR with 0 (top), 1 (middle), and 2 (bottom) jets, and $p_{\text{T}}^{\ell_2}$ below (left) or above (right) 20 GeV. The 2018 data set is shown for reference.

processes while simultaneously discard as few signal events as possible. All events are required to contain at least two charged leptons, with the leading (subleading) one having p_T above 25(10) GeV. If the subleading lepton is an electron, the threshold is moved to 13 GeV. For events in the 2017 and 2018 data sets the threshold on the subleading lepton is instead 13 GeV regardless of its flavor, due to the change in trigger conditions described in Section 4.1. Unless otherwise stated, jets are counted only if their p_T is above 30 GeV. The signal extraction is performed through a binned maximum likelihood fit, as will be detailed in Section 5.1, to a set of kinematic distributions chosen in each category to maximize the sensitivity to the signal.

Table 4.3: Overview of the selection criteria defining the analysis categories.

Category	$N_{leptons}$	N_{jets}	Subcategorization
ggH	2	-	(DF, SF) $\times$ (0 jets, 1 jet, ≥ 2 jets)
VBF	2	≥ 2	(DF, SF)
VH2j	2	≥ 2	(DF, SF)
WHSS	2	≥ 1	(DF, SF) $\times$ (1 jet, 2 jets)
WH3 ℓ	3	0	SF lepton pair with opposite or same sign
ZH3 ℓ	3	≥ 1	(1 jet, 2 jets)
ZH4 ℓ	4	-	(DF, SF)

1577

4.3.1 Different flavor ggH category

On top of the common lepton requirements, the ggH DF category is defined by requiring the two leptons in the event to form an oppositely charged $e\mu$ pair. Events containing additional leptons with $p_T > 10$ GeV are excluded from the selection. The main backgrounds in this category are non-resonant WW production, t quark production, DY production of τ leptons, and non-prompt leptons, arising mainly from W+jets events with a jet misidentified as a lepton. Additional sources of background include diboson and triboson production.

Events are required to have $m_{\ell\ell} > 12$ GeV to suppress QCD events with multiple misidentified jets. In order to select events with genuine missing momentum, p_T^{miss} is required to be at least 20 GeV. Contamination from events containing t quarks is reduced by rejecting any that contain a b tagged jet with p_T above 20 GeV. Events containing pairs of τ leptons are reduced by requiring the p_T of the dilepton system $p_T^{\ell\ell}$ to be above 30 GeV, since leptons from W bosons have, on average, higher p_T than those coming from τ lepton decays. The transverse mass of the H boson m_T^H is required to

be above 60 GeV. The transverse mass of the subleading lepton and $\vec{p}_T^{\text{miss}}$ ($m_T(\ell_2, \vec{p}_T^{\text{miss}})$), built following the definition in Eq. 3.5, is required to be at least 30 GeV. This further reduces contamination from τ lepton pairs and W+jets events. Finally, for events with at least two jets, the leading dijet pair is required to satisfy the condition $m_{jj} < 65$ GeV or $105 < m_{jj} < 120$ GeV, in order to make the selection orthogonal to the VBF and VH2j ones described in the following.

In order to enhance the sensitivity to the signal, events are further divided into subcategories. A first division is done based on the number of jets in the event, separating events with zero, one, and more than one jet, given the different background compositions: e.g., contamination from t quark production is subdominant in events with no jets, roughly equal to non-resonant WW production in events with one jet, and becomes dominant in events with more than one jet. Events with less than two jets are further subdivided based on the electric charge of the leading lepton. This aids in constraining W+jets events, since given the net positive charge of the initial p-p system in the collision, W^+ bosons are produced more abundantly than W^- bosons, while the signal is symmetric in this regard. Finally, a further division is performed depending on if the p_T of the subleading lepton ($p_T^{\ell_2}$) is above or below 20 GeV. The reason for this further subdivision is due to the fact that the low $p_T^{\ell_2}$ phase space is where the signal is concentrated (due to the virtuality of one of the W bosons in the $H \rightarrow WW$ decay), but it is also where the background from non-prompt leptons is more important, allowing for a tighter constraint on the latter.

To better constrain the main sources of background, additional regions of phase space enriched in a given background are defined. These control regions (CRs) enter the fit and serve as an in-situ measurement of the background yields in data, which are also propagated to the signal regions (SRs). Separate CRs are defined for the different jet multiplicity SRs. The selections are the same, with the following modifications:

- CRs for t quark production are defined by inverting the veto on events with b jets i.e., by requiring at least one of the jets in the event to be b tagged. In categories with no jets with p_T above 30 GeV, the CR is defined by requiring at least a b tagged jet with $20 < p_T < 30$ GeV. The threshold on $m_{\ell\ell}$ is increased to 50 GeV, and the m_T^H requirement is dropped.
- CRs for DY production of τ lepton pairs are defined by inverting the m_T^H requirement. To better select $Z \rightarrow \tau\tau$, events in this region are further required to have $40 < m_{\ell\ell} < 80$ GeV.

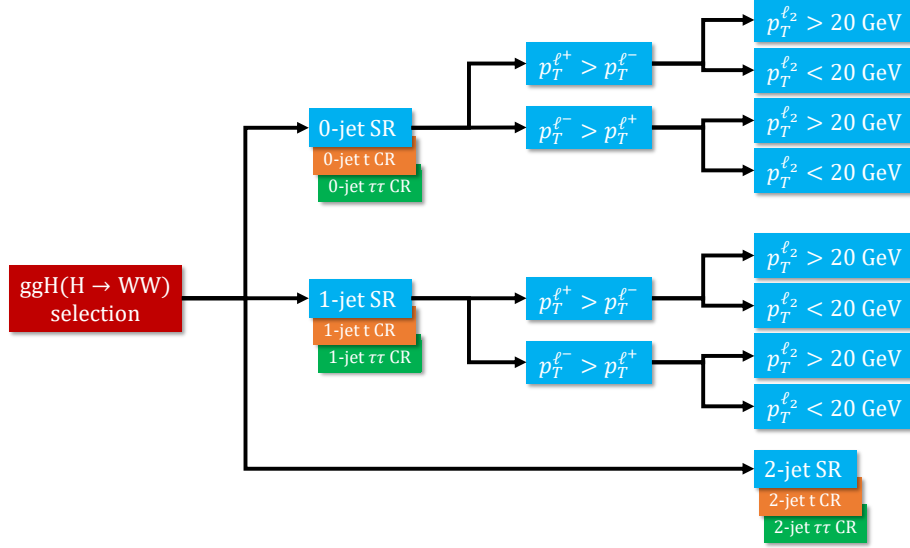


Figure 4.15: Visual overview of the categorization strategy in the ggH DF channel.

1634 The full selection and subcategorization scheme is given in Table 4.4 and
 1635 summarized in Fig. 4.15.

1636 As explained in Section 3.1, the scalar H resonance determines pecu-
 1637 liar distributions of the $m_{\ell\ell}$ and m_T^H observables in signal events. This fact
 1638 is exploited by using the two-dimensional $(m_{\ell\ell}, m_T^H)$ distribution itself as
 1639 the variable on which to perform the signal extraction fit. Reliance on the
 1640 $(m_{\ell\ell}, m_T^H)$ distribution has the additional advantage that signal and non-
 1641 resonant WW backgrounds are distributed differently in this space, as can
 1642 be seen in Figure 4.16. This allows to measure the two simultaneously, effec-
 1643 tively constraining the non-resonant WW yield in data without the need of
 1644 an additional CR. Observed distributions of the $m_{\ell\ell}$ and m_T^H observables are
 1645 shown in Fig. 4.17, 4.18, and 4.19. One-dimensional marginal distributions
 1646 are shown for the sake of clarity. The background and signal yields are scaled
 1647 to their post-fit values, as will be detailed in Chapter 5. The uncertainty
 1648 bands on the MC prediction include all systematic contributions, including
 1649 the uncertainty due to the finite sample size of the simulation. Error bars
 1650 on data points indicate the corresponding statistical uncertainty. The signal
 1651 is shown both stacked on the background as well as superimposed, in order
 1652 to better highlight its shape with respect to the background. The $m_{\ell\ell}$ and
 1653 m_T^H distributions in the t quark and $\tau\tau$ CRs are shown in Fig. 4.20 and 4.21,

Table 4.4: Summary of the selection used in different-flavor ggH category.

Subcategories	Selection
<u>Global selection</u>	
-	$p_T^{\ell_1} > 25 \text{ GeV}, p_T^{\ell_2} > 10 \text{ GeV}$ (2016) or 13 GeV $p_T^{\text{miss}} > 20 \text{ GeV}, p_T^{\ell\ell} > 30 \text{ GeV}, m_{\ell\ell} > 12 \text{ GeV}$ $e\mu$ pair with opposite charge
<u>0-jet ggH category</u>	
$\ell^\pm \ell^\mp, p_T^{\ell_2} \leq 20 \text{ GeV}$	$m_T^H > 60 \text{ GeV}, m_T(\ell_2, p_T^{\text{miss}}) > 30 \text{ GeV}$ $p_T^{\ell_2} \leq 20 \text{ GeV}$ No jet with $p_T > 30 \text{ GeV}$ No b-tagged jet with $p_T > 20 \text{ GeV}$
Top quark CR	As SR but with no m_T^H requirement, $m_{\ell\ell} > 50 \text{ GeV}$ At least 1 b-tagged jet with $20 < p_T < 30 \text{ GeV}$
$\tau\tau$ CR	As SR but with $m_T^H < 60 \text{ GeV}$ $40 < m_{\ell\ell} < 80 \text{ GeV}$
<u>1-jet ggH category</u>	
$\ell^\pm \ell^\mp, p_T^{\ell_2} \leq 20 \text{ GeV}$	$m_T^H > 60 \text{ GeV}, m_T(\ell_2, p_T^{\text{miss}}) > 30 \text{ GeV}$ $p_T^{\ell_2} \leq 20 \text{ GeV}$ 1 jet with $p_T > 30 \text{ GeV}$ No b-tagged jet with $p_T > 20 \text{ GeV}$
Top quark CR	As SR but with no m_T^H requirement, $m_{\ell\ell} > 50 \text{ GeV}$ At least 1 b-tagged jet with $p_T > 30 \text{ GeV}$
$\tau\tau$ CR	As SR but with $m_T^H < 60 \text{ GeV}$ $40 < m_{\ell\ell} < 80 \text{ GeV}$
<u>2-jet ggH category</u>	
SR	$m_T^H > 60 \text{ GeV}, m_T(\ell_2, p_T^{\text{miss}}) > 30 \text{ GeV}$ $p_T^{\ell_2} \leq 20 \text{ GeV}$ At least 2 jets with $p_T > 30 \text{ GeV}$ No b-tagged jet with $p_T > 20 \text{ GeV}$ $m_{jj} < 65 \text{ GeV}$ or $105 < m_{jj} < 120 \text{ GeV}$
Top quark CR	As SR but with no m_T^H requirement, $m_{\ell\ell} > 50 \text{ GeV}$ At least one b-tagged jet with $p_T > 30 \text{ GeV}$
$\tau\tau$ CR	As SR but with $m_T^H < 60 \text{ GeV}$ $40 < m_{\ell\ell} < 80 \text{ GeV}$

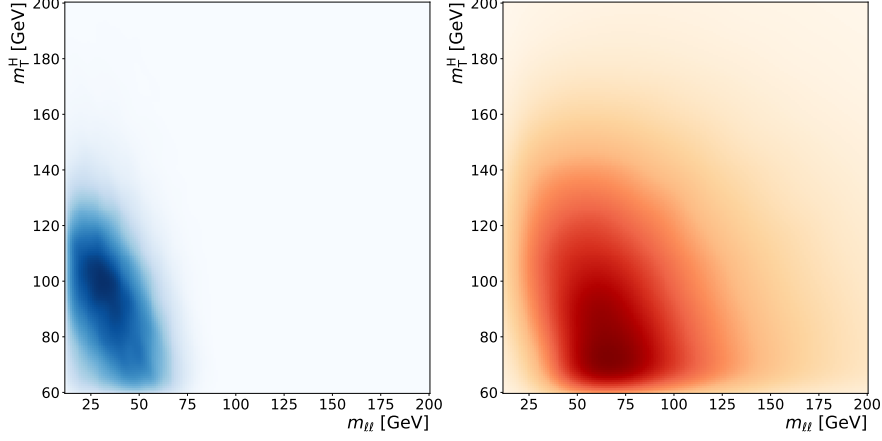


Figure 4.16: Distribution of ggH (left) and non-resonant WW events (right) in the two-dimensional $(m_{\ell\ell}, m_T^H)$ space.

1654 respectively.

1655 4.3.2 Same flavor ggH category

1656 The ggH SF categories require the presence of either an ee or a $\mu\mu$ pair with
 1657 opposite electric charge. Similarly to the DF case, events are required to
 1658 contain no b-tagged jets with $p_T > 20$ GeV. Low-mass resonances decaying
 1659 to a pair of electrons or muons are discarded by requiring $m_{\ell\ell} > 12$ GeV, and
 1660 the production of on-shell Z bosons is suppressed by requiring $|m_{\ell\ell} - m_Z| >$
 1661 15 GeV. As in the DF case, the selection is enriched in signal events by
 1662 requiring $p_T^{\ell\ell} > 30$ GeV and $p_T^{\text{miss}} > 20$ GeV, and events are divided into
 1663 subcategories based on the number of hadronic jets, as well as on the flavor
 1664 of the dilepton pair. All SRs also require $m_{\ell\ell} < 60$ GeV. Further requirements
 1665 on the azimuthal angle separation between the two leptons ($\Delta\phi_{\ell\ell}$) and m_T^H
 1666 are applied depending on the subcategory, as detailed in Table 4.5. Control
 1667 regions for t quark background events are defined by inverting the veto on b
 1668 jets analogously to the DF case. A second set of CRs enriched in non-resonant
 1669 WW events are defined by requiring $m_{\ell\ell} > 100$ GeV.

1670 By far the most important background contribution in this category is
 1671 DY production of dilepton pairs. A dedicated deep neural network (DNN),
 1672 referred to as DYMVA in the following, is used to reduce the DY contribution
 1673 in SF signal regions. The DYMVA is trained independently in each category,
 1674 and the threshold values are optimized separately in each category and data

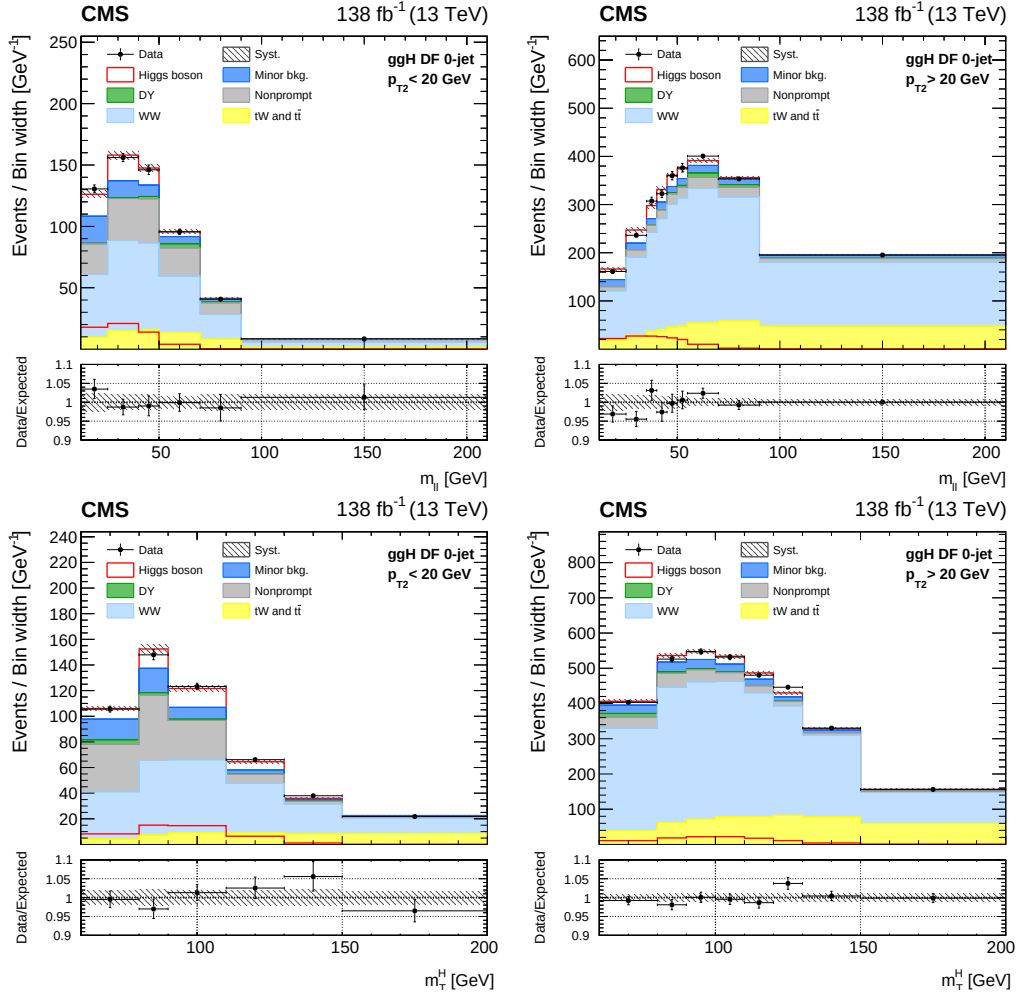


Figure 4.17: Observed distributions of the $m_{\ell\ell}$ (upper) and m_T^H (lower) fit variables in the 0-jet ggH $p_T^{\ell_2} < 20$ GeV (left) and $p_T^{\ell_2} > 20$ GeV (right) DF categories. The lower panel in each figure shows the ratio of the number of events observed in data to that of the total SM MC as extracted from the fit.

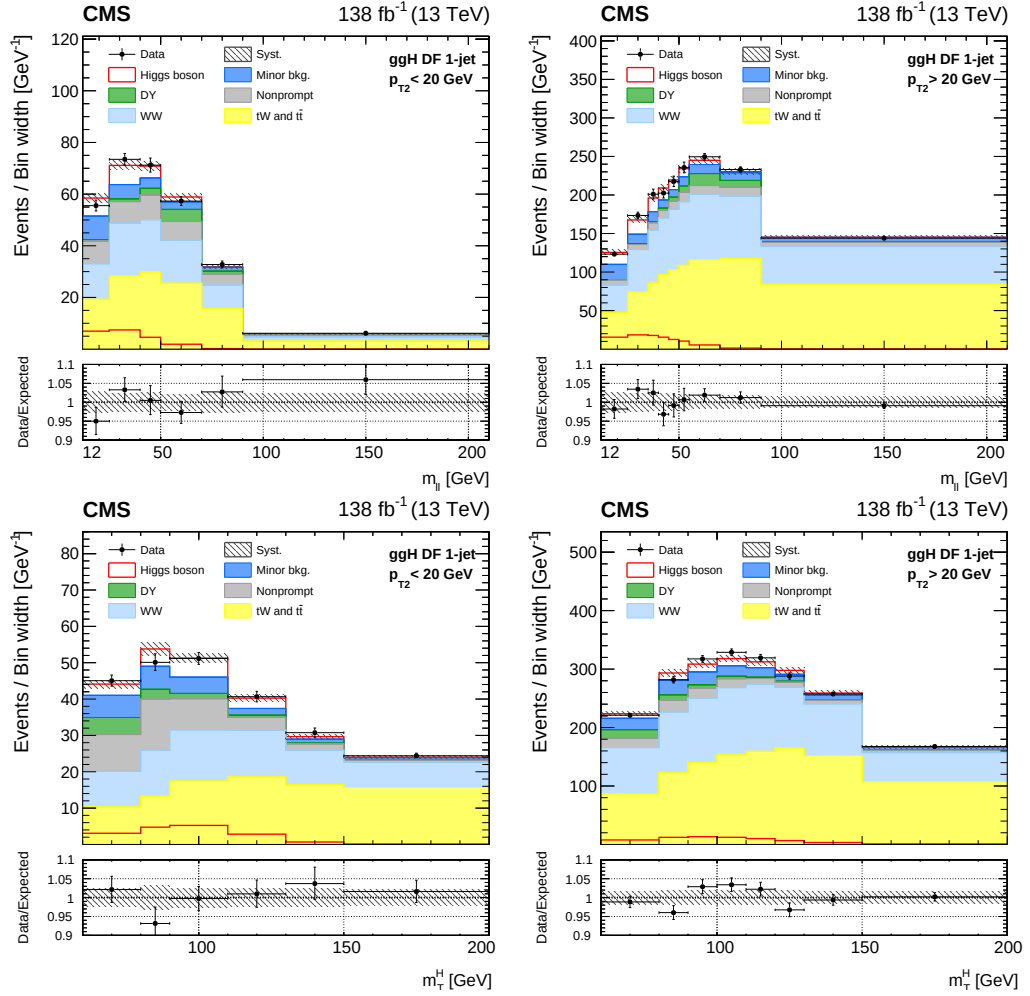


Figure 4.18: Observed distributions of the $m_{\ell\ell}$ (upper) and m_T^H (lower) fit variables in the 1-jet ggH $p_T^{\ell_2} < 20$ GeV (left) and $p_T^{\ell_2} > 20$ GeV (right) DF categories. The lower panel in each figure shows the ratio of the number of events observed in data to that of the total SM MC as extracted from the fit.

Table 4.5: Summary of the selection used in same-flavor ggH category. The DYMVA threshold is optimized separately in each subcategory and data set.

Subcategories	Selection
<u>Global selection</u>	
-	$p_T^{\ell_1} > 25 \text{ GeV}, p_T^{\ell_2} > 10 \text{ GeV}$ (2016) or 13 GeV $p_T^{\text{miss}} > 20 \text{ GeV}, p_T^{\ell\ell} > 30 \text{ GeV}$ ee or $\mu\mu$ pair with opposite charge $m_{\ell\ell} > 12 \text{ GeV}, m_{\ell\ell} - m_Z > 15 \text{ GeV}$
<u>0-jet ggH category</u>	
ee, $\mu\mu$	$m_{\ell\ell} < 60 \text{ GeV}, m_T^H > 90 \text{ GeV}, \Delta\phi_{\ell\ell} < 2.3$ No b-tagged jets with $p_T > 20 \text{ GeV}$ DYMVA above threshold
WW CR	As SR but with $m_{\ell\ell} > 100 \text{ GeV}$ $m_T^H > 60 \text{ GeV}, m_T(\ell_2, p_T^{\text{miss}}) > 30 \text{ GeV}$
Top quark CR	As SR but with $m_{\ell\ell} > 100 \text{ GeV}, m_T(\ell_2, p_T^{\text{miss}}) > 30 \text{ GeV}$ At least one b-tagged jet with $20 < p_T < 30 \text{ GeV}$
<u>1-jet ggH category</u>	
ee, $\mu\mu$	$m_{\ell\ell} < 60 \text{ GeV}, m_T^H > 80 \text{ GeV}, \Delta\phi_{\ell\ell} < 2.3$ No b-tagged jets with $p_T > 20 \text{ GeV}$ DYMVA above threshold
WW CR	As SR but with $m_{\ell\ell} > 100 \text{ GeV}$ $m_T^H > 60 \text{ GeV}, m_T(\ell_2, p_T^{\text{miss}}) > 30 \text{ GeV}$
Top quark CR	As SR but with $m_{\ell\ell} > 100 \text{ GeV}, m_T(\ell_2, p_T^{\text{miss}}) > 30 \text{ GeV}$ At least one b-tagged jet with $p_T > 30 \text{ GeV}$
<u>2-jet ggH category</u>	
ee, $\mu\mu$	$m_{\ell\ell} < 60 \text{ GeV}, 65 < m_T^H < 150 \text{ GeV}$ No b-tagged jets with $p_T > 20 \text{ GeV}$ DYMVA above threshold
WW CR	As SR but with $m_{\ell\ell} > 100 \text{ GeV}$ $m_T^H > 60 \text{ GeV}, m_T(\ell_2, p_T^{\text{miss}}) > 30 \text{ GeV}$
Top quark CR	As SR but with $m_{\ell\ell} > 100 \text{ GeV}, m_T(\ell_2, p_T^{\text{miss}}) > 30 \text{ GeV}$ At least one b-tagged jet with $p_T > 30 \text{ GeV}$

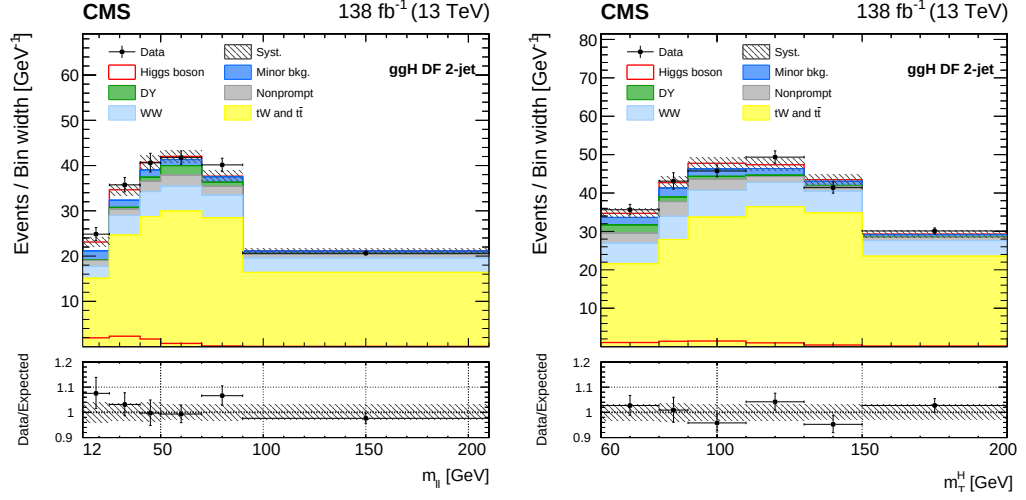


Figure 4.19: Observed distributions of the $m_{\ell\ell}$ (left) and m_T^H (right) fit variables in the 2-jet ggH DF category. The lower panel in each figure shows the ratio of the number of events observed in data to that of the total SM MC as extracted from the fit.

set. Because the instrumental nature of p_T^{miss} in DY production of e and μ pairs makes its estimation in MC unreliable, a data driven approach is used instead. The strategy consists in estimating the number of DY events in the SR by measuring the number of DY events $N_{\ell\ell}^{\text{in}}$ in the $m_{\ell\ell}$ region close to the resonant peak of the Z boson, by selecting a window $|m_{\ell\ell} - m_Z| < 7.5$ GeV. While mostly pure in $Z \rightarrow \ell\ell$ events, a residual contamination from t quark, WW, and W+jets production will be present. Exploiting the fact that these processes are symmetric in the flavor composition of the dilepton pair, their contribution can be measured in a DF region with the same $m_{\ell\ell}$ requirement and subtracted from the SF events in the Z boson peak. A correction factor $k_{\ell\ell}$ accounts for the difference in reconstruction efficiencies between electrons and muons. The number of DY events outside of the Z boson $m_{\ell\ell}$ peak, where the analysis is performed, can thus be rewritten as:

$$N_{\ell\ell}^{\text{out}} = RN_{\ell\ell}^{\text{in}} = R \left(N_{\ell\ell}^{\text{in}} - \frac{1}{2}k_{\ell\ell}(N_{e\mu}^{\text{in}} - N_{e\mu}^{\text{in}}(\text{VV})) - N_{\ell\ell}^{\text{in}}(\text{VV}) \right). \quad (4.14)$$

The transfer factor R is measured directly in data, and $N_{\ell\ell}^{\text{in}}(\text{VV})$ refers to non-DY events in the Z boson mass peak from sources other than t quark, WW, and W+jets production, which are estimated from MC. This procedure is carried out in the SR selection, modulo the cut on the DYMVA, which is relaxed in order to make the signal contribution negligible when measuring R . For this reason however, a second transfer factor has to be evaluated to

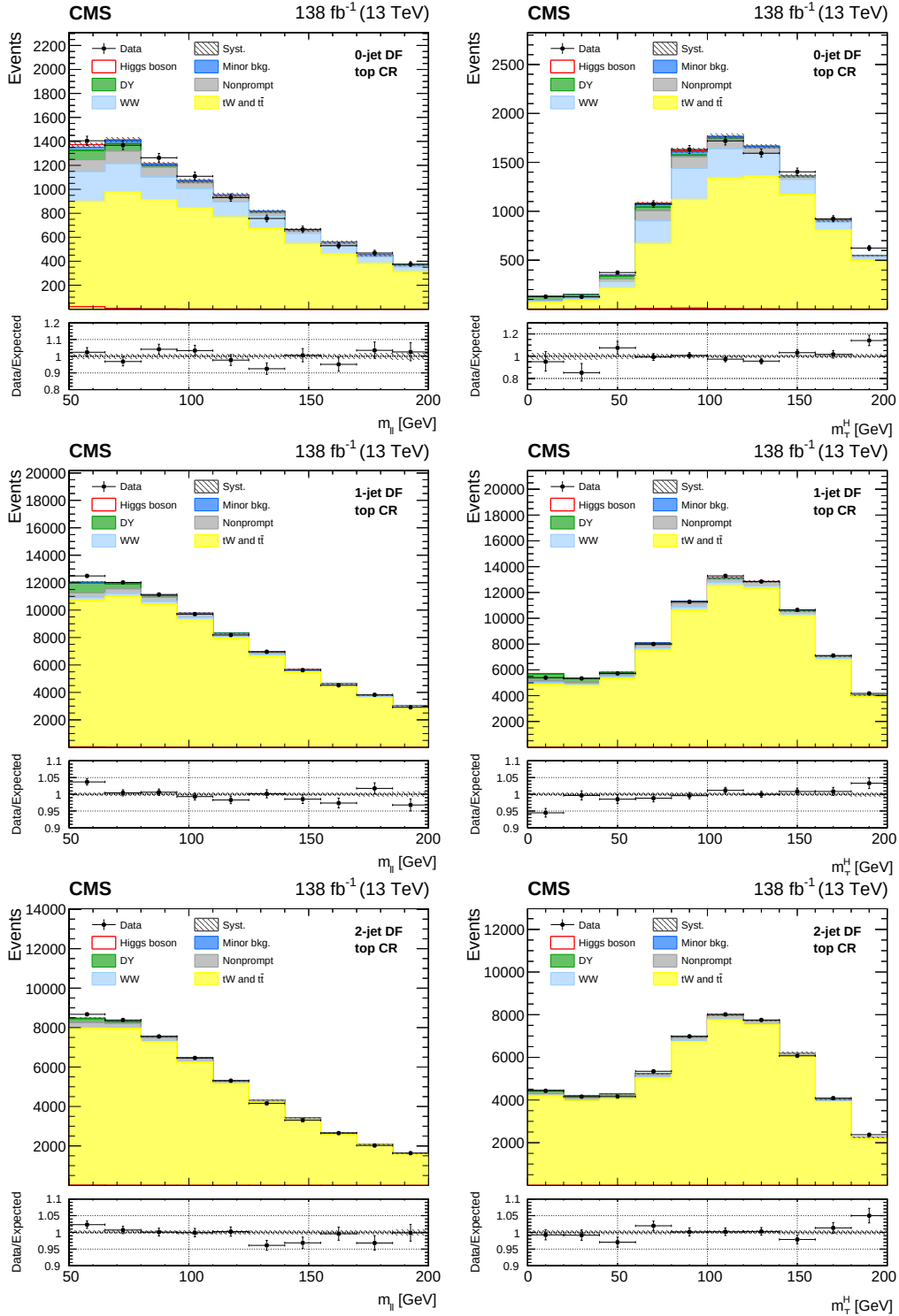


Figure 4.20: Observed distributions of the $m_{\ell\ell}$ (left) and m_T^H (right) fit variables in the 0- (top), 1- (center), and 2-jet (bottom) t quark CRs. The lower panel in each figure shows the ratio of the number of events observed in data to that of the total expected SM MC.

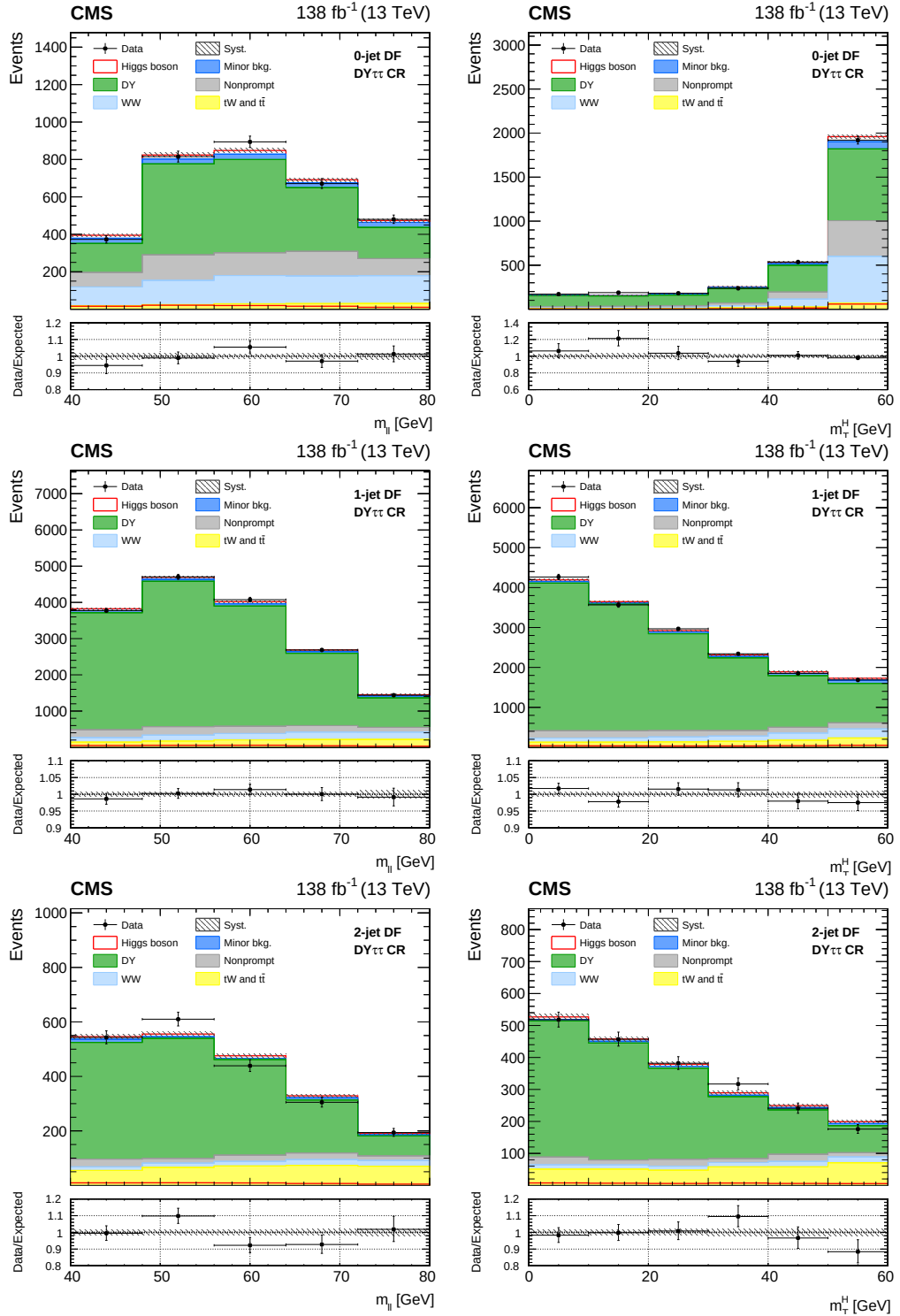


Figure 4.21: Observed distributions of the $m_{\ell\ell}$ (left) and m_T^H (right) fit variables in the 0- (top), 1- (center), and 2-jet (bottom) $\tau\tau$ CRs. The lower panel in each figure shows the ratio of the number of events observed in data to that of the total expected SM MC.

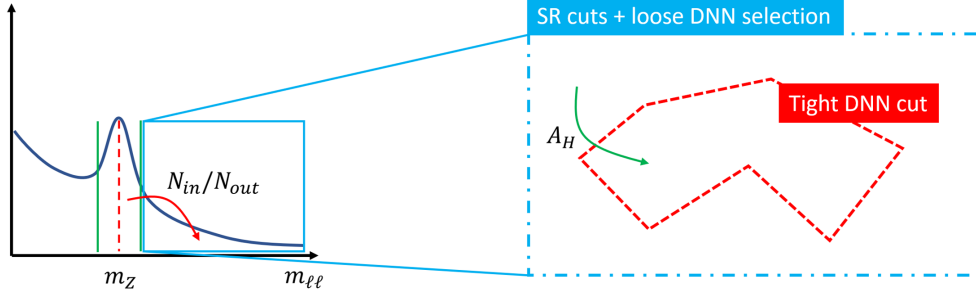


Figure 4.22: Visual summary of the procedure applied to estimate the DY contribution in SF categories.

1694 estimate the number of DY events once the cut on the DYMVA is tightened.
 1695 This acceptance factor is calculated in MC as:

$$A_H = \frac{N_{\ell\ell}^{\text{DYMVA loose}}}{N_{\ell\ell}^{\text{out}}}. \quad (4.15)$$

1696 The procedure is visualized in Fig. 4.22. The full procedure is validated in
 1697 a side band of the Z boson mass peak, defined by the requirement $7.5 < |m_{\ell\ell} - m_Z| < 15 \text{ GeV}$, and is shown in Fig. 4.23. While this procedure
 1698 makes the estimation of the DY background with instrumental p_T^{miss} much
 1699 more robust, only the total number of expected events can be measured
 1700 this way. For this reason, in this and other SF categories (described in the
 1701 following), the number of events in the SRs and CRs is used as a fit variable
 1702 i.e., a counting experiment is performed.
 1703

1704 4.3.3 Different flavor VBF category

1705 As for the DF ggH category, the target are events in which the W boson pair
 1706 originating from the H boson decays fully leptonically. The same require-
 1707 ments on the lepton p_T and p_T^{miss} are therefore applied. As anticipated in
 1708 Section 3.1, the distinguishing signature of VBF production comes from the
 1709 kinematics of the VBF jets. Events are therefore required to have exactly
 1710 two jets, whose invariant mass m_{jj} is required to be above 120 GeV. Events
 1711 containing any b tagged jet with $p_T > 20 \text{ GeV}$ are excluded from the defini-
 1712 tion of the SR. In this category, the DEEPFLAVOR tagger was used since it
 1713 was found to have better performance in this region. Two CRs are defined,
 1714 enriched in t quark and $\tau\tau$ events, respectively. The full selection is given in
 1715 Table 4.6.

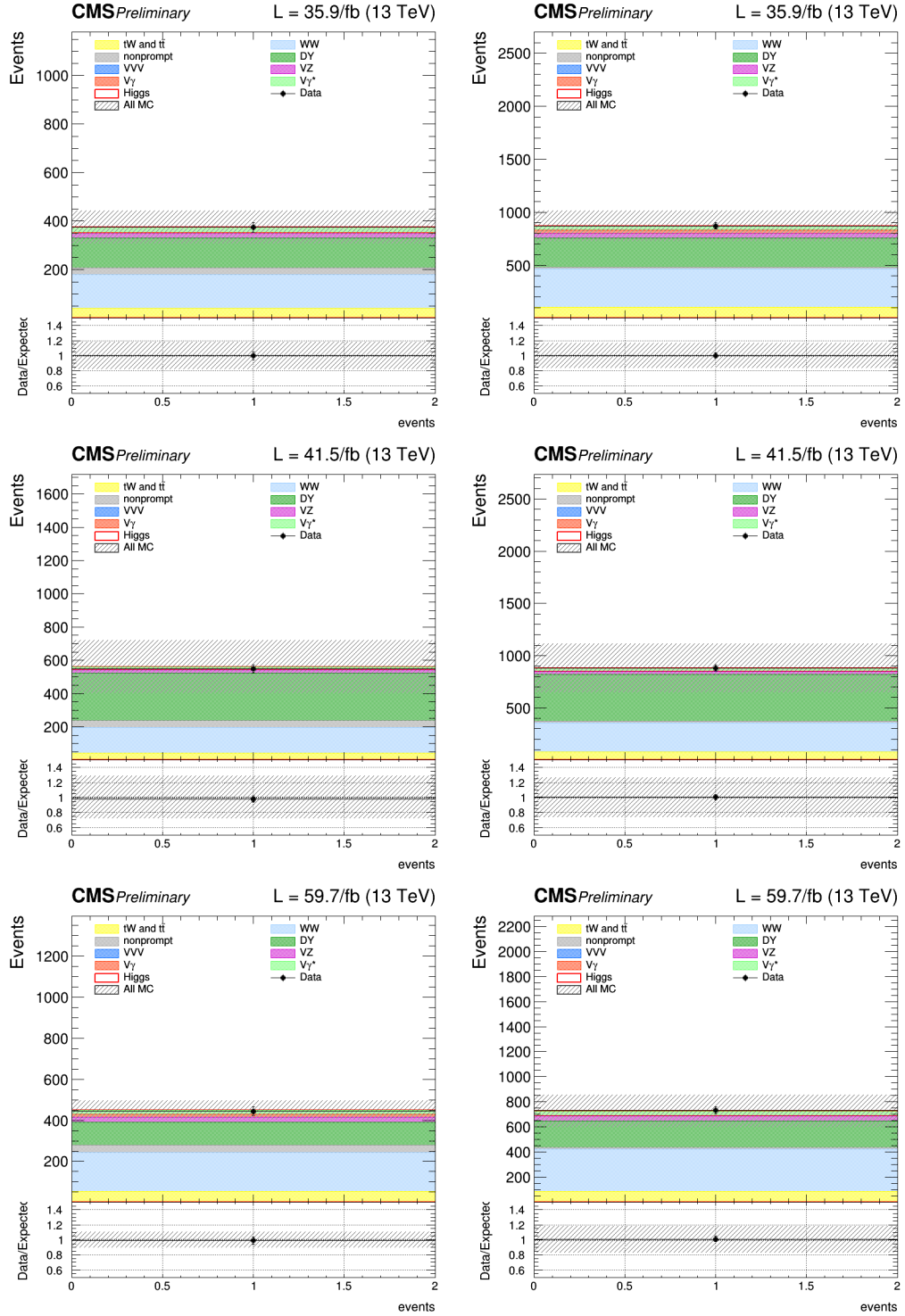


Figure 4.23: Validation of the DY background estimation method described above in a side band of the Z boson mass peak for ee (left) and $\mu\mu$ (right) events in the 2016 (top), 2017 (middle), and 2018 (bottom) data sets.

Table 4.6: Summary of the selection used in the different-flavor VBF category.

Subcategories	Selection
<u>Global selection</u>	
-	$p_T^{\ell_1} > 25 \text{ GeV}, p_T^{\ell_2} > 10 \text{ GeV}$ (2016) or 13 GeV $p_T^{\text{miss}} > 20 \text{ GeV}, p_T^{\ell\ell} > 30 \text{ GeV}, m_{\ell\ell} > 12 \text{ GeV}$ $e\mu$ pair with opposite charge
<u>2-jet VBF category</u>	
SR	$60 < m_T^H < 125 \text{ GeV}, m_T(\ell_2, p_T^{\text{miss}}) > 30 \text{ GeV}$ 2 jets with $p_T > 30 \text{ GeV}, m_{jj} > 120 \text{ GeV}$ No b-tagged jet with $p_T > 20 \text{ GeV}$
Top quark CR	As SR but with no m_T^H requirement, $m_{\ell\ell} > 50 \text{ GeV}$ At least one b-tagged jet with $p_T > 30 \text{ GeV}$
$\tau\tau$ CR	As SR but with $m_T^H < 60 \text{ GeV}$ $40 < m_{\ell\ell} < 80 \text{ GeV}$

Arguably the prime challenge in a successful measurement of the VBF cross section comes from ggH production. The ggH cross section is predicted to be one order of magnitude higher, and hadronic jets can arise due to initial state radiation. Since the decay kinematics of the H boson are common between the two processes, this gives rise to signatures that are not easy to separate with cut-based methods. For this reason, a multivariate approach has been developed in this category. A DNN has been trained to classify events in four mutually exclusive classes, targeting VBF, ggH, t quark, and WW events, respectively. This is achieved by having four output nodes O_i , normalized to $\sum_i O_i = 1$. The training is done on a MC sample containing a mixture of the four processes of interest in equal proportions. During training, the MC sample is split into three parts, labelled training, validation, and test samples, respectively. The training sample is the one on which the actual training is performed. The validation sample is used to monitor the loss function during the training: as long as the DNN learns genuine features of the target processes, the loss function will decrease on both the training and the validation sample. Once the network starts learning features of the training sample that do not generalize to the independent validation sample (overtraining), the loss on the latter will increase. The configuration of the network that is used in the analysis is chosen from the last training step in which the loss function on the validation set is still decreasing, a procedure known as early stopping. Finally, the performance of the network is evaluated

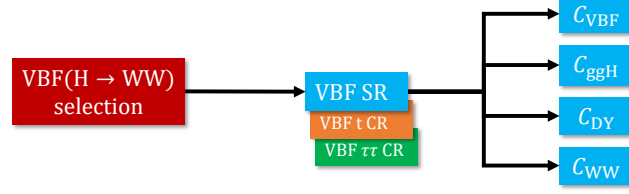


Figure 4.24: Visual overview of the categorization strategy in the VBF DF channel.

on the test sample. The analysis is then performed by evaluating the DNN on a statistically independent set of MC samples. The input to the DNN is a set of 26 kinematic variables, including 4-momenta of the leptons and jets, $\vec{p}_T^{\text{miss}}$, m_{jj} , $m_{\ell\ell}$, and $\Delta\eta_{jj}$. Events are divided into four classes C_j defined by the condition:

$$C_j = \left\{ x : \max_i (O_i(x)) = O_j \right\}. \quad (4.16)$$

The categorization scheme is visualized in Fig. 4.24. More details on the structure and tuning of the DNN can be found in Ref. [54]. The signal extraction fit is performed in each class C_i on the corresponding output distribution of the DNN O_i . The distributions of O_{VBF} and O_{ggH} are shown in Fig. 4.25 in the C_{VBF} and C_{ggH} classes, respectively, and the O_{VBF} distribution before categorizing events based on the DNN output is shown in Fig. 4.26.

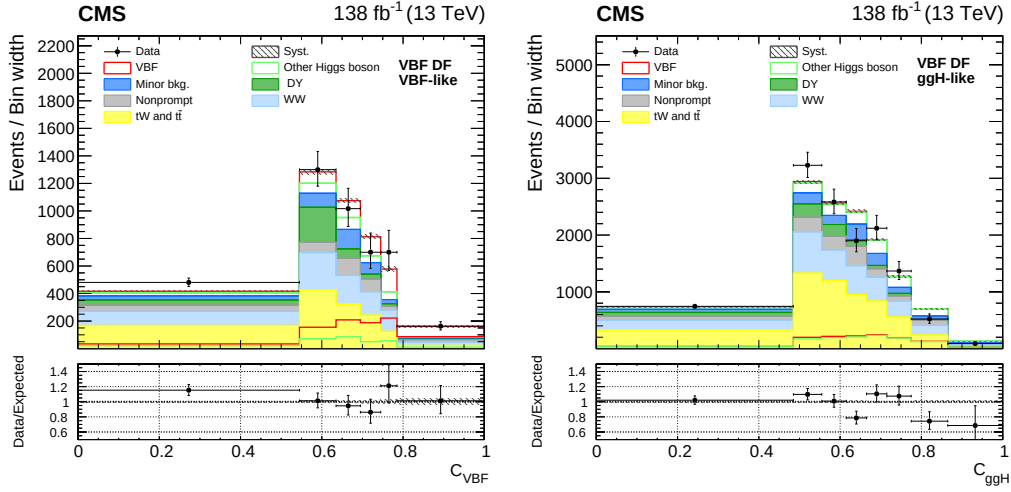


Figure 4.25: Distributions of the O_{VBF} (left) and O_{ggH} (right) output values in the C_{VBF} and C_{ggH} classes, respectively. The lower panel in each figure shows the ratio of the number of events observed in data to that of the total SM MC as extracted from the fit.

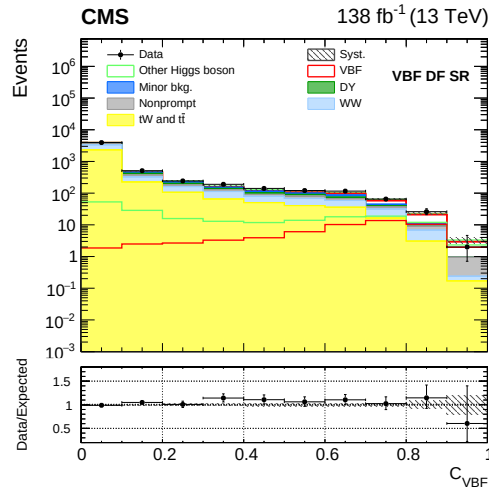


Figure 4.26: Distributions of the O_{VBF} output values after the VBF DF selection, before categorizing events based in the DNN output. The lower panel shows the ratio of the number of events observed in data to that of the total SM MC as extracted from the fit.

1750 4.3.4 Same flavor VBF category

1751 Most of the considerations given for the ggH SF category apply equally in the
 1752 VBF SF one. The selection is the same as described for the ggH SF category,
 1753 with the exception of requiring at least two jets and $m_{jj} > 120$ GeV. The same
 1754 data driven estimation for DY events is used, and control regions enriched
 1755 in t quark and WW events are defined. For the same reason outlined when
 1756 describing the ggH SF category, a counting experiment is performed for the
 signal extraction. The complete selection is given in Table 4.7.

Table 4.7: Summary of the selection used in the same flavor VBF categories. The DYMVA threshold is optimized separately in each subcategory and data set.

Subcategories	Selection
<u>Global selection</u>	
-	$p_T^{\ell_1} > 25$ GeV, $p_T^{\ell_2} > 10$ GeV (2016) or 13 GeV $p_T^{\text{miss}} > 20$ GeV, $p_T^{\ell\ell} > 30$ GeV ee or $\mu\mu$ pair with opposite charge $m_{\ell\ell} > 12$ GeV, $ m_{\ell\ell} - m_Z > 15$ GeV
<u>2-jet VBF category</u>	
ee, $\mu\mu$	$m_{\ell\ell} < 60$ GeV, $65 < m_T^H < 150$ GeV At least 2 jets with $p_T > 30$ GeV $ \Delta\phi_{\ell\ell} < 1.6$, $m_{jj} > 350$ GeV No b-tagged jets with $p_T > 20$ GeV DYMVA above threshold
WW CR	As SR but with $m_{\ell\ell} > 100$ GeV $m_T^H > 60$ GeV, $m_T(\ell_2, p_T^{\text{miss}}) > 30$ GeV
Top quark CR	As SR but with $m_{\ell\ell} > 100$ GeV, $m_T(\ell_2, p_T^{\text{miss}}) > 30$ GeV At least one of the leading jets b-tagged

4.3.5 VH2j categories

The VH2j categories target VH production with hadronic decay of the V boson and fully leptonic decay of the H boson. The aim is on events in the so-called resolved regime, with the V boson having sufficiently low p_T that the two jets from its decay are reconstructed separately (as opposed to being merged into a single jet). This allows to select VH events by requiring the presence of at least two jets, with the leading dijet pair having an invariant mass $65 < m_{jj} < 105$ GeV, therefore selecting W and Z resonances. The two jets are required to lie in the η region covered by the tracker to ensure a better reconstruction. Both DF and SF final states are considered, and the rest of the selection is in line with the ggH DF and SF ones, respectively. In the DF selection, given in Table 4.8, two CRs, enriched in t quark and $\tau\tau$ events respectively, are defined. The signal extraction fit is performed on the $m_{\ell\ell}$ distribution in the signal region, providing discrimination power also against the non-resonant WW background, and is shown in Fig. 4.27. In the SF selection, shown in Table 4.9, compared to the analogous ggH one, the threshold on $m_{\ell\ell}$ is moved to 70 GeV, m_T^H is required to be below 150 GeV, and the azimuthal angle separation between the two leptons $\Delta\phi_{\ell\ell}$ is required to be below 1.6. Two CRs are defined as for the 2-jet ggH SF selection, enriched in t quark and WW events, respectively. The signal extraction is done via a counting experiment.

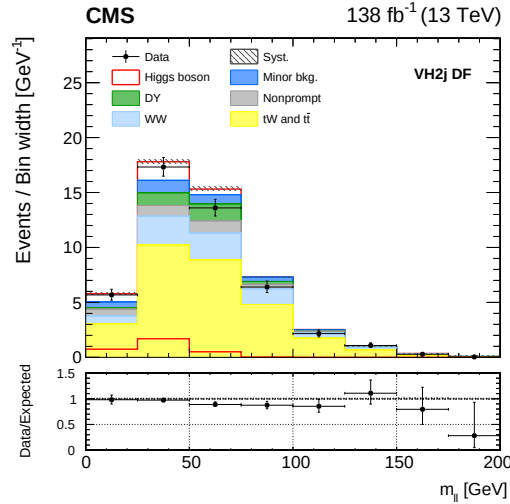


Figure 4.27: Distribution of the $m_{\ell\ell}$ fit variable in the VH2j DF SR. The lower panel in each figure shows the ratio of the number of events observed in data to that of the total SM MC as extracted from the fit.

Table 4.8: Summary of the selection applied to different-flavor VH2j categories.

Subcategory	Selection
<u>Global selection</u>	
-	$p_T^{\ell_1} > 25 \text{ GeV}$, $p_T^{\ell_2} > 10 \text{ GeV}$ (2016) or 13 GeV $p_T^{\text{miss}} > 20 \text{ GeV}$, $p_T^{\ell\ell} > 30 \text{ GeV}$, $m_{\ell\ell} > 12 \text{ GeV}$ $e\mu$ pair with opposite charge
<u>Signal region</u>	
-	At least 2 jets with $p_T > 30 \text{ GeV}$, $ \eta_{j1} , \eta_{j2} < 2.5$ $\Delta\eta_{jj} < 3.5$, $65 < m_{jj} < 105 \text{ GeV}$ $60 \text{ GeV} < m_T^H < 125 \text{ GeV}$, $\Delta R_{\ell\ell} < 2$ No b-tagged jet with $p_T > 20 \text{ GeV}$
<u>Control region</u>	
Top quark CR	As SR but with no m_T^H requirement, $m_{\ell\ell} > 50 \text{ GeV}$ At least 1 b-tagged jet with $p_T > 30 \text{ GeV}$
$\tau\tau$ CR	As signal region but with $m_T^H < 60 \text{ GeV}$ $40 < m_{\ell\ell} < 80 \text{ GeV}$

Table 4.9: Summary of the selection applied to same-flavor VH2j categories.

Subcategory	Selection
<u>Signal region</u>	
$ee, \mu\mu$	$m_{\ell\ell} < 70 \text{ GeV}$, $60 < m_T^H < 150 \text{ GeV}$ No b-tagged jets with $p_T > 20 \text{ GeV}$ $65 < m_{jj} < 105 \text{ GeV}$, $ \Delta\phi_{\ell\ell} < 1.6$ DYMVA above threshold
<u>Control region</u>	
WW CR	As SR but with $m_{\ell\ell} > 100 \text{ GeV}$ $m_T^H > 60 \text{ GeV}$, $m_T(\ell_2, p_T^{\text{miss}}) > 30 \text{ GeV}$
Top quark CR	As SR but with $m_{\ell\ell} > 50 \text{ GeV}$ $m_T(\ell_2, p_T^{\text{miss}}) > 30 \text{ GeV}$ At least one b-tagged jet with $p_T > 30 \text{ GeV}$

4.3.6 WHSS category

The WHSS category targets WH production with leptonic decay of the W boson and semileptonic decay of the H boson, resulting in a final state with two leptons and at least one jet. Because they come from different particles, the two leptons in the final state can have the same electric charge (SS). Since the number of background processes resulting in a SS final state is considerably lower than for oppositely charged leptons, only the former case is selected in the analysis. Moreover, it has been found that the ee final state adds negligible sensitivity to the measurement because of the higher background from non-prompt leptons, and these events are therefore not considered.

The leading (subleading) lepton is required to have $p_T > 25(20)$ GeV, p_T^{miss} is required to be at least 30 GeV, and the η separation between the two leptons ($\Delta\eta_{\ell\ell}$) is required to be below two. Events containing any b-tagged jets with $p_T > 20$ GeV are discarded. Events are subdivided based on the number of jets and flavor composition of the dilepton pair. For events with at least 2 jets, m_{jj} is required to be below 100 GeV.

To better isolate signal events, the $\tilde{m}_H$ variable is used as a proxy for m_H , which can not be fully reconstructed due to the neutrino from the semileptonic decay of the H boson. The $\tilde{m}_H$ variable is defined as the invariant mass of the four-momentum of the (di)jet system P_j and twice the four-momentum P_ℓ of the lepton closest to the (di)jet:

$$\tilde{m}_H = \sqrt{(P_j + 2P_\ell)^2}, \quad (4.17)$$

where factor 2 for P_ℓ serves as a proxy for four-momentum of the undetected neutrino. In events with more than one jet, P_j is taken as the four-momentum of the leading dijet pair. In events entering the 1-jet subcategory, if a second jet with $20 < p_T < 30$ GeV is present, it is added to the computation of P_j . All events in the WHSS category are required to have $\tilde{m}_H > 50$ GeV. The main sources of background in this category are WZ, W+jets with non-prompt leptons, $V\gamma$, and $V\gamma^*$ production. Minor backgrounds arise from t quark and other di- and triboson production processes. The WZ background is normalized in a dedicated CR common to the ZH3l category described later. The full event selection is given in Table 4.10. The signal extraction is performed by fitting the $\tilde{m}_H$ distribution in the SRs, as shown in Fig. 4.28. The WZ CRs, shared with the ZH3l selection, are shown in Fig. 4.29

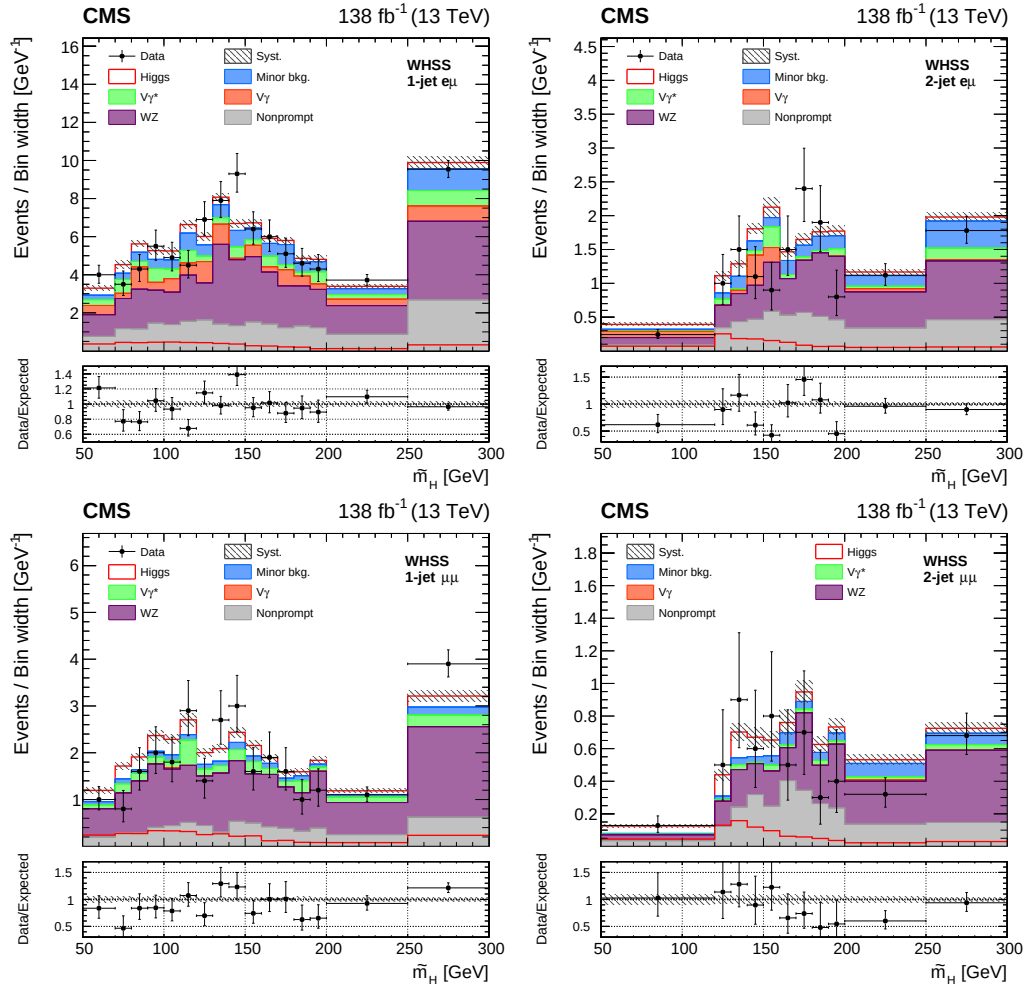


Figure 4.28: Distributions of the $\tilde{m}_H$ fit variable in the WHSS 1-jet $e\mu$ (upper left), 2-jet $e\mu$ (upper right), 1-jet $\mu\mu$ (lower left), and 2-jet $\mu\mu$ (lower right) SRs. The lower panel in each figure shows the ratio of the number of events observed in data to that of the total SM MC as extracted from the fit.

Table 4.10: Summary of the selection in the WHSS category.

Subcategories	Selection
<u>Global selection</u>	
-	$p_T^{\ell_1} > 25 \text{ GeV}, p_T^{\ell_2} > 20 \text{ GeV}$ $m_{\ell\ell} > 12 \text{ GeV}, \Delta\eta_{\ell\ell} < 2, p_T^{\text{miss}} > 30 \text{ GeV}$ $\tilde{m}_H > 50 \text{ GeV}$, no b-tagged jet with $p_T > 20 \text{ GeV}$
<u>Signal region</u>	
1-jet $e\mu(\mu\mu)$	One jet with $p_T > 30 \text{ GeV}$ $e\mu(\mu\mu)$ pair with same charge
2-jet $e\mu(\mu\mu)$	At least two jets with $p_T > 30 \text{ GeV}$, $m_{jj} < 100 \text{ GeV}$ $e\mu(\mu\mu)$ pair with same charge
<u>Control region</u>	
WZ	Shared with ZH3 ℓ

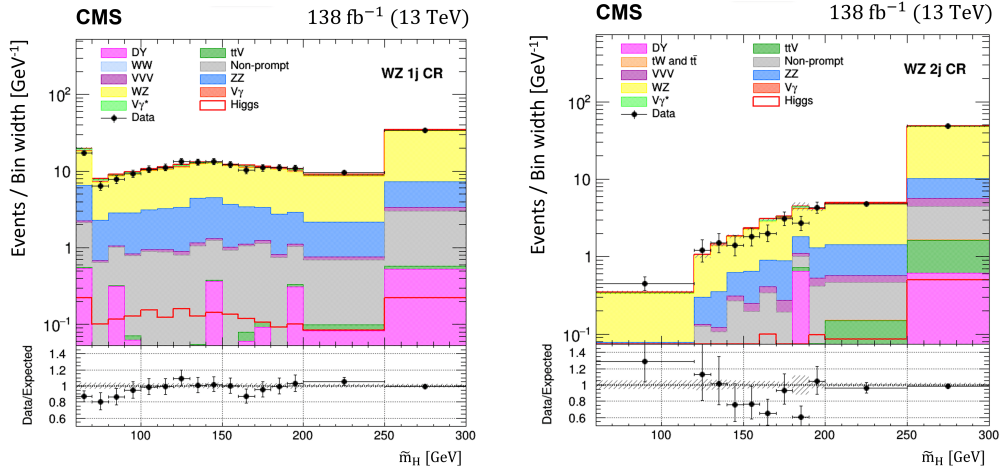


Figure 4.29: Distributions of the $\tilde{m}_H$ fit variable in the 1-jet (left) and 2-jet (right) WZ CR. The lower panel in each figure shows the ratio of the number of events observed in data to that of the total expected SM MC.

1813 4.3.7 WH3l category

1814 This category targets WH production with fully leptonic decays of both the
 1815 W and H bosons, resulting in a final state with three charged leptons. Events
 1816 are required to contain three leptons with p_T above 25, 20, and 15 GeV,
 1817 respectively, and p_T^{miss} of at least 30 GeV. The total charge $Q_{3\ell}$ of the trilepton
 1818 system is required to equal ± 1 . To remove contributions from low mass
 1819 resonances, events are discarded if any of the dilepton pairs has invariant mass
 1820 $m_{\ell\ell}$ above 12 GeV. Events are required to contain no jet with $p_T > 30$ GeV
 1821 and no b-tagged jet with $p_T > 20$ GeV.

1822 Events are subdivided based on the flavor and charge composition of
 1823 the lepton system. If at least a pair of leptons with opposite charge and
 1824 same flavor (OSSF) is found, the event is placed in the corresponding OSSF
 1825 subcategory. Since this subcategory is subject to contamination from DY
 1826 events, a requirement on the invariant mass of all dilepton pairs to satisfy
 1827 the condition $|m_{\ell\ell} - m_Z| > 20$ GeV is imposed. The p_T^{miss} threshold is also
 1828 increased to 40 GeV to further suppress events with instrumental missing
 1829 momentum. Events with no OSSF pair are placed in the same sign, same
 1830 flavor (SSSF) subcategory, with no additional requirements.

1831 The main backgrounds in this category are ZW, ZZ, $V\gamma$, and $V\gamma^*$ pro-
 1832 duction, as well as events with non-prompt leptons. Two CRs are defined,
 1833 enriched in WZ and $Z\gamma$ events, respectively. The WZ CR matches the SR,
 1834 but the requirement on $|m_{\ell\ell} - m_Z|$ is inverted, and the invariant mass of the
 1835 trilepton system $m_{3\ell}$ is required to be above 100 GeV. The $V\gamma$ CR is de-
 1836 fined similarly to the WZ one, but p_T^{miss} is required to be below 40 GeV and
 1837 $m_{3\ell}$ is required to be between 80 and 100 GeV. The full selection is given in
 1838 Table 4.11. To better separate signal from background sources two boosted
 1839 decision trees (BDTs) have been trained in the OSSF and SSSF regions, re-
 1840 spectively. To increase the size of the training sample, the requirement on
 1841 $|m_{\ell\ell} - m_Z|$ in the OSSF region is dropped during this step. Inputs to the BDTs
 1842 include kinematics of the leptons and jets, b-tagging scores on the leading
 1843 jets, and $\vec{p}_T^{\text{miss}}$. The signal extraction is performed by fitting the output of
 1844 the BDTs in the corresponding SRs. The output distributions are shown in
 1845 the SRs in Fig. 4.30, and in CRs in Fig. 4.31.

1846 4.3.8 ZH3l category

1847 This category targets ZH production with leptonic and semileptonic decays
 1848 of the Z and H bosons, respectively. The final state therefore contains three
 1849 charged leptons with $Q_{3\ell} = \pm 1$. Three leptons are required with p_T greater
 1850 than 25, 20, and 15 GeV, respectively. No requirement on p_T^{miss} is imposed,

Table 4.11: Summary of the selection in the WH3 ℓ category.

Subcategories	Selection
<u>Global selection</u>	
-	$p_T^{\ell_1} > 25 \text{ GeV}, p_T^{\ell_2} > 20 \text{ GeV}, p_T^{\ell_3} > 15 \text{ GeV}$ $Q_{3\ell} = \pm 1, \min(m_{\ell\ell}) > 12 \text{ GeV}, \Delta\eta_{\ell\ell} > 2.0$ $p_T^{\text{miss}} > 30 \text{ GeV}$, no jets with $p_T > 30 \text{ GeV}$ No b-tagged jets with $p_T > 20 \text{ GeV}$
<u>Signal region</u>	
OSSF	OSSF lepton pair, $ m_{\ell\ell} - m_Z > 20 \text{ GeV}, p_T^{\text{miss}} > 40 \text{ GeV}$
SSSF	No OSSF lepton pair
<u>Control region</u>	
WZ	OSSF lepton pair, $ m_{\ell\ell} - m_Z < 20 \text{ GeV}$ $p_T^{\text{miss}} > 45 \text{ GeV}, m_{3\ell} > 100 \text{ GeV}$
Z γ	OSSF lepton pair, $ m_{\ell\ell} - m_Z < 20 \text{ GeV}$ $p_T^{\text{miss}} < 40 \text{ GeV}, 80 < m_{3\ell} < 100 \text{ GeV}$

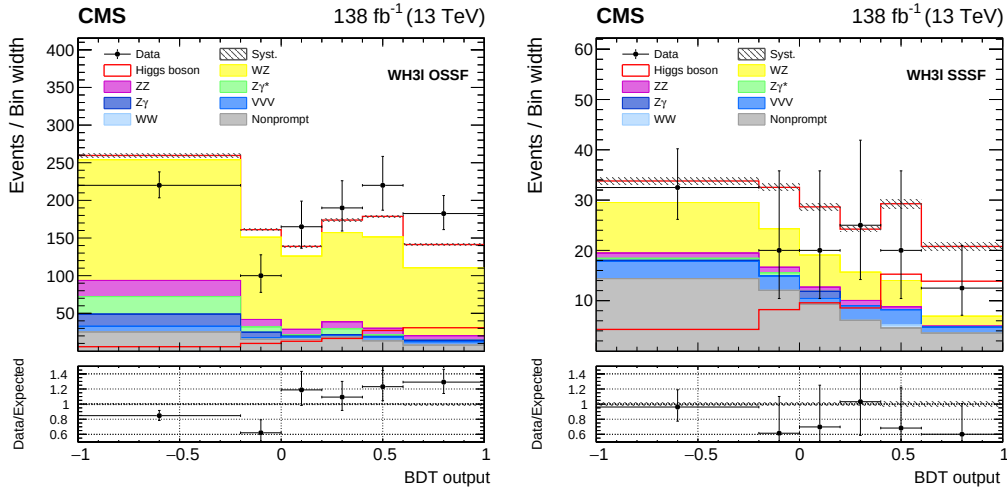


Figure 4.30: Distributions of the BDT score in the WH3 ℓ OSSF (left) and SSSF (right) SRs. The lower panel in each figure shows the ratio of the number of events observed in data to that of the total SM MC as extracted from the fit.

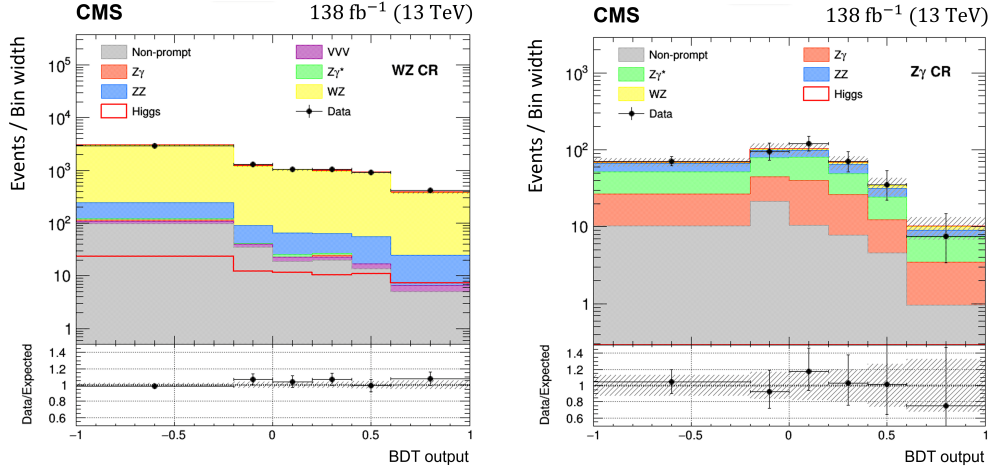


Figure 4.31: Distributions of the BDT score in the WZ (left) and $Z\gamma$ (right) CRs. The lower panel in each figure shows the ratio of the number of events observed in data to that of the total expected SM MC.

1851 since the presence of only one neutrino from the semileptonic H decays re-
 1852 sults in the signal having, on average, less missing momentum than some
 1853 of the backgrounds. Low mass resonances are suppressed by requiring $m_{\ell\ell}$
 1854 of all lepton pairs to be above 12 GeV. In order to select events with a Z
 1855 boson, at least one OSSF lepton pair with $|m_{\ell\ell} - m_Z| < 25$ GeV is required.
 1856 Contamination from $Z\gamma$ events, with the photon converting to a lepton in the
 1857 detector, are suppressed by requiring $|m_{3\ell} - m_Z| > 20$ GeV. Events are subdi-
 1858 vided based on jet multiplicity. Finally, events must have an azimuthal angle
 1859 separation $\Delta\phi < \pi/2$. The $\Delta\phi$ variable is defined as the angular separation
 1860 between the $\ell + \vec{p}_T^{\text{miss}}$ and (di)jet systems, corresponding to the leptonically
 1861 and hadronically decaying W bosons from the H decay, respectively, where ℓ
 1862 indicates the remaining lepton once the OSSF pair with $m_{\ell\ell}$ closest to m_Z is
 1863 removed.

1864 The main backgrounds in this category arise from WZ production, ZZ
 1865 production with one of the leptons emitted outside of the detector's accep-
 1866 tance, and Z+jets events with non-prompt leptons. A WZ CR is defined by
 1867 inverting the requirement on $\Delta\phi$, and is used to also constrain the WZ back-
 1868 ground in the WHSS category. The full selection is given in Table 4.12. The
 1869 signal extraction is performed by fitting the $m_T^H = m_T(\ell, p_T^{\text{miss}}, j(j))$ distribu-
 1870 tion in the SRs, shown in Fig. 4.32. The WZ CRs, shared with the WHSS
 1871 selection, are shown in Fig. 4.33

Table 4.12: Summary of the selection in the $ZH3\ell$ category.

Subcategories	Selection
<u>Global selection</u>	
-	$p_T^{\ell_1} > 25 \text{ GeV}, p_T^{\ell_2} > 20 \text{ GeV}, p_T^{\ell_3} > 15 \text{ GeV}$ $Q_{3\ell} = \pm 1, \min(m_{\ell\ell}) > 12 \text{ GeV}$ $ m_{\ell\ell} - m_Z < 25 \text{ GeV}, m_{3\ell} - m_Z > 20 \text{ GeV}$ No b-tagged jet with $p_T > 20 \text{ GeV}$
<u>Signal region</u>	
1-jet	$= 1 \text{ jet with } p_T > 30 \text{ GeV}, \Delta\tilde{\phi} < \pi/2$
2-jet	$\geq 2 \text{ jets with } p_T > 30 \text{ GeV}, \Delta\tilde{\phi} < \pi/2$
<u>Control region</u>	
1-jet WZ	$= 1 \text{ jet with } p_T > 30 \text{ GeV}, \Delta\tilde{\phi} > \pi/2$
2-jet WZ	$\geq 2 \text{ jets with } p_T > 30 \text{ GeV}, \Delta\tilde{\phi} > \pi/2$

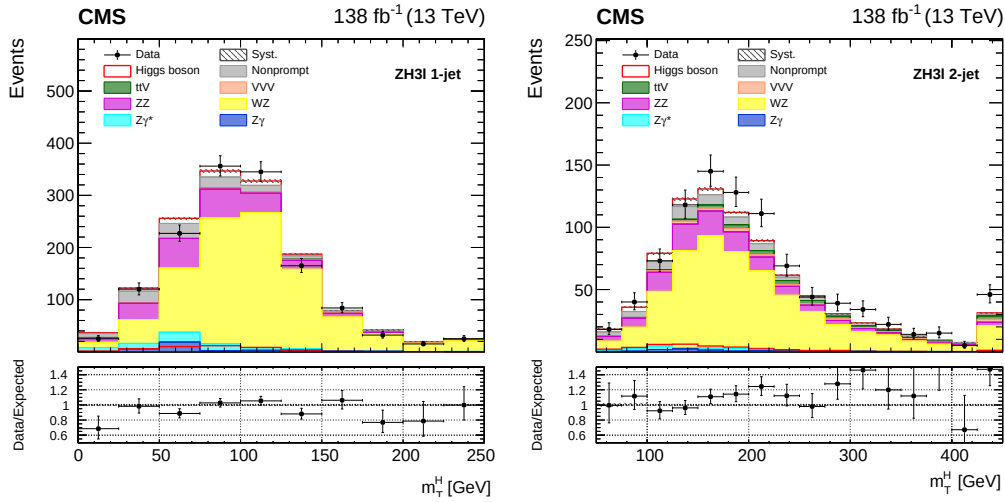


Figure 4.32: Distributions of the m_T^H fit variable in the $ZH3\ell$ 1-jet (left) and 2-jet (right) SRs. The lower panel in each figure shows the ratio of the number of events observed in data to that of the total SM MC as extracted from the fit.

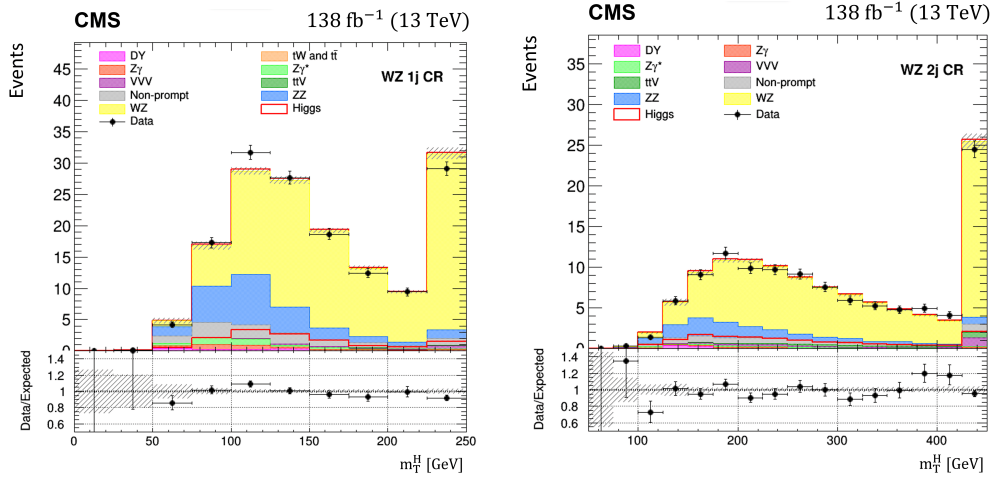


Figure 4.33: Distributions of the m_T^H fit variable in the 1-jet (left) and 2-jet (right) WZ CR. The lower panel in each figure shows the ratio of the number of events observed in data to that of the total expected SM MC.

4.3.9 ZH4l category

This category targets ZH production with both the Z and H bosons decaying leptonically. Events are required to contain four charged leptons with p_T greater than 25, 15, 10, and 10 GeV, respectively, and total electric charge $Q_{4\ell} = 0$. To reject low mass resonances, all dilepton pairs are required to have $m_{\ell\ell} > 12$ GeV. The OSSF lepton pair with $m_{\ell\ell}$ closest to m_Z is designated as the Z candidate, while the second pair is referred to as the X candidate (X being either the H boson or a second Z boson). The mass of the Z candidate is required to be within 15 GeV of m_Z . All events containing b-tagged jets with $p_T > 20$ GeV are rejected.

Events are subdivided based on the flavor composition of the leptons forming the X candidate, with events containing a DF (SF) X pair being assigned to the XDF (XSF) subcategory. Events in the XDF category are required to have $10 < m_X < 70$ GeV and $p_T^{\text{miss}} > 20$ GeV, while for the XSF region the requirements are placed at $10 < m_X < 60$ GeV and $p_T^{\text{miss}} > 35$ GeV. Additionally, events in the XSF are required to have the invariant mass of the full lepton system $m_{4\ell} > 140$ GeV.

By far the main background in this category arises from Z boson pair production, with minor contributions coming from $t\bar{t}Z$, VVV, and $V\gamma$ processes. A dedicated ZZ CR is defined by requiring $75 < m_X < 105$ GeV and $p_T^{\text{miss}} < 35$ GeV. The full event selection is given in Table 4.13. To better separate signal from background, a BTB is trained on events that pass the

Table 4.13: Summary of the selection in the $ZH4\ell$ category.

Subcategories	Selection
<u>Global selection</u>	
-	$p_T^{\ell_1} > 25 \text{ GeV}, p_T^{\ell_2} > 15 \text{ GeV}, p_T^{\ell_3} > 10 \text{ GeV}, p_T^{\ell_4} > 10 \text{ GeV}$ $Q_{4\ell} = 0, \min(m_{\ell\ell}) > 12 \text{ GeV}, m_{\ell\ell} - m_Z < 15 \text{ GeV}$ No b-tagged jet with $p_T > 20 \text{ GeV}$
<u>Signal region</u>	
XSF	Same flavor X pair, $m_{4\ell} > 140 \text{ GeV}$ $10 < m_X < 60 \text{ GeV}, p_T^{\text{miss}} > 35 \text{ GeV}$
XDF	Different flavor X pair, $10 < m_X < 70 \text{ GeV}$ $p_T^{\text{miss}} > 20 \text{ GeV}$
<u>Control region</u>	
ZZ	$75 < m_X < 105 \text{ GeV}, p_T^{\text{miss}} < 35 \text{ GeV}$

1894 global selection, in addition to having $p_T^{\text{miss}} > 20 \text{ GeV}$ and $10 < m_X < 70 \text{ GeV}$.
 1895 The inputs to the BDT include the kinematics of the four leptons and $\vec{p}_T^{\text{miss}}$.
 1896 The signal extraction is performed by fitting the output distribution of the
 1897 BDT in the SRs, shown in Fig. 4.34. The output of the BTD in the ZZ CR
 1898 is shown in Fig. 4.35.

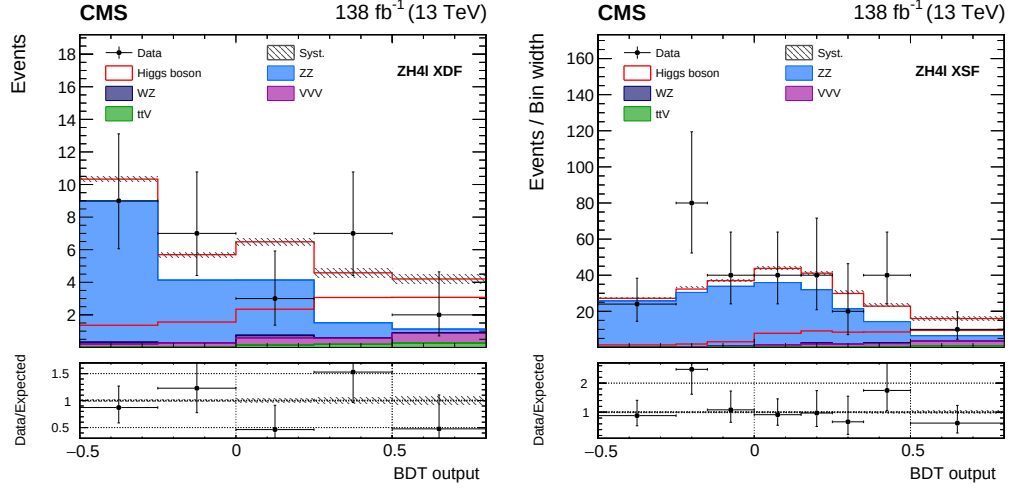


Figure 4.34: Observed distributions of the BDT score in the ZH4 ℓ XDF (left) and XSF (right) SRs. The lower panel in each figure shows the ratio of the number of events observed in data to that of the total SM MC as extracted from the fit.

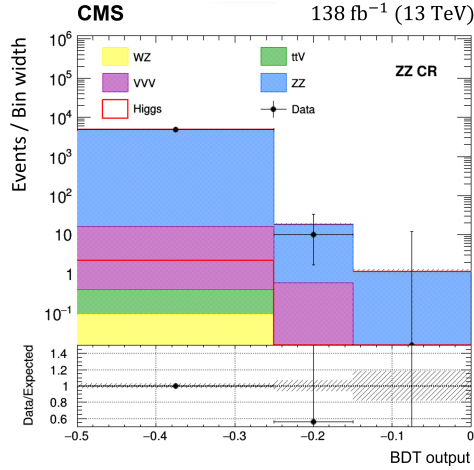


Figure 4.35: Observed distributions of the BDT score in the ZZ CR. The lower panel shows the ratio of the number of events observed in data to that of the total expected SM MC.

4.4 The simplified template cross section measurement

What has been described up to now is the setup that is used to extract results in terms of total production cross sections (i.e., extrapolated to the full particle-level phase space), as well as in terms of couplings of the H boson to vector bosons and fermions, as will be detailed in Chapter 6. A third set of results however is provided in the form of simplified template cross sections (STXS) [20], and the setup for this particular measurement is significantly different, warranting a dedicated discussion.

4.4.1 The STXS framework

Inclusive measurements, such as the overall cross section measurement performed in this work, have the advantage of allowing for a (relatively) straightforward combination of analyses across decay modes and experiments, yielding greatly increased sensitivity when compared to measurements in a single channel. This however also introduces a dependence on the SM signal prediction: when two analyses are combined, the signal yield is measured simultaneously in their respective selections. This implicitly assumes that the variation in signal yields in the two, often very different, regions of phase space selected by the analyses being combined, is the same. In other terms, the measurement done in a narrow slice of phase space (the one carved out by the analysis selection) is extrapolated to the full particle level phase space. While this assumption is sensible when measuring the parameters of a particular model, in this case the SM, it introduces severe limitations when trying to interpret the results in terms of other physics models.

The alternative approach is that of limiting the cross section measurement to a subset of particle level phase space, called fiducial volume, designed to overlap as closely as possible with the region of phase space selected in the analysis. While the fiducial approach greatly reduces model dependency, it also means that, since different analyses will have different selections (and thus, different fiducial volumes), combinations are not as well defined, limiting the sensitivity.

A second dimension in the spectrum of cross section analyses is the one that goes from fully inclusive cross section measurements to fully differential ones, like the one performed by CMS in the $H \rightarrow WW$ channel [55]. Differential cross section measurements provide more stringent constraints on the theory (since they are also sensitive to changes in the shape of the observable distributions, rather than their normalization only), but require larger data

1936 samples (given the split in different bins). Differential measurements are typ-
 1937 ically done in the fiducial framework, in order to allow an easier comparison
 1938 with theory; this however once again compromises the ability to combine
 1939 analyses in different decay modes, and in this case also gives rise to compli-
 1940 cations when correcting the observed distribution for distortions introduced
 1941 by the finite resolution of the detector (a procedure known as unfolding).

1942 The STXS framework provides a setup for what is essentially a non-
 1943 fiducial¹ differential cross section measurement:

- 1944 • it is inclusive in decay modes, allowing for combination of analyses in
 1945 different decay channels;
- 1946 • what is measured is a set of cross sections in different regions of particle-
 1947 level phase space, called bins for simplicity. Reconstructed event yields
 1948 are extrapolated back to particle level and translated to cross sections
 1949 (effectively performing an unfolding operation), allowing to directly
 1950 compare measured results with theoretical predictions;
- 1951 • STXS bins are designed to enhance the sensitivity to models beyond
 1952 the SM, which are expected to mainly affect e.g., regions with high p_T
 1953 of the H boson.

1954 In a nutshell then, the STXS framework allows to perform a set of differential
 1955 cross section measurements made in such a way that their combination is well
 1956 defined, while maintaining sensitivity for possible BSM effects manifesting in
 1957 specific corners of phase space (e.g., for high p_T of the H boson). For this
 1958 analysis, the stage 1.2 definitions are used. Different bins are defined for each
 1959 production mode of the H boson. For ggH, VBF, and VH with hadronic
 1960 decay of the V boson, the separation is based on the p_T of the H boson, the
 1961 number of hadronic jets, and, for events with at least two jets, m_{jj} , as shown
 1962 in Fig. 4.36 and 4.37. For VH production with leptonic decay of the V boson,
 1963 the split is instead done based on p_T of the V boson, as well as the number
 1964 of jets, as shown in Figure 4.38. Since the $H \rightarrow WW$ analysis by itself lacks
 1965 the constraining power to measure the full set of STXS bins, some bins are
 1966 merged i.e., they are measured together as a single bin. These are indicated
 1967 with solid colors in Fig. 4.36, 4.37, and 4.38.

1968 Ideally, to perform the STXS measurement it would be sufficient to cat-
 1969 egorize reconstructed events based on the same STXS bin definitions and
 1970 perform the analysis; because the resolution of the detector is finite however,
 1971 the reconstructed kinematic features of particles do not perfectly match those

¹Only a very loose requirement on the rapidity of the H boson $|y_H| < 2.4$ is applied at particle level.

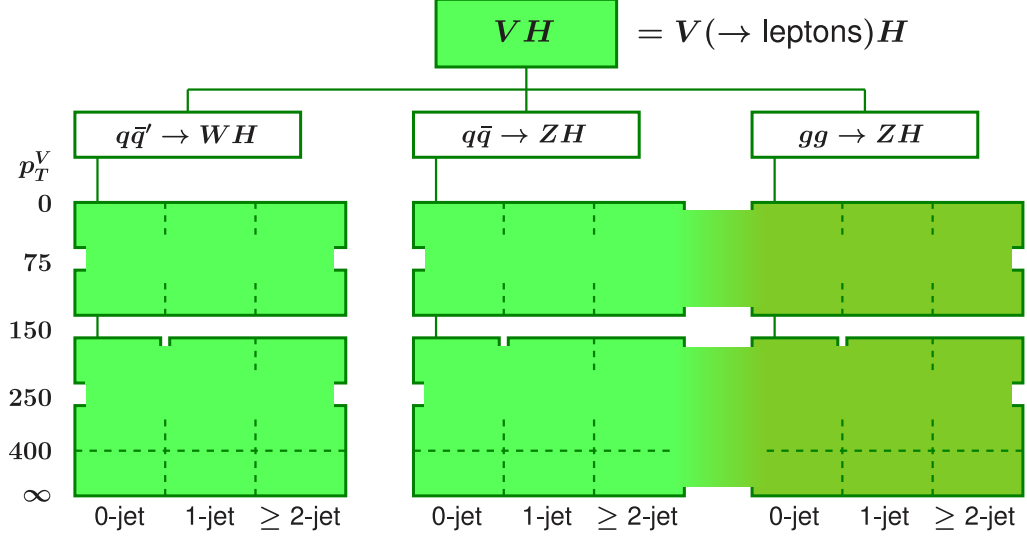


Figure 4.38: STXS bin definitions for the VH production mode with leptonic decay of the V boson. Bins measured together in this analysis are fused together with solid color.

before interacting with the detector (particle level). This gives rise to a migration effect: events belonging to a given STXS bin at particle level can be reconstructed in non-corresponding analysis categories. To account for this migration effect i.e., unfold the measurement to particle level, each signal MC sample is subdivided into sub-samples, each corresponding to a particular STXS bin at particle level. After simulating the interaction with the detector and the reconstruction, each of these sub-samples will contribute to more than one analysis category, thus quantifying the migration effect due to the detector smearing of kinematic observables. The unfolding is then done implicitly in the fit by treating each STXS sub-sample as a separate signal contribution.

4.4.2 The STXS analysis

In order to effectively perform the STXS measurement, the analysis described in Section 4.3 was modified. While the overall strategy is analogous to the inclusive measurement, all categories are adapted to match the target STXS bins. This additional sub-categorization is shown in Fig. 4.39, 4.40, and 4.41 for the ggH , VBF , and VH categories, respectively. The same CR setup as the inclusive measurement is maintained, with each CR being likewise divided to match each STXS bin. An additional complication arises in VH categories with a leptonic decay of the V boson: STXS bins are defined based on the

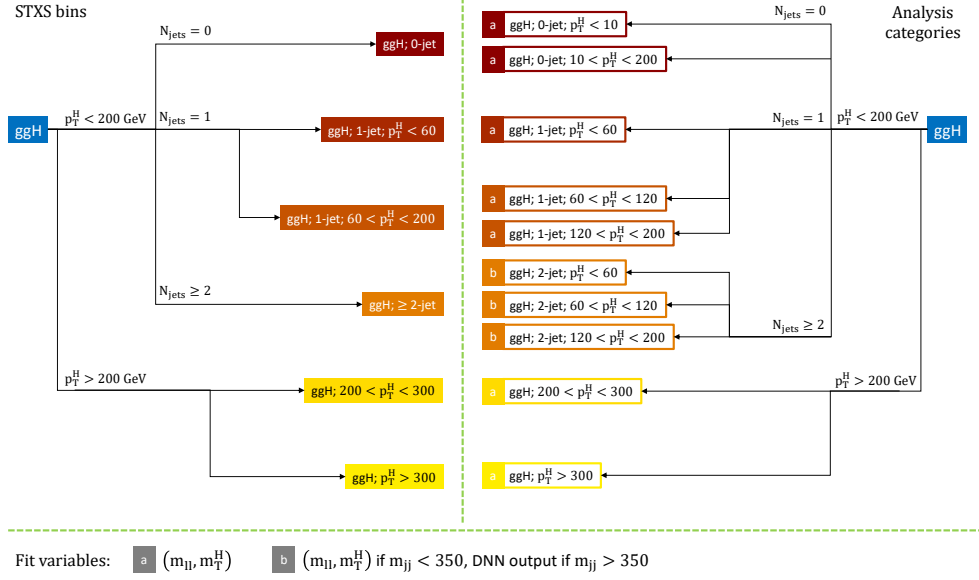


Figure 4.39: Schematic view of the categorization to target STXS bins for the ggH production mode.

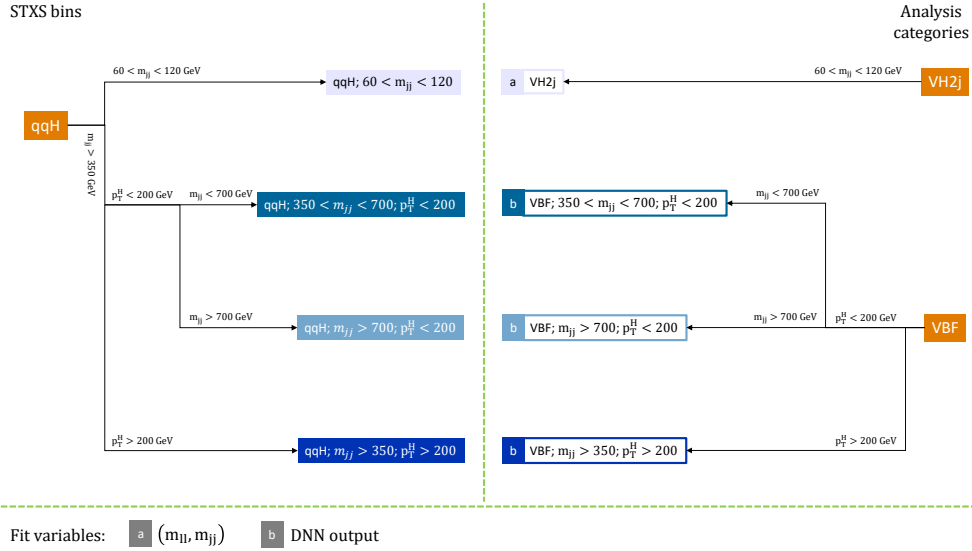


Figure 4.40: Schematic view of the categorization to target STXS bins for the VBF and VH (with hadronically decaying V boson) production modes.

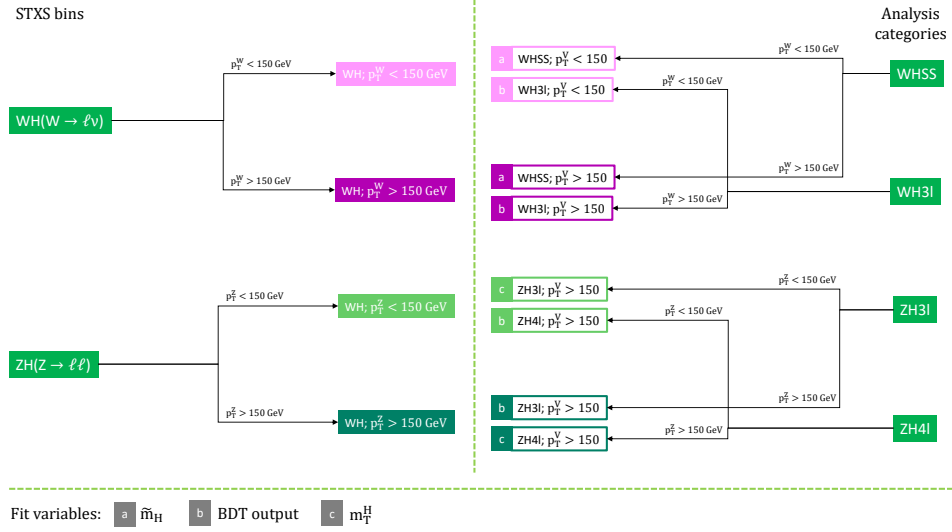


Figure 4.41: Schematic view of the categorization to target STXS bins for the VH production mode with leptonic decay of the V boson.

1992 p_T of the associated V boson, which in the case of the $H \rightarrow WW$ analysis
 1993 is not always possible to reconstruct given the additional neutrinos coming
 1994 from the H boson decay. The reconstructed V boson p_T is therefore defined
 1995 differently depending on the final state:

- 1996 • in the WHSS category, one of the neutrinos comes from the decay
 1997 of the H boson and the other from the decay of the associated W
 1998 boson. The lepton with the largest azimuthal separation from the dijet
 1999 system is identified as coming from the W boson (ℓ_W). The transverse
 2000 momentum of the neutrino arising from the H boson (ν_H) can then be
 2001 evaluated from the momenta of the other decay products (one lepton
 2002 (ℓ_H) and the dijet system) by imposing the constrain on the mass of
 2003 the H boson:

$$\vec{p}_T^{\nu_H} = \vec{p}_T^{\ell_H} \left(\frac{125 \text{ GeV}}{|\vec{p}^{\ell_H} + \vec{p}^{jj}|} - 1 \right). \quad (4.18)$$

2004 The transverse momentum of the neutrino associated with the W boson
 2005 can then be written as:

$$\vec{p}_T^{\nu_W} = \vec{p}_T^{\text{miss}} - \vec{p}_T^{\nu_H}, \quad (4.19)$$

2006 and finally added to the $\vec{p}_T$ of the lepton coming from the W boson, to
 2007 obtain the transverse momentum of the latter.

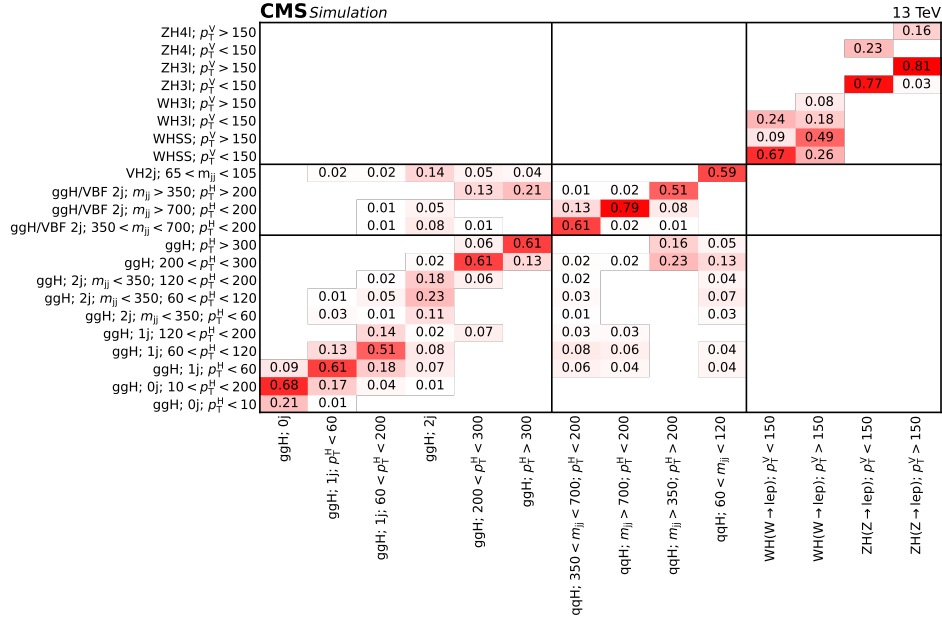


Figure 4.42: Migration matrix showing the expected signal composition in each analysis category of the STXS fit. Particle-level STXS bins are shown on the horizontal axis, and corresponding analysis categories on the vertical axis. All dimensional quantities are measured in GeV.

- In the WH3l category, given the three neutrinos in the final state, the constraint of the H boson mass is not sufficient to calculate $\vec{p}_T^W$ as above. The two leptons with smallest ΔR separation are identified as arising from the H boson, while the third lepton is identified as ℓ_W . The p_T of the W boson $\vec{p}_T^W$ is then approximated with $\vec{p}_T^{\ell_W}$.
- In the ZH3l and ZH4l categories, the $\vec{p}_T$ of the associated Z boson ($\vec{p}_T^Z$) is taken to be the transverse momentum of the OSSF lepton pair with $m_{\ell\ell}$ closest to m_Z .

Fig. 4.42 shows the migration effect between STXS bins and corresponding analysis categories, while Fig. 4.43 shows the expected composition of the signal in each category in terms of STXS bins.

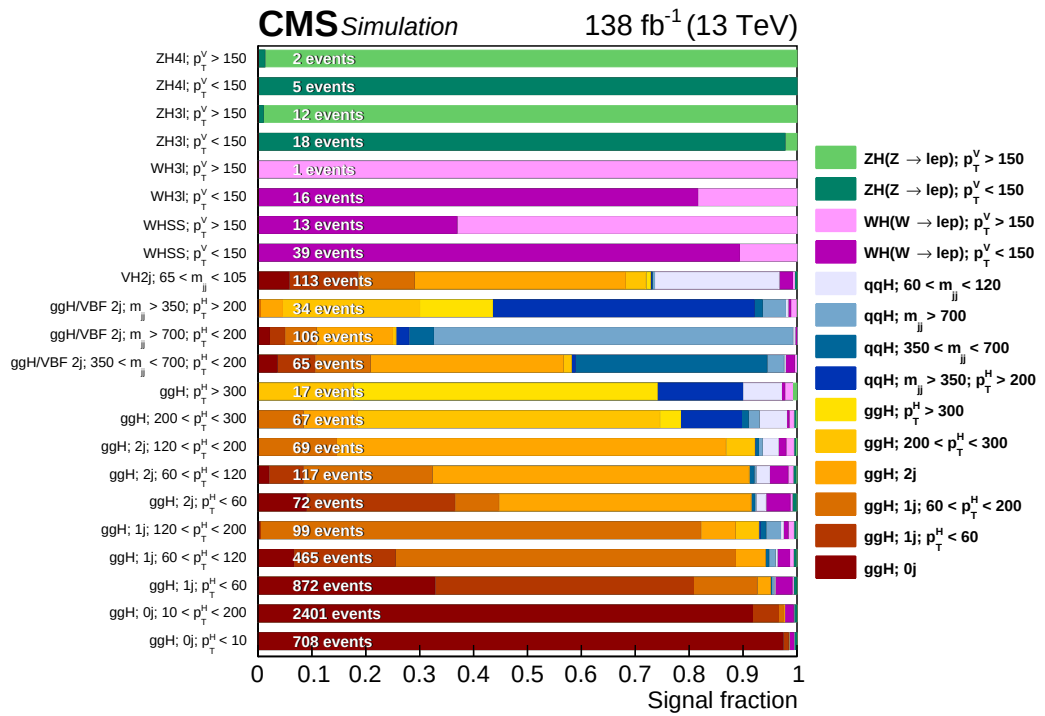


Figure 4.43: Expected fractions of particle-level STXS bins in each analysis category. The total number of expected $H \rightarrow WW$ events in each category is shown as well. All dimensional quantities are measured in GeV.

Chapter 5

Signal extraction and systematic uncertainties

5.1 Signal extraction fit

The signal is extracted by performing a binned maximum likelihood fit to data. Likelihood maximization is the standard procedure to estimate a set of unknown parameters of a distribution by sampling observations from it. In a nutshell, it proceeds as follows.

5.1.1 Maximum likelihood parameter estimation

Let $f(\mathbf{x}, \boldsymbol{\theta})$ be the probability distribution function (pdf) of the data $\mathbf{x}$, dependent on a set of unknown parameters $\boldsymbol{\theta}$. The problem is that of estimating $\boldsymbol{\theta}$ given an observation, or sampling, of the data $\mathbf{x}_{obs}$. This is done by evaluating the likelihood of $\boldsymbol{\theta}$ given the observation $\mathbf{x}_{obs}$:

$$L(\boldsymbol{\theta}) = f(\mathbf{x}_{obs}, \boldsymbol{\theta}), \quad (5.1)$$

where, crucially, the set of parameters $\boldsymbol{\theta}$ is now taken as the variable. Maximizing $L(\boldsymbol{\theta})$ gives the best estimate of $\boldsymbol{\theta}$ given the observation $\mathbf{x}_{obs}$:

$$\hat{\boldsymbol{\theta}} = \arg \max_{\boldsymbol{\theta}} (L(\boldsymbol{\theta})). \quad (5.2)$$

The estimator $\hat{\boldsymbol{\theta}}$ built in this fashion is called a maximum likelihood (ML) estimator. Generalization to multiple independent samples $\{\mathbf{x}_i^{obs}\}$ of $\mathbf{x}$ can be done trivially by noticing that the joint likelihood is simply the product of the likelihoods of each individual sample:

$$L(\boldsymbol{\theta}) = \prod_i L_i(\boldsymbol{\theta}). \quad (5.3)$$

In practice it is often useful to consider the negative logarithm of the likelihood (NLL) $-\log L(\boldsymbol{\theta})$. Monotonicity of the logarithm ensures that:

$$\arg \max_{\boldsymbol{\theta}} (L(\boldsymbol{\theta})) = \arg \min_{\boldsymbol{\theta}} (-\log L(\boldsymbol{\theta})). \quad (5.4)$$

Minimizing the NLL instead of maximising the likelihood achieves two effects. First, minimizing a function is computationally more convenient than maximizing one. Second, in the case of multiple observations given in Eq. 5.3, the logarithm turns the product into a sum, which is easier to handle. It can be shown [56] that the ML estimator satisfies a number of requirements for a reliable estimator:

- it is asymptotically unbiased, meaning that in the limit of infinite observations of $\boldsymbol{x}$ its expectation value $E[\hat{\boldsymbol{\theta}}]$ converges to the true values of the parameters:

$$\lim_{N_{obs} \rightarrow \infty} (E[\hat{\boldsymbol{\theta}}] - \boldsymbol{\theta}) = 0. \quad (5.5)$$

- it has the lowest possible variance in the limit of large number of observations i.e., it is always efficient in the large sample limit.
- it provides a natural way to incorporate sources of systematic uncertainty in the estimation of the parameters of interest (POI), as will be detailed in the following.

Naturally, a key aspect of the procedure lies in the choice of the distribution f which is believed to describe the data. If the analytical form of the pdf is known, its value for a given combination of $\boldsymbol{x}_{obs}$ and $\boldsymbol{\theta}$ can be calculated and used to evaluate the likelihood. This is commonly referred to as an unbinned ML fit. This is the case for e.g., the invariant mass spectrum of the decay products of a resonance, which has a distribution whose analytical form is known. In the case of the $H \rightarrow WW$ channel however, the two undetected neutrinos in the final state make the reconstruction of the invariant mass of the Higgs boson not possible experimentally. For this reason, other observables have to be used (as has been described in Section 4.3). An analytical form of the distributions of these observables is however not known; instead, a binned ML fit is done. The key distinction is that, as the name suggests, the observable of interest is binned i.e., a histogram of observations is built. This is done twice: once with the data sample, and once with a sample from MC simulation. This means that the parameters $\boldsymbol{\theta}$ and the observations $\boldsymbol{x}_i^{obs}$ are recast as the numbers of expected and observed entries in each histogram bin, labelled ν_i and n_i , respectively. The histogram effectively serves as an approximation of the unknown functional form of the pdf. The reason this

approach works, is that even if the underlying pdf is not know, the number of entries n_i in each bin of the histogram is known to follow a Poisson distribution $\mathcal{P}$, so that the likelihood can be explicitly written as:

$$L = \prod_{i=1}^{N_{\text{bins}}} \mathcal{P}(n_i, \nu_i). \quad (5.6)$$

Note that this formulation also accounts for the statistical uncertainty on the number of observed entries in each bin n_i . It is informative to rewrite the expected entries ν_i in terms of signal s_i and background b_i components:

$$L = \prod_{i=1}^{N_{\text{bins}}} \mathcal{P}(n_i, s_i + b_i). \quad (5.7)$$

The set of expected signal and background entries in each bin $\{s_i\}$ and $\{b_i\}$ are the signal and background histograms, or templates, in the observable of choice obtained from MC simulation (or other data driven methods, as specified in Chapter 4). Maximization of this likelihood function in terms of the normalization of the signal histogram is the essence of the so-called template fit. In the specific case of this analysis, the background contributions that are not measured from data (as detailed in the following) are set to their SM prediction (within the corresponding systematic uncertainties, as will also be detailed in the following). The same is true for the signal templates, apart from their normalization, which is left unconstrained by introducing the signal strength modifier μ :

$$L = \prod_{i=1}^{N_{\text{bins}}} \mathcal{P}(n_i, \mu \cdot s_i^{\text{SM}} + b_i^{\text{SM}}). \quad (5.8)$$

The signal strength parameter μ is the parameter with respect to which the likelihood is maximized i.e., the POI. Physically, it corresponds to the ratio of the number of signal events observed in data (i.e., extracted from the ML fit) to the number of signal events predicted by the SM. Both yields can be expresses as the product of the cross section σ of the process in question, the integrated luminosity $\mathcal{L}$ of the data set, and the branching ratio $\mathcal{B}$ to the final state of interest:

$$\mu = \frac{(\sigma \times \mathcal{B}) \times \mathcal{L}}{(\sigma \times \mathcal{B})_{\text{SM}} \times \mathcal{L}} = \frac{\sigma \times \mathcal{B}}{(\sigma \times \mathcal{B})_{\text{SM}}}, \quad (5.9)$$

meaning that μ can be interpreted in this context as the ratio between the observed and expected $\sigma \times \mathcal{B}$ of the signal process of interest.

2098 5.1.2 Treatment of systematic uncertainties

2099 Note that the construction up to now completely neglects any form of system-
 2100 atic uncertainty in the templates used to fit the data. These are accounted for
 2101 by adding so-called nuisance parameters to the likelihood function. These
 2102 control how the normalization, and possibly the shape, of the templates
 2103 changes when a quantity (e.g., the energy scale of muons) is varied within
 2104 its experimental uncertainty. This is achieved differently depending on the
 2105 effect of the systematic uncertainty in question:

- 2106 • if a source of uncertainty only changes the normalization of a given tem-
 2107 plate, its inclusion in the likelihood is done by adding a corresponding
 2108 multiplicative factor to all bins of the template. An example of such
 2109 a source of uncertainty is the one on the integrated luminosity: since
 2110 the yield of events in each bin of the template is proportional to the
 2111 integrated luminosity, changing its value directly scales the normaliza-
 2112 tion of the template. A corresponding nuisance parameter ν_{lumi} is thus
 2113 added to the likelihood:

$$L = \prod_{i=1}^{N_{\text{bins}}} \mathcal{P}(n_i, \mu \cdot s_i^{\text{SM}} \nu_{\text{lumi}} + b_i^{\text{SM}} \nu_{\text{lumi}}) \cdot \mathcal{N}(\nu_{\text{lumi}}). \quad (5.10)$$

2114 The additional term $\mathcal{N}$ encodes the uncertainty on ν_{lumi} : if the un-
 2115 certainty on the integrated luminosity is large, $\mathcal{N}$ will be a broad
 2116 distribution, allowing ν_{lumi} to vary over a wide range without signif-
 2117 icantly lowering the overall likelihood, ultimately translating in a loose
 2118 constrain on the normalization of the templates. This will naturally
 2119 increase the uncertainty on the measurement of μ , since the effect of
 2120 changing its value can be compensated for by adjusting ν_{lumi} , resulting
 2121 once again in a weaker constraint. Following the same logic, a small
 2122 uncertainty on the integrated luminosity will be codified in a narrow
 2123 $\mathcal{N}$, and thus in a tight constraint on the nuisance parameter and finally
 2124 a stronger constraint on μ . One crucial aspect to keep in mind is that,
 2125 while overestimating a source of systematic uncertainty will reduce the
 2126 final precision on μ , underestimating it can prevent the templates from
 2127 adjusting to describe the data, leading to an unstable fit procedure.
 2128 Nuisance parameters linked to normalization uncertainties are assigned
 2129 a log-normal constraint in the likelihood function:

$$\mathcal{N}(\nu) = \frac{1}{\sqrt{2\pi\nu \ln(\kappa)}} \exp\left(-\frac{(\ln \nu / \tilde{\nu})^2}{2(\ln \kappa)^2}\right), \quad (5.11)$$

2130 centered in $\nu = \tilde{\nu} = 1$, corresponding to the nominal normalization of
 2131 the template. The choice of a log-normal constraint on the nuisance

parameters is due to the fact that it vanishes when ν approaches 0. This is desirable when the quantity that is scaled by the nuisance parameter is physically bound to positive values, which is the case for event counts in histograms. The width of a log-normal distribution is controlled by the strictly positive parameter κ . The nuisance parameter ν will fall in the $[1/\kappa, \kappa]$ interval with a probability of 68%. In this sense, κ can be thought of as approximately encoding the relative uncertainty on the nuisance parameter:

$$\kappa = 1 + \frac{\sigma_\nu}{\bar{\nu}} \xrightarrow{\bar{\nu}=1} 1 + \sigma_\nu, \quad (5.12)$$

which if the uncertainty σ_ν is not too large, translates in a 68% probability for ν to be in the $[1 - \sigma_\nu, 1 + \sigma_\nu]$ interval, thus recovering the gaussian interpretation.

- The situation is somewhat more involved if the source of systematic uncertainty can also impact the shape of the templates, but the underlying idea is the same. An example of this second type of uncertainty is e.g., the lepton energy scale. In this case, the affected template is constructed three times: once with the nominal value of the energy scale, and twice more with the energy scale set to its nominal value plus or minus its uncertainty. An interpolation is then performed bin-by-bin, expressing the number of events in each bin N_i as a function of the nuisance parameter ν :

$$N_i(\nu) = N_i (1 + f_i(\nu)), \quad (5.13)$$

with the fractional change f_i defined as:

$$f_i(\nu) = \frac{1}{2} \left((\delta^+ - \delta^-) \nu + \frac{1}{8} (\delta^+ + \delta^-) (3\nu^6 - 10\nu^4 + 15\nu^2) \right). \quad (5.14)$$

The above expression is used for $|\nu| < 1$; for values of ν above (below) 1 (-1), a linear function with angular coefficient set to δ^+ ($-\delta^-$) is used instead. The parameters δ^+ and δ^- represent the difference in fractions f_i between the two variations and the nominal template:

$$\delta^\pm = f_i(\nu = \pm 1), \quad (5.15)$$

implicitly defining $\nu = \pm 1$ as codifying up and down variations of the parameter of interest by one time its corresponding uncertainty. This construction effectively uses three snapshots to build a template whose shape is allowed to shift continuously (and with continuous first and

second derivatives) as a function of the nuisance parameter ν . In this case, a gaussian constraint is used in the likelihood function:

$$\mathcal{N}(\nu) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\nu^2}{2}\right), \quad (5.16)$$

with standard deviation set to one, in accordance to the convention that the varied shapes correspond to a variation of the quantity in question by once its associated uncertainty.

The inclusion of nuisance parameters rapidly increases the number of parameters, and thus the dimensionality, of the likelihood. When calculating the uncertainty on the POI however, the dimensionality of the likelihood can be reduced by profiling it over the nuisance parameters. The profile likelihood function is defined as:

$$\ell(\mu) = \frac{L(\mu, \hat{\boldsymbol{\nu}})}{L(\hat{\mu}, \hat{\boldsymbol{\nu}})}, \quad (5.17)$$

where $\boldsymbol{\nu}$ is the array of nuisance parameters, $\hat{\boldsymbol{\nu}}$ is the array of best-fit values of the nuisance parameters, $\hat{\mu}$ is the best-fit value of the POI, and $\hat{\boldsymbol{\nu}}$ is the array of nuisance parameters that maximize L for a particular choice of μ . The 68% confidence level interval for μ can then be identified as the interval for which $-\ln \ell(\mu) < 0.5$.

5.1.3 Data-driven background normalization

As detailed in Section 4.3, the main backgrounds in the analysis are measured in dedicated control regions. The way this is done in practice is through so-called rate parameters in the likelihood function. These are essentially nuisance parameters in nature, but without a corresponding constraint term. These multiply a given background yield in both the SR and CR. Considering a simplified case of one background, this translates in a likelihood of the form:

$$L = \prod_{i=1}^{N_{\text{bins}}^{\text{SR}}} (n_i, s_i + \alpha \cdot b_i) \prod_{j=1}^{N_{\text{bins}}^{\text{CR}}} (n_j, \alpha \cdot b_j), \quad (5.18)$$

where α indicates the rate parameter. Since the CR is enriched in the background of interest, the corresponding likelihood term will determine the value of the rate parameter, which will then propagate the measured background normalization to the SR. In this analysis, only the number of events is used as a fit variable in the CRs i.e., $N_{\text{bins}}^{\text{CR}} = 1$.

5.1.4 Goodness of fit

The CMS collaboration enforces a strict blinding procedure on physics analyses in order to ensure as little bias as possible in the result. For this reason, this analysis as well has been optimized by relying entirely on MC simulation in the SR, and checking the agreement of the templates with data only in orthogonal CRs. Unblinding is done in steps, the first of which being evaluating the goodness of fit i.e., gauging how well the MC template describes the data after the fit is performed. This is done with a hypothesis test. In statistics, hypothesis tests are commonly formulated in terms of likelihood ratios, which are known to provide the best discriminating power through the Neyman-Pearson lemma. In a nutshell, given two hypotheses $H_0 : \boldsymbol{\theta} = \boldsymbol{\theta}_0$ and $H_1 : \boldsymbol{\theta} = \boldsymbol{\theta}_1$, a statistical test can be built by evaluating the likelihood under both hypotheses and taking their ratio:

$$\lambda = \frac{L(\boldsymbol{\theta}_1)}{L(\boldsymbol{\theta}_0)}, \quad (5.19)$$

and accepting or rejecting H_1 based on a threshold on λ . When evaluating goodness of fit however, we only have one hypothesis (the one whose ability to describe data we want to evaluate). The saturated model [57] provides an alternative hypothesis (playing the role of H_0 in Eq. 5.19) with a set of desirable properties. The saturated model is built by setting the expected values equal to the data points. In the case of a binned likelihood fit this is achieved by setting the expected event counts in each bin of each template equal to the observed number of events in that bin:

$$L_{\text{saturated}} = \prod_{i=1}^{N_{\text{bins}}} \mathcal{P}(n_i, n_i) \quad (5.20)$$

The saturated model thus has N_{bins} parameters. A particularly convenient feature of this construction is that $-2 \ln \lambda$ is distributed as a χ^2 with N_{bins} degrees of freedom. This can be used to evaluate a p-value for how well the set of parameters extracted by the fit describes the data:

$$p = \int_{\lambda}^{\infty} f_{\chi^2}, \quad (5.21)$$

where f_{χ^2} is the well known χ^2 probability distribution function. In practical terms, the presence of nuisance constraints in the likelihood function make defining the number of degrees of freedom (dof) of the χ^2 distribution challenging, and it is preferable to approximate f_{χ^2} via MC. This is done by generating a set of pseudo-data sets starting from the expected distribution

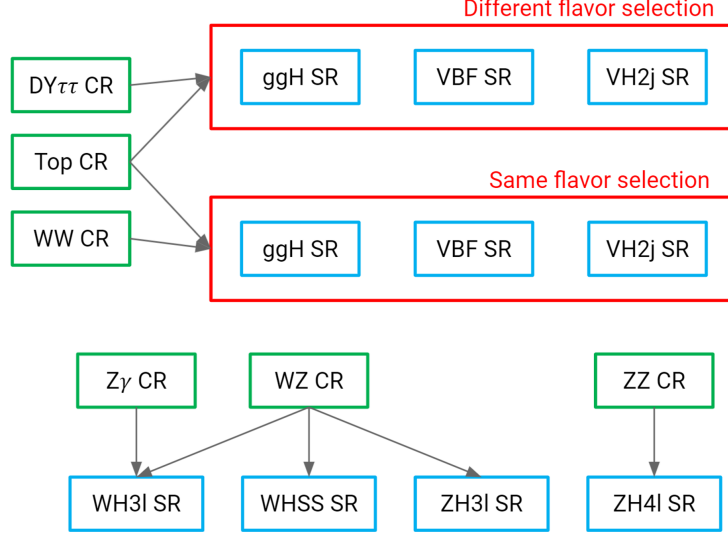


Figure 5.1: Summary of the CRs contributing to the background estimation in each analysis category.

of the data and evaluating λ on each. The approximate p-value can then be calculated as:

$$p = \frac{N_{\text{tries}}(\lambda > \lambda_{\text{data}})}{N_{\text{tries}}}. \quad (5.22)$$

5.2 Data-driven background estimation setup

As detailed in Section 4.3, the main backgrounds in each analysis category are estimated from data in dedicated CRs. This is achieved through rate parameters, as described in Section 5.1. An overview of the CRs that contribute to each category is given in Fig. 5.1. The setup is as follows:

- ggH DF category:** the t quark and $\tau\tau$ production processes are estimated in corresponding CRs. These are divided by jet multiplicity following the SR subdivision, and the yields in the 0- 1- and 2-jet regions are scaled independently. The non-resonant production of WW pairs is normalized directly in the SRs by exploiting the difference in the $(m_{\ell\ell}, m_T^H)$ distribution between this background and the signal, and also scales independently in different jet multiplicity bins.
- ggH SF category:** non-resonant WW pair and t quark production are normalized in dedicated CRs, independently in each jet multiplicity bin.

- 2234 The contribution from DY production of dilepton pairs is estimated via
 2235 the data-driven approach described in Section 4.3, but no CR is defined
 2236 at the time of fitting.
- 2237 • **VBF DF category:** the t quark and $\tau\tau$ production processes are
 2238 estimated in corresponding CRs, as well as by leveraging the dedicated
 2239 DNN classes. The non-resonant production of WW pairs is normalized
 2240 in the SR by exploiting the dedicated DNN class.
 - 2241 • **VBF SF category:** the setup is the same as in the ggH SF category,
 2242 although only one jet multiplicity bin is used.
 - 2243 • **VH2j DF category:** the setup is the same as in the ggH DF category,
 2244 although only one jet multiplicity bin is used.
 - 2245 • **VH2j SF category:** the setup is the same as in the ggH SF category,
 2246 although only one jet multiplicity bin is used.
 - 2247 • **WHSS category:** two WZ CRs are defined, for events with one and
 2248 two jets, respectively. These are common between the WHSS and ZH3l
 2249 channels, meaning that a single rate parameter scales the WZ yields in
 2250 the CR, the WHSS SR, and the ZH3l SR for 1-jet events, and a second
 2251 rate parameter does the same for 2-jet events in the aforementioned
 2252 regions.
 - 2253 • **WH3l category:** the WZ and $V\gamma$ production processes are normalized
 2254 in the two corresponding CRs.
 - 2255 • **ZH3l category:** the WZ production process is normalized in the 1-
 2256 and 2-jet CRs in common with the WHSS category.
 - 2257 • **ZH4l category:** the ZZ production process is normalized in a CR
 2258 common to both the XDF and XSF SRs.
- 2259 Background templates corresponding to the 2017 and 2018 data sets are
 2260 scaled with common rate parameters (if their normalization is data driven),
 2261 while templates pertaining to the 2016 data set are scaled with an indepen-
 2262 dent set of rate parameters. This is done to account for the different PS and
 2263 UE tunes in the two cases, which may give rise to tensions in the fit if the
 2264 background normalization was forced to be the same in all years.

2265 5.3 Sources of systematic uncertainty

2266 Sources of systematic uncertainty are accounted for in the form of nuisance
 2267 parameters, as detailed in Section 5.1. They can be roughly divided into
 2268 uncertainties of experimental and of theoretical origin. When not otherwise
 2269 stated, theoretical uncertainties are correlated between the MC templates
 2270 describing the three years of data taking. This is achieved by having a
 2271 common nuisance parameter scaling the templates (or regulating their shape)
 2272 across all data sets. Uncertainties of experimental nature on the other hand
 2273 are, unless otherwise stated, uncorrelated between the three years. This is
 2274 due to changes to the detector in between each of the three years of data
 2275 taking, and is achieved by assigning separate nuisance parameters to the
 2276 2016, 2017, and 2018 data sets.

2277 5.3.1 Sources of theoretical uncertainty

2278 Sources of theoretical uncertainty include:

- 2279 • **Parton distribution functions:** the effect of varying the PDFs within
 2280 their uncertainty. Two uncertainty schemes are used depending on the
 2281 process in question: for most backgrounds and all signals, the change
 2282 in shape of the templates resulting from varying the PDFs was found
 2283 to be negligible, with the only relevant effect being a change in the ac-
 2284 ceptance of the selection requirements, and thus in normalization only.
 2285 Uncertainty on the cross section calculation of signal processes due to
 2286 the uncertainty on the PDFs is also included as a separate normaliza-
 2287 tion uncertainty. For the non-resonant WW and t quark background,
 2288 the shape effect was found to be non-negligible and is thus included in
 2289 the analysis. The implementation follows the PDF4LHC recommenda-
 2290 tions [58].
- 2291 • **Missing higher-order effects:** the uncertainty due to cross section
 2292 calculations being done at fixed order in perturbation theory, which can
 2293 affect both the normalization and shape of MC templates. For back-
 2294 ground processes, the impact is evaluated by increasing and reducing
 2295 the renormalization and factorization scales by a factor of two. The en-
 2296 velope of all combinations (i.e., the largest bin-by-bin shifts across all
 2297 combinations of variations of the two scales, excluding those in which
 2298 one is scaled by two and the other by 0.5) is taken as the uncertainty for
 2299 all backgrounds except t quark production, for which the variations are
 2300 built by shifting both the factorization and renormalization scales up
 2301 or down by a factor two simultaneously. The reason for this difference

is that the envelope approach in this case was found to de-emphasize shape effects in favor of normalization ones, leading to a worse description of the data. In the case of the ggH and VBF processes, the uncertainty scheme is more complex, with separate uncertainties being assigned to the normalization scale, resummation scale, migration between jet multiplicity bins, H boson p_T , and m_{jj} spectrum (in the 2-jet phase space) [17]. Remaining signal processes are assigned uncertainties on their cross sections and acceptance, similarly to what is done for backgrounds.

- **Parton shower modelling:** the uncertainty in the number of hadronic jets from initial- and final-state radiation (ISR and FSR), evaluated by varying the ISR and FSR scales up and down by a factor of two in PYTHIA and assigning a separate uncertainty to each. The alternative templates obtained this way are reweighted to match the normalization of the nominal template, in order to only account for the variation in shape due to ISR and FSR.
- **Underlying event modelling:** obtained by shifting the UE tunes, and amounting to an uncertainty of about 1.5% on event yields. Since the CUETP8M1 tune is used in 2016 and the CP5 one for 2017 and 2018, the UE uncertainty in 2016 is uncorrelated from the others.

5.3.2 Sources of experimental uncertainty

Sources of experimental uncertainty include:

- **Integrated luminosity:** the three years of data taking have uncertainties on the integrated luminosity ranging between 1.2 and 2.5% individually, with the overall uncertainty on the full run 2 data set being of 1.6% [29, 37, 38]. This uncertainty is partially correlated among the three data sets, due to the different techniques that are used in the measurement. This is a normalization uncertainty.
- **Pileup:** the uncertainty in the amount of PU interactions, modelled by varying the p-p cross section, affecting both the shape and normalization of MC templates.
- **Lepton trigger efficiency:** the uncertainty in the evaluation of lepton trigger efficiencies, as described in Section 2.3. These are evaluated in bins of η and p_T by varying the $m_{\ell\ell}$ window when selecting Z boson candidates and the requirements on the probe lepton in the tag and

- 2337 probe procedure. They amount to about 2% for electrons and 1% for
 2338 muons on average, and can have both a normalization and shape effect
 2339 on the templates.
- 2340 • **Lepton identification and isolation efficiency:** the uncertainty on
 2341 the HWWID efficiency on electrons and muons, evaluated by varying
 2342 the fit function, number of bins, $m_{\ell\ell}$ window, and tag muon require-
 2343 ments in the tag and probe procedure described in Section 4.2.
 - 2344 • **Lepton, jet, and missing transverse momentum scales:** the un-
 2345 certainty on the momentum scale of reconstructed objects can affect
 2346 both the normalization (through acceptance effects) and shape of the
 2347 templates. Uncertainties on electron and muon scales arise primarily
 2348 from the finite statistics of the samples used to measure them, and are
 2349 therefore uncorrelated among the three years. Their normalization ef-
 2350 fect is between 0.6 and 1% for electrons and about 0.2% for muons. The
 2351 uncertainty on the energy scale of jets is modelled with 11 independent
 2352 nuisance parameters, six of which are correlated between data sets.
 2353 Their normalization effect varies between 1 and 10%, strongly depend-
 2354 ing on the jet multiplicity in the phase space of interest. Finally, the
 2355 uncertainty on $\vec{p}_T^{\text{miss}}$ includes both the propagation of the uncertainties
 2356 on the energy scale of jets, as well as an additional component from
 2357 unclustered energy i.e., PF candidates not clustered in any hadronic jet
 2358 and not identified with leptons or photons. The normalization effect of
 2359 the uncertainty on $\vec{p}_T^{\text{miss}}$ ranges from 1 to 10%.
 - 2360 • **Jet energy resolution:** the uncertainty in the energy resolution of
 2361 jets, which has both a shape and normalization effect on the templates.
 2362 The latter is found to be below 1% across the phase space of interest
 2363 to the analysis.
 - 2364 • **Pileup identification efficiency:** the uncertainty on the PUJetID
 2365 defined in Section 4.2. It affects both the shape and normalization of
 2366 the templates, with the former effect found to be below 1%.
 - 2367 • **b tagging:** the uncertainty on the efficiency of the b taggers is mod-
 2368 elled with 17 nuisance parameters, five of which are theoretical in na-
 2369 ture and thus correlated among the data sets. These uncertainties affect
 2370 both the shape and the normalization of the templates.
 - 2371 • **Finite MC sample size:** the statistical uncertainty on the number
 2372 of generated events in each bin of the MC templates. The Barlow-
 2373 Beeston-lite approach [59] is used, which relies on the assumption that

for sufficiently high number of entries in a bin, the distribution is gaussian. The advantage of this approach is that, if the number of simulated events in a bin is large enough, the uncertainty on the sum of events from all individual MC templates can be modelled by one gaussian uncertainty, instead of requiring one nuisance parameter per bin and per process. In bins where the number of simulated events is below 10, this approximation no longer holds, and individual poissonian uncertainties are assigned to each process.

5.3.3 Process-specific uncertainties

The following sources of uncertainty only apply to specific MC samples:

- **Non-prompt background:** experimental uncertainties on the estimation of non-prompt lepton rates described in Section 4.2 arise from the finite sample size of the pass-fail sample and from the different jet flavor composition in the sample used to evaluate the non-prompt rate and the SR. The statistical component has an effect below 5% for non-prompt electrons and 10% for non-prompt muons. The effect of the difference in jet flavor is more challenging to assess. The method used in this analysis consists in varying the p_T of the jet recoiling against the loose lepton object in the dijet sample. This amounts to varying the fraction of the lepton-like object p_T to the p_T of the jet that contains it, which translates into probing different jet flavor compositions. This procedure results in an uncertainty of about 8%. Finally, a 30% normalization uncertainty is assigned to the non-prompt lepton contribution, to cover any residual discrepancy between the prediction and data from a closure check in the same-sign CRs.
- **Quark-induced WW background:** the theoretical uncertainty on the NNLO + NNLL reweighting procedure is evaluated by varying the renormalization, factorization, and resummation scales. In categories where the normalization of the WW background is taken from data, only the effect on the shape of the template is included.
- **Gluon-induced WW background:** the gluon-induced WW sample is produced at LO and normalized to the NLO cross section with a factor of 1.4 [60]. A theoretical uncertainty of 15% on the scale factor is considered.
- **t quark background:** on top of varying the factorization and normalization scales, the ratio between production of $t\bar{t}$ pairs to production of single t quarks is assigned an 8% theoretical uncertainty.

- 2411 • **Tau-embedded sample:** the experimental uncertainty resulting from
2412 the embedding technique described in Section 4.1 is evaluated by vary-
2413 ing the electron and muon energy scales in the data sample used for
2414 embedding.

2415 **5.3.4 Theoretical uncertainties in the STXS measure-** 2416 **ment**

2417 Special care is taken with regards to the treatment of systematic uncertain-
2418 ties of theoretical nature in the STXS measurement. Since what is measured
2419 in this specific case is directly the cross section of the process of interest,
2420 theoretical uncertainties on the signal cross sections themselves should be re-
2421 moved from the fit. This is in contrast to the measurement of signal strength
2422 modifiers, since the latter contain the expected cross section in the denom-
2423 inator (Eq. 5.9). On the other hand, any shape uncertainty on the tem-
2424 plates, as well as on migration of events between different STXS bins, has
2425 to be accounted for. To achieve this, the alternative signal templates for the
2426 STXS fits corresponding to theoretical uncertainties have been reweighted,
2427 so that their normalization matches the nominal template. In this way, the
2428 uncertainty on the cross sections is effectively removed, while retaining the
2429 theoretical uncertainty on the event kinematics.

Chapter 6

Results

Results are extracted by performing a simultaneous maximum likelihood fit to all analysis categories outlined in Section 4.3. The subcategorization and the distributions used in each analysis category are summarized in Table 6.1. Four separate fits are done, corresponding to the four sets of results presented in this chapter:

- the signal strength modifier for the total production cross section of the H boson is measured by scaling all signal templates by a common POI i.e., keeping the relative fraction of signal events from the different production modes fixed to the SM expectation;
- the signal strength modifiers for the four target H boson production cross sections (ggH, VBF, WH, and ZH), are obtained by scaling each of the four signal templates by an independent POI. Note that this accounts for cross contamination of different processes in different categories e.g., the ggH contamination in the VBF-like phase space;
- the couplings of the H boson to vector bosons and fermions are extracted in the κ framework [61], as detailed in the following;
- the STXS cross sections are extracted by fitting the modified analysis template described in Section 4.3, scaling each of the 14 target STXS bins by an independent POI.

The number of expected and measured events for signal and background processes, as well as the number of observed events in each category, extracted from the per-production mode signal strength fit, are reported in Tables 6.2, 6.3, 6.4, and 6.5. The goodness of fit test as described in Section 5.1 has been performed, yielding a p-value ranging from 1 to 3% depending on the fit.

Table 6.1: Overview of the fit variables and SR subcategorization used in each analysis category. In all CRs, the number of events is used as fit variable. The number of subcategories shown in the last column includes both SRs and CRs.

Category	SR subcategorization	SR fit variable	$N_{\text{subcategories}}$
ggH DF	$(0j, 1j) \times (p_T^{\ell_2} \leq 20 \text{ GeV}) \times (\ell^\pm \ell^\mp), (\geq 2j)$	$(m_{\ell\ell}, m_T^H)$	15
ggH SF	$(0j, 1j, \geq 2j) \times (ee, \mu\mu)$	N_{events}	12
VBF DF	$\max_j C_j$	DNN output	6
VBF SF	$(ee, \mu\mu)$	N_{events}	4
WHSS	$(\text{DF}, \text{SF}) \times (1j, 2j)$	$\tilde{m}_H$	4
WH3l	SF lepton pair with opposite or same sign	BDT output	4
ZH3l	$(1j, 2j)$	m_T^H	4
ZH4l	(DF, SF)	BDT output	3
VH2j DF	-	$m_{\ell\ell}$	3
VH2j SF	$(ee, \mu\mu)$	N_{events}	4

Table 6.2: Number of events by process in the ggH DF categories after the fit to the data, scaling the ggH, VBF, WH, and ZH production modes separately. The ttH contribution is fixed to its SM expectation. Numbers in parenthesis indicate expected yields.

Process	0-jets ggH DF	1-jet ggH DF	2-jets ggH DF
ggH	1875 ± 45 (2157)	881 ± 28 (942)	67 ± 5 (71)
VBF	15 ± 2 (23)	62 ± 7 (92)	4 ± 1 (6)
WH	103 ± 7 (51)	124 ± 10 (60)	18 ± 2 (9)
ZH	38 ± 3 (19)	33 ± 3 (17)	7 ± 1 (4)
ttH	-	1 ± 1 (1)	1 ± 1 (1)
<i>Tot. sig.</i>	2032 ± 51 (2250)	1101 ± 31 (1111)	99 ± 6 (90)
WW	37297 ± 285 (34781)	12703 ± 307 (14932)	748 ± 121 (1101)
Top quark	10165 ± 179 (10204)	19711 ± 298 (19766)	3989 ± 123 (3868)
Nonprompt	4407 ± 225 (5888)	1999 ± 141 (2769)	252 ± 42 (262)
DY	495 ± 24 (563)	822 ± 12 (792)	87 ± 4 (86)
VZ/V γ^*	1464 ± 45 (1776)	1297 ± 44 (1531)	123 ± 7 (140)
V γ	1181 ± 19 (1273)	723 ± 18 (777)	57 ± 3 (56)
Triboson	38 ± 1 (39)	66 ± 1 (72)	13 ± 1 (14)
<i>Tot. backg.</i>	55045 ± 409 (54524)	37321 ± 453 (40639)	5269 ± 178 (5526)
<i>Tot. pred.</i>	57077 ± 412 (56773)	38422 ± 454 (41750)	5368 ± 178 (5616)
<i>Data</i>	57024	38373	5380

Table 6.3: Number of events by process in the ggH SF categories after the fit to the data, scaling the ggH, VBF, WH, and ZH production modes separately. The ttH contribution is fixed to its SM expectation. Numbers in parenthesis indicate expected yields.

Process	0-jets ggH SF	1-jet ggH SF	2-jets ggH SF
ggH	780 ± 31 (891)	397 ± 18 (422)	86 ± 7 (89)
VBF	5 ± 1 (7)	29 ± 4 (42)	10 ± 1 (13)
WH	24 ± 3 (11)	34 ± 4 (16)	12 ± 1 (6)
ZH	14 ± 1 (7)	16 ± 2 (8)	7 ± 1 (3)
ttH	-	-	1 ± 1 (1)
<i>Tot. sig.</i>	823 ± 31 (915)	476 ± 18 (489)	114 ± 7 (112)
WW	7034 ± 184 (6464)	2711 ± 128 (3064)	276 ± 61 (480)
Top quark	1345 ± 42 (1294)	3711 ± 75 (3524)	1879 ± 51 (1758)
Nonprompt	641 ± 88 (701)	366 ± 54 (412)	103 ± 18 (119)
DY	3149 ± 271 (2706)	4098 ± 197 (3284)	1403 ± 83 (829)
VZ/V γ^*	327 ± 13 (371)	270 ± 10 (301)	63 ± 4 (70)
V γ	138 ± 6 (145)	193 ± 15 (201)	48 ± 5 (47)
Triboson	4 ± 1 (5)	10 ± 1 (11)	6 ± 1 (6)
<i>Tot. backg.</i>	12639 ± 342 (11684)	11359 ± 253 (10797)	3777 ± 117 (3309)
<i>Tot. pred.</i>	13462 ± 343 (12599)	11835 ± 254 (11286)	3891 ± 117 (3421)
<i>Data</i>	13507	11976	3950

Table 6.4: Number of events by process in the VBF and VH2j categories after the fit to the data, scaling the ggH, VBF, WH, and ZH production modes separately. The ttH contribution is fixed to its SM expectation. Numbers in parenthesis indicate expected yields.

Process	VBF DF	VBF SF	VH2j DF	VH2j SF
ggH	114 ± 8 (115)	21 ± 2 (21)	36 ± 3 (39)	27 ± 2 (29)
VBF	62 ± 11 (91)	39 ± 5 (57)	2 ± 1 (3)	2 ± 1 (2)
WH	14 ± 1 (7)	1 ± 1 (1)	26 ± 4 (13)	16 ± 2 (8)
ZH	5 ± 1 (2)	1 ± 1 (0)	13 ± 2 (7)	8 ± 1 (4)
ttH	-	-	-	-
<i>Tot. sig.</i>	195 ± 14 (215)	62 ± 6 (79)	77 ± 5 (62)	53 ± 3 (43)
WW	1319 ± 57 (1368)	109 ± 17 (102)	98 ± 44 (205)	56 ± 22 (134)
Top quark	2875 ± 65 (3148)	267 ± 8 (249)	743 ± 32 (730)	539 ± 16 (514)
Nonprompt	404 ± 36 (399)	28 ± 4 (32)	81 ± 13 (113)	62 ± 10 (72)
DY	249 ± 4 (241)	402 ± 27 (465)	77 ± 3 (77)	555 ± 48 (479)
VZ/V γ^*	184 ± 9 (221)	11 ± 1 (12)	49 ± 3 (55)	23 ± 2 (27)
V γ	110 ± 4 (117)	10 ± 1 (10)	26 ± 3 (25)	16 ± 5 (17)
Triboson	11 ± 1 (11)	1 ± 1 (1)	6 ± 1 (7)	4 ± 1 (3)
<i>Tot. backg.</i>	5154 ± 94 (5505)	827 ± 33 (871)	1080 ± 56 (1212)	1255 ± 56 (1245)
<i>Tot. pred.</i>	5349 ± 95 (5720)	889 ± 34 (950)	1157 ± 56 (1274)	1308 ± 56 (1288)
<i>Data</i>	5254	862	1164	1318

6.1 Signals strength modifiers

The overall signal strength modifier for production of a H boson in 13 TeV p-p collisions in the WW decay channel is found to be:

$$\mu = 0.95^{+0.10}_{-0.09} = 0.95 \pm 0.05(\text{stat}) \pm 0.08(\text{syst}), \quad (6.1)$$

showing excellent agreement with the SM expectation of $\mu = 1$. The uncertainty has been broken down into its statistical and systematic components. The statistical component Δ_{stat} is calculated by performing a likelihood fit keeping all nuisance parameters fixed at their best-fit values. The systematic uncertainty Δ_{syst} is obtained as:

$$\Delta_{\text{syst}} = \sqrt{\Delta_{\text{tot}}^2 - \Delta_{\text{stat}}^2}. \quad (6.2)$$

The likelihood function profiled over all nuisance parameters is shown in Fig. 6.1, considering both the full set of uncertainties as well as the statistical component only. This marks one of the most precise signal strength measurements for production of a H boson in a single decay channel at the time of writing, with a $\sim 50\%$ improvement in relative uncertainty over the

Table 6.5: Number of events by process in the WHSS, WH3l, ZH3l, and ZH4l categories after the fit to the data, scaling the ggH, VBF, WH, and ZH production modes separately. The ttH contribution is fixed to its SM expectation. Numbers in parenthesis indicate expected yields.

Process	WHSS	WH3l	ZH3l	ZH4l
ggH	1 ± 1 (1)	-	-	-
VBF	-	-	-	-
WH	148 ± 12 (69)	44 ± 5 (20)	2 ± 1 (1)	-
ZH	10 ± 11 (5)	3 ± 1 (2)	74 ± 7 (36)	19 ± 2 (10)
ttH	1 ± 1 (1)	-	1 ± 1 (1)	-
<i>Tot. sig.</i>	159 ± 12 (76)	48 ± 5 (22)	76 ± 7 (38)	19 ± 2 (10)
WW	40 ± 1 (39)	-	-	-
Top quark	62 ± 1 (62)	-	-	-
Nonprompt	596 ± 37 (805)	55 ± 6 (85)	166 ± 16 (215)	-
DY	28 ± 7 (35)	-	30 ± 1 (29)	1 ± 1 (1)
VZ/V γ^*	1309 ± 26 (1355)	311 ± 10 (276)	1905 ± 25 (1796)	45 ± 1 (39)
V γ	135 ± 11 (162)	14 ± 3 (20)	36 ± 6 (40)	-
Triboson	41 ± 1 (41)	15 ± 1 (15)	30 ± 1 (30)	3 ± 1 (3)
<i>Tot. backg.</i>	2211 ± 47 (2498)	396 ± 12 (397)	2167 ± 30 (2110)	50 ± 1 (44)
<i>Tot. pred.</i>	2370 ± 49 (2574)	444 ± 13 (419)	2243 ± 31 (2148)	69 ± 2 (54)
<i>Data</i>	2359	423	2315	69

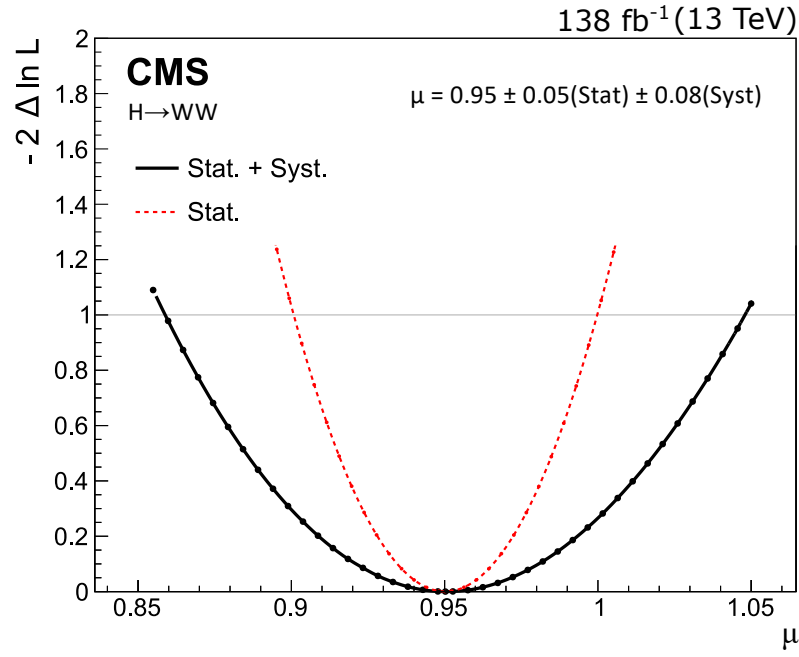


Figure 6.1: Observed NLL as a function of the overall signal strength modifier μ . The continuous black line shows the NLL with the full set of uncertainties. The dashed red line shows the NLL with only the statistical uncertainty on data. The uncertainty is taken as the interval of μ for which $\text{NLL}(\mu) - \text{NLL}(\mu = 0.95)$ is lower than 0.5.

previous CMS result in this channel [62], thanks in large part to the statistical power of the Run 2 data set, as well as the inclusion of additional channels and techniques. The systematic component is found to be the main contribution to the uncertainty.

The results for signal strength modifiers pertaining to each of the target production modes and the corresponding correlation matrix are shown in Fig. 6.2. They correspond to observed (expected) significances with respect to the background only hypothesis of 10.5 (11.8) σ , 3.15 (4.74) σ , 3.61 (1.82) σ , and 3.73 (2.19) σ for the ggH, VBF, WH, and ZH production processes, respectively. The correlation matrix between the four signal strength modifiers is shown in Fig. 6.3. The compatibility of the observed per-production mode signal strengths with the SM expectation is calculated by performing a second likelihood maximization while fixing all signal strength modifiers to their SM expectation of one. The difference in NLLs between the two fits is then used to calculate a p-value based on a χ^2 distribution with four degrees of freedom, corresponding to the four signal strength modifiers:

$$p = \int_{\Delta_{\text{NLL}}}^{\infty} f_{\chi^2}. \quad (6.3)$$

The observed p-value is 7%. Table 6.6 shows a more detailed breakdown of the impact of different sources of uncertainty on the result. The impact of each source is obtained as the difference in quadrature between the total uncertainty and the one obtained from a fit in which the nuisance parameters whose impact is being evaluated are fixed to their best-fit values. The agreement with the SM is found to be good, with only a slight excess in the WH result, of the order of two standard deviations from the expectation. The improvement with respect to the previous CMS result in this channel [62] is particularly noticeable in the VBF, WH, and ZH signal strength measurements, all exceeding the factor ~ 2 of reduction in uncertainty given by the larger data set available. This is due to the updated analysis strategy, in particular the switch to a DNN setup for the VBF measurement and the inclusion of the ZH3l and WHSS channels for the WH and ZH measurements, respectively.

6.2 Couplings to vector bosons and fermions

The couplings of the H boson are measured in the κ framework. Two coupling modifiers κ_V and κ_f are defined, for couplings to vector bosons and fermions, respectively. These are linked to the signal yields in the analysis categories

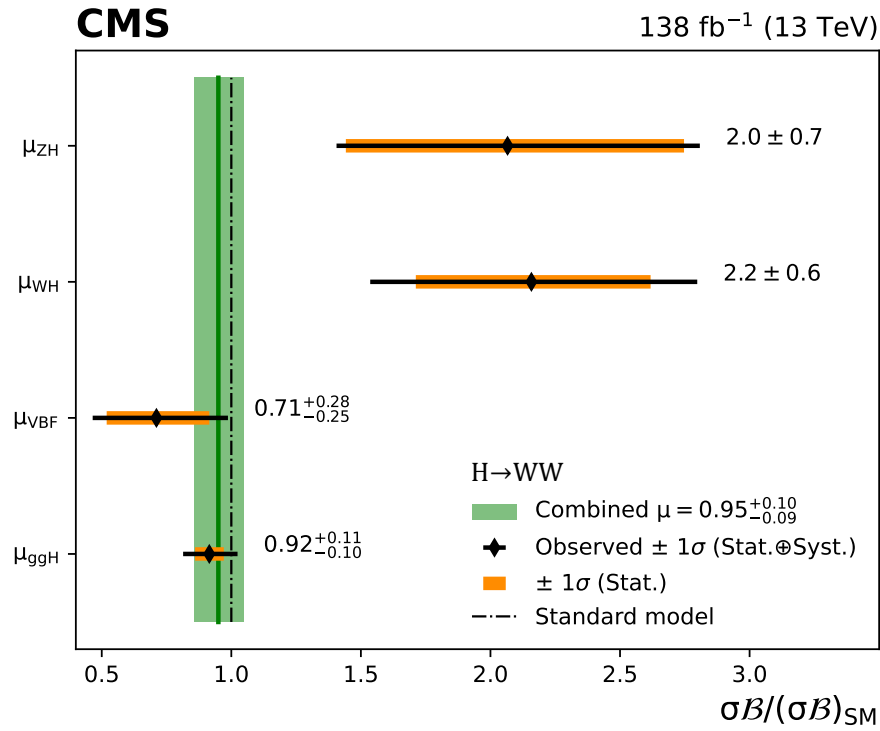


Figure 6.2: Observed signal strength modifiers for the ggH, VBF, WH, and ZH production modes. Black and orange bars show the total and statistical uncertainties, respectively. The dashed black line corresponds to the SM prediction of $\mu = 1$. The solid green line shows the overall signal strength modifier and the corresponding uncertainty band.

Table 6.6: Contributions of different sources of uncertainty in the signal strength measurement. The systematic component includes the combined effect from all sources besides data-driven background normalization and the size of the data set, which make up the statistical part.

Uncertainty source	$\Delta\mu/\mu$	$\Delta\mu_{\text{ggH}}/\mu_{\text{ggH}}$	$\Delta\mu_{\text{VBF}}/\mu_{\text{VBF}}$	$\Delta\mu_{\text{WH}}/\mu_{\text{WH}}$	$\Delta\mu_{\text{ZH}}/\mu_{\text{ZH}}$
Theory (signal)	4%	5%	13%	2%	<1%
Theory (background)	3%	3%	2%	4%	5%
Lepton misidentification	2%	2%	9%	15%	4%
Integrated luminosity	2%	2%	2%	2%	3%
b tagging	2%	2%	3%	<1%	2%
Lepton efficiency	3%	4%	2%	1%	4%
Jet energy scale	1%	<1%	2%	<1%	3%
Jet energy resolution	<1%	1%	<1%	<1%	3%
$p_{\text{T}}^{\text{miss}}$ scale	<1%	1%	<1%	2%	2%
PDF	1%	2%	<1%	<1%	2%
Parton shower	<1%	2%	<1%	1%	1%
Data-driven backg. norm.	3%	4%	6%	4%	6%
Stat. uncertainty	5%	6%	28%	21%	31%
Syst. uncertainty	9%	10%	23%	19%	11%
Total uncertainty	10%	11%	36%	29%	33%

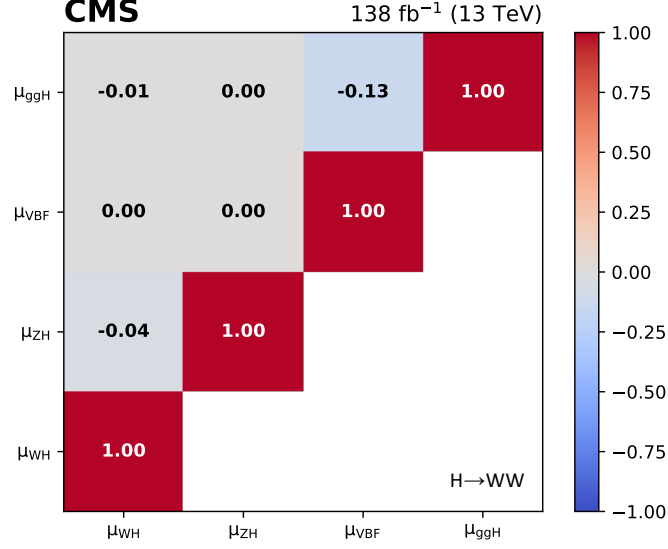


Figure 6.3: Correlation matrix between the signal strength modifiers of the four target production modes of the H boson.

as follows:

$$\sigma\mathcal{B}(X_i \rightarrow H \rightarrow WW) = \kappa_i^2 \frac{\kappa_V^2}{\kappa_H^2} \sigma_{SM} \mathcal{B}_{SM}, \quad (6.4)$$

where $\kappa_H = \kappa_H(\kappa_V, \kappa_f)$ is the modifier of the total decay width of the H boson, and X_i are the different production modes. The corresponding coupling modifiers κ_i are equal to κ_V for the VBF, WH and ZH modes, and κ_f for ggH, given that the H boson couples to gluons through a quark loop. This setup assumes universal scaling of the couplings to Z and W bosons i.e., $\kappa_Z = \kappa_W$, as well as (separately) to fermions. This assumption is made because the $H \rightarrow WW$ channel by itself does not have the statistical power to effectively separate e.g., κ_Z from κ_W . The best-fit values for the coupling modifiers are found to be:

$$\kappa_V = 0.99 \pm 0.05, \quad \kappa_f = 0.86^{+0.14}_{-0.11}. \quad (6.5)$$

The $H \rightarrow WW$ channel achieves 5% accuracy in the coupling to vector bosons in large part due to the constraint from the decay vertex. Good precision is achieved in the coupling to fermions too, mainly due to the precision in constraining the ggH production cross section. The two-dimensional likelihood as a function of the coupling modifiers is shown in Fig. 6.4, showing once again excellent agreement with the SM prediction of $\kappa_V = \kappa_f = 1$.

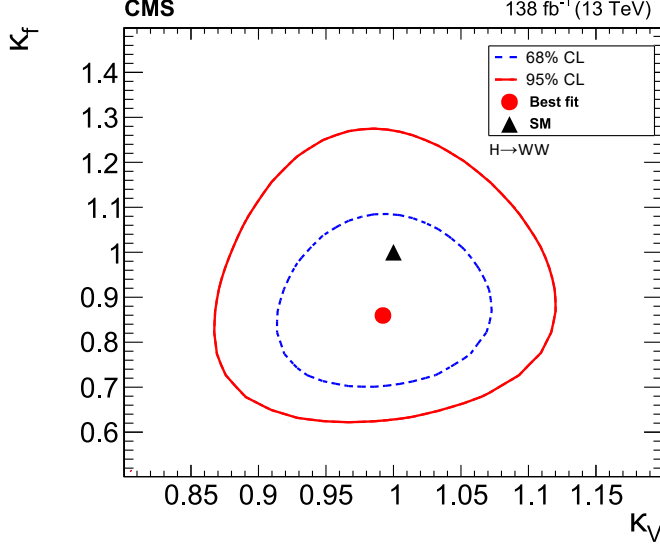


Figure 6.4: Two-dimensional likelihood as a function of κ_V and κ_f . The dashed blue and continuous red lines show the 68 and 95% confidence level contours, respectively.

6.3 STXS

Results in the STXS framework are reported both as cross sections, which form the final result, as well as signal strength modifiers in each STXS bins, in order to more easily gauge the agreement with the SM. Results are extracted by simultaneously fitting all 14 STXS bins, and are shown in Table 6.7 and Fig. 6.5. The correlation matrix between all measured STXS cross sections is shown in Fig. 6.6. The signal yields in two of the STXS bins are measured to be below zero due to an upwards fluctuation of the background. When reporting results in terms of cross sections, an upper limit at 68% confidence level is given in these cases, instead of a symmetric interval. The STXS measurement benefits heavily from the statistical power of the Run 2 data set. Comparing with the previous result by CMS [63], based on the 2016 dataset only, it allowed to move from constraining five cross sections (one per production mode), to the set of 14 STXS bins reported here. The precision achieved in the low p_T^H and low jet multiplicity bins, as well as in the WH production bins, is on par or better than results in other decay channels available at the time of writing [64, 65, 66, 67, 68]. The measurement in high p_T^H bins, as well as in ZH production bins, is still limited by the statistical uncertainty. The STXS measurement will once again benefit from the larger

Table 6.7: Observed cross sections of the $H \rightarrow WW$ process in each STXS bin. The uncertainties in the observed cross sections and their ratio to the SM expectation do not include the theoretical uncertainties on the latter. In cases where the ratio to the SM cross section is measured below zero, an upper limit at 68% confidence level on the observed cross section is reported. All dimensional quantities in STXS bin definitions are measured in GeV.

STXS bin	$\sigma(H \rightarrow WW)/\sigma(H \rightarrow WW)_{\text{SM}}$	$\sigma(H \rightarrow WW)$ [pb]	$\sigma(H \rightarrow WW)_{\text{SM}}$ [pb]
ZH ($Z \rightarrow \text{leptons}$); $p_T^V > 150$	$-0.1^{+1.2}_{-0.9}(\text{stat}) \pm 0.1(\text{theo})^{+0.4}_{-0.3}(\text{exp})$	< 0.03	0.139 ± 0.013
ZH ($Z \rightarrow \text{leptons}$); $p_T^V < 150$	$3.3^{+1.0}_{-0.9}(\text{stat}) \pm 0.1(\text{theo})^{+0.4}_{-0.3}(\text{exp})$	0.10 ± 0.03	0.030 ± 0.004
WH ($W \rightarrow \text{leptons}$); $p_T^V > 150$	$3.8^{+1.5}_{-1.3}(\text{stat}) \pm 0.1(\text{theo})^{+0.8}_{-0.7}(\text{exp})$	$0.8^{+0.4}_{-0.3}$	0.22 ± 0.02
WH ($W \rightarrow \text{leptons}$); $p_T^V < 150$	$1.6 \pm 0.8(\text{stat}) \pm 0.1(\text{theo})^{+0.7}_{-0.6}(\text{exp})$	0.06 ± 0.04	0.035 ± 0.005
qqH; $60 < m_{jj} < 120$	$4.1 \pm 2.6(\text{stat})^{+0.7}_{-0.6}(\text{theo}) \pm 2.2(\text{exp})$	1.5 ± 1.2	0.36 ± 0.01
qqH; $p_T^H > 200$	$1.1^{+0.7}_{-0.6}(\text{stat}) \pm 0.1(\text{theo}) \pm 0.3(\text{exp})$	$0.17^{+0.11}_{-0.10}$	0.15 ± 0.02
qqH; $p_T^H < 200$; $m_{jj} > 700$	$0.7 \pm 0.3(\text{stat}) \pm 0.1(\text{theo}) \pm 0.2(\text{exp})$	$0.023^{+0.011}_{-0.010}$	0.032 ± 0.004
qqH; $p_T^H < 200$; $350 < m_{jj} < 700$	$0.4^{+0.8}_{-0.7}(\text{stat}) \pm 0.2(\text{theo}) \pm 0.5(\text{exp})$	0.04 ± 0.10	0.11 ± 0.03
ggH; $p_T^H > 300$	$-2.1^{+1.7}_{-1.5}(\text{stat})^{+0.2}_{-0.3}(\text{theo})^{+1.6}_{-2.0}(\text{exp})$	< 0.04	0.028 ± 0.009
ggH; $200 < p_T^H < 300$	$2.3 \pm 0.9(\text{stat}) \pm 0.1(\text{theo}) \pm 0.6(\text{exp})$	0.22 ± 0.10	0.09 ± 0.02
ggH; $\geq 2j$	$1.8 \pm 0.6(\text{stat}) \pm 0.4(\text{theo}) \pm 0.4(\text{exp})$	1.5 ± 0.7	0.9 ± 0.4
ggH; l_j ; $p_T^H > 60$	$0.41 \pm 0.25(\text{stat})^{+0.10}_{-0.06}(\text{theo}) \pm 0.17(\text{exp})$	0.5 ± 0.4	1.15 ± 0.16
ggH; l_j ; $p_T^H < 60$	$1.7 \pm 0.3(\text{stat}) \pm 0.2(\text{theo}) \pm 0.2(\text{exp})$	$2.6^{+0.7}_{-0.6}$	1.5 ± 0.2
ggH; $0j$	$0.74 \pm 0.07(\text{stat}) \pm 0.04(\text{theo})^{+0.08}_{-0.07}(\text{exp})$	$4.2^{+0.7}_{-0.6}$	5.8 ± 0.3

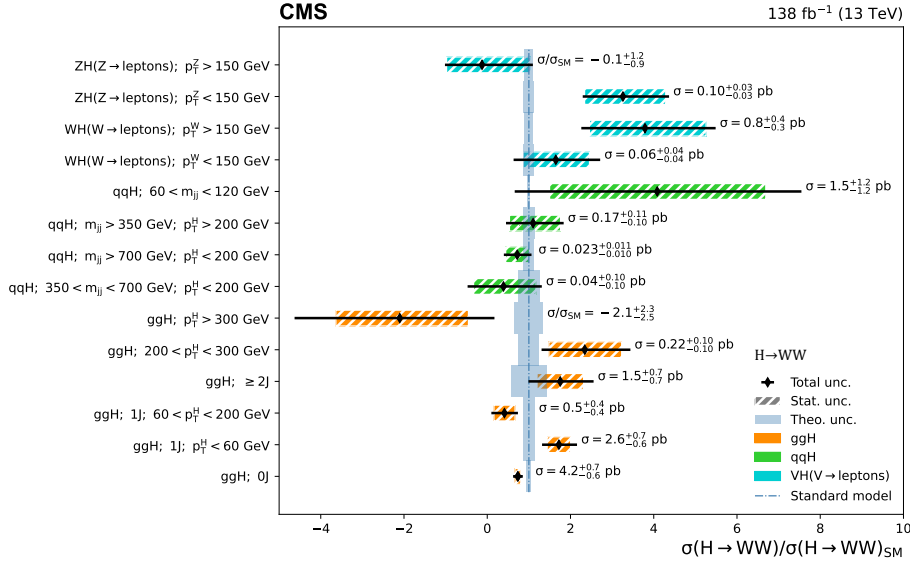


Figure 6.5: Observed signal strength modifiers and cross sections in each STXS bin. The dashed blue line shows the SM prediction with the corresponding theoretical uncertainty.

2539 data set gathered during Run 3 of the LHC, which is ongoing at the time of
 2540 writing, allowing for both increased precision as well as a finer splitting of
 2541 the phase space.

2542 6.4 Combination with other decay channels

2543 This analysis is part of the combined CMS measurement of the Higgs bo-
 2544 son production cross sections, branching ratios, and couplings with Run 2
 2545 data [69]. The combined analysis includes the $H \rightarrow WW$ (the work re-
 2546 ported in this thesis), $H \rightarrow ZZ$, $H \rightarrow \gamma\gamma$, $H \rightarrow Z\gamma$, $H \rightarrow \tau\tau$, $H \rightarrow bb$, and
 2547 $H \rightarrow \mu\mu$ decay channels. Results are reported in terms of per-production
 2548 mode and per-decay mode signal strength modifiers, shown in Fig. 6.7, as
 2549 well as couplings in the κ framework, shown in Fig. 6.8. While a compre-
 2550 hensive discussion of the combination procedure is beyond the scope of this
 2551 thesis, it is worth highlighting the impact of the $H \rightarrow WW$ analysis on the
 2552 combined results, with the μ_{WW} signal strength and coupling to W bosons
 2553 being among the most precise results.

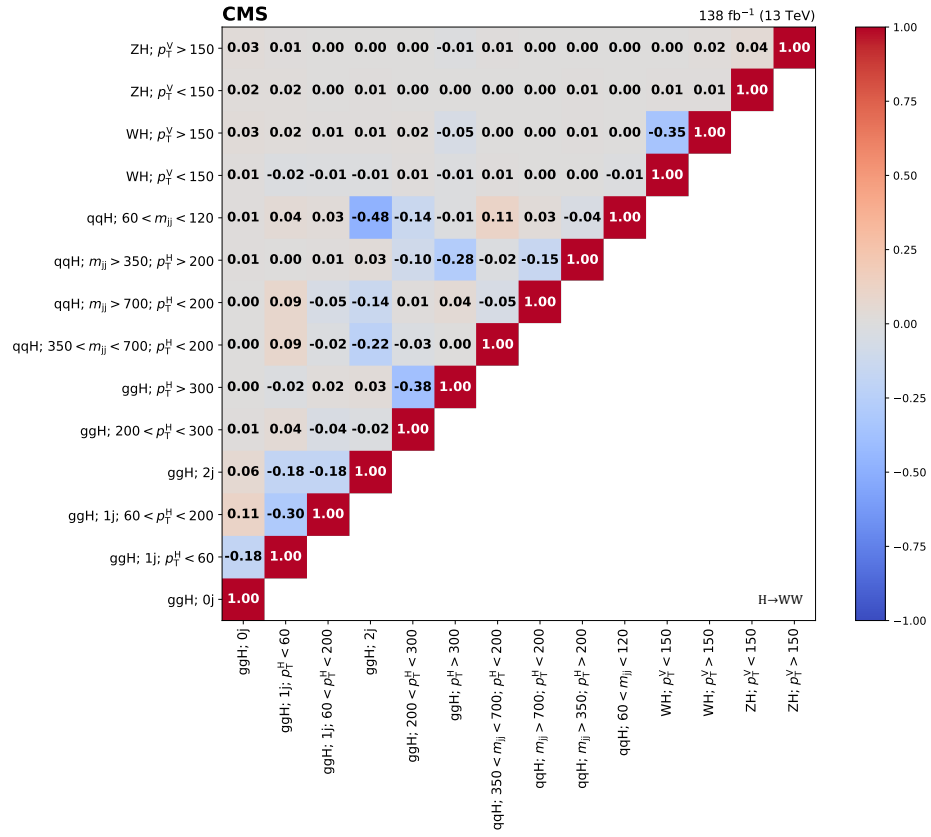


Figure 6.6: Correlation matrix between the 14 measured STXS cross sections.

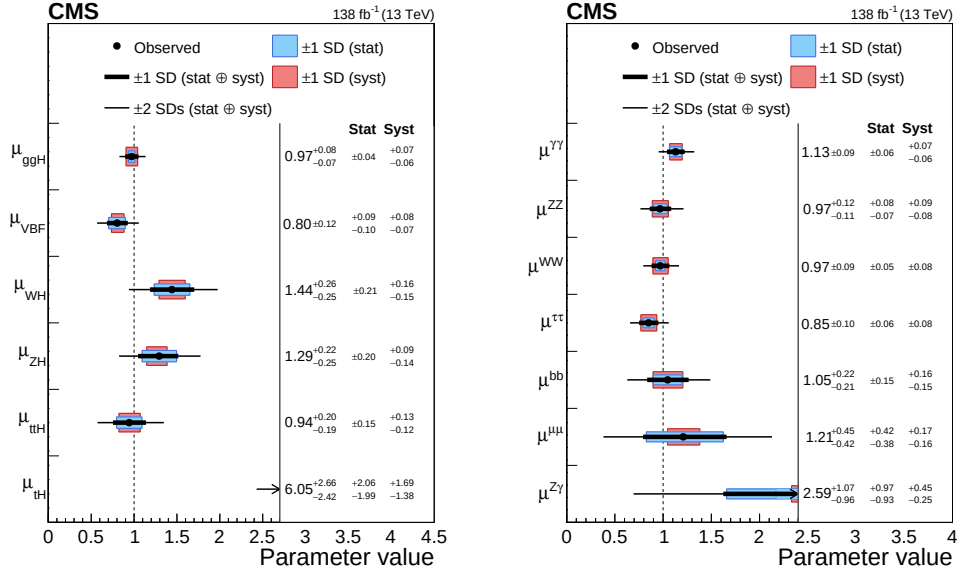


Figure 6.7: Per-production mode (left) and per-decay mode (right) signal strength modifiers, as extracted from the combined fit to all H boson decay modes [69].

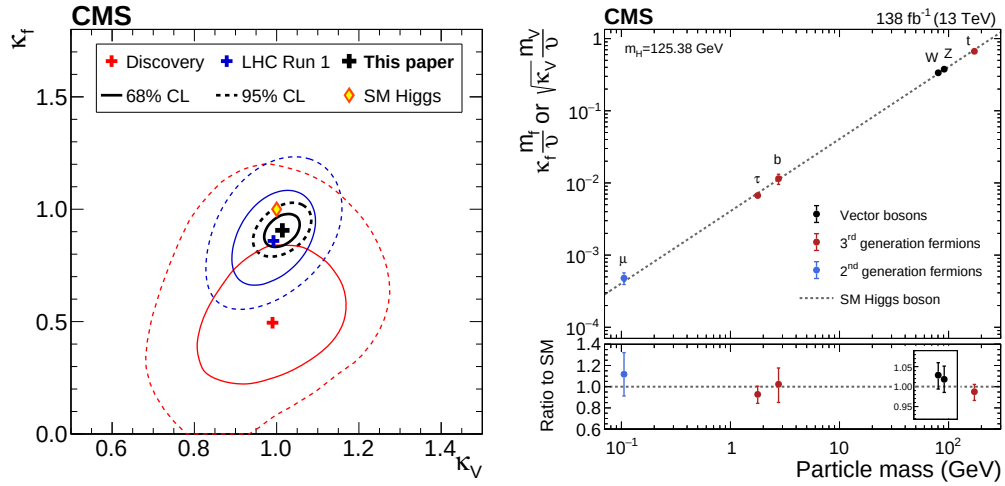


Figure 6.8: Left: two-dimensional likelihood as a function of κ_V and κ_f . Right: couplings of the H boson to SM particles as a function of their mass. Both results are as obtained from the combined fit to all H boson decay modes [69].

Conclusions

I have presented a measurement of the Higgs boson production cross sections and couplings in proton-proton collisions at the CERN LHC, using data gathered by CMS in the 2016–2018 timeframe. The $H \rightarrow WW$ decay channel is a prime candidate to accurately probe the dynamics of the Higgs sector of the SM, thanks to its high branching ratio and relatively low-background final state. The analysis targets Higgs boson production via gluon fusion, vector boson fusion, and in association with an electroweak W or Z boson, by designing a set of event selections tailored to each final state. The measurement is then carried out by performing a simultaneous fit to all analysis categories. The results I have presented mark a sizable improvement with respect to the previous CMS result in this channel, owing to the larger data set, additional final states, and improved analysis techniques, and are among the most precise results in a single decay channel at the time of writing. The analysis is also part of the combined CMS measurement of the Higgs boson production cross sections, branching ratios, and couplings with Run 2 data [69].

Bibliography

- 2571 [1] P. W. Higgs, “Broken symmetries and the masses of gauge bosons,” *Phys.*
2572 *Rev. Lett.*, vol. 13, pp. 508–509, Oct 1964.
- 2573 [2] F. Englert and R. Brout, “Broken symmetry and the mass of gauge
2574 vector mesons,” *Phys. Rev. Lett.*, vol. 13, pp. 321–323, Aug 1964.
- 2575 [3] The ATLAS Collaboration, “Observation of a new particle in the search
2576 for the standard model Higgs boson with the ATLAS detector at the
2577 LHC,” *Phys. Lett. B*, vol. 716, p. 1, 2012.
- 2578 [4] The CMS Collaboration, “Observation of a new boson at a mass of
2579 125 GeV with the CMS experiment at the LHC,” *Phys. Lett. B*, vol. 716,
2580 p. 30, 2012.
- 2581 [5] The CMS Collaboration, “Observation of a new boson with mass near
2582 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV,” *JHEP*, vol. 06, p. 081,
2583 2013.
- 2584 [6] P. Woit, *Quantum Theory, Groups and Representations: An Introduc-*
2585 *tion*. Springer International Publishing, 01 2017.
- 2586 [7] M. Peskin and D. Schroeder, *An Introduction To Quantum Field Theory*.
2587 Frontiers in Physics, Avalon Publishing, 1995.
- 2588 [8] S. L. Glashow, “Partial-symmetries of weak interactions,” *Nuclear*
2589 *Physics*, vol. 22, no. 4, pp. 579–588, 1961.
- 2590 [9] The CMS Collaboration, “A measurement of the higgs boson mass in the
2591 diphoton decay channel,” *Physics Letters B*, vol. 805, p. 135425, 2020.
- 2592 [10] R. L. Workman and Others, “Review of Particle Physics,” *PTEP*,
2593 vol. 2022, p. 083C01, 2022.
- 2594 [11] J. C. Collins, D. E. Soper, and G. Sterman, “Factorization of hard pro-
2595 cesses in qcd,” 2004.

- [12] G. Altarelli and G. Parisi, “Asymptotic freedom in parton language,” *Nuclear Physics B*, vol. 126, no. 2, pp. 298–318, 1977.
- [13] R. D. Ball, V. Bertone, S. Carrazza, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, and J. Rojo, “Parton distributions with qed corrections,” *Nuclear Physics B*, vol. 877, no. 2, pp. 290–320, 2013.
- [14] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, J. Rojo, and M. Ubiali, “Unbiased global determination of parton distributions and their uncertainties at nnlo and at lo,” *Nuclear Physics B*, vol. 855, no. 2, pp. 153–221, 2012.
- [15] R. D. Ball *et al.*, “Parton distributions from high-precision collider data,” *Eur. Phys. J. C*, vol. 77, p. 663, 2017.
- [16] C. Anastasiou, C. Duhr, F. Dulat, E. Furlan, T. Gehrmann, F. Herzog, A. Lazopoulos, and B. Mistlberger, “High precision determination of the gluon fusion Higgs boson cross-section at the LHC,” *JHEP*, vol. 05, p. 058, 2016.
- [17] The LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*. CERN Yellow Reports: Monographs, Geneva: CERN, Oct 2016.
- [18] The CMS Collaboration, “Measurement of the Higgs boson production rate in association with top quarks in final states with electrons, muons, and hadronically decaying tau leptons at $\sqrt{s} = 13$ TeV,” *Eur. Phys. J. C*, vol. 81, no. 4, p. 378, 2021.
- [19] The CMS Collaboration, “Observation of higgs boson decay to bottom quarks,” *Phys. Rev. Lett.*, vol. 121, p. 121801, Sep 2018.
- [20] LHC Higgs Cross Section Working Group, “Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector,” CERN Report CERN-2017-002-M, 2016.
- [21] D. J. Gross and F. Wilczek, “Ultraviolet behavior of non-abelian gauge theories,” *Phys. Rev. Lett.*, vol. 30, pp. 1343–1346, Jun 1973.
- [22] J. Greensite, *An introduction to the confinement problem*, vol. 821. 2011.
- [23] M. Thomson, *Modern Particle Physics*. Cambridge University Press, 2013.

- [24] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_t jet clustering algorithm,” *Journal of High Energy Physics*, vol. 2008, pp. 063–063, apr 2008.
- [25] The ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider,” *JINST*, vol. 3, p. S08003, 2008.
- [26] The CMS Collaboration, “The CMS Experiment at the CERN LHC,” *JINST*, vol. 3, p. S08004, 2008.
- [27] T. L. Collaboration, “The LHCb Detector at the LHC,” *JINST*, vol. 3, p. S08005, 2008.
- [28] The ALICE Collaboration, “The ALICE experiment at the CERN LHC,” *JINST*, vol. 3, p. S08002, 2008.
- [29] “CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV,” CMS Physics Analysis Summary CMS-PAS-LUM-18-002, 2019.
- [30] F. S. Cafagna, “Latest results for Proton-proton Cross Section Measurements with the TOTEM experiment at LHC,” *PoS*, vol. ICRC2019, p. 207, 2021.
- [31] The CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector,” *Journal of Instrumentation*, vol. 12, pp. P10003–P10003, oct 2017.
- [32] The CMS Collaboration, “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}=13$ TeV,” *Journal of Instrumentation*, vol. 13, pp. P06015–P06015, jun 2018.
- [33] The CMS Collaboration, “Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC,” *Journal of Instrumentation*, vol. 16, p. P05014, may 2021.
- [34] The CMS Collaboration, “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV,” *Journal of Instrumentation*, vol. 12, pp. P02014–P02014, feb 2017.
- [35] The CMS Collaboration, “Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector,” *Journal of Instrumentation*, vol. 14, pp. P07004–P07004, jul 2019.

- [36] D. Bertolini, P. Harris, M. Low, and N. Tran, “Pileup Per Particle Identification,” *JHEP*, vol. 10, p. 059, 2014.
- [37] The CSM Collaboration, “Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS,” *Eur. Phys. J. C*, vol. 81, p. 800, 2021.
- [38] “CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV,” CMS Physics Analysis Summary CMS-PAS-LUM-17-004, 2017.
- [39] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, “An introduction to PYTHIA 8.2,” *Comput. Phys. Commun.*, vol. 191, p. 159, 2015.
- [40] V. Khachatryan *et al.*, “Event generator tunes obtained from underlying event and multiparton scattering measurements,” *Eur. Phys. J. C*, vol. 76, p. 155, 2016.
- [41] A. M. Sirunyan *et al.*, “Extraction and validation of a new set of CMS PYTHIA8 tunes from underlying-event measurements,” *Eur. Phys. J. C*, vol. 80, p. 4, 2020.
- [42] P. Nason, “A new method for combining NLO QCD with shower Monte Carlo algorithms,” *JHEP*, vol. 11, p. 040, 2004.
- [43] G. Luisoni, P. Nason, C. Oleari, and F. Tramontano, “ $\text{HW}^\pm/\text{HZ} + 0$ and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO,” *JHEP*, vol. 10, p. 083, 2013.
- [44] R. Frederix and K. Hamilton, “Extending the MINLO method,” *JHEP*, vol. 05, p. 042, 2016.
- [45] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H. S. Shao, T. Stelzer, P. Torrielli, and M. Zaro, “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations,” *JHEP*, vol. 07, p. 079, 2014.
- [46] J. M. Campbell, R. K. Ellis, and C. Williams, “Vector boson pair production at the LHC,” *JHEP*, vol. 07, p. 018, 2011.
- [47] S. Agostinelli *et al.*, “GEANT4 — a simulation toolkit,” *Nucl. Instrum. Meth. A*, vol. 506, p. 250, 2003.

- [48] A. Sirunyan *et al.*, “An embedding technique to determine $\tau\tau$ backgrounds in proton-proton collision data,” *JINST*, vol. 14, p. P06032, 2019.
- [49] A. M. Sirunyan *et al.*, “Measurement of the Higgs boson production rate in association with top quarks in final states with electrons, muons, and hadronically decaying tau leptons at $\sqrt{s} = 13$ TeV,” *Eur. Phys. J. C*, vol. 81, p. 378, 2021.
- [50] K. Hamilton, P. Nason, E. Re, and G. Zanderighi, “NNLOPS simulation of Higgs boson production,” *JHEP*, vol. 10, p. 222, 2013.
- [51] K. Hamilton, P. Nason, and G. Zanderighi, “Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator,” *JHEP*, vol. 05, p. 140, 2015.
- [52] A. M. Sirunyan *et al.*, “Identification of heavy-flavour jets with the CMS detector in collisions at 13 TeV,” *JINST*, vol. 13, p. P05011, 2018.
- [53] “CMS Phase 1 heavy flavour identification performance and developments,” CMS Detector Performance Summary CMS-DP-2020-019, 2017.
- [54] R. Ceccarelli, *Development of Radiation Resistant Pixel Detectors for the Luminosity Frontier and Measurement of the Higgs Boson Production via Vector Boson Fusion with the CMS Experiment at the LHC*. 2022.
- [55] A. M. Sirunyan *et al.*, “Measurement of the inclusive and differential Higgs boson production cross sections in the leptonic WW decay mode at $\sqrt{s} = 13$ TeV,” *JHEP*, vol. 03, p. 003, 2021.
- [56] G. Cowan, *Statistical Data Analysis*. Oxford science publications, Clarendon Press, 1998.
- [57] R. D. Cousins, “Generalization of chisquare goodness-of-fit test for binned data using saturated models , with application to histograms,” 2013.
- [58] J. Butterworth, S. Carrazza, A. Cooper-Sarkar, A. D. Roeck, J. Feltesse, S. Forte, J. Gao, S. Glazov, J. Huston, Z. Kassabov, R. McNulty, A. Morsch, P. Nadolsky, V. Radescu, J. Rojo, and R. Thorne, “Pdf4lhc recommendations for lhc run ii,” *Journal of Physics G: Nuclear and Particle Physics*, vol. 43, p. 023001, jan 2016.
- [59] R. Barlow and C. Beeston, “Fitting using finite Monte Carlo samples,” *Comput. Phys. Commun.*, vol. 77, p. 219, 1993.

- 2727 [60] F. Caola, K. Melnikov, R. Röntsch, and L. Tancredi, “QCD corrections
2728 to W^+W^- production through gluon fusion,” *Phys. Lett.*, vol. B754,
2729 pp. 275–280, 2016.
- 2730 [61] LHC Higgs Cross Section Working Group, “Handbook of LHC Higgs
2731 cross sections: 3. Higgs properties: Report of the LHC Higgs Cross
2732 Section Working Group,” CERN Report CERN-2013-004, 2013.
- 2733 [62] The CMS Collaboration, “Measurements of properties of the Higgs boson
2734 decaying to a W boson pair in pp collisions at $\sqrt{s} = 13$ TeV,” *Physics*
2735 *Letters B*, vol. 791, pp. 96–129, 2019.
- 2736 [63] S. Chatrchyan *et al.*, “Measurement of Higgs boson production and
2737 properties in the WW decay channel with leptonic final states,” *JHEP*,
2738 vol. 01, p. 096, 2014.
- 2739 [64] “Measurements of gluon fusion and vector-boson-fusion production of
2740 the Higgs boson in $H \rightarrow WW^* \rightarrow e\nu\mu\nu$ decays using pp collisions at
2741 $\sqrt{s} = 13$ TeV with the ATLAS detector,” tech. rep., CERN, Geneva,
2742 2021.
- 2743 [65] G. Aad *et al.*, “Measurements of the Higgs boson inclusive and dif-
2744 ferential fiducial cross-sections in the diphoton decay channel with pp
2745 collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector,” *JHEP*, vol. 08,
2746 p. 027, 2022.
- 2747 [66] A. M. Sirunyan *et al.*, “Measurements of Higgs boson production cross
2748 sections and couplings in the diphoton decay channel at $\sqrt{s} = 13$ TeV,”
2749 *JHEP*, vol. 07, p. 027, 2021.
- 2750 [67] G. Aad *et al.*, “Measurements of the Higgs boson inclusive and differ-
2751 ential fiducial cross sections in the 4ℓ decay channel at $\sqrt{s} = 13$ TeV,”
2752 *Eur. Phys. J. C*, vol. 80, no. 10, p. 942, 2020.
- 2753 [68] A. M. Sirunyan *et al.*, “Measurements of production cross sections of the
2754 Higgs boson in the four-lepton final state in proton–proton collisions at
2755 $\sqrt{s} = 13$ TeV,” *Eur. Phys. J. C*, vol. 81, no. 6, p. 488, 2021.
- 2756 [69] “A portrait of the Higgs boson by the CMS experiment ten years after
2757 the discovery,” *Nature*, vol. 607, no. 7917, pp. 60–68, 2022.