

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

4 by

5 Jose Andres Monroy Montañez

# A DISSERTATION

7 Presented to the Faculty of

8 The Graduate College at the University of Nebraska

In Partial Fulfilment of Requirements

10 For the Degree of Doctor of Philosophy

11 Major: Physics and Astronomy

12 Under the Supervision of Kenneth Bloom and Aaron Dominguez

13 Lincoln, Nebraska

14 July, 2018

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

18 Jose Andres Monroy Montañez, Ph.D.

19 University of Nebraska, 2018

20 Adviser: Kenneth Bloom and Aaron Dominguez

# **<sup>21</sup> Table of Contents**

<b><sup>22</sup> Table of Contents</b>	<b>iii</b>
<b><sup>23</sup> List of Figures</b>	<b>vi</b>
<b><sup>24</sup> List of Tables</b>	<b>ix</b>
<b><sup>25</sup> 1 INTRODUCTION</b>	<b>1</b>
<b><sup>26</sup> 2 Theoretical approach</b>	<b>2</b>
27     2.1 Introduction . . . . .	2
28     2.2 Standard model of particle physics . . . . .	3
29     2.2.1 Fermions . . . . .	5
30     2.2.1.1 Leptons . . . . .	6
31     2.2.1.2 Quarks . . . . .	8
32     2.2.2 Fundamental interactions . . . . .	13
33     2.2.3 Gauge bosons . . . . .	18
34     2.3 Electroweak unification and the Higgs mechanism . . . . .	20
35     2.3.1 Spontaneous symmetry breaking (SSB) . . . . .	28
36     2.3.2 Higgs mechanism . . . . .	32
37     2.3.3 Masses of the gauge bosons . . . . .	35

38	2.3.4	Masses of the fermions . . . . .	36
39	2.3.5	The Higgs field . . . . .	37
40	2.3.6	Production of Higgs bosons at LHC . . . . .	38
41	2.3.7	Higgs boson decay channels . . . . .	42
42	2.4	Associated production of a Higgs boson and a single Top quark. . . . .	43
43	2.5	The CP-mixing in tH processes . . . . .	47
44	2.6	Experimantal status of the anomalous Higg-fermion coupling. . . . .	52
45	<b>3</b>	<b>The CMS experiment at the LHC</b>	<b>54</b>
46	3.1	Introduction . . . . .	54
47	3.2	The LHC . . . . .	55
48	3.3	The CMS experiment . . . . .	64
49	3.3.1	Coordinate system . . . . .	66
50	3.3.2	Pixels detector . . . . .	68
51	3.3.3	Silicon strip tracker . . . . .	70
52	3.3.4	Electromagnetic calorimeter . . . . .	71
53	3.3.5	Hadronic calorimeter . . . . .	73
54	3.3.6	Superconducting solenoid magnet . . . . .	74
55	3.3.7	Muon system . . . . .	76
56	3.3.8	CMS trigger system . . . . .	77
57	3.3.9	CMS computing . . . . .	79
58	<b>4</b>	<b>Event generation, simulation and reconstruction</b>	<b>83</b>
59	4.1	Event generation . . . . .	84
60	4.2	Monte Carlo Event Generators. . . . .	88
61	4.3	CMS detector simulation. . . . .	89
62	4.4	Event reconstruction. . . . .	91

63	4.4.1	Particle-Flow Algorithm.	92
64	4.4.1.1	Particle identification and reconstruction.	96
65	4.5	MVA methods, NN, BDT, boosting, overtraining, variable ranking	99
66	4.6	statistical inference, likelihood parametrization	99
67	4.7	nuisance parameters	99
68	4.8	exclusion limits	99
69	4.9	asymptotic limits	99
70	<b>5</b>	<b>Search for production of a Higgs boson and a single top quark in multilepton final states in pp collisions at <math>\sqrt{s} = 13</math> TeV</b>	<b>100</b>
71	5.1	Introduction	100
72	5.2	Data and MC Samples	102
73	5.2.1	Full 2016 dataset and MC samples	102
74	5.2.2	Triggers	105
75	5.2.2.1	Trigger efficiency scale factors	105
76	5.3	Object Identification and event selection	106
77	5.3.1	Jets and $b$ tagging	106
78	5.3.2	Lepton selection	107
79	5.3.3	Lepton selection efficiency	108
80	5.4	Background predictions	109
81	5.5	Signal discrimination	110
82	5.5.1	Classifiers response	114
83	5.6	Additional discriminating variables	117
85	<b>Bibliography</b>		<b>118</b>
86	<b>References</b>		<b>120</b>

# <sup>87</sup> List of Figures

88	2.1 Standard model of particle physics. . . . .	4
89	2.2 Transformations between quarks . . . . .	12
90	2.3 Fundamental interactions in nature. . . . .	13
91	2.4 SM interactions diagrams . . . . .	14
92	2.5 Neutral current processes . . . . .	21
93	2.6 Spontaneous symmetry breaking mechanism . . . . .	29
94	2.7 SSB Potential form . . . . .	30
95	2.8 Potential for complex scalar field . . . . .	31
96	2.9 SSB mechanism for complex scalar field . . . . .	32
97	2.10 Proton-Proton collision . . . . .	39
98	2.11 Higgs boson production mechanism Feynman diagrams . . . . .	40
99	2.12 Higgs boson production cross section and decay branching ratios . . . . .	41
100	2.13 Associated Higgs boson production mechanism Feynman diagrams . . . . .	43
101	2.14 Cross section for tHq process as a function of $\kappa_t$ . . . . .	46
102	2.15 Cross section for $tHW$ process as a function of $\kappa_{Htt}$ . . . . .	47
103	2.16 NLO cross section for $tX_0$ and $t\bar{t}X_0$ . . . . .	50
104	2.17 NLO cross section for $tWX_0$ , $t\bar{t}X_0$ . . . . .	51

105	2.18 Two dimentional $\kappa_t$ - $\kappa_V$ plot of the coupling modifiers. ATLAS and CMS 106 combination. . . . .	52
107	3.1 CERN accelerator complex . . . . .	55
108	3.2 LHC protons source. First acceleration stage. . . . .	56
109	3.3 The LINAC2 accelerating system at CERN. . . . .	57
110	3.4 LHC layout and RF cavities module. . . . .	58
111	3.5 LHC dipole magnet. . . . .	60
112	3.6 Integrated luminosity delivered by LHC and recorded by CMS during 2016	62
113	3.7 LHC interaction points . . . . .	63
114	3.8 Multiple $pp$ collision bunch crossing at CMS. . . . .	65
115	3.9 Layout of the CMS detector . . . . .	66
116	3.10 CMS detector coordinate system . . . . .	67
117	3.11 CMS pixel detector schematic view. . . . .	69
118	3.12 SST Schematic view. . . . .	70
119	3.13 CMS ECAL schematic view . . . . .	72
120	3.14 CMS HCAL schematic view . . . . .	74
121	3.15 CMS solenoid magnet . . . . .	75
122	3.16 CMS Muon system schematic view . . . . .	76
123	3.17 CMS Level-1 trigger architecture . . . . .	78
124	3.18 WLCG structure . . . . .	79
125	3.19 Data flow from CMS detector through hardware Tiers . . . . .	82
126	4.1 Event generation process. . . . .	84
127	5.1 The two leading-order diagrams of $tHq$ production. . . . .	101
128	5.2 Input variables to the BDT for signal discrimination normalized. . . . .	111

129	5.3	Input variables to the BDT for signal discrimination not normalized. . . . .	113
130	5.4	BDT inputs as seen by TMVA against $t\bar{t}$ . . . . .	114
131	5.5	BDT inputs as seen by TMVA against $t\bar{t}V$ . . . . .	115
132	5.6	Correlation matrices for the input variables in the TMVA. . . . .	116
133	5.7	MVA classifiers performance. . . . .	116
134	5.8	Additional discriminating variables distributions. . . . .	118

# <sup>135</sup> List of Tables

136	2.1	Fermions of the SM. . . . .	5
137	2.2	Fermion masses. . . . .	6
138	2.3	Leptons properties. . . . .	9
139	2.4	Quarks properties. . . . .	9
140	2.5	Fermion weak isospin and weak hypercharge multiplets. . . . .	11
141	2.6	Fundamental interactions features. . . . .	15
142	2.7	SM gauge bosons. . . . .	20
143	2.8	Higgs boson properties. . . . .	38
144	2.9	Predicted branching ratios for a SM Higgs boson with $m_H = 125 \text{ GeV}/c^2$ . . . . .	42
145	2.10	Predicted SM cross sections for $tH$ production at $\sqrt{s} = 13 \text{ TeV}$ . . . . .	44
146	2.11	Predicted enhancement of the $tHq$ and $tHW$ cross sections at LHC . . . . .	48
147	5.1	Signal samples and their cross section and branching fraction. . . . .	102
148	5.2	$\kappa_V$ and $\kappa_t$ combinations. . . . .	103
149	5.3	List of background samples used in this analysis (CMSSW 80X). . . . .	104
150	5.4	Leading-order $t\bar{t}W$ and $t\bar{t}Z$ samples used in the signal BDT training. . . . .	104
151	5.5	Table of high-level triggers that we consider in the analysis. . . . .	105
152	5.6	Trigger efficiency scale factors and associated uncertainties. . . . .	106
153	5.7	Requirements on each of the three muon selections. . . . .	107

154	5.8 Criteria for each of the three electron selections. . . . .	108
155	5.9 MVA input discriminating variables . . . . .	112
156	5.10 TMVA input variables ranking for BDTA_GRAD method . . . . .	117
157	5.11 TMVA configuration used in the BDT training. . . . .	117
158	5.12 ROC-integral for all the testing cases. . . . .	119

<sup>159</sup> Chapter 1

<sup>160</sup> INTRODUCTION

# <sup>161</sup> Chapter 2

## <sup>162</sup> Theoretical approach

### <sup>163</sup> 2.1 Introduction

<sup>164</sup> The physical description of the universe is a challenge that physicists have faced by  
<sup>165</sup> making theories that refine existing principles and proposing new ones in an attempt  
<sup>166</sup> to embrace emerging facts and phenomena.

<sup>167</sup>

<sup>168</sup> At the end of 1940s Julian Schwinger [1] and Richard P. Feynman [2], based in the  
<sup>169</sup> work of Sin-Itiro Tomonaga [3], developed an electromagnetic theory consistent with  
<sup>170</sup> special relativity and quantum mechanics that describes how matter and light inter-  
<sup>171</sup> act; the so-called “quantum eletrodynamics” (QED) had born.

<sup>172</sup>

<sup>173</sup> QED has become the guide in the development of theories that describe the universe.  
<sup>174</sup> It was the first example of a quantum field theory (QFT), which is the theoretical  
<sup>175</sup> framework for building quantum mechanical models that describes particles and their  
<sup>176</sup> interactions. QFT is composed of a set of mathematical tools that combines classical  
<sup>177</sup> fields, special relativity and quantum mechanics, while keeping the quantum point

178 particles and locality ideas.

179 This chapter gives an overview of the standard model of particle physics, starting  
 180 with a description of the particles and interactions that compose it, followed by a  
 181 description of the electroweak interaction, the Higgs boson and the associated pro-  
 182 duction of Higgs boson and a single top quark ( $tH$ ). The description contained in  
 183 this chapter is based on references [4–6].

## 184 2.2 Standard model of particle physics

185 Particle physics at the fundamental level is modeled in terms of a collection of in-  
 186 teracting particles and fields in a theory known as the “standard model of particle  
 187 physics (SM)”<sup>1</sup>.

188

189 The full picture of the SM is composed of three fields<sup>2</sup>, whose excitations are inter-  
 190 preted as particles called mediators or force-carriers; a set of fields, whose excitations  
 191 are interpreted as elementary particles, interacting through the exchange of those  
 192 mediators and a field that gives the mass to elementary particles. Figure 2.1 shows  
 193 an scheme of the SM particles organization. In addition to the particles in the scheme  
 194 (but not listed in it), their corresponding anti-particles, with opposite quantum num-  
 195 bers, are also part of the picture; some particles are their own anti-particles, like  
 196 photon or Higgs, or anti-particle is already listed like in the  $W^+$  and  $W^-$  case.

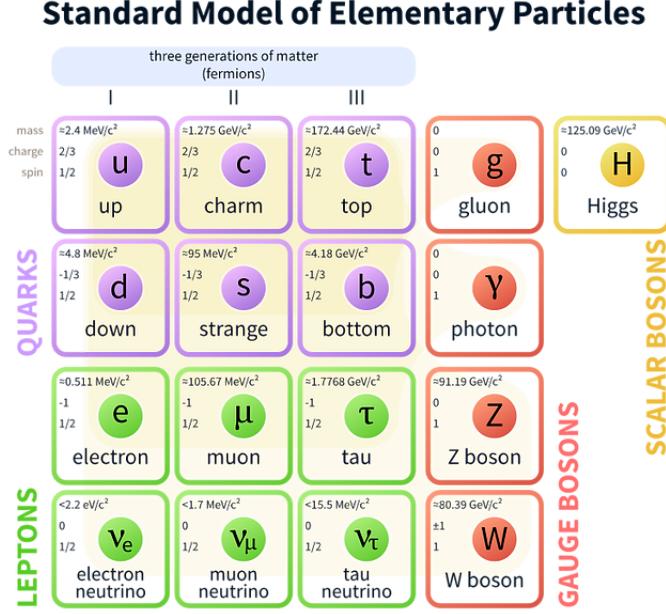
197

198 The mathematical formulation of the SM is based on group theory and the use of  
 199 Noether’s theorem [8] which states that for a physical system modeled by a Lagrangian

---

<sup>1</sup> 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 [?] is quite comprehensive and detailed.

<sup>2</sup> Note that gravitational field is not included in the standard model formulation



**Figure 2.1:** Schematic representation of the Standard model of particle physics. 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].

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 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 of conservation of electric charge is established.

207

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 2.2.2): strong interaction(SI), weak interaction(WI) and electromagnetic interactions (EI) in terms of symmetries associated to physical quantities:

- 212     • Strong:  $SU(3)_C$  associated to color charge
- 213     • Weak:  $SU(2)_L$  associated to weak isospin and chirality
- 214     • Electromagnetic:  $U(1)_Y$  associated to weak hypercharge and electric charge
- 215   It will be shown that the electromagnetic and weak interactions are combined in  
 216   the so-called electroweak interaction where chirality, hypercharge, weak isospin and  
 217   electric charge are the central concepts.

218   **2.2.1 Fermions**

219   The basic constituents of the ordinary matter at the lowest level, which form the set  
 220   of elementary particles in the SM formulation, are quarks and leptons. All of them  
 221   have spin 1/2, therefore they are classified as fermions since they obey Fermi-Dirac  
 222   statistics. There are six “flavors” of quarks and three of leptons organized in three  
 223   generations, or families, as shown in table 2.1.

224

		Generation		
		1st	2nd	3rd
Leptons	Charged	Electron (e)	Moun( $\mu$ )	Tau ( $\tau$ )
	Neutral	Electron neutrino ( $\nu_e$ )	Muon neutrino ( $\nu_\mu$ )	Tau neutrino ( $\nu_\tau$ )
Quarks	Up-type	Up (u)	Charm (c)	Top (t)
	Down-type	Down (d)	Strange (s)	Bottom (b)

**Table 2.1:** 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 WI between leptons is limited to the members of the same generation; WI between quarks is not limited but greatly favoured, to same generation members.

225

226   There is a mass hierarchy between generations (see table 2.2), where the higher gener-  
 227   ation particles decays to the lower one, which can explain why the ordinary matter is

made of particles in 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 <sup>2</sup> )	Quark	Mass (MeV/c <sup>2</sup> )
e	0.51	u	2.2
$\mu$	105.65	c	$1.28 \times 10^3$
$\tau$	1776.86	t	$173.1 \times 10^3$
$\nu_e$	Unknown	d	4.7
$\nu_\mu$	Unknown	s	96
$\tau_\mu$	Unknown	b	$4.18 \times 10^3$

**Table 2.2:** Fermion masses [9]. Generations differ by mass in a way that have been interpreted as a masss 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.

### 2.2.1.1 Leptons

A lepton is an elementary particle that is not subject to the SI. As seen in table 2.1, 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

245 by EI because they don't carry electric charge.

246

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

257

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

263

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

271

272 Each generation of leptons is seen as a weak isospin doublet.<sup>3</sup> The left-handed charged  
 273 lepton and its associated left-handed neutrino are arranged in doublets of weak isospin  
 274 T=1/2 while their right-handed partners are singlets:

$$\begin{pmatrix} \nu_l \\ l \end{pmatrix}_L, l_R := \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, e_R, \mu_R, \tau_R, \nu_{eR}, \nu_{\mu R}, \nu_{\tau R} \quad (2.1)$$

275 The isospin third component refers to the eigenvalues of the weak isospin operator  
 276 which for doublets is  $T_3 = \pm 1/2$ , while for singlets it is  $T_3 = 0$ . The physical meaning  
 277 of this doublet-singlet arrangement falls in that the WI couples the two particles in  
 278 the doublet by exchanging the interaction mediator while the singlet member is not  
 279 involved in WI. The main properties of the leptons are summarized in table 2.3.

280

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

288

### 289 2.2.1.2 Quarks

290 Quarks are the basic constituents of protons and neutrons. The way quarks join to  
 291 form bound states, called “hadrons”, is through the SI. Quarks are affected by all the

---

<sup>3</sup> 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.

Lepton	Q(e)	$T_3$	$L_e$	$L_\mu$	$L_\tau$	Lifetime (s)
Electron (e)	-1	-1/2	1	0	0	Stable
Electron neutrino( $\nu_e$ )	0	1/2	1	0	0	Unknown
Muon ( $\mu$ )	-1	-1/2	0	1	0	$2.19 \times 10^{-6}$
Muon neutrino ( $\nu_\mu$ )	0	1/2	0	1	0	Unknown
Tau ( $\tau$ )	-1	-1/2	0	0	1	$290.3 \times 10^{-15}$
Tau neutrino ( $\tau_\mu$ )	0	1/2	0	0	1	Unknown

**Table 2.3:** Leptons 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.

292 fundamental interactions which means that they carry all the four types of charges:  
 293 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 2.4:** Quarks 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.

294  
 295 Table 2.4 summarizes the features of quarks, among which the most particular is  
 296 their fractional electric charge. Note that fractional charge is not a problem, given  
 297 that quarks are not found isolated, but serves to explain how composed particles are  
 298 formed out of two or more valence quarks<sup>4</sup>.

299

---

<sup>4</sup> 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.

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

- 306     • one quark with a color charge is attracted by an anti-quark with the correspond-  
 307         ing anti-color charge forming a colorless particle called a “meson.”
- 308     • three quarks (anti-quarks) with different color (anti-color) charges are attracted  
 309         among them forming a colorless particle called a “baryon(anti-baryon).”

310 In practice, when a quark is left alone isolated a process called “hadronization” occurs  
 311 where the quark emits gluons (see section 2.2.3) which eventually will generate new  
 312 quark-antiquark pairs and so on; those quarks will recombine to form hadrons that  
 313 will decay into leptons. This proliferation of particles looks like a “jet” coming from  
 314 the isolated quark. More details about the hadronization process and jet structure  
 315 will be given in chapter4.

316 In the first version of the quark model (1964), M. Gell-Mann [12] and G. Zweig  
 317 [13, 14] developed a consistent way to classify hadrons according to their properties.  
 318 Only three quarks (u, d, s) were involved in a scheme in which all baryons have  
 319 baryon number  $B=1$  and therefore quarks have  $B=1/3$ ; non-baryons have  $B=0$ . The  
 320 scheme organizes baryons in a two-dimensional space ( $I_3 - Y$ );  $Y$  (hypercharge) and  $I_3$   
 321 (isospin) are quantum numbers related by the Gell-Mann-Nishijima formula [15, 16]:

$$Q = I_3 + \frac{Y}{2} \quad (2.2)$$

322 where  $Y = B + S + C + T + B'$  are the quantum numbers listed in table 2.4. Baryon  
 323 number is conserved in SI and EI which means that single quarks cannot be created  
 324 but in pairs  $q - \bar{q}$ .

325

326 There are six quark flavors organized in three generations (see table 2.1) following a  
 327 mass hierarchy which, again, implies that higher generations decay to first generation  
 328 quarks.

	Quarks			$T_3$	$Y_W$	Leptons			$T_3$	$Y_W$
Doublets	$(\begin{smallmatrix} u \\ d' \end{smallmatrix})_L$	$(\begin{smallmatrix} c \\ s' \end{smallmatrix})_L$	$(\begin{smallmatrix} t \\ b' \end{smallmatrix})_L$	$(\begin{smallmatrix} 1/2 \\ -1/2 \end{smallmatrix})$	1/3	$(\begin{smallmatrix} \nu_e \\ e \end{smallmatrix})_L$	$(\begin{smallmatrix} \nu_\mu \\ \mu \end{smallmatrix})_L$	$(\begin{smallmatrix} \nu_\tau \\ \tau \end{smallmatrix})_L$	$(\begin{smallmatrix} 1/2 \\ -1/2 \end{smallmatrix})$	-1
Singlets	$u_R$	$c_R$	$t_R$	0	4/3	$\nu_{eR}$	$\nu_{\mu R}$	$\nu_{\tau R}$	0	-2
	$d'_R$	$s'_R$	$b'_R$	0	-2/3	$e_R$	$\mu_R$	$\tau_R$		

**Table 2.5:** Fermion weak isospin and weak hypercharge multiplets. Weak hypercharge is calculated through the Gell-Mann-Nishijima formula 2.2 but using the weak isospin and charge for quarks.

329

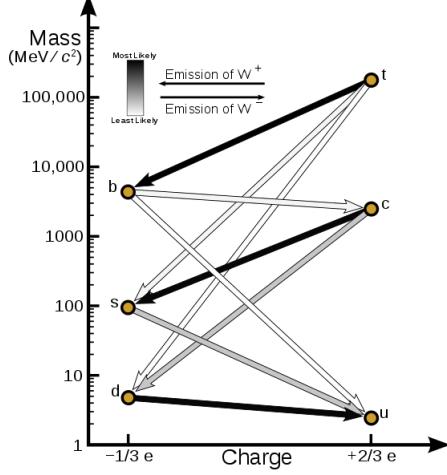
330 Isospin doublets of quarks are also defined (see table 2.5) and as for neutrinos, the  
 331 mass eigenstates are not the same as the WI eigenstates which means that members of  
 332 different quark generations are connected by the WI mediator; thus, up-type quarks  
 333 are coupled not to down-type quarks directly but to a superposition of down-type  
 334 quarks ( $q'_d$ ) via WI according to:

$$q'_d = V_{CKM} q_d$$

335

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (2.3)$$

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



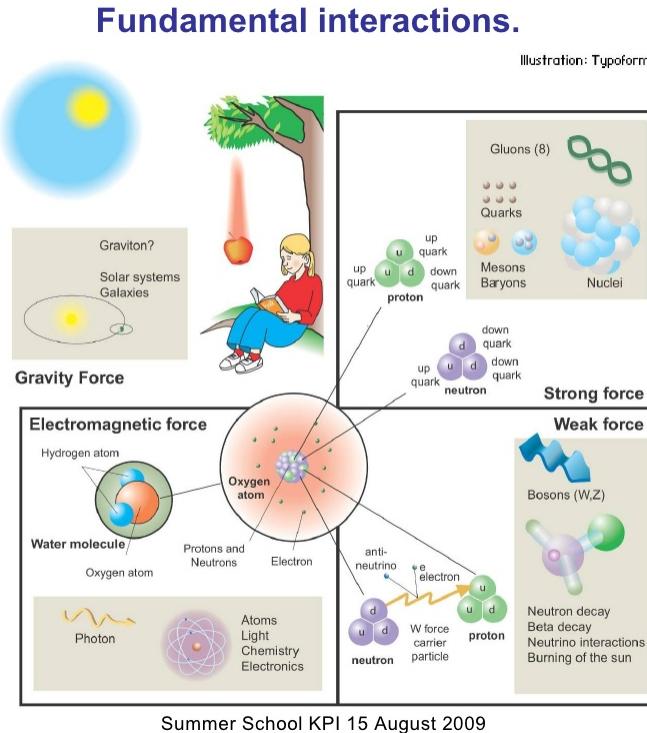
**Figure 2.2:** Transformations between quarks through the exchange of a WI. Higher generations quarks decay to first generation quarks by emitting a W boson. The arrow color indicates the likelihood of the transition according to the grey scale in the top left side which represent the CKM matrix parameters [19].

337 The weak decays of quarks are represented in the diagram of figure 2.2; again the  
 338 CKM matrix plays a central role since it contains the probabilities for the different  
 339 quark decay channels, in particular, note that quark decays are greatly favored be-  
 340 tween generation members.

341

342 CKM matrix is a  $3 \times 3$  unitary matrix parametrized by three mixing angles and  
 343 the *CP-mixing phase*; the latter is the parameter responsible for the Charge-Parity  
 344 symmetry violation (CP-violation) in the SM. The fact that the b quark decays almost  
 345 all the times to a top quark is exploited in this thesis when making the selection of  
 346 the signal events by requiring the presence of a jet tagged as a jet coming from a  
 347 b quark in the final state. The effect of the *CP-mixing phase* on the cross section of  
 348 associated production of Higgs boson and a single top process is also explored in this  
 349 thesis.

350 2.2.2 Fundamental interactions



**Figure 2.3:** 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.

351 Even though there are many manifestations of force in nature, like the ones repre-

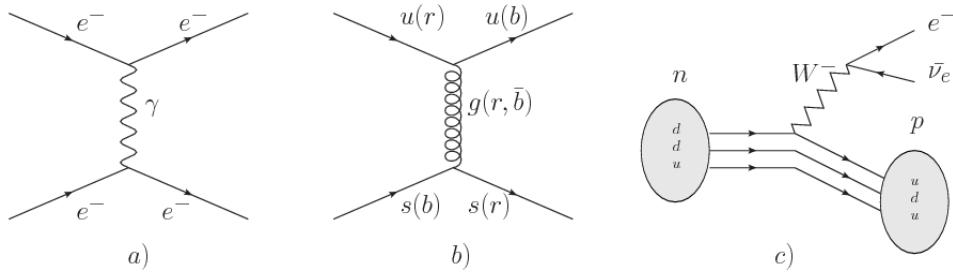
352 sented in figure 2.3, we can classify all of them into one of four fundamental interac-

353 tions:

- 354 • *Electromagnetic interaction (EI)* affects particles that are “electrically charged,”
- 355 like electrons and protons. It is described by QED combining quantum mechan-
- 356 ics, special relativity and electromagnetism in order to explain how particles
- 357 with electric charge interact through the exchange of photons, therefore, one
- 358 says that “Electromagnetic Force” is mediated by “photons”. Figure 2.4a. shows

359 a graphical representation, known as “feynman diagram”, of electron-electron  
 360 scattering.

- 361 • *Strong interaction (SI)* described by Quantum Chromodynamics (QCD). Hadrons  
 362 like proton and neutron have internal structure given that they are composed  
 363 of two or more valence quarks<sup>5</sup>. Quarks have fractional electric charge which  
 364 means that they are subject to electromagnetic interaction and in the case of the  
 365 proton they should break apart due to electrostatic repulsion; however, quarks  
 366 are held together inside the hadrons against their electrostatic repulsion by the  
 367 “Strong Force” through the exchange of “gluons.” The analog to the electric  
 368 charge is the “color charge”. Electrons and photons are elementary particles  
 369 as quarks but they don’t carry color charge, therefore they are not subject to  
 370 SI. The feynman diagram for gluon exchange between quarks is shown in figure  
 371 2.4b.



**Figure 2.4:** Feynman diagrams representing the interactions in SM; a) EI: e-e scattering; b) SI: gluon exchange between quarks ; c) WI:  $\beta$ -decay

- 372 • *Weak interaction (WI)* described by the weak theory (WT), is responsible, for  
 373 instance, for the radioactive decay in atoms and proton-proton (pp) fusion  
 374 within the sun. Quarks and leptons are the particles affected by the weak  
 375 interaction; they possess a property called “flavor charge” (see 2.2.1) which can  
 376 be changed by emitting or absorbing one weak force mediator. There are three

<sup>5</sup> particles made of four and five quarks are exotic states not so common.

377 mediators of the “weak force” known as “Z” boson in the case of electrically  
 378 neutral changes and “ $W^\pm$ ” bosons in the case of electrically charged changes.  
 379 The “weak isospin” is the WI analog to electric charge in EI, and color charge  
 380 in SI, and defines how quarks and leptons are affected by the weak force. Figure  
 381 2.4c. shows the feynman diagram of  $\beta$ -decay where a newtron (n) is transformed  
 382 in a proton (p) by emmiting a  $W^-$  particle. Since this thesis is in the frame  
 383 of the electroweak interaction, a more detailed description of it will be given in  
 384 section 2.3

385 • *Gravitational interaction (GI)* described by General Theory of Relativity (GR).  
 386 It is responsible for the structure of galaxies and black holes as well as the  
 387 expansion of the universe. As a classical theory, in the sense that it can be for-  
 388 mulated without even appeal to the concept of quantization, it implies that the  
 389 spacetime is a continuum and predictions can be made without limitation to the  
 390 precision of the measurement tools. The latter represent a direct contradiction  
 391 of the quantum mechanics principles. Gravity is deterministic while quantum  
 392 mechanics is probabilistic; despite that, efforts to develop a quantum theory of  
 393 gravity have predicted the “graviton” as mediator of the Gravitational force<sup>6</sup>.

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 2.6:** Fundamental interactions features [20].

394

---

<sup>6</sup> Actually a wide variety of theories have been developed in an attempt to describe gravity; some famous examples are string theory and supergravity.

395 Table 2.6 summarizes the main features of the fundamental interactions. The rela-  
 396 tive strength of the fundamental forces reveals the meaning of strong and weak; in  
 397 a context where the relative strength of the SI is 1, the EI is about hundred times  
 398 weaker and WI is about million times weaker than the SI. A good description on  
 399 how the relative strength and range of the fundamental interactions are calculated  
 400 can be found in references [20, 21]. In the everyday life, only EI and GI are explicitly  
 401 experienced due to the range of these interactions; i.e., at the human scale distances  
 402 only EI and GI have appreciable effects, in contrast to SI which at distances greater  
 403 than  $10^{-15}$ m become negligible.

404

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

- 408     • lorentz invariance: independence on the reference frame.
- 409     • locality: interacting fields are evaluated at the same space-time point to avoid  
     410       action at a distance.
- 411     • renormalizability: physical predictions are finite and well defined
- 412     • particle spectrum, symmetries and conservation laws already known must emerge  
     413       from the theory.
- 414     • gauge invariance.

415 The gauge invariance requirement reflects the fact that the fundamental fields cannot  
 416 be directly measured but associated fields which are the observables. Electric (“E”)  
 417 and magnetic (“B”) fields in CED are associated with the electric scalar potential

418 “V” and the vector potential “A”. In particular,  $\mathbf{E}$  can be obtained by measuring  
 419 the change in the space of the scalar potential ( $\Delta V$ ); however, two scalar potentials  
 420 differing by a constant “f” correspond to the same electric field. The same happens in  
 421 the case of the vector potential “A”; thus, different configurations of the associated  
 422 fields result in the same set of values of the observables. The freedom in choosing  
 423 one particular configuration is known as “gauge freedom”; the transformation law con-  
 424 necting two configurations is known as “gauge transformation” and the fact that the  
 425 observables are not affected by a gauge transformation is called “gauge invariance”.

426

427 When the gauge transformation:

$$\begin{aligned} \mathbf{A} &\rightarrow \mathbf{A} - \Delta f \\ V &\rightarrow V - \frac{\partial f}{\partial t} \end{aligned} \tag{2.4}$$

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

433

434 As mentioned in the first lines of section 2.2, QED has one symmetry group ( $U(1)$ )  
 435 with one group generator (the  $Q$  operator) and one gauge field (the electromagnetic  
 436 field  $A^\mu$ ). In CED there is not a clear definition, beyond the historical convention, of  
 437 which fields are the fundamental and which are the associated, but in QED it is clear  
 438 that the fundamental field is  $A^\mu$ . When a gauge theory is quantized, the gauge field

439 is quantized and its quanta is called “gauge boson”. The word boson characterizes  
 440 particles with integer spin which obvey Bose-einstein statistics.

441

442 As will be detailed in section 2.3, interactions between partcles in a system can be  
 443 obtained by considering first the Lagrangian density of free particles in the system,  
 444 which of course is incomplete because the interaction terms have been left out, and  
 445 demanding global phase transformation invariance. Global phase transformation in-  
 446 variance means that a gauge transformation is performed identically to every point  
 447 in the space<sup>7</sup> and the Lagrangian remains invariant. Then, the global transformation  
 448 is promoted to a local phase transformation (this time the gauge transformation de-  
 449 pends on the position in space) and again invariance is required.

450

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

458 This recipe was used to build QED and the theories that aim to explain the funda-  
 459 mental interactions.

### 460 **2.2.3 Gauge bosons**

461 The importance of the gauge bosons comes from the fact that they are the force  
 462 mediators or force carriers. The features of the gauge bosons reflect those of the

---

<sup>7</sup> Here space corresponds to the 4-dimensional space i.e. space-time.

463 fields they represent and they are extracted from the Lagrangian density used to  
 464 describe the interactions. In section 2.3, it will be shown how the gauge bosons of the  
 465 EI and WI emerge from the electroweak Lagrangian. The SI gauge bosons features  
 466 are also extracted from the SI Lagrangian but it is not detailed in this document. The  
 467 main features of the SM gauge bosons will be briefly presented below and summarized  
 468 in table 2.7.

- 469     • **Photon.** EI occurs when the photon couples to (is exchanged between) particles  
     470 carrying electric charge; however, the photon itself does not carry electric charge,  
     471 therefore, there is no coupling between photons. Given that the photon is  
     472 massless the EI is of infinite range, i.e., electrically charged particles interact  
     473 even if they are located far away one from each other; this also implies that  
     474 photons always move with the speed of light.
- 475     • **Gluon.** SI is mediated by gluons which, same as photons, are massless. They  
     476 carry one unit of color charge and one unit of anticolor charge which means that  
     477 gluons couple to other gluons. As a result, the range of the SI is not infinite  
     478 but very short due to the attraction between gluons, giving rise to the “color  
     479 confinement” which explains why color charged particles cannot be isolated but  
     480 live within composited particles, like quarks inside protons.
- 481     • **W, Z.** The WI mediators,  $W^\pm$  and Z, are massive which explains their short-  
     482 range. Given that the WI is the only interaction that can change the flavor  
     483 of the interacting particles, the W boson is the responsible for the nuclear  
     484 transmutation where a neutron is converted in a proton or vice versa with the  
     485 involvement of an electron and a neutrino (see figure 2.4c). The Z boson is the  
     486 responsible of the neutral weak processes like neutrino elastic scattering where

487 no electric charge but momentum transference is involved. WI gauge bosons  
 488 carry isospin charge which makes possible the interaction between them.

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

**Table 2.7:** SM gauge bosons main features [9].

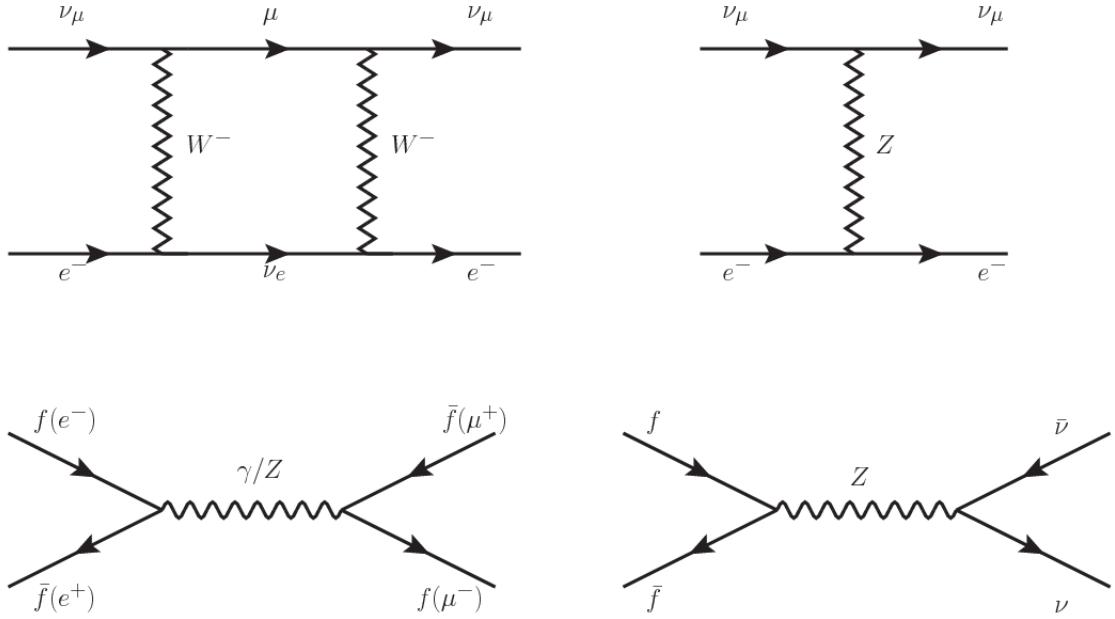
489

## 490 **2.3 Electroweak unification and the Higgs 491 mechanism**

492 Physicists dream of building a theory that contains all the interactions in one single  
 493 interaction, i.e., showing that at some scale in energy all the four fundamental in-  
 494 teractions are unified and only one interaction emerges in a “Theory of everything”.  
 495 The first sign of the feasibility of such unification comes from success in the con-  
 496 struction of the CED. Einstein spent years trying to reach that dream, which by  
 497 1920 only involved electromagnetism and gravity, with no success; however, a new  
 498 partial unification was achieved in the 1960’s, when S.Glashow [22], A.Salam [23] and  
 499 S.Weinberg [24] independently proposed that electromagnetic and weak interactions  
 500 are two manifestations of a more general interaction called “electroweak interaction  
 501 (EWT)”. Both, QCD and EWT, were developed in parallel and following the useful  
 502 prescription provided by QED and the gauge invariance principles.

503

504 The theory of weak interactions was capable of explaining the  $\beta$ -decay and in general  
 505 the processes mediated by  $W^\pm$  bosons. However, there were some processes like the



**Figure 2.5:** 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.

506 “ $\nu_\mu - e$  scattering” which would require the exchange of two  $W$  bosons (see figure 2.5  
507 top diagrams) giving rise to divergent loop integrals and then non finite predictions.  
508 By including neutral currents involving fermions via the exchange of neutral bosons  
509  $Z$ , those divergences are compensated and the predictions become realistic.

510

511 Neutral weak interaction vertices conserve flavor in the same way as the electromag-  
512 netic vertices do, but additionally, the  $Z$  boson can couple to neutrinos which implies  
513 that processes involving charged fermions can proceed through EI or WI but processes  
514 involving neutrinos can proceed only through WI.

515

516 The prescription to build a gauge theory of the WI consists of proposing a free field  
517 Lagrangian density that includes the particles involved; next, by requesting invari-

518 ance under global phase transformations first and generalizing to local phase trans-  
 519 formations invariance later, the conserved currents are identified and interactions are  
 520 generated by introducing gauge fields. Given that the goal is to include the EI and  
 521 WI in a single theory, the group symmetry considered should be a combination of  
 522  $SU(2)_L$  and  $U(1)_{em}$ , however the latter cannot be used directly because the EI treats  
 523 left and right-handed particles indistinctly in contrast to the former. Fortunately, the  
 524 weak hypercharge, which is a combination of the weak isospin and the electric charge  
 525 (eqn 2.2) is suitable to be used since it is conserved by the EI and WI. Thus, the  
 526 symmetry group to be considered is

$$G \equiv SU(2)_L \otimes U(1)_Y \quad (2.5)$$

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

529

530 Given the first generation of leptons

$$\psi_1 = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \quad \psi_2 = \nu_{eR}, \quad \psi_3 = e_R^- \quad (2.6)$$

531 the charged fermionic currents are given by

$$J_\mu \equiv J_\mu^+ = \bar{\nu}_{eL} \gamma_\mu e_L, \quad J_\mu^\dagger \equiv J_\mu^- = \bar{e}_L \gamma_\mu \nu_{eL} \quad (2.7)$$

532 and the free Lagrangian is given by

$$\mathcal{L}_0 = \sum_{j=1}^3 i\bar{\psi}_j(x) \gamma^\mu \partial_\mu \psi_j(x). \quad (2.8)$$

533 Mass terms are included directly in the QED and QCD free Lagrangians since they

534 preserve the invariance under the symmetry transformations involved which treat  
 535 left-handed and right-handed similarly, however mass terms of the form

$$m_W^2 W_\mu^\dagger(x) W^\mu(x) + \frac{1}{2} m_Z^2 Z_\mu(x) Z^\mu(x) - m_e \bar{\psi}_e(x) \psi_e(x) \quad (2.9)$$

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

538

539 Experiments have shown that the gauge fields are not massless; however, they have  
 540 to acquire mass through a mechanism compatible with the gauge invariance; that  
 541 mechanism is known as the “Higgs mechanism” and will be considered later in this  
 542 section. The global transformations in the combined symmetry group G can be  
 543 written as

$$\begin{aligned} \psi_1(x) &\xrightarrow{G} \psi'_1(x) \equiv U_Y U_L \psi_1(x), \\ \psi_2(x) &\xrightarrow{G} \psi'_2(x) \equiv U_Y \psi_2(x), \\ \psi_3(x) &\xrightarrow{G} \psi'_3(x) \equiv U_Y \psi_3(x) \end{aligned} \quad (2.10)$$

544 where  $U_L$  represent the  $SU(2)_L$  transformation acting only on the weak isospin dou-  
 545 blet and  $U_Y$  represent the  $U(1)_Y$  transformation acting on all the weak isospin mul-  
 546 tiplets. Explicitly

$$U_L \equiv \exp\left(i \frac{\sigma_i}{2} \alpha^i\right), \quad U_Y \equiv \exp(i y_i \beta) \quad (i = 1, 2, 3) \quad (2.11)$$

547 with  $\sigma_i$  the Pauli matrices and  $y_i$  the weak hypercharges. In order to promote the  
 548 transformations from global to local while keeping the invariance, it is required that

549  $\alpha^i = \alpha^i(x)$ ,  $\beta = \beta(x)$  and the replacement of the ordinary derivatives by the covariant  
 550 derivatives

$$\begin{aligned} D_\mu \psi_1(x) &\equiv \left[ \partial_\mu + ig\sigma_i W_\mu^i(x)/2 + ig'y_1 B_\mu(x) \right] \psi_1(x) \\ D_\mu \psi_2(x) &\equiv \left[ \partial_\mu + ig'y_2 B_\mu(x) \right] \psi_2(x) \\ D_\mu \psi_3(x) &\equiv \left[ \partial_\mu + ig'y_3 B_\mu(x) \right] \psi_3(x) \end{aligned} \quad (2.12)$$

551 introducing in this way four gauge fields,  $W_\mu^i(x)$  and  $B_\mu(x)$ , in the process. The  
 552 covariant derivatives (eqn 2.12) are required to transform in the same way as fermion  
 553 fields  $\psi_i(x)$  themselves, therefore, the gauge fields transform as:

$$\begin{aligned} B_\mu(x) &\xrightarrow{G} B'_\mu(x) \equiv B_\mu(x) - \frac{1}{g'} \partial_\mu \beta(x) \\ W_\mu^i(x) &\xrightarrow{G} W_\mu^{i\prime}(x) \equiv W_\mu^i(x) - \frac{i}{g} \partial_\mu \alpha_i(x) - \varepsilon_{ijk} \alpha_i(x) W_\mu^j(x). \end{aligned} \quad (2.13)$$

554 The G invariant version of the Lagrangian density 2.8 can be written as

$$\mathcal{L}_0 = \sum_{j=1}^3 i\bar{\psi}_j(x) \gamma^\mu D_\mu \psi_j(x) \quad (2.14)$$

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

$$B_{\mu\nu}(x) \equiv \partial_\mu B_\nu - \partial_\nu B_\mu \quad (2.15)$$

$$W_{\mu\nu}^i(x) \equiv \partial_\mu W_\nu^i(x) - \partial_\nu W_\mu^i(x) - g\varepsilon^{ijk}W_\mu^j W_\nu^k \quad (2.16)$$

558 the last term in eqn. 2.16 is added in order to hold the gauge invariance; therefore,

$$\mathcal{L}_G = -\frac{1}{4}B_{\mu\nu}(x)B^{\mu\nu}(x) - \frac{1}{4}W_{\mu\nu}^i(x)W_i^{\mu\nu}(x) \quad (2.17)$$

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

561

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

$$J_\mu^i(x) = \frac{1}{2}\bar{\psi}_1(x)\gamma_\mu\sigma^i\psi_1(x) \quad (2.18)$$

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

$$J_\mu^Y = \sum_{j=1}^3 \bar{\psi}_j(x)\gamma_\mu y_j\psi_j(x) \quad (2.19)$$

566 In order to evaluate the electroweak interactions modeled by an isos triplet field  $W_\mu^i$   
567 which couples to isospin currents  $J_\mu^i$  with strength  $g$  and additionally the singlet  
568 field  $B_\mu$  which couples to the weak hypercharge current  $J_\mu^Y$  with strength  $g'/2$ . The  
569 interaction Lagrangian density to be considered is

$$\mathcal{L}_I = -g J^{i\mu}(x) W_\mu^i(x) - \frac{g'}{2} J^{Y\mu}(x) B_\mu(x) \quad (2.20)$$

570 Note that the weak isospin currents are not the same as the charged fermionic currents  
 571 that were used to describe the WI (eqn 2.7), since the weak isospin eigenstates are  
 572 not the same as the mass eigenstates, but they are closely related

$$J_\mu = \frac{1}{2}(J_\mu^1 + iJ_\mu^2), \quad J_\mu^\dagger = \frac{1}{2}(J_\mu^1 - iJ_\mu^2). \quad (2.21)$$

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

$$W_\mu^+ = \frac{1}{\sqrt{2}}(W_\mu^1 - iW_\mu^2), \quad W_\mu^- = \frac{1}{\sqrt{2}}(W_\mu^1 + iW_\mu^2). \quad (2.22)$$

575 The fact that there are three weak isospin conserved currents is an indication that in  
 576 addition to the charged fermionic currents, which couple charged to neutral leptons,  
 577 there should be a neutral fermionic current that does not involve electric charge  
 578 exchange; therefore, it couples neutral fermions or fermions of the same electric charge.  
 579 The third weak isospin current contains a term that is similar to the electromagnetic  
 580 current ( $j_\mu^{em}$ ), indicating that there is a relation between them and resembling the  
 581 Gell-Mann-Nishijima formula 2.2 adapted to electroweak interactions

$$Q = T_3 + \frac{Y_W}{2}. \quad (2.23)$$

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

$$j_\mu^{em} = J_\mu^3 + \frac{1}{2} J_\mu^Y. \quad (2.24)$$

585 The neutral gauge fields  $W_\mu^3$  and  $B_\mu$  cannot be directly identified with the  $Z$  and the  
 586 photon fields since the photon interacts similarly with left and right-handed fermions;  
 587 however, they are related through a linear combination given by

$$\begin{aligned} A_\mu &= B_\mu \cos \theta_W + W_\mu^3 \sin \theta_W \\ Z_\mu &= -B_\mu \sin \theta_W + W_\mu^3 \cos \theta_W \end{aligned} \quad (2.25)$$

588 where  $\theta_W$  is known as the “Weinberg angle.” The interaction Lagrangian is now given  
 589 by

$$\mathcal{L}_I = -\frac{g}{\sqrt{2}}(J^\mu W_\mu^+ + J^{\mu\dagger} W_\mu^-) - \left(g \sin \theta_W J_\mu^3 + g' \cos \theta_W \frac{J_\mu^Y}{2}\right) A^\mu - \left(g \cos \theta_W J_\mu^3 - g' \sin \theta_W \frac{J_\mu^Y}{2}\right) Z^\mu \quad (2.26)$$

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

$$g \sin \theta_W = g' \cos \theta_W = e, \quad \frac{g'}{g} = \tan \theta_W \quad (2.27)$$

592 contained in the eqn.2.24; the third term is the neutral weak current.

593

594 Note that the neutral fields transformation given by the eqn. 2.25 can be written in  
 595 terms of the coupling constants  $g$  and  $g'$  as:

$$A_\mu = \frac{g' W_\mu^3 + g B_\mu}{\sqrt{g^2 + g'^2}}, \quad Z_\mu = \frac{g W_\mu^3 - g' B_\mu}{\sqrt{g^2 + g'^2}} \quad (2.28)$$

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

$$\mathcal{L}_{nmEWI} = \mathcal{L}_0 + \mathcal{L}_G \quad (2.29)$$

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

### 602 2.3.1 Spontaneous symmetry breaking (SSB)

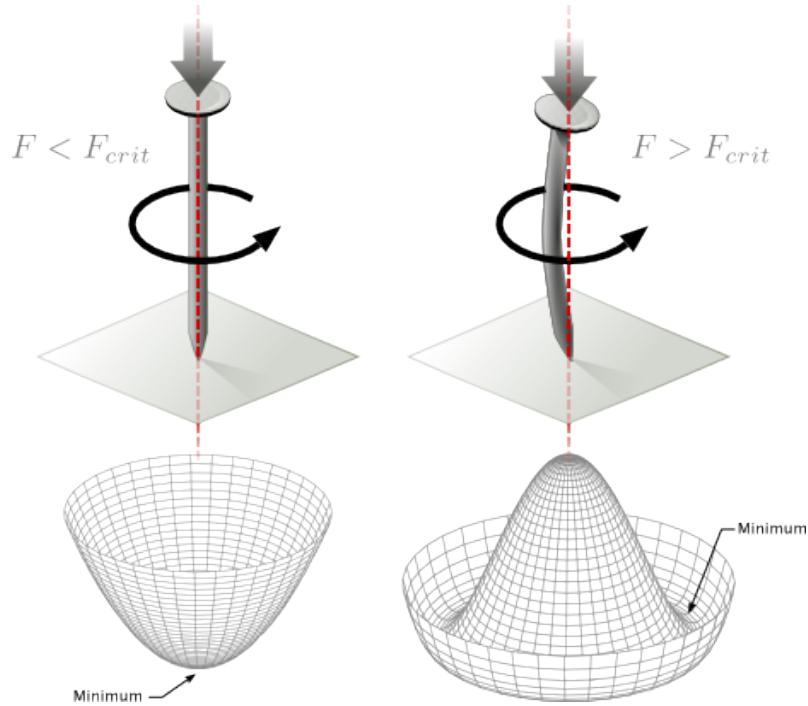
603 Figure 2.6 left shows a steel nail (top) which is subject to an external force; the form  
 604 of the potential energy is also shown (bottom).

605

606 Before reaching the critical force value, the system has rotational symmetry with re-  
 607 spect to the nail axis; however, after the critical force value is reached the nail buckles  
 608 (top right). The form of the potential energy (bottom right) changes, preserving its  
 609 rotational symmetry although its minima does not exhibit that rotational symmetry  
 610 any longer. Right before the nail buckles there is no indication of the direction the  
 611 nail will bend because any of the directions are equivalent, but once the nail bends,  
 612 choosing a direction, an arbitrary minimal energy state (ground state) is selected and  
 613 it does not share the system’s rotational symmetry. This mechanism for reaching an  
 614 asymmetric ground state is known as “*spontaneous symmetry breaking*”.

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

617



**Figure 2.6:** 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 [27].

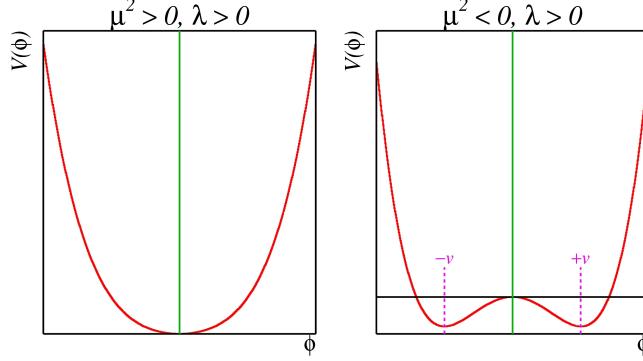
618 Figure 2.7 shows a plot of the potential  $V(\phi)$  in the case of a scalar field  $\phi$

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \quad (2.30)$$

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

622

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



**Figure 2.7:** 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. [27].

$$\phi(x) = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \quad (2.31)$$

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

$$\mathcal{L} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi), \quad V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2 \quad (2.32)$$

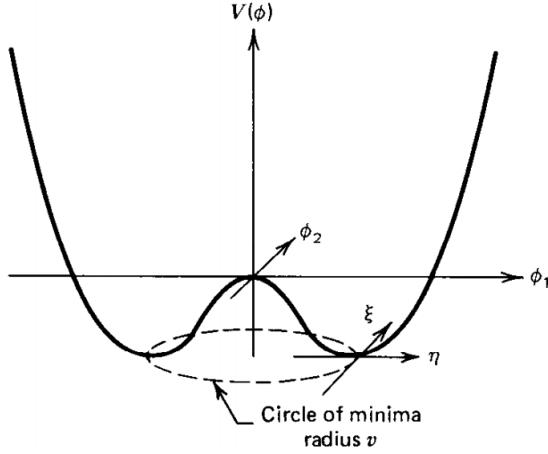
625 where an appropriate potential has been added in order to introduce the SSB.

626

627 As seen in figure 2.8, the potential has now an infinite number of minima circularly  
 628 distributed along the  $\xi$ -direction which makes possible the occurrence of the SSB by  
 629 choosing an arbitrary ground state; for instance,  $\xi = 0$ , i.e.  $\phi_1 = v, \phi_2 = 0$

$$\phi_0 = \frac{v}{\sqrt{2}} \exp(i\xi) \xrightarrow{\text{SSB}} \phi_0 = \frac{v}{\sqrt{2}} \quad (2.33)$$

630 As usual, excitations over the ground state are studied by making an expansion about



**Figure 2.8:** Potential for complex scalar field. There is a circle of minima of radius  $v$  along the  $\xi$ -direction [6].

631 it; thus, the excitation can be parametrized as:

$$\phi(x) = \frac{1}{\sqrt{2}}(v + \eta(x) + i\xi(x)) \quad (2.34)$$

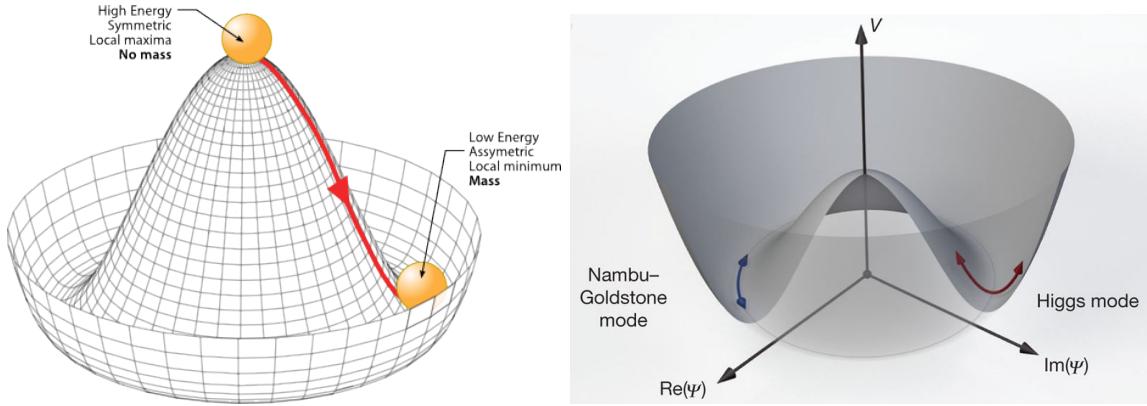
632 which when substituted into eqn. 2.32 produces a Lagrangian in terms of the new  
633 fields  $\eta$  and  $\xi$

$$\mathcal{L}' = \frac{1}{2}(\partial_\mu \xi)^2 + \frac{1}{2}(\partial_\mu \eta)^2 + \mu^2 \eta^2 - V(\phi_0) - \lambda v \eta (\eta^2 + \xi^2) - \frac{\lambda}{4}(\eta^2 + \xi^2)^2 \quad (2.35)$$

634 where the last two terms represent the interactions and self-interaction between the  
635 two fields  $\eta$  and  $\xi$ . The particular feature of the SSB mechanism is revealed when  
636 looking to the first three terms of  $\mathcal{L}'$ . Before the SSB, only the massless  $\phi$  field is  
637 present in the system; after the SSB there are two fields of which the  $\eta$ -field has  
638 acquired mass  $m_\eta = \sqrt{-2\mu^2}$  while the  $\xi$ -field is still massless (see figure 2.9).

639

640 Thus, the SSB mechanism serves as a method to generate mass but as a side effect a



**Figure 2.9:** SSB mechanism for a complex scalar field [27, 28].

641 *massless field is introduced in the system.* This fact is known as the Goldstone theorem  
 642 and states that a massless scalar field appears in the system for each continuous  
 643 symmetry spontaneously broken. Another version of the Goldstone theorem states  
 644 that “*if a Lagrangian is invariant under a continuous symmetry group  $G$ , but the*  
 645 *vacuum is only invariant under a subgroup  $H \subset G$ , then there must exist as many*  
 646 *massless spin-0 particles (Nambu-Goldstone bosons) as broken generators.*” [26] The  
 647 Nambu-Goldstone boson can be understood considering that the potential in the  $\xi$ -  
 648 direction is flat so excitations in that direction are not energy consuming and thus  
 649 represent a massless state.

### 650 2.3.2 Higgs mechanism

651 When the SSB mechanism is introduced in the formulation of the EWI in an attempt  
 652 to generate the mass of the so far massless gauge bosons and fermions, an interesting  
 653 effect is revealed. In order to keep the  $G$  symmetry group invariance and generate  
 654 the mass of the EW gauge bosons, a  $G$  invariant Lagrangian density ( $\mathcal{L}_S$ ) has to be  
 655 added to the non massive EWI Lagrangian (eqn. 2.29)

$$\mathcal{L}_S = (D_\mu \phi)^\dagger (D^\mu \phi) - \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2, \quad \lambda > 0, \mu^2 < 0 \quad (2.36)$$

$$D_\mu \phi = \left( i\partial_\mu - g \frac{\sigma_i}{2} W_\mu^i - g' \frac{Y}{2} B_\mu \right) \phi \quad (2.37)$$

656  $\phi$  has to be an isospin doublet of complex scalar fields so it preserves the G invariance;  
 657 thus  $\phi$  can be defined as:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \quad (2.38)$$

658 The minima of the potential are defined by

$$\phi^\dagger \phi = \frac{1}{2} (\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = -\frac{\mu^2}{2\lambda}. \quad (2.39)$$

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

$$\phi_0 \equiv \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad v^2 \equiv -\frac{\mu^2}{\lambda}. \quad (2.40)$$

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

666

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

668 The next step is to expand  $\phi$  about the chosen ground state as:

$$\phi(x) = \frac{1}{\sqrt{2}} \exp\left(\frac{i}{v} \sigma_i \theta^i(x)\right) \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \approx \frac{1}{\sqrt{2}} \begin{pmatrix} \theta_1(x) + i\theta_2(x) \\ v + H(x) - i\theta_3(x) \end{pmatrix} \quad (2.41)$$

669 to describe fluctuations from the ground state  $\phi_0$ . The fields  $\theta_i(x)$  represent the  
 670 Nambu-Goldstone bosons while  $H(x)$  is known as “higgs field.” The fundamental  
 671 feature of the parametrization used is that the dependence on the  $\theta_i(x)$  fields is  
 672 factored out in a global phase that can be eliminated by taking the physical “unitary  
 673 gauge”  $\theta_i(x) = 0$ . Therefore the expansion about the ground state is given by:

$$\phi(x) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix} \quad (2.42)$$

674 which when substituted into  $\mathcal{L}_S$  (eqn. 2.36) results in a Lagrangian containing the now  
 675 massive three gauge bosons  $W^\pm, Z$ , one massless gauge boson (photon) and the new  
 676 Higgs field ( $H$ ). The three degrees of freedom corresponding to the Nambu-Goldstone  
 677 bosons are now integrated into the massive gauge bosons as their longitudinal po-  
 678 larizations which were not available when they were massless particles. The effect  
 679 by which vector boson fields acquire mass after an spontaneous symmetry breaking,  
 680 but without an explicit gauge invariance breaking is known as the “*Higgs mechanism*”.

681

682 The mechanism was proposed by three independent groups: F.Englert and R.Brout  
 683 in August 1964 [29], P.Higgs in October 1964 [30] and G.Guralnik, C.Hagen and  
 684 T.Kibble in November 1964 [31]; however, its importance was not realized until  
 685 S.Glashow [22], A.Salam [23] and S.Weinberg [24], independently, proposed that elec-  
 686 tromagnetic and weak interactions are two manifestations of a more general interac-  
 687 tion called “electroweak interaction” in 1967.

688 **2.3.3 Masses of the gauge bosons**

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

$$\left| \left( \partial_\mu - ig \frac{\sigma_i}{2} W_\mu^i - i \frac{g'}{2} B_\mu \right) \phi_0 \right|^2 = \left( \frac{1}{2} v g \right)^2 W_\mu^+ W^{-\mu} + \frac{1}{8} v^2 (W_\mu^3, B_\mu) \begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix} \begin{pmatrix} W^{3\mu} \\ B^\mu \end{pmatrix} \quad (2.43)$$

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

$$M_W = \frac{1}{2} v g. \quad (2.44)$$

The second term in the right side of the eqn.2.43 comprises the masses of the neutral bosons, but it needs to be written in terms of the gauge fields  $Z_\mu$  and  $A_\mu$  in order to be compared to the typical mass terms for neutral bosons, therefore using eqn. 2.28

$$\begin{aligned} \frac{1}{8} v^2 [g^2 (W_\mu^3)^2 - 2gg' W_\mu^3 B^\mu + g'^2 B_\mu^2] &= \frac{1}{8} v^2 [g W_\mu^3 - g' B_\mu]^2 + 0[g' W_\mu^3 + g B_\mu]^2 \quad (2.45) \\ &= \frac{1}{8} v^2 [\sqrt{g^2 + g'^2} Z_\mu]^2 + 0[\sqrt{g^2 + g'^2} A_\mu]^2 \end{aligned}$$

692 and then

$$M_Z = \frac{1}{2} v \sqrt{g^2 + g'^2}, \quad M_A = 0 \quad (2.46)$$

### 693 2.3.4 Masses of the fermions

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

$$\mathcal{L}_{Yl} = -G_l \left[ (\bar{\nu}_l, \bar{l})_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} l_R + \bar{l}_R (\phi^-, \bar{\phi}^0) \begin{pmatrix} \nu_l \\ l \end{pmatrix} \right], \quad l = e, \mu, \tau. \quad (2.47)$$

696 After the SSB and replacing the usual field expansion about the ground state (eqn.2.40)  
 697 into  $\mathcal{L}_{Yl}$ , the mass term arises

$$\mathcal{L}_{Yl} = -m_l (\bar{l}_L l_R + \bar{l}_R l_L) - \frac{m_l}{v} (\bar{l}_L l_R + \bar{l}_R l_L) H = -m_l \bar{l} l \left( 1 + \frac{H}{v} \right) \quad (2.48)$$

698 where the additional term represents the lepton-Higgs interaction. The quark masses  
 $m_l = \frac{G_l}{\sqrt{2}} v$  (2.49)

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

$$\phi_c = -i\sigma_2 \phi^* = \begin{pmatrix} -\bar{\phi}^0 \\ \phi^- \end{pmatrix}. \quad (2.50)$$

702 Additionally, given that the quark isospin doublets are not constructed in terms of  
 703 the mass eigenstates but in terms of the flavor eigenstates, as shown in table2.5, the  
 704 coupling parameters will be related to the CKM matrix elements; thus the quark  
 705 Lagrangian is given by:

$$\mathcal{L}_{Yq} = -G_d^{i,j} (\bar{u}_i, \bar{d}'_i)_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} d_{jR} - G_u^{i,j} (\bar{u}_i, \bar{d}'_i)_L \begin{pmatrix} -\bar{\phi}^0 \\ \phi^- \end{pmatrix} u_{jR} + h.c. \quad (2.51)$$

706 with i,j=1,2,3. After SSB and expansion about the ground state, the diagonal form

707 of  $\mathcal{L}_{Yq}$  is:

$$\mathcal{L}_{Yq} = -m_d^i \bar{d}_i d_i \left(1 + \frac{H}{v}\right) - m_u^i \bar{u}_i u_i \left(1 + \frac{H}{v}\right) \quad (2.52)$$

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

### 710 2.3.5 The Higgs field

711 After the characterization of the fermions and gauge bosons as well as their interac-  
712 tions, it is necessary to characterize the Higgs field itself. The Lagrangian  $\mathcal{L}_S$  in eqn.  
713 2.36 written in terms of the gauge bosons is given by

$$\mathcal{L}_S = \frac{1}{4} \lambda v^4 + \mathcal{L}_H + \mathcal{L}_{HV} \quad (2.53)$$

714

$$\mathcal{L}_H = \frac{1}{2} \partial_\mu H \partial^\mu H - \frac{1}{2} m_H^2 H^2 - \frac{1}{2v} m_H^2 H^3 - \frac{1}{8v^2} m_H^2 H^4 \quad (2.54)$$

715

$$\mathcal{L}_{HV} = m_H^2 W_\mu^+ W^{\mu-} \left(1 + \frac{2}{v} H + \frac{2}{v^2} H^2\right) + \frac{1}{2} m_Z^2 Z_\mu Z^\mu \left(1 + \frac{2}{v} H + \frac{2}{v^2} H^2\right) \quad (2.55)$$

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

$$m_H = \sqrt{-2\mu^2} = \sqrt{2\lambda}v \quad (2.56)$$

718 however, it is not predicted by the theory either. The experimental efforts to find  
719 the Higgs boson, carried out by the “Compact Muon Solenoid (CMS)” experiment  
720 and the “A Toroidal LHC AppartuS (ATLAS)” experiments at the “Large Hadron  
721 Collider(LHC)”, gave great results by July of 2012 when the discovery of a new  
722 particle compatible with the Higgs boson predicted by the electroweak theory [32, 33]  
723 was announced. Although at the announcement time there were some reservations  
724 about calling the new particle the “Higgs boson”, today this name is widely accepted.

725 The Higgs mass measurement, reported by both experiments [34], is in table 2.8.

Property	Value
Electric charge	0
Colour charge	0
Spin	0
Weak isospin	-1/2
Weak hypercharge	1
Parity	1
Mass ( $\text{GeV}/c^2$ )	$125.09 \pm 0.21 \text{ (stat.)} \pm 0.11 \text{ (syst.)}$

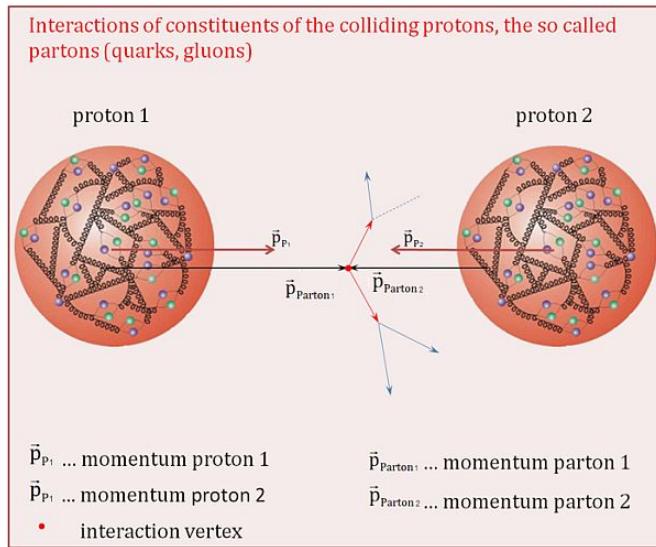
**Table 2.8:** Higgs boson properties. Higgs mass is not predicted by the theory and the value here corresponds to the experimental measurement.

726

### 727 2.3.6 Production of Higgs bosons at LHC

728 At LHC, Higgs boson is produced as a result of the collision of two counter-rotating  
 729 protons beams. A detailed description of the LHC machine will be presented in  
 730 chapter 3. “The total cross section” is a parameter that quantifies the number of pp  
 731 collisions that happen when a number of protons are fired at each other. Different  
 732 results can be obtained after a pp collision and for each one the “cross section” is  
 733 defined as the number of pp collisions that conclude in that particular result with  
 734 respect to the number of protons fired at each other.

735 Protons are composed of quarks and these quarks are bound by gluons; however,  
 736 what is commonly called the quark content of the proton makes reference to the  
 737 valence quarks. A sea of quarks and gluons is also present inside the proton as repre-  
 738 sented in figure 2.10. In a proton-proton (pp) collision, the constituents (quarks and  
 739 gluons) are those who collide. The pp cross section depends on the momentum of  
 740 the colliding particles, reason for which it is needed to know how the momentum is  
 741 distributed inside the proton. Quarks and gluons are known as partons and the func-  
 742 tions that describe how the proton momentum is distributed among partons inside it



**Figure 2.10:** 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. [35].

743 are called “parton distribution functions (PDFs)”; PDFs are determined from experi-

744 mental data obtained in experiments where the internal structure of hadrons is tested.

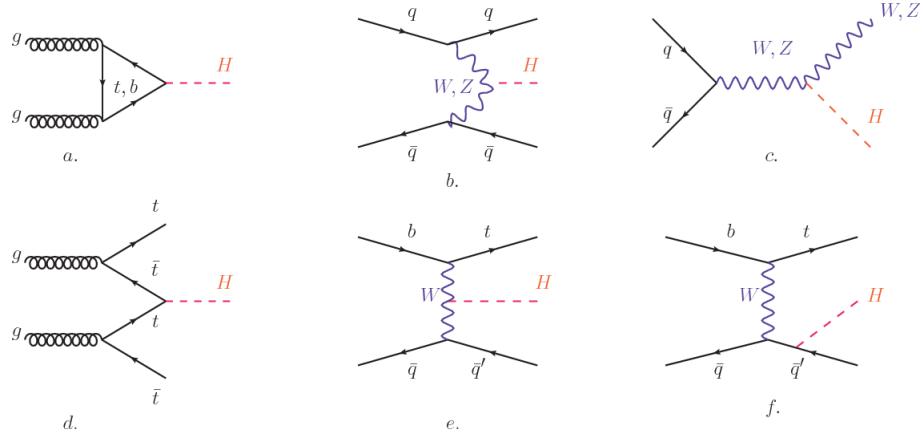
745

In addition, in physics, a common approach to study complex systems consists in starting with a simpler version of them, for which a well known description is available, and add 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; therefore, the more terms are considered in the series (the higher order in the perturbation series), the more precise is the description of the complex system.

754

755 This thesis explores the Higgs production at LHC; therefore the overview presented  
756 here will be oriented specifically to the production mechanisms after pp collisions at

757 LHC.



**Figure 2.11:** 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.

758 Figure 2.11 shows the Feynman diagrams for the leading order (first order) Higgs  
 759 production processes at LHC, while the cross section for Higgs production as a func-

760 tion of the center of mass-energy ( $\sqrt{s}$ ) for pp collisions is showed in figure 2.12 left.

761 The tags NLO (next to leading order), NNLO (next to next to leading order) and

762 N3LO (next to next to next to leading order) make reference to the order at which

763 the perturbation series have been considered.

764 As shown in eqns 2.47, 2.51 and 2.55, the strength of the Higgs-fermion interaction

765 is proportional to the fermion mass while the strength of the Higgs-gauge boson

766 interaction is proportional to the square of the gauge boson mass, which implies

767 that the Higgs production and decay mechanisms are dominated by couplings  $H -$

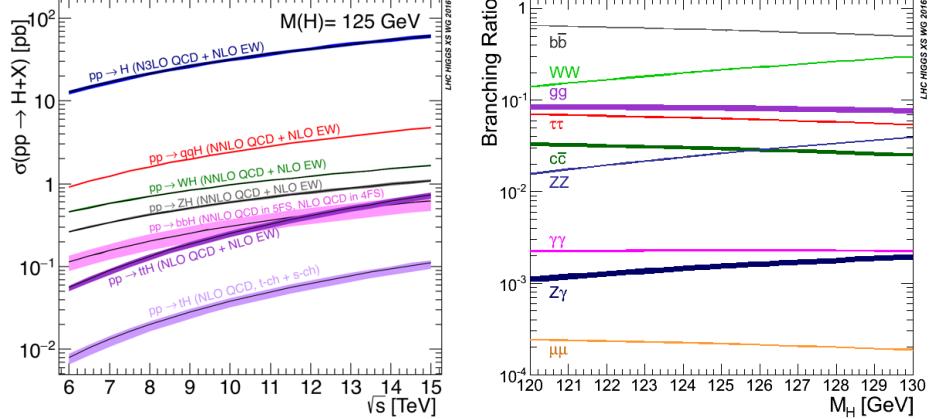
768 ( $W, Z, t, b, \tau$ ).

769 The main production mechanism is the gluon fusion (figure 2.11a and  $pp \rightarrow H$  in figure

770 2.12) given that gluons carry the highest fraction of momentum of the protons in pp

771 colliders. Since the Higgs boson does not couple to gluons, the mechanism proceeds

772 through the exchange of a virtual top-quark loop given that for it the coupling is



**Figure 2.12:** Higgs boson production cross sections (left) and decay branching ratios (right) for the main mechanisms. The VBF is indicated as  $qqH$  [36].

the biggest. Note that in this process, the Higgs boson is produced alone, which makes this mechanism experimentally clean when combined with the two-photon or the four-lepton decay channels (see section 2.3.7).

Vector boson fusion (figure 2.11b and  $pp \rightarrow qqH$  in figure 2.12) 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 a particular region of the detector which is used as a signature when analyzing the datasets provided by the experiments. More details about how to identify events of interest in an analysis will be given in chapter 5.

The next production mechanism is Higgs-strahlung (figure 2.11c and  $pp \rightarrow WH, pp \rightarrow ZH$  in figure 2.12) where two fermions annihilate to form a weak gauge boson. If the initial fermions have enough energy, the emergent boson eventually will emit a Higgs boson.

The associated production with a top or bottom quark pair and the associated production with a single top quark (figure 2.11d-f and  $pp \rightarrow bbH, pp \rightarrow t\bar{t}H, pp \rightarrow tH$  in figure 2.12) have a smaller cross section than the main three mechanisms above,

789 but they provide a good opportunity to test the Higgs-top coupling. The analysis  
 790 reported in this thesis is developed using these production mechanisms. A detailed  
 791 description of the  $tH$  mechanism will be given in section 2.4.

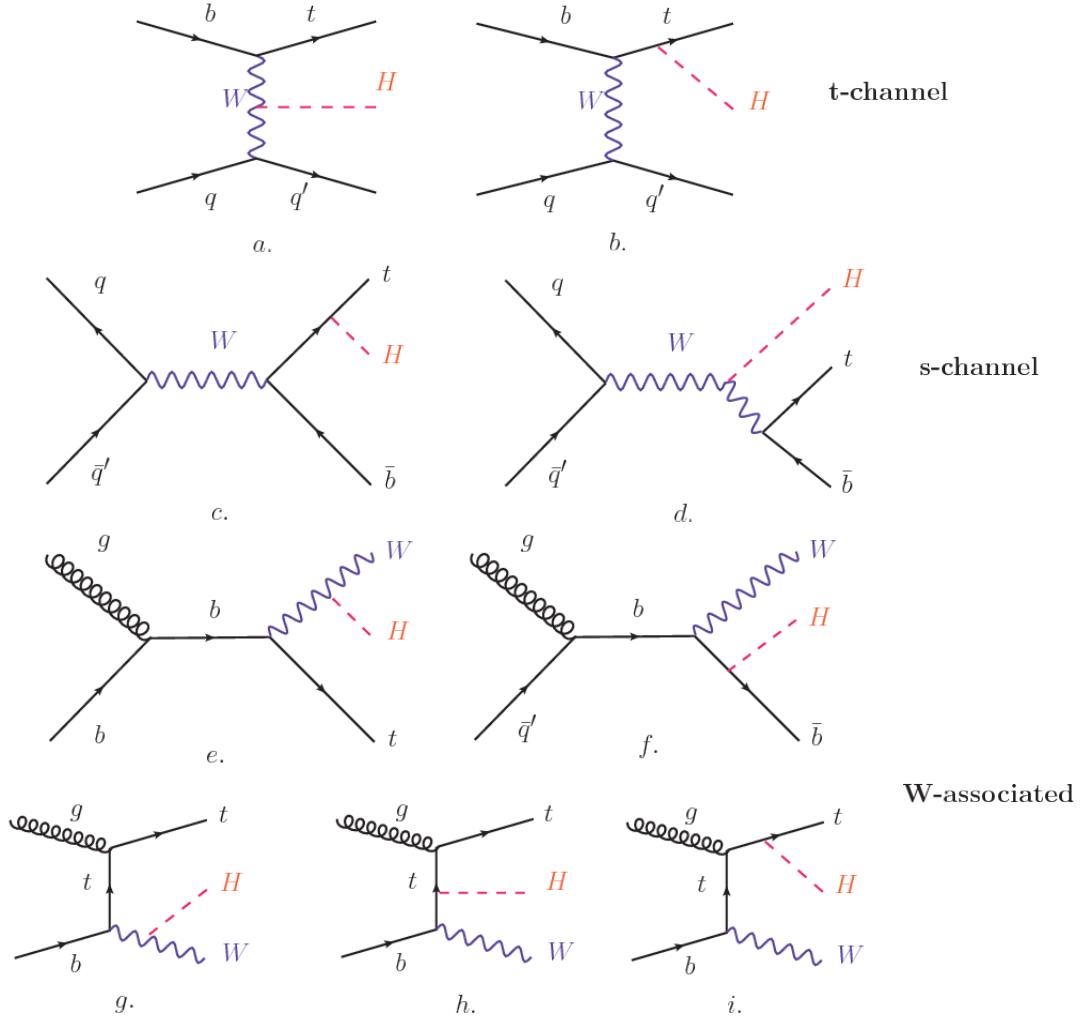
### 792 2.3.7 Higgs boson decay channels

793 When a particle can decay through several modes, also known as channels, the  
 794 probability of decaying through a given channel is quantified by the “branching ratio  
 795 (BR)” of the decay channel; thus, the BR is defined as the ratio of number of decays  
 796 going through that given channel to the total number of decays. In regard to the  
 797 Higgs boson decay, the BR can be predicted with accuracy once the Higgs mass is  
 798 known [37, 38]. In figure 2.12 right, a plot of the BR as a function of the Higgs mass  
 799 is presented. The largest predicted BR corresponds to the  $b\bar{b}$  pair decay channel (see  
 800 table 2.9).

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

**Table 2.9:** Predicted branching ratios and the relative uncertainty for a SM Higgs boson with  $m_H = 125 GeV/c^2$ . [9]

802 **2.4 Associated production of a Higgs boson and a  
803 single Top quark.**



**Figure 2.13:** Associated Higgs boson production mechanism Feynman diagrams. a.,b. t-channel (tHq), c.,d. s-channel (tHb), e-i. W-associated.

804 Associated production of Higgs boson has been extensively studied [39–43]. While  
805 measurements of the main Higgs production mechanisms rates are sensitive to the  
806 strength of the Higgs coupling to W boson or top quark, they are not sensitive to the  
807 relative sign between the two couplings. In this thesis, the Higgs boson production

808 mechanism explored is the associated production with a single top quark (*th*) which  
 809 offers sensitiveness to the relative sign of the Higgs couplings to W boson and to top  
 810 quark. The description given here is based on the reference [41]

811

812 A process where two incoming particles interact and produce a final state with two  
 813 particles can proceed in three ways also called channels (see, for instance, figure 2.13  
 814 ommiting the red line). The t-channel represents processes where an intermediate  
 815 particle is emitted by one of the incoming particles and absorbed by the other. The  
 816 s-channel represents processes where the two incoming particles merge into an inter-  
 817 mediate particle which eventually will split into the particles in the final state. The  
 818 third channel, u-channel, is similar to the t-channel but the two outgoing particles  
 819 interchange their roles.

820

821 The *th* production, where Higgs boson can be radiated either from the top quark or  
 822 from the W boson, is represented by the leading order Feynman diagrams in figure  
 823 2.13. The cross section for the *th* process is calculated, as usual, summing over  
 824 the contributions from the different feynman diagrams; therefore it depends on the  
 825 interference between the contributions. In the SM, the interference for t-channel (tHq  
 826 process) and W-associated (tHW process) production is destructive [39] resulting in  
 827 the small cross sections presented in table 2.10.

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

**Table 2.10:** Predicted SM cross sections for *tH* production at  $\sqrt{s} = 13$  TeV [44, 45].

828

829 While the s-channel contribution can be neglected, it will be shown that a deviation  
 830 from the SM destructive interference would result in an enhancement of the  $th$  cross  
 831 section compared to that in SM, which could be used to get information about the  
 832 sign of the Higgs-top coupling [41, 42]. In order to describe  $th$  production processes,  
 833 Feynman diagram 2.13b will be considered; there, the W boson is radiated by a  
 834 quark in the proton and eventually it will interact with the b quark. In the high  
 835 energy regime, the effective W approximation [46] allows to describe the process as  
 836 the emmision of an approximately on-shell W and its hard scattering with the b  
 837 quark; i.e.  $Wb \rightarrow th$ . The scattering amplitude for the process is given by

$$\mathcal{A} = \frac{g}{\sqrt{2}} \left[ (\kappa_t - \kappa_V) \frac{m_t \sqrt{s}}{m_W v} A \left( \frac{t}{s}, \varphi; \xi_t, \xi_b \right) + \left( \kappa_V \frac{2m_W s}{v} \frac{1}{t} + (2\kappa_t - \kappa_V) \frac{m_t^2}{m_W v} \right) B \left( \frac{t}{s}, \varphi; \xi_t, \xi_b \right) \right], \quad (2.57)$$

838 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 quan-  
 839 tify possible deviations of the couplings, Higgs-Vector boson (H-W) and Higgs-top  
 840 (H-t) respectively, from the SM couplings;  $s = (p_W + p_b)^2$ ,  $t = (p_W - p_H)^2$ ,  $\varphi$  is the  
 841 Higgs azimuthal angle around the  $z$  axis taken parallel to the direction of motion of  
 842 the incoming W; A and B are funtions describing the weak interaction in terms of  
 843 the chiral states of the quarks  $b$  and  $t$ . Terms that vanish in the high energy limit  
 844 have been neglected as well as the Higgs and  $b$  quark masses<sup>8</sup>.

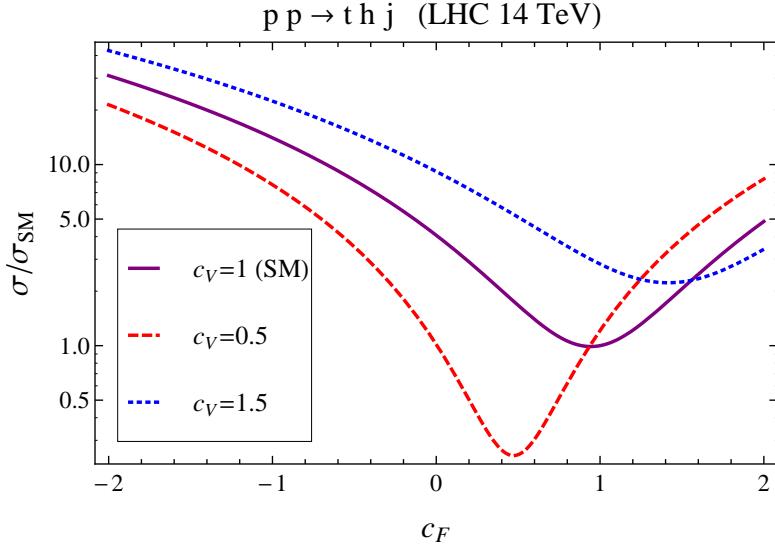
845

846 The scattering amplitude grows with energy like  $\sqrt{s}$  for  $\kappa_V \neq \kappa_t$ , in contrast to  
 847 the SM ( $\kappa_t = \kappa_V = 1$ ), where the first term in 2.57 cancels out and the amplitude  
 848 is constant for large  $s$ ; therefore, a deviation from the SM predictions represents an  
 849 enhancement in the  $tHq$  cross section. In particular, for a SM H-W coupling and a H-t

---

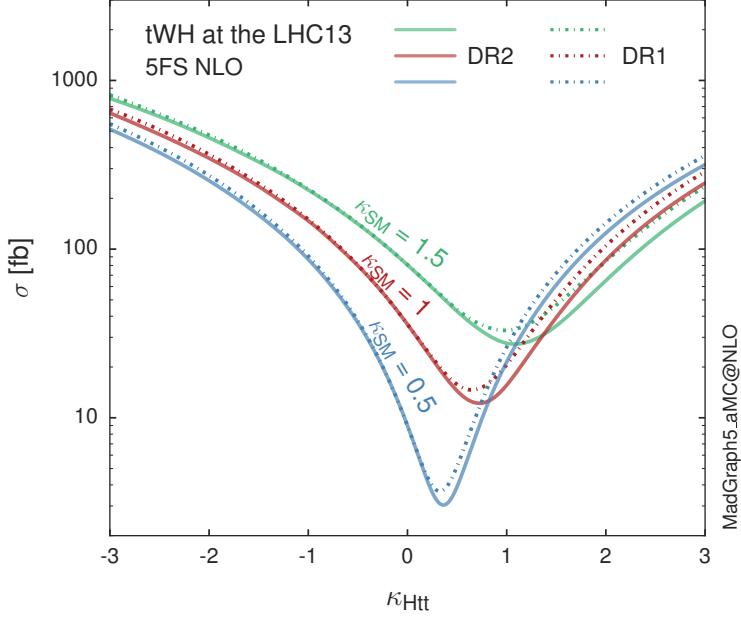
<sup>8</sup> A detailed explanation of the structure and approximations used to derive  $\mathcal{A}$  can be found in reference [41]

850 coupling of inverted sign with respect to the SM ( $\kappa_V = -\kappa_t = 1$ ) the  $tHq$  cross section  
 851 is enhanced by a factor greater 10 as seen in the figure 2.14 taken from reference [41];  
 852 reference [47] has reported similar enhancement results.



**Figure 2.14:** Cross section for  $tHq$  process as a function of  $\kappa_t$ , normalized to the SM, for three values of  $\kappa_V$ . In the plot  $c_f$  refers to the Higgs-fermion coupling which is dominated by the H-t coupling and represented here by  $\kappa_t$ . 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.

853 A similar analysis is valid for the W-associated channel but, in that case, the inter-  
 854 ference is more complicated since there are more than two contributions and an ad-  
 855 ditional interference with the production of Higgs boson and a top pair process( $t\bar{t}H$ ).  
 856 The calculations are made using the so-called Diagram Removal (DR) technique where  
 857 interfering diagrams are removed (or added) from the calculations in order to evaluate  
 858 the impact of the removed contributions. DR1 was defined to neglect  $t\bar{t}H$  interference  
 859 while DR2 was defined to take  $t\bar{t}H$  interference into account [48]. As shown in figure  
 860 2.15, the  $tHW$  cross section is enhanced from about 15 fb (SM:  $\kappa_{Htt} = 1$ ) to about  
 861 150 fb ( $\kappa_{Htt} = -1$ ). Differences between curves for DR1 and DR2 help to gauge the  
 862 impact of the interference with  $t\bar{t}H$ .



**Figure 2.15:** Cross section for  $tHW$  process as a function of  $\kappa_{Htt}$ , for three values of  $\kappa_{SM}$  at  $\sqrt{s} = 13$  TeV.  $\kappa_{Htt}^2 = \sigma_{Htt}/\sigma_{Htt}^{SM}$  is a simple rescaling of the SM Higgs interactions.

863 Results of the calculations of the  $tHq$  and  $tHW$  cross sections at  $\sqrt{s} = 13$  TeV can be  
 864 found in reference [49] and a summary of the results is presented in table 2.11.

865

## 866 2.5 The CP-mixing in tH processes

867 In addition to the sensitivity to sign of the H-t coupling,  $tHq$  and  $tHW$  processes have  
 868 been proposed as a tool to investigate the possibility of a H-t coupling that does not  
 869 conserve CP [43, 48, 50]. Current experimental results are consistent with SM H-V  
 870 and H-t couplings; however, negative H-t coupling is not excluded completely [52].

871

872 In this thesis, the sensitivity of  $th$  processes to CP-mixing is also studied in the  
 873 effective field theory framework and based in references [43, 48]; a generic particle

	$\sqrt{s}$ TeV	$\kappa_t = 1$	$\kappa_t = -1$
$\sigma^{LO}(tHq)(fb)$ [41]	8	$\approx 17.4$	$\approx 252.7$
	14	$\approx 80.4$	$\approx 1042$
$\sigma^{NLO}(tHq)(fb)$ [41]	8	$18.28^{+0.42}_{-0.38}$	$233.8^{+4.6}_{-0.0}$
	14	$88.2^{+1.7}_{-0.0}$	$982.8^{+28}_{-0.0}$
$\sigma^{LO}(tHq)(fb)$ [47]	14	$\approx 71.8$	$\approx 893$
$\sigma^{LO}(tHW)(fb)$ [47]	14	$\approx 16.0$	$\approx 139$
$\sigma^{NLO}(tHq)(fb)$ [49]	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	$21.72^{+6.52\%}_{-5.24\%}$	$\approx 150$
$\sigma^{NLO}DR2(tHW)(fb)$ [48]	13	$16.28^{+7.34\%}_{-15.13\%}$	$\approx 150$

**Table 2.11:** 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 than a factor of 10 is due to the flippling in the sign of the H-t coupling with respect to the SM one.

874 ( $X_0$ ) of spin-0 and a general CP violating interaction with the top quark, can couple  
 875 to scalar and pseudoscalar fermionic densities. The H-W interaction is assumed to  
 876 be SM-like. The Lagrangian modeling the H-t interaction is given by

$$\mathcal{L}_0^t = -\bar{\psi}_t (c_\alpha \kappa_{Htt} g_{Htt} + i s_\alpha \kappa_{Att} g_{Att} \gamma_5) \psi_t X_0, \quad (2.58)$$

877 where  $\alpha$  is the CP-mixing phase,  $c_\alpha \equiv \cos \alpha$  and  $s_\alpha \equiv \sin \alpha$ ,  $\kappa_{Htt}$  and  $\kappa_{Att}$  are real  
 878 dimensionless rescaling parameters<sup>9</sup>,  $g_{Htt} = g_{Att} = m_t/v = y_t/\sqrt{2}$  and  $v \sim 246$  GeV is  
 879 the Higgs vacuum expectation value. In this parametrization, it is easy to recover  
 880 three special cases

881 • CP-even coupling  $\rightarrow \alpha = 0^\circ$

882 • CP-odd coupling  $\rightarrow \alpha = 90^\circ$

883 • SM coupling  $\rightarrow \alpha = 0^\circ$  and  $\kappa_{Htt} = 1$

---

<sup>9</sup> analog to  $\kappa_t$  and  $\kappa_V$

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

$$\mathcal{L}_0^g = -\frac{1}{4} \left( c_\alpha \kappa_{Hgg} g_{Hgg} G_{\mu\nu}^a G^{a,\mu\nu} + s_\alpha \kappa_{Agg} g_{Agg} G_{\mu\nu}^a \tilde{G}^{a,\mu\nu} \right) X_0. \quad (2.59)$$

886 where  $g_{Hgg} = -\alpha_s/3\pi v$  and  $g_{Agg} = \alpha_s/2\pi v$ . Under the assumption that the top quark  
 887 dominates the gluon-fusion process at LHC energies,  $\kappa_{Hgg} \rightarrow \kappa_{Htt}$  and  $\kappa_{Agg} \rightarrow \kappa_{Att}$ ,  
 888 so that the ratio between the gluon-gluon fusion cross section for  $X_0$  and for the SM  
 889 Higgs prediction can be written as

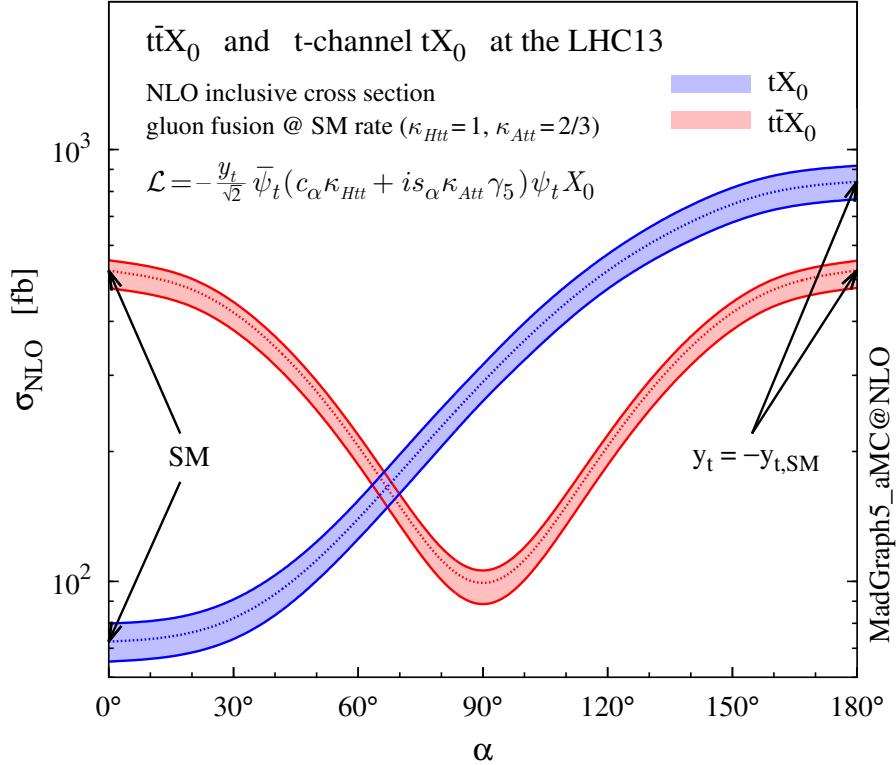
$$\frac{\sigma_{NLO}^{gg \rightarrow X_0}}{\sigma_{NLO,SM}^{gg \rightarrow H}} = c_\alpha^2 \kappa_{Htt}^2 + s_\alpha^2 \left( \kappa_{Att} \frac{g_{Agg}}{g_{Hgg}} \right)^2. \quad (2.60)$$

890 If the rescaling parameters are set to

$$\kappa_{Htt} = 1, \quad \kappa_{Att} = \left| \frac{g_{Hgg}}{g_{Agg}} \right| = \frac{2}{3}. \quad (2.61)$$

891 the gluon-fusion SM cross section is reproduced for every value of the CP-mixing  
 892 angle  $\alpha$ ; therefore, by imposing that condition to the Lagrangian density 2.58, the  
 893 CP-mixing angle is not constrained by current data. Figure 2.16 shows the NLO cross  
 894 sections for t-channel  $tX_0$  (blue) and  $t\bar{t}X_0$  (red) associated production processes as  
 895 a function of the CP-mixing angle  $\alpha$ .  $X_0$  is a generic spin-0 particle with top quark  
 896 CP-violating coupling. Rescaling factors  $\kappa_{Htt}$  and  $\kappa_{Att}$  have been set to reproduce  
 897 the SM gluon-fusion cross sections.

898 It is interesting to notice that the  $tX_0$  cross section is enhanced, by a factor of  
 899 about 10, when a continuous rotation in the scalar-pseudoscalar plane is applied; this  
 900 enhancement is similar to the enhancement produced when the H-t coupling is flipped  
 901 in sign with respect to the SM ( $y_t = -y_{t,SM}$  in the plot), as showed in section 2.4. In



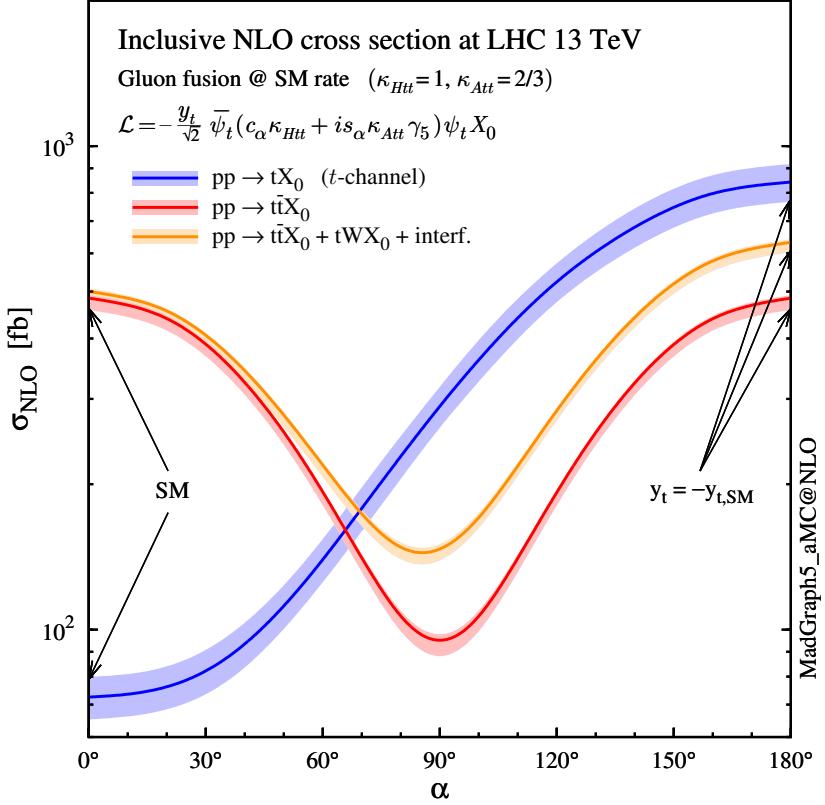
**Figure 2.16:** NLO cross sections for t-channel  $tX_0$ (blue) and  $t\bar{t}X_0$  (red) associated production processes as a function of the CP-mixing angle  $\alpha$ .  $X_0$  is a generic spin-0 particle with top quark CP-violating coupling [43].

902 contrast, the degeneracy in the  $t\bar{t}X_0$  cross section is still present given that it depends  
 903 quadratically on the H-t coupling, but more interesting is to notice that  $t\bar{t}X_0$  cross  
 904 section is exceeded by  $tX_0$  cross section after  $\alpha \sim 60^\circ$ .

905 A similar parametrization can be used to investigate the  $tHW$  process sensitivity to  
 906 CP-violating H-t coupling. As said in 2.4, the interference in the W-associated chan-  
 907 nel is more complicated because there are more than two contributions and also there  
 908 is interference with the  $t\bar{t}H$  production process.

909

910 Figure 2.17 shows the NLO cross sections for t-channel  $tX_0$ (blue),  $t\bar{t}X_0$  (red) asso-  
 911 ciated production and for the combined  $tWX_0 + t\bar{t}X_0 +$  interference (orange) as a

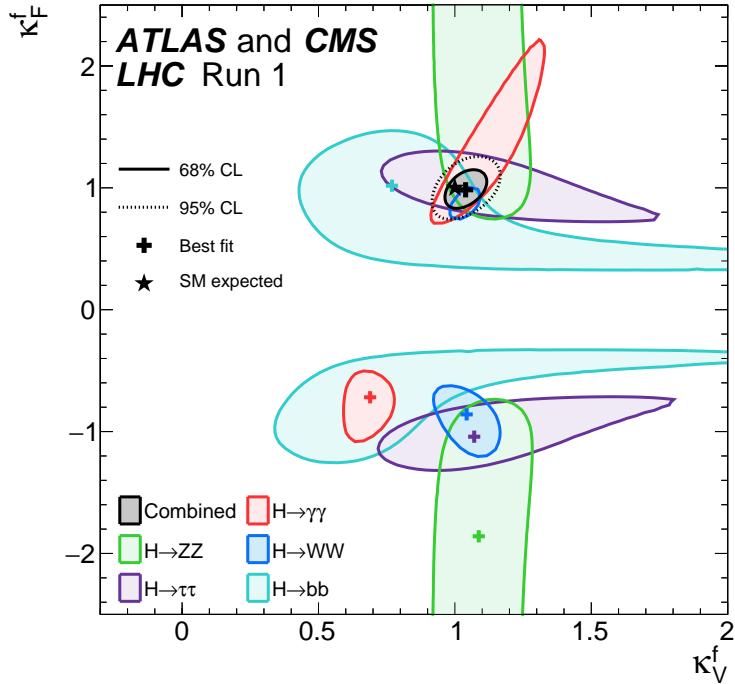


**Figure 2.17:** NLO cross sections for t-channel  $tX_0$  (blue),  $t\bar{t}X_0$  (red) associated production processes and combined  $tWX_0 + t\bar{t}X_0$  (including interference) production as a function of the CP-mixing angle  $\alpha$  [43].

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 ( $\alpha = 180^\circ \rightarrow y_t = -y_{t,SM}$ ).

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

918 **2.6 Experimantal status of the anomalous**  
 919 **Higg-fermion coupling.**



**Figure 2.18:** Combination of the ATLAS and CMS fits for coupling modifiers  $\kappa_t$ - $\kappa_V$ ; also shown the individual decay channels combination and their global combination. No assumptions have been made on the sign of the coupling modifiers [52].

920 ATLAS and CMS have performed analysis of the anomalous H-f coupling by making  
 921 likelihood scans for the two coupling modifiers,  $\kappa_t$  and  $\kappa_V$ , under the assumption that  
 922  $\kappa_Z = \kappa_W = \kappa_V$  and  $\kappa_t = \kappa_\tau = \kappa_b = \kappa_f$ . Figure 2.18 shows the result of the combination  
 923 of ATLAS and CMS fits; also the individual decay channels combination and the  
 924 global combination results are shown.

925 While all the channels are compatible for positive values of the modifiers, for negative  
 926 values of  $\kappa_t$  there is no compatibility. The best fit for individual channels is compatible  
 927 with negative values of  $\kappa_t$  except for the  $H \rightarrow bb$  channel which is expected to be the

928 most sensitive channel; therefore, the best fit for the global fit yields  $\kappa_t \geq 0$ . Thus,  
929 the anomalous H-t coupling cannot be excluded completely.

930 **Chapter 3**

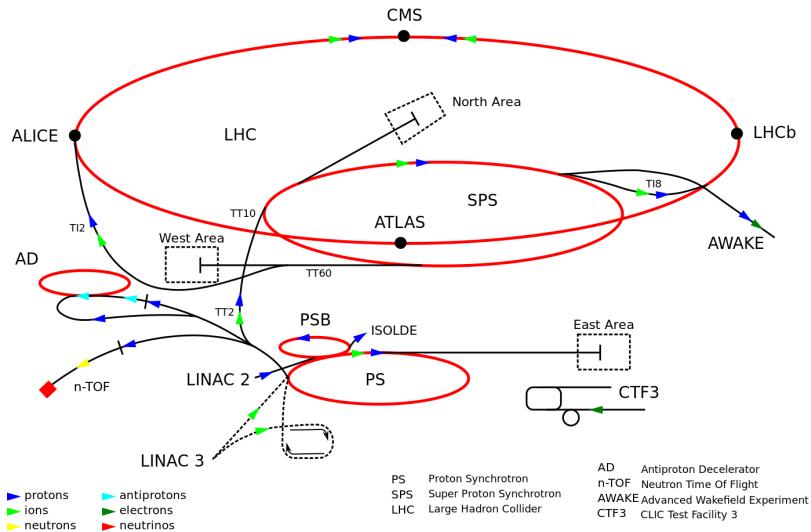
931 **The CMS experiment at the LHC**

932 **3.1 Introduction**

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

## 946 3.2 The LHC

947 With 27 km of circumference, the LHC is currently the largest and most powerful  
 948 circular accelerator in the world. It is installed in the same tunnel where the Large  
 949 Electron-Positron (LEP) collider was located, taking advantage of the existing infras-  
 950 tructure. The LHC is also the larger accelerator in the CERN's accelerator complex  
 951 and is assisted by several successive accelerating stages before the particles are in-  
 952 jected into the LHC ring where they reach their maximum energy (see figure 3.1).



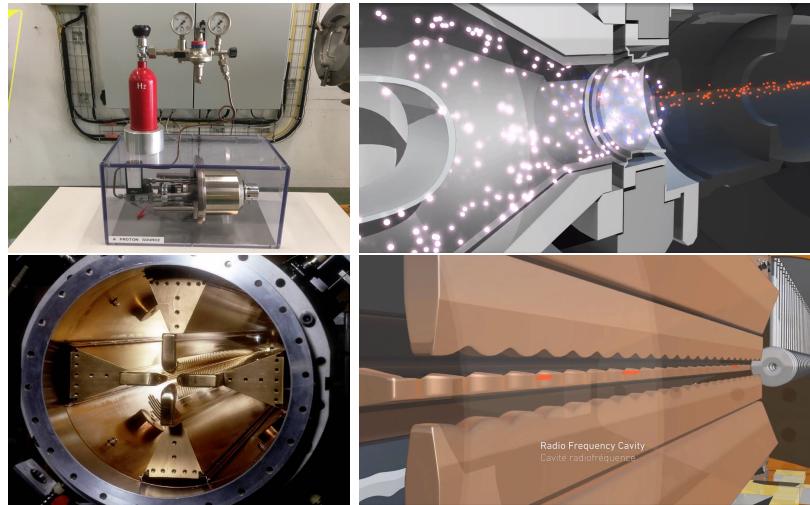
**Figure 3.1:** CERN accelerator complex. Blue arrows show the path followed by protons along the acceleration process [53].

953 LHC runs in three modes depending on the particles being accelerated

- 954 • Proton-Proton collisions ( $pp$ ) for multiple physics experiments.
- 955 • Lead-Lead collisions ( $Pb-Pb$ ) for heavy ion experiments.
- 956 • Proton-Lead collisions ( $p-Pb$ ) for quark-gluon plasma experiments.

957 In this thesis  $pp$  collisions will be considered.

958



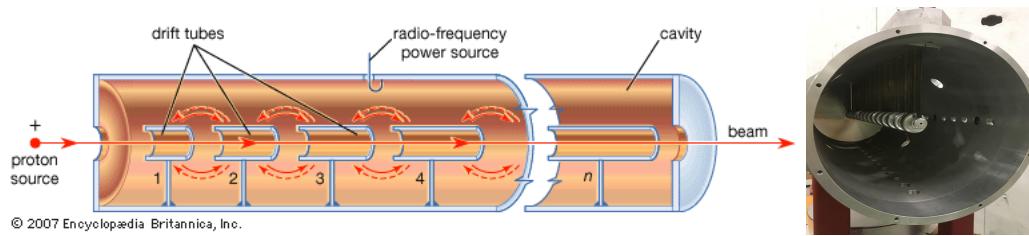
**Figure 3.2:** LHC protons source and the first acceleration stage. Top: the bottle contains hydrogen gas (white dots) which is injected into the metal cylinder to be broken down into electrons(blue dots) and protons(red dots); Bottom: the obtained protons are directed towards the radio frequency quadrupole which perform the first acceleration, focus the beam and create the bunches of protons. [57, 58]

959 Collection of protons starts with hydrogen atoms taken from a bottle, containing hy-  
 960 drogen gas, and injecting them in a metal cylinder; hydrogen atoms are broken down  
 961 into electrons and protons by an intense electric field (see figure3.2 top). The result-  
 962 ing protons leave the metal cylinder towards a radio frequency quadrupole (RFQ)  
 963 that focus the beam, accelerates the protons and creates the packets of protons called  
 964 bunches. In the RFQ, an electric field is generated by a RF wave at a frequency that  
 965 matches the resonance frequency of the cavity where the electrodes are contained.  
 966 The beam of protons traveling on the RFQ axis experiences an alternating electric  
 967 field gradient that generates the focusing forces.

968

969 In order to accelerate the protons, a longitudinal time-varying electric field component  
 970 is added to the system; it is done by giving the electrodes a sinus-like profile as shown  
 971 in figure 3.2 bottom. By matching the speed and phase of the protons with the  
 972 longitudinal electric field the bunching is performed; protons synchronized with the

973 RFQ (synchronous proton) do not feel an accelerating force, but those protons in the  
 974 beam that have more (or less) energy than the synchronous proton (asynchronous  
 975 protons) will feel a decelerating (accelerating) force; therefore, asynchronous protons  
 976 will oscillate around the synchronous ones forming bunches of protons [55]. From the  
 977 RFQ protons emerge with energy 750 keV in bunches of about  $1.15 \times 10^{11}$  protons [56].

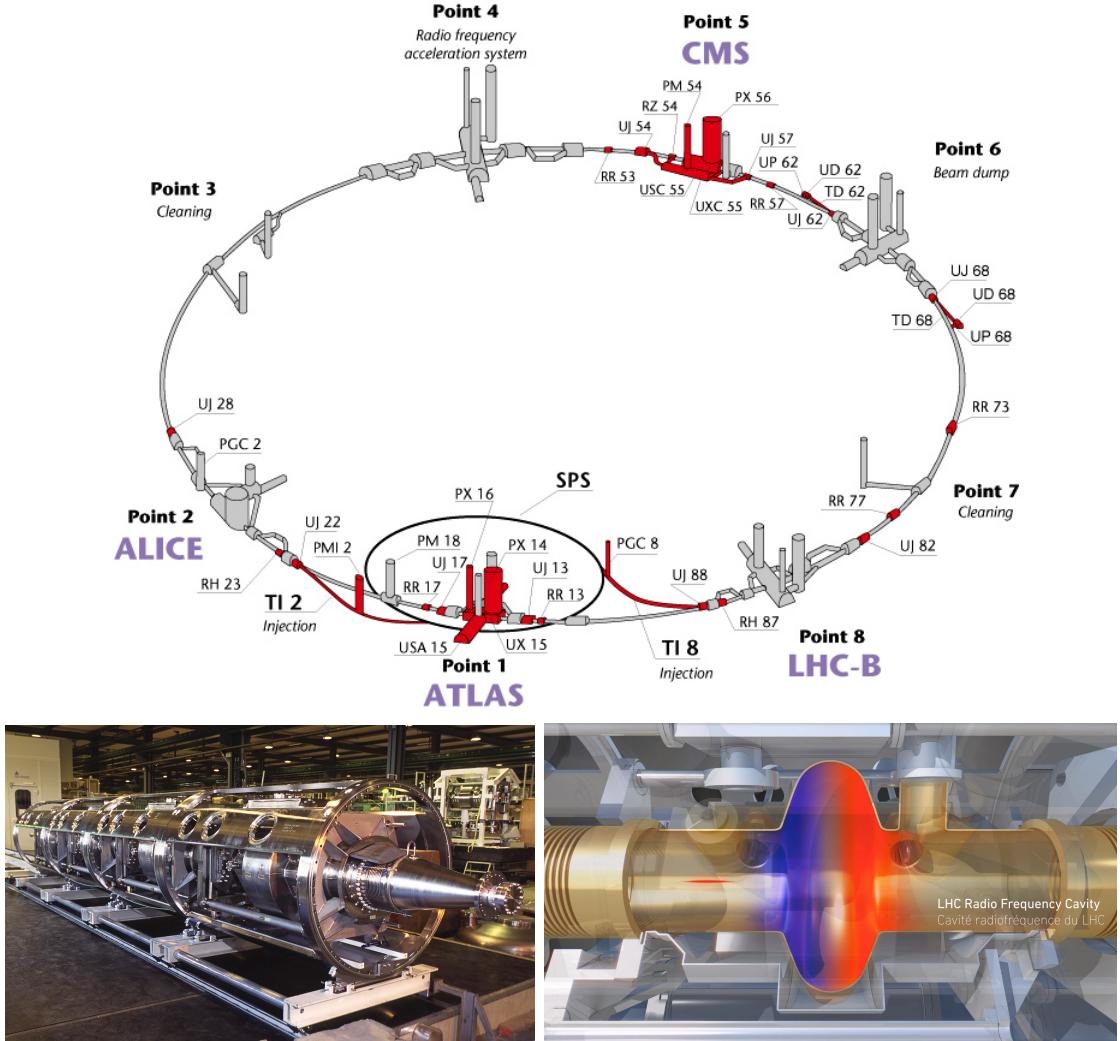


**Figure 3.3:** The LINAC2 accelerating system at CERN. Electric fields generated by radio frequency (RF) create acceleration and deceleration zones inside the cavity; deceleration zones are blocked by drift tubes where quadrupole magnets focus the proton beam. [59]

978 Proton bunches coming from the RFQ go to the linear accelerator 2 (LINAC2) where  
 979 they are accelerated to reach 50 MeV energy. In the LINAC2 stage, acceleration  
 980 is performed using electric fields generated by radio frequency which create zones  
 981 of acceleration and deceleration as shown in figure 3.3. In the deceleration zones,  
 982 the electric field is blocked using drift tubes where protons are free to drift while  
 983 quadrupole magnets focus the beam.

984

985 The beam coming from LINAC2 is injected into the proton synchrotron booster  
 986 (PSB) to reach 1.4 GeV in energy. The next acceleration is provided at the pro-  
 987 ton synchrotron (PS) up to 26 GeV, followed by the injection into the super proton  
 988 synchrotron (SPS) where protons are accelerated to 450 GeV. Finally, protons are  
 989 injected into the LHC where they are accelerated to the target energy of 6.5 TeV.  
 990 PSB, PS, SPS and LHC accelerate protons using the same RF acceleration technique  
 991 described before.



**Figure 3.4:** 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 [54]. Bottom: LHC RF cavities. A module accommodates 4 cavities that accelerate protons and preserve the bunch structure of the beam. [58, 60]

992 LHC has a system of 16 RF cavities located in the so-called point 4, as shown in  
 993 figure 3.4 top, tuned at a frequency of 400 MHz and the protons are carefully timed  
 994 so in addition to the acceleration effect the bunch structure of the beam is preserved.  
 995 Bottom side of figure 3.4 shows a picture of a Rf module composed of 4 RF cavities  
 996 working in a superconducting state at 4.5 K; also is showed a representation of the

997 accelerating electric field that accelerates the protons in the bunch.

998

999 While protons are accelerated in one section of the LHC ring, where the RF cavities  
 1000 are located, in the rest of their path they have to be kept in the curved trajectory  
 1001 defined by the LHC ring. Technically, LHC is not a perfect circle; RF, injection, beam  
 1002 dumping, beam cleaning and sections before and after the experimental points where  
 1003 protons collide are all straight sections. In total, there are 8 arcs 2.45 Km long each  
 1004 and 8 straight sections 545 m long each. In order to curve the proton's trajectory in  
 1005 the arc sections, superconducting dipole magnets are used.

1006

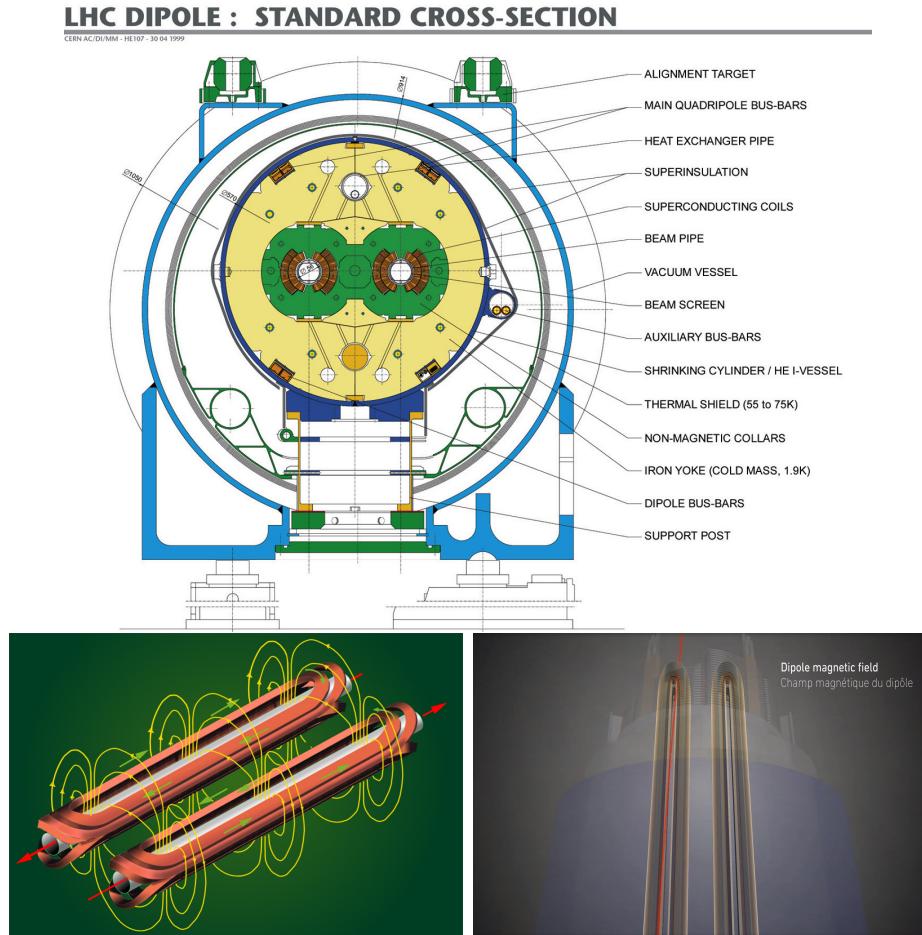
1007 Inside the LHC ring, there are two proton beams traveling in opposite directions in  
 1008 two separated beam pipes; the beam pipes are kept at ultra-high vacuum ( $\sim 10^{-9}$   
 1009 Pa) to ensure that there are no particles that interact with the proton beams. The  
 1010 superconducting dipole magnets used in LHC are made of a NbTi alloy, capable of  
 1011 transporting currents of about 12000 A when cooled at a temperature below 2K using  
 1012 liquid helium (see figure 3.5).

1013

1014 Protons in the arc sections of LHC feel a centripetal force exerted by the dipole  
 1015 magnets which is perpendicular to the beam trajectory; The magnitude of magnetic  
 1016 field needed can be found assuming that protons travel at  $v \approx c$ , using the standard  
 1017 values for proton mass and charge and the LHC radius, as

$$F_m = \frac{mv^2}{r} = qBv \quad \rightarrow B = 8.33T \quad (3.1)$$

1018 which is about 100000 times the Earth's magnetic field. A representation of the mag-  
 1019 netic field generated by the dipole magnets is shown on the bottom left side of figure



**Figure 3.5:** 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 [58, 61, 62].

1020 3.5. The bending effect of the magnetic field on the proton beam is shown on the  
 1021 bottom right side of figure 3.5. Note that the dipole magnets are not curved; the arc  
 1022 section of the LHC ring is composed of straight dipole magnets of about 15 m. In  
 1023 total there are 1232 dipole magnets along the LHC ring.

1024

1025 In addition to bending the beam trajectory, the beam has to be focused so it stays

1026 inside the beam pipe. The focusing is performed by quadrupole magnets installed in  
 1027 a different straight section; in total 858 quadrupole magnets are installed along the  
 1028 LHC ring. Other effects like electromagnetic interaction among bunches, interaction  
 1029 with electron clouds from the beam pipe, the gravitational force on the protons, dif-  
 1030 ferences in energy among protons in the same bunch, among others, are corrected  
 1031 using sextupole and other magnetic multipoles.

1032

1033 The two proton beams inside the LHC ring are made of bunches with a cylindrical  
 1034 shape of about 7.5 cm long and about 1 mm in diameter; when bunches are close  
 1035 to the collision point (CP), the beam is focused up to a diameter of about 16  $\mu\text{m}$  in  
 1036 order to maximize the number of collisions per unit area and per second, known as  
 1037 luminosity ( $L$ ). Luminosity can be calculated using

$$L = fn \frac{N_1 N_2}{4\pi\sigma_x\sigma_y} \quad (3.2)$$

1038 where  $f$  is the revolution frequency,  $n$  is the number of bunches per beam,  $N_1$  and  
 1039  $N_2$  are the numbers of protons per bunch ( $1.5 \times 10^{11}$ ),  $\sigma_x$  and  $\sigma_y$  are the gaussian  
 1040 transverse sizes of the bunches. The expected luminosity about

$$f = \frac{v}{2\pi r_{LHC}} \approx \frac{3 \times 10^8 \text{ m/s}}{27 \text{ km}} \approx 11.1 \text{ kHz},$$

$$n = 2808$$

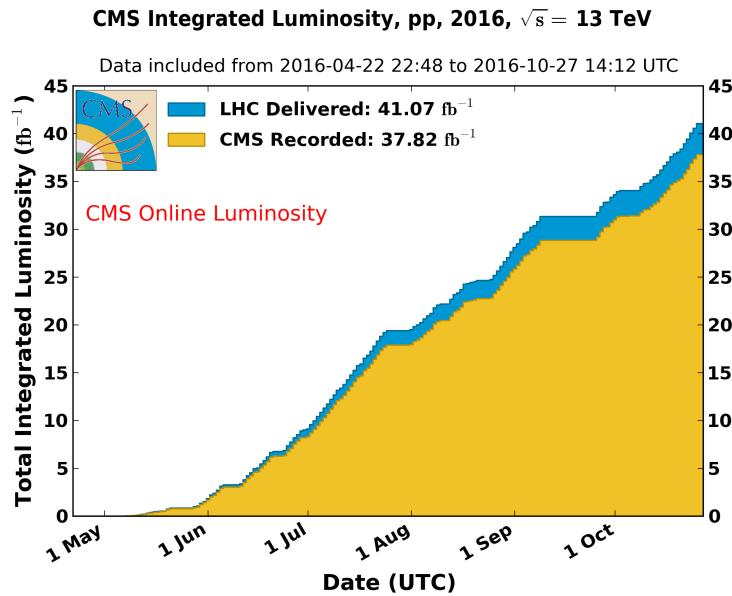
$$N_1 = N_2 = 1.5 \times 10^{11}$$

$$\sigma_x = \sigma_y = 16 \mu\text{m}$$

1041

$$L = 1.28 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1} = 1.28 \times 10^{-5} \text{ fb}^{-1} \text{ s}^{-1} \quad (3.3)$$

1042 Luminosity is a fundamental aspect of LHC given that the bigger luminosity the bigger  
 1043 number of collisions, which means that for processes with a very small cross section  
 1044 the number of expected occurrences is increased and so the chances of being detected.  
 1045 The integrated luminosity, i.e. the total luminosity, collected by the CMS experiment  
 1046 during 2016 is shown in figure 3.6; the data analyzed in this thesis corresponds to an  
 1047 integrated luminosity of  $35.9 \text{ fb}^{-1}$  at a center of mass-energy  $\sqrt{s} = 13 \text{ TeV}$ .  
 1048 A way to increase  $L$  is increasing the number of bunches in the beam. Currently, the  
 1049 separation between two consecutive bunches in the beam is 7.5 m which corresponds  
 1050 to a time separation of 25 ns. In the full LHC ring the allowed number of bunches is  
 1051  $n = 27\text{km}/7.5\text{m} = 3600$ ; however, there are some gaps in the bunch pattern intended  
 1052 for preparing the dumping and injection of the beam, thus, the proton beams are  
 1053 composed of 2808 bunches.

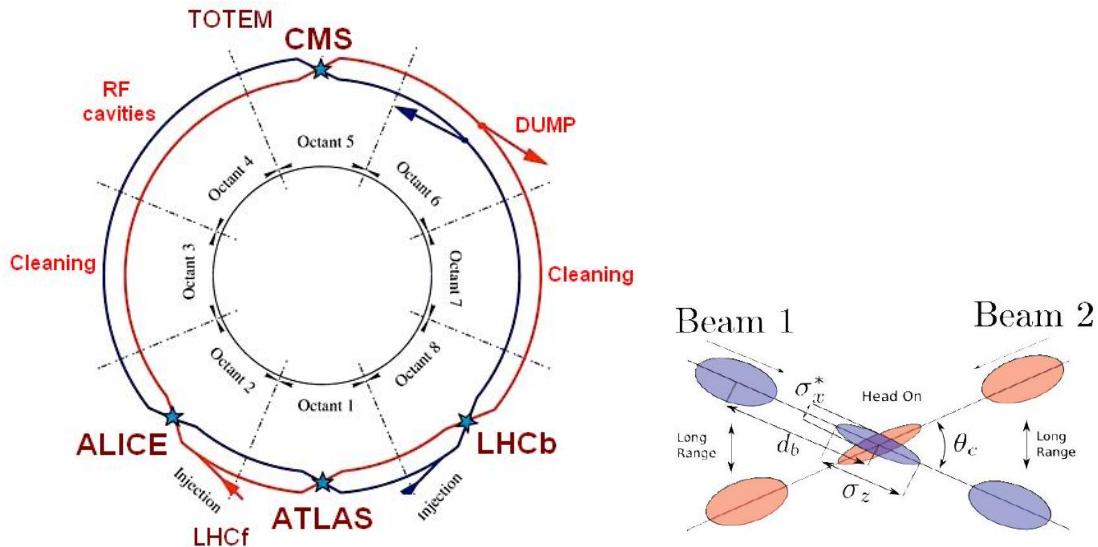


**Figure 3.6:** Integrated luminosity delivered by LHC and recorded by CMS during 2016. The difference between the delivered and the recorded luminosity is due to fails and issues occurred during the data taking in the CMS experiment [63].

1054 Once the proton beams reach the desired energy, they are brought to cross each other

1055 producing proton-proton collisions. The bunch crossing happens in precise places  
 1056 where the four LHC experiments are located, as seen in figure 3.7 left. In 2008, the  
 1057 first set of collisions involved protons with  $\sqrt{s} = 7$  TeV; the energy was increased to  
 1058 8 TeV in 2012 and to 13 TeV in 2015.

1059 CMS and ATLAS experiments, which are multi-purpose experiments, are enabled  
 1060 to explore physics in any of the collision modes. LHCb experiment is optimized  
 1061 to explore bottom quark physics, while ALICE is optimized for heavy ion collisions  
 1062 searches; TOTEM and LHCf are dedicated to forward physics studies; MoEDAL (not  
 1063 indicated in the figure) is intended for monopoles or massive pseudo stable particles  
 1064 searches.



**Figure 3.7:** Left: LHC interaction points. Bunch crossing occurs where the LHC experiments are located [64]. 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. Right: bunch crossing scheme. Since the bunch crossing is not perfectly head-on, the luminosity is reduced in a factor of 17%.

1065 At the CP there are two interesting details that need to be addressed. The first  
 1066 one is that the bunch crossing does not occur head-on but at a small crossing angle  
 1067 (280  $\mu$ rad in CMS and ATLAS) as shown in the right side of figure 3.7, affecting the

1068 overlapping between bunches; the consequence is a reduction of about 17% in the  
 1069 luminosity. The second one is the occurrence of multiple  $pp$  collisions in the same  
 1070 bunch crossing; this effect is called pile-up (PU). A fairly simple estimation of the  
 1071 PU follows from estimating the probability of collision between two protons, one from  
 1072 each of the bunches in course of collision; it depends roughly on the ratio of proton  
 1073 size and the cross section of the bunch in the interaction point, i.e.,

$$P(pp\text{-}collision) \sim \frac{d_{proton}^2}{\sigma_x \sigma_y} = \frac{(1\text{fm})^2}{(16\mu\text{m})^2} \sim 4 \times 10^{-21} \quad (3.4)$$

1074 however, there are  $N = 1.15 \times 10^{11}$  protons in a bunch, thus the estimated number of  
 1075 collisions in a bunch crossing is

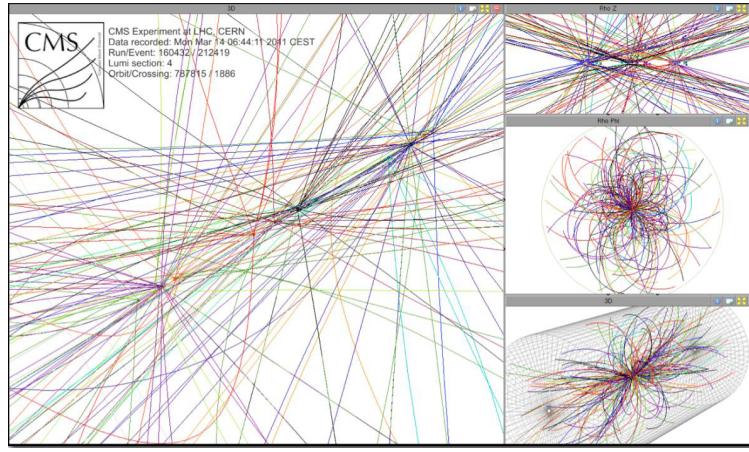
$$PU = N^2 * P(pp\text{-}collision) \sim 50 \text{ pp-collision per bunch crossing}, \quad (3.5)$$

1076 about 20 of those  $pp$  collisions are inelastic. Each collision generates a vertex, but  
 1077 only the most energetic is considered as a primary vertex; the rest are considered as  
 1078 PU vertices. A multiple  $pp$  collision event in a bunch crossing at CMS is showed in  
 1079 figure 3.8. Unstable particles outgoing from the primary vertex will eventually decay;  
 1080 this decay vertex is known as a secondary vertex.

1081 Next section presents a description of the CMS detector which it is the detector used  
 1082 to collect the data used in this thesis.

### 1083 3.3 The CMS experiment

1084 CMS is a general-purpose detector designed to conduct research in a wide range  
 1085 of physics from the standard model to new physics like extra dimensions and dark  
 1086 matter. Located at the point 5 in the LHC layout as shown in figure 3.4, CMS is

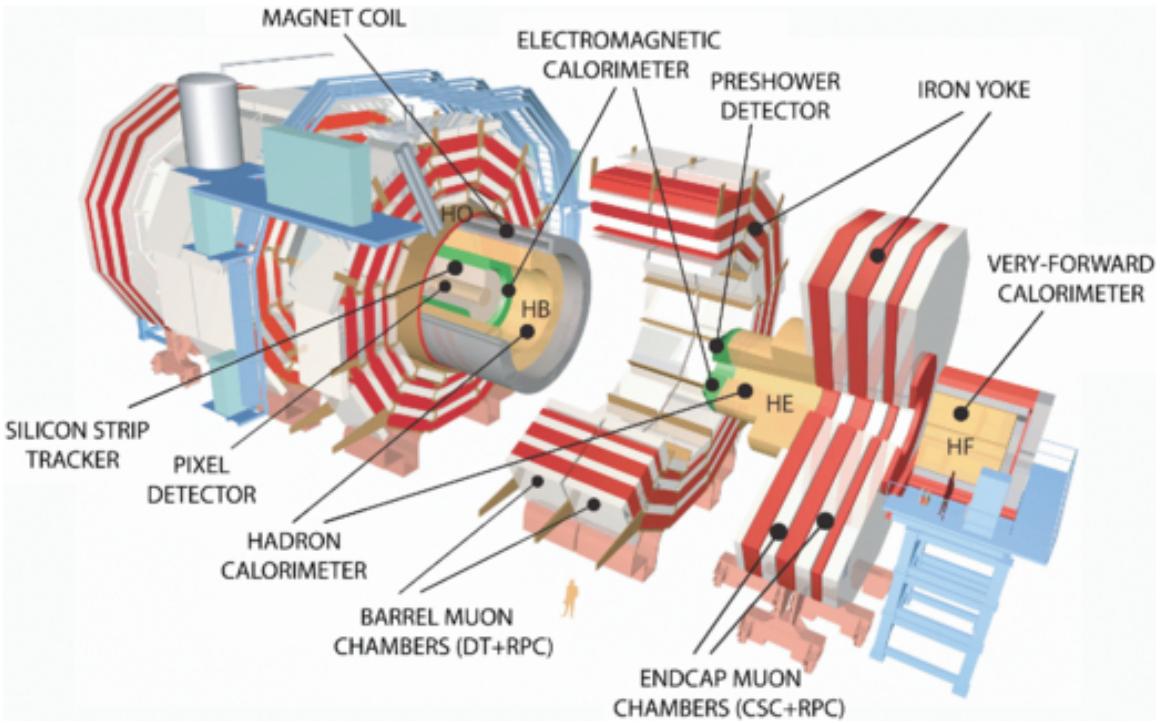


**Figure 3.8:** Multiple  $pp$  collision bunch crossing at CMS. Only the most energetic vertex is considered and the rest are cataloged as PU vertices [65].

1087 composed of several detection systems distributed in a cylindrical structure; in total,  
 1088 CMS weights about 12500 tons in a very compact 21.6 m long and 14.6 m diameter  
 1089 cylinder. It was built in 15 separate sections at the ground level and lowered to the  
 1090 cavern individually to be assembled. A complete and detailed description of the CMS  
 1091 detector and its components is given in reference [66] on which this section is based on.  
 1092

1093 Figure 3.9 shows the layout of the CMS detector. The design is driven by the require-  
 1094 ments on the identification, momentum resolution and unambiguous charge determi-  
 1095 nation of the muons; therefore, a large bending power is provided by the solenoid  
 1096 magnet made of superconducting cable capable to generate a 3.8 T magnetic field.  
 1097 The detection system is composed of (from the innermost to the outermost)

- 1098     • Pixel detector.
- 1099     • Silicon strip tracker.
- 1100     • Preshower detector.
- 1101     • Electromagnetic calorimeter.



**Figure 3.9:** 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. [67].

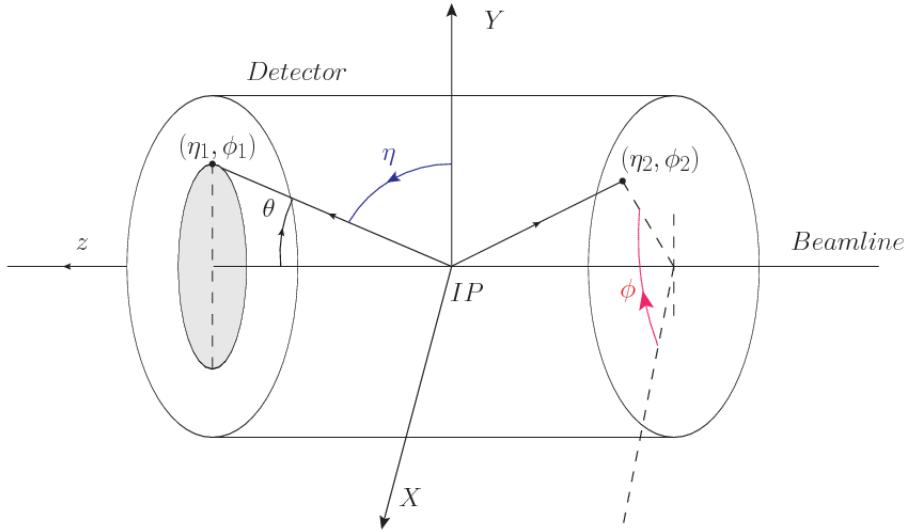
1102 • Hadronic calorimeter.

1103 • Muon chambers (Barrel and endcap)

1104 The central region of the detector is commonly referred as the barrel section while the  
 1105 endcaps are referred as the forward sections of the detector; thus, each subdetector  
 1106 is composed of a barrel section and a forward section.

### 1107 **3.3.1 Coordinate system**

1108 The coordinate system used by CMS is centered in the geometrical center of the  
 1109 detector which is the same as the CP as shown in figure 3.10. The  $z$ -axis is parallel  
 1110 to the beam direction, while the  $Y$ -axis pointing vertically upward, and the  $X$ -axis  
 1111 pointing radially inward toward the center of the LHC.



**Figure 3.10:** CMS detector coordinate system.

1112 In addition to the common cartesian and cylindrical coordinate systems, two coordi-  
 1113 nates are of particular utility in particle physics: rapidity( $y$ ) and pseudorapidity( $\eta$ ),  
 1114 defined in connection to the polar angle  $\theta$ , energy and longitudinal momentum com-  
 1115 ponent (momentum along the  $z$ -axis) according to

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \quad \eta = -\ln \left( \tan \frac{\theta}{2} \right) \quad (3.6)$$

1116 Rapidity is related to the angle between the  $XY$ -plane and the direction in which the  
 1117 products of a collision are emitted; it has the nice property that the difference between  
 1118 the rapidities of two particles is invariant with respect to Lorentz boosts along the  $z$ -  
 1119 axis. Thus, data analysis becomes more simple when based on rapidity; however, it is  
 1120 not simple to measure the rapidity of highly relativistic particles, as those produced  
 1121 after  $pp$  collisions. Under the highly relativistic motion approximation,  $y$  can be  
 1122 rewritten in terms of the polar angle, concluding that rapidity is approximately equal  
 1123 to the pseudorapidity defined above, i.e.  $y \approx \eta$ . Note that  $\eta$  is easier to measure than  $y$   
 1124 given the direct relationship between the former and the polar angle. Angular distance

1125 between two objects in the detector ( $\Delta R$ ) is defined in terms of their coordinates  
 1126  $(\eta_1, \phi_1), (\eta_2, \phi_2)$  as

$$\Delta R = \sqrt{(\Delta\eta)^2 - (\Delta\phi)^2} \quad (3.7)$$

1127 **3.3.2 Pixels detector**

1128 The CMS tracking system is designed to provide a precise measurement of the tra-  
 1129 jectory (*track*) followed by the charged particles created after the *pp* collisions; also,  
 1130 the precise reconstruction of the primary and secondary origins (*vertices*) is expected  
 1131 in an environment where, each 25 ns, the bunch crossing produce about 20 inelastic  
 1132 collisions and about 1000 particles. An increment in the luminosity is ongoing which  
 1133 implies that the PU will increase accordingly.

1134

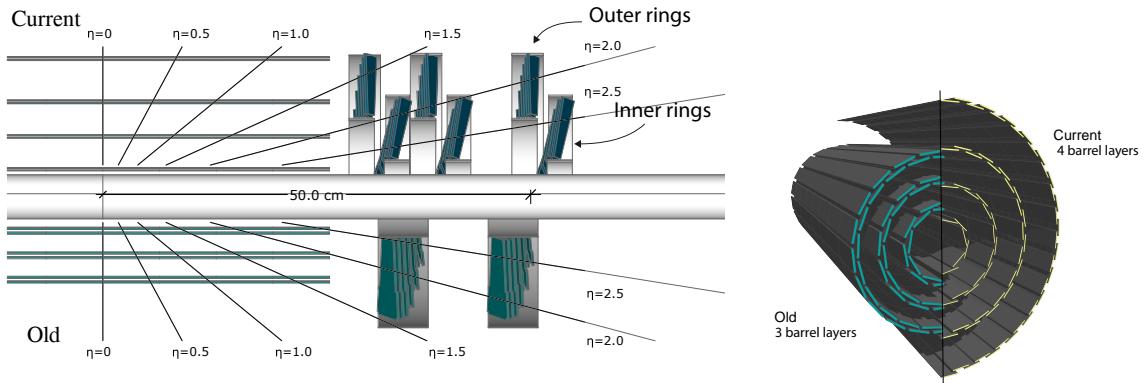
1135 The pixel detector was replaced during the 2016-2017 extended year-end technical  
 1136 stop, due to the increasingly challenging operating conditions like the higher particle  
 1137 flow and more radiation harsh environment, among others. The new one is respond-  
 1138 ing as expected, reinforcing its crucial role in the successful way to fulfill the new  
 1139 LHC physics objectives after the discovery of the Higgs boson. The last chapter of  
 1140 this thesis is dedicated to describe my contribution to the “Forward Pixel Phase 1  
 1141 upgrade”.

1142

1143 The current pixel detector is composed of 1856 silicon pixel detector modules orga-  
 1144 nized in four-barrel layers in the central region and three disks in the forward region;  
 1145 it is designed to record efficiently and with high precision, up to  $10\mu\text{m}$  in the *XY*-  
 1146 plane and  $20\mu\text{m}$  in the *z*-direction, the first four space-points (*hits*) near to the CP  
 1147 region (see figure 3.11 left side) in the range  $|\eta| \leq 2.5$ . The first barrel layer is located

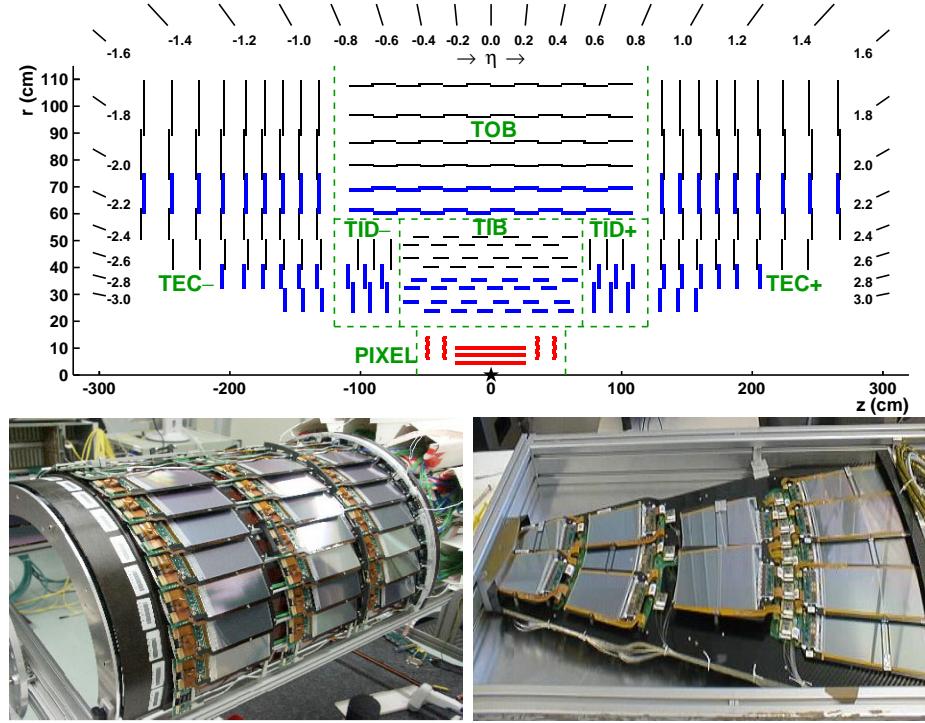
at a radius of 30 mm from the beamline, while the fourth layer is located at a radius of 160 mm closer to the strip tracker inner barrel layer (see section 3.3.3) in order to reduce the rate of fake tracks. The high granularity of the detector is represented in its about 123 Mpixels, each of size  $100 \times 150 \mu\text{m}^2$ , which is almost twice the channels of the old detector. The transverse momentum resolution of tracks can be measured with a resolution of 1-2% for muons of  $p_T = 100 \text{ GeV}$ .

1154



**Figure 3.11:** CMS pixel detector schematic view. Left: layout comparing the layers and disks in the old and current pixel detectors. Right: Transverse-oblique view comparing the pixel barrel layers in the two [69].

Some of the improvements with respect to the previous pixel detector include a higher average tracking efficiency and lower average fake rate as well as higher track impact parameter resolution which is fundamental in order to increase the efficiency in the identification of jets originating from b quarks (b-tagging). A significant source of improvement comes from the overall reduction in the material budget of the detector which results in fewer photon conversions and less multiple scattering from charged particles.



**Figure 3.12:** 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 [70–72].

### 1162 3.3.3 Silicon strip tracker

1163 The silicon strip tracker (SST) is the second stage in the CMS tracking system. The  
 1164 top side of figure 3.12 shows a schematic of the SST. The inner tracker region is com-  
 1165 posed of the tracker inner barrel (TIB) and the tracker inner disks (TID) covering the  
 1166 region  $r < 55$  cm and  $|z| < 118$  cm. The TIB is composed of 4 layers while the TID  
 1167 is composed of 3 disks at each end. The silicon sensors in the inner tracker are 320  
 1168  $\mu\text{m}$  thick, providing a resolution of about 13–38  $\mu\text{m}$  in the  $r\phi$  position measurement.  
 1169

1170 The modules indicated in blue in the schematic view of figure 3.12 are two modules

1171 mounted back-to-back and rotated in the plane of the module by a “stereo” angle of  
 1172 100 mrad; the hits from these two modules, known as “stereo hits”, are combined to  
 1173 provide a measurement of the second coordinate ( $z$  in the barrel and  $r$  on the disks)  
 1174 allowing the reconstruction of hit positions in 3-D.

1175

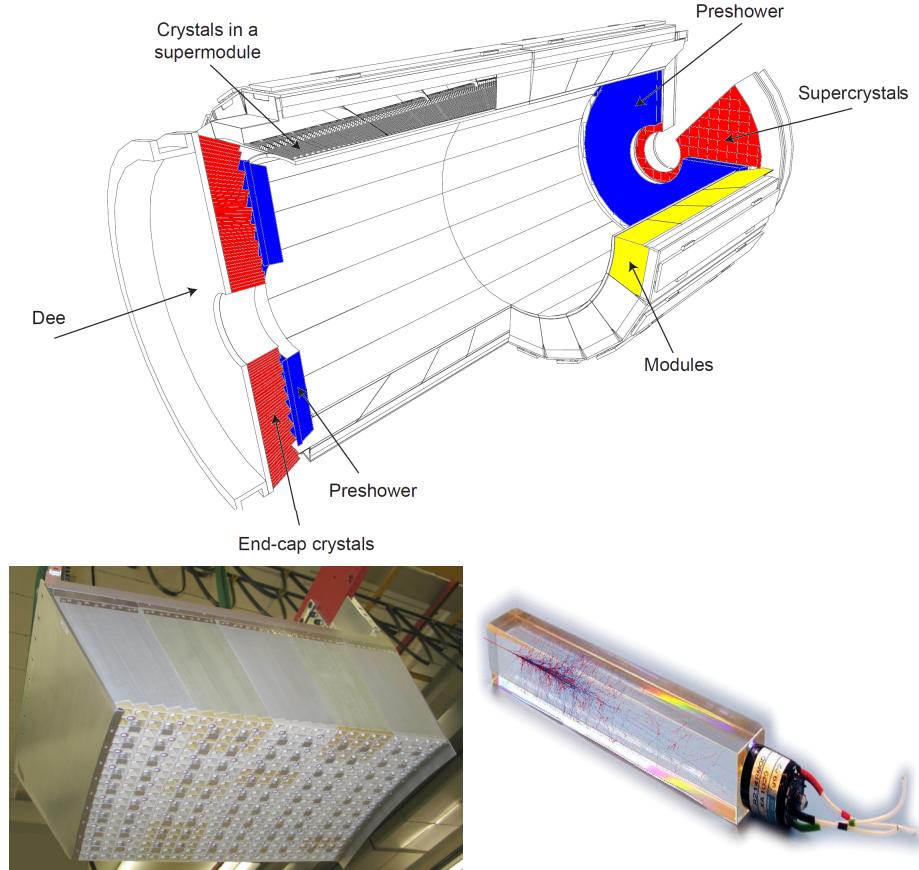
1176 The outer tracker region is composed of the tracker outer barrel (TOB) and the  
 1177 tracker endcaps (TEC). The six layers of the TOB offer coverage in the region  $r > 55$   
 1178 cm and  $|z| < 118$  cm, while the 9 disks of the TEC cover the region  $124 < |z| < 282$   
 1179 cm. The resolution offered by the outer tracker is about 13-38  $\mu\text{m}$  in the  $r\phi$  position  
 1180 measurement. The inner four TEC disks use silicon sensors 320  $\mu\text{m}$  thick; those in  
 1181 the TOB and the outer three TEC disks use silicon sensors of 500  $\mu\text{m}$  thickness. The  
 1182 silicon strips run parallel to the  $z$ -axis and the distance between strips varies from 80  
 1183  $\mu\text{m}$  in the inner TIB layers to 183  $\mu\text{m}$  in the inner TOB layers; in the endcaps the  
 1184 wedge-shaped sensors with radial strips, whose pitch range between 81  $\mu\text{m}$  at small  
 1185 radii and 205  $\mu\text{m}$  at large radii.

1186

1187 The whole SST has 15148 silicon modules, 9.3 million silicon strips and cover a total  
 1188 active area of about 198  $\text{m}^2$ .

### 1189 3.3.4 Electromagnetic calorimeter

1190 The CMS electromagnetic calorimeter (ECAL) is designed to measure the energy of  
 1191 electrons and photons. It is composed of 75848 lead tungstate crystals which have a  
 1192 short radiation length (0.89 cm) and fast response, since 80% of the light is emitted  
 1193 within 25 ns; however, they are combined with Avalanche photodiodes (APDs) as  
 1194 photodetectors given that crystals themselves have a low light yield ( $30\gamma/\text{MeV}$ ). A



**Figure 3.13:** Top: CMS ECAL schematic view. Bottom: Module equipped with the crystals (left); ECAL crystal(right).

1195 schematic view of the ECAL is shown in figure 3.13.

1196

1197 Energy is measured when electrons and photons are absorbed by the crystals which  
 1198 generates an electromagnetic “shower”, as seen in bottom right picture of the fig-  
 1199 ure3.13; the shower is seen as a *cluster* of energy which depending on the amount  
 1200 of energy deposited can involve several crystals. The ECAL barrel (EB) covers the  
 1201 region  $|\eta| < 1.479$ , using crystals of depth of 23 cm and  $2.2 \times 2.2 \text{ cm}^2$  transverse  
 1202 section; the ECAL endcap (EE) covers the region  $1.479 < |\eta| < 3.0$  using crystals  
 1203 of depth 22 cm and transverse section of  $2.86 \times 2.86 \text{ cm}^2$ ; the photodetectors used

1204 are vacuum phototriodes (VPTs). Each EE is divided in two structures called “Dees”.

1205

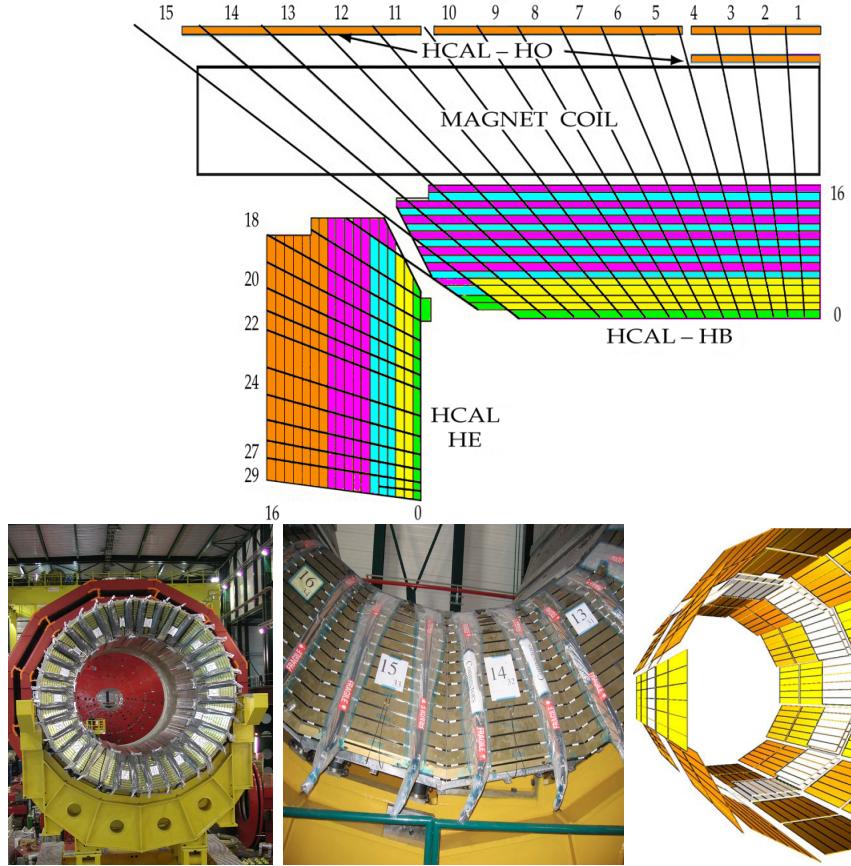
1206 In front of the EE, it is installed the preshower detector (ES) which covers the region  
 1207  $1.653 < |\eta| < 2.6$ . The ES provides a precise measurement of the position of electro-  
 1208 magnetic showers, which allows to distinguish electrons and photons signals from  $\pi^0$   
 1209 decay signals. The ES is composed of a layer of lead absorber followed by a layer of  
 1210 plastic scintillators

### 1211 3.3.5 Hadronic calorimeter

1212 Hadrons are not absorbed by the ECAL but by the hadron calorimeter (HCAL),  
 1213 which is made of a combination of alternating brass absorber layers and silicon photo-  
 1214 multiplier(SiPM) layers; therefore, particles passing through the scintillator material  
 1215 produce showers, as in the ECAL, as a result of the inelastic scattering of the hadrons  
 1216 with the detector material. Since the particles are not absorbed in the scintillator,  
 1217 their energy is sampled; therefore the total energy is not measured but estimated from  
 1218 the energy clusters, which reduce the resolution of the detector. Brass was chosen  
 1219 as the absorber material due to its short interaction length ( $\lambda_I = 16.42\text{cm}$ ) and its  
 1220 non-magnetivity. Figure 3.14 shows a schematic view of the CMS HCAL.

1221

1222 The HCAL is divided into four sections; the Hadron Barrel (HB), the Hadron Outer  
 1223 (HO), the Hadron Endcap (HE) and the Hadron Forward (HF) sections. The HB  
 1224 covers the region  $0 < |\eta| < 1.4$ , while the HE covers the region  $1.3 < |\eta| < 3.0$ . The HF,  
 1225 made of quartz fiber scintillator and steel as absorption material, covers the forward  
 1226 region  $3.0 < |\eta| < 5.2$ . Both the HB and HF are located inside the solenoid. The HO  
 1227 is placed outside the magnet as an additional layer of scintillators with the purpose

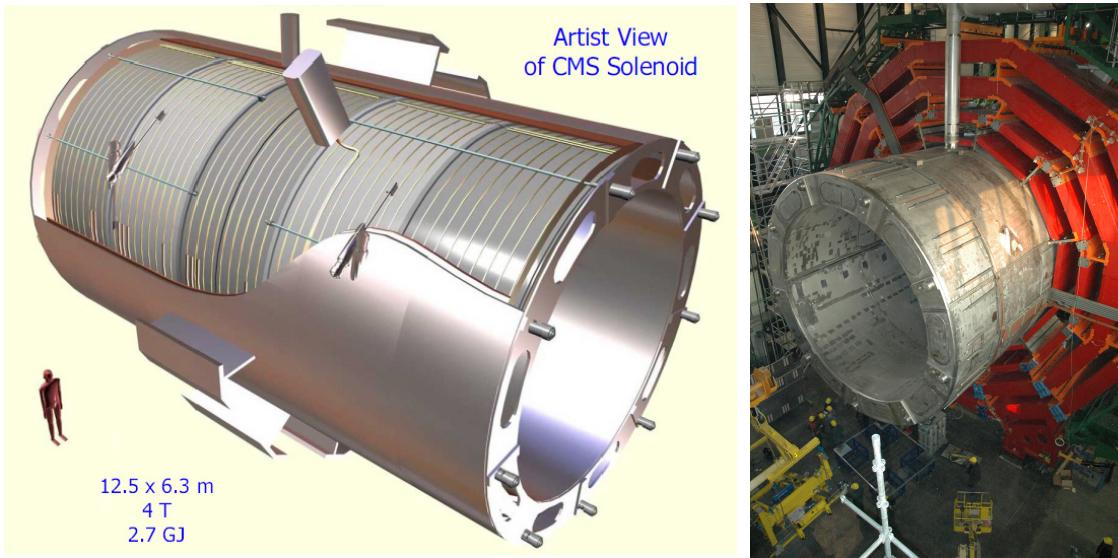


**Figure 3.14:** Top: CMS HCAL schematic view, the colors indicate the layers that are grouped into the same readout channels. Bottom: picture of a section of the HB; the absorber material is the golden region and scintillators are placed in between the absorber material (left and center). Schematic view of the HO (right). [73,74]

of measure the energy tails of particles passing through the HB and the magnet (see figure 3.14 top and bottom right). The upgrades made to the HCAL during the technical stop 2016-2017 consisted in the replacement of the photo transducer, in order to improve the efficiency.

### 3.3.6 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



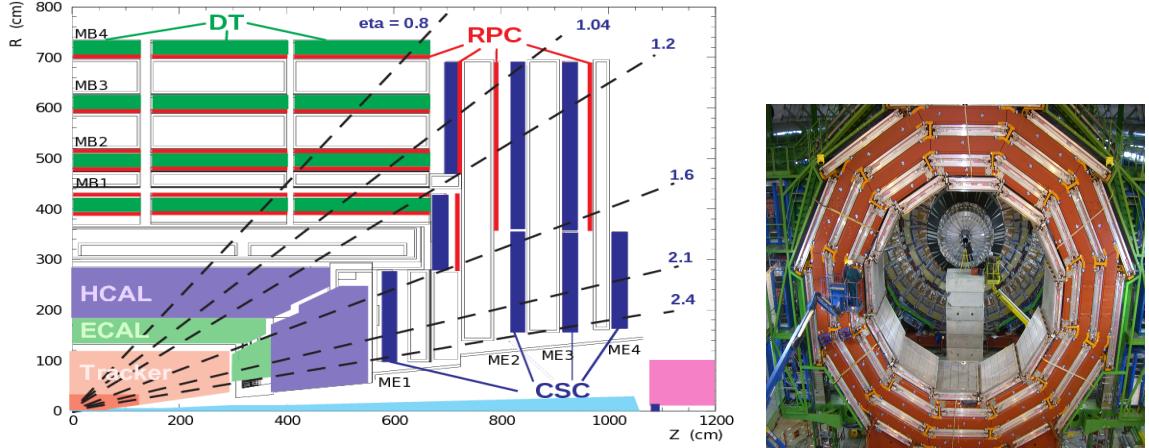
**Figure 3.15:** Artistic representation of the CMS solenoid magnet(left). The magnet is supported on an iron yoke (right) which also serves as the house of the muon detector and as mechanical support for the whole CMS detector [68].

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 unambiguous determination of the sign for high momentum muons was a driven 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 3.15 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.

1245

1246 The yoke (see figure 3.15), composed of 5 barrel wheels and 6 endcap disks made  
 1247 of iron, serves not only as the media for magnetic flux return but also provides the  
 1248 house for the muon detector system and structural stability to the full detector.

1249 **3.3.7 Muon system**



**Figure 3.16:** 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 [75].

1250 Muons are the only charged particles able to pass through all the CMS detector due  
 1251 to their low ionization energy loss; thus, muons can be separated easily from the  
 1252 high amount of particles produced in a  $pp$  collision. Also, muons are expected to be  
 1253 produced in the decay of several new particles; therefore, a good detection of muons  
 1254 was on the leading principles when designing the CMS detector.

1255

1256 The CMS muon detection system (muon spectrometer) is embedded in the return  
 1257 yoke as seen in figure 3.16. It is composed of three different detector types, the drift  
 1258 tube chambers (DT), Cathode strip chambers (CSC), and resistive plate chambers  
 1259 (RPC); DT are located in the central region  $\eta < 1.2$  arranged in four layers of drift  
 1260 chambers filled with an Ar/CO<sub>2</sub> gas mixture.

1261

1262 The muon endcaps are made of CSCs covering the region  $\eta < 2.4$  and filled with a  
 1263 mixture of Ar/CO<sub>2</sub>/CF<sub>4</sub>. The reason behind using a different detector type lies on  
 1264 the different conditions in the forward region like the higher muon rate and higher

1265 residual magnetic field compared to the central region.

1266

1267 The third type of detector used in the muon system is a set of four disks of RPCs

1268 working in avalanche mode. The RPCs provide good spatial and time resolutions.

1269 The track of  $high - p_T$  muon candidates is built combining information from the

1270 tracking system and the signal from up to six RPCs and four DT chambers.

1271 The muon tracks are reconstructed from the hits in the several layers of the muon

1272 system.

### 1273 3.3.8 CMS trigger system

1274 Under normal conditions, CMS expects  $pp$  collisions every 25 ns i.e. an interaction

1275 rate of 40 MHz for which it is not possible to store the recorded data in full. In order

1276 to handle this high event rate data, an online event selection, known as triggering, is

1277 performed; triggering reduce the event rate to 100 Hz for storage and further offline

1278 analysis.

1279

1280 The trigger system starts with a reduction of the event rate to 100 kHz in the so-called

1281 “level 1 trigger (L1)”. L1 is based on dedicated programmable hardware like Field

1282 Programmable Gate Arrays (FPGAs) and Application Specific Integrated Circuits

1283 (ASICs), partly located in the detector itself; another portion is located in the CMS

1284 under-ground cavern. Hit patterns information from the muon chambers and the en-

1285 ergy deposits in the calorimeter are used to decide if an event is accepted or rejected,

1286 according to selection requirements previously defined, which reflect the interesting

1287 physics processes. Figure 3.17 shows the L1 trigger architecture

1288

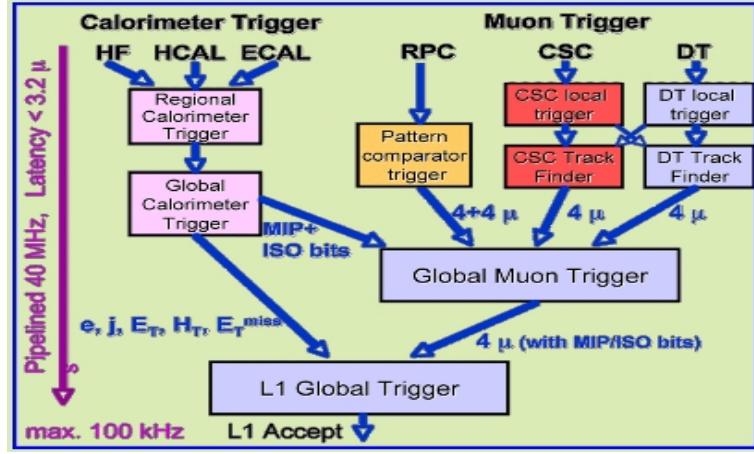


Figure 3.17: CMS Level-1 trigger architecture [76].

1289 The second stage in the trigger system is called “high-level trigger (HLT)”; events  
 1290 accepted by L1 are passed to HLT in order to make an initial reconstruction of them.  
 1291 HLT is software based and runs on a dedicated server farm, using selection algo-  
 1292 rithms and high-level object definitions; the event rate at HLT is reduced to 100 Hz.  
 1293 The first HLT stage takes information from the muon detectors and the calorimeters  
 1294 to make the initial object reconstruction; in the next HLT stage, information from  
 1295 the pixel and strip detectors is used to do first fast-tracking and then full tracking  
 1296 online. This initial object reconstruction is used in further steps of the trigger system.

1297

1298 Events and preliminary reconstructed physics objects from HLT are sent to be fully  
 1299 reconstructed at the CERN computing center. Again, the pixel detector information  
 1300 provides high-quality seeds for the track reconstruction algorithm offline, primary ver-  
 1301 tex reconstruction, electron and photon identification, muon reconstruction,  $\tau$  iden-  
 1302 tification, and b-tagging. After full reconstruction, data sets are made available for  
 1303 offline analyses.

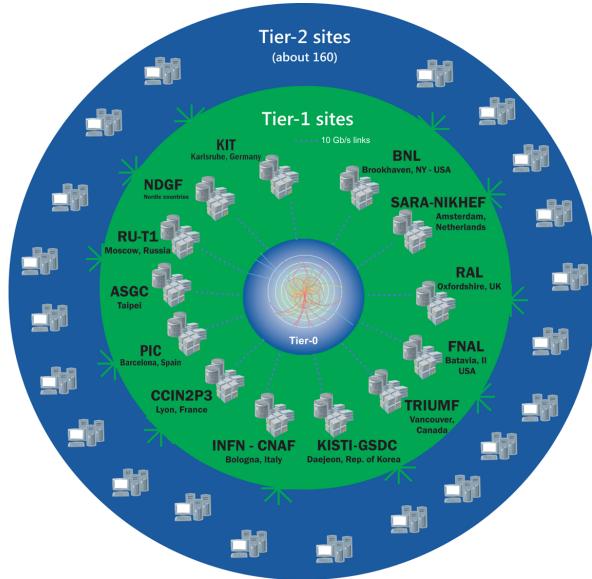
1304

1305 During the 2016-2017 technical stop, the L1 system was updated in order to improve

1306 the physics object identification by improving the algorithms and accounting for the  
 1307 increasing pile-up scenario.

1308 **3.3.9 CMS computing**

1309 After the data, coming from the experiment, are processed at several levels, they have  
 1310 to be stored and made available for further analysis; in order to cope all the tasks  
 1311 implied in the offline data processing, like transfer, simulation, reconstruction and  
 1312 reprocessing, among others, a big computing power is required. The CMS computing  
 1313 system is based on the distributed architecture concept, where users of the system  
 1314 and physical computer centers are distributed worldwide and interconnected by high-  
 1315 speed networks.



**Figure 3.18:** WLCCG 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. [77].

1316 The worldwide LHC computing grid (WLCCG) is the mechanism used to provide that  
 1317 distributed environment. WLCCG is a tiered structure connecting computing centers

1318 around the world, which provides the necessary storage and computing facilities. The  
1319 primary computing centers of the WLCG are located at the CERN and the Wigner  
1320 datacenter in Budapest and are known as Tier-0 as shown in figure 3.18. The main  
1321 responsibilities for each tier level are [77]

1322 • **Tier-0:** initial reconstruction of recorded events and storage of the resulting  
1323 datasets, the distribution of raw data to the Tier-1 centers.

1324 • **Tier-1:** provide storage capacity, support for the Grid, safe-keeping of a pro-  
1325 portional share of raw and reconstructed data, large-scale reprocessing and safe-  
1326 keeping of corresponding output, generation of simulated events, distribution  
1327 of data to Tier 2s, safe-keeping of a share of simulated data produced at these  
1328 Tier 2s.

1329 • **Tier-2:** store sufficient data and provide adequate computing power for specific  
1330 analysis tasks, provide analysis requirements and proportional share of simu-  
1331 lated event production and reconstruction.

1332 Aside from the general computing strategy to manage the huge amount of data pro-  
1333 duced by experiments, CMS uses a framework to perform a variety of processing,  
1334 selection and analysis tasks. The central concept of the CMS data model referred to  
1335 as “event data model” (EDM) is the “Event”; therefore, an event is the unit that con-  
1336 tains the information from a single bunch crossing as well as any data derived from  
1337 that information like the reconstructed objects, the details under which additional  
1338 data are derived.

1339

1340 Events are passed as the input to the “physics modules” that obtain information from  
1341 them and create new one; for instance, “event data producers” add new data into the

1342 events, “analyzers” produce an information summary from an event set, “filters” per-  
1343 form selection and triggering.

1344

1345 CMS uses several event formats with different levels of detail and precision

1346 • **Raw format:** events in this format contain the full recorded information from  
1347 the detector as well as trigger decision and other metadata. An extended version  
1348 of raw data is used to store information from the CMS Monte Carlo simulation  
1349 tools. Raw data are stored permanently, occupying about 2MB/event

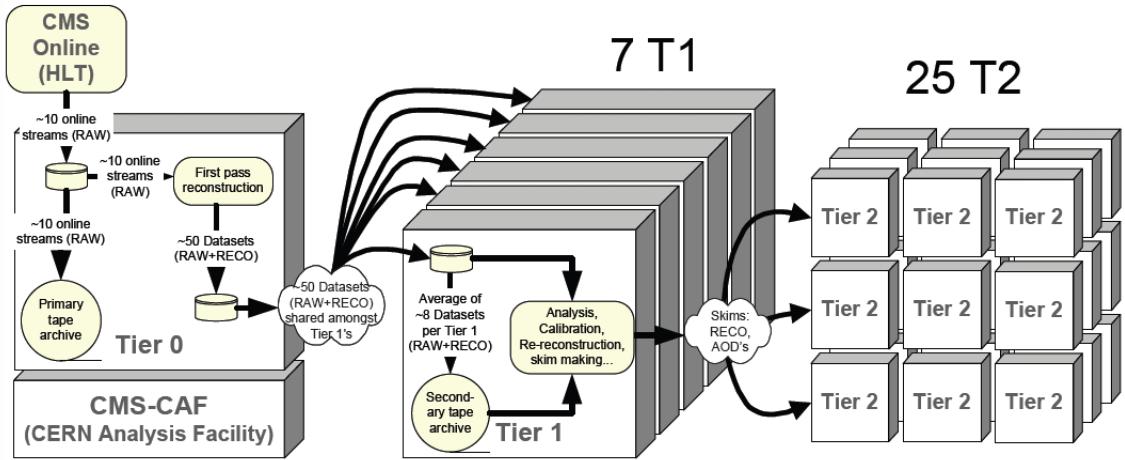
1350 • **RECO format:** events in this format correspond to raw data that have been  
1351 submitted to reconstruction algorithms like primary and secondary vertex re-  
1352 construction, particle ID, track-finding. RECO events contain physical objects  
1353 and all the information used to reconstruct them; average size is about 0.5  
1354 MB/event.

1355 • **AOD format:** Analysis Object Data (AOD) is the data format used in the  
1356 physics analyses given that it contains the parameters describing the high-level  
1357 physics objects in addition to enough information to allow a kinematic refitting if  
1358 needed. AOD events are filtered versions of the RECO events to which skimming  
1359 or other kind processes have been applied. Requires about 100 kB/event.

1360 • **Non-event data** are data needed to interpret and reconstruct events. Some  
1361 of the non-event data used by CMS contains information about the detector  
1362 contraction and condition data like calibrations, alignment, and detector status.

1363 Figure 3.19 shows the data flow scheme between CMS detector and hardware tiers.

1364 The whole collection of software built as a framework is referred to as “CMSSW”. This  
1365 framework provides the services needed by the simulation, calibration and alignment,



**Figure 3.19:** Data flow from CMS detector through hardware Tiers.

1366 and reconstruction modules that process event data, so that physicists can perform  
 1367 analysis. The CMSSW event processing model is composed of one executable, called  
 1368 cmsRun, and several plug-in modules which contain all the tools (calibration, recon-  
 1369 struction algorithms) needed to process an event. The same executable is used for  
 1370 both detector and Monte Carlo data [78].

# <sup>1371</sup> Chapter 4

## <sup>1372</sup> Event generation, simulation and <sup>1373</sup> reconstruction

<sup>1374</sup> The process of analyzing the data recorded by the CMS experiment involves several  
<sup>1375</sup> stages where the data are processed in order to interpret the information provided by  
<sup>1376</sup> all the detection systems; in those stages the particles produced after the  $pp$  collision  
<sup>1377</sup> are identified by reconstructing their trajectories and measuring their features. In  
<sup>1378</sup> addition, the SM provides a set of predictions that have to be compared with the  
<sup>1379</sup> experimental results; however, in most of the cases, theoretical predictions are not  
<sup>1380</sup> directly comparable to experimental results due to the diverse source of uncertainties  
<sup>1381</sup> introduced by the experimental setup and theoretical approximations among others.

<sup>1382</sup>

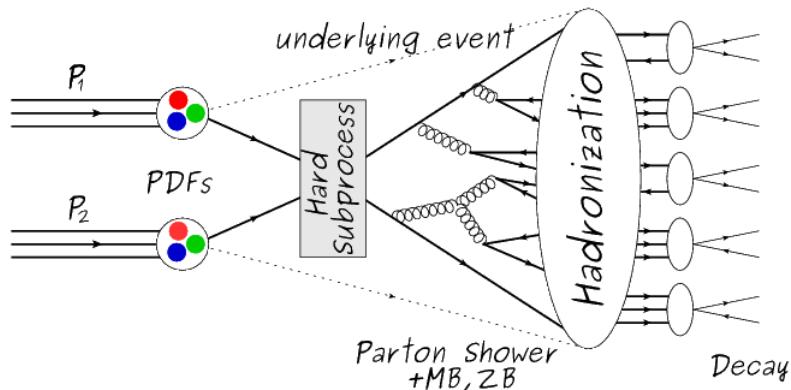
<sup>1383</sup> The strategy to face these conditions consist in using statistical methods implemented  
<sup>1384</sup> in computational algorithms to produce numerical results that can be contrasted with  
<sup>1385</sup> the experimental results. These computational algorithms are commonly known as  
<sup>1386</sup> Monte Carlo (MC) methods and, in the case of particle physics, they are designed to  
<sup>1387</sup> 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 for 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.

1393

1394 This chapter presents a description of the event generation strategy and the tools  
 1395 used to perform the detector simulation and physics objects reconstruction. A com-  
 1396 prehensive review on event generators for LHC physics can be found in reference [79]  
 1397 on which this chapter is based.

## 1398 4.1 Event generation



**Figure 4.1:** Event generation process. In the first step, the PDF of the colliding particles is considered so the specific interaction is described. The actual interaction is generated in the hard subprocess; the cross section of the process is calculated from the matrix element connecting the initial and final states. The parton shower describes the evolution of the partons from the hard subprocess according to the DGLAP equations. At this step the underlying event and PU effects are included in the generation. The resulting partons from the parton shower are recombined to form hadrons in the hadronization step; most of them are unstable, therefore, their decays are also generated in agreement to the known branching ratios. Modified from reference [80].

1399 The event generation is intended to create events that mimic the behavior of actual  
 1400 events produced in the collisions; they obey a sequence of steps from the particles colli-  
 1401 sion hard process to the decay process into the final state particles. Figure 4.1 shows  
 1402 an schematic view of the event generation process; the fact that the full process can  
 1403 be treated as several independent steps is based on the QCD factorization theorem.

1404

1405 Generation starts by taking into account the PDFs of the incoming particles. Event  
 1406 generators offer the option to chose from several PDF sets depending on the partic-  
 1407 ular process under simulation<sup>1</sup>; in the following  $pp$  collisions will be considered. The  
 1408 *hard subprocess* describes the actual interaction between partons from the incoming  
 1409 protons; it is represented by the matrix element connecting the initial and final states  
 1410 of the interaction. Normally, the matrix element can be written as a sum over Feyn-  
 1411 man diagrams and consider interferences between terms in the summation. During  
 1412 the generation of the hard subprocess, the production cross section is calculated.

1413

1414 The order to which the cross section is calculated depends on the order of the Feyn-  
 1415 man diagrams involved in the calculation; therefore, radiative corrections are included  
 1416 by considering a higher order Feynman diagrams where QCD radiation dominates.  
 1417 Currently, cross sections calculated to LO do not offer a satisfactory description of the  
 1418 processes, i.e., the results are only reliable for the shape of distributions; therefore,  
 1419 NLO calculations have to be performed with the implication that the computing time  
 1420 needed is highly increased.

1421

1422 The final parton content of the hard subprocess is subjected to the *parton shower*  
 1423 which generates the gluon radiation. Parton shower evolves the partons; i.e., glouns

---

<sup>1</sup> Tool in Reference [81] allows to plot different PDF sets under customizable conditions.

1424 split into quark-antiquark pairs and quarks of enough energy radiate gluons giving rise  
 1425 to further parton multiplication, following the DGLAP (Dokshitzer-Gribov-Lipatov-  
 1426 Altarelli-Parisi) equations. Showering continues until the energy scale is low enough  
 1427 to reach the non-perturbative limit.

1428

1429 In the simulation of LHC processes that involve  $b$  quarks like the single top quark or  
 1430 Higgs associated production, it is needed to consider that the  $b$  quark is heavier than  
 1431 the proton; in this sense, the QCD interaction description is made in two different  
 1432 schemes [82]

1433 • four-flavor (4F) scheme.  $b$  quarks appears only in the final state because they  
 1434 are heavier than the proton and therefore they can be produced only from the  
 1435 splitting of a gluon into pairs or singly in association with a  $t$  quark in high  
 1436 energy-scale interactions. During the simulation, the  $b$ -PDFs are set to zero  
 1437 because it cannot be part of the proton. Calculation in this scheme are more  
 1438 complicated due to the presence of the second  $b$  quark but the full kinematics is  
 1439 considered already at LO and therefore the accuracy of the description is better.

1440 • five-flavor (5F) scheme.  $b$  quarks are considered massless, therefore they can  
 1441 appear in both initial and final states since it can now be part of the proton; thus,  
 1442 during the simulation  $b$ -PDFs are not set to zero. In this scheme, calculations  
 1443 are simpler than in the 4F scheme and possible logarithmic divergences are  
 1444 absorbed by the PDFs through the DGLAP evolution.

1445 In this thesis, the  $tHq$  events are generated using the 4F scheme in order to reduce  
 1446 uncertainties, while the  $tHW$  events are generated using the 5F scheme to eliminate  
 1447 LO interference with the  $t\bar{t}H$  process [48].

1448

1449 Partons involved in the  $pp$  collision are the focus of the simulation, however, the rest  
 1450 of the partons inside the incoming protons are also affected because the remnants are  
 1451 colored objects; also, multiple parton interactions can occurs. The hadronization of  
 1452 the remnants and multiple parton interactions are known as “underlying event” and  
 1453 it has to be included in the simulation. In addition, multiple  $pp$  collisions in the same  
 1454 bunch crossing (pile-up mentioned in 3.2) occurs, actually in two forms

- 1455 • *in-time PU* which refers to multiple  $pp$  collision in the bunch crossing but that  
 1456 are not considered as primary vertices.
- 1457 • *Out-of-time PU* which refers to overlapping  $pp$  collisions from consecutive bunch  
 1458 crossings; this can occurs due to the time-delays in the detection systems where  
 1459 information from one bunch crossing is assigned to the next or previous one.

1460 While the underlying event effects are included in generation using generator-specific  
 1461 tools, PU effects are added to the generation by overlying Minimum-bias (MB) and  
 1462 Zero-bias (ZB) events to the generated events. MB events are inelastic events selected  
 1463 by using a loose (minimum bias) trigger with as little bias as possible, therefore ac-  
 1464 cepting a large fraction of the overall inelastic event; ZB events correspond to random  
 1465 events recorded by the detector when collisions are likely. MB model in-time PU and  
 1466 ZB model out-of-time PU.

1467

1468 The next step in the generation process is called “hadronization”. Since particles  
 1469 with a net color charge are not allowed to exits isolated, they have recombine to form  
 1470 bound states. This is precisely the process by which the partons resulting from the  
 1471 parton shower arrange themselves as color singlets to form hadrons. At this step, the  
 1472 energy-scale is low and the strong coupling constant is large, therefore hadronization  
 1473 process is non-perturbative and phenomenological model are used to describe the

1474 parton’s evolution. Most of the baryons and mesons produced in the hadronization  
 1475 are unstable and hence they will decay in the detector.

1476

1477 The last step in the generation process corresponds to the decay of the unstable  
 1478 particles generated during hadronization; it is also simulated in the hadronization  
 1479 step, based on the known branching ratios.

## 1480 4.2 Monte Carlo Event Generators.

1481 The event generation described in the previous section has been implemented in  
 1482 several software packages for which a brief description is given.

- 1483 • **PYTHIA 8.** It is a program designed to perform the generation of high en-  
 1484 ergy physics events which describe the collisions between particles such as elec-  
 1485 trons, protons. Several theories and models are implemented in it, in order to  
 1486 describe physical aspects like hard and soft interaction, parton distributions,  
 1487 initial and final-state parton showers, multiple parton interactions, beam rem-  
 1488 nants, hadronization<sup>2</sup> and particle decay. Thanks to extensive testing, several  
 1489 optimized parametrizations known as “tunnings” have been defined in order  
 1490 to improve the description of actual collisions to a high degree of precision; for  
 1491 analysis at  $\sqrt{s} = 13$  TeV, the underline event CUETP8M1 tune is employed [84].  
 1492 The calculation of the matrix element is performed at LO which is not enough  
 1493 for the current required level of precision; therefore, pythia is often used for  
 1494 parton shower, hadronization, decays, while other event generators are used to  
 1495 generate the matrix element at NLO.

---

<sup>2</sup> based in the Lund string model [83]

1496     • **MadGraph5\_aMC@NLO.** MadGraph is a matrix element generator which  
 1497       calculates the amplitudes for all contributing Feynman diagrams of a given pro-  
 1498       cess but does not provide a parton shower while MC@NLO incorporate NLO  
 1499       QCD matrix elements consistently into a parton shower framework; thus, Mad-  
 1500       Graph5\_aMC@NLO, as a merger of the two event generators MadGraph5 and  
 1501       aMC@NLO, is an event generator capable to calculate tree-level and NLO cross  
 1502       sections and perform the matching of those with the parton shower. It is one  
 1503       of the most frequently used matrix element generators; however, it has as par-  
 1504       ticular feature the presence of negative event weights which reduce the number  
 1505       of events used to reproduce the the properties of the objects generated [85].

1506

1507     • **POWHEG.** It is an NLO matrix element generator where the hardest emis-  
 1508       sion of color charged particles is generated in such a way that the negative event  
 1509       weights issue of MadGraph5\_aMC@NLO is overcome; however, the method re-  
 1510       quires an interface with  $p_T$ -ordered parton shower or a parton shower generator  
 1511       where this highest emission can be vetoed in order to avoid double counting of  
 1512       this highest-energetic emission. PYTHIA is a commonly matched to POWHEG  
 1513       event generator [86].

1514 Events resulting from the whole generation process are known as MC events.

### 1515 4.3 CMS detector simulation.

1516 After generation, MC events contain the physics of the collisions but they are not  
 1517 ready to be compared to the events recorded by the experiment since these recorded  
 1518 events correspond to the response of the detection systems to the interaction with the

1519 particles traversing them. The simulation of the CMS detector have to be applied on  
1520 top of the event generation; it is simulated with Geant4, a MC toolkit for the simula-  
1521 tion of particles passing though matter which is also able to simulates the electronic  
1522 signals that would be measured by all detectors inside CMS.

1523

1524 The simulation takes the generated particles contained in the MC events as input,  
1525 makes them to pass through the simulated geometry, and models physics processes  
1526 that particles experience during their passage through matter. The full set of results  
1527 from particle-matter interactions correspond to the simulated hit which contains in-  
1528 formation about the energy loss, momentum, position. Particles of the input event  
1529 are called “primary”, while the particles originating from GEANT4-modeled interac-  
1530 tions of a primary particle with matter are called a “secondary”. Simulated hits are  
1531 the input of subsequent modules that emulate the response of the detector readout  
1532 system and triggers. The output from the emulated detection systems and triggers is  
1533 known as digitization [87, 88].

1534

1535 The modeling of the CMS detector corresponds to the accurate modeling of the inter-  
1536 action among particles, the detector material and the magnetic field. This simulation  
1537 procedure includes the following standard steps

- 1538     • Modeling of the Interaction Region.
- 1539     • Modeling of the particle passage through the hierarchy of volumes that compose  
1540        CMS detector and of the accompanying physics processes.
- 1541     • Modeling of the effect of multiple interactions per beam crossing and/or the  
1542        effect of events overlay ( Pile-Up simulation).

1543       • Modeling of the detector’s electronics response, signal shape, noise, calibration  
 1544           constants (digitization).

1545 In addition to the full simulation, i.e. a detailed detector simulation, a faster simula-  
 1546 tions (FastSim) have been developed, that may be used where much larger statistics  
 1547 are required. In FastSim, detector material effects are parametrized and included in  
 1548 the hits; those hits are used as input of the same higher-level algorithms<sup>3</sup> used to an-  
 1549 alyze the recorded events. In this way, comparisons between fast and full simulations  
 1550 can be performed [90].

1551

1552 After the full detector simulation, the output events can be directly compared with  
 1553 events actually recorded in the CMS detector. The collection of MC events that  
 1554 reproduce the expected physics for a given process are known as MC samples.

## 1555 4.4 Event reconstruction.

1556 In contrast to MC samples for which all the particles’ information is available from  
 1557 it’s identity to it’s mass and energy, recorded events contain the electronic signals,  
 1558 provided by the CMS detection systems, encoding the interaction of physical parti-  
 1559 cles with the detector matter; these electronic signals have to be combined in order  
 1560 to identify these particles and measure their features i.e., particles have to be “recon-  
 1561 structed” using the signals provided by the detection systems. The CMS experiment  
 1562 use the “particle-flow event reconstruction algorithm (PF)” to do reconstruction of  
 1563 particles produced in  $pp$  collisions. Next sections will present a basic description of

---

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

1564 the *Elements* used by PF (tracker tracks, energy clusters and muon tracks), based in  
 1565 the references [91, 92] where more detailed descriptions can be found.

1566 **4.4.1 Particle-Flow Algorithm.**

1567 Each of the several subdetection systems of the CMS detector is dedicated to identify  
 1568 specific type of particles, i.e., photons and electrons are absorbed by the ECAL and  
 1569 their reconstruction is based on ECAL information; hadrons are reconstructed from  
 1570 clusters in the HCAL while muons are reconstructed from hits in the muon chambers.  
 1571 PF is designed to correlate signals from all the detector layers (tracks and energy  
 1572 clusters) in order to reconstruct and identify each final state particle and its properties.  
 1573 For instance, a charged hadron is identified by a geometrical connection, know as *link*  
 1574 between one or more calorimeter clusters and a track in the tracker provided there  
 1575 are no hits in the muon system; combining several measurements allows a better  
 1576 determination of the energy and charge sign of the charged hadron.

1577 **Charged-particle track reconstruction.**

1578 The strategy used by PF in order to reconstruct tracks is called “Iterative Tracking”  
 1579 which occurs in four steps

- 1580     • Seed generation where initial track candidates are found by looking for combina-  
       1581       tion of hits in the pixel detector, strip tracker and muon chambers. In total ten  
       1582       iterations are performed, each one with a different seeding requirement. Seeds  
       1583       are used to estimate the trajectory parameters and uncertainties at the time of  
       1584       the full track reconstruction. Seeds are also considered track candidates.
- 1585     • Track finding using a tracking software known as Combinatorial Track Finder  
       1586       (CTF) [93]. The seed trajectories are extrapolated along the expected flight

1587 path of a charged particle, in agreement to the the trajectory parameters ob-  
 1588 tained in the first step, in an attempt to find additional hits that can be assigned  
 1589 to the track candidates.

- 1590 • Track-fitting where the found tracks are passed as input to a module which  
 1591 provides the best estimate of the parameters of each trajectory.  
 1592 • Track selection where track candidates are submitted to a selection which dis-  
 1593 cards those that fail a set of defined quality criteria.

1594 Iterations differ in the seeding configuration and the final track selection as elaborated  
 1595 in references [91, 92]. In the first iteration, high  $p_T$  tracks and tracks produced near  
 1596 to the interaction region are identified and those hits are masked thereby reducing  
 1597 the combinatorial complexity. Next iterations search for more complicated tracks,  
 1598 like low  $p_T$  tracks and tracks from b hadron decays, which tend to be displaced from  
 1599 the interaction region.

1600 **Vertex reconstruction.**

1601 During the track reconstruction, an extrapolation toward to the calorimeters is per-  
 1602 formed in order to match energy deposits; that extrapolation is performed also toward  
 1603 the beam line in order to find the origin of the track known as *vertex*. The vertex re-  
 1604 construction is performed by selecting from the available reconstructed tracks, those  
 1605 that are consistent with being originated in the interaction region where  $pp$  collisions  
 1606 are produced. The selection involves a requirement on the number of tracker (pixel  
 1607 and strip) hits and the goodness of the track fit.

1608

1609 Selected tracks are clustered using a “deterministic annealing algorithm (DA)”<sup>4</sup>. A  
 1610 set of candidate vertices and their associated tracks, resulting from the DA, are then  
 1611 fitted with an “adaptive vertex fitter (AVF)” to produce the best estimate of the  
 1612 vertices locations.

1613

1614 The  $p_T$  of the several tracks associated to a reconstructed vertex is added, squared and  
 1615 used to organize the vertices; the vertex with the highest squared sum is designated  
 1616 as the *primary vertex (PV)* while the rest are designated as PU vertices.

1617 **Calorimeter clustering.**

1618 After traversing the CMS tracker system, electrons, photons and hadrons deposit their  
 1619 energy in the ECAL and HCAL cells. The PF clustering algorithm aims to provide  
 1620 a high detection efficiency even for low-energy particles and an efficient distinction  
 1621 between close energy deposits. The clustering runs independently in the ECAL barrel  
 1622 and endcaps, HCAL barrel and endcaps, and the two preshower layers, following two  
 1623 steps

1624     • cells with an energy larger than a given seed threshold and larger than the energy  
 1625         of the neighboring cells are identified as cluster seeds. The neighbor cells are  
 1626         those that either share a side with the cluster seed candidate, or the eight closest  
 1627         cells including cells that only share a corner with the seed candidate.

1628     • cells with at least a corner in common with a cell already in the cluster seed  
 1629         and with an energy above a cell threshold are grouped into topological clusters.

1630 Clusters formed in this way are known as *particle-flow clusters*. With this clustering  
 1631 strategy it is possible detect and measure the energy and direction of photons and

---

<sup>4</sup> DA algorithm and AVF are described in detail in references [95,96]

1632 neutral hadrons as well as differentiate these neutral particles from the charged hadron  
 1633 energy deposits. In cases involving charged hadrons for which the track parameters  
 1634 are not determined accurately, for instance low-quality and high-pT tracks, clustering  
 1635 helps in the energy measurements.

1636 **Electron track reconstruction.**

1637 Although the charged-particle track reconstruction described above works for elec-  
 1638 trons, they lose a significant fraction of their energy via bremsstrahlung photon radi-  
 1639 ation before reaching the ECAL; thus, the reconstruction performance depends on the  
 1640 ability to measure also the radiated energy. The reconstruction strategy in this case  
 1641 requires information from the tracking system and from the ECAL. Bremsstrahlung  
 1642 photons are emitted at similar  $\eta$  values to that of the electron but at different values  
 1643 of  $\phi$ ; therefore, the radiated energy can be recovered by grouping ECAL clusters in a  
 1644  $\eta$  window over a range of  $\phi$  around the electron direction. The group is called ECAL  
 1645 supercluster.

1646

1647 Electron candidates from the the track-seeding and ECAL superclustering are merged  
 1648 into a single collection which is submitted to a full electron tracking fit with a  
 1649 Gaussian-sum filter (GSF) [94]. The electron track and its associated ECAL su-  
 1650 percluster form a *particle-flow electron*.

1651 **Muon track reconstruction.**

1652 Given that the CMS detector is equipped with a muon spectrometer capable to iden-  
 1653 tify and measure the momentum of the muons traversing it, the muon reconstruction  
 1654 is not specific to PF; therefore, three different muon types are defined

- 1655     • *Standalone muon.* A clustering on the DTs or CSCs hits is performed to form  
 1656       track segments; those segments are used as seeds for the reconstruction in the  
 1657       muon spectrometer. All DTs, CSCs, and RPCs hits along the muon trajectory  
 1658       are combined and fitted to form the full track. The fitting output is called a  
 1659       *standalone-muon track*.
- 1660     • *tracker muon.* Each track in the inner tracker with  $p_T$  larger than 0.5 GeV and  
 1661       a total momentum  $p$  larger than 2.5 GeV is extrapolated to the muon system. A  
 1662       *tracker muon track* corresponds to the extrapolated tracks that match at least  
 1663       one muon segment.
- 1664     • *Global muon.* When tracks in the inner tracker (inner tracks) and standalone-  
 1665       muon tracks are matched and turns out being compatibles, their hits are com-  
 1666       bined and fitted to form a *global-muon track*.
- 1667     Global muons sharing the same inner track with tracker muons are merged into a  
 1668     single candidate. PF muon identification uses the muon energy deposits in ECAL,  
 1669     HCAL, and HO associated with the muon track to improve the muon identification.

1670     **4.4.1.1 Particle identification and reconstruction.**

1671     PF elements are connected by a linker algorithm that test the connection between any  
 1672       pair of elements, if they are found to be linked a geometrical distance that quantify  
 1673       the quality of the link is assigned; two elements may be linked indirectly through  
 1674       common elements. Linked elements form *PF blocks* and a PF block may contain  
 1675       elements originating in one or more particles. Links can be established between  
 1676       tracks, between calorimeter clusters, and between tracks and calorimeter clusters.  
 1677     The identification and reconstruction start with a PF block and proceeds as follows

- 1678     • Muons. An “isolated global muon” is identified by evaluating the presence of  
 1679       inner track and energy deposits close to the global muon track in the  $(\eta, \phi)$   
 1680       plane, i.e., in a particular point of the global muon track, inner tracks and  
 1681       energy deposits are sought within a radius of  $\Delta R = 0.3$  (see eqn. 3.7) from the  
 1682       muon track; if they exits and the  $p_T$  of the found track added to the  $E_T$  of the  
 1683       found energy deposit do not exceed 10% of the muon  $p_T$  then the global muon  
 1684       is an isolated global muon. This isolation condition is stringent enough to reject  
 1685       hadrons misidentified as muon.
- 1686       “Non-isolated global muons” are identified using additional selection require-  
 1687       ments on the number of track segments in the muon system and energy deposits  
 1688       along the muon track. Muons inside jets are identified with more stringent crite-  
 1689       ria in isolation and momentum as described in reference [97]. The PF elements  
 1690       associated to an identified muon are masked from the PF block.
- 1691     • Electrons are identified and reconstructed as described above plus some addi-  
 1692       tional requirements on fourteen variables like the amount of energy radiated,  
 1693       the distance between the extrapolated track position at the ECAL and the po-  
 1694       sition of the associated ECAL supercluster among others, which are combined  
 1695       in a specialized multivariate analysis strategy that improves the electron iden-  
 1696       tification. Tracks and clusters used to identify and reconstruct electrons are  
 1697       masked in the PF block.
- 1698     • Isolated photons are identified from ECAL superclusters with  $E_T$  larger than 10  
 1699       GeV, for which the energy deposited at a distance of 0.15, from the supercluster  
 1700       position on the  $(\eta, \phi)$  plane, does not exceed 10% of the supercluster energy;  
 1701       note that this is an isolation requirement. In addition, there must not be links  
 1702       to tracks. Clusters involved in the identification and reconstruction are masked

1703       in the PF block.

- 1704     ● Bremsstrahlung photons and prompt photons tend to convert to electron-positron  
 1705       pairs inside the tracker, therefore, a dedicated finder algorithm is used to link  
 1706       tracks that seem to originate from a photon conversion; in case those two tracks  
 1707       are compatible with direction of a bremsstrahlung photon, they are also linked  
 1708       to the original electron track. Photon conversion tracks are also masked in the  
 1709       PF block.
- 1710     ● The remaining elements in the PF block are used to identify hadrons. In the  
 1711       region  $|\eta| \leq 2.5$ , neutral hadrons are identified with HCAL clusters not linked  
 1712       to any track while photons from neutral pion decays are identified with ECAL  
 1713       clusters without links to tracks. In the region  $|\eta| > 2.5$  ECAL clusters linked to  
 1714       HCAL clusters are identified with a charged or neutral hadron shower; ECAL  
 1715       clusters with no links are identified with photons. HCAL clusters not used yet,  
 1716       are linked to one or more unlinked tracks and to an unlinked ECAL in order to  
 1717       reconstruct charged-hadrons or a combination of photons and neutral hadrons  
 1718       according to certain conditions on the calibrated calorimetric energy.
- 1719     ● Charged-particle tracks may be liked together when they converge to a “sec-  
 1720       ondary vertex (SV) ” displaced from the interaction point where the PV and  
 1721       PU vertices are reconstructed; at least three tracks are needed in that case,  
 1722       of which at most one has to be an incoming track with hits in tracker region  
 1723       between a PV and the SV.

1724

1725 Once individual particles are identified and reconstructed, they can be grouped into  
 1726 larger objects to form jets. The reconstruction of all the particles in an event is used

1727 to determine the presence of neutrinos represented by an imbalance in the transverse  
1728 energy.

1729 The linker algorithm, and the whole PF algorithm, has been validated and commis-  
1730 sioned; results from that validation are presented in the references [91].

1731 , jets reco, anti-kt algorithm, jet energy corrections, btagging, MET

1732 **4.5 MVA methods, NN, BDT, boosting,  
1733 overtraining, variable ranking**

1734 **4.6 statistical inference, likelihood  
1735 parametrization**

1736 **4.7 nuisance parameters**

1737 **4.8 exclusion limits**

1738 **4.9 asymptotic limits**

<sup>1739</sup> **Chapter 5**

<sup>1740</sup> **Search for production of a Higgs**

<sup>1741</sup> **boson and a single top quark in**

<sup>1742</sup> **multilepton final states in pp**

<sup>1743</sup> **collisions at  $\sqrt{s} = 13$  TeV**

<sup>1744</sup> **5.1 Introduction**

<sup>1745</sup> Dont forget to mention previous constrains to ct check reference ?? and references

<sup>1746</sup> <https://link.springer.com/content/pdf/10.1007%2FJHEP01>

<sup>1747</sup> A. Azatov, R. Contino and J. Galloway,  $\rightarrow$ IJModel-Independent Bounds on a

<sup>1748</sup> Light Higgs, $\rightarrow$ JHEP 1204 (2012) 127 [arXiv:1202.3415 [hep-ph]].

<sup>1749</sup> J. R. Espinosa, C. Grojean, M. Muhlleitner and M. Trott,  $\rightarrow$ IJFingerprinting

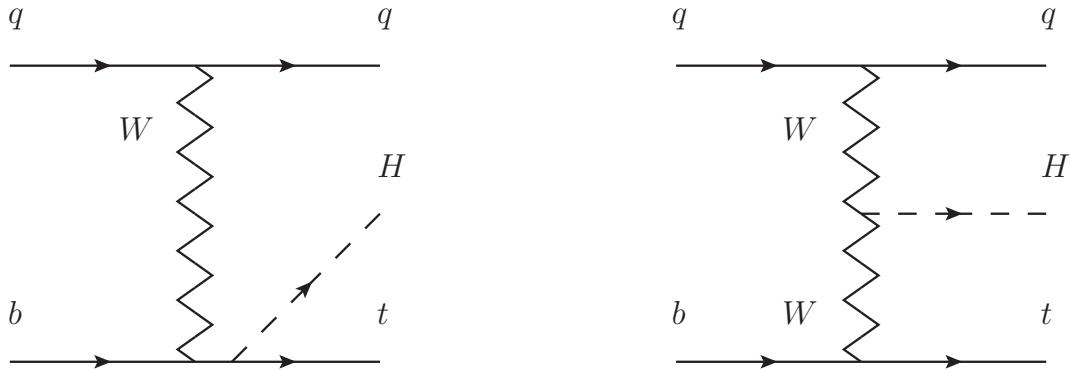
<sup>1750</sup> Higgs Suspects at the LHC, $\rightarrow$ JHEP 1205 (2012) 097 [arXiv:1202.3697 [hep-ph]].

<sup>1751</sup> This chapter present the search for the associated production of a Higgs boson and

<sup>1752</sup> a single top quark events with three leptons in the final state, targeting Higgs decay

1753 modes to  $WW$ ,  $ZZ$ , and  $\tau\tau$ . The analysis uses the 13 TeV dataset produced in 2016,  
 1754 corresponding to an integrated luminosity of  $35.9\text{fb}^{-1}$ . It is based on and expands  
 1755 previous analyses at 8 TeV [98, 99] and searches for associated production of  $t\bar{t}$  and  
 1756 Higgs in the same channel [100], and complements searches in other decay channels  
 1757 targeting  $H \rightarrow b\bar{b}$  [101].

1758 As showed in section 2.4, the cross section of the associated production of a Higgs  
 1759 boson and a single top quark ( $tHq$ ) process is driven by a destructive interference of  
 1760 two contributions (see Figure 5.1), where the Higgs couples to either the W boson or  
 1761 the top quark. Any deviation from the standard model (SM) in the Higgs coupling  
 1762 structure could therefore lead to a large enhancement of the cross section, making  
 1763 this analysis sensitive to such deviations. A second process, where the Higgs and  
 1764 top quark are accompanied by a W boson ( $tHW$ ) has similar behavior, albeit with a  
 1765 weaker interference pattern.



**Figure 5.1:** The two leading-order diagrams of  $tHq$  production.

1766 We selects events with three leptons and a  $b$  tagged jet in the final state. The  $tHq$   
 1767 signal contribution is then determined in a fit of the observed data to two multivariate  
 1768 classifier outputs, each trained to discriminate against one of the two dominant back-  
 1769 grounds of events with non-prompt leptons from  $t\bar{t}$  and of associated production of  $t\bar{t}$

1770 and vector bosons ( $t\bar{t}W$ ,  $t\bar{t}Z$ ). The fit result is then used to set an upper limit on the  
 1771 combined  $t\bar{t}H$ ,  $tHq$  and  $tHW$  production cross section, as a function of the relative  
 1772 coupling strengths of Higgs and top quark and Higgs and W boson, respectively.

## 1773 5.2 Data and MC Samples

1774 The data considered in this analysis were collected by the CMS experiment dur-  
 1775 ing 2016 and correspond to a total integrated luminosity of  $35.9\text{fb}^{-1}$ . Only periods  
 1776 when the CMS magnet was on were considered when selecting the data samples, that  
 1777 corresponds to the 23 Sep 2016 (Run B to G) and PromptReco (Run H) versions  
 1778 of the datasets. The MC samples used in this analysis correspond to the RunI-  
 1779 ISummer16MiniAODv2 campaign produced with CMSSW 80X. The two signal sam-  
 1780 ples (for  $tHq$  and  $tHW$ ) were produced with MG5\_aMC@NLO (version 5.222), in  
 1781 leading-order mode, and are normalized to next-to-leading-order cross sections,  
 1782 see Tab. 5.1. Each sample is generated with a set of event weights corresponding to  
 1783 different values of  $\kappa_t$  and  $\kappa_V$  couplings as shown in Tab. 5.2.

### 1784 5.2.1 Full 2016 dataset and MC samples

Sample	$\sigma$ [pb]	BF
/THQ_Hincl_13TeV-madgraph-pythia8_TuneCUETP8M1/	0.7927	0.324
/THW_Hincl_13TeV-madgraph-pythia8_TuneCUETP8M1/	0.1472	1.0

**Table 5.1:** Signal samples and their cross section and branching fraction used in this analysis. See Ref. [102] for more details.

1785 Different MC generators were used to generate the background processes. The  
 1786 dominant sources ( $t\bar{t}$ ,  $t\bar{t}W$ ,  $t\bar{t}Z$ ,  $t\bar{t}H$ ) were produced using AMC@NLO interfaced  
 1787 to PYTHIA8, and are scaled to NLO cross sections. Other processes are simulated

		<i>tHq</i>		<i>tHW</i>		
$\kappa_V$	$\kappa_t$	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 5.2:**  $\kappa_V$  and  $\kappa_t$  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). See also Ref. [102].

Sample	$\sigma$ [pb]
TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.2043
TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.2529
tthJetToNonbb_M125_13TeV_amcatnloFXFX_madspin_pythia8_mWCutfix /store/cmst3/group/susy/gpetrucc/13TeV/u/TTLL_m1to10_L0_NoMS_for76X/	0.2151 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

**Table 5.3:** List of background samples used in this analysis (CMSSW 80X). In the first section of the table are listed the samples of the processes for which we use the simulation to extract the final yields and shapes, in the second section the samples of the processes we will estimate from data. The MC simulation is used to design the data driven methods and derive the associated systematics.

Sample	$\sigma$ [pb]
ttWJets_13TeV_madgraphMLM	0.6105
ttZJets_13TeV_madgraphMLM	0.5297/0.692

**Table 5.4:** Leading-order  $t\bar{t}W$  and  $t\bar{t}Z$  samples used in the signal BDT training.

1788 using POWHEG interfaced to PYTHIA, or bare PYTHIA (see table 5.3 and [100]  
1789 for more details).

Three lepton and Four 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 5.5:** Table of high-level triggers that we consider in the analysis.

## 1790 5.2.2 Triggers

1791 We consider online-reconstructed events triggered by one, two, or three leptons.  
 1792 Single-lepton triggers are included to boost the acceptance of events where the  $p_T$  of  
 1793 the sub-leading lepton falls below the threshold of the double-lepton triggers. Ad-  
 1794 ditionally, by including double-lepton triggers in the  $\geq 3$  lepton category, as well  
 1795 as single-lepton triggers in all categories, we increase the efficiency, considering the  
 1796 logical “or” of the trigger decisions of all the individual triggers in a given category.  
 1797 Tab. 5.5 shows the lowest-threshold non-prescaled triggers present in the High-Level  
 1798 Trigger (HLT) menus for both Monte-Carlo and data in 2016.

### 1799 5.2.2.1 Trigger efficiency scale factors

1800 The efficiency of events to pass the trigger is measured in simulation (trivially using  
 1801 generator information) and in the data (using event collected by an uncorrelated

Category	Scale Factor
ee	$1.01 \pm 0.02$
e $\mu$	$1.01 \pm 0.01$
$\mu\mu$	$1.00 \pm 0.01$
3l	$1.00 \pm 0.03$

**Table 5.6:** Trigger efficiency scale factors and associated uncertainties, shown here rounded to the nearest percent.

1802 MET trigger). Small differences between the data and MC efficiencies are corrected  
 1803 by applying scale factors as shown in Tab. 5.6. The exact procedure and control plots  
 1804 are documented in [103] for the current analysis.

## 1805 5.3 Object Identification and event selection

### 1806 5.3.1 Jets and $b$ tagging

1807 The analysis uses anti- $k_t$  (0.4) particle-flow (PF) jets, corrected for charged hadrons  
 1808 not coming from the primary vertex (charged hadron subtraction), and having jet  
 1809 energy corrections (`Summer16_23Sep2016V3`) applied as a function of the jet  $E_T$  and  
 1810  $\eta$ . Jets are only considered if they have a transverse energy above 25GeV.

1811 In addition, they are required to be separated from any lepton candidates passing  
 1812 the fakeable object selections (see Tables 5.7 and 5.8) by  $\Delta R > 0.4$ .

1813 The loose and medium working points of the CSV b-tagging algorithm are used to  
 1814 identify  $b$  jets. Data/simulation differences in the  $b$  tagging performance are corrected  
 1815 by applying per-jet weights to the simulation, dependent on the jet  $p_T$ , eta,  $b$  tagging  
 1816 discriminator, and flavor (from simulation truth) [104]. The per-event weight is taken  
 1817 as the product of the per-jet weights, including those of the jets associated to the  
 1818 leptons. More details can be found in the corresponding  $t\bar{t}H$  documentation [100,103].

1819 **5.3.2 Lepton selection**

Cut	Loose	Fakeable object	Tight
$ \eta  < 2.4$	✓	✓	✓
$p_T$	$> 5\text{GeV}$	$> 15\text{GeV}$	$> 15\text{GeV}$
$ d_{xy}  < 0.05 \text{ (cm)}$	✓	✓	✓
$ d_z  < 0.1 \text{ (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 5.7:** 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. For the two  $p_T^{\text{ratio}}$  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.

1820 The lepton reconstruction and selection is identical to that used in the  $t\bar{t}H$  mul-  
 1821 tilepton analysis, as documented in Refs. [100, 103]. For details on the reconstruction  
 1822 algorithms, isolation, pileup mitigation, and a description of the lepton MVA discrim-  
 1823 inator and validation plots thereof, we refer to that document since they are out of  
 1824 the scope of this thesis. Three different selections are defined both for the electron  
 1825 and muon object identification: the *Loose*, *Fakeable Object*, and *Tight* selection. As  
 1826 described in more detail later, these are used for event level vetoes, the fake rate  
 1827 estimation application region, and the final signal selection, respectively. The  $p_T$  of  
 1828 fakeable objects is defined as  $0.85 \times p_T(\text{jet})$ , where the jet is the one associated to the  
 1829 lepton object. This mitigates the dependence of the fake rate on the momentum of  
 1830 the fakeable object and thereby improves the precision of the method.

1831 Tables 5.7 and 5.8 list the full criteria for the different selections of muons and  
 1832 electrons.

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

**Table 5.8:** 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  $\eta$  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<sub>1</sub>, value<sub>2</sub>, value<sub>3</sub>).

### 1833 5.3.3 Lepton selection efficiency

1834 Efficiencies of reconstruction and selecting loose leptons are measured both for muons  
 1835 and electrons using a tag and probe method on both data and MC, using  $Z \rightarrow \ell^+ \ell^-$ .  
 1836 Corresponding scale factors are derived from the ratio of efficiencies and applied to the  
 1837 selected These. Events are produced for the leptonic SUSY analyses using equivalent  
 1838 lepton selections and recycled for the  $t\bar{t}H$  analysis as well as for this analysis. The  
 1839 efficiencies of applying the tight selection as defined in Tables 5.7 and 5.8, on the  
 1840 loose leptons are determined again by using a tag and probe method on a sample of  
 1841 DY-enriched events. They are documented for the  $t\bar{t}H$  analysis in Ref. [103] and are  
 1842 exactly equivalent for this analysis.

## 1843 5.4 Background predictions

1844 The modeling of reducible and irreducible backgrounds in this analysis uses the exact  
1845 methods, analysis code, and ROOT trees used for the  $t\bar{t}H$  multilepton analysis. We  
1846 give a brief description of the methods and refer to the documentation of that analysis  
1847 in Refs. [100, 103] for any details.

1848 The backgrounds in three-lepton final states can be split in two broad categories:  
1849 irreducible backgrounds with genuine prompt leptons (i.e. from on-shell W and Z  
1850 boson decays); and reducible backgrounds where at least one of the leptons is “non-  
1851 prompt”, i.e. produced within a hadronic jet, either a genuine lepton from heavy  
1852 flavor decays, or simply mis-reconstructed jets.

1853 Irreducible backgrounds can be reliably estimated directly from Monte-Carlo sim-  
1854 ulated events, using higher-order cross sections or data control regions for the overall  
1855 normalization. This is done in this analysis for all backgrounds involving prompt lep-  
1856 tons:  $t\bar{t}W$ ,  $t\bar{t}Z$ ,  $t\bar{t}H$ ,  $WZ$ ,  $ZZ$ ,  $W^\pm W^\pm qq$ ,  $t\bar{t}t\bar{t}$ ,  $tZq$ ,  $tZW$ ,  $WWW$ ,  $WWZ$ ,  $WZZ$ ,  
1857  $ZZZ$ .

1858 Reducible backgrounds, on the other hand, are not well predicted by simulation,  
1859 and are estimated using data-driven methods. In the case of non-prompt leptons, a  
1860 fake rate method is used, where the contribution to the final selection is estimated by  
1861 extrapolating from a sideband (or “application region”) with a looser lepton definition  
1862 (the fakeable object definitions in Tabs. 5.7 and 5.8) to the signal selection. The tight-  
1863 to-loose ratios (or “fake rates”) are measured in several background dominated data  
1864 events with dedicated triggers, subtracting the residual prompt lepton contribution  
1865 using MC. Non-prompt leptons in our signal regions are predominantly produced in  $t\bar{t}$   
1866 events, with a much smaller contribution, from Drell–Yan production. The systematic  
1867 uncertainty on the normalization of the non-prompt background estimation is on the

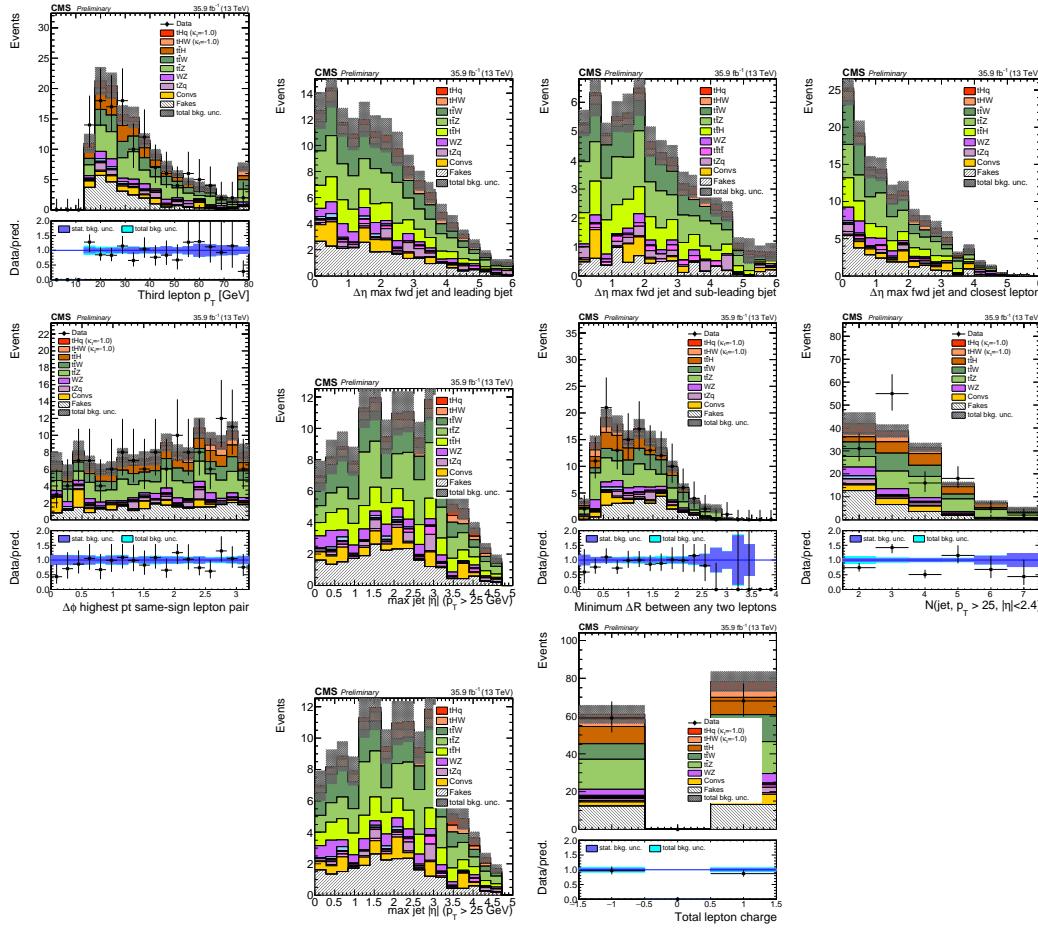
1868 order of 50%, and thereby one of the dominant limitations on the performance of  
 1869 multilepton analyses in general and this analysis in particular. It consists of several  
 1870 individual sources, such as the result of closure tests of the method using simulated  
 1871 events, limited statistics in the data control regions due to necessary prescaling of  
 1872 lepton triggers, and the uncertainty in the subtraction of residual prompt leptons  
 1873 from the control region.

1874 The fake background where the leptons pass the looser selection are weighted  
 1875 according to how many of them fail the tight criteria. Events with a single failing  
 1876 lepton are weighted with the factor  $f/(1-f)$  for the estimate to the tight selection  
 1877 region, where  $f$  is the fake rate. Events with two failing leptons are given the negative  
 1878 weight  $-f_i f_j / (1 - f_i)(1 - f_j)$ , and for three leptons the weight is positive and equal  
 1879 to the product of  $f/(1-f)$  factor evaluated for each failing lepton.

1880 Figures 5.2 show the distributions of some relevant kinematic variables, normalized  
 1881 to the cross section of the respective processes and to the integrated luminosity.

## 1882 5.5 Signal discrimination

1883 The  $tHq$  signal is separated from the main backgrounds using a boosted decision  
 1884 tree (BDT) classifier, trained on simulated signal and background events. A set of  
 1885 discriminating variables are given as input to the BDT which produces a output  
 1886 distribution maximizing the discrimination power. Table 5.9 lists the input variables  
 1887 used while Figures 5.3 show their distributions for the relevant signal and background  
 1888 samples, for the three lepton channel. Two BDT classifiers are trained for the two  
 1889 main backgrounds expected in the analysis: events with prompt leptons from  $t\bar{t}W$  and  
 1890  $t\bar{t}Z$  (also referred to as  $t\bar{t}V$ ), and events with non-prompt leptons from  $t\bar{t}$ . The datasets  
 1891 used in the training are the  $tHq$  signal (see Tab. 5.1), and LO MADGRAPH samples



**Figure 5.2:** Distributions of input variables to the BDT for signal discrimination, three lepton channel, normalized to their cross section and to  $35.9\text{fb}^{-1}$ .

of  $t\bar{t}W$  and  $t\bar{t}Z$ , in an admixture proportional to their respective cross sections (see Tab. 5.4).

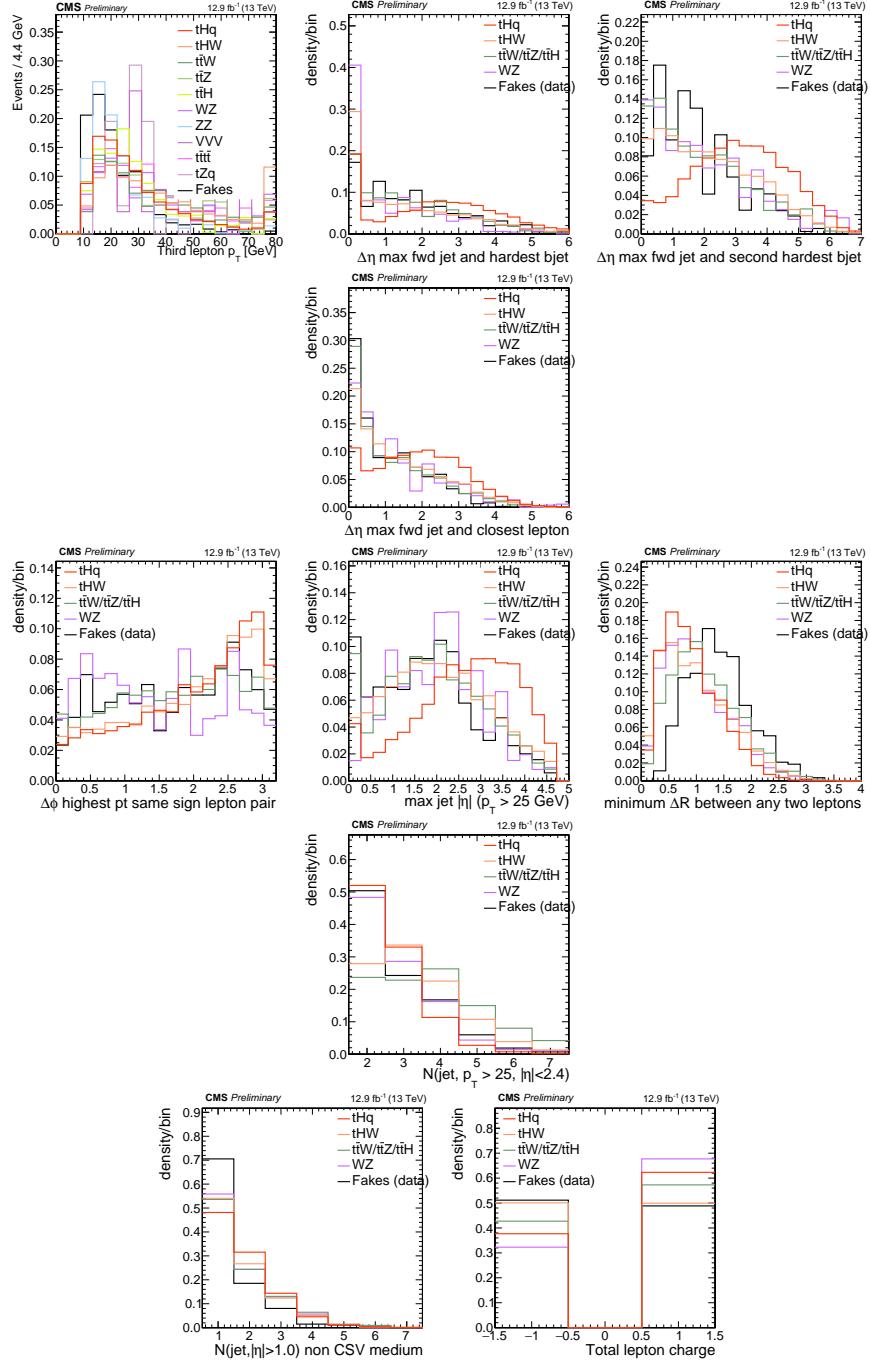
The MVA analysis consist of two stages: first a “training” where the MVA method is trained to discriminate between simulated signal and background events, then a “test” stage where the trained algorithm is used to classify different events from the samples. The sample is obtained from a pre-selection (see Tab. ?? with pre-selection cuts). Figures 5.4 show the input variables distributions as seen by the MVA algorithm. Note that in contrast to the distributions in Fig. 5.3 only the main backgrounds ( $t\bar{t}$  from simulation,  $t\bar{t}V$ ) are included.

Variable name	Description
nJet25	Number of jets with $p_T > 25$ GeV, $ \eta  < 2.4$
MaxEtaJet25	Maximum $ \eta $ of any (non-CSV-loose) jet with $p_T > 25$ GeV
totCharge	Sum of lepton charges
nJetEta1	Number of jets with $ \eta  > 1.0$ , non-CSV-loose
detaFwdJetBJet	$\Delta\eta$ between forward light jet and hardest CSV loose jet
detaFwdJet2BJet	$\Delta\eta$ between forward light jet and second hardest CSV loose jet (defaults to -1 in events with only one CSV loose jet)
detaFwdJetClosestLep	$\Delta\eta$ between forward light jet and closest lepton
dphiHighestPtSSPair	$\Delta\phi$ of highest $p_T$ same-sign lepton pair
minDRll	minimum $\Delta R$ between any two leptons
Lep3Pt/Lep2Pt	$p_T$ of the 3 <sup>rd</sup> lepton (2 <sup>nd</sup> for ss2l)

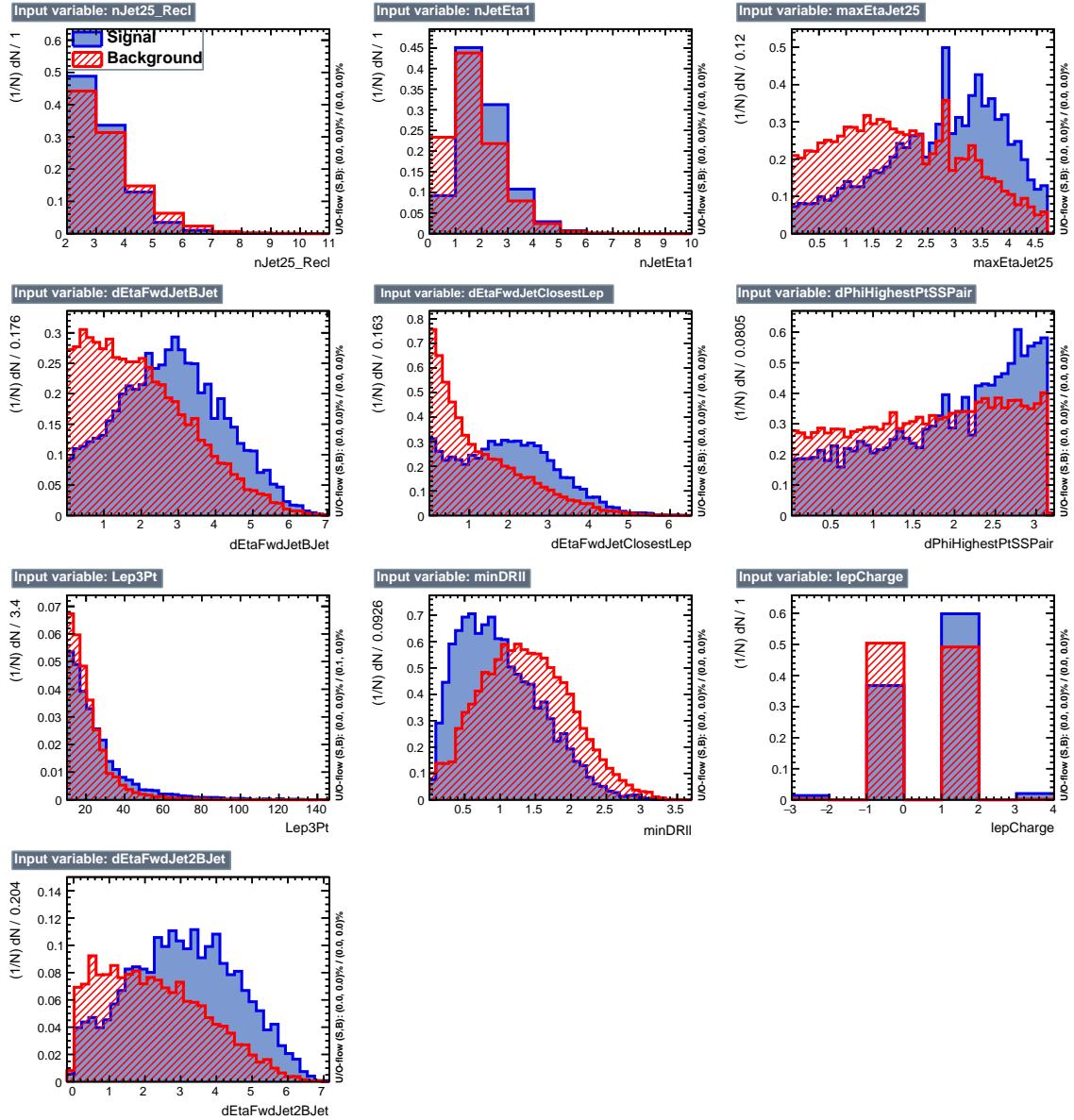
**Table 5.9:** MVA input discriminating variables

1901 Note that splitting the training in two groups reveals that some variables show  
 1902 opposite behavior for the two background sources; potentially screening the discrimi-  
 1903 nation power if they were to be used in a single discriminant. For some other variables  
 1904 the distributions are similar in both background cases.

1905 From table 5.9, it is clear that the input variables are correlated to some extend.  
 1906 These correlations play an important role for some MVA methods like the Fisher  
 1907 discriminant method in which the first step consist of performing a linear transfor-  
 1908 mation to an phase space where the correlations between variables are removed. In  
 1909 case a boosted decision tree (BDT) method however, correlations do not affect the  
 1910 performance. Figure 5.6 show the linear correlation coefficients for signal and back-  
 1911 ground for the two training cases (the signal values are identical by construction). As  
 1912 expected, strong correlations appears for variables related to the forward jet activity.  
 1913 Same trend is seen in case of the same sign dilepton channel in Figure ??.



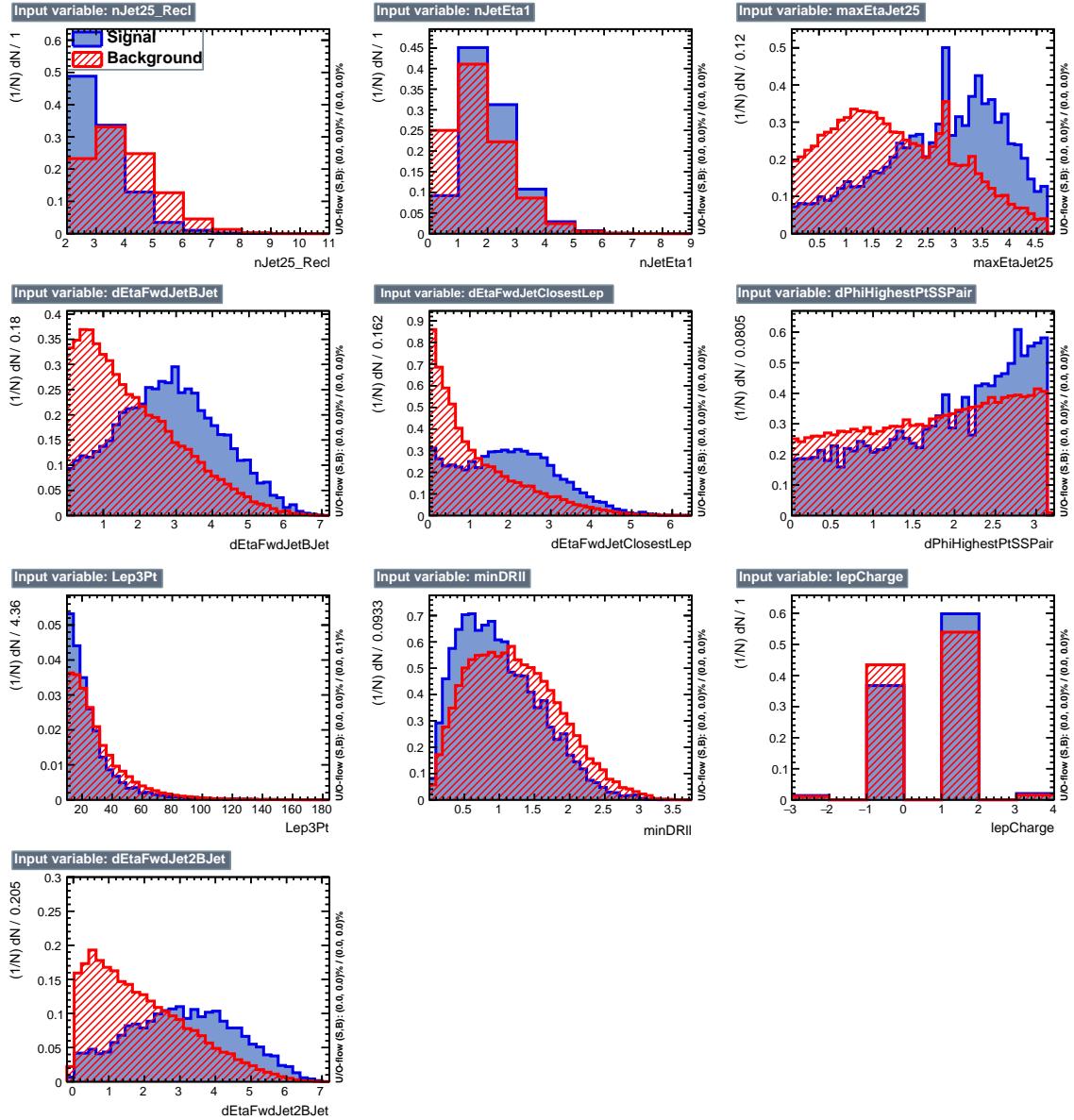
**Figure 5.3:** Distributions of input variables to the BDT for signal discrimination, three lepton channel.



**Figure 5.4:** BDT inputs as seen by TMVA (signal, in blue, is  $tHq$ , background, in red, is  $t\bar{t}$ ) for the three lepton channel, discriminated against  $t\bar{t}$  (fakes) background.

### 1914 5.5.1 Classifiers response

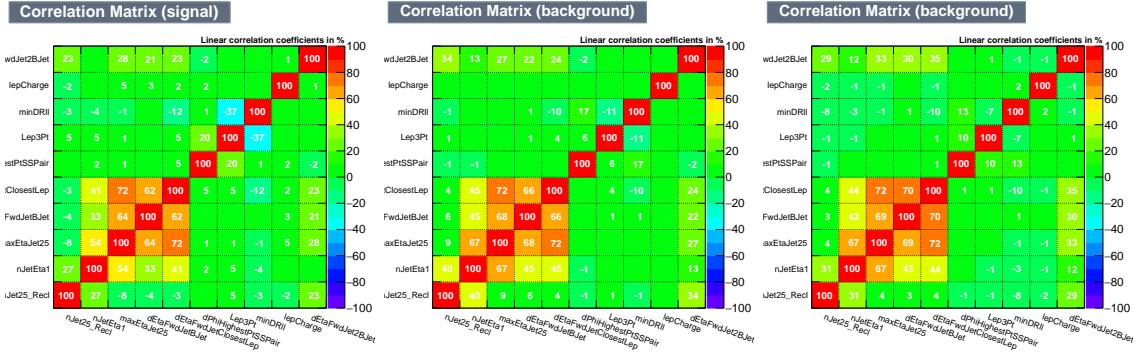
1915 Several MVA algorithms were evaluated to determine the most appropriate method  
 1916 for this analysis. The plots in Fig. 5.7 (top) show the background rejection as a  
 1917 function of the signal efficiency for  $t\bar{t}$  and  $t\bar{t}V$  trainings (ROC curves) for the different



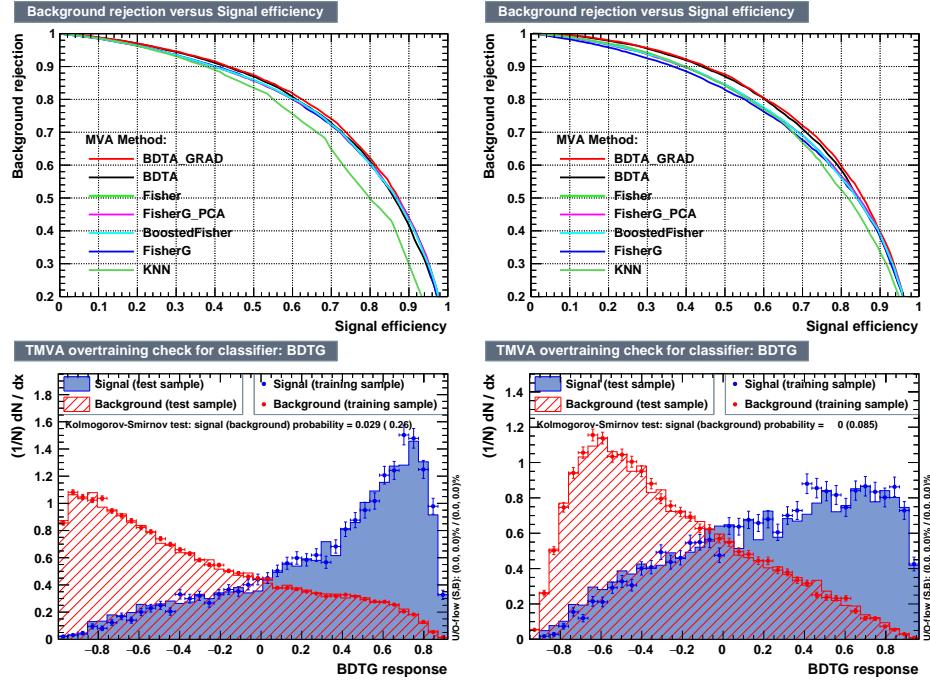
**Figure 5.5:** BDT inputs as seen by TMVA (signal, in blue, is  $tHq$ , background, in red, is  $t\bar{t}W+t\bar{t}Z$ ) for the three lepton channel, discriminated against  $t\bar{t}V$  background.

1918 algorithms that were evaluated.

1919 In both cases the gradient boosted decision tree (“BDTA\_GRAD”) classifier offers  
 1920 the best results, followed by an adaptive BDT classifier (“BDTA”). The BDTA\_GRAD  
 1921 classifier output distributions for signal and backgrounds are shown on the bottom of  
 1922 Fig. 5.7. As expected, a good discrimination power is obtained using default discrim-



**Figure 5.6:** Signal (left),  $t\bar{t}$  background (middle), and  $t\bar{t}V$  background (right.) correlation matrices for the input variables in the TMVA analysis for the three lepton channel.



**Figure 5.7:** Top: background rejection vs signal efficiency (ROC curves) for various MVA classifiers (top) in the three lepton channel against  $t\bar{t}V$  (left) and  $t\bar{t}$  (right). Bottom: classifier output distributions for the gradient boosted decision trees, for training against  $t\bar{t}V$  (left) and against  $t\bar{t}$  (right).

1923 inator parameter values, with minimal overtraining. TMVA provides a ranking of the  
 1924 input variables by their importance in the classification process, shown in Tab. 5.10.  
 1925 The TMVA settings used in the BDT training are shown in Tab. 5.11.

ttbar training			ttV training		
Rank	Variable	Importance	Variable	Importance	
1	minDRll	1.329e-01	dEtaFwdJetBJet	1.264e-01	
2	dEtaFwdJetClosestLep	1.294e-01	Lep3Pt	1.224e-01	
3	dEtaFwdJetBJet	1.209e-01	maxEtaJet25	1.221e-01	
4	dPhiHighestPtSSPair	1.192e-01	dEtaFwdJet2BJet	1.204e-01	
5	Lep3Pt	1.158e-01	dEtaFwdJetClosestLep	1.177e-01	
6	maxEtaJet25	1.121e-01	minDRll	1.143e-01	
7	dEtaFwdJet2BJet	9.363e-02	dPhiHighestPtSSPair	9.777e-02	
8	nJetEta1	6.730e-02	nJet25_Recl	9.034e-02	
9	nJet25_Recl	6.178e-02	nJetEta1	4.749e-02	
10	lepCharge	4.701e-02	lepCharge	4.116e-02	

**Table 5.10:** TMVA input variables ranking for BDTA\_GRAD method for the trainings in the three lepton channel. For both trainings the rankings show almost the same 5 variables in the first places.

---

```

TMVA.Types.kBDT
NTrees=800
BoostType=Grad
Shrinkage=0.10
!UseBaggedGrad
nCuts=50
MaxDepth=3
NegWeightTreatment=PairNegWeightsGlobal
CreateMVAPdfs

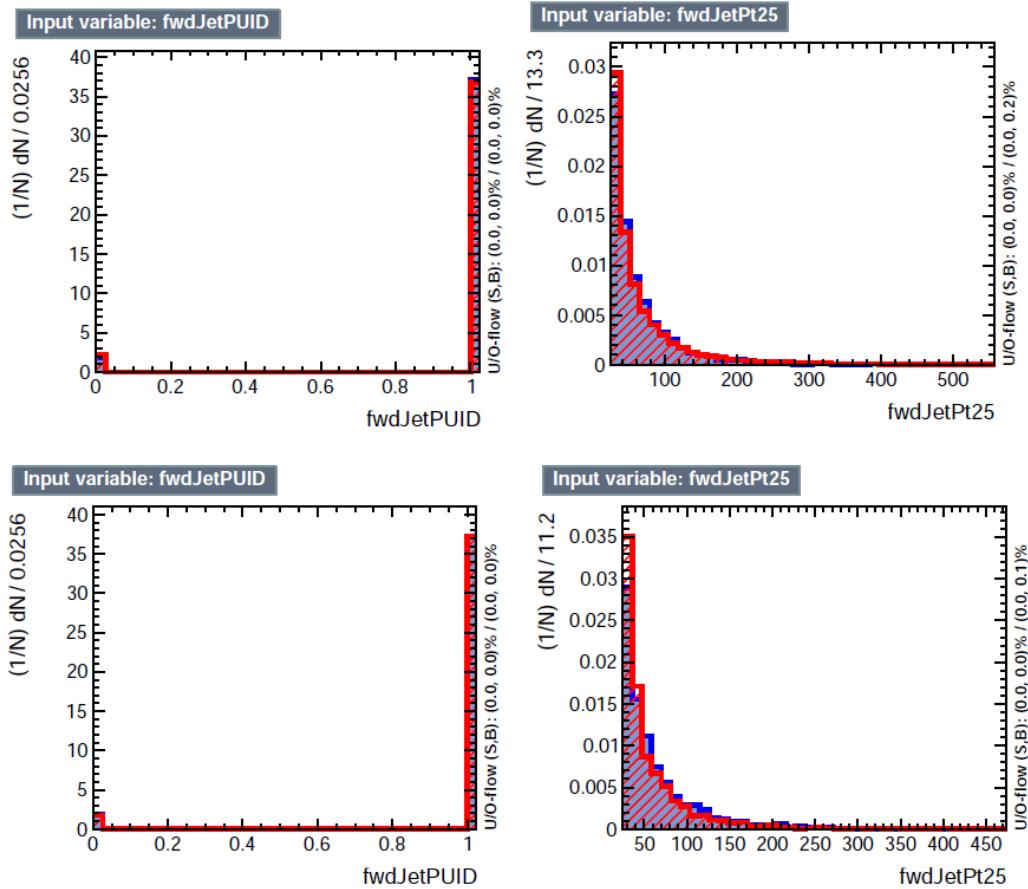
```

---

**Table 5.11:** TMVA configuration used in the BDT training.

## 1926 5.6 Additional discriminating variables

1927 Two additional discriminating variables were tested considering the fact that the  
 1928 forward jet in the background could come from the pileup; since we have a real  
 1929 forward jet in the signal, it could give some improvement in the discriminating power.  
 1930 The additional variables describe the forward jet momentum (fwdJetPt25) and the  
 1931 forward jet identification(fwdJetPUID). Distributions for these variables in the three  
 1932 lepton channel are shown in the figure 5.8. The forward jet identification distribution  
 1933 show that for both, signal and background, jets are mostly real jets.  
 1934 The testing was made including in the MVA input one variable at a time, so we



**Figure 5.8:** Additional discriminating variables distributions for ttV training (Top row) and tt training (bottom row) in the three lepton channel. The origin of the jets in the forward jet identification distribution is tagged as 0 for “pileup jets” while “real jets” are tagged as 1.

1935 can evaluate the discrimination power of each variable, and then both simultaneously.  
 1936 fwdJetPUID was ranked in the last place in importance (11) in both training (ttV  
 1937 and tt) while fwdJetPt25 was ranked 3 in the ttV training and 7 in the tt training.  
 1938 When training using 12 variables, fwdJetPt25 was ranked 5 and 7 in the ttV and tt  
 1939 trainings respectively, while fwdJetPUID was ranked 12 in both cases.

1940 The improvement in the discrimination performance provided by the additional  
 1941 variables is about 1%, so it was decided not to include them in the procedure. Table  
 1942 5.12 show the ROC-integral for all the testing cases we made.

ROC-integral	
base 10 var ttv	0.848
+ fwdJetPUID ttv	0.849
+ fwdJetPt25 ttv	0.856
12 var ttv	0.856
<hr/>	
base 10 var tt	0.777
+ fwdJetPUID tt	0.777
+ fwdJetPt25 tt	0.787
12 var	0.787

**Table 5.12:** ROC-integral for all the testing cases we made in the evaluation of the additional variables discriminating power. The improvement in the discrimination performance provided by the additional variables is about 1%

<sup>1943</sup> **References**

- 1944 [1] J. Schwinger. "Quantum Electrodynamics. I. A Covariant Formulation". Physical Review. 74 (10): 1439-61, (1948).
- 1946 [2] R. P. Feynman. "Space-Time Approach to Quantum Electrodynamics". Physical Review. 76 (6): 769-89, (1949).
- 1948 [3] S. Tomonaga. "On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields". Progress of Theoretical Physics. 1 (2): 27-42, (1946).
- 1950 [4] D.J. Griffiths, "Introduction to electrodynamics". 4th ed. Pearson, (2013).
- 1951 [5] F. Mandl, G. Shaw. "Quantum field theory." Chichester: Wiley (2009).
- 1952 [6] F. Halzen, and A.D. Martin, "Quarks and leptons: An introductory course in modern particle physics". New York: Wiley, (1984) .
- 1954 [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>.
- 1958 [8] E. Noether, "Invariante Variationsprobleme", Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, mathematisch-physikalische Klasse, vol. 1918, pp. 235-257, (1918).

- 1961 [9] C. Patrignani et al. (Particle Data Group), Chin. Phys. C, 40, 100001 (2016)  
 1962 and 2017 update.
- 1963 [10] M. Goldhaber, L. Grodzins, A.W. Sunyar “Helicity of Neutrinos”, Phys. Rev.  
 1964 109, 1015 (1958).
- 1965 [11] Palanque-Delabrouille N et al. “Neutrino masses and cosmology with Lyman-  
 1966 alpha forest power spectrum”, JCAP 11 011 (2015).
- 1967 [12] M. Gell-Mann. “A Schematic Model of Baryons and Mesons”. Physics Letters.  
 1968 8 (3): 214-215 (1964).
- 1969 [13] G. Zweig. “An SU(3) Model for Strong Interaction Symmetry and its Breaking”  
 1970 (PDF). CERN Report No.8182/TH.401 (1964).
- 1971 [14] G. Zweig. “An SU(3) Model for Strong Interaction Symmetry and its Breaking:  
 1972 II” (PDF). CERN Report No.8419/TH.412(1964).
- 1973 [15] M. Gell-Mann. “The Interpretation of the New Particles as Displaced Charged  
 1974 Multiplets”. Il Nuovo Cimento 4: 848. (1956).
- 1975 [16] T. Nakano, K, Nishijima. “Charge Independence for V-particles”. Progress of  
 1976 Theoretical Physics 10 (5): 581-582. (1953).
- 1977 [17] N. Cabibbo, “Unitary symmetry and leptonic decays” Physical Review Letters,  
 1978 vol. 10, no. 12, p. 531, (1963).
- 1979 [18] M.Kobayashi, T.Maskawa, “CP-violation in the renormalizable theory of weak  
 1980 interaction,” Progress of Theoretical Physics, vol. 49, no. 2, pp. 652-657, (1973).
- 1981 [19] File: Weak Decay (flipped).svg. (2017, June 12). Wikimedia Com-  
 1982 mons, the free media repository. Retrieved November 27, 2017 from

- 1983 https://commons.wikimedia.org/w/index.php?title=File:Weak\_Decay\_(flipped)  
 1984 .svg&oldid=247498592.
- 1985 [20] Georgia Tech University. Coupling Constants for the Fundamental  
 1986 Forces(2005). Retrieved January 10, 2018, from http://hyperphysics.phy-  
 1987 astr.gsu.edu/hbase/Forces/couple.html#c2
- 1988 [21] M. Strassler. (May 31, 2013).The Strengths of the Known Forces. Retrieved Jan-  
 1989 uary 10, 2018, from https://profmattstrassler.com/articles-and-posts/particle-  
 1990 physics-basics/the-known-forces-of-nature/the-strength-of-the-known-forces/
- 1991 [22] S.L. Glashow. “Partial symmetries of weak interactions”, Nucl. Phys. 22 579-  
 1992 588, (1961).
- 1993 [23] A. Salam, J.C. Ward. “Electromagnetic and weak interactions”, Physics Letters  
 1994 13 168-171, (1964).
- 1995 [24] S. Weinberg, “A model of leptons”, Physical Review Letters, vol. 19, no. 21, p.  
 1996 1264, (1967).
- 1997 [25] M. Peskin, D. Schroeder, “An introduction to quantum field theory”. Perseus  
 1998 Books Publishing L.L.C., (1995).
- 1999 [26] A. Pich. “The Standard Model of Electroweak Interactions”  
 2000 https://arxiv.org/abs/1201.0537
- 2001 [27] F.Bellaiche. (2012, 2 September). “What’s this Higgs boson anyway?”. Retrieved  
 2002 from: https://www.quantum-bits.org/?p=233
- 2003 [28] M. Endres et al. Nature 487, 454-458 (2012) doi:10.1038/nature11255

- 2004 [29] F. Englert, R. Brout. “Broken Symmetry and the Mass of Gauge  
2005 Vector Mesons”. Physical Review Letters. 13 (9): 321-23.(1964)  
2006 doi:10.1103/PhysRevLett.13.321
- 2007 [30] P.Higgs. “Broken Symmetries and the Masses of Gauge Bosons”. Physical Re-  
2008 view Letters. 13 (16): 508-509,(1964). doi:10.1103/PhysRevLett.13.508.
- 2009 [31] G.Guralnik, C.R. Hagen and T.W.B. Kibble. “Global Conservation Laws  
2010 and Massless Particles”. Physical Review Letters. 13 (20): 585-587, (1964).  
2011 doi:10.1103/PhysRevLett.13.585.
- 2012 [32] CMS collaboration. “Observation of a new boson at a mass of 125 GeV with  
2013 the CMS experiment at the LHC”. Physics Letters B. 716 (1): 30-61 (2012).  
2014 arXiv:1207.7235. doi:10.1016/j.physletb.2012.08.021
- 2015 [33] ATLAS collaboration. “Observation of a New Particle in the Search for the Stan-  
2016 dard Model Higgs Boson with the ATLAS Detector at the LHC”. Physics Letters  
2017 B. 716 (1): 1-29 (2012). arXiv:1207.7214. doi:10.1016/j.physletb.2012.08.020.
- 2018 [34] ATLAS collaboration; CMS collaboration (26 March 2015). “Combined Mea-  
2019 surement of the Higgs Boson Mass in pp Collisions at  $\sqrt{s}=7$  and 8 TeV with  
2020 the ATLAS and CMS Experiments”. Physical Review Letters. 114 (19): 191803.  
2021 arXiv:1503.07589. doi:10.1103/PhysRevLett.114.191803.
- 2022 [35] LHC InternationalMasterclasses“When protons collide”. Retrieved from  
2023 [http://atlas.physicsmasterclasses.org/en/zpath\\_protoncollisions.htm](http://atlas.physicsmasterclasses.org/en/zpath_protoncollisions.htm)
- 2024 [36] CMS Collaboration, “SM Higgs Branching Ratios and To-  
2025 tal Decay Widths (up-date in CERN Report4 2016)”).

- 2026        <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageBR>  
 2027        , last accessed on 17.12.2017.
- 2028 [37] R.Grant V. “Determination of Higgs branching ratios in  $H \rightarrow W^+W^- \rightarrow l\nu jj$  and  $H \rightarrow ZZ \rightarrow l^+l^-jj$  channels”. Physics Department, University of Tennessee (Dated: October 31, 2012). Retrieved from  
 2029        <http://aesop.phys.utk.edu/ph611/2012/projects/Riley.pdf>
- 2030         
 2031
- 2032 [38] LHC Higgs Cross Section Working Group, Denner, A., Heinemeyer, S. et al.  
 2033        “Standard model Higgs-boson branching ratios with uncertainties”. Eur. Phys.  
 2034        J. C (2011) 71: 1753. <https://doi.org/10.1140/epjc/s10052-011-1753-8>
- 2035 [39] F. Maltoni, K. Paul, T. Stelzer, and S. Willenbrock, “Associated production  
 2036        of Higgs and single top at hadron colliders”, Phys.Rev. D64 (2001) 094023,  
 2037        [hep-ph/0106293].
- 2038 [40] S. Biswas, E. Gabrielli, F. Margaroli, and B. Mele, “Direct constraints on the  
 2039        top-Higgs coupling from the 8 TeV LHC data,” Journal of High Energy Physics,  
 2040        vol. 07, p. 073, (2013).
- 2041 [41] M. Farina, C. Grojean, F. Maltoni, E. Salvioni, and A. Thamm, “Lifting de-  
 2042        generacies in Higgs couplings using single top production in association with a  
 2043        Higgs boson,” Journal of High Energy Physics, vol. 05, p. 022, (2013).
- 2044 [42] T.M. Tait and C.-P. Yuan, “Single top quark production as a window to physics  
 2045        beyond the standard model”, Phys. Rev. D 63 (2000) 014018 [hep-ph/0007298].
- 2046 [43] F. Demartin, F. Maltoni, K. Mawatari, and M. Zaro, “Higgs production in  
 2047        association with a single top quark at the LHC,” European Physical Journal C,  
 2048        vol. 75, p. 267, (2015).

- 2049 [44] CMS Collaboration, “Modelling of the single top-quark pro-  
 2050 duction in association with the Higgs boson at 13 TeV.”  
 2051 <https://twiki.cern.ch/twiki/bin/viewauth/CMS/SingleTopHiggsGeneration13TeV>,  
 2052 last accessed on 16.01.2018.
- 2053 [45] CMS Collaboration, “SM Higgs production cross sections at  $\sqrt{s} = 13$  TeV.”  
 2054 <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/CERNYellowReportPageAt13TeV>, last  
 2055 accessed on 16.01.2018.
- 2056 [46] S. Dawson, The effective W approximation, Nucl. Phys. B 249 (1985) 42.
- 2057 [47] S. Biswas, E. Gabrielli and B. Mele, JHEP 1301 (2013) 088 [arXiv:1211.0499  
 2058 [hep-ph]].
- 2059 [48] F. Demartin, B. Maier, F. Maltoni, K. Mawatari, and M. Zaro, “tWH associated  
 2060 production at the LHC”, European Physical Journal C, vol. 77, p. 34, (2017).  
 2061 arXiv:1607.05862
- 2062 [49] LHC Higgs Cross Section Working Group, “Handbook of LHC Higgs Cross  
 2063 Sections: 4.Deciphering the Nature of the Higgs Sector”, arXiv:1610.07922.
- 2064 [50] J. Ellis, D. S. Hwang, K. Sakurai, and M. Takeuchi.“Disentangling Higgs-Top  
 2065 Couplings in Associated Production”, JHEP 1404 (2014) 004, [arXiv:1312.5736].
- 2066 [51] CMS Collaboration, V. Khachatryan et al., “Precise determination of the mass  
 2067 of the Higgs boson and tests of compatibility of its couplings with the standard  
 2068 model predictions using proton collisions at 7 and 8 TeV,” arXiv:1412.8662.
- 2069 [52] ATLAS and CMS Collaborations, “Measurements of the Higgs boson produc-  
 2070 tion and decay rates and constraints on its couplings from a combined ATLAS

- 2071 and CMS analysis of the LHC pp collision data at  $sqrts = 7$  and 8 TeV,” (2016).
- 2072 CERN-EP-2016-100, ATLAS-HIGG-2015-07, CMS-HIG-15-002.
- 2073 [53] File:Cern-accelerator-complex.svg. Wikimedia Commons,
- 2074 the free media repository. Retrieved January, 2018 from  
 2075 <https://commons.wikimedia.org/wiki/File:Cern-accelerator-complex.svg>
- 2076 [54] J.L. Caron , “Layout of the LEP tunnel including future LHC infrastructures.”,  
 2077 (Nov, 1993). A C Collection. Legacy of AC. Pictures from 1992 to 2002. Re-  
 2078 trieval from <https://cds.cern.ch/record/841542>
- 2079 [55] M. Vretenar, “The radio-frequency quadrupole”. CERN Yellow Report CERN-  
 2080 2013-001, pp.207-223 DOI:10.5170/CERN-2013-001.207. arXiv:1303.6762
- 2081 [56] L.Evans. P. Bryant (editors). “LHC Machine”. JINST 3 S08001 (2008).
- 2082 [57] CERN Photographic Service.“Radio-frequency quadrupole, RFQ-1”, March  
 2083 1983, CERN-AC-8303511. Retrieved from <https://cds.cern.ch/record/615852>.
- 2084 [58] CERN Photographic Service “Animation of CERN’s accelerator net-  
 2085 work”, 14 October 2013. DOI: 10.17181/cds.1610170 Retrieved from  
 2086 <https://videos.cern.ch/record/1610170>
- 2087 [59] C.Sutton. “Particle accelerator”.Encyclopedia Britannica. July 17, 2013. Re-  
 2088 trieval from <https://www.britannica.com/technology/particle-accelerator>.
- 2089 [60] L.Guiraud. “Installation of LHC cavity in vacuum tank.”. July 27 2000. CERN-  
 2090 AC-0007016. Retrieved from <https://cds.cern.ch/record/41567>.
- 2091 [61] J.L. Caron, “Magnetic field induced by the LHC dipole’s superconducting coils”.  
 2092 March 1998. AC Collection. Legacy of AC. Pictures from 1992 to 2002. LHC-  
 2093 PHO-1998-325. Retrieved from <https://cds.cern.ch/record/841511>

- 2094 [62] AC Team. “Diagram of an LHC dipole magnet”. June 1999. CERN-DI-9906025  
2095 retrieved from <https://cds.cern.ch/record/40524>.
- 2096 [63] CMS Collaboration “Public CMS Luminosity Information”.  
2097 <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#2016>  
2098 \_proton\_proton\_13\_TeV\_collis, last accessed 24.01.2018
- 2099 [64] J.L Caron. “LHC Layout” AC Collection. Legacy of AC. Pictures  
2100 from 1992 to 2002. September 1997, LHC-PHO-1997-060. Retrieved from  
2101 <https://cds.cern.ch/record/841573>.
- 2102 [65] J.A. Coarasa. “The CMS Online Cluster: Setup, Operation and Maintenance  
2103 of an Evolving Cluster”. ISGC 2012, 26 February - 2 March 2012, Academia  
2104 Sinica, Taipei, Taiwan.
- 2105 [66] CMS Collaboration. “The CMS experiment at the CERN LHC” JINST 3 S08004  
2106 (2008).
- 2107 [67] CMS Collaboration. “CMS detector drawings 2012” CMS-PHO-GEN-2012-002.  
2108 Retrieved from <http://cds.cern.ch/record/1433717>.
- 2109 [68] R. Breedon. “View through the CMS detector during the cooldown of the  
2110 solenoid on February 2006. CMS Collection”, February 2006, CMS-PHO-  
2111 OREACH-2005-004, Retrieved from <https://cds.cern.ch/record/930094>.
- 2112 [69] A. Dominguez et. al. “CMS Technical Design Report for the Pixel Detector  
2113 Upgrade”, CERN-LHCC-2012-016. CMS-TDR-11.
- 2114 [70] CMS Collaboration. “Description and performance of track and primary-vertex  
2115 reconstruction with the CMS tracker,” Journal of Instrumentation, vol. 9, no.  
2116 10, p. P10009,(2014).

- 2117 [71] CMS Collaboration and M. Brice. “Images of the CMS Tracker Inner  
2118 Barrel”, November 2008, CMS-PHO-TRACKER-2008-002. Retrieved from  
2119 <https://cds.cern.ch/record/1431467>.
- 2120 [72] M. Weber. “The CMS tracker”. 6th international conference on hyperons, charm  
2121 and beauty hadrons Chicago, June 28-July 3 2004.
- 2122 [73] CMS Collaboration. “Projected Performance of an Upgraded CMS Detector at  
2123 the LHC and HL-LHC: Contribution to the Snowmass Process”. Jul 26, 2013.  
2124 arXiv:1307.7135
- 2125 [74] L. Veillet. “End assembly of HB with EB rails and rotation in-  
2126 side SX ”,January 2002. CMS-PHO-HCAL-2002-002. Retrieved from  
2127 <https://cds.cern.ch/record/42594>.
- 2128 [75] J. Puerta-Pelayo.“First DT+RPC chambers installation round in the  
2129 UX5 cavern.”. January 2007, CMS-PHO-OREACH-2007-001. Retrieved from  
2130 <https://cds.cern.ch/record/1019185>
- 2131 [76] X. Cid Vidal and R. Cid Manzano. “CMS Global Muon Trigger” web  
2132 site: Taking a closer look at LHC. Retrieved from [https://www.lhc-closer.es/taking\\_a\\_closer\\_look\\_at\\_lhc/0.lhc\\_trigger](https://www.lhc-closer.es/taking_a_closer_look_at_lhc/0.lhc_trigger)
- 2134 [77] WLCG Project Office, “Documents & Reference - Tiers - Structure,” (2014).  
2135 <http://wlcg.web.cern.ch/documents-reference> , last accessed on 30.01.2018.
- 2136 [78] CMS Collaboration. “CMSSW Application Framework”,  
2137 <https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookCMSSWFFramework>,  
2138 last accesses 06.02.2018

- 2139 [79] A. Buckleya, J. Butterworthb, S. Giesekec, et. al. “General-purpose event gen-  
 2140 erators for LHC physics”. arXiv:1101.2599v1 [hep-ph] 13 Jan 2011
- 2141 [80] A. Quadt. “Top Quark Physics at Hadron Colliders”. Advances in the Physics  
 2142 of Particles and Nuclei. Springer-Verlag Berlin Heidelberg. DOI: 10.1007/978-  
 2143 3-540-71060-8 (2007)
- 2144 [81] DurhamHep Data Project, “The Durham HepData Project - PDF Plotter.”  
 2145 <http://hepdata.cedar.ac.uk/pdf/pdf3.html> , last accessed on 02.02.2018.
- 2146 [82] F. Maltoni, G. Ridolfi, and M. Ubiali, “b-initiated processes at the LHC: a  
 2147 reappraisal,” Journal of High Energy Physics, vol. 07, p. 022, (2012).
- 2148 [83] B. Andersson, G. Gustafson, G. Ingelman and T. Sjostrand, “Parton fragmen-  
 2149 tation and string dynamics”, Physics Reports, Vol. 97, No. 2-3, pp. 31-145,  
 2150 1983.
- 2151 [84] CMS Collaboration, “Event generator tunes obtained from underlying event  
 2152 and multiparton scattering measurements;” European Physical Journal C, vol.  
 2153 76, no. 3, p. 155, (2016).
- 2154 [85] J. Alwall et. al., “The automated computation of tree-level and next-to-leading  
 2155 order differential cross sections, and their matching to parton shower simula-  
 2156 tions,” Journal of High Energy Physics, vol. 07, p. 079, (2014).
- 2157 [86] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with  
 2158 Parton Shower simulations: the POWHEG method,” Journal of High Energy  
 2159 Physics, vol. 11, p. 070, (2007).
- 2160 [87] S. Agostinelli et al., “GEANT4: A Simulation toolkit,” Nuclear Instruments  
 2161 and Methods in Physics, vol. A506, pp. 250–303, (2003).

- 2162 [88] J.Allison et.al.,“Recent developments in Geant4”, Nuclear Instruments and  
2163 Methods in Physics Research A 835 (2016) 186-225.
- 2164 [89] CMS Collaboration “Full Simulation Offline Guide”,  
2165 <https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideSimulation>, last  
2166 accessed 04.02.2018
- 2167 [90] A. Giammanco. “The Fast Simulation of the CMS Experiment” J. Phys.: Conf.  
2168 Ser. 513 022012 (2014)
- 2169 [91] A.M. Sirunyan et. al. “Particle-flow reconstruction and global event description  
2170 with the CMS detector”, JINST 12 P10003 (2017) <https://doi.org/10.1088/1748-0221/12/10/P10003>.
- 2172 [92] The CMS Collaboration. “ Description and performance of track and pri-  
2173 mary vertex reconstruction with the CMS tracker”. JINST 9 P10009 (2014).  
2174 doi:10.1088/1748-0221/9/10/P10009
- 2175 [93] P. Billoir and S. Qian, “Simultaneous pattern recognition and track fitting by  
2176 the Kalman filtering method”, Nucl. Instrum. Meth. A 294 219. (1990).
- 2177 [94] W. Adam, R. Fruhwirth, A. Strandlie and T. Todorov, “Reconstruction of  
2178 electrons with the Gaussian sum filter in the CMS tracker at LHC”, eConf  
2179 C 0303241 (2003) TULT009 [physics/0306087].
- 2180 [95] K. Rose, “Deterministic Annealing for Clustering, Compression, Classification,  
2181 Regression and related Optimisation Problems”, Proc. IEEE 86 (1998) 2210.
- 2182 [96] R. Fruhwirth, W. Waltenberger and P. Vanlaer, “ Adaptive Vertex Fitting”,  
2183 CMS Note 2007-008 (2007).

- 2184 [97] CMS collaboration, “Performance of CMS muon reconstruction in pp collision  
 2185 events at  $\sqrt{s} = 7$  TeV”, JINST 7 P10002 2012, [arXiv:1206.4071].
- 2186 [98] CMS Collaboration, “Search for the associated production of a Higgs boson  
 2187 with a single top quark in proton-proton collisions at  $\sqrt{s} = 8$  TeV”, JHEP 06  
 2188 (2016) 177, doi:10.1007/JHEP06(2016)177, arXiv:1509.08159.
- 2189 [99] B. Stieger, C. Jorda Lope et al., “Search for Associated Production of a Single  
 2190 Top Quark and a Higgs Boson in Leptonic Channels”, CMS Analysis Note CMS  
 2191 AN-14-140, 2014.
- 2192 [100] M. Peruzzi, C. Mueller, B. Stieger et al., “Search for ttH in multilepton final  
 2193 states at  $\sqrt{s} = 13$  TeV”, CMS Analysis Note CMS AN-16-211, 2016.
- 2194 [101] CMS Collaboration, “Search for H to bbar in association with a single top quark  
 2195 as a test of Higgs boson couplings at  $\sqrt{s} = 13$  TeV”, CMS Physics Analysis  
 2196 Summary CMS-PAS-HIG-16-019, 2016.
- 2197 [102] B. Maier, “SingleTopHiggProduction13TeV”, February, 2016.  
 2198 <https://twiki.cern.ch/twiki/bin/viewauth/CMS/SingleTopHiggsGeneration13TeV>.
- 2199 [103] M. Peruzzi, F. Romeo, B. Stieger et al., “Search for ttH in multilepton final1  
 2200 states at  $\sqrt{s} = 13$  TeV”, CMS Analysis Note CMS AN-17-029, 2017.
- 2201 [104] B. WG, “BtagRecommendation80XReReco”, February, 2017.  
 2202 <https://twiki.cern.ch/twiki/bin/view/CMS/BtagRecommendation80XReReco>.