SEARCH FOR PRODUCTION OF A HIGGS BOSON AND A SINGLE TOP QUARK IN MULTILEPTON FINAL STATES IN pp COLLISIONS AT $\sqrt{s} = 13$ TeV.

by

Jose Andres Monroy Montañez

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfilment of Requirements

For the Degree of Doctor of Philosophy

Major: Physics and Astronomy

Under the Supervision of Kenneth Bloom and Aaron Dominguez

Lincoln, Nebraska

July, 2018

SEARCH FOR PRODUCTION OF A HIGGS BOSON AND A SINGLE TOP QUARK IN MULTILEPTON FINAL STATES IN pp COLLISIONS AT $\sqrt{s} = 13$ TeV.

Jose Andres Monroy Montañez, Ph.D.

University of Nebraska, 2018

Adviser: Kenneth Bloom and Aaron Dominguez

Table of Contents

List of Figures

List of Tables

Chapter 1

Theoretical approach

1.1 Introduction

The physical description of the universe is a challenge that physicists have faced by making theories that refine existing principles and proposing new ones in an attempt to embrace emerging facts and phenomena.

At the end of 1940s Julian Schwinger [1] and Richard P. Feynman [2], based on the work of Sin-Itiro Tomonaga [3], developed an electromagnetic theory consistent with special relativity and quantum mechanics that describes how matter and light interact; the so-called *quantum electrodynamics* (QED) was born.

QED has become the blueprint for developing theories that describe the universe. It was the first example of a quantum field theory (QFT), which is the theoretical framework for building quantum mechanical models that describes particles and their interactions. QFT is composed of a set of mathematical tools that combines classical fields, special relativity and quantum mechanics, while keeping the quantum point particles and locality ideas.

This chapter gives an overview of the standard model of particle physics, starting with a description of the particles and their interactions, followed by a description of the electroweak interaction, the Higgs boson and the associated production of Higgs boson and a single top quark (tH). The description contained in this chapter is based on References [4–6].

1.2 Standard model of particle physics

The *standard model of particle physics (SM)* describes particle physics at the fundamental level in terms of a collection of interacting particles and fields. The full picture of the SM is composed of three fields[1] whose excitations are interpreted as particles called mediators or force-carriers, a set of fields whose excitations are interpreted as elementary particles interacting through the exchange of those mediators, and a field that gives the mass to elementary particles. Figure 1.1 shows a scheme of the SM particles' organization. In addition, for each of the particles in the scheme there exits an antiparticle with the same mass and opposite quantum numbers. The existence of antiparticles is a prediction of the relativistic quantum mechanics from the solution of the Dirac equation for which a negative energy solution is also possible. In some cases a particle is its own anti-particle, like photon or Higgs boson.

The mathematical formulation of the SM is based on group theory and the use of Noether's theorem [8] which states that for a physical system modeled by a Lagrangian that is invariant under a group of transformations a conservation law is expected. For instance, a system described by a time-independent Lagrangian is invariant (symmetric) under time changes (transformations) with the total energy conservation law as the expected conservation law. In QED, the charge operator (Q) is the generator of

[1] The formal and complete treatment of the SM is out of the scope of this document, however a plenty of textbooks describing it at several levels are available in the literature. The treatment in References [5, 6] is quite comprehensive and detailed. Note that gravitational field is not included in the standard model formulation

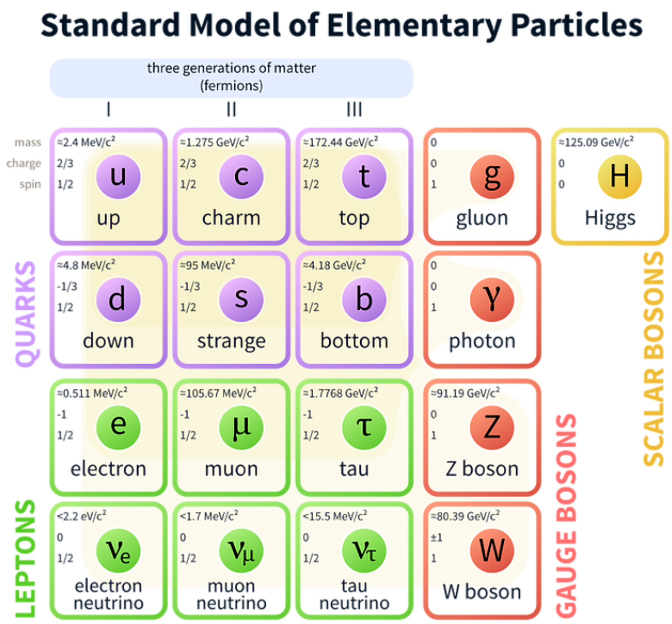

Figure 1.1: Schematic representation of the Standard Model of particle physics. The SM is a theoretical model intended to describe three of the four fundamental forces of the universe in terms of a set of particles and their interactions. [7].

the U(1) symmetry which according to the Noether's theorem means that there is a conserved quantity; this conserved quantity is the electric charge and thus the law conservation of electric charge is established.

In the SM, the symmetry group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ describes three of the four fundamental interactions in nature (see Section 1.2.2): strong interaction (SI), weak interaction (WI) and electromagnetic interactions (EI) in terms of symmetries associated to physical quantities:

- Strong: $SU(3)_C$ associated to color charge
- Weak: $SU(2)_L$ associated to weak isospin and chirality
- Electromagnetic: $U(1)_Y$ associated to weak hypercharge and electric charge

It will be shown that the electromagnetic and weak interactions are combined in the so-called electroweak interaction where chirality, hypercharge, weak isospin and electric charge are the central concepts.

1.2.1 Fermions

The basic constituents of the ordinary matter at the lowest level, which form the set of elementary particles in the SM formulation, are quarks and leptons. All of them have spin 1/2, therefore they are classified as fermions since they obey Fermi-Dirac statistics. There are six *flavors* of quarks and three of leptons organized in three generations, or families, as shown in Table 1.1.

		Generation		
	Type	1st	2nd	3rd
Leptons	Charged	Electron (e)	Moun(μ)	Tau (τ)
	Neutral	Electron neutrino (ν_e)	Muon neutrino (ν_μ)	Tau neutrino (ν_τ)
Quarks	Up-type	Up (u)	Charm (c)	Top (t)
	Down-type	Down (d)	Strange (s)	Bottom (b)

Table 1.1: Fermions of the SM. There are six flavors of quarks and three of leptons, organized in three generations, or families, composed of two pairs of closely related particles. The close relationship is motivated by the fact that each pair of particles is a member of an $SU(2)_L$ doublet that has an associated invariance under isospin transformations. WI between leptons is limited to the members of the same generation; WI between quarks is not limited but greatly favored, to same generation members.

There is a mass hierarchy between generations (see Table 1.2), where the higher generation particles decays to the lower one, which can explain why the ordinary matter is made of particles from the first generation. In the SM, neutrinos are modeled as massless particles so they are not subject to this mass hierarchy; however, today it is known that neutrinos are massive so the hierarchy could be restated. The reason behind this mass hierarchy is one of the most important open questions in particle physics, and it becomes more puzzling when noticing that the mass difference between first and second generation fermions is small compared to the mass difference with respect to the third generation.

Lepton	Mass (MeV/c^2)	Quark	Mass (MeV/c^2)
e	0.51	u	2.2
μ	105.65	c	1.28×10^3
τ	1776.86	t	173.1×10^3
ν_e	Unknown	d	4.7
ν_μ	Unknown	s	96
τ_μ	Unknown	b	4.18×10^3

Table 1.2: Fermion masses [9]. Generations differ by mass in a way that has been interpreted as a mass hierarchy. Approximate values with no uncertainties are used, for comparison purpose.

Usually, the second and third generation fermions are produced in high energy processes, like the ones recreated in particle accelerators.

1.2.1.1 Leptons

A lepton is an elementary particle that is not subject to the SI. As seen in Table 1.1, there are two types of leptons, the charged ones (electron, muon and tau) and the neutral ones (the three neutrinos). The electric charge (Q) is the property that gives leptons the ability to participate in the EI. From the classical point of view, Q plays a central role determining, among others, the strength of the electric field through which the electromagnetic force is exerted. It is clear that neutrinos are not affected by EI because they don't carry electric charge.

Another feature of the leptons that is fundamental in the mathematical description of the SM is the chirality, which is closely related to spin and helicity. Helicity defines the handedness of a particle by relating its spin and momentum such that if they are parallel then the particle is right-handed; if spin and momentum are antiparallel the particle is said to be left-handed. The study of parity conservation (or violation) in β-decay has shown that only left-handed electrons/neutrinos or right-handed positrons/anti-neutrinos are created [10]; the inclusion of that feature in the theory was achieved by using projection operators for helicity, however, helicity is frame dependent for massive particles which makes it not Lorentz invariant and then another related attribute has to be used: *chirality*.

Chirality is a purely quantum attribute which makes it not so easy to describe in graphical terms but it defines how the wave function of a particle transforms under certain rotations. As with helicity, there are two chiral states, left-handed chiral (L) and right-handed chiral (R). In the highly relativistic limit where $E \approx p \gg m$ helicity and chirality converge, becoming exactly the same for massless particles.

In the following, when referring to left-handed (right-handed) it will mean left-handed chiral (right-handed chiral). The fundamental fact about chirality is that while EI and SI are not sensitive to chirality, in WI left-handed and right-handed fermions are treated asymmetrically, such that only left-handed fermions and right-handed anti-fermions are allowed to couple to WI mediators, which is a violation of parity. The way to translate this statement in a formal mathematical formulation is based on the isospin symmetry group $SU(2)_L$.

Each generation of leptons is seen as a weak isospin doublet.[2] The left-handed charged lepton and its associated left-handed neutrino are arranged in doublets of weak isospin T=1/2 while their right-handed partners are singlets:

$$\begin{pmatrix} \nu_l \\ l \end{pmatrix}_L, l_R := \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, e_R, \mu_R, \tau_R, \nu_{eR}, \nu_{\mu R}, \nu_{\tau R} \tag{1.1}$$

The isospin third component refers to the eigenvalues of the weak isospin operator

[2] The weak isospin is an analogy of the isospin symmetry in strong interaction where neutron and proton are affected equally by strong force but differ in their charge.

which for doublets is $T_3 = \pm 1/2$, while for singlets it is $T_3 = 0$. The physical meaning of this doublet-singlet arrangement falls in that the WI couples the two particles in the doublet by exchanging the interaction mediator while the singlet member is not involved in WI. The main properties of the leptons are summarized in Table 1.3.

Although all three flavor neutrinos have been observed, their masses remain unknown and only some estimations have been made [11]. The main reason is that the flavor eigenstates are not the same as the mass eigenstates which implies that when a neutrino is created its mass state is a linear combination of the three mass eigenstates and experiments can only probe the squared difference of the masses. The Pontecorvo-Maki-Nakagawa-Sakata (PMNS) mixing matrix encodes the relationship between flavor and mass eigenstates.

Lepton	Q(e)	T_3	L_e	L_μ	L_τ	Lifetime (s)
Electron (e)	-1	-1/2	1	0	0	Stable
Electron neutrino(ν_e)	0	1/2	1	0	0	Unknown
Muon (μ)	-1	-1/2	0	1	0	2.19×10^{-6}
Muon neutrino (ν_μ)	0	1/2	0	1	0	Unknown
Tau (τ)	-1	-1/2	0	0	1	290.3×10^{-15}
Tau neutrino (τ_μ)	0	1/2	0	0	1	Unknown

Table 1.3: Lepton properties [9]. Q: electric charge, T_3: weak isospin. Only left-handed leptons and right-handed anti-leptons participate in the WI. Anti-particles with inverted T_3, Q and lepton number complete the leptons set but are not listed. Right-handed leptons and left-handed anti-leptons, neither listed, form weak isospin singlets with $T_3 = 0$ and do not take part in the weak interaction.

1.2.1.2 Quarks

Quarks are the basic constituents of protons and neutrons. The way quarks join to form bound states, called *hadrons*, is through the SI. Quarks are affected by all the fundamental interactions which means that they carry all the four types of charges: color, electric charge, weak isospin and mass.

Flavor	Q(e)	I_3	T_3	B	C	S	T	B'	Y	Color
Up (u)	2/3	1/2	1/2	1/3	0	0	0	0	1/3	r,b,g
Charm (c)	2/3	0	1/2	1/3	1	0	0	0	4/3	r,b,g
Top(t)	2/3	0	1/2	1/3	0	0	1	0	4/3	r,b,g
Down(d)	-1/3	-1/2	-1/2	1/3	0	0	0	0	1/3	r,b,g
Strange(s)	-1/3	0	-1/2	1/3	0	-1	0	0	-2/3	r,b,g
Bottom(b)	-1/3	0	-1/2	1/3	0	0	0	-1	-2/3	r,b,g

Table 1.4: Quark properties [9]. Q: electric charge, I_3: isospin, T_3: weak isospin, B: baryon number, C: charmness, S: strangeness, T: topness, B': bottomness, Y: hypercharge. Anti-quarks posses the same mass and spin as quarks but all charges (color, flavor numbers) have opposite sign.

Table 1.4 summarizes the features of quarks, among which the most remarkable is their fractional electric charge. Note that fractional charge is not a problem, given that quarks are not found isolated, but serves to explain how composed particles are formed out of two or more valence quarks[3].

Color charge is responsible for the SI between quarks and is the symmetry ($SU(3)_C$) that defines the formalism to describe SI. There are three colors: red (r), blue (b) and green (g) and their corresponding three anti-colors; thus each quark carries one color unit while anti-quarks carries one anti-color unit. As explained in Section 1.2.2, quarks are not allowed to be isolated due to the color confinement effect, hence, their features have been studied indirectly by observing their bound states created when

- one quark with a color charge is attracted by an anti-quark with the corresponding anti-color charge forming a colorless particle called a *meson.*
- three quarks (anti-quarks) with different color (anti-color) charges are attracted among them forming a colorless particle called a *baryon (anti-baryon).*

[3] Hadrons can contain an indefinite number of virtual quarks and gluons, known as the quark and gluon sea, but only the valence quarks determine hadrons' quantum numbers.

In practice, when a quark is left alone isolated a process called *hadronization* occurs where the quark emits gluons (see Section 1.2.4) which eventually will generate new quark-antiquark pairs and so on; those quarks will recombine to form hadrons that will decay into leptons. This proliferation of particles looks like a *jet* coming from the isolated quark. More details about the hadronization process and jet structure will be given in chapter3.

In the first version of the quark model (1964), M. Gell-Mann [12] and G. Zweig [13, 14] developed a consistent way to classify hadrons according to their properties. Only three quarks (u, d, s) were involved in a scheme in which all baryons have baryon number B=1 and therefore quarks have B=1/3; non-baryons have B=0. Baryon number is conserved in SI and EI which means that single quarks cannot be created but in pairs $q-\bar{q}$.

The scheme organizes baryons in a two-dimensional space (I_3 - Y); Y (hypercharge) and I_3 (isospin) are quantum numbers related by the Gell-Mann-Nishijima formula [15, 16]:

$$Q = I_3 + \frac{Y}{2} \tag{1.2}$$

where $Y = B + S + C + T + B'$ are the quantum numbers listed in Table 1.4.

There are six quark flavors organized in three generations (see Table 1.1) following a mass hierarchy which, again, implies that higher generations decay to first generation quarks.

	Quarks			T_3	Y_W	Leptons			T_3	Y_W
Doublets	$\binom{u}{d'}_L$	$\binom{c}{s'}_L$	$\binom{t}{b'}_L$	$\binom{1/2}{-1/2}$	1/3	$\binom{\nu_e}{e}_L$	$\binom{\nu_\mu}{\mu}_L$	$\binom{\nu_\tau}{\tau}_L$	$\binom{1/2}{-1/2}$	-1
Singlets	u_R	c_R	t_R	0	4/3	ν_{eR}	$\nu_{\mu R}$	$\nu_{\tau R}$		
	d'_R	s'_R	b'_R	0	-2/3	e_R	μ_R	τ_R	0	-2

Table 1.5: Fermion weak isospin and weak hypercharge multiplets. Weak hypercharge is calculated through the Gell-Mann-Nishijima formula 1.2 but using the weak isospin and charge for quarks.

Isospin doublets of quarks are also defined (see Table 1.5), and same as for neutrinos, the WI eigenstates are not the same as the mass eigenstates which means that members of different quark generations are connected by the WI mediator; thus, up-type quarks are coupled not to down-type quarks (the mass eigenstates) directly but to a superposition of down-type quarks ($q'_d; the weak eigenstates$) via WI according to:

$$q'_d = V_{CKM}\, q_d$$
$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.3}$$

where V_{CKM} is known as Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix [17,18] given by

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{pmatrix}. \tag{1.4}$$

The weak decays of quarks are represented in the diagram of Figure 1.2; again the CKM matrix plays a central role since it contains the probabilities for the different quark decay channels, in particular, note that quark decays are greatly favored between generation members.

CKM matrix is a 3×3 unitary matrix parametrized by three mixing angles and the *CP-mixing phase*; the latter is the parameter responsible for the Charge-Parity

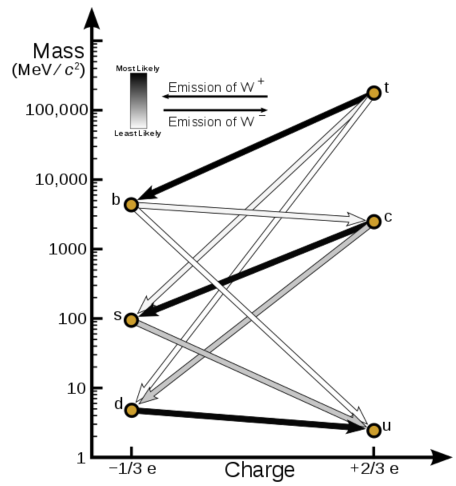

Figure 1.2: Transformations between quarks through the exchange of a WI. Higher generations quarks decay to first generation quarks by emitting a W boson. The arrow color indicates the likelihood of the transition according to the grey scale in the top left side which represent the CKM matrix parameters [19].

symmetry violation (CP-violation) in the SM. The fact that the top quark decays almost all the time to a bottom quark is exploited in this thesis when making the selection of the signal events by requiring the presence of a jet tagged as a jet coming from a b quark in the final state.

1.2.2 Fundamental interactions

Even though there are many manifestations of force in nature, like the ones represented in Figure 1.3, we can classify all of them in four fundamental interactions:

- *Electromagnetic interaction (EI)* affects particles that are *electrically charged,* like electrons and protons. Figure 1.4a. shows a graphical representation, known as *Feynman diagram*, of electron-electron scattering.

- *Strong interaction (SI)* described by Quantum Chromodynamics (QCD). Hadrons like the proton and the neutron have internal structure given that they are com-

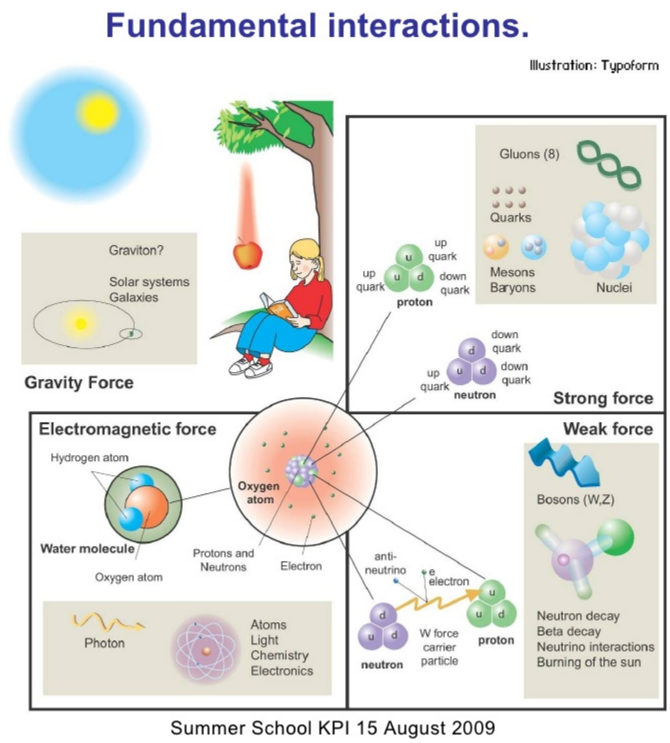

Figure 1.3: Fundamental interactions in nature. Despite the many manifestations of forces in nature, we can track all of them back to one of the fundamental interactions. The most common forces are gravity and electromagnetic given that all of us are subject and experience them in everyday life.

posed of two or more valence quarks[4]. Quarks have fractional electric charge which means that they are subject to electromagnetic interaction and in the case of the proton they should break apart due to electrostatic repulsion; however, quarks are held together inside the hadrons against their electrostatic repulsion by the *Strong Force* through the exchange of *gluons.* The analog to the electric charge is the *color charge.* Electrons and photons are elementary particles as quarks but they don't carry color charge, therefore they are not subject to SI. A Feynman diagram for gluon exchange between quarks is shown in Figure 1.4b.

- *Weak interaction (WI)* described by the weak theory (WT), is responsible, for instance, for the radioactive decay in atoms and the deuterium production

[4] Particles made of four and five quarks are exotic states not so common.

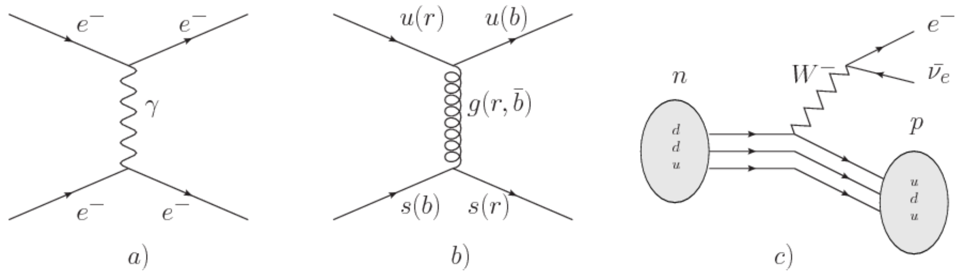

Figure 1.4: Feynman diagrams representing the interactions in SM; a) EI: e-e scattering; b) SI: gluon exchange between quarks ; c) WI: β-decay

within the sun. Quarks and leptons are the particles affected by the weak interaction; they possess a property called *flavor charge* (see 1.2.1) which can be changed by emitting or absorbing one weak force mediator. There are three mediators of the *weak force* known as Z boson in the case of electrically neutral flavor changes and $W^{\pm}$ bosons in the case of electrically charged flavor changes. The *weak isospin* is the WI analog to electric charge in EI, and color charge in SI, and defines how quarks and leptons are affected by the weak force. Figure 1.4c. shows the Feynman diagram of β-decay where a neutron (n) is transformed in a proton (p) by emitting a W^- particle.

- *Gravitational interaction (GI)* described by General Theory of Relativity (GR). It is responsible for the structure of galaxies and black holes as well as the expansion of the universe. As a classical theory, in the sense that it can be formulated without even appeal to the concept of quantization, it implies that the space-time is a continuum and predictions can be made without limitation to the precision of the measurement tools. The latter represents a direct contradiction of the quantum mechanics principles. Gravity is deterministic while quantum mechanics is probabilistic; despite that, efforts to develop a quantum theory of gravity have predicted the *graviton* as mediator of the gravitational force[5].

Interaction	Acts on	Relative strength	Range (m)	Mediators
Electromagnetic (QED)	Electrically charged particles	10^{-2}	Infinite	Photon
Strong (QCD)	Quarks and gluons	1	10^{-15}	Gluon
Weak (WI)	Leptons and quarks	10^{-6}	10^{-18}	$W^{\pm}$, Z
Gravitational (GI)	Massive particles	10^{-39}	Infinite	Graviton

Table 1.6: Fundamental interactions features [20].

Table 1.6 summarizes the main features of the fundamental interactions. The strength of the interactions is represented by the coupling constants which depend on the energy scale at which the interaction is evaluated, therefore, it is the relative strength of the fundamental forces that reveals the meaning of strong and weak; in a context where the relative strength of the SI is 1, the EI is about hundred times weaker and WI is about million times weaker than the SI. A good description on how the relative strength and range of the fundamental interactions are calculated can be found in References [20, 21]. In the everyday life, only EI and GI are explicitly experienced due to the range of these interactions; i.e., at the human scale distances only EI and GI have appreciable effects, in contrast to SI which at distances greater than 10^{-15}m become negligible. Is it important to clarify that the weakness of the WI is attributed to the fact that its mediators are highly massive which affects the propagators of the interaction, as a result, the effect of the coupling constant is reduced.

[5] Actually a wide variety of theories have been developed in an attempt to describe gravity; some famous examples are string theory and supergravity.

1.2.3 Gauge invariance.

QED was built successfully on the basis of the classical electrodynamics theory (CED) of Maxwell and Lorentz, following theoretical and experimental requirements imposed by

- Lorentz invariance: independence on the reference frame.
- Locality: interacting fields are evaluated at the same space-time point to avoid action at a distance.
- Renormalizability: physical predictions are finite and well defined.
- Particle spectrum, symmetries and conservation laws already known must emerge from the theory.
- Local gauge invariance.

The gauge invariance requirement reflects the fact that the fundamental fields cannot be directly measured but associated fields which are the observables. Electric ($\mathbf{E}$) and magnetic ($\mathbf{B}$) fields in CED are associated with the electric scalar potential V and the vector potential $\mathbf{A}$. In particular, $\mathbf{E}$ can be obtained by measuring the change in the space of the scalar potential (ΔV); however, two scalar potentials differing by a constant f correspond to the same electric field. The same happens in the case of the vector potential $\mathbf{A}$; thus, different configurations of the associated fields result in the same set of values of the observables. The freedom in choosing one particular configuration is known as *gauge freedom*; the transformation law connecting two configurations is known as *gauge transformation* and the fact that the observables are not affected by a gauge transformation is called *gauge invariance*.

When the gauge transformation:

$$\begin{aligned}\mathbf{A} &\rightarrow \mathbf{A} - \Delta f \\ V &\rightarrow V - \frac{\partial f}{\partial t}\end{aligned} \tag{1.5}$$

is applied to Maxwell equations, they are still satisfied and the fields remain invariant. Thus, CED is invariant under gauge transformations and is called a *gauge theory*. The set of all gauge transformations form the *symmetry group* of the theory, which according to the group theory, has a set of *group generators*. The number of group generators determine the number of *gauge fields* of the theory.

As mentioned in the first lines of Section 1.2, QED has one symmetry group (U(1)) with one group generator (the Q operator) and one gauge field (the electromagnetic field A^{μ}). In CED there is not a clear definition, beyond the historical convention, of which fields are the fundamental and which are the associated, but in QED the fundamental field is A^{μ}. When a gauge theory is quantized, the gauge fields are quantized and their quanta are called *gauge bosons*. The word boson characterizes particles with integer spin which obey Bose-Einstein statistics.

As will be detailed in Section 1.3, interactions between particles in a system can be obtained by considering first the Lagrangian density of free particles in the system, which of course is incomplete because the interaction terms have been left out, and demanding global phase transformation invariance. Global phase transformation invariance means that a gauge transformation is performed identically to every point in the space[6] and the Lagrangian remains invariant. Then, the global transformation is promoted to a local phase transformation (this time the gauge transformation depends on the position in space) and again invariance is required.

[6] Here space corresponds to the 4-dimensional space i.e. space-time.

Due to the space dependence of the local transformation, the Lagrangian density is not invariant anymore. In order to reinstate the gauge invariance, the gauge covariant derivative is introduced in the Lagrangian and with it the gauge field responsible for the interaction between particles in the system. The new Lagrangian density is gauge invariant, includes the interaction terms needed to account for the interactions and provides a way to explain the interaction between particles through the exchange of the gauge boson.

This recipe was used to build QED and the theories that aim to explain the fundamental interactions.

1.2.4 Gauge bosons

The importance of the gauge bosons comes from the fact that they are the force mediators or force carriers. The features of the gauge bosons reflect those of the fields they represent and they are extracted from the Lagrangian density used to describe the interactions. In Section 1.3, it will be shown how the gauge bosons of the EI and WI emerge from the electroweak Lagrangian. The SI gauge bosons features are also extracted from the SI Lagrangian but it is not detailed in this document. The main features of the SM gauge bosons will be briefly presented below and summarized in Table 1.7.

- **Photon**. EI occurs when the photon couples to (is exchanged between) particles carrying electric charge; however, The photon itself does not carry electric charge, therefore, there is no coupling between photons. Given that the photon is massless the EI is of infinite range, i.e., electrically charged particles interact even if they are located far away one from each other; this also implies that photons always move with the speed of light.

- **Gluon**. SI is mediated by gluons which just as photons are massless. They carry one unit of color charge and one unit of anticolor charge, hence, gluons can couple to other gluons. As a result, the range of the SI is not infinite but very short due to the attraction between gluons, giving rise to the *color confinement* which explains why color charged particles cannot be isolated but live within composite particles, like quarks inside protons.

- **W, Z**. $W^{\pm}$ and Z, are massive which explains their short-range. Given that the WI is the only interaction that can change the flavor of the interacting particles, the W boson is the responsible for the nuclear transmutation where a neutron is converted into a proton or vice versa with the involvement of an electron and a neutrino (see Figure 1.4c). The Z boson is the responsible for the neutral weak processes like neutrino elastic scattering where no electric charge but momentum transference is involved. WI gauge bosons carry isospin charge which makes interaction between them possible.

Interaction	Mediator	Electric charge (e)	Color charge	Weak Isospin	mass (GeV/c^2)
Electromagnetic	Photon (γ)	0	No	0	0
Strong	Gluon (g)	0	Yes -octet	No	0
Weak	$W^{\pm}$	± 1	No	± 1	80.385 $\pm$ 0.015
	Z	0	No	0	91.188 $\pm$ 0.002

Table 1.7: SM gauge bosons main features [9].

1.3 Electroweak unification and the Higgs mechanism

Physicists dream of building a theory that contains all the interactions in one single interaction, i.e., showing that at some scale in energy all the four fundamental interactions are unified and only one interaction emerges in a *Theory of everything*. The first sign of the feasibility of such unification came from success in the construction of the CED. Einstein spent years trying to reach that full unification, which by 1920 only involved electromagnetism and gravity, with no success; however, a new partial unification was achieved in the 1960's, when S.Glashow [22], A.Salam [23] and S.Weinberg [24] independently proposed that electromagnetic and weak interactions are two manifestations of a more general interaction called *electroweak interaction (EWI)*. EWI was developed by following the useful prescription provided by QED and the gauge invariance principles.

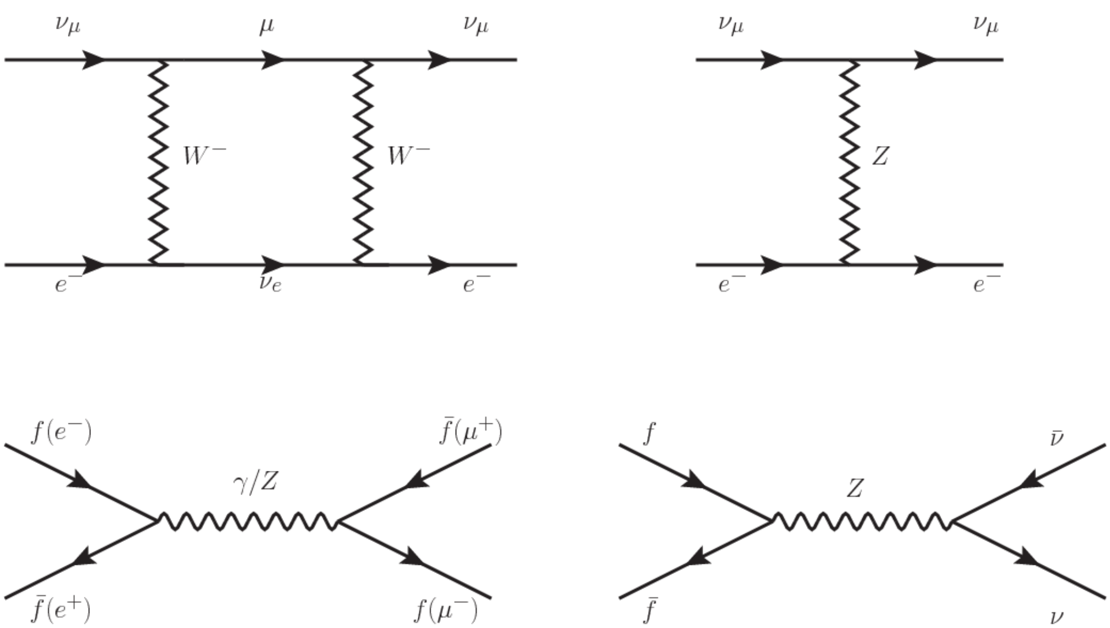

Figure 1.5: Top: $\nu_\mu - e^-$ scattering going through charged currents (left) and neutral currents (right). Bottom: neutral current processes for charged fermions (left) and involving neutrinos (right). While neutral current processes involving only charged fermions can proceed through EI or WI, those involving neutrinos can only proceed via WI.

The *classic* weak theory developed by Fermi, did not have the concept of the W boson but instead it was treated as a point interaction with the dimensionful constant G_F associated with it. It works really well at low energies very far off the W mass shell. When going up in energy, the theory of weak interactions involving the W boson is capable of explaining the β-decay and in general the processes mediated by $W^{\pm}$ bosons. However, there were some processes like the $\nu_\mu - e$ *scattering* which would require the exchange of two W bosons (see Figure 1.5 top diagrams) giving rise to divergent loop integrals and then non-finite predictions. The EWI theory, by including neutral currents involving fermions via the exchange of a neutral bosons Z, overcomes those divergences and the predictions become realistic.

Neutral weak interaction vertices conserve flavor in the same way as the electromagnetic vertices do, but additionally, the Z boson can couple to neutrinos which implies that processes involving charged fermions can proceed through EI or WI but processes involving neutrinos can proceed only through WI.

The prescription to build a gauge theory of the WI consists of proposing a free field Lagrangian density that includes the particles involved; next, by requesting invariance under global phase transformations first and generalizing to local phase transformations invariance later, the conserved currents are identified and interactions are generated by introducing gauge fields. Given that the goal is to include the EI and WI in a single theory, the group symmetry considered should be a combination of $SU(2)_L$ and $U(1)_{em}$, however the latter cannot be used directly because the EI treats left and right-handed particles indistinctly in contrast to the former. Fortunately, the weak hypercharge, which is a combination of the weak isospin and the electric charge (Eqn. 1.2) is suitable to be used since it is conserved by the EI and WI. Thus, the symmetry group to be considered is

$$G \equiv SU(2)_L \otimes U(1)_Y \tag{1.6}$$

The following treatment applies to any of the fermion generations, but for simplicity the first generation of leptons will be considered [5, 6, 25, 26].

Given the first generation of leptons

$$\psi_1 = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \qquad \psi_2 = \nu_{eR}, \qquad \psi_3 = e^-_R \tag{1.7}$$

the charged fermionic currents are given by

$$J_\mu \equiv J_\mu^+ = \bar{\nu}_{eL}\gamma_\mu e_L, \qquad J_\mu^\dagger \equiv J_\mu^- = \bar{e}_L\gamma_\mu \nu_{eL} \tag{1.8}$$

and the free Lagrangian is given by

$$\mathcal{L}_0 = \sum_{j=1}^{3} i\overline{\psi}_j(x)\gamma^\mu \partial_\mu \psi_j(x). \tag{1.9}$$

Mass terms are included directly in the QED free Lagrangians since they preserve the invariance under the symmetry transformations involved which treat left and right handed particles similarly, however mass terms of the form

$$m_W^2 W_\mu^\dagger(x) W^\mu(x) + \frac{1}{2} m_Z^2 Z_\mu(x) Z^\mu(x) - m_e \bar{\psi}_e(x)\psi_e(x) \tag{1.10}$$

which represent the mass of $W^\pm$, Z and electrons, are not invariant under G transformations, therefore the gauge fields described by the EWI are in principle massless.

Experiments have shown that the EWI gauge fields are not massless [27–30]; however, they have to acquire mass through a mechanism compatible with the gauge invariance; that mechanism is known as the *Higgs mechanism* and will be considered later in this Section. The global transformations in the combined symmetry group G can be written as

$$\begin{aligned}
\psi_1(x) \xrightarrow{G} \psi_1'(x) \equiv & U_Y U_L \psi_1(x), \\
\psi_2(x) \xrightarrow{G} \psi_2'(x) \equiv & U_Y \psi_2(x), \\
\psi_3(x) \xrightarrow{G} \psi_3'(x) \equiv & U_Y \psi_3(x)
\end{aligned} \tag{1.11}$$

where U_L represent the $SU(2)_L$ transformation acting only on the weak isospin doublet and U_Y represent the $U(1)_Y$ transformation acting on all the weak isospin multiplets. Explicitly

$$U_L \equiv \exp\left(i\frac{\sigma_i}{2}\alpha^i\right), \qquad U_Y \equiv \exp(iy_i\beta) \qquad (i = 1,2,3) \tag{1.12}$$

with σ_i the Pauli matrices and y_i the weak hypercharges. In order to promote the transformations from global to local while keeping the invariance, it is required that $\alpha^i = \alpha^i(x)$, $\beta = \beta(x)$ and the replacement of the ordinary derivatives by the covariant derivatives

$$\begin{aligned}
D_\mu\psi_1(x) &\equiv \left[\partial_\mu + ig\sigma_i W_\mu^i(x)/2 + ig' y_1 B_\mu(x)\right]\psi_1(x) \\
D_\mu\psi_2(x) &\equiv \left[\partial_\mu + ig' y_2 B_\mu(x)\right]\psi_2(x) \\
D_\mu\psi_3(x) &\equiv \left[\partial_\mu + ig' y_3 B_\mu(x)\right]\psi_3(x)
\end{aligned} \tag{1.13}$$

introducing in this way four gauge fields, $W_\mu^i(x)$ and $B_\mu(x)$, in the process. The covariant derivatives (Eqn. 1.13) are required to transform in the same way as fermion fields $\psi_i(x)$ themselves, therefore, the gauge fields transform as:

$$\begin{aligned} B_\mu(x) \xrightarrow{G} B'_\mu(x) \equiv & B_\mu(x) - \frac{1}{g'}\partial_\mu\beta(x) \\ W^i_\mu(x) \xrightarrow{G} W^{i\prime}_\mu(x) \equiv & W^i_\mu(x) - \frac{i}{g}\partial_\mu\alpha_i(x) - \varepsilon_{ijk}\alpha_i(x)W^i_\mu(x). \end{aligned} \tag{1.14}$$

The G invariant version of the Lagrangian density 1.9 can be written as

$$\mathcal{L}_0 = \sum_{j=1}^{3} i\overline{\psi}_j(x)\gamma^\mu D_\mu\psi_j(x) \tag{1.15}$$

where free massless fermion and gauge fields and fermion-gauge boson interactions are included. The EWI Lagrangian density must additionally include kinetic terms for the gauge fields ($\mathcal{L}_G$) which are built from the field strengths, according to

$$B_{\mu\nu}(x) \equiv \partial_\mu B_\nu - \partial_\nu B_\mu \tag{1.16}$$

$$W^i_{\mu\nu}(x) \equiv \partial_\mu W^i_\nu(x) - \partial_\nu W^i_\mu(x) - g\varepsilon^{ijk}W^j_\mu W^k_\nu \tag{1.17}$$

the last term in Eqn. 1.17 is added in order to hold the gauge invariance; therefore,

$$\mathcal{L}_G = -\frac{1}{4}B_{\mu\nu}(x)B^{\mu\nu}(x) - \frac{1}{4}W^i_{\mu\nu}(x)W^{\mu\nu}_i(x) \tag{1.18}$$

which contains not only the free gauge fields contributions, but also the gauge fields self-interactions and interactions among them.

The three weak isospin conserved currents resulting from the $SU(2)_L$ symmetry are given by

$$J^i_\mu(x) = \frac{1}{2}\bar{\psi}_1(x)\gamma_\mu\sigma^i\psi_1(x) \tag{1.19}$$

while the weak hypercharge conserved current resulting from the $U(1)_Y$ symmetry is given by

$$J_\mu^Y = \sum_{j=1}^{3} \overline{\psi}_j(x)\gamma_\mu y_j \psi_j(x) \tag{1.20}$$

In order to evaluate the electroweak interactions modeled by an isotriplet field W_μ^i that couples to isospin currents J_μ^i with strength g and additionally the singlet field B_μ which couples to the weak hypercharge current J_μ^Y with strength $g'/2$. The interaction Lagrangian density to be considered is

$$\mathcal{L}_I = -gJ^{i\mu}(x)W_\mu^i(x) - \frac{g'}{2}J^{Y\mu}(x)B_\mu(x) \tag{1.21}$$

Note that the weak isospin currents are not the same as the charged fermionic currents that were used to describe the WI (Eqn. 1.8), since the weak isospin eigenstates are not the same as the mass eigenstates, but they are closely related

$$J_\mu = \frac{1}{2}(J_\mu^1 + iJ_\mu^2), \qquad J_\mu^\dagger = \frac{1}{2}(J_\mu^1 - iJ_\mu^2). \tag{1.22}$$

The same happens with the gauge fields W_μ^i which are related to the mass eigenstates $W^\pm$ by

$$W_\mu^+ = \frac{1}{\sqrt{2}}(W_\mu^1 - iW_\mu^2), \qquad W_\mu^- = \frac{1}{\sqrt{2}}(W_\mu^1 + iW_\mu^2). \tag{1.23}$$

The fact that there are three weak isospin conserved currents is an indication that in addition to the charged fermionic currents, which couple charged to neutral leptons, there should be a neutral fermionic current that does not involve electric charge exchange; therefore, it couples neutral fermions or fermions of the same electric charge. The third weak isospin current contains a term that is similar to the electromagnetic current (j_μ^{em}), indicating that there is a relation between them and resembling the Gell-Mann-Nishijima formula 1.2 adapted to electroweak interactions

$$Q = T_3 + \frac{Y_W}{2}. \tag{1.24}$$

Just as Q generates the $U(1)_{em}$ symmetry, the weak hypercharge generates the $U(1)_Y$ symmetry as said before. It is possible to write the relationship in terms of the currents as

$$j_\mu^{em} = J_\mu^3 + \frac{1}{2} J_\mu^Y. \tag{1.25}$$

The neutral gauge fields W_μ^3 and B_μ cannot be directly identified with the Z and the photon fields since the photon interacts similarly with left and right-handed fermions; however, they are related through a linear combination given by

$$\begin{aligned} A_\mu =& B_\mu \cos\theta_W + W_\mu^3 \sin\theta_W \\ Z_\mu =& - B_\mu \sin\theta_W + W_\mu^3 \cos\theta_W \end{aligned} \tag{1.26}$$

where θ_W is known as the *Weinberg angle.* The interaction Lagrangian is now given by

$$\begin{aligned} \mathcal{L}_I = -\frac{g}{\sqrt{2}}(J^\mu W_\mu^+ + J^{\mu\dagger} W_\mu^-) &- \left(g \sin\theta_W J_\mu^3 + g' \cos\theta_W \frac{J_\mu^Y}{2} \right) A^\mu \\ &- \left(g \cos\theta_W J_\mu^3 - g' \sin\theta_W \frac{J_\mu^Y}{2} \right) Z^\mu \end{aligned} \tag{1.27}$$

the first term is the weak charged current interaction, while the second term is the electromagnetic interaction under the condition

$$g \sin\theta_W = g' \cos\theta_W = e, \quad \frac{g'}{g} = \tan\theta_W \tag{1.28}$$

contained in the Eqn.1.25; the third term is the neutral weak current.

Note that the neutral fields transformation given by the Eqn. 1.26 can be written in terms of the coupling constants g and g' as:

$$A_\mu = \frac{g' W_\mu^3 + g B_\mu}{\sqrt{g^2 + g'^2}}, \qquad Z_\mu = \frac{g W_\mu^3 - g' B_\mu}{\sqrt{g^2 + g'^2}}. \tag{1.29}$$

So far, the Lagrangian density describing the non-massive EWI is:

$$\mathcal{L}_{nmEWI} = \mathcal{L}_0 + \mathcal{L}_G \tag{1.30}$$

where fermion and gauge fields have been considered massless because their regular mass terms are manifestly non invariant under G transformations; therefore, masses have to be generated in a gauge invariant way. The mechanism by which this goal is achieved is known as the *Higgs mechanism* and is closely connected to the concept of *spontaneous symmetry breaking.*

1.3.1 Spontaneous symmetry breaking (SSB)

Figure 1.6 left shows a steel nail (top) which is subject to an external force; the form of the potential energy is also shown (bottom).

Before reaching the critical force value, the system has rotational symmetry with respect to the nail axis; however, after the critical force value is reached the nail buck-

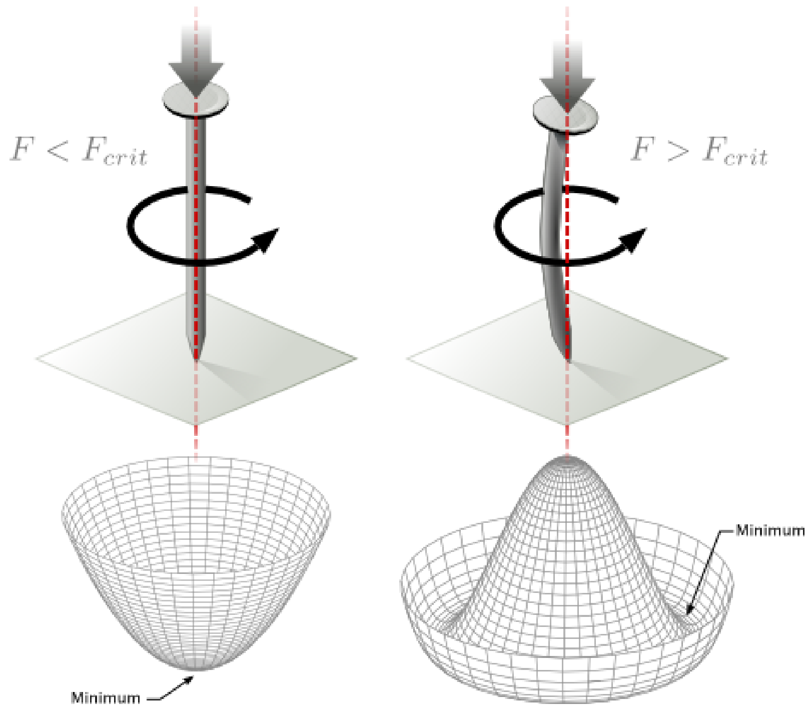

Figure 1.6: Spontaneous symmetry breaking mechanism. The steel nail, subject to an external force (top left), has rotational symmetry with respect to its axis. When the external force overcomes a critical value the nail buckles (top right) choosing a minimal energy state (ground state) and thus *breaking spontaneously the rotational symmetry.* The potential energy (bottom) changes but holds the rotational symmetry; however, an infinite number of asymmetric ground states are generated and circularly distributed in the bottom of the potential [31].

les (top right). The form of the potential energy (bottom right) changes appearing a set of infinity minima but preserving its rotational symmetry. Right before the nail buckles there is no indication of the direction the nail will bend because any of the directions are equivalent, but once the nail bends, choosing a direction, an arbitrary minimal energy state (ground state) is selected and it does not share the system's rotational symmetry. This mechanism for reaching an asymmetric ground state is known as *spontaneous symmetry breaking.*

The lesson from this analysis is that the way to introduce the SSB mechanism into a system is by adding the appropriate potential to it.

Figure 1.7 shows a plot of the potential $V(\phi)$ in the case of a scalar field ϕ

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2 \tag{1.31}$$

If $\mu^2 > 0$ the potential has only one minimum at $\phi = 0$ and describes a scalar field with mass μ. If $\mu^2 < 0$ the potential has a local maximum at $\phi = 0$ and two minima at $\phi = \pm\sqrt{-\mu^2/\lambda}$ which enables the SSB mechanism to work.

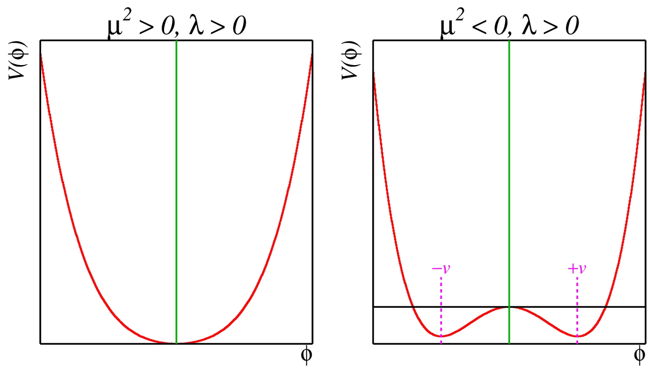

Figure 1.7: Shape of the potential $V(\phi)$ for $\lambda > 0$ and: $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right). The case $\mu^2 < 0$ corresponds to the potential suitable for introducing the SSB mechanism by choosing one of the two ground states which are connected via reflection symmetry. [31].

In the case of a complex scalar field $\phi(x)$

$$\phi(x) = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \tag{1.32}$$

the Lagrangian (invariant under global $U(1)$ transformations) is given by

$$\mathcal{L} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi), \qquad V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2 \tag{1.33}$$

where an appropriate potential has been added in order to introduce the SSB.

As seen in Figure 1.8, the potential has now an infinite number of minima circularly distributed along the ξ-direction which makes possible the occurrence of the SSB by choosing an arbitrary ground state; for instance, $\xi = 0$, i.e. $\phi_1 = v, \phi_2 = 0$

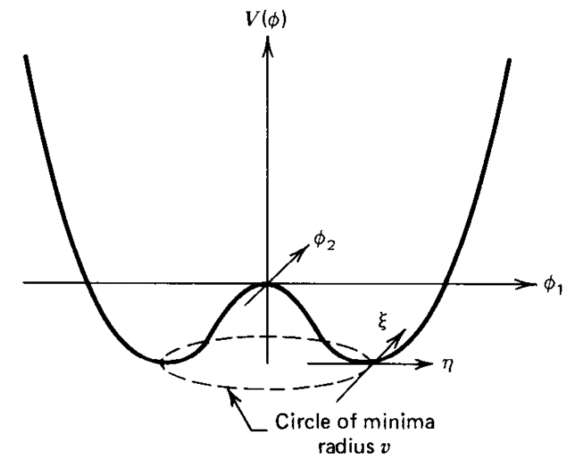

Figure 1.8: Potential for complex scalar field. There is a circle of minima of radius v along the ξ-direction [6].

$$\phi_0 = \frac{v}{\sqrt{2}}\exp(i\xi) \quad \xrightarrow{SSB} \quad \phi_0 = \frac{v}{\sqrt{2}} \tag{1.34}$$

As usual, excitations over the ground state are studied by making an expansion about it; thus, the excitations can be parametrized as:

$$\phi(x) = \frac{1}{\sqrt{2}}(v + \eta(x) + i\xi(x)) \tag{1.35}$$

which when substituted into Eqn. 1.33 produces a Lagrangian in terms of the new fields η and ξ

$$\mathcal{L}' = \frac{1}{2}(\partial_\mu\xi)^2 + \frac{1}{2}(\partial_\mu\eta)^2 + \mu^2\eta^2 - V(\phi_0) - \lambda v\eta(\eta^2+\xi^2) - \frac{\lambda}{4}(\eta^2+\xi^2)^2 \tag{1.36}$$

where the last two terms represent the interactions and self-interaction between the two fields η and ξ. The particular feature of the SSB mechanism is revealed when looking to the first three terms of $\mathcal{L}'$. Before the SSB, only the massless ϕ field is present in the system; after the SSB there are two fields of which the η-field has acquired mass $m_\eta = \sqrt{-2\mu^2}$ while the ξ-field is still massless (see Figure 1.9).

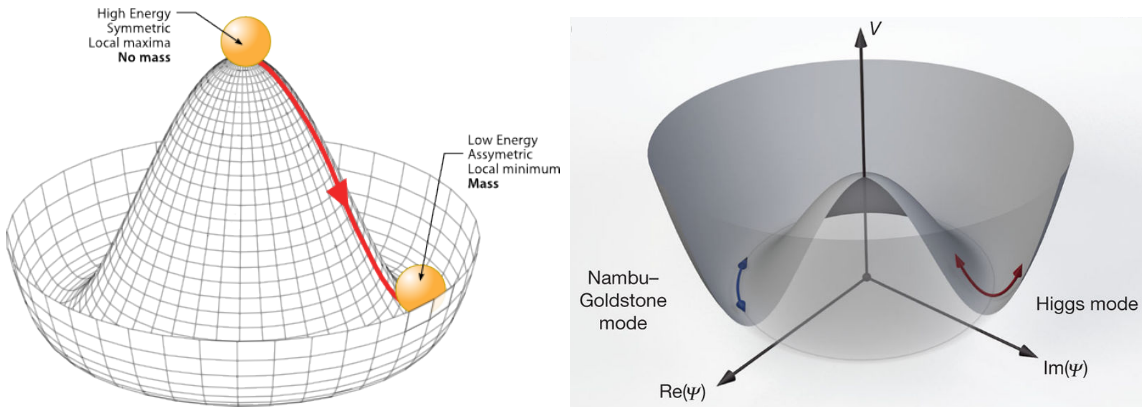

Figure 1.9: SSB mechanism for a complex scalar field [31, 32].

Thus, *the SSB mechanism serves as a method to generate mass but as a side effect a massless field is introduced in the system.* This fact is known as the Goldstone theorem and states that a massless scalar field appears in the system for each continuous symmetry spontaneously broken. Another version of the Goldstone theorem states that *"if a Lagrangian is invariant under a continuous symmetry group G, but the vacuum is only invariant under a subgroup $H \subset G$, then there must exist as many massless spin-0 particles (Nambu-Goldstone bosons) as broken generators."* [26] The Nambu-Goldstone boson can be understood considering that the potential in the ξ-direction is flat so excitations in that direction are not energy consuming and thus represent a massless state.

1.3.2 Higgs mechanism

When the SSB mechanism is introduced in the formulation of the EWI in an attempt to generate the mass of the so far massless gauge bosons and fermions, an interesting effect is revealed. In order to keep the G symmetry group invariance and generate the mass of the EW gauge bosons, a G invariant Lagrangian density ($\mathcal{L}_S$) has to be added to the non massive EWI Lagrangian (Eqn. 1.30)

$$\mathcal{L}_S = (D_\mu\phi)^\dagger(D^\mu\phi) - \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2, \qquad \lambda > 0, \mu^2 < 0 \tag{1.37}$$

$$D_\mu\phi = \left(i\partial_\mu - g\frac{\sigma_i}{2}W_\mu^i - g'\frac{Y}{2}B_\mu\right)\phi \tag{1.38}$$

ϕ has to be an isospin doublet of complex scalar fields so it preserves the G invariance; thus ϕ can be defined as:

$$\phi = \begin{pmatrix}\phi^+ \\ \phi^0\end{pmatrix} \equiv \frac{1}{\sqrt{2}}\begin{pmatrix}\phi_1 + i\phi_2 \\ \phi_3 + i\phi_4\end{pmatrix}. \tag{1.39}$$

The minima of the potential are defined by

$$\phi^\dagger\phi = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_1^4) = -\frac{\mu^2}{2\lambda}. \tag{1.40}$$

The choice of the ground state is critical. By choosing a ground state, invariant under $U(1)_{em}$ gauge symmetry, the photon will remain massless and the $W^\pm$ and Z bosons masses will be generated which is exactly what is needed. In that sense, the best choice corresponds to a weak isospin doublet with $T_3 = -1/2$, $Y_W = 1$ and $Q = 0$ which defines a ground state with $\phi_1 = \phi_2 = \phi_4$ and $\phi_3 = v$:

$$\phi_0 \equiv \frac{1}{\sqrt{2}}\begin{pmatrix}0 \\ v\end{pmatrix}, \qquad v^2 \equiv -\frac{\mu^2}{\lambda}. \tag{1.41}$$

where the vacuum expectation value v is fixed by the Fermi coupling G_F according to $v = (\sqrt{2}G_F)^{1/2} \approx 246$ GeV.

The G symmetry has been broken and three Nambu-Goldstone bosons will appear.

The next step is to expand ϕ about the chosen ground state as:

$$\phi(x) = \frac{1}{\sqrt{2}} \exp\left(\frac{i}{v}\sigma_i \theta^i(x)\right) \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \approx \frac{1}{\sqrt{2}} \begin{pmatrix} \theta_1(x) + i\theta_2(x) \\ v + H(x) - i\theta_3(x) \end{pmatrix} \tag{1.42}$$

to describe fluctuations from the ground state ϕ_0. The fields $\theta_i(x)$ represent the Nambu-Goldstone bosons while $H(x)$ is known as *Higgs field.* The fundamental feature of the parametrization used is that the dependence on the $\theta_i(x)$ fields is factored out in a global phase that can be eliminated by taking the physical *unitary gauge* $\theta_i(x) = 0$. Therefore the expansion about the ground state is given by:

$$\phi(x) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \tag{1.43}$$

which when substituted into $\mathcal{L}_S$ (Eqn. 1.37) results in a Lagrangian containing the now massive three gauge bosons $W^{\pm}, Z$, one massless gauge boson (photon) and the new Higgs field (H). The three degrees of freedom corresponding to the Nambu-Goldstone bosons are now integrated into the massive gauge bosons as their longitudinal polarizations which were not available when they were massless particles. The effect by which vector boson fields acquire mass after an spontaneous symmetry breaking, but without an explicit gauge invariance breaking is known as the *Higgs mechanism.*

The mechanism was proposed by three independent groups: F.Englert and R.Brout in August 1964 [33], P.Higgs in October 1964 [34] and G.Guralnik, C.Hagen and T.Kibble in November 1964 [35]; however, its importance was not realized until S.Glashow [22], A.Salam [23] and S.Weinberg [24], independently, proposed that electromagnetic and weak interactions are two manifestations of a more general interaction called *electroweak interaction* in 1967.

1.3.3 Masses of the gauge bosons

The masses of the gauge bosons are extracted by evaluating the kinetic part of Lagrangian $\mathcal{L}_S$ in the ground state (known also as the vacuum expectation value), i.e.,

$$\left|\left(\partial_\mu - ig\frac{\sigma_i}{2}W_\mu^i - i\frac{g'}{2}B_\mu\right)\phi_0\right|^2 = \left(\frac{1}{2}vg\right)^2 W_\mu^+ W^{-\mu} + \frac{1}{8}v^2(W_\mu^3, B_\mu)\begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix}\begin{pmatrix} W^{3\mu} \\ B^\mu \end{pmatrix} \tag{1.44}$$

comparing with the typical mass term for a charged boson $M_W^2 W^+ W^-$

$$M_W = \frac{1}{2}vg. \tag{1.45}$$

The second term in the right side of the Eqn.1.44 comprises the masses of the neutral bosons, but it needs to be written in terms of the gauge fields Z_μ and A_μ in order to be compared to the typical mass terms for neutral bosons, therefore using Eqn. 1.29

$$\begin{aligned} \frac{1}{8}v^2[g^2(W_\mu^3)^2 - 2gg'W_\mu^3 B^\mu + g'^2 B_\mu^2] &= \frac{1}{8}v^2[gW_\mu^3 - g'B_\mu]^2 + 0[g'W_\mu^3 + gB_\mu]^2 \\ &= \frac{1}{8}v^2[\sqrt{g^2+g'^2}Z_\mu]^2 + 0[\sqrt{g^2+g'^2}A_\mu]^2 \end{aligned} \tag{1.46}$$

and then

$$M_Z = \frac{1}{2}v\sqrt{g^2+g'^2}, \qquad M_A = 0 \tag{1.47}$$

1.3.4 Masses of the fermions

The lepton mass terms can be generated by introducing a gauge invariant Lagrangian term describing the Yukawa coupling between the lepton field and the Higgs field

$$\mathcal{L}_{Yl} = -G_l \left[(\bar{\nu}_l, \bar{l})_L \begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} l_R + \bar{l}_R(\phi^-, \bar{\phi}^0)\begin{pmatrix}\nu_l\\ l\end{pmatrix}_L \right], \qquad l = e, \mu, \tau. \tag{1.48}$$

After the SSB and replacing the usual field expansion about the ground state (Eqn.1.41) into $\mathcal{L}_{Yl}$, the mass term arises

$$\mathcal{L}_{Yl} = -m_l(\bar{l}_L l_R + \bar{l}_R l_L) - \frac{m_l}{v}(\bar{l}_L l_R + \bar{l}_R l_L)H = -m_l \bar{l} l \left(1 + \frac{H}{v}\right) \tag{1.49}$$

$$m_l = \frac{G_l}{\sqrt{2}} v \tag{1.50}$$

where the additional term represents the lepton-Higgs interaction. The quark masses are generated in a similar way as lepton masses but for the upper member of the quark doublet a different Higgs doublet is needed:

$$\phi_c = -i\sigma_2 \phi* = \begin{pmatrix}-\bar{\phi}^0\\ \phi^-\end{pmatrix}. \tag{1.51}$$

Additionally, given that the quark isospin doublets are not constructed in terms of the mass eigenstates but in terms of the flavor eigenstates, as shown in Table1.5, the coupling parameters will be related to the CKM matrix elements; thus, the quark Lagrangian is given by:

$$\mathcal{L}_{Yq} = -G_d^{i,j}(\bar{u}_i, \bar{d}'_i)_L \begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} d_{jR} - G_u^{i,j}(\bar{u}_i, \bar{d}'_i)_L \begin{pmatrix}-\bar{\phi}^0\\ \phi^-\end{pmatrix} u_{jR} + h.c. \tag{1.52}$$

with i,j=1,2,3. After SSB and expansion about the ground state, the diagonal form of $\mathcal{L}_{Yq}$ is:

$$\mathcal{L}_{Yq} = -m_d^i \bar{d}_i d_i \left(1 + \frac{H}{v}\right) - m_u^i \bar{u}_i u_i \left(1 + \frac{H}{v}\right) \tag{1.53}$$

Fermion masses depend on arbitrary couplings G_l and $G_{u,d}$ and are not predicted by the theory.

1.3.5 The Higgs field

After the characterization of the fermions and gauge bosons as well as their interactions, it is necessary to characterize the Higgs field itself. The Lagrangian $\mathcal{L}_S$ in Eqn. 1.37 written in terms of the gauge bosons is given by

$$\mathcal{L}_S = \frac{1}{4}\lambda v^4 + \mathcal{L}_H + \mathcal{L}_{HV} \tag{1.54}$$

$$\mathcal{L}_H = \frac{1}{2}\partial_\mu H \partial^\mu H - \frac{1}{2} m_H^2 H^2 - \frac{1}{2v} m_H^2 H^3 - \frac{1}{8v^2} m_H^2 H^4 \tag{1.55}$$

$$\mathcal{L}_{HV} = m_H^2 W_\mu^+ W^{\mu -} \left(1 + \frac{2}{v} H + \frac{2}{v^2} H^2\right) + \frac{1}{2} m_Z^2 Z_\mu Z^\mu \left(1 + \frac{2}{v} H + \frac{2}{v^2} H^2\right) \tag{1.56}$$

The mass of the Higgs boson is deduced as usual from the mass term in the Lagrangian resulting in:

$$m_H = \sqrt{-2\mu^2} = \sqrt{2\lambda} v \tag{1.57}$$

however, it is not predicted by the theory either. The experimental measurement of the Higgs boson mass have been performed by the *Compact Muon Solenoid (CMS)* experiment and the *A Toroidal LHC ApppartuS (ATLAS)* experiments at the *Large Hadron Collider(LHC)*, [36–38], and is presented in Table 1.8.

Property	Value
Electric charge	0
Color charge	0
Spin	0
Weak isospin	-1/2
Weak hypercharge	1
Parity	1
Mass (GeV/c^2)	125.09±0.21 (stat.)±0.11 (syst.)

Table 1.8: Higgs boson properties. Higgs mass is not predicted by the theory and the value here corresponds to the experimental measurement.

1.3.6 Production of Higgs bosons at LHC

At the LHC, Higgs bosons are produced as a result of the collision of two counter-rotating protons beams. A detailed description of the LHC machine will be presented in chapter 2.

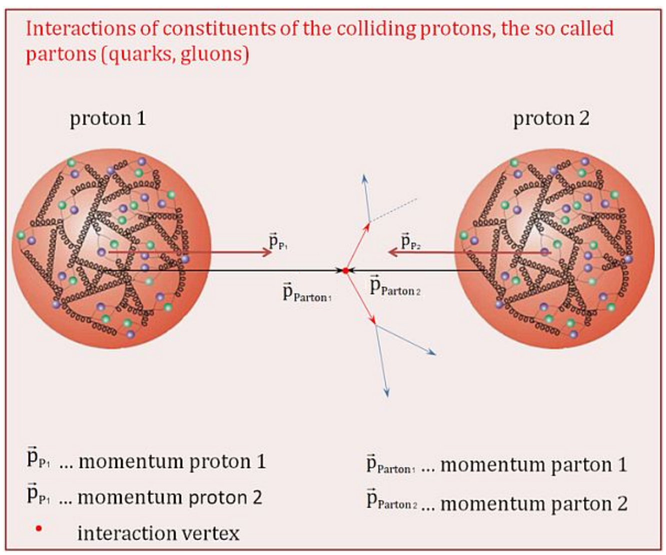

Figure 1.10: Proton-proton collision. Protons are composed of 3 valence quarks, a sea of quarks and gluons; therefore in a proton-proton collision, quarks and gluons are those who collide. [39].

Protons are composed of quarks and these quarks are bound by gluons; however, what is commonly called the quark content of the proton makes reference to the valence quarks. In fact, a proton is not just a rigid entity with three balls in it all tied up with springs, but the gluons exchanged by the valence quarks tend to split spontaneously into quark-antiquark pairs or more gluons, creating *sea of quarks and gluons* as represented in Figure 1.10.

In a proton-proton (pp) collision, the proton's constituents, quarks and gluons, are those that collide. The pp cross section depends on the momentum of the colliding particles, reason for which it is needed to know how the momentum is distributed inside the proton. Quarks and gluons are known as partons, hence, the functions that describe how the proton momentum is distributed among partons inside it are called *parton distribution functions (PDFs)*; PDFs are determined from experimental data obtained in experiments where the internal structure of hadrons is tested, and depend on the momentum transfer Q and the fraction of momentum x carried by an specific parton. Figure 1.11 shows the proton PDFs ($xf(x, Q^2)$) for two values of Q.

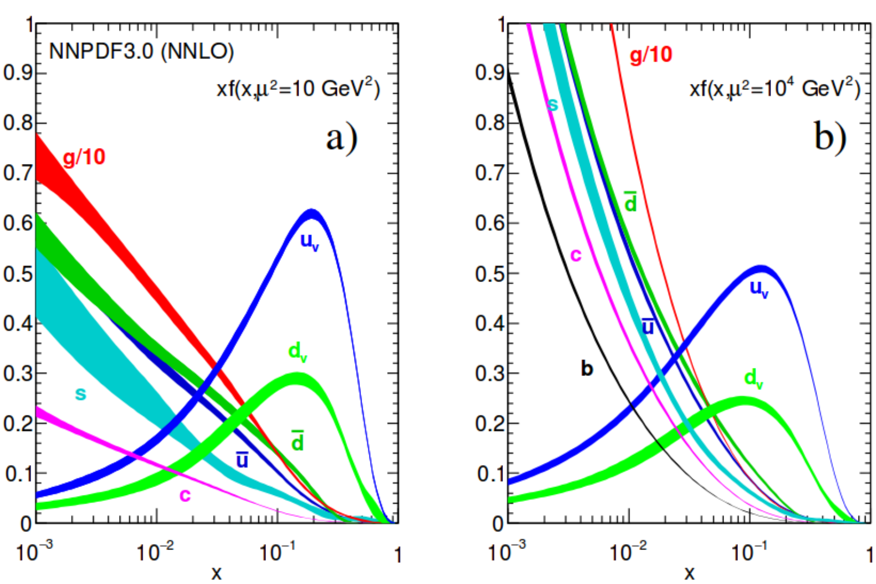

Figure 1.11: Proton PDFs for two values of Q^2: left.$\mu^2 = Q^2 = 10$ GeV^2 , right. $\mu^2 = Q^2 = 10$ GeV^2. u_v and d_v correspond to the u and d valence quarks, $s, c, b, \bar{u}, \bar{d}$ correspond to sea quarks, and g corresponds to gluons. Note that glouns carry a high fraction of the proton's momentum. [9]

In physics, a common approach to study complex systems consists of starting with a simpler version of them, for which a well known description is available, and adding an additional *perturbation* which represents a small deviation from the known behavior. If the perturbation is small enough, the physical quantities associated with the perturbed system are expressed as a series of corrections to those of the simpler system. The perturbation series corresponds to an expansion in power series of a small parameter, therefore, the more terms are considered in the series (the higher order in the perturbation series), the more precise is the the description of the complex system. If the perturbation does not get progressively smaller, the strategy cannot be applied and new methods have to be employed.

High energy systems, like the Higgs production at LHC explored in this thesis, usually can be treated perturbatively with the expansion made in terms of the coupling constants. The overview presented here will be oriented specifically to the Higgs boson production mechanisms in *pp* collisions at LHC.

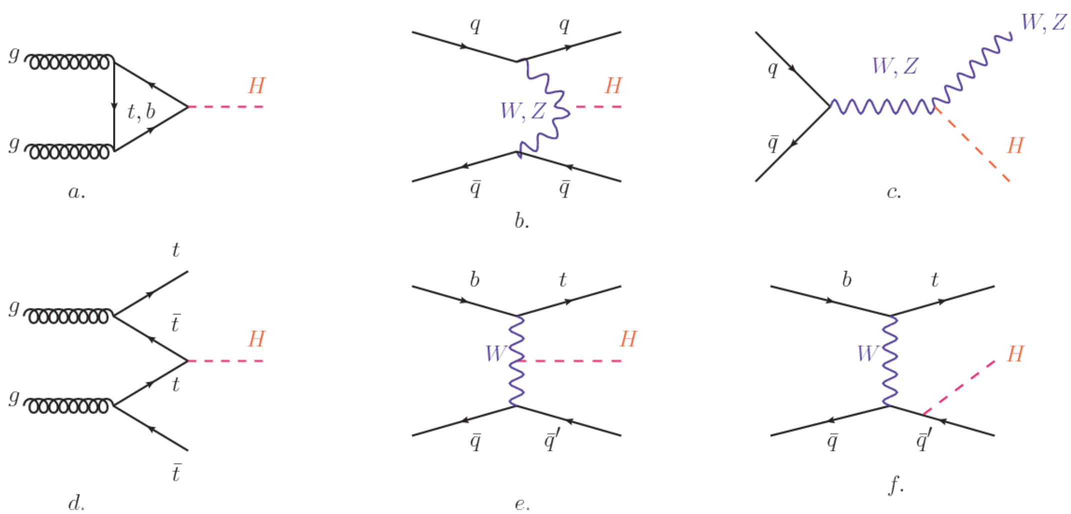

Figure 1.12: Main Higgs boson production mechanism Feynman diagrams. a. gluon-gluon fusion, b. vector boson fusion (VBF), c. Higgs-strahlung, d. Associated production with a top or bottom quark pair, e-f. associated production with a single top quark.

Figure 1.12 shows the Feynman diagrams for the leading order (first order) Higgs production processes at LHC; note that in these diagrams the incoming particles are not the protons themselves but the partons from the protons that actually participate in the interaction, hence, theorists typically calculate the cross section for the parton interaction, and then convolute that cross section with the information from the PDFs to get a production cross section that is actually measured in experiments. The cross section for Higgs production as a function of the center of mass-energy ($\sqrt{s}$) for pp collisions is showed in Figure 1.13 left. The tags NLO (next to leading order), NNLO (next to next to leading order) and N3LO (next to next to next to leading order) make reference to the order at which the perturbation series have been considered while the tags QCD and EW correspond to the strong and electroweak coupling constants respectively.

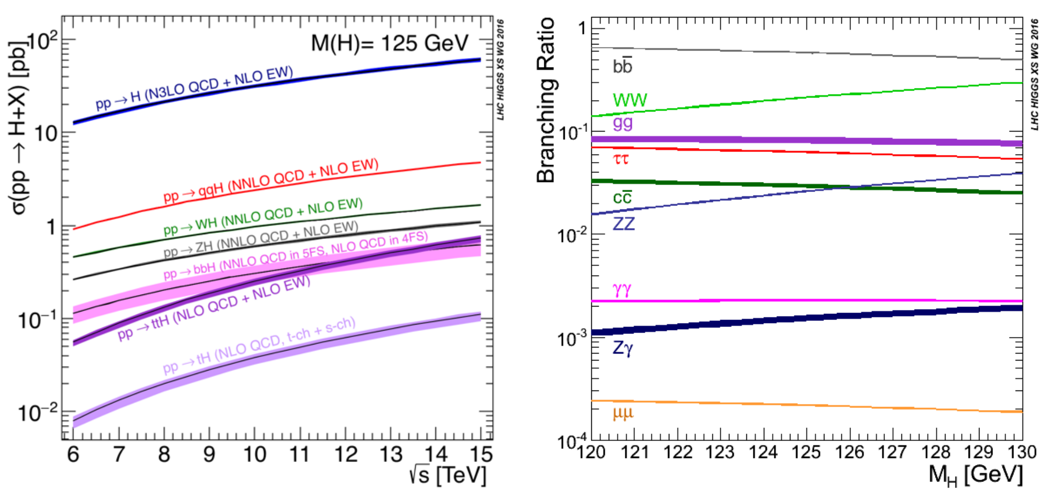

Figure 1.13: Higgs boson production cross sections (left) and decay branching ratios (right) for the main mechanisms. The VBF is indicated as qqH [40].

The main production mechanism is the gluon fusion (Figure 1.12a and $pp \to H$ in Figure 1.13) given that gluons carry the highest fraction of momentum of the protons in pp colliders (as shown in Figure **??**). Since the Higgs boson does not couple to gluons, the mechanism proceeds through the exchange of a virtual top-quark loop. Note that in this process the Higgs boson is produced alone, turning out to be problematic for some Higgs decays, because such absence of anything produced in association with the Higgs represent a trouble for triggering, however, this mechanism is experimentally clean when combined with the two-photon or the four-lepton decay channels (see Section 1.3.7).

Vector boson fusion (Figure 1.12b and $pp \to qqH$ in Figure 1.13) has the second largest production cross section. The scattering of two fermions is mediated by a weak gauge boson which later emits a Higgs boson. In the final state, the two fermions tend to be located in the central region of the detector; this kind of features are generally used as a signature when analyzing the datasets provided by the experiments[7].

In the Higgs-strahlung mechanism (Figure 1.12c and $pp \to WH, pp \to ZH$ in Figure 1.13) two fermions annihilate to form a weak gauge boson. If the initial fermions have enough energy, the emergent boson might emit a Higgs boson.

The associated production with a top or bottom quark pair and the associated production with a single top quark (Figure 1.12d-f and $pp \to bbH, pp \to t\bar{t}H, pp \to tH$ in Figure 1.13) have a smaller cross section than the main three mechanisms above, but they provide a good opportunity to test the Higgs-top coupling. The analysis reported in this thesis is developed using these production mechanisms. A detailed description of the tHmechanism will be given in Section 1.5.

1.3.7 Higgs boson decay channels

When a particle can decay through several modes, also known as channels, the probability of decaying through a given channel is quantified by the *branching ratio (BR)* of the decay channel; thus, the BR is defined as the ratio of number of decays going through that given channel to the total number of decays. In regard to the Higgs boson decay, the BR can be predicted with accuracy once the Higgs mass is known [41,42]. In Figure 1.13 right, a plot of the BR as a function of the Higgs mass

[7] More details about how to identify events of interest in this analysis will be given in chapter 6.

is presented; the largest predicted BR corresponds to the $b\bar{b}$ pair decay channel (see Table 1.9) given that it is the heaviest particle pair whose on-shell [8] production is kinematically allowed in the decay.

Decay channel	Branching ratio	Rel. uncertainty
$H \to b\bar{b}$	5.84×10^{-1}	$+3.2\% - 3.3\%$
$H \to W^+W^-$	2.14×10^{-1}	$+4.3\% - 4.2\%$
$H \to \tau^+\tau^-$	6.27×10^{-2}	$+5.7\% - 5.7\%$
$H \to ZZ$	2.62×10^{-2}	$+4.3\% - 4.1\%$
$H \to \gamma\gamma$	2.27×10^{-3}	$+5.0\% - 4.9\%$
$H \to Z\gamma$	1.53×10^{-3}	$+9.0\% - 8.9\%$
$H \to \mu^+\mu^-$	2.18×10^{-4}	$+6.0\% - 5.9\%$

Table 1.9: Predicted branching ratios and the relative uncertainty for some decay channels of a SM Higgs boson with $m_H = 125\text{GeV/c}^2$ [9]; the uncertainties are driven by theoretical uncertainties for the different Higgs boson partial widths and by parametric uncertainties associated to the strong coupling and the masses of the quarks which are the input parameters. Further details on these calculations can be found in Reference [43]

Decays to other lepton and quark pairs, like electron, strange, up, and down quark pairs not listed in the table, are also possible but their likelihood is too small to measure since they are very lightweight, hence, their interaction with the Higgs boson is very weak. On other hand, the decay to top quark pairs is heavily suppressed due to the top quark mass (≈ 173 GeV/c^2).

Decays to gluons proceed indirectly through a virtual top quark loop while the decays to photons proceed through a virtual W boson loop, therefore, their branching ratio is smaller compared to direct interaction decays. Same is true for the decay to a photon and a Z boson.

[8] In general, on-shell or real particles are those which satisfy the energy-momentum relation ($E^2 - |\vec{p}|^2c^2 = m^2c^4$); off-shell or virtual particles does not satisfy it which is possible under the uncertainty principle of quantum mechanics. Usually, virtual particles correspond to internal propagators in Feynman diagrams.

In the case of decays to pairs of W and Z bosons, the decay proceed with one of the bosons being on-shell and the other being off-shell. The likelihood of the process diminish depending on how far off-shell are the virtual particles involved, hence, the branching ratio for W boson pairs is bigger that for Z boson pairs since Z boson mass is bigger that W boson mass.

Note that the decay to a pair of virtual top quarks is possible, but the probability is way too small.

1.4 Experimental status of the anomalous Higgs-fermion coupling

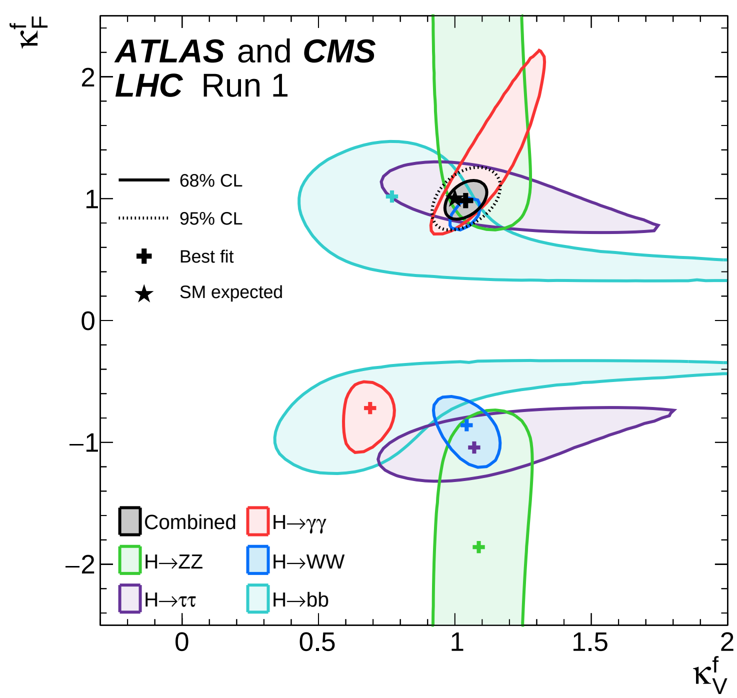

Figure 1.14: Combination of the ATLAS and CMS fits for coupling modifiers κ_t-κ_V; also shown the individual decay channels combination and their global combination. No assumptions have been made on the sign of the coupling modifiers [44].

ATLAS and CMS have performed analyses of the anomalous Higgs-fermion coupling by making likelihood scans for the two coupling modifiers, κ_f and κ_V, under the assumption that $\kappa_Z = \kappa_W \equiv \kappa_V$ and $\kappa_t = \kappa_\tau = \kappa_b \equiv \kappa_f$. Figure 1.14 shows the result of the combination of ATLAS and CMS fits; also the individual decay channels combination and the global combination results are shown. Note that from this plot there is limited information on the sign of the coupling since the only information available about the sign of the coupling comes from decays rather than production.

While all the channels are compatible for positive values of the modifiers, for negative values of κ_f there is no compatibility. The best fit for individual channels is compatible with negative values of κ_f except for the $H \rightarrow bb$ channel. The best fit for the combination yields $\kappa_f \geq 0$, in contrast to the yields from the individual channels; the reason of this yield resides in the H $\rightarrow \gamma\gamma$ coupling. H $\rightarrow \gamma\gamma$ decay proceeds through a loop of either top quarks or W bosons, hence, this channel is sensitive to κ_t thanks to the interference of these two amplitude contributions; under the assumption that no beyond SM particles take part in the loops, a flipped sign of κ_t will increase the H $\rightarrow \gamma\gamma$ branching fraction by a factor of ~ 2.4 which is not supported by measurements; thus, this large asymmetry between the positive and negative coupling ratios in the H $\rightarrow \gamma\gamma$ channel drives the yield of the global fit and would means that the anomalous H-t coupling is excluded as stated in Reference [44], but there is a caveat, this exclusion holds only if no new particles contribute to the loop in the main diagram for that decay.

Although the $H \rightarrow bb$ channel is expected to be the most sensitive channel and its best fit value of κ_t is positive, and then the global fit yield is still supported, the contributions from all the other decay channels, small compared to the $H \rightarrow bb$, indicate that the anomalous H-t coupling cannot be excluded completely, motivating to look at tHprocesses which can help with both, the limited information on the sign of the H-t coupling and the access to information from the Higgs boson production rather that from its decays.

It will be shown in Section 1.5 that the same interference effect enhance the tHproduction rate and could reveal evidence of direct production of heavy new particles as predicted in composite and little Higgs models [45], or new physics related to Higgs boson mediated flavor changing neutral currents [46] as well as probes the CP-violating phase of the H-t coupling [47, 48].

1.5 Associated production of a Higgs boson and a single top quark

The production of Higgs boson in association with a top quark has been extensively studied [47, 49–52]. While measurements of the main Higgs production mechanisms rates are sensitive to the strength of the Higgs coupling to W boson or top quark, they are not sensitive to the relative sign between the two couplings. In this thesis, the Higgs boson production mechanism explored is the associated production with a single top quark (tH) which offers sensitivity to the relative sign of the Higgs couplings to W boson and to top quark. The description given here is based on Reference [51]

A process where two incoming particles interact and produce a final state with two particles can proceed in three called channels (see, for instance, Figure 1.15 omitting the red line). The t-channel represents processes where an intermediate particle is emitted by one of the incoming particles and absorbed by the other. The s-channel represents processes where the two incoming particles merge into an intermediate particle which eventually will split into the particles in the final state. The third channel, u-channel, is similar to the t-channel but the two outgoing particles interchange their

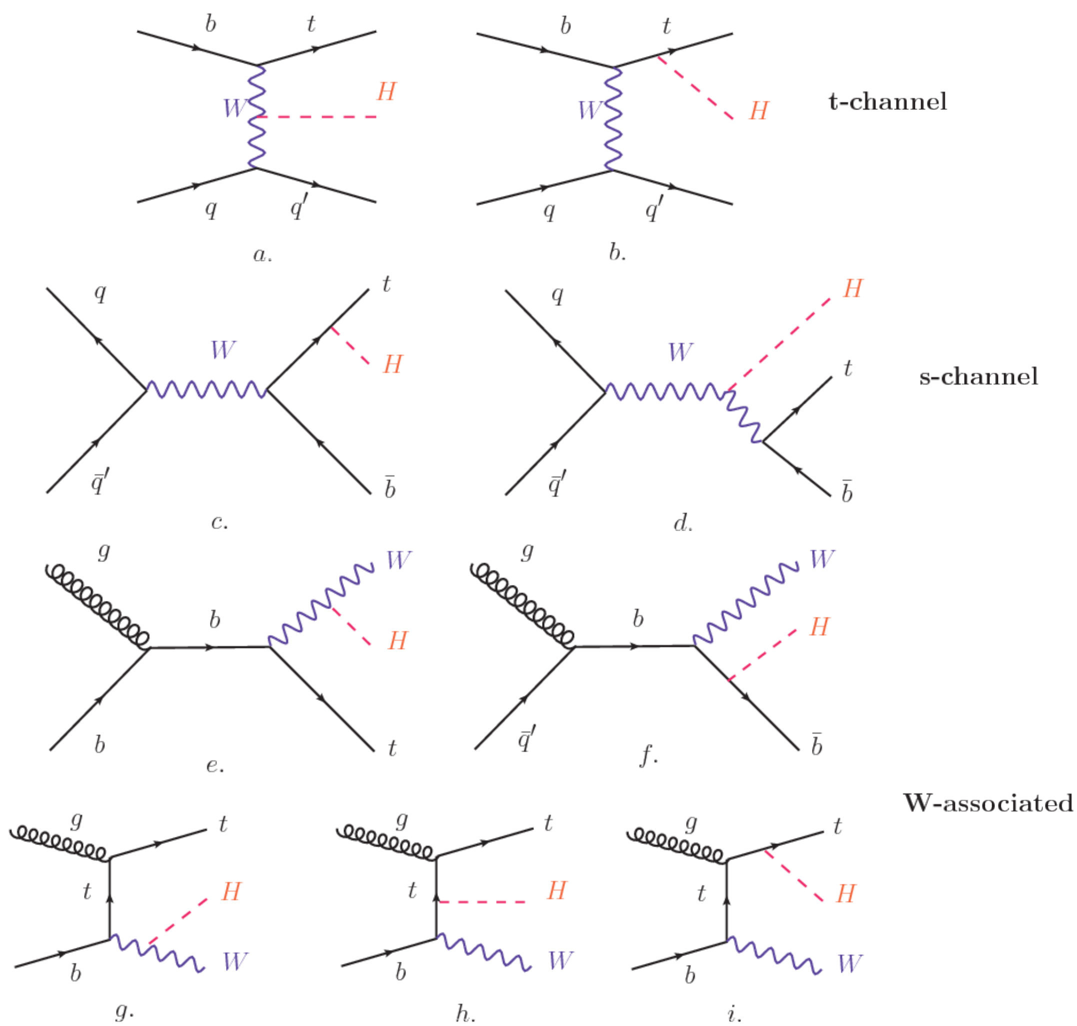

Figure 1.15: Associated Higgs boson production with a top quark mechanism Feynman diagrams. a.,b. t-channel (tHq), c.,d. s-channel (tHb), e-i. W-associated.

roles. These three channels are connected to the so-called Mandelstam variables

$$s = (p_1 + p_2)^2 = (p_1' + p_2')^2 \rightarrow \text{ square of the center mass-energy.} \tag{1.58}$$

$$t = (p_1 - p_1')^2 = (p_2' - p_2)^2 \rightarrow \text{ square of the four-momentum transfer.} \tag{1.59}$$

$$u = (p_1 - p_2')^2 = (p_1' - p_2)^2 \rightarrow \text{ square of the crossed four-momentum transfer.} \tag{1.60}$$

$$s + t + u = m_1^2 + m_2^2 + m_1'^2 + m_2'^2 \tag{1.61}$$

which relate the momentum, energy and the angles of the incoming and outgoing particles in an scattering process of two particles to two particles. The importance of the Mandelstam variables reside in that they form a minimum set of variables needed to describe the kinematics of this scattering process; they are Lorentz invariant which makes them very useful when doing calculations.

The tHproduction, where Higgs boson can be radiated either from the top quark or from the W boson, is represented by the leading order Feynman diagrams in Figure 1.15. The cross section for the tHprocess is calculated, as usual, summing over the contributions from the different Feynman diagrams; therefore it depends on the interference between the contributions. In the SM, the interference for t-channel (tHq process) and W-associated (tHWprocess) production is destructive [49] resulting in the small cross sections presented in Table 1.10.

tH production channel	Cross section (fb)
t-channel ($pp \to tHq$)	$70.79^{+2.99}_{-4.80}$
W-associated ($pp \to tHW$)	$15.61^{+0.83}_{-1.04}$
s-channel($pp \to tHb$)	$2.87^{+0.09}_{-0.08}$

Table 1.10: Predicted SM cross sections for tHproduction at $\sqrt{s} = 13$ TeV [53, 54].

The s-channel contribution can be neglected. It will be shown that a deviation from the SM destructive interference would result in an enhancement of the tHcross section compared to that in SM, which could be used to get information about the sign of the Higgs-top coupling [51, 52]. In order to describe tHproduction processes, Feynman diagram 1.15b will be considered; there, the W boson is radiated by a quark in the proton and eventually it will interact with the b quark. In the high energy regime, the effective W approximation [55] is used to describe the process as the emission of an approximately on-shell W and its hard scattering with the b quark; i.e. $Wb \to th$. The scattering amplitude for the process is given by

$$\mathcal{A} = \frac{g}{\sqrt{2}} \left[(\kappa_t - \kappa_V) \frac{m_t \sqrt{s}}{m_W v} A\left(\frac{t}{s}, \varphi; \xi_t, \xi_b\right) + \left(\kappa_V \frac{2m_W}{v}\frac{s}{t} + (2\kappa_t - \kappa_V)\frac{m_t^2}{m_W v}\right) B\left(\frac{t}{s}, \varphi; \xi_t, \xi_b\right) \right], \tag{1.62}$$

where $\kappa_V \equiv g_{HVV}/g_{HVV}^{SM}$ and $\kappa_t \equiv g_{Ht}/g_{Ht}^{SM} = y_t/y_t^{SM}$ are scaling factors that quantify possible deviations of the couplings from the SM values, Higgs-Vector boson (H-W) and Higgs-top (H-t) respectively, from the SM couplings; $s = (p_W + p_b)^2$, $t = (p_W - p_H)^2$, φ is the Higgs azimuthal angle around the z axis taken parallel to the direction of motion of the incoming W; A and B are functions describing the weak interaction in terms of the chiral states (ξ_t, ξ_b) of the quarks b and t. Terms that vanish in the high energy limit have been neglected as well as the Higgs and b quark masses[9].

The scattering amplitude grows with energy like $\sqrt{s}$ for $\kappa_V \neq \kappa_t$, in contrast to the SM ($\kappa_t = \kappa_V = 1$), where the first term in 1.62 cancels out and the amplitude is constant for large s; therefore, a deviation from the SM predictions represents an enhancement in the tHq cross section. In particular, for a SM H-W coupling and a H-t coupling of inverted sign with respect to the SM ($\kappa_V = -\kappa_t = 1$) the tHq cross section is enhanced by a factor greater 10 as seen in the Figure 1.16 taken from Reference [51]; Reference [56] has reported similar enhancement results.

A similar analysis is valid for the W-associated channel but, in that case, the interference is more complicated since there are more than two contributions and an additional interference with the production of Higgs boson and a top pair process($t\bar{t}H$). The calculations are made using the so-called Diagram Removal (DR) technique where

[9] A detailed explanation of the structure and approximations used to derive $\mathcal{A}$ can be found in Reference [51]

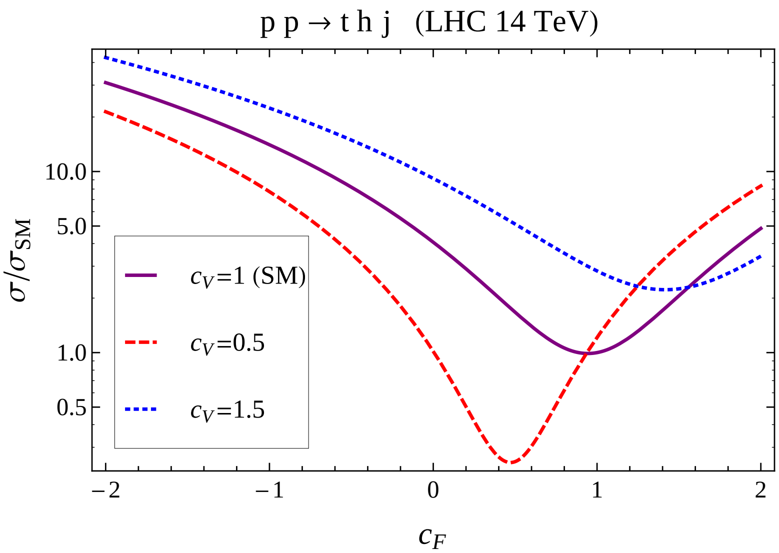

Figure 1.16: Cross section for tHq process as a function of κ_t, normalized to the SM, for three values of κ_V. In the plot c_f refers to the Higgs-fermion coupling which is dominated by the H-t coupling and represented here by κ_t. Solid, dashed and dotted lines correspond to $c_V \rightarrow \kappa_V = 1, 0.5, 1.5$ respectively. Note that for the SM ($\kappa_V = \kappa_t = 1$), the destructive effect of the interference is maximal.

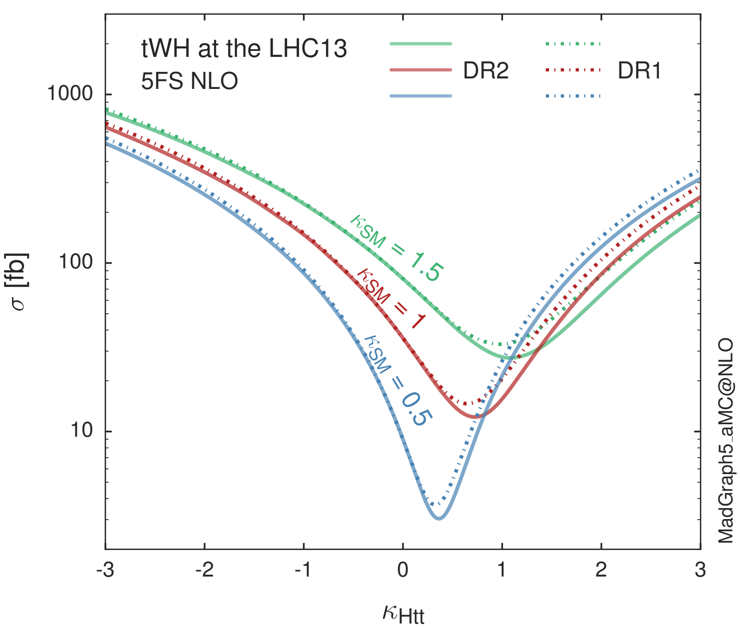

Figure 1.17: Cross section for tHW process as a function of κ_{Htt}, for three values of κ_{SM} at $\sqrt{s} = 13$ TeV. $\kappa^2_{Htt} = \sigma_{Htt}/\sigma^{SM}_{Htt}$ is a simple re-scaling of the SM Higgs interactions.

interfering diagrams are removed (or added) from the calculations in order to evaluate the impact of the removed contributions. DR1 was defined to neglect $t\bar{t}H$ interference while DR2 was defined to take $t\bar{t}H$ interference into account [48]. As shown in Figure 1.17, the tHW cross section is enhanced from about 15 fb (SM: $\kappa_{Htt} = 1$) to about 150 fb ($\kappa_{Htt} = -1$). Differences between curves for DR1 and DR2 help to gauge the impact of the interference with $t\bar{t}H$. Results of the calculations of the tHq and tHW cross sections at $\sqrt{s} = 13$ TeV can be found in Reference [57] and a summary of the results is presented in Table 1.11.

	$\sqrt{s}$ TeV	$\kappa_t = 1$	$\kappa_t = -1$
$\sigma^{LO}(tHq)$(fb) [51]	8	≈ 17.4	≈ 252.7
	14	≈ 80.4	≈ 1042
$\sigma^{NLO}(tHq)$(fb) [51]	8	$18.28^{+0.42}_{-0.38}$	$233.8^{+4.6}_{-0.0}$
	14	$88.2^{+1.7}_{-0.0}$	$982.8^{+28}_{-0.0}$
$\sigma^{LO}(tHq)$(fb) [56]	14	≈ 71.8	≈ 893
$\sigma^{LO}(tHW)$(fb) [56]	14	≈ 16.0	≈ 139
$\sigma^{NLO}(tHq)$(fb) [57]	8	$18.69^{+8.62\%}_{-17.13\%}$	-
	13	$74.25^{+7.48\%}_{-15.35\%}$	$848^{+7.37\%}_{-13.70\%}$
	14	$90.10^{+7.34\%}_{-15.13\%}$	$1011^{+7.24\%}_{-13.39\%}$
$\sigma^{LO}(tHW)$(fb) [48]	13	$15.77^{+15.91\%}_{-15.76\%}$	-
$\sigma^{NLO}DR1(tHW)$(fb) [48]	13	$21.72^{+6.52\%}_{-5.24\%}$	≈ 150
$\sigma^{NLO}DR2(tHW)$(fb) [48]	13	$16.28^{+7.34\%}_{-15.13\%}$	≈ 150

Table 1.11: Predicted enhancement of the tHq and tHW cross sections at LHC for $\kappa_V = 1$ and $\kappa_t = \pm 1$ at LO and NLO; the cross section enhancement of more that a factor of 10 is due to the flipping in the sign of the H-t coupling with respect to the SM one.

1.6 CP-mixing in tH processes

In addition to the sensitivity to sign of the H-t coupling, the tHq and tHW processes have been proposed as a tool to investigate the possibility of a H-t coupling that does not conserve CP [47, 48, 58].

In this thesis, the sensitivity of tHprocesses to CP-mixing is also studied on the basis of References [47, 48] using the effective field theory framework where a generic particle (X_0) of spin-0 and a general CP violating interaction with the top quark (Htt coupling), can couple to scalar and pseudo-scalar fermionic densities. The H-W interaction is assumed to be SM-like. The Lagrangian modeling the H-t interaction is given by

$$\mathcal{L}_0^t = -\bar{\psi}_t \left(c_\alpha \kappa_{Htt} g_{Htt} + i s_\alpha \kappa_{Att} g_{Att} \gamma_5 \right) \psi_t X_0, \tag{1.63}$$

where α is the CP-mixing phase, $c_\alpha \equiv \cos\alpha$ and $s_\alpha \equiv \sin\alpha$, κ_{Htt} and κ_{Att} are real dimensionless re-scaling parameters[10] used to parametrize the magnitude of the CP-violating and CP-conserving parts of the amplitude. The model defines $g_{Htt} = g_{Att} = m_t/v = y_t/\sqrt{2}$ with $v \sim 246$ GeV the Higgs vacuum expectation value. In this parametrization, three special cases can be recovered

- CP-even coupling $\rightarrow \alpha = 0^o$
- CP-odd coupling $\rightarrow \alpha = 90^o$
- SM coupling $\rightarrow \alpha = 0^o$ and $\kappa_{Htt} = 1$

The loop induced X_0 coupling to gluons can also be described in terms of the parametrization above, according to

$$\mathcal{L}_0^g = -\frac{1}{4} \left(c_\alpha \kappa_{Hgg} g_{Hgg} G^a_{\mu\nu} G^{a,\mu\nu} + s_\alpha \kappa_{Agg} g_{Agg} G^a_{\mu\nu} \widetilde{G}^{a,\mu\nu} \right) X_0. \tag{1.64}$$

where $g_{Hgg} = -\alpha_s/3\pi v$ and $g_{Agg} = \alpha_s/2\pi v$ and $G_{\mu\nu}$ is the gluon field strength tensors. Under the assumption that the top quark dominates the gluon-fusion process at LHC

[10] analog to κ_t and κ_V

energies, $\kappa_{Hgg} \to \kappa_{Htt}$ and $\kappa_{Agg} \to \kappa_{Att}$, so that the ratio between the gluon-gluon fusion cross section for X_0 and for the SM Higgs prediction can be written as

$$\frac{\sigma^{gg\to X_0}_{NLO}}{\sigma^{gg\to H}_{NLO,SM}} = c^2_\alpha \kappa^2_{Htt} + s^2_\alpha \left(\kappa_{Att} \frac{g_{Agg}}{g_{Hgg}}\right)^2. \tag{1.65}$$

If the re-scaling parameters are set to

$$\kappa_{Htt} = 1, \qquad \kappa_{Att} = \left|\frac{g_{Hgg}}{g_{Agg}}\right| = \frac{2}{3}. \tag{1.66}$$

the gluon-fusion SM cross section is reproduced for every value of the CP-mixing angle α; therefore, by imposing that condition to the Lagrangian density 1.63, the CP-mixing angle is not constrained by current data. Figure 1.18 shows the NLO cross sections for t-channel tX_0(blue) and $t\bar{t}X_0$ (red) associated production processes as a function of the CP-mixing angle α. X_0 is a generic spin-0 particle with top quark CP-violating coupling. Re-scaling factors κ_{Htt} and κ_{Att} have been set to reproduce the SM gluon-fusion cross sections.

It is interesting to notice that the tX_0 cross section is enhanced, by a factor of about 10, when a continuous rotation in the scalar-pseudoscalar plane is applied; this enhancement is similar to the enhancement produced when the H-t coupling is flipped in sign with respect to the SM ($y_t = -y_{t,SM}$ in the plot), as showed in Section 1.5. In contrast, the degeneracy in the $t\bar{t}X_0$ cross section is still present given that it depends quadratically on the H-t coupling, but more interesting is to notice that $t\bar{t}X_0$ cross section is exceeded by tX_0 cross section after $\alpha \sim 60^o$.

A similar parametrization can be used to investigate the tHWprocess sensitivity to CP-violating H-t coupling. As said in 1.5, the interference in the W-associated channel is more complicated because there are more than two contributions and also there is interference with the $t\bar{t}H$ production process.

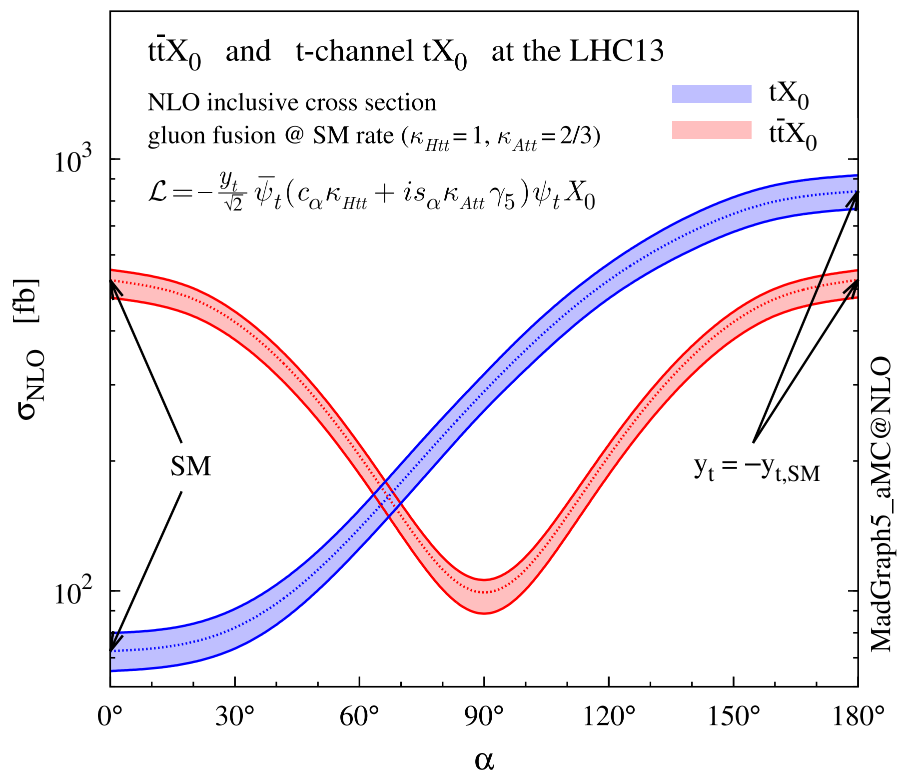

Figure 1.18: NLO cross sections for t-channel tX_0(blue) and $t\bar{t}X_0$ (red) associated production processes as a function of the CP-mixing angle α. X_0 is a generic spin-0 particle with top quark CP-violating coupling [47].

Figure 1.19 shows the NLO cross sections for t-channel tX_0(blue), $t\bar{t}X_0$ (red) associated production and for the combined $tWX_0+t\bar{t}X_0+interference$ (orange) as a function of the CP-mixing angle. It is clear that the effect of the interference in the combined case is the lifting of the degeneracy present in the $t\bar{t}X_0$ production. The constructive interference enhances the cross section from about 500 fb at SM ($\alpha = 0$) to about 600 fb ($\alpha = 180^o \rightarrow y_t = -y_{t,SM}$).

An analysis combining tHq and tHW processes will be made in this thesis taking advantage of the sensitivity improvement.

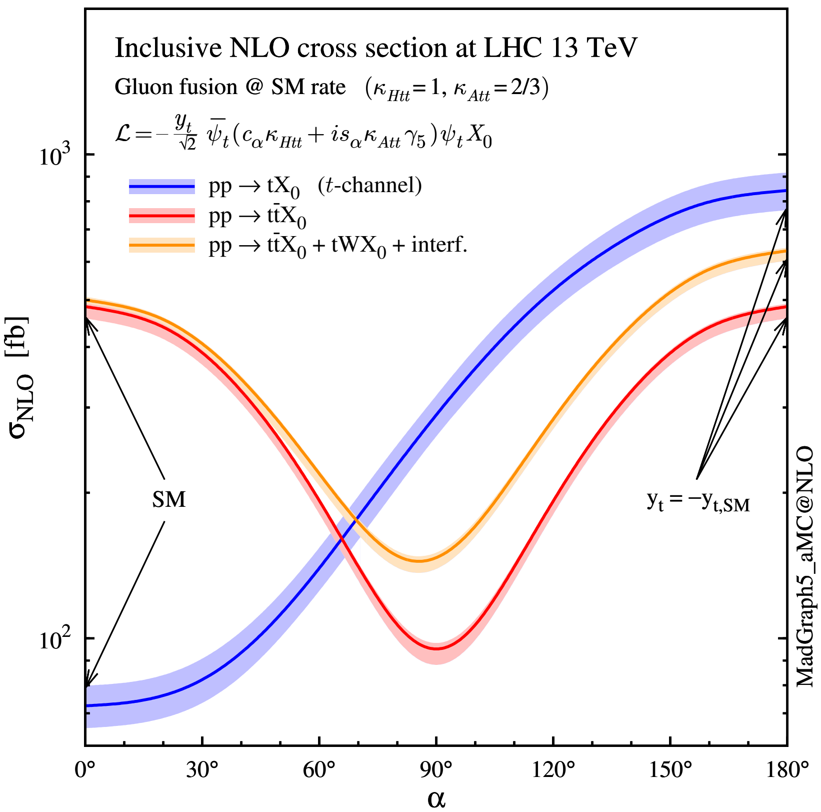

Figure 1.19: NLO cross sections for t-channel tX_0(blue), $t\bar{t}X_0$ (red) associated production processes and combined $tWX_0 + t\bar{t}X_0$ (including interference) production as a function of the CP-mixing angle α [47].

Chapter 2

The CMS experiment at the LHC

2.1 Introduction

Located on the Swiss-French border, the European Council for Nuclear Research (CERN) is the largest scientific organization leading particle physics research. About 13000 people in a broad range of roles including users, students, scientists, engineers, among others, contribute to the data taking and analysis, with the goal of unveiling the secrets of nature and revealing the fundamental structure of the universe. CERN is also the home of the Large Hadron Collider (LHC), the largest particle accelerator around the world, where protons (or heavy ions) traveling close to the speed of light, are made to collide. These collisions open a window to investigate how particles (and their constituents if they are composite) interact with each other, providing clues about the laws of nature. This chapter presents an overview of the LHC structure and operation. A detailed description of the CMS detector is offered, given that the data used in this thesis have been taken with this detector.

2.2 The LHC

With 27 km of circumference, the LHC is currently the most powerful circular accelerator in the world. It is installed in the same tunnel where the Large Electron-Positron (LEP) collider was located, taking advantage of the existing infrastructure. The LHC is part of the CERN's accelerator complex composed of several successive accelerating stages before the particles are injected into the LHC ring where they reach their maximum energy (see Figure 2.1).

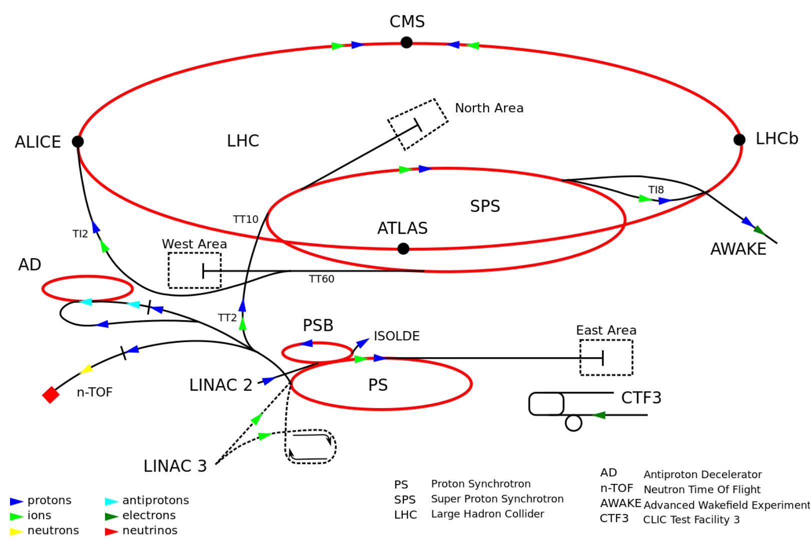

Figure 2.1: CERN accelerator complex. Blue arrows show the path followed by protons along the acceleration process [61].

The LHC runs in three collision modes depending on the particles being accelerated

- Proton-Proton collisions (*pp*) for multiple physics experiments.
- Lead-Lead collisions (Pb-Pb) for heavy ion experiments.
- Proton-Lead collisions (p-Pb) for quark-gluon plasma experiments.

In this thesis only *pp* collisions will be considered.

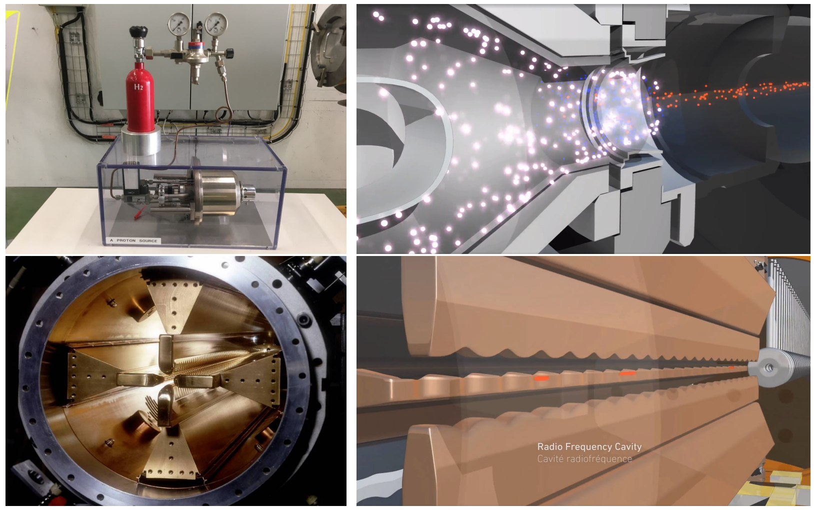

Figure 2.2: LHC protons source and the first acceleration stage. Top: the bottle contains hydrogen gas which is injected into the metal cylinder (white dots) to be broken down into electrons (blue dots) and protons (red dots); Bottom: the obtained protons are directed towards the radio frequency quadrupole which perform the first acceleration, focus the beam and create the bunches of protons. Letf images are real pictures while right images are drawings [65, 66].

Collection of protons starts with hydrogen atoms taken from a bottle, containing hydrogen gas, and injecting them in a metal cylinder; hydrogen atoms are broken down into electrons and protons by an intense electric field (see Figure 2.2 top). The resulting protons leave the metal cylinder towards a radio frequency quadrupole (RFQ) that focus the beam, accelerates the protons and creates the packets of protons called bunches. In the RFQ, an electric field is generated by a RF wave at a frequency that matches the resonance frequency of the cavity where the electrodes are contained. The beam of protons traveling on the RFQ axis experiences an alternating electric field gradient that generates the focusing forces.

In order to accelerate the protons, a longitudinal time-varying electric field component is added to the system; it is done by giving the electrodes a sine-like profile as shown in Figure 2.2 bottom. By matching the speed and phase of the protons with the longitudinal electric field the bunching is performed; protons synchronized with the RFQ (synchronous protons) do not feel an accelerating force, but those protons in the beam that have more (or less) energy than the synchronous proton (asynchronous protons) will feel a decelerating (accelerating) force; therefore, asynchronous protons will oscillate around the synchronous ones forming bunches of protons [63]. From the RFQ protons emerge with energy 750 keV in bunches of about 1.15×10^{11} protons [64].

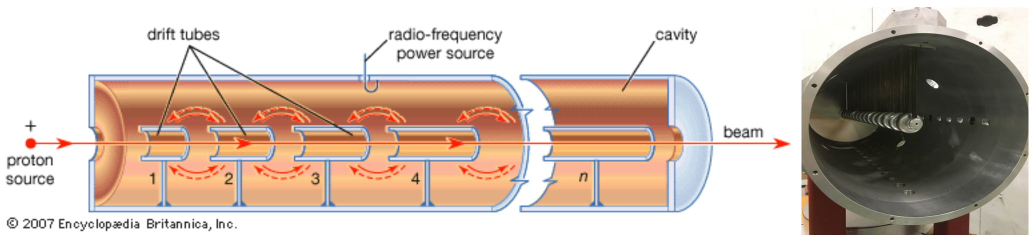

Figure 2.3: Left: drawing of the LINAC2 accelerating system at CERN. Electric fields generated by radio frequency (RF) create acceleration and deceleration zones inside the cavity; deceleration zones are blocked by drift tubes where quadrupole magnets focus the proton beam. Right: picture of a real RF cavity [67].

Proton bunches coming from the RFQ go to the linear accelerator 2 (LINAC2) where they are accelerated to reach 50 MeV energy. In the LINAC2 stage, acceleration is performed using electric fields generated by radio frequency which create zones of acceleration and deceleration as shown in Figure 2.3. In the deceleration zones, the electric field is blocked using drift tubes where protons are free to drift while quadrupole magnets focus the beam.

The beam coming from LINAC2 is injected into the proton synchrotron booster (PSB) to reach 1.4 GeV in energy. The next acceleration is provided at the proton synchrotron (PS) up to 26 GeV, followed by the injection into the super proton synchrotron (SPS) where protons are accelerated to 450 GeV. Finally, protons are injected into the LHC where they are accelerated to the target energy of 6.5 TeV.

PSB, PS, SPS and LHC accelerate protons using the same RF acceleration technique described before.

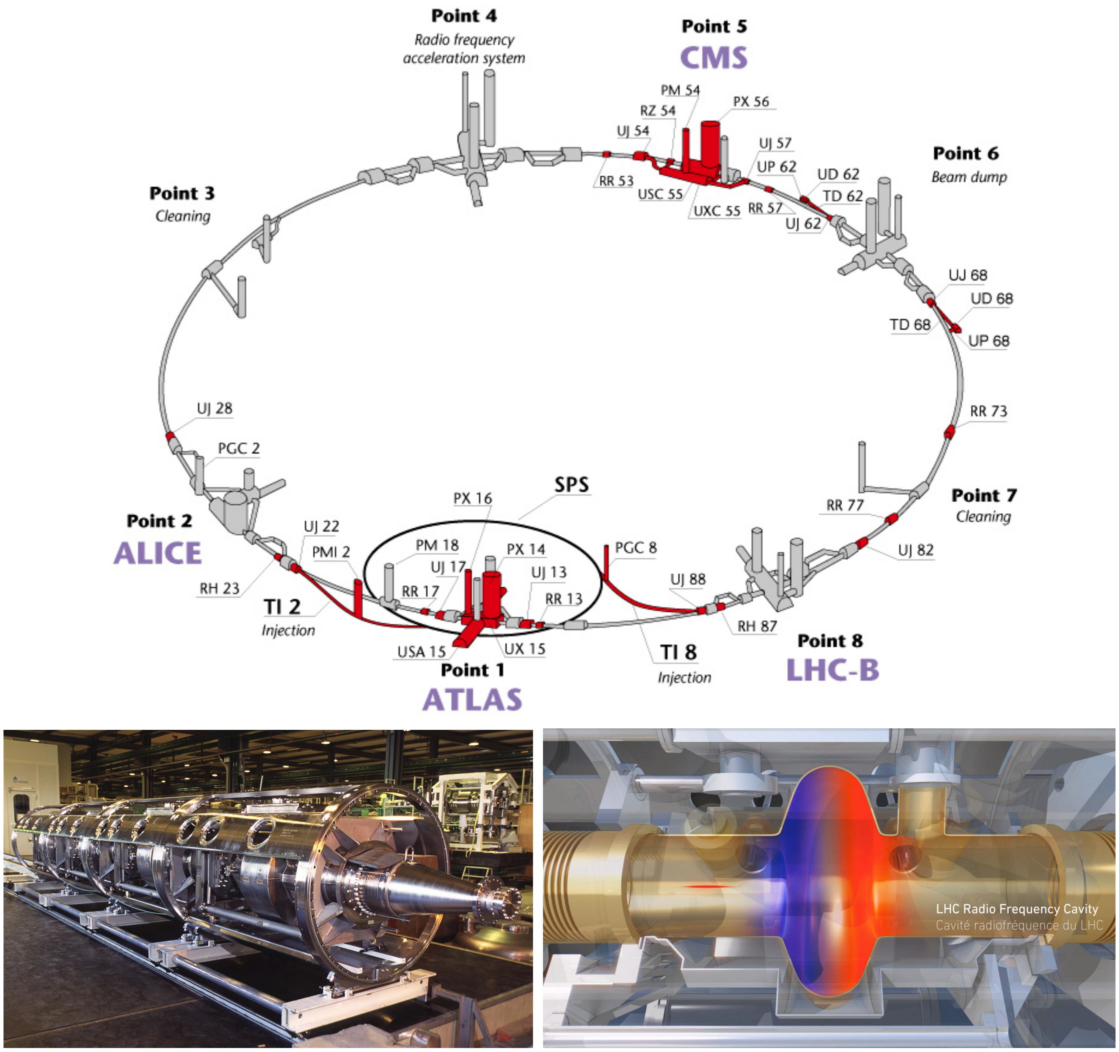

Figure 2.4: Top: LHC layout. The red zones indicate the infrastructure additions to the LEP installations, built to accommodate the ATLAS and CMS experiments which exceed the size of the former experiments located there [62]. Bottom: LHC RF cavities. A module (left picture) accommodates 4 cavities, each like the one in the right drawing, that accelerate protons and preserve the bunch structure of the beam. The color gradient pattern represents the strenght of the electric field inside the cavity; red zone corresponds to the maximum of the oscillation electric field while the blue zore corresponds to the minimum. [66, 68]

The LHC has a system of 16 RF cavities located in the so-called point 4, as shown in Figure 2.4 top, tuned at a frequency of 400 MHz. The bottom side of Figure 2.4 shows a picture of a RF module composed of 4 RF cavities working in a superconducting state at 4.5 K; also, a representation of the accelerating electric field that accelerates the protons in the bunch is shown. The maximum of the oscillating electric field (red region) picks the proton bunches at the the entrance of the cavity and keeps accelerating them through the whole cavity. The protons are carefully timed so that in addition to the acceleration effect the bunch structure of the beam is preserved.

While protons are accelerated in one section of the LHC ring, where the RF cavities are located, in the rest of their path they have to be kept in the curved trajectory defined by the LHC ring. Technically, LHC is not a perfect circle; RF, injection, beam dumping, beam cleaning and sections before and after the experimental points where protons collide are all straight sections. In total, there are 8 arcs 2.45 km long each and 8 straight sections 545 m long each. In order to curve the proton's trajectory in the arc sections, superconducting dipole magnets are used.

Inside the LHC ring, there are two proton beams traveling in opposite directions in two separated beam pipes; the beam pipes are kept at ultra-high vacuum ($\sim 10^{-9}$ Pa) to ensure that there are no particles that interact with the proton beams. The superconducting dipole magnets used in LHC are made of a NbTi alloy, capable of transporting currents of about 12000 A when cooled at a temperature below 2K using liquid helium (see Figure 2.5).

Protons in the arc sections of LHC feel a centripetal force exerted by the dipole magnets; the magnitude of magnetic field needed to keep the protons in the LHC curved trayectory can be found assuming that protons travel at $v \approx c$, using the standard values for proton mass and charge and the LHC radius, as

$$F_m = \frac{mv^2}{r} = qBv \quad \rightarrow B = 8.33T \tag{2.1}$$

which is about 100000 times the Earth's magnetic field. A representation of the

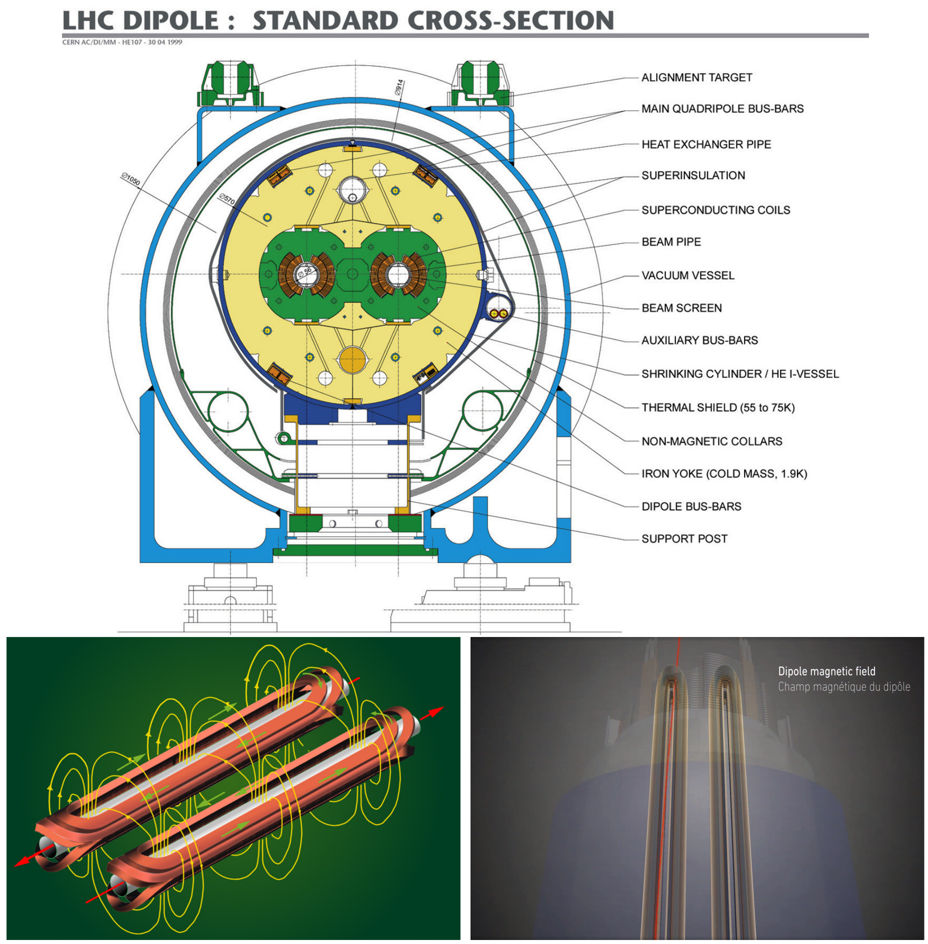

Figure 2.5: Top: LHC dipole magnet transverse view; cooling, shielding and mechanical support are indicated. Bottom left: Magnetic field generated by the dipole magnets; note that the direction of the field inside one beam pipe is opposite with respect to the other beam pipe which guarantee that both proton beams are curved in the same direction towards the center of the ring. The effect of the dipole magnetic field on the proton beam is represented on the bottom right side [66, 69, 70].

magnetic field generated by the dipole magnets is shown on the bottom left side of Figure 2.5. The bending effect of the magnetic field on the proton beam is shown on the bottom right side of Figure 2.5. Note that the dipole magnets are not curved; the arc section of the LHC ring is composed of straight dipole magnets of about 15 m. In total there are 1232 dipole magnets along the LHC ring.

In addition to the bending of the beam trajectory, the beam has to be focused. The focusing is performed by quadrupole magnets installed in a different straight section; in total 858 quadrupole magnets are installed along the LHC ring. Other effects like electromagnetic interaction among bunches, interaction with electron clouds from the beam pipe, the gravitational force on the protons, differences in energy among protons in the same bunch, among others, are corrected using sextupole and other magnetic multipoles.

The two proton beams inside the LHC ring are made of bunches with a cylindrical shape of about 7.5 cm long and about 1 mm in diameter; when bunches are close to the interaction point (IP), the beam is focused up to a diameter of about 16 μm in order to maximize the probability of collisions between protons. The number of collisions per second is proportional to the cross section of the bunches with the *luminosity* (L) as the proportionality factor, thus, the luminosity can be calculated using

$$L = fn\frac{N_1 N_2}{4\pi\sigma_x\sigma_y} \tag{2.2}$$

where f is the revolution frequency, n is the number of bunches per beam, N_1 and N_2 are the numbers of protons per bunch, σ_x and σ_y are the gaussian transverse sizes of the bunches. With the expected parameters, the LHC expected luminosity is about

$$\begin{aligned}
f =& \frac{v}{2\pi r_{LHC}} \approx \frac{3\times 10^8\text{m/s}}{27\text{km}} \approx 11.1\text{kHz},\\
n =& 2808\\
N_1 =& N_2 \sim 1.5\times 10^{11}\\
\sigma_x =& \sigma_y = 16\mu\text{m}
\end{aligned}$$

$$L = 1.28\times 10^{34}\text{cm}^{-2}\text{s}^{-1} = 1.28\times 10^{-5}\text{fb}^{-1}\text{s}^{-1} \tag{2.3}$$

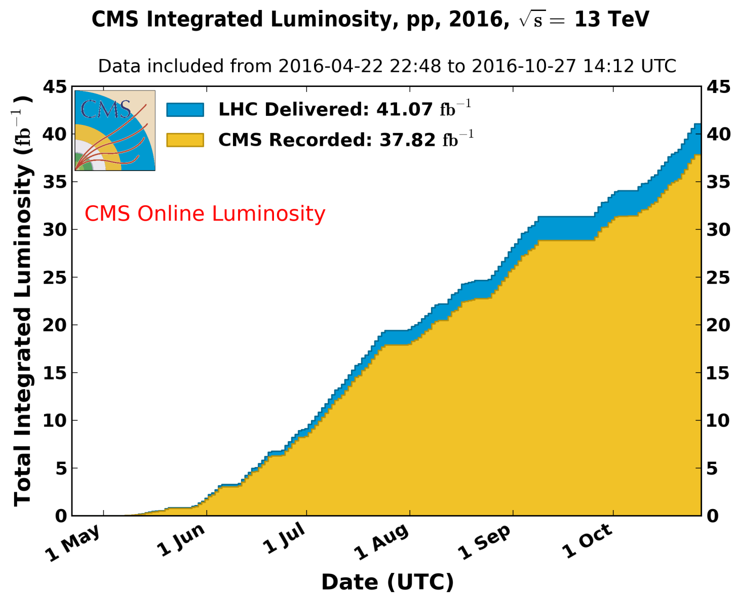

Figure 2.6: Integrated luminosity delivered by LHC and recorded by CMS during 2016. The difference between the delivered and the recorded luminosity is due to fails and issues occurred during the data taking in the CMS experiment [71].

Luminosity is a fundamental aspect of LHC given that the bigger luminosity the bigger number of collisions, which means that for processes with a very small cross section the number of expected occurrences is increased and so the chances of being detected. The integrated luminosity, collected by the CMS experiment during 2016 is shown in Figure 2.6; the data analyzed in this thesis corresponds to an integrated luminosity of 35.9 fb^{-1} at a center of mass-energy $\sqrt{s} = 13$ TeV.

One way to increase L is increasing the number of bunches in the beam. Currently, the separation between two consecutive bunches in the beam is 7.5 m which corresponds to a time separation of 25 ns. In the full LHC ring the allowed number of bunches is $n = 27\text{km}/7.5\text{m} = 3600$; however, there are some gaps in the bunch pattern intended for preparing the dumping and injection of the beam, thus, the proton beams are composed of 2808 bunches.

Once the proton beams reach the desired energy, they are brought to cross each other producing *pp* collisions. The bunch crossing happens in precise places where the four LHC experiments are located, as seen in the top of Figure 2.7. In 2008 *pp* collisions of $\sqrt{s} = 7$ TeV were performed; the energy was increased to 8 TeV in 2012 and to 13 TeV in 2015.

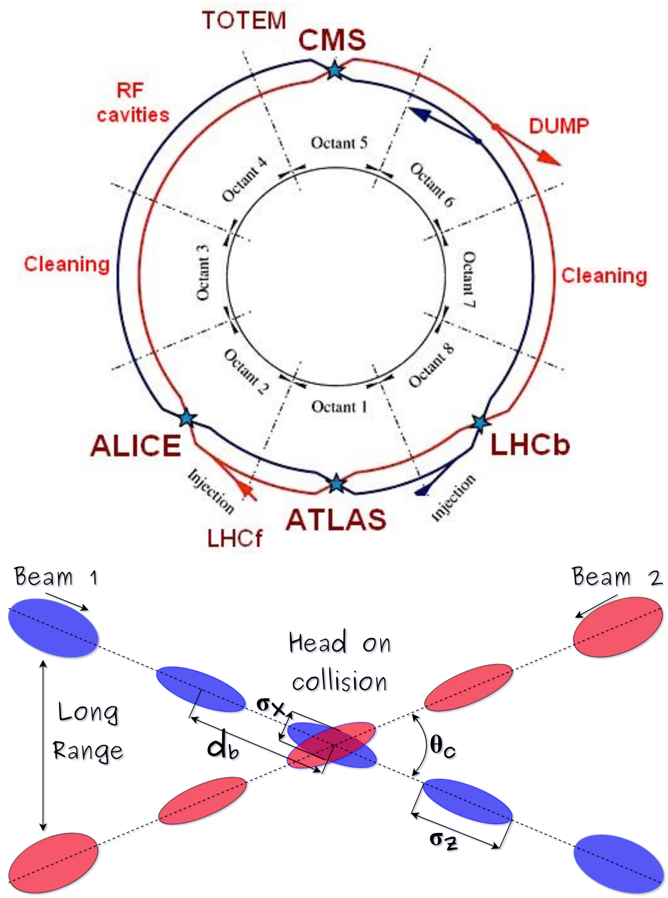

Figure 2.7: Top: LHC interaction points. Bunch crossing occurs where the LHC experiments are located [72]. Sections indicated as cleaning are dedicated to collimate the beam in order to protect the LHC ring from collisions with protons in very spreaded bunches. Bottom: bunch crossing scheme. Since the bunch crossing is not perfectly head-on, the luminosity is reduced in a factor of 17%; adapted from Reference [86].

The CMS and ATLAS experiments are multi-purpose experiments, hence, they are enabled to explore physics in any of the LHC collision modes. LHCb experiment is optimized to explore bottom quark physics, while ALICE is optimized for heavy ion collisions searches; TOTEM and LHCf are dedicated to forward physics studies; MoEDAL (not indicated in the Figure) is intended for monopole or massive pseudo stable particles searches.

At the IP there are two interesting details that need to be addressed. The first one is that the bunch crossing does not occur head-on but at a small crossing angle θ_c (280 μrad in CMS and ATLAS) as shown in the bottom side of Figure 2.7, affecting the overlapping between bunches; the consequence is a reduction of about 17% in the luminosity (represented by a factor not included in eqn. 2.2). The second one is the occurrence of multiple *pp* collisions in the same bunch crossing; this effect is called *pileup* (PU). A fairly simple estimation of the PU follows from estimating the probability of collision between two protons, one from each of the bunches in the course of collision; it depends roughly on the ratio of proton size and the cross section of the bunch in the IP, i.e.,

$$P(pp-collision) \sim \frac{d^2_{proton}}{\sigma_x \sigma_y} = \frac{(1\text{fm})^2}{(16\mu\text{m})^2} \sim 4 \times 10^{-21} \tag{2.4}$$

however, there are $N = 1.15 \times 10^{11}$ protons in a bunch, thus the estimated number of collisions in a bunch crossing is

$$PU = N^2 * P(pp-collision) \sim 50pp \text{ collision per bunch crossing,} \tag{2.5}$$

about 20 of which are inelastic. A multiple *pp* collision event in a bunch crossing at CMS is shown in Figure 2.8.

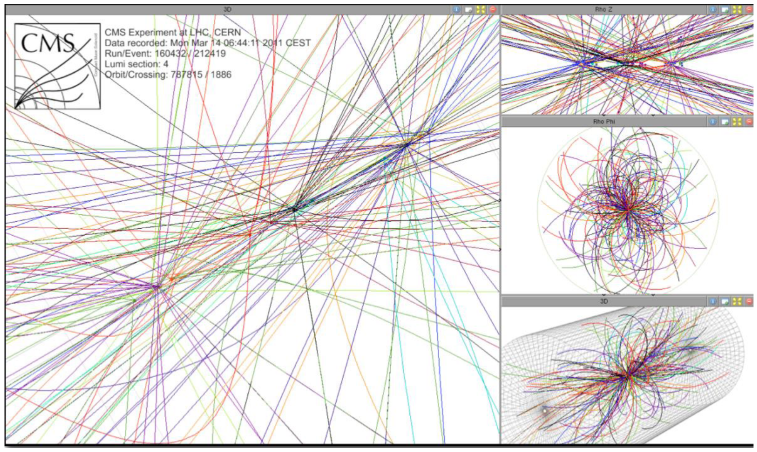

Figure 2.8: Multiple *pp* collision bunch crossing at CMS. [73].

2.3 The CMS experiment

CMS is a general-purpose detector designed to conduct research in a wide range of physics from the standard model to new physics like extra dimensions and dark matter. Located at Point 5 in the LHC layout as shown in Figure 2.4, CMS is composed of several detection systems distributed in a cylindrical structure; in total, CMS weights about 12500 tons in a very compact 21.6 m long and 14.6 m diameter cylinder. It was built in 15 separate sections at the ground level and lowered to the cavern individually to be assembled. A complete and detailed description of the CMS detector and its components is given in Reference [74] on which this section is based. Figure 2.9 shows the layout of the CMS detector. The detection system is composed of (from the innermost to the outermost)

- Pixel detector.
- Silicon strip tracker.
- Preshower detector.
- Electromagnetic calorimeter.

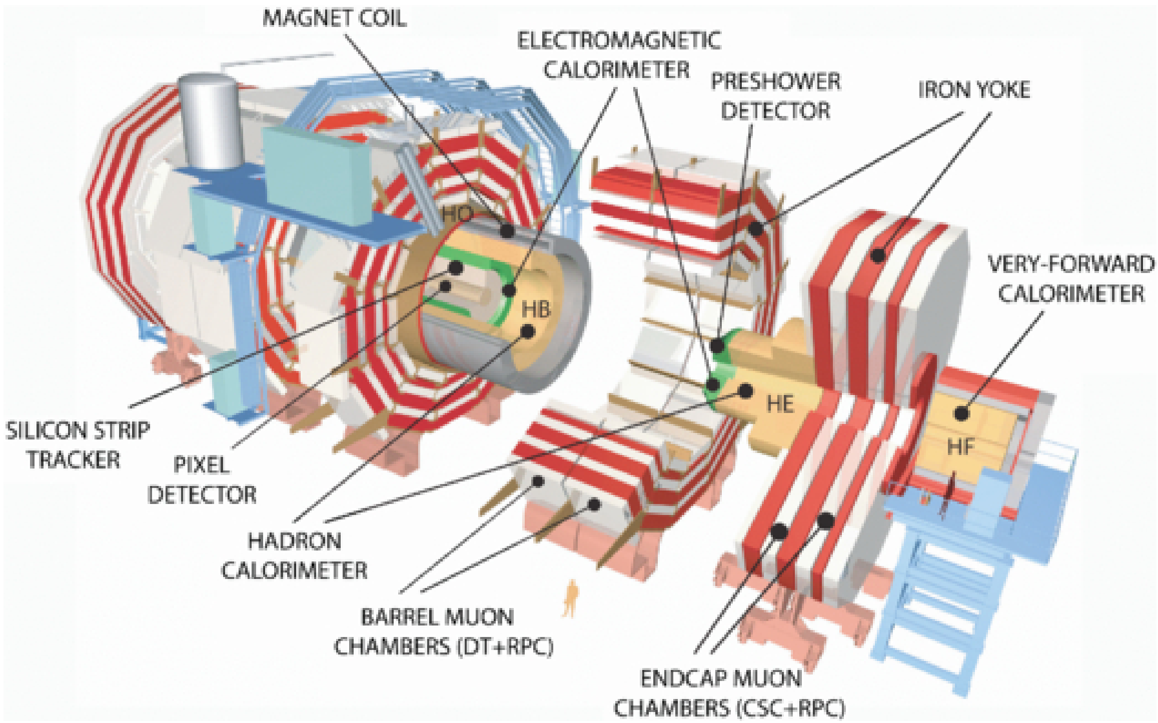

Figure 2.9: Layout of the CMS detector. The several subdetectors are indicated. The central region of the detector is referred as the barrel section while the endcaps are referred as the forward sections. [75].

- Hadronic calorimeter.
- Muon chambers (barrel and endcap)

The central region of the detector is commonly referred as the barrel section while the endcaps are referred as the forward sections of the detector; thus, each subdetector is composed of a barrel section and a forward section.

When a *pp* collision happens inside the CMS detector, many different particles are produced, but only some of them live long enough to be detected; they are electrons, photons, pions, kaons, protons, neutrons and muons; neutrinos are not detected by the CMS detector. Thus, the CMS detector was designed to detect those particles and measure their properties. Figure 2.10 shows a transverse slice of the CMS detector. The silicon tracker (pixel detector + strip tracker) is capable to register the track of the charged particles traversing it, while calorimeters (electromagnetic and hadronic) measure the energy of the particles that are absorbed by their materials. Considering the detectable particles, mentioned above, emerging from the IP, a basic description of the detection process is as follows.

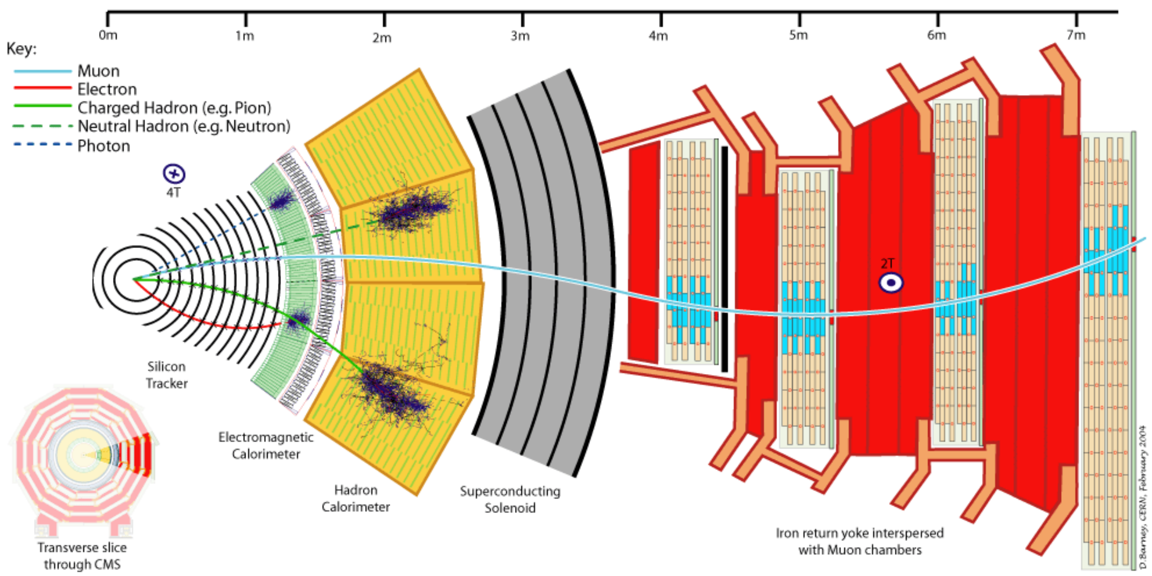

Figure 2.10: CMS detector transverse slice [76].

A muon emerging from the IP, will create a track on the silicon tracker and on the muon chambers. The design of the CMS detector is driven by the requirements on the identification, momentum resolution and unambiguous charge determination of the muons; therefore, a large bending power is provided by the solenoid magnet made of superconducting cable capable of generating a 3.8 T magnetic field. The muon track is bent twice since the magnetic field inside the solenoid is directed along the z-direction but outside its direction is reversed. Muons interact very weakly with the calorimeters, therefore, it is not absorbed but escape away from the detector.

An electron emerging from the IP will create a track along the tracker which will be bent due to the presence of the magnetic field, later, it will be absorbed in the electromagnetic calorimeter where its energy is measured.

A photon will not leave a track because it is neutral, but it will be absorbed in the electromagnetic calorimeter.

A neutral hadron, like the neutron, will not leave a track either but it will lose a small amount of its energy during its passage through the electromagnetic calorimeter and then it will be absorbed in the hadron calorimeter depositing the rest of its energy.

A charged hadron, like the proton or $\pi^{\pm}$, will leave a curved track on the silicon tracker, some of its energy in the electromagnetic calorimeter and finally will be absorbed in the hadronic calorimeter.

A more detailed description of each detection system will be presented in the following sections.

2.3.1 CMS coordinate system

The coordinate system used by CMS is centered on the geometrical center of the detector which is the nominal IP as shown in Figure 2.11[1]. The z-axis is parallel to the beam direction, while the Y-axis pointing vertically upward, and the X-axis pointing radially inward toward the center of the LHC.

In addition to the common cartesian and cylindrical coordinate systems, two coordinates are of particular utility in particle physics: rapidity (y) and pseudorapidity (η), defined in connection to the polar angle θ, energy and longitudinal momentum component (momentum along the z-axis) according to

$$y = \frac{1}{2} ln \frac{E + p_z}{E - p_z} \qquad \eta = -ln\left(tan\frac{\theta}{2}\right) \tag{2.6}$$

Rapidity is related to the angle between the XY-plane and the direction in which the products of a collision are emitted; it has the nice property that the difference

[1] Not all the pp interaction occur at the nominal IP because of the bunch lenght, therefore, each pp collision has it own IP location

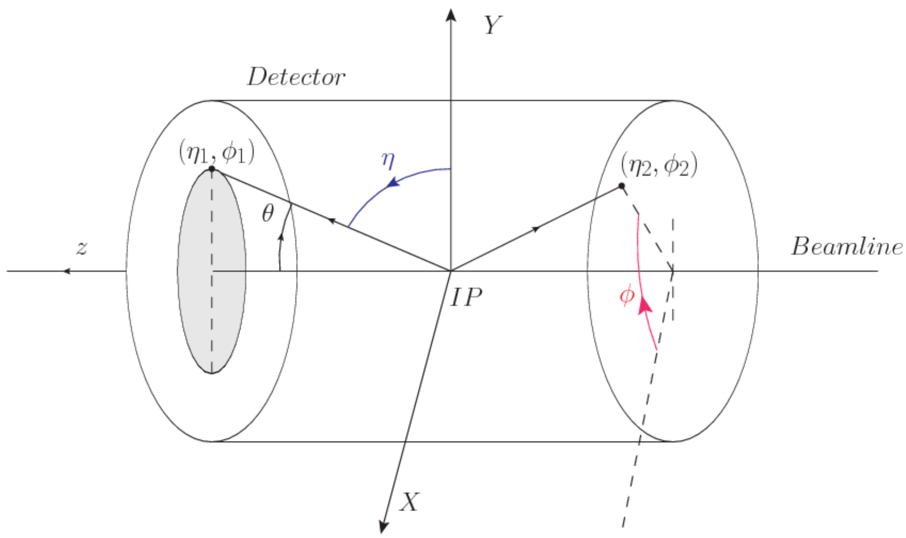

Figure 2.11: CMS detector coordinate system.

between the rapidities of two particles is invariant with respect to Lorentz boosts along the z-axis, hence, data analysis becomes more simple when based on rapidity; however, it is not simple to measure the rapidity of highly relativistic particles, as those produced after pp collisions. Under the highly relativistic motion approximation, y can be rewritten in terms of the polar angle, concluding that rapidity is approximately equal to the pseudorapidity defined above, i.e., $y \approx \eta$. Note that η is easier to measure that y given the direct relationship between the former and the polar angle.

The angular distance between two objects in the detector (ΔR) is commonly used to judge the isolation of those object; it is defined in terms of their coordinates (η_1, ϕ_1), (η_2, ϕ_2) as

$$\Delta R = \sqrt{(\Delta\eta)^2 - (\Delta\phi)^2} \tag{2.7}$$

2.3.2 Tracking system

The CMS tracking system is designed to provide a precise measurement of the trajectories (*track*) followed by the charged particles created after the pp collisions; also, the precise reconstruction of the primary and secondary origins of the tracks (*vertices*) is expected in an environment where, each 25 ns, the bunch crossing produces about 20 inelastic collisions and about 1000 particles.

Physics requirements guiding the tracking system performance include the precise characterization of events involving gauge bosons, W and Z, and their leptonic decays for which an efficient isolated lepton and photon reconstruction is of capital importance, given that isolation is required to suppress background events to a level that allows observations of interesting processes like Higgs boson decays or beyond SM events.

The ability to identify and reconstruct b-jets and B-hadrons within these jets is also a fundamental requirement, achieved through the ability to reconstruct accurately displaced vertices, given that b-jets are part of the signature of top quark physics, like the one treated in this thesis.

An schematic view of the CMS tracking system is shown in Figure 2.12

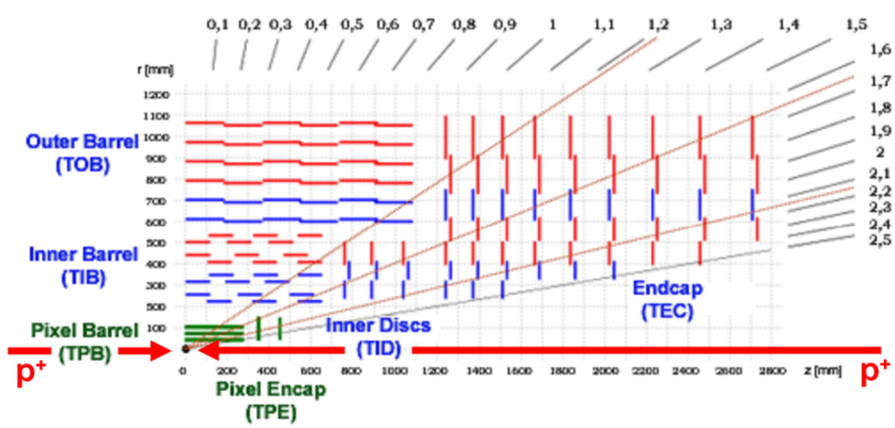

Figure 2.12: CMS tracking system schematic view [78].

In order to satisfy these performance requirements, the tracking system uses two different detector subsystems arranged in concentric cylindrical volumes, the pixel detector and the silicon strip tracker; the pixel detector is located in the high particle density region (r<20cm) while the silicon strip tracker is located in the medium and lower particle density regions 20cm<r<116cm.

Pixel detector

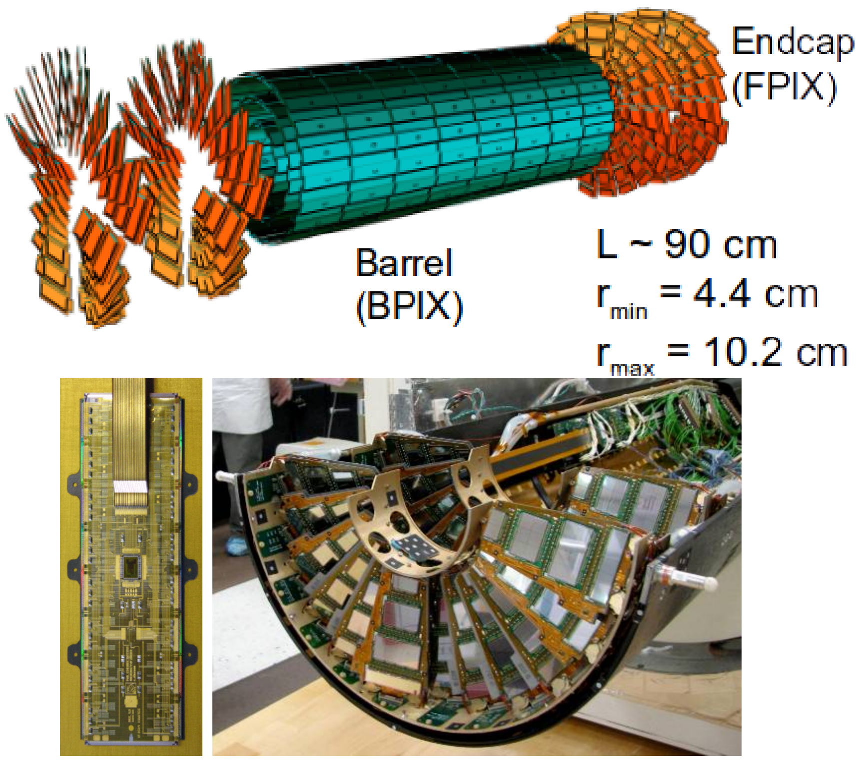

Figure 2.13: CMS pixel detector. Top: schematic view; Bottom: pictures of a barrel(BPIX) module(left) and forward modules(right) [74].

The pixel detector was replaced during the 2016-2017 extended year-end technical stop, due to the increasingly challenging operating conditions like the higher particle flux and more radiation harsh environment, among others. The new one is responding as expected, reinforcing its crucial role in the successful way to fulfill the new LHC physics objectives after the discovery of the Higgs boson. Since the data sets used in this thesis were produced using the previous version of the pixel detector, it will be the subject of the description in this section. The last chapter of this thesis is dedicated to describe my contribution to the *Forward Pixel Phase 1 upgrade*.

The pixel detector was composed of 1440 silicon pixel detector modules organized in three-barrel layers in the central region (BPix) and two disks in the forward region (FPix) as shown in the top side of Figure 2.13; it was designed to record efficiently and with high precision, up to 20μm in the XY-plane and 20μm in the z-direction, the first three space-points (*hits*) nearest to the IP region in the range $|\eta| \leq 2.5$. The first barrel layer was located at a radius of 44 mm from the beamline, while the third layer was located at a radius of 102 mm, closer to the strip tracker inner barrel layer (see Section 2.3.3) in order to reduce the rate of fake tracks. The high granularity of the detector is represented in its about 66 Mpixels, each of size $100 \times 150\mu\text{m}^2$. The transverse momentum resolution of tracks can be measured with a resolution of 1-2% for muons of $p_T = 100$ GeV.

A charged particle passing through the pixel sensors produce ionization in them, giving energy for electrons to be removed from the silicon atoms, hence, creating electron-hole pairs. The collection of charges in the pixels generates an electrical signal that is read out by an electronic readout chip (ROC); each pixel has its own electronics which amplifies the signal. Combining the signal from the pixels activated by a traversing particle in the several layers of the detector allows one to reconstruct the particle's trajectory in 3D.

Commonly, the charge produced by traversing of a particle is collected by and shared among several pixels; by interpolating between pixels, the spatial resolution is improved. In the barrel section the charge sharing in the rϕ-plane is due to the Lorentz effect. In the forward pixels the charge sharing is enhanced by arranging the blades in the turbine-like layout as shown in Figure 2.13 bottom left.

2.3.3 Silicon strip tracker

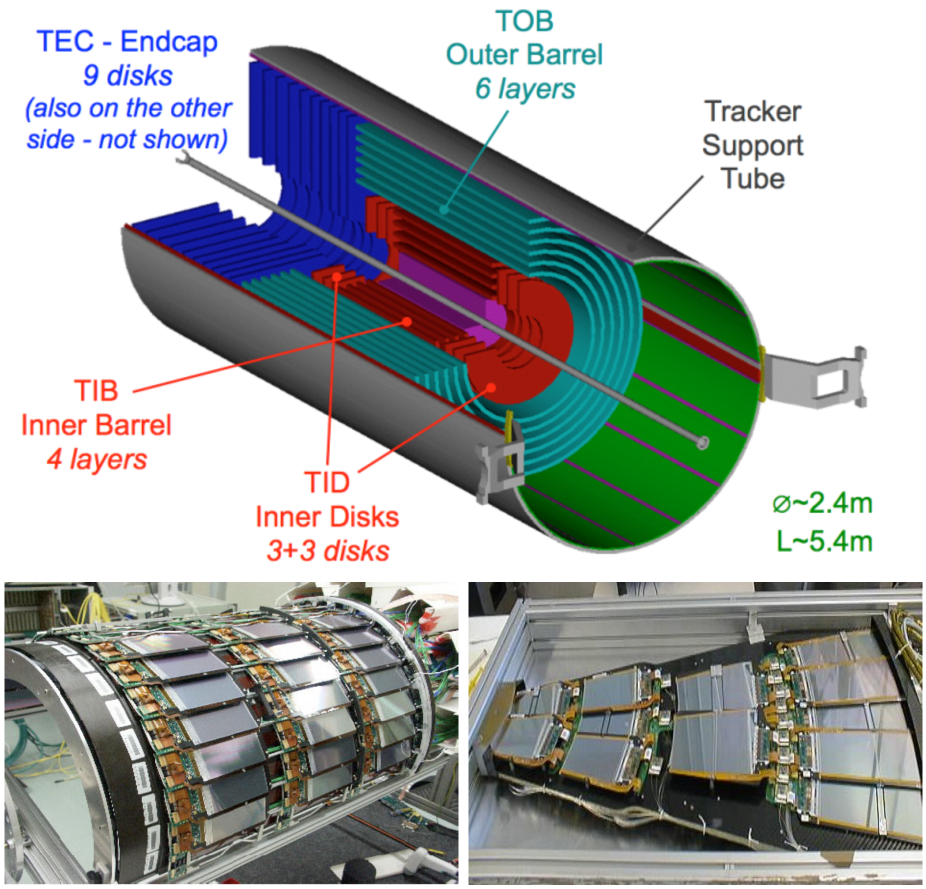

Figure 2.14: Top: CMS Silicon Strip Tracker (SST) schematic view. The SST is composed of the tracker inner barrel (TIB), the tracker inner disks (TID), the tracker outer barrel (TOB) and the tracker endcaps (TEC). Each part is made of silicon strip modules; the modules in blue represent two modules mounted back-to-back and rotated in the plane of the module by a stereo angle of 120 mrad in order to provide a 3-D reconstruction of the hit positions. Bottom: pictures of the TIB (left) and TEC (right) modules [80–82].

The silicon strip tracker (SST) is the second stage in the CMS tracking system. The top side of Figure 2.14 shows a schematic of the SST. The inner tracker region is composed of the tracker inner barrel (TIB) and the tracker inner disks (TID) covering the region $r < 55$ cm and $|z| < 118$ cm. The TIB is composed of 4 layers while the TID is composed of 3 disks at each end. The silicon sensors in the inner tracker are 320 μm thick, providing a resolution of about 13-38 μm in the $r\phi$ position measurement.

The modules indicated in blue in the schematic view of Figure 2.14 are two modules mounted back-to-back and rotated in the plane of the module by a *stereo* angle of 100 mrad; the hits from these two modules, known as *stereo hits*, are combined to provide a measurement of the second coordinate (z in the barrel and r on the disks) allowing the reconstruction of hit positions in 3-D.

The outer tracker region is composed of the tracker outer barrel (TOB) and the tracker endcaps (TEC). The six layers of the TOB offer coverage in the region $r > 55$ cm and $|z| < 118$ cm, while the 9 disks of the TEC cover the region $124 < |z| < 282$ cm. The resolution offered by the outer tracker is about 13-38 μm in the $r\phi$ position measurement. The inner four TEC disks use silicon sensors 320 μm thick; those in the TOB and the outer three TEC disks use silicon sensors of 500 μm thickness. The silicon strips run parallel to the z-axis and the distance between strips varies from 80 μm in the inner TIB layers to 183 μm in the inner TOB layers; in the endcaps the wedge-shaped sensors with radial strips, whose pitch range between 81 μm at small radii and 205 μm at large radii.

The whole SST has 15148 silicon modules, 9.3 million silicon strips and cover a total active area of about 198 m^2.

2.3.4 Electromagnetic calorimeter

The CMS electromagnetic calorimeter (ECAL) is designed to measure the energy of electrons and photons. It is composed of 75848 lead tungstate crystals which have a short radiation length (0.89 cm) and fast response, since 80% of the light is emitted within 25 ns; however, they are combined with Avalanche photodiodes (APDs) as

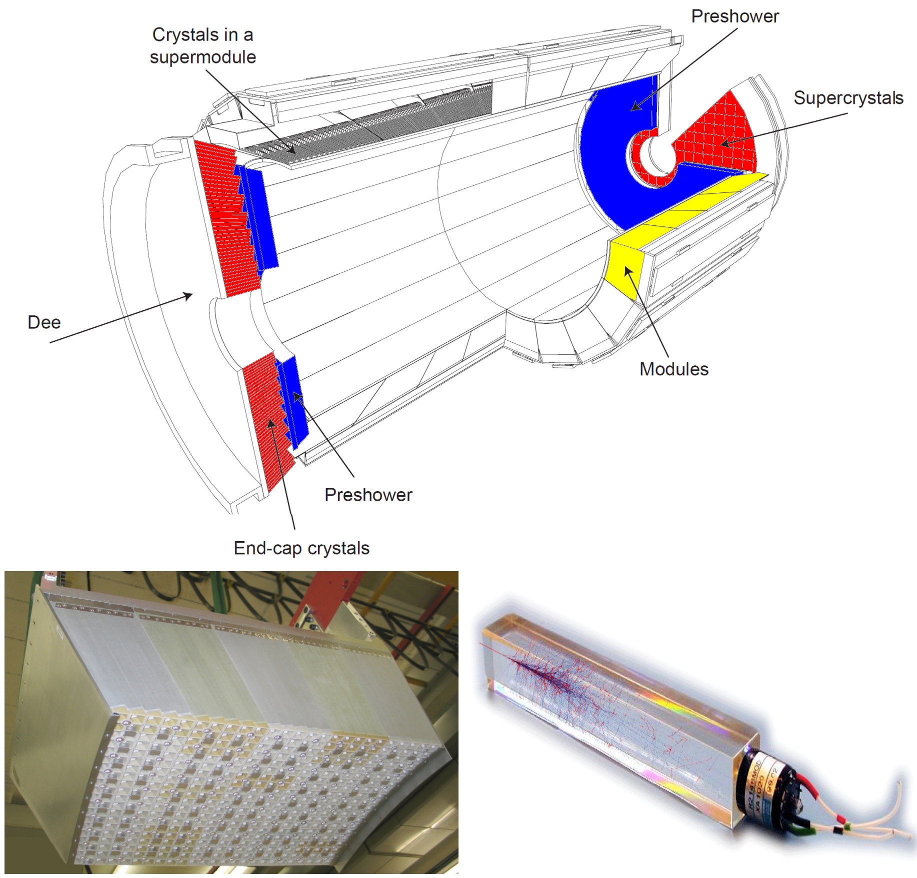

Figure 2.15: Top: CMS ECAL schematic view. Bottom: Module equipped with the crystals (left); ECAL crystal(right) with an artistic representation of an electromagnetic shower [74].

photodetectors given that crystals themselves have a low light yield (30γ/MeV). A schematic view of the ECAL is shown in Figure 2.15.

Energy is measured when electrons and photons are absorbed by the crystals which generates an electromagnetic *shower*, as seen in bottom right picture of the Figure 2.15; the shower is seen as a *cluster* of energy which depending on the amount of energy deposited can involve several crystals. The ECAL barrel (EB) covers the region $|\eta| < 1.479$, using crystals of depth of 23 cm and 2.2×2.2 cm^2 transverse section; the ECAL endcap (EE) covers the region $1.479 < |\eta| < 3.0$ using crystals of depth 22 cm and transverse section of 2.86×2.86 cm^2; the photodetectors used are vacuum phototriodes (VPTs). Each EE is divided in two structures called *Dees*.

The preshower detector (ES) is installed in front of the EE and covers the region $1.653 < |\eta| < 2.6$. The ES provides a precise measurement of the position of electromagnetic showers, which allows to distinguish electrons and photon signals from π^0 decay signals. The ES is composed of a layer of lead radiators followed by a layer of silicon strip sensors. The lead radiators initiate electromagnetic showers when reached by photons and electrons, then, the strip sensors measure the deposited energy and the transverse shower profiles. The full ES thickness is 20 cm.

2.3.5 Hadronic calorimeter

Hadrons are not absorbed by the ECAL[2] but by the hadron calorimeter (HCAL), which is made of a combination of alternating brass absorber layers and silicon photomultiplier(SiPM) layers; therefore, particles passing through the scintillator material produce showers, as in the ECAL, as a result of the inelastic scattering of the hadrons with the detector material. Since the particles are not absorbed in the scintillator, their energy is sampled; therefore the total energy is not measured but estimated from the energy clusters, which reduces the resolution of the detector. Brass was chosen as the absorber material due to its short interaction length ($\lambda_I = 16.42$cm) and its non-magnetivity. Figure 2.16 shows a schematic view of the CMS HCAL.

The HCAL is divided into four sections; the Hadron Barrel (HB), the Hadron Outer (HO), the Hadron Endcap (HE) and the Hadron Forward (HF) sections. The HB covers the region $0 < |\eta| < 1.4$, while the HE covers the region $1.3 < |\eta| < 3.0$. The HF, made of quartz fiber scintillator and steel as absorption material, covers the forward region $3.0 < |\eta| < 5.2$. Both the HB and HF are located inside the solenoid. The HO is placed outside the magnet as an additional layer of scintillators with the

[2] Most hadrons are not absorbed, but few low-energy ones might be.

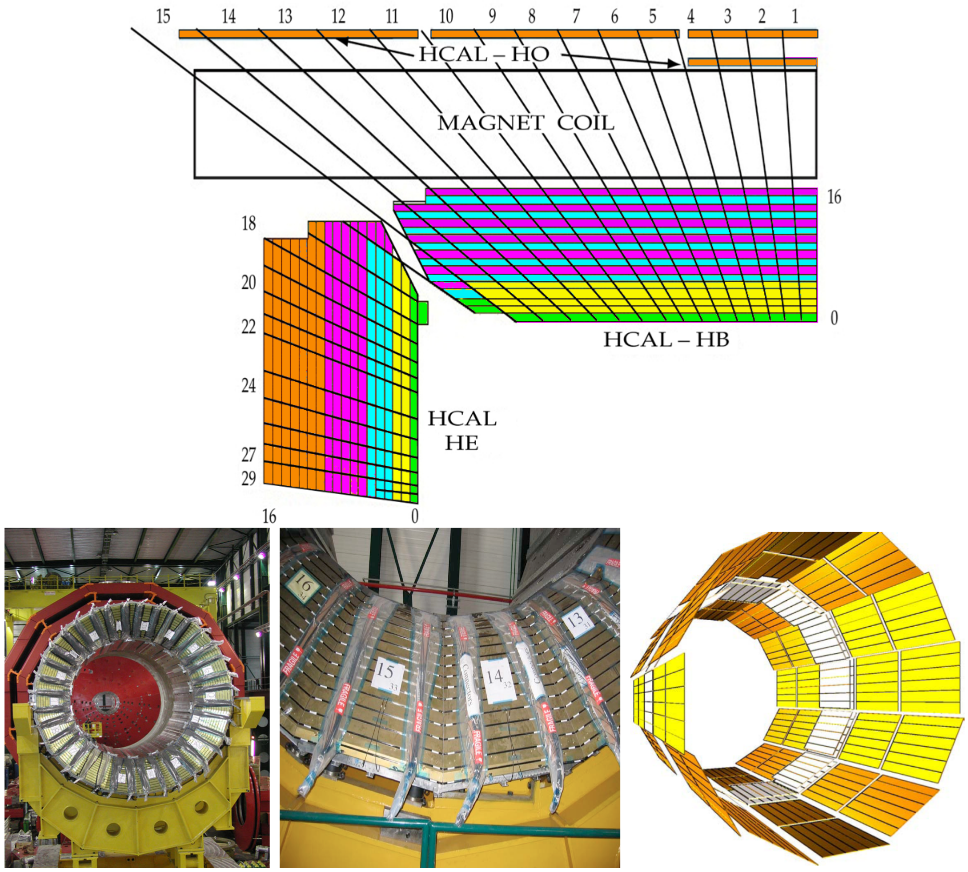

Figure 2.16: Top: CMS HCAL schematic view, the colors indicate the layers that are grouped into the same readout channels. Bottom: picture of a section of the HB; the absorber material is the golden region and scintillators are placed in between the absorber material (left and center). Schematic view of the HO (right). [83, 84]

purpose of measure the energy tails of particles passing through the HB and the magnet (see Figure 2.16 top and bottom right).

2.3.6 Superconducting solenoid magnet

The superconducting magnet installed in the CMS detector is designed to provide an intense and highly uniform magnetic field in the central part of the detector. In fact, the tracking system takes advantage of the bending power of the magnetic field to measure with precision the momentum of the particles that traverse it; the

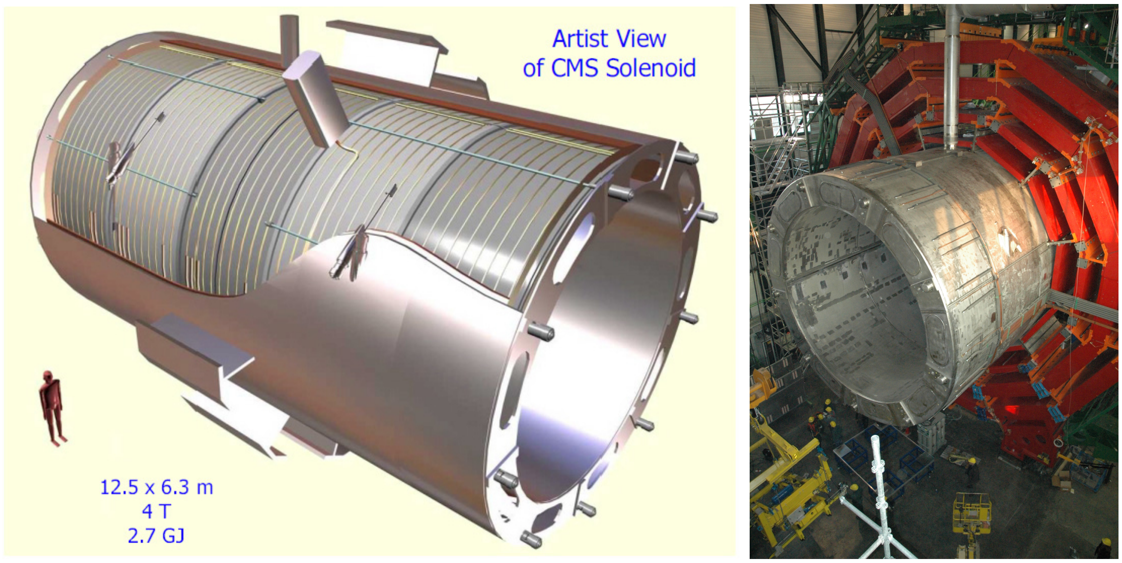

Figure 2.17: Artistic representation of the CMS solenoid magnet(left). The magnet is supported on an iron yoke (right) which also serves as the house of the muon detector and as mechanical support for the whole CMS detector [77].

unambiguous determination of the sign for high momentum muons was a driving principle during the design of the magnet. The magnet has a diameter of 6.3 m, a length of 12.5 m and a cold mass of 220 t; the generated magnetic field reaches a strength of 3.8T. Since it is made of Ni-Tb superconducting cable it has to operate at a temperature of 4.7 K by using a helium cryogenic system; the current circulating in the cables reaches 18800 A under normal running conditions. The left side of Figure 2.17 shows an artistic view of the CMS magnet, while the right side shows a transverse view of the cold mass where the winding structure is visible.

The yoke (see Figure 2.17), composed of 5 barrel wheels and 6 endcap disks made of iron, serves not only as the media for magnetic flux return but also provides housing for the muon detector system and structural stability to the full detector.

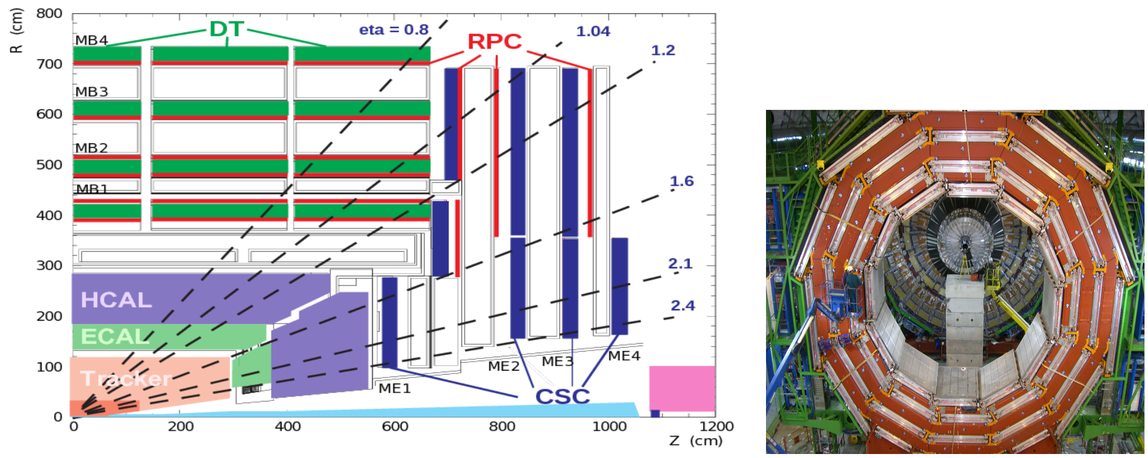

Figure 2.18: Left: CMS muon system schematic view; Right: one of the yoke rings with the muon DTs and RPCs installed; in the back it is possible to see the muon endcap [85].

2.3.7 Muon system

Muons are the only charged particles able to pass through all the CMS detector due to their low ionization energy loss; thus, muons can be separated easily from the high amount of particles produced in a *pp* collision. Also, muons are expected to be produced in the decay of several new particles; therefore, good detection of muons was one of the leading principles when designing the CMS detector.

The CMS muon detection system (muon spectrometer) is embedded in the return yoke as seen in Figure 2.18. It is composed of three different detector types, the drift tube chambers (DT), cathode strip chambers (CSC), and resistive plate chambers (RPC); DT are located in the central region $\eta < 1.2$ arranged in four layers of drift chambers filled with an Ar/CO_2 gas mixture.

The muon endcaps are made of CSCs covering the region $\eta < 2.4$ and filled with a mixture of $Ar/CO_2/CF_4$. The reason behind using a different detector type lies on the different conditions in the forward region like the higher muon rate and higher residual magnetic field compared to the central region.

The third type of detector used in the muon system is a set of four disks of RPCs working in avalanche mode. The RPCs provide good spatial and time resolutions. The track of high-p_T muon candidates is built combining information from the tracking system and the signal from up to six RPCs and four DT chambers.

The muon tracks are reconstructed from the hits in the several layers of the muon system.

2.3.8 CMS trigger system

CMS expects pp collisions every 25 ns, i.e., an interaction rate of 40 MHz for which it is not possible to store the recorded data in full. In order to handle this high event rate data, an online event selection, known as triggering, is performed; triggering reduces the event rate to 100 Hz for storage and further offline analysis.

The trigger system starts with a reduction of the event rate to 100 kHz in the so-called *the level 1 trigger* (L1). L1 is based on dedicated programmable hardware like Field Programmable Gate Arrays (FPGAs) and Application Specific Integrated Circuits (ASICs), partly located in the detector itself; another portion is located in the CMS underground cavern. Hit pattern information from the muon chambers and the energy deposits in the calorimeter are used to decide if an event is accepted or rejected, according to selection requirements previously defined, which reflect the interesting physics processes. Figure 2.19 shows the L1 trigger architecture.

The second stage in the trigger system is called *the high-level trigger* (HLT); events accepted by L1 are passed to HLT in order to make an initial reconstruction of them. HLT is software based and runs on a dedicated server farm, using selection algorithms and high-level object definitions; the event rate at HLT is reduced to 100 Hz. The first HLT stage takes information from the muon detectors and the calorimeters to make the initial object reconstruction; in the next HLT stage, information from the

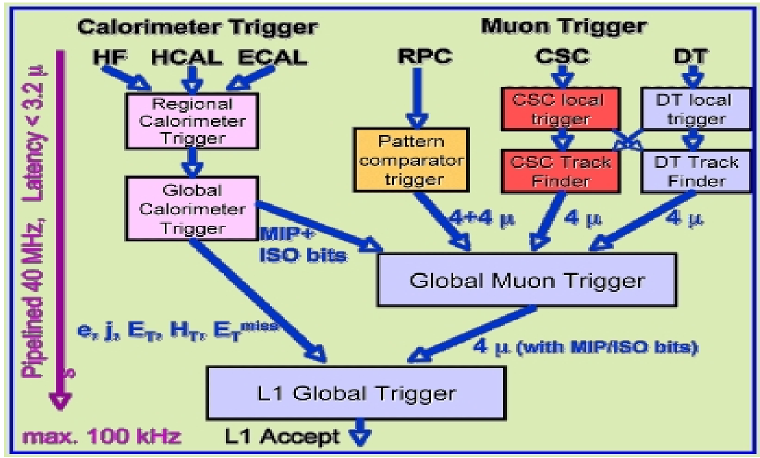

Figure 2.19: CMS Level-1 trigger architecture [86].

pixel and strip detectors is used to do first fast tracking and then full tracking online. This initial object reconstruction is used in further steps of the trigger system.

Events and preliminary reconstructed physics objects from HLT are sent to be fully reconstructed at the CERN computing center know also as Tier-0 facility. Again, the pixel detector information provides high-quality seeds for the track reconstruction algorithm offline, primary vertex reconstruction, electron and photon identification, muon reconstruction, τ identification, and b-tagging. After full reconstruction, data sets are made available for offline analyses.

2.3.9 CMS computing

Data coming from the experiment have to be stored and made available for further analysis; in order to cope with all the tasks implied in the offline data processing, like transfer, simulation, reconstruction and reprocessing, among others, a large computing power is required. The CMS computing system is based on the distributed architecture concept, where users of the system and physical computer centers are distributed worldwide and interconnected by high-speed networks.

The worldwide LHC computing grid (WLCG) is the mechanism used to provide

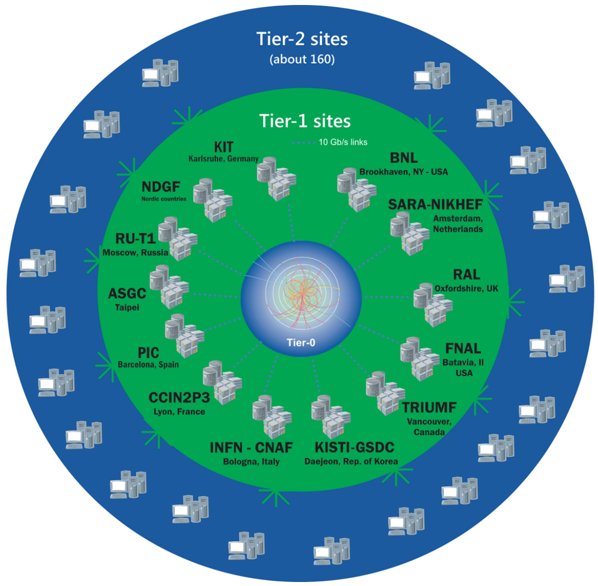

Figure 2.20: WLCG structure. The primary computer centers (Tier-0) are located at CERN (data center) and at the Wigner datacenter in Budapest. Tier-1 is composed of 13 centers and Tier-2 is composed of about 160 centers. [87].

that distributed environment. WLCG is a tiered structure connecting computing centers around the world, which provides the necessary storage and computing facilities. The primary computing centers of the WLCG are located at the CERN and the Wigner datacenter in Budapest and are known as Tier-0 as shown in Figure 2.20. The main responsibilities for each tier level are [87]

- **Tier-0**: initial reconstruction of recorded events and storage of the resulting datasets, the distribution of raw data to the Tier-1 centers.

- **Tier-1**: provide storage capacity, support for the Grid, safe-keeping of a proportional share of raw and reconstructed data, large-scale reprocessing and safe-keeping of corresponding output, generation of simulated events, distribution of data to Tier 2s, safe-keeping of a share of simulated data produced at these Tier 2s.

- **Tier-2**: store sufficient data and provide adequate computing power for specific analysis tasks and proportional share of simulated event production and reconstruction.

Aside from the general computing strategy to manage the huge amount of data produced by experiments, CMS uses a software framework to perform a variety of processing, selection and analysis tasks. The central concept of the CMS data model referred to as *event data model* (EDM) is the *Event*; therefore, an event is the unit that contains the information from a single bunch crossing, any data derived from that information like the reconstructed objects, and the details of the derivaction.

Events are passed as the input to the *physics modules* that obtain information from them and create new information; for instance, *event data producers* add new data into the events, *analyzers* produce an information summary from an event set, *filters* perform selection and triggering.

CMS uses several event formats with different levels of detail and precision

- **Raw format**: events in this format contain the full recorded information from the detector as well as trigger decision and other metadata. An extended version of raw data is used to store information from the CMS Monte Carlo simulation tools (see Chapter 3). Raw data are stored permanently, occupying about 2MB/event

- **RECO format**: events in this format correspond to raw data that have been submitted to reconstruction algorithms like primary and secondary vertex reconstruction, particle ID, and track finding. RECO events contain physics objects and all the information used to reconstruct them; average size is about 0.5 MB/event.

- **AOD format**: Analysis Object Data (AOD) is the data format used in the physics analyses given that it contains the parameters describing the high-level physics objects in addition to enough information to allow a kinematic refitting if needed. AOD events are filtered versions of the RECO events to which skimming or other filtering have been applied, hence AOD events are subsets of RECO events. Requires about 100 kB/event.

- **Non-event data** are data needed to interpret and reconstruct events. Some of the non-event data used by CMS contains information about the detector contraction and condition data like calibrations, alignment, and detector status.

Figure 2.21 shows the data flow scheme between CMS detector and tiers.

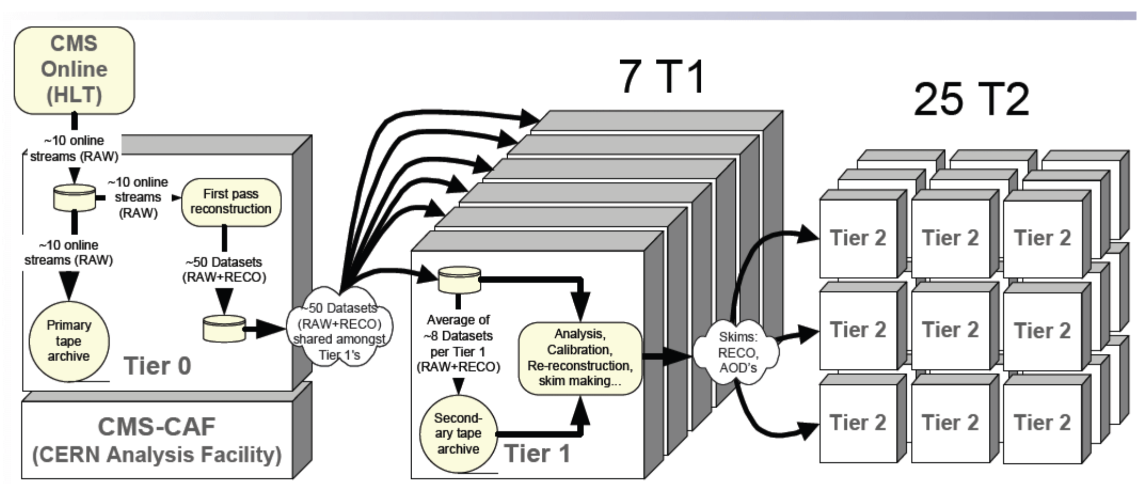

Figure 2.21: Data flow from CMS detector through tiers.

The whole collection of software built as a framework is referred to as *CMSSW*. This framework provides the services needed by the simulation, calibration and alignment, and reconstruction modules that process event data, so that physicists can perform analysis. The CMSSW event processing model is composed of one executable, called cmsRun, and several plug-in modules which contains all the tools (calibration, reconstruction algorithms) needed to process an event. The same executable is used for both detector data and Monte Carlo simulations [88].

Chapter 3

Event generation, simulation and reconstruction

The process of analyzing data recorded by the CMS experiment involves several stages where the data are processed in order to interpret the information provided by all the detection systems; in those stages, the particles produced after the *pp* collision are identified by reconstructing their trajectories and measuring their features. In addition, the SM provides a set of predictions that have to be compared with the experimental results; however, in most of the cases, theoretical predictions are not directly comparable to experimental results due to the diverse source of uncertainties introduced by the experimental setup and theoretical approximations, among others.

The strategy to face these conditions consists in using statistical methods implemented in computational algorithms to produce numerical results that can be contrasted with the experimental results. These computational algorithms are commonly known as Monte Carlo (MC) methods and, in the case of particle physics, they are designed to apply the SM rules and produce predictions about the physical observables measured in the experiments. Since particle physics is governed by quantum mechanics principles, predictions are not allowed from single events; therefore, a high number of events are *generated* and predictions are produced in the form of statistical distributions for the observables. Effects of the detector presence are included in the predictions by introducing simulations of the detector itself.

This chapter presents a description of the event generation strategy and the tools used to perform the detector simulation and physics objects reconstruction. A comprehensive review of event generators for LHC physics can be found in Reference [89] on which this chapter is based.

3.1 Event generation

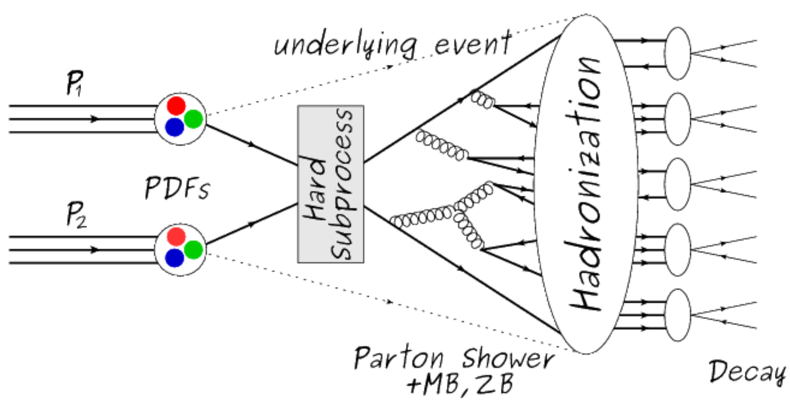

Figure 3.1: Event generation process. The actual interaction is generated in the hard subprocess. The parton shower describes the evolution of the partons from the hard subprocess.Modified from Reference [90].

The event generation is intended to create events that mimic the behavior of actual events produced in collisions; they obey a sequence of steps from the particles collision hard process to the decay process into the final state. Figure 3.1 shows a schematic view of the event generation process; the fact that the full process can be treated as several independent steps is motivated by the QCD factorization theorem.

Generation starts by taking into account the PDFs of the incoming particles. Event generators offer the option to chose from several PDF sets depending on the particular process under simulation[1]; in the following, *pp* collisions will be considered. The *hard subprocess* describes the actual interaction between partons from the incoming protons; it is represented by the matrix element connecting the initial and final states of the interaction. Normally, the matrix element can be written as a sum over Feynman diagrams and consider interferences between terms in the summation. During the generation of the hard subprocess, the production cross section is calculated.

The order to which the cross section is calculated depends on the order of the Feynman diagrams involved in the calculation; therefore, radiative corrections are included by considering a higher order Feynman diagrams where QCD radiation dominates. Currently, cross sections calculated to LO do not offer a satisfactory description of the processes, i.e., the results are only reliable for the shape of distributions; therefore, NLO calculations have to be performed with the implication that the computing time needed is highly increased.

The final parton content of the hard subprocess is subjected to the *parton shower* which generates the gluon radiation. Parton shower evolves the partons, i.e., glouns split into quark-antiquark pairs and quarks with enough energy radiate gluons giving rise to further parton multiplication, following the DGLAP (Dokshitzer-Gribov-Lipatov-Altarelli-Parisi) equations. Showering continues until the energy scale is low enough to reach the non-perturbative limit.

In the simulation of LHC processes that involve b quarks, like the single top quark or Higgs associated production, it is needed to consider that the b quark is heavier than the proton; hence, the QCD interaction description is made in two different

[1] Tool in Reference [91] allows to plot different PDF sets under customizable conditions.

schemes [95]

- four-flavor (4F) scheme. b quarks appear only in the final state because they are heavier than the proton and therefore they can be produced only from the splitting of a gluon into pairs or singly in association with a t quark in high energy-scale interactions; furthermore, during the simulation, the b-PDFs are set to zero. Calculations in this scheme are more complicated due to the presence of the second b quark but the full kinematics is considered already at LO and therefore the accuracy of the description is better.

- five-flavor (5F) scheme. b quarks are considered massless, therefore they can appear in both initial and final states since they can now be part of the proton; thus, during the simulation b-PDFs are not set to zero. In this scheme, calculations are simpler than in the 4F scheme and possible logarithmic divergences are absorbed by the PDFs through the DGLAP evolution.

In this thesis, the tHq events are generated using the 4F scheme in order to reduce uncertainties, while the tHW events are generated using the 5F scheme to eliminate LO interference with $t\bar{t}H$ process [48].

Partons involved in the pp collision are the focus of the simulation, however, the rest of the partons inside the incoming protons are also affected because the remnants are colored objects; also, multiple parton interactions can occur. The hadronization of the remnants and multiple parton interactions are known as *underlying event* and it has to be included in the simulation. In addition, multiple pp collisions in the same bunch crossing (pile-up mentioned in 2.2) occurs, actually in two forms

- *in-time PU* which refers to multiple pp collision in the bunch crossing but that are not considered as primary vertices.

- *Out-of-time PU* which refers to overlapping pp collisions from consecutive bunch crossings; this can occur due to the time-delays in the detection systems where information from one bunch crossing is assigned to the next or previous one.

While the underlying event effects are included in generation using generator-specific tools, PU effects are added to the generation by overlying Minimum-bias (MB) and Zero-bias (ZB) events to the generated events. MB events are inelastic events selected by using a loose trigger with as little bias as possible, therefore accepting a large fraction of the overall inelastic event; ZB events correspond to random events recorded by the detector when collisions are likely. MB models in-time PU and ZB models out-of-time PU.

The next step in the generation process is called *hadronization*. Since particles with a net color charge are not allowed to exits isolated, they have to recombine to form bound states. This is precisely the process by which the partons resulting from the parton shower arrange themselves as color singlets to form hadrons. At this step, the energy-scale is low and the strong coupling constant is large, therefore hadronization process is non-perturbative and the evolution of the partons is described using phenomenological models. Most of the baryons and mesons produced in the hadronization are unstable and hence they will decay in the detector.

The last step in the generation process corresponds to the decay of the unstable particles generated during hadronization; it is also simulated in the hadronization step, based on the known branching ratios.

3.2 Monte Carlo Event Generators.

The event generation described in the previous section has been implemented in several software packages for which a brief description is given.

- **PYTHIA 8**. It is a program designed to perform the generation of high energy physics events which describes the collisions between particles such as electrons and protons. Several theories and models are implemented in it, in order to describe physical aspects like hard and soft interaction, parton distributions, initial and final-state parton showers, multiple parton interactions, beam remnants, hadronization[2] and particle decay. Thanks to extensive testing, several optimized parametrizations, known as *tunings*, have been defined in order to improve the description of actual collisions to a high degree of precision; for analysis at $\sqrt{s} = 13$ TeV, the underline event CUETP8M1 tune is employed [97]. The calculation of the matrix element is performed at LO which is not enough for the current required level of precision; therefore, pythia is often used for parton shower, hadronization and decays, while other event generators are used to generate the matrix element at NLO.

- **MadGraph5_aMC@NLO**. MadGraph is a matrix element generator which calculates the amplitudes for all contributing Feynman diagrams of a given process but does not provide a parton shower while MC@NLO incorporates NLO QCD matrix elements consistently into a parton shower framework; thus, MadGraph5_aMC@NLO, as a merger of the two event generators MadGraph5 and aMC@NLO, is an event generator capable to calculate tree-level and NLO cross sections and perform the matching of those with the parton shower. It is one of the most frequently used matrix element generators; however, it has the particular feature of the presence of negative event weights which reduce the number of events used to reproduce the properties of the objects generated [98].

- **POWHEG**. It is an NLO matrix element generator where the hardest emis-

[2] based in the Lund string model [96]

sion of color charged particles is generated in such a way that the negative event weights issue of MadGraph5_aMC@NLO is overcome; however, the method requires an interface with p_T-ordered parton shower or a parton shower generator where this highest emission can be vetoed in order to avoid double counting of this highest-energetic emission. PYTHIA is a commonly matched to POWHEG event generator [100].

Events resulting from the whole generation process are known as MC events.

3.3 CMS detector simulation.

After generation, MC events contain the physics of the collisions but they are not ready to be compared to the events recorded by the experiment since these recorded events correspond to the response of the detection systems to the interaction with the particles traversing them. The simulation of the CMS detector has to be applied on top of the event generation; it is simulated with a MC toolkit for the simulation of particles passing through matter called Geant4 which is also able to simulate the electronic signals that would be measured by all detectors inside CMS.

The simulation takes the generated particles contained in the MC events as input, makes them pass through the simulated geometry, and models physics processes that particles experience during their passage through matter. The full set of results from particle-matter interactions corresponds to the simulated hit which contains information about the energy loss, momentum and position. Particles of the input event are called *primary*, while the particles originating from GEANT4-modeled interactions of a primary particle with matter are called a *secondary*. Simulated hits are the input of subsequent modules that emulate the response of the detector readout system and triggers. The output from the emulated detection systems and triggers is known as digitization [101, 102].

The modeling of the CMS detector corresponds to the accurate modeling of the interaction among particles, the detector material, and the magnetic field. This simulation procedure includes the following standard steps

- Modeling of the Interaction Region.
- Modeling of the particle passage through the hierarchy of volumes that compose CMS detector and of the accompanying physics processes.
- Modeling of the effect of multiple interactions per beam crossing and/or the effect of events overlay (Pile-Up simulation).
- Modeling of the detector's electronics response, signal shape, noise, calibration constants (digitization).

In addition to the full simulation, i.e., a detailed detector simulation, a faster simulation (FastSim) have been developed, that may be used where much larger statistics are required. In FastSim, detector material effects are parametrized and included in the hits; those hits are used as input of the same higher-level algorithms[3] used to analyze the recorded events. In this way, comparisons between fast and full simulations can be performed [104].

After the full detector simulation, the output events can be directly compared to events actually recorded in the CMS detector. The collection of MC events that reproduces the expected physics for a given process isknown as MC sample.

[3] track fitting, calorimeter clustering, b tagging, electron identification, jet reconstruction and calibration, trigger algorithms which will be considered in the next sections

3.4 Event reconstruction.

The CMS experiment use the *particle-flow event reconstruction algorithm (PF)* to do the reconstruction of particles produced in *pp* collisions. Next sections will present a basic description of the *Elements* used by PF (tracker tracks, energy clusters, and muon tracks), based in the References [105,106] where more detailed descriptions can be found.

3.4.1 Particle-Flow Algorithm.

Each of the several sub detection systems of the CMS detector is dedicated to identify an specific type of particles, i.e., photons and electrons are absorbed by the ECAL and their reconstruction is based on ECAL information; hadrons are reconstructed from clusters in the HCAL while muons are reconstructed from hits in the muon chambers. PF is designed to correlate signals from all the detector layers (tracks and energy clusters) in order to reconstruct and identify each final state particle and its properties as sketched in Figure 3.2.

For instance, a charged hadron is identified by a geometrical connection, known as *link*, between one or more calorimeter clusters and a track in the tracker, provided there are no hits in the muon system; combining several measurements allows a better determination of the energy and charge sign of the charged hadron.

Charged-particle track reconstruction.

The strategy used by PF in order to reconstruct tracks is called *Iterative Tracking* which occurs in four steps

- Seed generation where initial track candidates are found by looking for a combination of hits in the pixel detector, strip tracker, and muon chambers. In total

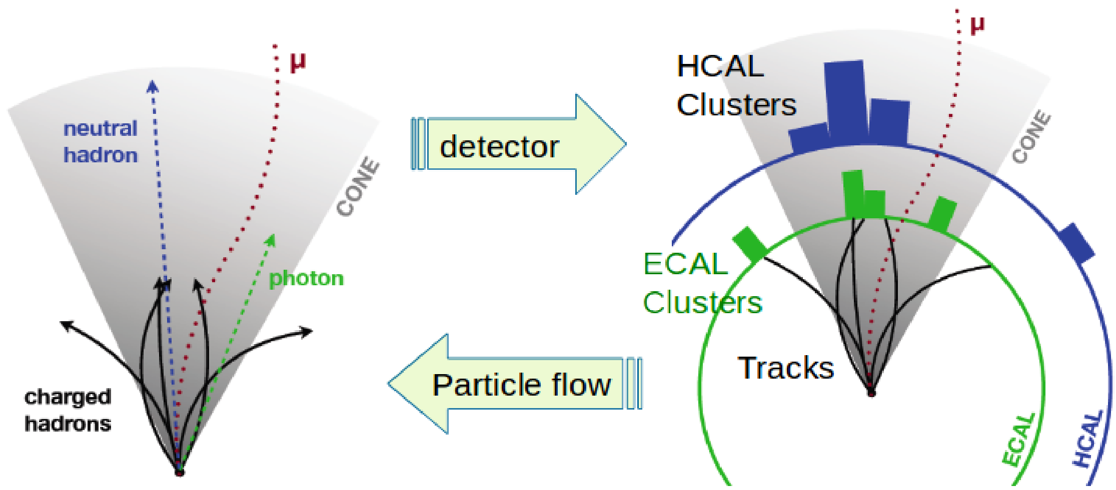

Figure 3.2: Particle flow algorithm. Information from the several CMS detection systems if provided as input to the algorithm which then combine it to identify and reconstruct all the particles in the final state and their properties. Reconstruction of simulated events is also performed by providing information from MC samples, detector and trigger simulation [107].

ten iterations are performed, each one with a different seeding requirement. Seeds are used to estimate the trajectory parameters and uncertainties at the time of the full track reconstruction. Seeds are also considered track candidates.

- Track finding using a tracking software known as Combinatorial Track Finder (CTF) [108]. The seed trajectories are extrapolated along the expected flight path of a charged particle, in agreement to the trajectory parameters obtained in the first step, in an attempt to find additional hits that can be assigned to the track candidates.

- Track-fitting where the found tracks are passed as input to a module which provides the best estimate of the parameters of each trajectory.

- Track selection where track candidates are submitted to a selection which discards those that fail a set of defined quality criteria.

Iterations differ in the seeding configuration and the final track selection as elaborated in References [105, 106]. In the first iteration, high p_T tracks and tracks produced near to the interaction region are identified and those hits are masked thereby reducing the combinatorial complexity. Next, iterations search for more complicated tracks, like low p_T tracks and tracks from b hadron decays, which tend to be displaced from the interaction region.

Vertex reconstruction.

During the track reconstruction, an extrapolation toward to the calorimeters is performed in order to match energy deposits; that extrapolation is performed also toward the beamline in order to find the origin of the track known as *vertex*. The vertex reconstruction is performed by selecting from the available reconstructed tracks, those that are consistent with being originated in the interaction region where *pp* collisions are produced. The selection involves a requirement on the number of tracker (pixel and strip) hits and the goodness of the track fit.

Selected tracks are clustered using a *deterministic annealing algorithm (DA)*[4]. A set of candidate vertices and their associated tracks, resulting from the DA, are then fitted with an *adaptive vertex fitter (AVF)* to produce the best estimate of the vertices locations.

The p_T of the tracks associated to a reconstructed vertex is added, squared and used to organize the vertices; the vertex with the highest squared sum is designated as the *primary vertex (PV)* while the rest are designated as PU vertices.

Calorimeter clustering.

After traversing the CMS tracker system, electrons, photons and hadrons deposit their energy in the ECAL and HCAL cells. The PF clustering algorithm aims to provide

[4] DA algorithm and AVF are described in detail in References [110, 111]

a high detection efficiency even for low-energy particles and an efficient distinction between close energy deposits. The clustering runs independently in the ECAL barrel and endcaps, HCAL barrel and endcaps, and the two preshower layers, following two steps

- cells with an energy larger than a given seed threshold and larger than the energy of the neighboring cells are identified as cluster seeds. The neighbor cells are those that either share a side with the cluster seed candidate, or the eight closest cells including cells that only share a corner with the seed candidate.

- cells with at least a corner in common with a cell already in the cluster seed and with an energy above a cell threshold are grouped into topological clusters.

Clusters formed in this way are known as *particle-flow clusters*. With this clustering strategy, it is possible to detect and measure the energy and direction of photons and neutral hadrons as well as differentiate these neutral particles from the charged hadron energy deposits. In cases involving charged hadrons for which the track parameters are not determined accurately, for instance, low-quality and high-p_T tracks, clustering helps in the energy measurements.

Electron track reconstruction.

Although the charged-particle track reconstruction described above works for electrons, they lose a significant fraction of their energy via bremsstrahlung photon radiation before reaching the ECAL; thus, the reconstruction performance depends on the ability to measure also the radiated energy. The reconstruction strategy, in this case, requires information from the tracking system and from the ECAL. Bremsstrahlung photons are emitted at similar η values to that of the electron but at different values of ϕ; therefore, the radiated energy can be recovered by grouping ECAL clusters in a

η window over a range of ϕ around the electron direction. The group is called ECAL supercluster.

Electron candidates from the track-seeding and ECAL super clustering are merged into a single collection which is submitted to a full electron tracking fit with a Gaussian-sum filter (GSF) [109]. The electron track and its associated ECAL supercluster form a *particle-flow electron*.

Muon track reconstruction.

Given that the CMS detector is equipped with a muon spectrometer capable to identify and measure the momentum of the muons traversing it, the muon reconstruction is not specific to PF; therefore, three different muon types are defined

- *Standalone muon.* A clustering on the DTs or CSCs hits is performed to form track segments; those segments are used as seeds for the reconstruction in the muon spectrometer. All DTs, CSCs, and RPCs hits along the muon trajectory are combined and fitted to form the full track. The fitting output is called a *standalone-muon track*.

- *Tracker muon.* Each track in the inner tracker with p_T larger than 0.5 GeV and a total momentum p larger than 2.5 GeV is extrapolated to the muon system. A *tracker muon track* corresponds to a extrapolated track that matches at least one muon segment.

- *Global muon.* When tracks in the inner tracker (inner tracks) and standalone-muon tracks are matched and turn out being compatibles, their hits are combined and fitted to form a *global-muon track*.

Global muons sharing the same inner track with tracker muons are merged into a single candidate. PF muon identification uses the muon energy deposits in ECAL, HCAL, and HO associated with the muon track to improve the muon identification.

Particle identification and reconstruction.

PF elements are connected by a linker algorithm that tests the connection between any pair of elements; if they are found to be linked, a geometrical distance that quantifies the quality of the link is assigned. Two elements may be linked indirectly through common elements. Linked elements form *PF blocks* and each PF block may contain elements originating in one or more particles. Links can be established between tracks, between calorimeter clusters, and between tracks and calorimeter clusters. The identification and reconstruction start with a PF block and proceed as follows

- Muons. An *isolated global muon* is identified by evaluating the presence of inner track and energy deposits close to the global muon track in the (η,ϕ) plane, i.e., in a particular point of the global muon track, inner tracks and energy deposits are sought within a radius of $\Delta R = 0.3$ (see eqn. 2.7) from the muon track; if they exit and the p_T of the found track added to the E_T of the found energy deposit does not exceed 10% of the muon p_T then the global muon is an isolated global muon. This isolation condition is stringent enough to reject hadrons misidentified as muons.

  Non-isolated global muons are identified using additional selection requirements on the number of track segments in the muon system and energy deposits along the muon track. Muons inside jets are identified with more stringent criteria in isolation and momentum as described in Reference [112]. The PF elements associated with an identified muon are masked from the PF block.

- Electrons are identified and reconstructed as described above plus some additional requirements on fourteen variables like the amount of energy radiated, the distance between the extrapolated track position at the ECAL and the position of the associated ECAL supercluster, among others, which are combined in an specialized multivariate analysis strategy that improves the electron identification. Tracks and clusters used to identify and reconstruct electrons are masked in the PF block.

- Isolated photons are identified from ECAL superclusters with E_T larger than 10 GeV, for which the energy deposited at a distance of 0.15, from the supercluster position on the (η,ϕ) plane, does not exceed 10% of the supercluster energy; note that this is an isolation requirement. In addition, there must not be links to tracks. Clusters involved in the identification and reconstruction are masked in the PF block.

- Bremsstrahlung photons and prompt photons tend to convert to electron-positron pairs inside the tracker, therefore, a dedicated finder algorithm is used to link tracks that seem to originate from a photon conversion; in case those two tracks are compatible with the direction of a bremsstrahlung photon, they are also linked to the original electron track. Photon conversion tracks are also masked in the PF block.

- The remaining elements in the PF block are used to identify hadrons. In the region $|\eta| \leq 2.5$, neutral hadrons are identified with HCAL clusters not linked to any track while photons from neutral pion decays are identified with ECAL clusters without links to tracks. In the region $|\eta| > 2.5$ ECAL clusters linked to HCAL clusters are identified with a charged or neutral hadron shower; ECAL clusters with no links are identified with photons. HCAL clusters not used yet, are linked to one or more unlinked tracks and to an unlinked ECAL in order to reconstruct charged-hadrons or a combination of photons and neutral hadrons according to certain conditions on the calibrated calorimetric energy.

- Charged-particle tracks may be liked together when they converge to a *secondary vertex (SV)* displaced from the IP where the PV and PU vertices are reconstructed; at least three tracks are needed in that case, of which at most one has to be an incoming track with hits in tracker region between a PV and the SV.

The linker algorithm, as well as the whole PF algorithm, has been validated and commissioned; results from that validation are presented in the Reference [105].

Jet reconstruction.

Quarks and gluons may be produced in the *pp* collisions, therefore, their hadronization will be seen in the detector as a shower of hadrons and their decay products in the form of a *jet*. Two classes of clustering algorithms have been developped based in their jet definition [113]:

- Iterative cone algorithms (IC). Jets are defined in terms of circles of fixed radius R in the η-ϕ plane, known as *stable cones*, for which the sum of the momenta of all the particles within the cone points in the same direction as the center of the circle. The seed of the iteration is the hardest non-isolated particle in the event, then, the resulting momentum direction is assigned as the new cone direction and a new iteration starts; iteration process stops when the cone if found to be stable.

- Sequential recombination algorithms. The distance between non-isolated particles is calculated; if that distance is below a threshold, these particles are recombined into a new object. The sequence is repeated until the separation between the recombined object and any other particle is above certain threshold; the recombined object is called a jet and the algorithm starts again with the remaining particles.

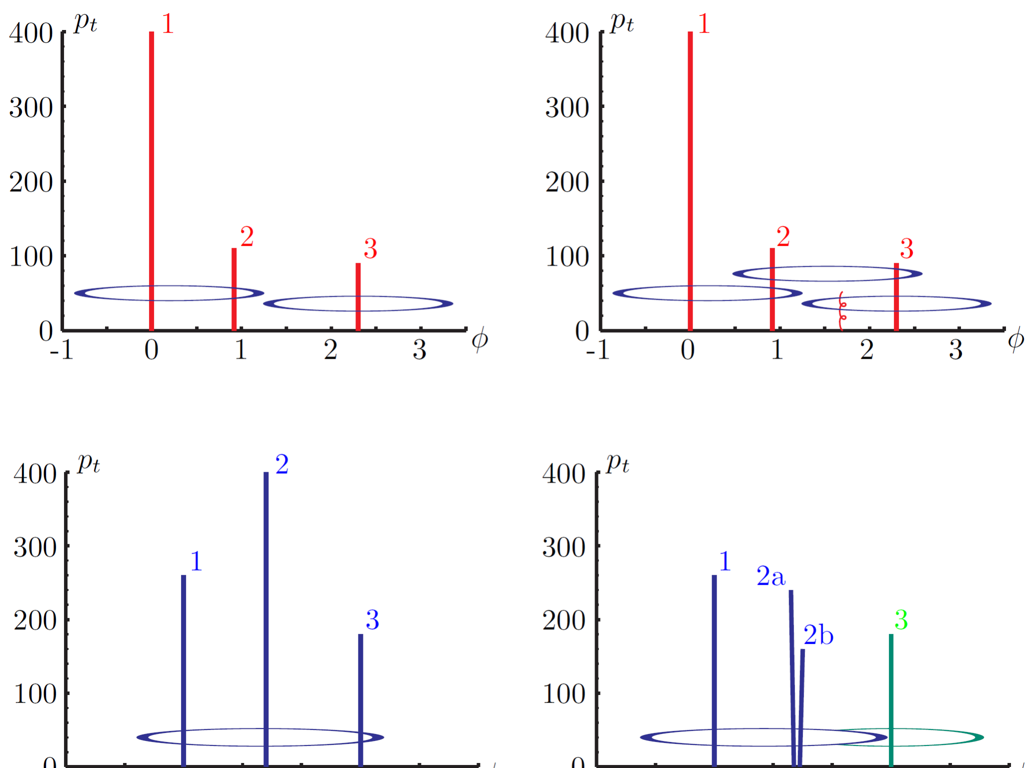

Figure 3.3: Stable cones identification using IC algorithms [113].

Two conditions are of particular importance for the clustering algorithms, *infrared and collinear (IRC) safety*. In order to explain the concept of infrared (IR) safety, consider an event with three hard particles as shown in the top left side of Figure 3.3, two stable cones are found and then two jets are identified; if a soft gluon is added, as shown in the top right side of Figure 3.3, three stable cones are found and the three hard particles are now clustered into a single jet. If the addition of soft particles change the outcome of the clustering, then it is said that the algorithm is IR unsafe. Soft radiation is highly likely in perturbative QCD, which dominates the physics of the jets, and then IR unsafe effect leads to divergences [113].

The concept of collinear safety can also be explained considering a three hard particles event, as shown in the bottom left side of Figure 3.3, where one stable cone containing all three particles is found and one jet is identified; if the hardest particle is split into two collinear particles (2a and 2b) in the bottom right side of Figure 3.3, then the clustering results in a different jet identification and the algorithm is said to be collinear unsafe. The collinear unsafe effect leads to divergences in jet cross section calculations [114].

It has been determined that IC algorithms are IRC unsafe, and therefore, they have to be replaced by algorithms that not only provide the finite perturbative results from theoretical computations, but also that are not highly dependent on underlying event and pileup effects which leads to significant corrections [113].

The sequential recombination algorithms arise as the IRC safe alternative used by the CMS experiment; in particular the anti-k_t algorithm which is a generalization of the previously existing k_t [115] and Cambridge/Aachen [116] jet clustering algorithms.

The anti-k_t algorithm is used to perform the jet reconstruction by clustering those PF particles within a cone (see Figure 3.4); previously, isolated electrons, isolated muons, and charged particles associated with other interaction vertices are excluded from the clustering.

The anti-k_t algorithm proceeds in a sequential recombination of PF particles; the distance between particles i and j (d_{ij}) and the distance between particles and the beam are defined as

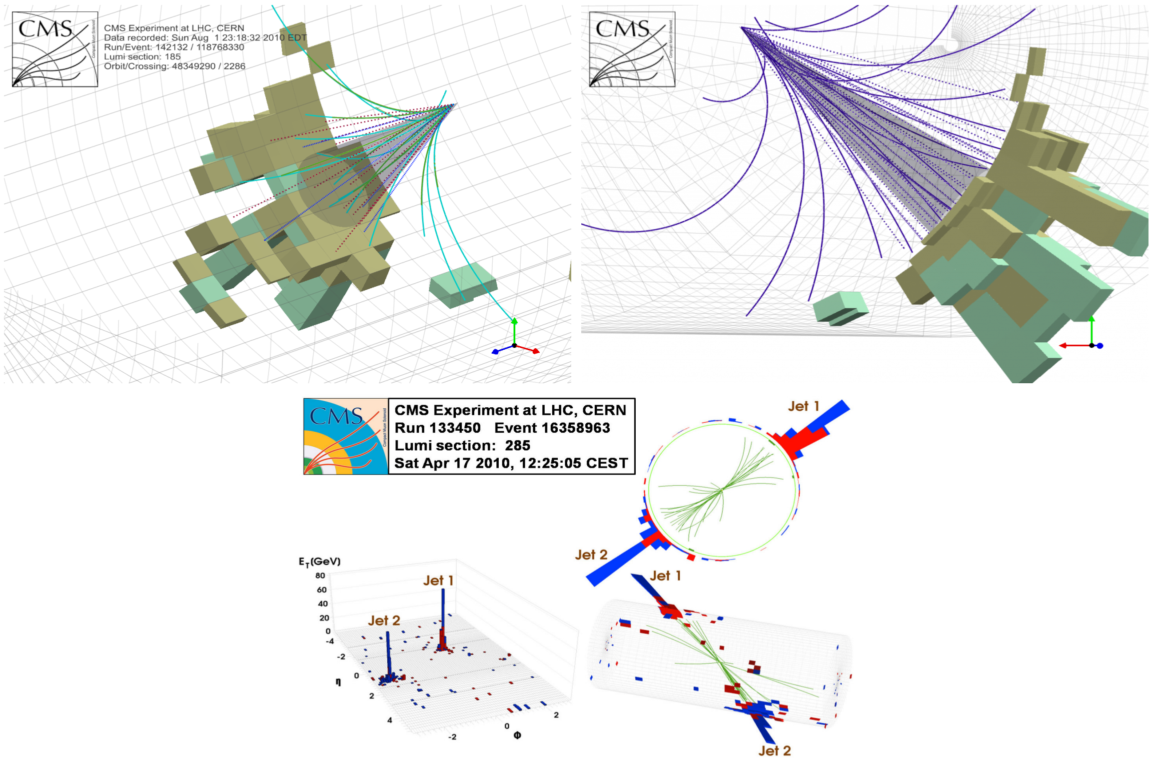

Figure 3.4: Jet reconstruction performed by the anti-k_t algorithm. Top: Two different views of a CMS recorded event are presented. Continuous lines correspond to tracks left by charged particles in the tracker while dotted lines are the imaginary paths followed by neutral particles. The green cubes represent the ECAL cells while the blue ones represent the HCAL cells; in both cases, the height of the cube represent the amount of energy deposited in the cells [117]. Bottom: Reconstruction of a recorded event with two jets [118].

$$
\begin{aligned}
d_{ij} =& \min\left(\frac{1}{k_{ti}^2}, \frac{1}{k_{tj}^2}\right)\frac{\Delta_{ij}^2}{R^2} \\
d_{iB} =& \frac{1}{k_{ti}^2}
\end{aligned}
\tag{3.1}
$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, k_{ti}, y_i and ϕ_i are the transverse momentum, rapidity and azimuth of particle i respectively and R is the called jet radius. For all the remaining PF particles, after removing the isolated ones, d_{ij} and d_{iB} are calculated[5] and the smallest is identified; if it is a d_{ij}, particles i and j are replaced with a new object whose momentum is the vectorial sum of the combined particles. If the smallest distance is a d_{iB} the clustering process ends, the object i (which at this stage should be a combination of several PF particles) is declared as a *Particle-flow-jet* (PF jet) and all the associated PF particles are removed from the detector. The clustering process is repeated until no PF particles remain. R is a free parameter that can be adjusted according to the specific analysis conditions; usually, two values are used, R=0.4 and R=0.5, giving the name to the so-called AK4-jet and AK5-jet respectively.

An advantage of the anti-k_t algorithm over other clustering algorithms is the regularity of the boundaries of the resulting jets. For all known IRC safe algorithms, soft radiation can introduce irregularities in the boundaries of the final jets; however, anti-k_t algorithm is soft-resilient, meaning that jets shape is not affected by soft radiation, which is a valuable property considering that knowing the typical shape of jets makes experimental calibration of jets more simple. In addition, that soft-resilience is expected to simplify certain theoretical calculations and reduce the momentum-resolution loss caused by underlying-event (UE) and pileup contamination [114].

The effect of the UE and pileup contamination over a jet identification, can be seen as if soft events are added to the jet; for instance, if a soft event representing UE or pileup is added to an event for which a set of jets J have been identified, and the clustering is rerun on that new extended event, the outcome will be different in two aspects: jets will contain some additional soft energy and the distribution of particles in jets may have change; that effect is called *back-reaction*. The back-reaction effect in the anti-k_t algorithm is suppressed not by the amount of momentum added to the jet but by the jet transverse momentum $p_{T,J}$, which means that this strong suppression leads to a smaller correction due to EU and pileup effect [114].

[5] Notice that this is a combinatorial calculation.

Jet energy Corrections

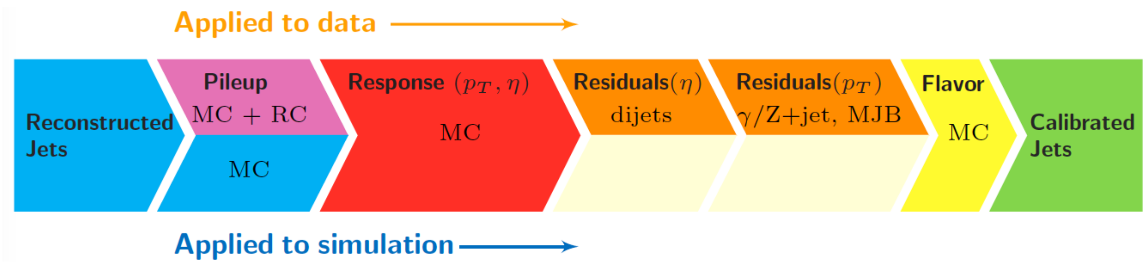

Figure 3.5: Jet energy correction diagram. Correction levels are applied sequentially in the indicated fixed order [120].

Even though jets can be reconstructed efficiently, there are some effects that are not included in the reconstruction and that lead to discrepancies between the reconstructed results and the predicted results; in order to overcome these discrepancies, a factorized model has been designed in the form of jet energy corrections (JEC) [119, 120] applied sequentially as shown in the diagram of Figure 3.5.

At each level, the jet four-momentum is multiplied by a scaling factor based on jet properties, i.e., η, flavor, etc.

- Level 1 correction removes the energy coming from pile-up. The scale factor is determined using a MC sample of QCD dijet (2 jets) events with and without pileup overlay; it is parametrized in terms of the offset energy density ρ, jet area A, jet η and jet p_T. Different corrections are applied to data and MC due to the detector simulation.

- MC-truth correction accounts for differences between the reconstructed jet energy and the MC particle-level energy. The correction is determined on a QCD dijet MC sample and is parametrized in terms of the jet p_T and η.

- Residuals correct remaining small differences within jet response in data and MC. The Residuals η-dependent correction compares jets of similar p_T in the barrel reference region. The Residuals p_T-dependent correct the jet absolute scale (JES vs p_T).

- Jet-flavor corrections are derived in the same way as MC-truth corrections but using QCD pure flavor samples.

b-tagging of jets.

A particular feature of the hadrons containing bottom quarks (b-hadrons) is that their lifetime is long enough to travel some distance before decaying, but it is not as long as those of light quark hadrons; therefore, when looking at the hadrons produced in *pp* collisions, b-hadrons decay typically inside the tracker rather than reaching the calorimeters as some light-hadrons do. As a result, a b-hadron decay gives rise to a displaced vertex (secondary vertex) with respect to the primary vertex as shown in Figure 3.6; the SV displacement is in the order of a few millimeters. A jet resulting from the decay of a b-hadron is called *b* jet; other jets are called light jets.

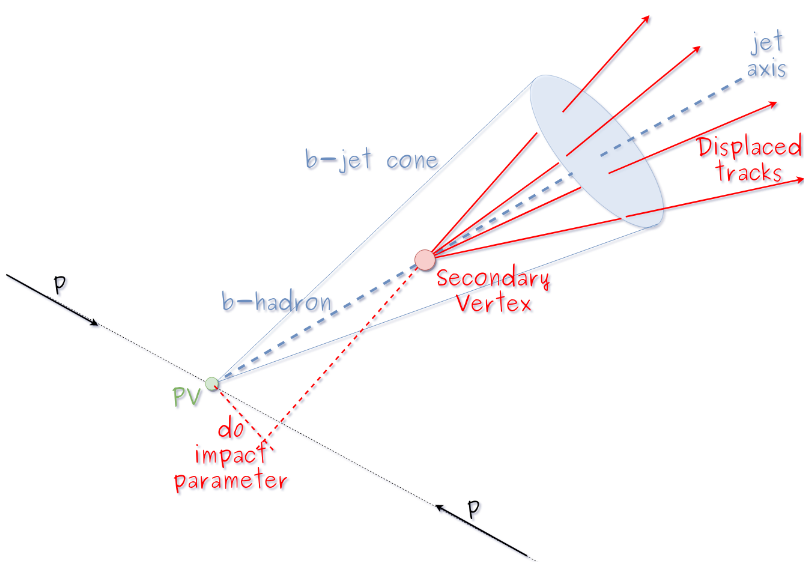

Figure 3.6: Secondary vertex in a b-hadron decay.

Several methods to identify *b*-jets (*b*-tagging) have been developed; the method used in this thesis is known as *Combined Secondary Vertex* algorithm in its second version (CSVv2) [121]. By using information of the impact parameter, the reconstructed secondary vertices, and the jet kinematics as input in a multivariate analysis that combines the discrimination power of each variable in one global discriminator variable, three working points (references): loose, medium and tight, are defined which quantify the probabilities of mistag jets from light quarks as jets from b quarks; 10, 1 and 0.1 % respectively. Although the mistagging probability decreases with the working point strength, the efficiency to correctly tag b-jets also decreases as 83, 69 and 49 % for the respective working point; therefore, a balance needs to be achieved according to the specific requirements of the analysis.

3.4.1.1 Missing transverse energy.

The fact that proton bunches carry momentum along the $z-$axis implies that for each event it is expected that the momentum in the transverse plane is balanced. Imbalances are quantified by the missing transverse energy (MET) and are attributed to several sources including particles escaping undetected through the beam pipe, neutrinos produced in weak interactions processes which do not interact with the detector and thus escaping without leaving a sign, or even undiscovered particles predicted by models beyond the SM.

The PF algorithm assigns the negative sum of the momenta of all reconstructed PF particles to the *particle-flow MET* according to

$$\vec{\not{E}}_T = -\sum_i \vec{p}_{T,i} \tag{3.2}$$

JEC are propagated to the calculation of the $\vec{\not{E}}_T$ as described in the Reference [122].

3.4.2 Event reconstruction examples

Figures 3.7-3.9 show the results of the reconstruction performed on 3 recorded events. Descriptions are taken directly from the source.

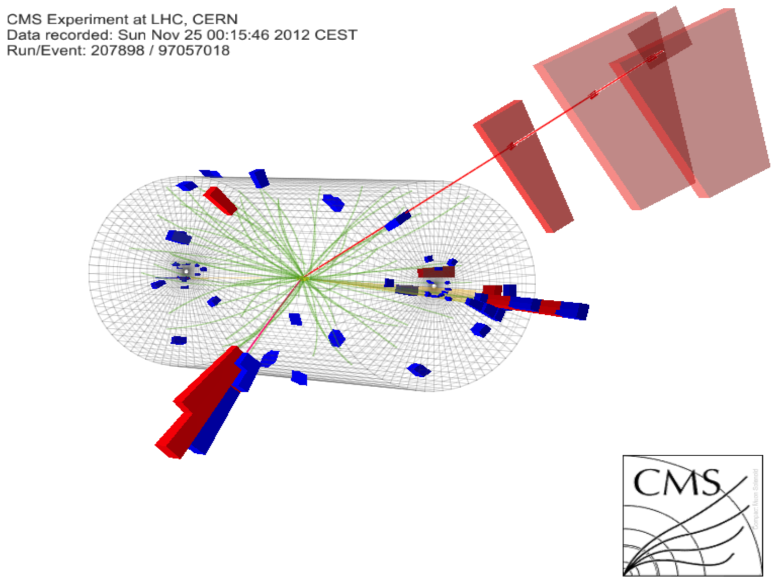

Figure 3.7: HIG-13-004 Event 1 reconstruction results; "HIG-13-004 Event 1: Event recorded with the CMS detector in 2012 at a proton-proton center-of-mass energy of 8 TeV. The event shows characteristics expected from the decay of the SM Higgs boson to a pair of τ leptons. Such an event is characterized by the production of two forward-going jets, seen here in opposite endcaps. One of the τ decays to a muon (red lines on the right) and neutrinos, while the other τ decays into a charged hadron and a neutrino." [123].

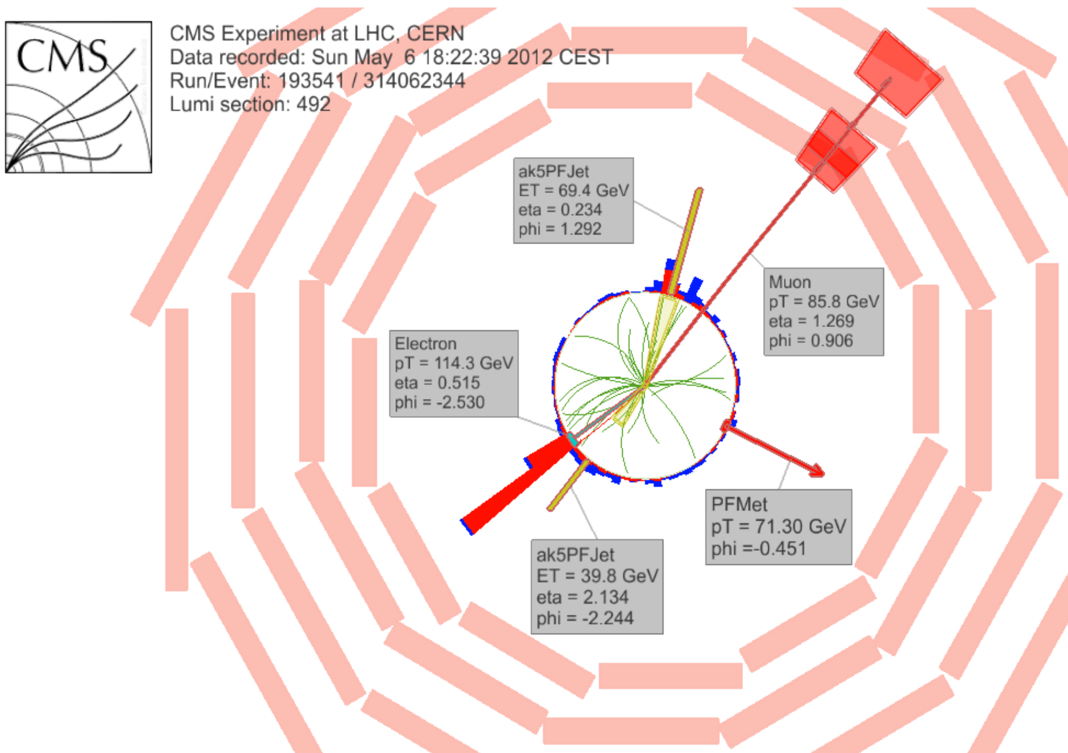

Figure 3.8: $e\mu$ event reconstruction results;"An $e\mu$ event candidate selected in 8 TeV data, as seen from the direction of the proton beams. The kinematics of the main objects used in the event selection are highlighted: two isolated leptons and two particle-flow jets. The reconstructed missing transverse energy is also displayed for reference" [124].

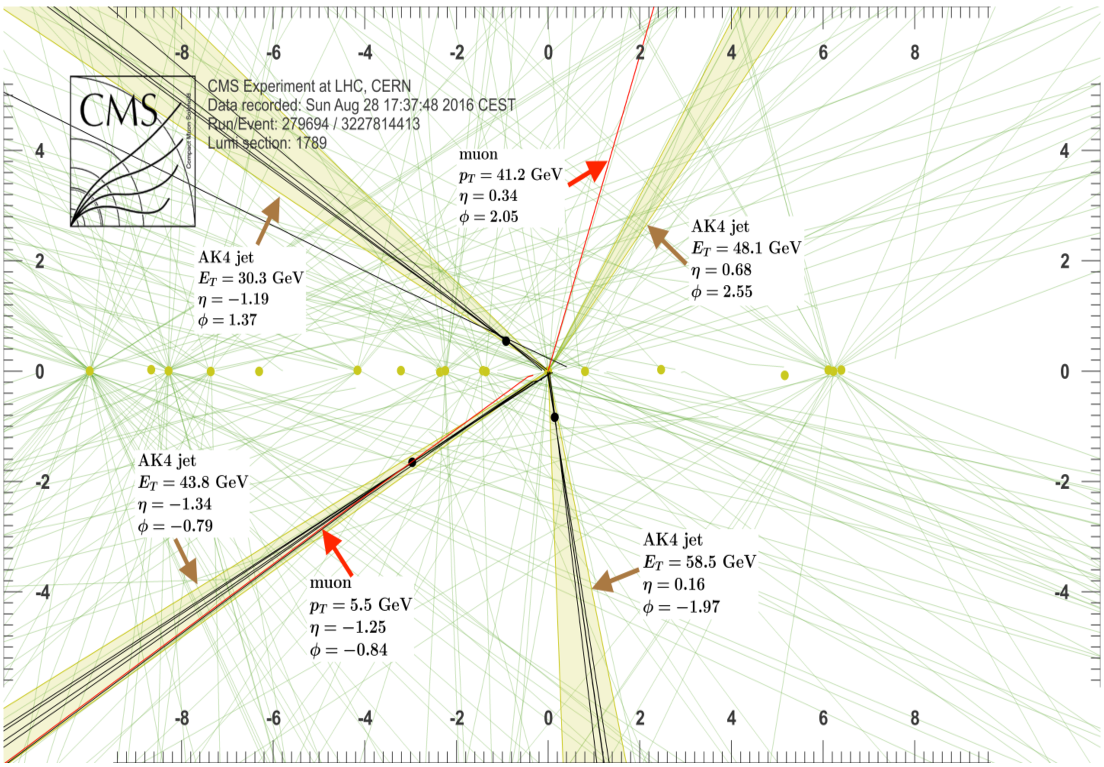

Figure 3.9: Recorded event reconstruction results;"Recorded event (ρ-z projection) with three jets with $p_T > 30$ GeV with one displaced muon track in 2016 data collected at 13 TeV. Each of the three jets has a displaced reconstructed vertex. The jet with $p_T(j) = 43.8$ GeV, $\eta(j) = -1.34$, $\phi(j) = -0.79$ contains muon with $p_T(\mu) = 5.5$ GeV, $\eta(\mu) = -1.25$, $\phi(\mu) = -0.84$. Event contains reconstructed isolated muon with $p_T(\mu) = 41.2$ GeV, $\eta(\mu) = 0.34$, $\phi(\mu) = 2.05$ and MET with $p_T = 72.5$ GeV, $\phi = -0.32$. Jet candidates for a b-jet from top quark leptonic and hadronic decays are tagged by CSVv2T algorithm. One of the other two jets is tagged by CharmT algorithm. Tracks with $p_T > 0.5$ GeV are shown. The number of reconstructed primary vertices is 18. Reconstructed m_T(W) is 101.8 GeV. Beam spot position correction is applied. Reconstructed primary vertices are shown in yellow color, while reconstructed displaced vertices and associated tracks are presented in black color. Dimensions are given in cm" [125].

Chapter 5

Statistical methods

In the course of analyzing the data sets provided by the CMS experiment and used in this thesis, several statistical tools have been employed; in this chapter, a description of these tools will be presented, starting with the general statement of the multivariate analysis methods, followed by the particularities of the Boosted Decision Trees (BDT) method and its application to the classification problem. Statistical inference methods used will also be presented. This chapter is based mainly on References [126–128].

5.1 Multivariate analysis

Multivariate data analysis (MVA) makes use of the statistical techniques developed to analyze more than one variable at once, taking into account all the correlations among variables. MVA is employed in a variety of fields like consumer and market research, quality control and process optimization. Using MVA it is possible to identify the dominant patterns in a data sample, like groups, outliers and trends, and determine to which group a set of values belong; in the particle physics context, MVA methods are used to perform the selection of certain type of events from a large data set.

Processes with small cross section, such as the tHq process ($\sigma_{SM}(\sqrt{s} = 13\text{TeV}) = 70.96$ fb), are hard to detect in the presence of the processes with larger cross sections, $\sigma^{t\bar{t}}_{SM}(\sqrt{s} = 13\text{TeV}) = 823.44$ fb for instance; therefore, only a small fraction of the data contains events of interest (signal), the major part is signal-like events, which mimic signal characteristics but belong to different processes, so they are a background to the process of interest. This implies that it is not possible to say with certainty that a given event is a signal or a background and statistical methods should be involved. In that sense, the challenge can be formulated as one where a set of events have to be classified according to certain special features; these features correspond to the measurements of several parameters like energy or momentum, organized in a set of *input variables*. The measurements for each event can be written in a vector $\mathbf{x} = (x_1, ..., x_n)$ for which

- $f(\mathbf{x}|s)$ is the probability density (*likelihood function*) that $\mathbf{x}$ is the set of measured values given that the event is a signal event (signal hypothesis).

- $f(\mathbf{x}|b)$ is the probability density (*likelihood function*) that $\mathbf{x}$ is the set of measured values given that the event is a background event (background hypothesis).

Figure 5.1 shows three ways to perform a classification of events for which measurements of two properties, i.e., two input variables x_1 and x_2, have been performed; blue circles represent signal events while red triangles represent background events. The classification on the left is *cut-based* requiring $x_1 < c_1$ and $x_2 < c_2$; usually the cut values (c_1 and c_2) are chosen according to some knowledge about the event process. In the middle plot, the classification is performed using a linear function of the input variables, hence the boundary is a straight line, while in the right plot the

the relationship between input variables is not linear thus the boundary is not linear either.

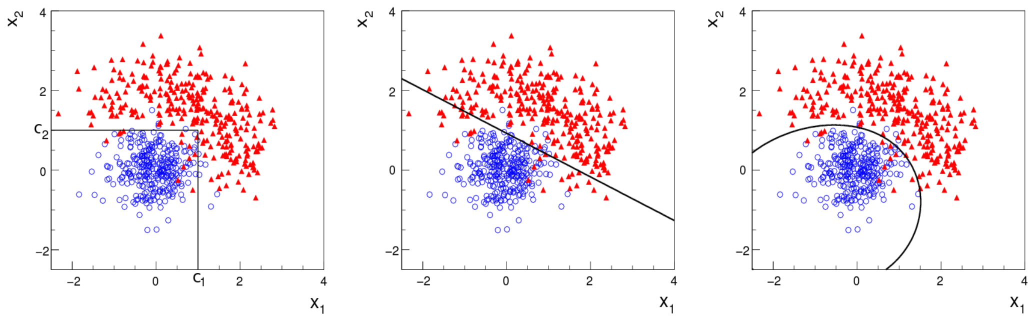

Figure 5.1: Scatter plots-MVA event classification. Distribution of two input variables x_1 and x_2 measured for a set of events; blue circles represent signal events and red triangles represent background events. The classification is based on cuts (left), linear boundary (center), and nonlinear boundary (right) [126]

.

In general, the boundary can be parametrized in terms of the input variables such that the cut is set on the parametrization instead of on the variables, i.e., $y(\mathbf{x}) = y_{cut}$ with y_{cut} being a constant; thus, the acceptance or rejection of an event is based on which side of the boundary the event is located. If $y(\mathbf{x})$, usually called *test statistic*, has functional form, it can be used to determine the probability distribution functions $p(y|s)$ and $p(y|b)$ and then perform a test statistic with a single cut on the scalar variable y.

Figure 5.2 shows an example of what would be the probability distribution functions under the signal and background hypotheses for a scalar test statistic with a cut on the classifier y. Note that the tails of the distributions indicate that some signal events fall in the rejection region and some background events fall on the acceptance region; therefore, it is convenient to define the *efficiency* with which events of a given type are accepted. The signal and background efficiencies are given by

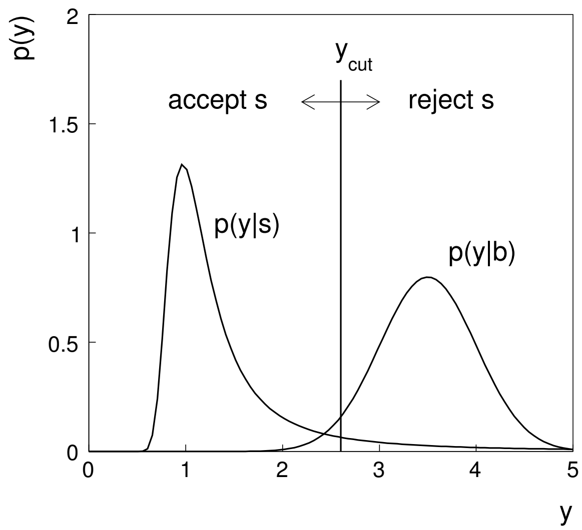

Figure 5.2: Distributions of the scalar test statistic $y(\mathbf{x})$ under the signal and background hypotheses. [126]

$$\varepsilon_{\mathrm{s}} = P(\text{accept event}|\mathrm{s}) = \int_{\mathrm{A}} f(\mathbf{x}|\mathrm{s})\, d\mathbf{x} = \int_{-\infty}^{y_{\mathrm{cut}}} p(y|\mathrm{s})\, dy\ , \tag{5.1}$$

$$\varepsilon_{\mathrm{b}} = P(\text{accept event}|\mathrm{b}) = \int_{\mathrm{A}} f(\mathbf{x}|\mathrm{b})\, d\mathbf{x} = \int_{-\infty}^{y_{\mathrm{cut}}} p(y|\mathrm{b})\, dy\ , \tag{5.2}$$

where A is the acceptance region. If the background hypothesis is the *null hypothesis (H_0)*, the signal hypothesis would be *alternative hypothesis (H_1)*; in this context, the background efficiency corresponds to the significance level of the test (α) and describes the misidentification probability, while the signal efficiency corresponds to the power of the test (1-β)[1] and describes the probability of rejecting the background hypothesis if the signal hypothesis is true. What is sought in an analysis is to maximize the power of the test relative to the significance level, i.e., set a selection with the largest possible selection efficiency and the smallest possible misidentification probability.

[1] β is the fraction of signal events that fall out of the acceptance region

5.1.1 Decision trees

For this thesis, the implementation of the MVA strategy, described above, is performed through decision trees by using the TMVA software package [127] included in the ROOT analysis framework [129]. In a simple picture, a decision tree classifies events according to their input variables values by setting a cut on each input variable and checking which events are on which side of the cut, just as proposed in the MVA strategy, but in addition, as a machine learning algorithm, decision trees offer the possibility to be trained and then perform the classification efficiently.

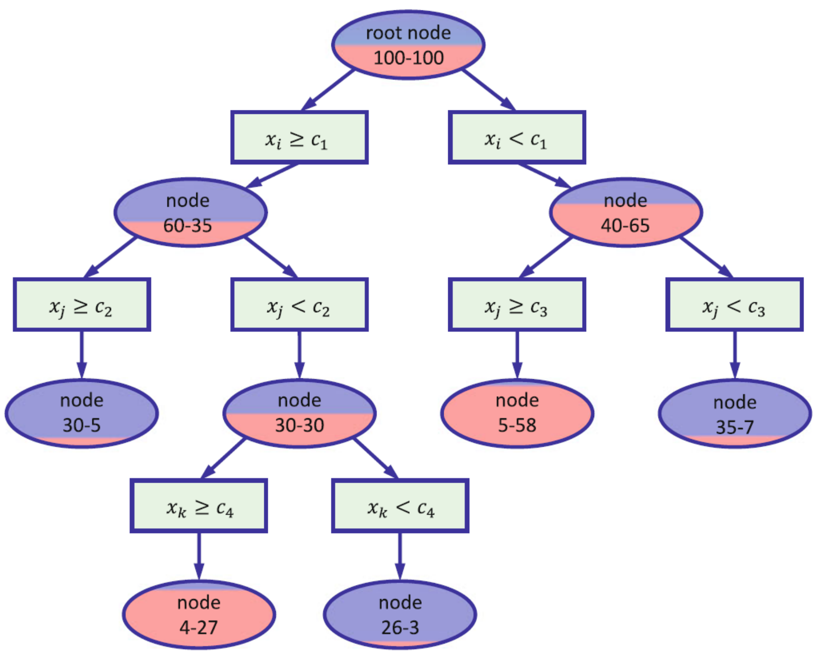

Figure 5.3: Example of a decision tree. Each node is fed with a MC sample mixing signal and background events (left-right numbers); nodes colors represent the relative number of signal/background events [128].

The training or growing of a decision tree is the process where the rules for classifying events are defined; this process is represented in Figure 5.3 and consists of several steps:

- take MC samples of signal and background events and split them into two parts each; the first parts will be used in the decision tree training, while the second parts will be used for testing the final classifier obtained from the training. Each event has associated a set of input variables $\mathbf{x} = (x_1, ..., x_n)$ which serve to distinguish between signal and background events. The training sample is taken in at the *root node*.

- Pick one variable, say x_i.
- Pick one value of x_i, each event has its own value of x_i, and split the training sample into two subsamples B_1 and B_2; B_1 contains events for which $x_i < c_1$ while B_2 contains the rest of the training events;
- scan all possible values of x_i and find the splitting value that provides the *best* classification[2], i.e., B_1 is mostly made of signal events while B_2 is mostly made of background events.
- It is possible that variables other than the picked one produce a better classification, hence, all the variables have to be evaluated. Pick the next variable, say x_j, and repeat the scan over its possible values.
- At the end, all the variables and their values will have been scanned, the *best* variable and splitting value will have been identified, say x_1, c_1, and there will be two nodes fed with the subsamples B_1 and B_2.

Nodes are further split by repeating the decision process until a given number of final nodes is obtained, nodes are largely dominated by either signal or background events, or nodes have too few events to continue. Final nodes are called *leaves* and they are classified as signal or background leaves according to the class of the majority of events in them. Each *branch* in the tree corresponds to a sequence of cuts.

[2] Quality of the classification will be treated in the next paragraph.

The quality of the classification at each node is evaluated through a separation criteria; there are several of them but the *Gini Index (G)* is the one used in the decision trees trained for the analysis in this thesis. G is written in terms of the purity (P), i.e., the fraction of signal events in the samples after the separation is made; it is given by

$$G = P(1-P) \tag{5.3}$$

note that P=0.5 at the root node while G=0 for pure leaves. For a node A split into two nodes B_1 and B_2 the G gain is

$$\Delta G = G(A) - G(B_1) - G(B_2). \tag{5.4}$$

The *best* classification corresponds to that for which the gain of G is maximized; hence, the scanning over all the variables in an event and their values is of great importance.

In order to provide a numerical output for the classification, events in a signal(background) leaf are assigned an score of 1(-1) each, defining in this way the decision tree *classifier/weak learner* as

$$f(\mathbf{x}) = \begin{cases} 1 & \mathbf{x} \quad \text{in signal region,} \\ -1 & \mathbf{x} \quad \text{in background region.} \end{cases}$$

Figure 5.4 shows an example of the classification of a sample of events, containing two variables, performed by a decision tree.

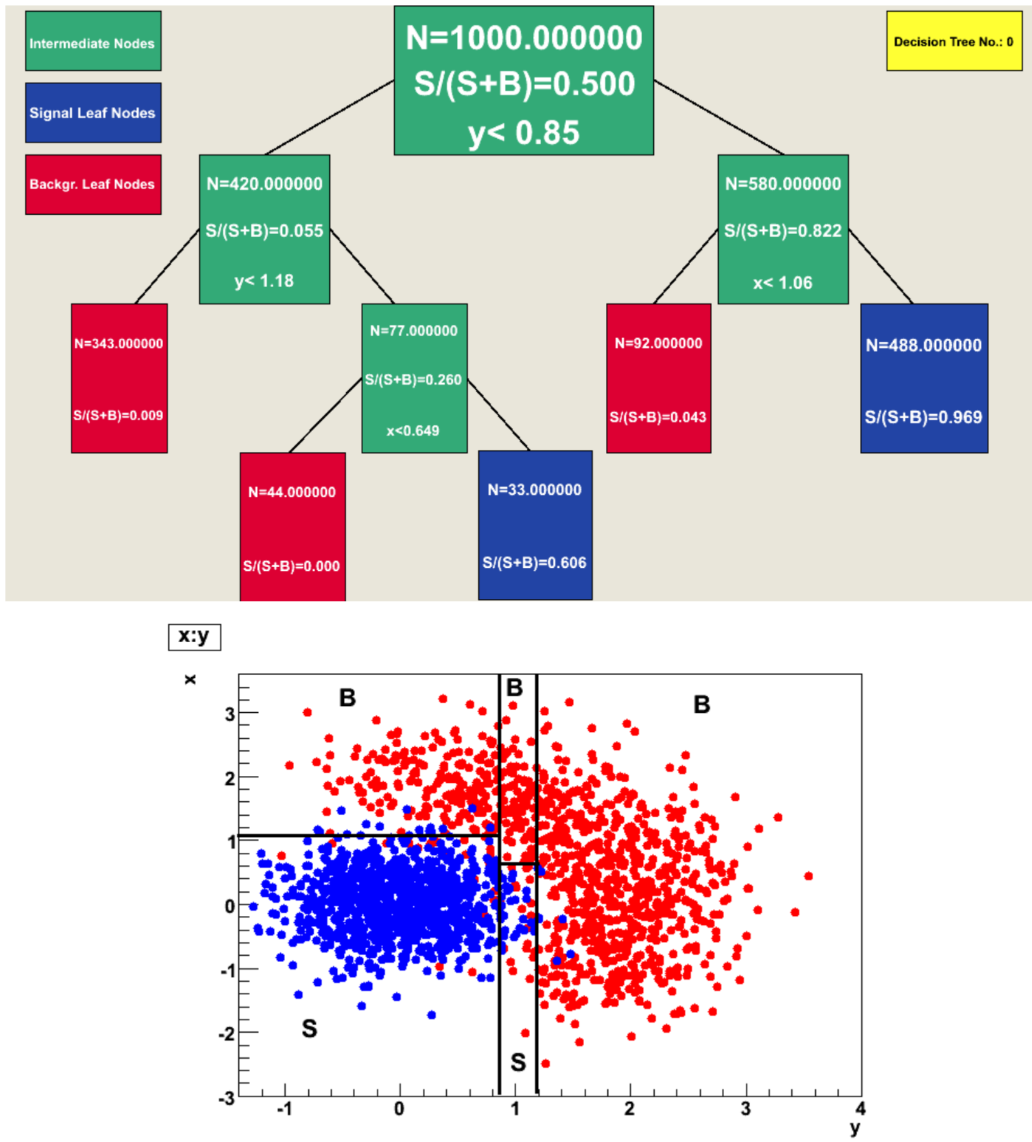

Figure 5.4: Example of a decision tree output. Each leaf, blue for signal events and red for background events, is represented by a region in the variables phase space [130].

5.1.2 Boosted decision trees (BDT).

Event misclassification occurs when a training event ends up in the wrong leaf, i.e., a signal event ends up in a background leaf or a background event ends up in a signal leaf. A way to correct it is to assign a weight to the misclassified events and train a second tree using the reweighted events; the event reweighting is performed by a boosting algorithm in such a way that when used in the training of a new decision tree the *boosted events* get correctly classified. The process is repeated iteratively adding a new tree to the forest and creating a set of classifiers, which are combined to create the next classifier; the final classifier offers more stability[3] and has a smaller misclassification rate than any individual ones. The resulting tree collection is known as a *boosted decision tree (BDT)*.

Thus, purity of the sample is generalized to

$$P = \frac{\sum_s w_s}{\sum_s w_s + \sum_b w_b} \tag{5.5}$$

where w_s and w_b are the weights of the signal and background events respectively; the Gini index is also generalized

$$G = \left(\sum_i^n w_i \right) P(1-P) \tag{5.6}$$

with n the number of events in the node. The final score of an event, after passing through the forest, is calculated as the renormalized sum of all the individual (possibly weighted) scores; thus, high(low) score implies that the event is most likely signal(background).

The boosting procedure, implemented in the *Gradient boosting* algorithm used in this thesis, produces a classifier $F(\mathbf{x})$ which is the weighted sum of the individual classifiers obtained after each iteration, i.e.,

$$F(\mathbf{x}) = \sum_{m=1}^{M} \beta_m f(\mathbf{x}; a_m) \tag{5.7}$$

where M is the number of trees in the forest. The *loss function* $L(F,y)$ represents the

[3] Decision trees suffer from sensitivity to statistical fluctuations in the training sample which may lead to very different results with an small change in the training samples.

deviation between the classifier $F(\mathbf{x})$ response and the true value y obtained from the training sample (1 for signal events and -1 for background event), according to

$$L(F,y) = \ln(1 + e^{-2F(\mathbf{x})y}) \tag{5.8}$$

thus, the reweighting is employed to ensure the minimization of the loss function; a more detailed description of the minimization procedure can be found in Reference [131]. The final classifier output is later used as a final discrimination variable, labeled as *BDT output/response*.

5.1.3 Overtraining

Decision trees offer the possibility to have as many nodes as desired in order to reduce the misclassification to zero (in theory); however, when a classifier is too much adjusted to a particular training sample, the classifier's response to a slightly different sample may leads to a completely different classification results; this effect is know as *overtraining*.

An alternative to reduce the overtraining in BDTs consists in prunning the tree by removing statistically insignificant nodes after the tree growing is completed but this option is not available for BDTs with gradient boosting in the TMVA-toolkit, therefore, the overtraining has to be reduced by tuning the algorithm, number of nodes, minimum number of events in the leaves, etc. The overtraining can be evaluated by comparing the responses of the classifier when running over the training and test samples.

5.1.4 Variable ranking

BDTs have a couple of particular advantages related to the input variables; they are relatively insensitive to the number of input variables used in the vector **x**. The ranking of the BDT input variables is determined by counting the number of times a variable is used to split decision tree nodes; in addition, the separation gain-squared achieved in the splitting and the number of events in the node are accounted by applying a weighting to that number. Thus, those variables with small or no power to separate signal and background events are rarely chosen to split the nodes, i.e., are effectively ignored.

In addition, variables correlations play an important role for some MVA methods like the Fisher discriminant algorithm in which the first step consist of performing a linear transformation to a phase space where the correlations between variables are removed; in the case of BDT algorithm, correlations do not affect the performance.

5.1.5 BDT output example

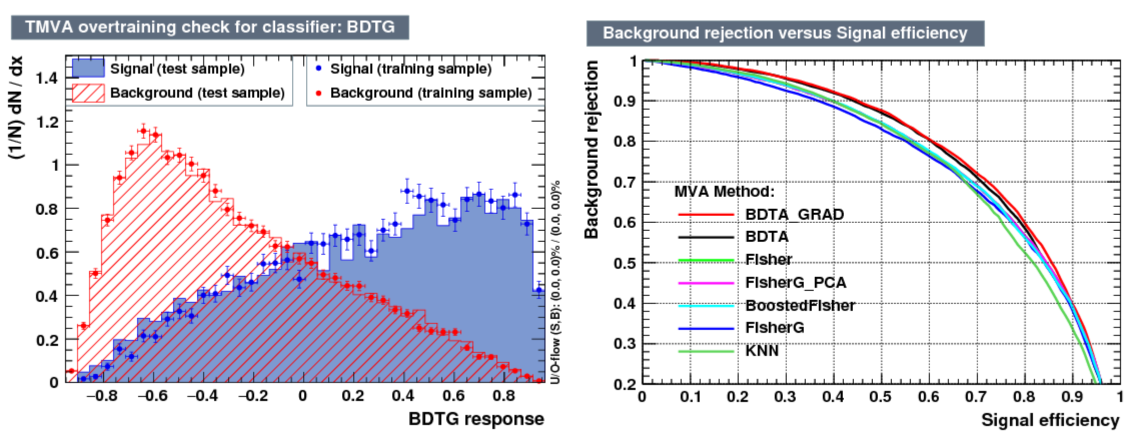

Figure 5.5: Left: Output distributions for the gradient boosted decision tree (BDTG) classifier using a sample of signal ($pp \to tHq$) and background ($pp \to t\bar{t}$) events. Right: Background rejection vs signal efficiency (ROC curves) for various MVA classifiers running over the same sample used to produce the plot on the left.

The left side of figure 5.5 shows the BDT output distributions for signal ($pp \to tHq$) and background ($pp \to t\bar{t}$) events; this plot is the equivalent to the one showed in Figure 5.2. A forest with 800 trees, maximum depth per tree = 3, and gradient boosting have been used as training parameters. The BDTG classifier offers a good separation power. There is a small overtraining in the signal distribution, while the background distribution is very well predicted which might indicate that the sample is composed of more background than signal events.

The right side of figure 5.5 shows the background rejection vs signal efficiency curves for several combinations of MVA classifiers-boosting algorithms running over the same MC sample; these curves are known as ROC curves and give an indication of the performance of the classifier. In this particular example, the best performance is achieved with the BDTG classifier (BDTA_GRAD), which motivate its use in this thesis.

5.2 Statistical inference

Once events are classified, the next step consists of finding the parameters that define the likelihood functions $f(\mathbf{x}|s), f(\mathbf{x}|b)$ for signal and background events respectively. In general, likelihood functions depend not only on the measurements but also on parameters (θ_m) that define their shapes; the process of estimating these *unknown parameters* and their uncertainties from the experimental data is called *inference*.

The statistical inference tools used in this analysis are implemented in the RooFit toolkit [132] and COMBINE package [133] included in the CMSSW software framework.

5.2.1 Nuisance parameters

The unknown parameter vector $\boldsymbol{\theta}$ is made of two types of parameters: those parameters that provide information about the physical observables of interest for the experiment or *parameters of interest*, and the *nuisance parameters* that are not of direct interest for the experiment but that need to be included in the analysis in order to achieve a satisfactory description of the data; they represent effects of the detector response like the finite resolutions of the detection systems, miscalibrations, and in general any source of uncertainty introduced in the analysis.

Nuisance parameters can be estimated from experimental data; for instance, data samples from a test beam are usually employed for calibration purposes. In cases where experimental samples are not availables, the estimation of nuisance parameters makes use of dedicated simulation programs to provide the required samples.

The estimation of the unknown parameters involves certain deviations from their true values, hence, the measurement of the nuisance parameter is written in terms of an estimated value, also called central value, $\hat{\theta}$ and its uncertainty $\delta\theta$ using the notation

$$\theta = \hat{\theta} \pm \delta\theta \tag{5.9}$$

where the interval $[\hat{\theta} - \delta\theta, \hat{\theta} + \delta\theta]$ is called *confidence interval*; it is usually interpreted, in the limit of infinite number of experiments, as the interval where the true value of the unknown parameter θ is contained with a probability of 0.6827 (if no other convention is stated); this interval represents the area under a Gaussian distribution in the interval $\pm 1\sigma$.

The uncertainties associated with nuisance parameters produce *systematic uncertainties* in the final measurement, while the uncertainties related only to fluctuations in data and that affect the determination of parameters of interest produce *statistical uncertainties*.

5.2.2 Maximum likelihood estimation method

The estimation of the unknown parameters that are in best agreement with the observed data is performed through a function of the data sample that returns the estimate of those parameters; that function is called an *estimator*. Estimators are usually constructed using mathematical expressions encoded in algorithms.

In this thesis, the estimator used is the likelihood function $f(\mathbf{x}|\boldsymbol{\theta})$[4] which depends on a set of measured variables $\mathbf{x}$ and a set of unknown parameters $\boldsymbol{\theta}$. The likelihood function for N events in a sample is the combination of all the individual likelihood functions, i.e.,

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{N} f(\mathbf{x}^i|\boldsymbol{\theta}) = \prod_{i=1}^{N} f(x_1^i, ..., x_n^i; \theta_1, ..., \theta_m) \tag{5.10}$$

and the estimation method used is the *Maximum Likelihood Estimation* method (MLE); it is based on the combined likelihood function defined by eqn. 5.10 and the procedure seeks for the parameter set that corresponds to the maximum value of the combined likelihood function, i.e., the *maximum likelihood estimator* of the unknown parameter vector $\boldsymbol{\theta}$ is the function that produces the vector of *best estimators* $\hat{\boldsymbol{\theta}}$ for which the likelihood function $L(\boldsymbol{\theta})$ evaluated at the measured $\mathbf{x}$ is maximum.

Usually, the logarithm of the likelihood function is used in numerical algorithm implementations in order to avoid underflow the numerical precision of the computers due to the product of low likelihoods. In addition, it is common to minimize the negative logarithm of the likelihood function, therefore, the negative log-likelihood

[4] analogue to the likelihood functions described in previous sections

function is

$$F(\boldsymbol{\theta}) = -\ln L(\boldsymbol{\theta}) = -\sum_{i=1}^{N} f(\mathbf{x}^i|\boldsymbol{\theta}). \tag{5.11}$$

The minimization process is performed by the software MINUIT [134] implemented in the ROOT analysis framework. In case of data samples with large number of measurements, the computational resources necessary to calculate the likelihood function are too big; therefore, the parameter estimation is performed using binned distributions of the variables of interest for which the *binned likelihood function* is given by

$$L(\mathbf{x}|r,\boldsymbol{\theta}) = \prod_{i=1} \frac{(r \cdot s_i(\boldsymbol{\theta}) + b_i(\boldsymbol{\theta}))^{n_i}}{n_i!} e^{-r \cdot s_i(\boldsymbol{\theta}) - b_i(\boldsymbol{\theta})} \prod_{j=1} \frac{1}{\sqrt{2\pi\sigma_{\theta_j}^2}} e^{-(\theta_j - \theta_{0,j})^2/2\sigma_{\theta_j}^2}, \tag{5.12}$$

with s_i and b_i the expected number of signal and background yields for the bin i, n_i is the observed number of events in the bin i and $r = \sigma/\sigma_{SM}$ is the signal strength. Note that the number of entries per bin follows a Poisson distribution. The effect of the nuisance parameters have been included in the likelihood function through the multiplication by a Gaussian distribution that models the nuisance. The three parameters, r, s_i and b_i are jointly fitted to estimate the value of r.

5.3 Upper limits

In this analysis, two hypotheses are considered; the background only hypothesis ($H_0(b)$) and the signal plus background hypothesis ($H_1(s+b)$), i.e., the sample of events is composed of background only events (r=0) or it is a mixture of signal plus background events (r=1). The exclusion of one hypothesis against the other means that the observed data sample better agrees with H_0 or rather with H_1. In order to discriminate these hypotheses, a test statistic is constructed on the basis of the likelihood function evaluated for each of the hypothesis.

The *Neyman-Pearson* lemma [135] states that the test statistic that provides the maximum power for H_1 for a given significance level (background misidentification probability α), is given by the ratio of the likelihood functions $L(\mathbf{x}|H_1)$ and $L(\mathbf{x}|H_0)$; however, in order to use that definition it is necessary to know the true likelihood functions, which in practice is not always possible. Approximate functions obtained by numerical methods, like the BDT method described above, have to be used, so that the *profile likelihood* test statistic is defined by

$$\lambda(\mathbf{r}) = \frac{L(\mathbf{x}|r, \hat{\hat{\boldsymbol{\theta}}}(r))}{L(\mathbf{x}|\hat{r}, \hat{\boldsymbol{\theta}})}, \tag{5.13}$$

where, $\hat{r}$ and $\hat{\boldsymbol{\theta}}$ maximize the likelihood function, and $\hat{\hat{\boldsymbol{\theta}}}$ maximizes the likelihood function for a given value of the signal strength modifier r. In practice, the test statistic t_r

$$t_r = -2\ln\lambda(r) \tag{5.14}$$

is used to evaluate the presence of signal in the sample, since the minimum of t_r at $r = \hat{r}$ suggests the presence of signal with signal strength $\hat{r}$. The uncertainty interval for r is determined by the values of r for which $t_r = +1$.

The expected probability density function (p.d.f) $f(t_r|r, \boldsymbol{\theta})$ of the test statistic t_r can be obtained numerically by generating MC samples where one hypothesis, $H_0(b)$ or $H_1(s+b)$, is assumed; thus, MC samples contain the possible values of t_r obtained from *pseudo-experiments* as shown in Figure 5.6. The probability that t_r takes a value equal or greater than the observed value ($t_{r,obs}$) when a signal with a signal modifier r is present in the data sample, is called the *p-value* of the observation; it can be calculated using

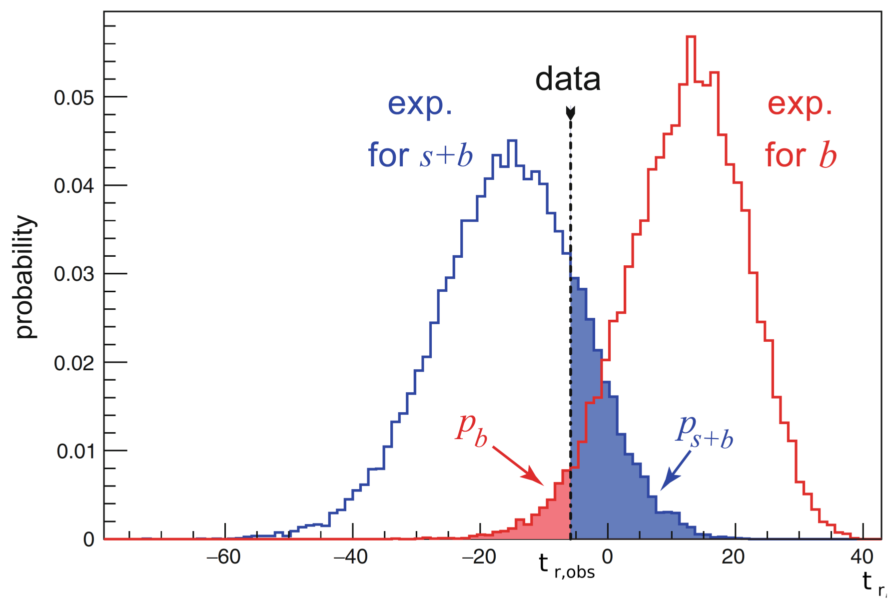

Figure 5.6: t_r p.d.f. from MC pseudo experiments assuming H_0 (red) and H_1 (blue). The black dashed line shows the value of the test statistic as measured from data. Adapted from Reference [128].

$$p_r = \int_{t_{r,obs}}^{\infty} f(t_r'|r,\boldsymbol{\theta})dt_r', \tag{5.15}$$

thus, $p_r < 0.05$ means that, for that particular value of r, H_1 could be excluded at 95% Confidence Level (CL). The corresponding background-only p-value is given by

$$1 - p_b = \int_{t_{r,obs}}^{\infty} f(t_r'|0,\boldsymbol{\theta})dt_r', \tag{5.16}$$

If the t_r p.d.f.s for both hypotheses are well separated, as shown in the top side of Figure 5.7, the experiment is sensitive to the presence of signal in the sample. If the signal presence is small, both p.d.f.s will be largely overlapped (bottom of Figure **??**) and either the signal hypothesis could be rejected with not enough justification because the experiment is not sensitive to the signal or a fluctuation of the background could be misinterpreted as presence of signal with the corresponding rejection of the

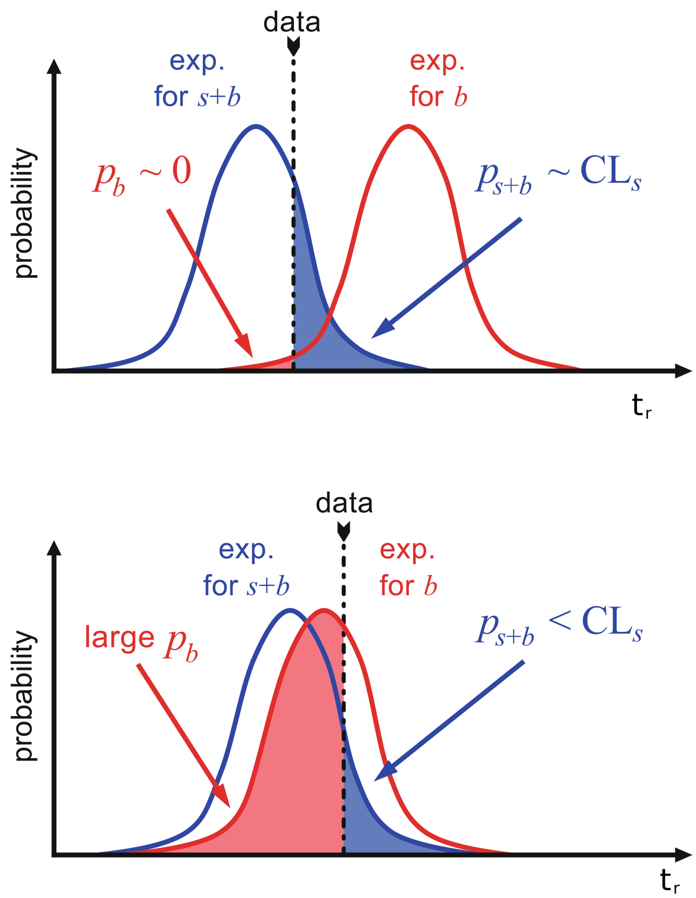

Figure 5.7: CL_s limit illustration. When the test statistic p.d.f. for the two hypotheses H_0 and H_1 are well separated (top) and when they are largely overlapped (bottom). Adapted from Reference [128].

background-only hypothesis. These issues are corrected by using the modified p-value [136]

$$p'_r = \frac{p_r}{1-p_b} \equiv CL_s. \tag{5.17}$$

If H_1 is true, then p_b is small, $CL_s \simeq p_r$ and H_0 is rejected; if there is large overlap and an statistical fluctuation cause that p_b is large, then both numerator and denominator in Eqn. 5.17 become small but CL_s would allow the rejection of H_1 even if there is poor sensitivity to signal.

The upper limit of the parameter of interest r^{up} is determined by excluding the range of values of r for which $CL_s(r,\boldsymbol{\theta})$ is lower than the confidence level desired, normally 90% or 95%, e.g, scanning over r and finding the value for which $p_r^{'up} =$ 0.05. The expected upper limit can be calculated using pseudo-experiments based on the background-only hypothesis and obtaining a distribution for r_{ps}^{up}; the median of that distribution corresponds to the expected upper limit, while the $\pm1\sigma$ and $\pm2\sigma$ deviations correspond to the values of the distribution that defines the 68% and 95% of the area under the distribution centered in the median. It is usual to present all the information about the expected and observed limits in the so-called *Brazilian-flag plot* as the one showed in Figure 5.8. The solid line represent the observed CL_s

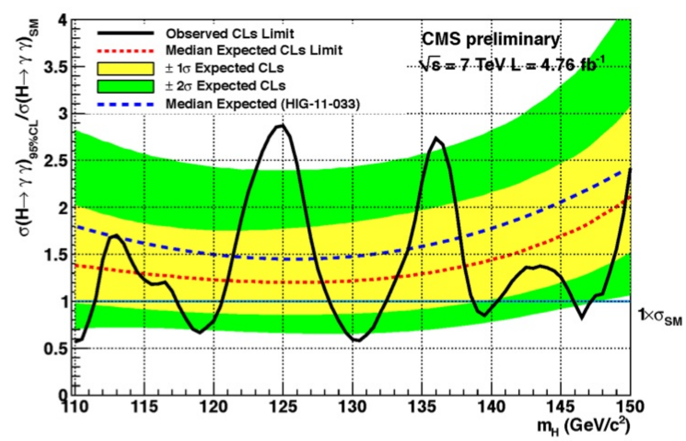

Figure 5.8: Brazilian flag plot of CMS experiment limits for Higgs boson decaying to photons [137].

5.4 Asymptotic limits

As said before, the complexity of the likelihood functions, the construction of test statistics, and the calculation of the limits and their uncertainties is not always manageable and requires extensive computational resources; in order to overcome those issues, asymptotic approximations for likelihood-based test statistics, like the ones described in previous sections, have been developed [138, 139] using Wilks' theorem. Asymptotic approximations replace the construction of the test statistics p.d.f.s using MC pseudo-experiments, with the approximate calculation of the test statistics p.d.f.s by employing the so-called *Asimov dataset*.

The Asimov dataset is defined as the dataset that produce the true values of the nuisance parameters when it is used to evaluate the estimators for all the parameters; it is obtained by setting the values of the variables in the dataset to their expected values [139].

Limits calculated by using the asymptotic approximation and the Asimov dataset are know as *asymptotic limits*.

Chapter 6

Search for production of a Higgs boson and a single top quark in multilepton final states in pp collisions at $\sqrt{s} = 13$ TeV

6.1 Introduction

The Higgs boson discovery, supported on experimental observations and theoretical predictions made about the SM, gives the clue of the way in that elementary particles acquire mass through the Higgs mechanism; therefore, knowing the Higgs mass, the Higgs-vector boson and Higgs-fermion couplings can be determined. In order to test the Higgs-top coupling, several measurements have been performed, as stated in the chapter 1, but they are limited in sensitivity to measure the square of the coupling. The production of a Higgs boson in association with a single top quark (tH) not only offers access to the sign of the coupling, but also, to the CP phase of the Higgs couplings.

This chapter presents the search for the associated production of a Higgs boson and a single top quark (tHq) events, focusing on leptonic signatures provided by the Higgs decay modes to WW, ZZ, and $\tau\tau$; the 13 TeV dataset produced in 2016, which corresponds to an integrated luminosity of $35.9 fb^{-1}$, is used.

As shown in Section 1.5, the SM cross section of tHq process is driven by a destructive interference between two contributions (see Figure 1.15), where the Higgs couples to either the W boson or the top quark; however, if the sign of the Higgs-top coupling is flipped with respect to the SM prediction, a large enhancement of the cross section occurs, making this analysis sensitive to such deviation. A second process, where the Higgs boson and top quark are accompanied by a W boson (tHW) has similar behavior, albeit with a weaker interference pattern and lower contribution to the cross section, therefore, a combination of both processes would increase the sensitivity to the sign of the coupling; in this analysis both contributions are combined and referred as tHchannel. A third contribution comes from $t\bar{t}H$ process. The purpose of this analysis is to investigate the exclusion of the presence of the tH+ $t\bar{t}H$ processes under the assumption of the anomalous Higgs-top coupling modifier (κ_t =-1). The analysis exploits signatures with two leptons of the same sign (*2lss*) channel and three leptons (*3l*) channel in the final state.

Constraints on the sign of the Higgs-top coupling (y_t) have been derived from the decay rate of Higgs boson to photon pairs [50] and from the cross section for associated production of Higgs and Z bosons via gluon fusion [141], with recent results disfavoring negative signs of the coupling [44, 59, 142], although the negative sign coupling have not been completely excluded.

The analysis presented here, expands previous analyses performed at 8 TeV [143, 144] and searches for associated production of $t\bar{t}$ pair and a Higgs boson in the multilepton final state channel [145]; it also complements searches in other decay channels targeting $H \to b\bar{b}$ [146].

The first sections present the characteristic tHq signature as well as the expected backgrounds. The MC samples, data sets, and the physics object definitions are then defined. Following, the background predictions, the signal extraction, and the statistical treatment of the selected events as well as the systematic uncertainties are described. The final section present the results for the exclusion limits as a function of the ratio of κ_t and the dimensionless modifier of the Higgs-vector boson coupling κ_{V}.

The analysis is designed to efficiently identify and select prompt leptons from on-shell W and Z boson decays and to reject non-prompt leptons from b quark decays and spurious lepton signatures from hadronic jets. Events are then selected in the $2lss$ and $3l$ channels, and are required to contain hadronic jets, some of which must be consistent with b quark hadronization. Finally, the signal yield is extracted by simultaneously fitting the output of two dedicated multivariate discriminants, trained to separate the tHq signal from the two dominant backgrounds, in all categories. The fit result is then used to set an upper limit on the combined $t\bar{t}H$ + tHproduction cross section, as a function of the relative coupling strengths of Higgs-top quark and Higgs-Vector boson. Figure 6.1 shows an schematic overview of the analysis strategy workflow.

With respect to the 8 TeV analysis, the object selections have been adjusted for the updated LHC running conditions at 13 TeV, the lepton identification has been improved, and more powerful multivariate analysis techniques are used for the signal extraction.

The analysis has been made public by CMS as a Physics Analysis Summary [147] combining the result for the three lepton and two lepton same-sign channels; the

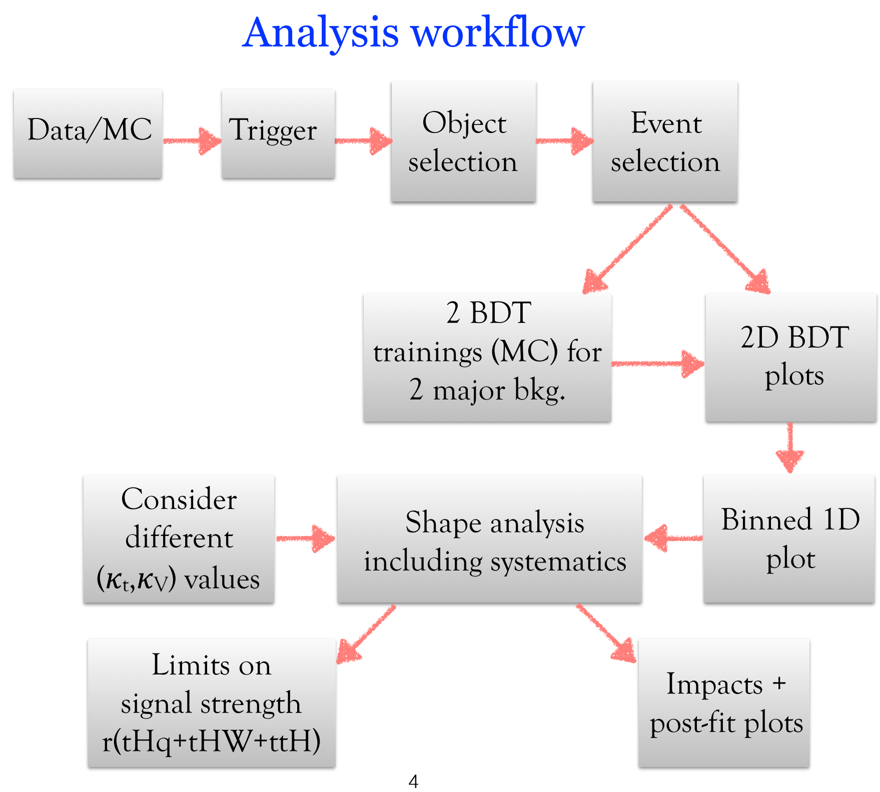

Figure 6.1: A schematic overview of the analysis workflow. Based on sets of optimized physics object definitions and selection criteria, signal and background events in a data sample are discriminated. The discrimination is performed by a BDT, previously trained using MC samples of the dominant backgrounds, using discriminant variables based on the b-jet multiplicity, the activity in the forward region of the detector, and the kinematic properties of leptons. The CL_s limits on the combined $t\bar{t}H + tH$production cross section, as a function of the relative coupling strengths are calculated.

content present in this chapter is based on that document and on References [145,149] unless other Reference is stated. Currently, an effort to turn the analysis into a paper combining the multilepton and $H \to b\bar{b}$ is ongoing.

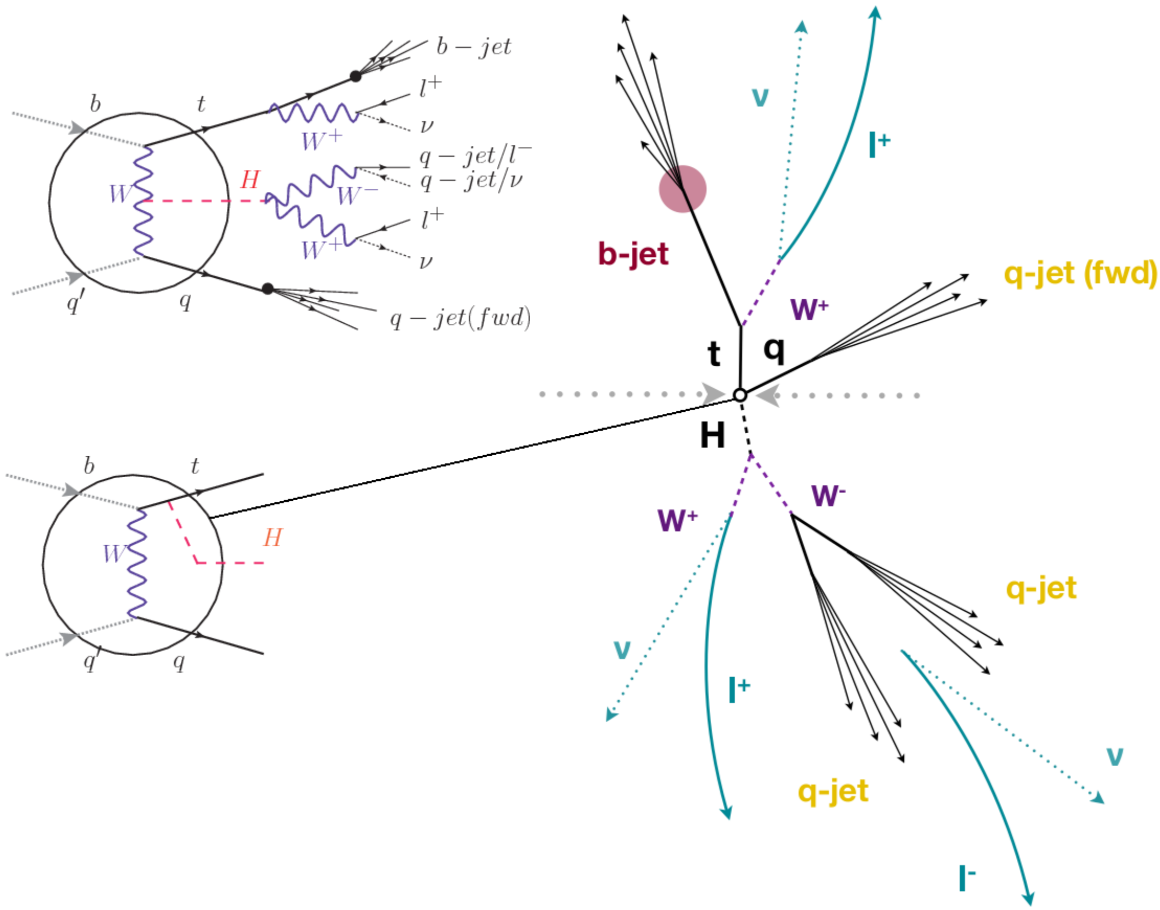

Figure 6.2: tHq event signature. Left: Feynman diagram including the whole evolution up to the final state for the case of the Higgs boson emitted by the W boson (top); Feynman diagram for the case where the Higgs boson is emitted by the top quark. Right: Schematic view as it would be seen in the detector; the circle in the Feynman diagrams on the left corresponds to the circle in the center of the schematic view as indicated by the line connecting them. In the $2lss$ channel, one of the W bosons from the Higgs boson decays to two light-quark jets while in the $3l$ channel both W bosons decay to leptons.

6.2 tHq signature

In order to select events of tHq process, its features are translated into a set of selection rules; Figure 6.2 shows the Feynman diagram and an schematic view of the tHq process from the pp collision to the final state configuration. A single top quark is produced accompanied by a light quark, denoted as q; this light quark is produced predominantly in the forward region of the detector. The Higgs boson can be either emitted by the exchanged W boson or directly by the singly produced top quark.

Due to their high masses/short lifetimes, top quark and Higgs boson decay after their production within the detector. The Higgs boson is required to decay into a W boson pair[1]. The top quark almost always decays into a bottom quark and a W boson, as encoded in the CMK matrix. The W bosons are required to decay leptonically either all the three in the $3l$ channel or the pair with equal electrical charge in the $2lss$ channel case; τ leptons are not reconstructed separately and only their leptonic decays into either electrons or muons are considered in this analysis.

In summary, the signal process is characterized by a the final state with

- one light-flavored forward jet,
- one central b-jet,
- $2lss$ channel $\rightarrow$ two leptons of the same sign, two neutrinos and two light (often soft) jets,
- $3l$ channel $\rightarrow$ three leptons, three neutrinos and no central light-flavored jets,

The presence of neutrinos is inferred from the presence of MET.

6.3 Background processes

The background processes are those that can mimic the signal signature or at least can be reconstructed as that as a result of certain circumstances. The backgrounds can be classified as

[1] ZZ and $\tau\tau$ decays are also include in the analysis but they are not separately reconstructed

- irreducible backgrounds: where genuine prompt leptons are produced in on-shell W and Z boson decays; they can be reliantly estimated directly from MC simulated events, using higher-order cross sections or data control regions for the overall normalization.

- reducible backgrounds: where at least one of the leptons is *non-prompt*, i.e., produced within a hadronic jet; genuine leptons from heavy flavor decays and misreconstructed jets, also known as *mis-ID leptons* or *fake leptons*, are considered non-prompt leptons. These non-prompt leptons leave tracks and hits in the detection systems as would a prompt lepton, but correlating those hits with nearby jets could be a way of removing them. The misassignment of electron charge in processes like $t\bar{t}$ or Drell-Yan, represent an additional source of background, but it is relevant only for the $2lss$ channel. Reducible backgrounds are not well predicted by simulation, hence, they are estimated using data-driven methods.

The main sources of background events for tHq process are $t\bar{t}$ process and $t\bar{t} + X(X = W, Z, \gamma)$ processes, the latter regarded together as $t\bar{t}$Vprocess. Figure 6.3 shows the signature for $t\bar{t}$ and $t\bar{t}W$processes.

The largest contribution to irreducible backgrounds comes from $t\bar{t}W$and $t\bar{t}Z$processes for which the number of $(b-)$jets ($(b-)$jet multiplicity) is higher than that of the signal events, while for other contributing background events, WZ, ZZ, and rare SM processes like $W^{\pm}W^{\pm}qq$, $t\bar{t}t\bar{t}$, tZq, tZW, WWW, WWZ, WZZ, ZZZ, the $(b-)$jet multiplicity is lower compared to that of the signal events. None of the irreducible backgrounds present activity in the forward region of the detector.

On the side of the reducible backgrounds, the largest contribution comes from the $t\bar{t}$ events which have a very similar signature to the signal events but does no present

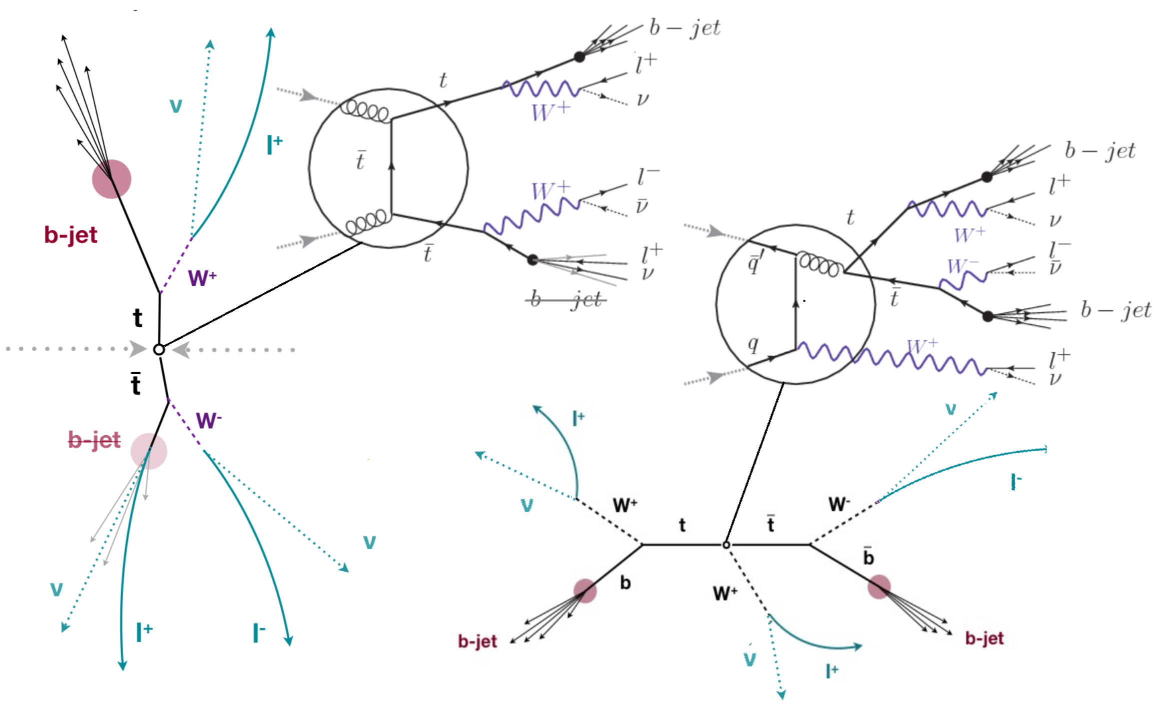

Figure 6.3: $t\bar{t}$ (left) and $t\bar{t}W$ (right) events signature as they would be seen in the detector; the Feynman diagrams including the whole evolution up to the final state are also showed. The $t\bar{t}$ process signature is very similar to that of the signal process with one fake lepton and non forward activity. The $t\bar{t}W$ process present a higher b-jet multiplicity compared to the signal process, a prompt lepton and no forward activity.

activity in the forward region of the detector either; A particular feature of the $t\bar{t}$ events is their charge-symmetry, which is also a difference with respect to the signal events.

6.4 Data and MC Samples

6.4.1 Full 2016 data set

The data set used in this analysis was collected by the CMS experiment during 2016 at while running at $\sqrt{s} = 13$TeV and corresponds to a total integrated luminosity of $35.9 fb^{-1}$. Only periods when the CMS magnet was on were considered when selecting the data samples; that corresponds to the `23Sep2016 (Run B to G)` and `PromptReco (Run H)` versions of the datasets.

Multilepton final states with either two same-sign leptons or three leptons target the case where the Higgs boson decays to a pair of W bosons, τ leptons, or Z bosons, and where the top quark decays leptonically, hence, the `SingleElectron`, `SingleMuon`, `DoubleEG`, `MuonEG`, `DoubleMuon` dataset (see Table A.1) compose the full dataset. The certified luminosity sections are selected using the golden JSON file defined by the CMS experiment [148].

6.4.2 Triggers

The events considered are those online-reconstructed events triggered by one, two, or three leptons. Single-lepton triggers are included in order to boost the acceptance of events where the p_T of the sub-leading lepton falls below the threshold of the double-lepton triggers. The trigger efficiency is increased by including double-lepton triggers in the $3l$ category, and single-lepton triggers in all categories; it is possible given the logical "or" of the trigger decisions of all the individual triggers in a given category. Table A.2 shows the lowest-threshold non-prescaled triggers present in the High-Level Trigger (HLT) menus for both Monte-Carlo and data in 2016.

Trigger efficiency scale factors

Trigger efficiency describes the ability of events to pass the trigger requirements. It is measured in simulated events using generator information given that there is no trigger bias with the MC sample. Measuring the trigger efficiency in data requires a more elaborated procedure; first, select a set of events collected by a trigger that is uncorrelated with the lepton triggers such that the selected events form an unbiased

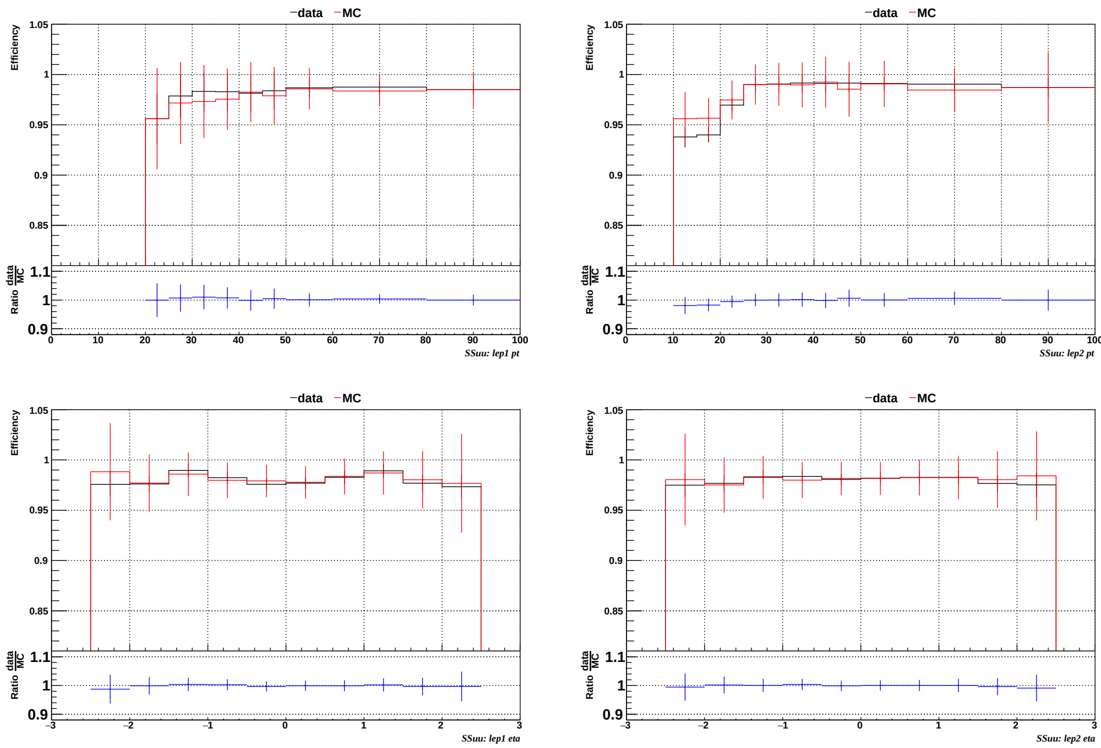

Figure 6.4: Comparison between data an MC trigger efficiencies in the same-sign $\mu\mu$ category, as as a function of the p_T and η of the leading lepton (left) and the sub-leading lepton (right) [149].

sample. In this analysis, that uncorrelated trigger is a MET trigger. Second step is looking for candidate events with exactly two good leptons (exactly three good leptons for the $3l$ channel). Finally, measure the efficiency for the candidate events to pass the logical "or" of triggers being considered in a given event category as defined in Table A.2.

Comparisons between the data and MC efficiencies for each category, showed in Figures 6.4, 6.5, and 6.6, reveal that they are in good agreement; the difference is corrected by applying scale factors derived from the ratio between both efficiencies.

Applied flat scale factors in each category are shown in Table 6.1; they have been inherited from Reference [149].

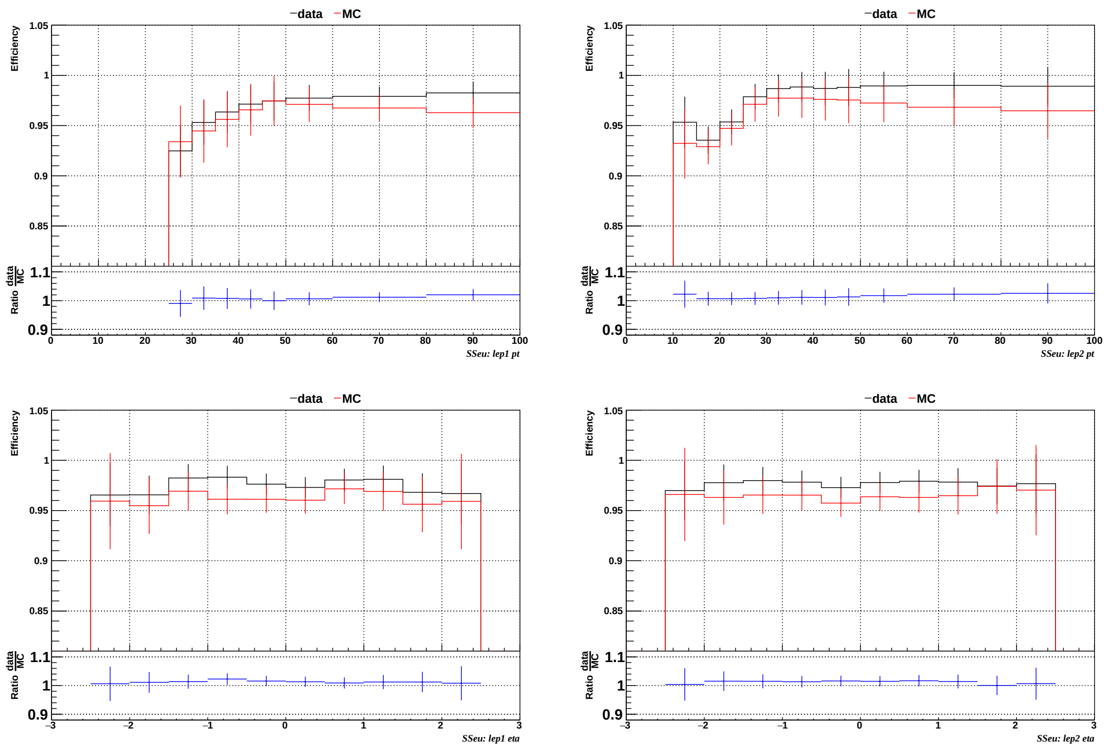

Figure 6.5: Comparison between data an MC trigger efficiencies in the same-sign $e\mu$ category as as a function of the p_T and η of the leading lepton (left) and the sub-leading lepton (right) [149].

Category	Scale Factor
ee	1.01 ± 0.02
eμ	1.01 ± 0.01
$\mu\mu$	1.00 ± 0.01
3l	1.00 ± 0.03

Table 6.1: Trigger efficiency scale factors and associated uncertainties, shown here rounded to the nearest percent.

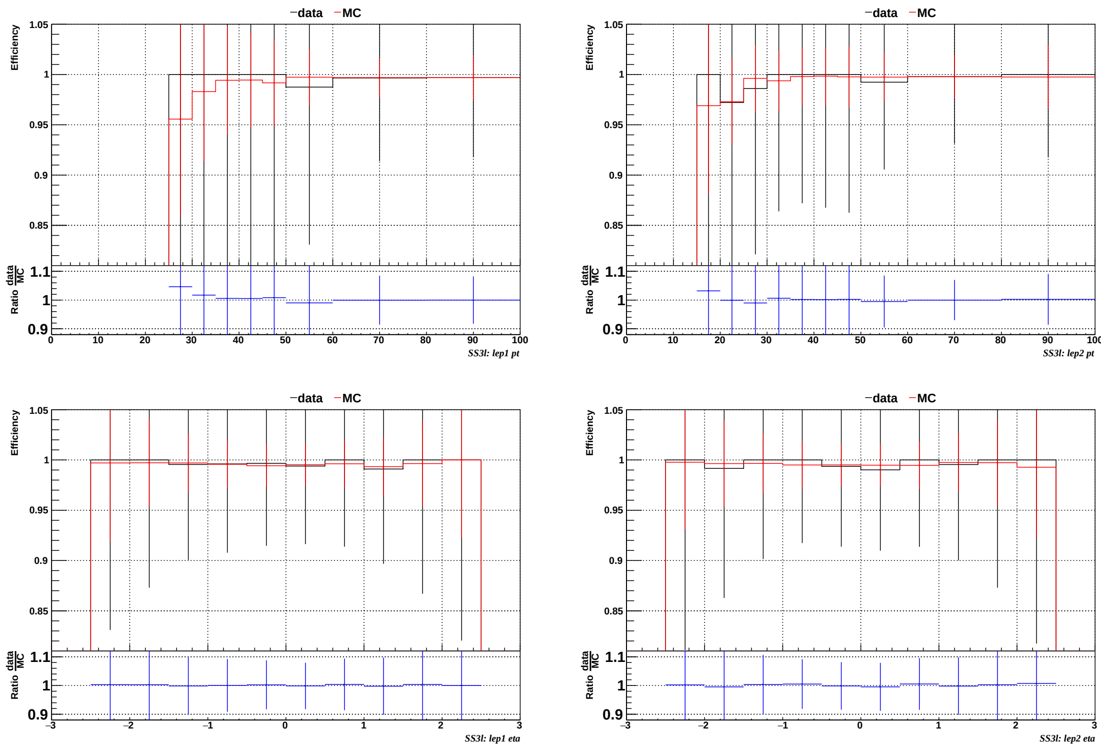

Figure 6.6: Comparison between data an MC trigger efficiencies in the $3l$ category, as as a function of the p_T and η of the leading lepton (left) and the sub-leading lepton (right) [149].

6.4.3 Signal modeling and MC samples

Current event generators allow for adjusting the kinematics of the generated events, based on an event-wise reweighting; in this way, several generation parameters phase spaces can be explored according to the experimental interests. The signal samples used in this analysis were generated in such a way that not only the case κ_t =-1, but an extended range of κ_t and κ_{V} values may be investigated.

tHq and tHW cross section in the κ_t-κ_{V} phase space are shown in Figure 6.7. As said in section 3.1, the tHq sample was generated using the 4F scheme which provides a better description of the additional b quark from the initial gluon splitting, while the tHW sample was generated using the 5F scheme in order to remove its the interference with $t\bar{t}H$ at LO.

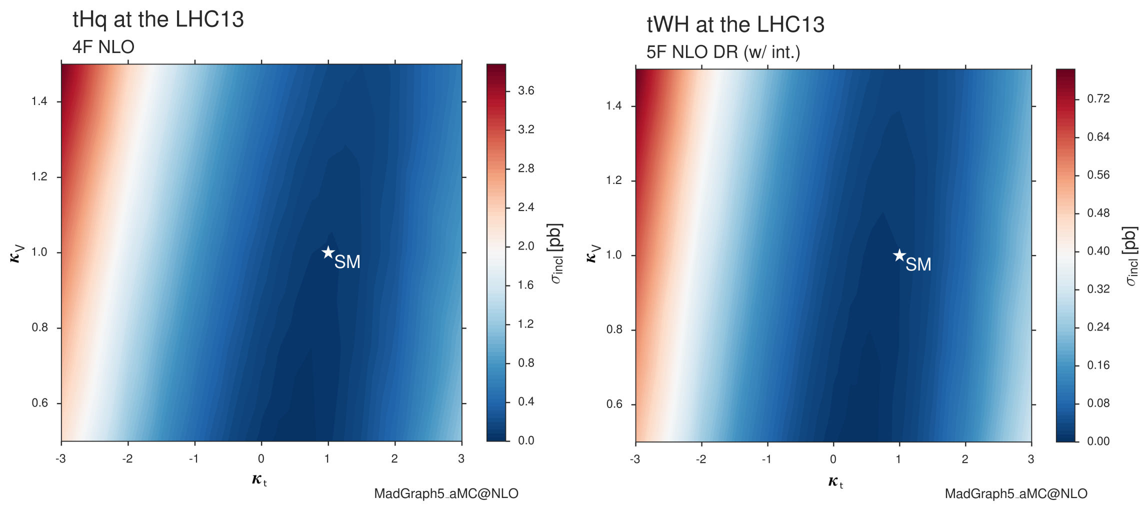

Figure 6.7: tHq and tHW cross section in the κ_t-κ_V phase space [150].

MC signal samples

The two signal samples, tHq and tHW, correspond to the `RunIISummer16MiniAODv2` campaign produced with CMSSW_80X; they were produced with MG5_aMC@NLO (version 5.2.2.3), in LO order mode at $\sqrt{s} = 13$ TeV, and are normalized to NLO cross sections (see Table 6.2). The Higgs boson is assumed to be SM-like except for the values of its couplings to the top quark and W boson. Each sample was generated with a set of event weights corresponding to 51 different values of (κ_t, κ_V) couplings, accessible in terms of LHE event weights as shown in Table A.3; however, the main interest is the (κ_t =-1,κ_V =1) case.

Sample	σ [pb]	BF
`/THQ_Hincl_13TeV-madgraph-pythia8_TuneCUETP8M1/`	0.7927	0.324
`/THW_Hincl_13TeV-madgraph-pythia8_TuneCUETP8M1/`	0.1472	1.0
`/ttHJetToNonbb_M125_13TeV_amcatnloFXFX_madspin_pythia8_mWCutfix/`	0.2151	1.0

Table 6.2: MC signal samples used in this analysis; cross section and branching fraction are also listed [150].

The $t\bar{t}H$ sample was produced using AMC@NLO interfaced to PYTHIA 8 for the parton shower, and is scaled to NLO cross sections. The $t\bar{t}H$ cross section depends quadratically on κ_t; however, in contrast to the tHq and tHW samples, the scaling is not performed during the sample generation process but in the analysis code since it was decided to include the $t\bar{t}H$ process as part of the signal in the course of the analysis.

MC background samples

Several MC generators were used to generate the samples of the background processes. The dominant background sources ($t\bar{t}$, $t\bar{t}W$, $t\bar{t}Z$) were produced using aMC@NLO interfaced to PYTHIA8, and are scaled to NLO cross sections. Other minor background processes are simulated using POWHEG interfaced to PYTHIA, or bare PYTHIA as stated in the sample names in Table A.4. Pileup interactions are included in the simulation in order to reflect the observed multiplicity in data; the simulated events are weighted according to the actual pileup in data, estimated from the measured bunch-to-bunch instantaneous luminosity and the total inelastic cross section, 69.2 mb. All events are finally passed through a full simulation of the CMS detector based on GEANT4, and reconstructed using the same algorithms as used for the data.

6.5 Object Identification

In this section, the specific definitions of the physical objects in terms of the numerical values assigned to the reconstruction parameters are presented; thus, the provided details summarize and complement the descriptions presented in previous chapters. The object reconstruction and selection strategy used in this thesis is inherited from the analyses in References [145,149], thus, the information in this section is extracted from those documents unless other References are stated.

6.5.1 Lepton reconstruction and identification

Two types of leptons are defined in this analysis: *signal leptons* are those coming from W, Z and τ decays which usually are isolated from other particles; *background leptons* are defined as leptons produced in b-jet hadron decays, light-jets misidentification, and photon conversions.

The process of reconstruction and identification of electron and muon candidates was described in chapter3, hence, the identification variables used in order to retain the highest possible efficiency for signal leptons while maximizing the rejection of background leptons are listed and described in the following sections [2].

The identification variables include not only observables related directly to the reconstructed leptons themselves, but also to the clustered energy deposits and charged particles in a cone around the lepton direction (jet-related variables); an initial loose preselection of leptons candidates is performed and then an MVA discriminator, referred to as *lepton MVA* discriminator, is used to distinguish signal leptons from background leptons.

Muons

The Physics Objects Groups (POG) at CMS, are in charge of studying and defining the set of selection criteria applied on the course of reconstruction and identification of particles. These selection criteria are implemented in the CMS framework in the form of several object identification working points according to the strength of the requirements.

The muon candidates are reconstructed by combining information from the tracker system and the muon detection system of CMS detector and the POG defined three

[2] the studies performed to optimize the identification are far from the scope of this thesis, therefore, only general descriptions are provided

working points for muon identification *MuonID* [153];

- *POG Loose Muon ID* is a particle identified as a muon by the PF event reconstruction and also reconstructed either as a global-muon or as an arbitrated tracker-muon. This identification criteria is designed to be highly efficient for prompt muons and for muons from heavy and light quark decays; it can be complemented by applying impact parameter cuts in analyses with prompt muon signals.

- *POG Medium Muon ID* is a Loose muon with additional track-quality and muon-quality (spatial matching between the individual measurements in the tracker and the muon system) requirements. This identification criteria is designed to be highly efficient in the separation of the muons coming from decay in flight of heavy quarks and muons coming from B meson decays as well as prompt muons. An additional category *MVA Prompt ID* is defined in this identification criteria directed to discriminated muons from B mesons and prompt muons (from W,Z and τ decays). The Medium ID provides the same fake rate as the Tight Muon ID but a higher efficiency on prompt and B-decays muons. [154]

- *POG Tight Muon ID* is a global muon with additional muon-quality requirements Tight Muon ID selects a subset of the PF muons.

Only muons within the muon system acceptance $|\eta| < 2.4$ and minimum p_T of 5 GeV are considered.

Electrons

Electrons are reconstructed using information from the tracker and from the electromagnetic calorimeter and identified by an MVA algorithm (*MVA eID* discriminant) using the shape of the calorimetric shower variables like the shape in η and ϕ, the cluster circularity, widths along η and ϕ; track-cluster matching variables like E_{tot}/p_{in}, E_{Ele}/p_{out}, $\Delta\eta_{in}$, $\Delta\eta_{out}$, $\Delta\phi_{in}$, $1/E-1/p$; and track quality variables like χ^2 of the GSF tracks, the number of hits used by the GSF filter [155].

A loose selection based on η-dependent cuts on this discriminant is used to preselect electron candidates, the full shape of the discriminant is used in the lepton MVA selection to separate signal leptons from background leptons (described in Section 6.5.1).

In order to reject electrons from photon conversions, electron candidates with missing hits in the pixel tracker layers or matched to a conversion secondary vertex are discarded. Electrons are selected for the analysis if they have $p_T > 7$ GeV and are located within the tracker system acceptance region ($|\eta| < 2.5$).

Lepton vertexing and pile-up rejection

The impact parameter in the transverse plane d_0 , impact parameter along the z-axis d_z , and the impact parameter significance in the detector space SIP_{3D}, are considered to perform the identification and rejection of pile-up, misreconstructed tracks, and background leptons from b-hadron decays; pile-up and misreconstructed track mitigation is achieved by imposing loose cuts on the impact parameter variables. The full shape of the those variables is used in a lepton MVA classifier to achieve the best separation between the signal and the background leptons.

Lepton isolation

PF is able to recognize leptons from two different sources: on one side, leptons from the decays of heavy particles, such as W and Z bosons, which are normally isolated in space from the hadronic activity in the event; on the other side, leptons from the decays of hadrons and jets misidentified as leptons, which are not isolated as the former. For highly boosted systems, like the lepton and the b-jet generated in the semileptonic decay of a boosted top, the decay products tend to be more closer and sometimes they even overlap; thus, the PF standard definition of isolation in terms of the separation between the lepton candidates and other PF objects in the η-ϕ plane,

$$\Delta R = \sqrt{(\eta^l - \eta^i)^2 + (\phi^l - \phi^i)^2} < 0.3 \tag{6.1}$$

which considers all the neutral, charged hadrons and photons in a cone around the leptons, is refocused to the local isolation of the leptons through the mini-isolation I_{mini} [156] defined as the sum of particle flow candidates p_T within a cone around the lepton, corrected for the effects of pileup and divided by the lepton p_T

$$I_{mini} = \frac{\sum_R p_T(h^\pm) - \max\left(0, \sum_R p_T(h^0) + p_T(\gamma) - \rho\mathcal{A}\left(\frac{R}{0.3}\right)^2\right)}{p_T(l)} \tag{6.2}$$

where ρ is the pileup energy density, $h^\pm, h^0, \gamma, l$, represent the charged hadron, neutral hadrons, photons, and the lepton, respectively. The radius R of the cone depends on the p_T of the lepton according to

$$R = \frac{10\mathrm{GeV}}{\min(\max(p_T(l), 50\mathrm{GeV}), 200\mathrm{GeV})}, \tag{6.3}$$

The p_T dependence of the cone size allows for greater signal efficiency. Setting a cut on I_{mini} below a given threshold ensures that the lepton is locally isolated, even in boosted systems. The effect of pileup is mitigated using the so-called effective area correction $\mathcal{A}$ listed in Table6.3.

A loose cut on I_{mini} is applied to pre-select the muon and electron candidates;

$\|\eta\|$ range	$\mathcal{A}$(e) neutral/charged	A (μ) neutral/charged
0.0 - 0.8	0.1607 / 0.0188	0.1322 / 0.0191
0.8 - 1.3	0.1579 / 0.0188	0.1137 / 0.0170
1.3 - 2.0	0.1120 / 0.0135	0.0883 / 0.0146
2.0 - 2.2	0.1228 / 0.0135	0.0865 / 0.0111
2.2 - 2.5	0.2156 / 0.0105	0.1214 / 0.0091

Table 6.3: Effective areas, for electrons and muons used to mitigate the effect of pileup by using the so-called effective area correction.

however, the full shape is used in the lepton MVA discriminator when performing the signal lepton selection.

Jet-related variables

In order to reject misidentified leptons from $b-$jets, mostly coming from tt+jets, Drell-Yan+jets, and W+jets events, the vertexing and isolation described in previous sections are complemented with additional variables related to the closest reconstructed jet to the lepton, i.e., the PF jets reconstructed[3] around the leptons with $\Delta R = \sqrt{(\eta^l - \eta^{jet})^2 + (\phi^l - \phi^{jet})^2} < 0.5$. The identification variables used in the lepton MVA discriminator are the ratio p_T^l/p_T^{jet}, the CSV b-tagging discriminator value of the jet, the number of charged tracks of the jet, and the relative p_T given by

$$p_T^{rel} = \frac{(\vec{p}_{jet} - \vec{p}_l)\cdot\vec{p}_l}{||\vec{p}_{jet} - \vec{p}_l||}. \tag{6.4}$$

LeptonMVA discriminator

Electrons and muons passing the basic selection process described above are referred to as *loose leptons*. Additional discrimination between signal leptons and background leptons is crucial considering that the rate of $t\bar{t}$ production is much larger than the signal, hence, an overwhelming background from $t\bar{t}$ production. To maximally ex-

[3] charged hadrons from PU vertices are not removed prior to the jet clustering.

ploit the available information in each event to that end, the dedicated lepton MVA discriminator, based on a boosted decision tree (BDT) algorithm, has been built so that all the identification variables can be used together.

The lepton MVA discriminator training is performed using simulated signal Loose leptons from the $t\bar{t}H$ MC sample and fake leptons from the $t\bar{t}$ +jets MC sample, separately for muons and electrons. The input variables used include vertexing, isolation and jet-related variables, the p_T and η of the lepton, the electron MVA eID discriminator and the muon segment-compatibility variables. An additional requirement known as *tight-charge* requirement, is imposed by comparing two independent measurement of the charge, one from the ECAL supercluster and the other from the tracker; thus, the consistency in the measurements of the electron charge is ensured so that events with a wrong electron charge assignment are rejected; this variable is particularly used in the $2lss$ channel to suppress opposite-sign events for which the charge of one of the leptons has been mismeasured. The tight-charge requirement for muons is represented by the requirement of a consistently well measured track transverse momentum given by $\Delta p_T/p_T < 0.2$. Leptons are selected for the final analysis if they pass a given threshold of the BDT output, and are referred to as *tight leptons* in the following.

The validation of the lepton MVA algorithm and the lepton identification variables is performed using data in various control regions; the details about that validation are not discussed here but can be found in Reference [149].

Selection definitions

Electron and muon object identification is defined in three different sets of selections criteria; the *Loose*, *Fakeable Object*, and *Tight* selection. These three levels of selection are designed to serve for event level vetoes, the fake rate estimation application region (see Section 6.7.2), and the final signal selection, respectively. The p_T of fakeable objects is defined as $0.85 \times p_T(\text{jet})$, where the jet is the one associated to the lepton object. This mitigates the dependence of the fake rate on the momentum of the fakeable object and thereby improves the precision of the method.

Tables 6.4 and 6.5 list the full criteria for the different selections of muons and electrons.

Cut	Loose	Fakeable object	Tight
$\lvert\eta\rvert < 2.4$	✓	✓	✓
p_T	$> 5GeV$	$> 15GeV$	$> 15GeV$
$\lvert d_{xy}\rvert < 0.05$ (cm)	✓	✓	✓
$\lvert d_z\rvert < 0.1$ (cm)	✓	✓	✓
$\text{SIP}_{3D} < 8$	✓	✓	✓
$I_{\text{mini}} < 0.4$	✓	✓	✓
is Loose Muon	✓	✓	✓
jet CSV	–	< 0.8484	< 0.8484
is Medium Muon	–	–	✓
tight-charge	–	–	✓
lepMVA > 0.90	–	–	✓

Table 6.4: Requirements on each of the three muon selections. In the cases where the cut values change between the selections, those values are listed in the table. Otherwise, whether the cut is applied is indicated.

In addition to the previously defined requirements for jets, they are required to be separated from any lepton candidates passing the fakeable object selections by $\Delta\text{R} > 0.4$.

6.5.2 Lepton selection efficiency

Efficiencies of reconstruction and selecting loose leptons are measured both for muons and electrons using a tag and probe method on both data and MC, using $Z \to \ell^+\ell^-$ [157]. The scale factors are derived from the ratio of efficiencies $\varepsilon_i(p_T, \eta)$ measured

Cut	Loose	Fakeable Object	Tight
$\lvert\eta\rvert < 2.5$	✓	✓	✓
p_T	$> 7GeV$	$> 15GeV$	$> 15GeV$
$\lvert d_{xy}\rvert < 0.05$ (cm)	✓	✓	✓
$\lvert d_z\rvert < 0.1$ (cm)	✓	✓	✓
$\mathrm{SIP}_{3D} < 8$	✓	✓	✓
$I_{\text{mini}} < 0.4$	✓	✓	✓
MVA eID $> (0.0, 0.0, 0.7)$	✓	✓	✓
$\sigma_{i\eta i\eta} < (0.011, 0.011, 0.030)$	–	✓	✓
H/E $< (0.10, 0.10, 0.07)$	–	✓	✓
$\Delta\eta_{in} < (0.01, 0.01, 0.008)$	–	✓	✓
$\Delta\phi_{\text{in}} < (0.04, 0.04, 0.07)$	–	✓	✓
$-0.05 < 1/E - 1/p < (0.010, 0.010, 0.005)$	–	✓	✓
p_T^{ratio}	–	$> 0.5^{\dagger}$ / –	–
jet CSV	–	$< 0.3^{\dagger}$ / < 0.8484	< 0.8484
tight-charge	–	–	✓
conversion rejection	–	–	✓
Number of missing hits	< 2	$== 0$	$== 0$
lepton MVA > 0.90	–	–	✓

Table 6.5: Criteria for each of the three electron selections. In cases where the cut values change between selections, those values are listed in the table. Otherwise, whether the cut is applied is indicated. In some cases, the cut values change for different η ranges. These ranges are $0 < |\eta| < 0.8$, $0.8 < |\eta| < 1.479$, and $1.479 < |\eta| < 2.5$ and the respective cut values are given in the form (value$_1$, value$_2$, value$_3$). For the two p_T^{ratio} and CSV rows, the cuts marked with a † are applied to leptons that fail the lepton MVA cut, while the loose cut value is applied to those that pass the lepton MVA cut.

for a given lepton in data/MC, according to

$$\rho(p_T, \eta) = \frac{\varepsilon_{data}(p_T, \eta)}{\varepsilon_{MC}(p_T, \eta)}. \tag{6.5}$$

The scale factor for each event is used to correct the weight of the event in the full sample; therefore, the full simulation correction is given by the product of all the individual scale factors. The scale factors used in this thesis are inherited from Reference [149] which in turns inherited them from leptonic SUSY analyses using equivalent lepton selections.

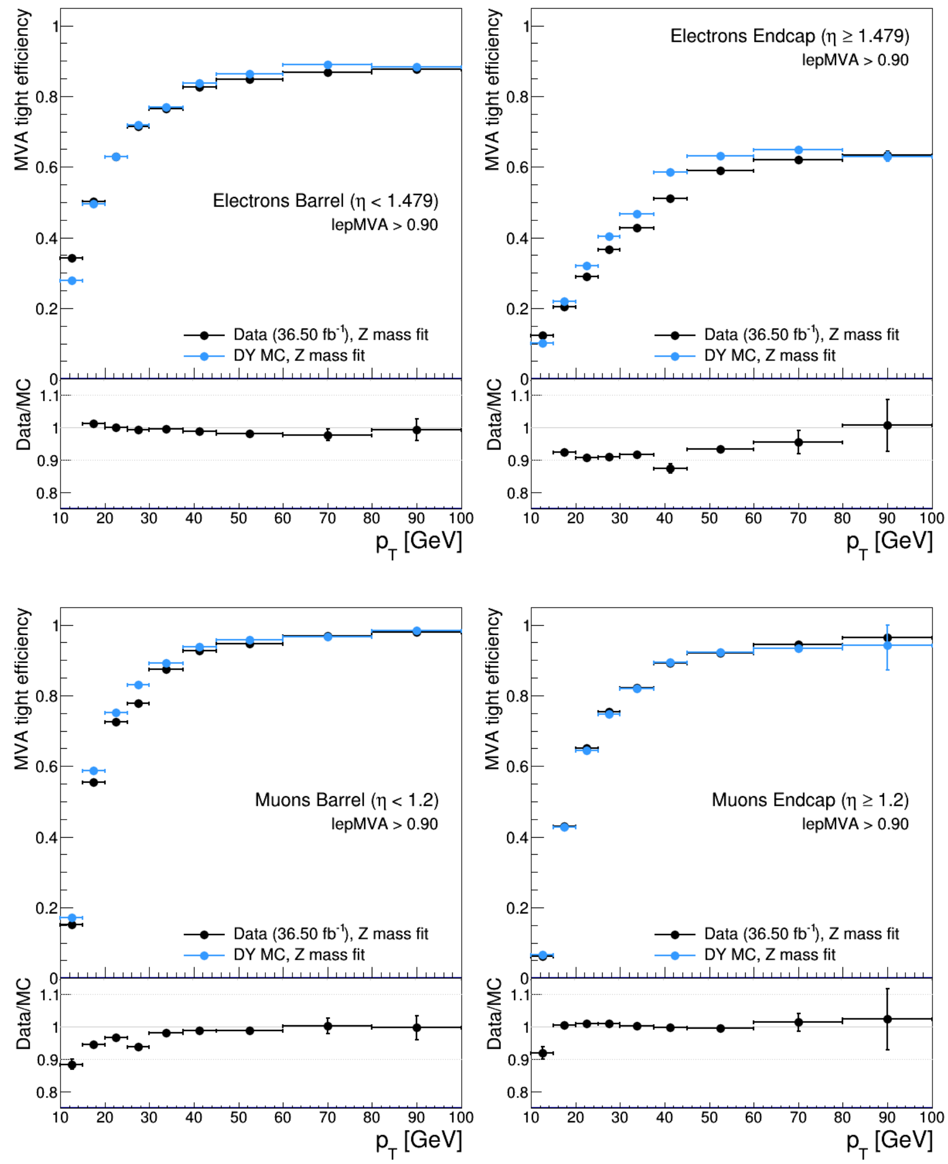

Figure 6.8: Tight vs loose selection efficiencies for electrons (top), and muons (bottom), for the $2lss$ definition, i.e., including the tight-charge requirement.

The efficiency of applying the tight selection as defined in Tables 6.4 and 6.5, on the loose leptons are determined by using a tag and probe method on a sample of Drell-Yan enriched events. Figures 6.8 and 6.9 show the efficiencies for the $2lss$ channel and $3l$ channel respectively. Efficiencies in the $2lss$ channel have been produced including the tight-charge requirement, while for the $3l$ channel it is not included. Number of passed and failed probes are determined from a fit to the invariant mass of the dilepton system. Simulation is corrected using these scale factors; note that they depends on η and p_T.

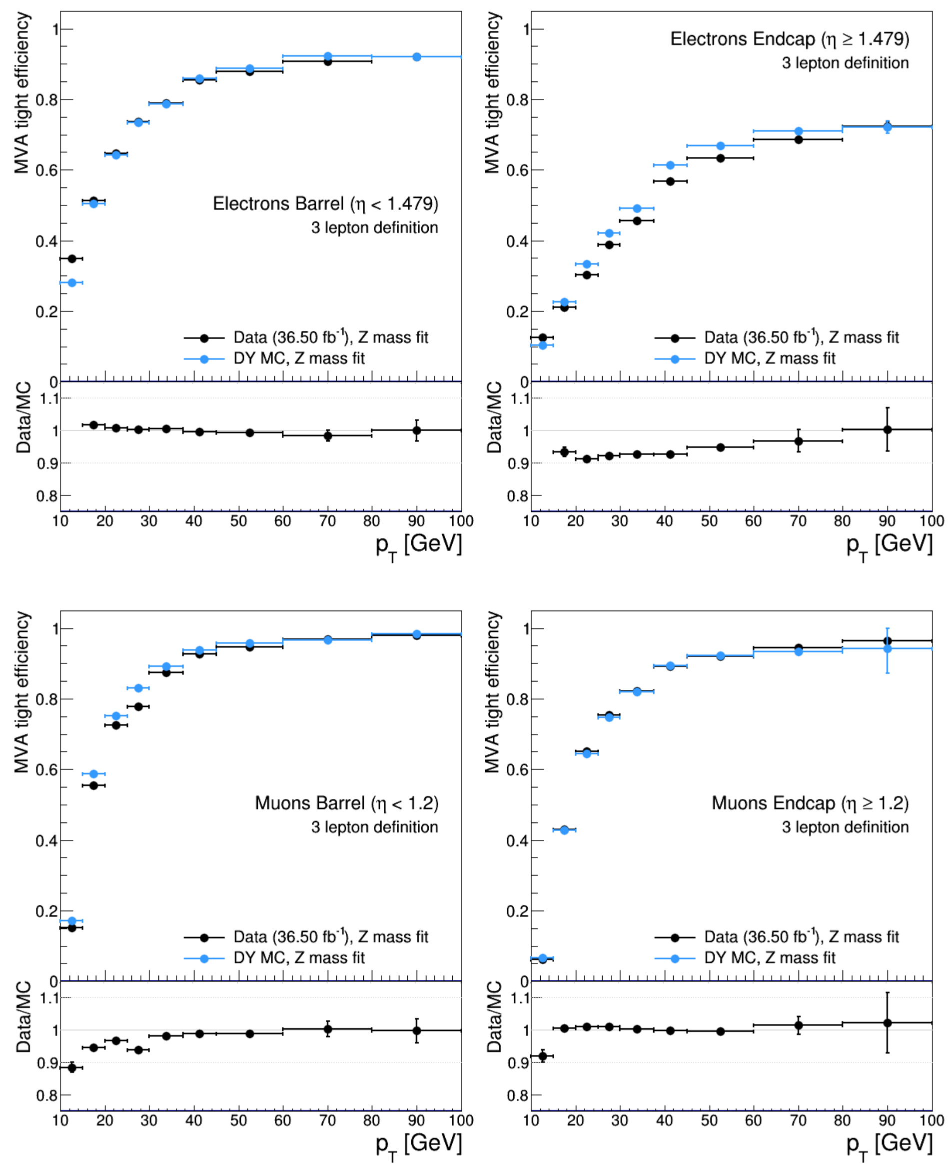

Figure 6.9: Tight vs loose selection efficiencies for electrons (top), and muons (bottom), for the $3l$ channel not including the tight-charge requirement.

6.5.3 Jets and b-jet tagging

In this analysis, jets are reconstructed by clustering PF candidates using the anti-k_t algorithm with parameter distance $\Delta R = 0.4$; those charged hadrons that are not consistent with the selected primary vertex are discarded from the clustering. The jet energy is then corrected for the varying response of the detector as a function of transverse momentum p_T and pseudorapidity η. Jets are selected for use in the analysis only if they have $p_T > 25$ GeV and are separated from any selected leptons by $\Delta R > 0.4$.

Jets coming from the primary vertex and jets coming from pile-up vertices are distinguished using a MVA discriminator based on the differences in the jet shapes, in the relative multiplicity of charged and neutral components, and in the different fraction of transverse momentum which is carried by the hardest components. Jet tracks are also required to be compatible with the primary vertex.

Jets originated from the hadronization of a b quark are selected using a MVA likelihood discriminant which uses track-based lifetime information and reconstructed secondary vertices (CSV algorithm). Only jets within the CMS tracker acceptance ($\eta < 2.4$) are identified with this tool. Data samples are used to measure the efficiency of the b-jet tagging and the probability to misidentify jets from light quarks or gluons; in both cases the measurements are parametrized as a function of the jet p_T and η and later used to correct differences between the data and MC simulation in the b tagging performance, by applying per-jet weights to the simulation, dependent on the jet p_T, η, b tagging discriminator, and flavor (from simulation truth) [151]. The per-event weight is taken as the product of the per-jet weights, including those of the jets associated to the leptons. The weights are derived on $t\bar{t}$ and Z+jets events.

Two working points are defined, based on the CSV algorithm output: *loose'* working point (CSV>0.46) with a b signal tagging efficiency of about 83% and a mistagging rate of about 8%; and *medium* working point (CSV>0.80) with $b-$tagging efficiency of about 69% and mistagging rate of order 1% [152]. Tagging of jets from charm quarks have efficiencies of about 40% and 18% for loose and medium working points respectively. Separate scale factors are applied to jets originating from bottom/charm quarks and from light quarks in simulated events to match the tagging efficiencies measured in the data.

6.5.4 Missing Energy MET

As stated in Section 3.4.1.1, the MET vector is calculated as the negative of the vector sum of transverse momenta of all PF candidates in the event and its magnitude is referred to as E_T^{miss}. Due to pile-up interactions, the performance in determining MET is degraded; in order to correct for that, the energy from the selected jets and leptons that compose the event is assigned to the variable H_T^{miss}. It is calculated in the same way as E_T^{miss} and although it has worse resolution than E_T^{miss}, it is more robust in the sense that it does not rely on the soft part of the event. The event selection uses a linear discriminator based on the two variables given by

$$E_T^{miss}{}_{LD} = 0.00397 * E_T^{miss} + 0.00265 * H_T^{miss} \tag{6.6}$$

taking advantage of the fact that the correlation between E_T^{miss} and H_T^{miss} is less for events with instrumental missing energy than for events with real missing energy. The working point $E_T^{miss}{}_{LD} > 0.2$ was chosen to ensure a good signal efficiency while keeping a good background rejection.

6.6 Event selection

Events are selected considering the features of the signal process and the decay signature as described in Section 6.2. At the trigger level, events are selected to contain either one, two, or three leptons with minimal p_T thresholds:

- single-lepton trigger $\rightarrow$ 24 GeV for muons and at 27 GeV for electrons
- double-lepton triggers $\rightarrow$ leading and sub-leading leptons: 17 and 8 GeV for muons and 23 and 12 GeV for electrons.
- three-lepton triggers $\rightarrow$ threshold on the third hardest lepton in the event: 5 and 9 GeV for muons and electrons, respectively.

The offline event selection level targets the specific topology of the tHq signal with $H \rightarrow WW$ and $t \rightarrow Wb \rightarrow l\nu b$; therefore, the resulting state is composed of three W bosons, one b quark, and a light spectator quark at high rapidity. The selection criteria for the two channels exploited in this analysis are summarized in Table 6.6. This selection includes contributions from $H \rightarrow \tau\tau$ and $H \rightarrow ZZ$ as well.

Same-sign $\ell\ell$ channel	**$\ell\ell\ell$ channel**
have fired one of the corresponding trigger paths	
No loose leptons with $m_{\ell\ell} < 12$GeV	
One or more b tagged jets (CSV medium) $\lvert\eta\rvert$<2.4	
One or more non-tagged jets: central $\rightarrow p_T$ >25 GeV, η <2.4; forward $\rightarrow p_T$ >40 GeV, η >2.4	
${E_T^{miss}}_{LD} > 0.2$	
Exactly two tight same-sign leptons	Exactly three tight leptons
Lepton $p_T > 25/15$GeV	Lepton $p_T > 25/15/15$GeV
Electrons are triple-charge consistent.	No OSSF lepton pair with $\lvert m_{\ell\ell} - m_Z\rvert < 15$GeV
Muon p_T resolution: $\Delta p_T/p_T < 0.2$.	
No ee pair with $\lvert m_{ee} - m_Z\rvert < 10$GeV	

Table 6.6: Summary of event pre-selection.

In the $2lss$ channel, events with additional tight leptons are vetoed as well as those for which a loose lepton pair has an invariant mass below 12 GeV. A threshold in p_T of the leading and sub-leading leptons is also required. Events where the two electrons have invariant mass within 10 GeV of the Z boson mass (*Z-veto*) are discarded in order to reject events from DY+jets production with charge misidentified electrons. In addition, contribution from the associated production of two W bosons of equal charge and two light jets $W^{\pm}W^{\pm}qq$ and from same-sign W boson pairs can also be produced in double parton scattering (DPS) processes, where each of the colliding protons gives two partons, resulting in two hard interactions.

In the $3l$ lepton channel, leptons are required to have respectively $p_T > 25$GeV, $>$ 15 GeV, and > 15 GeV. Events with an opposite-sign same-flavor lepton combination (OSSF) with invariant mass within 15 GeV of the Z boson mass are discarded in order to reject events from WZ +jets production.

6.7 Background modeling and predictions

The dominant background contribution is expected to arise from top quark production processes, either $t\bar{t}$ pair production or in $t\bar{t}$ associated production with a W/Z. Processes with production of single top quarks also contribute, mainly in the associated production with a Z boson (tZq) or when produced with both a W and a Z boson (tZW). Background contamination from diboson processes is strongly suppressed by imposing the Z-veto, vetoing additional leptons and requiring b-jets in the event.

The selection criteria in Table 6.6 represent a relatively loose selection that allows to maintain a large signal efficiency while suppressing the main backgrounds; thus that selection is called *pre-selection*. The events obtained from the pre-selection are then used to extract the signal contribution in a second analysis step, using BDT discriminators against the main backgrounds of $t\bar{t}W/t\bar{t}Z$ and non-prompt leptons from $t\bar{t}$. The shape of the discriminator variables is then fit to the observed data distribution to estimate the signal and background yields, simultaneously for all channels.

Irreducible backgrounds are reliantly estimated from MC simulated events; therefore, in this analysis all backgrounds involving prompt leptons are estimated in this way. Reducible backgrounds, like non-prompt lepton backgrounds, are not well predicted by simulation, hence, they are estimated using data-driven methods.

6.7.1 $t\bar{t}$V and diboson backgrounds

Backgrounds from $t\bar{t}W$and $t\bar{t}Z$ processes are estimated using simulated events, corrected for data/MC differences and inefficiencies (trigger and lepton selection) in the same way as signal events. Their production cross sections are calculated at NLO order of QCD and EWK, considering theoretical uncertainties from unknown higher orders of 12% for $t\bar{t}W$and 10% for $t\bar{t}Z$. Additional uncertainties arise from the knowledge of PDFs and α_s of about 4% each for $t\bar{t}W$and $t\bar{t}Z$.

The diboson contribution is also estimated from simulated events; however, the overall normalization of this process is obtained from a dedicated control region. The motivation behind that strategy is that even thought the measured inclusive cross section for diboson processes (WZ,ZZ) is in good agreement with the NLO calculations [149], that agreement is perturbed when leptonic Z decays and hadronic jets in the final state are required; those requirements are precisely the ones that make the diboson production a background for the tHq signal. Thus, by using a dedicated control region dominated by WZ production[4], the overall normalization is constrained.

[4] ZZ background is strongly reduced by the cut on MET.

The control region is defined by the presence of at least three leptons, of which one opposite-sign pair must be compatible with a Z boson decay, i.e., invert the Z-veto which makes the control region orthogonal to signal region; the b-jet tagging requirements is also inverted with respect to the signal region, i.e., require two not b-jets. A scale factor is extracted from the predicted distribution of WZ events in the control region, and the observed data, while keeping other processes fixed; this factor is used to scale the diboson prediction in the signal selection region. More details about the procedure used can be found in Reference [149] from where the scale factor is taken.

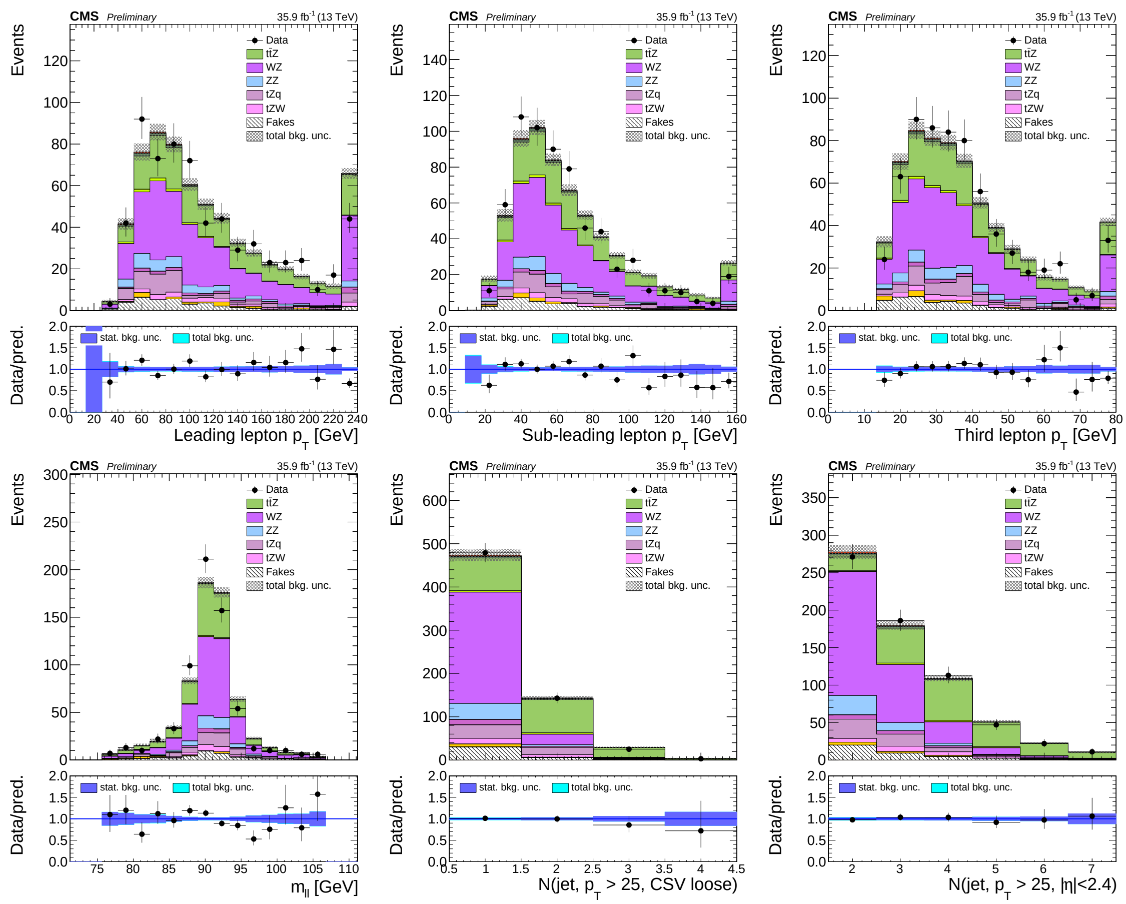

Figure 6.10: Kinematic distributions in the diboson control region.

In order to test the usability of the diboson background scale factor in this analysis, a Z-enriched control region[5] was defined by inverting the Z-veto and requiring exactly three tight leptons with $p_T > 25/15/15$ GeV, one or more jets passing the CSVv2 loose working point and less than four central jets. Figure 6.10 shows the distribution of three variables in the diboson control region; the good agreement between MC and data motivates the adoption of the diboson background scale factor.

Most of the diboson events passing the signal selection contain jets from light quarks and gluons that are incorrectly tagged as b-jets; it makes the estimate mainly sensitive to the experimental uncertainty in the mis-tag rate rather than the theoretical uncertainty in the jet flavor composition. The overall uncertainty assigned to the diboson prediction is estimated from the statistical uncertainty due to the limited sample size in the control region (30%), the residual background in the control region (20%), the uncertainties on the b-tagging rate (10-40%), and from the knowledge of PDFs and the theoretical uncertainties of the extrapolation (up to 10%).

6.7.2 Non-prompt and charge mis-ID backgrounds

The non-prompt lepton background contribution to the final selection is estimated using the fake factor method. The main idea of the method is to define a control region of events enriched in the background to estimate and determine a factor that relates (extrapolates) these events to those in the signal region. The method is data-driven in the sense that the control sample is selected from data, and the extrapolation factor is measured from data.

In the signal region of this analysis, non-prompt leptons are predominantly produced in $t\bar{t}$ events, with a much smaller contribution, from Drell-Yan production; therefore, the control region also know as *application region*, is defined by modifying the event selection criteria in such a way that most of the events after selection are

[5] This control region is different to the one used to find the scale factor.

$t\bar{t}$ events and thus the misidentification rate is increased. The application regions for electrons and muons are defined by the fakeable object definitions in Tables 6.4 and 6.5. Since the fakeable definition is a loosened version of the tight definition, in the context of fake rates the fakeable definition it becomes the loose selection.

The ratio between the number of events that pass both, the loose and tight selections, and the number of events that pass the loose selection but fail the tight one, corresponds to the *fake factor/fake rate (f)*. The measurement of the fake factor is made using two background dominated data samples, collected with dedicated triggers, as a function of p_T and $|\eta\,|$ and separately for muons and electrons:

- A sample dominated by QCD multijet events, collected using single lepton triggers at relatively high p_T thresholds. It is used to extract ratios for lepton candidates with p_T above 30 GeV.

- A sample dominated by Z + jets events, where the two high p_T leptons resulting from the Z decay are used to trigger the events without biasing the p_T spectrum of a third lepton at low transverse momentum. It is used to determine the ratios for low p_T leptons.

Processes like W + jets, Z + jets , WZ and ZZ produce prompt leptons that contaminate the samples; thus, they are suppressed by vetoing additional leptons in the selection, and the residual contamination is then subtracted using the transverse mass as a discriminating variable.

The extrapolation from the application region to the signal region is performed by weighting the events in the application region using the fake factor according to the following rules:

- events with one lepton failing the tight criteria are weighted with the factor $\frac{f}{(1-f)}$ for the estimate to the signal region.
- events with two leptons (i,j) failing the tight criteria are weighted with the factor $-\frac{f_i f_j}{(1-f_i)(1-f_j)}$ for the estimate to the signal region.
- events with three leptons (i,j,k) failing the tight criteria are weighted with the factor $\frac{f_i f_j f_k}{(1-f_i)(1-f_j)(1-f_k)}$ for the estimate to the signal region.

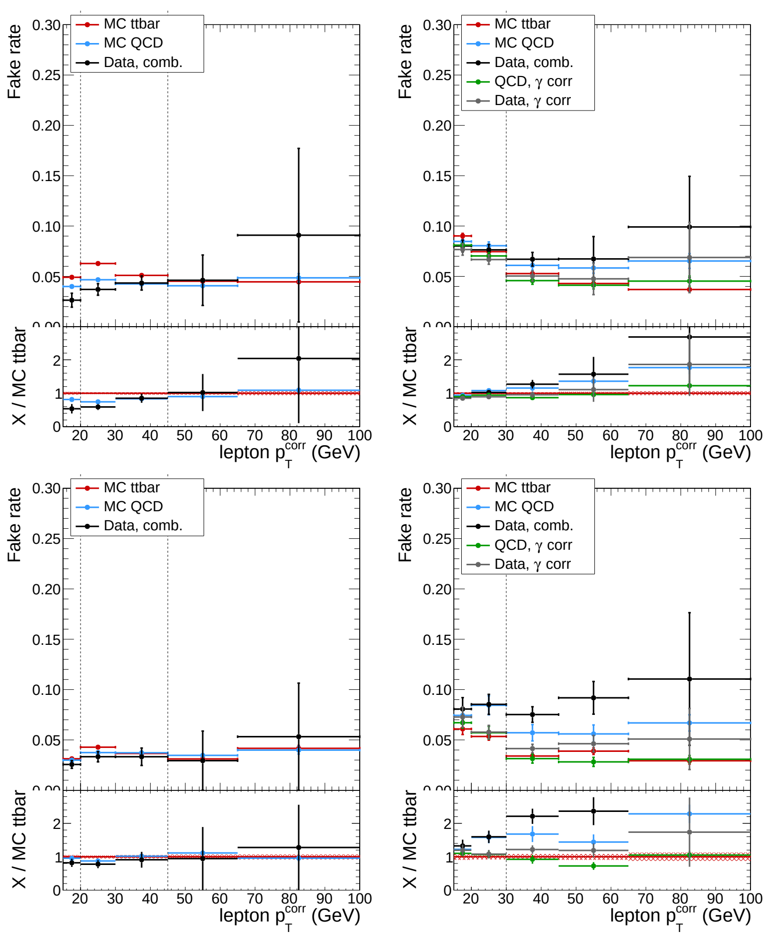

Figure 6.11: Fake rate measurement in events in data for muons (left column) and electrons (right column). Predictions from simulated events in the measurement region (blue) and from non-prompt leptons in $t\bar{t}$ (red) are included for comparison. Top row is for $|\eta|<2.5$ and bottom row for $|\eta|>2.5$.

Figure 6.11 shows the fake rates for electrons and muons used in this analysis which were taken from the studies in Reference [149].

The resulting prediction of the event yield in the signal selection carries an uncertainty of 30-50% which is composed of the statistical uncertainty in the measurement of the fake rates due to prescaling of lepton triggers, the uncertainty in the subtraction of residual prompt leptons from the control region, and from testing the closure of the method in simulated background events; hence, it is one of the dominant limitations on the performance of multilepton analyses in general and this analysis in particular.

Finally, an additional source of background arises in the $2lss$ channel from events with an originally opposite-sign lepton pair for which the charge of one of the leptons is misidentified (*charge mis-ID*); usually this happens because of the convertion of hard bremsstrahlung photons emitted from the initial lepton, hence, it is more likely to happen for electrons than for muons.

The charge mis-ID background is estimated from the yield of opposite-sign event in the signal region by measuring the charge mis-ID probability in same-sign and opposite-sign events compatible with a Z boson decay, in several bins of p_T and η, and weighting events with opposite-sign leptons in the signal selection.

Data	$10 \leq p_T < 25$ GeV	$25 \leq p_T < 50$ GeV	50 GeV$\leq p_T$
$0 \leq \eta < 1.48$	0.0442 ± 0.0011	0.0179 ± 0.0004	0.0262 ± 0.0020
$1.48 \leq \eta < 2.5$	0.1329 ± 0.0066	0.1898 ± 0.0014	0.3067 ± 0.0113
MC			
$0 \leq \eta < 1.48$	0.0378 ± 0.0016	0.0222 ± 0.0003	0.0233 ± 0.0015
$1.48 \leq \eta < 2.5$	0.0956 ± 0.0044	0.2108 ± 0.0027	0.3157 ± 0.0018

Table 6.7: Electron charge mis-ID probabilities (in percent), determined in data (top) and Drell-Yan MC (bottom) [149].

The charge mis-ID probability is found to be negligible for this analysis for muons, whereas for electrons it ranges from about 0.02% in the barrel section ($|\eta\ | < 1.48$)

up to about 0.35% in the detector endcaps ($1.48 < |\eta| < 2.5$). as shown in Table 6.7 and Figure 6.12.

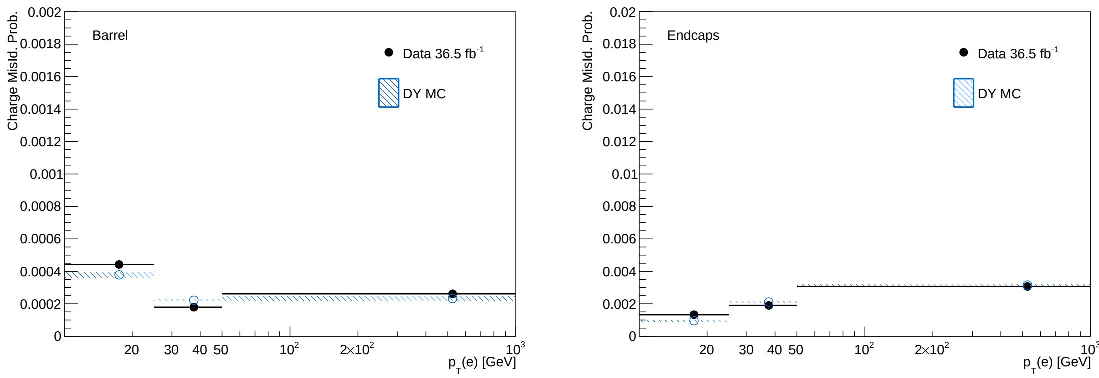

Figure 6.12: Electron charge mis-ID probabilities as a function of p_T for $|\eta|$<2.5 (left) and $|\eta|$<2.5 (right) [149].

The contribution from charge mis-ID electrons in signal selection of this analysis comes mainly from $t\bar{t}$ and Drell-Yan events. The systematic uncertainty of the normalization of the charge mis-id. estimate is evaluated at about 30%, arising from a slight disagreement of the mis-ID. probability between data and simulation. Given that it only affects the $e\mu$ channel, it impact on the final sensitivity is very limited.

6.8 Pre-selection yields

The expected and observed event yields of the pre-selection are shown in Table 6.8. For the tH and $t\bar{t}H$ processes, the largest contribution comes from Higgs decays to WW (about 75%), followed by $\tau\tau$ (about 20%) and ZZ (about 5%). Other Higgs production modes contribute negligible event yields (< 5% of the tH+$t\bar{t}H$ yield) as shown in Table 6.9.

Figure 6.13 shows the distributions of some relevant kinematic variables, normalized to the cross section of the respective processes and to the integrated luminosity.

	3ℓ	$\mu^{\pm}\mu^{\pm}$	$e^{\pm}\mu^{\pm}$	$e^{\pm}e^{\pm}$
$t\bar{t}W$	22.50 ± 0.35	68.03 ± 0.61	97.00 ± 0.71	29.63 ± 0.39
$t\bar{t}Z/\gamma^{*}$	32.80 ± 1.79	25.89 ± 1.12	64.82 ± 2.42	28.74 ± 1.70
WZ	8.22 ± 0.86	15.07 ± 1.19	26.25 ± 1.57	9.31 ± 0.93
ZZ	1.62 ± 0.33	1.16 ± 0.29	2.86 ± 0.45	1.09 ± 0.27
$W^{\pm}W^{\pm}qq$	–	3.96 ± 0.52	6.99 ± 0.69	2.19 ± 0.37
$W^{\pm}W^{\pm}$(DPS)	–	2.48 ± 0.42	4.17 ± 0.54	0.81 ± 0.24
VVV	0.42 ± 0.16	2.99 ± 0.34	4.85 ± 0.43	1.19 ± 0.21
tttt	1.84 ± 0.44	2.32 ± 0.45	4.06 ± 0.57	0.89 ± 0.31
tZq	3.92 ± 1.48	5.77 ± 2.24	10.73 ± 3.03	7.56 ± 1.72
tZW	1.70 ± 0.12	2.13 ± 0.13	3.91 ± 0.18	1.13 ± 0.10
γ conversions	7.43 ± 1.94	–	23.81 ± 6.04	9.87 ± 4.17
Non-prompt	25.61 ± 1.26	80.94 ± 2.02	135.34 ± 2.83	47.72 ± 1.79
Charge mis-ID	–	–	58.50 ± 0.31	44.52 ± 0.31
All backgrounds	106.05 ± 3.45	210.74 ± 3.61	443.30 ± 8.01	184.65 ± 5.29
tHq ($\kappa_t = -1.0$)	7.48 ± 0.14	18.48 ± 0.22	27.41 ± 0.27	8.47 ± 0.15
tHW ($\kappa_{\mathrm{V}} = -1.0$)	7.38 ± 0.16	7.72 ± 0.17	11.23 ± 0.20	3.66 ± 0.11
$t\bar{t}H$	18.29 ± 0.41	24.18 ± 0.48	35.21 ± 0.58	11.07 ± 0.32
Data ($35.9 fb^{-1}$)	127	280	525	208

Table 6.8: Expected and observed yields for $35.9 fb^{-1}$ after the pre-selection in all final states. Uncertainties are statistical only.

The remaining variables distributions are shown in Appendix B.3

A significant fraction of selected data events (about 50% in the dilepton channels, and about 80% in the trilepton channel) also passes the selection used in the dedicated search for ttH in multilepton channels [149]. This is particularly important when considering a possible combination of the measurements from both studies. More details about the overlap between this both analyses are presented in Appendix E.

6.9 Signal discrimination

The production cross section for the signal processes tHq, tHW, and $t\bar{t}H$ is only about 600 fb (the enhancement provided by inverted couplings, κ_t = -1 almost double it), resulting in a small signal to background ratio even for a tight selection. A

	3ℓ		$\mu^{\pm}\mu^{\pm}$	
tHq(Inclusive)	**6.57**	100.0%	**17.38**	100.0%
$tHq(H \to WW)$	4.84	73.9%	13.33	76.9%
$tHq(H \to \tau\tau)$	1.04	15.9%	3.62	20.6%
$tHq(H \to ZZ)$	0.48	7.2%	0.37	2.2%
$tHq(H \to \mu\mu)$	0.21	3.0%	0.04	0.2%
$tHq(H \to \gamma\gamma)$	< 0.01	0.1%	0.02	0.1%
$tHq(H \to bb)$	< 0.01	$< 0.1\%$	0.01	$< 0.1\%$
tHW(Inclusive)	**7.32**	100.0%	**7.62**	100.0%
$tHW(H \to WW)$	5.50	76.9%	5.60	74.1%
$tHW(H \to \tau\tau)$	1.40	20.6%	1.81	23.1%
$tHW(H \to ZZ)$	0.31	2.2%	0.21	2.7%
$tHW(H \to \mu\mu)$	0.12	0.2%	0.01	0.1%
$tHW(H \to \gamma\gamma)$	< 0.01	$< 0.1\%$	< 0.01	$< 0.1\%$
$tHW(H \to bb)$	< 0.01	$< 0.1\%$	< 0.01	$< 0.1\%$

Table 6.9: Signal yields split by decay channels of the Higgs boson. Forward jet p_T cut at 25 GeV.

multivariate method is hence employed to train discriminators to separate tH signal events from the dominant background events.

6.9.1 MVA classifiers evaluation

Several MVA classifier algorithms were evaluated in order to determine the most appropriate method for this analysis[6]. The comparison is based on the performance of the classifiers, encoded in the plot of the background rejection as a function of the signal efficiency (ROC curve). The top row of Figure 6.14 shows the ROC curves for the several methods evaluated; two separated training were performed in the $3l$ channel: against $t\bar{t}$ (right) and $t\bar{t}$V (left) processes.

In both cases, the gradient boosted decision tree *BDTG* (BDTA_GRAD in the plot) classifier offers the best results, followed by the adaptive BDT classifier (*BDTA*);

[6] The choice of the tested algorithms was based on the recommendations provided by the official TMVA user guide, the experience from previous analyses and considering the expertise of the members of the tHq and $t\bar{t}H$ analyses groups. Only the BDT classifier is described in this thesis and a detailed description of all available methods can be found in Reference [127]

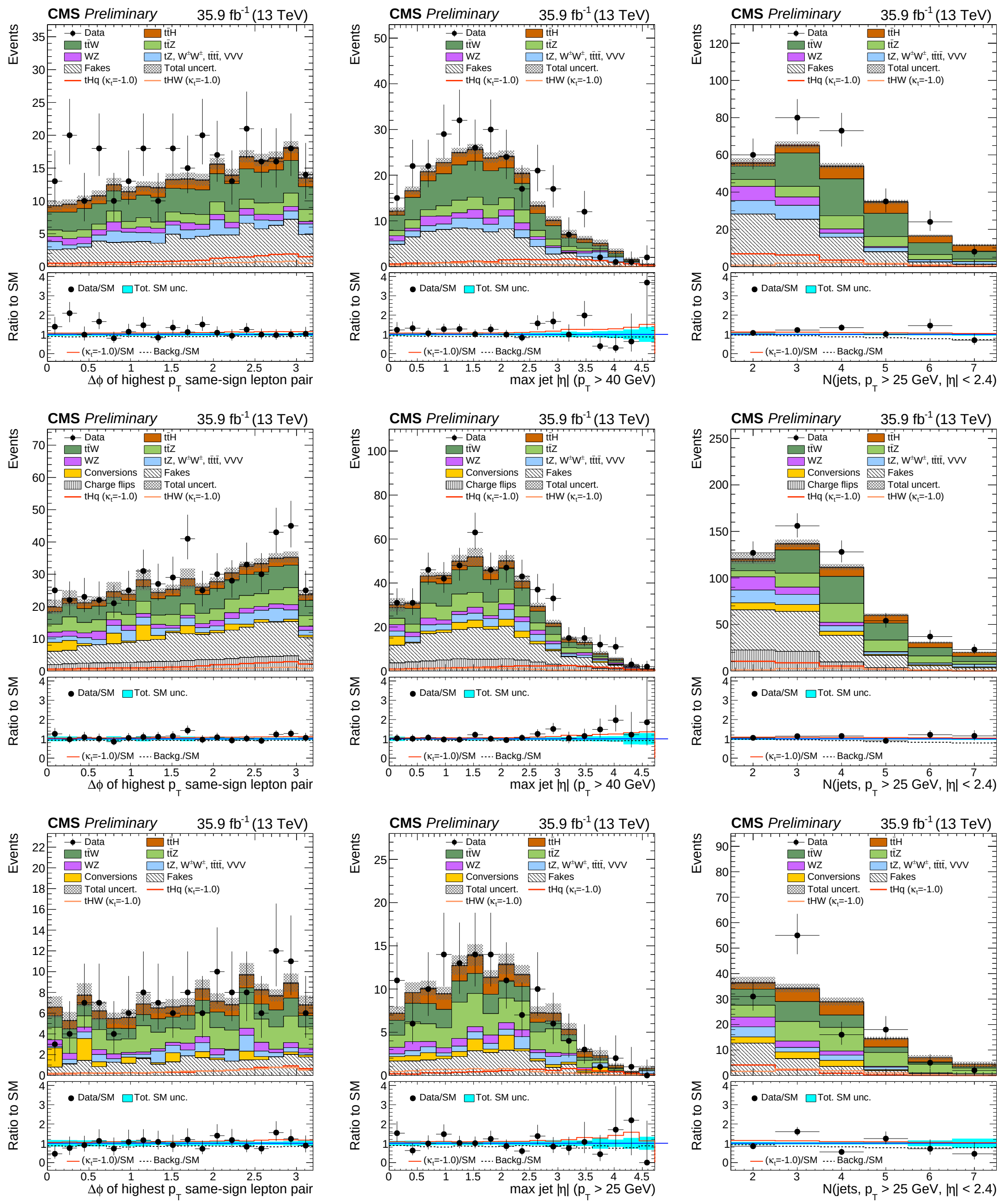

Figure 6.13: Distributions of discriminating variables for the event pre-selection for the same-sign $\mu^{\pm}\mu^{\pm}$ channel (top row), the same-sign $e^{\pm}\mu^{\pm}$ channel (middle row) and three lepton channel (bottom row), normalized to $35.9 fb^{-1}$, before fitting the signal discriminant to the observed data. Uncertainties are statistical and unconstrained (pre-fit) normalization systematics. The shape of the two tH signals for $\kappa_t = -1.0$ is shown, normalized to their respective cross sections for $\kappa_t = -1.0, \kappa_{\mathrm{V}} = 1.0$.

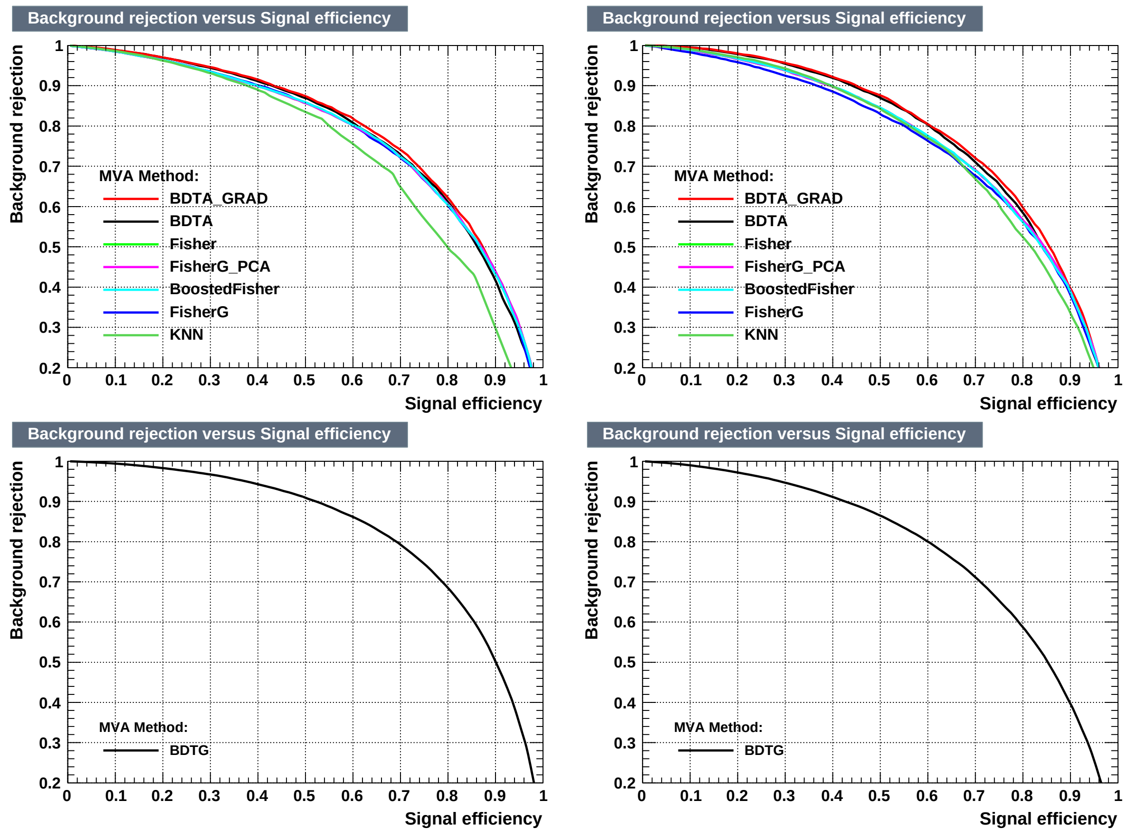

Figure 6.14: Top: Background rejection vs signal efficiency (ROC curves) for various MVA classifiers in the $3l$ channel for training against $t\bar{t}$V (left) and $t\bar{t}$ (right). Bottom: background rejection vs signal efficiency (ROC curve) in the $2lss$ channel for a single discriminator: BDTG, against $t\bar{t}$V (left) and $t\bar{t}$ (right).

the several Fisher classifiers tested, which differ in their parameters and/or boosting method, they offer similar performance among them, while the k-Nearest Neighbour (kNN) classifier performance is below the rest of the classifiers. The corresponding ROC curves and in the $2lss$ channel for trainings against $t\bar{t}$V(left) and $t\bar{t}$ (right) processes are shown in the bottom row of Figure 6.14; the BDTG performance is similar to that in the $3l$ channel.

6.9.2 Discriminating variables

The classifier chosen to separate the tHq signal from the main backgrounds is the BDTG classifier, trained on simulated signal and background events. The samples used in the training are the tHq sample in Table 6.2, the samples in the third section of table A.4 and the samples marked with an * in the same table.

As explained in Section 5.1.1, a set of discriminating variables are given as input to the BDTG which combines the individual discrimination power of each input variable to produce a discriminator with the maximum discrimination power. Table 6.10 lists the input variables used in the BDTG trainings for this analysis. The same set of input variables was used to produce the plots for MVA classifiers evaluation.

Variable name	Description
nJet25	Number of jets with $p_T > 25$ GeV, $\|\eta\| < 2.4$
nJetEta1	Number of jets with $\|\eta\| > 1.0$, non-CSV-loose
MaxEtaJet25	Max. $\|\eta\|$ of any (non-CSV-loose) jet with $p_T > 25$ GeV
detaFwdJetClosestLep	$\Delta\eta$ forward light jet and closest lepton
detaFwdJetBJet	$\Delta\eta$ forward light jet and hardest CSV loose jet
detaFwdJet2BJet	$\Delta\eta$ forward light jet and second hardest CSV loose jet
Lep3Pt/Lep2Pt	p_T of the 3^{rd} lepton (2^{nd} for ss2l)
totCharge	Sum of lepton charges
minDRll	Min ΔR any two leptons
dphiHighestPtSSPair	$\Delta\phi$ of highest p_T same-sign lepton pair

Table 6.10: BDTG input variables. First section lists variables related to jet multiplicities; second section lists variables related to forward jet activity, and third section lists variables related to lepton kinematics.

Plots in Figure 6.15 shows the BDTG input variables distributions for the signal and background samples, in the $3l$ channels.

All the input variables have some discrimination power, however, that power is bigger for some of them; for instance, the third lepton p_T plot (top left in Figure 6.15) shows some discrimination power against WZ and VVV backgrounds for which there is a peak around 30 GeV while tHq peak around 18 GeV; although the discrimination

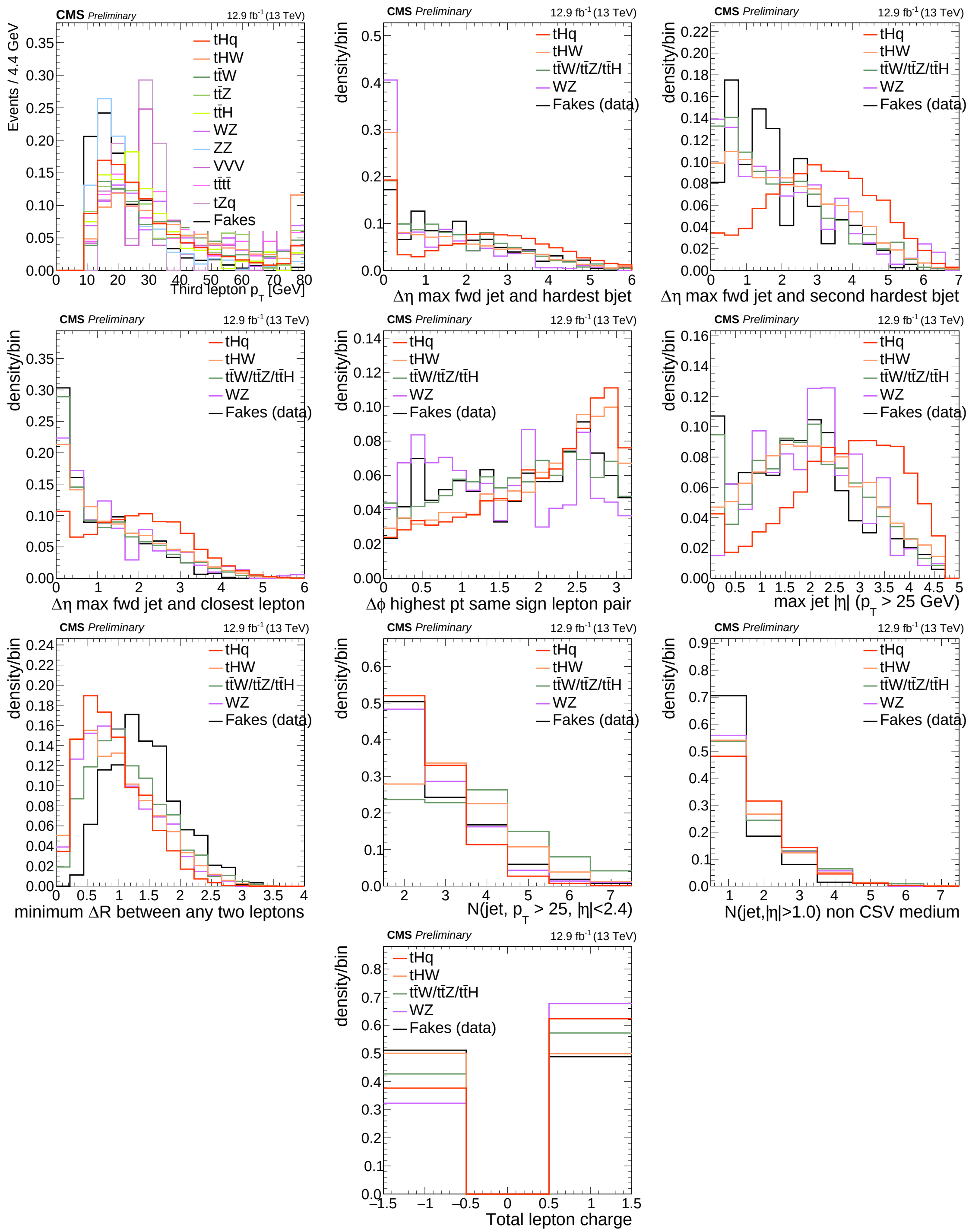

Figure 6.15: Distributions of the BDTG classifier input variables (not normalized) for signal discrimination in the $3l$ channel.

power does not cover all the backgrounds, it counts for the final discriminator. A similar situation can be seen in the plot for the number of jets (row three, column two); $t\bar{t}W$, $t\bar{t}Z$and $t\bar{t}H$ processes tend to have more jets compared to the tHq process. The discrimination power is more evident in other plots like in the plot of the maximum $|\eta|$ of the jets in the event (row two, column three). The same or equivalent input variables are found to be performing well for both $3l$ and $2lss$ channels. Figure B.1 shows the corresponding input variables distribution plots for the $2lss$ channel.

Discrimination power from BDTG classifier

The Discrimination power of the input variables can also be evaluated from the BDTG training, exclusively for the training samples, i.e., dominant backgrounds ($t\bar{t}$ and $t\bar{t}$V); the training samples are submitted to the selection cuts on Table 6.6.

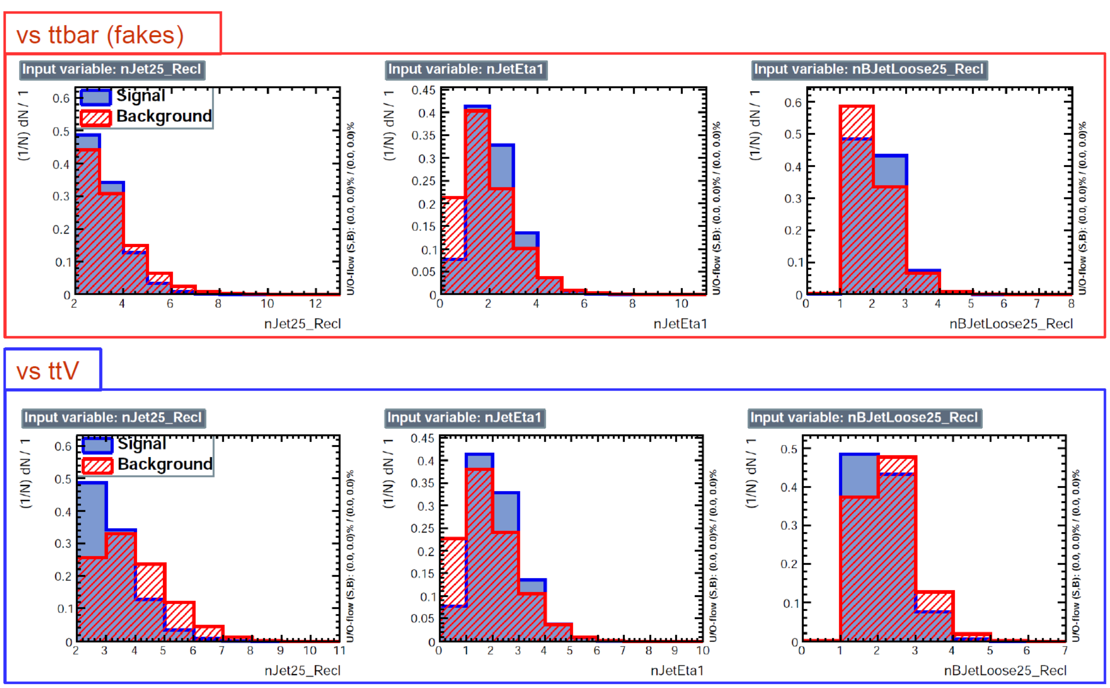

Figure 6.16: BDT input variables as seen by BDTG classifier for the $3l$ channel, tHq signal(blue) discriminated against $t\bar{t}$V background (red).

Figure 6.16 shows the comparison between input variables for the two trainings in the $3l$ channel; it reveals that some variables show opposite behavior for the two background sources, which results in potentially screening the discrimination power if they were to be used in a single discriminant, i.e., if the training would join $t\bar{t}$ and $t\bar{t}$V. For some other variables the distributions are similar in both background cases. In contrast to the distributions in Figure 6.15 only the dominant backgrounds are included; however, the discrimination power agrees among plots.

Figures in the Appendix B.2, B.3, B.4, and B.5 show the input variables distributions for the $2lss$ and $3l$ channel as seen by the BDTG classifier.

Input variables correlations

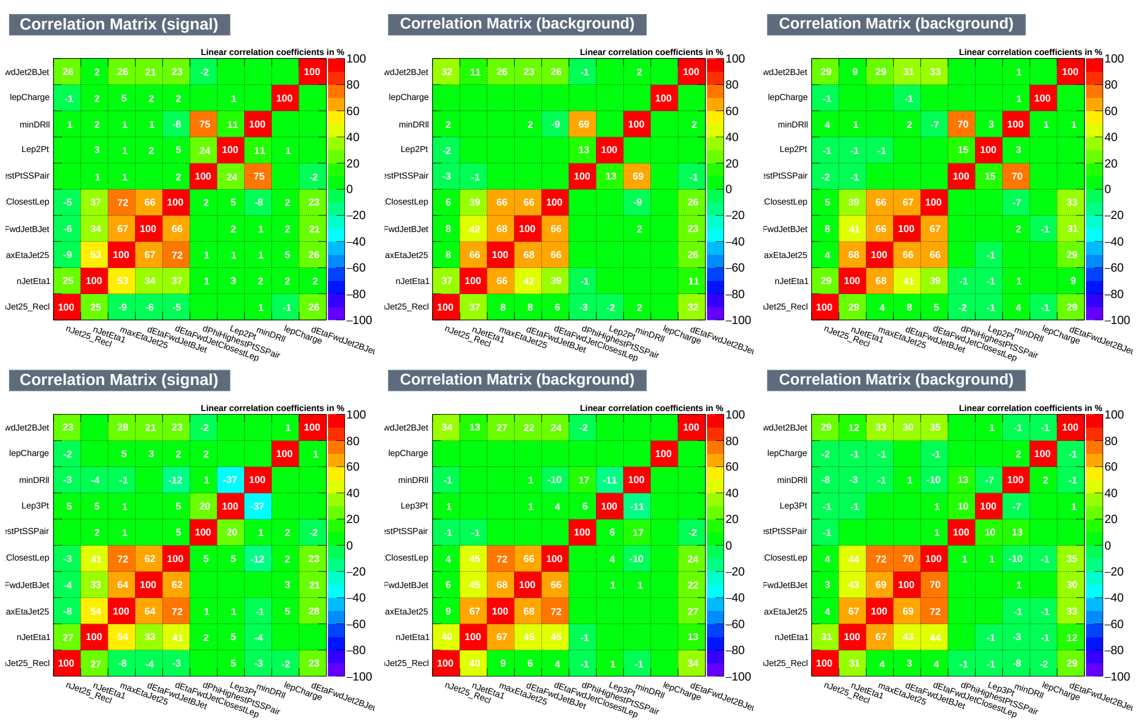

Figure 6.17: Signal (left), $t\bar{t}$ background (middle), and $t\bar{t}$V background (right.) correlation matrices for the input variables in the BDTG classifier for the $2lss$ (top) and $3l$ (bottom) channels.

From Table 6.10, it is clear that the input variables are correlated to some extend.

These correlations play an important role for some MVA methods like the Fisher discriminant method in which the first step consist of performing a linear transformation to an phase space where the correlations between variables are removed. In the case of BDT, correlations do not affect the performance. Figure 6.17 shows the linear correlation coefficients for signal and background for the two training cases (the signal values are identical by construction). As expected, strong correlations appears for variables related to the forward jet activity.

6.9.3 BDTG classifiers response

After the training stage, the BDTG classifier is tested to ensure its ability to discriminate between simulated signal and background events. The BDTG classifier output distributions for signal and backgrounds in the $3l$ channel are shown in Figure 6.18. As expected, a good discrimination power is obtained using default discriminator parameter values; some overtraining is also visible.

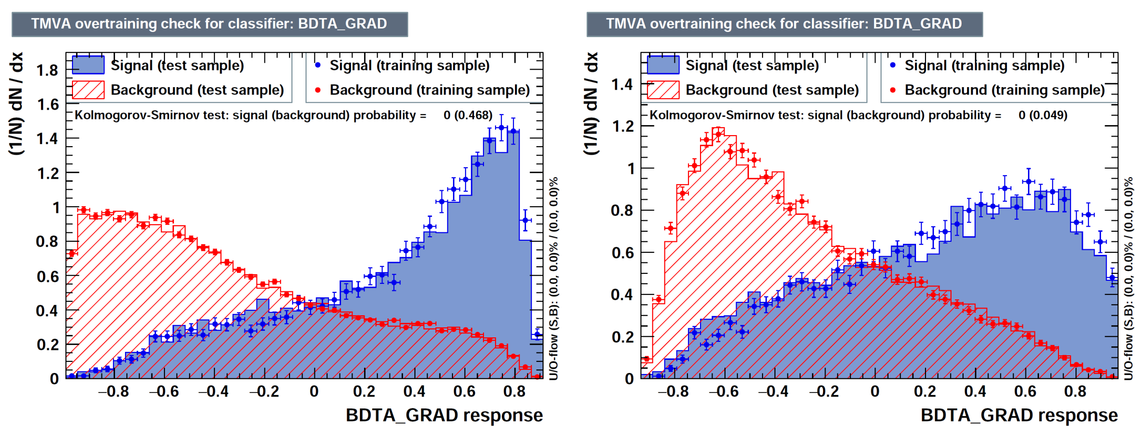

Figure 6.18: BDTG classifier output for trainings against $t\bar{t}$V(left) and $t\bar{t}$ (right). Default BDTG parameters have been used.

In order to explore further optimization in the BDTG performance, several changes from the default BDTG parameters were tested; Table 6.11 list the set of parameters found to be most discriminant with minimal overtraining as shown in Figure 6.19.

TMVA.Types.kBDT		
Option	Default	Used
NTrees	200	800
BoostType	AdaBoost	Grad
Shrinkage	1	0.1
nCuts	20	50
MaxDepth	3	

Table 6.11: Configuration used in the final BDTG training. Parameters not listed were not tested.

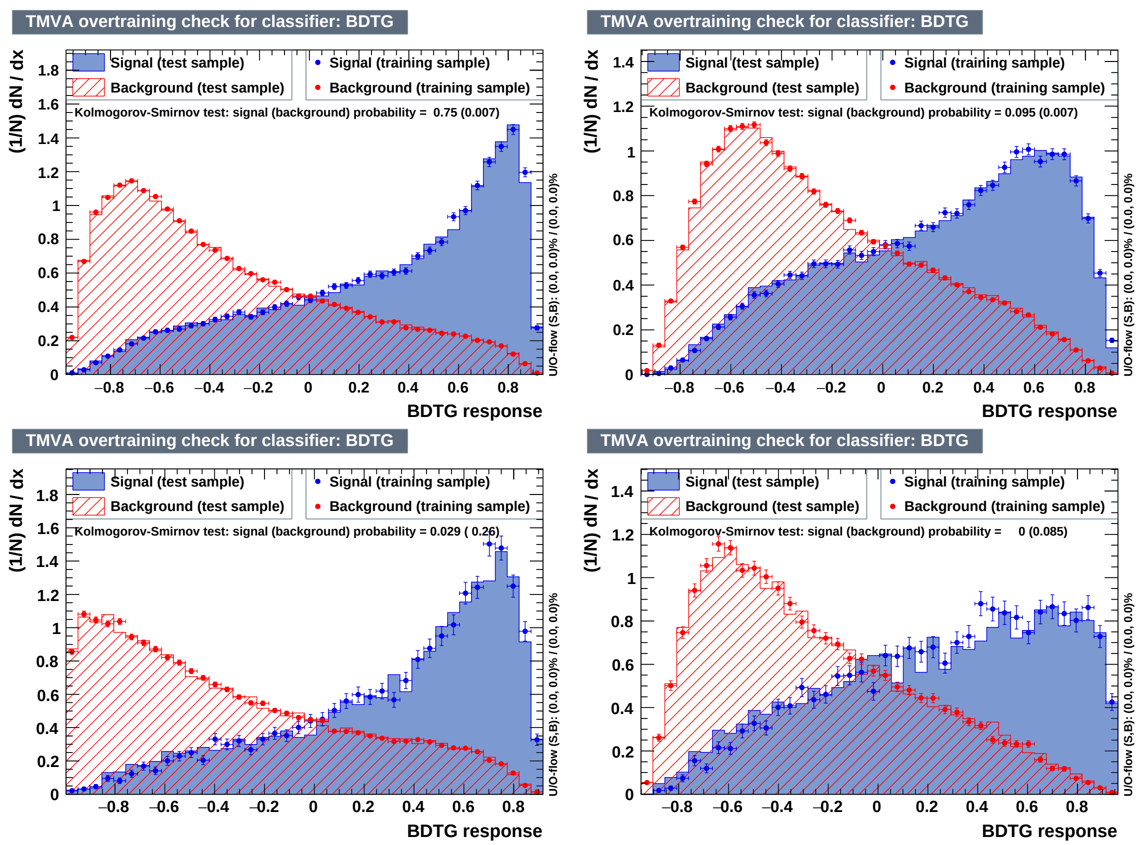

Figure 6.19: BDTG classifiers output for training against $t\bar{t}$V(left) and $t\bar{t}$ (right) for $2lss$ channel(top) and $3l$ channel (bottom) .

The ranking of the input variables by their importance in the classification process is shown in Table 6.12; for both trainings the rankings show almost the same 5 variables in the first places.

	$2lss$ channel		$3l$ channel	
	$t\bar{t}$ training	$t\bar{t}$Vtraining	$t\bar{t}$ training	$t\bar{t}$Vtraining
Rank	Variable	Variable	Variable	Variable
1	minDRll	dEtaFwdJetBJet	dEtaFwdJetClosestLep	maxEtaJet25
2	dEtaFwdJetClosestLep	Lep3Pt	minDRll	dEtaFwdJet2BJet
3	dEtaFwdJetBJet	maxEtaJet25	maxEtaJet25	dEtaFwdJetBJet
4	dPhiHighestPtSSPair	dEtaFwdJet2BJet	dPhiHighestPtSSPair	Lep2Pt
5	Lep3Pt	dEtaFwdJetClosestLep	Lep2Pt	dEtaFwdJetClosestLep
6	maxEtaJet25	minDRll	dEtaFwdJetBJet	minDRll
7	dEtaFwdJet2BJet	dPhiHighestPtSSPair	dEtaFwdJet2BJet	nJet25
8	nJetEta1	nJet25	nJetEta1	dPhiHighestPtSSPair
9	nJet25	nJetEta1	nJet25	nJetEta1
10	lepCharge	lepCharge	lepCharge	lepCharge

Table 6.12: Input variables ranking for BDTG classifiers for the trainings in the $3l$ channel and $2lss$ channel. In both trainings the rankings show almost the same 5 variables in the first places.

6.9.4 Additional discriminating variables

Given that the forward jet in background processes could be originated from pileup, two additional discriminating variables accounting for that were tested. These additional variables describe the forward jet momentum (fwdJetPt25) and the forward jet identification(fwdJetPUID); their distributions in the $3l$ channel are shown in Figure 6.20. The forward jet identification distribution show that for both, signal and background, jets are mostly originated in the primary vertex.

The testing was performed by including in the BDTG input one variable at a time, so the discrimination power of each variable can be evaluated individually, and then both simultaneously. fwdJetPUID was ranked in the last place in importance (11) in both training ($t\bar{t}$Vand $t\bar{t}$) while fwdJetPt25 was ranked 3 in the $t\bar{t}$Vtraining and 7 in the $t\bar{t}$ training. When training using 12 variables, fwdJetPt25 was ranked 5 and 7 in the $t\bar{t}$Vand $t\bar{t}$ trainings respectively, while fwdJetPUID was ranked 12 in both cases.

The improvement in the discrimination performance provided by the additional variables is about 1%, so it was decided not to include them in the procedure. Table 6.13 show the ROC-integral for all the testing cases performed.

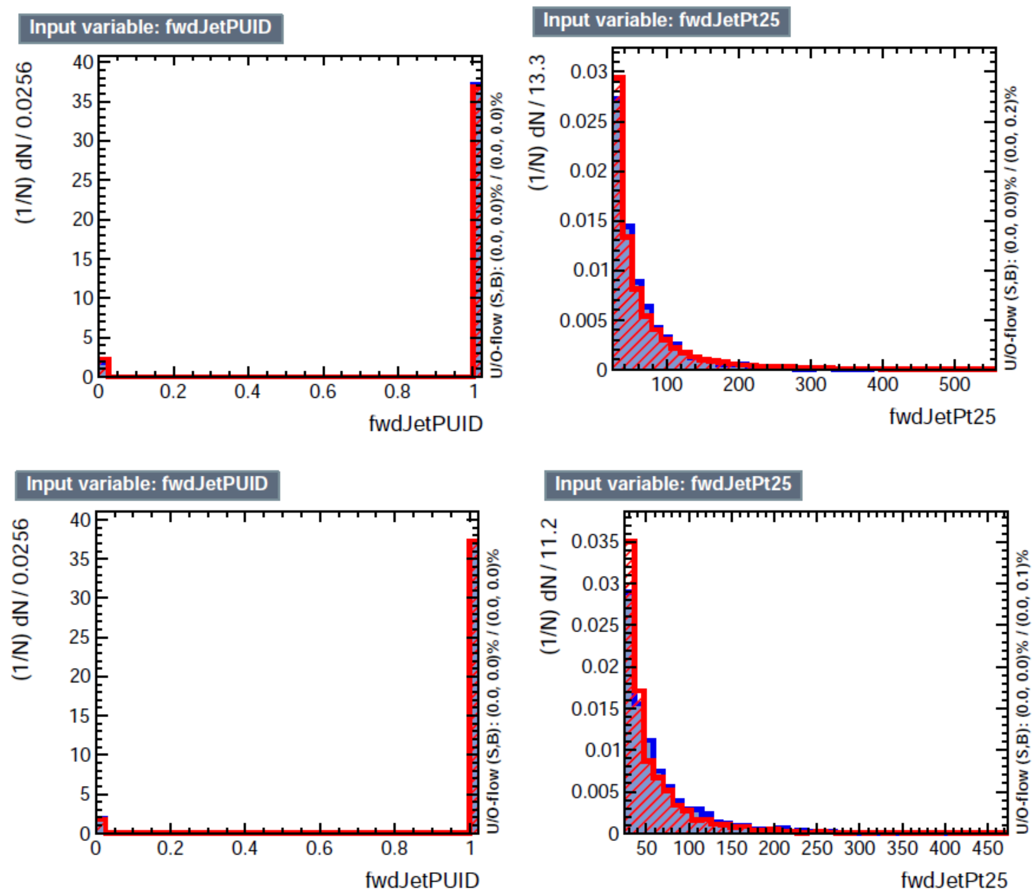

Figure 6.20: Additional discriminating variables distributions for $t\bar{t}$Vtraining (top row) and $t\bar{t}$ training (bottom row) in the $3l$ channel. The origin of the jets in the forward jet identification distribution is tagged as 0 for *pileup jets* while *primary vertex jets* are tagged as 1.

	ROC-integral	
	$t\bar{t}$V	$t\bar{t}$
base 10 var	0.848	0.777
+ fwdJetPUID	0.849	0.777
+ fwdJetPt25	0.856	0.787
12 var	0.856	0.787

Table 6.13: ROC-integral for all the testing cases performed in the evaluation of the additional variables discriminating power. The improvement in the discrimination performance provided by the additional variables is about 1% .

6.9.5 Signal extraction procedure

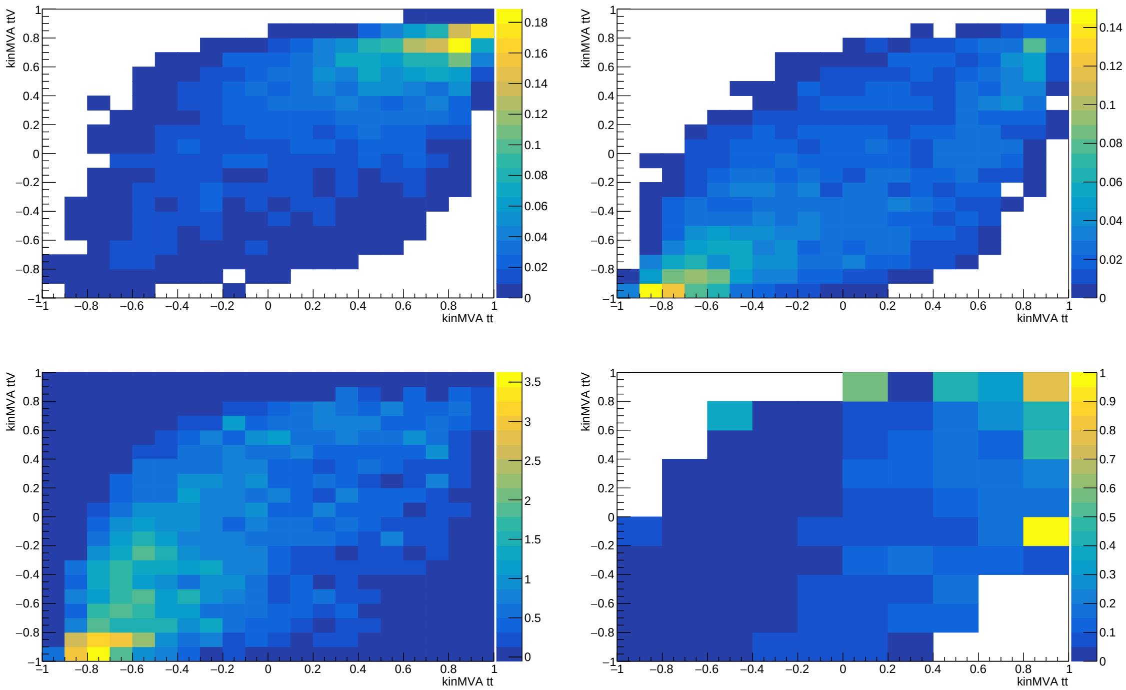

Figure 6.21: BDT classifier output planes (training vs $t\bar{t}$ on x-axis and vs $t\bar{t}$V on y-axis) for the tHq and tHW signals (top row), and for the combined backgrounds (bottom left). Bottom right: S/B ratio (combining tHq and tHW) in the same plane. Plots are for $3l$ channel.

Once the two BDTG classifiers, introduced in the previous section, are trained against the dominant backgrounds in each channel, they are used to classify the events in the samples; their outputs are then used to evaluate the signal cross section limits in a fit to the classifier shape. Figure 6.21 shows the expected output distributions in a 2D plane of one training vs. the other, i.e., $t\bar{t}$V vs. $t\bar{t}$. Top row shows the 2D planes for tHq and tHWsignals, while the bottom left plot shows the corresponding 2D plane for the combined backgrounds, which are evaluated as in the final background prediction, i.e., these are not the samples used in the BDTG training and this includes data-driven backgrounds. The signal (combining of tHq and tHW) to background ratio (S/B) is showed in the bottom right plot of Figure 6.21.

Each event is now classified into one of ten 2D-bins according to its position in the plane, as shown in Figure 6.22. The number of bins is chosen such that no bins are entirely empty for any process. The bin boundary positions and number of bins have been studied and optimized with respect to the expected limit on the signal strength (see Sec. 6.9.6).

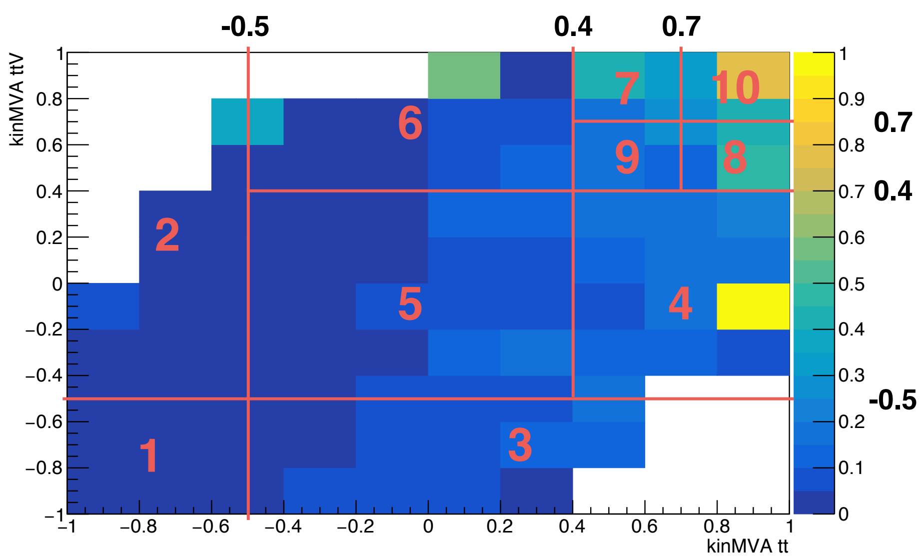

Figure 6.22: Binning overlaid on the S/B ratio map on the plane of classifier outputs.

From this event categorization, a 1D histogram of expected distribution is produced for each signal and background process, and fit to the observed data (or the Asimov dataset for expected limits).

6.9.6 Binning and selection optimization

The effect of the choice of pre-selection cuts and the number of bins of the 1D histogram on the cross section limit is evaluated by varying the most important cuts and re-calculating the limit in each case. In this analysis, the optimization was performed in the $3l$ channel, by evaluating the upper limits on the $tHq + tHW$ expected signal strength only (without $t\bar{t}H$ component), always evaluated at $\kappa_t = -1.0$, $\kappa_V = 1.0$.

Table 6.14 shows the several variations explored, compared with a baseline; the baseline is similar to the selection reported in Table 6.6 but only a loose CSV jet and a Z veto of ± 10 GeV are required.

Selection	Variation	Expected limit
Baseline		< 2.93
Loose CSV tags	$\geq 1 \rightarrow \geq 2$	< 3.81
Medium CSV tags	$\geq 0 \rightarrow \geq 1$	< 2.76
Light forward jet η	$\geq 0 \rightarrow \geq 1$	< 2.94
Light forward jet η	$\geq 0 \rightarrow \geq 1.5$	< 3.00
MET>30 GeV		< 2.91
Z veto ($\lvert m_{\ell\ell} - m_Z \rvert$)	> 10GeV $\rightarrow > 15$ GeV	< 2.79
One medium CSV + 15 GeV Z veto	combined	< 2.62

Table 6.14: Signal strength limit variation as a function of tighter cuts. The baseline selection corresponds to a looser selection compared to the one reported in Tab. 6.6 where only a CSV-loose b-jet is required, and the Z veto is loosened to ± 10 GeV. The optimal selection determined here corresponds to the baseline plus the two variations in the last row.

The optimal limit is found when requiring a slightly tighter selection with respect to the baseline. The optimal selection is reported in Table 6.6.

The signal strength limit also depends on the chosen binning in the 2D plane as the S/B ratio varies across the plane, hence, several sizes and binning combinations were tested in order to improve the limit. Figure 6.23 shows some of the binning combinations tested; in the default combination all the bins have the same size, while the best limit was found for a set of 10 bins. The bin borders and the resulting limits are shown in Table 6.15.

Combining the optimization of binning and using the tighter pre-selection cuts, the expected limit in the $3l$ channel alone reaches $\mathbf{r<2.59}$.

A similar binning optimization was made for $2lss$ channel, including other binning

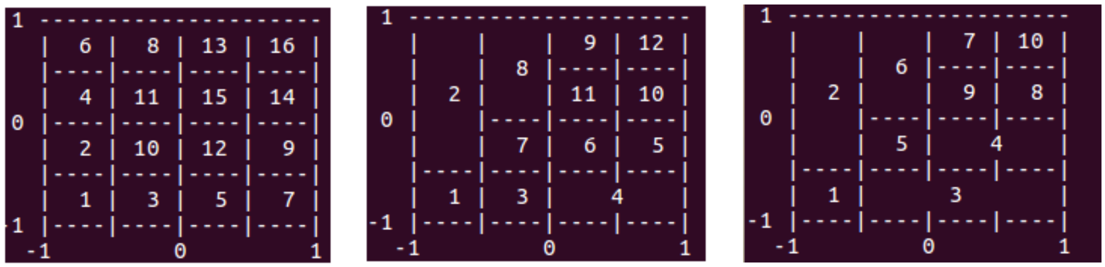

Figure 6.23: Binning combination scheme.

Number of bins	Bin borders						Expected limit
	x_1	x_2	x_3	y_1	y_2	y_3	
16 (default)	-0.5	0.0	0.5	-0.5	0.0	0.5	< 2.91
16	-0.5	0.3	0.7	-0.5	0.3	0.7	< 2.83
10	-0.5	0.0	0.5	-0.5	0.0	0.5	< 2.93
10	-0.5	0.0	0.7	-0.5	0.0	0.7	< 2.86
10	-0.5	0.0	0.7	-0.5	0.0	0.5	< 2.84
10	-0.5	0.0	0.5	-0.5	0.0	0.7	< 2.87
10	**-0.5**	**0.4**	**0.7**	**-0.5**	**0.4**	**0.7**	$<$ **2.81**

Table 6.15: Limit variation as a function of bin size. The final bin borders used in the $3l$ channel are indicated in bold.

combinations. First, the $3l$ channel binning was used to estimate the expected limit, then, bin borders were varied to obtain the best possible expected limit. The bin borders and the resulting signal strength limits for the same-sign dimuon channel are shown in Table 6.16:

The expected limit was found to be **r<1.69** for optimized bin borders in 10 bins and optimized pre-selection cuts.

Two additional binning strategies were tested, however, the obtained limits are degraded; they are documented in Appendix C.

. Multivariate techniques are used to discriminate the signal from the dominant backgrounds. The analysis yields a 95% confidence level (C.L.) upper limit on the combined tH + ttH production cross section times branching ratio of 0.64 pb, with

Number of bins	Bin borders						Expected limit
	x_1	x_2	x_3	y_1	y_2	y_3	
16	-0.5	0.4	0.7	-0.5	0.4	0.7	< 1.72
12	-0.5	0.4	0.7	-0.5	0.4	0.7	< 1.72
12	-0.3	0.4	0.7	-0.5	0.4	0.7	< 1.71
12	-0.3	0.3	0.7	-0.5	0.4	0.7	< 1.71
12	-0.3	0.3	0.7	-0.4	0.4	0.7	< 1.70
12	-0.3	0.3	0.7	-0.3	0.4	0.7	< 1.70
12	-0.3	0.3	0.7	-0.3	0.2	0.7	< 1.68
12	-0.3	0.3	0.7	-0.3	0.1	0.7	< 1.70
12	-0.3	0.3	0.7	-0.3	0.2	0.6	< 1.70
10	-0.5	0.4	0.7	-0.5	0.4	0.7	< 1.75
10	**-0.3**	**0.3**	**0.7**	**-0.3**	**0.2**	**0.6**	$< \mathbf{1.69}$

Table 6.16: Limit variation as a function of bin size in the same-sign dimuon channel. (In bold: the final bin borders used in the $2lss$ channel.)

an expected limit of 0.32 pb, for a scenario with kt = âĹŠ1.0 and kV = 1.0. Values of kt outside the range of âĹŠ1.25 to +1.60 are excluded at 95% C.L., assuming kV = 1.0.

Dont forget to mention previous constrains to ct check Reference **??** and References https://link.springer.com/content/pdf/10.1007%2FJHEP01%282013%29088.pdf (paragraph after eq 2)

6.10 Signal model

The goal of this analysis is to test the compatibility of points in the parameter space of Higgs-to-vector boson and Higgs-to-top quark couplings. The simulated tHq, tHW, and $t\bar{t}H$ signal events are used with event-by-event weights to reflect the impact of the couplings on kinematic distributions, and together with different predictions of the respective production cross sections and branching ratios, we can produce limits for different values of κ_{V} and κ_t. (See Tab. A.3 for the set of κ_t and κ_{V} values generated.) The slight shape-dependence of the BDT outputs as a function of the couplings is documented in Appendix D.

Apart from the $\kappa_t/\kappa_\mathrm{V}$ interference of the tHq and tHW production cross sections, the cross section of $t\bar{t}H$ scales as κ_t^2. Furthermore, the Higgs branching fractions to vector bosons depend on κ_V, and the overall Higgs decay width depend both on κ_t and κ_V when considering resolved top-quark loops in the $H \to \gamma\gamma$, $H \to Z\gamma$, and $H \to gg$ decays. The relative contributions from $H \to WW$, $H \to ZZ$, and $H \to \tau\tau$ changes with changing κ_V.

We hence set an upper limit on the combined cross section times branching ratio of tHq, tHW, and $t\bar{t}H$.

If we assume a modifier for the Higgs-to-tau coupling (κ_τ) to be equal to κ_t, the relative fractions of WW, ZZ, and $\tau\tau$ in our selection will only depend on the ratio of $\kappa_t/\kappa_\mathrm{V}$. Any limit set at any given value of $\kappa_t/\kappa_\mathrm{V}$ is thus valid for all values of κ_t and κ_V with that ratio, and could then be compared with theoretical predictions of cross sections at different values of either modifier. Rather than as a function of the $\kappa_t/\kappa_\mathrm{V}$ ratio, limits could (equivalently) be reported as a function of the relative strength of Higgs-top and Higgs-vector-boson couplings, multiplied by the relative sign. Such a parameter, further referred to as f_t, as defined in Equation 6.7, spans the entire possible parameter space between -1.0 and 1.0, with the SM expectation at 0.5. Absolute values of 1.0 or 0.0 would then correspond to purely Higgs-top and purely Higgs-V couplings, respectively.

$$f_t = \mathrm{sign}(\kappa_t/\kappa_\mathrm{V}) \times \frac{\kappa_t^2}{\kappa_t^2 + \kappa_\mathrm{V}^2}. \tag{6.7}$$

Table 6.17 shows the points in the $\kappa_t/\kappa_\mathrm{V}$ and f_t parameter space that are mapped by the 51 individual κ_t and κ_V points.

The overall higgs decay width (modified by both κ_t and κ_V) becomes irrelevant

f_t	$\kappa_t/\kappa_{\mathrm{V}}$	$\kappa_{\mathrm{V}} = 0.5$	$\kappa_{\mathrm{V}} = 1.0$	$\kappa_{\mathrm{V}} = 1.5$
-0.973	-6.000	-3.00		
-0.941	-4.000	-2.00		
-0.900	-3.000	-1.50	-3.00	
-0.862	-2.500	-1.25		
-0.800	-2.000	-1.00	-2.00	-3.00
-0.692	-1.500	-0.75	-1.50	
-0.640	-1.333			-2.00
-0.610	-1.250		-1.25	
-0.500	-1.000	-0.50	-1.00	-1.50
-0.410	-0.833			-1.25
-0.360	-0.750		-0.75	
-0.308	-0.667			-1.00
-0.200	-0.500	-0.25	-0.50	-0.75
-0.100	-0.333			-0.50
-0.059	-0.250		-0.25	
-0.027	-0.167			-0.25
0.000	0.000	0.00	0.00	0.00
0.027	0.167			0.25
0.059	0.250		0.25	
0.100	0.333			0.50
0.200	0.500	0.25	0.50	0.75
0.308	0.667			1.00
0.360	0.750		0.75	
0.410	0.833			1.25
0.500	1.000	0.50	1.00	1.50
0.610	1.250		1.25	
0.640	1.333			2.00
0.692	1.500	0.75	1.50	
0.800	2.000	1.00	2.00	3.00
0.862	2.500	1.25		
0.900	3.000	1.50	3.00	
0.941	4.000	2.00		
0.973	6.000	3.00		

Table 6.17: The 33 distinct values of $\kappa_t/\kappa_{\mathrm{V}}$ and f_t as mapped by the 51 κ_t and κ_{V} points.

if limits are quoted as absolute cross sections rather than multiples of the expected cross section (which depends on the overall Higgs decay width).

The 1D histograms of events as categorized in regions of the 2D BDT plane is then used in a maximum likelihood fit of signal and background shapes, where the tHq, tHW, and $t\bar{t}H$ signals are floating with a common signal strength modifier r, producing a 95% C.L. upper limit the observed cross section of $tHq + tHW + t\bar{t}H$.

This is done separately for each point of κ_t and κ_{V}, where the cross sections and branching fractions are scaled accordingly in each point. Limits at fixed values of $\kappa_t/\kappa_{\mathrm{V}}$ are by construction identical. Tables **??**–**??** and **??**–**??** in Appendix **??** show the scalings of cross section times branching fraction, as well as branching fractions alone for each of the Higgs decay modes and each of the signal components.

Systematic uncertainties on the signal selection efficiency arise from correction factors applied to the simulated events to better match the measured detector performance and also from theoretical uncertainties in the modeling of the signal process. Scale factors applied to correct for data/MC differences in the trigger efficiency, lepton reconstruction and identification performance, and lepton selection efficiency carry a combined uncertainty of about 5from jet energy corrections is evaluated by varying the correction factors within their uncertainty and propagating the effect to the final result by recalculating all kinematic quantities. Effects on the overall normalization of event yields and on the shape of kinematic properties are both taken into account. Jet energy resolution effects have negligible impact on this

analysis. Correction factors for data/MC differences in the b-tagging performance are applied depending on the pT and Îů, and on the flavor of the jet, and their effect on the signal efficiency is evaluated by varying the factors within their measured uncertainty and recalculating the overall event scale factors. The uncertainties from unknown higher orders of tHq and tHW production are estimated from a change in the Q2 scale of double and half the initial value, evaluated for each point of Îžt and ÎžV. The ttH signal component has an uncertainty of about +5.8/âĹŠ 9.2scale variations and a further 3.6Uncertainties related to the choice of PDF set and its scale are estimated to be about 3.7tHq and about 4.0

Appendix A

Datasets and triggers

Dataset name
/JetHT/Run2016X-23Sep2016-vY/MINIAOD
/HTMHT/Run2016X-23Sep2016-vY/MINIAOD
/MET/Run2016X-23Sep2016-vY/MINIAOD
/SingleElectron/Run2016X-23Sep2016-vY/MINIAOD
/SingleMuon/Run2016X-23Sep2016-vY/MINIAOD
/SinglePhoton/Run2016X-23Sep2016-vY/MINIAOD
/DoubleEG/Run2016X-23Sep2016-vY/MINIAOD
/MuonEG/Run2016X-23Sep2016-vY/MINIAOD
/DoubleMuon/Run2016X-23Sep2016-vY/MINIAOD
/Tau/Run2016B-23Sep2016-v3/MINIAOD
/JetHT/Run2016H-PromptReco-v3/MINIAOD
/HTMHT/Run2016H-PromptReco-v3/MINIAOD
/MET/Run2016H-PromptReco-v3/MINIAOD
/SingleElectron/Run2016H-PromptReco-v3/MINIAOD
/SingleMuon/Run2016H-PromptReco-v3/MINIAOD
/SinglePhoton/Run2016H-PromptReco-v3/MINIAOD
/DoubleEG/Run2016H-PromptReco-v3/MINIAOD
/MuonEG/Run2016H-PromptReco-v3/MINIAOD
/DoubleMuon/Run2016H-PromptReco-v3/MINIAOD

Table A.1: Full 2016 dataset used in the analysis. In the first section of the table are listed the 23Sep2016 samples; in the Run2016X-23Sep-vY label, X:B-G tag the run while Y:1,3 tag the version of the data sample. Second section list the PromptReco version of the dataset.

Same-sign dilepton (==2 muons)
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v*
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v*
HLT_IsoMu22_v*
HLT_IsoTkMu22_v*
HLT_IsoMu22_eta2p1_v*
HLT_IsoTkMu22_eta2p1_v*
HLT_IsoMu24_v*
HLT_IsoTkMu24_v*
Same-sign dilepton (==2 electrons)
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Ele27_eta2p1_WPLoose_Gsf_v*
HLT_Ele27_WPTight_Gsf_v*
HLT_Ele25_eta2p1_WPTight_Gsf_v*
Same-sign dilepton (==1 muon, ==1 electron)
HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_IsoMu22_v*
HLT_IsoTkMu22_v*
HLT_IsoMu22_eta2p1_v*
HLT_IsoTkMu22_eta2p1_v*
HLT_IsoMu24_v*
HLT_IsoTkMu24_v*
HLT_Ele27_WPTight_Gsf_v*
HLT_Ele25_eta2p1_WPTight_Gsf_v*
HLT_Ele27_eta2p1_WPLoose_Gsf_v*
Three lepton
HLT_DiMu9_Ele9_CaloIdL_TrackIdL_v*
HLT_Mu8_DiEle12_CaloIdL_TrackIdL_v*
HLT_TripleMu_12_10_5_v*
HLT_Ele16_Ele12_Ele8_CaloIdL_TrackIdL_v*
HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu23_TrkIsoVVL_Ele8_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_v*
HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v*
HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v*
HLT_IsoMu22_v*
HLT_IsoTkMu22_v*
HLT_IsoMu22_eta2p1_v*
HLT_IsoTkMu22_eta2p1_v*
HLT_IsoMu24_v*
HLT_IsoTkMu24_v*
HLT_Ele27_WPTight_Gsf_v*
HLT_Ele25_eta2p1_WPTight_Gsf_v*
HLT_Ele27_eta2p1_WPLoose_Gsf_v*

Table A.2: Table of high-level triggers considered in the analysis.

		tHq		tHW		
κ_V	κ_t	sum of weights	cross section [pb]	sum of weights	cross section [pb]	LHE weights
1.0	-3.0	35.700022	2.991	11.030445	0.6409	LHEweight_wgt[446]
1.0	-2.0	20.124298	1.706	5.967205	0.3458	LHEweight_wgt[447]
1.0	-1.5	14.043198	1.205	4.029093	0.2353	LHEweight_wgt[448]
1.0	-1.25	11.429338	0.9869	3.208415	0.1876	LHEweight_wgt[449]
1.0	-1.0		0.7927		0.1472	
1.0	-0.75	7.054998	0.6212	1.863811	0.1102	LHEweight_wgt[450]
1.0	-0.5	5.294518	0.4723	1.339886	0.07979	LHEweight_wgt[451]
1.0	-0.25	3.818499	0.3505	0.914880	0.05518	LHEweight_wgt[452]
1.0	0.0	2.627360	0.2482	0.588902	0.03881	LHEweight_wgt[453]
1.0	0.25	1.719841	0.1694	0.361621	0.02226	LHEweight_wgt[454]
1.0	0.5	1.097202	0.1133	0.233368	0.01444	LHEweight_wgt[455]
1.0	0.75	0.759024	0.08059	0.204034	0.01222	LHEweight_wgt[456]
1.0	1.0	0.705305	0.07096	0.273617	0.01561	LHEweight_wgt[457]
1.0	1.25	0.936047	0.0839	0.442119	0.02481	LHEweight_wgt[458]
1.0	1.5	1.451249	0.1199	0.709538	0.03935	LHEweight_wgt[459]
1.0	2.0	3.335034	0.2602	1.541132	0.08605	LHEweight_wgt[460]
1.0	3.0	10.516125	0.8210	4.391335	0.2465	LHEweight_wgt[461]
1.5	-3.0	45.281492	3.845	13.426212	0.7825	LHEweight_wgt[462]
1.5	-2.0	27.606715	2.371	7.809713	0.4574	LHEweight_wgt[463]
1.5	-1.5	20.476088	1.784	5.594971	0.3290	LHEweight_wgt[464]
1.5	-1.25	17.337465	1.518	4.635978	0.2749	LHEweight_wgt[465]
1.5	-1.0	14.483302	1.287	3.775902	0.2244	LHEweight_wgt[466]
1.5	-0.75	11.913599	1.067	3.014744	0.1799	LHEweight_wgt[467]
1.5	-0.5	9.628357	0.874	2.352505	0.1410	LHEweight_wgt[468]
1.5	-0.25	7.627574	0.702	1.789184	0.1081	LHEweight_wgt[469]
1.5	0.0	5.911882	0.5577	1.324946	0.08056	LHEweight_wgt[470]
1.5	0.25	4.479390	0.4365	0.959295	0.05893	LHEweight_wgt[471]
1.5	0.5	3.331988	0.3343	0.692727	0.04277	LHEweight_wgt[472]
1.5	0.75	2.469046	0.2558	0.525078	0.03263	LHEweight_wgt[473]
1.5	1.0	1.890565	0.2003	0.456347	0.02768	LHEweight_wgt[474]
1.5	1.25	1.596544	0.1689	0.486534	0.02864	LHEweight_wgt[475]
1.5	1.5	1.586983	0.1594	0.615638	0.03509	LHEweight_wgt[476]
1.5	2.0	2.421241	0.2105	1.170602	0.06515	LHEweight_wgt[477]
1.5	3.0	7.503280	0.5889	3.467546	0.1930	LHEweight_wgt[478]
0.5	-3.0	27.432685	2.260	8.929074	0.5136	LHEweight_wgt[479]
0.5	-2.0	13.956013	1.160	4.419093	0.2547	LHEweight_wgt[480]
0.5	-1.5	8.924438	0.7478	2.757611	0.1591	LHEweight_wgt[481]
0.5	-1.25	6.835341	0.5726	2.075247	0.1204	LHEweight_wgt[482]
0.5	-1.0	5.030704	0.4273	1.491801	0.08696	LHEweight_wgt[483]
0.5	-0.75	3.510528	0.2999	1.007273	0.05885	LHEweight_wgt[484]
0.5	-0.5	2.274811	0.1982	0.621663	0.03658	LHEweight_wgt[485]
0.5	-0.25	1.323555	0.1189	0.334972	0.01996	LHEweight_wgt[486]
0.5	0.0	0.656969	0.06223	0.147253	0.008986	LHEweight_wgt[487]
0.5	0.25	0.274423	0.02830	0.058342	0.003608	LHEweight_wgt[488]
0.5	0.5	0.176548	0.01778	0.068404	0.003902	LHEweight_wgt[489]
0.5	0.75	0.363132	0.03008	0.177385	0.009854	LHEweight_wgt[490]
0.5	1.0	0.834177	0.06550	0.385283	0.02145	LHEweight_wgt[491]
0.5	1.25	1.589682	0.1241	0.692099	0.03848	LHEweight_wgt[492]
0.5	1.5	2.629647	0.2047	1.097834	0.06136	LHEweight_wgt[493]
0.5	2.0	5.562958	0.4358	2.206057	0.1246	LHEweight_wgt[494]
0.5	3.0	14.843102	1.177	5.609519	0.3172	LHEweight_wgt[495]

Table A.3: κ_V and κ_t combinations generated for the two signal samples and their NLO cross sections. The tHq cross section is multiplied by the branching fraction of the enforced leptonic decay of the top quark (0.324) [150].

Sample	σ [pb]	
TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.2043	*
TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.2529	*
/store/cmst3/group/susy/gpetrucc/13TeV/u/TTLL_m1to10_LO_NoMS_for76X/	0.0283	
WGToLNuG_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	585.8	
ZGTo2LG_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	131.3	
TGJets_TuneCUETP8M1_13TeV_amcatnlo_madspin_pythia8	2.967	
TGJets_TuneCUETP8M1_13TeV_amcatnlo_madspin_pythia8	2.967	
TTGJets_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	3.697	
WpWpJJ_EWK-QCD_TuneCUETP8M1_13TeV-madgraph-pythia8	0.03711	
ZZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.01398	
WWZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.1651	
WZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.05565	
WW_DoubleScattering_13TeV-pythia8	1.64	
tZq_ll_4f_13TeV-amcatnlo-pythia8_TuneCUETP8M1	0.0758	
ST_tWll_5f_LO_13TeV-MadGraph-pythia8	0.01123	
TTTT_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.009103	
WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8	4.4296	
ZZTo4L_13TeV_powheg_pythia8	1.256	
TTJets_SingleLeptFromTbar_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	182.1754	*
TTJets_SingleLeptFromT_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	182.1754	*
TTJets_DiLept_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	87.3	*
DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	18610	
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6024	
WJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	61526.7	
ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.6	
ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.6	
ST_t-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	70.3144	
ST_t-channel_antitop_4f_leptonDecays_13TeV-powheg-pythia8_TuneCUETP8M1	26.2278	
ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	3.68064	
WWTo2L2Nu_13TeV-powheg	10.481	
ttWJets_13TeV_madgraphMLM	0.6105	
ttZJets_13TeV_madgraphMLM	0.5297/0.692	

Table A.4: List of background samples used in this analysis (CMSSW 80X). The first section of the table lists the samples used in simulation to extract the final yields and shapes; the second section lists the samples of the processes for which the yields are estimated from data. The MC simulation is used to design the data driven methods and in the derivation of the associated systematic uncertainties. The third section list the leading order $t\bar{t}W$and $t\bar{t}Z$samples, which in addition to the ones market with a *, where used in the BDT training.

Appendix B

Aditional plots

B.1 BDTG input variables distributions for $2lss$ channel

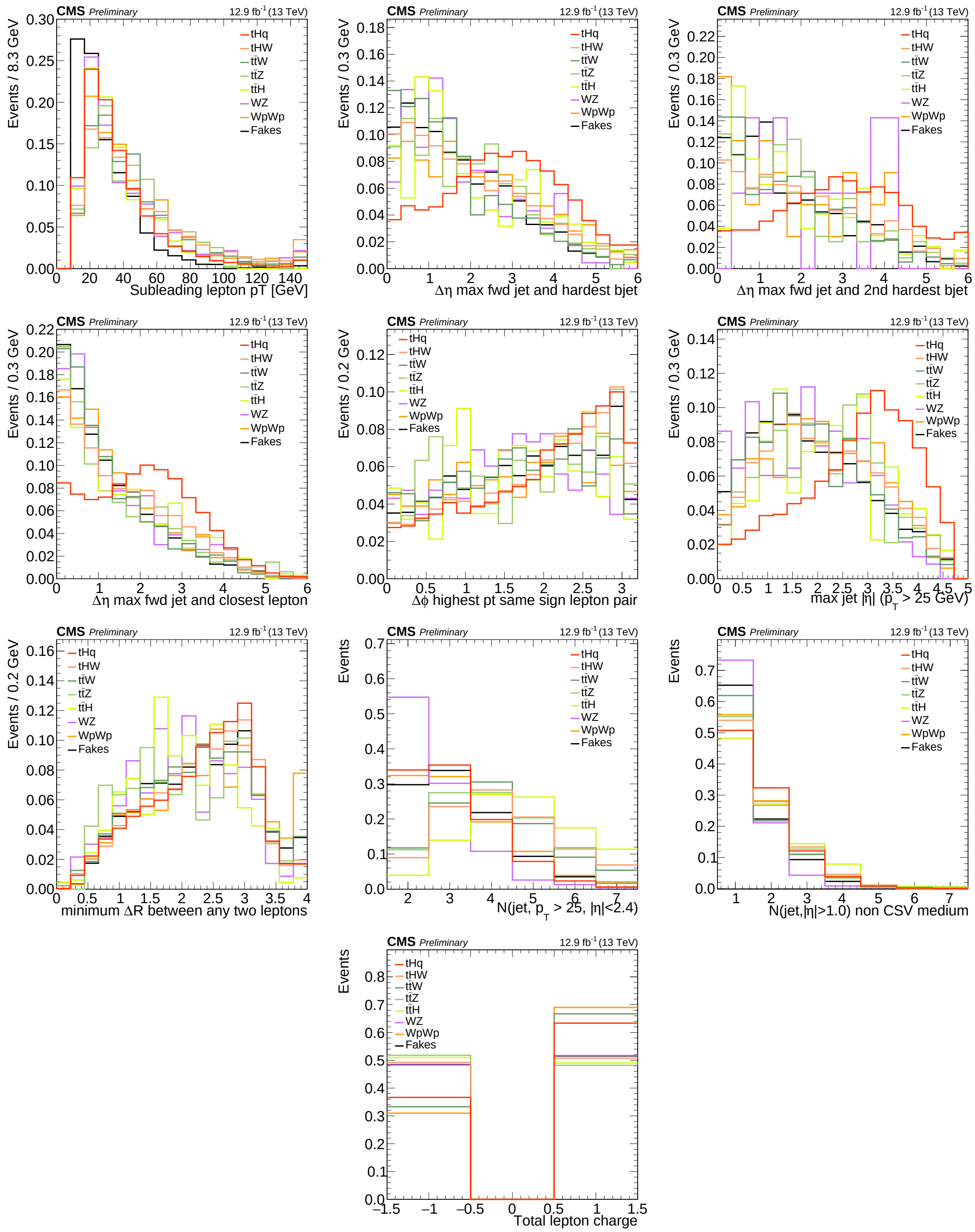

Figure B.1: Distributions of input variables to the BDT for signal discrimination, two lepton same sign channel.

B.2 Input variables distributions from BDTG classifiers

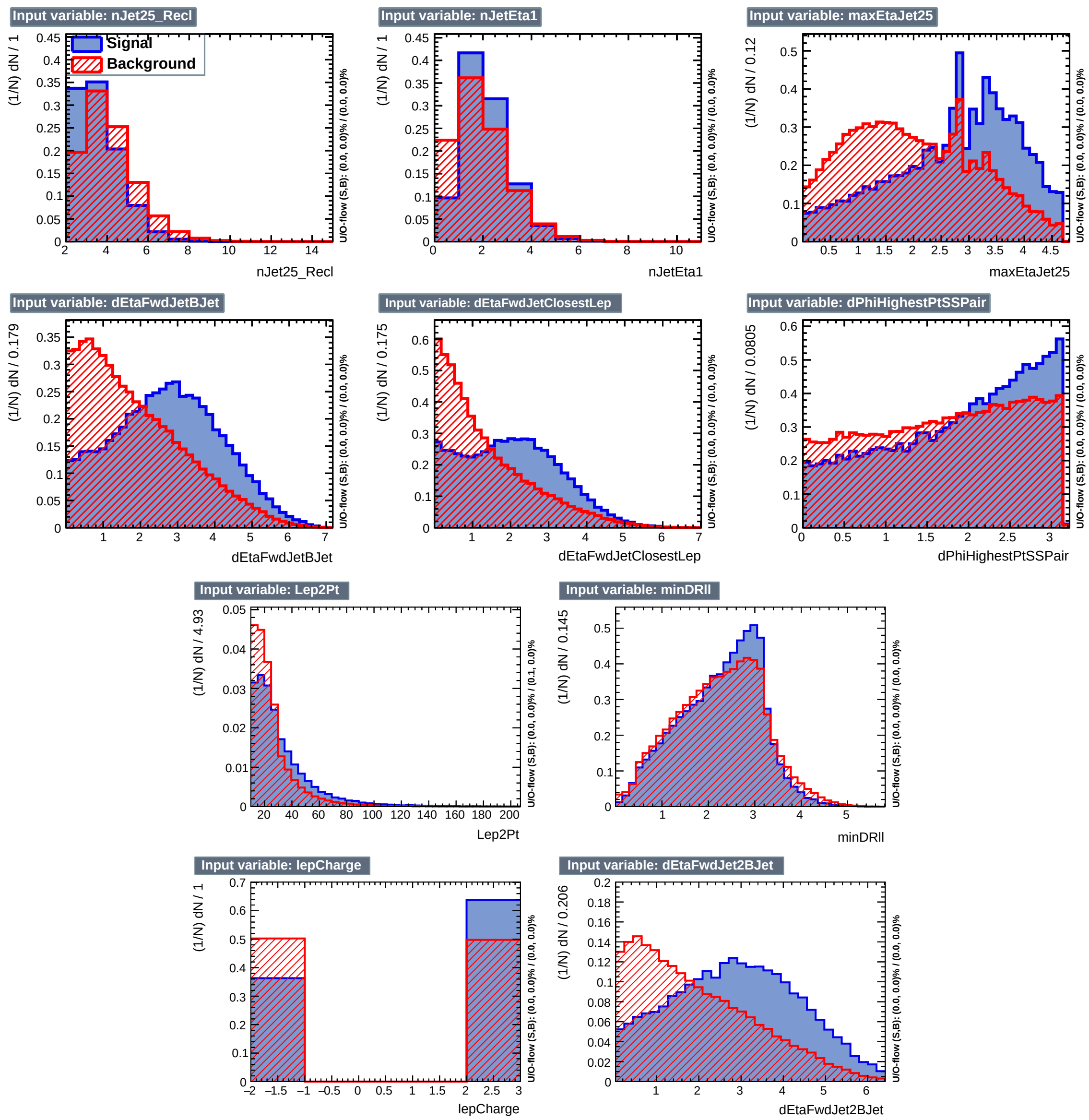

Figure B.2: BDT input variables as seen by BDTG classifier for the $2lss$ channel, tHq signal (blue) discriminated against $t\bar{t}$ background (red).

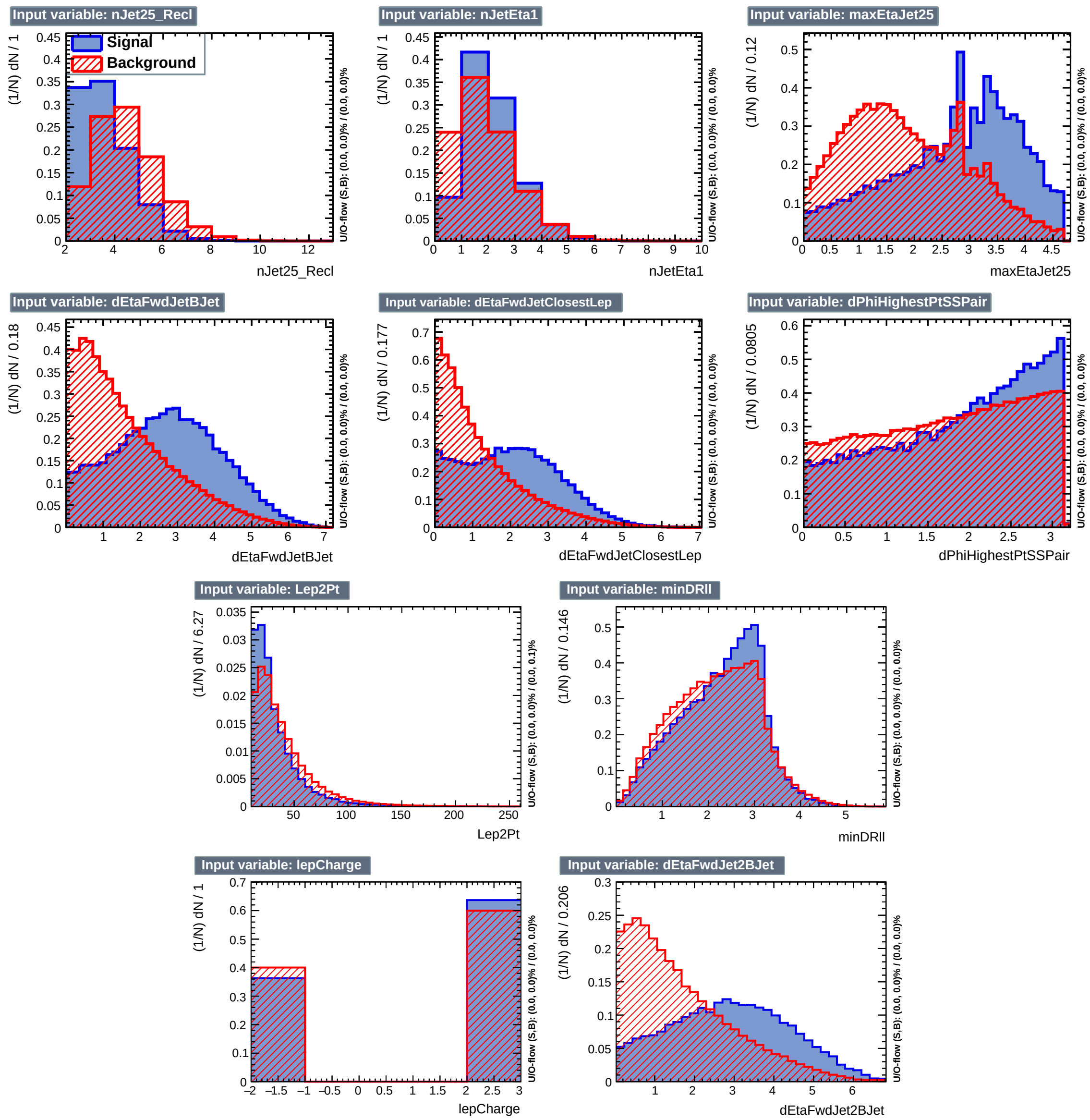

Figure B.3: BDT input variables as seen by BDTG classifier for the $2lss$ channel, tHq signal(blue) discriminated against $t\bar{t}$V background (red).

B.3 Pre-selection kinematic variables

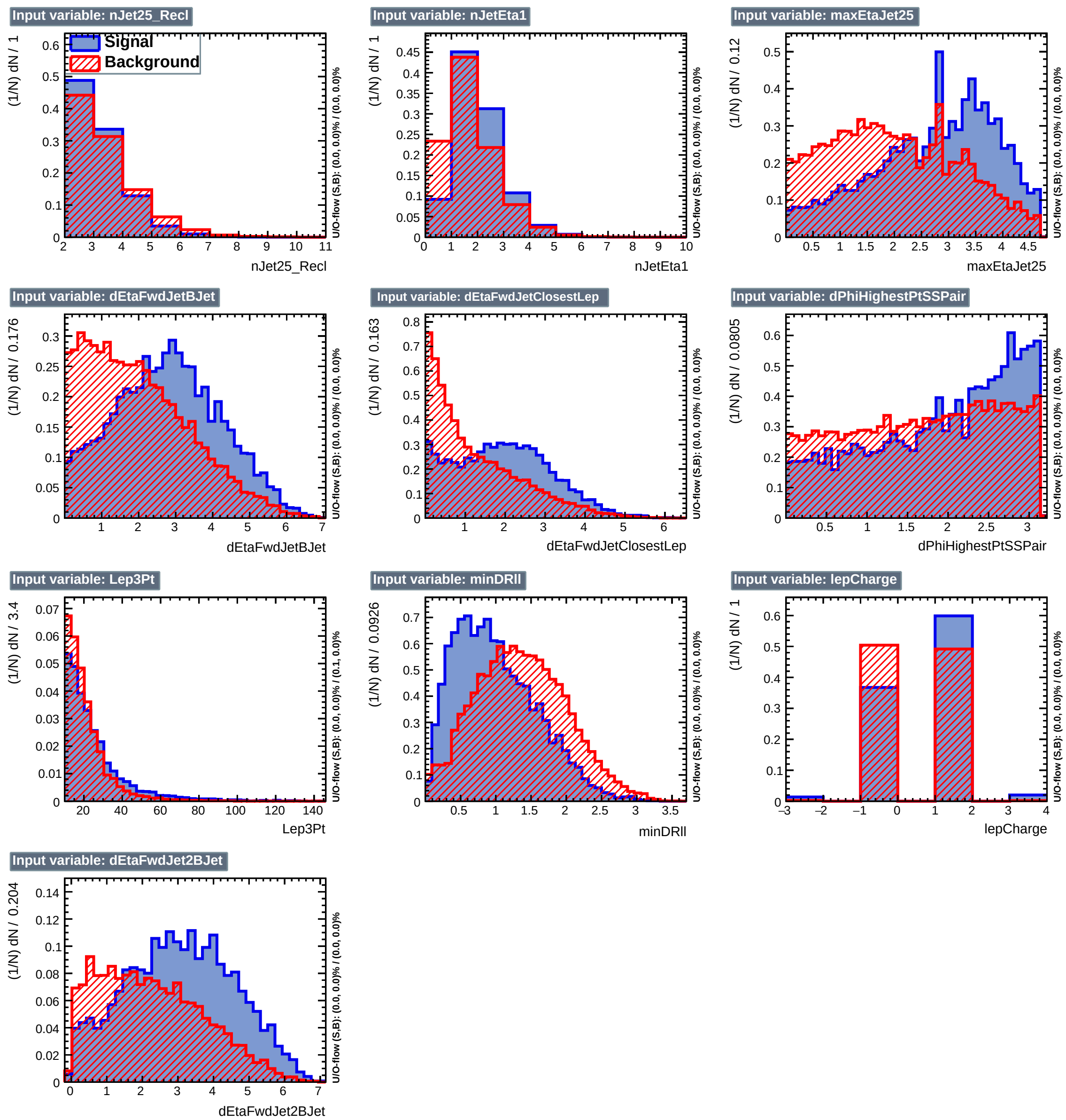

Figure B.4: BDT input variables as seen by BDTG classifier for the $3l$ channel, tHq signal (blue) discriminated against $t\bar{t}$ background (red).

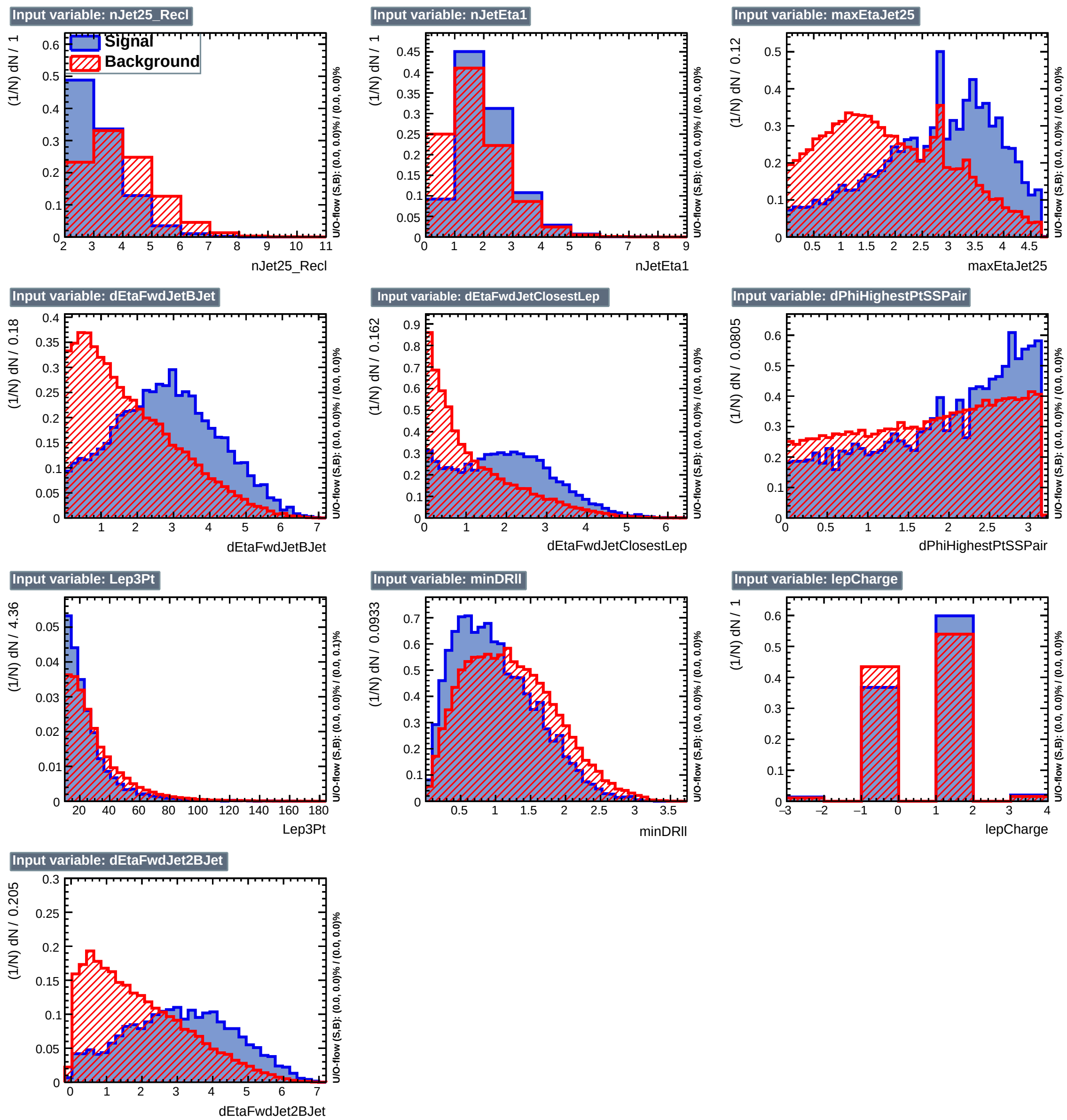

Figure B.5: BDT input variables as seen by BDTG classifier for the $3l$ channel, tHq signal (blue) discriminated against $t\bar{t}$V background (red).

Appendix C

Other binning strategies

Two additional strategies of clustering regions in the 2D plane of $BDTG_{tt}$ vs $BDTG_{ttV}$ into bins were attempted, following studies done and documented in great detail in Reference [149]. A brief description is provided in the following.

Clustering by S/B ratio In thin method, the 2D plane is clustered into a given number of bins corresponding to regions where S/B is within a certain range. The bin borders are determined such that the number of background events in each bin is approximately equal. The resulting regions for $2lss$ and $3l$ events are shown in Figure C.1, while the expected distribution of signal and dominant backgrounds are shown in Figure C.2.

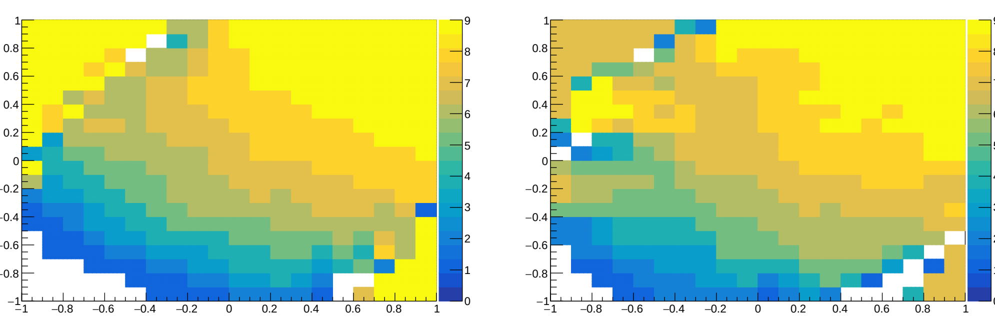

Figure C.1: Binning by S/B regions for $2lss$ (left) and $3l$ (right).

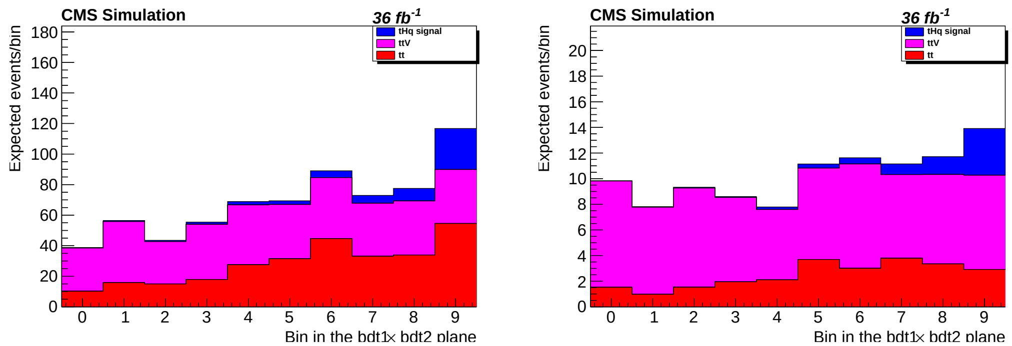

Figure C.2: Final bins (corresponding to S/B regions in the 2D plane) for $2lss$ and $3l$ (right).

Using this method, the resulting limits (for the $\kappa_t = -1, \kappa_\mathrm{V} = 1$ scenario) are about 20% worse than with the binning in Section 6.9.6: $\mu^\pm\mu^\pm$ changed from 1.82 to 2.15, $3l$ changed from 1.52 to 1.75.

***k*-Means geometric clustering** This method employs a recursive application of the k-means algorithm (see Appendix D in Reference [149]) to separate the 2D plane into geometric regions. The resulting clustering (using the $t\bar{t}H$ multilepton code on tHq signal and $t\bar{t}$ and $t\bar{t}$V background events) are shown in Figure C.3. The expected distribution of events for the signal and dominant backgrounds in these bins is shown in Fig. C.4.

Similarly to the S/B ratio binning, the limits using the k-means clustering are significantly worse than those of the bins described before. In the $\mu^\pm\mu^\pm$ channel, the limit deteriorates from 1.82 to 2.05, whereas in $3l$ it changes from 1.58 to 1.78.

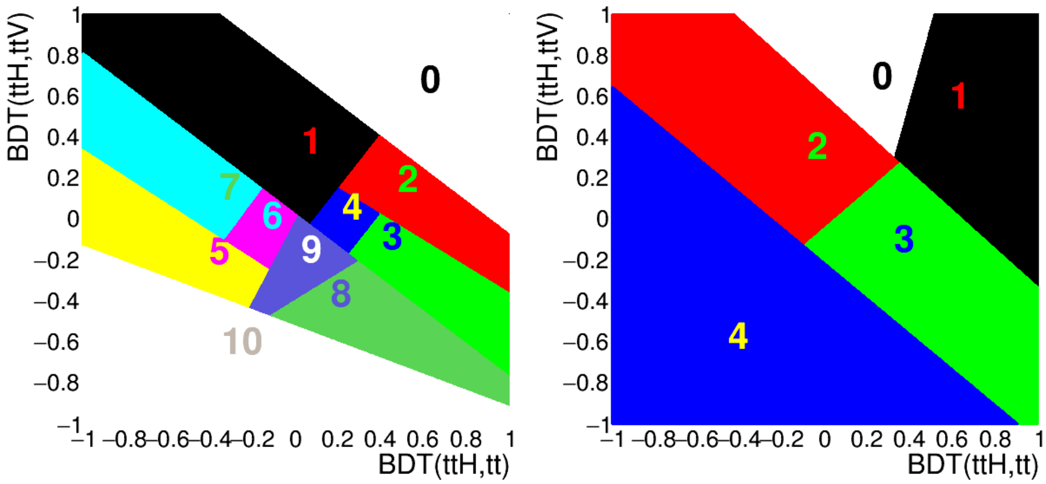

Figure C.3: Binning into geometric regions using a k-means algorithm for $2lss$ (left) and $3l$ (right).

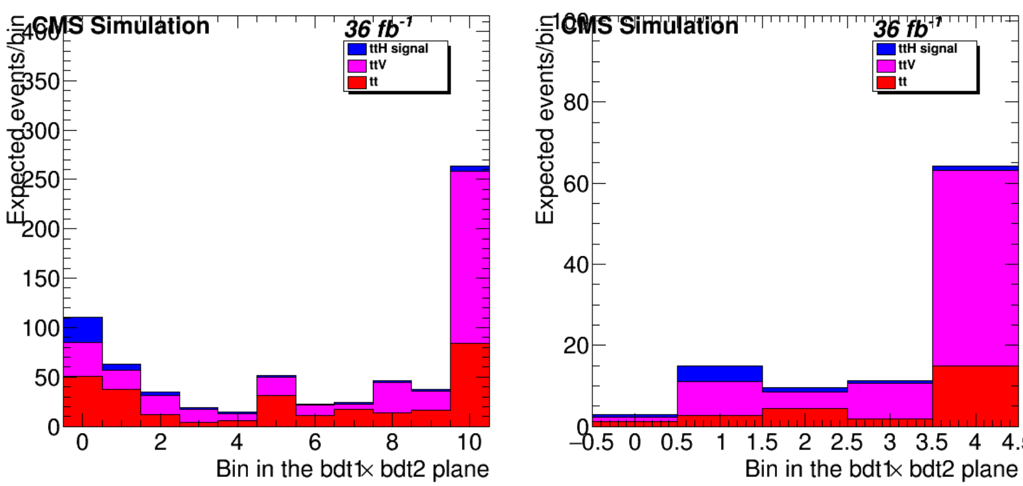

Figure C.4: Final bins using a k-means algorithm for $2lss$ (left) and $3l$ (right). Note that the bin numbering here is such that signal-like bins are lower.

Appendix D

BDTG output variation with κ_V/κ_t

The BDTG classifier output was described in Section in the κ_t =-1,κ_V =1 escenario; the change of BDTG classifiers output shape when varying the κ_V/κ_t coupling scenario is shown in Figure D.1 in the $3l$ channel for five different values of κ_t, with κ_V fixed at 1.0.

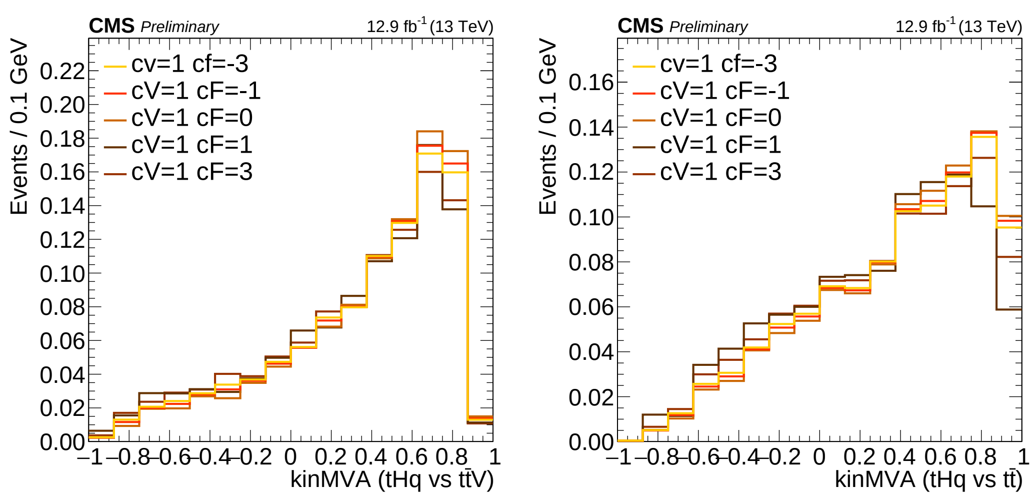

Figure D.1: Change of the BDTG classifiers output when varying κ_t coupling (κ_V is fixed at 1.0). Training vs. $t\bar{t}$V (right) and vs. $t\bar{t}$ (left).

Complete this section !!!!!!!! ask about this !

Appendix E

tHq-$t\bar{t}H$ overlap

This section provides a quick overview of the differences and commonalities in event selections between this analysis and the $t\bar{t}H$ multilepton search [149]. The object selections of the two analysis are perfectly synchronized due to shared frameworks and ROOT trees. The only exception is the usage of forward jets ($|\eta| > 2.4, p_T > 40$ GeV) in this analysis. Such jets are not considered in the $t\bar{t}H$ analysis.

Table E.1 gives an overview of the main differences in the event selections. Here, ${E_T^{miss}}_{\text{LD}}$ is defined as $E_T^{miss} \times 0.00397 + H_T^{miss} \times 0.00265$. Un-tagged jets in the tHq analysis are jets that do not pass the CSV loose working point and are either central ($|\eta < 2.4|, p_T > 25$ GeV) or forward ($|\eta < 2.4|, p_T > 40$ GeV). All jets in the $t\bar{t}H$ analysis are selected with $p_T > 25$ GeV. Lepton p_T cuts and the trigger selections are identical.

Channel	tHq	$t\bar{t}H$
3l	Z veto, 15bGeV	Z veto, 10 GeV
	$N_{\text{jets}}^{\text{b, med.}} \geq 1$	$N_{\text{jets}}^{\text{b, med.}} \geq 1$ OR $N_{\text{jets}}^{\text{b, loose}} \geq 2$
	≥ 1 un-tagged jet	${E_T^{miss}}_{\text{LD}} > 0.2$ OR $N_{\text{jets}}^{\text{centrl.}} \geq 4$
2lss	$N_{\text{jets}}^{\text{b, med.}} \geq 1$	$N_{\text{jets}}^{\text{b, med.}} \geq 1$ OR $N_{\text{jets}}^{\text{b, loose}} \geq 2$
	≥ 1 un-tagged jet	$N_{\text{jets}}^{\text{central}} \geq 4$

Table E.1: Differences in event selection between this analysis and the $t\bar{t}H$ multilepton analysis.

Table E.2 shows the total events yields in the individual channels, and the yield of shared events between each channel, for the tHq signal sample, the $t\bar{t}H$ signal sample, and the data. In the data, for the three lepton channel, about 80% of events passing the tHq selection also pass the $t\bar{t}H$ selection, constituting about 70% of that channel. In the same-sign dilepton channel, about 50% of data events passing the tHq selection also pass the $t\bar{t}H$ selection, but these events constitute almost 90% of the $t\bar{t}H$ selection in those channels. Similar overlaps are also seen in the tHq and $t\bar{t}H$ signal samples.

There is no migration between different channels and different selections, i.e. no events passing the selection of a given tHq channel pass the selection of any other channels of $t\bar{t}H$ and vice versa.

tHq **sample**	tHq	$t\bar{t}H$	Common	(% tHq)	(% $t\bar{t}H$)
$\mu^{\pm}\mu^{\pm}$	7400	2353	2166	29.3	92.1
$e^{\pm}\mu^{\pm}$	11158	3600	3321	29.8	92.2
$e^{\pm}e^{\pm}$	3550	1106	1025	28.9	92.7
$\ell\ell\ell$	3115	2923	2347	75.3	80.3

$t\bar{t}H$ **sample**	tHq	$t\bar{t}H$	Common	(% tHq)	(% $t\bar{t}H$)
$\mu^{\pm}\mu^{\pm}$	32612	28703	26547	81.4	92.5
$e^{\pm}\mu^{\pm}$	48088	42521	39164	81.4	92.1
$e^{\pm}e^{\pm}$	15476	12869	11896	76.9	92.4
$\ell\ell\ell$	26627	30598	25288	95.0	82.6

Data	tHq	$t\bar{t}H$	Common	(% tHq)	(% $t\bar{t}H$)
$\mu^{\pm}\mu^{\pm}$	280	160	140	50.0	87.5
$e^{\pm}\mu^{\pm}$	525	280	242	46.1	86.4
$e^{\pm}e^{\pm}$	208	90	79	38.0	87.8
$\ell\ell\ell$	126	154	104	82.5	67.5

Table E.2: Individual and shared event yields between this analysis (tHq) and $t\bar{t}H$ multilepton selections.

References

[1] J. Schwinger. “Quantum Electrodynamics. I. A Covariant Formulation”. Physical Review. 74 (10): 1439-61, (1948). https://doi.org/10.1103/PhysRev.74.1439

[2] R. P. Feynman. “Space-Time Approach to Quantum Electrodynamics”. Physical Review. 76 (6): 769-89, (1949). https://doi.org/10.1103/PhysRev.76.769

[3] S. Tomonaga. “On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields”. Progress of Theoretical Physics. 1 (2): 27-42, (1946). https://doi.org/10.1143/PTP.1.27

[4] D.J. Griffiths, “Introduction to electrodynamics”. 4th ed. Pearson, (2013).

[5] F. Mandl, G. Shaw. “Quantum field theory.” Chichester, Wiley (2009).

[6] F. Halzen, and A.D. Martin, “Quarks and leptons: An introductory course in modern particle physics”. New York: Wiley, (1984) .

[7] File: Standard_Model_of_Elementary_Particle_dark.svg. (2017, June 12) Wikimedia Commons, the free media repository. Retrieved November 27, 2017 from https://www.collegiate-advanced-electricity.com/single-post/2017/04/10/The-Standard-Model-of-Particle-Physics.

[8] E. Noether, “Invariante Variationsprobleme”, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, mathematisch-physikalische Klasse, vol. 1918, pp. 235-257, (1918).

[9] C. Patrignani et al. (Particle Data Group), Chin. Phys. C, 40, 100001 (2016) and 2017 update.

[10] M. Goldhaber, L. Grodzins, A.W. Sunyar “Helicity of Neutrinos”, Phys. Rev. 109, 1015 (1958).

[11] Palanque-Delabrouille N et al. “Neutrino masses and cosmology with Lyman-alpha forest power spectrum”, JCAP 11 011 (2015).

[12] M. Gell-Mann. “A Schematic Model of Baryons and Mesons”. Physics Letters. 8 (3): 214-215 (1964).

[13] G. Zweig. “An SU(3) Model for Strong Interaction Symmetry and its Breaking” (PDF). CERN Report No.8182/TH.401 (1964).

[14] G. Zweig. “An SU(3) Model for Strong Interaction Symmetry and its Breaking: II” (PDF). CERN Report No.8419/TH.412(1964).

[15] M. Gell-Mann. “The Interpretation of the New Particles as Displaced Charged Multiplets”. Il Nuovo Cimento 4: 848. (1956).

[16] T. Nakano, K, Nishijima. “Charge Independence for V-particles”. Progress of Theoretical Physics 10 (5): 581-582. (1953).

[17] N. Cabibbo, “Unitary symmetry and leptonic decays” Physical Review Letters, vol. 10, no. 12, p. 531, (1963).

[18] M.Kobayashi, T.Maskawa, “CP-violation in the renormalizable theory of weak interaction,” Progress of Theoretical Physics, vol. 49, no. 2, pp. 652-657, (1973).

[19] File: Weak Decay (flipped).svg. (2017, June 12). Wikimedia Commons, the free media repository. Retrieved November 27, 2017 from https://commons.wikimedia.org/w/index.php?title=File:Weak_Decay_(flipped)\.svg&oldid=247498592.

[20] Georgia Tech University. Coupling Constants for the Fundamental Forces(2005). Retrieved January 10, 2018, from http://hyperphysics.phy-astr.gsu.edu/hbase/Forces/couple.html#c2

[21] M. Strassler. (May 31, 2013).The Strengths of the Known Forces. Retrieved January 10, 2018, from https://profmattstrassler.com/articles-and-posts/particle-physics-basics/the-known-forces-of-nature/the-strength-of-the-known-forces/

[22] S.L. Glashow. “Partial symmetries of weak interactions”, Nucl. Phys. 22 579-588, (1961).

[23] A. Salam, J.C. Ward. “Electromagnetic and weak interactions”, Physics Letters 13 168-171, (1964).

[24] S. Weinberg, “A model of leptons”, Physical Review Letters, vol. 19, no. 21, p. 1264, (1967).

[25] M. Peskin, D. Schroeder, “An introduction to quantum field theory”. Perseus Books Publishing L.L.C., (1995).

[26] A. Pich. “The Standard Model of Electroweak Interactions” https://arxiv.org/abs/1201.0537

[27] G. Arnison et al. (UA1 Collaboration), Phys. Lett. B 122, 103 (1983).

[28] M. Banner et al. (UA2 Collaboration), Phys. Lett. B 122, 476 (1983).

[29] G. Arnison et al. (UA1 Collaboration), Phys. Lett. B 126, 398 (1983).

[30] P. Bagnaia et al. (UA2 Collaboration), Phys. Lett. B 129, 130 (1983).

[31] F.Bellaiche. (2012, 2 September). "What's this Higgs boson anyway?". Retrieved from: `https://www.quantum-bits.org/?p=233`

[32] M. Endres et al. Nature 487, 454-458 (2012) doi:10.1038/nature11255

[33] F. Englert, R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". Physical Review Letters. 13 (9): 321-23.(1964) doi:10.1103/PhysRevLett.13.321

[34] P.Higgs. "Broken Symmetries and the Masses of Gauge Bosons". Physical Review Letters. 13 (16): 508-509,(1964). doi:10.1103/PhysRevLett.13.508.

[35] G.Guralnik, C.R. Hagen and T.W.B. Kibble. "Global Conservation Laws and Massless Particles". Physical Review Letters. 13 (20): 585-587, (1964). doi:10.1103/PhysRevLett.13.585.

[36] CMS collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". Physics Letters B. 716 (1): 30-61 (2012). arXiv:1207.7235. doi:10.1016/j.physletb.2012.08.021

[37] ATLAS collaboration. "Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC". Physics Letters B. 716 (1): 1-29 (2012). arXiv:1207.7214. doi:10.1016/j.physletb.2012.08.020.

[38] ATLAS collaboration; CMS collaboration (26 March 2015). “Combined Measurement of the Higgs Boson Mass in pp Collisions at âĹŽs=7 and 8 TeV with the ATLAS and CMS Experiments”. Physical Review Letters. 114 (19): 191803. arXiv:1503.07589. doi:10.1103/PhysRevLett.114.191803.

[39] LHC InternationalMasterclasses“When protons collide”. Retrieved from `http://atlas.physicsmasterclasses.org/en/zpath_protoncollisions.htm`

[40] CMS Collaboration, “SM Higgs Branching Ratios and Total Decay Widths (update in CERN Report4 2016)”. `https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageBR` , last accessed on 17.12.2017.

[41] R.Grant V. “Determination of Higgs branching ratios in $H \to W^+W^- \to l\nu jj$ and $H \to ZZ \to l^+l^-jj$ channels”. Physics Department, University of Tennessee (Dated: October 31, 2012). Retrieved from `http://aesop.phys.utk.edu/ph611/2012/projects/Riley.pdf`

[42] LHC Higgs Cross Section Working Group, Denner, A., Heinemeyer, S. et al. “Standard model Higgs-boson branching ratios with uncertainties”. Eur. Phys. J. C (2011) 71: 1753. `https://doi.org/10.1140/epjc/s10052-011-1753-8`

[43] D. de Florian et al., LHC Higgs Cross Section Working Group, CERNâĂŞ2017âĂŞ002-M, arXiv:1610.07922[hep-ph] (2016).

[44] ATLAS and CMS Collaborations, “Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV,” (2016). CERN-EP-2016-100, ATLAS-HIGG-2015-07, CMS-HIG-15-002.

[45] J. A. Aguilar-Saavedra, R. Benbrik, S. Heinemeyer, and M. Perez-Victoria, “Handbook of vector-like quarks: Mixing and single production”, Phys. Rev. D 88 (2013) 094010, doi:10.1103/PhysRevD.88.094010, arXiv:1306.0572.

[46] A. Greljo, J. F. Kamenik, and J. Kopp, “Disentangling flavor violation in the top-Higgs sector at the LHC”, JHEP 07 (2014) 046, doi:10.1007/JHEP07(2014)046, arXiv:1404.1278.

[47] F. Demartin, F. Maltoni, K. Mawatari, and M. Zaro, “Higgs production in association with a single top quark at the LHC,” European Physical Journal C, vol. 75, p. 267, (2015). doi:10.1140/epj c/s10052-015-3475-9, arXiv:1504.00611.

[48] F. Demartin, B. Maier, F. Maltoni, K. Mawatari, and M. Zaro, “tWH associated production at the LHC”, European Physical Journal C, vol. 77, p. 34, (2017). arXiv:1607.05862

[49] F. Maltoni, K. Paul, T. Stelzer, and S. Willenbrock, “Associated production of Higgs and single top at hadron colliders”, Phys.Rev. D64 (2001) 094023, [hep-ph/0106293].

[50] S. Biswas, E. Gabrielli, F. Margaroli, and B. Mele, “Direct constraints on the top-Higgs coupling from the 8 TeV LHC data,” Journal of High Energy Physics, vol. 07, p. 073, (2013).

[51] M. Farina, C. Grojean, F. Maltoni, E. Salvioni, and A. Thamm, “Lifting degeneracies in Higgs couplings using single top production in association with a Higgs boson,” Journal of High Energy Physics, vol. 05, p. 022, (2013).

[52] T.M. Tait and C.-P. Yuan, “Single top quark production as a window to physics beyond the standard model”, Phys. Rev. D 63 (2000) 014018 [hep-ph/0007298].

[53] CMS Collaboration, “Modelling of the single top-quark production in association with the Higgs boson at 13 TeV.” https://twiki.cern.ch/twiki/bin/viewauth/CMS/SingleTopHiggsGeneration13TeV, last accessed on 16.01.2018.

[54] CMS Collaboration, “SM Higgs production cross sections at $\sqrt{s} =$ 13 TeV.” https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt13TeV, last accessed on 16.01.2018.

[55] S. Dawson, The effective W approximation, Nucl. Phys. B 249 (1985) 42.

[56] S. Biswas, E. Gabrielli and B. Mele, JHEP 1301 (2013) 088 [arXiv:1211.0499 [hep-ph]].

[57] LHC Higgs Cross Section Working Group, “Handbook of LHC Higgs Cross Sections: 4.Deciphering the Nature of the Higgs Sector”, arXiv:1610.07922.

[58] J. Ellis, D. S. Hwang, K. Sakurai, and M. Takeuchi.“Disentangling Higgs-Top Couplings in Associated Production”, JHEP 1404 (2014) 004, [arXiv:1312.5736].

[59] CMS Collaboration, V. 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,” arXiv:1412.8662.

[60] ATLAS Collaboration, G. Aad et al., “Updated coupling measurements of the Higgs boson with the ATLAS detector using up to 25 $\mathrm{fb}^{-}1$ of proton-proton collision data”, ATLAS-CONF-2014-009.

[61] File:Cern-accelerator-complex.svg. Wikimedia Commons, the free media repository. Retrieved January, 2018 from https://commons.wikimedia.org/wiki/File:Cern-accelerator-complex.svg

[62] J.L. Caron , “Layout of the LEP tunnel including future LHC infrastructures.”, (Nov, 1993). A C Collection. Legacy of AC. Pictures from 1992 to 2002. Retrieved from https://cds.cern.ch/record/841542

[63] M. Vretenar, “The radio-frequency quadrupole”. CERN Yellow Report CERN-2013-001, pp.207-223 DOI:10.5170/CERN-2013-001.207. arXiv:1303.6762

[64] L.Evans. P. Bryant (editors). “LHC Machine”. JINST 3 S08001 (2008).

[65] CERN Photographic Service.“Radio-frequency quadrupole, RFQ-1”, March 1983, CERN-AC-8303511. Retrieved from https://cds.cern.ch/record/615852.

[66] CERN Photographic Service “Animation of CERN’s accelerator network”, 14 October 2013. DOI: 10.17181/cds.1610170 Retrieved from https://videos.cern.ch/record/1610170

[67] C.Sutton. “Particle accelerator”.Encyclopedia Britannica. July 17, 2013. Retrieved from https://www.britannica.com/technology/particle-accelerator.

[68] L.Guiraud. “Installation of LHC cavity in vacuum tank.”. July 27 2000. CERN-AC-0007016. Retrieved from https://cds.cern.ch/record/41567.

[69] J.L. Caron, “Magnetic field induced by the LHC dipole’s superconducting coils”. March 1998. AC Collection. Legacy of AC. Pictures from 1992 to 2002. LHC-PHO-1998-325. Retrieved from https://cds.cern.ch/record/841511.

[70] AC Team. “Diagram of an LHC dipole magnet”. June 1999. CERN-DI-9906025 retrieved from https://cds.cern.ch/record/40524.

[71] CMS Collaboration "Public CMS Luminosity Information". https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#2016_proton_proton_13_TeV_collis, last accessed 24.01.2018

[72] J.L Caron. "LHC Layout" AC Collection. Legacy of AC. Pictures from 1992 to 2002. September 1997, LHC-PHO-1997-060. Retrieved from https://cds.cern.ch/record/841573.

[73] J.A. Coarasa. "The CMS Online Cluster:Setup, Operation and Maintenance of an Evolving Cluster". ISGC 2012, 26 February - 2 March 2012, Academia Sinica, Taipei, Taiwan.

[74] CMS Collaboration. "The CMS experiment at the CERN LHC" JINST 3 S08004 (2008).

[75] CMS Collaboration. "CMS detector drawings 2012" CMS-PHO-GEN-2012-002. Retrieved from http://cds.cern.ch/record/1433717.

[76] Davis, Siona Ruth. "Interactive Slice of the CMS detector", Aug. 2016, CMS-OUTREACH-2016-027, retrieved from https://cds.cern.ch/record/2205172

[77] R. Breedon. "View through the CMS detector during the cooldown of the solenoid on February 2006. CMS Collection", February 2006, CMS-PHO-OREACH-2005-004, Retrieved from https://cds.cern.ch/record/930094.

[78] Halyo, V. and LeGresley, P. and Lujan, P. "Massively Parallel Computing and the Search for Jets and Black Holes at the LHC", Nucl.Instrum.Meth. A744 (2014) 54-60, DOI: 10.1016/j.nima.2014.01.038"

[79] A. Dominguez et. al. "CMS Technical Design Report for the Pixel Detector Upgrade", CERN-LHCC-2012-016. CMS-TDR-11.

[80] CMS Collaboration. "Description and performance of track and primary-vertex reconstruction with the CMS tracker," Journal of Instrumentation, vol. 9, no. 10, p. P10009,(2014).

[81] CMS Collaboration and M. Brice. "Images of the CMS Tracker Inner Barrel", November 2008, CMS-PHO-TRACKER-2008-002. Retrieved from https://cds.cern.ch/record/1431467.

[82] M. Weber. "The CMS tracker". 6th international conference on hyperons, charm and beauty hadrons Chicago, June 28-July 3 2004.

[83] CMS Collaboration. "Projected Performance of an Upgraded CMS Detector at the LHC and HL-LHC: Contribution to the Snowmass Process". Jul 26, 2013. arXiv:1307.7135

[84] L. Veillet. "End assembly of HB with EB rails and rotation inside SX ",January 2002. CMS-PHO-HCAL-2002-002. Retrieved from https://cds.cern.ch/record/42594.

[85] J. Puerta-Pelayo."First DT+RPC chambers installation round in the UX5 cavern.". January 2007, CMS-PHO-OREACH-2007-001. Retrieved from https://cds.cern.ch/record/1019185

[86] X. Cid Vidal and R. Cid Manzano. "CMS Global Muon Trigger" web site: Taking a closer look at LHC. Retrieved from https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.lhc_trigger

[87] WLCG Project Office, “Documents & Reference - Tiers - Structure,” (2014). http://wlcg.web.cern.ch/documents-reference , last accessed on 30.01.2018.

[88] CMS Collaboration. “CMSSW Application Framework”, https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookCMSSWFramework, last accesses 06.02.2018

[89] A. Buckleya, J. Butterworthb, S. Giesekec, et. al. “General-purpose event generators for LHC physics”. arXiv:1101.2599v1 [hep-ph] 13 Jan 2011

[90] A. Quadt. “Top Quark Physics at Hadron Colliders”. Advances in the Physics of Particles and Nuclei. Springer-Verlag Berlin Heidelberg. DOI: 10.1007/978-3-540-71060-8 (2007)

[91] DurhamHep Data Project, “The Durham HepData Project - PDF Plotter.” http://hepdata.cedar.ac.uk/pdf/pdf3.html , last accessed on 02.02.2018.

[92] G. Altarelli and G. Parisi. “ASYMPTOTIC FREEDOM IN PARTON LANGUAGE”, Nucl.Phys. B126:298 (1977).

[93] Yu.L. Dokshitzer. Sov.Phys. JETP 46:641 (1977)

[94] V.N. Gribov, L.N. Lipatov. “Deep inelastic e p scattering in perturbation theory”, Sov.J.Nucl.Phys. 15:438 (1972)

[95] F. Maltoni, G. Ridolfi, and M. Ubiali, “b-initiated processes at the LHC: a reappraisal,” Journal of High Energy Physics, vol. 07, p. 022, (2012).

[96] B. Andersson, G. Gustafson, G.Ingelman and T. Sjostrand, “Parton fragmentation and string dynamics”, Physics Reports, Vol. 97, No. 2-3, pp. 31-145, 1983.

[97] CMS Collaboration, “Event generator tunes obtained from underlying event and multiparton scattering measurements;” European Physical Journal C, vol. 76, no. 3, p. 155, (2016).

[98] J. Alwall et. al., “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations,” Journal of High Energy Physics, vol. 07, p. 079, (2014).

[99] T. SjÃűstrand and P. Z. Skands, “Transverse-momentum-ordered showers and interleaved multiple interactions,” European Physical Journal C, vol. 39, pp. 129âĂŞ154, (2005).

[100] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method,” Journal of High Energy Physics, vol. 11, p. 070, (2007).

[101] S. Agostinelli et al., “GEANT4: A Simulation toolkit,” Nuclear Instruments and Methods in Physics, vol. A506, pp. 250âĂŞ303, (2003).

[102] J.Allison et.al.,“Recent developments in Geant4”, Nuclear Instruments and Methods in Physics Research A 835 (2016) 186-225.

[103] CMS Collaboration “Full Simulation Offline Guide”, `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideSimulation`, last accessed 04.02.2018

[104] A. Giammanco. “The Fast Simulation of the CMS Experiment” J. Phys.: Conf. Ser. 513 022012 (2014)

[105] A.M. Sirunyan et. al. “Particle-flow reconstruction and global event description with the CMS detector”, JINST 12 P10003 (2017) `https://doi.org/10.1088/1748-0221/12/10/P10003`.

[106] The CMS Collaboration. “ Description and performance of track and primary vertex reconstruction with the CMS tracker”. JINST 9 P10009 (2014). doi:10.1088/1748-0221/9/10/P10009

[107] J. Incandela. “Status of the CMS SM Higgs Search” July 4, 2012. Pdf slides. Retrieved from `https://indico.cern.ch/event/197461/contributions/1478917/attachments/290954/406673/CMS_4July2012_Final.pdf`

[108] P. Billoir and S. Qian, “Simultaneous pattern recognition and track fitting by the Kalman filtering method”, Nucl. Instrum. Meth. A 294 219. (1990).

[109] W. Adam, R. Fruhwirth, A. Strandlie and T. Todorov, “Reconstruction of electrons with the Gaussian sum filter in the CMS tracker at LHC”, eConf C 0303241 (2003) TULT009 [physics/0306087].

[110] K. Rose, “Deterministic Annealing for Clustering, Compression, Classification, Regression and related Optimisation Problems”, Proc. IEEE 86 (1998) 2210.

[111] R. Fruhwirth, W. Waltenberger and P. Vanlaer, “ Adaptive Vertex Fitting”, CMS Note 2007-008 (2007).

[112] CMS collaboration, “Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV ”, JINST 7 P10002 2012, [arXiv:1206.4071].

[113] Coco, Victor and Delsart, Pierre-Antoine and Rojo-Chacon, Juan and Soyez, Gregory and Sander, Christian, “Jets and jet algorithms”, Proceedings, HERA and the LHC Workshop Series on the implications of HERA for LHC physics: 2006-2008, pag. 182-204. `http://inspirehep.net/record/866539/files/access.pdf`, (2009), doi:10.3204/DESY-PROC-2009-02/54

[114] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm," Journal of High Energy Physics, vol. 04, p. 063, (2008).

[115] S. Catani, Y. L. Dokshitzer, M. H. Seymour, and B. R. Webber, "Longitudinally invariant K_t clustering algorithms for hadron hadron collisions", Nuclear Physics B, vol. 406, pp. 187âĂŞ224, (1993).

[116] Y.L. Dokshitzer, G.D. Leder, S.Moretti, and B.R. Webber, "Better jet clustering algorithms," Journal of High Energy Physics, vol. 08, p. 001, (1997).

[117] B. Dorney. "Anatomy of a Jet in CMS". Quantum Diaries. June 1st, 2011. Retrieved from `https://www.quantumdiaries.org/2011/06/01/anatomy-of-a-jet-in-cms/`

[118] The CMS Collaboration."Event Displays from the high-energy collisions at 7 TeV", May 2010, CMS-PHO-EVENTS-2010-007, Retrieved from `https://cds.cern.ch/record/1429614`.

[119] The CMS collaboration. "Determination of jet energy calibration and transverse momentum resolution in CMS". JINST 6 P11002 (2011). `http://dx.doi.org/10.1088/1748-0221/6/11/P11002`

[120] The CMS Collaboration, "Introduction to Jet Energy Corrections at CMS.". `https://twiki.cern.ch/twiki/bin/view/CMS/IntroToJEC`, last accessed 10.02.2018.

[121] CMS Collaboration Collaboration. "Identification of b quark jets at the CMS Experiment in the LHC Run 2". Tech. rep. CMS-PAS-BTV-15-001. Geneva: CERN, (2016). `https://cds.cern.ch/record/2138504`.

[122] CMS Collaboration Collaboration. “Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector”. Tech. rep. CMS-PAS-JME16-004. Geneva: CERN, 2016. https://cds.cern.ch/record/2205284.

[123] CMS Collaboration, “New CMS results at Moriond (Electroweak) 2013”, Retrieved from http://cms.web.cern.ch/sites/cms.web.cern.ch/files/styles/large/public/field/image/HIG13004_Event01_0.png?itok=LAwZzPHR

[124] CMS Collaboration, “New CMS results at Moriond (Electroweak) 2013”, Retrieved from http://cms.web.cern.ch/sites/cms.web.cern.ch/files/styles/large/public/field/image/TOP12035_Event01.png?itok=uMdnSqzC

[125] K. Skovpen. “Event displays highlighting the main properties of heavy flavour jets in the CMS Experiment”, Aug 2017, CMS-PHO-EVENTS-2017-006. Retrieved from https://cds.cern.ch/record/2280025.

[126] G. Cowan. “Topics in statistical data analysis for high-energy physics”. arXiv:1012.3589v1

[127] A. Hoecker et al., “TMVA-Toolkit for multivariate data analysis” arXiv:physics/0703039v5 (2009)

[128] L. Lista. “Statistical Methods for Data Analysis in Particle Physics”, 2nd ed. Springer International Publishing. (2017) https://dx.doi.org/10.1007/978-3-319-62840-0

[129] I. Antcheva et al., “ROOT-A C++ framework for petabyte data storage, statistical analysis and visualization ,” Computer Physics Communications, vol. 182, no. 6, pp. 1384âĂŞ1385, (2011).

[130] Y. Coadou. “Boosted decision trees”, ESIPAP, Archamps, 9 February 2016. Lecture. Retrieved from https://indico.cern.ch/event/472305/contributions/1982360/attachments/1224979/1792797/ESIPAP_MVA160208-BDT.pdf

[131] J.H. Friedman. “Greedy function approximation: A gradient boosting machine”. Ann. Statist. Volume 29, Number 5 (2001), 1189-1232. https://projecteuclid.org/download/pdf_1/euclid.aos/1013203451.

[132] W. Verkerke and D. Kirkby, “The RooFit toolkit for data modeling,” arXiv preprint physics, (2003).

[133] CMS Collaboration, “Documentation of the RooStats-based statistics tools for Higgs PAG”. https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideHiggsAnalysisCombinedLimit, last accessed on 08.04.2018.

[134] F. James, M. Roos, “MINUIT: Function minimization and error analysis”. Cern Computer Centre Program Library, Geneve Long Write-up No. D506, 1989

[135] J. Neyman and E. S. Pearson, “On the problem of the most efficient tests of statistical hypotheses”. Springer-Verlag, (1992).

[136] A.L. Read. “Modified frequentist analysis of search results (the CL_s method),” (2000). CERN-OPEN-2000-205.

[137] C. Palmer. “Searches for a Light Higgs with CMS”, CMS-CR-2012-215. https://cds.cern.ch/record/1560435.

[138] A. Wald, “Tests of statistical hypotheses concerning several parameters when the number of observations is large”, Transactions of the American Mathematical society, vol. 54, no. 3, pp. 426âĂŞ482, (1943).

[139] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, “Asymptotic formulae for likelihood-based tests of new physics”, European Physical Journal C, vol. 71, p. 1554, (2011).

[140] S. S. Wilks, “The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses”, Annals of Mathematical Statistics, vol. 9, pp. 60-62, (03, 1938).

[141] B. Hespel, F. Maltoni, and E. Vryonidou, “Higgs and Z boson associated production via gluon fusion in the SM and the 2HDM”, JHEP 06 (2015) 065, `https://dx.doi:10.1007/JHEP06(2015)065`, arXiv:1503.01656.

[142] ATLAS Collaboration, “Measurements of Higgs boson production and couplings in diboson final states with the ATLAS detector at the LHC”, Phys. Lett. B726 (2013) 88-119, doi:10.1016/j.physletb.2014.05.011,10.1016/j.physletb.2013.08.010, arXiv:1307.1427. [Erratum: Phys. Lett.B734,406(2014)].

[143] CMS Collaboration, “Search for the associated production of a Higgs boson with a single top quark in proton-proton collisions at $\sqrt{s} = 8$ TeV”, JHEP 06 (2016) 177,doi:10.1007/JHEP06(2016)177, arXiv:1509.08159.

[144] B. Stieger, C. Jorda Lope et al., “Search for Associated Production of a Single Top Quark and a Higgs Boson in Leptonic Channels”, CMS Analysis Note CMS AN-14-140, 2014.

[145] M. Peruzzi, C. Mueller, B. Stieger et al., "Search for ttH in multilepton final states at $\sqrt{s} = 13$ TeV", CMS Analysis Note CMS AN-16-211, 2016.

[146] CMS Collaboration, "Search for H to bbar in association with a single top quark as a test of Higgs boson couplings at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-HIG-16-019, 2016.

[147] CMS Collaboration, "Search for production of a Higgs boson and a single top quark in multilepton final states in proton collisions at $\sqrt{s} = 13$ TeV", CMS Physics Analysis Summary CMS-PAS-HIG-17-005, 2016.

[148] CMS Collaboration, "PdmV2016Analysis," (2016). `https://twiki.cern.ch/twiki/bin/viewauth/CMS/PdmV2016Analysis#DATA`, last accessed 11.04.2016.

[149] M. Peruzzi, F. Romeo, B. Stieger et al., "Search for ttH in multilepton final1 states at $\sqrt{s} = 13$ TeV", CMS Analysis Note CMS AN-17-029, 2017.

[150] B. Maier, "SingleTopHiggProduction13TeV", February, 2016. `https://twiki.cern.ch/twiki/bin/viewauth/CMS/SingleTopHiggsGeneration13TeV`.

[151] B. WG, "BtagRecommendation80XReReco", February, 2017. `https://twiki.cern.ch/twiki/bin/view/CMS/BtagRecommendation80XReReco`.

[152] CMS Collaboration, "Identification of b quark jets at the CMS Experiment in the LHC Run 2", CMS Physics Analysis Summary CMS-PAS-BTV-15-001, 2016.

[153] CMS Collaboration, "Baseline muon selections for Run-II." `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonIdRun2`, last accessed on 24.02.2018.

[154] G. Petrucciani and C. Botta, “Two step prompt muon identification”, January, 2015. https://indico.cern.ch/event/368007/contribution/2/material/slides/0.pdf.

[155] H. Brun and C. Ochando, “Updated Results on MVA eID with 13 TeV samples”, October, 2014. https://indico.cern.ch/event/298249/contribution/3/material/slides/0.pdf.

[156] K. Rehermann and B. Tweedie, “Efficient Identification of Boosted Semileptonic Top Quarks at the LHC”, JHEP 03 (2011) 059, https://dx.doi:10.1007/JHEP03(2011)059, arXiv:1007.2221.

[157] CMS Collaboration. “Tag and Probe”, https://twiki.cern.ch/twiki/bin/view/CMSPublic/TagAndProbe, last accessed on 02.03.2018.