BUAP

Benemérita Universidad Autónoma de Puebla
Facultad de Ciencias Físico Matemáticas

Search for Additional Higgs Bosons with the CMS Detector at the CERN LHC

Jan Eysermans

A thesis submitted to
Posgrado en Física
in fulfillment of the requirements for the degree of
Doctorado en Física

Supervised by **Dr. María Isabel Pedraza Morales**

Puebla, Pue.
January 2019

Acknowledgments

The process of my PhD couldn't succeed with the right people besides me. Many people were directly or indirectly involved, some of which deserve a special note listed here.

In the first place I would like to thank my supervisor Isabel, for giving me many opportunities through the past years in order to develop myself in various topics in experimental physics. Your continuous support in Mexico and at CERN was valuable to me and you gave me the right directions to go through this career.

My parents, Moeke and Vake, my sisters, Ann, Iris, Inne and their cute kids. Coming home in Belgium surrounded by my family was always something to look forward to. I want to thank all of them for their listening and continuous support.

Markus, Mariarosaria and Andrea, a big thanks for allowing me to work intensively with you in the MIT group. My knowledge of physics, analysis and statistics have been shaped thanks to all the discussions and coffees in your office. Our long efforts to understand the excesses were understood and solved, unfortunately it didn't result in any signature of a charged Higgs boson!

Alejandra and Diana, whom I need to be grateful to for all the nice moments in Puebla and conferences. The coffees in the Italian was always to look something forward to as a break during the day. Not to forget for guiding and solving most of my paperwork!

Two special persons, Eline and Alessia, which were close to me and gave me guidance and strength at the right moments. You gave me the necessary space outside the work and and life was more joyful in Belgium and Switzerland with you in my life.

Abstract

Title. Search for Additional Higgs Bosons with the CMS Detector at the CERN LHC

Abstract. A search for charged Higgs bosons decaying into a tau lepton and a neutrino is presented in the hadronic and leptonic final states. Emphasis is given on the leptonic final states, performed in the muon/electron+jets channels. The search is distinguised based on whether an hadronic tau lepton can be resolved or not. To constrain the backgrounds, the event phase space is categorized in number of jets, b jets and in bins of E_T^{miss}. The search is based on the 13 TeV dataset with 35.9 fb^{-1} of integrated luminosity collected with the CMS experiment in 2016. Results are presented for charged Higgs boson mass hypotheses ranging from 80 GeV to 3 TeV, where the intermediate mass range around the top quark mass is included. 95% CL upper limits are set on the charged Higgs boson production cross section. A combination with the fully hadronic final state is performed, where combined 95% CL upper limits are set on the charged Higgs boson production cross section. The model independent result is interpreted in the MSSM m_h^{mod-} benchmark scenario.

Contents

List of Figures

List of Tables

1 The Standard Model and Beyond

The Standard Model (SM) is a theory describing subatomic particle structure and the interactions between elementary particles. It contains a precise mathematical description based on symmetries and group theory to describe the electromagnetic, weak and strong nuclear forces as well as a mechanism of particle mass generation. The theoretical development of the SM took place from the early fifties and continued to develop for several decades. Collider experiments throughout the last half century precisely tested the SM with an astonishing precision, which leads the SM. Despite its success up to which level the theory was tested experimentally, yet it does not include a description of gravity nor it can explain phenomena observed on the cosmological scale. Even within the SM there are questions which cannot be answered, ultimately the SM is still a descriptive model based on assumptions without explaining the underlying structure.

1.1 Particle Spectrum of the Standard Model

The physical building blocks of the SM is a set of elementary particles, regarded as point-like, all having a certain mass, quantum numbers and interaction rules governed by the elementary forces. A first classification is done based on the particle's spin: fermionic particles have half-integer spin and obey the Fermi-Dirac statistics whereas bosonic particles have integer spin values and obey the Bose-Einstein statistics. Fermions and bosons behave differently in quantum systems as well as on macroscopic level as the Pauli exclusion principle only applies to fermions, hence they are described by two different statistics. It must be noted that the fermionic or bosonic nature is true for composite particles or quantum states.

As will be explained later, each fundamental force has its own force carrier (e.g. the photon is the force carrier for the electromagnetic force). In the SM, force carriers are bosons, whereas the particles where upon the force interacts are fermions (e.g. electrons).

Based to their interaction nature and quantum properties, the SM elementary particles are classified as follows (see figure **1-1**):

- Leptons: leptons are fermions which do not undergo the strong interactions. There are

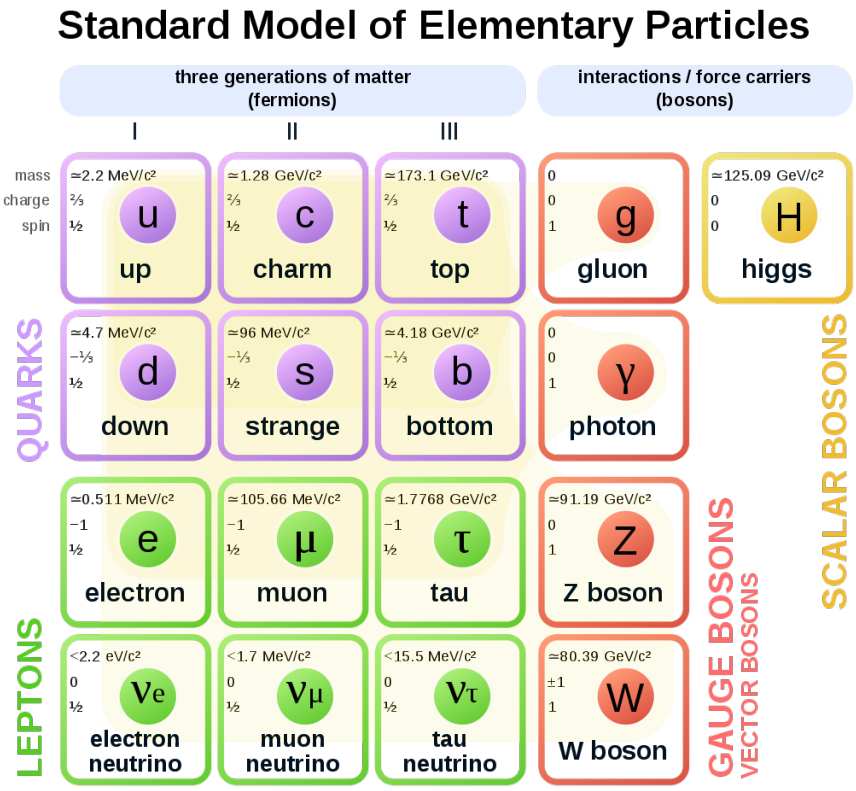


Figure **1-1**: Overview of the particle spectrum of the Standard Model.

three charged leptons (electron, muon and τ-lepton), and three neutral leptons, the neutrinos which correspond to the charged leptons (ν_e, ν_μ and ν_τ respectively).

- Quarks: quarks are fermions which are subjected to all the fundamental forces. Also six quarks exist, which are grouped in families by (up, down), (strange, charm) and (top, bottom).

- Vector bosons: for each fundamental force (except gravity), a vector boson exist which is acts as the force carrier.

- Scalar bosons: only one scalar boson exist which is responsible for the mass generation and electroweak symmetry breaking.

The particle spectrum of quarks and leptons is similar and are divided in three generations, which is justified by their mass and quantum numbers. For each of the fermions, also their anti-particles exist for which the charge change but all other quantum numbers remain the same.

1.2 Particle and Fields

Historically, the SM emerged from the unification of special relativity and quantum mechanics, which lead to a field description of the particles and forces. Each particle is represented as a complex field whose value depends on the spacetime coordinates: $\phi(x^\mu)$. The kinematics and dynamics of theory is captured in the so called Lagrangian density $\mathcal{L}$, for short Lagrangian, which is composed of different terms as function of the field and/or derivatives, each term describing the interactions and forces between particles. Field Lagrangians are similar to the classical Lagrangian in the sense that the derivatives in the Euler-Lagrange equations are not w.r.t. the general coordinates but directly to the fields themselves. By imposing basic principles for the terms in the Lagrangian (e.g. internal symmetries, gauge variance, action must be real, renormalizable, etc.), only a specific structure of terms is allowed. As a fermion with mass m is described by Dirac four-component spinor fields $\psi = \psi(x^\mu)$, which simultaneously describes particle-antiparticle and takes into account the spin. They are solutions of the free Dirac equation:

$$(i\gamma^\mu \partial_\mu - m)\psi = 0, \tag{1-1}$$

which follows from the Dirac Lagrangian:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi. \tag{1-2}$$

The fields entering the Lagrangian are quantized according to the laws of quantum mechanics (second quantization): the field is regarded as an operator which obeys a set of commutator rules, and is Fourier-expanded where the expansion coefficients ar creation and annihilation operators (cfr. quantum harmonic oscillator). The quanta of this quantized field are then identified with the individual particles. When interactions are involved between different fields, each term in the Lagrangian represents an interaction which can be depicted by a Feynman diagram.

1.3 Description of the Forces

1.3.1 Electromagnetic Force

Classically, the well-known electromagnetic force is described by the four Maxwell equations, which was a unification of the electric and magnetic forces by Maxwell in 1861-1862. Although these equations were discovered 40 years before the introduction of special relativity, the equations are invariant under a Lorentz transformation, as required by special relativity.

The Maxwell equations can be derived from field theory by using gauge theories. In short, a gauge theory is a (mathematical) transformation of the fields which does not alter the Lagrangian, hence changing the gauge does not change the physical properties of the Lagrangian. Starting from a (complex) scalar field $\phi = (\phi_1 + i\phi_2)/\sqrt{2}$, with $\phi_{1,2}$ two real fields, one can write down a Lagrangian:

$$\mathcal{L} = (\partial_\mu \phi)(\partial^\mu \phi^*) - m^2 \phi^* \phi. \tag{1-3}$$

This Lagrangian is invariant under a transformation $\phi \rightarrow \exp(-i\Lambda)\phi$ (*global* gauge transformation), with Λ a real constant, but such transformation clearly violates the principles of relativity as the gauge instantaneously should change for the entire space-time. To overcome this violation, the parameter Λ becomes as function of space-time $\Lambda(x^\mu)$ (*local* gauge transformation). In order to make eq. 1-3 local gauge invariant, extra terms must be added, which results in the following Lagrangian:

$$\mathcal{L}_{\text{QED}} = (D_\mu \phi)(D^\mu \phi^*) - m^2 \phi^* \phi - \tfrac{1}{4} F^{\mu\nu} F_{\mu\nu}, \tag{1-4}$$

with $D_\mu \phi = (\partial_\mu + ieA_\mu)\phi$ the covariant derivative and $F^{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ the electromagnetic field tensor which describes the electromagnetic fields and interactions (i.e. tensor formulation of the laws of Maxwell).

A gauge field A_μ and corresponding covariant derivative has been introduced to restore the gauge variance. When converting eq. 1-4 to the Hamiltonian formalism by using the quantum mechanical relationship $\mathbf{p} = -i\hbar\nabla$, one can recognize the classical equation of an electron with mass m in an electromagnetic field described by $A_\mu = (\phi, \mathbf{A})$. Therefore, one can regard ϕ as a field with charge e which couples to A_μ the same strength e. A mass term such as $mA_\mu A^\mu$ is not present in the Lagrangian and adding would break the gauge invariance. Therefore A_μ is interpreted as the massless photon propagator for the electromagnetic force.

In general, the field ϕ describes a fermion with charge e whereas ϕ^* describes its anti-fermion with charge $-e$. The intrinsic spin of the fermion is not taken into account with such complex scalar fields, but the Dirac spinor formalism takes the spin into account. In that formalism, the QED Lagrangian becomes:

$$\mathcal{L}_{\text{QED}} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \tfrac{1}{4} F^{\mu\nu} F_{\mu\nu}, \tag{1-5}$$

The Lagrangian 1-5 formulates the field theory of quantum electrodynamics. The internal symmetry leading to the gauge invariance is a rotation and the corresponding symmetry group is U(1). Therefore, QED is a U(1) field theory.

1.3.2 Strong Force

The phenomenology of the strong interaction was initially developed in electron-proton elastic and deep inelastic scattering experiments. The outcome of such experiments revealed the inner structure of the proton; it was found that the proton consists of point-like particles which are called partons. Parton models were built to describe the proton and neutron structure in terms of charge and spin. Those partons, later called quarks, are confined in the proton and carry each a fraction of the proton momentum, which is described by the parton distribution functions (PDF). In the sixties, a series of experiments at the Stanford linear collider observed hundreds of strong unstable states (hadrons). The mystery of this "particle zoo" was solved by theorists which found an underlying symmetry pattern which exhibit the same quantum numbers. This gave birth to the quantum chromodynamics (QCD), the quantum field theory of the strong interaction.

Two classes of hadrons exist:

- baryons: half-integer spin hadrons which are composed of three quarks $|q_1 q_2 q_3\rangle$ (e.g. proton, neutron). In reactions, the same number of baryons and/or anti-baryons is produced, therefore the baryon quantum number B is a conserved quantity ($B = 1(-1)$ for baryons(antibaryons), $B = 1/3(-1/3)$ for quarks(antiquarks), $B = 0$ for other particles).

- mesons: integer spin hadrons which are composed of quark-antiquark pair $|q_1 \bar{q_2}\rangle$ (e.g. pions). No meson number exist as mesons can decay e.g. to electrons and neutrinos through the weak force.

Crucial in the development of QCD was the Δ^{++} resonance, which is composed of three up quarks. Measurements of the spin ($J = 3/2$) and parity (positive) indicates the spin of the three up quarks must be parallel: $\Delta^{++} = |u^{\uparrow} u^{\uparrow} u^{\uparrow}\rangle$. This clearly violates the Pauli exclusion principle for fermions and a new property for quarks, the color, was invented to distinguish the three up quarks on quantum level. Each quark can have three colors: red, green or blue (and anticolor for the antiquarks). Color anti-symmetric wave functions for baryons and symmetric for mesons must be constructed for baryons and mesons:

$$B = \frac{1}{\sqrt{6}} \epsilon^{ijk} \left| q_i q_j q_k \right\rangle , \qquad M = \frac{1}{\sqrt{3}} \delta^{ij} \left| q_i \bar{q_j} \right\rangle . \tag{1-6}$$

The strong interaction is mediated by gluons, massless vector bosons with spin 1. In order to couple gluons to the colored quarks, they simultaneously carry color and anticolor. Because three colors exist, the color space is defined by the SU(3) group, which requires 8 generators for a fundamental representation of this group (i.e. with these octet states any

color in color space might be reconstructed). The remaining singlet state is colorless and therefore cannot participate in the interaction.

Similarly as for the QED case, the free Lagrangian is (locally) gauged. For each quark type, the quark spinors are described by a vector in color space: $q_f^T = (\phi_f^1\ \phi_f^2\ \phi_f^3)$, where $(1, 2, 3)$ denotes the color and f the quark flavor. One obtains, in spinor notation, the Lagrangian for quantum chromodynamics, invariant under *local* isospin rotations in SU(3):

$$\mathcal{L}_{\mathrm{QCD}} = \sum_f \bar{q}_f(i\gamma^\mu D_\mu - m_f)q_f - \tfrac{1}{4}G_a^{\mu\nu}G^a_{\mu\nu}. \tag{1-7}$$

The last term denotes the gluon kinetic term, which is a combination of the gluon fields. The covariant derivative D_μ contains now the eight gauge fields G_a^μ $(a = 1...8)$ which couples to the quark spinors with strong coupling constant g_s, and are identified with the gluon fields:

$$D^\mu q_f = \left(\partial^\mu - ig_s\frac{\lambda^a}{2}G_a^\mu\right) q_f, \tag{1-8}$$

The traceless matrices λ^a are the generators of the SU(3) group. No mass terms in form of $m_a\lambda^a G_a^\mu$ appear in the Lagrangian, which is confirmed by the massless gluons.

1.3.3 Weak Force and Electroweak Unification

The weak force was discovered in β decays of unstable nuclei. Experimentally, it was known the weak force had a very short range in the order of 10^{-18} m, very short compared to the infinite range of the electromagnetic force. In view of the uncertainty principle, the range is estimated to be $\hbar/(2mc)$, hence the force carrier mass must be in the order of 100 GeV/c^2. To preserve charge conservation in the observed interaction processes, at least two oppositely charged force carries must exist ($W^\pm$ bosons) to explain such "charged current" interactions. Later onward also "neutral currents" where observed which was identified by a neutral weak force carrier (Z boson).

In 1933, Fermi attempted to describe β decay by assuming a point-like interaction involving four particles (e.g. $\mu \to \nu_\mu\nu_e e$). The involving neutrinos were predicted two years earlier by Pauli to explain the continuous energy spectrum of the decay products. Nowadays, the Fermi theory is regarded in the limit of low momentum transfer $q \ll M_W$ at the interaction vertex with coupling constant g to the W$^\pm$ boson:

$$\frac{g^2}{M_W^2 - q^2} \approx \frac{g^2}{M_W^2} = 8\pi\alpha\frac{g^2}{e^2}\frac{1}{M_W^2} = 4\sqrt{2}G_F. \tag{1-9}$$

The coupling constant G_F, that appears in Fermi's theory, was experimentally measured to be several orders of magnitude lower than the electromagnetic one, hence the name weak

force. A full development of the electroweak theory was established many years later in 1973, by Weinberg, Salam and Glashow. It was based on several observations involving charged and neutral currents and extensively relies on the electroweak symmetry breaking mechanism.

Processes involving weak deays were studied in detail in the fifties and sixties. A crucial observation was that neutrinos, particles that only interact with the weak force and gravity, always found to be left-handed (i.e. the spin is always in the opposite direction than its momentum). They possess a helicity value of $h = -1$, where $h = \mathbf{s} \cdot \mathbf{p}/(|\mathbf{s}||\mathbf{p}|)$ is the helicity operator with $\mathbf{s}$ the spin and $\mathbf{p}$ the momentum of the particle. This is generalized such that charged currents only couple to left-handed fermions (and to right-handed anti-fermions). In addition, for weak decays involving quarks, experiments found that the coupling of weak quark decays was different among the quark families. The reason for this behavior is that for the weak interaction involving $\mathrm{W}^{\pm}$ boson exchange, the quark families mix and are therefore not eigenstates of the weak interaction. Such eigenstates can be found by rotating the quark-states, which is described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix:

$$\begin{pmatrix} |d'\rangle \\ |s'\rangle \\ |b'\rangle \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} |d\rangle \\ |s\rangle \\ |b\rangle \end{pmatrix}. \tag{1-10}$$

The meaning of this matrix is that the probability of transition from a quark q to q' is given by $|V_{qq'}|^2$.

These experimental observations are summarized in the electroweak theory by introducing two new quantum numbers: the weak isospin I and weak hypercharge Y_W. Left-handed quarks and leptons form doublets with a difference of one electrical charge unit, and can be transformed into each other by exchange of a $\mathrm{W}^{\pm}$ boson: they form an isospin doublet with $T = 1/2$ and $T_3 = 1/2$. Right-handed fermions are described as singlets ($T = T_3 = 0$). The weak hypercharge Y_W are singlets and are in relationship with the electric charge Q through

$$Q = T_3 + \tfrac{1}{2} Y_W. \tag{1-11}$$

A complete overview of the quantum numbers assigned to the SM fermions and gauge bosons is given in table **1-1**. Translating these ideas to field theory, the starting point is the free Dirac Lagrangian, but written in terms of the left-handed isospinor doublets $L^T = (\nu_e \;\; e_L)$ and right-handed isospinor singlets $R = e_R$:

$$\mathcal{L} = i\overline{R}\gamma \cdot \partial R + i\overline{L}\gamma \cdot \partial L. \tag{1-12}$$

One now gauges the Lagrangian to be invariant under SU(2) (weak isospin) and U(1)

				Q	T	T_3	Y_W
Leptons	$\begin{pmatrix}\nu_e\\ e\end{pmatrix}_L$	$\begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix}_L$	$\begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix}_L$	$\begin{matrix}0\\ -1\end{matrix}$	$1/2$	$\begin{matrix}+1/2\\ -1/2\end{matrix}$	-1
	e_R	μ_R	τ_R	-1	0	0	-2
Quarks	$\begin{pmatrix}u\\ d'\end{pmatrix}_L$	$\begin{pmatrix}c\\ s'\end{pmatrix}_L$	$\begin{pmatrix}t\\ b'\end{pmatrix}_L$	$\begin{matrix}+2/3\\ -1/3\end{matrix}$	$1/2$	$\begin{matrix}+1/2\\ -1/2\end{matrix}$	$\begin{matrix}+1/3\\ +1/3\end{matrix}$
	u_R	c_R	t_R	$+2/3$	0	0	$+4/3$
	d_R	s_R	b_R	$-1/3$	0	0	$-2/3$
Gauge bosons		$\mathrm{W}^\pm$		±1	1	±1	0
		Z		0	0	0	0
		γ		0	0	0	0
		h^0		0	$1/2$	$-1/2$	$+1$

Table 1-1: Overview of quantum numbers (electromagnetic charge Q, weak isospin I and weak hypercharge Y_W) for the electroweak theory.

(weak hypercharge) rotations. This requires to replace the derivatives by their covariant derivatives,

$$D_\mu L = \partial_\mu L - ig\widetilde{W}_\mu - iY_W g' B^0_\mu L, \tag{1-13}$$

$$D_\mu R = \partial_\mu R + iY_W g' B^0_\mu R, \tag{1-14}$$

with $2\widetilde{W}_\mu = \sigma_i W^i_\mu$ and σ^i the Pauli-matrices. New gauge fields and coupling constants for both symmetries are introduced: three gauge fields $(W^+_\mu, W^-_\mu, W^0_\mu)$ with coupling constant g for SU(2) and one B^0_μ gauge field with coupling constant g' for U(1). Appropriate gauge invariant kinetic terms should be added to the Lagrangian. A full derivation of the theory requires the introduction of electroweak symmetry breaking (see section 1.4), which takes into account the masses of the vector bosons. Crucial in the further development of the electroweak theory is that the photon and Z boson now are mixed states of the W^0 and B^0 states:

$$\begin{aligned} |\gamma\rangle &= \quad \cos\theta_W |B^0\rangle + \sin\theta_W |W^0\rangle \\ |Z^0\rangle &= -\sin\theta_W |B^0\rangle + \cos\theta_W |W^0\rangle . \end{aligned} \tag{1-15}$$

θ_W is the Weinberg angle and is defined by $\tan\theta_W = g'/g$ and related to the electromagnetic coupling constant by $e = g\sin\theta_W$. The value of θ_W can be obtained from various experiments and a combined value is measured to be $\sin^2\theta_W = 0.23122$ [6].

1.4 Higgs Mechanism and Particle Mass

1.4.1 Electroweak Symmetry Breaking

As discussed in section 1.3.1, adding a mass term to the Lagrangian breaks the (local) U(1) gauge invariance, therefore the photon field is massless. Similarly, for the electroweak Lagrangian, which possesses a U(1)×SU(2) symmetry, gauge invariance requires the gauge fields to be massless. This is in contradiction with experiment as the vector bosons are massive. An explanation why the gauge bosons acquire mass was investigated by Englert, Brout and Higgs in the early sixties [7, 8, 9]. To explain these ideas, consider the following Lagrangian which possesses a *global* U(1) symmetry:

$$\mathcal{L} = (\partial_\mu\phi)(\partial^\mu\phi^*) - m^2\phi^*\phi - \lambda(\phi^*\phi)^2 = (\partial_\mu\phi)(\partial^\mu\phi^*) - V(\phi), \tag{1-16}$$

with $V(\phi) = m^2\phi^*\phi + \lambda(\phi^*\phi)^2$ the potential. ϕ is a complex scalar field, λ a real self-interaction term and m a real parameter which can be negative. The ground state of this potential is calculated by taking its derivative w.r.t. the field. For $m^2 > 0$, the minimum occurs at $\phi = \phi^* = 0$ but for $m^2 < 0$ a local minimum is found at $|\phi|^2 = -m^2/(2\lambda) \equiv a^2$. This corresponds to a non-zero vacuum expectation value (vev) of ϕ:

$$|\langle 0|\phi|0\rangle|^2 = -\frac{m^2}{2\lambda} = a^2. \tag{1-17}$$

The physical fields are excitations above the vacuum state, therefore the fields must be expanded around this vacuum state. By writing the field as

$$\phi(x) = a + \frac{\phi_1 + i\phi_2}{\sqrt{2}}, \tag{1-18}$$

where ϕ_1 and ϕ_2 are real scalar fields with zero vacuum expectation value, the Lagrangian

in terms of the real fields ϕ_1 and ϕ_2 becomes:

$$\mathcal{L} = \tfrac{1}{2}(\partial_\mu \phi_1)^2 + \tfrac{1}{2}(\partial_\mu \phi_2)^2 - 2\lambda a^2 \phi_1^2 - \sqrt{2}\lambda \phi_1(\phi_1^2 + \phi_2^2) - \frac{\lambda}{4}(\phi_1^2 + \phi_2^2)^2. \tag{1-19}$$

The field ϕ_1 became massive with mass proportional to $\sqrt{\lambda}a$ whereas ϕ_2 has no mass term and is massless. The latter is referred as a Goldstone boson. As a consequence of the spontaneous U(1) symmetry breaking, two massive scalar fields (ϕ_1, ϕ_2) are transformed to a massive particle (ϕ_1) and massless Goldstone boson (ϕ_2).

When translating these ideas to the breaking of a *local* U(2) gauge symmetry as defined in (see Eq. 1-4), where initially two massive fields (ϕ_1, ϕ_2) and a massless field (A_μ) are present, after spontaneous symmetry breaking one remains a massive scalar field (ϕ_1) and a massive photon. The field ϕ_2 corresponding to the massless Goldstone boson is non-physical and can be eliminated in an appropriate gauge.

1.4.2 Higgs-Kibble Mechanism

The idea of spontaneous symmetry breaking is applied to the theory of electroweak unification [10, 11]. The starting point is the free Dirac Lagrangian (see eq. 1-12) which possesses two symmetries: SU(2) weak isospin and U(1) weak hypercharge. By gauging them, covariant derivatives are introduced with their gauge fields $\mathbf{W}_\mu$ and B^0_μ which are invariant under the respective transformations and are accompanied by two gauge coupling constants g and g' respectively (see section 1.3.3). The resulting Lagrangian describes the interaction of charged and neutral currents to their respective components (left vs right handed), but no mass terms for the fermions nor for the gauge fields are present. If such terms are added, gauge invariance is broken.

To give mass to the gauge bosons, a Higgs doublet is introduced to the theory:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \tag{1-20}$$

with ϕ^+ and ϕ^0 two complex scalar fields. The quantum numbers are chosen to be $I_W = 1/2$ and $Y_W = 1$, such that ϕ^0 has no charge and and ϕ^+ has unit charge. The Lagrangian in terms of this field takes the following form:

$$\mathcal{L} = (D_\mu \phi)^\dagger (D^\mu \phi) - \mu^2 \phi^\dagger \phi - \lambda(\phi^\dagger \phi)^2. \tag{1-21}$$

The minimum of the potential occurs at a vacuum expectation value of $\sqrt{\frac{-\mu^2}{2\lambda}} \equiv v/\sqrt{2}$.

Under the (local) electroweak SU(2)×U(1) symmetry, the covariant derivative is written as:

$$D_\mu \phi = \left(\partial_\mu - ig\widetilde{W}_\mu - \frac{i}{2} g' B^0_\mu \right) \phi, \tag{1-22}$$

By virtue of the fact that the Lagrangian is local gauge invariant, one can perform independent gauge transformations at each point in space-time. Hence, the Higgs doublet can be parameterized as follows (unitary gauge):

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \tag{1-23}$$

where $H(x)$ is called the Higgs field (neutral field). In this unitary gauge, the Lagrangian (eq. 1-21) can be written in as follows:

$$\mathcal{L} = \frac{1}{4} h v^4 + \mathcal{L}_H + \mathcal{L}_{HG^2} \tag{1-24}$$

where, with the use of eqns. 1-15,

$$\mathcal{L}_H = \frac{1}{2} \partial_\mu H \partial^\mu H - h v^2 H^2 - h v H^3 - \frac{1}{4} h H^4, \tag{1-25}$$

$$\mathcal{L}_{HG^2} = (v + H)^2 \left(\frac{g^2}{4} W^\dagger_\mu W^\mu + \frac{g^2}{8 \cos^2 \theta_W} Z_\mu Z^\mu \right). \tag{1-26}$$

From these equations, the masses for the gauge bosons are identified from the quadratic terms in the respective fields:

$$M_W = M_Z \cos \theta_W = \frac{1}{2} g v, \qquad m_H = \sqrt{2h} v. \tag{1-27}$$

The mass of the $W^\pm$ and Z bosons, and the vacuum expectation value v can be estimated from the Fermi theory (see eq. 1-9):

$$M_W^2 = \frac{4\pi\alpha}{8 \sin^2 \theta_W} \frac{\sqrt{2}}{G_F}, \tag{1-28}$$

leading to a vev of 246 GeV. One free parameter in the theory, the Higgs self coupling h, can be determined from the Higgs boson mass. Experimental verifications of the electroweak and Higgs gauge bosons are briefly discussed in section 1.5.

1.4.3 Fermion Mass

The Higgs field couples to the fermions with the following Yukawa interaction term (again, in the unitary gauge):

$$\mathcal{L} = -\frac{1}{\sqrt{2}}(v + H)(\lambda_d \bar{d}d + \lambda_u \bar{u}u + \lambda_e \bar{e}e), \tag{1-29}$$

with u, d, e the spinor fields for the up-quark, down-quark and electron field and the parameters λ_a ($a = (u, d, e)$) represents the coupling between the Higgs field H and the respective fermion a. The fermion mass follows directly from the coupling terms and can be written as:

$$m_u = v\lambda_u, \quad m_d = v\lambda_d, \quad m_e = v\lambda_e. \tag{1-30}$$

Similar terms can be added for the other lepton/quark families.

1.5 Experimental Tests of the Standard Model

During the early developments of the SM, theorists extensively relied on experimental observations for the understanding of the basic equations of the SM which ultimately lead to the Lagrangian description using a gauge theory. The electron, muon and photon were already known particles to describe QED, and their knowledge was extended and applied to the other forces and the prediction of other particles. Though the theoretical foundations of the SM were complete in the seventies, several particles were discovered later as they required the proper technological tools (e.g. τ-lepton, the neutrinos, top quark). At the end, to date, all those particles (and anti-particles) which enter the SM are discovered and their properties are measured with increasing accuracy.

In point of view of this thesis work, two interesting experimental discoveries are briefly discussed here: the discovery of the electroweak gauge bosons and the most recent Higgs boson discovery.

1.5.1 Discovery of the Electroweak Gauge Bosons

The massive electroweak gauge bosons were predicted in 1973 as unification of the electromagnetic and weak forces. Immediately after, in 1973, a neutrino experiment at CERN with the Gargamelle bubble chamber observed the first neutral current interactions. Though this observation was not a direct proof of the theory, it encouraged the experimental physicists at CERN to further focus on the $W^{\pm}$ and Z boson experimental verification. In 1976, several scientists at CERN proposed to transform the current Super Proton Synchrotron (SPS), which at that time was used for fixed target experiments, to a proton-antiproton collider, the $Sp\bar{p}S$. It involved technological challenges to produce and store anti-protons which ultimately lead to the development of stochastic cooling by S. van der Meer. Two major collaborations were formed in 1978 for the design, construction and operation of the two large Underground Experiments (UA1 and UA2), two large detectors which were installed along the accelerator interaction points. The accelerator project, lead by C. Rubbia, proceeded in parallel and first collisions at a center-of-mass energy of 540 GeV were delivered to the experiments. This energy was sufficient to produce the $W^{\pm}$ and Z bosons, whose mass has been accurately calculated using eq. 1-28.

In November 1982, the first valid W candidates were seen in the detector having a 40 GeV electron and missing transverse energy of the same order. More events of such types were collected (as well as in the muon channel), and tight selections were applied to the events. The transverse mass was reconstructed and was consistent with a boson of mass of 80 GeV. A paper was published soon after near the end of January 1983 stating the observation of the $W^{\pm}$ gauge bosons. Similarly, the first Z boson candidates were observed in the next run in 1983 where a back-to-back electron-positron pair were observed. On June 1983, a CERN press release announced the discovery of the Z boson.

1.5.2 Discovery of the Higgs Boson and Couplings

Opposite to the $W^{\pm}$ and Z bosons, the mass of the Higgs boson was not known and could be any value, but it should be around the same order than the gauge bosons. Furthermore, the production cross section was calculated to be very low. Searches in the SppS, the large Electron Positron collider (LEP) and the Tevatron quoted upper limits of the Higgs boson or excluded several mass ranges. It was therefore the Large Hadron Collider was designed and built (see section 3.1), which lead to an early discovery of the Higgs boson in the ATLAS and CMS experiments [12, 13, 14] (see section 3.2 for a description of the CMS detector). The discovery has been made in the "golden channels", $h \to ZZ \to 4l$ and $h \to \gamma\gamma$, with a mass of 125 GeV.

In the ongoing analyses of the LHC data, the Higgs mass is accurately measured and to

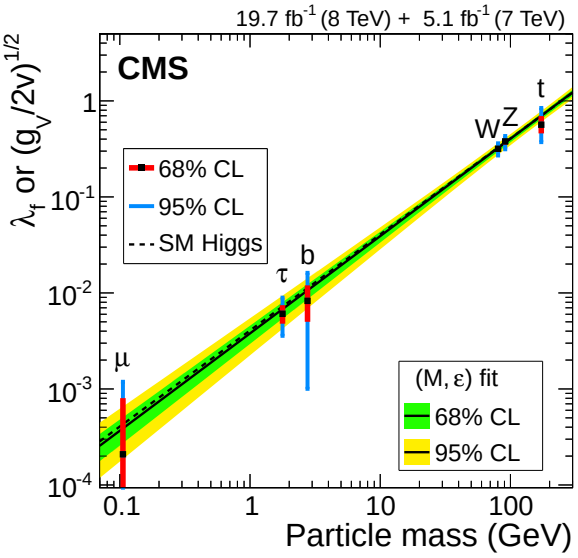


Figure **1-2**: Measured Higgs coupling constants in the appropriate channels versus the particle mass for the 7+8 TeV dataset [1].

date, a combined result of the ATLAS and CMS results quote a number of 125.09 ±0.21(stat.) ±0.11(syst.) GeV[15]. This proved the existence of the Higgs field which couples gives mass not only to the massive gauge bosons, but also to the fermions (except neutrinos). Hence the LHC data is extensively used to measure the couplings to the Higgs field, by measuring the Higgs decay to those particles: $H^{\pm} \to ZZ$, $H^{\pm} \to WW$, $H^{\pm} \to b\bar{b}$, $H^{\pm} \to \tau\tau$, $H^{\pm} \to \mu\mu$. Higgs couplings to the top quark are measured at production level, however recent measurements with the CMS experiment quote a direct coupling to the top quark in the $t\bar{t}H$ associated production [16]. From eqns. 1-30, the fermions and gauge boson couple directly to the Higgs field with a factor proportional to the particle mass. Those couplings λ_f are directly measured for each channel and are plotted in figure **1-2** versus the particle mass. The dashed line represents the SM calculations with a Higgs mass of 125 GeV and $v = 246$ GeV, and the solid line is a fit to the data points, which agrees perfectly with the theory. The couplings to other particles (light quarks and electron) is very difficult as their mass, and coupling, are very low.

An interesting coupling and, and in effect the final proof of the Higgs particle, is the Higgs self coupling. From the EWK Lagrangian, self-interaction terms appear which will eventually generate the Higgs mass (see eq. 1-25). To directly probe such couplings, di-Higgs production is necessary. Despite the very low cross section, this will be achievable in the High-Luminosity phase of the LHC (HL-LHC), which will produce roughly 10 times more data and capable to observe di-Higgs production.

1.6 Missing Pieces and Beyond the Standard Model

Besides the two crucial discoveries briefly mentioned in the previous section, the SM has been tested up to TeV scale with astonishing precision. During the LHC Run II, with enhanced collision energy to 13 TeV, the observations were in good agreement with the SM predictions. However, still several fundamental questions are not yet solved or the SM cannot give an appropriate answer. Some of them are listed below:

- Free parameters: the SM is a model describing the forces between elementary particles, but it relies on a set of axioms (e.g. only specific terms in the Lagrangian) and constants (couplings, masses, etc.). In total, 19 free parameters enter the Lagrangian of the SM whose values had to be obtained from experiment. Those constants are rather arbitrary and uncorrelated.

- Neutrino mass: the Higgs mechanism does not allow a mass term for neutrinos, despite experimental observations that neutrinos are not found to be massless. Upper mass limits have been measured for the three neutrino flavors, with the heaviest ν_τ lower than 18 MeV [6]. Ad-hoc mechanisms for neutrino masses within the SM will attribute to an additional set of free parameters, needed to be determined from experiment.

- Gravitation: the SM does not include the gravitational force. Despite unification attempts, general relativity is difficult to treat as a quantum theory and quantum gravity to first extent is not renormalizable. Moreover, the associated graviton as gravity force carrier is not observed.

- Hierarchy problem: related to gravitation is the hierarchy problem, which questions the fact why the weak force is 10^{24} times stronger than the gravitational force. This huge discrepancy of coupling constants is not observed in the SM.

- Baryon-asymmetry: despite the fact that the SM only allows a small deviation of matter-antimatter ($\mathcal{CP}$-violation in kaon decays), is not sufficient to explain wht the visible universe is made out of matter only. Higher degrees of $\mathcal{CP}$ violation or baryon-number violation are needed, but not observed, to explain the baryon-asymmetry.

- Dark matter and dark energy: from cosmological experiments only 5% of the total energy in the universe is accounted for the visible mass, as described by particles and interactions in the SM. The remaining mass and energy is accounted for dark energy (69%) and dark matter (26%).

- Anomalous magnetic dipole moment: the anomalous magnetic dipole moment describes quantum loop effects to the (classical) magnetic dipole moment. For the electron, precise calculations are in agreement with observation up to 10 significant digits, which serves as a successful precision test of QED. For the muon, however, the calculated value disagrees with 3.5 standard deviations from recent observations [17].

It is clear the SM is not able to describe (all) the physics phenomena on a fundamental level. New theories, models and ideas are needed, whether or not based on the SM formalism. Such theories were proposed during the development of quantum mechanics and the standard model. Nowadays, two theories are under wide attention: string theory and supersymmetry. Both were developed in the sixties and seventies, aiming to incorporate gravity, explaining dark matter and several other shortcomings of the SM.

In particular for supersymmetry, which has direct implications on the existence of additional Higgs bosons, it provided a direct way for observation. Indeed, it introduces an additional symmetry, on top of the SM symmetries, between fermions and bosons. For each fermion an associated boson is predicted and vice versa. Such particles can be candidates for dark matter and could be directly produced in collider experiments, if kinematically allowed. Searches are ongoing at LHC for several SM "superpartners", however to date only experimental constrains the theory are set.

2 Phenomenology of the Charged Higgs Sector

Charged Higgs bosons appear naturally in extensions in the Higgs sector of the Standard Model. Such extensions exhibit a more complex Higgs sector yielding additonal Higgs bosons with different masses, charges, and other properties. Whereas the 125 GeV Higgs boson is compatible with both the SM and extensions, the detection of a charged Higgs boson would unequivocally point to new physics beyond the SM. The properties of the observed 125 GeV Higgs boson constrain, but not exclude, several parameters in the extended models.

2.1 Two Higgs Doublet Model

The most natural extension of the SM Higgs sector is adding a second doublet to the electroweak symmetry breaking theory. As described in section 1.4.2, the SM Higgs theory relies on a Higgs doublet that breaks the Lagrangian symmetry. In the Two Higgs Doublet Model (2HDM) [18], two identical Higgs doublets $\phi_{1,2}$ with weak isospon $I_W = 1/2$ weak hypercharge $Y_W = \pm 1$ are introduced. Accordingly, the Higgs potential is modified to allow all possible quartic interactions:

$$\begin{aligned} V(\Phi_1,\Phi_2) = {} & m_{11}^2\Phi_1^\dagger\Phi_1 + m_{22}^2\Phi_2^\dagger\Phi_2 - \left[m_{12}^2\Phi_1^\dagger\Phi_2 + \text{h.c.}\right] + \tfrac{1}{2}\lambda_1(\Phi_1^\dagger\Phi_1)^2 + \tfrac{1}{2}\lambda_2(\Phi_2^\dagger\Phi_2)^2 \\ & + \lambda_3(\Phi_1^\dagger\Phi_1)(\Phi_2^\dagger\Phi_2) + \lambda_4(\Phi_1^\dagger\Phi_2)(\Phi_2^\dagger\Phi_1) \\ & + \left[\tfrac{1}{2}\lambda_5(\Phi_1^\dagger\Phi_2)^2 + [\lambda_6(\Phi_1^\dagger\Phi_1) + \lambda_7(\Phi_2^\dagger\Phi_2)](\Phi_1^\dagger\Phi_2) + \text{h.c.}\right]. \end{aligned} \tag{2-1}$$

In this equation, the parameters m_{11}, m_{22}, m_{12} and $\lambda_{1,2,3,4}$ are real, whereas the remaining parameters $\lambda_{5,6,7}$ could be complex. Both doublets couple to quarks and leptons through Yukawa couplings. When expanding the full Yukawa Lagrangian, interaction terms describing $\mathcal{CP}$-violation and tree-level flavor changing neutral currents (FCNCs) are present. Such terms can be removed when a Z_2 symmetry is imposed on the Lagrangian: $\Phi_1 \to \Phi_1$ and $\Phi_2 \to -\Phi_2$. By requiring this symmetry on 2-1, the constants $\lambda_{5,6,7}$ must be zero. The

resulting Lagrangian contains 8 degrees of freedom for the coupling terms.

The doublet fields are expanded around their vacuum expectation values (where, as before by virtue of the local gauge principle, the direction of the vev is assumed to be along the second component):

$$\langle \Phi_1 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_1 \end{pmatrix}, \qquad \langle \Phi_2 \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_2 \end{pmatrix}, \tag{2-2}$$

and the global vacuum expectation value should be equal to the one found in the SM: $v = \sqrt{v_1^2 + v_2^2} \approx 246$ GeV. Together with the definition $\tan\beta = v_2/v_1$, the doublets can be written with a proper translation of the vevs:

$$\Phi_1 = \begin{pmatrix} \phi_1^+ \\ \frac{1}{\sqrt{2}}(v\cos\beta + \phi_1^0) \end{pmatrix}, \qquad \Phi_2 = \begin{pmatrix} \phi_2^+ \\ \frac{1}{\sqrt{2}}(v\sin\beta + \phi_2^0) \end{pmatrix}, \tag{2-3}$$

with $\phi_1^+, \phi_1^0, \phi_2^+$ and ϕ_2^0 complex scalar fields. From the initial 8 degrees of freedom, three are absorbed to give mass to the $W^\pm$ and Z bosons, leaving 5 degrees of freedom which are identified with 5 massive physical Higgs states:

- h: $\mathcal{CP}$-even neutral scalar,
- H: $\mathcal{CP}$-even neutral, scalar, with, per definition, $m_H > m_h$,
- A: $\mathcal{CP}$-odd scalar,
- $H^\pm$: pair of charged Higgs bosons.

The five physical Higgs states can be written in terms of the complex scalar fields [19]. In particular for the charged Higgs boson, the physical state is obtained from the charged field component orthogonal to the vev:

$$H^+ = -\phi_1^+ \sin\beta + \phi_2^+ \cos\beta. \tag{2-4}$$

Similarly, the pseudoscalar Higgs boson can be written as:

$$A = -\mathrm{Im}(\phi_1^0)\sin\beta + \mathrm{Im}(\phi_2^0)\cos\beta. \tag{2-5}$$

The masses of both states are in relationship w.r.t. each other and are as follows:

$$M_A^2 = \frac{2m_{12}^2}{\sin 2\beta} - \lambda_5 v^2, \qquad M_{H^\pm} = M_A^2 + \frac{1}{2}v^2(\lambda_5 - \lambda_4). \tag{2-6}$$

The remaining two $\mathcal{CP}$-even neutral scalars (h, H) mix with an angle α according to:

$$\begin{pmatrix} h \\ H \end{pmatrix} = \begin{pmatrix} -\sin\alpha & \cos\alpha \\ \cos\alpha & \sin\alpha \end{pmatrix} \begin{pmatrix} \mathrm{Re}(\phi_1^0) \\ \mathrm{Re}(\phi_2^0) \end{pmatrix}, \tag{2-7}$$

and their masses are obtained by diagonalization of the mass-matrix. To relate the neutral scalar states with the observed 125 GeV Higgs boson, it is convenient to work in the Higgs basis, defined as a rotation over β:

$$H_1 = \begin{pmatrix} H_1^+ \\ H_1^0 \end{pmatrix} \equiv \Phi_1 \cos\beta + \Phi_2 \sin\beta, \qquad H_2 = \begin{pmatrix} H_2^+ \\ H_2^0 \end{pmatrix} \equiv -\Phi_1 \sin\beta + \Phi_2 \cos\beta. \tag{2-8}$$

The vacuum expectation values of these states are $\langle H_1^0 \rangle = v/\sqrt{2}$ and $\langle H_2^0 \rangle = 0$, such that the physical state H_1^0 is identified by the SM Higgs boson (see section 1.4.2). An explicit representation of the neutral scalar states in the Higgs basis is written as:

$$\begin{pmatrix} h \\ H \end{pmatrix} = \begin{pmatrix} -\sin(\beta-\alpha) & \cos(\beta-\alpha) \\ \cos(\beta-\alpha) & \sin(\beta-\alpha) \end{pmatrix} \begin{pmatrix} \mathrm{Re}H_1^0 - v \\ \mathrm{Re}H_2^0 \end{pmatrix}, \tag{2-9}$$

from which it become clear that the SM Higgs boson coincides with h if $\cos(\beta-\alpha) \to 0$ or with H if $\sin(\beta-\alpha) \to 0$.

Type	U_R	U_R	U_R	λ_{UU}	λ_{DD}	λ_{LL}
I	Φ_2	Φ_2	Φ_2	$\cot\beta$	$\cot\beta$	$\cot\beta$
II	Φ_2	Φ_1	Φ_1	$\cot\beta$	$-\tan\beta$	$-\tan\beta$
Lepton-specific	Φ_2	Φ_2	Φ_1	$\cot\beta$	$\cot\beta$	$-\tan\beta$
Flipped	Φ_2	Φ_1	Φ_2	$\cot\beta$	$-\tan\beta$	$\cot\beta$

Table 2-1: Four different 2HDM types and their couplings to the doublets. Per definition the up-quark types couples to the second doublet ϕ_2.

2.1.1 Fermion Couplings

The physical Higgs boson state couple to fermions described by Yukawa interactions. A wide variety of coupling schemes possible, resulting in several types that are summarized in table **2-1**. Per definition, the up-quark type couples to the second doublet ϕ_2, leaving 4 possibilities:

- Type I: the down-type quark and leptons couple to the same doublet ϕ_2;
- Type II: the down-type quark and leptons couple to the other doublet ϕ_1;
- Lepton-specific: the down-type quark couples to ϕ_2 whereas the leptons couple to ϕ_1;
- Flipped: the down-type quark couples to ϕ_1 whereas the leptons couple to ϕ_2.

When writing the full Yukawa interactions in terms of the charged Higgs fields (equation 2-4), one obtains a set of 9 free parameters describing the couplings to the fermions: $\lambda_{uu}, \lambda_{dd}, \lambda_{ss}, \lambda_{cc}, \lambda_{bb}, \lambda_{tt}, \lambda_{ee}, \lambda_{\mu\mu}$ and $\lambda_{\tau\tau}$. In general, their hierarchy is constrained by the mass of fermions, e.g. $m_{\mu\mu}|\lambda_{\mu\mu}| \ll m_{\tau\tau}|\lambda_{\tau\tau}|$.

In the assumption that one doublet couples to each class of fermions (i.e. prevent flavor changing neutral currents), the charged Higgs states couples to fermions $f\bar{f}$ with the following coupling constants [19]:

$$g_{\mathrm{H}^\pm \bar{u}_i d_j} = \frac{iV_{ij}^{\mathrm{CKM}}}{\sqrt{2}M_W}\left(\lambda_{UU}\frac{1-\gamma_5}{2} + \lambda_{DD}\frac{1+\gamma_5}{2}\right) \tag{2-10}$$

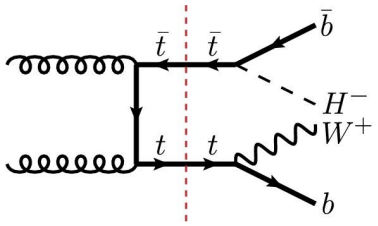
$$g_{\mathrm{H}^\pm \bar{\nu}_\ell \ell} = \frac{i}{\sqrt{2}M_W}\lambda_{LL}\frac{1+\gamma_5}{2}, \tag{2-11}$$

with V_{ij}^{CKM} the CKM matrix and λ_{FF} the fermion-specific coupling depending on $\tan\beta$ (see table **2-1**).

2.2 Charged Higgs Production

In general, the production of charged Higgs bosons depend on the mass and parameters of the model (2HDM and MSSM). In the latter, production cross sections can differ whether superymmetric particles are involved or not. Depending on the mass, three regimes are distinguished: light, intermediate and heavy mass regime.

2.2.1 Light Mass Regime



The light mass regime is defined as charged Higgs production below the top quark mass ($m_{\mathrm{t}} \approx 172.5$ GeV). Typical experimental searches range from 80 to 160 GeV, as below 80 GeV the charged Higgs is mainly excluded by LEP measurements (see section 2.5), and above 160 GeV top-interference effects start to become important (see section 2.2.2). In this regime, charged Higgs bosons are dominantly produced from top quark decays. Hence, searches for light charged Higgs bosons are focused on the measurement of the branching ratio $\mu = \mathcal{B}(t \to \mathrm{H}^{\pm} b)$, or any deviation of the $t\bar{t}$ cross section. From the known $t\bar{t}$ cross section $\sigma_{t\bar{t}}$ (831.76 pb at 13 TeV), the charged Higgs cross section can be calculated as:

$$\sigma_{\mathrm{H}^+} = \mu^2 \sigma_{t\bar{t}} + 2\mu(1-\mu)\sigma_{t\bar{t}} \approx 2\mu(1-\mu)\sigma_{t\bar{t}} \tag{2-12}$$

The approximation is valid because the branching ratio is typically to be considered below percent level, as no larger deviations in the $t\bar{t}$ cross section have been observed. (within the uncertainties).

2.2.2 Intermediate Mass Regime

For charged Higgs boson masses around the top quark mass, interference effects due to the top-width become important as well as resonant and non-resonant diagrams must be taken into account. Hence, the full process $pp \to \mathrm{H}^{\pm}\mathrm{W}^{\pm} b\bar{b}$ must be taken into account in the calculations with massive bottom quarks. Due to this additional diagrams the computation becomes rather complex, especially at NLO accuracy, yet recently the results were calculated and summarized in [2].

The contributing LO diagrams are shown in Fig. **2-1**: non-resonant top-quark contribution (a), single resonant top quark contribution (b), double-resonant top quark contribution (c) and contributions with neutral scalars (d). In the latter case, many additional diagrams

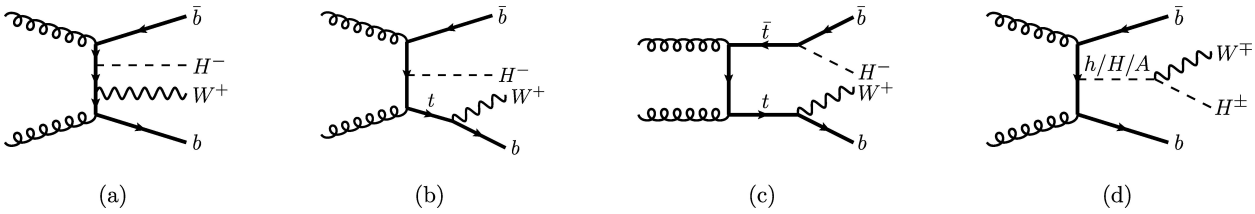


Figure **2-1**: Leading order diagrams for the full process $pp \to \mathrm{H}^{\pm}\mathrm{W}^{\pm}b\bar{b}$ [2]: (a) non-resonant top quark contribution, (b) single resonant top-quark contribution, (c) double-resonant top-quark contribution and (d) contributions with neutral scalars.

could be present depending on the neutrals involved. Restricting to the SM, only the 125 GeV Higgs neutral Higgs boson can contribute but in the general 2HDM also the heavy and pseudoscalar Higgs bosons (H, A) can contribute. As a result the production cross section becomes model dependent as it relies on the assumptions on the couplings and masses of these neutral scalars. With appropriate mass, the A/H contributions can become resonant, which is better described with $pp \to A/Hb\bar{b}$, where A/H decays to charged Higgs bosons. In general, is was shown that the contributions of the neutral scalars is almost negligible [2].

2.2.3 Heavy Mass Regime

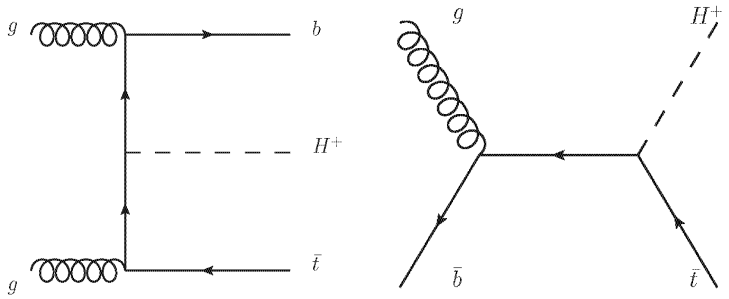


Figure **2-2**: Production of a heavy charged Higgs boson according to 4FS (left) and 5FS (right). Both contributions are added using the Santander matching scheme.

The charged Higgs boson with a mass larger than the top quark mass is produced in association with a top quark through the process $gg \to \mathrm{H}^{\pm}\mathrm{t(b)}$ as illustrated in Figure **2-2**. The associated b-quark in the production is typical for the so-called *four-flavor scheme* (4FS, left figure) but is not present in the *five-flavor scheme* (5FS, right figure). The difference between both at QCD tree level is that in the 4FS the bottom quark takes part of the hard scale of the process whereas in the 5FS the bottom quark is part of the proton wave function.

Total inclusive cross sections of both descriptions (σ^{4FS} and σ^{5FS}) are calculated up to NLO accuracy, and they are combined using the Santander Matching weighted average [20]:

$$\sigma^{\text{matched}} = \frac{\sigma^{\text{4FS}} + w\sigma^{\text{5FS}}}{1 - w}, \tag{2-13}$$

where the weight w is defined as:

$$w = \ln\left(\frac{\text{m}_{\text{H}^\pm}}{m_b}\right) - 2, \tag{2-14}$$

with m_b the mass of the bottom quark (≈ 4.75 GeV). Considering that $\text{m}_{\text{H}^\pm} \gg m_b$, the 5FS cross section is dominant for the entire heavy mass range. This is illustrated in Fig. **2-3** (left), where the solid lines represents the full inclusive cross section for the 4FS (red) and 5FS (blue) and Santander-matched (black). It must be noted that no SUSY QCD correctiond are applied (see section 2.3).

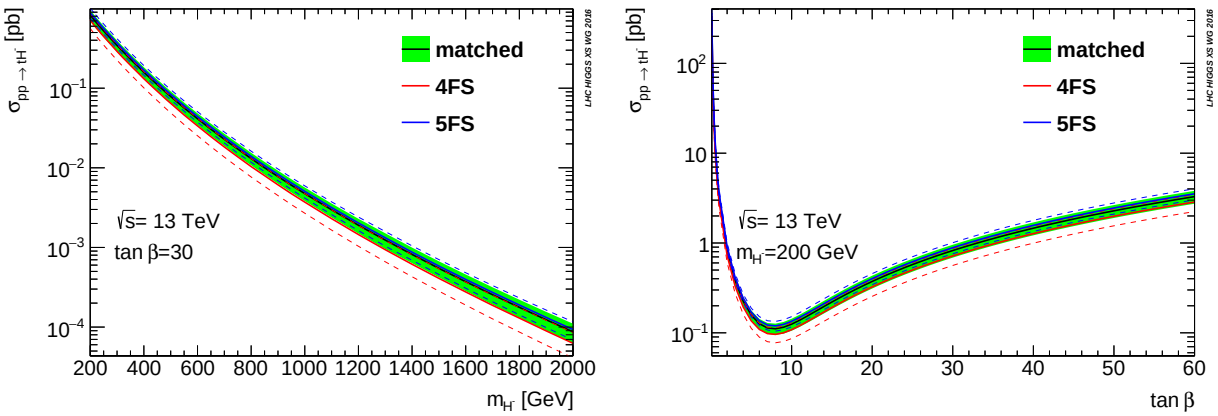


Figure **2-3**: Production cross sections in the 4FS and 5FS with the Santander-matched results as function of $\text{m}_{\text{H}^\pm}$ (left) or $\tan\beta$ (right) [3], for Type-II MSSM like. No SUSY QCD corrections are incorporated in these results.

2.3 2HDM Scenarios and the Minimal Supersymmetric Model

In the 2HDM described above, in total there are seven free parameters:

- masses of the Higgs states: $\mathrm{m_{H^\pm}}$, m_h, m_H, m_A;
- $\tan\beta$, the ratio of the vacuum expectation values;
- $\sin(\beta-\alpha)$, with α the mixing angle between the $\mathcal{CP}$-even Higgs states;
- m_{12}, the diagonal term of the mass matrix of the Higgs doublets.

Scenarios are built which impose specific relations among the free parameters in order to reduce them such that experimental constraints become possible. Among the Type-II models, one of the most popular models is the minimal supersymmetric extension to the standard model (MSSM) [21]. This model predicts the minimum amount of super symmetric particles necessary in SUSY theories, in particular the Higgs sector is similar to the one described in the 2HDM. At tree level, the MSSM-like model can be described with two free parameters (see e.g. [19] for a description): $\tan\beta$ and $\mathrm{m_{H^\pm}}$. The other parameters are constrained by: $\mathrm{m^2_{H^\pm}} = m_A^2 + m_W^2 \approx M_A^2$ in case of heavy SUSY particles, $\sin(\beta-\alpha)=1$ which means the Higgs state h is identified with the observed 125 GeV Higgs boson, hence $m_h = 125$ GeV, $m_H = m_A$ and $m_{12}^2 = m_A^2 \tan\beta/(1+\tan\beta^2)$ (which implies $\lambda_5 = 0$).

The phenomenology of the 2HDM and charged Higgs in particular strongly depends on the Yukawa sector. For the Type-II models, the fermion couplings introduce 9 additional free parameters which can be further relaxed under certain assumptions. For the MSSM, adding supersymmetric particles to the model requires radiative corrections which are referred as SUSY QCD corrections. The free choice of these parameters give rise to several scenarios within the MSSM [22, 23]. For a given set of input parameters, the charged Higgs production cross section is calculated together with the branching rations of possible decays (see next section). Often such scenarios are used as benchmark to interpret several searches within the same model (e.g. constraints from $H \to \tau\tau$ searches can be compared to constraints obtained from $\mathrm{m_{H^\pm}} \to \tau\nu_\tau$). Under the assumption of no SUSY particles (i.e. absence of SUSY QCD corrections), example charhed Higgs production cross section as function of $\tan\beta$ or $\mathrm{m_{H^\pm}}$ is given in figure **2-3**.

2.4 Charged Higgs Decay

Decays of charged Higgs to quarks (u, d), fermions (ν_ℓ, ℓ) or supersymmetric particles if incorporated in the model. In the Type-II 2HDM, the light charged Higgs dominantly decays to $\tau\nu_\tau$, whereas for the heavy charged Higgs the top-bottom final state is dominant. This can be understood due to enhanced couplings to the third generation SM particles (t, τ) as a result of their larger masses. However, as the leptons couple directly to $\tan\beta$ in Type-II (see equation 2-11), the $\tau\nu_\tau$ final state also has a sizable branching ratio for high charged Higgs masses. As a result, the tb and $\tau\nu$ final states are interesting for direct charged Higgs searches and put constraints on the 2HDM parameter space in Type-II.

In figure **2-4**, the branching ratios of the charged Higgs decay to tb, cs, $\tau\nu_\tau$ and $\mu\nu_\mu$ is plotted as function of $m_{H^\pm}$ for $\tan\beta = 10$ (left) and $\tan\beta = 50$ (right). The values have been calculated for the $m_h^{\text{mod}-}$ MSSM scenario, which assumes SUSY particles at the TeV scale and is compatible with the observed 125 GeV Higgs boson [22]. From the plot one can see that indeed the $\tau\nu_\tau$ is dominant in the light mass regime, and the tb takes over at high mass, with a non-negligible contribution of the $\tau\nu_\tau$ final state. For higher charged Higgs masses, SUSY decays become more dominant, especially at low $\tan\beta$ (not shown on the plot).

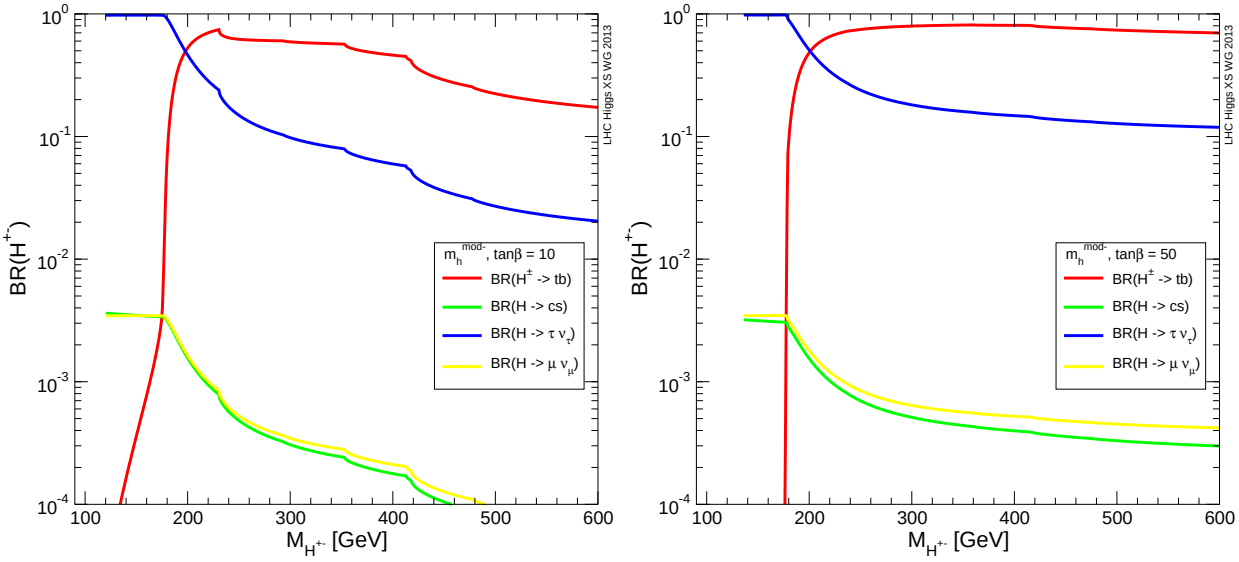


Figure **2-4**: Charged Higgs branching ratio as function of $m_{H^\pm}$ for $\tan\beta = 10$ (left) and $\tan\beta = 50$ (right) [3]. The $m_h^{\text{mod}-}$ MSSM scenario is used to obtain the branching ratios for the several decay channels. Invisible SUSY decay mode(s) are not shown on the plot.

2.5 Overview of Past Searches

The search for charged Higgs particles was performed in the CERN LEP, Fermilab Tevatron and CERN LHC Run I/II experiments. Exclusion limits were set on the production cross section as function of the charged Higgs mass, within the kinematic reach of each experiment.

At LEP, charged Higgs bosons could be produced pairwise in Drell-Yan events ($e^+e^- \to \gamma/Z \to H^+H^-$). The dominant decay modes are $\mathrm{H}^\pm \to \tau\nu$ and $\mathrm{H}^\pm \to cs$. 95% CL limits on the charged Higgs mass were set, excluding a charged Higgs mass below 80 GeV [24].

For the Fermilab Tevatron [25, 26] and CERN LHC hadron experiments, the searches focused on charged Higgs masses above 80 GeV. Mainly the LHC Run I results are dominant due to the large luminosity and higher energy than the Tevatron. Several searches were covered such as $\tau\nu_\tau$ [27, 28, 29, 30, 31, 32, 33], tb [31, 34], $c\bar{s}$ [35, 36] and $c\bar{b}$ [37]. Additionally, indirect searches by the ATLAS and CMS collaborations put constraints on the 2HDM (MSSM) parameter space [38, 39, 40, 19]. In such searches, charged Higgs bosons have been searched from 80 GeV to the TeV scale, putting upper limits on the production times branching fraction.

3 The Tools for the Search

To probe the laws of Nature on quantum level, typical high energy experiments are conducted. Since the initial developments of quantum mechanics and related field theories, experiments in high energy physics used cosmic rays and accelerated particles bombarding fixed-targets to study the interactions between elementary particles. In particular the accelerator-driven experiments were very useful because they provided a controlled way of experiments (e.g. the kinematics of the experiments were fixed). Later on, in the seventies, technologies became available to collide two particle beams head-on, which roughly increased the interaction energy with a factor of two.

Nowadays experiments are only based on colliding two particle beams head-on. Upon collision, several particles can be created according to the laws of quantum mechanics, either short or long-lived, which eventually decay in known SM particles or new particles wich are not yet discovered. Such particles eventually decay to stable particles which can be measured by installing detectors around the interaction point (IP). Large multi-purpose detectors are installed around the IP to measure the properties of each stable, visible particles such as momentum, mass and charge. Based on this reconstructed information, the underlying physics can be revealed by reconstructing those particles to their initial states, which comprises the physics theory.

Due to the probabilistic nature of quantum mechanics, several particles could be created and several final states are possible. Each process is described by a cross section σ, a number which is proportional to the interaction probability and can be calculated from the underlying theory. For processes with low cross section (e.g. Higgs boson, new physics), many collisions are necessary to be able to produce such processes. Hence, the aim of accelerators is to produce as many colissions as possible, to enhance the production rate of such rare processes. To deal with such amount of events, statistical methods are necessary to separate the to-be-studied process (the signal) from the other processes (the backgrounds). This can be done by having the knowledge of each background process, which can be achieved through event simulation.

3.1 Particle Accelerators and the LHC

As already stated in the introduction, nowadays in HEP experiments particles are accelerated to high energies and collided head on, preferentially with a high frequency for the observation or discovery of rare processes. The largest accelerator to date is the Large Hadron Collider [41] (LHC) at the European Organization for Nuclear Research (CERN), a circular accelerator with 27 km in circumference, on average 100 m beneath the France-Switzerland border near Geneva. It accelerates bunches of protons to 13 TeV center-of-mass energy and let them collide head-on in four interaction points (IP) along the ring. Around each IP, detectors are installed to study and reconstruct the particle collisions: A Toroidal LHC ApparatuS (ATLAS) and the Compact Muon Solenoid (CMS) are multi-purpose dectectors; the Large Hadron Collider beauty experiment (LHCb) focuses on b-physics and A Large Ion Collider Experiment (ALICE), specialized heavy-ion (Pb-Pb nuclei) interactions. The original idea of the LHC and experiments goes back to the late eigthies and the main purpose is to give answers to the SM electroweak sector, in particular to the existence of the Higgs boson.

To operate the LHC with such high energy and intensity, it relies on a series of injectors, which originally served other experiments in the past, but upgraded for LHC purposes. To start, the protons are obtained by stripping electrons of hydrogen atoms. The protons are injected in LINAC2, a linear accelerator that boosts the energy up to 50 MeV before injection in the (circular) Proton Synchrotron Booster (PSB). The PSB plays a crucial role in the formation of the proton bunches, each containing $\approx 10^{11}$ protons, accelerates the protons further to 1.4 GeV. Proton bunches are then injected to the Proton Synchrotron (PS) to accelerate the proton bunches to 25 GeV and followed by Super Proton Synchrotron (SPS) to increase the energy to 450 GeV. The proton bunches from the SPS are injected clockwise and anti-clockwise to the LHC ring with a 25 ns spacing, yielding a total $N_b \approx 2500$ bunches per beam. When the LHC is occupied with the 2500 bunches at 450 GeV, the energy is further ramped up to the design energy of 6500 GeV. To achieve the acceleration in the LHC and injectors, pulsed electrical fields are used (radiofrequency (RF) cavities). They are steered along the ring(s) using dipole magnets, which can go op to 8.3 Tesla for the LHC dipoles. The necessary current to deliver such magnetic fields is in the order of 10 kA, therefore the dipole magnets are made of superconducting materials and cooled to a few degree Kelvin. Apart from the RF cavities and dipole magnets, also quadrupole magnets are installed along the ring for beam focusing purposes. Once the beams are ramped to 6.5 Tev, steering magnets on both sides of the IP further squeeze the beam and are eventually brought to head-on collision. With a bunch spacing of 25 ns, a collision frequency of $f = 40$ MHz correspond, which means 40 million collisions per second.

From the physics point of view, it is desired to have high energy and intensity beams which can deliver a large amount of collisions in order to study processes with low cross section or to study the SM with high precision. The collision energy (center-of-mass energy) is directly related to the beam energy through $\sqrt{s} = 2E$, which is valid when two beams collide head on. The beam intensity and collision frequency is captured in a single parameter called luminosity. When the beam transverse beam profiles are Gaussian in shape with widths σ_x and σ_y (for the LHC, $\sigma_{x,y} \approx 17$ mm), the luminosity $\mathcal{L}$ can be calculated as [42]:

$$\mathcal{L} = \frac{N_1 N_2 f N_b}{4\pi\sigma_x\sigma_y}, \tag{3-1}$$

with $N_1(N_2)$ the amount of protons per bunch in beam 1(2), f the revolution frequency and N_b the amount of bunches. For the nomimnal LHC parameters, a design luminosity of 10^{34} cm^{-2}s^{-2} is expected. Although all parameters in Eq. 3-1 depend solely the accelerator capabilities (see discussion above), it has a direct relationship between the number of reactions per unit time R of a certain process and the corresponding cross section σ:

$$R = \mathcal{L} \cdot \sigma. \tag{3-2}$$

During the time of collisions, the amount of protons in the bunch is reduced as they take part in the collisions (proton burnup). Hence, the luminosity $\mathcal{L}$ decreases in time and can be described by an exponential decay. When the luminosity dropped below a certain value (typically after 12-20 hours after the first collision), the beams are dumped and the LHC is refilled with a fresh amount of proton bunches.

3.1.1 LHC Run I and II

The LHC became operational in 2009 with a reduced center-of-mass energy of 7 TeV in 2010-2011 and 8 TeV in 2012 due to magnet constraints. For this Run I, LHC delivered a total of $\int \mathcal{L}dt = 29.4$ fb^{-1} of integrated luminosity to the CMS experiment. In 2013-2014, a long shutdown was planned for accelerator and experiment upgrades, which allowed the accelerator to increase the beam energy to 6500 GeV. In 2015, the LHC was recommissioned and the first *pp* collisions at 13 TeV were delivered and recorded by the experiments. In the three successive years, 2016, 2017 and 2018, thanks to the high LHC availiblity and excellent performance, a total integrated luminosity of $\int \mathcal{L}dt = 163.2$ fb^{-1} was delivered to CMS with a peak luminosity up to 2×10^{34} cm^{-2}s^{-2}, twice its design value.

3.2 Detectors and the CMS Experiment

The Compact Muon Solenoid (CMS) detector is one of the four experiments at the LHC [43]. With its cylindrical shape coaxial to the beam axis, the entire space is almost instrumented with detectors. Being a general purpose detector able to run at the highest LHC luminosities, a wide physics program can be conducted, including physics at the TeV scale, Higgs physics, BSM physics and heavy ion collisions.

A sketch of the CMS detector is shown in figure **3-1**. The heart of CMS is composed of a superconducting magnet with radius of 3.15 m and length of 12.5 m [44]. A uniform field of 3.8 Tesla is generated coaxial to the beam axis generated by a current of ≈ 18 kA. Massive iron return yokes are installed around the coil to precisely confine the magnetic field lines yielding a concentrated field just outside the cylinder. Charged particles produced in the IP are bend in the magnetic field from where the momentum and charge can be calculated.

Several detectors are installed in- and outside the magnet to detect different kind of particles. Mainly they are classified as follows:

- **Tracker** detectors measure the position and timing of charged particles and are combined to tracks. They are located close to the beam axis and interaction point in order to capture all charged particles.

- **Calorimeters** absorbs hadrons, electrons and photons. Energy deposits in the calorimeters are measured and clustered for particle reconstruction. An electromagnetic calorimeter (ECAL) which is surrounded by the hadronic calorimeter (HCAL).

- **Muon spectrometer** registers the muon hits in the outer layers of the detector, embedded in the return yokes. Muons do not interact with the inner material but leave hits in the tracker. embedded in the return yokes

The CMS coordinate system is defined as follows: the origin is chosen to be the IP, the x-axis points to the center of the LHC ring, the y-axis points vertically upwards and the z-axis along with the beam line to obtain a right-handed coordinate system. A point is described by the radius r, the azimuthal angle ϕ and the pseudorapidity, defined as:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right) \tag{3-3}$$

Due to collision symmetry around the beam axis, the ϕ-direction is invariant.

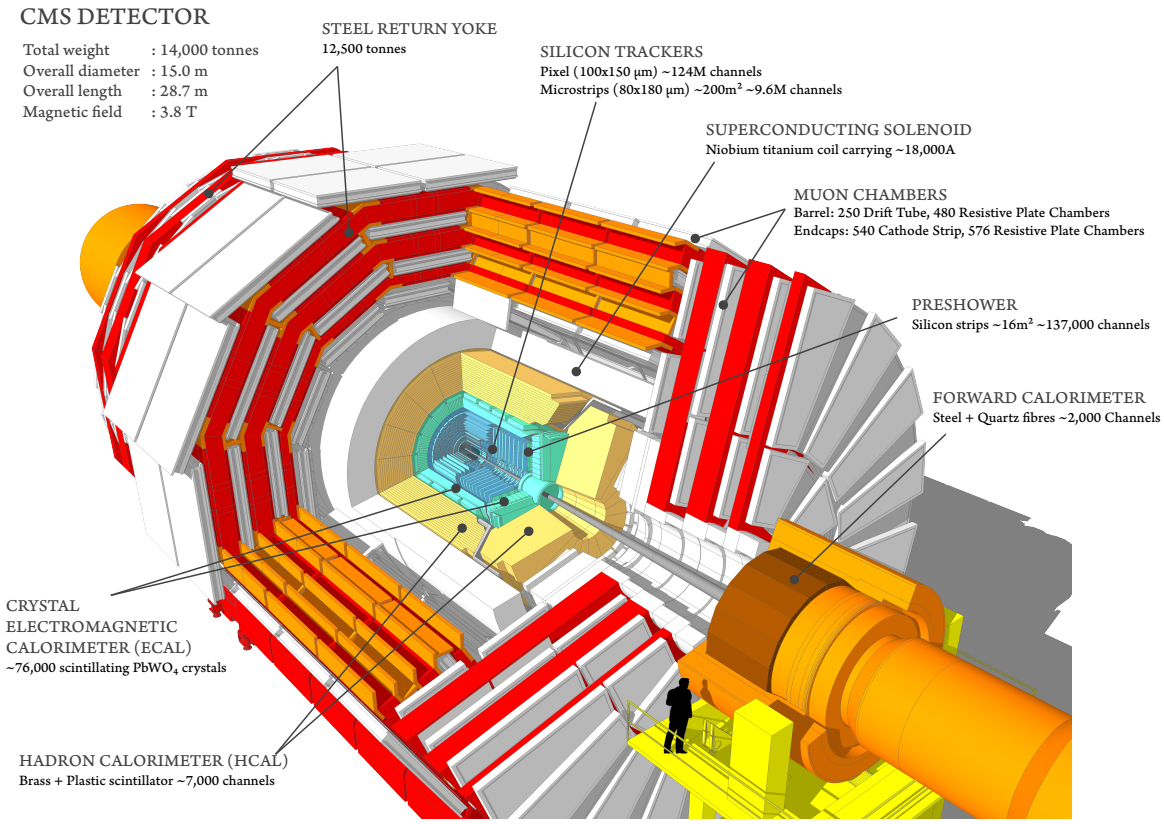


Figure **3-1**: Sketch of the Compact Muon Solenoid (CMS) detector [4].

3.2.1 Tracker

The CMS tracking system is composed out of two systems: a high granular pixel detector and silicon strip tracker [45]. They both are made of *np*-type silicium as active material and is operated at a temperature of -10 degC to suppress leakage currents induced by radiation.

The pixel detector is located near the beampipe and is comprised out of three barrel layers (with distances of 4.4 cm, 7.3 cm, and 10.2 cm respectively to the beamline), and two endcap disks (with distances of 34.5 cm and 46.5 cm from the IP) which covers an active area up to $|\eta| = 2.5$. In total, 66 million pixels of size 100 μm by 150 μm are installed resulting in an excellent position resolution. This is crucial for the tracking performance, secondary vertexing jet reconstruction. During the winter shutdown of 2016-2017, the pixel detector was replaced within the pixel Phase-I upgrade to accommodate with the higher luminosity and pileup conditions [46]. This new pixel detector consists of an additional barrel layer and three endcap disks to provide for hits for the track seeding algorithms. The readout system has been upgraded to reduce dead time and mimimize radiation degradation.

Surrounded by the pixel detector is the silicon strip tracker, whose barrel consists of 10 detection layers up to an outer radius of 1.1 m. The endcaps consist of 3+9 disks on each side of the barrel, extending the pseudorapidity to $|\eta| = 2.5$. It is divided in inner/outer barrel (4/6 layers) and inner disks (3 layers), supplied by the tracker endcap disks (9 layers). The silicon strips vary in size from 80 to 180 μm for each of the layers, yielding in total 9.3 million of strips providing at least 9 hits in the range of $|\eta| = 2.4$

3.2.2 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) identifies and measures the energy of electrons and photons, which are completely absorbed in the active material [47]. The active material is composed of lead tungstate crystals with a high degree of transparency. When an electron or photon hits the crystal, an electromagnetic shower is developed inside the crystal which produces scintillation light, proportional to the energy of the particle. Photodetectors at the back of the crystals converts the scintillation light to a measurable current, proportional to the incident scintillation light. Hence, the amplitude of the current pulse in the photodetectors is directly proportional to the energy of the incident particle.

The barrel ECAL consists of 61200 crystals with average face size of 2.4 cm by 2.4 cm with length of 23 cm, and covers up to $\eta = 1.479$. The endcap ECAL extends the pseudorapidity coverage up to $\eta = 3.0$ and contains in total around 7200 crystals with similar dimensions as the barrel crystals. In the region $1.653 < |\eta| < 2.6$, the preshower detector is installed mainly to detect neutral pions. It is a sampling calorimeter with lead layers as active material to initiate the shower and silicon layers for shower detection.

Both scintillation light production and the gain of the photodetectors are strongly dependent on the temperature. A stable operation temperature of 18 degC is maintainted with precision less than 0.05 degC to preserve an excellent energy resolution.

Continuous calibration of the ECAL is necessary to preserve a good energy resolution. Several effects contribute to the degradation of the ECAL response to incident particles, but radiation damage is a dominant contribution. Radiation dislocates atoms in the crystal that affects the transparency, hence the scintillation light. A dedicated laser monitoring system is in place to provide an in-time calibration for each crystal.

3.2.3 Hadronic Calorimeter

The hadronic calorimeter (HCAL) aims to detect hadrons and is crucial for jet reconstruction and the estimation of missing transverse energy [48]. It is a sampling calorimeter (cfg. electromagnetic preshower), where layers of active material of brass (alloy of copper and

zinc) or steel are alternated with plastic scintillation material to absorb and guide the light. The blue-violet scintillation light is finally collected through optical fibers and guided to photodetectors which transforms it to a measurable current. As is the case for ECAL, the amplitude and shape of the current pulses are directly proportional to the energy deposit of the particle in the active material.

Four regions are distinguished in the HCAL, which are in general designed to minimize the detection dead zone such that no particles are left unseen (except for the very forward region):

- Inner barrel (HB): located within the magnetic coil and the ECAL. It consists of 36 wedges in azumithal direction up to $|\eta| < 3.0$. Each wedge has different dimensions along or η and is composed alternating layers of brass and scintilalating material.
- Outer barrel (HO): a thin single layer of brass-scintillator material outside just outside the magnet coil are placed to measure any hadronic energy leaks from the HB. 5 rings are installed along the η-direction (following the geometry of the muon barrel chambers); the central ring 0 consists of two single layers.
- Endcap (HE): the hadronic calorimeter is extended on both endcaps to $1.3 < |\eta| < 3.0$, and the geometry has been optimized to reduce the dead zones between the HB and HE.
- Forward hadronic calorimeter (HF): the HF measures the hadronic activity in the high pseudorapidity region $3.0 < |\eta| < 5.0$. Due to the strong radiation environment in the forward region, steel is used as active material instead of brass.

3.2.4 Muon Spectrometer

Muons are registered in gaseous detectors embedded in the return yokes. Three different technologies provide the detection of muons to span the large volumes surrounding the magnetic coil in the barrel and endcaps up to $\eta = 2.4$ [49]. A muon, as minimum ionizing particle (MIP), ionizes the gas and by applying a strong electric field the electrons will cause an avalanche, inducing a signal on the readout materials. Custom electronics amplifies the signals and a digitization is performed before sending the data to the trigger and DAQ.

In the barrel, **drift tubes (DTs)** are installed, rectangular cells typically 4 cm wide and 1.3 cm in height and contains a wire under potential difference w.r.t. the cell edges. During ionization, electrons are accelerated and multiplied towards the wire and the signal is induced. Spatial resolution in both transverse plane and coaxial to the wire is provided by measuring the drift velocity of the electrons in the gas and the signal propagation velocity along the wire and respectively.

In the endcaps, **cathode strip chambers (CSCs)** are installed and are composed of a grid of dense positively charged wires and negatively charged copper strips. Ionization causes an avalanche in the direction of the strips, where a signal is induced on both wires (by the electrons) and strips (by the ions). Therefore it provides spatial resolution in two directions and are relatively fast due to the closely spaced wires.

A redundant system covering both barrel and endcaps is provided by the **resistive plate chambers (RPCs)**, who mainly contribute in the muon trigger. RPCs are parallel plate gaseous detectors made of two gas gaps, with a strip readout plane in between, providing an extra coordinate along the ϕ direction. RPCs have an excellent time resolution in the order of $2-3$ ns, making them suitable for muon triggering.

3.2.5 Trigger and Data Acquisition

The total pp inelastic cross section at 13 TeV is roughly 70 mb. Operating with a luminosity of 10^{34} cm^{-2}s^{-2}, the event rate is equal to 7×10^8 Hz (see eq. 3-2). As a result, the average number of pp collisions per bunch crossing (i.e. each 25 ns) is then calculated to be 17.5. During the LHC Run II data taking, this number is measured to be ≈ 30, because not all bunches are occupied and the instantaneous luminosity can reach higher levels. The readout of the channels of each subdetector and the storage typically requires 1-2 mb per event. With such high interaction rates, a data transfer rate beyond TByte/s is needed to store each event, which is currently (and in the nearby future), not achievable. Clearly, both high transfer rates and the amount of interactions per bunch crossing are practically difficult to deal with. Fortunately, the absolute majority of events are not of interest as for a hadron collider, the SM governed QCD cross sections are typically very high and such events have soft jets in their final states. It is the hard interaction, where two quarks in the protons contain sufficient energy to create heavy resonances or particles which is interesting to study. For example, at 13 TeV, the process $W \to \ell\nu$ occurs at a rate of 100 Hz and $t\bar{t}$ is produced at 10 Hz.

An online trigger system is implemented using coarse detector information to quickly analyze each event and decides whether the event is interesting or not [50, 51]. It consists of a two-level system, to systematically reduce the rates for event storage:

- Level-1 trigger: algorithms are encoded in hardware for fast reconstruction of jets (including τ-jets), jet counters, electrons photons, muons and missing transverse energy. It consists of a global muon trigger (DT+CSC+RPC) and global calorimeter trigger (ECAL+HCAL+HF) to reduce the event rates from 40 MHz to 100 kHz.

- High-level trigger (HLT): events selected by the level-1 trigger are further validated by the HLT, processed in computer farms. Its main purpose is to further reduce the event rates from 100 kHz to 100 Hz which is possible to store. It includes the tracker information and the reconstruction algoritms and techniques are similar to the offline reconstruction. Therefore the HLT provides accurate information of the event and efficient event selection is possible. The reconstruction outcome is also used for an online data-quality monitoring (DQM).

The trigger rules to decide whether an event is kept or not are encoded in trigger menus, which can be changed online during data taking. Trigger menus basically contains thresholds of the object momenta and are accurately studied in simulation to not and are driven by the physics. Several trigger menus are deployed online. Once an event passes the HLT criteria, the detector information is concentrated and stored to disc by the data acquisition system (DAQ). During online data-taking, a prompt reconstruction is performed for further data quality monitoring (DQM), used by online and offline shifters. The central service of data storage and computing power makes use of the Worldwide LHC Computing Grid (WLCG).

3.3 Event Reconstruction and Simulation

The HLT and offline reconstruction is done by a common framework called CMSSW. It has as main purpose to interpret the detector data to reconstruct the visible particles by using sophisticated algorithms. A detailed prescription of each subdetector and the magnetic field mapping is stored, and is continuously updated to be consistent with the detector or for improvements of reconstruction algorthms and DQM purposes. Furthermore, it contains the necessary tools for calibration and alignment of detectors, which is necessary to finally provide accurate information about charge, energy and momenta of the reconstructed particles. Besides the recorded data, physics analysis heavily relies on Monte Carlo generated events, necessary to evaluate the signature of known SM and unknown physics processes in the CMS detector.

3.3.1 Event Reconstruction

Depending on the particle type, the hits in spacetime are reconstructed and connected with each layer of detector. For fast processing, the L1 and HLT use coarse information of the detector which are formed by clustering hits). The full offline event reconstruction is based on the Particle Flow (PF) algorithm [52], which as as objectibe to reconstruct all visible objects such as photons, electrons, muons, jets and missing transverse momentum muons, electrons, photons, and charged and neutral hadrons.. Crucial information is delivered by the bending of charged particles due to the magnetic field. A visualisation of the The PF relies on local detector calibrations which are propagated as corrections to the objects. Associated with such corrections are their uncertainties, which should be propagated to the final statistical treatment of the analysis (see section 3.4). The success of the PF algorithm is devoted due to the fact that all layers in the detector are correlated.

For the relevant objects used in this thesis/analysis, the reconstruction and corrections are described in detail.

Tracking and Primary Vertex

Tracks based on tracker hits are clustered and fitted using a Kalman-Filter [53] (KF) to obtain the track origin, momentum and direction. A list of primary vertices is then reconstructed from these tracks using the *Deterministic Annealing* (DA) clustering algorithm [54]. The vertex distance to the nominal interaction point is required to be originate from the beam interaction point (typically few mm within the transverse plane and 20 cm in the z-direction) and the number of degrees of freedom of the vertex fit is required to be larger than 4. Out of the selected vertices, the one with the largest summed squared p_T of the tracks associated to that vertex is chosen as the corresponding to the hard scattering process for the event. Tracks which do not originate from the primary vertex are referred as pile-up.

Muons

Muons are identified using tracks in the silicon tracker which should be compatible with hits in the muon chambers, and in addition compatibility with calorimeter deposits [55] is required. A global fit of the tracks and muon path curvature is performed to estimate the charge and momentum. The momentum resolution for low and intermediate transverse momenta ($p_T < 200$ GeV) is measured to be 1% in the barrel and 3% in the endcaps with an average uncertainty of 5%, and is mainly driven by the performance of the inner tracker [55]. As the muon system covers up to $|\eta| < 2.4$, the muon reconstruction is restricted to this range of pseudorapidity.

Two working points for muon identification are defined: a medium working point with 97% identification efficiency and a loose working point with >99% identification efficiency [55]. Correction factors and uncertainties for muon identification, trigger, isolation and reconstruction are factorized and briefly given in section 4.3.2.

In addition, isolation criteria are required using the *mini-isolation* $I_{\text{mini-iso}}$ in order to suppress nonprompt and misidentified muons. It is defined as the p_T sum of the the PF candidates in a cone around the lepton, divided by the transverse momentum of the lepton. The cone size ΔR is varied with the lepton p_T as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 10 \text{ GeV}/\min(\max(p_T(\ell), 50 \text{ GeV}), 200 \text{ GeV}), \quad (3\text{-}4)$$

leading to cone radii from 0.05 to 0.20. Indeed, at higher lepton energies the objects are more boosted (resulting in more) and the accidental overlap of jet activity around a lepton is lowered when reducing the cone size. A tight (loose) isolation criterion with discriminant $I_{\text{mini-iso}} < 0.1$ (< 0.4) is defined.

Electrons

Electrons are reconstructed starting from the clusterization of ECAL single-cell hits and building superclusters (SC) around a seed cluster. SCs are narrow in η but extended in ϕ in order to collect possible bremsstrahlung photons which are radiated by the electron in the tracker material. The SCs are then matched to track seeds, which are built from pairs or triplet hits in the tracking system. Based on these track seeds, electron tracks are built by re-fitting the tracks with a Gaussian-sum filter (GSF) algorithm to estimate properly the electron track parameters. The GSF algorithm is preferred over the KF algorithm in case the electron energy losses partly its energy through bremsstrahlung. From Z decays, the momentum resolution for electrons with $p_T \approx 45$ GeV ranges from 1.7% for nonshowering electrons in the barrel region to 4.5% for showering electrons in the endcaps [56]. The ECAL fiducial region is defined by $|\eta_{SC}| < 1.4442$ (barrel) or $1.566 < |\eta_{SC}| < 2.5$ (endcap), where $|\eta_{SC}|$ is the pseudo-rapidity of the supercluster.

The selection of electrons used in the analyses presented in this thesis is based on a multivariate (MVA) identification algorithm. It combines several variables describing the shape of the energy deposits in the ECAL, as well as the direction and quality of the associated tracks [57]. Custom MVA working points are defined based on the resulting identification efficiency: the tight(loose) identification efficiencies are measured to be 88(95)% in in $t\bar{t}$ events. Identical isolation working points based on mini-isolation are defined.

Jets

Hadronic jets are clustered from the reconstructed PF candidates. The anti-k_T algorithm [58, 59] with a distance parameter of 0.4 is being used. The basic idea of this algorithm is to loop over the PF candidates and cluster them if they are within a certain distance defined by a custom distance-function. Since all the PF candidates are used as input to the clustering algorithm, loose identification criteria are applied to the jets, in order to distinguish them from leptons. After the clusterization, the jet momentum is defined as the vectorial sum of the clustered PF candidates. From simulation, it is found that the jet momentum is within 5 to 10% of the true momentum over the whole p_T spectrum and detector acceptance. The highest jet performance is obtained within the tracker acceptance ($|\eta| < 2.4$, *central jets*) which can provide charged PF candidates.

To mitigate the effect of pileup, charged hadrons that do not arise from the primary vertex are removed; a procedure known as *charged hadron subtraction* [60]. Jet energy corrections, derived from simulation, aim to correct the detector response A residual offset correction for pileup is applied to the jet momentum.

Jet energy corrections are derived from simulation to bring the measured response of jets to that of particle level jets on average. Such corrections include a offset correction for pileup, simulated response corrections and flavor corrections. In situ measurements of the momentum balance in dijet, photon + jets, Z + jets, and multijet events are used to account for any residual differences in jet energy scale between data and simulation [61]. The jet energy resolution amounts typically to 15% at 10 GeV, 8% at 100 GeV, and 4% at 1 TeV [62].

Hadronic Tau Decays

In about 65% of the cases, tau leptons decay hadronically to neutral and charged pions which give signatures in tracker and calorimeters. Hadronic taus (hereafter referred as τ_h) are reconstructed using the Hadron-Plus-Strips (HPS) algorithm [63, 64] which uses clustered anti-k_T jets as seeds. PF charged hadron candidates and PF photons (originating from $\pi^0 \to \gamma\gamma$) located in η-strips in the vicinity of the charged hadrons are combined to form categories: single charged hadron (one-prong τ_h), charged hadron plus one or two strips (to reconstruct the one prong τ_h decays in association with π^0) and three charged hadrons (to reconstruct the three-prong τ_h decays). The τ_h reconstruction is typically performed in the following kinematic regime: $p_T > 20$ GeV and $|\eta| < 2.3$.

The jets originating from the hadronization of quarks or gluons misidentified as τ_h are suppressed using a multivariate discriminant [64]. Typical input variables to the MVA discriminant are τ_h isolation parameters (based on the surrounding hadronic activity) and the τ_h lifetime, estimated from tracks of the τ_h decay products. A loose working point

is used for this discriminant, corresponding to ≈50% identification efficiency, determined from $Z/\gamma^* \to \tau^+\tau^-$ events, and a probability of 3×10^{-3} for misidentifying a jet as a τ_h, determined from QCD multijet events. τ_h candidates are rejected if they are consistent with the hypothesis of being muons or electrons misidentified as τ_h.

b-Jets

Jets are identified as originating from the fragmentation of b quarks by using the Combined Secondary Vertex (CSV) algorithm [65, 66]. A b-quark hadronizes to b-hadrons which have relatively large lifetime before decaying into other particles and therefore they can be regarded as a jet with secondary (displaced) vertex. The CSV algorithm combines the information about the track impact parameters and identified secondary vertices within jets, even when full vertex information is not available. Additional categories for jets where a pseudo vertex is found, or no vertex at all is identified, can be defined and combined in a neural network discriminant to provide maximal separation of b jets from the much larger background of jets arising from the fragmentation of charm quarks, light quarks and gluons. The algorithm provides a continuous discriminator output, that is used to select optimal working points. The medium working point corresponds to a 65% selection efficiency of genuine b jets, and the probability to misidentify jets originating from light-flavor quarks or gluons as b jets is around 1% [67]. As the tracking information is crucial to the CSV algorithm, the b jet identification is only possible up to a pseudorapidity of $|\eta| < 2.4$.

Since the b tagging efficiencies measured in data are somewhat different than those predicted by the simulation, scale factors are applied to simulated events. Such scale factors depend in general on the jet flavour (b, c, light jets), jet p_T, and jet η.

Missing Transverse Energy

The Missing Transverse Energy (E_T^{miss}) is a measure of the imbalance of the transverse momentum in the event and is computed as the negative vector sum of the p_T of all reconstructed PF candidates [68]. The jet energy corrections (and eventual τ_h energy scale corrections) described above are propagated to the E_T^{miss} calculation to correct for the imperfect knowledge of the scale of the jet energy and the unclustered energy.

Additional event cleaning to account for instrumental noise in the detector and to reduce the impact of the non collisional background, E_T^{miss} filters are applied as summarized below:

3.3.2 Event and Detector Simulation

Physics analyses heavily rely on event simulation, which for a given process (e.g. $pp \to W \to \ell\nu$) estimates the detector response. It is crucial for the understanding of the SM processes in data, but also to estimate the expected outcome of new physics processes (e.g. $H+ \to \tau\nu$) in data. Monte Carlo techniques are deployed for massive event generation campaigns such that sufficient statistics is available to perform a solid statistical analysis.

The full chain of event simulation is based on three steps:

1. **Event generation.** Several tools are available for event generation, but CMS mainly adopts the MadGraph5_aMC@NLO [69] and POWHEG [70, 71, 72, 73, 74] event generators. Both generators can calculate SM and non-SM processes up to NLO accuracy but use different techniques to handle the hard scattering and the matrix element partons. Due to their different implementations, POWHEG fully takes into account spin correlations of the outgoing particles and generates events solely with a positive weight. On the other hand, MadGraph5_aMC@NLO the spin correlations are not fully taken into account and a substantial amount of generated events is negative. A third event generator PYTHIA [75] is also deployed is mainly interfaced with the other event generators to model the parton showering, fragmentation and decay of the tau leptons.

2. **Detector response simulation.** A full simulation of the CMS detector is provided by Geant4 v9.4 software [76], which processes the generated events. Pileup events are modeled by generating minimum bias collision events with PYTHIA and added to the simulated hard scattering events. Both multiple inelastic pp collisions per bunch crossing (in-time pileup) and interference collisions from preceding and subsequent bunch crossings (out-of-time pileup) are taken into account.

3. **Event reconstruction.** The outcome of the simulated detector response is very similar to what is obtained in data. The reconstruction is therefore quasi identical between data and simulated events.

3.3.3 Overview of Monte Carlo Samples

The simulated SM processes for the charged Higgs analyses are listed below. The event generators used are MadGraph5_aMC@NLO v2.2.2 [69] or POWHEG v2.0 [70, 71, 72, 73, 74], and interfaced with PYTHIA 8.212 for hadronization. The PYTHIA parameters affecting the description of the underlying event are set to the CUETP8M1 tune [77] for all processes except $t\bar{t}$, for which a customized CUETP8M2T4 tune [78] is used. The NNPDF3.0 parton

distribution functions (PDFs) [79] are used for the generation. To match the pileup distribution with the one observed in data, the simulated events are weighted per event. For the LHC Run II in 2016, the average interactions per bunch crossings was measured to be 23.

- top-pair production ($t\bar{t}$): two event generators are used: MadGraph5_aMC@NLO and POWHEG. The MadGraph5_aMC@NLO is used for a more accurate jet modeling as it takes up to 3 partons in the matrix element. the POWHEG campaign has substantially more statistics and is generated with FxFx jet matching and merging [80].

- $W^{\pm}$+jets, Z/γ^*+jets, QCD: events are generated at LO using MadGraph5_aMC@NLO v2.2.2 with up to four noncollinear partons in the final state [81].

- single top: the POWHEG generator is used to model single top quark production via t-channel and the associated tW production [82, 83]. For the s-channel production, the MadGraph5_aMC@NLO v2.2.2 generator is used.

- diboson (WW, ZZ, WZ): the diboson processes are simulated using PYTHIA 8.212.

- top-pair+$W^{\pm}$/Z ($t\bar{t}$ + X): the MadGraph5_aMC@NLO v2.2.2 generator is used.

Each process is normalized to the amount of generated events (with a weighthed procedure to correctly taken into account any negative events) and their theoretical cross sections. For the $t\bar{t}$ background and the single top quark background in the s and tW channels, the cross sections are calculated at next-to-next-to-leading-order precision [84, 85], yielding a $t\bar{t}$ cross section of $\sigma_{t\bar{t}} = 831.76$ pb. NLO precision calculations are used for single top quark production in the t-channel, and for the $W^{\pm}$+jets, Z/γ^*+jets, and diboson processes [86, 85, 87, 88].

As explained in section 2.2, three mass ranges are distinguished for the search of a charged Higgs boson. Their cross sections are normalized to 1 pb (unless stated otherwise), and similar event normalizations are applied as for the processes described above. Separated samples were produced up to 3 TeV for tb and $\tau\nu_\tau$ decays, and the amount of generated events as function of the charged Higgs mass sample is shown figure **3-2** (left). A substantial amount of events is negative due to the generator used, which can go up to 33%, as shown in figure **3-2** (right)

- Light regime (80 to 160 GeV): the light mass samples are generated at NLO with the MadGraph5_aMC@NLO v2.3.3 [69] generator, assuming the $H^{\pm}$ production occurs through top decays. Only $\tau\nu_\tau$ decays are simulated.

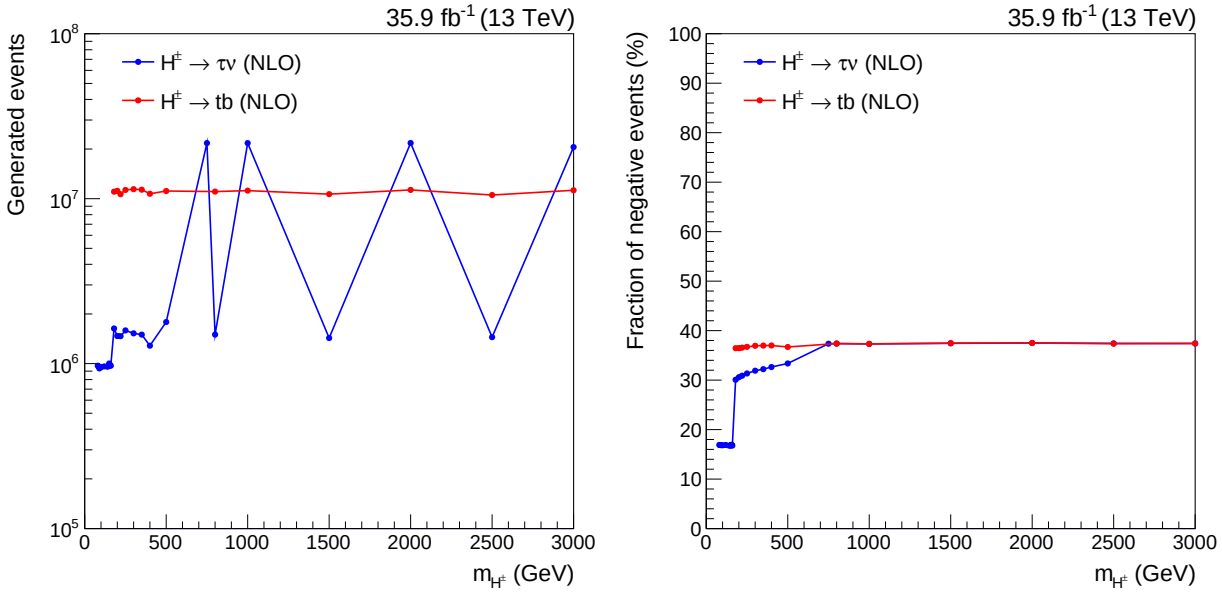


Figure **3-2**: Left: amount of generated events for each mass sample and for both tb and $\tau\nu_\tau$ campaigns. Right: fraction in % of negative events per mass point.

- Intermediate regime (165-180 GeV): the samples are generated at LO using the MadGraph5_aMC@NLO v2.3.3 with the model described in Ref. [89] (no NLO prescription is available to date). A more dedicated approach to treat this LO-NLO ambiguity is given in section 4.11. Only $\tau\nu_\tau$ decays are simulated.

- Heavy regime (180 GeV to 3 TeV): samples are generated at NLO with the MadGraph5_aMC@NLO v2.3.3 [69] generator, for both tb and $\tau\nu_\tau$ decays.

3.4 Analysis Strategy and Statistical Methods

Due to the high integrated luminosity delivered by the LHC, the physics datasets are very large and overwhelmed by SM processes, especially $t\bar{t}$ production the production of vector bosons ($W^{\pm}$, Z) in association with jets. New physics phenomena (e.g. a charged Higgs boson) manifest themselves as an excess of data in an isolated phase space which is sensitive to the signal. The signal is defined as the detector response of the hypothetical particle under certain assumptions (e.g. mass and width of the particle, cross section, etc.), and is expressed as number of events or as a (kinematic) distribution. The signal strengths are typically very low compared to the SM production rate and statistical methods are used to fit the signal and the known SM backgrounds to the data. With backgrounds it is meant the sum of all the known SM processes, either estimated from Monte Carlo simulation or data-driven techniques.

Several sources of systematic uncertainties must be taken into account in the statistical analysis. The simulation, used to model the signal (charged Higgs) and backgrounds (SM processes), is affected by two different sources of systematic uncertainty. The first source is related to the uncertainty on the correction factors applied to the simulation in order to better reproduce the detector conditions and performance in data. The second source copes for assumptions made in the theoretical models that were used to produce the simulation. Moreover, when data-driven techniques are deployed, specific systematic uncertainties must be derived to cope for any mis-modeling of the underlying prescription. Each event phase space has their own set of experimental uncertainties. Upon combination several event phase spaces (e.g. several final states), the correlation of such uncertainties must be carefully studied.

To optimize the sensitivity of the search, typically one tries to optimize the event phase space by maximizing S/B, where S represents the signal yield and B the total SM background yields. The traditional way of maximizing S/B is by cutting the event phase space based on variables which are sensitive to the signal. Experimental deficiencies must be taken into account (e.g. b tagging mis-identification, τ_h mis-identification) as well as acceptance effects. Nowadays, with large datasets and increased computer resources, more sophisticated techniques can be used rather than cutting on single variables. Most common are the multivariate analysis classifiers (MVA). Based on a set of input variables which have a higher discrimination power between signal and background, an MVA is trained against signal and backgrounds (binary classifier), based on a subset of signal and background events. When applying on the the remaining events and the data, the MVA returns a value per event (typically between 0 and 1) based on the per-event input variables, whose value is higher for more signal-like events. One can then use this output discriminator directly for fitting or to

provide an additional cut.

Once the event phase space, or signal region, is defined, the remaining events are fitted to the data, taking into account the necessary systematic uncertainties (commonly denoted as nuisance parameters). This can be done either by just fitting the absolute event yields (counting experiment), or by fitting a shape distribution which has an optimal discrimination between signal and background. The advantage of fitting such distributions, usually represented by histograms, is that the shape has more information than a simple counting experiment as more bins are involved in the fit, each bin with their respective S/B ratio. Moreover, regions with low S/B (background abundant) can be used to constrain the nuisances which can be dominant for the sensitivity of the analysis. Sometimes entire control regions, regions with negligible signal yields, are used in a simultaneous fit in order to precisely constrain the backgrounds.

The techniques described above are extensively used and evaluated in the search of a charged Higgs boson. Several additional techniques are deployed and elaborated in detail in the dedicated chapters. Because the fitting part takes a crucial role in the analysis, the basic concepts are introduced below. Especially, when no charged Higgs signal is observed in the data, one usually quotes an upper limit on the production cross section, which can also serve as a benchmark of the analysis.

3.4.1 Likelihood and Nuisance Parameters

A set of histogram representing the signal, backgrounds and data are the outcome of the analysis optimization and used to perform a statistical analysis (a continuous representation rather than histograms can also be used). The expected event content in bin i can be written as $E[n_i] = \mu s_i + b_i$, where $i \in [0, N]$ where N is the amount of bins. The total background events b_i can be composed of several sources, each representing a SM process (e.g. $t\bar{t}$, W+jets, DY+jets, etc.):

$$b_i = \sum_{j}^{\text{sources}} L\sigma_j\epsilon_{ij}, \tag{3-5}$$

with L the integrated luminosity (representing the size of the dataset), σ_j the cross section of process j (in pb) and ϵ_{ij} the efficiency in bin i of process j after all the event selections. s_i is the expected signal yield, obtained with the same formula as above but with a normalized cross section (usually one normalizes the signal to 1 pb). The parameter μ is the signal strength and must be obtained from the fit ($\mu = 0$ is no signal, $\mu = 1$ equals a given normalized amount of signal events).

Poisson Likelihood

Auxiliary measurements in control regions represented by M bins have an expectation value $E[m_i] = u_i$, with u_i the total predicted background events. The Poisson likelihood is then defined as the product of Poisson probabilities in each bin:

$$L(\mu) = \prod_{i=1}^{N} \frac{(\mu s_i + b_i)^{n_i}}{n_i!} \exp(-(\mu s_i + b_i)) \prod_{j=1}^{M} \frac{u_j^{m_j}}{m_j!} \exp(-u_j). \tag{3-6}$$

The free parameters in the Poisson likelihood are obtained by maximizing the likelihood w.r.t. these parameters (in this case μ), or equivalently minimizing the negative log likelihood, defined as

$$-\ln L(\mu) = \sum_i (-n_i \ln(\mu s_i + b_i) + (\mu s_i + b_i)) + \sum_j (-m_j \ln u_j + u_j), \tag{3-7}$$

where the constant terms have been dropped (except for the term in u_j). The minimization is usually performed with numerical methods such as implemented in MINUIT [90].

Nuisance Parameters

The experimental and theoretical uncertainties described above can be represented by two types of nuisance parameters: rate and shape uncertainties [91]. In general, for each nuisance a free parameter is introduced in the likelihood. Therefore, one often writes the likelihood as $L = L(\mu, \boldsymbol{\theta})$ with $\boldsymbol{\theta} = (\theta_1, \theta_2, ..., \theta_n)$ representing the n nuisances; the minimization is performed simultaneously over all the parameters.

Rate uncertainties or multiplicative uncertainties directly change the event rate of a process j by adding a factor κ_j in eq. 3-5. κ_j follows a certain probability density $\mathcal{R}(\kappa_j|1, \sigma_j)$ with mean 1 and a prior standard deviation σ_j. It enters the likelihood as a multiplication factor:

$$L(\mu) = \prod_{i=1}^{N} \frac{(\mu s_i + b_i)^{n_i}}{n_i!} \exp(-(\mu s_i + b_i)) \prod_{j=1}^{M} \frac{u_j^{m_j}}{m_j!} \exp(-u_j) \mathcal{R}(\kappa_j|1, \sigma_j), \tag{3-8}$$

and can be regarded as a penalty term in the negative log likelihood. The floating parameters κ_j are simultaneously minimized, and can be constrained in the control regions. The physical bound on κ_j is that the parameter remains positive, which is not strictly the case when using Gaussian priors. Therefore, often one uses log normal priors which cannot result in negative values of κ_j.

Shape uncertainties alter the shape of the underlying distribution and directly change the

bin efficiencies ϵ_{ij}. For a given uncertainty, one typically provides the up and down efficiencies $\epsilon_{ij}^{\pm}$ which are compared to the nominal efficiency ϵ_{ij}^{0}. Estimations of the uncertainty are obtained by changing event or object scale factors (e.g. pileup, energy scale, etc.). In order to implement them in the likelihood, a continuous parameterization of the uncertainty as function of $\epsilon_{ij}^{0}, \epsilon_{ij}^{\pm}$ is needed, that can be achieved by writing the efficiency as ϵ_{ij} function of a vertical morphing parameter f, whose nominal value is zero. The uncertainty is then usually expanded quadratically for $|f| < 1$, and linearly beyond. The morphing parameter is then minimized with the other set of parameters in the likelihood.

Statistical Uncertainties

Statistical uncertainties are inherent on the limited size of the Monte Carlo samples. For each process in each bin, a statistical uncertainty is estimated. The Barlow-Beeston prescription [92] defines for each bin and process a Gaussian multiplicative uncertainty, which is minimized together with the other parameters. Considering the large amount of bins and processes in the analyses, this method becomes very computationally expensive and is practically not suitable. A Barlow-Beeston lite procedure is adopted [91], where one Gaussian multiplicative uncertainty is introduced, and assigned to the process with the largest background yield. The σ of the Gaussian constraint term should represent the total statistical uncertainty for all processes. Besides the drastic reduction of the free parameters in the likelihood, the uncertainty constraint terms can be solved analytically and no computational minimization is necessary.

3.4.2 The CL_s Method for Upper Limits

Statistical methods are used to test whether the data is compatible with a given signal hypothesis. In high energy physics, one usually refers to the background-only and signal+background hypotheses, and both are tested against each other. For discovery, one usually rejects the background-only hypothesis w.r.t. the signal+background one, whereas for setting upper limits, the signal+background hypothesis is rejected w.r.t. the background-only. The level of agreement of a given hypothesis is quantified by the p-value, where the hypothesis is excluded if the p-value is below a certain threshold. Usually this is translated in an equivalent significance Z, by computing the Gaussian distributed standard deviation corresponding to the same p-value, i.e. $Z = \Phi^{-1}(1-p)$, with Φ^{-1} the inverse cumulative distribution. To express the sensitivity of an experiment, one usually quotes the median (expected) significance

To test and eventually reject the hypotheses, a test statistic has to be defined. It is based

on the profile likelihood ratio, defined as:

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\boldsymbol{\theta}}})}{L(\hat{\mu}, \hat{\boldsymbol{\theta}})}. \tag{3-9}$$

It is bounded as function of μ with $0 \leq \lambda(\mu) \leq 1$ and higher values imply good compatibility between data and the signal with strength μ. In the numerator, the likelihood is evaluated for a given signal strength μ with $\hat{\hat{\boldsymbol{\theta}}}$ the nuisance parameters that maximizes L for that specified μ. The denominator is the unconditional likelihood, maximized for both μ and $\hat{\hat{\boldsymbol{\theta}}}$. The test-statistic q_μ for upper limits is then defined as:

$$q_\mu = -2 \ln \lambda(\mu) \quad \text{for} \quad \hat{\mu} \leq \mu; \qquad q_\mu = 0 \quad \text{otherwise.} \tag{3-10}$$

Higher values of the test statistic represents less compatibility with between the data and the signal with strength μ. Following modified frequentist CL_s criterion [93, 94], a p-value for the signal+background and background-only are computed according to:

$$p_\mu = \int_{q_{\mu,\text{obs}}}^{+\infty} f(q_\mu|\mu) \mathrm{d}q_\mu, \tag{3-11}$$

$$1 - p_b = \int_{q_{\mu,\text{obs}}}^{+\infty} f(q_\mu|0) \mathrm{d}q_\mu, \tag{3-12}$$

where $q_{\mu,\text{obs}}$ is the value of the test statistic observed from data (for a given μ) and $f(q_\mu|\mu)$ the probability distribution of the test statistic for a given value of μ. Both p-values are then combined in a single variable:

$$CL_s(\mu) = \frac{p_\mu}{1 - p_b}. \tag{3-13}$$

The value of μ is then searched such that $CL_s(\mu_\alpha) = 1 - \alpha$, and is denoted as the confidence level for the upper limit. Usually one takes $\alpha = 0.05$ and limits are quoted with 95% confidence level.

Crucial in the above formulas is the pdf of the test statistic, that can be obtained by Monte Carlo toys but is usually computationally expensive. Wald [95] showed that the profile likelihood can be estimated as follows, assuming a single parameter of interest:

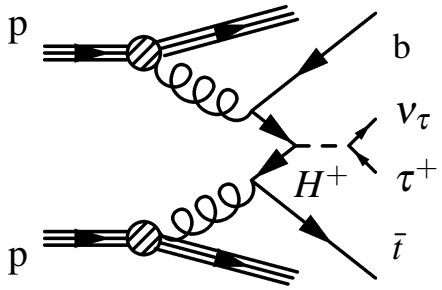
$$-2\lambda(\mu) = \frac{(\mu - \hat{\mu})^2}{\sigma^2} + \mathcal{O}(1/\sqrt{N}), \tag{3-14}$$

with $\hat{\mu}$ following a Gaussian distribution centered around μ' and standard deviation σ. N

represents the sample size, and higher order terms can be neglected in case the sample size is large. The standard deviation can be obtained by the Fisher information matrix constructed using Asimov dataset [96], defined as the dataset where the observed datapoints are replaced by the expectation values obtained from Monte Carlo.

4 Searches in the $\tau\nu_\tau$ Leptonic Final States

On the right figure the Feynman diagram of the 4FS production and decay process in the $H^\pm \rightarrow \tau^\pm\nu_\tau$ final state is shown. From this intermediate final state, several final states are possible: fully hadronic, semi-leptonic and the di-lepton final state. The fully hadronic final state requires the isolation of an hadronic tau with an additional lepton veto. In the semi-leptonic final state, either the tau lepton or the associated top quark decays leptonically, in the latter case also an isolated hadronic tau needs to be resolved. At last, in the di-lepton final state both tau and top decay leptonically.



In this chapter the $H^\pm \rightarrow \tau^\pm\nu_\tau$ final state is described, evaluated in the leptonic final states. Three mass regimes are considered in this analysis (see section 3.3.3 for details on the Monte Carlo sample production and event generators):

- light regime: charged Higgs hypotheses ranging from 80 to 160 GeV. NLO Monte Carlo samples are used;
- intermediate regime: charged Higgs hypotheses around the top mass ranging from 160 to 180 GeV. LO Monte Carlo samples are used taking into account interference with neutral hadrons. A dedicated study on the intermediate mass regime and the LO to NLO transformation is given in section 4.11;
- heavy regime: charged Higgs hypotheses ranging from 180 GeV to 3 TeV. NLO Monte Carlo samples are used.

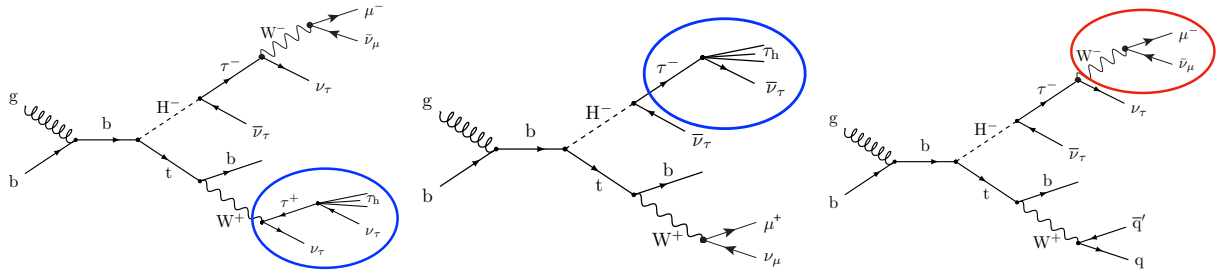


Figure **4-1**: Feynman diagrams of the possible leptonic decay modes.

4.1 Overview

The final state considered in this analysis is semi-leptonic and is conducted in two channels: $\ell + \tau_h$ where an hadronic tau is resolved and $\ell+$ no τ_h where a veto on the hadronic tau is applied. This categorization reflects the different signal and background natures and increases the sensitivity of the analysis. Indeed, in the former category two leptons are present which has mainly the di-leptonic $t\bar{t}$ background whereas the latter case consists mainly of semi-leptonic $t\bar{t}$ and $W^{\pm}$+jets. Other minor backgrounds such as Z/γ^*, single top and diboson are added to the analysis. In the $\ell+$ no τ_h, QCD multijet backgrounds leading to mis-identified leptons or taus can contaminate the event phase space and care has to be taken to suppress such QCD multijet events to a negligible amount (see Section 4.9.2).

The backgrounds are validated with data in dedicated control regions with low signal sensitivity, opposed to the blinded signal regions with higher signal sensitivity. Afterwards, a simultaneous fit is performed on all signal and control regions to extract the signal strength, based on transverse mass templates. Model independent limits as function of the charged Higgs boson mass are presented, and can be reinterpreted into model specific limits, where the underlaying free parameters are usually different (e.g. m_A, $\tan\beta$) and model specific.

4.2 Event Topology

The Feynman diagrams for the signal is shown in figure **4-1**. The event topology in the leptonic final state consists of an isolated lepton, missing transverse energy, jets where at least one is b-tagged and optionally an hadronic tau required in the $\ell + \tau_h$ category. Having such objects in the final state, the main corresponding background is $t\bar{t}$ production. The expected event signatures for the charged Higgs decay and $t\bar{t}$ decays is schematically shown in table **4-1**. It is clear that the semi-leptonic $t\bar{t}$ background is dominant in the $\ell+$ no τ_h category whereas the $t\bar{t}$ di-lepton background is dominant in the $\ell + \tau_h$ category.

category	**signal H^+**	$t\bar{t}$ **background**
$\ell+$ no τ_h	$tH^+ \to Wb + \tau\nu \to qqb + \ell\nu\nu\nu$	$t\bar{t} \to WbWb \to qqb + \ell\nu b$
	- lepton from τ-lepton decay	- lepton from $W^\pm$ decay
	- E_T^{miss} exclusively from H^+ decay	- E_T^{miss} from $W^\pm$ decay
	- jets from associated top	- jets from other top decay
$\ell + \tau_h$	$tH^+ \to Wb + \tau\nu \to \ell\nu b + \tau_h\nu\nu$	$t\bar{t} \to WbWb \to \tau_h\nu b + \ell\nu b$
	- lepton from $W^\pm$ decay	- lepton from $W^\pm$ decay
	- E_T^{miss} shared between $H^+W^\pm$ decay	- E_T^{miss} shared between $W^\pm$ decays

Table 4-1: Schematic overview of the categories used in this analysis with their signal and $t\bar{t}$ background signatures.

4.3 Object Definitions

4.3.1 Jets, τ-leptons and E_T^{miss}

Central jets are used with $|\eta| < 2.4$, compatible with the tracker coverage necessary for the secondary vertex b tagging algorithm. A p_T of at least 30 GeV is required. For hadronic τ-leptons, the criteria $|\eta| < 2.3$ and $p_T > 20$ GeV have been used. The value of E_T^{miss} is investigated in terms of QCD multijet production (see section 4.4).

4.3.2 Muons and Electrons

Events are selected with the single lepton triggers for muons and electrons. Several (non-isolated) high threshold paths are included to recover efficiency for high p_T muons and electrons, which is important for the charged Higgs at high mass:

- Muon triggers: `HLT_IsoMu24_v`, `HLT_IsoTkMu24_v*`, `HLT_Mu50_v*`;
- Electron triggers: `HLT_Ele27_eta2p1_WPTight_Gsf_v*`, `Ele35_WPLoose_Gsf_v*`, `Ele105_CaloIdVT_GsfTrkIdT`, `Photon165_HE10`.

The efficiency recovered by adding such triggers in *or* combination is shown in figure **4-2**. The recovery of the single high energy egamma triggers is significant in the electron channel and motivates the use of them.

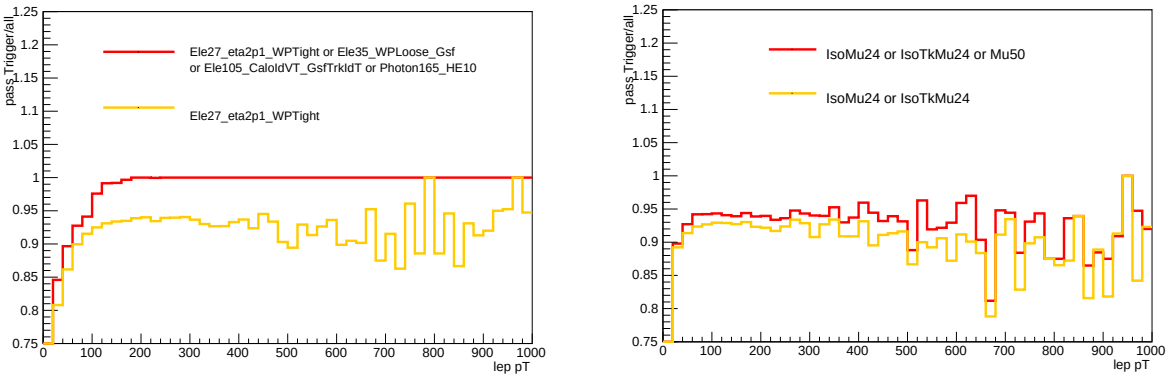


Figure **4-2**: The impact of the higher threshold triggers is shown for the single electron (left) and single muon (right) final state.

The muon (electron) is required to pass a *medium* (*tight*) working point and to have $p_T > 30$ GeV (for muon) and $p_T > 35$ GeV (for electron) in order to satisfy the trigger requirement. The electron is required to have also the $|\eta| < 2.1$ to satisfy the trigger requirement whereas the requirement on the muon is relaxed to have $|\eta| < 2.4$. The trailing lepton used for vetoing additional leptons is identified as *loose* muon(electron) and must exceed $p_T > 10$ GeV and have a coverage up to $|\eta| < 2.4$. The tight and loose working points are discussed in section 3.3.1. The lepton efficiencies are estimated in data and are factorized as follows:

$$\epsilon_\ell = \epsilon_{\text{RECO}} \times \epsilon_{\text{ID}} \times \epsilon_{\text{ISO}} \times \epsilon_{\text{trigger}}. \tag{4-1}$$

Those efficiencies are used to determine scale factors between data and simulation as function of p_T and η and are applied to each lepton used in this analysis. The scale factors

for the muon identification and reconstruction efficiency are available and applied [55, 56]; custom trigger and isolation scale factors shown were measured using the tag-and-probe technique [64].

4.4 Baseline Selection

Based on the different event signatures in both categories, their baseline selection is slightly different. This is mainly due to the fact that in the $\ell + \tau_h$ category the amount of jets is lower and the QCD fakes are largely suppressed due to the di-leptonic background. For the $\ell + \tau_h$ category, the baseline selection is as follows:

- one triggered electron or muon and veto on any additional leptons;
- one, two or three jets;
- at least one of the selected jets must be b-tagged;
- missing transverse energy larger than 70 GeV;
- angular cuts: $\Delta\phi(\ell, \mathrm{E_T^{miss}}) > 0.5$.

For the $\ell+$ no τ_h category, the baseline selection is summarized as follows:

- one triggered electron or muon and veto on any additional lepton;
- two or three jets;
- at least one of the selected jets must be b-tagged;
- missing transverse energy larger than 100 GeV;
- angular cuts: $\Delta\phi(\ell, \mathrm{E_T^{miss}}) > 0.5$, $\Delta\phi^{min}(\ell, \mathrm{jet}) > 0.5$ and $\Delta\phi(\mathrm{jet1}, \mathrm{E_T^{miss}}) > 0.5$.

The relative high thresholds on the missing transverse energy and the additional angular cuts are applied to suppress QCD such that it becomes quasi negligible (see Section 4.9.2). The constraints on the number of jets is motivated in the next section.

The cut flow for both channels and signal acceptance yields for each cut are shown in figure **4-3**. It is clear that the τ_h requirement in the $\ell + \tau_h$ category causes a drastic decrease of the event yields whereas this it not noticed in the $\ell+$ no τ_h category, which points to a possible inefficiency of the hadronic tau identification (a similar inefficiency is observed in the hadronic final state). Furthermore, the $\mathrm{E_T^{miss}}$ requirement in both categories drastically reduces the signal acceptance as well as the njet constraint in the $\ell+$ no τ_h category. The high $\mathrm{E_T^{miss}}$ cut can be removed when having a proper estimation of the QCD and fakes.

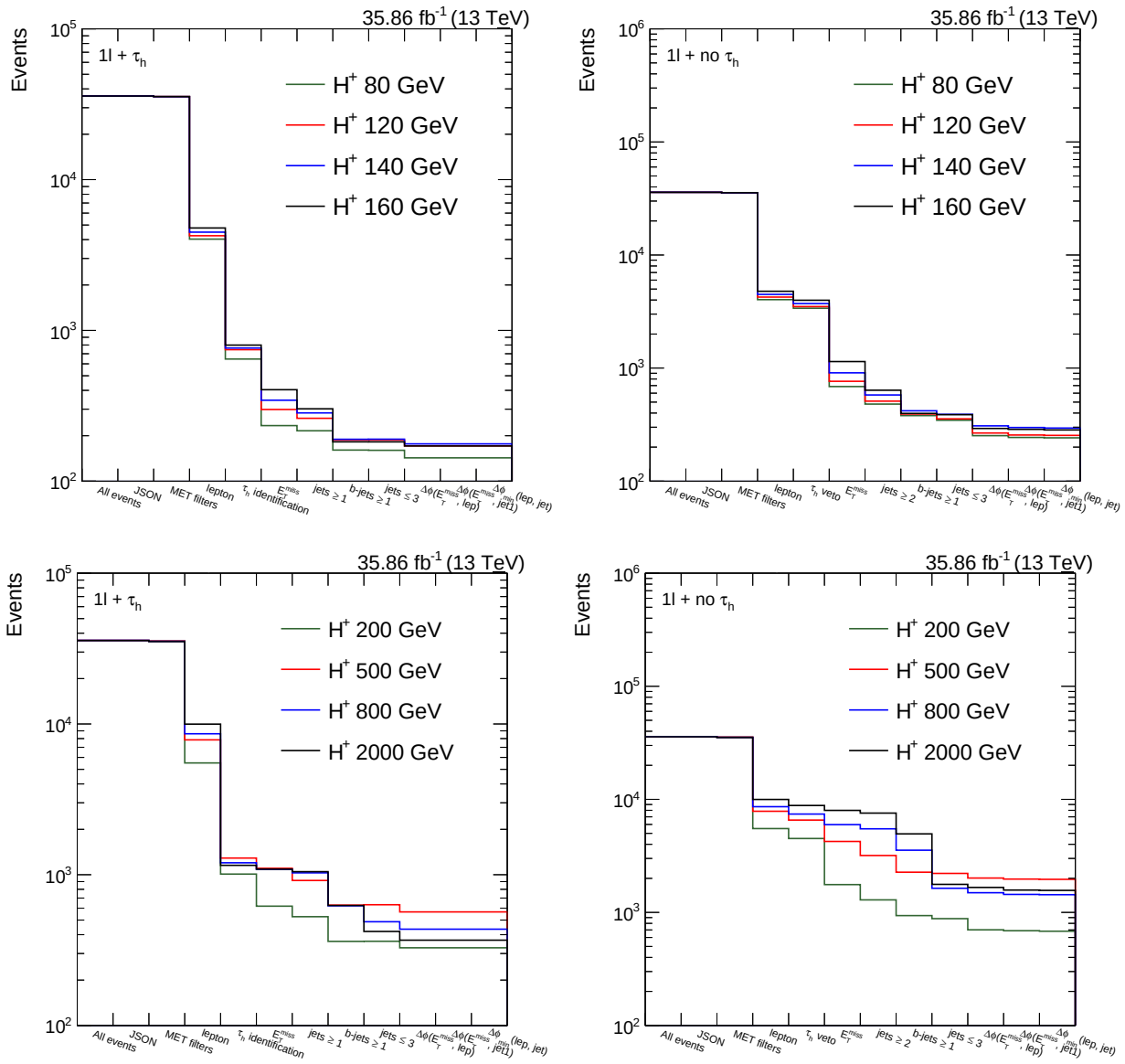


Figure **4-3**: Cut flow for the $\ell + \tau_h$ category (left) and $\ell+$ no τ_h category (right) and for the light mass regime (top) and heavy mass regime (bottom).

4.5 Orthogonality with $H^\pm \to t\bar{b}$ Leptonic Final State

The final state of the charged $H^\pm \to \tau^\pm\nu_\tau$ leptonic channel overlaps in large amounts with the charged Higgs to $H^\pm \to t\bar{b}$ leptonic channel. In order to perform a solid combination of these (and other) channels, one has to ensure there is no event overlap in the event phase space, i.e. both phase spaces need to be orthogonal. As the $H^\pm \to t\bar{b}$ leptonic channel is dominant over the $H^\pm \to \tau^\pm\nu_\tau$ leptonic channel in terms of charged Higgs production branching ratio, priority has been given to the former analysis to fully exploit and optimize the event phase space. Remaining orthogonal regions, characterized with a lower jet multiplicity, are devoted to the phase space of the $H^\pm \to \tau^\pm\nu_\tau$ leptonic final state.

Because the $H^\pm \to t\bar{b}$ single leptonic analysis applies a veto on hadronic taus using the same loose working point, any overlap is only possible in the $\ell+$ no τ_h category. The event phase space of both analyses is shown in figure **4-4** (left) for the single leptonic final states. Shown in gray are the regions used in the $H^\pm \to t\bar{b}$ final state having at least 4 jets in the selection. In orange are the orthogonal regions used in this analysis which are limited to the use of maximum three jets. This constraint on the number of jets affects the heavy mass acceptance (especially above charged Higgs masses above 500 GeV) as the number of extra jets increases with the mass of the charged Higgs due to extra radiation effects. However, any re-interpretation of the charged Higgs to $H^\pm \to \tau^\pm\nu_\tau$ in the $H^\pm \to t\bar{b}$ phase space is considered in the final combination of all the channels.

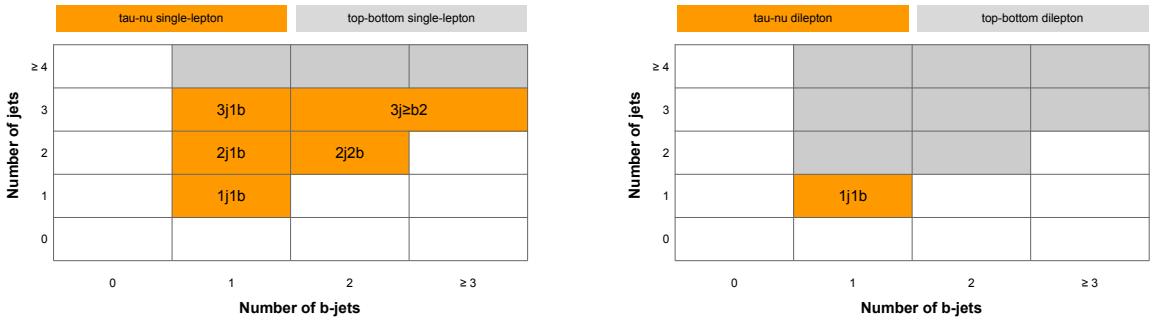


Figure **4-4**: Cartoon of the regions used in this analysis: single lepton final state (left) and di-lepton final state (right). The gray boxes are used by the $H^\pm \to t\bar{b}$ leptonic analysis and are therefore not accessible. The regions in orange in the single lepton final state are used in this analysis.

For the light and intermediate mass regime as well as for the $\ell + \tau_h$ category, no such constraints are necessary as there cannot be any event overlap. However, these regimes are characterized by a lower jet multiplicity and therefore the number of jets is still kept to a maximum of three to avoid the large bulk of $t\bar{t}$ (characterized with number of jets around 4).

Besides the single lepton final state, the di-lepton state can be considered as well where the associated top decays leptonically as well as the tau from the charged Higgs. W.r.t. the single lepton final state, this event topology has two jets less and an additional neutrino from the top decay. The orthogonality constraint for the di-lepton final state results in a phase space with exactly one jet which is required to be b-tagged (i.e. mono b), as depicted in figure **4-4** (right). Taken into account the b-tagging efficiency, the low branching ratio of the decay of the associated top to leptons and the suppression of any additional jet activity (e.g. gluon radiation especially for high mass charged Higgs), this final state is rather insensitive and therefore not taken into account.

4.6 Signal Extraction and Control Regions

The triggered lepton in the events originate from the tau decay in the $\ell+$ no τ_h category and from the associated top decay in the $\ell + \tau_h$ category. In the former category, the missing transverse energy is confined as decay products on the charged Higgs and together with the lepton the invariant mass can be reconstructed:

$$\mathrm{m_T} = \sqrt{2\mathrm{p_T}(\ell)\mathrm{E_T^{miss}}(1 - \cos\Delta\phi(\mathbf{p_T}(\ell), \mathbf{p_T^{miss}}))}. \tag{4-2}$$

In this equation, $\mathrm{p_T}(\ell)$ is the momentum of the triggered lepton, $\mathrm{E_T^{miss}}$ is the missing transverse energy and $\Delta\phi$ the azimuthal angle between them. Because the presence of 3 neutrinos and detector resolution effects, the invariant mass spectrum is broadened a bit rather than a sharp cut off at the end point.

Concerning the $\ell+\tau_h$ category, the missing transverse energy is a mixture of the neutrinos from the charged Higgs plus the W neutrinos from the associated top decay. Nevertheless the invariant transverse mass can be constructed but it is expected the distribution is broadened.

This transverse mass distribution for several charged Higgs masses is shown in figure **4-5**, after application of all the selection criteria. A typical shape is observed for the background which peaks around 80 GeV, representing the $\mathrm{W^{\pm}}$ boson which can either come from top decay or direct $\mathrm{W^{\pm}}$ boson production. On the other hand, the signal does not exhibit this peak and is more flat. This distinct difference between signal and background makes the transverse mass a good distribution to extract the signal.

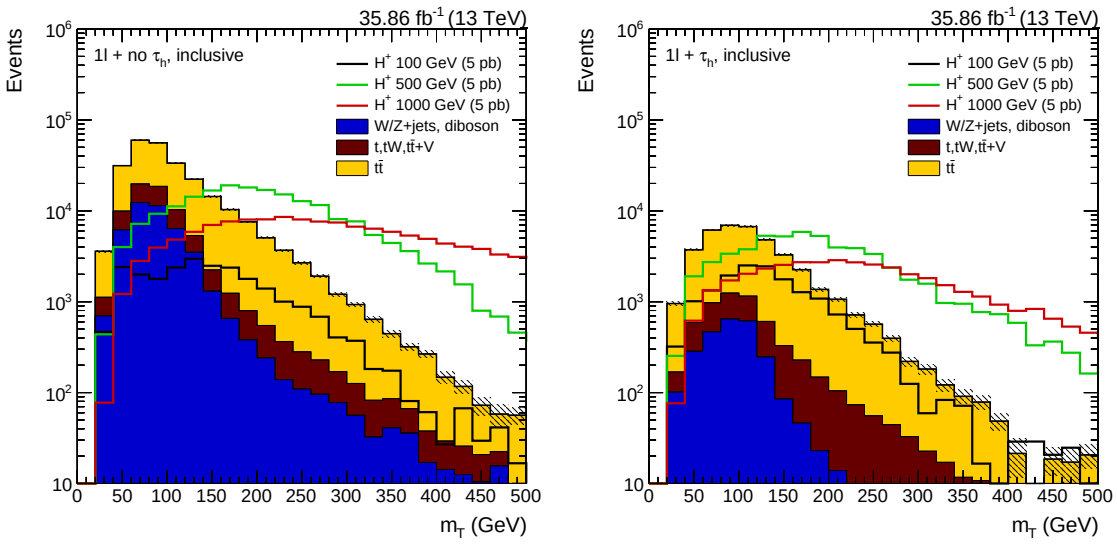


Figure **4-5**: m_T distribution for several mass hypotheses and the corresponding backgrounds. Left: $\ell+$ no τ_h category, right: $\ell + \tau_h$ category.

Because the background in the low m_T regime is several orders of magnitude larger than a possible signal (with cross section of order 1 pb or less), the low m_T regime is used as control region for unblinding to check the data/Monte Carlo agreement. A value of $m_T = 120$ GeV is chosen as cut-off between control and signal region. The final fit uses both control and signal regions by performing a simultaneous fit.

4.7 Signal Acceptance

The signal acceptance is defined as the absolute event yields as function of the charged Higgs mass, and is shown after several cuts in figure **4-6** for the $\ell + \tau_h$ category (left) and $\ell+$ no τ_h category (right). The entire mass range is considered from 80 GeV to 3 TeV except for the intermediate mass range. For low charged Higgs masses up to 500 GeV, a quasi monotonic increase of acceptance is observed besides a lower acceptance towards 160 GeV. This effect is due to the fact that for light mass samples close to the top mass, the additional b-quark is very soft and got out of acceptance.

For higher masses above 500 GeV, a saturation effect is noticed in both $\ell + \tau_h$ and $\ell+$ no τ_h categories which is due to the constraint of the jet multiplicity. For the $\ell + \tau_h$ category, the drop in yields is more pronounced because the tau identification tend to saturate for charged Higgs masses above 500 GeV (see plots in appendix **4-6**). This is explained by the fact that the τ-lepton identification becomes less efficient for high p_T τ-leptons.

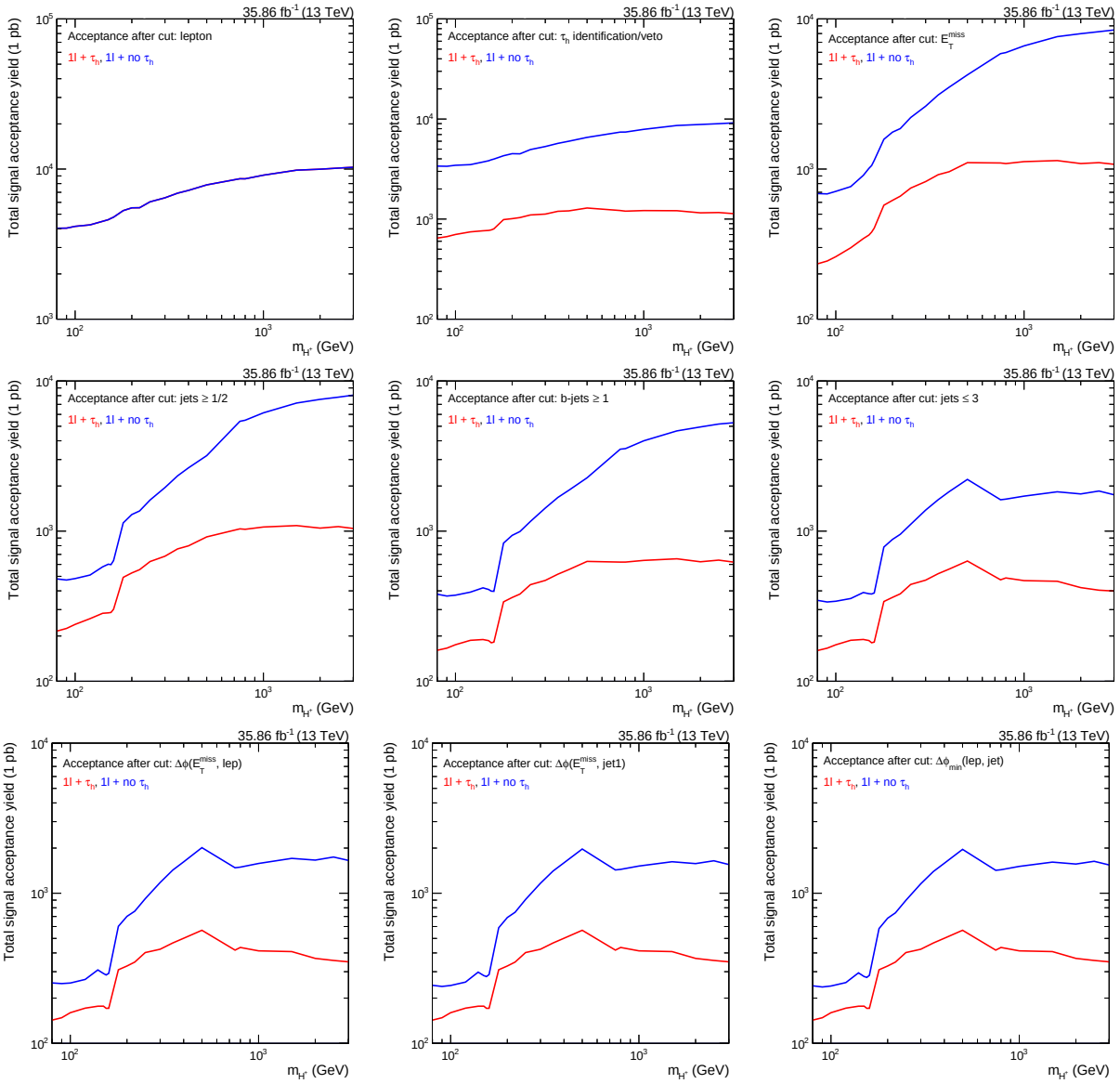


Figure **4-6**: Signal acceptance after each cut (according to the cut flow), where the signal cross section is normalized to 1 pb. It is clear that the constraint in number of jets and the tau identification are responsible for the drop in yields at higher charged Higgs masses.

4.8 Analysis Categorization

The loose event phase space is divided in several disjoint categories. Categories enriched in signal are sensitive for the statistical component whereas background abundant regions are used to constrain the systematics.

4.8.1 Categorization in Jet and b-Jet Multiplicity

The event phase space is split in regions based on the number of jets and b-jets (see figure **4-4**). Due to the different event kinematics in these regions and different background constraints upon performing the fit, the expected upper limit is more sensitive w.r.t. a full inclusive fit. The effect on the upper limit when applying the categorization is shown in figure **4-7**. Especially for the low mass range where the systematics are dominant, a gain up to 50% is obtained.

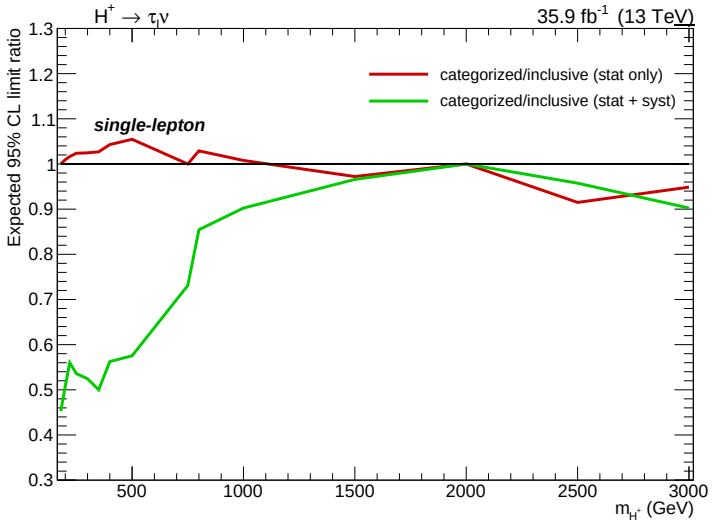


Figure **4-7**: Comparison of upper limits when splitting the m_T templates into the njet/nbjet categories (see text), for both statistical only and statistical + systematics.

Shown in figure **4-8** is the significance signal/background yields ($S/\sqrt{B}$) for several charged Higgs masses and for each of the jet and b-jet multiplicities (inclusive over $\ell+\tau_h$ and $\ell+$ no τ_h categories). For low mass charged Higgs hypotheses, the significance is confined to the allowed jet multiplicities whereas for higher mass charged Higgs hypotheses the significance is pushed towards higher jet multiplicities. This is due to the fact that the objects in the event have a high energy and therefore the probability of extra radiation increases leading to more jets in the event.

Also the 2b and 3b categories are equally significant to the 1b category, mostly due to the stronger suppression of the backgrounds. Generator level studies shows that in the 1b category indeed the b-tagged jet originates from top decays, the extra b-jet in the 2b category originates from the associated production (5FS) and the third b-tagged jet in the 3b category is mis-tagged.

For the $\ell+$ no τ_h category, 4 regions have been chosen: $2j, 1b$, $2j, 2b$, $3j, 1b$ and $3j, \geq 2b$. For the $\ell + \tau_h$ category, which has naturally two jets less, the mono-b region is added ($1j1b$) and the 2-3 jet categories are merged due to low signal statistics in the 3jet regions. As a result, the $\ell + \tau_h$ category uses the following three regions: $1j, 1b$, $2-3j, 1b$ and $2-3j, 2b$.

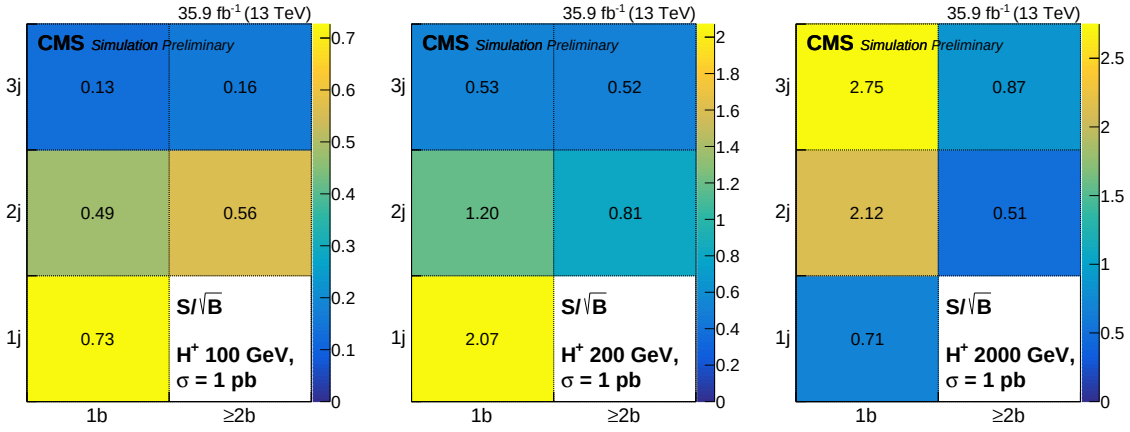


Figure **4-8**: Significance $S/\sqrt{B}$ for each of the jet categories for a charged Higgs mass of 200 GeV (left) and 2 TeV (right). $S(B)$ represents the total signal(background) yield in that category; the signal is normalized to 1 pb [5].

4.8.2 Categorization in E_T^{miss}

Based on the categorization in the previous subsection, a further categorization is elaborated in terms of missing transverse energy (E_T^{miss}). The events are split in different E_T^{miss} categories to further constrain systematics and be sensitive to different event kinematics in those categories. The effect on the upper limit is shown in figure **4-9**, where the ratio of the upper limits (based on the m_T discriminant) is calculated in two MET categories (for each region R2 to R5) ranging from 100 to 150 GeV and MET $>$ 150 GeV for the $\ell+$ no τ_h category. From the plots it is clear the sensitivity at charged Higgs mass goes up to 20% for the statistical only ratio and to 40% when taking into account systematics, as expected as the systematics are more dominant for lower charged Higgs masses. This clearly shows the effect of different event kinematics and more power constraint when the categorization is present.

The categorization in E_T^{miss} is chosen based on the statistics of the signal samples:

- $\ell + \tau_h$ is split three E_T^{miss} categories:
  1. $70 < E_T^{miss} < 100$ GeV
  2. $100 < E_T^{miss} < 150$ GeV
  3. $E_T^{miss} > 150$ GeV
- ℓ+ no τ_h is split two E_T^{miss} categories:
  1. $100 < E_T^{miss} < 150$ GeV
  2. $E_T^{miss} > 150$ GeV

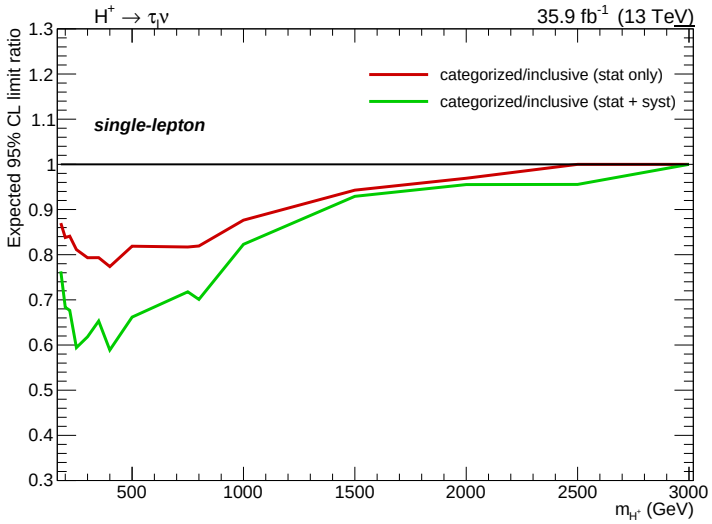


Figure **4-9**: Comparison of upper limits when splitting the m_T templates into two MET categories (see text), for both statistical only and statistical + systematics.

4.8.3 Summary of Event Categorization

Based on the optimizations examined in this section, the following configuration has been chosen for optimal signal sensitivity:

- $\ell + \tau_h$: simultaneous fit of the m_T templates in both electron and muon final states, categorized in number of jets (1 and ≥ 2), number of b-tagged jets (1 and ≥ 2) and 3 MET categories ($70 < E_T^{miss} < 100$, $100 < E_T^{miss} < 150$ and $E_T^{miss} > 150$). This results in 18 templates;

- ℓ+ no τ_h: simultaneous fit of the m_T templates in both electron and muon final states, categorized in number of jets (2 and 3), number of b-tagged jets (1 and ≥ 2) and 2 MET categories ($100 < E_T^{miss} < 150$ and $E_T^{miss} > 150$). This results in 16 templates.

Hence, in total, the search is performed in 34 bins of m_T templates.

4.9 Background Estimation

The dominant backgrounds are $t\bar{t}$ and $W^{\pm}$+jets for the ℓ+ no τ_h category. All the backgrounds, as well as the minor backgrounds (Z/γ^*+jets, single top, $t\bar{t}+X$ and diboson) are directly estimated from simulation. Backgrounds are modeled using simulated samples corrected to account for known theoretical and experimental deficiencies (see section 4.10). For a complete list of background samples were summarized in section 3.3.3. The QCD samples are used to estimate the overall QCD/fakes contamination (see section 4.9.2), but are not propagated as templates to the fit.

In particular, for $t\bar{t}$ the tt+jets sample is used which uses the MG5_aMC@NLO generator. It is found that this sample predicts more accurately the jet kinematics for which the ℓ+ no τ_h category is sensitive due to the constraint on the jet multiplicity. Due to the limited b hadron content in the event phase space, the $t\bar{t}$ is not split in heavy flavor components as was done for the $H^{\pm} \to t\bar{b}$ leptonic search.

4.9.1 Control Plots

As already explained in section 4.6, the control and signal regions are distinguished by the value of m_T: for $m_T < 120$ GeV the backgrounds are abundant, henceforth defined as control regions, whereas for $m_T > 120$ GeV the backgrounds are suppressed and therefore labeled as signal regions.

Such control regions were used in the blind analysis to validate the data/Monte Carlo agreement in so called control plots. During unblinding phase, the control plots are extended to the signal regions as well to validate the entire event phase space with data and to perform the final fit.

Shown in Figs. **4-11**, **4-11**, **4-12** and , **4-13**, the basic event kinematic variables (m_T, E_T^{miss}, H_T (i.e. sum of the jet p_T) and the lepton p_T) in the extended control regions, inclusive in the entire event selection (i.e. the E_T^{miss} and jet categories are merged). The electroweak processes ($W^{\pm}$+jets, Z/γ^*+jets and diboson) are all merged into one electroweak category. To validate the event kinematics in the different jet regions, more control plots were checked emphasized on jet and lepton kinematics.

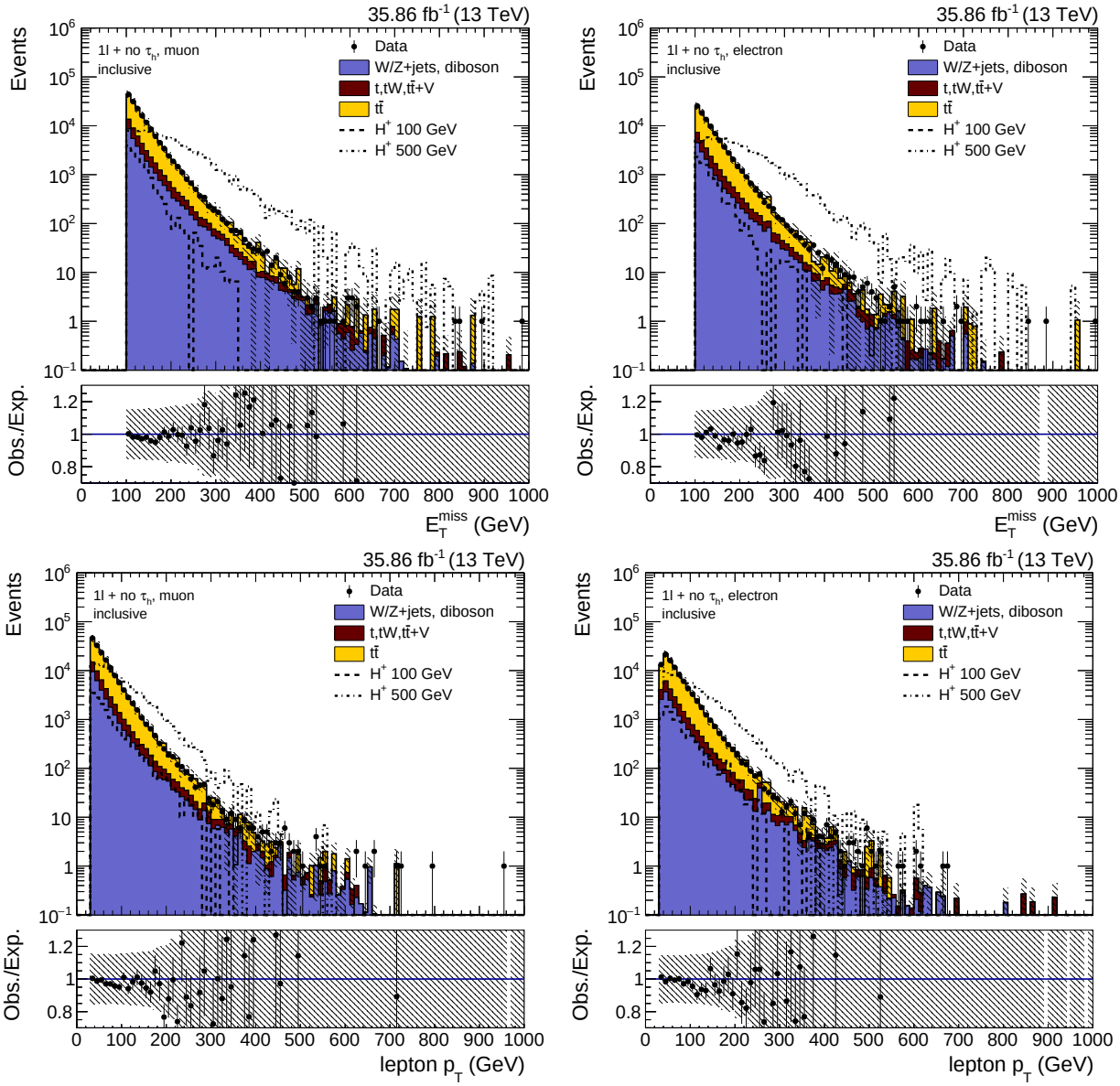


Figure **4-10**: Inclusive pre-fit control plots for 1Mu (left) and 1El (right) for the ℓ+ no τ_h category: E_T^{miss} (top) and lepton p_T (bottom). The error band includes both statistical and systematic components.

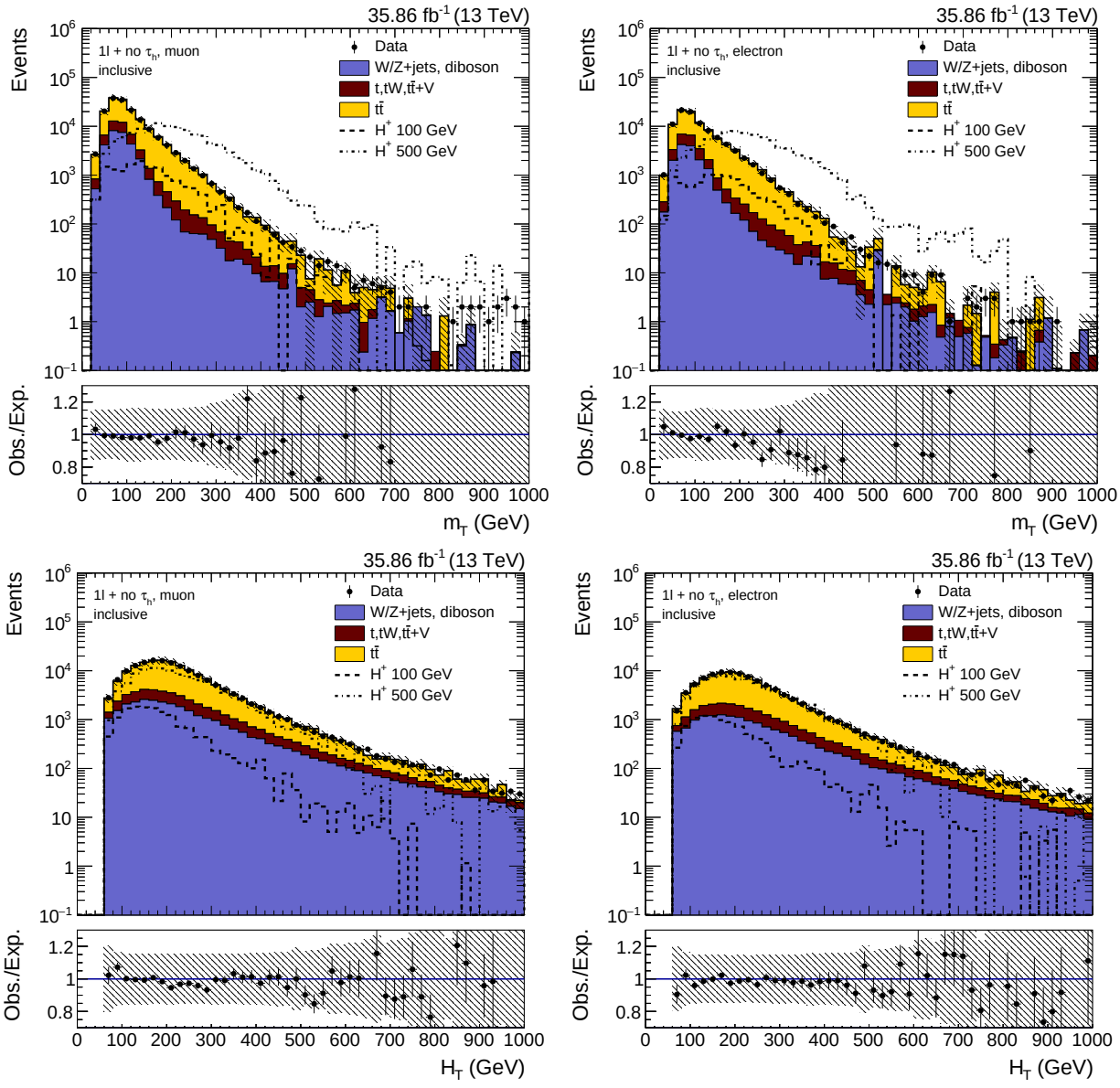


Figure **4-11**: Inclusive pre-fit control plots for 1Mu (left) and 1El (right) for the ℓ+ no τ_h category: m_T (top) and H_T (bottom). The error band includes both statistical and systematic components.

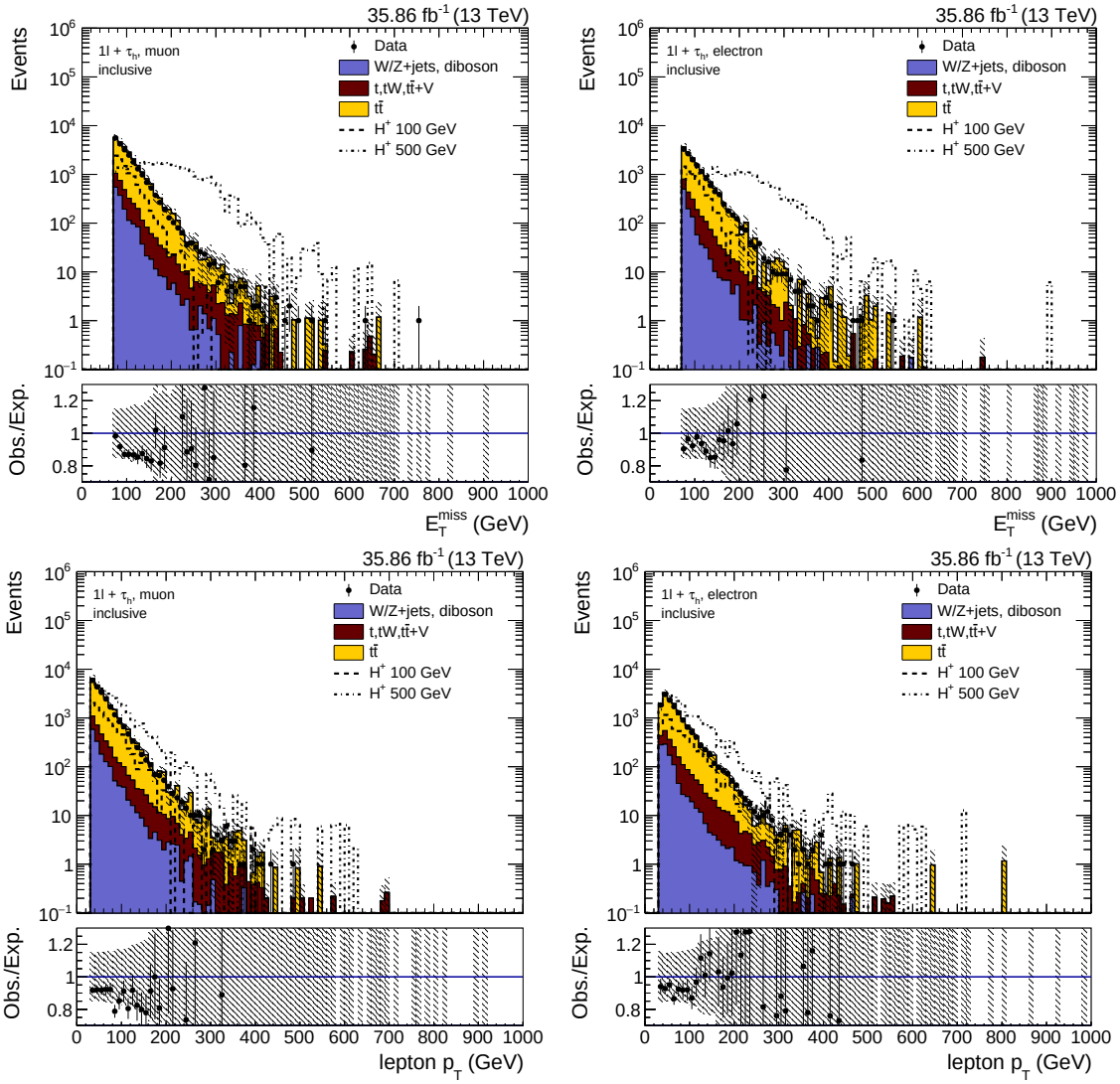


Figure **4-12**: Inclusive pre-fit control plots for 1Mu (left) and 1El (right) for the $\ell + \tau_h$ category: E_T^{miss} (top) and lepton p_T (bottom). The error band includes both statistical and systematic components.

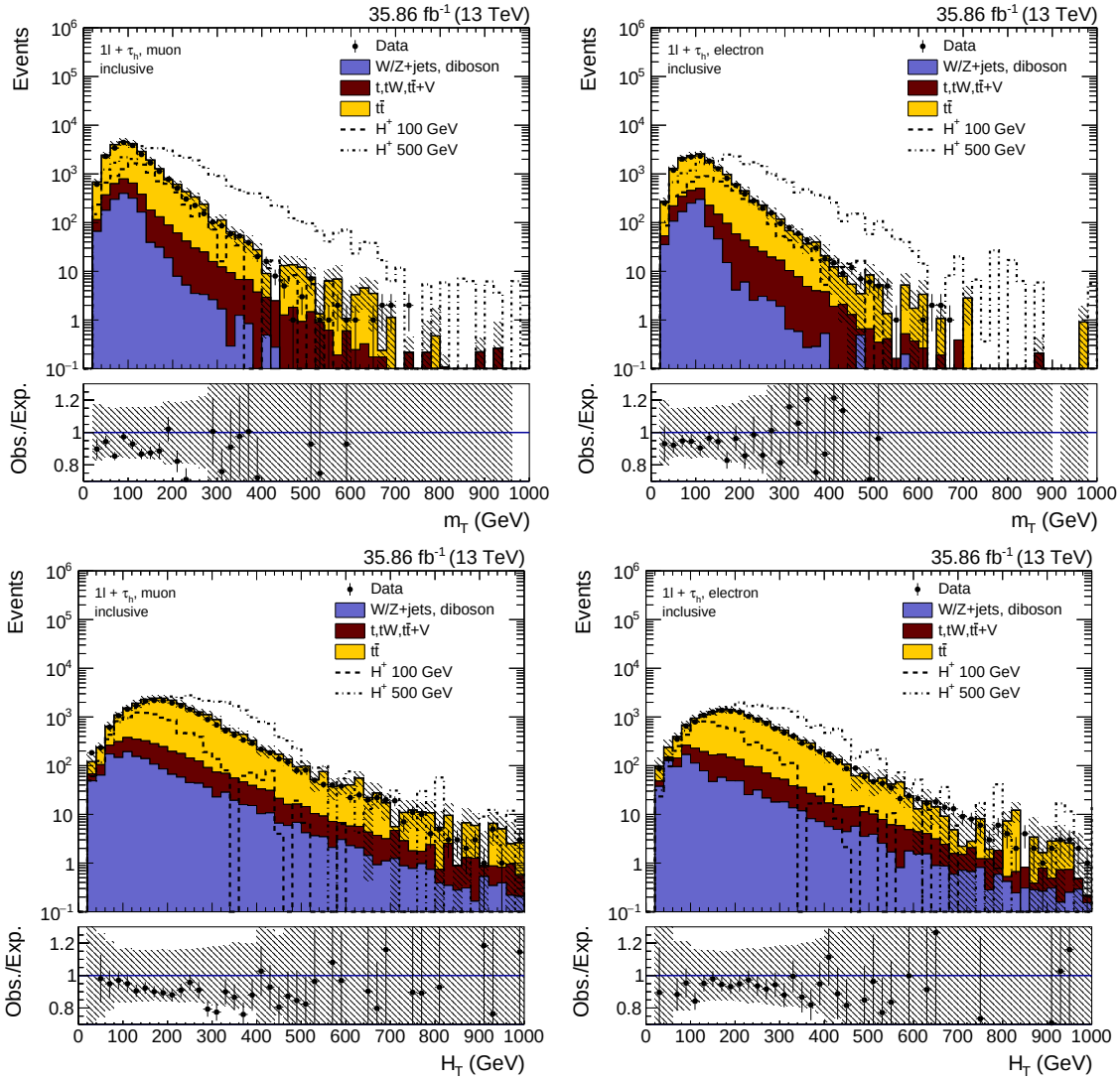


Figure **4-13**: Inclusive pre-fit control plots for 1Mu (left) and 1El (right) for the $\ell + \tau_h$ category: m_T (top) and H_T (bottom). The error band includes both statistical and systematic components.

4.9.2 QCD Estimation and Fakes

QCD multi-jet production can contaminate the event phase space through jet faking leptons and jet mis-measurements. Such events are more abundant in regions with low jet multiplicities and with relatively energetic jets which can lead to fake leptons which enter the event selection. Specifically for the muon final state, fakes can arise from leptonic b-hadron decays which are non-prompt, therefore labeled as fakes. Despite the tight mini-isolation requirement for the leptons (mini-iso < 0.1), it is found that fakes are sensitive in this analysis, especially in the high energy tails of the distributions. From the design point of view of the analysis, there are two ways to deal with QCD multi-jet production:

- estimation from data: by using data-driven methods (fake-rate method) one can estimate the QCD multi-jet production in the event phase space by inverting the isolation cuts and propagate scale factors in this inverted region to the event phase space;
- suppress QCD multi-jet production to a negligible amount: apply cuts on the event phase space such that the QCD multi-jet production becomes negligible.

In this analysis the second approach has been chosen, whereas for the hadronic analysis a data-driven method has been used to estimate the fake-tau rate. In the $\ell+$ no τ_h category, QCD multi-jets can enter the event selection by a jet faking a lepton or the leptonic b-hadron decay and especially in this category the QCD multi-jet production and fakes needs to be investigated. For the $\ell + \tau_h$ category, QCD multi-jets can enter the event selection by faking a lepton and the hadronic tau, which is less probable.

An effective way to kill QCD multi-jet production is to cut on the missing transverse energy as it is expected for multi-jet events the $\mathrm{E_T^{miss}}$ comes from jet mis-measurements and is typically low. Therefore a $\mathrm{E_T^{miss}}$ cut of 100 GeV is applied in the $\ell+$ no τ_h category. For the $\ell + \tau_h$ category the $\mathrm{E_T^{miss}}$ cut is lowered to 70 GeV, chosen such that the tau fakes is also kept to a minimum.

Angular Cuts

Several $\Delta\phi$ distributions have been checked to look for any pileup of QCD events in the $\ell+$ no τ_h category. Because the fakes can also be present in the high lepton $\mathrm{p_T}$ spectrum which makes part of the signal regions, the distributions are fully unblinded but kept inclusively for the entire selection. To test whether the any pile-up of events is originating from QCD multi-jet events, QCD samples (see section 3.3.3) have been used for comparison to check the data/Monte Carlo agreement.

Several angular distributions among the lepton, $\mathrm{E_T^{miss}}$ and jets have been cross checked in data and Monte Carlo. From this analysis three angular distributions have been chosen to apply additional cuts which are listed below. The inclusive angular distributions are shown in Fig. **4-14** for the $\ell+$ no τ_h category, where on the left column the distributions are shown without QCD samples and on the right column with QCD samples for comparison.

- $\Delta\phi(\text{lepton}, \mathrm{E_T^{miss}})$ (top row): especially in the muon channel a discrepancy is observed at low values of $\Delta\phi$, which originates from muonic b-hadron decays. A cut of $\Delta\phi(\text{lepton}, \mathrm{E_T^{miss}}) > 0.5$ is applied (both in electron and muon channels). This cut is also applied in the $\ell + \tau_h$ category.

- $\Delta\phi(\text{leading jet}, \mathrm{E_T^{miss}})$ (middle row): in both electron and muon channel fakes pile-up at low values of $\Delta\phi$ especially in the signal regions with high values of m_T. A cut of $\Delta\phi(\text{leading jet}, \mathrm{E_T^{miss}}) > 0.5$ is applied (both in electron and muon channels).

- $\Delta\phi_{\text{min}}(\text{lepton}, \text{jet})$ (bottom row): in both electron and muon channel at high values of $\Delta\phi$ a discrepancy is observed. Because this is in the lepton-jet back-to-back regime, this accounts for jet mis-measurements. A cut of $\Delta\phi_{\text{min}}(\text{lepton}, \text{jet}) < \pi - 0.5$ is applied (both in electron and muon channels).

Summary

To summarize, the following additional angular cuts have been applied in the $\ell+$ no τ_h category:

- $\Delta\phi(\text{lepton}, \mathrm{E_T^{miss}}) > 0.5$;
- $\Delta\phi(\text{leading jet}, \mathrm{E_T^{miss}}) > 0.5$;
- $\Delta\phi_{\text{min}}(\text{lepton}, \text{jet}) < \pi - 0.5$.

For the $\ell + \tau_h$ category, the following angular cut has been applied:

- $\Delta\phi(\text{lepton}, \mathrm{E_T^{miss}}) > 0.5$.

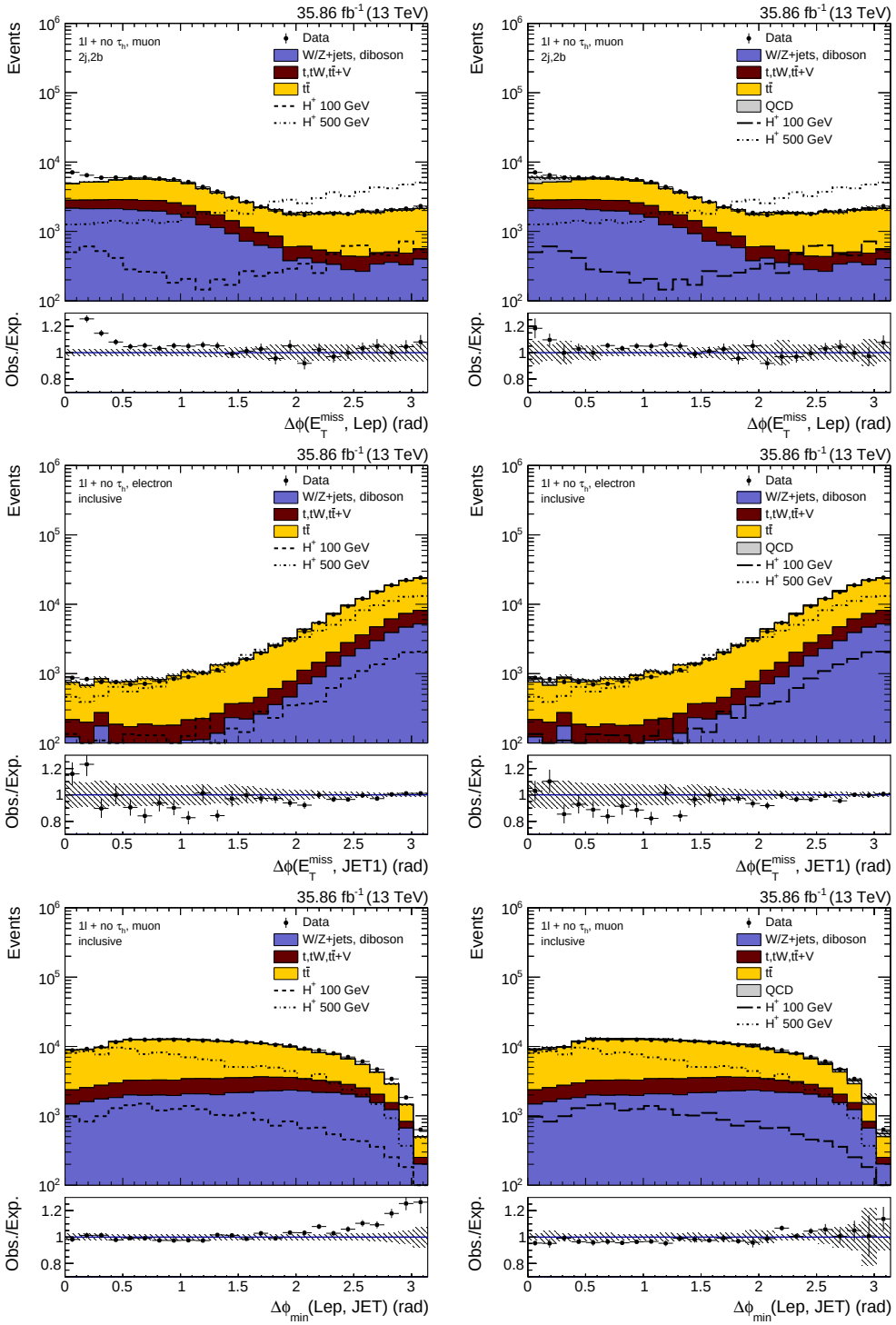


Figure **4-14**: Angular distributions $\Delta\phi(\text{lepton}, \text{E}_\text{T}^\text{miss})$ (top row), $\Delta\phi(\text{leading jet}, \text{E}_\text{T}^\text{miss})$ (middle row) and $\Delta\phi_\text{min}(\text{lepton}, \text{jet})$ fully inclusive for the ℓ+ no τ_h category. On the left the distributions are shown without QCD samples whereas on the right the QCD samples are added for comparison. Error bars account for statistical uncertainty only.

4.10 Systematic Uncertainties

The primary physics result described in this note is a 95% CL upper limit on the cross section of the production of a charged Higgs boson as function of its mass. Several sources of systematic uncertainty must be taken into account when building the test statistic used to test the background-only hypothesis versus the presence of a charged Higgs boson (see section 3.4). Such sources are modeled in the test statistic as *nuisance parameters*, with the effect of either changing the event rate of the considered signal and background processes (*rate uncertainties*) or changing the shape of the transverse mass template distributions (*shape uncertainties*). Depending on their type, experimental and theoretical uncertainties are propagated as shape or normalization uncertainties to the m_T and other signal-extraction distributions.

The simulation, used for the signal and background modeling, is affected by two different sources of systematic uncertainty. The first source of uncertainty is correction factors applied to the simulation in order to better reproduce the detector conditions and performance in data. The second source is assumptions made in the theoretical models that were used to produce the simulation. We account for uncertainties from both sources.

All the sources of systematic uncertainty applied to the analysis and their correlations are discussed below. Table **4-2** summarizes the systematic uncertainties assessed on the signal and backgrounds for this analysis. An estimation of the pre-fit size, together with the hadronic final state, is summarized in table **5-1**.

Luminosity and Pileup An overall uncertainty of 2.5%, due to the measurement of the 2016 data luminosity, is applied [97]. The pileup uncertainty is evaluated by changing the minimum bias cross section with ±5%, which is propagated to the an event weight [98]. The difference between the nominal and the altered distributions is taken as the uncertainty. The changes in the weight factor are propagated to the final discriminant distributions and treated as rate+shape variation in the final fit. Each of these uncertainties is treated as fully correlated among all processes.

Jet A scale factor is applied to each reconstructed jet, to correct its energy scale (JES). The uncertainties associated to the scale factor are applied to obtain its upwards and downwards variation that are then applied to the jet transverse momentum and balanced by the missing transverse energy [61]. The uncertainty factors depend on the jet $\mathrm{p_T}$ and η. Upwards and downwards shape uncertainties on the jet energy resolution (JER) are applied and also propagated to the missing transverse energy. The events are then re-analyzed, including reapplying the jet-based selection and recalculating all kinematic quantities in order to extract appropriate variations for the final yields and

template distributions.

Missing Transverse Energy The uncertainties on the unclustered E_T^{miss}, defined as the total missing transverse momentum minus the non-varied jets p_T, are propagated to the different distributions. The JES are used to correct the MET.

b-jet scale factors The b-tag and mistag uncertainties are obtained by varying the corresponding per-jet scale factors up and downwards [66]. Light jet (u,d,s) mistag scale factors are considered to be uncorrelated with the b tag scale factor, whereas the c mistag scale factor is correlated and therefore should be varied simultaneously with the b tag SF. Both the b tag/c mistag and light jet mistag uncertainties are then treated as uncorrelated. The b tag and mistag uncertainties are doubled whenever they occur outside the p_T,η-range of the scale factors (validity for light jets: $20 < p_T < 1000$ GeV and $|\eta| < 2.4$, for b/c-jets: $30 < p_T < 670$ and $|\eta| < 2.4$,). The impact of statistical uncertainties during the scale factor determination, is propagated to an alternative set of scale factors.

The b tag uncertainty associated with the JES is evaluated within the overall JES uncertainty as the b tag scale factor values are also shifted accordingly when the jet kinematics is shifted up or down.

Lepton selection and trigger Electron/Muon ID + scale factors are applied to the analysis (see appendix **??**). The uncertainties of the trigger efficiency scale factors are varied up and down as function of the lepton transverse momentum. The overall impact of is of 4% for the muon and 3% for the electron.

Tau ID efficiency and veto The agreement of the τ_h identification efficiency between data and simulated samples is measured using the tag-and-probe technique [64]. A 5% tau efficiency uncertainty is applied on genuine taus in the selection for the chosen loose working point. Convoluted with the tau efficiency in the $\ell + \tau_h$ category and deficiency in the ℓ+ no τ_h category, the uncertainties are estimated as 3.1% and 0.7% respectively. Both uncertainties are anti-correlated in the fit.

MC normalization Theory uncertainties consisting of parton density function (PDF) and scale uncertainties are applied as shape and normalization uncertainties to the MC samples. The QCD scale and PDF normalization uncertainties cover for the uncertainty of the cross section and are applied for each process individually (except signal).

The uncertainty in the choice of renormalization and factorization scale is assessed by separately varying the renormalization and factorization scales by a factor of 2.0 and

Source	**Uncertainty**	**Name in datacard**
Luminosity	2.5%	lumi_13TeV
Pileup	shape	CMS_pileup
MC statistics	shape	autoMCStats 0
Electron efficiency	3%	CMS_eff_e
Muon efficiency	4%	CMS_eff_m
Tau efficiency ($\ell + \tau_h$)	3.1%	CMS_eff_t
Tau veto (ℓ+ no τ_h)	-0.7%	CMS_eff_t
MET unclustered	shape	CMS_scale_uncluster
Jet Energy Scale (JES)	shape	CMS_scale_j
Jet Energy Resolution (JER)	shape	CMS_res_j
Parton Shower ISR	$\sim +4/\sim -8\%$	CMS_Hptn_ISR
Parton Shower FSR	$\sim \pm 20\%$	CMS_Hptn_FSR
Parton Shower hDAMP	$\sim +8/\sim -2\%$	CMS_Hptn_HDAMP
Parton Shower UE	$\sim \pm 2\%$	CMS_Hptn_UE
b-tagging - HF	shape	CMS_eff_b_HF
b-tagging - HFStat1	shape	CMS_eff_b_HFstat1
b-tagging - HFStat2	shape	CMS_eff_b_HFstat2
b-tagging - LF	shape	CMS_eff_b_LF
b-tagging - LFStat1	shape	CMS_eff_b_LFstat1
b-tagging - LFStat2	shape	CMS_eff_b_LFstat2
b-tagging - CFerr1	shape	CMS_eff_b_CFerr1
b-tagging - CFerr1	shape	CMS_eff_b_CFerr2
QCD scale (acc.) signal	$\pm 4.8\%$ or $\pm 1.2\%$	CMS_Hptn_mu_RF_Hptn
QCD scale (acc.) $t\bar{t}$, single top	$\pm 2\%$	CMS_Hptn_mu_RF_top
QCD scale (acc.) electroweak	$\pm 5\%$	CMS_Hptn_mu_RF_ewk
QCD scale (xsec) $t\bar{t}$	$+2.4/-3.5\%$	QCDscale_ttbar
QCD scale (xsec) single top	$\pm 2.5\%$	QCDscale_singletop
QCD scale (xsec) $W^{\pm}$+jets	$+0.8/-0.4\%$	QCDscale_wjets
QCD scale (xsec) diboson	$\pm 3.2\%$	QCDscale_vv
QCD scale (xsec) Z/γ^*+jets	$0.7/-0.4\%$	QCDscale_dy
PDF (acc.) signal	$+1.7/-0.4\%$	CMS_Hptn_pdf_Hptn
PDF (acc.) $t\bar{t}$, single top	$+0.27/-2\%$	CMS_Hptn_pdf_top
PDF (acc.) electroweak	$+4.6/-3.3\%$	CMS_Hptn_pdf_ewk
PDF (xsec) $t\bar{t}$	$\pm 4.2\%$	CMS_pdf_ttbar
PDF (xsec) single top	$\pm 4.7\%$	CMS_pdf_singletop
PDF (xsec) $W^{\pm}$+jets	$\pm 3.8\%$	CMS_pdf_wjets
PDF (xsec) diboson	$\pm 4.4\%$	CMS_pdf_vv
PDF (xsec) Z/γ^*+jets	$\pm 3.7\%$	CMS_pdf_dy
Mass top	2.7% ($t\bar{t}$), 2.5% (single top)	CMS_mass_top

Table 4-2: Summary of systematics.

0.5 [99]. The unphysical anti-correlated variations are discarded, yielding a total of 7 combinations of the renormalization and factorization scales. The event weight is determined for each variation and the envelope is taken as the scale uncertainty. These variations are applied using the event weights provided in the LHE eventInfo. This exercise was performed in the hadronic channel and a set of conservative normalization uncertainties were derived for signal, $t\bar{t}$, single top, $W^{\pm}$+jets, Z/γ^*+jets and diboson. Those uncertainties are also used in the hadronic channel to the corresponding processes.

To model the PDF uncertainty, 100 MC replicas are drawn from the PDF probability distribution [99]. The standard deviation of these weights is taken as the PDF uncertainty. This exercise was performed in the hadronic channel and a set of conservative normalization uncertainties were derived for signal, $t\bar{t}$, single top, $W^{\pm}$+jets, Z/γ^*+jets and diboson. Those uncertainties are also used in the hadronic channel to the corresponding processes.

For the MC samples involving a top quark, and additional uncertainty from the top mass is added. Uncertainties assigned to the signal samples are not applied as the normalization does not affect the upper limit calculation. All these uncertainties are considered to be correlated.

Uncertainty due to the size of the MC samples The limited size of the background and signal MC samples results in statistical fluctuations of the nominal prediction. The content of the bin of each final discriminant distribution is varied by its statistical uncertainty (except for the signal). A Barlow-Beeston lite approach has been implemented for automatic treatment of the statistical uncertainties (see section 3.4). A threshold of 0 has been applied to take into account all the statistical uncertainties. As this configuration introduces a large amount of uncorrelated nuisances, the analytic minimizer is used to run the asymptotic limits in a rather fast way.

$t\bar{t}$ parton shower Additional uncertainties based on the jet multiplicity are taken into account, which arises from the uncertainty at generator level on the additional jet production. Four parton showering variations are included by perturbing the initial- and final-state parameters [100], the matching of jets from matrix element calculations and from parton shower, and the underlying event tune [78]. This uncertainty is estimated in each category (jet and b-jet multiplicity, MET), based on Monte-Carlo samples which are enriched/reduced in initial and final state radiation (ISR/FSR), tuning of the damping parton showering factor (hDAMP) and underlying event perturbations (UE).

Only $t\bar{t}$ samples are considered as this background is by far the most dominant one in this analysis. The jet multiplicity is compared to the nominal sample which has the same tune and generator (POWHEG).

In order to estimate the uncertainty for this analysis, the event yields for each perturbation in each category of this analysis is evaluated. The deviation w.r.t. the nominal sample is calculated propagated as a fully correlated normalization uncertainty to the datacards.

It is found that the FSR uncertainty in the $\ell + \tau_h$ category is the dominant one which goes up to 10% for the 1 jet category and 20% in the 2-3 jet category. For convenience this uncertainty per final state and in each category is shown in figure **4-15**. Clearly this effect scales with the jet multiplicity and accounts for jet migrations.

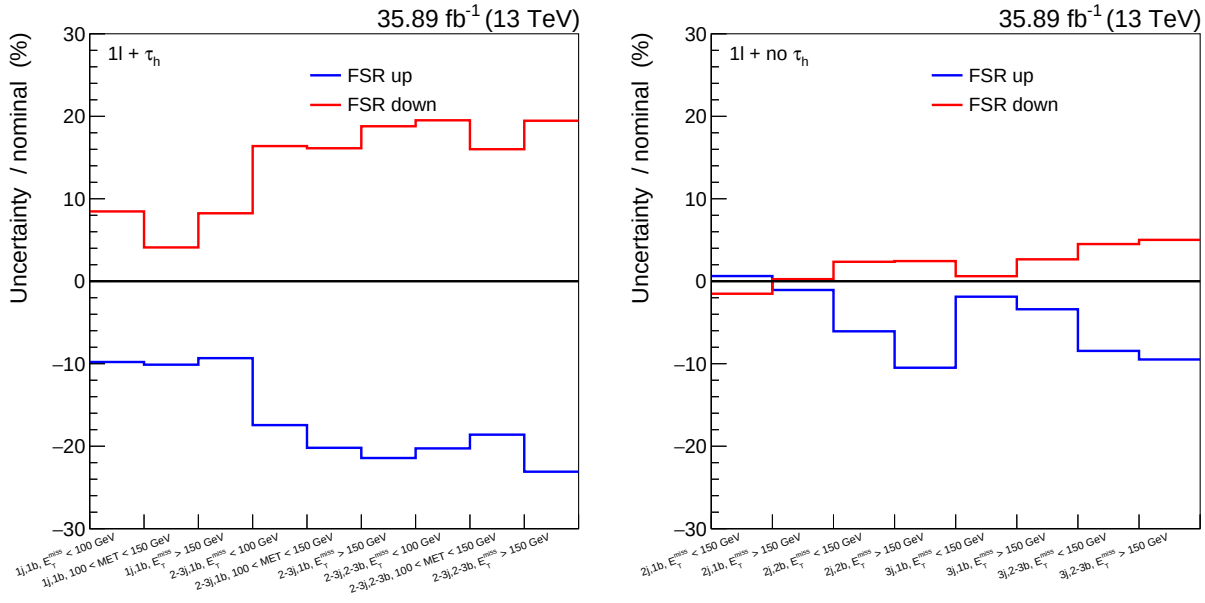


Figure **4-15**: $t\bar{t}$ parton shower uncertainties (final state radiation, FSR) for the $\ell + \tau_h$ (left) and $\ell+$ no τ_h (right) category.

4.10.1 Template Binning

A dedicated binning strategy has been developed to cope with limited Monte Carlo statistics. Especially in the m_T tails and less abundant background regions (e.g. signal regions), enough statistics needs to be ensured to obtain a reliable fit. Furthermore, due to the MG5_aMC@NLO signal event generator and $t\bar{t}$, about 33% of the events have a negative weight and such events can lead to negative yields in the template bins.

The binning of the m_T templates starts with a fine 10 GeV grid from 0 to 6000 GeV. Merging of adjacent bins (looping from the high m_T values to low m_T values) is then performed using the following criteria:

- statistical uncertainty of the total expected background must be less than 20%;
- statistical uncertainty of the signal must be less than 20%;
- no negative event bins in signal nor background should be present.

As a result, the binning becomes mass dependent. The statistical thresholds have been chosen to have a solid estimation of the signal and background in all the templates.

4.10.2 Shape Systematics and Template Smoothing

Nuisances as jet energy scale or resolution directly alter the jet p_T and can cause category migration of events in the n-jet direction. Indeed, by moving up or down jets close to the 30 GeV threshold, jets can enter or leave the event phase space leading to a redistribution of the events in the categories. Besides category migration, more strongly such nuisances induce bin migrations in the rather fine binned templates. Especially in the high p_T tails where the uncertainties are larger (they scale with the p_T), strong bin migrations have been observed.
As a result, the uncertainty of the scale or resolution is ill defined and the up/down shape templates exhibit spiky behavior. In turn, this can cause artificial constraints in the fit which can leas to wrong or biased results. To overcome these artificial constraints, the nuisance shape templates which are ill defined are smoothed out in order to remove the spikes while preserving a conservative estimation of the uncertainty.

The smoothing is performed on the ratio up/nom and down/nom binned templates separately before applying the binning procedure (see section 4.10.1). A Gaussian kernel smoothing has been implemented where the parameters are tuned to have a reliable smoothing without loss of shape variations. The algorithm consists of two steps:

1. initial re-binning: the up and down templates are initially re-binned such that the bins in the nominal template has a maximal statistical uncertainty of 20%. In this way it is ensured the up and down variation are defined in each bin(e.g. avoid empty bins) and the very spiky behavior due to lack of statistics is removed;

2. smoothing: a Gaussian kernel using a width of 10 bins performs the smoothing by altering the up/down ratio. The algorithm assigns a value for each bin i based on the neighboring values, with a Gaussian weight factor defined as:

$$w_{i,j} = \exp\left(-\frac{(i-j)^2}{2\sigma}\right), \tag{4-3}$$

   with σ the width (taken to be 10 bins), and the index j run over all the neighboring bins. The event contents in each bin are reweighted and a new smoothed value is assigned for bin i.

The smoothing procedure has been applied on all nuisances and validated by a visual cross check of the fitted templates, except for the pile-up which has found to have a highly ranked impacted on the signal when combining with the hadronic final state (see section 4.12.3) and therefore is not smoothed out in order to not lose sensitive shape information.

4.11 Treatment of the Intermediate Mass Regime

For charged Higgs masses around the top quark, the interference of the heavy and light charged Higgs production schemes must be taken into account as well as top-width effects. Several resonances contribute to the total cross section of the intermediate charged Higgs masses, where also resonances with neutral scalars (H, A) - if they exist - must be taken into. Due to the interplay of the neutral scalars, the charged Higgs production mode becomes model dependent.

Currently, from theory point of view it is difficult to calculate the interference effects up to NLO accuracy level, therefore only the leading order results are available. Several LO samples ranging from 145 to 200 GeV are simulated, both with and without neutrals, which take into account the interference effect (see section 3.3.3 for a list of the processed samples). Because this analysis aims to quote a model-independent result, the neutral samples have not been used.

In this analysis, all the light (top decay production mode) and heavy (associated top production mode) charged Higgs samples are simulated up to NLO accuracy. The transition from NLO to LO will be non-smooth due to different acceptance effects between LO and NLO samples. This is shown in figure **4-16** where the expected upper limit is shown for the NLO light/heavy samples with the intermediate LO samples superimposed. For the combined result, the expected upper limit for the intermediate LO samples is roughly 30% higher. The non-smooth behavior in the $\ell+$ no τ_h category is mainly driven by the constraint in jet multiplicity (maximum 3), for which the NLO/LO is very sensitive to.

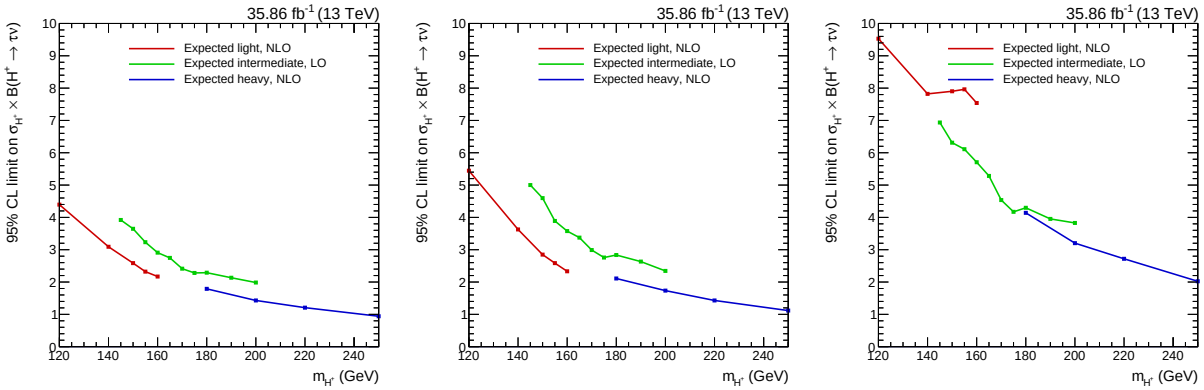


Figure **4-16**: Upper limit of the light/heavy NLO samples with the intermediate LO samples superimposed. The left plot shows the combined results of the middle $(\ell + \tau_h)$ and right $(\ell+$ no $\tau_h)$ result.

Derivation of LO to NLO Scale Factors

In order to have a smooth transition between the light and heavy mass regime, scale factors have been derived to rescale the intermediate signal from LO to NLO accuracy. The scale factors are defined as the ratio of NLO to LO event yields based on the signal samples. An identical approach is used in the hadronic analysis.

Differences between LO and NLO are studied based on LO samples in overlap with the light NLO (140–160 GeV) and heavy NLO (180–220 GeV) samples. It has been found their m_T shapes are similar, though deviations in acceptance from jet multiplicity and production mode (light vs. heavy) have been observed. This is shown in Fig. **4-17** where the normalized generator jet multiplicity distributions are shown for both LO and NLO, for charged Higgs masses of 150 GeV (light production mode) and 180 GeV (heavy production mode). To cope with these acceptance differences, NLO/LO scale factors are derived using the overlapping samples:

- derived for both light (140–160 GeV) and heavy mass (180–220 GeV) regime;
- derived in each category of the final states.

The scale factors are applied to the intermediate mass samples for 165 GeV (light scale factors) and 170/175 GeV (heavy scale factors). Statistical uncertainties in the scale factors are taken into account as normalization uncertainties.

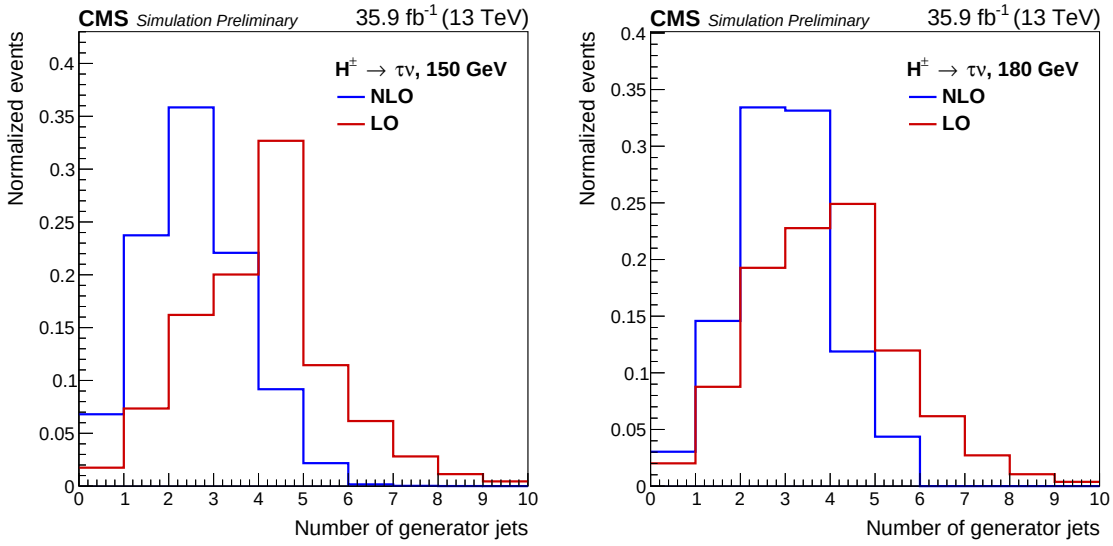


Figure **4-17**: Normalized generator jet multiplicity distributions for both LO (red) and NLO (blue), for charged Higgs masses of 150 GeV (light production mode, left) and 180 GeV (heavy production mode, right) [5].

The NLO/LO ratio is computed for each category in this analysis, and are averaged over the light and heavy mass regime in order to obtain a light and heavy LO to NLO scale factor per category. All scale factors and uncertainties are summarized in Table **4-3**, the detailed ratios per mass point and average are shown in Figures **4-18** and **4-19**.

Category	Light scale factors	Heavy scale factors
$\ell+$ no τ_h, muon+electron		
2j,1b, E_T^{miss} 150 GeV	1.28 ± 0.03	1.45 ± 0.04
2j,1b, E_T^{miss} > 150 GeV	0.95 ± 0.03	1.19 ± 0.04
2j,2b, E_T^{miss} 150 GeV	1.78 ± 0.10	3.30 ± 0.21
2j,2b, E_T^{miss} > 150 GeV	1.34 ± 0.10	2.79 ± 0.20
3j,1b, E_T^{miss} 150 GeV	0.40 ± 0.01	0.85 ± 0.03
3j,1b, E_T^{miss} > 150 GeV	0.34 ± 0.02	0.66 ± 0.03
3j,2b, E_T^{miss} 150 GeV	0.73 ± 0.04	2.22 ± 0.11
3j,2b, E_T^{miss} > 150 GeV	0.58 ± 0.05	1.64 ± 0.10
$\ell + \tau_h$, muon+electron		
1j,1b, E_T^{miss} 100 GeV	2.04 ± 0.06	1.78 ± 0.07
1j,1b, 100 < E_T^{miss} 150 GeV	1.76 ± 0.06	1.80 ± 0.07
1j,1b, E_T^{miss} > 150 GeV	1.72 ± 0.11	1.58 ± 0.09
2-3j,1b, E_T^{miss} 100 GeV	0.17 ± 0.02	0.68 ± 0.05
2-3j,1b, 100 < E_T^{miss} 150 GeV	0.18 ± 0.02	0.67 ± 0.04
2-3j,1b, E_T^{miss} > 150 GeV	0.14 ± 0.02	0.46 ± 0.04
2-3j,2b, E_T^{miss} 100 GeV	0.78 ± 0.07	2.08 ± 0.18
2-3j,2b, 100 < E_T^{miss} 150 GeV	0.71 ± 0.06	2.24 ± 0.17
2-3j,2b, E_T^{miss} > 150 GeV	0.61 ± 0.08	1.46 ± 0.13

Table 4-3: Derived scale factors as average for light and heavy mass regime, per category.

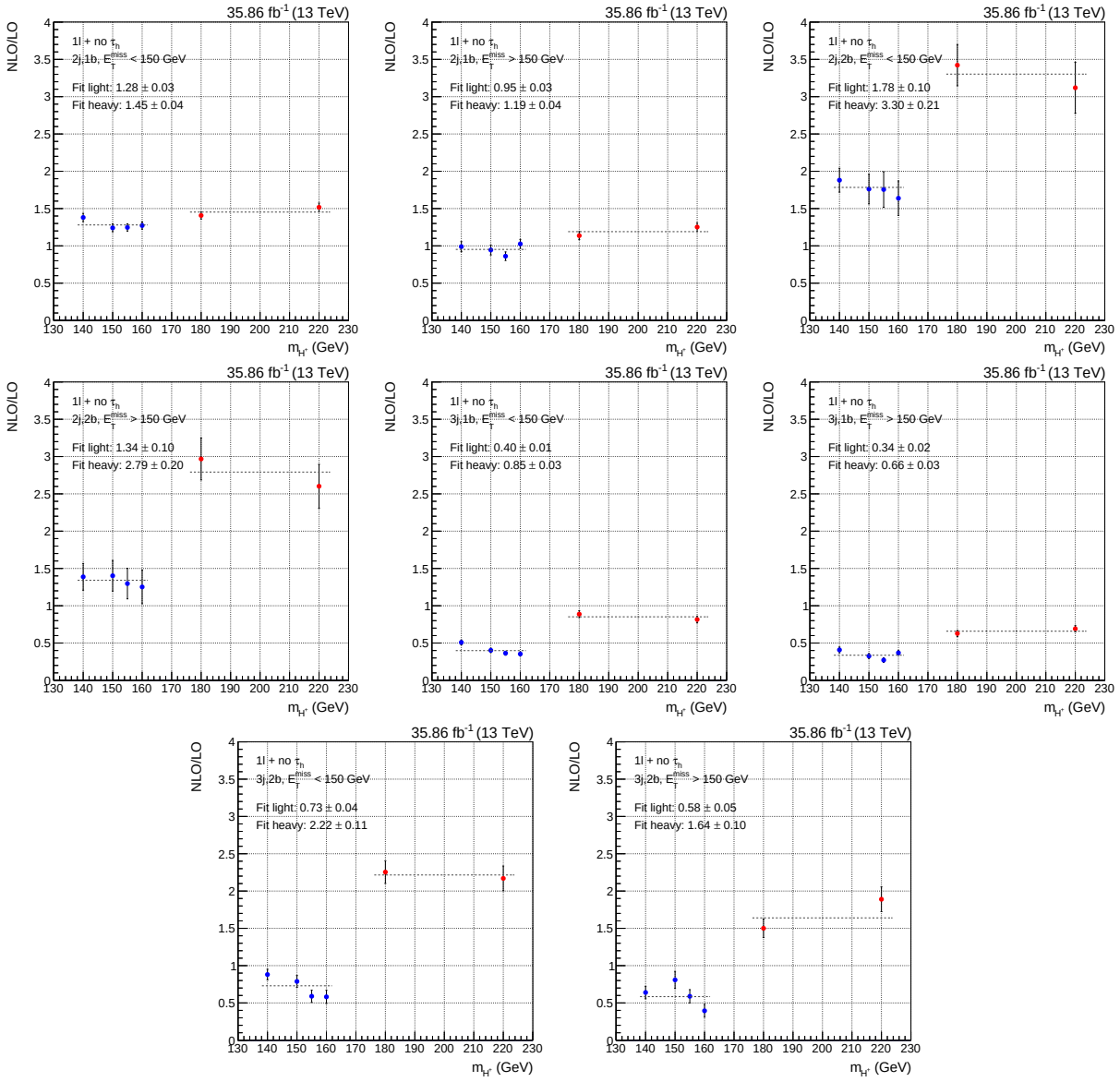


Figure **4-18**: NLO/LO scale factors per mass point with the fitted average and uncertainty for the ℓ+ no τ_h category.

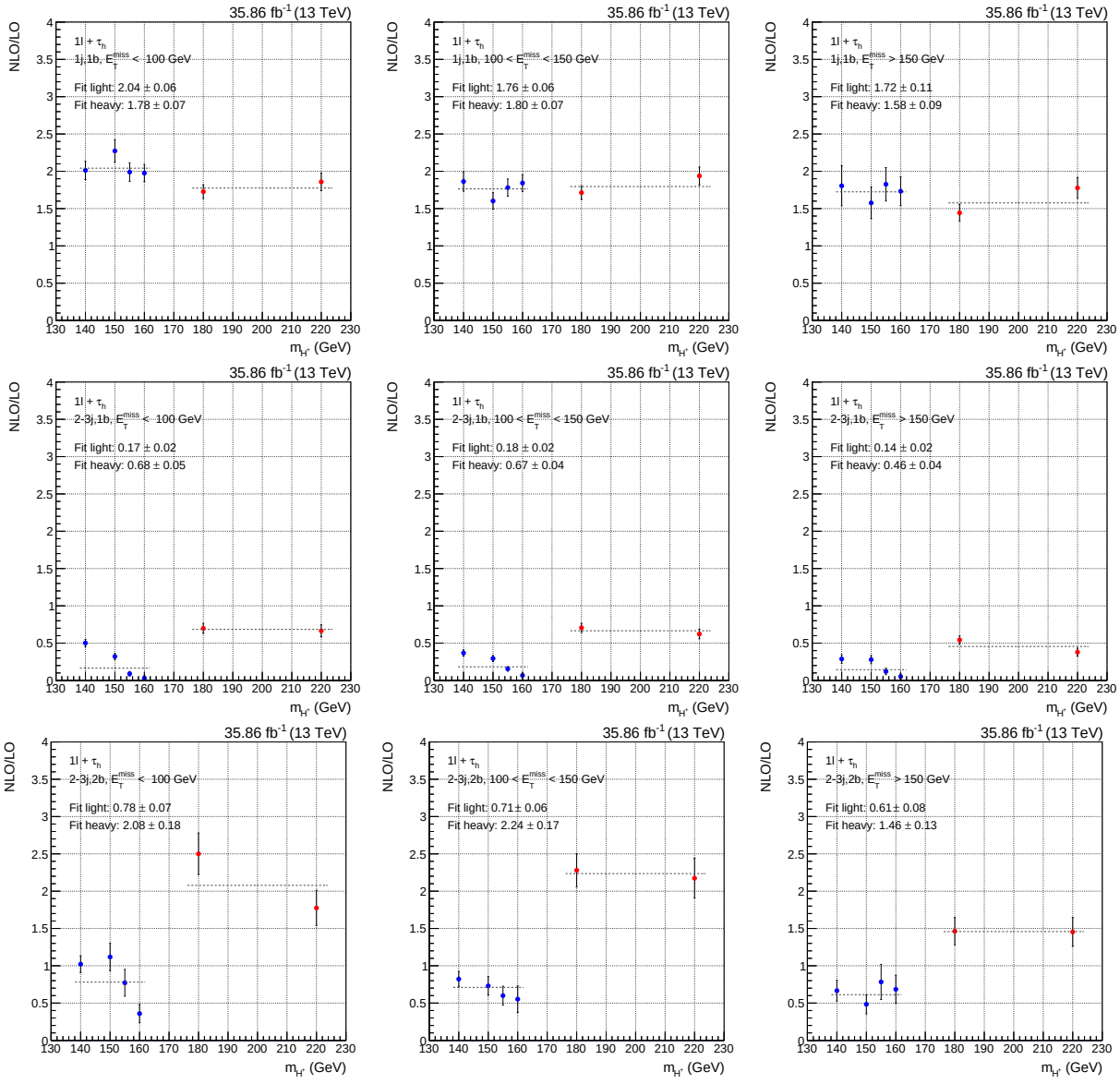


Figure **4-19**: NLO/LO scale factors per mass point with the fitted average and uncertainty for the $\ell + \tau_h$ category.

Closure of LO Samples

It is expected the interference effect becomes negligible for masses far from the top mass. This has been verified with the LO samples only. In Fig. **4-20**, the intermediate LO samples (without scale factors) and the LO light/heavy mass samples are shown. The closure is only present for the heavy mass regime, whereas for the light mass regime the intermediate mass samples are 20% lower w.r.t. the non-intermediate light LO samples.

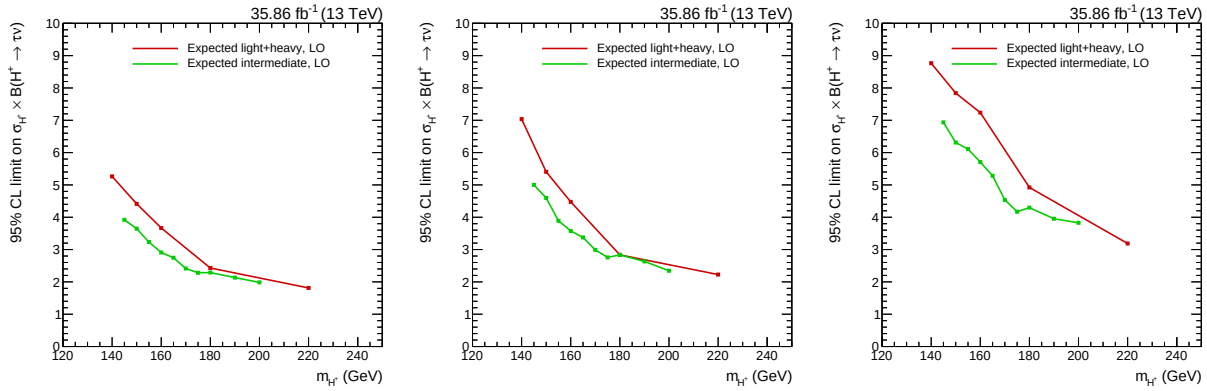


Figure **4-20**: Closure test for the intermediate LO and light/heavy LO samples. The left plot shows the combined results of the middle ($\ell + \tau_h$) and right (ℓ+ no τ_h) result.

Results Upon Application of LO to NLO Scale Factors

The obtained scale factors are applied to the intermediate samples: the light scale factors for the charged Higgs samples below 173 GeV and the heavy scale factors for charged Higgs samples above 173 GeV. The uncertainties on the scale factors are propagated as normalization uncertainties in the fit. The expected limits are shown in Fig. **4-21**. On average the intermediate and NLO samples in both mass regimes converge to each other. As already stated before, the NLO/LO scale factor in the $\ell+$ no τ_h is very sensitive.

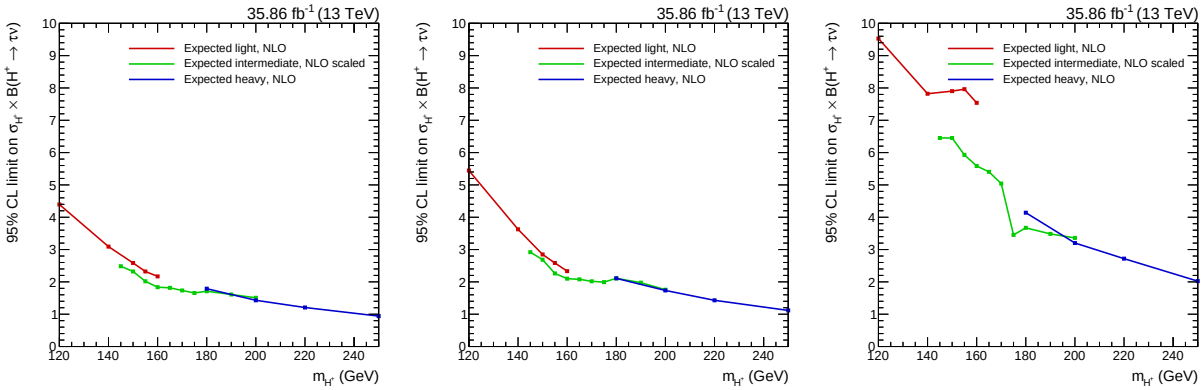


Figure **4-21**: Upper limit of the light/heavy NLO samples with the intermediate samples rescaled to NLO. The left plot shows the combined results of the middle ($\ell+\tau_h$) and right ($\ell+$ no τ_h) result.

4.12 Results

The Standard Model is assumed to contribute exclusively to the yields in each signal and control region. This hypothesis is tested against the additional presence of a charged Higgs boson production and subsequent decay, that would result in increased yields (mostly) in the signal regions. 95% CL expected upper limits are derived on a charged Higgs boson production cross section, for different charged Higgs boson mass hypotheses. The statistical method used to report the results is the modified frequentist approach, also known as CLs (see section 3.4).

A simultaneous fit is performed on both $\ell + \tau_h$ and $\ell+$ no τ_h categories. As explained in section 4.8.3, the fits are based on $m_T(\ell, E_T^{miss})$ shape templates which both consist of sub-categories to enhance the signal sensitivity:

- $\ell + \tau_h$: simultaneous fit of the m_T templates in both electron and muon final states, categorized in number of jets (1 and ≥ 2), number of b-tagged jets (1 and ≥ 2) and 3 MET categories ($70 < E_T^{miss} < 100$, $100 < E_T^{miss} < 150$ and $E_T^{miss} > 150$). This results in 16 templates;
- $\ell+$ no τ_h: simultaneous fit of the m_T templates in both electron and muon final states, categorized in number of jets (2 and 3), number of b-tagged jets (1 and ≥ 2) and 2 MET categories ($100 < E_T^{miss} < 150$ and $E_T^{miss} > 150$). This results in 18 templates.

The m_T templates are binned such that sufficient Monte Carlo statistics is present in each bin and to avoid any negative bin contributions due to the MadGraph5_aMC@NLO event generators for signal and $t\bar{t}$ (see section 4.10.1), resulting in a mass dependent binning. All systematics as explained in section 4.10 are included in the fit and fully correlated among both categories.

Results are computed for the entire charged Higgs mass range from 80 GeV to 3 TeV with the luminosity of the full 2016 dataset being 35.9 fb^{-1}. For charged Higgs masses below the top quark mass, the dominant decay mode is through top decays which is incorporated in the fit using a modified physics model which uses as POI is the branching ratio of the $t\bar{t}$ decay to a charged Higgs. The fitted branching BR ratio is transformed to an upper limit cross section σ using the following formula (cfr. equation 2-12):

$$\sigma = 2.0 \times BR \times (1 - BR) \times 831.76 \text{ (pb)}. \tag{4-4}$$

4.12.1 Post-fit Templates and Yields

To extract the post-fit distributions and yields, the background-only fit has been performed using the charged Higgs 200 GeV templates. During the fit the following processes are taken into account: $t\bar{t}$, top (single top and $t\bar{t}+\mathrm{X}$), $\mathrm{W}^{\pm}$+jets, Z/γ^{*}+jets and diboson (WW, ZZ, WZ). Two signal-sensitive post-fit $\mathrm{m_T}$ templates are shown in figure **4-22**. The other categories are summarized in the bar plot in figure **4-23**, where the post-fit background only event yields are shown for each category in this analysis (inclusive in the lepton flavor). From these results, the fit is compatible with the SM hypotheses. A summary of the post-fit event yields is discussed together with the hadronic final state, summarized in table **5-2**.

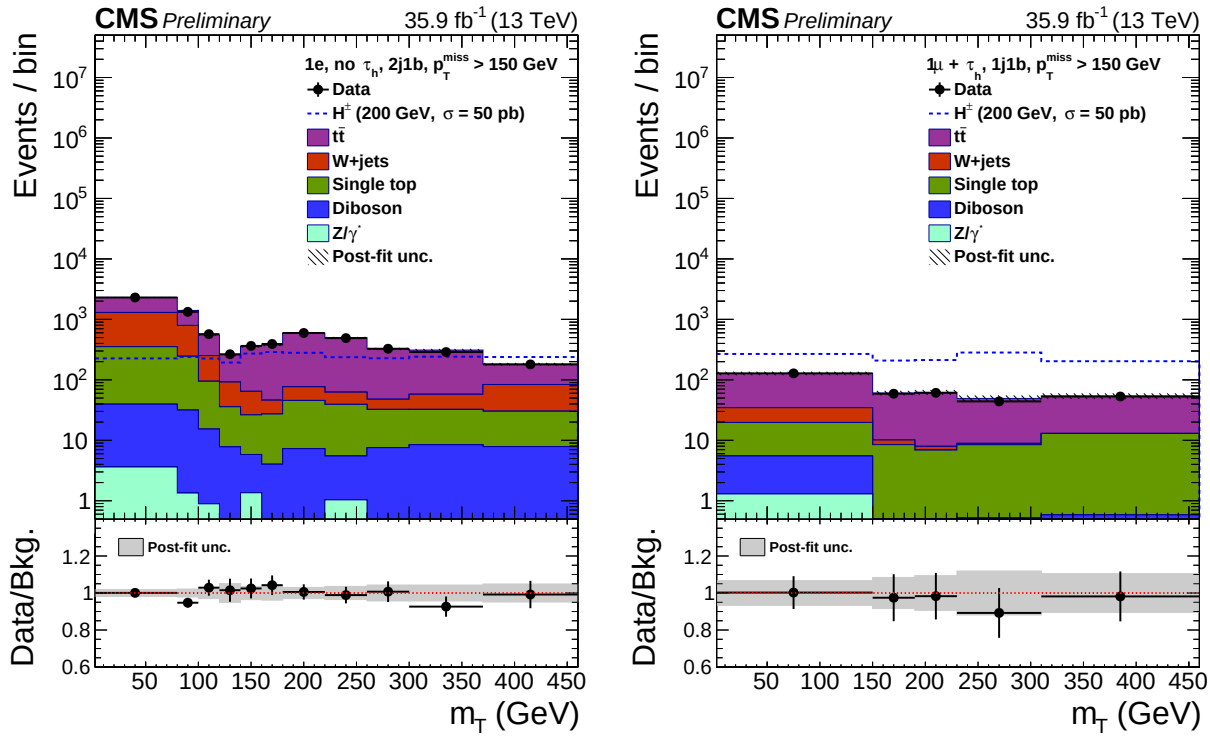


Figure **4-22**: The transverse mass distributions for two categories with high signal sensitivity after a background-only fit to the data. Left: category with one electron, ℓ+ no τ_h, two jets where one identified as a b jet, and $\mathrm{E_T^{miss}} > 150$ GeV. Right: category with one muon, $\ell + \tau_h$, one jet identified as a b jet and $\mathrm{E_T^{miss}} > 150$ GeV. In both categories, the last bin shown extends to 5 TeV [5].

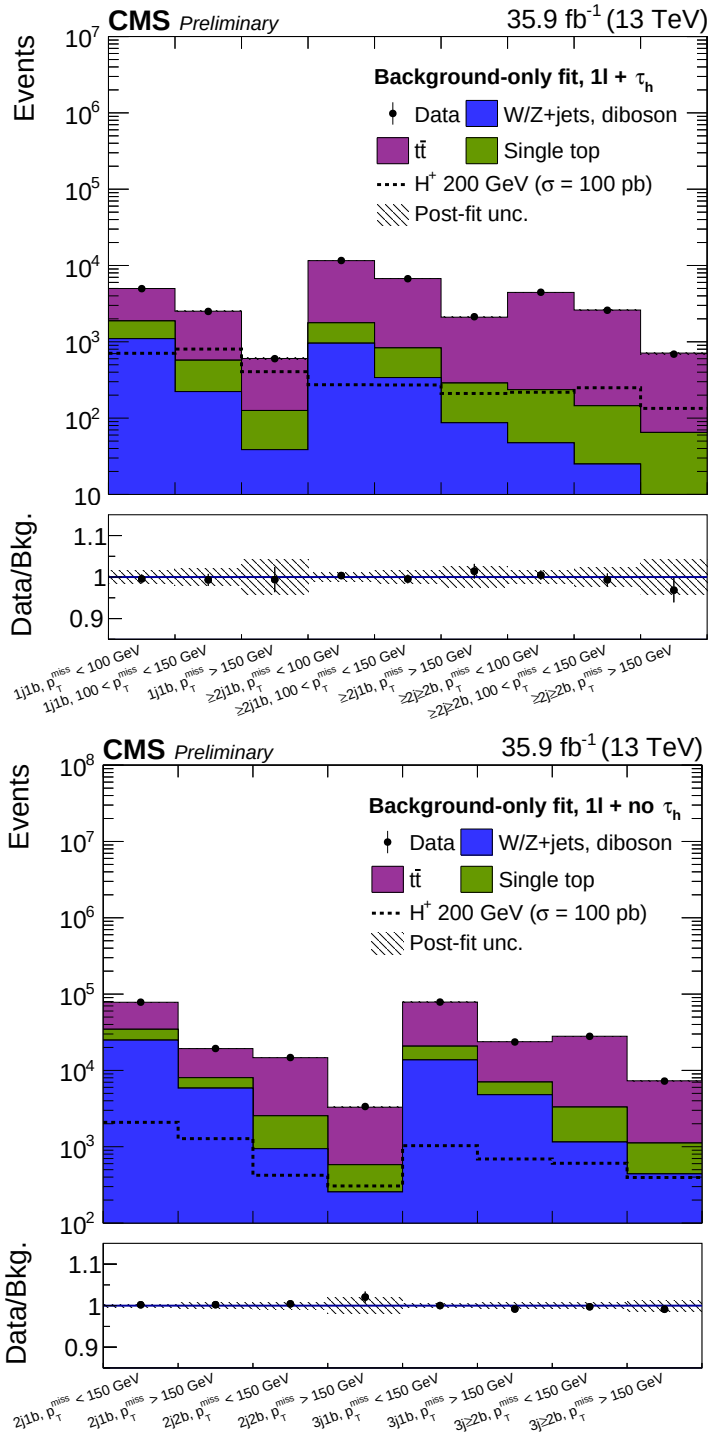


Figure **4-23**: Post-fit event yields in all the categories after a background-only fit to the data. Top: $\ell + \tau_h$ category, bottom: $\ell+$ no τ_h category. The yields are shown inclusive in lepton flavor [5].

4.12.2 Expected Limits

The expected and observed limits obtained with the m_T templates using the asymptotic approach are shown in figure **4-24** (left). No excess is observed across the entire mass range meaning that the fit is compatible with the SM expectations.

The relative contributions of the individual categories to the fit is shown in figure **4-24** (right), as the ratio of the upper limit over the combined expected upper limit. This ratio is obtained from the expected upper limit for the individual $\ell + \tau_h$ and $\ell+$ no τ_h categories, as shown in figure **4-25**. At high mass clearly the $\ell+$ no τ_h is dominant whereas towards lower masses the $\ell + \tau_h$ becomes dominant. This is in agreement with the signal acceptance yields as shown in figure **4-6**.

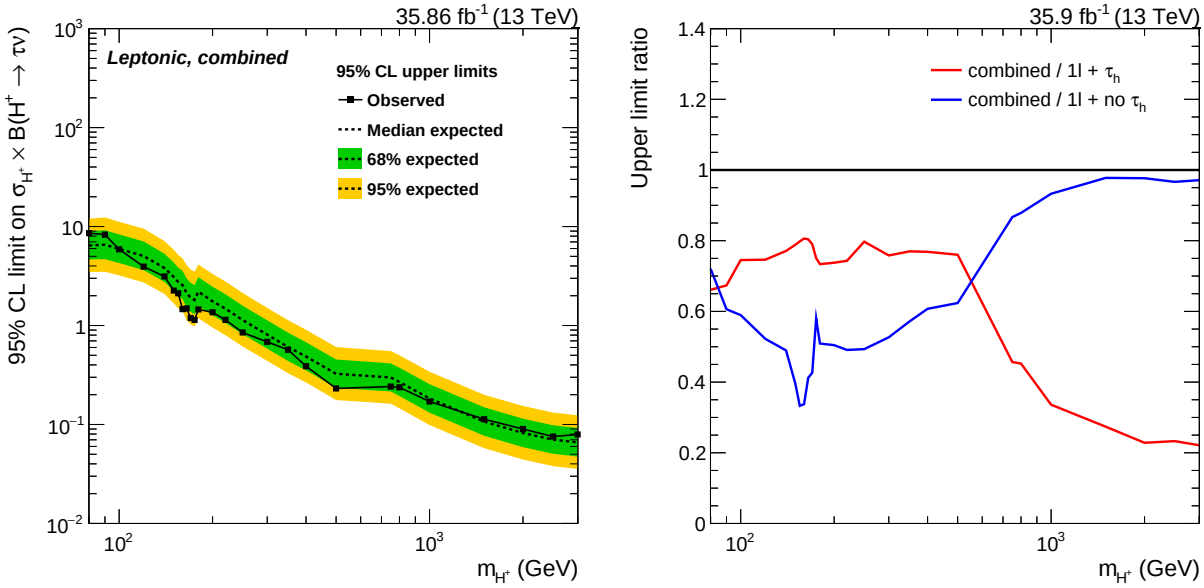


Figure **4-24**: Left: 95% CL expected upper limits on the cross section times branching fraction for charged Higgs production to $H^\pm \to \tau^\pm\nu_\tau$ for the combined $\ell + \tau_h$ and $\ell+$ no τ_h channels. Right: upper limit ratio plot.

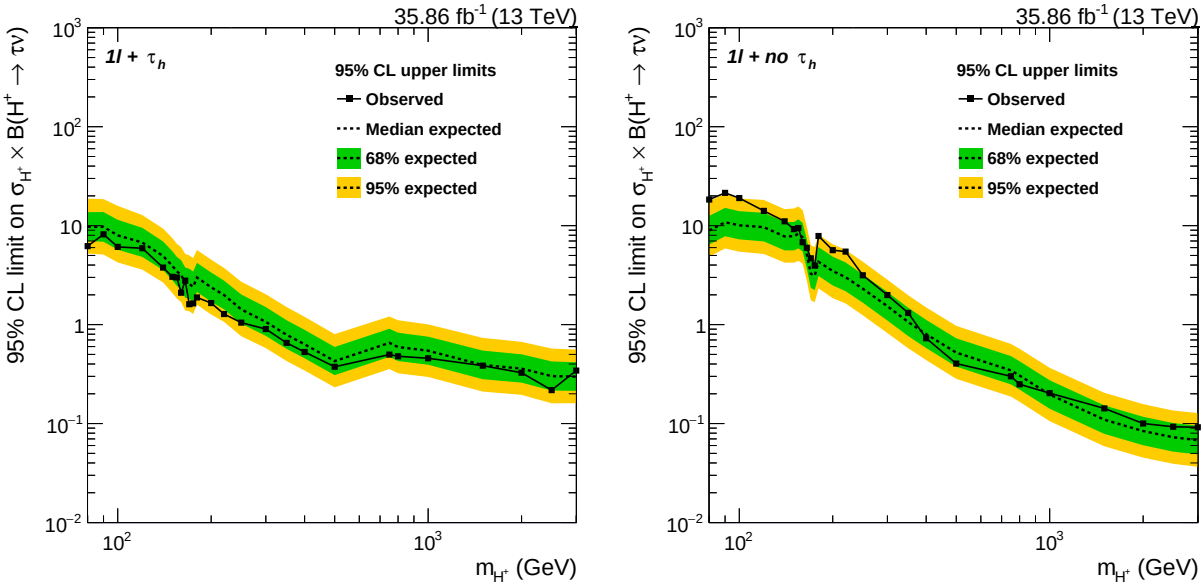


Figure **4-25**: 95% CL expected upper limits on the cross section times branching fraction for charged Higgs production to $H^{\pm} \to \tau^{\pm}\nu_{\tau}$ for the $\ell + \tau_h$ (left) and ℓ+ no τ_h (right) channels.

4.12.3 Fit Tests

Several fit tests have been performed throughout the analysis in order to check the consistency of the fit and the nuisances. During the design of the analysis, the validity of the nuisances (and expected limits) have been tested thoroughly on the Asimov dataset. In particular, for each nuisance the constraint, pulls and impact on the signal sensitivity is validated, both on the Asimov and later on the observed dataset. Especially the validation of the Gaussian kernel smoothing algorithm was partially based on the constraints on the Asimov dataset, where to constraint $\Delta\theta$ is defined as the uncertainty of the nuisance θ after the fit w.r.t. its pre-fit value. To check the impact of a nuisance on the signal sensitivity, the shift $\Delta\mu$ is computed, defined as follows

$$\Delta\mu(\pm) = \hat{\mu}(\hat{\theta} \pm \Delta_{\theta}) - \hat{\mu}(\hat{\theta}), \tag{4-5}$$

where $\hat{\theta}$ is the best-fit value with all nuisances floating and $\hat{\theta} \pm \Delta_{\theta}$ the fit where the nuisance is brought to its $\pm\sigma$ value. It is obtained by performing two separate fits per nuisance in addition to the initial global fit. It contains the direction where the signal strength μ is going when a nuisance is altered in a certain direction.

The impacts are tested on the Asimov dataset and verified with the observed dataset,

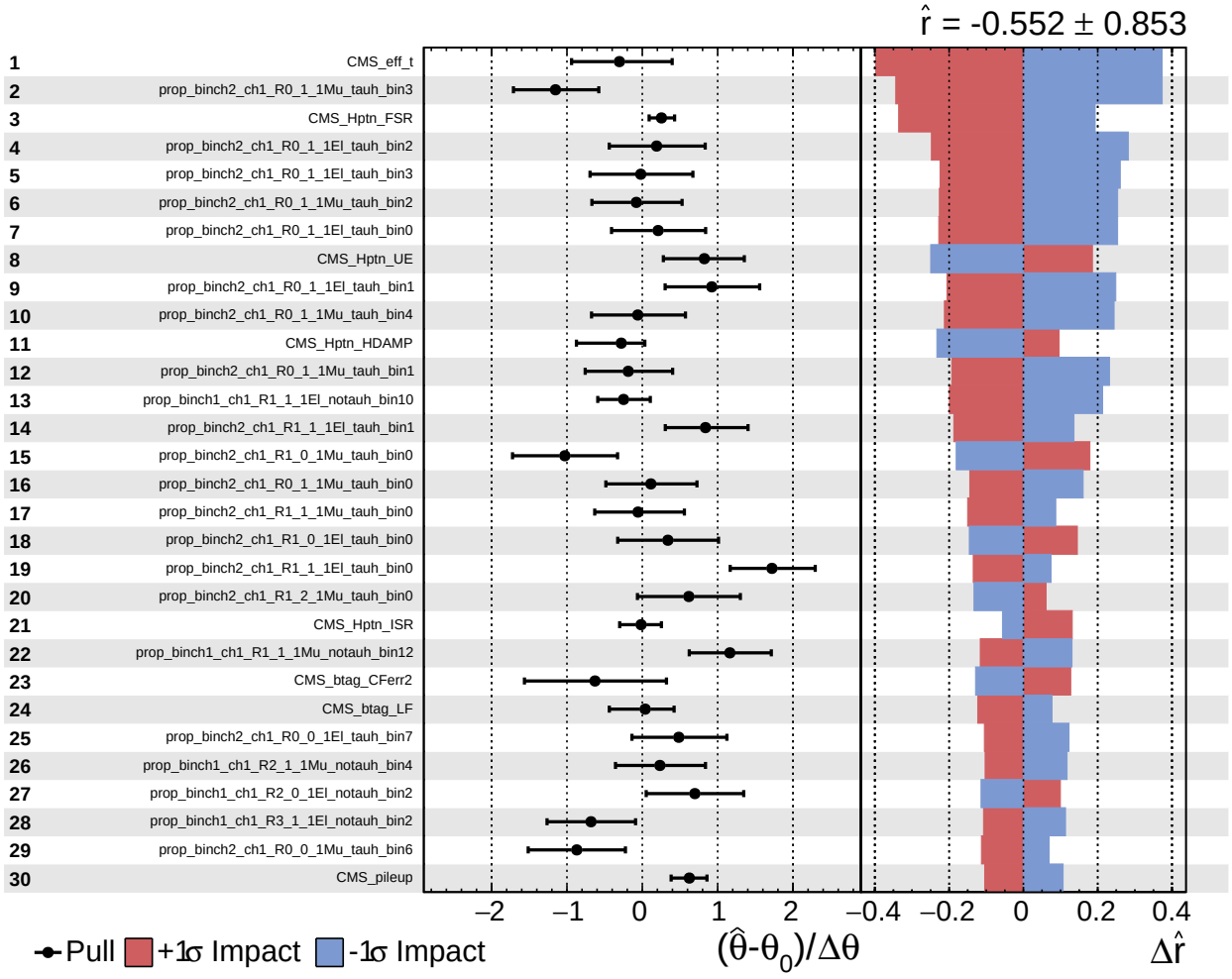


Figure **4-26**: 30 highest nuisance impacts and pull distributions for mass point $m_{H^\pm}$ = 200 GeV based on the observed dataset.

where no large differences among them are expected. The observed impacts for the combined fit are shown in the right column of figure **4-26**, for the fit of $m_{H^\pm}$ 200 GeV. Only the 30 highest impacting nuisances are shown, including both statistical and systematic nuisances. The highest impacting nuisances are the τ_h identification uncertainty, which migrates events between the $\ell+\tau_h$ and $\ell+$ no τ_h category, hence it is sensitive to the signal: if it is brought to -1σ, events from the $\ell+\tau_h$ category are migrated to the $\ell+$ no τ_h category, yielding a higher sensitivity to the signal strength, which is confirmed from figure **4-24** where the $\ell+\tau_h$ has the highest sensitivity at 200 GeV. The third impacting nuisance is the FSR parton shower uncertainty, which can be understood in a similar reasoning. Other high impacting nuisances are related to sensitive bin statistics.

The pull ξ_θ of each nuisance is also carefully studied, which is defined as

$$\xi_\theta = \frac{\hat{\theta} - \theta_0}{\Delta\theta}, \tag{4-6}$$

with θ_0 the pre-fit position of the nuisance (usually taken as 0). It has been calculated both on the Asimov dataset and using the observed dataset. For the Asimov dataset (taken with signal strength zero), per definition, no nuisance should be pulled after the fit as the Asimov dataset respresents the total Monte Carlo sample itself. Also the signal strength, which is left floating during the fit, should be zero. Both necessities were confirmed on the Asimov dataset.

The pull distributions based on the observed dataset are also shown in figure **4-26** (middle column), but in figure **4-27** the pull distributions for the non-statistical nuisances are shown. A signal strength of zero was injected in the fit, and the signal is free to float during the fit. As can be seen from this figure, some nuisances are constrained which can be explained as follows:

- `CMS_scale_j/CMS_res_j/CMS_uncluster`: after smoothing the distributions the artificial constraint has been removed, but still it got constrained by the fit. This is an artifact due to the high categorization in number of jets and b-jets, which is sensitive to the jet energy scale, especially for the dominant $t\bar{t}$ background;

- `CMS_btag_LF/HF`: the same rationale as the `CMS_scale_j` is applicable to explain the constraints. Furthermore, the b-tagging uncertainties are split in 8 components where each component is sensitive to a region with particular b-hadrons (light, charm, b).

- `ISR/FSR`: these uncertainties are sensitive to category migrations in the number of jets. Also the pre-fit uncertainty is taken from additional samples and is an overestimation of the true uncertainty. Hence, constraints can be expected.

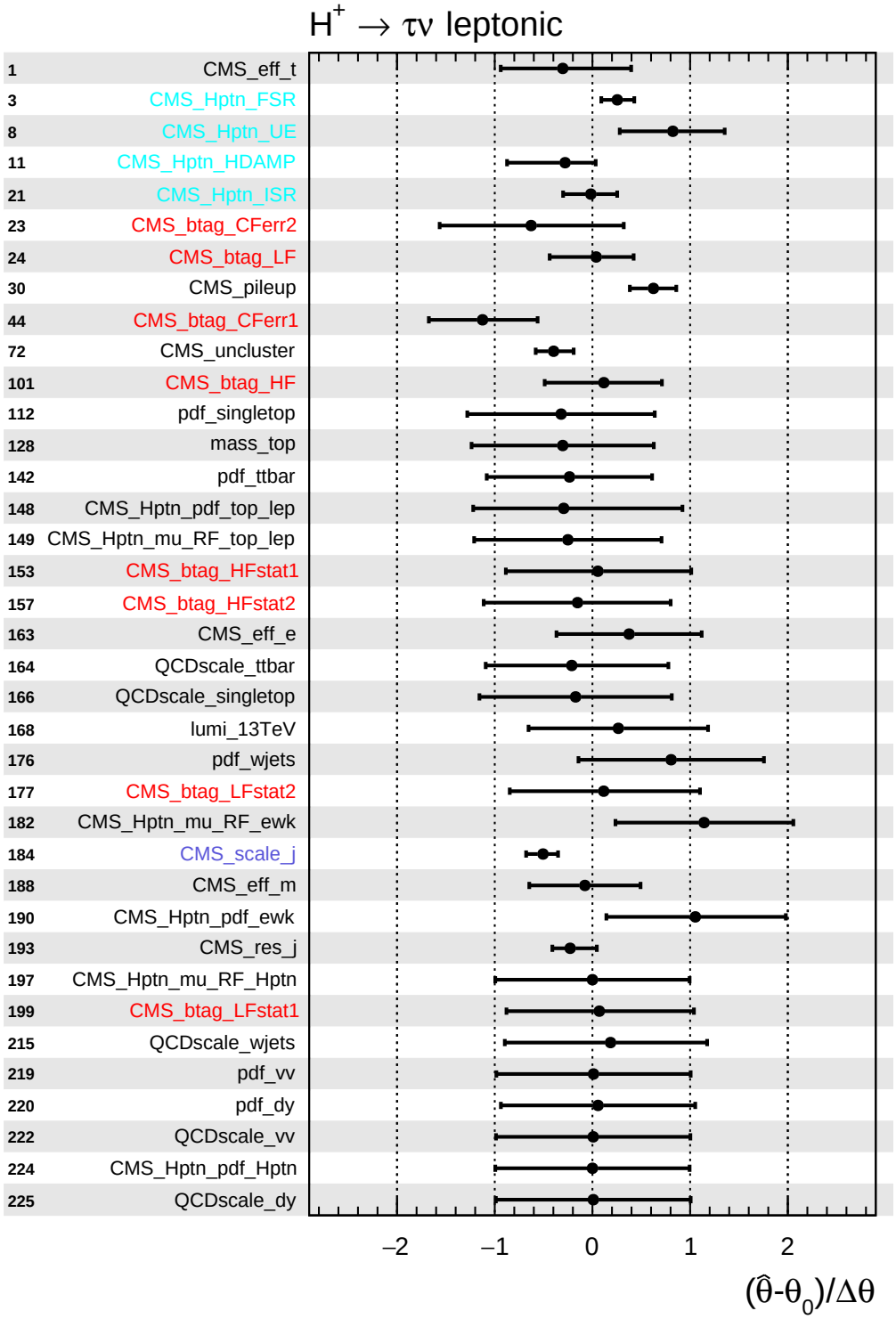


Figure **4-27**: Non-statistical pull distributions for mass point $m_{H^\pm} = 200$ GeV based on the observed dataset.

5 Searches in the $\tau\nu_\tau$ Hadronic Final State and Combination

The final result of the charged Higgs to $\tau\nu_\tau$ search involves both the leptonic and hadronic final state. Before describing the combination of both final states, the hadronic final state is briefly introduced. Both final states were developed in parallel in order to avoid any event overlap in their event phase spaces, such that the model independent combination can be performed smooth. At last, the model independent results are interpreted in several MSSM benchmark models, excluding $\tan\beta$ as function of the charged Higgs mass.

5.1 Hadronic Final State

The hadronic final state requires the presence of a hadronically decaying τ-lepton (τ_h) in association with large $\mathrm{E_T^{miss}}$ from the neutrinos (see figure **5-1** for a descriptive Feynman diagram of the process). Additional jets are present from the hadronic top quark decays. The presence of any lepton (electron, muon) is vetoed. Because all neutrinos originate from the charged Higgs decay, the transverse mass reconstruction is possible using $\tau_h + \mathrm{E_T^{miss}}$ system:

$$\mathrm{m_T} = \sqrt{2\mathrm{p_T}(\tau_h)\mathrm{E_T^{miss}}(1 - \cos\Delta\phi(\mathbf{p}_\mathrm{T}(\tau_h), \mathbf{p}_\mathrm{T}^{\mathrm{miss}}))}. \tag{5-1}$$

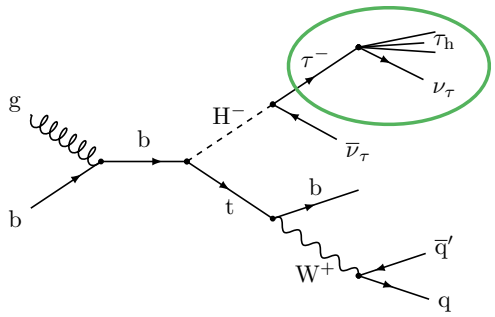


Figure **5-1**: Feynman diagram of the hadronic final state.

5.1.1 Event Selection and Categorization

Events are selected using an online trigger composed of a τ_h and $\mathrm{E_T^{miss}}$ leg with respective $\mathrm{p_T}$ thresholds of 50 and 90 GeV, with an additional selection on the τ_h leading track $\mathrm{p_T}$ of 30 GeV. Further offline selection criteria are applied which are optimized to enhance signal sensitivity:

- hadronically decaying τ-lepton (τ_h), 1-prong only;
- $\mathrm{E_T^{miss}} > 90$ GeV;
- at least 3 jets, one b-tagged jet;
- veto on electrons or muons.

An additional angular cut to suppress the jet $\rightarrow$ fake τ_h background (mainly from QCD multijets) is applied, defined as follows:

$$R_{\mathrm{bb}}^{\mathrm{min}} = \min\left\{\sqrt{\left(180^\circ - \Delta\phi(\tau_h, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})\right)^2 + \left(\Delta\phi(\mathrm{jet}_n, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})\right)^2}\right\} > 40^\circ. \tag{5-2}$$

This formula is justified by considering that the majority of QCD multijet events have a back-to-back di-jet pair, where one of the jets is mis-identified as a τ jet. Hence the $\mathrm{E_T^{miss}}$ is overestimated in the mis-identified τ direction which simultaneously minimizes $\Delta\phi(\tau_h, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})$ and maximizes $\Delta\phi(\mathrm{jet}_n, \mathbf{p}_{\mathrm{T}}^{\mathrm{miss}})$, where n runs over the reconstructed jets. An optimal working point of 40° has been chosen across entire mass range (see Fig. **5-2** (left)).

In order to further enhance the signal sensitivity, the event phase space is divided in two categories by exploiting the difference in polarization states of the τ_h originating either from $\mathrm{H^\pm}$ decays (scalar, left-handed τ_h) or from $\mathrm{W^\pm}$ decays [101]. (vector, right-handed τ_h). Because longitudinal polarization states tend to have a higher leading track $\mathrm{p_T}$, the variable R_τ defined as

$$R_\tau = \frac{\mathrm{p_T(\text{leading track})}}{\mathrm{p_T}(\tau_h)}, \tag{5-3}$$

is sensitive to the τ_h polarization state, as can be seen from Fig. **5-2** (right). Indeed, for low values of R_τ the backgrounds are dominant whereas at high values of R_τ the distribution is more pure in signal. Events are categorized based on $R_\tau > 0.75$ and $R_\tau < 0.75$ which is also optimized across entire mass range.

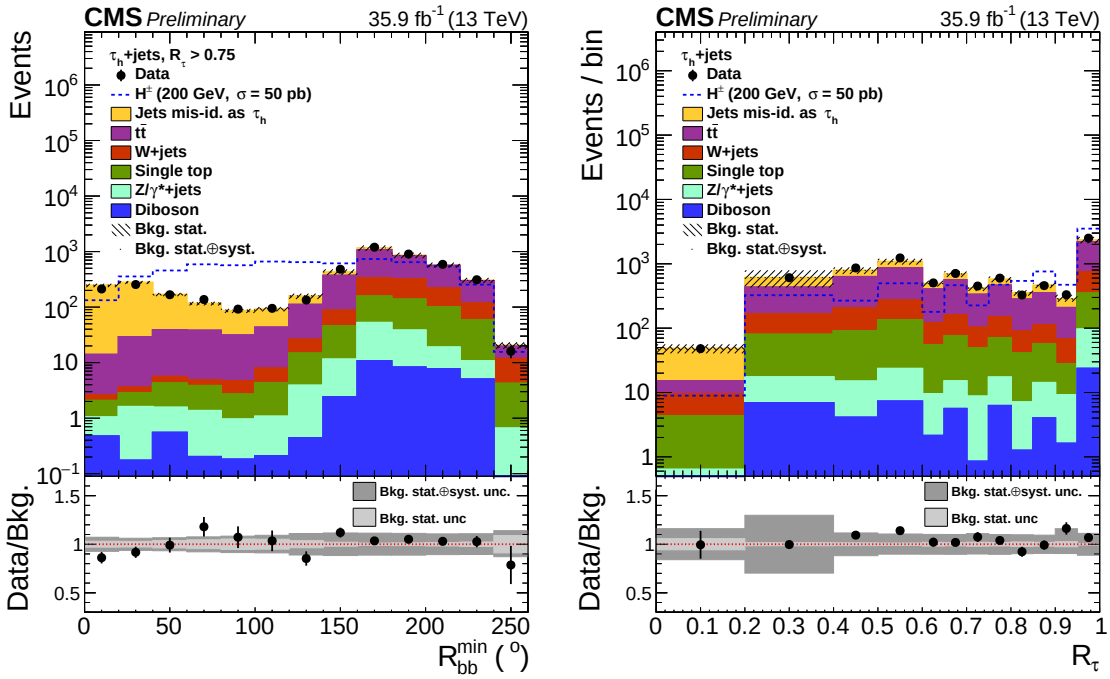


Figure **5-2**: Left: distribution of the $R_{\mathrm{bb}}^{\mathrm{min}}$ variable where it can be seen that the jet $\rightarrow$ fake τ_h background piles up at low values of $R_{\mathrm{bb}}^{\mathrm{min}}$. Right: R_τ distribution which separates the signal (high R_τ) and background (low R_τ) τ_h polarization states [5].

5.1.2 Background Estimation

The major backgrounds are jet $\rightarrow$ fake τ_h , $t\bar{t}$ and $\mathrm{W}^{\pm}$+jets. Minor backgrounds such as single top, Z/γ^*+jets and diboson are also considered in the analysis. All backgrounds are estimated from simulation, except the jet $\rightarrow$ fake τ_h which is estimated using a data-driven technique (fake rate).

The data-driven method in the hadronic final state consists of deriving transfer factors in a loose region (pure in fake τ_h) between the nominal and inverted τ_h isolation (see figure **5-3**). The transfer factors are derived in bins of τ_h $\mathrm{p_T}$ and η to minimize correlations and mitigation of detector inefficiencies. Small contaminations of $t\bar{t}$ and $\mathrm{W}^{\pm}$+jets are estimated from simulation and subtracted. To estimate the jet $\rightarrow$ fake τ_h in signal region, the transfer factors are applied in the application region. Systematic uncertainties related to the transfer factors are propagated to the final fit.

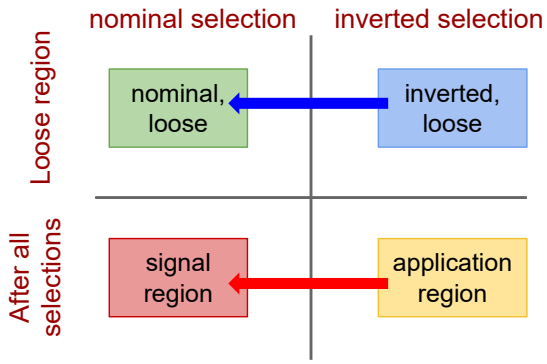


Figure **5-3**: Schematic representation of the fake rate method, a data-driven technique to estimate the jet $\rightarrow$ fake τ_h .

5.1.3 Systematics

Similar experimental and theoretical uncertainties (where relevant) as described in section 4.10 are taken into account. Most of them have similar implementations between hadronic and leptonic analyses such that a possible correlation among them is valid. Additional, dominant uncertainties in the hadronic final state are listed below. For an estimation of the pre-fit size, see table **5-1**.

- jet $\rightarrow$ fake τ_h : uncertainties related to the fake rate method are threefold: 1) statistical uncertainties in data an simulated samples are propagated as a normalization uncertainty; 2) differences in m_T shapes between the signal and control regions due to limited statistics is incorporated as a shape uncertainty; 3) experimental and theoretical uncertainties assigned to the subtracted simulated events are propagated to the fake τ_h background, and correspond to the contribution from simulated events in the control region after all selections (anticorrelated between the fake τ_h background and the other background processes).
- τ_h identification: a 5% tau efficiency uncertainty is applied on genuine taus in the selection for the chosen loose working point (same as in the leptonic final state, see section 4.10). An additional uncertainty of $^{+5}_{-35}\%p_T$ TeV is applied on τ_h with large transverse momentum to account for possible differences arising in the extrapolation of the measured efficiencies to the high-p_T range, and is propagated as a shape uncertainty.
- τ_h energy scale: $\pm 1.2\%$ for $p_T < 400$ GeV and $\pm 3\%$ otherwise [64]. The variations of the τ_h energy scales are propagated to E_T^{miss}.

5.2 Combination Hadronic and Leptonic Final States

Both hadronic and leptonic channels are combined in a single fit to improve the upper limit cross section exclusion. In practice, the input histograms are combined where several nuisances are renamed for correlation purposes, followed by a simultaneous fit of both channels on the signal strength.

By design of the analysis, it is expected that the leptonic becomes more important at low mass. Indeed, due to the jet multiplicity constraints in the leptonic channel, significance at high mass is lost and the hadronic channel becomes fully dominant. However, even if the combination impact is negligible in this regime, the combination is still performed to have a uniform approach and the benefit of a possible combination at later stage with other charged Higgs searches (e.g. $H^{\pm} \to t\bar{b}$ in the leptonic final state).

5.2.1 Event Overlap

During the design of both analysis, care has been taken to remove any event overlap. Exact phase space orthogonality is achieved by using the mini-isolation property of the leptons:

- the hadronic final state vetoes electrons and muons with a mini-isolation < 0.15;
- the leptonic final state uses tight electrons and muons having a mini-isolation of < 0.1.

5.2.2 Processes and Correlations

The dominant background in both channels is $t\bar{t}$ production, followed by jet $\to$ fake τ_h in the hadronic channel and $W^{\pm}$+jets production in both hadronic and leptonic channels. The $t\bar{t}$ is treated differently in both channels: in the hadronic channel the POWHEG sample is used whereas in the leptonic channel the MADGRAPH5_aMC@NLO is used which has a more precise estimation of the jet kinematics. Despite this difference the ttbar are considered as one process in the fit and the theoretical nuisances (pdf, QCD scale, mass uncertainty) specifically to the $t\bar{t}$ background are fully correlated among both channels, except for the acceptance uncertainties on the cross section (pdf and QCD scale). Other backgrounds minor backgrounds (single top, $W^{\pm}$+jets, Z/γ^*+jets and diboson) use the same Monte Carlo samples and these processes are correlated in the fit as well as the assigned theoretical uncertainties. The jet $\to$ fake τ_h background is only present in the hadronic final state.

Source	Shape	$H^\pm$ (200 GeV)	jet $\rightarrow$ fake τ_h	$t\bar{t}$	top	electroweak
$\tau_h + E_T^{miss}$ trigger efficiency	✓	1.4	2.0	0.2	0.2	0.2
τ_h identification	✓	1.8	0.6	1.1	1.0	0.9
Lepton selection efficiency		2.3	–	2.7	2.7	2.7
Jet energy scale and resolution	✓	4.7	0.4	5.1	9.2	13.4
τ_h energy scale	✓	0.2	0.6	< 0.1	< 0.1	< 0.1
Unclustered E_T^{miss} energy scale	✓	2.6	< 0.1	3.2	5.2	7.2
b jet identification	✓	3.6	0.8	3.1	3.4	13.8
Integrated luminosity		2.5	0.4	2.5	2.5	2.5
Pileup	✓	1.1	< 0.1	0.8	1.2	4.0
Jets misid. as τ_h estimation	✓	–	6.1	–	–	–
Cross section (scales, PDF)		–	0.8	5.5	5.3	3.6
Top quark mass		–	0.4	2.7	2.2	–
Acceptance (scales, PDF)		5.1	0.5	2.8	2.8	6.8
$t\bar{t}$ parton showering		–	–	6.1	–	–
Total		9.4	6.6	12.1	13.5	22.7

Table 5-1: Pre-fit systematic uncertainties in %, summed over all final states and categories. For the $H^\pm$ signal, the values for $m_{H^\pm} = 200$ GeV are shown.

All the nuisance parameters in both final states have been studied and compared in detail, based on Asimov and observed datasets. The pre-fit uncertainties for the combined fit are summarized in table **5-1**. Although both final state cover a different event phase space, all the experimental uncertainties (`CMS_scale_j`, `CMS_res_j`, `CMS_eff_t`, `CMS_pileup`, `lumi_13TeV` are fully correlated in the fit. Such correlations have been studied by examining the (individual) unblinded pull distributions and validated on the observed dataset. Due to a different usage of the b tagging scale factors in the hadronic and leptonic channels, the b tagging related uncertainties are also de-correlated.

5.2.3 Post-fit Templates and Yields

To extract the post-fit distributions and yields, the background-only fit has been performed using the charged Higgs 200 GeV templates. During the fit the following processes are taken into account among both hadronic and leptonic final state: $t\bar{t}$, top (single top and $t\bar{t} + X$), $W^\pm$+jets, Z/γ^*+jets and diboson (WW, ZZ, WZ). The post-fit event yields are summarized in table **5-2** and the post-fit m_T distributions for the hadronic final state are shown in figure. **5-4**. The post-fit m_T templates for the leptonic final state are similar to the ones discussed in section 4.12.1. From these results, the fit is compatible with the SM hypotheses.

Process	τ_h+jets	$\ell + \tau_h$	ℓ+ no τ_h
jet $\rightarrow$ fake τ_h	4619 ± 35	–	–
$t\bar{t}$	1455 ± 13	30556 ± 466	174744 ± 346
top	801 ± 13	3006 ± 49	26128 ± 261
electroweak	1739 ± 18	2760 ± 37	52314 ± 223
total expected from the SM	8614 ± 42	36322 ± 500	253187 ± 398
observed	8647	36277	253236
$H^{\pm}$ signal, $m_{H^{\pm}} = 100$ GeV	20 ± 3	160 ± 20	241 ± 26
$H^{\pm}$ signal, $m_{H^{\pm}} = 200$ GeV	156 ± 22	327 ± 37	682 ± 61
$H^{\pm}$ signal, $m_{H^{\pm}} = 2000$ GeV	1627 ± 310	369 ± 24	1571 ± 99

Table 5-2: Post-fit background only event yields and corresponding uncertainties, summed over all categories in each final state. The signal yields are normalized to the $H^{\pm}$ production cross section of 1 pb and the total pre-fit uncertainties are shown.

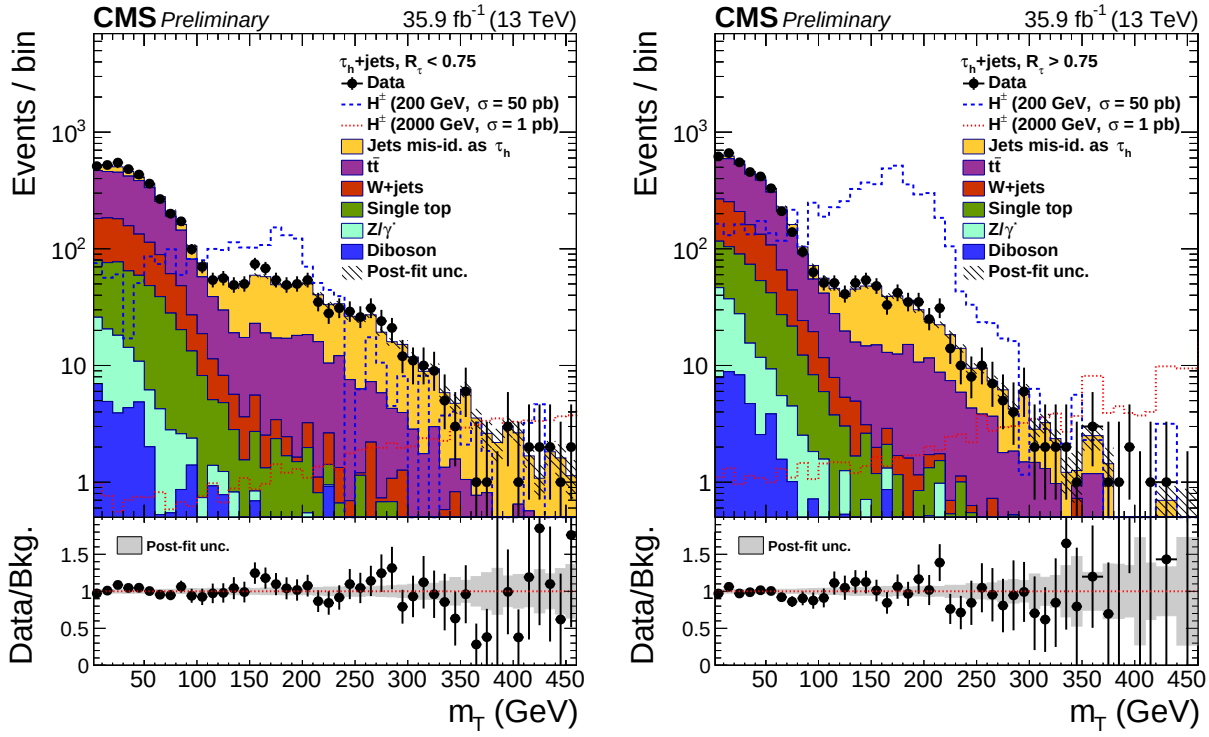


Figure **5-4**: The transverse mass distributions for the τ_h+jets channel for the two categories with $R_\tau < 0.75$ (left) and $R_\tau > 0.75$ (right) [5].

5.2.4 Expected and Observed Limits

The combined expected limits obtained with the are shown in Fig. **5-5** (left) for the full mass range from 80 GeV to 3 TeV. The masses 165, 170 and 175 GeV are computed with the LO intermediate samples without LO to NLO correction. On Fig. **5-5** (right), the upper limit ratios combined/hadronic only and combined/leptonic only are shown. From this plot it is clear that the leptonic channel is dominant in the light mass regime whereas the hadronic channel is fully dominant in the heavy mass regime. Indeed, at low mass the hadronic final state is constrained by the relative high trigger thresholds for τ_h p_T and E_T^{miss} whereas the leptonic final state benefits from a large categorization and constraints of backgrounds due to high statistics. Going to higher masses, the hadronic final state becomes fully dominant due to its clear m_T shape (no smearing as all E_T^{miss} comes from the charged Higgs) and due to the fact the leptonic final state is limited in jet multiplicity.

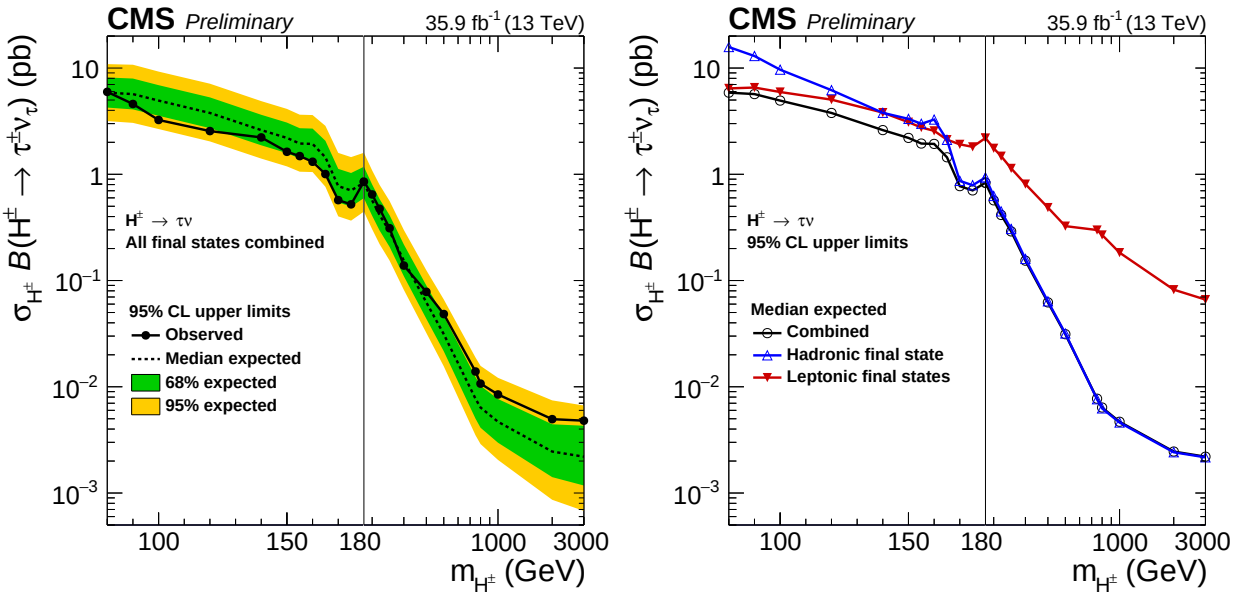


Figure **5-5**: Left: 95% CL expected upper limits on the cross section times branching fraction for charged Higgs production to $\tau\nu_\tau$ for the combined hadronic and leptonic channels. The masses 165, 170 and 175 GeV are computed with the LO intermediate samples with LO to NLO correction. Right: expected upper limit for the hadronic final state (blue), leptonic final states (red) and combined result (black) [5].

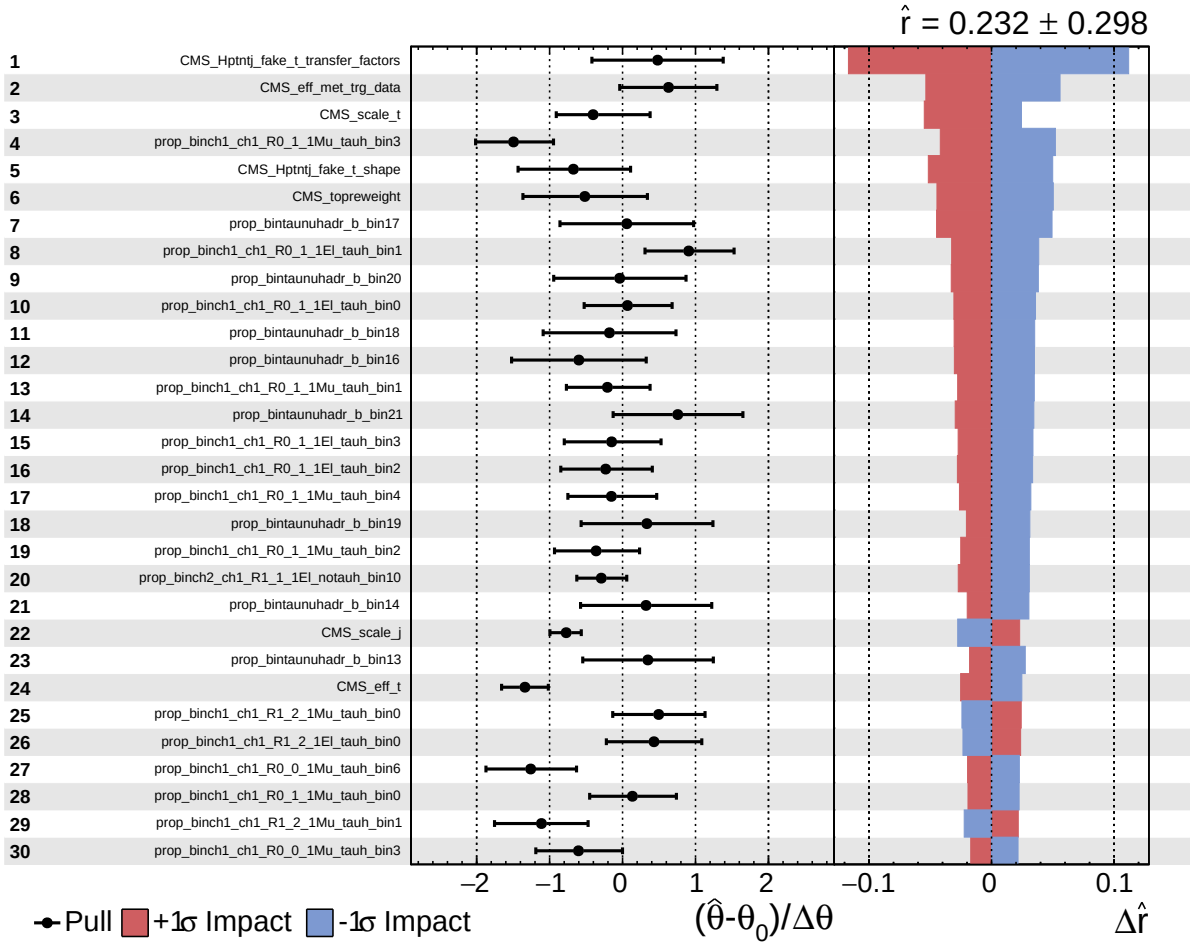


Figure **5-6**: 30 highest nuisance impacts and pull distributions for mass point $m_{H^{\pm}} = 200$ GeV based on the observed dataset for the combined hadronic and leptonic final states.

5.2.5 Fit ests

The fit tests as explained in section 4.12.3 have been used extensively to study the combination. The 30 highest impact nuisances for the combined fit are shown in figure **5-6** and the non-statistical pull distributions are shown in figure **5-7**. At a charged Higgs mass of 200 GeV, the hadronic final state is dominant and therefore it is expected the nuisances related to the jet $\rightarrow$ fake τ_h method are highly ranked, as well as the τ_h identification uncertainty. The latter is explained within the leptonic context in section 4.12.3. Other high impacting nuisances are related to sensitive bin statistics.

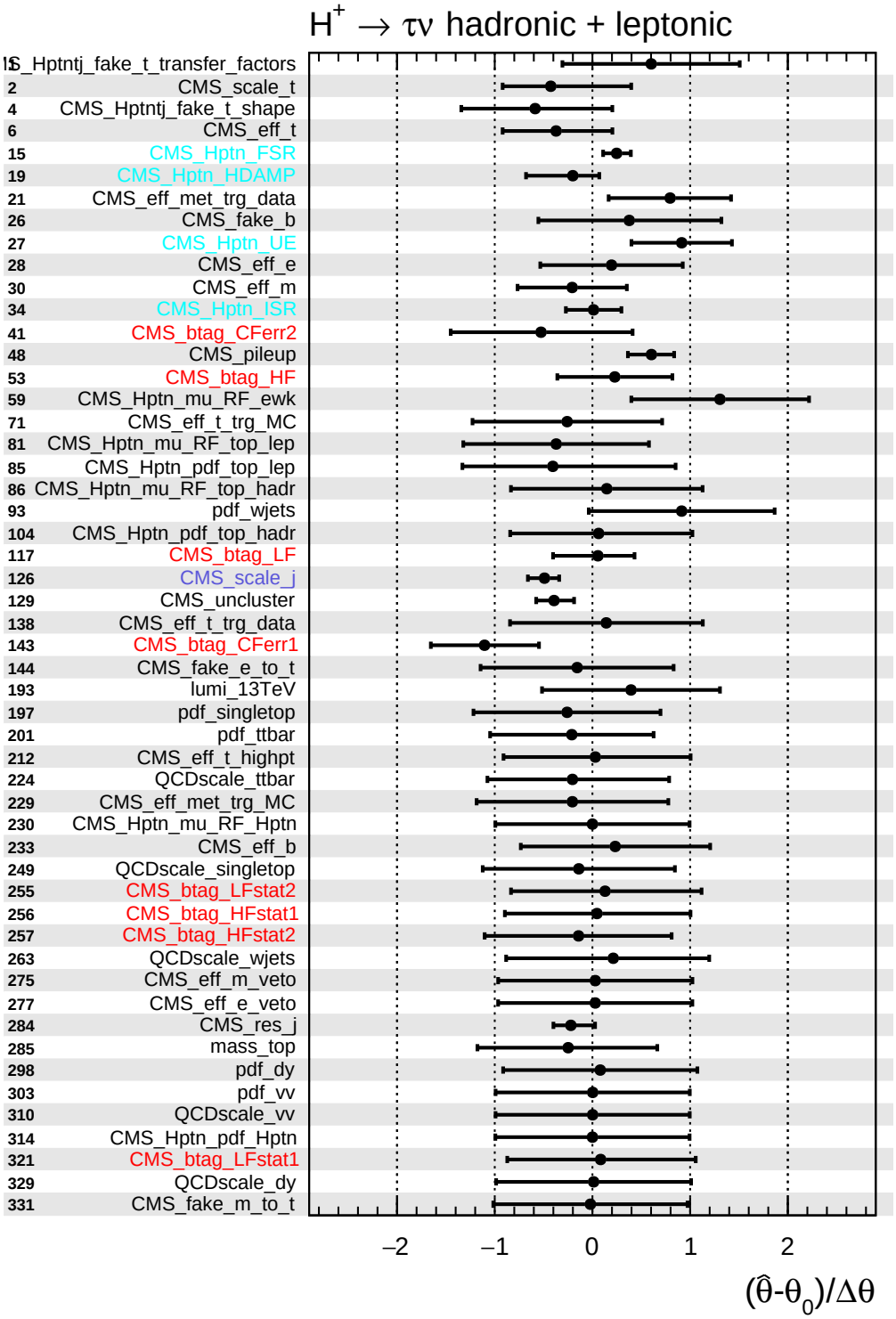


Figure **5-7**: Non-statistical pull distributions for mass point $m_{H^\pm} = 200$ GeV based on the observed dataset for the combined hadronic and leptonic final states.

5.3 Model Dependent Interpretation

The model independent upper limits are interpreted in a Minimal Supersymmetric Standard Model (MSSM) benchmark scenario. The framework of the MSSM within the 2HDM is explained in section 2.3. In such benchmark scenarios, the model independent upper limits are transformed to an exclusion of $\tan\beta$ as function of $m_{H^\pm}$. Shown in Fig. **5-8** is the exclusion of the value of $\tan\beta$ as function of $m_{H^\pm}$ for the MSSM $m_h^{\text{mod-}}$ scenario. The branching ratio of this model for several decay channels is shown in figure **2-4**, where it is clear that the $\tau\nu_\tau$ is dominant at low values of charged Higgs mass. Indeed, below $m_{H^\pm} \approx 160$ GeV, a charged Higgs is completely excluded for the entire range of $\tan\beta$. The exclusion becomes weaker at higher mass as the branching ratio of the $\tau\nu_\tau$ channel reduces whereas the tb final state takes over.

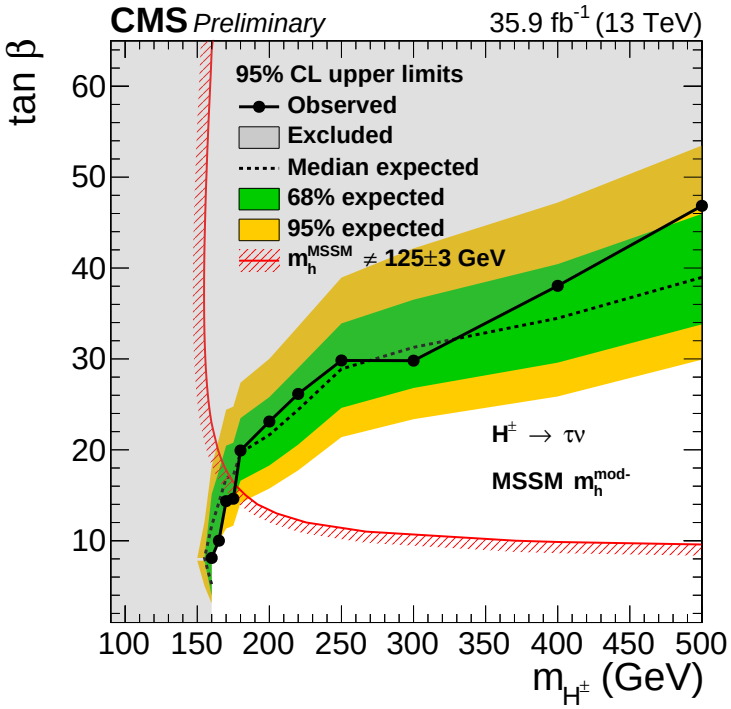


Figure **5-8**: Model dependent exclusion of $\tan\beta$ as function of $m_{H^\pm}$ for the MSSM $m_h^{\text{mod-}}$ scenario [5].

6 Conclusions and Outlook

Summary and conclusions. This thesis comprise the search of a charged Higgs boson in the $\tau\nu_\tau$ final state. Emphasis has been given on the leptonic final state, which had been analyzed independently based on the 13 TeV dataset with 35.9 fb^{-1} of integrated luminosity collected with the CMS experiment in 2016. The analysis is performed in the muon/electron+jets channel, using triggered isolated leptons, and is performed in two categories whether an hadronic tau lepton can be resolved or not. The transverse mass distribution of the lepton and $\text{E}_\text{T}^\text{miss}$ is used as a discriminant between the charged Higgs signal and background, the latter mainly composed of $t\bar{t}$ production. The event phase space is extensively categorized in number of jets, b jets and in bins of $\text{E}_\text{T}^\text{miss}$, leading to 34 regions which are used in a combined fit to the observed data, in order to constrain the backgrounds and enhance the signal sensitivity. This rather complex fit is studied in detail during the development phase of the analysis, as well as during the unblinding phase where the validation of all the nuisances was performed. In particular, due to the large amount of categories, an automatic template binning method has been implemented for a proper estimation of signal and background events in each bin. Furthermore, artificial constraints due to do ill defined shape nuisances were removed by implementing a smoothing technique based on a Gaussian kernel density. Results are presented for charged Higgs boson mass hypotheses ranging from 80 GeV to 3 TeV, where the intermediate mass range around the top quark mass is included. A procedure to reweight the leading order intermediate samples has been established in order to interpret them on the same footing as the NLO samples used in the light and heavy mass regime. The results are compatible with the Standard Model predictions and 95% CL upper limits are set on the charged Higgs boson production cross section.

A second chapter of this thesis involved the combination with the $\tau\nu_\tau$ hadronic final state, which was developed in parallel with the leptonic final state. The main challenges were the understanding and correlation of the experimental and theoretical nuisances between both leptonic and hadronic final states, such that an optimal combined fit could be performed, maximizing the signal sensitivity. Again, the results are compatible with the Standard Model and 95% CL upper limits are set on the charged Higgs boson production cross section and the model independent result is interpreted in the MSSM $\text{m}_\text{h}^\text{mod-}$ benchmark scenario.

Outlook. Several aspects of the analysis can be optimized and are in the pipeline for a global Run-II result. The first aspect concerns the m_T discriminant, which is sub-optimal as the additional neutrinos present in the leptonic final state smear the m_T distribution. Studies will be performed to switch to multivariate techniques, which can be optimized per mass regime or for each available mass point. Training against the $t\bar{t}$ background can be optimized in the several regions in the analysis. Secondly, the constraint on the number of jets will be removed which will make the leptonic final state more sensitive at higher masses. This constraint was necessary for a possible combination with the tb final state, though preliminary studies on the combination points to a weak overlap between the $\tau\nu_\tau$ and tb final states w.r.t. branching ratio. Therefore, a possible full combination will not gain w.r.t. the individual limits. However, to keep the option open, the categorization will be defined such that by removing single categories with high jet multiplicity at combination stage, a full combination will still be possible. Thirdly, it was shown that the signal acceptance drops significantly after the E_T^{miss} cut. This cut was necessary to reduce QCD multijet events to a negligible amount. Therefore, it will be examined whether one can gain from a data-driven QCD estimation, in order to be able to relax the E_T^{miss} cuts. This will benefit for the search in the light and intermediate regime. However, the penalty is the impact of the associated systematic uncertainties. A fourth aspect, which also involves the hadronic final state, is to have a more generic approach for the intermediate mass range LO to NLO correction. At the moment, the correction is defined on the event phase space, and a more generic approach based on gen-level quantities will be more appropriate.

Bibliography

[1] Vardan Khachatryan et al. Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV. *Eur. Phys. J.*, C75(5):212, 2015.

[2] Celine Degrande, Rikkert Frederix, Valentin Hirschi, Maria Ubiali, Marius Wiesemann, and Marco Zaro. Accurate predictions for charged higgs production: Closing the mh+- mt window. *Physics Letters B*, 772:87 – 92, 2017.

[3] D. de Florian et al. Handbook of LHC Higgs cross sections: 4. deciphering the nature of the Higgs sector. 2016.

[4] Tai Sakuma and Thomas McCauley. Detector and Event Visualization with SketchUp at the CMS Experiment. *J. Phys. Conf. Ser.*, 513:022032, 2014.

[5] CMS Collaboration. Search for charged Higgs bosons with the $H^{\pm} \to \tau^{\pm}\nu_{\tau}$ decay channel in proton-proton collisions at $\sqrt{s} = 13$ TeV. 2018.

[6] Particle Data Group. Review of particle physics. *Phys. Rev. D*, 98:030001, Aug 2018.

[7] Peter W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Lett.*, 12:132, 1964.

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

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

[10] T. W. B. Kibble. Symmetry breaking in non-Abelian gauge theories. *Phys. Rev.*, 155:1554, 1967.

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

[12] Georges Aad et al. Observation of a new particle in the search for the standard model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716:1, 2012.

[13] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett. B*, 716:30, 2012.

[14] Serguei Chatrchyan et al. Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV. *JHEP*, 06:081, 2013.

[15] ATLAS and CMS Collaborations. 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.*, 114:191803, 2015.

[16] CMS Collaboration. Observation of $t\bar{t}H$ production. *Phys. Rev. Lett.*, 120(CMS-HIG-17-035. CMS-HIG-17-035-003):231801. 17 p, Apr 2018. Replaced with the published version. Added the journal reference and the DOI. All the figures and tables, including additional supplementary figures and tables, can be found at http://cms-results.web.cern.ch/cms-results/public-results/publications/HIG-17-035 (CMS Public Pages).

[17] D. Giusti, V. Lubicz, G. Martinelli, F. Sanfilippo, and S. Simula. Strange and charm HVP contributions to the muon $(g-2)$ including QED corrections with twisted-mass fermions. *JHEP*, 10:157, 2017.

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

[19] A. Arbey, F. Mahmoudi, O. Stal, and T. Stefaniak. Status of the Charged Higgs Boson in Two Higgs Doublet Models. *Eur. Phys. J.*, C78(3):182, 2018.

[20] Robert Harlander, Michael Kramer, and Markus Schumacher. Bottom-quark associated Higgs-boson production: reconciling the four- and five-flavour scheme approach. 2011.

[21] Abdelhak Djouadi. The anatomy of electro-weak symmetry breaking. II. the Higgs bosons in the minimal supersymmetric model. *Phys. Rept.*, 459:1, 2008.

[22] M. Carena, S. Heinemeyer, O. Stal, C. E. M. Wagner, and G. Weiglein. MSSM Higgs Boson Searches at the LHC: Benchmark Scenarios after the Discovery of a Higgs-like Particle. *Eur. Phys. J.*, C73(9):2552, 2013.

[23] Henning Bahl, Elina Fuchs, Thomas Hahn, Sven Heinemeyer, Stefan Liebler, Shruti Patel, Pietro Slavich, Tim Stefaniak, Carlos E. M. Wagner, and Georg Weiglein. MSSM Higgs Boson Searches at the LHC: Benchmark Scenarios for Run 2 and Beyond. 2018.

[24] ALEPH, DELPHI, L3, OPAL and LEP Collaborations. Search for charged Higgs bosons: combined results using LEP data. *Eur. Phys. J. C*, 73:2463, 2013.

[25] T. Aaltonen et al. Search for Higgs bosons predicted in two-Higgs-doublet models via decays to tau lepton pairs in 1.96 TeV p anti-p collisions. *Phys. Rev. Lett.*, 103:201801, 2009.

[26] V. M. Abazov et al. Search for Higgs bosons of the minimal supersymmetric standard model in $p\overline{p}$ collisions at $\sqrt{s} = 1.96$ TeV. *Phys. Lett. B*, 710:569, 2012.

[27] Georges Aad et al. Search for charged Higgs bosons decaying via $H^+ \to \tau\nu$ in top quark pair events using pp collision data at $\sqrt{s} = 7$ TeV with the ATLAS detector. *JHEP*, 06:039, 2012.

[28] Serguei Chatrchyan et al. Search for a light charged Higgs boson in top quark decays in pp collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 07:143, 2012.

[29] Georges Aad et al. Search for charged Higgs bosons through the violation of lepton universality in $t\bar{t}$ events using pp collision data at $\sqrt{s} = 7$ TeV with the ATLAS experiment. *JHEP*, 03:076, 2013.

[30] Georges Aad et al. Search for charged Higgs bosons decaying via $H^\pm \to \tau^\pm\nu$ in fully hadronic final states using pp collision data at $\sqrt{s} = 8$ TeV with the ATLAS detector. *JHEP*, 03:088, 2015.

[31] Vardan Khachatryan et al. Search for a charged Higgs boson in pp collisions at $\sqrt{s} = 8$ TeV. *JHEP*, 11:018, 2015.

[32] Morad Aaboud et al. Search for charged Higgs bosons produced in association with a top quark and decaying via $H^\pm \to \tau\nu$ using pp collision data recorded at $\sqrt{s} = 13$ TeV by the ATLAS detector. *Phys. Lett. B*, 759:555, 2016.

[33] Morad Aaboud et al. Search for charged Higgs bosons decaying via $H^\pm \to \tau^\pm\nu_\tau$ in the τ+jets and τ+lepton final states with 36 fb^{-1} of pp collision data recorded at $\sqrt{s} = 13$ TeV with the ATLAS experiment. *JHEP*, 09:139, 2018.

[34] Georges Aad et al. Search for charged Higgs bosons in the $H^\pm \to tb$ decay channel in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector. *JHEP*, 03:127, 2016.

[35] Georges Aad et al. Search for a light charged Higgs boson in the decay channel $H^+ \to c\bar{s}$ in $t\bar{t}$ events using pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *Eur. Phys. J. C*, 73:2465, 2013.

[36] Vardan Khachatryan et al. Search for a light charged Higgs boson decaying to $c\bar{s}$ in pp collisions at $\sqrt{s} = 8$ TeV. *JHEP*, 12:178, 2015.

[37] Albert M Sirunyan et al. Search for a charged Higgs boson decaying to charm and bottom quarks in proton-proton collisions at $\sqrt{s} = 8$ TeV. Submitted to *JHEP*, 2018.

[38] Morad Aaboud et al. Search for additional heavy neutral Higgs and gauge bosons in the ditau final state produced in 36 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. *JHEP*, 01:055, 2018.

[39] Albert M Sirunyan et al. Search for additional neutral MSSM Higgs bosons in the $\tau\tau$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 09:007, 2018.

[40] Albert M Sirunyan et al. Search for beyond the standard model Higgs bosons decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV. *JHEP*, 08:113, 2018.

[41] Oliver Sim BrÃ$\frac{1}{4}$ning, Paul Collier, P Lebrun, Stephen Myers, Ranko Ostojic, John Poole, and Paul Proudlock. *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004.

[42] Werner Herr and B Muratori. Concept of luminosity. 2006.

[43] CMS Collaboration. The CMS experiment at the CERN LHC. The Compact Muon Solenoid experiment. *JINST*, 3:S08004. 361 p, 2008. Also published by CERN Geneva in 2010.

[44] *The CMS magnet project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.

[45] V Karimaki, M Mannelli, P Siegrist, H Breuker, A Caner, R Castaldi, K Freudenreich, G Hall, R Horisberger, M Huhtinen, and A Cattai. *The CMS tracker system project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.

[46] CMS Collaboration. CMS Technical Design Report for the Pixel Detector Upgrade. Technical Report CERN-LHCC-2012-016. CMS-TDR-11, Sep 2012. Additional contacts: Jeffrey Spalding, Fermilab, Jeffrey.Spalding@cern.ch Didier Contardo, Universite Claude Bernard-Lyon I, didier.claude.contardo@cern.ch.

[47] *The CMS electromagnetic calorimeter project: Technical Design Report.* Technical Design Report CMS. CERN, Geneva, 1997.

[48] *The CMS hadron calorimeter project: Technical Design Report.* Technical Design Report CMS. CERN, Geneva, 1997. The following files are from ¡a href=.

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

[50] CMS Collaboration. *CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems.* Technical Design Report CMS.

[51] Sergio Cittolin, Attila RÃ¡cz, and Paris Sphicas. *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project.* Technical Design Report CMS. CERN, Geneva, 2002.

[52] A. M. Sirunyan et al. Particle-flow reconstruction and global event description with the CMS detector. *JINST*, 12:P10003, 2017.

[53] Wolfgang Adam, Boris Mangano, Thomas Speer, and Teddy Todorov. Track Reconstruction in the CMS tracker. Technical Report CMS-NOTE-2006-041, CERN, Geneva, Dec 2006.

[54] Kenneth Rose. Deterministic annealing for clustering, compression, classification, regression, and related optimization problems. *Proceedings of the IEEE*, 86:2210, 1998.

[55] A. M. Sirunyan et al. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV. *JINST*, 13(06):P06015, 2018.

[56] Vardan Khachatryan et al. Performance of electron reconstruction and selection with the CMS detector in proton-proton collisions at $\sqrt{s} = 8$ TeV. *JINST*, 10:P06005, 2015.

[57] Andreas Hoecker, Peter Speckmayer, Joerg Stelzer, Jan Therhaag, Eckhard von Toerne, and Helge Voss. TMVA: Toolkit for multivariate data analysis. *PoS*, ACAT:040, 2007.

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

[59] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. FastJet user manual. *Eur. Phys. J. C*, 72:1896, 2012.

[60] Pileup Removal Algorithms. Technical Report CMS-PAS-JME-14-001, CERN, Geneva, 2014.

[61] Vardan Khachatryan et al. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *JINST*, 12:P02014, 2017.

[62] CMS Collaboration. Jet algorithms performance in 13 TeV data. CMS Physics Analysis Summary CMS-PAS-JME-16-003, 2017.

[63] Vardan Khachatryan et al. Reconstruction and identification of τ lepton decays to hadrons and ν_τ at CMS. *JINST*, 11:P01019, 2016.

[64] Albert M Sirunyan et al. Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s} = 13$ TeV. *JINST*, 13:P10005, 2018.

[65] Serguei Chatrchyan et al. Identification of b-quark jets with the CMS experiment. *JINST*, 8:P04013, 2013.

[66] A. M. Sirunyan et al. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. *JINST*, 13:P05011, 2018.

[67] Performance of b tagging algorithms in proton-proton collisions at 13 TeV with Phase 1 CMS detector. Jun 2018.

[68] CMS Collaboration. Performance of missing transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV using the CMS detector. CMS Physics Analysis Summary CMS-PAS-JME-17-001, 2018.

[69] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *JHEP*, 07:079, 2014.

[70] Paolo Nason. A new method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 11:040, 2004.

[71] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with parton shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.

[72] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010.

[73] Tomáš Ježo, Jonas M. Lindert, Paolo Nason, Carlo Oleari, and Stefano Pozzorini. An NLO+PS generator for $t\bar{t}$ and Wt production and decay including non-resonant and interference effects. *Eur. Phys. J. C*, 76:691, 2016.

[74] Stefano Frixione, Paolo Nason, and Giovanni Ridolfi. A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction. *JHEP*, 09:126, 2007.

[75] Torbjörn Sjöstrand, Stefan Ask, Jesper R. Christiansen, Richard Corke, Nishita Desai, Philip Ilten, Stephen Mrenna, Stefan Prestel, Christine O. Rasmussen, and Peter Z. Skands. An introduction to PYTHIA 8.2. *Comput. Phys. Commun.*, 191:159, 2015.

[76] S. Agostinelli et al. GEANT4: A simulation toolkit. *Nucl. Instrum. Meth. A*, 506:250, 2003.

[77] Vardan Khachatryan et al. Event generator tunes obtained from underlying event and multiparton scattering measurements. *Eur. Phys. J. C*, 76:155, 2016.

[78] CMS Collaboration. Investigations of the impact of the parton shower tuning in PYTHIA 8 in the modelling of $t\bar{t}$ at $\sqrt{s} = 8$ and 13 TeV. CMS Physics Analysis Summary CMS-PAS-TOP-16-021, 2016.

[79] Richard D. Ball, Valerio Bertone, Francesco Cerutti, Luigi Del Debbio, Stefano Forte, Alberto Guffanti, Jose I. Latorre, Juan Rojo, and Maria Ubiali. Unbiased global determination of parton distributions and their uncertainties at NNLO and at LO. *Nucl. Phys. B*, 855:153, 2012.

[80] Rikkert Frederix and Stefano Frixione. Merging meets matching in MC@NLO. *JHEP*, 12:061, 2012.

[81] J. Alwall, S. Hoche, F. Krauss, N. Lavesson, L. Lonnblad, F. Maltoni, M.L. Mangano, M. Moretti, C.G. Papadopoulos, F. Piccinini, S. Schumann, M. Treccani, J. Winter, and M. Worek. Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *Eur. Phys. J. C*, 53:473, 2008.

[82] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. NLO single-top production matched with shower in POWHEG: s- and t-channel contributions.

[83] Emanuele Re. Single-top Wt-channel production matched with parton showers using the POWHEG method. *Eur. Phys. J. C*, 71:1547, 2011.

[84] Michal Czakon and Alexander Mitov. Top++: A program for the calculation of the top-pair cross-section at hadron colliders. *Comput. Phys. Commun.*, 185:2930, 2014.

[85] Nikolaos Kidonakis. Top quark production. In *Proceedings, Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013)*, page 139, 2014.

[86] P. Kant, O. M. Kind, T. Kintscher, T. Lohse, T. Martini, S. Mölbitz, P. Rieck, and P. Uwer. HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions. *Comput. Phys. Commun.*, 191:74, 2015.

[87] Kirill Melnikov and Frank Petriello. Electroweak gauge boson production at hadron colliders through $O(\alpha_s^2)$. *Phys. Rev. D*, 74:114017, 2006.

[88] John M. Campbell, R. Keith Ellis, and Ciaran Williams. Vector boson pair production at the LHC. *JHEP*, 07:018, 2011.

[89] Celine Degrande, Rikkert Frederix, Valentin Hirschi, Maria Ubiali, Marius Wiesemann, and Marco Zaro. Accurate predictions for charged Higgs production: Closing the $m_{\mathrm{h}^\pm} \sim m_{\mathrm{t}}$ window. *Phys. Lett. B*, 772:87, 2017.

[90] F and James. MINUIT: Function Minimization and Error Analysis Reference Manual. 1998. CERN Program Library Long Writeups.

[91] J.S. Conway. Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra. (arXiv:1103.0354):115–120. 6 p, Mar 2011. Comments: Presented at PHYSTAT 2011, CERN, Geneva, Switzerland, January 2011, to be published in a CERN Yellow Report.

[92] Roger J. Barlow and Christine Beeston. Fitting using finite Monte Carlo samples. *Comput. Phys. Commun.*, 77:219, 1993.

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

[94] Alexander L. Read. Presentation of search results: The $\mathrm{CL_s}$ technique. *J. Phys. G*, 28:2693, 2002.

[95] Abraham Wald. Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, 54(3):426–482, 1943.

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

[97] CMS Collaboration. CMS luminosity measurements for the 2016 data taking period. CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.

[98] M. Aaboud et al. Measurement of the inelastic proton-proton cross section at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC. *Phys. Rev. Lett.*, 117:182002, 2016.

[99] Jon Butterworth et al. PDF4LHC recommendations for LHC Run II. *J. Phys. G*, 43:023001, 2016.

[100] Peter Skands, Stefano Carrazza, and Juan Rojo. Tuning PYTHIA 8.1: the Monash 2013 tune. *Eur. Phys. J. C*, 74:3024, 2014.

[101] D. P. Roy. The hadronic tau decay signature of a heavy charged Higgs boson at LHC. *Phys. Lett. B*, 459:607, 1999.