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 new physics in final state with top quark plus invisible particles at the CMS experiment with ML techniques

Relatore:
Prof. Alberto Orso Maria Iorio
Correlatore:
Prof.ssa Antonia Di Crescenzo

Candidato:
Antimo Cagnotta
Matr. N94000633

Anno Accademico 2020/2021

Alla mia famiglia.

Contents

Introduction

The Standard Model of Particle Physics (SM) is the theory that most accurately describes elementary particles and their interactions. All the predictions of the SM have been experimentally verified, and in particular all the particles it predicted have been discovered, the last one being the Higgs Boson in 2012. Despite its successes, the SM is not a complete theory, as it does not explain phenomena such as the gravitational interaction, the masses of the neutrinos, the hierarchy problem or the existence of dark matter and dark energy.

In order to overcome its shortcomings, several theories have been proposed to extend the SM, such as supersymmetry, extra dimension, or composite Higgs. A common feature of many of these theories is the prediction of the existence of Vector-Like Quarks (VLQs) with masses at TeV scale. The VLQs, differently from the SM quarks, are characterized by a chirality symmetry under the electroweak interactions.

Another class of models aims at explaining the existence of Dark Matter (DM) as measurable predictions at the LHC. The DM takes its name from the fact that it does not appear to interact with the electromagnetic field, which means it does not absorb, reflect, or emit electromagnetic radiation, if it exists, it must barely interact with ordinary baryonic matter and radiation through interactions with small couplings like for example gravity and/or weak force. Some models predict the interaction beetwen SM particles and DM allowing the production of DM particles in a proton-proton collision.

In this work is presented the search for a final state common to both VLQ and DM models, characterized by the presence of a single top quark in association with invisible particles. In the case of VLQ, this final state is generated via resonant production of a single VLQ T, in the case of DM models, it is produced in association with a scalar mediator. The considered final state is characterized by the hadronic decay of the top quark, and the Z decay into a couple of neutrinos or the couple production of DM particles thorough a boson mediator.

In the selection of the event with the final state of interest a Machine Learning approch is proposed for the top quark candidate reconstruction. The Machine Learning is used to improve the selection of the final state conteining a single top quark, the hadronic top quark decay can be reconstructed in three different and complementary ways, the ML approch can facilitate the choice of the best one.

The search is performed with the Compact Muon Solenoid (CMS) experi-

ment, one of the four experiments at the Large Hadron Collider (LHC), that thanks to its high centre-of-mass energy of $\sqrt{s} = 13$ TeV and the instantaneous luminosity of 10^{34} cm^{-2}s^{-1}, could find evidences for VLQs and DM production.
The analysis is performed with the CMS data taking of 2018, with different mass hypothesis for VLQ T and for the boson that could mediate the interaction beetwen top quark and DM. The content of this thesis is organised in five chapters:

- Chapter 1, an introduction of the Standard Model and the open issues;
- Chapter 2, the description of the Large Hadron Collider and the Compact Muon Solenoid experiment;
- Chapter 3, a brief introduction on the theories that predict the existence of the VLQs and the DM, and an overview on the different current searches.
- Chapter 4, the definition of the final state searched for and the description of the objects observed in the detector;
- Chapter 5, the description of the data sets used, the event selection performed. Finally the fit procedure is described and the results are presented.

Chapter 1

Standard Model of particle physics

The Standard Model (SM) is an $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ quantum field theory that provides precise predictions of the fundamental particles and their interactions: *electromagnetic*, *weak*, and *strong*. The SM has been confirmed by the experimental evidence collected throughout the past century in particle physics.

The development of SM has began in the last century, and one of the first notable contribution came from Chen Ning Yan and Robert Mills[1], they extended the concept of gauge theory for abelian groups, e.g. quantum electrodynamics, to non abelian groups to provide an explanation for strong interactions. During the 60s and 70s, the description of strong interaction continued to develop further thanks to contributions from Murray Gell-Mann and George Zweig [2], who suggested the existence of quarks with different flavours, and Moo-Young Han with Yoichiro Nambu [3] and Oscar W. Greenberg[4], who suggested the existence of new quantum number called *colour* introducing the simmetry group $SU(3)_{col}$. The description of strong interactions in the SM reached its present form when, in 1973, David Politzer [5] and David Gross [6] together with Frank Wilczek suggested the theory of the asymptotic freedom of QCD.

In 1961, Sheldon Glashow [7] combined the electromagnetic and weak interactions, and the unification was ultimately formalized in 1967 by Steven Weinberg [8] and Abdus Salam [9] with the inclusion of the Higgs mechanism. Over the years, the SM received other contributions aimed at the inclusion of the strong interaction in the Glashow, Weinberg and Salam framework.

The main predictions of the SM were confirmed by measurements performed throughout the years, which notably include the discovery of the three massive bosons (W^+, W^- and Z) at CERN in 1983 by the UA1 and UA2 experiments [10–13], the discovery of the top quark at Tevatron in 1994 by the D0 and CDF experiments [14], and finally the discovery of the Higgs boson at CERN in 2012 by the ATLAS and CMS experiments[15, 16]. Thanks also to these compelling experimental evidences, the SM is widely accepted by the scientific community as the main model to describe fundamental interactions in particle physics. Nevertheless, it doesn't include the gravitational interaction, doesn't predict neutrino masses, that are confirmed by neutrino

Bosons	Interaction	Spin	Electric charge
γ	Electromagnetic	1	0
W^+, W^-, Z	Weak	1	$+1, -1, 0$
$8\ gluons(g)$	strong	1	0

Table 1.1: Standard Model mediator bosons

oscillation experiment, and the existence of dark matter and the dark energy,that is predicted by the ΛCDM cosmological model.

1.1 Standard Model overview

The dynamics equations for the SM are obtained from a gauge principle: a free particle lagrangian $\mathcal{L}$ is requested to be invariant under a local transformation of the symmetry group:

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

where,

- $U(1)_Y$ is the unitary group, the transformation is represented by a unitary scalar complex operator multiplied by its quantum number called the *weak hypercharge* (Y), and the associated vector field is called B_μ;
- $SU(2)_L$ is the special unitary group with $n = 2$, the trasformation is rappresented by the 2×2 Pauli matrices ($\sigma_{1,2,3}$) multiplied by the *weak isospin* (I_3), and the vector field are $W_\mu^{1,2,3}$;
- $SU(3)_C$ is the special unitary group with $n = 3$, the transformation is rappresented by the eight 3×3 Gell-Mann matrices ($\lambda_{1,\ldots,8}$), multiplied by the colour charge (r, g, b), and the eight vector field $G_\mu^{1,\cdots,8}$.

To each interaction an absolute conserved quantum number and a multiplet are associated. The components of such multiplet are called vector bosons because they obey the Bose-Einstein statistics and they have spin equals to one. They are also referred to as mediator bosons as they are considered the *mediators* of the respective interaction. The number of the mediator bosons is determined by the number of the generator of the symmetry group associated to each interaction. The SM has 12 vector fields associated to the three gauge symmetries. A list of the physical bosons is reported in table 1.1.

The SM contains 12 fields with an half-integer spin, that obey the Fermi-Dirac statistics and they are called fermions. Fermions are divided into lepton and quark fields, both are listed in three families or generations. Each generation is a doublet of particles associated to an isospin quantum number and each particle carries electroweak charge, and only the quarks have strong charge. In table 1.2,1.3 a list of SM leptons and quarks is reported.

Particles	Spin	Electric charge
$\begin{pmatrix}\nu_e\\ e\end{pmatrix} \begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix} \begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix}$	$\frac{1}{2}$	0 $-\frac{1}{2}$

Table 1.2: Standard Model leptons

Particles	Spin	Electric charge
$\begin{pmatrix}u\\ d'\end{pmatrix} \begin{pmatrix}c\\ s'\end{pmatrix} \begin{pmatrix}t\\ b'\end{pmatrix}$	$\frac{1}{2}$	$\frac{2}{3}$ $-\frac{1}{3}$

Table 1.3: Standard Model quarks

Each lepton and quark has an associated anti-particle with the same mass but opposite quantum numbers.
The last field in the SM is a complex scalar doublet field ϕ, named *Higgs field* after the physicist who predicted its existence in 1964, together with Brout and Englert [17, 18].

1.2 Quantum Electrodynamics

Quantum electrodynamics (QED) is the quantum field theory of the electromagnetic interaction. The Lagrangian density for the QED can be obtained starting from the free Lagrangian density of the Dirac field ψ and imposing its local invariance under $U(1)$ transformation group,

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

where the first is the kinetic term and the second is the mass term, m is the fermion mass, γ^μ are the Dirac matrices and ψ $\bar{\psi}$ are the 4-components spinor and its adjoint. It is worthy to precise that ψ spinor can be decompose in term of the chiral components,

$$\psi = \begin{pmatrix}\psi_L\\ \psi_R\end{pmatrix}. \tag{1.2}$$

The free, i.e. non-interacting, Lagrangian density of electromagnetic field is the lorentz-invariant

$$\mathcal{L}_\gamma = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \tag{1.3}$$

where $F^{\mu\nu}$ is the field strength tensor, defined by $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$. It can be obtained the free Lagrangian density for QED by adding 1.1 and 1.3,

$$\mathcal{L}_{QED}^{free} = \mathcal{L}_D + \mathcal{L}_\gamma = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\psi\bar{\psi} - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{1.4}$$

This Lagrangian density describes the kinematics of a non-interacting fermion in an electromagnetic field and it is globally invariant under a $U(1)_q$ transformation

$$\psi \to \psi' = e^{i\theta}\psi. \tag{1.5}$$

Through *Gauge principle*, e.g. by imposing invariance over the local version of the $U(1)$ transformation above,

$$\psi \to \psi' = e^{i\theta(x)}\psi, \tag{1.6}$$

where the parameter θ, now, depends on the space-time coordinates, arise a term of interaction beetwen fermion field and electromagnetic field

$$\mathcal{L}_{int} = -e\bar{\psi}\gamma^\mu A_\mu \psi = -J^\mu A_\mu. \tag{1.7}$$

This additive term can be reabsorbed by redefinition of the standard derivative with the covariante derivative defined as

$$D_\mu = \partial_\mu - ieA_\mu, \tag{1.8}$$

in a way to re-establish the invariance of Lagrangian density. The final form of Lagrangian QED is

$$\begin{aligned}\mathcal{L}_{QED} &= i\bar{\psi}(\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} \\ &= i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\psi\bar{\psi} - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - e\bar{\psi}\gamma^\mu A_\mu\psi,\end{aligned}$$

where the fields transform under gauge transformation as

$$\begin{aligned}A_\mu \to A'_\mu &= A_\mu - \frac{1}{e}\partial_\mu\theta(x) \\ \psi \to \psi' &= e^{-i\theta(x)}\psi.\end{aligned}$$

1.3 The Electroweak theory

The first phenomenological description of the weak interaction is due to Enrico Fermi [19], he provided a current-current lagrangian density,

$$\mathcal{L}_F = \frac{G_F}{\sqrt{2}} J^\dagger_\mu(x) J^\mu(x), \tag{1.9}$$

where G_F is Fermi's constant, $G_F = 1.16638 \times 10^{-5}$ GeV^{-2}, and J_μ is the weak current composed to leptoni current and hadronic current.
It is possible to observe weak interaction of three kinds:

- pure leptonic (e.g. $\mu^- \to e^- + \bar{\nu}_e + \nu_\mu$),

$$\mathcal{L}_F = \frac{G_F}{\sqrt{2}} \ell^\dagger_\mu(x)\ell^\mu(x); \tag{1.10}$$

- semileptonic (e.g. $n \rightarrow p + e^- + \bar{\nu_e}$),

$$\mathcal{L}_F = \frac{G_F}{\sqrt{2}}[\ell_\mu^\dagger(x)h^\mu(x) + h_\mu^\dagger(x)\ell^\mu(x)]; \tag{1.11}$$

- pure hadronic(e.g. $\lambda \rightarrow p + \pi^-$),

$$\mathcal{L}_F = \frac{G_F}{\sqrt{2}}h_\mu^\dagger(x)h^\mu(x). \tag{1.12}$$

These terms represent contact interactions, meaning that all particles involved in the process are in same space-time point when the interaction occurs. In the leptonic case, which only involves elementary particles, it is possible to observed the vector and spin structure of current is of the V-A type

$$l^\mu = \bar{e}(x)\gamma^\mu(1-\gamma^5)\nu_e(x) + \bar{\mu}(x)\gamma\mu(1-\gamma^5)\nu_\mu(x), \tag{1.13}$$

where $(1-\gamma^5)$ is the left-chiral projector of Dirac field, therefore only left-chiral particles are involved in the weak interaction. However the description is not complete, because it can be show the Fermi-like cross section is divergent.

1.3.1 The GSW model

In the 60's and 70's Glashow, Weinberg, and Salam elaboreted the Electroweak theory, unifying electromagnetic and weak interactions in the GSW model.

The GWS theory is an improvement of Fermi's theory, with the introduction of an adimensional coupling constant and three intermediate vector bosons. It is an $SU(2)_L$ quantum field theory, where L stands for left-handed chiral component.

Since the generators of $SU(2)_L$ are Pauli matrices, it is useful to adopt the formalism of angolar momentum, therefore the particles eigenstates of weak interaction are divided in six doublets of weak isospin with an important difference between lepton and quark sectors.

For the lepton sector

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

while, for the quark sector

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

where d', s', b' are eigenstates of weak interaction that are obtained as linear combinations of mass eigenstates, and the mixing of different states is regulated by the *CKM matrix*,

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{1.14}$$

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}, \tag{1.15}$$

where V_{CKM} was named after Cabibbo[20]-Kobayashi-Maskawa[21] and can be parametrized with 4 real parameters. In the so-called *Wolfestein parametrization* those parameters are named λ, η, ρ, A and the CKM matrix is written as:

$$V_{CKM} = \begin{pmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{pmatrix} + \mathcal{O}(\lambda^4).$$

The simultaneus fit to all SM input measurements results in [22] $\lambda = 0.22453 \pm 0.00044$, that means the matrix is almost diagonal and quarks are inclined to be associated with quarks to the same family in the interactions.

The GWS model is develped in a similar way with respect to the QED theory. Appling the Gauge principle on $SU(2)_L$ local transformation,

$$\begin{pmatrix} \nu_\ell \\ \ell^- \end{pmatrix} \rightarrow \begin{pmatrix} \nu_\ell \\ \ell^- \end{pmatrix}' = e^{-\frac{i}{2}\vec{\alpha}(x)\cdot\vec{\tau}} \begin{pmatrix} \nu_\ell \\ \ell^- \end{pmatrix}, \tag{1.16}$$

where $\vec{\tau}$ are the Pauli matrices and $\vec{\alpha}$ are the real parameters of transformation, the interaction terms can be obtained. The GSW model includes three gauge fields, named W^1, W^2, W^3, transforming as a multiplet under gauge group, two of this can be combined together in order to obtain two vector bosons $W^{\pm}$, that are electrically charged and can induce transitions between the members of the weak isospin doublets. The third boson should be electrically neutral and provides a *Neutral current* (NC) interaction analogously to *Charge Current* (CC) interactions mediated by $W^{\pm}$ bosons. Experimental evidences showed that NC interactions permit also couplings between right-handed components, so W^3 is not the only component of neutral boson but exists another one that admit right-handed coupling, similar to photon.

This feature can be exploited to provide an unified description of electromagnetic and weak forces by introducing a new $SU(2)_L \times U(1)_Y$ simmetry group, the new $U(1)_Y$ abelian group is associated with the weak hypercharge (Y), and Glashow proposed that the Gell-Mann, Nishijima relation

Particles	I	I_3	$Y^{(}W)$	Q
$\begin{pmatrix}\nu_e \\ e^-\end{pmatrix}_L$	$\frac{1}{2}$	$\frac{1}{2}$	-1	0
	$\frac{1}{2}$	$-\frac{1}{2}$	-1	0
e_R	0	0	-2	-1
$\begin{pmatrix}u \\ d'\end{pmatrix}_L$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{2}{3}$
	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{3}$	$-\frac{1}{3}$
u_R	0	0	$\frac{4}{3}$	$\frac{2}{3}$
d'_R	0	0	$-\frac{2}{3}$	$-\frac{1}{3}$

Table 1.4: Weak isospin and weak hypercharge of same elemanty particles

for charges should also hold for these weak analogues, giving

$$eQ = I_3^{(weak)} + \frac{Y^{(weak)}}{2}. \tag{1.17}$$

The quantum number of same elementary particles are reported in table 1.4.

When the gauge principle is applied, the charged vector bosons appear as before but there are, now, two neutral vector bosons, whose combination will be responsible for the weak NC processes and for electromagnetism. The covariant derivative, that is needed to be introduced in according to Gauge principle, is

$$D_\mu = \partial_\mu + ig\frac{\vec{\tau}}{2}W_\mu + ig'\frac{Y}{2}B_\mu, \tag{1.18}$$

where g and g' are the coupling constants for the two interactions that are related to G_F and electromagnetic coupling constant.
Neglecting the mass term for now, the electroweak Lagrangian density musts include a Dirac term for fermions

$$\mathcal{L}_{fermions} = \sum_f \bar{\psi}\gamma^\mu D_\mu \psi, \tag{1.19}$$

and a term for the dynamics of the gauge boson fields

$$\mathcal{L}_{gauge} = -\frac{1}{4}W_i^{\mu\nu}W^i_{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}, \tag{1.20}$$

where $W_i^{\mu\nu}$ and $B^{\mu\nu}$ are the tensor fields,

$$\begin{aligned} W_i^{\mu\nu} &= \partial^\mu W_i^\nu + \partial^\nu W_i^\mu \\ B^{\mu\nu} &= \partial^\mu B^\nu + \partial^\nu B^\mu. \end{aligned}$$

The complete Lagrangian density for the electroweak processes therefore is

$$\begin{aligned}\mathcal{L}_{EW} = -i\psi_L\gamma^\mu(\partial_\mu + ig\frac{\vec{\tau}}{2} \times W_\mu + id'\frac{Y}{2}B_\mu)\psi_L + \\ - i\psi_R\gamma^\mu(\partial_\mu + ig'\frac{Y}{2}B_\mu)\psi_R + -\frac{1}{4}W_i^{\mu\nu}W_{\mu\nu}^i - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} + \\ + \frac{1}{2}g\epsilon_{ijk}W_i^{\mu\nu}W_{j\mu}W_{k\nu} + \frac{1}{4}g^2\epsilon_{ijk}\epsilon_{imn}W_{j\mu}W_{k\nu}W_m^\mu W_n^\nu, \quad (1.21)\end{aligned}$$

where ψ_L and ψ_R are the left and right-handed chiral components of the particles, and the terms 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 $SU(2)_L$ group. The vector fields for the physical $W^\pm, Z$ bosons and for the photon can be obtain as linear combination of the four gauge filds

$$\begin{aligned}W_\mu^\pm &= \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \\ Z_\mu &= -\sin\theta_W B_\mu + \cos\theta_W W_\mu^3 \\ A_\mu &= \cos\theta_W B_\mu + \sin\theta_W W_\mu^3,\end{aligned} \quad (1.22)$$

where θ_W is the Weinberg angle and its value is not predict by SM, but it is related to g, g' and e

$$\begin{aligned}\theta_W &= \frac{g}{\sqrt{g^2 + g'^2}} \\ e &= g'\cos\theta_w = g\sin\theta_W \\ \frac{e^2}{g^2} &= \sin^2\theta_w \approx 0.23146 \pm 0.00012.\end{aligned}$$

Here the gauge fields are all massless, while the observed $W^\pm$ and Z vector bosons have a non-zero mass, as confirmed by the UA1 and UA2 collaborations. It is therefore necessary to introduce a new term for a boson field, invariant under an $SU(2)_L \otimes U(1)_Y$ transformation, that coherently represents the masses of the particles.

1.4 Spontaneous symmetry breaking and Higgs mechanism

The simplest and most elegant way in for $W^\pm$ and Z to acquire mass is the spontaneus simmetry breaking (SSB). The SSB explain such feature through the introduction of a new particle, the *Higgs* boson.
The Higgs boson field is a doublet of complex scalar fields,

$$\begin{pmatrix}\phi^+ \\ \phi^0\end{pmatrix} = \begin{pmatrix}\phi_1 + i\phi_2 \\ \phi_3 + i\phi_4\end{pmatrix} \quad (1.23)$$

and the associated Lagrangian density is

$$\mathcal{L}_H = (D^\mu\phi)^\dagger(D_\mu\phi) - V(\phi) = (D^\mu)^\dagger(D_\mu\phi) - \frac{1}{2}\mu^2\phi^\dagger\phi - \frac{1}{4}\lambda(\phi^\dagger\phi)^2, \quad (1.24)$$

where $V(\phi)$ is the potential responsible for simmetry breaking and λ is assumed positive. Acting with 1.18 on the Higgs field, it is obtain the follow terms,

$$\mathcal{L}_H = (D^\mu)^\dagger(D_\mu\phi) - \frac{1}{2}\mu^2\phi^\dagger\phi - \frac{1}{4}\lambda(\phi^\dagger\phi)^2 - \frac{1}{4}W^{\mu\nu}W_{\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu}, \quad (1.25)$$

with

$$D^\mu\phi = (\partial^\mu + ig\frac{\vec{\tau}}{2}W^\mu + ig'\frac{Y}{2}B^\mu)\phi \quad (1.26)$$

$$W^{\mu\nu} = \partial^\mu W^\nu - \partial^\nu W^\mu - gW^\mu \times W^\nu \quad (1.27)$$

$$B^{\mu\nu} = \partial^\mu B^\nu - \partial^\nu B^\mu. \quad (1.28)$$

By requiring that $\mu^2 < 0$ and $\lambda > 0$ the minimum of potential $V(\phi)$ is not unique and assumes the value

$$\phi_0^2 = -\frac{\mu^2}{2\lambda} \equiv \frac{v^2}{2}, \quad (1.29)$$

which identified a circumference in the complex plane. Without losing generality a minimum can be chosen, among the infinite possibilities,

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (1.30)$$

Considering the fluctuation of the ϕ field around the vacuum, the complex field becomes:

$$\phi = \begin{pmatrix} 0 \\ \frac{1}{\sqrt{2}}(v + H(x)) \end{pmatrix}. \quad (1.31)$$

By substitution of the 1.31 in 1.25, considering the expressions in 1.22 leads to

$$\begin{aligned} \mathcal{L}_{G\phi} = {} & \frac{1}{2}\partial^\mu H\partial_\mu H - \mu^2 H^2 \\ & - \frac{1}{4}(\partial_\mu W^1_\nu - \partial_\nu W^1_\mu)(\partial^\mu W^{1\nu} - \partial^\nu W^{1\mu}) + \frac{1}{8}g^2v^2W^1_\nu W^{1\nu} \\ & - \frac{1}{4}(\partial_\mu W^2_\nu - \partial_\nu W^2_\mu)(\partial^\mu W^{2\nu} - \partial^\nu W^{2\mu}) + \frac{1}{8}g^2v^2W^2_\nu W^{2\nu} \\ & - \frac{1}{4}(\partial_\mu Z_\nu - \partial_\nu Z_\mu)(\partial^\mu Z^\nu - \partial^\nu Z^\mu) + \frac{1}{8}(g^2 + g'^2)v^2 Z_\nu Z^\nu \\ & - \frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu). \quad (1.32) \end{aligned}$$

The first line is the Lagrangian density of Higgs scalar field with $m = \sqrt{2}\mu$.The next two lines show that the components W^1, W^2 of the triplet W_μ acquire mass,

$$M_1 = M_2 = \frac{1}{2}gv \equiv M_W,$$

the fourth line shows the mass of the field Z

$$M_Z = \frac{1}{2}v\sqrt{g^2 + g'^2} = \frac{M_w}{\cos\theta_W},$$

and the last line contains the field of the photon that not contain mass term.

1.4.1 Leptons masses

The fermion mass term

$$\begin{aligned} -m\psi\bar{\psi} &= -m(\psi_L + \psi_R)(\bar{\psi}_L + \bar{\psi}_R) \\ &= -m(\psi\bar{\psi}_L + \psi_R\bar{\psi}_R + \psi_L\bar{\psi}_R + \psi_R\bar{\psi}_L) \end{aligned}$$

is not invariant under the $SU(2)_L \otimes U(1)_Y$ group because of the different way the right- and left-handed chiral components of the fields transform. A gauge invariant mass term with a Yukawa coupling between the fermion field and the Higgs field can be introduced:

$$\mathcal{L}_Y = g_f(\bar{\psi}_L\phi\psi_R)(\bar{\psi}_R\phi\psi_L), \tag{1.33}$$

where g_f is the Yukawa coupling constant. Using 1.31 the term becomes

$$\begin{aligned} \mathcal{L}_Y = \frac{g_\ell}{\sqrt{2}}\left[(\nu_\ell, \ell)\begin{pmatrix} 0 \\ v+H \end{pmatrix}\ell_R + \ell_R(0, v+H)\begin{pmatrix} \nu_\ell \\ \ell \end{pmatrix}\right] \\ = \frac{g_f}{\sqrt{2}}(v+H)(\ell_L\ell_R + \ell_R\ell_L), \end{aligned} \tag{1.34}$$

and the constant befor term $(\ell_L\ell_R + \ell_R\ell_L)$ represents the mass of lepton

$$m_\ell = \frac{v}{\sqrt{2}}g_\ell.$$

The mass terms do not arise from a gauge principle, the g_l values is not predict by the SM, and all of them can in fact assume different values as it is shown by experimental evidence.

1.4.2 Quarks masses

The same approch used for leptons failed with quarks masses because the upper terms of doublets are not massless, in this case it can be written:

$$\begin{aligned} \mathcal{L}_Y = \frac{1}{\sqrt{2}}\left[g^d_{ij}(u_{iL}, d_{iL})\begin{pmatrix} 0 \\ v+H \end{pmatrix}d_{jR} + g^u_{ij}(u_{iL}, d_{iL}\begin{pmatrix} -(v+H)^* \\ 0 \end{pmatrix}u_{jR} + h.c.\right] \\ = \frac{1}{\sqrt{2}}(v+H)[g^u_{ij}(u_{iL}u_{jR} + u_{jR}u_{iL}) + g^d_{ij}(d_{iL}d_{jR} + d_{jR}d_{iL})], \end{aligned} \tag{1.35}$$

where $u_i = (u, c, t)$ and $d_i = (d, s, b)$, each mass term is

$$\begin{aligned} m^u_{ij} &= -\frac{v}{\sqrt{2}} g^u_{ij} \\ m^d_{ij} &= -\frac{v}{\sqrt{2}} g^d_{ij}. \end{aligned} \tag{1.36}$$

The masses matrix is not diagonal, but it can diagonalize with a series of transformations:

$$\begin{aligned} u_{\alpha L} &= (\mathcal{U}^u_L)_{\alpha i} u_{iL} \qquad & u_{\alpha R} &= (\mathcal{U}^u_R)_{\alpha i} u_{iR} \\ d_{\alpha L} &= (\mathcal{U}^d_L)_{\alpha i} d_{iL} \qquad & d_{\alpha R} &= (\mathcal{U}^d_R)_{\alpha i} d_{iR} \end{aligned} \tag{1.37}$$

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

$$\mathcal{L}_Y = \frac{1}{\sqrt{2}}(v + H)[m^u u\bar{u} + m^d d\bar{d} + m^c c\bar{c} + m^s s\bar{s} + m^b b\bar{b} + m^t t\bar{t}]. \tag{1.38}$$

The same transformations must to be applied to the interacting term invariant under the $SU(2)_L \otimes U(1)_Y$ 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 1.37 are unitary, instead the term of the coupling with the W boson, i.e. charged current coupling term, is

$$\begin{aligned} \mathcal{L}_{CC} &= -\frac{g}{\sqrt{2}}(\bar{u}_{iL}, \bar{d}_{iL})\gamma^\mu \tau_+ W^+_\mu \begin{pmatrix} u_{iL} \\ d_{iL} \end{pmatrix} + h.c. \\ &= -\frac{g}{\sqrt{2}} \bar{u}_{iL} \gamma^\mu d_{iL} W^+_\mu + h.c. \\ &= -\frac{g}{\sqrt{2}} \bar{u}_{\alpha L} \Big[(\mathcal{U}^u_L)_{\alpha i} (\mathcal{U}^d_L)^\dagger_{\beta i}\Big] \gamma^\mu d_{\beta L} W^+_\mu + h.c., \end{aligned} \tag{1.39}$$

where $V^{\alpha\beta}_{CKM} = \Big[(\mathcal{U}^u_L)_{\alpha i} (\mathcal{U}^d_L)^\dagger_{\beta i}\Big]^{\alpha\beta}$, the CKM matrix defined in 1.15.

1.5 Quantum cromodynamics

The Quantum Cromodynamics (QCD) is the non-Abelian quantum gauge field theory which describes the strong interactions. The simmetry group for QCD is $SU(3)_C$, where C stands for the charge associated with this simmetry, named *colour*. The only particles color-charged fermions are the quarks.

Due to non-Abelian nature of $SU(3)_C$ the generetors do not commute themselves, this means that in the Lagrangian density there would be a self-interaction terms between the gauge fields, as well as for electroweak interaction (1.21).

The QCD is invariant under $SU(3)_C$ transformation,

$$\psi \to \psi' = e^{i g_S \vec{\alpha}(x) \cdot \vec{\mathbf{T}}} \psi, \tag{1.40}$$

where g_S is the coupling constant, $\vec{\alpha}$ are the parameters of transformation and $\vec{\mathbf{T}}$ are the vectors made by the eight generators of the group, which are related to Gell-Mann 3×3 matrices (λ^a),

$$T^a = \frac{\lambda^a}{2}, \tag{1.41}$$

which follows the following commutation rules

$$\left[T^a, T^b\right] = i f_{abc} T^c, \tag{1.42}$$

f_{abc} are the structure constants of the group. Since the transformation acts on a 3-dimensional space, the ψ have three more additional degrees of freedom, these are rappresented as three states of colour,labelled as *red, green* and*blue*. The $SU(3)_C$ local phase transformation corresponds to rotating states in this colour space about an axis whose direction is different at every point in space-time. The Gauge principle leads to definition of the covariant derivative:

$$D_\mu = \partial_\mu + i g_S T_a G^a_\mu, \tag{1.43}$$

where G^a_μ are the eight gluon fields, that transform as:

$$G^a_\mu \rightarrow G'^a_\mu = G^a_\mu + i g_S f^{abc} \theta_b(x) G_{c\mu}. \tag{1.44}$$

From these, by adding the contributions of gluons, the complete Lagrangian density for QCD is obtained:

$$\mathcal{L}_{QCD} = \bar{\psi}\gamma^\mu \partial_\mu \psi - m\bar{\psi}\psi - i g_S \bar{\psi}\gamma^\mu \frac{\lambda_a}{2}\psi G^a_\mu - \frac{1}{4} G^{\mu\nu}_a G^a_{\mu\nu}, \tag{1.45}$$

where $G^{\mu\nu}_a$ is the tensor field defined as

$$G^{\mu\nu}_a = \partial^\mu G^\nu_a + \partial^\nu G^\mu_a - g_S f_{abc} G^{\mu b} G^{\nu c}.$$

The QCD theory exhibits two relevant properties that significantly differentiate it from the electroweak theory: *colour confinement* and *asymptotic freedom*.

Both these properties derive from experimental evidence and are successfully described in the context of the SM explanation of the strong interactions.

The latter property has been formulated to deal with the experimental evidence that no coloured hadron is observed in Nature. Hadrons are interpreted in the QCD parton model as bound states of quarks, and they must be colour singlets. This imposes restrictions on the types of quark bound state configurations can exist. All this can be summarized by saying that the quark colour degree of freedom must be confined.

The asymptotic freedom can be explained by looking at the running coupling constant form of strong interactions,

$$\alpha_{strong}(\left|q^2\right|) = \frac{\alpha_S(\mu^2)}{\left[1 + \alpha_S(\mu^2)\frac{33-2N_f}{12\pi} ln(\frac{|q^2|}{\mu^2})\right]}, \tag{1.46}$$

where $|q^2|$ is the transferred 4-momentum, μ is a scale parameter for the strength of the coupling, N_f is the number of fermions capable of strong interactions at the scale considered. It can be observed from 1.46 that α_S decreases when $|q^2|$ increases, in particular for $|q^2| \approx 200MeV$ its value is large enough that any perturbative approch cannot be applied. In this transferred momentum range the calculations are carried on with other approches, one of the most famous being the *lattice QCD* [23].

1.6 Unsolved issues of Standard Model

The SM succeeds in giving an exhaustive explanation of the three fundamental interactions: electromagnetic, weak, and strong. The electroweak theory successfully predicts the existence and the features of the weak neutral and charged currents, the existence and masses of the W and Z bosons. Another great success of the SM is the prediction of the existence of the Higgs boson, which is the last experimental validation of the electroweak theory. The discovery of charm quark, as required by the GIM mechanism, is another important experimental validation of the SM. Finally, the existence of the top quark, and the precise measurement of its mass has also been crucial to confirm the predictions of the SM, as its contributions to radiative corrections, driven by its large mass, which are a further validation of the mathematically-consistence of renormalizable field theory. Despite all of this compelling evidence, the SM model fails in giving explanation of many other phenomena, for example the original formulation did not provide for massive neutrinos, which are expected for the evidences of neutrino oscillation phenomena. Some notable criticality still subject to studies are:

- **Gravity**, the SM doesn't include in any way General Relativity;
- **Hierarchy in fermion masses**, the masses of known fermions vary on 5 orders of magnitude, between the top quark and the electron, and even less clear is the mass hierarchy of neutrinos, which are many orders of magnitude lighter still. There is no explanation or prediction of fermion masses that occur in a hierarchical pattern.
- **matter-antimatter asymmetry**, the SM CP-violation in quark sector and lepton sector are not sufficient to justify the observed matter - antimatter asymmetry in the universe;
- **Flavour Changing Neutral Current** (FCNC), the SM doesn't predict the observed suppression of flavour changing due to neutral current, as supposed by Glashow, Iliopulos and Maiani (GIM mechanism). Also the presence of only three families of quarks and leptons is not specifically predicted by the SM;
- **Higgs mass fine-tuning**, the tree-level (bare) Higgs mass receives corrections from fermion loop diagrams which are quadratically-divergent

and that are not cancelled by the boson loop diagrams. A precise tuning between fermions and scalars, known as *fine-tuning*, is required in order to cancel divergencies. In general, there is a tendency to naturalness of the theory, e.g. all parameters included in the theory have to be of the same order of magnitude, this tendency is opposite to fine-tuning approch, so one can propose ways to remedy the fine-tuning, which usually requires adding new particles manually;

- **Dark Energy and Dark Matter**, from astophysical observations there's evidences according to the Universe is made up only for the 5% of ordinary matter. The rest of the Universe is composed partially by the so-called *Dark Matter*, which constitutes 24% of the universe, while the remaining 71% is ascribed to a constant vacuum energy called *Dark Energy*. The existence of the Dark Energy would account for cosmological observations, in particular for the red shift measurement, which through it is discovered that the Universe is accelerated.

Chapter 2

CMS experiment at LHC

The *Compact Muon Solenoid* (CMS) is one of the main experiments at CERN's *Large Hadron Collider* (LHC). The acronym CERN stands for *Conseil européen pour la recherche nucléaire*, it is a European research organization, whose main site is located near Geneva, that operates the largest particle physics laboratory in the world. It was established in 1954, today it counts 23 Member States, and the name changed in *Organisation Européenne pour la Recherche Nucléaire*, although it has been decided that the acronym remains CERN.

2.1 Large Hadron Collider

The Large Hadron Collider [24] is the World's largest and most complex accelerator. It is a two-ring-superconducting-hadron accelerator and collider installed in the existing $26.7km$ tunnel that was constructed between 1984 and 1989 for the CERN *Large Electron Positron* accelerator.
The aim of the LHC is to reveal the physics beyond the Standard Model emerging from proton-proton collisions at centre-of-mass collision energies of up to 14 TeV. The number of events per second generated in the LHC collisions is given by:

$$N_{event} = L\sigma_{event}, \tag{2.1}$$

where σ_{events} is the cross section for the process under study and L the machine luminosity. The luminosity depends only on the beam parameters and, for a Gaussian beam distribution in the two axes orthogonal to the beam direction, can be written as:

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

where:

- N_b is the number of particles per bunch;
- n_b is the number of bunches per beam;
- f_{rev} is the revolution frequency;

Parameter	Value
N_b	1.15×10^{11}
n_b	2256
$epsilon[\mu m]$	2
$beta[cm]$	$30 - 25$
F	0.6

Table 2.1: The LHC parameters value

- γ_r is the relativistic Lorentz factor ($\frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$);
- ϵ is the normalized transverse beam emittance;
- β^* is the beta function at the collision point, it is a measure of how narrow the beam is at the interaction point, that could be seen as product of two separate function $sigma_x^*$ and σ_y^*, beam sizes in the horizontal and vertical plane divided by $\frac{\epsilon}{\gamma_r}$;
- F is the geometric luminosity reduction factor due to the crossing angle at the interaction point (IP):

  $$F = \left(1 + \left(\frac{\theta_c \sigma_z}{2\sigma^*}\right)^2\right)^{-\frac{1}{2}},$$

  where θ_c is the full crossing angle at the IP, σ_z the RMS bunch length, and σ^* the transverse RMS beam size at the IP.

Table 2.1 report the design values of the LHC beam parameters of the 2018 Data taking [25].The design luminosity of LHC is $10^{34} cm^{-2} s^{-1}$, however the peak luminosity achieved so far has been $2 \times 10^{34} cm^{-2} s^{-1}$ reached in several data taking periods of 2018. In fig. 2.1 it is compared with the values reached between 2011 and 2018. Margins in the design of the nominal LHC are expected to allow about two times the nominal design performance. The schedule for next years is summarised in fig. 2.2.

The LHC tunnel is located between 45 and 170 meters below the surface, crossing the border between France and Switzerland, and it is connected to the CERN accelerating complex by two tunnels. The CERN accelerating complex is made up of several steps represented schematically in fig. 2.3. The proton source used is an hydrogen gas bottle, the electrons of hydrogen atoms are stripped through an electric field, providing protons for injection in the accelaration complex:

- *Linac2*, a linear accelerator, that can accelerate protons up to 50 MeV;
- *Proton Synchrotron Booster* (PSB), which accelerates protons to 1.4 GeV;

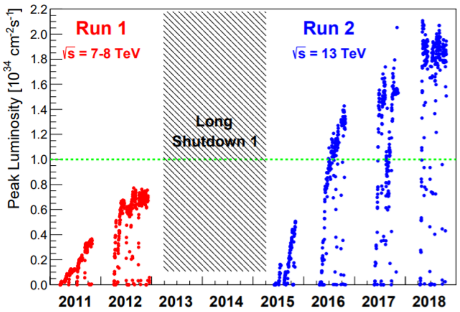


Figure 2.1: Evolution of the LHC peak luminosity between 2011 and 2018 [24]

Figure 2.2: Project schedule for LHC and HL-LHC [26]

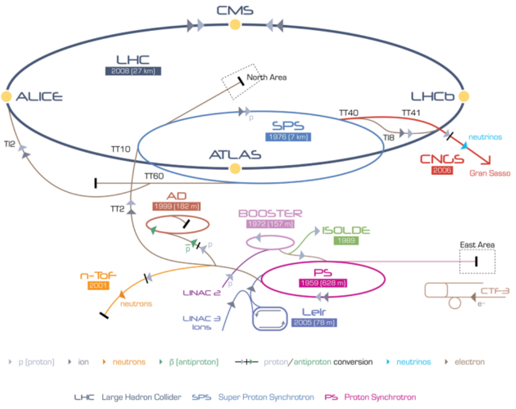


Figure 2.3: Schematic view of the acceleration complex at CERN [24]

- *Proton Synchrotron* (PS), which increases the protons energy to 25 GeV;
- *Super Proton Synchrotron* (SPS), last step before LHC , that pushes the beam to 450 GeV.

The protons are injected in LHC where the two beams travel in opposite directions, in two different beam pipes. In the LHC anular ring the beams are guided through the structures and are accelerated using a strong magnetic field, realized with superconducting electromagnets.
After reaching the target energy, beams are collided in four point where the four main experiments are located:

- ALICE (*A Large Ion Collider Experiment*) is a general-purpose heavy-ion detector. It is designed to address the physics of strongly interacting matter and the quark-gluon plasma at extreme values of energy density and temperature in nucleus-nucleus collisions. The energy in center-of-mass per nucleon pair is 2.76 GeV;
- ATLAS (*A Toroidal LHC ApparatuS*) is a general-purpose detector that aims at studyng a vast array of phenomena, designed for precision measurements of SM, the search and the study of Higgs boson, and searches for new physics. It is 46 m long and has a 25 m diameter and it is the biggest experiment at LHC;
- CMS (*Compact Muon Solenoid*)is also a general-purpose detector, sharing the same goals as ATLAS but designed with different technologies. It is described in section 2.2;
- LHCb (*LHC beauty*) is dedicated to heavy flavour physics. Its primary goal is to look for indirect evidence of new physics in CP violation and rare decays of beauty and charm hadrons. It is the only one out of the four main experiment that 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).

2.2 CMS experiment

CMS [27] is a general-purpose detector at the Large Hadron Collider (LHC). It has a broad physics programme ranging from studying the Standard Model to searching for extra dimensions and particles that could make up dark matter. An unusual feature of the CMS detector is that instead of being built in-situ like the other giant detectors of the LHC experiments, it was constructed in 15 sections at ground level before being lowered into an underground cavern near Cessy in France and reassembled.
The detector is built around a huge solenoid magnet, from which it takes the name. The magnet is a cylindrical coil of superconducting cable that generates a field of 3.8 T. The field is confined by a steel yoke that forms the bulk of the detector.

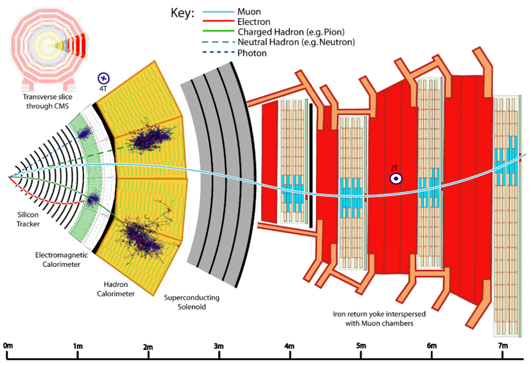


Figure 2.4: Schematic view of CMS sub-detectors [28].

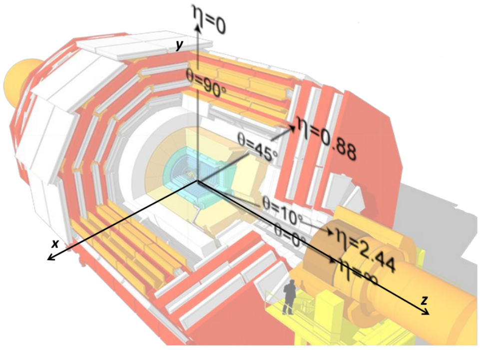


Figure 2.5: The coordinate system of CMS [27].

CMS has a complex system of sub-detectors, as can be seen in fig. 2.4, that allows to identify different particles with an high momentum resolution on a wide energy and angular coverage. It has a component coaxial to the beam, named *Barrel*, and two components to close it from each side, named *Endcaps*. CMS has a cylindrical symmetry and it is 21.6 m long with a diameter of 14.6 m and a total weight of about 14500 tons.

Fig. 2.5 shows the CMS coordinates system used to describe the detector and in the reconstruction of the tracks of particles. The coordinates system used to describe the detector is a right-handed Cartesian frame, centered in the IP, with x-axis points towards the center of LHC ring and z-axis along the beam direction toward the Jura mountains from LHC Point 5, also referred to as longitudinal direction. The y-axis is vertical and pointing upwards. The x-y plane is referred to as transverse plane. However, given the cylindrical symmetry of the CMS design and of the underlying physics, the physics objects are described in a cylindrical coordinate system (r, ϕ, θ):

- r is the distance from z-axis;
- ϕ is the azimuthal angle in the transverse plane, with $\phi = 0$ at x-axis;
- θ is the polar angle, measured from z-axis towards x-y plane.

Instead of using θ, it is convenient to use the *pseudorapidity* η, which is defined as:

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

Using (ϕ, η) parameters, a distance between two particle directions can be defined as:

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

Referring to the Cartesian system, the momentum of a particle is usually divided in two components, the longitudinal momentum p_z and the transverse momentum p_T, defined as:

$$\vec{p_T}^2 = \vec{p_x}^2 + \vec{p_y}^2.$$

The curvature in *x*-*y* plane, due to the magnetic field, allows for measurement of the particles p_T. It can be defined, for a particle of energy E, a variable called *rapidity*:

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

this variable, for an ultrarelativistic particle, can be approximated by the pseudorapidity. The differences in rapidity (Δy) and p_T are used because the centre-of-mass of parton-parton collisions can be boosted along the z direction; both these quantities have invariance properties under this kind of boost.

In the next subsections the CMS sub-detector systems are described, starting from the IP and going outwards, that consists of: the inner tracking system, the electromagnetic calorimeter, the hadron calorimeter, the superconducting magnet, and, finally, the iron return yoke interspersed with muon chambers.

2.2.1 The tracking system

The inner tracking system of CMS is designed to provide a precise and efficient measurement of the trajectories of charged particles emerging from the LHC collisions, as well as a precise reconstruction of secondary vertices. It surrounds the interaction point and has a length of 5.8 m and a diameter of 2.5 m. The CMS solenoid provides a homogeneous magnetic field of 3.8 T over the full volume of the tracker.

The CMS tracker is composed of a pixel detector with three barrel layers at radii between 4.4 cm and 10.2 cm and a silicon strip tracker with 10 barrel detection layers extending outwards to a radius of 1.1 m. Each system is completed by Endcaps which consist of 2 disks in the pixel detector and 3 plus 9 disks in the strip tracker on each side of the barrel, extending the acceptance of the tracker up to a pseudorapidity of $|\eta| < 2.5$.

With about 200 GeV^2 of active silicon area the CMS tracker is the largest silicon tracker ever built [29, 30].

2.2.2 The electromagnetic calorimeter

The electromagnetic calorimeter of CMS (ECAL) is a hermetic homogeneous calorimeter made of 61200 lead tungstate ($PbWO_4$) crystals mounted in the central barrel part, closed by 7324 crystals in each of the two Endcaps. The characteristics of the $PbWO_4$ crystals [31] make them an appropriate choice for operation at LHC. The high density(8.28 $\mathrm{g/cm^3}$), short radiation length (0.89 cm) and small Moliére radius (2.2 cm) result in a fine granularity and a compact calorimeter.

The scintillation decay time of these production crystals is of the same order of magnitude as the LHC bunch crossing time: about 80% of the light is emitted in 25 ns.

The barrel part of the ECAL (EB) covers the pseudorapidity range $|\eta| < 1.479$, occupies a volume of $8.14m^3$ and its weight is 67.4 tons.

The Endcaps (EE) cover the rapidity range $1.479 < |\eta| < 3.0$, occupy a volume of 2.90 $\mathrm{m^3}$ and the weight is 24.0 tons.

The energy resolution of the ECAL of CMS is:

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

with E in GeV.

2.2.3 The hadron calorimeter

The CMS detector is designed to study a wide range of high-energy processes involving diverse signatures of final states. The hadron calorimeters (HCAL) are particularly important for the measurement of hadron jets momenta and neutrinos or exotic particles resulting in apparent missing transverse energy.

The hadron calorimeter barrel and Endcaps sit behind the tracker and the electromagnetic calorimeter as seen from the IP. The hadron calorimeter barrel is radially restricted between the outer extent of the electromagnetic calorimeter (r = 1.77 m) and the inner extent of the magnet coil (r = 2.95 m). This constrains the total amount of material which can be put in to absorb the hadronic shower. Therefore, an outer hadron calorimeter or tail catcher is placed outside the solenoid complementing the barrel calorimeter. Beyond $|\eta| = 3$, the forward hadron calorimeters placed at 11.2 m from the IP extend the pseudorapidity coverage down to $|\eta| = 5.2$ using a Cherenkov-based, radiation-hard technology.

Since the calorimeter is inserted into the ends of a 3.8 T solenoidal magnet, the absorber must be made from a non-magnetic material. It must also have a maximum number of interaction lengths to contain hadronic showers, good mechanical properties and reasonable cost, leading to the choice of C26000 cartridge brass. The properties of this material are reported in table 2.2.

Properties of C26000	
Chemical composition	70% Cu, 30% Zn
Density	8.53 $\mathrm{g/cm^3}$
Radiation length	1.49 cm
Interaction length	16.42 cm

Table 2.2: Physical properties of the HCAL brass absorber, known as C26000/cartridge brass.

The energy resolution of the HCAL of CMS is:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{0.8470}{\sqrt{E}}\right)^2 + (7.40\%),$$

with E in GeV.

2.2.4 The superconducting magnet

The superconducting magnet [32] embraces the tracker detector and the calorimeters. It provides a magnetic field of 3.8 T and it is composed by a superconducting coil. The magnet's goal is to curve the particles produced in the collisions to measure their charge and momentum. Thanks to the iron Return Yoke, border effect are avoided. Moreover the yoke introduces a 1.8 T approximately constant field outside the magnet. The yoke is made up of alternating layers interspersed with the muon detectors.

2.2.5 The muon system

Muon detection is a powerful tool for recognizing signatures of interesting processes over the very high background rate expected at the LHC at design luminosity luminosity.

The muon system [33] is of central importance to CMS because in many physics processes muons result to be object reconstructable in the final, therefore precise and robust muon measurement could be used to distinguish interesting processes from backgrounds. The CMS muon system has 3 functions: muon identification, momentum measurement and triggering. Good muon momentum resolution and trigger capability are enabled by the high-field solenoidal magnet and its flux-return yoke. The latter also serves as a hadron absorber for the identification of muons.

CMS uses 3 types of gaseous particle detectors for muon identification. Due to the shape of the solenoid magnet, the muon system was naturally driven to have a cylindrical, barrel section and 2 planar Endcap regions. Because the muon system consists of about $25 \times 10^3\ \mathrm{m}^2$ of detection planes, the muon chambers had to be inexpensive, reliable, and robust.

In the barrel region, where the neutron-induced background is small, the muon rate is low, and the $3.8T$ magnetic field is uniform and mostly contained in the steel yoke, drift chambers with standard rectangular drift cells are used. The barrel *drift tube* (DT) chambers cover the pseudorapidity region $|\eta| < 1.2$ and are organized into 4 stations interspersed among the layers of the flux return plates. The first 3 stations each contain 8 chambers, in 2 groups of 4, which measure the muon coordinate in the $r - \phi$ bending plane, and 4 chambers which provide a measurement in the z direction. The fourth station does not contain the z-measuring planes. The 2 sets of 4 chambers in each station are separated as much as possible to achieve the best angular resolution. The drift cells of each chamber are offset by a half-cell width with respect to their neighbour to eliminate dead spots in the efficiency. This arrangement also provides a convenient way to measure the muon time with excellent time resolution, using simple mean timer circuits, for efficient, standalone bunch crossing identification. The number of chambers in each station and their orientation were chosen to provide good efficiency for linking together muon hits from different stations into a single muon track and for rejecting background hits.

In the 2 Endcap regions of CMS, where the muon rates and background levels are high and the magnetic field is large and non-uniform, the muon system uses *cathode strip chambers* (CSC). With their fast response time, fine segmentation, and radiation resistance, the CSCs identify muons between $|\eta|$ values of 0.9 and 2.4. There are 4 stations of CSCs in each Endcap, with chambers positioned perpendicular to the beam line and interspersed between the flux return plates. The cathode strips of each chamber run radially outward and provide a precision measurement in the $r - \phi$ bending plane. The anode wires run approximately perpendicular to the strips and are also read out in order to provide measurements of η and the beam-crossing time of a muon. Each 6-layer CSC provides robust pattern recognition for rejection of non-muon backgrounds and efficient matching of hits to those in other stations and to the CMS inner tracker. Because the muon detector elements cover the full pseudorapidity interval $|\eta| < 2.4$ with no acceptance gaps, muon identification is ensured over the range corresponding to $10^\circ < \theta < 170^\circ$. Offline reconstruction efficiency of simulated single-muon samples is typically $95 - 99\%$ except in the regions around $|\eta| = 0.25$ and 0.8 (the regions between 2 DT wheels) and $|\eta| = 1.2$ (the transition region between the DT and CSC systems), where the efficiency drops. Negligible punchthrough reaches the system due to the amount of material in front of the muon system, which exceeds 16 interaction lengths.

Due to multiple-scattering in the detector material before the first muon station, the offline muon momentum resolution of the standalone muon system is about 9% for small values of η and p for transverse momenta up to $200\ \mathrm{GeV}$. At $1\ \mathrm{TeV}$ the standalone momentum resolution varies between 15% and 40%, depending on $|\eta|$. A global momentum fit using also the inner tracker improves the momentum resolution by an order of magnitude at low momenta. At high momenta ($1\ \mathrm{TeV}$) both detector parts together yield a momentum resolution of about 5%. Note that the muon system and the

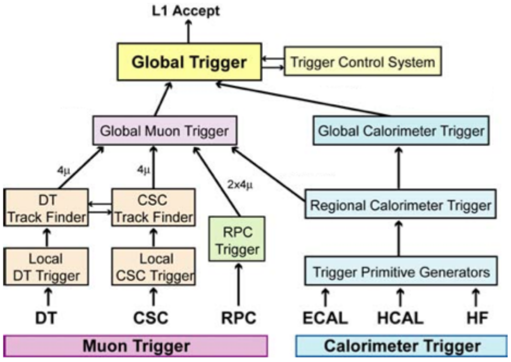


Figure 2.6: Architecture of Livel-1 Trigger [27].

inner tracker provide independent muon momentum measurements; this redundancy enhances fault finding and permits cross-checking between the systems. A crucial characteristic of the DT and CSC subsystems is that they can each trigger on the p_T of muons with good efficiency and high background rejection, independent of the rest of the detector. The Level-1 trigger p_T resolution is about 15% in the barrel and 25% in the Endcap.
Because of the uncertainty in the eventual background rates and in the ability of the muon system to measure the correct beam-crossing time when the LHC reaches full luminosity, a complementary, dedicated trigger system consisting of *resistive plate chambers* (RPC) was added in both the barrel and Endcap regions. The RPCs provide a fast, independent, and highly-segmented trigger with a sharp p_T threshold over a large portion of the rapidity range ($|\eta| < 1.6$) of the muon system. The RPCs are double-gap chambers, operated in avalanche mode to ensure good operation at high rates. They produce a fast response, with good time resolution but rougher position resolution than the DTs or CSCs. They also help to resolve ambiguities in attempting to make tracks from multiple hits in a chamber. A total of 6 layers of RPCs are embedded in the barrel muon system, 2 in each of the first 2 stations, and 1 in each of the last 2 stations. The redundancy in the first 2 stations allows the trigger algorithm to work even for low-p_T tracks that may stop before reaching the outer 2 stations. In the Endcap region, there is a plane of RPCs in each of the four stations in order for the trigger to use the coincidences between stations to reduce background, to improve the time resolution for bunch crossing identification, and to achieve a good p_T resolution.
Finally, a sophisticated alignment system measures the positions of the muon detectors with respect to each other and to the inner tracker, in order to optimize the muon momentum resolution.

2.2.6 The trigger system

The LHC provides proton-proton and heavy-ion collisions at high interaction rates. For protons the beam crossing interval is 25 ns, corresponding to a crossing frequency of 40 MHz. Depending on luminosity, several collisions occur at each crossing of the proton bunches (approximately 20 simultaneous pp collisions at the nominal design luminosity of $10^{34}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$). Since it is impossible to store and process the large amount of data associated with the resulting high number of events, a drastic rate reduction has to be achieved. This task is performed by the trigger system, which is the start of the physics event selection process. The rate is reduced in two steps called *Level-1* (L1) Trigger [34] and *High-Level Trigger* (HLT) [35], respectively. The Level-1 Trigger consists of customdesigned, largely programmable electronics, whereas the HLT is a software system implemented in a filter farm of about one thousand commercial processors. The rate reduction capability is designed to be at least a factor of 10^6 for the combined L1 Trigger and HLT. The design output rate limit of the L1 Trigger is 100 kHz, which translates in practice to a calculated maximal output rate of 30 kHz, assuming an approximate safety factor of three.

The L1 Trigger uses coarsely segmented data from the calorimeters and the muon system, while holding the high-resolution data in pipelined memories in the front-end electronics. The HLT has access to the complete readout data and can therefore perform complex calculations similar to those made in the the analysis off-line software if required for specially interesting events. For reasons of flexibility the L1 Trigger hardware is implemented in FPGA technology where possible, but ASICs and programmable memory lookup tables (LUT) are also widely used where speed, density and radiation resistance requirements are important. A software system, the Trigger Supervisor, controls the configuration and operation of the trigger components. The L1 Trigger has local, regional and global components. At the bottom end, the Local Triggers, also called Trigger Primitive Generators (TPG), are based on energy deposits in calorimeter trigger towers and track segments or hit patterns in muon chambers, respectively. Regional Triggers combine their information and use pattern logic to determine ranked and sorted trigger objects such as electron or muon candidates in limited spatial regions. The rank is determined as a function of energy or momentum and quality, which reflects the level of confidence attributed to the L1 parameter measurements, based on detailed knowledge of the detectors and trigger electronics and on the amount of information available. The Global Calorimeter and Global Muon Triggers determine the highest-rank calorimeter and muon objects across the entire experiment and transfer them to the Global Trigger, the top entity of the Level-1 hierarchy. The latter takes the decision to reject an event or to accept it for further evaluation by the HLT. 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 is depicted in fig. 2.6. The L1 Trigger has to analyse every bunch crossing. 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 $\mu\mathrm{s}$. The processing must therefore be pipelined in order to enable a quasi-deadtimefree operation. The L1 Trigger electronics is housed partly on the detectors, partly in the underground control room located at a distance of approximately 90 GeV from the experimental cavern.

Chapter 3

BSM physics with single top quark

As discussed in section 1.6, the SM does not provide an explanation for all observed phenomena in high energy physics. Several promising theories have been developed to provide explanations in scenarios where the SM fails to provide predictions, and foresee the existence of new particles in the LHC energy reach. Many theories that aim at solving the hierarchy problem and realizing the naturalness, predict the existence of new particles, the Vector Like Quarks (VLQs) [36] .

Another argument supporting the idea that new particles beyond the SM might exist arises from astrophysical measurements, such as the rotational speed of stars in galaxies and gravitational lensing. These observations point to the existence of non-light-emitting matter, a dominant fraction of which is of non-baryonic form, usually referred to as Dark Matter (DM). Even if there are no viable candidates in the SM for particles which could explain DM, pp collisions at the Large Hadron Collider (LHC) may possibly produce new particles that couple both to SM particles and to these DM candidates.

In this chapter we will describe the phenomenology of VLQs, in particular the VLQ named T, and of Dark Matter.

3.1 Vector-Like Quarks

Vector-like quarks are hypothetical spin 1/2 particles that transform as triplets under the colour gauge group and whose left- and right-handed components have the same colour and electroweak quantum numbers.

VLQs are the simplest example of coloured fermions still allowed by experimental evidence after the discovery of the Higgs boson. Extra quarks with chiral couplings, such as fourth generation quarks, are now in fact excluded [37] by the recent measurements of Higgs-mediated cross sections, when combined with direct searches at the LHC [38, 39].

VLQs, on the other hand, do not receive their masses from Yukawa couplings to a Higgs doublet, and are consistent with existing Higgs data.

Secondly, they can mix with the Standard Model (SM) quarks and thereby modify their couplings to the Z, W, and Higgs boson. Indeed, the addition of

VLQs to the SM is the simplest way of breaking the GIM mechanism, giving rise for example to tree-level flavour-changing neutral currents and potentially striking new effects in low energy physics, none of which have been observed, however. In this respect, new VLQs also introduce new sources of CP violation, as it typically occurs in most SM extensions.
In the third place, they can be analysed in a model-independent approach in terms of just a few free parameters.
Finally, VLQs at the TeV scale are strongly motivated by at least two theoretical ideas: they are required if the Higgs is a pseudo-Goldstone boson to induce electroweak breaking and explain the observed lightness of the Higgs [40, 41], and they emerge as fermion resonances in the partial-compositeness theory of flavour [42].
Due to the large Yukawa coupling of the top quark, both mechanisms give rise to a sizable mixing of the new quarks with the third family of SM quarks, hence the name *top partners*. Of course, VLQs do arise in all explicit models that implement these ideas, such as little Higgs and composite Higgs models, a brief description of which will be reported in the next subsection 3.1.1. They also appear in any model with quarks propagating in the bulk of extra dimensions and in grand unified and string theories based on the group E_6 [43], although in this case they are not guaranteed to be near the TeV scale.

3.1.1 Composite Higgs models

In 1984 Georgi and Kaplan proposed the Composite Higgs Model [44], in which the Higgs boson is a composite pseudo Nambu-Goldston Boson (pNGB), therefore it is a bound state of a new strong interaction. The Higgs particle is realized as a pNGB associated to the breaking $SO(5) \to SO(4)$ at a scale $f > v$, where v is the vacuum expectation value, this is usually referred as *minimal* extension since it adds only one field in the scalar sector. The $SO(5)$ symmetry has to be extended to the top quark sector by adding new vector-like quarks in order to reduce the UV sensitivity of m_h to the top quark loop. In principle new heavy vectors should also be included in order to cut-off the gauge boson loops, however here only the quark sector will be studied because the dominant contribution comes from the top.
Considering the vector Ψ_L of $SO(5)$ as an extension of the SM left handed doublet q_L for the third generation of quarks, it is possible to report a general model. The full fermionic content of the third quark generation is now:

$$\Psi_L = \begin{pmatrix} q = \begin{pmatrix} t \\ b \end{pmatrix}, X = \begin{pmatrix} X^{5/3} \\ X \end{pmatrix}, T \end{pmatrix}_L, t_R, X_R = \begin{pmatrix} X^{5/3} \\ X \end{pmatrix}_R, T_R.$$

The SM gauge group $GSM = SU(2)_L \times U(1)$ is here given by $SU(2)_L$ and the τ_3 of $SU(2)_R$ of the subgroup $SO(4) = SU(2)_L \times SU(2)_R \subset SO(5)$, since SO$(5)$ breaks up as $(2, 2) + 1$ under a $SU(2)_L \times SU(2)_R$ transformation. The right handed states have been introduced to give mass to the new fermions and their hypercharges have been fixed to obtain the correct electric charges. The Yukawa Lagrangian of the fermion sector consists of an $SO(5)$ symmetric mass term (this guarantees the absence of quadratic divergences in the contribution to m_h) for the top and the most general (up to redefinitions) gauge invariant mass terms for the heavy X and T:

$$\mathcal{L}_{top} = \lambda_1 \bar{\Psi}_L \phi t_R + \lambda_2 f \bar{T}_L T_R + \lambda_3 f \bar{T}_L t_R + M_X \bar{X}_L X_R + h.c.,$$

where $\lambda_{1,2,3}$ are the coupling constants, ϕ is is the scalar 5-plet containing the Higgs Field and M_X is the mass of the heavy X quark.
Since $\phi = \begin{pmatrix} H \\ H^c \end{pmatrix}$, the previous Lagrangian becomes:

$$\begin{aligned}\mathcal{L}_{top} = \bar{q}_L H^c(\lambda_t t_R + \lambda_T T_R) + \bar{X}_L H(\lambda_t t_R + \lambda_T T_R) \\ + M_T \bar{T}_R T_L + M_X \bar{X}_L X_R + h.c. \quad .\end{aligned}$$

Through diagonalization of the mass matrix we obtain the physical fields.It can be check the cancellation of the quadratically divergent contribution to m_h through the introduction of the VLQs considering the potential:

$$V = \lambda(\phi^2 - f^2)^2 - Af^2\vec{\phi}^2 + Bf^3\phi_5,$$

where $\vec{\phi}$ are the first four components of ϕ. The Higgs boson mass can be shown to be controlled by the A parameter, that is by the $SO(5)$-breaking term, $m_h = 2v\sqrt{A}$ for big λ. This is reasonable since if SO(5)-symmetry is not broken, the Higgs particle would be a massless Goldstone boson. Setting $v = 0$, the divergent part of the one loop correction to A could be written as:

$$\begin{aligned}\delta A = -\frac{12f^2}{64\pi^2}\lambda^2\left(\frac{M_X^2}{f^2} - 4(\lambda_1 + \lambda_3) - 2\lambda_2^2\right)\ln\lambda^2 \\ = -\frac{3}{16\pi^2 f^2}(\lambda_t^2 + \lambda_T^2)\left(M_X^2 + M_T^2\left(\frac{2}{1 + \frac{\lambda_T^2}{\lambda_t^2}}\right)\right)\ln\lambda^2.\end{aligned}$$

It is shown that there is no quadratic divergence and the two mass of the VLQs, M_X and M_T , could be seen as the cut-off Λ in the original top-loop contribution. However, the two masses can not be too much above 2 TeV, or the logarithmic term could produce a δm_h of the same order of the weak-scale expectation value v, and therefore the naturalness problem is not solved.
This simplified model introduces the VLQs [45], but often models like it fail to explain the fermion masses origin or they do not pass the electroweak precision tests. The Composite Higgs are not only ones theories that predict the VLQs in the Composite scenario. The Composite Top theories, in a similar way, suppose that the top quark is a composite state. The SM particles,

X	+5/3
T	+2/3
B	-1/3
Y	-4/3

Table 3.1: Charge assignment for VLQs.

thanks to a new strong sector, get their masses by mixing themselves with composite states. The top quark, due to its large mass, could show the properties of Compositeness, having a sizeable admixture of the composite state. The electroweak precision tests further models of right-handed composite top quark models.

3.1.2 Phenomenology of VLQs

The new vector-like quarks, which are presented in table 3.1 with their electric charges, coupling to the SM ones with renormalisable couplings can only appear in seven gauge-covariant multiplets with definite $SU(3)_C \times SU(2)_L \times U(1)_Y$ quantum numbers:

$$\begin{array}{cccc} & singlets & T^0_{L,R} \quad B^0_{L,R} & \\ doublets & \left(X, T^0\right)_{L,R} & (T^0, B^0)_{L,R} & (B^0, Y)_{L,R} \\ triplets & (X, T^0, B^0)_{L,R} & (T^0, B^0, Y)_{L,R}, & \end{array}$$

where T_L and T_R are the left- and right- handed components, while the superscript 0 indicates weak eigenstates. The weak and the mass eigenstates for X and Y coincide, since they cannot mix with the SM quarks due to their exotic charges.

When the new fields $T^0_{L_R}$ are added to the SM lagrangian, the up-type eigenstates (u, c, t, T) can contain components of the new fields, changing the coupling of the quarks to the Z boson. However, the mixing is proportional to the ratio between the mass of the SM quark and the VLQ, and this allow to consider the VLQs mixing only with the third generation quarks, being m_q/m_{VLQ} negligible for the first two generations. The top quark, thanks to these considerations and to its large Yukawa coupling, is linked to new physics, related with electroweak symmetry breaking and the fermion mass hierarchy.

The signatures of VLQs have been analyzed both in model independent and in specific model-dependent scenarios. For the latter case, the possibility of flavour changing neutral currents processes for VLQs interactions leads to a wide range of potential final states, being sought after at the LHC.

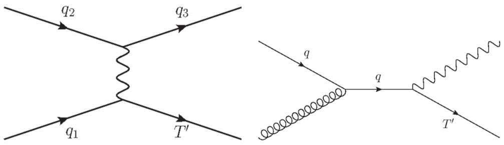


Figure 3.1: Feynman diagrams for single production of Vector-like quarks T.

Production mechanism

The VLQs 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 divided into :

◇ the single production, via electroweak processes. The single production process depends on the fermion mass, on the mixing parameters with SM particles and on the couplings between the new quarks and the W and Z bosons:

$$qq' \xrightarrow{V^*} qQ \qquad V = W, Z.$$

Due to the small masses of the light quarks, the contributions of the Higgs bosons are always suppressed.
In fig. 3.1 is shown the single electroweak interaction mediated by a vector boson in association with a SM quark and the one mediated by SM quark in association with a vector boson.

◇ The pair production is dominated by QCD processes via gluons. Since the cross section of this kind of process only depends on the mass of the new fermion, it is model independent and it 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:

$$gg, q\bar{q} \to Q\bar{Q} \qquad Q = T, B, X.Y.$$

A small contribution of electroweak gauge bosons, subleading in terms of cross section, is still present. This contribution leads to interesting channels when the production is reached trought charged current process:

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

However, the electroweak cross sections are strongly suppressed for large masses and their effects on the search strategies can be safely neglected. Rather than the pair production due to QCD interactions, the cross sections of electroweak pair production are model-dependent as

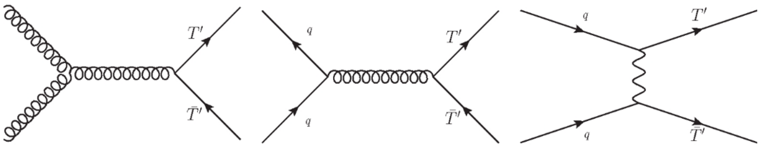


Figure 3.2: Feynman diagrams for pair production of Vector-like quarks T [36].

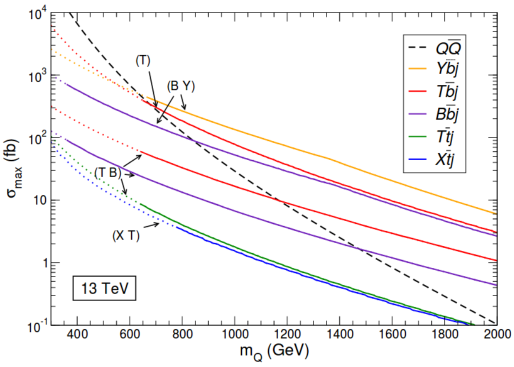


Figure 3.3: Production cross section for Vector-Like Quarks in pp collisions at $\sqrt{s} = 13TeV$ as a function of their mass, for pair production and for single production in different channels. The black dashed line represents VLQ pair production, while the colored lines represent the singly produced VLQs. The dashed coloured lines correspond to the values of cross section excluded by previous studies [36].

> they depend on the representation the VLQ belong to. Another relevant electroweak production process is mediated by the 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 fig. 3.2.

The dependence of the pair production and the single production cross sections on the energy is illustrated in fig. 3.3. The plot shows that pair production processes are the dominant ones for masses below $m_Q \sim 800 - 1000$ GeV, while they becomes less important for higher masses due to their phase-space suppression.

Decay mechanism

Vector-like quarks can decay through electroweak interactions into SM particles, also trought flavour-changing neutral currents process since they break the GIM mechanism, or into other VLQs. Considering the the decay chan-

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

Table 3.2: Allow decay modes for Vector-like singlets.

nels into SM particles, the main allowed channels are:

$$\begin{gathered} T \to W^+b, Zt, Ht \\ B \to W^-t, Zb, Hb \\ X \to W^+t \\ Y \to W^-b. \end{gathered}$$

The branching ratio of T and B are different in the three scenarios of multiplets.
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:

$$Br(Q \to Qq') + Br(Q \to Zq) + Br(Q \to Hq) = 1,$$

where $(Q, q, q^0) = (T, t, b), (B, b, t)$.
The scenario is different for the isospin doublets and triplets. First of all, looking at the small mass difference beetwen VLQs, the decays to other VLQ are usually suppressed for doublets and triplets, impling that the only allowed decays are into vector bosons and the Higgs boson plus a t or b quark. The decays of the doublet (T, B) depend on the mixing factor of the extended CKM matrix V_{Tb} and V_{tB}, if $V_{Tb} \sim V_{tB}$ implies that the T and B quarks have the same decays as the corresponding singlets. They have different angular distributions since only the right-handed component of (T, B) couples to the SM quarks. However, due to constraints on the b quark mixing and by mass hierarchy, $m_t \gg m_b$ where $V_{Tb} \ll V_{tB}$, the mixing of the heavy quarks with the SM top quark is much stronger. Therefore the $T \to Wb$, $B \to Hb$ and $B \to Zb$ decays are suppressed.
The possible decays of vector-like quarks are reported in tables 3.2,3.3 and 3.4. On the other hand, the branching ratios of the vector-like quarks are model-dependent and they also depend on the heavy quark mass themselves. In fig. 3.4 are illustrated the branching ratios of the decays of T and B in the case of a VLQ coming from the singlet or doublet of $SU(2)$.

Search for VLQs

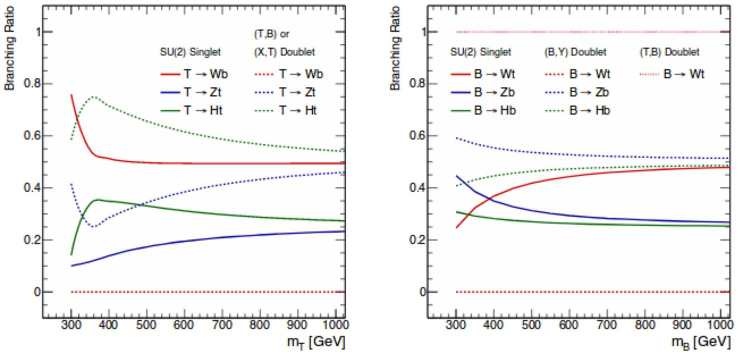
Various searches of VLQs have been undertaken both at Tevatron and at LHC. The main studies at Tevatron have been done during Run II with a

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 3.3: Allow 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 3.4: Allow decay modes for Vector-like triplets.



(a) (b)

Figure 3.4: 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 [36].

center-of-mass energy of 1.96 TeV by the CDF and D0 experiments. The studies have been done and are still going on at LHC by CMS and ATLAS experiments.
The D0 collaboration performed a search for single production of VLQs at 5.4 fb^{-1} [46].The analysis searches for final states with either a W or Z boson and two jets. One jet comes from the VLQ decay and the other is produced in association with the VLQ. The gauge boson is then required to decay leptonically and events with exactly one lepton or exactly two leptons and more than 2 jets are selected. Results are given for different choices of the coupling parameters and assumptions on BRs:

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

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

The CDF collaboration has performed two analyses on VLQs at integrated luminosity of 5.7 fb^{-1}:

- The search for pair production of a heavy particle T decaying to tX, where X is an invisible dark matter, the search considers events in the full hadronic channel [47]. The event selection has been made requirinig a number of jet among 5 and 10 and a lot of missing transverse energy. The bounds have been provided 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.
- The search for single production of heavy quarks where it is expected to decay 100% in Wq, where q is a SM quark of the first generation. The signal event has the topology $W + 2j$ where the W is required to decay leptonically. Due to the fact that single production is model dependent, 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.

The main serches at LHC by ATLAS and CMS are based on QCD pair production of charge $5/3$, $2/3$, and $-1/3$ partners and their decay into third family quarks and W, Z, H bosons. The analysis strategies have required a lot of different final states: all-hadronic searches, single- or multi-lepton final state, with or without transverse missing energy.
The possible decays that have been studied for the TT^{-} pair production are:

- The decay of at least one T in Ht or a T in tZ, with $H \to bb$ and $Z \to \nu\nu$ respectively.It was required in the final states at least one lepton, from the top decay, multi-jets and missing transverse energy [48].

- The decay of $T\bar{T}$ both in the Wb channel, where one W decays to leptons and one decays to quarks, assuming a $Br(T \to Wb) = 100\%$ [49].
- The study of the $Zt+X$ final state with exactly one charged lepton and $Z \to \nu\nu$. The analysis has provided an upper limits on the T mass of 0.85 (1.05) TeV, considering the weak-isospin singlet (doublet) model; while an upper limit of $m_T \leq 1.16$ TeV for the pure Zt decay mode [50].
- The decay of the T pair to $bWbW \to bl\nu\bar{b}qq'$. In the final state are required just one charged lepton, at least 4 jets and a boosted W-tagged. The assuming a branching fraction $Br(T \to bW) = 100\%$ and the reconstructing the mass of the T quark, thanks to a kinematic fit, have provided an upper limit on the T quark mass of 1295 GeV [51].
- The pair production of T or B in fully hadronic final states, using a multiclassification algorithm to improve the reconstruction of the events [52].
- The decay of a T pair, where one decays via $T \to tZ$ and the other one via $T \to tZ, bW, tH$ with two oppositely charged leptons in the final state coming from the Z boson [53]. This final state is studied also for the B pair production.
- The decay of a T pair considering three channels, corresponding to final states with a single lepton, two leptons with the same sign of the electric charge, or at least three leptons [54].

The searches for $B\bar{B}$ pair production have studied the decays into tW, bZ, bH. The analysis strategies follow the same of the $T\bar{T}$ searches.
The studies of single production of VLQs have analyzed the following channels:

- production of a T quark decaying to tZ with $Z \to ll$ and $t \to hadrons$ [55];
- production of $B \to bH$ with $H \to bb$ [56].
- production of a T quark decaying to tZ or tH in fully hadronic final states [57].

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 samesign channels [56, 58].
Fig. 3.5 shows the results of the searches for the single and the pair production of VLQs by CMS in July 2019.

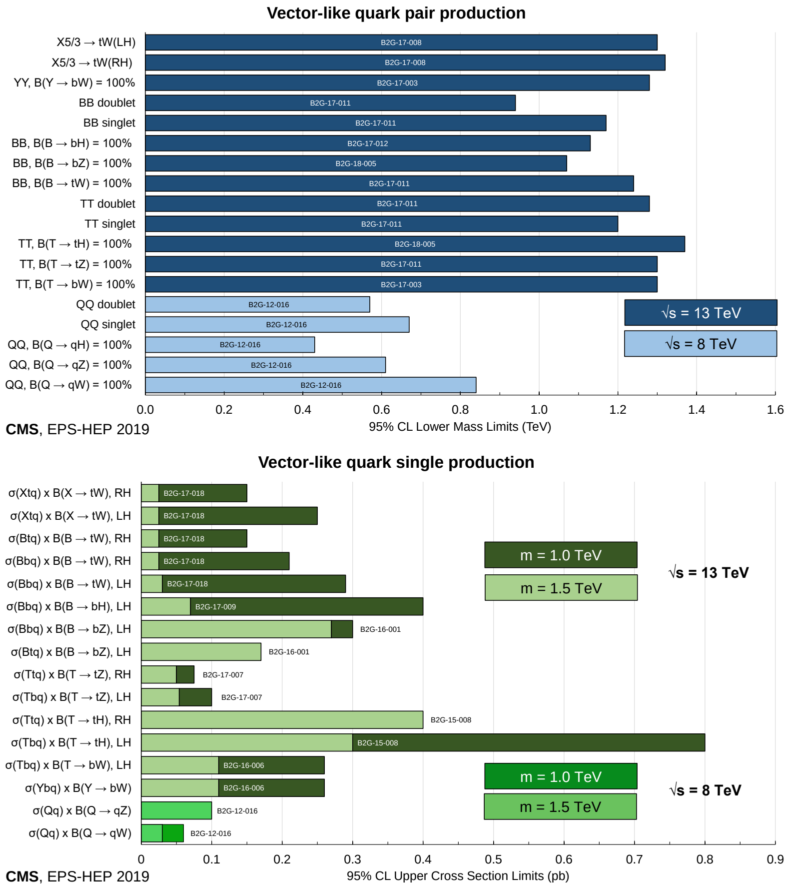


Figure 3.5: Summary of the results of the VLQs single and pair production [59].

3.2 Dark Matter

Dark matter [60] (DM) is a hypothetical form of matter thought to account for approximately 84.4% of the matter in the universe, and about 26.4% of its total mass-energy density. Its presence is implied in a variety of astrophysical observations, including gravitational effects that cannot be explained by accepted theories of gravity unless more matter is present than can be seen. For this reason, it is common to think that dark matter is abundant in the universe and that it has had a strong influence on its structure and evolution.

Dark Matter is called *dark* because it does not appear to interact with the electromagnetic field, which means it does not absorb, reflect, or emit electromagnetic radiation, and is therefore difficult to detect. Because DM has not yet been observed directly, if it exists, it must barely interact with ordinary baryonic matter and radiation, i.e. through interactions with small couplings like for example gravity. Most dark matter is thought to be non-baryonic in nature; it may be composed of some as-yet undiscovered subatomic particles. The primary candidate for DM is some new kind of elementary particle that has not yet been discovered, ofeten referred to as *Weakly Interacting Massive Particles* (WIMPs).

Although the existence of dark matter is generally accepted by the scientific community, some astrophysicists, intrigued by certain observations which are not well-explained by standard DM, argue for various modifications of the standard laws of general relativity, such as *Moddified Newtonian Dynamics* (MOND) [61], *tensor-vector-scalar gravity*, or *entropic gravity*. These models attempt to account for all observations without invoking supplemental non-baryonic matter.

3.2.1 Observational evidence

The hypothesis of DM has an elaborate history. In a talk given in 1884, Lord Kelvin [62] estimated the number of dark bodies in the Milky Way from the observed velocity dispersion of the stars orbiting around the center of the galaxy. By using these measurements, he estimated the mass of the galaxy, which he determined is different from the mass of visible stars.

The first to suggest the existence of Dark Matter using stellar velocities was Dutch astronomer Jacobus Kapteyn [63] in 1922. Fellow Dutchman and radio astronomy pioneer Jan Oort [64] also hypothesized the existence of dark matter in 1932.

In 1933, Swiss astrophysicist Fritz Zwicky [65], who studied galaxy clusters while working at the California Institute of Technology, made a similar inference. Zwicky applied the virial theorem to the Coma Cluster and obtained evidence of unseen mass he called dunkle Materie ('dark matter'). Zwicky estimated its mass based on the motions of galaxies near its edge and compared that to an estimate based on its brightness and number of galaxies. He estimated the cluster had about 400 times more mass than was visually observable. The gravity effect of the visible galaxies was far too small for such fast orbits, thus mass must be hidden from view. Based on these conclusions, Zwicky inferred some unseen matter provided the mass and associated gravitation attraction to hold the cluster together.

Galaxy rotation curves

The first historical evidence to support the existence of Dark Matter comes from the observations of the rotation curves of galaxies, namely the graph of circular velocities of stars and gas as a function of their distance from the galactic center. Rotation curves are usually obtained by combining observations of the 21 cm line with optical surfacephotometry. Observed rotation curves usually exhibit a characteristic flat behavior at large distances, i.e.out towards, and even far beyond, the edge of the visible disks.
In Newtonian dynamics the circular velocity is expected to be:

$$v(r) = \sqrt{\frac{GM(r)}{r}} \tag{3.1}$$

where $M(r) = 4\pi \int \rho(r) r^2 dr$ and $\rho(r)$ is the mass density profile. In case of uniform density:

$$\rho(r) = \begin{cases} \rho_0 & \text{for } r \leq R \\ 0 & \text{for } r > R \end{cases},$$

that implies

$$\begin{cases} M(r) \propto r^3 \text{ and } v(r) \propto r & \text{for } r \leq R \\ M(r) \sim constant \text{ and } v(r) \propto \frac{1}{\sqrt{r}} & \text{for } r > R \end{cases},$$

thus, Newtonian dynamics does not explain the observed flat behaviour of velocity at large distances.
Considering $v(r) \sim constant$, it can be obtained:

$$\rho \propto r^{-2} \tag{3.2}$$

The proportionality of 3.2 for larger r is fequently parameterized in the following form:

$$\rho(r) = \rho_0 \frac{R_0^2 + a^2}{r^2 + a^2} \tag{3.3}$$

where r is the galactocentric distance, $R_0 = 8.5$ kpc (for the Milky Way) is the galactocentric radius of the Sun, and a = 5 kpc is the radius of the halo nucleus, $\rho_0 = 0.3\ \mathrm{GeV/cm^3}$ is the local energy density in the solar system.

Gravitational lensing

Another important evidence in favor of DM existence is the observed phenomenon of gravitational lensing. Following Einstein's theory of general relativity, light propagates along geodesics which deviate from straight lines when passing near intense gravitational fields. The distortion of the images of background objects due to the gravitational mass of a cluster can be used to infer the shape of the potential well and thus the mass of the cluster.

Strong lensing is the observed distortion of background galaxies into arcs when their light passes through such a gravitational lens. It has been observed around many distant clusters. By measuring the distortion geometry, the mass of the intervening cluster can be obtained. In the dozens of cases where this has been done, the mass-to-light ratios obtained correspond to the dynamical dark matter measurements of clusters. Lensing can lead to multiple copies of an image. By analyzing the distribution of multiple image copies, the distribution of dark matter around the galaxy clusters can be obtained.

Weak gravitational lensing investigates minute distortions of galaxies, using statistical analyses from vast galaxy surveys. By examining the apparent shear deformation of the adjacent background galaxies, the mean distribution of dark matter can be characterized. The mass-to-light ratios correspond to dark matter densities predicted by other large-scale structure measurements. Dark matter does not bend light itself; mass (in this case the mass of the dark matter) bends spacetime. Light follows the curvature of spacetime, resulting in the lensing effect.

The phenomena is observed by several experiments, such as MACHO, EROS and OGLE, that detected a dozen of *MAssive Compact Halo Objects* (MACHOs).

The baryonic non-luminous matter observed through gravitational lensing is not sufficient to explain all DM required by the other evidences, although Gravitational lensing studies of the *Bullet Cluster* are claimed to provide the best evidence to date for the existence of DM.

The Bullet Cluster consists of two colliding clusters of galaxies. Strictly speaking, the name refers to the smaller subcluster, moving away from the larger one. The major components of the cluster pair behave differently during collision, allowing them to be studied separately. The stars of the galaxies, observable in visible light, were not greatly affected by the collision, and most passed right through, gravitationally slowed but not otherwise altered. The hot gas of the two colliding components, seen in X-rays, represents most of the baryonic matter in the cluster pair. The gases interact electromagnetically, causing the gases present in the two clusters interact electromagnetically, causing them to slow down much more with respect to the stars. The third component, the Dark Matter, was detected indirectly by the gravitational lensing of background objects. In theories without DM, such as MOND, the lensing would be expected to follow the baryonic matter; i.e. the X-ray gas. However, the lensing is strongest in two separated regions near (possibly coincident with) the visible galaxies. This provides

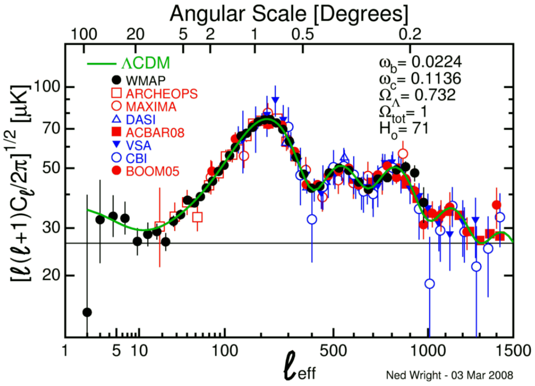


Figure 3.6: The observed power spectrum of CMB anisotropies.

support for the idea that most of the gravitation in the cluster pair is in the form of two regions of dark matter, which bypassed the gas regions during the collision.

Angular power spectrum of CMB

The existence of Cosmic Microwave Background (CMB) originating from the propagation of photons in the early Universe, once they decoupled from matter, was predicted by George Gamow and his collaborators in 1948 and inadvertently discovered by Arno Penzias and Robert Wilson in 1965. After many decades of experimental effort, the CMB is known to be isotropic at the 10^{-5} level and to follow with extraordinary precision the spectrum of a black body corresponding to a temperature T=2.726 K.
Today, the analysis of CMB anisotropies enables accurate testing of cosmological models and puts stringent constraints on cosmological parameters. The observed temperature anisotropies in the sky are usually expanded as:

$$\frac{\delta T}{T}(\phi,\theta) = \sum_{l=2}^{\infty} a_{\ell m} Y_{\ell m}(\phi,\theta), \tag{3.4}$$

where $Y(\phi,\theta)$ are the spherical harmonics. The variance C_ℓ of $a_{\ell m}$ is given by:

$$C_\ell \equiv \langle |a_{\ell m}|^2 \rangle = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2. \tag{3.5}$$

If the temperature fluctuations are assumed to be Gaussian, as appears to be the case, all of the information contained in CMB maps can be compressed into the power spectrum, essentially giving the behavior of $C\ell$ as a function of ℓ. Fig. 3.6 shows the quantity $\ell(\ell+1)C_\ell/2\pi$ commonly used to visulize the power spectrum. Information from CMB anisotropy maps can be extracted, starting from a cosmological model with a fixed number of parameters, usually 6 or 7, the best-fit parameters are determined from the peak of the N-dimensional likelihood surface. The following values are, at the moment, found for the abundance of baryons and matter in the Universe:

$$\Omega_{baryons} = 0.02237 \pm 0.00015 \qquad \Omega_{matter} = 0.315 \pm 0.007\,. \tag{3.6}$$

The observed CMB angular power spectrum provides powerful evidence in support of dark matter, as its precise structure is well fitted by the Λ-CDM model, but difficult to reproduce with any competing model such as MOND.

Main DM candidate in particle physics

Particle DM model building is deeply intertwined with the question of the nature of physics BSM. Directions in this area have followed two main strategies: either by pursuing DM candidates embedded in frameworks that include solutions to other open issues in particle physics or bottom-up models built with the intent of addressing or explaining a putative signal, coming either from (e.g. particle physical anomalies) or observational (e.g. astronomical). Follow some examples of both of these types of approches.

- *Axions* are an especially compelling example of a broad category of DM candidates encompassing very light scalar or pseudoscalar fields. The QCD axion provides a solution to the strong CP problem, and it is at present a viable DM candidate.
- *Sterile Neutrinos*, assumed to share a Dirac mass term with ordinary $SU(2)_L$-active neutrinos, have long been considered viable DM candidates. The mostly-sterile mass eigenstate participates in $SU(2)_L$ interactions via a mixing parameter $\theta \ll 1$ that controls much of the particle's phenomenology.
- *Dark Photons*, light vector bosons such as a dark photon V with a mass below $m_V < 2m_e$, can be cosmologically stable, depending upon its kinetic mixing coupling with the visible photon, and be a viable DM candidate.
- *WIMPs* has been a preferred framework chiefly because it often arises in BSM scenarios that address the hierarchy problem whilst also providing a simple mechanism to explain the observed relic abundance via the *WIMP miracle* [66].

3.2.2 DM production at LHC

As discuss in 3.2.1, a compelling candidate for DM is the so called WIMP. This new particle is predicted to have weak interactions with SM particles, allowing for direct- and indirect-detection experiments, as well as for searches at collider experiments.

Dark Matter production [67] via thermal mechanism provides a tantalizing possibility for collider search. Should these interactions exist, DM would be produced in proton-proton collisions at LHC and studied in detail. Collider data could access a different range of possible interactions between dark matter and SM particles than current direct and indirect searches, allowing an important interplay and complementarity among experiments in the quest of discovering dark matter.

A WIMP candidate must be electrically neutral, noninteracting via the strong nuclear force, and stable or metastable, with a decay lifetime larger than the age of the Universe ($\sim 10^{18}$ s). Therefore, dark matter is invisible at the LHC, and its production at colliders can be only inferred indirectly by a large momentum imbalance in the transverse plane of the detector. The specific nature of these visible particles is model-dependent, and given our lack of knowledge of the physics of the dark sector, it is necessary to consider all experimentally-viable possibilities.

In view of the center-of-mass energy reached at LHC in Run II, $\sqrt{s} = 13\ \mathrm{TeV}$, an interpretation of searches for DM production requires the inclusion of additional on-shell mediating particles is considered. Adding the mediator back in to the particle spectrum, however, requires specifying the details of the interactions, which were safely ignored in the previous searches at colliders, includinf RunI of LHC, based on Effective Field Theories (EFTs) approach. As an intermediate step between EFTs and a fully-specific UV theory, *simplified models* which have a small number of assumptions about the dark matter and contain the minimal particle content were developed to allow for accurate interpretation.

It is of particular interest to investigate interactions mediated by a new neutral scalar or pseudoscalar particle that decays into DM particles, as these can be easily accommodated in models containing extended Higgs boson sectors.

Given that couplings between spin-0 particles and SM fermions requires some amount of $SU(2)_L \otimes U(1)_Y$ breaking, it is reasonable to construct models where the scalar or pseudoscalar coupling to the SM fermions is weighted by the SM Yukawa couplings [68]. Assuming minimal flavor violation [69, 70] (MFV), the discovery potential for scalar and pseudoscalar interactions in the monojet channel, mediated by top-quark loops similar to gluon-fusion Higgs production, is significantly improved when considering processes where the dark matter couples to massive third generation quarks, in particular top quarks. This has motivated analyses searching for events in which the dark matter particles are produced in association with a pair of top quarks ($t\bar{t} + DM$) or with one or two bottom quarks ($b(b) + DM$). The top quark pair channel represents the most sensitive channel at low mediator masses, and directly probes couplings of the mediator to top quarks, while monojet processes are sensitive to such couplings only through loops, this model predicts additional production mechanisms for dark matter particles, created along with a single top quark ($t/\bar{t} + DM$), rather than a pair. The main production diagrams for this single top process are shown in fig. 3.7. The production of the single top quark is obtained through processes mediated by a virtual t-channel or s-channel W boson, (a) and (b) respectively, or through the associated production with a W boson, (c) and (d). So far, final states involving a single top quark and missing energy transverse from dark matter particles have been studied only considering flavor-changing neutral interactions.

Assuming this DM scenario the interactions of the new spin-0 mediator par-

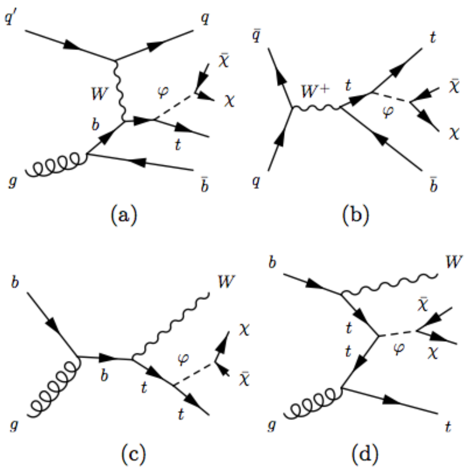


Figure 3.7: Main production diagrams for the associated production of dark matter with a single top at the LHC: (a) s-channel W boson production, (b) t-channel W boson production, and (c)-(d) associated tW production [67].

ticle follow the same Yukawa coupling structure as in the SM. Therefore, the mediator would couple preferentially to heavy third-generation quarks. Assuming the DM particles to be Dirac fermions, the interaction Lagrangian terms for the production of a scalar (φ) or pseudoscalar (a) mediator particle can be expressed as:

$$\mathcal{L}_\varphi \supset g_\chi \varphi_{\chi\bar{\chi}} + \frac{g_v \varphi}{\sqrt{2}} \sum_f (y_f \bar{f} f), \tag{3.7}$$

$$\mathcal{L}_a \supset i g_\chi a \bar{\chi} \gamma^5 \chi + i \frac{g_v a}{\sqrt{2}} \sum_f (y_f \bar{f} \gamma^5 f), \tag{3.8}$$

where, the sum runs over the SM fermions f, $y_f = \sqrt{2} m_f / v$ are the Yukawa couplings with the Higgs field vacuum expectation value $v = 246$ GeV, g_χ is the DM-mediator coupling, and g_v is the fermion-mediator coupling. Under the MFV assumption, this simplified model has a minimal set of four free parameters: $(m_\chi, m_\varphi, g_\chi, g_v)$, it is important to emphasize that these four variables represent the minimum set of parameters necessary for the comparison of collider and DD experiments.

3.2.3 Search for DM

Searches for DM can be roughly classified in *direct* experiments, *indirect* experiments, and searches at *colliders*.

- Direct detection experiments mostly aim to observe elastic or inelastic scattering of Galactic DM particles with atomic nuclei, or with electrons in the detector material. Predicted event rates assume a certain mass and scattering cross section, as well as a set of astrophysical parameters. For DM scattering off nuclei, the differential scattering rate

R as a function of nuclear recoil energy E_R is:

$$\frac{dR(E_R,t)}{dE_R} = N_T \frac{\rho_0}{m_{DM}} \int_{v>v_{min}} v f(\vec{v}+\vec{v}_E(t)) \frac{d\sigma(E_R,v)}{dE_R} d^3v, \tag{3.9}$$

where N_T is the number of target nuclei, m_{DM} is the mass of the DM particles, $v = |\vec{v}|$ is the speed of the particle in the experiment's rest frame, $f(\vec{v}+\vec{v}_E(t))$ is the velocity distribution in the Earth's frame, v_{min} is the minimum speed of the DM particles that can cause a recoil energy E_R and σ is the scattering cross section on the nucleus. For DM particle masses below the GeV-scale, most searches for DM-nucleus scattering rapidly lose sensitivity, due to energy thresholds around $10^2 - 10^3$ eV.

Another strategy is to search for DM scattering off bound electrons, allowing for all of the kinetic energy (50 eV in the above case) to be transferred to the material. The leading possibilities are ionisation, excitation, and molecular dissociation processes, which typically require energies of (1- 10) eV, and thus allow to probe scattering of DM particles with masses down to the $\mathcal{O}(\text{ MeV})$ range. The signal depends on the material, and can consist of one or more electrons (in semiconductors, noble liquids, graphene), one or more photons (in scintillators) or phonons (in superconductors and superfluids) and quasiparticles (in superconductors). As an example, the differential event rate for ionisation in atoms is given by

$$\frac{dR_{ion}}{d\ln E_R} = N_T \frac{\rho_0}{m_{DM}} \frac{d\langle\sigma_{ion} v\rangle}{d\ln E_R}, \tag{3.10}$$

where E_R is the recoil energy transferred to the electron, $\langle\sigma_{ion} v\rangle$ is the thermally averaged ionisation cross section and N_T is the number of target atoms per unit mass.

Direct detection experiments aim to observe the small, keV-scale and below, and rare, fewer than ~ 1 event/(kg $\cdot$ y), signals which are induced by DM particle scatters in a detector, mostly in the form of ionisation, scintillation or lattice vibrations. A majority of experiments detects more than one signal, which allows to distinguish between scattering off of electrons (electronic recoils, ER) and off of atomic nuclei (nuclear recoils, NR). A 3D position resolution is required to define central detector regions, or fiducial volumes, with low background rates from surrounding materials, and the distinction between single- versus multiple-scatters rejects a significant fraction of backgrounds, given that DM will scatter at most once. Fig. 3.8 shows the best constraints for SI couplings in the cross section versus DM mass parameter space, above masses of 0.3 GeV.

- Indirect DM detection refers to the search for the annihilation or decay debris from DM particles, resulting in detectable species, including especially gamma rays, neutrinos, and antimatter particles. The production rate of such particles depends on the annihilation, or decay, rate,

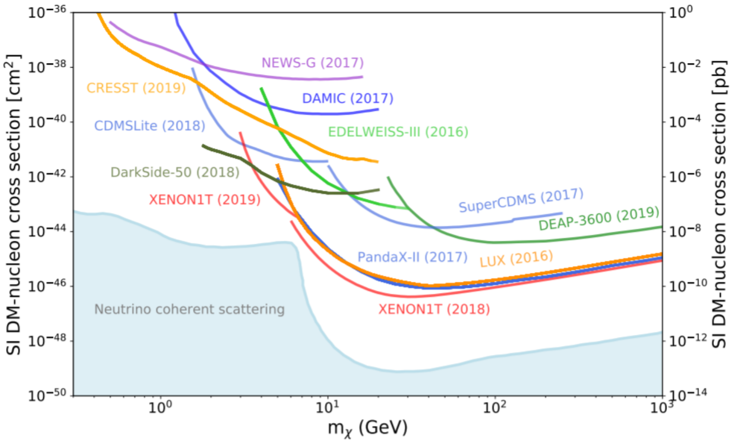


Figure 3.8: Current exclusion limits and region of interest of DM obtained with direct searches [22].

on the density of pairs in the region of interest, and on the number of final-state particles produced in one annihilation, decay, event.

DM annihilation to virtually any final state produces gamma rays: emission processes include the dominant two-photon decay mode of neutral pions resulting from the hadronization of strongly-interacting final states; final state radiation; and internal bremsshtralung, the latter two including, possibly, the emission of massive gauge or Higgs bosons subsequently producing photons via their decay products. Similarly, neutrinos are produced from charged pion decay and from radiative processes.

DM can be captured in celestial bodies in significant amounts, depending on the DM scattering cross section off of nucleons, the DM mass, and the DM flux incident on the celestial body of interest. For DM masses at or around the GeV scale, evaporation from the celestial body plays an important role. If enough DM accumulates, DM annihilation inside the celestial body can then lead to the production of Standard Model particles. Such particles can heat up the body if they lose most of their energy before escaping. Utilizing models for heat production in planets, or stellar interior models in the case of stars, constraints can be put on DM particle properties. Alternately, DM annihilation in celestial bodies can result in the production of particles that can escape the body. Within the Standard Model, the only such instance is annihilation to neutrinos, but, similarly to the boosted DM case the DM can annihilate to a (stable or unstable) dark-sector particles, whose decay or interactions can be detected on Earth.

Stable charged particles produced by decays as consequence of processes involving DM annihilation or decay populate the cosmic radiation and are a prime target for indirect DM searches. To maximize rate, searches focus on relatively rare particle species, such as positrons, antiprotons, and antinuclei. While in certain models the production of particles and antiparticles is not symmetric [71], generally DM annihilation or decay produces as many particles as antiparticles in the final state.

- Searches for DM with the LHC and other colliders have targeted DM models that interact with the SM via Higgs or Z boson exchange, effective field theories with heavy mediators, UVcomplete models such as supersymmetry, models with long-lived particles, and models with rich dark sectors. Various searches for dark matter have been carried out by the CMS and ATLAS collaborations at the LHC in pp collisions [72, 73].
  The experimental program correspondingly includes searches for invisible particle production mediated by a SM boson, generic searches for invisible particles produced via new particle mediators,as seen in 3.2.2, and specific searches for complete models. There are a variety of signals for DM: the imbalance in the transverse momentum in an event due to the presence of DM particles, produced together with one Standard Model particle, a bump in the di-jet or di-lepton invariant mass distributions, or an excess of events in the di-jet angular distribution, produced by a dark matter mediator. In this thesis we are interested in the production of invisible particles in association with a single top-quark, such a final state is characterised by a significant missing transverse momentum, which is due to the undetected particles. Background contributions from SM processes [74] are expected to be small. Similar searches were previously conducted by the CDF Collaboration using 7.7 fb^{-1} of Tevatron $p\bar{p}$ collisions at $\sqrt{s} = 1.96\ \mathrm{TeV}$ [75] and by the ATLAS and CMS collaborations using $\sqrt{s} = 8\ \mathrm{TeV}$ [76, 77] and $\sqrt{s} = 13\ \mathrm{TeV}$ [78] LHC data.
  No signal for DM has been observed in the LHC experiments so far. Instead limits are set on masses, couplings, and cross-sections.

Fig. 3.9 shows a comparison beetween DM search at CMS experiment, through production in association with a pair of top quarks, and the results of XENON1T, LUX, PandaX-II, CDMSlite, and CRESST-II, which provide the strongest constrains for DM direct searches.

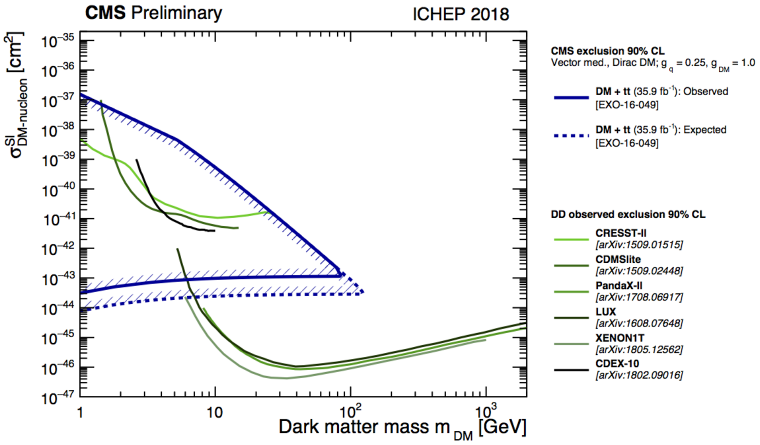


Figure 3.9: A comparison of CMS results to the $m_{DM} - \sigma_{SI}$ plane .The CMS contour in the SI plane is for a Scalar mediator, Dirac DM and couplings $g_q = 0.25$ and $g_{DM} = 1.0$, and for DM particles produced in association with a pair of top quarks. The CMS SI exclusion contour is compared with the XENON1T 2017, LUX 2016, PandaX -II 2016 , CDMSLite 2015 and CRESST-II 2015 limits, which constitutes the strongest documented constraints in the shown mass range [79] .

Chapter 4

Physics objects selection and recostruction

The main aim of this thesis is the search for a single top quark production in association with missing energy in the transverse plane with respect to the beam axis, in particular considering the hadronic decay of top quark. The search has been performed using pp collision data collected in 2018 and corresponding to an integrated luminosity of 59.7 fb^{-1} and its performances have been extrapolated to the full RunII.
Two different interpretations are considered for the signature of interest:

- the decay of a heavy Vector-Like quark T in the channel $T \rightarrow t + Z$, followed by the Z boson decaying into neutrinos. In fig. 4.1 is shown the Feynman diagram of the process. Higher masses for the T are accessible with respect to the RunI. As a consequence, the Z boson exhibits a higher Lorentz boost with respect to 8 TeV collisions of *RunI*, and the corresponding neutrinos stem very collimated along the Z boson axis. This results in signal events characterized by large missing energy. In addition, the branching ratio of $Z \rightarrow \nu\bar{\nu}$ is ~ 3 times the branching ratio of the dilepton channel, which makes this channel a very good probe for T quark production searches;

- DM production in association with a single top quark. In fig. 4.2 is shown the Feynman diagram of the process. This channel is considered to improve the previous searches that considered the channel $t\bar{t} + DM$. Different mass hypothisis for the scalar or pseudoscalar mediator can be probed, according to centre-of-mass energy. At every φ mass value corresponds an allowed range of χ mass. In the model considered, χ mass is suppose to be equal to 1 GeV because the cross-section of the process depends mainly from the mediator mass and its spin nature. The aim of this work is to expand the probed mass of φ until 1000 GeV in the scalar mediator hypothesis.

Two channels for the single-top quark signature are considered, targeting the case in which the W boson originating from the top-quark decays into a pair of quarks, *hadronic channel*. The objects in the detector are, therefore, stable hadrons, clustering into $jets$, and invisible particles.

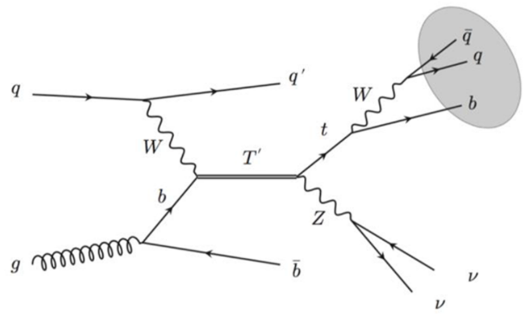


Figure 4.1: Feynman diagram for the T production and decay channel searched for.

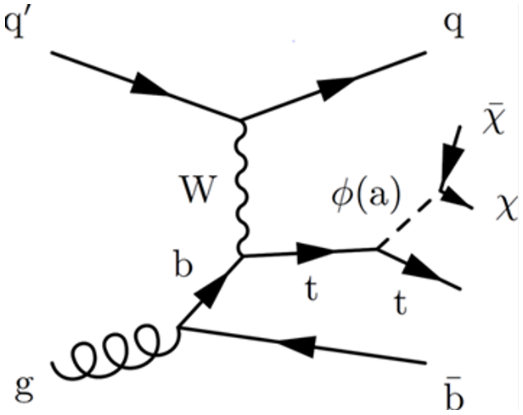


Figure 4.2: Feynman diagram for $t + DM$ channel searched for.

In this chapter we will describe the reconstruction of the final state objects in the detector.

4.1 Physics objects selection

Starting from the beam interaction region, particles first enter a silicon tracker, which allows to reconstruct charged-particle trajectories, i.e. tracks, and their originating points, i.e. vertices. The vertices and tracks are reconstructed from signals in the sensitive layers. The magnetic field bends the trajectories and allows the electric charges and momenta of charged particles to be measured. Electrons and photons are then absorbed in the ECAL. The corresponding electromagnetic showers are detected as clusters of energy recorded in neighbouring cells, from which the energy and direction of the particles can be determined. Charged and neutral hadrons may initiate a hadronic shower in the ECAL as well, which is subsequently fully absorbed in the HCAL. The corresponding clusters are used to estimate their energies and directions. Muons and neutrinos traverse the calorimeters with little or no interactions. While neutrinos escape undetected, muons produce signals in the muon system, located outside the calorimeters.
All the stable particles are reconstructed with the Particle Flow [28, 80] (PF) algorithm, it combines the information from the CMS sub-detectors to iden-

tify the particles and also to reconstract their 4-momenta.

4.2 Jets

The search is focused on the hadronic channel, therefore in the final state we want to to reconstruct the hadronization products of three quarks, two originating from the W boson and one b-quark from the direct top quark decay.

The quarks and gluons produced in the final state originate a shower of particles, through a process called *hadronisation*, which takes place before the particle directly interacts with the detector. The quarks in the final state are reconstructed indirectly through objects called *jets*, obtained by clustering the products of the hadronisation of quark and gluons through appropriate algorithms. The interaction between constituent partons and the showering into stable particles is well described by the perturbative theory and hadronization model.

The jets used for this analysis are provided by the anti-k_T algorithm [81]. Anti-k_T algorithm is the default choice at LHC, because of its robustness against pile up and underlying event contributions, moreover, the algorithm is IRC safe, soft-resilient jet algorithm, that leads to jets whose shape is not influenced by soft radiation.

In the momentum space are considered: d_{ij},which is the distance between the particle of the PF candidate i and the particle j, and d_{iB}, which is the distance between the particle i and the beam B. In formulae:

$$d_{ij} = min(\frac{1}{p_{T,i}^2}, \frac{1}{p_{T,j}^2})\frac{(\Delta R_{ij})^2}{R^2} \tag{4.1}$$

$$d_{iB} = \frac{1}{p_{T,i}^2}, \tag{4.2}$$

where R is the radius parameter, ΔR_{ij} is the distance in the (ϕ, η) plane of the particles i and j, and $p_{T,h}$ is the transverse momentum of particle h.

The algorithm proceeds by identifying the smallest among distances and if it is a d_{ij} recombining entities i and j, while if it is d_{iB} calling i a jet and removing it from the list of entities. The distances are recalculated and the procedure repeated until no entities are left. Two different kinds of jets have been used for the analysis, jets with $R = 0.4$ are the AK4 jets, also referred as *narrow jets*, and jets with $R = 0.8$ are the AK8, or *fat jets*. A set of identification criteria provided by the CMS dedicated group that works on Jets and Missing Energy Transverse (MET), is applied on AK4 jets collection. Moreover, different corrections are applied while calculating the energy of the jets at various levels: Jet Energy Corrections (JEC) are used to scale the jet energy by a factor that describes the detector response depending on the transverse energy and the pseudorapidity of the jet; 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 and Pile-Up Per particle Identification (PUPPI) used with CHS for pileup-rejection, PUPPI PF jets have improved jet p_T resolution for low jet p_T where pileup plays an important role; 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.

4.2.1 b-tagging

AK4 jets may be *tagged* as arising from a b quark (*b-jet*), indeed a b-jet is expected in the final state considered.
Several algorithms for identification of jets originating from b quarks are available in CMS, one of these is the Combined Secondary Vertex (CSV) algorithm [82]. 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, fig. 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 $c\tau \sim 450\ \mu\mathrm{m}$. In CMS the Impact Parameter can be measured with precision between $30\ \mu\mathrm{m}$ and hundreds $\mu\mathrm{m}$. The AK4 jets are usually considered coming from b-quark if they pass a given threshold on the value of DeepCSV, which is defined by the CMS group working on b-tagging based on the selection efficiency of b-originated jets and the mis-tagging efficiency of light quark-originated jets.
In this work DeepJet algorithm [83] is used for b-tagging discrimination, it is based on DNN algorithm and it is trained with more variables with respect to CSV algorithm, such as global variables (jet kinematics, the number of tracks in the jet, etc...), charged PF candidate features (containing information on the track kinematics, track fit, quality, displacement, and displacement uncertainty), neutral PF candidate features (such as the flight distance significance), and SV features associated with the jet.

4.2.2 Deep tagging

An efficient classification of hadronic decays of heavy SM particles that are reconstructed within a single jet can be done. An algorithm implemented in CMS allows to distinguish hadronically decaying massive SM particles, namely W, Z, H bosons, and top quarks, from other jets originating from lighter quarks or gluons.
In this thesis we used recently developed ML-based CMS heavy-object tagging method [84] to exploit the full potential of the detector and event reconstruction, *DeepAK8 algorithm*. The algorithm is a multiclass classifier based on CNN algorithm for the identification of hadronically decaying particles with five main categories: W, Z, H, t, other.
In the DeepAK8 algorithm, two lists of inputs are defined for each jet. The first list consists of up to 100 jet constituent particles, sorted by decreasing

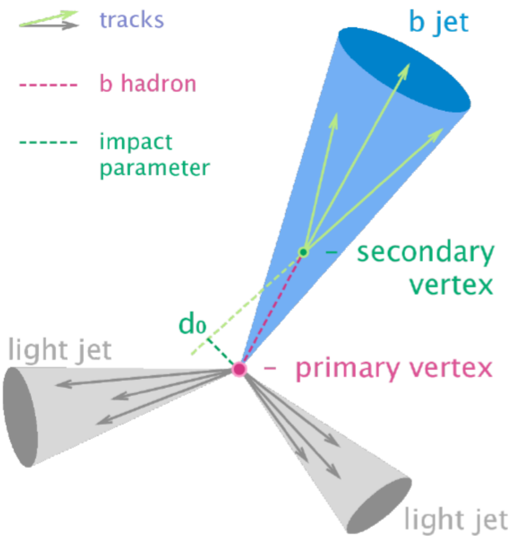


Figure 4.3: Impact parameter [82].

p_T. The list includes measured properties of each particle, such as the p_T, the energy deposit, the charge, the angular separation between the particle and the jet axis or the subjet axes, etc., these are included to help the algorithm extract features related to the substructure of the jet.

The second list called secondary vertex (SV) list consists of up to 7 SVs, each with 15 features, such as the SV kinematics, the displacement, and quality criteria. The SV list helps the network to extract features related to the heavy-flavor content of the jet. The Receiving Operating Curves (ROCs) of DeepAK8 algorithm are shown in fig. 4.4 with the ROCs of other algorithms.

4.3 Muons

The final state of interest for this thesis is characterised exclusively by quarks and invisible particles, therefore no high-energetic muon or electron is expected to come from the hard interaction. One of the possible backgrounds arises from $t\bar{t}$ events and electroweak processes, like single or double vector boson production. The contribution from $t\bar{t}$ events arises in the semi-leptonic decays, when one top quark decays in $t \rightarrow bW \rightarrow b\ell\nu$, therefore a lepton veto is applied to reject event with the production of an electron or a muon near the Interaction Point. The identification and isolation criteria used for muons are described here, while for the electrons in the next section.

The muons are detected both in the tracking system and in the muon system [85]. In the first detector they are reconstructed using a technique based on the Kalman filter algorithm, in the second one the hits from DTs, CSCs and RPCs are fitted reconstructing the Stand Alone Muons. The muons could be reconstructed as Tracker Muon if the muons are reconstructed in the tracker system, and the track is matched with a single hit in the muon system.

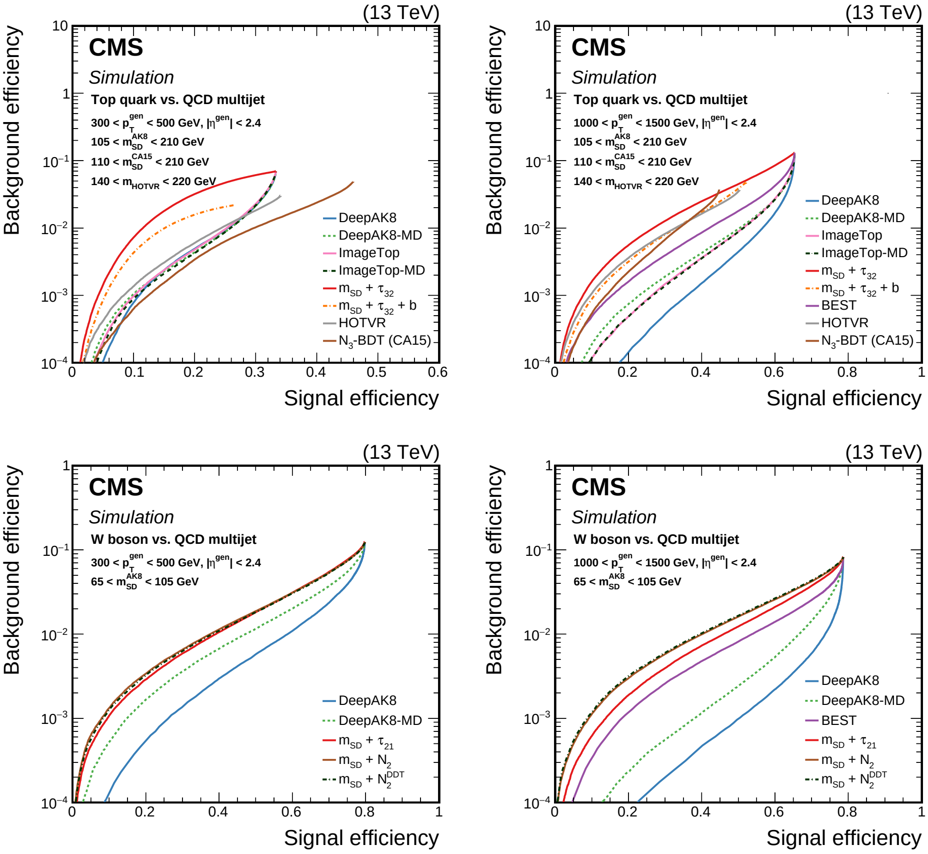


Figure 4.4: Comparison of the identification algorithms for hadronically decaying t quark and W boson in terms of ROC curves in two regions based on the p_T of the generated particle; Left: $300 < p_T < 500$ GeV, and Right: $1000 < p_T < 1500$ GeV [84].

The Global Muon reconstruction is given by the refitting of two tracks: a Tracker Muon, and Stand Alone Muon, that are matched together. This different kind of reconstruction defined identification criteria used in CMS, together with other variables such as the number of pixels hit.
Muon candidates are selected according to the veto selection criteria defined in CMS according to identification criteria with high efficiency but low purity. The isolation variable I^{μ}_{rel} is defined for muon candidates as:

$$I^{\mu}_{rel} = \frac{I^{ch.h} + max\big((I^{\gamma} + I^{n,h,} - 0.5 \times I^{PU}), 0\big)}{p_T}, \tag{4.3}$$

where $I^{ch.h}$, I^{γ}, $I^{n.h.}$, 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 $\Delta R = 0.4$ around the muon direction. For the lepton veto only tight muons, table 4.1, are considered.

Tight muon		
p_T	$\lvert\eta\rvert$	Loose Id
$> 30GeV$	< 2.4	

Table 4.1: Tight muon.

Tight electron	
p_T	Loose Id
$> 30GeV$	

Table 4.2: Tight electron.

4.4 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.
The isolation variable I^{e}_{rel} for electron candidates is defined as:

$$I^{e}_{rel} = \frac{I^{ch.h} + max\big((I^{\gamma} + I^{n,h,} - \rho \times A)\big)}{p_T}, \tag{4.4}$$

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 $\Delta R = 0.3$ around the electron direction. For the lepton veto only tight electrons, table 4.2, are considered.

4.5 MET

At the LHC, and in general at any hadron collider, the energy imbalance is computed in the transverse view only. The reason for this is that only the longitudinal momentum of the whole protons can be measured, while the one of the individual colliding partons is not known. The particle flow algorithm reconstructs the missing transverse energy, MET,as the opposite of the vectorial sum of all the candidates transverse momenta[*].
The MET is a crucial variable for this search because the final state is characterized by particles that do not leave any trace in the detector.
The presence of these non-interacting particles in the detector can be measured with MET since they create an energy imbalance when escaping the detector, not allowing kinematic closure in the transverse plane.

---

[*] It is worthy to observe that, what usually refer to as MET is actually the missing transverse momentum, for this reason in the next we will mention the p_T^{miss} as p_T^{MET}.

The MET is also one of the most complex variables to reconstruct, since it is very sensitive to detector malfunctions, energy resolution of all particles (hadrons, leptons and photons), 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, and are described in the next chapter.

4.6 Top quark reconstruction

In order to analyse the kinematics of singly produced top quarks, the four-vector of the top quarks momenta have to be reconstructed from the decay products. Top quark has a peculiar signature thanks to its very short lifetime ($\sim 10^{-25}$ s), as it can always be assumed to decay $t \rightarrow bW$ before hadronisation. The reconstruction is done through the hadronic channel, which means the one involving a W decaying in a couple of quarks.
Three reconstruction methods for the top quark are used in this analysis:

- top merged, where all three quarks are reconstructed in only one fatjet (top-tagged jet), see also fig. 4.6 ;
- top partially-merged, where the quarks are reconstructed in a fatjet originated from the W (W-tagged jet), and a narrow jet from the b quark (b-tagged jet), see also fig. 4.7;
- top resolved, where the quarks are reconstructed in three distinct narrow jets, one per each quark, with at least one b-tagged jet, see also fig. 4.10.

Top quark decay products can overlap completely or partially. In order to consider all possible cases, this analysis looks for events with AK8 jets and AK4 jets, identified as originating from top quarks, using a Machine Learning approch to improve the selection provided by the tagging algorithms.

4.6.1 Boosted Decision Trees

For the purpose of reconstructing and identifyng top quarks starting from the jets elaborated by the anti-k_T algorithm a Machine Learning approach was used.
A Machine Learning algorithm receives as input a set of discriminating variables, each of which individually does not allow to reach an optimal selection power. The algorithm computes an output that combines the input variables. The discriminant output value is taken as a test statistic and is then adopted to perform the signal selection, which is implemented as a cut on the value of the discriminant.
The computation of the output value, given the input variables, is based on

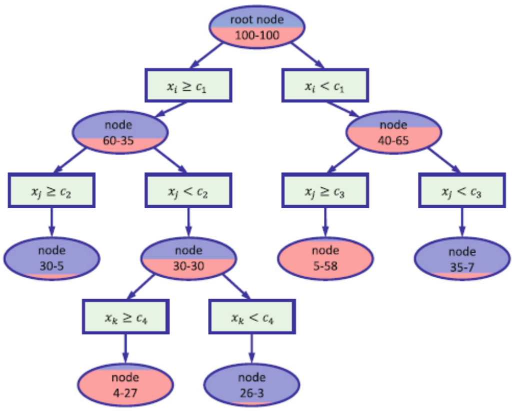


Figure 4.5: An example of a decision tree. Each node represented as an ellipse contains a different number of signal (left number) and background (right number) observations. Applied requirement are represented as rectangular boxes [87].

a number of parameters which can often be very large. The choice of the parameter values is a key task of the algorithm since an optimal choice allows achieving the best possible performances. The usual strategy consists in tuning the discriminant parameters providing as input to the algorithm large datasets distributed according to the hypothesis allowed. By comparing the discriminant output to the true origin of the dataset, the parameters are modified. This process is called training, and algorithms that use such training samples are called *supervised algorithms*.

For Top quark reconstruction a *Boosted Decision Tree* [86] was employed, it is a supervised algorithm based on elementary units called decision trees.

A decision tree is a sequence of selection cuts that are applied in a specified order on a given variable datasets. Each cut splits the sample into nodes, each of which corresponds to a given number of observations classified as signal or as background. A node may be further split by the application of the subsequent cut in the tree. Nodes in which either signal or background is largely dominant are classified as leafs, and no further selection is applied. A node may also be classified as leaf, and the selection path is stopped, in case too few observations per node remain, or in case the total number of identified nodes is too large, and different criteria have been proposed and applied in real implementations. A schematic view of the process is showed in fig. 4.5. Selection cuts can be tuned in order to achieve the best split level in each node according to some metrics, for example Gini index or cross entropy, and then the gain due to the splitting of a node A into the nodes B_1 and B_2, which depends on the chosen cut, is given by:

$$\Delta I = I(A) - I(B_1) - I(B_2), \tag{4.5}$$

where I denoted the adopted metric.

A Boosted Decision tree improves the learning process of a Decision Tree iteratively adding a new tree to a *forest* accordly to the *Boosting procedure*. The boosted procedure consists in different steps:

- Training observations are reweighted using the results of the previous trees.
- A new tree is built and optimized using the reweighted observations as a training sample.
- A score is given to each tree.
- The output of the final BDT classifier is the weighted average over all trees in the forest:

$$y(\vec{x}) = \sum_{k=1}^{N_{trees}} w_i C^{(i)}(\vec{x}), \tag{4.6}$$

  where $y(\vec{x})$ is the final output and $C^{(k)}(\vec{x})$ is the output of the k-tree.

The boosted procedure used is called the *XGBoost*, the different boosting depend on the different methods used to compute the weight w.
To improve the efficiency of the BDT algorithm the following parameters have been set:

- maximum depth of a tree = 4;
- learning rate = 0.1;
- the minimum sum of instance weight needed in a child = 4;
- number of trees equal = 100;
- L1 regularization term on weights (α) = 0.01.

Three separate training sessions have been performed for the configuration in 4.6, in addition each configuration is divides in three sub-categories based on p_T of the Top quark candidates: low p_T (p_T^{Top} <350 GeV), medium p_T (350< p_T^{Top} <700 GeV), high p_T (p_T^{Top} >700 GeV).
For the training MC datasets are used, in particular three VLQ T left-handed samples with three different mass hypotheses (700 GeV, 1000 GeV, and 1800 GeV) to ensure a large enogh dataset for each of Top quark configurations. As background also simulated signal events are used, jets did not match with a Top quark are selected.

4.6.2 Top quark merged

For the top quark merged configuration the entire AK8 jet collection is used. In the training samples, at most one fatjet for event is matched with a top quark that is labelled as *true* event, any other fatjet is labelled as *false* event[†].
The sets of variables used as input for BDT algorithm are:

[†] It is worthy to emphasize that the term *event* may lead to same confusion. In the context of the training *true* and *false* event refers to a single fatjet used during the learining process of ML algorithm, while in the context of the physics analysis usually refers to a collision event containing several jets and fatjets.

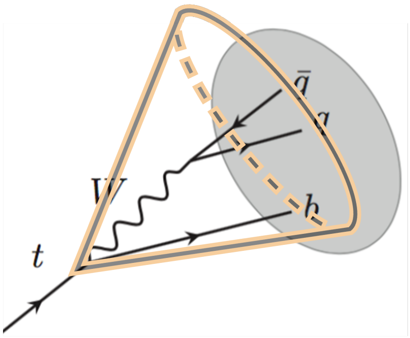


Figure 4.6: Top quark merged configuration.

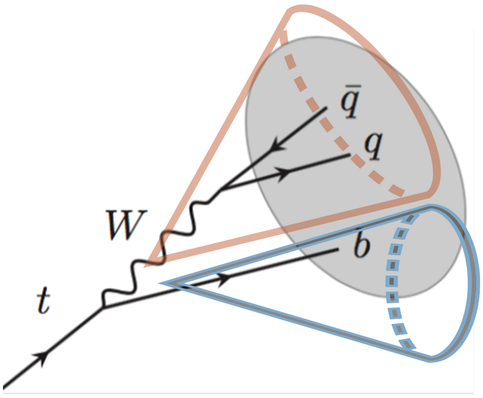


Figure 4.7: Top quark partially-merged configuration.

- Fatjet 4-momenta: p_T, ϕ, η, M;
- Fatjet top-tag discriminator.

The training results are shown in fig. 4.8 and 4.9. As one can expect, the top quark merged reconstruction works quite well because in this configuration the top-tag variable is an important discriminator that leads the training process facilitating its work, especially in the high p_T category when the quarks are produced more collimated.

4.6.3 Top quark partially-merged

For the top quark partially-merged configuration the entire AK8 jet collection is selected and one-by-one matched with an AK4 jet. Since narrow jets can be contained in a fatjet, every AK4 jet in the $R = 0.8$ cone around the

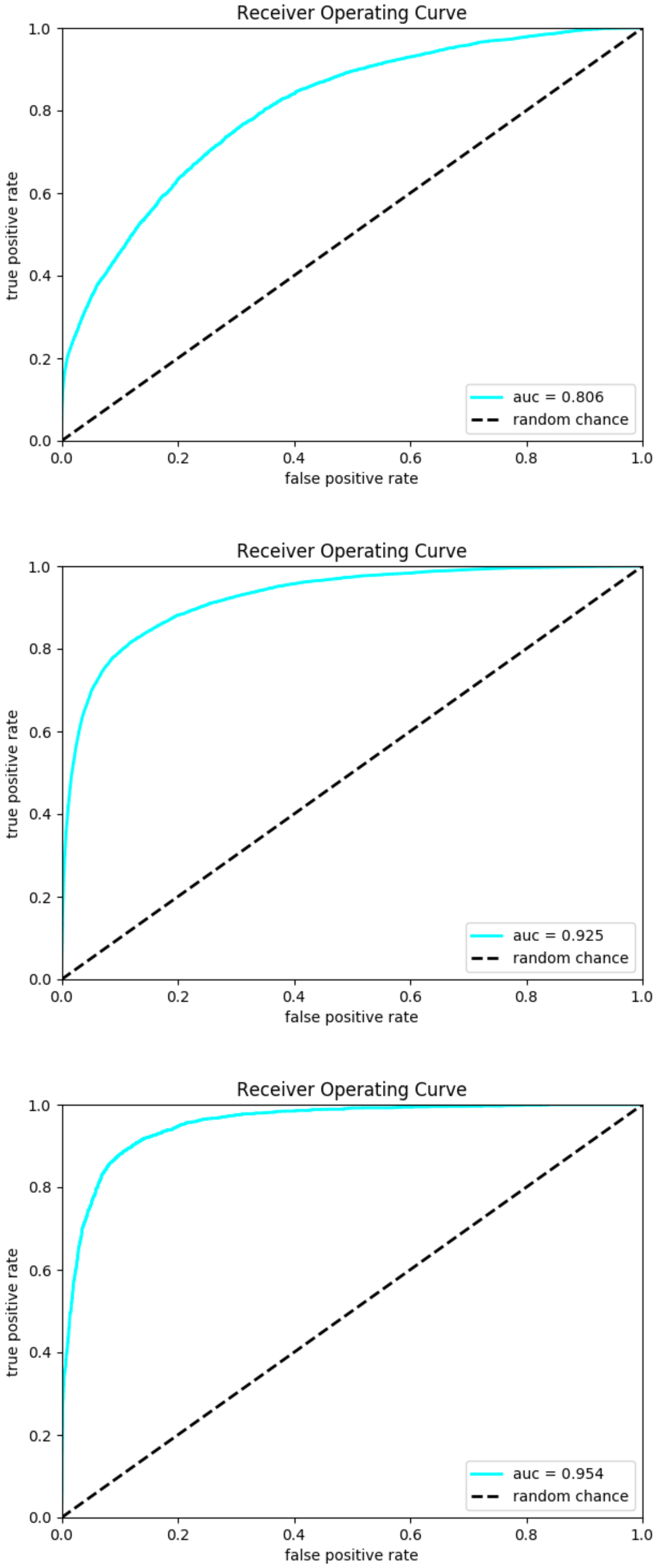


Figure 4.8: ROCs for the top quark merged configuration, from up: low p_T, medium p_T, high p_T.

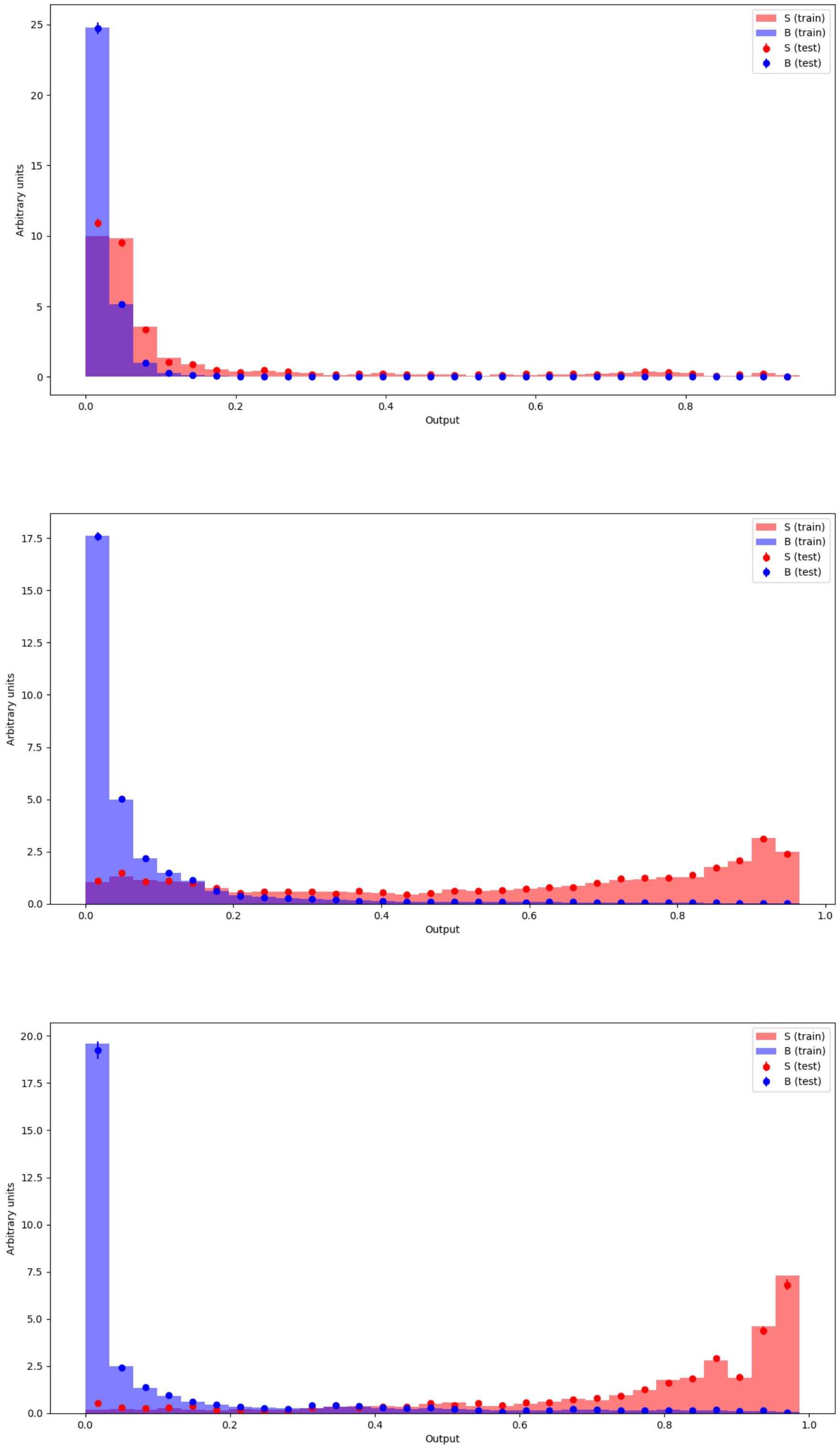


Figure 4.9: Histograms of output values of BDT models for Top merged. The output of train (filled bins) and test (points) datasets is reported distinguishing signal (red) and background (blue). From up to down low p_T, medium p_T and, high p_T.

fatjet centroid is rejected to avoid the overlap.
The fatjet has to be W-tagged while the narrow one b-tagged, the deepAK8 algorithm provide two continuous variables on which can be applyed a cut during the selection, but here we choose to used these variables as input in the BDT algorithm. The following variables are used:

- Fatjet 4-momenta;
- Fatjet W-tag discriminator;
- Narrow jet 4-momenta;
- Narrow jet b-tag discriminator;
- the summed momenta of the narrow jet and the fatjet.

In fig. 4.11 and 4.12 are shown the training results. This is the more complex category to reconstruct, how can be seen from the output value behavior. The descrimination is acceptable in low and medium p_T range.

4.6.4 Top quark resolved

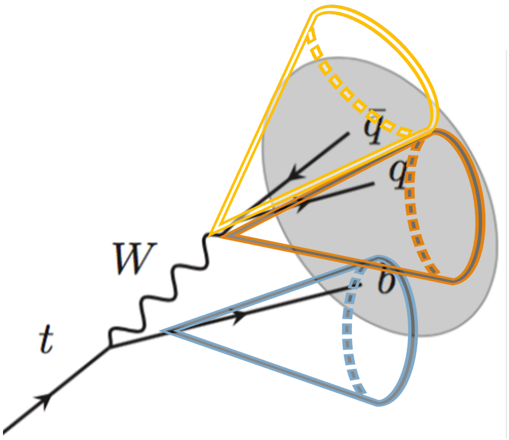


Figure 4.10: Top quark resolved configuration.

The last category described is the top quark resolved. For this one only AK4 jets are used, every triplet of jets is selected as top candidate.
For the reconstruction at least one b-tagged jet is required, the best b-tag variable of the three jets is used as input variables for the BDT algorithm. The variables adoperated are:

- First narrow jet 4-momenta;
- Second narrow jet 4-momenta;
- Third narrow jet 4-momenta;

- the summed 4-momenta of the three jets;
- the highest b-tag discriminator of the three jets.

Also in this case the training results are reported , fig. 4.14 and 4.13. The discrimination acts well enough in medium p_T range, but very poorly at high p_T. In the case however, this configuration is much less likely with respect to the other two.

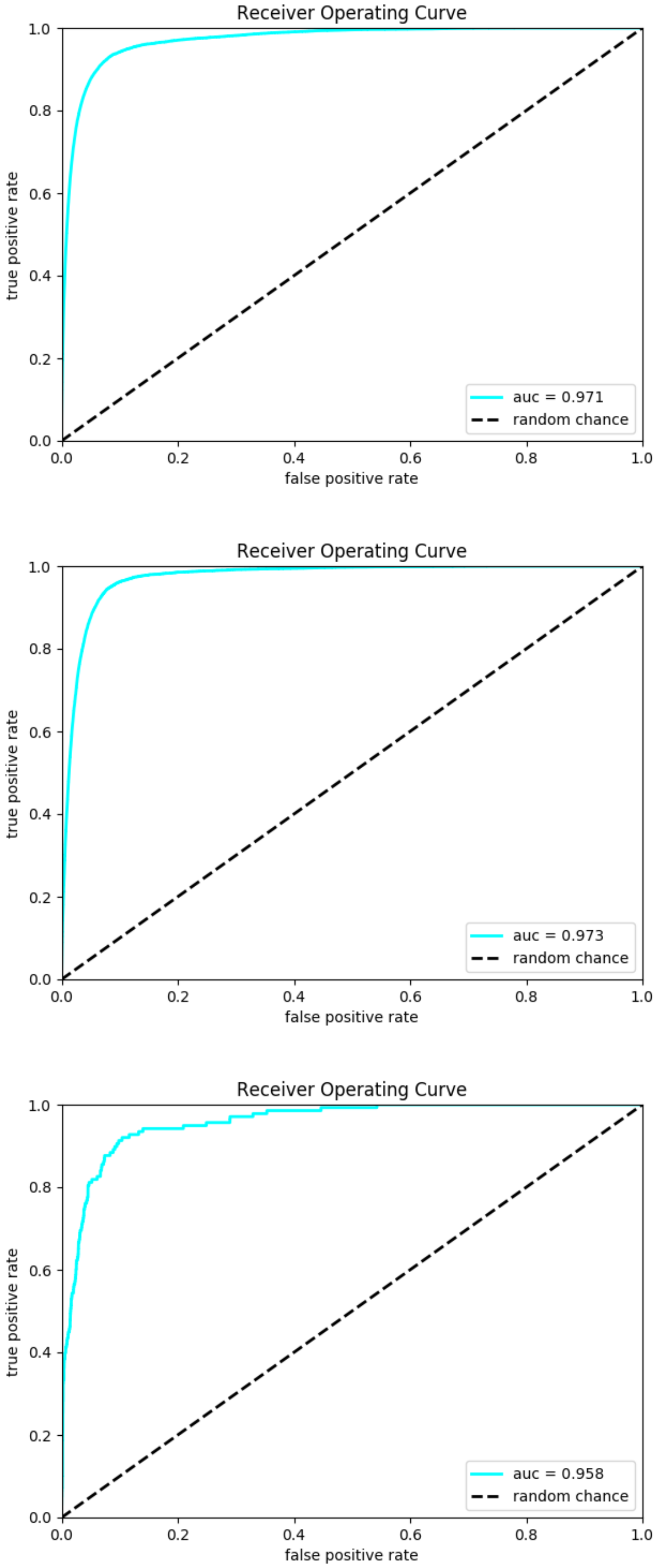
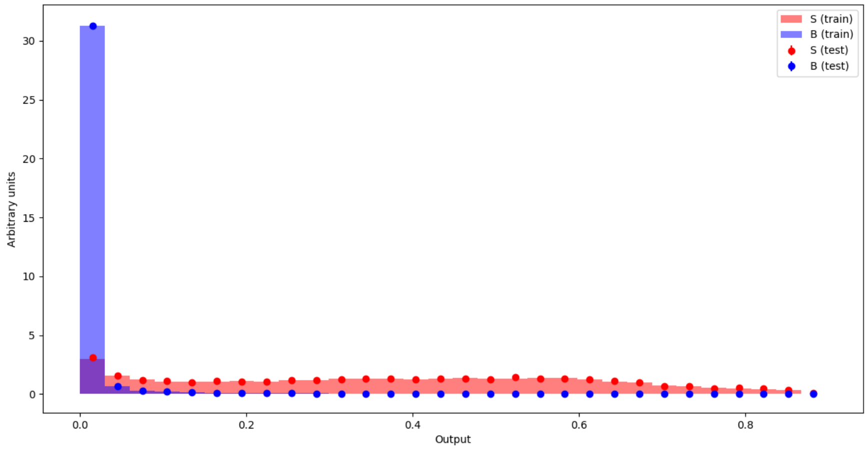
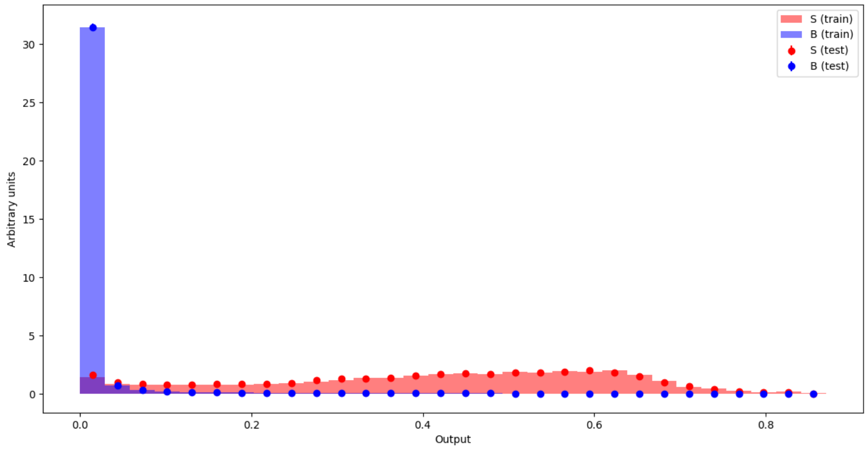


Figure 4.11: ROCs for the top quark partially-merged configuration, from up: low p_T, medium p_T, high p_T.





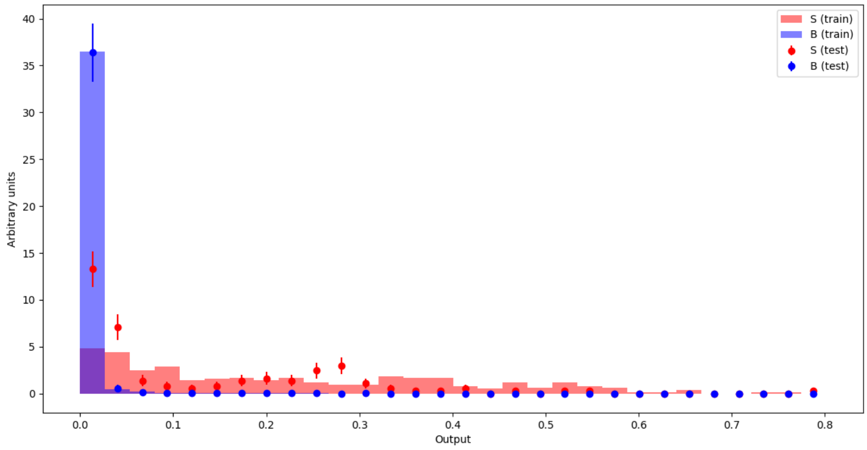


Figure 4.12: Histograms of output values of BDT models for top partially-merged. The output of train (filled bins) and test (points) datasets is reported distinguishing signal (red) and background (blue). From up: low p_T, medium p_T and, high p_T.

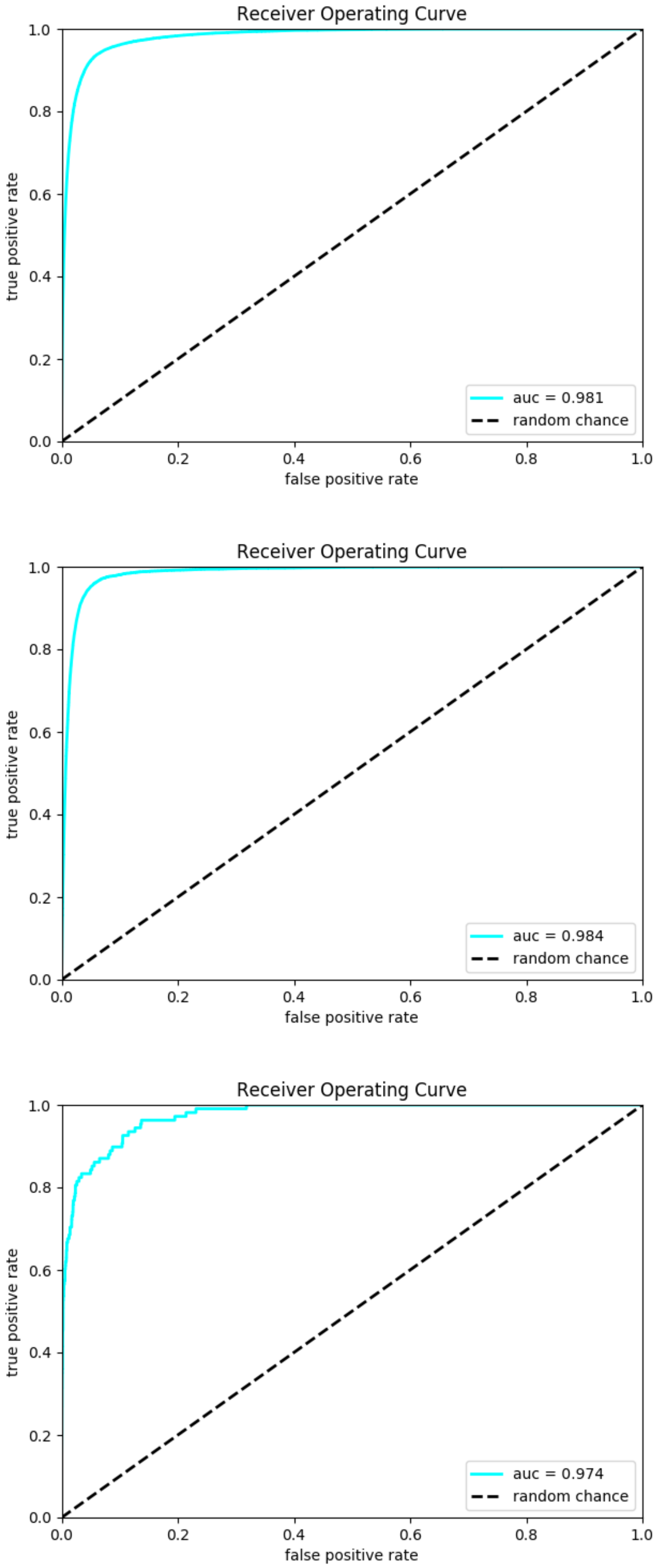


Figure 4.13: ROCs for the top quark resolved configuration, from up: low p_T, medium p_T, high p_T.

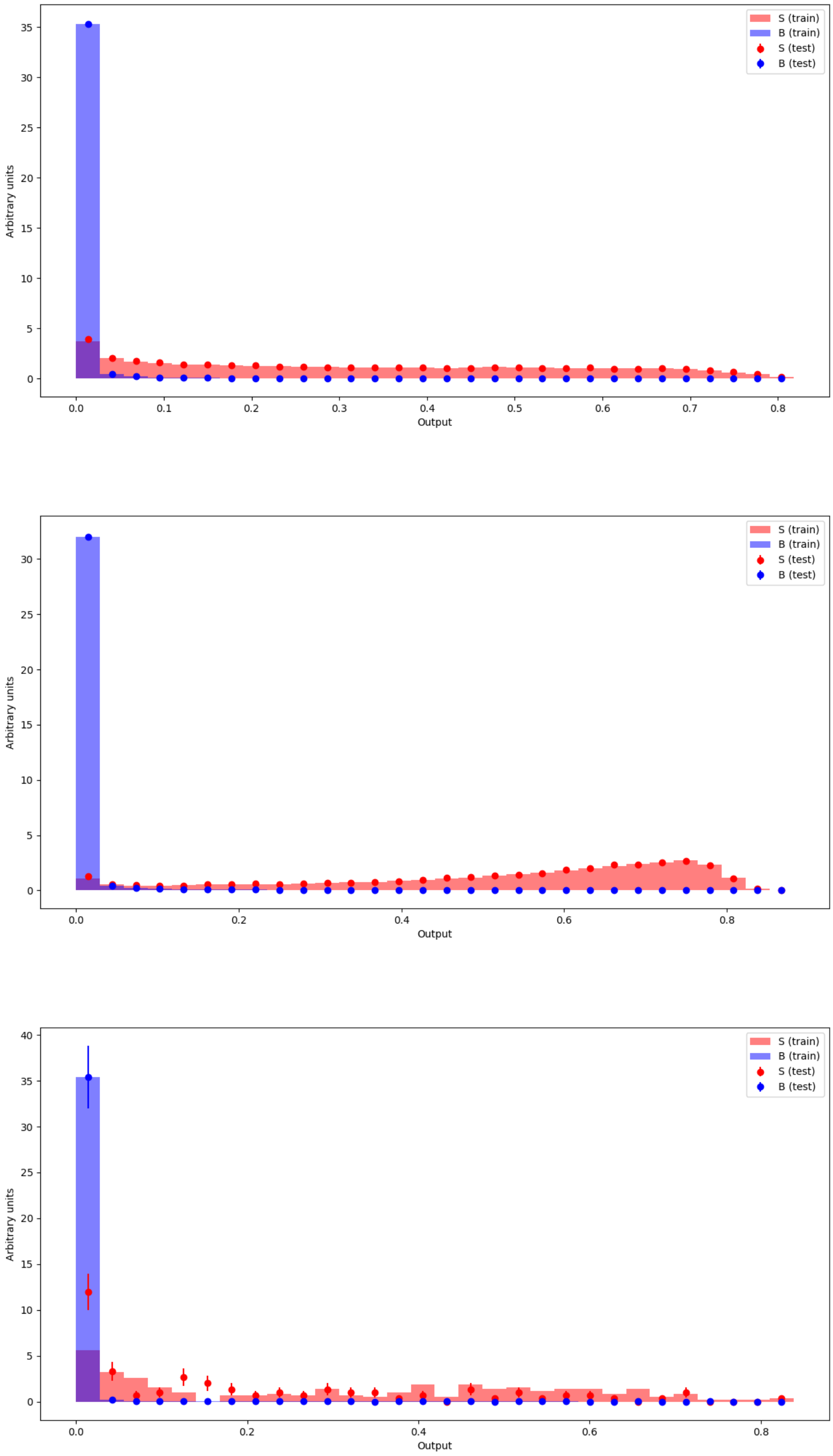


Figure 4.14: Histograms of output values of BDT models for top resolved. The output of train (filled bins) and test (points) datasets is reported distinguishing signal (red) and background (blue). From up: low p_T, medium p_T and, high p_T.

Chapter 5

Analysis strategy

In this chapter will be described the analysis strategy used to extract the number of signal events.

The data sample used has been collected by CMS in 2018, while simulated samples with the Monte Carlo (MC) technique have been used to estimate the background and the signal contribution. For the signal, several simulated samples have been used to probe different mass hypotheses for the Left-Handed VLQ T and for the φ scalr mediator in DM production.

First an event selection optimized for the signature of interest is described. The reconstructed transverse mass, defined later on in this chapter, is then used as discriminant to preform a maximum likelihood fit to data in the signal region in order to extract signal event yield.

In the end, 95% confidence-level upper limits on the corresponding production cross-sections are obtained.

5.1 Data and simulation samples

The analysis is performed using data collected by CMS in 2018. The data sets used are shown in table 5.1, where *JetHT* refers to the trigger used to select the data.

Dataset	Integrated luminosity (fb^{-1})
/JetHT/Run2018A-02Apr2020-v1/NANOAOD	14.00
/JetHT/Run2018B-02Apr2020-v1/NANOAOD	7.10
/JetHT/Run2018C-02Apr2020-v2/NANOAOD	6.94
/JetHT/Run2018D-02Apr2020-v1/NANOAOD	31.93

Table 5.1: List of pp collsion datasets collected by CMS in 2018 with corresponding luminosity.

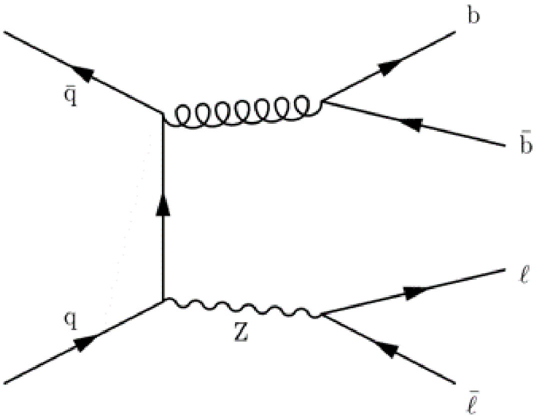


Figure 5.1: Feynman diagram of $Z + Jets$ background.

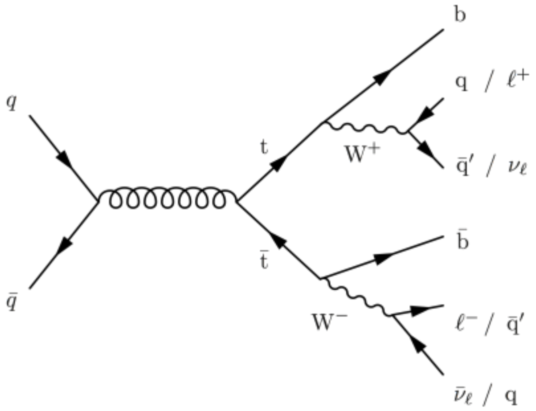


Figure 5.2: Feynman diagram of $t\bar{t}$ process.

5.1.1 Background description

The main background sources considered for this analysis are:

- $Z + Jets$, processes where a couplr of jets are produced in association with neutrinos coming from the decay of Z boson, fig. 5.1. This process rappresent the main background because the jets produced can be mistaken for a top quark;

- $t\bar{t}$ processes, pair production of a top quark and a top antiquark shown in fig. 5.2. In particular when one top quark decays through the leptonic channel, $t \to bW \to b\ell\nu$, the charged lepton might in some cases end failing the reconstruction or identification requirements mimicking, with the neutrino, the invisible part of the signal event.

- $W + Jets$ processes, fig. 5.3, can fake the signal topology in the same way as $t\bar{t}$ processes, when 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.

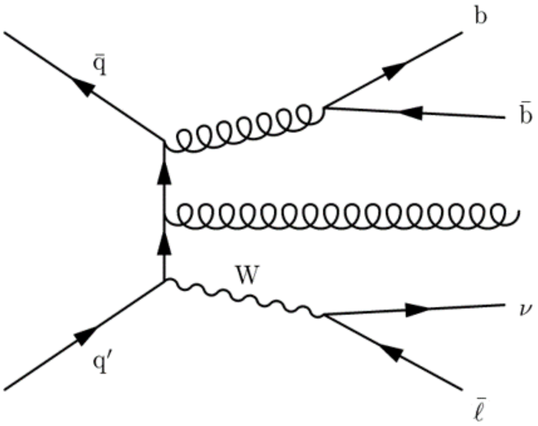


Figure 5.3: Feynman diagram of $W + Jets$ background.

5.1.2 Signal and background simulation

The signal and the main background events are taken from MC simulations using different software frameworks in order to:

- generate matrix elements either at leading order corrections (LO) or at next-to-leading order corrections (NLO), with Madgraph [88] or POWHEG [89];
- generate and simulate the hadronization of particles in the final state produced in pp collisions, in order to reproduce accurately the event properties of a wide range of processes, with Pythia [90] or Herwig [91];
- simulate particle interaction with CMS sub-detectors, with GEANT 4 [92].

The signal event samples are generated using Madgraph and Pythia, in order to have processes at leading-order (LO). Different event samples have been used in the analysis, in tables 5.2 and 5.3 are shown the datasets with LO cross-sections.

The $t\bar{t}$ pair production processes are generated with POWHEG matched to Pythia, evaluating their cross sections at the next-to-next-to leading order (NNLO) in perturbative QCD.

$W + Jets$, and $Z + Jets$ samples are generated with the Madgraph tree-level matrix-element generator matched to Pythia, their cross section are calculated at leading order (LO). The simulation of both samples is divided into ranges of Transverse Hadronic (HT) energy,i.e. the linear sum of hadronic energy of the event $HT = \sum_{jets} \left|p_T^{jet}\right|$, to increase the statistics in the regions with lower cross section, and therefore in high $\sqrt{s}$-region.

Simulated MC samples for background processes used in this thesis are listed in table 5.4 with the corresponding cross-sections.

T mass (GeV)	Cross-section (pb)
700	7.8×10^{-2}
1000	1.4×10^{-2}
1300	3.25×10^{-3}
1600	9.4×10^{-4}
1800	4.4×10^{-4}

Table 5.2: Simulated T samples employed in the analysis with their corresponding production cross sections.

φ mass (GeV)	Cross-section (pb)
50	0.70
500	43.85×10^{-4}
1000	24.99×10^{-5}

Table 5.3: Simulated DM samples employed in the analysis with their corresponding production cross sections.

5.2 Event selection

In the following subsections the selection strategy to extract signal event is presented.

5.2.1 Preselection

The first step of the analysis consists in a coarse selection of signal-like event, composed of a series of requirements to reject topologies significantly different from the searched signature. The first requirement on data includes a set of HLT conditions based on MET and HT. The request consists of the logical *or* between four conditions:

- *PFMETNoMu120_PFMHTNoMu120_IDTight*;
- *PFMET120_PFMHT120_IDTight*;
- *PFMETNoMu110_PFMHTNoMu110_IDTight*;
- *PFMETNoMu110_PFMHTNoMu110_IDTight*.

After trigger requirements, a set of filters are applied to reject events with anomalous MET signals coming from bad reconstruction. After this step, a selection is applied to identify the category it belongs to: event with one top quark *merged*, event with one top quark *partially-merged*, event with one top quark *resolved*.

Sample	Cross-section (pb)
$t\bar{t}$ (semilep)	364.3
$t\bar{t}$ (HT 700-1000)	80.5
$t\bar{t}$ (HT 1000-Inf)	21.3
$Z + Jets$ (HT 100-200GeV)	280.35×1.37
$Z + Jets$ (HT 200-400GeV)	77.67×1.52
$Z + Jets$ (HT 400-600GeV)	10.73×1.37
$Z + Jets$ (HT 600-800GeV)	2.56×1.04
$Z + Jets$ (HT 800-1200GeV)	1.18×1.14
$Z + Jets$ (HT 1200-2500GeV)	0.29×0.88
$Z + Jets$ (HT 2500-Inf GeV)	0.007×0.88
$W + Jets$ (HT 70-100GeV)	1353.0×1.21
$W + Jets$ (HT 100-200GeV)	1345×1.21
$W + Jets$ (HT 200-400GeV)	359.7×1.21
$W + Jets$ (HT 400-600GeV)	48.91×1.21
$W + Jets$ (HT 600-800GeV)	12.05×1.21
$W + Jets$ (HT 800-1200GeV)	5.501×1.21
$W + Jets$ (HT 1200-2500GeV)	1.329×1.21
$W + Jets$ (HT 2500-Inf GeV)	0.03216×1.21

Table 5.4: SM background samples and their cross-sections. The $W + Jets$ samples cross sections is multiplied by the scale factor obtained from the ratio NLO over LO cross section.

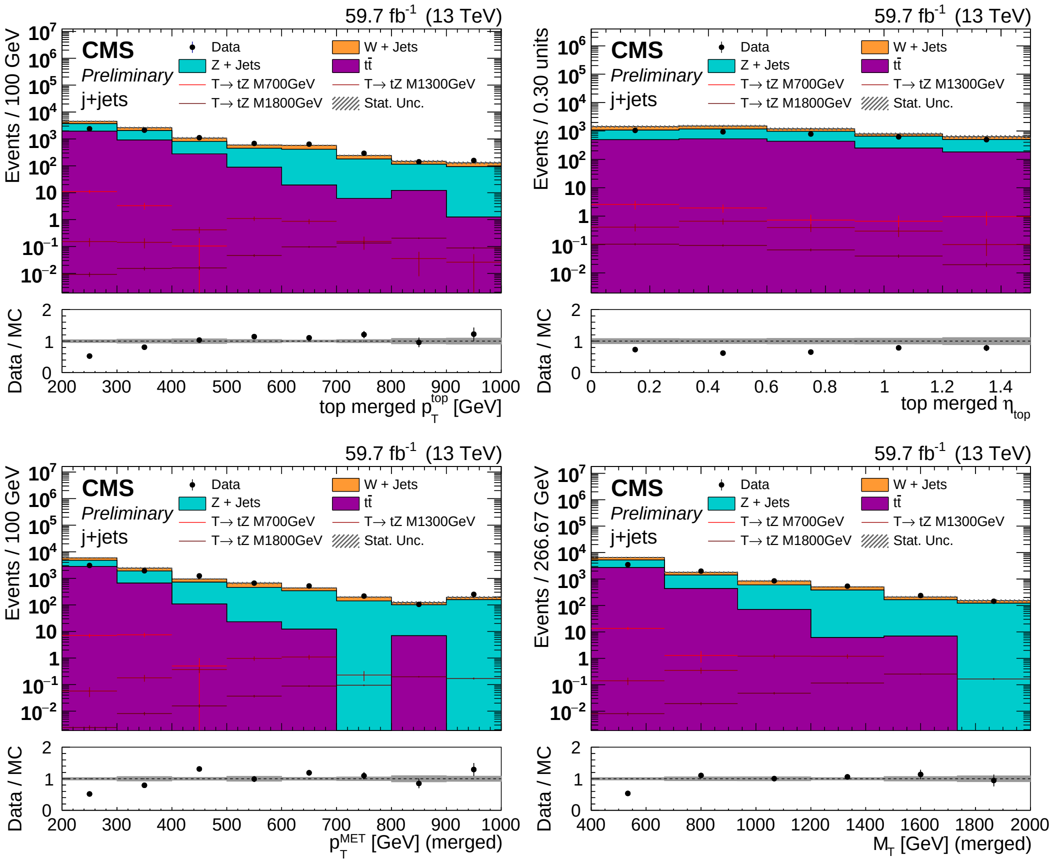


Figure 5.4: M_T, p_T^{top}, η_{top}, and p_T^{MET} distributions in CR for top quark merged category with VLQ T signal.

The aim of the search is to select event with high MET values due to generated invisible particles. Every event collects MET due to different reasons, one of these can be the not perfect jet energy reconstruction: a request on the angle in the transverse plane between MET and jets are made.
Moreover, no leptons are present in the final state, therefore a veto based on reconstructed electrons and muons are added.
The preselection cuts used are:

- $p_T^{MET} > 200$ GeV;
- $min\big(\Delta\phi_{(jet,MET)}\big) < 0.6$;
- lepton veto, no muon or electron candidates passing tight ID (Tables 4.1, 4.2).

During the data taking in 2018 there was a malfunction in a detector region, to take this into consideration some events have been rejected and the MET has been reweighted, after the preselection of the events.

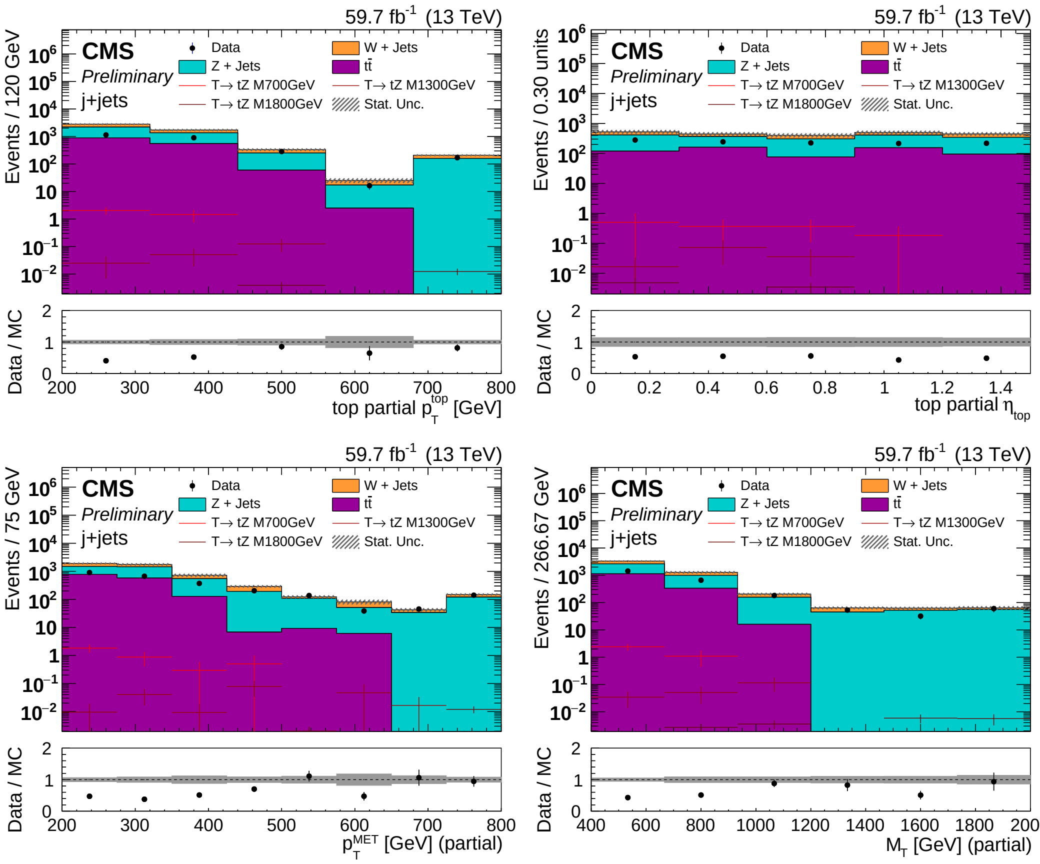


Figure 5.5: M_T, p_T^{top}, η_{top}, and p_T^{MET} distributions in CR for top quark partially-merged category with VLQ T signal.

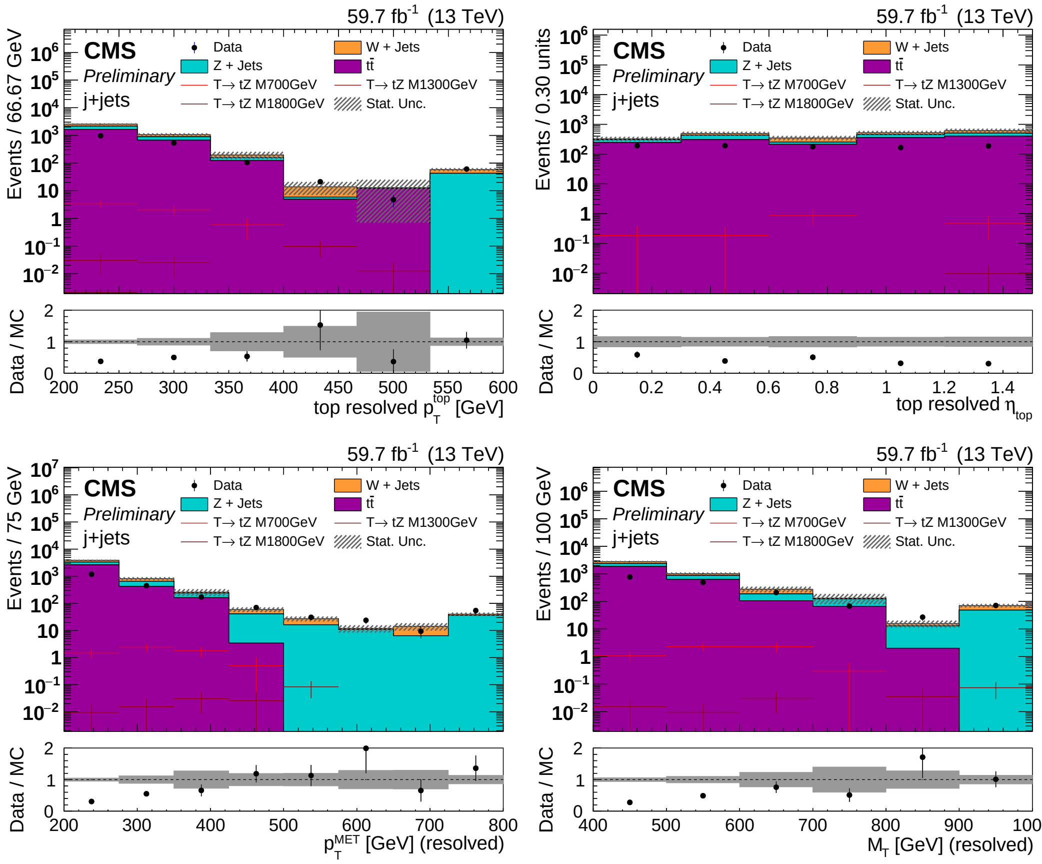


Figure 5.6: M_T, p_T^{top}, η_{top}, and p_T^{MET} distributions in CR for top quark resolved category with VLQ T signal.

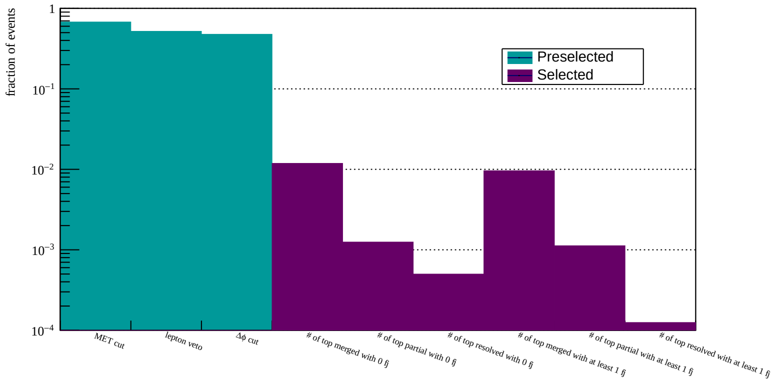


Figure 5.7: Cut flow histogram for T samples with T mass of 1300 GeV. In the first 3 bins are reported the fraction of events passing the correspondent pre-cut, in the last 6 bins are reported the number of selected event per each category of signal (merged, partially-merged, resolved) in SR and CR.

5.2.2 Top quark selection

After the preselection cuts described above, the event is submit to a selection to identify its category: top quark *merged*, top quark *partially-merged*, top quark *resolved*.

Firstly all top quark candidates are reconstructed, and the BDT selection is applied. For each of BDT model a threshold value is found to discriminate signal (*true* top quark) from background (*false* top quark) that reduce the background rate to 10% (*Loose*) and 1% (*Tight*) of the pre-cut yield.

The BDT scores of all top merged candidates are evaluated; In order to classify the event one top quark is selected:

1. the BDT score for top merged candidates are evaluated and compared to the correspondent threshold value;
2. if no top merged is selected, all top partially-merged are subsequently looked through;
3. finally, if not even one top partially-merged is identified, a top resolved is searched for.

The procedure is applied with both loose cut and tight cut at BDT selection level.

5.3 Signal and Control regions

The three regions taken into account (one top quark merged, one top quark partially-merged, one top quark resolved) have been divided into further two regions:

- *Signal Region* (SR), enriched in signal events;

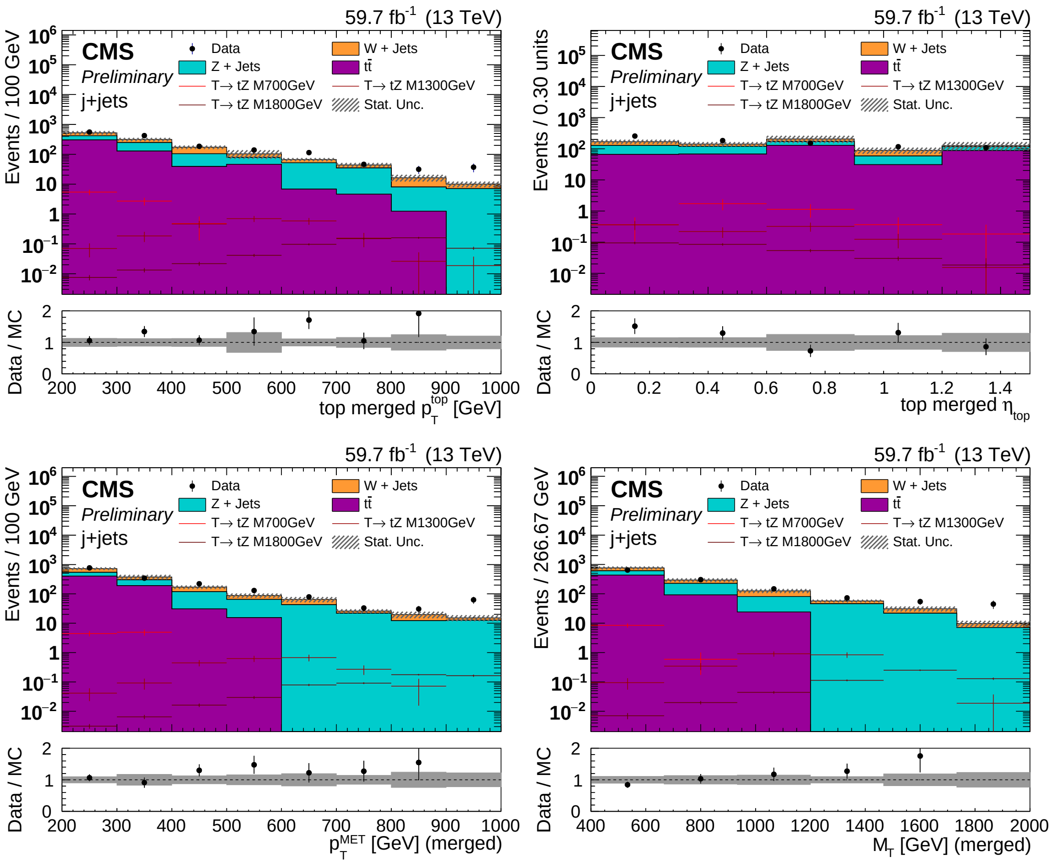


Figure 5.8: M_T, p_T^{top}, η_{top}, and p_T^{MET} distributions in SR for top quark merged category with VLQ T signal.

- *Control Region* (CR), enriched in background events.

These two regions are characterised by the number of jets in the forward region of the detector, *forward jets*. A forward jet is a narrow jet with the following features:

- $p_T^{jet} > 40$ GeV;
- $2.4 < |\eta_{jet}| < 4.0$;
- $jet\_Id \leq 2$,

where jet_Id is a set of requirements by the CMS Jet reconstruction group to reject fake jets. The SR is made of events with at least 1 forward jet, while the CR collects events with 0 forward jets. As can be seen in the diagrams in fig. 4.1 and 3.7, in the signal events a spectator quark is present and it can be detected in the forward region.

We do not apply any requirement but study separately events with 0 and at least 1 forward jets. The selection results are reported for different reconstructed variables:

- p_T of the reconstructed top;

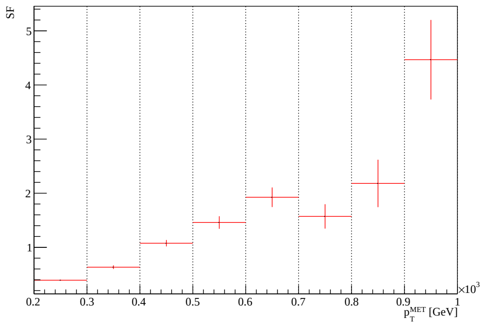


Figure 5.9: Scale Factors for each bin of MET.

- η of the reconstructed top;
- p_T^{MET};
- reconstructed transverse mass,

  $$m_{transverse} = \sqrt{2p_T^{top} \cdot p_T^{MET}(1 - c\cos\Delta\phi_{top,MET})}, \tag{5.1}$$

  that in VLQ T case corresponds to the reconstructed T mass, while in DM case to the reconstructed φ mass.

In fig. 5.7 is rapresented the cut flow histogram for simulated T sample in the hypothesis mass of 1300GeV, in which the number of survived and selected events is reported. The survived events are divided according to the pre-cut, obviously the number depends also by the order in which the cuts are applied. The selected events are divided in CR and in SR, as well as top quark reconstruction categories.

Background extraction

Systematic effects can affect both the yield and shape of the background and the signal predictions. To take into account possible background uncertainties, the correction has been derived from data in the CR.

Firstly, a requirementto reduce the contribution from jets in the region with $\eta > 1.6$ has been applied, to remove fake jets that came as results of a known malfunctioning of a section of the hadronic calorimeter in 2018. Then reconstructed top quarks in the same region are also rejected.

Moreover, a reweighting based on p_T^{MET} distribution is done. Considering the two thresholds in BDT score used for top quark reconstruction 5.2.2, we extract events with 1 top quark loose and 0 top quark tight, where loose (tight) is referred to top quark selected with the corresponding treshold, in CR. In this region, almost exclusively background events is expected, therefore, the p_T^{MET} of selected data have to be compatible with simulated background. Finally, a scale factor is extracted applying the compatibility request:

$$SF_i = \frac{n_{data}}{n_{bkg,i}}, \tag{5.2}$$

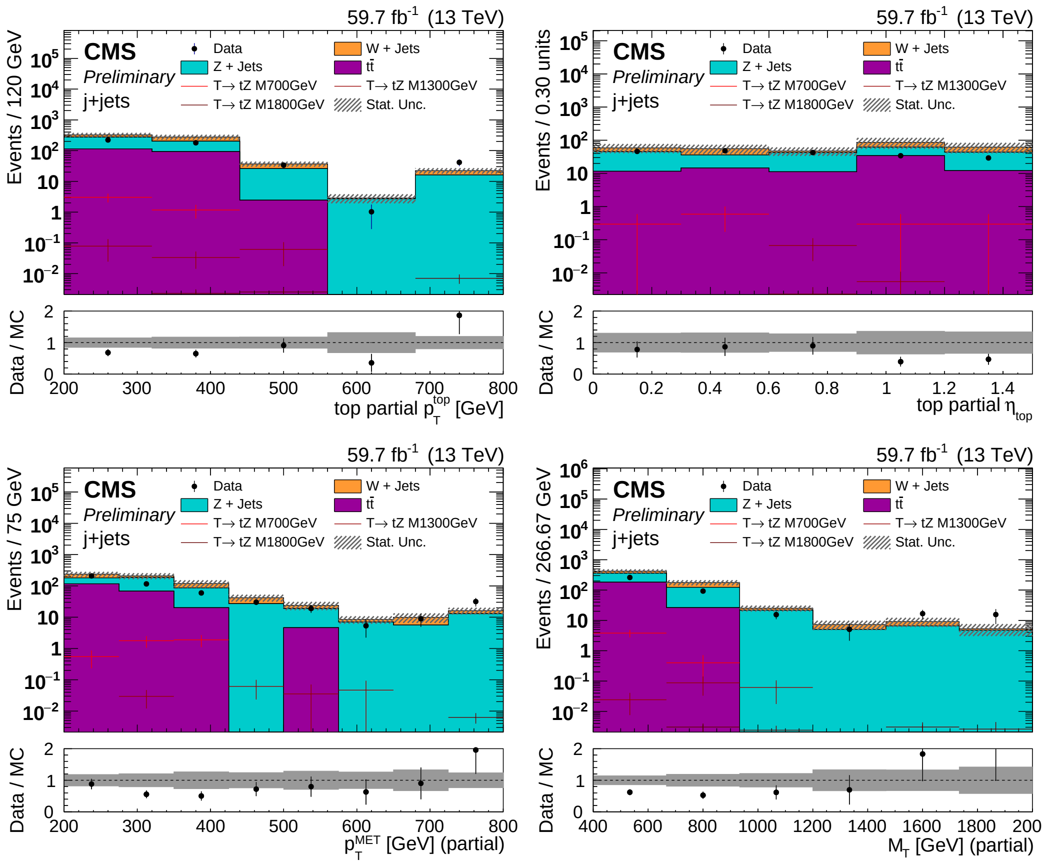


Figure 5.10: M_T, p_T^{top}, η_{top}, and p_T^{MET} distributions in SR for top quark partially-merged category with VLQ T signal.

where i is the i^{th} bin of the MET distribution. This weight is then applied to all background events. In fig. 5.9 is shown the SFs for each bin.
The comparison beetwen the distribution of M_T before these data driven correction and after can be observed in fig. 5.11 . The final distribution obtained for event reconstructed with tight cut is reported for the chosen variables for each category of reconstructed top quark and both considered signals in the CR (5.4-5.6, 5.13-5.15) and SR (5.8-5.12, 5.16, and 5.17). For the readability of the histograms only three signal samples are rapresented instead of the five used in the analysis.

5.4 Fit procedure

For the fit two hypothesis have been considered:

- null hypothesis, $\mathcal{H}_0$, assumes absence of new physics, therefore the signal is absent or too little to be detected;
- alternative hypothesis, $\mathcal{H}_1$, assumes the presence of new signal.

An extended binned Maximum Likelihood fit to data [87] has been performed on reconstructed transverse mass M_T simultaneusly in the 3 Signal

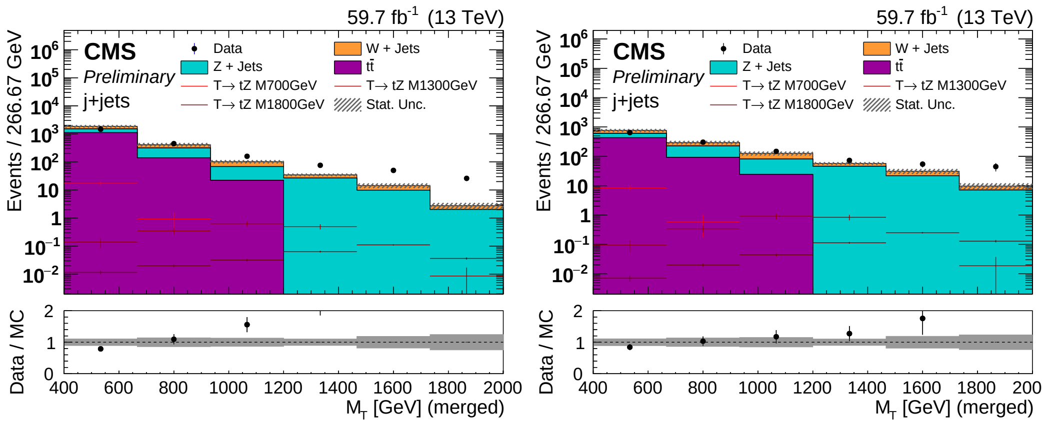


Figure 5.11: M_T distribution in SR for merged category with VLQ T signal, on the left is showed the histogram before background extraction and on the right the same histogram after.

and Control Regions chosen, considering 5 points of T mass from 700 GeV to 1800 GeV and for 3 points of φ mass from 50 GeV to 1000 GeV.
The likelihood $\mathcal{L}(obs|\mu,\theta)$ is the product of a Poissonian distribution multiplied a second term that represents the signal and background models:

$$\mathcal{L}(m|\mu_{signal},\theta) = \mathcal{P}(n_{obs},\mu_{signal}\cdot s(\theta)+b(\theta))\prod_{k=1}^{N_{events}} f(m_k|\mu_{signal},\theta), \qquad (5.3)$$

where $s(\theta)$ and $b(\theta)$ are the expected yields of the signal and background, that depend on the set θ of distribution parameters. The quantity μ_{signal} is the signal strength, it is expected to be 1 if $\mathcal{H}_0$ is false, while it is equal to 0 if $\mathcal{H}_1$ is false. The function $f(m_k|_{signal},\theta)$ is a combination of Probability Densitiy Functions (PDFs) for the signal and for the background:

$$f(m_k|\mu,\theta) = \frac{\mu\cdot s(\theta)}{\mu\cdot s(\theta)+b(\theta)} f_s(m_k|\theta) + \frac{b(\theta)}{\mu\cdot s(\theta)+b(\theta)} f_b(m_k|\theta). \qquad (5.4)$$

In this analysis rate parameters have been introduced for the backgrounds to take into account systematic effects that might alter the total yield of the backgrounds. Variations on the shape are taken into account by the shape extraction described in section (5.3). With the rate parameters the function $f(m_k|\mu,\theta)$ becomes:

$$\begin{aligned} f(m_k|\mu,\mu_{t\bar{t}},\mu_{Z+Jets},\mu_{W+Jets},\theta) &= \frac{\mu\cdot s(\theta)}{\mu\cdot s(\theta)+b(\theta)} f_s(m_k|\theta) \\ &+ \sum_i \left[\frac{\mu^i_{t\bar{t}}\cdot b^i_{t\bar{t}}}{\mu\cdot s(\theta)+b(\theta)} f^i_{b_{t\bar{t}}}(m_k|\theta) + \frac{\mu^i_{Z+Jets}\cdot b^i_{Z+Jets}}{\mu\cdot s(\theta)+b(\theta)} f^i_{b_{Z+Jets}}(m_k|\theta) \right. \\ &\left. + \frac{\mu^i_{W+Jets}\cdot b^i_{W+Jets}}{\mu\cdot s(\theta)+b(\theta)} f^i_{b_{W+Jets}}(m_k|\theta) \right], \end{aligned} \qquad (5.5)$$

where i corresponds to 6 regions defined before, in order to have that every parameters are independent.

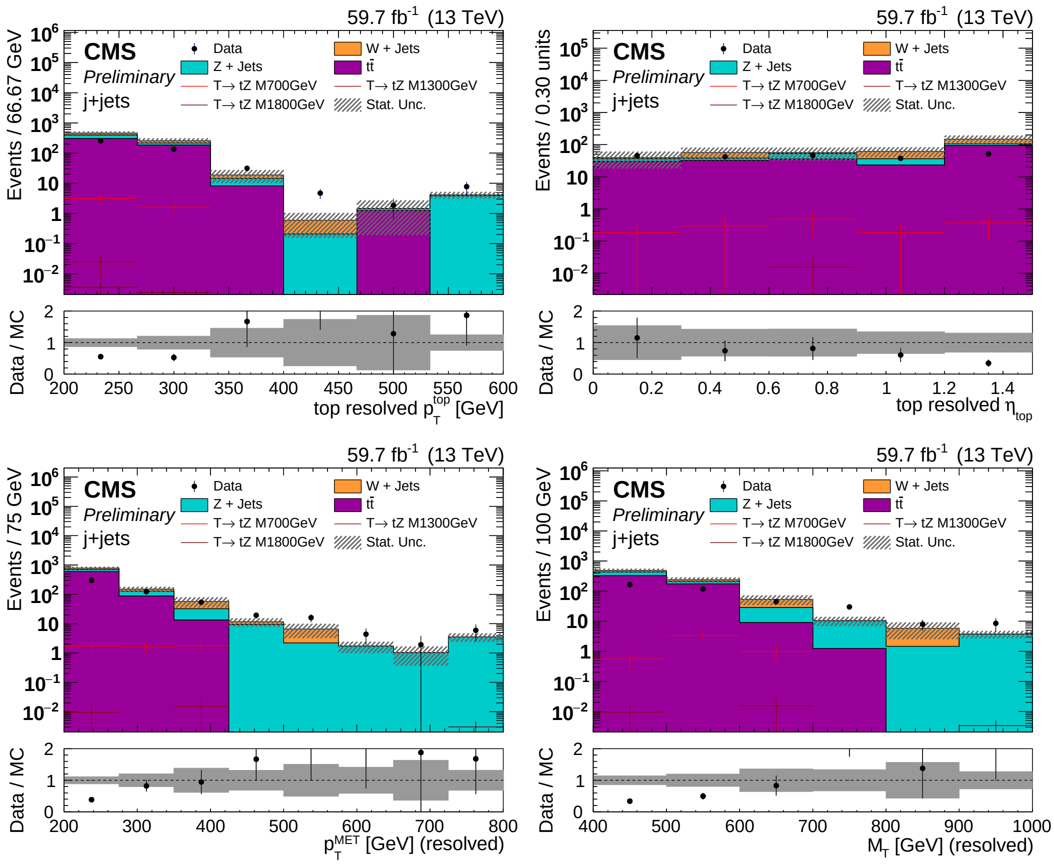


Figure 5.12: M_T, p_T^{top}, η_{top}, and p_T^{MET} distributions in SR for top quark resolved category with VLQ T signal.

Since no excess related to the production of a new particle is observed, and the data are compatible with the $\mathcal{H}_0$ hypothesis of only background, the upper limits are estimated on the production cross sections and the branching ratio of the process $T \to tZ$ and $t + DM$. Expected and observed upper limits on the signal cross-section as a function of the mass of the scalar mediator φ and the VLQ T mass are derived at 95% ctonfidence level (CL).
The limits are evaluated with the Combine tool [93] dedicated software, which is able to provide a measure of the level of incompatibility of data with a signal hypothesis. The modified frequentist method [87] has been used to measure the expected and observed upper limits. As test statistic is used the profile Likelihood ratio:

$$q_\mu = -2 \ln \lambda(\mu_{sig}) \tag{5.6}$$

where,

$$\lambda(\mu) = \frac{\mathcal{L}(m|\mu, \hat{\hat{\theta}}(\mu))}{\mathcal{L}(m|\hat{\mu}, \hat{\theta})}, \tag{5.7}$$

$\hat{\mu}$ and $\hat{\theta}$ are the best fit values observed from the data sample, while $\hat{\hat{\theta}}$ are the best fit values of the nuisance parameters obtained for a fixed value μ. The

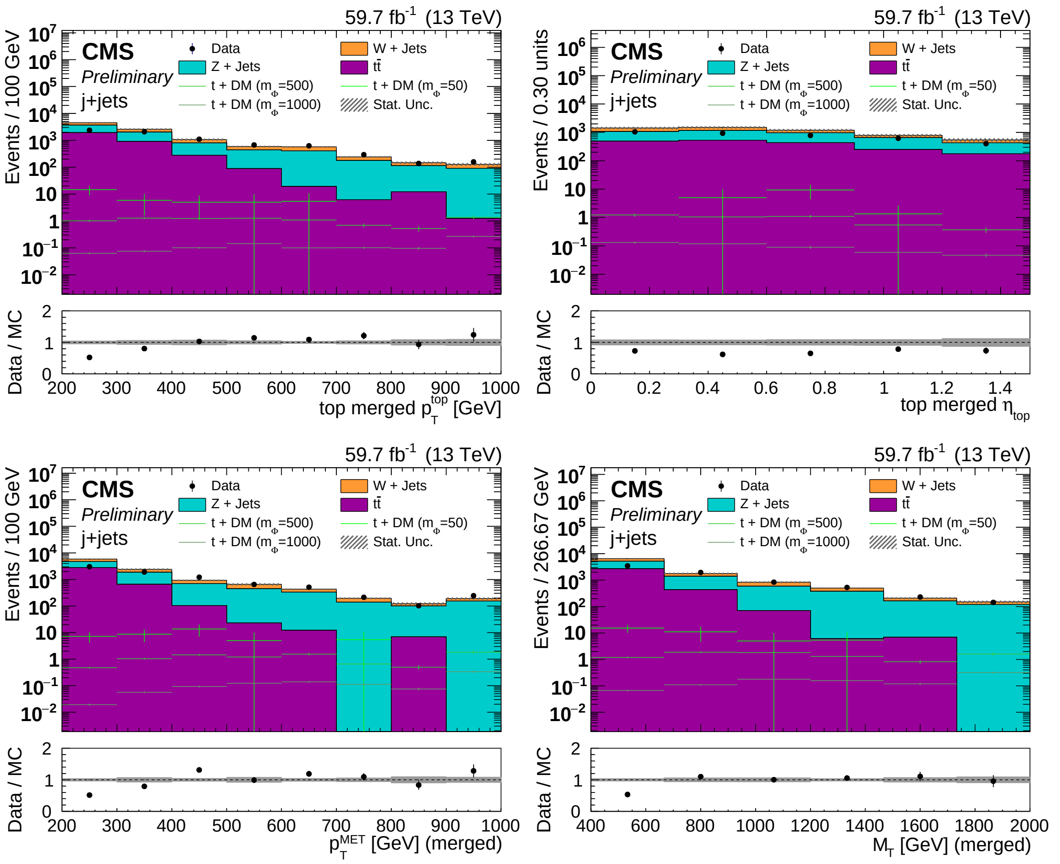


Figure 5.13: M_T, p_T^{top}, η_{top}, and p_T^{MET} distributions in CR for top quark merged category with DM signal.

profile Likelihood test statistic is the most powerful discriminator, minimizing the error of type-II, i.e. $\mathcal{H}_1$ is accepted while $\mathcal{H}_0$ is true, at a significance level α. If q_μ is higher than 0 the events appear to be under the $\mathcal{H}_0$ hypothesis, while if it is lower than 0, they are more compatible with the $\mathcal{H}_1$ hypothesis.

The limits on signal events are estimated using modified frequentist method CLs. The confidence level for the signal can be defined as the ratio between the confidence level observed for the signal and background hypothesis, CL_{s+b}, and the confidence level observed for the background-only hypothesis, CL_b:

$$CL_s = \frac{CL_{s+b}}{CL_b}, \tag{5.8}$$

CL_b is the probability to have for a given value of $_{sig}$ a value of the test statistics equal or larger than the value observed in the experiment, under the hypothesis $\mathcal{H}_0$:

$$CL_b = P_b(q_\mu \geq q_\mu^{obs}),$$

and 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

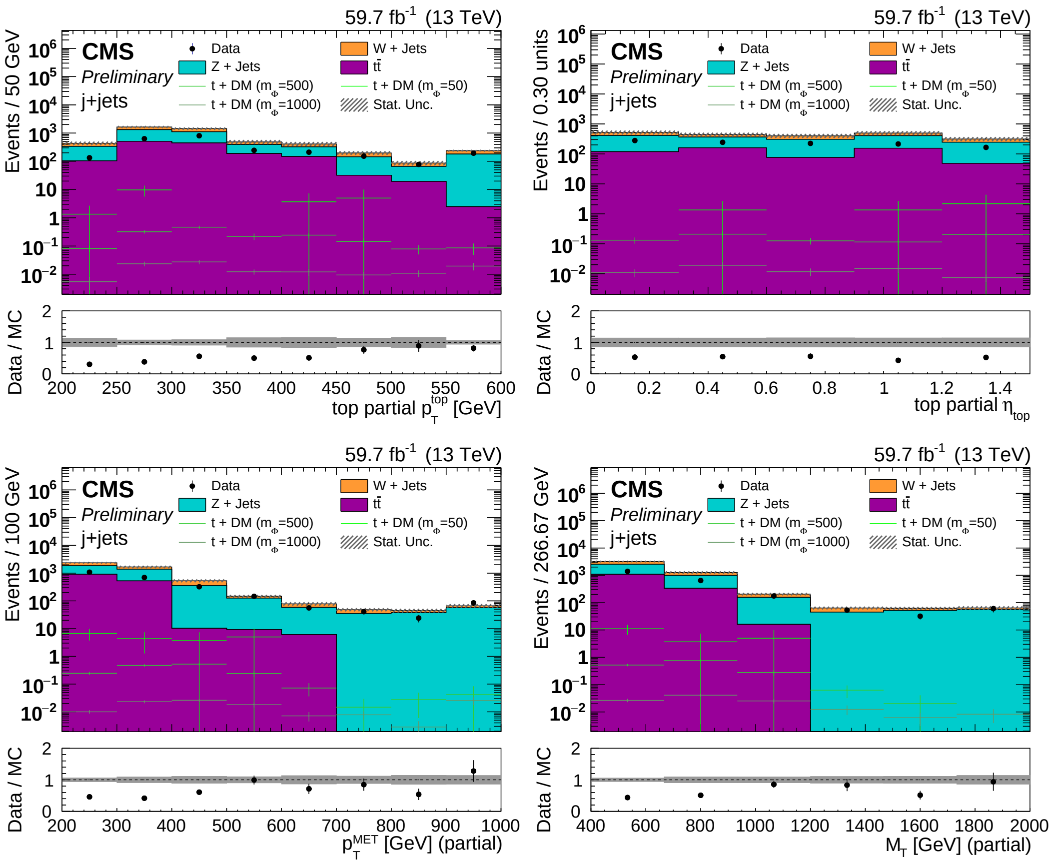


Figure 5.14: M_T, p_T^{top}, η_{top}, and p_T^{MET} distributions in CR for top quark partially-merged category with DM signal.

experiment, under the hypothesis $\mathcal{H}_1$:

$$CL_{s+b} = P_{s+b}(q_\mu \geq q_\mu^{obs}).$$

The limits have been computed at 95% of CL and shown in fig. 5.18 and 5.19.

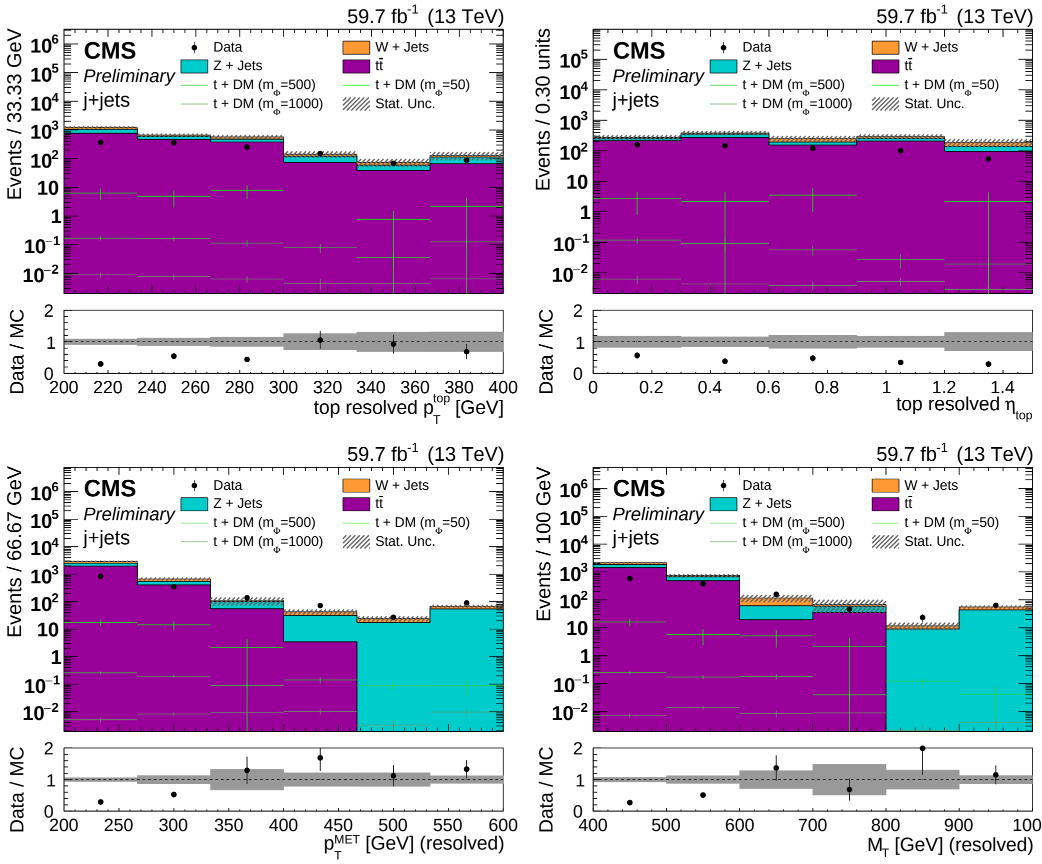


Figure 5.15: M_T, p_T^{top}, η_{top}, and p_T^{MET} distributions in CR for top quark resolved category with DM signal.

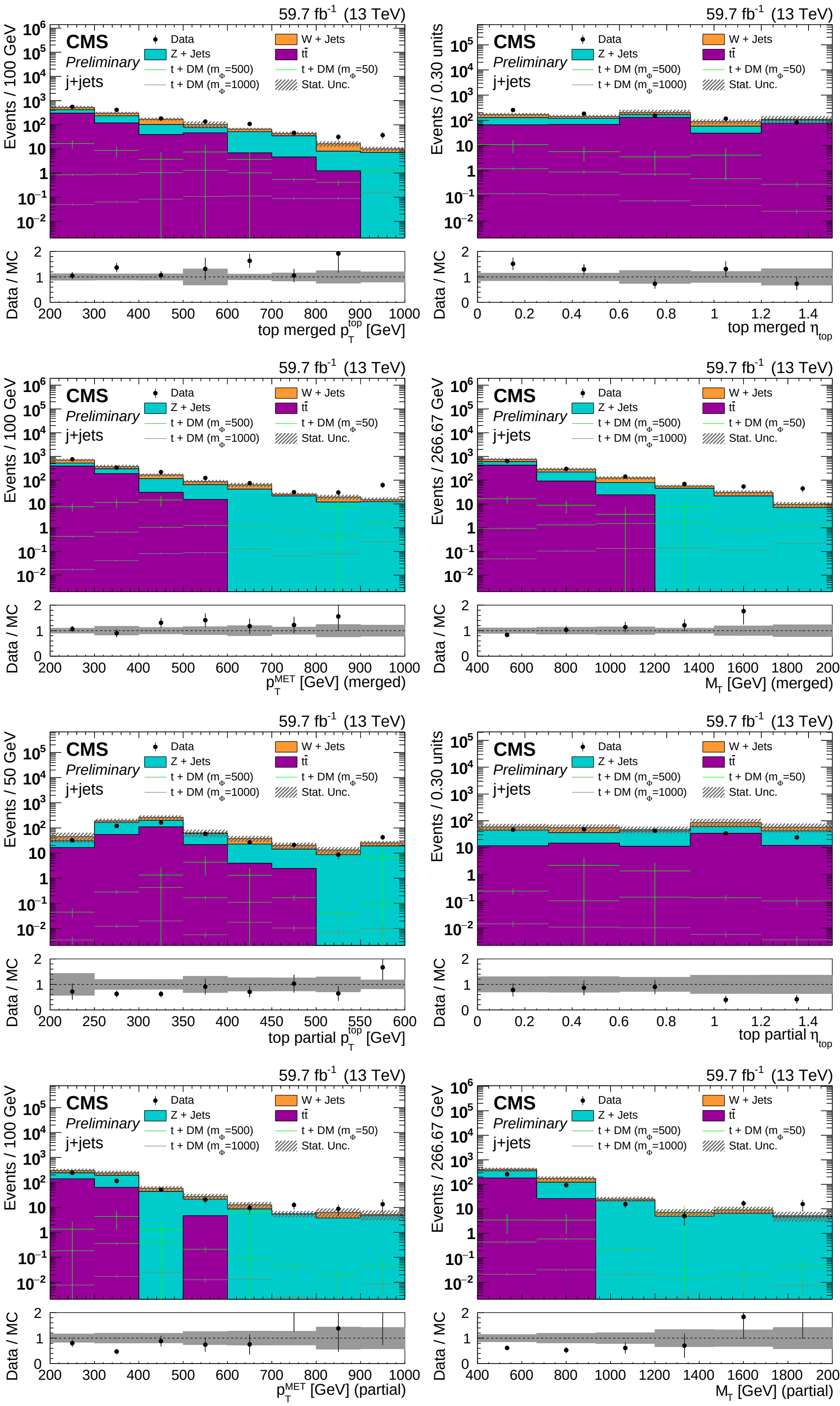


Figure 5.16: In the first four histograms: M_T, p_T^{top}, η_{top}, and p_T^{MET} distributions in SR for top quark merged category with DM signal. In the last four histograms: M_T, p_T^{top}, η_{top}, and p_T^{MET} distributions in SR for top quark partially-merged category with DM signal.

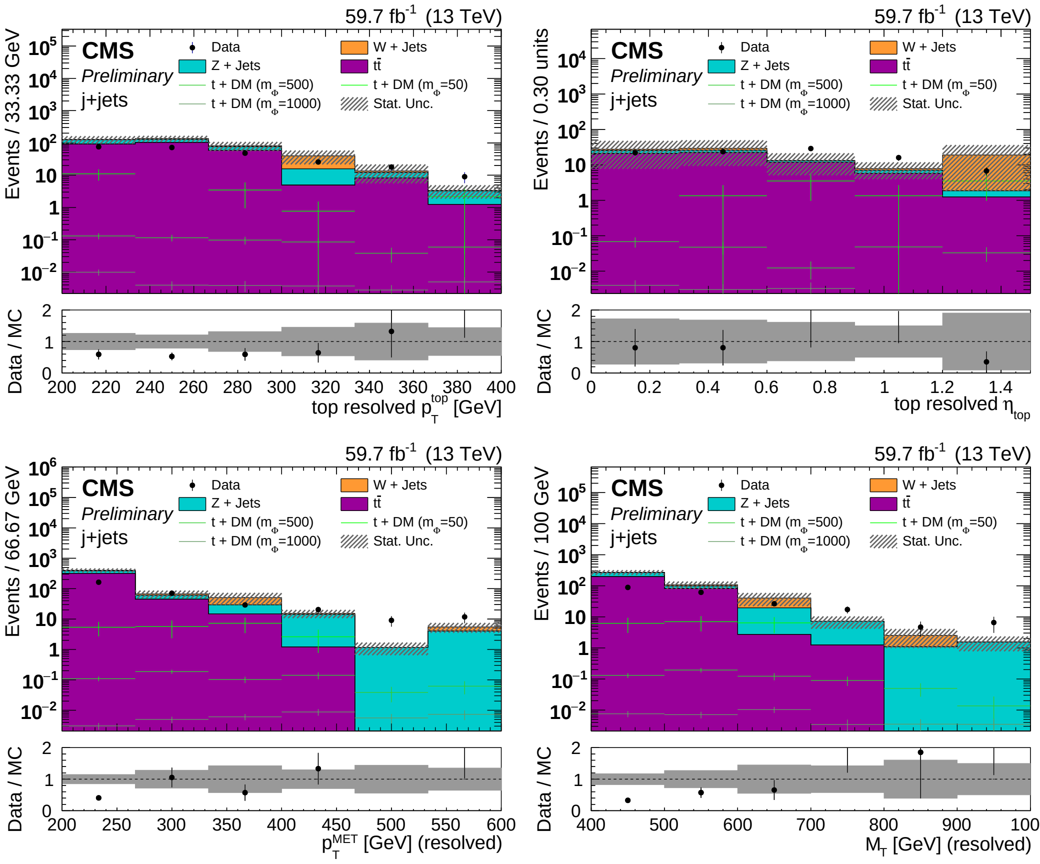


Figure 5.17: M_T, p_T^{top}, η_{top}, and p_T^{MET} distributions in SR for top quark resolved category with DM signal.

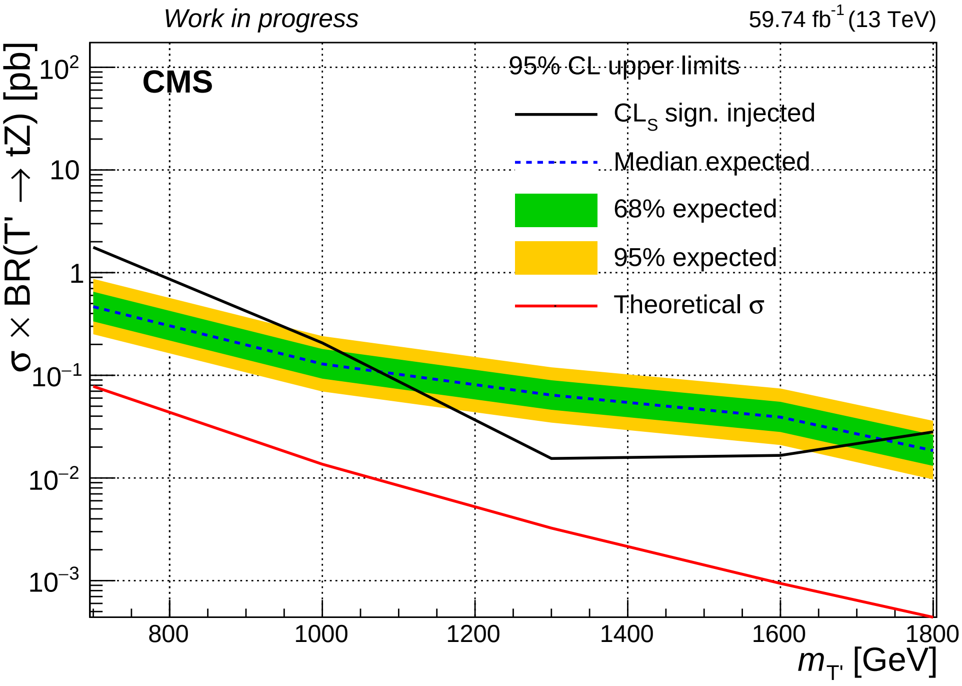


Figure 5.18: The observed and expected 95% CL upper limits of the VLQ T production cross section, with T decaying to tZ, as a function of the signal mass, resulting from the fit.

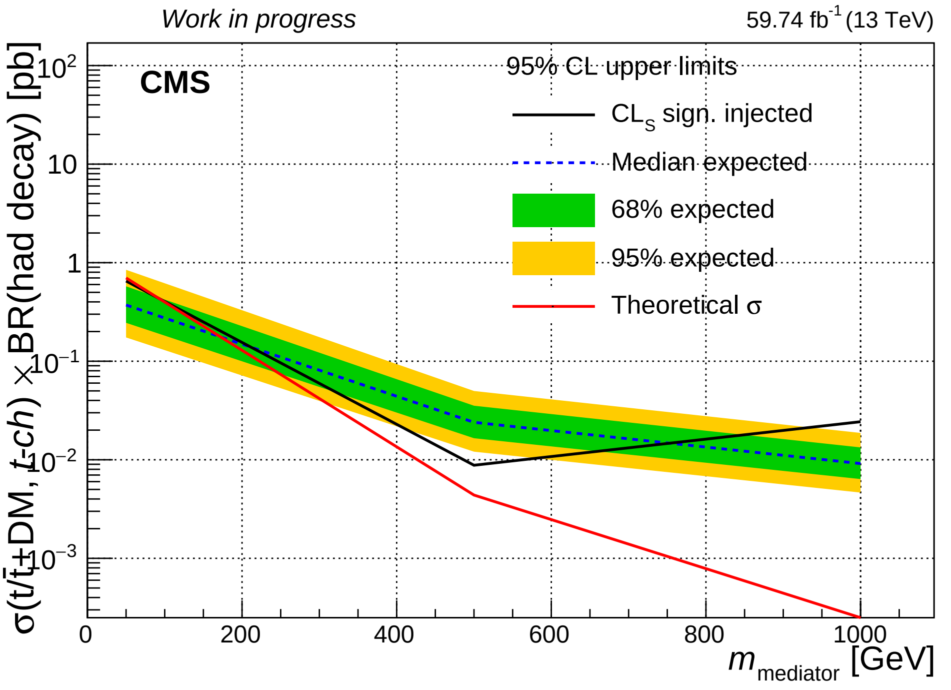


Figure 5.19: The observed and expected 95% CL upper limits of the DM production cross section, in φ scalar mediator hypothesis, as a function of the mediator mass, resulting from the fit.

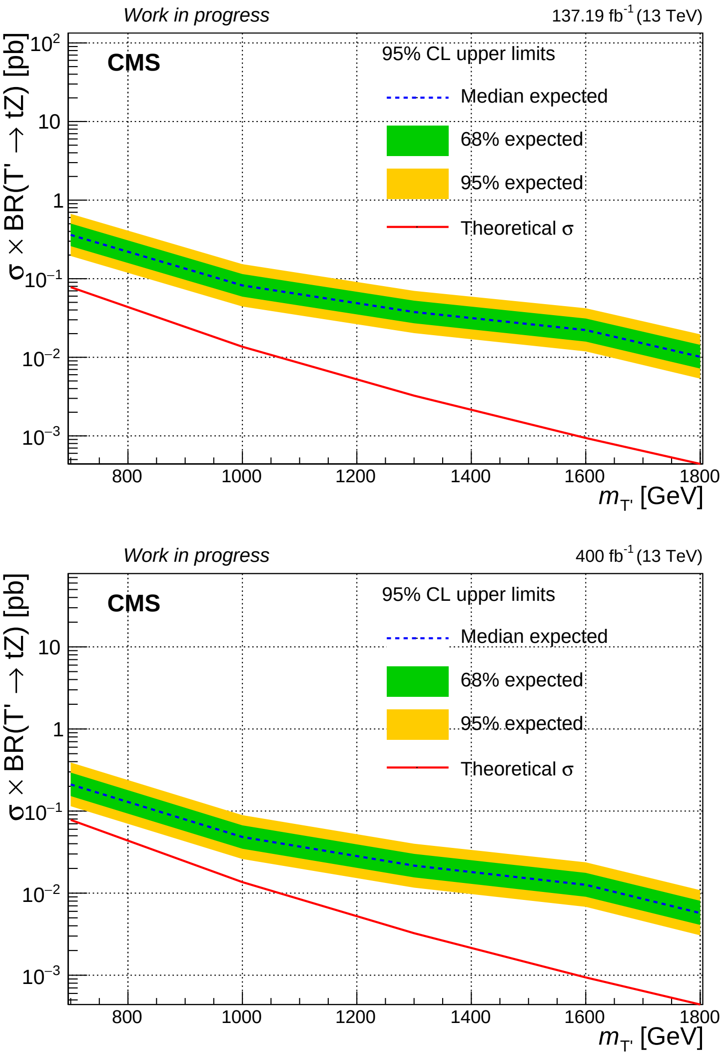


Figure 5.20: The expected 95% CL upper limits projected at 137.19 fb^{-1} and 400 fb^{-1} of the VLQ T production cross section, with T decaying to tZ, as a function of the signal mass.

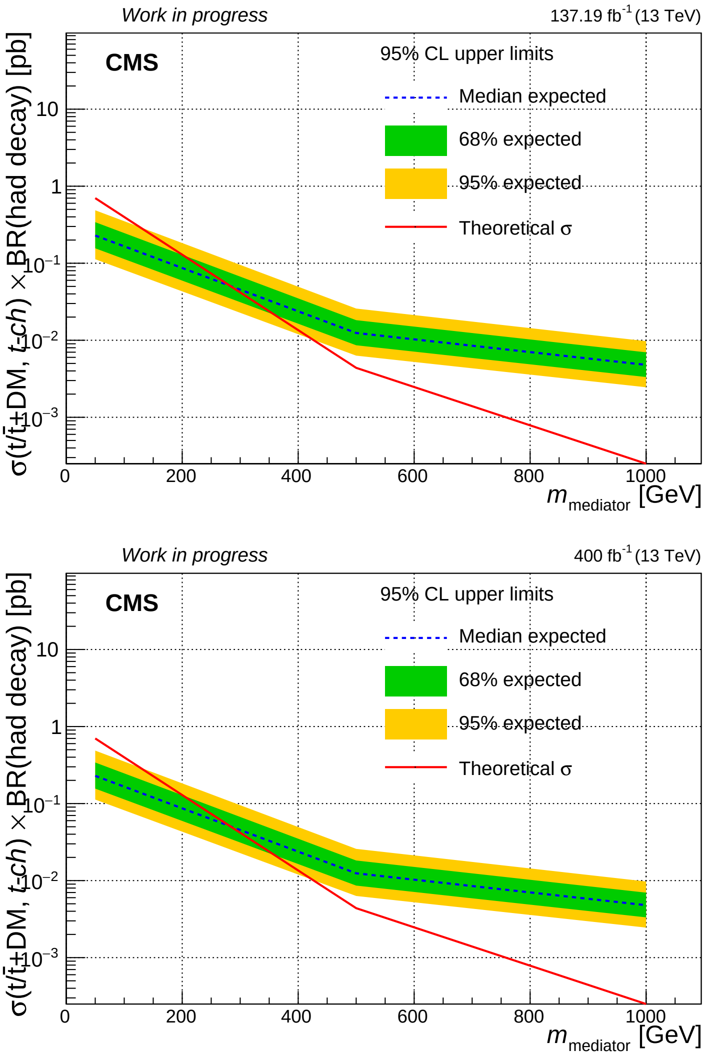


Figure 5.21: The expected 95% CL upper limits projected at 137.19 fb^{-1} and 400 fb^{-1} of the DM production cross section as a function of the mediator mass.

Conclusions

In this thesis a search for a final state composed by a singly produced top quark in association with invisible particles is presented. The search is performed using a data set collected by the CMS experiment at the LHC in 2018, at a centre-of-mass energy of 13 TeV, and corresponding to an integrated luminosity of 59.7 fb^{-1}.

The final state considered is motivated by several theoretical frameworks aiming at extending the Standard Model of particle physics, and the work presented focuses on two of those particular: the single production of the VLQ T and its decays into a top quark and Z boson, and the DM production through its coupling to the top quark.

In both cases a striking feature of the final state is the presence missing energy, due to the couple production of DM, $\chi\bar{\chi}$, or to the couple of neutrinos coming from Z decay, $\nu\bar{\nu}$. The second striking feature is the presence of a top quark decaying to $t \to bW \to bq\bar{q}'$. The goal of this work was to define new idetification criteria for the top quark reconstruction for the hadronic decaying of top quark making use of Machine Learning algorithm, to enhance the performances of the search and increase its discovery potential. Three different ways of reconstruction have been taken into account depending on the number of jets in the final state. The *top quark merged* is characterized by a single large radius, or AK8, jet; the *top quark partially-merged* by one Ak8 W-jet and one narrow radius, or AK4, b-jet; finally the *top quark resolved* is characterized by three distinct AK4 jet with at least one coming form a b-quark hadronization.

Two different methods can be used to apply top quark selection in one of the three categories: a standard "cut-based" selection making use of traditional jet reconstruction and identification criteria, or a ML selection that can take into account the deep tagging and kinematic variables of jets.

In this thesis the ML approch was used, a binary classification was implemented to select the true top quark candidate in each of the three categories. Events were selected where at least one top quark defined according to the ML identification criteria was present. In case more than one top quark was reconstructed, a hierarchy was established according to the following order: top quark merged, then top quark partially-merged, and finally top quark resolved.

The transverse mass, M_T was reconstructed in the events labelled as signal by the selection. The signal regions have been chosen considering the number of forward jets, which is expected be greater than one for the signals and equal to zero for the background.

A binned maximum likelihood fit to data is performed simultaneously on the M_T distribution reconstructed in all regions and for different mass hypotheses for both signal models considered. Upper limits on the cross sections have been estimated at 95% C.L., since no significant excess related to the production of new particles has been observed.

The first future development of this analysis is the introduction of systematic uncertainties to take into account the dependence from simulated data sets and experimental effects.

The analysis is also expected to be extended to the data collected by the LHC in RunII and in the upcoming years. In fig. 5.20 and 5.21 are showed the projected 95% C.L. upper limits when the whole RunII data set is considered, with the integrated luminosity of 137.19 fb^{-1}, and when also the future RunIII, expected for the three-year period 2022-2024, will be taken into account.

Further developments include using a different ML approach, in paricular using a classifier algorithm that will not only reconstruct the top quarks, but also choose optimal reconstruction strategy.

The analysis may be extended in both cases considering the leptonic channel of top quark decay, $t \to bW \to b\ell\nu$, in order to consider all the possible top quark reconstruction ways. Also the frameworks taken into account allow different interpretations of the results, such as a pseudoscalar o vector mediator boson can be considered in DM production case.

Acknowledgements

Sono giunto al termine di un importante capitolo della mia vita, questi ultimi 5 anni hanno fortemente condizionato la persona che sono e che sarò, hanno richiesto grandi sacrifici ma mi hanno portanto anche grandi gioie e soddisfazioni. A termine di questo percorso voglio ringraziare innanzitutto la mia famiglia per avermi sopportato nonostante i miei continui lamenti, per avermi dato la possibilità di dare sempre il massimo, per essere sempre stati d’aiuto nei momenti più difficili.
In particolar modo voglio ringraziare il mio relatore il prof. Iorio, o meglio Orso, per l’incredibile impegno che ha messo in questo lavoro, per essere stato sempre disponibile, per la sua contagiosa allegria che è stata fondamentale per superare i momenti di crisi. Voglio ringraziare la professoressa Di Crescenzo per i suoi utili consigli e l’ interesse mostrato verso questo lavoro.
Ringrazio Agostino, Francesco, Oriana, Cristina per aver reso più felici gli ultimi mesi di questo percorso, per il supporto dato a questo lavoro e per i pranzi sul tetto, i caffé, ma soprattuo le tazzine da lavare. Spero di poter continuare con voi e con tutto il gruppo CMS di Napoli questo viaggio iniziato con la tesi e che durerà ancora, almeno, un paio di anni.
Ringrazio gli amici di sempre, Giuseppe, Francesco C., Maria, Giovanni C., Chiara, Francesco S., Teresa, Giovanni G., Giovanna, Vincenzo, Giovinale, Simona, perché non riesco a trovare ricordi in cui non siete presenti ed ancora oggi siete sempre qui nonostante gli impegni e la lontananza.
Ringrazio Francesco L., Formicola, Luca, Luigi, Peppe, Vincenzo L., Emilia, perché, anche se mi sono inserito dopo, mi sono sempre sentito parte di voi, perché siete sempre lì ad ascoltare i miei sfoghi, per darmi supporto ed aiuto.
Ringrazio i ragazzi conosciuti all’università, Veronica, Francesca, Antonio, Mara, Valeria, perché abbiamo condiviso stesse ansie e paure, stessi dolori, e li abbiamo superati dandoci sostegno l’un l’altro. Anche se negli ultimi anni abbiamo scelto percorsi diversi è sempre bello rivedervi e passare le serate insieme. Grazie a Riccardo e Vigilante perché anche se distanti è un piacere condividere con voi ogni giornata tra meme e aneddoti assurdi, per le discussioni infinite di cui perdo puntualmente il filo e per gli infiniti viaggi proposti e mai organizzati.
Grazie a tutti voi e a chi mi ha sostenuto, supportato che non ho menzionato esplicitamente. É stato un viaggio bellissimo che rifarei volentieri.

Bibliography

[1] C. N. Yang and R. L. Mills. “Conservation of Isotopic Spin and Isotopic Gauge Invariance”. In: *Phys. Rev.* 96 (1 Oct. 1954), pp. 191–195. DOI: `10.1103/PhysRev.96.191`. URL: `https://link.aps.org/doi/10.1103/PhysRev.96.191`.

[2] George Zweig. *An SU $_3$ model for strong interaction symmetry and its breaking*. Tech. rep. CM-P00042884, 1964.

[3] Moo-Young Han and Yoichiro Nambu. “Three-triplet model with double SU (3) symmetry”. In: *Physical Review* 139.4B (1965), B1006.

[4] Oscar W. Greenberg. “Spin and unitary-spin independence in a paraquark model of baryons and mesons”. In: *Physical Review Letters* 13.20 (1964), p. 598.

[5] H. David Politzer. “Reliable perturbative results for strong interactions?” In: *Physical Review Letters* 30.26 (1973), p. 1346.

[6] David J. Gross and Frank Wilczek. “Ultraviolet behavior of non-abelian gauge theories”. In: *Physical Review Letters* 30.26 (1973), p. 1343.

[7] Sheldon L. Glashow. “Partial-symmetries of weak interactions”. In: *Selected Papers on Gauge Theory of Weak and Electromagnetic Interactions* 171 (1981).

[8] Steven Weinberg. “A model of leptons”. In: *Physical review letters* 19.21 (1967), p. 1264.

[9] Abdus Salam. “Weak and electromagnetic interactions”. In: *Selected Papers Of Abdus Salam: (With Commentary)*. World Scientific, 1994, pp. 244–254.

[10] G Arnison et al. “Experimental observation of lepton pairs of invariant mass around $95GeV/c^2$ at the CERN SPS collider”. In: *Physics Letters B* 126.5 (1983), pp. 398–410.

[11] G Arnison et al. “Experimental observation of isolated large transverse energy electrons with associated missing energy at s= 540 GeV”. In: *Physics letters B* 122.1 (1983), pp. 103–116.

[12] M Banner et al. “Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the CERN pp collider”. In: *Physics Letters B* 122.5-6 (1983), pp. 476–485.

[13] Po Bagnaia et al. “Evidence for $Z^0 \rightarrow e^+e^-$ at the CERN pp Collider”. In: *Physics Letters B* 129.1-2 (1983), pp. 130–140.

[14] F Abe et al. "Evidence for top quark production in p p collisions at $\sqrt{s} = 1.8$ TeV". In: *Physical Review D* 50.5 (1994), p. 2966.

[15] Atlas Collaboration et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *arXiv preprint arXiv:1207.7214* (2012).

[16] Serguei Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Physics Letters B* 716.1 (2012), pp. 30–61.

[17] Peter W Higgs. "Broken symmetries and the masses of gauge bosons". In: *Physical Review Letters* 13.16 (1964), p. 508.

[18] François Englert and Robert Brout. "Broken symmetry and the mass of gauge vector mesons". In: *Physical Review Letters* 13.9 (1964), p. 321.

[19] Enrico Fermi. "Tentativo di una teoria dei raggi β". In: *Il Nuovo Cimento (1924-1942)* 11.1 (1934), pp. 1–19.

[20] Nicola Cabibbo. "Unitary symmetry and leptonic decays". In: *Physical Review Letters* 10.12 (1963), p. 531.

[21] Makoto Kobayashi and Toshihide Maskawa. "CP-violation in the renormalizable theory of weak interaction". In: *Progress of theoretical physics* 49.2 (1973), pp. 652–657.

[22] M PDG. "Tanabashi et al". In: *Review of particle physics, Phys. Rev* (2018).

[23] R. Gupta. "Introduction to Lattice QCD". In: (1998). DOI: `https://arxiv.org/abs/hep-lat/9807028`.

[24] Lyndon Evans and Philip Bryant. "LHC Machine". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08001–S08001. DOI: `10.1088/1748-0221/3/08/s08001`. URL: `https://doi.org/10.1088/1748-0221/3/08/s08001`.

[25] Jorg Wenninger. "SISSA: LHC status and performance". In: *PoS* (2018), p. 001.

[26] CERN. *Project Schedule*. URL: `https://project-hl-lhc-industry.web.cern.ch/content/project-schedule`.

[27] Roman Adolphi et al. "The CMS experiment at the CERN LHC". In: *Jinst* 803 (2008), S08004.

[28] Aruna Nayak et al. "Particle-flow reconstruction and global event description with the CMS detector". In: (2017).

[29] V. Karimäki. "The CMS tracker system project: Technical Design Report". In: (1997). Ed. by M. Mannelli et al.

[30] "The CMS tracker: addendum to the Technical Design Report". In: (2000).

[31] S. Eidelman et al. "Review of Particle Physics". In: *Physics Letters B* 592.1 (2004). Review of Particle Physics, pp. 1–5. ISSN: 0370-2693. DOI: `https://doi.org/10.1016/j.physletb.2004.06.001`. URL: `https://www.sciencedirect.com/science/article/pii/S0370269304007579`.

[32] F Kircher et al. "Final design of the CMS solenoid cold mass". In: *IEEE transactions on applied superconductivity* 10.1 (2000), pp. 407–410.

[33] JG Layter. "The CMS muon project: Technical design report". In: *Technical Design Report CMS. CERN, Geneva* (1997).

[34] S. Dasu et al. "CMS. The TriDAS project. Technical design report, vol. 1: The trigger systems". In: (Dec. 2000).

[35] "CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger". In: (Dec. 2002). Ed. by P. Sphicas.

[36] Juan Antonio Aguilar-Saavedra et al. "Handbook of vectorlike quarks: Mixing and single production". In: *Physical Review D* 88.9 (2013), p. 094010.

[37] Abdelhak Djouadi and Alexander Lenz. "Sealing the fate of a fourth generation of fermions". In: *Physics Letters B* 715.4-5 (Sept. 2012), pp. 310–314. ISSN: 0370-2693. DOI: `10.1016/j.physletb.2012.07.060`. URL: `http://dx.doi.org/10.1016/j.physletb.2012.07.060`.

[38] G. Aad et al. "Search for pair production of heavy top-like quarks decaying to a high-p_T W boson and a b quark in the lepton plus jets final state at $\sqrt{s} = 7TeV$ with the ATLAS detector with the ATLAS detector". In: *Physics Letters B* 718.4-5 (Jan. 2013), pp. 1284–1302. ISSN: 0370-2693. DOI: `10.1016/j.physletb.2012.11.071`. URL: `http://dx.doi.org/10.1016/j.physletb.2012.11.071`.

[39] S. Chatrchyan et al. "Search for pair produced fourth-generation up-type quarks in pp collisions at $\sqrt{s} = 7TeV$ with a lepton in the final state with a lepton in the final state". In: *Physics Letters B* 718.2 (Dec. 2012), pp. 307–328. ISSN: 0370-2693. DOI: `10.1016/j.physletb.2012.10.038`. URL: `http://dx.doi.org/10.1016/j.physletb.2012.10.038`.

[40] Maxim Perelstein, Michael E. Peskin, and Aaron Pierce. "Top quarks and electroweak symmetry breaking in little Higgs models". In: *Physical Review D* 69.7 (Apr. 2004). ISSN: 1550-2368. DOI: `10.1103/physrevd.69.075002`. URL: `http://dx.doi.org/10.1103/PhysRevD.69.075002`.

[41] Roberto Contino, Leandro Da Rold, and Alex Pomarol. "Light custodians in natural composite Higgs models". In: *Physical Review D* 75.5 (Mar. 2007). ISSN: 1550-2368. DOI: `10.1103/physrevd.75.055014`. URL: `http://dx.doi.org/10.1103/PhysRevD.75.055014`.

[42] Roberto Contino et al. "Warped/composite phenomenology simplified". In: *Journal of High Energy Physics* 2007.05 (May 2007), pp. 074–074. ISSN: 1029-8479. DOI: `10.1088/1126-6708/2007/05/074`. URL: `http://dx.doi.org/10.1088/1126-6708/2007/05/074`.

[43] JoAnne L. Hewett and Thomas G. Rizzo. "Low-energy phenomenology of superstring-inspired E6 models". In: *Physics Reports* 183.5 (1989), pp. 193–381. ISSN: 0370-1573. DOI: `https://doi.org/10.1016/0370-1573(89)90071-9`. URL: `https://www.sciencedirect.com/science/article/pii/0370157389900719`.

[44] David B Kaplan, Howard Georgi, and Savas Dimopoulos. "Composite higgs scalars". In: *Physics Letters B* 136.3 (1984), pp. 187–190.

[45] Paolo Lodone. "Vector-like quarks in a"composite"Higgs model". In: *Journal of High Energy Physics* 2008.12 (2008), p. 029.

[46] Victor Mukhamedovich Abazov et al. "Search for single vector-like quarks in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV". In: (Oct. 2010). URL: `https://www.osti.gov/biblio/991304`.

[47] T. Aaltonen et al. "Search for a Heavy Toplike Quark in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV". In: *Phys. Rev. Lett.* 107 (26 Dec. 2011), p. 261801. DOI: `10.1103/PhysRevLett.107.261801`. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.107.261801`.

[48] Morad Aaboud et al. "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". In: *Journal of High Energy Physics* 2018.7 (2018), pp. 1–68.

[49] M Aaboud et al. "Search for pair production of heavy vector-like quarks decaying to high-p TW bosons and b quarks in the lepton-plus-jets final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *Journal of High Energy Physics* 2017.10 (2017), pp. 1–40.

[50] Morad Aaboud et al. "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". In: *Journal of High Energy Physics* 2017.8 (2017), pp. 1–40.

[51] Albert M Sirunyan et al. "Search for pair production of vector-like quarks in the $bW\bar{b}W$ channel from proton–proton collisions at $\sqrt{s} = 13TeV$". In: *Physics Letters B* 779 (2018), pp. 82–106.

[52] Albert M Sirunyan et al. "Search for pair production of vectorlike quarks in the fully hadronic final state". In: *Physical Review D* 100.7 (2019), p. 072001.

[53] CMS Collaboration. "Search for vector-like quarks in events with two oppositely charged leptons and jets in proton–proton collisions at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 79 (2019), p. 364.

[54] A. M. Sirunyan et al. "Search for pair production of vector-like T and B quarks in single-lepton final states using boosted jet substructure in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Journal of High Energy Physics* 2017.11 (Nov. 2017). ISSN: 1029-8479. DOI: `10.1007/jhep11(2017)085`. URL: `http://dx.doi.org/10.1007/JHEP11(2017)085`.

[55] Albert M Sirunyan et al. "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". In: *Physics letters B* 781 (2018), pp. 574–600.

[56] CMS collaboration et al. *Search for a singly produced vector-like quark B decaying to ab quark and a Higgs boson in a fully hadronic final state using boosted topologies*. Tech. rep. CMS-PAS-B2G-17-009, 2017.

[57] Albert M Sirunyan et al. "Search for electroweak production of a vector-like T quark using fully hadronic final states". In: *Journal of high energy physics* 2020.1 (2020), pp. 1–64.

[58] CMS collaboration et al. *Search for heavy vector-like quarks decaying to same-sign dileptons*. Tech. rep. CMS-PAS-B2G-16-019, 2016.

[59] CMS collaboration. URL: `https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsB2G#B2G_Preliminary_Results`.

[60] P.A. Zyla et al. "Review of Particle Physics". In: *PTEP* 2020.8 (2020), pp. 474–489. DOI: `10.1093/ptep/ptaa104`.

[61] Mordehai Milgrom. "MOND theory". In: *Canadian Journal of Physics* 93.2 (2015), pp. 107–118. DOI: `10.1139/cjp-2014-0211`. eprint: `https://doi.org/10.1139/cjp-2014-0211`. URL: `https://doi.org/10.1139/cjp-2014-0211`.

[62] Lord William Thomson Kelvin. *Baltimore lectures on molecular dynamics and the wave theory of light*. CUP Archive, 1904.

[63] JC Kapteyn. "First attempt at a theory of the arrangement and motion of the sidereal system, May 1922". In: *ApJ* 55 ().

[64] Jan H Oort. "The force exerted by the stellar system in the direction perpendicular to the galactic plane and some related problems". In: *Bulletin of the Astronomical Institutes of the Netherlands* 6 (1932), p. 249.

[65] Fritz Zwicky. "On the Masses of Nebulae and of Clusters of Nebulae". In: *The Astrophysical Journal* 86 (1937), p. 217.

[66] Nathaniel Craig and Andrey Katz. "The fraternal WIMP miracle". In: *Journal of Cosmology and Astroparticle Physics* 2015.10 (2015), p. 054.

[67] Deborah Pinna et al. "Single top quarks and dark matter". In: *Phys. Rev. D* 96 (3 Aug. 2017), p. 035031. DOI: `10.1103/PhysRevD.96.035031`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.96.035031`.

[68] Matthew R. Buckley, David Feld, and Dorival Gon çalves. "Scalar simplified models for dark matter". In: *Phys. Rev. D* 91 (1 Jan. 2015), p. 015017. DOI: `10.1103/PhysRevD.91.015017`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.91.015017`.

[69] G. D'Ambrosio et al. "Minimal flavour violation: an effective field theory approach". In: *Nuclear Physics B* 645.1-2 (Nov. 2002), pp. 155–187. ISSN: 0550-3213. DOI: `10.1016/s0550-3213(02)00836-2`. URL: `http://dx.doi.org/10.1016/S0550-3213(02)00836-2`.

[70] Gino Isidori and David M. Straub. "Minimal flavour violation and beyond". In: *The European Physical Journal C* 72.8 (Aug. 2012). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-012-2103-1`. URL: `http://dx.doi.org/10.1140/epjc/s10052-012-2103-1`.

[71] Yue Zhao and Kathryn M. Zurek. "Indirect detection signatures for the origin of asymmetric dark matter". In: *Journal of High Energy Physics* 2014.7 (July 2014). ISSN: 1029-8479. DOI: `10.1007/jhep07(2014)017`. URL: `http://dx.doi.org/10.1007/JHEP07(2014)017`.

[72] Albert M Sirunyan et al. "Search for dark matter produced in association with a single top quark or a top quark pair in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Journal of high energy physics* 2019.3 (2019), pp. 1–45.

[73] ATLAS collaboration et al. "Search for large missing transverse momentum in association with one top-quark in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *arXiv preprint arXiv:1812.09743* (2018).

[74] J. Andrea, B. Fuks, and F. Maltoni. "Monotops at the LHC". In: *Physical Review D* 84.7 (Oct. 2011). ISSN: 1550-2368. DOI: `10.1103/physrevd.84.074025`. URL: `http://dx.doi.org/10.1103/PhysRevD.84.074025`.

[75] T. Aaltonen et al. "Search for a Dark Matter Candidate Produced in Association with a Single Top Quark in $p\bar{p}$Collisions at $\sqrt{s} = 1.96$ TeV". In: *Physical Review Letters* 108.20 (May 2012). ISSN: 1079-7114. DOI: `10.1103/physrevlett.108.201802`. URL: `http://dx.doi.org/10.1103/PhysRevLett.108.201802`.

[76] V. Khachatryan et al. "Search for Monotop Signatures in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV". In: *Physical Review Letters* 114.10 (Mar. 2015). ISSN: 1079-7114. DOI: `10.1103/physrevlett.114.101801`. URL: `http://dx.doi.org/10.1103/PhysRevLett.114.101801`.

[77] G. Aad et al. "Search for invisible particles produced in association with single-top-quarks in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *The European Physical Journal C* 75.2 (Feb. 2015). ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-014-3233-4`. URL: `http://dx.doi.org/10.1140/epjc/s10052-014-3233-4`.

[78] A. M. Sirunyan et al. "Search for dark matter in events with energetic, hadronically decaying top quarks and missing transverse momentum at $\sqrt{s} = 13$ TeV". In: *Journal of High Energy Physics* 2018.6 (June 2018). ISSN: 1029-8479. DOI: `10.1007/jhep06(2018)027`. URL: `http://dx.doi.org/10.1007/JHEP06(2018)027`.

[79] CMS collaboration. URL: `https://twiki.cern.ch/twiki/bin/view/CMSPublic/SummaryPlotsEXO13TeV`.

[80] A.M. Sirunyan et al. "Particle-flow reconstruction and global event description with the CMS detector". In: *Journal of Instrumentation* 12.10 (Oct. 2017), P10003–P10003. ISSN: 1748-0221. DOI: `10.1088/1748-0221/12/10/p10003`. URL: `http://dx.doi.org/10.1088/1748-0221/12/10/P10003`.

[81] Matteo Cacciari, Gavin P Salam, and Gregory Soyez. "The anti-ktjet clustering algorithm". In: *Journal of High Energy Physics* 2008.04 (Apr. 2008), pp. 063–063. DOI: `10.1088/1126-6708/2008/04/063`. URL: `https://doi.org/10.1088/1126-6708/2008/04/063`.

[82] The CMS collaboration. "Identification of b-quark jets with the CMS experiment". In: *Journal of Instrumentation* 8.04 (Apr. 2013), P04013–P04013. DOI: `10.1088/1748-0221/8/04/p04013`. URL: `https://doi.org/10.1088/1748-0221/8/04/p04013`.

[83] E. Bols et al. "Jet flavour classification using DeepJet". In: *Journal of Instrumentation* 15.12 (Dec. 2020), P12012–P12012. ISSN: 1748-0221. DOI: `10.1088/1748-0221/15/12/p12012`. URL: `http://dx.doi.org/10.1088/1748-0221/15/12/P12012`.

[84] A.M. Sirunyan et al. "Identification of heavy, energetic, hadronically decaying particles using machine-learning techniques". In: *Journal of Instrumentation* 15.06 (June 2020), P06005–P06005. DOI: `10.1088/1748-0221/15/06/p06005`. URL: `https://doi.org/10.1088/1748-0221/15/06/p06005`.

[85] A.M. Sirunyan et al. "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}$=13 TeV". In: *Journal of Instrumentation* 13.06 (June 2018), P06015–P06015. DOI: `10.1088/1748-0221/13/06/p06015`. URL: `https://doi.org/10.1088/1748-0221/13/06/p06015`.

[86] Byron P. Roe et al. "Boosted decision trees as an alternative to artificial neural networks for particle identification". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 543.2 (2005), pp. 577–584. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/j.nima.2004.12.018`. URL: `https://www.sciencedirect.com/science/article/pii/S0168900205000355`.

[87] Luca Lista. *Statistical methods for data analysis in particle physics*. Vol. 941. Springer, 2017.

[88] 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". In: *Journal of High Energy Physics* 2014.7 (July 2014). ISSN: 1029-8479. DOI: `10.1007/jhep07(2014)079`. URL: `http://dx.doi.org/10.1007/JHEP07(2014)079`.

[89] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with parton shower simulations: the POWHEG method". In: *Journal of High Energy Physics* 2007.11 (2007), p. 070.

[90] Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. "PYTHIA 6.4 physics and manual". In: *Journal of High Energy Physics* 2006.05 (2006), p. 026.

[91] Manuel Bähr et al. "Herwig++ physics and manual". In: *The European Physical Journal C* 58.4 (Nov. 2008), pp. 639–707. ISSN: 1434-6052. DOI: `10.1140/epjc/s10052-008-0798-9`. URL: `http://dx.doi.org/10.1140/epjc/s10052-008-0798-9`.

[92] Sea Agostinelli et al. "GEANT4-a simulation toolkit". In: *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 506.3 (2003), pp. 250–303.

[93] CMS collaboration. *Documentation of the RooStats-based statistics tools for Higgs PAG*. URL: `https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideHiggsAnalysisCombinedLimit?rev=106`.