

UNIVERSITÀ DEGLI STUDI DI NAPOLI
“FEDERICO II”



Scuola Politecnica e delle Scienze di Base

Area Didattica di Scienze Matematiche Fisiche e Naturali

Dipartimento di Fisica “Ettore Pancini”

Laurea Magistrale in Fisica

**Search for a W' resonance decaying to a t and
a b quark with machine learning techniques**

Relatori:

Dott. Alberto Orso Maria Iorio

Candidato:

Michele Mormile
Matr. N94000565

Anno Accademico 2019/2020

Contents

Introduction	3
1 Standard Model of Particle Physics	5
1.1 Overview of the Standard Model	5
1.2 Quantum Electro-Dynamics	6
1.3 The Electroweak theory	8
1.3.1 Fermi theory of the weak interaction	8
1.3.2 The GSW Model	8
1.4 The Higgs mechanism	10
1.5 Quantum Chromo-Dynamics	12
2 The CMS experiment	14
2.1 The Large Hadron Collider	14
2.2 The Compact Muon Solenoid	16
2.2.1 The subdetector system of CMS	19
3 Physics BSM and W' models	24
3.1 The effective lagrangian	24
3.2 Topflavor model	26
3.3 W' decay channels	27
3.4 Search for the W' boson	28
4 Physics object selection and reconstruction	34
4.1 Physics objects selection	34
4.2 Muons	35
4.3 Jets	35
4.3.1 b-tagging	36
4.4 Missing transverse energy	36
4.5 Top quark reconstruction	37
4.6 Boosted Decision Trees	38
4.7 Top tagging with BDTs	40
4.8 Top tagging performances with BDTs	41
4.9 Deep Neural Networks	50
4.10 Top tagging with DNNs	52
4.11 Top tagging performances with DNNs	54

5	Application of ML algorithm and predictions on W' analysis	62
5.1	Simulated samples	62
5.1.1	Background samples	63
5.2	W' reconstruction	63
5.3	Analysis predictions	65
5.3.1	W' mass distributions	66
5.3.2	Sensitivity calculation	66
	Conclusions	73
	Bibliography	76

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 all the predicted particles have been discovered; the last one being the Higgs Boson, discovered 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 theories, or composite Higgs theories. Some of these theories, for instance, the left-right symmetrical models or the topflavour model, predict the existence of a W' boson: a hypothetical particle with the same quantum numbers of SM W boson, the mediator of the electroweak interactions, but holding a much higher mass. In the present thesis work, we have developed machine learning-based techniques to employ in the search for the W' boson, considering final states where it decays in a bottom quark and a top quark, with the latter decaying leptonically in a bottom quark, a muon and a muonic neutrino. Two machine learning algorithms have been designed in order to recognize and reconstruct (or *tag*) the top quark from its decay products: a deep neural network (DNN) and a boosted decision training (BDT). Machine learning is a class of algorithms able to perform classification, regression and clustering tasks through experience: in our cases, they have been trained on simulated samples where the correct top quark assignment was exactly known, before being used for the analysis on the W' . DNNs are algorithms based on artificial neurons: devices which try to imitate the functioning of biological neurons. On the other hand, BDTs are decision-making algorithms based on a series of selections (*cuts*) on the features of the object to classify. In the last part of this thesis work, we have developed an analysis strategy for the search for the W' boson, making use of Monte Carlo simulations, and made predictions on the potential outcome of the analysis, if performed on data taken at the Compact Muon Solenoid (CMS) experiment. CMS is one of the four main experiments present at the Large Hadron Collider at CERN, nearby Geneva. It is a general-purpose particle detector, i.e., it is designed to observe the products of the collisions produced by the accelerator in order to study a broad spectrum of physics. Its peculiarity is the large superconductive electromagnet which provides the experiment with a 3.8 T strong magnetic field. This thesis work is structured as follows:

- **Chapter 1** is dedicated to the SM, its features and its open problems;

- in **Chapter 2**, the LHC and the CMS experiment are discussed;
- **Chapter 3** is about the Beyond Standard Model (BSM) theories which predict the existence of the W' boson;
- **Chapter 4** discusses the physical object selection and reconstruction and treats the design of the machine learning algorithms;
- **Chapter 5** describes the analysis strategy developed for the search of the W' boson and its predictions.

Chapter 1

Standard Model of Particle Physics

The Standard Model of Particle Physics (SM) is a quantum field theory which accurately explains three of the four interactions between elementary particles: the electromagnetic, strong, and weak forces.

The SM describes particles and their interactions as quantum fields, and their dynamics is ruled by a characteristic Lagrangian Density function. The lagrangian arises from the gauge principle, after imposing it to be locally symmetric under the action of the group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, where the subscripts refer respectively to the colour charge, the left-hand chirality and the hypercharge. Furthermore, the SM includes the spontaneous symmetry breaking mechanism of the Higgs field, which attributes masses to the particles.

The SM predictions have been verified to a great degree of accuracy, but despite its remarkable success, it is not a complete theory: it does not predict several phenomena such as neutrino masses, and therefore neutrino oscillations, gravitational force or the existence of dark matter and dark energy.

1.1 Overview of the Standard Model

The SM describes the dynamics of particles and their interactions in terms of the lagrangian density \mathcal{L} , which is required to be locally invariant under the action of the group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. The three symmetry groups are:

- the special unitary group $SU(3)_C$, whose generators are the eight Gell-Mann matrices λ_i . The conserved quantity under these transformations is the color charge;
- the special unitary group $SU(2)_L$, whose generators are the three Pauli matrices τ_i . The quantity that is conserved under the transformations of this group is called weak isospin. Particles with left-handed chirality are put in weak isospin doublets, while particles with right-handed chirality represent a singlet;
- the unitary group $U(1)_Y$, whose generator is the identity matrix I and the conserved quantity is the hypercharge Y .

To each generator for every symmetry group corresponds a gauge boson which mediates the interaction between particles. The SM has therefore 12 gauge bosons, also described in tab. 1.1:

- 8 gluons which mediate the strong interaction between quarks and arise from the $SU(3)_C$ gauge symmetry;
- 3 vector bosons W^\pm , Z and the photon γ , which are generated through the spontaneous symmetry breaking mechanisms from the bosons $W_{1,2,3}$ and B which are the gauge bosons of the group $SU(2)_L \otimes U(1)_Y$.

Boson	Electric charge	Mass[1]	Spin	Interaction
Gluon	0	0	1	Strong
Photon	0	0	1	Electromagnetic
W^\pm	± 1	$80.379 \pm 0.012 \text{ GeV}$	1	Weak
Z	0	$91.1876 \pm 0.0021 \text{ GeV}$	1	

Table 1.1: The gauge bosons of the Standard Model.

The SM also includes 12 fermions (tab. 1.2), of spin $\frac{1}{2}$, divided into three families of quarks and leptons, whose bound states allow to form stable matter. All particles are subjected to the weak interaction, charged leptons and quarks also have electromagnetic charge, while quarks also interact through the strong interaction. The colour confinement, typical of the strong interaction, prevents us to observe quarks as free particles, as they are always found in bound states called hadrons.

For each particle, the Standard Model also includes an antiparticle, which has the same mass of the corresponding particle and opposite quantum numbers.

1.2 Quantum Electro-Dynamics

Quantum Electro-Dynamics (QED) is the theory within the Standard Model which describes the electromagnetic interaction. The lagrangian density for the QED can be obtained from the free lagrangian density by imposing it to be locally invariant under the action of the unitary group \mathcal{U}_1 . The free lagrangian density is:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi, \quad (1.1)$$

where ψ is a Dirac-bispinor field and m is the mass of the field itself. The group \mathcal{U}_1 is responsible for the transformation

$$\psi \rightarrow \psi' = e^{i\theta(x)}\psi. \quad (1.2)$$

The free lagrangian, under this gauge transformation becomes

$$\mathcal{L}' = i\bar{\psi}'\gamma^\mu\partial_\mu\psi' - m\bar{\psi}'\psi'. \quad (1.3)$$

Particle	Electric charge	Mass [1]	
u	$+\frac{2}{3}$	$2.16^{+0.49}_{-0.26} \text{ MeV}$	quarks
d	$-\frac{1}{3}$	$4.67^{+0.48}_{-0.17} \text{ MeV}$	
c	$+\frac{2}{3}$	$1.27 \pm 0.02 \text{ GeV}$	
s	$-\frac{1}{3}$	$93^{+11}_{-5} \text{ MeV}$	
t	$+\frac{2}{3}$	$172.76 \pm 0.30 \text{ GeV}$	
b	$-\frac{1}{3}$	$4.18^{+0.03}_{-0.02} \text{ GeV}$	
e	-1	0.511 MeV	leptons
μ	-1	105.658 MeV	
τ	-1	$1776.86 \pm 0.12 \text{ MeV}$	
$\nu_{e,\mu,\tau}$	0	$< 1.1eV$	

Table 1.2: The fermions of the Standard Model. Electron and muon mass are known with a relative uncertainty of order 10^{-9} , here we report only the most significant digits.

By expanding the transformed terms and taking the derivatives one finds:

$$\mathcal{L}' = i\bar{\psi}\gamma^\mu\partial_\mu\psi - \partial_\mu\theta(x)\bar{\psi}\gamma^\mu\psi - m\bar{\psi}\psi. \quad (1.4)$$

The transformed lagrangian density differs from the original as it presents an additional term, hence it is not locally invariant. We can use the gauge principle in order to add a new interaction field to the lagrangian, whose gauge properties are the ones needed to impose the local invariance of the lagrangian. The electromagnetic four-potential A^μ does satisfy these requirements, as its transformation law is:

$$A^\mu \rightarrow A'^\mu = A^\mu - \partial^\mu\Lambda, \quad (1.5)$$

where Λ is any twice continuously differentiable function that depends on position and time. The local invariance is then obtained by replacing the derivative ∂_μ with the covariant derivative

$$D_\mu = \partial_\mu + ieA_\mu, \quad (1.6)$$

where e is the electric charge of the bispinor field. Substituting the 1.6 into the 1.1:

$$\mathcal{L} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - e\bar{\psi}\gamma^\mu A_\mu\psi - m\bar{\psi}\psi. \quad (1.7)$$

The new interaction term in the 1.7, under a local gauge transformation, cancels the extra term in the 1.4, and grants local gauge invariance to the lagrangian density.

The lagrangian density for the QED is completed by a kinetic term for the electromagnetic field:

$$\mathcal{L}_{QED} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - e\bar{\psi}\gamma^\mu A_\mu\psi - m\bar{\psi}\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}, \quad (1.8)$$

where $F^{\mu\nu}$ is the electromagnetic field tensor

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \quad (1.9)$$

1.3 The Electroweak theory

The weak force is responsible for various diverse phenomena in particle physics, usually rare decays of heavy particles, such as the τ lepton or hadrons. The β -decay of nuclei is an important example of event due to this interaction. In fact, Enrico Fermi, in 1934, first formulated a description of the weak force in an attempt to describe the this phenomenon.

1.3.1 Fermi theory of the weak interaction

Fermi described the weak interaction as a current-current contact interaction without a mediator boson. The lagrangian term for the interaction in the Fermi theory is:

$$\mathcal{L} = \frac{G_F}{\sqrt{2}} J_\mu^\dagger(x) J^\mu(x), \quad (1.10)$$

where $G_F = 1.16638 \times 10^{-5} GeV^{-2}$ is the Fermi constant, which is equivalent to the charge of the fermions in the 1.7 and $J^\mu(x)$ is the weak current, which can be either leptonic or hadronic.

In the case of leptonic weak currents, which only involve elementary particles, experimental data show that the current has a Vectorial - Axial structure. For instance, the current associated to the transition from an electron to an electronic neutrino is:

$$J^\mu(x) = \bar{e}(x)\gamma^\mu(1 - \gamma^5)\nu(x). \quad (1.11)$$

Each Dirac bispinor can be decomposed in the sum of a left-handed and a right-handed chiral bispinor

$$\psi = \psi_L + \psi_R. \quad (1.12)$$

The left-chiral projector P_L is proportional to $(1 - \gamma^5)$, hence, according to the Fermi theory, only left-handed particles take part in weak interaction. However, this has been disproven by experimental data for charged leptons, as they do couple to weak neutral currents, even though the coupling constants are different for the different chiralities.

Furthermore, Fermi theory is less and less accurate as we consider higher energy transferred in the interaction, as it violates the unitarity of the scattering matrix. Nevertheless, Fermi theory is a useful device to describe low-energy weak interactions, and it also paved the way for later theories such as the GSW Model.

1.3.2 The GSW Model

The GSW Model is a gauge theory proposed by Sheldon Glashow, Abdus Salam and Steven Weinberg to overcome the problems encountered by the Fermi theory of the weak interaction. The lagrangian for this model is obtained making use of the gauge principle, by requesting the free one to be locally invariant under the action of the group

$$SU(2)_L \otimes U(1)_Y, \quad (1.13)$$

where the subscript L refers to the fact that only left-handed chirality particles undergo the interaction and the subscript Y refers to the hypercharge, which is the quantity conserved by the interaction.

In this model, left-handed fermions are coupled in doublets which are eigenstates of the weak isospin I , shown in tab. 1.3, where ℓ^- represents the charged lepton and ν_ℓ its associated neutrino; U represents an up-type quark and D' the associated down-type quark weak eigenstate, which is a linear combination of the mass eigenstates. I_3 is the third component of the weak isospin. Right-handed fermions are weak isospin singlets, hence are eigenstates with eigenvalue $I = 0$. The unitary mixing matrix between mass and weak eigenstates of the down-type

$$\begin{pmatrix} \nu_\ell \\ \ell^- \end{pmatrix}_L \quad \begin{pmatrix} U \\ D' \end{pmatrix}_L \quad \begin{matrix} I_3 = +\frac{1}{2} \\ I_3 = -\frac{1}{2} \end{matrix}$$

Table 1.3: The weak isospin doublets

quark is called CKM (Cabibbo-Kobayashi-Maskawa):

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \quad (1.14)$$

The CKM matrix can be expressed in the Wolfenstein parametrization, which uses 4 parameters:

$$V_{CKM} = \begin{pmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{pmatrix} + O(\lambda^4). \quad (1.15)$$

The measurements of A and λ are [1]:

$$\lambda = 0.22650 \pm 0.00048 \quad A = 0.790_{-0.012}^{+0.017}. \quad (1.16)$$

. This shows the quasi-diagonal nature of the CKM matrix, which causes up-type quarks to have the highest probability to interact with the down-type quark of the same generation.

The lagrangian for the electroweak interaction can again be obtained applying the gauge principle to the free lagrangian, that is by replacing the derivative with the covariant derivative

$$D_\mu = \partial_\mu + ig\frac{\vec{\tau}}{2} \cdot \vec{W}_\mu + ig'\frac{Y}{2}B_\mu, \quad (1.17)$$

where g and g' correspond to the electric charge in the 1.6 and the fields $W_\mu^{1,2,3}$ and B_μ have the gauge properties to ensure the local invariance of the lagrangian under the action of the group $SU(2)_L \otimes U(1)_Y$. An $SU(2)_L$ transformation can be expressed as:

$$\psi_L \rightarrow \psi'_L = e^{ig\frac{\vec{\tau}}{2} \cdot \vec{\alpha}} \psi_L, \quad (1.18)$$

where $\vec{\alpha}$ is the parameter of the transformation. The gauge property for the W_μ^i field is thus

$$W_\mu^i \rightarrow W_\mu'^i = W_\mu^i - \partial_\mu \alpha^i - g \epsilon^{ijk} \alpha^j W_\mu^k. \quad (1.19)$$

The gauge properties for the group $\mathcal{U}(1)$ are been discussed in sec. 1.2. The gauge transformation law for the B_μ field is:

$$B_\mu \rightarrow B'_\mu = B_\mu - \partial_\mu \Lambda. \quad (1.20)$$

Hence the interaction terms between the fermions and the fields:

$$\mathcal{L}_{EW}^I = -g \psi_L \gamma^\mu \frac{\vec{\tau}}{2} \cdot \vec{W}_\mu \psi_L - g' \frac{Y}{2} \psi \gamma^\mu B_\mu \psi, \quad (1.21)$$

where the ψ are Dirac bispinors and the ψ_L are doublets of left-handed Dirac bispinors.

The lagrangian for the model is again completed by the kinetic terms for the interaction fields. The one for the \vec{W}_μ is

$$\mathcal{L}_K = -\frac{1}{4} \vec{W}^{\mu\nu} \cdot \vec{W}_{\mu\nu} \quad (1.22)$$

$$\vec{W}_{\mu\nu} = \partial_\mu \vec{W}_\nu - \partial_\nu \vec{W}_\mu - g \vec{W}_\mu \times \vec{W}_\nu, \quad (1.23)$$

while the one for B_μ is analogous to the one in 1.8. The cross product in the definition of $\vec{W}_{\mu\nu}$ causes the lagrangian to include W-boson self-interaction terms which contain triple and quartic gauge boson vertices.

This formulation still presents some discrepancies with respect to the experiments. For instance, it does not predict a mass term for the mediator bosons of the weak interaction, as it would break the gauge invariance of the lagrangian. Nevertheless, experimental data show these bosons to be massive. Furthermore, this theory does not explain the interaction of right-handed chirality charged particles via neutral weak current. These problems are solved through the spontaneous symmetry breaking mechanism of the Higgs field.

1.4 The Higgs mechanism

The Higgs mechanism is an elegant theory, developed by Peter Higgs, Robert Brout, and François Englert in 1964 [2] [3], which explains the mass of the W^\pm bosons as the result of their interaction with the, at the time only hypothetical, Higgs boson field, whose corresponding particle was discovered only in 2012 [4] [5].

The Higgs field consists of a doublet of complex scalar fields:

$$\phi = \begin{pmatrix} \phi^+(x) \\ \phi^0(x) \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1(x) + i\phi_2(x) \\ \phi_3(x) + i\phi_4(x) \end{pmatrix}. \quad (1.24)$$

The lagrangian for the Higgs field is

$$\mathcal{L} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi) \quad (1.25)$$

$$= (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2. \quad (1.26)$$

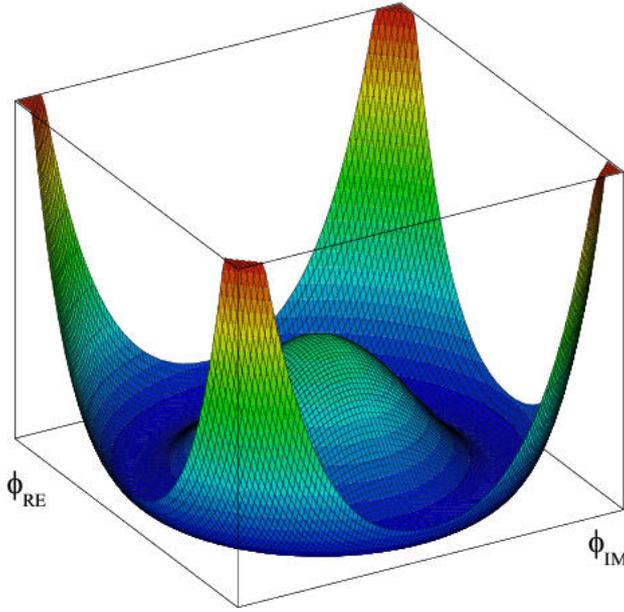


Figure 1.1: The potential of the Higgs field for $\mu^2 < 0$.

The potential (fig. 1.1), for $\mu^2 < 0$, assumes a minimum

$$\phi^{02} = -\frac{\mu^2}{2\lambda} = -\frac{v}{2}, \quad (1.27)$$

which represents a circumference in the complex plane. We can then choose a reference minimum without loss of generality and expand the field about it in order to get

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1(x) + i\phi_2(x) \\ v + \phi_3(x) + i\phi_4(x) \end{pmatrix}. \quad (1.28)$$

After the symmetry breaking, the photon is required to remain massless, and therefore the minimum of the potential must correspond to a non-zero vacuum expectation value only of the neutral scalar field ϕ^0 . We can then express the field as

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \quad (1.29)$$

By requesting the lagrangian in the 1.26 to be locally invariant under the action of the symmetry in 1.13 and then diagonalizing the mass matrices that appear, we obtain the physical bosons with their mass terms, which depend on

the interaction of the gauge bosons with the Higgs field:

$$\mathcal{L} = \frac{1}{2} \partial_\mu H \partial^\mu H + \quad (1.30)$$

$$- \frac{1}{4} (\partial_\mu W_\nu^1 - \partial_\nu W_\mu^1) (\partial^\mu W^{1\nu} - \partial^\nu W^{1\mu}) + \frac{1}{8} v^2 g^2 W_\mu^1 W^{1\mu} + \quad (1.31)$$

$$- \frac{1}{4} (\partial_\mu W_\nu^2 - \partial_\nu W_\mu^2) (\partial^\mu W^{2\nu} - \partial^\nu W^{2\mu}) + \frac{1}{8} v^2 g^2 W_\mu^2 W^{2\mu} + \quad (1.32)$$

$$- \frac{1}{4} (\partial_\mu Z_\nu - \partial_\nu Z_\mu) (\partial^\mu Z^\nu - \partial^\nu Z^\mu) + \frac{1}{8} v^2 (g^2 + g'^2) Z_\mu Z^\mu + \quad (1.33)$$

$$- \frac{1}{4} (F_{\mu\nu} F^{\mu\nu}). \quad (1.34)$$

Hence we find that the mass for the physical W^\pm bosons is

$$m_W = \frac{1}{2} g v, \quad (1.35)$$

while the mass for the Z boson is:

$$m_Z = \frac{1}{2} v \sqrt{g^2 + g'^2} = \frac{m_W}{\cos\theta_W}, \quad (1.36)$$

where $\cos\theta_W = \frac{g}{\sqrt{g^2 + g'^2}}$ is the Weinberg's angle, which represents the rotation needed to obtain the physical Z and γ from the fields W^3 and B

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix} \quad (1.37)$$

This mechanism also explains why the weak neutral current is able to couple to both left-handed and right-handed chirality charged particles, as the physical Z boson is a combination of the neutral bosons W^3 and B , the latter coupling to particles with no distinction in chirality.

1.5 Quantum Chromo-Dynamics

Quantum Chromo-Dynamics (QCD) is the part of the Standard Model which describes the behavior of quarks under the influence of the strong interaction. It is a gauge theory and the symmetry group is $\mathcal{SU}(3)_C$, whose infinitesimal generators are

$$T_{1,\dots,8} = \frac{\lambda_{1,\dots,8}}{2}, \quad (1.38)$$

where the $\lambda_{1,\dots,8}$ are the the Gell-Mann matrices. The commutation rules for the group are:

$$[T_\alpha, T_\beta] = f_{\alpha\beta\gamma} T_\gamma, \quad (1.39)$$

where the $f_{\alpha\beta\gamma}$ are the fine structure constants for the group $\mathcal{SU}(3)$. A local transformation for this group is:

$$\psi \rightarrow \psi' = e^{ig_s \vec{\theta}(x) \cdot \vec{T}} \psi, \quad (1.40)$$

where g_s is the strong coupling constant, $\vec{\theta}(x)$ are the parameters of the transformation. Since the \vec{T} are 3x3 matrices, ψ represents a triplet of Dirac bispinor. This additional degree of freedom is called "color", hence the name Chromo-Dynamics.

The gauge principle leads us to definition of the covariant derivative for this symmetry group

$$D_\mu = \partial_\mu + ig_s \frac{\vec{\lambda}}{2} \cdot \vec{G}_\mu, \quad (1.41)$$

where \vec{G}_μ are the eight gluon fields which are needed to ensure the local invariance of the lagrangian and respect the following gauge transformation law:

$$G_\mu^i \rightarrow G_\mu'^i = G_\mu^i + ig_s f_{ijk} \theta^j(x) G_\mu^k. \quad (1.42)$$

Once again, we need to complete the lagrangian with a kinetic term for the gluon field and obtain the full lagrangian for the QCD

$$\mathcal{L}_{QCD} = \bar{\psi} \gamma^\mu \partial_\mu \psi - m \bar{\psi} \psi - g_s \bar{\psi} \gamma^\mu \frac{\vec{\lambda}}{2} \cdot G_\mu \psi - \frac{1}{4} \vec{G}_{\mu\nu} \vec{G}^{\mu\nu}, \quad (1.43)$$

with

$$G_{\mu\nu}^i = \partial_\mu G_\nu^i - \partial_\nu G_\mu^i - g_s f^{ijk} G_\mu^j \times G_\nu^k. \quad (1.44)$$

The kinetic term in the 1.43 hence contains terms of self-interactions between gluons, as it was the case for the weak interaction as well. Both these effects are due to the non-abelian nature of the symmetry groups $SU(2)$ and $SU(3)$.

The QCD presents two peculiar properties, which are asymptotic freedom and colour confinement. Asymptotic freedom causes the strong interaction to become asymptotically weaker as the transferred momentum $|q|$ increases. At $|q| \sim 1 \text{ GeV}$, $g_s \sim O(1)$, which does not allow perturbation theory to be used, while at $|q| \sim 100 \text{ GeV}$, $g_s \sim O(0.1)$, which is a perturbative regime, instead. This is due to the renormalization of the QCD theory, which bears the following law for the running of g_s :

$$g_s(q^2) = \frac{g_s(\mu^2)}{1 + \frac{11N_c - 2N_f}{12\pi} g_s(\mu^2) \ln\left(\frac{q^2}{\mu^2}\right)}, \quad (1.45)$$

where $g_s(\mu^2)$ is the strong coupling constant at a certain energy scale μ^2 , $N_c = 3$ colors and N_f is the number of flavours available at transferred momentum q^2 .

Color confinement is an hypothesis with yet no analytic proof, that tries to explain why no free quarks has ever been observed. Under this assumption, colored objects are always confined color singlet states and therefore objects with non-zero colour charge cannot propagate as free particles. This property results in the fact that if one is able to break apart a hadron, like a proton in LHC, in the constituent quarks, they would immediately generate jets of hadrons in order to create color singlet states. There are theories beyond Standard Model which try to explain hadronization through non-perturbative QCD.

Chapter 2

The CMS experiment

The Compact Muon Solenoid (CMS) is one of the four experiments which take place at the Large Hadron Collider (LHC) at the European Organization for Nuclear Research (CERN) in Geneva. CERN was founded in 1954 and throughout its history it has been one of the most important laboratories for high energy physics. In fact, important discoveries were made at CERN, such as the discovery of the W and Z bosons at UA1 and UA2 experiments in 1983 [6] [7] [8] and the discovery of the Higgs boson at the LHC experiments ATLAS and CMS in 2012 [4] [5].

2.1 The Large Hadron Collider

The Large Hadron Collider is the world's largest and highest-energy particle collider, where two beams of particles travel at about the speed of light and are made collide in 4 points of intersections, where the four main experiments are located. The LHC can accelerate both protons and lead ions. These experiments are: Compact Muon Solenoid (CMS), A Toroidal LHC ApparatuS (ATLAS), A Large Ion Collider Experiment (ALICE), LHC-beauty (LHC-b).

LHC is the last stage of a complex of different accelerators: each machine brings the energy of the particle beams to a certain threshold before injecting them into the next stage. The protons which are accelerated in the complex are obtained from a hydrogen tank, using an electric field to strip the atoms of their electrons. The lead ions are instead obtained from a source of vaporized lead. The protons are injected into:

- Linac 2, a linear accelerator, which accelerates them up to 50 MeV of energy;
- Proton Synchrotron Booster (PBS), which brings them to 1.4 GeV;
- Proton Synchrotron (PS), which accelerates them to 25 GeV;
- Super Proton Synchrotron (SPS), which raises the energy to 450 GeV before finally injecting the beam into LHC.

A schematic view of LHC can be seen in fig. 2.1. In LHC, two particle beams travel

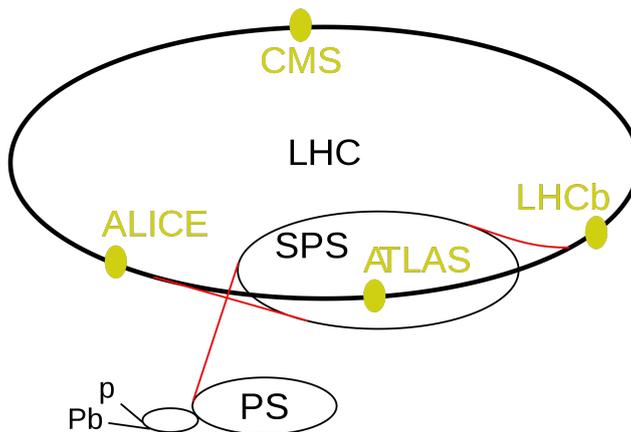


Figure 2.1: A schematic view of the accelerator complex at CERN.

in opposite directions in two different beam pipes. The accelerator is composed of a 27-km-circumference ring filled with superconductive magnets, which provide the accelerator with the high magnetic field necessary to keep the particles in their path and to accelerate them to the requested energy. The maximum energy per proton beam at LHC is 6.5 TeV, which brings the energy of the centre of mass of the collisions to 13 TeV, making it the world's highest-energy particle collider.

An important parameter for a particle accelerator is the Luminosity \mathcal{L} , which is related to the rate of events R in the machine through the total cross section:

$$R = \mathcal{L}\sigma, \quad (2.1)$$

where σ is the total cross section for the collision.

The parameter \mathcal{L} only depends on the characteristics of the accelerator and can be calculated as

$$\mathcal{L} = \frac{N_b^2 n_b f_{rev} \gamma}{4\pi\epsilon\beta^*} F \quad (2.2)$$

where

- N_b is the number of protons in a bunch;
- n_b is the number of bunches per beams;
- f_{rev} is the frequency of revolutions of the beams;
- γ is the relativistic Lorentz factor;
- ϵ is the normalized emittance;
- β^* is a measure of how narrow and collimated the beams are at the interaction point: it can be seen as the product of the standard deviations of the particle positions in the x-y plane;

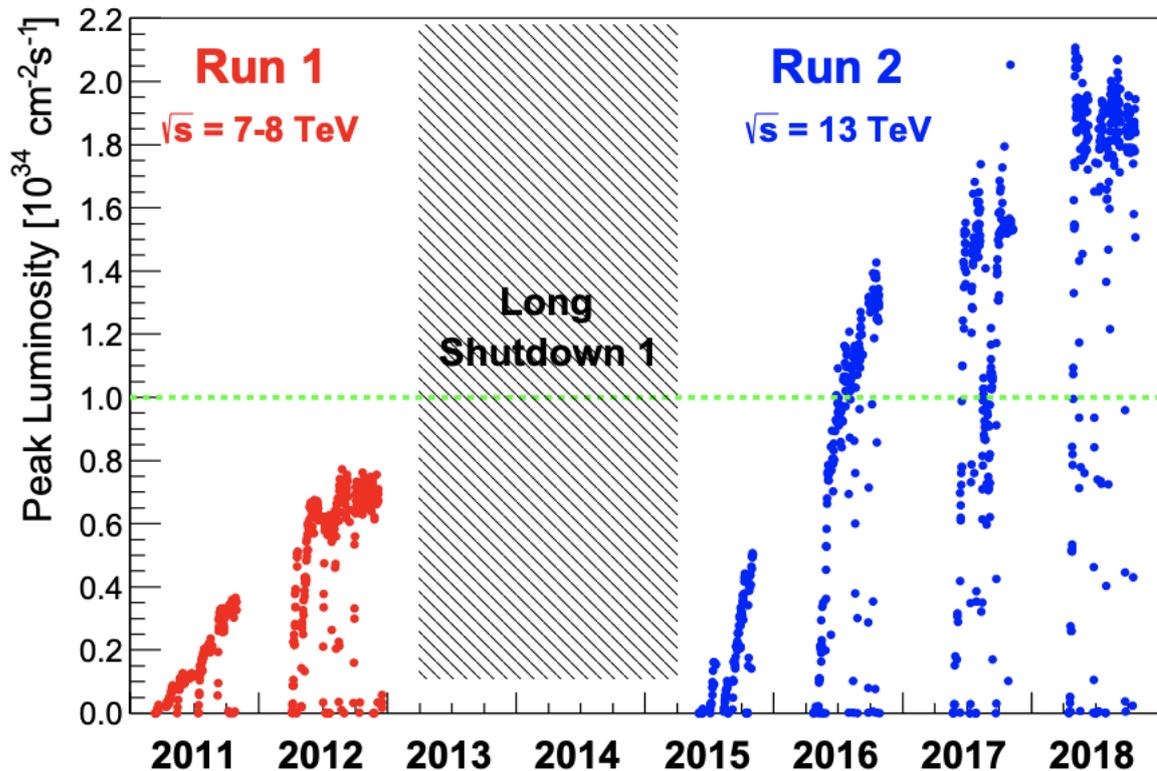


Figure 2.2: LHC peak luminosity for the 2011-2018 period.

- F is the geometric luminosity reduction factor due to the crossing angle at the interaction point.

The trend of the LHC peak luminosity between 2011 and 2018 is shown in fig. 2.2, where notably the peak luminosity achieved has already reached its design value of $2.0 \times 10^{34} \text{cm}^{-2} \text{s}^{-1}$ [9].

Furthermore, another important quantity is the integrated luminosity:

$$L = \int_{t_1}^{t_2} \mathcal{L}(t) dt, \quad (2.3)$$

where t_1 and t_2 are the time extremes in which the collider is in function. It is related to the number of total events:

$$N = L\sigma. \quad (2.4)$$

In the CMS Run II 137fb^{-1} equivalent of data have been collected and validated.

2.2 The Compact Muon Solenoid

CMS is a general-purpose particle detector, i.e., it is designed to observe the products of the collisions produced by the accelerator in order to study a broad

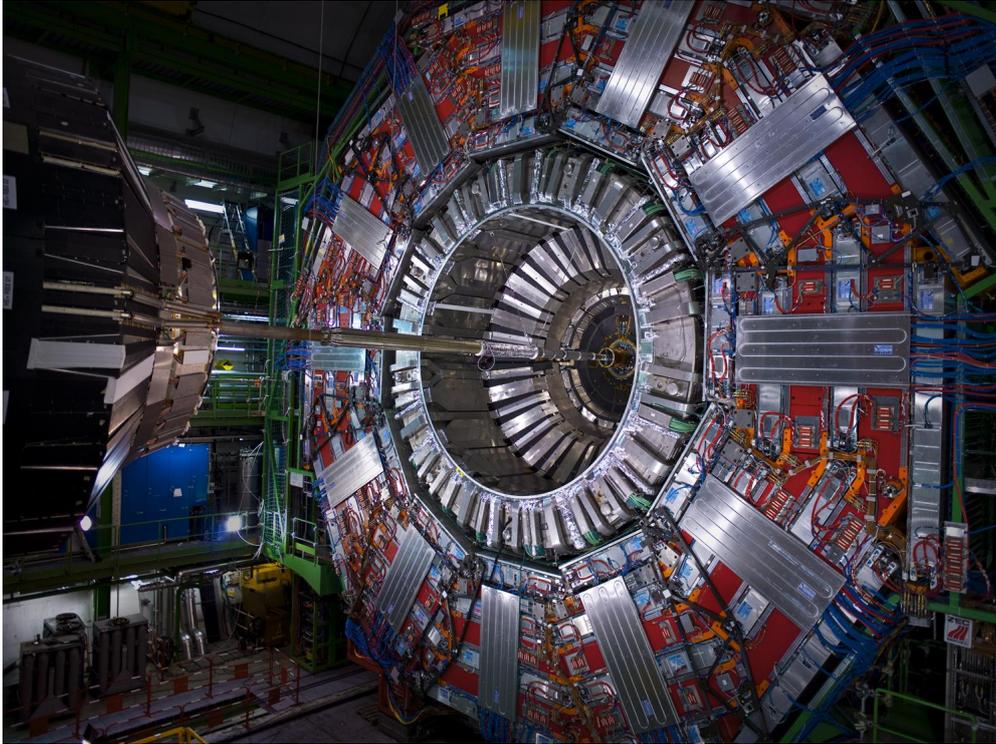


Figure 2.3: A picture of CMS.

spectrum of physics. It is able to detect Standard Model phenomena as well as events that could lead to new physics beyond standard model. It is built around a superconductive electromagnet (the "Solenoid" in the CMS acronym) which provides the apparatus with a very strong magnetic field of 3.8 T, allowing to bend the trajectory of charged particles and to therefore precisely measure their momentum. The detector is in the shape of a cylinder 21 meters high and with a 15-meter-long diameter, disposed horizontally and centered around the collision point to exploit the cylindrical symmetry of the interaction. Fig. 2.3 shows a picture of CMS opened during maintenance in Run 2. CMS is divided in different subdetectors, each designed for specific measurements and with precise requirements. Some general requirements are however shared within all the subdetectors, such as the high radiation hardness, because the detector is designed to work for a long time in a high radiation environment, due to high luminosity of the accelerator; high spatial resolution, needed to distinguish particles which cross the detector in close positions.

The experiment uses a right-handed cartesian system of coordinates to describe the collision events (fig. 2.4):

- **x-axis** points to the center of the LHC ring;
- **y-axis** points upwards, orthogonal to the ring;
- **z-axis** is along the beam pipe in counterclockwise direction.

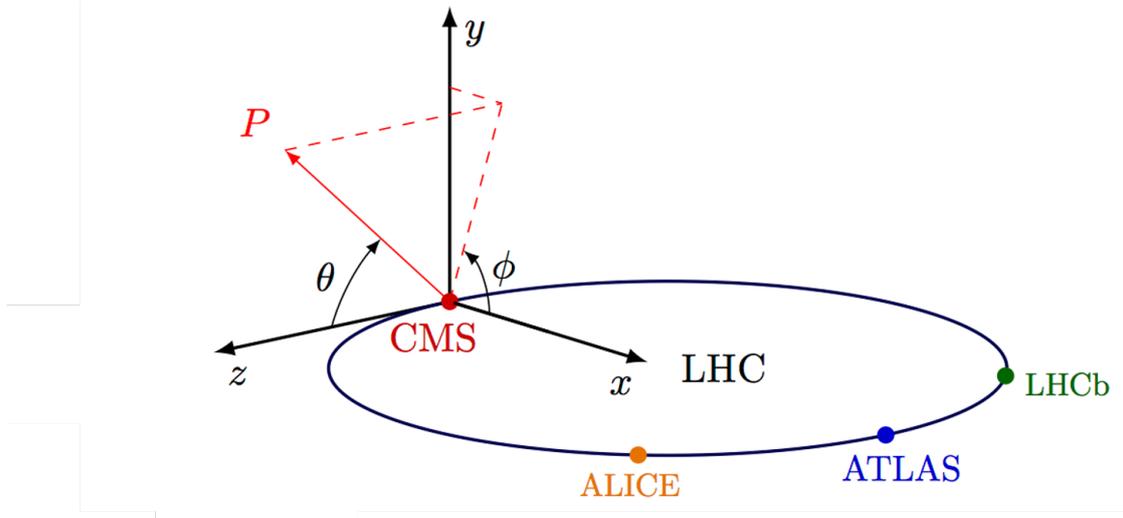


Figure 2.4: The coordinate system in CMS.

Furthermore, a cylindrical coordinate system is used to describe the quantities related to the particles arising from the p-p collision:

- the radial distance r from the z -axis;
- the azimuthal angle ϕ around the z -axis, increasing from the x -axis clockwise in the positive direction of the z ;
- the polar angle θ around the x -axis, increasing from the z -axis clockwise in the positive direction of the x .

Usually, an alternate variable is used instead of θ : the pseudo-rapidity

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right), \quad (2.5)$$

which transforms linearly under the effect of boosts along the z -axis, in ultra-relative approximation. Also, the difference in pseudorapidity between two particles is invariant under this kind of boosts.

Other important invariants under a Lorentz boost along the z -axis are:

- the radial distance:

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}; \quad (2.6)$$

- the transverse momentum \vec{p}_t and its module p_t :

$$p_t = \sqrt{p_x^2 + p_y^2}; \quad (2.7)$$

- the transverse energy:

$$E_t = E \sin\theta. \quad (2.8)$$

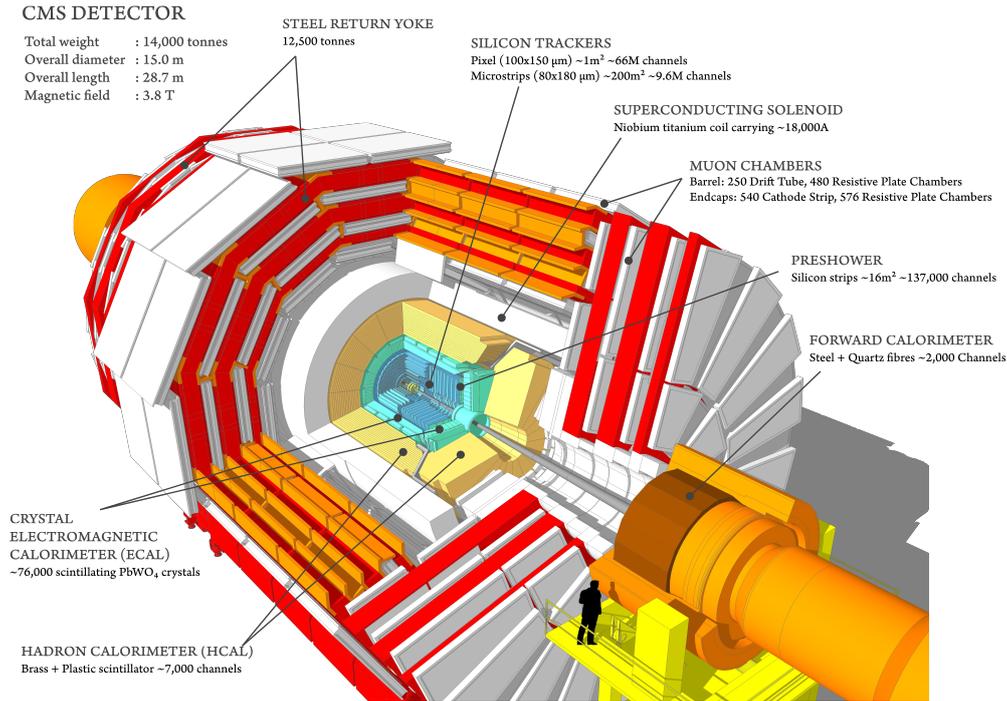


Figure 2.5: The subdetector structure of CMS.

2.2.1 The subdetector system of CMS

As displayed in fig. 2.5, CMS is composed of several subdetectors and components:

- tracking system;
- electromagnetic calorimeter (ECAL);
- hadronic calorimeter (HCAL);
- superconductive solenoid;
- muon system.

Each subdetector is composed of a barrel: a cylindrical layer which is coaxial with the beam pipe; and two endcaps: plane layers placed at the two ends of the cylinder to ensure the detection of particles which travel close to the z -axis.

The tracker

The tracker is designed to measure the position of charged particles arising from the collision. It is the closest subdetector to the interaction point, so to allow to best-possible accuracy in the reconstruction of the secondary vertices of interaction. The tracker has a diameter of 2.5 m and a length of 5.8 m, with an acceptance in η of $|\eta| < 2.4$. It has an active area of 200m^2 and is composed of the inner tracker, made of silicon pixel detectors, and the outer tracker, made of silicon strip detectors. The inner detector has an accuracy of $10 \mu\text{m}$ and of

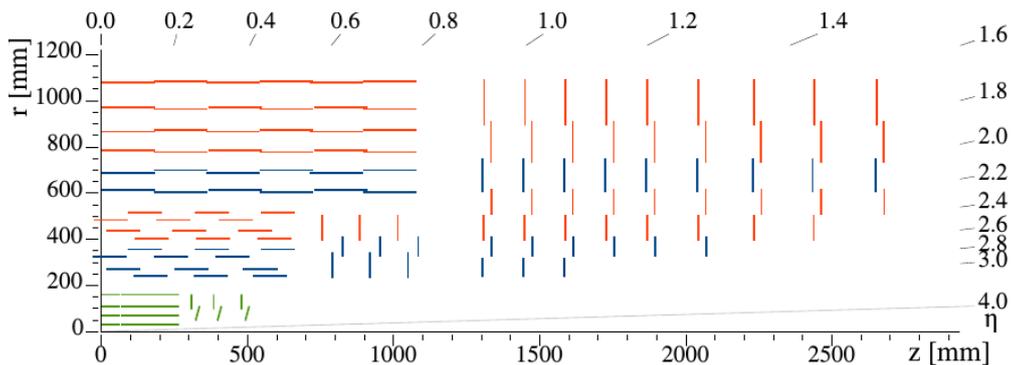


Figure 2.6: Sketch of one quarter of the tracker of CMS in r-z view. Green indicates the inner tracker, while blue and red indicate the outer tracker.

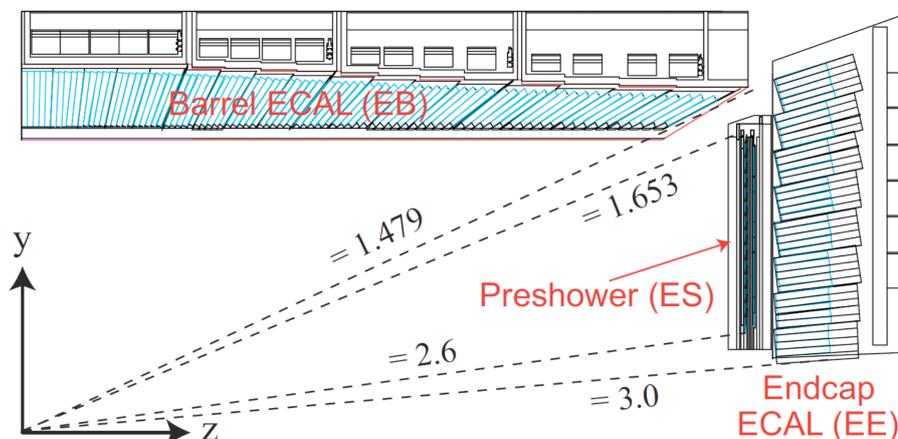


Figure 2.7: Layout of the ECAL of CMS.

20 μm respectively for the radial and transverse positions and was designed with 60 millions pixels distributed in three barrel layers and the two endcaps. After an upgrade, performed between the years 2016 and 2017, it now features 124 millions of pixel detectors distributed in four barrel layers and three endcaps [10]. Additionally, the outer tracker has an accuracy between 35 μm and 52 μm for the radial position and 530 μm in the transverse position. It is composed of 4 barrel layers and 3 endcaps layers. The layout of the tracker is shown in fig. 2.6.

The electromagnetic calorimeter

The ECAL is a hermetic homogeneous calorimeter made of lead tungstate ($PbWO_4$) crystals. It is used for the measurement of the total energy of electrons and photons which enter the detector. The ECAL is placed at a radial distance between 1.25 m and 1.8 m. The photons produced by the passage of the particles in the crystals are collected by different photodetectors: avalanche photodiodes (APDs) in the barrel and vacuum phototriodes in the endcaps. APDs cannot be used in the endcaps as the radiation damage in that area is too high for APDs. The layout for this subdetector is shown in fig. 2.7.



Figure 2.8: Schematic view of CMS where the HCAL is highlighted.

The hadronic calorimeter

The HCAL is sampling calorimeter made of layers of active material and absorber. The active material is made of fluorescent scintillator, the absorber is Cartridge Brass, which possesses a radiation length $X_0 = 1.49 \text{ cm}$ and a nuclear interaction length $\lambda_I = 16.42 \text{ cm}$, enough to contain the development of the hadronic shower in the thickness of the detector. The main purpose of the hadronic calorimeter is the measurement of the energy of hadrons produced in the collision event. The main sections of the HCAL are the Barrel (HB) and the Endcap (HE), using the technology described so far, and covering $|\eta| < 3.0$. In addition to those, an Outer Barrel section (HO) and a Forward Section (HF) are presents. A view of all four HCAL section is displayed in fig. 2.8.

The HF are 11.2 m away from the interaction point along the z-axis, covering the pseudorapidity range $3 < |\eta| < 5.2$. It is made of quartz fibres embedded within a 165 cm long steel absorber and exploits a Cherenkov-based technology. The HO, instead, is composed by additional scintillators that are placed outside the solenoid to ensure adequate sampling depth and to measure late shower development. A fundamental parameter useful to describe the performances of both the calorimeters is the energy resolution, which can be seen as:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{a}{\sqrt{E}}\right)^2 + \left(\frac{b}{E}\right)^2 + c^2, \quad (2.9)$$

where

- a is a stochastic term which takes into account the fluctuations in the shower containment or in the number of primary particles and the number of photons produced by the passage of the charged particles;

Parameter	ECAL	HCAL
a	0.0280	0.8470
b	0.12	0
c	0.003	0.074

Table 2.1: Parameters for the ECAL and the HCAL, reported in GeV.

- b is the noise term to which contribute the electric noise and pile-up energy;
- c is a constant term which considers calibration errors, the non-full containment of the shower and energy leakage.

For CMS, such parameters for both calorimeters are reported in tab. 2.1.

Magnet

The magnet is a superconductive coil surrounding the tracker and calorimeters, which provides the detector with a 3.8 T magnetic field. The purpose of the magnet is to bend the trajectory of charged particles in the detector, so to measure the respective transverse momentum. An iron return yoke is located around the magnet to avoid border effects, and to bend the lines of force in a way to provide a circa constant magnetic field of 1.8 T magnitude also outside of the Solenoid bore.

Muon system

The muon system is an important part of the detector which is used for muon identification, momentum measurement and triggering. Muons are the final products in many processes, and can pass through multiple layers of material without stopping, preventing the measurement of their energy. This causes the muon detector to present a large active surface. In fact, the muon system of CMS covers an active area of about 25000 m^2 . Three types of gas detectors are used for this purpose: drift tubes, cathode strip chambers and resistive plate chambers.

The drift tubes (DTs) are placed in the barrel and cover a pseudorapidity range between 0 and 1.2. They are arranged in four stations, three of these measure the $r - \phi$ coordinate of the muons, while the last one measures z coordinate.

The cathode strip chambers (CSCs) are placed in the endcaps, covering a range $0.9 < |\eta| < 2.4$. They have good segmentation, fast response time and high radiation distance, properties needed in the endcaps, where the rate of muons is higher. The cathode strips of each chamber provide position measurements in the $r - \phi$ plane. The anode wires allow to measure the pseudorapidity and the beam-crossing time of every muon.

The resistive plate chambers (RPCs) are placed both in the barrel and the endcaps and have very fast responses and very good time resolution. This feature makes them ideal to be used for the trigger purposes.

Both the DTs and CSCs have very good efficiency and background rejections in triggering on the p_t of the muons. Nevertheless, the fast response of the RPCs

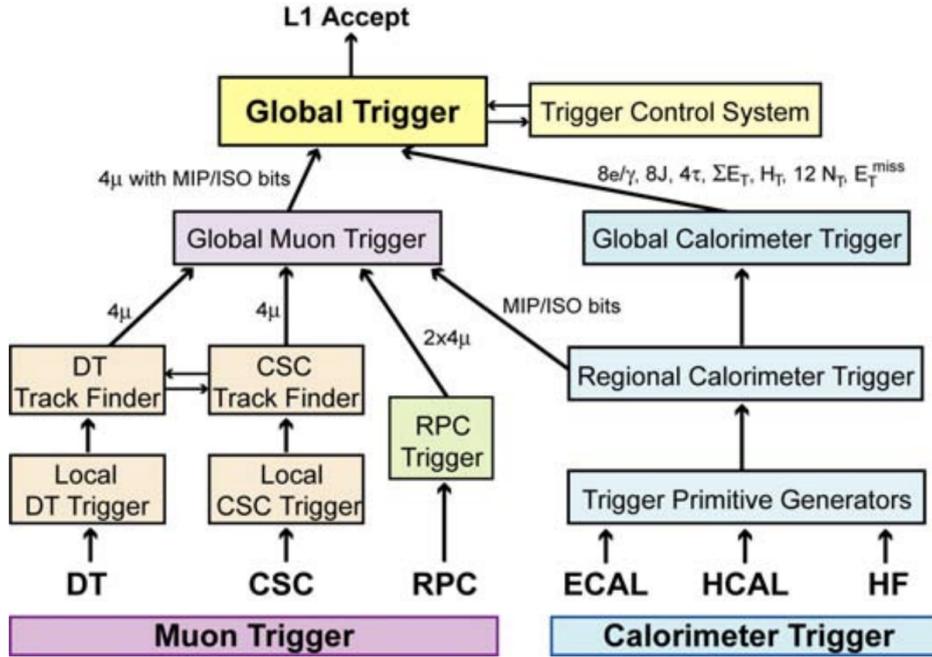


Figure 2.9: Architecture of the L1 trigger.

also allow low p_T tracks that may stop before reaching the outer two stations to be triggered.

Trigger

The LHC provides proton-proton collisions at high rates. It is not possible to store all the information on all the collision events, so a drastic reduction in the rate of stored events must be achieved. This task is performed by the trigger, which is composed of Level-1 trigger (L1 trigger) and High Level Trigger (HLT) and reduces the rate of a factor of at least 10^6 .

The L1 trigger is composed of custom-designed electronics, both FPGAs and ASICs. It has local, regional and global components. The local components are based on energy deposits in the calorimeters and track segments or hit patterns in the muon chambers. Regional triggers combine the information from the local ones before ranking and sorting trigger objects such as electron and muon candidates, before passing this information on to the Global Muon Trigger or the Global Calorimeter Trigger. These latter ones are then connected to the Global Trigger, which takes the decision to reject the event or to pass it to the HLT for further evaluation. The architecture of the L1 trigger is shown in fig. 2.9. The HLT, on the other hand, is a software offline trigger which runs quality reconstruction modules and filters to process and finally select events for storage.

Chapter 3

Physics BSM and W' models

As mentioned in the beginning of chapter 1, the Standard Model predictions have been verified at a very high level of accuracy. Nevertheless, the SM does not provide explanations for the existence of Dark Matter, the origin of non-zero neutrino masses, or the cause of the baryon asymmetry in the universe. Furthermore, the Standard Model holds no explanation for the fine tuning of the mass of the Higgs boson: the tree-level mass of the Higgs boson has higher order corrections which are quadratically divergent, which should lead to a very large measurable mass. This is not the case, thanks to the fact that the corrections of opposite signs are fine-tuned to be cancel each other. Moreover, the SM does not predict a large number of free parameters, such as the coupling constants of the interactions or the masses of the fermions, which also give rise to the hierarchy problem of masses: there is no explanation to the fact that elementary particles have masses as different as the electron's and the top quark's, reported in sec. 1.1. Numerous physics models Beyond Standard Model (BSM) predict the existence of a new spin-1 gauge boson W' , that would have the same quantum numbers as the standard model W boson, but could differ in terms of the couplings depending on the model: it could couple to either left-handed or right-handed currents, or it could have favorite couplings with some generations, as it is for the Topflavor model.

For instance, a W' which couples only to charm and bottom quarks and tau lepton could explain anomalies regarding lepton universality raised by the experiments BaBar, Belle and LHCb [11].

The W' is foreseen by theories such as left-right symmetric models, alternate left-right models and little Higgs models.

3.1 The effective lagrangian

The most general Lagrangian of the lowest possible dimension describing the couplings of W' bosons to fermions is given by:

$$\mathcal{L} = \frac{W'_\mu}{\sqrt{2}} [q'_i (C_{q_{ij}}^R P_R + C_{q_{ij}}^L P_L) \gamma^\mu q_j + \bar{\nu}_i (C_{l_{ij}}^R P_R + C_{l_{ij}}^L P_L) \gamma^\mu l_j], \quad (3.1)$$

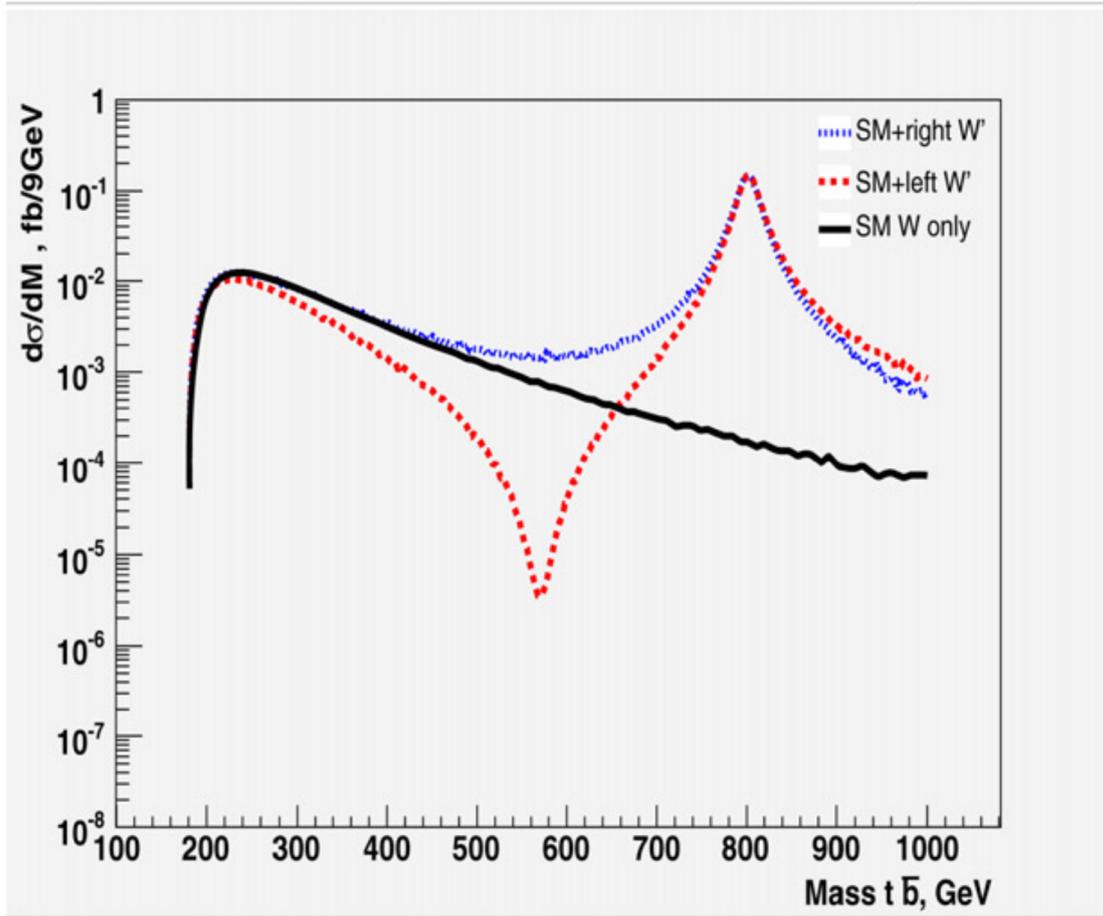


Figure 3.1: Differential cross section for the process $p p \rightarrow W/W' \rightarrow t \bar{b}$ with respect to the invariant mass of the couple $t \bar{b}$ for a simulated W' of mass 800 GeV, for the cases of: SM only, SM + right W' , and SM + left W' [12].

where q, q', l, ν are SM fermions in mass eigenstates, while i and j are the generation indexes, running from 1 to 3. For the SM W boson, the coefficients are $C_{q_{ij}}^R = C_{l_{ij}}^R = 0$, $C_{q_{ij}}^L = g_w V_{CKM}$, $C_{l_{ij}}^L = g_w$. The simplest extension of the SM which predicts the existence of a W' boson is $\mathcal{SU}(2)_1 \times \mathcal{SU}(2)_2 \times \mathcal{U}(1)$, which also includes a mixing between the SM W and the W' , if the W' couples with left-handed currents. An example of this is shown in fig. 3.1, which displays the trend of the differential cross section for the process $p p \rightarrow W/W' \rightarrow t \bar{b}$ with respect to the invariant mass of the couple $t \bar{b}$, for a simulated W' of mass 800 GeV. One can notice that the case which considers the SM W and a W' with left-handed couplings presents a local minimum due to the interference between the two bosons [12]. In the next section, we will discuss the W' arising from the Topflavor model, a theory which conjectures nonflavor universality and could give a solution to the hierarchy problem.

3.2 Topflavor model

The Topflavor model is an extension of the electroweak theory based on a symmetry group $\mathcal{SU}(2)_1 \times \mathcal{SU}(2)_2 \times \mathcal{U}(1)_Y$ where the first two generations of fermions couple to $\mathcal{SU}(2)_1$ and the third generation couples to $\mathcal{SU}(2)_2$. In this model, fermions of the first two generations have the following representations under $\mathcal{SU}(2)_1, \mathcal{SU}(2)_2, \mathcal{U}(1)_Y$:

$$(U, D)_L \rightarrow (2, 1, \frac{1}{3}), \quad U_R \rightarrow (1, 1, \frac{4}{3}), \quad D_R \rightarrow (1, 1, -\frac{2}{3}) \quad (3.2)$$

$$(\nu_l, l)_L \rightarrow (2, 1, 1), \quad l_R \rightarrow (1, 1, -2), \quad (3.3)$$

while the fermions in the third generations are:

$$(t, b)_L \rightarrow (1, 2, \frac{1}{3}), \quad t_R \rightarrow (1, 1, \frac{4}{3}), \quad b_R \rightarrow (1, 1, -\frac{2}{3}) \quad (3.4)$$

$$(\nu_\tau, \tau)_L \rightarrow (1, 2, 1), \quad \tau_R \rightarrow (1, 1, -2). \quad (3.5)$$

We can hence consider the covariant derivative:

$$D_\mu = \partial_\mu - i\frac{g'}{2}YB_\mu - ig_1\vec{T} \cdot \vec{W}_\mu - ig_2\vec{T} \cdot \vec{W}_\mu, \quad (3.6)$$

where g_1 and g_2 , \vec{W}_μ and \vec{W}_μ are respectively the coupling constants and the interaction fields for the groups $\mathcal{SU}(2)_1$ and $\mathcal{SU}(2)_2$.

The symmetry for this group is broken in two steps: in the first one, we get the Standard Model $\mathcal{SU}(2)_L \times \mathcal{U}(1)_Y$ and the following step is the one already discussed for the SM in sec. 1.4. The first stage is achieved with the introduction of a Higgs field Φ which transforms as a doublet under both $\mathcal{SU}(2)$ groups, with vacuum expectation value (vev):

$$\Phi = \frac{1}{\sqrt{2}} \begin{pmatrix} u & 0 \\ 0 & u \end{pmatrix}. \quad (3.7)$$

The second stage is then obtained introducing a doublet Higgs field H with vev v . One obtains the following mass matrix for the neutral sector of the theory:

$$\frac{1}{2} \begin{pmatrix} g_1^2 u^2 & -g_1 g_2 u^2 & 0 \\ -g_1 g_2 u^2 & g_2^2 (v^2 + u^2) & -g' g_2 v^2 \\ 0 & -g' g_2 v^2 & g'^2 v^2 \end{pmatrix}, \quad (3.8)$$

with the basis ordered as W, \tilde{W}, B . One can diagonalize the matrix by means of an appropriate orthogonal matrix R :

$$\begin{pmatrix} A \\ Z_l \\ Z_h \end{pmatrix} = R \begin{pmatrix} W_3 \\ \tilde{W}_3 \\ B \end{pmatrix} \quad (3.9)$$

with the mass eigenstates being A, Z_l, Z_h . The coupling for this theory result to be:

$$g_1 = \frac{e}{\cos\phi \sin\theta_w}, \quad g_2 = \frac{e}{\sin\phi \sin\theta_w}, \quad g' = \frac{e}{\cos\theta_w}, \quad (3.10)$$

where θ_w is the weak mixing angle and ϕ is an additional mixing angle.

The eigenstate A has zero mass and can be identified as the photon, while the masses for Z_h and Z_l can be found solving the following equation:

$$M_Z^4 - \frac{1}{2}u^2(g_1^2 + g_2^2 + g'^2\varepsilon + g_2^2\varepsilon)M_Z^2 + \frac{1}{4}u^4\varepsilon(g_1^2g'^2 + g_1^2g_2^2 + g_2^2g'^2) = 0, \quad (3.11)$$

where $\varepsilon = v^2/u^2$. Z_l is the eigenstate with lower mass and is identified as the SM Z . We can repeat this procedure for the charged sector of theory, which has the following mass matrix:

$$\frac{1}{2} \begin{pmatrix} g_1^2u^2 & -g_1g_2u^2 \\ -g_1g_2u^2 & g_2^2(v^2 + u^2) \end{pmatrix}, \quad (3.12)$$

with the basis ordered as W, \tilde{W} . We can then get the mass eigenstates using an appropriate orthogonal matrix R' :

$$\begin{pmatrix} W_l \\ W_h \end{pmatrix} = R' \begin{pmatrix} W \\ \tilde{W} \end{pmatrix}. \quad (3.13)$$

The masses for the bosons can be obtained by means of solving the following equation:

$$M_W^4 - \frac{1}{2}u^2[g_1^2 + g_2^2(1 + \varepsilon)]M_W^2 + \frac{1}{4}u^4g_1^2g_2^2\varepsilon. \quad (3.14)$$

W_l is identified with the SM W and W_h corresponds to an example W' .

3.3 W' decay channels

In this section, we will discuss the W' decay modes. The decay width, considering only cases with decays to fermions, is divided in three partial widths, depending on the final products:

$$\Gamma_{tot}(W') = \Gamma(W' \rightarrow t\bar{q}') + \Gamma(W' \rightarrow q\bar{q}') + \Gamma(W' \rightarrow l\bar{n}u). \quad (3.15)$$

We separate the partial widths containing a top quark because it is decay channel which we analyze in the remainder of this thesis' work. The leading order partial widths are:

$$\Gamma_{LO}(W' \rightarrow t\bar{q}') = \frac{g^2\beta^2}{16\pi m_{W'}} |V'_{tq'}|^2 (m_{W'}^2 + m_t^2/2), \quad (3.16)$$

$$\Gamma_{LO}(W' \rightarrow q\bar{q}') = \frac{g^2}{16\pi} |V'_{qq'}|^2 m_{W'}, \quad (3.17)$$

$$\Gamma_{LO}(W' \rightarrow l\bar{n}u) = \frac{g^2}{16\pi} |V'_{l\nu}|^2 \frac{m_{W'}}{3}, \quad (3.18)$$

where $\beta = 1 - \frac{m_t^2}{m_{W'}^2}$ and g is assumed to be the SM coupling. In this way, the W' boson partial widths have the same form of the SM W boson's and the new

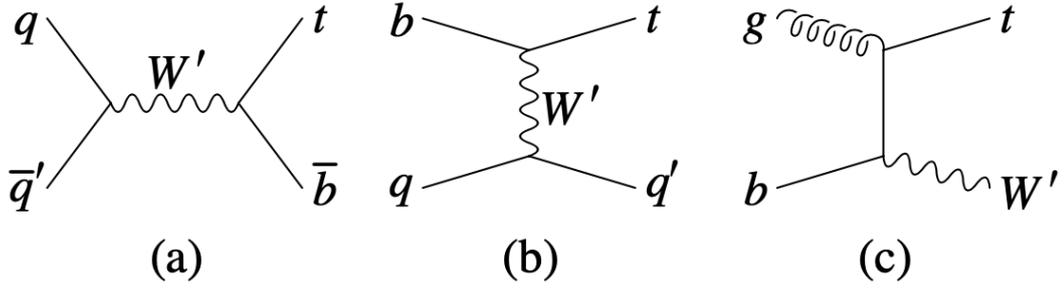


Figure 3.2: Feynman diagrams for single top production involving a W' boson: a) s-channel production, b) t-channel production, c) W' -t associated production

couplings and the CKM elements are absorbed in the $V'_{f_i f_j}$ matrix elements. The 3.16 shows that W' bosons tend to have a large branching ratio into top quarks, which in turn has an effect on the single-top-quark production cross section at hadron colliders, such as LHC. This happens through the three channels show in fig. 3.2. The cross sections for the t-channel production and the associated W' -t production are negligible at the LHC scale of energy [13].

A W' boson of suitable mass is also able to decay into two SM Ws or a Z boson and a Higgs boson. Fig. 3.3 shows the upper limits at 95% CL in several decay channels of the two bosons in which the W' decays [14].

3.4 Search for the W' boson

Searches for W' bosons have taken place both at Fermilab's Tevatron and at LHC. Among the searches at CERN, conducted by the ATLAS and CMS experiments, we consider W' bosons decaying to a top quark and a bottom quark. Assuming that the matrix $V_{q_i q_j}$, described in sec. 3.3, is the SM V_{CKM} , the W' boson is expected to decay to a top quark and a bottom quark. The experimental searches look for excesses of events in the invariant mass spectrum of top and bottom quarks (m_{tb}). The various searches can differ for the considered decay products of the top quark, which can decay either leptonically (fig. 3.4):

$$t \rightarrow b W \rightarrow b l^+ \nu_l, \quad (3.19)$$

or hadronically:

$$t \rightarrow b W \rightarrow b q q'. \quad (3.20)$$

The Branching Ratios are:

$$t \rightarrow b e^+ \nu_e, \quad BR = (11.10 \pm 0.30)\%, \quad (3.21)$$

$$t \rightarrow b \mu^+ \nu_\mu, \quad BR = (11.40 \pm 0.20)\%, \quad (3.22)$$

$$t \rightarrow b \tau^+ \nu_\tau, \quad BR = (11.1 \pm 0.9)\%, \quad (3.23)$$

$$t \rightarrow b q q', \quad BR = s(66.5 \pm 1.4)\%. \quad (3.24)$$

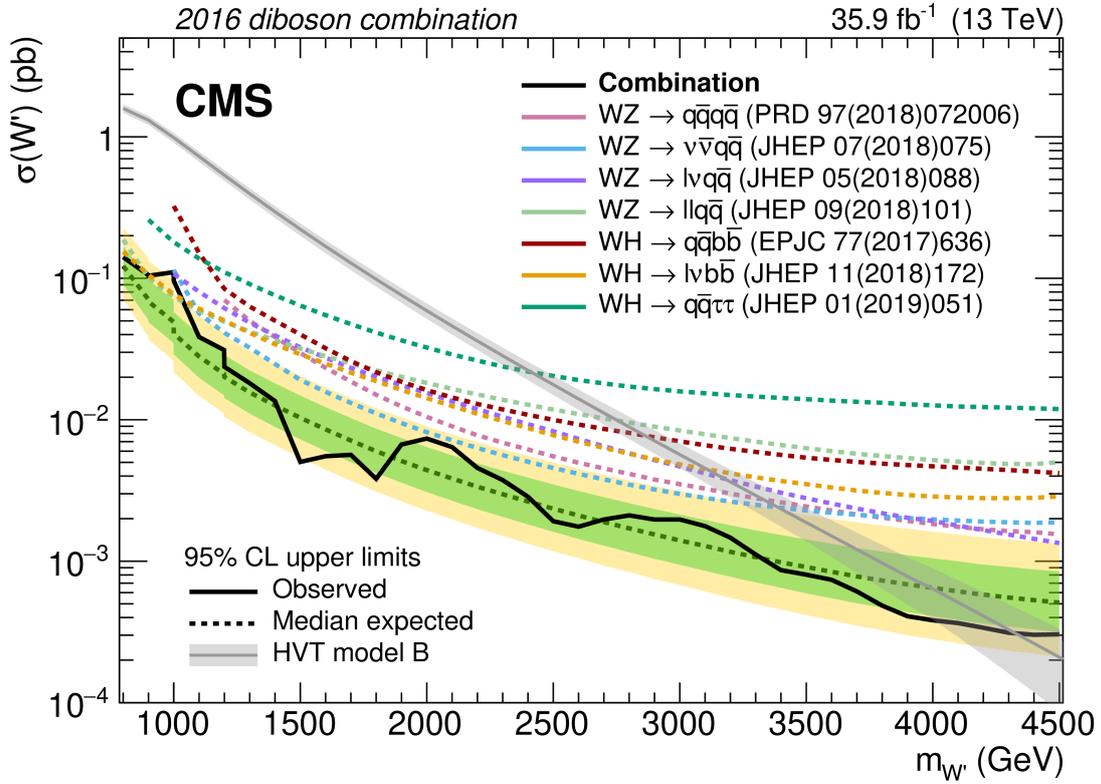


Figure 3.3: Upper limits at 95% CL for the W' boson production cross section at CMS in the diboson channels [14].

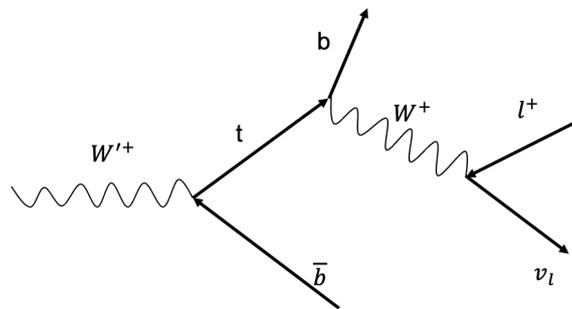


Figure 3.4: $W' \rightarrow t b$ decay, with t decaying leptonically.

A CMS search published in 2017 [15] considered the leptonic decay of the top quark in the e-channel and the μ -channel, using data collected at $\sqrt{s} = 13 \text{ TeV}$, with an integrated luminosity of 35.9 fb^{-1} . The results of the selections are displayed in fig. 3.5, which shows that the expected signal events are comparable with the background predictions, and therefore the search for higher mass points would be limited by the presence of background in the signal region that is not rejected by standard cuts. It excluded the production of a right-handed W' bosons at 95% confidence level for masses up to 3.6 TeV , as shown in fig. 3.6, in which two different theoretical production cross sections are considered as function of the potential sterile neutrino mass m_{ν_R} . If there is a right-handed neutrino ν_R with $m_{W'} > m_{\nu_R}$ the BR for $W' \rightarrow t b$ would decrease in order to account for the ulterior channel $W' \rightarrow \nu_R l$. During this thesis work, we developed machine-learning-based selections to try and improve the signal efficiency and background rejection for this analysis, retaining a good rejection of the background.

A similar search performed by the ATLAS experiment, with an integrated luminosity of 36.1 fb^{-1} , obtained analogous results, excluding at 95% confidence level the existence of a right-handed W' boson for masses up to 3.15 TeV , as shown in fig. 3.7 [16]. Another search performed at CMS considered all-hadronic final states and exploited a Deep Neural Network to recognize the hadronic jets which originated from a top quark. It used data equivalent to an integrated luminosity of 137 fb^{-1} collected at $\sqrt{s} = 13 \text{ TeV}$. This analysis excluded both right-handed and left-handed W' with mass below 3.4 TeV at 95% confidence level, as shown in fig. 3.8 [17].

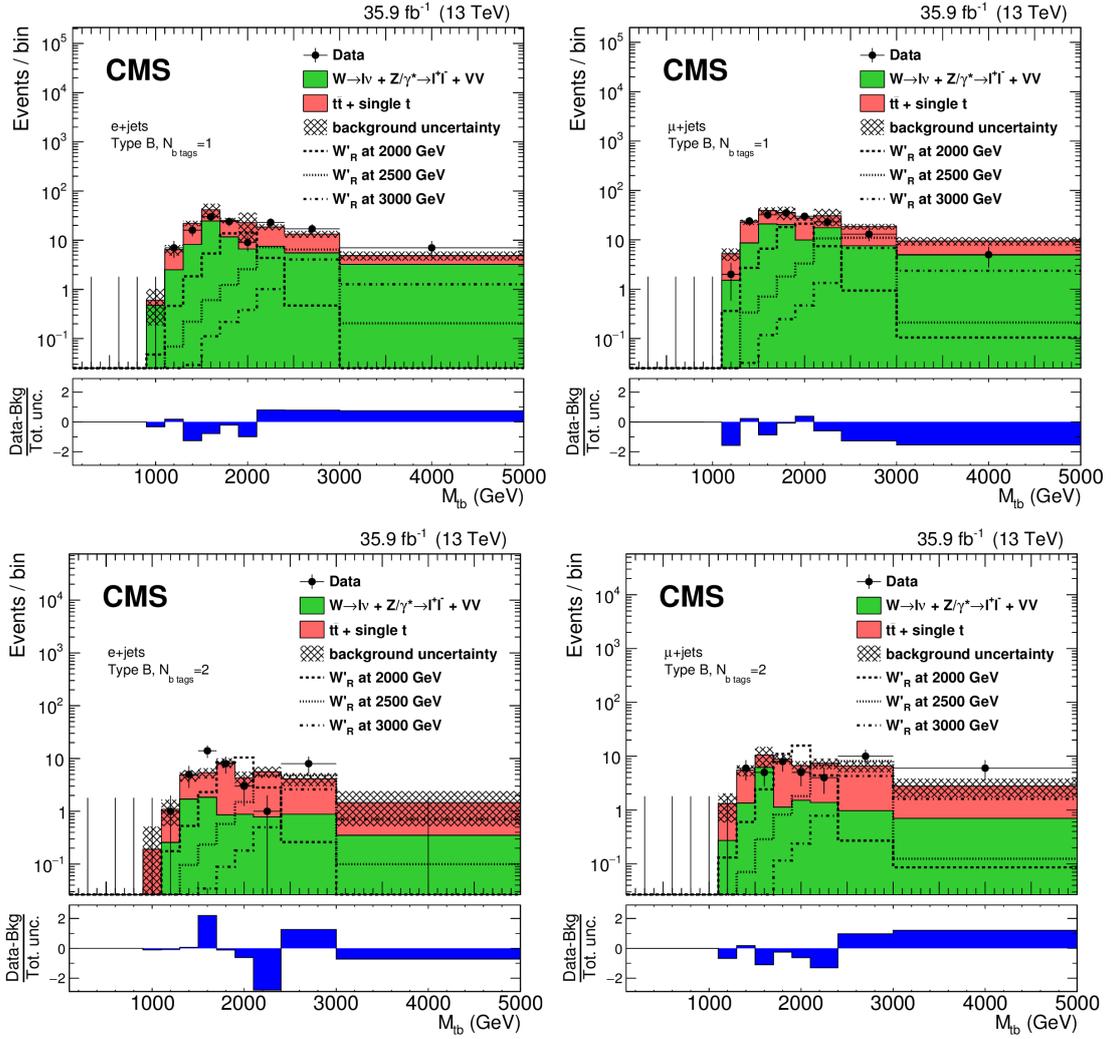


Figure 3.5: Reconstructed invariant mass of the couple b-jet and top quark with 1 b-tagged jet (up) or 2 (down) in the e-channel (left) and μ -channel (right) after selections. Distributions for W'_R bosons with masses of 2, 2.5, and 3 TeV are shown.

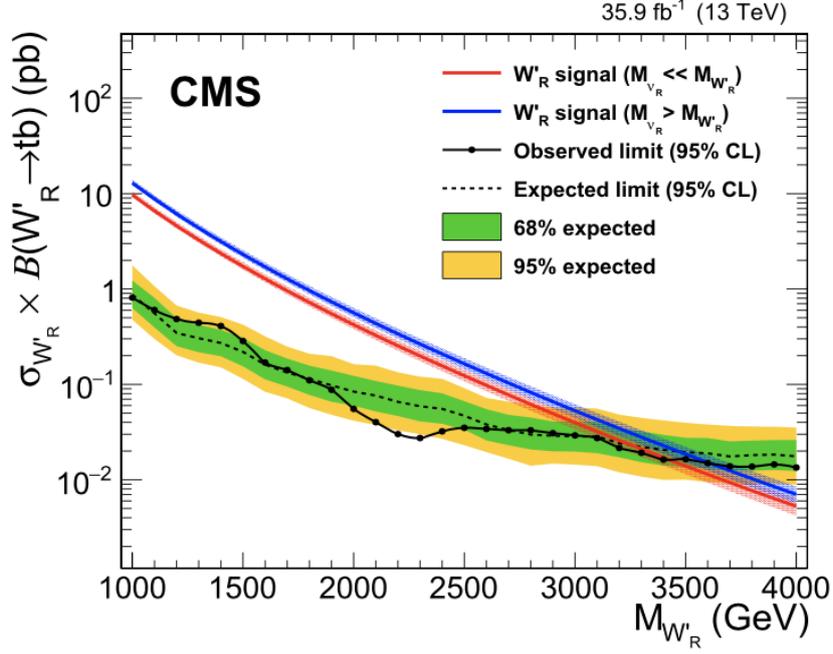


Figure 3.6: Upper limit at 95% CL on the W'_R boson production cross section for the combined electron and muon channels. Values for which the theoretical cross section observe exceeds the upper limit are excluded.

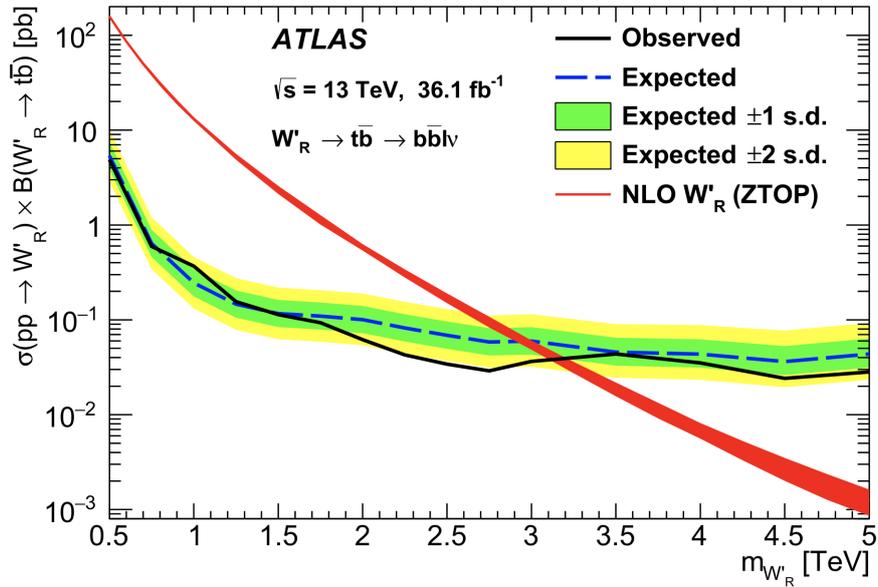


Figure 3.7: Upper limit at 95% CL on the W'_R boson production cross section for the combined electron and muon channels.

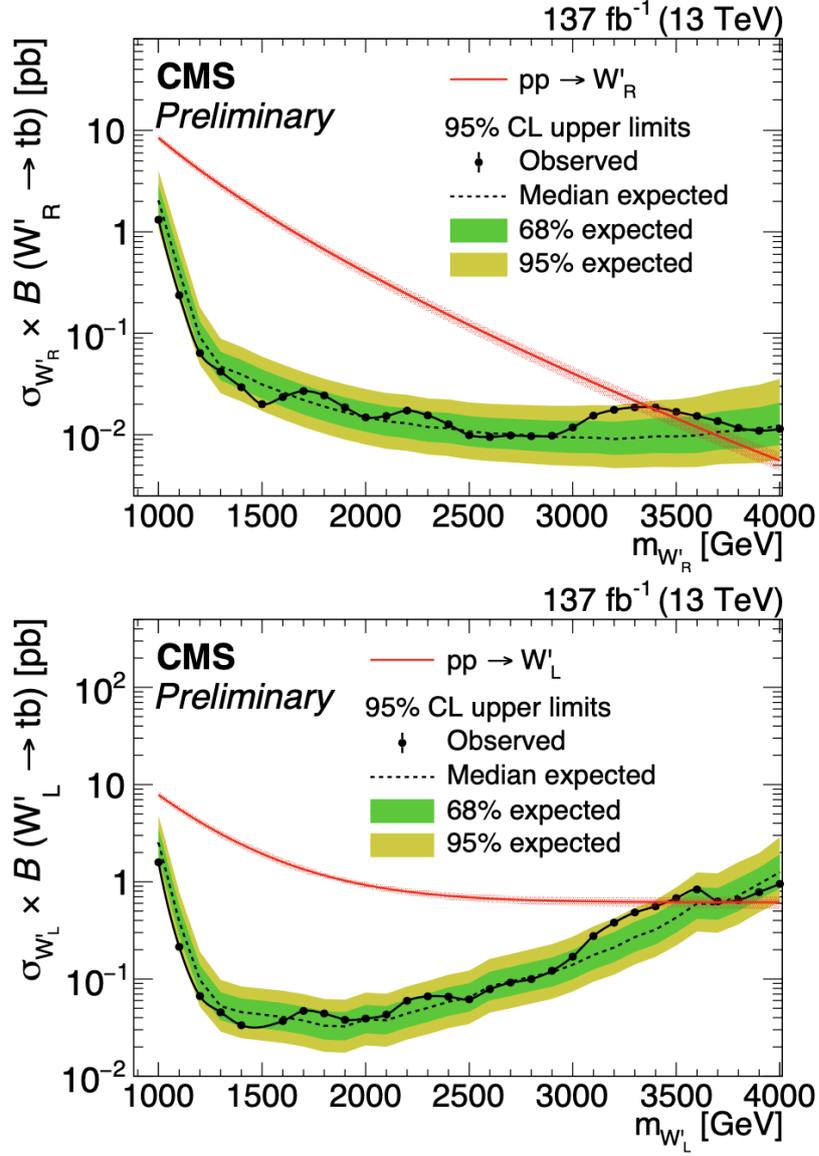


Figure 3.8: Upper limits at 95% CL for the production of W'_R boson (top) and W'_L boson (bottom). The two cross sections differ because a left-handed W' boson would undergo interference with the SM W boson. [12]

Chapter 4

Physics object selection and reconstruction

The goal of this thesis work is the search for a W' boson analyzing the channel $W' \rightarrow t b$, in which the top quark decays to a b quark, a muon, and a neutrino. A Feynman diagram which shows the process is displayed in fig. 3.4. This channel has been chosen in virtue of its importance in several models as described in Chapter 3, for which the decay width $\Gamma(W' \rightarrow t b)$ (sec. 3.3) is large. Moreover, the leptonic decay of the top quark, despite its lower branching fraction with respect to the hadronic one, also has low contamination from QCD multijet background. Because of the large mass of the W' boson, the top quark and its decay products are expected to have a large Lorentz boost in the reference frame of the laboratory, so that the final products tend to be collimated in a single jet aligned with the top quark momentum.

4.1 Physics objects selection

To reconstruct the W' boson in the channel of our interest, we look for b jets, muons and missing transverse energy in the final state of the collisions. The

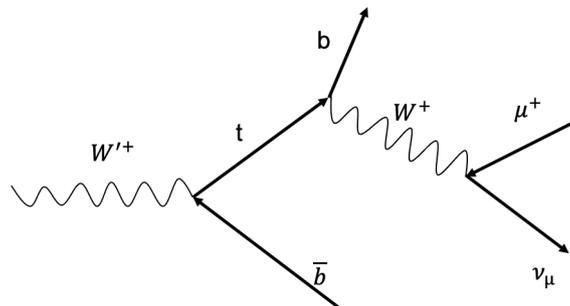


Figure 4.1: An example of Feynman diagram for the $W' \rightarrow t b$ decay with the top decaying to a muon.

final products of a collision event in CMS pass through the various subdetectors of the experiment and are identified by the Particle Flow (PF) algorithm [18]. Such algorithm combines the basic elements from all detector layers (tracks and clusters) to identify each final-state particle. For each collision event, it follows a precise order in the reconstruction of physics objects:

- first, **muons** are identified by considering tracks in the tracker and in the muon detectors and combines them with energy deposits in the calorimeter. The tracks and clusters are then taken out of further processing. More details on muon reconstruction are given further in this thesis;
- **electrons** and **isolated photons** are reconstructed at the same time, considering energy deposits in the ECAL and tracks in the tracker system: a cluster in the ECAL linked to a charged particle track in the tracker is identified as an electron; a cluster unmatched with any tracks is identified as an isolated photon. Tracks and clusters used in this step are then masked against further processing;
- the last objects to be reconstructed are the **hadrons** and the **non-isolated photons**. For this last step, information from both the ECAL and HCAL are used: clusters unmatched to any tracks are considered neutral hadrons or non-isolated photons, otherwise the candidates are considered charged hadrons;
- Lastly, the previous information is used to reconstruct high-level objects such as the **missing transverse energy** and the **hadronic jets**.

4.2 Muons

Muons in CMS are detected in all the subdetectors, and the leading role in muon identification and reconstruction is played by the inner tracker and muon system. Those two are used, alongside PF reconstruction, to identify different types of muon depending on what part of CMS is involved. The hits in the tracker are fitted with a Kalman filter to obtain a track, while the fit of the hits in the muon system returns a Stand Alone Muon. Muons are considered Tracker Muons if they are reconstructed in the tracker and matched with at least an hit in the muon system. A Global Muon is given by fitting together of a Tracker Muon with a Stand Alone Muon. In this thesis work, we considered only muons with $p_t > 10$ to reconstruct the top quark.

4.3 Jets

Quarks and gluons which arise from the collisions at LHC cannot be observed as free particles, because of color confinement, and as they propagate from the interaction point to the detector, they undergo hadronization, creating hadronic jets, which are clusters of color-neutral hadrons. The analysis of the jets can return

information regarding the parton which originated it. Jets are reconstructed using the anti- k_t clustering algorithm[19] with radius parameter 0.4. For this reason the reconstructed jets are called AK4 jets. This algorithm provides an infrared-safe and collinear-safe clustering for the jets. It introduces a distance d_{ij} between entities (PF candidates, pseudojets) i and j , and a distance d_{iB} between the entity i and the Beam (B). The distances are defined as:

$$d_{ij} = \min(k_{ti}^{-2}, k_{tj}^{-2}) \frac{\Delta_{ij}^2}{R^2}, \quad (4.1)$$

$$d_{iB} = k_{ti}^{-2}, \quad (4.2)$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and k_{ti} , y_i and ϕ_i are respectively the transverse momentum, the rapidity and the azimuth of the entity i . R is a radius parameter which modules the size of the jet. For our analysis it has been used $R = 0.4$. The clustering proceeds by finding the smallest distance for each identity i , against all others: if the minimum is d_{ij} for some j , entities i and j are recombined; if it is d_{iB} , then entity i is called a jet and is removed from the list of the entities. The distances are then recalculated and the procedure repeated until no entities are left. The entities considered for jet clustering are all the Particle Flow candidates, including muons and electrons, that thus can be reconstructed inside a jet.

4.3.1 b-tagging

In order to reconstruct the W boson, we need to identify the hadronic jets which originated from a b -quark, called *b-jets*. This operation is called b -tagging, which we performed exploiting the DeepFlavour discriminant [20], provided by CMS. It is an algorithm based on a deep neural network, which assigns to a jet its probability of being a b -jet. The DeepFlavour employs convolutional neural layers and features regarding charged and neutral particles in the jet as well as secondary vertices. A secondary vertex is the decay vertex of particles arising from the p - p collision. They grant insight of the origin of the hadronic jets as we can measure the distance travelled by the hadron generated in the collision before decaying and, knowing that it would move at about the speed of light, calculate its the decay time, so to identify it as a known hadron.

4.4 Missing transverse energy

At LHC, the z -component of the momentum of the two particle beams is known as well as the transverse component, which is circa zero. Momentum conservation in the trasverse plane allows us to measure the overall transverse momentum of the particles which escape detection, such as neutrinos. The energy corresponding to this momentum is called Missing Transverse Energy (MET) \vec{E}_t . The same procedure cannot be applied to the longitudinal component of the momentum, as it is not possible to measure the momentum of the proton fragments which escape through the beam pipe after the collision. The base definition of MET (called

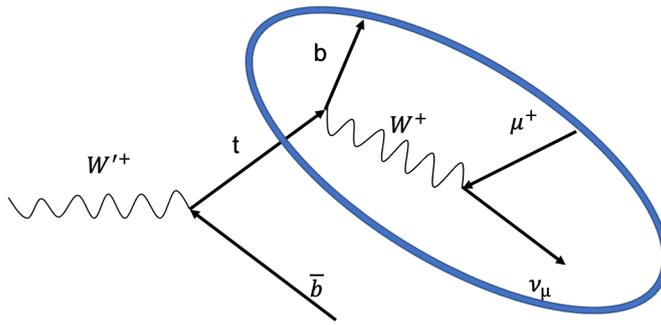


Figure 4.2

also *raw MET* is:

$$\vec{p}_t = - \sum_i \vec{p}_{ti}, \quad (4.3)$$

where the sum is on all the PF candidates of the events. The PF MET actually considered for the analyses in CMS is further corrected, resulting in a reweighted MET which takes into account energy corrections to the jets, the non-compensating nature of calorimeters and potential misalignments in the detectors. For our analysis, we considered the MET in order to have information on the muonic neutrino in which the top quark decays.

4.5 Top quark reconstruction

The top quark is reconstructed with a muon, a b-jet and a neutrino, as shown in fig. 4.2. We need to add together the 4-momenta of these objects in order to obtain the 4-momentum of the top quark. In fact, this is not a straightforward operation, as we can consider the MET as the transverse momentum of the ν , but have no means to measure its z-component of momentum. Nevertheless, we can calculate it by imposing $\sqrt{s}(\mu, \nu) = m_W$, with m_W being the mass of the W boson, as both the leptons come from a real W boson. We can approximate the decay width of the W to be negligible with respect to the experimental resolutions involved with the top quark reconstruction, and hence obtain the following equation:

$$p_{\nu,z} = \frac{\Lambda p_{\mu,z}}{p_{\mu,t}^2} \pm \frac{1}{|\vec{p}_{\mu,t}|} \sqrt{2\Lambda^2 - E_\mu^2 \vec{E}_t^2}, \quad (4.4)$$

$$\Lambda = \frac{m_W}{2} + \vec{p}_{\mu,t} \cdot \vec{\not{p}}_t, \quad (4.5)$$

where \vec{p}_μ and E_μ are respectively the momentum and the energy of the muon, $\vec{\not{p}}_t$ and \vec{E}_t are the missing transverse momentum and the missing transverse energy. The argument of the root in the 4.4 is usually positive and the solution chosen is the one with the smaller absolute value. If it is not positive, then the imaginary component is eliminated by imposing that the square root is null, obtaining a

quadratic relation between $p_{\nu,x}$ and $p_{\nu,y}$, with two possible solutions and one remaining degree of freedom. The solution with the minimum vectorial distance between $\vec{p}_{\nu,t}$ and \vec{p}_t is chosen.

Having calculated the momentum of the neutrino, we can reconstruct a top quark candidate as the sum of a jet, a muon, and a neutrino. In order to recognize the jets which come from the decay of a top quark, we developed two machine learning algorithms: a boosted decision tree (BDT) and a deep neural network (DNN).

4.6 Boosted Decision Trees

A boosted decision tree is a machine learning classification algorithm based on the concept of decision tree. A decision tree is a sequence of selection requirements (cuts) that are applied in a specified order on a given variable datasets [21]. The cuts split the dataset into nodes, which correspond to a certain number of samples classified as signal or background. Each node can then be split by other cuts. When there are too few samples of a category in a node, it can no longer be split up and it is called a leaf. An example of decision tree is shown in fig. 4.3. Before a decision tree is able to make predictions on data, it must be trained (in jargon, "grown") on a training dataset of already known classifications. One way to grow a decision tree is maximize the gain of Gini index at each node after splitting. The Gini index is defined as:

$$G = P(1 - P), \quad (4.6)$$

where P is the purity of the node, i.e., the fraction of signal samples in the node. The Gini index is 0 for nodes which contain only signal samples or background samples. The gain in Gini index for a node A split into two nodes B_1 and B_2 is defined as:

$$\Delta G = G(A) - G(B_1) - G(B_2). \quad (4.7)$$

Decision trees are useful instruments, as they can be easily interpreted, but are not reliable and have problems with overtraining, i.e., when a machine learning algorithm learns the fluctuations in the training dataset, instead of the right patterns in the features, and so it does not generalize well to new data. To improve the robustness of a decision tree, they are often used in ensembles, which can be created via various procedures, such as boosting. A boosted decision tree algorithm is created as follows:

- Training observations are reweighted using the previous iteration's classifier result;
- A new tree is grown and optimized using the reweighted observations as a training sample;
- Each tree is given a score;

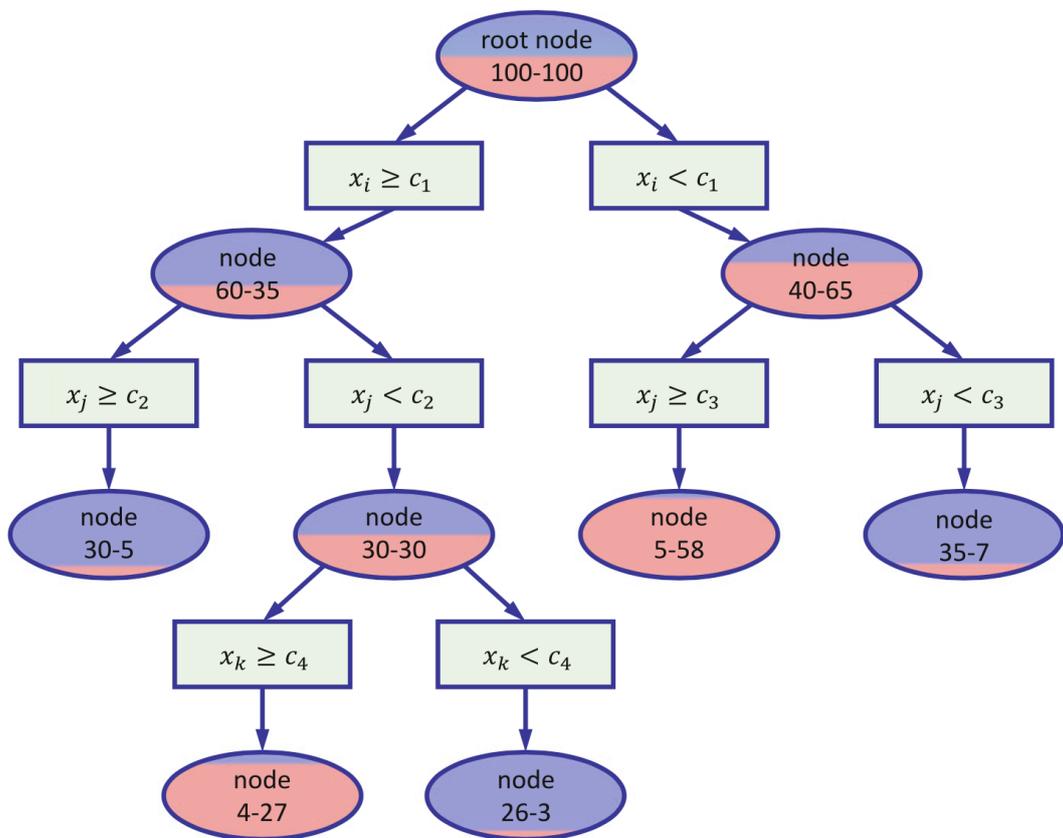


Figure 4.3: Schematic description of a decision tree [21].

- The final output of the BDT classifier is the weighted average of every tree in the forest:

$$y = \sum_{i=1}^{N_{trees}} w_k y_k, \quad (4.8)$$

where w_k and y_k are respectively the score and the prediction of the k -th tree.

4.7 Top tagging with BDTs

During this thesis work, six BDTs have been developed to work on the top-tagging task, using the XGBoost module in Python. The dataset used for this classifiers is based on triplets of hadronic jets, muons and MET in simulated collision events at LHC. The variables analyzed are summarized in tab. 4.1, where we labelled "Top candidates" the sum of the muons and the jet, while the Top_ν candidates also consider the neutrino. The label *ub* (unboosted) refers to the variables when calculated in the top-quark candidate center-of-mass frame for every candidate. Some other variables which is worth to highlight are:

- **miniiso**: It is an powerful variable which helps discriminating whether or not a muon close to an hadronic jet is a byproduct of hadronization. It is calculated as:

$$\sum \frac{p_{t,\gamma+had}}{p_{t,\mu}}, \quad (4.9)$$

in a cone centered around the direction of the momentum of the muon. The cone has a variable radius R in the metric (ϕ, η) :

$$R \propto p_{t,\mu}^{-1}, \quad (4.10)$$

- $\theta_{\mu,b}$ is the angle between the muon momentum direction and the jet momentum direction;
- $p_{t,rel}$ is defined as:

$$p_{t,rel} = \frac{|\vec{p}_\mu \times \vec{p}_{jet}|}{|\vec{p}_\mu|} \quad (4.11)$$

and also helps to discriminate between prompt muons and muons which come from hadronization.

- the **DeepCSV score** is the probability, given by a CMS-developed deep neural network algorithm, that the jet is a b-jet;
- d_{xy} and d_z are the impact parameters respectively in the transverse plane and in the longitudinal direction.

Furthermore, we divided the top candidates in two categories, based on the angular distance ΔR between the muon and the jet:

- **merged**: if $\Delta R < 0.4$;
- **resolved**: if $0.4 < \Delta R < 2$.

Jet	Muon	Jet _{ub}	Muon _{ub}	Top	Top _ν
p_t	# tracker hits	p_t	p_t	mass	p_t
ϕ	is Global	ϕ	ϕ	$\cos\theta_{\mu,b}$	ϕ
η	miniiso	η	η	$p_{t,rel}$	η
DeepCSV score	dxy			mass	
	dz				

Table 4.1: Variables employed by the BDT top tagger.

4.8 Top tagging performances with BDTs

In this section, we discuss the top-tagging performances of the BDT we developed. We trained and tested the performance of the algorithm on data obtained from Monte Carlo simulations of W' and Z' production at LHC. The samples were created with different mass hypotheses of 4, 5 and 6 TeV. We prepared the dataset by picking all the triplets of jet, muon and MET which respect the criteria to reconstruct either a merged or resolved top candidate. We also required the muon's miniiso to be less than 0.5. Finally, the dataset has been split between merged and resolved triplets and in three bins of p_t of the top candidate (here considered as jet + muon):

- **low p_t :** $p_t < 1000$ GeV;
- **medium p_t :** $p_t > 1000$ and $p_t < 2000$ GeV;
- **high p_t :** $p_t > 2000$ GeV.

Each of the six datasets has then been split in a training set and a validation set; the latter one is first of all used to check if the algorithm has the same performance of the training even on new data. The performance on training set and validation set being different is a proof of overtraining. We then trained six BDTs with identical hyperparameters and input variables: one for each dataset. This was done to try and minimize the correlation between the output of the machine learning algorithm and the p_t of the couple of jet and muon. For each one of the six BDTs, we considered the distribution of the output of test and training sets. From the distribution of the output, we can obtain the Receiver Operating Characteristic (ROC) curve, which gives us information on the trend of signal efficiency ϵ_s and the probability of misidentification ϵ_b with respect to the threshold for selection, defined as:

$$\epsilon_s = \frac{TP}{P}, \quad (4.12)$$

$$\epsilon_b = \frac{FP}{N}, \quad (4.13)$$

where P , N are respectively the total number of signal samples and background samples in the training set, while TP and FP are the number of signal samples and background samples selected by the algorithm, also on the training set.

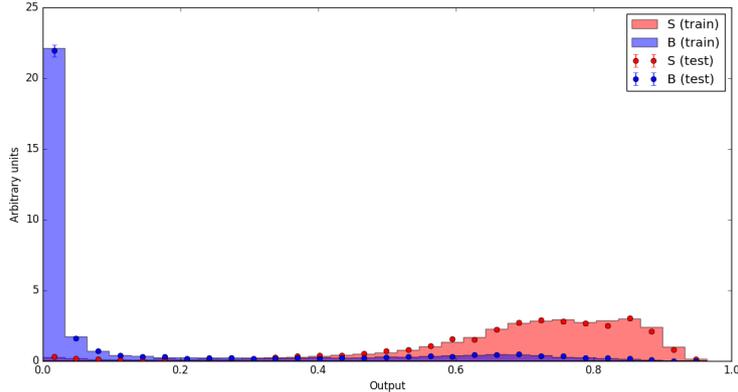


Figure 4.4: Output of the low-pt BDT for the merged signal.

In the following we will describe the results of the BDTs in terms of the distribution of the output of the BDTs and the related ROC curves. Fig. 4.4 shows the output of the merged signal for the the low-pt configuration. No signs of overtraining are displayed, but the background still presents a non-negligible tail, and in some cases even an accumulation of events at the high output of the BDT. The relative ROC curve is displayed in fig. 4.5, with an area under the curve of 0.940. The output and ROC curve for the resolved signal in the low-pt configuration are shown respectively in figg. 4.6, 4.7, which again show a non-ideal discrimination between signal and background. The output and ROC curve of the medium-pt BDT for the merged signal are shown in figg. 4.8, 4.9, which show again an overlap between the output for the signal and for the background. The output and ROC curve of the medium-pt BDT for the resolved signal are shown in figg. 4.10, 4.11 and display signs of overtraining as the output distributions for the training set and the validation set are different. This can be do the fact that there are few events in this sample, so the algorithm cannot learn the patterns in the features and merely *memorize* the training set, instead. Finally, the output and ROC curve of the high-pt BDT for the merged signal are shown in figg. 4.12, 4.13 and they show a very good discrimination power. The performances of the high-pt configuration for the resolved signal are not reported as there are very few signal events in this sample. This is due to the fact that the large p_t of the jet is related to a larger p_t of the top quark, so that its decay products are very collimated and the angular distance between muon and jet is virtually never greater than 0.4.

We can define three types of background samples:

- **B0:** if neither the muon nor the jet come from a top quark;
- **B1:** if only the jet comes from a top quark;
- **B2:** if only the muon comes from a top quark.

Fig. 4.14 shows the distribution of the different types of background for the low-pt resolved BDT, in which the adjacent bins of different colors refer to the same output of the BDT. They have not been superimposed for graphical clearness. The

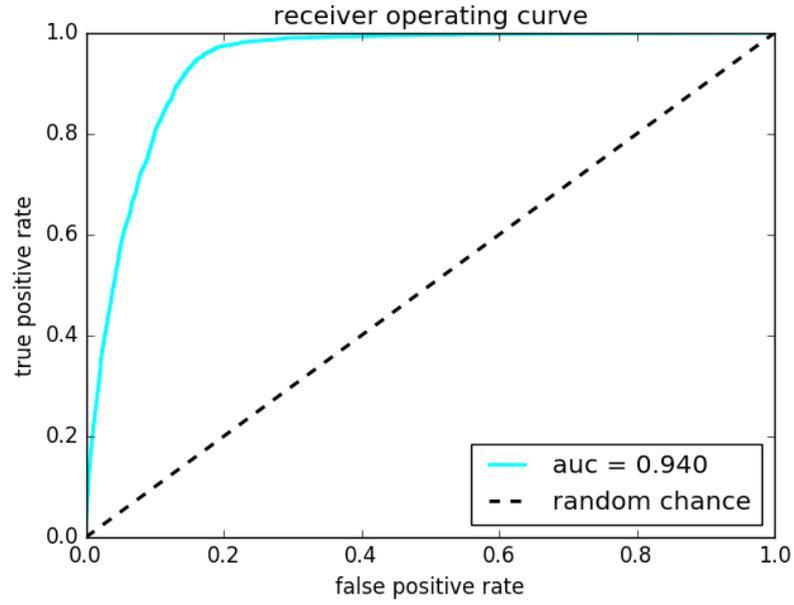


Figure 4.5: ROC curve of the low-pt BDT for the merged signal.

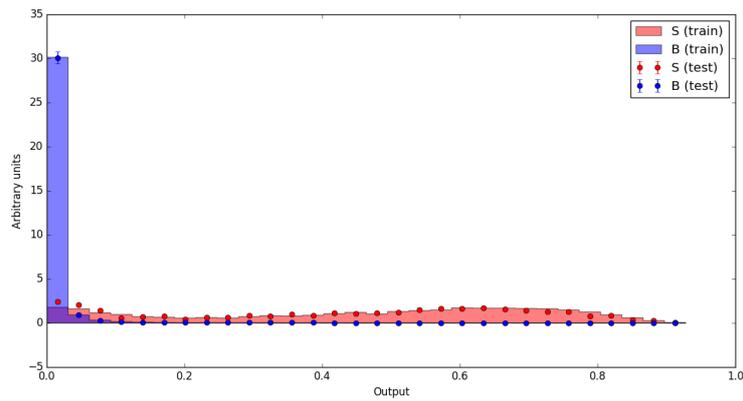


Figure 4.6: Output of the low-pt BDT for the resolved signal.

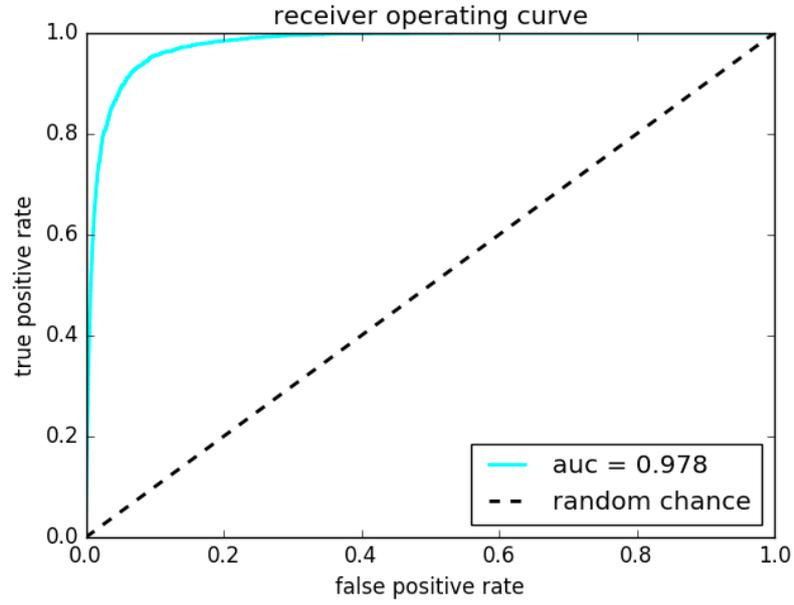


Figure 4.7: ROC curve of the low-pt BDT for the resolved signal.

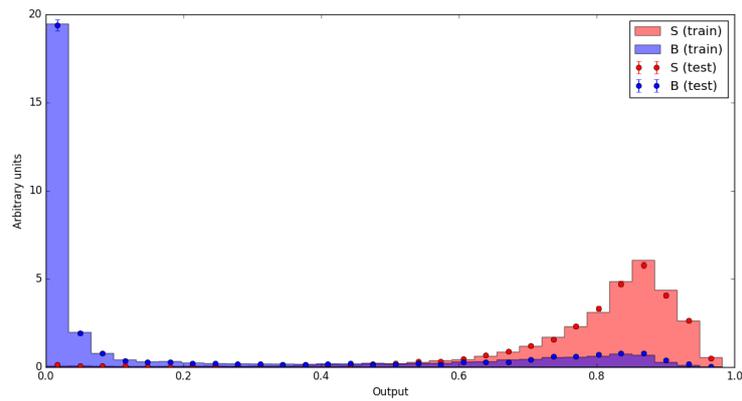


Figure 4.8: Output of the medium-pt BDT for the merged signal.

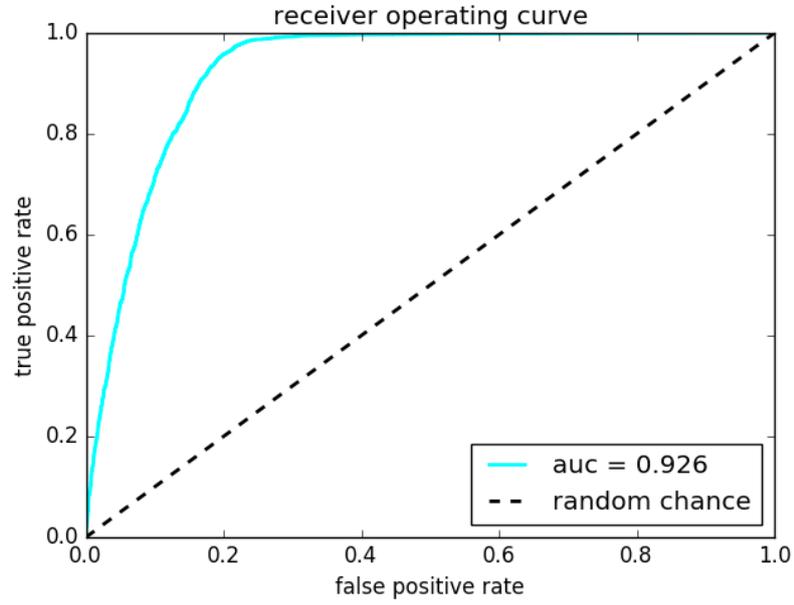


Figure 4.9: ROC curve of the medium-pt BDT for the merged signal.

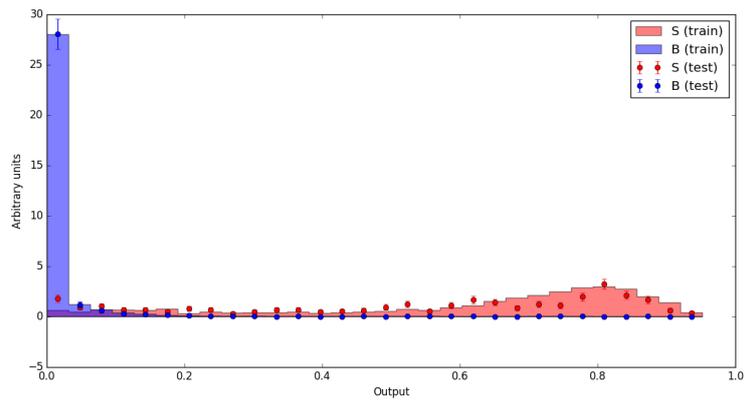


Figure 4.10: Output of the medium-pt BDT for the resolved signal.

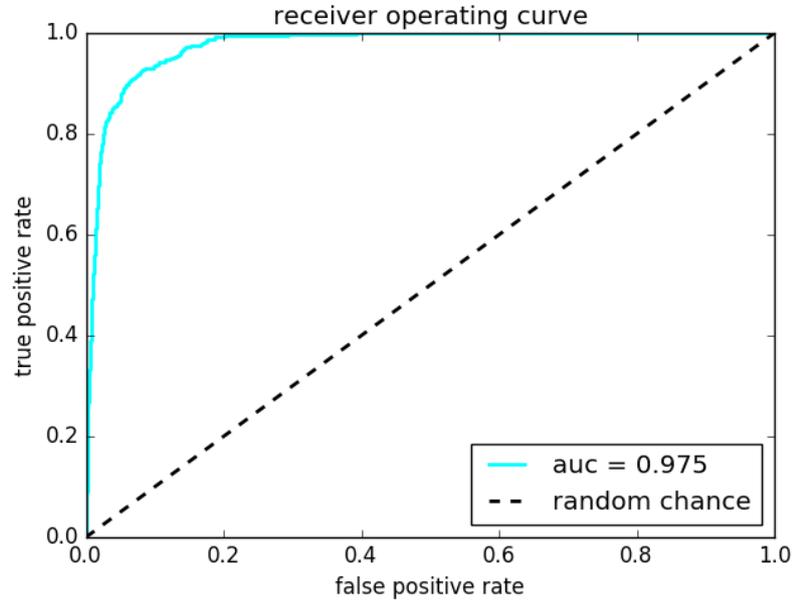


Figure 4.11: ROC curve of the medium-pt BDT for the resolved signal.

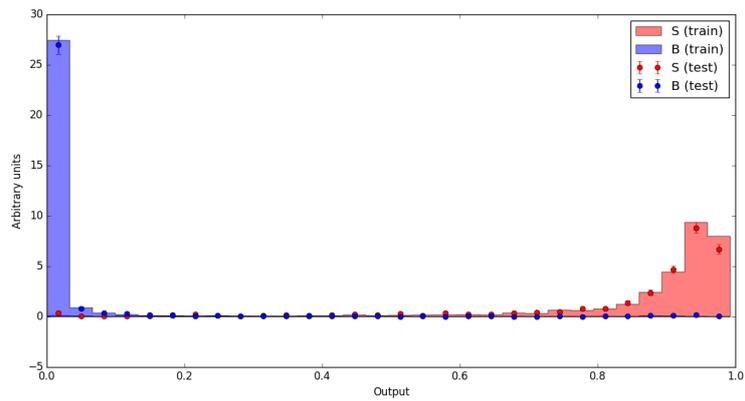


Figure 4.12: Output of the high-pt BDT for the merged signal.

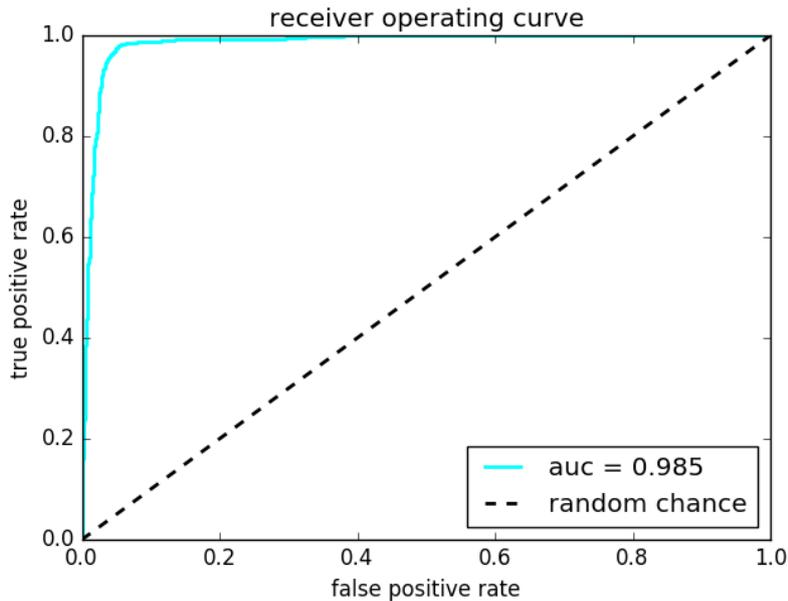


Figure 4.13: ROC curve of the high-pt BDT for the merged signal.

graph shows that most of the background at high output of the BDT is composed of B1 background. This is due to the way we prepared the dataset: for each jet, we consider several triplets with several muons, but only one of them is a signal triplet. The other ones constitute a type of background which the BDT is not able to classify properly. In order to overcome this problems, we developed a DNN and formatted the dataset in a different fashion, that we will describe further down this chapter in sec. 4.10.

We also compared our top-tagging BDT to a cut and count selection, which required:

- the muon to be a tight muon, i.e., it must comply with some strict criteria designed by the CMS group which works on muons;
- the muon's miniiso to be less than 0.1;
- the jet to be b-tagged, with different possible working points.

We plotted the selection efficiency and the background efficiency with respect to the p_t of the jet, for two working points of the BDT and the three working points of the DeepCSV b-tagger used in the cut and count selection. The plots, for merged and resolved configuration, are shown in fig. 4.15, 4.16, 4.17, 4.18. We can notice how, compared to the standard cuts, the BDT have comparable background efficiencies and higher signal efficiencies, in particular at high p_t of the jet, where conventional b-taggers tend to work poorly.

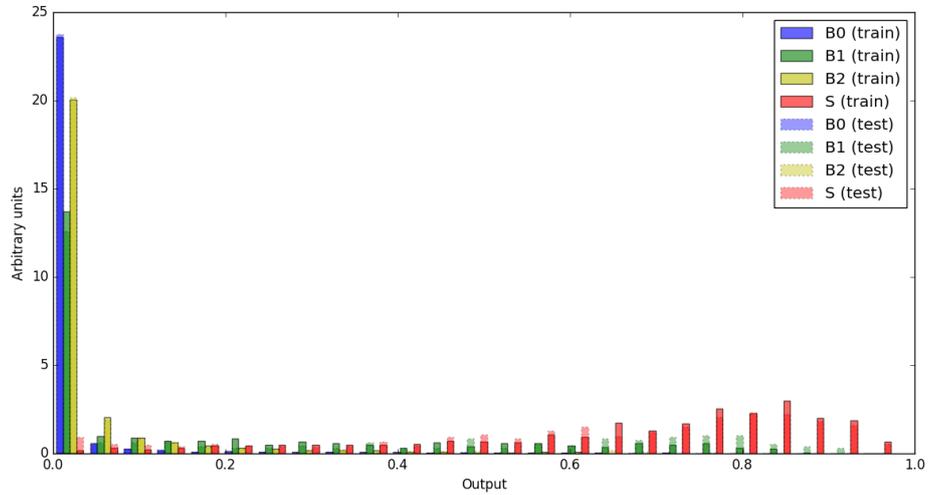


Figure 4.14: Distribution of the three different types of background for the low-pt resolved configuration. Each bin for each distribution has size 0.03 and the bins have not been superimposed for graphical clarity.

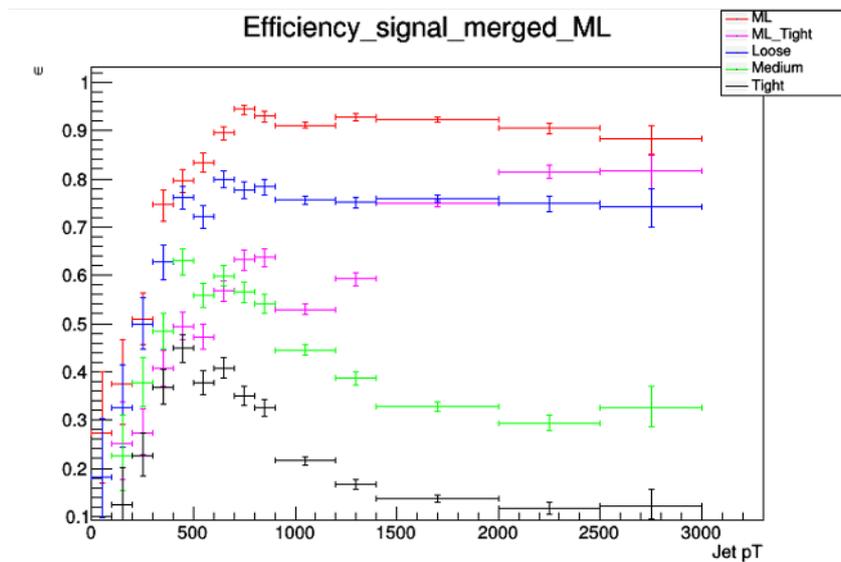


Figure 4.15: Signal efficiency for the merged configuration.

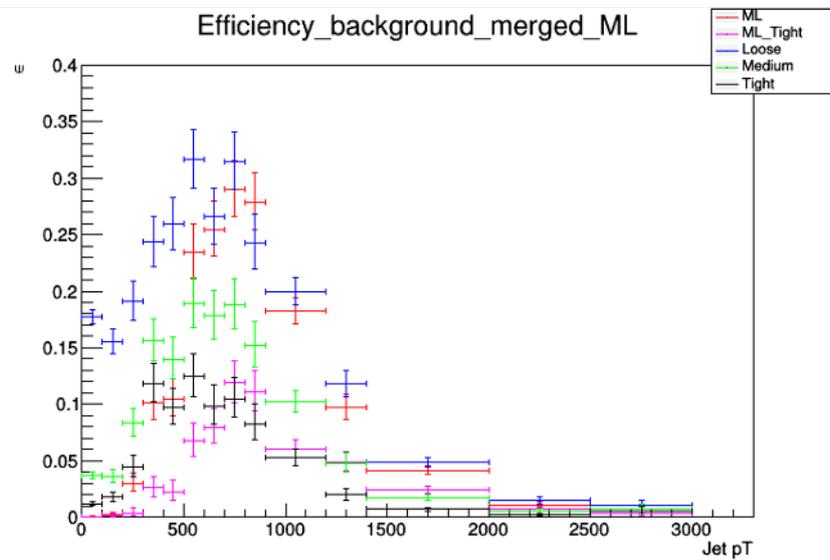


Figure 4.16: Signal background for the merged configuration.

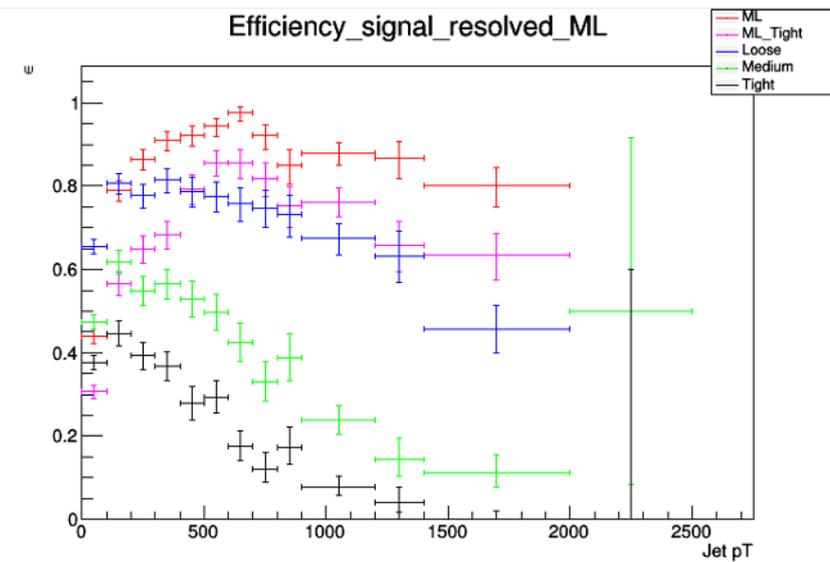


Figure 4.17: Signal efficiency for the resolved configuration.

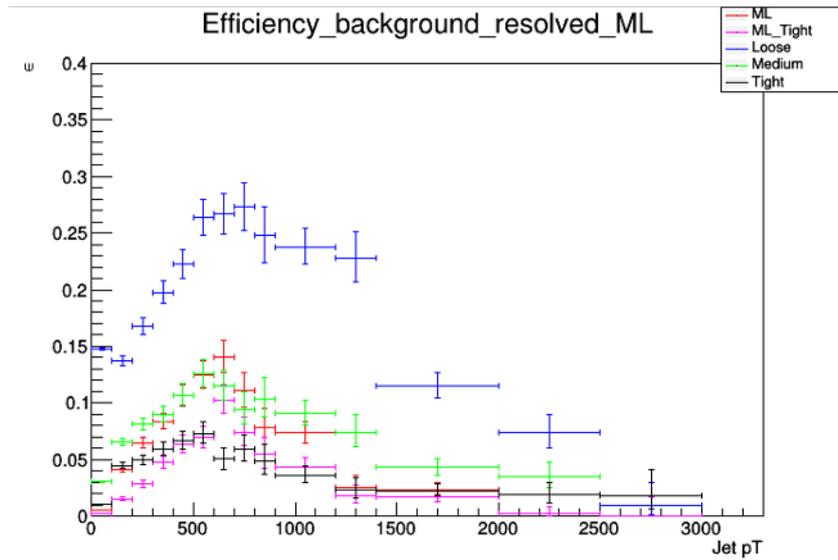


Figure 4.18: Signal background for the resolved configuration.

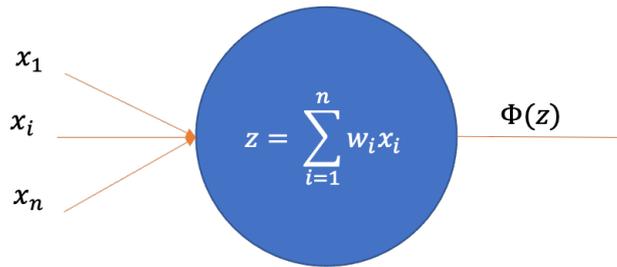


Figure 4.19: Schematic display of the operation principles of an artificial neuron.

4.9 Deep Neural Networks

A Deep Neural Network (DNN) is a machine learning classification algorithm whose concept was built upon hypotheses and models of how the human brain works to solve complex problem tasks. The building blocks of a DNN are artificial neurons, whose operation principles are displayed in fig. 4.19. The neuron receives n input features x_i which are put in a weighted sum $z = \sum_{i=1}^n w_i x_i$. Hence the neuron computes an activation function $\Phi(z)$, which is output by the neuron itself. A DNN is constituted of layers of artificial neurons (fig. 4.20):

- the first layer is composed of neurons which receive the input features of the algorithm and output the outcome of an activation function;
- the hidden layers are composed of neurons which receive the outputs of the neurons of the previous layers as input;

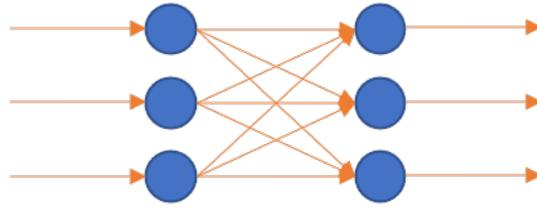


Figure 4.20: Schematic layout of a Deep Neural Network.

- the last layer outputs the result of the algorithm and terminates the classification task.

So far, we have described what is called a "Dense" neuron layer. Our top-tagging DNN employs a wider range of layers:

- **dropout:** it is a layer only used during the training phase, which randomly deactivates some of the connections between the layers it is interposed between. It is used to create more robust connections and minimize overtraining;
- **batch normalization:** it rescales the features it receives in input into a normal distribution;
- **long short-term memory:** a recurrent neural network (RNN), which has feedback connection, i.e., the output of the layer is also passed to the input of itself together with the following input;
- **convolutional neural network:** the neurons of this layer perform a convolution between the input features and a filter, which contains the weights of the neurons;
- **max pooling:** it reduces the dimensionality of the output space of a convolutional neural network by taking the maximum output between an arbitrary number of outputs;
- **masking:** it is used to discard a sample when some data is missing.

This type of machine learning algorithm belongs to the category of the supervised machine learning algorithms, i.e., it must be trained on a set of already classified data before it can be used to make predictions. The training phase consists in the iterative readjustments of the weights w_i of every neuron in the DNN in order to minimize a loss function which is a measure of how inefficient the algorithm is at performing its classification task. For our algorithm, the classification task is categorical, i.e., the DNN must predict which certain category a sample belongs to, as opposed to a binary classification task in which the algorithm needs to classify in only two groups.

The loss function chosen for our algorithm is the categorical crossentropy. Let us suppose that a dataset is composed of data classifiable in three categories. Then, for every sample of data i , it is possible to consider a three-dimensional

vector $\hat{y}_j^{(i)}$, with $j = 1, 2, 3$ which is 1 if the sample belongs to category j and 0 otherwise. Then we can express the categorical crossentropy loss calculated on n data samples as:

$$L = - \sum_{i,j} \hat{y}_j^{(i)} \log(y_j^{(i)}), \quad (4.14)$$

where $y_j^{(i)}$ is the prediction of the algorithm, which includes the dependency on the weights of every neuron in the DNN.

4.10 Top tagging with DNNs

As mentioned in sec. 4.5, three DNNs have been developed for the top-tagging of hadronic jets in this thesis work, using the Keras and Tensorflow modules in Python. The dataset used for the training and application of the DNN has been formatted in a way to define a physics object called *lepto-jet*. The lepto-jet is a "container datum", constructed on the basis of an hadronic jet, which includes variables regarding:

- the jet itself;
- the Particle Flow candidates which are clustered in the jet;
- the secondary vertices included in the jet;
- the top candidates (both merged and resolved) constructed with the jet, the MET of the event and all the muons which are in an angular distance $\Delta R < 2$ from the jet;
- the muons considered for the top candidates.

Furthermore, we redefined the two categories of top candidate for the use with the DNN:

- **merged:** if the muon is part of the component of the jet. In this case, the top's 4-momentum is obtained by taking the sum of the jet and the neutrino;
- **resolved:** if the muon is not reconstructed inside the jet, so we obtain the top candidate as the sum of a jet, a muon and a neutrino.

The variables analyzed are summarized in tab. 4.2.

The top-tagging DNN has been designed to analyze properly each part of the leptojet. Its architecture is displayed in fig. 4.21. The features which regard the jets are analyzed by a dense layer after a batch normalization layer. The features regarding the PF candidates and the SV are analyzed by a convolutional layer whose output is maxpooled before going through an LSTM layer. On the other hand, the features which refer to the muons, the top candidates and the jet unboosted are all normalized and analyzed by an LSTM layer. At the end of the DNN the outputs of the separate branches are put together with a concatenate

Jet	PF candidates	Secondary vertices	Top	Top $_{\nu}$	Muons
mass	mass	mass	mass	mass	miniiso
p_t	p_t	p_t	$p_{t,rel}$	p_t	dxy
ϕ	ϕ	ϕ_{rel}	$cos\theta_{\mu,b}$	ϕ	dxy $_{err}$
η	η	# tracks		η	dz
$p_{t,ub}$	d0	χ^2_{norm}			is Global
η_{ub}	dz	dxy			# tracker hits
ϕ_{ub}					$p_{t,ub}$
energy $_{ub}$					ϕ_{ub}
mass $_{ub}$					η_{ub}

Table 4.2: Variables employed by the DNN top tagger.

layer, before going through two pairs of dense and dropout layers. The last layer has three neurons, which output the probability that the jet is background, top merged or top resolved. It is important to notice that this top-tagger is capable of recognizing top jets with good performances (sec. 4.11), but, as it is designed, it does not define which top quark candidate is the right one to be reconstructed among the candidates associated to the same leptojet. This choice is made by picking the top candidate whose mass is closer to $m_{top} = 172.76 \pm 0.30 \text{ GeV}$ [1].

4.11 Top tagging performances with DNNs

In this section, we discuss the top-tagging performances of the DNN we developed. We trained and tested the performance of the algorithm on data obtained from Monte Carlo simulations of W' and Z' production at LHC. The samples were created with different mass hypotheses of 4, 5 and 6 TeV. We prepared the dataset by picking jets which have at least one merged or resolved muon with miniiso less than 0.5. We then split the dataset in three smaller ones, based on the p_t of the jets:

- **low p_t :** $p_t < 1000 \text{ GeV}$;
- **medium p_t :** $p_t > 1000$ and $p_t < 2000 \text{ GeV}$;
- **high p_t :** $p_t > 2000 \text{ GeV}$.

As already done for the BDTs (4.8), we split the three datasets in training and validation set and then trained the three DNNs. For each one of the three networks, we considered the distribution of the output of the DNN for test and training sets, which represent the probability of the jet being either merged signal or resolved signal.

The remaining part of this section is dedicated to the display of the distribution of the output of the DNNs and the related ROC curves. Fig. 4.22 shows the output of the merged signal for the the low-pt configuration. It is displayed a good discrimination between the signal and the background, with no signs of

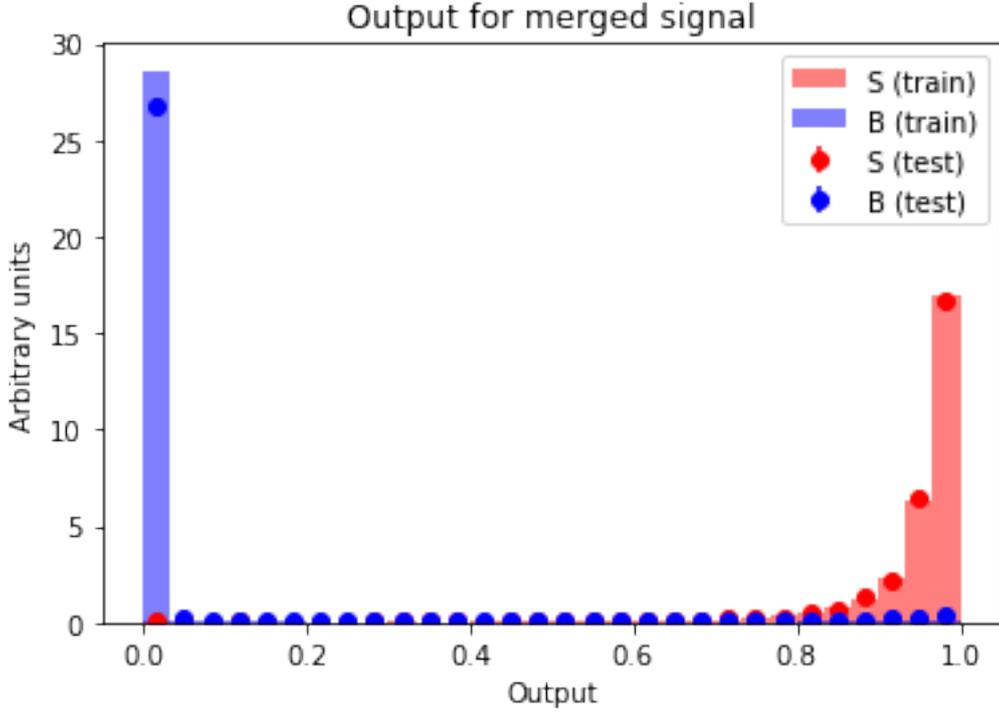


Figure 4.22: Output of the low-pt DNN for the merged signal.

overtraining. The ROC curve is displayed in fig. 4.23, with an area under the curve of 0.984, which is very good for this task. The output and ROC curve for the resolved signal in the low-pt configuration are shown respectively in fig. 4.24, 4.25. The output shows some deviations between the output of the training set and the test set. This is not due to overtraining, but rather to the fact that we had to undersample the background in the training set, so that the distributions of signal and background are different between the two sets. This was necessary because DNNs do not train properly with unbalanced classes and the dataset was excessively unbalanced, as the background made up 90% of it.

The output and ROC curve of the medium-pt DNN for the merged signal are shown in fig. 4.26, 4.27. In this case, we can see that the area under the ROC curve is less than the one for the low-pt configuration. This is due to an excessive number of background events, supposedly because of the kinematic properties of such jets mimicking the ones of top quarks. The output and ROC curve of the medium-pt DNN for the resolved signal are shown in fig. 4.28, 4.29 and display a good discrimination power. Finally, the output and ROC curve of the high-pt DNN for the merged signal are shown in fig. 4.30, 4.31 and they also show a good discrimination power. We do not report the graphs for the resolved signal of the high-pt configuration, because even redefining the resolved category, there are very few signal samples, for the same reasons discussed in sec. 4.8. After the development of the DNN, two working points have been chosen: a medium one (with misidentification of 1%) and a tight one (with misidentification 0.1%). The efficiencies are shown in tab 4.3. It is worth noticing that the top-tagger also

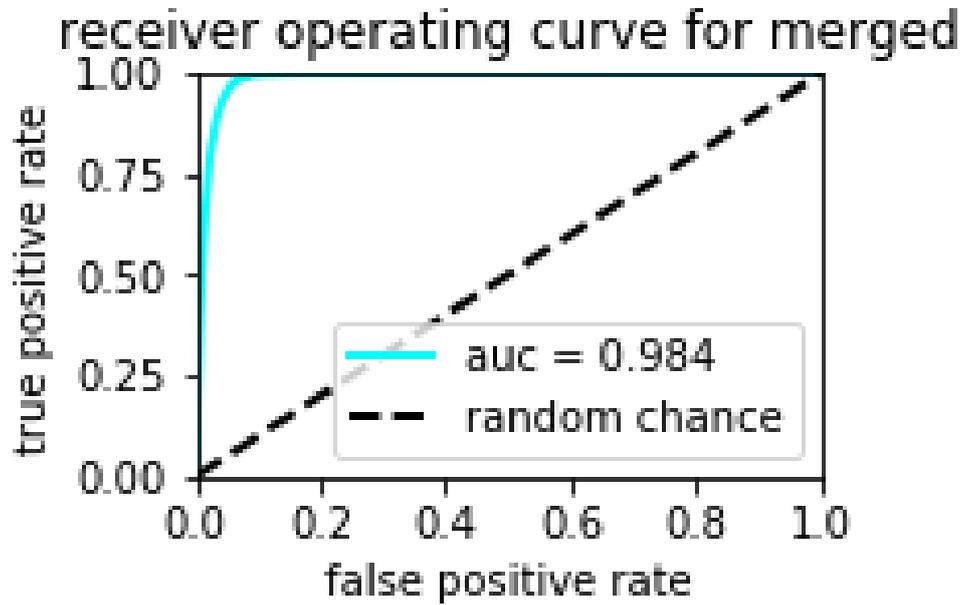


Figure 4.23: ROC curve of the low-pt DNN for the merged signal.

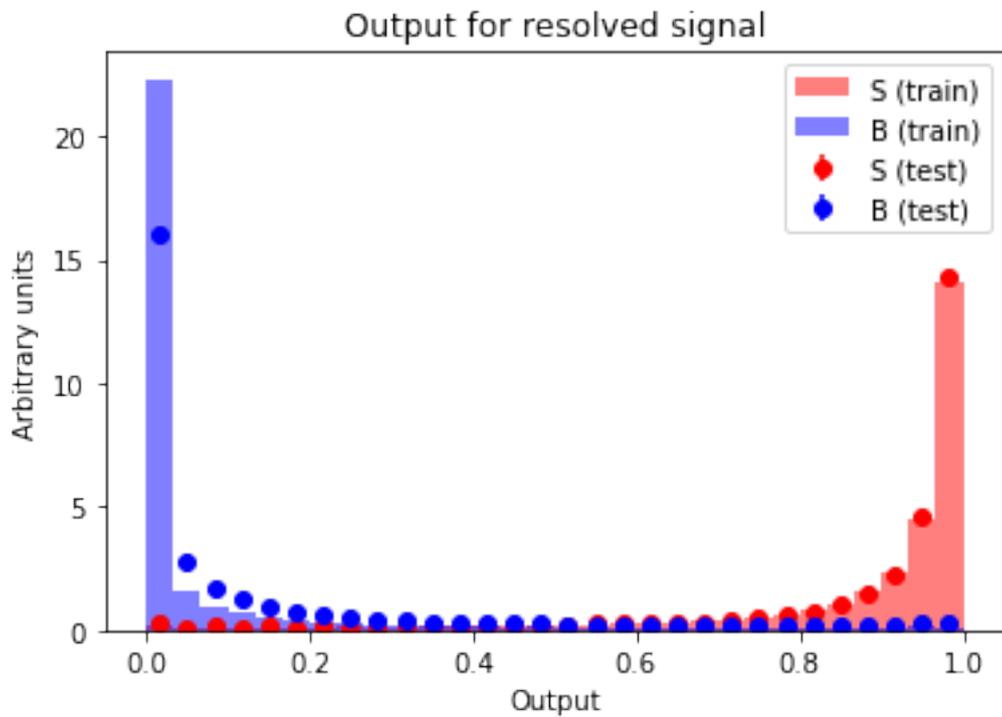


Figure 4.24: Output of the low-pt DNN for the resolved signal.

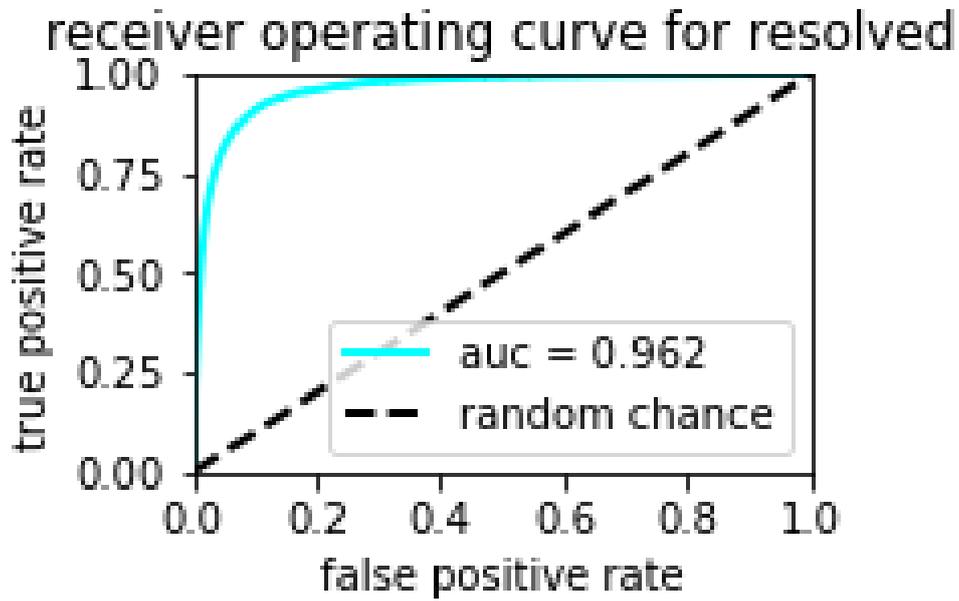


Figure 4.25: ROC curve of the low-pt DNN for the resolved signal.

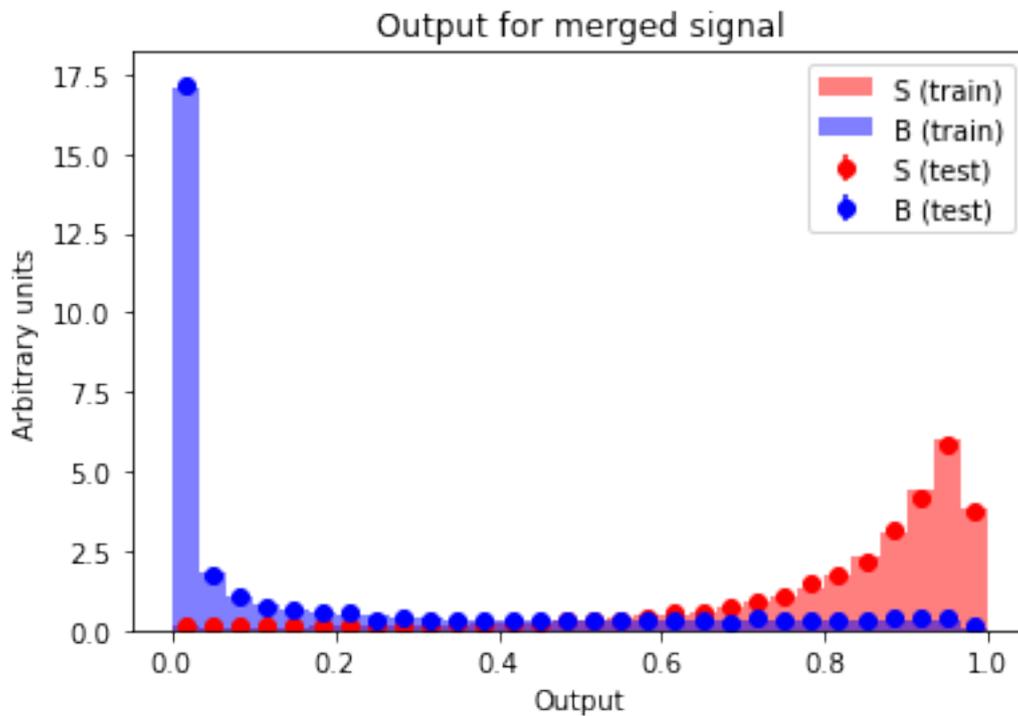


Figure 4.26: Output of the medium-pt DNN for the merged signal.

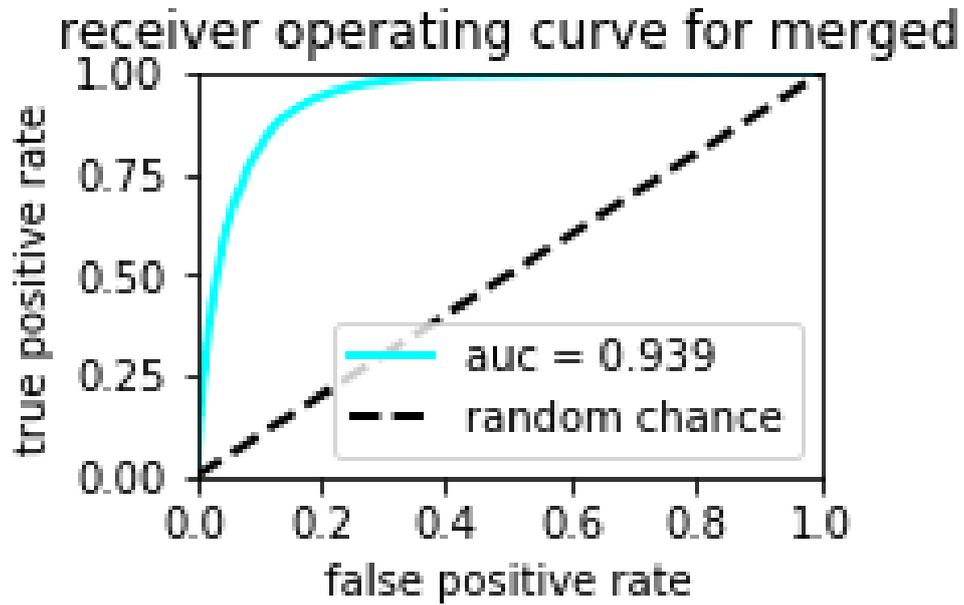


Figure 4.27: ROC curve of the medium-pt DNN for the merged signal.

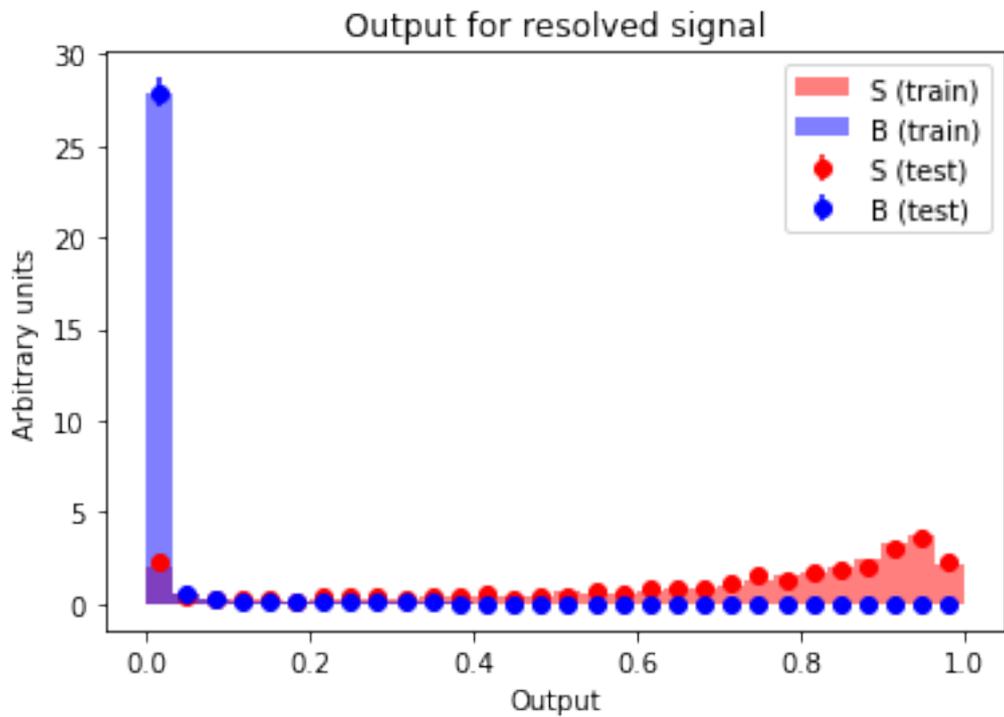


Figure 4.28: Output of the medium-pt DNN for the resolved signal.

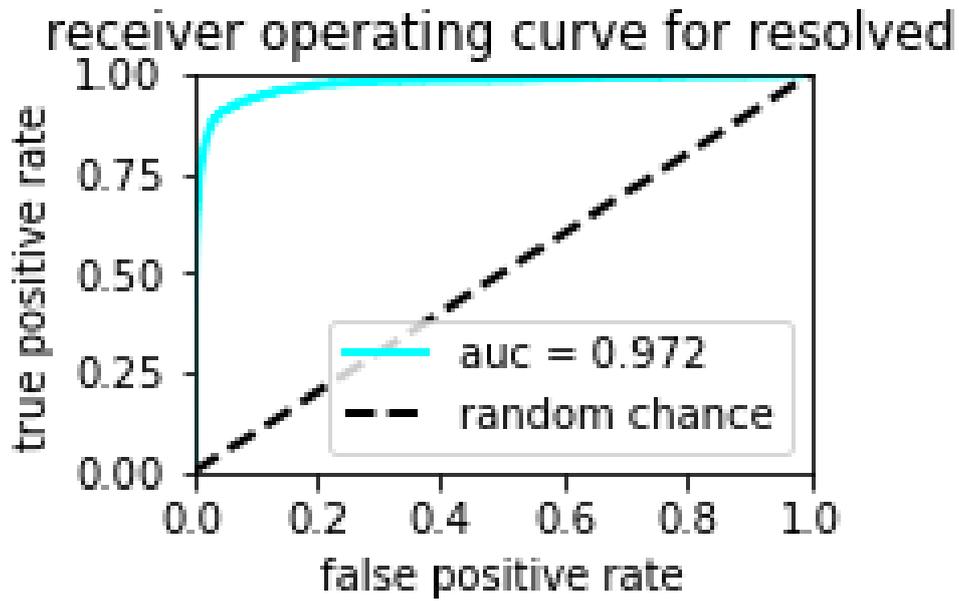


Figure 4.29: ROC curve of the medium-pt DNN for the resolved signal.

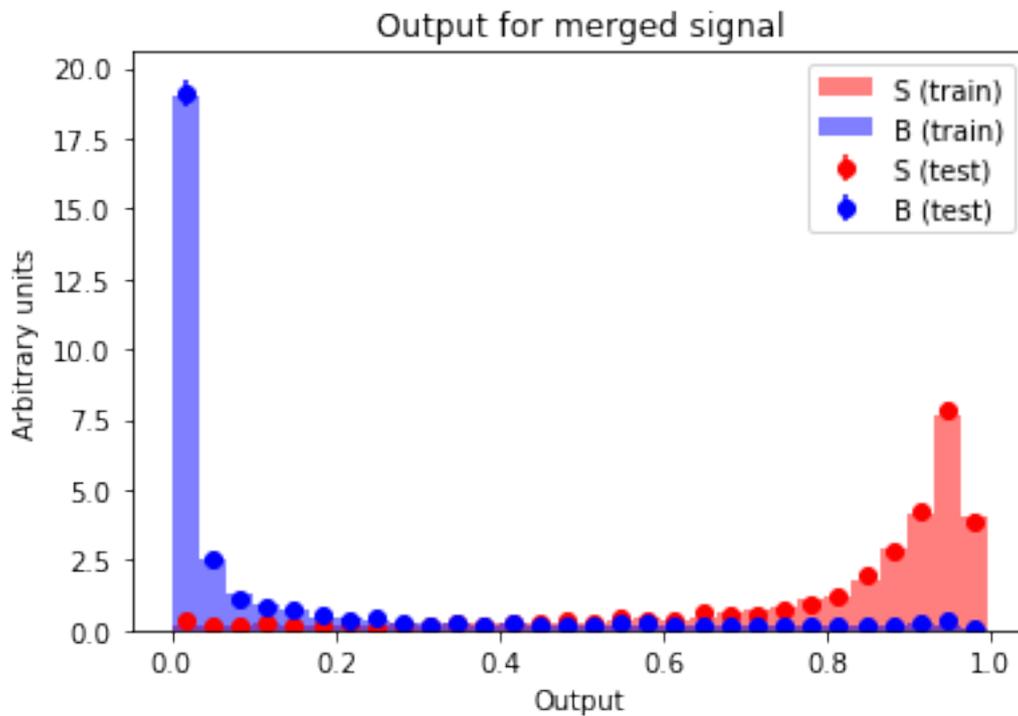


Figure 4.30: Output of the high-pt DNN for the merged signal.

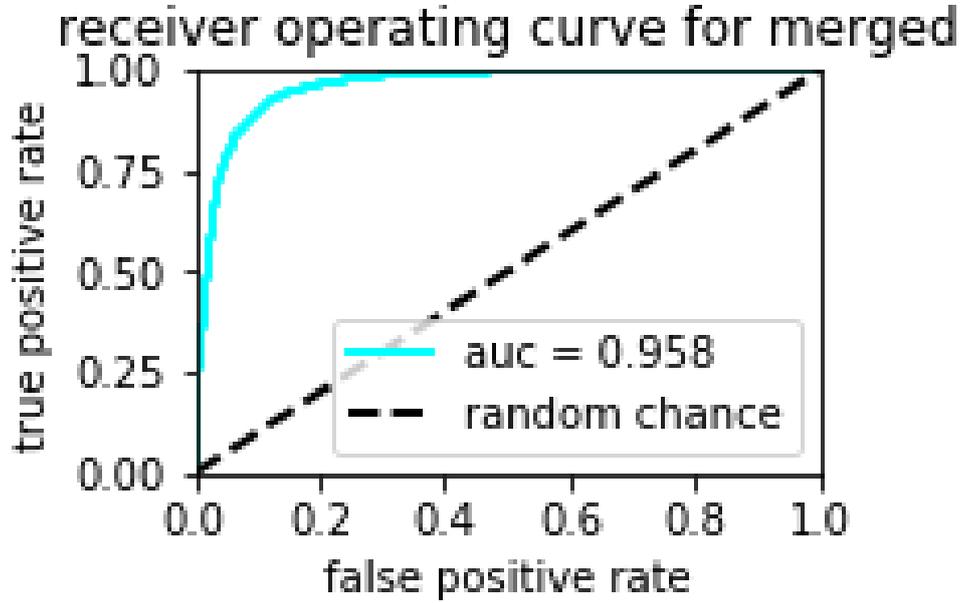


Figure 4.31: ROC curve of the high-pt DNN for the merged signal.

Configuration	Medium Signal Efficiency	Tight Signal Efficiency
low merged	88.7 %	57.4 %
low resolved	82.7 %	47.1 %
medium merged	82.1 %	30.2 %
medium resolved	75.9 %	59.9 %
high merged	84.7 %	60.3 %
high resolved	32.2 %	30.9 %

Table 4.3: Efficiencies for the three DNNs at the two working points.

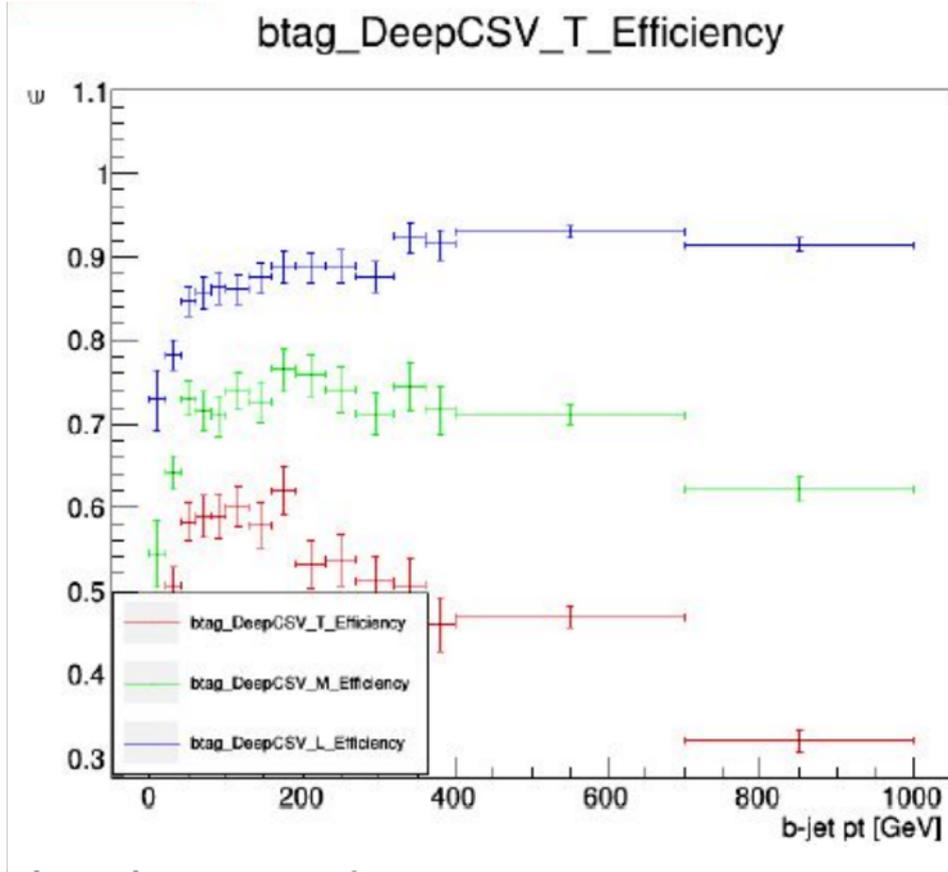


Figure 4.32: Efficiency of the DeepCSV b-tagger with respect to the p_t of the jet.

performs a b-tagging task, as only b-jets can be the decay products of a top quark. Thanks to the fact that only b-jets make up the signal samples for the DNNs we designed, we can compare the efficiencies in tab. 4.3 with the efficiency of the DeepCSV b-tagger, shown in fig. 4.32. It is noticeable how the efficiency rapidly decreases with increasing p_t of the jet. For instance, the medium working point presents an efficiency of about 60 % at 1000 GeV of p_t of the jet. On the other hand, the DNN object of this thesis has a significantly higher efficiency even at higher p_t of the jet. This fact explains the choice of not restricting the sample used for the DNN to only b-tagged jets, as the total efficiency would be the product of the top-tagging efficiency and the b-tagging efficiency, which would be lower of the efficiencies that have been obtained in this thesis work.

Chapter 5

Application of ML algorithm and predictions on W' analysis

In this chapter, we will describe the analysis we performed for the reconstruction of the W' boson from simulated signal samples. As mentioned in Chapter 4, we are looking for the decay of the W' into a bottom and top quarks, with the latter decaying leptonically into a muon and a neutrino, as shown in fig. 5.1. The W' 4-momentum is obtained as the sum of the 4-momentum of a b-jet and top quark, the latter decaying leptonically into a muon. The top quark is selected by the DNN described in ch. 4.

5.1 Simulated samples

We performed the analysis on Monte Carlo-simulated signals of production of W' which decays:

$$W' \rightarrow t \bar{b} \rightarrow l^+ \nu_l b \bar{b}. \quad (5.1)$$

Three cases of left-handed W' mass have been considered: 4000, 5000 and 6000 GeV, with a decay width of 1% of the mass.

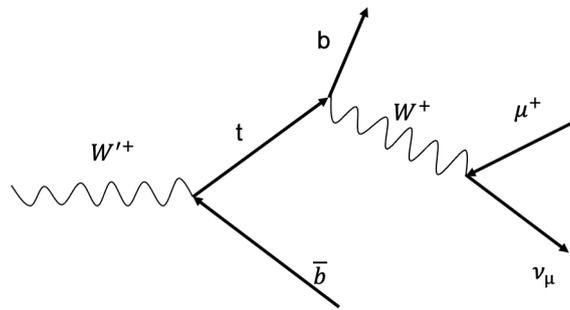


Figure 5.1: An example of Feynman diagram for the $W' \rightarrow t b$ process with the top decaying to a muon.

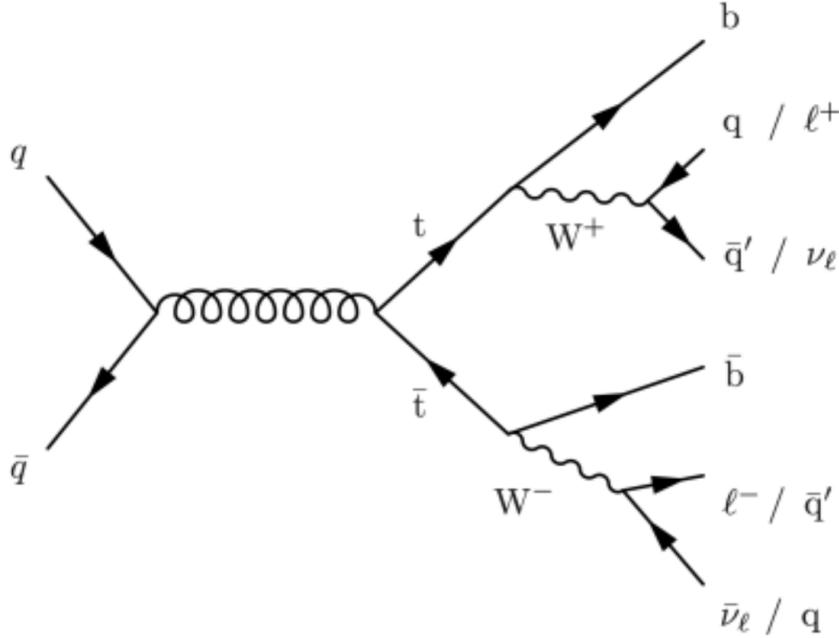


Figure 5.2: Feynman diagram of a process of $t \bar{t}$ quark-pair creation.

5.1.1 Background samples

We have considered two type of background samples which can mimic the final state we are looking for:

- $t\bar{t}$: a top quark-antiquark is produced, as shown in fig. 5.2. If at least one of the two top quarks decays leptonically, this process can fake a W' decay.
- $W + jets$: the associated production of a W boson and two jets can mimic the signal final state if the W decays $W \rightarrow \mu \nu_\mu$ and the jets are b-jets, as show in fig. 5.3.

Tab. 5.1 describes the employed simulates samples and their respective cross sections multiplied for the branching ratios. $M_{t \bar{t}}$ refers to the invariant mass of the pair $t\bar{t}$, and HT refers to the hadronic tranverse energy of the jets.

5.2 W' reconstruction

In order to reduce the background from the sources which are less similar to the signal and to apply the top tagger to a restricted number of jets, a first set of selection is applied, called *preselection*. We request that the event presents at least a muon with $miniiso < 0.5$ and at least a jet with $p_T > 200 GeV$. The first step of the reconstruction is the selection of the jet with the largest DNN score, on the basis of which, we have constructed three regions, based on the working points defined in sec. 4.11:

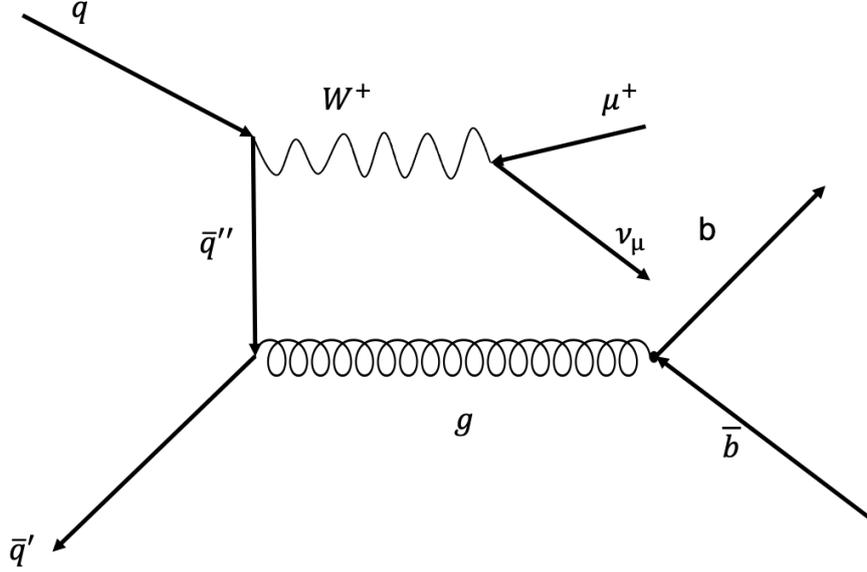


Figure 5.3: Feynman diagram of a $W + \text{jets}$ process, with W decaying to a muon and a neutrino and two b-jets.

Sample	$\sigma \times BR$ (pb)
$W'(4TeV)$	0.01679
$W'(5TeV)$	0.003381
$W'(6TeV)$	0.0008807
$t\bar{t}_1, 700GeV < M_{t\bar{t}} < 1000GeV$	80.5
$t\bar{t}_2, M_{t\bar{t}} > 1000GeV$	21.3
$W + jets_1, 1200GeV < HT < 2500GeV$	1.6081
$W + jets_1, HT > 2500GeV$	0.038914

Table 5.1: Simulated samples and respective cross sections multiplied for the branching ratios.

- **Medium:** if the jet's DNN score is over the medium threshold;
- **Tight:** if the jet's DNN score is over the tight threshold;
- **Low purity:** a region which is filled with jets whose DNN score is over the medium threshold, but under the tight threshold, that could potentially be simultaneously fit with the Tight region without overlap.

From this jet, we reconstruct the best top as the one with the mass closer to the measured mass of the top quark (see sec. 4.10). Thereafter, we need to determine the b-jet which is the decay product of the W' boson. We request it to be the jet with the highest p_t and check that its features respect the following criteria:

- $p_t > 200\text{GeV}$;
- **DeepFlavor Medium Working Point:** the working point of the Deep-Flavour b-tagger which presents a misidentification probability of 1%;
- **Tight Id:** it has to respect certain criteria of quality of the reconstruction, defined by the CMS group which works on the reconstruction of hadronic jets.

We compare the described selection flow with another algorithm which does not make use of machine learning, which we will refer to as *standard cut*. We ask for the muon to have the following features:

- **Tight Id:** it has to respect certain criteria of quality of the reconstruction, defined by the CMS group which works on the reconstruction of muons;
- $miniiso < 0.1$;

While we require both the jets to have:

- **Tight Id;**
- $p_t > 100\text{GeV}$;
- **DeepFlavor Medium Working Point.**

We consider the jet with the leading p_t to be the decay product of the W' and the jet with the second highest p_t to be the decay product of the top quark.

5.3 Analysis predictions

This section will be dedicated to the prediction of the results that could be obtained employing our analysis strategy on a CMS dataset of luminosity $L = 137\text{fb}^{-1}$.

Sample	Tight(%)	Medium(%)	Low purity(%)	Standard cut(%)
$W'(4TeV)$	2.41 ± 0.03	8.67 ± 0.06	6.26 ± 0.05	3.51 ± 0.04
$W'(5TeV)$	2.20 ± 0.03	7.80 ± 0.06	5.61 ± 0.05	2.99 ± 0.04
$W'(6TeV)$	1.82 ± 0.03	6.80 ± 0.06	4.98 ± 0.05	2.50 ± 0.04
$W + jets_1$	0.428 ± 0.003	2.320 ± 0.007	1.892 ± 0.006	0.436 ± 0.003
$W + jets_2$	0.690 ± 0.005	3.23 ± 0.01	2.55 ± 0.01	0.890 ± 0.006
$t\bar{t}_1$	0.317 ± 0.001	1.225 ± 0.002	0.908 ± 0.002	0.721 ± 0.001
$t\bar{t}_2$	0.863 ± 0.002	2.762 ± 0.004	1.899 ± 0.003	1.760 ± 0.003

Table 5.2: Selection efficiencies for each working point for every sample.

5.3.1 W' mass distributions

As mentioned in sec. 5.2, we reconstruct the 4-momentum of the W' boson as the sum of the 4-momenta of the reconstructed b-jet and top quark. We graphed the invariant mass m_{tb} of the couple for signal and background samples and calculated the selection efficiencies. First, we will show the distribution of m_{tb} at all working points for each sample. The selection efficiencies are shown in tab. 5.2 and are the integrals of histograms in figs. 5.4 - 5.10, in which the number of selected events is scaled for the number of total events. One can notice how the selection efficiency, for both the signal samples and the background samples, increases in the order:

- tight;
- standard cut;
- low purity;
- medium.

It is also noticeable that the selection efficiencies for the tight working point is close to the standard cut's, even though the machine learning-based tight working point presents a fairly better rejection of the $t\bar{t}$ background.

5.3.2 Sensitivity calculation

We then proceeded to calculate the sensitivities for the analyses in the four working point, considering the m_{tb} distributions for the diverse working points showed in figs. 5.11 - 5.14. The sensitivity is a test statistic used to perform statistical hypothesis testings in the cases in which the null hypothesis is represented by the existence of the signal, while the alternative hypothesis is represented by its absence. It is a test statistic used to optimize the power of search, as it can be considered as the probability that an experiment is able to observe a signal, if it does exist. For a counting experiment, it is defined as:

$$Z = \frac{s}{\sqrt{s+b}}, \quad (5.2)$$

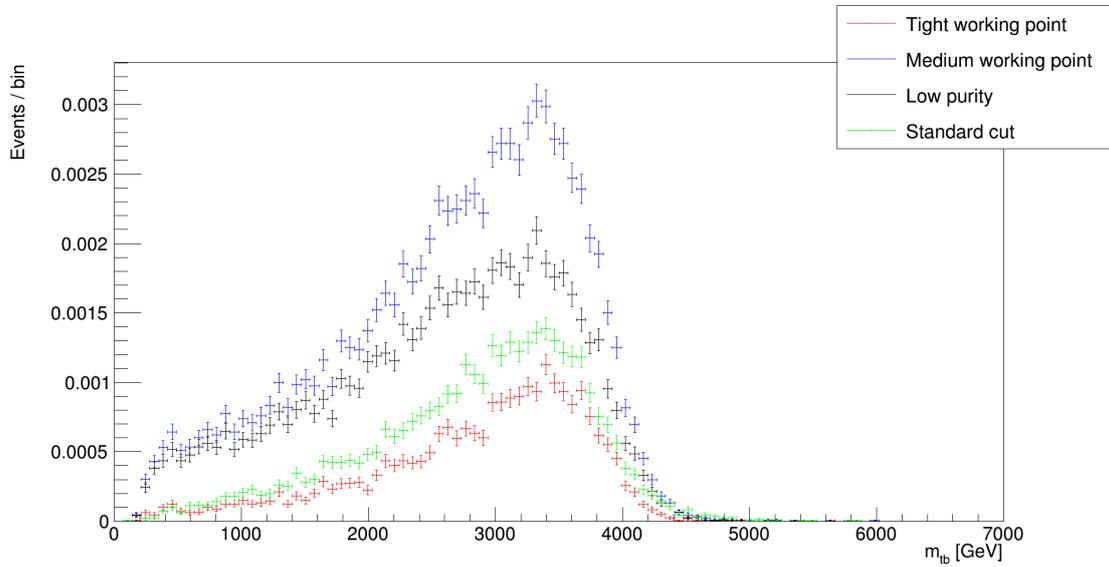


Figure 5.4: Selection efficiencies at the various working points for the W' (4 TeV) sample as a function of m_{tb} .

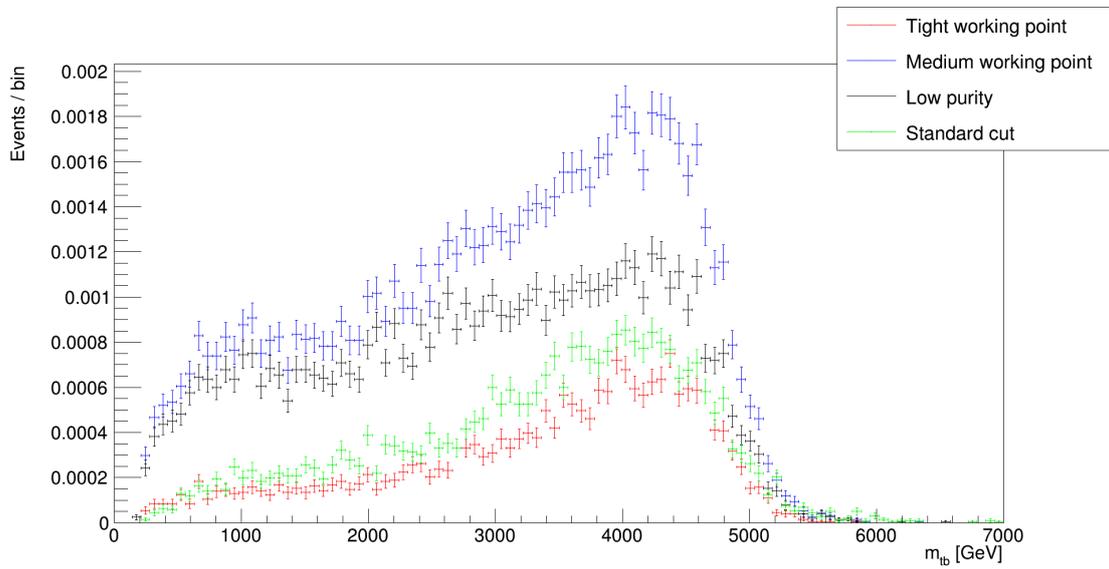


Figure 5.5: Selection efficiencies at the various working points for the W' (5 TeV) sample as a function of m_{tb} .

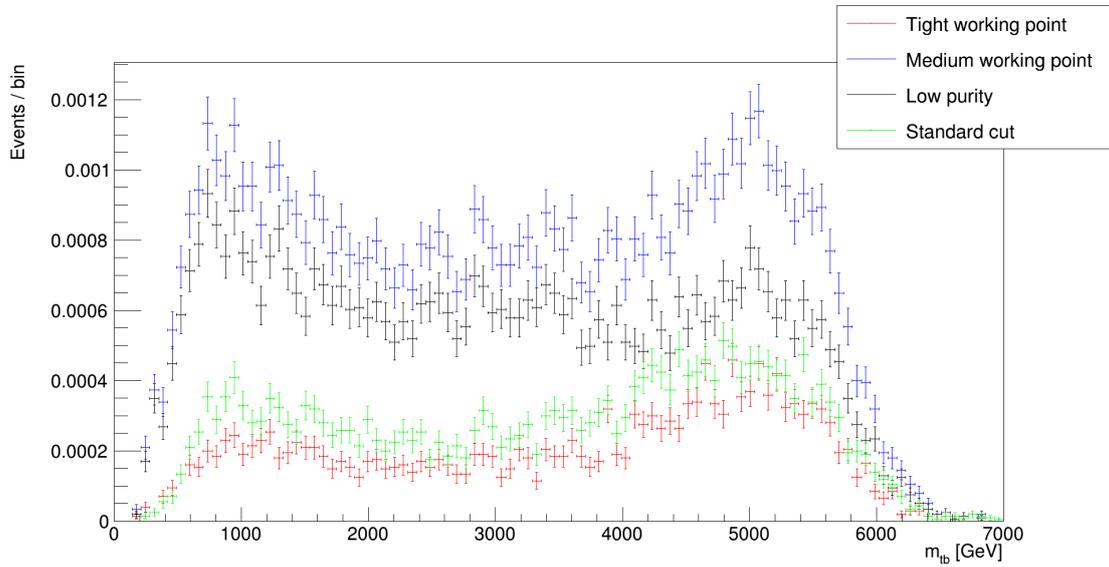


Figure 5.6: Selection efficiencies at the various working points for the W' (6 TeV) sample as a function of m_{tb} .

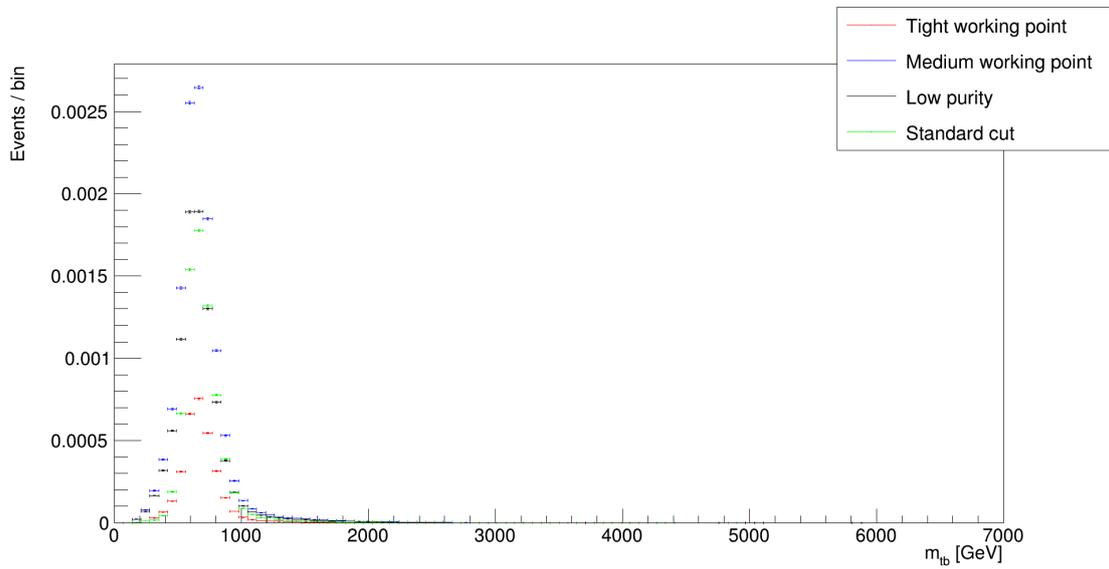


Figure 5.7: Selection efficiencies at the various working points for the $t\bar{t}$, $700\text{GeV} < M_{t\bar{t}} < 1000\text{GeV}$ sample as a function of m_{tb} .

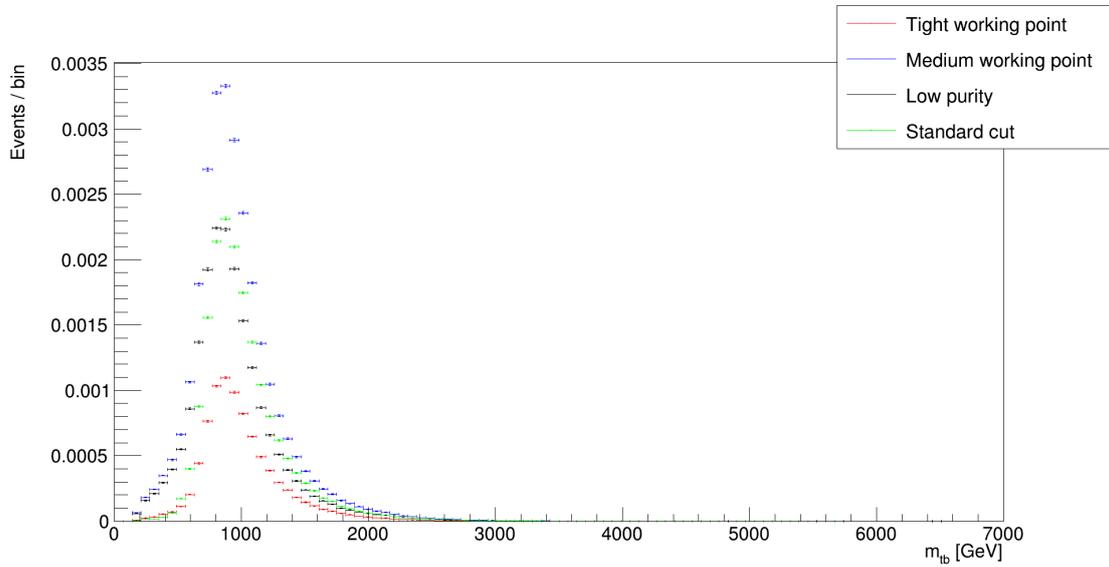


Figure 5.8: Selection efficiencies at the various working points for the $t\bar{t}$, $M_{t\bar{t}} > 1000\text{GeV}$ sample as a function of $m_{t\bar{b}}$.

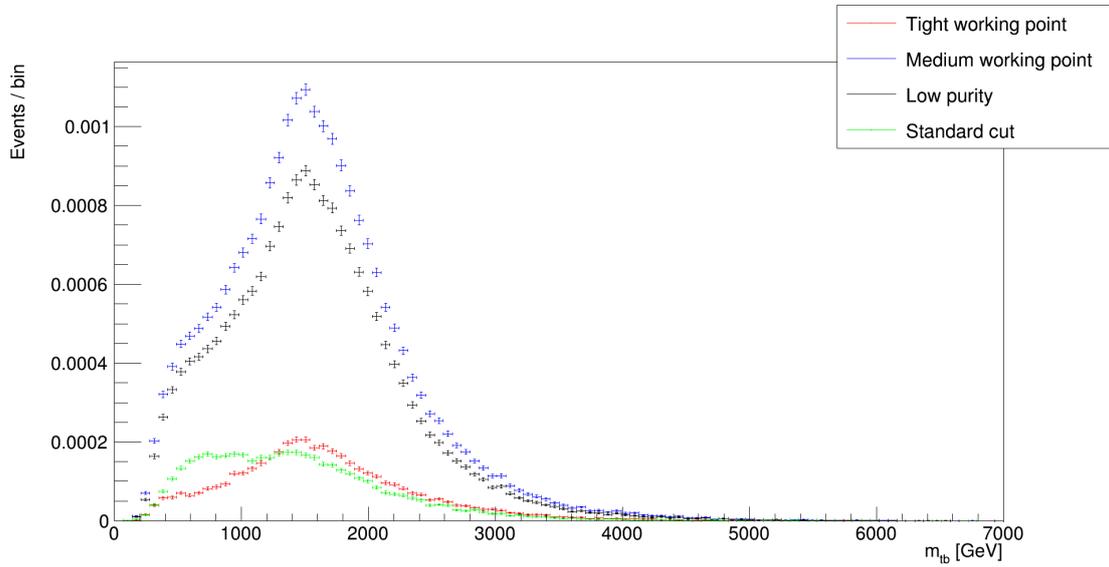


Figure 5.9: Selection efficiencies at the various working points for the $W + jets$, $1200\text{GeV} < HT < 2500\text{GeV}$ sample as a function of $m_{t\bar{b}}$.

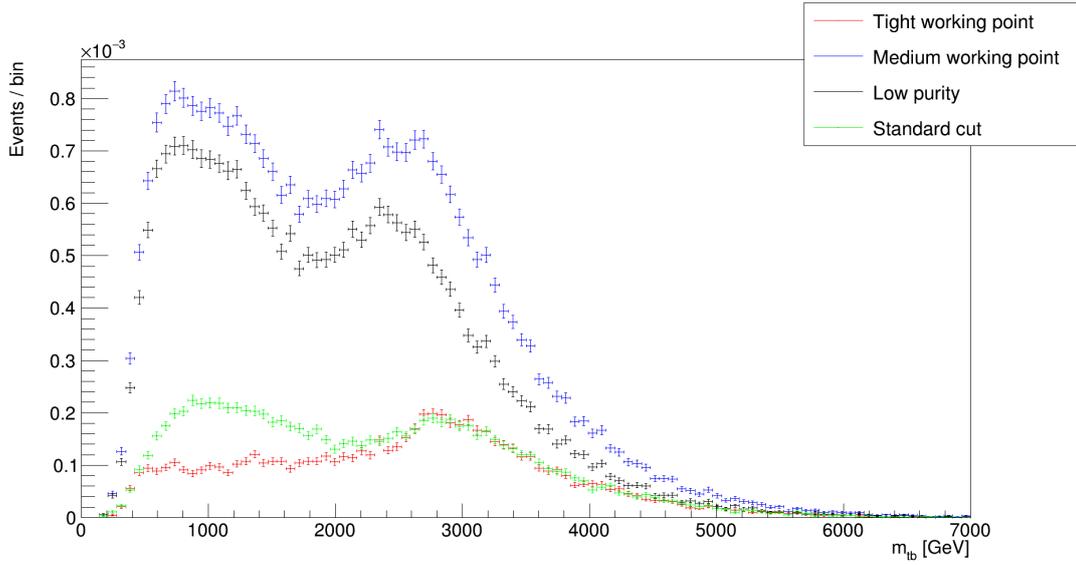


Figure 5.10: Selection efficiencies at the various working points for the $W + jets, HT > 2500 GeV$ sample as a function of m_{tb} .

Mass hypothesis	Tight	Medium	Low purity	Standard cut
4 TeV	2.65	4.04	3.12	3.16
5 TeV	0.66	0.99	0.75	0.80
6 TeV	0.13	0.22	0.17	0.16

Table 5.3: Sensitivity for the diverse mass hypotheses of the W' boson and the four working points.

where s is the number of signal events in the considered region and b is the number of background events. The sensitivities for the diverse working points and signal mass hypotheses are reported in tab. 5.3 and refer to the region for $m_{tb} > 3000 GeV$. They show that, for all mass hypotheses, the medium working point is the most suitable to perform the search for a W' boson, because it selects the most signal events without selecting an excessive number of background events at the same time. A further development in this search may be a more complex statistical analysis which could use the combination of the tight working point and a potential loose working point, which could be designed by requesting a misidentification probability of 10% for the DNN-based top-tagging algorithm.

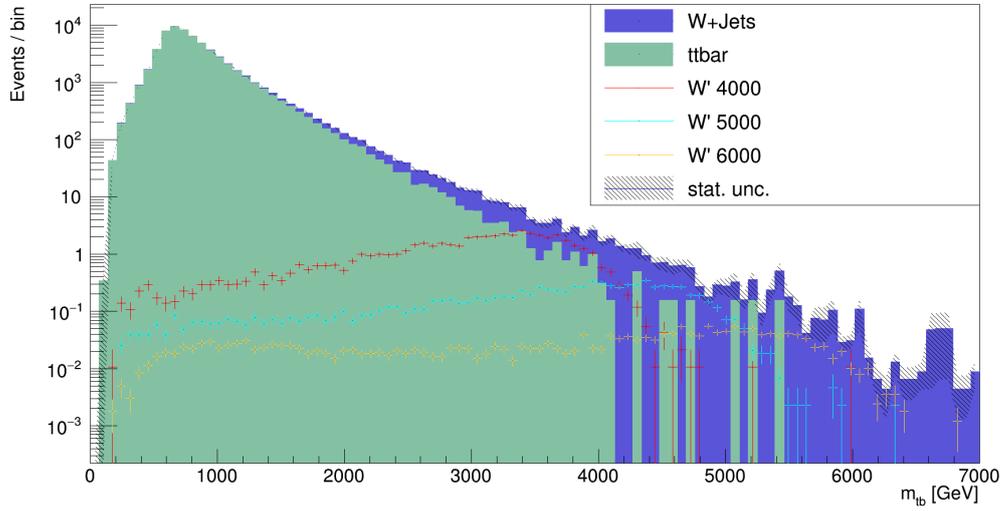


Figure 5.11: m_{tb} distribution of signal and background for the tight working point.

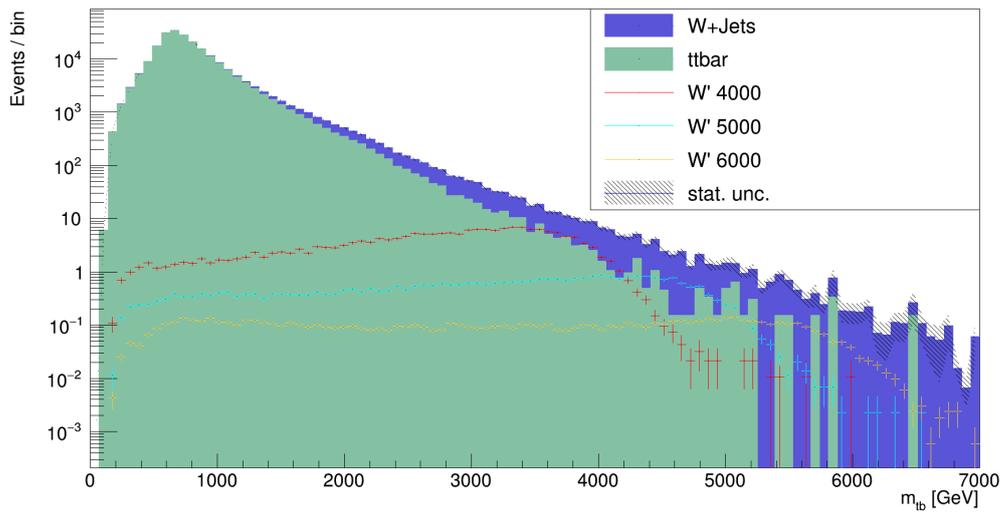


Figure 5.12: m_{tb} distribution of signal and background for the medium working point.

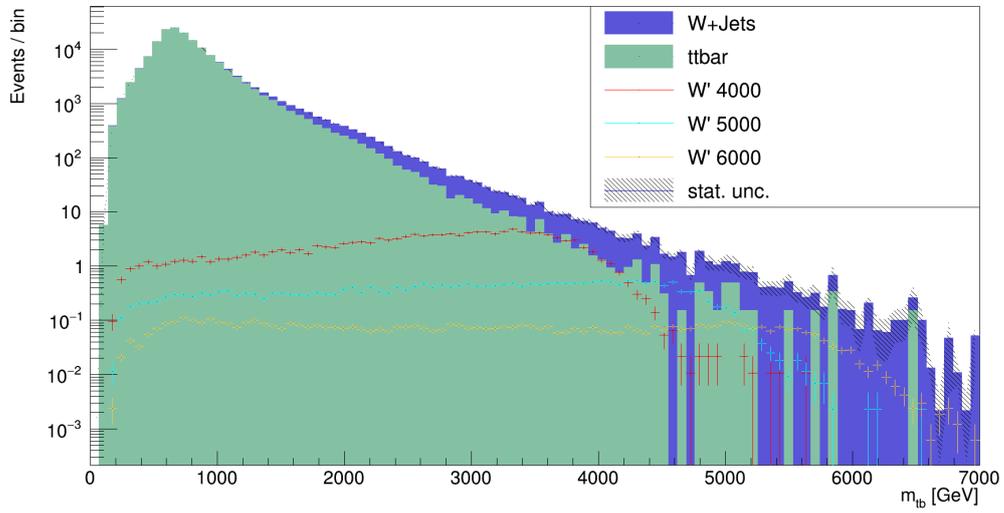


Figure 5.13: m_{tb} distribution of signal and background for the low purity working point.

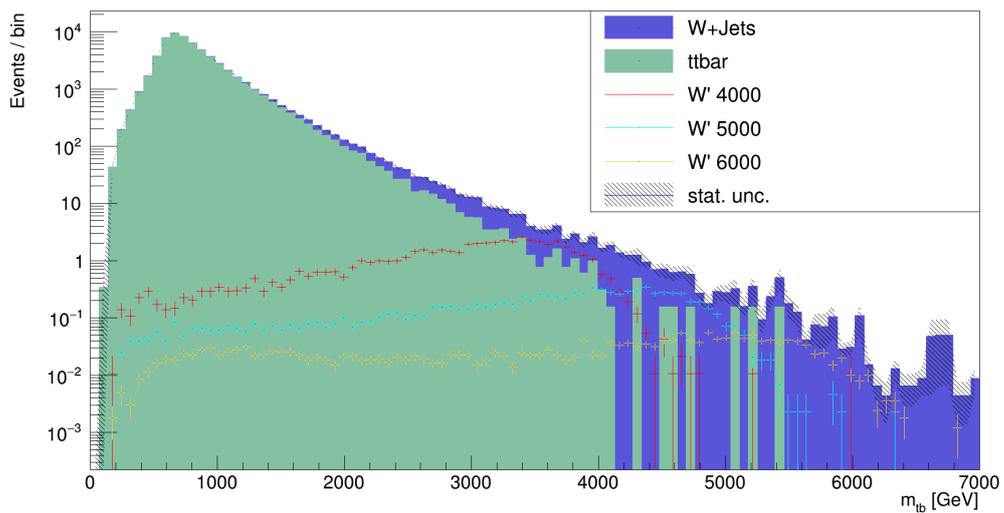


Figure 5.14: m_{tb} distribution of signal and background for the standard cut analysis.

Conclusions

In this thesis elaborate, we developed an analysis strategy based on a DNN algorithm for the search for a left-handed W' boson. We considered the decay channels $W' \rightarrow t \bar{b} \rightarrow \mu^+ \nu_\mu b \bar{b}$ making use of simulated samples of W' events produced in LHC proton-proton collisions at 13 TeV, as they would be collected by the CMS experiment.. The purpose of this thesis was to search for an alternative to cut and count selection which could result in a better sensitivity to the W' production process. A DNN algorithm was developed with the purpose of the identification of a top quark, or top-tagging, from its decay products, so to reconstruct a branch of the W' decay. In order to do this, we defined a type of datum, the "lepto-jet", which, for each hadronic jets, contains information on the jet itself, on all the muons either clustered in the jet (merged configuration) or external to it (resolved configuration) and, the top quark candidates built with said muons. We designed the top-tagging DNN so that it could analyze each set of information of the lepto-jet in a proper fashion and with the best-possible performances. Thereafter, two working points have been established for the DNN:

- **Tight:** with a mis-identification probability of 0.1%;
- **Medium:** with a mis-identification probability of 1%.

Making use of these two working points, we considered three possible analysis selection strategies for the W' search and compared them with the cut and count selection:

- **Medium:** if the jet's DNN score is over the medium threshold;
- **Tight:** if the jet's DNN score is over the tight threshold;
- **Low purity:** a region which is filled with jets whose DNN score is over the medium threshold, but under the tight threshold, that could potentially be simultaneously fit with the Tight region without overlap.

Besides the signal, we considered two background processes: the $t\bar{t}$ pair production and the $W + jets$ associated production, and then we deployed the algorithm on such samples to assess its performance in a realistic scenario with background present. We then estimated a sensitivity for the analysis after applying a selection to the top quark and the b-jet stemming from the other leg of the W' decay. Finally, we evaluated that a search performed on data corresponding to a luminosity $L = 137 fb^{-1}$ with the medium working point of our DNN-based algorithm

would have a better sensitivity with respect to the same search performed with the cut and count selection rules defined in chapter 5. Further developments for this type of search may include a more complex statistical analysis which could use the combination of the selection with tight working point with the low-purity working point, and the addition of a potential loose working point, which could be designed by requesting a misidentification probability of 10% for the DNN-based top-tagging algorithm. Another improvement could be represented by the development of a top-tagging DNN which could reconstruct top quarks decaying in the channel $t \rightarrow b e^+ \nu_e$.

Bibliography

- [1] Zyla, P. et al. Review of Particle Physics. *PTEP*, 2020(8):083C01 (2020). doi:10.1093/ptep/ptaa104.
- [2] Englert, F. and Brout, R. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323 (1964). doi:10.1103/PhysRevLett.13.321.
- [3] Higgs, P. W. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509 (1964). doi:10.1103/PhysRevLett.13.508.
- [4] Aad, G., et al. Observation of a new particle in the search for the standard model higgs boson with the atlas detector at the lhc. *Physics Letters B*, 716(1):1–29 (2012). ISSN 0370-2693. doi:https://doi.org/10.1016/j.physletb.2012.08.020.
- [5] Chatrchyan, S., et al. Observation of a new boson at a mass of 125 gev with the cms experiment at the lhc. *Physics Letters B*, 716(1):30–61 (2012). ISSN 0370-2693. doi:https://doi.org/10.1016/j.physletb.2012.08.021.
- [6] Arnison, G., et al. Experimental observation of isolated large transverse energy electrons with associated missing energy at s=540 gev. *Physics Letters B*, 122(1):103–116 (1983). ISSN 0370-2693. doi:https://doi.org/10.1016/0370-2693(83)91177-2.
- [7] Banner, M., et al. Observation of single isolated electrons of high transverse momentum in events with missing transverse energy at the cern pp collider. *Physics Letters B*, 122(5):476–485 (1983). ISSN 0370-2693. doi:https://doi.org/10.1016/0370-2693(83)91605-2.
- [8] Arnison, G., et al. Experimental observation of lepton pairs of invariant mass around 95 gev/c² at the cern sps collider. *Physics Letters B*, 126(5):398–410 (1983). ISSN 0370-2693. doi:https://doi.org/10.1016/0370-2693(83)90188-0.
- [9] Wenninger, J. LHC status and performance. *PoS*, CHARGED2018:001 (2019). doi:10.22323/1.339.0001.
- [10] of the CMS Collaboration, T. T. G. The cms phase-1 pixel detector upgrade (2020).

- [11] Abdullah, M., et al. Probing a simplified w' model of $r(d^{(*)})$ anomalies using b tags, τ leptons, and missing energy. *Physical Review D*, 98(5) (2018). ISSN 2470-0029. doi:10.1103/physrevd.98.055016.
- [12] Boos, E., et al. Interference between w' and w in single-top quark production processes. *Physics Letters B*, 655(5-6):245–250 (2007). ISSN 0370-2693. doi:10.1016/j.physletb.2007.03.064.
- [13] Sullivan, Z. Fully differential w' production and decay at next-to-leading order in qcd. *Physical Review D*, 66(7) (2002). ISSN 1089-4918. doi:10.1103/physrevd.66.075011.
- [14] Sirunyan, A., et al. Combination of cms searches for heavy resonances decaying to pairs of bosons or leptons. *Physics Letters B*, 798:134952 (2019). ISSN 0370-2693. doi:10.1016/j.physletb.2019.134952.
- [15] Sirunyan, A., et al. Search for heavy resonances decaying to a top quark and a bottom quark in the lepton+jets final state in proton–proton collisions at 13. *Physics Letters B*, 777:39–63 (2018). ISSN 0370-2693. doi:10.1016/j.physletb.2017.12.006.
- [16] Aaboud, M., et al. Search for vector-boson resonances decaying to a top quark and bottom quark in the lepton plus jets final state in pp collisions at $s=13$ tev with the atlas detector. *Physics Letters B*, 788:347–370 (2019). ISSN 0370-2693. doi:10.1016/j.physletb.2018.11.032.
- [17] Search for a W' boson decaying to a top and a bottom quark at $\sqrt{s} = 13$ TeV. Technical Report CMS-PAS-B2G-20-005, CERN, Geneva (2020).
- [18] Sirunyan, A., et al. Particle-flow reconstruction and global event description with the cms detector. *Journal of Instrumentation*, 12(10):P10003–P10003 (2017). ISSN 1748-0221. doi:10.1088/1748-0221/12/10/p10003.
- [19] Cacciari, M., Salam, G. P., and Soyez, G. The anti-ktjet clustering algorithm. *Journal of High Energy Physics*, 2008(04):063–063 (2008). doi:10.1088/1126-6708/2008/04/063.
- [20] Verzetti, Mauro. Machine learning techniques for jet flavour identification at cms. *EPJ Web Conf.*, 214:06010 (2019). doi:10.1051/epjconf/201921406010.
- [21] Lista, L. *Statistical Methods for Data Analysis in Particle Physics*, volume 909. Springer (2016). ISBN 978-3-319-20175-7, 978-3-319-20176-4. doi:10.1007/978-3-319-20176-4.