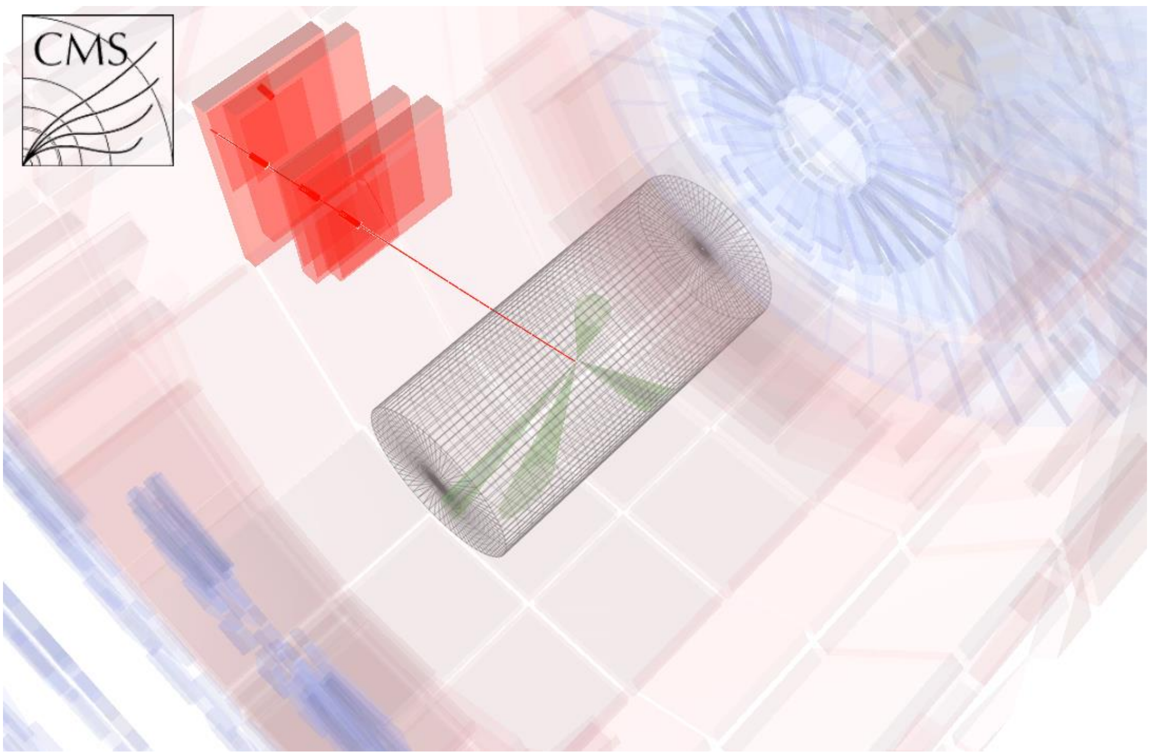
Università degli Studi di Pavia
Dottorato di Ricerca in Fisica – XXIX Ciclo

Search for a charged Higgs boson in top antitop pair events at CMS

Alice Magnani



Tesi per il conseguimento del titolo

Università di Pavia

Università degli Studi di Pavia
Dipartimento di Fisica

INFN Sez. di Pavia

DOTTORATO DI RICERCA IN FISICA – XXIX CICLO

Search for a charged Higgs boson in top antitop pair events at CMS

Alice Magnani

Submitted to the Graduate School of Physics in partial fulfillment of the requirements for the degree of

DOTTORE DI RICERCA IN FISICA

DOCTOR OF PHILOSOPHY IN PHYSICS

at the

University of Pavia

Supervisor: Prof. Cristina Riccardi

Cover: Simulated top pair event in the CMS detector, with the two top quarks decaying respectively into a W boson and charged Higgs boson

Search for a charged Higgs boson in top antitop pair events at CMS
Alice Magnani
PhD thesis - University of Pavia
Pavia, Italy, December 2016

Contents

List of Figures

Introduction

The Compact Muon Solenoid (CMS) is one of the two general purpose detectors currently working at the Large Hadron Collider (LHC). Since March 2010, when the experiment's operation took off, the CMS collabration has been giving notable contributions to high energy physics, publishing hundreds of physics papers in internationl journals. Huge effors have been spent to probe the standard model (SM) of particle physics, whose predictions have been found to be in fair agreement with experimental evidences. In this context, the discovery of the Higgs boson is the most remarkable cornerstone, but the SM robustness is backed up also by a number of analyses involving different processes.

Nevertheless, it's a widely shared opinion that the SM should be considered as an effective theory valid only at presently accessible energies. As a matter of fact, the need for an extended particle theory arises from various hints. For instance, the SM does not describe the gravitational force, it does not explain the pattern of fermion masses and its particle spectrum does not include any possible dark matter candidate. In addition, among the SM pitfalls, one should mention the so called unification and hierarchy problems. The search for new physics beyond the SM is then one of the primary purposes of the LHC detectors.

The wide range program of searches for new physics include investigation of possible extentions of the Higgs sectors. Many theoretical scenarios, including Supersymmetry, or more generally the two Higgs doublet models (2HDM), predict the existence of extra Higgs states.

Charged Higgs bosons have been looked for in a number of final states both at LEP and Tevaton and at the LHC, in a mass range spanning from the W mass to several hundreds GeV.

This thesis deals with the search for a light charged Higgs boson, having mass lower then the top quark. In this mass range, charged Higgs bosons can be produced via top quark decay, and thus they can be looked for in top quark pair events.

Searches for charged Higgs in top quark decay have been published so far both by ATLAS and CMS and are focused on the $\tau\nu$ and $c\bar{s}$ final states. No signal has been observed, and the observed upper limits on the branching ratio of the top quark into H^+ are at the level of percent.

Recently, the first attept to search for a charged Higgs in the $c\bar{b}$ has been made public by CMS. In the Minimal Supersymmetric Standard Model (MSSM), this channel has similar branching ratio with respect to the $c\bar{s}$ final state. Both channels are relevant in a particular region of the SUSY parameter space, defined by low values of the $tan\beta$ parameter. The branching fractions are at the level of some percent for $tan\beta \approx 2-3$ and can grow roughly to $\approx$10% as $tan\beta$ approaches 1. In addition to that, the $c\bar{b}$ decay mode is particularly enhanced in the so called flipped-2HDM scenario for the high and intermediate $tan\beta$ regime, where it can reach branching fractions of 60-80%.

The $H^+ \to c\bar{b}$ and $H^+ \to c\bar{s}$ analyses were carried out with a similar strategy, consisting in looking for semileptonic $t\bar{t}$ events with one top decaying to H^+b instead of Wb and subsequently going to $c\bar{b}$ or $c\bar{s}$, while the other top decays leptonically in the electron or muon final state ($\bar{t} \to W^-\bar{b} \to l\bar{\nu}\bar{b}$). The final state then consists of one lepton, missing energy and four jets, with three or two b-tagged jets. The main observable used in the analysis is the invariant mass of two jets indentified as the products of the charged Higgs decay. This dijet pair is selected from the jets in the event by a kinematic fitter.

The main source of background is due to semileptonic $t\bar{t}$ pair production for both the final states. However, since the branching ratio $W \to c\bar{b}$ is far suppressed with respect to $W \to c\bar{s}$, the $c\bar{b}$ final state can reach higher sensitivity, since the main $t\bar{t}$ background can be constrained applying an additional b-tagging with respect to the $c\bar{s}$ analysis.

The full Run I dataset collected in proton-proton collisions at a centre-of-mass energy of 8 TeV, corresponding to an integrated luminosity of 19.7 fb^{-1} was considered. No signal for the presence of a charged Higgs boson was observed in the $c\bar{b}$ final state and upper limits ranging from 1.1 to 0.4% were set on the branching fraction of the top quark to H^+b in the 90-150 GeV mass range.

Tuesday August 2 I presented the results to the CMS Collaboration. The analysis has been approved and results were presented at the International Conference of High Energy Physics held in Chicago in August 2016.

This thesis aims at presenting my personal contribution to the study, consisting in the development of a full and independent analysis flow for the search of charged Higgs to $c\bar{b}$ in the muonic final state. The study includes all the analysis steps, from Monte Carlo generation and sample processing to event selection, charged Higgs mass reconstruction, statistical analysis and evaluation of systematic uncertainties. The work was developed in a completely independent way with respect to the CMS result, thus providing a cross-check. Moreover, it allowed a better understanding of the SM background, in particular for the $t\bar{t}$ pair production in association with additional jets.

The thesis is organized as follows:

After a brief introduction about the SM of elementary particle physics, Chapter1 presents the extended Higgs sectors in the MSSM and 2HDM. A summary of the current status of charged Higgs searches is included.

Chapter2 describes the experimental setup, the design and operational parameters of the LHC and the CMS detector.

Chapter3 reports the CMS track reconstruction.

Chapter4 presents the search for a light charged Higgs boson decaying into a charm quark and a bottom antiquark, including the analysis strategy, the event selection, background estimation, the kinematical technique used for the charged Higgs mass reconstructon, statistical interpretation of the results and systematic uncertainties.

Chapter 1

The Charged Higgs Boson

1.1 The Standard Model of particle physics

1.1.1 Fundamental particles

Our understanding of matter and energy dynamics lies on a set of fundamental theories grown up during 1960s and 1970s. This description, called "Standard Model" (SM) reduces all the known phenomena (except gravity) to simple interactions between elementary particles.

In the current view, all the visible matter is made up by *fermions*, particles with half spin, that interact through the exchange of *bosons*, particles with integer spin, as a result of gauge-invariant theories.

Fermions can be divided into two main groups, *leptons* and *quarks* according to their different behavior with respect to fundamental interactions. Their classification is given in Table 1.1 and 1.2. Both leptons and quarks fall naturally under three generations. All the stable matter around us is made by fermions belonging to the first one, while due to their high masses, second and third generation particles can be produced just in accelerators and in the extreme conditions in primordial universe.

The SM theory describes three fundamental forces:

- The electromagnetic interacton, mediated by the massless and chargeless spin-one photon. Since the mass of the photon is zero, it can mediate interactions to infinite distances.

- The weak interaction between *fermions*, mediated by the charged spin-one $W^{\pm}$ and the neutral spin-one Z bosons, discovered at CERN in 1983. Since they carry mass, the weak interaction is short range.

- The strong interaction, responsible for actractive force between quarks. It is mediated by the massless, chargeless spin-1 gluons. Due to the asymptotic freedom of strong interactions, whose intensity decreases with the energy, quarks are never observed in free state, but only in bound states of two (*mesons*) or three (*baryons*) quarks, and any attempt to isolate single quarks give rise to a new quark pair.

The main parameters of the electromagnetic, weak and strong interactions are summarized in Table 1.3

Table 1.1: The three families of leptons (spin 1/2). Numerical values are taken from [1]. Neutrino masses are extremely small, and for most purposes can be taken to be zero

Generation	Flavor	Charge	Mass
first	e (electron)	-1	511 keV
	ν_e (e neutrino)	0	<2 eV
second	μ (muon)	-1	105.6 MeV
	ν_μ (μ neutrino)	0	<2 eV
third	τ (tau)	-1	1.78 GeV
	ν_τ (τ neutrino)	0	<2 eV

Table 1.2: The three families of quarks (spin 1/2). Numerical values are taken from [1]

Generation	Flavor	Charge	Mass
first	d (down)	-1/3	4.8 MeV
	u (up)	+2/3	2.3 MeV
second	s (strange)	-1/3	95 MeV
	c (charm)	+2/3	1.275 GeV
third	b (bottom)	-1/3	4.18 GeV
	t (top)	+2/3	173.2 GeV

Table 1.3: The elementary bosons of the SM. Numerical values are taken from [1]

Force	Mediator	Charge	Mass (GeV)	Couplng Constant
Strong	g (8 gluons)	0	0	$\alpha_S = 0.1184$
Electromagnetic	γ (photon)	0	0	$\alpha = 1/137$
Weak	$W^\pm$ (charged)	± 1	80,385	$\alpha_W = 1.02 \times 10^{-5}$
	Z (neutral)	0	91.187	$\alpha_W = 1.02 \times 10^{-5}$

1.1.2 The Standard Model

The SM is a quantum field theory based on the symmetry groups $SU(3)_C \times SU(2)_L \times U(1)_Y$. The electroweak theory (EW) describes the electromagnetic and weak interactions between quarks and leptons. It is a Yang-Mills theory based on the gauge symmetry group $SU(2)_L \times U(1)_Y$ of weak left-handed isospin and hypercharge [2, 3, 4]. Quantum chromodynamics (QCD), describing strong interactions between quarks and gluons, is based on $SU(3)_C$ gauge symmetry [5].

The SM before electroweak symmetry breaking In the SM formalism [6], the matter fields are represented by three generations of left-handed and right-handed chiral quarks and leptons $f_{L,R} = \frac{1}{2}(1 \mp \gamma_5) f$[1]. The left-handed fermions are in weak isodoublets, while the right-handed fermions are in weak isosinglets

$$
\begin{aligned}
& L_1 = \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, e_{R_1} = e_R^-, Q_1 = \begin{pmatrix} u \\ d \end{pmatrix}_L, u_{R_1} = u_R, d_{R_1} = d_R \\
I_f^{3L,3R} = \pm\frac{1}{2}, 0 : \; & L_2 = \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \mu_{R_1} = \mu_R^-, Q_1 = \begin{pmatrix} c \\ s \end{pmatrix}_L, c_{R_1} = c_R, s_{R_1} = s_R \\
& L_3 = \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L, \tau_{R_1} = \tau_R^-, Q_1 = \begin{pmatrix} t \\ b \end{pmatrix}_L, t_{R_1} = t_R, b_{R_1} = b_R
\end{aligned}
\tag{1.1}
$$

The fermion hypercharge Y_f, defined in terms of the third component of the weak isospin I_f^3 and the electric charge Q_f in units of the proton charge $+e$, is given by:

$$Y_f = 2Q_f - 2I_f^3. \tag{1.2}$$

The SM lagrangian, without mass terms for fermions and gauge bosons, reads:

$$L = -\frac{1}{4} W_{\mu\nu}^a W_a^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu} + \bar{L}_i i D_\mu \gamma^\mu L_i + \bar{e}_{R_i} i D_\mu \gamma^\mu e_{R_i}, \tag{1.3}$$

where:

- B_μ is a real vectorial field corresponding to the generator Y of the group $U(1)_Y$
- W_μ^a is a real vectorial field triplet corresponding to the generators of the $SU(2)_L$ group T_a

[1]Here γ_5 stands for one of the Gamma matrices [7]. $\frac{1}{2}(1 \mp \gamma_5)$ are the left- and right-handed projector operators

- the covariant derivative D_μ is defined as:

$$D_\mu\psi =_s (\partial_\mu - ig_3 T_a G^a_\mu - ig_1\frac{Y_q}{2}B_\mu)\psi, \tag{1.4}$$

being g_1 , g_2 and g_3 the coupling constants of the $SU(3)_C$, $U(1)_Y$ and $SU(2)_L$ groups respectively.

The lagrangian 1.4 is invariant under local gauge transformations of the $SU(2)_L \times U(1)_Y$ group of the form:

$$\begin{aligned} L_i \rightarrow & e^{i\alpha_a(x)T^a + i\beta(x)Y} L_i \\ e_{R_i} \rightarrow & e^{i\beta(x)\frac{Y}{2}e_{R_i}} \\ B_\mu \rightarrow & B_\mu - \frac{1}{g_1}\partial_\mu\alpha(x) - \overrightarrow{\alpha}(x) \times \overrightarrow{W}_\mu(x) \end{aligned} \tag{1.5}$$

However, it can be seen that this invariance is broken if one tries to incorporate mass terms for the weak vector bosons $\frac{1}{2}M_V^2 W_\mu W^\mu$. In addition, if one tries to include a mass term $-m_f\bar{\psi}\psi$, for each SM fermion f in the lagrangian, the result is manifestly not invariant under isospin symmetry transformations. Therefore, apparently, there is no way to account for the mass of the weak bosons and fermions, without giving up the principle of exact unbroken gauge symmetry.

The EWSM mechanism and the SM Higgs boson A cornerstone of the SM is the mechanism of spontaneous symmetry breaking (EWSB) proposed in 1964 by Higgs, and independently by Brout and Englert, Guralnik, Hagen and Kibble. The theory allows to generate the weak vector boson masses without violating the $SU(2) \times U(1)$ invariance.

The trick consists in introducing a new SU(2) doublet of complex scalar fields with hypercharge +1:

$$\Phi = \begin{pmatrix}\phi^+ \\ \phi^-\end{pmatrix}, Y_\phi = +1. \tag{1.6}$$

The gauge invariant lagrangian for this field reads:

$$L_S = (D^\mu\Phi)^\star(D_\mu\Phi)) - V(\Phi), V(\Phi) = \mu^2|\Phi|^2 + \lambda^2|\Phi|^4. \tag{1.7}$$

As indicatively shown in Figure 1.1, if the mass term μ^2 is positive, the potential $V(\Phi)$ is also positive and L becomes simply the lagrangian of a spin-zero particle of mass μ. In turn, for $\mu^2 < 0$, the neutral component of the doublet field ϕ develops a non-zero vacuum expectation value. The minimum

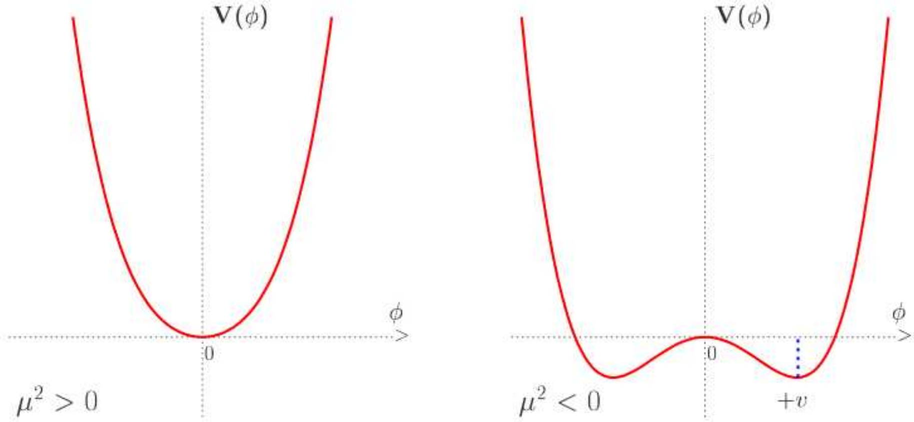


Figure 1.1: The potential V of the scalar field ϕ in the case $\mu^2 > 0$ (left) and $\mu^2 < 0$ (right)

of the $V(\Phi)$ potential can be arbitrarly chosen among all the points satisfying the condition:

$$|\phi^+|^2 + |\phi_0|^2 = -\frac{\mu^2}{2\lambda} = \frac{v^2}{2}, \tag{1.8}$$

for instance one can take

$$< \Phi_0 >=< 0|\Phi|0 >= \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}v \end{pmatrix}. \tag{1.9}$$

Redefining Φ in terms of a small vibration H around Φ_0

$$< \Phi >= \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v + H(x)) \end{pmatrix} \tag{1.10}$$

in the L_S lagrangian, one gets terms that are bilinear in the fields $W^{\pm}$, Z, A:

$$M_W^2 W_\mu^+ W^{-\mu} + \frac{1}{2} M_Z^2 Z_\mu Z^\mu + \frac{1}{2} M_A^2 A_\mu A^\mu \tag{1.11}$$

with

$$M_W = \frac{1}{2} v g_2, M_Z = \frac{1}{2} v \sqrt{g_1^2 + g_2^2}, M_A = 0 \tag{1.12}$$

The electroweak $SU(2)_L \times U(1)_Y$ symmetry is spontaneously broken to the electromagnetic $U(1)_C$ symmetry. Three of the four degrees of freedom of the doublet scalar field are absorbed by the $W^{\pm}$ and Z weak vector bosons to form their longitudinal polarizations and to acquire masses.

The fermion masses can be generated introducing in the lagrangian a new Yukawa interaction with the same scalar field Φ and its conjugate field:

$$L_F = -\lambda_e \bar(L) \Phi e_R - \lambda_d \bar(Q) \Phi d_R - \lambda_u \bar(Q) \Phi u_R + h.c. \tag{1.13}$$

The remaining degree of freedom corresponds to a scalar particle, the Higgs boson, whose lagrangian reads:

$$L_H = \frac{1}{2}(\delta^\mu H)^2 - \lambda v^2 H^2 - \lambda v H^3 - \frac{\lambda}{4} H^4 \tag{1.14}$$

Elementary particles couple to the Higgs boson with a strenght that is dependent on their mass, the coupling constants being $i\frac{m_f}{v}$ for fermions and $-2i\frac{M_V^2}{v}$ for gauge bosons.

The Higgs boson mass is given by $m_H^2 = 2\lambda v^2 = -2\mu^2$. Besides the lack of prediction from the SM theory on its value, the Higgs mass can be constrained from EW precision measurements. As a matter of fact, the Higgs boson contributes to radiative corrections on the top quark and W boson masses. Therefore, precision measurements of electroweak parameters, like the top quarks and the W and Z vector boson masses can be combined to perform a $\Delta\chi^2$ fit on the Higgs boson mass (Figure 1.2). The preferred value for the Higgs boson mass results 87 GeV, corresponding to the minimum of the fitting curve, with an uncertaity of +35 GeV ad -26 GeV at 68% of CL.

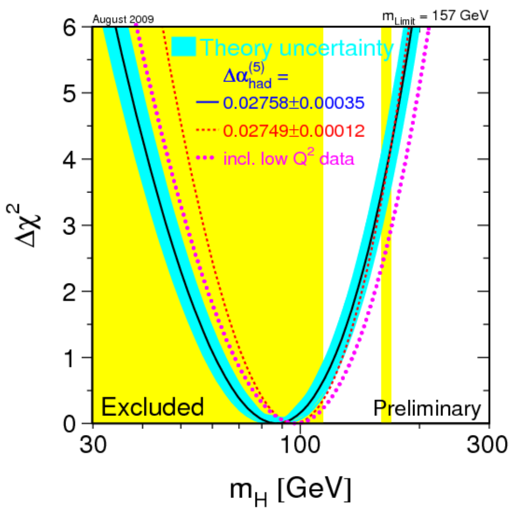


Figure 1.2: $\Delta\chi^2$ fit to the Higgs boson mass from electroweak precision measurements. The blue band represents the LEP exclusion up to a Higgs mass of 114.4 GeV, the yellow band to the right represents the Tevatron exclusion of Higgs masses between 162-166 GeV. Both exclusions are made at 95% [8].

The experimental observation at the LHC of a new resonance compatible with the Higgs boson was finally announced in July 2012 by the CMS and ATLAS collaboration [9, 10]. The mass of the observed Higgs state, around 125 GeV, is fully compatible with expectations from the electroweak fit.

1.2 Experimental tests on the SM at the LHC

The SM has been extensively studied at colliders in the last decades and it has been prooved to be an extremely successful theory, whose predictions can explain the whole set of measurements performed at LEP, Tevatron, LHC and elsewhere within the uncertainties.

Figure 1.3 shows a summary of cross sections measured at CMS with LHC data collected at center-of-mass energy of 7 and 8 TeV. Full agreement with SM expectations is observed over the whole range of processes, including production of electroweak bosons W and Z, diboson VV , top quark and Higgs boson.

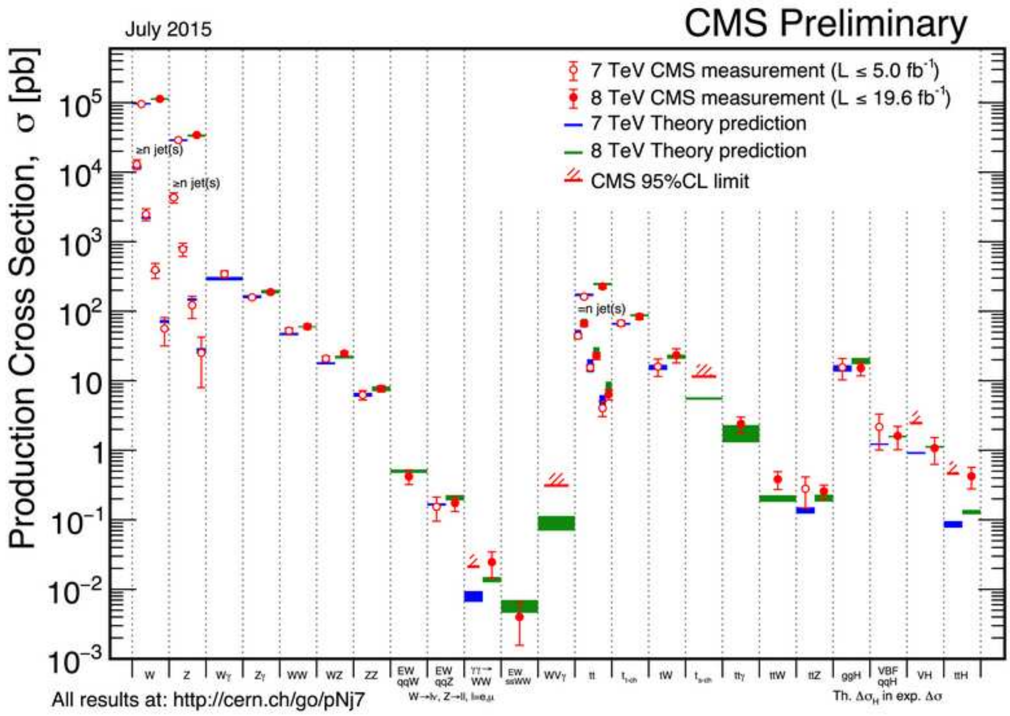


Figure 1.3: Summary of the SM cross sections measured at CMS and comparison with theoretical expectation. [11].

Concerning the production of electroweak bosons, Figure 1.4 shows a summary of the total W^+, W^-, W and Z cross sections times branching fractions as a function of the center-of-mass energy for CMS and other experiments at lower-energy colliders. The predicted behavior of the cross sections as a function of the center-of-mass energy is in remarkable agreement with experimental measurements done at different colliders, in a broad energy range.

Figure 1.5 shows cross section for the Drell-Yan process at a fixed center-of-mass energy of 8 TeV as a function of the invariant mass of the dilepton system. The dilepton invariant mass spectrum is in fair agreement with theory expectation over a wide energy range spanning three order of magnitues from ≈ 10 GeV to 1 TeV.

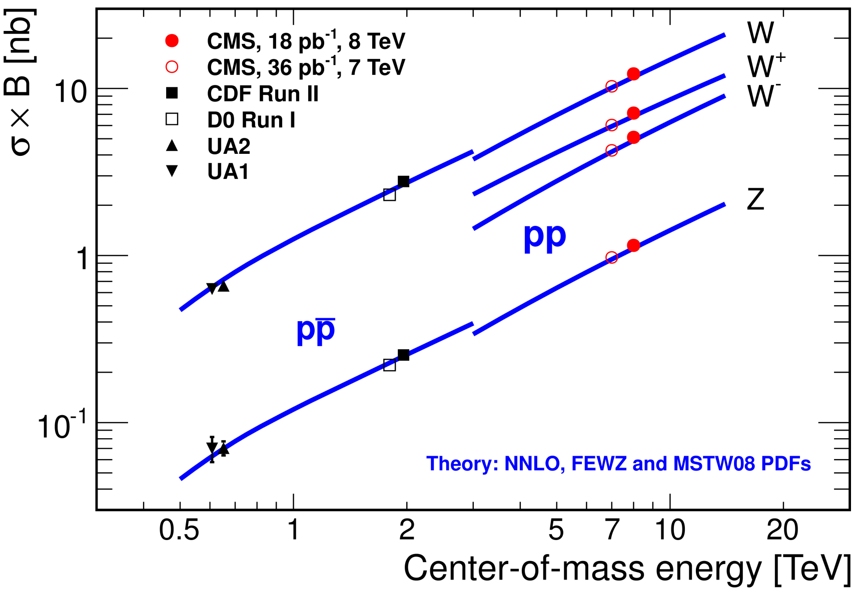


Figure 1.4: Total W^+, W^-, W and Z production cross sections times branching fractions in the center-of-mass energy for CMS and experiments at lower-energy colliders [12].

SM predictions have been confirmed not only wih measurements of inclusive cross sections, but also studying differential production processes. Measurements of differential jets production rates in association with W and Z bosons can provide a test of perturbative-QCD calculations and are sensitive to the possible presence of new physics. To reduce systematic uncertainties associated with the integrated luminosity measurement, the jet energy scale, the lepton reconstruction and trigger efficiencies, the V + n jets cross sections can be measured relatively to the inclusive W and Z production cross sections, as $\sigma(V+\geq n\,\text{jets})/\sigma(V)$. This measurement performed with CMS data collected at center-of-mass energy of 7 TeV is presented in Figure 1.6. The results are again in agreement with SM calculations.

Measurements of diboson production processes have been also performed and can provide stringent tests for the SM. Under the assumption that the new physics scale Λ is greater than the energies currently accessible at the LHC, the effects of the presence of BSM particles can be described by operators with mass dimensions larger than four in an effective field theory (EFT) framework. The higher-dimensional operators of the lowest order from purely electroweak

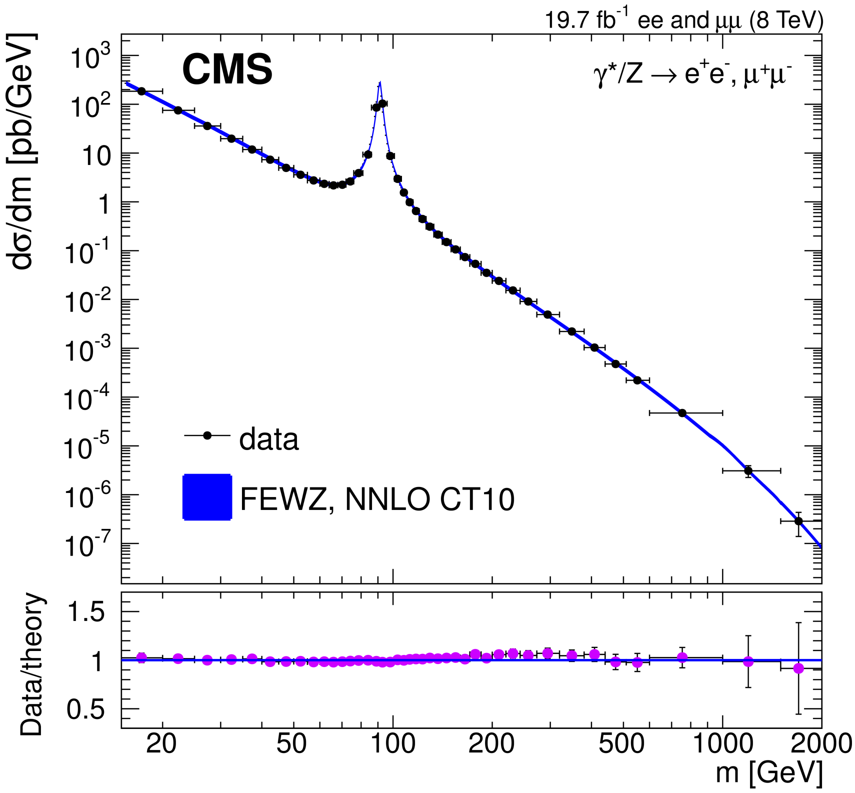


Figure 1.5: The Drell-Yan cross section as a function of the invariant mass of the dilepton system measured at CMS in the combined dilepton channel compared to next-to-next-to-leading order calculation. [13].

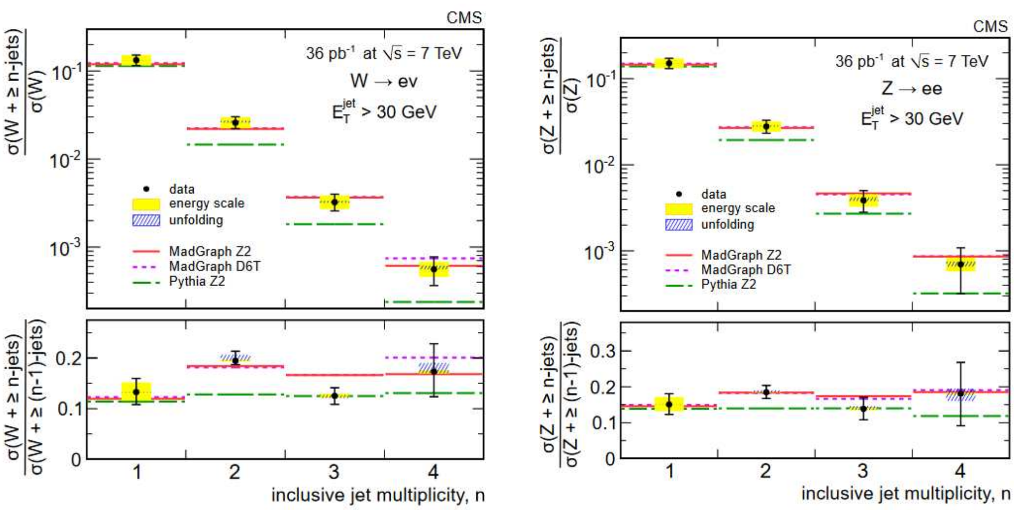


Figure 1.6: Left: The ratios $\sigma(W+\geq n$ jets$)/\sigma(W)$ (top) and $\sigma(W+\geq n$ jets$)/\sigma(W+\geq(n-1)$ jets$)$ (bottom) in the electron channnel. Right: The ratios $\sigma(Z+\geq n$ jets$)/\sigma(Z)$ (top) and $\sigma(Z+\geq n$ jets$)/\sigma(Z+\geq(n-1)$ jets$)$ (bottom) in the electron channnel. [14].

processes have dimension six, and can be written in the form:

$$\begin{aligned}\frac{c_{WWW}}{\Lambda^2}O_{WWW} &= \frac{c_{WWW}}{\Lambda^2}Tr[W_{\mu\nu}W^{\nu\rho}W^{\mu}_{\rho}] \\ \frac{c_W}{\Lambda^2}O_W &= \frac{c_W}{\Lambda^2}(D^\mu\Phi)^\dagger W_{\mu\nu}(D^\nu\Phi) \\ \frac{c_B}{\Lambda^2}O_B &= \frac{c_B}{\Lambda^2}(D^\mu\Phi)^\dagger B_{\mu\nu}(D^\nu\Phi)\end{aligned} \tag{1.15}$$

These operators generate anomalous trilinear gauge couplings at three level and modify the VV production cross section including the W^+W^- one. In particular, the invariant mass of the two leptons produced in the W^+W^- decay would be sensitive to these anomalous terms and can been used to extract constraints on the anomalous coupling constants [15]. Figure 1.7 shows the results of these kind of fits performed using CMS data at center-of-mass energy of 8 TeV. The fit values for the anomalous coupling are consistent with the SM hypothesis.

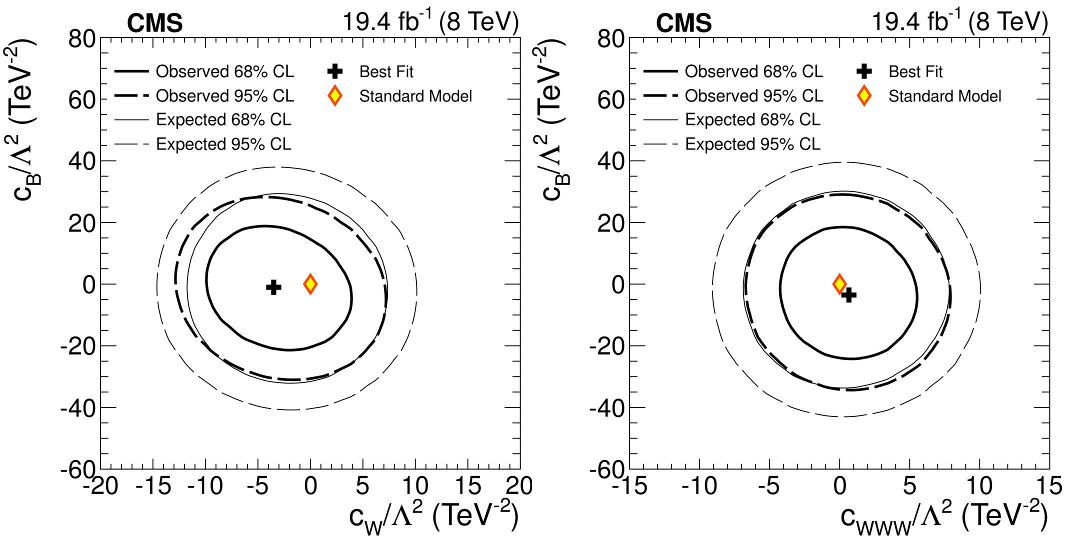


Figure 1.7: Two-dimentional observed (thick lines) and expected (thin lines) 68% and 96% CL contours for anomalous coupling constants $c_{WWW}/\Lambda^2 \times c_W$ and $c_W/\Lambda^2 \times c_B$ [15].

The discovery of the Higgs boson at the LHC provided another proof of the SM success and is considered a milestone in particle physics, since at the LHC it is now possible to probe with direct investigations and measurements the SM Higgs sector, which was not accessible at other colliders. At now, all the measurements performed are in agreement with the SM predictions within uncertainties.

The experimental value of the Higgs boson mass, obtained from a combination of the results of the ATLAS and CMS experiments, is [11]:

$M_H = 125.09 \pm 0.21(stat) \pm 0.11(sys) GeV/c^2$

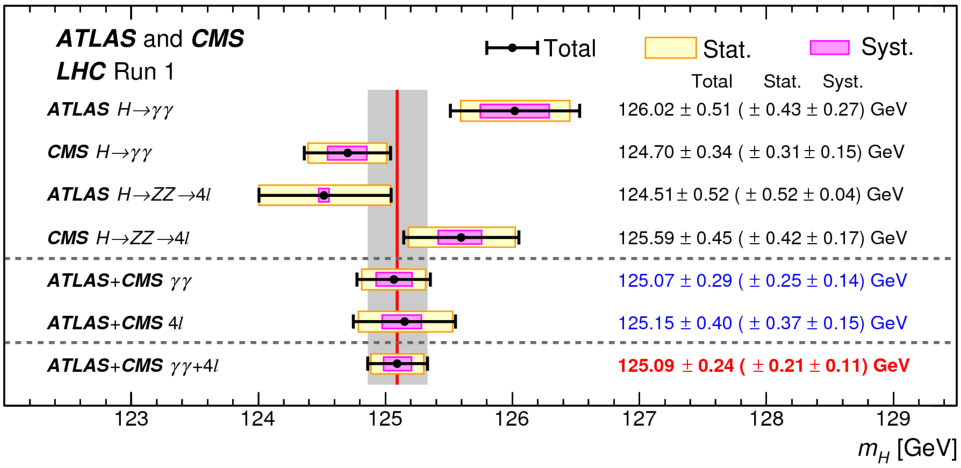


Figure 1.8: Summary of Higgs boson mass measurement from individual Run1 analyses of ATLAS and CMS and in the $H \to \gamma\gamma$ and $H \to 4l$ final states and their combination [16].

The Higgs boson cross sections in the different production modes measured at CMS are included in Figure 1.3. Results for the gluon gluon fusion, the vector boson fusion and associated production are in agreement with calculations for a Higgs mass value of 125 GeV. A small, non-significative deviation is observed only in the ttH production mode. The experimental Higgs signal strenghts, obtained in a combination of ATLAS and CMS results are shown in Figure 1.9. For each decay channel the result is consistent with SM expectations, though the large experimental uncertainties, of the order of $\approx 10\%$ leave room for BSM effects that could be highlighted by probing the Higgs couplings with increasing precision.

1.3 Beyond the SM

1.3.1 Motivation for physics beyond the SM

The high-precision measurements carried out at LEP, SLC, Tevatron, LHC and elsewhere have provided a decisive test of the SM and firmly established that it provides the correct effective description of the strong and electroweak interactions. Nevertheless, the SM is widely believed to be an effective theory valid only at presently accessible energies. Besides the fact that it does not include the gravitational force and the fermions masses are just free parameters of the theory, it has at least three issues that still require explanation:

- The SM is based on the $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetry, the direct product of three simple groups with different coupling constants and, in this sense, does not provide a true unification of the electroweak

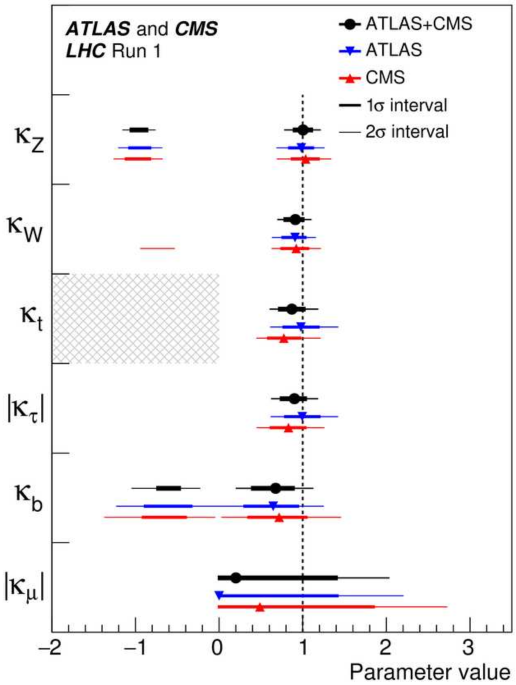


Figure 1.9: Best fit values of the Higgs boson couplings for the combination of ATLAS and CMS data, and separately for each experiment, for the parameterization assuming the absence of BSM particles in the loops [16].

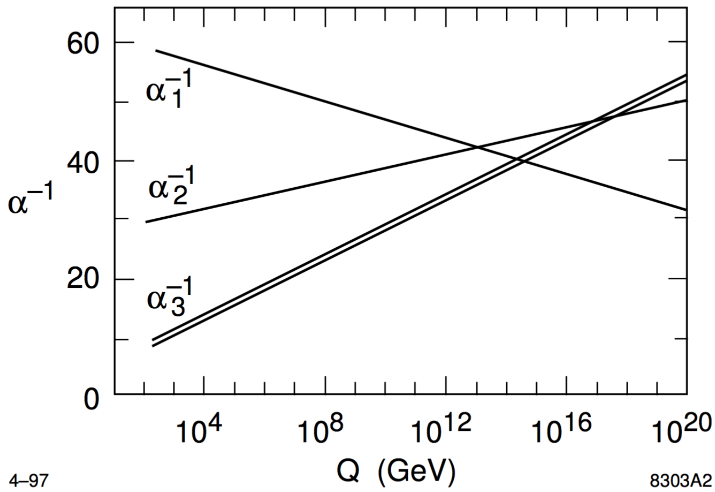


Figure 1.10: Evolution of the SM couplings $\alpha_i = \frac{g_i^2}{4\pi}$ as a funnction of the energy scale [17]

and the strong interactions. Therefore, one expects the existence of a more fundamental Grand Unified Theory (GUT), which describes the three forces within a single group, such as $SU(5)$ or $SU(10)$, with just one coupling constant. However, given the high-precision measurements at LEP and elsewhere, the coupling constants fail to meet at the GUT scale [17] (see Figure 1.10).

- Astronomical observations show that a large contribution to the critical density of the universe, about 25%, must be due to some kind of non-baryonic, non-luminous matter [18]. A particle that is stable, massive electrically neutral is required. The SM does not have any dark matter candidate.

- In the SM, when calculating the radiative corrections to the Higgs boson mass squared, one encounters divergences quadratic in the cut-off scale Λ at which New Physics should appear:

$$\Delta M_H^2 = N_f \frac{\lambda_f^2}{8\pi^2}\left[-\Lambda^2 + 6m_f log\frac{\Lambda}{m_f} - 2m_f\right]. \tag{1.16}$$

If the cutoff scale Λ is set to the GUT scale $\approx 10^{16}$ GeV, or the Planck scale, $\approx 10^{18}$ GeV, the Higgs mass would prefer to be close to very high scale, and thus huge. The existence of the Higgs boson with a mass of approximately 125 GeV embodies the problem of an unnatural cancellation among the quantum corrections to its mass [17] . This fine tuning could be solved assuming the existence of a number $N_S = 2N_f$ of scalar particles with a symmetry relating their couplings λ_S to the ones of standard fermions: $\lambda_f^2 = -\lambda_S$. The correction 1.16 would become:

$$\Delta M_H^2 = N_f \frac{\lambda_f^2}{4\pi^2}\left[(m_f^2 - m_S^2 log(\frac{\Lambda}{m_S}) + 3m_f^2 log(\frac{m_S}{m_f})\right]. \tag{1.17}$$

and quadratic divergences would disappear. The logaritmic divergence would still be present, but even for values $\Lambda \approx M_P$ of the cutoff, the contribution would be rather small. It would disappear under the assumption that the fermion and the two scalars have exactly the same mass.

1.4 Extended Higgs sector in the MSSM

Many extentions of the SM have been proposed in the last years, in order to address its open questions. Most of BSM theories would introduce modifications in the EWSB mechanism of the SM. The observed 125 GeV Higgs boson may be part of an extended Higgs sector. In the following paragraph, the two most compelling classes of BSM theories will be briefly presented: the Composite Higgs models and two Higgs doublet model. Both of them can accomodate the observed 125 GeV Higgs boson, and predict the existence of additional resonances. In particular, the 2HDM spectrum includes charged Higgs states $H^{\pm}$.

1.4.1 Composite Higgs model

A light Higgs boson could emerge as the bound state of a strongly interacting sector, rather than being an elementary field. A composite Higgs would solve the hierarchy problem of the SM, as its mass is not sensitive to virtual effects above the compositness scale, in the same way as the mass of the QCD pion does not receive corrections at the Planck scale [19].

The starting point of composite Higgs theories consists in considering a strongly interacting sector with a global symmetry G dynamically broken to H_1 at the scale f. The subgroup $H_0 \subset G$ is gauged by external vector bosons. The global symmetry breaking $G \to H_1$ gives rise to $n = dim(G) - dim(H_0)$ degrees of freedom. Calling H the unbroken gauge group $H = H_0 \cap H_1$, a number $n_0 = dim(H_0) - dim(H)$ of the n degrees of freedom are absorbed to give mass to as many vector bosons. The remaining $n - n_0$ are pseudo Nambu-Goldston bosons.

This construction introduces new massive states with the quantum numbers of (t, W, Z). Diagrams with these particles cancel the usual diagrams that give quadratic divergences in the Higgs mass.

Composite models can be probed at the LHC both in direct and in indirect searches. Indirect constraints can be derived by measurements in the field of flavor physics (e.g. $B\bar{B}$ mixing, angular observables in $B \to K^{\star}\mu\mu$, $BR(B_s \to \phi\mu\mu$, rare B decays, etc.), Z decays (e.g. with precision measurements of its branching ratios), SM Higgs production and decay. On the other hand, compositness would give rise to new resonances that could decay to fermions, weak vector bosons W and Z, or in the SM Higgs boson.

ATLAS and CMS are following an extensive program for the search of new resonances. No evidence for new physics has been found so far. Figures 1.11 and 1.12 show the current state of the art for searches of heavy resonances decaying to weak vector bosons pair or SM Higgs pairs with CMS data.

1.4.2 2HDM models

The simplest way to extend the SM Higgs sector, and the most extensively studied at colliders, consists in adding to the SM lagrangian two doublet of complex scalar fields instead of one [22, 23]. A broad class of models can be framed within 2HDMs, Supersimmetry (SUSY) is an example. SUSY has been for long considered one of the most attractive extentions of the SM, as it could potentially solve the dark matter, unification and hierarchy problems simultaneously.

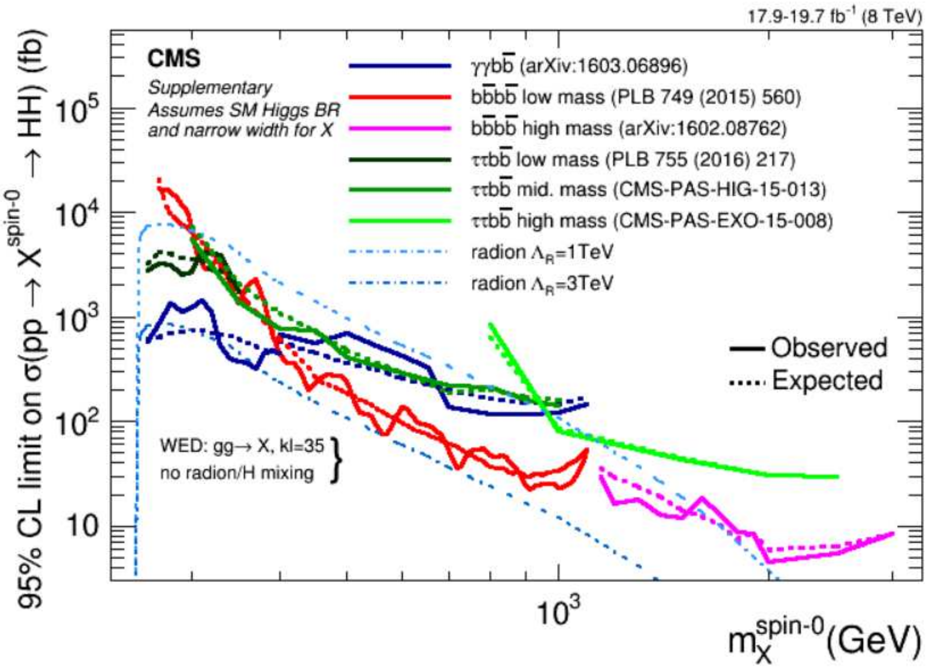


Figure 1.11: Summary of the exclusion limits set by CMS in di-higgs searches with 8 TeV data [20].

The 2HDM formalism consists in defining two Higgs doublets

$$\begin{aligned} \Phi_1 &= \frac{1}{\sqrt{2}}\begin{pmatrix} \phi_1^0 + ia_1^0 \\ \sqrt{2}\phi_1^- \end{pmatrix} \\ \Phi_2 &= \frac{1}{\sqrt{2}}\begin{pmatrix} \sqrt{2}\phi_2^+ \\ \phi_2^0 + ia_2^0 \end{pmatrix} \end{aligned} \tag{1.18}$$

with hypercharge $Y = -1$ and $Y = 1$ respectively. The Higgs potential reads:

$$\begin{aligned} V =& m_1^2\Phi_1^\dagger\Phi_1 + m_2^2\Phi_2^\dagger\Phi_2 - m_3^2(\Phi_1^T i\sigma_2\Phi_2 + h.c.) \\ &+ \frac{1}{2}\lambda_1(\Phi_1^\dagger\Phi_1)^2 + \frac{1}{2}\lambda_2(\Phi_2^\dagger\Phi_2)^2 + \lambda_3(\Phi_1^\dagger\Phi_1)(\Phi_2^\dagger\Phi_2) \\ &+ \lambda_4|\Phi_1^T i\sigma_2\Phi_2|^2 + \frac{1}{2}\lambda_5[(\Phi_1^T i\sigma_2\Phi_2)^2 + h.c.] \\ &+ [[\lambda_6(\Phi_1^\dagger\Phi_1) + \lambda_7(\Phi_2^\dagger\Phi_2)]\Phi_1^T i\sigma_2\Phi_2 + h.c.] \end{aligned} \tag{1.19}$$

where $m_i^2 = \mu^2 + m_{H_i}^2$, with μ being the supersymmetric Higgsino mass parameter and m_{H_i} (for i=1,2) the Higgs doublet soft supersymmetric breaking mass parameter, $m_3^2 \equiv B_\mu$ is associated to the B-term soft SUSY breaking parameter and λ_i are the Higgs quartic coupling.

After symmetry breaking, the two doublet fields lead to five physical Higgs particles: two CP even states, h and H, a CP-odd scalar A and one charged Higgs pair $H^\pm$.

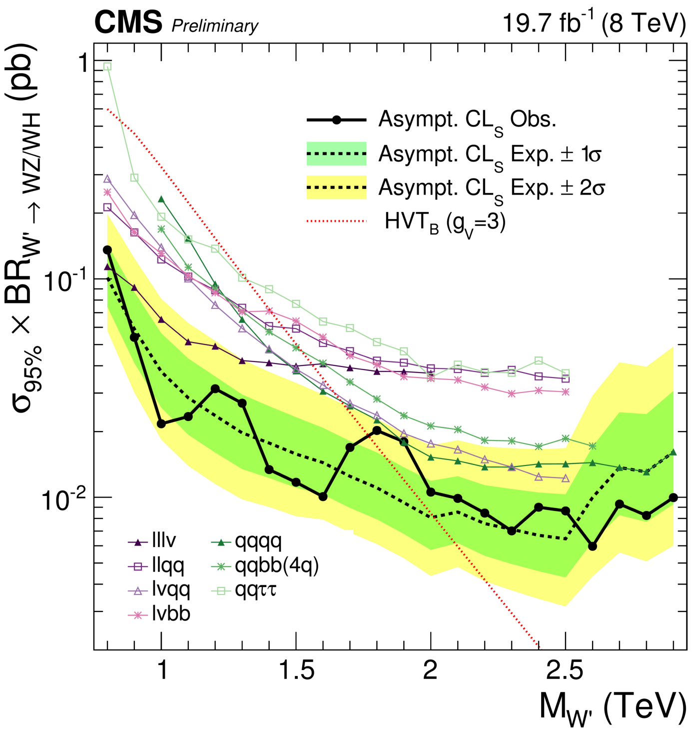


Figure 1.12: Observed and expected exclusion limit at 95% CL on $\sigma(pp \to V\prime \to WV/VH)$ as a function of the resonance mass obtained combining the results of diboson analyses with CMS data at 8 TeV [21].

Imposing the constraint of natural flavor conservation, that implies that fermions with a common electric charge are generated through couplings to exactly one Higgs doublet [24], there are four different ways to couple the SM fermions to the Higgs doublet, as summarized in table 1.4.

Table 1.4: The four possible assignments of fermion couplings to two Higgs doublets that satisfy natural flavor conservation. Here u, d, and l represent up- and down-type quarks and charged leptons, respectively

Model	Type I	Type II	Lepton-specific	Flipped
Φ_1	-	d, l	l	d
Φ_2	u, d, l	u	u, d	u, l

1.4.3 The MSSM

SUSY is a symmetry relating particles of integer spin (spin-0 and spin-1 bosons) and particles of spin $\frac{1}{2}$ (fermions). The SUSY generators Q transform fermions into bosons and vice versa

$$\begin{aligned} Q|Fermion> &= |Boson> \\ Q|Boson> &= |Fermion> \end{aligned} \tag{1.20}$$

When the symmetry is exact, the bosonic fields and the fermionic fields have the same masses and quantum numbers, except for the spin. However, since there are no experimental evidences for scalar particles having the same mass as known fermions, SUSY must be a broken symmetry. Usually one assumes SUSY-breaking to occur in such a way that the supersymmetric particles are not too heavy, in order to solve the aforementioned problem about the unnatural fine-tuning in the quantum correction on the Higgs mass. For the radiative corrections to be of the same order as the tree level M_h the SUSY-breaking scale should be around 1 TeV.

The Minimal Supersymmetric Standard Model (MSSM) [25, 26] is defined as the SUSY minimal realization in terms of gauge symmetries and particle content. SUSY demands the existence of two Higgs doublets such that one doublet couples to up-type quarks and the other to down-type quarks and charged leptons.

This Higgs-fermion coupling structure is the one identified as type-II 2HDM and assures that masses for both up and down-type quarks can be generated in a supersymmetric and gauge invariant way.

The Higgs sector is in principle described by 6 parameters, the masses of the Higgs states, the angle α that diagonalizes the mass eigenvalue matrix h and H and the β angle, whose tangent is defined as the ratio of the Higgs doublets vacuum expectation values $tan\beta = v_2/v_1$. However, at tree level, the parameters are related through the following equations:

$$M_{H^\pm}^2 = M_A^2 + M_W^2 \tag{1.21}$$

$$M_{h,H}^2 = \frac{1}{2}[M_A^2 + M_Z^2 \mp \sqrt{(M_A^2 + M_Z^2)^2 - 4M_A^2 M_Z^2 cos2\beta}] \tag{1.22}$$

$$\alpha = \frac{1}{2}arctan\big(tan2\beta \frac{M_A^2 + M_Z^2}{M_A^2 - M_Z^2}\big), -\frac{\pi}{2} \le \alpha \le 0 \tag{1.23}$$

It follows that only two parameters are needed to describe the system at tree level. Usually one relies on the mass of the pseudoscalar M_A, expected to lie in the range between M_Z and the SUSY breaking scale, and the ratio $tan\beta$, which is expected to take values in the range:

$$1 \lesssim tan\beta \lesssim \bar{m_t}/\bar{m_b} \approx 60 \tag{1.24}$$

with $\bar{m_t}$ and $\bar{m_b}$ the running top and bottom quark masses in the ($\bar{MS}$) renormalization scheme [27] evaluated at a scale close to the SUSY scale M_S. Equations 1.21-1.22 imply a strict hierarchy on the mass spectrum, in particular:

$$\begin{aligned} M_H &> max(M_A, M_Z) \\ M_{H_\pm} &> M_W \\ M_h \le min(M_Z, M_A|cos2\beta|) &\le M_Z \end{aligned} \tag{1.25}$$

Thus at tree leel, the CP-even h boson mass is bound to be lighter than the Z boson.

This simple pattern changes once radiative corrections are included. At one loop level, 1.21 is still valid , while for the mass of the CP-even bosons one has:

$$M_{h,H}^2 = \frac{1}{2}(M_A^2 + M_Z^2 + \epsilon)[1 \mp \sqrt{1 - 4\frac{M_A^2 M_Z^2 cos^2 2\beta + \epsilon(M_A^2 sin^2\beta + M_Z^2 cos2\beta)}{(M_A^2 + M_Z^2 + \epsilon)^2}}] \tag{1.26}$$

where ϵ is defined as:

$$\epsilon = \frac{3\bar{m}_t^4}{2\pi^2 v^2 sin^2\beta}\big[log\frac{M_S^2}{\bar{m}_t^2} + \frac{X_t^2}{M_S^2}\big(1 - \frac{X_t^2}{12M_S^2}\big)\big], \tag{1.27}$$

and X_t is the so-called stop mixing parameter.

The radiative corrections to the Higgs mass can lead to a mass value of at most 130 GeV. The observed mass of 125 GeV is very close to the upper limit. It follows that if we want to interpret the observed boson as the lightest Higgs state in the MSSM, we need to maximize the tree level h mass requiring $M_A >> M_Z$ and large $tan\beta$ values, and to be in the so-called maximal mixing scenario $X_t = \sqrt{6}M_S$ with the highest possible values of the SUSY breaking scale to maximize the radiative corrections.

Table 1.5: The couplings of the neutral MSSM Higgs bosons, collectively denoted by Φ, to fermions and gauge bosons when normalized to the SM Higgs couplings [28].

Φ	$g_{\Phi u\bar{u}}$	$g_{\Phi d\bar{d}}$	$g_{\Phi VV}$	$g_{\Phi AZ}/g_{\Phi H^+W^-}$
h	$cos\alpha/sin\beta$	$-sin\alpha/cos\beta$	$sin(\beta-\alpha)$	$cos(\beta-\alpha)$
H	$sin\alpha/sin\beta$	$cos\alpha/cos\beta$	$cos(\beta-\alpha)$	$sin(\beta-\alpha)$
A	$cot\beta$	$tan\beta$	0	0/

Couplings and decay pattern for the MSSM Higgs bosons The dependence of the neutral Higgs couplings to the α and $tan\beta$ parameters is reported in Table 1.5. Concerning the charged Higgs boson, the coupling to fermions is given by:

$$\begin{aligned} g_{H^+u\bar{d}} &= -\frac{i}{\sqrt{2}v}V_{ud}[m_d tan\beta(1-\gamma_5) + m_u cot\beta(1+\gamma_5)] \\ g_{H^+l\bar{\nu}_l} &= -\frac{i}{\sqrt{2}v}m_l tan\beta(1+\gamma_5) \end{aligned} \tag{1.28}$$

It follows that in the high $tan\beta$ regime, the non SM-like Higges couple strongly to b quarks and τ leptons, while the couplings to the top quark are suppressed. One is left with a SM-like light Higgs h plus three Higgs states $A, H, H^{\pm}$ almost degenerate in mass. In particular A and H would have the same couplings and branching ratios. The charged Higgs particles decay into $\tau\nu_\tau$ final states with a branching fraction of almost 100% for $H^{\pm}$ masses below the tb threshold, and a branching fraction of only 10% for $H^{\pm}$ masses above this threshold. The dominant channel in the latter case is $H^{\pm} \rightarrow t\bar{b}$ which occurs with a $\approx 90\%$ probability.

These prediction have driven the searches for charged Higgs states at colliders, that have been mostly focused on the $t\bar{b}$ and $\tau\nu$ final states.

1.5 Searches for Charged Higgs boson $H^{\pm}$

1.5.1 Charged Higgs production modes

At e^+e^- colliders charged Higgs bosons can be pair-produced in the s-channnel via γ or Z boson exchange. This process is dominant in the LEP centre-of-mass energies range *i.e.* up to 209 GeV. At higher centre-of-mass energies, other processes can play an important role such as the production in top quark decays via $t \rightarrow b + H^+$ is $m_{H^{\pm}} < m_t - m_b$ or via the one loop process $e^+e^- \rightarrow W^{\pm}H^{\mp}$.

At hadron colliders, charged Higgs bosons can be produced in decays of the top quark $t \rightarrow b + H^+$ if $m_{H^{\pm}} < m_t - m_b$. The production of top-quark pairs results from $q\bar{q}$ annihilation and gg fusion, with the former (latter) process being largely dominant at the Tevatron (LHC).

The cross section times branching ratio $\sigma(pp \to t\bar{t}) \times BR(t \to bH^+)$ for the MSSM scenario is shown in Figure 1.13 as a function of the $H^\pm$ mass for different values of $tan\beta$. As can be seen, for small ($\lesssim 3$) or large ($\gtrsim 30$) values of $tan\beta$, the production rates are huge if the charged Higgs boson is light enough. For intermediate values ($tan\beta \sim 10$) the $H^\pm tb$ coupling is not enough enhanced and the rates are rather small. The rate for H^- is the same and the cross sections for the two process have to be added. In principle, if the branching ratio was larger than 1%, the decay to bH^+ would lead to more than 10^6 charged Higgs particles in 100 fb^{-1} of integrated luminosity at the nominal LHC.

If $m_{H^\pm} > m_t - m_b$, then charged Higgs boson production occurs mainly through radiation from a third generation quark. Charged Higgs bosons may also be produced singly in association with a top quark via the $2 \to 3$ partonic processes $gg, q\bar{q} \to t\bar{b}H^-$ [1] (Figure 1.14). The cross sections for these processes are shown in figure 1.15.

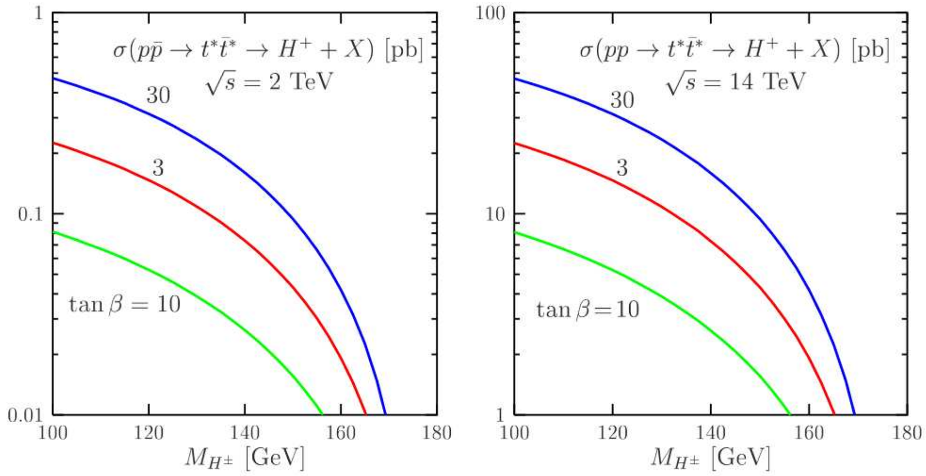


Figure 1.13: Production cross sections for the charged Higgs boson from top decays $\sigma(pp \to t\bar{t}) \times BR(t \to bH^+)$ as functions of the H^+ mass for different values of $tan\beta$ at the Tevatron (left) and the LHC (right) [29].

1.5.2 Charged Higgs searches before LHC

Charged Higgs bosons have been searched for at the LEP, where the combined data of the four experiments, ALEPH, DELPHI, L3 and OPAL, were sensitive to masses of up to about 90 GeV in two decay channels, the $\tau\nu$ ad $c\bar{s}$ [30]. The exclusion limit independent of the admixture of the two above mentioned branching fractions was 78.6 GeV.

The CDF and D0 collaborations at Tevatron have also searched for charged Higgs bosons in top quark decays with subsequent decays to $\tau\nu$ or to $c\bar{s}$ in a

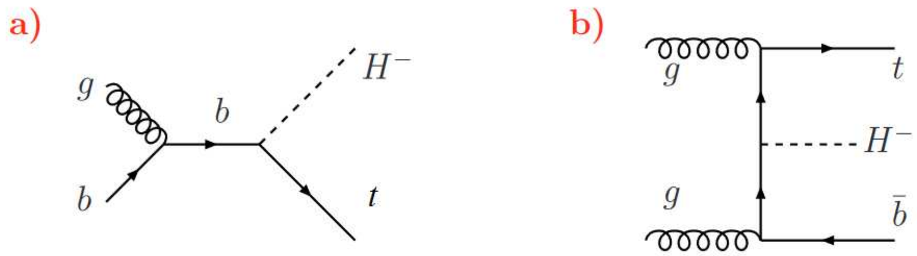


Figure 1.14: Feynman diagrams for the processes $bg \to H^{-}t$ and $gg \to t\bar{b}H^{-}$

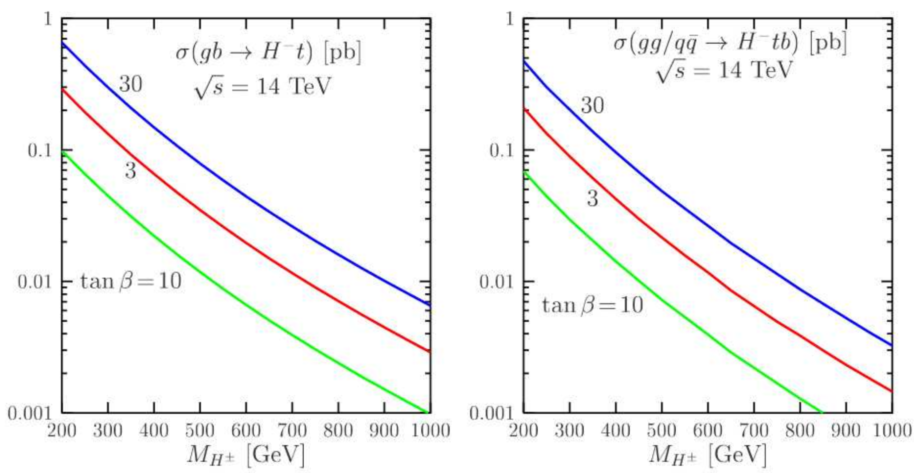


Figure 1.15: The production cross sections for the charged Higgs boson at the LHC as functions of the $H^{\pm}$ mass for different values of $tan\beta$ in the processes $bg \to H^{-}t$ (left) and $gg \to t\bar{b}H^{-}$ (right) [29].

complementary energy range. The limits on BR($t \rightarrow H^+b$)) from CDF and D0 are about 20% in a mass window ranging form 90 GeV to 160 GeV and assuming a branching fraction of 100% on each specific state [31, 32, 33].

1.5.3 Charged Higgs searches at the LHC

At the LHC, the sensitive mass domain is much larger and the variety of search channels wider.

The CMS collaboration has exploited the full data sample collected in proton-proton collisions at $\sqrt{s} = 8$ TeV corresponding to the luminosity of 19.7 fb^{-1} to search for charged Higgs bosons in top quark decays for $m_{H^+} < m_t - m_b$ and in the direct production $pp \rightarrow t\bar{b}H^+$ for $m_{H^+} > m_t - m_b$ [34]. The $H^+ \rightarrow \tau\nu$ and $H^+ \rightarrow tb$ decay modes in the final states $\tau_h + jets$, $\mu\tau_h$, $l + jets$ and ll ($l = e, \mu$) have been considered in the search. No signal has been observed and 95% confidence level upper limits have been set on the charged Higgs production. A model-independent upper limit on the product branching ratio $BR(t \rightarrow bH^+)BR(H^+ \rightarrow \tau\nu)$ =1.2-0.15% is obtained in the mass range m_{H^+}=80-160 GeV, while the upper limit on the cross section times branching fraction $\sigma(pp \rightarrow t(b)H^+)B(H^+ \rightarrow \tau\nu)$=0.38-0.25 pb is set in the mass range m_{H^+}=180-600 GeV. Assuming $BR(H^+ \rightarrow tb) = 1$, an upper limit on $\sigma(pp \rightarrow t(b)H^+)$ of 2.0-0.13 pb is set for m_{H^+}=180-600 GeV.

In addition, a search for a light charged Higgs boson originating from the decay of a top quark and subsequently decaying to a charm quark and a strange antiquark was performed [35]. The search was done in semileptonic $t\bar{t}$ events in the final state comprising an isolated lepton, at least four jets and large missing transverse energy. No significant deviation was observed with respect to SM predictions, and an upper limit on the branching fraction $B(t \rightarrow H^+b)$ ranging from 1.2 to 6.5% was set for a charged Higgs with mass between 90 and 160 GeV, under the assumption that $BR(H^+ \rightarrow c\bar{s})$=100%.

Similar searches using the 8 TeV dataset have been performed by the ATLAS collaboration. The search for a charged Higgs boson in the $\tau\nu$ fully hadronic final state [36] provided 95% confidence level upper limits on the product branching ratios $BR(t \rightarrow bH^+)BR(H^+ \rightarrow \tau\nu)$ between 0.23% and 1.3% for the charged Higgs boson mass range 80-160 GeV. In the mass range 180-1000 GeV an upper limit on the production cross section times branching ratio $\sigma(pp \rightarrow t(b)H^+)B(H^+ \rightarrow \tau\nu)$ between 0.76 pb and 4.5 fb was found.

For the tb final state, the production of a charged Higgs boson in association with top quark was explored in the mass range 200 to 600 GeV using multi-jet final states with one electron or muon [37]. Upper limits ranging from 7 pb to 0.25 pb were set on the $gb \rightarrow tH^+$ production cross section times the branching fraction $B(H^+ \rightarrow tb)$. Additionally, the complementary s-channel production, $q\bar{q} \rightarrow H^+$ was investigated. Final state with one electron or muon were relevant for H^+ masses from 0.4 to 2.0 TeV, whereas the all-hadronic final state was considered for the range 1.5 to 3.0 TeV. Upper limits of 6-0.09 pb were placed on the $q\bar{q} \rightarrow H^+$ cross section times the branching fraction

$B(H^+ \to tb)$.

An early search for charged Higgs bosons produced in association with a top quark and decaying to $\tau\nu$ using a 3.2 fb^{-1} dataset at $\sqrt{s} = 13$ TeV was presented by ATLAS [38]. The final state with both the τ lepton and the top quark decaying hadronically was considered. An upper limit on the cross section times branching fraction $\sigma(pp \to t(b)H^+)B(H^+ \to \tau\nu)$ ranging from 1.9 pb and 15 fb has been set in the mass range m_{H^+}=200-2000 GeV.

The $c\bar{s}$ final state was considered by ATLAS only with the $\sqrt{s} = 7$ TeV 2010 dataset, corresponding to an integrated luminosity of 35 pb^{-1} [39]. Upper limits between 25% and 14% were set on the branching ratio $B(t \to H^+b)$.

1.6 Charged Higgs to $c\bar{b}$

Though the final state into $c\bar{b}$ has so far received less attention than the other decay modes described before, it can be particularly interesting in the frame of the low $tan\beta$ region of the MSSM and in the Flipped-2HDM.

1.6.1 Charged Higgs to $c\bar{b}$ in the low $tan\beta$ region

The discussion presented in 1.4, leading to the requirement of high $tan\beta$ values to accomodate a 125 GeV Higgs boson in the frame of the MSSM, holds if the SUSY breaking scale is at the order of $M_S \approx 1TeV$, which is the natural choice in order to exploit SUSY to solve the hierarchy problem. On the other hand, in light of the fact that no evidence of any sfermions was found so far at the LHC, also theories with higher breaking scale were proposed, such as split SUSY [40] or high-scale SUSY [41]. These models can in principle accomodate a 125 GeV h boson, provided suitable values are chosen for the M_S and $tan\beta$ parameter. The contours for the allowed regions in the $[tan\beta, M_S]$ parameter space are shown in figures 1.16, under different assuptions for the SM-like Higgs mass M_h. As M_S grows, lower and lower values of $tan\beta$ can be reopened.

It is then possible to distinguish two regimes, approximately corresponding to high ($\gtrsim 3$) or low ($\lesssim 3$) values of $tan\beta$ respectively. The first one is the most natural and the most extensively studied in the MSSM, though also the latter can be envisaged for high values of the M_S scale. Since the couplings of the Higgs states are dependent on $tan\beta$, these regimes give rise to a different phenomenology for the extended Higgs sector.

In the low $tan\beta$ scenario, the decay pattern becomes more complex. The branching fractions for the $H/A/H^\pm$ decays are shown in Figure 1.17 as a function of their masses at $tan\beta = 2.5$. For charged Higgs boson, the tb final state is again dominant above the tb mass threshold. For light mass values, although the $\tau\nu_\tau$ final state is still enhanced, other channels have sizeble branching fractions. The $H^\pm \to c\bar{s}$ and $H^\pm \to c\bar{b}$ have similar branching ratio to the level of percent, that can grow to $\approx 10\%$ as $tan\beta$ approaches 1.

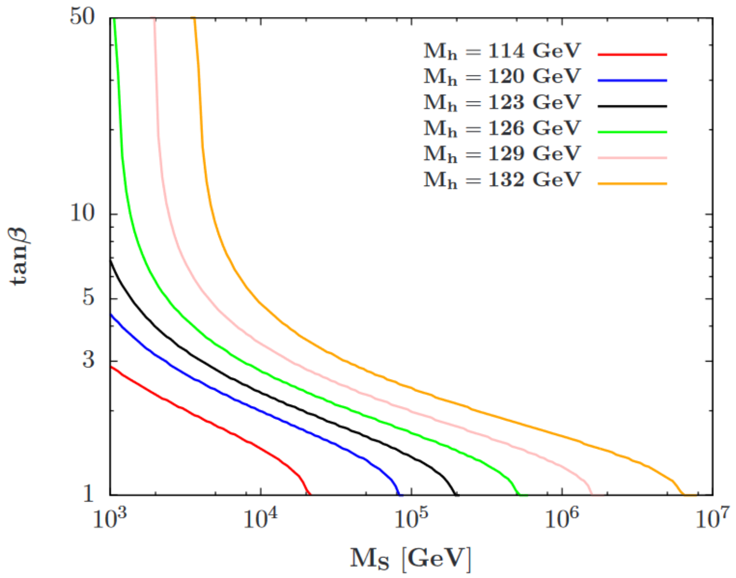


Figure 1.16: Contours for fixed values of M_h in the $[tan\beta, M_S]$ plane in the limit $M_A >> M_Z$ [28]

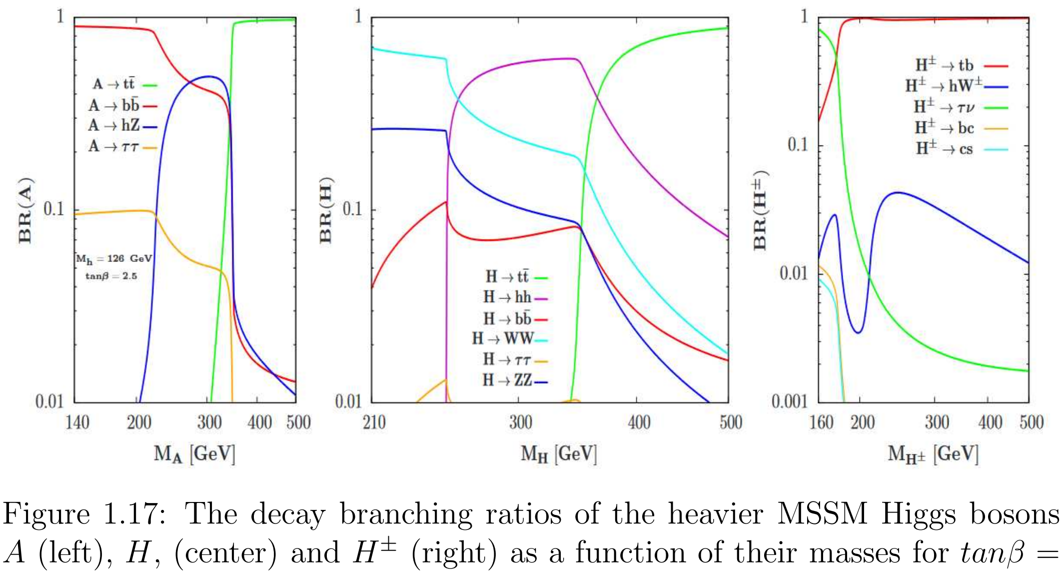


Figure 1.17: The decay branching ratios of the heavier MSSM Higgs bosons A (left), H, (center) and $H^{\pm}$ (right) as a function of their masses for $tan\beta = 2.5$ [28].

1.6.2 Charged Higgs to $c\bar{b}$ in the Flipped-2HDM

In the Flipped-2HDM, one doublet gives mass to up-type quarks and charged leptons and the other doublet to down type quarks. This model is particularly interesting for charged Higgs, since below the $H^+ \to tb$ its decay pattern is remarkably different with respect to the type-II scenario presented in previous paragraphs. The MSSM couplings in equations 1.28 are replaced by [42]:

$$\begin{aligned} g_{H^+u\bar{d}} &= -\frac{i}{\sqrt{2}v} V_{ud}[m_d tan\beta(1+\gamma_5) + m_u cot\beta(1-\gamma_5)] \\ g_{H^+l\bar{\nu}_l} &= -\frac{i}{\sqrt{2}v} m_l cot\beta(1+\gamma_5) \end{aligned} \tag{1.29}$$

The consequence is that the usual $\tau\nu$ decay is replaced by decays to quarks ($c\bar{b}$ and $c\bar{s}$) in most of the parameter space. The branching fractions of the charged Higgs in the flipped model are shown as a function of M_{H^+} in Figures 1.18-1.19 for different values of $tan\beta$. For comparison, the corresponding branching fractions in the type-II model are also shown for $tan\beta \neq 1$. For $tan\beta = 5$ decays to $\tau\nu$ reach at most $\approx$5% in the flipped model, while they dominate below the tb threshold in the type-II model. For $tan\beta = 50$, the branching fraction to leptons is below 10^{-4}. Instead, the dominant decay mode for $tan\beta \gtrsim 3$ is into $c\bar{b}$ with a branching fraction of about 2/3, followed by $c\bar{s}$ with a branching fraction of about 1/3. The relative strenght of these two decays at moderate to large $tan\beta$ is controlled by the $V_{cb}m_b/V_{cs}m_s$, that is greater than 1.

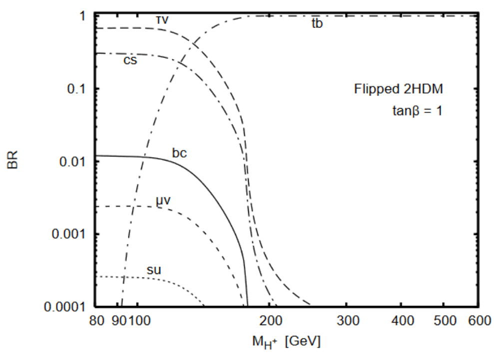


Figure 1.18: Charged Higgs branching ratios as a function of $M_{h^\pm}$ for $tan\beta = 1$ in the flipped 2HDM. Branching ratio in the type-II 2HDM are identical [42].

1.6.3 The first search for a light charged Higgs to $c\bar{b}$

Recently, the first attept to search for a charged Higgs in the $c\bar{b}$ has been made public by CMS [43]. I personally contributed to this study, that is the topic of my thesis. The analysis consisted in looking for semileptonic $t\bar{t}$

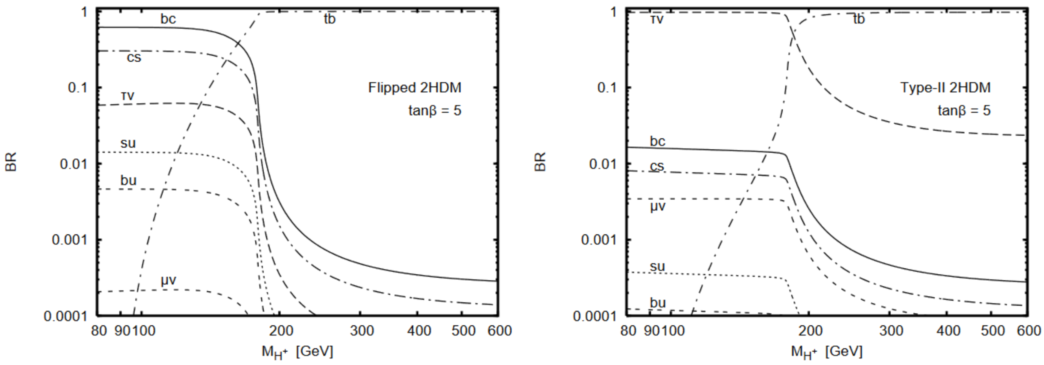


Figure 1.19: Charged Higgs branching ratios as a function of $M_{h^{\pm}}$ for $tan\beta = 5$ in the flipped (left) and in the type-II (right) 2HDM).

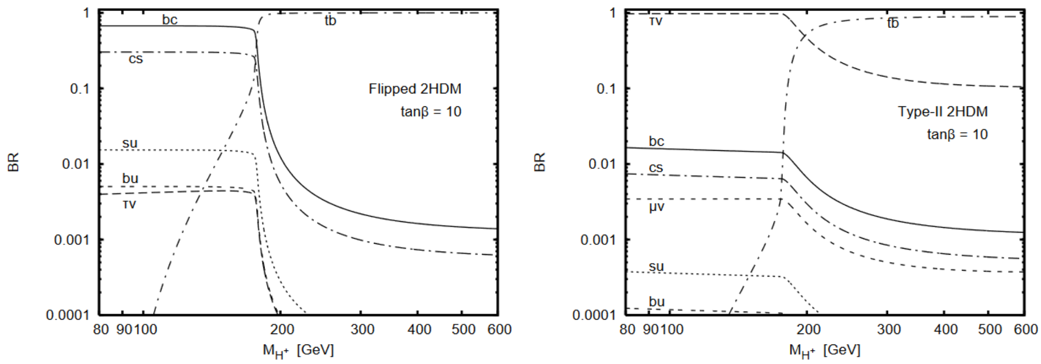


Figure 1.20: As in Figure 1.19, for $tan\beta = 10$

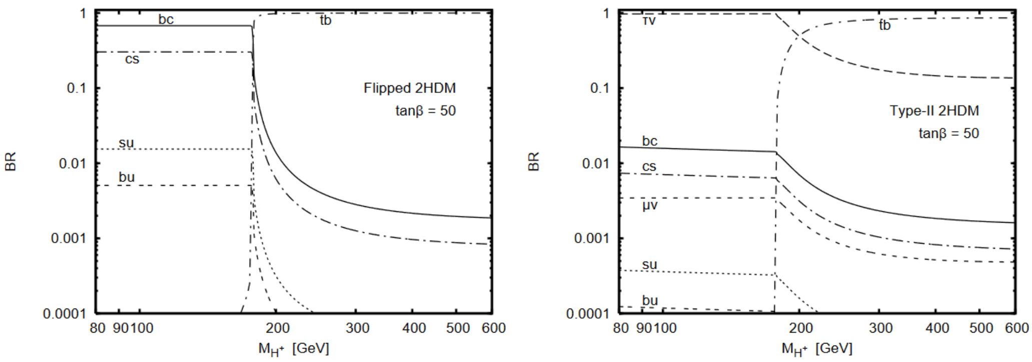


Figure 1.21: As in Figure 1.19, for $tan\beta = 50$

events with one top decaying to H^+b instead of Wb and subsequently going to $c\bar{b}$, while the other top decays leptonically in the electron or muon final state ($\bar{t} \rightarrow W^-\bar{b} \rightarrow l\bar{\nu}\bar{b}$). The full Run I dataset collected in proton-proton collisions at a centre-of-mass energy of 8 TeV was used, corresponding to an integrated luminosity of 19.7 fb^{-1}. A first result, presented at the International Conference of High Energy Physics held in Chicago in August 2016, shows no signal for the presence of a charged Higgs boson and upper limits ranging from 1.1 to 0.4% were set on the branching fraction of the top quark to H^+b in the 90-150 GeV mass range.

Chapter 2

The CMS experiment

The Compact Muon Solenoid (CMS) experiment is one of two general-purpose particle physics detectors at the Large Hadron Collider LHC. This chapter introduces LHC and presents the general design of CMS and its subdetectors.

2.1 The LHC collider

The Large Hadron Collider (LHC) [44] is a proton-proton superconducting accelerator and collider installed in the existing 26.7 km tunnel that was constructed between 1984 and 1989 to host the Large Electron Positron collider (LEP).

The LHC lies between 45 m and 170 m below the surface and is divided in eight arcs and eight straight sections, of which four house equipment needed for the accelerator and the other four contain the interaction points where the two beams are brought into collision in the four main experiments. ATLAS (A ToroidaL ApparatuS) [45] and CMS (Compact Muon Solenoid) [46] are two big independently designed general-purpose detectors designed to investigate the largest range of physics possible. ALICE (A Large Ion Collider Experiment) [47] and LHCb (Large Hadron Collider beauty experiment) [48] are medium-size experiments dedicated to specific phenomena.

In a circular collider of radius R, the energy loss per turn due to synchrotron radiation is proportional to $(E/m)^4/R$, where E and m are respectively the energy and mass of the particles accelerated. Protons, due to their higher mass with respect to electrons, imply a smaller energy loss for synchrotron radiation.

The high beam intensity required by the experiments excludes the use of antiproton beams, and hence excludes the particle-antiparticle collider configuration of a common vacuum and magnet system for both circulating beams, as used for example at Tevatron. Colliding two counter-rotating proton beams requires opposite magnetic dipole fields in both rings. The LHC is therefore designed as a proton-proton collider with separate magnetic fields and vacuum chambers in the main arcs and with common sections only at the intersection

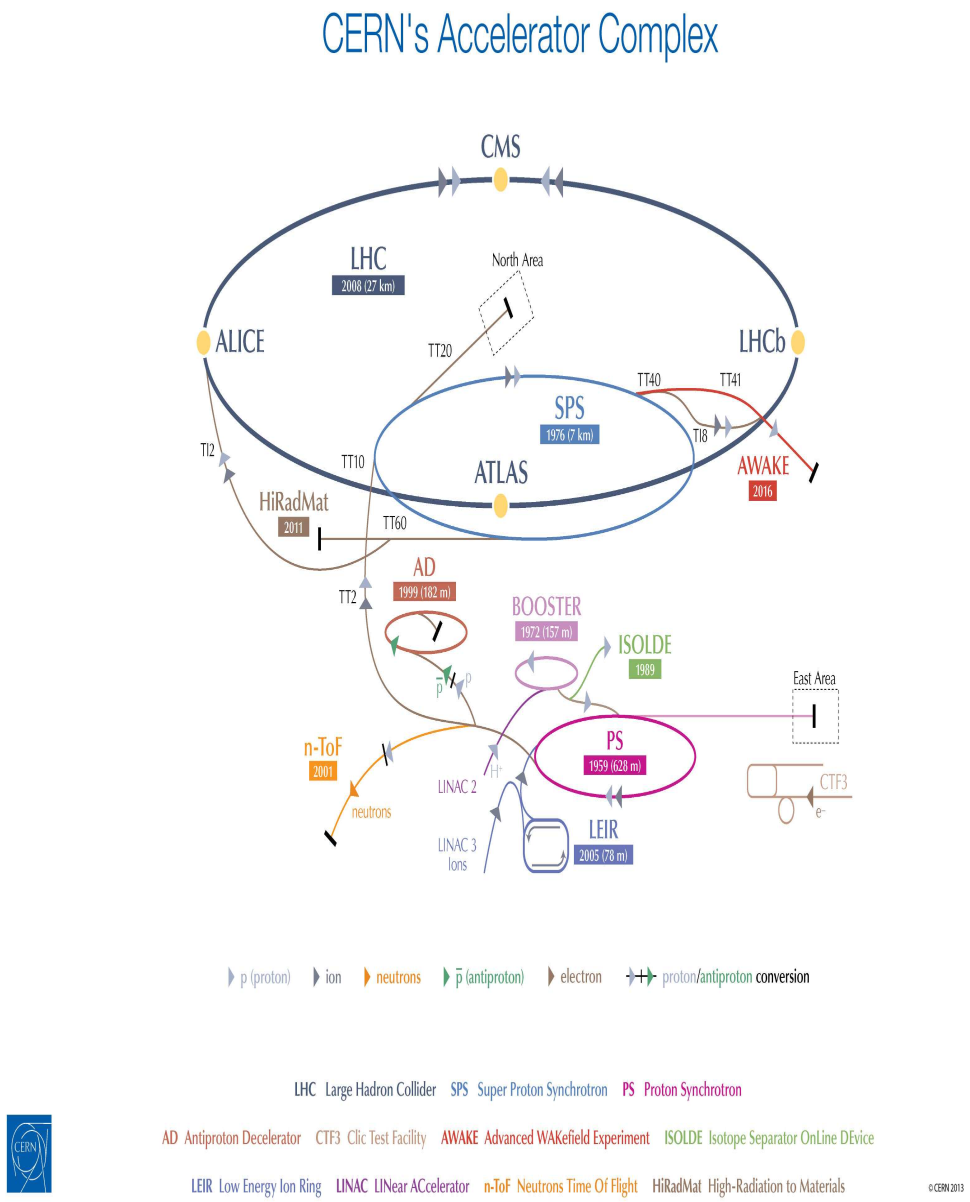


Figure 2.1: The CERN accelerator complex.

regions where the experimental detectors are located.

The existing CERN infrastructure, shown in Figure 2.1, is used for injecting the protons into LHC (Linac, Booster, Proton Synchroton (PS), Super Proton Synchroton (SPS)). The SPS accelerates protons to an energy of 450 GeV, and the remaining acceleration is done by the LHC during the first 20 minutes after beam injections.

The machine has 1232 dipole magnets and is designed to have an energy per proton beam of 7 TeV, which results in a center-of-mass energy of $\sqrt{s} =$ 14 TeV, 2808 bunches per ring, and a 25 ns time between two bunch crossings in an impact point (IP), which spaces the bunches about 7.5 m apart along the beam axis.

In the years 2010 and 2011 the LHC was operated with proton beam energies of 3.5 TeV. In 2012, the beam energy of 4 TeV was reached, resulting in a proton-proton (pp) center-of-mass energy of 8 TeV and a bunch spacing of 50 ns. This LHC running period is called Run-1. In spring 2013, the LHC was shut down for about 2 years to allow consolidation and upgrade of numerous machine systems.

In July 2015 LHC started to collide proton beams with a center-of-mass energy of 13 TeV (LHC running period called Run-2). After a short period of 50 ns operation (Run2015B), the machine collected data with a bunch spacing of 25 ns (Run2015C and D, Run2016A to G).

2.1.1 Luminosity and design conditions

At the LHC, the number of events per second generated in the collisions is proportional to the cross section σ_{event}:

$$R_{event} = \sigma_{event} \times \mathcal{L}, \tag{2.1}$$

where $\mathcal{L}$ is the machine instantaneous luminosity, defined as the number of collisions per unit time and cross-sectional area of the beams:

$$\mathcal{L} = \frac{N_1 N_2 n_b f_{rev}}{A}. \tag{2.2}$$

N_1 and N_2 are the number of particles in the two colliding bunches, A is the overlap area of the two bunches transverse to the beam, n_b is the number of bunches in one beam, and f_{rev} is the revolution frequency of one bunch (with a design value of 11245 Hz). At the LHC proton-proton collisions $N_1 = N_2 = N_p$, and, since the area of overlap is difficult to measure directly in an accelerator, for a Gaussian beam distribution $\mathcal{L}$ can be written as :

$$\mathcal{L} = N_p^2 n_b f_{rev} \frac{\gamma}{4\pi\epsilon_n\beta^*} F \tag{2.3}$$

where γ is the relativistic Lorentz factor, ϵ_n is the normalized transverse beam emittance (with a design value of 3.75 μm), β^* is the so called betatron function at the IP [49], and F is the geometric luminosity reduction factor due to the crossing angle at the IP.

The maximum number of bunches per beam and the revolution frequency are defined by the circumference of the LHC. In order to get as many events of interest as possible, one can either increase the number of particles in a bunch or focus the two beams on a smaller area for the interaction.

The values for the LHC machine parameters are listed in Table 2.1.

		Design	Run 2015
Centre-of-mass energy [TeV]	$\sqrt{s}$	14	13
Luminosity [cm^{-2}s^{-1}]	$\mathcal{L}$	10^{34}	10^{33}
Num. of bunches	n_b	2808	2244
Bunch spacing [ns]		25	25
Num. of protons/bunch	N_p	1.15×10^{11}	1.1×10^{11}
Norm. Rms. Emittance [μm]	ϵ_n	3.75	3.5
β^* **at the IP [m]**	β^*	0.55	0.8

Table 2.1: Machine parameters of LHC

During collisions, the number of particles in a bunch, and thus also the instantaneous luminosity, decreases exponentially from the initial peak luminosity. The peak luminosity of the LHC in 2016 is shown in Figure 2.2. In general, after about ten hours, the instantaneous luminosity has decreased so much that it is more efficient to abort the fill and refill the machine with new beams.

The integrated luminosity cumulated for all the pp fills collected during 2016 is shown in Figure 2.3

2.1.2 Proton-proton interactions

Several independent proton-proton interactions can take place in a bunch crossing in the interaction point. The interaction of two protons forms a primary vertex, from which the particles, that were created in the interaction, originate. The number of primary vertices created on average depends on the beam parameters, e.g. how many particles are in a bunch and how small is the focusing area. In 2012 this number has been measured by the CMS experiment and corresponds to, on average, 21 interactions per bunch crossing, as shown in Figure 2.4. The presence of many primary vertices per bunch crossing is a challenge for the event reconstruction, since the particles originating from different primary vertices can be superimposed in the detector. Interactions besides the interaction of interest are referred to as *pileup* interactions.

Different kind of processes can take place in an event.

In a large distance collision only a small momentum is transferred and particle scattering at large angle is suppressed. The final state particles have

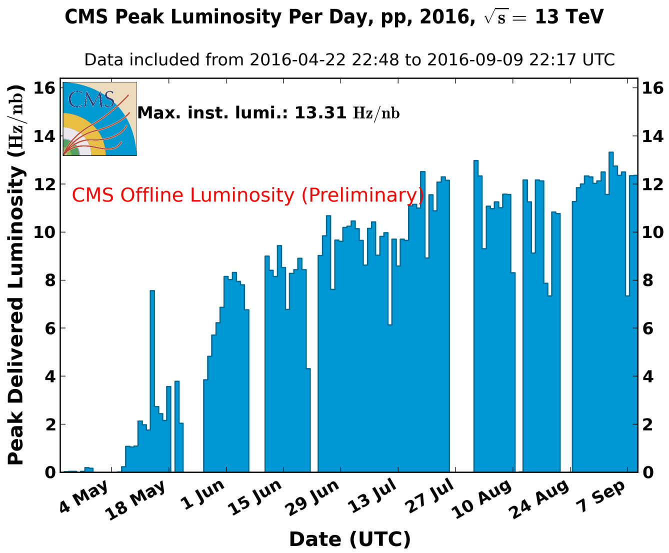


Figure 2.2: Peak delivered luminosity per day for 2016 as measured by the CMS experiment.

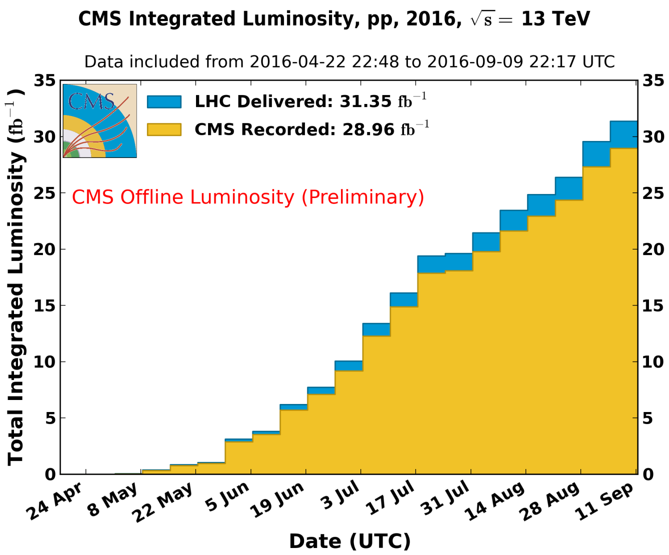


Figure 2.3: Cumulative offline luminosity versus day delivered to (blue), and recorded by CMS (orange) during stable beams and for p-p collisions at 13 TeV centre-of-mass energy in 2016 [50].

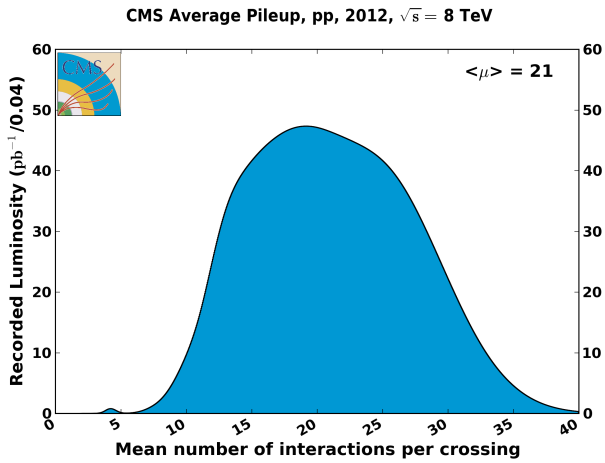


Figure 2.4: Number of reconstructed vertices per event for p-p collisions at 8 TeV centre-of-mass energy and 50 ns bunch spacing in 2012.

small transverse momentum ($\sim 10^2$ MeV), so that most of them escape down the beam pipe.

However, then two protons collide, two of their partons (quarks and gluons) can also take part in a hard interaction with high transferred p_T. The effective centre-of-mass energy of the hard scattering, $\sqrt{\hat{s}}$, is proportional to the fractional energies x_a and x_b carried by the two interacting partons:

$$\sqrt{\hat{s}} = \sqrt{x_a x_b s} \ , \tag{2.4}$$

where $\sqrt{s}$ is the centre-of-mass energy of the proton beams.

The probability density $f_p(x_p, Q^2)$ to find a parton p, with the fraction x of the longitudinal proton momentum in the proton-proton center-of- mass frame, depends on the squared four-momentum transfer Q^2 between the partons of the collision, and is described by the Parton Distribution Function (PDF). PDFs are different for gluons, u and d valence quarks and low-momentum sea quark-antiquark pairs of all flavours and depend on the energy scale at which the interaction between the partons takes place; for higher exchanged momenta a shorter distance scale is probed and the contribution of gluons and sea quarks becomes higher.

PDFs are measured in Deep Inelastic Scattering (DIS) experiments of leptons on hadrons and different models are available such as CTEQ [51, 52], MSTW [53], or NNPDF [54].

An example for parton distribution functions is shown in Figure 2.5 for two different values of the invariant momentum transfer Q^2.

To probe physics at a certain energy scale, the value for Q^2 has to be taken in the range of the squared effective centre-of-mass energy $\hat{s}^2$ of the hard scattering which corresponds to the squared invariant mass M^2 of the system.

Since the two partons interact with unknown energies, the total energy of an

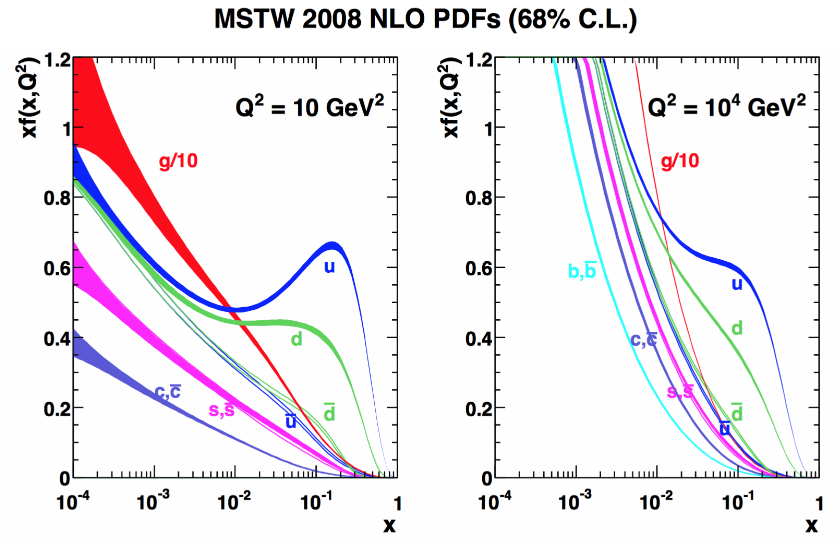


Figure 2.5: Parton density functions, including the one sigma uncertainty bands, for the partons in a proton for two different invariant momentum transfers Q^2=20 GeV^2 (left) and Q^2=10^4 GeV^2 (right) [53].

event is unknown, because the proton remnants, that carry a sizable fraction of the proton energy, are scattered at small angles and are predominantly lost in the beam pipe, escaping detection. For this reason it is not possible to define the total and missing energy of the event, but only the total and missing transverse energies (in the plane transverse to the beams).

Secondary particles created in an hard interaction, which in turn can decay, form the final state of an event that can be detected. The rate of hard interactions, though, is several orders of magnitude lower than that of soft interactions. The probability for one particular hard interaction in an event, as expressed in Equation (2.1), depends on the cross section of that particular process. Figure 2.6 shows the cross section for different SM processes as a function of the centre-of-mass energies in pp collisions.

Before the two partons interact with each other they can radiate other partons. Similar to this process also the decay products of the hard interaction can radiate partons or photons. This radiation of particles is called *initial state radiation* (ISR) when it happens before the hard interaction, and *final state radiation* (FSR) if it occurs with the decay products of the hard interaction. When quarks and gluons are involved in the ISR and FSR, one speaks also of *parton showering*.

If the final state of a hard interaction contains particles that carry a colour charge like e.g. quarks, they have to form new particles in order to become colour neutral. This process is called *hadronisation* and results in showers of particles that form a cone along the initial particles direction and are called *jets*. The exception to this is the top quark, which has a lifetime shorter than

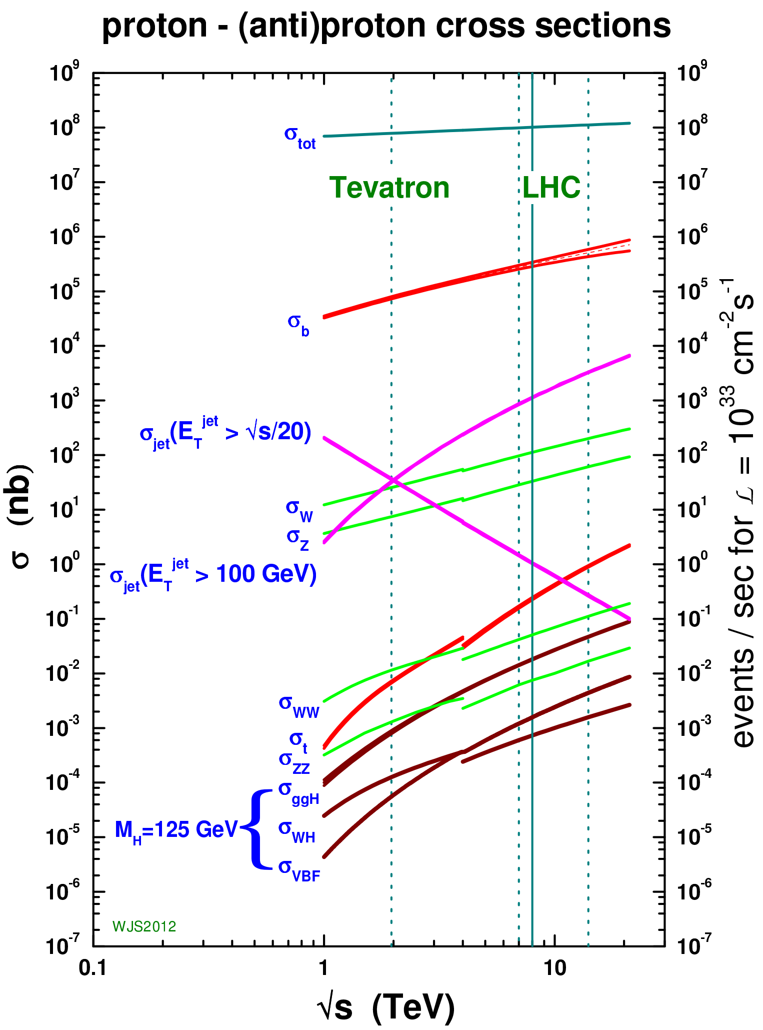


Figure 2.6: Cross section of SM processes as a function of the center-of-mass energy of proton-(anti)proton collisions. The vertical lines mark the center-of-mass energies of the Tevatron and the LHC [55].

the timescale at which the hadronisation takes place, and, therefore, decays before it hadronises. If the particles created in ISR and FSR carry a colour charge they hadronise as well. After the hard interaction, the remnants of the two protons are not colour neutral anymore and have to hadronise as well, forming jets that fly along the beam axis.

Coordinate system

Since the two partons interact with unknown energies, the centre of mass may be boosted along the beam direction. Therefore it is very useful to use experimental quantities that are invariant under such boosts.

We indicate the beam direction as z axis, referred to as longitudinal, and the $x - y$ plane, orthogonal to the beam line, is called transverse plane. Based in these definitions, the momentum of a particle can be divided in two components: the longitudinal momentum p_z and the transverse momentum p_T, defined as

$$p_T = \sqrt{p_x^2 + p_y^2} \; . \tag{2.5}$$

The rapidity is defined as:

$$y = \frac{1}{2} \ln \frac{E + p_z}{E - p_z} \; , \tag{2.6}$$

and has the property of being additive under Lorentz boosts along the z direction, i.e. it is simply shifted by a constant when subjected to such transformations. For ultrarelativistic particles ($p \gg m$) the rapidity is approximated by the pseudorapidity:

$$\eta = - \ln \tan \frac{\theta}{2} \; , \tag{2.7}$$

where θ is the angle between the particle momentum and the z axis. The pseudorapidity can be reconstructed from the measurement of the θ angle and can be also defined for particles whose mass and momentum are not measured.

The origin of the coordinate system of the detector lies in the center at the nominal collision point. The x-axis points radially inward to the center of the LHC ring and the y-axis points vertically upward. The coordinate system is right-handed and so the z-axis points horizontally along the counter clockwise beam direction. Since the products of the collisions will fly outward from the collision point, it makes sense to use cylindrical coordinates for the description (used by reconstruction algorithms) based on the azimuthal angle ϕ, defined as the angle measured from the x-axis in the $x - y$ plane, the radial coordinate r is also measured in the $x - y$ plane and finally, the polar angle θ measured from the z-axis. Instead of the polar angle the pseudorapidity η is used, which is zero in the $x - y$ plane and goes to positive and negative infinity, respectively, towards the positive and negative z-axis. The forward regions of the detector mean regions of higher $|\eta|$, close to the z-axis or about $|\eta| > 3$.

2.2 Components of the CMS detector

The CMS design [46] was driven by the goals of the LHC physics program. The overall layout of CMS, illustrated in Figure 2.7, is typical for a general purpose high energy particle detector. The detector has a cylindrical shape with an overall length of 28.7 m, of which 21.6 m make the main cylinder with a diameter of 15 m, and the rest of the length comes from the forward calorimeter. The total mass is 14000 t. The main detector is made of a central barrel section that is closed with an endcap section on both ends to cover most of the 4π solid angle.

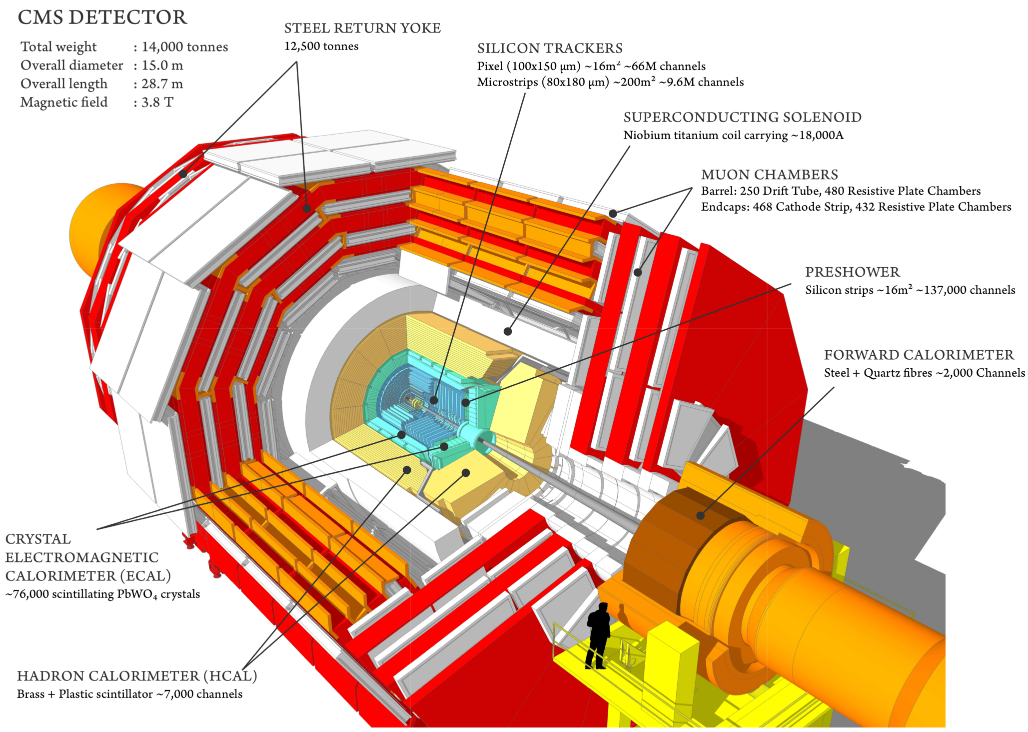


Figure 2.7: Sectional view of the CMS detector. The LHC beams travel in opposite directions along the central axis of the CMS cylinder colliding in the middle of the CMS detector.

The main features of the CMS detector are: the high-field ($\approx$ 3.8 T) solenoid in the barrel part, the full-silicon-based inner tracker, and the homogeneous electromagnetic calorimeter. In particular, the large bending power, needed to measure precisely the momentum of high-energy charged particles, forced a choice of superconducting technology for the magnets. Inside the 6 m diameter bore of the magnet are the silicon tracking system, the Electromagnetic CALorimeter (ECAL), and the Hadronic CALorimeter (HCAL). Outside of the solenoid, the muon tracking system is sandwiched in between the layers of the steel return yoke for the magnetic field. The high magnetic field not only provides a large bending power within a compact spectrometer, but also avoids stringent demands on muon-chambers resolution and alignment. The return field is large enough to saturate 1.5 m of iron, allowing 4 muon stations to be integrated to ensure robustness and full geometric coverage.

The muon spectrometer is composed by 4 stations of Drift Tube (DT) detectors in the barrel region (MB) and 4 stations of Cathode Strip Chambers (CSCs) in the endcaps (ME). Both the barrel and the endcaps muon chambers are coupled to Resistive Plate Chambers (RPCs) to ensure redundancy and robustness to the muon trigger and reconstruction.

2.2.1 Magnet

The solenoid of the CMS detector gives a uniform field in the axial direction, while the flux return is assured by an external iron yoke with three layers, in between which the muon system is installed. The momentum analysis of charged particles is performed by measurement of particles trajectories inside the solenoid and the momentum resolution is given by:

$$\frac{\Delta p_T}{p_T} = \Delta s \frac{8 p_T}{0.3 B R^2} \ , \tag{2.8}$$

where $p = \gamma m v$ is the particle momentum, B is the magnetic induction, s is the sagitta and R is the solenoid radius. Therefore strong field and large radius are an efficient approach to reach optimal momentum resolution: in the case of CMS the solution of a high field within a compact region was adopted.

The superconducting magnet of the CMS detector has a length of 12.5 m and a diameter of the cold bore of 6.3 m. It is made from a 4-layer winding of $NbTi$ cable reinforced with aluminium, weighting a total of 220 t, and kept at a temperature of 4.5 K with liquid helium. It was designed to produce a field of 4 T but operate at a lower field of 3.8 T. The magnetic field is generated by a 18 kA current circulation in the cables. The magnet system stores an energy of 2.5 GJ.

2.2.2 Inner tracker

The inner tracker reconstructs the trajectories of all charged particles in the region $|\eta| < 2.5$ with high momentum resolution and efficiency. It provides a measurement of their impact parameter, and reconstructs secondary vertices. For the tracker a detector technology with high granularity and fast response is required. On the other hand it is important to keep the minimum the amount of material in order to limit multiple scattering, bremsstrahlung, photon conversion, and nuclear interactions, since this detector is the closest to the beam line.

The longitudinal view of one quarter of the tracker is shown in Figure 2.8. The innermost tracker closest to the IP is made of three layers of silicon pixel detectors named Tracker Pixel Barrel (TPB), ranging from 8.8 cm to 20.4 cm diameters, and two wheels of Tracker Pixel Endcap (TPE), covering the pseudorapidity range up to $|\eta| = 2.5$. TPB and TPE contain 48 million and 18 million pixels, respectively. The pixels have a size of 100 $\times$ 150 μm^2.

Thanks to the large Lorentz drift angle in the magnetic field, with a charge interpolation from the analog pulse heights, the measured hit resolution in the TPB is 9.4 μm in the $r - \phi$ coordinate and 20-40 μm in the longitudinal direction. The longitudinal resolution depends on the angle of the track relative to the sensor. For longer clusters, sharing of charge among pixels improves the resolution, with optimal resolution reached for interception angles of $\pm 30°$.

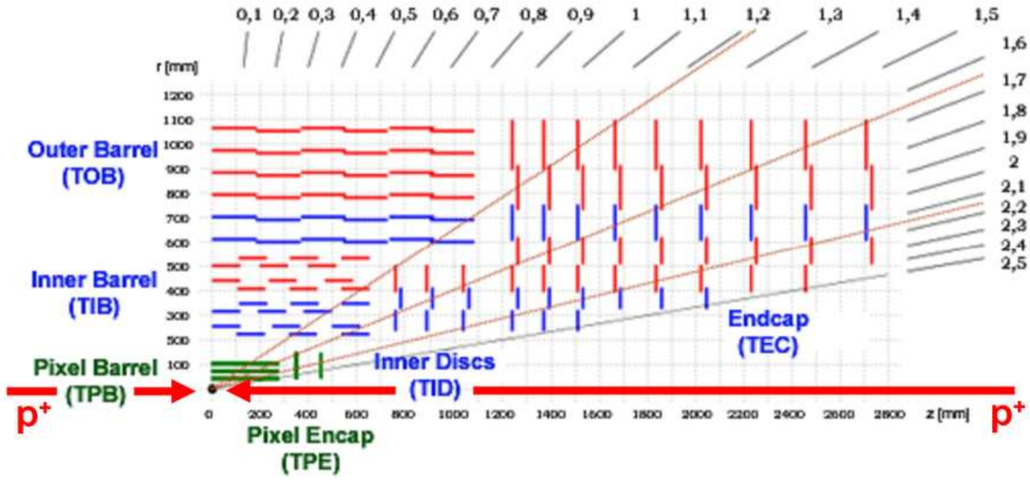


Figure 2.8: Schematic overview of the inner tracker.

The silicon strip tracker is placed outside of the pixel tracker. The barrel part of the strip tracker is divided in the 4-layers of the Tracker Inner Barrel (TIB) and the 6-layers of the Tracker Outer Barrel (TOB). Coverage in the forward region is provided by the 3 Tracker Inner Discs (TID), and the 9 disks of the tracker endcap (TEC) on each side. The pitch of the strips varies between 80 μm in the innermost layers of the TIB, and 183 μm in the outer layers of the TOB. In the disks the pitch varies between 97 μm and 184 μm. Some of the modules are composed by two detectors mounted back-to-back with the strips rotated by 100 mrad. These double-sided (stereo) modules will also provide a measurement in the coordinate orthogonal to the strips. The single point resolution that can be achieved depends strongly on the size of the cluster and on the pitch of the sensor and varies not only as a function of the cluster width, but also as a function of pseudorapidity, as the energy deposited by a charged particle in the silicon depends on the angle at which it crosses the sensor plane. The measured hit resolution in the barrel strip detector varies between ~20 μm and ~30 μm in $r - \phi$ in the TIB and TOB.

2.2.3 Calorimeter

The energy of hadronic jets and electromagnetic cascades induced by photons and electrons is measured by the CMS calorimeter system, which gives also a hermetic coverage to allow missing transverse energy measurement.

Electromagnetic calorimeter

The CMS ECAL is a scintillating crystal calorimeter, with lead-tungstate ($PbWO_4$) as the crystal material. Lead-tungstate is a fast, radiation-hard scintillator characterised by a small Moliere radius (21.9 mm) and a short radiation length (8.9 mm), that allows good shower containment in the limited space available for the ECAL. The scintillation decay time of these crystals is of the same order of magnitude as the LHC bunch crossing time: $\sim 80\%$ of the light is emitted within 25 ns.

The longitudinal view of one quarter of the ECAL is shown in Figure 2.9. The ECAL consists of 61200 crystals in the barrel (EB), covering a pseudora-

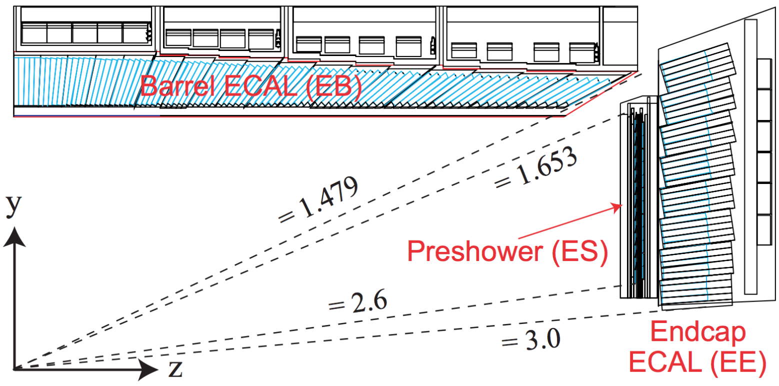


Figure 2.9: Geometry of the ECAL and the preshower detector, for a quadrant of the CMS detector.

pidity range of $|\eta| < 1.5$, and 14648 crystals in the endcaps (EE), which cover a pseudorapidity range of $1.5 < |\eta| < 3.0$. The length of the crystals is 230 mm in the barrel and 220 mm in the endcaps, corresponding to 25.8 and 24.7 radiation lengths respectively. Crystals are trapezoidal, with a square front face of 22×22 mm^2 in the barrel and 30×30 mm^2 in the endcaps, matching the Moliere radius. Scintillator light is collected by silicon avalanche photo-diodes in the case of barrel crystals, and vacuum photo-triodes for endcaps crystals.

In front of the EE there is a pre-shower detector (ES) that covers the pseudorapidity region of $1.65 < |\eta| < 2.6$ and consists of two lead radiators to initiate electromagnetic showers from incoming electrons and photons and two planes of silicon strip detectors to measure the energy and transverse shower profile. The ES is designed to identify photons coming from neutral pion decays and improve the estimation of the direction of photons, to improve the measurement of the two-photon invariant mass.

The energy resolution of a calorimeter can be parametrised as the quadratic sum of a stochastic term ($\sigma_s/\sqrt{E}$), a noise term (σ_n/E) and a constant term (c) [56]:

$$\frac{\sigma_E}{E} = \frac{\sigma_s}{\sqrt{E}} \oplus \frac{\sigma_n}{E} \oplus c. \tag{2.9}$$

The theoretical parametrization of the different contributions as a function of the energy are shown in Figure 2.10. The stochastic term includes the effects of fluctuations in the number of photo-electrons as well as in the shower containment, the noise term consists of electronic noise, digitisation noise, and

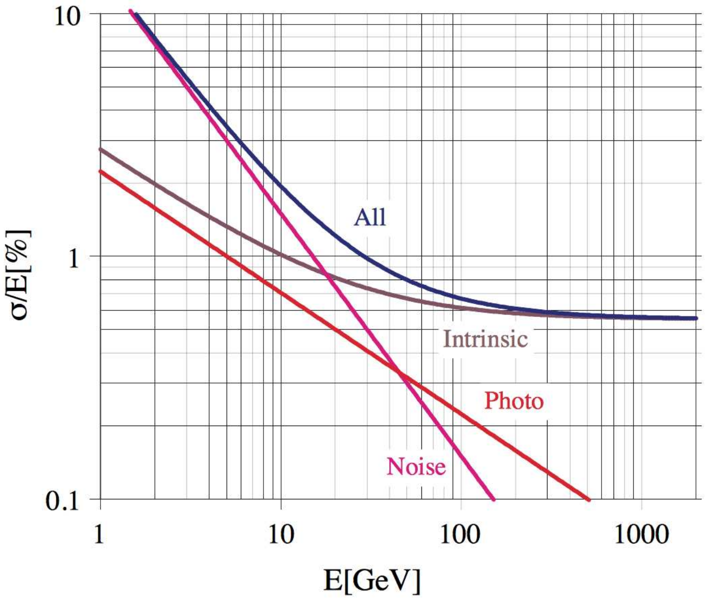


Figure 2.10: Theoretical parametrization of the different contributions to the energy resolution of the ECAL. The noise term contains the contributions from electronic noise and pileup energy. The curve labelled "intrinsic" includes the shower containment and a constant term of 0.55%.

noise from additional pp interactions (pileup), and the constant term related to the calibration of the calorimeter and non-uniformity of the longitudinal light collection of the crystals. The parameters, measured in an electron test beam, for incident electrons of seven energies from 20 to 250 GeV, with a 3×3 crystal configuration, considering E is in GeV, correspond to σ_s =0.028 GeV$^{1/2}$, σ_n=0.12 GeV, and $c = 0.003$ [57].

Hadronic calorimeter

CMS chose as a hadronic calorimeter a sampling calorimeter with brass as absorber, plastic scintillator tiles as active medium, and wavelength shifting fibers to transfer the light to the detector. This absorber material has been chosen as it has a reasonably short interaction length, and is non-magnetic. Most of the HCAL is located inside the bore of the cryostat, and consists in a barrel (HB) that extends to $|\eta| <1.4$, and two endcaps (HE) ranging from $1.3 < |\eta| < 3$. Since the absorber depth of the ECAL barrel and the HCAL barrel in the solenoid is not sufficient to contain the complete particle shower, an additional calorimeter (HO) is placed as a tail catcher outside the cryostat, using it as an additional absorber. In the central ring of the CMS barrel, the HO has two layers, one on each side of the first layer of iron of the yoke, while in the other 4 rings there is only one HO layer. Figure 2.11 shows a quadrant of the HCAL with the segmentation in calorimeter towers.

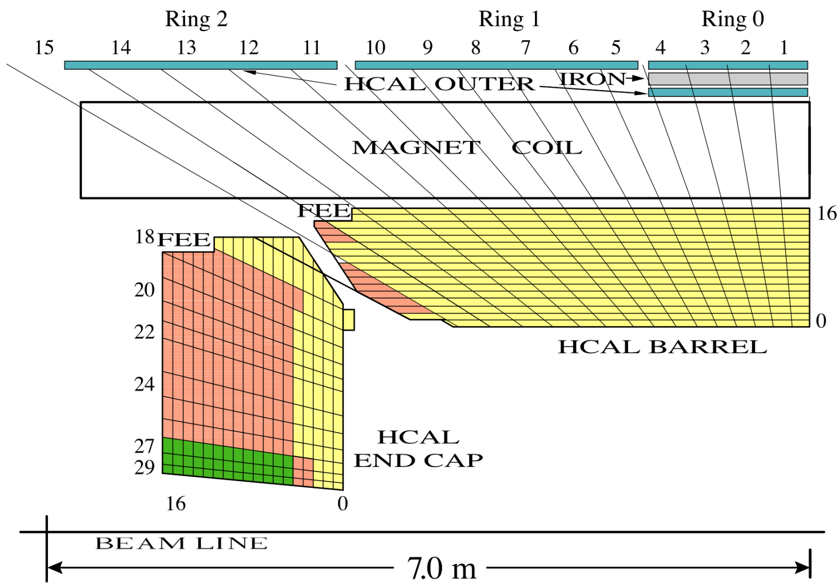


Figure 2.11: A quadrant of the HCAL with the segmentation in calorimeter towers in the $r - z$ plane. The colours indicate the optical grouping of the readout channels [46].

Since the identification of forward jets is very important for the rejection of many backgrounds, the barrel and the endcap parts, which cover up to $|\eta| < 3.0$, are complemented by a very forward calorimeter (HF), placed at ± 11.2 m from the interaction point, which extends the pseudorapidity range of the calorimetry up to $|\eta| < 5.2$. As the particle flux in this very forward region is extremely high, a radiation hard technology, using Cherenkov light in quartz fibers, was chosen with steel as an absorber. The HF detector is also used as a real-time monitor for the luminosity on a bunch-by-bunch basis.

The HCAL baseline single-particle energy resolution is

$$\frac{\sigma_E}{E} = \frac{65\%}{\sqrt{E}} \oplus 5\% \tag{2.10}$$

in the barrel,

$$\frac{\sigma_E}{E} = \frac{83\%}{\sqrt{E}} \oplus 5\% \tag{2.11}$$

in the endcaps, and

$$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 5\% \tag{2.12}$$

in the forward calorimeter (where E is expressed in GeV).

2.2.4 Muon system

In order to provide an independent muon identification, reliable trigger and precise momentum measurement and charge, for muons in a broad momentum range from a few GeV to a few TeV, the muon spectrometer is placed outside the magnet. It consists of four stations of detectors integrated into the iron return yokes so that the 3.8 T magnetic field inside the solenoid and the 1.8 T average return field can be used as bending field.

The muon spectrometer is composed of three independent sub-detectors used both for tracking and for trigger in order to guarantee robustness and redundancy. The layout of the system is presented in Figure 2.12. In the

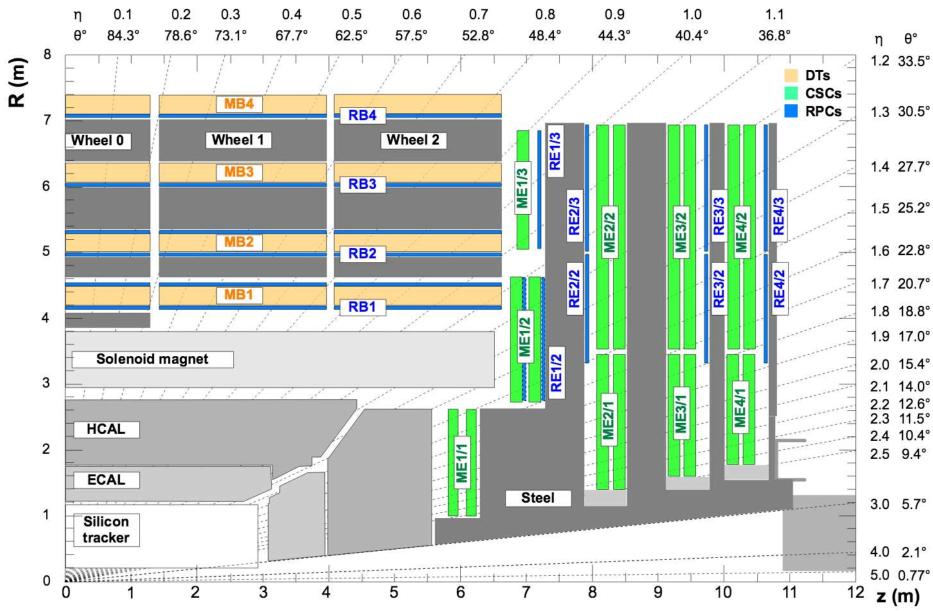


Figure 2.12: A quadrant of the CMS detector with the different muon sub-detectors highlighted [58].

barrel ($|\eta| < 1.2$), where the track occupancy and the residual magnetic field are low, DT detectors are installed. In the endcaps, where the particle rate is higher and a large residual magnetic field is present, CSCs are used. The coverage of the DT and the CSC system goes up to $|\eta| < 2.4$. In the region $|\eta| <2.1$ RPCs are present.

The muon identification is guaranteed by the amount of material in front of the chambers and in the return yoke of the magnet which shields the spectrometer from charged particles other than muons: more than 10 interaction length and 110 radiation length are present before the first measurement station of the spectrometer, at least 16 interaction length of material present up to $\eta = 2.4$ with no acceptance losses.

The magnetic field inside the iron of the yoke bends the tracks in the transverse plane thus allowing the measurement of their p_T. The high field is fundamental for the momentum resolution of the spectrometer but it also sets the environment in which the detector operates. The innermost endcap CSCs, the ME1/1 chambers, are exposed to the full field which, in this region, is almost entirely axial and uniform. In the following CSC stations the field is no longer axial and uniform, however, the small drift space allows these detectors to limit the degradation of the chamber resolution. In the barrel region most of the flux is contained within the iron plates of the yoke where the axial component of the field reaches ≈1.8 T. The space where the DT chambers are placed should ideally be field-free. However in the iron gaps and at the end of the coil the residual magnetic field is far from being negligible: there are spatially limited regions where the field in the radial direction can reach 0.8 T.

The robustness of the spectrometer is also guaranteed by the different sensitivity of DT, RPC and CSC to the backgrounds. The main sources of background particles in the LHC environment will be represented by secondary muons produced in π and K decays, from punch-through hadrons and from low energy electrons originating after slow neutron capture by nuclei with subsequent photon emission. This neutron induced background will be the responsible of the major contribution to the occupancy level in the muon detectors. The total background rate at high pseudorapidity reaches up to 1 kHz/cm^2 in the innermost part of the ME1/1 station. In the barrel the fluences are much lower being everywhere less than 10 Hz/cm^2. As described in the following sub-sections, CSC and DT chambers, in contrast with RPC detectors, are characterized by a layout which helps in reducing the effect of background hits: the request of correlation between consecutive layers is particularly effective against background hits affecting only a single layer.

Drift tubes

The choice of the drift tube detector in the barrel is motivated by the relatively low particle rates and magnetic field intensity in this region. The barrel section of the CMS iron yoke is divided into 5 wheels, forming 3 concentric layers of iron. Each wheel is divided into 12 sectors. The muon chambers are installed on the outer and inner sides of the yoke and in the pockets between layers, arranged in four stations at different radii, named MB1, MB2, MB3 and MB4. Each station consists of 12 chambers, one per sector, except for MB4 where 14 chambers are present.

The basic detector unit in this setup is a drift cell: a gas-filled tube with rectangular cross-section showed in Figure 2.13. The two shorter sides of the rectangle form cathodes, while an anode wire is strung through the middle. A charged particle passing through the detector volume ionizes the gas, producing a cloud of electrons that drifts toward the wire. The drift time is measured and converted to distance using the knowledge of drift velocity. A single drift cell has a cross-section of 42 × 13 mm^2 and wire length 2-3 m. It is filled with a 85%/15% mixture of Ar/CO_2, giving a 350 ns maximum drift time. Single wire measurement resolution is of the order of 200 μm.

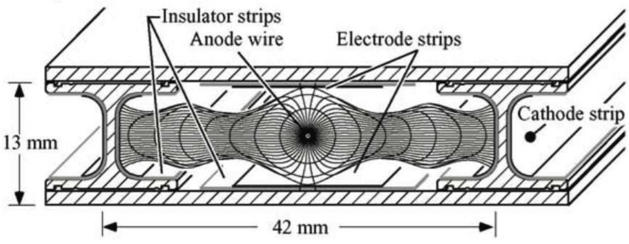


Figure 2.13: Drift tube layout.

The drift tubes in a chamber are grouped into SuperLayers (SL) consisting of four layers of tubes, staggered by half a tube. In each chamber there are two SLs with wires parallel to the beam direction, measuring muon position in the bending plane of the magnetic field. These are separated by a 128 mm thick aluminium honeycomb spacer, providing good angular resolution within one chamber. Additional SL measuring the η coordinate of the muon is present in the three inner stations. Each SL is equipped with fast pattern-recognition electronics, providing bunch crossing identification, and measuring the track segment position and angle.

Cathode strip chambers

Measurement of muon trajectories in the endcap part of the CMS muon system is performed mainly by CSCs. This type of detector has been chosen because of its capability to provide precise time and position measurement in the presence of a high and inhomogeneous magnetic field, and high particle rates.

The detector is a multi-wire proportional chamber with one of the cathode planes being segmented in strips running orthogonally to the wires. The principle of operation is shown in Figure 2.14: a muon crossing the chamber produces an avalanche in the gas (a 40%/50%/10% mixture of $Ar/CO_2/CF_4$) collected by the wire. This induces an electrical charge on several adjacent cathode strips. Fitting the measured distribution of charge picked up by the strips gives an estimate of the position of the muon along the wire.

There are four muon stations integrated into each endcap of the CMS detector (ME1-ME4). The chambers are grouped into rings, with the first station (ME1) consisting of three rings, and the remaining three (ME2-ME4) having two rings of chambers. The rings are formed by 18 or 36 trapezoidal chambers that overlap in ϕ in every ring except the outermost ring of the first station (ME1/3), giving geometrical coverage close to 100%.

Each individual chamber has a trapezoidal shape and is made of seven cathode panels stacked together, forming six gas-gaps each containing an array of anode wires. The gaps are 9.5 mm thick and one of the two cathode planes for each gap is segmented into radial strips orthogonal to the wires. The strips cover a constant area in ϕ (2.33-4.65 mrad, depending on the disk).

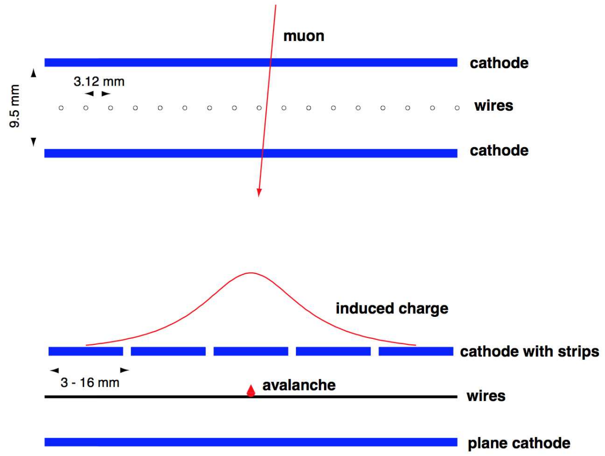


Figure 2.14: The principle of operation of a cathode strip chamber, with cross-section across the wires (top) and across the strips (bottom).

The orthogonal coordinate (r) is measured by the wires which, to reduce the number of channels, are read out in groups of 5 to 16. The inner-most CSC detector lies inside the solenoid, so the wires have to be rotated to compensate for the Lorentz drift.

Resistive plate chambers

RPCs are used throughout the CMS muon system, with the main goal of providing fast trigger signal. They are installed both in the barrel and in the endcaps and are thus complementary to the CSC and the DT systems. These detectors are characterized by an excellent time resolution and fast response providing unambiguous bunch crossing identification, but they have a limited spatial resolution (least one order of magnitude lower than DTs and CSCs) and therefore their impact on muon reconstruction performance is very low.

A single chamber consists of two bakelite planes externally coated with graphite separated by a 2 mm wide gas gap, as shown in Figure 2.15.

Charged particles crossing the gap generate avalanches by ionizing the 96.2% $C_2H_2F_4$ (freon) + 3.5% iC_4H_10 (isobutane) + 0.3% SF_6 + water vapour gas mixture. The signal is read out from detector by a set of aluminium strips, insulated from the electrode with a thin film. In CMS the efficiency of the detector is improved by combining two gas gaps with a common readout plane. This increases the charge induced on the strips. Moreover RPCs operate in "avalanche" mode rather than in the more common "streamer" mode, thus allowing the detectors to sustain higher rates. This mode is obtained with a lower electric field, thus the gas multiplication is reduced and an improved electronic amplification is required.

The barrel RPC chambers follow the segmentation of DT chambers. There are six layers of RPCs, two in the first and second muon station (MB1 and

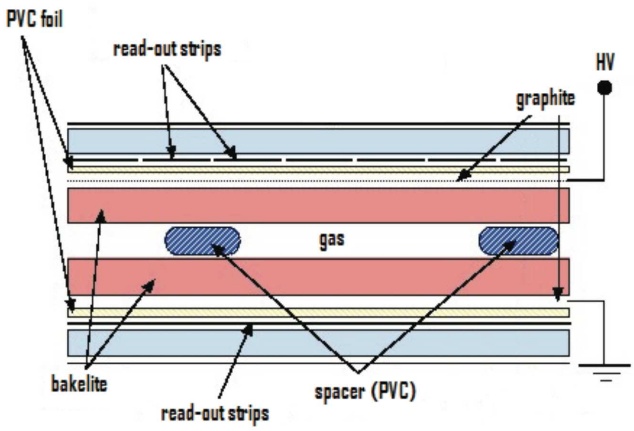


Figure 2.15: Single gap Resistive Plate Chambers layout.

MB2), and one in the third and fourth (MB3 and MB4). The barrel RPCs are rectangular, with dimensions 210-375$\times$ 85 cm. A total number of 96 readout strips run parallel to the beam, with pitch increasing from the inner to the outer muon station: from 2.1 cm in the inner MB1 plane to 4.1 cm in the MB4 planes.

In the endcaps, there are four stations, covering the region up to $\eta = 1.6$. Endcap RPC chambers are trapezoids, with strips running in the radial direction. The strips are also trapezoidal in shape, with a width changing to recover a constant angle in ϕ. The dimensions of the strips vary strongly from detector to detector: they are about 25 cm long and have a pitch of 0.7 cm in the lowest detector of the ME1 chambers (at $\eta = 2.1$) whereas in the chambers at highest r in ME2,3,4 they are about 80 cm in length and have a pitch of roughly 3 cm.

2.2.5 Trigger

The recognizing of the interesting signatures, among the high track multiplicity produced at every LHC collision, is for sure one of the most challenging tasks for the CMS detector. The bunch crossing frequency at CMS interaction point is 40 MHz (bunch spacing of 25 ns) while technical difficulties in handling, storing and processing extremely large amounts of data impose a limit of about 600 Hz on the rate of events that can be written to permanent storage, as the average event size will be of about 1 MB. At the LHC nominal luminosity the total event rate for inelastic interactions is expected to be of the order of 10^9 Hz while the rate of interesting events is very small (see Figure 2.16). A sophisticated trigger system selects events of interest. The time available for the selection is very small since the bunch crossing time is 25 ns. This interval of time is not enough to read out all raw data from the detectors, and for this reason CMS uses a multi-level trigger design, where each step of the selection uses only part of the available data. In this efficient way higher trigger levels have to process fewer events and have more time available; they can go into finer detail and use more refined algorithms. The two steps of the CMS selection chain are: the Level-1 (L1) trigger, built from custom hardware, which reduces the rate to a maximum of 100 kHz, and High Level Trigger (HLT), running the CMS reconstruction software on a processor farm, which performs higher level reconstruction and reduces the rate of events selected by the L1 trigger to about 400 Hz before the events are stored on disk.

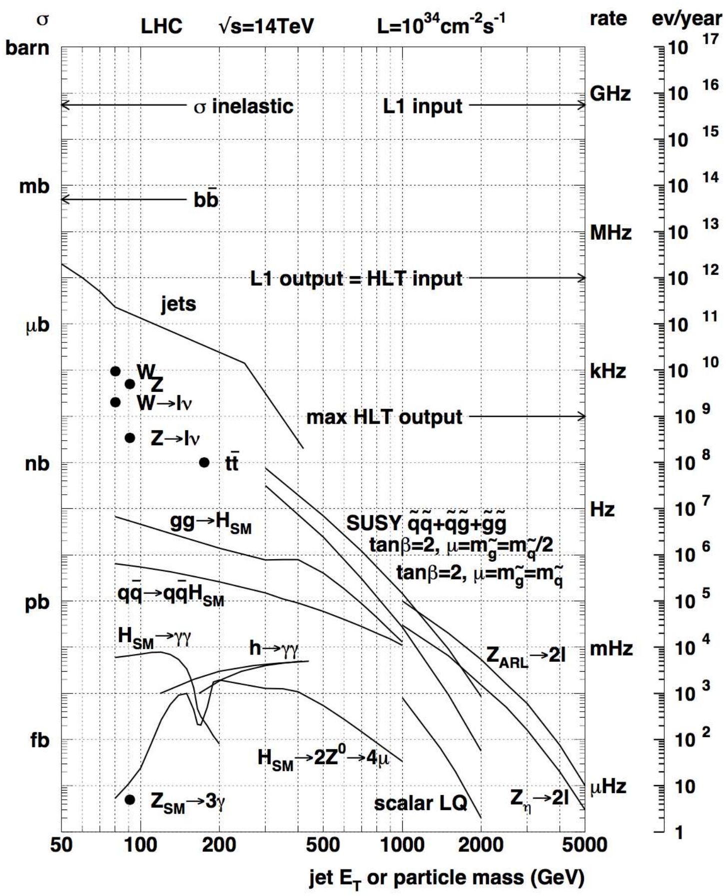


Figure 2.16: Event cross sections and rates of selected processes for the LHC design luminosity of $10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ as a function of the mass of produced objects.

Level-1 trigger

The L1 system [59] is built from custom designed, programmable electronics, and is located underground, both in the service and the experiment caverns. Within a time budget of 3.2 μs, it desides if an event is discarded or kept, and transfer this decision back to the subdetectors, which keep the high resolution data in memory in the meantime. Since the L1 trigger processing time is far greater than the bunch crossing time, the event information is pipelined into a FIFO buffer memory, able to host 128 events.

Since short processing times are required, the L1 system takes into account just a fraction of the whole information coming from subdetectors, ignoring calibration data. Hardware implementation makes use of Field Programmable Gate Array circuits (FPGA), Application Specific Integrated Circuits (ASICS) technology and programmable memory Lookup Tables (LUT).

The L1 is divided in a muon trigger and a calorimeter trigger, which classify and rank interesting event candidates, reconstructed from low resolution data read out from the subdetectors. The rank of a candidate is determined by energy or momentum, and quality of the data. The calorimeter and muon triggers do not perform any selection themselves. They identify "trigger objects" of different types: e/γ (isolated and not), jets and muons. Based on the input from the muon trigger and the calorimeter trigger, the global trigger calculates the final trigger decision. Up to 128 trigger algorithms can be executed in parallel to generate a decision. The simplest triggers are in general those based on the presence of one object with an E_T or p_T above a predefined threshold (single-object triggers) and those based on the presence of two objects of the same type (di-object triggers) with either symmetric or asymmetric thresholds. Other requirements are those for multiple objects of the same or different types ("mixed" and multiple-object triggers). The high resolution data from the inner tracker are not used to generate the L1 decision, which means that there is no information about the vertices and no distinction between electrons and photons available at this level.

High level trigger

Once the L1 trigger has accepted an event, the data of this event are transferred from the buffer memory to the surface, where they are reconstructed in the HLT [60]. The HLT is a special part of the CMS software and runs on a farm of several thousand processors. Each processor works on the reconstruction of one event at a time, to get to a trigger decision within on average 100 ms. Since the time budget for one event is much larger than at the L1 trigger, more complicated algorithms, including tracking, can be executed at the HLT. Once an event is accepted, it is stored on disk and fully reconstructed offline at a later time. The goal of the HLT is to reduce the event rate from the maximum Level-1 output to 600 Hz which is the maximum rate for mass storage.

The use of standard software techniques and languages makes it possible to benefit from the continuous improvements in the reconstruction software. In particular the algorithms used in the HLT, which access data with full resolution and granularity from any part of the detector, is identical to those used in the off-line reconstruction. However, in order to discard uninteresting events as soon as possible, the selection is organized in a sequence of logical steps: the Level-2 and Level-3. The Level-2 uses the full information from calorimeters and muon detectors and reduces the event rate by roughly one order of magnitude. The data from the silicon tracker represent almost 80% of the event size and require complex and time consuming algorithms for the reconstruction. For this reason this information is used only during the Level-3 selection.

The HLT consists of approximately 400 trigger paths, which, starting from the seed of the L1 trigger, look for different objects and signatures in an event. One trigger path is built from reconstruction modules and filter modules. After some parts of the data are reconstructed, a filter module decides if the reconstructed objects pass the thresholds and the next step in reconstruction is started, or if the event is not accepted by the path. In the later case, the execution of the path is stopped and the following reconstruction steps and filter steps are not performed to save computation time. Following this concept to save computation time, the less computation intense reconstruction steps (e.g. unpacking the data from the ECAL and measuring the energy deposit) are done first. The reconstruction steps that take a lot of time, e.g. the tracking, are done at the end of a path for objects that have already passed the previous steps. If an event is not accepted by a path, it can still be accepted by a different path.

If, for some paths with low thresholds, the acceptance rate is too high, they can be prescaled to lower the rate. A prescale value of ten means, for example, that the path is executed only for every tenth event that was accepted by the L1 trigger, and, consequently, the trigger rate for that path is ten times smaller. The prescale value for one trigger path has several predefined levels, depending on the instantaneous luminosity of the LHC machine. During an LHC fill, the instantaneous luminosity decreases, and the prescale values can be changed during a CMS run to keep the global trigger rate at an optimal level.

Chapter 3

CMS event reconstruction

3.1 Physics objects reconstruction

Data collected in each of the CMS sub-detectors are used to reconstruct physics objects through offline software algorithms which allow the identification of particles passing through CMS and identify their physics parameters of interest, like the charge and 4-momentum. The main physics objects used in CMS are: Muons, Electrons, Jets, Photons, Missing Energy, Taus. Figure 3.1 shows a pictorical view of the different objects reconstructed in CMS and of the signature they give in the different CMS subdetectors.

3.1.1 Charged particles track reconstruction

Charged particles are detected in the inner tracking system. Their trajectory bends in the CMS magnetic field, moving along an helix whose pace is related to the transverse momentum:

$$p_T \propto B \cdot r_{curl} \tag{3.1}$$

where B is the magnetic filed and r_{curl} is the curling radius of the circumference obtained projecting the helix in the $x - y$ plane. Therefore a precise reconstruction of the tracks is crucial for precise momentum measurements. The tracks in CMS inner tracking system are reconstructed with a fit using as input the position of the strips or pixel fired (*hits*) in the detectors. A pattern recognition is performed based on the Kalman filter method [61]. First of all a starting point (*seed*) is found looking at all hits in the tracker. Each seed is composed of a small subset of the position measurements in the tracker itself. Since five parameters (including the trajectory curvature) are needed to start trajectory building, at least 3 hits, or 2 hits and a beam constraint, are necessary to properly define a seed. The Kalman filter then proceeds iteratively from the layer where the seed is located starting from a coarse estimate of the track parameters provided by the trajectory seed, and includes the information

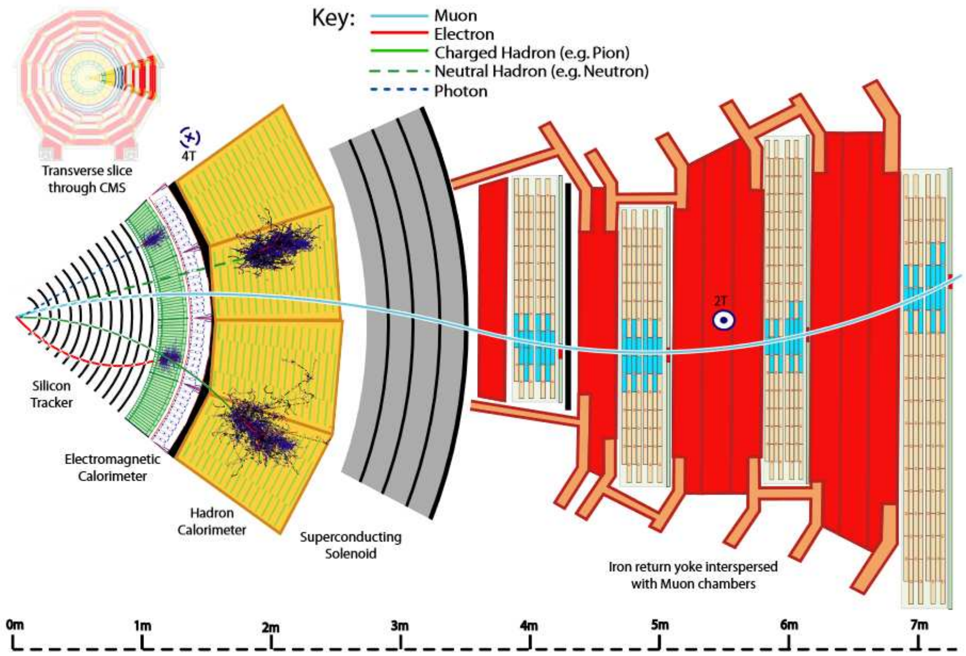


Figure 3.1: Slice through the CMS detector, illustrating the signatures of different types of particles.

of the successive detection layers one by one. On each layer, i.e. with every new measurement, the track parameters are known with a better precision, up to the last point, where they include the full tracker information. The Kalman filter is initialized at the location of the innermost hit with an estimate obtained during seeding. The corresponding covariance matrix is scaled by a large factor in order to avoid any bias. The fit then proceeds in an iterative way through the full list of hits. For each valid hit the position estimate is re-evaluated again using the current values of the track parameters. This first filter is complemented with the smoothing stage: a second filter is initialized with the result of the first one (except for the covariance matrix, which is scaled with a large factor) and is run backward toward the beam line. This filtering and smoothing procedure yields optimal estimates of the parameters at the surface associated with each hit and, specifically, at the first and the last hit of the trajectory. Estimates on other surfaces, e.g., at the impact point, are then derived by extrapolation from the closest hit. On top of the standard Kalman filter, an iterative tracking procedure [62] has been developed in CMS to preserve high tracking efficiency while minimizing the fake rate. For each iteration, the following steps are applied:

- Seed finding is performed on the available hits. The seeding configuration is the main difference between iterative steps.

- Track reconstruction (building, filtering, fitting, smoothing) is performed using the available hits. Parameters in each stage can be tuned separately at each iteration to improve performance.

- The track collection is cleaned according to quality criteria and the collection of tracks which pass the cleaning stage is stored.

3.1.2 Muon reconstruction

Muons are the charged particles that are best reconstructed in the Tracker. They mainly interact with the silicon detector through ionization and their energy loss by bremsstrahlung is generally negligible, except when muons are produced with an initial energy higher than about 100 GeV. Therefore these particles usually cross the whole volume of the tracking system, producing detectable hits on all the sensitive layers of the apparatus. To identify muons, informations from the outer CMS Muon System are used in combination with the tracker information.

The reconstruction starts with the so-called trajectory seeding. In CMS trajectory seeds are hit-based seeds (or state-based seeds, using momentum information). Hit-pairs (or hit-triplets) are required to be compatible with beam spot (further criteria can be added, for instance imposing the hit position is placed in a given region). The Seed Generator is based on DT and CSC segments: the former provides track segments in the ϕ projection (being $\Delta\phi$ the bending angle with respect to the vertex direction) and hit patterns in η projection; the latter delivers three dimensional track segments. Trajectory building then starts in the direction specified by seed, towards subsequent layers: in the standard configuration parameters are propagated from outer detector layers toward the innermost compatible ones. Compatible hits are searched and the track finding and fitting is accomplished by an iterative Kalman filter technique. Material effects, mainly due to random Coulomb scattering, are included in the iterative steps, since they introduce a gaussian-distributed uncertainty on scattering angle. Particle propagation is a very time consuming phase. Step by step, along trajectory propagation new hits informations are included in trajectory description using an outside-in reconstruction, as well as the knowledge of the magnetic field and detector material. The process is stopped when the innermost compatible layer of muon detectors is reached.

Since this procedure may give rise to a number of trajectory that may share the same hits, a process resolves all the ambiguities, keeping a number of track candidates. Finally the stakes of any remaining trajectories are removed a backward fitting is performed. Once the hits are fitted and the fake trajectories removed, the remaining tracks are extrapolated to the point of closest approach to the beam line. In order to improve the p_T resolution a beamspot constraint is applied.

Muon reconstructed tracks are classified in three categories, depending on the detectors used for muon reconstruction:

- StandAlone Muons: the track is reconstructed only in the muon detector.
- Global Muons: Those muons are reconstructed both in the Tracker and in the Muon System. A matching criteria is adopted to match inner tracks to tracks in the Muon System, and then the Kalman Filter is applied again on hits from both tracks to get a better estimate of muon parameters.
- Tracker Muons: Those muons are reconstructed in the inner Tracker and then matched to a segment in either the DT or the CSC. The matching criteria between the inner track and the muon is tighter than for Global Muons since the muon system segment lacks the robustness of a full track reconstruction

The majority of muons are constructed either as Global Muon or Tracker Muons. Just 1% of muons are reconstructed as Standalone-muon tracks only.

The momentum resolution as a function of the pseudorapidity is presented in Figure 3.2

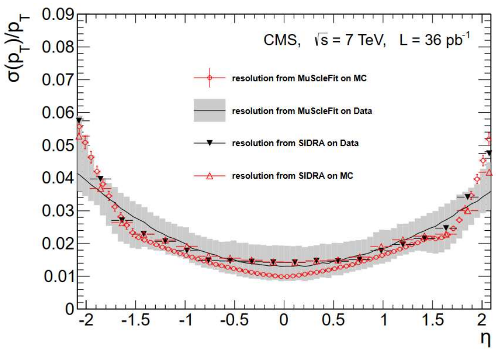


Figure 3.2: Relative transverse momentum resolution $\sigma(p_T)/p_T$ for muons in the decay of Z boson [63]

Muon Identification Variables A number of variables can be used for muon identification and quality selection in muon analysis, which have been described and reported in detail in [63]. Some of them are

- The number of track segments built from hits in muon chambers with the inner track extrapolation; such a quantity can be useful to reject muon from light flavors decays;
- The transverse impact d_0 parameter in the x-y plane, defined as the distance between the point of closest approach to the beamline and the beamline itself. The transverse impact parameter distribution tails are dominated by pion and kaon decays in flight. A longitudinal impact parameter d_z can be also defined as the z-coordinate of the point of closest approach along the trajectory).

- The number of valid hits of muon hits both in tracker and muons system;
- The χ^2 of the track fit both for the silicon tracker tracks or for global track;
- The combined isolation variable , which is able to distinguish prompt muons from non-prompt muons, for example coming from jets. Such a variable is calculated building a cone around muon trajectory, with radius $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$. A standard value for ΔR is 0.3. The scalar sum of tracks p_T inside the cone as well as the the energy loss inside ECAL and HCAL is calculated, excluding contribution from the candidate itself. The relative isolation is defined as the ratio of the total energy inside the cone to the transverse momentum of the candidate. In fact a more refined isolation variable based on reconstructed particles, named Particle Flow Isolation variable can be used to identify good muon candidates. Such a variable will be defined in Chapter 4, while the Particle Flow algorithm is described in Paragraph 3.1.4. The efficiency for the isolation variables as a function of the isolation threshold are shown in Figure 3.3.

3.1.3 Electron and Photon reconstruction

Electrons and photons energies are measured in the CMS ECAL [64]. To collect the photons and electrons energy in the ECAL, local deposits (*basic clusters*) are summed into superclusters (SCs) which are extended in ϕ. After applying small energy corrections the superclusters are used to reconstruct photons and electrons, and to seed electron track reconstruction. Two complementary algorithms are used at the track seeding stage: *tracker driven* seeding, more suitable for low p_T electrons as well as performing better for electrons inside jets and *ECAL driven* seeding. The ECAL driven algorithm starts by the reconstruction of ECAL superclusters of transverse energy $E_T > 4$ GeV and is optimized for isolated electrons in the p_T range relevant for Z or W decays and down to $p_T \approx$5 GeV.

Photons are reconstructed from the energy corrected superclusters, assigning the candidate momentum to the location of the reconstructed primary vertex. The energy of each photon candidate is estimated based on an observable called $r9$ which is the ratio of the energy contained within the 3×3 array of crystals centered on the seed crystal of the photon candidate's supercluster to the total energy contained in the supercluster. This quantity is used to determine if the photon is converted or unconverted. If the $r9$ of the candidate is above 0.94 (0.95) in the barrel (endcap), the energy of the 5×5 crystals around the highest energy crystal is used. Otherwise, the supercluster energy is used.

Electron identification variables Electron selection variables are defined and used to discriminate between real and fake electrons:

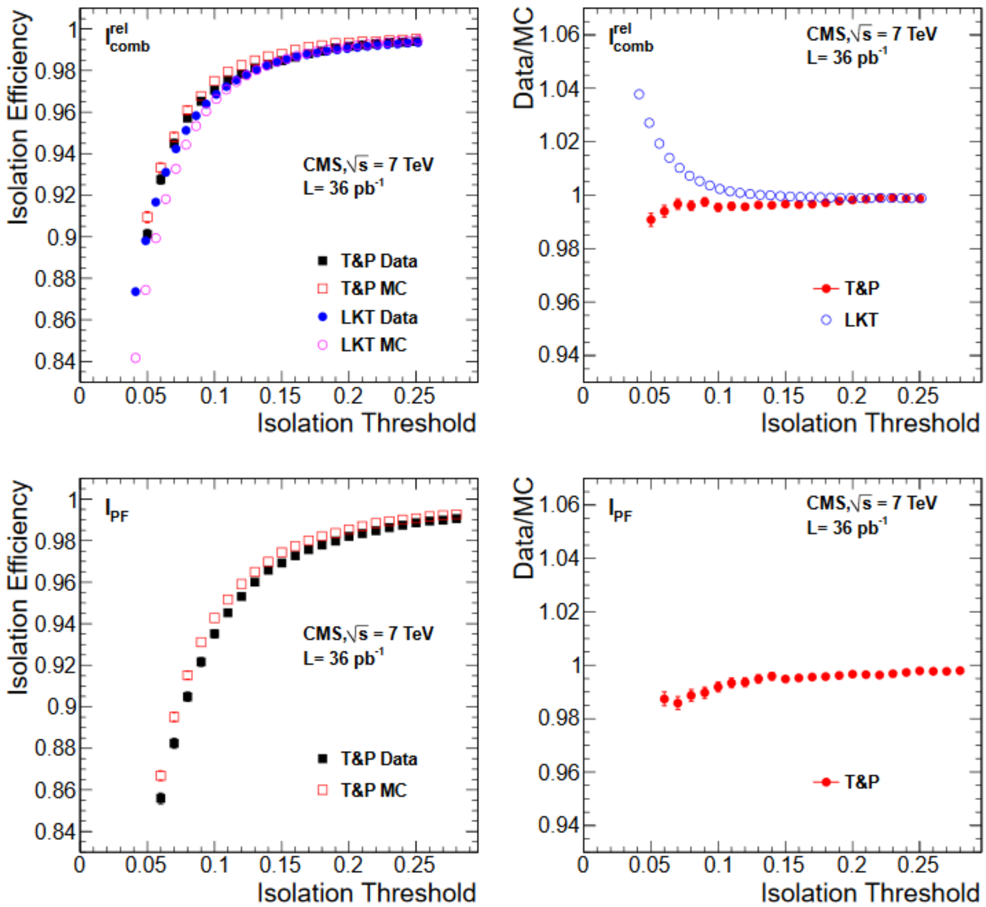


Figure 3.3: Efficiencies for the combined and particle flow isolation algorithms for muons from Z decays as a function of the isolation threshold [63]. Result are shown for both data and simulation using the tag-and-probe (T&P) and Lepton Kinematic Template (LKT) methods; the LKT method is not used for the particle flow algorithm.

- $\Delta\eta_{in} = \eta_{sc} - \eta_{in}^{extrap}$, where η_{sc} is the energy-weighted centroid position of the supercluster, while η_{in}^{extrap} is the associated track pseudorapidity at ECAL surface as extrapolated from the innermost track layer
- $\Delta\phi_{in} = \phi_{sc} - \phi_{in}^{extrap}$, where ϕ_{sc} is the energy-weighted centroid position of the supercluster, while ϕ_{in}^{extrap} is the associated track pseudorapidity at ECAL surface as extrapolated from the innermost track layer
- $\sigma_{i\eta i\eta} = \sqrt{\Sigma_i^{5\times5}\omega_i(\eta_i - \bar{\eta}_{5\times5})^2/\Sigma_i\omega_i}$ where the index i runs in the η position of i^{th} crystal in a 5×5 block on crystal centered on the seed crystal, η_i is the η position of the i^{th} crystal, $\eta_{5\times5}$ is the energy weighted mean η of the block and ω_i is the weight of the i^{th} crystal, defined as $\omega_i = 4.7 + ln(E_i/E_{5\times5})$, being E_i and $E_{5\times5}$ the energy of the i^{th} crystal and the block respectively.
- H/E where H is the energy deposited in HCAL towers in a cone of radius $\Delta R = 0.15$ centered on the electromagnetic supercluster position, while E is the energy of the electromagnetic supercluster.
- the transverse and longitudinal inpact parameters d_{xy} and d_z

3.1.4 Particle flow

Particle flow algorithm aims to reconstruct all the stable particles in the event, namely electron, muons, photons, charged and neutral hadrons, combining the information from all the CMS subdetectors. It worth to note that most of stable constituents have usually low p_T values (as the final products of exotic particle decay chains): thus an accurate, efficient and low-fake rate reconstruction must be performed. The list of particles is then used to build jets, to determine missing transverse energy E_T^{miss}, to reconstruct and identify taus from their decay products, to give an estimate of lepton isolation, to b-tag jets etc. The information of the basic reconstruction objects are combined and linked through a linking algorithm to form physical objects. Particle Flow algorithm is roughly composed by three steps

- Iterative tracking: it is based on the information coming from the tracker detector, which is able to provide an accurate measurement of charged particle direction at the production vertex. Track are first reconstructed with very tight requirements, with a moderate efficiency but also a very low fake rate. Once a tight trajectory has been built, assigned hits are removed, seeding criteria are loosened. Combinatorics is thus reduced and fake rate is kept low. In the first three iterations, 99.5% of isolated muons and 90% of charged hadrons are identified. In the subsequent iterations, a relaxed constraint on the origin vertex is chosen to allow the reconstruction of secondary charged particles produced in photon conversion or nuclear interactions in the tracker material.

- Calorimeter clustering: this step aims to four objectives, detecting and measuring neutral particle, separating them from charged hadrons reconstructing and identifying electrons and bremsstrahlung radiation, helping energy measurements of charged hadrons, and low quality or high-p_T tracks. A specific clustering algorithm is performed separately on different components of ECAL, HCAL and PS. Clustering starts from an energy maxima in calorimeter cell (seed); cells with common side and with a signal two standard deviations above the electronic noise are then aggregated in topological clusters, which give rise to as many particle-flow seeds as clusters. An iterative procedure determines position and energy of clusters.

- Link algorithm: it connects each element to fully reconstruct objects and single particles avoiding double counting, providing, for each pair of elements, a distance which is used to quantify the quality of the link. Some *bloks* containing two or three elements are produced, being the base of particle reconstruction and identification. Links can connect charged particle track and calorimeter clusters, or two calorimeter clusters, charged particle tracks in the tracker and muon track in muon system. More details about link algorithm can be found in [65].

Once blocks have been built, particle flow algorithm performs the reconstruction and identification step of the muon, electrons and all the remaining tracks, making available a full event description for the analysis. Some example of the performance of the PF algorithm is presented in Figure 3.4.

3.1.5 Jet reconstruction

Anti-kt algorithm Jets represent the signature of quarks and gluons emissions, which hadronize and give rise to a number of hadrons as a consequence of quark confinement predicted by QCD. Hadrons fly in the same direction of the parton object which they are generated from and release their energy mainly in ECAL and HCAL cells. A calotower, namely the combination of consecutive ECAL and HCAL cells, define jet energy in the $\eta - \phi$ plane.

In this analysis the clustering algorithm deputed to jet reconstruction is the anti-kt with cone size 0.5, which in few years has become the most used algorithm in CMS analysis. Energy reconstruction and calibration has been performed combining subdetectors information through a PF algorithm. Anti-kt algorithm is based on the generalization of Cambridge/Aachen algorithms. Two distance measures d_{ij} and d_{iB} are defined:

$$\begin{aligned} d_{ij} &= min(k_{ti}^{2p}, k_{tj}^{2p})\frac{\Delta_{ij}^2}{R^2} \\ d_{iB} &= k_{ti}^{2p} \end{aligned} \tag{3.2}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ denotes a distance in the $\eta - \phi$ space, while k_{ti}, y_i, ϕ_i are respectively the transverse momentum, rapidity and azimuthal

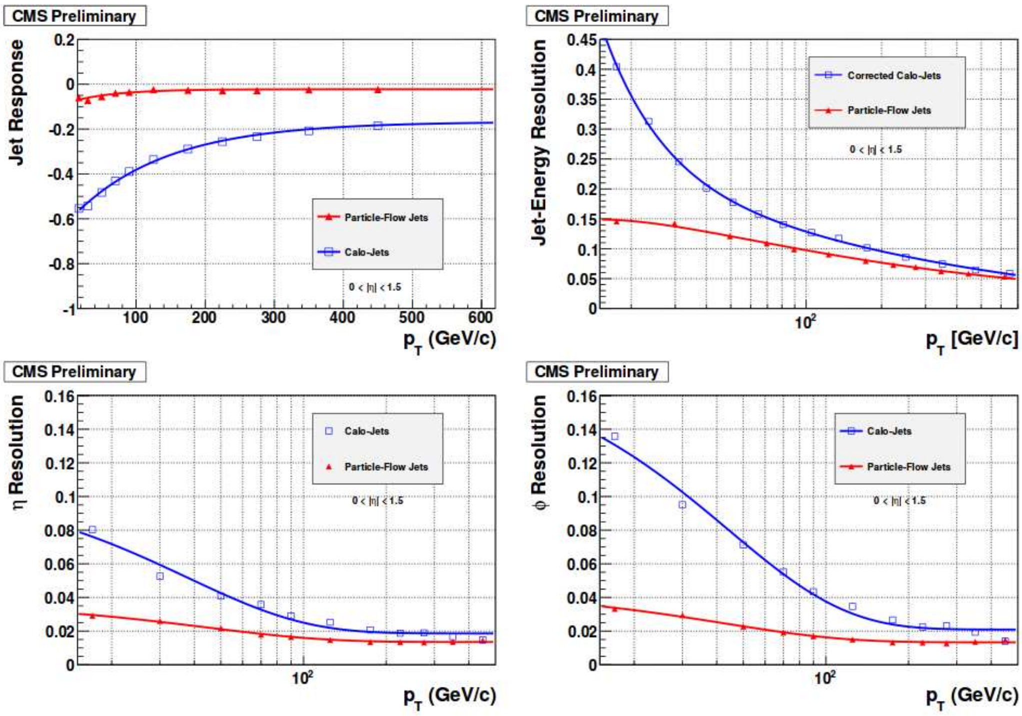


Figure 3.4: Jet response (upper left) energy resolution (upper right), η resolution (lower left), and ϕ resolution (lower right) for jets using calorimeter clusters only or particle flow candidates as input [65].

angle of the $i-th$ particle. R is a radius parameter (set to 0.5 in AK5 jets), while p governs the relative power of the energy versus geometrical scale Δ_{ij}. Anti-kt algorithm holds for $p=-1$.

Both d_{iB} and d_{ij} are calculated for all input objects or pairs of objects, respectively. If the smallest one is a d_{ij} those both objects are clustered together. If a d_{iB} is smallest, the corresponding object is considered to be a jet and excluded from further clustering. The clustering stops when no objects are left.

The functionality of the algorithm can be understood by considering an event with a few well-separated hard particles with transverse momenta $k_{t1}, k_{t2}, ...$ and many soft particles. The d_{1i} between a hard particle 1 and a soft particle i is exclusively determined by the transverse momentum of the hard particle and the Δ_{1i} separation. The d_{ij} between similarly separated soft particles will instead be much larger. Therefore soft particles will tend to cluster with hard ones long before they cluster among themselves. If a hard particle has no other hard particles within a distance of 2R, it just accumulates soft particles around itself. If another hard particle 2 is present within a distance $R < \Delta_{12} < 2R$ two jets can be produced, even though at least one of the two will not be perfectly conical; if $\Delta_{12} < R$ the two hard particles cluster in a single jet. It can be shown that soft particles soft particles do not change significantly jet shape, while hard ones do, making this algorithm quite stable also in pileup environments. More details about anti-kt algorithm can be found in [66].

Jet energy corrections The purpose of the jet energy calibration is to relate, on average, the energy, or equivalently the transverse momentum measured for the detector jet to the energy of the corresponding true particle jet. A true particle jet results from the clustering (with the same clustering algorithm applied to detector jets) of all stable particles originating from the fragmenting parton, as well as of the particles from the underlying event (UE) activity. The correction is applied as a multiplicative factor C to each component of the raw jet four-momentum vector p_μ^{raw} (components are indexed by μ in the following):

$$p_\mu^{cor} = C \cdot p_\mu^{raw}. \tag{3.3}$$

The correction factor C is composed of the offset correction C_{offset}, the MC calibration factor C_{MC}, and the residual calibrations C_{rel} and C_{abs} for the relative and absolute energy scales, respectively. The JEC scheme used in CMS [67] is shown in Figure 3.5.

The offset correction, named L1, removes the extra energy due to noise and pile-up. Additional pileup events are found to add a transverse momentum of about 0.6 GeV per unit area in $\eta - \phi$ space and additional interaction. This amounts to 0.5 GeV on average interaction for anti-k_T jets with $R = 0.5$. The

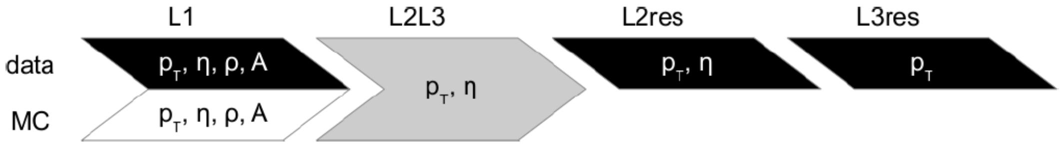


Figure 3.5: Jet energy correction scheme used in CMS analyses [68]

impact of pileup events is reduced by rejecting tracks that stem from other vertices than the seleted primary one. This charged hadron subtraction (CHS) is applied in this analysis. The so-called L1 pileup offset correction adds a correction factor depending on some jet parameters such as the transverse momentum p_T, the pseudorapidity η, the momentum density ρ and the jet area A. The correction has been determined in MC simulation by comparing identical events with and without pileup events, and has a size of -15% or -25% for 30 GeV jets in the barrel with and without CHS, respectively.

The L2L3 MC correction removes the bulk of the non-uniformity in η and the non-linearity in p_T. The L2L3 MC truth response R is defined as the ratio between the reconstructed and the generated jet transverse momentum. The corrections are evaluated using a QCD dijet MC sample after detector simulation and L1 correction. The ratio R is computed in bins of transverse momentum and pseudorapidity, and leads to correction factors of 5-15% in the barrel and up to 70% in the endcaps and forward regions, due to the increase of the inactive material that affects the linear response of the calorimeter and the traking efficiency. Both the L1 and L2L3 MC truth corrections are applied to both data and simulated events. Finally, the residual corrections referred to as L2Relative and L3Absolute account for the small differences between data and simulation, and are applied to simulated jets only. The various components are applied in sequence as described by the equation below:

$$C = C_{offset}(p_T^{raw}) \cdot C_{MC}(p'_T, \eta) \cdot C_{rel}(\eta) \cdot C_{abs}(p''_T), \tag{3.4}$$

where p'_T is the transverse momentum of the jet after applying the offset correction and p''_T is the p_T of the jet after all previous corrections.

Jet Energy Resolutions The jet energy resolution (JER) was measured with the *dijet asymmetry* method [67, 69]. As shown in Figure 3.6 (left), the resolution in data was found to be worse than in simulation by about 10% on average and up to 40% at most, depending on the detector region. The p_T dependence of the resolution in data was found to be well modeled by the simulation, as reported in Figure 3.6 (right). Also shown is the impact of pileup events on the jet energy resolution.

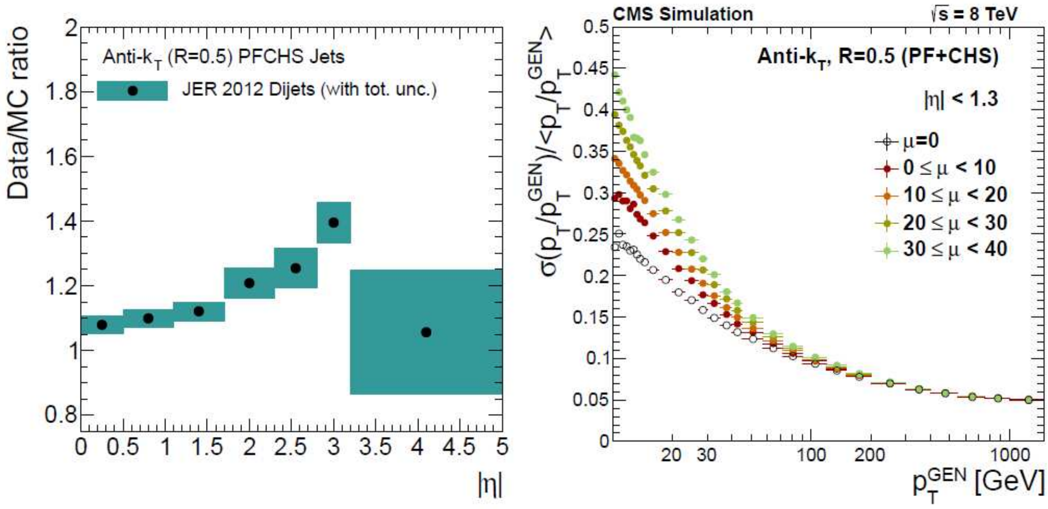


Figure 3.6: Corrections for the jet energy resolution at CMS in dependency of η (left) [69], and resolution in dependency of p_T and pileup events μ (right) [70]

b tagging Several algorithms are defined at CMS with the purpose to tag jets stemming from b quarks hadronization, or b-jets. Such jets usually contain B-hadrons which present several characteristics which allow to discriminate between b-jets and jets stemming from light quark hadronization (also referred to as light jets). First of all, the tracks produced by long lived particle decays (such as B-hadrons) are expected to have a non negligible impact parameter (IP). The IP is invariant with respect to changes of the long lived particle kinetic energy, this is due to the cancellation of the boost effects on the flight path (scaling as $\approx \gamma$) and the average angle of the decay products with respect to the flight direction (scaling as$\approx 1/\gamma$). The typical scale of the IP is the one of the decaying particle $c\tau$; for a B-hadron this corresponds to about 450 μm. In CMS the IP can be measured with a precision between 30 μm and few hundreds of μm. Given that the uncertainty can be of the same order of magnitude as the IP, a better observable for b-tagging is the impact parameter significance defined as

$$S = \frac{IP}{\sigma_{IP}} \tag{3.5}$$

The IP in CMS is *life time signed*: tracks orginating from the decay of particles travelling in the same direction of the jet are signed as positive, while those in opposite direction are tagged as negative. This is obtained by using the sign of the scalar product of the IP segment with the jet direction. On the other hand, it is possible to reconstruct the secondary vertices from B hadron decays inside of jets. To do this an adaptive vertex fit is performed. b-jet tagging algorithms can be divided in the following categories:

- Track Counting algorithm [71, 72]. This is the most simple algorithm, exploiting the long lifetime of B hadrons. It calculates the signed impact parameter significance of all good tracks, and orders them by decreasing significance. Its b tag discriminator is defined as the significance of the N^{th} track. It comes in two variations, with $N = 2$, to obtain a high efficiency, or $N = 3$, to guarantee high purity.

- Jet Probability algorithm [71]. Its b tag discriminator is equal to the negative logarithm of the confidence level that all the tracks in the jet are consistent with originating from the primary vertex. This confidence level is calculated from the signed impact parameter significances of all good tracks.

- Lepton based algorithms [71, 73]. Algorithms based on the presence of a lepton stemming from the B-hadron decay close to the jet's axis are present for both muons and electrons. The Soft Muon tagger takes into account the presence of a well reconstructed muon close to the jet's axis and uses as discriminator variable either the $p_{T,rel}$ of the muon with respect to the jet axis (soft muon by $p_{T,rel}$) or the muon IP significance. The Soft Electron tagger checks the presence of an electron close to the jet's axis and uses as discriminator a neural network variable based on the electron IP significance, the electron $p_{T,rel}$ with respect to the jet axis, the ΔR between the electron and the jet, and the ratio between the electron momentum, as reconstructed in the tracker, and the calorimetric jet energy.

- Simple secondary vertex algorithms. These class of algorithms reconstructs the B decay vertex using an adaptive vertex finder, and then uses variables related to it, such as decay length significance to calculate its b tag discriminator. It has been found to be more robust to tracker misalignment than the other lifetime-based tags.

- Combined Secondary Vertex algorithm [74]. This sophisticated and complex tag exploits all known variables which can distinguish b from non-b jets. Its goal is to provide optimal b tag performance, by combining information about impact parameter significance, the secondary vertex and jet kinematics. More precisely, the CSV algorithm workflow starts with the so-called Trimmed Kalman Vertex Finder [75] that reconstructs secondary vertices in an inclusive way inside the jet. This algorithm begins by using all tracks in the jet and subsequently rejects outliers which then are used to reconstruct additional vertices. Different cuts are applied to the reconstructed vertices to select good secondary vertex candidates. The secondary vertex reconstruction and selection classifies the vertices into three categories defined as *RecoVertex, PseudoVertex, or NoVertex*. When at least one secondary vertex candidate is reconstructed and satisfies the selection criteria, one has a RecoVertex. All tracks from all accepted vertices are used for the computation of the vertex related variables if there is more than one accepted secondary vertex. If no good reconstructed secondary vertex candidate is found, a so-called PseudoVertex is created using charged particle tracks not compatible with the primary vertex, if at least two such tracks are present in the jet. If neither the RecoVertex nor the Pseudovertex conditions are fullfilled, one has a so-called NoVertex. Optimal performance is achieved by combining several topological and kinematical variables related to the secondary vertex reconstruction, as well as variables related to the impact parameter signicances of charged particle tracks. The choice of variables entering into the combination depends on the vertex category. The track impact parameter signicances of accepted tracks enter into the discriminator for all categories. Examples of the other variables considered are the invarant mass and the multiplicity of charged particles associated to the secondary vertex, the distance between the primary vertex and the secondary vertex in the transverse plane, divided by its error, the energy of the charged particle associated to the secondary vertex relatively to the energy of all charged particles associated to the jet, and so on. The variables are combined using a likelihood ratio technique to compute the b tag discriminator. Different operating points were defined, loose (L), medium (M), and tight (T), corresponding to a misidentification probability for light-parton jets of close to 10%, 1% and 0.1% respectively, with an average jet p_T of about 80 GeV. This analysis makes use of the CSV b-tagging algorithm with medium working point (CSVM). Performance plots for this tagger are shown in Figure 3.7.

3.1.6 Missing transverse energy reconstruction

Neutrinos interact only weakly and thus leave no direct evidence in the detector. Still their p_T can be estimated by summing the vectorial transverse momenta of all other objects in the event. The imbalance

$$\overrightarrow{E}_T^{miss} = -\Sigma_{particles}\overrightarrow{p}_{T,i} \tag{3.6}$$

is called the missing transverse momentum or energy (MET) [77]. The MET reconstruction is improved by the so-called $Type - I$ corrections, where effectively the momenta of the uncalibrated jets in the expression above are replaced with the momenta of the corresponding calibrated jets. This analysis uses the MET as estimate for the transverse momentum of the undetected neutrino from the leptonic W decay.

3.1.7 Vertices and pileup

Given the high instantaneous luminosity in the 2012 data-taking, the average number of interactions per bunch crossing was 21. The interaction vertices are reconstructed in CMS as follows [78]: charged particle tracks are selected that stem from the interaction region, having an impact parameter significance

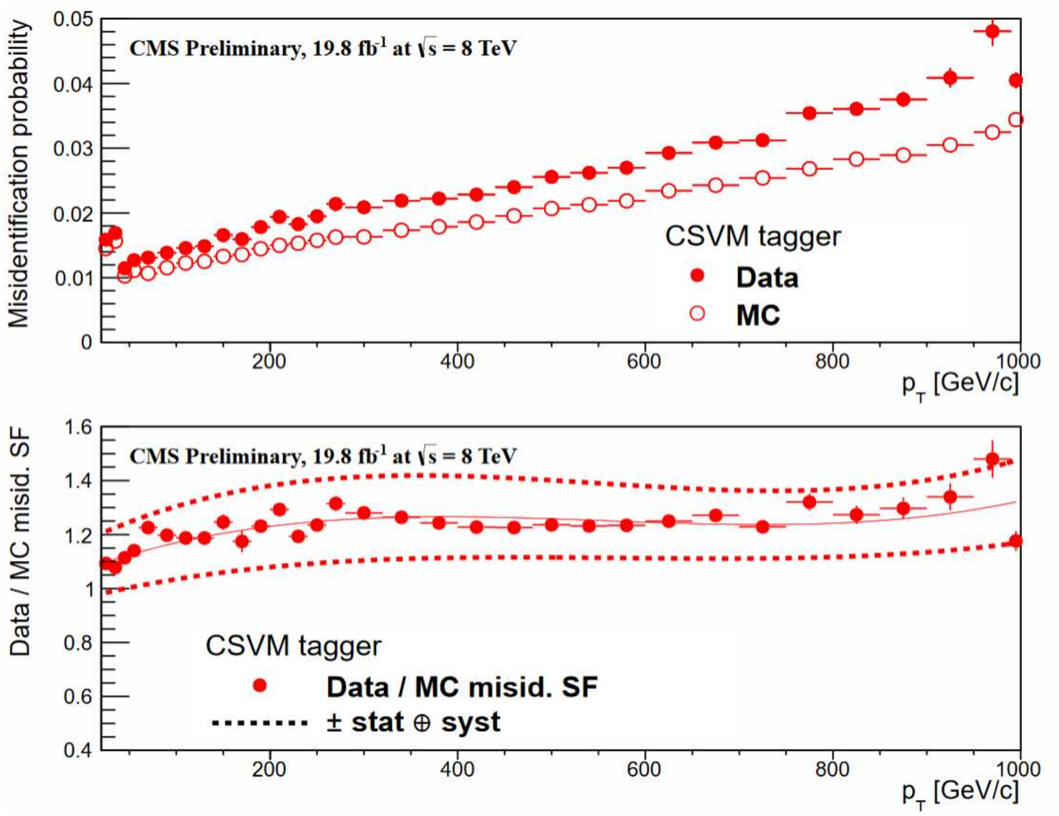


Figure 3.7: Misidentification probability in data and simulation (top), and data-MC scale factor for the misidentification probability (bottom) of the Combined Secondary Vertex tagger at the medium working point (CSVM) in dependence of the jet p_T. [76]

below 5σ. Each track needs to be fitted from at least five hits in the inner tracking system, with at least two hits in the pixel tracker. The normalized χ^2 of the track fit is required to be below 20. The adaptive vertex fitter [78] is used to cluster the tracks to vertex candidates. The algorithm tries to find as many vertices as possible without splitting true vertices. It assigns a weight (or probability) w between 0 and 1 to each connection between a track and a vertex candidate, based on their positions on the z-axis and the corresponding uncertainties. A minimal weight of 0.5 is required for each connection, and at least two tracks of a candidate vertex must be incompatible (w<0.5) with other vertices. As figure of merit for a fitted vertex, the number of degrees of freedom is defined as

$$n_{dof} = -3 + 2\Sigma_{i=1}^{tracks} w_i, \tag{3.7}$$

so that a large number of compatible tracks results in a high value. Using CMS data at a center-of-mass of 7 TeV, the primary vertex resolution was found to be below 50 μm in the x/y and z directions for vertices with more than ten associated tracks. The mean number of fitted primary vertex candidates per event in 2012 data is 14.55, corresponding to a vertex identification efficiency of about 70%, including the loss from vertices with neutral particles only. For each event, the vertex with the highest Σp_T^2 of the associated tracks is regarded as the primary interaction vertex. It is required to be within 24 cm in longitudinal and 2 cm in transverse direction from the nominal interaction point, and to have $n_{dof} > 4$.

Chapter 4

Analysis

4.1 Signal and backgrounds

The analysis aims at investigating possible production of a charged Higgs boson in the decay of the top quark where the former decays to a charm and an antibottom quark.

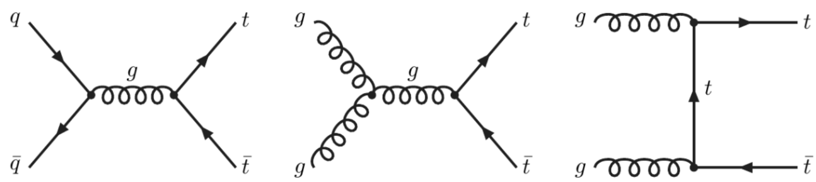


Figure 4.1: Leading order $t\bar{t}$ pair production modes.

The top quark can be produced via $q\bar{q}$ interactions or gluon fusion according to the diagrams in Figure 4.1. The $t\bar{t}$ final states depend on the W boson decay modes, and three channels are usually identified and considered in physics analyses: dileptonic, when both W bosons decay in one lepton (electron or muon) and a neutrino (BR≈5%); single lepton, when one W boson decays in two quarks, while the other decays leptonically into electron or muon (BR≈30%); fully hadronic, when both W bosons decay hadronically (BR≈44%).

If the top decay to charged Higgs takes place ($t \to H^+b \to c\bar{b}b$), the best channel to find it is $t\bar{t}$ production, with one top decaying to the charged Higgs boson and subsequently giving three jets, two of which originate from b-quarks, while the other top quark via the SM leptonic decay $t \to W^-\bar{b} \to l\bar{\nu}b$. This thesis is focused on the muonic channel.

The major irreducible background is the SM $t\bar{t}$+jets process, where both top decay to a b quark and W boson, followed by one W decaying to $\mu\nu$ and the other decaying to quarks. Feynman diagrams for signal and SM $t\bar{t}$ backgrounds are shown in Figure 4.2. The final states are identical, except for the number of b-jets, which is three in case of signal, two in case of SM $t\bar{t}$.

Requiring three b-jets in the final state thus helps to reduce SM backgrund. In addition to that, using proper kinematic assumptions, one can try to pick up, for each event, the two jets that are most likely to come from the decay of the boson (W or H^+) in the hadronic branch of the $t\bar{t}$ event. In the SM decay, the invariant mass of these two jets will obviously peak at the W mass. However, in the case of $t \to H^+b \to c\bar{b}b$ decays, the dijet invariant mass will peak at the mass of the charged Higgs boson. The analysis searches for a charged Higgs boson by looking for a second peak in the dijet mass spectrum of the $t\bar{t}$ decays.

Top quarks can also be produced in single top channel, through the following production modes (Figure 4.3):

- s-channel: $qq \to t\bar{b}$ the rarest production process (approx NNLO cross section: 5.55 pb at $\sqrt(s) = 8TeV$ [79])
- t-channel: $bq \to tq'$ which has the bigger PDF uncertainties due to the presence of gluons, but it has a larger cross section with respect to s-channel (approx. NNLO cross section: 84.69 pb at $\sqrt(s) = 8$ TeV [79]).
- tW: $gb \to t\bar{W}$ t starts with one gluon and one b-quark, and it represents the 20% of the total cross section in single top production (approx NNLO cross section 22.37 pb at $\sqrt(s) = 8$ TeV [79]).

Single top quark events with additional jets, where the top decays to bW and the W to $\mu\nu$ also contribute as a background.

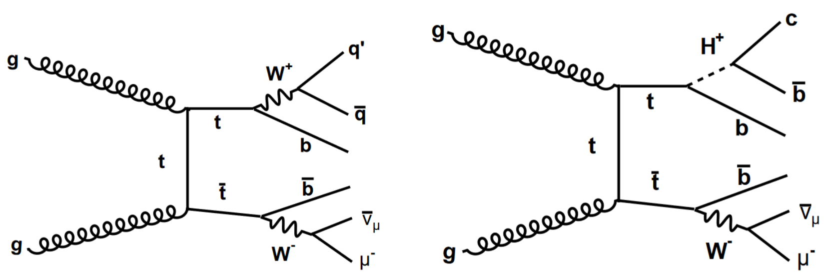
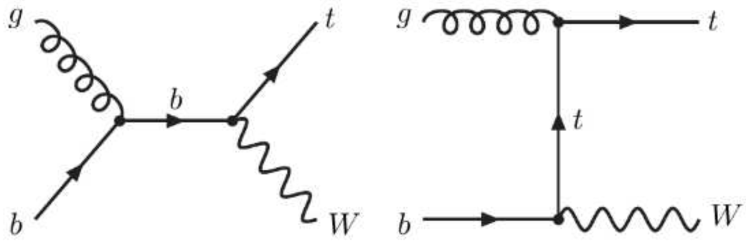


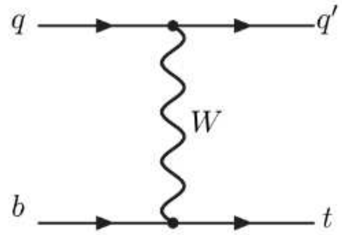
Figure 4.2: Feynman diagrams for the SM background (left) and for signal sample

Additional backgrounds are Z+jets, where the Z decays to two muons, one of which goes unidentified, and the production of a W boson along with four or more jets, with the W boson decaying to a muon and a neutrino. Another minor sources of background are QCD multijet events with muons arising from a heavy-flavored meson decay, diboson production with additional jets, production of a SM Higgs boson, or of an electroweak boson W or Z in association with a $t\bar{t}$ pair.

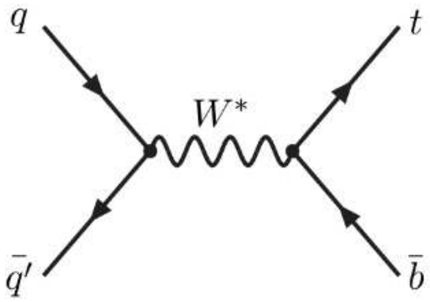
All sources of background, including QCD, are estimated using dedicated MC samples.



(a) Top production associated with W boson.



(b) t-channel.



(c) s-channel.

Figure 4.3: Feynman diagrams for single top production.

4.2 Data and Simulation Samples

The data collected with the CMS experiment at the center of mass energy of 8 TeV are used in the analysis. The analyzed runs are listed in Table 4.1 with their integrated luminosity. The whole dataset corresponds to a total integrated luminosity of 19.7 fb^{-1}.

Table 4.1: List of CMS datasets for pp collisions at $\sqrt{s} = 8$ TeV

Data set	Luminosity (fb^{-1})
Run2012A	0.886
Run2012B	4.435
Run2012C	7.125
Run2012D	7.246

Signal samples for the process $t\bar{t} \to bH^{+}\bar{b}W^{-}$ are generated using PYTHIA [80]. The generated process is $pp \to t\bar{t} \to H^{+}bW^{-}\bar{b} \to c\bar{b}b\bar{b}X$, where one of the top quarks is forced to decay to H^{+} and the other one to W^{-}; the H^{+} is then forced to decay to $c\bar{b}$. Since the decay $H^{+} \to c\bar{b}$ final state is not directly implemented in PYTHIA, the signal samples were produced as follows.

PYTHIA was used to generate events at parton level where the charged Higgs was forced to decay to $c\bar{s}$ instead of $c\bar{b}$. These events were converted in Les Houches Event (LHE) format [81]. In each event, the $\bar{s}$ quark was then replaced with a $\bar{b}$, and the four momenta of the charged Higgs decay products were adjusted in order to preserve energy and momentum conservation laws. The privately prepared LHE files were then used as a starting point for CMS official production of the samples. The signal samples were produced for H^{+} masses of 90, 100, 110, 120, 130, 150, 155 and 160 GeV.

For the backgrounds, standard centrally-produced Monte Carlo samples are used. A list of the backgrounds considered in the analysis with the corresponding cross section is presented in Table 4.2.

The $t\bar{t}$ sample is of particular importance because it represents by far the largest background. The default centrally produced Monte Carlo sample is used for this background. The Madgraph 5 generator [82] is used to produce the matrix-element ME at LO for $t\bar{t}$ plus up to three extra partons, which are then showered by PYTHIA. In the analysis this sample is split into five exclusive subsamples depending on the number and the flavor of additional jets produced along with the $t\bar{t}$ pair, in order to allow the study of the behavior of the different $t\bar{t}$ components separately.

To perform the splitting, all the reconstructed jets having pt > 20 GeV and $|\eta|$ <2.4 not originated by the decay of the $t\bar{t}$ pair are considered, the flavour of the corresponding matched parton was checked. The following subsamples are identified:

- $t\bar{t}bb$ - events with at least two additional b-jets

- $t\bar{t}bj$ - events with one additional b-jet and at least one additional jet which is not a b-jet
- $t\bar{t}cc$ - events with at least one additional c-jet. $t\bar{t}cj$ events are included here because the separation of light and c-jets is difficult.
- $t\bar{t}qq$ - events with at least two additional jets which are not b or c-jets.
- $t\bar{t}$ *other* - events not belonging to any of the above

The W+jets and Z+jets processes are produced using Madgraph [82]. In order to maximize the available statistics, exclusive W+bb jets and Z+bb jets samples are used along with the corresponding inclusive samples. Events with a W/Z boson plus two b jets are vetoed in the inclusive W/Z+jets samples in order to avoid double-counting. The NNLO cross sections are used to normalize the MC samples. The single top processes are produced using POWHEG [83], while the diboson samples are produced using PYTHIA.

Table 4.2: List of Monte Carlo estimated backgrounds with their cross sections

Process	Cross section (pb)
ttbar	259.2
Single top $\bar{t}$ prod. s-channel	1.76
Single top $\bar{t}$ prod. t-channel	30.7
Single top $\bar{t}$ prod. tW-channel	11.1
Single top t prod. s-channel	3.79
Single top t prod. t-channel	56.4
Single top t prod. tW-channel	11.1
W+jets	36257
DY+jets	3504
W+bb	377.6
Z+bb	76.7
Diboson WZ	12.63
Diboson ZZ	5.196
Diboson WW	33.61
$t\bar{t}$+Z	0.2057
$t\bar{t}$+W	0.232
$t\bar{t}$+H(125GeV)	0.133

4.3 Event Selections

Signal events $t\bar{t} \to b(H^+)\bar{b}W^- \to b\bar{b}(c\bar{b})\mu\nu$ contain one prompt muon, one neutrino, and four quarks, three of which are b-quarks. Therefore a final state consisting of a muon, at least four jets and missing transverse energy is expected. This section describes the selection of the final state objects and events to be used in the analysis.

4.3.1 Basic Selections

Events are selected using an isolated single muon trigger with a transverse momentum p_T threshold of 24 GeV and upper pseudorapidity $|\eta|$ threshold of 2.1.

For each event, a number of quality requirements have been applied. Events are first required to contain a well reconstructed primary vertex, having longitudinal position along the z axis less than 24 cm and impact parameter less than 2 cm.

Due to the high luminosity of 2012 running conditions, more than one proton-proton interactions per bunch are expected, giving rise to the so-called pileup interactions. The pileup activity causes an increase in the number of primary vertices in the events. The average number of reconstructed vertices during the integrated 2012 data taking period was approximately 21. The vertex with the maximal Σp_T^2 of associated tracks associated is chosen as the primary vertex related to the hard scattering.

Pileup interactions are usually simulated in MC as multiple minimum bias events overlayed to the hard scattering interactions. The expected pileup distribution can be calculated using the information provided by CMS about the bunch-by-bunch luminosity measurements and the total p-p interaction cross section. Although Monte Carlo samples are produced with a distribution of the number of pileup interactions that is meant to roughly cover the experimental conditions for the 2012 data-taking, the number of pp interactions per bunch-crossing in simulations does not match perfectly the data. As a consequence, there is a difference in the number of reconstructed primary vertices between data and MC simulated samples. To bring them to consistency, each sample must be reweighted to match the number of vertices distribution found in data.

For each events, further quality cuts are used in the analysis: For events with at least 10 tracks, at least 25% of high purity tracks has been required [1]. Moreover, events with high calorimeter noise in the HCAL barrel or endcap have been discarded.

4.3.2 Muon reconstruction, identification and isolation

Muons must pass a set of offline quality cuts, listed in Table 4.3. Muons are required to have a minimum p_T of 25 GeV with a pseudorapidity $|\eta| < 2.1$. To reject non-prompt muons, (cosmics, muons from b hadron decays etc.) the transverse impact parameter relative to the beam axis is required to be smaller than 0.2 cm, while the longitudinal position of the muon track at its closest

[1] Due to the dense environment in CMS events, a preliminary track cleaning is needed to reduce fake reconstructed tracks; several quality cuts are imposed on tracks, the most important ones being the track fit-χ^2, the longitudinal and transverse impact parameters (and their significance), the number of crossed layers with measurements, track η and p_T. The full *high purity* requirements are described in detail in [65]

Table 4.3: Muon identification criteria

ID	Threshold
p_T	25 GeV
$\lvert\eta\rvert$	<2.1
d_{xy}	<0.02 cm
d_z	<1 cm
Global and Particle flow muon	-
number of matched stations	>2
χ^2 global track fit	<10

approach to the beam line is required to lie within 1 cm from the position of the hard scattering vertex. The isolation of muon candidates is defined according to the particle flow (PF) isolation algorithm. In such approach, a cone with radius ΔR, calculated as

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \tag{4.1}$$

is defined around the muon track. Inside the cone the energy deposited by charged hadrons (E^l_{CH}), neutral hadrons (E^l_{NH}) and by photons (E^l_γ) are computed and a new variable I^l_{rel} is defined:

$$I^l_{rel} = \frac{E^l_{CH} + E^l_{NH} + E^l_\gamma}{p^l_T} \tag{4.2}$$

Due to the not negligible pileup contribution a correction must be applied to the isolation variable to account for effects of additional interactions. The contribution of energy deposited in the isolation cone by charged particles not associated to the primary vertex is first calculated, to correct jet energy from pileup contribution. This amount is multiplied by the ratio of neutral to charged hadron production in the hadronization process of pileup interactions and subtracted by the isolation variable [63]. Corrected isolation is required to be less than 0.12, with ΔR being 0.3. In Figure 4.4 the histogram for the muon relative isolation is shown for data and simulation samples after all selection steps.

4.3.3 Second lepton veto criteria

To veto any additional muon in the event, a loose muon identification criterion is used, consisting in selecting any PF muon, as well as a global or a tracker muon, satisfying the isolation requirement $I_{rel} < 0.3$.

Electrons are also selected for veto purposes. A loose electron identification criterion is used that corresponds to an identification efficiency of 95%. The following cuts on the variables previously defined in Chapter 3 are used for the selection:

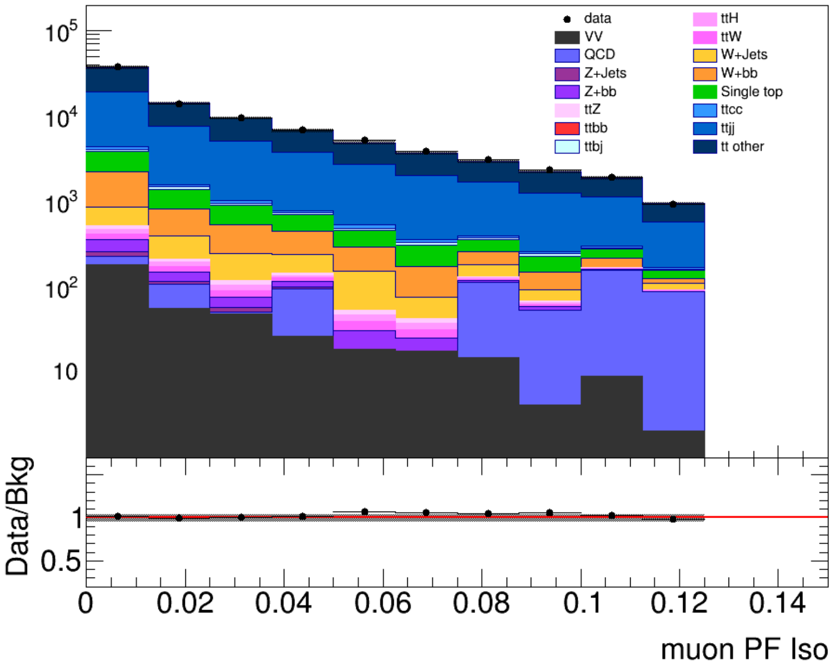


Figure 4.4: Muon isolation in data and MC after the full set of selections

- $\Delta\eta_{in} < 0.007$ (barrel),0.01 (endcap)
- $\Delta\phi_{in} < 0.8$ (barrel),0.7 (endcap)
- $\sigma_{i\eta i\eta} < 0.01$ (barrel),0.03 (endcap)
- $H/E < 0.15$ (for barrel only)
- d_{xy} with respect to the primary vertex <0.4 mm
- d_z with respect to the primary vertex <2 mm

The electron isolation is calculated as in Eq. 4.2 with $\Delta R < 0.3$. Any event with at least one further muon or electron passing the selection criteria above, having p_T >10 GeV, $|\eta|$ <2.5, I_{rel} <0.3 is vetoed. This second lepton veto rejects most of the events from Z+jets and from dileptonic SM $t\bar{t}$ events.

4.3.4 Jet reconstruction and selection

Jets clustering is performed using the anti-kt algorithm [66] with a size parameter of 0.5. Charged hadrons identified by PF algorithm and associated to pileup (PU) activity, isolated muons and electrons are not passed as input of clustering algorithm. Jet energy is corrected using the factorized approach described in section 3.1.5, where each level of correction takes care of different effects; each correction implies the application of a momentum scale factor, which depends on various related jet quantities (p_T, η...), aiming to make energy response function flat as function of jet features and correct the disagreement in energy resolution observed in data and MC. Finally, jets energy is also corrected taking into account neutral particle contribution from pileup collisions for which no particle subtraction is performed, due to large uncertainties on the originating primary vertex.

A set of minimal quality cuts is applied to the selected PF jets:

- Jet identified as not coming from PU with loose working point
- Neutral Hadron Energy Fraction < 0.99
- Neutral electromagnetic (EM) Energy Fraction < 0.99
- More than 1 constituent constructs a jet
- Charged hadron energy fraction > 0
- Charged multiplicity >0
- Charged EM energy fraction < 0.99
- Muon energy fraction < 0.8

Four jets passing these cuts and having p_T above 25 GeV and $|\eta| < 2.4$ are requested for each event. Moreover, at least three of the selected jets must be tagged as b-quark initiated jet by the Combined Secondary Vertex algorithm, configured at a medium working point (CSVM), corresponding to a discriminating threshold of 0.679. This configuration corresponds to a misidentification probability for light parton jets of approximately 1% for the jet p_T in the range 80 to 120 GeV. The corresponding b-tagging efficiency is about 70% [84]. This requirement strongly suppresses W+jets and QCD multijet backgrounds.

4.3.5 Missing transverse energy

In particle flow reconstruction missing energy E_T^{miss} is calculated as in equation 3.6. Jet energy corrections must be propagated for E_T^{miss} computation. Events are required to have E_T^{miss} >20 GeV. This cut suppresses the Z+jets and the QCD multijet backgrounds.

4.3.6 Data Monte Carlo corrections

Simulated events are assigned event-dependent weights which correct the efficiency and acceptance predicted by the simulation for a variety of effects. These include:

Generator-level corrections: following Ref. [85], the $t\bar{t}$ events simulated by the Madgraph generator are reweighted depending on transverse momentum of the generated top and antitop quarks, as to match the measured differential top-p_T distribution. This reweighting affects both the shape and the normalization of the $t\bar{t}$ sample, and comes along with a systematic error.

Pile-up: as mentioned in Section 4.3.1 the pile-up simulation in the Monte Carlo samples does not reproduce the pile-up profile integrated over the whole data-taking period. For this reason, a per-event weight is assigned to all simulated events based on the true-level number of pileup interactions. The pileup distribution in the MC samples used in this analysis before reweighting is shown in figure 4.5(left). The target pileup distribution in data is derived from the measurement of the istantaneous luminosity over all the data-taking period and is shown in figure 4.5(center). The pileup distribution in MC after reweighting is shown in figure 4.5(right). The agreement between data and MC in the number of reconstructed primary vertices has been checked after different selection steps. For instance in figure 4.6 the distribution of reconstructed vertices is shown for data and MC after the requirement for at least four jets. The pileup reweighting procedure improves consistently the data/MC agreement in the reconstructed vertices distribution. The remaining discrepancy is accounted for by including the uncertainty on the pileup weight among the systematic sources included in the analysis, as described in section 4.5.

Data/MC correction to muon reconstruction and identification efficiency: the muon efficiency in the simulation is corrected using multiplicative scale-factors, measured differentially in the pseudo-rapidity and transverse momentum of the leptons

Trigger efficiency: The trigger efficiency in the Monte Carlo simulation is accounted for by applying the measured trigger-turn-on curve as a per-event weight. The weight is measured differentially in the pseudo-rapidity and transverse momentum of the triggering muons. For single-muon event selected by a single-muon trigger, the weight is defined as $w_{tr} = f_l(p_T^l, \eta_T^l)$, where f_l is measured using a tag-and-probe technique [86].

Data/MC correction to the momentum of jets: as mentioned in section 4.3.4, the jet energy scale (JES) is corrected for data/MC discrepancy in the detector response. An extra-smearing of the jet momentum in simulated events is necessary to reproduce the jet energy resolution (JER) measured in data [87].

Data/MC correction to b-tagging: the b-tagging is a key variable for this analysis. For each jet within the tracker volume, the CSV tagger calculates a continuous output. Since b-tagging efficiencies are found to be different in MC and in data, it is necessary to apply scale factors to MC events in order to predict correctly the event yield surviving a certain selection on the number of b-tagged jets. The probability of having a given number of b-tagged and not b-tagged jets in a MC event is given by

$$P(MC) = \prod_{i=tagged} \epsilon_i \prod_{j=nottagged} (1 - \epsilon_j), \tag{4.3}$$

where ϵ_i is the b-tagging efficiency of the i-th jet and is dependent not only on the jet flavor (the probability for the b-tagging algorithm to tag a b-jet is of course higher than for a c-jet or a light jet), but also on the kinematical parameters of the jet, i.e. the jet transverse momentum and pseudorapidity. On the other hand, the same probability in data can be written as:

$$P(DATA) = \prod_{i=tagged} SF\epsilon_i \prod_{j=nottagged} (1 - SF\epsilon_j), \tag{4.4}$$

being $SF = \epsilon_{DATA}/\epsilon_{MC}$ the ratio between b-tagging efficiency in data and in MC. Each simulated event must be weighted by a global scale-factor defined by:

$$w = \frac{P(DATA)}{P(MC)}. \tag{4.5}$$

The scale factors SF have been centrally measured by CMS, separately for light and heavy-quark jets, in a Drell-Yan and $t\bar{t}$-enriched sample respectively. The b-tagging efficiencies and the mistagging efficiency in MC for c-jets and light jets have been measured for each background and signal sample, before applying any selections, in bins of transverse momentum and pseudorapidity. Figure 4.7 shows the b-tagging efficiency and the mistagging efficiency for c-jets and light-jets measured on the $t\bar{t}$ background sample.

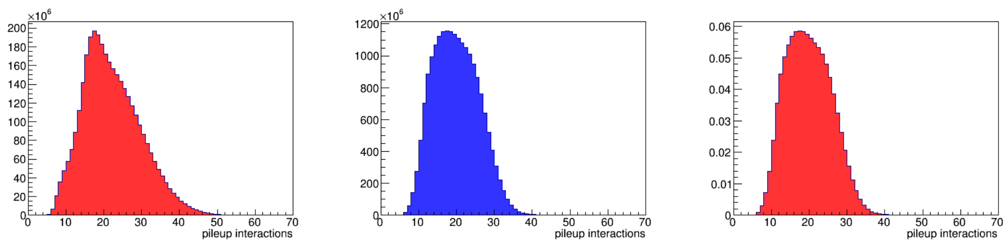


Figure 4.5: Number of pileup interactions in MC(left), in data (center) and in MC after reweighing (right)

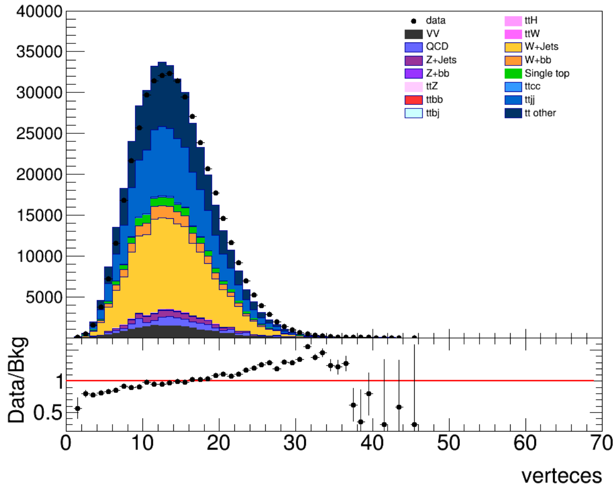


Figure 4.6: Number of reconstructed primary vertices in data and MC after requiring the presence of at least four jets in the event

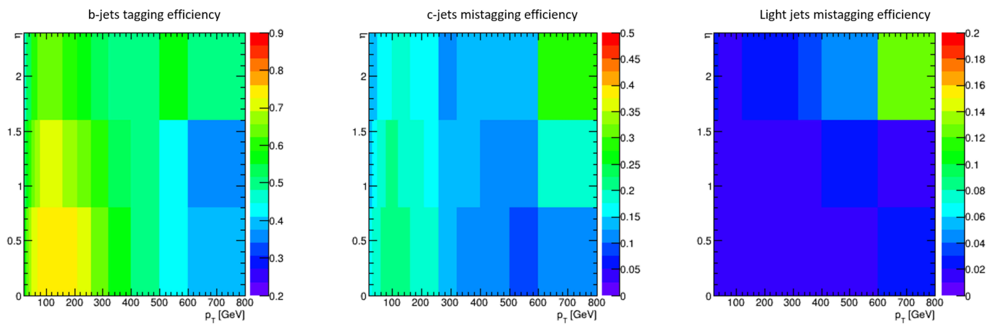


Figure 4.7: b-tagging efficiency and the mistagging efficiency for c-jets and light-jets measured on the $t\bar{t}$ background sample.

4.3.7 Selection study

In Figure 4.8 the event yields for the background samples and data are compared at different selection stages. In Figure 4.9 the same cutflow plot is presented for signal samples under different mass hypotheses for the charged Higgs boson. The signal is normalized assuming a branching fraction x of the top quark to charged Higgs of 10%. Since in a $t\bar{t}$ pair event both the top and antitop quark can decay into a charged Higgs boson with a probability x, the MC signal sample is normalized to a cross section given by:

$$\sigma_{signal} = 2x(1-x)\sigma_{t\bar{t}} \tag{4.6}$$

The event yields after each selection step are also given in Table 4.4 and Table 4.5 for backgrounds and signal samples.

The selection efficiency for signal samples decreases as the charged Higgs mass grows. This is due to the fact that the heavier the charged Higgs is, the softer will be the b-jet produced in the $t \to H^+ b$ decay. As the charged Higgs mass increases, this jet is more likely to be below the momentum threshold of 25 GeV imposed for event selection. Moreover, the softer the jet is, the less efficient the b-jet identification will be.

In order to check data-MC agreement, the shapes of different variables in data and MC after the second and third b-tagging requirement were compared, including the p_T and η distribution for muons, for the 4 leading jets in the events, the E_T^{miss} distribution, the jet multiplicity and the distribution of reconstructed primary vertices. Some results can be observed in Figures 4.10 and 4.11.

In general, we observe a good data-MC agreement after every selection steps. The small (approximately 5%) discrepancy observed after the muon selection is mostly due to uncertainties on the W+jets and QCD backgrounds.

After requiring at least 3 b-tagged jets an excess in the data with respect to MC of approximately 15% is observed. This is due to theoretical uncertainties that affect MC $t\bar{t}$ sample. Since in SM $t\bar{t}$ events only two b-jets are produced in the MC decay chain, the third b-tagged jet must come either from the mistagging of one of the light jets produced in the decay of the W boson, or from extra jets produced along with the $t\bar{t}$ pair. The simulation of these extra jets in MC brings in large uncertainties on the parton shower process and its matching with matrix element simulation. Similar discrepancies have already been observed in SM $t\bar{t}$ analyses such as [88].

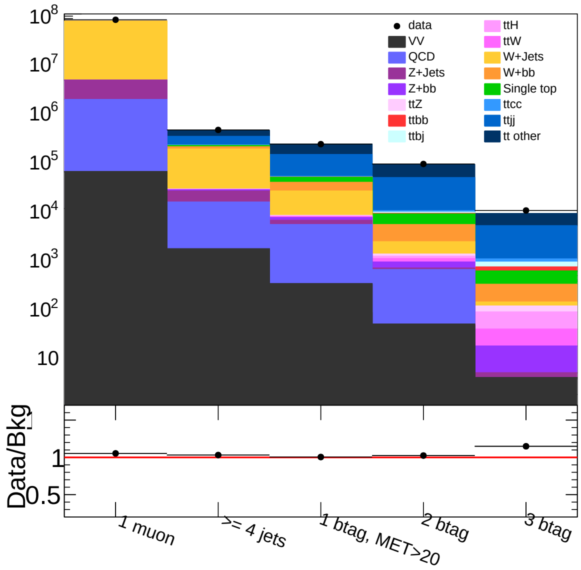


Figure 4.8: Event yields after different selection steps in MC and data

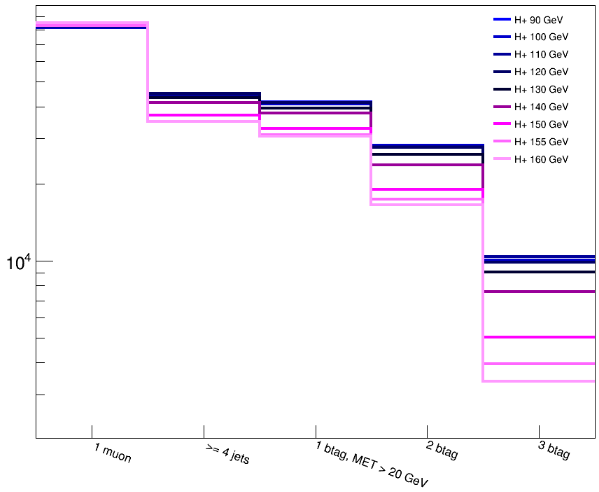


Figure 4.9: Event yields after different selection steps for MC signal sample. Signal is normalized assuming a branching fraction of the top quark to charged Higgs of 10%.

Table 4.4: Event Yields for MC background samples and data.

	1 muon	$\geq$ 4 jets	MET$\geq$20 , $\geq$ 1 CSVM	$\geq$ 2 CSVM	$\geq$ 3 CSVM
tt bb	274.82	256.939	248.582	204.292	116.003
tt bj	856.452	717.436	660.521	449.699	166.186
tt cc	1795.29	1537.1	1282.1	636.753	141.914
tt jj	140766	106645	85587.2	35371.8	3681.37
tt other	308937	98505.2	83380.4	38609.6	3673.99
tt	452629	207662	171159	75272.2	7779.46
diboson	59730.2	1576.04	305.388	45.6953	3.68693
single top	123401	12520.3	9517.54	3238.3	260.513
W+jets	6.61931e+07	145791	16340.4	967.182	22.1149
W+bb	786288	19735.5	12048.5	2720.12	167.192
Z+jets	2.61807e+06	9826.51	1095.5	50.1257	0.910375
Z+bb	53352	1269.42	824.134	202.859	11.6486
QCD	1.72464e+06	12612.8	4614.56	544.668	0
ttZ	383.523	284.134	238.683	118.053	25.9219
ttH	255.804	212.954	189.54	119.05	45.0074
ttW	557.201	402.328	329.638	144.87	19.6184
total bkg	7.20124e+07	411893	216663	83423.1	8336.07
data	7.5865e+07	424761	217736	85474	9578

Table 4.5: Event Yields for MC signal samples

	1 muon	$\geq$ 4 jets	MET$\geq$ 20 , $\geq$ 1 CSVM	$\geq$ 2 CSVM	$\geq$ 3 CSVM
H^+ 90 GeV	84722.8	45061.2	41752	28216.7	10016.3
H^+ 100 GeV	81583.1	44094.2	40848	27935.5	10069
H^+ 110 GeV	81642	44345.1	41393.6	28136	10354.4
H^+ 120 GeV	83467.4	45064.3	41532.1	27661.2	9887.06
H^+ 130 GeV	83126.1	43328.8	39467.4	26024.7	9018.13
H^+ 140 GeV	83952.1	41556.4	37683.2	23610.5	7560.85
H^+ 150 GeV	83584.8	36940.3	32864.6	18995.8	5029.97
H^+ 155 GeV	83068.5	34907.1	31130.6	17361	3969.83
H^+ 160 GeV	84952.7	34912.3	30538.2	16441.2	3375.23

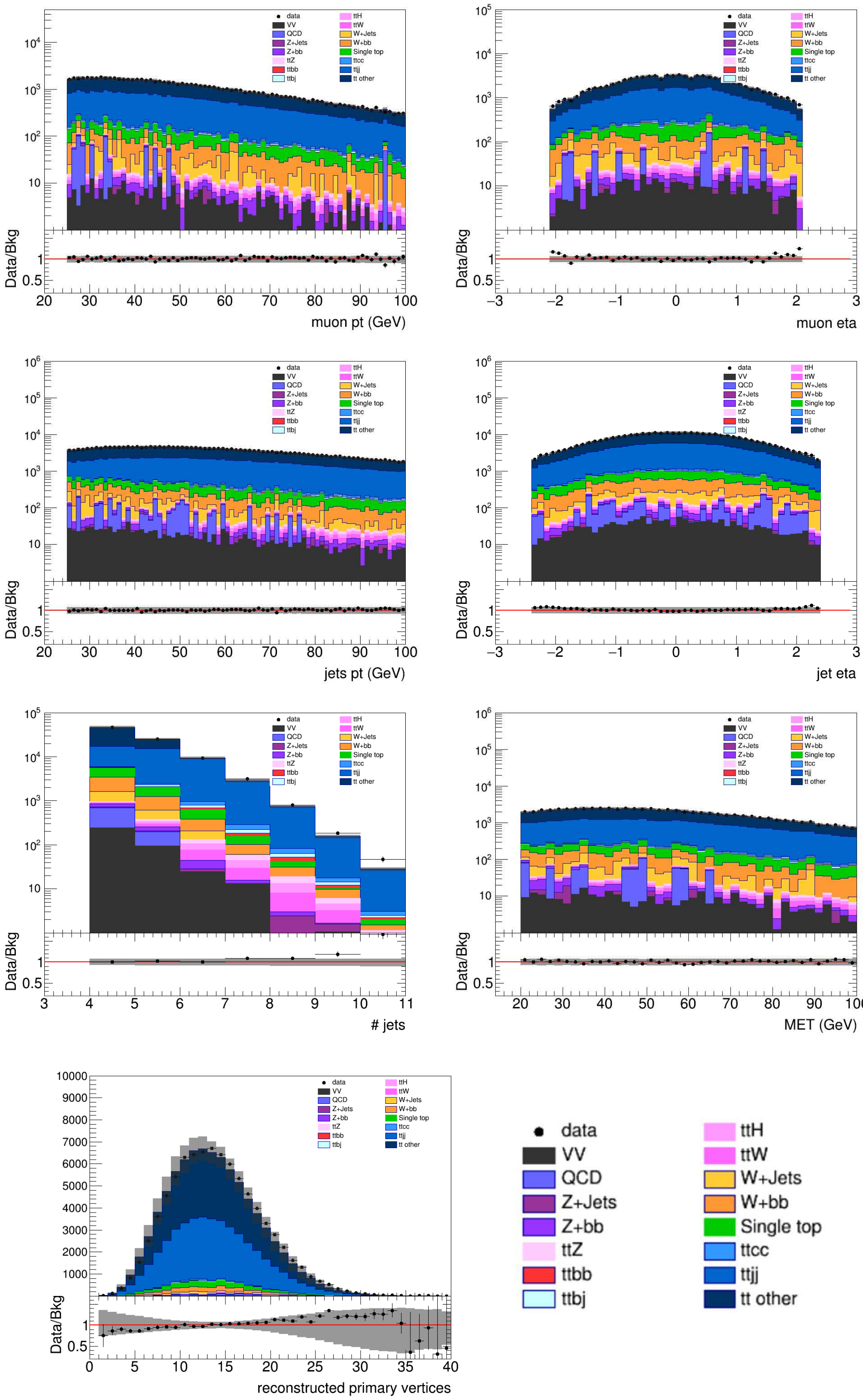


Figure 4.10: Control plots after the second b-tagging requirement

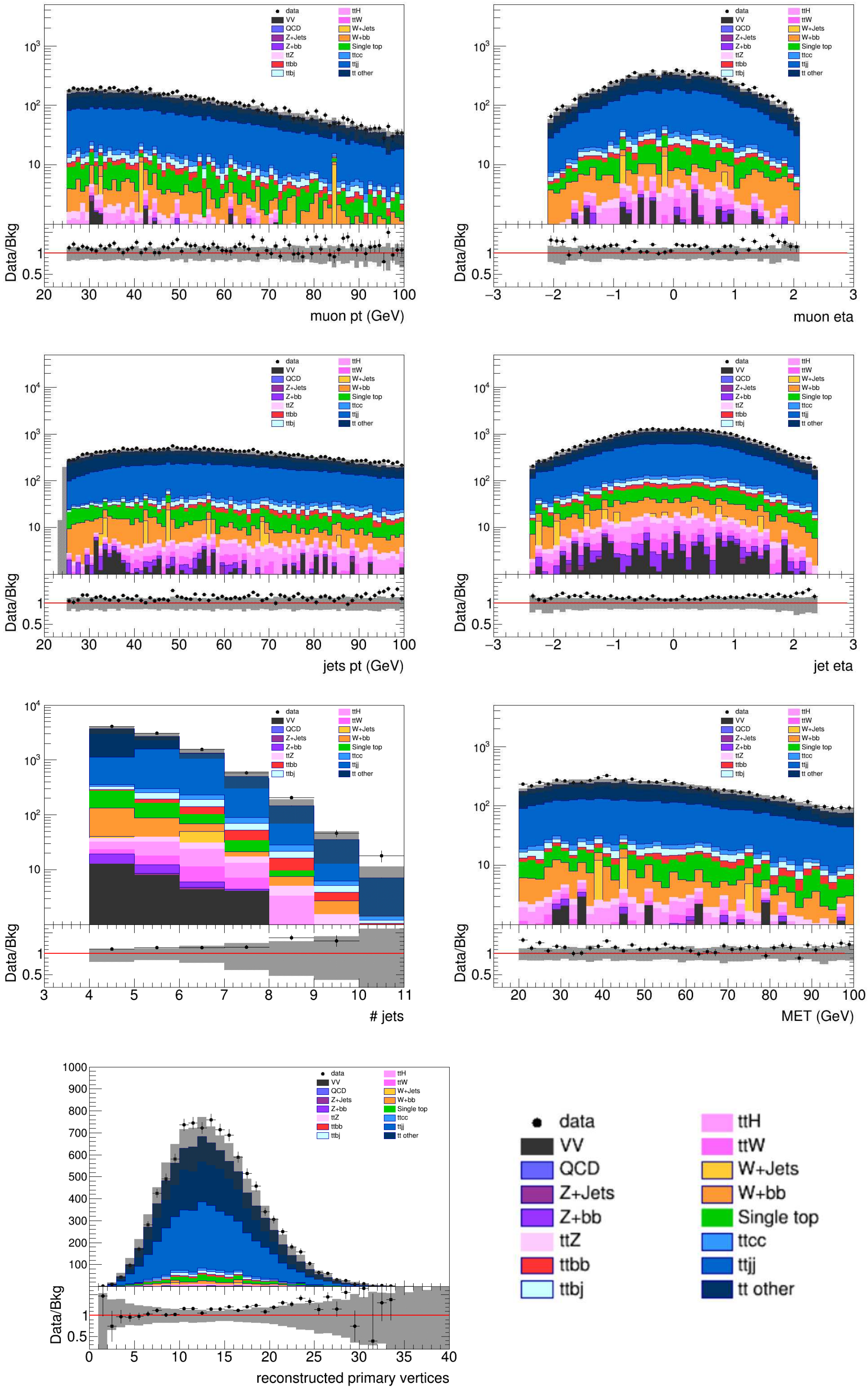


Figure 4.11: Control plots after all selections

4.4 Mass reconstruction

4.4.1 Kinematic Fit

An advanced kinematic fit is employed to fully reconstruct $t\bar{t}$ event kinematics from the final state and to improve the W/H^+ mass resolution. The core of the fit algorithm was implemented in the official CMSSW software, and it is usually adopted in $t\bar{t}$ analyses to improve the resolution on the measurement of the top mass. The detailed description is available in Ref. [89]. In this analysis the fit was adapted and optimized to reconstruct the charged W/H^+ masses.

The fit constrains the event to a hypothesis for the production of two top quarks, each one decaying to a W boson and a b quark. One of the W boson decays into a muon-neutrino pair, while the other W boson (H^+ in case of signal) decays into a quark-antiquark pair. When reconstructing an event, it is not possible to know a priori which observed jet corresponds to which parton in the $t\bar{t}$ event topology. Moreover due to QCD radiative effects, jet merging and splitting during reconstruction, and jet reconstruction inefficiencies, the observed jets may have no one-to-one correspondence with the unfragmemted partons from the $t\bar{t}$ decay. Nevertheless, the fitted mass m_{fit} constructed from the observed jets is correlated with the true W mass and thus can be used for a measurement. The input to the kinematic fit are the four-momenta of the muon and jets passing the selection requirements, the missing transverse energy, and their respective resolutions. Since we do not know the correspondence between jets and partons, all the possible permutations of jet assignments are considered, with the following constraints:

- all the reconstructed jets having pt$>$ 24 GeV and $|\eta| < 2.4$ are considered for the kinematic fit
- only jets that pass the b-tagging requirement are considered as possible candidates for the b quarks in the $t\bar{t}$ hypothesis.
- since we are looking for a H^+ decaying to a $c\bar{b}$ pair, we consider only jet permutations where exactly one of the jets associated to the W/H^+ decay passes the b-tagging requirement

Since we are interested in the reconstruction of the mass of the hadronically-decaying W/H^+ , the reconstructed mass of the two top quarks are constrained in the fit to 172.5 GeV, while the mass of the leptonic W is constrained to 80.3 GeV. The reconstruction of the mass of the hadronically-decaying W/H^+ is instead left with no constraints, and it is the final observable of this analysis.

For each possible jet permutation, the measured momenta and directions of jets, of the muon and the missing energy are arranged in a vector x^m. The fit program then minimizes the following χ^2:

$$\chi^2 = (x - x^m)^T G(x - x^m), \tag{4.7}$$

where x is the vector of the fit observables and G is the inverse error matrix that is given by the resolution of the observables. This χ^2 is minimized subject to the kinematic constraints on the top and the leptonic W masses, using the method of Lagrange multipliers. If the minimization does not converge, the combination is rejected. To increase the fraction of correct permutations, after the fit, we consider only combinations with goodness of fit (GOF) probability $P_{GOF} = exp(-\chi^2/NDF) > 0.2$, where the number of degrees of freedom (NDF) of the fit, given the number of constraints used in this analysis, is equal to one. Moreover, the combinations are taken into account only if the p_T of the jets returned by the fit is greater than the selection threshold of 25 GeV. These conditions restrict the number of possible parton-to-jet assignments per event.

For each surviving permutation, this method gives a fitted mass for the hadronic W boson and a χ^2.

Different strategies can be used to choose for each event the best jet permutation and thus the best fitted value for the W/H^+ mass. The most trivial would consist in selecting the jet combination having the best χ^2. However, the two jet permutations obtained by swapping the b-tagged jet associated to the decay of the H^+ candidate and the one associated to the decay of the top quark in the hadronic branch of the $t\bar{t}$ couple have a similar χ^2. As a consequence, considering just one of these two would lead to a loss of information.

The strategy that we have chosen thus consists in considering the two leading permutation in terms of χ^2, and choosing out of these two the one where the jet associated to top decay in the hadronic branch is softer. This choice is driven by the fact that for the signal, the p_T spectrum of this jet is softer than that of the jets coming from the decay of the H^+ (Figure 4.12).

In Figure 4.13, the expected upper limit on the branching fraction of top quark into charged Higgs obtained with this configuration of the kinematic fit is compared to the result obtained by choosing the best permutation only according to the χ^2. Considering the transverse momentum of the b-jet provides a slight improvement of the signal to background discrimination for high H^+ masses. This choice is therefore used to derive the expected limit shown in subsection 4.6.

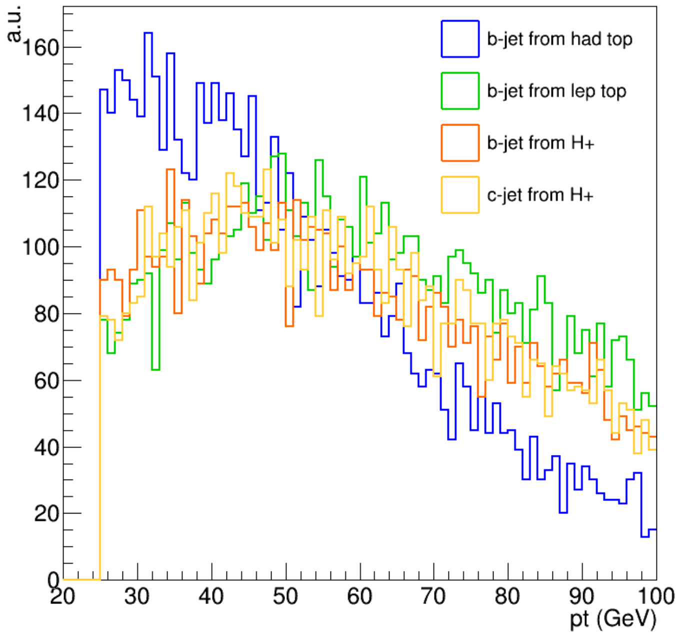


Figure 4.12: Reconstructed p_T spectrum for the 4 jets from $t\bar{t}$ semileptonic decay at truth-matched level, for a charged Higgs mass of 120 GeV. The momentum distribution for the b quark from the hadronic top decay and leptonic top decay are shown in blue and green respectively. The distribution for the charm and bottom quarks from charged Higgs decay are shown in orange and yellow respectively.

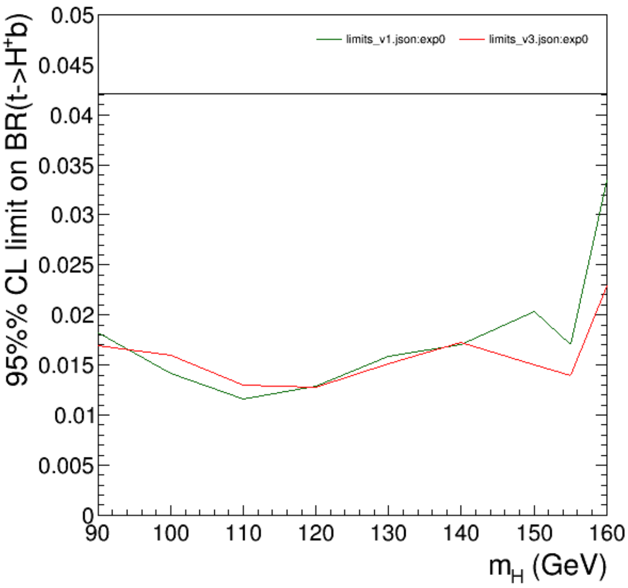


Figure 4.13: Expected 95% C.L. upper limit on the branching fraction of top quark into charged Higgs, with two different configurations of the kinematic fit: chosing the best permutation according to the fit probability (green), chosing the best permutation according to the transverse momentum of the jet associated to the hadronic top decay (red).

The fit probability distribution obtained in data and MC is shown in figure 4.14. Figure 4.15 shows the result of the fit on the backgrounds and on the signal samples under different hypotheses for the H^+ mass. The signal is normalized assuming a branching ratio for the top quark to charged Higgs equal to 10%, and assuming that the charged Higgs can decay exclusively to $c\bar{b}$ ($BR(t \to H^+b) = 0.1$, $BR(H^+ \to c\bar{b}) = 1$). These templates are used for the binned maximum-likelihood fit to extract possible signal, as described in subsection 4.6. Looking at the templates, it is clear that the yield of signal events after the fit decreases as the charged Higgs mass grows. This happens because the the b-jet coming from the top decay in the hadronic branch gets softer and softer at high charged Higgs mass, thus making less efficient the b-tagging procedure. Since the b-tagging is involved also in the kinematic fit algorithm, the fit efficiency decreases at high H^+ mass.

Figure 4.16 shows the comparison of the di-jet mass shape obtained through the kinematic fit in data and in the MC sample.

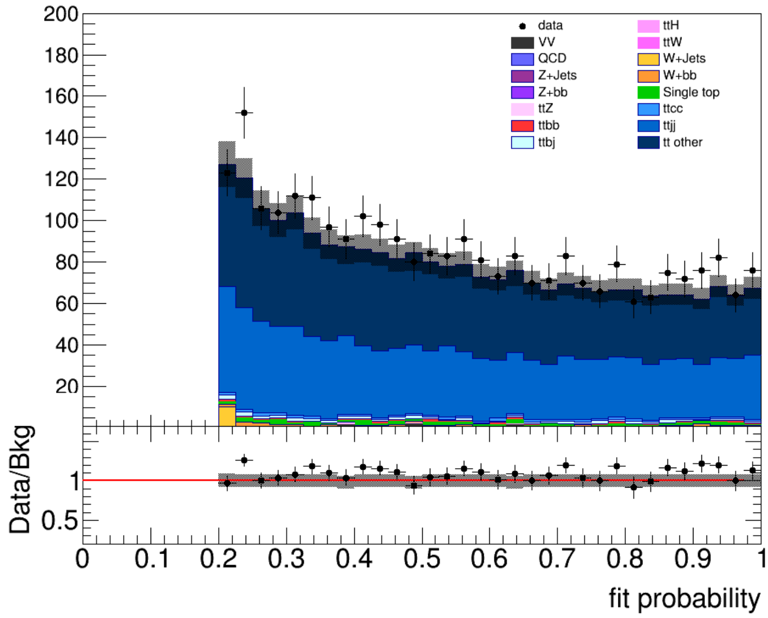


Figure 4.14: Fit probability distribution.

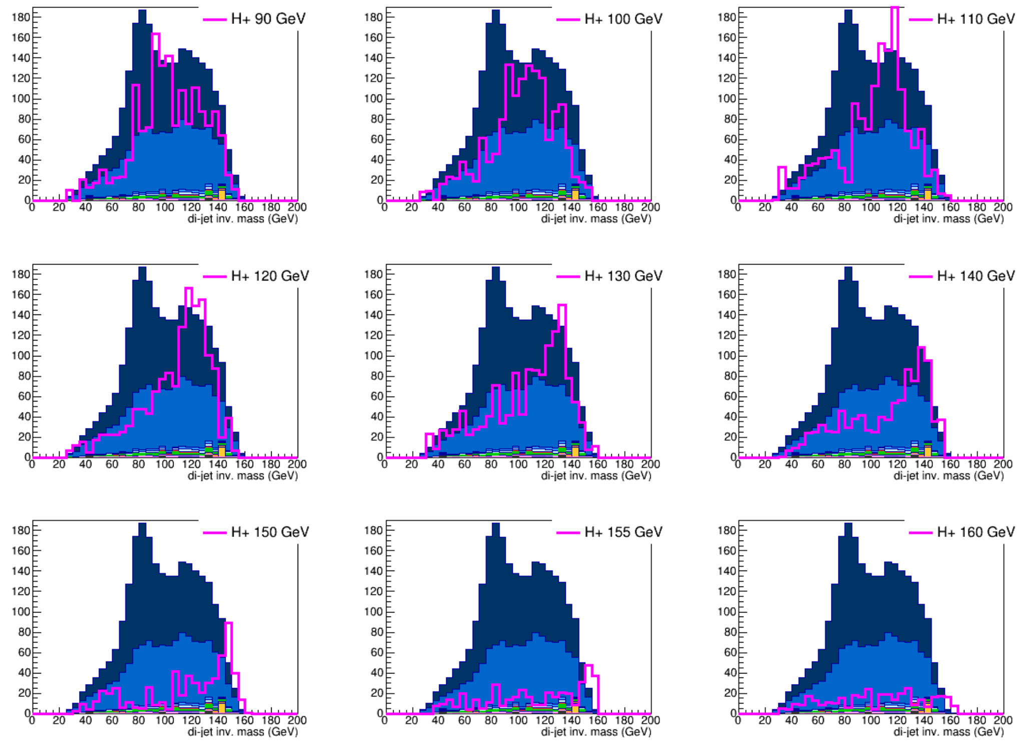


Figure 4.15: Mass templates used for the kinematic fit. Each plot shows the invariant mass of the charged Higgs candidate resulting from the kinematic fit of the MC background samples. In addition, the magenta line shows the same distribution for signal samples under different mass hypothesis for the charged Higgs boson, ranging from 90 to 160 GeV. Signal is normalized assuming a branching ratio of the top quark to charged Higgs equal to 10%, and a branching ratio of the charged Higgs to $c\bar{b}$ of 100%.

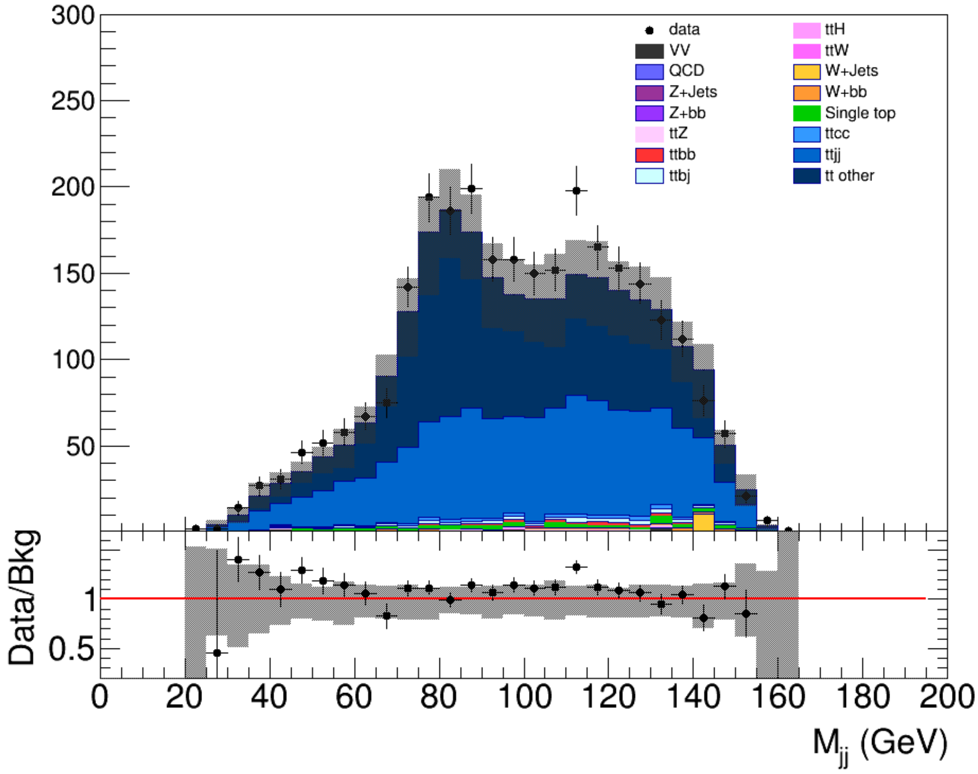


Figure 4.16: Dijet mass reconstructed through the kinematic fit in data and MC.

Table 4.6: Fraction of categories in $t\bar{t}$ sample

	after all selections	after the fit
ttbb	0.015	0.010
ttbj	0.021	0.014
ttcc	0.018	0.012
ttjj	0.473	0.431
ttother	0.472	0.532

4.4.2 Control plots after the kinematic fit

For the events surviving the kinematic fit, the data-MC consistency is checked again looking at kinematic variables (Figure 4.17).

The fit has different efficiencies on the events belonging to different $t\bar{t}$ categories. In particular, the fraction of events with additional jets is reduced after the fit (Table 4.6). Since these categories are the most affected by MC theoretical uncertainties, an improvement of the data-MC consistency is observed with respect to control plots previously shown after the whole selection chain but before the kinematic fit. In particular, the 15% excess of events in data with respect to MC observed before the kinematic fit is reduced to 9% after the fit procedure.

4.5 Systematic uncertainties

- JES,JER, E_T^{miss} scale uncertainty. Jet energy scale is not exactly reproduced by MC samples thus jet energy scale (JES) in simulation it has to be corrected to fix these undesired effects. The uncertainties on jet energy corrections have in general a complex dependence on the jet p_T and η. The jet energy scale is one of the sources with the largest impact on likelihood parameters. Jet energy resolution is known to be underestimated of about 10% in MC samples. To take into account this effect, the reconstructed jet transverse momentum p_T^R is oversmeared, defining a new transverse momentum

$$p_T = max\big[0.0, p_T^{GEN} + f(p_T^R - p_T^{GEN})\big] \tag{4.8}$$

  where p_T^{GEN} is the jet momentum at generation level, while f is a scale factor measured as the resolution ratio between data and MC, and depends on the reconstructed jet pseudorapidity. Systematic uncertainties are taken into account replacing in Equation 4.8 a value for f which corresponds to +1/-1 σ variation of of jet energy resolution. Finally both the JES and JER uncertainties are propagated for the calculation of E_T^{miss}.

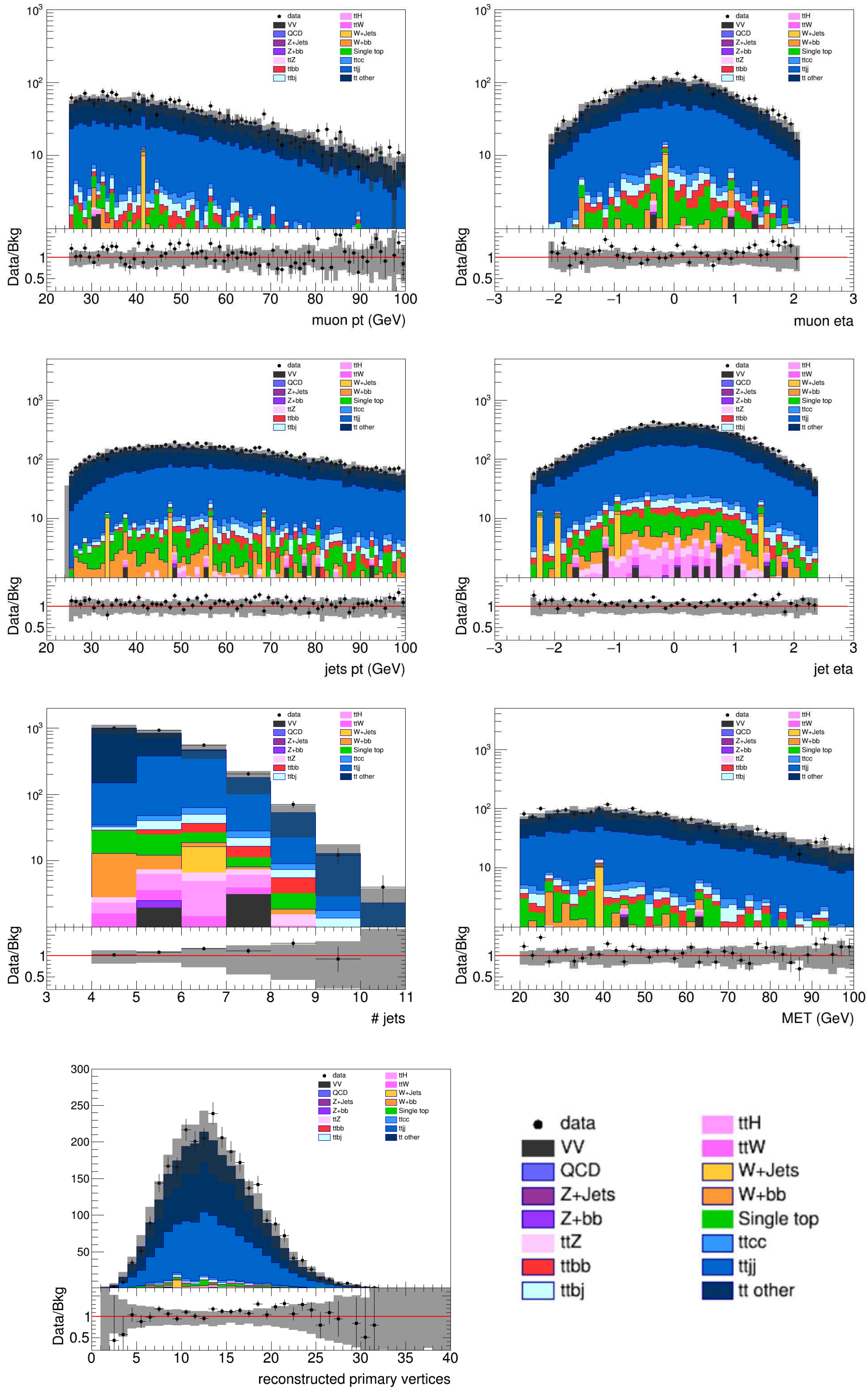


Figure 4.17: Control plots after the kinematic fit.

- uncertainty on b-tagging. The b-tagging scale factor uncertainty is one of the most important sources of systematics and it is treated as a shape uncertainty. In this analysis a $b-$ tagging CSV algorithm has been used; however the model of b-tagging is not perfect. The bias can be corrected reweighting events on the basis of their actual content in term of b, light quark and gluons, inferred using MC thruth. The weights are functions of jet p_T, η and flavor.

- Muon trigger, identification and isolation. The efficiency on the muon trigger, isolation and identification is biased. A unique value of 2% has been used to describe relative uncertainty in the efficiency of the muon trigger, identification and isolation as estimated by MC.

- Uncertainty on PU reweighting. MC is produced overlaying pileup interactions to each events. However each MC sample must be reweighted to match the true number of vertices distribution found in data. The weight uncertainty is dominated by the uncertainty on the total inelastic cross section and on the measured luminosity (both quantities are taken into account in the reweighting procedure). The total uncertainty can be obtained shifting the overall mean of interaction distribution in data and then performing the reweighting. One can do it recalculating the actual vertex distribution by varying the total inelastic cross section of 5% around the central value.

- Uncertainty due to top pt reweighting. It was estimated removing the top pt reweighting procedure.

- Uncertainty on the top quark mass. The top-quark mass is set to be 172.5 GeV in this analysis, and we use two different top mass samples to estimate this systematics. Two samples are generated with $m_t = 171.5$ GeV and 173.5GeV.

- ME-PS matching. In the $t\bar{t}$ simulation sample, matching threshold that interfaces the matrix element to the parton-showering is shifted up/down as described in Ref [89].

- $t\bar{t}$ event generator. Uncertainty on the Monte Carlo generator was estimated using two additional $t\bar{t}$ samples, the first one generated with MCatNLO and showered with HERWIG, the second generated with POWHEG and showered with Pythia.

- factorization/normalization scales. The uncertainty on the dijet mass template of the $t\bar{t}$ background due to missing beyond-LO terms is estimated by varying the renormalization and factorization scale by the conventional factor 2 up and down.

- Luminosity uncertainty. The uncertainty on the luminosity measurement is taken to be 2.6% .

- $t\bar{t}$ cross section. The theoretical value of the $t\bar{t}$ cross section is used to estimate the backgrounds of the SM $t\bar{t}$ with an uncertainty of 6.5%.

A summary of the different systematic uncertainties on signal ad backgrounds is presented in table 4.7. The effect of systematic uncertainties on the reconstructed di-jet mass is shown in Figures 4.18- 4.25 for a signal sample with charged Higgs mass equal to 120 GeV, for the $t\bar{t}$ background and for all the other sources of non-$t\bar{t}$ background. The uncertainty on the b-tagging scale factors is one of the leading source of systematics, although it changes mostly the overall event rate, while the di-jet mass shape is very slightly affected. Effects on the mass template is more evident for the theoretical uncertainties on the $t\bar{t}$ background, related to the generators (Figure 4.24), the top quark mass, the renormalization and factorization scales (Figure 4.25).

Table 4.7: Summary of the sources of systematic uncertainties for signal, $t\bar{t}$ background and non-$t\bar{t}$ background. The numbers show the size of the variation in the event rate corresponding to each systematic source. A symbol (s) is shown for the uncertainties whose effect on the di-jet mass shape has been accounted for in the fit.

Uncertainty	signal $m_{H+} = 120$ GeV	$t\bar{t}$	non-$t\bar{t}$
$t\bar{t}$ cross section	6.5%	6.5%	-
top quark mass	-	0.7%(s)	-
top pt reweighting	1.2%(s)	0.6%(s)	-
NLO-vs-LO shape	-	3%(s)	-
POWHEG vs madgraph	-	11%(s)	-
ME-PS matching	-	0.4%(s)	-
Renormalization and factorization scale	-	4%(s)	-
JES	3.5%(s)	3%(s)	10% (s)
JER	0.7%(s)	0.7%(s)	19%(s)
Btag scale for b/c jets	6.5%(s)	7%(s)	7%(s)
Btag scale for light jets	0.4%(s)	3.2%(s)	5%(s)
Pile-up reweighting	0.1%(s)	0.2%(s)	1%(s)
Muon scale factors	2%	2%	2%
luminosity	2.6%	2.6%	2.6%

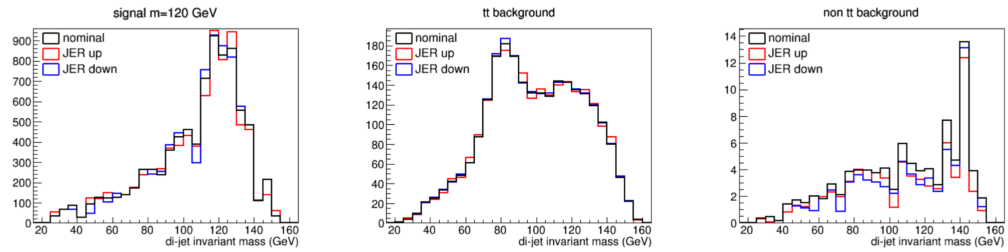


Figure 4.18: Effect of jet energy resolution uncertainty on the di-jet shape for the MC signal with charged Higgs mass of 120 GeV, on the $t\bar{t}$ background and on the non-tt background.

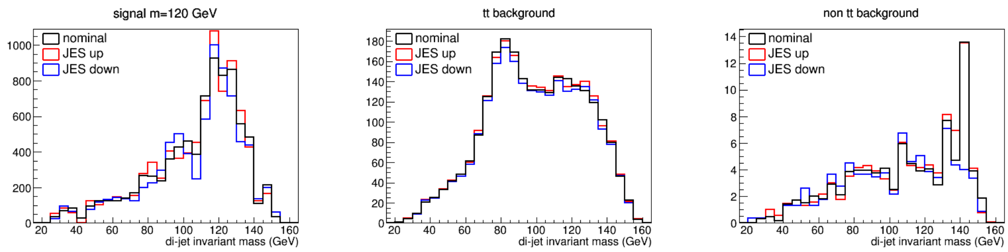


Figure 4.19: Effect of jet energy scale uncertainty on the di-jet mass shape for the MC signal with charged Higgs mass of 120 GeV, on the $t\bar{t}$ background and on the non-tt background.

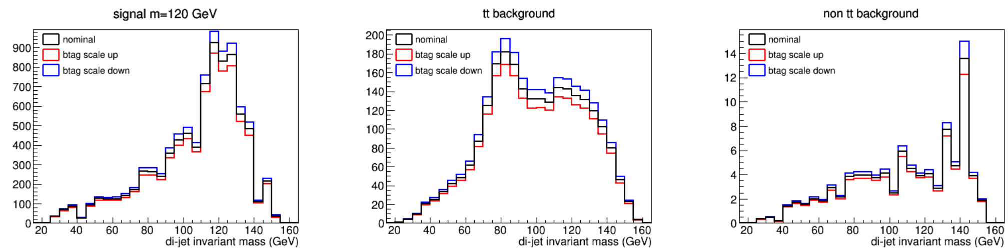


Figure 4.20: Effect of the uncertainty on the b-tagging scale factor for b and c-jets on the di-jet mass shape for the MC signal with charged Higgs mass of 120 GeV, on the $t\bar{t}$ background and on the non-tt background.

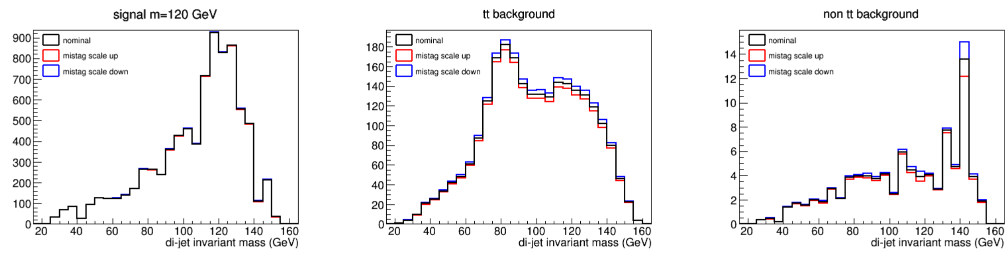


Figure 4.21: Effect of the uncertainty on the mistagging scale factor for light jets on the di-jet shape for the MC signal with charged Higgs mass of 120 GeV, on the $t\bar{t}$ background and on the non-tt background.

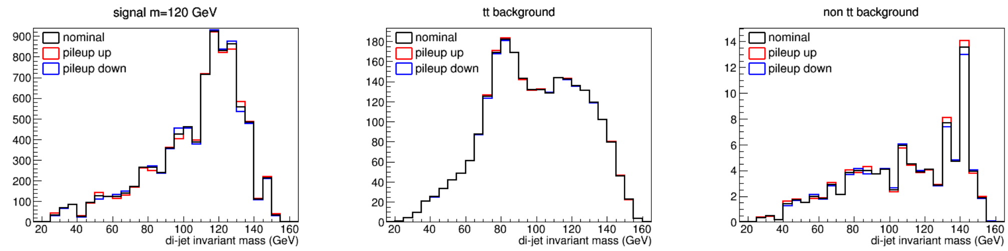


Figure 4.22: Effect of the uncertainty on the pileup reweighting scale factor on the di-jet mass shape for the MC signal with charged Higgs mass of 120 GeV, on the $t\bar{t}$ background and on the non-tt background.

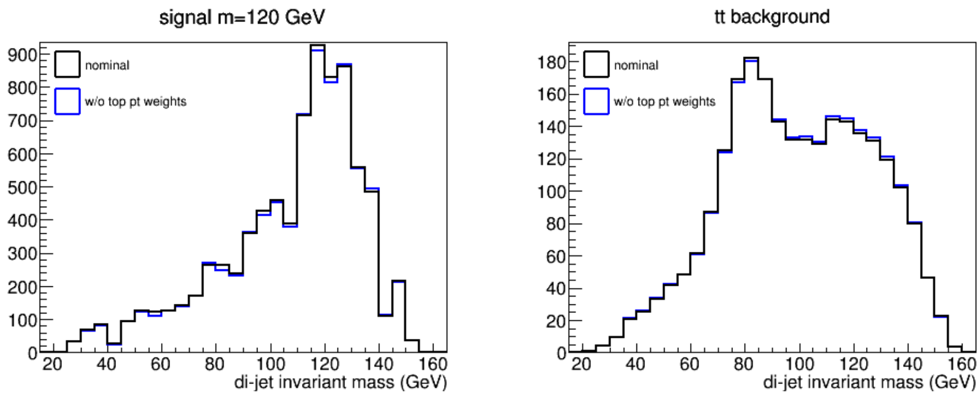


Figure 4.23: Effect of the uncertainty on the top pt reweighting scale factor on the di-jet mass shape for the MC signal with charged Higgs mass of 120 GeV and on the $t\bar{t}$ background.

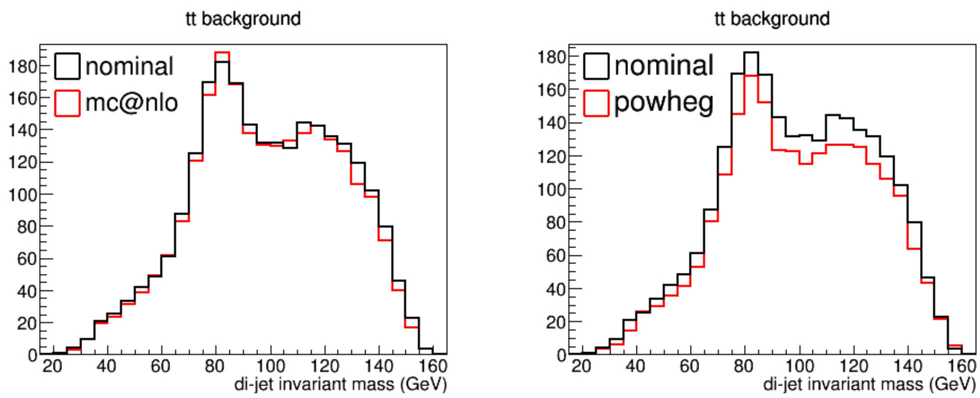


Figure 4.24: Effect of different MC samples on the di-jet mass shape for the $t\bar{t}$ background.

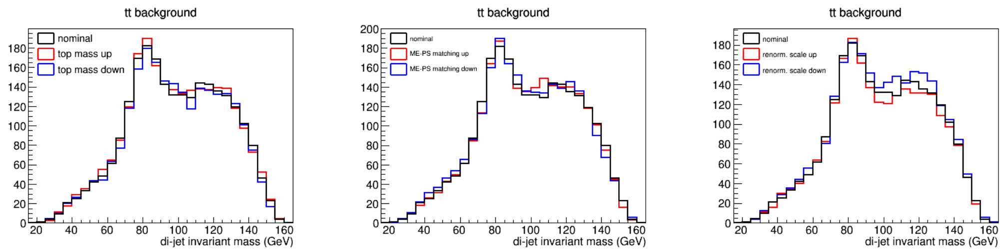


Figure 4.25: Effect of the uncertainty of the top mass, on the matching between ME and PS, and on the renormalization and factorization scale on di-jet mass shape for the $t\bar{t}$ background.

4.6 Limit

The dijet mass distributions obtained with the kinematic fit are used in a binned maximum-likelihood fit to extract a possible signal.

The calculation of upper limits on the charged Higgs branching ratio is based on the modified frequentist CL_s criterion with a test statistic based on profile likelihood ratio, and incorporating systematic uncertainties via nuisance parameters following the frequentist paradigm.

The upper limits are determined as a binned maximum likelihood fit on the dijet mass distribution. The likelihood is determined as a product of Poisson probabilities times a product of the probability density functions of nuisance parameters Θ

$$L(data|\mu,\Theta) = \Pi_i Poisson(n_i|s_i(\mu,\Theta)+b_i(\Theta)) \times \Pi_j p(\tilde{\Theta}_j|\Theta_j), \tag{4.9}$$

where n_i is the number of observed data events in the i-th bin, while $s_i(\mu,\Theta)$ and $b_i(\mu,\Theta)$ are the number of expected signal and background events in the i-th bin. The Poisson probability for the i-th bin is defined as

$$Poisson(n_i|s_i(\mu,\Theta)+b_i(\Theta)) = \frac{(s_i(\mu,\Theta)+b_i(\Theta))^{n_i}}{n_i!} exp(-s_i(\mu,\Theta)-b_i(\Theta)). \tag{4.10}$$

The probability density functions for the nuisance parameters define the probability for the true value of a nuisance parameter to be equal to Θ_i when the best estimate for it is $\tilde{\Theta}_j$ obtained by some measurement. The μ parameter is the signal strenght modifier, i.e. the quantity on which the limit is calculated.

To quantify how compatible the search is with the background only hypothesis, the test statistic $\tilde{q}_\mu$ defined as the profile likelihood ratio

$$\tilde{q}_\mu = -2ln\frac{L(data|\mu,\hat{\Theta}_\mu)}{L(data|\hat{\mu},\hat{\Theta})} \tag{4.11}$$

is used with the constraint $0 < \hat{\mu} < \mu$ to ensure a physical one-sided confidence interval. Here $\hat{\Theta}_\mu$ are conditional maximum likelihood estimators of Θ. The estimators $\hat{\mu}$ and $\hat{\Theta}$ correspond to the global maximum of the likelihood function in Eq. 4.9. The observed value $\tilde{q}_\mu^{obs}$ for tested signal strenght parameter μ is evaluated from Eq. 4.11. Furthermore, Eq. 4.9 is maximized to obtain the set of nuisance parameters values $\tilde{\Theta}_\mu^{obs}$ and $\tilde{\Theta}_0^{obs}$ best describing the signal plus background and background only hypotheses, respectively.

To obtain the upper limit, the p-values for the signal plus background (s+b) and background only (b) hypotheses need to be defined to quantify the degree of compatibility with the value $\tilde{q}_\mu^{obs}$.

They can be written as

$$CL_{s+b} = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs} | signal + background) = \int_{\tilde{q}_\mu^{obs}}^{\infty} f(\tilde{q}_\mu | \mu, \hat{\Theta}_\mu^{obs}) d\tilde{q}_\mu \quad (4.12)$$

$$CL_b = P(\tilde{q}_\mu \geq \tilde{q}_\mu^{obs} | background - only) = \int_{\tilde{q}_0^{obs}}^{\infty} f(\tilde{q}_\mu | 0, \hat{\Theta}_0^{obs}) d\tilde{q}_\mu \quad (4.13)$$

where $f()$ are the probability distribution functions of $\tilde{q}_\mu$ for the hypotheses. Their distributions are obtained with separated toy experiments simulated with Monte Carlo method while keeping the nuisance parameters fixed to the maximum likelihood values $\hat{\Theta}_\mu^{obs}$ and $\hat{\Theta}_0^{obs}$.

The test CL_s is defined as the ratio of CL_{s+b} and CL_b

$$CL_s(\mu) = \frac{CL_{s+b}(\mu)}{CL_b(\mu)} \leq \alpha \quad (4.14)$$

Since $CL_s(\mu)$ limits are one-sided by definition, the 95% confidence level upper limit on μ is found, when $CL_s(\mu) = 0.05$. To evaluate $CL_s(\mu) = 0.05$, a decaying exponential function is fitted to the tested μ values.

In this analysis, log-normal nuisance probability density functions are used for the nuisance parameters.

Since the H^+ is expected to be produced in $t\bar{t}$ decays, the number of signal events is quantified by

$$s_i(\mu, \Theta) = \mu^2 \times s_{HH,i}(\Theta) + 2\mu(1-\mu) \times s_{HW,i}(\Theta) + (1-\mu^2) \times s_{WW,i}(\Theta), \quad (4.15)$$

where μ is defined as

$$\mu = B(t \to bH^+)B(H^\pm \to c\bar{b}) \quad (4.16)$$

Here $s_{HW,i}(\Theta)$ and $s_{WW,i}(\Theta)$ are the number of expected events for the $t\bar{t} \to bH^\pm bW^\mp$ and $t\bar{t} \to bW^\pm bW^\mp$ processes. Since the expected μ values are less than some percent, the contribution from $t\bar{t} \to bH^\pm bH^\mp$ is neglected. Assuming for the charged Higgs boson a branching ratio $B(H^\pm \to c\bar{b}) = 1$, the parameter of interest μ becomes simply equal to the branching ratio of the top quark to charged Higgs. In this sense, it is possible to express the result of the statistical analysis as un upper limit on the probability for the top quark to decay to a charged Higgs state.

The expected and observed limits on the branching fraction $B(t \to H^+b)$, calculated as a function of the charged Higgs mass are shown in Figure 4.26. The observed limit is compared with the expected one, evaluated taking into account all the systematic uncertainties. Error bands at 1σ and 2σ on the expected limit are also shown. The expected limit ranges between 1.2% and 2.2%.

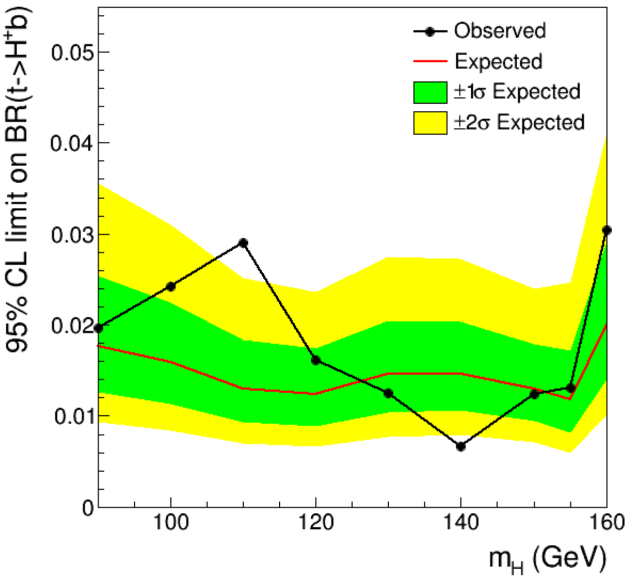


Figure 4.26: Expected and observed upper limit on the branching ratio of the top quark to H^+, calculated assuming for the H^+ a branching ratio $B(H^\pm \to c\bar{b}) = 1$

Conclusions

This thesis is framed within the current studies on beyond Standard Model physics at the CMS experiment and in particular it is focused on searches for a charged Higgs boson. In this sector, most of results were concerned with the $H^+ \to \tau\nu$ and $H^+ \to tb$, that within the MSSM are mostly sensitive to the high H^+ mass, high $tan\beta$ region of the parameter space.

The search for a light $H^+ \to c\bar{s}$ and $H^+ \to c\bar{b}$ were developed to extend the sensitivity of charged Higgs analyses to low H^+ masses and low $tan\beta$. Moreover, they can be the dominant final states for a light charged Higgs in the Flipped 2HDM, which so far has been less extensively probed by LHC results. Both ATLAS and CMS have provided results in the $c\bar{s}$ is final state. More recently, CMS has provided the first result for the $c\bar{b}$ decay.

This thesis is concerned with this last study. The analysis was carried out of the full 8 TeV data sample collected by CMS during run-1. An ad-hoc strategy, devised in collaboration with theoretical colleagues, for the production of the $H^+ \to c\bar{b}$ signal sample was adopted. Event selections have been studied in the comparison beetween data and MC samples. The $t\bar{t}$ background has been split in different contributions depending on the presence and the number of additional jets, in order to isolate and understand the impact of the $t\bar{t}$ categories with additonal b-jets. This was particularly important in an analysis using a three b-tag selection. For the reconstruction of the charged Higgs mass a kinematic fit technique was used, whose parameters have been checked and optimized using a SM $t\bar{t}$ sample. Different configurations of the kinematic fit where studied in order to optimize mass reconstruction in the whole H^+ mass range. Confidence limits on the branching fraction of the top quark to charged Higgs were calculated. My analysis was focused on the muon channel, and strongly contributed to the achievement of the CMS results presented at the International Conference of High Energy Physics held in Chicago in August 2016. Assuming that the charged Higgs boson decays exclusively into $c\bar{b}$, the expected limit for the branching fraction of the top quark into charged Higgs ranges between 1.2% and 2.2%, while the observed limit is between 0.8% and 2.7%. No evidence for a charged Higgs boson in the mass range 90-160 GeV was found.

Bibliography

[1] Particle Data Group, Chin.Phys. C38 090001 (2014)

[2] S.L. Glashow, Nucl. Phys. 22, 579 (1961)

[3] S.Weimberg, Phys. Rev. Lett. 19, 1264 (1967)

[4] S.L. Glashow, J.Iliopulos, and L. Maiani, Phys. Rev. D2, 1285, (1970)

[5] R.K. Ellis W.J. Stirling and B.R. Webber, *QCD and collider physics*, Camb. Monogr. Part. Phys. Cosmol. 81, ed. Cambridge University Press (1996)

[6] A.Djouadi, Phys. Rep. 457, 1-216 (2008)

[7] M. E. Peskin and D. V. Schroeder, *An introduction to quantum field theory*, ed. Perseus Books Publishing L.L.C (1995)

[8] LEP-Tevatron Electroweak Working Group http://lepwwg.web.cern.ch/LEPWWG

[9] ATLAS Collaboration, Phys. Lett. B 716, 1-29 (2012)

[10] CMS Collaboration, Phys. Lett. B 716, 30-61 (2012)

[11] ATLAS Collaboration, CMS Collaboration, Phys. Rev. Lett. 114, 191803 (2015)

[12] CMS Collaboration, Phy. Rev. Lett. 112, 191802 (2014)

[13] CMS Collaboration, Eur. Phys J. C 75, 147 (2015)

[14] CMS Collaboration, J. High Energy Physics 01, 010 (2012)

[15] CMS Collaboration, Eur. Phys J. C 76, 401 (2016)

[16] ATLAS and CMS Collaborations, J. High Energy Physics 08, 045 (2016)

[17] M.E. Peskin, arXiv:hep-ph/9705479 (1997)

[18] G. Bertone, D. Hooper and J. Silk, Phys. Rep. 405, 279 (2005)

[19] R.Contino, arXiv:1005.4269v1 (2010)

[20] CMS Physics Analysis Group Summary Plots, https://twiki.cern.ch/twiki/bin/viewauth/CMSPublic/SummaryResultsHIG

[21] CMS Collaboration, Technical Report CMS-PAS-B2G-16-007 (2016)

[22] R.M. Barnett et al., Phys. Lett. B 136, 191 (1984)

[23] R.M. Janovic and D. Wyler, Phys. Rev. D 30, 1529 (1984)

[24] S.L. Glashow and S. Weimberg, Phys Rev. D 15, 1985 (1977)

[25] H.E. Haber and G.L. Kane, Phys. Reports 117, 75 (1985)

[26] I.J.R. Aitchison, *Supersymmetry in particle physics: an elementary introduction*, ed. Cambridge University Press (2007)

[27] J.C Collins, *Renormalization*, Cambridge Monographs on Mathematical Physics, ed. Cambridge University Press (1984)

[28] A. Djouadi and J. Quevillon, J. High Energy Physics 10, 028 (2013)

[29] A. Djouadi, Phys. Rep. 459, 1-241 (2008)

[30] ALEPH, DELPHI, OPAL Collaborations and LEP Working Group for Higgs Boson Searches, Eur. Phys. J. C 74, 547 (2006)

[31] D0 Collaboration, Phys. Rev. Lett. 82, 4975 (1999)

[32] CDF Collaboration, Phys. Rev. Lett. 96, 042003 (2006)

[33] D0 Collaboration, Phys. Rev. Lett. B682, 278 (2009)

[34] CMS Collaboration, J. High Energy Physics 11, 018 (2015)

[35] CMS Collaboration, J. High Energy Physics 12, 1 (2015)

[36] ATLAS Collaboration, J. High Energy Physics B 03, 088 (2015)

[37] ATLAS Collaboration, J. High Energy Physics 03, 127 (2016)

[38] ATLAS Collaboration, Phys. Lett. B 759, 555-574 (2016)

[39] ATLAS Collaboration, Conference Report ATLAS-CONF-2011-094 (2011)

[40] N.Arkani-Hamed and S.Dimopoulos, JHEP 0506, 073 (2005); G.F.Giudice and A. Romanino, Nucl. Phys. B B699, 65 (2004); J.D. Wells, Phys. Rev. D 71, 015013 (2005)

[41] L.J. Hall and Y. Nomura, J. High Energy Physics 1003, 076 (2010)

[42] H.E. Logan and D. MacLennan, Phys. Rev. D 81, 075016 (2010)

[43] CMS Collaboration, Technical Report CMS-PAS-HIG-16-030 (2016)

[44] L. Evans and P. Bryant, *LHC Machine*, JINST 3 (2008)

[45] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST 3 (2008)

[46] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST 3 (2008)

[47] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST 3 (2008)

[48] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST 3 (2008)

[49] A. W. Chao, K. H. Mess, M. Tigner, F. Zimmermann, *Handbook of Accelerator Physics and Engineering*, ed. World Scientific (2013)

[50] CMS Collaboration, *CMS Luminosity - Public Results*, https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults

[51] H.L. Lai et al., Phys.Rev. D 82 (2010)

[52] J. Pumplin et al., J. High Energy Physics 0207 (2002)

[53] A. Martin, W. Stirling, R. Thorne, and G. Watt, Eur.Phys.J. C 63, 189 (2009)

[54] R. D. Ball et al., Nucl.Phys. B 867, 244(2013)

[55] W.J. Stirling, http://www.hep.ph.ic.ac.uk/~wstirlin/plots/crosssections2012_v5.pdf

[56] CMS Collaboration, Technical Design Report CERN/LHCC 97 - 033, CMS TDR 4 (1997)

[57] P. Adzic et al., JINST 2 (2007)

[58] G. Abbiendi for the CMS Collaboration, arXiV:1510.5424 (2015)

[59] CMS Collaboration, Technical Design Report CERN/LHCC 2000 - 38, CMS TDR 6.1 (2000)

[60] CMS Collaboration, Technical Design Report CERN/LHCC 02 - 26, CMS TDR 6 (2002)

[61] P. Billoir, et al., Nucl Instr. Meth. A 241, 115 (1985)

[62] M. Pioppi, CMS Internal Note 2007/065 (2008).

[63] The CMS Collaboration, JINST 7, P10002 (2012)

[64] The CMS Collaboration, JINST 10, P10002 (2015)

[65] CMS Collaboration, Technical Report CMS-PAS-PFT-09-001 (2009)

[66] M. Cacciari, G. P. Salam, G. Soyez, J. High Energy Physics 04, 063 (2008)

[67] CMS Collaboration, JINST, 6, 11002 (2011)

[68] D. Rathjens, *Jet energy calibration and a search for supersymmetry with vector boson fusion channel like sign di-τ_h final states*, PhD thesis

[69] K. Goebel, *Probing supersymmetry based on precise jet measurements at the CMS experiment*, PhD Thesis.

[70] CMS Collaboration, *Jet Energy Corrections for Multiple Cone Sizes*, https://twiki.cern.ch/twiki/bin/view/CMSPublic/MultipleConeSizes14

[71] CMS Collaboration, Technical Report CMS-PAS-BTV-09-001 (2009)

[72] CMS Collaboration, CMS Note 2006/019 (2006)

[73] CMS Collaboration, CMS Internal Note 2006/043 (2006)

[74] C. Weiser, CMS Note 2006/014 (2006)

[75] T. Speer, K.Prokofiev , R. Fruhwirth, W. Waltenberger , P. Vanlaer, CMS Internal Note 2006/032

[76] CMS Collaboration, Technical Report CMS-PAS-BTV-13-001 (2013)

[77] G. Landsberg, F. Moortgat, MET Reconstruction, Performance and Validation, CMS Analysis Note 2008/089

[78] CMS Collaboration, JINST 9, P10009 (2014)

[79] N. Kidonakis, arXiV:1311.0283v1 (2013)

[80] T. Sjostrand, S. Mrenna, and P.Z. Skands, J. High Energy Physics 05, 026 (2006)

[81] J. Alwall et al., Comput. Phys. Commun. 176, 300-304 (2007)

[82] J.Alwall et al., J. High Energy Physics 06, 128 (2011)

[83] S. Aioli, P. Nason, C. Oleari and E. Re, J. High Energy Physics 06, 43 (2010)

[84] CMS Collaboration, JINST 8, P04013 (2013)

[85] CMS Collaboration, Technical Report CMS-PAS-TOP-12-028 (2012)

[86] CMS Collaboration, Technical Report CMS-PAS-MUO-10-002

[87] CMS Collaboration, arXiV:1607.03663v1 (2016), submitted to the Journal of Instrumentation

[88] CMS Collaboration, Technical Report CMS-PAS-TOP-13-016.

[89] CMS Collaboration, Phys. Rev. D93, 072004 (2016)

Acknowledgements

I would like to thank my supervisor Cristina Riccardi for giving me the opportunity of a PhD in CMS, and for the confidence she gave me during these years. Special thanks to Fulvio Piccinini for a countless number of discussions and for his continuous help and encouragement.

I am most grateful to Maria Cepeda Hermida and Mario Pelliccioni for their help and support throughout the development of the analysis presented in this thesis. Thanks to Alexandre Nikitenko for all his valuable suggestions and encouragement during these past three years.

I would also like to thank the charged Higgs analysis group of the Seoul National University, in particular Geum Bong Yu for her collaboration.

I am indebted to prof. Vitaliano Ciulli and prof. Sandra Leone for the careful revision of this thesis, for their patience and helpfulness.

Thanks to the colleagues of the CMS group at the University of Pavia for their support and friendship, to all my officemates in Pavia and at CERN, especially to Marta Torti and Ilaria Vai, whom I had the pleasure of sharing the PhD with, from the beginning to the very end.

Finally I am deeply indebted to my family and all my friends, for their endless patience, and for their determination to make me eat, sleep and laugh every time I needed to. Thank you ♡.