

<sup>1</sup>

José D. Ruiz-Álvarez

<sup>2</sup>

Compiled on 26/05/2015 at 14:00



<sup>5</sup> **Outline**



# **1 The Standard Model**

Since the Greeks, different theories about the composition and structure of the world have been formulated. At ancient Greece these theories were elaborated from a philosophical point of view. Nowadays, we count with a very sophisticated set of tools and concepts that allowed us to build up a general vision of nature, its components and structure. Moreover, on the subject of the constituents, or elemental constituents, a theory capable of describing the majority of known phenomena has been developed. This theory is the Standard Model (SM) of particle physics.

This SM relies in two of the more elegant constructs of modern physics and mathematics. From the physics side, the quantum field theory; from mathematics, group theory. Quantum field theory has born from the understanding of processes that take place at very small spatial scales and in a regime where special relativity play an important role. To describe such, a major part of the most brilliant minds of the 20th century dedicated their life, Paul Dirac, Richard Feynman, Enrico Fermi among them. The theory of quantum fields has set in a common place two extraordinary achievements of physics: special relativity and quantum mechanics. With it we have been capable to describe many phenomena:  $\beta$  and  $\alpha$  decay, solid state, among many other.

From the mathematics side, group theory has become one of the most powerful tools for particle physicist. However, their development began quite early, with Galois around 1830, and was used in other parts of physics, it's with Lie algebras and the possibility of describing continuous symmetries that the most important step were given. Also, this would have not been possible with the amazing connection found by Emmy Noether in 1918. She found that for every conserved quantity there is a preserved symmetry. Group theory can be seen, roughly speaking, a way to mathematically describe symmetries, group theory became the tool to describe systems with conserved quantities.

In this chapter, we present the basics of the SM. We describe its seminal ideas, its structure and content and its ultimate consequences. Finally, we close with its limitations.

## <sup>37</sup> 1.1 Fields, symmetries and interactions

<sup>38</sup> From the very beginning of physics, one of the most fundamental questions has  
<sup>39</sup> been how does bodies interact, and what is exactly an interaction. On the first  
<sup>40</sup> type of interaction ever studied by physics, gravity, Newton proposed the con-  
<sup>41</sup> cept of distant interaction, the idea that bodies could interact without being in  
<sup>42</sup> direct contact. But the question on how exactly that distant action was performed  
<sup>43</sup> remained unanswered.

<sup>44</sup> During the 19th and 20th century new phenomena were discovered pointing to  
<sup>45</sup> brand new interactions, electricity, magnetism and radioactivity. The very precise  
<sup>46</sup> and complete description of electromagnetism developed by Gauss, Faraday, Am-  
<sup>47</sup> père and finished by Maxwell arrived to describe electricity and magnetism under  
<sup>48</sup> the formalism of only one interaction within the mathematical formalism of classi-  
<sup>49</sup> cal fields. Further works addressed radioactivity, driving to a deeper understanding  
<sup>50</sup> of nature and its composition.

<sup>51</sup> For the following discussion, and later, we are going to work in natural units  
<sup>52</sup> for simplicity. In these units the speed of light  $c$  is normalized to unity, as well as  
<sup>53</sup> electron electric charge  $e$ , reduced Planck constant  $\hbar$  and Boltzmann constant  $k_B$ .  
<sup>54</sup> Then, masses and temperature are expressed in energy units, i.e.  $eV$ , and time  
<sup>55</sup> and length in inverse energy units,  $eV^{-1}$ .

<sup>56</sup> A classical field is an assignment of a quantity to every point in space and time.  
<sup>57</sup> For physics, the quantity that is attributed is a physical quantity such as mass,  
<sup>58</sup> electrical charge or probability. This quantity can be scalar or vector, giving rise to  
<sup>59</sup> the notion of scalar or vector field, correspondingly. The simplest example, is the  
<sup>60</sup> temperature in a gas, that is a scalar quantity assigned to every point. Another  
<sup>61</sup> example, a fluid can be described in terms of fields, being the velocity of the fluid  
<sup>62</sup> a vector field and its pressure a scalar field. Generic classical electromagnetic  
<sup>63</sup> interactions can be described with the help of one vector field  $\vec{A}(x)$ , the vector  
<sup>64</sup> potential, and one scalar field  $\phi(x)$ , the scalar potential. In the formalism of  
<sup>65</sup> four-vectors from relativistic dynamics one can organize these two quantities in the  
<sup>66</sup> four-potential  $A_\mu = (-\phi, \vec{A})$ . This can be used to define the strength field tensor  
<sup>67</sup>  $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ , where  $\partial_\mu = \left(-\frac{\partial}{\partial t}, \nabla\right)$  is the covariant derivative. From the  
<sup>68</sup> tensor it is possible to obtain in a very generic and elegant way the equations of  
<sup>69</sup> motion of the free field using the Lagrangian formalism, as in equation 1.1. With  
<sup>70</sup> the Lagrangian density defined in equation 1.2.

$$\partial_\mu \left( \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \right) - \frac{\partial \mathcal{L}}{\partial A_\nu} = 0 \quad (1.1)$$

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} \quad (1.2)$$

71 It's very important to notice that the equations of motion of the free field are  
 72 invariant under the choice of the four-potential. More precisely, the covariant  
 73 potential is not unique and we can always add the covariant derivative of a scalar  
 74 field,

$$A'_\mu = A_\mu + \partial_\mu \Lambda(x) \leftrightarrow \partial^\mu A_\mu = 0 \quad (1.3)$$

75 and describe the same physics. This non-uniqueness corresponds to the choice  
 76 of a zero-point of the potential very well known in non-Lagrangian formalism of  
 77 electrodynamics. When we choose a specific value for this scalar field,  $\Lambda(x)$ , we  
 78 say that the gauge has been fixed.

79 One can also define a four current vector,  $J_\mu = (\rho, \vec{J})$  with  $\rho$  the electric charge  
 80 density and  $\vec{J}$  the current charge density. Then, plugging in this four current in  
 81 the Lagrangian of the free field, defined in equation 1.2,

$$\mathcal{L} = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu} - A_\mu J^\mu \quad (1.4)$$

82 we can obtain the complete set of equations of motion of the field with charges  
 83 and currents.

84 The transformation stated from equation 1.3 can be understood as a transfor-  
 85 mation of the field. These type of transformations are mathematically understood  
 86 under the group  $U(1)$ , where the generic transformation operator can be written  
 87 as  $U = e^{i\theta(x)}$ . It's said then that the electromagnetic vector potential is *invariant*  
 88 under  $U(1)$  transformations. This property identifies an essential characteristic of  
 89 electromagnetism, its symmetric behavior under  $U(1)$ .

90 From this reasoning the most interesting results are drawn when the same sym-  
 91 metry is imposed to another fields. For example, the kinetic Lagrangian for a  
 92 complex scalar field is  $\mathcal{L} = (\partial^\mu \phi)^* \partial_\mu \phi$ . To perform the transformation on the  
 93 scalar field, it is sufficient to apply the operator as  $\phi' = U\phi$  and  $\phi'^* = \phi^* U^{-1}$ . But  
 94 it's evident that the Lagrangian is not the same after applying such transforma-  
 95 tion. Then, in order to preserve the Lagrangian under  $U(1)$  is necessary to change  
 96 at the same time the derivative. Such transformation is given in equation 1.5,  
 97 where  $g$  is a constant.

$$\mathcal{D}^\mu = \partial^\mu - ig A^\mu \quad (1.5)$$

98 Then, the proposed Lagrangian can be rewritten, including the vector field, as

$$\mathcal{L} = (\mathcal{D}^\mu \phi)^* \mathcal{D}_\mu \phi - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} \quad (1.6)$$

99 that is invariant under  $U(1)$ . An interaction term, of the form  $ig A^\mu \phi^* \partial_\mu$ , between  
 100 the scalar and the vector field, can be derived from the kinematic part of the  
 101 Lagrangian  $\mathcal{D}^\mu \phi)^* \mathcal{D}_\mu \phi$ . This shows that the requirement of the invariance under

102 103 104 105 106 107  $U(1)$  of the scalar field lead to the introduction of an interaction with a vector field controlled by the constant  $g$ . We have also seen that electromagnetic interaction is described precisely by a vector field and that preserves  $U(1)$  symmetry, which implies that this symmetry is the connection to electromagnetic interaction, identifying the interaction itself with the  $U(1)$  symmetry. In addition, using Noether theorem one can show that  $g$  is a conserved quantity, as the electric charge is.

108 109 110 111 112 113 114 115 116 117 But not only electromagnetism can be described via a continuous symmetry as  $U(1)$ . On 1896 radioactivity was discovered by the french physicist Henri Becquerel. Three years after, Marie and Pierre Curie studied in more detail the phenomenon and found Polonium and Radium elements. And later on, Ernst Rutherford was able to describe radioactivity as coming in three types, alpha ( $\alpha$ ), beta ( $\beta$ ) and gamma ( $\gamma$ ). He also noticed that radioactivity was able to change matter, which allowed him, with also other experiences, to propose an atomic model, describing elements as basically an external core of negative charges and a nucleus positively charged. Consequently, This findings implied the existence of interactions different to electromagnetism, acting at the atomic scale.

118 119 120 121 122 123 The interaction that undergoes radioactivity, beta decay, is called the weak interaction. In 1933 Enrico Fermi made a first theoretical description of this interaction, but only in 1968 Sheldon Glashow, Abdus Salam and Steven Weinberg were able to describe weak interaction with a symmetry group:  $SU(2)$ . Finally, the interaction that keeps the nucleus components together, the strong interaction, was described with  $SU(3)$  group mainly by Murray Gell-Mann in 1963.

124 125 126 There have been many attempts to describe gravity with the same formalism, but up to present such attempts have been unsuccessful. Such question remains one of the most important problems for modern particle physics.

## 127 1.2 Quantum fields and particles

128 129 130 131 132 Classical fields, introduced and described in last section 1.1, can be extended to a quantum theory. Such procedure is known as the quantization of fields and allow to unify special relativity and quantum mechanics in one theory, Quantum Field Theory (QFT), to describe the dynamics of systems in such regimes: speed close to the speed of light on the atomic or smaller scales.

133 134 135 136 137 138 139 Quantum mechanics introduced two fundamental concepts: first, the description of the system by its states; and second, the identification of an observable with an operator. The state of a system is identified with a set of quantum numbers that tell us the characteristics of the system when it is at some state. For example, the hydrogen atom system has energy as quantum number, such that each state has a value for the energy describing the potential energy contained in the system. Quantum states are mathematically noted in Dirac notation as a *ket*,

$$|\alpha\rangle = |i, j, k, \dots\rangle \quad (1.7)$$

with  $\alpha$  the set of quantum numbers  $i, j, k, \dots$ . This mathematical object lives in Hilbert space (a complex space  $\mathbb{C}$  of functions), which conjugate, a *bra*, is noted  $\langle\alpha|$ , and their internal product  $\langle\beta|\alpha\rangle$ . The numerical value of  $|\langle\beta|\alpha\rangle|^2$  gives the transition probability of the system from state  $\beta$  to state  $\alpha$ , and  $|\langle\alpha|\alpha\rangle|^2$  is the probability to find the system in state the  $\alpha$ .

Physical observables as position, energy or momentum are described by complex operators such that to measure their value for a given state, one just have to calculate  $|\langle\alpha|\hat{O}|\alpha\rangle|^2$ . The identification of observables and operators is called *first quantization*. In addition, Schrodinger equation describes the evolution of states,

$$\hat{\mathcal{H}}|\alpha\rangle = i\frac{d}{dt}|\alpha\rangle \quad (1.8)$$

with  $\mathcal{H}$  the Hamiltonian of the system. The whole formalism is able to explain *quantized* systems, where the quantum numbers are discrete, such as hydrogen atom or black body.

Several functions or fields can be related to a given state. These functions, wave functions, can be used as the states to calculate probabilities. In *second quantization* wave functions are upgraded into field operators. This procedure gives rise to the quantization of the state of the field, which is described by the quantum number  $n$  which is definite positive.  $n = 0$  for the fundamental state and  $n > 0$  for the excited states. Such excitations of the field are understood as physical particles that propagates in space-time, which means that  $n = 0$  is vacuum.

The first QFT ever created was born from the quantization of the electromagnetic field. Quantum Electro Dynamics (QED) is the quantized version of classical electrodynamics, that was developed by Tomonaga, Schwinger and Feynman around 1960. This theory describes electromagnetic interactions of a charged field and the electromagnetic vector field. The charged field excitations correspond to electrons and the excitations of the vector fields are photons, responsible of light. Electrons are a particle with negative electric charge and orbit around the nucleus in atoms. Discovered in 1897 by J. J. Thomson, it was fully described by P. A. Dirac in 1928 with the Dirac equation that is the Schrodinger equation for a relativistic particle of spin  $1/2$ . Spin, the intrinsic angular momentum carried by a particle, can be integer  $(0, 1, 2, \dots)$  or semi-integer  $(\frac{1}{2}, \frac{3}{2}, \dots)$ . The particles with semi-integer spin, as electrons, are called *fermions* and particles with integer spin, as photons, are called *bosons*. Dirac equation predicted the existence of a particle identical to the electron but with positive charge, the positron. It was discovered on 1932 by Carl David Anderson.

175 Up to present days we have found 12 fundamental fermions and 5 fundamental  
 176 bosons. Fermions are organized in *leptons*, that don't interact strongly, and *quarks*,  
 177 that do interact strongly. Leptons are as well organized in three families, the  
 178 electron ( $e^-$ ) and electron neutrino ( $\nu_e$ ), muon ( $\mu^-$ ) and muon neutrino ( $\nu_\mu$ ) and  
 179 tau ( $\tau$ ) and tau neutrino ( $\nu_\tau$ ). Electron, muon and tau are electrically charged  
 180 while neutrinos are neutral. Their respective anti-particles are equally organized,  
 181 positron ( $e^+$ ) with electron anti-neutrino ( $\bar{\nu}_e$ ), anti-muon ( $\mu^+$ ) with muon anti-  
 182 neutrino ( $\bar{\nu}_\mu$ ) and anti-tau ( $\tau^+$ ) with tau anti-neutrino ( $\bar{\nu}_\tau$ ). Quarks also come in  
 183 three families, with the respective anti-quarks: up ( $u, \bar{u}$ ) and down ( $d, \bar{d}$ ), charm  
 184 ( $c, \bar{c}$ ) and strange ( $s, \bar{s}$ ), top ( $t, \bar{t}$ ) and bottom ( $b, \bar{b}$ ). The fundamental bosons  
 185 are the photon ( $A$ ), the W (positively and negatively charged) and Z that mediate  
 186 the electroweak interaction, the gluon ( $g$ ) mediating the strong interaction and  
 187 the Higgs ( $H$ ). The weak bosons were discovered at CERN in 1983 at the UA1  
 188 and UA2 collaborations over the SPS accelerator (described on section 2.1.1). The  
 189 Higgs boson has been discovered recently on 2012 by ATLAS and CMS experiments  
 190 at the LHC. The 2013 physics Nobel prize was awarded to Francois Englert and  
 191 Peter W. Higgs "for the theoretical discovery of a mechanism that contributes to  
 192 our understanding of the origin of mass of subatomic particles, and which recently  
 193 was confirmed through the discovery of the predicted fundamental particle, by the  
 194 ATLAS and CMS experiments at CERN's Large Hadron Collider".

### 195 **1.2.1 The mass problem**

196 Using the concepts developed on later sections about QFT and symmetries it's  
 197 possible to construct a whole theory giving rise to a precise description of particles  
 198 and interactions between them. But such a theory does not allow to have massive  
 199 bosons, whereas masses for fermions are allowed. A mass term for a fermion  $\psi$  is  
 200 of the form

$$m_\psi \bar{\psi} \psi$$

201 where  $m_\psi$  is the mass of the field.

202 Under  $U(1)$  transformations,  $\psi' = U\psi$ , the mass term remains the same, what  
 203 means that is invariant under  $U(1)$  transformations. The same is not true for a  
 204 boson. A mass term for a boson  $A$ , can be written as

$$m_A A^\mu A_\mu \quad (1.9)$$

205 where  $m_A$  is its mass.

206 The  $U(1)$  transformation for the boson is  $A'^\mu = A^\mu + \delta^\mu \theta(x)$ . Applying such  
 207 transformation on the mass term 1.10 one can obtain the transformed term

$$m_A (A^\mu A_\mu + A^\mu \delta_\mu \theta(x) + \delta^\mu \theta(x) A_\mu + \delta^\mu \theta(x) \delta_\mu \theta(x)) \quad (1.10)$$

where the last three terms make this term not invariant under  $U(1)$  transformations. Consequently, one can say that a mass term for the boson destroy the invariance of the theory under  $U(1)$  symmetry.

Nonetheless there is no need for a theory of a massive photon, there is a need to have massive bosons for weak interactions. There is a relation between the mass of a boson and the range of the interaction mediated by it. Massless bosons transmit long range interactions, as electromagnetism, but short range interactions, as the weak interaction, are mediated by massive bosons. More precisely the interaction range is inversely proportional to the mass of the boson, higher the mass shortest the range. Such relation can be seen from the structure of the propagator, which is a mathematical entity that describes the probability a particle has to travel a distance in a given time. Such propagator, for a vector boson, has a generic form given in equation 1.11, where  $k_\mu$  is the momentum carried by the boson. It's clear from this structure that a massive boson has less probability to travel a long distance than a massless boson in a given time.

$$\frac{g_{\mu\nu}}{k^2 - m^2 + i\epsilon} \quad (1.11)$$

Then, as massive bosons are requires for weak interaction, somehow the  $SU(2)$  symmetry has to be broken. There are basically two ways to broke a symmetry:

- Explicit symmetry breaking: By the introduction of a symmetry breaking term in the Lagrangian, as a mass term for the bosons.
- Spontaneous symmetry breaking: When the ground state of one field fail to be invariant under the symmetry.

Explicit symmetry breaking is not an option, because the symmetry needs to be preserved in the Lagrangian in order to introduce the interaction.

## 1.2.2 Spontaneous Symmetry Breaking

Several physical systems exhibit an spontaneous symmetry breaking. For example, a pencil balanced on its tip is perfectly symmetric system around the vertical axis, however, because of the instability of the system the pencil will eventually fall over. The final state is stable but not symmetrical. This transition also decreased the potential energy of the system, driving the system to its ground state. This means that whereas the system had a symmetry the ground state does not show the symmetry. In general, symmetry breaking is linked to phase transitions, as liquid to gas transition or magnetization of a ferromagnet, covering a plethora of

240 physical processes. To a greater extent, in recent studies, [8], the emergence of life  
 241 has been understood as a phase transition of matter.

242 To achieve an spontaneous symmetry breaking of  $SU(2)$  in QFT one should  
 243 choose a field for which its ground state, vacuum, will fail the symmetry. This  
 244 means, in practical terms, that such field will have a non-zero value in vacuum,  
 245 leading to a presence of particles coming from the field on theory vacua. If a  
 246 fermion field is chosen, the vacuum will show a preference on directionality de-  
 247 pending on its spin orientation, what breaks Poincaré symmetry imposed by spe-  
 248 cial relativity. The same is true if a spin-1 bosonic field is chosen for the task. In  
 249 order to avoid this problem a spin-0 field should be used. In addition, this field  
 250 should be electrically neutral to avoid having a charged vacuum.

251 Whit all this properties in mind, taking a scalar doublet of  $SU(2)$ , defined on  
 252 equation 1.12 where  $\phi^0$  and  $\phi^+$  are complex fields, the most general potential can  
 253 be written from two auto-interaction terms, in equation 1.13.

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \phi^+ \\ \phi_{RE} - i\phi_{IM} \end{pmatrix} \quad (1.12)$$

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \quad (1.13)$$

254 Such potential has a unique minimum for  $\lambda > 0$  and  $\mu^2 > 0$ , but for  $\lambda > 0$  and  
 255  $\mu^2 < 0$  has a set of minima with the shape of “Mexican hat”, shown in figure 1.1.  
 256 Under  $\lambda > 0$ ,  $\mu^2 < 0$  configuration the field breaks spontaneously the symmetry  
 257 reaching the ground state, acquiring an expectation value on vacuum different  
 258 from zero,  $v$ .

### 259 1.2.3 Englert-Brout-Higgs mechanism

260 After the spontaneous symmetry breaking, the scalar doublet transforms into the  
 261 form given in equation 1.14, where  $G^+$  and  $G^0$  are the Goldstone bosons product  
 262 of the breaking of the  $SU(2)$  symmetry, and  $H$  is the Englert-Brout-Higgs boson.  
 263 From Goldstone’s theorem when a symmetry is spontaneously broken a massless  
 264 boson appear for each broken generator. In our specific case, the three generators  
 265 of  $SU(2)$  are broken giving rise to three Goldstone bosons:  $G^+$ ,  $G^-$  and  $G^0$ . This  
 266 massless bosons are “eaten” by the  $W^+$ ,  $W^-$  and  $Z^0$  giving them an additional  
 267 degree of freedom, the longitudinal polarization.

$$\Phi = \begin{pmatrix} G^+ \\ \frac{1}{\sqrt{2}}(H + v - iG^0) \end{pmatrix} \quad (1.14)$$

268 By this mechanism, the  $W$  and  $Z$  bosons acquire mass, being its value set by the  
 269 coupling constant of  $SU(2)$  group and the vacuum expectation value of Englert-

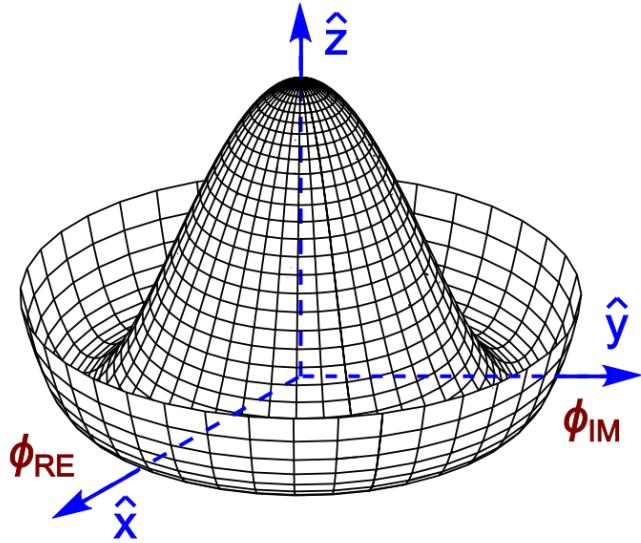


Figure 1.1: Higgs potential

270 Brout-Higgs boson. In addition, the fermions on the theory also acquire a mass  
271 from the interactions with the scalar doublet. Such masses are in general of the  
272 form  $m_f = \lambda_f v / \sqrt{2}$ , where  $\lambda_f$  sets the interaction between the Englert-Brout-  
273 Higgs boson and the fermion. Finally, also the Englert-Brout-Higgs boson has a  
274 mass  $m_H^2 = -2\mu^2$ .

275 In summary, with this mechanism the weak interaction bosons and fermions of  
276 the theory are given a mass on the price of introducing an additional scalar field  
277 to spontaneously break the  $SU(2)$  symmetry.

### 278 1.3 Top production at LHC

279 Discovered in 1995 by DØ and CDF collaborations at Tevatron, it's the heaviest  
280 fundamental particle known. As heaviest particle, many models beyond the SM  
281 predict a coupling of the top quarks with a heavier new physics sector. It forms  
282 a  $SU(2)_L$  weak isospin doublet with the b-quark, discovered in 1977. Their mass  
283 difference, two orders of magnitude, is one fundamental question in the SM. Pre-  
284 cision measurements of the top quark are fundamental input to test the SM and  
285 possibly find new physics.

286 The LHC can be seen as a top-factory, being the accelerator where the most of  
287 top-quarks can be produced. During run 1, taking into account 7 TeV and 8 TeV  
288 data, 5.6 millions of top pairs events and 2.7 millions of single top events were

289 delivered by the LHC to CMS and ATLAS experiments. Taking into account the  
290 different cross sections of the production processes, that will be discussed in 1.3.1  
291 and 1.3.2, and the instantaneous luminosity, cited in table 2.1.2, for 8 TeV center  
292 of mass energy, at LHC there are produced around 6 tops per second where 5 of  
293 them come from top-pair events and one from single-top events.

### 294 **1.3.1 Pair production**

295 Description of pair production and the relative importance of each channel.

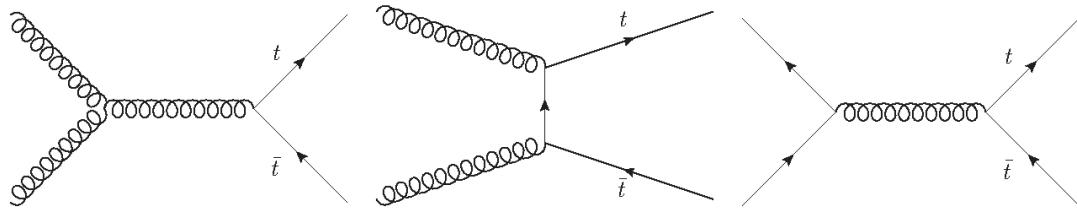


Figure 1.2: Top pair production processes Feynman diagrams for proton-proton collisions, via gluon fusion and quark-antiquark annihilation

### 296 **1.3.2 Single $t$ production**

297 Description of single production and the relative importance of each channel.

### 298 **1.3.3 Decay channels**

299 Description of possible decays of top-quark to show different channels of searches.

### 300 **1.3.4 Top properties**

301 Description of top properties and the measurement of them.

#### 302 **Electric charge**

303 Top charge related to its decay.

#### 304 **Lifetime**

305 Discussion on the importance of measurements of top as only quark decaying before  
306 hadronization time.

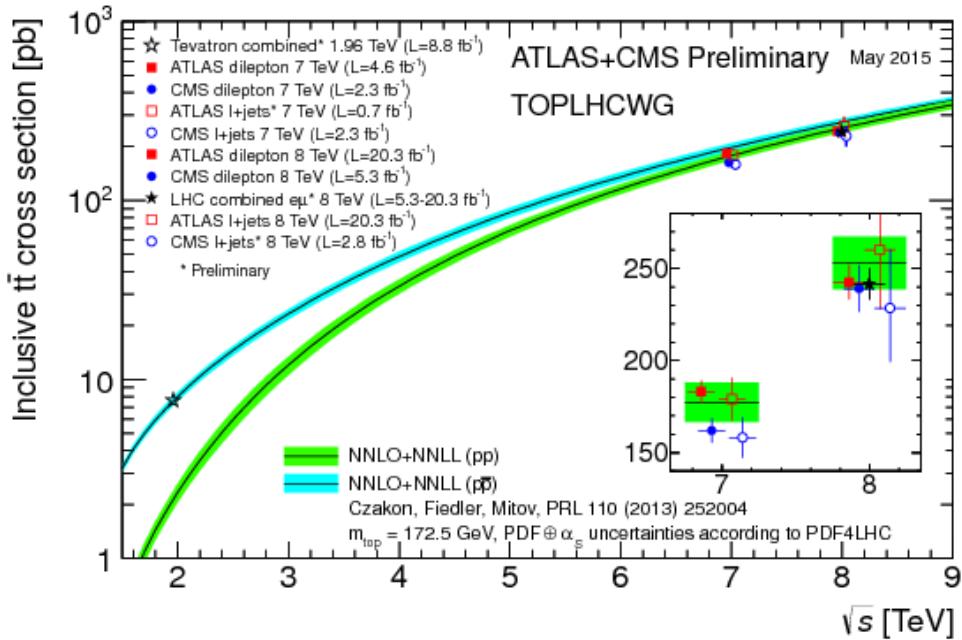


Figure 1.3:  $t\bar{t}$  production cross section as function of the center of mass energy in  $p\bar{p}$  and  $pp$  collisions compared to theoretical predictions.

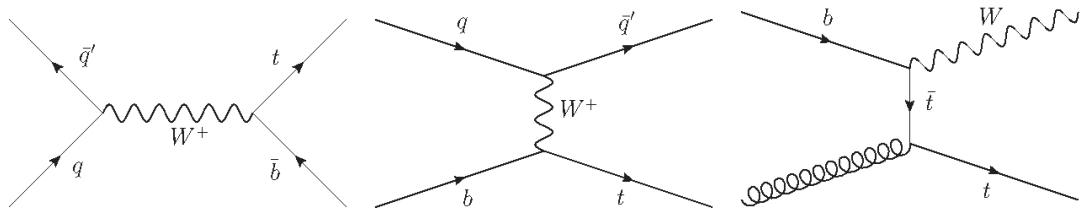


Figure 1.4: Single top production processes Feynman diagrams for proton-proton collisions, from left to right s-channel, t-channel and associated  $W$  production.

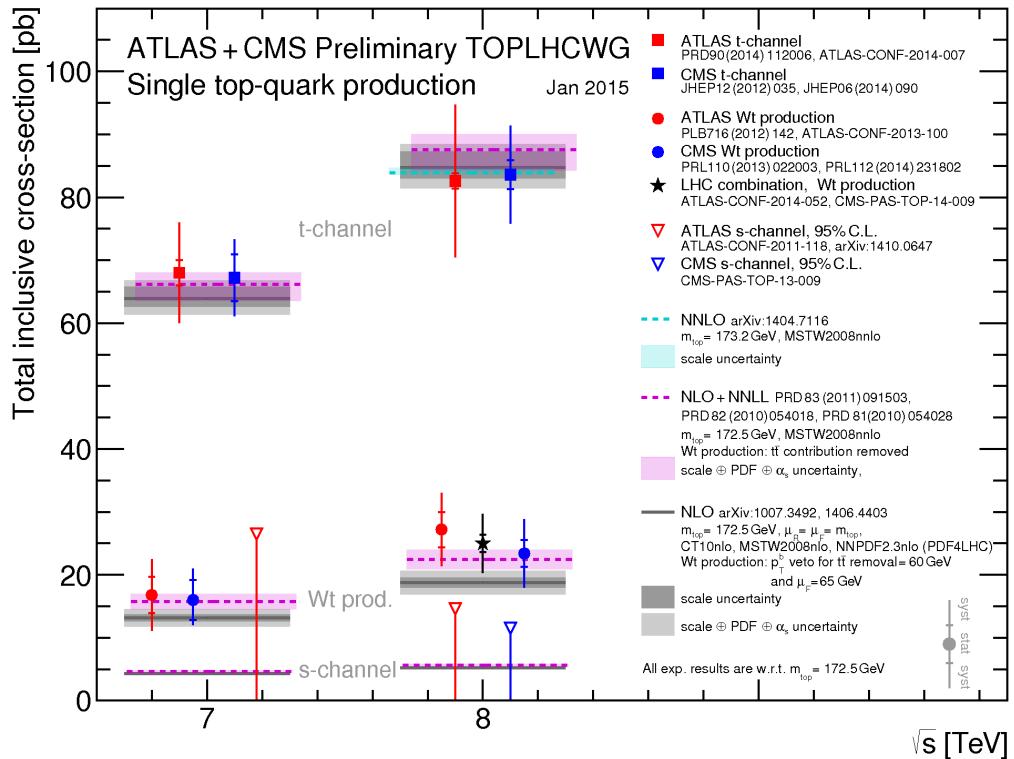


Figure 1.5: Single top production cross section as function of the center of mass energy in  $pp$  collisions compared to theoretical predictions for each production channel by ATLAS and CMS collaborations.

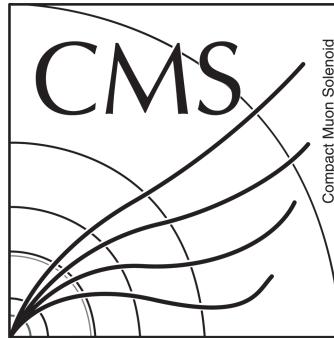


Figure 1.6: Feynman diagrams for top decay channels with respective branching ratios.

307 **Mass and width**

308 Measurements of top mass and width.

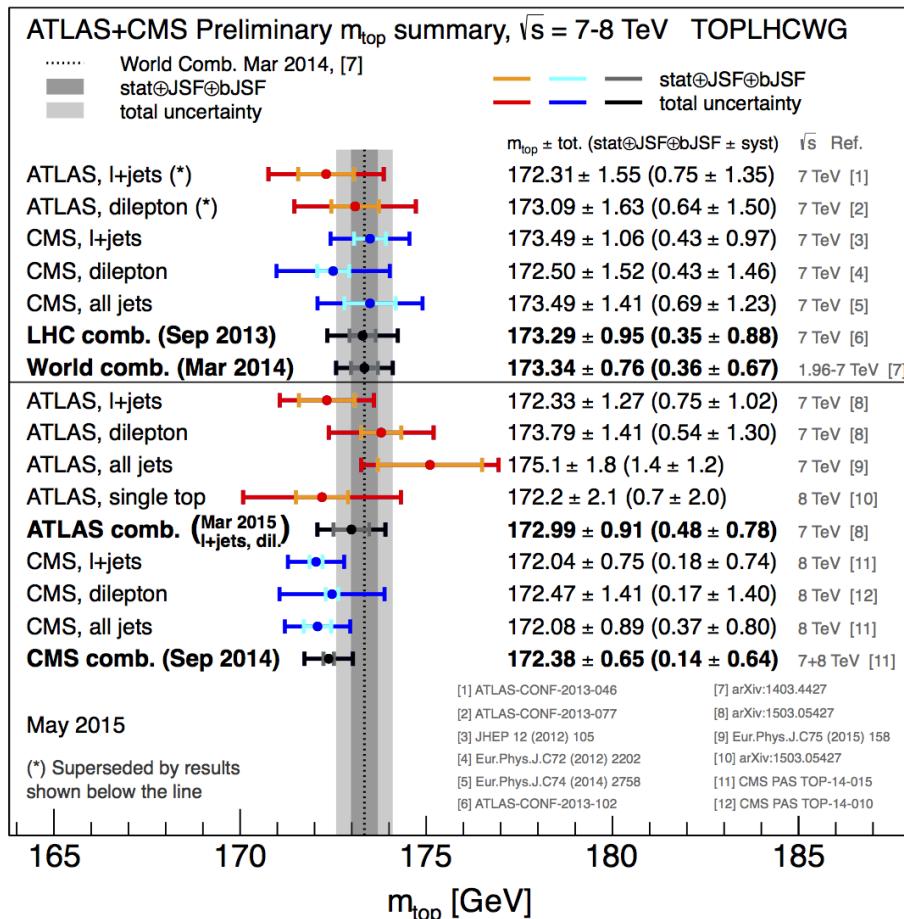


Figure 1.7: Top mass measurements from ATLAS and CMS collaborations and world combination including Tevatron results.

309 **Spin correlation**

310 Discussion on how  $t\bar{t}$  system have spin correlation that will be important for  
311 precision measurements.

## 312 1.4 Higgs production at LHC

313 Discussion of Higgs production channels and relative importance.

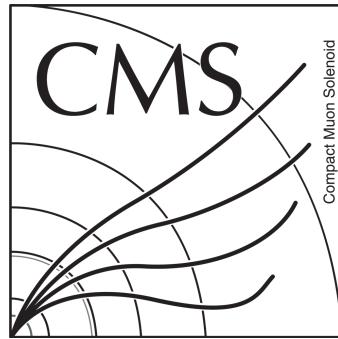


Figure 1.8: Higgs production Feynman diagrams for proton-proton collisions.

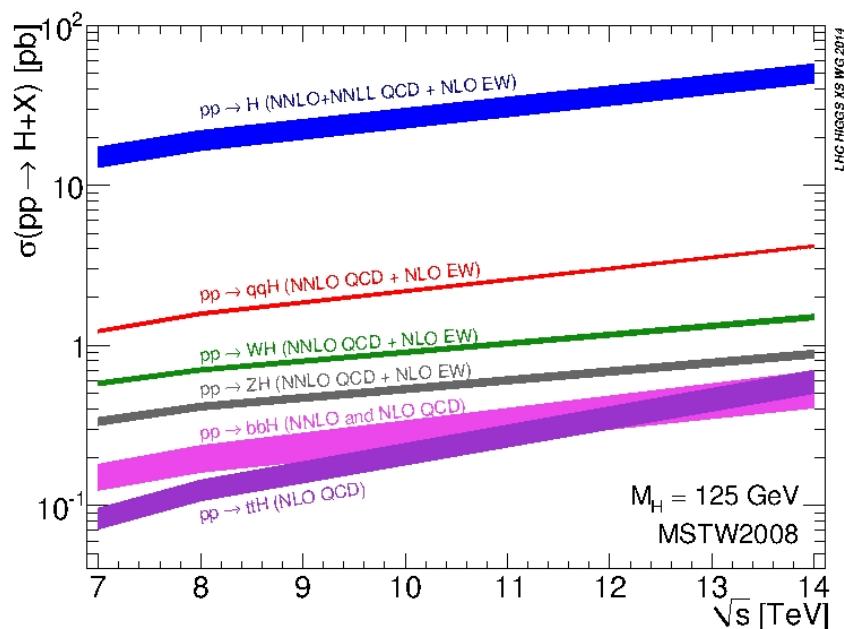


Figure 1.9: Higgs production cross section theoretical predictions as function of center of mass energy.

### <sup>314</sup> 1.4.1 Decay channels

- <sup>315</sup> Discussion of Higgs decays and relative importance also related to the resolution  
<sup>316</sup> of channels.

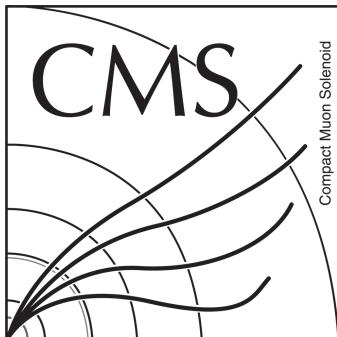


Figure 1.10: Feynman diagrams of Higgs decay:  $b\bar{b}$  [left], diphoton [center] and golden channels [right]

### <sup>317</sup> 1.4.2 Higgs properties

#### <sup>318</sup> Electric charge

- <sup>319</sup> Higgs neutrality description.

#### <sup>320</sup> Spin

- <sup>321</sup> Importance of spin parity determination to know if it's the SM one.

#### <sup>322</sup> Mass and width

- <sup>323</sup> Mass and width measurements.

## <sup>324</sup> 1.5 Hierarchy problem and other limitations

- <sup>325</sup> The SM has been one of the most successful theories on the history of physics.  
<sup>326</sup> With only 19 free parameters, is able to make thousands of predictions that have  
<sup>327</sup> been measured and tested over the last seventy to eighty years. However some  
<sup>328</sup> aspects in the model are not completely understood. The most important one  
<sup>329</sup> is the so-called hierarchy mass problem. At tree level, the Englert-Brout-Higgs  
<sup>330</sup> boson has a mass  $m_H^2 = -2\mu^2$ , but the physical mass also contain the one-loop

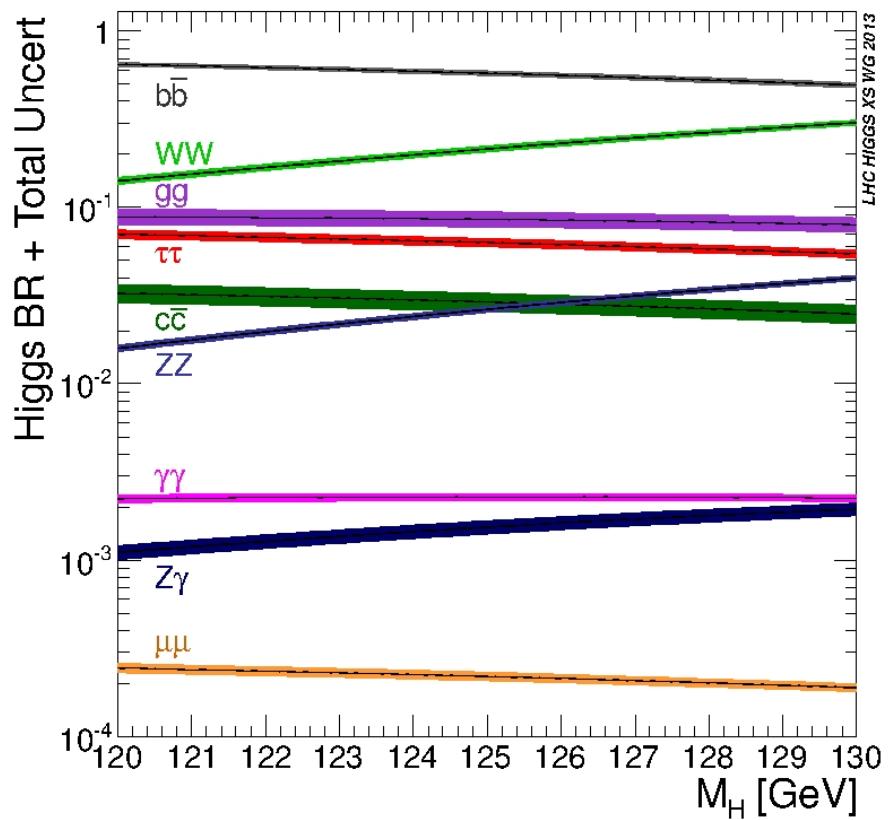


Figure 1.11: Higgs decay branching ratios.

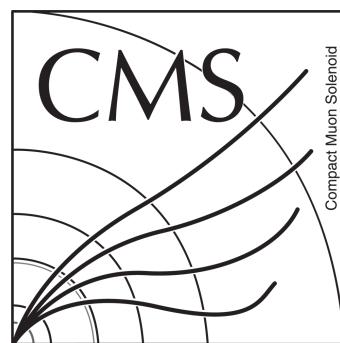


Figure 1.12: Higgs spin parity measurement

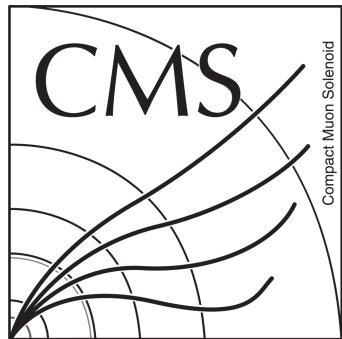


Figure 1.13: ATLAS and CMS combination of Higgs mass measurement [left] and signal strength for searches performed by CMS in different Higgs channels [right]



Figure 1.14: Higgs mass observation by CMS collaboration in the diphoton [left] and golden channels [right].

331 contributions from fermions that interact with it, as the top quark. The Feynman  
332 diagram for such contribution can be seen in figure 1.15.

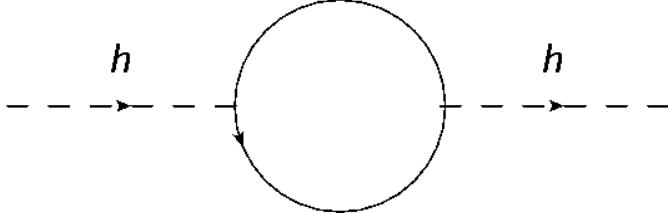


Figure 1.15: One loop diagram for contributions to the mass of the Englert-Brout-Higgs boson from interactions with fermions

333 Such contributions add up giving a mass greater than simple tree level mass.  
334 Each fermion contributes proportionally to its mass, what means that the top  
335 quark contributes the most. Moreover, if there are in nature heavier fermions that  
336 also interact with the Englert-Brout-Higgs boson they will also contribute to its  
337 mass. With such considerations one can expect the Englert-Brout-Higgs boson to  
338 be much greater than 125 GeV, and in principle not even of the order of 100 GeV  
339 but greater than 1 TeV. However the real relevance or significance of this problem  
340 at theoretical level has been discussed extensively, for example at [9], the majority  
341 of the community agrees there is something to be understood on the subject.

342 The most famous proposed solution to this problem is supersymmetry (SUSY) [10].  
343 It proposes the existence of an additional symmetry between fermions and bosons,  
344 at a given point of the history of universe nature didn't distinguish between  
345 fermions and bosons. However, we know this does not happen at the present,  
346 and then this symmetry should be broken. Such symmetry implies the existence  
347 of a super-symmetric partner for each particle, a super-partner. A fermion for each  
348 boson and vice versa. This SUSY procedure doubles the particle content of the  
349 model where it's applied. Before breaking SUSY, a particle and it's partner have  
350 the same mass. In this feature is where the hierarchy problem is solved. On fig-  
351 ure 1.16 one can see the one loop diagrams for the mass of the Englert-Brout-Higgs  
352 boson from the top and its super-partner the stop. Whereas, the top contribution  
353 is positive, the stop contribution is negative but equal in value, then cancelling  
354 between them.

355 But this solution works exactly only if SUSY is not broken. As we know SUSY  
356 has to be broken, there has been developed in the literature different ways to brake  
357 SUSY and still offer a solution to the hierarchy problem, leading normally to solu-  
358 tions that need a fine adjustment of the parameters of the theory. This represents  
359 for some theoreticians a problem itself: Fine-tuning or Naturalness. Extensive

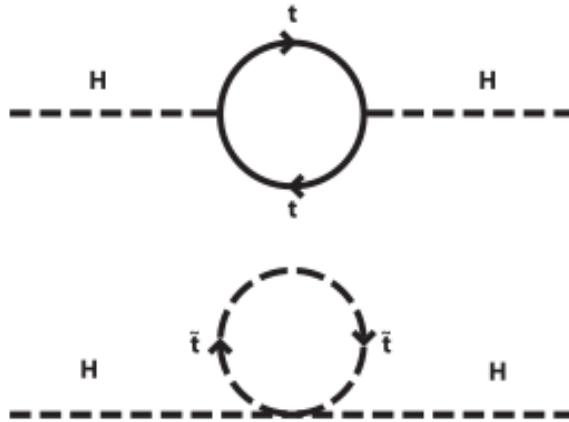


Figure 1.16: One loop diagrams for contributions to the mass of the Englert-Brout-Higgs boson from the top and the stop

360 searches for SUSY particles have been performed, accordingly to different model  
 361 realizations MSSM [11, 12], CMSSM, etc.

362 While hierarchy problem is an internal problem of the SM, there are several  
 363 questions that have not been solved. For example, how gravity is understood  
 364 in the frame of QFT's, why there is only 3 generations of leptons and quarks,  
 365 why there is only 4 fundamental forces among others. In addition, there have  
 366 been experimental questionings to the SM. The mos important one is the masses  
 367 of neutrinos. In the SM neutrinos are massless, careful measurements, [13, 14],  
 368 have shown that neutrinos can oscillate between different flavors, phenomenon  
 369 only possible if neutrinos have a mass. Measurements of solar and atmospheric  
 370 neutrino oscillations have been the most important proof of physics beyond the  
 371 SM.

372 From cosmological measurements, the Wilkinson Microwave Anisotropy Probe  
 373 have shown that the universe is not only made by visible matter, but suggests that  
 374 around 24% its made of dark matter. A type of matter not visible by means of  
 375 light. It has also shown that 71% of the universe is composed of dark energy, what  
 376 makes the universe to be in an accelerated state of expansion. These results can  
 377 be seen in [15, 16]. Also the Planck probe has shown similar results, for example  
 378 in [17]. The SM does not have any answer to this open problems so far.

379 Finally, there is known that the universe presents an asymmetry between mat-  
 380 ter and antimatter, being the first much more abundant than the second. Such  
 381 asymmetry can be obtained by CP-violating processes (C for charge and P for

382 parity). However the amount of CP violation present in the SM is not compatible  
383 with the huge matter-antimatter asymmetry in nature. This problem, known as  
384 baryon asymmetry, represent an additional huge challenge for particle physics.

385 In conclusion, the SM has been a formidable model that has helped us to under-  
386 stand a huge amount of physics. It has done thousands of predictions that have  
387 been measured and corroborated one by one in the last half-century. However,  
388 this is not the end of the story, perhaps only the beginning. There are theoretical  
389 and experimental motivations that lead us think that the SM is not the “final”  
390 theory that could explain all subatomic phenomena in nature. Currently, there is  
391 a mayor effort, both theoretical and experimentally, to understand and explain all  
392 the remaining pieces. The present work is one of them.

393 In the next chapter, we present an extension of the SM that looks for a solution  
394 to the discussed hierarchy problem.

## <sup>395</sup> 2 The CMS experiment at LHC

<sup>396</sup> The CMS experiment is one of the biggest particle physics experiments on the  
<sup>397</sup> world. It is located at the ring of the LHC that is the main accelerator managed  
<sup>398</sup> by CERN, the European Organization for Nuclear Research or Centre Européenne  
<sup>399</sup> pour la Recherche Nucléaire by its french name. This center constitutes the biggest  
<sup>400</sup> center for research on particle physics all over the world. All along its 60 years of  
<sup>401</sup> existence, from 1954, 21 member states have been joining it, but an overall of 113  
<sup>402</sup> countries participate in different ways on this center.

<sup>403</sup> On the present chapter we discuss in detail different aspects of the LHC acceler-  
<sup>404</sup> ator and the CMS experiment. In particular we make some emphasis in the CMS  
<sup>405</sup> sub-detectors related to jets, objects that play the main role on the search that is  
<sup>406</sup> the main subject of the present work. We also discuss the present state of both  
<sup>407</sup> machines, their achievements and the challenges that were overcome. Finally, also  
<sup>408</sup> the expectations and goals for the upcoming run II are mentioned.

### <sup>409</sup> 2.1 The Large Hadron Collider

<sup>410</sup> The Large Hadron Collider, or LHC [1], is a machine that accelerates and collides  
<sup>411</sup> protons and lead. This machine is the biggest particle collider nowadays with a  
<sup>412</sup> circumference of 27 km. It also achieves the highest energy by a collider up to  
<sup>413</sup> present, planned to be 14 TeV at the center of mass of the collision. On the first  
<sup>414</sup> run of the machine only 8 TeV were achieved, and next run is planned to start  
<sup>415</sup> with 13 TeV. It's located in French-Swiss border near to Geneva. The tunnel for  
<sup>416</sup> the machine was carved around 100 m under the ground, 45 m under the Jura  
<sup>417</sup> mountains and 170 m under the Léman lake with an inclination of around 1.4%,  
<sup>418</sup> sloping down towards the lake . This machine has used as much as possible old  
<sup>419</sup> LEP buildings and sites, that was an electron-positron collider built between 1984  
<sup>420</sup> and 1989.

<sup>421</sup> The protons and heavy ions accelerated by the machine collide in different points  
<sup>422</sup> where dedicated experiments are located to detect and study the product from the  
<sup>423</sup> collisions. The four main experiments located on the LHC ring are CMS [2, 3],  
<sup>424</sup> ATLAS [4], LHCb [5] and ALICE [6]. The first two are experiments of generic  
<sup>425</sup> purpose where searches for new physics and also precision measurements are per-  
<sup>426</sup> formed. LHCb is dedicated to the physics of the b-quark, and ALICE focuses on

427 the study of the quark-gluon plasma produced from heavy ions collisions. Even if  
428 one of the principal objectives of the construction of the LHC was the search for  
429 the Higgs boson, generic searches on new physics have been conducted from the  
430 very beginning of the first data taking in 2009. Moreover, after the Higgs discov-  
431 ery in 2012 there is a growing effort on the searches for new physics and precision  
432 measurement of the properties of the Higgs.

433 The LHC is a complex machine composed of several parts. The two principal  
434 parts are the injector chain and the main ring. A diagram of the whole CERN  
435 complex can be seen in figure 2.1. The injector chain has different stages that  
436 pre-accelerate protons and heavy ions to be injected into the main ring of LHC.

### 437 **2.1.1 Injector chain**

438 The injector chain begins with the proton source. Protons are extracted via ioniza-  
439 tion of Hydrogen gas in the Duoplasmatron Proton Ion Source. Such extraction is  
440 pulsed, what makes up the first bunch structure. The extracted protons are then  
441 accelerated up to 50 MeV in the linear accelerator, Linac2, that dates from 1978.  
442 After this first stage several steps are followed:

- 443 1. Linac2 injects proton bunches in the Proton Synchrotron Booster (PSB)  
444 where they are accelerated to 1.4 GeV.
- 445 2. From PSB, the protons are delivered to the Proton Synchrotron (PS) where  
446 they reach an energy of 25 GeV. In the PS the bunches are also split from 6  
447 initial bunches to 72 spaced by 25 ns.
- 448 3. Finally, the pre-acceleration chain is finished by the SPS, Super Proton Syn-  
449 chrotron. There the bunches are accelerated up to 450 GeV right before  
450 being inserted into the main LHC ring.

451 The whole pre-acceleration chain has been optimized to obtain the best possible  
452 performance on the final acceleration in the LHC main ring. All parameters are  
453 carefully controlled, for example the number of bunches, the separation between  
454 bunches, the separation between trains of bunches or the injection energy to each  
455 subsystem. It's also remarkable to notice the level of control achieved in the  
456 bunches manipulation, from old subsystems as the PS from 1959 or the newest,  
457 the SPS that dates from 1976.

458 Some recent plans for future accelerator have been studied using the LHC main  
459 ring as injector for a bigger accelerator, for example the so called FCC (Future Cir-  
460 cular Collider) at CERN. The FCC could be built perform proton-proton, electron-  
461 positron or electron-proton collisions, versions that are called respectively FCC-hh,  
462 FCC-ee and FCC-he. The FCC-hh is being designed to achieve 100 TeV of center  
463 of mass energy in a tunnel of 80-100 km of circumference.

## CERN Accelerator Complex

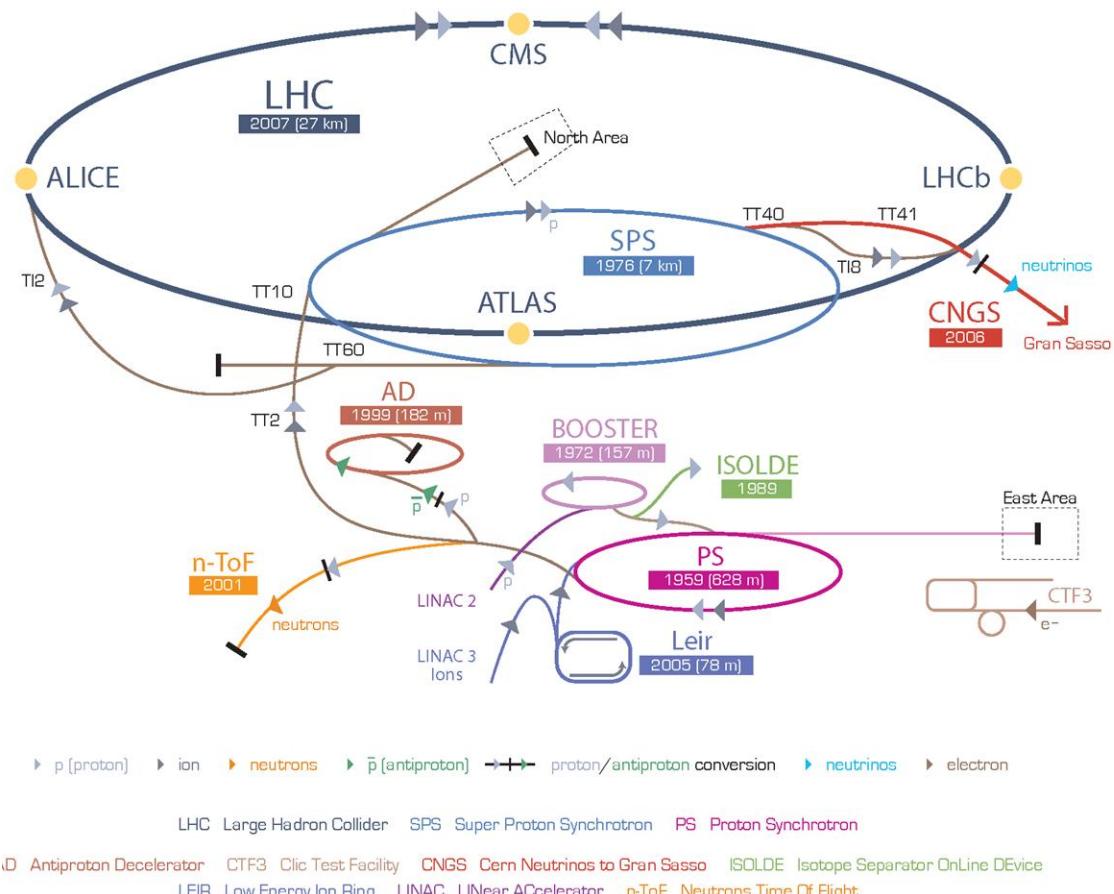


Figure 2.1: Organization of CERN accelerator complex

464 **2.1.2 Main ring**

465 The main ring is composed of two rings that accelerate the proton bunches in  
466 opposite directions, clock-wise and counter clock-wise. An schematic view of the  
467 design of the main ring can be seen in figure 2.2. The rings crosses in different  
468 points in order to collide the protons and they are divided in eight straight sections  
469 and eight arcs. In each octant bunches are controlled by dipole magnets. These  
470 complex magnets, in figure 2.3, need to produce a very strong magnetic field in  
471 order to be able to bend a 7 TeV beam of protons. This intense magnetic field, 8.33  
472 T, in opposite directions, is produced by electrical currents that are only achievable  
473 by means of superconductivity. All the 1232 dipoles operate at a temperature of  
474 1.9 K, under cooling by liquid helium. They also operate under ultra-high-vacuum.  
475 The beam lines with a pressure less than  $10^{-9}$  mbar and the whole dipole system  
476 with  $10^{-6}$  mbar, that serves also as insulating system from the surroundings. In  
477 addition, the LHC main ring has other magnets that focus and correct different  
478 characteristics of the beam: 520 quadrupoles, 2464 sextupoles, 1232 octupoles.

479 **Luminosity**

480 In collider physics, such as the LHC, the figure of merit is the luminosity, given in  
481 equation 2.1. The number of events per second is proportional to the luminosity,  
482 hence is the quantity to be maximized by the design and operation of the acceler-  
483 ator. The collider characteristics depend on the number of bunches in the ring  $n_b$ ,  
484 the number of protons per bunch  $N_b$ , the revolution frequency  $f_{rev}$ , the relativistic  
485 gamma factor  $\gamma$ , the normalized rms transverse beam emittance  $\epsilon_n$  and the beta  
486 function at the interaction point  $\beta^*$ . The denominator on 2.1 can also be rewritten  
487 in terms of the horizontal and vertical width of the bunches at the crossing,  $\sigma_x^*$   
488 and  $\sigma_y^*$ . In addition, there is the geometric reduction factor ( $R$ ) that introduces  
489 a dependence on the crossing angle of the bunches at the interaction points. In  
490 table 2.1.2 can be found the LHC beam parameters at injection and collision.

$$L = \frac{n_b N_b^2 f_{rev} \gamma}{4\pi \epsilon_n \beta^*} R = \frac{n_b N_b^2 f_{rev}}{4\pi \sigma_x^* \sigma_y^*} R \quad (2.1)$$

491 At the crossing points, the number of events coming from collisions and produced  
492 via a specific process, is directly proportional to the luminosity provided by the  
493 collider, as in equation 2.1.

$$N_{events} = L \sigma_{process} \quad (2.2)$$

494 where  $\sigma_{process}$  is the cross section of the process.

495 The total cross section of a proton-proton collision from the crossing of two  
496 bunches at 14 TeV is 100-110 mb [7], from three different scattering processes:

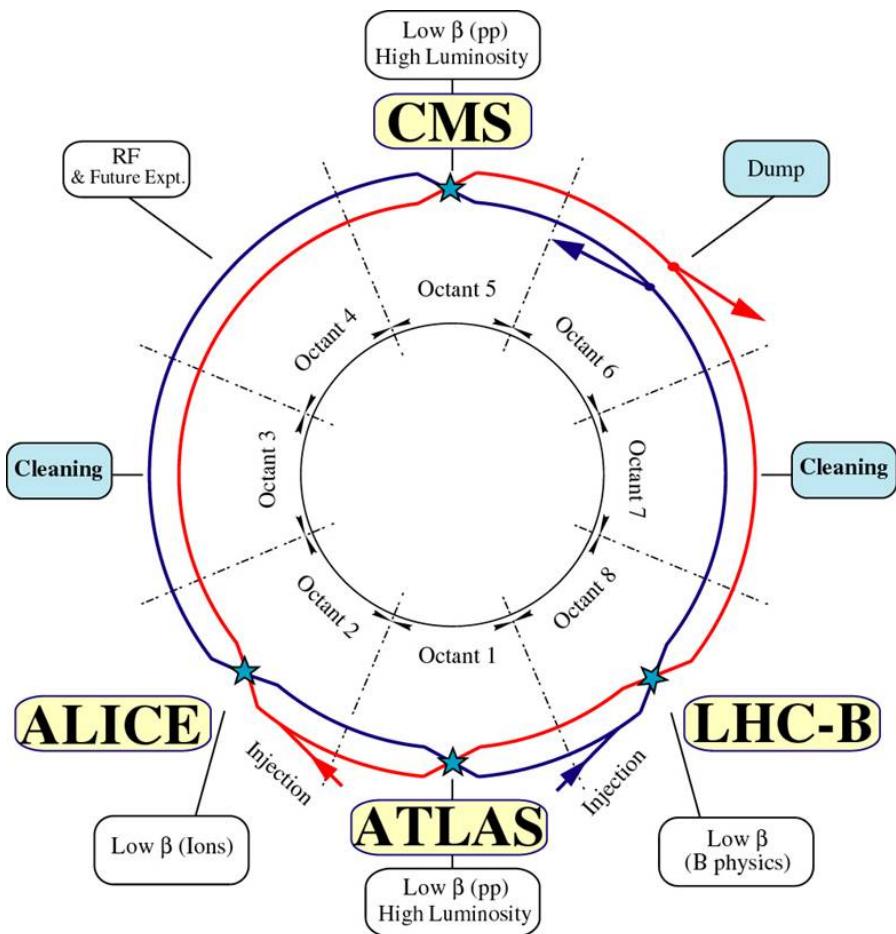


Figure 2.2: Schematic of the LHC main ring design.

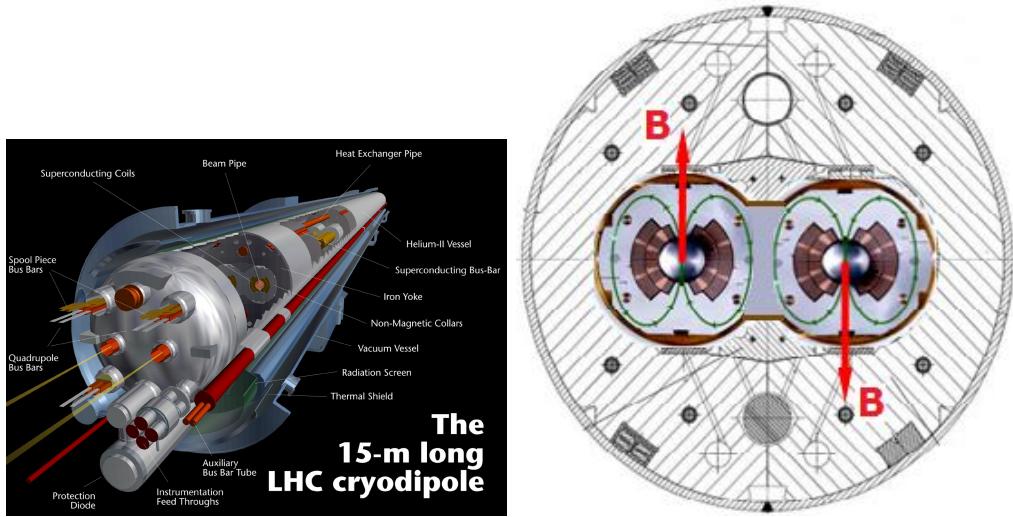


Figure 2.3: Design of LHC cryodipole and the magnetic field that bends the beam in the main ring.

497 elastic, diffractive and inelastic. In the elastic scattering the protons only ex-  
 498 change momenta but their structure remain unchanged, that is the case for the  
 499 majority of collisions. In diffractive scattering momenta is exchanged and also new  
 500 particles are produced in addition to the two final protons. Finally, in inelastic  
 501 scattering, the constituents of the protons, the partons, interchange a big amount  
 502 of momentum and produce a large quantity of particles. The inelastic processes  
 503 contribute less than diffraction to the total cross section. While inelastic collisions  
 504 produce particles in the central rapidity (defined in 2.2.1) region, diffractive and  
 505 elastic final products have a large rapidity. Only in the hard interactions, inelas-  
 506 tic scattering, color is exchanged, being the reason to fill up the central rapidity

Parameter/units	Injection	Collision
Energy [GeV]	450	7000
Luminosity [ $\text{cm}^{-2}\text{s}^{-1}$ ]		$10^{34}$
$k_b$ Number of bunches		2808
Bunch spacing [ns]		24.95
$N_b$ intensity per bunch [protons/bunch]		$1.15 \times 10^{11}$
Beam current [A]		0.58
$\epsilon_n$ normalized rms transverse beam emittance [ $\mu\text{m}$ ]	3.5	3.75
$f_{rev}$ revolution frequency [kHz]		11.25

507 region.

508 From the crossing of two bunches not only one proton-proton interaction is  
509 expected. In average, 25 interactions are expected for each crossing. From them,  
510 only one is coming from an inelastic collision, that is the type of process of more  
511 interest for detectors as CMS or ATLAS. This fact puts an additional difficulty  
512 to the detectors in order to extract the hard interaction from all the elastic and  
513 diffractive collisions happening at same time. Such phenomena is known as Pile-  
514 Up, an illustration of a collision with high pile-up can be found on figure 2.4 as  
515 seen by the CMS detector.

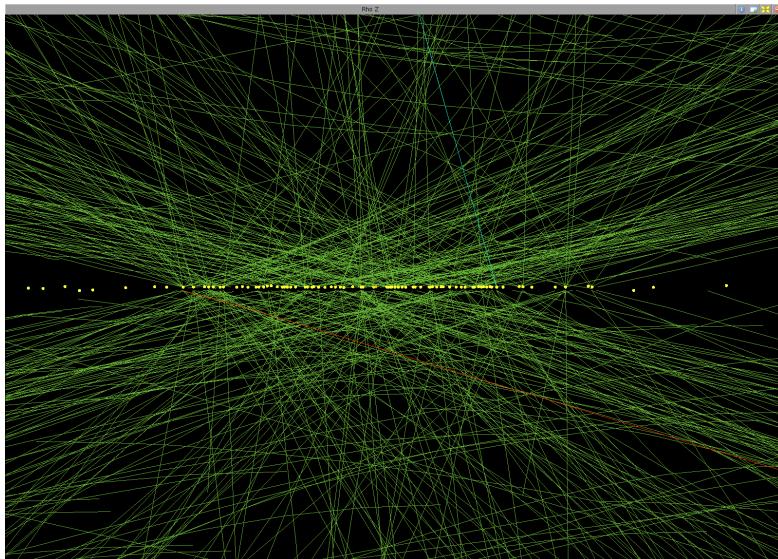


Figure 2.4: High pile-up event (78 interactions) seen by CMS detector. Event 35655522, from 198609 run, lumi 56, recorded on 2012. Image credit: Andre Holzner ©CERN

### 516 2.1.3 Run 1

517 On February 10 of 2013 the first stable run of the LHC reached an end. This run,  
518 now called Run 1, started on November 20 of 2009. LHC was originally planned  
519 to start in 2008, but an incident on one of the electric connections of one of the  
520 magnets forced to stop on the 19th of September of the same year. From the  
521 restart in 2009, the energy was augmented from 450 GeV to 4 TeV per beam. The  
522 23th of September 2009 the first collisions were detected by the experiments. One  
523 week after, the achieved center of mass energy was  $\sqrt{s} = 2.36$  TeV, already higher  
524 than Tevatron (0.98 TeV).

525 In 2010, from 30th March to 6th December 3.5 TeV per beam were reached de-  
 526 livering near  $50 \text{ pb}^{-1}$ . With the same energy, approximately  $6 \text{ fb}^{-1}$  were delivered  
 527 in 2011.

528 In 2012, the center of mass energy reached one additional TeV,  $\sqrt{s} = 8 \text{ TeV}$ ,  
 529 and around  $20 \text{ fb}^{-1}$  of integrated luminosity were delivered between April and  
 530 December. On figure 2.5 can be seen the progress of the recorded luminosity by  
 531 CMS for 2010-2012 period. The first six weeks of 2013 were devoted to proton-lead  
 532 collisions.

533 After this very successful run, the LHC has been stopped for more than a year  
 534 for repair and maintenance of different systems in the experiments and in the LHC  
 535 itself to achieve higher energies. After this period, known as Long Shutdown or  
 536 LS1, the LHC is planned to restart a new run on the early spring of 2015.

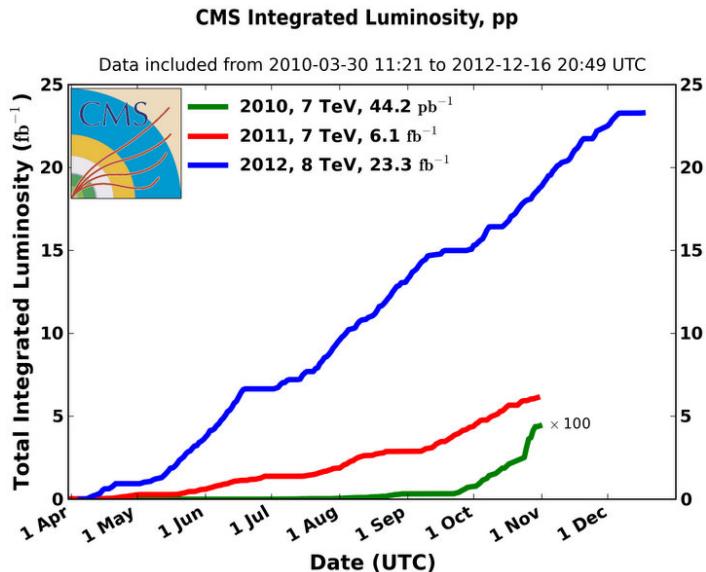


Figure 2.5: CMS integrated luminosity for proton-proton collisions delivered by LHC. ©CERN

### 537 2.1.4 Other experiments at LHC

538 On the main ring there are several experiments depending on the collisions de-  
 539 livered by the LHC main ring. The biggest are CMS [2] and ATLAS [4], both  
 540 of them generalist experiments designed to do precision measurements as well as  
 541 new physics searches. Mainly recording proton-proton collisions, they have also

542 recorded lead-lead and proton-lead collisions during the run 1. Both of them were  
543 designed for high instantaneous luminosity,  $L = 10^{34} \text{cm}^2\text{s}^{-1}$ .

544 In addition, there are two other experiments designed for specific purposes. The  
545 LHCb [5] that focus on the study of the physics of the b-hadrons, specially related  
546 to the CP violation, and ALICE [6] built for the study of strongly interacting matter.  
547 The first of them record proton-proton collisions at an instantaneous luminosity  
548 of  $10^{32} \text{cm}^2\text{s}^{-1}$  and the second record ion-ion collision with  $L = 10^{27} \text{cm}^2\text{s}^{-1}$ .

549 The CMS experiment is going to be described in detail in section 2.2. In the  
550 following sections we are going to present very briefly the other three experiments  
551 mentioned above.

## 552 ATLAS

553 The ATLAS experiment (A Toroidal LHC ApparatuS) is the biggest LHC experiment.  
554 It's located at point one, as displayed on figure 2.2, on the LHC main ring.  
555 It's a cylindrical detector similar to CMS, about 45 meter long, 25 meter high, and  
556 weights around 7000 tons. ATLAS main components are, from inside to outside, a  
557 tracking system, an electromagnetic calorimeter, a hadron calorimeter and muon  
558 chambers. In between these subsystems there is an internal solenoidal magnet and  
559 a set of external toroidal magnets. The detector design is presented on figure 2.6.

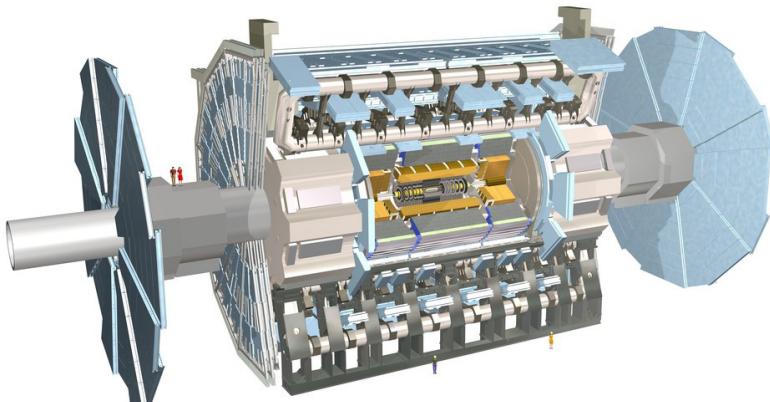


Figure 2.6: ATLAS detector internal view. ©CERN

560 On the human resources side, ATLAS experiment configures a collaboration  
561 of around 3000 persons, coming from 117 universities around the world, from 38  
562 countries. Thirty percent of the collaboration are students.

563 **LHCb**

564 LHCb detector, hosted at point 8 of the LHC main ring, has a different design than  
565 CMS and ATLAS. Smaller than these, its design mainly focus to be able to detect  
566 particles produced close to the beam direction. This is the reason why it is not  
567 cylindrically but conically shaped, in two detection arms 2.7. It also has the same  
568 main parts, a tracking system with a very precise vertex locator, electromagnetic  
569 and hadron calorimeters, muon chambers and magnets. Its major specificity is  
570 a system that allows to identify different hadrons, the RICH detectors, a crucial  
571 feature for the study of strong interacting matter. It measures 21 m long, 10 m  
572 high and 13 m wide, and weights 4500 tons. A view of the detector can be found on  
573 figure 2.7. The LHCb collaboration groups around 700 persons from 50 different  
574 universities over 15 countries.

575 **ALICE**

576 The ALICE experiment (A Large Ion Collider Experiment) is located at point 2  
577 of the LHC main ring, measures 16 m high, 16 m wide and 26 m long, and weights  
578 10000 tons. Designed for heavy ion physics, it is able to detect an extremely high  
579 number of tracks per event. Its main subsystem is the Time Projection Chamber  
580 (TPC), a 90 m<sup>3</sup> gas chamber filled with a mixture of Ne, CO<sub>2</sub> and N<sub>2</sub> operated in  
581 a solenoid of 0.5 T. It allows to measure leptonic and hadronic charged particles in  
582 a momentum range from 0.5 to 10 GeV/c. The experiment structure can be seen  
583 on figure 2.8. ALICE collaboration counts around 1500 people, from 154 physics  
584 institutes in 37 countries.

585 **2.2 The Compact Muon Solenoid (CMS)**  
586 **experiment**

587 The CMS detector, hosted at point 5 of the LHC main ring (see figure 2.2), is the  
588 second biggest LHC experiment. Cylindrically shaped, measures 15 m of diameter  
589 and 28.7 m long, and weights 14000, making it the heaviest LHC experiment. Its  
590 subsystems are concentrically located from the collision point in the beam line. It's  
591 called compact because the whole calorimetry is inside the solenoid magnet, and  
592 muon solenoid because it has a very precise muon detection. Its main characteristic  
593 is the strong 3.8 T solenoid magnet. A representation of the detector can be found  
594 in figure 2.9. The CMS collaboration is formed by around 2600 scientists, of which  
595 900 are students, from 181 institutes over 41 countries.

596 CMS has been designed to be able to do very precise identification of parti-  
597 cles originated from the collisions and their properties. For the measurement of

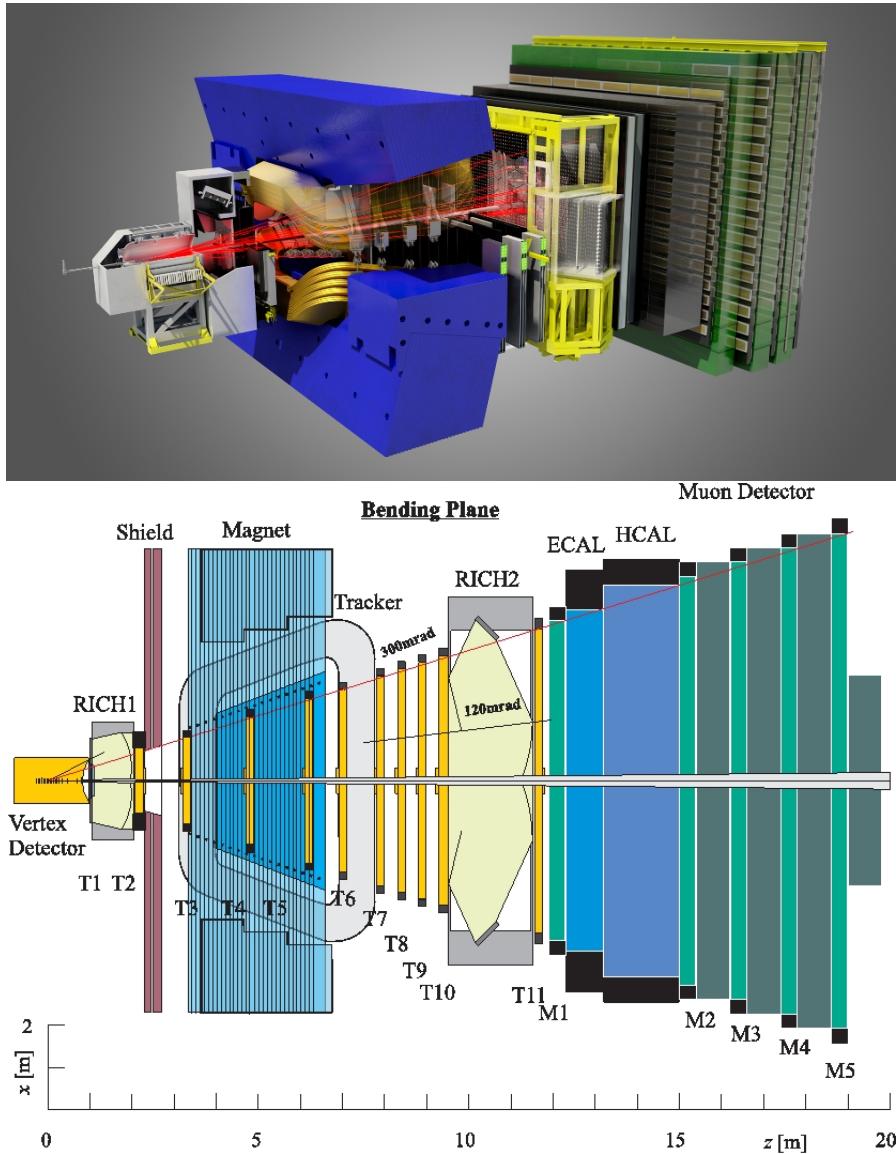


Figure 2.7: LHCb detector internal view [top] and view from the top [bottom].  
©CERN

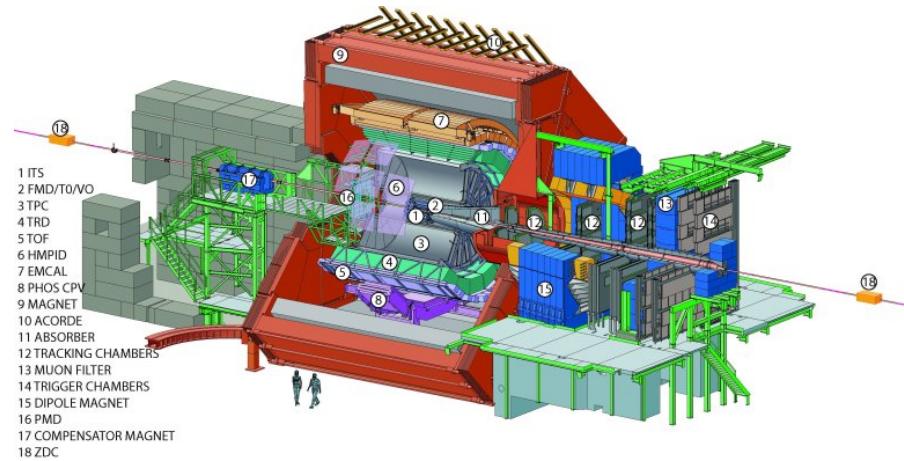


Figure 2.8: ALICE detector internal view. ©CERN

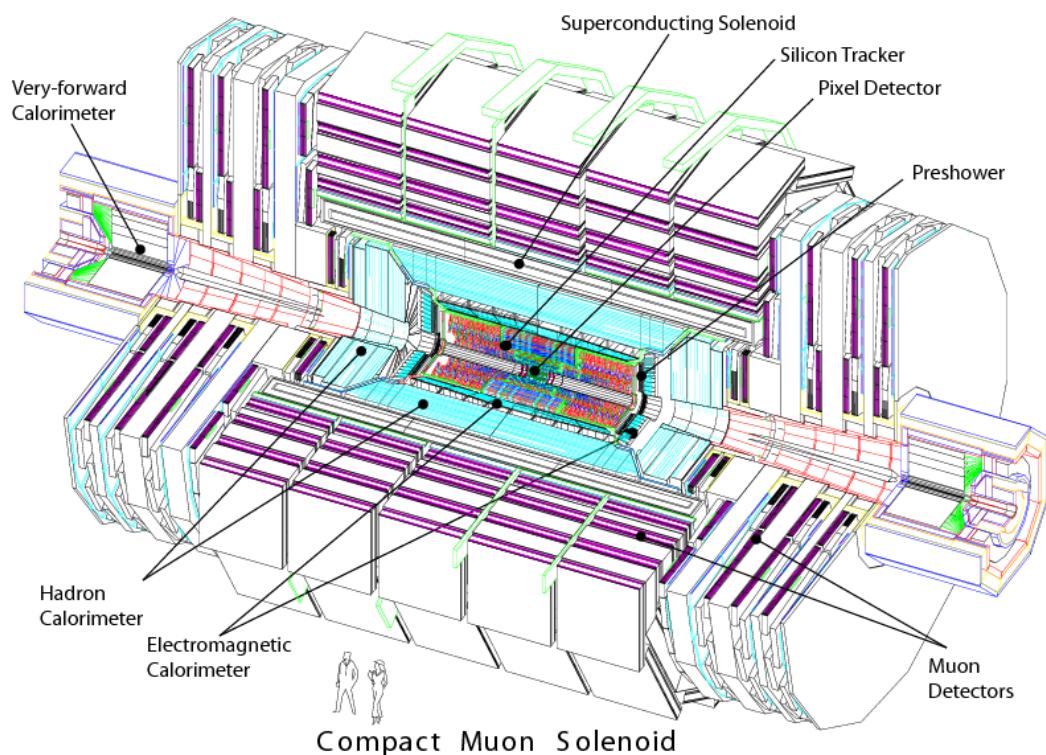


Figure 2.9: CMS detector internal view. ©CERN

598 the momentum of the charged particles, CMS counts with a very strong magnet  
 599 that allows to bend very energetic particles. In addition, the calorimeters allow to  
 600 measure accurately the energy from hadrons, electrons and photons. At the most  
 601 external layer, the muons chambers measuring muons properties, and in the in-  
 602 nermost the tracking system that reconstructs the collision points and the charged  
 603 particles tracks. In figure 2.10 can be found a representation of the different sub-  
 604 systems of CMS and how particles are reconstructed from them.

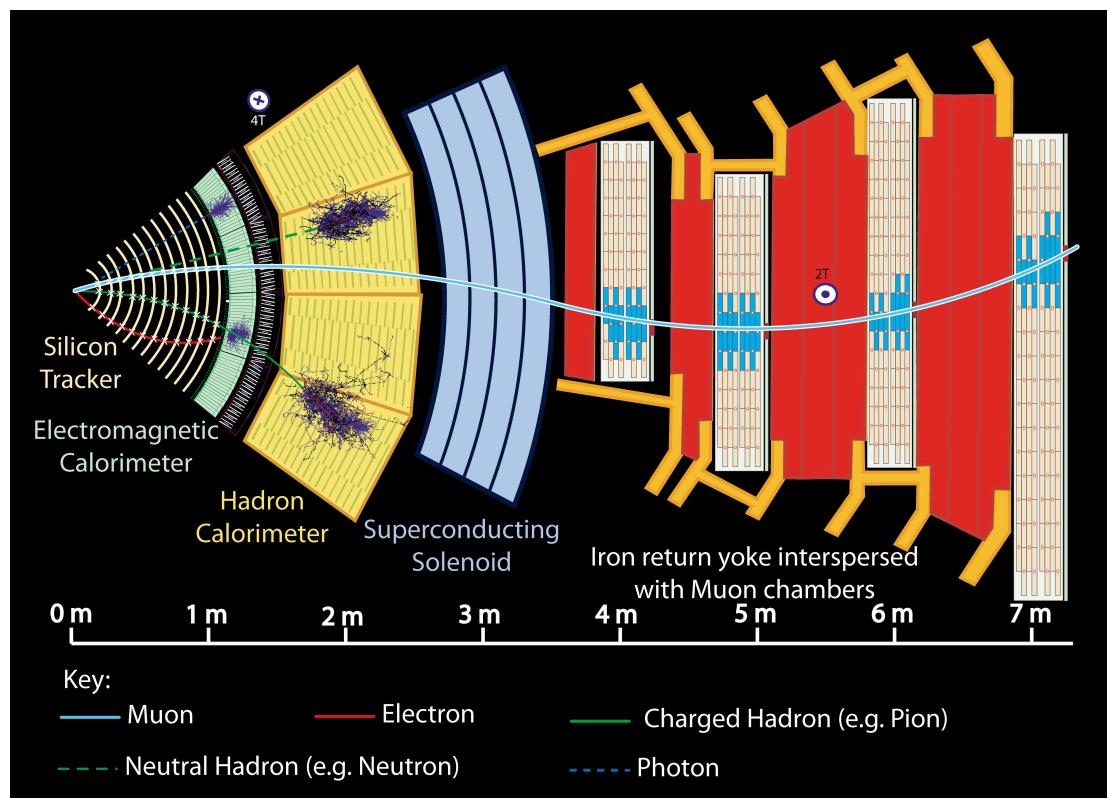


Figure 2.10: CMS sub-detectors and particle identification. ©CERN

### 605 2.2.1 Coordinate system

606 The origin of coordinates defined on the CMS detector is located on the nominal  
 607 collision point, the “interaction point”. From there, the z-axis is defined along the  
 608 beam pipe line pointing towards the Jura mountains. The positive/negative z-axis  
 609 directions define the positive/negative sides of the detector. The y-axis is defined  
 610 towards the zenith and the x-axis towards the center of the LHC ring. Due to the

611 inclination of the LHC plane, this coordinate system is slightly tilted with respect  
 612 to the true vertical. A representation of the coordinate system definition can be  
 613 found in figure 2.11.

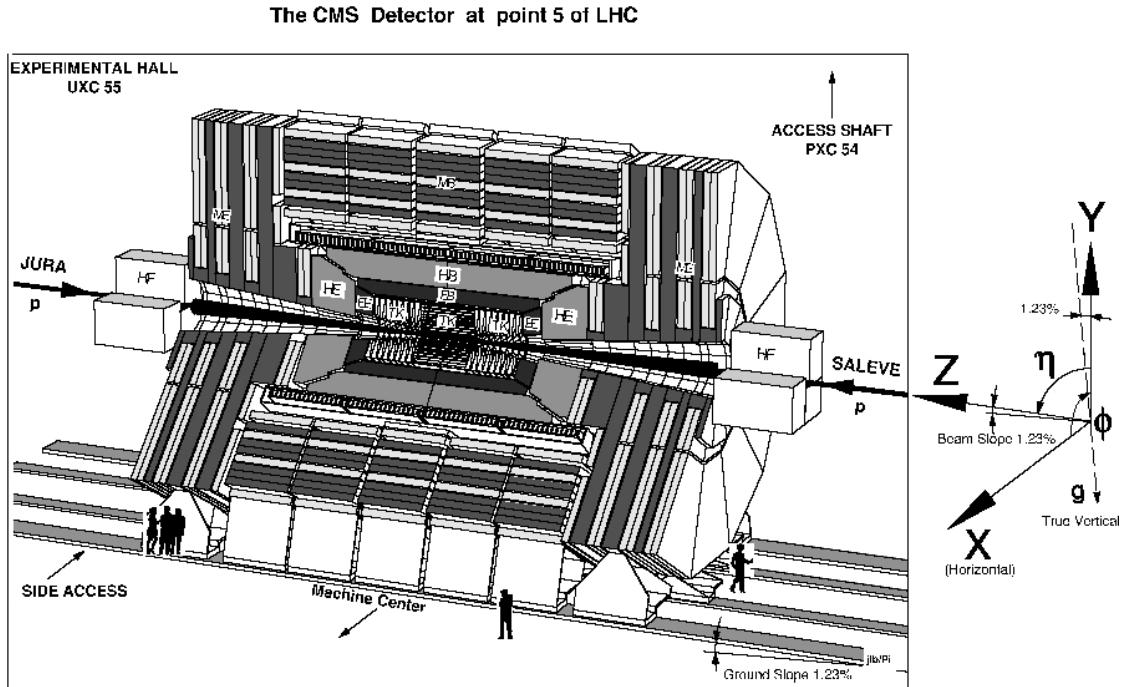


Figure 2.11: CMS coordinate system. ©CERN

614 We can also define two angles: the  $\phi$  angle in the x-y plane from the x-axis  
 615 towards the positive y-axis, and the  $\theta$  angle in the z-y plane from z-axis towards  
 616 the positive y-axis. In experimental particle physics is preferred to work with  
 617 relativistic invariant quantities, reason why instead of working with  $\theta$  we define  
 618 the pseudorapidity  $\eta$ , equation 2.3.

$$\eta = -\ln \left( \tan \left( \frac{\theta}{2} \right) \right) \quad (2.3)$$

619 One can define another relativistic invariant quantity, the rapidity  $y$  as from  
 620 equation 2.4. With  $\mathbf{p}$  being the momentum vector and  $E$  the energy of a given  
 621 particle,  $p_L$  denotes its longitudinal component, that in our case is the same z-  
 622 component.

$$y = \frac{1}{2} \ln \left( \frac{E + p_L}{E - p_L} \right) \quad (2.4)$$

623 On the limit that the mass of the particle is very small compared to its momentum,  
 624 one can replace approximate the particle energy by the momentum magnitude,  
 625 giving rise to the definition of the pseudorapidity in terms of the momentum of  
 626 the particle  $\eta = \frac{1}{2} \ln \left( \frac{p+p_L}{p-p_L} \right)$

627 We define also the radial coordinate over the x-y plane, plane that is called  
 628 the transverse plane being orthogonal to the longitudinal direction, the z-axis. In  
 629 such plane are also defined the transverse quantities of particles, as the transverse  
 630 momentum  $p_T$ . Finally, for any two objects an angular distance can be defined in  
 631 the  $\eta - \phi$  plane, as in equation 2.5.

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad (2.5)$$

## 632 2.2.2 Magnet

633 In order to measure the momentum of the charged particles going inside the de-  
 634 tector is crucial to apply the correct magnet field, sufficiently strong to bend very  
 635 energetic particles. The momentum of a charged particle inside an uniform mag-  
 636 netic field can be written as

$$p = \gamma mv = qBr \quad (2.6)$$

637 where  $B$  is the magnitude of the magnetic field,  $\gamma$  the usual relativistic factor,  $m$   
 638 the mass of the particle,  $v$  its rapidity,  $q$  its charge and  $r$  the bending radius. The  
 639 sagitta of the arc is

$$s = \frac{L^2}{8r} = \frac{qBL^2}{8p} \quad (2.7)$$

640 with  $L$  the trajectory length that the particle moved inside the magnetic field.  
 641 Inside a solenoid  $L$  is equal to the radius of it.

642 From relation 2.7 it is possible to deduce that the resolution on the momentum  
 643 of the particle has an inverse dependence with the magnetic field and the radius  
 644 of the solenoid, as shown in equation 2.8. From there, for a better resolution it is  
 645 needed to increase the magnetic field and the radius of the solenoid.

$$\frac{dp}{p} \propto \frac{p}{BL^2} \quad (2.8)$$

646 The design of CMS magnet target both features, it utilizes a large solenoid of 6  
 647 m of diameter and 13 m long. It's made of 4 layers of windings of NbTi cable that  
 648 is cooled to 4.45 K in order to achieve the superconducting state. This magnet is  
 649 able to produce an uniform magnetic field inside of it of 3.8 T. Outside the magnet  
 650 5 wheels and 3 disks of iron are placed in order to return the flux of the magnetic

651 field, inducing just a 2 T radial magnetic field outside the solenoid. This iron yoke  
652 is the main contribution to the detector weight, 10000 tons. The muon chambers  
653 are located in between the iron yoke.

### 654 **2.2.3 Tracker system**

655 The tracker system has been designed to specifically address the reconstruction  
656 of high  $p_T$  leptons, with particular interest in the isolation of electrons and, as a  
657 consequence, to isolate photons. Also the tracker fulfill granularity requirements  
658 to reconstruct secondary vertexes to tag and reconstruct B-hadrons. The tracker  
659 system is able to reconstruct tracks of particles with at least 2 GeV of  $p_T$  with  
660  $|\eta| < 2.5$ . Charged hadrons are reconstructed with an efficiency of at least 85%  
661 for  $p_T = 1$  GeV and up to 95% for  $p_T$  above 10 GeV. Another important point  
662 that was taken into account is the fact that the tracker is the part of CMS most  
663 exposed to radiation as it is the closest subsystem to the interaction point. The  
664 tracker system was built highly resistant to radiation damage and is expected to  
665 last for around 10 years. The pixel detector only lasts 2 years and was replaces  
666 during LS1.

667 The tracker has been built with two different technologies: Pixels and Silicon  
668 Strips. They are arranged concentrically in cylindrical volumes being the pixel  
669 detector the innermost. The CMS tracker extends to a radius of 1.1 m and a  
670 around 2.7 m on each  $z$  direction, reaching a coverage in  $|\eta|$  between 0 and 2.5.  
671 The pixel system is in the region with a radius below  $\approx 20$  cm and the silicon  
672 detector surrounding the pixel system.

673 The pixel system is formed by three barrel layers ( $r = 4.3$  cm,  $r = 7.3$  cm and  
674  $r = 10.4$  cm) and four end cap disks ( $z = \pm 35.5$  cm and  $z = \pm 46.5$  cm). It has  
675 an approximate active surface of one square meter with approximately  $70 \times 10^6$   
676 pixels with a cell size of 100  $\mu\text{m}$  by 150  $\mu\text{m}$ . This pixel system allows to obtain  
677 three highly precise points that are mainly used for reconstructing vertexes. Its  
678 resolution in  $r - \phi$  is shown in figure 2.12. In the same figure is shown the efficiency  
679 to find hits per each barrel layer and end cap disk.

680 The Silicon Strip system is formed by an inner ( $20 \text{ cm} < r < 55 \text{ cm}$ ) and outer  
681 tracker ( $55 \text{ cm} < r < 116 \text{ cm}$ ). The inner tracker is composed of a 4-layer barrel  
682 (TIB for Tracker Inner Barrel) and 3 disks (TID for Tracker Inner Disks) in each  
683 end cap. The outer tracker systems is composed of 6 layers in the barrel (TOB for  
684 Tracker Outer Barrel) and 9 disks in each end cap (TEC for Tracker EndCap).

685 In figure 2.13, can be seen the disposition of all the tracker subsystems. From  
686 the design of the tracker system the best resolution on the  $p_T$  measurement is  
687 achieved in the  $|\eta| < 1.6$  region, this due to the presence of more layers of detector  
688 in the different subsystems, as shown in figure 2.14.

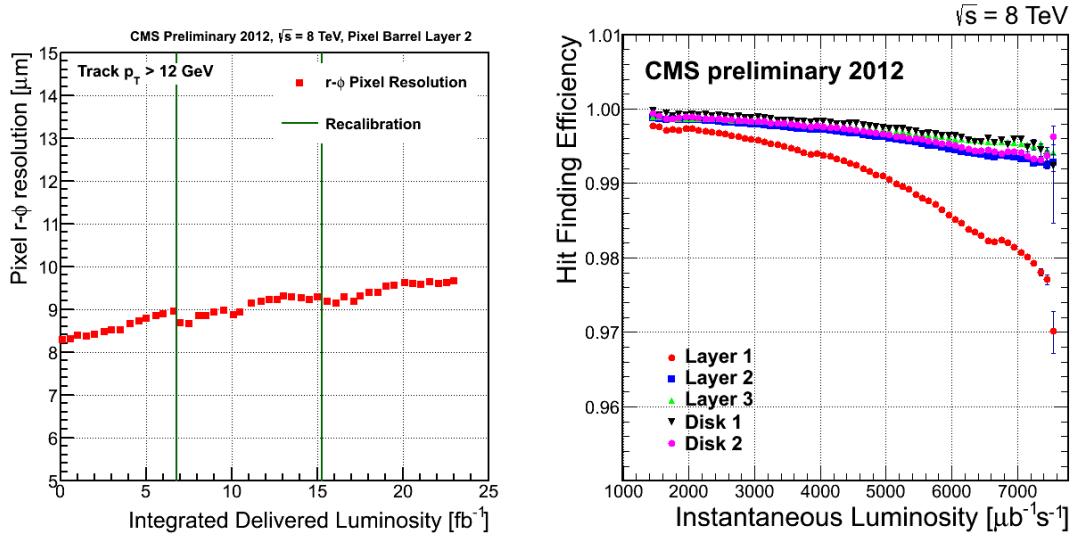


Figure 2.12: Resolution of the pixel detector in  $r - \phi$  as function of the luminosity delivered to CMS detector [left] and efficiency of finding hits by the same detector [right].

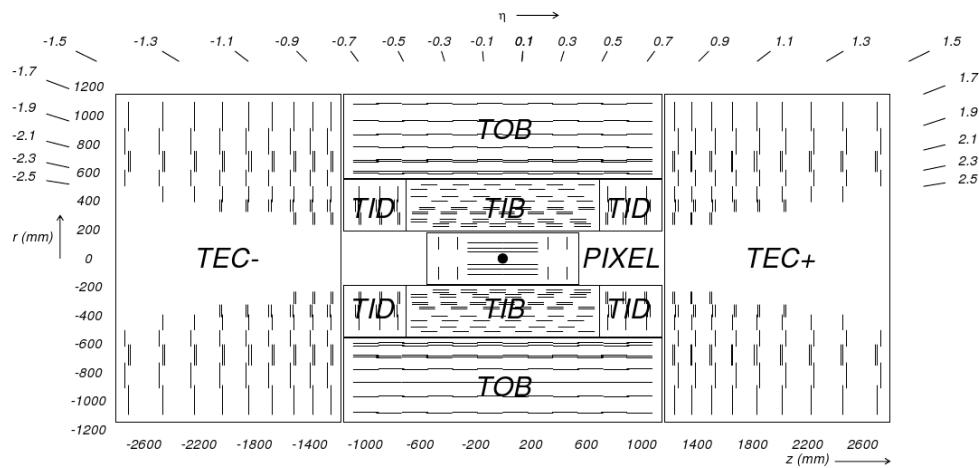


Figure 2.13: CMS tracker system configuration. ©CERN

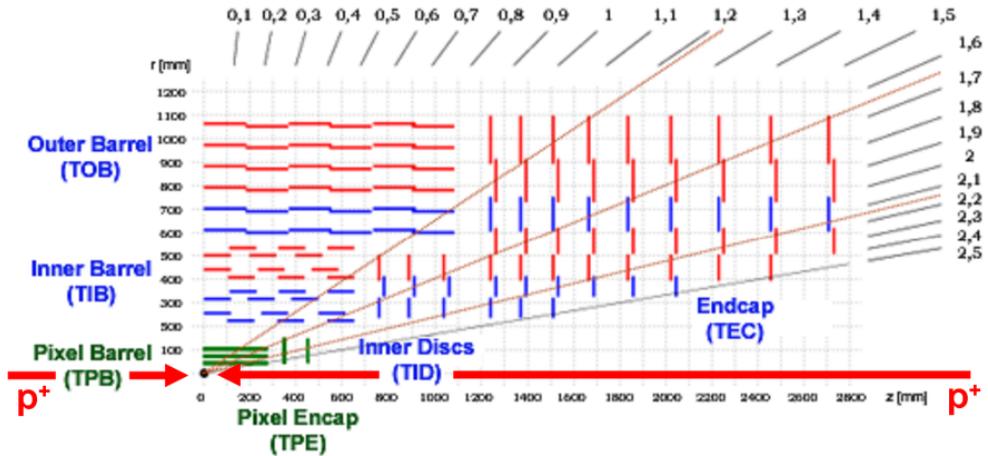


Figure 2.14: Tracker resolution with  $\eta$ . ©CERN

## 689 2.2.4 Electromagnetic calorimeter

690 The CMS ECAL (Electromagnetic CALorimeter) is the detector subsystem de-  
 691 signed to stop photons and electrons to measure their energy. It's an hermetic  
 692 cylindrical calorimeter made of 61200 crystals in the barrel ( $|\eta| < 1.479$ ) and 7324  
 693 in each end cap ( $1.48 < |\eta| < 3$ ). The crystals material is lead-tungstate that  
 694 is transparent, very dense ( $8.28 \text{ g/cm}^3$ ), has a small Moliere radius (2.2 cm) and  
 695 a short radiation depth ( $X_0 = 0.89 \text{ cm}$ ). This material has been chosen for the  
 696 characteristics already described, but also because of its fast emitting scintillation  
 697 light (about 25 ns) and its good energy resolution. The main disadvantage of this  
 698 material is its high dependent response to the temperature (about 2% /°C), mak-  
 699 ing crucial to maintain a stable temperature in the ECAL system. The crystals are  
 700 distributed in 36 super-modules, 1700 crystals each, in the barrel (EB for ECAL  
 701 Barrel) and in four 'Dee's, of 3662 crystals each, in the end caps (EE for ECAL  
 702 End cap). In the EB the scintillation light is collected by Avalanche Photo-Diodes,  
 703 or APD, and by Vacuum Photo-Triodes, or VPT, in the EE. A preshower system is  
 704 installed in face of each end cap to allow a better discrimination between photons  
 705 and  $\pi^0$ 's. A representation of the CMS ECAL can be found on figure 2.15.

706 The ECAL system requires constant correction to the lost of transparency due  
 707 to aging by radiation. To correct the crystals performance the peak of the Z from  
 708 di-electron events is used. In figure 2.16 can be seen the relative response of the  
 709 crystals during LHC run I and the di-electron mass from electrons measured in the  
 710 ECAL barrel. The aging of the crystals is controlled by a laser monitoring system  
 711 that measures the transparency of the crystals. To test the corrections derived

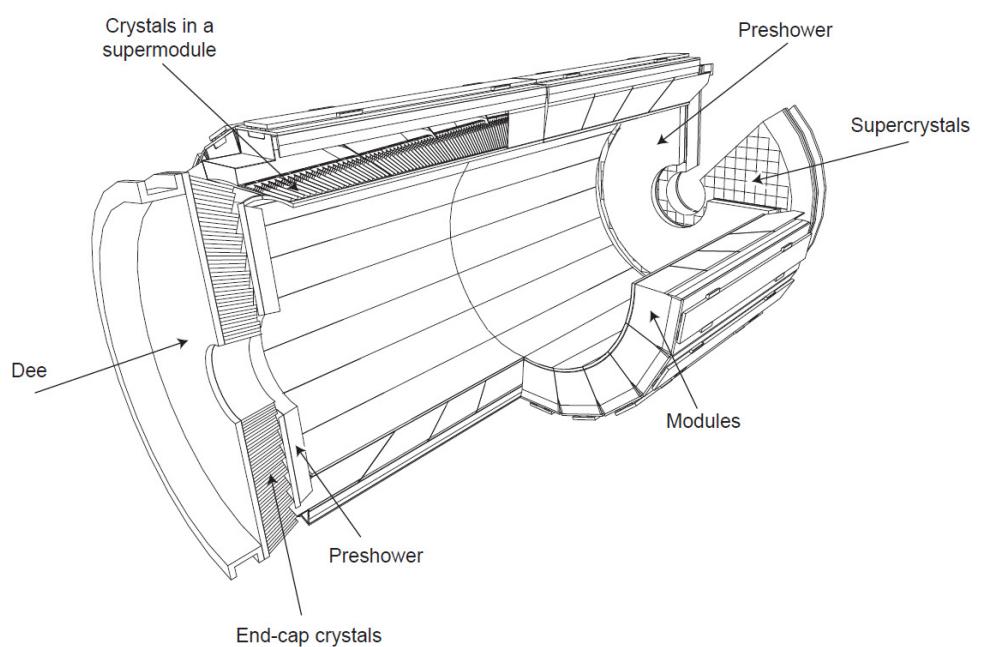


Figure 2.15: CMS ECAL representation. ©CERN

712 from the laser monitoring a comparison is performed between the reconstruction  
 713 of electrons by the tracker and the ECAL. Such comparison in function of time  
 714 can be seen in figure 2.17.

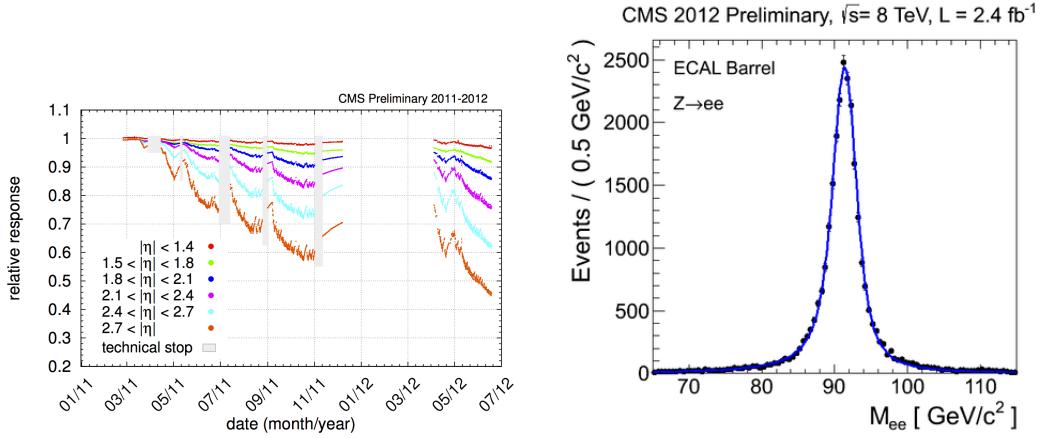


Figure 2.16: Relative response of the ECAL crystals as function of time, for different  $\eta$  regions [left]. Di-electron system mass from electrons identified in the ECAL barrel [right].

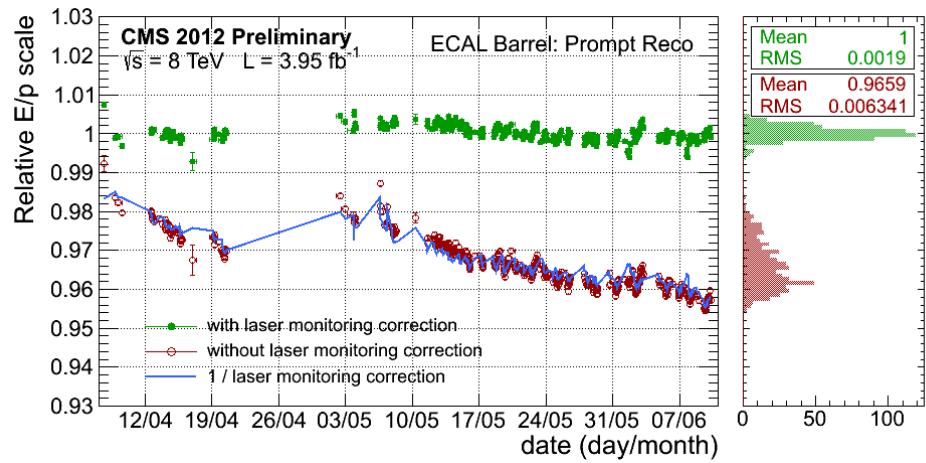


Figure 2.17: Relative  $E/p$  for electrons with and without laser monitoring as function of time. The electron energy  $E$  is measured by the ECAL while its momentum  $p$  is measured using the tracker information.

## 715 2.2.5 Hadronic Calorimeter

716 The CMS HCAL, for Hadronic CALorimeter, is the subdetector designed to mea-  
717 sure the energy of hadrons produced in the collisions, mainly the neutral hadrons  
718 because the charged hadrons are already traced by the tracker. It's also designed  
719 to measure the missing energy coming from particles not being detected by any  
720 of the subsystems, as neutrinos. It's an hermetic set of subsystems covering up to  
721  $|\eta| < 5.2$ :

- 722 • Hadron Barrel Calorimeter (HB): Covering  $|\eta| < 1.4$  is located between the  
723 ECAL barrel and the magnet.
- 724 • Hadron Endcap Calorimeter (HE): Extends the coverage of the barrel on the  
725 region  $1.4 < |\eta| < 3$ .
- 726 • Hadron Outer Calorimeter (HO): Located outside the magnet, uses it as an  
727 additional absorber.
- 728 • Hadron Forward Calorimeter (HF): Completes the coverage of the system  
729 from  $|\eta| = 3$  up to  $|\eta| = 5.2$ .

730 The CMS HCAL layout is shown in figure 2.18. The system is made of two  
731 parts, an absorber to develop the hadronic showers and a scintillator to measure  
732 the particles energy. The length scale of hadronic calorimetry is defined as the  
733 interaction length corresponding to the mean free path of an hadron in a material.  
734 The HB absorber is made of 40 mm thick steel plate, eight layers of brass plates of  
735 50.5 mm thick, six brass plates of 56.5 mm thick and a steel plate of 75 mm thick.  
736 The HE uses the same absorber but with thicker plates, of 79 mm. Between the  
737 absorber plates a plastic scintillator, Kuraray SCSN81, 3.7 mm thick, is placed.  
738 In the region with  $|\eta| < 1.6$  the achieved granularity is  $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$   
739 and  $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$  in the region with  $|\eta| > 1.6$ . This design gives a total  
740 of 70000 tiles. The produced light in the HB is collected by optical fibers and  
741 transferred to the Hybrid Photo Diodes (HPDs). This diodes were chosen thanks  
742 to their small sensitivity to the magnetic field, an important feature because HCAL  
743 is inside the solenoid magnet.

744 The HF design is very different from the rest of the HCAL subsystems. The  
745 most important challenge for the HF is the high resistance to radiation, while in  
746 the central rapidity region 100 GeV are deposited in average in the forward region  
747 is 760 GeV. For this reason it was chosen a Cherenkov detector made of quartz  
748 fibers with a steel absorber. The light produced in the HF is collected by photo  
749 multipliers.

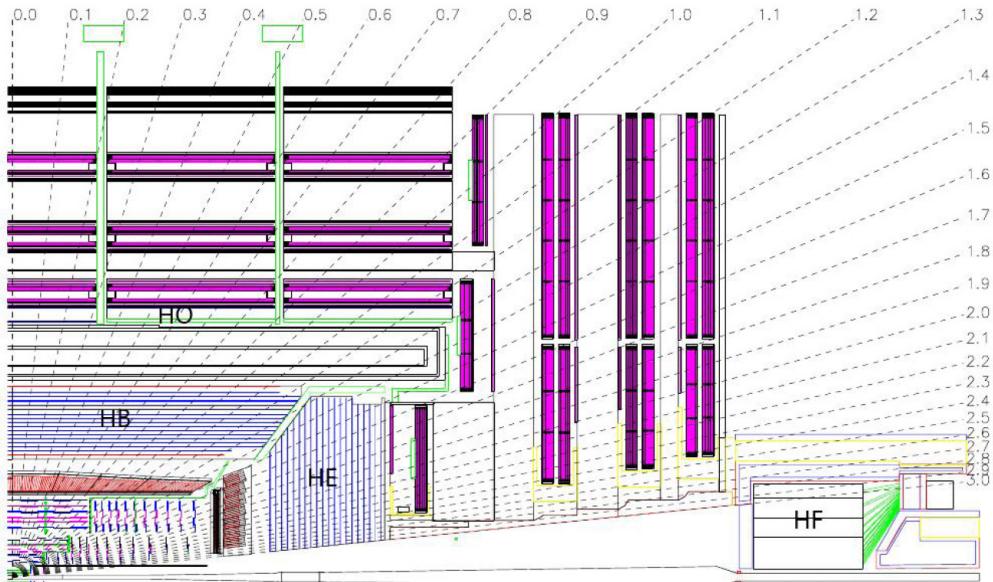


Figure 2.18: CMS HCAL representation. ©CERN

## 750 2.2.6 Muon chambers

751 The muon system of CMS is located at the most exterior layer of the detector due to  
 752 the penetration power of this particle. Muons are not stopped by the calorimeters  
 753 and, with neutrinos, they are able to escape the detector. The muon chambers  
 754 are placed in a cylinder around the HO and in disks on the end caps. Three main  
 755 characteristics have been fulfilled from the design: efficient muon identification,  
 756 precision measurement of muon charge and momentum and fast measurement  
 757 to provide trigger capabilities. The moun chambers are made of three different  
 758 subsystems:

- 759 • Drift Tubes Chambers (DT): Located in the region with  $|\eta| < 1.2$  and dis-  
 760 posed in four layers. They consist of individual drift tubes of  $50 \mu\text{m}$  of  
 761 diameter anode wire with two electrode plates creating a drift electric field.  
 762 The wall of the cell act as cathode. The cells are filled with a gas mixture,  
 763 85% Argon and 15%  $\text{CO}_2$ . The tubes are organized in plaques that are also  
 764 organized in SuperLayers (SL) each one made of 4 plaques. The barrel is  
 765 made of 250 DT's disposed in four cylinders separated by iron yokes.
- 766 • Cathode Strip Chambers (CSC): Installed in the end caps, provide a coverage  
 767 up to  $|\eta| = 2.4$  from  $|\eta| = 0.9$ . These chambers are multi-wire proportional  
 768 chambers made of six planes of anode wires with 7 cathode planes. Four

769 CSC stations are placed in each end cap. The wires are oriented in azimuthal  
 770 direction while the cathode planes are radially oriented, allowing a complete  
 771 measurement of the position of the particle. This system is able to measure  
 772 with a precision between the  $75 \mu\text{m}$  and  $150 \mu\text{m}$ .

- 773 • Resistive Plate Chambers (RPC): This subsystem is made of gaseous parallel  
 774 plate detectors. This detector is specially useful for triggering as it is very  
 775 fast and have a good position resolution. There are 480 RPC distributed in  
 776 6 layers in the barrel with the DT and in 3 layers in the end caps with the  
 777 CSC, and covers the region with  $|\eta| < 1.6$ .

778 On figure 2.19 can be found a representation of the muon system with the  
 779 different components. The DT and CSC system cover  $|\eta| < 2.4$  without any gap.

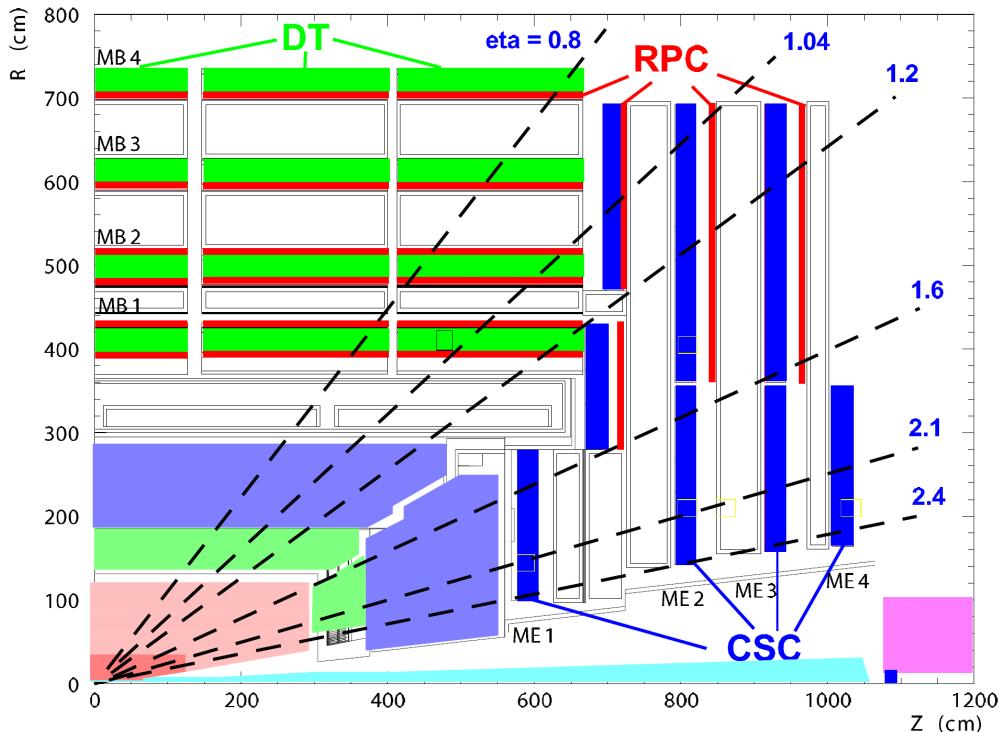


Figure 2.19: CMS muon chambers representation. ©CERN

780 The RPC subsystem efficiency depends on the voltage used and in the pressure  
 781 of the injected gas. The efficiency on the RPC barrel during the run I data taking  
 782 can be seen in figure 2.20.

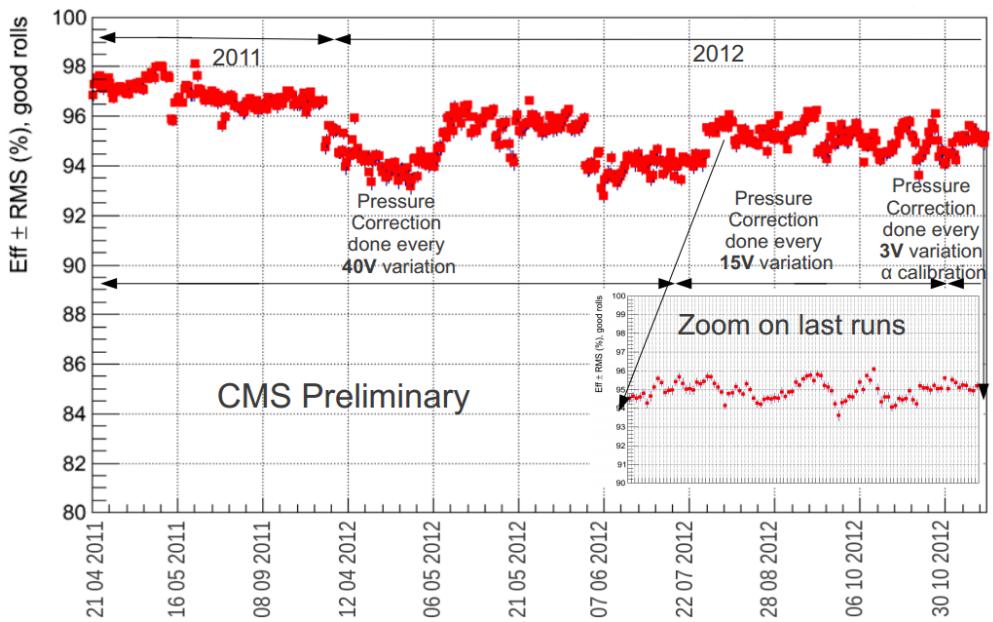


Figure 2.20: RPC Barrel efficiency during two years of data taking

### <sup>783</sup> 2.2.7 Trigger

<sup>784</sup> LHC has been designed to collect data from proton-proton collisions every 25 ns,  
<sup>785</sup> meaning a frequency of 40 MHz. Each recorded event by CMS has a nominal size  
<sup>786</sup> between 0.5 and 1 MB, corresponding to a data flow of around  $10^9$  MB/s = 1PB/s  
<sup>787</sup> that is extremely big for transfer and for storing. Therefore, an on-line selection  
<sup>788</sup> of events has to be done. The trigger system of CMS does this task in two fold,  
<sup>789</sup> a level 1 (L1) and a high level trigger (HLT). The L1 is hardware based and the  
<sup>790</sup> HLT is software based.

<sup>791</sup> From the searches conducted at CMS, the interesting events produced on proton-  
<sup>792</sup> proton collisions for new physics searches are very rare. The enormous majority of  
<sup>793</sup> events coming from proton-proton collisions correspond to well understood phe-  
<sup>794</sup> nomena, while new physics events are 'exotic' with regards to the most common  
<sup>795</sup> type of events. Then it is interesting to keep only a part of the events, what  
<sup>796</sup> actually makes easier the analyses afterward done over the data.

<sup>797</sup> The CMS trigger system is designed to keep only 100 kHz by the L1 and 300 Hz  
<sup>798</sup> by the HLT. L1 is reducing the data flux by 2 orders of magnitude and the HLT  
<sup>799</sup> another 3 orders of magnitude.

800 **Level 1 trigger**

801 The L1 is designed to trigger over coarse data coming from the calorimeters and  
 802 muon chambers, holding data in pipe-lined memories in the front-end electronics.  
 803 Therefore, relies on very fast reconstruction of objects coming from this subsystems:  
 804 muons, electrons, photons, jets and missing energy. This reconstruction  
 805 differs from the final reconstruction of the objects, for example a jet for the L1  
 806 consists on successive energy deposits in the ECAL and HCAL, while the off-line  
 807 reconstruction takes into account also the tracker information.

808 The L1 starts from regional data coming from the subsystems which is after-  
 809 ward combined in order to build ranked trigger objects in localized regions of the  
 810 detector. Global Muon and Calorimeter triggers sort the objects and send the best  
 811 ranked to the Global Trigger (GT). Before the GT no events are rejected, it is only  
 812 with the GT that the selection is applied. The GT combines the information and  
 813 can apply topological requirements and take a decision on keeping or disregarding  
 814 the event. On figure 2.21 can be found the work-flow of the L1.

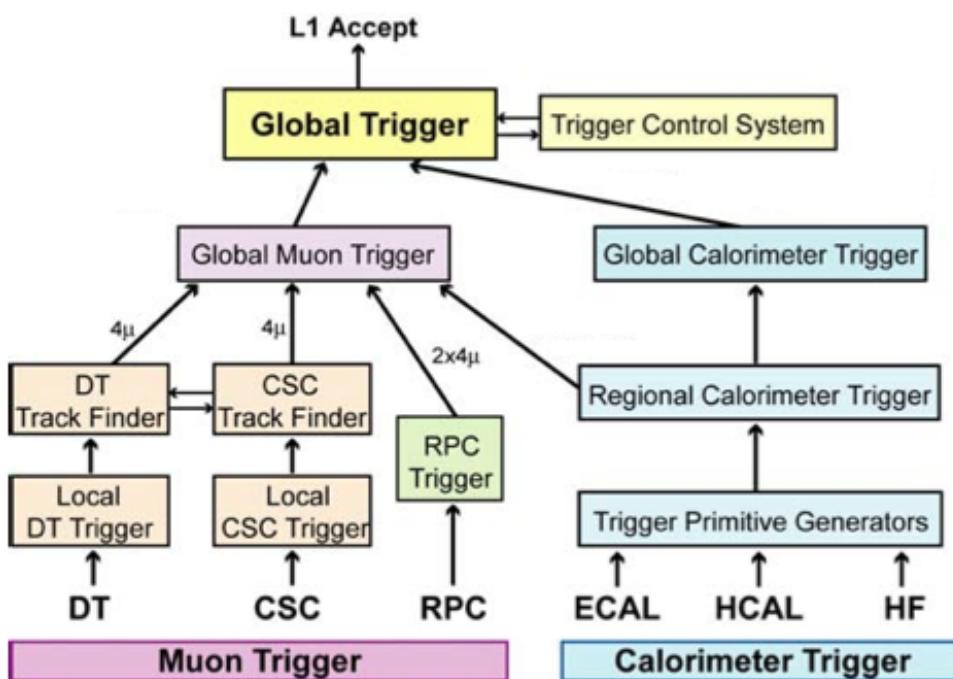


Figure 2.21: L1 architecture. ©CERN

815 The L1 cards are distributed between the detector and an adjoin cavern at 90  
 816 m distance from the detector. The latency time that the L1 disposes between the

817 collision and the taking of the decision is about  $3.2 \mu\text{s}$ . Therefore, the front-end  
818 memory in the cards should be able to keep in memory up to 128 bunch crossings.

819 **High Level Trigger**

820 The HLT takes as input the events accepted by the L1 and processes them using  
821 farms of commercial processors. The HLT does additional operations on the se-  
822 lected events making it much slower than L1 processing. In particular, the HLT  
823 takes also into account the tracker information. Consequently, this system is able  
824 to take into consideration the whole information of the detector. However, the  
825 reconstruction of objects done by the HLT differs slightly from the final off-line  
826 reconstruction. The decision taking process takes around 40 ms,  $10^4$  times more  
827 than for L1.

828 The events selected by HLT are finally stocked on disks under several paths  
829 depending on the selection performed. There is a constant development of HLT  
830 paths focusing on different analysis requirements in order to obtain the best pos-  
831 sible selection efficiency for specific signal types.

832 **2.3 Object reconstruction**

833 The different subsystems of CMS play specific roles in order to identify particles  
834 from collisions and measure their properties. In order to achieve such measure-  
835 ments, the information collected by the different parts of a subsystem should be  
836 combined to reconstruct particle information. Also for some particles the infor-  
837 mation between different subsystems should be connected to achieve a successful  
838 identification. The easiest example is the reconstruction of tracks in the pixel de-  
839 tector. In order to identify a track of a charged particle that traversed the pixel  
840 barrel the hits of such particle in each layers should be considered at the same  
841 time. A track in the pixel barrel is formed then by the three hits a particle leaves  
842 in each barrel layer.

843 **2.3.1 Track and vertex reconstruction**

844 Track and vertices reconstruction in CMS are done by the tracker system.

845 **Vertex reconstruction**

846 Description of vertex reconstruction procedure.

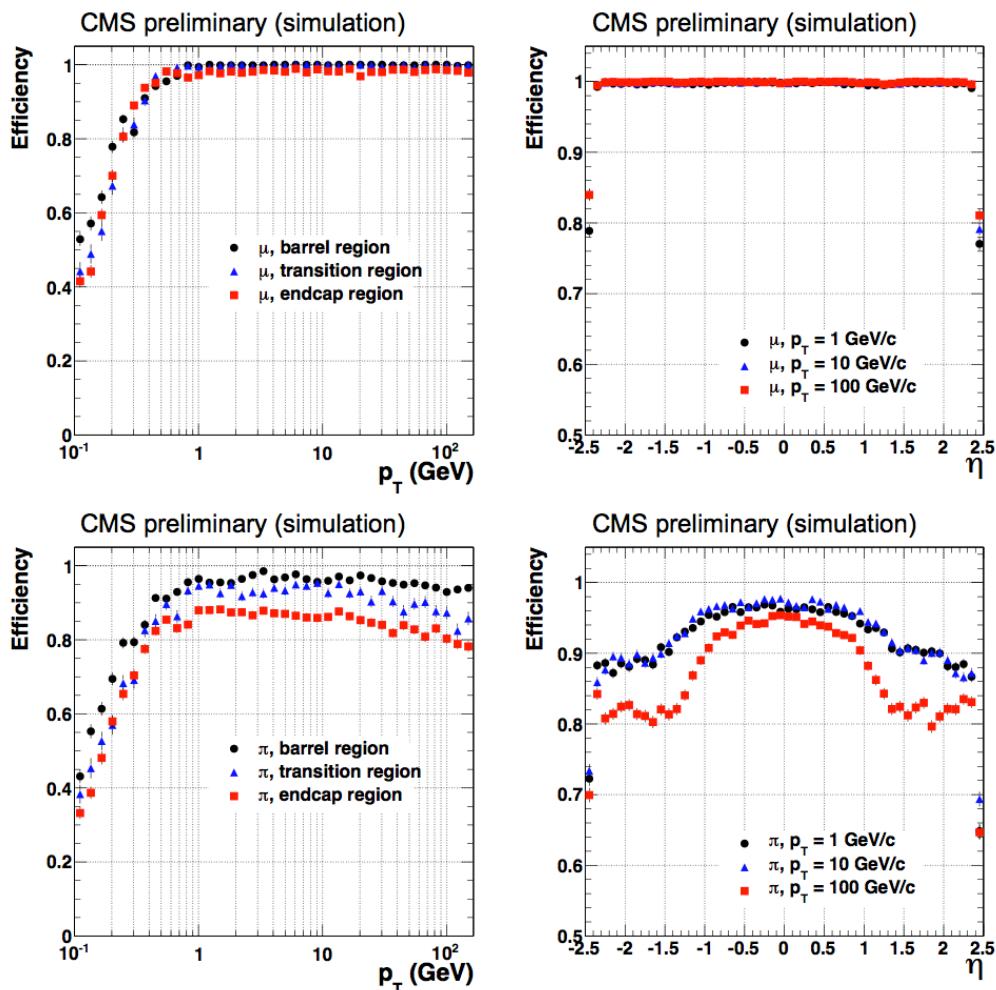


Figure 2.22:

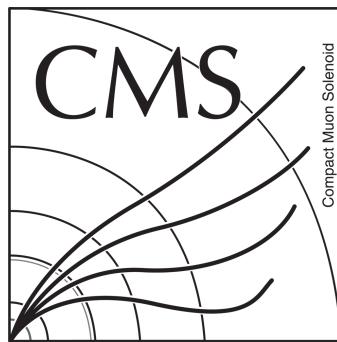


Figure 2.23: Resolution [left] and efficiency [right] of reconstruction of primary vertices.

### <sup>847</sup> 2.3.2 Particle Flow (PF) algorithm

<sup>848</sup> Description of Particle flow algorithm.

#### <sup>849</sup> Calorimeter clustering

<sup>850</sup> Description on how calorimeter deposits are clustered for particle reconstruction.

#### <sup>851</sup> Subdetectors link

<sup>852</sup> How different subdetectors information are related to reconstruct different particles.  
<sup>853</sup>

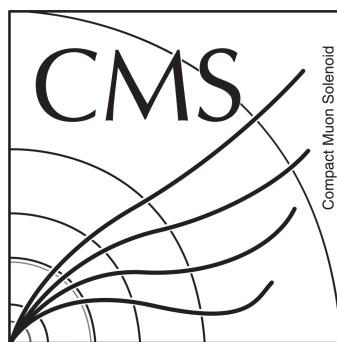


Figure 2.24: Connection between tracks reconstructed in the tracker (in green) and energy clusters in the ECAL [left] and HCAL clusters [right].

854 **2.3.3 Electron reconstruction**

855 Description of variables used to reconstruct electrons.

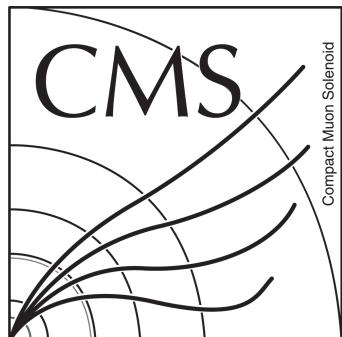


Figure 2.25: Reconstruction efficiency of electrons.

	xxxxxxx		xxxxxxx		xxxxxxx	
--	---------	--	---------	--	---------	--

Table 2.1: Identification requirements for electrons

856 **2.3.4 Muon reconstruction**

857 Muon reconstruction description and goodness of reconstruction from dimuon mass  
858 spectra showing resonances.

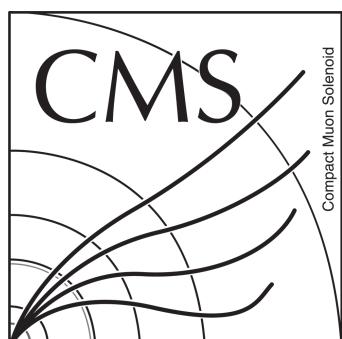


Figure 2.26: Reconstruction efficiency of muons.

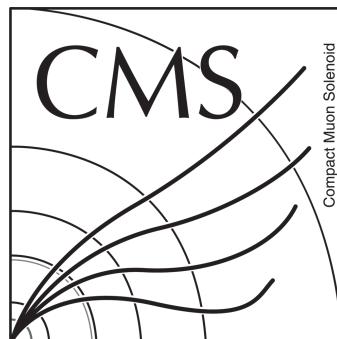


Figure 2.27: Dimuon mass spectra from reconstructed muons in CMS.

### **2.3.5 Jet reconstruction**

#### **Clustering algorithms**

Description of basis of clustering algorithms.

**SIScone** SIScone algorithm definition. Characteristics discussion.

**kT** KT algorithm definition. Characteristics discussion.

**Cambridge-Aachen** CA algorithm definition. Characteristics discussion.

**Anti-kT** AK algorithm definition. Characteristics discussion.

**Infrared and collinear safety** Definition of infrared and collinear safety. Description of which algorithms are safe and why.

**Jet area** Jet area definition. Importance of jet area.

#### **Jet energy corrections**

Why jet energy corrections. Brief how.

JES description. JER description.

#### **b-jets identification**

Description of basic parameters that help to differentiate b-jets from light jets.

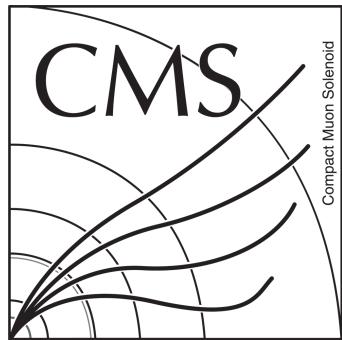


Figure 2.28: Jet areas for different jet algorithms related to the reconstructed  $p_T$ .



Figure 2.29: JEC uncertainties as function of  $p_T$  [left] and  $\eta$  [right].

- 874 **Identification algorithms and working points** Description of CSV algorithm.  
875 Variables used. Efficiency of different working points.

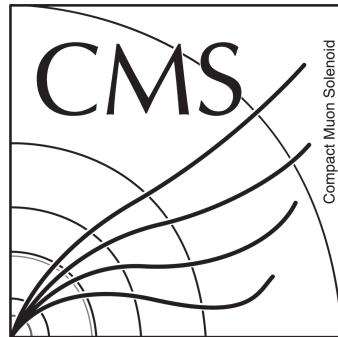


Figure 2.30: Number of secondary vertices [left] and CSV final discriminator variable [right] for each jet flavor.

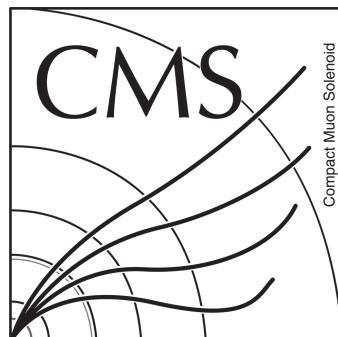


Figure 2.31: Efficiency of CSV b-tagging algorithm in the middle working point as function of jet  $p_T$  and  $\eta$ .

# <sup>876</sup> 3 Vector Like Quarks: Generic <sup>877</sup> model

<sup>878</sup> From chapter 1 we have seen how there are some parts in the SM that does  
<sup>879</sup> not work very well. From such internal issues some further models/theories have  
<sup>880</sup> been developed. All this theories are commonly grouped under the term Beyond  
<sup>881</sup> Standard Model or simply BSM. One of the most famous BSM theory is super-  
<sup>882</sup> symmetry (SUSY). This theory postulates a symmetry that does not distinguish  
<sup>883</sup> between fermions and bosons. This idea have given birth to a plethora of model  
<sup>884</sup> realizations and physics predictions. So far, nothing of the new consequences of  
<sup>885</sup> this theory have been confirmed but the experiments have an enormous invest-  
<sup>886</sup> ment on their search. But not only SUSY have seen the day light, there is on the  
<sup>887</sup> market an astonishing amount of BSM theories addressing different issues of the  
<sup>888</sup> SM. Extra dimensions, fourth families, composite Higgs are a few of them.

<sup>889</sup> In this chapter we will describe a bunch of models that introduce additional  
<sup>890</sup> heavy quarks, heavier than the top, in order to solve the hierarchy problem, de-  
<sup>891</sup> scribed on section 1.5.

## <sup>892</sup> 3.1 Motivation

- <sup>893</sup> Why to introduce extra quarks.
- <sup>894</sup> Plausible solution of hierarchy problem.
- <sup>895</sup> Reference to models with extra quarks. (?)

## <sup>896</sup> 3.2 Generic Formulation

<sup>897</sup> Formalism: Generic Langranian. Description of mixings with SM quarks.

$$| \text{xxxxxx} | \text{xxxxxx} | \text{xxxxxx} |$$

Table 3.1: Possible VLQ representations and correponding  $SU(2)_L \times U(1)$  charges.

### **3.2.1 Production modes**

- 898 Description of pair and single production. Parallel to production modes of top.  
900 Comparison between cross sections for pair and single production.

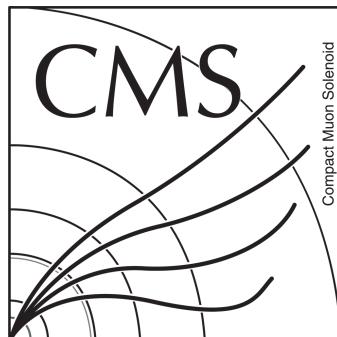


Figure 3.1: Feynman diagrams of  $T$  production in pairs [left] and single [right]

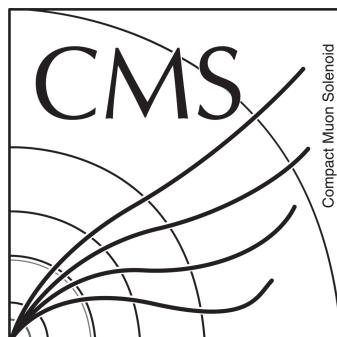


Figure 3.2:  $T$  production cross section for pair and single case as function of  $T$  mass for different center of mass energy in proton-proton collisions.

### **3.2.2 Decay modes**

- 902 Description of possible decay channels of  $T$ . Relative importance of decay channels depending on the mass.  
903

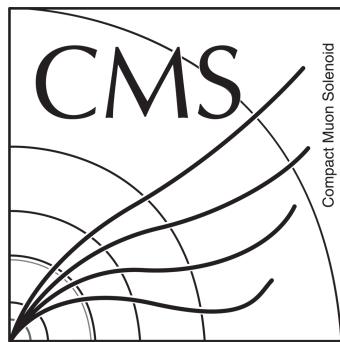


Figure 3.3:  $T$  branching ratios as function of its mass.

### 904 **3.3 Fesability study for a search of a $T$ at LHC at 8 905 TeV**

906 Discussion on selection of full hadronic channel.

#### 907 **3.3.1 Stragey for the full hadronic final state**

908 General discussion of strategy used for selection.

xxxxxxxx	xxxxxxxx	xxxxxxxx
----------	----------	----------

Table 3.2: Cross sections and expected number of events for background processes and signal.

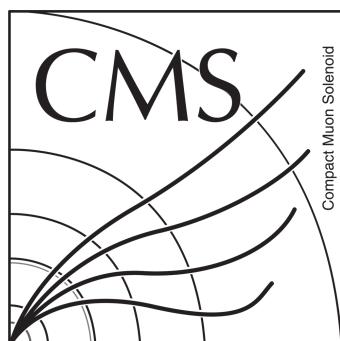


Figure 3.4: Forward jet produced in association with  $T$ .

<sub>909</sub> **3.3.2 Event selection**

<sub>910</sub> Description of selection. Paragraph per variable.

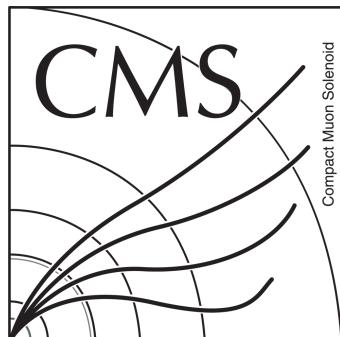


Figure 3.5:  $p_T$  of the six leading jets for backgrounds (stacked) and signal (over-imposed)

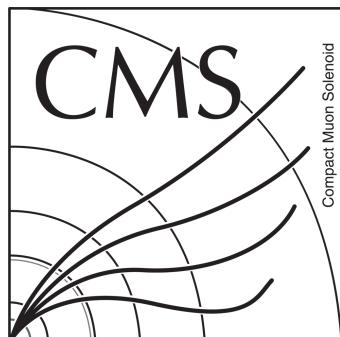


Figure 3.6: Total hadronic energy for backgrounds (stacked) and signal (over-imposed)

<sub>911</sub> **3.3.3 Results**

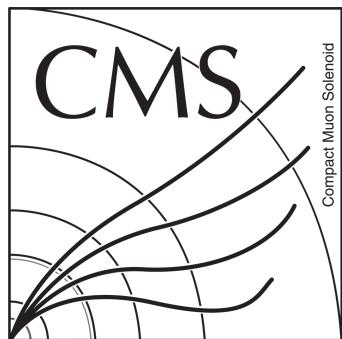


Figure 3.7:  $\Delta R$  between the reconstructed Higgs and W for backgrounds (stacked) and signal (over-imposed)

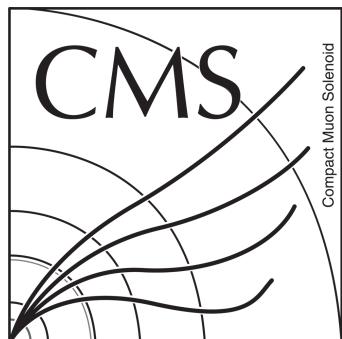


Figure 3.8: Mass of the reconstructed Higgs for backgrounds (stacked) and signal (over-imposed)

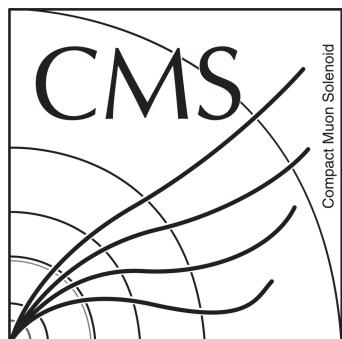


Figure 3.9: Relative total hadronic energy for backgrounds (stacked) and signal (over-imposed)

xxxxxxxx	xxxxxxxx	xxxxxxxx
----------	----------	----------

Table 3.3: Number of events for signal and backgrounds after the first selection cut (Cut 0), and efficiencies of each stage of the cutting procedure. The errors indicated are statistical only, based on the number of events.

	unweighted events after Cut 10	in mass window	weight	weighted events
Signal	8601	3780	0.03	$113 \pm 2$
$t\bar{t}$	409	57	7.7	$437 \pm 58$
$W+\text{jets}$	24	3	132	$395 \pm 228$
$QCD$	235	34	6.48	$220 \pm 38$
$tW$	18	3	11.3	$34 \pm 20$
$t+\text{jet}$	75	7	3.55	$25 \pm 9$
total background				$1112 \pm 352$

Table 3.4: Number of signal and background events from our simulation: in the first column the simulated events that pass all kinematic cuts, in the second column the events that fall in the mass window  $710 < M_{jjjj} < 750$  GeV, finally in the fourth column the number of weighted events in the mass window normalized to the physical cross section (the applied weight is listed in the third column). All the errors are statistical only. For the total background, we conservatively consider linear sum of errors.

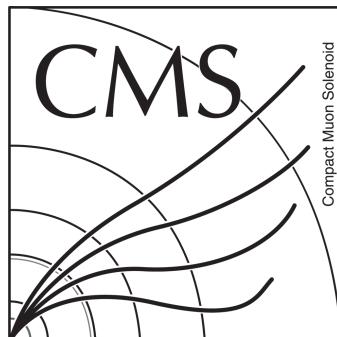


Figure 3.10: Reconstructed T mass after all cuts for backgrounds and signal (stacked)

# **912 4 Understanding theory predictions 913 via Monte-Carlo event generation**

**914** Although we have nowadays a very elegant and complete theoretical description  
**915** of particle physics, is not always evident how to translate this theory in actual  
**916** predictions, to compare with measurements. Moreover, on the case of hadronic  
**917** colliders, as the LHC, it's even more difficult due to the particularities of strong  
**918** interaction. On this subject, a set of tools and approaches have been developed in  
**919** order to be able to make accurate predictions from theory that could be directly  
**920** researched for on the experiments, as CMS or ATLAS for example. In the present  
**921** chapter, we describe such tools and formalisms and a set of studies comparing the  
**922** predictions these tools to data.

## **923 4.1 Mote-Carlo simulations**

**924** The Monte-Carlo simulations use random numbers and large samplings to calcu-  
**925** late mathematical quantities in complex configurations, as integrals or probabili-  
**926** ties. The typical example is on how to calculate the integral of a one-dimensional  
**927** function. One can throw several random coordinates pair in the Cartesian plane  
**928** and count how many of them are under the function. Then the integral of the  
**929** function will be proportional to the fraction of points under the curve to the total  
**930** thrown points. Larger the number of points, closer the estimation to the real value.  
**931** An illustration of the procedure can be seen in figure 4.1.

**932** A similar method is used to simulate proton-proton collisions. This simulation is  
**933** used to generate “random” events and to calculate quantities, as the cross section,  
**934** for a given physical process. Each event represent the final state of a collision, i.e.  
**935** the set of particles produced from the collision and seen by a detector. Such simu-  
**936** lations comprehend different stages: first, the partonic processes making reference  
**937** to the interaction between the partons inside the proton; second, the hadroniza-  
**938** tion of the particles produce from parton interactions; and third, the simulation of  
**939** the interaction between the hadrons (from second step) and the detector material.  
**940** Such events are used to evaluate predictions from theory in the frame of a specific  
**941** experiment. Whereas the hadronization and detector simulation are well-known  
**942** physical processes, new theories predictions rely basically on the partonic level,

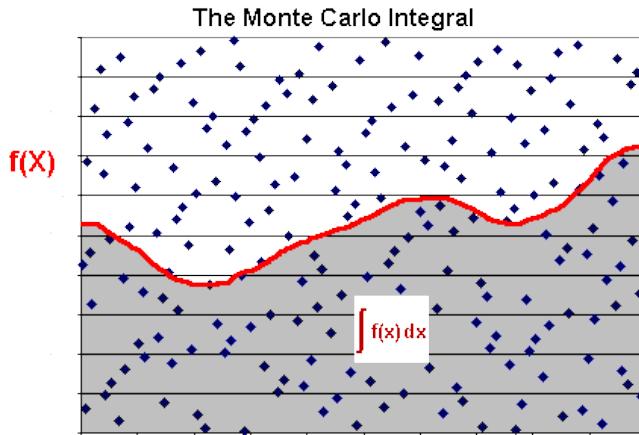


Figure 4.1: Integration using Monte-Carlo methods

943 where the fundamental interaction processes take part.

#### 944 4.1.1 Parton simulation

945 The parton model was initially proposed by Richard Feynman in 1969, as a method  
 946 to understand collisions of non-fundamental particles. The model consider a non-  
 947 fundamental particle, as a proton or a neutron, composed of a given number of  
 948 point-like fundamental particles. When a collision occur the point-like particles  
 949 inside have a major probability to scatter. For example, when an electron is fired  
 950 against a proton the most of the interactions will between the electron and the  
 951 fundamental components of the proton,  $u$  and  $d$  quarks. This “hard” components  
 952 are called *valence* quarks. Surrounding them there are the *sea* quarks and gluons.

953 However, as the energy of the collision increases the probability to scatter a sea  
 954 component, quark or gluon, increases. In addition, even if the valence quarks of a  
 955 proton are the  $u$  and  $d$  quarks, heavier quarks can appear in the sea, as the  $b$ ,  $c$  or  $s$   
 956 quarks. The probability to interact with a component, valence or sea, is described  
 957 by parton distribution function, commonly called PDF. A PDF  $f \equiv f(x, Q^2)$  rep-  
 958 resent the number density of a given quark or gluon as a function of the energy  
 959 scale  $Q^2$  and the fraction of momentum carried by the parton  $x$ . The determina-  
 960 tion of a PDF is done via a fit of large data samples from experiments specifically  
 961 designed to test the inner structure of nucleons. The DIS (Deep Inelastic Scat-  
 962 tering) experiment at SLAC (Stanford Linear Accelerator Center), in California,  
 963 United States, first probed the existence of partonic structure inside nucleons us-  
 964 ing leptons as probes scattered against nucleons. Another important experiment  
 965 was the HERA accelerator at DESY in Hamburg, Germany, which used electrons

966 to study the inner structure of protons.

967 In figure 4.2 is shown the Martin-Stirling-Thorne-Watt [18] (MSTW) PDF for  
 968 two energy scales. The MSTW PDF is one of the experimental fits combining data  
 969 from DIS and HERA. In this PDF can be seen that  $u$  and  $d$  quarks carry the most  
 970 of the momentum of the proton. The rest of the momentum is spread mainly over  
 971 a huge amount of gluons and some, less probable, sea quarks as  $\bar{u}, \bar{d}$  or  $c$  and  $s$ .  
 972 One important feature is that the composition of the proton changes depending  
 973 on the energy scale. At  $Q^2 = 10 \text{ GeV}^2$  there is no  $b$ -quark in the proton while at  
 974  $Q^2 = 10^4 \text{ GeV}^2$  there is a non-negligible probability to find it in the proton.

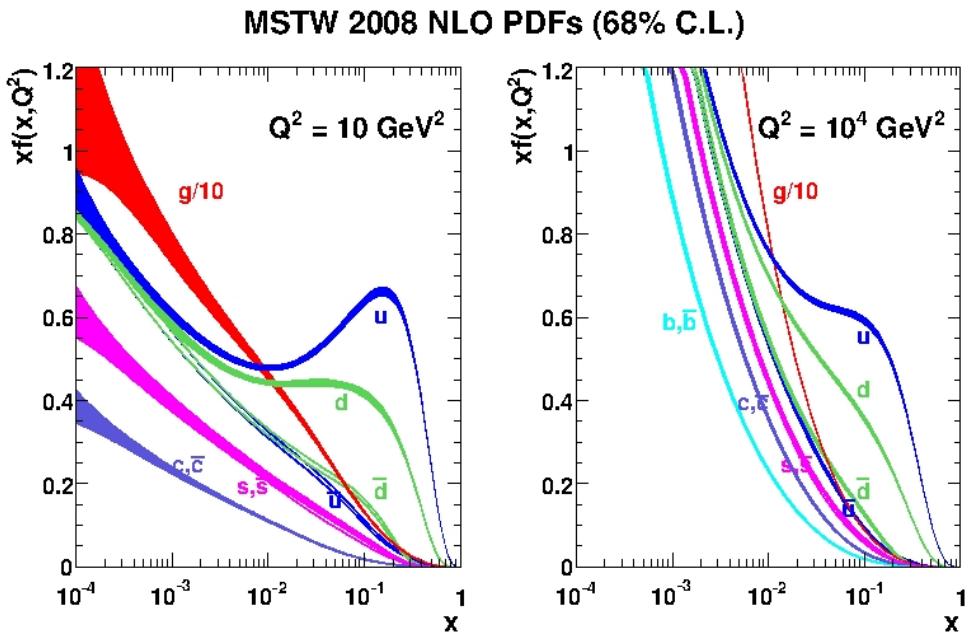


Figure 4.2: Martin-Stirling-Thorne-Watt proton PDF for  $Q^2 = 10 \text{ GeV}^2$  [left] and  $Q^2 = 10^4 \text{ GeV}^2$  [right]. From [18]

975 Two other important PDF fits are CTEQ [19] and NNPDF [20]. Together with  
 976 MSTW, they are the most used PDF sets in the CMS experiment for MC produc-  
 977 tion.

978 For the hard process, the differential cross section can be written as,

$$\begin{aligned}
d\sigma_{ij \rightarrow lm} &= \left( \int_0^1 \int_0^1 f_i(x_i, Q^2) f_j(x_j, Q^2) dx_i dx_j \right) \\
&\times \frac{d^3 p_l}{(2\pi)^2 2E_l} \frac{d^3 p_m}{(2\pi)^2 2E_m} \delta^4(p_i + p_j - p_l - p_m) \\
&\times |\mathcal{M}_{ij \rightarrow lm}|^2
\end{aligned} \tag{4.1}$$

where  $f_{i,j}$  correspond to the PDF's of the initial partons.  $\mathcal{M}_{ij \rightarrow lm}$  is the matrix element of the process which is the part of the S-matrix that contains the amplitude of the process, and modules the transition from the initial to the final state [21]. The matrix element could account effectively for all processes mediating the transition from the initial to the given final state, but in practice it is calculated only including a given number of processes. The calculation can achieve different levels, usually tree level or Leading Order (LO), but modern calculation could arrive, depending on the process, to one loop or Next-to-Leading-Order (NLO) or even two loops the Next-to-Next-to-Leading-Order (NNLO). This limit depends exclusively on the feasibility of the theoretical calculations. In figure 4.3 is shown an example of a leading order plus its corresponding NLO diagrams for a fermion scattering.

## 4.1.2 Hadron simulation

Description of how partons can't be seen free but inside hadrons.

## 4.1.3 Detector simulation

Description on how the detector is simulated, and how the detector response is simulated to be the closest possible to data with Geant4.

# 4.2 Tools

## 4.2.1 Matrix-element generators

Description of calculations performed by the MadGraph.

## 4.2.2 Hadron generators

Hadronization procedures performed by Pythia and Herwig.

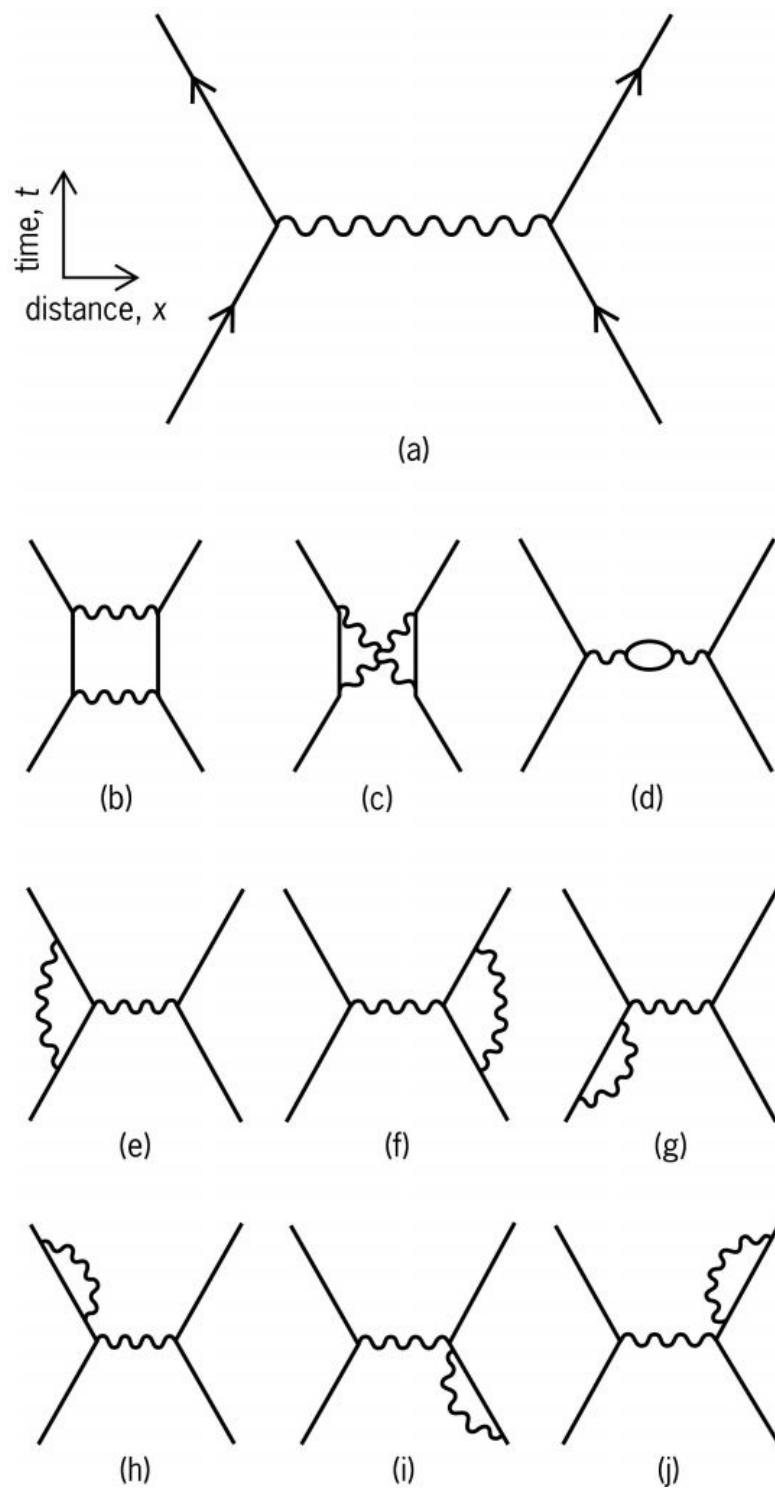


Figure 4.3: LO (a) and NLO (b)-(j) processes contributing to fermions scattering

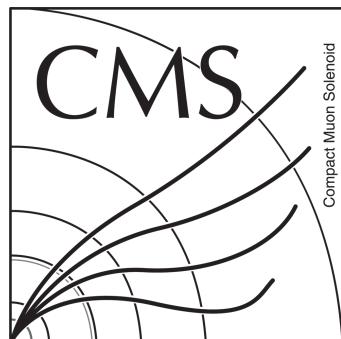


Figure 4.4: Graphical representation of hadronization process of partons resulting from a proton-proton collision.

### 1000 **4.3 Validation on data**

1001 Description of validation process of known processes against different MC predictions.  
1002



Figure 4.5:  $W + \text{jets}$  simulated by several MC generators compared to data

1003 Observed differences on the measured top pt and MC predictions.

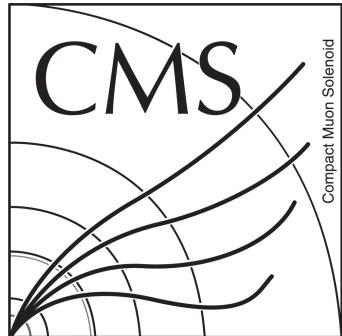


Figure 4.6:  $Z+jets$  simulated by several MC generators compared to data

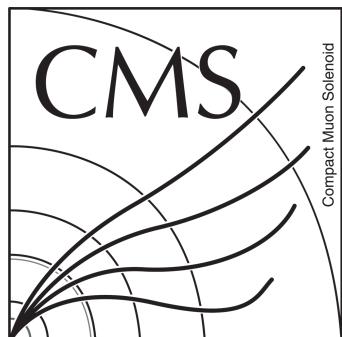


Figure 4.7: Data MC comparison of ratio of normalized diff cross section as function of the  $p_T$  of the top.



1004 **5 Search for a single produced T'**  
1005 **decaying into top and Higgs in**  
1006 **the full hadronic final state**

1007 In the present chapter we describe in full detail the search performed using 2012  
1008 data collected by CMS for a T' in the full hadronic final state. The theoretical  
1009 formalism for such object has been described on chapter 3.

1010 **5.1 Analysis Strategy**

1011 The strategy of the analysis is based on keeping large signal efficiency while keeping  
1012 under control the background. Main background of all hadronic final states is  
1013 multijet production. This background should not present any resonance in the  
1014 5-jets invariant mass variable but purely a continuum. In order to keep high  
1015 signal efficiency while constraining background, strategy to optimize the selection  
1016 is based on high signal efficiency criteria (around 80-90%) and on multiplication  
1017 of them.

1018 **5.2 Datasets**

1019 The analysis is based on the MultiJet primary dataset processed with the 22Jan2013  
1020 reconstruction.

Dataset name	Int. Luminosity ( $\text{pb}^{-1}$ )
/MultiJet/Run2012A-22Jan2013-v1/AOD	889.4
/MultiJet1Parked/Run2012B-05Nov2012-v2/AOD	4427.9
/MultiJet1Parked/Run2012C-part1-05Nov2012-v2/AOD	494.6
/MultiJet1Parked/Run2012C-part2-05Nov2012-v2/AOD	6654.9
/MultiJet1Parked/Run2012D-part1-10Dec2012-v1/AOD	5764.9
/MultiJet1Parked/Run2012D-part2-17Jan2013-v1/AOD	924.2
/MultiJet1Parked/Run2012D-part2-PixelRecover-17Jan2013-v1	538.4

Table 5.1: List of Multijet Primary Dataset used in the analysis and the corresponding integrated luminosity calculating using the golden JSON

Samples	Cross-Section (pb)	Number of events
QCD_Pt-120to170_TuneZ2star_8TeV_pythia6	$16 \times 10^4$	5.9M
QCD_Pt-170to300_TuneZ2star_8TeV_pythia6	$34 \times 10^3$	5.8M
QCD_Pt-300to470_TuneZ2star_8TeV_pythia6	$18 \times 10^2$	5.9M
QCD_Pt-470to600_TuneZ2star_8TeV_pythia6	110	3.9M
QCD_Pt-600to800_TuneZ2star_8TeV_pythia6	27	3.9M
QCD_Pt-800to1000_TuneZ2star_8TeV_pythia6	3.5	3.9M
QCD_HT-500To1000_TuneZ2star_8TeV-madgraph-pythia6	$84 \times 10^2$	30M
QCD_HT-1000ToInf_TuneZ2star_8TeV-madgraph-pythia6	$2 \times 10^2$	14M
DYToCC_M_50_TuneZ2star_8TeV_pythia6	$31 \times 10^2$	2M
DYToBB_M_50_TuneZ2star_8TeV_pythia6	$38 \times 10^2$	2M
TTJets_MSDecays_central_TuneZ2star_8TeV-madgraph-tauola	245.8 [NNLO]	60M
TT_CT10_TuneZ2star_8TeV-powheg-tauola	245.8 [NNLO]	22M
T_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola	11.1 [NNLO]	497k
T_s-channel_TuneZ2star_8TeV-powheg-tauola	1.76 [NNLO]	260k
T_t-channel_TuneZ2star_8TeV-powheg-tauola	30.7 [NNLO]	3.7M
Tbar_tW-channel-DR_TuneZ2star_8TeV-powheg-tauola	11.1 [NNLO]	492k
Tbar_s-channel_TuneZ2star_8TeV-powheg-tauola	3.79 [NNLO]	140k
Tbar_t-channel_TuneZ2star_8TeV-powheg-tauola	56.4 [NNLO]	1.9M
WZ_TuneZ2star_8TeV_pythia6_tauola	33.6 [NLO]	10M
ZZ_TuneZ2star_8TeV_pythia6_tauola	7.6 [NLO]	9.8M
WW_TuneZ2star_8TeV_pythia6_tauola	56 [NLO]	10M
TTH_Inclusive_M-125_8TeV_pythia6	0.13 [NLO]	100K

Table 5.2: List of Monte-Carlo background samples used in the analysis, their corresponding cross-section and their number of events.

Sample	T' Mass (GeV/c <sup>2</sup> )	Cross-Section (fb)
TprimeJetToTH_M-600_TuneZ2star_8TeV-madgraph_tauola	600	215.4
TprimeJetToTH_M-650_TuneZ2star_8TeV-madgraph_tauola	650	177.8
TprimeJetToTH_M-700_TuneZ2star_8TeV-madgraph_tauola	700	143.7
TprimeJetToTH_M-750_TuneZ2star_8TeV-madgraph_tauola	750	118.6
TprimeJetToTH_M-800_TuneZ2star_8TeV-madgraph_tauola	800	100
TprimeJetToTH_M-850_TuneZ2star_8TeV-madgraph_tauola	850	84.3
TprimeJetToTH_M-900_TuneZ2star_8TeV-madgraph_tauola	900	72.6
TprimeJetToTH_M-950_TuneZ2star_8TeV-madgraph_tauola	950	62.6
TprimeJetToTH_M-1000_TuneZ2star_8TeV-madgraph_tauola	1000	53.9

Table 5.3: List of Monte-Carlo background signal used in the analysis, their corresponding cross-section and mass of the T'.

1021 **5.3 Event selection**

1022 This section describes first the event processing details and then the event selection  
1023 applied. At the end we present the efficiencies measurement performed and/or  
1024 scale factors applied.

1025 **5.3.1 Event processing**

1026 Description of processing details of events.

1027 **5.3.2 Basic selection**

1028 Description of basic selection and justification of basic selection cuts.



Figure 5.1: Distribution of number of vertices reconstructed in the events before selection (except patuple creation) for data and for the sum of Monte-Carlo samples on which a weight have been applied. At this stage,  $H_T$  selection is not yet applied and as the QCD samples does not include events with  $H_T$  lower than 500 GeV/c

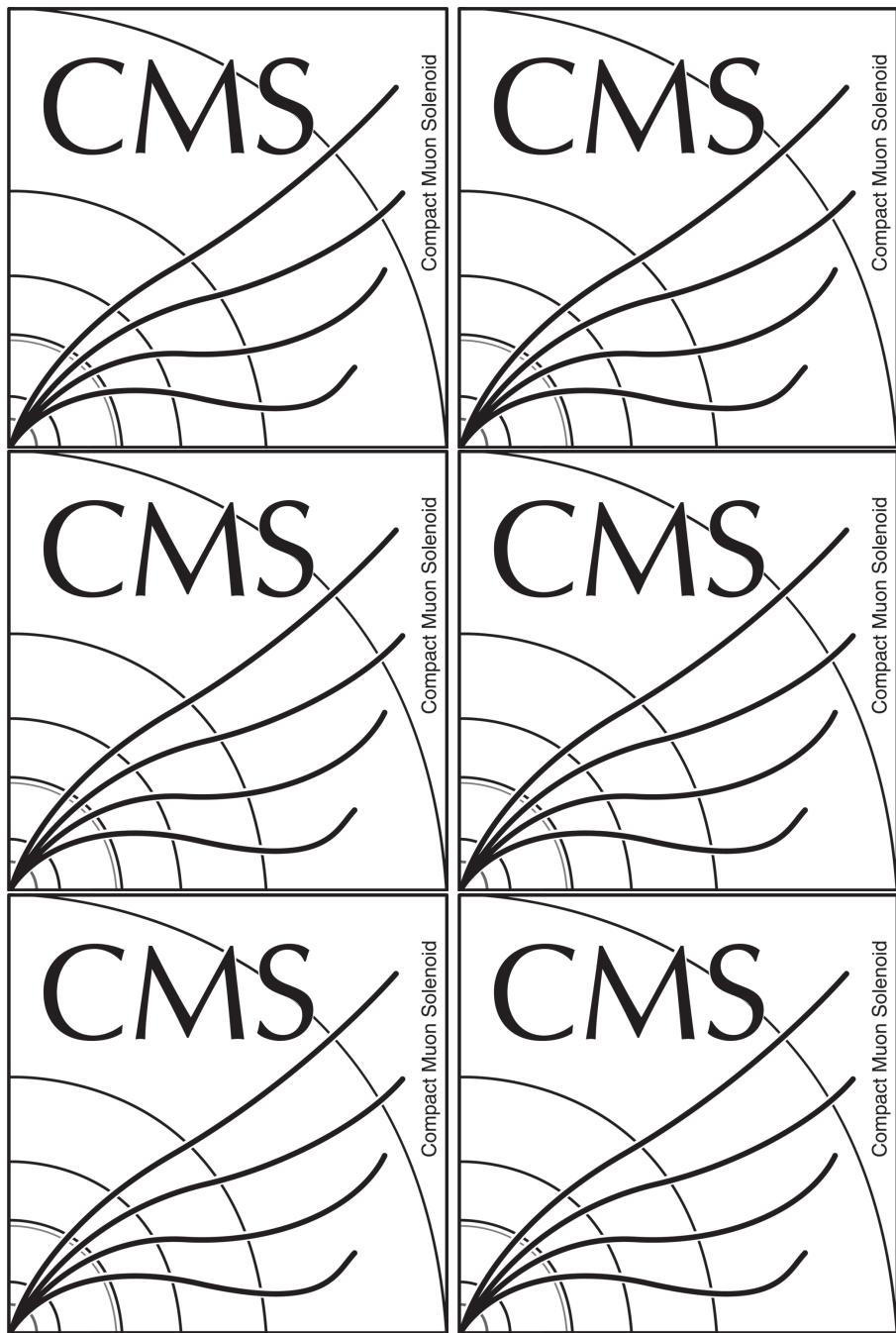


Figure 5.2: Distribution of transverse momentum of the 6 leading jets in the events. Data is compared to the sum of the MC samples normalized to luminosity after basic selection. The gray band represents the statistical uncertainties from the sum of the MC background. Reasonable agreement is observed, multijet process is the dominant process at this stage.

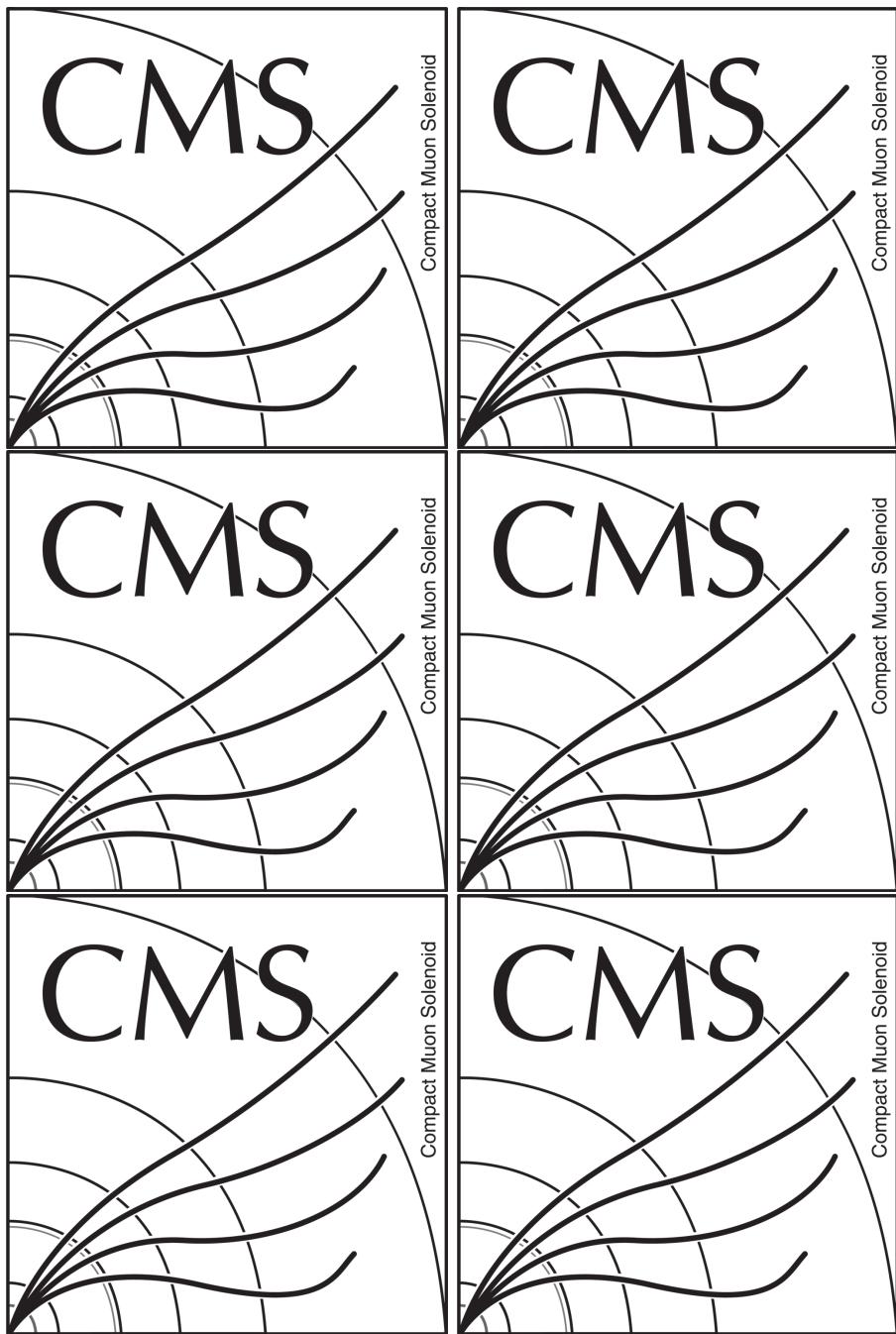


Figure 5.3: Distribution of  $\eta$  of the 6 leading jets in the events. Data is compared to the sum of the MC samples normalized to luminosity after basic selection. The gray band represents the statistical uncertainties from the sum of the MC background. Reasonable agreement is observed, multijet process is the dominant process at this stage.

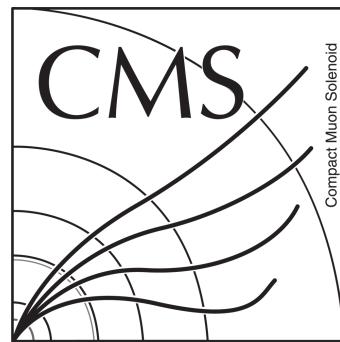


Figure 5.4: Distribution of  $H_T$  variable for data and the sum of the Monte Carlo samples normalized to luminosity after basic selection. The signal sample is simply overlay on top of the stack of the MC samples. It has a  $T$  mass of  $700 \text{ GeV}/c^2$ . The gray band represents the statistical uncertainties from the sum of the MC background. Reasonable agreement is observed, multijet process is the dominant process at this stage.

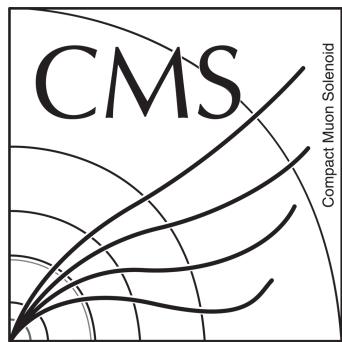


Figure 5.5: B-tag CSVM jet multiplicity for data and Monte Carlo samples after all previous criteria [left] and after requiring at least 3 b-tagged jets [right]. The signal has 3 b-tag while most of the background have mainly 2 b-tag. The disagreement observed can be understood as no scale factors are applied on the MC in left plots. In right plot, scale factor of 3 b-tag jets is applied. This criteria is the last criteria definition the basic selection for the analysis. The sum of MC is normalized to the integrated luminosity in the data.

<sub>1029</sub> **5.3.3 T' reconstruction with a  $\chi^2$  sorting algorithm**

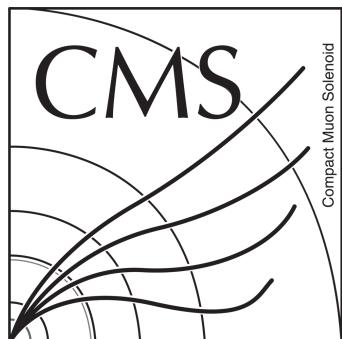


Figure 5.6: Reconstruction efficiency taking into account only events where jets could be matched to partons [left] and to total number of events [right]

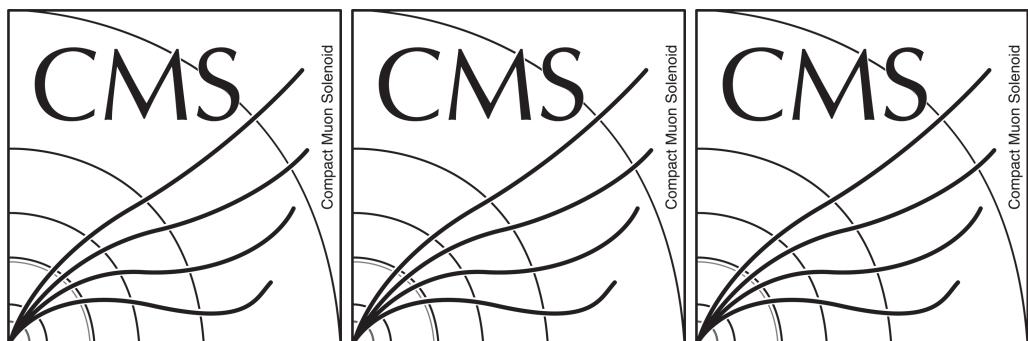


Figure 5.7: Reconstructed top, W and H for the T mass point of  $700 \text{ GeV}/c$

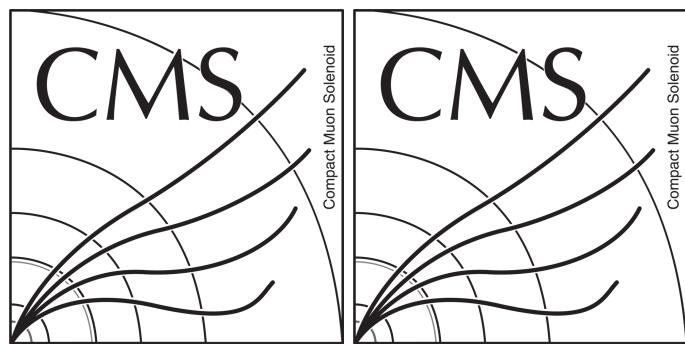


Figure 5.8: Reconstructed T mass for all mass points.

Generated		Reconstructed		
Mass ( $\text{GeV}/c^2$ )	Width ( $\text{GeV}/c^2$ )	Mass ( $\text{GeV}/c^2$ )	Width ( $\text{GeV}/c^2$ )	$\chi^2/\text{ndf}$
600	0.62	$616.68 \pm 14.17$	$30.73 \pm 10.52$	15.14/17
650	0.80	$649.48 \pm 13.55$	$37.32 \pm 10.67$	37.82/17
700	1.02	$697.54 \pm 14.25$	$42.98 \pm 12.28$	4.39/17
750	1.27	$740.23 \pm 15.53$	$44.62 \pm 13.68$	14.48/17
800	1.56	$787.13 \pm 16.20$	$49.95 \pm 12.18$	58.67/27
850	1.89	$835.79 \pm 16.30$	$48.94 \pm 12.18$	36.35/27
900	2.26	$882.308 \pm 18.34$	$49.91 \pm 13.84$	34.97/27
950	2.67	$933.30 \pm 23.30$	$54.78 \pm 18.24$	16.00/27
1000	3.13	$976.01 \pm 29.76$	$58.27 \pm 24.04$	21.17/27

Table 5.4: Reconstructed mass and width for T candidate after full analysis selection from a gaussian fit for each signal mass generated.

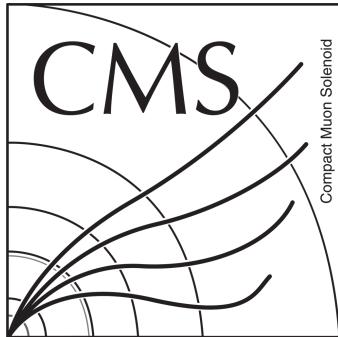


Figure 5.9: Normalized distribution of  $\chi^2$  variable for signal compared to multijet ( $H_T$  [500, 1000] only) and  $t\bar{t}$ . The signal sample used has a T' mass of 700  $\text{GeV}/c^2$ . Signal events are peaking at 0 with a falling distribution while background have larger tails.

1030 **5.3.4 Selection based on reconstructed objects**

1031 Description of the selection in general and detailed description of each variable  
1032 used.



Figure 5.10:  $\Delta R$  of the 2 b-tag jets used to reconstruct the Higgs candidate after basic selection. The signal which is simply overlaid prefers low  $\Delta R$  while backgrounds have larger distribution at higher value. The gray band represents the statistical uncertainties from the sum of the MC background.

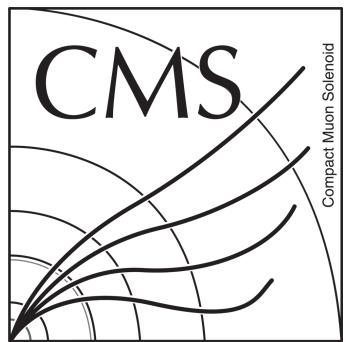


Figure 5.11: Distribution for  $\Delta R(W_{cand}H_{cand})$  for data and the sum of Monte Carlo samples. All others criteria are applied up to this one.

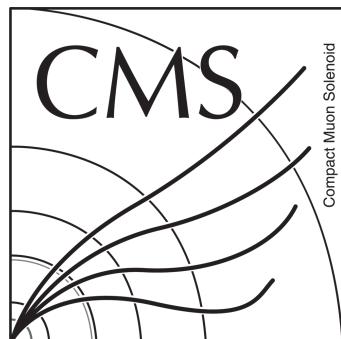


Figure 5.12: Distributions for  $\Delta R(T'j^6)$  for data and the sum of Monte Carlo samples. All others criteria are applied up to this one. The low statistics in the multijet (QCD) MC sample is visible at this stage. The gray band represents the statistical uncertainties from the sum of the MC background and it is dominated by QCD samples.

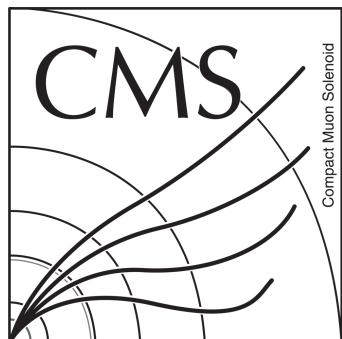


Figure 5.13: Distribution of Relative  $H_T$  for data and the sum of the Monte Carlo samples. Signal can be separated from background in this variable. All others criteria are applied up to this one. The low statistics in the multijet (QCD) MC sample is visible at this stage. The gray band represents the statistical uncertainties from the sum of the MC background and it is dominated by QCD samples.

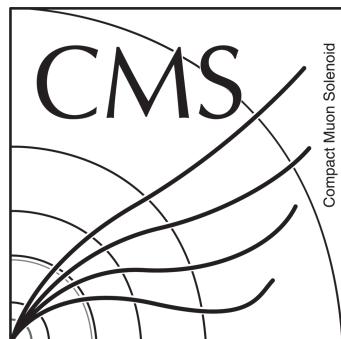


Figure 5.14: Distribution of Top-Higgs mass asymmetry for data and the sum of the Monte Carlo samples. Signal can be separated from background in this variable. All others criteria are applied before this one. The low statistics in the multijet (QCD) MC sample is visible at this stage. The gray band represents the statistical uncertainties from the sum of the MC background and it is dominated by QCD samples.



Figure 5.15: Distribution of  $(M(\text{top}^{2nd}) + M(W^2))/M(H)$  for data and the sum of the Monte Carlo samples. Signal can be separated from background in this variable. All other criteria are applied before this one. The low statistics in the multijet (QCD) MC sample is visible at this stage. The gray band represents the statistical uncertainties from the sum of the MC background and it is dominated by QCD samples.

	xxxxxxx		xxxxxxx		xxxxxxx	
--	---------	--	---------	--	---------	--

Table 5.5:  $S/B$  and  $S/\sqrt{S+B}$  from MC samples for each step of the selection.

1033 **5.3.5 Efficiencies**

1034 **Trigger**

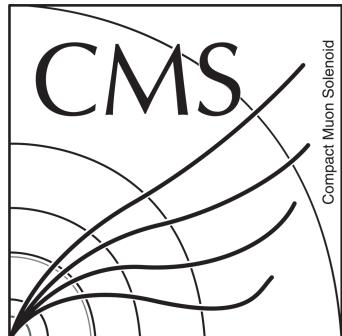


Figure 5.16: Efficiency in data and the MC signal samples for events passing trigger bit HLT\_Dijet80\_Dijet60\_Dijet20 with respect to trigger bit HLT\_HT400 after standard selection up to DEFINE SELECTION [included]. In the middle stage of the selection, we observe discrepancies between 10% and 6% at higher  $p_T$ . This efficiency is parametrized as function of the 6<sup>th</sup> jet  $p_T$ .

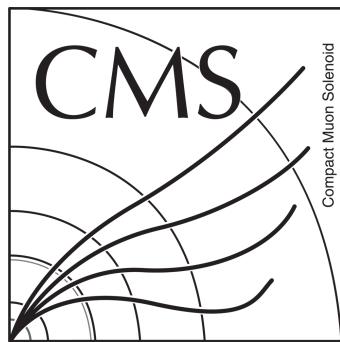


Figure 5.17: Efficiency in data and the MC signal samples for events passing trigger bit HLT\_Dijet80\_Dijet60\_Dijet20 with respect to trigger bit HLT\_HT400 after standard selection up to the Higgs-candidate mass selection (one cut before the final). This efficiency is parametrized as function of the 6<sup>th</sup> jet  $p_T$ . The dispersion observed is mainly 6%.

1035 **Selection**

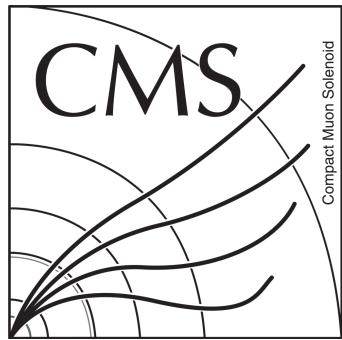


Figure 5.18: Selection efficiency for all signal mass points.

Selection	Cut	Signal (M=700 GeV/c <sup>2</sup> )	Multijet	t̄t + single top	Diboson
Basic	Trigger and $p_T, \eta$ selection	52.66	27.40	36.54	16.12
	$j^1 > 150$ GeV/c	47.65	24.05	24.50	11.55
	$H_T > 630$ GeV/c	44.29	20.57	20.57	8.98
	$n_b^{CSV} \geq 3$	13.74	0.16	1.73	0.11
Analysis	$\chi^2 < 800$	13.38	0.14	1.71	0.11
	$\Delta R(bb) < 1.2$	11.06	0.06	0.67	0.06
	$1.6 < \Delta R(W_{cand} H_{cand}) < 4.0$	10.44	0.05	0.56	0.05
	$\frac{M(\text{top}_{cand}^{2nd}) + M(W_{cand}^{2nd})}{M(H_{cand})} > 7.5$	5.82	0.03	0.17	0.02
	$\Delta R(T' j^6) > 4.5$	3.62	$3 \times 10^{-3}$	0.03	$3 \times 10^{-3}$
	$105 \text{ GeV}/c^2 < M(H_{cand}) < 145 \text{ GeV}/c^2$	2.97	$1 \times 10^{-3}$	0.01	$4 \times 10^{-4}$

Table 5.6: Cumulative efficiencies, in %, for signal and main background as a function of cuts.

<sub>1036</sub> **5.4 Background estimation from data**

<sub>1037</sub> Description of difficulties related to use MC for estimation of backgrounds.

<sub>1038</sub> **5.4.1 Known difficulties and tried methods**

<sub>1039</sub> Brief description of tried methods. One paragraph for matrix method one para-  
<sub>1040</sub> graph for tight-loose method.

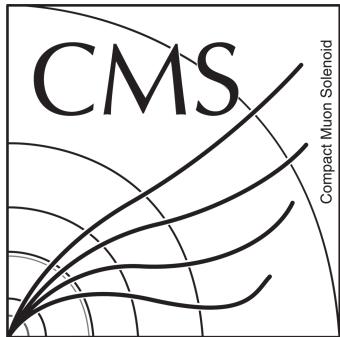


Figure 5.19:  $p_T(H)$  and  $p_T(\text{top})$  correlation for backgrounds [left] and signal [right].

1041 **5.4.2 Method**

1042 Description of control sample.

1043 **5.4.3 Validation**

1044 Description of validation procedure.

1045 Description on test to identify possible bias from chosen WP.

1046 Exclusive WP additional test.

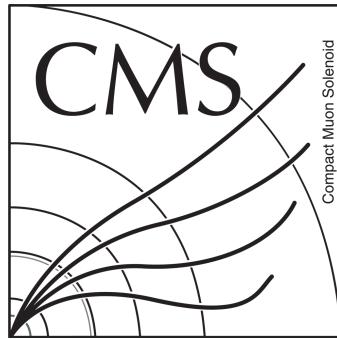


Figure 5.20: Left: Distribution of the agreement between control sample and analysis sample in data at early stage of the selection as a function of the  $\chi^2$  selection. The y-axis represents the chi2/ndf from a shape comparison made in data between control sample and analysis sample. A clear minimum is observed at 620. Right: Corresponding p-value of the  $\chi^2$  test, showing the compatibility with null hypothesis in the comparison between control and signal samples.

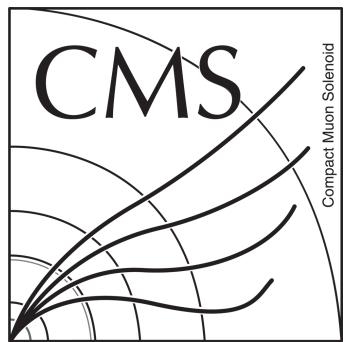


Figure 5.21: Distribution of the 5-jets invariant mass in the control sample for  $t\bar{t}$  for exclusive regions of  $\chi^2$ . Combinations with small values of  $\chi^2$  tend to populate  $M_{5J}$  between [600, 1000]

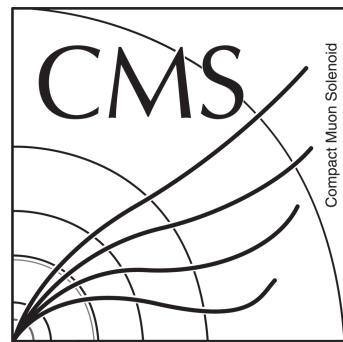


Figure 5.22: Number of Combinations entering the control sample for  $t\bar{t}$  and QCD HT 500-1000 for  $\chi^2 < 140$ . The majority of events entering the control sample only enter with one combination.

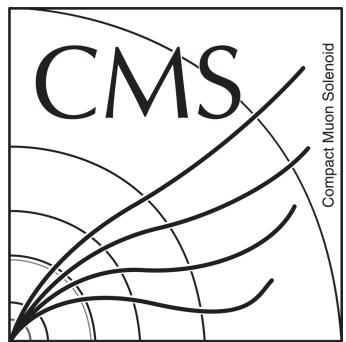


Figure 5.23: Signal contamination in the control region comparing 5-jets invariant mass between data and signal.

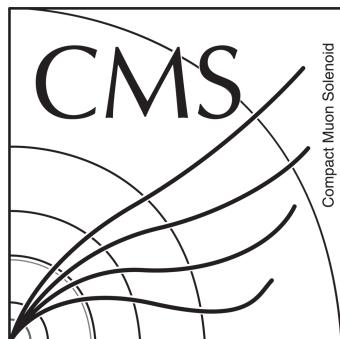


Figure 5.24: 5-jets invariant mass in the control sample for different b-tagging working point for  $t\bar{t}$  Monte Carlo samples within 4 stages of selection: A [top left], B [top right], C [bottom left] and D [bottom right]. The 3 working points are given in different color. Within statistical error, the 3 shapes are in agreement at all stages. A dip can be noticed around  $700 \text{ GeV}/c^2$ .  $t\bar{t}$  MC as all the other Monte-Carlo samples are purely used for illustration. The shape is not visible when the same exercise is made on the sum of MC background or in the data.

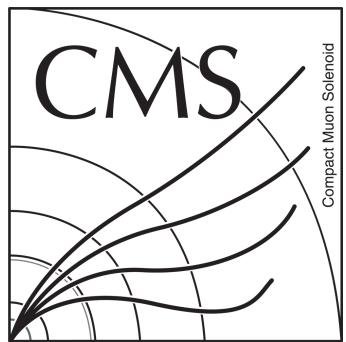


Figure 5.25: 5-jets invariant mass in the control sample for different b-tagging working point for QCD Monte Carlo samples within 4 stages of selection: A [top left], B [top right], C [bottom left] and D [bottom right]. The 3 working points are given in different color. Quickly a lack of statistics is visible.

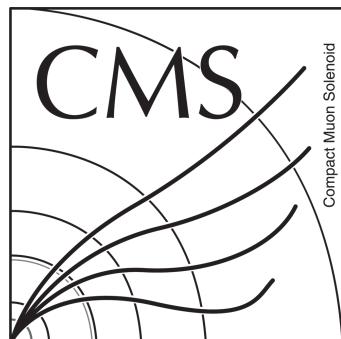


Figure 5.26: 5-jets invariant mass in the control sample for different b-tagging working point for the weighted sum of background Monte Carlo samples within 4 stages of selection: A [top left], B [top right], C [bottom left] and D [bottom right]. The 3 working points are given in different color. Within statistical error, the 3 shapes are in agreement at all stages.

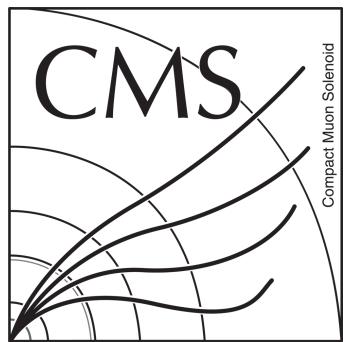


Figure 5.27: 5-jets invariant mass in the control sample for different b-tagging working point in data within 3 stages of selection: A [top left], B [top right] and C [bottom]. The 3 working points are given in different color. Within statistical error, the 3 shapes are in agreement at all stages.

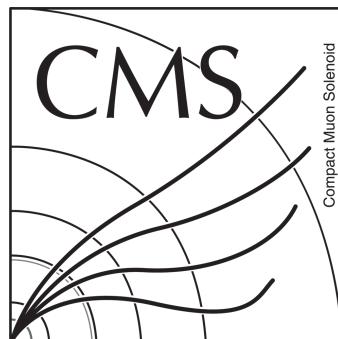


Figure 5.28: 5-jets invariant mass in the control sample for different b-tagging working point, in exclusive regions, for  $t\bar{t}$  Monte Carlo samples within 4 stages of selection: A [top left], B [top right], C [bottom left] and D [bottom right]. The 3 working points are given in different color. Within statistical error, the 3 shapes are in agreement at all stages. A dip can be noticed around  $700 \text{ GeV}/c^2$ .  $t\bar{t}$  MC as all the other Monte-Carlo samples are purely used for illustration. The shape is not visible when the same exercise is made on the sum of MC background or in the data.

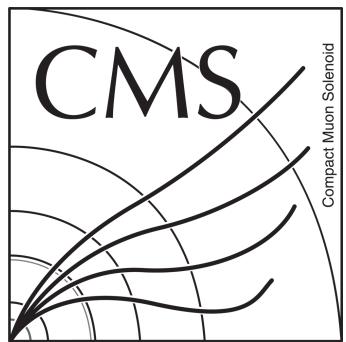


Figure 5.29: 5-jets invariant mass in the control sample for different b-tagging working point, in exclusive regions, for QCD Monte Carlo samples within 4 stages of selection: A [top left], B [top right], C [bottom left] and D [bottom right]. The 3 working points are given in different color. Quickly a lack of statistics is visible.

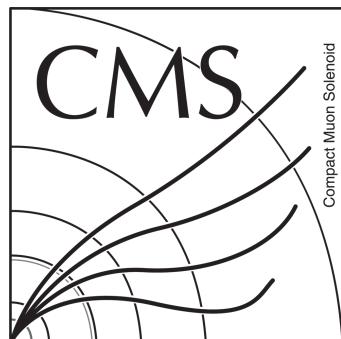


Figure 5.30: 5-jets invariant mass in the control sample for different b-tagging working point, in exclusive regions, for the weighted sum of background Monte Carlo samples within 4 stages of selection: A [top left], B [top right], C [bottom left] and D [bottom right]. The 3 working points are given in different color. Within statistical error, the 3 shapes are in agreement at all stages.

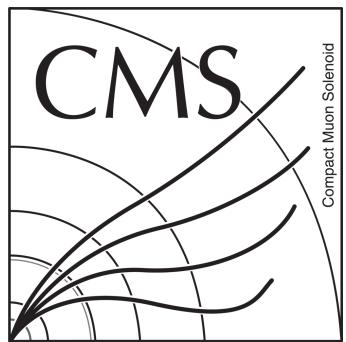


Figure 5.31: 5-jets invariant mass in the control sample for different b-tagging working point, in exclusive regions, in data within 3 stages of selection: A [top left], B [top right] and C [bottom]. The 3 working points are given in different color. Within statistical error, the 3 shapes are in agreement at all stages.

## <sup>1047</sup> 5.5 Systematics

<sup>1048</sup> Description by paragraphs of each systematic source for background estimation  
<sup>1049</sup> and signal.

Sample Name	b or c quark		Light flavours	
	up	down	up	down
$Tj \rightarrow tHj$ 600 $\text{GeV}/c^2$	7.70%	7.33%	0.84%	0.84%
$Tj \rightarrow tHj$ 650 $\text{GeV}/c^2$	7.94%	7.54%	0.65%	0.64%
$Tj \rightarrow tHj$ 700 $\text{GeV}/c^2$	7.75%	7.36%	0.56%	0.56%
$Tj \rightarrow tHj$ 750 $\text{GeV}/c^2$	7.82%	7.43%	0.57%	0.57%
$Tj \rightarrow tHj$ 800 $\text{GeV}/c^2$	7.89%	7.50%	0.65%	0.65%
$Tj \rightarrow tHj$ 850 $\text{GeV}/c^2$	8.14%	7.72%	0.67%	0.66%
$Tj \rightarrow tHj$ 900 $\text{GeV}/c^2$	8.34%	7.89%	0.68%	0.68%
$Tj \rightarrow tHj$ 950 $\text{GeV}/c^2$	8.58%	8.10%	0.65%	0.65%
$Tj \rightarrow tHj$ 1000 $\text{GeV}/c^2$	8.66%	8.17%	0.66%	0.66%

Table 5.7: B-tagging uncertainties for signal samples

Sample Name	JER		JES	
	up	down	up	down
$Tj \rightarrow tHj$ 600 $\text{GeV}/c^2$	0.28%	0.19%	11.48%	12.32%
$Tj \rightarrow tHj$ 650 $\text{GeV}/c^2$	0.55%	1.71%	10.01%	11.32%
$Tj \rightarrow tHj$ 700 $\text{GeV}/c^2$	0.63%	0.05%	6.32%	9.33%
$Tj \rightarrow tHj$ 750 $\text{GeV}/c^2$	0.29%	0.14%	5.37%	9.50%
$Tj \rightarrow tHj$ 800 $\text{GeV}/c^2$	0.0%	0.64%	5.24%	8.65%
$Tj \rightarrow tHj$ 850 $\text{GeV}/c^2$	0.29%	0.68%	5.99%	9.40%
$Tj \rightarrow tHj$ 900 $\text{GeV}/c^2$	0.42%	0.69%	6.04%	8.53%
$Tj \rightarrow tHj$ 950 $\text{GeV}/c^2$	0.95%	0.33%	5.52%	8.53%
$Tj \rightarrow tHj$ 1000 $\text{GeV}/c^2$	0.66%	0.20%	7.70%	7.24%

Table 5.8: JEC uncertainties for signal samples

Sample Name	CTEQ6.6		MSTW2008		NNPDF2.0
	up	down	up	down	up,down
$Tj \rightarrow tHj$ 600 $\text{GeV}/c^2$	2.11%	1.61%	2.78%	1.94%	1.97%
$Tj \rightarrow tHj$ 650 $\text{GeV}/c^2$	2.17%	1.61%	2.90%	2.00%	2.24%
$Tj \rightarrow tHj$ 700 $\text{GeV}/c^2$	2.19%	1.62%	2.90%	2.00%	2.24%
$Tj \rightarrow tHj$ 750 $\text{GeV}/c^2$	2.30%	1.68%	2.85%	1.97%	2.41%
$Tj \rightarrow tHj$ 800 $\text{GeV}/c^2$	2.35%	1.72%	2.94%	2.04%	2.21%
$Tj \rightarrow tHj$ 850 $\text{GeV}/c^2$	2.45%	1.73%	2.96%	2.07%	2.37%
$Tj \rightarrow tHj$ 900 $\text{GeV}/c^2$	2.62%	1.81%	3.04%	2.11%	2.70%
$Tj \rightarrow tHj$ 950 $\text{GeV}/c^2$	2.78%	1.88%	3.11%	2.18%	2.90%
$Tj \rightarrow tHj$ 1000 $\text{GeV}/c^2$	2.76%	1.88%	3.09%	2.15%	2.86%

Table 5.9: PDF+ $\alpha_s$  uncertainties for signal samples

Sample Name	b-tagging		JEC		PDF+ $\alpha_S$	
	up	down	up	down	up	down
$Tj \rightarrow tHj$ 600 $\text{GeV}/c^2$	7.52%	7.16%	18.2%	17.3%	3.60%	3.60%
$Tj \rightarrow tHj$ 650 $\text{GeV}/c^2$	7.84%	7.44%	10.9%	7.7%	3.11%	2.08%
$Tj \rightarrow tHj$ 700 $\text{GeV}/c^2$	7.73%	7.34%	8.7%	4.6%	3.20%	2.34%
$Tj \rightarrow tHj$ 750 $\text{GeV}/c^2$	7.68%	7.31%	9.8%	2.3%	3.15%	3.15%
$Tj \rightarrow tHj$ 800 $\text{GeV}/c^2$	7.80%	7.41%	8.2%	5.3%	3.07%	2.42%
$Tj \rightarrow tHj$ 850 $\text{GeV}/c^2$	8.12%	7.70%	6.8%	4.6%	3.06%	2.44%
$Tj \rightarrow tHj$ 900 $\text{GeV}/c^2$	8.24%	7.80%	9.4%	5.5%	3.18%	2.45%
$Tj \rightarrow tHj$ 950 $\text{GeV}/c^2$	8.60%	8.12%	10.0%	6.0%	3.22%	2.70%
$Tj \rightarrow tHj$ 1000 $\text{GeV}/c^2$	8.73%	8.23%	10.1%	9.1%	3.18%	2.87%

Table 5.10: Summary of uncertainties for signal samples

Systematics Name	Signal	Background
Theory	5%	—
PDF	+2.90% / -2.24%	—
Luminosity	2.6%	—
Trigger	10%	—
B-tag	+7.77% / -7.38%	—
JEC	+6.35% / -9.33%	—
Background Shape determination	—	12%
Background Normalization	—	20%

Table 5.11: Summary of uncertainties in case of signal mass point at  $700 \text{ GeV}/c^2$  and for background.

<sub>1050</sub> **5.6 Results**

<sub>1051</sub> Discussion of results and interpretation.

T' Mass $\text{GeV}/c^2$	Signal	Background	Observed Data
600	3.57	5.14	—
650	6.03	8.69	—
700	8.00	12.10	—
750	7.61	11.22	—
800	7.32	11.36	—
850	7.00	9.96	—
900	5.71	9.58	—
950	4.33	9.44	—
1000	3.36	9.44	—

Table 5.12: Expected number of events for the signal, estimated background and observed data after full selection

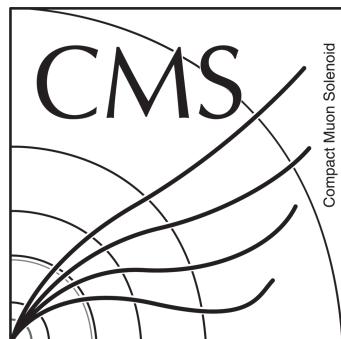


Figure 5.32:  $M_{jjjjj}$  after full selection, estimated background from data and signal mass point of  $700 \text{ GeV}/c^2$



Figure 5.33: Expected and observed limits



1052

# Bibliography

- 1053 [1] O. S. Brüning, P. Collier, P. Lebrun, S. Myers, R. Ostoja, J. Poole, and  
1054 P. Proudlock, *LHC Design Report*. CERN, Geneva, 2004.
- 1055 [2] **CMS Collaboration** Collaboration, G. L. Bayatian *et al.*, *CMS Physics:*  
1056 *Technical Design Report Volume 1: Detector Performance and Software*.  
1057 Technical Design Report CMS. CERN, Geneva, 2006. There is an error on  
1058 cover due to a technical problem for some items.
- 1059 [3] **CMS Collaboration** Collaboration, G. L. Bayatian *et al.*, “CMS Physics:  
1060 Technical Design Report Volume 2: Physics Performance,” *J. Phys. G* **34**  
1061 no. CERN-LHCC-2006-021. CMS-TDR-8-2, (2007) 995–1579. 669 p. revised  
1062 version submitted on 2006-09-22 17:44:47.
- 1063 [4] **ATLAS Collaboration** Collaboration, A. Collaboration, *ATLAS detector  
1064 and physics performance: Technical Design Report*. Technical Design Report  
1065 ATLAS. CERN, Geneva, 1999. Electronic version not available.
- 1066 [5] **LHCb Collaboration** Collaboration, J. Alves, A. Augusto *et al.*, “The  
1067 LHCb Detector at the LHC,” *JINST* **3** (2008) S08005.
- 1068 [6] **ALICE Collaboration** Collaboration, P. Cortese, C. W. Fabjan,  
1069 L. Riccati, K. Safarík, and H. de Groot, *ALICE physics performance:  
1070 Technical Design Report*. Technical Design Report ALICE. CERN, Geneva,  
1071 2005. revised version submitted on 2006-05-29 15:15:40.
- 1072 [7] **UA4/2 Collaboration** Collaboration, C. Augier *et al.*, “Predictions on the  
1073 total cross-section and real part at LHC and SSC,” *Phys.Lett.* **B315** (1993)  
1074 503–506.
- 1075 [8] C. Mathis, T. Bhattacharya, and S. Imari Walker, “The Emergence of Life  
1076 as a First Order Phase Transition,” *ArXiv e-prints* (Mar., 2015) ,  
1077 arXiv:1503.02776 [nlin.AO].
- 1078 [9] F. Jegerlehner, “The hierarchy problem of the electroweak Standard Model  
1079 revisited,” arXiv:1305.6652 [hep-ph].

- 1080 [10] S. P. Martin, “A Supersymmetry primer,” *Adv.Ser.Direct.High Energy Phys.*  
 1081 **21** (2010) 1–153, [arXiv:hep-ph/9709356 \[hep-ph\]](#).
- 1082 [11] **CMS** Collaboration, V. Khachatryan *et al.*, “Search for neutral MSSM  
 1083 Higgs bosons decaying to a pair of tau leptons in pp collisions,” *JHEP* **1410**  
 1084 (2014) 160, [arXiv:1408.3316 \[hep-ex\]](#).
- 1085 [12] **ATLAS** Collaboration, G. Aad *et al.*, “Search for neutral Higgs bosons of  
 1086 the minimal supersymmetric standard model in pp collisions at  $\sqrt{s} = 8$  TeV  
 1087 with the ATLAS detector,” *JHEP* **1411** (2014) 056, [arXiv:1409.6064 \[hep-ex\]](#).
- 1088 [13] **Super-Kamiokande** Collaboration, Y. Ashie *et al.*, “Evidence for an  
 1089 oscillatory signature in atmospheric neutrino oscillation,” *Phys.Rev.Lett.* **93**  
 1090 (2004) 101801, [arXiv:hep-ex/0404034 \[hep-ex\]](#).
- 1091 [14] C. Weinheimer and K. Zuber, “Neutrino Masses,” *Annalen Phys.* **525**  
 1092 no. 8-9, (2013) 565–575, [arXiv:1307.3518](#).
- 1093 [15] C. L. Bennett, D. Larson, J. L. Weiland, N. Jarosik, G. Hinshaw,  
 1094 N. Odegard, K. M. Smith, R. S. Hill, B. Gold, M. Halpern, E. Komatsu,  
 1095 M. R. Nolta, L. Page, D. N. Spergel, E. Wollack, J. Dunkley, A. Kogut,  
 1096 M. Limon, S. S. Meyer, G. S. Tucker, and E. L. Wright, “Nine-year  
 1097 Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps  
 1098 and Results,” *The Astrophysical Journal, Supplement* **208** (Oct., 2013) 20,  
 1099 [arXiv:1212.5225 \[astro-ph.CO\]](#).
- 1100 [16] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett,  
 1101 J. Dunkley, M. R. Nolta, M. Halpern, R. S. Hill, N. Odegard, L. Page,  
 1102 K. M. Smith, J. L. Weiland, B. Gold, N. Jarosik, A. Kogut, M. Limon, S. S.  
 1103 Meyer, G. S. Tucker, E. Wollack, and E. L. Wright, “Nine-year Wilkinson  
 1104 Microwave Anisotropy Probe (WMAP) Observations: Cosmological  
 1105 Parameter Results,” *The Astrophysical Journal, Supplement* **208** (Oct.,  
 1106 2013) 19, [arXiv:1212.5226 \[astro-ph.CO\]](#).
- 1107 [17] **Planck** Collaboration, P. Ade *et al.*, “Planck 2015 results. XIII.  
 1108 Cosmological parameters,” [arXiv:1502.01589 \[astro-ph.CO\]](#).
- 1109 [18] A. Martin, W. Stirling, R. Thorne, and G. Watt, “Parton distributions for  
 1110 the LHC,” *Eur.Phys.J.* **C63** (2009) 189–285, [arXiv:0901.0002 \[hep-ph\]](#).
- 1111 [19] P. M. Nadolsky, H.-L. Lai, Q.-H. Cao, J. Huston, J. Pumplin, *et al.*,  
 1112 “Implications of CTEQ global analysis for collider observables,” *Phys.Rev.*  
 1113 **D78** (2008) 013004, [arXiv:0802.0007 \[hep-ph\]](#).
- 1114

- 1115 [20] R. D. Ball, L. Del Debbio, S. Forte, A. Guffanti, J. I. Latorre, *et al.*, “A first  
1116 unbiased global NLO determination of parton distributions and their  
1117 uncertainties,” *Nucl.Phys.* **B838** (2010) 136–206, arXiv:1002.4407  
1118 [[hep-ph](#)].
- 1119 [21] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory*.  
1120 Advanced book program. Westview Press Reading (Mass.), Boulder (Colo.),  
1121 1995. <http://opac.inria.fr/record=b1131978>. Autre tirage : 1997.



# **Contents**

1	<b>The Standard Model</b>	<b>5</b>
1.1	Fields, symmetries and interactions . . . . .	6
1.2	Quantum fields and particles . . . . .	8
1.2.1	The mass problem . . . . .	10
1.2.2	Spontaneous Symmetry Breaking . . . . .	11
1.2.3	Englert-Brout-Higgs mechanism . . . . .	12
1.3	Top production at LHC . . . . .	13
1.3.1	Pair production . . . . .	14
1.3.2	Single $t$ production . . . . .	14
1.3.3	Decay channels . . . . .	14
1.3.4	Top properties . . . . .	14
1.4	Higgs production at LHC . . . . .	18
1.4.1	Decay channels . . . . .	19
1.4.2	Higgs properties . . . . .	19
1.5	Hierarchy problem and other limitations . . . . .	19
2	<b>The CMS experiment at LHC</b>	<b>25</b>
2.1	The Large Hadron Collider . . . . .	25
2.1.1	Injector chain . . . . .	26
2.1.2	Main ring . . . . .	28
2.1.3	Run 1 . . . . .	31
2.1.4	Other experiments at LHC . . . . .	32
2.2	The Compact Muon Solenoid (CMS) experiment . . . . .	34
2.2.1	Coordinate system . . . . .	37
2.2.2	Magnet . . . . .	39
2.2.3	Tracker system . . . . .	40
2.2.4	Electromagnetic calorimeter . . . . .	42
2.2.5	Hadronic Calorimeter . . . . .	45
2.2.6	Muon chambers . . . . .	46
2.2.7	Trigger . . . . .	48
2.3	Object reconstruction . . . . .	50
2.3.1	Track and vertex reconstruction . . . . .	50
2.3.2	Particle Flow (PF) algorithm . . . . .	52

1155	2.3.3	Electron reconstruction . . . . .	53
1156	2.3.4	Muon reconstruction . . . . .	53
1157	2.3.5	Jet reconstruction . . . . .	54
1158	<b>3</b>	<b>VLQ models</b>	<b>57</b>
1159	3.1	Motivation . . . . .	57
1160	3.2	Generic Formulation . . . . .	57
1161	3.2.1	Production modes . . . . .	58
1162	3.2.2	Decay modes . . . . .	58
1163	3.3	Fesability study for a search of a $T$ at LHC at 8 TeV . . . . .	59
1164	3.3.1	Stragey for the full hadronic final state . . . . .	59
1165	3.3.2	Event selection . . . . .	60
1166	3.3.3	Results . . . . .	60
1167	<b>4</b>	<b>MC event generation</b>	<b>63</b>
1168	4.1	Mote-Carlo simulations . . . . .	63
1169	4.1.1	Parton simulation . . . . .	64
1170	4.1.2	Hadron simulation . . . . .	66
1171	4.1.3	Detector simulation . . . . .	66
1172	4.2	Tools . . . . .	66
1173	4.2.1	Matrix-element generators . . . . .	66
1174	4.2.2	Hadron generators . . . . .	66
1175	4.3	Validation on data . . . . .	68
1176	<b>5</b>	<b>Single VLQ search</b>	<b>71</b>
1177	5.1	Analysis Strategy . . . . .	71
1178	5.2	Datasets . . . . .	71
1179	5.3	Event selection . . . . .	75
1180	5.3.1	Event processing . . . . .	75
1181	5.3.2	Basic selection . . . . .	75
1182	5.3.3	$T'$ reconstruction with a $\chi^2$ sorting algorithm . . . . .	80
1183	5.3.4	Selection based on reconstructed objects . . . . .	84
1184	5.3.5	Efficiencies . . . . .	91
1185	5.4	Background estimation from data . . . . .	95
1186	5.4.1	Known difficulties and tried methods . . . . .	95
1187	5.4.2	Method . . . . .	96
1188	5.4.3	Validation . . . . .	96
1189	5.5	Systematics . . . . .	108
1190	5.6	Results . . . . .	113