
Evidence for Higgs boson decay to muons with the CMS experiment

Dissertation zur Erlangung des Doktorgrades
an der Fakultät für
Mathematik, Informatik und Naturwissenschaften
Fachbereich Physik
an der Universität Hamburg

vorgelegt von
Jan Oliver Rieger

Hamburg
2021

Gutachter/innen der Dissertation:

Prof. Dr. Peter Schleper
Prof. Dr. Elisabetta Gallo

Zusammensetzung der Prüfungskommission:

Prof. Dr. Elisabetta Gallo
Prof. Dr. Gudrid Moortgat-Pick
Prof. Dr. Arwen Pearson
Prof. Dr. Peter Schleper
Prof. Dr. Kerstin Tackmann

Vorsitzende der Prüfungskommission:

Prof. Dr. Arwen Pearson

Datum der Disputation:

30.06.2021

Vorsitzender Fach-Promotionsausschuss:

Prof. Dr. Günter Sigl

Leiter des Fachbereichs Physik:

Prof. Dr. Wolfgang Hansen

Dekan der Fakultät MIN:

Prof. Dr. Heinrich Graener

Abstract

This thesis presents the first search for standard model (SM) Higgs boson decays to muons produced in association with a top quark-antiquark pair, $t\bar{t}H$ ($H \rightarrow \mu^+\mu^-$). The analysis is based on proton-proton collision data recorded by the CMS experiment at a center-of-mass energy of 13 TeV corresponding to an integrated luminosity of 137 fb^{-1} .

Signal events contain two oppositely charged muons and the decay products of the top quark-antiquark pair. Therefore, events are selected based on the presence of one or more b quark jets. To fully exploit the topology of the top quark decay, two event categories are defined which aim to identify either leptonic or hadronic top quark decays. The leptonic category targets events with one or two charged leptons in addition to the muon pair from the Higgs boson decay. In the hadronic category, the top quark decay to three resolved jets is reconstructed using a constrained kinematic fit combined with a multivariate discriminant based on a boosted decision tree (BDT). Both event categories are individually optimized using BDT-based multivariate discriminants which include several sensitive kinematic observables to distinguish the $t\bar{t}H$ ($H \rightarrow \mu^+\mu^-$) signal from SM backgrounds. The analyzed events are classified into subcategories according to the BDT output score. A simultaneous maximum-likelihood fit is performed to the observed dimuon mass distributions in all subcategories to quantify the presence of a potential signal. The observed (expected) significance over the SM background prediction for $t\bar{t}H$ ($H \rightarrow \mu^+\mu^-$) is 1.2σ (0.5σ). The corresponding best-fit value of the signal strength, defined as the ratio of the observed signal rate to that predicted by the SM, is $\hat{\mu}_{t\bar{t}H(H \rightarrow \mu^+\mu^-)} = 2.32^{+2.27}_{-1.95}$.

This result contributes to a combination of four complementary Higgs boson analyses. When combined with the analyses targeting the Higgs boson production via gluon fusion, via vector boson fusion, and in association with a vector boson, an excess of events over the background prediction is observed in data with a significance of 3.0σ , where the expectation for the SM Higgs boson is 2.5σ . The combined signal strength, relative to the SM expectation, is $\hat{\mu}_{H \rightarrow \mu^+\mu^-} = 1.19^{+0.44}_{-0.42}$. This result constitutes the first evidence for the Higgs boson decay to second generation fermions and provides the most precise measurement of the Higgs boson coupling to muons to date. The results are compared to those of a recently published analysis by the ATLAS Collaboration.

Kurzfassung

Die vorliegende Dissertation präsentiert die erste Suche nach dem Higgs-Boson des Standardmodells (SMs) welches in Assoziation mit einem Top-Quark-Antiquark-Paar produziert wird, und in Myonen zerfällt, $t\bar{t}H$ ($H \rightarrow \mu^+\mu^-$). Die Analyse basiert auf Daten, die mit dem CMS Experiment in Proton-Proton-Kollisionen bei einer Schwerpunktsenergie von 13 TeV aufgezeichnet wurden und einer integrierten Luminosität von 137 fb^{-1} entsprechen.

Signalereignisse enthalten zwei entgegengesetzt geladene Myonen und die Zerfallsprodukte des Top-Quark-Antiquark-Paares. Deshalb werden Ereignisse selektiert, die mindestens einen b-Quark-Jet aufweisen. Um die Topologie des Top-Quark-Zerfalls bestmöglich zu nutzen, werden zwei Ereigniskategorien definiert, die darauf optimiert sind, entweder leptonische oder hadronische Top-Quark-Zerfälle zu identifizieren. Die leptonische Kategorie zielt auf Ereignisse ab, die zusätzlich zu dem Myonen-Paar aus dem Zerfall des Higgs-Bosons ein oder zwei geladene Leptonen aufweisen. In der hadronischen Kategorie wird der Top-Quark-Zerfall in drei Jets mit Hilfe eines kinematischen Fits und einer multivariaten Analyse basierend auf einem Boosted Decision Tree (BDT) rekonstruiert. Beide Ereigniskategorien werden individuell mit BDT-basierten multivariaten Analysen optimiert, die mehrere sensitive kinematische Observablen enthalten, um Signal von Untergrundereignissen zu unterscheiden. Die analysierten Ereignisse werden entsprechend der Ausgabewerte der BDTs in Unterkategorien eingeordnet. Ein Maximum-Likelihood-Fit an die beobachteten Verteilungen der invarianten Masse des Myon-Paares wird in allen Unterkategorien durchgeführt, um die Stärke eines potenziellen Signals zu quantifizieren.

Die beobachtete (erwartete) Signifikanz für $t\bar{t}H$ ($H \rightarrow \mu^+\mu^-$) über der SM-Untergrundvorhersage beträgt $1,2\sigma$ ($0,5\sigma$). Die entsprechende, gemessene Signalstärke, definiert als das Verhältnis der beobachteten Signalrate zur im SM erwarteten Rate, ist $\hat{\mu}_{t\bar{t}H(H \rightarrow \mu^+\mu^-)} = 2,32_{-1,95}^{+2,27}$. Dieses Ergebnis trägt zu einer Kombination von vier komplementären Higgs-Boson Analysen bei. Kombiniert mit den Analysen, die auf die Produktion eines Higgs-Bosons in Gluon-Fusion, in Vektor-Boson-Fusion und in assoziierter Produktion mit einem Vektor-Boson optimiert wurden, wird ein Überschuss an Ereignissen in Daten mit einer Signifikanz von $3,0\sigma$ beobachtet, wobei die Erwartung für das SM Higgs-Boson $2,5\sigma$ beträgt. Die kombinierte Signalstärke, relativ zur SM Erwartung, ist $\hat{\mu}_{H \rightarrow \mu^+\mu^-} = 1,19_{-0,42}^{+0,44}$. Dieses Ergebnis ist der erste Nachweis für den Zerfall des Higgs-Bosons in Fermionen der zweiten Generation und zugleich die bisher präziseste Messung der Kopplung des Higgs-Bosons an Myonen. Diese Ergebnisse werden mit denen einer kürzlich veröffentlichten Analyse der ATLAS-Kollaboration verglichen.

List of publications

The work presented in this thesis was published as a part of a larger paper by the CMS Collaboration combining four Higgs boson analyses. I contributed the $t\bar{t}H$ ($H \rightarrow \mu^+\mu^-$) analysis described in Chapter 5, while others provided the ggH, VBF, VH analyses, and the combination of results presented in Chapter 6. Furthermore, I optimized a method to recover final state radiation photons which was used by all other analyses. The method is documented in Section 5.5.1.

Publications

- CMS Collaboration, "Evidence of Higgs boson decay to a pair of muons", *JHEP* 01 (2021) 148, doi:10.1007/JHEP01(2021)148
- CMS Collaboration, "Measurement of Higgs boson decay to a pair of muons in proton-proton collisions at $\sqrt{s} = 13$ TeV", CMS-PAS-HIG-19-006 (2020), <https://cds.cern.ch/record/2725423>

Internal documents

- CMS Collaboration, "Analysis of Higgs decays to muons with the full Run2 dataset", CMS Internal Analysis Note CMS AN-2019/124 (2020)

Explanatory notes

Figures and tables which are exactly those contained in the publication are referenced in the caption, and additionally labeled "CMS". All other figures which use CMS data or CMS resources are labeled either "CMS private work" or "CMS simulation • private work".

Contents

Abstract	i
Kurzfassung	iii
List of publications	v
Introduction	1
1 Physics of the Higgs boson	5
1.1 The Standard Model of particle physics	5
1.1.1 Particles & forces	5
1.1.2 Quantum electrodynamics	7
1.1.3 Quantum chromodynamics	8
1.1.4 Weak interactions	9
1.1.5 Electroweak unification	10
1.2 The Brout-Englert-Higgs mechanism	11
1.3 Shortcomings of the Standard Model	12
1.4 Experimental profile of the Higgs boson	13
1.4.1 The Higgs boson at the LHC	13
1.4.2 Previous search for $H \rightarrow \mu^+\mu^-$ with the CMS experiment	15
1.4.3 Recent results on $H \rightarrow \mu^+\mu^-$ with the ATLAS detector	17
2 Experimental setup	21
2.1 The Large Hadron Collider	21
2.2 The CMS experiment	24
2.2.1 Coordinate system & kinematic quantities	26
2.2.2 Tracking system	27
2.2.3 Calorimetry	29
2.2.4 Muon system	31
2.2.5 Magnet	33
2.2.6 Trigger system	33

3	Event simulation	35
3.1	Event generation	35
3.2	Simulated events	38
4	Event reconstruction	41
4.1	Detector signatures	41
4.2	Event reconstruction with the particle-flow algorithm	42
4.3	Selection of physics objects	44
4.3.1	Primary vertex and beam spot	44
4.3.2	Muons	45
4.3.3	Electrons	46
4.3.4	Lepton isolation	48
4.3.5	Hadronic jets	48
4.3.6	B-tagging	52
4.3.7	Quark-gluon discriminant	53
4.3.8	Missing transverse momentum	53
5	Higgs boson decays to muons produced in association with top quarks	55
5.1	General $H \rightarrow \mu^+\mu^-$ analysis strategy	56
5.2	Event topology in the $t\bar{t}H$ category	58
5.3	Trigger & dataset	61
5.4	Efficiency corrections & event reweighting	63
5.4.1	Lepton efficiencies	63
5.4.2	L1 ECAL prefiring correction	65
5.4.3	Pileup reweighting	66
5.4.4	B jet discriminator & quark-gluon likelihood reshaping	66
5.4.5	Z boson p_T correction	67
5.4.6	Muon momentum scale and resolution	72
5.4.7	Per-event Higgs boson mass resolution	76
5.5	Improvements to dimuon mass reconstruction	79
5.5.1	Final state photon recovery	79
5.5.2	Geometric track fit	86
5.6	Multivariate lepton identification	90
5.7	Resolved hadronic top quark tagger	94
5.7.1	Kinematic top quark fit	94
5.7.2	BDT-based top quark identification	96
5.8	$t\bar{t}H$ event selection	102
5.8.1	$t\bar{t}H$ leptonic selection	104
5.8.2	$t\bar{t}H$ hadronic selection	106
5.9	Multivariate analysis optimization of the $t\bar{t}H$ channel	108
5.9.1	Kinematic input observables to the BDT classifiers	108
5.9.2	Training of the BDT classifier	117
5.9.3	BDT performance & validation in the $t\bar{t}H$ leptonic channel	120
5.9.4	BDT performance & validation in the $t\bar{t}H$ hadronic channel	123

5.10	Event categorization	126
5.11	Signal extraction	127
5.11.1	Analytical signal and background model	128
5.11.2	Systematic uncertainties	134
5.12	Results for the $t\bar{t}H$ channel	137
6	Combined analysis of Higgs boson decays to muons	143
6.1	Higgs boson production categories	143
6.1.1	The VBF production category	144
6.1.2	The ggH production category	148
6.1.3	The VH production category	151
6.2	Evidence for Higgs boson decay to muons	154
6.3	Scientific impact	158
6.4	Comparison to previous CMS analysis	159
6.5	Comparison to ATLAS analysis	160
6.6	Future prospects	163
7	Conclusions	165
	Appendix A Method to calibrate event-by-event mass uncertainties	169
	Appendix B Correlation model of systematic uncertainties in the combined fit	173
	Bibliography	175

Introduction

The field of elementary particle physics is continuously evolving to elucidate the laws of nature. Experimental progress in the past decades has allowed deep insights into the structure of the fundamental constituents of matter and their interactions. This has been accomplished based on the predictions made within the *Standard Model* (SM) of particle physics [1–4]. Developed in the second half of the last century, the SM describes all known elementary particles and three out of four fundamental forces combining the theories of electroweak and strong interactions. So far it has predicted a variety of new elementary particles and successfully explained experimental results over a large energy range spanning from a few eV up to the TeV scale.

Understanding the mechanism that spontaneously breaks electroweak symmetry and generates the masses of the SM particles remains one of the main objectives in particle physics. The *Brout-Englert-Higgs* (BEH) mechanism [5–8] introduces a complex doublet scalar field whose neutral component acquires a non-zero vacuum expectation value after *electroweak symmetry breaking* (EWSB). This results in massive W^\pm and Z gauge bosons, while the photon remains massless. The fourth component of the complex doublet becomes a new, fundamental scalar particle, the Higgs boson. The mechanism of EWSB, therefore, provides an explanation for the observed masses of the W^\pm and Z bosons [9–12], whereas the masses of all fermions are free parameters of the theory. Fermions acquire a mass through Yukawa interactions with the Higgs field where the coupling strength is proportional to the fermion mass. Apart from generating all particle masses, the BEH mechanism preserves the perturbative unitarity of the SM. Without the Higgs boson, the amplitude of longitudinal vector boson scattering would diverge at high energies [13–15]. Thus, the importance of the BEH mechanism for the consistency of the SM cannot be emphasized enough.

The search for the postulated Higgs boson has been a leading motivation to build the *Large Hadron Collider* (LHC), located at the *European Organization for Nuclear Research* (CERN). After the experimental observation of the Higgs boson by the ATLAS [16] and CMS [17, 18] Collaborations in 2012, several measurements of its properties and interactions with other SM particles have been performed [19–28]. The first observation of a Yukawa coupling has been accomplished by a combined result of the ATLAS and CMS Collaborations searching for Higgs boson decays to τ leptons in 2016 [29]. Two years later, the observation of the associated production of a Higgs boson with a top quark-antiquark pair has directly confirmed the Yukawa coupling to top quarks [30]. With the observed Higgs boson decay to a bottom quark-antiquark pair, the last remaining Yukawa coupling of the third generation fermions has been validated [31]. The observed decays of the Higgs boson to the electroweak gauge bosons and

the charged fermions of the third generation have been found to be consistent with the SM expectation [32]. These results constitute that the major source of EWSB is indeed the postulated Higgs boson.

Despite the agreement between experiments and theoretical predictions, the SM is known to be an incomplete theory. For instance, the SM fails to explain the observed pattern of fermion masses, or a mechanism which protects the Higgs boson mass from large quantum corrections at the Planck scale, known as the hierarchy problem [33, 34]. These and other shortcomings of the SM have led to the development of new theories such as supersymmetry [35, 36], or composite Higgs models [37–40], which can alter the properties of the Higgs boson. Thus, differences in the Higgs boson coupling structure, with respect to the SM, might offer insights into new physics beyond the SM. Furthermore, the question of whether the Higgs mechanism is responsible at all for the masses of the first and second generation fermions is still to be answered. While the Yukawa coupling to first generation fermions is experimentally not accessible at the LHC, the analysis of Higgs boson decays to muons provides an opportunity to probe the coupling of the Higgs boson to second generation fermions.

This thesis reports the first search for Higgs boson decays to muons produced in association with a top quark-antiquark pair, $t\bar{t}H$ ($H \rightarrow \mu^+\mu^-$). For this search, proton-proton collision data recorded by the CMS experiment between 2016 and 2018 at a center-of-mass energy of 13 TeV corresponding to an integrated luminosity of 137 fb^{-1} are analyzed. The analysis is designed to complement three further Higgs boson analyses targeting the production of a Higgs boson via *gluon fusion* (ggH), via *vector boson fusion* (VBF), and in *association with a vector boson* (VH). Furthermore, a method to recover final state radiation photons is optimized to improve the resolution of the reconstructed Higgs boson candidate.

The considered final state of the search for $t\bar{t}H$ ($H \rightarrow \mu^+\mu^-$) contains two oppositely charged muons from the decay of the Higgs boson, and the decay products of the top quark-antiquark pair. Top quarks decay almost exclusively into a b quark and a W boson, where the W boson decays subsequently into a lepton and a neutrino, or into two quarks. Thus, events are selected based on the presence of one or more b quark jets. Two event categories are defined targeting either leptonic or hadronic top quark decays. Events with one or two charged leptons in the final state in addition to the muon pair from the Higgs boson decay are assigned to the leptonic category, while events with at least three additional jets are grouped in the hadronic category. To identify the hadronic decay of a top quark, a dedicated algorithm is developed combining a constrained kinematic fit with a multivariate discriminant based on a *boosted decision tree* (BDT). Both, the leptonic and the hadronic event categories are individually optimized using BDT-based multivariate discriminants. The analyzed events are classified into subcategories according to the BDT output score. A simultaneous maximum-likelihood fit is performed to the observed dimuon mass distributions in all subcategories to quantify the presence of a potential signal. The results of the $t\bar{t}H$ ($H \rightarrow \mu^+\mu^-$) analysis as well as the combined results including the ggH, VBF, and VH analyses are discussed and compared to previous results by the CMS Collaboration [41] and to recent results by the ATLAS Collaboration [42].

The previous search for $H \rightarrow \mu^+\mu^-$ decays performed by the CMS Collaboration analyzed proton-proton collision data collected at a center-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 35.9fb^{-1} [41]. In combination with data recorded during LHC Run 1 at center-of-mass energies of 7 and 8 TeV, the observed (expected) upper limit at 95% confidence level on the production cross section times branching fraction for a Higgs boson with $m_H = 125.09\text{GeV}$ [43] was set to 2.9 (2.2) times the SM prediction. The measured signal strength, relative to the SM expectation, was $\hat{\mu}_{H \rightarrow \mu^+\mu^-} = 1.0 \pm 1.0$.

In parallel to the analysis presented in this thesis, the ATLAS Collaboration has published a search for $H \rightarrow \mu^+\mu^-$ decays [42] based on proton-proton collision data recorded at a center-of-mass energy of 13 TeV corresponding to an integrated luminosity similar to the one collected with the CMS experiment.

This thesis is organized as follows: In Chapter 1, relevant aspects of the SM are discussed with an emphasis on electroweak symmetry breaking and its consequences. This is complemented by a phenomenological description of the Higgs boson production at the LHC and its decay channels followed by a brief overview of recent experimental results. Chapter 2 describes the LHC accelerator and the CMS detector. A brief review of the event simulation is given in Chapter 3. The reconstruction and selection of physics objects is detailed in Chapter 4. The search for Higgs boson decays to muons produced in association with a top quark-antiquark pair is presented in Chapter 5. In Chapter 6, the combined results are discussed and compared to the previous CMS result and the latest result by the ATLAS Collaboration. Chapter 7 concludes this thesis.

Physics of the Higgs boson

The *Standard Model* (SM) of particle physics provides the theoretical framework for the Higgs boson analysis presented in this thesis. The elementary particles and the fundamental forces described within the SM are introduced along with the underlying theoretical concepts in Section 1.1. The mechanism of *electroweak symmetry breaking* (EWSB), responsible for the generation of particle masses and the origin of the Higgs boson is described in Section 1.2. Shortcomings of the SM are briefly discussed in Section 1.3. Resulting phenomenological aspects for the experiment as well as previous experimental results on the Higgs boson decays to muons are reported in Section 1.4. The first sections of this chapter which introduce the SM and EWSB are based on Refs. [44–49].

1.1 The Standard Model of particle physics

The SM of particle physics [1–4] describes all currently known elementary particles and the fundamental forces which govern them. It is formulated as a *quantum field theory* (QFT) based on the $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge group. In QFT all elementary particles are interpreted as excitations of their underlying quantum fields. One key element of the SM is the local gauge invariance of the Lagrangian density \mathcal{L} under particular symmetry group transformations. For each force, a symmetry group as well as conserved charges can be identified, as described in the following.

1.1.1 Particles & forces

The particle content of the SM is illustrated in Fig. 1.1. Particles contained in the SM are classified as *fermions* carrying half-integer spin, and *bosons* carrying integer spin.

Fermions are the fundamental constituents of matter. Thus, they are indivisible, and not composed of other particles. They can be further categorized into *leptons* and *quarks*, which are arranged in three generations. Each generation contains two quarks with electric charge of $+2/3$ and $-1/3$, and two leptons with electric charge of -1 and 0 . Neutral leptons are referred to as *neutrinos*. The first generation is composed of the electron e and the electron neutrino ν_e as well as the up u and down quark d . These particles constitute ordinary matter. The electron is known for orbiting the atomic nuclei, which is composed of three valence quarks, two up quarks and one down quark. The second and third generation contain heavier copies of the leptons and quarks of the first generation. The leptons of the second and third generation are the muon μ and the τ lepton as well as their corresponding neutrinos ν_μ and ν_τ .

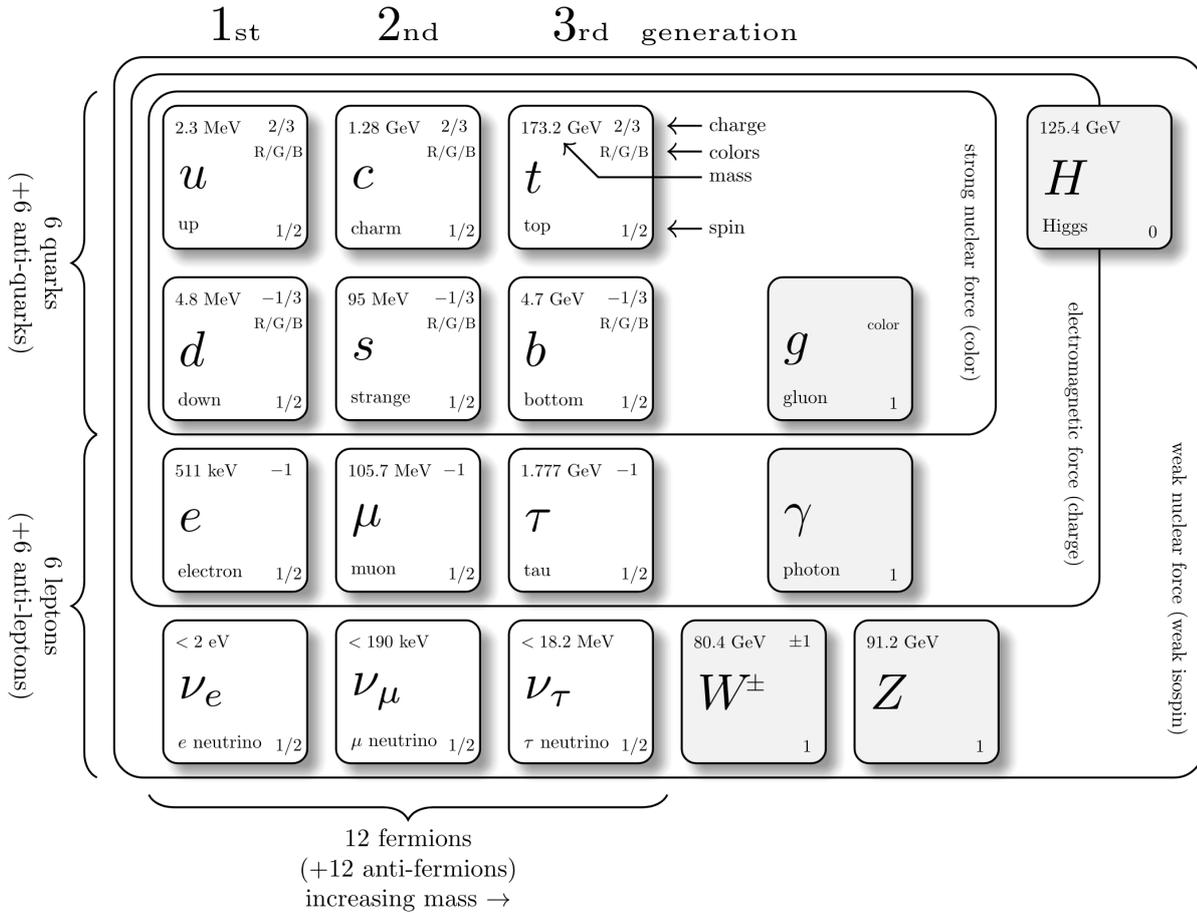


Figure 1.1: Illustration of the standard model of particle physics. The particles are presented along with their charge, color, mass and spin. Particle interactions mediated by the force carriers Z, W^\pm, γ, g are indicated by the different boxes. Modified from Ref. [50].

The quarks of the subsequent generations are the charm c and strange quark s in the second generation and the top t and bottom quark b in the third generation. For each fermion in the SM, there is a corresponding antiparticle with same quantum numbers except for an opposite-sign charge.

A consequence of local gauge invariance is the existence of gauge bosons, which are also referred to as *force carriers* as they mediate the interactions between particles. Three fundamental forces are described by the SM, namely the electromagnetic force, the weak force, and the strong force. The fourth fundamental force, Gravity, has not been formulated according to the principles of quantum field theory and is therefore not included in the SM. The interaction of particles which carry electric charge is mediated by the massless photon γ . The massive W^\pm and Z bosons are the mediators of the weak force, which is responsible for certain particle decays, for example, the β -decay of a neutron inside the atomic nucleus. The strong force binds neutrons and protons to create atomic nuclei and is mediated by eight massless gluons.

The only fundamental scalar particle in the SM is the Higgs boson. In contrast to the gauge bosons, it does not result from local gauge invariance. Instead, the Higgs boson naturally arises from the mechanism of EWSB, which explains how the massive gauge bosons acquire their mass. EWSB is also the origin of the fermion masses, generated via Yukawa interactions with the Higgs field.

1.1.2 Quantum electrodynamics

The electromagnetic interaction is described by *quantum electrodynamics* (QED), which explains interactions between electrically charged particles by means of the exchange of a photon. The Lagrange density in QED is invariant under local gauge transformations of the abelian $U(1)_{\text{em}}$ group. The fermion field ψ transforms as

$$\psi \rightarrow \psi' = \psi e^{iq\alpha(x)}, \quad \bar{\psi} \rightarrow \bar{\psi}' = \bar{\psi} e^{-iq\alpha(x)} \quad (1.1)$$

with a charge q and a real phase $\alpha(x) = \alpha(t, x_1, x_2, x_3)$. Since the Lagrange density for the Dirac equation contains the derivative $\partial_\mu \psi$, which is not invariant under such local gauge transformations, the derivative is redefined. The *covariant derivative*

$$D_\mu = \partial_\mu + iqA_\mu(x) \quad (1.2)$$

introduces a new vector field $A_\mu(x)$, which can be interpreted as the field of the massless photon. The covariant derivative and the vector field A_μ transform as

$$D_\mu \psi \rightarrow D'_\mu \psi' = e^{iq\alpha(x)} D_\mu \psi, \quad A_\mu(x) \rightarrow A'_\mu(x) = A_\mu(x) - \partial_\mu \alpha(x). \quad (1.3)$$

Thus, the overall QED Lagrangian can be written as

$$\mathcal{L}_{\text{QED}} = \underbrace{i\bar{\psi}\gamma^\mu\partial_\mu\psi}_{\text{kinetic}} - \underbrace{m\bar{\psi}\psi}_{\text{mass}} - \underbrace{q\bar{\psi}\gamma^\mu\psi A_\mu}_{\text{interaction}} - \underbrace{\frac{1}{4}F^{\mu\nu}F_{\mu\nu}}_{\text{free photon}}. \quad (1.4)$$

The first and the second term correspond to the Lagrangian of a massive fermion field. The third term follows from the covariant derivative and describes the interaction between a fermion and the photon. The coupling strength of this interaction is proportional to the electric charge q . The kinetic term for the free vector field A_μ is expressed by means of the field strength tensor

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad (1.5)$$

and corresponds to the fourth term in Eq. (1.4). The unification of the electromagnetic and weak interaction is explained in Section 1.1.5.

1.1.3 Quantum chromodynamics

The interaction between particles which carry *color* charge is described by *quantum chromodynamics* (QCD). The underlying gauge group of this theory is the non-abelian $SU(3)_C$, where the subscript C refers to color. Quarks (gluons) are represented in terms of color triplets (octets). Quarks carry *red*, *blue*, or *green* color charge, while gluons carry a combination of two color charges in a superposition of states. The eight generators of this group can be represented by hermitian 3×3 matrices denoted as *Gell-Mann* matrices λ_a . Following the same arguments as in QED, the assumption of local gauge invariance under $SU(3)_C$ transformations requires a covariant derivative, which introduces eight gauge vector fields G_μ according to the eight generators of the $SU(3)_C$ group. These are identified by eight color charged gluons. The coupling strength of the strong interaction is parameterized by g_s , usually specified in terms of $\alpha_s = g_s^2/4\pi$. The full Lagrangian of QCD encodes the structure of strong interactions:

$$\mathcal{L}_{\text{QCD}} = \underbrace{i\bar{\psi}\gamma^\mu\partial_\mu\psi}_{\text{kinetic}} - \underbrace{m\bar{\psi}\psi}_{\text{mass}} - \underbrace{g_s\bar{\psi}\gamma^\mu\frac{\lambda_a}{2}\psi G_{a,\mu}}_{\text{interaction}} - \underbrace{\frac{1}{4}G_a^{\mu\nu}G_{a,\mu\nu}}_{\text{gluon kin. \& self interaction}}. \quad (1.6)$$

The first three terms are analogue to those described for QED. However, the last term with the tensor

$$G_{\mu\nu}^a = \partial_\mu G_\nu^a - \partial_\nu G_\mu^a - g_s f_{abc} G_\mu^b G_\nu^c \quad (1.7)$$

is a direct consequence of the non-abelian nature of $SU(3)_C$. The structure constants f_{abc} of $SU(3)_C$ are defined by the commutation relation $[\lambda_a, \lambda_b] = i f_{abc} \lambda_c$. The cubic $\sim g_s(G_\mu^a)^3$ and quartic terms $\sim g_s(G_\mu^a)^4$ introduce self interactions between the gauge fields, which explains why gluons carry color charge.

An interesting physical ramification of QCD is that at short distances the coupling strength of the strong interaction is small, and thus quarks or gluons can propagate freely. This property is referred to as *asymptotic freedom*. However, the coupling strength increases at large distances. Consequently, also the field energy between two colored particles increases with distance. If the distance becomes sufficiently large, a new quark-antiquark pair appears. Therefore, quarks will always form color-neutral bound states of either a quark-antiquark pair (*mesons*) or a three quark system (*baryons*). Together, mesons and baryons are referred to as *hadrons*. As a result of *color confinement*, quarks and gluons can not be observed as isolated particles in experiments. Instead they form a collimated spray of hadrons, referred to as *jet*.

1.1.4 Weak interactions

All fermions described within the SM take part in weak interactions. The corresponding gauge theory is based on the $SU(2)_L$ group. The subscript L denotes the *chirality* of a fermion field. In the famous experiment by C. S. Wu et al. [51] it had been observed that weak interactions violate parity. Consequently, the theory of the weak force has to differentiate between the left-handed and right-handed fermion components. Formally, the parity violation is therefore explained by weak interactions only coupling to left-handed particles and right-handed antiparticles. The charge conserved by the weak interaction is the third component of the *weak isospin* I_3 . Fermion fields are represented by left-handed $SU(2)_L$ doublets ψ_L with $I_3 = \pm 1/2$, and right-handed singlets ψ_R with $I_3 = 0$. They are grouped in three generations:

$$\begin{aligned}
 l_L &= \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L & l_R &= e_R, \mu_R, \tau_R, \nu_{eR}, \nu_{\mu R}, \nu_{\tau R} \\
 q_L &= \begin{pmatrix} u \\ d \end{pmatrix}_L, \begin{pmatrix} c \\ s \end{pmatrix}_L, \begin{pmatrix} t \\ b \end{pmatrix}_L & q_R &= d_R, s_R, b_R, u_R, c_R, t_R.
 \end{aligned}$$

Local gauge invariance of the Lagrangian under $SU(2)_L$ symmetry transformations requires the existence of new vector fields, which mediate the weak interaction. The generators of the non-abelian $SU(2)_L$ gauge group can be represented as hermitian 2×2 matrices denoted as *pauli* matrices σ_a . As explained before, the number of generators is related to the number of gauge bosons, which in case of $SU(2)_L$ leads to three gauge boson fields $W_\mu^{a=1,2,3}$. The weak Lagrangian can be expressed as

$$\mathcal{L}_{\text{weak}} = \underbrace{i\bar{\psi}_L \gamma^\mu \partial_\mu \psi_L}_{\text{kinetic}} - \underbrace{g\bar{\psi}_L (\gamma^\mu \frac{\sigma_a}{2} W_\mu^a) \psi_L}_{\text{interaction}} - \underbrace{\frac{1}{4} W^{\mu\nu, a} W_{\mu\nu}^a}_{\text{gauge kin. \& self interaction}} \quad (1.8)$$

where g corresponds to the weak coupling constant. Similar to the self interaction described for gluons, the weak gauge bosons also interact among themselves induced by cubic $\sim g(W_\mu^a)^3$ and quartic terms $\sim g(W_\mu^a)^4$ of the Lagrangian. The gauge fields $W_\mu^{1,2,3}$, however, do not correspond to the physical fields of the observed W^\pm and Z bosons. The physical W^\pm bosons are identified with linear combinations of $W_\mu^{1,2}$, while the Z boson corresponds to a mixed state explained in the following Section 1.1.5.

Explicit mass terms are forbidden as they violate local gauge invariance of the Lagrangian. To keep the structure of gauge interactions and still generate the observed masses of the W^\pm and Z bosons [9–12], the mechanism of electroweak symmetry breaking was introduced as explained in Section 1.2.

1.1.5 Electroweak unification

Before the electromagnetic and the weak force can be unified, the gauge group of QED as introduced in Section 1.1.2 is replaced by the $U(1)_Y$ group with *hypercharge* Y , an associated gauge field B_μ , which again follows from local gauge invariance, and the coupling constant g' . The new field B_μ is not the photon field described earlier. Instead the electromagnetic field A_μ and the field Z_μ correspond to a mixed state of the fields W_μ^3 and B_μ , obtained via a rotation by the weak mixing angle θ_W :

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix}. \quad (1.9)$$

The photon field A_μ couples to the electric charge q , which is related to the hypercharge Y using the Gell-Mann-Nishijima formula

$$Q = I_3 - \frac{Y}{2}. \quad (1.10)$$

The Lagrangian of the unified electroweak theory based on the direct product of the gauge groups $SU(2)_L \times U(1)_Y$ reads:

$$\mathcal{L}_{EW} = \sum i\bar{\psi}_L \gamma^\mu D_\mu \psi_L + \sum i\bar{\psi}_R \gamma^\mu D_\mu \psi_R - \frac{1}{4} W_a^{\mu\nu} W_{a\mu\nu} - \frac{1}{4} B^{\mu\nu} B_{\mu\nu} \quad (1.11)$$

where the first (second) sum runs over all left-handed doublets ψ_L (right-handed singlets ψ_R) of quarks and leptons. The covariant derivatives are given by

$$D_L^\mu \psi_L = (\partial^\mu + ig \frac{\sigma_a}{2} W_a^\mu + ig' \frac{Y}{2} B^\mu) \psi_L, \quad (1.12)$$

$$D_R^\mu \psi_R = (\partial^\mu + ig' \frac{Y}{2} B^\mu) \psi_R. \quad (1.13)$$

With Eq. (1.11) all electroweak interactions between fermions, the electroweak gauge bosons, and the photon are quantitatively predicted. Electroweak interactions are grouped in those involving the W^\pm bosons corresponding to charged currents and those involving the Z boson and the photon corresponding to the neutral currents. The charged currents exclusively couple to left-handed particles, while the neutral currents act on both left-handed and right-handed particles. The photon, which mediates the electromagnetic force, is not sensitive to the chirality of the particles. The W^\pm bosons mediate transitions from a charged lepton to the corresponding neutrino and vice-versa, as well as from up-type quarks to down-type quarks and vice-versa. In the quark sector, the possible transitions are described by the *Cabibbo-Kobayashi-Maskawa* (CKM) matrix [52, 53]

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \quad (1.14)$$

The transition probability from flavor i to a different flavor j is proportional to $|V_{ij}|^2$. The diagonal elements of the CKM matrix are close to unity [54]. Thus, transitions within the same generation are favored. In the lepton sector, the *Pontecorvo-Maki-Nakagawa-Sakata* (PMNS) matrix [55] is used to explain neutrino oscillations.

1.2 The Brout-Englert-Higgs mechanism

F. Englert and R. Brout [5], as well as P. Higgs [6, 7] proposed a solution to generate the observed masses of the W^\pm and Z bosons without violating local gauge invariance of the electroweak theory. Their concept is based on spontaneous symmetry breaking and referred to as the *Brout-Englert-Higgs* (BEH) mechanism. For this purpose, a new complex $SU(2)_L$ doublet with hypercharge $Y = 1$ is introduced

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \quad (1.15)$$

where $\phi_i (i = 1, 2, 3, 4)$ are real scalar fields. The interaction with the scalar Higgs field is governed by the Klein-Gordon equation

$$\mathcal{L}_{\text{Higgs}} = (D^\mu \phi)^\dagger (D_\mu \phi) - V(\Phi), \quad (1.16)$$

where D^μ are the covariant derivatives in Section 1.1.5, and $V(\Phi)$ is a potential of the form

$$V(\Phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2, \quad (1.17)$$

with the coefficients μ^2 and λ . Vacuum stability demands $\lambda > 0$. For $\mu^2 > 0$ the ground state remains at zero, while the potential develops a non-zero *vacuum expectation value* v for $\mu^2 < 0$. An infinite number of degenerate ground states of the vacuum is given for

$$\phi^\dagger \phi = \frac{v^2}{2} = \frac{-\mu^2}{2\lambda}. \quad (1.18)$$

Experimentally the vacuum is electrically neutral. Therefore, only the neutral component of ϕ can take non-zero values. The ground state is chosen such that

$$\phi_{\text{vacuum}} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (1.19)$$

The choice of a ground state spontaneously breaks electroweak symmetry. Expanding the Higgs field around the ground state yields excitations interpreted as the real physical Higgs boson:

$$\phi \rightarrow \phi' = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \quad (1.20)$$

Inserting Eq. (1.20) into the Higgs Lagrangian in Eq. (1.16) results in a mass term for the Higgs boson $m_H = \sqrt{2\lambda}v$, as well as different terms describing the interactions between the Higgs boson and the electroweak gauge bosons, and Higgs self-interactions. Comparing the terms of the Lagrangian which describe the interaction between the gauge bosons and the Higgs boson to the terms for the gauge boson masses, one finds, that the couplings of the gauge bosons to the Higgs boson are proportional to the masses of the gauge bosons:

$$m_W = \frac{gv}{2}, \quad m_Z = \frac{v}{2} \sqrt{(g^2 + g'^2)}. \quad (1.21)$$

Three out of four generators of the $SU(2)_L \times U(1)_Y$ gauge group are broken. According to the *Goldstone theorem*, this induces the existence of three massless Goldstone bosons, which become the longitudinal polarizations of the W^\pm and Z bosons. The fourth generator, which is associated to the $U(1)_{\text{em}}$ gauge symmetry, is unbroken and thus the photon remains massless.

Yukawa couplings

From the initial four degrees of freedom of the Higgs doublet, three are absorbed by the electroweak gauge bosons. The remaining fourth degree results in a real, massive Higgs boson. However, fermions remain massless unless a new gauge invariant term is included in the SM Lagrangian. A Dirac mass term $-m\bar{\psi}\psi$ would violate local gauge invariance, since the right-handed and the left-handed components transform differently. Instead, fermions acquire mass through Yukawa interactions, where the left-handed doublets and the right-handed singlets couple to the Higgs field. These new terms $\mathcal{L}_{\text{Yukawa}} \sim \bar{\psi}_L \phi \psi_R$ are $SU(2)_L$ singlets. The Lagrangian of the Yukawa interaction is given by

$$\mathcal{L}_{\text{Yukawa}} = -\hat{\lambda}_d^{ij} \bar{q}_L^i \phi d_R^j - \hat{\lambda}_u^{ij} \hat{q}_L^i \tilde{\phi} u_R^j - \hat{\lambda}_\ell^{ij} \bar{\ell}_L^i \phi e_R^j + h.c., \quad (1.22)$$

where $\tilde{\phi}$ corresponds to the charge conjugate Higgs doublet and $\hat{\lambda}$ corresponds to 3×3 matrices which describe the Yukawa couplings for the up-type, and down-type quarks as well as the charged leptons. It is possible to choose a basis such that the matrices are diagonalized $\hat{\lambda}_f^{ij} \rightarrow \lambda_f^i \delta^{ij}$. After EWSB, the Lagrangian for the first generation becomes

$$\mathcal{L}_{\text{Yukawa}} = -\frac{y_d v}{\sqrt{2}} \bar{d} d - \frac{y_u v}{\sqrt{2}} \bar{u} u - \frac{y_e v}{\sqrt{2}} \bar{e} e - \frac{y_d}{\sqrt{2}} (\bar{d} d) h - \frac{y_u}{\sqrt{2}} (\bar{u} u) h - \frac{y_e}{\sqrt{2}} (\bar{e} e) h, \quad (1.23)$$

where y_i ($i = u, d, e$) are the Yukawa couplings of the first generation. The three terms at the beginning of this Lagrangian are identified as mass terms with $m_f = \frac{y_f v}{\sqrt{2}}$, while the three terms at the end describe interactions between the Higgs boson and the first generation fermions. The Lagrangian for fermions of the second and third generation is defined accordingly.

1.3 Shortcomings of the Standard Model

Although the SM has been remarkably successful in describing experimental data, there are a few indications that the SM is incomplete. A few selected shortcomings of the SM are listed below:

- **Gravity and dark matter** Since no renormalizable quantum theory of gravity has been formulated so far, the gravitational force cannot be included in the SM. Additionally, astrophysical measurements indicate that most of the mass in the universe consists of *dark matter* [56, 57], which also cannot be explained by the SM.
- **Free parameters** The SM includes 26 free parameters, the Higgs boson mass, the vacuum expectation value, three gauge couplings, six lepton masses, six quark masses, one QCD vacuum angle, as well as three mixing angles plus one CP-violating phase for each the quark sector and the lepton sector. All these parameters can be measured experimentally but cannot be derived from fundamental principles. It is worth noting that most of these free parameters are related to the Higgs sector. Although the BEH mechanism clarifies the origin of the particle masses, it does not give insights into the underlying structure explaining the mass hierarchy of the fermions within the SM.

- **Hierarchy problem** The Higgs boson mass is highly sensitive to quantum loop corrections. In a theory which is valid up to the Planck scale, $\Lambda_{\text{Planck}} \sim 10^{19}$ GeV, the corresponding corrections would either lead to a Higgs boson mass in the order of the Planck scale itself or require an extremely fine-tuned cancellation of corrections. The SM is therefore to be considered as an effective theory valid up to a certain energy scale at which additional particles and/or forces are manifested.

Thus, the Higgs boson might be a portal to new physics beyond the SM. There are many possible extensions of the SM which modify the Higgs boson sector, and a variety of these models also directly affect the Higgs boson coupling to muons [58–60]. Probing the Higgs Yukawa coupling to muons is therefore an important test of the SM description of the Higgs sector.

1.4 Experimental profile of the Higgs boson

In 2012, the ATLAS and CMS Collaborations announced the observation of a new boson with a mass around 125 GeV compatible with the SM predictions. One year later F. Englert and P. Higgs were jointly awarded the Nobel prize in physics "for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider" [61]. Since its discovery various studies have been performed to examine the properties of the Higgs boson. So far, the Higgs boson has been confirmed to be an electrically neutral, scalar particle which couples to the electroweak gauge bosons and all charged fermions of the third generation.

In the following, the phenomenology of Higgs boson production at the *Large Hadron Collider* (LHC) is discussed followed by a brief summary on existing searches for Higgs boson decays to muons by the ATLAS and CMS Collaborations.

1.4.1 The Higgs boson at the LHC

At the LHC, a high-energy proton-proton collider, the dominant Higgs boson production modes are gluon fusion (ggH), vector boson fusion (VBF), associated production with a vector boson (VH), and associated production with a heavy quark-antiquark pair ($t\bar{t}H/b\bar{b}H$) or with a single top quark (tH). Figure 1.2 shows the corresponding cross sections as well as the branching fraction for the decays of the SM Higgs boson as a function of the Higgs boson mass m_H . The branching fraction of the Higgs boson decay to a pair of muons is $\mathcal{B}(H \rightarrow \mu^+\mu^-) = 2.18 \cdot 10^{-4}$ (for $m_H = 125$ GeV). Representative LO Feynman diagrams of the described Higgs boson production modes are shown in Fig. 1.3.

- **ggH** The Higgs boson production mechanism with the largest production rate is the gluon fusion process, which is mediated by a virtual, heavy top quark loop. Since the Higgs boson couplings to fermions are proportional to the fermion mass, loop contributions from lighter quarks are suppressed proportionally to m_q^2 .

- **VBF** The vector boson fusion process is the production mode with the second largest production rate at the LHC, which is, however, about 10 times smaller compared to ggH. The two quarks from the protons radiate W or Z bosons, which interact to produce a Higgs boson. The scattered quarks result in two hadronic jets in the forward and backward regions of the detector, and in reduced gluon radiation from the central region.
- **VH** The third most prevalent Higgs boson production mode at the LHC is the associated production with a W or Z boson. A vector boson, which is produced in quark-antiquark annihilation, radiates a Higgs boson. The production rate for WH is about twice as large as for ZH.
- **t \bar{t} H/b \bar{b} H** The associated production of a Higgs boson with a heavy quark-antiquark pair has a production rate which is nearly two orders of magnitude smaller than the production rate of ggH. The t \bar{t} H/b \bar{b} H production is initiated by gluons producing a heavy quark-antiquark pair via t-channel exchange. The Higgs boson is radiated from the internal quark line.
- **tH** The Higgs boson production in association with a single top quarks has the smallest production rate of all here described processes. A single top quark is produced via t-channel exchange of a W boson. The Higgs boson is radiated from the top quark.

The dominant Higgs boson production modes at the LHC, ggH, VBF, VH, and t \bar{t} H, as well as the decays of the Higgs boson to the electroweak gauge bosons, and the charged fermions of the third generation have been established by the ATLAS and CMS Collaborations. So far the measured production and decay rates have been in agreement with the SM expectations.

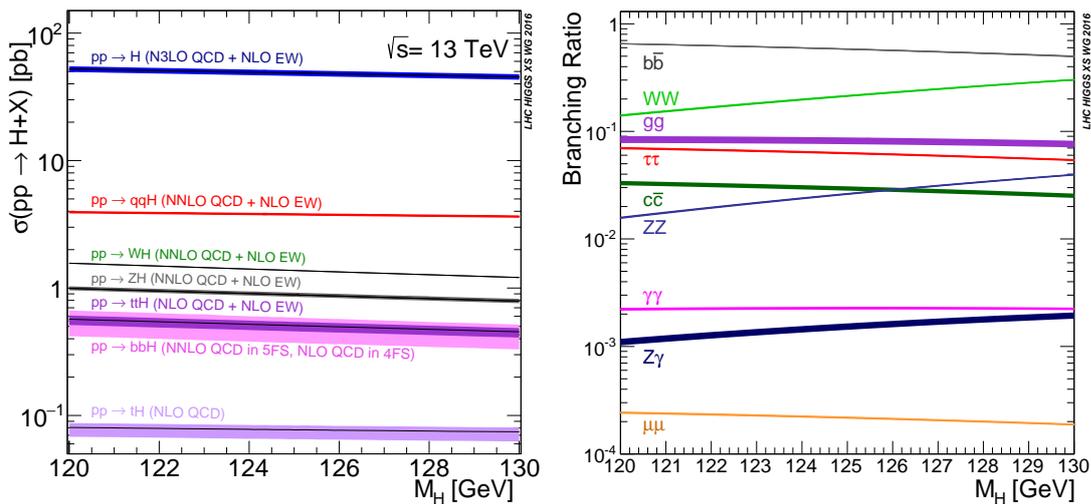


Figure 1.2: The Higgs boson production cross section for different Higgs boson production modes (left) and the branching fraction for the main decays of the Higgs boson (right) as a function of the Higgs boson mass m_H . Taken from Ref. [62].

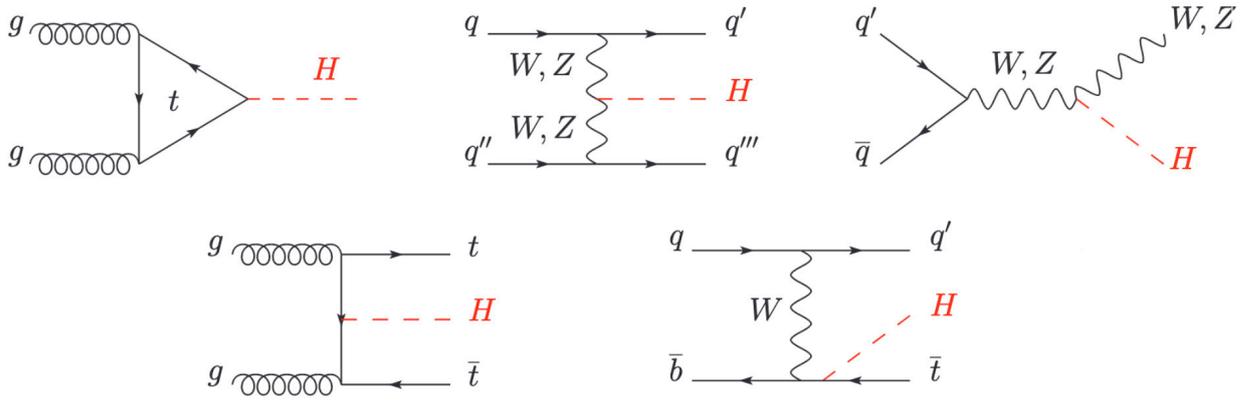


Figure 1.3: Main LO Feynman diagrams contributing to the Higgs boson production at the LHC. Shown are gluon fusion (top left), vector boson fusion (top middle), associated production with a vector boson (top right), associated production with a top quark-antiquark pair (bottom left) (similar diagram for bottom quarks), and production in association with a single top quark (bottom right). Modified from Ref. [54].

To further explore the Higgs boson sector, and especially the Yukawa interactions of the second-generation fermions, the Higgs boson decay to muons has been investigated by the ATLAS and CMS Collaboration. Due to the small expected branching fraction for the decay of the Higgs boson to muons, the signal rate is very small compared to the large SM backgrounds dominated by the Drell-Yan process. Thus, the analysis of Higgs boson decays to muons is experimentally challenging. The results of a previously published $H \rightarrow \mu^+\mu^-$ analysis by the CMS Collaboration [41], and the recent results by the ATLAS Collaboration [42] are briefly described in the following. The ATLAS analysis has been performed in parallel to the analysis presented in this thesis.

1.4.2 Previous search for $H \rightarrow \mu^+\mu^-$ with the CMS experiment

The search was based on proton-proton collision data recorded in 2016 at a center-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 35.9 fb^{-1} . Events were accepted by a single muon trigger. Events which contained two isolated and oppositely charged muons with a transverse momentum $p_T > 20 \text{ GeV}$ and a pseudorapidity $|\eta| < 2.4$ were selected. The Higgs boson mass was reconstructed from the two muons with the highest single muon p_T . The $H \rightarrow \mu^+\mu^-$ signal is expected to appear as a narrow resonance in the dimuon mass $m_{\mu\mu}$ over a smoothly falling background spectrum, dominated by Drell-Yan and dileptonic $t\bar{t}$ background. To extract the signal, a maximum likelihood fit to the observed dimuon mass $m_{\mu\mu}$ distribution was performed. The signal and background components were modeled using analytical functions. In this way, the background component in the signal region around the Higgs boson mass was estimated from signal-depleted $m_{\mu\mu}$ sidebands in data.

The analysis was optimized using a multivariate discriminant based on a boosted decision tree (BDT). The BDT classifier included several observables, which described the kinematic properties of the dimuon system and additional hadronic jets to distinguish the $H \rightarrow \mu^+\mu^-$ signal from the SM backgrounds. To avoid bias in the final signal extraction fit, the observables were chosen to be largely uncorrelated with $m_{\mu\mu}$. The output of the BDT classifier was transformed such that the sum of all signal events has a uniform distribution, as shown in Fig. 1.4. Events were categorized according to the BDT output score as well as the maximum $|\eta|$ of the

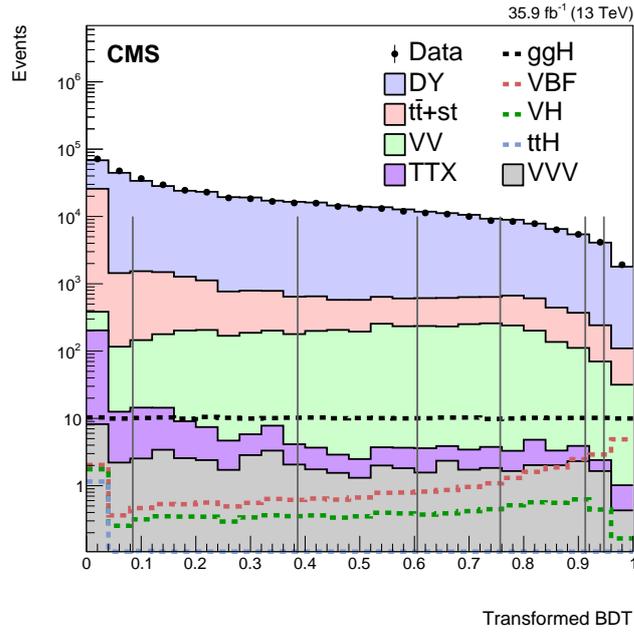


Figure 1.4: Transformed BDT output distribution of the previous CMS $H \rightarrow \mu^+\mu^-$ analysis for data (solid points) and simulation (histograms). The solid histograms represent the different background contributions. The dashed histograms represent the signal contributions. The vertical lines indicate category boundaries. Taken from Ref. [41].

muons in the event, accounting for differences in the muon momentum resolution between the central part of the CMS detector and the forward region. The category boundaries were determined using a two dimensional optimization algorithm. Finally, the signal extraction fit of the $m_{\mu\mu}$ distribution was performed simultaneously in 15 subcategories. The $m_{\mu\mu}$ distribution for the weighted combination of all event categories is shown in Fig. 1.5.

Results were presented for a Higgs boson with mass of 125.09 GeV [43]. When combined with the LHC Run 1 results for $H \rightarrow \mu^+\mu^-$, an observed (expected) upper limit at 95% confidence level was set on the production cross section times branching fraction of 2.9 (2.2) times the SM prediction. The observed signal strength modifier was $\hat{\mu} = 1.0 \pm 1.0$ (stat) ± 0.1 (syst).

In comparison, the new CMS $H \rightarrow \mu^+\mu^-$ analysis presented in this thesis is based on the full LHC Run 2 dataset, corresponding to nearly four times the integrated luminosity of the 2016

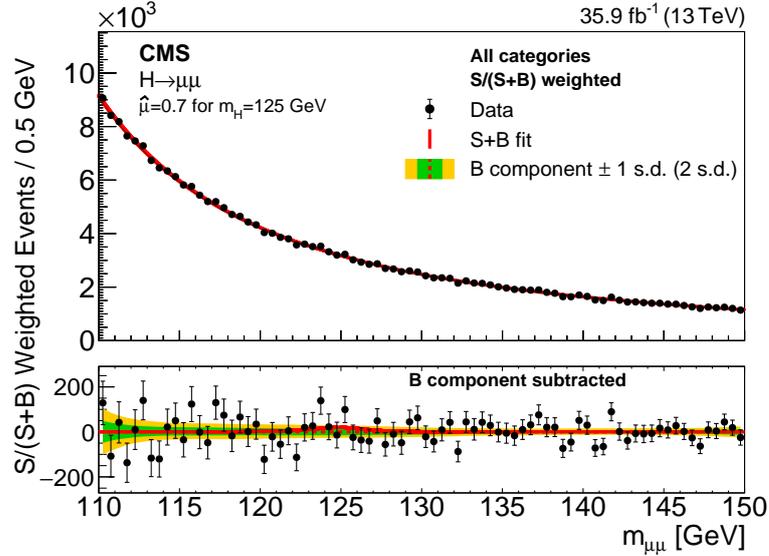


Figure 1.5: The $m_{\mu\mu}$ distribution for the weighted combination of all event categories in the previous CMS $H \rightarrow \mu^+\mu^-$ analysis. Events are weighted according to the expected $S/(S+B)$ in the category to which they belong. The lower panel shows the difference between data and the background component of the fit. Taken from Ref. [41].

dataset. The new analysis introduces dedicated Higgs boson production categories, which individually target ggH, VBF, VH and $t\bar{t}H$ to increase the overall sensitivity. The development of the $t\bar{t}H$ ($H \rightarrow \mu^+\mu^-$) analysis strategy is described in Chapter 5. The combined results, as well as an explicit comparison of the new CMS analysis with the analysis presented here, is given in Chapter 6.

1.4.3 Recent results on $H \rightarrow \mu^+\mu^-$ with the ATLAS detector

The ATLAS Collaboration performed a search for Higgs boson decays to muons using the full proton-proton collision dataset recorded during the LHC Run 2 period at a center-of-mass energy of 13 TeV corresponding to an integrated luminosity of 139fb^{-1} . Events were recorded by single muon triggers with $p_T > 26\text{GeV}$ for isolated muons and $p_T > 50\text{GeV}$ for non-isolated muons. The analysis selected events with two loosely isolated, oppositely charged muons, which were categorized into different Higgs boson production categories according to the presence of additional leptons, the number of jets, and the number of jets originating from the hadronization of b quarks referred to as b-tagged jets. Muons were selected within the detector acceptance of $|\eta| < 2.7$ using a p_T threshold of 27GeV for the highest p_T muon, while the subleading muon p_T threshold was 15GeV . In a dedicated three lepton category, which targeted the VH production, this threshold was lowered to 10GeV . Up to one final-state photon is considered when the invariant mass of the Higgs boson candidate is reconstructed. The p_T^γ threshold increases linearly from 3 to 8GeV as a function of the distance between the muon and the photon.

A $t\bar{t}H$ enriched phase space region was defined based on events with at least one lepton with $p_T > 15$ GeV in addition to the muon pair from the Higgs boson decay and at least one b-tagged jet. This category aimed at the dileptonic or semileptonic decays of the top quark-antiquark pair. The two highest- p_T muons with opposite charge were chosen as the Higgs boson candidate. Twelve kinematic observables were combined in a BDT classifier. These particularly focused on the reconstruction of the top quark decay products. Only events in the high BDT score region were included in the final $t\bar{t}H$ event category which contributes to the combined results. Events where both top quarks decay hadronically are not considered in any of the analysis categories.

Events with at least one isolated lepton in addition to the muon pair associated with the Higgs boson candidate and without b-tagged jets were assigned to the VH category. This category was tailored to signal events where the Higgs boson is produced in association with a leptonically decaying W or Z boson leading to three and four leptons in the final state. The assignment of muons to the Higgs boson and the Z (W) boson candidates was based on charge requirements and a dedicated χ^2 criterion resulting in a correct pairing rate of 97% (93%) in the three (four) lepton category. Two BDT classifiers were used to optimize the signal-to-background ratio separately for the WH and ZH production category. Similar to the $t\bar{t}H$ event category, only events with high to medium BDT score were used in the combined results. Two VH event categories were defined with three leptons in the final state. One category accounted for the four lepton final state.

Events, which were selected neither in the $t\bar{t}H$ nor the VH event categories, were further categorized according to the number of jets into three exclusive jet multiplicities with zero, one, or two and more jets. Dedicated BDT classifiers were used to define four VBF and twelve ggH enriched event categories according to the BDT classifier score. Several observables characterizing the dimuon system, as well as observables describing the kinematics of dijet systems identifying VBF signal events were included in the BDT classifiers. Figure 1.6 presents the output of the BDT classifiers for 0-jet events, 1-jet events, and 2-jet events.

After event categorization, the signal was extracted by a simultaneous maximum-likelihood fit in all 20 event categories to the $m_{\mu\mu}$ distributions in the range 110–160 GeV. The Higgs boson was assumed to have a mass of $m_H = 125.09 \pm 0.24$ GeV. The distributions of $m_{\mu\mu}$ for the weighted combination of all analysis categories are shown in Fig. 1.7. Events are weighted by $\ln(1 + S/B)$, where S is the observed signal yield and B is the background yield derived from the fit to data in $120 < m_{\mu\mu} < 130$ GeV.

The best-fit value of the signal strength parameter was $\hat{\mu} = 1.2 \pm 0.6$, corresponding to an observed (expected) significance of 2.0σ (1.7σ) with respect to the background-only hypothesis. Figure 1.8 summarizes the best-fit values of the signal strength parameters for the five main event categories and the combination.

1.4 Experimental profile of the Higgs boson

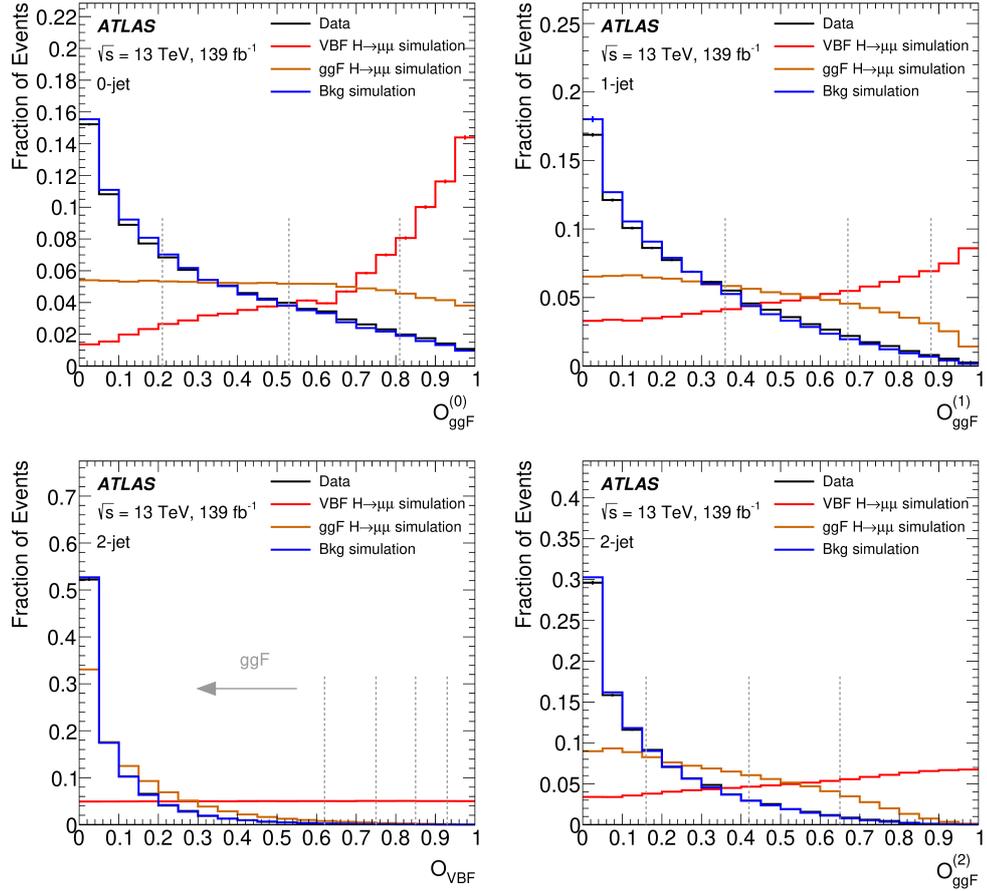


Figure 1.6: BDT classifier outputs of the ATLAS $H \rightarrow \mu^+\mu^-$ analysis for 0-jet events (top left), 1-jet events (top right), and 2-jet events (bottom). In each panel the normalized distributions are shown for the VBF (red) and ggH production (orange) as well as the simulated background used in the BDT classifier training (blue) and the observed data (black). The vertical lines indicate the category boundaries. Taken from Ref. [42].

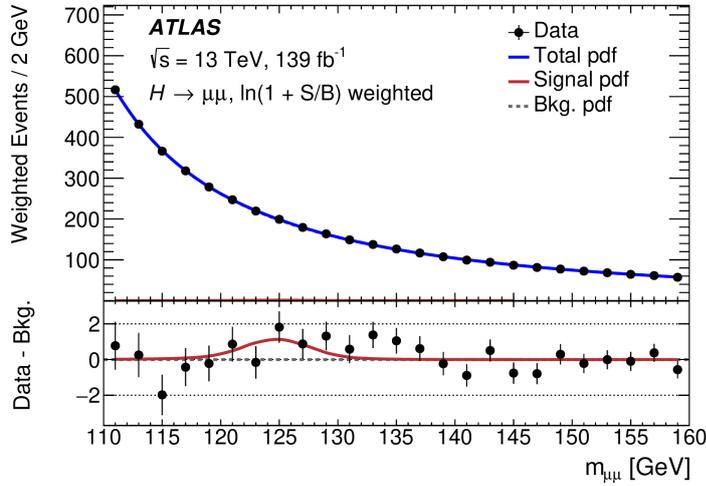


Figure 1.7: The $m_{\mu\mu}$ distribution for the weighted combination of all event categories in the ATLAS $H \rightarrow \mu^+\mu^-$ analysis. Events are weighted by $\ln(1 + S/B)$. The background and signal model are derived from the fit to data, with S normalized to its best-fit value. The lower panel shows the difference between the data and the background model. Taken from Ref. [42].

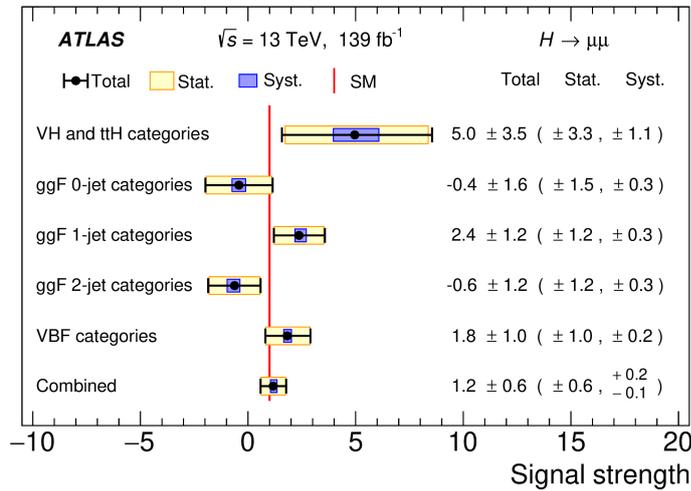


Figure 1.8: The best-fit values of the signal strength parameters of the ATLAS $H \rightarrow \mu^+\mu^-$ analysis for the five major groups of categories are shown together with the combined value. Taken from Ref. [42].

Experimental setup

Particle accelerators have become one of the key tools for producing new particles and probing the fundamental structure of matter and interactions as formulated in the SM. The world's most powerful particle accelerator, the *Large Hadron Collider* (LHC) [63, 64], is operated by the *European Organization for Nuclear Research* (CERN) near Geneva at the border between France and Switzerland. At the LHC, four particle physics experiments are employed to detect and record particles created in collisions of highly energetic particle beams: ATLAS [65], CMS [66], LHCb [67] and ALICE [68]. The two multipurpose experiments ATLAS and CMS aim at a broad physics research program reaching from precision measurements to searches for new physics beyond the SM. However, they differ in terms of the technologies used for particle detection and identification. Furthermore, they follow independent and individual analysis strategies. The cross validation of results by CMS and ATLAS is a key aspect of CERN's particle physics research program. The LHCb experiment is mainly designed to study CP violation and possible indirect signs of new physics in b-quark decays. The ALICE detector is specialized for the analysis of heavy-ion collisions recorded in specific runs of the LHC.

This chapter provides an overview of the experimental setup which is the basis of the physics data analysis performed in this thesis. First, the CERN accelerator complex and, in particular the LHC, are introduced in Section 2.1. In addition, a few general concepts of proton-proton colliders are briefly described. Second, the design of the CMS experiment and the detector technologies of each component are presented in Section 2.2.

2.1 The Large Hadron Collider

The LHC is a circular hadron collider designed to deliver proton beams for collisions at a center-of-mass energy of up to 14 TeV. Accelerated protons counter-revolve in two evacuated beam pipes until they are collided at certain interaction points. When the proton beams collide, new particles are created, and the corresponding events are recorded by the experiments. The accelerator and the experiments lie 45 m to 170 m below the surface in a tunnel with a circumference of 26.7 km between the lake Geneva and the Jura mountains. The underground infrastructure served previously for the *Large Electron-Positron Collider* (LEP) [69, 70].

Before the proton beams can enter the LHC, their energy needs to be successively increased. Therefore, they are run through a chain of preaccelerators. The proton source is a duoplasmatron [71] where protons are obtained by the ionization of hydrogen gas. In a first step, these protons are accelerated in a *linear accelerator* (LINAC). Then, the beam energy is further in-

Chapter 2. Experimental setup

created in the *proton synchrotron booster* (PSB), the *Proton Synchrotron* (PS), and the *Super Proton Synchrotron* (SPS) whereby the beam energy reaches 450 GeV. Finally, the beams are injected into the LHC and accelerated to their maximum energy. These and other features of the CERN accelerator complex are depicted in Fig. 2.1. The particle beams consist of up to 2808 bunches with 25 ns bunch spacing. A single bunch contains about 10^{11} protons. The beams are accelerated by superconducting cavities operated at 400 MHz. In the eight arcs of the LHC, superconducting niobium titanium dipole magnets are used to guide the beams on a curved track. The beams are focused by quadrupole magnets. Octopole and sextupole magnets correct the beam optics for energy-dependent tune shifts or the influence of irregularities in the magnetic field.

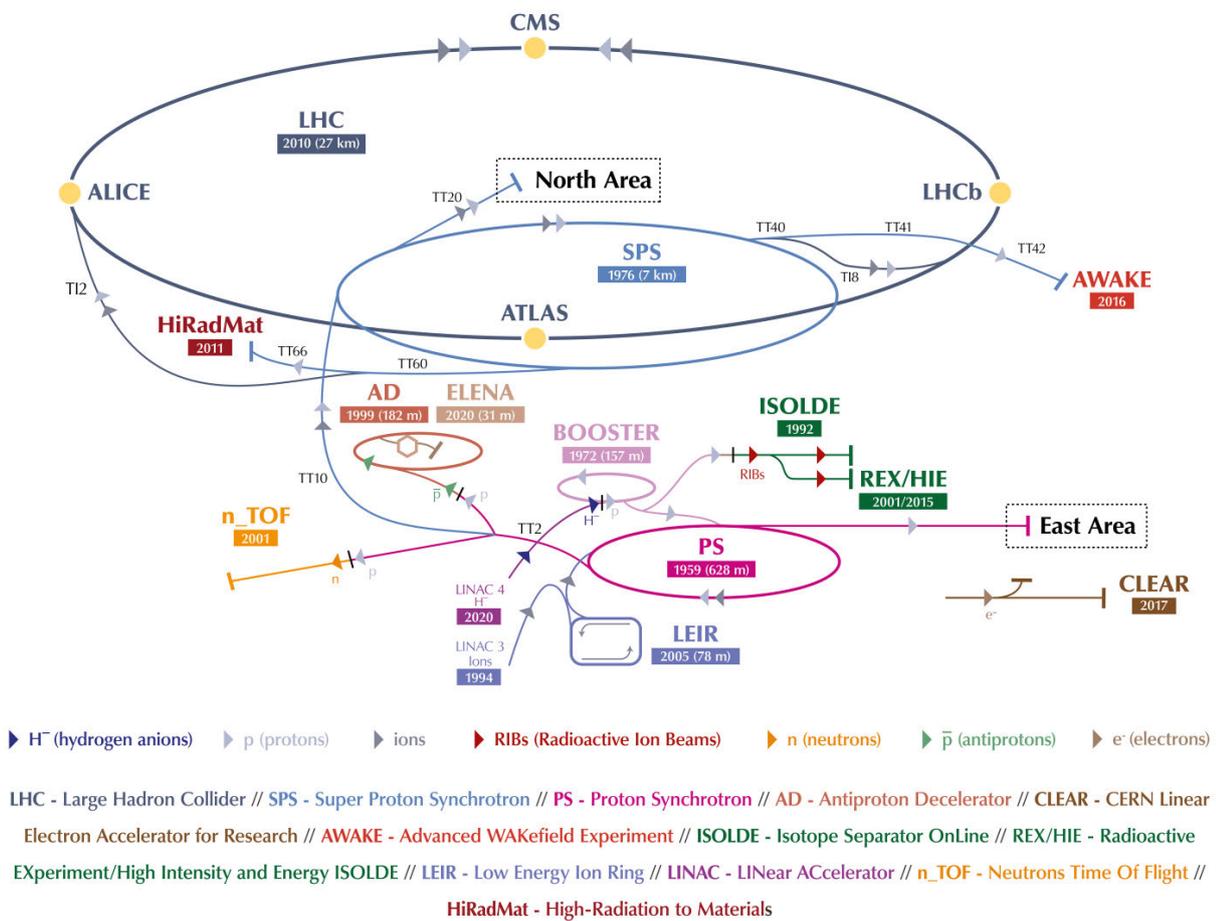


Figure 2.1: Depiction of the CERN accelerator complex. Taken from Ref. [72].

The integrated luminosity \mathcal{L}_{int} is a measure to predict the number of events N that take place in the beam collisions according to

$$N = \sigma \cdot \mathcal{L}_{\text{int}} = \sigma \cdot \int L dt \quad (2.1)$$

where σ denotes the cross section of a given process and L corresponds to the instantaneous luminosity which is determined by beam properties. The instantaneous luminosity L is given by

$$L = \frac{\gamma f k_b N_p^2}{4\pi\epsilon_n \beta^*} F \quad (2.2)$$

in terms of the Lorentz factor γ , the revolution frequency f , the number of bunches k_B , the number of protons per bunch N_p , the normalized transverse emittance ϵ_n , the betatron function at the interaction point β^* , and the hourglass reduction factor F . The emittance and the value of the betatron function at the interaction point are a measure of the transversal beam size at the interaction point. The hourglass factor is applied to consider the crossing angle between the two beams.

The designed instantaneous peak luminosity of 10 Hz/nb has already been surpassed. The LHC delivered a maximum peak luminosity of 15.3 Hz/nb in 2016 and up to 21 Hz/nb in 2017 and 2018 as presented in Fig. 2.2.

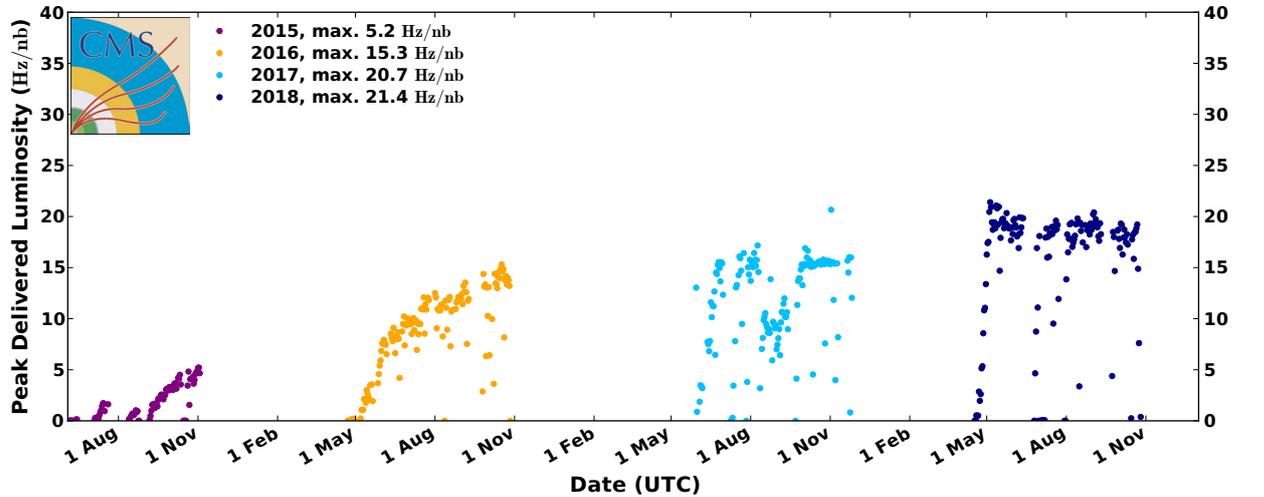


Figure 2.2: The instantaneous luminosity in proton-proton collisions delivered by the LHC for each data-taking period of the LHC Run 2. Taken from Ref. [73].

Figure 2.3 presents an overview of the delivered luminosity per data-taking period. The certified dataset with fully operational detectors used for particle physics analysis is usually slightly smaller compared to the values of the delivered luminosity. In 2010 and 2011 the LHC was run with a center-of-mass energy of 7 TeV. In this period of time, CMS collected and certified 5.6 fb^{-1} of proton-proton collision data. In the following year, the center-of-mass energy

was increased to 8 TeV and the verified data accumulated to a total amount of 21.8 fb^{-1} . The data-taking period until 2012 refers to LHC Run 1. Since 2015 the LHC has been operated at a center-of-mass energy of 13 TeV. The analysis presented in this thesis is based on LHC Run 2 data recorded in proton-proton collisions with the CMS experiment between 2016 and 2018 [74–76]. The certified dataset corresponds to an integrated luminosity of 137 fb^{-1} .

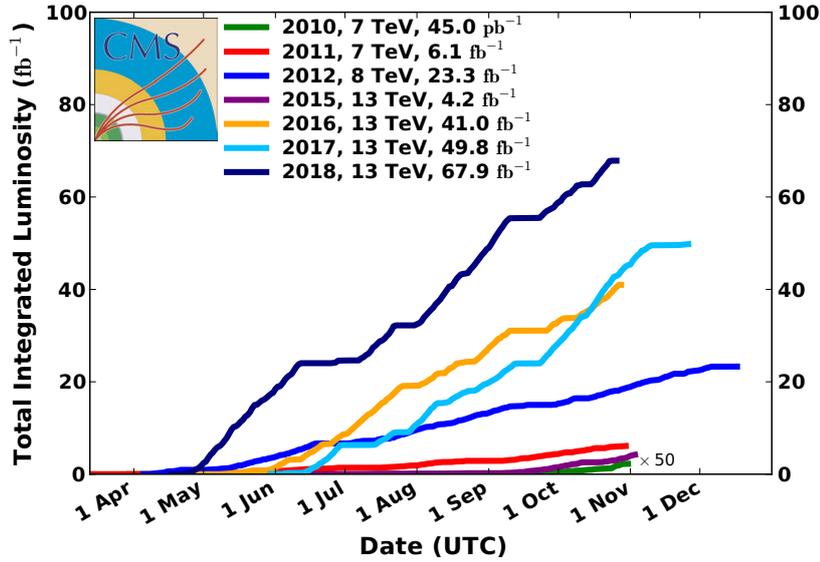


Figure 2.3: The integrated luminosity in proton-proton collisions delivered by the LHC for each data-taking period of the LHC Run 1 and Run 2. Taken from Ref. [73].

2.2 The Compact Muon Solenoid experiment

The *CMS* (Compact Muon Solenoid) experiment is one of the largest international scientific collaborations in the world. It involves 5000 particle physicists, engineers, technicians, students and support staff to operate the detector and analyze the recorded proton-proton collision data. Data-taking at hadron colliders entails many challenging tasks that lead to specific detector requirements. To meet these requirements, multipurpose detectors are usually composed of several sub-detector components. Each is specialized for the identification and measurement of specific particles and their properties. The CMS detector is constructed in a cylindrical shape around the interaction point. Due to its hermetical design, it provides a comprehensive picture of the physics taking place in the particle collisions. The dimensions of CMS are 21.6 m in length and 14.6 m in diameter with a weight of 14000 tons. The detector is divided into a *barrel* and *endcap* part. The components of the barrel are aligned in layers parallel to the beam pipe, while the endcap detectors are arranged perpendicular to it. Figure 2.4 presents an illustration of the CMS detector with its sub-detector components. Key features of the design and operation of the detector are described in the following.

CMS DETECTOR

Total weight : 14,000 tonnes
Overall diameter : 15.0 m
Overall length : 28.7 m
Magnetic field : 3.8 T

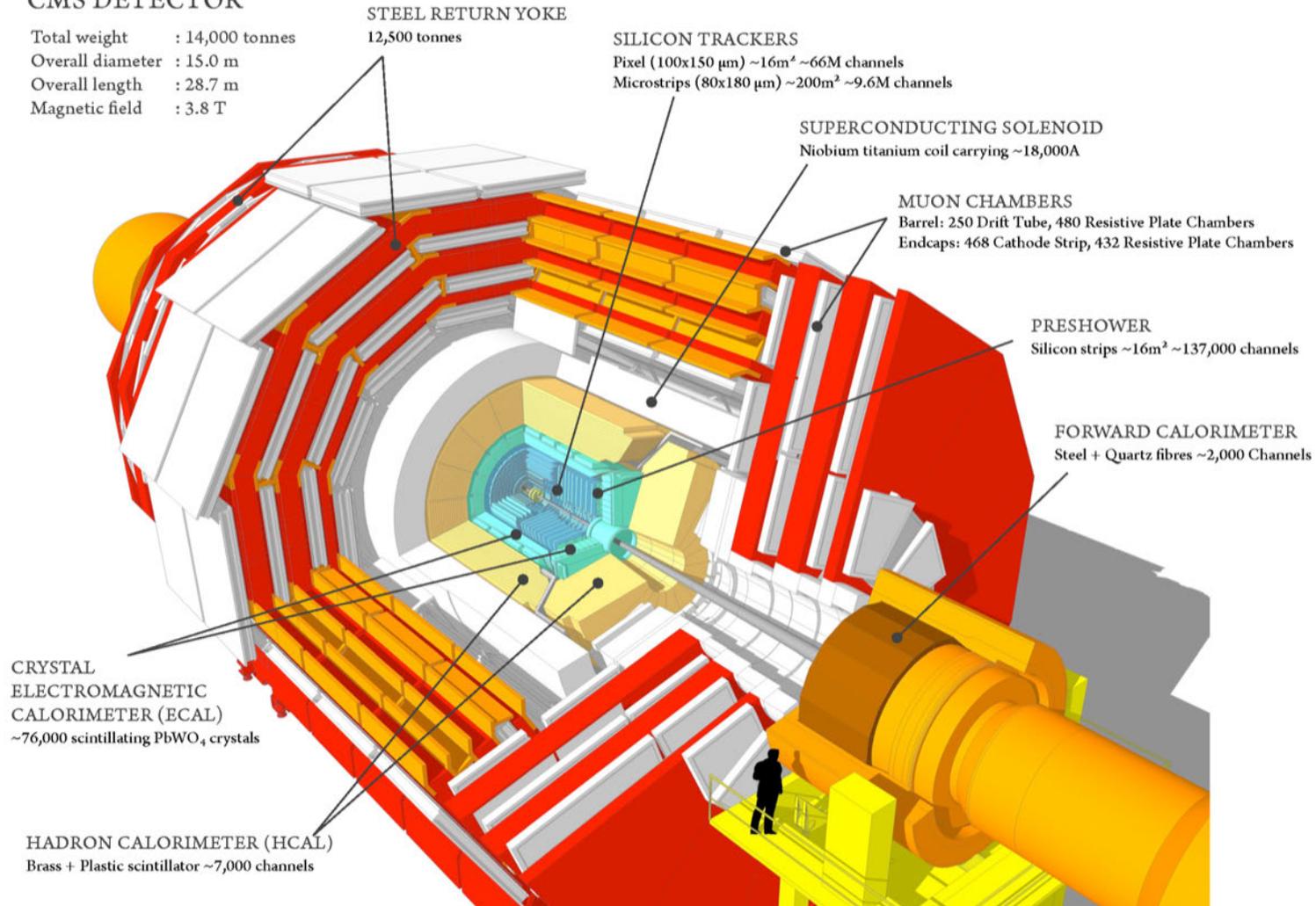


Figure 2.4: Depiction of the CMS experiment with details of the sub-detector components from the inside to the outside: the silicon tracking system, the electromagnetic calorimeter, the hadron calorimeter, the superconducting solenoid, the muon chambers, the steel return yoke, and the forward calorimeter. Taken from Ref. [77].

2.2.1 Coordinate system & kinematic quantities

The CMS coordinate system is illustrated in Fig. 2.5. The origin of the right-handed CMS coordinate system is located inside the detector at the nominal interaction point. The x -axis points towards the center of the LHC. The y -axis is aligned perpendicular to the x -axis, and points upwards. In the transverse plane defined by these axes, the azimuthal angle $\phi \in [-\pi, \pi]$ is measured from the x -axis. Accordingly, the z -axis is oriented parallel to the proton-beam. The polar angle $\theta \in [0, \pi]$ is defined with respect to the z -axis.

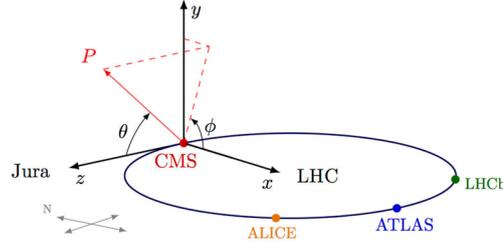


Figure 2.5: Coordinate system of the CMS detector at the LHC. Taken from Ref. [78].

The reconstruction of particle interactions at a proton-proton collider involves specific challenges due to the composite nature of the proton. The hard interaction, which is characterized by a large momentum transfer, takes place between one parton of each of the protons. Each parton carries only an unknown fraction $x_{i=1,2}$ of the proton's momentum, and in general $x_1 \neq x_2$. As a result, the center-of-mass frame of the colliding partons is boosted with respect to the laboratory frame along the z -axis. Particles created in beam collisions are highly relativistic, $v \approx c$, and therefore, Lorentz invariant quantities are preferred for their description. Differences of the rapidity y are Lorentz invariant. The rapidity y is defined as

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}. \quad (2.3)$$

In the limit of large momentum in the transverse plane $p_T = \sqrt{p_x^2 + p_y^2}$, the pseudorapidity η becomes equal to the rapidity

$$y \approx \eta - \frac{\cos \theta}{2} \left(\frac{m}{p_T} \right)^2 \xrightarrow{p_T \gg m} y \approx \eta. \quad (2.4)$$

The pseudorapidity can be expressed in terms of the polar angle θ with

$$\eta = -\ln \tan \left(\frac{\theta}{2} \right). \quad (2.5)$$

Using this definition, a nearly Lorentz invariant measure of the distance between two high momentum objects can be formulated for boosts along the z -axis by

$$\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}. \quad (2.6)$$

2.2.2 Tracking system

The tracking system [79–81] is designed to precisely reconstruct the trajectories of charged particles. The tracker is the innermost detector component of CMS. The immense particle flux at the LHC necessitates a high granularity, a quick response and, in particular, radiation hardness of the tracking system. Therefore, a design entirely based on silicon detectors is chosen. While the inner part consists of silicon pixel sensors, the outer part uses slightly more coarse silicon strip sensors. The sensors measure the position of the traversing particles due to ionization in the silicon. From signals in different layers of the detector a particle trajectory can be reconstructed. The tracker aims also to locate the primary interaction vertex, as well as secondary vertices, which result for example from the decay of b hadrons. An illustration of the tracker is shown in Fig. 2.6. The tracking system has a cylindrical shape with a length of 5.8 m and a diameter of 2.5 m. It consists of several barrel and endcap layers which cover a region up to $|\eta| = 2.5$.

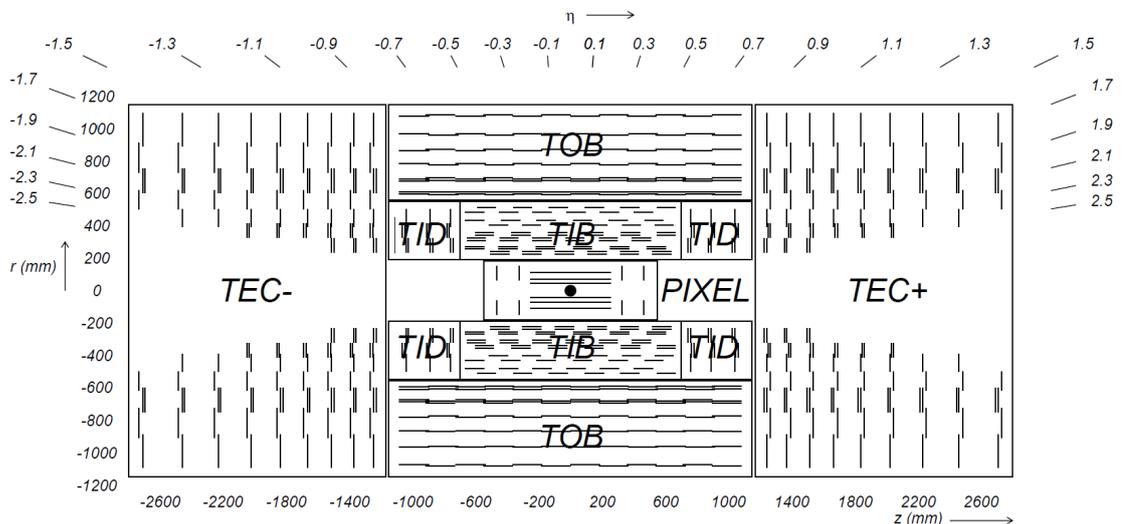


Figure 2.6: Schematic cross section of the CMS tracking system. Shown are the phase-0 pixel detector and the different components of the strip tracker system. Taken from Ref. [66].

Pixel tracker

In 2016 the pixel detector consisted of three barrel layers of modules equipped with silicon pixel sensors [82]. The layers were placed at 4.4 cm, 7.3 cm, and 10.2 cm radial distance to the beam pipe. Additionally, two endcap disks with a radius of 6 cm and 15 cm were mounted on each side. The pixel tracker contained in total 66 million pixels, each with a size of $100 \times 150 \mu\text{m}^2$. This resulted in a precise hit-position resolution of 10–15 μm in the (r, ϕ) plane and 15–20 μm in z direction. To preserve the high tracking quality, a new phase-1 pixel detector [83] has replaced the described layout in-between the 2016 and 2017 data-taking periods. Figure 2.7 presents a comparison between both layouts. The new phase-1 pixel detector comprises 124 million pixel sensors arranged in four layers in the barrel and three disks in the endcaps. The

barrel layers are placed at 2.9 cm, 6.8 cm, 10.9 cm and 16.0 cm radial distance to the beam pipe. Compared to the previous layout the inner barrel layer is now placed closer to the beam pipe while the outer layer is placed further outside. The dimensions of the sensors are unchanged. The fourth layer provides an additional measurement which improves the tracking robustness and precision, while the new inner layer at small radial distance improves the capability to identify secondary vertices from for example the decay of b hadrons.

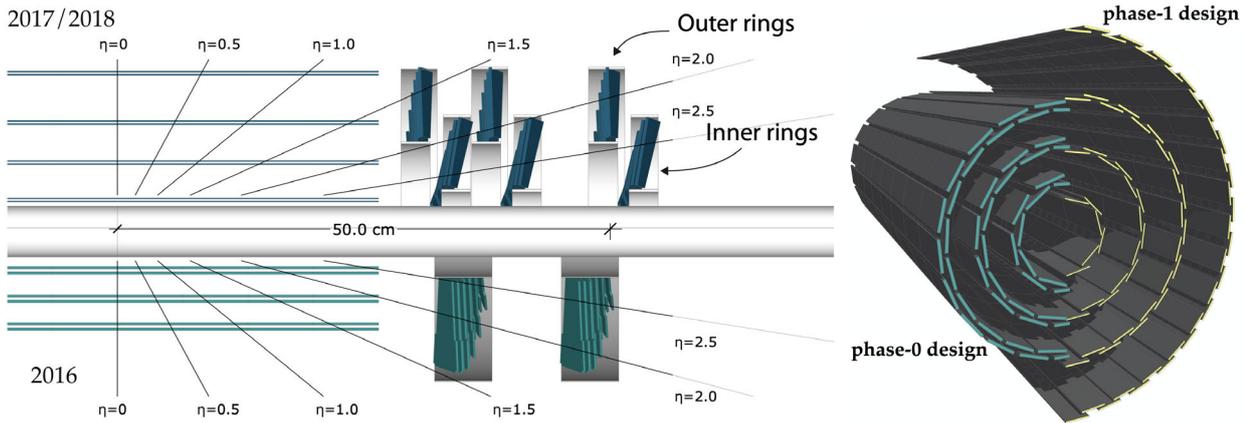


Figure 2.7: Depiction of the CMS pixel detector phase-1 upgrade. A comparison of the initial CMS pixel detector layout and the phase-1 design is presented by a cross section (left) and a three-dimensional model (right). The phase-1 design comprises four layers in the barrel and three disks in the endcaps corresponding to the layout on the top of the cross section and the layout on the right in the three-dimensional model. Modified from Ref. [83].

Strip tracker

The strip tracker [84] encloses the beam pipe and the pixel detector. It consists of ten barrel layers and twelve endcap disks. The modules of silicon strip sensors extend the sensitive tracking volume. The size and shape of the 9.6 million strip sensors differ depending on their location in the tracker. Due to the reduced particle flux at large radii, the size of silicon strip sensors can be enlarged while keeping the occupancy at a low level. The strip width varies between 80 and 180 μm in the inner and outer part, respectively. The spacial resolution is approximately 23 to 53 μm in the (r, ϕ) plane and 230 to 530 μm in z direction.

2.2.3 Calorimetry

Calorimeters are dense blocks of material in which the incoming particles deposit most of their energy. In a cascade of inelastic interactions of the incoming charged particle with the detector material, the energy of the incoming particle is distributed to many secondary particles, which leads to the formation of a particle shower. The initial particle energy can be determined from the deposited energy in the detector material. Two basic calorimeter designs can be distinguished. In a homogeneous calorimeter the sensitive material acts at the same time as absorber material, while a sampling calorimeter is based on alternating layers of sensitive material and absorber material.

Electromagnetic calorimeter

The *electromagnetic calorimeter* (ECAL) [85–87] is a homogeneous crystal calorimeter surrounding the inner tracking system. It is mainly designed to measure electron and photon energies. Lead tungstate (PbWO_4) crystals with a radiation length X_0 of 0.89 cm and a Molière radius R_M of 2.2 cm act as scintillating material. The high density, short radiation length, small Molière radius and fast scintillation decay time of the lead tungstate crystals allow for a compact calorimeter design with good position and time resolution. The layout of the ECAL is illustrated in Fig. 2.8. The ECAL barrel (EB) consists of 61200 lead tungstate crystals arranged in a cylindrical shape with an inner radius of 1.3 m covering a region up to $|\eta| < 1.479$. Each crystals cross section is 0.0174×0.0174 in $\eta \times \phi$ which corresponds to $22 \times 22 \text{ mm}^2$ in the inner part of the detector. The crystal length of 23 cm is equivalent to $25.5 X_0$. Due to the rather low light yield of approximately 30 photons per MeV, photodetectors with high amplification are required. For this purpose, the lead tungstate crystals are instrumented with avalanche photodiodes (APD). The ECAL endcaps (EE) are composed of 7324 lead tungstate crystals in each of the endcap discs. They extend the coverage to $|\eta| < 3$. They are slightly shorter compared to the crystals used in the barrel as they are only 22 cm long, corresponding to $24.7 X_0$. Their front cross section is $28.6 \times 28.6 \text{ mm}^2$. Vacuum phototriodes (VPTs) are employed to detect the scintillation light. The preshower (ES) sampling calorimeter is installed in front of the EE. It covers the pseudorapidity range of $1.7 < |\eta| < 2.6$ and consists of two planes of lead followed by silicon sensors. Due to its high granularity, the preshower is used to identify the background from neutral pions which decay into two collimated photons.

The performance of the ECAL was measured in a test beam [88]. The resulting energy resolution of the ECAL is parameterized by

$$\left(\frac{\sigma(E)}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{12\%}{E}\right)^2 + (0.3\%)^2, \quad (2.7)$$

where the first term corresponds to statistical effects in the shower development, the second term describes the electronic noise, and the third constant term is induced by non-uniform light collection and possibly calibration errors.

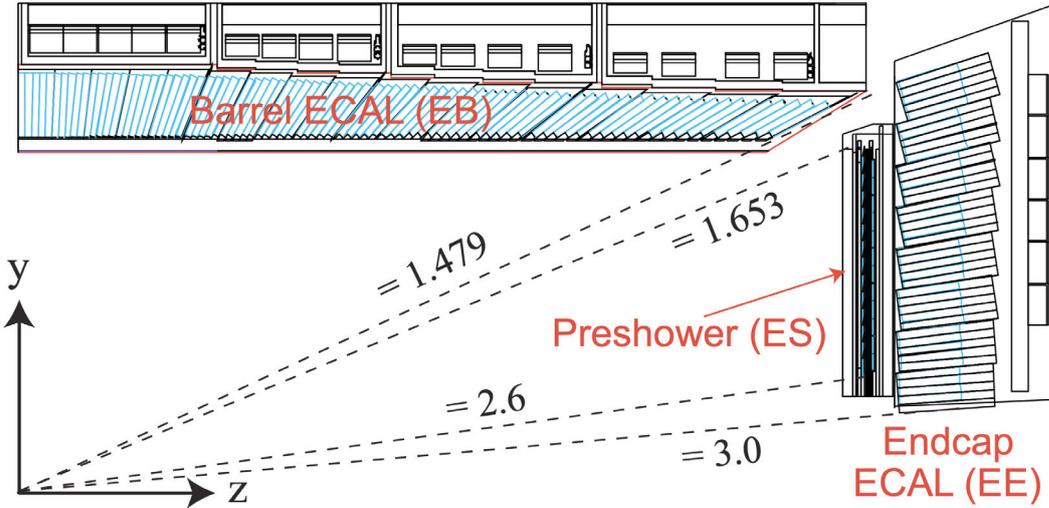


Figure 2.8: Schematic cross section of the CMS electromagnetic calorimeter (ECAL). Shown are the size and location of the barrel and endcap lead tungstate crystals. The preshower is situated right in front of the endcap calorimeter. The numbers indicate the pseudorapidity η values. Taken from Ref. [89].

Hadron calorimeter

The *hadronic calorimeter* (HCAL) [90] is a sampling calorimeter designed to measure the energy of hadrons. The layout of the HCAL is presented in Fig. 2.9. The *hadron barrel* (HB) and the *hadron endcap* (HE) are placed between the ECAL and the solenoid magnet. They cover a pseudorapidity region of $|\eta| < 1.3$ and $1.3 < |\eta| < 3.0$, respectively. Plastic scintillator tiles are mounted between brass plates which serve as absorber material with an interaction length of $\lambda_l = 16.4$ cm. The front and back panel consist of stainless steel to support the structural strength. The plastic scintillator tiles are segmented into parts covering 0.087×0.087 in $\Delta\eta \times \Delta\phi$ for $|\eta| < 1.6$, and 0.17×0.17 in $\Delta\eta \times \Delta\phi$ for larger pseudorapidities. The total thickness of the absorber material is restricted to fit into the solenoid magnet. It ranges from $5.8\lambda_l$ at $|\eta| = 0$ to a depth of about $10\lambda_l$ at $|\eta| = 1.3$. The barrel region is complemented by additional scintillator layers outside the solenoid magnet, referred to as *hadron outer* (HO) calorimeter. It increases the thickness of the absorber material to about $11.8\lambda_l$ within $\eta < 1.3$. To collect optical signals from the scintillator-based detectors in HB, HE and HO, wavelength-shifting (WLS) fibers are embedded in the scintillator tiles to convert the scintillation light, which is detected by multi-channel *hybrid photodiodes* (HPDs). In the forward region at $3 < |\eta| < 5.2$ the calorimeter is exposed to an enormous particle flux. The *hadron forward* (HF) is therefore instrumented with radiation hard quartz fiber scintillators embedded into steel absorber material. These fibers generate Cherenkov radiation when charged particles above the Cherenkov threshold pass through the material. Thus, the HF is mainly sensitive to the electromagnetic shower component. The Cherenkov light is detected by photomultiplier tubes. The described HCAL readout was upgraded throughout LHC Run 2. A key feature of the upgrade was the replacement of the old HPDs in the HB/HE/HO with new *silicon photomultipliers* (SiPMs) which offer some advantages like a higher photon detection efficiency and larger gain.

The energy resolution has been measured in a test beam [91, 92] for a combined ECAL and HCAL system resulting in a relative energy resolution of

$$\left(\frac{\sigma(E)}{E}\right)^2 = \left(\frac{120\%}{\sqrt{E}}\right)^2 + (7\%)^2. \quad (2.8)$$

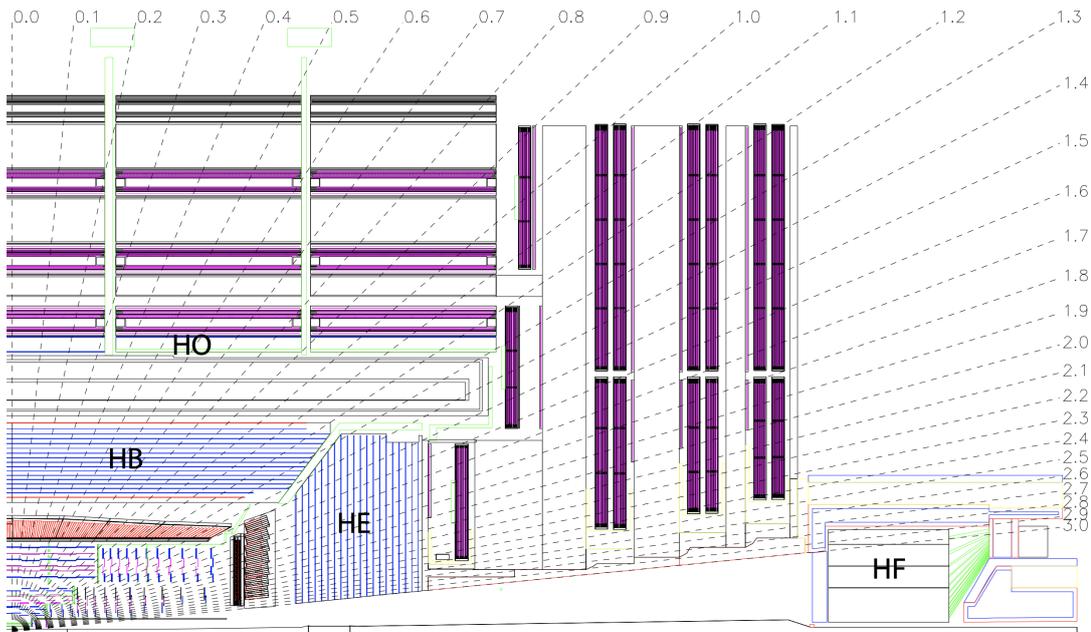


Figure 2.9: Schematic cross section of the CMS hadron calorimeter (HCAL). Shown are the size and location of the four main components of the HCAL: hadron barrel (HB), hadron endcap (HE), hadron forward (HF), and hadron outer (HO). Taken from Ref. [66].

2.2.4 Muon system

The muon system [93, 94] forms the outermost part of CMS. Similar to the inner tracking system, charged particle trajectories are reconstructed from signals in different layers of the muon system. In contrast to the silicon-based inner tracker, however, the muon system uses gaseous detectors. High momentum muons are minimum ionizing particles. They pass the tracker and also the calorimeters without showering. For this reason, muons are supposed to be the only particles in the SM besides neutrinos which reach the muon system. This allows a pure and efficient muon identification. Furthermore, the muon system is used to improve the muon momentum measurement of the inner tracking system, and for triggering purposes.

Three different types of gaseous tracking detectors are installed: Drift Tube (DT) Chambers, Cathode Strip Chambers (CSCs), and Resistive Plate Chambers (RPCs). They are arranged in four concentric layers and four disc-shaped caps as presented in Fig. 2.10.

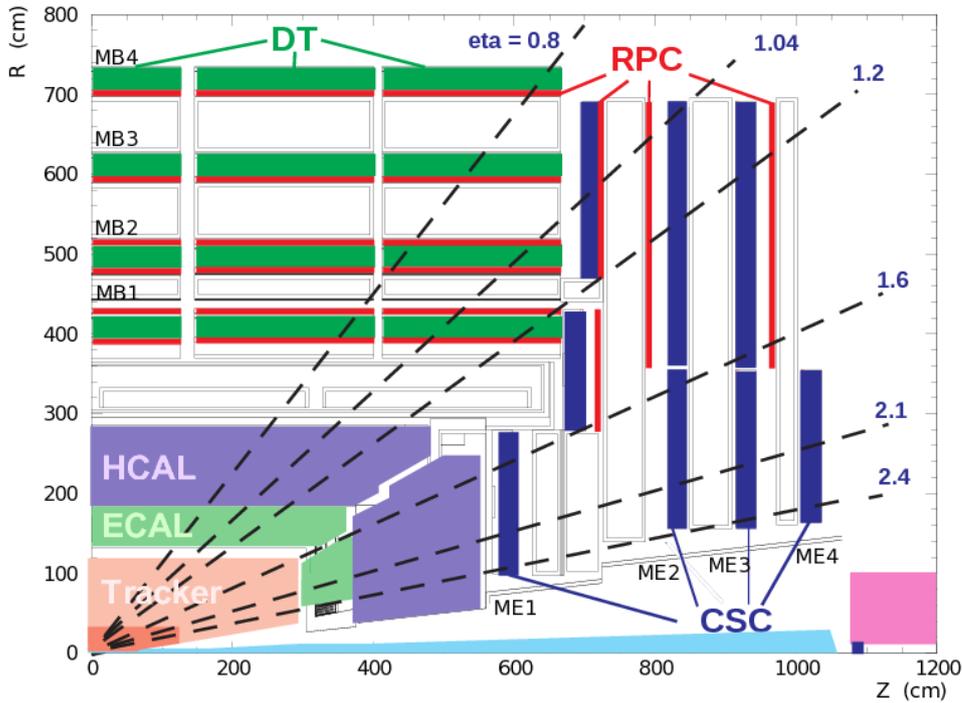


Figure 2.10: Schematic cross section of the CMS muon system. Shown are the main components of the muon system: Drift Tube (DT) Chambers, Cathode Strip Chambers (CSCs), and Resistive Plate Chambers (RPCs). Taken from Ref. [95].

The DT Chambers are installed in the muon barrel ($|\eta| < 1.2$) of the detector where the muon flux and neutron-induced background is small. Particles that traverse these detectors ionize a gas along their track. Charge carriers are separated due to the applied voltage in the drift tube and an avalanche of free electrons migrates to the anode, inducing a signal. With a maximum drift length of 2 cm, the barrel single point resolution is about $100 \mu\text{m}$ in position and about 1 mrad in ϕ -space.

The muon endcap ($0.9 < |\eta| < 2.4$) is instrumented with CSCs which can be operated at high particle flux and non-uniform magnetic fields. Closely placed anode wires between two cathodes are used to collect the avalanche electrons. CSCs provide high time resolution with adequate spatial resolution in the order of $200 \mu\text{m}$ in position and about 10 mrad in ϕ -space.

RPCs are gaseous parallel-plate detectors with adequate spatial resolution but excellent time resolution. They complement the DTs and CSCs in the region up to $|\eta| < 1.6$. Due to their fast response time being much shorter than 25 ns, they provide the required accuracy to unambiguously assign signals to the correct bunch crossing.

Each of the three CMS muon detector systems contribute to the level-1 muon trigger decision to ensure good coverage and redundancy [96]. The trigger system is described in Section 2.2.6. Trigger primitives are built from the signal hit information in the DT and CSC systems. These

are used by pattern recognition algorithms to identify muon candidates. The RPCs provide a complementary, dedicated triggering system with excellent time resolution. In contrast to CSCs and DTs, the RPC system does not form trigger primitives, instead the detector hits are used directly for trigger muon candidate recognition.

2.2.5 Magnet

When particles traverse the detector, their trajectory is bent due to a magnetic field. This field is provided by a superconductive solenoid magnet [97] which surrounds the tracker and both calorimeters. The solenoid magnet has an inner diameter of 5.9 m and a length of 12.9 m. From the curvature of the trajectories in the tracker, the particle momenta are measured and the sign of charge can be determined. A strong magnetic field of 3.8 T is delivered inside the magnet. The return of the magnetic flux in the iron yoke which is interposed between the muon chambers provides a field up to 2 T outside of the magnet. Since the orientation of the magnetic field changes outside of the solenoid, particle trajectories are bent in the opposite direction. Therefore, it is possible to combine two independent measurements from the tracker and the muon chambers to improve the muon momentum resolution.

2.2.6 Trigger system

The LHC bunch crossing rate of 40 MHz leads to very large quantities of data, which are impossible to read out and store in its entirety, since the hardware used to process these data is restricted to a transfer rate of $\mathcal{O}(1\text{ kHz})$. Because not all events contain signs of interesting highly energetic processes, a trigger system performs a first online selection of potentially interesting events. This reduces the number of stored events by several orders of magnitude. The trigger system at the CMS experiment [96] is divided into two subsequent steps.

The *level-1* (L1) trigger [98] utilizes information at hardware level from the calorimeters and the muon system. It performs a simplified event reconstruction, and selects events based on the presence of photons, electrons, muons, or jets above a certain transverse momentum threshold with a rate of $\mathcal{O}(100\text{ kHz})$. Events that pass the L1 trigger are further processed by the software-based *high-level trigger* (HLT) [99]. Here, a more sophisticated reconstruction is possible considering information from all sub-detector components. Interesting events are selected based on more detailed event properties. Several HLT paths which require similar event properties such as the presence of a single, isolated muon, are grouped and used to define the primary dataset. The HLT reduces the final output rate further to $\mathcal{O}(1\text{ kHz})$. Finally, all events passing the selection criteria by the L1 trigger and the HLT are written to data storage and are available for data analysis.

Event simulation

A reliable simulation of particle interactions is essential for probing or confirming theoretical predictions. Event generators are able to model interactions between colliding particles as well as additional radiation, particle showers, and prompt decays of unstable particles. A dedicated simulation of particle interactions with the detector material allows also for taking the finite detector resolution into account. Together, the event generator and the detector simulation serve as the basis of several aspects of particle physics analyses. Searches for new, yet undiscovered physics rely on simulation to model the signal process and to define the analysis strategy. Background processes can either be obtained from simulation, or estimated directly from data. In the former case, the simulation can be used in conjunction with data control regions to establish estimates of the background yields. In the latter case, the data-driven background estimation is validated using simulated events.

The analysis presented in this thesis relies on simulated signal and background events to optimize the analysis sensitivity. Furthermore, simulated events are used to validate selection efficiencies of signal processes, and to determine the corresponding systematic uncertainties. Section 3.1 introduces the different steps of the event generation and the common event generators that are used in this thesis. A comprehensive list of relevant simulated processes is presented in Section 3.2.

3.1 Event generation

The production of new heavy particles in proton-proton collisions, $pp \rightarrow X$, takes place in processes with large momentum transfer Q between the contributing partons $a, b \in \{q, \bar{q}, g\}$. These processes are denoted as the *hard interaction*. According to the factorization theorem [100–103], the scattering process at a hadron collider can be described in terms of the high-energy parton-parton interaction, based on perturbation theory, combined with a phenomenological description of the soft processes.

The initial state of the proton-proton collision is not precisely determined, due to the composite structure of protons. However, it can be described by the *Parton Distribution Functions* (PDFs) $f_i(x_i, Q^2)$, which correspond to the probability to find a parton i with the momentum fraction x_i in the proton given a certain energy scale, typically chosen to be Q^2 . The hard interaction itself, however, can be evaluated in perturbation theory. The cross section at parton level depends mainly on the parton flux $p(\hat{s}) \propto 1/\hat{s} = 1/(x_a x_b s)$ and the *matrix-element* (ME) $\mathcal{M}_{ab \rightarrow X}$ squared. The ME is evaluated by summing over Feynman diagrams. Hence, gener-

ators used to simulate the hard interaction are also referred to as *matrix-element generators*. Given the aforementioned PDFs and the partonic cross section

$$\hat{\sigma}_{ab \rightarrow X} \propto p(\hat{s}) \cdot |\mathcal{M}_{ab \rightarrow X}|^2 \quad \text{with} \quad \mathcal{M}_{ab \rightarrow X} = \sum_i \mathcal{M}_{ab \rightarrow X}^i, \quad (3.1)$$

the inclusive proton-proton cross section $\sigma_{pp \rightarrow X}(s)$ can be calculated as a function of the center-of-mass energy \sqrt{s} :

$$\sigma_{pp \rightarrow X}(s) = \sum_{a,b \in \{q, \bar{q}, g\}} \int_0^1 \int_0^1 dx_1 dx_2 \int f_a(x_1, Q^2) f_b(x_2, Q^2) d\hat{\sigma}_{ab \rightarrow X}(\hat{s}, Q^2). \quad (3.2)$$

For the calculation of the partonic cross section, in principle all existing Feynman diagrams would need to be taken into account to compute the matrix element. As the number of Feynman diagrams is infinite, the cross section is determined using a perturbation series of the coupling constant. The *leading-order* (LO) cross section includes only Feynman diagrams that have the smallest possible number of vertices, also referred to as *tree-level* Feynman diagrams. In the *next-to-LO* (NLO) cross section, the tree-level contribution and virtual loop corrections, as well as corrections due to real emission of an additional parton are considered. These calculations often provide a sufficiently accurate description of the momenta of the final state particles, although higher-order processes can still have a significant impact on physical observables.

Accompanying particles caused by additional radiation in the initial state (*initial-state radiation* (ISR)) or the final state (*final-state radiation* (FSR)) are modeled by dedicated *parton shower* (PS) algorithms, which are interfaced with the ME generator. The commonly used PS algorithms mainly differ in their shower ordering scheme, which can be based, for example, on the dipole mass, angular requirements, or transverse momentum requirements. With decreasing energy in the PS, the strong coupling α_s gets larger and, therefore, the rate of additional radiation increases until the energy reaches the QCD energy scale Λ_{QCD} , where particles start to confine to hadrons. This procedure is denoted as *hadronization*. In this non-perturbative regime, two different phenomenological models are typically used to simulate the hadronization. The string model performs a direct transition of the parton system into hadrons, while the cluster model introduces an intermediate step where objects with a mass of order 1 GeV are clustered before the hadron states are formed.

The event activity, which takes place in addition to the hard interaction, is denoted as the *Underlying Event* (UE). Typical processes are gluon-gluon scattering, interactions of the colored proton remnants, and further parton interactions referred to as *multiple parton interactions* (MPIs). In particular, MPIs induce a significant amount of soft particles which need to be considered. The UE activity is described in terms of empirical models which are tuned on the basis of collision data.

Additional proton-proton interactions that occur in the bunch crossing of interest are referred to as *pileup* (PU). This effect is taken into account in simulation, however, the number of PU

vertices depends strongly on the beam properties, which vary over time. Typically a correction is applied which is derived directly from data.

The interaction of generated particles with the detector material is simulated in the final step of the simulation procedure using the GEANT4 package [104]. The detector response in the different sub-detector components is modeled based on the implemented CMS detector geometry. The resulting detector signals are propagated to the event reconstruction algorithm described in Chapter 4. In this way, simulated events provide the same set of information as the data events. Furthermore, generator information can be added to compare the kinematic properties of particles in the event before and after reconstruction. This allows the unfolding of detector effects from the underlying physics.

At the quantum level physics becomes non-deterministic, and the results of experiments need to be interpreted in terms of statistical methods. Thus, the generation of events is performed using *Monte Carlo* (MC) techniques [54, 105–107] to produce a large number of collision events following the previously described procedure. The three most important MC event generators used in this thesis are briefly described in the following.

MADGRAPH5_AMC@NLO [108], which combines the event generators MADGRAPH5 [109] and AMC@NLO [110], provides a framework for ME calculations at NLO precision as well as matching of the ME and the PS simulation. At NLO, possible double counting of ME and PS contributions is prevented by introducing a subtraction scheme, which leads to negative event weights for a small fraction of events. Given a certain initial and final state, all possible Feynman diagrams are automatically calculated.

POWHEG [110–112] computes MEs at NLO precision in QCD. Since the hardest radiation is determined first, and the subsequent PS calculations are ordered by p_T , any overlap between NLO and LO calculations is avoided. Thus, no subtraction scheme needs to be applied, and all event weights are positive by definition. In contrast to MADGRAPH5_AMC@NLO, only predefined processes can be generated.

PYTHIA [113] is a general-purpose event generator, which provides in addition to ME calculations also a framework specialized to simulate the PS, the hadronization and the UE. Since the ME is only computed at LO accuracy, PYTHIA is often used for the PS and the subsequent simulation steps, when it is interfaced with NLO event generators.

The generator event weights are applied to all simulated events independent of the choice of the event generator. In addition, simulated events are scaled to the integrated luminosity corresponding to the recorded dataset. The per-event weight w_{event} is defined as

$$w_{\text{event}} = \sigma \cdot \mathcal{B} \cdot \frac{\mathcal{L} \cdot w_{\text{gen}}}{\sum_i w_{\text{gen},i}} \quad (3.3)$$

in terms of the cross section σ times branching fraction \mathcal{B} of the simulated process, the corresponding integrated Luminosity \mathcal{L} , and the per-event generator weight w_{gen} normalized to the sum of generator weights over all events i .

3.2 Simulated events

The samples of simulated events which are considered in this thesis are simulated using the MADGRAPH5_AMC@NLO, or the POWHEG (v2) ME event generators. The ME events are then interfaced with PYTHIA (v8.2 or greater) to model the PS and hadronization of partons in the initial and final states along with the UE description and the decay of the Higgs boson. The UE is modeled using the CUETP8M1 tune [114] in 2016 and the CP5 tune [115] in 2017 and 2018. Events simulated at NLO(LO) in QCD from ME calculations characterized by different parton multiplicities are merged using the Fx-Fx(MLM) prescription [116, 117]. The PDFs are obtained from the NNPDF [118, 119] set v3.0 for 2016 (NLO) and v3.1 for 2017 and 2018 (NNLO).

The main Higgs boson production modes are listed in Table 3.1 ordered according to their production cross section. The signal simulation is normalized to the production cross section times branching fraction $\mathcal{B}(H \rightarrow \mu^+\mu^-)$ as predicted by the SM based on the most accurate computation obtained from Ref. [120]. The dominant Higgs boson production at the LHC via ggH is simulated at NLO accuracy in QCD using the MADGRAPH5_AMC@NLO v2.4.2 event generator which includes up to two additional quarks or gluons in the final state at ME level. The corresponding cross section is determined at *next-to-next-to-NLO* (N3LO) accuracy in QCD, and at NLO in electroweak theory [121]. The transverse momentum of the generated Higgs boson produced in ggH events is reweighted to reflect the POWHEG NNLOPS prediction [122, 123]. The VBF, the WH/ZH, and the t \bar{t} H production are simulated using POWHEG v2.0 [124, 125] at NLO accuracy in QCD. The cross sections are computed at *next-to-NLO* (NNLO) or NLO in QCD, including NLO electroweak corrections. Apart from the five main Higgs boson production modes, which already have been experimentally observed, also yet undiscovered Higgs boson production modes are considered according to their SM prediction. The b \bar{b} H process is generated using POWHEG at NLO accuracy in QCD, and the Higgs boson production in association with a single top quark via tHW and the t-channel tHq are simulated with the MADGRAPH5_AMC@NLO event generator. Both cross sections are determined at NLO in QCD

Table 3.1: Higgs boson production modes with their corresponding cross sections at a given order in perturbative QCD and electroweak theory. For each production mode the used ME generator is listed. In addition, the theoretical uncertainties in the QCD scale and the combined uncertainty in the PDF and the strong coupling α_s are reported.

Process	σ (pb)	Perturbative Order		Generator	Theo. uncertainties	
		QCD	electroweak		\pm QCD scale (%)	\pm (PDF+ α_s) (%)
ggH	48.58	N3LO	NLO	MADGRAPH5_AMC@NLO	+4.6 / -6.7	+3.2 / -3.2
VBF	3.78	NNLO	NLO	POWHEG	+0.4 / -0.3	+2.1 / -2.1
WH	1.37	NNLO	NLO	POWHEG	+0.5 / -0.7	+1.9 / -1.9
ZH	0.88	NNLO	NLO	POWHEG	+0.5 / -0.6	+1.9 / -1.9
t \bar{t} H	0.51	NLO	NLO	POWHEG	+5.8 / -9.2	+3.6 / -3.6
b \bar{b} H	0.49	NLO	—	POWHEG	+20.2 / -23.9	—
tHq	0.074	NLO	—	MADGRAPH5_AMC@NLO	+6.5 / -14.9	+3.7 / -3.7
tHW	0.015	NLO	—	MADGRAPH5_AMC@NLO	+4.9 / -6.7	+6.3 / -6.3

without including higher order electroweak corrections. Further minor contributions from single top quark associated Higgs boson production via s-channel $t\bar{H}b$ or from Higgs boson production with W and Z bosons in association with a top quark-antiquark pair $t\bar{t}WH/t\bar{t}ZH$ are negligible and, therefore, not considered in this thesis. The partial decay width for the Higgs boson decay to muons is computed using `HDECAY` [126, 127] at NLO in QCD and electroweak theory.

An overview of the simulated background processes is given in Table 3.2. They are categorized in several groups of processes with similar kinematic properties. Background processes involving the production of top quarks are the top quark pair production $t\bar{t}$, the single top quark production (in association with a W boson, t-channel, s-channel) as well as processes including multiple top quarks. They are simulated at NLO precision in QCD using either `POWHEG` or `MADGRAPH5_AMC@NLO`. The corresponding cross sections are taken from the `TOP++ v2.0` [128] and `HATOR` [129] predictions, computed at NNLO or NLO in QCD. Contributions from the $t\bar{t}Z$, tZ , $t\bar{t}W$, and $t\bar{t}WW$ are simulated at NLO or LO accuracy in QCD using the `MADGRAPH5_AMC@NLO` generator. The $t\bar{t}Z$ cross section is taken from Ref. [130]. The Drell-Yan production is simulated using `MADGRAPH5_AMC@NLO` at NLO precision in QCD. Since the off-shell production of the Z boson is of particular interest for this analysis, a sample of simulated events was produced, which only includes events with a muon pair in the mass range $105 < m_{\mu\mu} < 160 \text{ GeV}$. The large number of simulated events in this phase space significantly improves the statistical precision of the background prediction. The corresponding cross section is calculated with `FEWZ v3.1` [131] at NNLO in QCD and NLO precision in electroweak theory. In addition, the electroweak production of a Z boson in association with two jets is simulated at LO in QCD using `MADGRAPH5_AMC@NLO`. The diboson processes WZ , ZZ , and WW , as well as the triboson processes WWW , WWZ , WZZ , and ZZZ are simulated at NLO in QCD using either the `POWHEG` or `MADGRAPH5_AMC@NLO` generators. The diboson processes which only include leptonic decays are listed separately. Their production cross sections are modified with NNLO/NLO k factors taken from Refs. [132], [133], and [134].

Table 3.2: Overview of simulated background processes. Processes with similar kinematic properties are summarized. For each simulated process the used ME generator and the cross section times branching fraction is reported.

Notation	Process	Generator	$\sigma(\text{pb}) \times \mathcal{B}$
Top quark	$\bar{t}\bar{t}$ + jets (dileptonic decays)	POWHEG	86.61
	$\bar{t}\bar{t}$ + jets (semileptonic decays)	POWHEG	358.57
	Single top quark (anti-quark) W-associated	POWHEG	35.90
	Single top quark t-channel	POWHEG	136.02
	Single top anti-quark t-channel	POWHEG	80.95
	Single top quark (anti-quark) s-channel	MADGRAPH5_AMC@NLO	3.40
	tttt (inclusive decays)	MADGRAPH5_AMC@NLO	$9.1 \cdot 10^{-3}$
	tttW (inclusive decays)	MADGRAPH5_AMC@NLO	$7.3 \cdot 10^{-4}$
	tttq (inclusive decays)	MADGRAPH5_AMC@NLO	$4.0 \cdot 10^{-4}$
$\bar{t}\bar{t}Z$ + tZ	$\bar{t}\bar{t}Z$ ($Z \rightarrow \ell\bar{\ell}, \nu\bar{\nu}$ $M_{\ell\bar{\ell}} > 10 \text{ GeV}$)	MADGRAPH5_AMC@NLO	0.32
	$\bar{t}\bar{t}Z$ ($Z \rightarrow \ell\bar{\ell}, \nu\bar{\nu}$ $M_{\ell\bar{\ell}} < 10 \text{ GeV}$)	MADGRAPH5_AMC@NLO	$9.3 \cdot 10^{-2}$
	tZq ($Z \rightarrow \ell\bar{\ell}$)	MADGRAPH5_AMC@NLO	$7.6 \cdot 10^{-2}$
	tZW (leptonic decays)	MADGRAPH5_AMC@NLO	$1.3 \cdot 10^{-3}$
$\bar{t}\bar{t}W$ (W)	$\bar{t}\bar{t}W$ ($W \rightarrow \ell\nu$)	MADGRAPH5_AMC@NLO	0.20
	$\bar{t}\bar{t}WW$ (inclusive decays)	MADGRAPH5_AMC@NLO	$7.0 \cdot 10^{-3}$
DY+Z-EWK	$Z \rightarrow \ell\bar{\ell}$ + jets ($M_{\ell\bar{\ell}} > 50 \text{ GeV}$)	MADGRAPH5_AMC@NLO	6225.40
	$Z \rightarrow \ell\bar{\ell}$ + jets ($105 < M_{\ell\bar{\ell}} < 160 \text{ GeV}$)	MADGRAPH5_AMC@NLO	47.12
	$Z \rightarrow \ell\bar{\ell}(q\bar{q})$ ($M_{\ell\bar{\ell}} > 50 \text{ GeV}$, $M_{q\bar{q}} > 120 \text{ GeV}$)	MADGRAPH5_AMC@NLO	1.61
WZ/ZZ lep.	WZ ($Z \rightarrow \ell\bar{\ell}, W \rightarrow \ell\nu$)	MADGRAPH5_AMC@NLO	5.25
	$q\bar{q} \rightarrow ZZ$ ($Z \rightarrow \ell\bar{\ell}$)	POWHEG	1.26
	$gg \rightarrow ZZ \rightarrow 2e2\mu$ ($2\mu2\tau$) ($2e2\tau$)	PYTHIA	$3.2 \cdot 10^{-3}$
	$gg \rightarrow ZZ \rightarrow 4\mu$ (4τ)	PYTHIA	$1.6 \cdot 10^{-3}$
VV(V)	WW ($W \rightarrow \ell\nu$)	POWHEG	12.18
	WZ ($Z \rightarrow \ell\bar{\ell}, W \rightarrow q\bar{q}$)	MADGRAPH5_AMC@NLO	5.61
	ZZ ($Z \rightarrow \ell\bar{\ell}, Z \rightarrow q\bar{q}$)	POWHEG	3.66
	ZZ ($Z \rightarrow \ell\bar{\ell}, Z \rightarrow \nu\bar{\nu}$)	POWHEG	0.56
	WWZ (inclusive decays)	MADGRAPH5_AMC@NLO	0.17
	WWW (inclusive decays)	MADGRAPH5_AMC@NLO	0.21
	WZZ (inclusive decays)	MADGRAPH5_AMC@NLO	$5.6 \cdot 10^{-2}$
	ZZZ (inclusive decays)	MADGRAPH5_AMC@NLO	$1.4 \cdot 10^{-2}$

Event reconstruction

The analysis described in this thesis relies on the efficient reconstruction of the particles produced in the proton-proton collisions. Each of these particles leaves a characteristic trace in the detector that is used to identify the different particle types and to measure their properties. The *particle-flow* (PF) algorithm [135–137] is a global event reconstruction method that combines information from the different subdetector components of the CMS detector to identify muons, electrons, photons, as well as neutral and charged hadrons.

The different detector signatures are briefly presented in Section 4.1. The event reconstruction and particle identification based on the PF algorithm is introduced in Section 4.2. The reconstructed physics objects and specific selection requirements used for data analysis are reported in Section 4.3.

4.1 Detector signatures

Charged-particle trajectories, referred to as *tracks*, are reconstructed from the signals in different layers of the silicon-tracking detector. The interaction point, identified to be the origin of the high- p_T particles in the event, is referred to as the *primary vertex* (PV). A less event dependent definition of the interaction point is obtained by the average position of the PV over many events. The corresponding luminous region in the detector is referred to as the *beam spot* (BS).

Electrons and photons are absorbed in the ECAL as they deposit most of their energy due to multiple electromagnetic interactions. These electromagnetic showers are detected as energy clusters in the ECAL crystals from which the particle energies are determined. Since electrons are charged, they are identified by a track in the silicon tracker pointing to an energy cluster in the ECAL. Photons are electrically neutral particles, and thus do not interact with the tracking detector. Hence, energy clusters without associated tracks are identified as photons. Hadrons additionally lose energy through strong interactions, initiating hadronic showers in the HCAL. This leads to a characteristic signature used to identify charged and neutral hadrons. Muons at high momentum are minimum ionizing particles so they are the only charged SM particles that typically pass through the calorimeters with little or no interactions and produce a signal in the muon chambers. Neutrinos only interact weakly, therefore they escape the detector undetected. An illustration of the described detector signatures is given in Fig. 4.1.

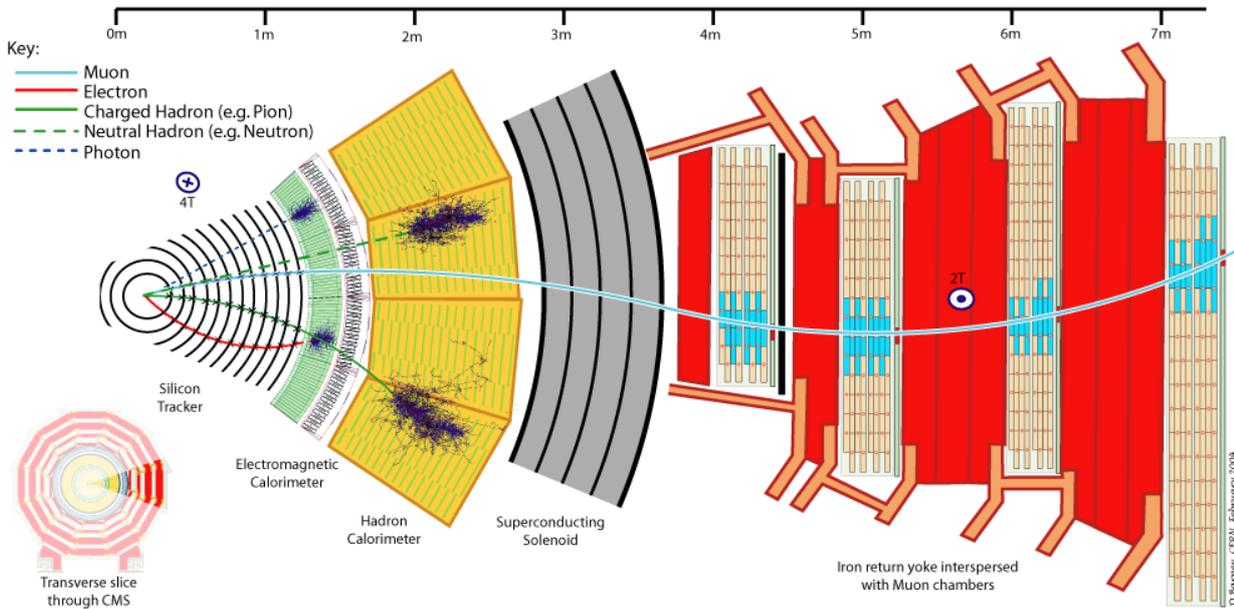


Figure 4.1: Cross-sectional view of a slice through the CMS detector. Typical detector signatures of electrons, muons, photons, neutral and charged hadrons are presented in the individual sub-detector systems. Taken from Ref. [138].

4.2 Event reconstruction with the particle-flow algorithm

The CMS experiment follows the *particle-flow* (PF) approach for particle reconstruction and identification. Particles can be well distinguished from each other due to their specific signatures in the different detector components. The PF algorithm significantly improves the event description by combining information from all sub-detector components. Figure 4.2 illustrates the principle of PF event reconstruction. The detector signatures in the sub-detectors are combined and interpreted to identify different particle types.

Reconstructed tracks are used as inputs, denoted as *PF elements*, to the PF algorithm. When a particle passes through layers of the pixel and strip tracking detectors a trajectory is built from the combination of associated signal hits. The tracking procedure is divided into the track seed generation, the track finding, the track fitting and the final track selection. In the high PU regime with large track multiplicities, a misreconstruction of unrelated hits becomes more probable. To provide high reconstruction efficiencies and low misreconstruction rates, the CMS tracking software denoted as *Combinatorial Track Finder* (CTF) [139] exploits an iterative tracking strategy. The CTF is a modified version of a combinatorial Kalman filter [140–142] which provides a combined framework for pattern recognition and track fitting. Initial track candidates, called seeds, are built from short tracks in the inner layers of the silicon-tracker. These are required to meet tight selection criteria on the distance to the PV, for example. In the next iteration, hits assigned to the tracks from the previous iteration are removed. The corresponding track seeding criteria are progressively relaxed, which increases the tracking ef-

efficiency. The hit removal reduces the possible hit combinations, which leads to a decreasing rate of misidentified tracks. The seeded trajectories are extrapolated along their expected path to find additional signal hits which can be used to build track candidates. To improve the precision of the track parameters, all reconstructed track candidates are refitted. For this purpose, a Kalman filter [143] and smoother are employed. The information of the BS position used for seeding is removed in this step. Finally, a track selection is applied to increase the quality of reconstructed tracks.

The second type of PF elements are energy deposits in the ECAL and HCAL calorimeters. A clustering algorithm is employed to detect and measure the energy and direction of stable neutral particles, such as photons or neutral hadrons. Furthermore, information from the ECAL and the HCAL in combination with information from the tracking detectors is used to identify charged hadrons and electrons. The cluster algorithm is performed in each of the calorimeter components separately. Energy deposits in localized calorimeter cells that meet the required energy threshold are considered as seeds of the cluster. Neighboring cells above a certain energy threshold are added to the cluster in an iterative procedure.

The PF reconstruction proceeds with a link algorithm that combines the PF elements from different subdetector components. To reduce the number of possible combinations of elements the link is restricted to the closest object in the (η, ϕ) plane. For each pair of elements the distance between them is determined and used to evaluate the quality of the link. A reconstructed track is linked to calorimeter clusters by extrapolating the track to the clusters. Electrons often emit significant bremsstrahlung along their trajectory. To account for the corresponding energy loss, tangents to the tracks are extrapolated to the ECAL from the intersection points between the track and each of the tracker layers. An energy cluster is linked to the track and considered to be a potential bremsstrahlung photon if the extrapolated tangents point to an ECAL cluster. Calorimeter links between HCAL clusters and ECAL clusters are established when the cluster position in the more granular calorimeter matches the cluster in the less granular calorimeter. A charged-particle track is linked to a track in the muon chambers if the corresponding track fit fulfills certain quality criteria.

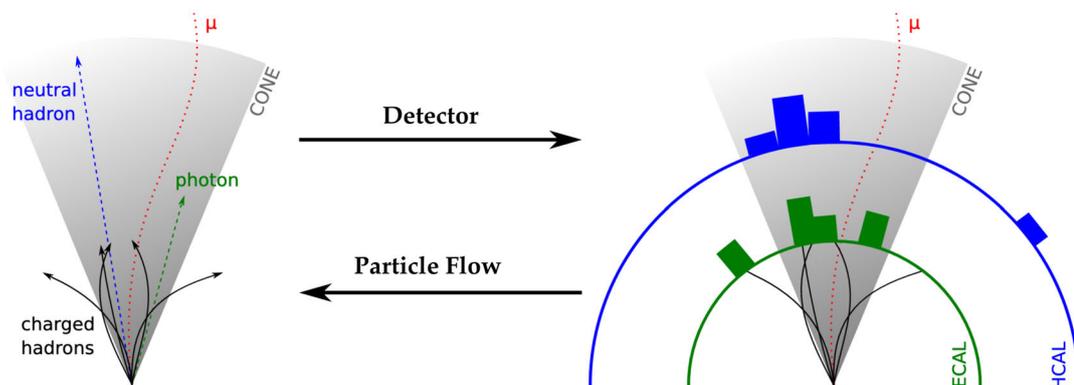


Figure 4.2: Illustration of the the PF event reconstruction. The PF algorithm combines information provided by the ECAL, HCAL, the silicon tracker, and muon chambers to reconstruct neutral and charged hadrons, electrons, muons or photons. Modified from Ref. [144].

The linked information are referred to as PF *blocks*. These blocks are masked against further processing, which means they are not used as elements for other particles. Finally, the event is reconstructed with the following global approach: First, tracks associated to an identified muon are removed from the blocks. Second, the track and ECAL cluster associated to an electron, as well as the ECAL cluster associated to an isolated photons, are removed from the blocks. Lastly, the remaining elements to be identified are charged and neutral hadrons as well as non-isolated photons from decays of neutral pions. These particles typically correspond to the experimental signature of highly-energetic quarks or gluons, which hadronize and produce a collimated bunch of particles. Recombination algorithms are employed to combine these particles into a common physics object, denoted as a jet.

4.3 Selection of physics objects

The PF algorithm allows for data analysis at the level of reconstructed physics objects identified as muons, electrons, jets, or missing transverse momentum from the presence of neutrinos. These physics objects are of particular interest as they potentially emerge from the decay of the Higgs boson or heavy particles produced in association with the Higgs boson.

In the following, the reconstruction and selection of physics objects relevant for this thesis is discussed. Note that from now on only muons and electrons are referred to as leptons. The hadronic decay of the τ -lepton is not considered by the physics objects selection in this thesis.

4.3.1 Primary vertex and beam spot

The vertex reconstruction [139, 145] aims to measure the interaction point of all proton-proton interactions in the event based on the reconstructed tracks. In a first step, tracks are pre-selected by imposing a set of quality criteria on the number of hits in the tracking layers, and the fit quality of the tracks, for example. All preselected tracks, which appear to have the same origin, are clustered by a *deterministic annealing* algorithm [146] to build vertex candidates. Thereby it is important to keep balance between resolving closeby interaction vertices and splitting a single, genuine vertex by mistake. The reconstructed vertices are fitted using the *adaptive vertex fitter* [147], which assigns a weight $w_i \in [0, 1]$ for each track i , to estimate the position and determine quality measures of the vertices. A large weight corresponds to a high compatibility between the track and the vertex position. The reconstruction efficiency for vertices built from more than two tracks is close to 100%. The vertex with the largest square sum of the transverse momentum of all tracks $\sum p_{\perp}^2$ is identified as the PV.

The *beam spot* (BS) [139] describes the luminous region where the LHC beams collide in the CMS detector. The position of the BS is important to estimate the position of the interaction point for vertex reconstruction as well as for triggering purposes. Two independent methods can be used to determine the center of the BS. The first method defines a 3D region from the position of primary vertices over many collisions. The mean position of x, y and z, as well as the size of the BS can be derived through a fit of a likelihood to the 3D profile. The second method

exploits a correlation between the transverse impact parameter d_0 and ϕ of tracks belonging to the PV. This provides consistent results with respect to the first method.

The analysis described in this thesis uses additional selection criteria on the quality of the PV. Events are selected, if at least one PV meets the following quality requirements:

- $|z| < 24$ cm distance between PV and the center of the detector,
- $|\rho_{xz}| < 2$ cm distance between PV and the center of the detector,
- number of free parameters $n_{\text{dof}} = 3 - 2 \sum_{i=1}^{N_{\text{track}}} w_i > 4$.

4.3.2 Muons

The reconstruction of muons [148] is based on tracks in the inner tracking system and tracks in the muon chambers with the aim to combine both track pieces using a Kalman filter technique. Depending on the reconstruction method, the resulting muon candidate is referred to as a *global muon* or *tracker muon*. These two types of reconstructed muons are distinguished as follows:

Global muons (outside-in) The matching of tracks from the muon chamber to the inner tracks is performed by comparing the track parameters. If both sets of track parameters match within certain boundaries, the combined track is re-fitted and the muon candidate is identified as a global muon.

Tracker muons (inside-out) Tracks in the inner tracking system are extrapolated to the muon chambers. In this procedure, the magnetic field, the average energy loss, and multiple scattering in the detector material are taken into account. The muon candidate is reconstructed as tracker muon if at least one muon track segment matches the inner track.

The efficiency for prompt muons to be reconstructed as a global muon or tracker muon is about 99%. In many cases muon candidates are reconstructed as both. The momentum resolution of the track is dominated by the inner track measurement up to $p_T < 200$ GeV. For very high- p_T muons the combined track fit of the inner track with the track in the muon chambers improves the momentum resolution.

The reconstructed muons are required to meet certain criteria in order to be selected for the analysis presented in this thesis. These criteria are summarized in Table 4.1. The muon selection is particularly designed to ensure a high efficiency of well-reconstructed muons while keeping the misidentification rate at a low level. Muons are required to be reconstructed as a global or tracker muon with transverse momentum $p_T > 20$ GeV inside the detector acceptance of $|\eta| < 2.4$. Additional quality criteria are imposed on these muons, which are collectively referred to as the *medium ID* [148] within the CMS Collaboration. These are based on the number of measured hits in the silicon tracker and in the muon system, the properties of the fitted muon track, and the impact parameters of the track with respect to the primary vertex of the

event. Beyond these standard selection criteria, more specific vertex requirements are employed to select prompt muons created in the hard process and to reject those from decays of hadrons. Furthermore, particles surrounding the muons are required to be incompatible with the hypothesis of a b hadron decay, based on the algorithm described in Section 4.3.6.

To ensure that the efficiency correction factors for the selected trigger can be applied without bias, the muon candidate which triggered the event is required to meet the selection criteria denoted as the *tight ID* [148]. These include, in particular, an increased minimal number of signal hits in the inner tracking detector and the muon chambers, as well as additional quality criteria on the compatibility of the muon track with the reconstructed PV.

Table 4.1: Summary of requirements used to select muons for data analysis. The b hadron likelihood is explained in Section 4.3.6.

Muon selection criteria	
Transverse momentum	$p_T > 20 \text{ GeV}$
Pseudorapidity	$ \eta < 2.4$
Muon identification	medium ID (tight ID for trigger muon)
Distance between muon track and PV	
in the transverse plane	$d_{xy} < 0.05 \text{ cm}$
in longitudinal direction	$d_z < 0.1 \text{ cm}$
Significance of 3D impact parameter	$\text{SIP}_{3D} < 8$
B hadron likelihood of closest jet	$\text{deepCSV}_{\text{closest}} < 0.8958$ in 2016 $\text{deepCSV}_{\text{closest}} < 0.8001$ in 2017/18

4.3.3 Electrons

Electrons are reconstructed [149, 150] by a track in the inner tracker and an associated energy cluster in the ECAL. Due to their small mass, electrons are highly prone to emit bremsstrahlung, while crossing the beam pipe and the layers of the inner tracking detector. The radiation of bremsstrahlung photons subsequently reduces the energy of electrons along their trajectory and changes the curvature of the track. The average energy loss ranges from 33% in the center of the detector to 85% in the forward region. A *Gaussian Sum Filter* (GSF) [151, 152], which embeds a dedicated model of the energy loss in the track fitting procedure, is employed to provide an improved measurement of the electron track parameters. The radiated photons are mainly spread in azimuthal direction. Therefore, the ECAL clustering algorithm collects the energy deposits induced by bremsstrahlung photons in a small area in η and an extended window in ϕ together with the primary energy deposit of the electron. These energy deposits are used to construct the *supercluster* (SC). The seeding of the electron reconstruction follows two complementary approaches. The ECAL-driven approach extrapolates the electron trajectory towards the interaction point. This backward propagation is based on the SC energy and its energy-weighted average position, which intersects with the helix corresponding to the initial electron energy. Then, both charge hypotheses are tested when the helices are tried to be

matched to a track in the silicon tracker. The tracker-driven seeding is particularly optimized for very soft electrons which are surrounded by additional particles, i.e., for non-isolated electrons. The track is extrapolated towards the ECAL and matched to the closest energy cluster.

Selection criteria are employed to increase the purity of electrons used for data analysis presented in this thesis. Table 4.2 provides an overview of all requirements. Electrons are selected in the detector acceptance of $|\eta| < 2.5$ with a transverse momentum of $p_T > 20 \text{ GeV}$. Furthermore, requirements are imposed on the cluster shape together with criteria on the quality of the matching between the SC and the extrapolated track. Furthermore, the ratio between the deposited energy in the HCAL and the ECAL is restricted to small values to reject charged hadron background. Selected electrons are required to provide high consistency of the electron energy measurement based on the ECAL and the momentum measured in the tracker. At maximum one missing hit is allowed for electrons reconstructed in the barrel region of the detector ($|\eta| < 1.4442$). A conversion veto is applied to reject electrons originating from the conversion of a photon. The same selection criteria are applied, which are used for muons to suppress background originating from the decay of b hadrons.

Table 4.2: Summary of requirements used to select electrons for data analysis. The first (second) term in parenthesis corresponds to the value used for electrons reconstructed in the barrel ($0 < |\eta| < 1.4442$ (endcap $1.556 < |\eta| < 2.5$)) region of the detector.

Electron selection criteria	
Transverse momentum	$p_T > 20 \text{ GeV}$
Pseudorapidity of SC	$ \eta_{\text{SC}} < 1.4442, 1.556 < \eta_{\text{SC}} < 2.5$
Pass conversion veto	✓
Distance between muon track and PV	
in the transverse plane	$d_{xy} < 0.05 \text{ cm}$
in longitudinal direction	$d_z < 0.1 \text{ cm}$
Significance of 3D impact parameter	$\text{SIP}_{3\text{D}} < 8$
	$\sigma_{i\eta i\eta} < (0.011, 0.030)$
Energy cluster	$ \Delta\eta(\text{SC}, \text{track}) < (0.010, 0.008)$ $ \Delta\phi(\text{SC}, \text{track}) < (0.04, 0.07)$
Hadronic energy/EM energy	$H/E < (0.10, 0.07)$
Number of missing hits	$N_{\text{missing}}^{\text{hits}} < 2$
Energy in SC & track p_T	$-0.05 < (1/E_{\text{SC}} - 1/p_{\text{track}}) < (0.010, 0.005)$
B hadron likelihood of closest jet	$\text{deepCSV}_{\text{closest}} < 0.8958$ in 2016 $\text{deepCSV}_{\text{closest}} < 0.8001$ in 2017/18

4.3.4 Lepton isolation

Leptons in signal events are expected to be isolated. The so-called mini-isolation I^{mini} [153] of a lepton is defined as the sum of scalar p_T values of charged particles, neutral hadrons, and photons, that are reconstructed within a narrow cone with radius R centered around the lepton direction not taking into account the presumed lepton itself. Residual PU contributions to the neutral component of the isolation sum are taken into account using the *effective area* (EA) \mathcal{A} correction [154], which is measured for electrons and muons in several η regions. The lepton isolation is given by

$$I^{\text{mini}} = \sum_{\text{charged}} p_T + \max\left(0, \sum_{\text{neutrals}} p_T - \rho \mathcal{A}\left(\frac{R}{0.3}\right)^2\right)$$

where ρ is the energy density of neutral particles reconstructed within the geometric acceptance of the tracking detectors.

To increase the efficiency of leptons reconstructed in events with high hadronic event activity, in which accidental overlap of particles is more probable, the cone size R decreases dynamically with the p_T of the lepton:

$$R = \begin{cases} 0.05 & \text{if } p_T > 200 \text{ GeV} \\ 10 \text{ GeV}/p_T & \text{if } 50 < p_T < 200 \text{ GeV} \\ 0.2 & \text{if } p_T < 50 \text{ GeV.} \end{cases}$$

Muons and electrons are required to have an isolation of less than 40% relative to their p_T . Again, a tighter selection is imposed in case the muon triggered the event. The isolation is required to be less than 15% relative to the muon p_T inside a cone with fixed radius $R = 0.4$.

4.3.5 Hadronic jets

Quarks and gluons in the final state confine to hadrons on a very short time scale, which can be estimated by the inverse of the QCD energy scale $\Delta\tau \approx \Lambda_{\text{QCD}}^{-1}$. Due to their nearly immediate hadronization, they can not be observed as isolated particles in the detector. The bunch of collimated particles, which is produced during the hadronization, consists on average of 65% charged hadrons, 25% photons, and 10% neutral hadrons. These constituents are reconstructed by the PF algorithm and clustered with specific sequential jet recombination algorithms. All jet algorithms are required to be *infrared and collinear safe* (IRC-safe), i.e., they are insensitive to soft gluon radiation and gluon-splitting, since in these cases perturbative QCD calculations diverge.

The reconstruction of jets relies on certain distance measures between particles, which are used to decide if a particle is assigned to a jet. The distance parameter d_{ij} between the particles i and j is defined by

$$d_{ij} = \min\left(p_{T,i}^{2p}, p_{T,j}^{2p}\right) \frac{\Delta_{ij}^2}{R^2} \quad (4.1)$$

in terms of the transverse momentum of the jet candidate particle p_T , the cone radius of the jet R , a free parameter of the algorithm p and the distance $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$. The jet reconstruction follows an iterative procedure. The parameter d_{ij} is evaluated for all particle pairs, and the pair with the smallest distance d_{ij} is merged into a temporary pseudo-jet, which is used in the following iteration until no d_{ij} can be found smaller than $p_{T,i}^{2p}$. The pseudo-jet i is removed from the set of particles and is considered as the final jet. The parameter p determines the properties of the reconstructed jets and, in particular, the geometrical shape of jets in the detector. Three classes of IRC-safe jet reconstruction algorithms can be identified: The k_T algorithm [155] corresponding to $p = 1$, the *Cambridge-Aachen* algorithm [156] corresponding to $p = 0$, and the *anti- k_T* algorithm [157] corresponding to $p = -1$. Figure 4.3 shows the resulting reconstructed jets using the anti- k_T algorithm.

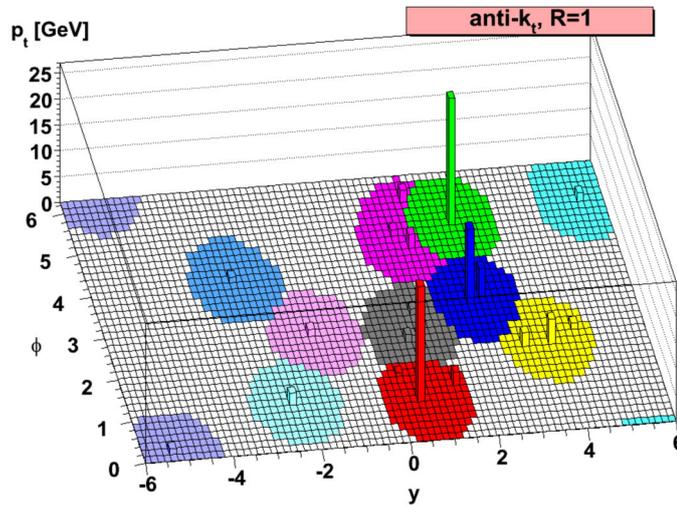


Figure 4.3: Reconstructed jets in simulated events using the anti- k_T algorithm. Shown are the transverse momenta of particles in the $y - \phi$ plane and their combination into a jet with cone radius $R = 1$. Taken from Ref. [157].

In this thesis the anti- k_T algorithm with a cone size $R = 0.4$ is used. Particles from PU vertices can have a significant impact on the jet clustering. To reduce their influence, charged hadrons which are identified to originate from a PU vertex are excluded from the jet clustering. The jet momentum is determined from the vectorial momentum sum of all associated particles. Jets are required to be reconstructed with $p_T > 25 \text{ GeV}$ inside the detector acceptance of $|\eta| < 4.7$. They are rejected if they overlap with a selected lepton by a geometrical separation requirement of $\Delta R(j, \ell) > 0.4$. Furthermore, a set of quality criteria is used to suppress the contamination from misreconstructed jets arising from calorimeter or electronic noise. To distinguish noise jets from genuine jets, a set of η -dependent selection criteria is imposed on the hadronic and electromagnetic energy fractions carried by neutral particles, the hadronic energy fraction carried by charged hadrons, the total number of constituents as well as the multiplicities of charged and neutral particles [158]. These selection criteria are reported in Table 4.3.

To further reduce the contamination of jets originating from a PU vertex, a dedicated *pile up identification* (PU ID) discriminator is built based on a combination of vertex and jet shape information [159, 160]. All jets with $p_T < 50$ GeV are required to pass the *loose* PU ID. The detector region $2.6 < |\eta| < 3.0$ was strongly affected by ECAL noise during the 2017 data-taking period. This led to a significantly increased number of noise jets in the corresponding region. Jets from noise or PU are typically more diffuse compared to genuine hadronic jets. To mitigate the impact of jet misreconstruction due to ECAL endcap noise in 2017, jets that lie in the region are required to pass the *tight* PU ID.

Table 4.3: Summary of requirements used to select jets for data analysis. The requirements depend on the year of data taking, and on the pseudorapidity of the jet. The jet ID is particularly designed to reduce jets originating from detector noise.

Jet ID in 2016	$ \eta \leq 2.6$	$2.4 < \eta \leq 2.7$	$2.7 < \eta \leq 3.0$	$ \eta > 3.0$
Number of constituents	> 1	> 1	—	—
Neutral hadron fraction	< 0.9	< 0.9	< 0.98	—
Neutral EM fraction	< 0.9	< 0.9	> 0.01	< 0.9
Number of neutral particles	—	—	> 2	> 10
Charged hadron fraction	> 0	—	—	—
Charged EM fraction	< 0.99	—	—	—
Charged multiplicity	> 0	—	—	—
Jet ID in 2017	$ \eta \leq 2.4$	$2.4 < \eta \leq 2.7$	$2.7 < \eta \leq 3.0$	$ \eta > 3.0$
Number of constituents	> 1	> 1	—	—
Neutral hadron fraction	< 0.9	< 0.9	—	> 0.2
Neutral EM fraction	< 0.9	< 0.9	> 0.02 and < 0.99	< 0.9
Number of neutral particles	—	—	> 2	> 10
Charged hadron fraction	> 0	—	—	—
Charged multiplicity	> 0	> 0	—	—
Jet ID in 2018	$ \eta \leq 2.6$	$2.6 < \eta \leq 2.7$	$2.7 < \eta \leq 3.0$	$ \eta > 3.0$
Number of constituents	> 1	—	—	—
Neutral hadron fraction	< 0.9	< 0.9	—	> 0.2
Neutral EM fraction	< 0.9	< 0.99	> 0.02 and < 0.99	< 0.9
Number of neutral particles	—	—	> 2	> 10
Charged hadron fraction	> 0	—	—	—
Charged multiplicity	> 0	> 0	—	—

Jet energy scale & resolution

The reconstructed jet energy is influenced by various effects. Within CMS a procedure was developed to calibrate the energy of reconstructed jets with the aim to provide the correct *jet energy scale* (JES) for jets in simulation and data [161]. The correction procedure is factorized into several stages, where each of these corrections accounts for a particular effect. In the first step a PU correction is applied in data and simulation, which accounts for the energy offset due to neutral hadrons and photons from PU. The next step is the simulated response correction, which is derived from simulation and applied as a function of jet η and p_T . Finally, residual corrections are applied to data only. These are measured in situ from the energy balance in di-jet, multi-jet, γ +jets, and Z +jets events to correct for the remaining differences between data and simulation. The JES correction is centrally provided by CMS and the latest recommendations [162] are applied in this thesis.

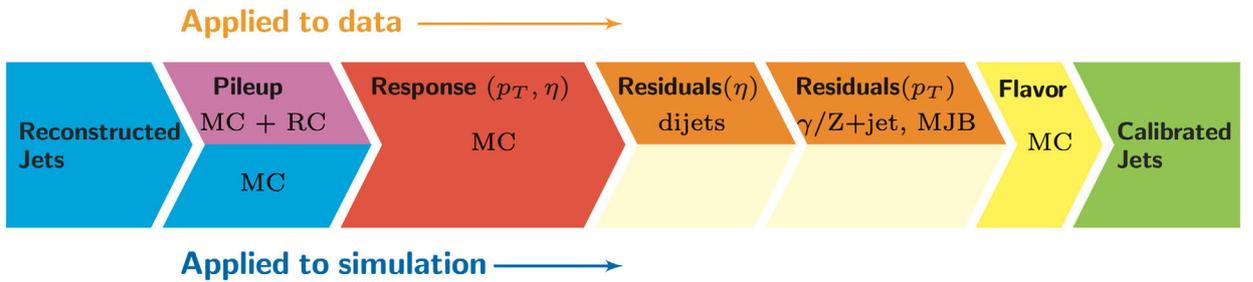


Figure 4.4: Consecutive stages of the procedure used to correct the JES in simulation and data. Taken from Ref. [163].

The *jet energy resolution* (JER) in simulation is corrected to reflect the resolution measured in data. For this purpose, jets in simulated events are smeared by a hybrid method. When a reconstructed jet can be matched to a generated jet, the four-momentum of the reconstructed jet is scaled with

$$c_{\text{JER}} = 1 + (s_{\text{JER}} - 1) \frac{p_T - p_T^{\text{gen}}}{p_T} \quad (4.2)$$

in terms of the reconstructed jet p_T , the generated jet p_T^{gen} , and the data-to-simulation core resolution scale factor s_{JER} . In cases where no match of a reconstructed jet to a generated jet can be established, a stochastic smearing is applied. The core resolution scale factors are centrally provided by CMS [164]. These factors are derived in samples enriched with high- p_T jets, while a large fraction of jets used in this thesis has smaller p_T . In the regime with jet $p_T < 50$ GeV the JER smearing does not improve the agreement between data and simulated events. For this reason, the JER smearing is not applied by default. Instead it is used to estimate the systematic uncertainty in the signal rate due to variations in the JER.

4.3.6 B-tagging

The identification of jets initiated by b quarks is referred to as *b-tagging*. These jets are of particular interest for the analysis presented in this thesis since b-tagged jets will be used to identify the top quarks produced in association with the Higgs boson. Top quarks dominantly decay to a b quark and a W boson. B-tagging algorithms exploit different properties of jets and their constituents to distinguishing b jets from u,d,s,g (light-flavor+gluon) jets. Due to their large mass, hadrons containing b quarks have a large lifetime, which leads to a displacement for b hadrons of a few millimeter with respect to the primary interaction point. The displaced tracks can be used to reconstruct a *secondary vertex* (SV) as presented in Fig. 4.5. Since the tracker is necessary to reconstruct the SV, the selection of b-tagged jets is restricted to the tracker acceptance of $|\eta| < 2.5$.

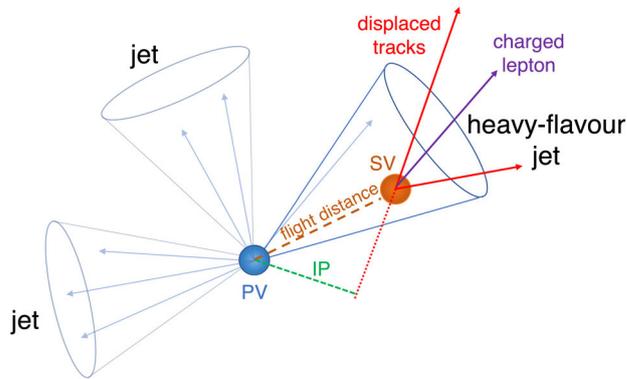


Figure 4.5: Illustration of a b quark jet with a secondary vertex resulting in charged-particle tracks that are displaced with respect to the primary interaction vertex. Taken from Ref. [165].

The *Deep Combined Secondary Vertex* (DeepCSV) algorithm [165] based on a *deep neural network* (DNN) is employed for b-tagging. Compared to previous versions of this algorithm the selection efficiencies and misidentification rates for b-tagged jets could be improved by using the advances in the field of deep machine learning [166]. A fully connected DNN was trained on preselected jets with $p_T > 20$ GeV using a set of already established input observables, which comprise the quality of the SV, different track properties, as well as the p_T and η of the jet. The working points used in this thesis are listed for each data-taking period in Table 4.4.

Table 4.4: Selected b-tagging working points for each data-taking period. The chosen medium (loose) working point of the DeepCSV discriminator output corresponds to a b-tagging selection efficiency of 68% (84%) with a misidentification rate of 1.1% (11%) [167].

Year	DeepCSV working point	
	medium	loose
2016	0.6321	0.2217
2017	0.4941	0.1522
2018	0.4184	0.1241

4.3.7 Quark-gluon discriminant

Hadronic jets initiated by gluons show different properties with respect to jets from light flavor quarks. Gluon jets have higher charged particle multiplicities, and are less collimated compared to quark jets. These differences are used to build a single discriminant able to distinguish gluon jets from quark jets. The discriminant developed by the CMS Collaboration is referred to as *quark-gluon likelihood* (QGL) [168, 169]. It is based on three observables reconstructed from the jet constituents. Namely, the total number of particles reconstructed within the jet, the ellipse minor axis of the jet constituents in the $(\eta-\phi)$ plane, and the fragmentation function defined as

$$p_T D = \frac{\sqrt{\sum_i p_{T,i}^2}}{\sum_i p_{T,i}}, \quad (4.3)$$

in terms of the transverse momenta p_T of all jet constituents i . Gluon jets tend to have small values of $p_T D$, while quarks have large values. The QGL is constructed using the product of three probability density functions corresponding to the three observables. A large value of the QGL implies that a jet originates from the hadronization of a light flavor quark, while a small value of the QGL indicates a gluon jet. This is later used to identify the decay of a top quark to three resolved jets as described in Section 5.7.

4.3.8 Missing transverse momentum

The reconstruction of particles is based on signals in the detector. However, neutrinos interact only weakly which means they traverse the detector with little or no signal. Therefore, an indirect method is used to identify neutrinos and recover their transverse momentum referred to as missing transverse momentum \vec{p}_T^{miss} [170]. The momentum sum in the transverse plane is conserved. Assuming neutrinos to be the only source of \vec{p}_T^{miss} , any imbalance in the transverse momentum sum indicates the presence of neutrinos in the event. Therefore, \vec{p}_T^{miss} is defined by the negative vectorial sum of all PF particles that pass a certain p_T threshold

$$\vec{p}_T^{\text{miss}} = - \sum \vec{p}_T. \quad (4.4)$$

The magnitude of this vector corresponds to the *missing transverse energy* (MET) E_T^{miss} . The indirect measurement of \vec{p}_T^{miss} relies on a precise p_T measurement of all detected particles in the event. In particular, wrongly measured jet p_T can have a significant impact on the \vec{p}_T^{miss} . Therefore, the type-1 corrected definition of \vec{p}_T^{miss} is employed in this thesis, which accounts for a mismeasurement of the jet p_T by replacing non-calibrated jets with the calibrated jets. In this way, the JEC is propagated to the definition of \vec{p}_T^{miss} . A set of additional event filters is used to reject events which are affected by detector noise in the HCAL, by low quality crystals in the ECAL endcap, and by beam halo effects [171].

Higgs boson decays to muons produced in association with top quarks

The previous CMS search for Higgs boson decays to muons HIG-17-019 [41], as introduced in Section 1.4, primarily focused on the two largest Higgs boson production modes ggH and VBF . To suppress the background contribution from top quark-antiquark pair production $t\bar{t}$, the number of b -tagged jets was included as input to the final multivariate discriminant. Since $t\bar{t}$ as well as the $t\bar{t}H$ production both feature b jets in the final state, the $t\bar{t}H$ signal was rejected in the same way as the background from $t\bar{t}$. Both, $t\bar{t}$ and $t\bar{t}H$ events accumulated in the least sensitive analysis category with very small signal-to-background ratio. Thus, the final signal extraction fit was insensitive to the signal contribution from $t\bar{t}H$.

For the recent CMS analysis HIG-19-006 [172] based on the full Run 2 dataset, an analysis strategy dedicated to the $t\bar{t}H$ production mode is developed. This analysis is the central subject of this thesis and presented in this chapter.

The analysis outline is provided in Section 5.1 followed by a description of the signal and background event topology in Section 5.2. Thereafter, Section 5.3 motivates the trigger selection, presents an overview of the used datasets, and summarizes the data-taking conditions. Section 5.4 provides a review of corrections, which are applied to account for differences between data and simulation. In particular, the muon momentum scale and resolution and their impact on the reconstructed Higgs boson mass are discussed. Section 5.5 introduces two methods that aim to improve the dimuon mass resolution of the reconstructed Higgs boson candidate. A multivariate lepton identification algorithm is explained in Section 5.6 which helps to distinguish between prompt leptons and non-prompt leptons from jets. The reconstruction of top quarks decaying into three resolved jets is performed by a dedicated algorithm presented in Section 5.7. The event selection and the optimization strategy based on a multivariate discriminant is discussed in Section 5.8 and Section 5.9, respectively. After the application of the event selection criteria, events are sub-divided according to the output of the multivariate discriminant into several subcategories. The category boundaries, determined by an iterative algorithm, are reported in Section 5.10. The analytical functions that model the dimuon mass distribution of the signal and background component in the final maximum likelihood fit are described along with the considered systematic uncertainties in Section 5.11. Lastly, Section 5.12 provides the results of this analysis.

5.1 General $H \rightarrow \mu^+ \mu^-$ analysis strategy

The analysis reported in this thesis is designed to contribute to a combined search for Higgs boson decays to muons targeting all main Higgs boson production modes at the LHC: gluon-fusion (ggH), vector boson fusion (VBF), production in association with a vector boson (VH), and production in association with a top quark-antiquark pair ($t\bar{t}H$). To avoid any overlap between the analyzed events, four exclusive event categories are defined. Each of the event categories is enriched in one of the Higgs boson production modes. A comprehensive overview of the analysis strategy is provided in Fig. 5.1. Several aspects of the analyses follow similar approaches, although each event category is particularly optimized to achieve the best possible results.

The analysis proceeds as follows. Events are accepted by the single muon trigger. The baseline dimuon event selection includes all events containing a pair of oppositely charged muons, where at least one of these muons matches the single muon trigger object. The corresponding muon pair defines the Higgs boson candidate. The mass resolution of the Higgs boson candidate is improved using two independent methods referred to as *final-state radiation* (FSR) photon recovery and GEOFIT. Thereafter, events are categorized according to their additional final state particles corresponding to the characteristics of a Higgs boson production mode.

Events consistent with $t\bar{t}H$ production are identified using b jets. The $t\bar{t}H$ category comprises all events that contain at least two jets passing the loose working point of the b-tagging algorithm, or one jet passing the medium working point. The $t\bar{t}H$ category is sub-divided into a $t\bar{t}H$ leptonic and a $t\bar{t}H$ hadronic category based on the presence of additional leptons. Events without b-tagged jets that contain three (four) leptons compatible with the WH (ZH) production mode are assigned to the VH category. All remaining events represent the Higgs boson production modes ggH and VBF. These cannot directly be distinguished by their final state particles because ggH as well as VBF can feature two jets in addition to the muon pair. However, the VBF production mode is identified by two jets predominantly radiated in forward direction leading to a large dijet mass $m_{jj} > 400$ GeV, and large difference of the pseudorapidity between the two jets $\Delta\eta_{jj} > 2.5$. Events containing at least two jets, which meet these criteria, are assigned to the VBF category. Events, which are not selected by the $t\bar{t}H$, VH, or VBF event category are assigned to the ggH category.

The ggH, VH, and $t\bar{t}H$ categories exploit the dimuon mass distribution of the Higgs boson candidate to extract the signal. Therefore, multivariate techniques, which are designed to distinguish signal from background events, only include observables largely uncorrelated with the Higgs boson candidate mass. Based on the resulting multivariate discriminants, events are classified into several subcategories. These define sensitive phase-space regions with enhanced signal-to-background ratio. The background shape is modeled using empirical functions in an extended window around the Higgs boson mass between 110 and 150 GeV, referred to as the *sideband* (SB) region. The background contribution in the *signal region* (SR) close to the Higgs boson mass in the range between 120 and 130 GeV is thereby estimated from the mass sidebands which almost entirely consist of background events.

The characteristic event topology of VBF events allows strong suppression of the dominant background contributions, while signal events are selected with high efficiency. The large expected sensitivity is limited by the statistical precision of the background estimation due to the small number of expected data events in the dimuon mass sideband. For this reason, the VBF category follows a different signal extraction approach. In contrast to the described data driven background estimation, the VBF analysis relies on simulated events to model the background. Simulated events can be produced corresponding to a multiple of the actual luminosity, which significantly reduces the uncertainty of the background estimation, and improves the sensitivity. The signal extraction is performed by a fit to the output distribution of a multivariate discriminant, which includes the dimuon mass information.

The combined result is obtained by a simultaneous maximum likelihood fit in all 21 subcategories. Further details on the combined analyses, and its results are given in Chapter 6. The development of the $t\bar{t}H$ analysis is described in the following sections.

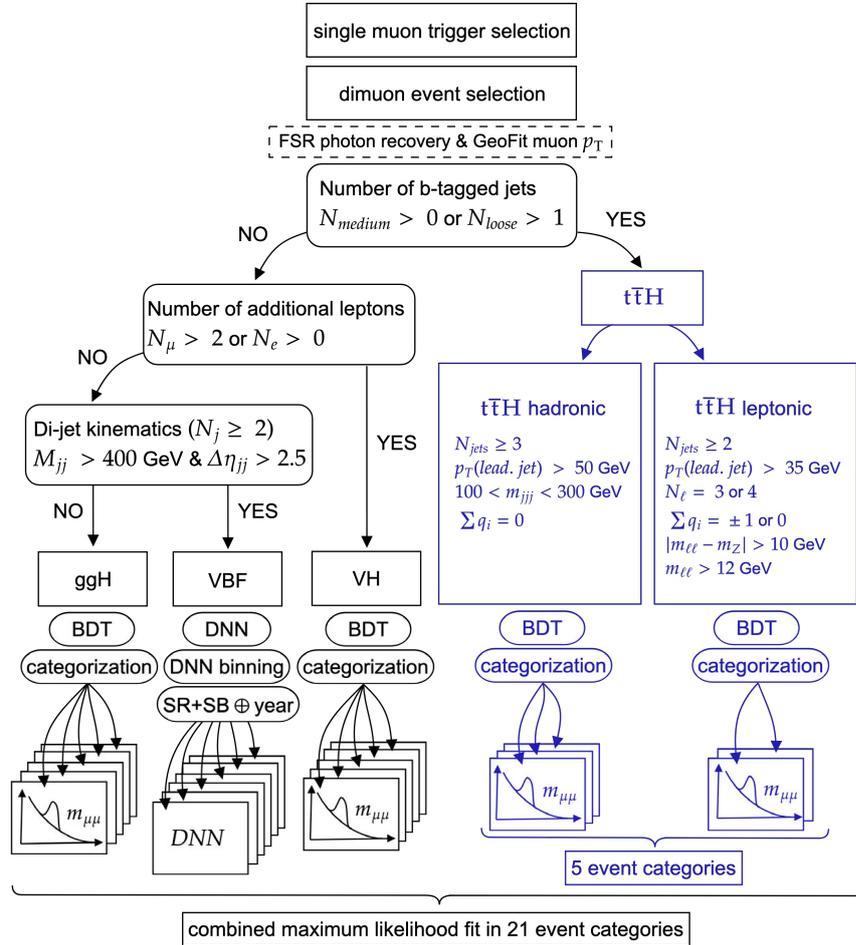


Figure 5.1: Illustration of the analysis strategy for the CMS $H \rightarrow \mu^+\mu^-$ analysis using the full dataset of LHC Run 2. The preselected dimuon events are divided into four exclusive event categories targeting the main Higgs boson production modes. Each analysis is optimized using multivariate techniques. The final results are obtained by a simultaneous maximum likelihood fit in all event categories. The $t\bar{t}H$ analysis presented in this thesis is highlighted.

5.2 Event topology in the $t\bar{t}H$ category

Massive particles like the H,Z,W bosons, or the top-quark, which are produced in proton-proton collisions at the LHC, decay on very short timescales and are identified by their decay products. Signal events are therefore distinguished from background events using differences between the kinematic properties of the final state particles. An accurate understanding of the event topology is necessary to achieve a sufficient discrimination between signal and background events. In the following, the final state signatures of the $t\bar{t}H$ production are described along with all considered background processes.

Signal

The $t\bar{t}H$ production has the smallest cross section among the Higgs boson production modes which are targeted in the analyses. This process features two oppositely charged muons from the decay of the Higgs boson accompanied by the decay products from the top quark-antiquark pair. The top quark decays predominantly into a b quark and a W boson. The corresponding W boson can either decay hadronically into two quarks, or leptonically into a lepton and neutrino. An exemplary illustration of the $t\bar{t}H$ signal process where one top quark decays leptonically and the other top quark decays hadronically is shown in Fig. 5.2. The cases where both top quarks decay hadronically (leptonically) are considered as well.

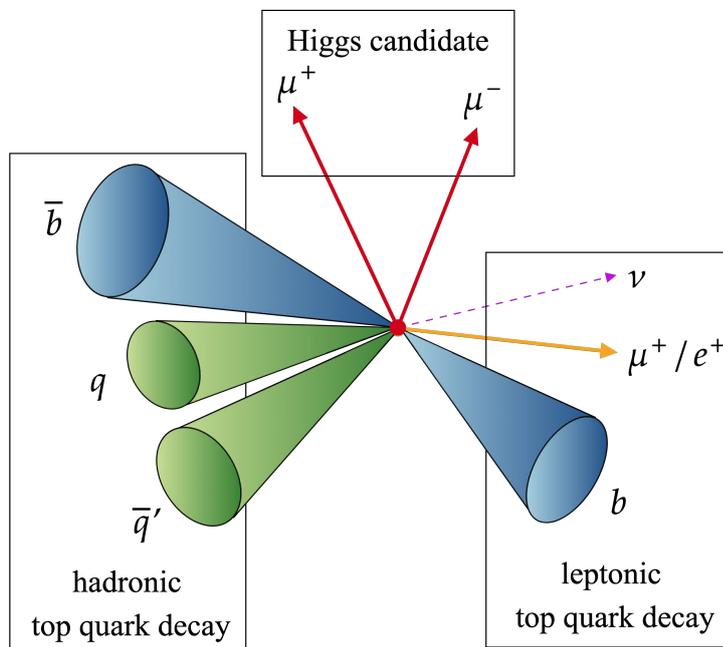


Figure 5.2: Depiction of the experimental signatures expected for $t\bar{t}H$ production with Higgs boson decay to muons. Both top quarks, produced in association with the Higgs boson, decay to a b quark and a W boson. The corresponding W boson can either decay hadronically into two quarks (left), or leptonically into a lepton and neutrino (right).

Backgrounds

A certain fraction of the different processes that take place in the proton-proton collisions produces a final state similar to the one expected from the signal process. The considered background processes contain two oppositely charged muons and at least one b tagged jet. Additional jets or leptons can mimic the decay products of the top quarks. Example Feynman diagrams of possible background processes are presented in Fig. 5.3. Further details, including their cross section times branching fraction are provided in Section 3.2. The event yields of all background contributions after the final selection are reported in Section 5.8. The different background processes and their potential importance are discussed below.

- **Top quark** The processes containing one, two, or multiple top quarks are summarized into the "Top quark" background. The largest contribution is represented by the top quark-antiquark pair production with both top quarks decaying leptonically into muons. Due to the very large cross section of this process, which can also include additional jets from radiation or leptons originating from the decay of heavy b hadrons, it constitutes the dominant background for the $t\bar{t}H$ event category.
- **$t\bar{t}Z + tZ$** The production of top quark-antiquark pairs or single top quarks in association with a Z boson are considered to be an irreducible background, since the final state particles provide very similar kinematic properties compared to signal events. The cross section of $t\bar{t}Z$ production is twice as large as the cross section of $t\bar{t}H$. However, the ratio of the branching fractions for decays to muons of the Higgs boson is 10^{-3} smaller than the one of the Z boson. Considering also the off-shell Z boson production, which drastically reduces the $t\bar{t}Z$ production rate, this background is expected to have event yields around 5 times larger than the signal in the relevant dimuon mass range around the Higgs boson mass.
- **$t\bar{t}W(W)$** The top quark-antiquark pair production in association with one or two W bosons comprises a slightly smaller background contribution compared to $t\bar{t}Z$. Although its cross section is similar, the decay products from the W boson and the top quarks only form by chance a pair of muons compatible with the decay of a Higgs boson, which reduces the effective background rate.
- **DY+Z-EWK** The background with the largest cross section, producing a pair of oppositely charged muons in the final state, is the Drell-Yan process. It refers to the production of a Z boson or virtual photon in quark-antiquark annihilation. Additionally, the electroweak production of the Z boson with much smaller cross section is considered. The DY+Z-EWK background contribution is strongly suppressed due to the absence of b quarks at LO, which are only produced by ISR.
- **WZ/ZZ lep.** The SM production of multiple weak vector bosons occurs in all combinations of W and Z bosons. The "WZ/ZZ lep." background summarizes the cases, where a Z boson is produced along with an additional W and Z boson, and both vector bosons decay leptonically. Similar to the Drell-Yan process, b quarks are only produced by ISR, which significantly reduces the rate of this background.

- VV(V)** All remaining contributions of possible diboson and triboson combinations are included in the "VV(V)" background. Although potentially many additional leptons and jets are produced in these processes, they represent one of the smallest backgrounds, due to their small cross sections.

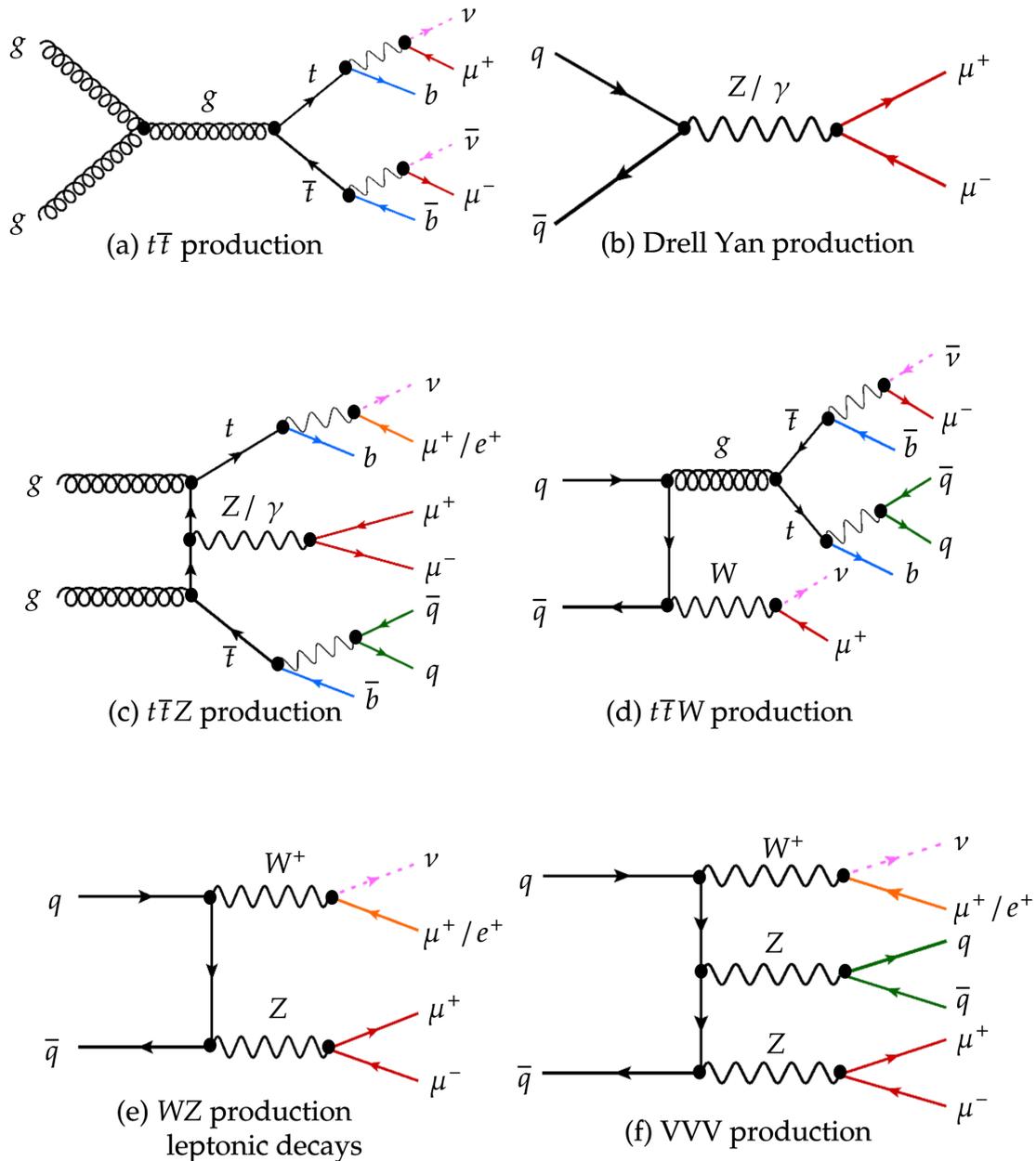


Figure 5.3: Example LO Feynman diagrams of the six main background contributions, $t\bar{t}$ production (top left), Drell-Yan (top right), $t\bar{t}Z$ production (center left), $t\bar{t}W$ production (center right), WZ production (bottom left), and VVV production (bottom right).

5.3 Trigger & dataset

The analyzed dataset is based on proton-proton collisions at a center-of-mass energy of 13 TeV recorded by the CMS experiment between 2016 and 2018, and corresponds to an integrated luminosity of 137 fb^{-1} . The datasets are split by time periods corresponding to different beam intensities and operating conditions. Table 5.1 summarizes the used datasets for all three years along with their integrated luminosity and trigger properties. Only data, where the CMS detector was fully operational ("CMS certified high quality data") is included.

Muons emerging from the decay of the Higgs boson are expected to have large transverse momentum in the order of half of the Higgs boson mass $\frac{1}{2}m_H \approx 62 \text{ GeV}$. Generally, the online reconstructed trigger muons are required to pass a certain p_T threshold to limit the trigger rate. Signal events are selected with a high selection efficiency using single muon triggers, which require triggered muon candidates to pass loose isolation criteria, and a p_T threshold of 24 GeV in 2016/2018, and 27 GeV in 2017. At least one offline reconstructed muon has to be matched within a cone of $\Delta R = 0.1$ to the muon candidate that triggered the event record.

To measure trigger and selection efficiencies, a *tag-and-probe* method [173] is used. For this purpose, the well-known Z boson mass resonance is exploited. Lepton pairs are selected based on two different selection types. One of the leptons, referred to as the *tag*, is required to fulfill very tight selection criteria with a misidentification rate much smaller than 1%. A set of lepton candidates passing very loose selection requirements, referred to as *probes*, is selected by pairing these with the tags such that the invariant mass is compatible with the Z boson mass. The efficiency of a desired selection criterion is determined by the number of probes that pass the selection normalized to the total number. Typically these efficiencies are measured as a function of the lepton p_T and η .

Figure 5.4 shows the single muon trigger efficiency depending on the muon p_T and η for each data-taking period. To avoid the turn-on region of the trigger efficiency in p_T , where the efficiency rapidly increases until it reaches a plateau, slightly larger p_T thresholds are required for the offline reconstructed muons. The offline p_T thresholds are 26 GeV in 2016/2018 and 29 GeV in 2017. Due to the high p_T of muons in signal events, this has only a small impact on this

Table 5.1: Summary of the used datasets collected between 2016 and 2018. The selected trigger and the corresponding trigger thresholds, the dataset names, and the integrated luminosity for each year of data taking are reported.

Year	Datasets	Trigger	Trigger threshold		Integrated luminosity
			online	offline	
2016	Run B,C,D,E,F,G,H	Single muon	$p_T > 24 \text{ GeV}$	$p_T > 26 \text{ GeV}$	35.9 fb^{-1}
2017	Run B,C,D,E,F	Single muon	$p_T > 27 \text{ GeV}$	$p_T > 29 \text{ GeV}$	41.5 fb^{-1}
2018	Run A,B,C,D	Single muon	$p_T > 24 \text{ GeV}$	$p_T > 26 \text{ GeV}$	59.7 fb^{-1}

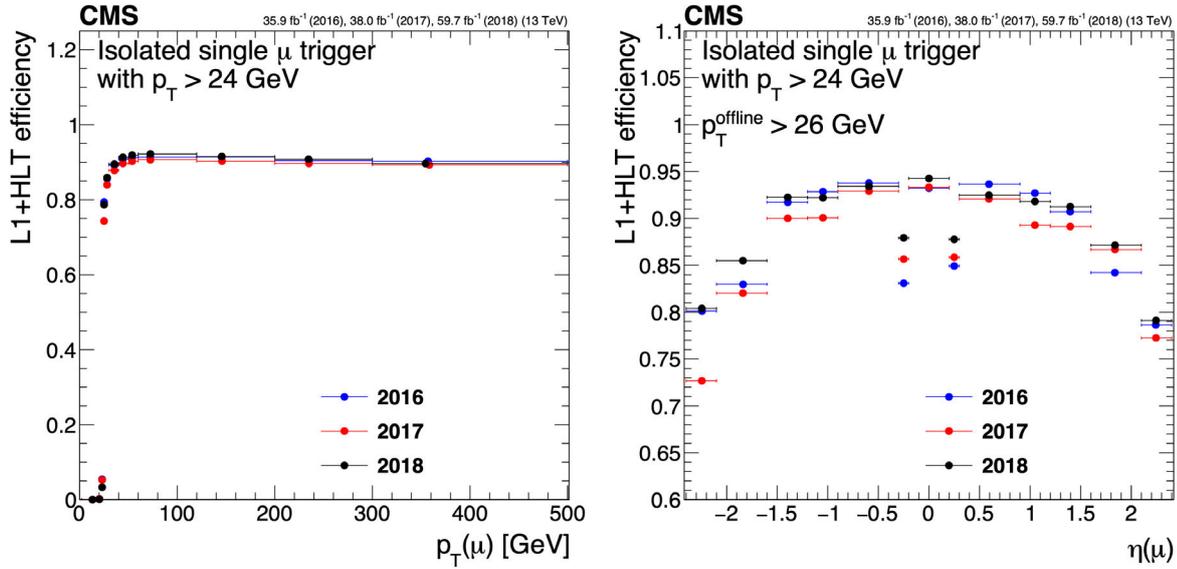


Figure 5.4: Single muon trigger efficiency in data as a function of muon p_T (left) and η (right) for each year of data taking. Taken from Ref. [174].

analysis. The trigger efficiency is mostly constant in the barrel region of the detector, except for a narrow detector gap region $|\eta| \approx 0.2$. Towards large $|\eta|$ the efficiency decreases due to limitations in the L1 trigger reconstruction algorithms. The overall trigger efficiency for muons is 90% for 2016 and 2018 while in 2017 the efficiency is slightly lower. This is partially caused by the introduction of a new trigger object reconstruction algorithm. The performance stabilizes throughout the 2018 data-taking period, and the trigger efficiency in 2018 reaches a value similar to the one in 2016 [174]. In simulated events, the trigger is emulated to provide a similar selection efficiency.

Experimental circumstances that influenced the data taking are discussed in the following. In 2016 and 2017, a slight shift in the timing of the ECAL readout was not properly propagated to the L1 trigger. This affected the trigger efficiency and is corrected as explained in Section 5.4.2. A powering issue led to inactive pixel detector modules in 2017. The general efficiency corrections, described in Section 5.4, account for any differences in efficiencies between data and simulation. By the end of 2017, the ECAL experienced a loss of transparency due to continuous radiation damage. To cope with the less sensitive ECAL crystals, signal amplification factors were increased, which led to considerable detector noise. Specific selection criteria are applied to mitigate the performance loss for jets detected in the forward region of the detector as described in Section 4.3.5. Moreover, jets with low p_T that are reconstructed in the corresponding detector region are excluded from the calculation of the E_T^{miss} . An adapted ECAL calibration scheme eliminated this effect in 2018. The power supplies of some HCAL modules broke in the middle of the 2018 data-taking period, which affected the jet energy measurement. The overall impact on this analysis is found to be negligible.

5.4 Efficiency corrections & event reweighting

This analysis aims to evaluate the $H \rightarrow \mu^+\mu^-$ signal based on the measured efficiency in simulated events. Any mismodeling of physical observables or bias in the detector response, which is not accounted for in the simulation, leads to discrepancies between the selection efficiency in data and simulation. To correct for these differences and to ensure that physical observables are modeled correctly, simulated events are reweighted using multiplicative *scale factors* (SFs). Standard corrections are provided centrally by the corresponding CMS experts, who are organized in the *Physics Object Groups* (POGs). The applied corrections are discussed in the following.

5.4.1 Lepton efficiencies

Differences in the selection efficiency of leptons between data and simulation are corrected using SFs which correspond to the ratio of the efficiency measured in data and simulation

$$\text{SF} = \frac{\varepsilon_{\text{data}}}{\varepsilon_{\text{simulation}}}. \quad (5.1)$$

The muon trigger efficiency ε_i measured in a tag-and-probe method describes the probability for a single muon i to trigger the event. Since at least two muons are present in all selected events and both can potentially trigger the event record, the probability for events where at least one of these muons triggered is to be determined. This probability is complementary to the case where none of the muons triggered the event. Therefore, the per-event trigger efficiency is computed by

$$\varepsilon_{\text{trigger}} = 1 - \prod_{i \in \mu} (1 - \varepsilon_i). \quad (5.2)$$

The per-event trigger SF corresponds to the ratio of the per-event trigger efficiency in data and simulation. The per-event lepton *identification and isolation* (id+iso) SFs are computed using the SF product of all selected leptons in the event. The total lepton SF, which is defined by the product of the trigger SF and the id+iso SF, is applied as an event weight depending on the muon p_T and η to correct the lepton selection efficiencies in simulated events

$$\text{SF}_{\text{lepton}} = \text{SF}_{\text{trigger}} \cdot \text{SF}_{\text{id+iso}}. \quad (5.3)$$

The SFs depend on the chosen trigger, and on the specific isolation and identification criteria used to select muons and electrons. The identification and isolation efficiencies, as well as the muon trigger efficiencies are provided by the $t\bar{t}Z/tZq$ analysis group [175, 176], and the Muon POG [177–180], respectively.

Figure 5.5 presents an overview of the single muon trigger SFs for all three years. The SFs are shown as a function of p_T and η . In general, the trigger selection efficiencies are well modeled in simulated events even before any correction, since most SFs are in the range between 0.95 and 1. Larger differences between the efficiencies in data and simulation are observed for leptons at large $|\eta|$, where the SFs vary between 0.90 and 1.02.

Figure 5.6 presents the electron and muon id+iso SFs for 2016 and 2017 in different $|\eta|$ regions as a function of p_T . Except for the high η region in 2016, the muon id+iso SFs yield values in the range between 0.98 and 1.01 in both years and for all p_T and η regions. In contrast, electron id+iso SFs tend to have smaller values, which vary between 0.85 and 1.02 depending on the electron p_T . The larger corrections result from the more challenging simulation of the electron reconstruction, which is affected by bremsstrahlung for example.

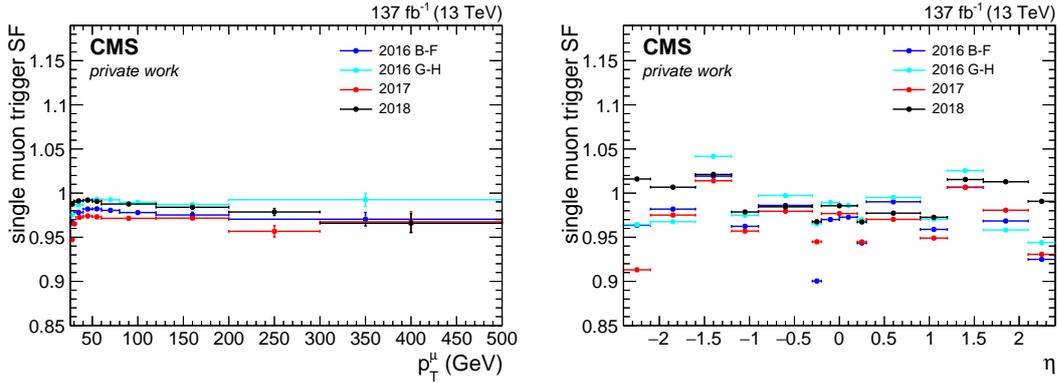


Figure 5.5: Single muon trigger SFs as a function of muon p_T (left) and η (right) for each year of data taking. The SFs are provided by the Muon POG [177].

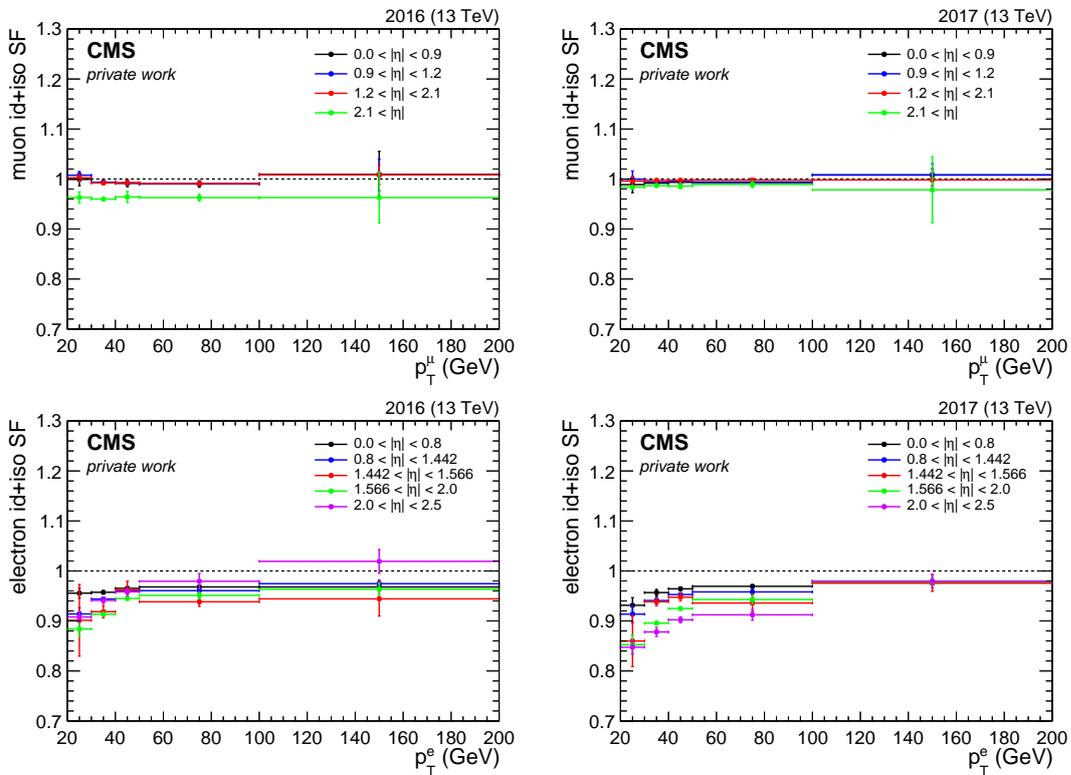


Figure 5.6: Lepton identification and isolation SFs as a function of p_T . The muon (top) and electron (bottom) SFs are shown independently for 2016 (left) and 2017 (right). The 2017 SFs are used for 2018 as well. The SFs are provided by the $t\bar{t}Z/tZq$ analysis group [175, 176].

5.4.2 L1 ECAL prefiring correction

During the data taking in 2016 and 2017 a fraction of the L1 trigger objects, which are generated by energy deposits in the high $|\eta|$ region of the ECAL, are associated to the previous bunch crossing by mistake, which is referred to as prefiring. This issue is caused by a drifted time alignment in the ECAL readout. The L1 trigger system of CMS is designed to prevent the selection of events in consecutive bunch crossings, which induces a systematic trigger inefficiency due to the the trigger prefiring.

A procedure to correct simulated events according to their prefiring probability is carried out by the JetMET POG [181]. The prefiring probabilities of jets are assumed to be uncorrelated. They are measured using a set of unbiased single jet events. Figure 5.7 provides the resulting probability maps for a single jet depending on the pseudorapidity and the electromagnetic component of the jet transverse momentum. Events in 2017 are more affected by prefiring than those in 2016.

The event efficiency factor $\varepsilon_{\text{non-prefire}}$ is obtained as the product of the non-prefiring probability of all jets in the event

$$\varepsilon_{\text{non-prefire}} = \prod_j 1 - \mathcal{P}_{\text{prefire}}(\eta, p_{\text{T}}^{\text{EM}}).$$

Events containing forward jets in the pseudorapidity region $2.6 < |\eta| < 3$ are affected the most. As jets in the non-VBF event categories are primarily within the tracker acceptance ($|\eta| < 2.5$), the impact on $t\bar{t}H$ signal events is rather small. The inefficiency leads to a reduction of the $t\bar{t}H$ signal yield by 2% in 2016 and 3% in 2017. The considered uncertainty corresponds to 20% of the prefiring probability.

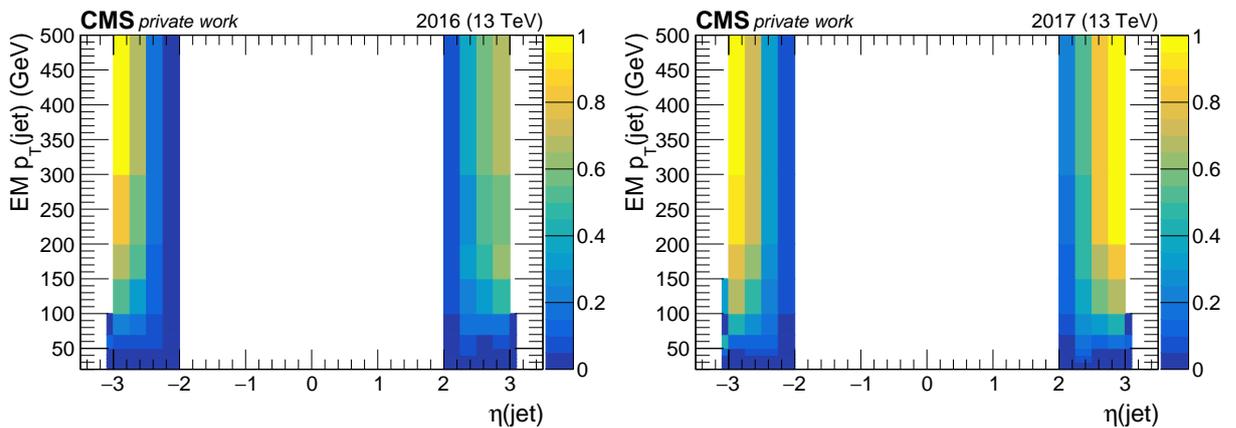


Figure 5.7: Prefiring probability as function of the electromagnetic transverse momentum and the pseudo rapidity of jets for 2016 (left) and 2017 (right). Prefiring probabilities provided by the JetMET POG [181].

5.4.3 Pileup reweighting

During the Run 2 data taking, the instantaneous luminosity was continuously increased. This influences the shape of the pileup distribution and the average number of pileup interactions per event for each year. The time consuming production of simulated events is performed using a preliminary pileup profile, which needs to be reweighted to match the pileup distribution in data. For this purpose, the ratio of the profiles in data and simulation is used to compute an event weight. The pileup profile in data is determined based on a total inelastic proton-proton cross section of $69.2 \text{ mb} \pm 4.6\%$, which is provided by independent measurements involving the hit occupancy of the pixel detector. In case of simulated events, the pileup profile is obtained using the true number of pileup interactions per event.

The pileup weights are independently derived for each year of data taking and their impact on the number of reconstructed vertices is presented in Fig. 5.8. The corresponding events are selected in the $m_{\mu\mu}$ sideband region using the baseline dimuon selection. After pileup reweighting, the agreement between data and simulation is slightly improved. However, a remaining disagreement between data and simulation in the number of reconstructed vertices is clearly visible. The mismodeling mainly influences physical observables sensitive to PU. In particular, events without genuine E_T^{miss} from the presence of neutrinos are not well modeled in simulated events. Since possibly affected observables are not considered in the optimization procedure presented in Section 5.8 and Section 5.9, and the background is estimated from the sideband region in data only, the impact on the $t\bar{t}H$ analysis is expected to be small.

5.4.4 B jet discriminator & quark-gluon likelihood reshaping

In this thesis different jet properties are used to reconstruct the hadronic decay of a top quark. For this, three jets are tested to be compatible with the expected top quark decay products. These comprise one b jet and two light-flavor jets. To distinguish between the different jet types, the b jet discriminator and the quark-gluon likelihood as introduced in Section 4.3.6 and Section 4.3.7 are employed. Each of them provides a single discriminator output distribution. The hadronic top quark reconstruction relies on the full distribution of the discriminator values in both cases. Therefore, the full range of the discriminator needs to be corrected to account for differences in the discriminator shape between data and simulation.

The b jet discriminator shape is calibrated using event weights [182] depending on the b jet discriminator value, the jet p_T and η , as well as the jet flavor. These weights are derived in an iterative way based on a tag-and-probe method. The tag jet is required to pass the medium working point of the b-tagging algorithm, while the discriminator distribution of the probe jet is calibrated. To extract the weight for b jets, the procedure selects dileptonic $t\bar{t}$ events which yield an exceptionally pure sample of genuine b jets. The contribution from light-flavor jets is subtracted using simulation. This allows for a direct calibration of the response of true b jets.

To correct the b jet discriminator shape of light-flavor jets, Drell-Yen+jets events are used. Since the calibration for b jets and light-flavor jets depend on each other, they are extracted

iteratively. Fig. 5.9 presents the DeepCSV discriminator for the leading jet in p_T , separately for each year of data taking. The ratio panel indicates general shape differences between the distributions in data and simulation before the calibration. In particular, small values which represent light-flavor jets are not well modeled. The b jet discriminator shape calibration can significantly improve the agreement of data and simulation for light-flavor jets in all three years.

The distribution of the QGL discriminant for jets is found to be generally in good agreement between data and simulation. However, the description of gluon jets modeled from the PS using PYTHIA does not entirely reproduce the shape observed in data. To correct this disagreement, a reweighting procedure is developed by the JetMET POG [183] based on a three-dimensional polynomial reweighting function derived from a data-to-simulation comparison. The normalization itself is corrected to guarantee that the event yield is unchanged. A comparison of the QGL discriminant before and after the reweighting is provided in Fig. 5.10. The agreement between data and simulation is improved, in particular, in the region of small values of the QGL discriminant, which describes gluon-like jets.

5.4.5 Z boson transverse momentum correction

A specific correction to the Drell-Yan background simulation is performed with the aim to resolve an observed disagreement between data and simulation for small transverse momenta of the muon pair $p_T^{\mu\mu}$. This is potentially caused by missing resummation effects, which are not modeled in the simulation. The $p_T^{\mu\mu}$ correction is derived from a control region orthogonal to the signal region. For this purpose, events are selected in a window around the Z boson mass, $70 < m_{\mu\mu} < 110$ GeV. The ratio of the normalized $p_T^{\mu\mu}$ distributions in data and simulation is used to define a reweighting factor depending on the $p_T^{\mu\mu}$ up to 60 GeV. The correction is designed to preserve the number of events, i.e., the normalization of events is unchanged. Figure 5.11 presents the $p_T^{\mu\mu}$ spectrum before and after the correction is applied. The $p_T^{\mu\mu}$ distribution in simulated events peaks at smaller values compared to data. This shift is corrected by the reweighting procedure in all three years.

Chapter 5. Higgs boson decays to muons produced in association with top quarks

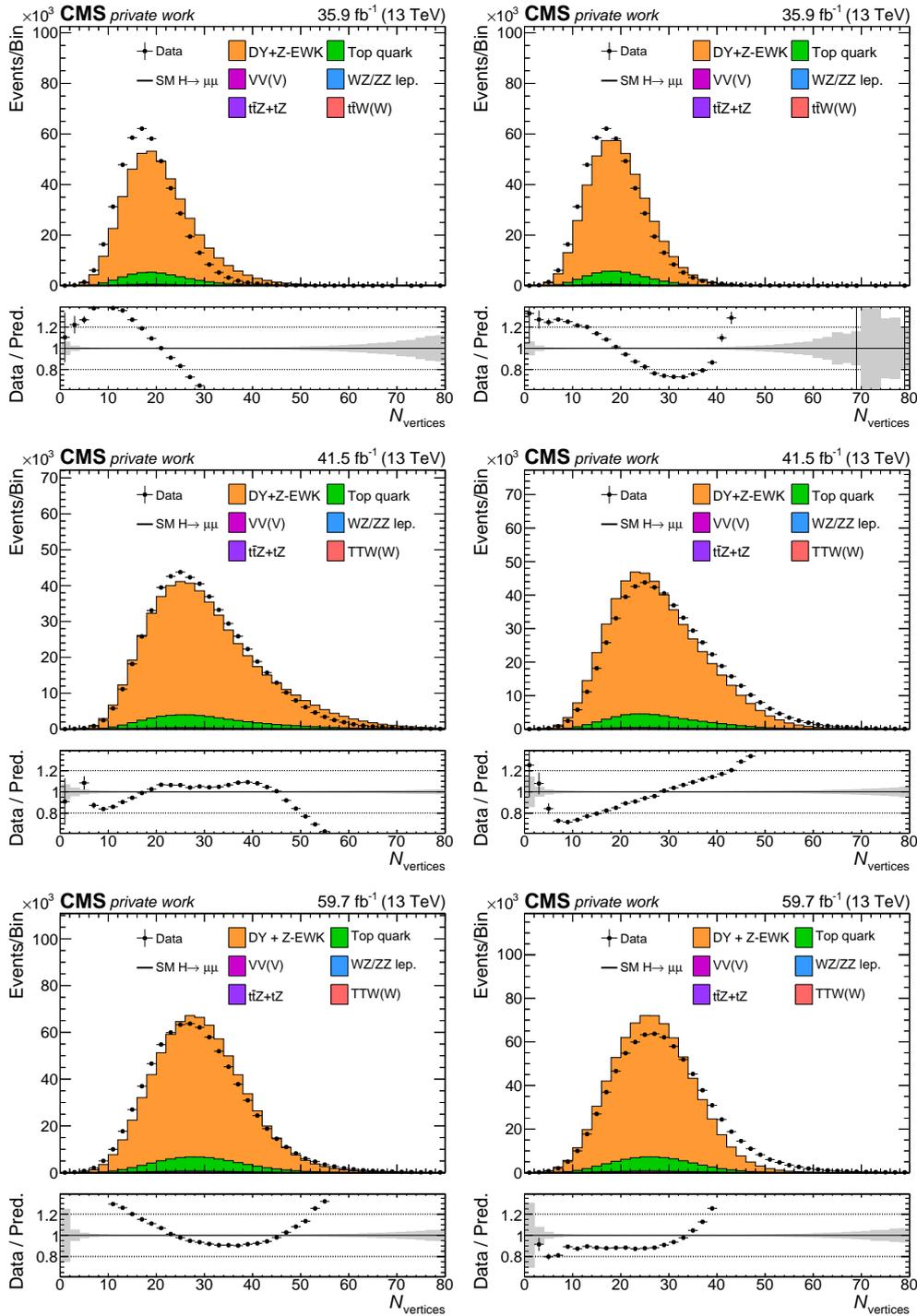


Figure 5.8: Comparison of the number of primary vertices before (left) and after (right) PU reweighting. The corresponding events are selected in the $m_{\mu\mu}$ sideband region using the baseline dimuon selection. The comparison is provided for each year of data-taking: 2016 (top), 2017 (center), 2018 (bottom). All additional event weights considered in this thesis are applied. The prediction is scaled to the observed number of events.

5.4 Efficiency corrections & event reweighting

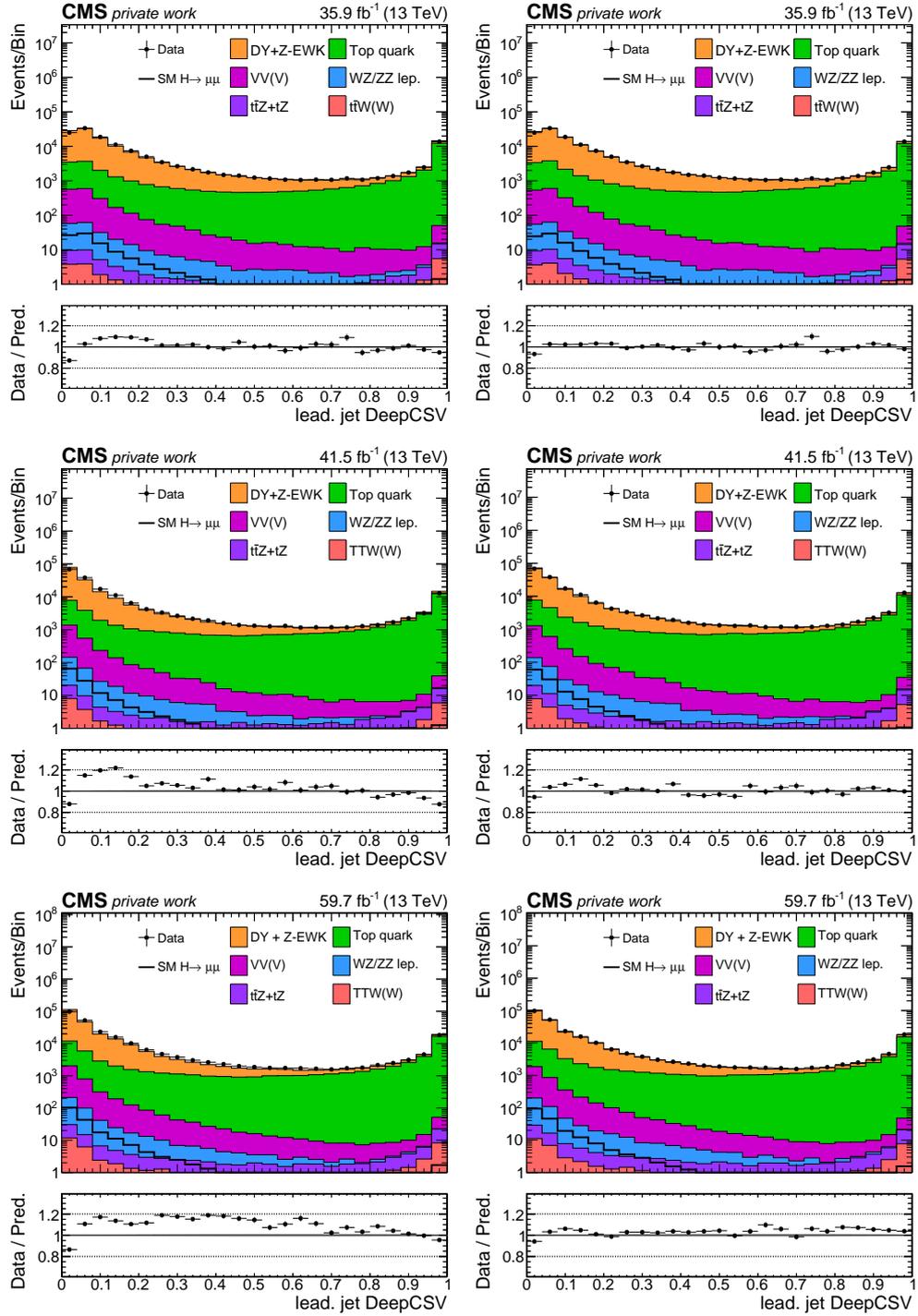


Figure 5.9: Comparison of the leading jet DeepCSV discriminant before (left) and after (right) the discriminator reshaping. The distributions are shown in the sideband region using the baseline dimuon selection. The comparison is provided for each year of data-taking: 2016 (top), 2017 (center), 2018 (bottom). All additional event weights considered in this thesis are applied. The prediction is scaled to the observed number of events.

Chapter 5. Higgs boson decays to muons produced in association with top quarks

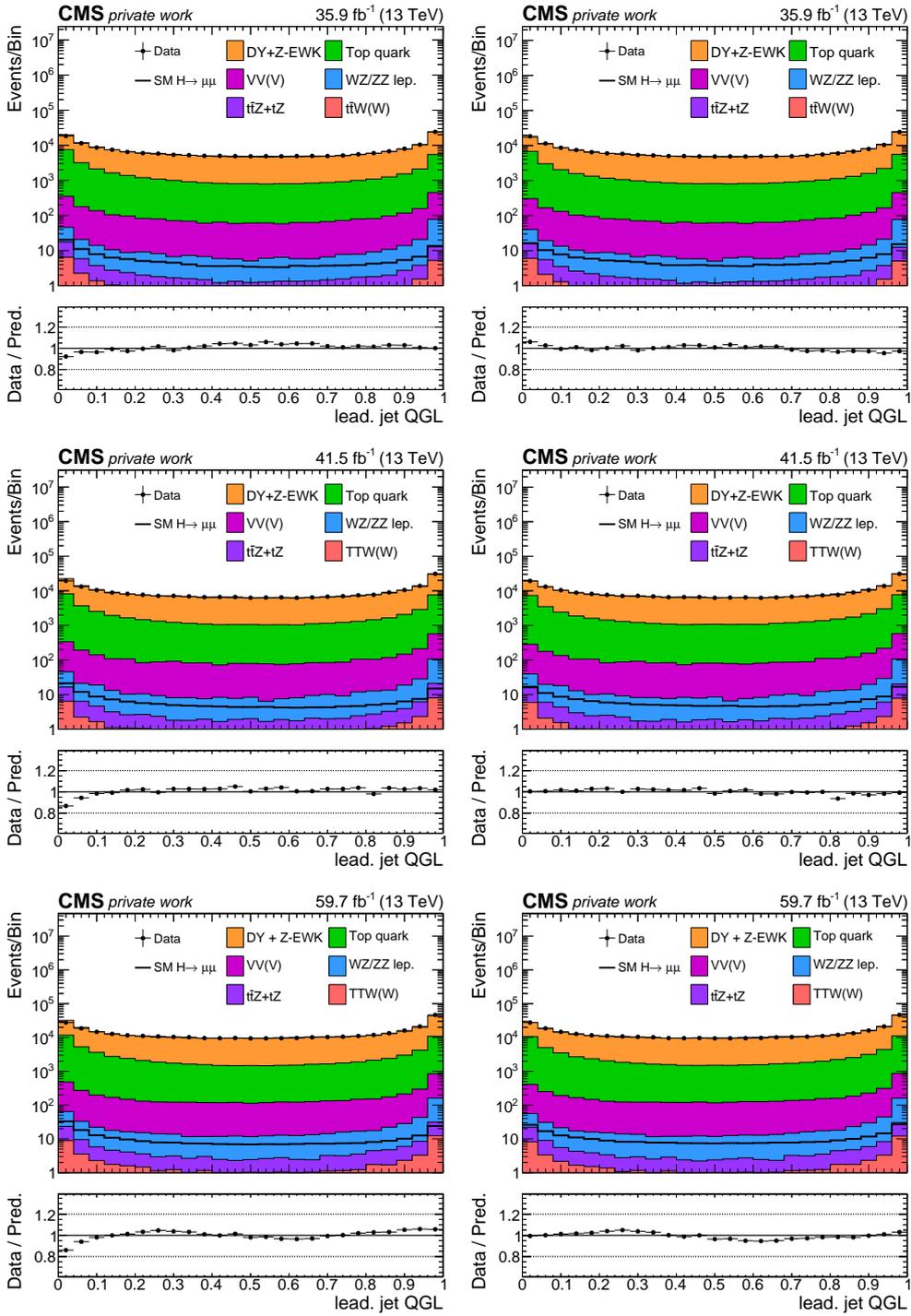


Figure 5.10: Comparison of the leading jet QGL before (left) and after (right) reweighting. The distributions are shown in the sideband region using the baseline dimuon selection. The comparison is provided for each year of data-taking: 2016 (top), 2017 (center), 2018 (bottom). All additional event weights considered in this thesis are applied. The prediction is scaled to the observed number of events.

5.4 Efficiency corrections & event reweighting

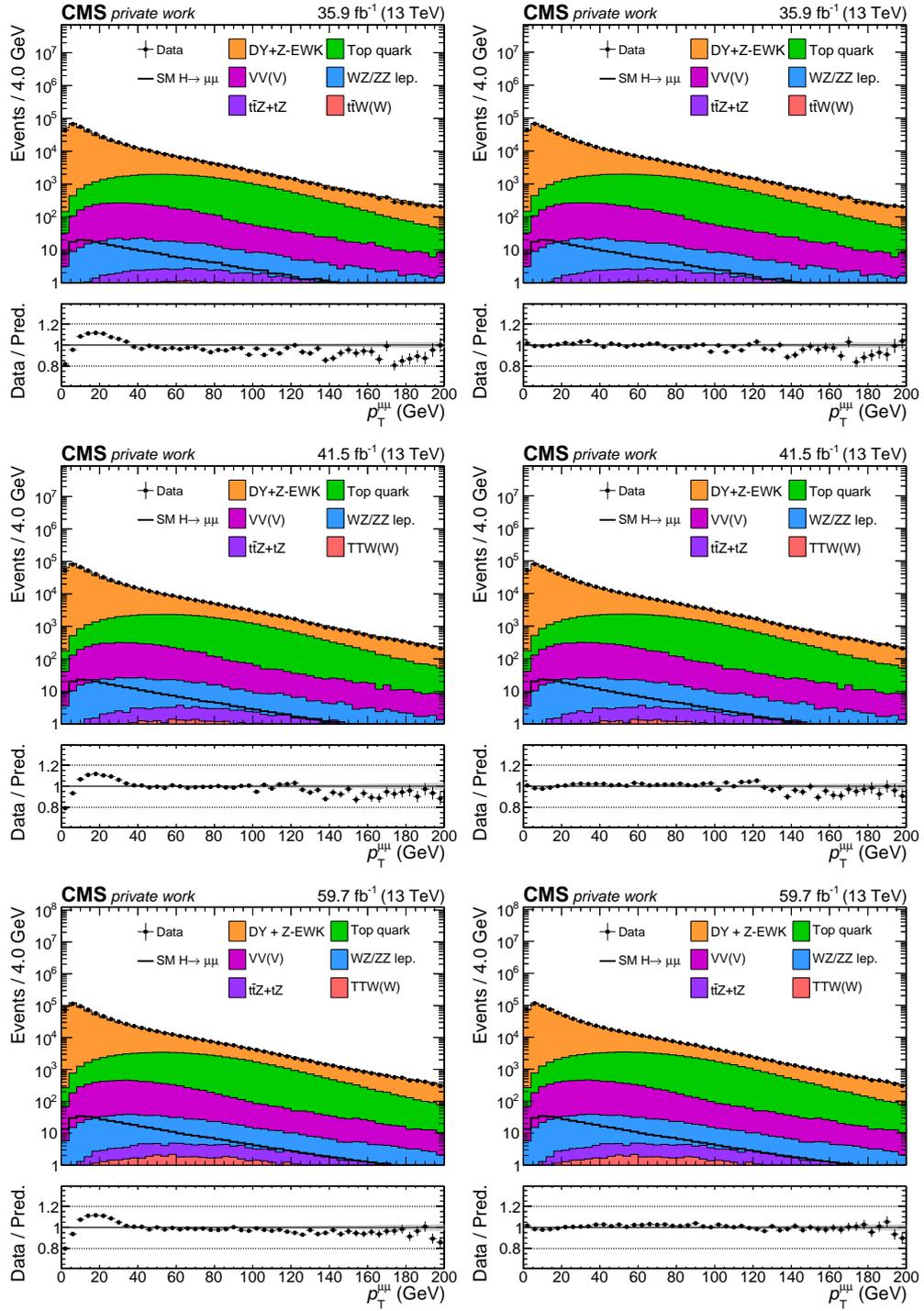


Figure 5.11: Comparison of the transverse dimuon momentum $p_T^{\mu\mu}$ before (left) and after (right) the $p_T^{\mu\mu}$ reweighting. The distributions are shown in the sideband region using the baseline dimuon selection. The comparison is provided for each year of data-taking: 2016 (top), 2017 (center), 2018 (bottom). All additional event weights considered in this thesis are applied. The prediction is scaled to the observed number of events.

5.4.6 Muon momentum scale and resolution

The sensitivity of the analysis relies particularly on the resolution and scale of the Higgs boson mass reconstruction, which in turn depend on the momentum scale and resolution of reconstructed muons. Following the method described in Ref. [184], the muon momentum is corrected to account for any bias in the muon momentum measurement which potentially arises from detector misalignment, software reconstruction effects, and uncertainties in the magnetic field configuration. The calibration of the correction is performed at the Z boson mass peak, which provides a very similar kinematic event topology compared to signal events. The method, which is referred to as *Rochester correction*, proceeds in two steps. First, the mean $\langle \frac{1}{p_T} \rangle$ for muons from Z boson decays is used to compute a momentum scale correction depending on the η , ϕ and charge of the muons. Second, the correction is further adjusted using the average dimuon mass $\langle m_{\mu\mu} \rangle$ of $Z \rightarrow \mu^+\mu^-$ events. The muon momentum scale is corrected in both data and simulation with the aim to align the reconstructed Z boson peak position with the precisely known Z boson mass value. The Z boson resolution in simulation is slightly underestimated compared to data. Therefore, an additional smearing of the muon momentum is applied in simulated events.

The resulting muon momentum scale and resolution for $H \rightarrow \mu^+\mu^-$ events after Rochester correction are presented in Fig. 5.12. They correspond to the mean and width of a Gaussian fit to the ratio of the true muon p_T , and the reconstructed muon p_T . The true muon p_T is obtained directly from the event generator. They are presented for all three data-taking periods collectively. To demonstrate the dependency on $|\eta|$ of the momentum resolution, three different distributions are provided based on an inclusive muon selection, on muons detected in the central detector region $0 < |\eta| < 0.9$, and on muons in the forward region $1.9 < |\eta| < 2.4$.

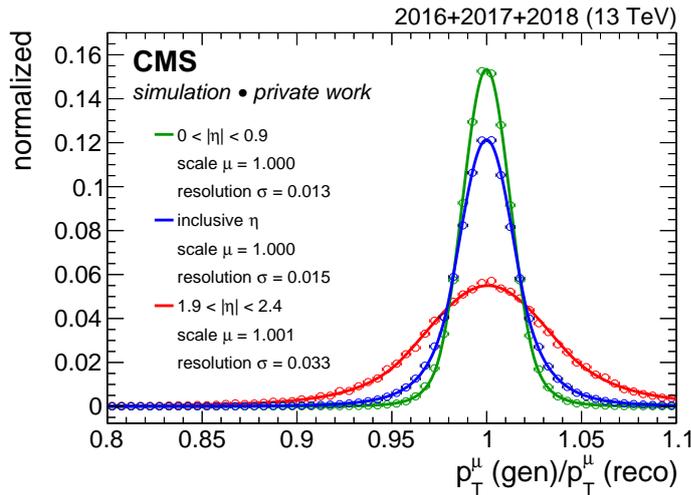


Figure 5.12: Ratio of the generated and reconstructed transverse muon momentum after Rochester correction for $H \rightarrow \mu^+\mu^-$ events summarized for all three data-taking periods. The muon momentum scale and resolution are extracted from a Gaussian fit. Three distributions represent the momentum scale and resolution of all selected muons (blue), central muons (green), and forward muons (red). Each distribution is normalized to have unit area.

The muon momentum resolution is found to be 1.3% in the central part of the detector up to $|\eta| = 0.9$. It degrades to 3.3% for muons in the forward region due to a reduced lever arm and increased multiple scattering within the tracking detector. The muon momentum scale of one implies that the generated and reconstructed muon momentum are accurately aligned within uncertainties for all $|\eta|$ -regions.

To avoid any bias on the final results of the analysis, the muon momentum scale and resolution correction can not be validated directly in the signal region at the Higgs boson mass peak. Instead the validation of the Rochester correction is performed at the Z boson mass peak. Figure 5.13 presents the dimuon mass distributions for events passing the dimuon baseline selection in a Z boson control region with a dimuon mass in the range between $80 < m_{\mu\mu} < 100$ GeV. These distributions are provided independently for each year of data taking before and after the Rochester correction is applied. Before correction, a clear shift of the peak position between data and simulation is observed in all three years. After the application of the Rochester muon momentum correction, this disagreement disappears completely, and simulated events precisely reproduce the distribution in data.

The remaining differences between data and simulation are investigated to estimate a systematic uncertainties on the signal peak position and resolution. The muon p_T resolution depends primarily on the η of the muon. The central barrel region of the detector has the best p_T resolution, while the forward region at large $|\eta|$ has a significant worse resolution. To take this dependency into account, events are categorized according to the muon $|\eta|$ in the *barrel* (B) region with $|\eta| < 0.9$, the *overlap* (O) region with $0.9 < |\eta| < 1.8$, and the *endcap* (E) region with $1.8 < |\eta| < 2.4$. Considering each muon individually, this yields nine resolution categories in total. In each of these categories, the Z boson peak is fitted independently in data and simulation using a Breit-Wigner function convoluted with a Gaussian. This accounts for the natural width of the Z boson, and the detector resolution, respectively. The peak and width parameters of the Breit-Wigner distribution are set to the nominal Z boson mass and width, while the Gaussian fit parameters are freely floating. In addition, the residual background component from for example $t\bar{t}$ events is modeled using an exponential function.

Figure 5.14 presents the fit values of the Z boson peak position and resolution after the application of the Rochester correction in data and simulation for each of the nine resolution categories in all three data-taking periods. The muon momentum scale, which is represented by the Z boson peak position, agrees within 0.1% between data and simulation across all resolution categories in all data-taking periods. The dimuon mass resolution is estimated from the fitted width σ of the Gaussian. Depending on the resolution category, the dimuon mass resolution varies between 0.8 GeV and 2.2 GeV in all three years, and generally agrees well between data and simulation. The remaining differences still indicate a slightly underestimated mass resolution in simulation compared to data. The maximal discrepancy is found to be 5% corresponding to muons reconstructed during the 2016 data-taking period in the center of the detector. In the final signal extraction fit these differences are completely covered by a systematic uncertainty in the signal peak position and resolution of 0.1% and 10%, respectively. Further discussion on systematic uncertainties is provided in Section 5.11.

Chapter 5. Higgs boson decays to muons produced in association with top quarks

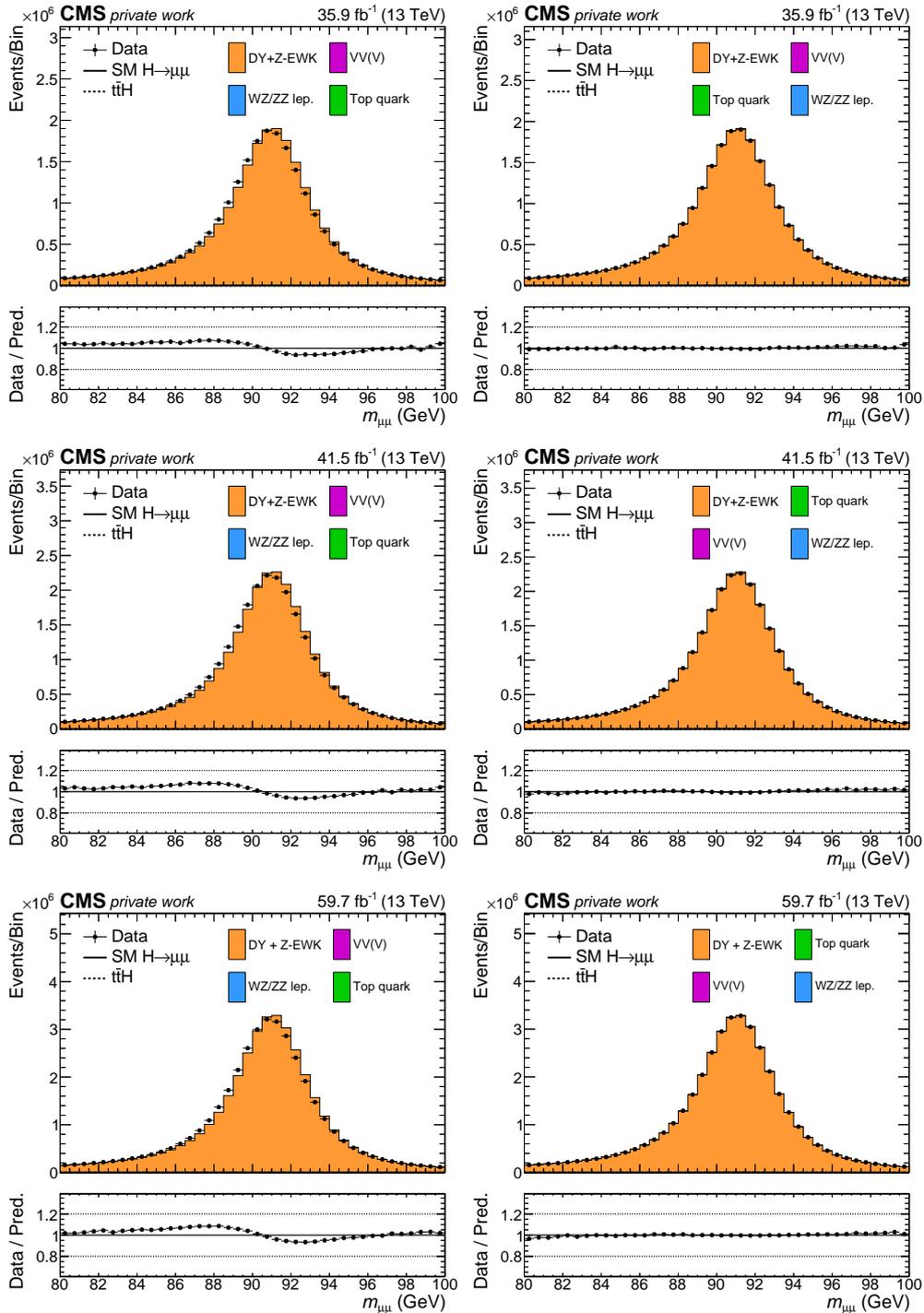


Figure 5.13: Comparison of the dimuon mass $m_{\mu\mu}$ before (left) and after (right) Rochester correction. The distributions are shown in a Z boson control region using the baseline dimuon selection. The comparison is provided for each year of data taking: 2016 (top), 2017 (center), 2018 (bottom). All additional event weights considered in this thesis are applied. The prediction is scaled to the observed number of events.

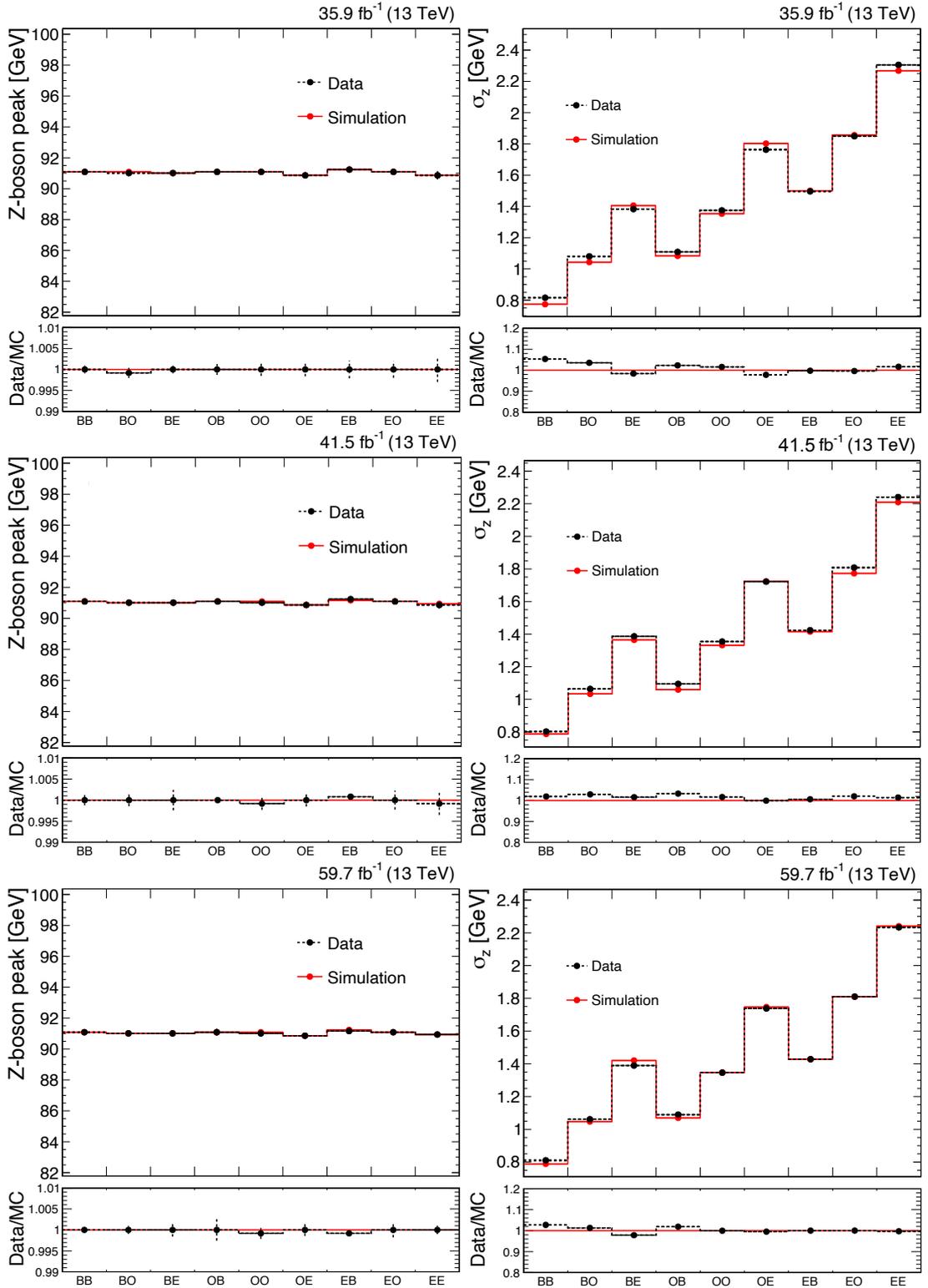


Figure 5.14: Peak position (left) and resolution (right) of $Z \rightarrow \mu^+\mu^-$ events extracted from an analytical fit to the dimuon mass distribution after Rochester correction. Events are grouped in nine exclusive resolution categories according to the muon $|\eta|$ with $|\eta| < 0.9$ (B), $0.9 < |\eta| < 1.8$ (O), and $1.8 < |\eta| < 2.4$ (E). Values are extracted independently for data and simulated events for each year of data taking: 2016 (top), 2017 (center), 2018 (bottom). Taken from Refs. [185, 186].

5.4.7 Per-event Higgs boson mass resolution

The sensitivity of the analysis depends on the resolution of the Higgs boson mass peak. The more precise the mass measurement, the more signal events concentrate in a narrow window around the Higgs boson mass, which increases the signal-to-background ratio. An estimate of the per-event dimuon mass resolution can therefore be helpful to improve the sensitivity of this analysis.

The previous CMS analysis HIG-17-019 [41] classified events into three resolution categories according to the maximum muon $|\eta|$ in the event as described in Section 1.4. This provides a rough estimate of the actual dimuon mass resolution for each event, due to the correlation of the muon $|\eta|$ and the muon momentum resolution. The three resolution categories were combined with five categories based on the output of a multivariate event discriminant. In total, this led to 15 event categories used to extract the signal.

The new analysis HIG-19-006 [172] follows an optimized procedure to account for the dimuon mass resolution. The uncertainty in the measured dimuon mass $\sigma_{\mu\mu}$ is derived individually for each dimuon pair using the covariance matrix of the muon track fits. This provides a more accurate measure of the per-event dimuon mass resolution compared to an approximation based on the muon $|\eta|$. Figure 5.15 presents the relative dimuon mass uncertainty for the previously introduced nine resolution categories in Section 5.4.6. Depending on the $|\eta|$ of the muons, the relative mass uncertainty varies between 0.8% and 5%. The best resolution is found for events where both muons are reconstructed in the central part of the detector. The muon momentum resolution of the leading muon in p_T dominates the dimuon mass resolution, which is implied by a comparison of the resolution categories "BO/OB" for example. On average the resolution degrades with increasing $|\eta|$ of the muons.

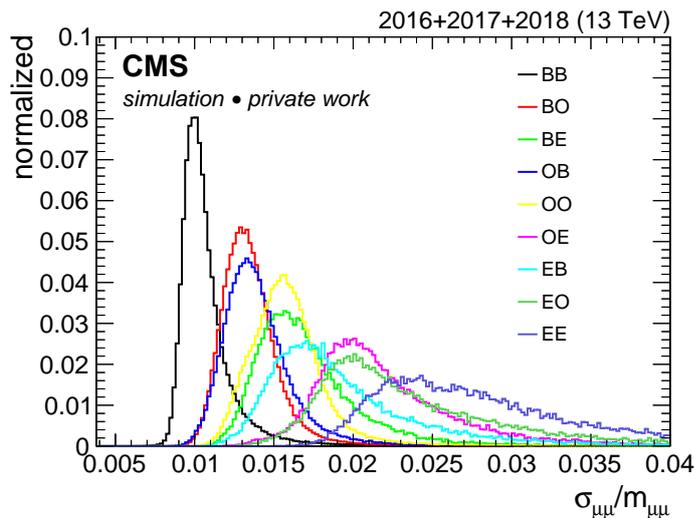


Figure 5.15: Relative per-event dimuon mass uncertainty of $H \rightarrow \mu^+\mu^-$ events for all three years of data taking. The distributions are shown for different combinations of the leading and sub-leading muon $|\eta|$ as indicated in the figure. Each distribution is normalized to have unit area.

Apart from the more precise definition of the per-event dimuon mass resolution, also the way is changed how this information is incorporated. Instead of a cut-based definition of resolution categories, the dimuon mass resolution information is embedded directly into the training of the multivariate discriminant described in more detail in Section 5.9.

As discussed in Section 5.4.6, the muon momentum measurement before Rochester correction might have some bias. Since the Rochester correction only calibrates the muon p_T but not the underlying track fit parameters, the mass uncertainty determined from the covariance matrix of the muon track fit is potentially affected in the same way. To ensure that the true dimuon mass resolution is modeled correctly, a calibration of the per-event mass uncertainty is performed using $Z \rightarrow \mu^+\mu^-$ events independently for data and simulation. The calibration method is described in Appendix A.

Figure 5.16 presents the distribution of the per-event mass resolution in data and simulation for events passing the baseline dimuon selection with a dimuon mass in the range between $110 < m_{\mu\mu} < 150$ GeV. Before calibration of the per-event mass resolution, simulated events underestimate the number of events in the tail of the distribution with a relative mass resolution of 1.8% and larger values. While the disagreement between data and simulation is in the order of 10-15% in 2016 and 2018, it becomes larger than 20% in 2017. The performance of the new pixel tracking detector may have not yet been modeled as precisely in 2017 as in 2018, which can be a reason for the differences between the data-taking periods. The calibration allows to significantly improve the agreement between data and simulation in all three years, and simulated events can accurately reproduce the shape of the central part of the distribution observed in data. In general, the shape of the per-event mass resolution distribution changes slightly after calibration and tends to have larger values. A remaining small disagreement between data and simulation is visible in the tail of the distribution at large values of the relative per-event mass resolution in 2017. Since only a very small number of events is affected, the overall impact of this disagreement on the analysis is negligible.

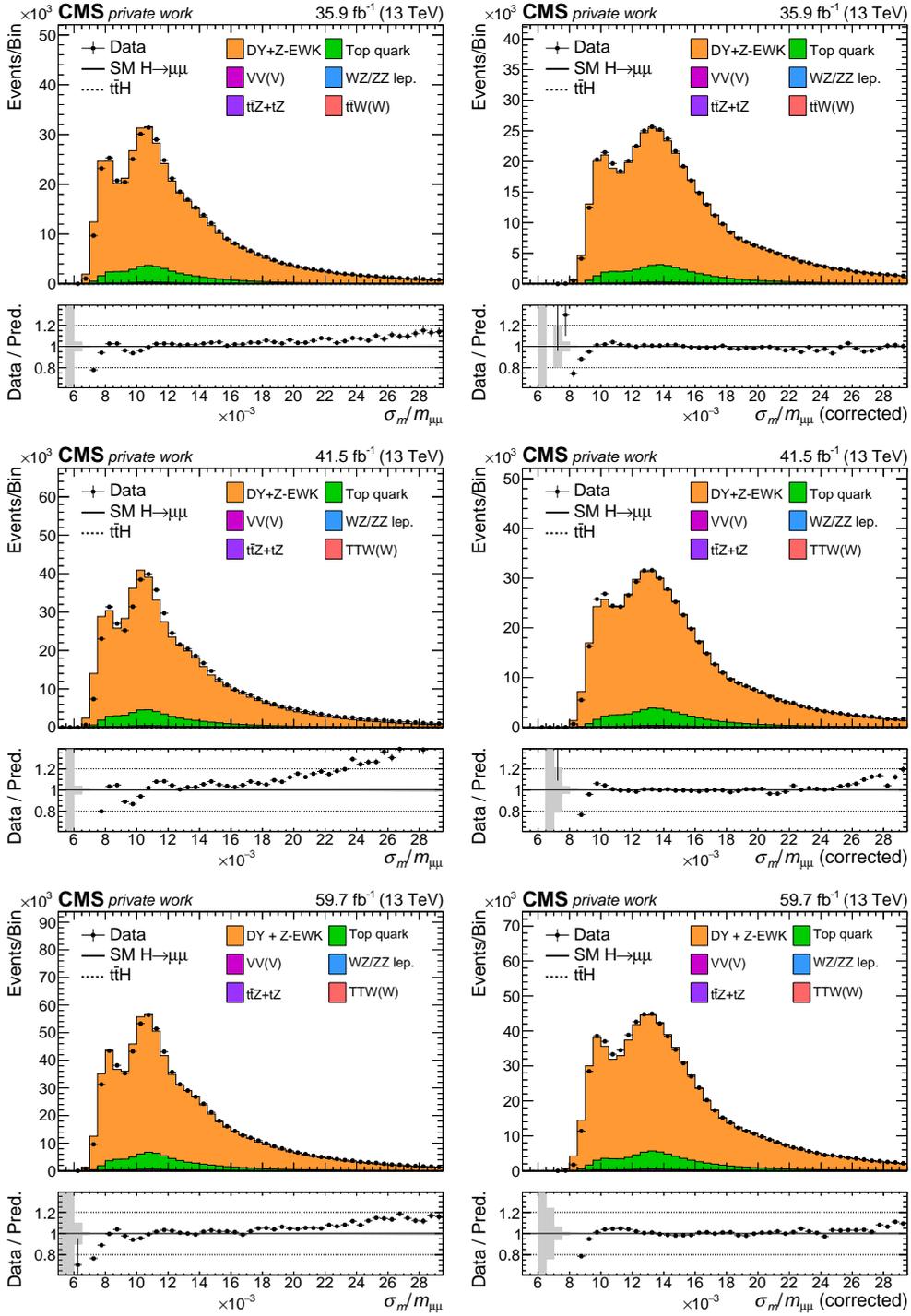


Figure 5.16: Comparison of the per-event mass resolution before (left) and after (right) calibration. The distributions are shown in the sideband region using the baseline dimuon selection. The comparison is provided for each year of data taking: 2016 (top), 2017 (center), 2018 (bottom). All additional event weights considered in this thesis are applied. The prediction is scaled to the observed number of events.

5.5 Improvements to dimuon mass reconstruction

Dimuon mass reconstruction methods which improve the signal acceptance and the signal peak resolution provide a natural way to optimize the analysis sensitivity independent of the Higgs boson production mode. Two methods, which are employed in all four Higgs boson production categories, are described in the following. The first method aims to recover the initial momentum of muons, which radiate a photon. The second method improves the dimuon mass resolution by an empirical correction to the reconstructed muon track momentum.

5.5.1 Final state photon recovery

The muons emerging from the Higgs boson decay may lose some momentum due to the radiation of a photon in the final state. The corresponding photons are referred to as *FSR photons*. This process is a QED correction to the $H \rightarrow \mu^+\mu^-$ decay. The radiation of FSR photons leads to a bias in the dimuon mass reconstruction if not taken into account. This effect can be mitigated by including FSR photons to the invariant mass of the reconstructed Higgs boson candidate. Furthermore, FSR photons can induce a muon selection inefficiency, if the muon isolation criterion is failed due to an FSR photon inside the isolation cone. The acceptance of muons can therefore be increased by excluding FSR photons from the muon isolation. Due to different background sources such as detector noise, underlying event, and pileup, the selection of FSR photons is experimentally challenging.

Figure 5.17 demonstrates the impact of FSR photons on the reconstructed signal mass peak, which is shown for events passing the baseline dimuon selection. To identify genuine FSR photons, the reconstructed FSR photon candidates are matched to photons radiated by a muon on

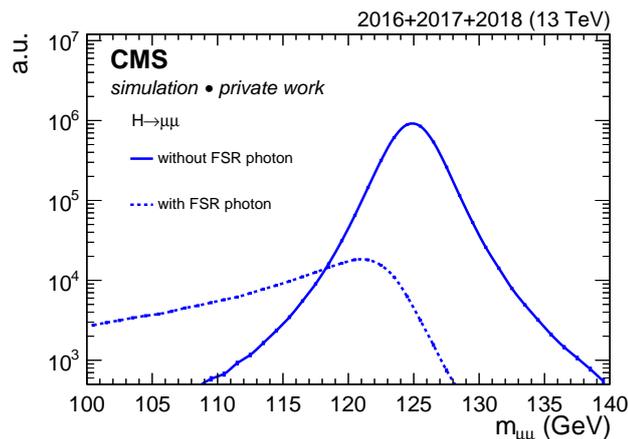


Figure 5.17: Invariant dimuon mass of simulated $H \rightarrow \mu^+\mu^-$ events for all three data-taking periods combined. The fraction of events containing reconstructed FSR photons with $E_T^Y > 2$ GeV matched to photons radiated by a muon on event generator level, and events without FSR photons are presented.

event generator level. A match is established, if a reconstructed photon is found inside a cone of $\Delta R = 0.2$ around the generated photon. In total, 5.5% of signal events contain at least one reconstructed FSR photon with a transverse energy $E_T^\gamma > 2 \text{ GeV}$. Since the FSR photon carries a certain energy fraction of the initially produced Higgs boson, only events with a dimuon mass smaller than the mass of the Higgs boson are affected. This results in an asymmetric smearing of the dimuon mass distribution towards smaller mass values.

First studies on the Z boson which include a method to recover FSR photons [12, 187] were carried out at the UA1 [188] and UA2 [189] experiments, and similar approaches have also been implemented in different CMS analyses [190, 191]. The latest CMS result using an FSR photon recovery strategy is provided by the measurement of $H \rightarrow ZZ^* \rightarrow 4\ell$ [192], where the baseline photon selection used here is taken from. The FSR photon recovery strategy presented in the following is particularly optimized for the $H \rightarrow \mu^+\mu^-$ analysis and implemented in all Higgs boson production categories used for the combined result.

Baseline FSR photon selection

Photons are reconstructed using the PF algorithm from local, isolated energy deposits in the ECAL, which are not associated to any track. The PF algorithm also allows to reconstruct photons that convert into electron-positron pairs in the detector material.

Among all reconstructed PF photons, FSR photon candidates are selected with $E_T^\gamma > 2 \text{ GeV}$ and $|\eta| < 1.4$, $1.6 < |\eta| < 2.4$. This avoids the selection of photons close to the gap between the ECAL barrel and endcap. FSR photon candidates are required to lie in a cone with $\Delta R = 0.5$ around a muon. Photons are rejected, if they are matched to a reconstructed electron, or to one of the associated bremsstrahlung photons. Each photon is assigned to the closest muon, and the transverse energy of a photon is restricted to values smaller than the transverse energy of the corresponding muon. In case multiple FSR photon candidates per muon are found, the one with the smallest energy weighted distance $\Delta R(\mu, \gamma)/E_{T,\gamma}^2$ to the muon is chosen. The selection of multiple photons per muon is found to spoil the dimuon mass resolution, since the probability to pick a photon induced by PU gets larger.

Optimization of FSR photon selection

Additional selection criteria are employed to reject misidentified photons, photons from PU vertices, and photons inside jets, which are collectively referred to as non-prompt photons in the following. Since these non-prompt photons are not correlated with the muon kinematics, they induce a non-resonant background. In contrast, the background contribution from $H \rightarrow Z\gamma \rightarrow \mu^+\mu^-\gamma$ where a Higgs boson decays to a Z boson and a prompt photon γ is to be considered as a resonant background, since the three-body mass of the muon pair from the Z boson decay and the prompt photon yields the Higgs boson mass.

The FSR photon selection is optimized in two steps. In the first step, three different kinematic observables are exploited to distinguish FSR photons from prompt photons occurring

in $H \rightarrow Z\gamma \rightarrow \mu^+\mu^-\gamma$ events. This study aims to find a single observable able to suppress the resonant background to a negligible level. In the second step, a cut-based optimization of two commonly used observables is carried out, which improves the FSR photon purity and reduces the background from non-prompt photons.

The optimization of the FSR photon selection criteria is performed on simulated $H \rightarrow \mu^+\mu^-$ events using the baseline dimuon selection. Event generator information are used to extract the fraction of events containing genuine FSR photons as described above.

Figure 5.18 presents characteristic kinematic features of FSR photons in comparison to prompt photons. The distributions of the transverse energy of the photon E_T^γ and the distance between the photon and the muon $\Delta R(\mu, \gamma)$ imply that most of the muons radiate soft, i.e., low-energetic FSR photons collinear to the muon, while prompt photons have larger energies and are well separated from the muon. The energy ratio between the photon and the muon E_T^γ/E_T^μ shows that the FSR photon energy is typically only a small fraction of the associated muon energy as expected. In contrast, prompt photons from $H \rightarrow Z\gamma \rightarrow \mu^+\mu^-\gamma$ have in most cases an energy of around 80 – 90% compared to the muon energy.

In the first step, a cut-based optimization of the three observables is performed to define a pure selection of FSR photons, which strongly reduces the $H \rightarrow Z\gamma$ background. Since the $H \rightarrow Z\gamma$ process has not been observed yet, this processes could in principle be modified by physics beyond the SM, leading to a larger production rate compared to the SM prediction. The latest observed limits by CMS (ATLAS) on the cross section times branching fraction are 3.9 (3.6) times the SM cross section [193, 194]. For this reason, the product of the signal selection efficiency and the background rejection $\epsilon_S \times (1 - \epsilon_B)$ is used to optimize the cut value of the different observables, instead of absolute signal and background rates. The best cut value corresponds to the maximum of $\epsilon_S \times (1 - \epsilon_B)$. The results of the cut optimization are shown in Fig. 5.18. Among these three observables, the energy ratio E_T^μ/E_T^γ is the most discriminating observable with a best cut value at 0.4, as it provides the largest value for $\epsilon_S \times (1 - \epsilon_B)$. This allows the FSR photon candidate to have a maximal energy of 40% of the muon energy. The corresponding cut efficiency with respect to the baseline FSR photon selection is about 85% for FSR photons in signal events while the background rejection is 92%. The contribution of SM $H \rightarrow Z\gamma$ events, which pass the full dimuon event selection and the FSR photon requirements, is reduced to around 0.1% of the overall $H \rightarrow \mu^+\mu^-$ signal and is thus neglected in the further analysis.

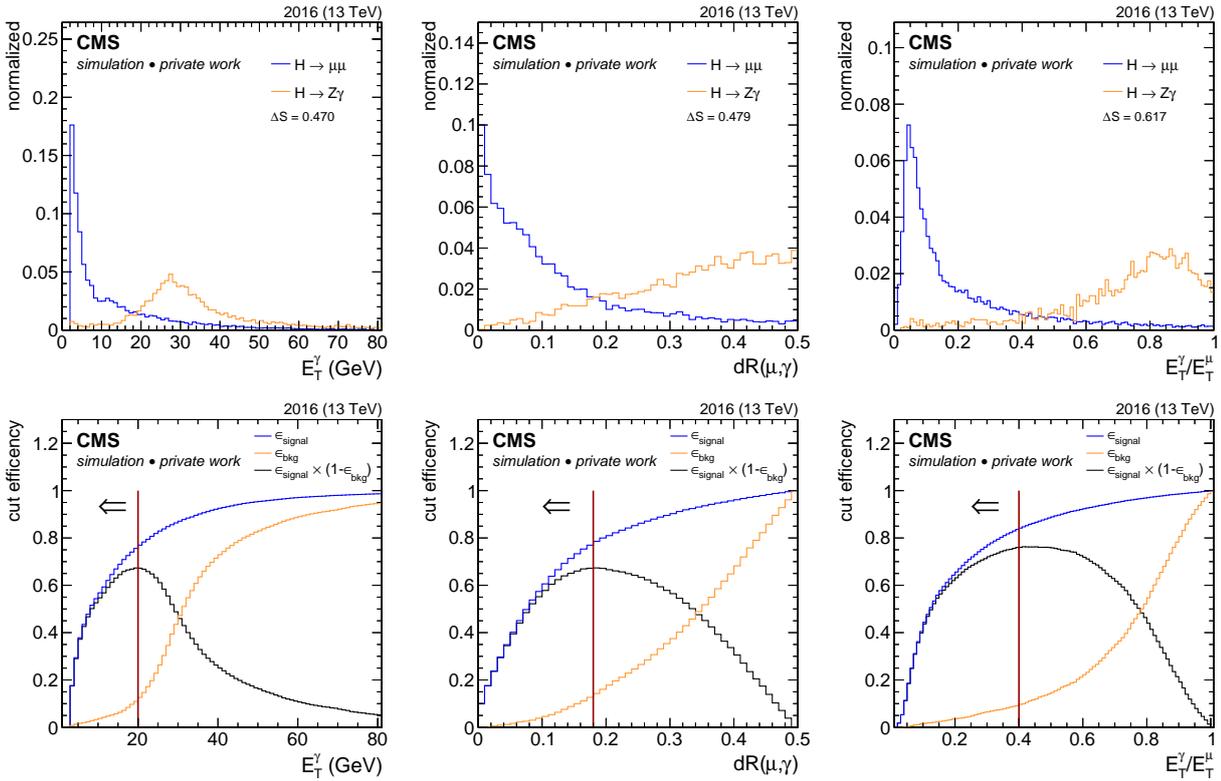


Figure 5.18: Kinematic observables of FSR photons from $H \rightarrow \mu^+\mu^-$ events and prompt photons from $H \rightarrow Z\gamma$ events using the 2016 data-taking conditions. Photons are reconstructed using the baseline photon selection, and matched to generated photons. Shown are the transverse photon energy E_T^γ (top left), the distance between the muon and the photon $\Delta R(\mu, \gamma)$ (top center), and the energy ratio of the photon and the muon E_T^μ/E_T^γ (top right). The corresponding optimization of cut values is presented for each of the observables in the bottom.

After the first optimization step, the FSR photon selection is further improved using the relative photon isolation I_γ and the energy weighted angular distance between the muon and the photon $\Delta R(\mu, \gamma)/E_{T,\gamma}^2$, which combines $\Delta R(\mu, \gamma)$ and E_T^γ into a single observable. The isolation of a photon is defined inside a cone with radius $R = 0.3$, where hadrons as well as photons are considered including the contribution from PU vertices. The p_T threshold on charged (neutral) hadrons is set to 0.2 (0.5) GeV. Both observables particularly aim to distinguish FSR photons from non-prompt photons and have already been used in previous analysis [192]. To illustrate the kinematic properties of FSR photons and non-prompt photons, simulated $Z \rightarrow \mu^+\mu^-$ events are selected with a dimuon mass in the range between $70 < m_{\mu\mu} < 110$ GeV. Photons are identified as FSR photons or non-prompt photons based on event generator information. Figure 5.19 shows the distributions of $\Delta R(\mu, \gamma)/E_{T,\gamma}^2$ and rel. I_γ for FSR photons and non-prompt photons. The $\Delta R(\mu, \gamma)/E_{T,\gamma}^2$ distribution of FSR photons is steeply falling towards larger values while for non-prompt photons the tail of the distribution is rather flat. The distribution of the relative isolation demonstrates that FSR photons tend to be more isolated compared to non-prompt photons from hadronic jets for example.

5.5 Improvements to dimuon mass reconstruction

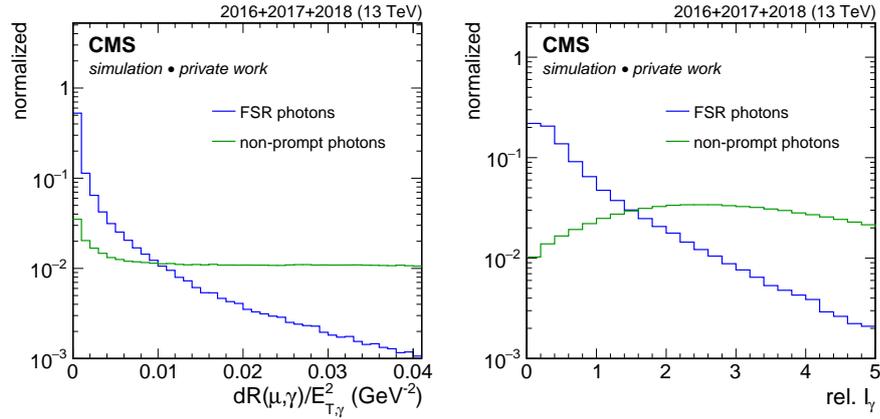


Figure 5.19: Distance between the muon and the photon normalized to the photon transverse energy squared $\Delta R(\mu, \gamma)/E_{T,\gamma}^2$ (left), and the relative photon isolation I_γ (right) for FSR photons and non-prompt photons in simulated $Z \rightarrow \mu^+\mu^-$ events with $70 < m_{\mu\mu} < 110$ GeV. The distributions are scaled to have unit area.

The cut values of both observables are optimized simultaneously in terms of a 2D scan. For this purpose, $H \rightarrow \mu^+\mu^-$ signal events and the dominant background processes from Drell-Yan, dileptonic $t\bar{t}$, and different diboson processes for all three data-taking periods are used. The processes are weighted according to their cross section. To account for the impact of the FSR photon recovery on the dimuon mass resolution, the significance is calculated as

$$Z_{\Delta m=500 \text{ MeV}} = \sum_i^{n=20} S_i^2 / B_i, \quad (5.4)$$

where S_i and B_i are the number of expected signal and background events in the i -th mass bin from 120 to 130 GeV. This corresponds to a 500 MeV mass bin width which is slightly smaller than the dimuon mass resolution. Figure 5.20 presents the optimization of $Z_{\Delta m=500 \text{ MeV}}$ and the resulting optimal cut values $\Delta R(\mu, \gamma)/E_{T,\gamma}^2 < 0.012$ and $\text{rel. } I_\gamma < 1.8$.

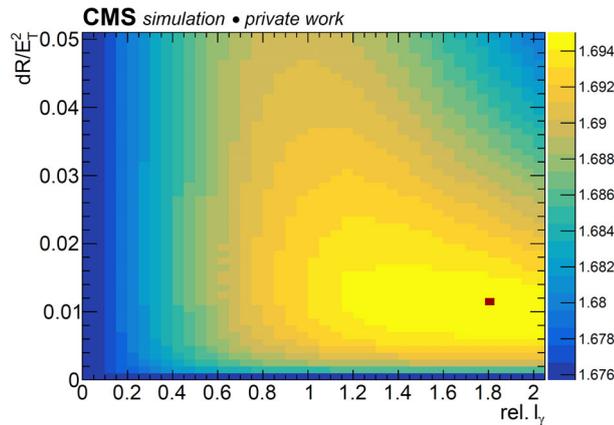


Figure 5.20: Significance $Z_{\Delta m=500 \text{ MeV}}$ as a function of different cut values for the FSR photon section criteria $\Delta R(\mu, \gamma)/E_{T,\gamma}^2$ and $\text{rel. } I_\gamma$. The optimization is performed using simulated $H \rightarrow \mu^+\mu^-$ events and the dominant background contributions for all three data-taking periods. The chosen working point is indicated by the red square.

Performance and validation

The final performance of the FSR photon recovery based on the optimized selection criteria is presented in Figs. 5.21 and 5.22. The distribution of FSR-tagged events on the left in Fig. 5.21 demonstrates an accurate recovery of the Higgs boson mass peak after FSR photons are included in the calculation of the invariant Higgs boson mass. The overall impact of FSR photon recovery on the signal peak and the signal acceptance is presented in the distribution on the right in Fig. 5.21. Around 3% of all signal events are affected by the recovery of FSR photons. The *root-mean-square* (RMS), which is a measure of the width of the dimuon mass distribution, is improved by 3% while the signal acceptance is increased by about 2% due to the removal of FSR photons from the muon isolation. The final selection requirements provide a selection efficiency of 58% with a purity of 85%, where the FSR photon efficiency is the fraction of selected photons among all reconstructed FSR photon candidates, and the FSR photon purity is the fraction of true FSR photons among the selected photons.

Finally, the FSR photon recovery strategy is validated in data using dimuon events in a Z boson control region with $70 < m_{\mu\mu} < 110$ GeV. Figure 5.22 shows the reconstructed Z boson mass for dimuon events with at least one selected FSR photon before and after FSR photon recovery. The Z boson mass peak is accurately recovered in data and simulation. In general, simulated events can reproduce the observed distribution in data before and after FSR photon recovery. The small peak in the dimuon mass distribution at 91 GeV, which is visible even before FSR photons are recovered, indicates events where a photon is misidentified as FSR photon. This effect seems to be more pronounced in data compared to simulation. Simulated events tend to underestimate the number of events in data for $m_{\mu\mu} > 95$ GeV by 15%. These differences are observed before as well as after the recovery of FSR photons. Since this discrepancy only affects the small fraction of events which include FSR photons, the overall impact of this mis-modeling on the signal shape is negligible.

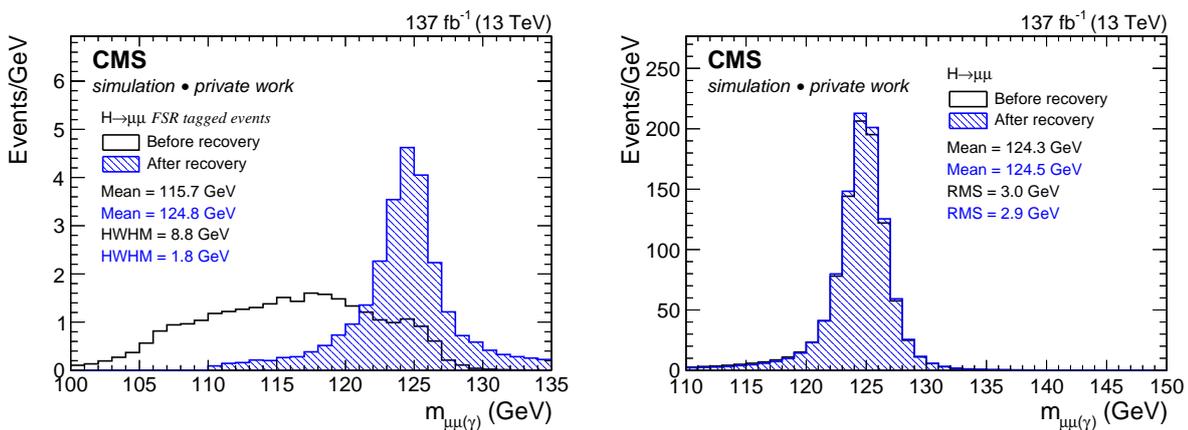


Figure 5.21: Performance of the FSR photon recovery in simulated $H \rightarrow \mu^+\mu^-$ events. The reconstructed Higgs boson candidate mass is shown before and after FSR photon recovery for events with at least one selected FSR photon (left) and for all events (right).

5.5 Improvements to dimuon mass reconstruction

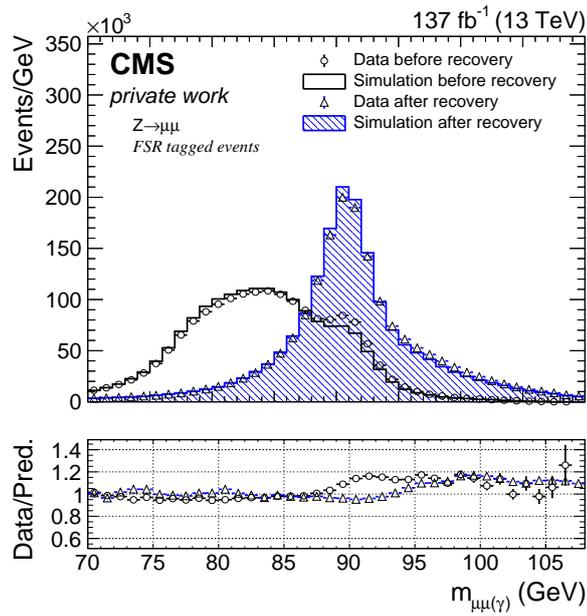


Figure 5.22: Validation of the FSR photon recovery in data using events selected in a Z boson control region with $70 < m_{\mu\mu} < 110$ GeV. The reconstructed Z boson mass is shown before and after FSR photon recovery for events with at least one selected FSR photon.

The presented FSR photon recovery method is implemented in all Higgs boson production categories included in the final combined $H \rightarrow \mu^+\mu^-$ analysis, and improves the expected analysis sensitivity by about 3%.

5.5.2 Geometric track fit

Higgs bosons created in proton-proton collisions decay on a very short time scale, $\sim 10^{-22}$ s. Thus, the muon pair emerging from the decay of the Higgs boson is produced directly at the interaction point. The final track fit, however, is performed without a constraint to the primary interaction point to ensure a possible reconstruction of displaced secondary vertices. Therefore, reconstructed muon tracks might still show a small, residual displacement from the primary interaction point. This knowledge can be used to improve the resolution of the reconstructed dimuon mass by including the position of the interaction point in the final track fit. To prevent a computationally intensive track re-fitting procedure, a new method was developed within the CMS $H \rightarrow \mu^+\mu^-$ working group [172, 195] to adjust the muon p_T , which is on average equivalent to a full track re-fit. This novelty, referred to as GEOFIT, is explained here for completeness.

Track geometry in the transverse plane

The minimum distance in the transverse plane between a reconstructed muon track and the interaction point d_0 is found to be proportional to the muon momentum resolution Δp_T . This relation can be derived for high- p_T tracks using the geometry of a true track originating from the primary interaction point, and a reconstructed track, which is slightly displaced in the transverse plane. The position of the primary interaction point is estimated using the beam spot. The underlying geometric consideration is illustrated in Fig. 5.23. A rectangular triangle can be formed between the reconstructed track and the true track assuming high- p_T tracks and small distances d_0 and x . One obtains

$$d_0 = x \cdot \sin \alpha. \quad (5.5)$$

Applying the sine law to the triangle spanned by s , R , and $R - \Delta R$ yields

$$\frac{s}{\sin \alpha} = \frac{R}{\sin(\pi - \beta)} = \frac{R}{\sin \beta}. \quad (5.6)$$

Using again trigonometric ratios one obtains

$$\sin \beta = \frac{L}{2(R - \Delta R)}. \quad (5.7)$$

The Pythagorean theorem implies

$$s = \sqrt{R^2 - \left(\frac{L}{2}\right)^2} - \sqrt{(R - \Delta R)^2 - \left(\frac{L}{2}\right)^2}, \quad (5.8)$$

which is simplified assuming the radius R to be much larger than the length L to $s \approx \Delta R$. Assuming additionally only small relative differences between the radii of the true tracks and the reconstructed tracks $R \gg \Delta R$, Eqs. (5.6) to (5.8) yield a relation between d_0 and $\Delta R/R^2$. Since

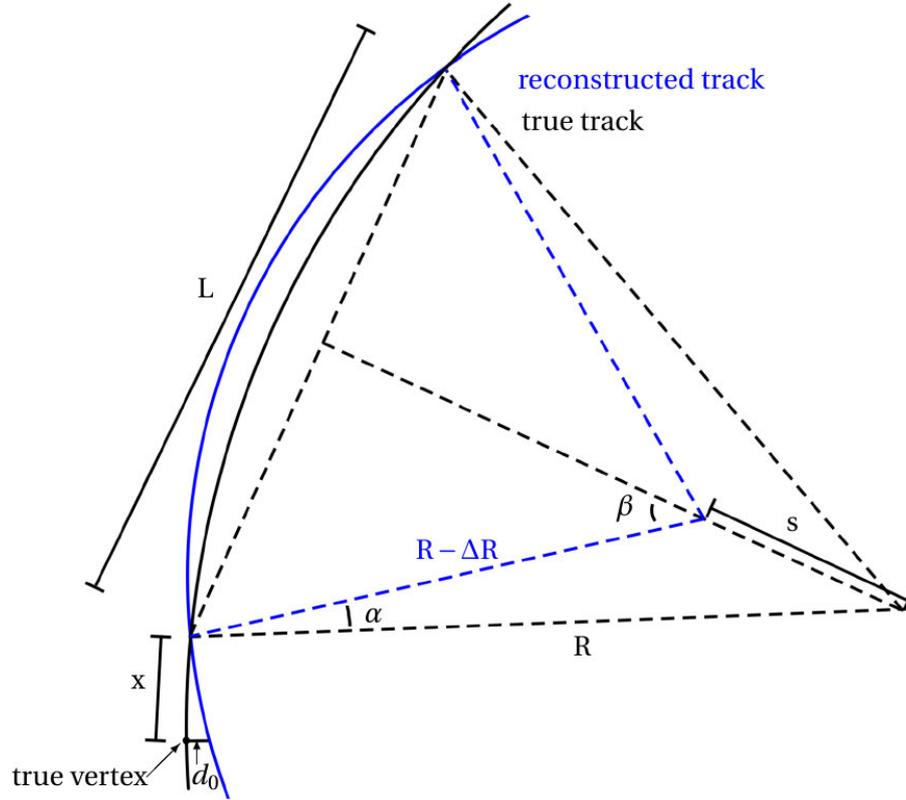


Figure 5.23: Illustration of the track geometry used to derive a muon p_T correction based on the minimum distance in the transverse plane between the reconstructed track and the interaction point.

the radius of a particle trajectory in a magnetic field is proportional to its momentum, one finally obtains

$$d_0 \simeq \frac{xL}{2} \cdot \frac{\Delta R}{R^2} \longrightarrow d_0 \propto \frac{\Delta p_T}{p_T^2}. \quad (5.9)$$

The constant of proportionality consists of unknown quantities, which vary event-by-event. Therefore, the average value of the constant of proportionality is extracted from a linear fit in three η regions, and separately for each data-taking period using simulated $Z \rightarrow \mu^+\mu^-$ events. Since tracks are bend in different direction depending on the muon charge, the correction has reversed sign for muons compared to anti-muons. To take this into account, the value of d_0 is multiplied by the muon charge. The fit results are used to adjust the muon p_T .

Performance and validation

The impact of the GEOFIT correction on data and simulation is presented in Fig. 5.24 using events in a Z boson control region. The d_0 distribution in $Z \rightarrow \mu^+\mu^-$ events is approximately Gaussian, and centered at zero. Around 95% of muons have a $|d_0|$ smaller than $40 \mu\text{m}$. The Z boson mass is estimated by an analytical fit to the $m_{\mu\mu}$ distribution in different bins of $|d_0|$ using a convolution of a Breit-Wigner and a Gaussian function. Before the GEOFIT is applied, a trend in the measured Z boson mass proportional to d_0 is observed, which degrades the dimuon mass resolution. The maximal deviation of the fitted peak value from the nominal Z boson mass is about 1 GeV at $|d_0| = 70 \mu\text{m}$. After the GEOFIT correction is applied, this trend is much reduced while the level of agreement between data and simulation is preserved. The alignment of the peak position intrinsically improves the resolution of the dimuon mass.

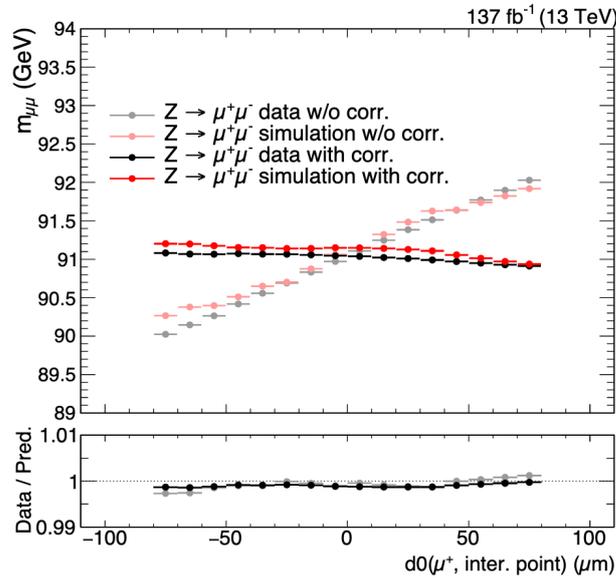


Figure 5.24: Evolution of the measured Z boson mass in data (black points) and simulation (red points) as a function of the minimum distance in the transverse plane between the track of the positive charged muon in the event and the interaction point ($d_0(\mu^+, \text{IP})$). The solid (shaded) points indicate the results obtained after (before) the correction. Based on data provided by author of Ref. [195].

Figure 5.25 presents the dimuon mass resolution for simulated $H \rightarrow \mu^+\mu^-$ events before and after applying GEOFIT corrections. The resolution is examined in nine resolution categories, independently for the four main Higgs production modes ggH , VBF , VH , and $t\bar{t}H$. The results are reported for a combination of all three data-taking periods. After GEOFIT the resolution is improved in all resolution categories and for all Higgs boson production modes. The improvement seems to be stable across the different $|\eta|$ regions. The average improvement varies depending on the Higgs boson production mode between 6.8–9.4%. The largest improvement is observed in $t\bar{t}H$ events. Muons from the decay of the Higgs boson produced via $t\bar{t}H$ tend to have larger p_T . Since high- p_T muons are affected the most, the overall improvement is larger for $t\bar{t}H$ events compared to the other Higgs boson production modes.

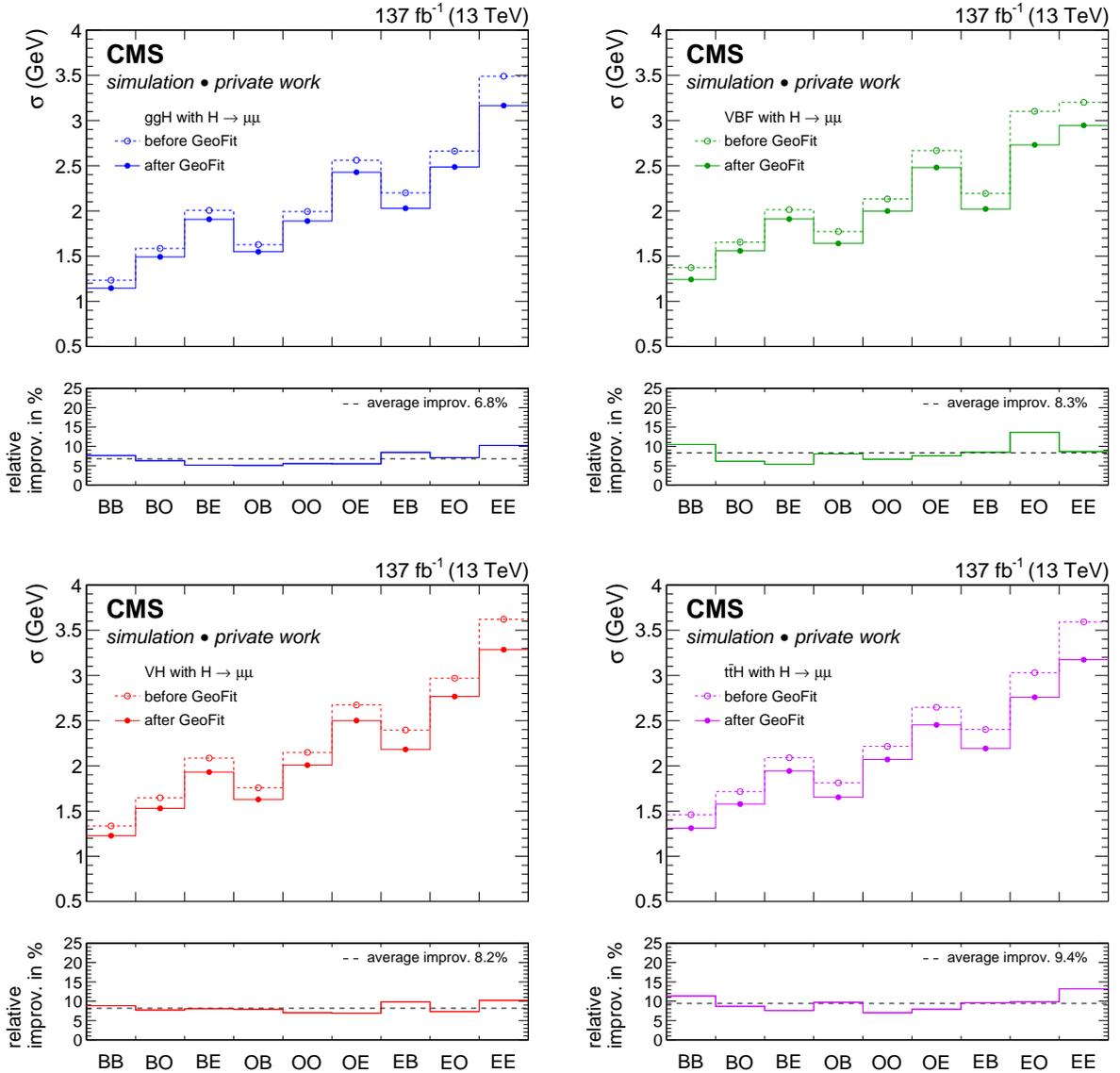


Figure 5.25: Comparison of the dimuon mass resolution before (dashed line) and after (solid line) GEOFIT correction in $H \rightarrow \mu^+\mu^-$ signal events. The resolution is extracted from an analytical fit to the dimuon mass distribution independently for ggH (top left), VBF (top right), VH (bottom left), and ttH (bottom right) signal based on a combination of all three data-taking periods. Events are grouped in nine exclusive resolution categories depending on the leading and sub-leading muon $|\eta|$. The considered regions are the barrel (B) region with $|\eta| < 0.9$, the overlap (O) region with $0.9 < |\eta| < 1.8$, and the endcap (E) region with $1.8 < |\eta| < 2.4$. The solid line in the lower panel shows the relative improvement in each of the resolution categories. The dashed line corresponds to the average improvement.

5.6 Multivariate lepton identification

Leptons originating from the decay of a massive particle like the Higgs boson or electroweak gauge bosons are denoted as *prompt* or *signal* leptons, while leptons produced in the decay of heavy hadrons are denoted as *non-prompt* or *background* leptons. The identification of prompt and non-prompt leptons is one of the key challenges in the analysis of events with multi-lepton final states. Likewise, this applies to events with high hadronic activity, where jets are more probable to randomly overlap with a genuine prompt lepton. Background processes which contain non-prompt leptons can be reduced using sophisticated lepton identification techniques which are applied in addition to the common selection requirements presented in Section 4.2. For this purpose, a multivariate lepton identification method referred to as LEPTONMVA is employed. Several analyses within CMS have used the LEPTONMVA approach. For the analysis presented in this thesis, the version developed in context of the search for single top quark production in association with a Z boson [175, 196] is applied and discussed in the following.

A *boosted decision tree* (BDT) [197, 198] is employed to create a single discriminant based on a set of kinematic observables. These observables are chosen to reflect the kinematic differences between prompt and non-prompt leptons. The BDT classifies the lepton candidates according to various properties into signal-like leptons and background-like leptons. The BDT output distribution, i.e., the final discriminant is defined on the set $[-1, 1]$. More positive values of the output are assigned to signal-like leptons, while background-like leptons receive more negative values. The training procedure was performed separately for electrons and muons. The different detector geometries in 2016 and 2017 are taken into account to address the improved vertex reconstruction resolution due to the additional fourth pixel tracking layer. Signal leptons were extracted from simulated events where a Z boson is produced in association with a single top quark, or a top quark-antiquark pair. Background leptons are obtained from inclusive top quark-antiquark pair production.

The presence of a jet close to the lepton is an important feature used to distinguish prompt leptons from non-prompt leptons. For these jets a maximum distance between the jet and the lepton of $\Delta R < 0.4$ and a lower threshold on the transverse jet momentum of $p_T > 15 \text{ GeV}$ are required. The jet p_T is computed by applying *lepton-aware* JES corrections. In this case the JES correction is applied to the difference of the uncorrected jet and the lepton, and subsequently adding the lepton momentum to the corrected jet again. In case no matching jet is found, the jet related observables are constructed using the PF particle candidates inside a cone $\Delta R = 0.4$. Overall 10 different input observables are used in the LEPTONMVA BDT training:

- $I_{\text{rel,charged}}^{\text{mini}}$: The mini-isolation of the lepton with respect to charged particles.
- $I_{\text{rel,neutral}}^{\text{mini}}$: The mini-isolation of the lepton with respect to neutral particles.
- p_T^{ratio} : The ratio of the lepton p_T to the p_T of the nearest jet.

- $\text{DeepCSV}^{\text{closest}}$: The discriminant value of the DeepCSV algorithm of the nearest jet.
- $N_{\text{tracks}}^{\text{jet}}$: The number of tracks of charged particles within the nearest jet.
- $p_{\text{T}}^{\text{rel}}$: The component of the lepton momentum in direction transverse to the nearest jet.
- d_{xy} and d_{z} : The transverse and longitudinal distance of closest approach between the track and the PV.
- $\text{SIP}_{3\text{D}}$: The three dimensional impact parameter of the lepton track with respect to the PV, divided by its uncertainty, corresponding to its significance.
- MVA_{ID} : The discriminant value of the multivariate electron identification used to separate electrons from genuine jets (*electrons only*).
- Comp_{seg} : The compatibility of track segments in the muon system with the pattern expected for a minimum ionizing particle (*muons only*).

To probe if a deep neural network could improve the capability to distinguish prompt and non-prompt leptons using the same set of input observables, a deep feed-forward densely-connected neural network was trained using KERAS [199] with TENSORFLOW [200] as backend. This leads to a similar result in terms of signal selection efficiency and background rejection compared to the presented BDT-based approach. Since no further improvement could be achieved, the BDT-based algorithm was preferred.

Three different working points, which correspond to a certain identification efficiency and purity, are tested to obtain the best performance for $t\bar{t}H$ signal events in the three and four lepton final states. The medium working point provides the best balance between prompt lepton selection efficiency and non-prompt lepton rejection. It corresponds to a LEPTONMVA BDT output of 0.4 or larger values. The medium working point is applied to all electrons and muons in addition to the selection requirements presented in Section 4.2. The corresponding selection efficiency for muons (electrons) with $p_{\text{T}} > 20 \text{ GeV}$ is 95% (92%) as indicated in Fig. 5.26. The impact on the signal event selection and the rejection of background processes containing non-prompt leptons are further discussed in Section 5.8.

The selection efficiencies in data and simulation are measured independently for 2016 and 2017. Since no dedicated training is available for the 2018 data-taking period, the version of the training from 2017 and the corresponding selection efficiencies are used also for 2018. The modeling of the input observables between 2017 and 2018 is in most cases very similar, some slight differences in jet related observables are found. Therefore, an additional systematic uncertainty is introduced to account for any potential bias due to differences in the input observables between 2017 and 2018. The overall impact of this systematic uncertainty on the analysis performance is very small as the statistical uncertainty dominates as discussed in Section 5.11.2.

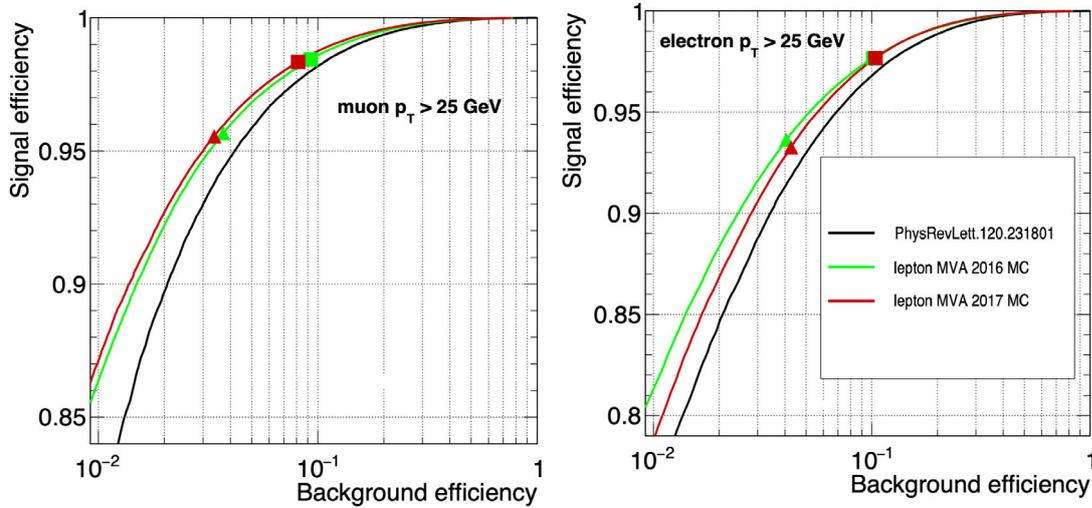


Figure 5.26: LEPTONMVA selection efficiencies as a function of the background efficiency for muons (left), and electrons (right) with $p_T > 25 \text{ GeV}$. The efficiency curves are presented for 2016 (green) and 2017 (red). A comparison to an earlier version used in Ref. [30] is shown in black. Two different working points are presented and indicated by a square and a triangle. The chosen medium working point corresponds to the triangle and a BDT output larger than 0.4. Modified from Ref. [196].

The LEPTONMVA method is validated in a dileptonic $t\bar{t}$ control region using events with one electron and one muon with opposite charge and invariant mass between $110 < m_{e\mu} < 150 \text{ GeV}$. These events mostly feature prompt leptons as expected from $t\bar{t}H$ signal events. The LEPTONMVA output distribution is presented for electrons and muons and all three data-taking periods in Fig. 5.27. Simulated events can reproduce the LEPTONMVA output distribution for values larger than 0.4, which corresponds to the chosen working point. This region is particularly dominated by prompt leptons. An increasing disagreement is observed towards small values of the LEPTONMVA output. This is expected, since non-prompt leptons are known to be less precisely modeled in simulated events than prompt leptons. The agreement between data and simulation is similar in 2016 and 2018, while in 2017 larger discrepancies are observed. These originate from a mismodeling of some jet-related observables in 2017. The discrepancy, however, does not affect the final result because the background contribution containing non-prompt leptons is estimated using a dimuon mass sideband fit to data.

The LEPTONMVA is one of the key features of the $t\bar{t}H$ analysis. The chosen medium working point allows for a high selection efficiency of signal muons in $t\bar{t}H$ events, while the background from non-prompt leptons is substantially reduced. Further information on the signal yields is provided in Section 5.8.1.

5.6 Multivariate lepton identification

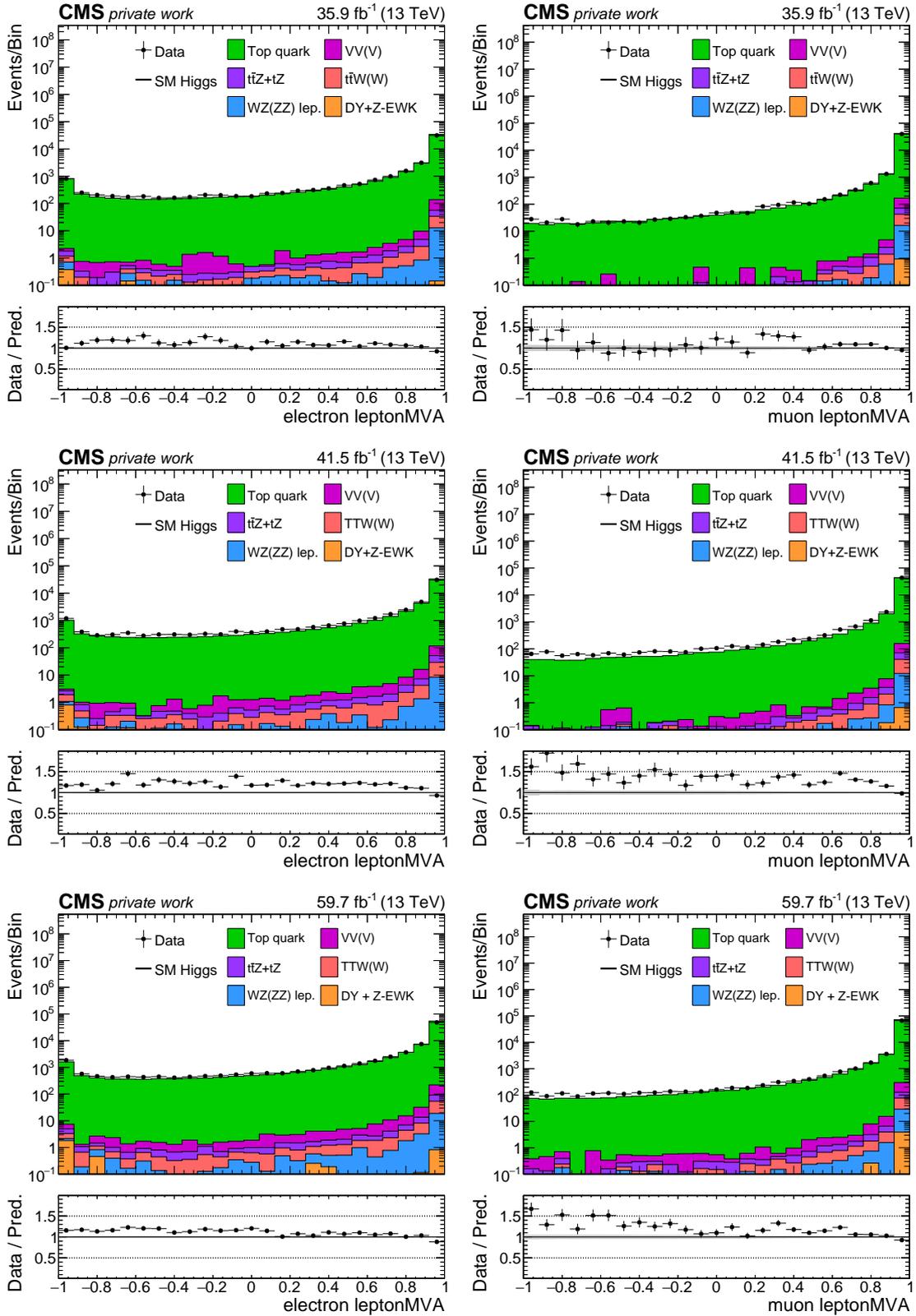


Figure 5.27: The BDT output distribution of the LEPTONMVA is presented for electrons (left) and muons (right) independently for all three data-periods. Events are selected in a $t\bar{t}$ control region with exactly one oppositely charged $e\mu$ pair.

5.7 Resolved hadronic top quark tagger

The top quark is the only quark which decays faster than it hadronizes. A dedicated algorithm is used to identify the hadronic decay of a top quark to three resolved jets. One of these jets originates from a b quark and the remaining two jets are initiated by the W boson decay to two quarks. The *resolved hadronic top quark tagger* (RH TT) combines a constrained kinematic fit [201] with a particle-level BDT. Different kinematic jet properties are used to distinguish genuine top quarks from random jet permutations. The purpose of the kinematic fit is to probe the kinematic compatibility of a jet-triplet with the hypothesis of a hadronic top quark decay, while the BDT allows to exploit additional intrinsic jet properties, for example, the b-tag probability. Although the RH TT follows a similar strategy as reported in Ref. [202], the algorithm is developed independently and optimized to identify low to moderately boosted hadronic top quark decays in $t\bar{t}H$ signal events. Among all tested jet-triplets in the event, the one with the highest RH TT score is selected as the top quark candidate.

5.7.1 Kinematic top quark fit

The compatibility of the measured final state objects with a proposed kinematic hypothesis can be tested using a general least squares fitting technique. In this thesis, a constrained kinematic fit is performed to identify hadronically decaying top quarks. For this purpose, the three-momenta (p_T, η, ϕ) of three jets are varied within their uncertainties. The best estimate of the fitted parameters is obtained by minimizing a $\chi^2(\vec{p}_T^{\text{fit}})$ function, which describes deviations between the measured parameters x_{reco} , and the fitted parameters x_{fit} normalized to their resolution:

$$\chi^2(\vec{p}_T^{\text{fit}}) = \sum_{j \in \{\text{jets}\}} \frac{(p_T^{\text{reco}} - p_T^{\text{fit}})^2}{\sigma_{p_T}^2} + \frac{(\eta^{\text{reco}} - \eta^{\text{fit}})^2}{\sigma_\eta^2} + \frac{(\phi^{\text{reco}} - \phi^{\text{fit}})^2}{\sigma_\phi^2}. \quad (5.10)$$

The assumed resolution for p_T , η , and ϕ are the same as used in the CMS top quark mass measurement [203, 204]. The noise term of the resolution used during Run 1 has been inflated by quadratically adding 3 GeV to account for increased pileup in Run 2. To select events consistent with the kinematic topology expected from a hadronic decay of a top quark, two Gaussian constraints have to be fulfilled when the minimum of the $\chi^2(\vec{p}_T^{\text{fit}})$ function is determined. First, the fitted mass of the jet-triplet is constrained to the top quark mass $m_t = 172.8 \pm 1.4$ GeV [54]. Second, the fitted dijet mass of jets originating from the decay of the W boson is constrained to the W boson mass $m_W = 80.4 \pm 2.1$ GeV [54]. Note that for both the top quark and the W boson mass, the central values correspond to recent measurements, whereas the uncertainties are taken from the theoretical values of the natural width.

A general and very conventional approach to solve a minimization problem of a multidimensional function $\chi^2(\vec{p}_T^{\text{fit}})$, which satisfies m constraints $f_k(\vec{p}_T^{\text{fit}})$, is the method of Lagrange Multipliers. The best parameter estimates are obtained by the fitted jet three-momenta and Lagrange multipliers λ_k that least modify the reconstructed jets, i.e., minimize the Lagrange function

$$L(\vec{p}^{\text{fit}}, \lambda_k) = \chi^2(\vec{p}^{\text{fit}}) + 2 \sum_{k=1}^m \lambda_k f_k(\vec{p}^{\text{fit}}). \quad (5.11)$$

All possible parton-jet assignments of the six leading jets in p_T with a mass in the range between $100 < m_{jjj} < 300 \text{ GeV}$ are tested. The jet which is assigned to the b quark candidate is required to pass the loose b-tagging working point of the DeepCSV algorithm. The permutations of the two jets associated with the W boson decay are equivalent either way. For each tested jet permutation the minimized χ^2 value normalized to the number of degrees of freedom is used as an input to the particle-level BDT described in Section 5.7.2.

To validate the correctness and accuracy of the kinematic fit, the fitted parameter values are compared with the reconstructed ones using a pull distribution [205]. The pull for every fitted parameter is defined as the ratio between the shift in the parameter and the variance of the shift:

$$\text{Pull}(x) = \frac{x_{\text{reco}} - x_{\text{fit}}}{\sqrt{\sigma_{\text{reco}}^2 - \sigma_{\text{fit}}^2}}. \quad (5.12)$$

If the assumed resolutions in the fit are modeled correctly and the constraints can be fulfilled by the jet-triplet, the pull is expected to be distributed as a standard Gaussian with mean zero and unit width for the correct hypothesis. For this consistency test only correct jet assignments are considered where all jets are matched unambiguously to the right partons. Figure 5.28 shows the mean and width of the pull distributions for all fitted parameters in $t\bar{t}H$ signal events summarized for all three data-taking periods. The pull distributions of all fitted parameters are sufficiently close to a Gaussian with mean of zero and width of one. Small parameter shifts are observed in the p_T of the b-tagged jet as well as the p_T of the subleading W boson jet within 10% of their experimental resolution. The widths of the pull distributions tend to be larger than one. This indicates some inaccuracies in the matching procedure, or slightly underestimated values of the assumed resolutions by up to 5%. However, none of these differences is expected to influence the results, since the JES and JER are already corrected in simulated events to match the observed distribution in data. Thus, any resolution scaling alters the χ^2 distribution for data and simulation in the same way.

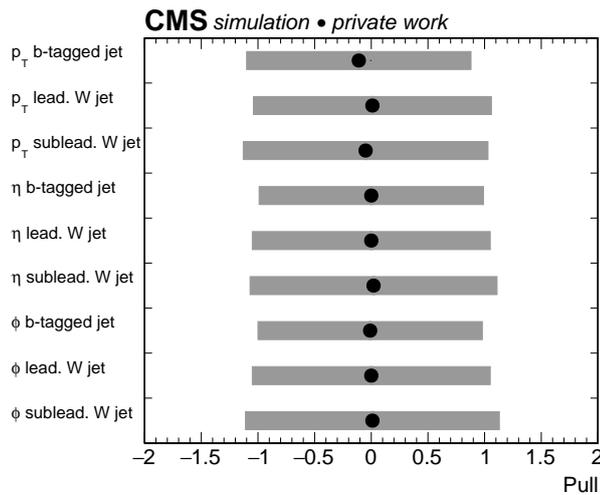


Figure 5.28: Pull distributions for all fitted parameters used in a constrained kinematic fit to reconstruct hadronically decaying top quarks in $t\bar{t}H$ signal events. The mean (solid marker) and width (gray band) of the pull distributions are extracted using Gaussian fits.

5.7.2 BDT-based top quark identification

The RHTT is a particle-level BDT discriminator, which includes the information of the constrained kinematic top quark fit. The jet-triplets used in the training are obtained from $t\bar{t}H$ signal samples with generated Higgs boson masses m_H of 120 and 130 GeV. The shifted Higgs boson mass values compared to the nominal Higgs boson mass of 125 GeV is found to have a negligible impact on the kinematic distributions of the top quarks. To ensure a correct parton-jet assignment in signal events for training, the reconstructed jets are matched to the top quark decay products using event generator information. The background, however, consists of random jet combinations in dileptonic $t\bar{t}$ and Drell-Yan events.

The hyperparameters of the BDT training are optimized in an iterative procedure where different hyperparameters are tested and the set which provides the best performance is chosen. The Gradient Boost method [197, 198] is used to train a BDT with 1000 trees and a maximum depth of four. A shrinkage parameter, which corresponds to the learning rate for the Gradient Boost algorithm, of 0.05 is chosen. The simulated event samples used in the training are approximately half of the available sample size, the other half is used for testing purposes. The Drell-Yan background simulation contains negative weighted events. Three alternative strategies on how to incorporate these events in the training are tested. First, both positive and negative weighted events are used in the training. Second, events with a negative weight are discarded. Third, events are weighted with the absolute value of the per-event weight. A similar performance on the testing events is found for all three approaches, and the first option is used in the final training. The training is based on a combined set of simulated events from all three data-taking periods.

A loose preselection is employed to optimize the discrimination between signal and background jet triplets in the relevant phase-space region. Events are required to have at least one oppositely charged pair of muons with a mass in the range $110 < m_{\mu\mu} < 150$ GeV as well as at least three jets with $p_T > 25$ GeV and $|\eta| < 4.7$ where at least one passes the medium or at least two the loose working point of the DeppCSV b-tagging algorithm. Additionally, only top quark candidates with a mass in the range between $100 < m_{jjj} < 300$ GeV, and $\chi^2/\text{ndf} < 40$ are considered in the training. This preselection primarily affects background jet-triplets, while the selection efficiency for genuine top quarks in signal events is close to 100%.

To distinguish signal from background jet-triplets, 17 observables are included as input to the BDT classifier. The most important observable is the χ^2/ndf , which describes the compatibility of a jet-triplet with the kinematic topology expected from a hadronic top quark decay. The signal and background distributions of the χ^2/ndf , as well as of the reconstructed top quark mass m_{top} and of the W boson mass m_W are shown in Fig. 5.29. Signal events dominate over the background for small χ^2/ndf as expected. The output of the DeepCSV b-tagging algorithm and the quark-gluon likelihood of each jet help to correctly assign jets to the b quark and the light-flavour quarks from the W boson decay. Furthermore, they allow a reduction in background from gluon-induced jets. These single jet properties are shown in Fig. 5.30 for the b jet and for both jets from the W boson decay. The modified transverse momentum of each jet

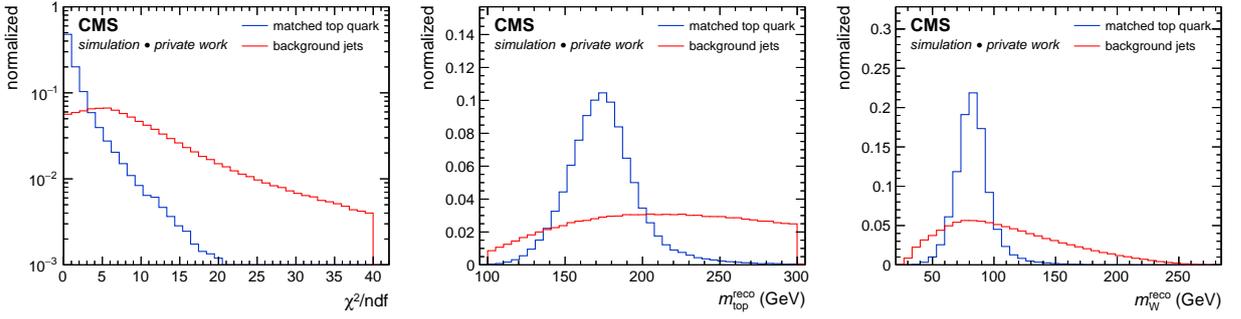


Figure 5.29: Distributions of input observables employed by the BDT classifier of the RHTT. The χ^2/ndf (left), reconstructed top quark mass m_t (middle), and reconstructed W boson mass m_W (right) are presented. The generator matched signal (blue) and background (red) contributions are scaled to have the same area.

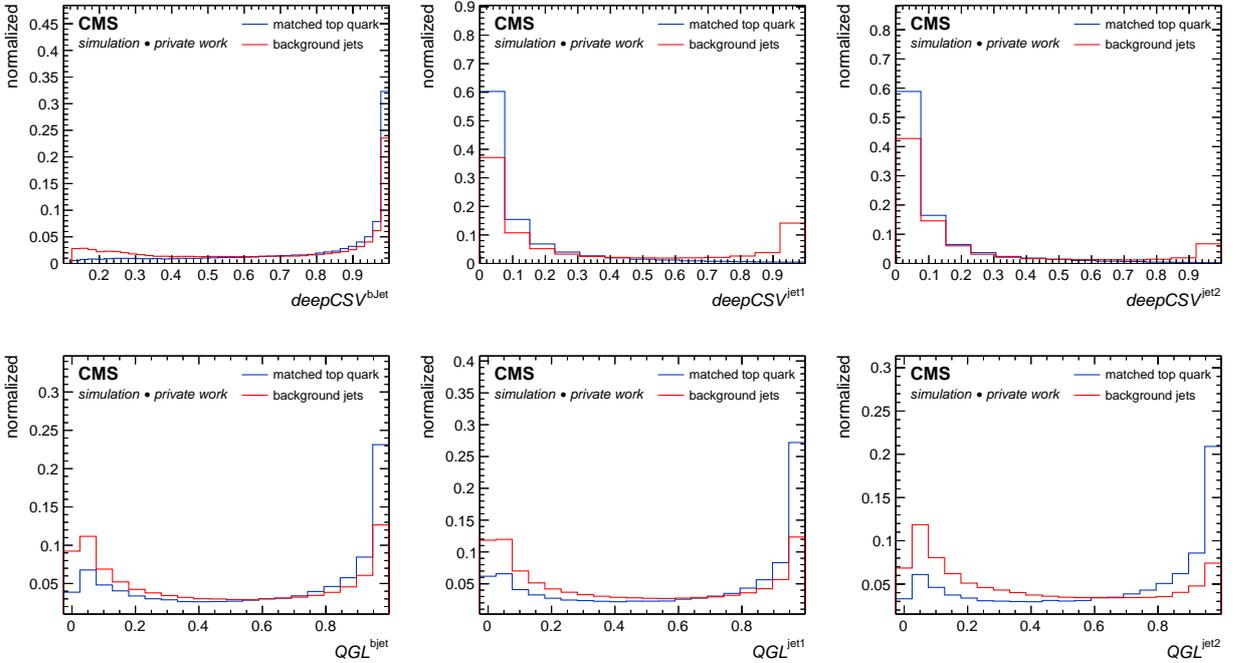


Figure 5.30: Distributions of input observables employed by the BDT classifier of the RHTT. The DeepCSV score (top), and quark-gluon likelihood (bottom) are shown for the jet assigned to the b quark candidate (left) and both jets assigned to the decay of the W boson (center, right). The generator matched signal (blue) and background (red) contributions are scaled to have the same area.

as well as different angular distributions exploit the correlation between jets. A full list of all input observables and their importance ranking is provided in Table 5.2. The ranking is based on two quantities, the separation and the relative importance. Both are commonly used to describe the impact of input observables on the BDT performance. The relative importance

Table 5.2: Description of the input observables employed by the BDT classifier of the RHTT. The observables are subdivided in groups and ranked according to their importance during the training.

Definition	Observables	Separation [%]	Importance [%]
normalized χ^2	χ^2/ndf	43.8	8.2
Reconstructed W boson mass	m_W^{reco}	34.4	8.0
Reconstructed top quark mass	$m_{\text{top}}^{\text{reco}}$	31.9	7.9
Transverse momentum of top quark candidate	p_T^{top}	22.0	6.2
$\Delta\eta$ difference between b quark and W boson	η_{bw}	12.4	4.6
$\Delta\phi$ difference between b quark and W boson	ϕ_{bw}	7.0	5.1
$\Delta\eta$ difference between jets from W boson	η_{jj}	19.4	5.1
$\Delta\phi$ difference between jets from W boson	ϕ_{jj}	9.1	5.4
Transverse momentum of b jet	p_T^{bjet}	11.3	5.4
Transverse momentum of lead. jet from W boson	p_T^{jet1}	12.1	4.9
Transverse momentum of sublead. jet from W boson	p_T^{jet2}	21.5	5.7
DeepCSV score of b jet	DeepCSV ^{bjet}	4.8	5.1
DeepCSV score of lead. jet from W boson	DeepCSV ^{jet1}	15.0	5.8
DeepCSV score of sublead. jet from W boson	DeepCSV ^{jet2}	7.7	5.0
Quark-gluon likelihood of b jet	QGL^{bjet}	4.3	6.5
Quark-gluon likelihood of lead. jet from W boson	QGL^{jet1}	6.9	5.6
Quark-gluon likelihood of sublead. jet from W boson	QGL^{jet2}	7.7	5.4

reflects how often an observable is used during the BDT training. The separation is a measure of the overlap between two normalized distributions. Large values indicate a strong discrimination power of an observable [197, 198]. Figure 5.31 (left) presents the BDT output distributions of the training and testing samples for signal and background jet-triplets. The BDT training should be designed to avoid the selection of specific phase-space regions corresponding to statistical fluctuations in the training sample. This effect is referred to as *overtraining*, and its impact is examined by comparing the BDT output distribution of the training sample and the testing sample. In particular the high-BDT score region is modeled very similarly in the training and the testing sample. The signal contribution in the low-BDT score region with values less than -0.2 is affected by considerable statistical fluctuations. Overall, no significant overtraining is observed. Fig. 5.31 (right) shows the *Receiver Operating Characteristic* (ROC) diagram of the RHTT, which describes the signal efficiency versus the background rejection for a given working point of the RHTT algorithm. Three example working points are selected and the obtained signal efficiencies between the training and testing samples are compared in Table 5.3. The absolute difference between the signal efficiencies is in the order of a few permille.

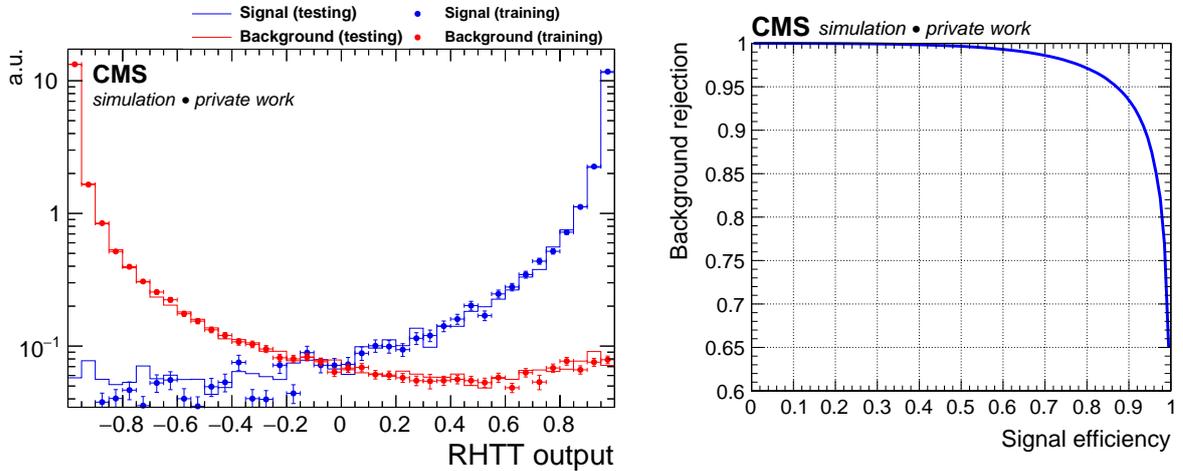


Figure 5.31: BDT output distribution (left) and ROC diagram (right) of the RHTT algorithm. The BDT output distributions of the signal (blue) and background (red) are shown for the training sample (marker) and the test sample (solid line). The ROC diagram presents the signal efficiency for genuine hadronic top quark decays versus the background rejection of the RHTT algorithm.

Table 5.3: Comparison of the signal efficiency for three example working points of the RHTT algorithm between the testing sample and the training sample. The signal corresponds to reconstructed hadronic top quark decays matched to partons on event generator level.

Background rejection	Signal efficiency	
	testing sample	training sample
99%	64.9%	65.3%
90%	93.9%	93.8%
70%	99.1%	99.1%

$t\bar{t}Z$ control region Since the RHTT algorithm is tailored to identify $t\bar{t}H$ signal events where top quarks decay hadronically, a control region containing genuine hadronic top quark decays is defined to validate the signal efficiency of the RHTT in data. Therefore events with a pair of oppositely charged muons in the $m_{\mu\mu}$ range around the Z boson mass between 70–110 GeV are selected. These events are further required to contain one additional lepton, and at least three jets, where at least one of them needs to pass the medium working point or two the loose working point of the DeepCSV b-tagging algorithm. The additional lepton is required to increase the purity of $t\bar{t}Z$ events and suppress the background contribution from $t\bar{t}$ and Drell-Yan events. Thus, the $t\bar{t}Z$ control region particularly aims to select $t\bar{t}Z$ events where one of the top quarks decays hadronically to three resolved jets, while the other one decays to a b quark, a charged lepton and a neutrino. Among all jet-triplets in the event, the one with the highest RHTT score is selected as the top quark candidate. Figure 5.32 presents the RHTT score of the selected top quark candidate in the $t\bar{t}Z$ enriched control region. Although the statistical precision of data events in this control region is limited, simulated events reproduce the distribution observed in data within the statistical uncertainties. The signal efficiency in data is thus confirmed.

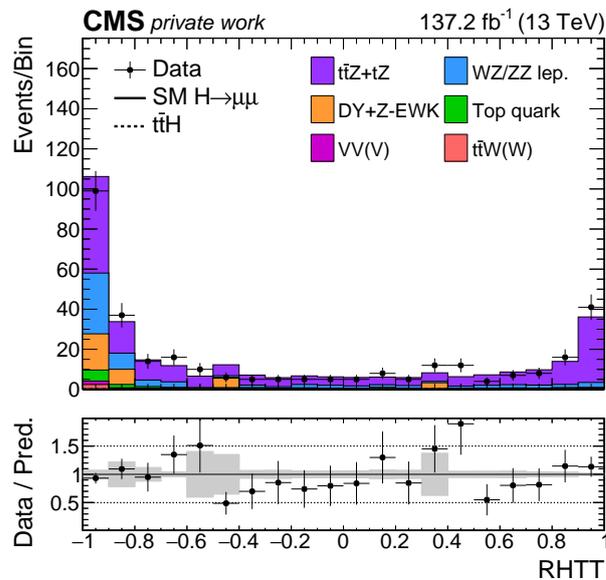


Figure 5.32: Validation of the RHTT output distribution in data. The RHTT output of the top quark candidate is shown in a $t\bar{t}Z$ enriched control region which provides genuine hadronic top quark decays with a similar kinematic topology as expected from $t\bar{t}H$ signal events.

Dimuon mass sideband region The events are selected if exactly one oppositely charged muon pair with dimuon mass in the range between 110 – 120 GeV or 130 – 150 GeV is found. Moreover, these events are required to contain at least three jets where at least one of them needs to pass the medium working point or at least two the loose working point of the DeepCSV b-tagging algorithm. Figure 5.33 shows the RHTT score in the $m_{\mu\mu}$ sideband region dominated by dileptonic $t\bar{t}$ events due to the required presence of b-tagged jets. Simulated events can precisely model the normalization and the shape of the distribution observed in data. Discrepancies between data and simulation are typically smaller than 5%. In contrast to the $t\bar{t}Z$ control region, events in the dimuon mass sidebands are dominated by combinatorial background of jet-triplets from dileptonic $t\bar{t}$ and Drell-Yan events.

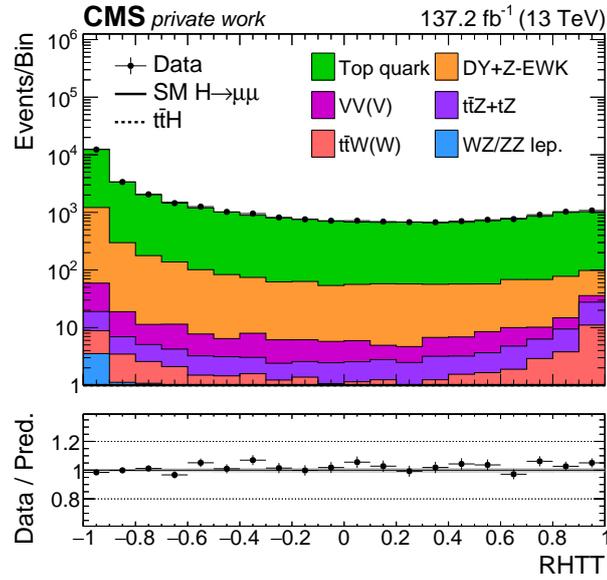


Figure 5.33: Validation of the RHTT output distribution in data. The RHTT output of the top quark candidate is presented in the dimuon mass sideband region. Events from the signal region are excluded. This control region is dominated by combinatorial background from dileptonic $t\bar{t}$ and Drell-Yan events.

A top quark candidate is defined by the jet-triplet with the highest RHTT score in the event. This top quark candidate is used in a multivariate event classifier which is trained to distinguish $t\bar{t}H$ signal events where at least one top quark decays hadronically to three resolved jets from the SM backgrounds as described in Section 5.9.1.

5.8 $t\bar{t}H$ event selection

The selection of events is based on the identified final state particles. Muons, electrons, and jets are required to pass the selection criteria detailed in Sections 4.3.2 to 4.3.6. For convenience, a few main features of the selected particles are briefly summarized:

- Muons and electrons from the decay of a Higgs boson or a heavy top quark tend to have large p_T . Therefore, muons (electrons) are selected with $p_T > 20$ GeV inside the detector acceptance of $|\eta| < 2.4$ (2.5).
- Baseline muon and electron identification criteria are used along with the LEPTONMVA medium working point to suppress non-prompt lepton background.
- A loose lepton isolation $I_{\text{rel}}^{\text{mini}} < 0.4$ ensures a high selection efficiency for leptons in $t\bar{t}H$ signal events with high hadronic activity.
- Jets are selected with $p_T > 25$ GeV and $|\eta| < 4.7$. The DeepCSV algorithm is employed to identify jets originating from b hadrons inside the tracker acceptance of $|\eta| < 2.5$.

Events are recorded by single muon triggers as explained in Section 5.3. The dimuon baseline selection comprises events with two oppositely charged muons. The resulting data shown in Fig. 5.34 is strongly dominated by events from the Drell-Yan process as seen by the dimuon mass distribution of the Z boson mass resonance and the tail of the mass spectrum with the blinded Higgs boson signal region. The signal region in data is blinded to avoid a bias in the optimization of the event selection.

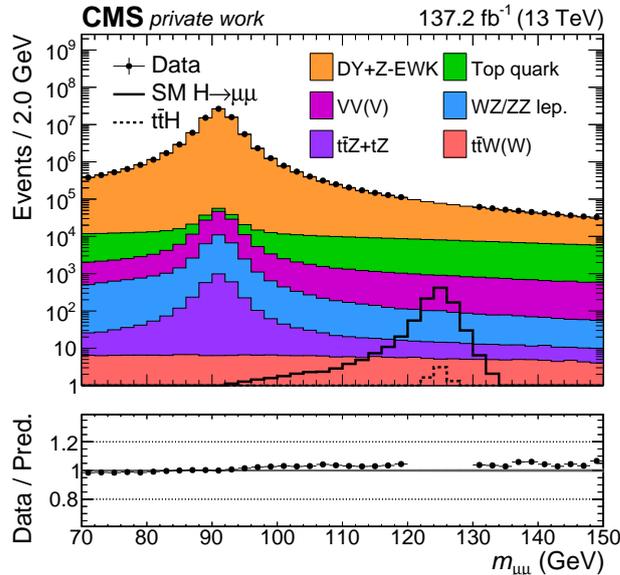


Figure 5.34: Dimuon mass distribution of events passing the dimuon baseline selection with the full Run 2 dataset. The SM expectation for the $H \rightarrow \mu^+\mu^-$ signal is represented by the solid black line. The $t\bar{t}H$ signal contribution is indicated by the dashed line.

Apart from the selected muon pair, the decay products of the top quarks are taken into account to define a dedicated $t\bar{t}H$ production category. Top quarks decay predominantly into a b quark and a W boson. Therefore, events with at least one medium or two loose b-tagged jets are selected. This provides a $t\bar{t}H$ baseline event selection, which guarantees mutual exclusivity between the $t\bar{t}H$ category and other Higgs boson production categories used for a combined $H \rightarrow \mu^+\mu^-$ result. Figure 5.35 shows the corresponding dimuon mass distribution in the dimuon mass sideband region. After the selection of events that contain b-tagged jets,

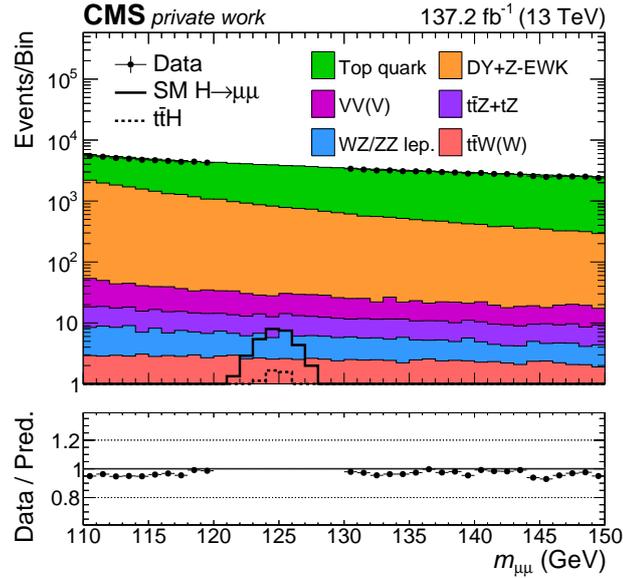


Figure 5.35: Dimuon mass distribution of events selected using the baseline $t\bar{t}H$ event selection with the full Run 2 dataset. The expected $H \rightarrow \mu^+\mu^-$ signal is represented by the solid black line. The $t\bar{t}H$ signal contribution is indicated by the dashed line.

the background composition changes. Events from dileptonic $t\bar{t}$ decays become the dominant background, while the Drell-Yan background contribution is reduced. The purity of $t\bar{t}H$ signal events after the $t\bar{t}H$ baseline selection yields 22%, and ggH events still constitute a large fraction of the signal. The shape of the dimuon mass distribution in data is well modeled by simulated events. The normalization between simulation and data differs by about 3%.

To further increase the $t\bar{t}H$ signal component, events are classified according to the decay products of the W boson produced in the top quark decay. The W boson decays either leptonically to a lepton and a neutrino or hadronically to two quarks. The corresponding subcategories are denoted as $t\bar{t}H$ leptonic, and $t\bar{t}H$ hadronic, respectively. They are described in detail in Sections 5.8.1 and 5.8.2. The $t\bar{t}H$ leptonic and $t\bar{t}H$ hadronic subcategories are designed to provide a highly efficient and pure selection of events based on the expected kinematic properties of the $t\bar{t}H$ signal process. The background can be reduced by exploiting a number of discriminating observables. To achieve an optimum analysis sensitivity, multivariate techniques are employed. The corresponding multivariate optimization of the analysis is presented in Section 5.9.

5.8.1 $t\bar{t}H$ leptonic selection

Events with one or two charged leptons in addition to the muon pair are considered for the $t\bar{t}H$ leptonic channel. This includes events with semi- and dileptonic decays of the top quark-antiquark pair. Additional criteria are applied which more purely select $t\bar{t}H$ signal events. At least two jets are required with a leading jet $p_T > 35$ GeV. One of the key features of the $t\bar{t}H$ leptonic event selection is the muon pair finding. In events with more than two muons, at least one opposite charged muon pair is required with an invariant mass between 110 and 150 GeV. If two muon pairs fall into this mass window, the pair with the higher dimuon transverse momentum is chosen as the Higgs candidate. This strategy allows in 94% of all selected events to assign the correct muon pair to the Higgs boson decay. Backgrounds from quarkonium decays and combinatorial background are reduced by a veto on events where the invariant mass of any oppositely charged same flavour lepton pair is close to the mass of the Z boson with $81 \text{ GeV} < m_Z < 101 \text{ GeV}$ or below a threshold of 12 GeV. All selection criteria are summarized in Table 5.4. The impact of the different requirements on the signal composition and the signal event yield are given in Table 5.5. Figure 5.36 presents the dimuon mass distribution for all events in the $t\bar{t}H$ leptonic event category. After the $t\bar{t}H$ leptonic event selection the signal

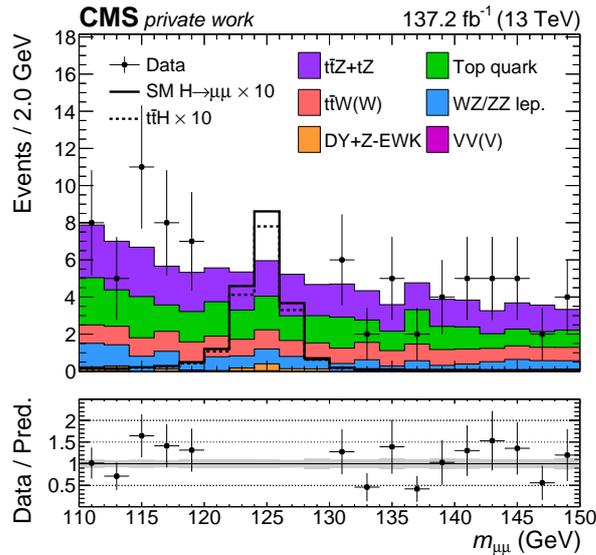


Figure 5.36: Dimuon mass distribution in events selected by the $t\bar{t}H$ leptonic event category using the full Run 2 dataset. The expected $H \rightarrow \mu^+\mu^-$ signal is represented by the solid black line and scaled by a factor of 10. The $t\bar{t}H$ signal component is indicated by the dashed line.

purity for events produced by the $t\bar{t}H$ production mechanism increases to 89%. The dominant background processes are $t\bar{t}Z$ and dileptonic $t\bar{t}$ with additional non-prompt leptons, which can originate from the decay of heavy b hadrons. The $t\bar{t}$ + non-prompt lepton background still contributes a considerable amount of the total background although the LEPTONMVA significantly reduces this contribution by 96% compared to a selection without LEPTONMVA. The measured data yields as well as the predicted event yields for all background processes are reported in Table 5.6. These numbers are given for events in the dimuon mass sideband region

where events from the signal region are excluded. The overall predicted event yield is compatible with the event yield observed in data below 120 and above 130 GeV within 1σ of the statistical uncertainty.

Table 5.4: Summary of the event selection criteria for the $t\bar{t}H$ leptonic category. The selection aims at the semileptonic and dileptonic final states of the top quarks.

Event selection criteria	$t\bar{t}H$ leptonic	
	semileptonic	dileptonic
Number of leptons	3	4
Lepton charge requirements	$\sum q_i = \pm 1$	$\sum q_i = 0$
Number of b-tagged jets	$N_{\text{b-tagged}} > 0$ (medium) or > 1 (loose)	
Jet multiplicity	≥ 2	
Leading jet p_T	≥ 35 GeV	
Z mass veto	$ m_{\ell\bar{\ell}} - m_Z > 10$ GeV	
Low mass resonance veto	$m_{\ell\bar{\ell}} > 12$ GeV	

Table 5.5: Comparison of the signal composition using the $t\bar{t}H$ baseline selection to define an exclusive $t\bar{t}H$ production category, and the final $t\bar{t}H$ leptonic event selection. The event counts are based on simulated signal events in the mass range between $120 < m_{\mu\mu} < 130$ GeV.

Event selection	signal event yield					
	ggH	VBF	VH	$t\bar{t}H$	tH	Σ
$t\bar{t}H$ baseline (trigger + $\mu^+\mu^-$ + b-tag)	17.3	3.0	3.8	7.0	0.9	32.0
$t\bar{t}H$ leptonic (baseline + add. lepton + jets + $m_{\ell\ell}$ vetos)	–	–	0.1	1.7	0.1	1.9

Table 5.6: Expected and observed event yields of the $t\bar{t}H$ leptonic event category in the dimuon mass sideband region. Events from the signal region are excluded. The event yields are shown for all three data-taking periods and their combination. The statistical uncertainty is reported for data and the predicted sum of simulated background events.

Background contribution	$t\bar{t}H$ leptonic			
	2016	2017	2018	Run 2
$t\bar{t}Z + tZ$	7.0	7.2	12.4	26.6
$t\bar{t}W(W)$	3.4	3.7	6.1	13.2
WZ/ZZ lep.	2.2	2.8	4.6	9.6
Top quark	6.8	5.1	9.6	21.4
DY	0	0.1	0.2	0.3
VV(V)	0.1	0.1	0.3	0.5
Σ prediction	19.4 ± 0.7	19.0 ± 0.8	33.1 ± 1.3	71.5 ± 1.7
observed data	14 ± 4	29 ± 5	36 ± 6	79 ± 9

5.8.2 $t\bar{t}H$ hadronic selection

Events with exactly two muons and without additional leptons are assigned to the $t\bar{t}H$ hadronic channel. This event category is tailored to events where both top quarks decay hadronically. Since the reconstruction and selection of all of the six jets from the two hadronic top quark decays would drastically reduce the signal selection efficiency, the event selection is tuned to identify at least one high-quality top quark candidate per event. To ensure the presence of at least one potential top quark candidate, events are required to contain at least three jets with a leading jet $p_T > 50\text{ GeV}$. Furthermore, one of the jet-triplets needs to be reconstructed with an invariant mass in the range between $100 < m_{jjj} < 300\text{ GeV}$, and the kinematic top quark fit of this jet triplet is required to converge. The convergence efficiency of the kinematic fit for genuine top quarks is nearly 100%. An overview of the applied selection criteria is given in Table 5.7. The impact of the different requirements on the signal composition and the signal event yield are reported in Table 5.8. Figure 5.37 shows the dimuon mass distribution for all events in the sideband region of the $t\bar{t}H$ hadronic event category.

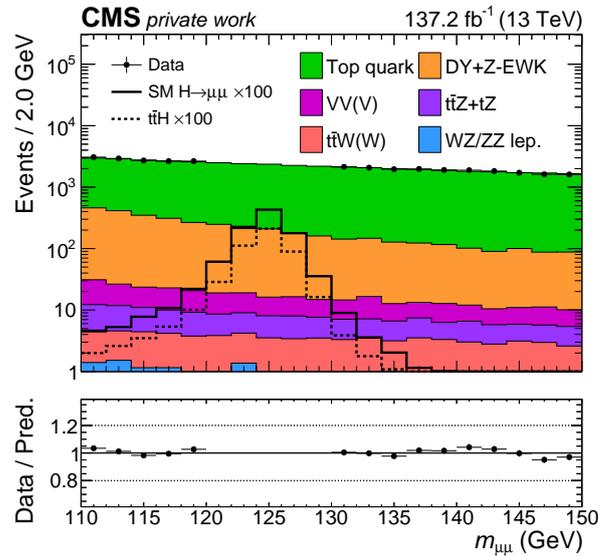


Figure 5.37: Dimuon mass distribution of events selected by the $t\bar{t}H$ leptonic event category using the full Run 2 dataset. The expected $H \rightarrow \mu^+\mu^-$ signal is represented by the solid black line and scaled by a factor of 100. The $t\bar{t}H$ signal component is indicated by the dashed line.

A signal purity for events produced by the $t\bar{t}H$ production mechanism of 49% is achieved in the $t\bar{t}H$ hadronic event category. The dominant background processes are dileptonic $t\bar{t}$ and Drell-Yan events with additional jets from ISR or FSR. The event yields in data as well as the predicted yields for all background processes in the dimuon mass sideband region excluding the signal region are summarized in Table 5.9. The overall predicted event yield in simulated events of the full Run 2 data-taking period is compatible with the event yield observed in data within 1σ of the statistical uncertainty.

Table 5.7: Summary of the event selection criteria for the $t\bar{t}H$ hadronic categories. The selection aims at the fully hadronic final states of the top quarks.

Event Selection	$t\bar{t}H$ hadronic
Number of b-tagged jets	$N_{b\text{-tagged}} > 0$ (medium) or > 1 (loose)
Number of muons	2
Muon charge requirement	$\sum q_i = 0$
Jet multiplicity	≥ 3
Leading jet p_T	≥ 50 GeV
Jet-triplet convergence	$100 \text{ GeV} < m_{j\bar{j}j} < 300 \text{ GeV}$

Table 5.8: Comparison of the signal composition using the $t\bar{t}H$ baseline selection to define an exclusive $t\bar{t}H$ production category, and the final $t\bar{t}H$ hadronic event selection. The event counts are based on simulated signal events in the mass range $120 < m_{\mu\mu} < 130$ GeV.

Event selection	signal event yield					
	ggH	VBF	VH	$t\bar{t}H$	tH	Σ
$t\bar{t}H$ baseline (trigger + $\mu^+\mu^-$ + b-tag)	17.3	3.0	3.8	7.0	0.9	32.0
$t\bar{t}H$ hadronic (baseline + lepton veto + jet-triplet)	2.7	0.2	1.2	4.6	0.6	9.3

Table 5.9: Expected and observed event yields of the $t\bar{t}H$ hadronic event category in the dimuon mass sideband region where events from the signal region are excluded. The event yields are shown for all three data-taking periods and their combination. The statistical uncertainty is reported for data and the predicted sum of simulated background events.

Background contribution	$t\bar{t}H$ hadronic			
	2016	2017	2018	Run 2
$t\bar{t}Z + tZ$	16.8	18.5	30.2	65.5
$t\bar{t}W(W)$	10.3	11.2	18.5	40.1
WZ/ZZ lep.	3.9	2.3	6.4	12.6
Top quark	7633.4	8177.0	13823.8	29634.2
DY	672.7	657.8	1343.0	2673.5
VV(V)	42.7	28.7	60.0	131.3
Σ prediction	8379.9 ± 23.1	8895.5 ± 26.2	15281.9 ± 43.8	32557.3 ± 56.1
observed data	8125 ± 90	9406 ± 97	15193 ± 123	32724 ± 181

5.9 Multivariate analysis optimization of the $t\bar{t}H$ channel

After the $t\bar{t}H$ event selection and the categorization into the $t\bar{t}H$ leptonic and $t\bar{t}H$ hadronic channels, two BDT-based multivariate discriminants are trained to select specific phase-space regions with enhanced signal-to-background ratio using the TMVA software package [197]. These BDT discriminants exploit several common input observables, as well as some specific observables that target the kinematic properties of either the leptonic or the hadronic top quark decay. Due to many similar features of the BDTs used in the $t\bar{t}H$ leptonic and $t\bar{t}H$ hadronic channels, the input observables and the BDT setup for both channels are described in parallel in Sections 5.9.1 and 5.9.2, while the phase-space selection and the validation of the BDT classifiers in data are discussed individually in Sections 5.9.3 and 5.9.4.

5.9.1 Kinematic input observables to the BDT classifiers

The predicted Higgs boson signal forms a narrow peak on top of a falling background distribution in the dimuon mass spectrum. The invariant mass of the Higgs boson candidate is the most discriminating observable between signal and background events. For this reason, the dimuon mass serves as the final discriminant in the signal and background extraction. To ensure that the background shape of the dimuon mass distribution is not sculpted due to the BDT-based optimization, input observables are chosen which are hardly correlated with the dimuon mass.

The observables sensitive to the characteristic properties of the $t\bar{t}H$ production mechanism which are used as input to the $t\bar{t}H$ leptonic and $t\bar{t}H$ hadronic BDT classifiers are divided into three groups: First, observables describing the production and decay of the Higgs boson; Second, observables that contain information on the hadronic event activity or the presence of neutrinos. Third, observables aiming to identify the top quark decay products. The first two groups represent a set of common observables used as input to both the $t\bar{t}H$ leptonic as well as the $t\bar{t}H$ hadronic BDT.

In $t\bar{t}H$ signal events where a Higgs boson recoils from a heavy top quark, the Higgs boson is boosted in the transverse plane. This leads to a much harder dimuon p_T spectrum in signal events compared to the background. Thus, the dimuon transverse momentum $p_T^{\mu\mu}$, as well as the dimuon rapidity $y_{\mu\mu}$ are included as input observables to the BDT. Muons originating from the decay of a Higgs boson are correlated, which results in characteristic angular distributions. This correlation can be used to distinguish a muon pair from the decay of the Higgs boson from mostly uncorrelated muons in decays of top quark-antiquark pairs. Therefore, the azimuthal angle ϕ_{CS} and the cosine of the polar angle $\cos\theta_{CS}$ between both muons computed in the Collins–Soper rest frame [206] are employed. Additionally, the η of the two muons corresponding to the Higgs boson candidate and their single muon p_T relative to $m_{\mu\mu}$ are used. Figures 5.38 and 5.39 show the corresponding normalized distributions of these observables for signal and background events in the $t\bar{t}H$ hadronic and $t\bar{t}H$ leptonic event categories.

To enhance the fraction of events with a precise dimuon mass resolution to the high-BDT score region, signal events are weighted with the inverse of the per-event dimuon mass resolution which is estimated from the muon track fit as described in Section 5.4.7. The mass resolution is also tested as an input observable, but this does not significantly change the BDT performance. The reason for this might be that the distribution of the per-event dimuon mass resolution is very similar between signal and background events.

To account for the large expected hadronic activity in $t\bar{t}H$ signal events, the number of jets N_{jets} , the maximum DeepCSV value among all selected jets $\text{DeepCSV}_{\text{max}}$, and the scalar (vectorial) transverse momentum sum H_T (H_T^{miss}) of all selected leptons and jets with $|\eta| < 2.5$ are included. The single jet p_T and η of the leading jets are considered as well. The p_T^{miss} and the $\Delta\zeta$ observable [207], which corresponds to the projection of \vec{p}_T^{miss} onto the bisector of the dimuon system in the transverse plane, are also included as input to the BDT. These observables allow to probe the magnitude and direction of neutrinos with respect to the dimuon system in the event. Figures 5.40 and 5.41 present some of the corresponding normalized distributions of these observables for signal and background events in the $t\bar{t}H$ hadronic and $t\bar{t}H$ leptonic event categories.

Additionally, observables related to the characteristic kinematic properties of either the leptonic or the hadronic top quark decay are considered. In the $t\bar{t}H$ leptonic event category, these are represented by the azimuthal separation $\Delta\phi$ between the reconstructed Higgs boson candidate and the additional charged lepton with highest p_T , the invariant mass of the jet with the highest DeepCSV score and the additional lepton $m_{\ell b}$, and the transverse mass of the additional lepton and the missing transverse momentum $m_T^{\ell, \text{MET}}$ [208]. To distinguish the three muon final-state from the two muon and one electron final state, the flavour of the additional lepton is included as input to the BDT. In the $t\bar{t}H$ hadronic event category, the RHTT algorithm is used to identify top quark decays to three resolved jets as explained in Section 5.7. The jet triplet with the highest RHTT score and a mass in the range between $100 < m_{\text{jjj}} < 300 \text{ GeV}$ is chosen as a hadronic top quark candidate, and the corresponding RHTT score is used as input to the BDT. The p_T of the top quark candidate and the balance of the top quark and the Higgs boson candidate

$$B(\text{H, top}) = \frac{|\vec{p}_T^H + \vec{p}_T^{\text{top}}|}{p_T^H + p_T^{\text{top}}} \quad (5.13)$$

are sensitive to the kinematic configuration of the Higgs boson with respect to the top quark and, therefore, added to the list of input observables as well. Figures 5.42 and 5.43 present the corresponding normalized distributions of these observables for signal and background events in the $t\bar{t}H$ hadronic and $t\bar{t}H$ leptonic event categories.

A comprehensive overview and ranking of all mentioned input observables used in the $t\bar{t}H$ hadronic and $t\bar{t}H$ leptonic BDT classifiers is given in Table 5.10.

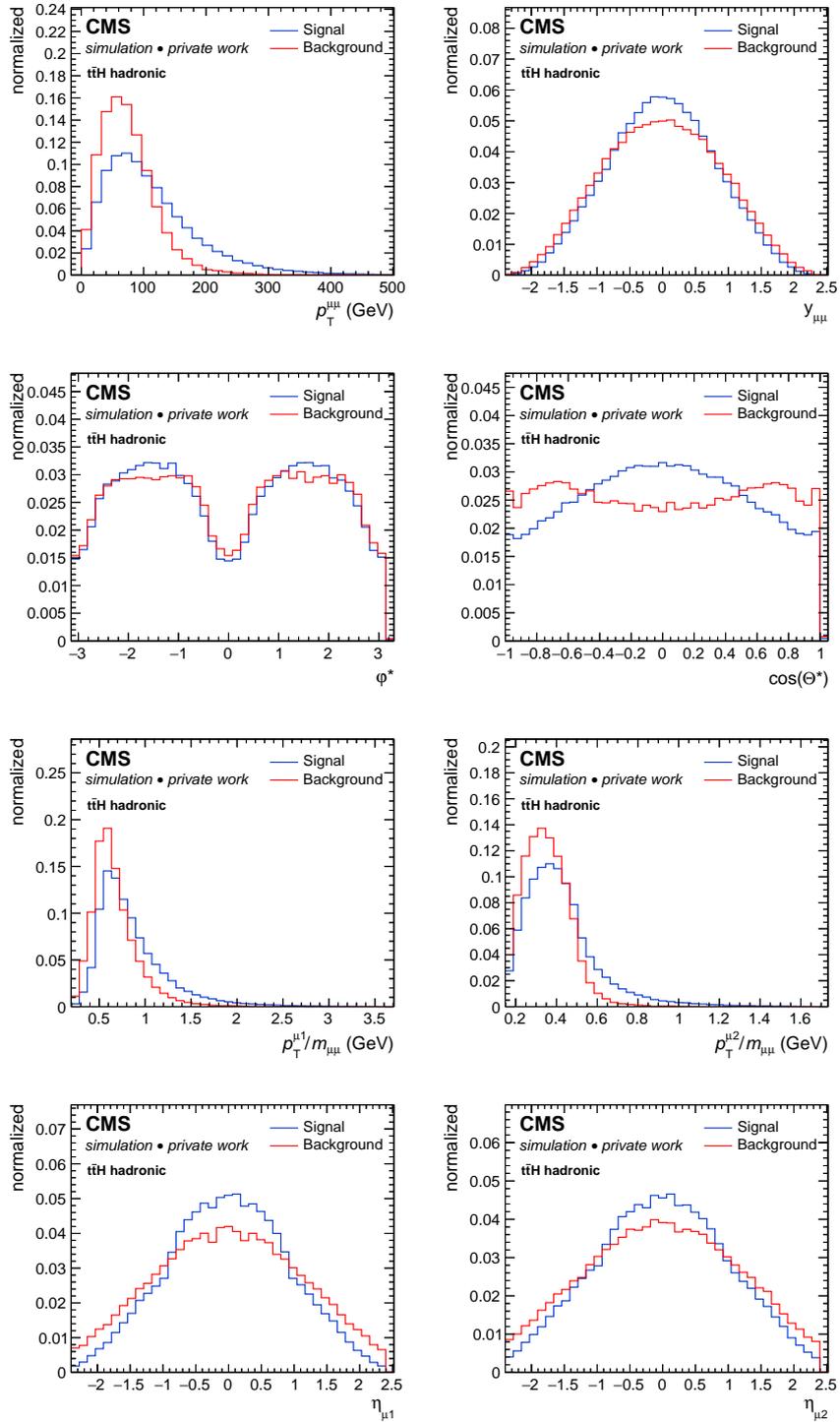


Figure 5.38: Normalized distributions of the BDT input observables used to describe the kinematic properties of the Higgs boson candidate decaying into two muons in the $t\bar{t}H$ hadronic category. The distributions are presented for signal (blue) and background (red) events.

5.9 Multivariate analysis optimization of the $t\bar{t}H$ channel

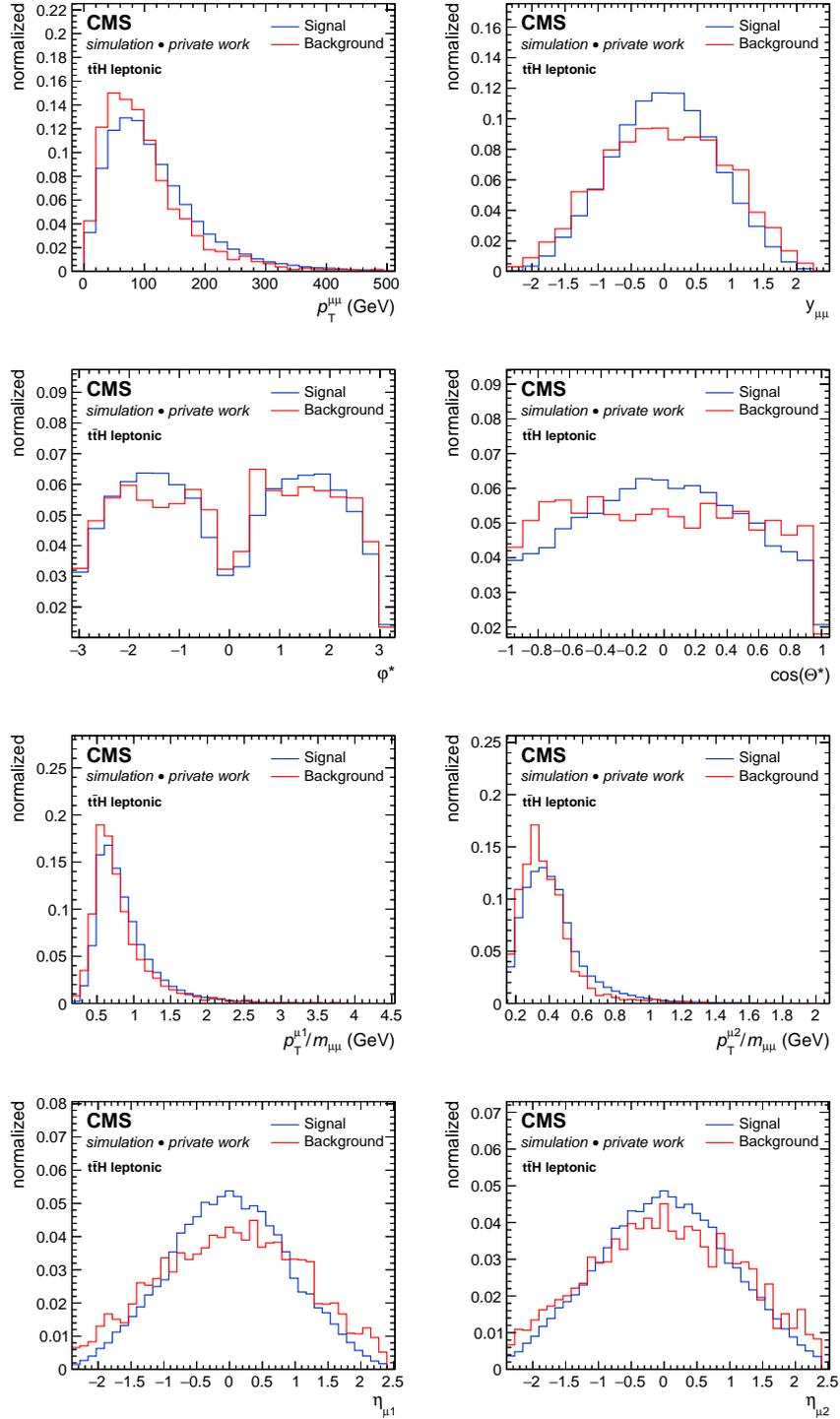


Figure 5.39: Normalized distributions of the BDT input observables used to describe the kinematic properties of the Higgs boson candidate decaying into two muons in the $t\bar{t}H$ leptonic category. The distributions are presented for signal (blue) and background (red) events.

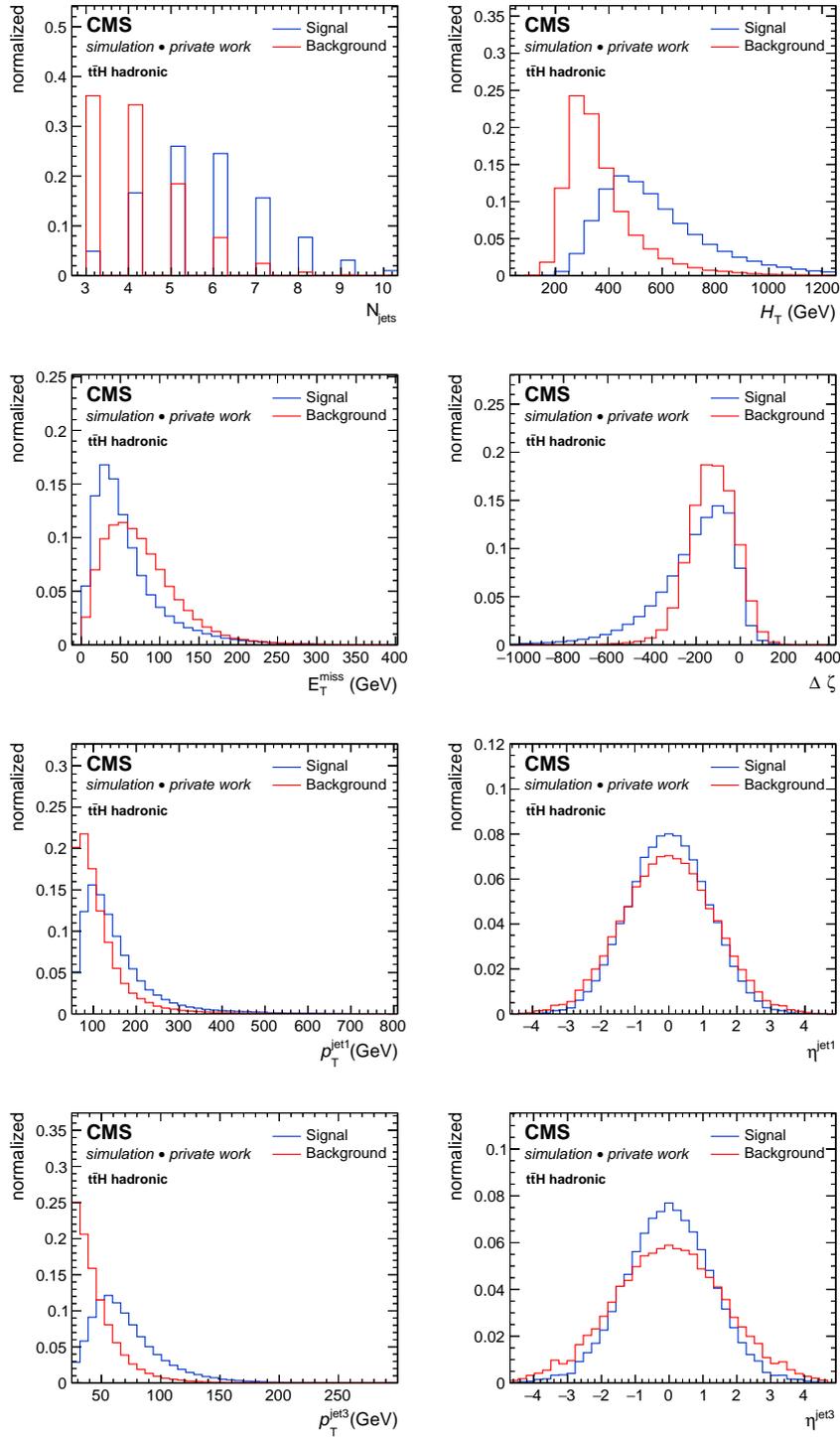


Figure 5.40: Normalized distributions of the BDT input observables used to describe the hadronic event activity in the $t\bar{t}H$ hadronic category. The distributions are presented for signal (blue) and background (red) events.

5.9 Multivariate analysis optimization of the $t\bar{t}H$ channel

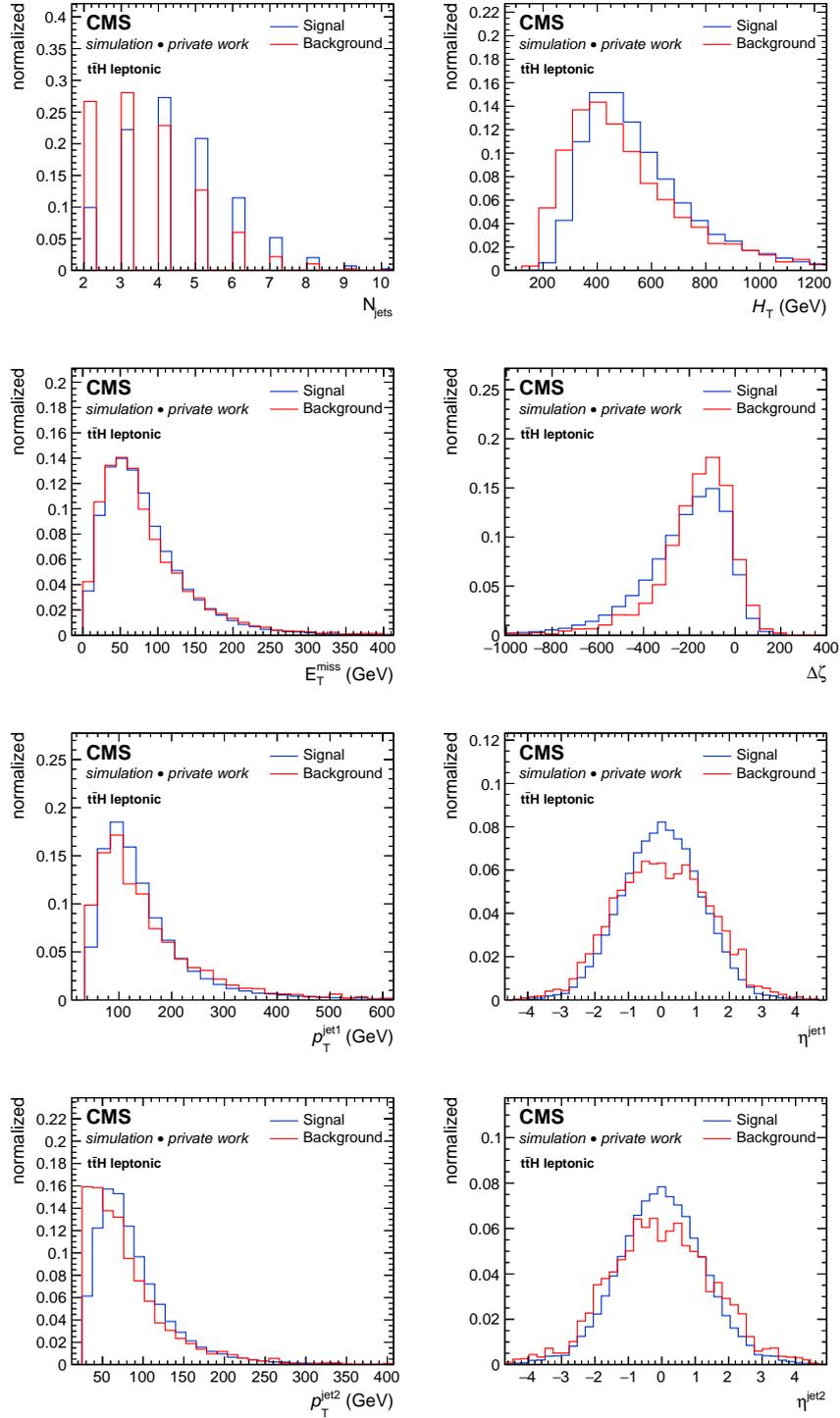


Figure 5.41: Normalized distributions of the BDT input observables used to describe the hadronic event activity in the $t\bar{t}H$ leptonic category. The distributions are presented for signal (blue) and background (red) events.

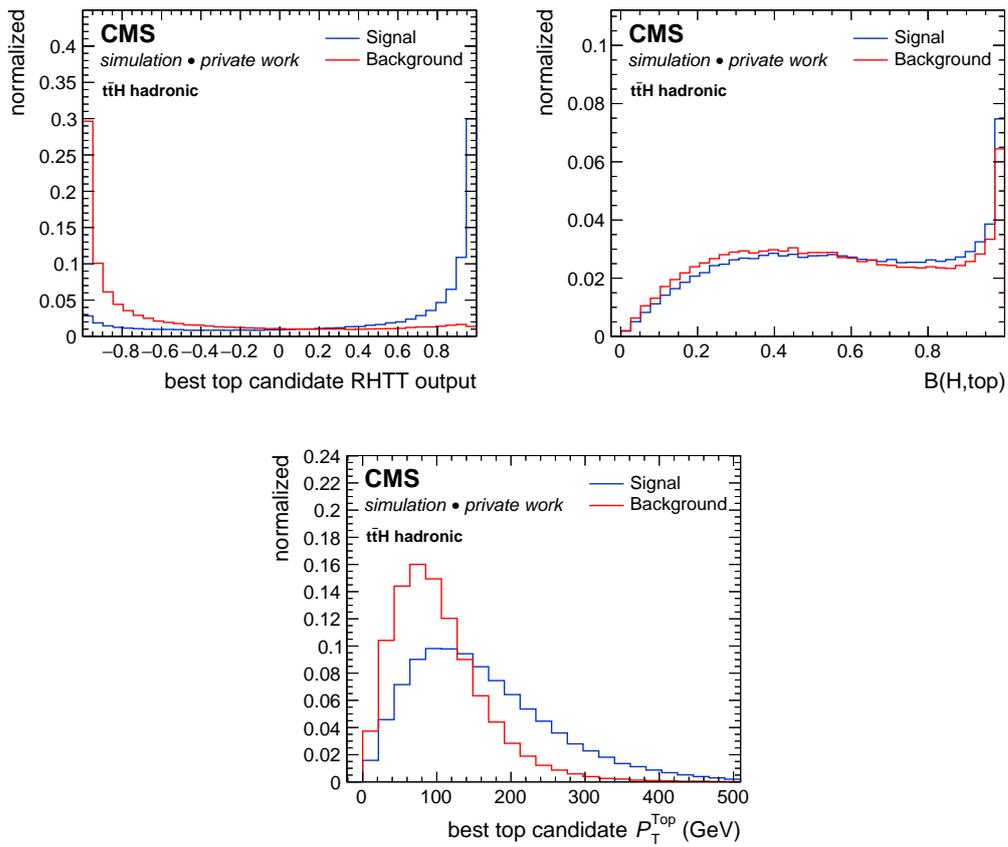


Figure 5.42: Normalized distributions of the BDT input observables used to describe the kinematic properties of the top quark decay products in the $t\bar{t}H$ hadronic category. The distributions are presented for signal (blue) and background (red) events.

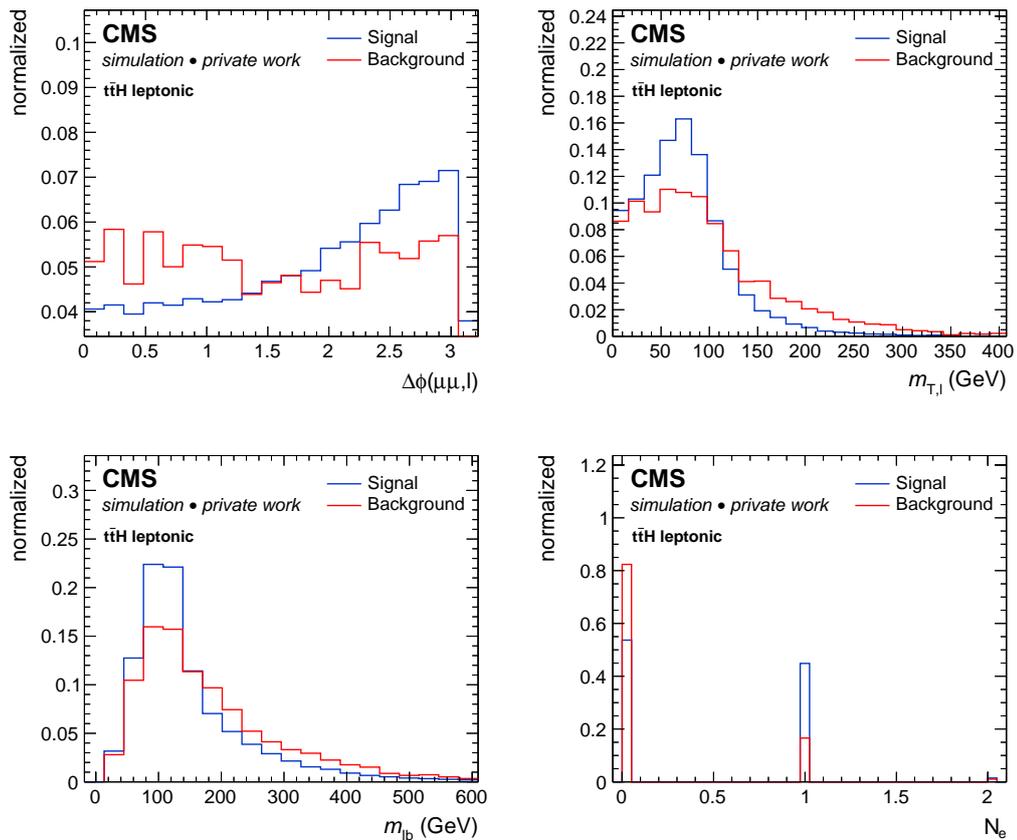


Figure 5.43: Normalized distributions of the BDT input observables used to describe the kinematic properties of the top quark decay products in the $t\bar{t}H$ leptonic category. The distributions are presented for signal (blue) and background (red) events.

Table 5.10: Description of the input observables employed by the $t\bar{t}H$ hadronic (top) and $t\bar{t}H$ leptonic (bottom) BDT classifiers. The observables are subdivided in groups and ranked according to their importance during the training.

$t\bar{t}H$ hadronic				
	Definition	Observable	Separation	Importance
Higgs boson	Transverse momentum of Higgs boson candidate	$p_T^{\mu\mu}$	9.0 %	4.4 %
	Single muon pseudorapidity	η_μ	1.9 %	4.1 %
	Polar angle in CS-frame	$\cos\theta_{CS}$	1.2 %	4.1 %
	Single muon transverse momentum	$p_T^\mu / m_{\mu\mu}$	8.4 %	3.7 %
	Rapidity of Higgs boson candidate	$y_{\mu\mu}$	0.5 %	3.7 %
	Azimuthal angle in CS-frame	$\Delta\Phi_{CS}$	0.09 %	3.0 %
Hadronic activity	Missing transverse momentum	E_T^{miss}	5.4 %	6.0 %
	E_T^{miss} projection onto dimuon bisector	$\Delta\zeta$	8.7 %	5.8 %
	Momentum sum (vectorial)	H_T^{miss}	1.7 %	5.6 %
	Momentum sum (scalar)	H_T	30.8 %	5.5 %
	Single jet η of leading jets	η^{jet}	5.5 %	5.5 %
	Max. DeepCSV of jets	max. DeepCSV	0.08 %	5.2 %
	Number of jets	N_{jets}	35.2 %	4.7 %
	Single jet transverse momentum of leading jets	p_T^{jet}	23.5 %	4.0 %
Top quark	Max. RHTT output	max. RHTT	39.9 %	6.4 %
	Transverse momentum of top candidate	p_T^{top}	20.3 %	3.1 %
	Balance of Higgs and Top candidate momentum	$p_T^{\text{H,top}}$	12.0 %	2.8 %

$t\bar{t}H$ leptonic				
	Definition	Observable	Separation	Importance
Higgs boson	Polar angle in CS-frame	$\cos\theta_{CS}$	0.7 %	5.6 %
	Single muon pseudorapidity	η_μ	1.3 %	5.2 %
	Single muon transverse momentum	$p_T^\mu / m_{\mu\mu}$	1.7 %	4.6 %
	Rapidity of Higgs boson candidate	$y_{\mu\mu}$	2.0 %	3.7 %
	Transverse momentum of Higgs boson candidate	$p_T^{\mu\mu}$	1.7 %	3.3 %
	Azimuthal angle in CS-frame	$\Delta\Phi_{CS}$	0.7 %	3.3 %
Hadronic activity	Max. DeepCSV of jets	max. DeepCSV	4.3 %	6.2 %
	Single jet η of leading jets	η^{jet}	2.4 %	5.9 %
	Number of jets	N_{jets}	7.1 %	5.6 %
	Momentum sum (scalar)	H_T	4.4 %	4.3 %
	E_T^{miss} projection onto dimuon bisector	$\Delta\zeta$	2.4 %	4.3 %
	Single jet transverse momentum of leading jets	p_T^{jet}	1.4 %	4.2 %
	Missing transverse momentum	E_T^{miss}	0.4 %	3.8 %
	Momentum sum (vectorial)	H_T^{miss}	0.6 %	2.7 %
Top quark	Number of electrons	N_e	9.6 %	5.7 %
	Transverse mass of additional lepton and E_T^{miss}	$m_T^{\ell, \text{MET}}$	5.2 %	5.4 %
	$\Delta\Phi$ Higgs candidate and add. leading lepton	$\Delta\Phi(\mu\mu, \ell)$	1.4 %	5.1 %
	Mass of b-tagged jet and add. leading lepton	$m_{\ell b}$	3.3 %	4.1 %

5.9.2 Training of the BDT classifier

The BDT training in the $t\bar{t}H$ leptonic and $t\bar{t}H$ hadronic event categories are performed using simulated $t\bar{t}H$ signal events generated with the nominal Higgs boson mass of $m_H = 125\text{ GeV}$. Other Higgs boson production modes are not considered in order to fully exploit the specific kinematic properties of the $t\bar{t}H$ production mechanism. Event generator information is used to correctly assign reconstructed muons to the Higgs boson decay. All background contributions described in Section 5.2 are taken into account. Simulated events are divided in two independent event samples for training and testing purposes. Events which enter the BDT training are not used in the final signal extraction to avoid overtraining and to guarantee an unbiased result. The final set of BDT hyperparameters is determined in an iterative procedure which optimizes the ROC integral. A BDT with 600 (200) trees and a maximum depth of three is trained in the $t\bar{t}H$ hadronic (leptonic) channel using the Gradient Boost method. The shrinkage parameter is set to 0.1.

To focus on background events in the sensitive signal-like phase-space region, only events with a dimuon mass between 115 and 135 GeV are considered in the training. Since the statistical precision of the simulated background samples is limited in this phase-space region, events from all three data-taking periods are included in the BDT training of the $t\bar{t}H$ leptonic and $t\bar{t}H$ hadronic BDT classifiers. To further increase the available number of training events in the $t\bar{t}H$ leptonic category, the event selection is relaxed by extending the dimuon mass window to a range between 110 and 140 GeV and by removing the Z boson mass veto, which is applied to all same flavour opposite charge lepton pairs. This leads to a larger background contribution from $t\bar{t}Z$ and other processes with genuine Z boson decays to muons. Although the background composition slightly changes, no significant impact on the final performance of the BDT training is found. Different options to handle the fraction of negative weighted events are tested in the same way as explained in Section 5.7. Both, positive and negative weighted events are used in the training procedure.

Figure 5.44 shows the $t\bar{t}H$ leptonic and $t\bar{t}H$ hadronic BDT output distributions. The impact of overtraining on signal and background events is studied by comparing the BDT output distributions of the training and the testing samples. The overall shape of the BDT output distributions is modeled very similarly, and differences are within the statistical precision. In addition, two example working points with 90% and 70% background rejection are selected, and the corresponding signal efficiencies are compared between the training and testing samples in Table 5.11. The working points used to define the best performing subcategories in the $t\bar{t}H$ leptonic and $t\bar{t}H$ hadronic channels correspond to a background rejection of around 90%. The absolute differences between the signal efficiencies measured in the training and the testing samples at this working point are in the order of a few permille. Thus, no significant overtraining is observed. Figure 5.45 presents the ROC diagrams of the $t\bar{t}H$ leptonic and $t\bar{t}H$ hadronic BDT discriminants with a ROC integral of 83% and 93%, respectively. The difference in performance can be explained by a larger fraction of reducible backgrounds in the $t\bar{t}H$ hadronic category compared to the $t\bar{t}H$ leptonic category. Overall, both BDT discriminants provide good performance, which allows an increase in the analysis sensitivity as presented in Section 5.10.

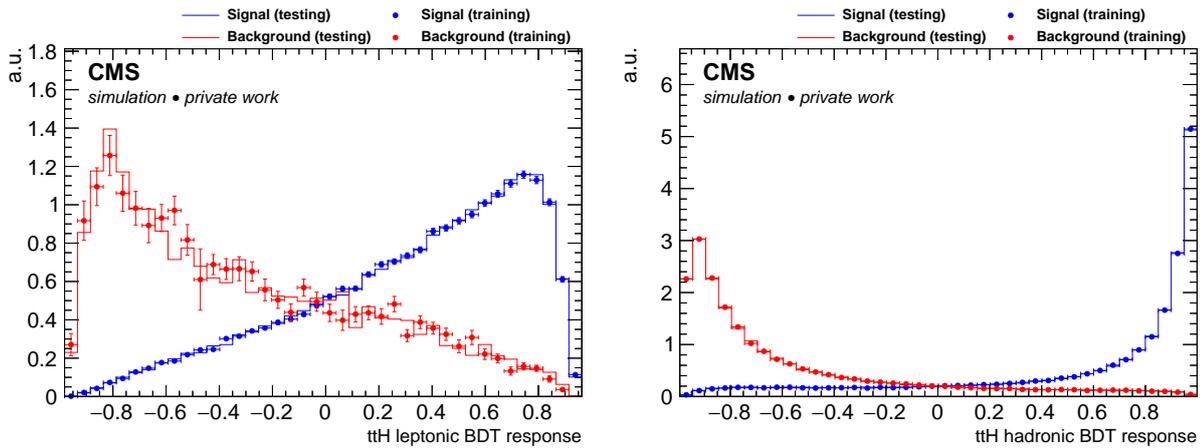


Figure 5.44: The BDT output distributions of the $t\bar{t}H$ leptonic (left) and $t\bar{t}H$ hadronic (right) categories for signal (blue) and background (red) events. The distributions obtained from the training samples (points) are compared to the corresponding test samples (solid lines).

Table 5.11: Comparison of the signal efficiency obtained from the testing sample and the training samples for two example working points of the $t\bar{t}H$ leptonic and $t\bar{t}H$ hadronic BDT discriminants.

Background rejection	$t\bar{t}H$ leptonic signal eff.		$t\bar{t}H$ hadronic signal eff.	
	testing sample	training sample	testing sample	training sample
90%	50.8%	50.3%	81.8%	81.9%
70%	79.8%	80.4%	93.6%	93.6%

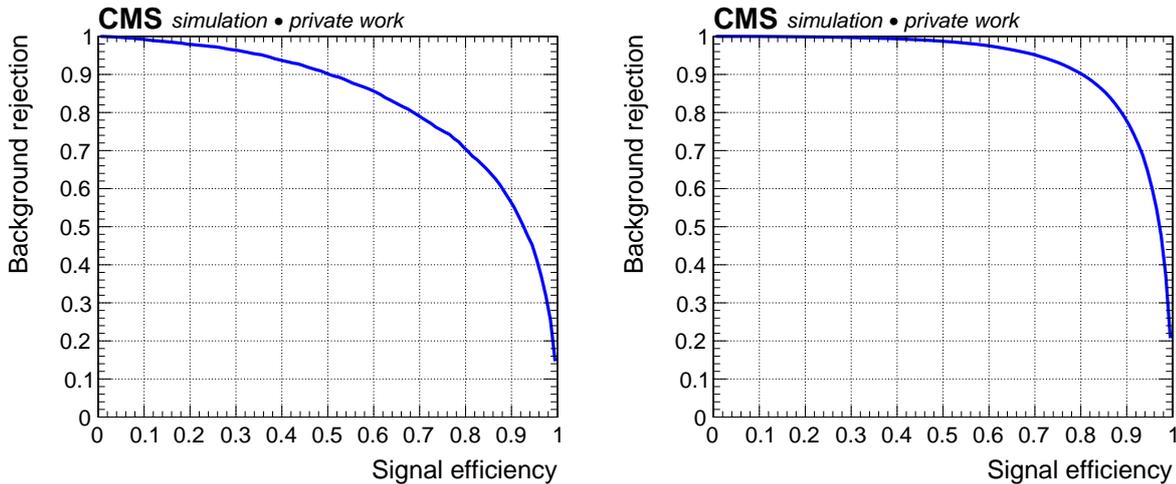


Figure 5.45: ROC diagrams indicating the signal efficiency versus background rejection for the $t\bar{t}H$ leptonic (left) and $t\bar{t}H$ hadronic (right) BDT discriminants.

Although all input observables are carefully chosen to be uncorrelated with the dimuon mass, there might be hidden correlations between input observables which allow the BDT to learn the mass of the Higgs boson. This could lead to a strong bias in the final results, and thus, it is important to exclude as far as possible a correlation between the final BDT output and the dimuon mass. For this purpose, two independent cross checks are performed. First, the evolution of the dimuon mass shape for background events is examined as a function of the BDT score as presented in Fig. 5.46. Second, the evolution of the BDT output shape for signal events is analyzed as a function of the Higgs boson mass hypothesis as shown in Fig. 5.47. No differences in the BDT output as a function of the Higgs boson mass are found. No sculpting of the dimuon mass shape is observed. This confirms that the BDT output is not significantly correlated with the dimuon mass.

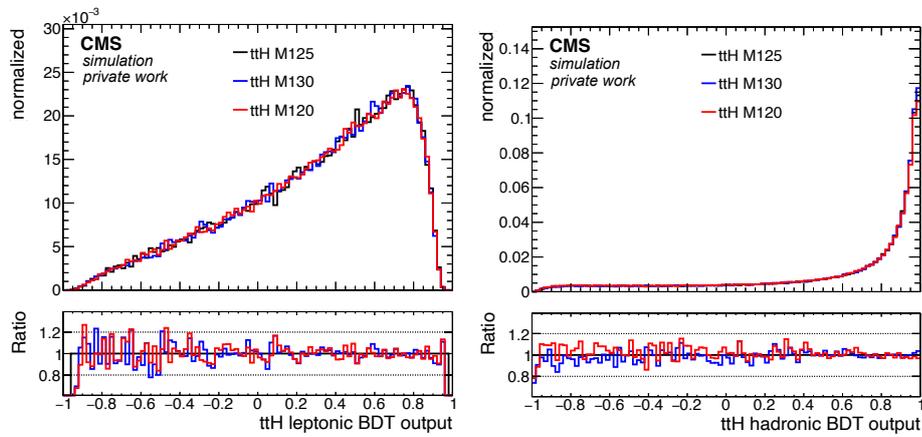


Figure 5.46: Evolution of the BDT output distributions in simulated $t\bar{t}H$ signal events as a function of the Higgs boson mass hypothesis in the $t\bar{t}H$ leptonic (left) and $t\bar{t}H$ hadronic (right) category. In the lower panel, the ratios between the nominal Higgs boson mass hypothesis of $m_H = 125$ GeV (black) and the mass probes at $m_H = 120$ GeV (red) and $m_H = 130$ GeV (blue) are shown. The distributions are normalized to have unit area.

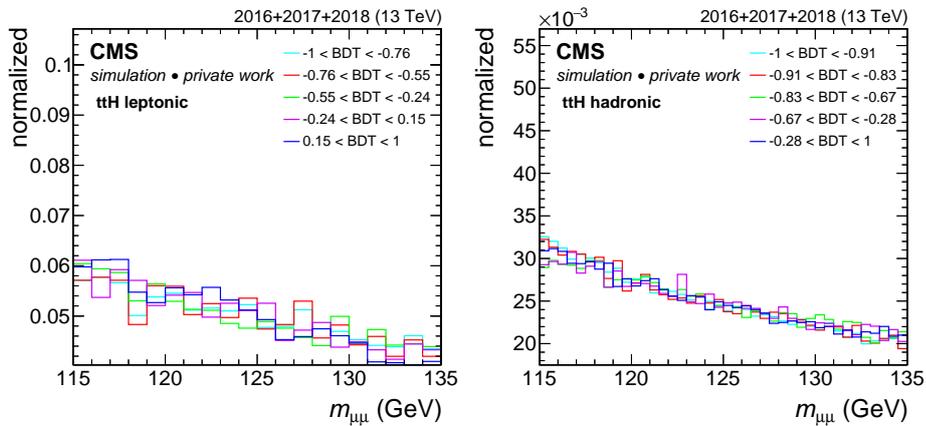


Figure 5.47: Evolution of the dimuon mass shape for five quantiles of the BDT output in simulated background events for the $t\bar{t}H$ leptonic (left) and $t\bar{t}H$ hadronic (right) category. The distributions are normalized to have unit area.

5.9.3 BDT performance & validation in the $t\bar{t}H$ leptonic channel

Important input observables and the output of the $t\bar{t}H$ leptonic BDT classifier are discussed and validated in data. The validation of the $t\bar{t}H$ leptonic BDT is performed in a signal-depleted control region. For this purpose, events in the dimuon mass sidebands passing the $t\bar{t}H$ leptonic event selection are used. Events in the signal region are excluded. The predicted distributions in simulation are compared to the observed distributions in data to ensure an unbiased event selection based on the BDT score as described in Section 5.10.

Figure 5.48 shows the $t\bar{t}H$ leptonic BDT output in data and simulation for events in the dimuon mass sideband region where events from the signal region are excluded. Important observables used as input to the $t\bar{t}H$ leptonic BDT are shown in Fig. 5.49. Overall, the shape of the different kinematic distributions is well modeled within the statistical precision of the data. The observed deviation in normalization between data and simulation corresponds to 9.5% which is compatible within 1σ of the statistical uncertainty in data. Since the background in this analysis is directly estimated from data, this small discrepancy in normalization does not affect the final results. In contrast to the relaxed selection used during the BDT training, the Z boson mass veto is included here. This particularly reduces the $t\bar{t}Z$ background contribution in the low-BDT score regime, which otherwise would be enriched in events with three muon final-state. This is also the reason for the rather flat BDT background distribution. In the high-BDT score region all reducible backgrounds from for example $t\bar{t}W(W)$, WZ, and ZZ are suppressed, while the $t\bar{t}Z$ background is still dominant. The unblinded distribution including events from the signal region is presented in Fig. 5.59.

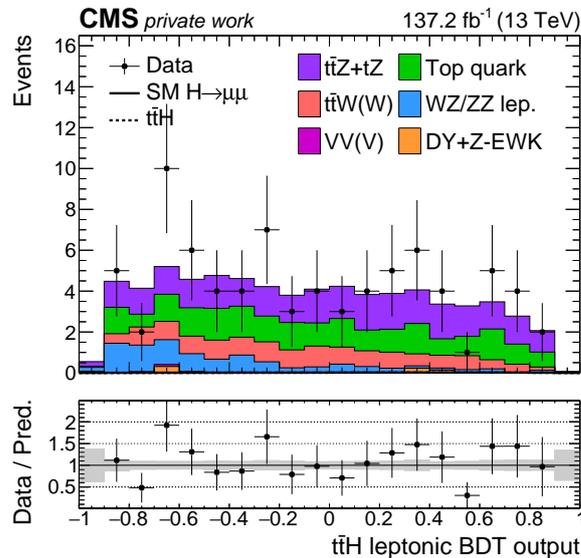


Figure 5.48: The $t\bar{t}H$ leptonic BDT output distribution for events in the dimuon mass sideband region where events from the signal region are excluded. In the lower panel, the ratio between data and the expected background is shown. The gray band indicates the statistical uncertainty. Systematic uncertainties as described in Section 5.11.2 are not included here since they are much smaller than the statistical uncertainty.

5.9 Multivariate analysis optimization of the $t\bar{t}H$ channel

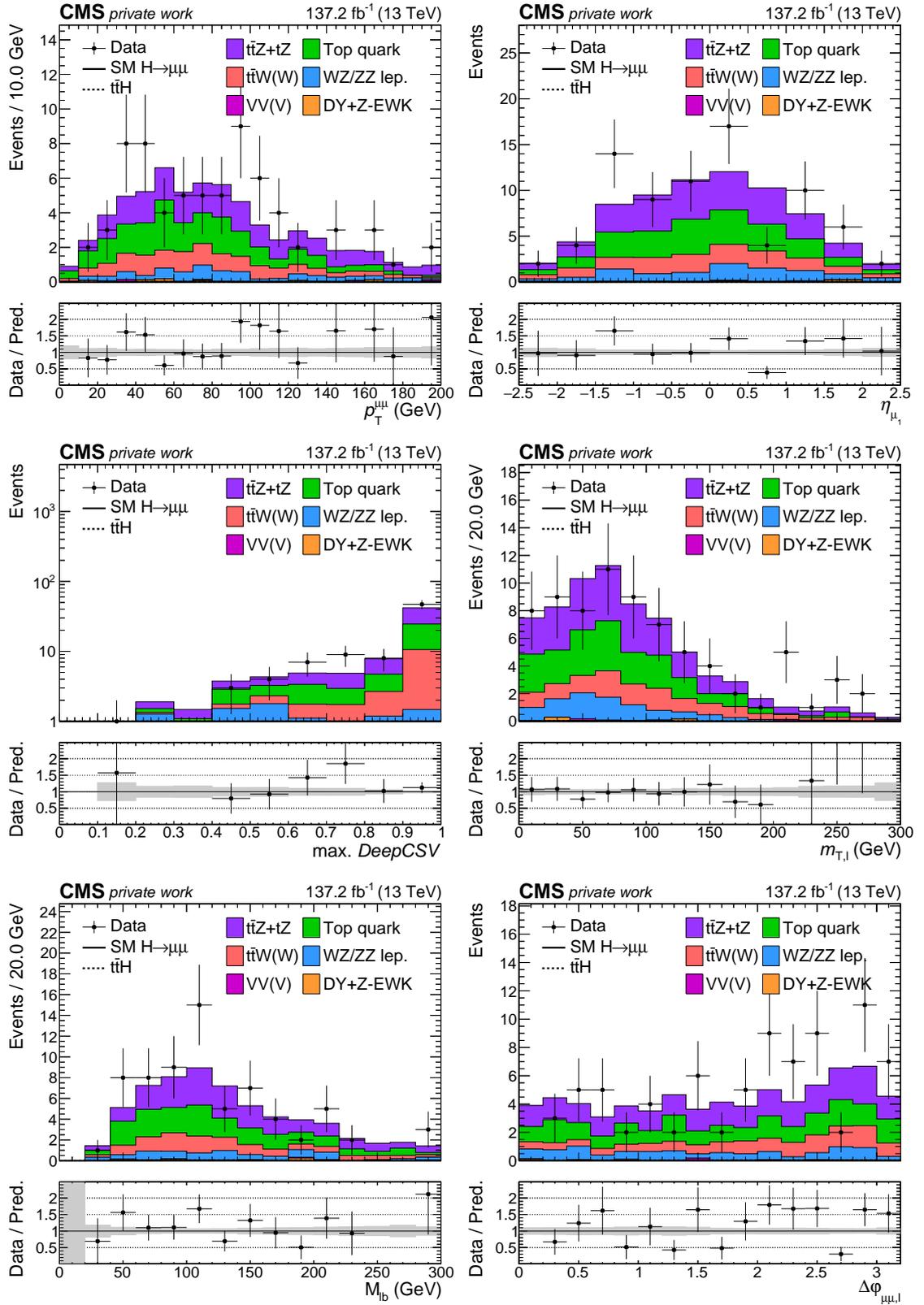


Figure 5.49: Distributions of six important observables used as input to the $t\bar{t}H$ leptonic BDT. The distributions are shown for events in the dimuon mass sideband region where events from the signal region are excluded.

To illustrate the kinematic properties of signal events as a function of the BDT score, the $t\bar{t}H$ leptonic BDT is divided into five quantiles. Figure 5.50 shows the shape evolution of six important input observables as a function of the $t\bar{t}H$ leptonic BDT output. The high-BDT score region is enriched in events with a large dimuon p_T induced by a large hadronic recoil of the Higgs boson. Since the BDT is aware of the dimuon mass resolution, in particular central muons with accurate transverse momentum resolution are selected in the high-BDT score region. Furthermore, events with high-BDT output tend to contain at least one jet with a high DeepCSV score to increase the fraction of genuine b jets. The reconstructed $m_{\ell b}$ mass of the b jet and the additional lepton peaks well below the top quark mass. Larger values of $m_{\ell b}$ are suppressed for events with high-BDT score, since larger values indicate events where the selected b jet and lepton do not originate from the same top quark. The transverse mass of the additional lepton and the missing transverse momentum m_T^ℓ is employed to reconstruct the leptonic decay of the W boson. Events with a high-BDT score tend to have values of m_T^ℓ around the W boson mass. Furthermore, events in the high-BDT score prefer a kinematic configuration where the Higgs boson and the additional lepton from the decay of the top quark are scattered back-to-back in the transverse plane.

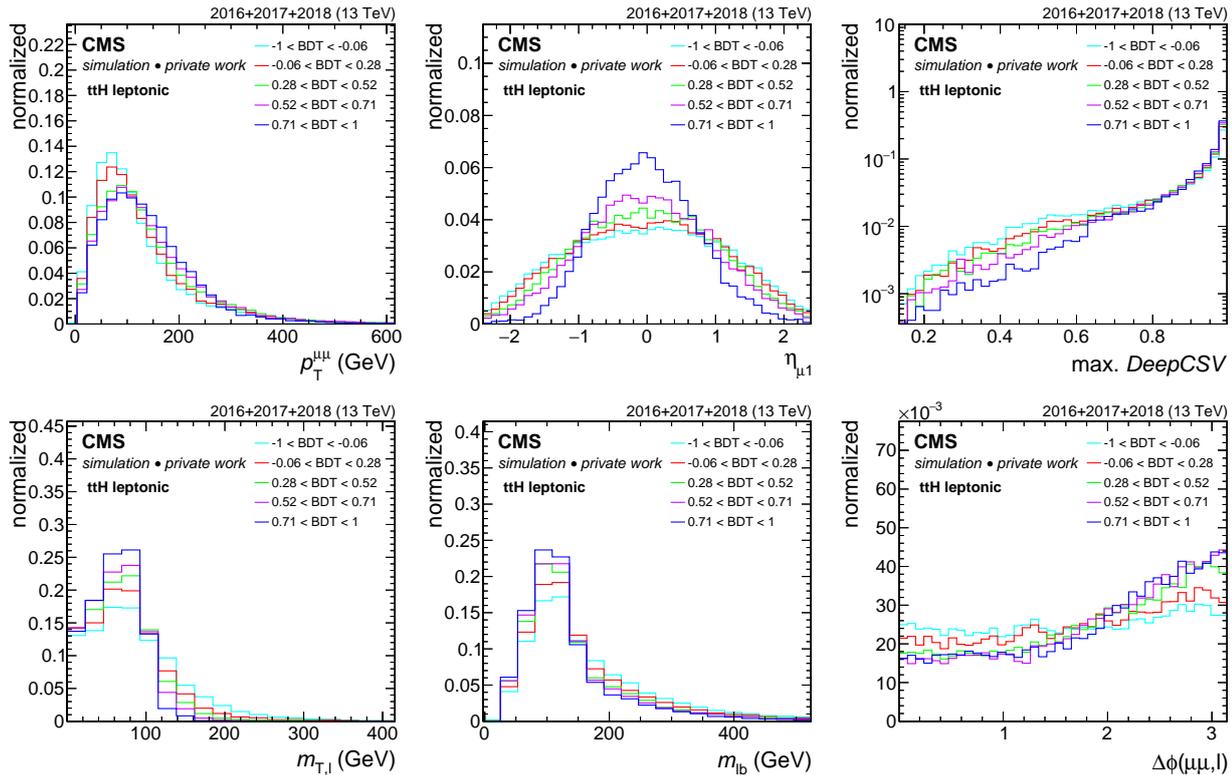


Figure 5.50: Shape evolution of six important input observables for five quantiles of the $t\bar{t}H$ leptonic BDT output in simulated $t\bar{t}H$ signal events. Each distribution is scaled to have unit area.

5.9.4 BDT performance & validation in the $t\bar{t}H$ hadronic channel

The most important input observables and the output of the $t\bar{t}H$ hadronic BDT classifier are discussed and validated in data in the same way as presented in the previous section for the $t\bar{t}H$ leptonic BDT. The validation of the $t\bar{t}H$ hadronic BDT is performed using events in the dimuon mass sideband region which pass the $t\bar{t}H$ hadronic event selection. Events in the signal region are excluded.

Figure 5.51 shows the transformed output of the $t\bar{t}H$ hadronic BDT discriminant. The original BDT output is transformed to expand the high-BDT score region using $\text{arctanh}((\text{BDT} + 1)/2)$. Important input observables of the $t\bar{t}H$ hadronic BDT are shown in Fig. 5.52. The predicted normalization and the shape of the different kinematic distributions in simulated events is generally in good agreement with the observed distribution in data. Most discrepancies are well covered by the considered systematic uncertainties, which are described in Section 5.11.2. The largest differences between data and simulation are observed in the number of selected jets. In particular, for very high jet multiplicities the data yield is underestimated by the simulation. In this region of the phase-space both dominant backgrounds from $t\bar{t}$ and Drell-Yan rely on a significant number of jets from the PS, which are known to not entirely reproduce the data distribution [209]. Since the high-BDT score region of the $t\bar{t}H$ hadronic category is enriched in events with large jet multiplicity, this disagreement is also the reason for the excess of data events in the $t\bar{t}H$ hadronic BDT output. However, the background prediction is estimated using a data-driven method and the prediction of jets in signal events relies largely on the much more precise ME calculation. Thus, the observed differences between data and simulation do not affect the final fit result. The unblinded distribution including events from the signal region is presented in Fig. 5.59.

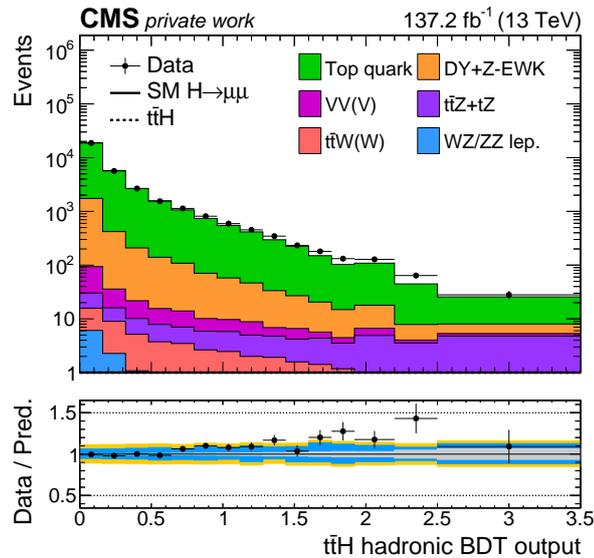


Figure 5.51: The transformed $t\bar{t}H$ hadronic BDT output distribution for events in the dimuon mass sideband region. Events in the signal region are excluded. The gray band indicates the statistical uncertainty of the simulated background samples. The azure band corresponds to the sum in quadrature between the statistical uncertainty and experimental systematic uncertainties. The orange band includes theoretical uncertainties affecting the background.

Chapter 5. Higgs boson decays to muons produced in association with top quarks

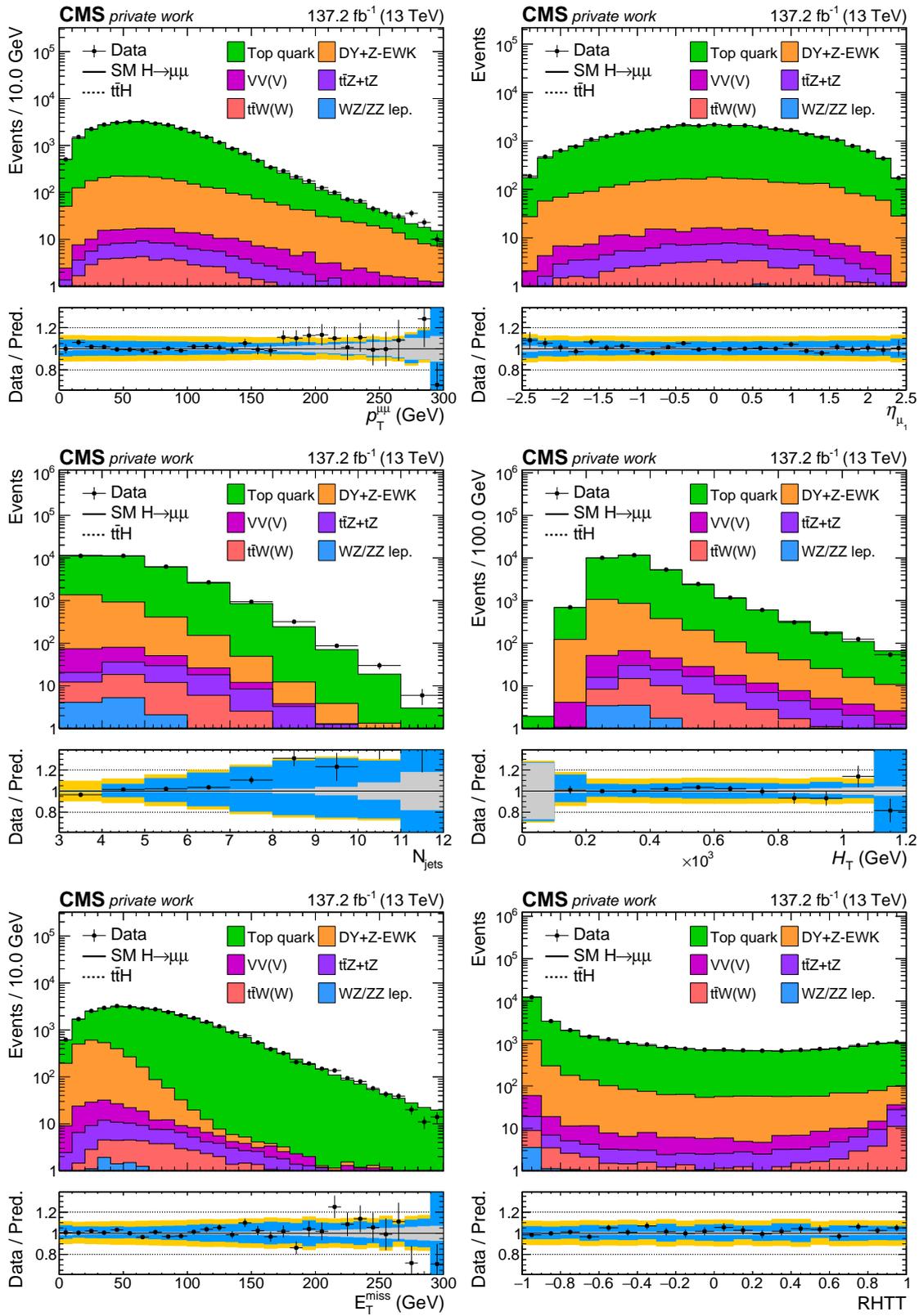


Figure 5.52: Distributions of six important observables used as input to the $t\bar{t}H$ hadronic BDT. The distributions are shown for events in the dimuon mass sideband region where events from the signal region are excluded.

The $t\bar{t}H$ hadronic BDT is divided into five quantiles to illuminate kinematic properties of signal events as a function of the BDT output. Figure 5.53 shows the shape evolution of six important input observables as a function of the $t\bar{t}H$ hadronic BDT score. Similar to the $t\bar{t}H$ leptonic BDT, also the high-BDT regime of the $t\bar{t}H$ hadronic BDT primarily selects events with large dimuon p_T , and muons, that are reconstructed in the central part of the detector. Furthermore, events with high BDT score are characterized by large jet multiplicities and large H_T . In $t\bar{t}H$ signal events where both top quarks decay hadronically, imperfect measurements of the jet momenta lead to small E_T^{miss} . Since the dominant background from dileptonic $t\bar{t}$ decays contains two neutrinos, much larger E_T^{miss} is expected in background events. Therefore, events with high BDT score tend to have small values of E_T^{miss} . Finally, events in the high-BDT region contain a top quark candidate with high RHTT score. This ensures the presence of at least one well reconstructed hadronic top quark in the event.

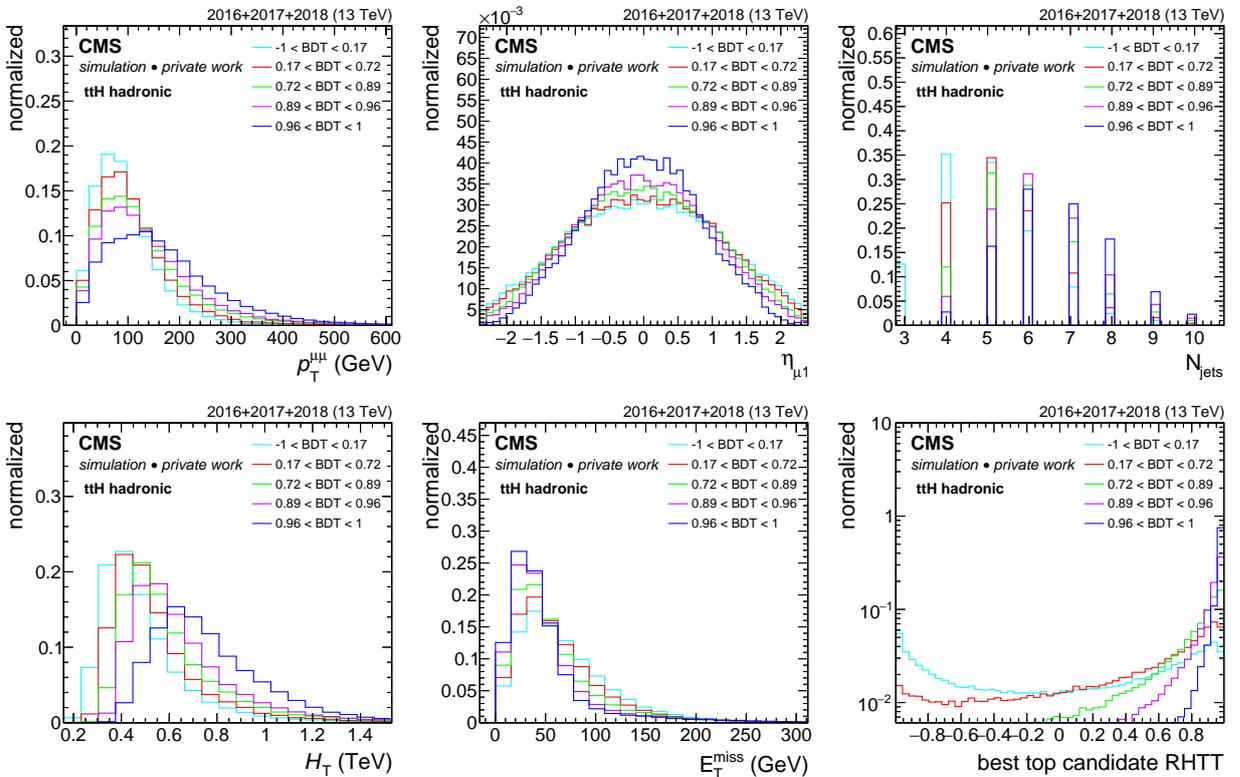


Figure 5.53: Shape evolution of six important input observables for five quantiles of the $t\bar{t}H$ hadronic BDT output in simulated $t\bar{t}H$ signal events. The prediction for all three data-taking periods is summed. All distribution are scaled to have unit area.

5.10 Event categorization

To improve the overall sensitivity of the $t\bar{t}H$ analysis, events are classified according to their BDT score, and grouped into several subcategories. The optimization of the subcategory boundaries is based on an iterative algorithm. The aim of this algorithm is to maximize the analysis sensitivity in terms of the expected significance. For this purpose, a simplified version of the final signal extraction fit described in Section 5.11 is performed. The expected significance is estimated by an analytical fit to the dimuon mass distribution of simulated signal and background events based on all three data-taking periods. A Gaussian is chosen to model the signal contribution, while the total expected background is modeled using an exponential function. Systematic uncertainties are neglected here.

The category boundaries are determined by a scan of the BDT distribution. Events are divided into two subcategories at a certain BDT value, and the combined expected significance is evaluated. A stepwidth of 0.025 in the BDT score is chosen in the $t\bar{t}H$ hadronic category, while a larger stepwidth of 0.05 is used in the $t\bar{t}H$ leptonic category to be less affected by statistical fluctuations. The BDT boundary which provides the largest combined expected significance is selected, and the procedure is repeated to define additional subcategory boundaries until the gain in the expected significance is less than 1%. The BDT boundary optimization results in two $t\bar{t}H$ leptonic subcategories, and three $t\bar{t}H$ hadronic categories, which are shown in Fig. 5.54. A summary of the BDT subcategories is reported in Table 5.12.

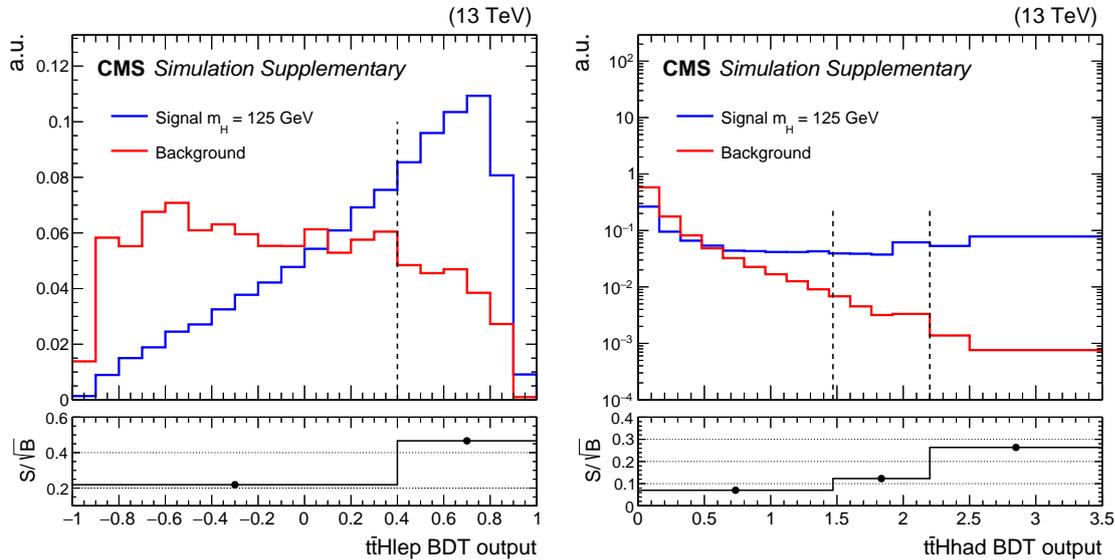


Figure 5.54: Distribution of the $t\bar{t}H$ leptonic (left) and $t\bar{t}H$ hadronic (right) BDT output in simulated signal (blue) and background (red) events. The dashed vertical lines indicate the boundaries of the optimized event categories. The lower panel provides the S/\sqrt{B} obtained by integrating signal (S) and background (B) events inside the half width half maximum of the signal peak. This result is part of the publicly available supplementary material of Ref. [172].

The expected significance due to this subcategorization improves by a factor 1/3 in the $t\bar{t}H$ leptonic category, and by a factor of 3 in the $t\bar{t}H$ hadronic category. This difference is explained by a much larger contribution of reducible backgrounds in the $t\bar{t}H$ hadronic category compared to the $t\bar{t}H$ leptonic category, where the largest fraction of reducible background is already rejected by the LEPTONMVA selection criteria.

Table 5.12: Summary of $t\bar{t}H$ event categories defined along the $t\bar{t}H$ hadronic and $t\bar{t}H$ leptonic BDT output. For each category the optimized category boundaries, and the corresponding signal efficiency quantiles are given.

Event category	Signal efficiency	BDT boundaries
$t\bar{t}H_{lep}$ -cat2	>52%	≥ 0.40
$t\bar{t}H_{lep}$ -cat1	0-52%	< 0.40
$t\bar{t}H_{had}$ -cat3	>86%	≥ 0.95
$t\bar{t}H_{had}$ -cat2	70-86%	[0.80, 0.95)
$t\bar{t}H_{had}$ -cat1	0-70%	< 0.80

5.11 Signal extraction

The $H \rightarrow \mu^+\mu^-$ signal is extracted by a fit of an analytical *signal-plus-background* (S+B) model to the shape of the $m_{\mu\mu}$ spectrum in data across all five $t\bar{t}H$ subcategories. Signal events appear as a narrow resonance in the dimuon mass spectrum on top of a monotonically falling background distribution. This allows to estimate the background contribution under the signal peak in a data-driven way using signal-depleted mass sidebands as illustrated in Fig. 5.55. This signal extraction method is typically referred to as *sideband analysis*.

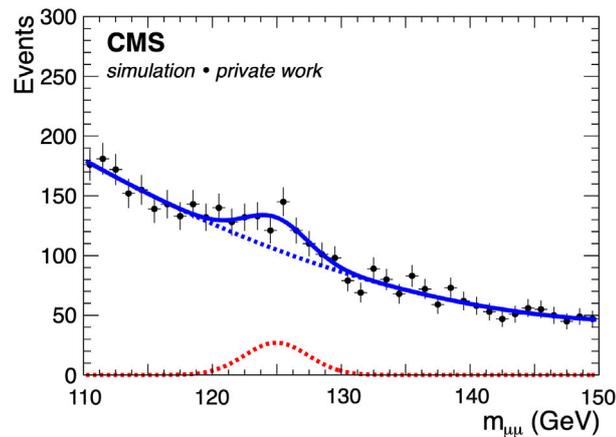


Figure 5.55: Illustration of the dimuon mass sideband fit method using pseudo data. The signal (dashed red line), and the background model (dashed blue line) are combined, and the signal-plus-background model (solid blue line) is fit to a pseudo dataset (black points). For the purpose of this example, the signal is amplified.

Signal shape parameters are determined from simulated events using all three data-taking periods together, but independently for each Higgs boson production mode. In the final S+B fit, these signal shape parameters are treated as constants or as constrained parameters to account for systematic variations of the signal peak position and resolution. Background shape parameters are bounded to ensure a reasonable modeling of the falling $m_{\mu\mu}$ background spectrum. Within these bounds the background parameters are constrained by data in the S+B fit. The overall normalization of the signal and the background processes is allowed to float freely in the fit to data.

5.11.1 Analytical signal and background model

The ratio between the natural expected width of the SM Higgs boson $\Gamma_H \approx 4$ MeV [54] and the Higgs boson mass, Γ_H/m_H , is much smaller than the typical relative momentum resolution of high- p_T muons from the decay of the Higgs boson. Thus, the shape of the measured Higgs boson mass distribution is driven by the detector resolution which induces a nearly Gaussian smearing of the reconstructed Higgs boson mass. For this reason, the signal is modeled using a *double-sided crystal ball* (DCB) function [210] which is based on a Gaussian core with mean \hat{m} and a width σ , and two power-law functions for the tails:

$$\text{DCB}(m_{\mu\mu}) = \begin{cases} e^{-(m_{\mu\mu}-\hat{m})^2/2\sigma^2}, & -\alpha_L < \frac{m_{\mu\mu}-\hat{m}}{\sigma} < \alpha_R \\ \left(\frac{n_L}{|\alpha_L|}\right)^{n_L} e^{-\alpha_L^2/2} \left(\frac{n_L}{|\alpha_L|} - |\alpha_L| - \frac{m_{\mu\mu}-\hat{m}}{\sigma}\right)^{-n_L}, & \frac{m_{\mu\mu}-\hat{m}}{\sigma} \leq -\alpha_L \\ \left(\frac{n_R}{|\alpha_R|}\right)^{n_R} e^{-\alpha_R^2/2} \left(\frac{n_R}{|\alpha_R|} - |\alpha_R| + \frac{m_{\mu\mu}-\hat{m}}{\sigma}\right)^{-n_R}, & \frac{m_{\mu\mu}-\hat{m}}{\sigma} \geq \alpha_R \end{cases}. \quad (5.14)$$

The power-law functions with the parameters α_L and n_L for the low-mass regime, and α_R and n_R for the high-mass regime, provide additional flexibility to better model non-Gaussian tails.

Compared to other signal models, as for example the sum of three Gaussians, which was used in an earlier iteration of the CMS $H \rightarrow \mu^+\mu^-$ analysis [41], the DCB has only one parameter associated to the peak and one to the width of the lineshape. This allows for a straightforward assignment of shape uncertainties on the peak position and the width of the Gaussian core.

Figure 5.56 shows the fitted signal models for all five $t\bar{t}H$ subcategories. The DCB function is able to accurately model the peak position and width, as well as the low- and high-mass tails in all categories. The total expected event yields, the signal composition, and the signal resolution are given in Table 5.13. The high BDT score categories $t\bar{t}H_{\text{lep-cat2}}$ and $t\bar{t}H_{\text{had-cat3}}$ provide a very high $t\bar{t}H$ signal purity of around 95%. In contrast, $t\bar{t}H_{\text{had-cat1}}$ contains a large fraction of ggH events with additional ISR jets, and VH events where the gauge boson decays into two quarks. Since the information about the dimuon mass resolution is included in the BDT training, the dimuon mass resolution improves as a function of the BDT score. The signal resolution in the $t\bar{t}H$ leptonic channel differs between the high and the low BDT score categories by about 10%, while this difference is much smaller in the $t\bar{t}H$ hadronic categories. This has two reasons. First, the relative importance of the muon η , which is strongly correlated with

the dimuon mass resolution, is significantly larger in the $t\bar{t}H$ leptonic BDT compared to the $t\bar{t}H$ hadronic BDT. Second, the signal in the low-BDT score region in the $t\bar{t}H$ leptonic category consists primarily of events with three muons in the final state. Thus, an occasional wrong assignment of muons to the Higgs boson candidate leads to a worse dimuon mass resolution compared to events with two muons and one electron in the final state.

Table 5.13: Summary of the total expected signal yield, the signal composition of different Higgs boson production modes, and the signal resolution in terms of the *half width at half maximum* (HWHM) for each $t\bar{t}H$ event category. Modified version published in Ref. [172].

Event category	Total signal	$t\bar{t}H$ (%)	ggH (%)	VH (%)	tH,VBF,bbH (%)	HWHM (GeV)
$t\bar{t}H$ lep-cat2	0.99	94.7	—	1.0	4.3	1.75
$t\bar{t}H$ lep-cat1	1.06	85.8	—	4.7	9.5	1.92
$t\bar{t}H$ had-cat3	1.33	94.0	0.3	1.3	4.4	1.80
$t\bar{t}H$ had-cat2	1.62	84.3	3.8	5.6	6.2	1.81
$t\bar{t}H$ had-cat1	6.87	32.3	40.3	17.2	10.2	1.85

The dominant background in the $t\bar{t}H$ leptonic categories are $t\bar{t}Z$ events, and dileptonic $t\bar{t}$ events with additional non-prompt leptons. The background in the $t\bar{t}H$ hadronic categories dominantly consists of dileptonic $t\bar{t}$ events with additional ISR jets and minor contributions from Drell-Yan and $t\bar{t}Z$. To model the monotonically decreasing $m_{\mu\mu}$ spectrum expected from these backgrounds, three different families of analytical functions are considered:

Bernstein polynomials

$$\text{Bern}(m_{\mu\mu}) = \sum_i^n a_i \cdot \binom{n}{i} m_{\mu\mu}^i (1 - m_{\mu\mu})^{n-1}, \quad (5.15)$$

Sum of power law

$$\text{SumPow}(m_{\mu\mu}) = \sum_i^n a_i \cdot m_{\mu\mu}^{b_i}, \quad (5.16)$$

Sum of exponentials

$$\text{SumExp}(m_{\mu\mu}) = \sum_i^n a_i \cdot \exp(b_i \cdot m_{\mu\mu}). \quad (5.17)$$

Due to the rather small number of events in most of the $t\bar{t}H$ subcategories, one to two degrees of freedom are sufficient to model the $m_{\mu\mu}$ background shape. In the following, the tested background functions are referred to as "SumofExp" corresponding to a sum of two exponential functions, "SingleExp" corresponding to a single exponential function, "SumofPowerlaw" corresponding to a sum of two power law functions, "SinglePowerlaw" corresponding to a single power law function, and "Bernstein" corresponding to a second-order Bernstein polynomial.

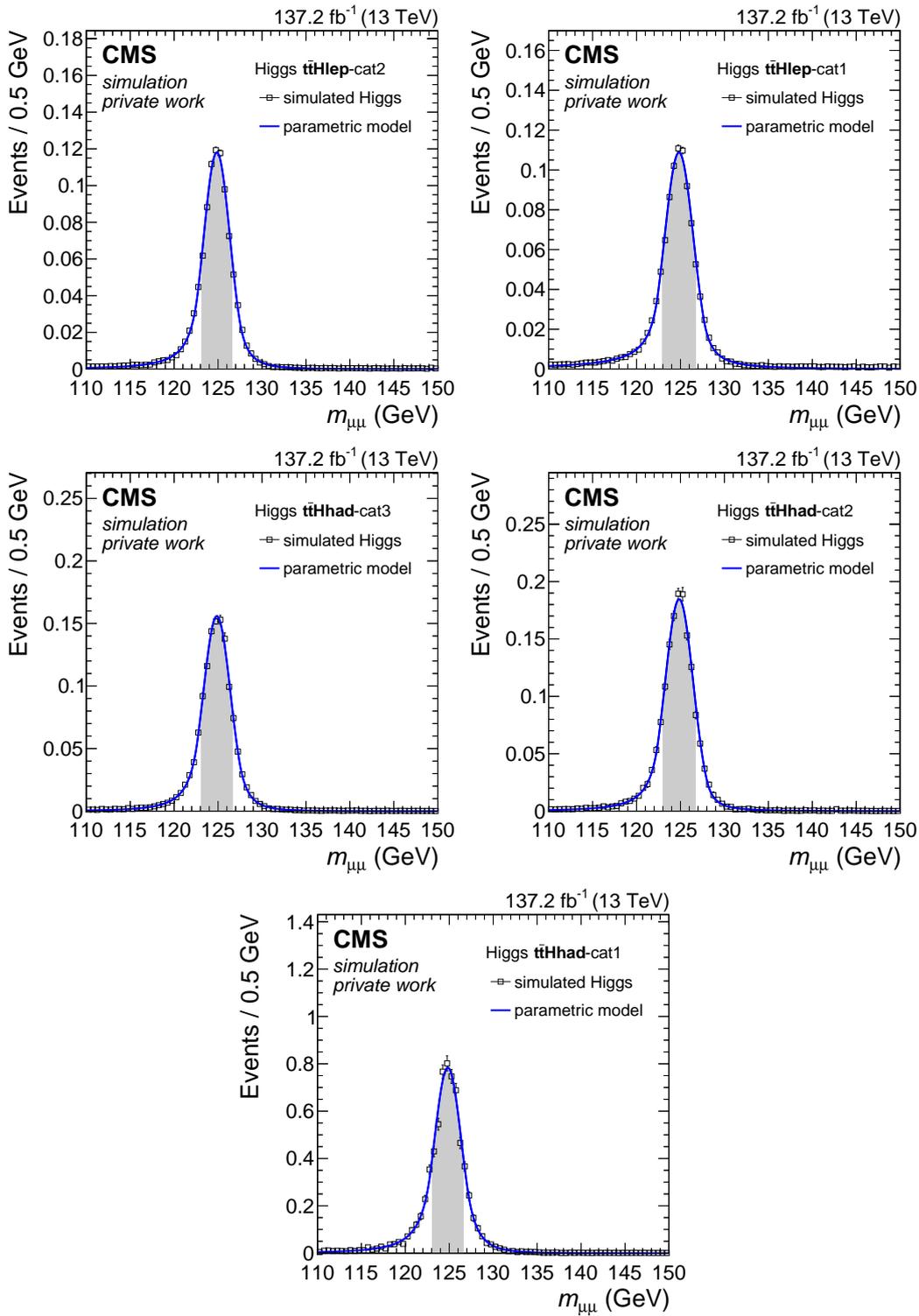


Figure 5.56: Distributions of the dimuon mass for $H \rightarrow \mu^+\mu^-$ signal events and the corresponding signal models in all $t\bar{t}H$ subcategories. The shaded area indicates to the full width at half maximum of the lineshape. The presented signal expectation corresponds to the sum of all Higgs boson production modes in all three data-taking periods.

For each of the five $t\bar{t}H$ subcategories, a single function among the considered ones is chosen to model the background. Since the underlying $m_{\mu\mu}$ shape of the background is unknown, the choice of a particular background function can induce a bias in the measured signal yield. This bias is examined using pseudodata of the $m_{\mu\mu}$ background shape. In each subcategory, background-only fits are performed to the $m_{\mu\mu}$ distribution in data for all considered background functions. Based on these background-only models, pseudodata are generated by injecting signal events with a rate corresponding to the SM prediction. The amount of signal events is allowed to vary within the statistical uncertainty. To measure the bias with a precision of about 2%, 2000 pseudo datasets are generated for each background function in each category. Finally, S+B fits are performed to each of the pseudodata, where the background component is modeled by one of the considered background functions. The signal strength $\hat{\mu}$ describes the ratio of the measured signal rate and the signal rate expected from the SM prediction. The pull of the signal strength is distributed like a Gaussian, which is centered at zero with a width of one. The bias is defined by any shift of the peak position of the pull

$$\text{Pull}(\hat{\mu}_{\text{fit}}, \sigma_{\hat{\mu}_{\text{fit}}}) = \frac{\hat{\mu}_{\text{fit}} - 1}{\sigma_{\hat{\mu}_{\text{fit}}}}, \quad (5.18)$$

where $\hat{\mu}_{\text{fit}}$ corresponds to the measured signal strength extracted from the fit, and $\sigma_{\hat{\mu}_{\text{fit}}}$ corresponds to the post-fit uncertainty. The peak value of the pull distribution is obtained by a Gaussian fit. A background model is considered as unbiased, if the additional uncertainty induced by the choice of the background function is less than 20% of the statistical uncertainty on the measured signal strength. This guarantees a negligible impact of less than 2% on the final results due to the choice of the background model.

The background function which provides the largest expected sensitivity to the $H \rightarrow \mu^+\mu^-$ signal, and at the same time a robust and unbiased description of the $m_{\mu\mu}$ spectrum is chosen to model the background. The results of the bias study are presented in Fig. 5.57. In general, all considered background functions are able to model the generated $m_{\mu\mu}$ distributions without inducing a strong bias in the measured signal strength. In a few cases, however, a bias larger than 20% is observed which excludes the corresponding background function. Among the unbiased background models, the best performance is achieved using a second-order Bernstein polynomial in $t\bar{t}H_{\text{had-cat1}}$ and $t\bar{t}H_{\text{had-cat2}}$, a sum of two exponentials in $t\bar{t}H_{\text{had-cat3}}$, and a single exponential in both $t\bar{t}H$ leptonic categories. These background models are employed in the final signal extraction fit.

Figure 5.58 shows the observed $m_{\mu\mu}$ distributions in all $t\bar{t}H$ subcategories where the signal region around the expected Higgs boson mass peak is still blinded. A background-only fit is performed in each of the subcategories using all considered background functions. The ratio panel presents the residual between the different models with respect to the chosen model in the corresponding subcategory.

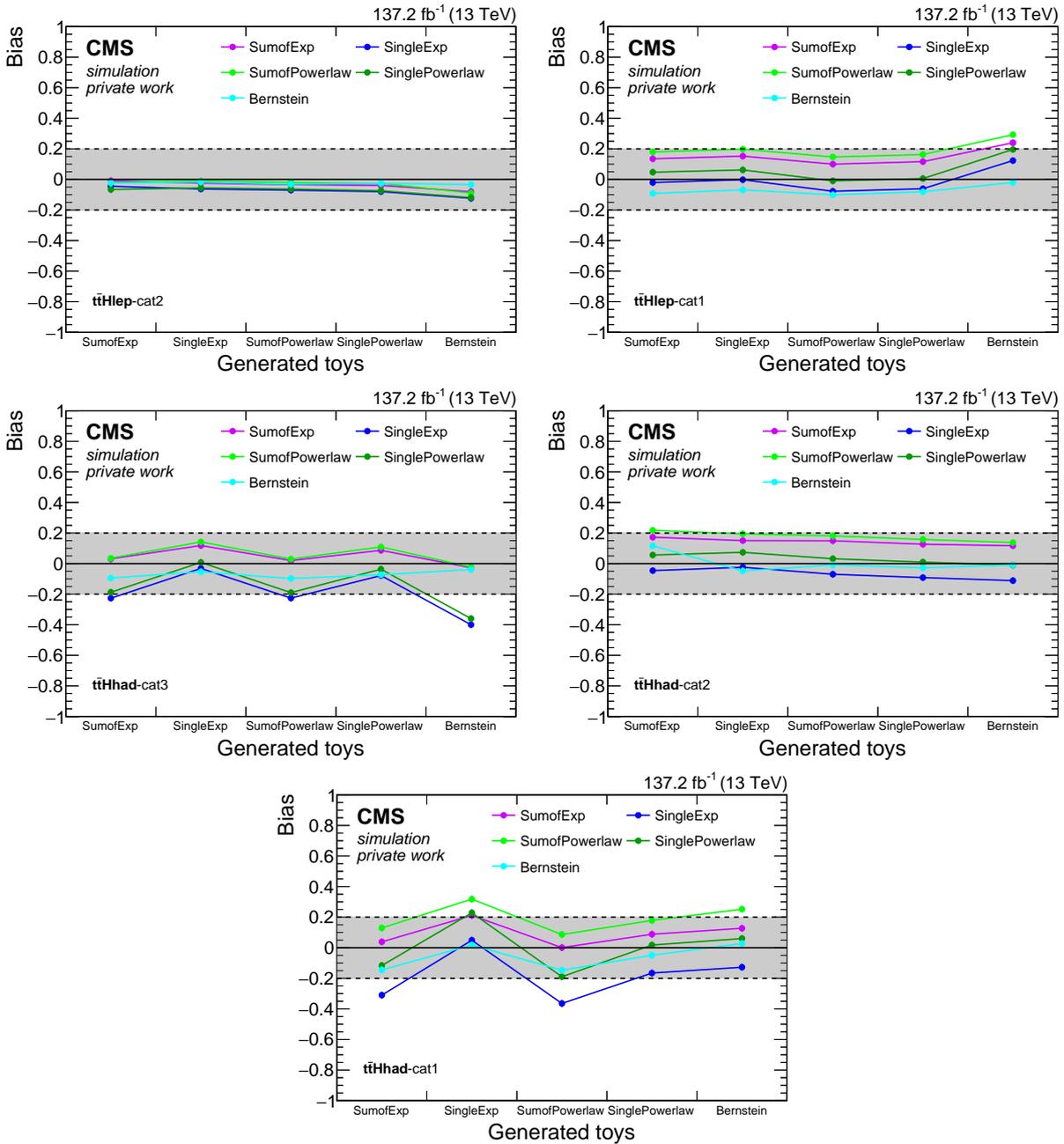


Figure 5.57: Results of the study to estimate the bias induced due to the choice of the background model in the final signal extraction fit. The study is performed using pseudodata in all $t\bar{t}H$ subcategories. The gray area indicates a bias of less than 20% compared to the statistical uncertainty.

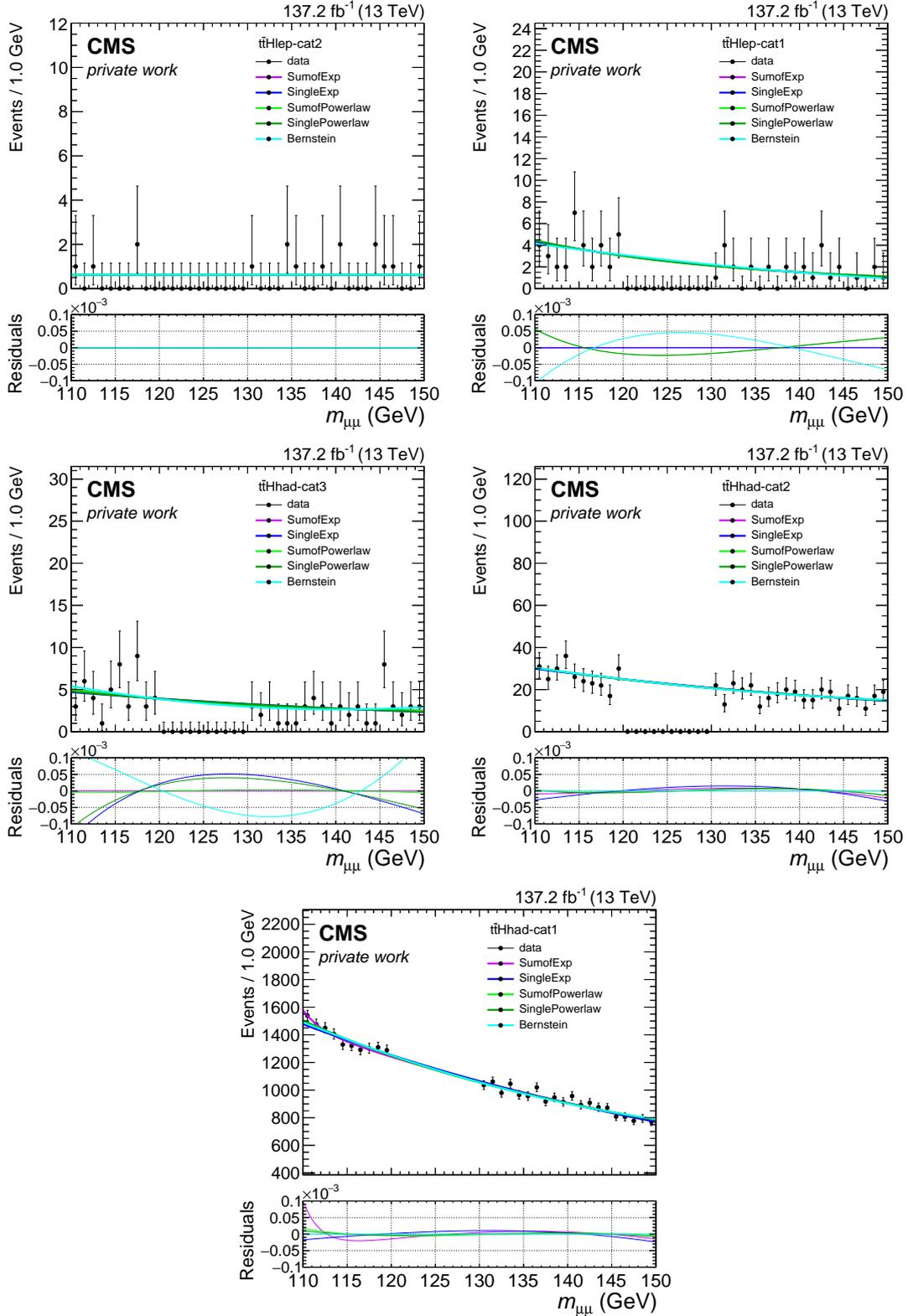


Figure 5.58: Distributions of the dimuon mass for events in the sideband region and the corresponding analytical background models for all $t\bar{t}H$ subcategories. In the lower panel, the normalized residual shape differences between the different background models and the chosen background function are presented.

5.11.2 Systematic uncertainties

Due to the small number of events selected in the most sensitive phase-space region, the statistical uncertainty in the dimuon mass signal region and in the sidebands is the limiting factor of this analysis. To obtain a reliable description of the total uncertainty on the final results, various theoretical and experimental sources of systematic uncertainty are taken into account as well. These systematic uncertainties can either have an effect on the signal rate or affect the signal shape in addition. All systematic uncertainties are considered as nuisance parameters in the fit. The different sources of systematic uncertainty considered in this analysis and their impact are discussed in the following. Note that neither theoretical nor experimental uncertainties are assigned to the estimated background component besides the statistical uncertainty of the sideband fit.

Shape uncertainties

The expected $m_{\mu\mu}$ shape is strongly correlated with the muon momentum scale and resolution. The corresponding uncertainties comprise variations of the Rochester corrections and the calibration of the dimuon mass resolution as well as an uncertainty associated with the geometrical track fit and the FSR photon recovery. The impact on the signal peak position is estimated to be 0.1% in all $t\bar{t}H$ subcategories. The uncertainty in the $m_{\mu\mu}$ resolution is estimated conservatively to be 10%. The signal peak position and the resolution are the only signal shape parameters allowed to vary in the fit. They are modelled by Gaussian priors with widths corresponding to the estimated uncertainty in the $m_{\mu\mu}$ scale (0.1%) and resolution (10%). Both, the uncertainty in the signal peak position and in the resolution are correlated across Higgs boson production modes and data-taking periods, but uncorrelated across categories.

Experimental and theoretical rate uncertainties

Experimental sources of systematic uncertainty can change the acceptance of events and thus the predicted rate of events in each category. To take differences between the efficiency in data and simulation into account, centrally provided *scale factors* (SFs) are used to correct the efficiencies in simulated events as described in Section 5.4. The systematic uncertainty in these corrections is propagated to the analysis by varying the corresponding SFs by one standard deviation, and evaluating the impact on the predicted signal rate. Theoretical uncertainties directly affecting the prediction of the Higgs boson signal are also taken into account. The different experimental and theoretical sources of systematic uncertainty are listed below:

- **Lepton efficiency** The lepton selection efficiency is subject to the uncertainty in the trigger efficiency, lepton identification and isolation efficiency, as well as the lepton momentum scale. All muon uncertainties are assumed to be fully correlated between muons, as well as across categories, Higgs boson production modes, and data-taking periods. The rate uncertainty due to variations of the muon identification and isolation efficiency translates into an overall effect of 0.5–1.5%. The impact of uncertainties corresponding to the multivariate lepton identification efficiency [175, 196] ranges from 0.2% to 0.5%. Since

the efficiency correction extracted on the 2017 data is also used to correct the signal prediction in 2018, the corresponding uncertainty is inflated by 100% for signal events corresponding to the 2018 data-taking period. The single muon trigger as well as the muon momentum scale and resolution uncertainties induce only a small effect of 0.1–0.8%.

Similar sources of uncertainty are taken into account for electrons. The dominant effect on the predicted signal rate results from the uncertainty in the reconstruction and identification of electrons. The corresponding values vary between 0.2–1%, depending on the category and the Higgs boson production process. The impact of the electron energy scale uncertainty is determined to be $\lesssim 0.1\%$. Migration effects between the $t\bar{t}H$ leptonic and the $t\bar{t}H$ hadronic categories due to the selection of a lepton in addition to the muon pair are studied and found to be $< 0.1\%$. The main reason for this is the tighter jet selection in the $t\bar{t}H$ hadronic category compared to the $t\bar{t}H$ leptonic category. The different uncertainties are correlated between electrons, across categories, Higgs boson production processes and data-taking periods.

- **Jet energy corrections** Jet energy uncertainties affecting the jet selection are accounted for by varying the *jet energy scale* and *resolution* (JES/JER) following the recommendations of the Jet/MET POG [163, 211]. Both, the JES and JER uncertainties depend on the p_T and η of the jets. The jet energies of all selected jets in the event are scaled up and down according to the different sources of the JES uncertainty. Then the entire event selection in each category is repeated and the BDT classifiers are evaluated again to account for changes in the jet kinematics and the \vec{p}_T^{miss} due to the shifted jet energies. The corresponding variations of the expected rate for each process are modeled by 21 nuisance parameters which are correlated across Higgs boson production modes and categories. Depending on the source of the uncertainty they are either correlated or uncorrelated across data-taking periods. In addition, the nominal value of the JER correction is used to define a one-sided systematic variation, which is correlated across categories and Higgs boson production modes, but uncorrelated across data-taking periods. This choice is motivated by the selected phase space used to determine the JER smearing factors. They are measured for high- p_T jets, while the phase space explored by the $H \rightarrow \mu^+\mu^-$ analysis is dominated by jets with p_T below 100 GeV. Since the outputs of the $t\bar{t}H$ BDT classifiers are correlated with the number of selected jets in the event, JES uncertainties cause migration of events between categories. The total uncertainty in the $t\bar{t}H$ signal rate due to variations of the JES is found to be between 1–6% depending on the category. The total uncertainty due to the jet energy resolution has a slightly smaller impact of 0.2–4% per event category.
- **Identifying b jets** The efficiency of the b jet identification is affected mainly by the systematic uncertainty of the JES and the purity of heavy- or light-flavour jets. This uncertainty is taken into account in terms of eight uncorrelated nuisance parameters and one nuisance parameter correlated with the JES uncertainty. The overall impact is estimated to be 1–3% per event category.

- **Quark-gluon likelihood (QGL) reweighting** The reconstruction of the QGL is described in Section 4.3.7. A systematic uncertainty is assigned to the reweighting of the QGL distribution for each jet as explained in Section 4.3.7. The uncertainty is taken as the full correction difference following the CMS prescription [183]. The down variation corresponds to the QGL distribution without reweighting and the up variation corresponds to the QGL distribution applying twice the QGL reweighting. The impact is determined to be 0.5–2% per event category.
- **Pile up (PU)** The uncertainty in the measured minimum bias cross section can affect the PU reweighting method. The minimum bias cross section is varied by $\pm 4.6\%$ to create alternative PU profiles, which are used to repeat the PU reweighting. The effect on the signal yield is found to be 0.2–1% depending on the category.
- **Luminosity** The uncertainty in the luminosity measurement of the 2016, 2017, and 2018 data-taking periods [74–76] are individually estimated to be 2.3–2.5%, while the uncertainty in the total integrated luminosity is 1.8%.
- **ECAL prefireing** The prefireing inefficiency is emulated in simulated events by a reweighting using precomputed prefireing probability maps as a function of the jet η and p_T . The corresponding uncertainty is obtained from a variation of these probabilities by $\pm 20\%$, which results in an overall effect of about 0.3–0.5% per category.
- **Signal cross section & branching fraction** The $t\bar{t}H$ signal cross section is calculated at NLO accuracy with a total uncertainty of $+7\%/ - 10\%$ including the effect of QCD scale and PDF variations. An uncertainty of 1.2% is assigned to the Higgs boson decay rate to muons [120, 212, 213].
- **Underlying event tune** The description of non-perturbative QCD effects in PYTHIA is improved using measurements of the underlying event. The corresponding parameters in simulated events are varied within their uncertainties, leading to an impact on the signal rate of 1–2%.
- **ISR/FSR PS scale** The PS scale value of initial-state and final-state radiation in PYTHIA is independently scaled up by a factor 2 and down by a factor 0.5, affecting the emission of additional jets. The impact on the signal yield is 0.3–1%.
- **Renormalization and factorization scales** An uncertainty is assigned due to the choice of the renormalization scale μ_r and the factorization scale μ_f in simulated events. Both scales are varied by a factor of 0.5 and 2 with respect to the nominal value. Variations where both are scaled in the same direction are omitted. This leads to a set of six configurations in addition to the nominal one. The total uncertainty corresponds to the envelope of the variations ranging from 0.5% to 1.5%.
- **Parton distribution function** The last source of uncertainty corresponds to the choice of the PDF used for event generation. The impact is estimated by considering 100 PDF replicas of the NNPDF3.0/3.1 set [118]. The variance of all 100 variations is taken as the total PDF uncertainty resulting in an overall impact on the signal rate of 0.1–0.5% per event category.

5.12 Results for the $t\bar{t}H$ channel

The presence of a $H \rightarrow \mu^+\mu^-$ signal in events where a Higgs boson is produced in association with a top quark-antiquark pair is quantified by a statistical procedure using the COMBINE software package [214]. A simultaneous *maximum-likelihood* (ML) fit is performed to the dimuon mass distribution over the range $110 < m_{\mu\mu} < 150\text{GeV}$ in all five $t\bar{t}H$ event categories based on the $t\bar{t}H$ leptonic and $t\bar{t}H$ hadronic BDT. The corresponding BDT output distributions for all events in the signal fit region are shown in Fig. 5.59. The $H \rightarrow \mu^+\mu^-$ signal expected from all Higgs boson production modes is scaled by a common signal strength modifier $\hat{\mu}$ which is free to float in the fit. The signal strength modifier is defined as the ratio between the observed Higgs boson rate and its SM expectation:

$$\hat{\mu} = \frac{(\sigma \cdot \mathcal{B}(H \rightarrow \mu^+\mu^-))_{\text{obs}}}{(\sigma \cdot \mathcal{B}(H \rightarrow \mu^+\mu^-))_{\text{SM}}}. \quad (5.19)$$

The corresponding confidence intervals on the signal strength are evaluated using a profile likelihood ratio test statistic [215]. Systematic uncertainties discussed in Section 5.11.2 are incorporated into the analysis via nuisance parameters following a modified frequentist approach [216]. A log-normal prior distribution is used for uncertainties only affecting the rate, while shape uncertainties are modeled by Gaussian priors. Binned fits of the dimuon mass distribution are employed to improve the speed of the computation. A bin size of 50 MeV is chosen, which is over one magnitude smaller than the signal peak resolution. Figure 5.60 shows the $m_{\mu\mu}$ distributions, and the results of the S+B fit in the $t\bar{t}H$ leptonic and $t\bar{t}H$ hadronic subcategories using the analytical models described in Section 5.11.1.

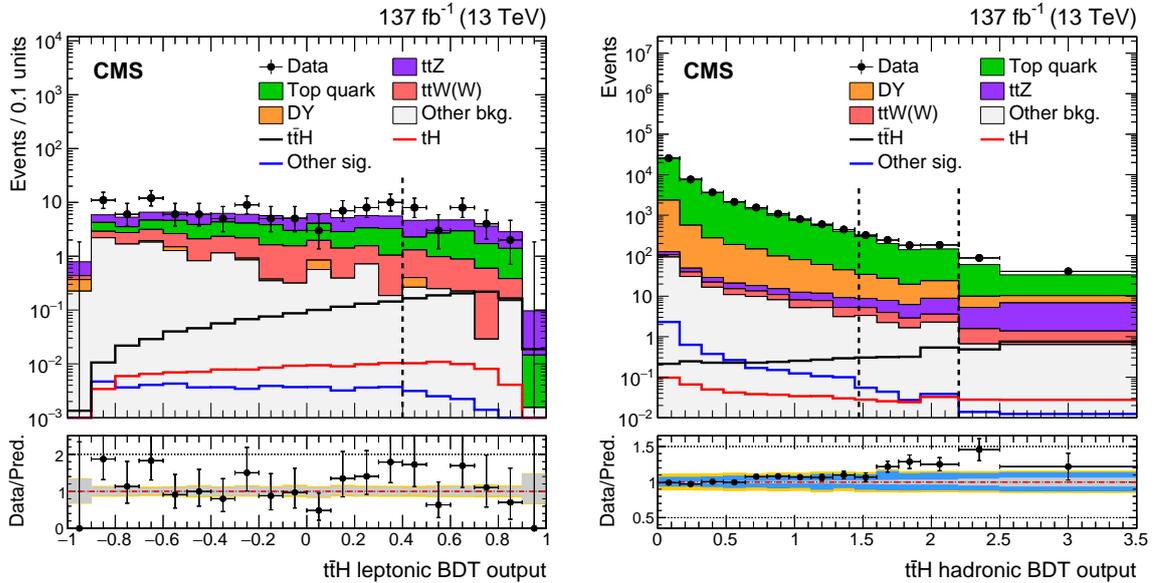


Figure 5.59: The BDT output distribution in the $t\bar{t}H$ leptonic (left) and hadronic (right) categories for events with $110 < m_{\mu\mu} < 150\text{GeV}$. The signal distributions expected from different production modes of the Higgs boson with $m_H = 125\text{GeV}$ are overlaid. The dashed vertical lines indicate the boundaries of the optimized event categories. The description of uncertainties is the same as in Fig. 5.51. Published also in Ref. [172].

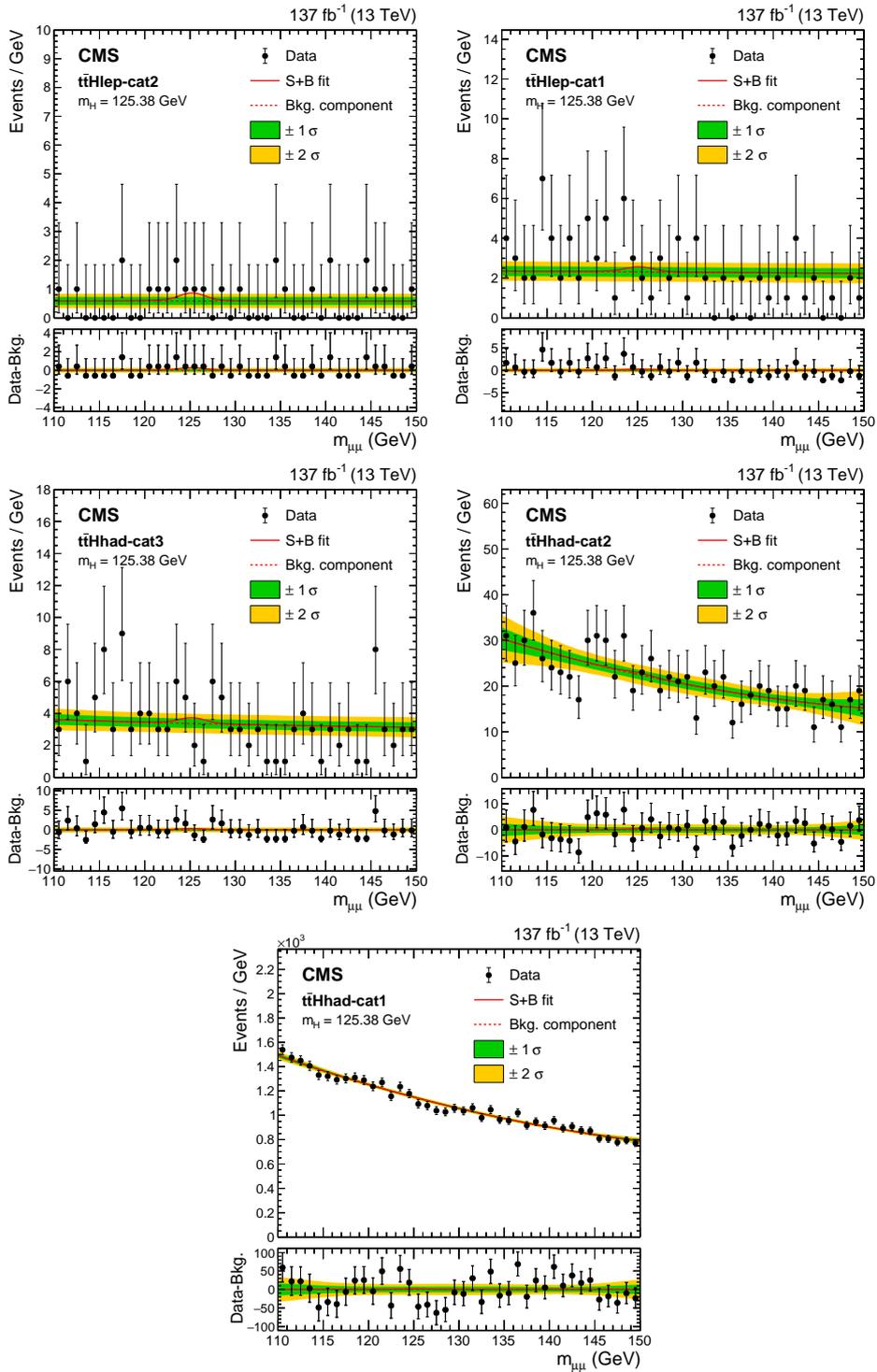


Figure 5.60: Comparison between the data and the total background extracted from a S+B fit performed across the various $t\bar{t}H$ hadronic and $t\bar{t}H$ leptonic event subcategories. The one (green) and two (yellow) standard deviation bands include the uncertainty in the background component of the fit. The lower panel shows the residuals after the background subtraction, where the red line indicates the signal with $m_H = 125.38 \text{ GeV}$ extracted from the fit. Published also in Ref. [172].

The total expected signal yield (S), the estimated background yield (B) as well as the observed number of events are presented along with the signal purity $S/(S+B)$ and the sensitivity S/\sqrt{B} for all $t\bar{t}H$ subcategories in Table 5.14.

Table 5.14: Summary of the total expected signal yield, the HWHM of the signal peak, the used background model, the estimated number of background events and the observed number of events, and the $S/(S+B)$ and S/\sqrt{B} ratios computed within the HWHM of the signal peak for each of the $t\bar{t}H$ subcategories. Modified version published in Ref. [172].

Event category	Total signal	HWHM (GeV)	Bkg. fit function	Bkg. @HWHM	Data @HWHM	$S/(S+B)$ (%) @HWHM	S/\sqrt{B} @HWHM
$t\bar{t}H_{lep}$ -cat2	0.99	1.75	SingleExp	2.08	4	24.5	0.47
$t\bar{t}H_{lep}$ -cat1	1.06	1.92	SingleExp	9.00	13	7.09	0.22
$t\bar{t}H_{had}$ -cat3	1.33	1.80	SumofExp	12.3	12	6.87	0.26
$t\bar{t}H_{had}$ -cat2	1.62	1.81	Bernstein	82.0	89	1.32	0.12
$t\bar{t}H_{had}$ -cat1	6.87	1.85	Bernstein	4298	4251	1.07	0.07

Signal strength The ML fit is performed in all five $t\bar{t}H$ subcategories combined and, in addition, separately for each of the subcategories. The resulting best-fit values of the signal strength are presented in Fig. 5.61. The best-fit signal strength from the combined ML fit is:

$$\hat{\mu}_{t\bar{t}H(H \rightarrow \mu^+\mu^-)} = 2.32_{-1.95}^{+2.27} \text{ (stat + syst),}$$

where the total uncertainty is dominated by the statistical component. This result is compatible with the SM expectation ($\hat{\mu} = 1$) for $t\bar{t}H$ ($H \rightarrow \mu^+\mu^-$) within 1σ .

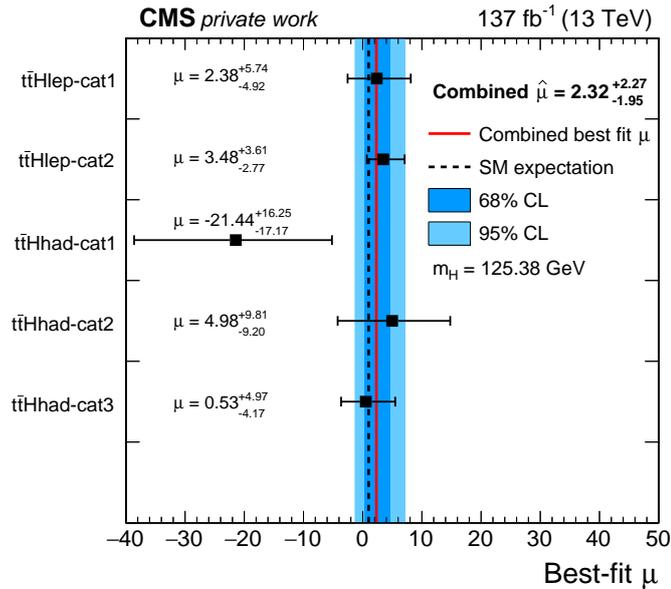


Figure 5.61: The measured signal strength modifiers for $m_H = 125.38$ GeV in each $t\bar{t}H$ subcategory (black points) are compared to the result of the combined $t\bar{t}H$ fit (solid red line) and to the SM expectation (dashed line).

Significance A small excess of events over the SM background prediction is observed in data. The observed (expected) significance for a Higgs boson with $m_H = 125.38 \text{ GeV}$ of the incompatibility with the background-only hypothesis is 1.2σ (0.5σ). The expected significance is determined using an Asimov dataset [215].

Exclusion limits Since the measured signal strength also allows values smaller than one, a 95% *confidence level* (CL) upper limit is set on the signal strength using the asymptotic CL_s method [215, 217, 218]. The observed 95% CL upper limit on the signal strength derived from the combined fit is 6.5, with an expected upper limit in the absence (presence) of a SM Higgs boson of 4.2 (4.9). The current dataset, therefore, allows the rejection of new physics models which predict a signal rate about four times larger than in the SM. The observed and expected upper limits in each subcategory as well as the combined limit are illustrated in Fig. 5.62.

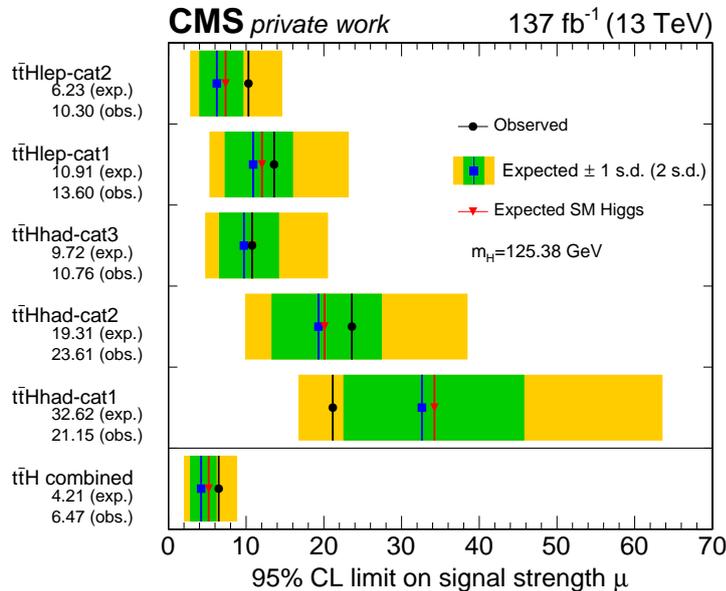


Figure 5.62: Expected (blue) and observed (black) 95% CL upper limits on $\hat{\mu}$. The expected limits are displayed together with their 68% and 95% CL intervals, and with the expectation for an injected SM Higgs signal (red).

The analysis sensitivity is limited by the statistical precision of the recorded dataset. Thus, more data will help to improve these results. To increase the sensitivity to the $H \rightarrow \mu^+\mu^-$ signal given the existing dataset, the $t\bar{t}H$ analysis is combined with three analyses dedicated to the ggH , VBF and VH production modes as described in Chapter 6. Event displays of potential $t\bar{t}H$ signal events where a Higgs boson decays to a pair of muons are shown in Fig. 5.63.

Comparison to ATLAS sensitivity Being the first CMS analysis which targets the $t\bar{t}H$ production in the Higgs boson decay channel to muons, the results presented in this thesis set the first benchmark. However, the $H \rightarrow \mu^+\mu^-$ analysis performed by ATLAS also includes a dedicated $t\bar{t}H$ category, which can be used for a comparison based on the estimated significance S/\sqrt{B} . The $t\bar{t}H$ category employed by ATLAS focuses on events with leptonic decays of the top quarks, while events where both top quarks decay hadronically are not taken into account. The corresponding estimated significance by ATLAS is 0.36. A similar phase-space region is selected by the $t\bar{t}H$ leptonic category in this thesis with an estimated significance of 0.47 for $t\bar{t}H_{lep-cat2}$ and 0.22 for $t\bar{t}H_{lep-cat1}$. The combined $t\bar{t}H$ leptonic significance of both subcategories is about 30% stronger than the ATLAS result. Furthermore, the $t\bar{t}H$ combination which includes the $t\bar{t}H$ hadronic category is about 40% more sensitive compared to ATLAS. Therefore, the presented thesis provides the most sensitive analysis on Higgs boson decays to muons in the associated production with top quarks to date. The reasons for these differences are discussed in Section 6.5.

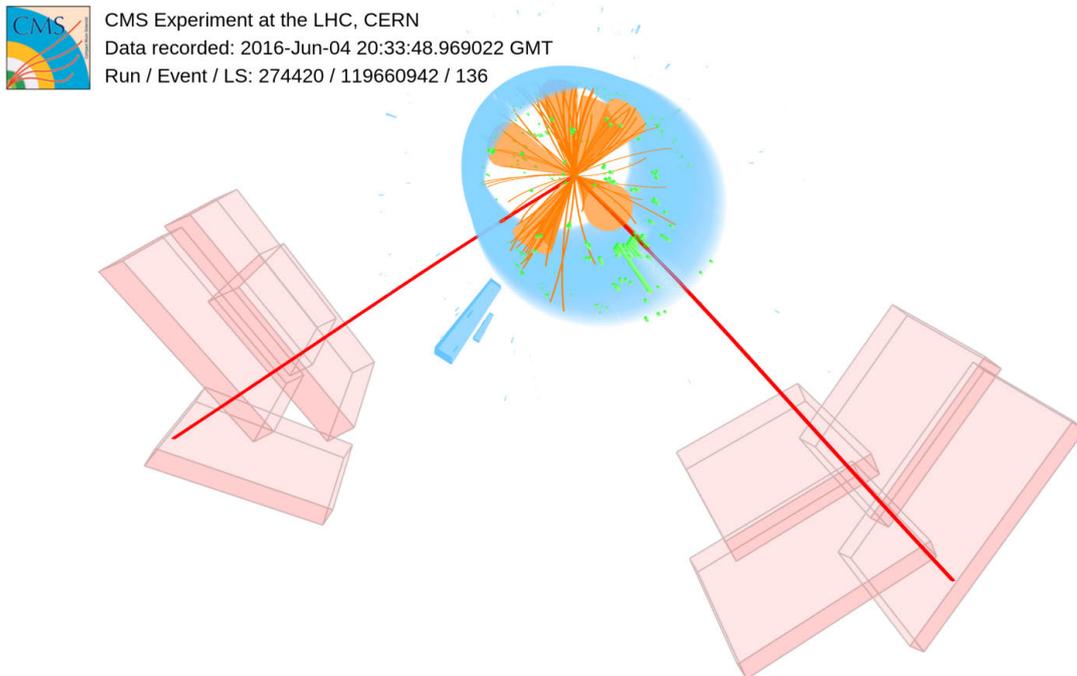


Figure 5.63: Measured event of a candidate for a Higgs boson produced in association with a top quark-antiquark pair where the Higgs boson decays into two muons, indicated by the red lines. In this event, both top quarks decay hadronically. The total of six jets are represented by the orange cones. The dimuon invariant mass of the Higgs candidate is 125.40 ± 1.24 GeV. Published as part of the supplementary material of Ref. [172].

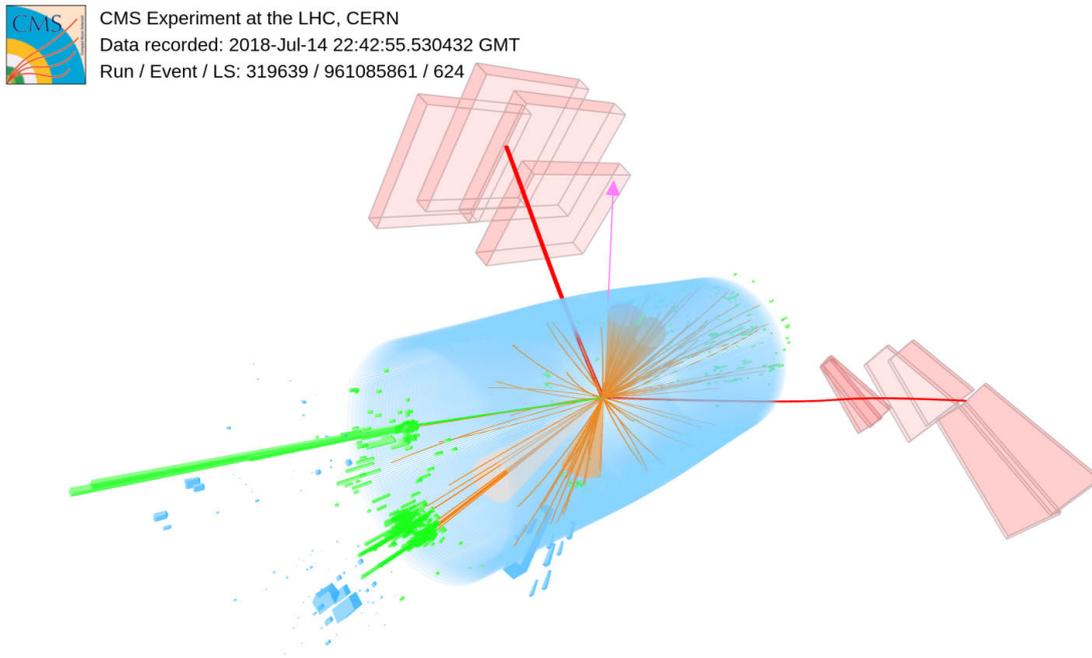


Figure 5.64: Measured event of a candidate for a Higgs boson produced in association with a top quark-antiquark pair where the Higgs boson decays into two muons, represented by the red lines. One top quark decays hadronically into three jets, indicated by the orange cones. The other top quark decays leptonically into a jet, an electron (indicated by the green line), and a neutrino (indicated by the purple arrow). The reconstructed mass of the Higgs boson candidate is 125.30 ± 1.22 GeV. Published as part of the supplementary material of Ref. [172].

Combined analysis of Higgs boson decays to muons

The analysis of Higgs boson decays to muons produced in association with a pair of top quarks described in Chapter 5 is part of a combined result based on four complementary analyses which address the ggH, VBF, VH, and t̄tH production modes. The results of this combination are published in Ref. [172], and are presented in this chapter.

For completeness, the main features of the analyses targeting the ggH, VBF and VH production are briefly described in Section 6.1, followed by a summary of the combined $H \rightarrow \mu^+\mu^-$ results and the scientific impact in Sections 6.2 and 6.3. The new analysis is compared to a previous CMS analysis, and a recent ATLAS analysis in Sections 6.4 and 6.5. Finally, this chapter concludes with an outlook on future prospects for the $H \rightarrow \mu^+\mu^-$ analysis in Section 6.6. The description of the individual Higgs boson production categories, and the combined results in Sections 6.1 and 6.2 are largely based on information from Ref. [172] unless stated otherwise.

6.1 Higgs boson production categories

Events are separated into independent event categories based on the final state expected from the Higgs boson production modes ggH, VBF, VH, and t̄tH. To avoid overlap between the selected events, the event categories are fully separated by the number of b-tagged jets, the number of leptons, and the dijet kinematics. An overview of the event selection is given in Table 6.1.

Table 6.1: Summary of the kinematic requirements used to define mutually exclusive Higgs boson production categories. The categorization is based on the number of b-tagged jets passing the loose (L) or medium (M) working point of the DeepCSV algorithm, the number of leptons, and the dijet kinematics.

Observable	t̄tH cat.		VH cat.		VBF cat.	ggH cat.
	t̄tH had.	t̄tH lep.	WH	ZH		
b-tagged jets	≥2(L) or 1(M)		≤1(L) and 0(M)		≤1(L) and 0(M)	≤1(L) and 0(M)
Selected leptons	=2	=3,4	=3	=4	=2	=2
Jet multiplicity	≥3	≥2	—	—	≥2	0,1 (≥2)
Dijet kinematics	—	—	—	—	$m_{jj} \geq 400 \text{ GeV}$ $ \Delta\eta_{jj} \geq 2.5$	$m_{jj} < 400 \text{ GeV}$ or $ \Delta\eta_{jj} < 2.5$ or $\geq 25 \text{ GeV}$
Leading jet p_T	≥50 GeV	≥35 GeV	—	—	≥35 GeV	—
$ m_{\mu\mu(ee)} - m_Z $	—	> 10 GeV	> 10 GeV	< 10(20) GeV	—	—
Dilepton mass $m_{\ell\ell}$	—	> 12 GeV	> 12 GeV	> 12 GeV	—	—
Jet triplet mass	100 – 300 GeV	—	—	—	—	—

The lepton selection criteria in the ggH, VBF, and VH categories differ slightly from the ones used in the $t\bar{t}H$ event category and described in Sections 4.3.2 and 4.3.3. In the ggH, VBF, and VH categories the lepton isolation cone does not depend on the lepton p_T , instead it is defined with constant cone radius $R = 0.4$ around the muon track. Muons are required to pass the medium ID and to have a relative isolation of less than 25% of the muon p_T . However, this does not affect the categorization of $t\bar{t}H$ signal events, since these are separated by the presence of b-tagged jets which are selected in the same way in all categories as explained in Section 4.3.5.

6.1.1 The VBF production category

Events with two muons and at least two jets are selected. The jets are required to have $|\eta|$ less than 4.7 and a p_T larger than 35(25) GeV for the leading (sub-leading) jet, where jets are ordered by p_T . Since VBF jets are expected to be radiated in forward direction, the system of the two leading jets is required to have $m_{jj} > 400$ GeV and $|\Delta\eta_{jj}| > 2.5$.

A DNN is trained using kinematic observables which distinguish VBF signal events generated with $m_H = 125$ GeV from the main backgrounds Drell-Yan, Z-EWK, dileptonic $t\bar{t}$, and diboson production VV. The implementation of the DNN is based on the KERAS framework [199] with TENSORFLOW [200] as backend. The DNN architecture comprises four intermediate networks. Two networks aim to distinguish VBF signal events from the dominant background processes Z-EWK or Drell-Yan. The remaining two networks are used to discriminate VBF signal events from the total expected background either based on all input observables except for the dimuon mass, or based only on the dimuon mass and its resolution. Each of these networks contains three or four hidden layers where the last hidden layers are combined to provide a single classifier output. Alternatively, a BDT was trained using the same set of input observables which led to a slightly worse performance. For this reason, the DNN was chosen as the final discriminant.

The input observables included in the DNN aim to describe the kinematic properties of the dimuon system, as well as the dijet system of both leading jets. The most sensitive observable is the invariant mass of the muon pair, and the corresponding per-event mass resolution. Additional input observables are the p_T and the rapidity y of the dimuon system, the angular distributions ϕ_{CS} , and $\cos\theta_{CS}$, as well as both full momentum vectors of the two highest p_T jets, their mass m_{jj} , and angular separation $\Delta\eta_{jj}$. In particular, correlations between the muon pair and the VBF jets are exploited. These are the minimum separation in η between the dimuon system and each of the two highest p_T jets, the Zeppenfeld observable [219] and the p_T -balance [172]. Since VBF jets originate from quarks, the QGL of both VBF jets is used to distinguish them from gluon initiated jets. Moreover, the quarks in VBF events are well separated. Thus, the hadronic activity in the rapidity region between these quarks is suppressed. This information is included in the BDT in terms of the number and the scalar p_T sum of all jets reconstructed using soft tracks.

The signal is extracted from a binned *maximum-likelihood* (ML) fit to the output of the DNN where the various background processes are estimated from simulation. The fit is performed simultaneously in a signal region and a control region for each data-taking period. Events with invariant dimuon mass between 115 and 135 GeV correspond to the signal region (VBF-SR), while events with a dimuon mass in $110 < m_{\mu\mu} < 115$ GeV or $135 < m_{\mu\mu} < 150$ GeV correspond to the sideband control region (VBF-SB), which is used to validate and constrain the uncertainty in the background estimation. Due to variations in the detector response of forward jets between the years, the fit is performed independently for each data-taking period. A similar strategy was used to measure the cross section of the electroweak production of two jets in association with a Z boson Z-EWK [220] which is the main background contribution in the VBF event category. Therefore, the background is known to be well modeled within uncertainties in simulated events. Compared to a sideband analysis where the signal is extracted by fitting an analytical function to the observed $m_{\mu\mu}$ distribution, this strategy leads to an improved expected significance by about 20% in the VBF category, and about 10% in the combined $H \rightarrow \mu^+\mu^-$ result.

The influence of experimental and theoretical systematic uncertainties on the DNN output is nontrivial. Systematic variations affect in most cases not only the expected rate, but also the shape of the signal and background templates. The main sources of systematic uncertainty are listed in the following. A comprehensive overview of all uncertainties and their correlation is given in Table B.1.

- **Parton shower modeling** The VBF signal and the dominant background from Z-EWK are very sensitive to the PS model [220, 221]. The impact on the normalization and the shape of the DNN templates is estimated by varying the parameters that determine the properties of ISR and FSR jets produced by PYTHIA. In addition, a conservative PS uncertainty is assigned to account for differences between PS algorithms. The full symmetrized difference between the prediction using PYTHIA with dipole shower mode and HERWIG with angular-ordered showers is assigned as uncertainty in each DNN bin. The corresponding uncertainty varies as a function of the DNN score between 2 to 10%.
- **Drell-Yan background** The normalization of Drell-Yan events with less than two generator matched jets is floating in the fit and constrained by experimental data. Due to variations in the detector response between the years, these normalization parameters are uncorrelated across the years.
- **JES+JER uncertainty** Forward jets are particularly affected by the uncertainty in the measurement of the JES and JER. The p_T of each jet is varied according to a set of partially correlated uncertainty sources for both, the VBF signal and the background.
- **Statistical precision of event samples** The statistical precision of the simulated samples is taken into account by allowing each bin of the total background template to vary within the corresponding statistical uncertainty. These bin-by-bin uncertainties are uncorrelated across the DNN templates.

Figure 6.1 presents the observed and predicted post-fit distribution of the DNN discriminant in the VBF-SB and VBF-SR region. The post-fit distribution of the background precisely reproduces the observed distribution in data. In the VBF-SR a small excess of data over the SM background expectation is observed, which is compatible with the prediction of a SM Higgs boson decay to muons. The best-fit value of the signal strength for a Higgs boson with $m_H = 125.38 \text{ GeV}$ is

$$\hat{\mu}_{\text{VBF}(H \rightarrow \mu^+\mu^-)} = 1.36_{-0.61}^{+0.69} \text{ (stat + syst).}$$

The observed (expected) significance of the incompatibility with the background-only hypothesis is 2.4σ (1.8σ). The VBF analysis provides the most sensitive analysis channel among the Higgs boson production categories used in the combination.

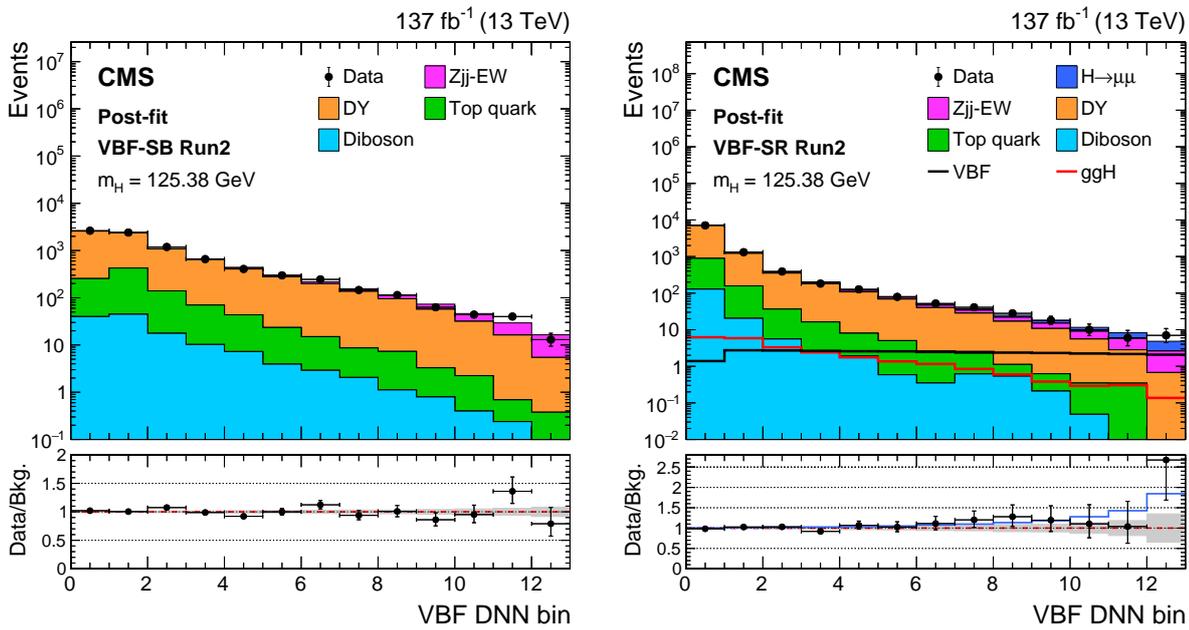


Figure 6.1: The DNN output distribution in the sideband VBF-SB (left) and the signal region VBF-SR (right) for all three data-taking periods combined. The observed distribution in data is compared to the post-fit background prediction. Their ratio is given in the lower panel including the total background uncertainty (gray band). The ggH (solid red) and VBF (solid black) Higgs boson signals are overlaid. The best-fit $H \rightarrow \mu^+\mu^-$ signal contribution for $m_H = 125.38 \text{ GeV}$ is presented in the upper panel (blue histogram), as well as in the lower panel (blue solid line). Taken from Ref. [172].

To visualize the results, the $m_{\mu\mu}$ distribution for the $S/(S+B)$ weighted combination of all events in the VBF-SR and VBF-SB regions is presented in Fig. 6.2. The $S/(S+B)$ weights are determined in bins of a mass-decorrelated DNN which is obtained by replacing $m_{\mu\mu}$ with a fixed value of 125 GeV when the DNN is evaluated. S and B are the number of expected signal and background events with $m_{\mu\mu}$ within $\pm \text{HWHM}$ of the expected signal peak with $m_H = 125.38 \text{ GeV}$. The best-fit values for the nuisance parameters and signal strength are propagated to the $m_{\mu\mu}$ distribution. This distribution is not used for any of the measurements.

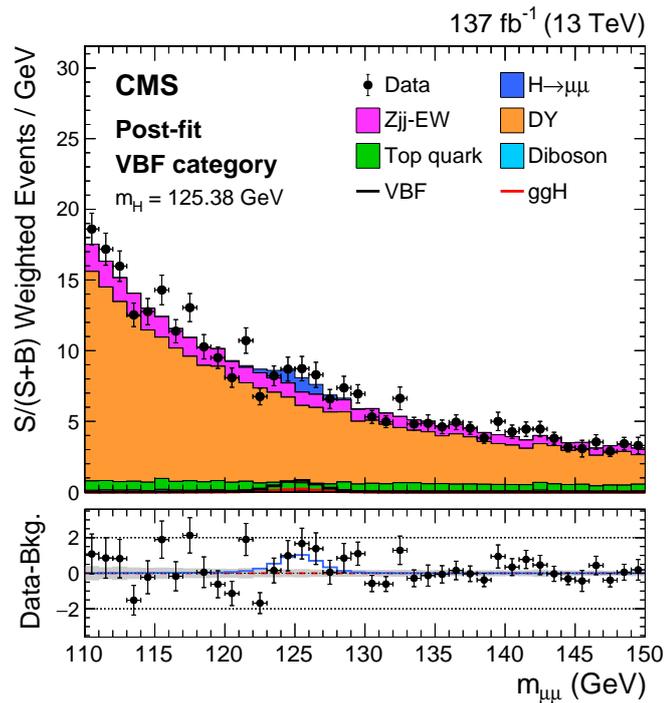


Figure 6.2: The $m_{\mu\mu}$ distribution for the $S/(S+B)$ weighted combination of events in the side-band VBF-SB and the signal region VBF-SR. The best-fit $H \rightarrow \mu^+\mu^-$ signal contribution is indicated in the upper panel by the blue histogram and in the lower panel by the blue line. The lower panel shows the residuals after subtracting the background prediction from the $S+B$ fit. The gray band corresponds to the total background uncertainty from the background-only fit. Taken from Ref. [172].

6.1.2 The ggH production category

The ggH event category collects all dimuon events which are not selected by the $t\bar{t}H$, VH, and VBF production categories. These events contain exactly two oppositely charged muons, no additional leptons and at most one b-tagged jet passing the loose working point of the DeepCSV algorithm. To avoid overlap of the selected phase-space between the ggH and the VBF category, events containing two or more jets with $p_T > 25$ GeV are only considered if the leading jet has $p_T < 35$ GeV, the invariant mass of the two highest p_T jets is smaller than 400 GeV, or the absolute angular separation $\Delta\eta_{jj}$ is smaller than 2.5.

Similar to the strategy described in Chapter 5, events are classified into subcategories using a multivariate discriminant based on a BDT, which is trained separately for the three data-taking periods using the TMVA software package [197, 198]. The signal simulation considered in the training comprise the ggH, VBF, VH, and $t\bar{t}H$ processes generated at $m_H = 125$ GeV. Several kinematic features are used to discriminate signal events from the Drell-Yan, $t\bar{t}$, single top quark, diboson, and Z-EWK background processes.

The p_T and the rapidity y of the dimuon system, as well as the angular distributions ϕ_{CS} , and $\cos\theta_{CS}$ distinguish ggH signal events from the dominant Drell-Yan background. The muon η and the ratio of the muon p_T to $m_{\mu\mu}$ are sensitive to the dimuon mass resolution. Since gluons are more likely to initiate additional radiation compared to quarks, different jet related observables help to further reduce the Drell-Yan background. These are the p_T and η of the highest p_T jet in the event and the absolute distance in η and ϕ between the jet and the muon pair, as well as the total number of jets in the event. The residual signal component from VBF is identified using the m_{jj} , $\Delta\eta_{jj}$, and $\Delta\phi_{jj}$ of the two highest p_T jets, as well as the Zeppenfeld observable and the angular separation in η and ϕ between the dimuon system and each of the two jets. All these observables are used as input to the BDT. The resulting BDT output distribution is shown in Fig. 6.3. Five event subcategories are defined based on the output of the BDT discriminants for each data-taking period. The subcategory boundaries are determined using the iterative algorithm described in Section 5.10. Events in the highest BDT score category are characterized by a large p_T of the dimuon system, and one high- p_T jet detected in the endcaps. The excess in data observed in the high BDT score region originates from the modeling of forward jets in 2017, and 2018. Since the background is estimated using the dimuon mass sidebands in data, and the discrepancy between data and simulation is nearly fully covered by the total systematic uncertainty, the impact on the final result is negligible.

To extract the $H \rightarrow \mu^+\mu^-$ signal, a simultaneous ML fit is performed to the observed $m_{\mu\mu}$ distribution as described in Sections 5.11 and 5.12. However, in contrast to the $t\bar{t}H$ category, which uses independent functions to model the $m_{\mu\mu}$ background shape in each subcategory, the ggH category employs a correlated background model. This approach is referred to as *core-pdf method*. Since the background is strongly dominated by the Drell-Yan process and the background composition is rather stable along the BDT output, the $m_{\mu\mu}$ background shape is expected to be very similar in the ggH subcategories. The core background shape is defined as the envelope of a modified Breit-Wigner, a sum of two exponentials, and the product of a

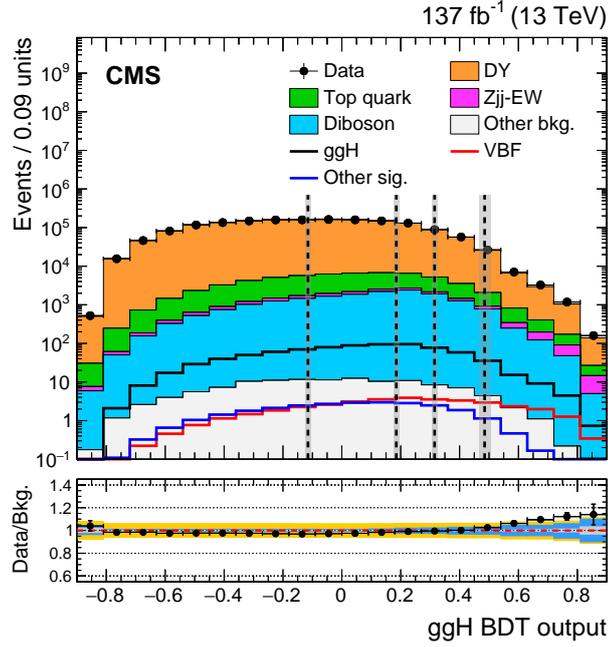


Figure 6.3: The BDT output distribution in the ggH category for events in the signal fit region with $110 < m_{\mu\mu} < 150$ GeV. The signal distributions expected from different production modes of the Higgs boson with $m_H = 125$ GeV are overlaid. The gray vertical bands indicate the boundary range of the BDT output values used to define subcategories for the different data-taking periods. In the lower panel, the uncertainties are shown as in Fig. 5.51. Taken from Ref. [172].

nonanalytical shape derived from the FEWZ v3.1 generator [131] and a third-order Bernstein polynomial. To account for differences in the slope of the $m_{\mu\mu}$ spectrum across the subcategories, the core shape is modified using a third- or a second-order polynomial with parameters uncorrelated across event categories. A discrete profiling method [222] is employed to treat the choice of the background function as a discrete nuisance parameter in the likelihood function used to extract the result. The theoretical and experimental sources of systematic uncertainty affecting the expected signal rate in each ggH subcategories are similar to those described in Section 5.11.2. A comprehensive overview of all uncertainties and their correlation is presented in Table B.1.

Figure 6.4 shows the $m_{\mu\mu}$ distribution for a $S/(S+B)$ weighted combination of all five ggH subcategories, and the corresponding $S+B$ fits. Events are weighted proportional to the $S/(S+B)$ ratio of the subcategory to which they belong. S and B are the number of expected signal and background events with mass within $\pm\text{HWHM}$ of the expected signal peak. A slight excess of data is observed at $m_H = 125.38$ GeV, which is compatible with the hypothesis of a SM Higgs boson decaying to a pair of muons, as well as the B-only hypothesis. The best-fit value of the signal strength is

$$\hat{\mu}_{\text{ggH}(H \rightarrow \mu^+\mu^-)} = 0.63_{-0.64}^{+0.65} \text{ (stat + syst).}$$

The observed (expected) significance for a Higgs boson with $m_H = 125.38 \text{ GeV}$ of the incompatibility with the background-only hypothesis is 1.0σ (1.6σ). Due to the largest signal rate of all considered Higgs boson production modes, the ggH event category is the second most sensitive analysis channel in the combination and determines together with the VBF channel the overall analysis sensitivity.

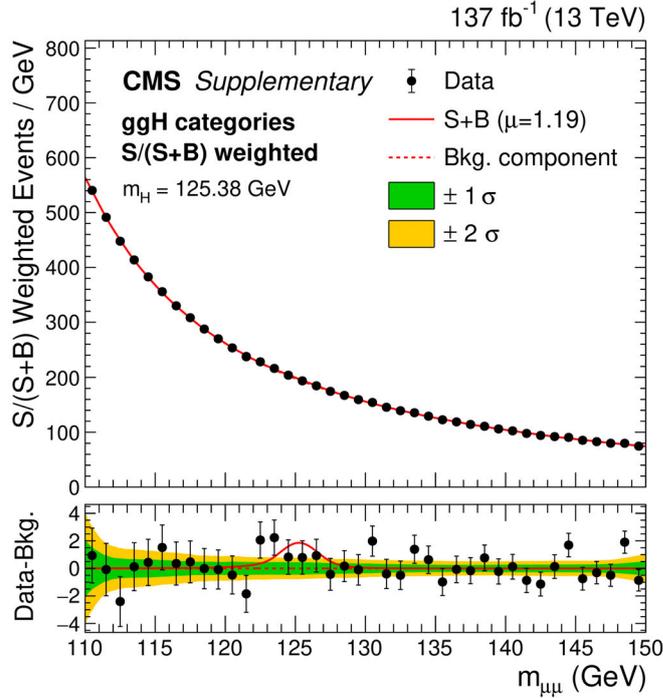


Figure 6.4: The $m_{\mu\mu}$ distribution for the $S/(S+B)$ weighted combination of ggH event categories. The solid red line represents the best-fit $H \rightarrow \mu^+\mu^-$ signal, while the red dashed line corresponds to the background component. The background uncertainty is indicated by the green and yellow bands. In the lower panel, the residuals after background subtraction are shown. Supplementary material of Ref. [172].

6.1.3 The VH production category

The VH category aims to select events where a W boson decays leptonically to a charged lepton and a neutrino, or a Z boson decays to a pair of oppositely charged leptons. Therefore, events with two oppositely charged muons and one (two) additional charged lepton(s) are assigned to the WH (ZH) category. Events are restricted to no more than one b-tagged jet passing the loose working point of the DeepCSV algorithm, which ensures a nonoverlapping phase-space selection between the VH category and the $t\bar{t}H$ leptonic category. Similar to the lepton selection in the $t\bar{t}H$ category, the LEPTONMVA approach is used to suppress non-prompt lepton background.

For events with three selected muons, at least one of the two oppositely charged muon pairs is required to have a dimuon mass in the range between 110–150 GeV. To reduce the combinatorial background from quarkonium or Z boson decays in the WH category, events with dilepton masses $m_{\ell\ell} < 12$ GeV or $|m_Z - m_{\ell\ell}| < 10$ GeV are rejected. The dimuon pair finding strategy is identical to the one used in the $t\bar{t}H$ leptonic event category.

Events in the ZH category are required to contain two oppositely charged, same-flavour lepton pairs. A potential Z boson candidate is reconstructed by a pair of same-flavour, opposite-sign muons (electrons) compatible with the Z boson mass within 10 (20) GeV. For events with four muons, the muon pair with $m_{\mu\mu}$ closer to m_Z is chosen as the Z boson candidate, while the other muon pair is selected as the Higgs boson candidate. The invariant mass of each same-flavour lepton pair of oppositely charged leptons is required to be larger than 12 GeV.

Two independent BDT discriminants are trained using simulated events from all three data-taking periods to distinguish WH and ZH signal events from the background. Similar to the ggH and $t\bar{t}H$ event category the training is performed using the TMVA software package [197]. The dominant background consists of the WZ and ZZ processes where both vector bosons decay leptonically. The BDT discriminant employed in the WH (ZH) category is based on observables describing the kinematic properties of the three (four) charged leptons in the event.

The observables used as input to the WH BDT include the angular separation $\Delta\phi$ and $\Delta\eta$ between the additional lepton and the Higgs boson candidate, between each muon from the Higgs boson candidate and the additional lepton, and between the negative vector p_T sum of all jets in the event \vec{H}_T^{miss} and the additional lepton. Also included is the transverse mass calculated from the \vec{H}_T^{miss} and the additional lepton. Finally, the flavour and p_T of the additional lepton, as well as the p_T^{miss} are considered.

The BDT discriminant trained in the ZH category aims to identify events where one lepton pair can be assigned to the decay of a Z boson, and the other one to the decay of the Higgs boson. For this purpose, the p_T and η of both, the Z and Higgs boson candidates are included as input to the BDT. Furthermore, several observables exploit the differences in angular distributions between the Higgs and the Z boson in ZH and ZZ events. Among other, these are the angular separation $\Delta\phi$ between the muons associated to the Higgs candidate, and the ΔR between the

leptons associated to the Z boson candidate, and the cosine of the polar angle between the Higgs and Z boson candidates. Lastly, the flavour of the lepton pair from the Z boson decay is included as an input observable. Figure 6.5 shows the BDT discriminants in the WH and ZH categories. The shapes and normalizations of both distributions are well modeled in sim-

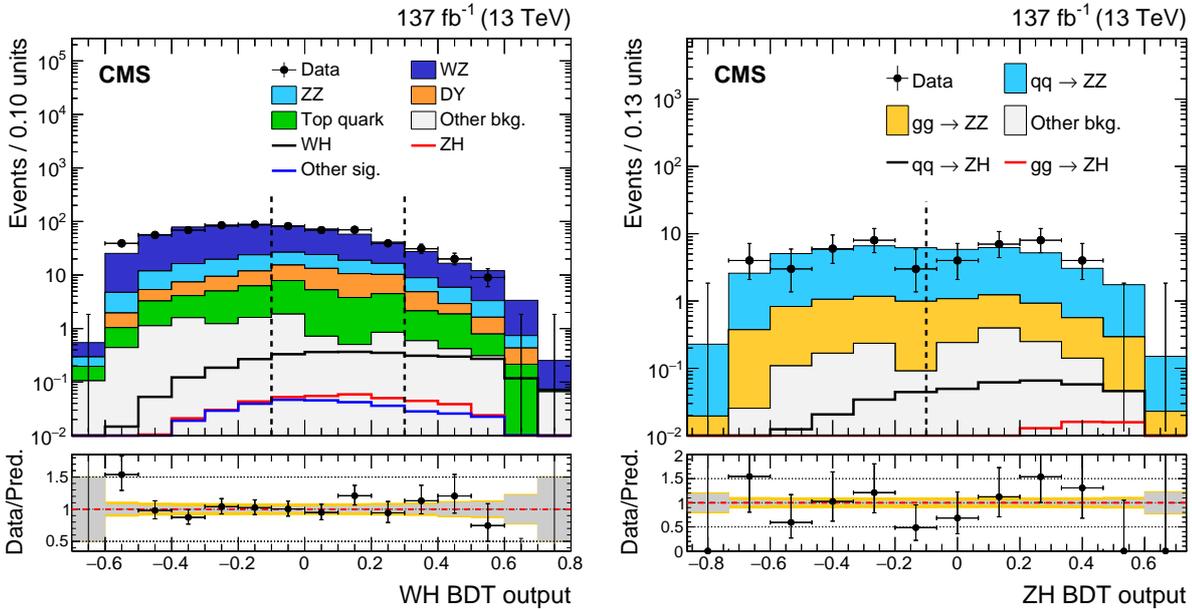


Figure 6.5: The BDT output distribution in the WH (left) and ZH (right) categories for events in the signal fit region with $110 < m_{\mu\mu} < 150$ GeV. The signal distributions expected from different production modes of the Higgs boson with $m_H = 125$ GeV are overlaid. The dashed vertical lines indicate the subcategory boundaries. Taken from Ref. [172].

ulation and agree with the measured distribution in data within the statistical uncertainties. Events in the WH (ZH) category are classified according to their BDT score into three (two) subcategories. The subcategory boundaries are determined using the same optimization algorithm employed in the ggH and t \bar{t} H categories.

The results are extracted using a simultaneous ML fit to the observed $m_{\mu\mu}$ distribution in each VH event category following the same approach as described in Sections 5.11 and 5.12. The background models in the VH subcategories are based on modified Breit-Wigner distributions. Any potential bias due to the choice of the background function is found to be negligible. The theoretical and experimental sources of systematic uncertainty which account for variations of the signal shape and rate are similar to those described in Section 5.11.2. A comprehensive overview of all uncertainties and their correlation is reported in Table B.1.

Figure 6.6 shows the $m_{\mu\mu}$ distribution for a $S/(S+B)$ weighted combination of all five VH subcategories, and the corresponding S+B fits. The observed (expected) significance for a Higgs boson with $m_H = 125.38 \text{ GeV}$ of the incompatibility with the background-only hypothesis is 2.0σ (0.4σ). The best-fit value of the signal strength extracted from the combined fit in all VH subcategories is

$$\hat{\mu}_{\text{VH}(H \rightarrow \mu^+\mu^-)} = 5.48_{-2.83}^{+3.10} \text{ (stat + syst).}$$

Due to the large irreducible backgrounds from the WZ and ZZ processes, and a rather small signal rate, the WH production category is limited in the expected sensitivity.

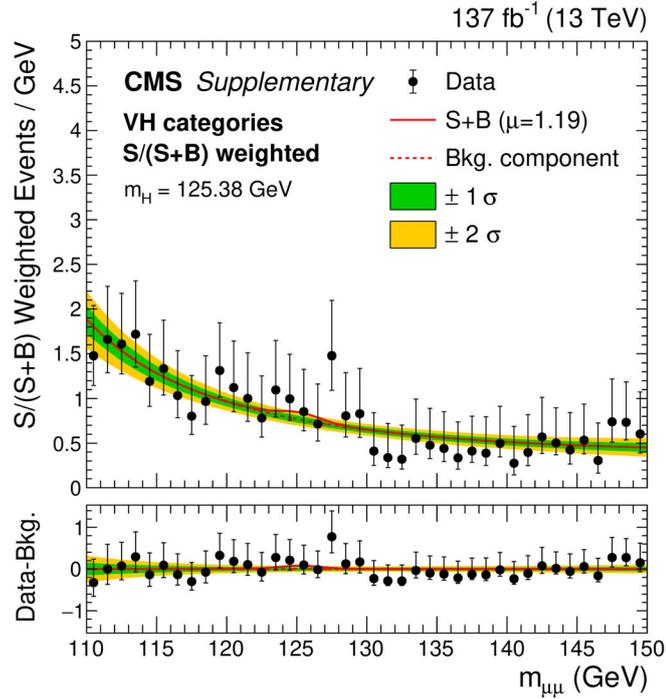


Figure 6.6: The $m_{\mu\mu}$ distribution for the $S/(S+B)$ weighted combination of VH event categories. The solid red line represents the best-fit $H \rightarrow \mu^+\mu^-$ signal, while the red dashed line corresponds to the background component. The background uncertainty is indicated by the green and yellow bands. In the lower panel, the residuals after background subtraction are shown. Supplementary material of Ref. [172].

6.2 Evidence for Higgs boson decay to muons

To probe the presence of a $H \rightarrow \mu^+\mu^-$ signal, a binned ML fit is performed simultaneously in all categories targeting the Higgs boson production modes, ggH, VBF, VH, and $t\bar{t}H$, as described in Chapter 5 and Section 6.1. Results are presented within a small window around the Higgs boson mass $120 < m_H < 130$ GeV, and at a mass value of $m_H = 125.38 \pm 0.14$ GeV, which corresponds to the most precise Higgs boson mass measurement to date [223].

The ggH, VH, and $t\bar{t}H$ analyses perform a fit to the shape of the observed $m_{\mu\mu}$ distribution. To obtain a signal model for a test mass m' , the signal shape parameters and the expected signal rate are interpolated using $H \rightarrow \mu^+\mu^-$ signal samples generated with a m_H of 120, 125, and 130 GeV. In contrast, the VBF category employs a ML template fit to the output of a DNN which includes $m_{\mu\mu}$ as one of the input observables. The results for a discrete target mass m' are extracted by shifting the mass value by $\Delta m = 125$ GeV $- m'$ when the DNN is evaluated, and fitting the resulting DNN templates. This procedure is applied to the background DNN templates and the observed distribution in data, while for the signal templates only the expected signal yields need to be updated according to the production cross section and decay rate for a Higgs boson with mass m' .

The probability, that a statistical fluctuation of the SM background is as large or larger than the apparent signal in a specified search range, is quantified by the local p -value. The observed and expected local p -values are presented for the combined fit, and for each Higgs boson production category as a function of the Higgs boson mass in Fig. 6.7.

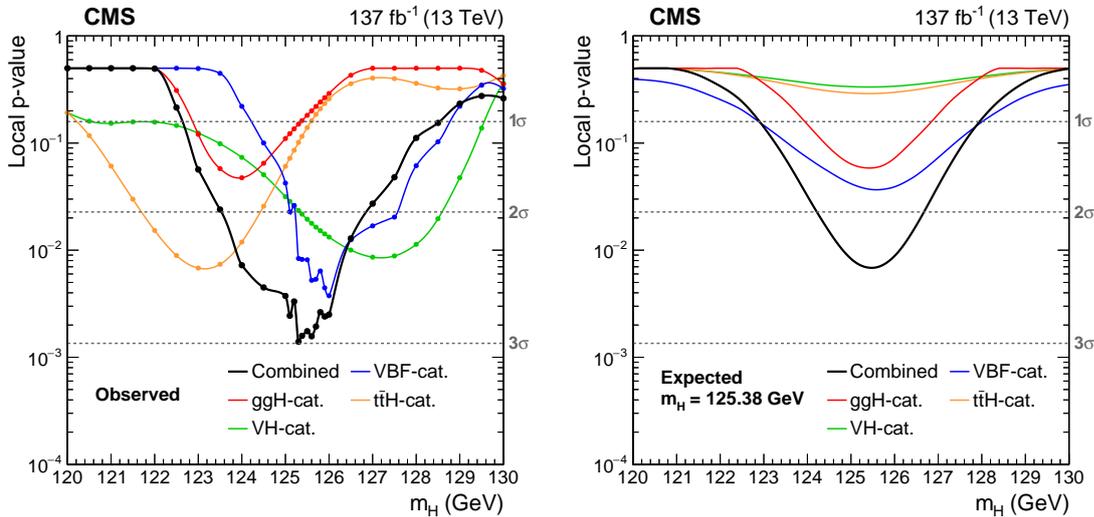


Figure 6.7: Observed (left) and expected (right) local p -values as a function of m_H . The p -values are shown for the combination as well as each individual Higgs boson production category. The solid markers indicate discrete test masses used to compute the observed p -values. Taken from Ref. [172].

The expected p -values are determined using the background expectation obtained from the S+B fit and injecting a SM $H \rightarrow \mu^+\mu^-$ signal at $m_H = 125.38$ GeV. The discrete fluctuations in the observed p -value are caused by event migration between bins in the observed DNN distribution when different mass hypotheses are tested.

Significance An excess of events over the SM background prediction is observed in data with a significance of 3.0σ , where the expectation for the SM Higgs boson with mass of 125.38 GeV is 2.5σ .

Signal strength The best-fit signal strength for the Higgs boson with mass of 125.38 GeV extracted from the combined fit is

$$\hat{\mu}_{H \rightarrow \mu^+\mu^-} = 1.19^{+0.41}_{-0.40} (\text{stat})^{+0.17}_{-0.16} (\text{syst}),$$

where the total uncertainty is separated into a statistical and systematic component. The statistical component is estimated by fixing nuisance parameters associated with systematic uncertainties to their best-fit value. The systematic component is then determined as the difference in quadrature between the total uncertainty and the statistical component. A breakdown of uncertainties and their impact on the result is given in Table 6.2. The combined results are dominated by statistical uncertainties due to the limited number of events in data.

Table 6.2: Uncertainty sources in the measurement of the signal strength $\hat{\mu}$ and their impact on the result. The uncertainty corresponding to the statistical precision of the simulation only affects the VBF category. Modified from Ref. [172].

Uncertainty source	$\Delta\hat{\mu}$	
Total post-fit uncertainty	+0.44	-0.42
Statistical uncertainty	+0.41	-0.40
Systematic uncertainty	+0.17	-0.16
Experimental syst.	+0.12	-0.11
Theoretical syst.	+0.10	-0.11
Simulation stat.	+0.07	-0.06

Figure 6.8 shows the best-fit values of the signal strength and the corresponding 68% CL intervals for each production category and the combined result. The best-fit signal strengths in each production category are consistent with the combined fit result as well as with the SM expectation.

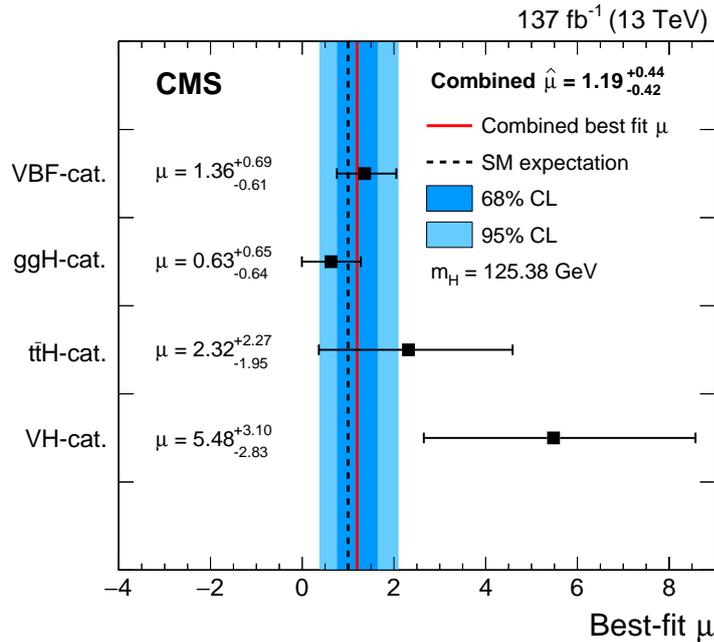


Figure 6.8: The measured signal strength modifiers for $m_H = 125.38 \text{ GeV}$ in each production category (black points) are compared to the result of the combined fit (solid red line) and the SM expectation (dashed line). Taken from Ref. [172].

Exclusion limits The observed upper limit at 95% CL on the signal strength is 1.9, while the expected limit in absence of a signal is 0.8. Assuming SM production cross sections for the Higgs boson, the branching ratio for decays to muons is constrained at 95% CL to be within $0.8 \cdot 10^{-4} < \mathcal{B}(H \rightarrow \mu^+\mu^-) < 4.5 \cdot 10^{-4}$. New physics scenarios without a Yukawa coupling of the Higgs boson to muons can therefore be excluded at more than 95% CL.

Dimuon mass resonance The $m_{\mu\mu}$ distribution for the $S/(S+B)$ weighted combination of all event categories is presented in Fig. 6.9. A small excess over the SM background prediction is observed in the weighted data distribution which is consistent with the resonant mass distribution expected from a $H \rightarrow \mu^+\mu^-$ signal with m_H near 125 GeV.

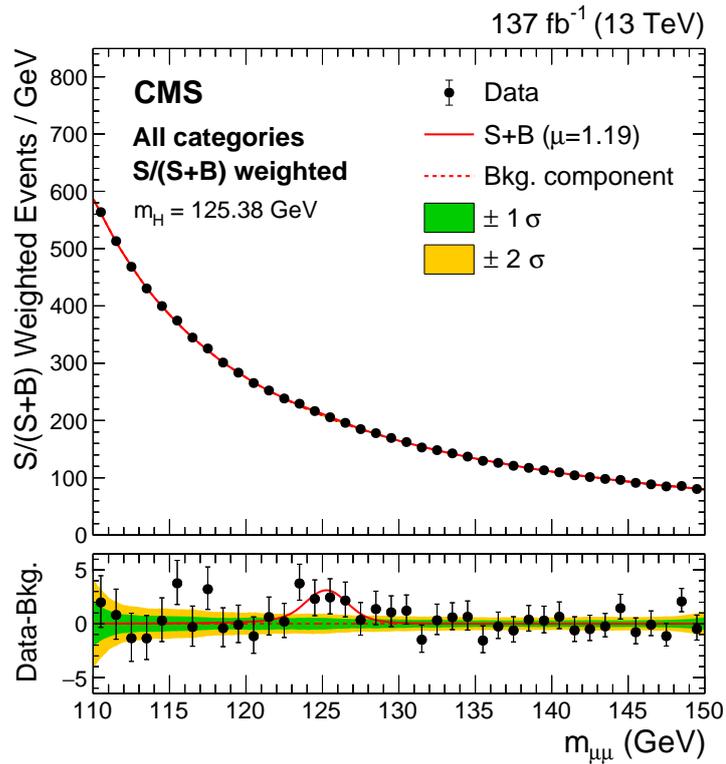


Figure 6.9: The $m_{\mu\mu}$ distribution for the S/(S+B) weighted combination of all event categories. The solid red line represents the best-fit $H \rightarrow \mu^+\mu^-$ signal, while the red dashed line corresponds to the background component. The background uncertainty is indicated by the green and yellow bands. In the lower panel, the residuals after background subtraction are shown. Taken from Ref. [172].

6.3 Scientific impact

The compatibility of this new result with the predicted SM Higgs boson coupling structure is probed within the κ -framework [213] following the approach presented in Ref. [32]. A set of coupling modifiers κ_i is introduced to parameterize deviations in the Higgs boson couplings to the particles i from the SM prediction. The best-fit value for the coupling modifier to muons is

$$\kappa_\mu = 1.07^{+0.29}_{-0.12}.$$

To illustrate the results of the combined κ -framework fit including all Higgs boson couplings which have been established so far, the reduced coupling strength modifiers are defined for weak bosons $y_V = \sqrt{\kappa_V} m_V / v$ and fermions $y_F = \kappa_F m_F / v$ in terms of the vacuum expectation value $v = 246.22 \text{ GeV}$ [54], and the particle mass m . The best-fit values for the reduced coupling strength modifiers of vector bosons, third generation fermions and muons are presented in Fig. 6.10 as a function of particle mass. The compatibility between the measured coupling strength modifiers and their SM expectation is quantified by an additional fit which fixes the coupling modifiers to the SM prediction. A comparison to the best-fit yields a p -value of 44%.

The results of the combined $H \rightarrow \mu^+\mu^-$ analysis are consistent with the SM expectation for a Higgs boson decay to a pair of muons and constitute the first evidence for a Yukawa coupling to the second generation fermions. They also represent the most precise measurement of the Higgs boson coupling to muons reported to date.

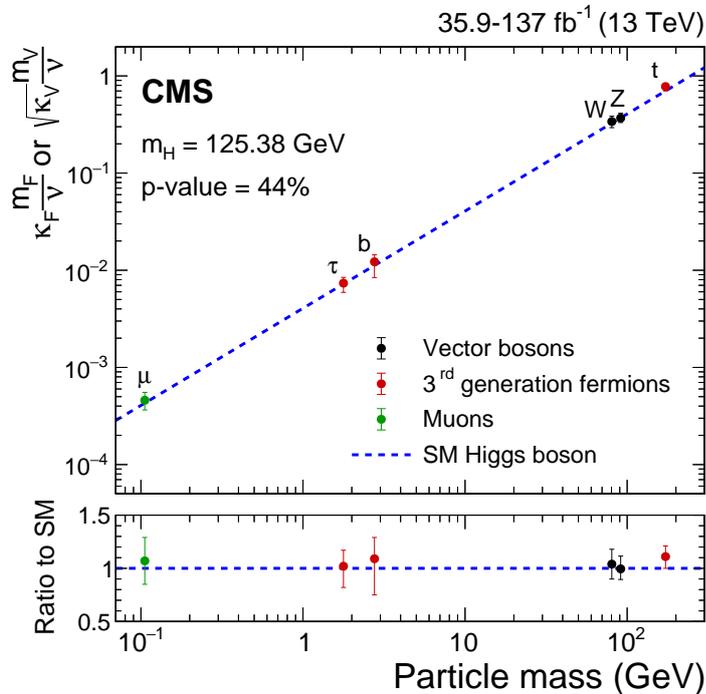


Figure 6.10: The reduced coupling modifiers for fermions and weak bosons estimated within the resolved κ -framework. The uncertainties represent 68% CL intervals of the measured parameters. The measured values are compared to the SM prediction. Taken from Ref. [172].

6.4 Comparison to previous CMS analysis

The presented $H \rightarrow \mu^+\mu^-$ analysis strategy HIG-19-006 [172] is improved compared to the strategy followed in the 2016 data analysis HIG-17-019 [41]. The new strategy leads to a substantial enhancement of the expected significance by about 35% in addition to the expected gain from the larger analyzed dataset. The main reasons are listed below:

- **Event classification** The previous analysis HIG-17-019 [41] employs a single BDT which was trained on the inclusive Higgs boson signal. This analysis was primarily sensitive to the ggH and VBF production of the Higgs boson. In contrast, the new analysis strategy HIG-19-006 [172] is based on four exclusive Higgs boson production categories targeting ggH, VBF, VH, and $t\bar{t}H$. In each category, multivariate discriminants are trained to distinguish the expected $H \rightarrow \mu^+\mu^-$ signal from the SM background processes. The input observables are chosen to account for the characteristic kinematic features of the targeted Higgs boson production mode. The new VH and $t\bar{t}H$ production categories increase the expected significance by about 4% with respect to the combined result.
- **FSR photon recovery + geometric track fit** Two independent methods described in Section 5.5 are implemented in HIG-19-006 [172] to improve the dimuon mass resolution. The FSR photon recovery strategy aims to reconstruct the original dimuon mass of the Higgs boson candidate in events affected by the radiation of a final state photon. Furthermore, an empirically muon momentum correction is performed to reduce the impact of a residual muon track displacement from the primary interaction point. These methods improve the overall analysis sensitivity by about 6%.
- **Per-event mass uncertainties** In HIG-17-019 [41], the BDT discriminant was insensitive to the dimuon mass resolution. To include this information in the final signal extraction fit, subcategories were defined based on the maximum $|\eta|$ of both muons in the event. The new analysis HIG-19-006 [172] embeds the dimuon mass resolution directly in the BDT training by reweighting events with the inverse of the per-event mass uncertainty estimated from the covariance matrix of the muon tracks as explained in Section 5.4.7. This approach allows the BDT to learn how the mass resolution evolves with the muon p_T and η . As a result events with precise dimuon mass resolution are propagated to the high BDT score.
- **Optimized categorization** Fifteen event categories were defined in HIG-17-019 [41] using the maximum $|\eta|$ of both muons and the BDT score. HIG-19-006 [172] employs a new iterative algorithm to choose the boundary values. Since the information on the dimuon mass resolution is included in the BDT, the optimization of the category boundaries is reduced to a single variable, i.e., the BDT score. This results in less event categories, while improving the overall sensitivity of the analysis.
- **Correlated background modeling in ggH** The background in the ggH category is dominated by Drell-Yan events. The shape of the $m_{\mu\mu}$ distribution is expected to be very similar across all ggH subcategories. For this reason, a new strategy with partially correlated

background parameters is implemented. This allows to keep the bias due to the choice of the background function under control and, at the same time, reduces the number of freely floating parameters in the final fit. This approach improves the sensitivity on $\hat{\mu}$ of the ggH category by about 10% with respect to the background model considered in the previous result HIG-17-019 [41].

- **Background estimation based on simulation in VBF category** The distinct kinematic properties of VBF signal events allow for a strong suppression of the Drell-Yan background contribution. Due to the small number of data events in the $m_{\mu\mu}$ sideband region, the large expected sensitivity of this channel would be limited by the statistical precision of a data-driven background estimation. To improve the precision of the background estimation, simulated events are used to model the background, and the signal is extracted in terms of a simultaneous ML template fit to the DNN output distribution. The new VBF analysis improves the expected significance of the combined result by about 10%.

6.5 Comparison to ATLAS analysis

The analysis strategy and the results of the search for Higgs boson decays to muons by the ATLAS Collaboration are compared to the new results from CMS. In general, both analyses follow a similar approach. The main differences in the event selection and categorization, the analysis optimization, as well as the signal extraction are described in the following:

- **Trigger selection** ATLAS employs a single muon trigger with p_T threshold of up to 26 GeV for isolated muons similar to the CMS analysis. In addition, however, ATLAS uses a trigger for non-isolated muons with $p_T > 50$ GeV. Since both muons from the decay of the Higgs boson are expected to be isolated, this additional trigger recovers probably only signal events, where a muon overlaps with a jet by accident. Based on lepton isolation studies, the impact of a relaxed isolation on the CMS analysis is estimated to increase the acceptance by about 2% for both, signal and background events.
- **Lepton selection** A further difference between both detectors is the muon detector coverage. ATLAS selects muons up to $|\eta| < 2.7$, which increases the muon acceptance compared to CMS. Furthermore, different lepton p_T thresholds are chosen in the ATLAS and CMS analyses. While ATLAS imposes subsequently relaxed thresholds on leptons down to 6 GeV in multilepton events, CMS uses a constant p_T threshold for all leptons across all event categories. The impact of relaxed p_T thresholds was studied in the CMS $t\bar{t}H$ leptonic event category. The estimated overall gain was rather small, due to an increasing background component from misidentified leptons in events with three lepton final state.
- **Jet selection** A jet p_T threshold of 25 GeV is applied by ATLAS for jets in the central part of the detector. This threshold is tightened in the forward region to 30 GeV for $2.4 < |\eta| < 4.5$. The ATLAS selection may lead to a slightly smaller acceptance for VBF jets compared to

the VBF jet selection by CMS which requires a leading (sub-leading) jet p_T of 35 (25) GeV for $|\eta| < 4.7$.

- **Event categorization & analysis optimization** The ATLAS and CMS analyses both define exclusive event categories to address the dominant Higgs boson production modes ggH, VBF, VH and $t\bar{t}H$.
 - In ATLAS, events are classified according to the number of reconstructed jets into three jet multiplicity categories with zero, one, or two and more jets. These categories are optimized using BDT-based multivariate discriminants to distinguish the ggH and VBF signal from the dominant Drell-Yan background. Finally, twelve ggH subcategories and four VBF categories are defined based on the BDT outputs. In contrast, the CMS analysis categorizes events into a dedicated VBF and ggH channel using characteristic dijet kinematics of VBF jets. Instead of three BDT classifiers for different jet multiplicities, a single BDT discriminant which includes the jet multiplicity as input observable is trained for each year of data taking. This approach results in five ggH event categories, while the VBF category is not further subdivided.
 - The VH categories employed by ATLAS and CMS perform very similarly. One of the main differences is the method to assign the correct muon pair to the Higgs boson candidate. The approach followed by ATLAS is based on a χ^2 criterion, which allows for a correct pairing in 97% of the selected events with three leptons in the final state. The strategy followed in the CMS VH analysis provides a similar performance [195].
 - The $t\bar{t}H$ analysis presented in this thesis aims at all possible decays of top quarks produced in association with the Higgs boson, while the ATLAS analysis only focuses on leptonic decays. All $t\bar{t}H$ hadronic events are removed from the ATLAS event selection. Moreover, in the ATLAS analysis the Higgs boson candidate is built from the two muons with highest single muon p_T , which provides a correct pairing in 80% of the events compared to 94% using the muon pair selection employed in the $t\bar{t}H$ leptonic analysis presented here for CMS. Lastly, an approach to specifically reduce non-prompt lepton background using a multivariate lepton identification algorithm is only followed in the analysis by CMS.
 - The ATLAS analysis does not explicitly exploit the dimuon mass resolution, whereas the CMS analysis implements a per-event mass uncertainty estimate in the BDT training, which allows the BDTs to learn the correlation between the dimuon mass resolution and other kinematic event properties. Consequently, events with a precise dimuon mass resolution are propagated to the high-BDT score. The signal models of the event categories with the best and the worst dimuon mass resolution are shown for ATLAS and CMS in Fig. 6.11. The dimuon mass resolution in the best (worst) CMS category is about 50% (40%) more precise than the mass resolution in the respective ATLAS category due to a better momentum resolution for central muons of the CMS detector compared to ATLAS [224]. The reason for this is mainly the larger magnetic field strength of the CMS magnet in the inner detector.

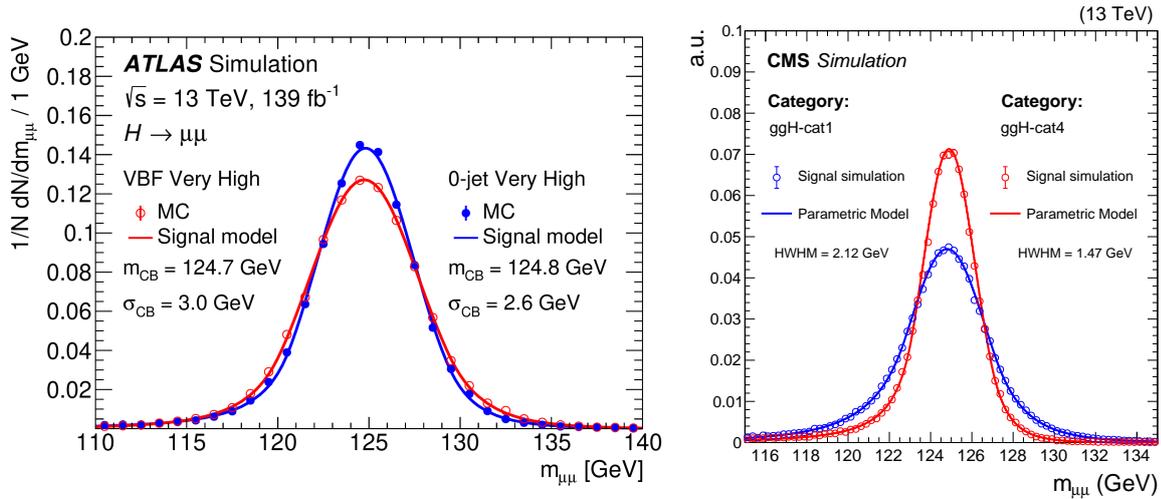


Figure 6.11: Dimuon mass distributions of $H \rightarrow \mu^+\mu^-$ signal events in two specific categories of the ATLAS (left) and CMS (right) analyses. The categories correspond to those with the best and worst mass resolution, respectively. For both categories the distribution from the signal simulation and the parametric signal model is shown using a DCB function. The resolution is represented in terms of the Gaussian width of the DCB or the HWHM. Taken from Ref. [42] and Ref. [172].

- **Signal extraction** Apart from the VBF category, ATLAS and CMS follow a similar strategy to extract the signal. A simultaneous ML fit is performed to the $m_{\mu\mu}$ distribution using analytical functions. While ATLAS uses the core-pdf method in all analysis categories including $t\bar{t}H$ and VH, uncorrelated background models are employed in the $t\bar{t}H$ and VH categories by CMS.
- **Results** ATLAS observes a small excess over the SM backgrounds with an observed (expected) significance of 2.0σ (1.7σ) at a Higgs boson mass of $m_H = 125.09 \pm 0.24$ GeV [43]. This value is compatible within uncertainties with a more recent measurement of the Higgs boson mass [223] used in the CMS $H \rightarrow \mu^+\mu^-$ analysis. The measured signal strength by ATLAS is

$$\hat{\mu}_{H \rightarrow \mu^+\mu^-} = 1.2 \pm 0.6 \text{ (stat + syst).}$$

The observed excess in data by ATLAS is compatible with the CMS result. Moreover, the best-fit signal strengths in each production category as well as the combined fit are consistent within the experimental precision between both experiments. Compared to the ATLAS, the expected significance obtained by the CMS Collaboration is about 30% stronger due to a more precise measurement of the muon momentum in the inner CMS tracking detector and a few differences in the analysis strategy which were described in this section.

6.6 Future prospects

The future of the $H \rightarrow \mu^+\mu^-$ analysis appears very promising. Being statistically limited, the analysis will profit from larger datasets that will be collected in the coming years. Between 2022 and 2025 (LHC Run 3), the LHC will provide proton-proton collisions with a center-of-mass energy of 13 or even 14 TeV. Within this time period, the ATLAS and CMS detector are each anticipated to collect a dataset corresponding to an integrated luminosity of about 160 fb^{-1} , roughly doubling the dataset analyzed in the present work. During the second long shutdown currently underway, and in a long shutdown after LHC Run 3, all CMS detector components will be upgraded [225–230] to prepare the CMS detector for the demanding operational conditions at the HL-LHC [231, 232]. This upgrade is referred to as the CMS *Phase-2 Upgrade*. The instantaneous luminosity at the HL-LHC will be increased by a factor of 5 to 7.5 beyond the LHC’s nominal value, leading to much larger data samples, but also increased pile up interactions. Apart from several improvements to the detector components, the Phase-2 Upgrade especially aims at providing the detector with the radiation hardness required for the HL-LHC operation conditions. The largest impact on the $H \rightarrow \mu^+\mu^-$ analysis is expected from the silicon tracker upgrade which will improve the relative resolution of the $H \rightarrow \mu^+\mu^-$ signal peak from now 0.92% to 0.65% for muons in the central region [226]. Additionally, the signal acceptance will be increased by the tracker and muon detector upgrade that will extend the acceptance for muons to $|\eta| = 2.8$. These detector improvements, and the larger dataset will particularly help to increase the sensitivity of the $H \rightarrow \mu^+\mu^-$ analysis.

A naive estimate of the expected significance at higher integrated luminosities is obtained by scaling the recorded dataset according to Poisson statistics with $f = \sqrt{\mathcal{L}^{\text{target}}/137 \text{ fb}^{-1}}$ in terms of a target integrated luminosity $\mathcal{L}^{\text{target}}$. This projection does not take into account further improvements on the analysis strategy, the CMS Phase-2 Upgrade, a center-of-mass energy of 14 TeV, the increased pile up or systematic effects. The projected results are reported in Table 6.3. By convention the observation of a new process is claimed when the observed significance reaches 5σ . Given a SM like Higgs boson coupling to muons, the expected significance reaches this threshold using $\mathcal{L}^{\text{target}} = 565 \text{ fb}^{-1}$. The CMS experiment will collect this amount of data by the end of this decade.

Table 6.3: Luminosity projection of the expected significance obtained by the CMS $H \rightarrow \mu^+\mu^-$ analysis. The reported values of the expected significance are approximated using a luminosity scaling of the LHC Run 2 results based on an integrated luminosity of 137 fb^{-1} .

Expected significance	Integrated luminosity		
	137 fb^{-1}	300 fb^{-1}	565 fb^{-1}
CMS $t\bar{t}H, H \rightarrow \mu^+\mu^-$	0.5σ	0.8σ	1.1σ
CMS combined $H \rightarrow \mu^+\mu^-$	2.5σ	3.6σ	5.0σ

Applying this prescription to the current ATLAS and CMS results assuming $\mathcal{L}^{\text{target}} = 300 \text{ fb}^{-1}$ for each of the experiments, a combined expected significance of 4.4σ is derived. Considering

the observed signal strength of 1.2 for both experiments, and potential further improvements to the analysis strategies in the future, the observation of Higgs boson decays to muons might already be feasible by a combined ATLAS and CMS result using the full LHC Run 2 and Run 3 data.

A mass measurement in the $H \rightarrow \mu^+\mu^-$ channel has not been performed yet. Given the current statistical uncertainty, a mass measurement is not competitive to the precise measurements in the $H \rightarrow \gamma\gamma$ [223] and $H \rightarrow ZZ^* \rightarrow 4\ell$ [192] channels. However, after a potential observation of the $H \rightarrow \mu^+\mu^-$ decay, and with the upgraded CMS detector at the HL-LHC, the $H \rightarrow \mu^+\mu^-$ channel will yield a further Higgs boson mass measurement due to its precise $m_{\mu\mu}$ resolution in addition to the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ channels.

The presented CMS $H \rightarrow \mu^+\mu^-$ analysis includes several new methods, and a completely new analysis strategy compared to the previous 2016 data analysis as described in Section 6.4. All analysis aspects have been optimized to achieve the best possible results. However, the following iteration of the $H \rightarrow \mu^+\mu^-$ analysis may include these improvements:

- A background estimation based on simulated events in the $t\bar{t}H$ and VH categories can be considered. Due to the small number of events in the $m_{\mu\mu}$ data sidebands, both analyses are limited by large statistical uncertainties. A new background estimate might improve the sensitivity by about 20% in each production category, corresponding to an overall improvement of the combined expected significance of about 1%. It should be noted that this background estimation method would require a fake rate method to estimate the misidentified lepton background from data, as well as additional studies on how to model the background for large jet multiplicities.
- Dedicated three and four lepton channels in the $t\bar{t}H$ and VH categories, which account for hadronically decaying τ -leptons τ_{had} , targeting the final states with $\mu^+\mu^- + \tau_{\text{had}}^\pm$ and $\mu^+\mu^- + \ell^\pm\tau_{\text{had}}^\mp$ can be used in addition to the final states considered in the presented analysis.
- Improvements of the physics object identification techniques such as DeepCSV and LEPTONMVA would be expected to result in a further optimized event selection with increased signal efficiency and reduced backgrounds.

Conclusions

The discovery of the Higgs boson, announced in 2012, has been one of the most important results of particle physics in the last decade. Ever since then, the properties of the Higgs boson have been examined to probe the SM, and to explore the physics landscape beyond it. Thus far, the Higgs boson production via gluon fusion (ggH), via vector boson fusion (VBF), in association with a vector boson (VH), and in association with a top quark-antiquark pair ($t\bar{t}H$) as well as the decays of the Higgs boson to the electroweak gauge bosons and the charged fermions of the third generation have been observed. In all these production and decay modes, the measured rates have been found to be consistent with the SM predictions within the experimental and theoretical uncertainties. The analysis of Higgs boson decays to muons extends the investigation of Yukawa couplings to the light fermions of the second generation which is of particular importance to gain a deeper understanding of the nature of electroweak symmetry breaking.

In this thesis, the first search for Higgs boson decays to a pair of muons, produced in association with a top quark-antiquark pair ($t\bar{t}H$), has been presented using proton-proton collision data recorded with the CMS experiment at a center-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 137 fb^{-1} . The analysis has been designed to complement three further Higgs boson analyses targeting the ggH, VBF, and VH production modes. Furthermore, a method to recover final state radiation photons has been refined to improve the resolution of the reconstructed Higgs boson mass in all four categories, resulting in an increased overall sensitivity.

The search for $t\bar{t}H$ ($H \rightarrow \mu^+\mu^-$) was conducted in the final state with two muons from the decay of the Higgs boson and additional decay products of the top quarks. Two $t\bar{t}H$ categories which focus on either the leptonic or the hadronic top quark decay were individually optimized. Observables sensitive to the characteristic properties of the $t\bar{t}H$ production mechanism and the decay of the Higgs boson to muons were combined in multivariate discriminants based on boosted decision trees (BDTs). In particular, observables which identify the top quark decay products were helpful to reduce background events. A dedicated algorithm was implemented and optimized to identify the hadronic decay of a top quark to three resolved jets. The algorithm is based on a BDT classifier which includes the information of a constrained kinematic top quark fit, and different additional jet properties to probe the compatibility of a jet-triplet with the expected signature from a hadronic top quark decay. Due to the impact of the Higgs boson mass resolution on the analysis sensitivity, new techniques were employed to improve the resolution and to exploit the correlation between the resolution and the observables used in the BDT classifiers. Events were classified according to the BDT output score into five $t\bar{t}H$

subcategories which provide phase-space regions with enhanced signal-to-background ratio. The presence of a potential $t\bar{t}H$ ($H \rightarrow \mu^+\mu^-$) signal was quantified by a simultaneous maximum likelihood fit to the observed dimuon mass distributions in all $t\bar{t}H$ subcategories. This allowed for a data-driven estimation of the background contribution in the signal region around the Higgs boson mass from signal-depleted $m_{\mu\mu}$ sidebands.

The observed (expected) significance over the SM prediction for a Higgs boson with a mass of 125.38 GeV is 1.2σ (0.5σ). The observed upper limit at 95% confidence level on the cross section times branching fraction for $t\bar{t}H$ ($H \rightarrow \mu^+\mu^-$) is 6.5 times the SM prediction, while the expected limit is 4.2 (4.9) assuming the absence (presence) of a SM signal. The measured value of the signal strength modifier, defined as the ratio of the observed signal yield to that predicted by the SM, is

$$\hat{\mu}_{t\bar{t}H(H \rightarrow \mu^+\mu^-)} = 2.32^{+2.27}_{-1.95} \text{ (stat + syst).}$$

When this result is combined with the Higgs boson production categories targeting ggH, VBF, and VH, an excess of events over the background prediction is observed in data with a significance of 3.0σ , where 2.5σ are expected for a SM Higgs boson with $m_H = 125.38$ GeV. The corresponding measured signal strength is

$$\hat{\mu}_{H \rightarrow \mu^+\mu^-} = 1.19^{+0.41}_{-0.40} \text{ (stat)}^{+0.17}_{-0.16} \text{ (syst),}$$

consistent with the SM expectation within a relative experimental precision of about 40%. This result constitutes the first evidence for the decay of the Higgs boson to second generation fermions and is the most precise measurement of the Higgs boson coupling to muons reported to date. The observed upper limit at 95% confidence level on the signal strength is 1.9, while the expected limit in absence of a signal is 0.8. Assuming SM production cross sections for the Higgs boson, the branching fraction for decays to muons is constrained at 95% confidence level to be within $0.8 \cdot 10^{-4} < \mathcal{B}(H \rightarrow \mu^+\mu^-) < 4.5 \cdot 10^{-4}$. New physics scenarios without a Yukawa coupling of the Higgs boson to muons can therefore be excluded at more than 95% confidence level.

The observed excess of events in data by the CMS Collaboration is compatible with the recently published result by the ATLAS Collaboration, which observes a slight excess over the SM backgrounds with an observed (expected) significance of 2.0σ (1.7σ) at a Higgs boson mass of $m_H = 125.09$ GeV. Moreover, the measured signal strengths in each of the Higgs boson production categories as well as the combination are consistent between both experiments within the experimental and theoretical uncertainties.

Within this thesis, the Yukawa coupling of the muon, which is more than 1000 times smaller than that of the top quark, the heaviest SM particle, has been probed. The new result on Higgs boson decays to muons constitutes the first evidence for the Yukawa coupling to second generation fermions. Yet, the LHC has just delivered approximately 5% of its projected full dataset. The expected datasets to be collected during LHC Run 3 and at the HL-LHC will allow the investigation of rare processes with even smaller production rates, which might shed light on new physics phenomena and the fundamental laws of nature kept hidden so far.

Appendix **A**

Method to calibrate event-by-event mass uncertainties

The calibration of the estimated per-event Higgs boson mass resolution was developed in Ref. [185] and is reported for completeness. A Z boson control region is used for this purpose. Events with dimuon mass in the range between $75 < m_{\mu\mu} < 105$ GeV are selected and classified according to the $|\eta|$ of the muons. Each of the $|\eta|$ categories is further subdivided into four p_T regions 30–45, 45–52, 52–62 and 62–200 GeV of the leading muon p_T .

In events with a leading muon $p_T > 52$ GeV the mass peak is modeled using a Voigtian function. The dimuon mass distribution for events with a leading muon p_T smaller than 52 GeV shows slightly asymmetric tails which are better modeled using a convolution of a Breit-Wigner function and a double-sided Crystal Ball function.

The peak and width parameters of the Breit-Wigner function are fixed to the nominal mass and width of the Z boson, while all other parameters are allowed to float freely in the fit. The width of the Gaussian core σ is compared to the median value of the estimated Higgs boson mass uncertainty for all events in the corresponding category. Their ratio is used to define a calibration factor for each (η, p_T) category. The calibration factors are measured independently for data and simulation, as well as for each data-taking period. The resulting calibration factors are listed in Table A.1.

The calibration factors, in both data and simulation, and in all (η, p_T) bins are found to be in the range between 1.05 and 1.30. This indicates that the per-event dimuon mass resolution estimated from the covariance matrix of the track fit tends to overestimate the mass resolution of the Z boson mass peak by about 5–30%. In particular, events high- p_T muons with at least one muon detected in the barrel region require large corrections. The calibration factors extracted from simulation are slightly larger than those from data, and they progressively decrease from 2016 to 2018 in both, data and simulation.

A closure test of the calibration is performed using the same set of events divided into several bins depending on the estimated dimuon mass resolution. The resolution is again extracted from a fit to the dimuon mass distribution in the corresponding category using the same analytical model described previously. Figure A.1 presents the comparison between the fitted value of the mass resolution in each bin with the corresponding median value of the estimated dimuon mass uncertainty from the track fit for both data and simulation. The calibrated per-event dimuon mass uncertainty and the measured resolution agree within 10% in both data and simulation.

Table A.1: Measured calibration factors using a fit to the dimuon mass distribution in a Z boson control region. The calibration factors are reported separately for data and simulation for each data-taking period and for all (η, p_T) bins. Taken from Ref. [185].

Leading muon p_T -bin (GeV)	Muon $ \eta $ -bin	Data			Simulation		
		2016	2017	2018	2016	2017	2018
30–45	BB+OB+EB	1.24	1.19	1.16	1.25	1.16	1.17
30–45	BO+OO+EO	1.18	1.13	1.11	1.19	1.11	1.12
30–45	BE+OE+EE	1.15	1.11	1.07	1.17	1.07	1.09
45–52	BB	1.15	1.19	1.16	1.14	1.16	1.10
45–52	BO	1.16	1.13	1.12	1.17	1.12	1.13
45–52	BE	1.16	1.10	1.08	1.18	1.08	1.12
45–52	OB	1.14	1.12	1.08	1.16	1.08	1.10
45–52	OO	1.12	1.09	1.07	1.11	1.11	1.09
45–52	OE	1.07	1.07	1.06	1.10	1.09	1.09
45–52	EB	1.11	1.05	1.03	1.15	1.09	1.05
45–52	EO	1.13	1.03	1.05	1.16	1.10	1.04
45–52	EE	1.00	0.98	0.96	1.04	1.02	1.00
52–62	BB	1.25	1.21	1.19	1.26	1.24	1.21
52–62	BO	1.26	1.24	1.24	1.28	1.28	1.26
52–62	BE	1.19	1.14	1.15	1.22	1.16	1.16
52–62	OB	1.24	1.24	1.22	1.26	1.27	1.24
52–62	OO	1.17	1.16	1.15	1.19	1.18	1.17
52–62	OE	1.17	1.17	1.15	1.20	1.21	1.17
52–62	EB	1.19	1.18	1.13	1.23	1.23	1.18
52–62	EO	1.19	1.15	1.13	1.23	1.20	1.17
52–62	EE	1.11	1.10	1.08	1.15	1.14	1.12
62–200	BB	1.24	1.21	1.18	1.26	1.25	1.19
62–200	BO	1.27	1.24	1.22	1.29	1.28	1.23
62–200	BE	1.15	1.12	1.14	1.17	1.17	1.16
62–200	OB	1.23	1.22	1.21	1.26	1.27	1.23
62–200	OO	1.16	1.15	1.14	1.18	1.18	1.16
62–200	OE	1.18	1.15	1.15	1.21	1.18	1.18
62–200	EB	1.15	1.10	1.12	1.17	1.16	1.16
62–200	EO	1.16	1.11	1.10	1.20	1.18	1.13
62–200	EE	1.09	1.08	1.06	1.14	1.15	1.10

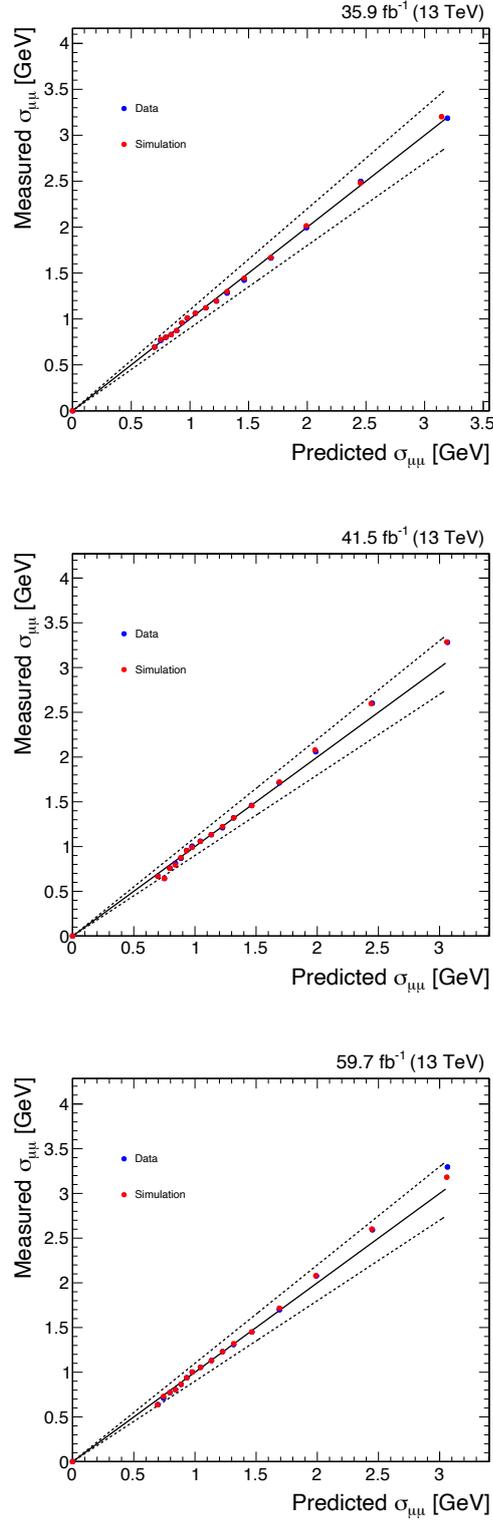


Figure A.1: Closure test of the per-event dimuon mass uncertainty calibration. The closure test is performed separately for data (blue) and simulation (red) in 2016 (top), 2017 (center), and 2018 (bottom). The x-axis corresponds to the median calibrated mass uncertainty in each bin, while the y-axis shows the resolution of the Z boson mass peak extracted from an analytical fit. Taken from Ref. [185].

Appendix B

Correlation model of systematic uncertainties in the combined fit

The considered systematic uncertainties and the corresponding correlation scheme used in the combined CMS $H \rightarrow \mu^+\mu^-$ analysis are given in Table B.1.

Table B.1: Summary of the systematic uncertainties and the correlation scheme considered in the combined $H \rightarrow \mu^+\mu^-$ analysis. Modified and taken from supplementary material of Ref. [172].

Source of uncertainty	Categories and processes	Type	Correlation between cat.	Correlation between year
Theoretical uncertainties				
μ_R and μ_F for ggH	ggH in all cat.	Rate	Correlated	Correlated
μ_R and μ_F for VBF	VBF in all cat.	Rate	Correlated	Correlated
μ_R and μ_F for $t\bar{t}H$	$t\bar{t}H$ in all cat.	Rate	Correlated	Correlated
μ_R and μ_F for VH	VH in all cat.	Rate	Correlated	Correlated
PDF for ggH	ggH in all cat.	Rate	Correlated	Correlated
PDF for VBF	VBF in all cat.	Rate	Correlated	Correlated
PDF for $t\bar{t}H$	$t\bar{t}H$ in all cat.	Rate	Correlated	Correlated
PDF for VH	VH in all cat.	Rate	Correlated	Correlated
ggH accept. vs $(p_T(H), N_j, m_{jj})$	ggH in all cat.	Shape in VBF, rate in others	Correlated	Correlated
VBF accept. vs $(p_T(H), N_j, m_{jj})$	VBF in all cat.	Shape in VBF, rate in others	Correlated	Correlated
$t\bar{t}H$ accept. from μ_R and μ_F	$t\bar{t}H$ in all cat.	Rate	Correlated	Correlated
VH accept. from μ_R and μ_F	VH in all cat.	Rate	Correlated	Correlated
$t\bar{t}H$ accept. from PDF	$t\bar{t}H$ in all cat.	Rate	Correlated	Correlated
VH accept. from PDF	VH in all cat.	Rate	Correlated	Correlated
PYTHIA ISR and FSR	Sig. in all cat., bkg. in VBF	Shape in VBF, rate in others	Correlated	Correlated
PYTHIA vs HERWIG	VBF and Z-EWK in VBF cat.	Shape	Correlated	Correlated
μ_R and μ_F for Drell–Yan	VBF cat.	Shape	Correlated	Correlated
μ_R and μ_F for Z-EWK	VBF cat.	Shape	Correlated	Correlated
μ_R and μ_F for top bkg.	VBF cat.	Shape	Correlated	Correlated
μ_R and μ_F for diboson	VBF cat.	Shape	Correlated	Correlated
PDF for Drell–Yan	VBF cat.	Shape	Correlated	Correlated
PDF for Z-EWK	VBF cat.	Shape	Correlated	Correlated
PDF for top bkg.	VBF cat.	Shape	Correlated	Correlated
PDF for dibosons	VBF cat.	Shape	Correlated	Correlated
Size of simulated samples	VBF cat.	Bin-by-bin	—	Uncorrelated
Experimental uncertainties				
Integrated luminosity	Sig. in all cat., bkg. in VBF	Rate	Correlated	Partial
Muon efficiency	Sig. in all cat., bkg. in VBF	Rate	Correlated	Correlated
Electron efficiency	Sig. in $t\bar{t}H$ and VH	Rate	Correlated	Correlated
Muon trigger	Sig. in all cat., bkg. in VBF	Rate	Correlated	Correlated
Muon p_T scale	Sig. in all cat., bkg. in VBF	Shape in VBF, rate in others	Correlated	Correlated
Nonprompt leptons	Sig. in $t\bar{t}H$ and VH	Rate	Correlated	Correlated
Pileup model	Sig. in all cat., bkg. in VBF	Shape in VBF, rate in others	Correlated	Uncorrelated
L1 inefficiency	Sig. in all cat., bkg. in VBF	Shape in VBF, rate in others	Correlated	Uncorrelated
B-tagging efficiency	Sig. in all cat., bkg. in VBF	Shape in VBF, rate in others	Correlated	Correlated
Jet energy scale	Sig. in all cat., bkg. in VBF	Shape in VBF, rate in others	Correlated	Partial
Jet energy resolution	Sig. in all cat., bkg. in VBF	Shape in VBF, rate in others	Correlated	Uncorrelated

Bibliography

- [1] S. L. Glashow, “Partial Symmetries of Weak Interactions”, *Nucl. Phys.* **22** (1961) 579–588, doi:10.1016/0029-5582(61)90469-2.
- [2] S. Weinberg, “A model of leptons”, *Phys. Rev. Lett.* **19** (Nov, 1967) 1264–1266, doi:10.1103/PhysRevLett.19.1264.
- [3] S. L. Glashow, J. Iliopoulos, and L. Maiani, “Weak Interactions with Lepton-Hadron Symmetry”, *Phys. Rev. D* **2** (1970) 1285–1292, doi:10.1103/PhysRevD.2.1285.
- [4] A. Salam, “Elementary particle theory”, *Eighth Nobel Symposium*. Stockholm.
- [5] F. Englert and R. Brout, “Broken symmetry and the mass of gauge vector mesons”, *Phys. Rev. Lett.* **13** (1964) 321, doi:10.1103/PhysRevLett.13.321.
- [6] P. W. Higgs, “Broken symmetries and the masses of gauge bosons”, *Phys. Rev. Lett.* **13** (Oct, 1964) 508–509, doi:10.1103/PhysRevLett.13.508.
- [7] P. W. Higgs, “Spontaneous symmetry breakdown without massless bosons”, *Phys. Rev.* **145** (1966) 1156, doi:10.1103/PhysRev.145.1156.
- [8] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, “Global Conservation Laws and Massless Particles”, *Phys. Rev. Lett.* **13** (1964) 585–587, doi:10.1103/PhysRevLett.13.585.
- [9] G. Arnison et al., “Experimental observation of isolated large transverse energy electrons with associated missing energy at $\sqrt{s} = 540$ GeV”, *Physics Letters B* **122** (1983), no. 1, 103–116, doi:https://doi.org/10.1016/0370-2693(83)91177-2.
- [10] M. Banner et al., “Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider”, *Physics Letters B* **122** (1983), no. 5, 476–485, doi:https://doi.org/10.1016/0370-2693(83)91605-2.
- [11] UA1 Collaboration, “Experimental Observation of Lepton Pairs of Invariant Mass Around 95-GeV/c² at the CERN SPS Collider”, *Phys. Lett. B* **126** (1983) 398–410, doi:10.1016/0370-2693(83)90188-0.
- [12] UA2 Collaboration, “Evidence for $Z^0 \rightarrow e^+ e^-$ at the CERN $\bar{p}p$ Collider”, *Phys. Lett. B* **129** (1983) 130–140, doi:10.1016/0370-2693(83)90744-X.
- [13] P. W. Graham, D. E. Kaplan, and S. Rajendran, “Cosmological relaxation of the electroweak scale”, *Physical Review Letters* **115** (Nov, 2015) doi:10.1103/physrevlett.115.221801.

BIBLIOGRAPHY

- [14] J. Espinosa et al., “Cosmological higgs-axion interplay for a naturally small electroweak scale”, *Physical Review Letters* **115** (Dec, 2015)
doi:10.1103/physrevlett.115.251803.
- [15] G. Dvali, “Cosmological relaxation of higgs mass before and after lhc and naturalness”, 2019.
- [16] ATLAS Collaboration, “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”, *Phys. Lett. B* **716** (2012) 1,
doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.
- [17] CMS Collaboration, “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, *Phys. Lett. B* **716** (2012) 30,
doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.
- [18] CMS Collaboration, “Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV”, *JHEP* **06** (2013) 081,
doi:10.1007/JHEP06(2013)081, arXiv:1303.4571.
- [19] CMS Collaboration, “Observation of the Higgs boson decay to a pair of τ leptons with the CMS detector”, *Phys. Lett. B* **779** (2018) 283,
doi:10.1016/j.physletb.2018.02.004, arXiv:1708.00373.
- [20] CMS Collaboration, “Measurements of Higgs boson properties in the diphoton decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JHEP* **11** (2018) 185,
doi:10.1007/JHEP11(2018)185, arXiv:1804.02716.
- [21] CMS Collaboration, “Measurements of properties of the Higgs boson decaying to a W boson pair in pp collisions at $\sqrt{s} = 13$ TeV”, *Phys. Lett. B* **791** (2019) 96,
doi:10.1016/j.physletb.2018.12.073, arXiv:1806.05246.
- [22] ATLAS Collaboration, “Cross-section measurements of the Higgs boson decaying into a pair of τ -leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *Phys. Rev D.* **99** (2019) 072001, doi:10.1103/PhysRevD.99.072001,
arXiv:1811.08856.
- [23] ATLAS Collaboration, “Observation of $H \rightarrow b\bar{b}$ decays and VH production with the ATLAS detector”, *Phys. Lett. B* **786** (2018) 59,
doi:10.1016/j.physletb.2018.09.013, arXiv:1808.08238.
- [24] ATLAS Collaboration, “Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector”, *Phys. Lett. B* **784** (2018) 173,
doi:10.1016/j.physletb.2018.07.035, arXiv:1806.00425.
- [25] ATLAS Collaboration, “Measurements of gluon-gluon fusion and vector-boson fusion Higgs boson production cross-sections in the $H \rightarrow WW^* \rightarrow e\nu\mu\nu$ decay channel in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *Phys. Lett. B* **789** (2019) 508,
doi:10.1016/j.physletb.2018.11.064, arXiv:1808.09054.

- [26] ATLAS Collaboration, “Measurement of the Higgs boson coupling properties in the $H \rightarrow ZZ^* \rightarrow 4\ell$ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *JHEP* **03** (2018) 095, doi:10.1007/JHEP03(2018)095, arXiv:1712.02304.
- [27] ATLAS Collaboration, “Measurements of Higgs boson properties in the diphoton decay channel with 36fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector”, *Phys. Rev. D* **98** (2018) 052005, doi:10.1103/PhysRevD.98.052005, arXiv:1802.04146.
- [28] ATLAS Collaboration, “Combined measurements of Higgs boson production and decay using up to 80fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS experiment”, *Phys. Rev. D* **101** (2020), no. 1, 012002, doi:10.1103/PhysRevD.101.012002, arXiv:1909.02845.
- [29] ATLAS, CMS Collaboration, “Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV”, *JHEP* **08** (2016) 045, doi:10.1007/JHEP08(2016)045, arXiv:1606.02266.
- [30] CMS Collaboration, “Observation of $t\bar{t}H$ production”, *Phys. Rev. Lett.* **120** (2018) 231801, doi:10.1103/PhysRevLett.120.231801, arXiv:1804.02610.
- [31] CMS Collaboration, “Observation of Higgs boson decay to bottom quarks”, *Phys. Rev. Lett.* **121** (2018) 121801, doi:10.1103/PhysRevLett.121.121801, arXiv:1808.08242.
- [32] CMS Collaboration, “Combined measurements of Higgs boson couplings in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Eur. Phys. J. C* **79** (2019) 421, doi:10.1140/epjc/s10052-019-6909-y, arXiv:1809.10733.
- [33] K. G. Wilson, “The Renormalization Group and Strong Interactions”, *Phys. Rev. D* **3** (1971) 1818, doi:10.1103/PhysRevD.3.1818.
- [34] G. 't Hooft et al., eds., “Recent Developments in Gauge Theories. Proceedings, Nato Advanced Study Institute, Cargese, France, August 26 - September 8, 1979”, volume 59. (1980). doi:10.1007/978-1-4684-7571-5.
- [35] J. Wess and B. Zumino, “A Lagrangian Model Invariant Under Supergauge Transformations”, *Phys. Lett. B* **49** (1974) 52, doi:10.1016/0370-2693(74)90578-4.
- [36] S. P. Martin, “A Supersymmetry primer”, *Adv. Ser. Direct. High Energy Phys.* **21** (2010) 1–153, doi:10.1142/9789812839657_0001, arXiv:hep-ph/9709356.
- [37] D. B. Kaplan and H. Georgi, “SU(2) x U(1) Breaking by Vacuum Misalignment”, *Phys. Lett. B* **136** (1984) 183–186, doi:10.1016/0370-2693(84)91177-8.

BIBLIOGRAPHY

- [38] B. Bellazzini, C. Csáki, and J. Serra, “Composite Higgses”, *Eur. Phys. J. C* **74** (2014), no. 5, 2766, doi:10.1140/epjc/s10052-014-2766-x, arXiv:1401.2457.
- [39] G. Panico and A. Wulzer, “The Composite Nambu-Goldstone Higgs”, volume 913. Springer, 2016. doi:10.1007/978-3-319-22617-0.
- [40] C. Csaki, C. Grojean, and J. Terning, “Alternatives to an Elementary Higgs”, *Rev. Mod. Phys.* **88** (2016), no. 4, 045001, doi:10.1103/RevModPhys.88.045001, arXiv:1512.00468.
- [41] CMS Collaboration, “Search for the Higgs boson decaying to two muons in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Phys. Rev. Lett.* **122** (2019) 021801, doi:10.1103/PhysRevLett.122.021801, arXiv:1807.06325.
- [42] ATLAS Collaboration, “A search for the dimuon decay of the Standard Model Higgs boson with the ATLAS detector”, *Phys. Lett. B* **812** (2021) 135980, doi:10.1016/j.physletb.2020.135980, arXiv:2007.07830.
- [43] ATLAS, CMS Collaboration, “Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments”, *Phys. Rev. Lett.* **114** (2015) 191803, doi:10.1103/PhysRevLett.114.191803, arXiv:1503.07589.
- [44] F. Halzen, “Quarks and Leptons: An Introductory Course in Modern Particle Physics”. Wiley, Chichester, England, UK, Jan, 1984. ISBN 978-0-47188741-6.
- [45] P. Schmöser, “Feynman-Graphen und Eichtheorien für Experimentalphysiker”. Springer, Berlin, Germany, Oct, 2013. ISBN 978-3-54058486-5.
- [46] M. D. Schwartz, “Quantum Field Theory and the Standard Model”. Cambridge University Press, Cambridge, England, UK, Mar, 2014. ISBN 978-1-10703473-0.
- [47] T. Lancaster, “Quantum Field Theory for the Gifted Amateur”. Oxford University Press, Oxford, England, UK, Jun, 2015. ISBN 978-0-19969933-9.
- [48] D. Griffiths, “Introduction to Elementary Particles”. Wiley-VCH, Weinheim, Germany, Aug, 2008. ISBN 978-3-52740601-2.
- [49] C. Sander and A. Schmidt, “Advanced Particle Physics”. lecture notes.
- [50] D. Galbraith, “Standard Model of the Standard Model”. <http://davidgalbraith.org/portfolio/ux-standard-model-of-the-standard-model/>.
- [51] C. S. Wu et al., “Experimental test of parity conservation in beta decay”, *Phys. Rev.* **105** (Feb, 1957) 1413–1415, doi:10.1103/PhysRev.105.1413.
- [52] N. Cabibbo, “Unitary Symmetry and Leptonic Decays”, *Phys. Rev. Lett.* **10** (1963) 531–533, doi:10.1103/PhysRevLett.10.531.

- [53] M. Kobayashi and T. Maskawa, “CP Violation in the Renormalizable Theory of Weak Interaction”, *Prog. Theor. Phys.* **49** (1973) 652–657, doi:10.1143/PTP.49.652.
- [54] Particle Data Group et al., “Review of Particle Physics”, *Progress of Theoretical and Experimental Physics* **2020** (08, 2020) doi:10.1093/ptep/ptaa104.
- [55] Z. Maki, M. Nakagawa, and S. Sakata, “Remarks on the unified model of elementary particles”, *Prog. Theor. Phys.* **28** (1962) 870–880, doi:10.1143/PTP.28.870.
- [56] M. Persic, P. Salucci, and F. Stel, “The Universal rotation curve of spiral galaxies: 1. The Dark matter connection”, *Mon. Not. Roy. Astron. Soc.* **281** (1996) 27, doi:10.1093/mnras/278.1.27, arXiv:astro-ph/9506004.
- [57] D. Clowe et al., “A direct empirical proof of the existence of dark matter”, *Astrophys. J. Lett.* **648** (2006) L109–L113, doi:10.1086/508162, arXiv:astro-ph/0608407.
- [58] T. Abe, R. Sato, and K. Yagyu, “Muon specific two-Higgs-doublet model”, *JHEP* **07** (2017) 012, doi:10.1007/JHEP07(2017)012, arXiv:1705.01469.
- [59] D. Ghosh, R. S. Gupta, and G. Perez, “Is the Higgs Mechanism of Fermion Mass Generation a Fact? A Yukawa-less First-Two-Generation Model”, *Phys. Lett. B* **755** (2016) 504–508, doi:10.1016/j.physletb.2016.02.059, arXiv:1508.01501.
- [60] A. Dery, A. Efrati, Y. Hochberg, and Y. Nir, “What if $B(h \rightarrow \mu\mu)/B(h \rightarrow \tau\tau) \neq m_\mu^2/m_\tau^2$?”, *JHEP* **05** (2013) 039, doi:10.1007/JHEP05(2013)039, arXiv:1302.3229.
- [61] Nobel Prize, “The Nobel Prize in Physics 2013”.
<https://www.nobelprize.org/prizes/physics/2013/summary>.
- [62] LHC Higgs cross section working group, “SM Higgs boson properties”.
https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCHWG?redirectedfrom=LHCPhysics.LHCHXSWG#SM_Higgs.
- [63] “LHC Machine”, *JINST* **3** (2008) S08001, doi:10.1088/1748-0221/3/08/S08001.
- [64] “LHC Design Report Vol.1: The LHC Main Ring”, doi:10.5170/CERN-2004-003-V-1.
- [65] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider”, *JINST* **3** (2008) S08003, doi:10.1088/1748-0221/3/08/S08003.
- [66] CMS Collaboration, “The CMS experiment at the CERN LHC”, *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.
- [67] LHCb Collaboration, “The LHCb Detector at the LHC”, *JINST* **3** (2008) S08005, doi:10.1088/1748-0221/3/08/S08005.

BIBLIOGRAPHY

- [68] ALICE Collaboration, “The ALICE experiment at the CERN LHC”, *JINST* **3** (2008) doi:10.1088/1748-0221/3/08/S08002.
- [69] “LEP design report v1”, technical report, Geneva, 1983.
- [70] “LEP design report v2”, technical report, Geneva, 1984.
- [71] R. Scrivens et al, “Overview of the status and developments on primary ion sources at CERN*”, Sep, 2011. <https://cds.cern.ch/record/1382102>.
- [72] E. Mobs, “The CERN accelerator complex - 2019. Complexe des accélérateurs du CERN - 2019”, Jul, 2019. <https://cds.cern.ch/record/2684277>.
- [73] CMS Collaboration, “Luminosity - Public Results”. https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#Multi_year_plots.
- [74] CMS Collaboration, “CMS Luminosity Measurements for the 2016 Data Taking Period”, technical report CMS-PAS-LUM-17-001, CERN, Geneva, 2017.
- [75] CMS Collaboration, “CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV”, technical report CMS-PAS-LUM-17-004, CERN, Geneva, 2018.
- [76] CMS Collaboration, “CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV”, technical report CMS-PAS-LUM-18-002, CERN, Geneva, 2019.
- [77] T. Sakuma and T. McCauley, “Detector and Event Visualization with SketchUp at the CMS Experiment”, *J. Phys. Conf. Ser.* **513** (2014) 022032, doi:10.1088/1742-6596/513/2/022032, arXiv:1311.4942.
- [78] I. Neutelings, “CMS coordinate system”. https://wiki.physik.uzh.ch/cms/latex:example_spherical_coordinates.
- [79] V. Veszpremi, “Operation and performance of the CMS tracker”, *JINST* **9** (2014) C03005, doi:10.1088/1748-0221/9/03/c03005.
- [80] CMS Collaboration, “The CMS tracker system project: Technical Design Report”, technical report, Geneva, 1997.
- [81] CMS Collaboration, “The CMS tracker: addendum to the Technical Design Report”, technical report, Geneva, 2000.
- [82] A. Dominguez, “The CMS pixel detector”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **581** (2007), no. 1, 343 – 346, doi:<https://doi.org/10.1016/j.nima.2007.07.156>.
- [83] CMS Collaboration, “CMS Technical Design Report for the Pixel Detector Upgrade”, doi:10.2172/1151650.

- [84] J.-L. Agram, “CMS silicon strip tracker performance”, *Physics Procedia* **37** (2012) 844 – 850, doi:<https://doi.org/10.1016/j.phpro.2012.02.423>. Proceedings of the 2nd Internat.Conference on Technology and Instrumentation in Particle Physics.
- [85] CMS Collaboration, “The CMS electromagnetic calorimeter project: Technical Design Report”, technical report, Geneva, 1997.
- [86] CMS Collaboration, “Changes to CMS ECAL electronics: addendum to the Technical Design Report”, technical report, Geneva, 2002.
- [87] A. Zghiche, “Performance of the CMS electromagnetic calorimeter during the LHC run II”, *Journal of Physics: Conference Series* **1162** (jan, 2019) 012001, doi:[10.1088/1742-6596/1162/1/012001](https://doi.org/10.1088/1742-6596/1162/1/012001).
- [88] CMS Collaboration, “Energy resolution of the barrel of the CMS electromagnetic calorimeter”, *JINST* **2** (2007), no. 04, P04004–P04004, doi:[10.1088/1748-0221/2/04/p04004](https://doi.org/10.1088/1748-0221/2/04/p04004).
- [89] A. Benaglia, “The CMS ECAL performance with examples”, *JINST* **9** (2014) doi:[10.1088/1748-0221/9/02/C02008](https://doi.org/10.1088/1748-0221/9/02/C02008).
- [90] CMS Collaboration, “The CMS hadron calorimeter project: Technical Design Report”, technical report, Geneva, 1997.
- [91] V. D. Elvira, “Measurement of the Pion Energy Response and Resolution in the CMS HCAL Test Beam 2002 Experiment”, technical report CMS-NOTE-2004-020, CERN, Geneva, Sep, 2004.
- [92] CMS, ECAL/HCAL Collaboration, “The CMS barrel calorimeter response to particle beams from 2-GeV/c to 350-GeV/c”, *Eur. Phys. J. C* **60** (2009) 359–373, doi:[10.1140/epjc/s10052-009-0959-5](https://doi.org/10.1140/epjc/s10052-009-0959-5). [Erratum: *Eur.Phys.J.C* 61, 353–356 (2009)].
- [93] CMS Collaboration, “The CMS muon project: Technical Design Report”, technical report, Geneva, 1997.
- [94] M. Fouz, “The CMS Muon System”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **446** (2000), no. 1, 366 – 372, doi:[https://doi.org/10.1016/S0168-9002\(00\)00082-6](https://doi.org/10.1016/S0168-9002(00)00082-6).
- [95] CMS Collaboration, “CMS reconstruction improvement for the muon tracking by the RPC chambers”, *PoS RPC2012* (2012) 045, doi:[10.1088/1748-0221/8/03/T03001](https://doi.org/10.1088/1748-0221/8/03/T03001), arXiv:1209.2646.
- [96] CMS Collaboration, “The CMS trigger system”, *JINST* **12** (2017), no. 01, P01020, doi:[10.1088/1748-0221/12/01/P01020](https://doi.org/10.1088/1748-0221/12/01/P01020), arXiv:1609.02366.

BIBLIOGRAPHY

- [97] CMS Collaboration, “The CMS magnet project: Technical Design Report”, technical report, Geneva, 1997.
- [98] CMS Collaboration, “CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems”, technical report.
- [99] CMS Trigger and Data Acquisition Group Collaboration, “The CMS high level trigger”, *Eur. Phys. J.* **C46** (2006) 605–667, doi:10.1140/epjc/s2006-02495-8, arXiv:hep-ex/0512077.
- [100] S. Drell and T.-M. Yan, “Partons and their Applications at High-Energies”, *Annals Phys.* **66** (1971) 578, doi:10.1016/0003-4916(71)90071-6.
- [101] J. C. Collins and D. E. Soper, “The Theorems of Perturbative QCD”, *Ann. Rev. Nucl. Part. Sci.* **37** (1987) 383–409, doi:10.1146/annurev.ns.37.120187.002123.
- [102] J. C. Collins, D. E. Soper, and G. F. Sterman, “Factorization for Short Distance Hadron - Hadron Scattering”, *Nucl. Phys. B* **261** (1985) 104–142, doi:10.1016/0550-3213(85)90565-6.
- [103] J. C. Collins, D. E. Soper, and G. F. Sterman, “Soft Gluons and Factorization”, *Nucl. Phys. B* **308** (1988) 833–856, doi:10.1016/0550-3213(88)90130-7.
- [104] GEANT4 Collaboration, “GEANT4—a simulation toolkit”, *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.
- [105] S. Alioli et al., “Monte Carlo event generators for high energy particle physics event simulation”, arXiv:1902.01674.
- [106] A. Buckley et al., “General-purpose event generators for LHC physics”, *Phys. Rept.* **504** (2011) 145–233, doi:10.1016/j.physrep.2011.03.005, arXiv:1101.2599.
- [107] N. Metropolis and S. Ulam, “The monte carlo method”, *Journal of the American Statistical Association* **44** (1949), no. 247, 335–341.
- [108] J. Alwall et al., “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”, *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.
- [109] J. Alwall et al., “MadGraph 5 : Going Beyond”, *JHEP* **06** (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522.
- [110] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”, *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.
- [111] P. Nason, “A new method for combining NLO QCD with shower Monte Carlo algorithms”, *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

- [112] S. Alioli, P. Nason, C. Oleari, and E. Re, “A general framework for implementing NLO calculations in shower monte carlo programs: the POWHEG BOX”, *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.
- [113] T. Sjostrand, S. Mrenna, and P.Z. Skands, “A Brief Introduction to PYTHIA 8.1”, *Comput. Phys. Commun.* **178** (2008) 852, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.
- [114] CMS Collaboration, “Event generator tunes obtained from underlying event and multiparton scattering measurements”, *Eur. Phys. J. C* **76** (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.
- [115] CMS Collaboration, “Extraction and validation of a new set of CMS PYTHIA 8 tunes from underlying-event measurements”, *Eur. Phys. J. C* **80** (2020) 4, doi:10.1140/epjc/s10052-019-7499-4, arXiv:1903.12179.
- [116] J. Alwall et al., “Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions”, *Eur. Phys. J. C* **53** (2008) 473, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.
- [117] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO”, *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.
- [118] NNPDF Collaboration, “Parton distributions for the LHC Run II”, *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.
- [119] NNPDF Collaboration, “Parton distributions from high-precision collider data”, *Eur. Phys. J. C* **77** (2017) 663, doi:10.1140/epjc/s10052-017-5199-5, arXiv:1706.00428.
- [120] LHC Higgs Cross Section Working Group, “Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector”, *CERN* (2016) doi:10.23731/CYRM-2017-002, arXiv:1610.07922.
- [121] C. Anastasiou et al., “High precision determination of the gluon fusion Higgs boson cross-section at the LHC”, *JHEP* **05** (2016) 058, doi:10.1007/JHEP05(2016)058, arXiv:1602.00695.
- [122] K. Hamilton, P. Nason, E. Re, and G. Zanderighi, “NNLOPS simulation of Higgs boson production”, *JHEP* **10** (2013) 222, doi:10.1007/JHEP10(2013)222, arXiv:1309.0017.
- [123] K. Hamilton, P. Nason, and G. Zanderighi, “Finite quark-mass effects in the NNLOPS POWHEG+MiNLO Higgs generator”, *JHEP* **05** (2015) 140, doi:10.1007/JHEP05(2015)140, arXiv:1501.04637.

BIBLIOGRAPHY

- [124] P. Nason and C. Oleari, “NLO Higgs boson production via vector-boson fusion matched with shower in POWHEG”, *JHEP* **02** (2010) 037, doi:10.1007/JHEP02(2010)037, arXiv:0911.5299.
- [125] G. Luisoni, P. Nason, C. Oleari, and F. Tramontano, “ $HW^\pm/HZ + 0$ and 1 jet at NLO with the POWHEG BOX interfaced to GoSam and their merging within MiNLO”, *JHEP* **10** (2013) 083, doi:10.1007/JHEP10(2013)083, arXiv:1306.2542.
- [126] A. Djouadi, J. Kalinowski, and M. Spira, “HDECAY: A Program for Higgs boson decays in the standard model and its supersymmetric extension”, *Comput. Phys. Commun.* **108** (1998) 56, doi:10.1016/S0010-4655(97)00123-9, arXiv:hep-ph/9704448.
- [127] M. Spira, “QCD effects in Higgs physics”, *Fortsch. Phys.* **46** (1998) 203, doi:10.1002/(SICI)1521-3978(199804)46:3<203::AID-PROP203>3.0.CO;2-4, arXiv:hep-ph/9705337.
- [128] M. Czakon and A. Mitov, “Top++: A program for the calculation of the top-pair cross-section at hadron colliders”, *Comput. Phys. Commun.* **185** (2014) 2930, doi:10.1016/j.cpc.2014.06.021, arXiv:1112.5675.
- [129] P. Kant et al., “HATHOR for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions”, *Comput. Phys. Commun.* **191** (2015) 74, doi:10.1016/j.cpc.2015.02.001, arXiv:1406.4403.
- [130] CMS Collaboration, “Measurement of the cross section for top quark pair production in association with a W or Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JHEP* **08** (2018) 011, doi:10.1007/JHEP08(2018)011, arXiv:1711.02547.
- [131] Y. Li and F. Petriello, “Combining QCD and electroweak corrections to dilepton production in FEWZ”, *Phys. Rev D.* **86** (2012) 094034, doi:10.1103/PhysRevD.86.094034, arXiv:1208.5967.
- [132] M. Grazzini, S. Kallweit, D. Rathlev, and M. Wiesemann, “ $W^\pm Z$ production at the LHC: fiducial cross sections and distributions in NNLO QCD”, *JHEP* **05** (2017) 139, doi:10.1007/JHEP05(2017)139, arXiv:1703.09065.
- [133] M. Grazzini, S. Kallweit, and D. Rathlev, “ZZ production at the LHC: fiducial cross sections and distributions in NNLO QCD”, *Phys. Lett. B* **750** (2015) 407, doi:10.1016/j.physletb.2015.09.055, arXiv:1507.06257.
- [134] T. Gehrmann et al., “ W^+W^- production at hadron colliders in Next to Next to Leading Order QCD”, *Phys. Rev. Lett.* **113** (2014) 212001, doi:10.1103/PhysRevLett.113.212001, arXiv:1408.5243.

- [135] CMS Collaboration, “Particle-flow reconstruction and global event description with the CMS detector”, *JINST* **12** (2017), no. 10, P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.
- [136] CMS Collaboration, “Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET”, technical report CMS-PAS-PFT-09-001, CERN, Geneva, Apr, 2009.
- [137] CMS Collaboration, “Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector”, technical report CMS-PAS-PFT-10-001, CERN, 2010.
- [138] D. Barney, “CMS Detector Slice”, Jan, 2016. CMS Collection. <https://cds.cern.ch/record/2120661>.
- [139] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, *JINST* **9** (2014), no. 10, P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.
- [140] P. Billoir, “Progressive track recognition with a kalman-like fitting procedure”, *Computer Physics Communications* **57** (1989), no. 1, 390 – 394, doi:10.1016/0010-4655(89)90249-X.
- [141] P. Billoir and S. Qian, “Simultaneous pattern recognition and track fitting by the kalman filtering method”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **294** (1990), no. 1, 219 – 228, doi:10.1016/0168-9002(90)91835-Y.
- [142] R. Mankel, “A concurrent track evolution algorithm for pattern recognition in the hera-b main tracking system”, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **395** (1997), no. 2, 169 – 184, doi:10.1016/S0168-9002(97)00705-5.
- [143] R. Fruhwirth, “Application of Kalman filtering to track and vertex fitting”, *Nucl. Instrum. Meth.* **A262** (1987) 444–450, doi:10.1016/0168-9002(87)90887-4.
- [144] CMS Collaboration, “Mini-AOD: A New Analysis Data Format for CMS. A new analysis data format for CMS”, technical report CMS-CR-2015-052. 7, CERN, Geneva, May, 2015. doi:10.1088/1742-6596/664/7/072052, 6 pages, 4 figures, presented at the 21st International Conference on Computing in High Energy and Nuclear Physics (CHEP2015).
- [145] T. Speeret et al., “Vertex Fitting in the CMS Tracker”, technical report CMS-NOTE-2006-032, CERN, Geneva, Feb, 2006.
- [146] E. Chabanat and N. Estre, “Deterministic Annealing for Vertex Finding at CMS”, doi:10.5170/CERN-2005-002.287.

BIBLIOGRAPHY

- [147] R. Frühwirth, W. Waltenberger, and P. Vanlaer, “Adaptive Vertex Fitting”, technical report CMS-NOTE-2007-008, CERN, Geneva, Mar, 2007.
- [148] CMS Collaboration, “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JINST* **13** (2018), no. 06, P06015, doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.
- [149] S. Baffioni et al., “Electron reconstruction in CMS”, doi:10.1140/epjc/s10052-006-0175-5.
- [150] CMS Collaboration, “Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV.”, *JINST* **10** (Feb, 2015) P06005. 63 p, doi:10.1088/1748-0221/10/06/P06005.
- [151] W. Adam, R. Frühwirth, A. Strandlie, and T. Todor, “Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC”, technical report CMS-NOTE-2005-001, CERN, Geneva, Jan, 2005.
- [152] R. Frühwirth, “Track fitting with non-Gaussian noise”, *Comput. Phys. Commun.* **100** (1997) 1–16, doi:10.1016/S0010-4655(96)00155-5.
- [153] A. Dishaw et al., “Alternatives to Standard Isolation for 13 TeV”.
- [154] Folguras, Xiw, and Calderon, “Isolation Studies”. https://indico.cern.ch/event/188494/contributions/338726/attachments/264683/370707/IsolationStudies_Apr30_MuonPOG.pdf.
- [155] S. D. Ellis and D. E. Soper, “Successive combination jet algorithm for hadron collisions”, *Phys. Rev. D* **48** (Oct, 1993) 3160–3166, doi:10.1103/PhysRevD.48.3160.
- [156] M. Wobisch and T. Wengler, “Hadronization corrections to jet cross-sections in deep inelastic scattering”, in *Workshop on Monte Carlo Generators for HERA Physics (Plenary Starting Meeting)*, pp. 270–279. 4, 1998. arXiv:hep-ph/9907280.
- [157] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- k_T jet clustering algorithm”, *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
- [158] CMS Collaboration, “Jet identification”. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID>.
- [159] CMS Collaboration, “Performance of the pile up jet identification in CMS for Run 2”, Apr, 2020. <https://cds.cern.ch/record/2715906>.
- [160] CMS Collaboration, “Pileup mitigation at CMS in 13 TeV data”, technical report CMS-PAS-JME-18-001, CERN, Geneva, 2019. arXiv:2003.00503.
- [161] V. Khachatryan et al., “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV”, *Journal of Instrumentation* **12** (feb, 2017) P02014–P02014, doi:10.1088/1748-0221/12/02/p02014.

- [162] CMS JetMET POG, “Reduced set of jet energy uncertainties for Run2 searches”.
https://twiki.cern.ch/twiki/bin/viewauth/CMS/JECUncertaintySources#Run_2_reduced_set_of_uncertainty.
- [163] CMS Collaboration, “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV”, *JINST* **12** (2017), no. 02, P02014,
doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.
- [164] “Jet energy resolution”, Feb, 2021. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution?rev=72>.
- [165] CMS Collaboration, “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV”, *JINST* **13** (2018) P05011,
doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.
- [166] D. Guest et al., “Jet flavor classification in high-energy physics with deep neural networks”, *Phys. Rev. D* **94** (Dec, 2016) 112002,
doi:10.1103/PhysRevD.94.112002.
- [167] CMS Collaboration, “Recommendations for b-tag objects in Physics Analyses”. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation>.
- [168] CMS Collaboration, “Performance of quark/gluon discrimination using pp collision data at $\sqrt{s}=8$ TeV”, CMS Physics Analysis Summary CMS-PAS-JME-13-002, CERN, 2013.
- [169] CMS Collaboration, “Performance of quark/gluon discrimination in 13 TeV data”, technical report, Nov, 2016.
- [170] CMS Collaboration, “Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector”, *JINST* **14** (2019), no. 07, P07004, doi:10.1088/1748-0221/14/07/P07004,
arXiv:1903.06078.
- [171] CMS JetMET POG, “Recommended MET filters”. <https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFiltersRun2>.
- [172] CMS Collaboration, “Evidence for Higgs boson decay to a pair of muons”, *JHEP* **01** (2021) 148, doi:10.1007/JHEP01(2021)148, arXiv:2009.04363.
- [173] N. Adam et al., “Generic Tag and Probe Tool for Measuring Efficiency at CMS with Early Data”.
- [174] CMS Collaboration, “Performance of the CMS muon trigger reconstruction during Run 2 of the LHC”, technical report.
- [175] A. M. Sirunyan et al., “Measurement of top quark pair production in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Journal of High Energy Physics* **2020** (Mar, 2020) doi:10.1007/jhep03(2020)056.

BIBLIOGRAPHY

- [176] CMS Collaboration, “Observation of single top quark production in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13\text{ TeV}$ ”, *Phys. Rev. Lett.* **122** (2019) 132003, doi:10.1103/PhysRevLett.122.132003, arXiv:1812.05900.
- [177] CMS Muon POG, “Muon reference efficiencies 2016”, Feb, 2021. <https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2016LegacyRereco>.
- [178] CMS Muon POG, “Muon reference trigger efficiencies”, Feb, 2021. <https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffsRun2Legacy>.
- [179] CMS Muon POG, “Muon reference efficiencies 2018”, Feb, 2021. <https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2018>.
- [180] CMS Muon POG, “Muon reference efficiencies 2017”, Feb, 2021. <https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs2017>.
- [181] CMS JetMET POG, “L1 ECAL prefiring corrections”. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/L1ECALPrefiringWeightRecipe>.
- [182] CMS BTV POG, “B-tagging discriminant shape calibration using event weights with a tag-and-probe method”. <https://twiki.cern.ch/twiki/bin/view/CMS/BTagShapeCalibration>.
- [183] CMS JetMET POG, “Quark-Gluon Likelihood uncertainty prescription”. <https://twiki.cern.ch/twiki/bin/viewauth/CMS/QuarkGluonLikelihood#Systematics>.
- [184] A. Bodek et al., “Extracting muon momentum scale corrections for hadron collider experiments”, *Eur. Phys. J. C* **72** (2012) 2194, doi:10.1140/epjc/s10052-012-2194-8, arXiv:1208.3710.
- [185] CMS Collaboration, “Search for Higgs boson decays into muons in proton-proton collisions at 13 TeV using full Run2 data”, technical report AN-19-090, Geneva, 2019.
- [186] CMS Collaboration, “Search for the Higgs boson decaying to two muons in pp collisions at 13 TeV using the full Run 2 data”, technical report AN-19-185, Geneva, 2019.
- [187] Riverside-Rome-Rutherford-Saclay(CEN)-Victoria-Vienna-Wisconsin Collaboration, Aachen-Amsterdam(NIKHEF)-Annecy-Birmingham-CERN-Harvard-Helsinki-Kiel-London(IC)-London(QMC)-Madrid(CIEMAT)-MIT-Padua-Paris(CdF)-Rutherford-Saclay(CEN) Collaboration, “Studies of intermediate vector boson production and decay in UA1 at the CERN proton-antiproton collider”, *Z. Phys. C* **44** (Nov, 1988) 15–61. 135 p, doi:10.1007/BF01548582.
- [188] UA1 Collaboration, J. Timmer, “The UA1 Detector”, in *3rd Moriond Workshop: Antiproton-Proton Physics*, pp. 593–608. 1983.

- [189] K. Jakobs, “The physics results of the UA2 experiment at the CERN $p\bar{p}$ collider”, *Int. J. Mod. Phys. A* **9** (Jan, 1994) 2903–2977. 78 p, doi : 10.1142/S0217751X94001163.
- [190] CMS Collaboration, “Study of Final-State Radiation in Decays of Z Bosons Produced in pp Collisions at 7 TeV”, *Phys. Rev. D* **91** (2015), no. 9, 092012, doi : 10.1103/PhysRevD.91.092012, arXiv : 1502.07940.
- [191] CMS Collaboration, “Evidence for a new state in the search for the standard model Higgs boson in the H to ZZ to 4 leptons channel in pp collisions at $\sqrt{s} = 7$ and 8 TeV”, technical report CMS-PAS-HIG-12-016, CERN, Geneva, 2012.
- [192] CMS Collaboration, “Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s}=13$ TeV”, *JHEP* **11** (2017) 047, doi : 10.1007/JHEP11(2017)047, arXiv : 1706.09936.
- [193] CMS Collaboration, “Search for the decay of a Higgs boson in the $\ell\ell\gamma$ channel in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *JHEP* **11** (2018) 152, doi : 10.1007/JHEP11(2018)152, arXiv : 1806.05996.
- [194] G. Aad et al., “A search for the $z\gamma$ decay mode of the higgs boson in pp collisions at 13 TeV with the ATLAS detector”, *Physics Letters B* **809** (2020) 135754, doi : <https://doi.org/10.1016/j.physletb.2020.135754>.
- [195] X. Zuo, “A Search for the Standard Model Higgs Boson Decaying into Two Muons in the Vector Boson Associated Production Mode at the CMS Experiment”, 2021. <https://cds.cern.ch/record/2790970>.
- [196] CMS Collaboration, “Measurement of the inclusive and differential cross section of top pair-production in association with a Z boson at 13 TeV”, technical report.
- [197] A. Hoecker et al., “TMVA - Toolkit for multivariate data analysis”, (2007). arXiv : physics/0703039.
- [198] H. Voss, A. Höcker, J. Stelzer, and F. Tegenfeldt, “TMVA, the Toolkit for Multivariate Data Analysis with ROOT”, *PoS ACAT* (2009) 040, doi : 10.22323/1.050.0040.
- [199] F. Chollet et al., “Keras”, 2015. Software available from keras.io. <https://keras.io>.
- [200] M. Abadi et al., “TensorFlow: Large-scale machine learning on heterogeneous systems”, 2015. Software available from tensorflow.org. <http://tensorflow.org/>.
- [201] J. D’Hondt et al., “Fitting of Event Topologies with External Kinematic Constraints in CMS”, technical report CMS-NOTE-2006-023, CERN, Geneva, Jan, 2006.
- [202] CMS Collaboration, “Search for dark matter particles produced in association with a top quark pair at $\sqrt{s} = 13$ TeV”, *Phys. Rev. Lett.* **122** (Jan, 2019) 011803, doi : 10.1103/PhysRevLett.122.011803.

BIBLIOGRAPHY

- [203] A. M. Sirunyan et al., “Measurement of the top quark mass in the all-jets final state at $\sqrt{s} = 13$ TeV and combination with the lepton+jets channel”, doi:10.1140/epjc/s10052-019-6788-2.
- [204] J. Lange, “Measurement of the top quark mass in the all-jets final state at $\sqrt{s} = 13$ TeV and combination with the lepton+jets channel”. Dissertation, Universität Hamburg, Hamburg, 2020. Dissertation, Universität Hamburg, 2020. doi:10.3204/PUBDB-2020-00969.
- [205] L. Demortier and L. Lyons, “Everything you always wanted to know about pulls”, technical report.
- [206] J. C. Collins and D. E. Soper, “Angular distribution of dileptons in high-energy hadron collisions”, *Phys. Rev. D* **16** (1977) 2219, doi:10.1103/PhysRevD.16.2219.
- [207] ATLAS Collaboration, “Search for the neutral Higgs bosons of the minimal supersymmetric standard model”, *JHEP* **02** (2013) 095, doi:10.1007/JHEP02(2013)095, arXiv:1211.6956.
- [208] J. Smith, W. van Neerven, and J. Vermaseren, “The Transverse Mass and Width of the W Boson”, *Phys. Rev. Lett.* **50** (1983) 1738, doi:10.1103/PhysRevLett.50.1738.
- [209] CMS Collaboration, “Measurements of $t\bar{t}$ differential cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV using events containing two leptons”, *JHEP* **02** (2019) 149, doi:10.1007/JHEP02(2019)149, arXiv:1811.06625.
- [210] M. Oreglia, “A Study of the Reactions $\psi' \rightarrow \gamma\gamma\psi$ ”. PhD thesis, SLAC, 1980.
- [211] CMS JetMET POG, “Sources of uncertainty affecting jet energy scale in CMS”. <https://twiki.cern.ch/twiki/bin/view/CMS/JECUncertaintySources>.
- [212] LHC Higgs Cross Section Working Group Collaboration, “Handbook of LHC Higgs Cross Sections: 1. Inclusive Observables”, doi:10.5170/CERN-2011-002, arXiv:1101.0593.
- [213] LHC Higgs Cross Section Working Group Collaboration, “Handbook of LHC Higgs Cross Sections: 3. Higgs Properties”, doi:10.5170/CERN-2013-004, arXiv:1307.1347.
- [214] CMS Collaboration, “Roostats-based statistic tool for higgs pag”, CMS Twiki page.
- [215] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics”, *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727. [Erratum: *Eur. Phys. J. C* **73**, 2501 (2013)].
- [216] The ATLAS and CMS Collaborations, and the LHC Higgs Combination Group, “Procedure for the LHC Higgs boson search combination in Summer 2011”, technical report CMS-NOTE-2011-005, ATL-PHYS-PUB-2011-11, 2011.

- [217] T. Junk, “Confidence level computation for combining searches with small statistics”, *Nucl. Instrum. Meth. A* **434** (1999) 435, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.
- [218] A. L. Read, “Presentation of search results: the CL_s technique”, *J. Phys. G* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.
- [219] F. Schissler and D. Zeppenfeld, “Parton shower effects on W and Z production via vector boson fusion at NLO QCD”, *JHEP* **04** (2013) 057, doi:10.1007/JHEP04(2013)057, arXiv:1302.2884.
- [220] CMS Collaboration, “Electroweak production of two jets in association with a Z boson in proton-proton collisions at $\sqrt{s} = 13$ TeV”, *Eur. Phys. J. C* **78** (2018) 589, doi:10.1140/epjc/s10052-018-6049-9, arXiv:1712.09814.
- [221] B. Jäger et al., “Parton-shower effects in Higgs production via Vector-Boson Fusion”, (2020). arXiv:2003.12435.
- [222] P. D. Dauncey, M. Kenzie, N. Wardle, and G. J. Davies, “Handling uncertainties in background shapes”, *JINST* **10** (2015) P04015, doi:10.1088/1748-0221/10/04/P04015, arXiv:1408.6865.
- [223] CMS Collaboration, “A measurement of the Higgs boson mass in the diphoton decay channel”, *Phys. Lett. B* **805** (2020) 135425, doi:10.1016/j.physletb.2020.135425, arXiv:2002.06398.
- [224] ATLAS Collaboration, “Muon reconstruction performance of the ATLAS detector in proton-proton collision data at $\sqrt{s} = 13$ TeV”, *Eur. Phys. J. C* **76** (2016), no. 5, 292, doi:10.1140/epjc/s10052-016-4120-y, arXiv:1603.05598.
- [225] D. Contardo et al., “Technical Proposal for the Phase-II Upgrade of the CMS Detector”, technical report CERN-LHCC-2015-010. LHCC-P-008. CMS-TDR-15-02, Geneva, Jun, 2015. Upgrade Project Leader Deputies: Lucia Silvestris (INFN-Bari), Jeremy Mans (University of Minnesota) Additional contacts: Lucia.Silvestris@cern.ch, Jeremy.Mans@cern.ch.
- [226] CMS Collaboration, “The Phase-2 Upgrade of the CMS Tracker”, technical report CERN-LHCC-2017-009. CMS-TDR-014, CERN, Geneva, Jun, 2017.
- [227] CMS Collaboration, “The Phase-2 Upgrade of the CMS Barrel Calorimeters”, technical report CERN-LHCC-2017-011. CMS-TDR-015, CERN, Geneva, Sep, 2017. This is the final version, approved by the LHCC.
- [228] CMS Collaboration, “The Phase-2 Upgrade of the CMS Endcap Calorimeter”, technical report CERN-LHCC-2017-023. CMS-TDR-019, CERN, Geneva, Nov, 2017.

BIBLIOGRAPHY

- [229] CMS Collaboration, “The Phase-2 Upgrade of the CMS Muon Detectors”, technical report CERN-LHCC-2017-012. CMS-TDR-016, CERN, Geneva, Sep, 2017. This is the final version, approved by the LHCC.
- [230] CMS Collaboration, “Expected performance of the physics objects with the upgraded CMS detector at the HL-LHC”, technical report CMS-NOTE-2018-006. CERN-CMS-NOTE-2018-006, CERN, Geneva, Dec, 2018.
- [231] L. Rossi and O. Brüning, “Introduction to the HL-LHC Project”, *Adv. Ser. Direct. High Energy Phys.* **24** (2015) 1–17, doi : 10.1142/9789814675475_0001.
- [232] G. Apollinari et al., “High-Luminosity Large Hadron Collider (HL-LHC): Preliminary Design Report”. CERN Yellow Reports: Monographs. CERN, Geneva, 2015. doi : 10.5170/CERN-2015-005.

Acknowledgments

Throughout my time as a PhD student I have received a lot of support and assistance. I have been very fortunate to be part of the engaging working group at the University of Hamburg.

I would first and foremost like to thank Prof. Peter Schleper for giving me the opportunity to work on this highly interesting research topic. Your support and insightful feedback was invaluable in successfully completing this project.

Further I would like to acknowledge Prof. Elisabetta Gallo as second referee of this thesis, and for her patience and guidance.

I would also like to thank my supervisors Dr. Adrian Perieanu, Dr. Sam Bein and Dr. Hartmut Stadie for their careful review and important feedback from the initial stages to the final results.

I would like to acknowledge the whole CMS $H \rightarrow \mu^+\mu^-$ working group. Special thanks to Raffaele, Stephane, Vivek, Andrea M., Pier, Andrea R., Andrew, Darin, and Xunwu for the vivid discussions and all your help during all phases of the analysis.

In addition, I would like to thank my friends and colleagues Tobias, Viktor, Johannes, Moritz, Alex, Malte, Torben, Christoph, who provided stimulating discussions as well as happy distractions.

Finally, I could not have completed this dissertation without the constant support of my friends and family!

Eidesstattliche Versicherung

Declaration on oath

Hiermit versichere ich an Eidesstatt, die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt zu haben. Die eingereichte schriftliche Fassung entspricht der auf dem elektronischen Speichermedium.
Die Dissertation wurde in der vorgelegten oder einer ähnlichen Form nicht schon einmal in einem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt.

.....
Ort, Datum

.....
Jan Oliver Rieger