Università degli Studi di Napoli “Federico II”

Scuola Politecnica e delle Scienze di Base
Area Didattica di Scienze Matematiche Fisiche e Naturali

Dipartimento di Fisica “Ettore Pancini”

Laurea Magistrale in Fisica

Search for a singly produced Vector-Like Quark T decaying to a top quark and a Z boson in proton-proton collisions at 13 TeV with the Compact Muon Solenoid detector at the LHC

Relatore:
Dott. Alberto Orso Maria Iorio
Dott. Luca Lista
Correlatore:
Prof. Fausto Guarino

Candidato:
Carmen Giugliano
Matricola N94/311

A.A. 2017/2018

To my family.

"Rome wasn't built in a day."

Contents

Introduction

The Standard Model (SM) of particle physics is the most consolidated and successful theory for the description of fundamental particles and their interactions. It has been thoroughly validated since all particles it predicts have been directly observed, and the features of the fundamental interactions it encompasses have been described with accuracy. Despite its extraordinary success, the Standard Model is not believed to give a complete picture of nature because there are many unsolved issues: from the large set of parameters not postulated by the theory to the lack of a quantistic description of gravity, passing through the lack of an explanation for the many astronomical and cosmological evidences of dark matter and dark energy in the Universe. A theoretical hint towards new physics is given also by the hierarchy problem, that is related to the divergences introduced by the corrections to the Higgs mass. Several new physics theories have been formulated to solve it, and, among them, some of the most promising predict the existence of new particles called Vector-Like quarks (VLQs). VLQs appear at the TeV scale, and behave differently under electroweak transformation with respect to the SM quarks since they present a chirality symmetry.
In this scenario, it is clear that a laboratory such as the Large Hadron Collider (LHC), whose energy reaches span up to the tens of TeV, is the right place. The LHC is an hadronic collider that provides proton-proton collisions at a design centre-of-mass energy of $\sqrt{s} = 14$ TeV, and a design luminosity of 10^{34} cm^{-2}s^{-1}. Around the collision points the LHC is equipped with four main experiments: ALICE, ATLAS, CMS and LHCb in order to detect and validate the SM physics and the physics beyond it, also known as BSM.
The aim of this thesis is to search for a singly produced Vector-Like quark T, in channels with a top quark and a Z boson, to verify the presence of new physics. The final state under investigation consists of a top quark that decays hadronically, and a Z boson that decays to two neutrinos. In this thesis, data collected by the Compact Muon Solenoid experiment at the LHC during 2016 have been analysed, and simulated samples have been used in order to emulate the VLQ T signal under investigation and the backgrounds from the SM that can mimic its signature in the detector.
The content of this thesis is organised in five Chapters:

Chapter 1 gives an overview on the theoretical framework of the Standard Model and describes its unsolved issues.
Chapter 2 introduces the LHC accelerator machine and provides a description of the CMS detector.
Chapter 3 presents the theoretical backgrounds that lead to the introduction of VLQs, with a description of their production mechanism and their decays channels. Finally a summary of the experimental state of the art concerning the Vector-Like quarks physics at the hadronic colliders is given.
Chapter 4 describes the final state sought after and provides a description of the algorithms used to identify and reconstruct physics objects employed for the presented physics analysis.
Chapter 5 provides a description of the data set and the MC simulation used, of the the event selection, and of the background estimation. Finally the fit procedure is described and the results are presented.

Chapter 1

The Standard Model of particle physics

The Standard Model (SM) of particle physics is the quantum field theory that describes the fundamental components of matter and their interactions observed in nature thus far. Its first formulation in the current form dates back to 1950s, and since then it has been tested to an ever increasing level of precision. The SM has provided precise predictions of fundamental interactions that have been confirmed by experiments throughout the years (see [12–17]). The SM is however not sufficient to provide a full picture of the known universe, as it fails to explain several phenomena, like the evidence for neutrino masses, dark energy and dark matter, and does not include any interpretation of gravitational field.

1.1 Standard Model overview

The SM manages to describe three of the four known fundamental interactions, as it define in a coherent model the unification of the electromagnetic and weak interaction [1–3], the Higgs mechanism [4–6], the strong interaction [10, 11] within the framework of the Quantum Chromodynamics (QCD) [7], including its features and the asymptotic freedom, by introducing a new degree named *color* [8, 9] (so the $SU(3)_C$ symmetry). The SM is a quantum field theory and, in addition to the usual spacetime symmetries, it exhibits invariance under transformations of the three gauge groups related to the fundamental interactions. The three gauge group are related to three of the four known fundamental interactions. In principle each group is associated with a set of massless spin-1 (vector) fields, called bosons, which obey the Bose-Einstein statistics:

- the circle group $U(1)_Y$, whose transformation can be represented by a unitary scalar complex operator multiplied by its quantum number, acting as the equivalent of the electromagnetic charge, called the weak hypercharge Y. The associated

Interaction	Mediators	Spin	Charge
Electromagnetic	γ	1	0
Weak	W^+, W^-, Z	1	$1, -1, 0$
Strong	8 gluons (g)	1	0

Table 1.1: Standard Model mediator bosons and fundamental interactions.

vector field is called B_μ ;

- the n = 2 special unitary group $SU(2)_L$, whose three transformations can be represented by the 2×2 Pauli matrices σ_i (i = 1, 2, 3) multiplied by a quantum number called the weak isospin I_3. The three associated vector fields are $W_\mu^{1,2,3}$;
- the n = 3 special unitary group $SU(3)_C$, whose eight operations can be represented by the 3×3 Gell-Mann matrices λ^j (j = 1, ..., 8) multiplied by a charge called the colour $C = (r, g, b)$. The eight associated vector fields are $G_\mu^{1,...,8}$.

In total, that makes 12 vector fields associated with three gauge symmetries, all of which can be summarized by the tensor product of groups that defines the gauge symmetry of the Standard Model:

$$SU(3)_C \otimes SU(2)_I \otimes U(1)_Y$$

Then, as will be shown in section 1.3.1, three of these massless fields acquire mass and become the known physics boson: W^+, W^-, Z. The other nine fields remain without mass: the eight gluons and the photon. The bosons, listed in Table 1.1, are the mediators of the known forces and regulate the interactions among other spin$-1/2$ particles of the SM, called fermions. For example the electromagnetic force is carried by spin-1 photons (γ) and acts between electrically charged particles, the weak interaction is ruled by three gauge vector bosons, $W^\pm$ and Z, discovered in 1983 by the UA1 and UA2 experiments [12–14] , and is responsible for phenomena like nuclear β-decays, and the strong interaction responsible of holding together nuclei its gauge bosons are the gluons (g).
The SM also has 12 fermion fields with half-integer spin which obey to the Fermi-Dirac statistics and are classified as lepton or quark fields. Among the lepton fields, fermions are listed in three generations or families, as can be seen in Table 1.2, in which each one is a doublet of particles associated to an isospin quantum number and carry electromagnetic and weak charges. Quark fields are fermions of spin 1/2 which carry electromagnetic, weak and strong charges. In particular the colour for the strong interactions plays the same role of the charge for the electromagnetic interactions. Quarks are grouped in three families and are listed in Table 1.3. According to Dirac equation all these fermions have an associated anti-particle that has the same mass but opposite quantum numbers.

Particles			Spin	Charge
$\begin{pmatrix}\nu_e\\ e\end{pmatrix}$	$\begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix}$	$\begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix}$	1/2	$\begin{matrix}0\\ -1\end{matrix}$

Table 1.2: Standard Model leptons.

Particles			Spin	Charge
$\begin{pmatrix}\mathrm{u}\\ \mathrm{d}\end{pmatrix}$	$\begin{pmatrix}\mathrm{c}\\ \mathrm{s}\end{pmatrix}$	$\begin{pmatrix}\mathrm{t}\\ \mathrm{b}\end{pmatrix}$	1/2	$\begin{matrix}2/3\\ -1/3\end{matrix}$

Table 1.3: Standard Model quarks.

Fermions have also a property called chirality, and they can either be left-chiral, or right-chiral, which are sometimes referred to as left- or right- handed. It is experimentally verified that left-handed fermion fields transform differently under the $SU(2)_L$ gauge symmetry than do the right-handed fermion fields, which results in an asymmetry in the properties of fermions. Specifically, there are left-handed and right-handed charged leptons, left-handed neutrinos, but no right-handed neutrinos. This does not yet have an explanation from prime principles. The last field in the Standard Model is a complex scalar doublet field (ϕ), named the Higgs field after one of the theorists who predicted its existence in 1964 [5]. In Figure 1.1 the particles predicted by the SM are listed.

1.2 Quantum electrodynamics

The first relativistic quantum field theory to be developed was the the Quantum Electrodynamics (QED). It is an Abelian gauge theory that describes the dynamics and interactions of fermions and the electromagnetic field.
The lagrangian density for the QED can be obtained starting from the free lagrangian density of the Dirac field ψ with mass m:

$$\mathcal{L}_D = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi, \tag{1.1}$$

where the first is a kinetic term and the last is a mass term, with γ^μ the Dirac matrices and ψ and $\bar{\psi}$ the 4-components spinor and its adjoint, requiring to satisfy the "local gauge invariance" principle, i.e. has to be invariant under the local gauge transformation. In this case the symmetry group involved is $U(1)_q$ where the subscript indicates the charge as conserved quantum number, that assumes the meaning of the charge of the particles. The Equation (1.1) describes the kinematics of a free (non-interacting) fermion in an

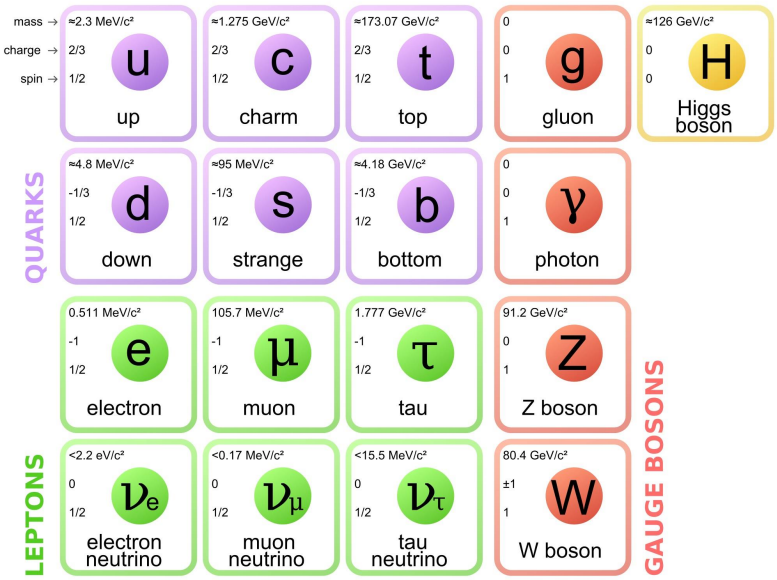


Figure 1.1: Standard Model particles: in violet the quarks, in green the leptons, in red the gauge bosons, and in yellow the Higgs boson.

electromagnetic field and it is globally invariant under a $U(1)_q$ transformation:

$$\psi \to \psi' = e^{i\theta}\psi \tag{1.2}$$

with θ arbitrary constant. Instead by imposing the locally invariance under the transformation $U(1)_q$ one obtains:

$$\begin{aligned} \psi &\to \psi' = e^{i\theta(x)}\psi \\ \bar{\psi} &\to \bar{\psi}' = e^{-i\theta(x)}\bar{\psi} \end{aligned} \tag{1.3}$$

where $\theta(x)$ is an arbitrary function depending on the spacetime coordinates. In this case the (1.1) is not invariant under local gauge transformation, indeed, while the term with mass is invariant, the derivative is not;

$$\partial_\mu\psi \to \partial_\mu\psi' = e^{iq\theta(x)}\partial_\mu\psi(x) + iqe^{iq\theta(x)}\psi(x)\partial_\mu\theta(x) \tag{1.4}$$

the second term destroys the invariance, to restore which is necessary to introduce the covariant derivative, a particular derivative that undergoes the same phase transformation of the field:

$$\begin{aligned} \mathcal{D}_\mu\psi &\to e^{iq\theta(x)}\mathcal{D}_\mu\psi \\ \mathcal{D}_\mu &\equiv \partial_\mu + iqA_\mu(x) \end{aligned}$$

as long as the vectorial field A_μ becomes like:

$$A_\mu(x) \to A_\mu(x) - \partial_\mu \theta(x) \tag{1.5}$$

obtained using the minimal substitution. The lagrangian now is invariant under local gauge transformation and can be written as:

$$\mathcal{L}_D = i\bar{\psi}\gamma^\mu \mathcal{D}_\mu \psi - m\bar{\psi}\psi,$$

and making explicit the covariant derivative:

$$\mathcal{L}_D = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi - q\bar{\psi}\gamma^\mu \psi A_\mu = \mathcal{L}_0 - J^\mu A_\mu \tag{1.6}$$

where the latter term contains the interaction between the Dirac particle and the electromagnetic field. The quantity J^μ is interpreted as the charge current, i.e. the probability current of the particle multiplied by its charge.
To complete the lagrangian of the QED has to be added the term related to kinetic energy:

$$\mathcal{L}_\gamma = -\frac{1}{4} F^{\mu\nu} F_{\mu\nu}, \tag{1.7}$$

it describes the propagation of free photons and it is invariant for local gauge transformation and $F^{\mu\nu}$ is the field strength tensor which can be written in terms of 4-vector electromagnetic field A^μ in the following way:

$$F^{\mu\nu} = -F^{\nu\mu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \tag{1.8}$$

So the QED lagrangian for a Dirac particle in a electromagnetic field can be obtained by adding 1.6 and 1.7:

$$\mathcal{L}_{QED} = \mathcal{L}_D + \mathcal{L}_\gamma = i\bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi - \frac{1}{4} F^{\mu\nu} F_{\mu\nu}. \tag{1.9}$$

To each lagrangian it is possible to associate a Feynman diagram; in this case the interaction term is used to obtain the amplitude for all electromagnetic processes and in Figure 1.2 all the fundamental transitions associated to the lagrangian in (1.9) can be seen.

1.3 The Electroweak theory

Fermi's theory of the β decay leads to a phenomenological description of the weak interactions determined by the current-current lagrangian, known as Fermi-like interac-

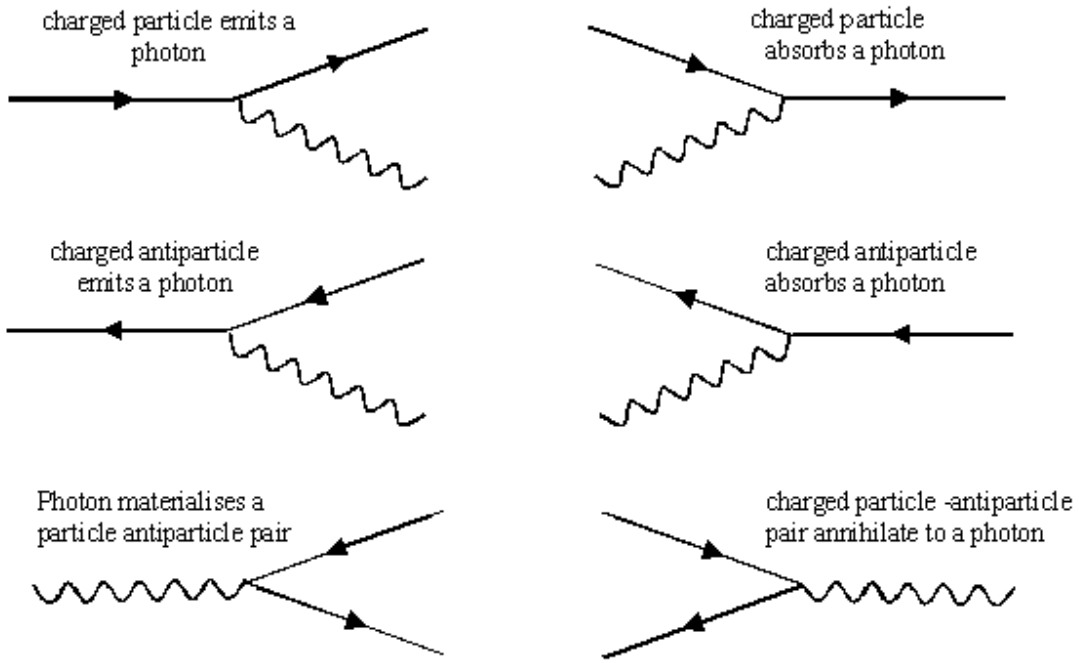


Figure 1.2: Basics vertices in Quantum Electrodynamics.

tion:

$$\mathcal{L}_F = \frac{G}{\sqrt{2}} J^{\dagger\mu}(x) J_\mu(x) \tag{1.10}$$

where G is the Fermi constant

$$G = 1.16638 \times 10^{-5}\ \text{GeV}^{-2}$$

and J^μ is the weak current (analogue of the electromagnetic one) and is the sum of a weak leptonic current, $l^\mu(x)$ and a weak hadronic current $h^\mu(x)$. This interpretation, although it manages to describe successfully the short-range interaction approximation, leads to divergences, which manifest themselves in the violation of the unitarity limit of the Fermi-like cross-section, due to the non-dimensionality of Fermi's constant. The divergent behaviour of the cross section can be avoided by perfecting the analogy with the electromagnetic interaction, i.e. introducing an adimensional coupling constant and an intermediate vector boson, analogous to the photon, which should be the mediator of the weak interaction. The Electroweak theory so it is the natural development of Fermi's theory and has been proposed by S. Glashow, A. Salam and S. Weinberg in 60's and 70's [1–3] and is also known as the GWS model of the weak interactions.

1.3.1 GSW Model

The GWS Model is a quantum field theory based on the symmetry group $SU(2)_L$, where the subscript L means that only the left-handed chiral components of the fields can take part in the weak interactions. The generators of $SU(2)_L$ are the Pauli 2×2 matrices σ_i ($i = 1, 2, 3$), so it is useful adopt the formalism of the angular momentum, in this way the particles, eigenstates of the weak interaction, are arranged in six doublets of weak isospin. By naming I the weak isospin and I_3 its observed component, the six doublets for leptons are:

$$I = 1/2 \qquad \begin{matrix} I_3 = +1/2 \\ I_3 = -1/2 \end{matrix} \qquad \begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L ,$$

and for quarks are:

$$I = 1/2 \qquad \begin{matrix} I_3 = +1/2 \\ I_3 = -1/2 \end{matrix} \qquad \begin{pmatrix} \mathrm{u} \\ \mathrm{d}' \end{pmatrix}_L \begin{pmatrix} \mathrm{c} \\ \mathrm{s}' \end{pmatrix}_L \begin{pmatrix} \mathrm{t} \\ \mathrm{b}' \end{pmatrix}_L .$$

for quark pairs, d′, s′ and b′ are the weak interaction eigenstates that are obtained as linear combination the strong interaction eigenstates (or mass eigenstates) and the mixing of different flavours is given by:

$$\begin{pmatrix} \mathrm{d}' \\ \mathrm{s}' \\ \mathrm{b}' \end{pmatrix} = V \begin{pmatrix} \mathrm{d} \\ \mathrm{s} \\ \mathrm{b} \end{pmatrix},$$

where V is the complex unitary matrix named Cabibbo-Kobayashi-Maskawa (CKM) matrix:

$$V = \begin{pmatrix} \mathrm{V_{ud}} & \mathrm{V_{us}} & \mathrm{V_{ub}} \\ \mathrm{V_{cd}} & \mathrm{V_{cs}} & \mathrm{V_{cb}} \\ \mathrm{V_{td}} & \mathrm{V_{ts}} & \mathrm{V_{tb}} \end{pmatrix}.$$

In Sec. 1.3.4 is reported the derivation of the CKM matrix with a focus on its parametrization and its most significant properties. As for the QED development also in the GSW model a local gauge transformation invariance is required, the action of the $SU(2)_L$ in the weak isospin doublets leads to:

$$\begin{pmatrix} \nu_\ell \\ \ell^- \end{pmatrix}'_L = e^{-i\vec{\alpha}(x)\cdot\vec{\tau}} \begin{pmatrix} \nu_\ell \\ \ell^- \end{pmatrix}_L , \tag{1.11}$$

where $\vec{\tau}$ are the Pauli matrices divided by 2 and $\vec{\alpha}(x)$ is the vector of real parameters of the transformation that depends on the spacetime coordinates.
Only the left-handed components can take part to the weak charged-current processes,

while the right-handed of charged fermions can take part to the neutral current weak process. The right-handed components are singlet of $SU(2)_I$:

$$I = 0 \qquad e_R^-, \mu_R^-, \tau_R^-, d_R, u_R, s_R, c_R, b_R, t_R.$$

The request of invariance under the $SU(2)_L$ group leads to the introduction of an isospin triplet of Yang-Mills fields: $W^\mu_{(i)}$ with $i = (1, 2, 3)$. These gauge fields are not yet the physics boson known as $W^\pm$ and Z but they can obtained by combination of the W^μ_i with another gauge field, the singlet B^μ that is the field associated at the additional local gauge symmetry related to the Abelian group $U(1)_Y$. In this way it was possible incorporate the electromagnetic interaction in the weak one. The quantum number associated to the new group is the weak hypercharge Y, defined by:

$$Q = I_3 + \frac{Y}{2} \tag{1.12}$$

which represent the electric charge Q (in units of e) of the I_3 member of a weak isomultiplet, assigned a weak hypercharge Y. The definition (1.12) was proposed by Glashow and is an extension of the Gell–Mann-Nishijima relation for charges valid also for these weak quantum numbers. So the symmetry group of transformation is $SU(2)_I \otimes U(1)_Y$ and leads to 4 gauge fields, 3 as said before coming from the $SU(2)_L$ and 1 coming from the group $U(1)_Y$.
The procedure to obtain the Electroweak lagrangian is analogous to the electromagnetic case, indeed the requirement of gauge local invariance under the $SU(2)_I \otimes U(1)_Y$ leads to the introduction of the covariant derivative:

$$D_\mu = \partial_\mu + ig\frac{\vec{\tau}}{2}W_\mu + ig'\frac{Y}{2}B_\mu, \tag{1.13}$$

where g and g' are the two coupling constants for the two interactions. Neglecting the mass term for now and introducing the (1.13), the electroweak lagrangian for fermions can be written:

$$\mathcal{L}_{fermions} = \sum_f \bar{\psi}\gamma^\mu D_\mu \psi \tag{1.14}$$

and a term for the dynamics of the gauge boson fields

$$\mathcal{L}_{gauge} = -\frac{1}{4}W_i^{\mu\nu}W^i_{\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} \tag{1.15}$$

with $W_i^{\mu\nu}$ and $B^{\mu\nu}$ the tensor fields:

$$\begin{aligned} W_i^{\mu\nu} &= \partial^\mu W_i^\nu - \partial^\nu W_i^\mu \\ B_i^{\mu\nu} &= \partial^\mu B^\nu - \partial^\nu B^\mu \end{aligned} \tag{1.16}$$

The complete electroweak lagrangian therefore is:

$$\begin{aligned}\mathcal{L}_{EW} = &- i\psi_L\gamma^\mu\left(\partial_\mu + ig\frac{\vec{\tau}}{2}\cdot W_\mu + ig'YB_\mu\right)\psi_L + \\ &- i\psi_R\gamma^\mu\left(\partial_\mu + ig'YB_\mu\right)\psi_R + \\ &- \frac{1}{4}W_i^{\mu\nu}W^i_{\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} + \\ &+ \frac{1}{2}g\,\epsilon_{ijk}W_i^{\mu\nu}W_{j\mu}W_{k\nu} + \frac{1}{4}g^2\,\epsilon_{ijk}\epsilon_{imn}W_{j\mu}W_{k\nu}W^\mu_m W^\nu_n,\end{aligned} \tag{1.17}$$

where ψ_L and ψ_R are the left and right-handed chiral components of the particles, and the term in the last line describes the three and four-point self interactions of the vector bosons that arise because of the non-Abelian nature of the $SU(2)_I$ group. The four gauge fields can be combined to produce the physical vector fields for the $W^\pm, Z$ bosons and the photon:

$$W^\pm_\mu = \frac{1}{\sqrt{2}}\left(W^1_\mu \mp iW^2_\mu\right) \tag{1.18}$$

and therefore one transforms into the physical bosons in the following way:

$$\begin{pmatrix} Z^\mu \\ A^\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix}\begin{pmatrix} W^\mu \\ B^\mu \end{pmatrix} \tag{1.19}$$

with

$$\cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}}, \quad \text{and} \quad \sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}} \tag{1.20}$$

The parameter θ_W has to be determined experimentally an is called Weinberg angle or weak mixing angle. The electromagnetic charge therefore is:

$$q = g'\cos\theta_W = g\sin\theta_W \tag{1.21}$$

Two of these can be combined together in order to give two vector bosons $W^\pm$, that are electrically charged and can induce transitions between the members of the weak isospin doublets. The third gauge boson of the triplet should be electrically neutral.

1.3.2 Spontaneous symmetry breaking and Higgs mechanism

The gauge fields found in subsection 1.3.1 are all massless, the introduction of a mass term ah hoc like

$$(1/2)m^2B_\mu B^\mu$$

in Equation (1.17) would break the gauge symmetry. To give mass to gauge fields, the simplest and most elegant way is the spontaneous symmetry breaking (SSB) and the

Higgs mechanism. The Higgs boson field is a doublet of complex scalar fields that can be written as:

$$\begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{1.22}$$

in the form of $SU(2)_I \otimes U(1)_Y$ multiplet to ensure that the lagrangian invariance remains. The complex scalar field ϕ^+ destroys positive charged particles and created negative particles while ϕ^0 destroys neutral particles and creates neutral antiparticles. The lagrangian for the ϕ field is:

$$\begin{aligned} \mathcal{L}_H &= (D^\mu \phi)^\dagger D_\mu \phi - V(\phi) = \\ &= (D^\mu \phi)^\dagger D_\mu \phi - \frac{1}{2}\mu^2 \phi^\dagger \phi - \frac{1}{4}\lambda (\phi^\dagger \phi)^2, \end{aligned} \tag{1.23}$$

where $V(\phi)$ is the potential responsible of the symmetry breaking, and the parameter λ is assumed to be positive. The potential is minimised to determine the ground state, ϕ_0. For $\mu^2 > 0$ the potential V assumes a unique minimum at $\phi_0 = 0$ and the ground state is symmetric under $SU(2)_I$. Instead, for $\mu^2 < 0$, the shape of the potential is modified as can be seen in the Figure 1.3, and V assumes a non-trivial minimum:

$$\phi_0^2 = -\frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2}.$$

The vacuum expectation value, defined as the absolute value of the field at the minimum of the potential, is non-zero and corresponds to the radius of a circumference in the complex plane $Re(\phi) - Im(\phi)$. Without any loss of generality a reference minimum can be chosen among all possible ground states:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}$$

Adding to the lagrangian of the gauge field sector the potential in Equation (1.23), when the covariant derivative acts, one has:

$$\mathcal{L}_H = (D^\mu \phi)^\dagger D_\mu \phi - \frac{1}{2}\mu^2 \phi^\dagger \phi - \frac{\lambda}{4}(\phi^\dagger \phi)^2 - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} - \frac{1}{4} B^{\mu\nu} B_{\mu\nu}, \tag{1.24}$$

with:

$$D^\mu \phi = \left(\partial^\mu + i g \frac{\vec{\tau}}{2} W^\mu + i g' Y B^\mu \right) \phi \tag{1.25}$$

$$F^{\mu\nu} = \partial^\mu W^\nu - \partial^\nu W^\mu - g W^\mu \times W^\nu \tag{1.26}$$

$$B^{\mu\nu} = \partial^\mu B^\nu - \partial^\nu B^\mu. \tag{1.27}$$

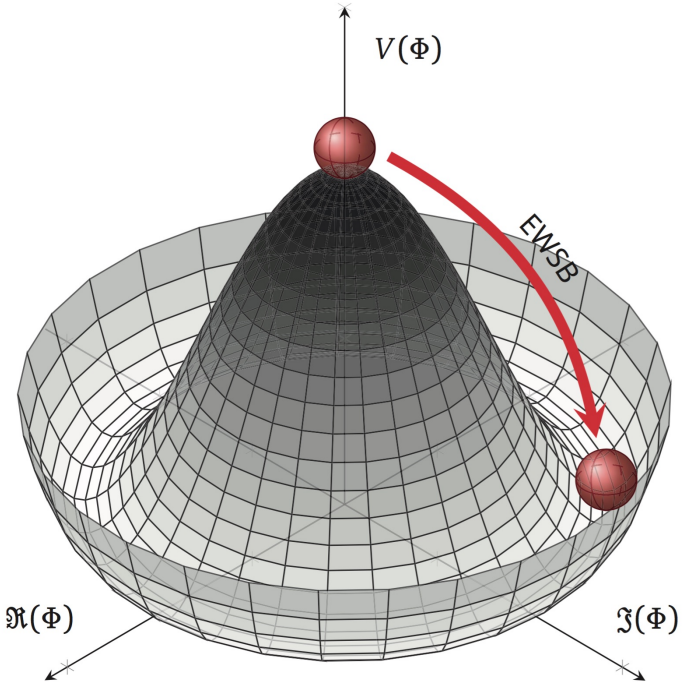


Figure 1.3: Shape of the Higgs potential $V(\phi) = \mu^2\phi^2 + \lambda\phi^4$ for $\lambda > 0$ and $\mu^2 < 0$.

By the parametrization of the fluctuations of the ϕ field around the vacuum, the Higgs field assumes the value:

$$\phi = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}\left(v + H(x)\right), \end{pmatrix} \tag{1.28}$$

with:

$$v = \sqrt{-\frac{\mu^2}{\lambda}}. \tag{1.29}$$

By substituting (1.28) in (1.24), one finds that:

$$\begin{aligned} \mathcal{L}_{G\Phi} = & \frac{1}{2}\partial_\mu H \partial^\mu H - \mu^2 H^2 + \\ & - \frac{1}{4}(\partial_\mu W^1_\nu - \partial_\nu W^1_\mu)(\partial^\mu W^{1\nu} - \partial^\nu W^{1\mu}) + \frac{1}{8}g^2 v^2 W^1_\nu W^{1\nu} \\ & - \frac{1}{4}(\partial_\mu W^2_\nu - \partial_\nu W^2_\mu)(\partial^\mu W^{2\nu} - \partial^\nu W^{2\mu}) + \frac{1}{8}g^2 v^2 W^2_\nu W^{2\nu} \\ & - \frac{1}{4}(\partial_\mu Z_\nu - \partial_\nu Z_\mu)(\partial^\mu Z^\nu - \partial^\nu Z^\mu) + \frac{1}{8}(g^2 + g'^2)v^2 Z_\nu Z^\nu \\ & - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \end{aligned} \tag{1.30}$$

The first line of (1.30) is the lagrangian of massive scalar field, the Higgs one, with mass $\sqrt{2}\mu$. The next two lines show that the components W^μ_1 and W^μ_2 of the triplet $\mathbf{W}^\mu$

acquire mass:

$$M_1 = M_2 = \frac{1}{2}gv \equiv M_W. \tag{1.31}$$

The fourth line shows that the field Z^μ acquires a mass:

$$M_Z \equiv \frac{1}{2}v\sqrt{g^2 + g'^2} = \frac{M_W}{\cos\theta_W} \tag{1.32}$$

and the last line shows that the field A^μ has a mass:

$$M_A = 0. \tag{1.33}$$

1.3.3 Masses of leptons

Fermion masses can be also generated through the spontaneous breaking of the $SU(2)_I \otimes U(1)_Y$ gauge symmetry. The procedure is different from the which used in the boson case, indeed the fermion mass term $-m\bar{\psi}\psi$ is not invariant under the $SU(2)_I \otimes U(1)_Y$ group because of the different transformation of the right and left-handed chiral components of the fields. However in a theory where the symmetry is spontaneously broken it is not necessary introduce in the lagrangian mass terms in an explicit way, but it can be done by coupling them to a scalar field that acquires a vacuum expectation value. This can be achieved by introducing a Yukawa coupling between the fermion field and the Higgs field, and can be written as:

$$\mathcal{L}_Y = g_f(\bar{\psi}_L\phi\psi_R - \bar{\psi}_R\phi^\dagger\psi_L), \tag{1.34}$$

where g_f is the Yukawa coupling constant. By substituting 1.28 in 1.34 one obtains:

$$\begin{aligned}\mathcal{L}_Y =& \frac{g_f}{\sqrt{2}}\left[(\nu_\ell,\ \ell)\begin{pmatrix}0\\ v+H\end{pmatrix}\ell_R + \ell_R(0,\ v+H)\begin{pmatrix}\nu_\ell\\ \ell\end{pmatrix}\right] = \\ =& \frac{g_f}{\sqrt{2}}(v+H)(\ell_L\ell_R + \ell_R\ell_L)\end{aligned} \tag{1.35}$$

The second line of (1.35) is a mass term "a la Dirac" and allows to identify the constant coefficient of $(\ell_L\ell_R + \ell_R\ell_L)$ as the mass term for leptons:

$$m_f = \frac{v}{\sqrt{2}}g_f. \tag{1.36}$$

This is the least satisfactory part of the model because even if this kind of Yukawa coupling solves the problem of leptons' masses, it does not arise from a gauge principle, it is purely phenomenological and needs a specific coupling constant for each fermion-Higgs interaction. Moreover the couplings are very different given the wide range of fermion masses experimentally observed.

1.3.4 Masses of quarks

The same mechanism adopted to give mass to leptons can be used also in the quarks case:

$$\begin{aligned}\mathcal{L}_Y =& \frac{1}{\sqrt{2}}\left[g^d_{i,j}(u_{i,L},\ d_{i,L})\begin{pmatrix}0\\ v+H\end{pmatrix}d_{j,R} + g^u_{i,j}(u_{i,L},\ d_{i,L})\begin{pmatrix}-(v+H)^*\\ 0\end{pmatrix}u_{j,R} + h.c.\right] = \\ =& \frac{1}{\sqrt{2}}(v+H)[g^u_{ij}(u_{i,L}u_{j,R} + u_{j,R}u_{i,L}) + g^d_{ij}(d_{i,L}d_{j,R} + d_{j,R}d_{i,L}) + h.c.]\end{aligned} \tag{1.37}$$

with u_i = (u, c, t) and d_i = (d, s, b). The matrix of mass terms is not diagonal, meaning:

$$m^u_{ij} = -\frac{v}{\sqrt{2}}g^u_{ij} \quad m^d_{ij} = -\frac{v}{\sqrt{2}}g^d_{ij}. \tag{1.38}$$

but it can be made diagonal with four different transformations on the family triplets $u_{i,L}$, $u_{i,R}$, $d_{i,L}$, and $d_{i,R}$ through:

$$u_{\alpha,L} = (\mathcal{U}^u_L)_{\alpha i}u_{i,L} \quad u_{\alpha,R} = (\mathcal{U}^u_R)_{\alpha i}u_{i,R} d_{\alpha,L} = (\mathcal{U}^d_L)_{\alpha i}d_{i,L} \quad d_{\alpha,R} = (\mathcal{U}^d_R)_{\alpha i}d_{i,R} \tag{1.39}$$

where α is the index in the mass diagonal basis and i is the index in the non-diagonal weak interaction basis.

$$\mathcal{L}_Y = \frac{1}{\sqrt{2}}(v+H)\left[m^u \mathrm{u}\bar{\mathrm{u}} + m^d \mathrm{d}\bar{\mathrm{d}} + m^s \mathrm{s}\bar{\mathrm{s}} + m^c \mathrm{c}\bar{\mathrm{c}} + m^t \mathrm{t}\bar{\mathrm{t}} + m^b \mathrm{b}\bar{\mathrm{b}}\right] \tag{1.40}$$

The same transformations must be applied to the interacting term, invariant under the $SU(2)_I \otimes U(1)_Y$ symmetry, that still contains the eigenkets of the weak interaction. When this operation is worked out the term of the coupling with the Z boson, i.e. neutral current coupling term, is diagonal also in the mass basis if the transformations of Eq. 1.39 are unitary, instead the term of the coupling with the W boson, i.e. charged current coupling term, is:

$$\begin{aligned}\mathcal{L}_{CC} =& -\frac{g}{\sqrt{2}}(\bar{u}_{i,L},\ \bar{d}_{i,L})\gamma^\mu\tau_+ W^+_\mu \begin{pmatrix}u_{Li}\\ d_{Li}\end{pmatrix} + h.c. \\ =& -\frac{g}{\sqrt{2}}\bar{u}_{iL}\gamma^\mu d_{Li}W^+_\mu + h.c. \\ =& -\frac{g}{\sqrt{2}}\bar{u}_{\alpha_L}\left[(\mathcal{U}^u_L)_{\alpha i}(\mathcal{U}^d_L)^\dagger_{\beta i}\right]\gamma^\mu d_{L\beta}W^+_\mu + h.c.,\end{aligned} \tag{1.41}$$

where the matrix:

$$V_{\alpha\beta} = \left[\mathcal{U}^u_L \mathcal{U}^{d\dagger}_L\right]_{\alpha\beta} \tag{1.42}$$

is unitary but not diagonal. V is the CKM [1] matrix, it is a 3×3 unitary matrix and display the mismatch between the weak eigenstates and the mass eigenstates leads to transitions between quark generations through flavour changing interactions. The definition of the CKM matrix, up to a non-eliminable phase, leads to CP violation. The charged current lagrangian can be ultimately written as:

$$\mathcal{L}_{CC} = -\frac{g}{\sqrt{2}}\bar{u}_{L\alpha}\gamma^{\mu}V_{\alpha\beta}d_{L\beta}W^{+}_{\mu} - \frac{g}{\sqrt{2}}\bar{d}_{L\alpha}\gamma^{\mu}V^{\dagger}_{\alpha\beta}d_{L\beta}W^{-}_{\mu} \tag{1.43}$$

1.4 Quantum Chromodynamics

The Quantum Chromodynamics, QCD, is the gauge theory of strong interactions. It is a gauge theory analogous to the QED, where the $U(1)_Y$ symmetry group with the symmetry group $SU(3)_C$, where the subscript C stands for the charge associated with this symmetry, named colour. In the strong interactions, the colour is identified with the strong charge, so as the source of the cromodynamic field. The most relevant difference among $U(1)_Y$ and $SU(3)_C$ is that $U(1)_Y$ is an abelian group, while $SU(3)_C$ is not. The generators of the symmetry group $SU(3)_C$ do not commute and this leads to the introduction in the QCD lagrangian of interaction terms among the gauge fields, called gluons, that bring the charge of the group, the coloured charge. Instead in QED, photons do not have electrical charge and therefore a self-interaction terms do not exist in the QED. The QCD is invariant under local gauge transformations of $SU(3)_C$ group, i.e.:

$$\psi \rightarrow \psi' = e^{ig_s\vec{\alpha}(x)\cdot\vec{T}}\psi, \tag{1.44}$$

where g_s is the strong coupling constant, $\vec{\alpha}(x)$ are eight functions of the space-time coordinates and $\vec{T} = T^{\alpha}$ with $\alpha = (1, ..., 8)$ are generators of the $SU(3)_C$ group. The generators are related to the Gell-Mann matrices:

$$T^{a} = \frac{1}{2}\lambda^{a} \tag{1.45}$$

and follow the commutation rules:

$$[T\alpha, T\beta] = if_{\alpha\beta\gamma}T\gamma, \tag{1.46}$$

where $f_{\alpha\beta\gamma}$ are the structure constants of the group $SU(3)_C$ and the indices run from 1 to 8. Since the generators of $SU(3)_C$ are represented by 3×3 matrices, a new degree of freedom is needed, the colour. In this way the field ψ has three possible states labelled

[1] Analogously a mixing matrix can be introduced also for the neutrino sector, the PMNS matrix from Pontecorvo, Maki, Nakagawa e Sakata.

as red, green, and blue. The lagrangian of free quarks, assuming massless quarks, is:

$$\mathcal{L} = \sum_{f=1}^{6} \bar{\psi}^f i\gamma^\mu \partial_\mu \psi^f \tag{1.47}$$

The local gauge invariance under $SU(3)_C$ introduces 8 massless fields of gauge, gluons and the covariant derivative, $\mathcal{D}_\mu$, given by:

$$D_\mu = \partial_\mu + ig_s T_\alpha G_\mu^\alpha. \tag{1.48}$$

where G_μ^α are the 8 gluon fields that transform as:

$$G_\mu^\alpha \to G_\mu'^\alpha = G_\mu^\alpha + ig_s f^{\alpha\beta\gamma}\theta_b(x) G_{\gamma,\mu} \tag{1.49}$$

By adding the contribution of the kinetic energy for each gluon, one obtains the complete Lagrangian density for the QCD:

$$\mathcal{L}_{QCD} = \bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi - igs\bar{\psi}\gamma^\mu \lambda_a \psi G_\mu^\alpha - \frac{1}{4} G_\alpha^{\mu\nu} G_{\mu\nu}^\alpha \tag{1.50}$$

with $G_\alpha^{\mu\nu}$ the tensor field defined as

$$G_\alpha^{\mu\nu} = \partial^\mu G_\alpha^\nu - \partial^\nu G_\alpha^\mu - g_s f_{\alpha\beta\gamma} G^{\beta,\mu} G^{\gamma,\nu} \tag{1.51}$$

It is possible to associate each lagrangian interaction term with a Feynman diagram; Figure 1.4 shows all the fundamental transitions associated to the lagrangian in (1.50). Two relevant properties, which stem from experimental evidence and are well described in SM in the strong interactions sector, significantly differentiate the QCD from QED: colour confinement and asymptotic freedom. The asymptotic freedom has to cope with the experimental fact that no coloured hadrons are observed in nature. Hadrons therefore are colour singlets, since they are interpreted as bound states of quarks in the QCD parton model. This imposes restrictions on the types of bound quark state configurations that can exist. All this can be summarized by saying that the quark colour degree of freedom must be confined.

The colour confinement can be explained by looking at the running coupling constant form of strong force:

$$\alpha_s\left(|q^2|\right) = \frac{\alpha_s(\mu^2)}{\left[1 + \alpha_s(\mu^2)\frac{33-2N_f}{12\pi} \ln \frac{|q^2|}{\mu^2}\right]}, \tag{1.52}$$

with:

- q^2, transferred 4-momentum;

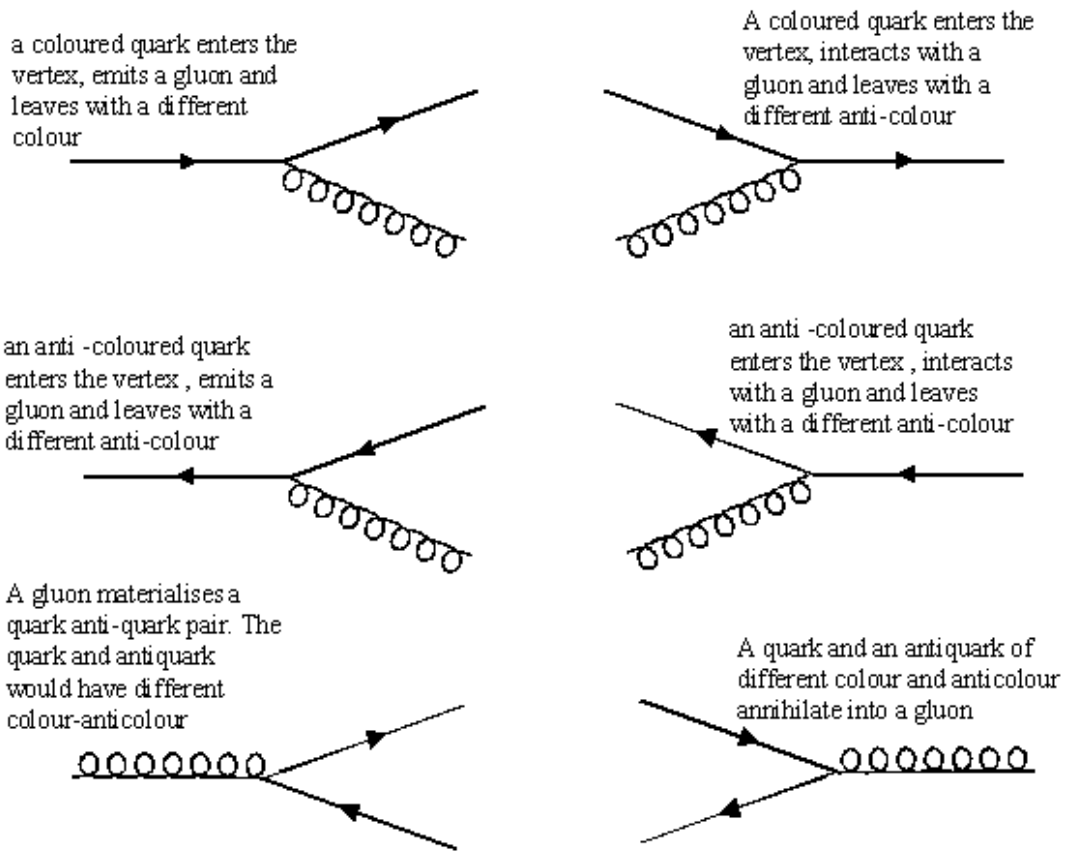


Figure 1.4: Basics vertices in Quantum Cromodynamics, QCD.

- μ, scale parameter for the strength of the coupling;
- N_f, number of fermions capable of strong interactions at the scale considered.

It can be seen that $\alpha_s(|q^2|)$ decreases as $|q^2|$ increases. For $|q| \sim 200$ MeV the value of α_s is large enough that any perturbative approach cannot be applied. In this region the calculations are carried on with the QCD-lattice approach. For increasing values of $|q^2|$, $\alpha_s(|q^2|)$ decreases and one moves towards a regime in which perturbative approach is a good approximation.

1.5 Unsolved issues in the Standard Model

Experiments in the last 30 years have tested the Standard Model theory in many ways: every predicted particle has been found, features of the electromagnetic and strong interactions have been described with accuracy, and in many cases, as for instance for the case of the electron gyromagnetic ratio [19], predictions of the SM were verified with a precision up to 12 orders of magnitude. However, the SM model fails in giving

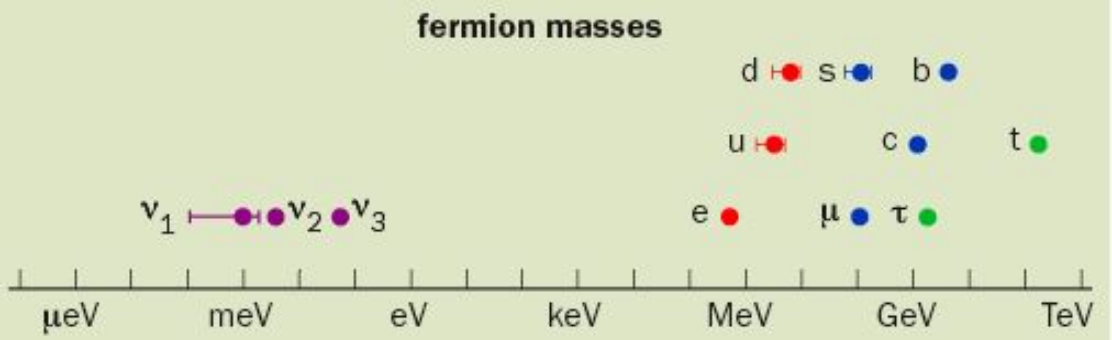


Figure 1.5: Pictorical view of the fermion masses differences.

explanation of other phenomena and observation in nature, for that it is not an all-encompassing and fully satisfactory theory of the known universe. Some of the most important questions that do not find satisfactory answer within the SM are:

- **Electroweak unification:** the GWS model cannot be considered a real unification theory because the symmetry group is the product of two different groups each one with its own constant, g and g' not linked by the theory. The ratio:

$$\frac{g'}{g} = \tan\theta_W$$

  has to be determined experimentally.

- **Large set of parameters:** in the model there are many parameters that are not postulated by theory and must therefore be obtained via measurements:
  - 3 coupling constants: g, g' and g_s;
  - fermion masses;
  - mass of Higgs boson;
  - CKM matrix elements;
  - PMNS matrix elements.

- **Fermion masses:** since there is no explanation or prediction of fermion masses a problem arise in this sector because of the differences of magnitudes, till 5 order of magnitudes between the top quark and the electron as can be seen in the Figure 1.5.

- **Hierarchy problem:** also known as naturalness, this problem is related to the radiative corrections which the Higgs mass receives through boson and fermionic loops, as can be seen in the Figure 1.6. Self-interaction terms arise such as:

$$m_H^2 \approx M_{H,0}^2 + \frac{\lambda}{4\pi^2}\Lambda^2 \tag{1.53}$$

where $M_{H,0}$ is the bare mass of the Higgs, λ is the strength of the coupling and Λ^2 is the scale of new physics. Since the Higgs boson has to couple to every massive particle, one can recalculate the one-loop corrections to Higgs mass as:

$$m_H^2 \approx M_{H,0}^2 + \frac{g_f}{4\pi^2}\left(\Lambda^2 + m_f^2\right) + \frac{g_s}{4\pi^2}\left(\Lambda^2 + m_S^2\right) \tag{1.54}$$

with g_f and g_S the coupling constant of fermion and scalar particles to the Higgs and m_f and m_S the related mass term. These corrections are quadratically divergent with the cut-off (that represents the scale beyond which new physics needs to be considered). Usually, a typical scale is the Planck scale where the corrections are $\sim 10^{30}\ m_H^2$, many orders of magnitude larger than the Higgs mass at tree level. A precise tuning between fermions and scalars, known as fine tuning, is required in order to reduce or cancel this divergences. This balance should also keep into account the tight constraints on the Higgs mass as shown in Figure 1.7.

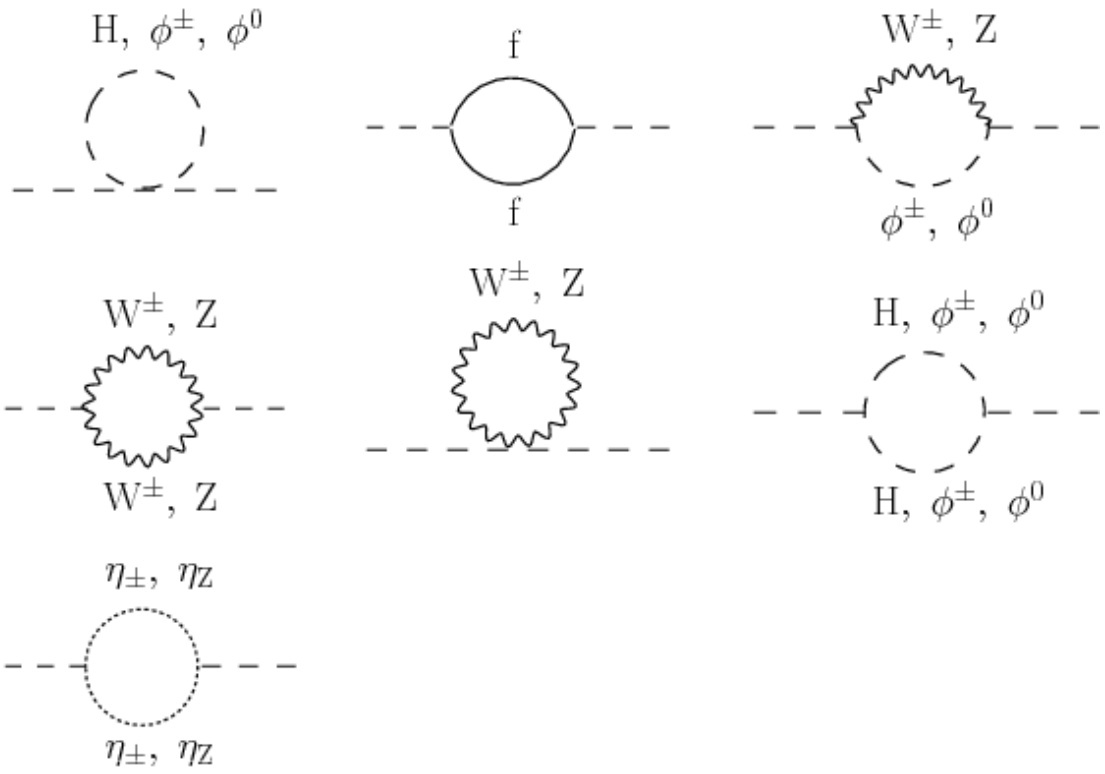


Figure 1.6: One-loop self-energy corrections to the Higgs mass.

- **Flavour Changing Neutral Currents (FCNC):** another observation not predicted by the SM is the presence of just three families of quarks and leptons that can be identified by the flavour quantum number. The suppression of flavour changing neutral currents at tree level, as expected from the Glashow, Iliopoulos and Maiani mechanism, GIM, is also entered by hand in the full theory.

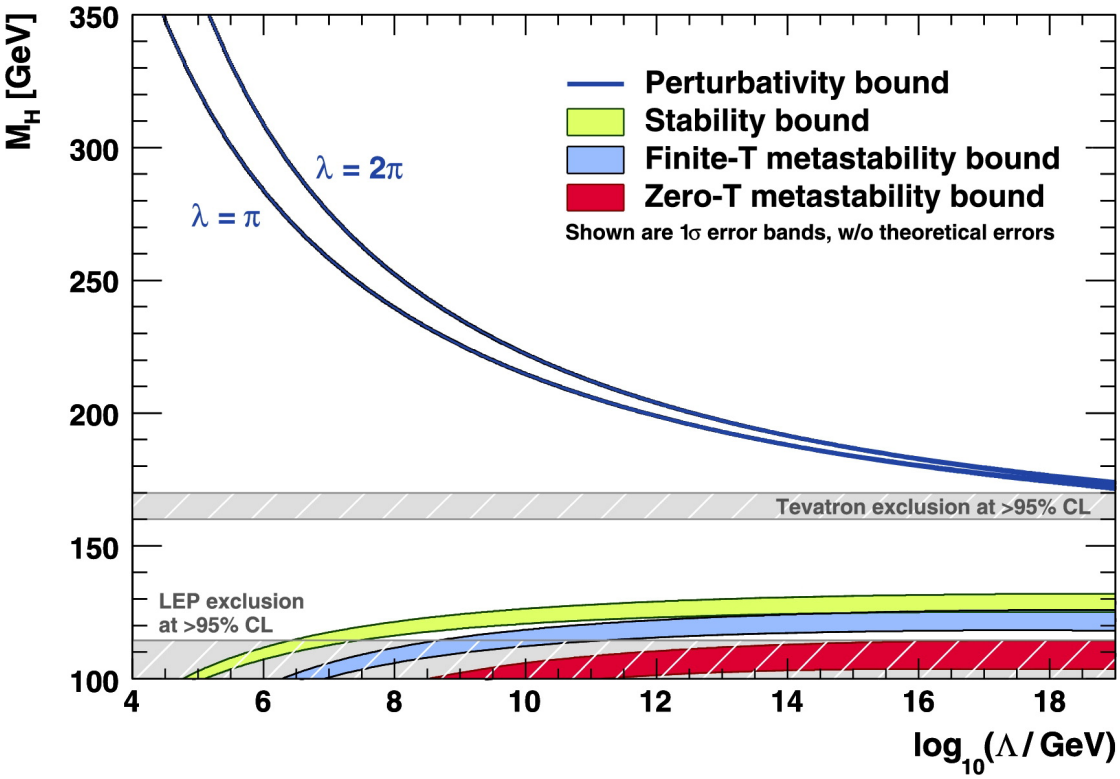


Figure 1.7: The scale Λ at which the two-loop renormalisation-group equations(RGEs) drive the quartic SM Higgs coupling non-perturbative, and the scale Λ at which the RGEs create an instability in the electroweak vacuum ($\lambda < 0$). The width of the bands indicates the errors induced by the uncertainties in and (added quadratically). The perturbativity upper bound (sometimes referred to as "triviality" bound) is given for $\lambda = \pi$ (the blue lower bold line) and $\lambda = 2\pi$ (the blue upper bold line). Their difference indicates the size of the theoretical uncertainty in this bound. The absolute vacuum stability bound is displayed by the light shaded in green band, while the less restrictive finite-temperature and zero-temperature metastability bounds are medium, in blue, and dark shaded, the red one, respectively. The grey hatched areas indicate the LEP [20] and Tevatron [21] exclusion domains.

- **Symmetries and conservation:** there are some experimental facts that have to be inserted ad hoc in the theory such as: electrical charge quantization, the proton stability and the conservation of the baryonic number which does not come from a symmetry. On the other side there are other charges such as the coloured one and the electrical one that are related to exact symmetries because they come from $U(1)_Y$ and $SU(3)_C$ groups.

- **Neutrino sector:** the SM assumes neutrinos to be massless, while observations of flavour oscillations can only be explained by massive neutrinos, via a mixing of the EW eigenstates. The seesaw mechanism, indeed, incorporates neutrino masses into the SM by introducing heavy Majorana neutrinos, whose masses are inversely

coupled to the light SM neutrino masses, hence motivating their small values of O(1 eV). Despite extensive searches, unfortunately the experimental proof is still pending.

- **Cosmological problem**: the SM in not able, trough the SM CP-violation in the quark sector to justify the actual matter-antimatter unbalance measured in the universe, neither to predict and include some results from astroparticle physics.

- **Gravity:** one of the most striking shortcoming of the Standard Model is that it lacks a description for gravity. Gravity was the first force to be fully understood over large distances, but it will likely be the last to be fully understood at very short distances. This is because the coupling strength of gravity is very weak if compared to other interactions, with a coupling constant that is 10^{34} times smaller than the electromagnetic coupling α. Gravitational effects would not be observable in particle collisions below a centre of mass energy close to the Planck scale (10^{19} GeV), bendingly larger than the energetic reach of modern particle colliders. Most theoretical models that could describe the Standard Model and gravity, sometimes referred to as theories of everything or ultraviolet (UV) completion models, manifest new phenomena only above a large energy called Λ_{UV}, roughly in the vicinity of the Planck scale.

- **Dark matter and dark energy:** astronomical evidences show that the universe is made up of only for the 5% of ordinary matter, while the rest does not correspond to the matter known. Indeed cosmological observation of galaxy rotation profiles provides evidence for a large amount of non SM-matter that weakly interact with SM particles. This matter is referred to with the name of dark matter, and should represent the 24% of the universe. The remaining 71% is ascribed to a constant vacuum energy called dark energy. The existence of the dark energy would account for the accelerating expansion of the universe.

Much thought has been put into how the Standard Model might be modified to understand these puzzles. The vast majority of the possible answers to the questions above fall into one or more of three broad classes of proposals:

- One consider the known fundamental fields but introducing new interactions. This road, which seems to be favoured when reference is made to the acquired experimental results, leads to Great Unification, Supersymmetry, string theory.

- Second parties consider new fundamental fields with new interactions ("compositeness", i.e. that some of the known particles might be composites of still-smaller things, condensed fermion-anti-dermal, "technicolour", "extended technicolour", preons and more).

- Finally, third parties consider the existence of "large extra dimensions", meaning that there could be more than the standard three dimensions to space (a likely possibility in string theory), and that some of these are large enough to be seen in high-energy accelerators.

Chapter 2

The CMS experiment at LHC

2.1 Physics motivation

Theoretical models, such as the Standard Model, indicate the most promising directions in the experimentation. In particle physics, accelerators are instruments of effective investigation to probe processes on a sub-atomic scale. Particles injected in accelerators, usually protons or electrons, are collimated into beams, accelerated, and then collided, either against a fixed target or against other beams. Hadron colliders, in particular, are well suited to the task of exploring new energy domains where new physic models, like supersymmetry, dark matter or extra dimensions, could manifest. Experimental hints point towards those new physics phenomena manifesting at the scale of TeV, which can be explored thoroughly if the proton energy and the luminosity are high enough. As shown in the Livingston diagram, Figure 2.1, the history of the accelerators is characterized by a continuous increasing of beam's energy to study new processes at higher energies. It can also be seen from the diagram, the Large Hadron Collider (LHC) is able to collide two proton beams at the design energy in the centre of mass equal to $\sqrt{s} = 14$ TeV. The beam energy and the design luminosity of the LHC have been chosen in order to study physics at the TeV energy scale. A wide range of physics processes is potentially at the reach of LHC with the seven-fold increase in energy and a hundred-fold increase in integrated luminosity over the previous hadron collider experiments. The usage of new technologies has led to an increase in the size of the machines and to the improvement of their properties: in particular, the use of superconductivity was decisive.

2.2 Large Hadron Collider

The Large Hadron Collider (LHC) [22] at CERN (European Centre for Nuclear Research) is a circular particle accelerator devoted to accelerating and colliding protons and heavy ions. It is located in the former Large Electron Positron (LEP) collider cave

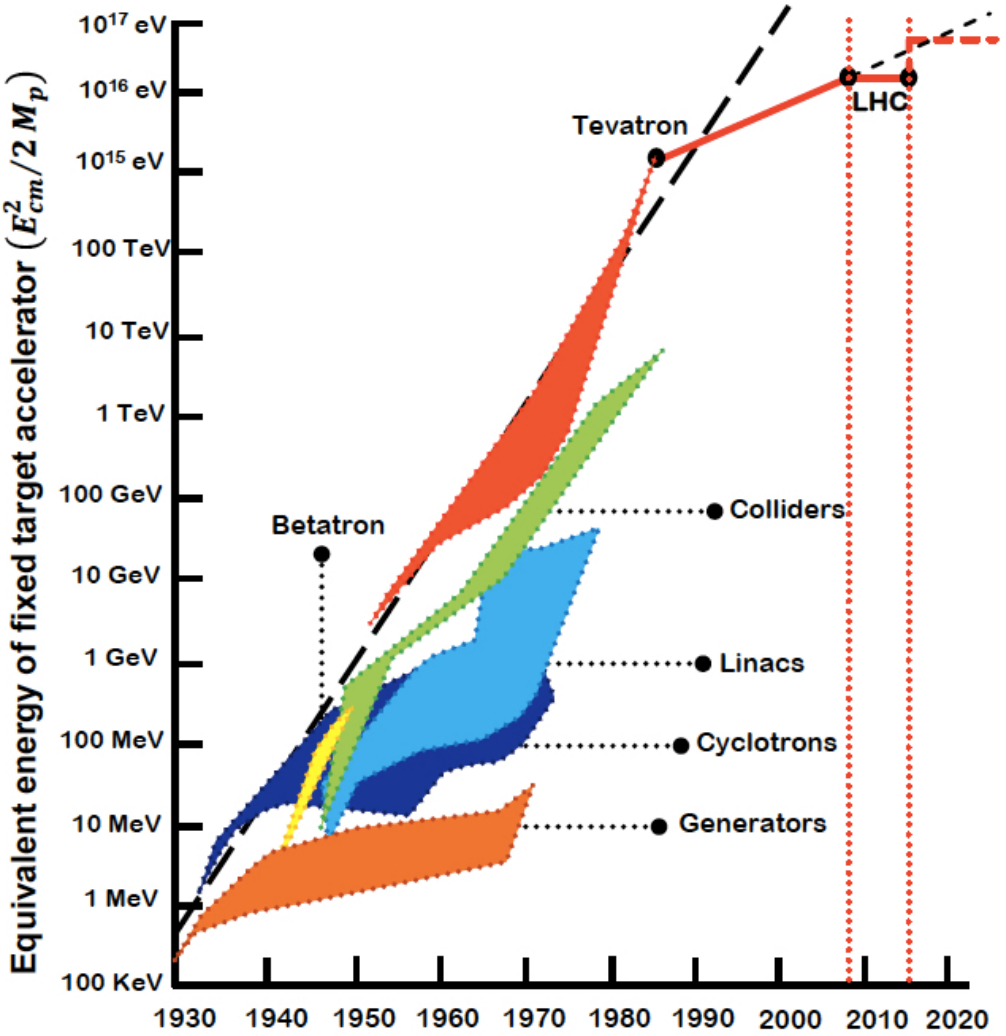


Figure 2.1: The Livingston chart depicting progress in collision energy through 2020. The LHC is the largest collision energy to date, but also represents the first break in the log-linear trend.

and consists of a 27 km tunnel in which there are two ring of superconducting magnets with a number of accelerating structures to boost the energy of the particles along the way. The tunnel is located between 45 m and 170 m below the ground surface, between Geneva airport and the Jura mountains, spanning the Swiss-French border. The LHC is the highest energy collider ever built because it is designed to provide proton-proton collisions with a center-of-mass energy of 14 TeV, with an unprecedent instantaneous luminosity of 10^{34} cm^{-2}s^{-1}. As said before, in addition to the proton-proton collisions, the LHC is also designed to provide lead ion collisions at a center-of-mass energy of 2.76 TeV per nucleon and a peak luminosity of 10^{27} cm^{-2}s^{-1}.

2.2.1 The accelerator complex

Thanks to the succession of LHC and of pre-existent machines it is possible to accelerate protons to an energy of 7 TeV, starting from an initial energy of 450 GeV.
A simple bottle of hydrogen gas is used as the proton source, then an electric field allow to strip hydrogen atoms of their electrons to yield protons. As shown in Figure 2.2 protons are injected into the first element of the chain, LINAC2, a linear accelerator which accelerates the protons to the energy of 50 MeV. The LINAC2 is followed by the Proton Synchrotron Booster (PSB) and by the Proton Synchrotron (PS), which push the beam to 1.4 and 25 GeV respectively. The Super Proton Synchrotron (SPS) is then used to reach 450 GeV, and inject the protons into the LHC as bunches of 1.15×10^{11} protons. LHC provides to accelerate bunches thanks to radio frequency (RF) cavities which accelerate the beam by 485 KeV at each turn. Each beam has 2808 circulating proton bunches, which are arranged in groups of 3 and 4 trains of 72 bunches, with 25 ns spacing within the train corresponding to 8 empty bunches between two trains. At every bunch crossing occour the collisions between the beams so the resulting maximum collisions rate is 40 MHz. The accelerator complex is also made up of superconducting magnets: 1232 dipoles which allows to keep the beams in the circular ring and 392 quadrupoles which are employed to focus the beams and sextupole, octupole and decapole which are used as spool piece correction. The magnetic field is generated by an electric current of 11.700 A for each dipole. The entire magnetic system is based on the niobium-titaniun (NbTi) Rutherford cables technology and works at a temperature of 2 K, obtained using the superfluid helium. The nominal strenght of such a magnetic field is of 8.33 T. Since the space in the tunnel is limited, a twin-bore design with a very complicated structure is used, in such a way that it is possible to use only one cryogenic structure with proton rings in the same cryostat, but this ultimately requires the presence of oppositely oriented magnetic fields to allow the coexistence of two proton beams along the same circumference. In the Figure 2.3 the twin-bore structure can be seen. The two beams are kept on parallel orbits and are brought together in a single beam pipe only near the interaction point (IP). The LHC complex system also requires vacuum systems: one for the insulation of the cryomagnets, another one for the helium distribution (QRL) and a a beam vacuum. The requirements for the beam vaccum are very stringent to guarantee the beam lifetime and to minimise the background at the experiments. The typical vacua at cryogenic temperatures in the IP requires a pressure around in the range 10^{-10} to 10^{-11} mbar.
The beams with protons rotate for many hours in the LHC beam pipes and are brought in collision in four interaction points (the yellow dots in Figure 2.2), in correspondence of the particle detectors in order to analyse the products of the collisions. The accelerator complex includes also other facilities like Antiproton Decelerator and the Online Isotope Mass Separator (ISOLDE) and feeds different physics project like the CERN Neutrinos

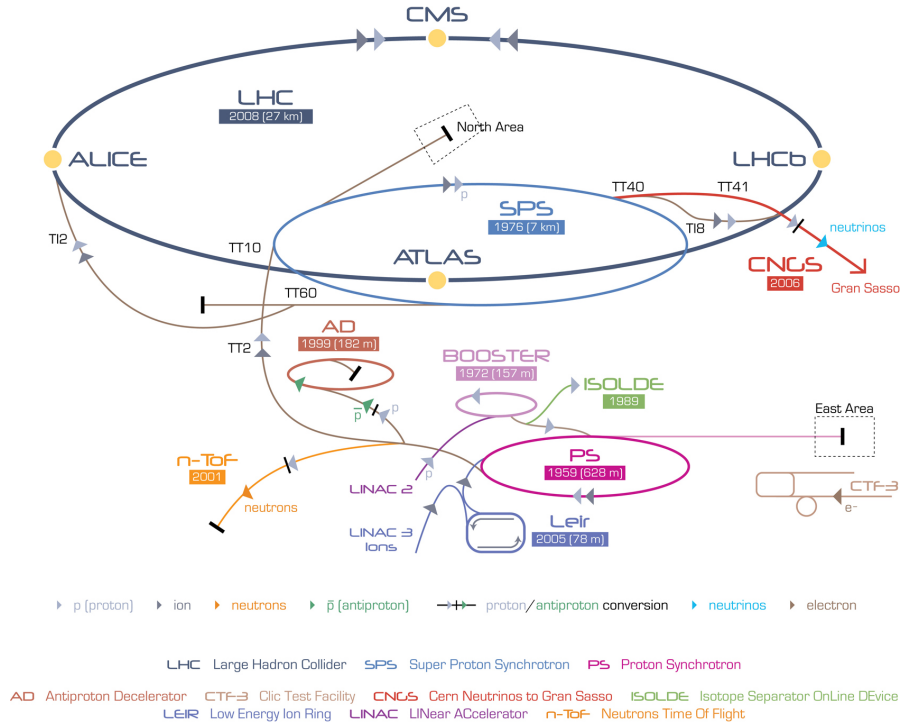


Figure 2.2: The CERN accelerator complex.

to Gran Sasso (CNGS) and the Compact Linear Collider test area.

2.2.2 Luminosity

One important parameter in particle experiments at colliders is the *instantaneous luminosity* $\mathcal{L}(t)$ which relates the cross section of a given process with the number of expected events N per unit of time in the collisions.

$$\frac{dN_{event}}{dt} = \mathcal{L} \cdot \sigma_{event} \tag{2.1}$$

where σ_{event} is the cross section for the process under study. In the LHC case, the beams which take part to the collisions have the same energy, and the assumption of a Gaussian distribution for protons in the transverse directions with respect to the beam yields. Therefore it is possible to write the instantaneous luminosity as a function of the accelerator parameters in the following way:

$$\mathcal{L} = \frac{N_b^2 n_b f_{rev} \gamma_r}{4\pi\epsilon_n \beta^*} F, \tag{2.2}$$

where:

- N_b is the number of particles per bunch,

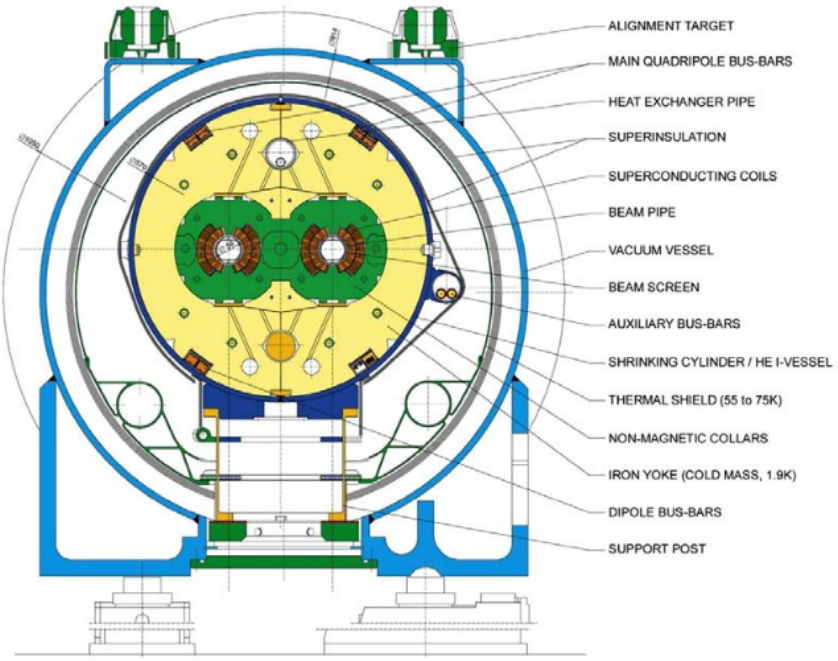


Figure 2.3: Standard cross section of the LHC superconducting dipole magnet.

- n_b the number of bunches per beam,
- f_{rev} the revolution frequency,
- γ_r the relativistic Lorentz factor $(1 - v^2/c^2)^{-1/2}$,
- ϵ_n the normalized transverse beam emittance,
- β^* the beta function at the collision point, is a measure of how narrow the beam is at the IP,
- F the geometric luminosity reduction factor due to the crossing angle at the interaction point.

The geometric luminosity reduction can be written as follow:

$$F = \left(1 + \left(\frac{\theta_c \sigma_Z}{2\sigma^*}\right)^2\right)^{1/2} \tag{2.3}$$

with θ_c the full crossing angle of the beams at the interaction point,σ_z bunch length and σ^* transverse RMS beam size at the interaction point.
The values of the above parameters are summarized for the operating period of 2016 in Table 2.1.

Parameter	Value
N_b	1.6 $\times 10^{11}$
n_b	1.6 2200
$f_{rev}[MHz]$	40 MHz
γ_r	4260
$\epsilon_n[\mu m]$	2.5
$\beta^*[m]$	0.6
$\theta_c[\mu rad]$	290
$\sigma_z[cm]$	9.4
$\sigma^*[\mu m]$	19

Table 2.1: The LHC machine parameters.

It is also possible to measure of how many collisions have occurred with the integrated luminosity:

$$L = \int \mathcal{L} dt \tag{2.4}$$

In this way the number of events for a process with given cross section in a data sample is given by:

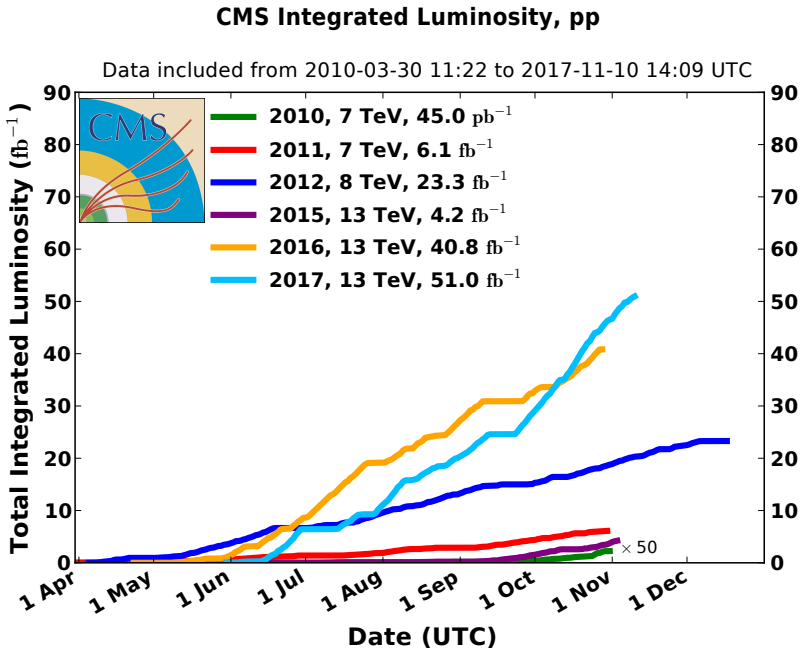
$$N_{process} = L \cdot \sigma_{process} \tag{2.5}$$

2.2.3 LHC Run 1 and Run 2

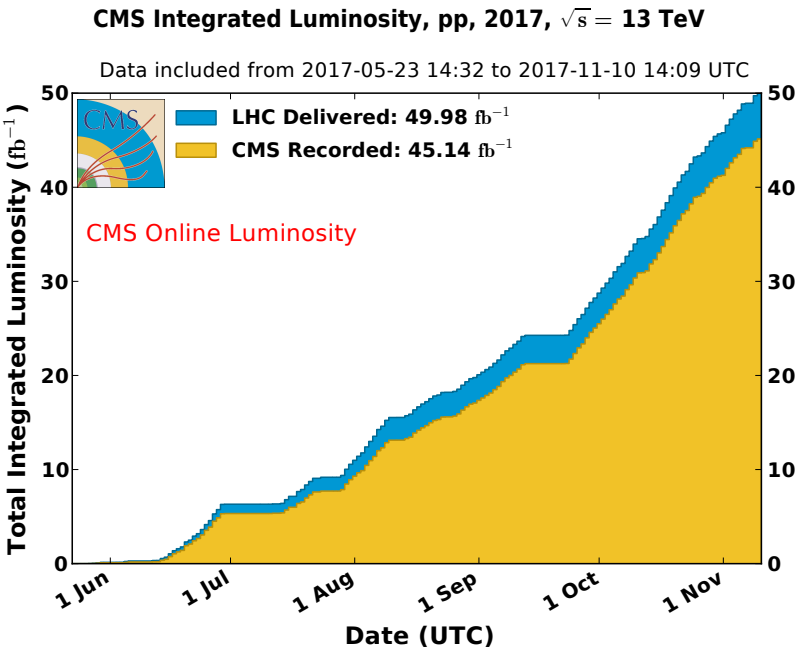
The first injections of beams started in September 2008, but due to a faulty resistance of an interconnection between two magnets, an accident occurred which forced to stop the LHC for more than one year for repairs and for introducing further safety measures. The operations restarted in November 2009 with collisions at a centre-of-mass energy of 900 GeV, then the energy of the beams was raised up to 3.5 TeV. The data taking proceeded in the years 2010 and 2011 and the LHC continued to run at $\sqrt{s} = 7$ TeV and then reach the 8 TeV in the 2013. The 2010-2013 data taking period is referred to as Run 1. After the end of Run 1, in 2013, the LHC stopped for detector upgrade and maintenance operations, and restarted, after a long shutdown, in 2015, with collisions at a centre-of-mass energy of 13 TeV. The phase that started in 2015 until nowadays is referred to as Run 2. In Table 2.2 details of the LHC luminosity recorded by CMS and certified as good for physics analysis. In Figure 2.4(a) and 2.4(b) the total luminosity delivered by LHC and collected by CMS during 2017 and a comparison of the luminosity of pp collision data collected in Run 1 and Run 2 are shown.

2.2.4 LHC experiments

As mentioned, the beams interact at four different points along the ring, where the largest experiments are located in order to study the the products of the high-energy



(a) Cumulative luminosity versus day delivered to CMS during stable beams for pp collisions at nominal center-of-mass energy. This is shown for data-taking in 2010 (green), 2011 (red), 2012 (blue), 2015 (purple), 2016 (orange), and 2017 (light blue).



(b) The plot show the measured luminosity delivered by the LHC to CMS (blue) and recorded by CMS (orange) during stable beams and for proton-proton collisions at a 13 TeV center-of-mass energy in 2017.

Period	$\sqrt{s}$ [TeV]	LHC delivered [fb^{-1}]	CMS Recorded [fb^{-1}]	CMS Validated [fb^{-1}]
Run 1 (2010)	7	40.22×10^{-2}	40.76×10^{-2}	34.68×10^{-2}
Run 1 (2011)	7	6.13	5.55	5.09
Run 1 (2012)	8	23.30	21.79	19.79
Run 2 (2015)	13	4.22	3.81	2.39
Run 2 (2016)	13	40.82	37.76	35.92
Run 2 (2017)	13	49.98	45.14	41.86

Table 2.2: The cumulative luminosity delivered by LHC, recorded by CMS and certified as good for physics analysis, starting from 2010 to 2017.

collisions. As illustrated in the Figure 2.2 in a clockwise order there are

- ATLAS: *A Toroidal LHC ApparatuS* a general-purpose detector whose targets are precision measurements of SM, the search and the study of the Higgs boson, and mechanisms due to new physics. Its lenght is 44 m, has a 25 m diameter and it is the biggest experiment at LHC with a weight of 7000 tons.
- ALICE: *A Large Ion Collider Experiment* an ion-ion collision experiment whose purpose is to explore the initial state of matter. To form the quark-gluon plasma are required high energy densities, which are achieved by colliding lead ions with $\sqrt{s} = 2.67$ TeV at a peak luminosity of $L = 10^{27}$ cm^{-2}s^{-1}.
- CMS: *Compact Muon Solenoid* is described in section 2.3.
- LHCb: *LHC-beauty* an experiment built for the study of the b quark properties, its production mechanism, and to probe rare decays in B mesons including the CP -violating processes. LHCb requires clean events with low pile-up while it is not necessary a large amount of luminosity, and it works with asymmetric beams: one is at the LHC full energy (up to 7 TeV) and the other is at the injection energy (450 GeV).

2.3 CMS experiment

The Compact Muon Solenoid experiment is one of the four great experiments at the pp accelerator LHC. It is a so-called "general-purpose" experiment since its main goal is to explore physics at the TeV energy scale from the Standard Model measurements, including the Higgs boson and its properties, high precision tests of QCD, flavour physics and electroweak interactions, to new physics searches in channels like Supersymmetry or searches for new vector bosons and quarks (the so-called exotic searches).
CMS has a compact design because of its relatively small size considering its weight:

about 14.000 tonnes for 15.00 meters of diameter and 28.7 meters of length 2.4. It takes its name from the huge superconducting solenoid which generates an internal magnetic field of 3.8 Tesla, about 10 5 times the magnetic field of the Earth. While working at high luminosities such as the LHC ones does indubitably allow to study rare processes, but also introduces problems related to the presence of inelastic collisions and therefore of multiple scattering. For these reasons, the detector has been designed to be capable of operating in a high radiation environment and to be able to distinguish different process and different particles of interest from the background, so to be able to design analyses where the signal to background production rate is strongly disfavoured even by several orders of magnitude.
In order to cope with these challenges, CMS was designed to have:

- **Radiation robustness:** the high-rate radiation implies consequence on the detectors which have to sustain a severe amount of radiation, and have to maintain good performances over the course of several years of data taking.
- **Geometrical coverage:** full azimuthal coverage is required to make hermetic the detector. In this way it is possible the kinematic closure of the events in the plane transverse to the beams collision plane and therefore the measurement of the missing transverse energy (MET) is possible.
- **Trigger efficiency:** the huge number of events that happen in a bunch crossing has to face with limitations in the bandwidth at which data can be transferred to the storage facilities; this implies a reduction which has to be performed with an electronic triggering system.
- **High granularity and good time resolution:** the huge amount of particles to be detected requires detectors with high granularity in order to avoid, or at least to limit, the overlap between different particles of the same event or coming from interactions in the same bunch crossing. The effect of this pile-up can be reduced also using fast electronic elements.

These challenges were coped with by making use of a system of several sub-detectors to identify different particles on a wide energy and angular coverage, and influenced the choice of detector layout and geometry that is made up of cylindrical layers coaxial to the beam pipe, called barrel layers, and two endcaps that ensure hermitical closure of the detector.

2.3.1 The coordinate frame at CMS

The CMS reference frame is defined as follows:

- The coordinate frame is centred at the nominal interaction point;

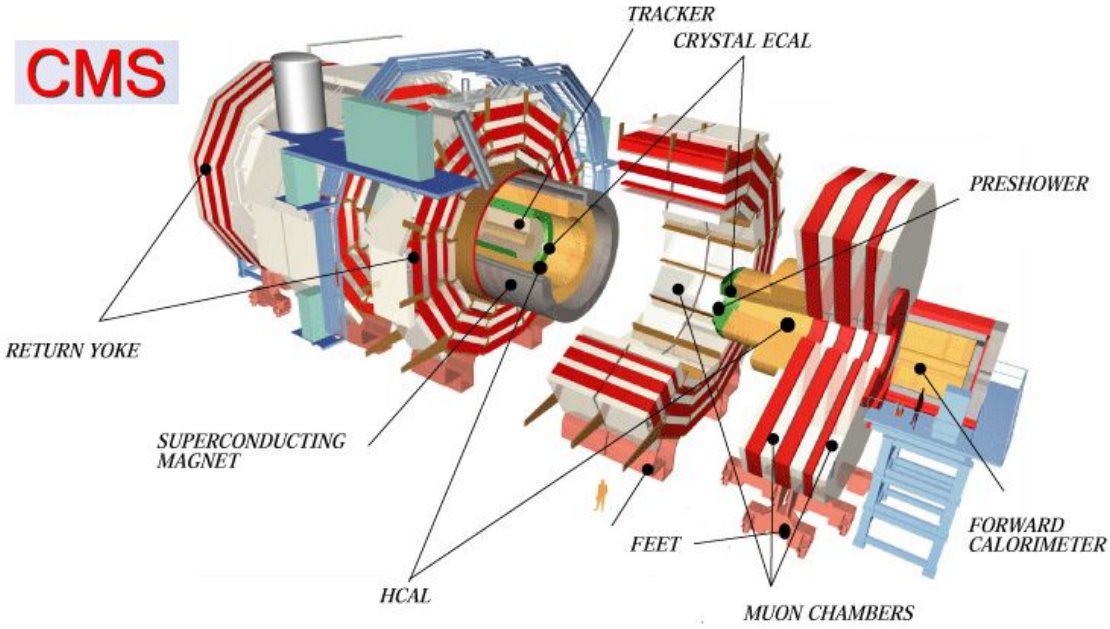


Figure 2.4: CMS 3D view.

- $x - axis$ points to the centre of the LHC;
- $y - axis$ points upwards, perpendicular to the LHC plane;
- $z - axis$ along the anticlockwise-beam direction.

By exploiting the cylindrical symmetry of the detector it is possible to introduce a pseudo-angular reference system, as seen in Figure 2.5, defining: the radial distance from the $z - axis$, r; the azimuthal angle taken from the $x - axis$, ϕ; and the pseudorapidity, η, defined as:

$$\eta = -ln\left(tan\frac{\theta}{2}\right) \tag{2.6}$$

where θ is the polar angle. The pseudorapidity comes close to the rapidity:

$$y = \frac{1}{2}\ln\left(\frac{E + p_z c}{E - p_z c}\right) \tag{2.7}$$

where E is the energy of the particle and p_z is the particle momenta along the $z-$axis, when $E \gg m$, so the pseudorapidity is used for ultra relativistic particles. The pseudorapidity and the rapidity are both natural variables for describing angles in a system where the initial momentum along the $z-$axis is unknown, and differences in rapidity Δy (or in pseudorapidity$\Delta\eta$ in the limit of massless particles), are invariant under boosts along the $z-$axis. Referring to this system, the distance between two particle direction

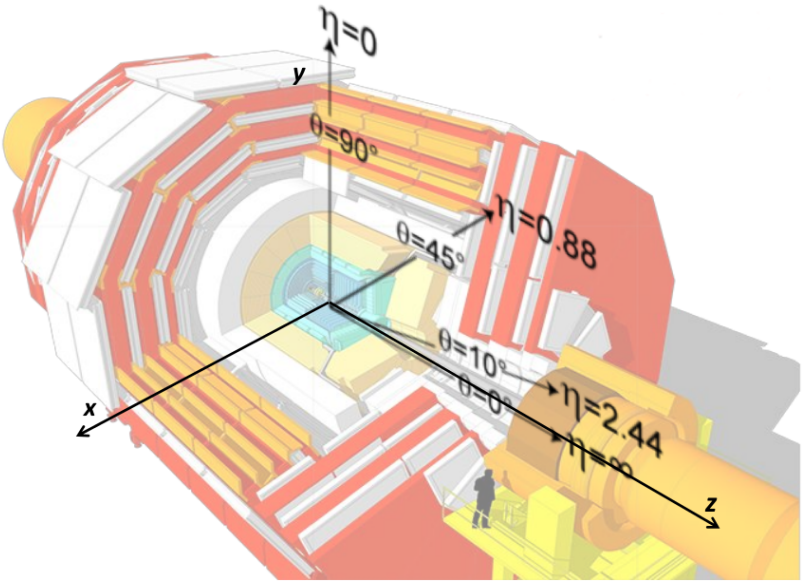


Figure 2.5: The CMS coordinate system: The $x-axis$ points to the centre of the LHC, the $y-axis$ points upwards, perpendicular to the LHC plane, and the $z-axis$ along the anticlockwise-beam direction. In this figure plot of pseudorapidity as a function of the polar angle, θ, is shown: as angle increases from zero, pseudorapidity decreases from infinity.

can be written as another Lorentz invariant variable, in the following way:

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} \tag{2.8}$$

Usually two important variables are p_{T} and E_{T}, referred to the Cartesian system, they are respectively the transverse momentum and the transverse energy, defined as:

$$\vec{p_T} = \sqrt{\vec{p_x}^2 + \vec{p_y}^2} \tag{2.9}$$

$$E_T = E \sin\theta \tag{2.10}$$

2.3.2 The CMS subdetectors layout

The CMS detector is made up of several layers of detectors centred on the interaction point as can be seen in Figure 2.6. From the inner to the outer part of the detector we have:

- **The Tracker System**: designed to provide a precise and efficient measurements of the trajectories of charged particles emerging from the LHC collisions. The CMS tracker consist of a silicon pixel detector and a silicon strip detector.

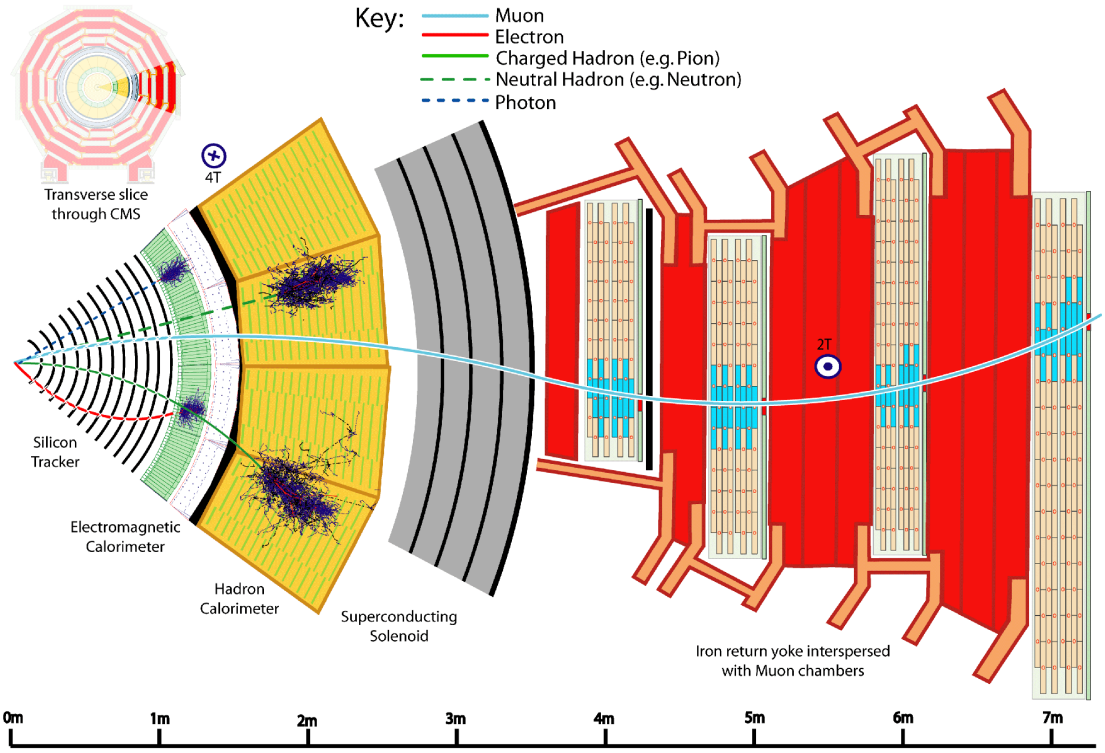


Figure 2.6: The CMS subdetectors system.

- **The Electromagnetic Calorimeter (ECAL)**: for accurate electron and photon energy measurement.
- **The Hadronic Calorimeter (HCAL)**: crucial for energy measurements of jets and missing energy, provides energy measurements for charged and neutral hadrons. Thanks to the tracker it manages to distinguish between charged and neutral.
- **The Superconducting Solenoid**: the coil generating an internal constant magnetic field of 3.8 Tesla in the direction of the beam axis.
- **The Return Yoke**: interspersed with the Muon system, sustain the structure and it is studied to allow magnetic field lines of the solenoid.
- **The Muon System**: designed to have the capability of reconstructing the momentum and charge of muons over the entire kinematic range of the LHC. CMS uses three types of particle detectors for muon identification:
  - Drift Tube detectors;
  - Resistive Plate Chambers;
  - Cathode Strip Chambers.

2.3.3 The Tracking System

The CMS detector has been designed to detect cleanly the diverse signatures of Standard Model as well as new physics at the LHC, and this is done by identifying and precisely measuring energy and momentum, where is possible, of muons, electrons, photons, and jets over a large energy range. Robust tracking and detailed vertex reconstruction within a strong magnetic field are the two main performances requirements which the tracker must satisfy. The Tracker system must at the same time be sufficiently radiation hard in order to guarantee the operation of all sub-detectors which his is made up of over the full data taking period of the experiment. Due to its high segmentation the CMS Tracker produces high quality seeds for the track reconstruction algorithm offline, and manages to identify unambiguously tracks coming from multiple vertices, and is also used to perform fast tracking online in the high level trigger (HLT) for primary vertex reconstruction, electron/photon identification, muon reconstruction, tau identification, and b-tagging. The CMS Tracker [23, 24] has a radius of 1.25 m and a length of 5.8 m, and its acceptance is of $|\eta|$ <2.5. It is composed by two main sub-detector elements:

- The Pixel vertex detector system, Figure 2.7: used for accurate measurement of the vertex position, provides a two-dimensional measurements of the hit position in the module planes.
- The Silicon Strip (SST), Figure 2.8: Used for accurate track reconstruction.

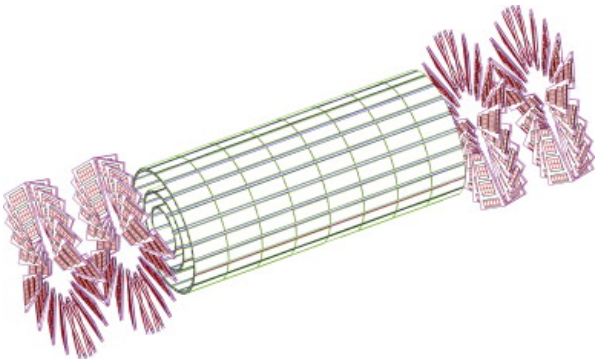


Figure 2.7: Schematics of CMS Silicon pixel system.

These two detector technologies are arranged in concentric cylindrical volumes, as can seen in the Figure 2.9 and are all fast on the scale of 25 ns to limit the PU events to a single bunch crossing. The region closest to the beams interaction point is surrounded by silicon Pixel detectors, and the cell of each one is 150μm $\times$ 150μm. They are disposed in a barrel geometry, in the central rapidity region, with three barrel layers (BPIX) and with two forward/backward disks (FPIX) at higher values of the rapidity into end-caps. In general the silicon Pixel, due to their position, have to stand a very high level of radiation

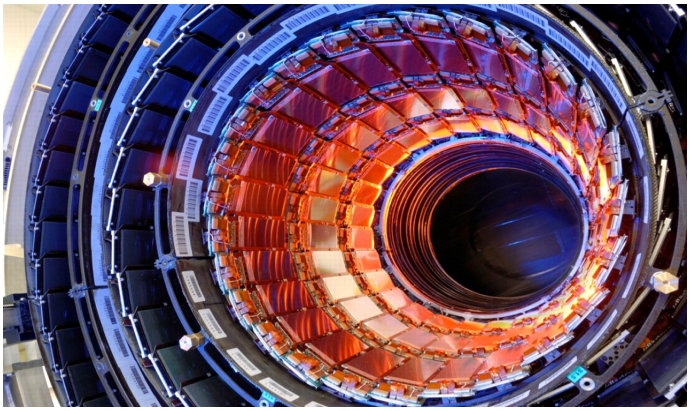


Figure 2.8: Tracker showing silicon strips detectors in the barrel module.

and also a very high track rate and particle fluencies. Going outwards with respect to the interaction point, the intermediate part is instrumented with Silicon Strip and it is made up of 4 barrel layers parallel to the beam axis, Tracker Inner Barrel (TIB), and 3 disks at each end of the TIB, the Tracker Inner Disks, called TID. The outer volume, outside the TIB/TID, is also composed by SST but with different thick and pitches and is called Tracker Outer Barrel, TOB. At both ends of the TOB are located other Tracker EndCaps named TEC+ and TEC-, whose signs indicate the location along the $z-$axis. The region covered by TEC corresponds to the $|\eta| < 2.5$. One of the problems which could happen at higher luminosity as discussed before, is a pile-up increase, which can cause an increase of fake rates in the tracker. To cope with this higher fake rates of events an upgrade is foreseen [25] whose main goal is to maintain the detector performances even in case of 150-200 of PU interactions per bunch crossing PU. The upgrade program includes:

- Replacement of the pixel detector with a four-layer high-data-rate design;
- Improvements to the L1-Trigger system with higher granularity and additional processing capabilities.

2.3.4 The Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) [26] is an hermetic and homogeneous calorimeter made up of 75848 scintillating crystals of Lead tungstate ($PbWO_4$). The choice of such material allows to ultimately achieve high granularity, fast radiation, radiation hardness and a very compact system thanks to their high density (8.28 g/cm^3), short radiation length (0.89 cm), and small Molière radius (2.2 cm). The crystals are arranged, as shown in Figure 2.10, in a barrel section, forming the ECAL barrel (EB), and two endcaps, the ECAL EndCaps EE+ and EE- , where the sign indicates the location along the $z-$axis. The barrel part covers the pseudorapidity range $|\eta|$ <1.479 and

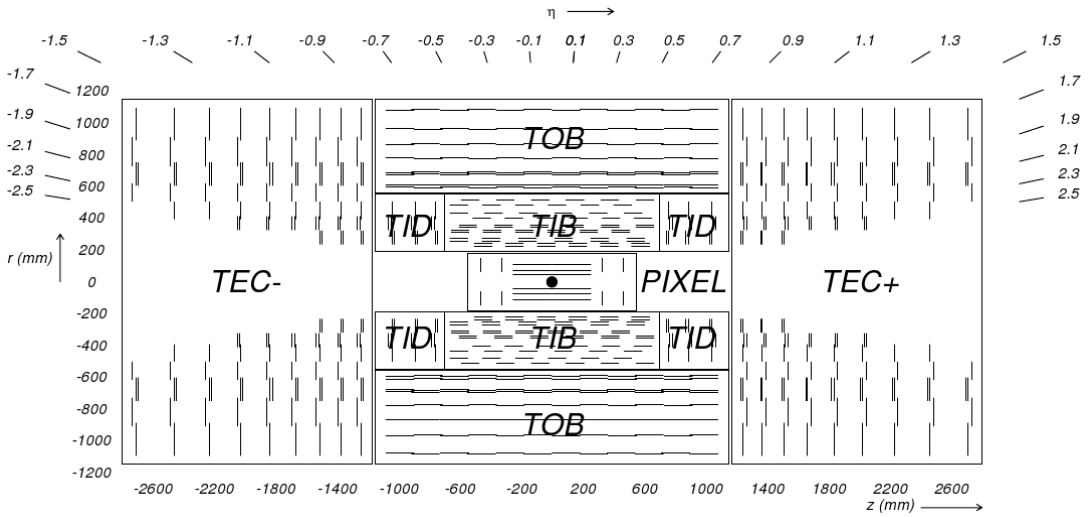


Figure 2.9: Schematic view of the CMS tracker.

each crystal in this part has a cross-section of ∼22×22 mm^2 and 23 cm corresponding to 25 X_0. The endcap part cover a pseudorapidity range from 1.479 to 3.0, allowing for high precision energy measurements up to $|\eta|$ =2.6. In this region crystals have a cross-section of ∼30×30 mm^2 and 22 cm corresponding to 24.7 X_0 each, and are grouped together in 36 units called supercrystals, each containing 1700 crystals. Since the yield

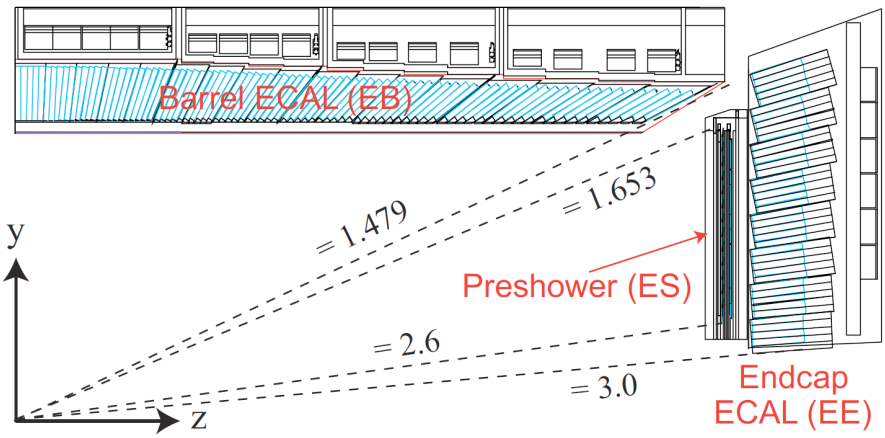


Figure 2.10: Schematic view of the CMS ECAL.

of light depends on the temperature, all the ECAL is maintained at the temperature of 0.1 ℃ from a dedicated cooling system. The light is collected by photodetectors, converted to an electrical signal and then amplified. The amplification is achieved in the barrel region with avalanche photodiodes (APDs), since they are able to provide gain in the presence of the high transverse magnetic field, while in the endcaps by vacuum phototriodes (VPTs) because there the radiation is too high to use a silicon photodiode.

The ECAL contains also a Preshower detector instrumented in front of the endcaps which allows to increase precision in the position determination, so to distinguish between single high-energy photons and close pairs of low-energy photons and therefore to prevent false signals. It forms a disk and is made up of two lead planes followed by silicon sensors strip arranged in a grid in the endcaps in order to cover the crystal endcap.
The ECAL energy resolution can be parametrized as:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{a}{\sqrt{E}}\right)^2 + \left(\frac{b}{E}\right)^2 + (c)^2 \tag{2.11}$$

where:

- a represents the stochastic term: it takes in account the statistical fluctuations of the signal in the shower containments such as fluctuations in the number of primary particles and/or the number of photons which includes fluctuations in the shower containments collected by a photomultiplier;
- b is the noise term which contains the contributions from electronic noise and PU energy, negligible at low luminosity;
- c is the constant term which takes in account the fluctuations of the longitudinal leakage, of intercalibration errors and of leakage of energy from the back of the crystal.

The energy resolution of the Electromagnetic Calorimeter of CMS is:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E}}\right)^2 + \left(\frac{0.12}{E}\right)^2 + (0.30\%)^2\,, \tag{2.12}$$

where E is in GeV.

2.3.5 The Hadron Calorimeter

The Hadron Calorimeter (HCAL) [27] not only does measure the energy of hadrons such as protons, neutrons, pions, and kaons, but also provide to an indirect measurement of non-interacting, uncharged and almost massless, particles such as neutrinos. Like the ECAL, also the HCAL is an hermetic calorimeter to ensure the capture of every emerging particle from the collisions, but by converse it is a sampling calorimeter and allows to determine a particle's position, energy and arrival time. The HCAL is made up of alternating layers of active material and of absorber, but since it is inserted into a magnetic field the absorber must be made from a non-magnetic material. While the active material is fluorescent scintillator, the absorber is made of brass. The brass absorber is commonly known as cartridge brass (C26000) and is composed of 70% Cu

and 30% Zn, its density is 8.83 g/cm^3 with radiation lenght of X_0=1.49 cm and nuclear interaction length of λ_I=16.42 cm. The HCAL cover the pseudorapidity range $|\eta|$ <3 and at higher η values it is complemented by the Forward Calorimeter detectors. It is organised into Hadron Calorimeter Barrel (HB and HO), Endcap (HE) and Forward (HF) sections as can be seen in Figure 2.11. The HB and HE regions are located around the ECAL respecting the concentric cylindrical geometry. The Hadron Calorimeter Barrel covers the region η <1.3, and is divided into two half-barrel sections, HB- and HB+ in the verse of the $z-$axis. The HB is made up of alternating plates of brass absorber, segmented into four azimuthal angle sectors and has 16.42 cm interaction length, and scintillator tiles embedded with wavelenght shifting (WLS) fibers which are spliced into 16 $|\eta|$ sectors also segmented in η and ϕ coordinates. The Hadron Calorimeter Endcaps cover the pseudorapidity range 1.3< $|\eta|$ <3, its design had to keep into account high particles rates of the order of the MHz. The HF is located 11.2 m away from the interaction point, allowing to include the pseudorapidity range 3< $|\eta|$ <5.2, and is made up of quartz fibres embedded within a 165 cm long steel absorber. The HF uses a Cherenkov-based technology, due to its location in a very high radiation area and a very high rate environment. Additional scintillators compose the Hadron Calorimeter Outer(HO), and are placed outside the solenoid to ensure adequate sampling depth and to measure late shower development: they are considered as "tail catchers".

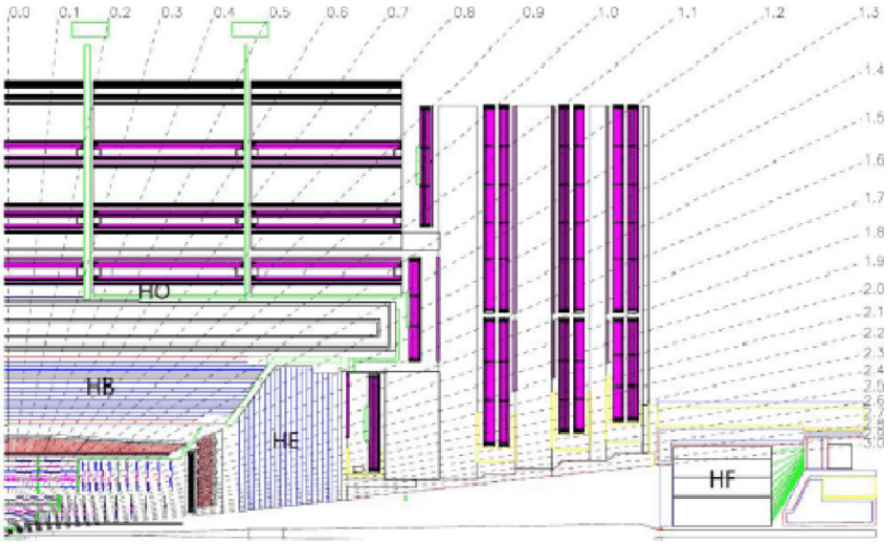


Figure 2.11: View of the CMS detector showing the HCAL subsystems (HB, HE, HO, and HF).

2.3.6 The Superconducting Magnet and the Return Yoke

The CMS Magnet is a solenoid made up of coils of wire that produce a uniform magnetic field of 3.8 T when electricity flows through them, and works in superconducting conditions. The Magnet is the instrument around which the whole experiment is built and its main task is to bend the paths of particles emerging from LHC high-energy collisions. Once the trajectory is known it is possible to measure the particle's momentum and its charge. The Tracker and the Calorimeters are instrumented inside the magnet while the Muon System is located outside the coil. Moreover, to ensure uniform field lines with an intense and constant magnetic field, outside and surrounding the coil, the iron Return Yoke is located. The Return Yoke is composed of alternating layers, interspersed with the muon detectors.

2.3.7 The Muon System

As suggested from the name of the detector itself, CMS, the muon detection is one of CMS's main design goals. Muons are indeed important particles as they have a clean signature, and can appear in final states of many new physics processes. Their properties make them good candidates for precision measurements: muons with energy below the TeV scale lose energy mainly due to ionization, so they can penetrate several meters without interacting with the matter in the tracker and in the calorimeters, in this way also a muon of about 5 gev can reach the outside of the CMS detector. Moreover, muons are charged particles so in the tracking system a precise measurement of their momentum can be performed.
The Muon System provides muon identification, triggering, and momentum reconstruction, and is located in the region outside the magnet. In this outer region, the presence of the return yoke allows the presence of a magnetic field of 1.8 T in the opposite direction with respect to the one inside the magnet. In this way, by collecting hits in the muon system, it is possible to reconstruct the muon path also adding the information coming from the tracker system. The particle path is bent from the magnetic field, thus allowing for a measurement of its momentum, as the path can be totally reconstructed and the value of the magnetic field is known.
The Muon System is made up of different detection technologies:

- 250 drift tubes (DTs);
- 549 cathode strip chambers (CSCs)
- 610 resistive plate chambers (RPCs)

chosen because they are relatively inexpensive with respect to other tracking technologies given the volume they need to occupy, reliable and robust to the radiation, even if the

radiation rate they have to sustain is inferior with respect to the innermost detectors. DTs are used in the barrel, covering the region with pseudorapidity $|\eta|$ <1.2, and are placed into 4 stations located among the layers of iron yoke plates. The first 3 stations each contain 2 layer of 4 chambers which measure the muon coordinate in $r-\phi$ bending plane, and one layer of 4 chambers measuring the z coordinate. Each cell of each chamber is separated by a half-cell width with respect to their neighbour to eliminate dead spots in the efficiency. In this way it is possible to measure passing muon with excellent time resolution, the number of chambers and their orientation were chosen to allow good efficiency for linking together muon hits from different stations into a single muon track and for rejecting hits coming from the background. In the endcap, where both muon and background rates are higher and the magnetic field is not uniform, the CSC are chosen, because of their good segmentation, their fast response time, and their radiation resistance. The CSCs cover the range 0.9< $|\eta|$ <2.4 and the cathode strips of each chamber provide position measurements in the $r-\phi$ bending plane. The anode wires allow to measure the pseudorapidity and the beam-crossing time of a muon. Muon identification is ensured over the range corresponding to $10^\circ < \theta < 170^\circ$, since the muon detector cover the pseudorapidity region $|\eta| < 2.4$ with no acceptance gaps. Tipically the offline reconstruction efficiency of simulated single-muon samples is $95-99\%$, while in the regions around $|\eta| = 0.25$ and 0.8, i.e. the regions between 2 DT wheels, and $|\eta| = 1.2$: which is the transition region between the DT and CSC systems, the efficiency drops. Negligible punchthrough reaches the Muon system due to the amount of material in front of the Muon system, which exceeds 16 interaction lengths. The multiple-scattering in the detector material before the first muon station influences the offline muon momentum resolution of the standalone muon system, which comes about 9% for small values of η and p for transverse momenta up to the hundreds of GeV. At higher energies $\sim$ TeV the standalone momentum resolution varies between 15% and 40%, depending on $|\eta|$. So if at low momenta a global fit, using also the inner tracker, improves the momentum resolution by an order of magnitude, at high momenta (1 TeV) both detector parts yield a momentum resolution of about 5%. A crucial property of the DT and CSC subsystems is that they have each a good efficiency in p_{T} of muons triggering and high background rejection. The Level-1 trigger p_{T} resolution is about 15% in the barrel and 25% in the endcap. To take in account the high background rate, the identification and triggering system was complemented with the RPC that is placed both in the barrel and in the endcap. RPCs have fast response, excellent time resolution but coarser position resolution than DTs and CSCs. 6 RPC layers are located in the barrel region, 2 on each of the first two stations of DT and 1 on the last 2 stations. In this way it is possible to use the trigger algorithm even for low-p_{T} tracks that may stop before reaching the outer 2 stations. To use the trigger coincidences between stations in order to reduce the background, to improve the time resolution for bunch crossing identification, and to

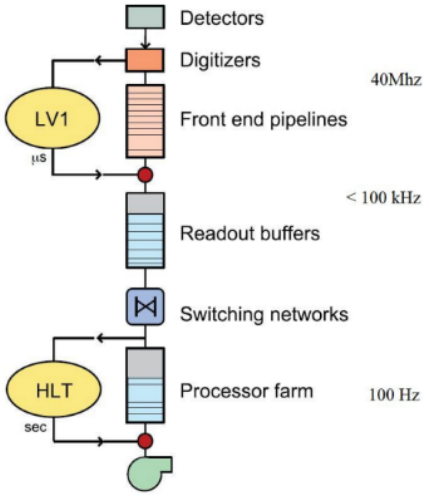


Figure 2.12: Architecture of the CMS trigger system, including the ingoing/rates relative to each step.

achieve a good p_T resolution, there is a plase or RPC in each of the 3 stations in the endcap region. An alignment system is used to optimize the muon momentum resolution is used an alignment system to measure positions of among the muon detectors and with respect to the inner tracker.

2.3.8 The Trigger System

Given the high interaction rates at the LHC(see also Section 2.2.2), a drastic reduction is needed, as it would be impossible to store entirely the information from all LHC collision events. So a drastic rate reduction is needed because is impossible to store and process the large amount of data associated with the resulting high number of events. In this way it is possible to select the potentially interesting events and reduce the rate to just a few hundred "events" per second, which can be read out and stored on computer disk for subsequent analysis.
This important task is performed by the trigger system in two steps called Level-1 (L1) Trigger [29] and High-Level Trigger (HLT) [30], Figure 2.12 shows the scheme of the overall architecture of the CMS Trigger system. These two levels are implemented largely with programmable electronics, for the former trigger, and in a software system with filter farm of about thousand commercial processors in the latter case. Together they manage to reduce the rate by a factor of 10^6. The Level-1 Trigger system is designed for a fast, automatic response by looking for simple signs of interesting physics (e.g. particles with a large amount of energy). It uses coarsely segmented data from the calorimeters and the muon system, while holding the high-resolution data in pipelined memories in the front-

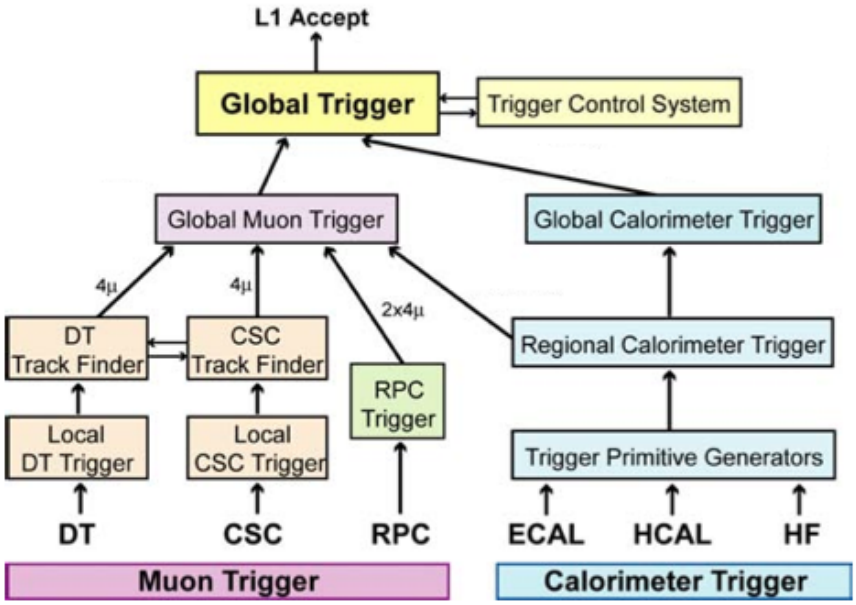


Figure 2.13: Architecture of the Level-1 Trigger.

end electronics, it is implemented with the FPGA technology, ASICs and programmable memory lookup tables (LUT) in order to guarantee speed and radiation resistance since it has to analyse every bunch crossing. The Level-1 trigger is organized in local, regional and global components. The local components are also called Trigger Primitive Generators (TPG), and they are based on energy deposits in calorimeter trigger towers and track segments or hit patterns in muon chambers. The regional triggers, using a pattern logic, based on the detailed knowledge of the detectors and trigger electronics, manage to combine their information to sort trigger objects like electron, photon or muon, in limited spatial regions. The global components, Global Calorimeter and Global Muon Triggers, determine the highest-rank calorimeter and muon objects across the entire experiment and transfer them to the the top entity of the Level-1 hierarchy. The decision is based on algorithm calculations and on the readiness of the sub-detectors and the DAQ, which is determined by the Trigger Control System (TCS). The Level-1 Accept (L1A) decision is communicated to the sub-detectors through the Timing, Trigger and Control (TTC) system. The architecture of the L1 Trigger can be seen in Figure 2.13. The allowed L1 Trigger latency, between a given bunch crossing and the distribution of the trigger decision to the detector front-end electronics, is 3.2 μs. The processing must therefore be pipelined in order to enable a quasi-deadtime-free operation. The L1 Trigger electronics is housed partly on the detectors, partly in the underground control room located at a distance of approximately 90 m from the experimental cavern. The final decision to reject or to accept an event is taken by the HLT that has access to the complete read-out data and can perform an analysis quite similar to the analysis off-line.

Chapter 3

New physics Models and Vector Like Quarks

As it was discussed in Chapter 1, the Standard Model does not manage to explain all observed phenomena in high energy physics. Several promising theories do however exist that provide explanations in scenarios where the SM fails to provide predictions, and foresee the existence of new particles in the LHC energy reach. In this Chapter in particular two broad scenarios will be described that concretely realize naturalness: Supersymmetry and Compositeness, and one of their observable consequences, the presence of new particles named Vector Like Quarks (VLQs).

In the Supersymmetry case, there is a chiral symmetry protecting from large radiative corrections which give rise to fermionic "superpartners", for bosons. The boson masses are related to the ones of their fermionic superpartners. The Composite Higgs models propose another solution, as the mechanism that protect the Higgs mass in the composite scenario is dimensional transmutation: the Higgs is supposed to be a composite state of a new-strongly interacting sector. In this case, the dimensionality of the Higgs mass operator can be $d > 4$, and in this scenarios there are no problems in explaining why the Higgs mass is relatively light.

Both theories, Supersymmetry and Compositness, identify a new sector, the strong sector, from which arise new particles such as the VLQs. These particles are coloured fermions called partners of the SM quarks, and, specially in the top sector, they do play an important role. They are also known as "top-partners" and they manage to help in the resolution of the hierarchy problem, as they help to the computation of the radiative corrections to the Higgs mass with the mechanisms described above.

3.1 Supersymmetry

The hierarchy problem is solved by Supersymmetry theory with the introduction of new particles. The Supersymmetry predicts for each SM particle a new particle, a "supersymmetric partner", with spin differing by 1/2. A supersymmetric transformation turns a bosonic state in a fermionic one and viceversa. So for example for a SM fermion is introduced a new boson S. All of these particles couples with the Higgs boson and affect the Higgs corrections such as:

$$\Delta m_H^2 = \frac{g_S}{16\pi^2}\left[2\Lambda^2 + \mathcal{O}\left(m_S^2 ln\left(\frac{\Lambda}{m_S}\right)\right)\right] \tag{3.1}$$

with g_S is the Yukawa coupling of the new boson S, Λ is a cutoff and m_S is the new boson mass. This correction has opposite sign to the fermion contribution therefore, assuming $\mid g_f \mid = g_S$, all the fermion terms have a counter term that naturally cancels the quadratic divergence introduced. Moreover, ignoring the logarithmic correction the residual term to the Higgs mass would be:

$$\Delta m_H^2 = \frac{y_f^2}{16\pi^2} \mid m_S^2 - m_f^2 \mid . \tag{3.2}$$

The naturalness arguments require that the size of the correction is expected to be smaller than the Higgs mass, so:

$$\mid m_S^2 - m_f^2 \mid \leq 1\text{TeV}^2 \tag{3.3}$$

This result can be taken as the range of validity of the Standard Model, implying that at the TeV scale the superpartners of the SM particles can be produced and the Supersymmetry extends the SM at those energies. This range, clearly, does not fix the upper bound of the supersymmetric extensions but is just a way to stabilize the correction to the Higgs mass.

Through the Supersymmetry it is possible to accomplish the cancellation of the SM corrections to the first order. It is predicted that the mass of the superpartners is the same as the SM particles but, since there is not any observation yet, Supersymmetry is considered as a broken symmetry, in this way supersymmetric particles masses have to be above the reach of current experiments.

The extension of the SM through a Supersymmetry is not unique: the number of generators in the symmetry group, as well as the composition and arrangement of the SM particles into supermultiplets allow many possibilities. Many SM extensions can be derived from the Supersymmetry such as the Minimal Supersymmetric Standard Model (MSSM), which introduces the minimal amount of new particles amongst all possible supersymmetric extensions and consists of one single operator in the symmetry group.

In the MSSM every SM particle is paired with one single superpartner:

- fermion partners are denoted with the prefix "s", e.g. *stop* is the superpartner of the top quark;
- partners of the SM bosons are labelled with the suffix "ino", e.g. *gluino* is the superpartner of the gluon.

In the Higgs sector the MSSM requires the introduction of an additional complex doublet, in order to produce five particles after giving mass to the SM bosons. Table 3.1 shows the arrangement and notation of the MSSM particles content. Moreover the MSSM can

Names	Spin	P_R	Gauge Eigenstates	Mass Eigenstates
Higgs bosons	0	+1	$H_u^0\, H_d^0\, H_u^+\, H_d^-$	$h^0\, H^0\, A^0\, H^\pm$
Squarks	0	-1	$\widetilde{u}_L\, \widetilde{u}_R\, \widetilde{d}_L\, \widetilde{d}_R$ $\widetilde{s}_L\, \widetilde{s}_R\, \widetilde{c}_L\, \widetilde{c}_R$ $\widetilde{t}_L\, \widetilde{t}_R\, \widetilde{b}_L\, \widetilde{b}_R$	same same $\widetilde{t}_1\ \widetilde{t}_2\ \widetilde{b}_1\ \widetilde{b}_2$
Sleptons	0	-1	$\widetilde{e}_L\, \widetilde{e}_R\, \widetilde{\nu}_e$ $\widetilde{\mu}_L\, \widetilde{\mu}_R\, \widetilde{\mu}_\nu$ $\widetilde{\tau}_L\, \widetilde{\tau}_R\, \widetilde{\nu}_\tau$	same same $\widetilde{\tau}_1\ \widetilde{\tau}_2\ \widetilde{\nu}_\tau$
Neutralinos	1/2	-1	$\widetilde{B}^0\, \widetilde{W}^0\, \widetilde{H}_u^0\, \widetilde{H}_d^0$	$\widetilde{\chi}_1^0\, \widetilde{\chi}_2^0\, \widetilde{\chi}_3^0\, \widetilde{\chi}_4^0$
Charginos	1/2	-1	$\widetilde{W}^\pm\ \widetilde{H}_u^+\ \widetilde{H}_d^-$	$\widetilde{\chi}_1^\pm\ \widetilde{\chi}_2^\pm$
Gluino	1/2	-1	$\widetilde{g}$	same

Table 3.1: Particles foreseen from MSSM.

contain also operators that violate baryon and/or lepton number, allowing for example the proton decays, but the non-observation of such decays forbids the existence of those operators and imposes very stringent upper limits on their coefficients. The introduction of the discrete symmetry called *R-parity* is a way to avoid these operators. The quantum number associated is P_R:

$$P_R = (-1)^{3(B-L)+2s} \tag{3.4}$$

with B and L are respectively the baryon and the lepton quantum numbers and s the spin of the particle. P_R is a multiplicative quantum number and it is set to be $+1$ for all the particles and -1 for the supersymmetric particles, s-particles. If P_R is exactly conserved in the scattering and in the decay processes, a mixing amongst particles and s-particles cannot take place, therefore in each vertex there could be only an even number of s-particles. This fact has important implications on the phenomenology:

- Superpartners can only be produced in pairs.
- The Lightest Supersymmetic Particle (LSP) have to be stable. If it is not charged, it interacts only weakly with the ordinary matter and it is a good candidate for the non baryonic dark matter required in the cosmological models. The lightest neutralino is the most promising candidate for the LSP, even if at the moment cannot be excluded scalar neutrinos, gravitino or other particles considered in models that do not retain the R-symmetry.
- Each s-particle different from the LSP can eventually decay in a state that contains an odd number of LSP.

The supersymmetric partners, mix giving rise to the mass eigenstates after electroweak symmetry breaking. The neutral higgsinos ($\widetilde{H}_u^0$ and $\widetilde{H}_d^0$) and the neutral gauginos ($\widetilde{B}$ and $\widetilde{W}^0$) combine to form four mass eigenstates named neutralinos. The charged higgsinos ($\widetilde{H}_u^+$ and $\widetilde{H}_d^-$) and the winos ($\widetilde{W}^+$ and $\widetilde{W}^-$) mix to form two mass eigenstates with electric charge ± 1, named charginos.
In the sfermion sector, mixing across generations can cause large contributions to flavour changing neutral current that are excluded by observations, so it is usually removed. However, the mixing between the left-handed and the right-handed sfermions of the same generation depends on the mass of the SM fermion, and therefore cannot be neglected for the third generation superpartners. The mass eigenstates obtained through the mixing are labelled as $\widetilde{q}_1$, $\widetilde{q}_2$.
The requirement of R-parity conservation in MSSM if on one hand provides a solution to the hierarchy problem and contains a good candidate for dark matter on the other hand leads to the introduction of 105 new parameters, to be added to the 19 parameters of the SM. In order to reduce the number of parameters to be considered, several simplifications are introduced in collider searches: usually, only the s-particles that contribute to a particular final state are considered and the rest are considered heavy enough so that they can be completely decoupled.

3.2 The Composite theories

Many theories propose that particles currently considered as elementary are made up of other, yet unknown, constituents which are strongly coupled forming new heavy reso-

nances. For example, in the QCD theory the flavour chiral symmetry of the Lagrangian is broken spontaneously, generating three massless scalar pseudo-Goldstone bosons. A further explicit symmetry breaking is operated by the quark masses, and gives mass also to the pseudo-Goldstone bosons which is, however, much lighter than other mesons in QCD. The three pseudo-Goldstone bosons are the $\pi^{\pm}$ and π^{0} particles, which are not elementary but composed of a quark-antiquark pair.

3.2.1 The Composite Higgs boson theory

The Model of a composite Higgs boson is one of the composite theories mentioned, it was proposed by Georgi and Kaplan [32, 33] in 1984 and takes into account that the Higgs boson is a composite pseudo-Goldstone Boson, i.e. it is a bound state of a new strongly interacting sector. In the composite sector thanks to the breaking of the symmetry, also due to a small mixing with the SM sector, a pseudo-Goldstone boson is produced, the Higgs boson. The Higgs boson produced in this way is much lighter than the scale of the new sector and, in this scenario, its mass is "protected" from Planck corrections and the Hierarchy problem is surpassed. In fact the Higgs boson reveals its composite nature at the energy scale of the new strong sector, so the radiative corrections to its mass do not have to reach the Planck scale.
Many problems characterize these theories for strong interactions, such as the fact that usually they hardly pass the electroweak precision tests, or explain the origin of fermion masses. Throughout these years the models of pseudo-Goldstone Higgs in the framework of a five-dimensional theory have received increasing attention since the theory is weakly coupled, making it possible to perform calculations, and it can satisfy the bounds from electroweak data.

3.2.2 The Composite Top quark theory

The Compositeness it is also extended to the top sector indeed certain models propose that the top quark is not an elementary particle, but rather a composite or condensate state. As said before, in models with composite particles SM particles can get their masses by mixing with composite states thanks to the new strong sector. Given the large mass of the top quark it would be natural to expect the top quark to have a sizeable admixture of the composite state and therefore to show properties of Compositeness. Moreover, the results of the electroweak precision test strongly constrains the possibility of a composite left-handed top, so most models focus on right-handed composite top quarks [34].

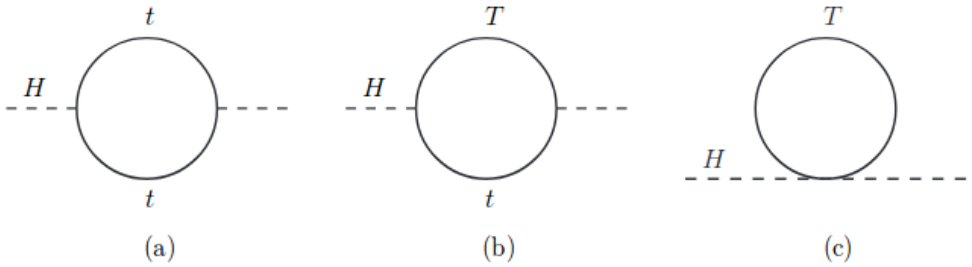


Figure 3.1: One-loop contributions to the Higgs boson mass term from (a) the top quark and (b,c) a vector-like top partner.

3.3 Vector Like Quarks

Several models predict the existence of Vector-Like Quarks (VLQs), defined as spin 1/2 fermions that transform as triplets under the colour gauge group, and whose left- and right-handed chiral components have the same transformation properties under the weak-isospin gauge group.
At present there is no evidence of the existence of vector-like quarks, nevertheless they are key ingredients of physics beyond the Standard Model (BSM), since from a theoretical point of view they are considered in different models.
At the TeV scale the vector-like quarks are required, if the Higgs boson is a pseudo-Goldstone boson, to induce electroweak breaking and explain the observed lightness of the Higgs. They also emerge as fermion resonances in the partial-Compositeness theory of flavour. Due to the large Yukawa coupling of the top quark, both mechanisms give rise to a sizeable mixing of the new quarks with the third family of SM quarks, hence their are usually called "top partners".
VLQ also arise as Kaluza–Klein excitations as excitations in warped extra-dimension scenarios of SM quarks propagating in the bulk and in grand unified theories based on the E_6 group, but in this case they are not guaranteed to be near the TeV scale.
The introduction of vector-like quarks also stabilizes the Higgs boson mass since thanks to their contributions the quadratic divergences disappear and only remains only the logarithmic divergence. The one-loop contributions to the Higgs boson mass are shown in Figure 3.1. VLQs are the simplest example of coloured fermions, since they present the same colour charges of SM quarks, still allowed by experimental data, moreover since they do not receive their masses from Yukawa couplings to a Higgs doublet are consistent with existing Higgs data.
In fact, since they do not receive mass from Yukawa couplings, their mass term can be directly inserted into the Lagrangian without breaking the gauge symmetry, and this makes them unique because their coupling to the Higgs field is unrelated to their mass.

Moreover there are no constraints on the existence of vector-like quarks arising from the measured production cross section of the Higgs because the contribution to loop-induced Higgs boson couplings, i.e. ggH and $\gamma\gamma H$, is suppressed by the heavy quark mass. The values of VLQs masses can naturally have values higher than the electroweak symmetry breaking scale.
They can mix with the Standard Model quarks and then modify their coupling to the Z, W, and Higgs bosons. In fact the addition of vector-like quarks to the SM is the simplest way to break the the Glashow-Illiopoulos-Maiani mechanism allowing the flavour-changing neutral currents at the tree-level. Under this aspect they also can be considered as a new way to introduce new sources of CP violation.
The VLQs come in four different forms depending on their charge as can be seen in the Table 3.2. It is also possible to classify VLQs in multiplets of $SU(2)_I$ and so to write

VLQ	**Electric charge**
X	+5/3
T	+2/3
B	-1/3
Y	-4/3

Table 3.2: Charge assignment for VLQs.

the gauge-invariant terms for the representations. These new particles can appear as $SU(2)_I$ singlets, doublets or triplets, as can be seen in the Table 3.3. Another way to represent the VLQs is the following:

$$\begin{array}{ll} Singlets & T^0_{L,R}, B^0_{L,R} \\ Doublets & \left(XT^0\right)_{L,R}, \left(T^0B^0\right)_{L,R}, \left(B^0Y\right)_{L,R} \\ Triplets & \left(XT^0B^0\right)_{L,R}, \left(T^0B^0Y\right)_{L,R} \end{array}$$

The subscript indicates the left- and right- handed components, while the zero superscript indicates weak eigenstates to distinguish them from mass eigenstates. For X and Y the weak and the mass eigenstates coincide since they cannot mix via the mass matrix due to their exotic charges. The extensions of the Stadard Model can be considered taking into account all the VLQs multiplets, assuming that no other new physics modifies the electroweak observables, i.e.: when new fields $T^0_{L,R}$ are added to the SM, the resulting up−type mass eigenstates (u, c, t, T) might then contain non-zero components of the $T^0_{L,R}$ fields; this could translate in a variation of the couplings with the Z boson. These deviations are constrained for u and c quarks by the LEP experiments while are much weaker for the top quark. Moreover the mixing is proportional to the ratio between the mass of the SM and the VLQ, m_Q/M_{VLQ}, that is negligible for the first two generations, these facts, added to the large Yukawa coupling of the top quark, which highlights a link

Multiplet	**Hypercharge**
Singlets	
(T)	+2/3
(B)	-1/3
Doublets	
(X,T)	+7/6
(T,B)	+1/6
(B,Y)	-5/6
Triplets	
(X,T,B)	+2/3
(T,B,Y)	-1/3

Table 3.3: Hypercharge assignment for Vector-like quarks in different $SU(2)_I$ representations.

between the top quark and new physics related with electroweak symmetry breaking and the fermion mass hierarchy, allow to assume that the VLQs mix only with the third generation of quarks.
Phenomenologically the signatures of VLQ have been analysed in literature using both independent perspective and specific scenarios. As said a distinctive feature of VLQ is the presence of flavour changing neutral currents which lead to a wide range of possible final states, which are still analysed in detail in order to drive the experimental search of these new states.

3.3.1 VLQ production mechanisms

The VLQ production cross section in pp collisions strongly depends on their mixing with SM quarks, particularly on the square of the couplings to the W or Z bosons. The production mechanisms of VLQ can be grouped in the following way:

- Pair Production: dominated by QCD production via gluons. The cross section of this type of process is model independent as it only depends on the mass of the new fermion and decreases for higher masses due to PDF suppression. The process of pair production through QCD interactions is completely analogous to pair production of SM top quarks, and only depends on α_S and the mass of the heavy quark:

$$gg, q\bar{q} \rightarrow Q\bar{Q}$$

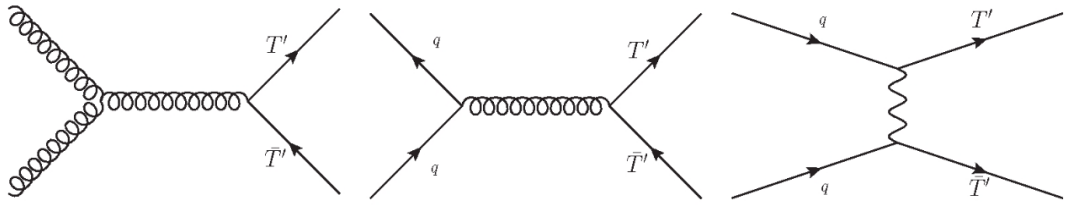


Figure 3.2: Feynman diagrams for pair production of vector-like top quarks via gluon and W,Z or Higgs bosons.

where $Q = T, B, X, Y$. There is also a small contribution of electroweak gauge bosons, which are sub-leading in terms of cross section. If on one hand production via Z and γ is completely negligible because have the same final states as QCD production, on the other hand, production via charged current, so via $W^{\pm}$ boson, leads to interesting channels like:

$$\bar{q}q' \to W^{+} \to \bar{T}X, \bar{B}T, \bar{Y}B$$

$$\bar{q}q' \to W^{-} \to T\bar{X}, B\bar{T}, Y\bar{B}$$

The electroweak cross sections are very small and also strongly suppressed for large masses, therefore their impact on the search strategies can be safely neglected. In any case the cross section are model-dependent as they depend on the representation the VLQ belong to. Another potentially relevant production process is mediated by the W or Z or Higgs in the $t - channel$ and is represented by the production of a pair of VLQ, QQ'. This process is completely absent in QCD and, depending on subsequent decays, it can give rise same-sign dileptons or to the following final states: TT, BB, XB, TB, TY with peculiar kinematics. Some Feynman diagrams for pair production of T can be seen in Figure 3.2.

- Single production: it is an electroweak interaction dominated process and can happen in association with top quarks, jets coming from any light quark or boson (including $W^{\pm}, Z$, and the Higgs H). It is sensitive to both the fermion mass and its mixing parameters with SM particles. Single production via electroweak interaction is subdominant process for masses below $m_Q \sim 800 - 1000$ GeV, but becomes important for higher masses due to phase-space suppression of pair production. It also depends on the couplings between the new quarks and the W and Z bosons:

$$qq' \xrightarrow{V^*} qQ$$

with $V = W, Z$ while the contributions of the Higgs will always be suppressed by the small masses of the light quarks. Figure 3.3 shows the single electroweak

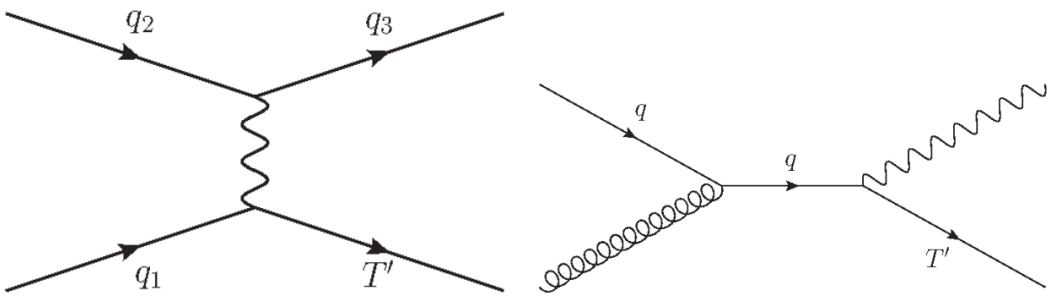


Figure 3.3: Feynman diagrams for single production of Vector-like top quarks.

interaction mediated by a vector boson in t-channel in association with a SM quark and the one mediated by SM quark in association with a vector boson.

The behaviour of the cross section with the energy, and so with the mass of the vector-like quarks and their multiplets, for the pair production and the single production can be seen in Figure 3.4.

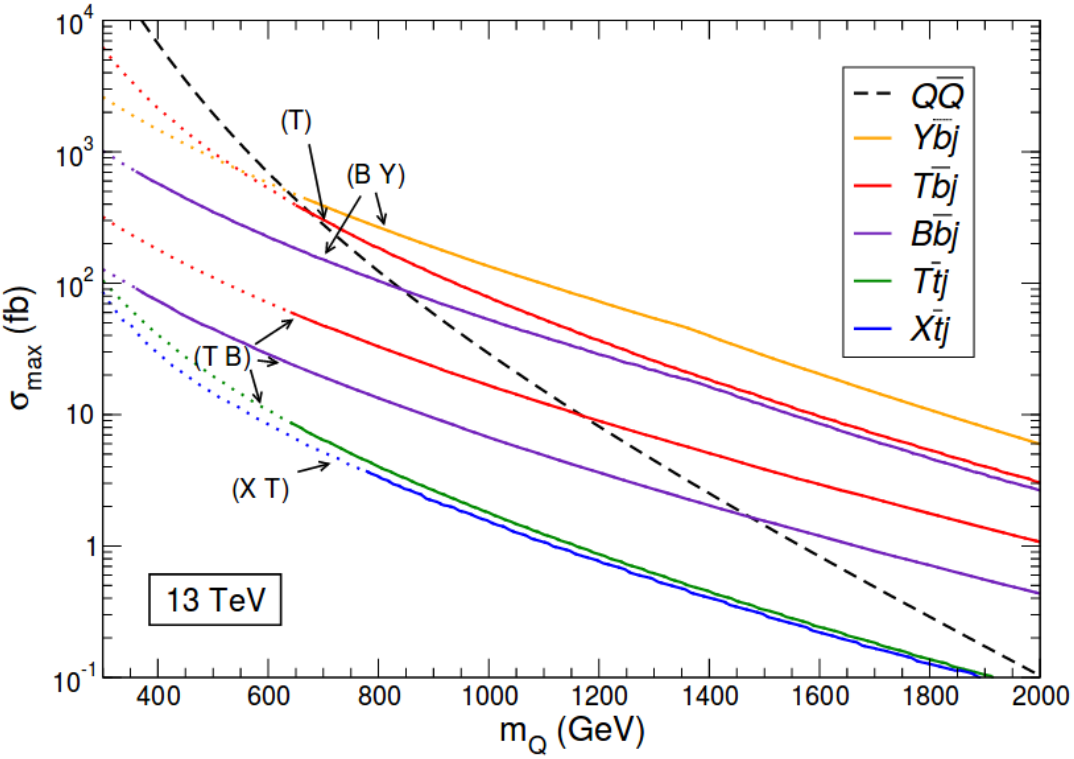


Figure 3.4: Production cross section for heavy quarks in pp collisions at $\sqrt{s} = 8$ TeV as a function of their mass, for pair production and for single production in different channels. The black dashed line represents VLQ pair production, while the colored lines represent the singly produced VLQs. The dashed coloured lines correspond to the values of cross section excluded by previous studies.

3.3.2 VLQ decay channels

Vector-like quarks decay through electroweak interactions into SM particles and generally the allowed decays are:

$$T \to W^+b, Zt, Ht$$

$$B \to W^-t, Zb, Hb$$

$$X \to W^+t$$

$$Y \to W^-b.$$

In the doublets and triplets the decays to other VLQ are usually suppressed, given the small mass difference, and the only allowed decays are into vector bosons and the Higgs boson plus a t or b quark.
They also can decay via flavour-changing neutral currents since they break the GIM mechanism [31]. The new quarks are expected to couple to the third generation, but couplings to lighter generations, although not favoured, are not excluded and can lead to flavor-changing neutral top interactions. For the isospin singlets T and B all three decays are possible, the branching ratios for the three channels depend on the mass of VLQ and are not inter-dependent as shown in the following relation:

$$Br\left(Q \to Wq'\right) + Br\left(Q \to Zq\right) + Br\left(Q \to Hq\right) = 1$$

with $(Q, q, q') = (T, t, b), (B, b, t)$.
The scenario is different for the isospin doublets: in the case of a (T, B) doublet, the mixing with t and b quarks induces a splitting of VLQ masses and it can be deduced that $m_T \geq m_X$, $m_B \geq m_Y$, and that T quark can be lighter or heavier than B and decays in other VLQ are suppressed since the small difference between their masses. The decays strongly depend on the mixing factors of the extended CKM matrix V_{Tb} and V_{tB}. If $V_{Tb} \sim V_{tB}$ then the T and B quarks have the same decays as the corresponding singlets but different angular distributions since only the right-handed component of (T, B) couples to the SM quarks. In the most natural case, induced by constraints on the b quark mixing and by mass hierarchy, i.e. $m_t \gg m_b$, where $V_{Tb} \ll V_{tB}$, then the mixing of the heavy quarks with the SM top quark is much stronger, and the $T \to Wb$ decay is suppressed, as well as $B \to Hb$ and $B \to Zb$.
Tables 3.4, 3.5 and 3.6 summarize the possible decays of vector-like quarks. The branching ratios of the vector-like quarks depend on the model but also on the heavy quark mass. In Figure 3.5 are shown the branching ratios of the decays of T and B considering VLQ coming from the singlet or doublet of $SU(2)_I$.

Singlets	**Decay modes**
X	W^+t
T	W^+b, Ht, Zt
B	W^-t, Hb, Zb
Y	W^-b

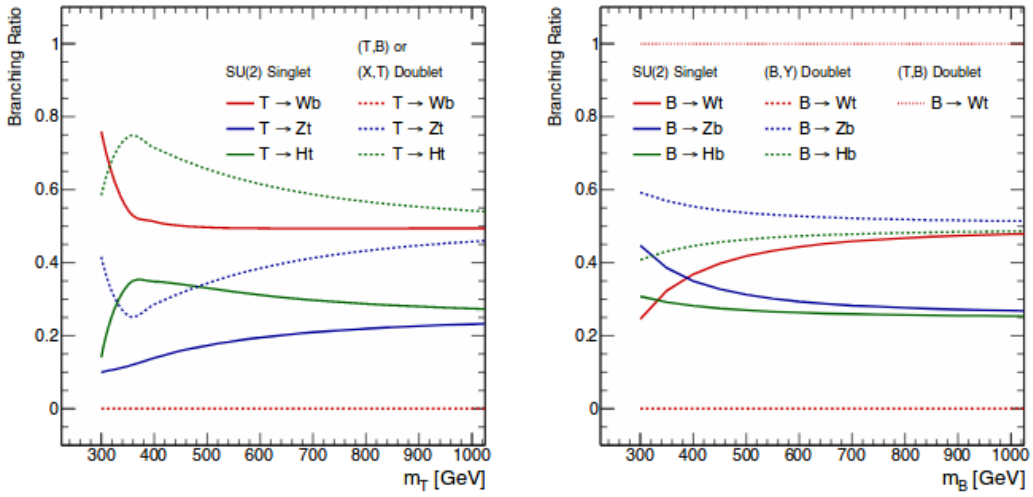
Table 3.4: Allowed decay modes for Vector-like singlets.

Doublets	**Decay modes**
$\begin{pmatrix} X \\ T \end{pmatrix}$	W^+t Ht, Zt
$\begin{pmatrix} T \\ B \end{pmatrix}$	Ht, Zt W^-t
$\begin{pmatrix} B \\ Y \end{pmatrix}$	Hb, Zb W^-b

Table 3.5: Allowed decay modes for Vector-like doublets.

Triplets	**Decay modes**
$\begin{pmatrix} X \\ T \\ B \end{pmatrix}$	W^+t W^+b, Ht, Zt Hb, Zb
$\begin{pmatrix} T \\ B \\ Y \end{pmatrix}$	Ht, Zt W^-t, Hb, Zb W^-b

Table 3.6: Allowed decay modes for Vector-like triplets.



(a) (b)

Figure 3.5: Branching ratio of vector-like top (a) and bottom (b) partners as a function of the heavy quark mass m_T and m_B respectively for isosinglets and isodoublets.

3.4 Searches on VLQ at hadronic colliders

Various searches of new heavy states have been undertaken both at Tevatron and at the LHC, though no evidence for the existence of other quarks, beside those of the SM, has been obtained. As their masses are expected to be in the vicinity of the TeV scale, these particles are accessible at the Large Hadron Collider and their search is therefore of prime importance. VLQs are a very promising field for searches of new physics since it has a rich phenomenology and many possible channels to explore.

3.4.1 Tevatron searches on VLQ

The main studies on VLQ at the Tevatron Run II were made with a $\sqrt{s} = 1.96$ TeV by the two experiments CDF and D0. At the first was studied the B Vector-like quark and only when it was reach a luminosity greater than 5 fb^{-1} also the T. The D0 Collaboration performed a search for single production of VLQ at 5.4 fb^{-1}. They considered a model where there were two degenerate VL doublets and searches for final states with either W or Z boson and two jets, one coming from the VLQ decay and the other produced in association with VLQ [37]. They also required a leptonic decay of the gauge boson and events with exactly one lepton or two, depending from the W or Z cases. Moreover they also required a number of jet selected equal or grater than two. Results are given

for different values of couplings parameters and branching ratios:

$$\left.\begin{array}{l} m_B > 693 \text{ GeV } at\ 95\% C.L.\ \ BR(B \to Wq) = 100\% \\ m_T > 551 \text{ GeV } at\ 95\% C.L.\ \ BR(B \to Zq) = 100\% \end{array}\right\} \text{no coupling with down quark}$$

$$\left.\begin{array}{l} m_B > 430 \text{ GeV } at\ 95\% C.L.\ \ BR(B \to Zq) = 100\% \\ m_T > 403 \text{ GeV } at\ 95\% C.L.\ \ BR(T \to Wq) = 100\% \end{array}\right\} \text{no coupling with up quark}$$

The CDF Collaboration on the other hand arrived at integrated luminosity of 5.7 fb^{-1} and managed to perform two analyses:

- Search for pair production of heavy particles T decaying to tX with X is an invisible dark matter particle using the full hadronic channel, Ref [35]. The request was of a number of jet among five and ten and missing transverse energy. Bounds are given for the combination of T and X masses, excluding the presence of T with $m_T \leq 400$ GeV for $m_X \leq 70$ GeV.

- Search for single production of heavy quarks expected to decay 100% in Wq, where q is a SM quark of the first generation, Ref [36]. The signal event searched have the topology $W + 2j$, where the W is required to decay leptonically. In this case the bounds on the cross section and couplings of the heavy quarks with SM quarks are given for different masses of the heavy quarks, ranging from 300 GeV to 600 GeV.

3.4.2 LHC searches on VLQ

At the LHC Collider it is possible to study the signatures of single and pair production of VLQ since they are accessible at current $\sqrt{s}$ and luminosity. Both ATLAS and CMS performed numerous dedicated searches for 3rd family quark partners at the LHC. The searches mostly focus on QCD pair production of charge 5/3, 2/3, and -1/3 partners and their decay into 3rd family quarks and W/Z/H bosons. Searches are performed in various final states, including all-hadronic searches, single- or multi-lepton final state searches with or without transverse missing energy, and – for charge 5/3 partner searches – also same-sign dilepton final states.
In the case of the searches for the pair production of charge 5/3 VLQs have been studied the decays into $WtWt$, the analysis covers the single lepton and dilepton same-sign channels [38, 39]. For the single lepton final state have been identified 16 channel using the lepton flavor, the number of b-tagged, W-tagged and top tagged jets while in the case of the dilepton channel 3 different search regions have been identified based on the lepton flavor. For $T\bar{T}$ pair production have been studied the possible decays:

- with at least one $T \to Ht$ and $H \to bb$, or $T \to tZ$ and $Z \to \nu\nu$. The final states searched for were with one lepton and/or missing transverse energy [40].

- Two Wb pairs and with one W to leptons and one to quarks, the analysis was sensitive to the other two decay modes as well as to Vector-like B quarks, assuming $Br(T \to Wb) = 100\%$ [41].
- $Zt + X$ with exactly one charged lepton and $Z \to \nu\nu$, the upper limits on T mass are 0.85 (1.05)TeV for the weak-isospin singlet (doublet) model; 1.16 TeV for the pure Zt decay mode [42].
- Into $bWbW \to bl\nu\bar{b}qq'$, where a kinematic fit has been used to fully reconstruct the mass of the T quark, assuming a branching fraction $Br(T \to bW) = 100\%$ [43]. Is required exactly one charged lepton (e or μ), at least 4 jets and a boosted W-tagging is used for categorization

In the case of the production of $B\bar{B}$ the decays into $tW/bZ/bH$ have been studied and follow the same analysis strategy of the $T\bar{T}$, while for the single production of VLQ the following channels have been investigated :

- production of a T quark decaying to tZ with $Z \to ll$ and $t \to hadrons$, it has been also studied the production of $Z' \to Tt$ [44];
- production of $B \to bH$ with $H \to bb$ and at least one Higge-tagged wide jet [45].

In the Table 3.7 are reported the most recent results obtained from both the collaborations ATLAS and CMS and in the Figures 3.6 and 3.7 can be seen the different results for the single and the pair production of VLQs.

Model	**Observed Exclusion**		**Expected Exclusion**	
	Left Handed	*Right Handed*	*Left Handed*	*Right Handed*
$T \to tZ$	1.2 TeV	-	1.25TeV	-
$B \to Hb$	-	-	-	-
$X_{5/3}X_{5/3} \to tWtW$	1.30TeV	1.28TeV	1.33TeV	1.30 TeV
$T\bar{T} \to bW/tZ/tH$	1.20 TeV	1.28TeV	1.16TeV	1.24TeV
$B\bar{B} \to tW/bZ/bH$	1.17TeV	0.94TeV	1.13TeV	0.92TeV

Table 3.7: VLQ limit summary table.

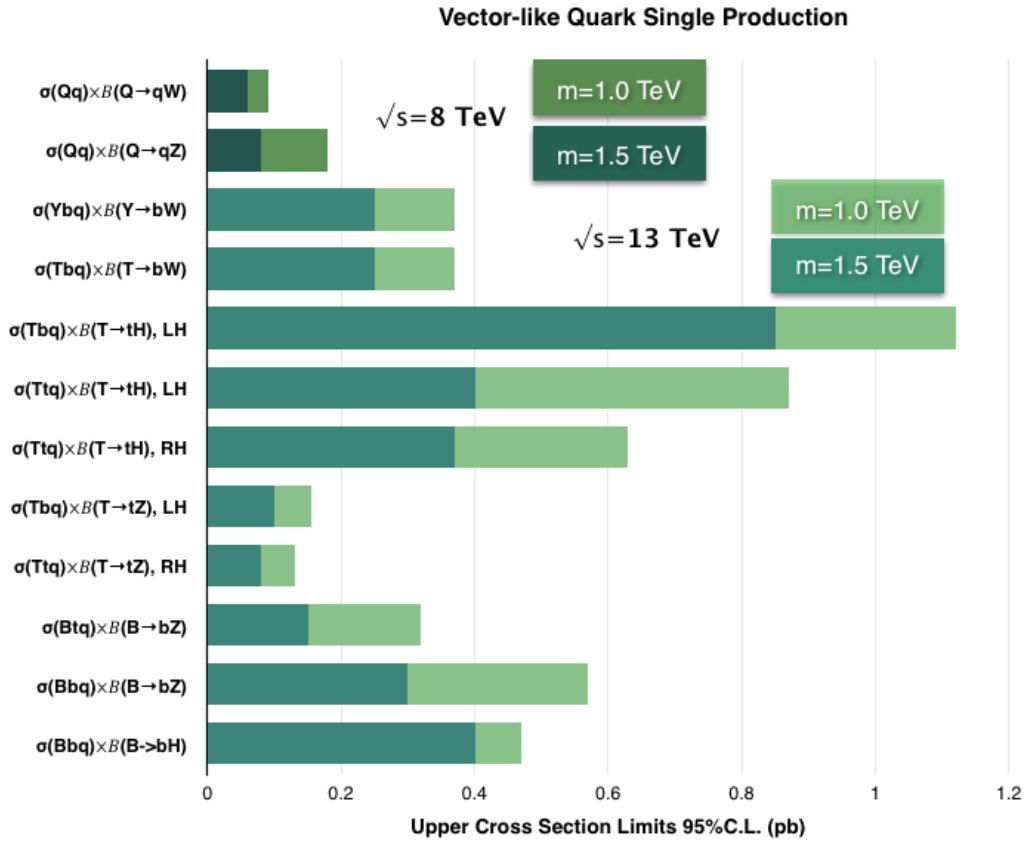


Figure 3.6: Summary of the results of the VLQs single production.

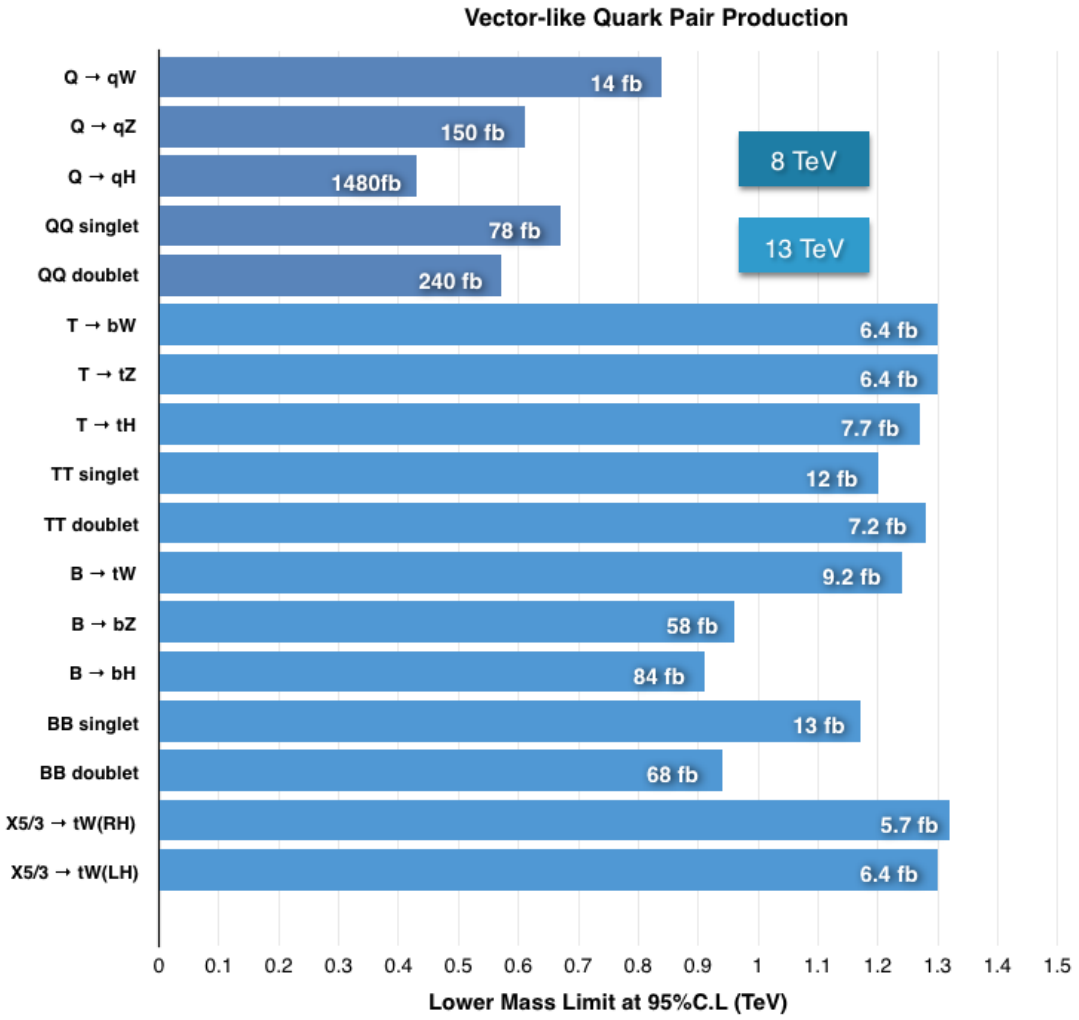


Figure 3.7: Summary of the result of the VLQs pair production.

Chapter 4

Physics objects selection and reconstruction

A search for the single production of a heavy vector-like quark T of charge + 2e/3 is presented in this thesis. The search is performed in the $T \rightarrow$ tZ decay channel, with a top quark decaying hadronically and a Z boson decaying to a pair of neutrinos, using 35.86 fb^{-1} of pp collision data delivered by the LHC at a centre-of-mass energy of $\sqrt{s} = 13$ TeV and collected by the CMS experiment in 2016. An example of Feynman diagram for the single production of a T quark is shown in Figure 4.1. With the increase in centre-of-

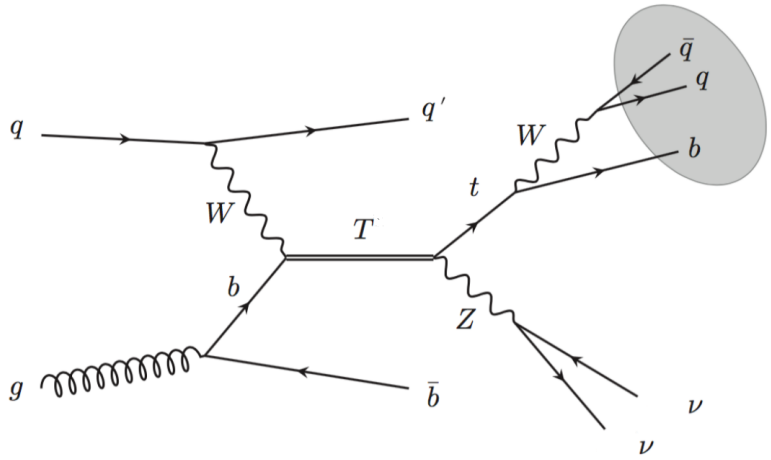


Figure 4.1: Feynman graph for the signal.

mass energy, the Z is produced more boosted with respect to 8 TeV collisions. This implies that neutrinos, produced as collimates, resulting in signals with large missing energy. The choice of this decay channel over the cleaner di-lepton channels is justified by a factor ~ 2 of the branching ratio for $Z \rightarrow \nu\nu$ than the one for $Z \rightarrow ll$ and it has been investigated in [46]:

- $Z \to \nu\nu$: $(20.00 \pm 0.06)\,\%$
- $Z \to ll$: $(10.099 \pm 0.019)\,\%$

The search for a vector-like quark T is performed looking for missing transverse energy, jets compatible with the signature of a top quark or a W boson, as well as b-jets and forward jets. The requirements applied in order to identify these objects are listed in this chapter and in Table 4.1.

	jets	bjet	fwd jet	top	W
p_{T} >(GeV)	30	39	30	400	200
$\lvert\eta\rvert$	<4	<2.4	[2.4,4]		
b-tag		>0.8484			
Soft Drop mass				[105,220]	
Pruned mass					[65,105]
τ_{32}				<0.81	
τ_{21}					<0.35

Table 4.1: List of requirements applied to identify the objects necessary for the analysis.

4.1 Physics objects selection

The reconstruction and the identification of all stable particles produced through the pp collisions, such as electrons, muons, photons, charged hadrons and neutral hadrons is done through the Particle Flow (PF) algorithm [47]. The PF algorithm exploits a set of information coming from all CMS sub-detectors towards an optimal identification and determination of the particles, their direction, energy and type. This list of individual particles is then used, as if it came from a MC event generator, to:

- build jets from which the quark and gluon energies and directions are inferred;
- determine the missing transverse energy MET which gives an estimate of the direction and energy of the neutrinos and other invisible particles;
- quantify charged lepton isolation with respect to other particles;
- tag jets coming from the hadronisation of b-quarks or top-quarks, and, to a minor extent, c-quarks or gluons.

Thanks to an interactive tracking algorithm all the reconstructed elements are linked to each other by identifying blocks of elements that are compatible, e.g. the track of a charged particle is linked with calorimeter deposits if the extrapolated position of the

track is within the cluster boundaries. PF candidates are reconstructed in the following order:

- *muons*: global muon, reconstructed from the hits in the tracker and the track in the muon system, gives rise to a PF muon. If such particle is identified, the corresponding tracks are removed from the block;
- *electrons*: a particle is identified as an electron if the link between a charged-particle track and one or more ECAL clusters are compatible. The tracks are then removed from the list;
- *charged hadrons*: the remaining tracks are used to reconstruct charged hadrons. The tracks can be linked to ECAL and HCAL clusters while the momentum is directly taken from the tracker and corrected taking into account information from the calorimeters;
- *photons and neutral hadrons*: clusters in the ECAL that are not compatible with charged tracks give rise to PF photons; in a similar way, neutral hadrons are identified if there are unmatched clusters in HCAL;

When the list of PF candidates is complete, PF jets can be reconstructed using jet reconstruction algorithms and the Missing Transverse Energy of the event can be estimated.

4.2 Primary vertices

Vertices, reconstructed by a best fit to the intersection of tracks reconstructed in the tracking system, are defined as good if they satisfy the following requirements:

- more than 4 degrees of freedom in the fit
- less than 2 cm away in the x-y plane from the interaction point
- less than 24 cm away in the z direction from the interaction point

Those requirements guarantee that the reconstructed vertices are in the luminous region. In the thesis only events where at least on good primary vertex is found are selected. Among the good vertices, it is defined as primary vertex the one corresponding to the highest value of the sum of the squared transverse momentum of all the tracks associated with it.

4.3 Leptons

The final state addressed by this analysis is characterised by quarks and neutrinos. The major source of background arises from $t\bar{t}$ events and electroweak processes. The

contribution from $t\bar{t}$ events arises mostly from decays where a lepton is not identified, therefore a lepton veto is applied. This veto does additionally avoid overlap in the signal dilepton final state. The identification and isolation criteria used are described in the following sections.

4.3.1 Electrons

Electron candidates are reconstructed from a collection of electromagnetic clusters with matched pixel tracks. Electron tracks are fitted using a Gaussian Sum Filter (GSF) algorithm along its trajectory taking into account the possible emission of bremsstrahlung photons in the silicon tracker. All electrons are required to have p_{T} > 30 GeV and $|\eta|$ <2.5. Different requirements are applied to electrons reconstructed in the calorimeter barrel and endcaps. Electron candidates are selected according to the "veto" selection criteria defined in CMS according to identification criteria with high efficiency but low purity [48]. The isolation variable, $I^{\mathrm{e}}_{\mathrm{rel}}$, in the electrons case, is defined as:

$$I^{\mathrm{e}}_{\mathrm{rel}} = \frac{I^{\mathrm{ch.h}} + \max[(I^{\gamma} + I^{\mathrm{n.h}} - \rho \times A), 0]}{p_{\mathrm{T}}}, \tag{4.1}$$

where ρ is the average energy density not clustered in jets, measured event-by-event, by the cone area A. The sums are computed in a cone of $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.3$ around the electron direction.

4.3.2 Muons

Muons are, at first, reconstructed in the tracking system using a technique based on Kalman filter algorithm then in the muon chamber, combining informations from the drift tubes(DT), cathode strip chambers(CSC), and resistive plate chambers(RPC). The reconstructed track in the muon chamber is named *Stand-alone muon*. Starting from these these objects, two approaches can be used:

- The *Global Muon reconstruction* that for each stand-alone track a matching track is found in order to reject the background of muons coming from hadronisation and from the particles that interact around the beam pipe. This approach is especially useful at large transverse momenta because it improves the momentum resolution.

- The *Tracker Muon reconstruction* that consider a track in the tracker as a muon candidate with $p_{\mathrm{T}} > 0.5$ GeV and a total momentum $p > 2.5$ GeV. If some hits deposited in the muon chambers can match the tracker track, also considering the energy losses due to the Coulomb scattering, the corresponding track is identified as a Tracker Muon. This method is efficient at low momentum ($p < 5$ GeV), because it requires only a hit in one of the components of the muon chambers.

In the cases when muons are identified both as Global and Tracker muons, in the case they share the same track in the tracking system, they are merged together. Muons with $p_{\mathrm{T}} > 30$GeV and $|\eta| < 2.4$ and passing the "loose" identification criteria defined in CMS [49] are vetoed. The I^{μ}_{rel} variable in the muon case, is defined as:

$$I^{\mu}_{\mathrm{rel}} = \frac{I^{\mathrm{ch.h}} + \max[(I^{\gamma} + I^{\mathrm{n.h}} - 0.5 \times I^{\mathrm{PU}}), 0]}{p_{\mathrm{T}}}, \tag{4.2}$$

where $I^{\mathrm{ch.h}}$, I^{γ}, $I^{\mathrm{n.h}}$, and I^{PU} are respectively, the scalar p_{T} sums of the charged hadrons, photons, neutral hadrons, and charged hadrons associated with pileup vertices. The sums are computed in a cone of $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2} = 0.4$ around the muon direction.

Tight Muons

An additional category is defined in order to study the trigger efficiency with "tight" selection requirements. Any event with a muon, within $|\eta| < 2.4$, having $p_{\mathrm{T}} > 60$ GeV/c is considered without isolation requirement.

4.4 Jets

The majority of the LHC events contains partons in the final state that cannot be directly observed. Quarks and gluons manifest themselves trough hadronization to stable particles which can be detected in the tracking chambers and calorimeters. The interaction between constituent partons and the showering into stable particles is well described from perturbative theory and hadronization model. In Figure 4.2 the evolution can be observed of a jet from hard interaction to observable energy deposits. The jet reconstruction algorithm employed at CMS is based on clustering all the PF candidates reconstructed. The so identified particles are clustered by applying the anti-k_{T} algorithm [50]. The Charged Hadron Subraction (CHS) is a correction applied on the group of hardest particle in order to remove the energetic contribution coming from the pile-up interaction before clustering. The distance between the particle i and the particle j, d_{ij}, and d_{iB}, the one between the particle i and the beam, can be written as follows:

$$d_{ij} = min\left(\frac{1}{p^2_{T,i}}, \frac{1}{p^2_{T,j}}\right)\frac{(\Delta R)^2_{ij}}{R^2}$$

$$d_{iB} = \frac{1}{p^2_{T,i}}$$

where:

$$(\Delta R)^2_{ij} = (\eta_i - \eta_j)^2 + (\phi_i\phi_j)^2$$

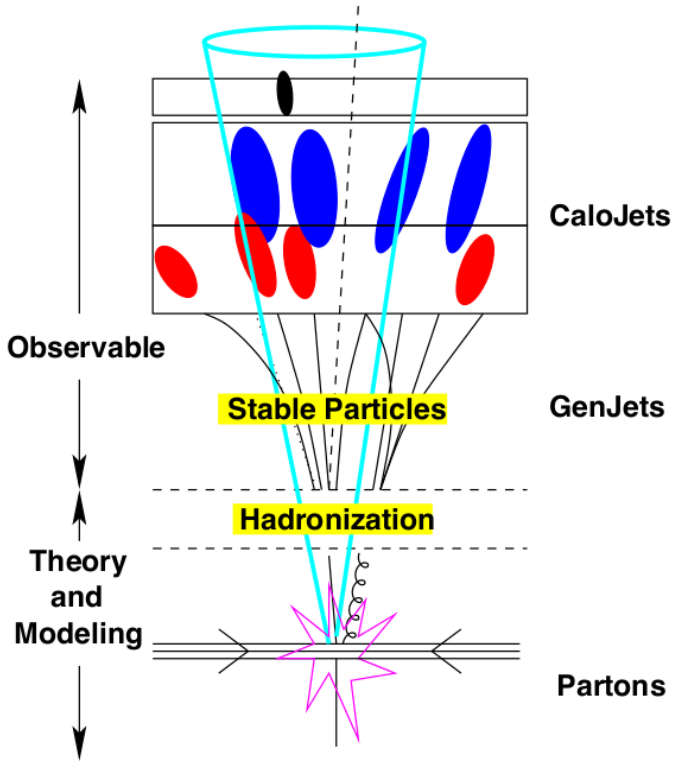


Figure 4.2: Pictorical view of the evolution of a jet.

with R radius parameter used to define the portion of solid angle covered by the jet, $\eta_{i(j)}$ the pseudorapidity of the particle $i(j)$, and $\phi_{i(j)}$ the angle in the transverse plane of the particle $i(j)$. Jets with cone parameter of $R = 0.4$ are named *narrow jets* (AK4), while jets clustered with a cone parameter of $R = 0.8$ are defined as *fat jets* (AK8). The latter are considered for the selection of boosted top quark candidates or W boosted candidates while the former are used to select a forward light-flavour jet produced in association to T quark production. Jets in both collections are required to pass a set of identification criteria ("loose" working point) according to the guidelines of the dedicated group that works on Jets and MET. AK8 jets are selected with p_{T} above 300 GeV while AK4 jets are selected with p_{T} above 30 GeV. There are some correction to be taken in account calculating the energy of the jets at various level and are known as Jet Energy Corrections (JEC). These correction have been introduced to reproduce the energy at particle level, the jet energy is scaled by a factor that describes the detector response depending on the transverse energy and the pseudorapidity of the jet [53]. The corrections pass trough the following steps:

- removal of the energy due to the contribution of pileup events;
- correction of the response of jets in order to achieve a flat distribution in η;
- correction of the jet response for p_{T}.

The correction factor Jet Energy Scale (JES) is calibrated on the η, p_{T}, energy density, and area of the jet; the Jet Energy Resolution (JER) for the simulated jets is degraded to reproduce the resolution observed in data.

4.4.1 Top tagging

Highly boosted top quark candidates decaying hadronically are reconstructed from AK8 jets and identified according to the jet top tagging working point corresponding to 3% of mistag rate and based on CHS inputs. The mass of the AK8 jet is required to be consistent with the top quark mass, so the jet mass computed using the soft drop algorithm [54, 55] must satisfy the condition: $105 < M_{softdrop} < 220$ GeV. To increase the efficiency of detecting a top originated jet, the fat-jet has to have a 3 component structure to mimic the number of the top quark decay. The $N-subjettines$ variable τ_i estimates the compatibility of a jet with the hypotesis that it is composed of i constituents and it is defined as: :

$$\tau_N = \frac{\sum_k p_{T,k} min\left(\Delta R_{1,k}, \Delta R_{2,k}, ..., \Delta R_{N,k}\right)}{\sum_k p_{T,k} R_0} \tag{4.3}$$

where k indicates all the constituents of the jet, $\Delta R_{j,k}$ is the distance of the candidate subjet j and the particle k and R_0 is the jet radius. In the case of having N components or fewer the value of τ_i goes to zero. If τ_N larger than zero, it indicates that part of the jet energy is spread at a larger angle, and that it is more likely that the jet has at least N+1 subjets. The ratio $\tau_{32} = \tau_3/\tau_2$ is used to discriminate between top quark jets and jets from quarks and gluons and is required to be less than 0.81. The flavour of the top quark decay products is exploited by selecting only AK8 jets containing one b-tagged subjet. These selection constraints correspond to a top mistag rate of 3%.

4.4.2 W tagging

The AK8 jets may also be tagged as coming from a W boson decaying to $q'\bar{q}$ (denoted "W jets"). For the W jets a p_{T} grater than 200 GeV is required, and a pruning algorithm [56] is applied. The mass of the jet, after the pruning is performed, is used as a discriminant to select W bosons and reject quark and gluon jets. The discrimination between W jets and jets from quarks and gluons is further improved by requiring the $N-subjettiness$ ratio τ_{21} to be less than 0.35, where $\tau_{21} = \tau_2/\tau_1$ [57], and the mass of the pruned AK8 jet to be within the range 65–105 GeV.

4.4.3 b-Tagging

AK4 jets may be tagged as arising from a b quark ("b jets"), indeed a b-jet is expected in the final state originated from the gluon splitting and acting as a spectator. Several

algorithms for identification of jets originating from b quarks are available in CMS, one of these is the Combined Secondary Vertex algorithm [58, 59]. The tracks produced by long lived particle decays such as b-hadrons makes possible for the hadron to travel a considerable distance from the primary vertex before decaying. The variable used to define the distance between the two vertices is the Impact Parameter (IP), Figure 4.3, that is a Lorentz invariant, so it is invariant with respect to changes of the long lived particle kinetic energy, and for the B-hadrons this corresponds to $c\tau \sim 450\ \mu m$. In CMS the IP can be measured with precision between 30 μm and hundreds μm. Also the uncertainty can be of the same order of magnitude as the IP, so the IP significance $IP/\sigma(IP)$ is a better observable and takes into account also the resolution. It can be positive or negative, depending on the signs of the scalar product of the IP-vector and the jet direction. It is expected to be symmetric around 0 for decays with short life and mostly positive for particles with longer lifetime. The AK4 considered stemming from a

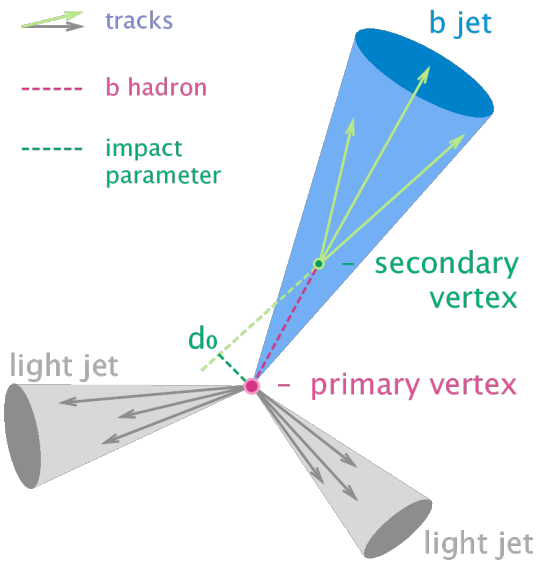


Figure 4.3: Impact Parameter.

b quark if they have:

- pseudorapidity in a central region: $|\eta| < 2.4$;
- discriminant value of the Combined Secondary Vertex greater than 0.8484 (that CMS Collaboration define as a medium working point).

Moreover since the top quark decays to a bottom quark and a W boson, must be satisfied also the requirement that b-jets selected not overlap with the top tagged jets to avoid selecting b-jets from top decay as a narrow b-jet; the separation distance between b-tagged jet and top R(bjet,top) is required to be larger than 1.2.

4.4.4 Forward jets

An interesting feature of the direct production of a single vector-like T quark is the presence of an additional jet that is produced in the forward direction. Forward jets are reconstructed as AK4 jets using the same selections and corrections as defined above, but have 2.4< $|\eta|$ <5.0 and p_{T} >30 GeV.

4.5 Missing transverse energy

At the LHC, and in general at any hadron collider, the energy imbalance is computed in the transverse view only. Indeed only the longitudinal momentum of the whole proton can be measured, the one of the individual colliding partons is not known, and with large precision can be assumed to be zero their transverse momentum. The particle flow algorithm reconstructs the missing transverse energy, MET, from the vectorial sum of all the candidates transverse momentum. The MET is one of the most important variables for discriminating leptonic decays of W bosons and top quarks from background events which do not contain neutrinos. It is also a crucial variable to search for Beyond Standard Model physics, since many models predicts particles that would not leave any trace in the detector. The presence of particles that do not interact in the detector can be measured with MET since they create an energy imbalance. In this thesis the MET is fundamental to identify events where the Z boson produced in association with a top quark decays into neutrinos. The MET is also one of the most complex variables to reconstruct, since it is very sensitive to detector malfunctions and particles crossing poorly-instrumented regions of the detector. To obtain the best definition of MET the energy corrections applied to the jets(JEC) must be propagated to the MET, as well as corrections to the scale of the particles that are not clustered into jets. Filters to the MET are applied that remove pathological events from data, referred in this thesis as MET $filters$.

4.6 Background description

The most important backgrounds mimicking the final state under study, sketched in Figure 4.4, are:

- $\mathrm{t\bar{t}}$: processes where a $\mathrm{t\bar{t}}$ quark pair is produced are the dominant background. In particular events where one t quark decays semi-leptonically, namely through the chain $\mathrm{t} \to Wb \to l^{+}\nu b$, and the other top quark decays hadronically, namely through the chain $\mathrm{t} \to Wb \to b\mathrm{q\bar{q}}$, have several features in common with signal events. Indeed the top quark signal behaviour can be mimicked by the top that decays hadronically while the Z boson plus b jet can be mimicked by the other top quark that decays semi-leptonically in the case in wich the charged lepton has not

been reconstructed or it was not in the detector acceptance region. Such events are also called semi-leptonic $t\bar{t}$ events in jargon.

- W + jets: in the same way as $t\bar{t}$, the lepton decayed from W is not reconstructed, and if the gluon is emitted close enough to the couple of b-jets it might be misidentified as a top jets.

- Z+jets: similarly the couple of quarks and a gluon emitted in the next-to-leading order in Z+Jets processes can be mistaken for a T quark decay.

- QCD: produces a high number of jets that can also simulate the signature of the process where large MET arises from mis-reconstructed jets.

- Single top and VV: single top quark processes and diboson processes like WW, WZ, and ZZ processes are also minor sources of reducible background which can reproduce in some cases the signal topology. However they can be reduced through event selection and appear as a negligible fraction in the final region.

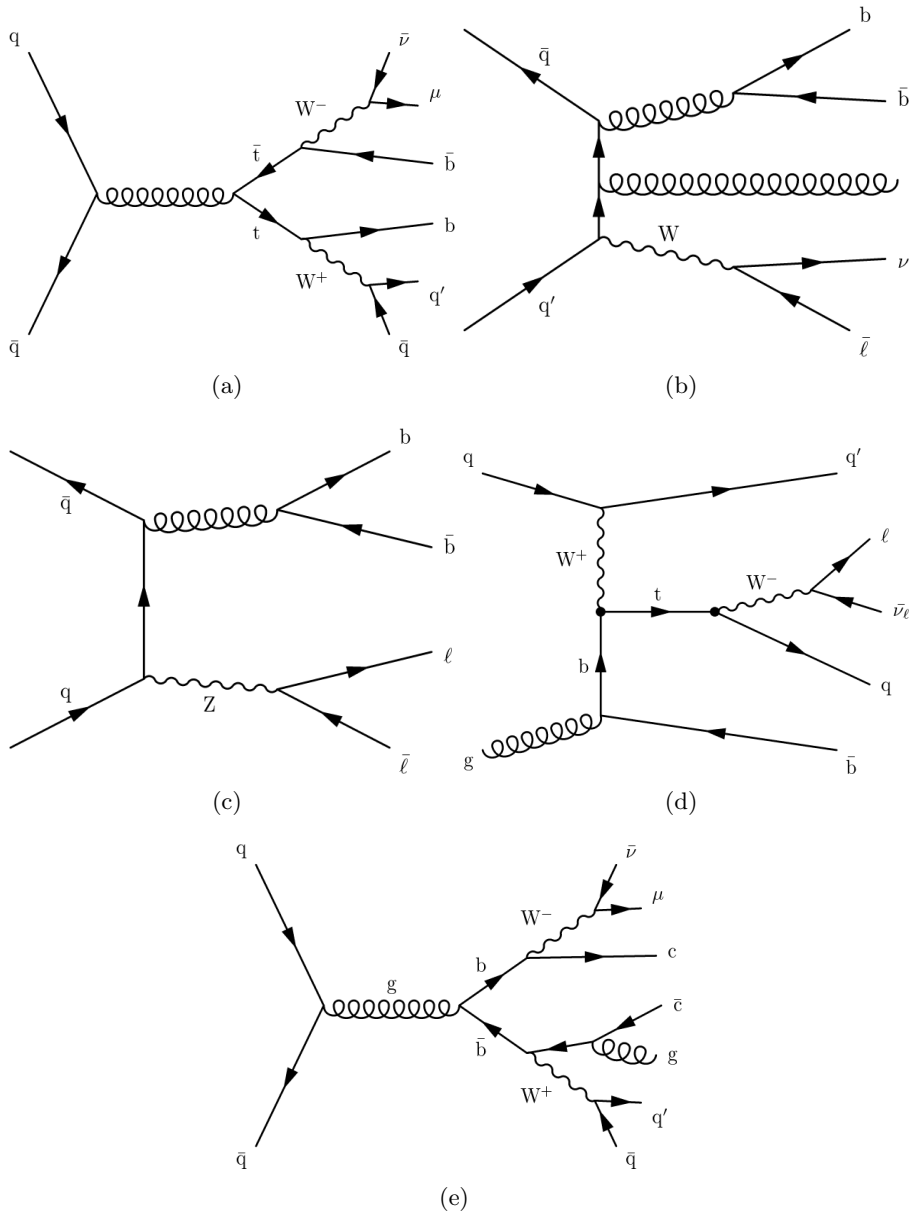


Figure 4.4: Feyman diagrams of background processes: $t\bar{t}$ (a),W + jets (b), Z+jets(c), Single top(d) and QCD(e).

Chapter 5

Analysis strategy

This Chapter describes the analysis strategy to extract the T signal, by constraining the SM background in a signal enriched region directly from data. First an event selection optimized for T signal is described: after a preselection step with basic cuts to define a topology close to the one of signal events, several categories are studied by dividing events by jet content, number of b-jets and forward jets, and by top quark reconstruction method. Control regions have been identified with different number of leptons in order to extract background from data with a *data-driven* method. A simultaneous maximum likelihood fit is performed in all signal regions.

5.1 Data and simulated samples

In this work the pp collision data recorded by CMS in 2016 and delivered by the LHC at a centre-of-mass energy $\sqrt{s} = 13$TeV are used. In Table 5.1 the list of data sets employed in the analysis is reported; the label MET refers to the set of triggers used to select the data set, the label Run2016B through H3 refers to the data taking period, the label 3Feb2017 refers to the date in which the data sets have been processed. The final MINIAOD refers to the file format and content according to CMS standard definitions. The data set used corresponds to an integrated luminosity of 35.86 fb^{-1}. Signal and the main background processes are taken from simulation. Three steps can be identified in the simulation of events at the LHC:

- the hard scattering process,
- the parton-shower of quarks that hadronize,
- the interaction of particles with the detector.

The simulation of such levels it is possible thanks to Monte-Carlo event generators that manage to:

Data set	Integrated luminosity [pb^{-1}]
/MET/Run2016B-3Feb2017-v3/MINIAOD	5788
/MET/Run2016C-3Feb2017-v1/MINIAOD	2573
/MET/Run2016D-3Feb2017-v1/MINIAOD	4248
/MET/Run2016E-3Feb2017-v1/MINIAOD	4007
/MET/Run2016F-3Feb2017-v1/MINIAOD	3101
/MET/Run2016G-3Feb2017-v1/MINIAOD	7540
/MET/Run2016H-3Feb2017-v2/MINIAOD	8391
/MET/Run2016H-3Feb2017-v3/MINIAOD	225

Table 5.1: List of data sets of pp collision data produced at $\sqrt{s} = 13$ TeV and collected by CMS in 2016, employed in the thesis. Data collected correspond to an integrated luminosity of 35.86 fb^{-1}.

- generate matrix elements either at leading order corrections (LO) or at next-to-leading order corrections (NLO), with programs like Madgraph5_aMC@NLO [61] or*POWHEG 2.0* [64];
- generate and simulate the hadronization of outgoing particles produced in proton-proton collisions, in order to reproduce accurately the event properties of a wide range of processes, such as Pythia [62] or Herwig [63];
- simulate particle interaction with CMS sub-detectors, such as GEANT 4 [65].

Signal event samples have been generated using the leading order Monte Carlo event generator Madgraph matched to the Pythia generator for the parton-shower simulation, with different benchmark models based on the hypothesis made on the T quark mass, ranging from 700 up to 1800 GeV in steps of 100 GeV, and on the width of the resonance, ranging from a narrow width approximation to 30%, of the resonance mass, in steps of 10%. The generated signal samples for the T process and the corresponding LO cross sections are reported in Table 5.2. Simulated Monte Carlo samples for background processes used in this thesis are listed in Table 5.3 with the corresponding theory cross sections [69]. The $t\bar{t}$ pair production process is generated with *POWHEG 2.0* [64, 66, 67] and its cross section is calculated at the next-to-next-to leading order (NNLO) in perturbative QCD. Two additional simulated samples are generated to increase the $t\bar{t}$ process statistics in the signal region, by generating events where the mass of the $t\bar{t}$ pair is greater 700 GeV, and the corresponding cross sections are evaluated at NLO. To simulate Single Top events is used a match of *POWHEG 2.0* and Pythia as well. Multijet QCD production, W + jets, and Z+Jets samples are generated with the MADGRAPH 5 [61] tree-level matrix-element generator matched to PYTHIA 8 [68] for the parton-shower simulation, their cross section are instead calculated at leading order (LO). Pythia is

used for the generation of vector boson pair, or Diboson, samples also for the matrix element component.

5.1.1 Pile-up reweighting

To take into account the presence of additional interactions with respect to the primary one, known as pile-up (PU), that create energy deposit in the tracker and in the calorimeters, CMS computes the number of primary vertices from the instantaneous luminosity. The pile-up distribution in simulation does not reproduce the one observed in data. Therefore a reweighting is needed, consisting in applying multiplicative factors to MC sample events to correct for the discrepancy with the distribution of primary vertices in the data.

Table 5.2: Simulated $Tb \rightarrow$ tZ samples employed in the analysis with their corresponding production cross sections.

Mass [GeV]	Cross section $[fb]$			
	Width 1%	Width 10%	Width 20%	Width 30%
700	78.04			
800	41.54		791.86	1153.83
900	23.34			
1000	13.62	131.82	251.47	360.22
1100	8.23			
1200	5.11	49.18	92.68	131.44
1300	3.25			
1400	2.12	20.25	37.94	53.49
1500	1.41			
1600	0.94	9.02	16.81	23.55
1700	0.64			
1800	0.44	4.25	7.91	11.04

5.2 Event selection

5.2.1 Trigger

Signal events are characterised by the presence of two highly energetic neutrinos that manifest as missing transverse energy as described in Section 4.5. Events are therefore required to satisfy trigger conditions based on high values of MET. The request consists of the logical "OR" between two conditions: the online reconstructed MET has to pass the threshold of either 110 GeV or 120 GeV. This combination of trigger conditions is henceforth referred to as the hadronic trigger.

Table 5.3: Standard-model background samples and their corresponding production cross sections. Many of the samples listed are split in order to increase the available statistics by generating more events in more extreme regions of the phase space. $t\bar{t}$ samples are divided considering different ranges of $t\bar{t}$ per mass: $0-700$ GeV, $700-1000$ GeV and $1000-Inf$ GeV, instead other samples such as QCD or W + jets and Z + jets are split taking in account different values of the total transverse momenta.

Sample	Cross section [pb]
$t\bar{t}$ (0-700)	831.76 *
$t\bar{t}$ (700-1000)	80.5
$t\bar{t}$ (1000-Inf)	21.3
QCD (HT100-200)	27990000
QCD (HT200-300)	1712000
QCD (HT300-500)	347700
QCD (HT500-700)	32100
QCD (HT700-1000)	6831
QCD (HT1000-1500)	1207
QCD (HT1500-2000)	119.9
QCD (HT2000-Inf)	25.24
W+Jets (lν,HT 600-800)	12.05×1.21 **
W+Jets (lν,HT 800-1200)	5.501×1.21 **
W+Jets (lν,HT 1200-2500)	1.329×1.21 **
W+Jets (lν, HT 2500-Inf)	0.03216×1.21 **
Z+Jets ($\nu\nu$, HT 100-200)	280.35×1.23 **
Z+Jets ($\nu\nu$, HT 200-400)	77.67×1.23 **
Z+Jets ($\nu\nu$, HT 400-600)	10.73×1.23 **
Z+Jets ($\nu\nu$, HT 600-800)	2.559×1.23 **
Z+Jets ($\nu\nu$, HT 800-1200)	1.1796×1.23 **
Z+Jets ($\nu\nu$, HT 1200-2500)	0.28833×1.23 **
Z+Jets ($\nu\nu$, HT 2500-Inf)	0.006945×1.23 **
ZZ	16.523
WZ	47.13
WW	116.7
Single top ($\bar{t}$, t-channel)	80.95
Single top (t, t-channel)	136.02
Single top (tW)	71.17
Single top (s-channel)	10.32

* This is an inclusive cross section, then it is correct to remove the 80.5 pb e 21.3 pb of the other two $t\bar{t}$ samples.

** The cross section is multiplied by the scale factor obtained from the ratio NLO / LO.

5.2.2 Pre-selection

The presence of neutrinos in the final state allows to have large values of the offline reconstructed MET, therefore only events with MET p_{T} greater than 200 GeV are selected. Loose leptons as identified in Section 4.3 are vetoed. Since the signal typically presents jets from the spectator quarks as well as fat jets coming from the boosted top quark, events are selected with at least one narrow jet with p_{T} >30 GeV and at least one fat jet with p_{T} >300 GeV. Moreover, to reduce the background events of low mass QCD, a further condition is applied on the fat jet by requiring its mass to be grater than 40 GeV. This set of requirements is referred to as *pre-selection* and summarized here:

- MET p_{T} >200 GeV;
- veto on loose muons;
- veto on loose electrons;
- at least one narrow jet with p_{T} >30 GeV;
- at least one fat jet with p_{T} >300 GeV and with soft drop mass or pruned mass (as defined in Section 4.4) >40 GeV.

5.2.3 Angular variable selection

The angular distance in the transverse plane between MET p_{T} and narrow jets p_{T} has been investigated, to further reduce the contribution from multijet QCD background events. This selection criterion is motivated by the fact that it is expected that multijet QCD events are expected to be characterized by a smaller distance of the jets from the missing energy, as most of the MET arise from neutrinos originating from hadronic decays of the QCD jet components. Therefore the missing energy typically arises from neutrinos travelling alongside the jets, rather than from ones originating from the hard scattering of proton constituents (so called prompt neutrinos), that are naturally produced back-to-back with the rest of the proton collision remnants in the transverse plane. In Figure 5.1 the minimum value between the angular distances of MET p_{T} and the narrow jets of the event up to 10 jets displayed shows a good discrimination power against multijet QCD events. Events have been selected if $min\Delta\phi(MET, jet) > 0.6$.

5.2.4 Forward jet selection and categories identification

The presence of a forward jet that results from the light flavour quark produced in association with the vector-like quark T is a feature that can be used to further separate the background events from signal events. Figure 5.2 shows the multiplicity of forward jets ($|\eta| > 2.4$) for several signal hypotheses and backgrounds. Two categories are

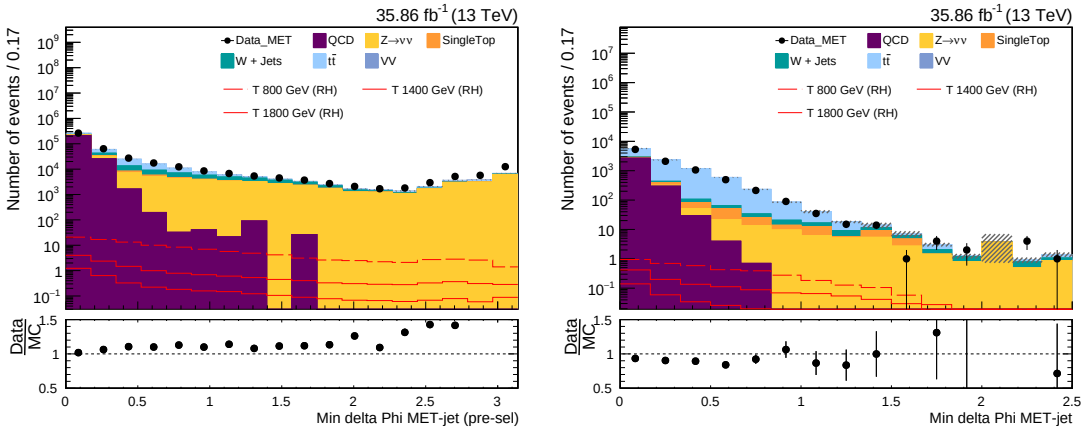


Figure 5.1: Distribution of the minimum distance $min\Delta\phi$ between MET p_T and narrow-jet p_T filled after pre-selection (left) and after pre-selection, b-tagging and top tagging requirements (right).

defined, one requiring at least one forward jet, enriched in signal, since the background is largely reduced, and one requiring 0 forward jets in final state. The latter category allows to improve the final sensitivity because it has higher statistics, with a lower purity with respect to the former category.

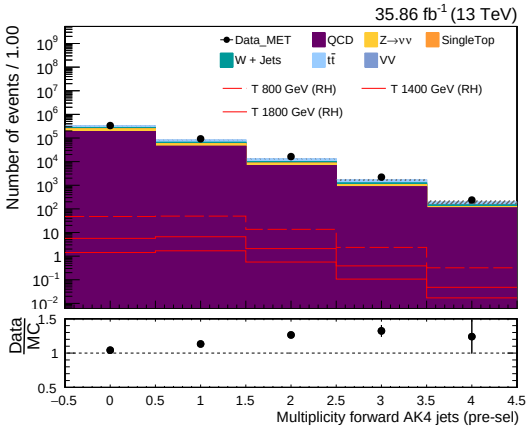


Figure 5.2: Expected distribution, for both data and simulation, of forward jet multiplicity after pre-selection.

5.2.5 Top quark reconstruction

The top quark is one of the products of the VLQ T, and can be reconstructed as a single jet or two jets, one being from the W boson and one from the b jet. In the case in which the top quark is at higher energy regimes, it can be identified with a single top tagged fat jet as described in SubSection 4.4.1, this case is referred in this thesis as

"merged". The full selection applied in the merged case is the following:

1. MET $p_T > 200$ GeV,
2. veto on leptons,
3. at least 1 b-jet, from the spectator b quark,
4. at least 1 top tagged fat jet, as defined in SubSection 4.4.1,
5. $min\Delta\phi(MET, jet) > 0.6$,
6. Jet selection, 2 jet categories:
   (a) at least 1 forward jet with $2.4 < |\eta| < 4$;
   (b) 0 forward jets.

Instead, in order to consider top quarks in a lower energy regimes, favoured by the lower−mass T signals, a second case is considered where the fat jet passing the pre-selection requirements to be tagged as W, according to the criterion described in the SubSection 4.4.2. In this case the top quark is not boosted enough, for its constituent to cluster in a single fat jet, and it can be reconstructed starting from its decay products: for its constituent to cluster in a single fat jet, a W boson and a b quark, as fat jet the first, and a narrow jet the last. This case is henceforth referred as "semi-resolved".

1. MET $p_T > 200$ GeV,
2. veto on leptons,
3. 1 b-jet, from the top quark decay
4. at least 1 W tagged fat jet, as defined in SubSection 4.4.2
5. $min\Delta\phi(MET, jet) > 0.6$,
6. Jet selection, 2 jet categories:
   (a) at least 1 forward jet with $2.4 < |\eta| < 4$;
   (b) exactly 0 forward jets.

The selection with at least 2 b-jet was not considered since the statistic in a region so made was very low.

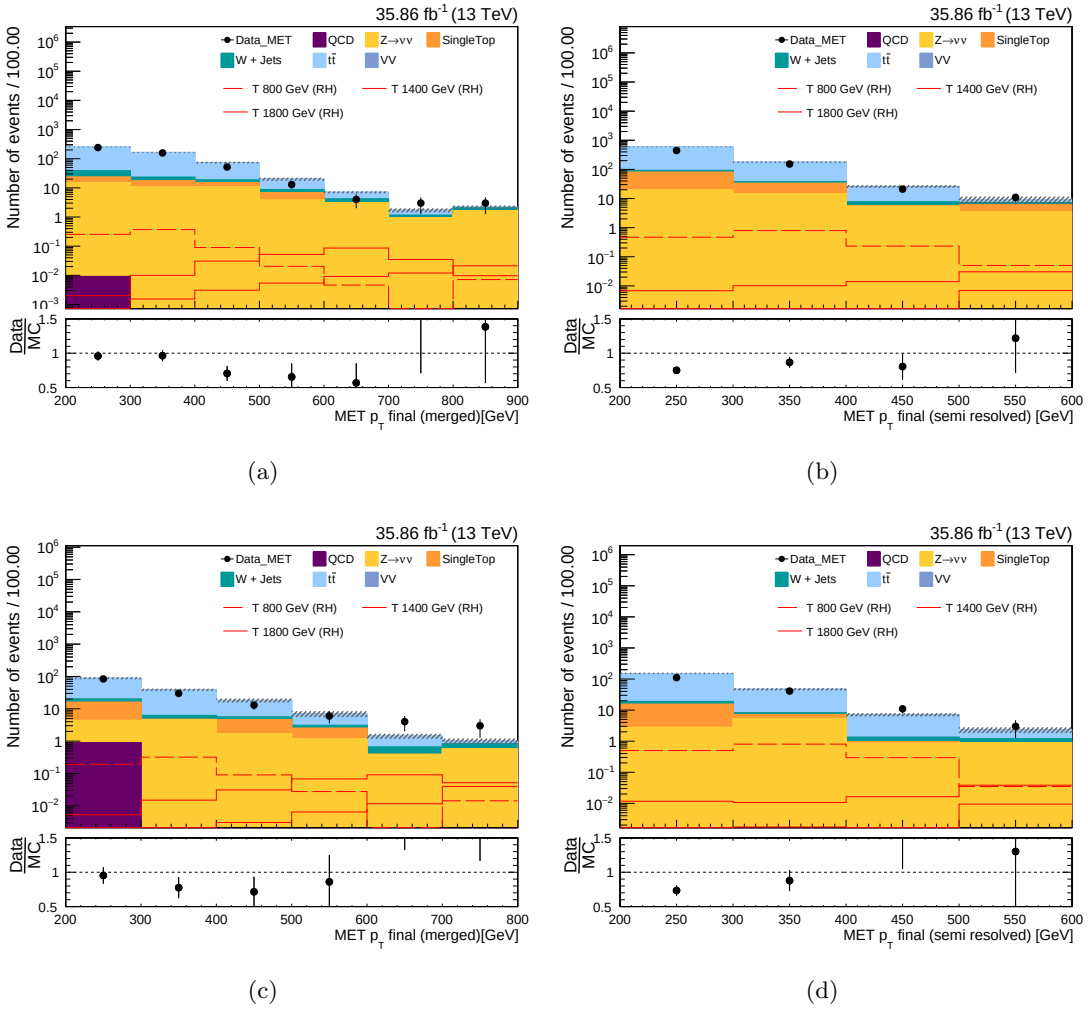


Figure 5.3: MET p_{T} spectrum expectation as from MC simulation and data after the full selection with 0 forward jets in merged case(a), and in semi resolved case(b). The MET p_{T} distribution is also shown for the category with at least 1 forward jet both for mergerd(c) and semi resolved case(d).

5.2.6 Discriminating variables

The main variable considered until now has been the MET p_{T}, to take in account the presence of neutrinos in the final state. The final MET p_{T} spectra for the forward and not-forward categories and for both merged and semi resolved cases, are shown in Figure 5.3, where all the backgrounds are taken from simulation. To increase the discriminating power between signal and background a new variable has been introduced, the transverse mass defined as:

$$M_T = \sqrt{2p_{\mathrm{T}}^{top} \cdot p_{\mathrm{T}}^{MET}\,(1 - cos\Delta\phi)} \tag{5.1}$$

where:

- p_{T}^{top} is the transverse momenta of the top quark;
- p_{T}^{MET} is the transverse momenta of the missing transverse energy;
- $\Delta\phi$ is the angular difference between the top quark p_{T} and the MET p_{T}.

In Figure 5.4 the distribution of the transverse mass for both categories and cases are shown. The range of the x-axis and the bin size have been chosen in order to avoid the

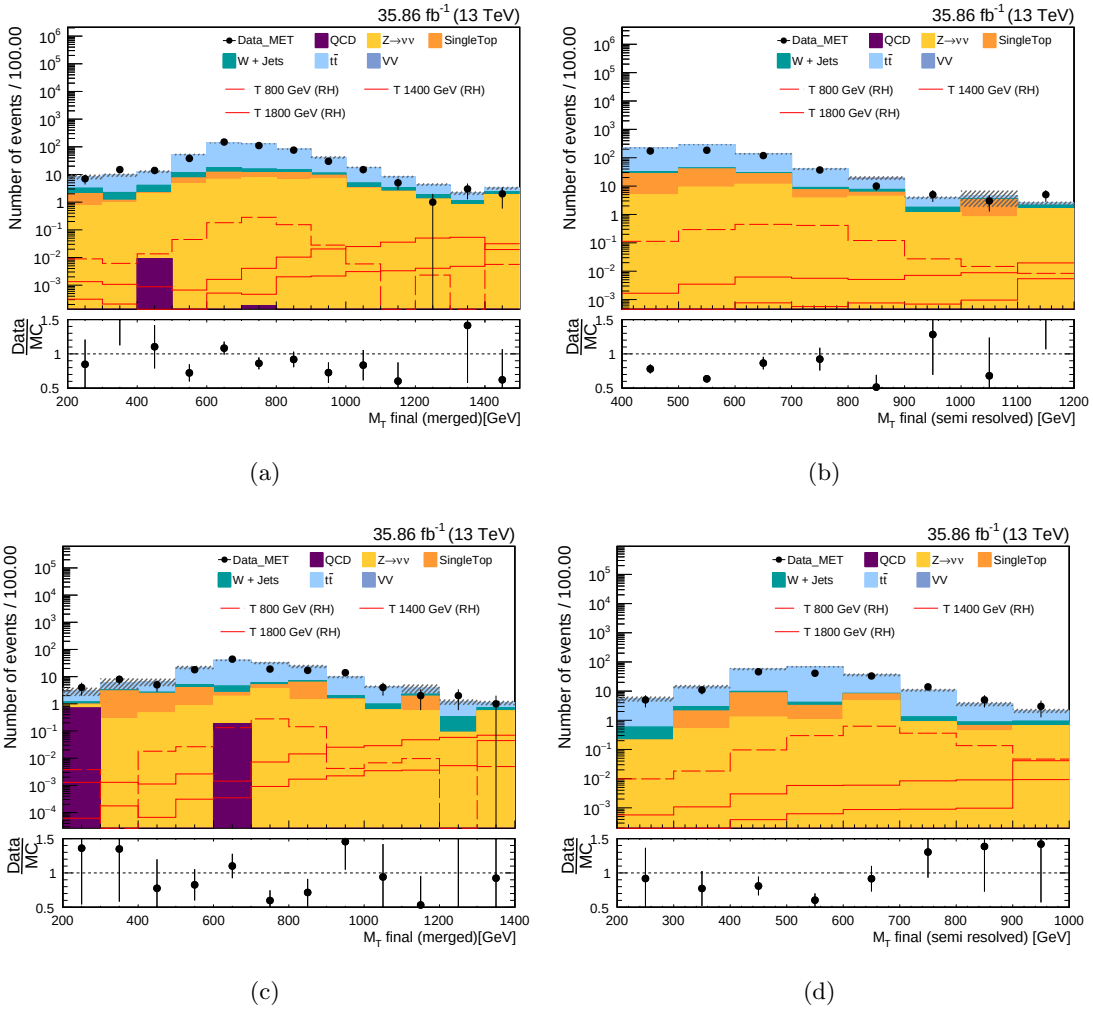


Figure 5.4: Transverse mass distribution for data and MC samples after the full selection with 0 forward jets, for the merged(a) and semi resolved case(b), and for at least 1 forward jet category in the merged(c) and in the semi resolved case(d).

lack of statistics at higher values of the variables considered. In all the plots shown the last bin is inclusive of the overflow.

5.3 Background estimation

Looking at the Tables 5.4 and 5.5, referring to the merged and semi-resolved case respectively, it can be noticed that the dominant contribution in both cases comes from the $t\bar{t}$. In order to avoid to rely solely on simulation it will be estimated from data thanks to a Data-Driven method, as will be explained in the following. The main irreducible background is given by $t\bar{t}$ events, where one top quark decays hadronically and the other leptonically, and the lepton is not reconstructed or falls outside the detector acceptance. In order to rely as little as possible on the simulation, the dominant irre-

	MET p_{T}		M_T	
	Category 0	Category 1	Category 0	Category 1
QCD	0.01 ± 0.07	0.89 ± 1.59	0.01 ± 0.07	0.89 ± 1.59
$t\bar{t}$	423.73 ± 6.17	118.19 ± 3.31	414.26 ± 6.12	116.41 ± 3.30
W + Jets	27.67 ± 1.05	7.54 ± 0.55	27.24 ± 1.04	7.46 ± 0.55
Single Top	21.71 ± 5.27	15.95 ± 5.37	21.59 ± 5.27	15.95 ± 5.37
$Z \rightarrow \nu\nu$	46.26 ± 4.00	11.86 ± 2.34	45.47 ± 4.00	11.72 ± 2.34
VV	0.49 ± 0.53	0.00 ± 0.00	0.49 ± 0.53	0.00 ± 0.00
Backgrounds	519.87 ± 9.12	154.43 ± 6.94	509.06 ± 9.09	152.44 ± 6.93
m_T 800 GeV	0.83 ± 0.07	0.71 ± 0.06	0.81 ± 0.07	0.69 ± 0.06
m_T 1200 GeV	0.47 ± 0.02	0.55 ± 0.02	0.47 ± 0.02	0.54 ± 0.02
m_T 1400 GeV	0.23 ± 0.01	0.27 ± 0.01	0.23 ± 0.01	0.27 ± 0.01
m_T 1600 GeV	0.11 ± 0.00	0.13 ± 0.00	0.11 ± 0.00	0.13 ± 0.00
m_T 1800 GeV	0.06 ± 0.00	0.07 ± 0.00	0.05 ± 0.00	0.07 ± 0.00
Data	474.00 ± 21.77	140.00 ± 11.83	466.00 ± 21.59	138.00 ± 11.75

Table 5.4: Number of expected data, signal and background events after the final selection in the merged case for both categories and variables considered.

ducible background is partially estimated from data using the alpha ratio method. The method consists of an extrapolation of the background nomalization and shape of the discriminating variable from a background-enriched region, control region (CR), into the signal region (SR). The distribution of a variable, e.g. MET p_{T} or M_T, will be used to estimate the background contribution in the SR. In this thesis were studied SR and CR for the two cases merged and semi resolved, and the background has been estimated from the distribution of the MET p_{T} and from the distribution of the M_T. In the merged case, the signal region is the one in which the main selection is made, its corresponded CR has been determined requiring exactly one lepton, instead of vetoing it, and removing the cut on the angular variable $min\Delta\phi(MET, jet)$, in order to increase the statistics of the sample. The selection applied is the merged case is the following:

- MET $p_{\mathrm{T}} > 200$ GeV;
- exactly 1 lepton;

	MET p_{T}		M_T	
	Category 0	Category 1	Category 0	Category 1
QCD	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.00 ± 0.00
$t\bar{t}$	653.58 ± 10.16	176.43 ± 4.98	583.64 ± 9.39	166.80 ± 4.79
W + Jets	20.42 ± 0.98	5.16 ± 0.50	15.85 ± 0.86	4.74 ± 0.48
Single Top	81.12 ± 11.07	14.53 ± 4.67	76.86 ± 10.80	14.38 ± 4.67
$Z \to \nu\nu$	43.39 ± 5.51	9.96 ± 2.55	37.15 ± 5.45	9.73 ± 2.55
VV	5.85 ± 1.91	1.24 ± 0.86	4.30 ± 1.67	1.24 ± 0.86
Backgrounds	804.37 ± 16.15	207.33 ± 7.36	717.80 ± 15.43	196.91 ± 7.23
m_T 800 GeV	1.71 ± 0.10	1.81 ± 0.10	1.59 ± 0.09	1.76 ± 0.10
m_T 1200 GeV	0.17 ± 0.01	0.19 ± 0.01	0.17 ± 0.01	0.19 ± 0.01
m_T 1400 GeV	0.06 ± 0.00	0.08 ± 0.00	0.06 ± 0.00	0.08 ± 0.00
m_T 1600 GeV	0.03 ± 0.00	0.03 ± 0.00	0.03 ± 0.00	0.03 ± 0.00
m_T 1800 GeV	0.01 ± 0.00	0.01 ± 0.00	0.01 ± 0.00	0.01 ± 0.00
Data	630.00 ± 25.10	166.00 ± 12.88	536.00 ± 23.15	158.00 ± 12.57

Table 5.5: Number of expected data, signal and background events after the final selection for the categories with at least 1 forward jet (Category 1) and zero jets (Category 0) for both variables considered, MET p_{T} and M_T, in the semi resolved case.

- at least 1 b-jet;
- at least 1 top tagged fat jet.

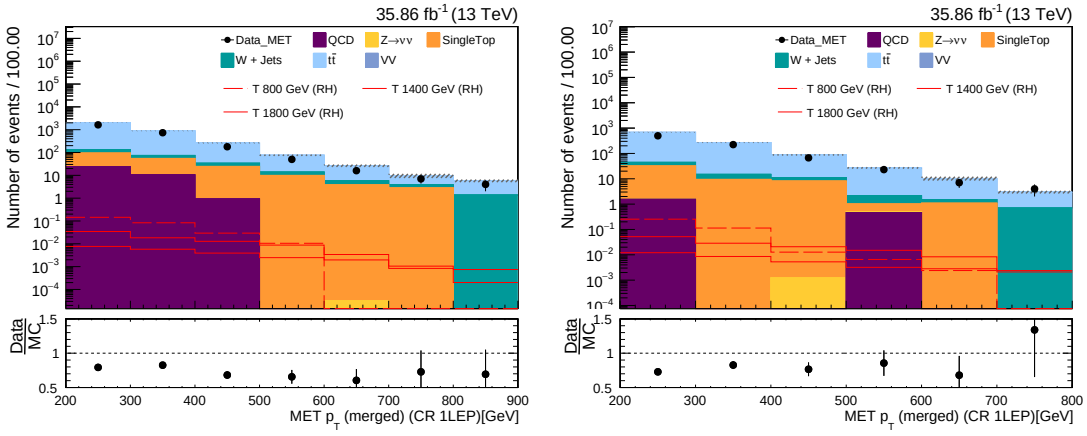
The distributions of the MET p_{T} and the M_T in the control regions are shown in Figures 5.5 and 5.6, with the comparison between simulation and data for both the categories with zero forward jets and with at least one forward jet in the merged case. In the semi resolved case the signal region is again the one obtained with the full selection described above and the CR investigated is been obtained requiring:

- MET $p_{\mathrm{T}} > 200$ GeV;
- 0 leptons;
- $min\Delta\phi(MET, jet) < 0.6$;
- exactly 1 b-jet;
- at least 1 W tagged fat jet.

Figures 5.7 and 5.8 show the distribution of METp_{T} and M_T, respectively, for the CRs for both the categories with 0 and at least 1 forward jet in the semi-resolved case.

5.3.1 Summary of signal and control regions selections

A summary of the selections used to identify the control regions and the signal regions is listed in Table 5.6.



(a) (b)

Figure 5.5: Distribution of MET p_{T} in control region, in the merged case, for the category with zero forward jets(a) and for the one with at least one forward jet(b).

Requirement	**merged**		**semi resolved**	
	SR	CR	SR	CR
MET p_{T}	>200 GeV	>200 GeV	>200 GeV	>200 GeV
leptons	vetoed	1 lepton	vetoed	vetoed
b-jet	at least 1	at least 1	1	1
fat jet	top tagged	top tagged	W tagged	W tagged
$min\Delta\phi(MET, jet)$	>0.6	–	>0.6	<0.6
Forward Jet selection	at least 1 0 jets	at least 1 0 jets	at least 1 0 jets	at least 1 0 jets

Table 5.6: Summary table of the full selection used to identify the control region, (CR), and the signal region, (SR), for both the cases merged and semi resolved.

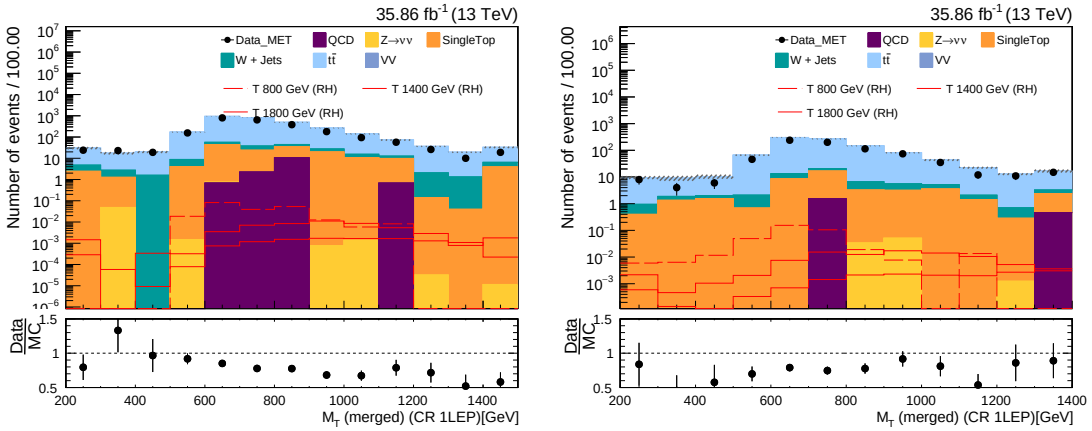
5.3.2 Background extraction

Once identified the control and the signal regions, the background in the signal regions is evaluated as follow:

$$N_{SR} = \alpha \cdot N_{CR}^{data-(other\ background)}$$
$$\alpha = \frac{N_{SR}^{t\bar{t}}}{N_{CR}^{t\bar{t}}} \tag{5.2}$$

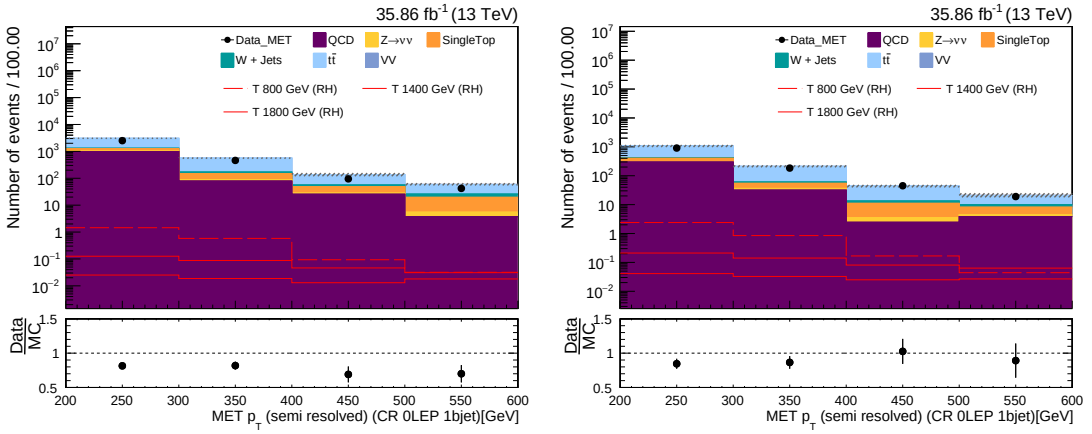
where:

- N_{SR} is the resulting number of backgorund events in the SR for the background that have to be estimated;



(a) (b)

Figure 5.6: M_T distribution in control region, in the merged case, for the category with zero forward jets(a) and for the one with at least one forward jet(b).

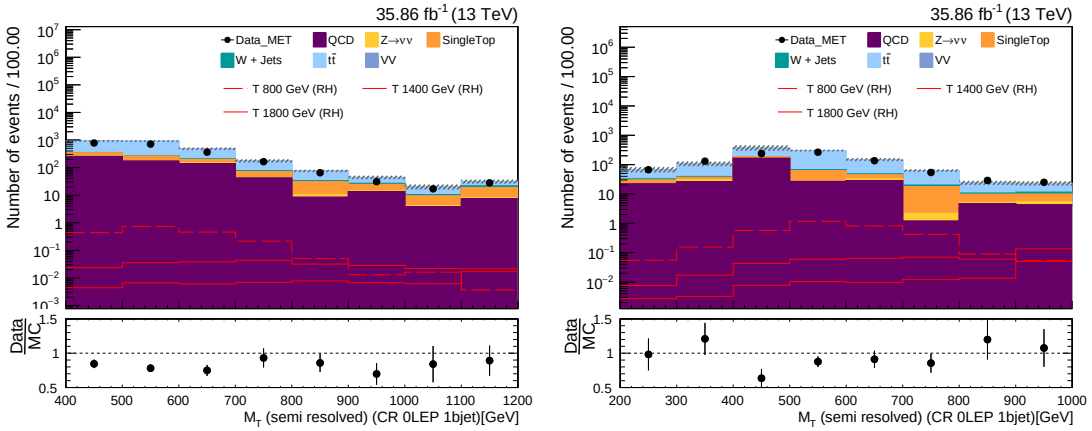


(a) (b)

Figure 5.7: MET p_{T} distribution in control region, in the semi resolved case, for the category with zero forward jets(a) and for the one with at least one forward jet(b).

- $N_{CR}^{data-(other\ background)}$ is the number of data events in the CR, subtracted of the number of MC events of QCD, Single Top, W+jets, Z+jets, and Dibosons;
- $N_{SR}^{(\mathrm{t\bar{t}})}$ is the number of events of background in the SR according to simulation;
- $N_{CR}^{(\mathrm{t\bar{t}})}$ is the number of events of background in the CR according to simulation.

The shape for the $\mathrm{t\bar{t}}$ is then obtained from the distribution of data in the CR. In Figures 5.9 and 5.10 the alpha ratio plots can be seen for the MET p_{T} and M_T in both cases merged and semi resolved, for both categories, with 0 forward jet and at least one forward jet, with a comparison of the $\mathrm{t\bar{t}}$ data driven with $\mathrm{t\bar{t}}$ from simulation. The plots



(a) (b)

Figure 5.8: M_T distribution in control region, in the semi-resolved case, for the category with zero forward jets(a) and for the one with at least one forward jet(b).

in Figures 5.11 and 5.12 show a comparison between data, background taken from MC simulation and background partially estimated with the data driven method. The good agreement of data with the background partially estimated with the data driven method confirms the possibility to use the CR to estimate $t\bar{t}$ background sample.

5.4 Systematics uncertainties

This section includes a list of the relevant systematic uncertainties for this analysis and a description on how they are estimated. Systematic uncertainties do affect the background and signal prediction, and have to be taken into account when extracting the signal value from the fit. The effect of the systematics can be of two types:

- **Yield effect only**, modify just the integral of the distribution of the variable used in the fit, changing in a flat way the distribution.
- **Yield and shape effect**, modify the integral and the shape of the distribution of the variable used in the fit, changing not only in a flat way the distribution but also re-shaping it.

In the following Table 5.7 are summarized the uncertainties used and is specified their effect on the variables considered in the fit. The uncertainties have been considered to have an effect on the shape of the MET p_{T} and in the shape of the M_T distributions, in the Appendix A several plots for the MET p_{T} distribution in the merged case are shown.

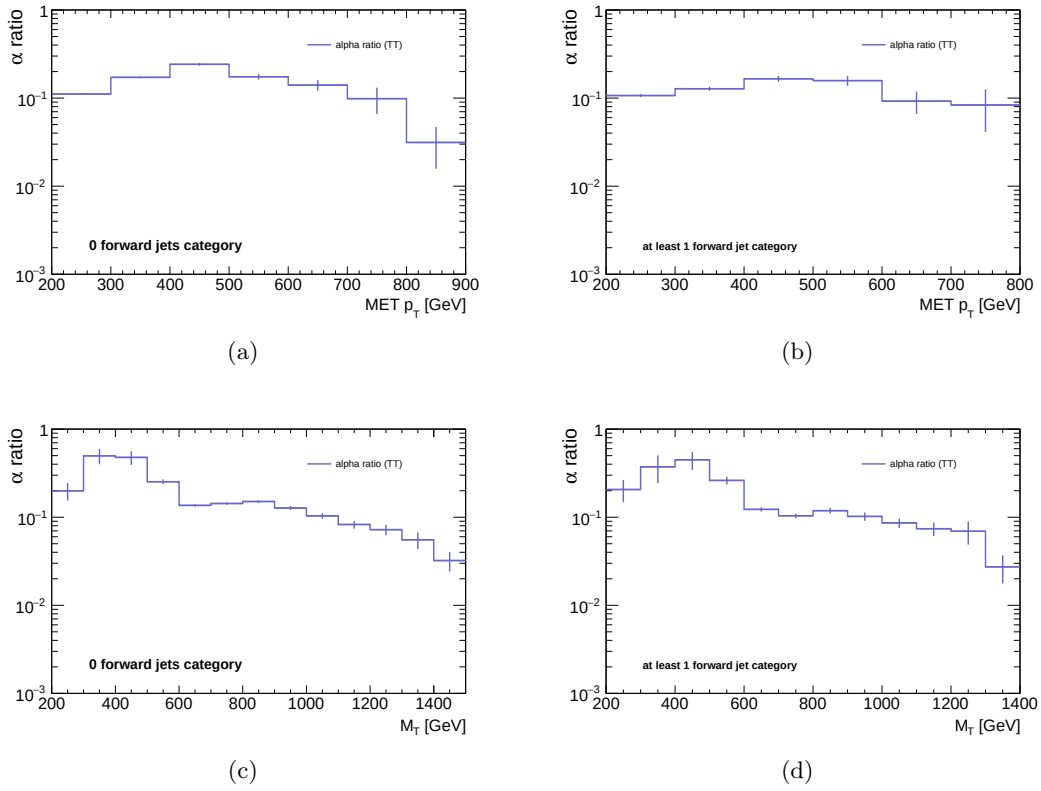


Figure 5.9: Alpha ratio distributions as function of MET p_{T}, in the merged case, for the category with 0(a) forward jets and for at least 1 forward jet(b), and M_T for the category with 0 forward jets(c) and for at least 1 forward jet(d).

Luminosity

A standard uncertainty of 2.6% on the amount of integrated luminosity of the data sample at 13 TeV, 35.89 $\mathrm{cm}^{-2}\mathrm{s}^{-1}$, is used [70].

b-tagging and mis-tag efficiency scale factors

b tagging and misidentification efficiencies are estimated from control samples in 13 TeV data [60]. Scale factors are applied to MC samples to reproduce efficiencies observed in data, and to consider the systematic effect of the scale factors, the nominal values are varied "up"$(+1\sigma)$and "down"(-1σ) both for the b tagging and mistag efficiency. The uncertainty of the b-tagging efficiency SF is estimated in different p_{T} bins. For jets originated by a c quark, the uncertainty is considered as twice the error for b-jets. The uncertainty, instead, of the SF for light jets is estimated as a flat scale factor. The uncertainty of the b tagging SF is assumed to be fully correlated for b and c jets, while it is assumed to be uncorrelated with light jets.

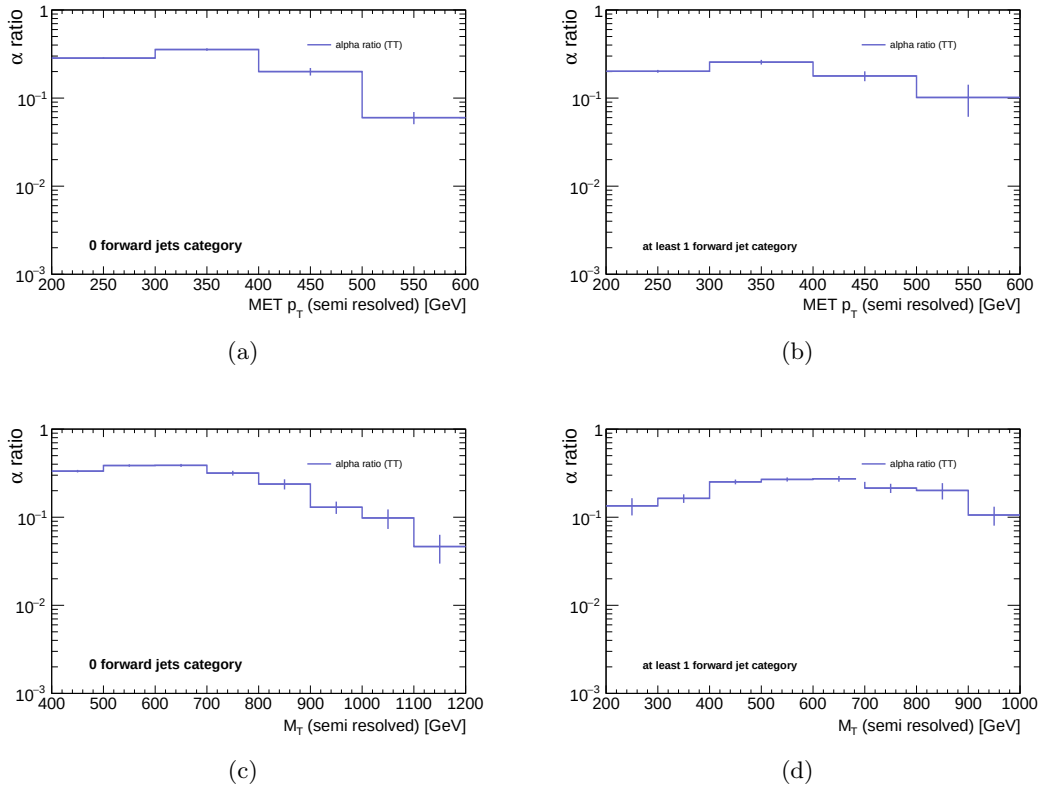


Figure 5.10: Alpha ratio distributions as function of MET p_{T}, in the semi resolved case, for the category with 0 forward jets(a) and for at least 1 forward jet(b), and M_T for the category with 0 forward jets(c) and for at least 1 forward jet(d).

Pileup modeling

In order to correct for the number of primary vertices in data and in the simulated MC samples, a scale factor is applied. The systematic uncertainties related to pileup modelling are taken into account by varying by ±4.6% the minimum bias cross section of 69.2 mb, used to calculate the data pileup distributions.

Parton Distribution Function (PDF)

The theoretical uncertainty due to the choice of the parton distribution function for MC samples is estimated by reweighting all distributions for the backgrounds with the different NNPDF3.0 [71] replicas. For signal events only the effect of this uncertainty on acceptance is considered.

Factorisation and renormalisation scales (q^2)

The theoretical uncertainty introduced by the choice of the factorisation and renormalisation scales is estimated by using distributions obtained by halving or doubling

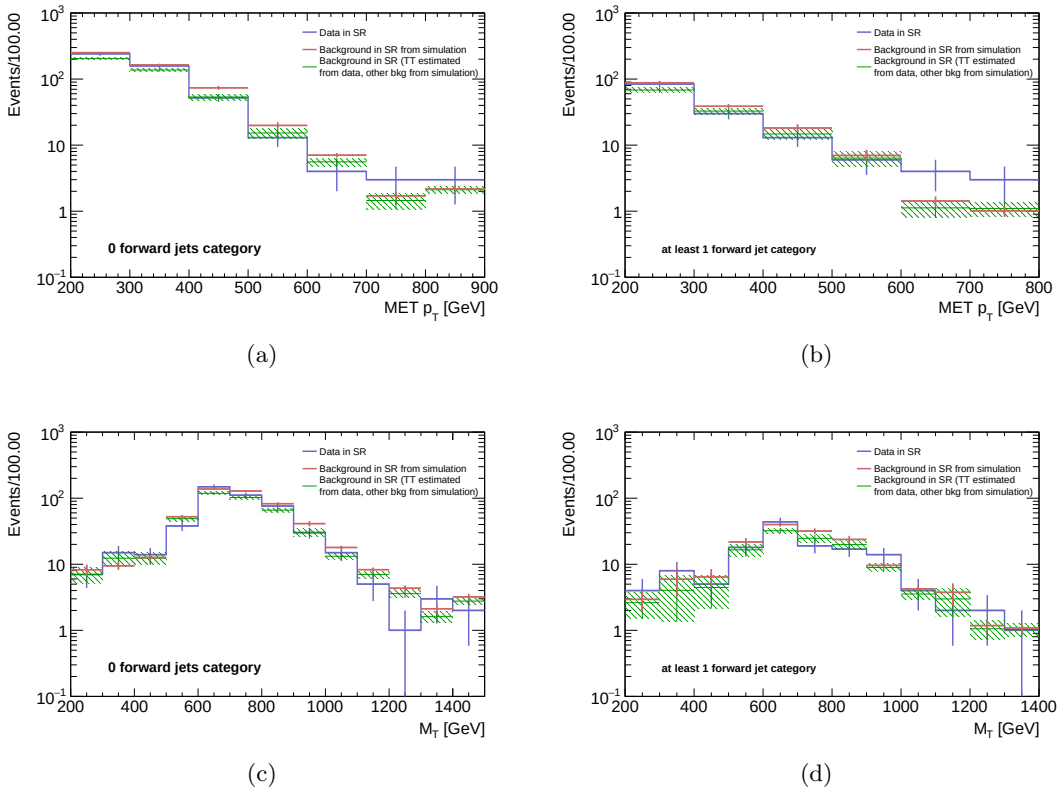


Figure 5.11: Distributions of data, background partially estimated from data and background taken all from simulation, of MET p_{T}, in the merged case, for the category with 0 forward jets(a) and for at least 1 forward jet(b), and M_T for the category with 0 forward jets(c) and for at least 1 forward jet(d).

the scales. The uncertainties for the different backgrounds are considered uncorrelated. Only the effect on acceptance is considered for signals MC samples, while also the cross section is considered for the minor backgrounds.

top tagging scale factors

In order to correct the efficiency in tagging a boosted fat jet as top quark, need to be considered scale factors that have to be applied on MC events. The errors introduced with the use of scale factor are propagated summing or subtracting the associated uncertainty to the nominal value. The variations from the central value obtained in this way represent the systematic uncertainties for the top quark tagging.

W tagging scale factors

The procedure adopted to calculate the systematic uncertainties for the top quark tagging are implemented in the case of W tagging of a fat jet in order to reconstruct a

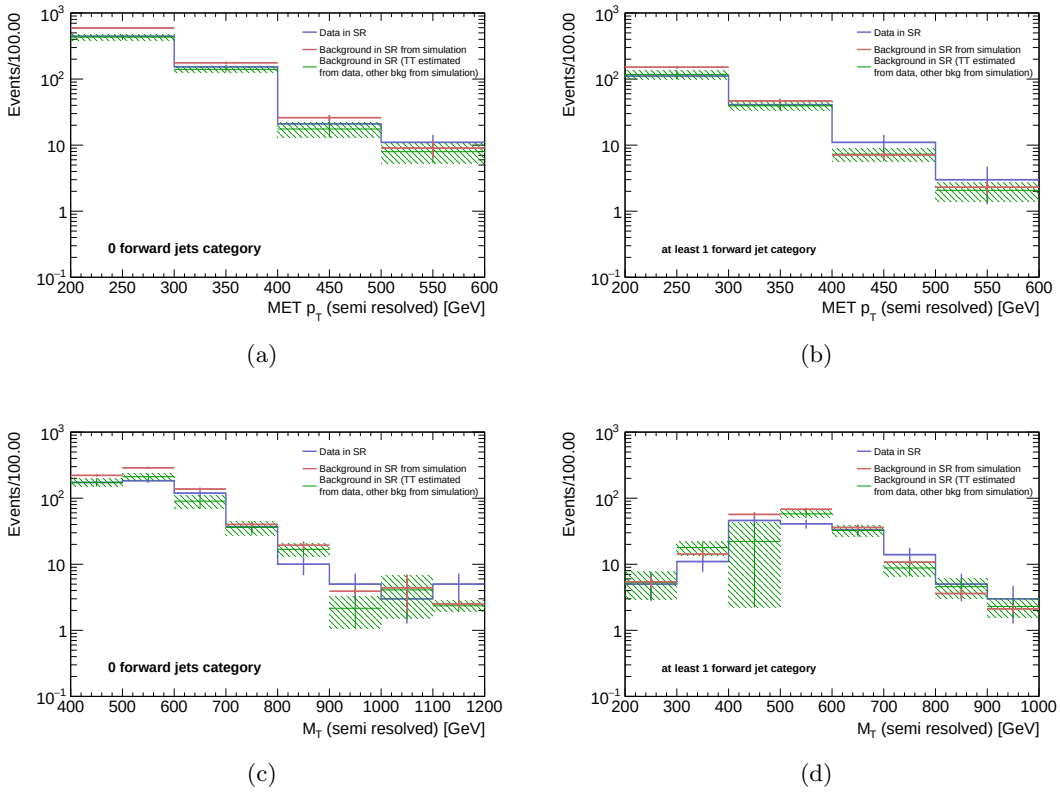


Figure 5.12: Distributions of data, background partially estimated from data and background taken all from simulation, of MET p_{T}, in the semi resolved case, for the category with 0 forward jets(a) and for at least 1 forward jet(b), and M_T for the category with 0 forward jets(c) and for at least 1 forward jet(d).

less boosted top quark.

Simulation statistics

Another source of uncertainty is due to the limited number of events in the simulated samples. The uncertainty is included by allowing each bin of each variable taken in account, MET p_{T} and M_T, to fluctuate up and down, around the nominal value, with the value of its uncertainty, which stems from the poisson uncertainty of the bin. To make this treatment possible, the choice of binning has been made in a way to have a minimum of 10 events per each bin.

Trigger scale factor

The scale factor coming from the trigger selection has been measured implemented a new trigger requiring an orthogonal selection to the one made for the signal extraction. The new trigger implemented is labelled as SingleMuon trigger. The online trigger

Uncertainty	Yield effect only	Yield and shape effect
Luminosity	✓	
b-tagging and mis-tag	✓	
Pileup modeling	✓	
PDF	✓	
q^2		✓
top tagging	✓	
W tagging	✓	
Simulation statistics		✓
Trigger Scale factor	✓	
jes		✓
jer		✓
Top p_T reweighting	✓	

Table 5.7: Summary table of the uncertainties considered in this thesis. The uncertainties are classified to take in account their effect on the distribution of the variables used in the fit.

requirement is to have one muon with at least 50 GeV reconstructed with the online trigger algorithms. To match it and to be sure to be on the plateau of the single muon trigger, the selection described previously has to be modified. Instead of vetoing the muons, one requires:

- at least one Tight muon;
- p_T of the Tight muon>60 GeV.

The variable in the offline selection that corresponds to the hadronic trigger is actually defined as METp_T $no\mu$:

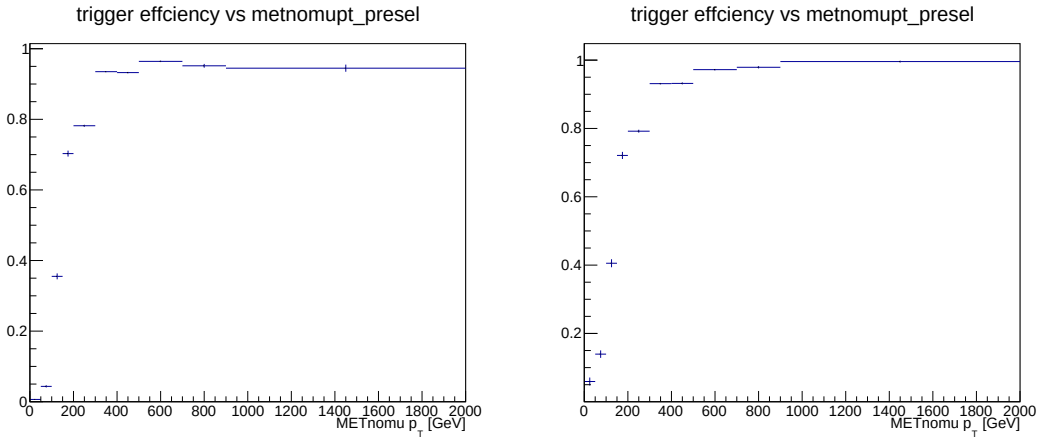
$$METp_T\ no\mu = \mid \overrightarrow{MET\ p_T} + \overrightarrow{loose\mu p_T} \mid$$

i.e. is the MET p_T considered till now but subtracted of the contribution of the loose muons. At this point the efficiency is measured in data and in a MC sample composed of the two main backgrounds populating this control regions, $t\bar{t}$ and W + jets. The efficiency is defined as:

$$efficiency = \frac{\text{N hadronic trigger \&\& N single} - \mu \text{ trigger}}{\text{N single} - \mu \text{ trigger}}$$

where the numerator is given by the "AND" logic operation of the events that passes the hadronic and the single$-\mu$ trigger while the denominator is given by the number of events that passes only the single$-\mu$ trigger. In Figure 5.13 are shown the efficiencies as a function of the variable METp_T $no\mu$ both for data and MC main backgrounds using a

variable binning in order to avoid lack of statistics in the region with high value of p_{T}. The scale factor in Figure 5.14, is then obtained, as:



(a) Efficiency versus MET$p_T no\mu$ spectrum for data. (b) Efficiency versus MET$p_T no\mu$ spectrum for MC main backgorund, $\mathrm{t\bar{t}}$ and $\mathrm{W}+\mathrm{jets}$.

Figure 5.13: Efficiency versus METp_{T} $no\mu$ calculated for data(a) and MC(b).

$$Scale\ factor = \frac{\text{data efficiency vs METp}_{\mathrm{T}}\ \text{no}\mu}{\text{MC efficiency vs METp}_{\mathrm{T}}\ \text{no}\mu}$$

The uncertainties introduced are due to variation "up" and "down" evaluated by sum-

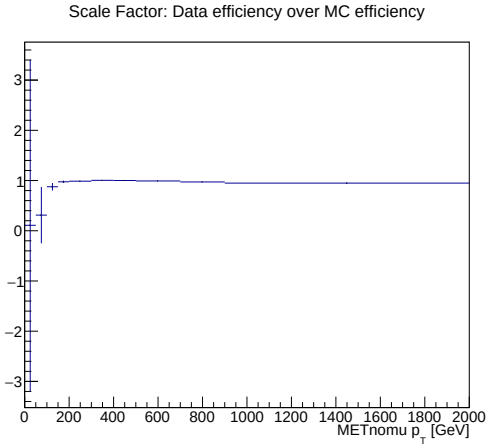


Figure 5.14: Scale Factor versus METp_{T} $no\mu$ with variable binning.

ming or subtracting to the nominal value the uncertainty on scale factor, obtained by a statistical propagation of the error, respectively.

Jet energy scale (jes)

All reconstructed jet four-momenta in simulated events are simultaneously varied according to the η- and p_{T}-dependent uncertainties in the JES [72]. This variation in jet four-momenta is also propagated to METp_{T}.

Jet energy resolution (jer)

A smearing is applied to account for the difference in the JER between simulation and data, increasing or decreasing the resolutions by their uncertainties.

Top p_{T} reweighting

The top quark cross section shows that the p_{T} spectra of its decay products (leptons and jets) are harder in the MC simulation than in data; this seems to be caused by a mismodeling of the simulated top quark p_{T} distribution. To account for this discrepancy, simulated samples events are reweighted to match the data.

The top quark transverse momentum is reweighted in order to better describe the p_{T} distribution observed in the data since the MC simulation is affect by a mismodelling of the simulated top quark p_{T} distribution. The uncertainties are evaluated as the difference with respect to the value obtained with the non-reweighted samples.

5.5 Fit procedure

Once selected the events that maximise the ratio of signal over background, the hypothesis of existence of the singly produced VLQ T has been tested. Two hypotheses were considered:

- $\mathbf{H}_0$ that assumes the absence of new physics (i.e. the signal is absent or too little to be detected);
- $\mathbf{H}_1$ that assumes the presence of the VLQ T, as predicted by theoretical models introduced in Chapter 3.

As shown in Sections 5.3.1 and 5.3, the distributions of the two variables differ for signal and background, and this fact made them good candidates for the statistical analysis. Both variables were tested for signal discrimination, and in the end the M_T was chosen. A simultaneous binned Maximum Likelihood fit is done for the two variables distribution, MET p_{T} and M_T for both the categories, with 0 forward jets and with at least 1 forward jet, and both cases, merged and semi resolved, in different scenarios: the first assuming all background coming from the MC simulation, labelled with the name of the variable considered plus the acronym MC, the latter takes into account a background that is partially estimate from data, thanks to the data driven method described in Section 5.3,

and is identified with the name of the variable considered plus the abbreviation DD. The fit categories identified are listed here:

- MET p_{T} MC;
- MET p_{T} DD;
- M_T MC;
- M_T DD.

Since for both variables, MET p_{T} and M_T, signal distributions depend on the T mass, the models considered include 12 mass points ranging from 700 GeV up to 1800 GeV in steps of 100 GeV and for 3 different width scenarios for the T resonance, ranging from a narrow width approximation to 30% of the resonance mass in steps of 10%. The fit has been performed for both variables MET p_{T} and M_T, and the latter consistently yields better expected limits, therefore is chosen for the final result.
The likelihood [73], $\mathcal{L}\left(obs \mid \mu, \theta\right)$, is the product of a Poissonian distribution multiplied a second term that represents the signal and background models,

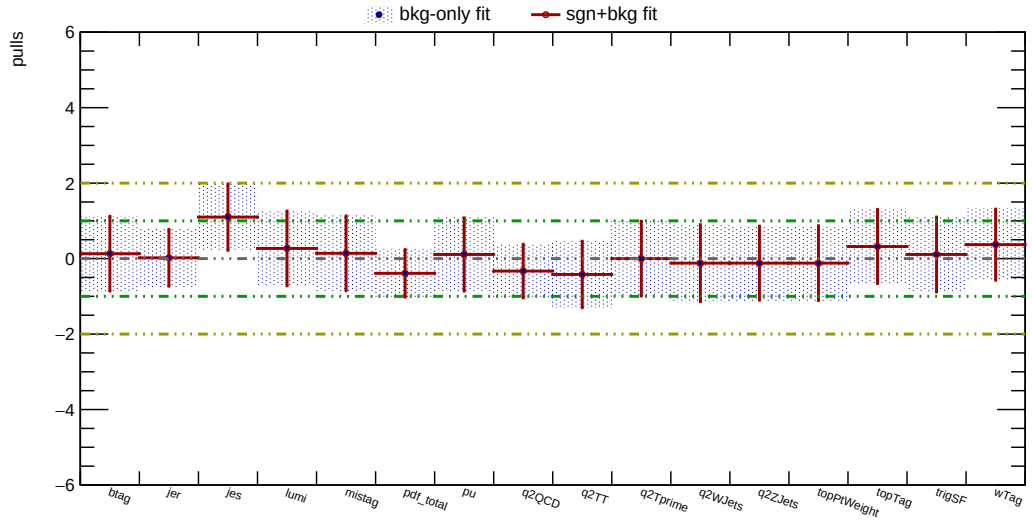
$$\mathcal{L}\left(obs \mid \mu, \theta\right) = Poisson\left(n, \mu \cdot s\left(\theta\right) + b\left(\theta\right)\right) \prod_{k=1}^{Nevent} f\left(x_k \mid \mu, \theta\right) \tag{5.3}$$

where μ is the *signal strength*, defined as the ratio of the value of the fitted parameter before and after the fit, θ is the set of systematics uncertainties, named nuisance parameters, n is the number of events, $s(\theta)$ and $b(\theta)$ are the expected yields of signal and background depending from the unknown parameters, respectively, while the probability distribution function $f\left(x_k \mid \mu, \theta\right)$ is defined as a combination of two PDFs one for signal f_s and one for background:

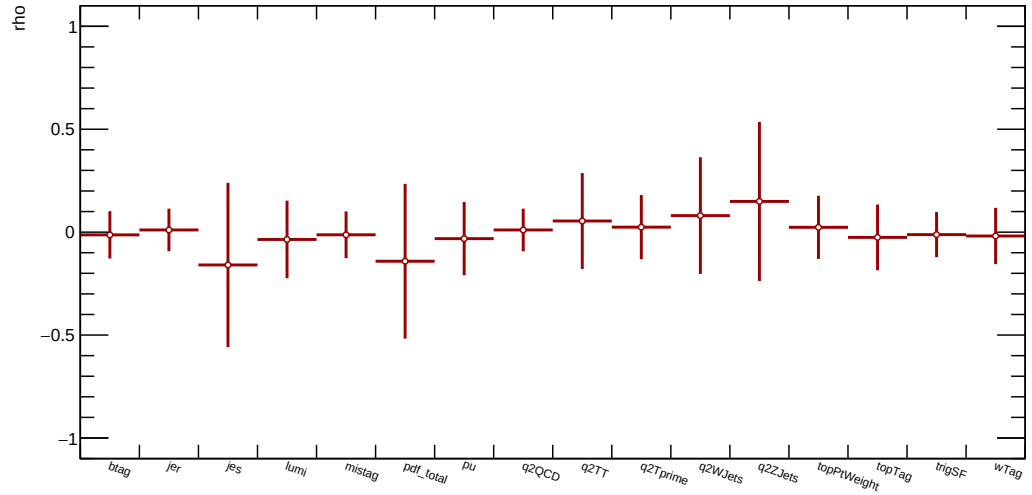
$$f\left(x_k \mid \mu, \theta\right) = \frac{\mu \cdot s\left(\theta\right)}{\mu \cdot s\left(\theta\right) + b\left(\theta\right)} f_s\left(x, \theta\right) + \frac{b\left(\theta\right)}{\mu \cdot s\left(\theta\right) + b\left(\theta\right)} f_b\left(x, \theta\right).$$

Using the binned maximum likelihood technique a shape analysis is performed, and for each independent source of systematic uncertainty, a nuisance parameter is considered. The background normalization uncertainties are modeled with a coefficient for the background templates with a log-normal prior. The shape uncertainties, instead, are modeled by choosing a Gaussian prior for the nuisance parameters, and are used to interpolate between the nominal template and other two, obtained shifting up and down the nuisance parameters by 1σ.
The fit results are returned in term of *signal strength* μ, and in Table 5.8 are reported the results obtained with the M_T fit at m = 1200 GeV using a partially data driven background. In Table 5.8 can be read the the values found for the central value and uncertainty on nuisance parameters ("pulls") for the main nuisance parameters, the ratio

of error in the model before and after the fit, and the correlation coefficient ρ between the signal strength μ and each nuisance parameter. With Δx is indicated the difference between the value of the parameter before and after the fit, instead, in the last two columns is reported the error in the model before and after the fit, and are referred to as σ_{in} and σ_{out}, respectively. In Figure 5.15 are plotted the same parameters reported in the Table 5.8 for the signal sample at value of mass of 1200 GeV in case of background partially estimated from data. Since no excess that could be ascribed to the production



(a)



(b)

Figure 5.15: Pulls with pseudo-data for each nuisance parameter and the correlation coefficient between the signal strength μ and for each nuisance parameter.

Table 5.8: Pulls with pseudo-data for the main nuisance parameter, the ratio of error in the model before and after the fit and the correlation coefficient ρ between the signal strength μ and each nuisance parameter.

Systematic uncertainty	b-only fit		$b+s$ fit		$\rho(\theta\mu)$
	$\Delta x/\sigma_{\text{in}}$	$\sigma_{\text{out}}/\sigma_{\text{in}}$	$\Delta x/\sigma_{\text{in}}$	$\sigma_{\text{out}}/\sigma_{\text{in}}$	
btag	+0.13	0.99	+0.13	0.99	-0.0132
jer	+0.02	0.75	+0.02	0.75	+0.0107
jes	+1.12	0.87	+1.10	0.88	-0.1594
lumi	+0.27	0.99	+0.27	0.99	-0.0355
mistag	+0.14	0.99	+0.14	0.99	-0.0129
pdf_total	-0.39	0.63	-0.39	0.63	-0.1413
pu	+0.12	0.97	+0.11	0.97	-0.0316
q^2_{QCD}	-0.33	0.71	-0.33	0.71	+0.0107
q^2_{TT}	-0.42	0.88	-0.42	0.88	+0.0542
q^2_{Tprime}	+0.00	0.99	-0.00	0.99	+0.0243
q^2_{WJets}	-0.13	1.02	-0.12	1.02	+0.0803
q^2_{ZJets}	-0.13	0.97	-0.12	0.98	+0.1493
topPtWeight	-0.13	0.99	-0.12	0.99	+0.0235
topTag	+0.32	0.98	+0.32	0.98	-0.0255
trigSF	+0.11	0.99	+0.11	0.99	-0.0120
wTag	+0.38	0.94	+0.37	0.94	-0.0187

of a new particle is observed, and data are compatible with the only background hypothesis, H_0, upper limits are set on production cross section of the benchmark models. The limits are calculated with a dedicated software: the *Combine tool* [77] in order to quantify the level of incompatibility of data with a signal hypothesis. The expected and observed upper limits are calculated with the modified frequentist method [74–76], and the systematic uncertainties described in Section 5.4, are taken into account as nuisance parameters, affecting both the signal and the background. The upper limits are determined on the product of the cross section and branching fraction of a VLQ T decaying to tZ. The introduction of such limits allows to take in account the possibility of variations, not yet considered, introduced by new physics that could influence the production cross section rates in the final state investigated. The test statistics used for the calculation of the limits is the ratio of the likelihood of the hypothesis H_0 and H_1, and allows to discriminate signal-like events from background-like ones. The Neyman-Pearson lemma states that the ratio of the likelihood, without any other undetermined parameter:

$$\lambda(x) = \frac{\mathcal{L}\left(x \mid H_1\right)}{\mathcal{L}\left(x \mid H_0\right)} \tag{5.4}$$

is the most powerful discriminator since minimises at a significance level α the so called error of type-II, i.e. the probability of not rejecting H_1 if H_0 is true. When is verified the case in which H_0 is true, H_0 is rejected if the ratio is smaller than a chosen constant value depending on α. Considering the nuisance parameters of the two hypotheses, a profile likelihood ratio, used as test statistic, as described by the Neyman-Pearson lemma, is:

$$q_\mu = -2\ ln \frac{\mathcal{L}\left(obs \mid \mu, \theta_\mu\right)}{\mathcal{L}\left(obs \mid \mu' + b,\ \theta\right)} \tag{5.5}$$

where θ_μ is the set of nuisance parameters, that maximises the numerator for a given value of μ, while θ give a constraint on the estimation of the likelihood maximum since it is the set of nuisance parameters that maximises the denominator, and μ' the value of μ that maximises the denominator. Events with $q_\mu \geq 0$ appear to be under the background only hypothesis, viceversa for the background plus-signal hypothesis.
To place a bound for the production of VLQ T, the Confidence Level method is used. The confidence level of the signal can be written as the ratio between the confidence level observed for the signal+background hypothesis, CL_{s+b}, and the confidence level observed for the background-only hypothesis, CL_b:

$$CL_s = \frac{CL_{s+b}}{CL_b} \tag{5.6}$$

where CL_{s+b} is defined as the probability to have for a given value of μ, a value of the test statistics equal or larger than the value observed in the experiment, under the hypothesis of signal+background H_1:

$$CL_{s+b} = P_{s+b}\left(q_\mu \geq q_\mu^{obs}\right) \tag{5.7}$$

while CL_b is the probability to have for a given value of μ a value of the test statistics equal or larger than the value observed in the experiment, under the hypothesis of background only, H_0:

$$CL_b = P_b\left(q_\mu \leq q_\mu^{obs}\right).$$

5.6 Results

The total cross section for the singly produced VLQ T decaying in a generic final state X can be written as:

$$\sigma\left(C_1, C_2, m_T, \Gamma_T, X\right) = C_1^2\ C_2^2\ \hat{\sigma}_{AW}\left(m_T, \Gamma_T\right) \tag{5.8}$$

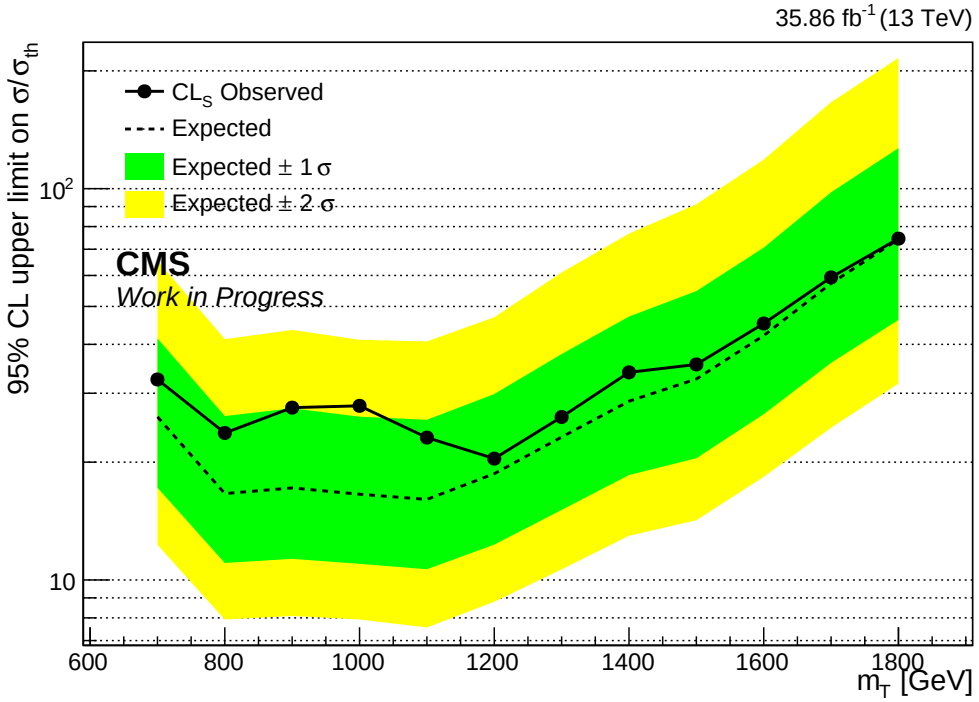
with C_1 and C_2 the production and the decay couplings corresponding to the interactions through which a T quark is produced and decays, and $\hat{\sigma}_{AW}$ is the reduced

cross section for a resonance of arbitrary width (AW). The width can be written as $\Gamma_T = \Gamma\left(C_i, m_T, m_{decays}\right)$, as it depends on the T quark mass, on the masses of all its decay products, and on its couplings to all decay channels, C_i. For the process under study it can be set $C_2 \equiv c_Z$ and $C_1 \equiv c_W$, where: $c_Z = e/(2cos\theta_W\ sin\theta_W\ \kappa_Z)$ and $c_W = e/(\sqrt{2}\ sin\theta_W\ \kappa_W)$, with e the electic charge of the proton, θ_W the weak mixing angle, and κ a coupling strenght that can be fixed to obtain the desired width. Numerically $c_W = 0.458$ and $e/2cos\theta_W\ sin\theta_W = 0.370$, while κ_Z assumes values depending on the width of the mass resonance. The values that κ_Z assumes in the theoretical framework taken in consideration [78, 79] are reported in Table 5.9 and give the cross sections written in Table 5.2. To interpret the results in a model-independent way, the mechanism through which the VLQ T achieve large widths is not specified, and Γ_T is considered as a free parameter. The results in case of narrow width approximation are

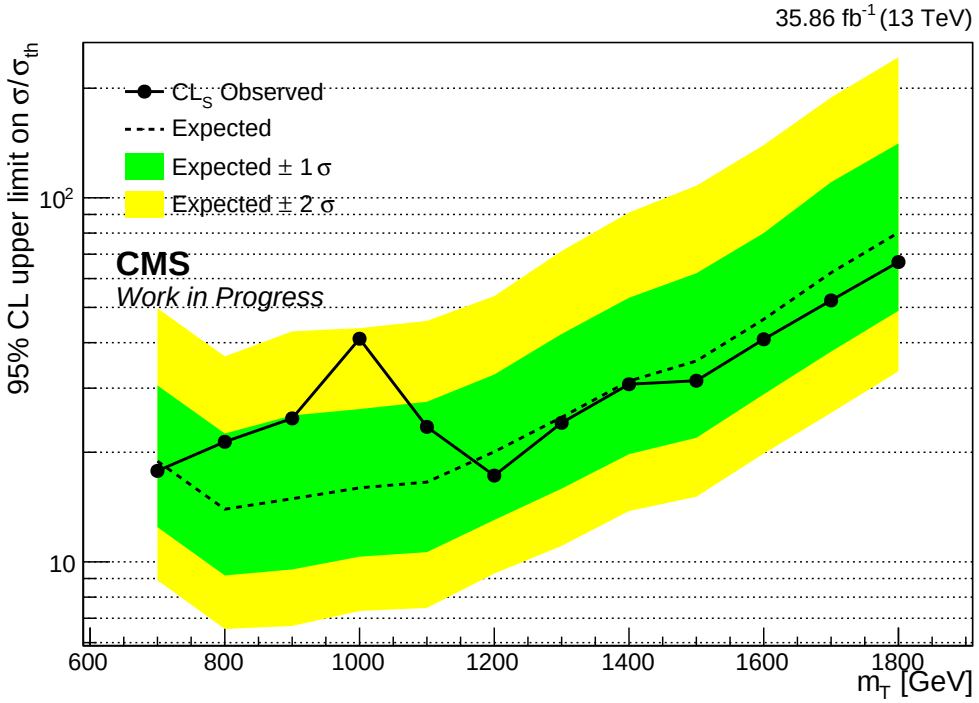
Table 5.9: κ_Z values reported for different values of mass and different values of mass resonance width.

Mass [GeV]	κ_Z			
	Width 1%	Width 10%	Width 20%	Width 30%
700	0.18			
800	0.15		0.71	0.86
900	0.14			
1000	0.12	0.39	0.56	0.69
1100	0.11			
1200	0.10	0.33	0.46	0.57
1300	0.09			
1400	0.08	0.28	0.40	0.49
1500	0.08			
1600	0.07	0.24	0.34	0.42
1700	0.07			
1800	0.06	0.21	0.31	0.37

reported in Figures 5.16 for the cases of background partially estimated from data and background coming from simulation. The limit on the ratio of σ/σ_{th} is given as function of m_T. Model-independent limits are also shown directly on the cross section for the case of narrow width approximation, shown in Figures 5.17, as well as for variable widths in Figures 5.18. In the case of narrow width approximation, cross sections exclusion upper limits are set in the range 1.5 pb to 30 fb depending on the mass. For the 10%, 20%, and 30% cases the exclusion upper limits range from 250 fb to 30 fb, from 450 fb to 40 fb, and from 300 fb to 40 fb, respectively.

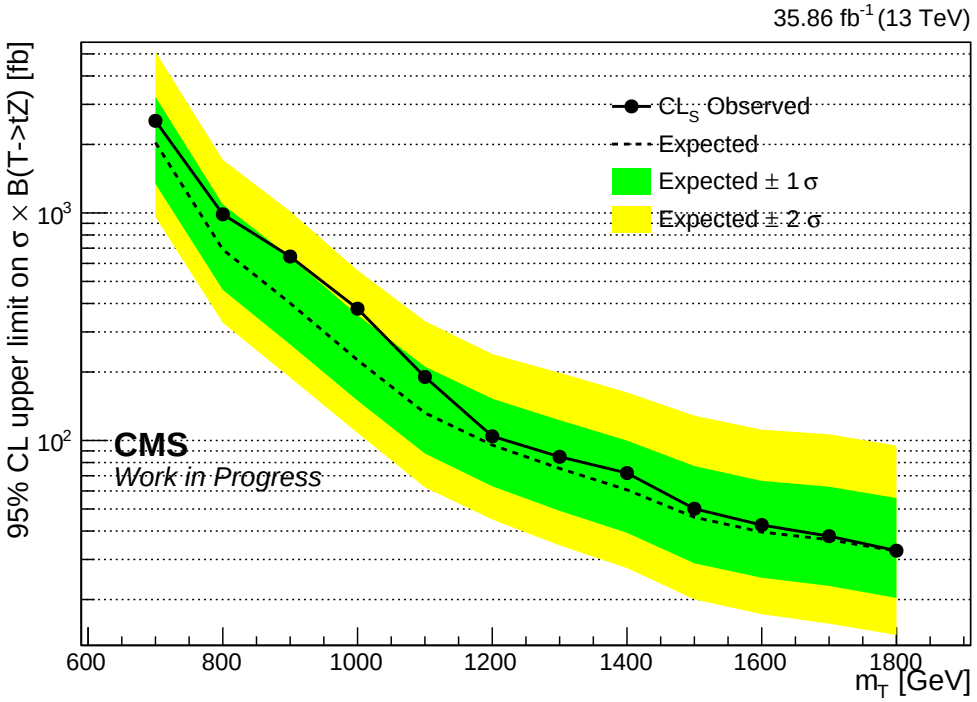


(a)

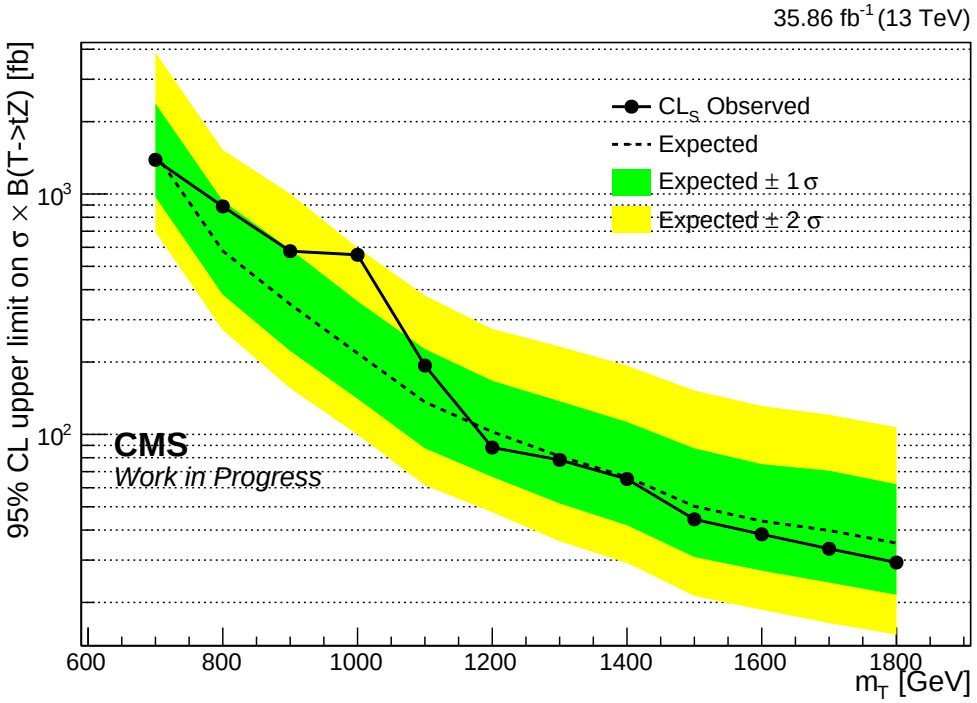


(b)

Figure 5.16: The observed and expected 95% CL upper limits of the VLQ T quark ratio of the observed over expected production cross section as a function of the signal mass, assuming narrow width resonances, resulting from the fit with partially estimated background from data(a) and background taken from MC simulations(b).

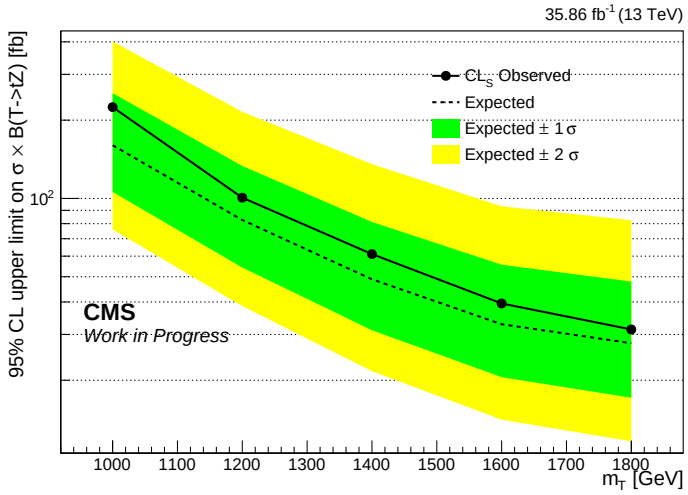


(a)

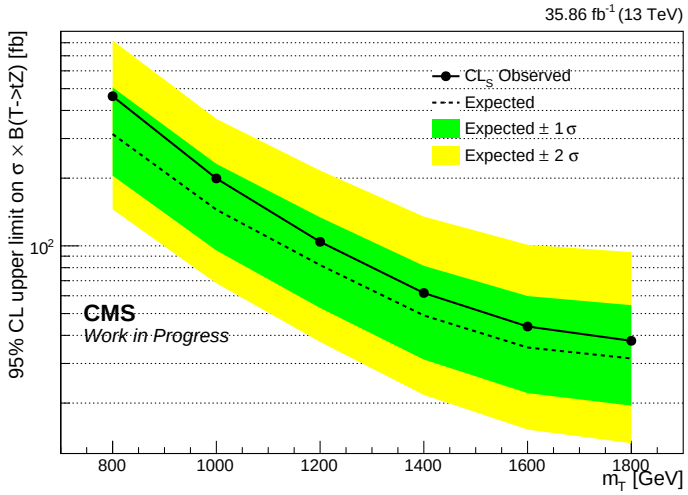


(b)

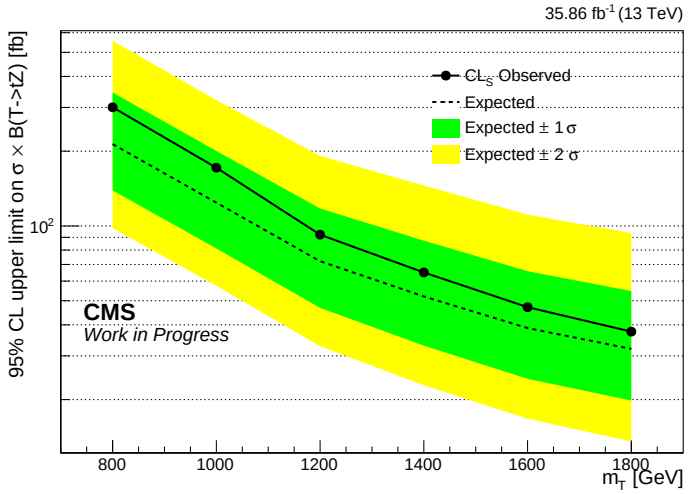
Figure 5.17: The observed and expected 95% CL upper limits on the VLQ T quark production cross section, with T decaying to tZ, as a function of the signal mass, assuming narrow width resonances resulting from the fit with partially estimated background from data(a) and background taken from MC simulations(b)



(a)



(b)



(c)

Figure 5.18: The observed and expected 95% CL upper limits on the VLQ T quark production cross section, with T decaying to tZ, as a function of the signal mass resulting from the fit with partially estimated background from data assuming width resonance mass of: 10%(a), 20%(b), and 30%(c).

Conclusions

In this thesis a search for a singly singly produced Vector-Like Quark T decaying to a top quark and a Z boson is presented. The search is performed using a data set collected in 2016 by the CMS experiment at the LHC, at a centre-of-mass energy of 13 TeV, and corresponding to an integrated luminosity of 35.86 fb^{-1}.
Final states are investigated where the top quark and the Z boson decay hadronically and to two neutrinos, respectively. The final topology is therefore characterized by missing transverse energy and jets.
The full selection is performed by identifying multiple event categories, differing by the number of jets in the forward region of the detector, and cases, differing by the top quark reconstruction method. The presence of forward jets in the final state coming from spectator quarks is exploited by identifying two categories, the first with zero forward jets and the second with at least one forward jet.
The top quark reconstruction method performed in the final state, instead, allows to take into account two cases: the merged case, where the top at higher energy regimes is reconstructed from a single "top-tagged" fat jet, and the semi resolved case, where the top at lower energy regimes is reconstructed with a "W-tagged" fat jet and a narrow b jet. These cases and categories allow to define multiple signal enriched regions where two main discriminating variables are identified, MET p_{T} and M_T.
A data driven method is applied in order to estimate the main irreducible background and to rely as little as possible on the simulation. The method consists in the extrapolation of the background from a control region into the signal regions.
A simultaneous maximum likelihood binned fit is performed on the distributions of the variables MET p_{T} and M_T taking into account systematic uncertainties.
Different benchmark models are tested based on the hypothesis made on the T quark mass, ranging from 700 up to 1800 GeV, and on the width of the resonance, ranging from a narrow width approximation to 30%, of the resonance mass.
No significant deviation from the Standard Model expectations is observed, and therefore the results obtained with the presented analysis allow to set limits upper limits on the cross sections at 95% C.L..
This search could be improved by considering a data driven method also to extract the

minor backgrounds. The determination of the systematics could be further improved by considering additional control samples to reduce even more the dependency of the data-based method on simulation assumptions. The analysis developed on 2016 data can be applied also to the data set collected by CMS in 2017 and to the data being currently collected throughout 2018.

Appendix A

Systematics

This Appendix includes the plots with comparison between nominal distributions and the same distribution shifted "up" or "down" accordingly to the considered uncertainties, for signal and background.

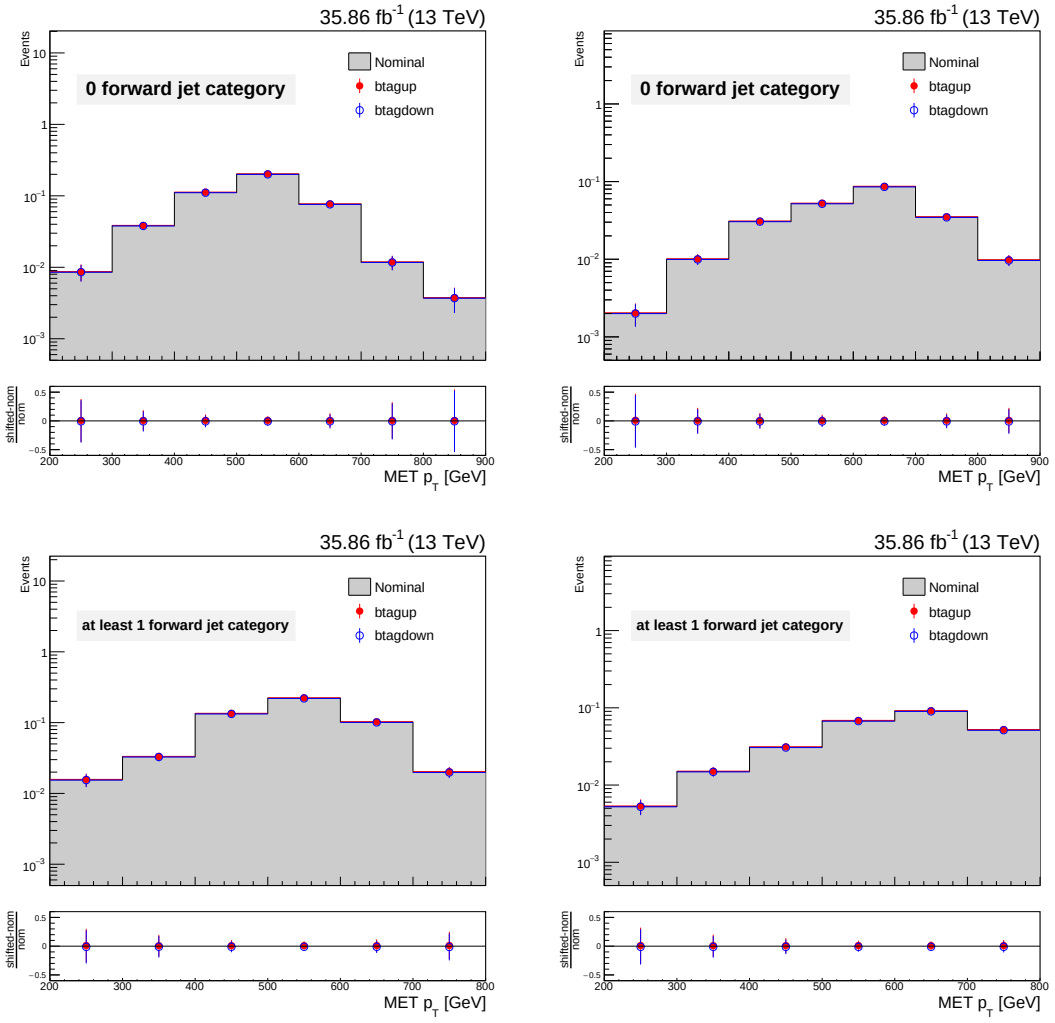


Figure A.1: Effect of b-tagging efficiency uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the signal sample with T mass hypothesis of 1200 GeV, while second one is for 1400 GeV.

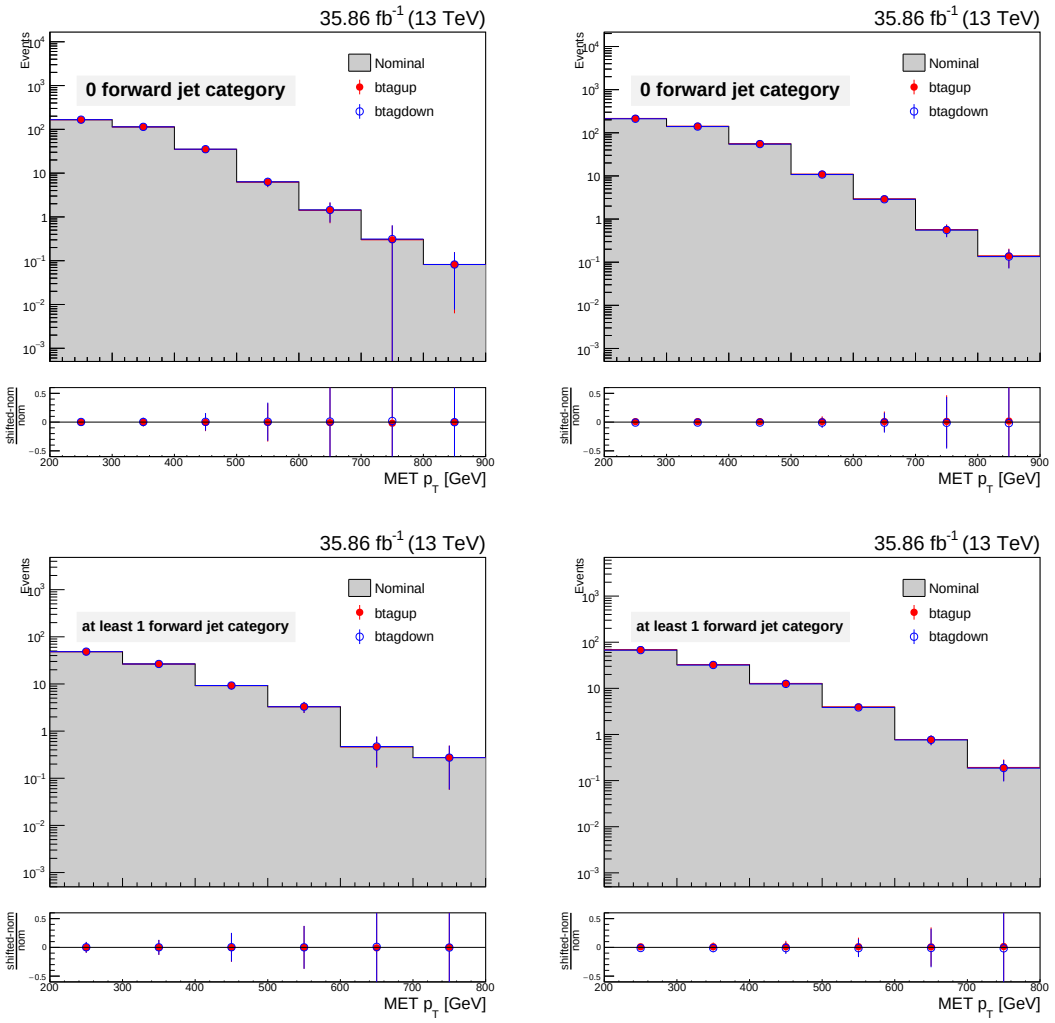


Figure A.2: Effect of b-tagging efficiency uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the $\mathrm{t\bar{t}}$ estimated from data, while second one is for $\mathrm{t\bar{t}}$ taken from MC simulation.

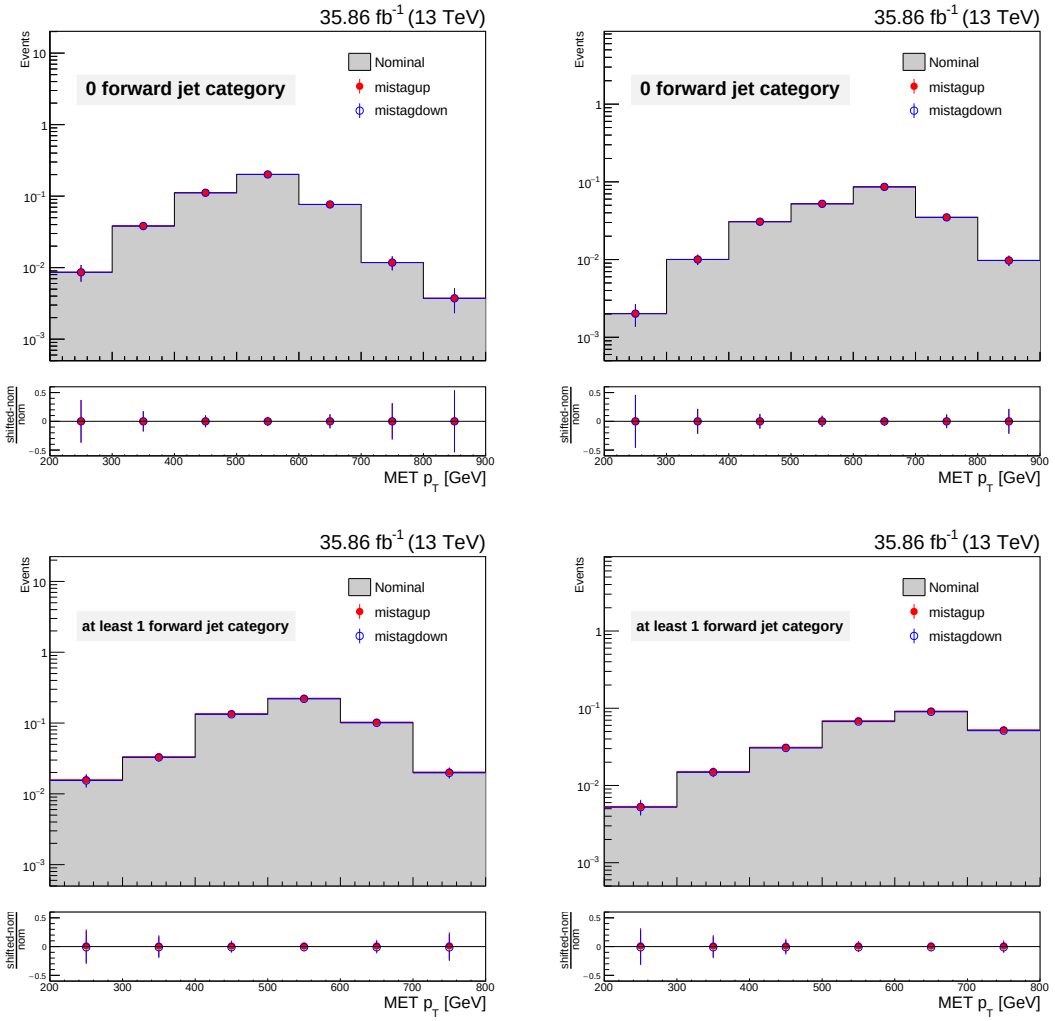


Figure A.3: Effect of mistag efficiency uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the signal sample with T mass hypothesis of 1200 GeV, while second one is for 1400 GeV.

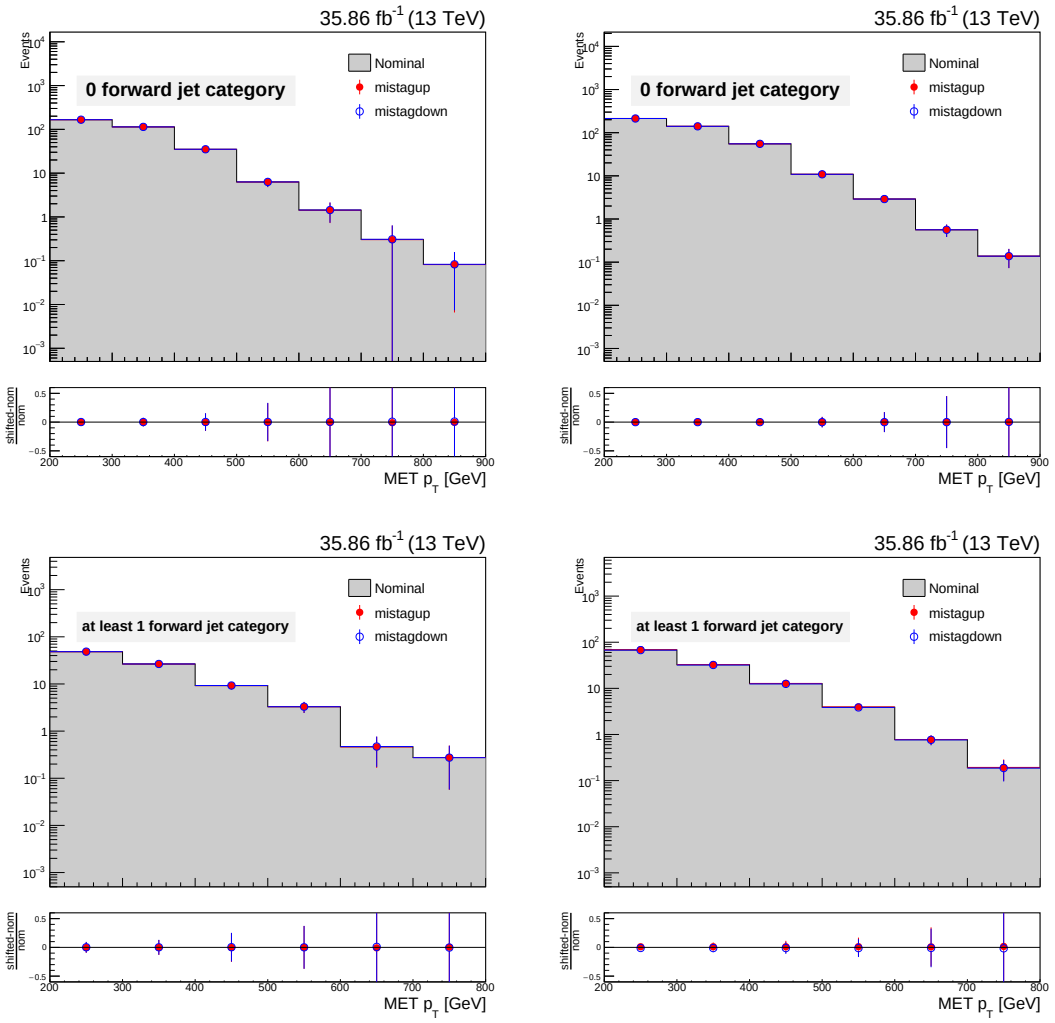


Figure A.4: Effect of mistag efficiency uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the $\mathrm{t\bar{t}}$ sample estimated from data, while second one is for $\mathrm{t\bar{t}}$ taken from MC simulation.

Pileup modeling

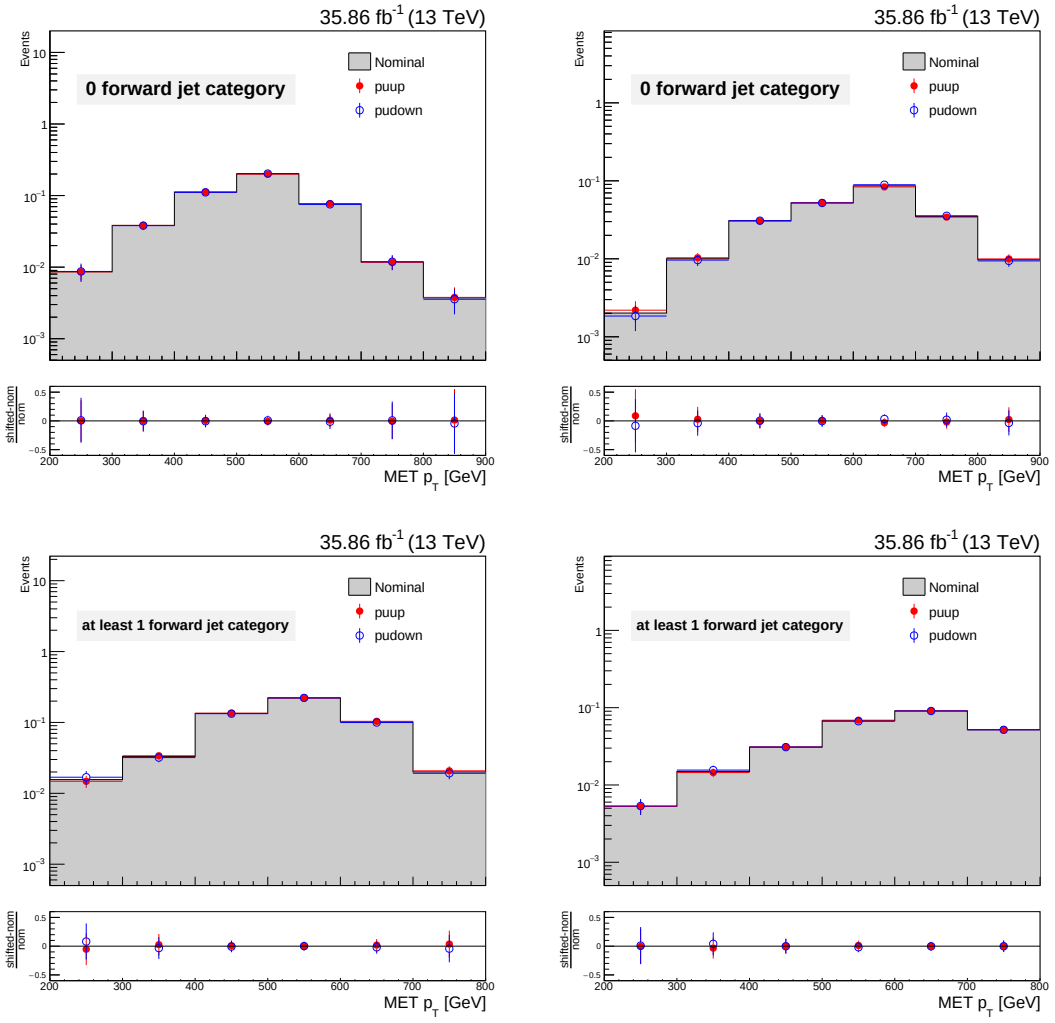


Figure A.5: Effect of pile up uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the signal sample with T mass hypothesis of 1200 GeV, while second one is for 1400 GeV.

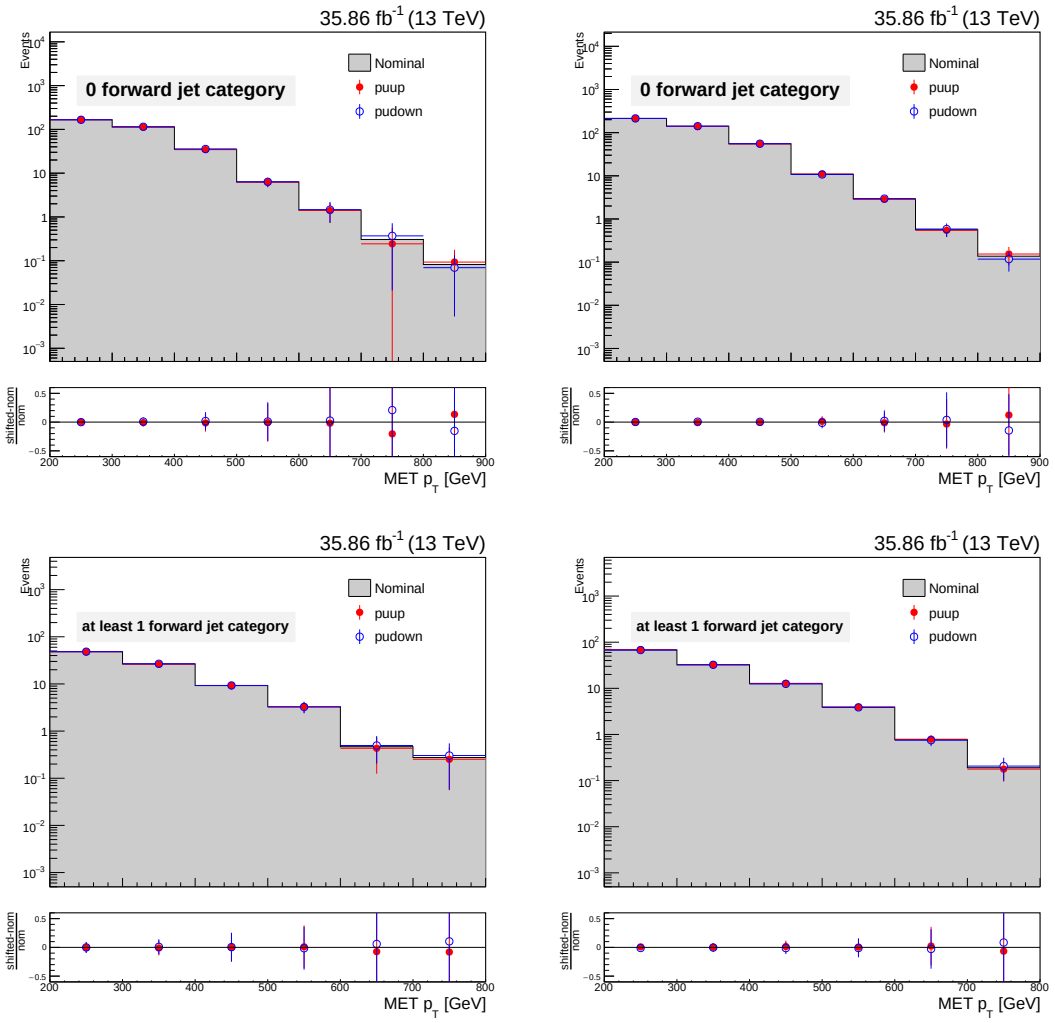


Figure A.6: Effect of pile up uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the $\mathrm{t\bar{t}}$ sample estimated from data, while second one is for $\mathrm{t\bar{t}}$ taken from MC simulation.

Parton Distribution Function (PDF)

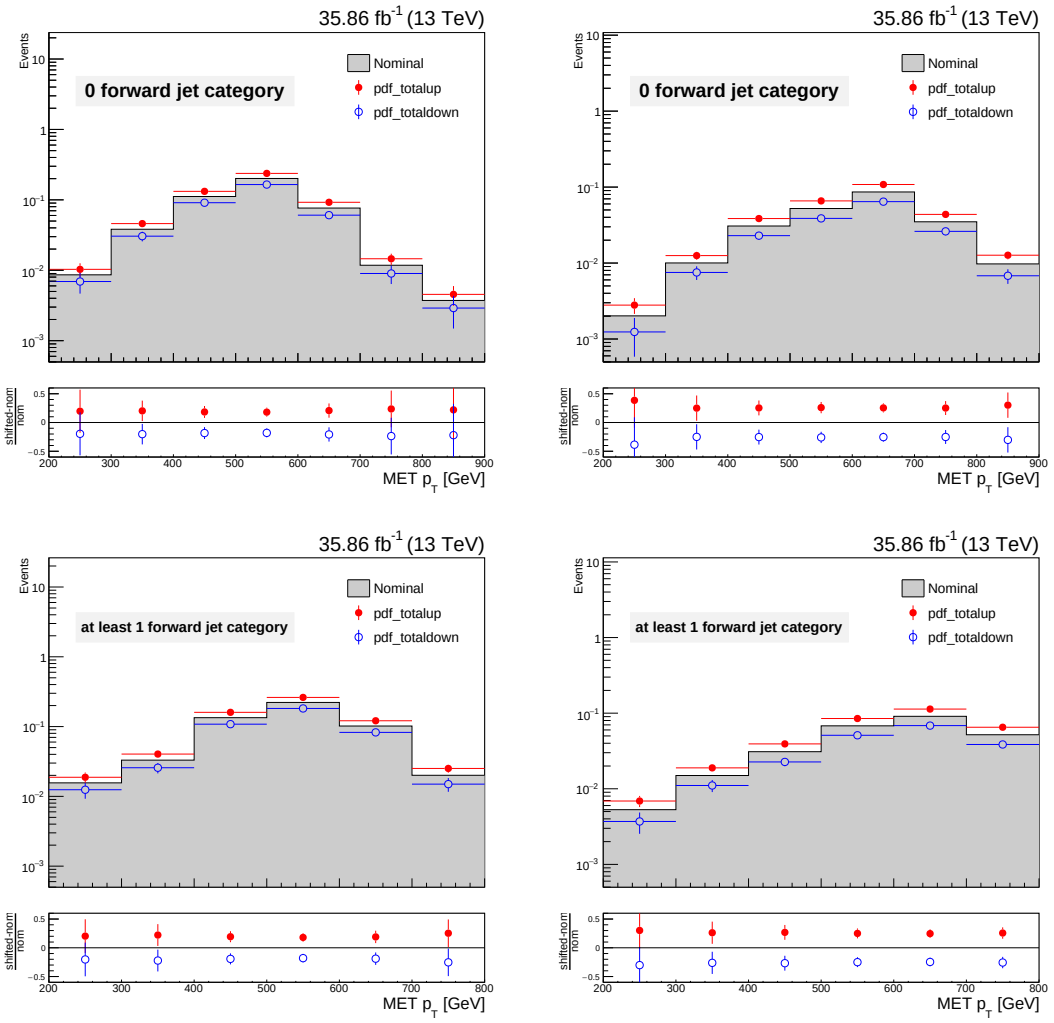


Figure A.7: Parton distribution function uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the signal sample with T mass hypothesis of 1200 GeV, while second one is for 1400 GeV.

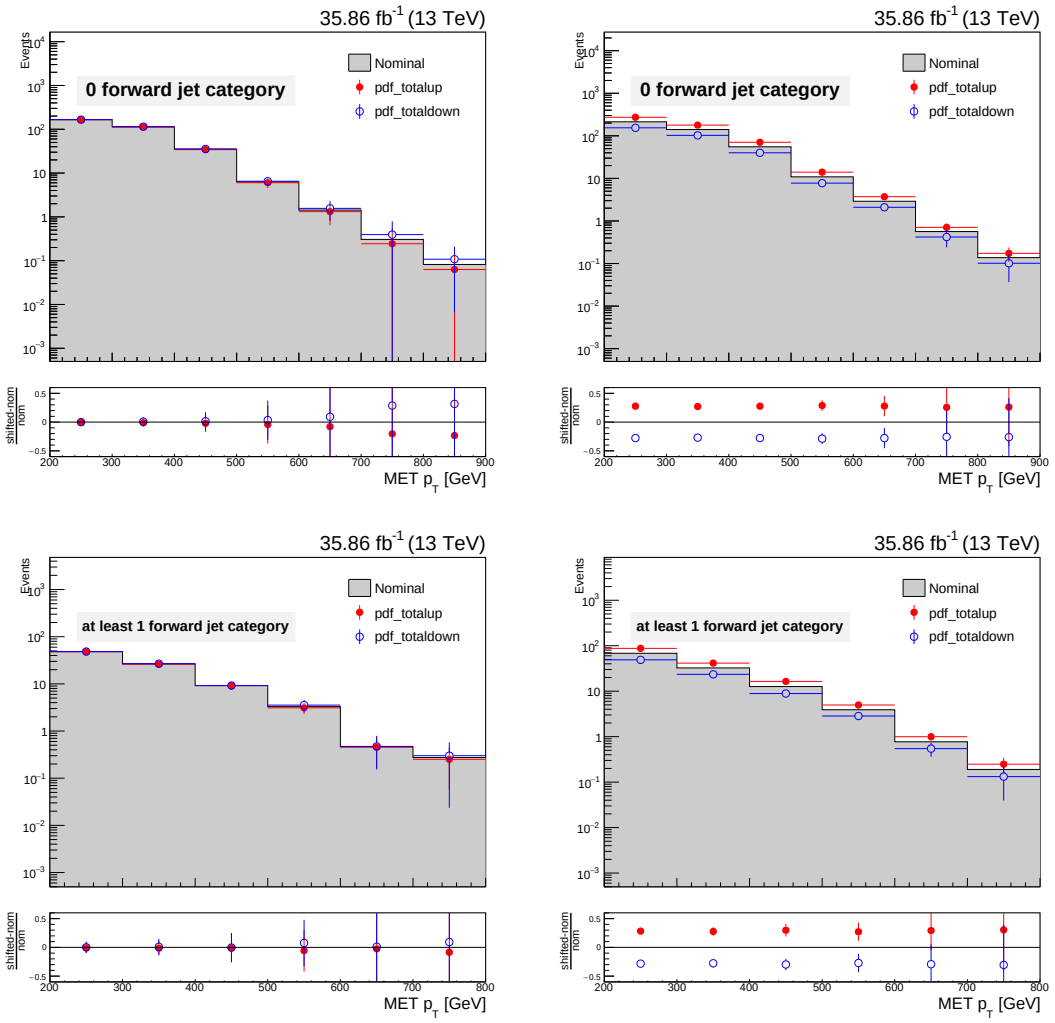


Figure A.8: Parton distribution function uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the $\mathrm{t\bar{t}}$ sample estimated from data, while second one is for $\mathrm{t\bar{t}}$ taken from MC simulation.

Factorisation and renormalisation scales (q^2)

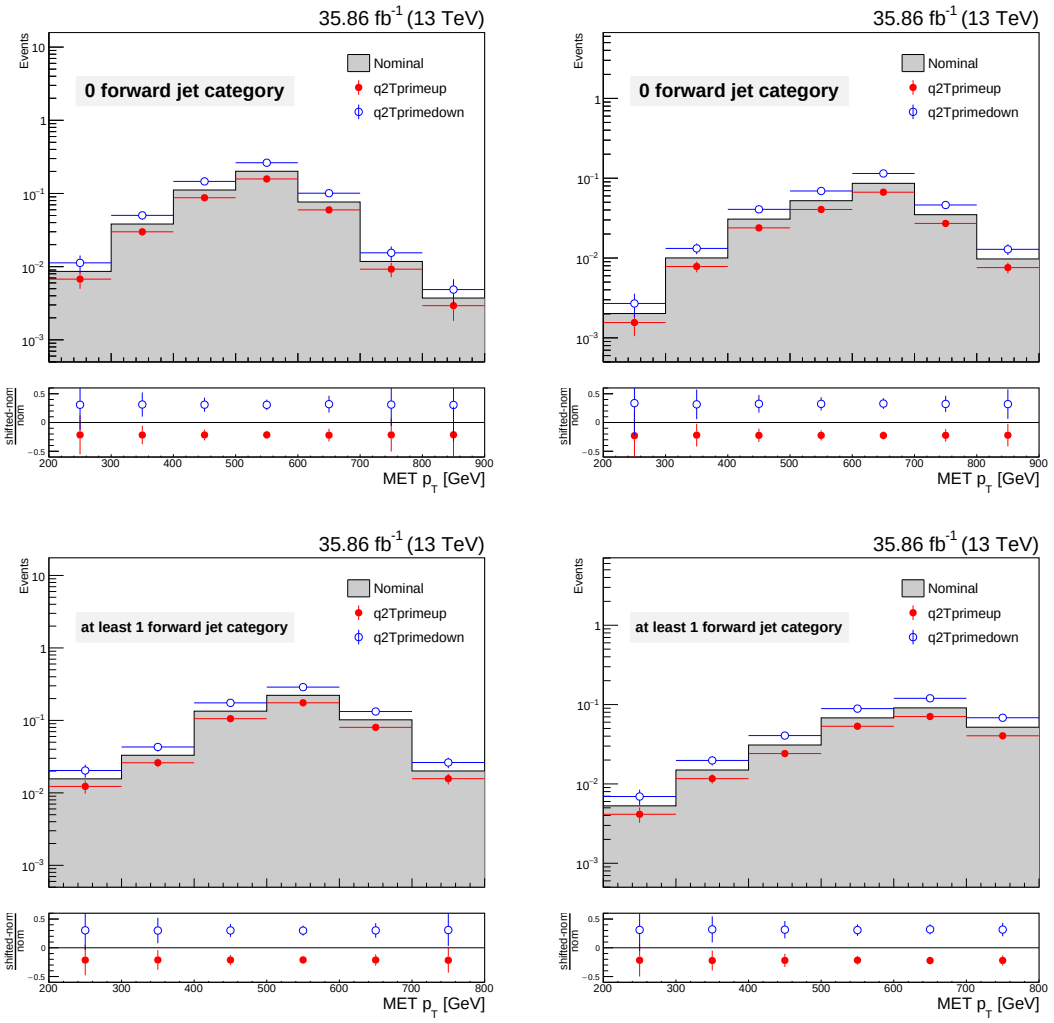


Figure A.9: Renormalization and factorization scale uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the signal sample with T mass hypothesis of 1200 GeV, while second one is for 1400 GeV.

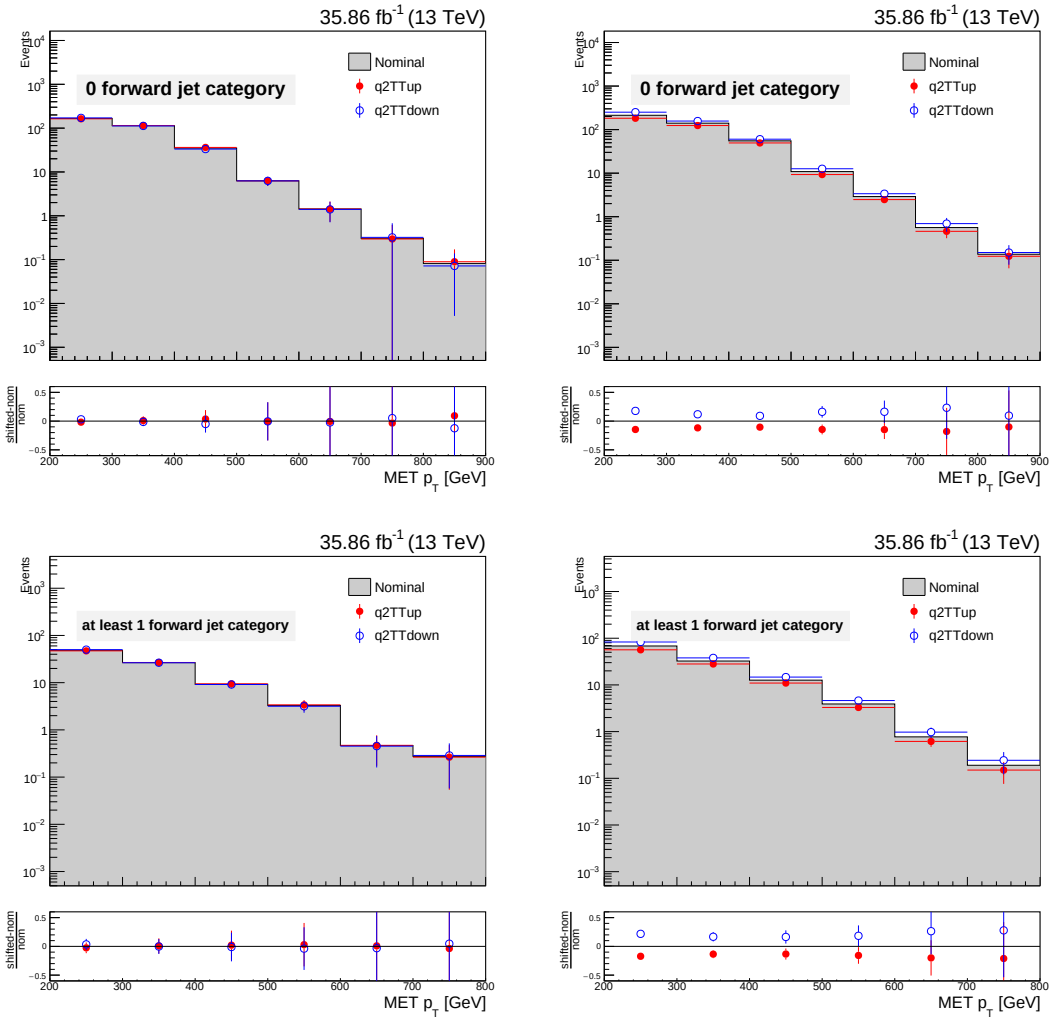


Figure A.10: Renormalization and factorization scale uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the $\mathrm{t\bar{t}}$ sample estimated from data, while second one is for $\mathrm{t\bar{t}}$ taken from MC simulation.

top tagging scale factors

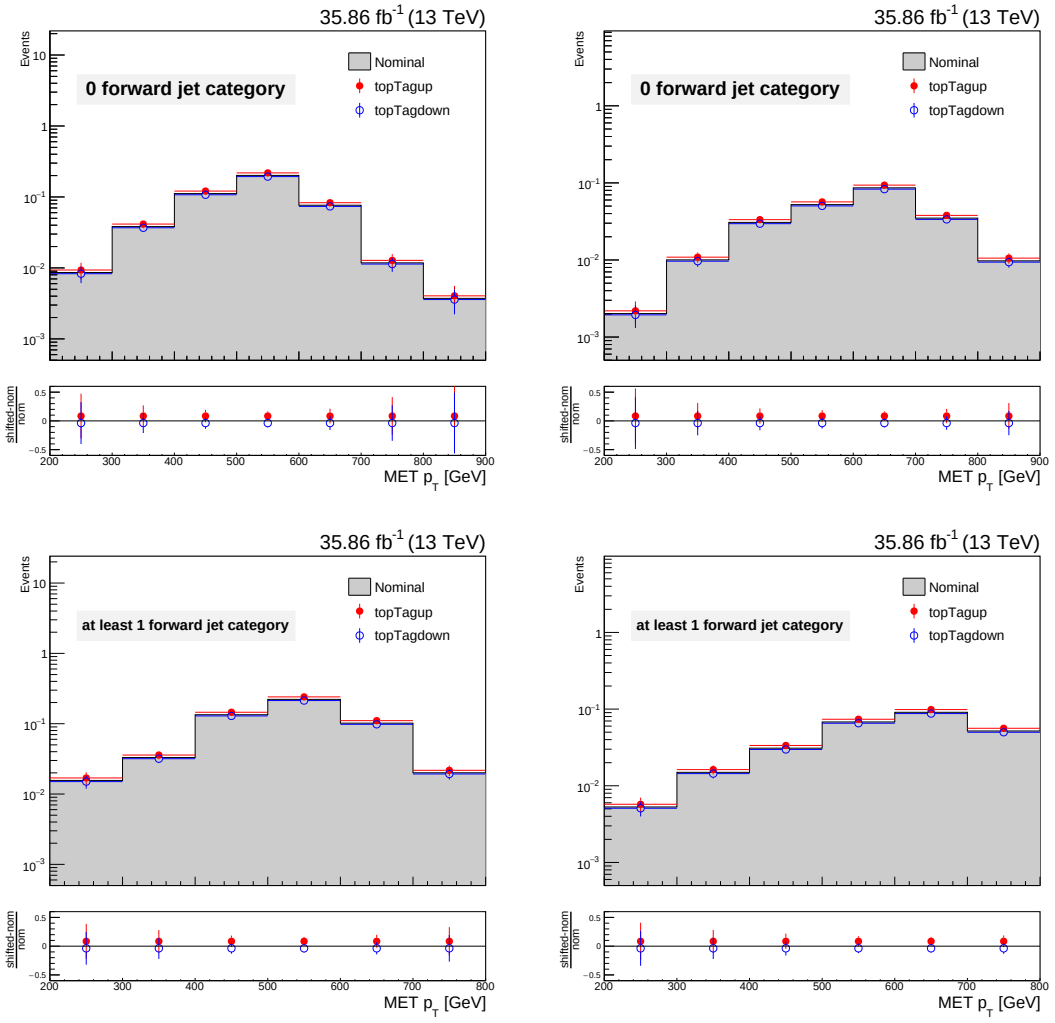


Figure A.11: Effect of top Tag scale factor uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the signal sample with T mass hypothesis of 1200 GeV, while second one is for 1400 GeV.

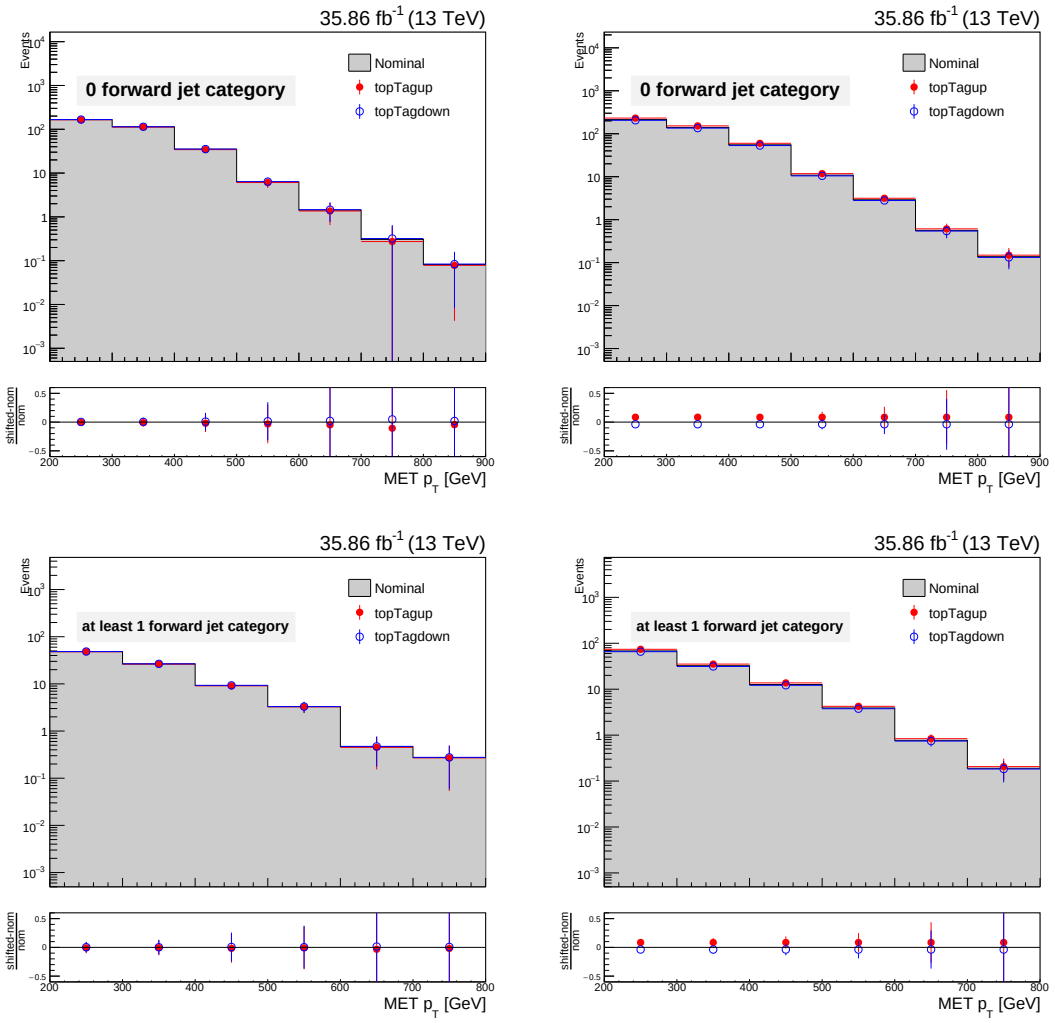


Figure A.12: Effect of top Tag scale factor uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the $\mathrm{t\bar{t}}$ sample estimated from data, while second one is for $\mathrm{t\bar{t}}$ taken from MC simulation.

W tagging scale factors

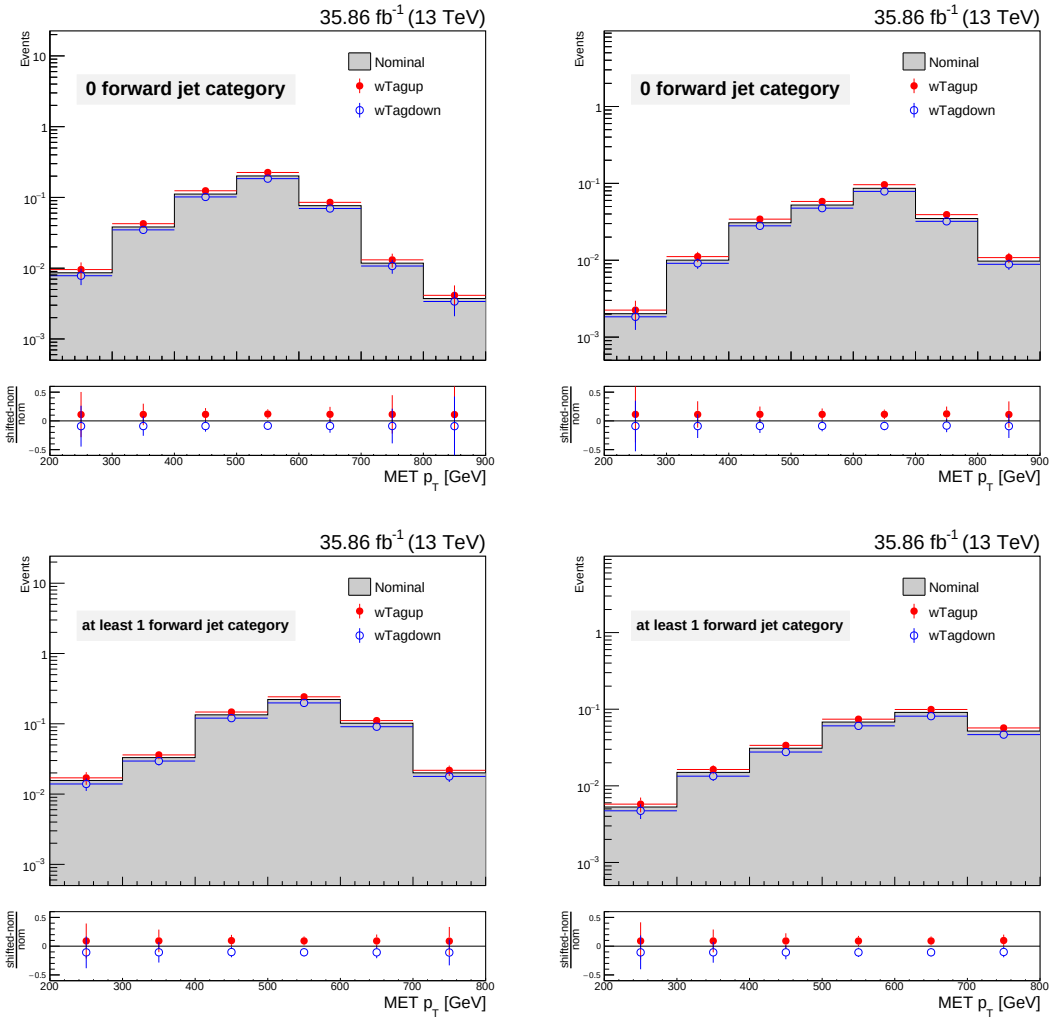


Figure A.13: Effect of wTag scale factor uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the signal sample with T mass hypothesis of 1200 GeV, while second one is for 1400 GeV.

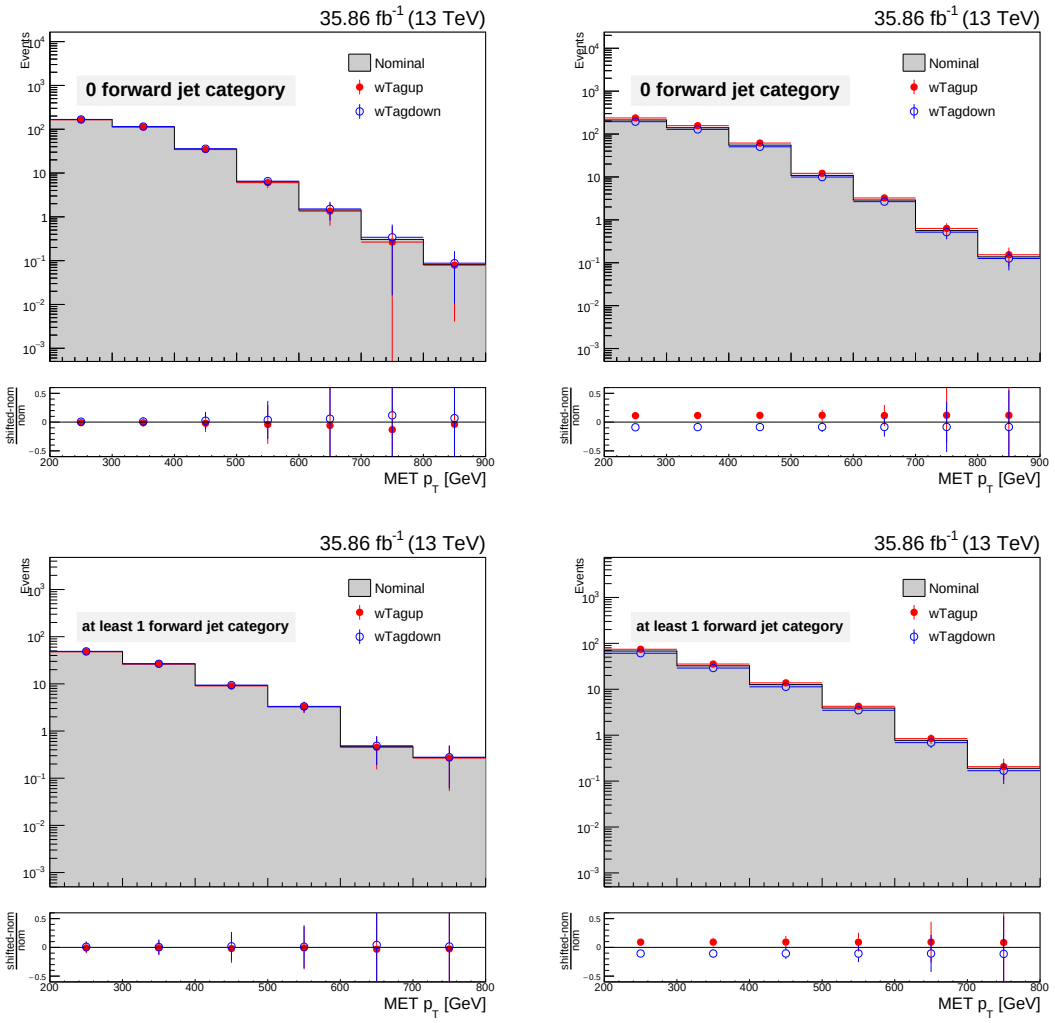


Figure A.14: Effect of w Tag scale factor uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the $\mathrm{t\bar{t}}$ sample estimated from data, while second one is for $\mathrm{t\bar{t}}$ taken from MC simulation.

Trigger scale factor

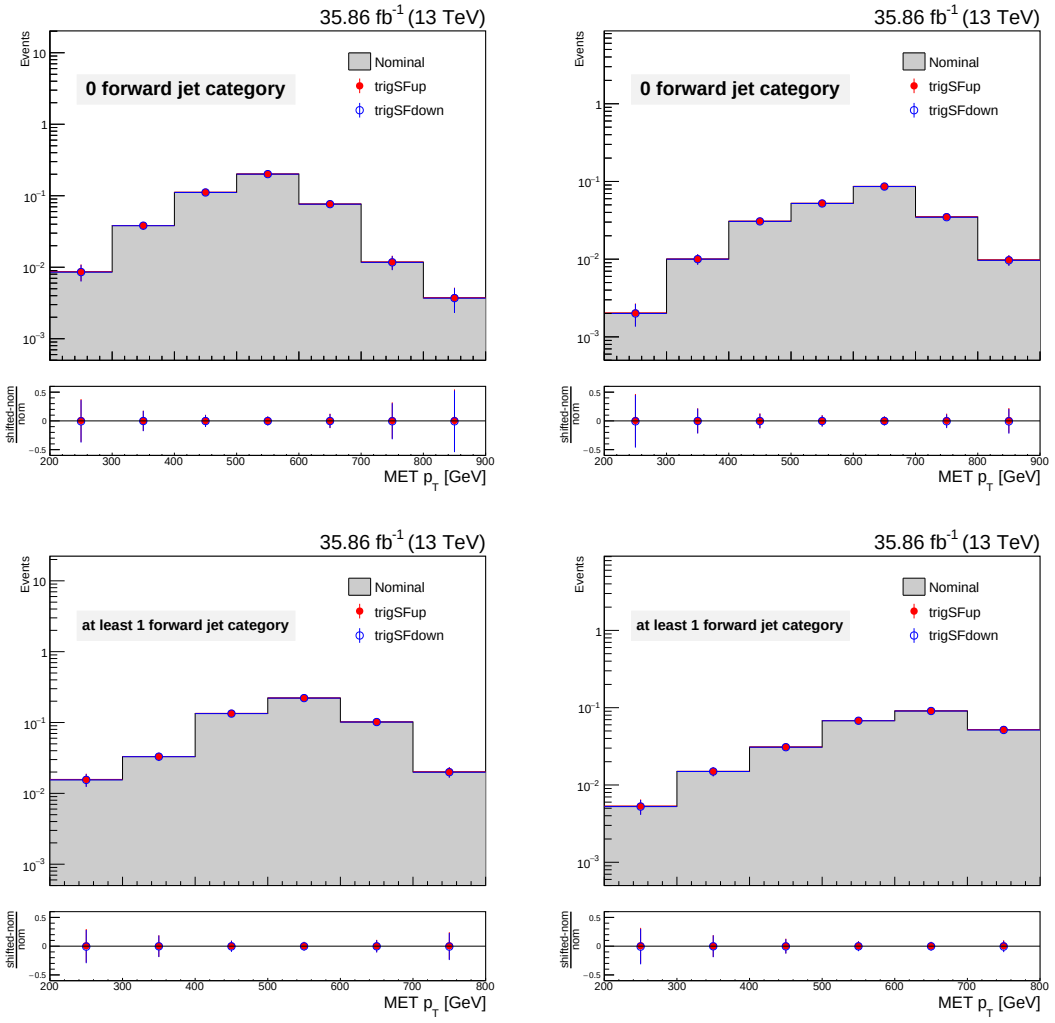


Figure A.15: Trigger scale factor uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the signal sample with T mass hypothesis of 1200 GeV, while second one is for 1400 GeV.

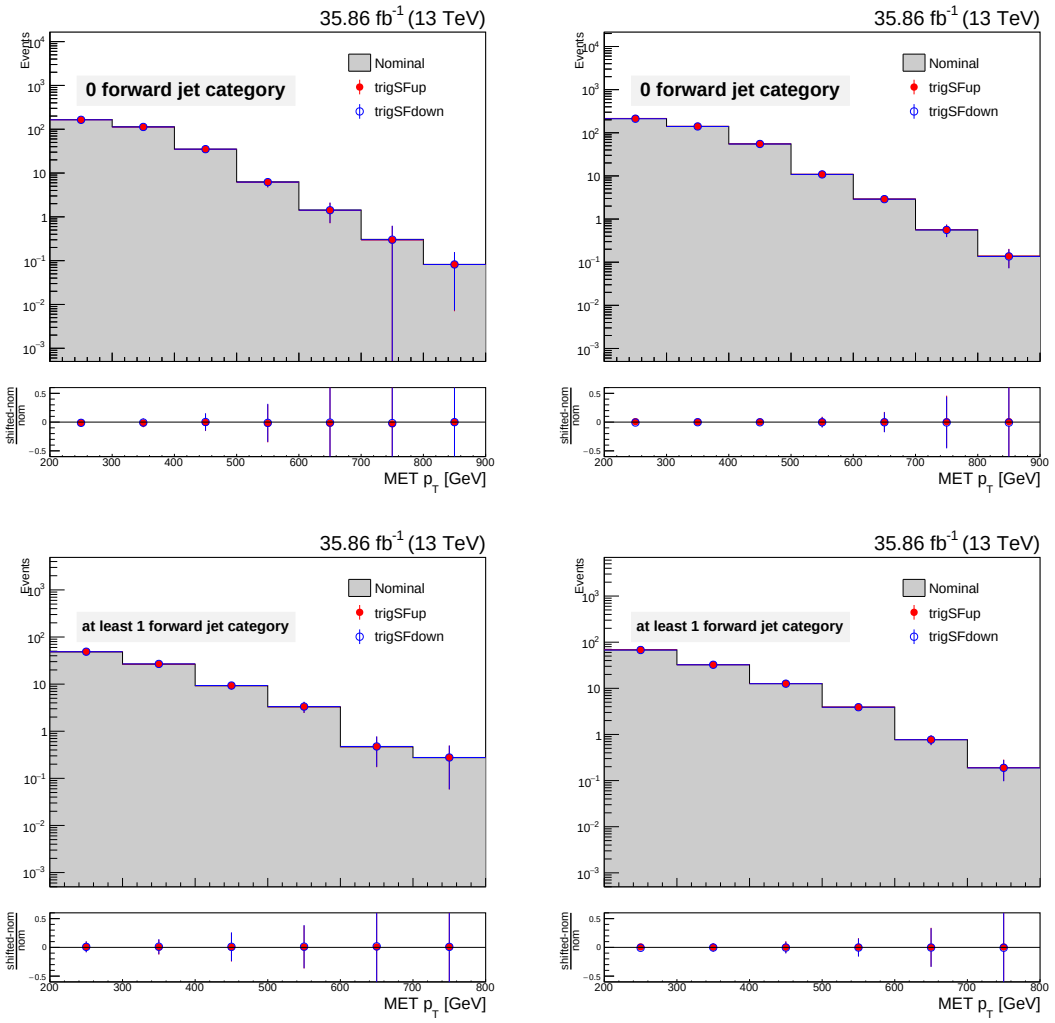


Figure A.16: Trigger scale factor uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the $\mathrm{t\bar{t}}$ sample estimated from data, while second one is for $\mathrm{t\bar{t}}$ taken from MC simulation.

jes

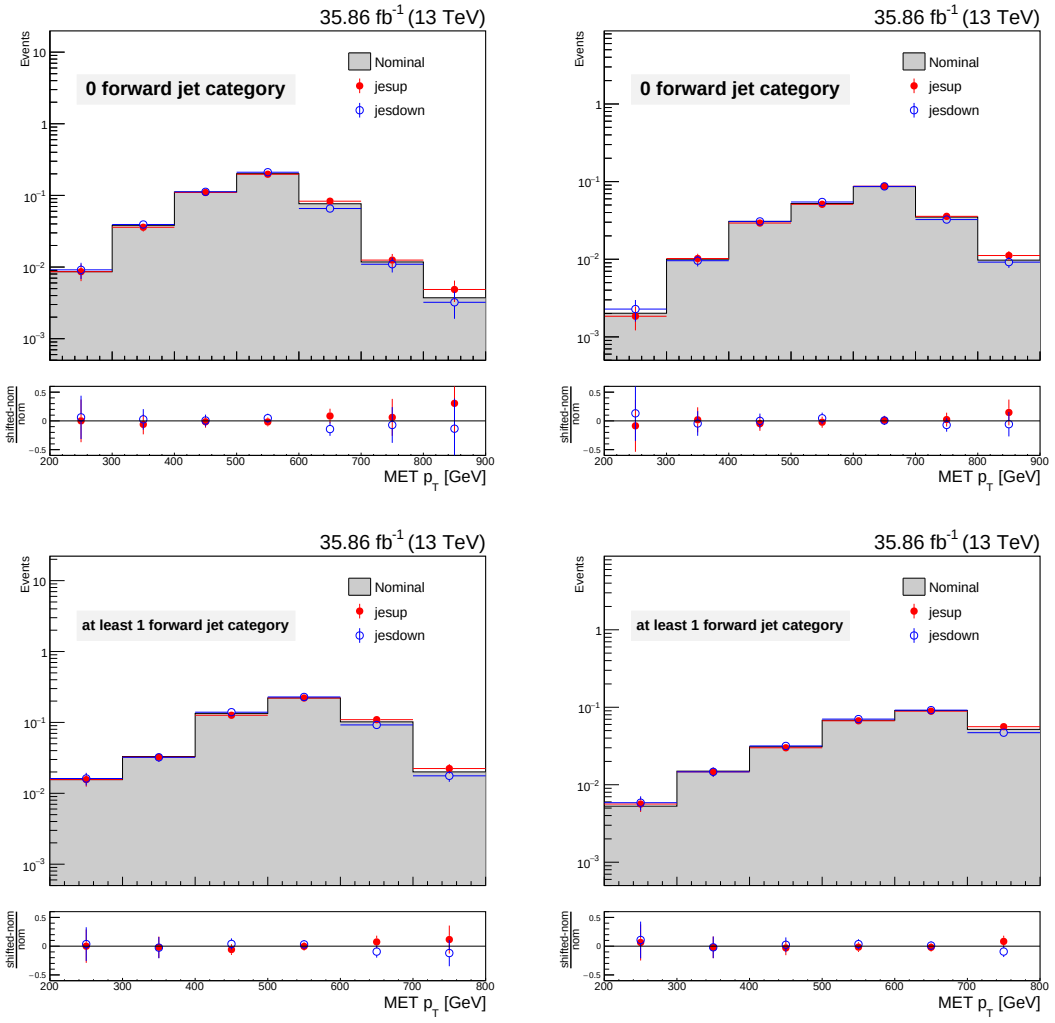


Figure A.17: Jet energy scale uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the signal sample with T mass hypothesis of 1200 GeV, while second one is for 1400 GeV.

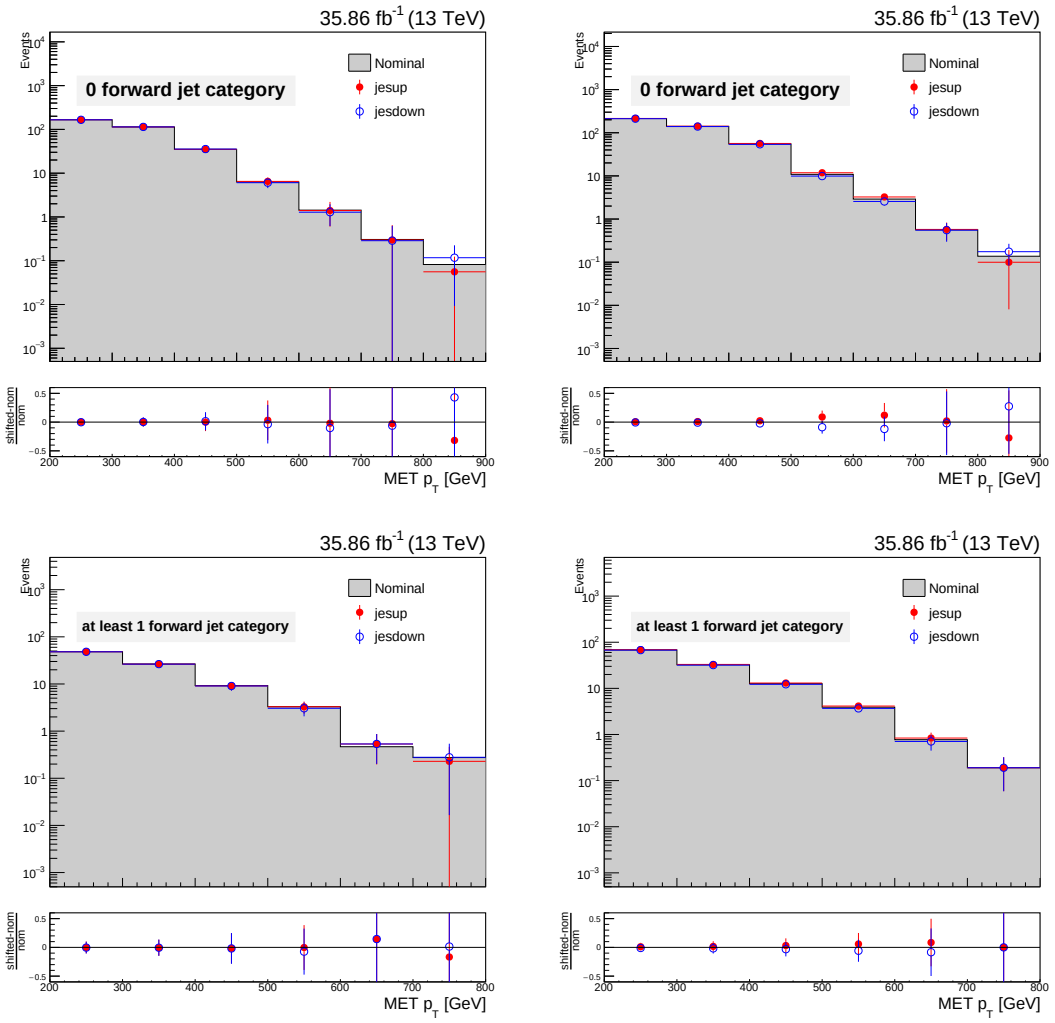


Figure A.18: Jet energy scale uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the $\mathrm{t\bar{t}}$ sample estimated from data, while second one is for $\mathrm{t\bar{t}}$ taken from MC simulation.

jer

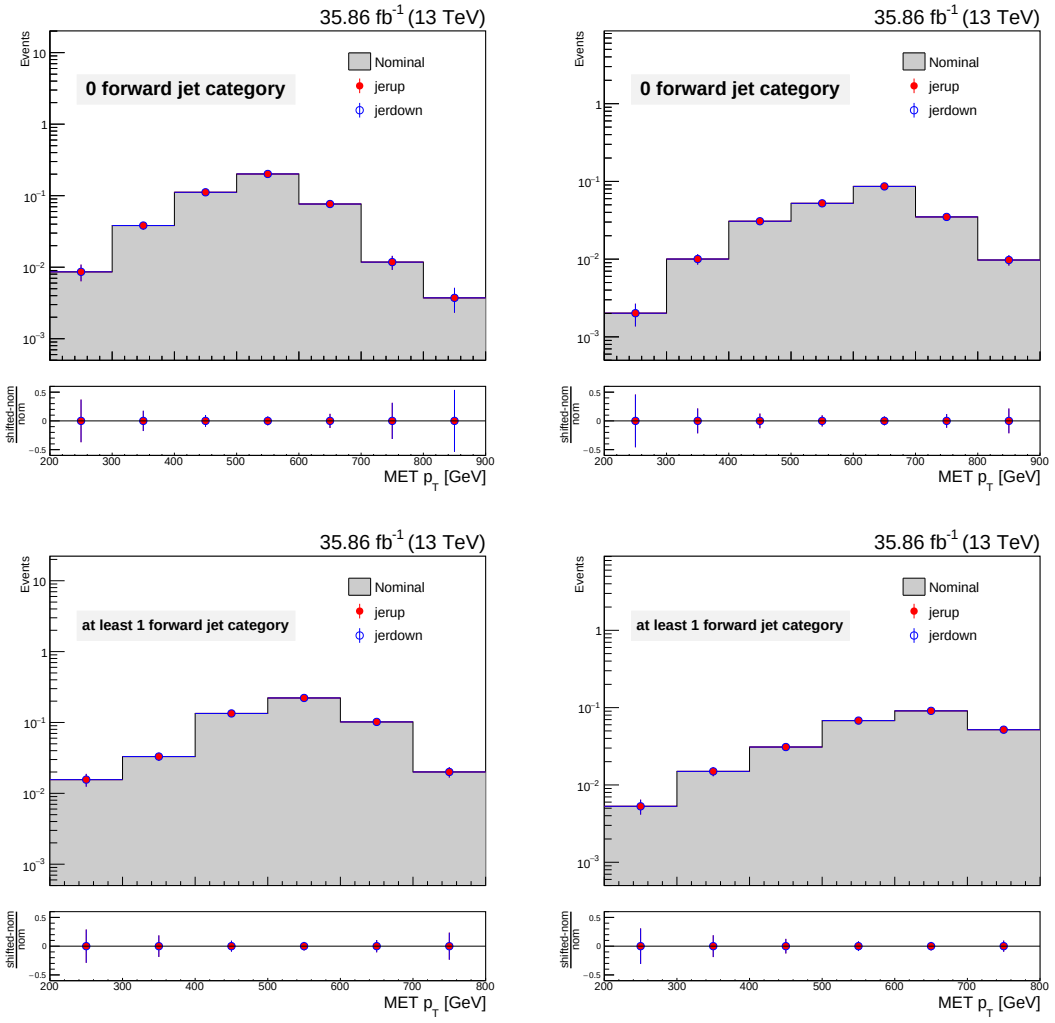


Figure A.19: Jet energy resolution uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the signal sample with T mass hypothesis of 1200 GeV, while second one is for 1400 GeV.

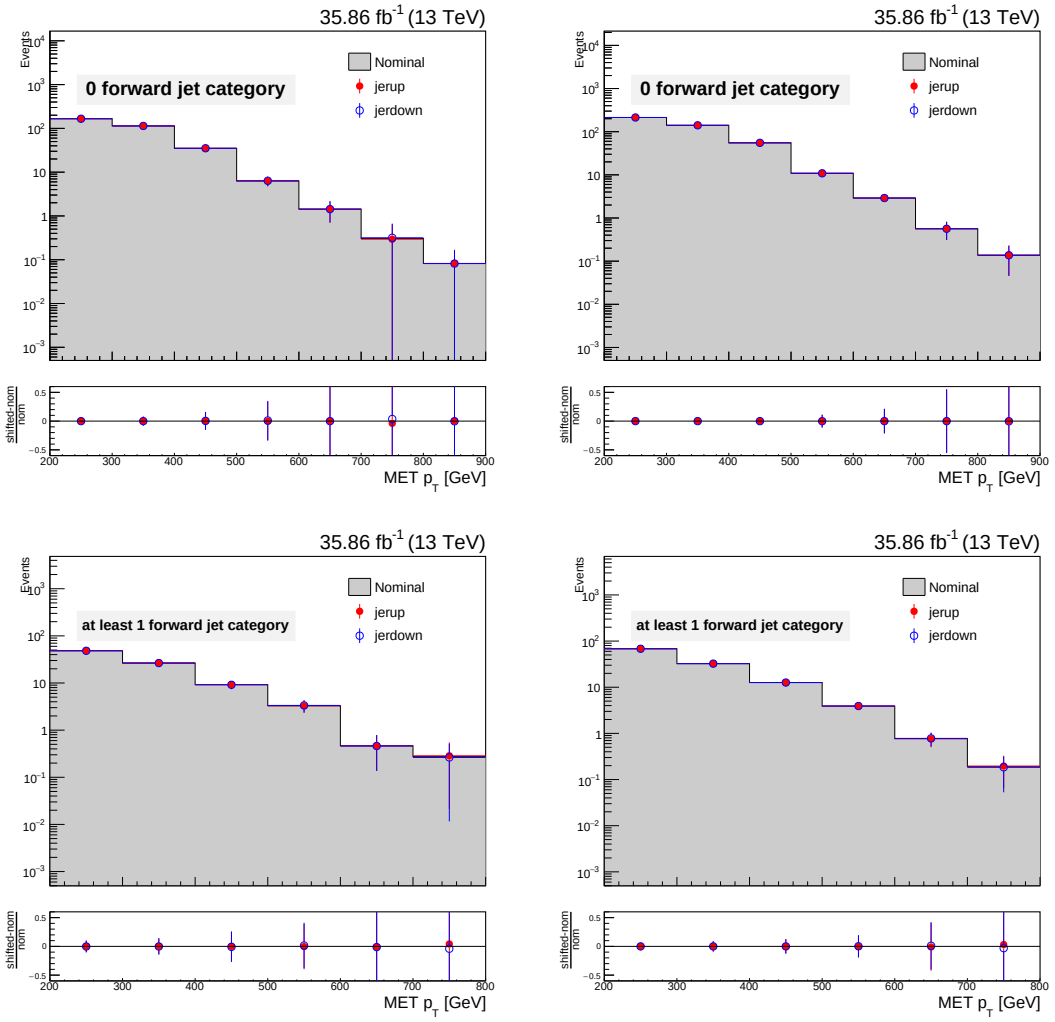


Figure A.20: Jet energy resolution uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the $\mathrm{t\bar{t}}$ sample estimated from data, while second one is for $\mathrm{t\bar{t}}$ taken from MC simulation.

Top p_{T} reweighting

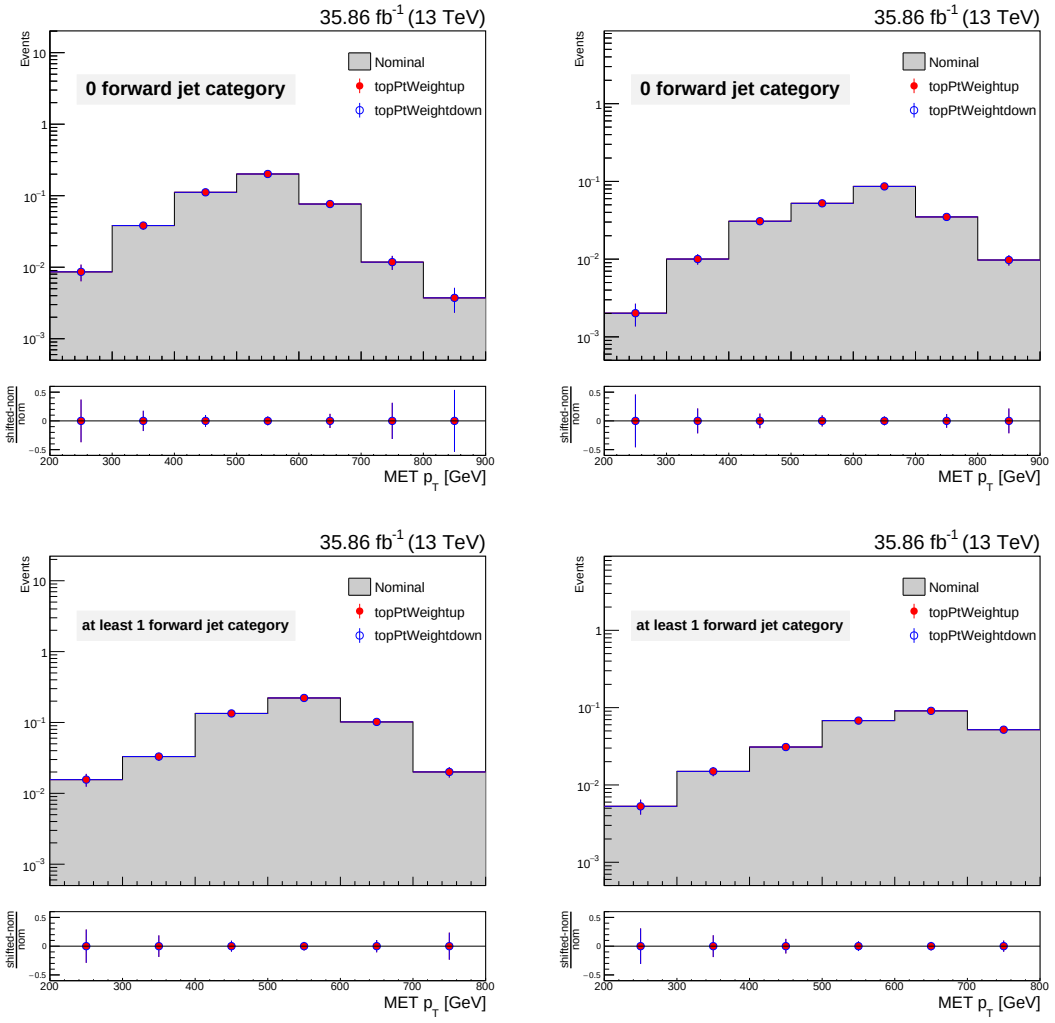


Figure A.21: Top p_{T} Weight uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the signal sample with T mass hypothesis of 1200 GeV, while second one is for 1400 GeV.

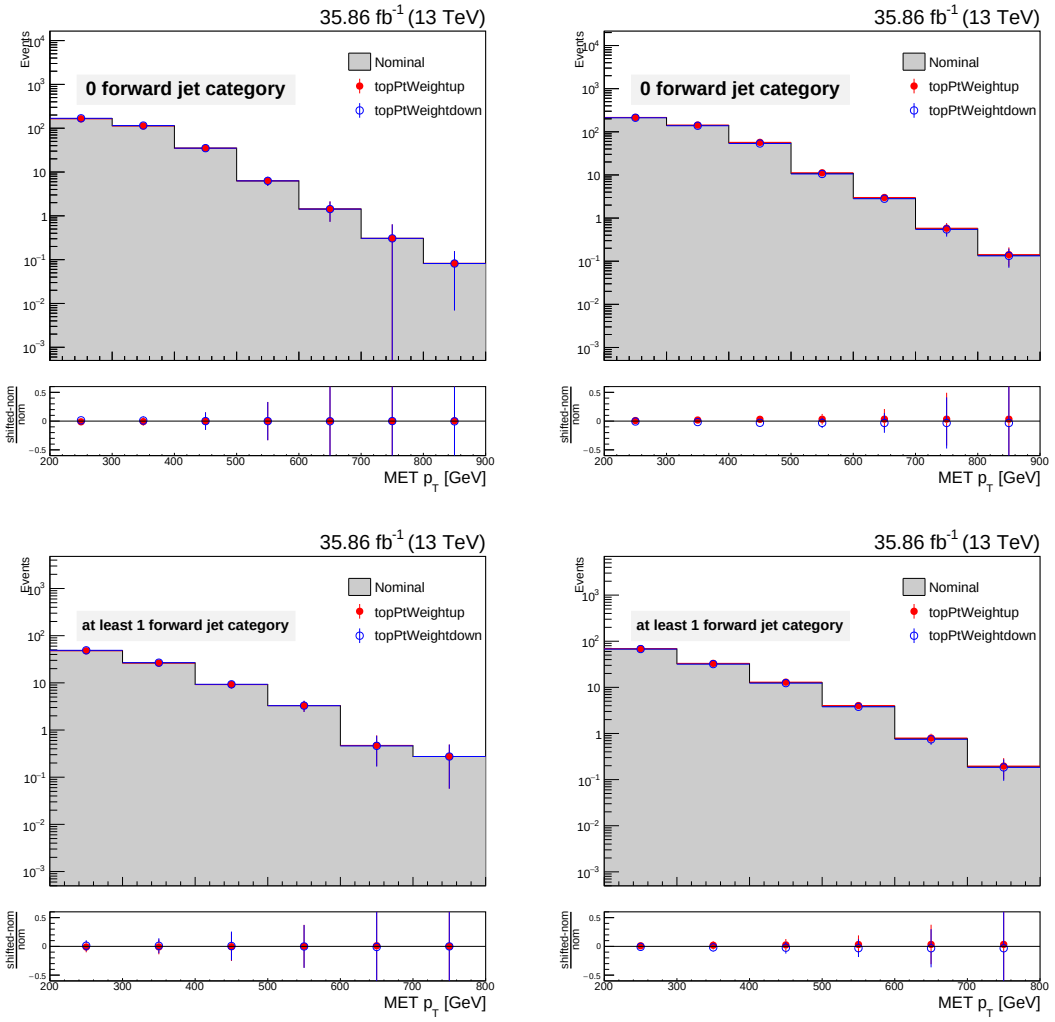


Figure A.22: Top p_{T} Weight uncertainty shift up and down with respect to nominal MET p_{T} distribution in the 0 (upper row) and at least 1 (lower row) forward jet category. The first column shows the effect on the $\mathrm{t\bar{t}}$ sample estimated from data, while second one is for $\mathrm{t\bar{t}}$ taken from MC simulation.

Acknowledgements

It has been hard arrive at this goal, but if for me has been hard for people that have been by my side has been harder, for that I need to spend some lines to thank them all because without them this thesis might not have been written.
I would to like to thank Orso for welcoming me in the CMS Naples group, he encouraged me, paying attention, willingness, and really a lot of patience.
I would to like also to thank my other two supervisors Dr. Luca Lista and Prof. Fausto Guarino for their proficient advice given to me during this thesis.
I would like to thank Agostino for the support he gave to me for the whole duration of the thesis, for having been always ready to leave what he was doing to help me, for all the things he taught me, for his patience and for the laughs.
I would like to thank Dr. Wajid Ali Khan and Dr. Biagio Rossi to be a constant presence, for helping me to explain better what I was doing, and for sharing with me their tricks and their funny stories.
I would like to thank my family for supporting me and inciting me to do my best. They have been teach to me the values of the culture, of the study, of the work and of the perseverance.
I would like to thank Fabio for taking care of me in the last years, being by my side on every day without being afraid of my folly and inciting me to became better and better.
I would like to thank Martina for sharing the unforgettable never ending days of hard and desperate study.
I would like to thank my historical friends to understand the reasons of my absences and to have been a secure harbour in the hard times and a happy island in the funny moments.
Many thanks to you guys for having been for me the best cheerleaders of ever.

Bibliography

[1] S. Glashow, "Partial-symmetries of weak interactions", Nuclear Physics, 22:579-588, (1961).

[2] S. Weinberg, "A model of leptons", Phys. Rev. Lett., 19:1264-1266, (1967).

[3] A. Salam, N. Svalthrom ed., "Elementary particle theory: relativistic groups and analyticity", Nobel symposium. (1968)

[4] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons", Phys. Rev. Lett., 13:321-323, (1964).

[5] P.W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons", Phys. Rev. Lett., 13:508-509, (1964).

[6] G.S. Guralnik, C.R. Hagen and T.W.B. Kibble, "Global Conservation Laws and Massless Particles", Phys. Rev. Lett., 13: 585–587, (1964).

[7] G. Zweig, "An SU(3) model for strong interaction symmetry and its breaking", Technical Report CERN-TH-401, CERN, Geneva, (1964).

[8] M.Y. Han and Y. Nambu, "Three-triplet model with double SU(3) symmetry", Phys. Rev., 139:B1006-B1010, (1965).

[9] O.W. Greenberg, "Spin and Unitary Spin Independence in a Paraquark Model of Baryons and Mesons", Phys. Rev. Lett., 13:598-602, (1964).

[10] H.D. Politzer, "Reliable perturbative results for strong interactions", Phys. Rev. Lett., 30:1346–1349, (1973).

[11] D.J. Gross and F. Wilczek, "Ultraviolet behavior of non-abelian gauge theories", Phys. Rev. Lett., 30:1343–1346, (1973).

[12] G. Arnison et al., Phys.Lett. 122B (1983) 103; 126B 398, (1983).

[13] M. Banner et al., Phys. Lett. 122B 476, (1983).

[14] P. Bagnaia et al., Phys. Lett. B129 130, (1983).

[15] Abe F. et al, CDF Collaboration, Phys Rev. D50 2966 and Phys. Rev. Lett. 73 225, (1994).

[16] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". Phys.Lett.B, (2012).

[17] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". Phys.Lett.B, (2012).

[18] P. Langacker, "Introduction to the Standard Model and Electroweak Physics", Phys.Lett. High Energy Physics-Phenomenology, (2009), http://arxiv.org/abs/0901.0241.

[19] B. Odom et al., "New Measurement of the Electron Magnetic Moment Using a One-Electron Quantum Cyclotron", Phys. Rev. Lett. 97 (3 2006) 030801, http://link.aps.org/doi/10.1103/PhysRevLett.97.030801

[20] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration, LEP Working Group for Higgs Boson Searches, Phys. Lett. B, 565 (2003), http://arXiv:hep-ex/0306033

[21] Tevatron, " New Phenomena and Higgs Working Group for the CDF and DØ Collaborations", FERMILAB-PUB-09-060-E,http://arXiv:0903.4001

[22] "LHC machine", Journal of Instrumentation 3 no.8, (2008) S08001. http://iopscience.iop.org/1748-0221/3/08/S08001.

[23] CMS collaboration, "The CMS tracker system project: technical design report", CERN-LHCC-98-006, http://cdsweb.cern.ch/record/368412.

[24] CMS collaboration, "The CMS tracker: addendum to the technical design report", CERN-LHCC-2000-016, http://cdsweb.cern.ch/record/490194.

[25] CMS collaboration, "CMS Technical Design Report for the Pixel Detector Upgrade", CERN-LHCC-2012-016, https://cds.cern.ch/record/1481838.

[26] CMS collaboration, "The CMS electromagnetic calorimeter project : Technical Design Report", CERN-LHCC-97-033, https://cds.cern.ch/record/349375.

[27] CMS collaboration, "The CMS hadron calorimeter project : Technical Design Report", CERN-LHCC-97-031, https://cds.cern.ch/record/357153.

[28] F. Kircher et al., "Final design of the CMS solenoid cold mass", IEEE Trans. Appl. Supercond., 10 (2000) 407.

[29] CMS collaboration, "The TriDAS project, technical design report. Volume 1: The level-1 trigger", CERN-LHCC-2000-038, http://cdsweb.cern.ch/record/706847.

[30] CMS collaboration, "The TriDAS project, technical design report. Volume 2: Data acquisition and high-level trigger technical design report", CERN-LHCC-2002-026, http://cdsweb.cern.ch/record/578006.

[31] Luciano Maiani, "The GIM Mechanism: origin, predictions and recent uses", https://arxiv.org/abs/1303.6154.

[32] D.B. Kaplan and H. Georgi, Phys. Lett. B136 (1984).

[33] D.B. Kaplan, Nucl. Phys B365 (1991).

[34] H. Georgi, L. Kaplan, D. Morin, and A. Schenk,"Effects of top compositeness",Phys. Rev. D51(1995) 3888,https://arxiv.org/abs/hep-ph/9410307.

[35] T. Aaltonen et al.[CDF Collaboration],"Search for a Heavy Top-Like Quark in pp Collisions at $\sqrt{s} = 1.96$ TeV", Phys. Rev. Lett. 107 (2011) 261801 https://arxiv.org/pdf/1107.3875.pdf

[36] CDF public note:https://arxiv.org/pdf/www-cdf.fnal.gov/physics/new/top/2010/tprop/HQ_public/HQpub.pdf

[37] V. M. Abazov et al. [D0 Collaboration],"Search for single vector-like quarks in pp collisions at $\sqrt{s} = 1.96$ TeV", Phys. Rev. Lett. 106 (2011) 081801 https://arxiv.org/pdf/1010.1466.pdf.

[38] CMS Collaboration, "Search for heavy vector-like quarks decaying to same-sign dileptons", http://cds.cern.ch/record/2256747?ln=en

[39] CMS Collaboration, "Search for top quark partners with charge 5/3 in the single-lepton final state at $\sqrt{s} = 13$ TeV",http://cds.cern.ch/record/2264686?ln=en.

[40] ATLAS Collaboration, "Search for pair production of up-type vector-like quarks and for four-top-quark events in final states with multiple b-jets with the ATLAS detector",https://arxiv.org/abs/1803.09678

[41] "Search for pair production of heavy vector-like quarks decaying to high-p_T W bosons and b quarks in the lepton-plus-jets final state in pp collisions at $\sqrt{s} =$ 13 TeV with the ATLAS detector", https://arxiv.org/abs/1707.03347.

[42] ATLAS Collaboration,"Search for pair production of vector-like top quarks in events with one lepton, jets, and missing transverse momentum in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector",https://arxiv.org/abs/1705.10751.

[43] CMS Collaboration, "Search for pair production of vector-like quarks in the $bW\bar{b}W$ channel from proton-proton collisions at $\sqrt{s} = 13$ TeV",https://arxiv.org/abs/1710.01539.

[44] CMS Collaboration, " Search for single production of a vector-like T quark decaying to a Z boson and a top quark in proton-proton collisions at $\sqrt{s} = 13$ TeV", https://arxiv.org/abs/1708.01062.

[45] CMS Collaboration, "Search for a singly produced vector-like quark B decaying to a b quark and a Higgs boson in a fully hadronic final state using boosted topologies ",https://cds.cern.ch/record/2273909.

[46] Mihailo Backovi et al.,"Discovering heavy new physics in boosted Z channels:$Z \to l^+l^-$ vs $Z \to \nu\bar{\nu}$",https://arxiv.org/abs/1501.07456

[47] CMS Collaboration, "Particle–Flow Event Reconstruction in CMS and Performance for Jets, Taus, and $\not{E}_{\mathrm{T}}$", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, (2009).

[48] CMS Collaboration, "Electron cut-based Identification",https://twiki.cern.ch/twiki/bin/view/CMSCutBasedElectronIdentificationRun2#Electron_.

[49] CMS Collaboration, "Baseline Muon Selection", https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideMuonIdRun2#Tight_Muon.

[50] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_{T} jet clustering algorithm", JHEP 04 063, https://arxiv.org/abs/0802.1189, (2008)

[51] CMS Collaboration, "Jet Performance in pp Collisions at $\sqrt{s} = 7$ TeV", CMS Physics Analysis Summary CMS-PAS-JME-10-003, CERN, 2010.

[52] "Particle Flow Jet ID",https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID.

[53] T. C. collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", Journal of Instrumentation 6, no. 11, P11002, (2011).

[54] M. Dasgupta, A. Fregoso, S. Marzani, and G. P. Salam, "Towards an understanding of jet substructure", JHEP 09 (2013) 029, https://arxiv.org/abs/1307.0007.

[55] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft Drop", JHEP 05 (2014) 146, https://arxiv.org/abs/1402.2657.

[56] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, "Techniques for improved heavy particle searches with jet substructure", Phys. Rev. D 80 (2009) 051501, https://arxiv.org/abs/0903.5081.

[57] CMS Collaboration, "Identification techniques for highly boosted W bosons that decay into hadrons", JHEP 12 (2014) 017, https://arxiv.org/abs/1410.4227.

[58] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", JINST 8 (2013) P04013,https://arxiv.org/abs/1211.4462.

[59] CMS Collaboration, "Identification of b quark jets at the CMS experiment in the LHC Run 2", CMS Physics Analysis Summary CMS-PAS-BTV-15-001, 2016.

[60] CMS Collaboration, "Identification of b quark jets at the CMS Experiment in the LHC Run 2", CMS Physics Analysis Summary CMS-PAS-BTV-15-001, CERN, http://cds.cern.ch/record/2138504, (2015).

[61] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", JHEP 1407 (2014) 079, https://arxiv.org/abs/1405.0301.

[62] T. Sjostrand, "PYTHIA 6.4 Physics and Manual", 2006, https://arxiv.org/pdf/hep-ph/0603175.pdf.

[63] M. Bähr et al., "Herwig++ physics and manual", Eur. Phys. J. C 58 (2008) 639,https://arxiv.org/abs/0803.0883.

[64] S. Frixione et al., "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", (2007),https://arxiv.org/abs/0709.2092v1.

[65] S. Agostinelli et al., "Geant4. A simulation toolkit", Nuclear Instruments and Methods in Physics Research Section A: Ac- celerators, Spectrometers, Detectors and Associated Equipment, no. 3, volume 506 (pag. 250−303), 2003, https://www.sciencedirect.com/science/article/pii/S0168900203013688.

[66] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms", JHEP 0411 (2004) 040,https://arxiv.org/abs/hep-ph/0409146.

[67] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", JHEP 1006 (2010) 043,https://arxiv.org/abs/1002.2581.

[68] T. Sjostrand, S. Mrenna, and P. Z. Skands, "A Brief Introduction to PYTHIA 8.1", Comput.Phys. Commun. 178 (2008) 852–867, https://arxiv.org/abs/0710.3820.

[69] CMS Collaboration, "Summary table of samples produced for the 1 Billion campaign, with 25 ns bunch-crossing", https://twiki.cern.ch/twiki/bin/viewauth/CMS/SummaryTable1G25ns.

[70] CMS Collaboration, "Cms luminosity based on pixel cluster counting – summer 2013 update", CMS Physics Analysis Summary CMS-PAS-LUM-13-001, https://arxiv.org/abs/1012.2466., (2013).

[71] M. Botje et al., "The PDF4LHC Working Group interim recommendations", https://arxiv.org/abs/1101.0538, (2011).

[72] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", JINST 6 P11002, https://arxiv.org/abs/1107.4277.

[73] L. Lista, "Practical Statistics for Particle Physicists", https://arxiv.org/pdf/1609.04150.pdf

[74] A. L. Read., "Modified Frequentist Analysis of Search Results (The CLs Method)", CERN OPEN, 2000-205, https://cds.cern.ch/record/451614/files/open-2000-205.pdf.

[75] T. Junk, "Confidence level computation for combining searches with small statistics", Nucl. Instrum. Meth. A 434 (1999) 435, https://arxiv.org/abs/hep-ex/9902006.

[76] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", Eur. Phys. J. C 71 (2011) 1554, https://arxiv.org/abs/1007.1727.

[77] CMS Collaboration, "Documentation of the RooStats-based statistics tools for Higgs PAG (rev. 106)", CMS twiki page, (2013), https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideHiggsAnalysisCombinedLimit?rev=106.

[78] A. Carvalho, S. Moretti, et al., "Single production of vector-like quarks with large width at the Large Hadron Collider", https://arxiv.org/abs/1805.06402v1.

[79] CMS Collaboration, "Search for single production of vector-like quarks decaying to a b quark and a Higgs boson", https://arxiv.org/abs/1802.01486.