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

Measurements of the CKM matrix elements in single top events at CMS with machine learning techniques

Relatore:
Dott. Alberto Orso Maria Iorio
Correlatore:
Prof. Fabio Ambrosino

Candidato:
Valeria D'Amante
Matr. N94000410

Anno Accademico 2018/2019

A Cosmo e Isabella,
Per aspera
ad astra.

Contents

Introduction

The Standard Model of particles physics (SM) is the current framework that describes particles and their interactions (the electromagnetic, weak and strong forces, with the gravitational force being excluded). It is verified with high precision for most processes involving elementary particles. All particles foreseen by the SM have been discovered.

The top quark is the most massive fundamental particle in the SM, with a mass of $m_t \approx 173\,\mathrm{GeV}/c^2$. It belongs to the third generation of quarks and, together with the bottom quark, to the corresponding weak-isospin doublet. The top quark has an unique phenomenology: it is the only quark that can be observed outside of bound states, as it decays to visible final states with a W boson and a down-type quark. In addition its decay products allow to measure the top quark features, like the mass or the spin information.
The top quark is a gateway to measure some of the fundamental parameters of the SM, among which there are some of the Cabibbo-Kobayashi-Maskawa (CKM) matrix elements. Such matrix does allow quark families to transform into each other. The top-related CKM matrix elements, $|V_{tb}|$, $|V_{ts}|$ and $|V_{td}|$, can be indirectly constrained from experiments in the B and K meson sectors, but those determinations rely crucially on model assumptions such as the existence of only three generations of quarks and the absence of particles beyond the SM in the loops.
The measured value of $|V_{tb}|$ (~ 0.999) when applying the standard model constraints is much larger than the $|V_{ts}|$ and $|V_{td}|$ ones (~ 0.041 and ~ 0.009 respectively). This leads to a strong suppression of processes involving the mixing between the third quark generation with the other two. The study of the top-quark production and decay allows to directly measure the CKM matrix elements $|V_{tq}|$, with q being a d, s or b quark.

The Large Hadron Collider is a circular proton-proton (pp) accelerator designed to reach a luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and a center-of-mass energy of $\sqrt{s} = 14\,\mathrm{TeV}$. It has been built in order to probe the SM and to search for clues of Beyond Standard Model (BSM) physics processes at its operating energy scale.

At LHC, top quarks are produced in proton-proton collisions through strong interaction, via gluon fusion that leads to $t\bar{t}$ pair production, and through the electroweak interaction, where the leading mechanism is the production of a single top quark through the exchange of a W boson between a sea quark and a valence quark in the t-channel. The dominant decay process for top quarks is to a W boson and a b quark, via an electroweak charged-current interaction.

In order to directly measure the CKM matrix element $|V_{tb}|$, two complementary approaches have been pursued, so far, by Tevatron and LHC experiments: measuring the branching fraction $\mathcal{B}(t \to Wb)$ in $t\bar{t}$ events, or exploiting the cross section of single-top-quark electroweak production. The former is sensitive to the CKM through the decay of top quarks, and is reliant on the unitarity hypothesis of the CKM matrix.
The method based on single-top-quark production, explored in this thesis, is less dependent on model assumptions, and allows the study of the top-quark coupling with the other quarks through the tWq vertex, with q being b, s or d quark, both in production and in decay, and with a reduced combinatorial background with respect to the $t\bar{t}$ channel. To disentangle the effects at the production and decay vertices, in past measurements from both Tevatron and LHC, $|V_{tb}|$ was extracted in the t-channel under the assumption that the values of $|V_{ts}|$ and $|V_{td}|$ are negligible, obtaining a global average of $|V_{tb}| = 1.019 \pm 0.025$ [1]. Nevertheless, if one considers $|V_{td}|$ and $|V_{ts}|$ too, a direct measurement of all the top quark related CKM matrix elements can be performed in this channel.

This thesis aims at a direct and model independent measurement of $|V_{tb}|$, $|V_{ts}|$ and $|V_{td}|$, by considering their respective contributions to t-channel production and top-quark decay. Three different signal processes are considered: the first one where $|V_{tb}|$ element is involved in both production and decay, and other two where a $|V_{ts}|$ or a $|V_{td}|$ element is involved either in production or decay. The processes with the $|V_{ts}|$ or $|V_{td}|$ elements both in production and in decay are not considered since they are highly suppressed.

The Compact Muon Solenoid (CMS) is one of the four main experiments hosted at LHC. It is a general-purpose detector, that has a broad physics programme ranging from studying the Standard Model to searching Beyond Standard Model physics.

This analysis is performed on LHC data collected during 2016 at the CMS experiment, at a centre-of-mass energy of 13 TeV, corresponding to an integrated luminosity of $35.92\,\text{fb}^{-1}$; the selected events are required to have one muon and two or three jets, one of which must pass the b-tagging requirement, i.e. it is identified as coming from a b quark, in the final state. Both the signals and the main background sources are simulated with the state-of-art simulation techniques too.

For each event, jets are sorted according to the probability to be coming from a top quark decay. A first approach to classify the jets exploits a likelihood ratio built from the 2D distributions (extracted from the simulated samples) of the variables that operate the best discrimination between the jets coming and not coming from top-quark decay. This is then compared with a more complex approach, based on different machine learning techniques (Boosted Decision Trees and Artificial Neural Network), that exploits information from many

variables in order to gain a better discrimination power. The classifier output quantifies the probability for each jet to come from a top-quark decay; the most powerful classifier is then chosen to sort the jets.
Three signal enriched regions are defined, according to their sensitivity to the different processes. A simultaneous maximum-likelihood fit to data is then performed in all the regions. The variables chosen for the fit in the signal regions are the outputs of a machine-learning algorithm, trained specifically for the events that have the $|V_{td}|$ in production or the $|V_{tq}|$, where q stays for d or s, in decay.

This is the first time a top-tagging approach based on a multi-variate classifier is applied in this type of measurement, allowing a dedicated $|V_{td}|$ enriched sample to be defined.

This thesis is organized in five chapters. Chapter 1 gives a brief introduction to the SM and the particles representation. In Chapter 2 the LHC accelerator and the CMS detector are described. The events selection and categorization are described in Chapter 3. Chapter 4 presents the different methods used for the jet ordering and their comparisons. Finally, in Chapter 5 are described the statistical methods adopted for the analysis and the results.

1 The top quark in the SM

1.1. The Standard Model of particle physics

The Standard Model (SM) of particle physics is a renormalizable quantum field theory that describes three out of the four fundamental interactions: the strong, the electromagnetic and the weak forces. The SM provides currently the best representation of fundamental particle physics, describing relativistic particles in terms of quantum fields. In a first instance, for the standard model to respect locality and to be Lorentz-invariant, i.e. it must be invariant under the combination of charge, parity, and time reversal symmetry (C, P, and T). The properties of fundamental interactions are successfully described by imposing additional, local symmetries to the lagrangian, according to the so-called gauge principle. In particular, the SM requires the lagrangian to be invariant under the non-abelian group of symmetry $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, where the letters "C", "L", and "Y" stand for colour charge, left-handed chirality and hypercharge, resepctively. $SU(2)_L \otimes U(1)_Y$ is the symmetry of electroweak interactions, as in the Glashow-Weinberg-Salam theory [2, 3, 4]. Finally, the spontaneous symmetry breaking mechanism allows to attribute a different mass to each particle in the standard model.

1.1.1 Particle representation in the SM

Particles are divided in two main categories: *fermions* are half-integer spin particles, which obey to the Fermi-Dirac statistics, and *bosons* are integer spin particles, which obey to the Bose-Einstein statistics.

There are three generations of fermions. Each fermion has an antiparticle, with opposite quantum numbers but same mass.
For each fermion generation there are five representations of the $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$

group:

$$Q_{L,i} = \begin{pmatrix} u_{L,i} \\ d_{L,i} \end{pmatrix}, \quad u_{R,i}, \quad d_{R,i},$$
$$\mathscr{L}_{L,i} = \begin{pmatrix} \nu_{L,i} \\ \ell_{L,i} \end{pmatrix}, \quad \ell_{R,i},$$

where i is the generation index, L and R refer the chirality (left-handed and right-handed, respectively). It is possible to classify the chiral component for the fermion fields, corresponding to:

- *left-handed* particles ($Q_{L,i}$, $\mathscr{L}_{L,i}$) organize in $SU(2)_L$ doublets and interact via weak force, with W^i bosons;
- *right-handed* particles are $SU(2)_L$ singlets and do not interact with W^i bosons.

Quarks ($Q_{L,i}$, $u_{R,i}$, $d_{R,i}$) are $SU(3)_C$ triplets and therefore interact via strong force, with gluons. Leptons ($\mathscr{L}_{L,i}$, $\ell_{R,i}$) are $SU(3)_C$ singlets, so they do not interact via strong force. The three quarks and leptons generation are listed in aTab. 1.1. It is important to notice that right-handed neutrinos do not interact at all since they do not have any charge, thus the neutrino right-handed singlet is omitted.

Generation	Lepton	Lepton Mass	Lepton Charge[e]	Quark	Quark Mass	Quark Charge [e]
first	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}$	$< 225\,\mathrm{eV}/c^2$ $0.511\,\mathrm{MeV}/c^2$	0 -1	$\begin{pmatrix} u \\ d \end{pmatrix}$	$2.16\,\mathrm{MeV}$ $4.67\,\mathrm{MeV}$	$+2/3$ $-1/3$
second	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}$	$< 0.19\,\mathrm{MeV}/c^2$ $0.106\,\mathrm{GeV}/c^2$	0 -1	$\begin{pmatrix} c \\ s \end{pmatrix}$	$93\,\mathrm{MeV}$ $1.27\,\mathrm{GeV}$	$+2/3$ $-1/3$
third	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}$	$< 18.2\,\mathrm{MeV}/c^2$ $1776.82\,\mathrm{MeV}/c^2$	0 -1	$\begin{pmatrix} c \\ s \end{pmatrix}$	$4.18\,\mathrm{GeV}$ $172.9\,\mathrm{GeV}$	$+2/3$ $-1/3$

Table 1.1: Lepton and quark generations and their properties. The neutrino mass limits are at 90%$C.L.$.

Fermions interactions can be described as the exchange of a integer-spin particle, a boson, that is the *mediator* of the interaction. From each of the group generators in the SM arises a vectorial field multiplet, whose components physically represent the interaction mediators, spin 1 bosons, that are listed in Tab. 1.2.

Section 1.1.2 describes the correspondence between the physical bosons with vectorial fields arising from the interactions term in the lagrangian density.

1.1.2 The SM lagrangian symmetries

A particle, represented by a field, is required to be invariant under a local phase transformation of this group of symmetry. The free lagrangian density for the generic spinor

Interaction	Mediator(s)		Spin	Charge
Electromagnetic	photon	γ	1	0
Strong	8 gluons	g	1	0
Weak	Positive charged W	W^+	1	$+1$
	Negative charged W	W^-	1	-1
	Neutral	Z^0	1	0

Table 1.2: Lepton and quark generations.

represented by the ψ Dirac field can be written as follows:

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

By imposing the theory to be invariant under local transformations of this group leads to interaction terms arising in the lagrangian density which physically correspond to the interaction with the mediator bosons; there is a boson gauge field for each of the group generator.

Quantum Electrodynamics The free lagrangian density in Eq. 1.1 is not invariant under a local phase transformation.
The electromagnetic interaction can be derived by imposing the requirement of invariance under $U(1)_{em}$ to a given field ψ:

$$\psi \rightarrow \psi' = e^{iq\theta(x)}, \tag{1.2}$$

where the parameter θ is a function of the coordinate in the space-time. This is a transformation under the symmetry group $U(1)$. By requiring this invariance, in the lagrangian density appears an interaction term:

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

where the last term represents lagrangian density for the electromagnetic field and $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$ is the electromagnetic tensor field. This is invariant under the transformation in Eq. 1.2 since the gauge field A_{μ} is invariant under the following transformation:

$$A_{\mu} \rightarrow A'_{\mu} = A_{\mu} - \partial_{\mu}\theta(x).$$

The A_{μ} fields physically represents the photon, the mediator boson of the electromagnetic force.

Quantum Chromodynamics The strong interaction between quarks arises by imposing the requirement of invariance under $SU(3)_c$ to a given field ψ:

$$\psi \rightarrow \psi' = e^{ig_s\vec{\theta}(x)\cdot\vec{T}}, \tag{1.3}$$

where $\vec{\theta}(x)$ are eight function of the 4-position in the space-time x, g_s is the coupling constant and $\vec{T} = T_\alpha$ are:

$$T_\alpha = \frac{1}{2}\lambda_\alpha, \tag{1.4}$$

where λ_a are 3×3 Gell-Mann matrices, the $SU(3)_c$ generators in the standard representation. By requiring the invariance under the transformation of the field in Eq. 1.3, in the lagrangian density appears an interaction term:

$$\mathcal{L}_{QCD} = i\bar{\psi}\gamma^\mu\partial_\mu\psi - m\bar{\psi}\psi - ig_s\bar{\psi}\gamma^\mu\lambda_a\psi G_{a,\mu} - \frac{1}{4}G_a^{\mu\nu}G_{a,\mu\nu},$$

where the last term represents lagrangian density for the strong field and:

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f^{abc} G^b_\mu G^c_\nu,$$

here, f^{abc} are the structural constants of the group. This is invariant under the local phase transformation if the gauge field G^a_μ is invariant under the following transformation:

$$G^a_\mu \rightarrow G'^a_\mu = G^a_\mu - \partial_\mu\alpha^a(x) - g_s f^{abc}\alpha^b G^c_\mu.$$

So there are eight gluon boson fields G^α linked to the $SU(3)_C$ generators.

Electro-weak interaction In 1961 Glashow, Salam and Weinberg developed an unified theory for the weak and the electromagnetic forces, by requiring the local invariance under the $SU(2)_L \otimes U(1)_Y$ symmetry group.

The lagrangian density that is invariant under the total group of symmetry $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$ can be written as follows:

$$\mathcal{L}_{kinetic} = i\bar{\psi}\gamma^\mu D_\mu\psi,$$

where there is no mass term, since the one in the free lagrangian density in Eq. 1.1 is not invariant. D_μ is the standard derivative is replaced by the covariant derivative:

$$D_\mu = \partial_\mu + ig_s G^a_\mu T_a + igW^b_\mu I_b + ig'B_\mu Y, \tag{1.5}$$

where:

- T_a are the $SU(3)_C$ generators: Eq. 1.4 shows the T_a in the standard representation;
- I_b are $SU(2)_L$ generators, that in the standard representation are given by $\frac{1}{2}\sigma_b$, where σ_b are 2×2 Pauli matrices;
- Y is the $U(1)_Y$ charge.

The three weak interaction boson fields, W^i, linked to the $SU(2)_L$ generators, and the single hypercharge boson field B, from the $U(1)_Y$ generator, do not correspond to physical observed particles, $W^\pm$ and Z^0. These ones will emerge after the spontaneous symmetry breaking.

1.1.3 Spontaneous symmetry breaking

No mass terms for bosons arise in Eq. 1.1, neither after the requirements of the SM symmetry group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$; nevertheless, weak bosons $W^{\pm}$ and Z^0 are expermentally found to have mass.

The way particles acquire mass is described by the *Spontaneous Symmetry Breaking* (SSB). It introduces a scalar complex doublet fields ϕ:

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

with a corresponding lagrangian density:

$$\begin{aligned} \mathcal{L}_{Higgs} &= (D^\mu \phi)^\dagger (D_\mu \phi) - V(\phi) \\ &= (D^\mu \phi)^\dagger (D_\mu \phi) - \frac{1}{2}\mu^2(\phi^\dagger \phi) - \frac{1}{4}\lambda(\phi^\dagger \phi)^2. \end{aligned} \tag{1.6}$$

D_μ is the covariant derivative in Eq. 1.5, $\lambda > 0$ and $\mu^2 < 0$.
The minimum of $V(\phi)$ is not unique, so the choice of a particular minimum for ϕ leads to to the spontaneous symmetry breaking of the gauge group:

$$\langle \phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \tag{1.7}$$

where $v = \sqrt{\frac{-\mu^2}{\lambda}}$. Developing the field around its Vacuum Expectation Value (VEV) v, it can be written in the form:

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H(x) \end{pmatrix}, \tag{1.8}$$

where $H(x)$ is the neutral scalar field corresponding to the Higgs boson. By substituting Eq. 1.8 in Eq. 1.6 one obtains the expression of the lagrangian desity for $H(x)$, the higgs field; it contains mass terms for the electro-weak bosons W^i, for the B boson and for the Higgs boson $H(x)$.
Finally, the physical bosons $W^{\pm}$ and Z^0 acquire mass:

$$\begin{aligned} M_{W^\pm} &= \frac{1}{2} g v, \\ M_{Z^0} &= \frac{1}{2} v \sqrt{g^2 + g'^2} = \frac{M_{W^\pm}}{\cos\theta_W}, \end{aligned}$$

where g and g' are the coupling constants of W^i and B, respectively, and θ_W is the Weinberg angle.
The A_μ field, which represents the photon, is massless, such as the gluons.

Fermion masses The fermion mass term introduced in Eq. 1.1

$$-m\bar{\psi}\psi = m\left(\bar{\psi_L}\psi_R + \bar{\psi_R}\psi_L\right)$$

is not invariant under $SU(2)_L \otimes U(1)_Y$ group because of the different behaviour of right-handed and left-handed chiral components of the fields. An invariant mass term arises by introducing the Yukawa coupling between ψ and Higgs fields:

$$\mathcal{L}_{Yukawa} = -\left(g^d_{ij}\bar{Q}_{L,i}\phi d_{R,j} + g^u_{ij}\bar{Q}_{L,i}(i\sigma^2\phi^\dagger)u_{R,j} + g^\ell_{ij}\bar{\mathcal{L}}_{L,i}\phi\ell_{R,j}\right) + h.c.\,, \tag{1.9}$$

where $g^{u,d,\ell}_{ij}$ are 3×3 complex matrices, i , j are generation indices. When ϕ acquires a vacuum expectation value as in Eq. 1.7, from Eq. 1.9 arise mass terms for fermions:

$$\mathcal{L}_{masses} = -M^\ell_{ij}\bar{\ell}_{L,i}\ell_{R,j} - M^d_{ij}\bar{d}_{L,i}d_{R,j} - M^u_{ij}\bar{u}_{L,i}u_{R,j},$$

where:

$$M^{\ell,d,u} = \frac{v}{\sqrt{2}}g^{f\ell,d,u}.$$

Mass eigenstates The physical observed states are obtained by diagonalizing M^f_{ij} with pairs of unitary matrices $\mathcal{U}^f_L$ and $\mathcal{U}^f_R$ such that:

$$\mathcal{U}^f_L M^f \mathcal{U}^{f,\dagger}_L = \mathcal{M}^f,$$

where $\mathcal{M}^f$ is the diagonal mass matrix. In this way mass eigenstates are expressed as a rotation of the weak interaction basis:

$$\begin{aligned}
u_{L,\alpha} &= (\mathcal{U}^u_L)_{\alpha,i}u_{L,i}, \quad u_{R,\alpha} = (\mathcal{U}^u_R)_{\alpha,i}u_{R,i},\\
d_{L,\alpha} &= (\mathcal{U}^d_L)_{\alpha,i}d_{L,i}, \quad d_{R,\alpha} = (\mathcal{U}^d_R)_{\alpha,i}d_{R,i},\\
\ell_{L,\alpha} &= (\mathcal{U}^\ell_L)_{\alpha,i}\ell_{L,i}, \quad \ell_{R,\alpha} = (\mathcal{U}^\ell_R)_{\alpha,i}\ell_{R,i},\\
\nu_{L,\alpha} &= (\mathcal{U}^\nu_L)_{\alpha,i}\nu_{L,i}.
\end{aligned}$$

The pedices with greek letters indicate mass eigenstates, whereas the roman letter are weak eigenstates pedices. In this basis the mass matrices are diagonal.

1.1.4 The CKM Matrix

Let's consider the complete lagrangian density:

$$\mathcal{L} = \mathcal{L}_{kinetic} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa}.$$

The charged current (CC) interaction term between quarks and $W^\pm_\mu = \frac{1}{\sqrt{2}}(W^1_\mu \pm iW^2_\mu)$ $SU(2)$ physical gauge bosons in terms of weak eigenstates is:

$$\begin{aligned}
\mathcal{L}_{CC} &= -\frac{g}{\sqrt{2}}\overline{Q}_{L,i}\gamma^\mu W^+_\mu Q_{L,i}\\
&= \frac{g}{\sqrt{2}}\overline{u}_{L,i}\gamma^\mu W^+_\mu d_{L,i} + h.c..
\end{aligned}$$

There are only the left-handed terms since $W^{\pm}$ preserve the $SU(2)_L$ symmetry, therefore they only couple with left-handed particles (in spite of Z^0 that, as a result of W^3 and B mixing, couples with right-handed chiral fields too). The same term can be expressed as function of mass eigenstates:

$$\begin{aligned}\mathcal{L}_{CC} &= -\frac{g}{\sqrt{2}}\overline{u}_{L,\alpha}(\mathcal{U}_L^u)_{\alpha,i}\gamma^\mu W_\mu^+(\mathcal{U}_L^u)^\dagger_{\beta,j}d_{L,\beta} + h.c. \\ &= -\frac{g}{\sqrt{2}}\overline{u}_{L,\alpha}\left[\mathcal{U}_L^u\mathcal{U}_L^{d\,\dagger}\right]_{\alpha\beta}\gamma^\mu W_\mu^+ d_{L,\beta} + h.c..\end{aligned} \tag{1.10}$$

The resulting matrix

$$V_{\alpha\beta} = \left[\mathcal{U}_L^u\mathcal{U}_L^{d\,\dagger}\right]_{\alpha\beta}$$

is a 3×3 unitary, not diagonal matrix, also known as the Cabibbo-Kobayashi-Maskawa (CKM) matrix [5, 6]. Equation 1.10 can be written as:

$$\mathcal{L}_{CC} = -\frac{g}{\sqrt{2}}(\overline{u}_L, \overline{c}_L, \overline{t}_L)\gamma^\mu W_\mu^+ V_{CKM}\begin{pmatrix} d_L \\ s_L \\ b_L\end{pmatrix} + h.c..$$

The weak interaction eigenstates are therefore linear combination of the mass ones, and the transformation matrix is:

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

Since this is a complex matrix, it can be written in terms of $2\times 3^2 = 18$ real parameters. The unitarity condition:

$$V^\dagger V = \mathbb{1} \tag{1.11}$$

adds 8 contraints, therefore reducing the number of free parameters to 9. Each of the six quark fields is invariant under phase shifting. It is possible to define a common phase for all of them, and the other parameters can be removed through a phase transformation without altering the physics.
The number of parameters then is reduced to four.

Mixing in the lepton sector Among the seven $\mathcal{U}_{L,R}^{u,d,\ell,\nu}$, only $\mathcal{U}_L^u$ and $\mathcal{U}_L^d$ have been considered, and they enter in the CKM matrix and rule quarks mixing and CP violation in CC interactions.
Therefore, as in the quark sector, one can define a mixing matrix for left-handed neutrinos and leptons. Nevertheless in standard model neutrinos are mass-less and $\mathcal{U}_L^\nu$ can be chosen such that $\mathcal{U}_L^\nu = \mathcal{U}_L^\ell$, hence the mixing matrix is now the unit. In the Standard Model extension that allows neutrinos to have mass, the lepton mixing matrix V_{PMNS} (known as PontecorvoMakiNakagawaSakata matrix [7]) leads to mixing and CP violation in the lepton interaction with $W^{\pm}$.

Standard parametrization In the standard parametrization, the CKM matrix parameters are three rotation angles (θ_{12}, θ_{23}, θ_{13}) ad a complex phase (δ) [8]:

$$V_{CKM} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ & -s_{12} & c_{12} \quad 0 \\ 0 & 0 & 1 \end{pmatrix}$$
$$= \begin{pmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ c_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{pmatrix},$$

where $c_{ij} = \cos\theta_{ij}$, $s_{ij} = \sin\theta_{ij}$. The three rotation angles can be chosen in the first quadrant, such that:

$$s_{ij}, c_{ij} \geq 0$$

With this parametrization the V_{ud}, V_{us}, V_{cb}, and V_{tb} parameters do not have imaginary component. The consequence of a non-zero phase factor in the mixing matrix is the CP-Violation (CPV). The CP invariance implies $V = V^*$. It is experimentally known that the diagonal elements' modulus is numerically close to unity and the mixing angles satisfy the hierarchy:

$$s_{13} \ll s_{23} \ll s_{12} \ll 1. \tag{1.12}$$

Wolfenstein parametrization Another convenient parametrization of the CKM matrix is due to Wolfenstein, where the hierarchy in Eq. 1.12 is more evident. The CKM matrix is expressed in terms of four real parameters, λ, A, ρ, η, defined by the following expressions [9]:

$$s_{12} = \lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}},$$
$$s_{23} = A\lambda^2 = \lambda|\frac{V_{cb}}{V_{us}}|,$$
$$s_{13}e^{i\delta} = A\lambda^3(\rho + i\eta).$$

The CKM matrix is written as an expansion in terms of λ:

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

One intuitive feature of this parametrisation is that the farther a parameter is from the diagonal, the higher the power of λ is, thus giving the "hierarchy structure" of the matrix. The SM does not provide any explanation to this structure. The η parameter is linked to the CP-violating phase: if it was zero, the CKM matrix would be completely real.

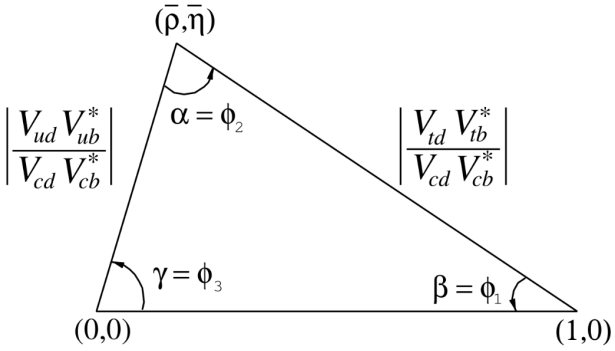


Figure 1.1: Unitarity triangle.

Unitarity triangles The Eq. 1.11 imposes the orthogonality among rows and columns, resulting in:

$$\sum_i V_{ij}V_{ik}^* = \delta_{jk},$$
$$\sum_i V_{ij}V_{kj}^* = \delta_{ik}.$$

These relationship can be represented in the complex plane as *unitarity triangles*. The most relevant unitarity triangle is given by the relationship:

$$V_{ud}V_{ub}^* + V_{cd}V_{cb}^* + V_{td}V_{tb}^* = 0$$

divinding all the terms by $V_{cd}V_{cb}^*$ since it is the best experimentally known term, one gets:

$$\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*} + 1 + \frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*} = 0.$$

another alternative parametrization ρ and η is:

$$\overline{\rho} + i\overline{\eta} = -\frac{V_{ud}}{V_{ub}^*}V_{cd}V_{cb}^*,$$

because with this definition, the vertices of the triangle are exactly $(0,0)$, $(1,0)$ and $(\overline{\rho},\overline{\eta})$, as it is shown in Fig. 1.1 In the Wolfenstein parametrization, the η parameter quantifies the height of the unitarity triangle; if $\eta = 0$ the triangle becomes segments on the real axis.

Current experimental status The CKM matrix elements are fundamental parameters for the standard model, so their determination is object of many experiments. Their precise measurements can put strong constraints on Beyond Standard Model (BSM) theories. They are measured through processes where the relevant quark couplings appear at tree level process in order to directly extract the value of $|V_{ij}|$. In many cases such processes are suppressed or hard to measure, so they are indirectly measured by the products $|V_{ij}V_{ik}|$.

A global fit on the SM parameters is performed in order to extract the elements of the CKM matrix. In figure Fig. 1.2 are shown many measurements conveniently represented in the $\overline{\rho} - \overline{\eta}$ plane. The 95% CL regions overlap consistently around the global fit region. The fit results for the Wolfenstein parameters gives[1] [10]:

$$\lambda = 0.224747^{+0.000254}_{-0.000059}, \qquad A = 0.8403^{+0.0060}_{-0.0201}, \qquad \bar{\rho} = 0.1577^{+0.0096}_{-0.0074}, \qquad \bar{\eta} = 0.3493^{+0.0095}_{-0.0071}.$$

The fit results for the magnitudes of all nine CKM elements are [10]:

$$|V_{CKM}| = \begin{pmatrix} 0.974410^{+0.000014}_{-0.000058} & 0.224745^{+0.000254}_{-0.000059} & 0.003746^{+0.000090}_{-0.000062} \\ 0.224608^{+0.000254}_{-0.000060} & 0.973526^{+0.000050}_{-0.000061} & 0.04240^{+0.00030}_{-0.00115} \\ 0.008710^{+0.000086}_{-0.000246} & 0.04169^{+0.00028}_{-0.00108} & 0.999093^{+0.000049}_{-0.000013} \end{pmatrix}.$$

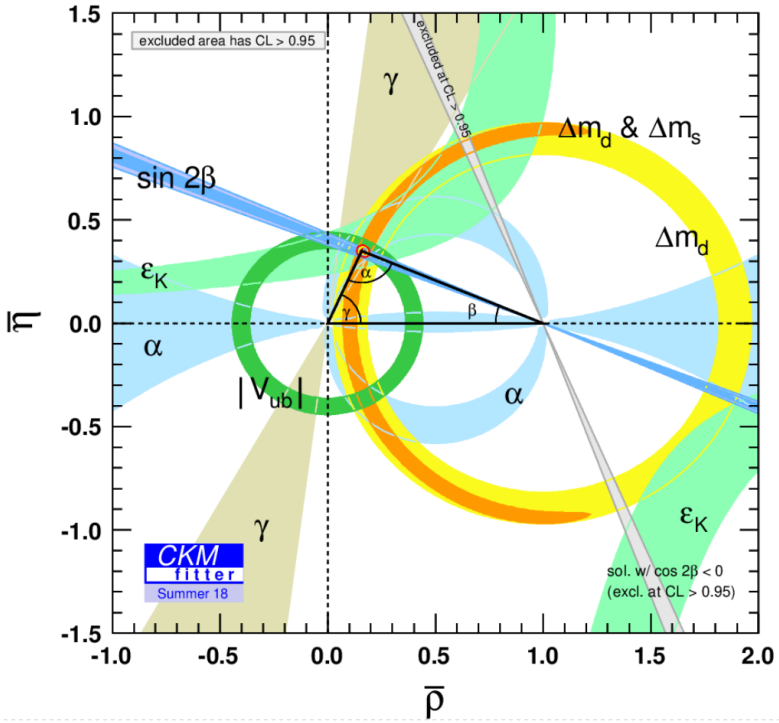


Figure 1.2: Global fit to the SM parameters. The shaded areas have 95% CL [10].

[1] Here, the unitarity constraint is applied.

1.2. The Top quark

The top quark was discovered by the DØ and CDF experiments at the Fermilab Tevatron $p\overline{p}$ collider [11]. After the Higgs discovery in 2012, this is one of the most recent SM proof, and gives a powerful confirmation to the three-generation structure of fermions.

Together with the bottom quark, the top quark belongs to the third quark generation, and to the corresponding weak-isospin doublet. It is the heaviest elementary particle in the SM, with a mass of approximately $173\,\mathrm{GeV}/c^2$. Such a high mass value for a fundamental fermion drives a peculiar phenomenlogy for the top quark:

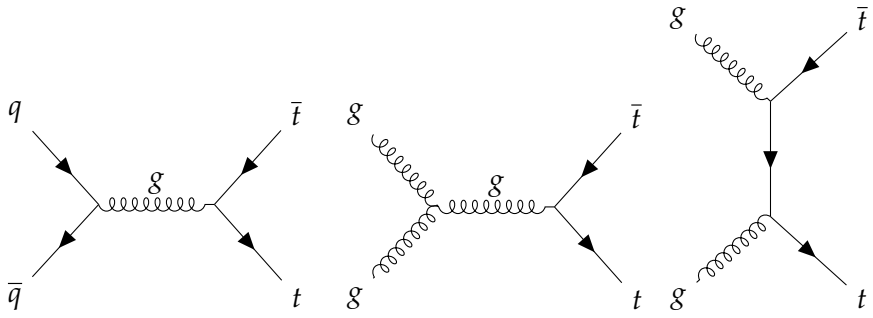
- it is the only quark with a Yukawa coupling value of unity order: according to the description of the origin of fermion masses provided by the SM (see Sec. 1.1.3), the quark mass values are related to the strength of the Yukawa coupling with the Higgs boson;
- the large mass value of the top quark makes it sensitive to BSM physics;
- the top quark has a lifetime of order 10^{-25} s, a couple of orders of magnitude shorter than the hadronization process timescale, that arranges colored quarks into color-neutral hadrons. This means that no hadrons can ever contain the top quark, since it decays before it can hadronize, giving us an unique opportunity to study a "naked" quark;
- it is the only quark kinematically allowed to weakly decay into a real W boson (and a down-type quark), since its mass is greater than the W boson's one;
- due to the V-A nature of the electroweak interaction, the angular distributions of the decay products depend the top-quark spin orientation. This information can be exploited to study the structure of the tWb vertex.

An accurate knowledge of its properties (mass, couplings, production cross sections, decay branching ratios, etc.) can bring key information on fundamental interactions at the electroweak symmetry-breaking scale and beyond [1].

1.2.1 Top quark production

Top-antitop ($t\bar{t}$) pair production In hadron collisions, top quarks are dominantly produced in pairs through $q\overline{q} \to t\bar{t}$ and $gg \to t\bar{t}$ processes at leading order in QCD, as shown in Fig. 1.3. Approximately 85% of the production cross section at the Tevatron ($p\overline{p}$ at 1.96 TeV) is from qq annihilation, with the remainder from gluon-gluon fusion, while at LHC (pp) about 90% of the production is from the latter process at $\sqrt{s} = 14\,\mathrm{TeV}$.

Predictions for the top-quark production total cross sections are available at next-to-next-to leading order (NNLO) [12, 13], also including next-to-next-to-leading-log (NNLL) soft gluon re-summation. Table 1.3 summarizes the resulting theoretical prediction of the top-quark pair cross-section at NNLO+NNLL accuracy for different centre-of-mass energies



(a) $q\bar{q} \to t\bar{t}$ (b) $gg \to t\bar{t}$ in t-channel and s-channel

Figure 1.3: Top pair production Feynman diagrams at tree level.

at Tevatron and LHC, assuming a top-quark mass of 173.3 GeV/c^2 (close to the Tevatron + LHC average), and of 172.5 GeV/c^2, respectively [12].

Accelerator	$\sqrt{s}$ [TeV]	$\sigma_{t\bar{t}}$ [pb]
Tevatron ($p\overline{p}$)	1.96	$7.16^{+0.11+0.17}_{-0.20-0.12}$
LHC (pp)	7	$177.3^{+4.6+9.0}_{-6.0-9.0}$
LHC (pp)	8	$252.9^{+6.4+11.5}_{-8.6-11.5}$
LHC (pp)	13	$831.8^{+19.8+35.1}_{-29.2-35.1}$
LHC (pp)	14	$984^{+23.2+41.3}_{-34.7-41.3}$

Table 1.3: Theoretical predictions at NNLO+NNLL for $t\bar{t}$ pair production cross sections at different centre-of-mass energy and different hadron colliders. The first uncertainty is from factorization and renormalization scales set in the calculation and the second is from parton distribution functions.

Single-top-quark production It was first observed in 2009 by DØ and CDF at Tevatron. This is a charged current electroweak process that involves the tWq vertex in the production of the top quark and in its decay, where q is a down-type quark, and negligible contributions from Flavor-Changing Neutral Currents (FCNC).
In the SM, single-top-quark production mostly proceeds via three mechanisms:

- the ***t*-channel** mode;
- the production of top quark in association with a W boson, **tW** mode;
- the ***s*-channel** mode.

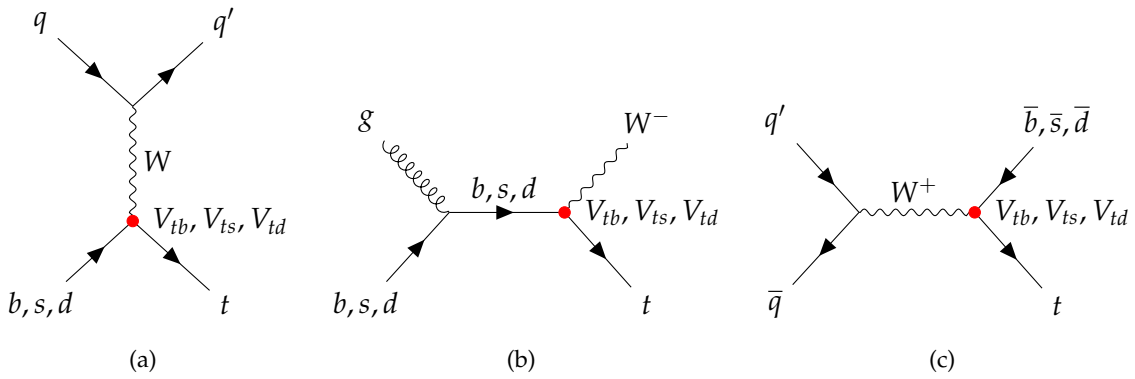


Figure 1.4: Feynman diagrams for single top-quark production in t-channel (left), tW (middle) and s-channel (right) modes in the presence of non-zero $|V_{tb}|$, $|V_{ts}|$ and $|V_{td}|$ values.

Figure 1.4 shows the tree level Feynman diagrams for t-channel, tW and s-channel in the five-flavor scheme, i.e., considering all quarks up to the b as incoming partons.

Furthermore, top-quark production in association with a Z boson, **tZ**, has been observed [14] and evidence of top-quark production in association with a photon, **tγ**, recently appeared [15]; those processes have predicted cross section of $\sigma(tZq \rightarrow t\ell\overline{\ell}q) = 94.2 \pm 3.1$ fb and $\sigma_{t\gamma} = 2.95^{+0.13+0.03}_{-0.13-0.03}$ pb.

At the Tevatron, t-channel and s-channel production differ by roughly a factor 2 (with the t-channel being the largest), while the tW production cross section is negligible. At the LHC, the t-channel mode has the largest cross section followed by tW while the s-channel production has a small cross section. The cross section for t-channel single-top-quark production therfore differs from Tevatron to LHC: this is due of course to different centre-of-mass energy, but an important difference is due to parton distribution functions and could be understood by observing the left diagram of Fig. 1.4: the incoming quark q must be u, c, $\overline{d}$, or $\overline{s}$ (d, s, $\overline{u}$, or $\overline{c}$) in order to create a top (anti-top) quark(anti-quark). Since the LHC is a pp collider and the valence composition of the proton is uud, more top quarks are produced than anti-top anti-quarks. Nevertheless, the inclusive single (top+anti-top) cross section for the t-channel mode is the same for Tevatron and LHC.
As the partonic collision happens at large values of Bjorken x, where valence quarks dominate over the "sea" quarks, the charge asymmetry observable, defined as:

$$R_t = \frac{\sigma^{top}_{t-channel}}{\sigma^{anti-top}_{t-channel}},$$

gets values close to two. Similar considerations apply to s-channel production. The value of R_t can be enhanced significantly when $|V_{td}|$ is non-zero due to the valence-valence quark interaction in the diagram in Fig. 1.4 (left). In the tW case, the SM predicts complete charge symmetry when $|V_{td}| = 0$, as can be understood from the middle panel of Fig. 1.4, since there are no valence s and b quarks. Figure 1.5 shows the measured and the predicted cross section at Tevatron and LHC.

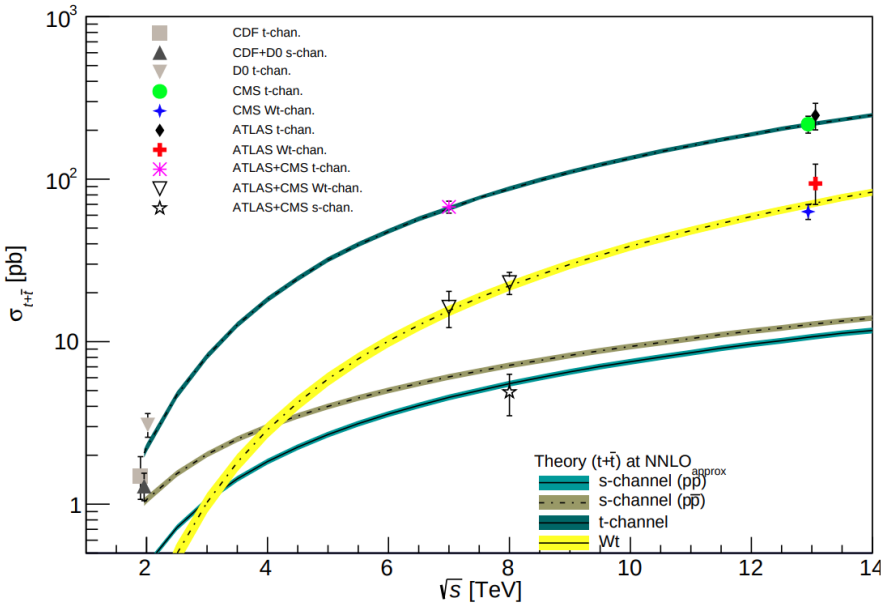


Figure 1.5: Measured and predicted single-top-quark production cross sections from Tevatron ($p\overline{p}$ collisions) to LHC (pp collisions) [1].

The predicted cross section for the single-top-quark t-channel, s-channel and tW associated productions are shown in Tab. 1.4, Tab. 1.5, and Tab. 1.6, respectively. In particular, the tW process is considered to have equal propotion for t and $\bar{t}$, thus in Tab. 1.6 only inclusive $\sigma_{t+\bar{t}}$ cross section is reported.

Accelerator	$\sqrt{s}$ [TeV]	σ^{t}_{t-ch} [pb]	$\sigma^{\bar{t}}_{t-ch}$ [pb]	$\sigma^{t+\bar{t}}_{t-ch}$ [pb]
Tevatron ($p\overline{p}$)	1.96	$\frac{1}{2}\sigma^{t+\bar{t}}_{t-ch}$	$\frac{1}{2}\sigma^{t+\bar{t}}_{t-ch}$	$2.08^{<0.01+0.12}_{-0.04-0.12}$
LHC (pp)	7	$41.80^{+1.24+1.28}_{-0.82-1.28}$	$22.02^{+0.67+1.08}_{-0.43-1.08}$	$63.89^{+1.92+2.19}_{-1.25-2.19}$
LHC (pp)	8	$54.87^{+1.64+1.60}_{-1.09-1.60}$	$29.74^{+0.92+1.39}_{-0.59-1.39}$	$84.69^{+2.56+2.76}_{-1.68-2.76}$
LHC (pp)	13	$136.02^{+4.09+3.52}_{-2.92-3.52}$	$80.95^{+2.53+3.18}_{-1.71-3.18}$	$216.99^{+6.62+6.16}_{-4.64-6.16}$
LHC (pp)	14	$154.76^{+4.66+3.96}_{-3.39-3.96}$	$93.28^{+2.92+3.58}_{-2.01-3.58}$	$248.09^{+7.58+6.98}_{-5.40-6.98}$

Table 1.4: Theoretical predictions at NNLO+NNLL for single-top-production cross sections in t-channel at different centre-of-mass energies and different hadron colliders. The first uncertainty is from factorization and renormalization scales set in the calculation and the second one is from parton distribution functions.

Accelerator	$\sqrt{s}$ [TeV]	σ^{t}_{s-ch} [pb]	$\sigma^{\bar{t}}_{s-ch}$, [pb]	$\sigma^{t+\bar{t}}_{s-ch}$ [pb]
Tevatron ($p\overline{p}$)	1.96	$\frac{1}{2}\sigma^{t+\bar{t}}_{s-ch}$	$\frac{1}{2}\sigma^{t+\bar{t}}_{s-ch}$	$1.046^{+0.002+0.060}_{-0.010-0.056}$
LHC (pp)	7	$2.76^{+0.08+0.08}_{-0.07-0.08}$	$1.57^{+0.04+0.07}_{-0.04-0.07}$	$4.29^{+0.12+0.14}_{-0.10-0.14}$
LHC (pp)	8	$3.34^{+0.09+0.09}_{-0.08-0.09}$	$1.90^{+0.05+0.08}_{-0.04-0.08}$	$5.24^{+0.15+0.16}_{-0.12-0.16}$
LHC (pp)	13	$6.35^{+0.18+0.14}_{-0.15-0.14}$	$3.97^{+0.11+0.15}_{-0.09-0.15}$	$10.32^{+0.29+0.27}_{-0.24-0.27}$
LHC (pp)	14	$6.98^{+0.20+0.15}_{-0.16-0.15}$	$4.42^{+0.13+0.16}_{-0.10-0.16}$	$11.39^{+0.32+0.29}_{-0.26-0.29}$

Table 1.5: Theoretical predictions at NNLO+NNLL for single-top-production in s-channel cross sections at different centre-of-mass energies and different hadron colliders. The first uncertainty is from factorization and renormalization scales set in the calculation and the second is from parton distribution functions.

Accelerator	$\sqrt{s}$ [TeV]	σ_{tW} [pb]
Tevatron ($p\overline{p}$)	1.96	—
LHC (pp)	7	$15.74^{+0.40+1.10}_{-0.40-1.14}$
LHC (pp)	8	$22.37^{+0.60+1.40}_{-0.60-1.40}$
LHC (pp)	13	$71.7^{+1.8+3.4}_{-1.8-3.4}$
LHC (pp)	14	$84.4^{+2.0+3.0}_{-2.0-4.8}$

Table 1.6: Theoretical predictions at NNLO+NNLL for single-top-quark production in tW channel cross sections at different centre-of-mass energies and different hadron colliders. The first uncertainty is from factorization and renormalization scales set in the calculation and the second is from parton distribution functions.

1.2.2 Top quark decay

At tree level, the dominant decays of the top quark expressed by the process $t \to W^{+}q_d$, where q_d indicates a down-type quark. Assuming $|V_{tb}| \gg |V_{ts}|, |V_{td}|$, and neglecting terms of order m_q^2/m_t^2, α_s^2 and $(\alpha_s/\pi)M_W^2/m_t^2$, the SM predicted (NLO) is [16]:

$$\Gamma_t = \frac{G_F m_t^3}{8\pi\sqrt{2}}\left(1-\frac{M_W^2}{m_t^2}\right)^2\left(1+2\frac{M_W^2}{m_t^2}\right)\left[1-\frac{2\alpha_s}{3\pi}\left(\frac{2\pi^2}{3}-\frac{5}{2}\right)\right], \tag{1.13}$$

where m_t refers to the pole mass. Taking $m_t = 173.3\,\mathrm{GeV}/c^2$ and $\alpha_s(M_Z) = 0.118$, one gets $\Gamma_t = 1.35\,\mathrm{GeV}/c^2$. With its correspondingly short lifetime of $\approx 0.5 \times 10^{-24}$ s the top quark is expected to decay before top-flavored hadrons or tt-quarkonium-bound states can form [17]. Since the decay of the top quark occurs through the $t \to W^+ q$ process, there are two different decay chains:

- $t \to W^+ q_d \to q_d \bar{\ell} \nu_\ell$ where the W^+ decays in a lepton and a neutrino. The inclusive branching ratio is $\sim 32\%$
- $t \to W^+ q_d \to q_d \overline{q'_d} \nu_{q''_u}$, where the W^+ decays hadronically. The inclusive branching ratio is $\sim 68\%$.

For the $t\bar{t}$ pair there are all the combinatorial possibilities, and are defined as follows:

- the lepton+jets channel is: $t\bar{t} \to W^+ q_d W^- \overline{q'_d} \to \bar{\ell} \nu_\ell q_d \overline{q''}_u q'''_d \overline{q'}_d + q''_u \overline{q'''}_d q_d \ell \overline{\nu}_\ell \overline{q'}_d$ has one of the two W boson leptonically decaying, and the other one hadrinically decays. The branching ratio of this channel is 43.8% ;
- the dilepton channel has both of the W boson leptonically decaying: $t\bar{t} \to W^+ q_d W^- \overline{q'_d} \to \bar{\ell} \nu_\ell q_d \ell' \overline{\nu}_\ell \overline{q'}_d$. The branching ratio is 10.5%
- the all-jets channel $t\bar{t} \to W^+ q_d W^- \overline{q'_d} \to \overline{q''} q''' q_d q^{\mathrm{iv}} \overline{q^{\mathrm{v}}} \overline{q'}_d$ where both the W bosons hadronically decay.

Here, q_i indicates a i-th type quark, and q is of course a down-type quark.

W boson helicity in the decay The SM states that the top quark has the $(V-A)$ CC weak interactions:

$$-i\frac{g}{\sqrt{2}} V_{tq} \gamma^\mu \frac{1}{2}(1-\gamma^5),$$

as all the other fermions. In the SM, the fraction of top-quark decays to longitudinally polarized W bosons is proportional to its Yukawa coupling and hence enhanced with respect to the weak coupling. Since the W is on-mass-shell, its helicity is quite different with respect to the CC interactions where the W is off-mass-shell. Since in the CC interactions only the left-handed fermions are involved, in the top-quark rest frame the $+1$ value is prohibited for the W helicity. The W helicity comes from a superposition of helicity states with $\lambda = 0$ or $\lambda = -1$. The ratio among these two components is:

$$\frac{\Gamma(t \to b_L W_{\lambda=0})}{\Gamma(t \to b_L W_{\lambda=-1})} \simeq \frac{m_t^2}{2M_W^2} = x$$

In the SM, at tree level, the fraction of longitudinal polarizated ($\lambda = 0$) W in the top-quark rest frame, $\mathcal{F}_0$, is expected to be [18]:

$$\mathcal{F}_0^{SM} \approx x/(1+x),\ x = m_t^2/2M_W^2\ (\mathcal{F}_0^{SM} \sim 70\% \text{ with } m_t = 175\,\mathrm{GeV}/c^2).$$

Fractions of left-handed, right-handed, or longitudinal W bosons are denoted as $\mathcal{F}_-$, $\mathcal{F}_+$, and $\mathcal{F}_0$ respectively.
In the SM, $\mathcal{F}_-$ is expected to be $\approx 30\%$ and $\mathcal{F}_+ \approx 0\%$. The study of the angular distribution of the top.quark decay products provides a tool for a direct proof of the $(V-A)$ structure of the tWb coupling. The Tevatron and the LHC experiments use various techniques to measure the helicity of the W boson in top-quark decays in $t\bar{t}$ production, both in the lepton+jets and in dilepton channels. All results are in agreement with the SM expectations [1].

1.2.3 Top quark properties

Mass As the heaviest elementary particle in the standard model, the top quark plays an important role in the precision fits of the SM and in loop corrections, where contributions depending on quark masses arise: for example, M_W measurement gets corrections that are proportional to m_t^2 and $\ln m_H$. In flavor physics, many observables contain quadratic terms of m_t/M_W. On the other hand, the top quark has the potential to provide insights into physics beyond the SM (BSM). Many BSM models introduce changes within the top quark sector, which can be constrained by precise measurements of the cross sections and properties of top-quark production channels [19].

The most precise measurement among top quark properties is its mass: it has been measured in the lepton+jets, the dilepton, and the all-jets channel[2] by CDF and DØ experiments at Tevatron as well as by CMS and ATLAS experiments at LHC.

The lepton+jets channel yields the most precise single measurements because of good signal to background ratio (in particular thanks to b-tagging) and the presence of a single neutrino in the final state. The momentum of a single neutrino can be reconstructed (up to a quadratic ambiguity) via the missing transverse energy measurement and applying the constraint of W boson mass.

In the di-lepton channel, the signal to background is typically very good, but reconstruction of the mass is non-trivial because there are two neutrinos in the final state, yielding a kinematically unconstrained system. A variety of techniques have been developed to handle this problem.

The most recise measurements of the top quark mass are [1]:

$$m = 172.9 \pm 0.4\,\mathrm{GeV}/c^2 \text{ from direct measurements}$$
$$m = 160^{+5}_{-4}\,\mathrm{GeV}/c^2 \text{ from cross-section measurements}$$
$$m = 173.1 \pm 0.9\,\mathrm{GeV}/c^2 \text{ Pole from cross-section measurements}$$

Width Early measurements made at CDF [20] and CMS [21] established confidence-level intervals for the width, but do not have the sensitivity to make a direct measurement because

[2]These definitions arise from the different decay mechanisms of the W bosons from the tWq vertices, see Sec. 1.2.2.

the experimental resolutions are much larger than the width itself. A recent measurement from ATLAS, using 139 fb^{-1} at $\sqrt{s} = 13\,\text{TeV}$ found [22]:

$$\Gamma_t = 1.9 \pm 0.5\ \text{GeV}/c^2,$$

that is the most precise measurements up to date.

Electric Charge The top quark is the only quark whose electric coupling has not been measured through production at threshold in e^+e^- collisions. Since the CDF and DØ analyses on top-quark production did not associate the b, $\overline{\text{b}}$, and $W^\pm$ uniquely to top or anti-top anti-quarks, decays such as $t \to W^+\bar{b}$,$\bar{t} \to W^-b$ were not excluded. A charge 4/3 quark of this kind could be consistent with current electroweak precision data. Nevertheless, measurements of Tevatron [23, 24] and LHC experiments [25, 26] exclude these scenarios with high confidence level. The top-quark coupling to a photon provides the opportunity to directly measure the electromagnetic charge of the top quarks. By measuring the associated production cross section of a top-antitop pair and a photon ($t\bar{t} + \gamma$), the coupling of the top quark and the photon is probed [27, **?**]. Any deviation of the measured cross section value from the SM prediction would be an indication of BSM physics, such as the production of an exotic quark with electric charge of 4/3, or a top quark with an anomalous electric dipole moment [28, 29]. A different coupling would change directly the cross section, therefore cross section measurements in this channel are a direct proof of the top quark charge. Measurements of the production cross section of $t\bar{t} + \gamma$ have been performed by the CDF Collaboration at the Tevatron using pp collisions at $\sqrt{s} = 1.96\,\text{TeV}$ [30], and by the ATLAS Collaboration at the LHC using pp collisions at $\sqrt{s} = 7\,\text{TeV}$ [31] and $\sqrt{s} = 8\,\text{TeV}$ [32]. These results are in agreement with the SM predictions within uncertainties.

Spin and polarization The top quark decays before its spin can be flipped by the strong interaction. Thus its polariztion is directly observable via the angular distribution of its decay products and it is possible to define and measure observables sensitive to the top-quark spin and its production mechanism. Although the top and anti-top quarks produced by strong interactions in hadron collisions are essentially unpolarized, the spins of t and $\bar{\text{t}}$ are correlated. Let's consider the process

$$q\bar{q} \to t\bar{t}.$$

In the ultra-relativistic limit ($E \gg m$), helicity conservation imposes q and $\bar{q}$ to have opposite helicities. Thus, t and $\bar{t}$ can be produced with either the opposite or the same helicity values. The direction of the top-quark spin is 100% correlated to the angular distributions of the down-type fermions (charged leptons or d-type quarks) in the decay.

At Tevatron the dominating process for $\text{t}\bar{\text{t}}$ production is $\text{q}\overline{\text{q}} \to \text{t}\bar{\text{t}}$, whereas at LHC it is $gg \to \text{t}\bar{\text{t}}$, hence the basis choices that optimize the spin correlation analyses are different. At Tevatron the off-diagonal basis is used since it optimizes the correlations, while at LHC the

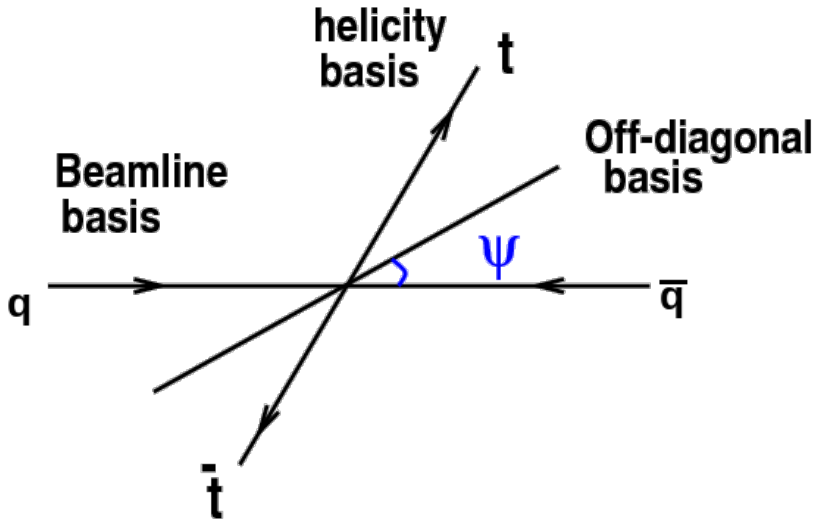


Figure 1.6: Definitions of the off-diagonal, helicity and beamline basis, used in top-quark spin correlations studies.

more convenient basis is the helicity one, where the quantization axis is along the top-quark direction in the $t\bar{t}$ rest frame. Figure 1.6 shows a schematic view of off-diagonal and helicity basis with respect to the beamline basis. In the off-diagonal basis the spin quantization axis makes an angle ψ with the direction beam, that can be expressed in function of the diffusion angle:

$$\tan\psi = \frac{\beta^2 \sin\theta\cos\theta}{1-\beta^2\sin^2\theta},$$

where β is the top-quark (and anti-top anti-quark) velocity in the centre-of-mass ($q\bar{q}$) rest frame. In the limit $E \gg m$ the off-diagonal basis coincides with the helicity one.

A very sensitive observable is the joint angular distribution:

$$\frac{1}{\sigma}\frac{d^2\sigma}{d\cos\theta_+\, d\cos\theta_-} = \frac{(1 + B_+\cos\theta_+ + B_-\cos\theta_- + \kappa\cos\theta_+\cos\theta_-)}{4}, \tag{1.14}$$

where θ_+ and θ_- are the angles of the daughters in the top-quark rest frame with respect to a particular spin quantization axis (assumed here to be the same for both the angles), is a very sensitive observable. The maximum value for the parameter κ, is found in the offdiagonal basis [33]. The coefficients B_+ and B_- are near zero in the SM because the top quarks are unpolarized in $t\bar{t}$ production.

In electroweak t-channel single-top-quark production, the standard model predicts that produced top quarks are highly polarized along the direction of the momentum of the spectator quark [34, 35]. However, new physics models could also lead to a depolarisation in production by altering the coupling structure [36, 37, 38, 39]. In this channel is possible to

define the spin asymmetry as follows:

$$A_X = P_t \cdot \alpha_X = \frac{N(\uparrow) - N(\downarrow)}{N(\uparrow) + N(\downarrow)},$$

is used to probe the coupling structure, where Pt represents the top-quark polarization in production and α_X denotes the degree of the angular correlations of one of its decay products, X, with respect to the spin of the top quark, the spin analyzing power. The variables $N(\uparrow)$ and $N(\downarrow)$ are defined, for each top-quark decay product from the decay chain $t \to qW \to q\ell\nu$, as the number of instances in which that decay product is aligned or anti-aligned, respectively, relative to the direction of the recoiling spectator quark momentum [40].

2 The CMS experiment at the LHC

2.1. The Large Hadron Collider

The Large Hadron Collider (LHC), hosted at CERN, is the largest proton-proton and heavy ions (Pb) accelerator and collider. It is located in a 27 km circular tunnel underground, at an average depth of 100 m across the Swiss-French border near Geneva. Its centre-of-mass design energy for the protons is 14 TeV,for heavy ions is 2.4 GeV, for a peak luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and of $10^{27}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, respectively.

The proton source is a bottle of hydrogen gas, from where they are extracted; to reach centre-of-mass energy for collisions, protons go through some acceleration stages that are schematized in Fig. 2.1:

- protons are first accelerated up to 50 MeV by a 36 m **Linear Accelerator** (**LINAC 2**); they are pulsed from the hydrogen bottle;
- then they enter the **Proton Synchrotron Booster** (**BOOSTER**), the next step in CERN's accelerator chain, a circular accelerator of 50 meter diameter, and reach 1.5 GeV;
- the beams are then injected in the **Proton Synchrotron** (**PS**), a circular accelerator whose circumference is 628 m which accelerates them to 25 GeV;
- the last acceleration stage before the injection in the LHC is through the **Super Proton Synchrotron** (**SPS**), a 6.9 km accelerator, where protons reach 450 GeV.

Finally, the beams are injected in the Large Hadron Collider (**LHC**), where they circulate in opposite directions and reach the energy for the collisions, that during Run 2 (2015-2018) was 6.5 TeV per beam; they are accelerated by superconductive radio-frequency (RF) cavities. Bemas are curved by 1232 superconductive Nb-Ti dipole magnets, and focused by 392 quadrupole magnets and other magnets whose purpose is to correct the field geometry (sextupoles, octupoles). The LHC magnet system exploits the well-proven technology based on Nb-Ti Rutherford cables and cools the magnets to a temperature below 2 K thanks to the use of the superfluid helium, and operates with fields above 8 T [41]. The beam design

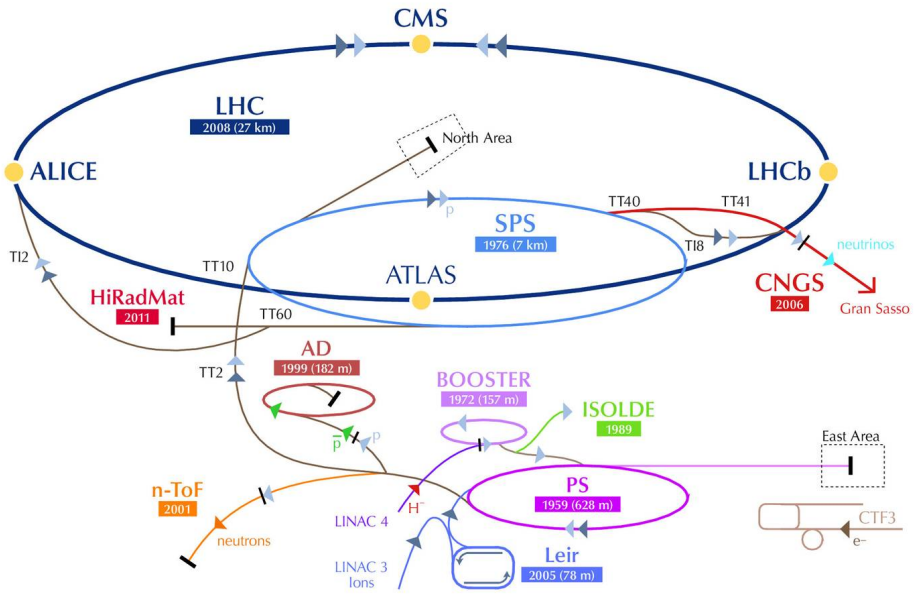


Figure 2.1: Schematic representation of CERN accelerator complex.

energy for LHC is 7 times greater the Tevatron one, for a 100 gain factor in luminosity. The LHC has two beam pipes with the beams that travel in opposite directions, and collide in four intersection points, where the particles are detected by the four main experiments: ATLAS, CMS, LHCb and ALICE.

2.2. Luminosity at the LHC

LHC officially started in 2008, but physics data taking lasted from 2010 to 2012, during the so-called Run 1. In this phase the centre-of-mass enery was first kept at $\sqrt{s} = 7$ TeV, then, in 2012, it was raised to 8 TeV. From 2013 to 2014 there was the first Long Shutdown (LS 1), in order to allow upgrades and checks to the magnets' system. During Run 2, from 2015 to the end of 2018, the energy of each beam was 6.5 TeV, for a centre-of-mass energy of 13 TeV. Currently, it's in the LS 2 phase. Run 3 will start in 2021, during which the beams will reach the design energy of 8 TeV, i.e. $\sqrt{s} = 14$ TeV.
The event rate in LHC collisions is:

$$R \equiv \frac{dN}{dt} = \mathcal{L} \cdot \sigma, \tag{2.1}$$

where σ is the cross section of the event of interest, and $\mathcal{L}$ is the instantaneous luminosity of the collider. For a circular collider and assuming a Gaussian beam distribution, it depends

only on the beam parameters and can be written as follows:

$$\mathcal{L} = \frac{k}{4\pi} \cdot \frac{f n_1 n_2}{\sigma_x \sigma_y} \frac{1}{\sqrt{1 + \left(\frac{\sigma_z}{\sigma_y} \cdot tan(\frac{\phi}{2})\right)^2}}, \tag{2.2}$$

where:

- k is the number of bunches per beam. For LHC $k \sim 1200$ during Run 1 and $k \sim 2500$ during Run 2;
- n_1, n_2 are the numbers of the particles in the bunches, respectively. In particular, at the beginning of the fill, $n_1 = n_2 = n$, with $n = 1.15 \times 10^{11}$ for the protons and 7×10^7 for the ions (during Run 2);
- f is the revolution frequency: $f = 11\,245$ kHz, corresponding to a round trip time of 90 µs;
- σ_x, σ_y and σ_z are the standard deviations of the Gaussian beam distributions in the x, y and z directions, respectively, with $\sigma_x = \sigma_y \simeq 16$ µm and $\sigma_z \gg \sigma_x$, with $\sigma_z \sim$ cm;
- ϕ is the intersection angle of the 2 beams.

An important quantity to be computed is the bunch separation, that quantifies the time elapsed between the transition of two bunches; it was of 50 ns during Run 1, and of 25 ns during Run 2, that correspond to frequencies of 20 MHz and 40 MHz, respectively. So the electronic devices must be capable to readout at these frequencies. Nevertheless, the beams are injected from the SPS to LHC, and this leads to technical issues there are some missing bunches and to decrease the average crossing frequency to 15 MHz and 30 MHz, for Run 1 and Run 2 respectively.

The CMS experiment was designed to reach a maximum peak luminosity of $\sim 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. Figure 2.2 shows the values of the integrated luminosity delivered at CMS, and Tab. 2.1 shows the values of the integrated luminosity delivered at LHC, and the one recorded and validated by CMS.

2.3. The CMS detector

The Compact Muon Solenoid (CMS) detector is one of the four main detectors around LHC; it is 21.6 m long and has a total diameter of 14.6 m.
One of the most important aspects driving the design of the detector is the requirement to measure with high precision the momentum of muons, and to do it within a compact spectrometer and without stringent demands on muon-chamber resolution and alignement; so a large bending power, hence a high magnetic field, is needed. The choice of the magnetic field configuration led to the use of superconducting technology for the magnets.
The structure of the detector, from the inside to the outside, is the following:

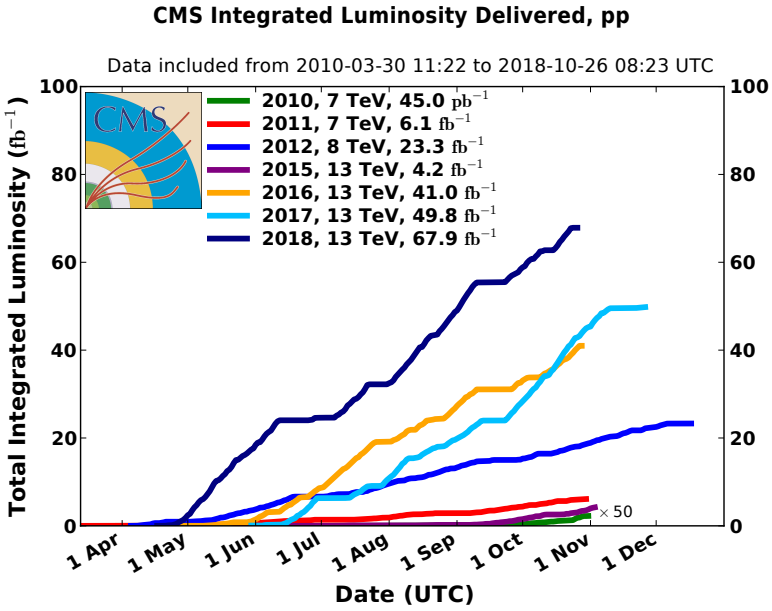


Figure 2.2: CMS Integrated luminosity, $L = \int \mathcal{L} dt$, delivered during Run 1 and Run 2.

Year	Centre-of-mass-energy $\sqrt{s}$ (TeV)	LHC Delivered luminosity (fb^{-1})	CMS recorded luminosity (fb^{-1})	CMS validated luminosity (fb^{-1})
2010 (Run 1)	7	44.22×10^{-2}	40.76×10^{-2}	34.68×10^{-2}
2011 (Run 1)	7	6.13	5.55	5.32
2012 (Run 1)	8	23.30	21.79	19.62
2015 (Run 2)	13	4.22	3.81	2.39
2016 (Run 2)	13	40.82	37.76	35.92
2017 (Run 2)	13	49.28	44.51	41.53
2018 (Run 2)	13	67.86	63.67	59.74

Table 2.1: Luminosity delivered by LHC, recorded by CMS and certified as good for physics analysis, during Run 1 and Run 2.

- In the innermost part there is the **Tracking System** (that will be discussed in Sec. 2.3.2): it employs silicon micro-strip detectors, which provide the required granularity and precision, and silicon pixel detectors, placed close to the interaction region, to improve the measurement of the impact parameter of charged-particle tracks and the position of secondary vertices.
- The **Electromagnetic calorimeter (ECAL)**, discussed in Sec. 2.3.3, uses lead tungstate ($PbWO_4$) crystals; the scintillation light is detected by silicon avalanche photodiodes (APDs) in the barrel region and vacuum phototriodes (VPTs) in the endcap region. A preshower system is installed in front of the endcap ECAL for π_0 rejection.
- ECAL is surrounded by a brass/scintillator sampling **Hadron calorimeter (HCAL)**, where the scintillation light is converted by wavelength-shifting (WLS) fibers embedded in the scintillator tiles and channeled to photodetectors via clear fibers. This light is detected by hybrid photodiodes (HPDs) that can provide gain and operate in high axial magnetic fields.
- **Forward Calorimeters**. The hadron calorimeters complex, placed in the central region, is complemented by an iron/quartz-fiber calorimeter in the forward region. The Cerenkov light emitted in the quartz fibers is detected by photomultipliers. The forward calorimeters ensure full geometric coverage for the measurement of the transverse energy in the event.The HCAL and the forward hadronic calorimetry are discussed in Sec. 2.3.4.
- A 13-meter long **superconducting solenoid**, that produces a 4 T magnetic field. Its 5.9 m inner diameter is large enough to accommodate the inner tracker and the calorimeters inside.
- **Muon system**: The return field is large enough to saturate 1.5 m of iron, allowing 4 muon stations to be integrated to ensure robustness and full geometric coverage. Each muon station consists of several layers of aluminum drif tubes (DT) in the barrel region and cathode strip chambers (CSCs) in the endcap region, complemented by resistive plate chambers (RPCs).

2.3.1 The CMS Coordinate System convention

The coordinate system adopted by CMS has the origin centered at the nominal collision point inside the experiment, the y-axis pointing vertically upward, and the x-axis pointing radially inward toward the center of the LHC ring [42]. Thus, the z-axis points along the beam direction toward the Jura mountains from LHC Point 5. A schematic picture of the coordiate system is shown in Fig. 2.4. Since CMS has a cylindrical symmetry, it is convenient to reconstruct the particles' tracks with the cylindrical coordinates: the azimuthal angle ϕ, measured from the x-axis in the x-y plane (also called the *transverse plane*), and the polar angle θ, measured from the z-axis. The polar angle can be expressed more conveniently in

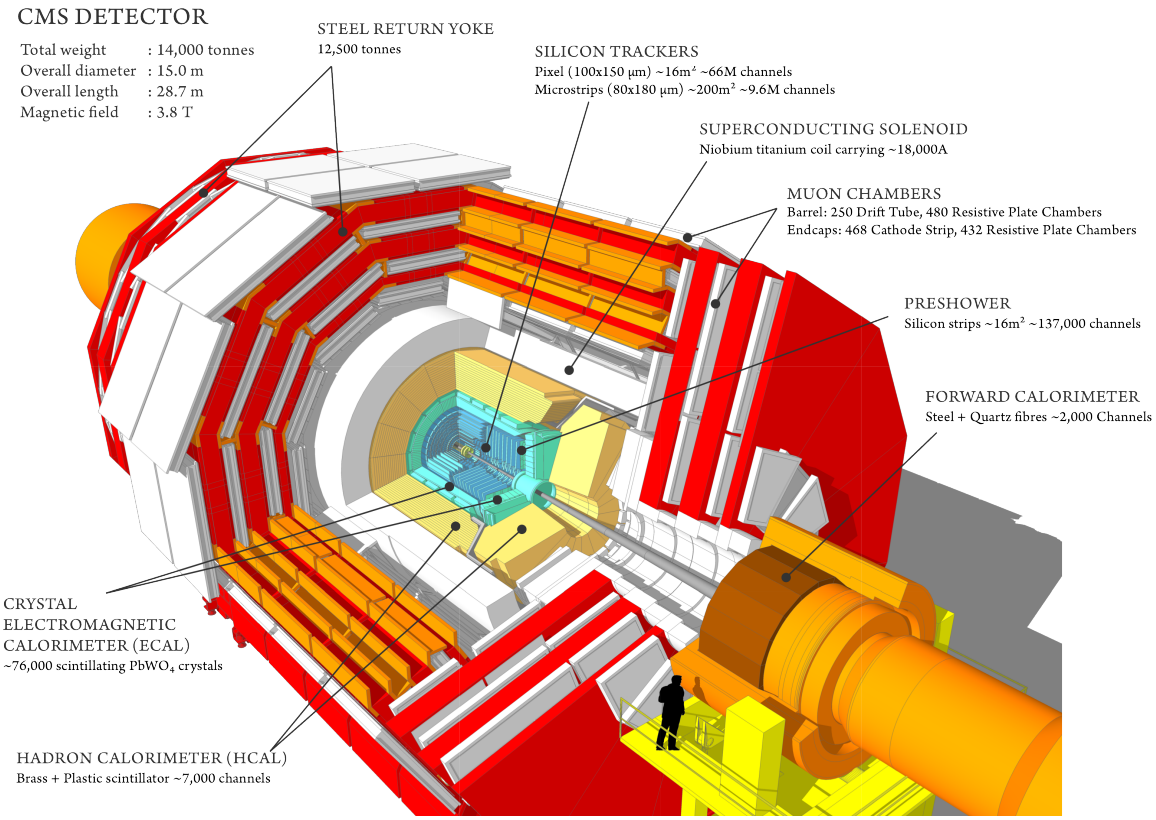


Figure 2.3: A schematic view of the CMS detector.

terms of *pseudorapidity*, η:

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

The pseudorapidity can be defined also in function of the momentum of the particle:

$$\eta = \frac{1}{2}\ln\left(\frac{|\mathbf{p}| + p_z}{|\mathbf{p}| - p_z}\right), \tag{2.4}$$

where p_z is the component of the momentum along the beam axis. For ultra-relativistic particles, where $E \approx |\mathbf{p}|$, the pseudorapidity converges to the definition of *rapidity*, y, defined as:

$$y = \frac{1}{2}\ln\left(\frac{E + p_z c}{E - p_z c}\right), \tag{2.5}$$

Through these parameters it is possible to define a distance between two particle directions as follows:

$$\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}\,. \tag{2.6}$$

It is possible to identify three regions of the sub-detectors: the *barrel region* identifies the central part of the CMS apparatus, where the sub-detectors are coaxial to the beam, and

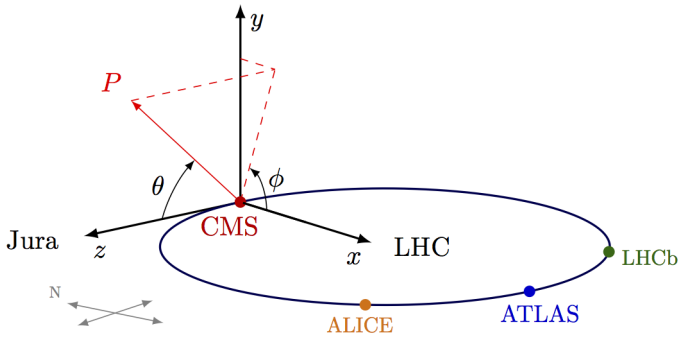


Figure 2.4: Scheme of CMS coordinate system.

extends in $|\eta| < 0.9$; the *overlap regions*, between the barrel and the endcap regions, covering the range $0.9 < |\eta| < 1.2$; the *endap regions*, that extend in $|\eta| > 1.2$. A longitudinal view of the CMS detector, expressed in terms of the pseudorapidity, is shown in figure Fig. 2.5.

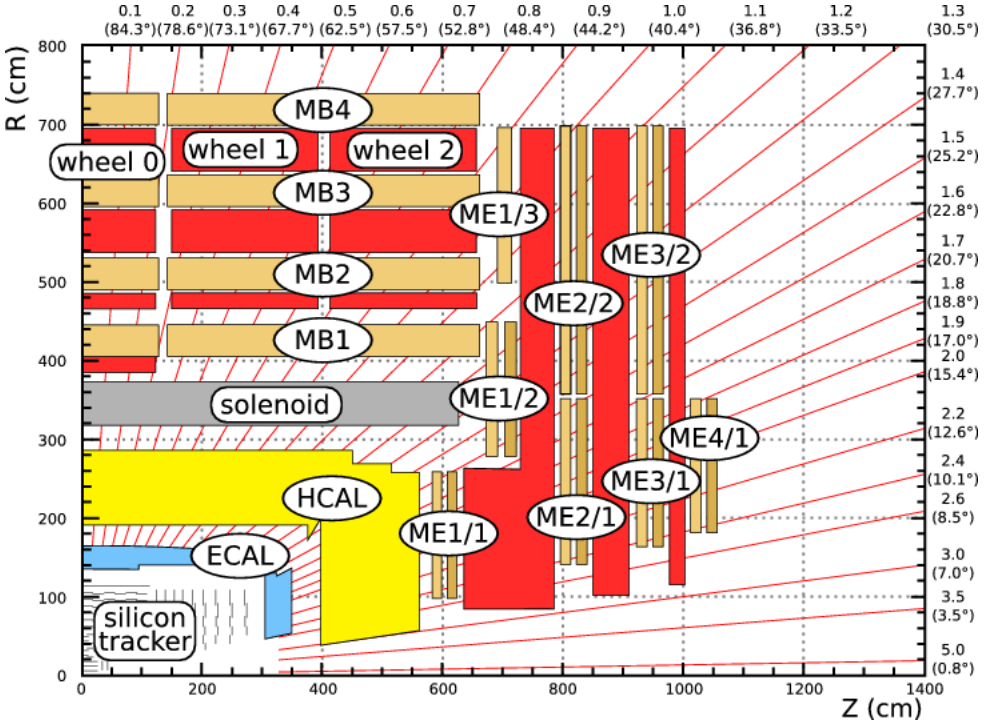


Figure 2.5: Longitudinal view of the CMS detector; the red lines denote the values of η and θ.

The momentum of a particle in the cartesian representation, can be divided in the *longitudinal* component, along the beam axis, that in the CMS coordinate system is p_z, and in the *transverse* component, $\vec{p}_T$, whose modulus is:

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

Differences in η and p_T are invariant under boost in th z direction: this is useful because the centre-of-mass of parton-parton collisions is often boosted along the z direction.

2.3.2 Tracker

The CMS tracking volume is a $\sim 5.6\,\text{m}$ long cilinder, with a diameter of $\approx 2.2\,\text{m}$. One of the main features of the CMS tracker is that it is based entirely on the silicon technology.
The entire silicon strip detector consists of more than 15000 modules mounted on carbon-fiber structures and housed inside a temperature controlled outer support tube. The operating temperature is aroung $-20\,^{\circ}\text{C}$.

A schematic view of the CMS tracker system is shown in Fig. 2.6. The tracker covers the region $|r| < 110\,\text{cm}$, $|\eta| < 2.4$ and $|z| < 280\,\text{cm}$.

Barrel region

In the barrel, it is possible to identify three regions, from the innermost to the outermost:

- **Pixel Barrel tracker**: this is the closest region to the interaction vertex, where the particle flux is very high, with $R < 11\,\text{cm}$. It is constituted by silicon pixel detectors of area $100 \times 150\,\mu\text{m}^2$ per pixel, for a total area of about $1\,\text{m}^2$ and $\sim 66 \times 10^6$ pixels arranged in three cylindrical layers. The spatial resolution is $10\,\mu\text{m}$ for R and ϕ, and $20\,\mu\text{m}$ for z coordinate. From Run 2 there is an additional layer.
- **Tracker Inner Barrel (TIB)**, which covers the region $20 < R < 55\,\text{cm}$ and $|z| < 55\,\text{cm}$: it's made of four layers of $320\,\mu\text{m}$ thick silicon microstrips along the z direction, with a pitch between each other in the range from $80\,\mu\text{m}$ to $120\,\mu\text{m}$. The first two modules are "stereo" in order to provide a measurement in both ϕ and z coordinates, for a stereo angle of 100 mrad and a single point resolution of $23 - 34\,\mu\text{m}$ in ϕ direction and of $23\,\mu\text{m}$ in z.
- **Tracker Outer Barrel (TOB)**, that is made of 6 layers that extend in the region of $|z| < 110\,\text{cm}$. Since the radiation levels are smaller in this region, thicker silicon sensors ($500\,\mu\text{m}$), a longer strip length and wider pitch, that varies from $120\,\mu\text{m}$ to $180\,\mu\text{m}$, are used. The first 2 layers provide a stereo measurement in both ϕ and z coordinates. The stereo angle is again 100 mrad and the single-point resolution is about $35 - 52\,\mu\text{m}$ in ϕ direction and of $52\,\mu\text{m}$ in z.

Endcap regions

The TOB is longer than the TIB, in the transition zone between the TIB and the endcap regions there are 3 layers of **Tracker Inner Disks (TID)**, whose center is along the beam axis; they have strips that point toward the beam axis, with a variable pitch. The thickness of the sensors is of $320\,\mu\text{m}$, and first two rings of the TID have stereo modules.
In the endcap regions there are the **Tracker Endcap (TEC)**, extending from $120\,\text{cm} < |z| <$

280 cm. The TECs are made by 9 disks centered on the beam axis and of variable thickness of 320 µm for the 3 innermost rings and 500 µm for the others. The fifth ring has stereo modules.

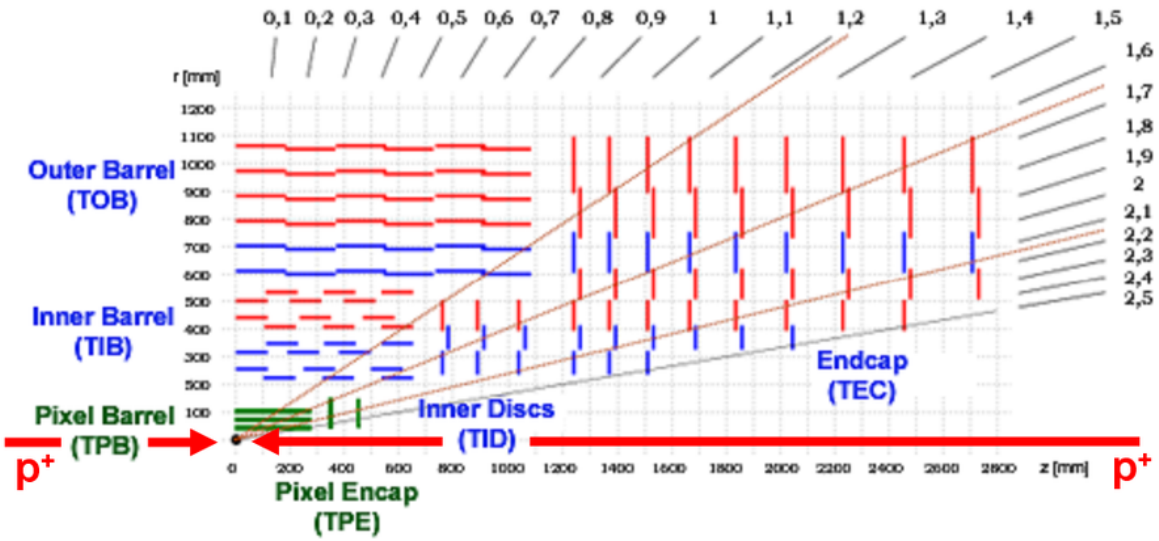


Figure 2.6: A schematic view of the CMS tracker.

Momentum resolution

In general, the momentum resolution can be written as follows:

$$\left(\frac{\sigma_p}{p}\right)^2 = \left(\frac{\sigma_{p_T}}{p_T}\right)^2 + \left(\frac{\sigma_\theta}{\sin\theta}\right)^2. \tag{2.7}$$

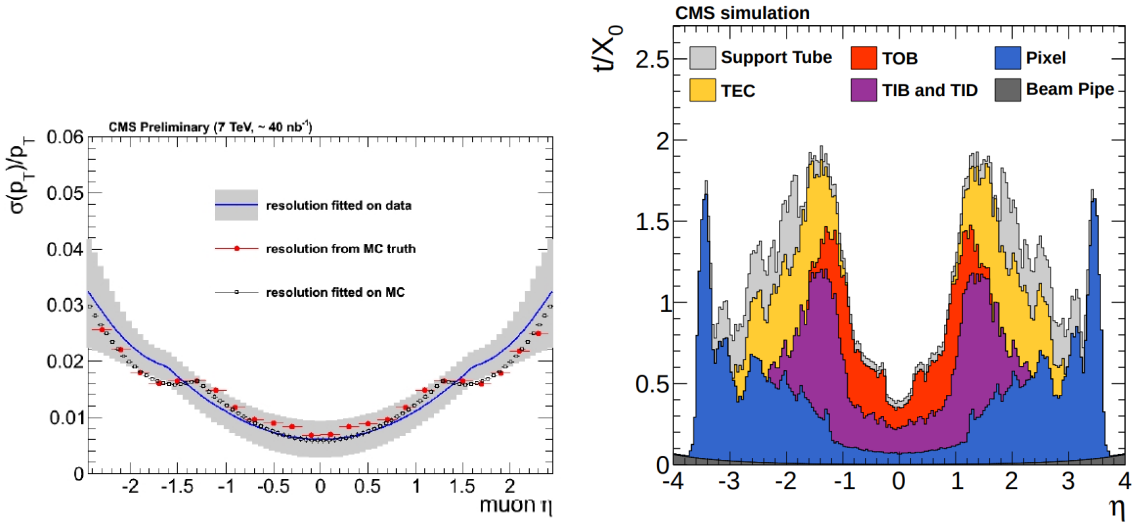
Where the first term is the relative uncertainty on the track reconstruction and the second one is the multiple scattering contribution. In particular, by considering the features of CMS one can write:

$$\left(\frac{\sigma_{p_T}}{p_T}\right)^2 = c_1 \cdot \left(\frac{p_T}{BL^2}\right)^2 + c_2 \cdot \left(\frac{1}{B\sqrt{LX_0}}\right)^2 \simeq (13\% \cdot p_T(\mathrm{TeV}/c))^2 + (0.7\%)^2, \tag{2.8}$$

where the latter coefficients reflect the average resolution in CMS. It is important however to notice that the momentum resolution depends on η too. This is due to the number and types of sub-detectors the particle crossed, and the material budget, i.e. the quantity of matter along the particle's path, as shown in Fig. 2.7(b); Fig. 2.7(a) shows the resolution dependencies from η comparing data and simulated J/ψ decay.

2.3.3 The Electromagnetic Calorimeter (ECAL)

The Electromagnetic Calorimeter (ECAL) is homogeneous and hermetic, made of 61200 $PbWO_4$ crystals in the barrel, and 7324 in each endcap. The truncated-pyramid shaped



(a) (b)

Figure 2.7: (a) shows the transverse momentum resolution dependency from η measured through J/ψ decays; (b) shows the quantity of matter in function of η.

crystals cross-section corresponds to approximately $\Delta\eta = 0.0174$ and $\Delta\phi = 0.0174$ The ECAL, hosted inside the superconducting magnet, must be compact, the crystals must reach the required granularity and precision, be fast and resistant to the radiations in a relatively small volume.
The choice of the $PbWO_4$ is due to the high density $\rho(PbWO_4) = 8.28\,g/cm^3$, small values of the radiation length and Moliere radius, $X_0 = 0.89\,cm$, and $\lambda_M = 2.2\,cm$, respectively; in addition, the 80% of the light is emitted in $\sim 25\,ns$ and this material has a huge radiation emission power ($\sim 10\,Mrad$). Nevertheless, it has a relatively small light yield of $30\,\gamma/MeV$: so it's necessary to couple those crystals with high gain photo-detectors, that are capable to operate in a high magnetic field: Silicon Avalanche Photo-diodes (SiAPD) are thus used in the barrel region and Vacuum Photo-diodes are used in the endcaps. The sensitivity of these devices to the temperature requires a temperature stability of $0.1\,°C$.

It is possible to separate two regions:

- **ECAL Barrel (ECAL B)**, with an internal radius of $129\,cm$, is divided in 36 super-modules, each long half of the barrel length, covering up the region $|\eta| < 1.48$. The crystal length is $23\,cm$, corresponding to $25.8\,X_0$.
- **ECAL Endcaps (ECAL E)**, that are $314\,cm$ far from the nominal interaction vertex, covering the range $1.48 < |\eta| < 3.0$. Each ECAL E is constituted by two aluminum semicircular plates, where there are structural unities of 5×5 crystals. They are $22\,cm$ long, corresponding to $\sim 24.7\,X_0$, and have a superficial area of $28.6 \times 28.6\,mm^2$.

An additional **Pre-shower devices**, made by two planes of silicon strip detectors with $1.9\,mm$

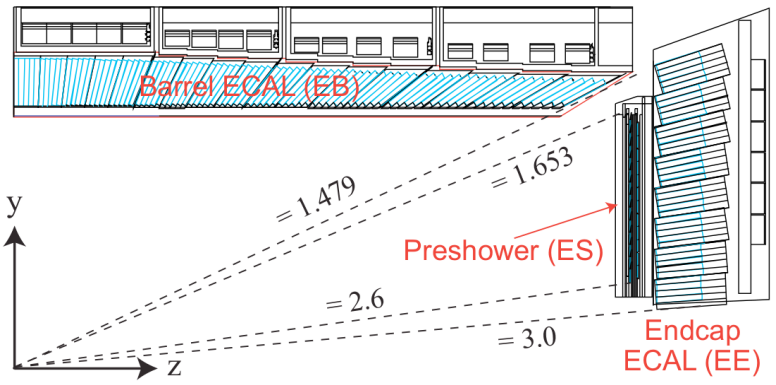


Figure 2.8: Transverse section through the ECAL, showing geometrical configuration.

pitch and absorber lead disks, of $2\,X_0$ and $3\,X_0$ depth, respectively, is placed in front of the endcap crystals, to enhance photon identification capabilities, in particular to optimize the π^0 detection. The ECAL barrel energy resolution for electrons in beam tests has been measured to be [43]:

$$\frac{\sigma_E}{E} = \frac{2.8\%}{\sqrt{E\,(\mathrm{GeV})}} \oplus \frac{12\%}{E\,(\mathrm{GeV})} \oplus 0.3\% \tag{2.9}$$

where the three contributions correspond to the stochastic fluctuations, electronic noise, and constant terms. The stochastic term includes contributions from the shower containment, the number of photoelectrons and the fluctuations in the gain process. The noise term of 12% at 1 GeV corresponds to a single channel noise of about 40 MeV. The constant term, which dominates the energy resolution for high-energy electron and photon showers, depends on non-uniformity of the longitudinal light collection, energy leakage from the back of the calorimeter, single-channel response uniformity and stability.

2.3.4 The Hadronic Calorimeter (HCAL)

The **Hadronic Calorimeter (HCAL)** design was driven by the choice of the magnet parameters since most of the CMS calorimetry is located inside the magnet coil and surrounds the ECAL system.

The HCAL is a sampling calorimeter, with brass as absorber material, since it has short interaction length and it is non-magnetic. An important requirement is to minimize the non-Gaussian tails of the energy resolution and to provide a containment and a good hermiticity in order to perform a good E_T^{miss} measurement. This leads to maximize the absorber material inside the magnet coil, in terms of radiation lengths, and thus to minimize the space for the active medium. The active medium is made by plastic scintillator tiles read out with embedded wavelength-shifting (WLS) fibers, that are spliced to high-attenuation-length clear fibers outside the scintillator that carry the light to the readout system; this is based on multi-channel Hybrid Photo-diodes (HPDs). The whole structure enables the HCAL to be

built with essentially no uninstrumented cracks or dead areas in ϕ. It is possible to identify three region inside the magnet coil for the HCAL:

- the **Hadron Barrel (HB)** covers the region $|\eta| < 1.4$, and is constituted by 32 towers with $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ segmentation, for a total of 2304 towers. It is divided in two half barrels, 15 brass plates of thickness 5 cm and two external stainless steel plates. Particles leaving the ECAL volume first see a scintillator plate with a thickness of 0.9 cm rather than 0.37 cm for the other plates. The light collected by the first layer is optimized to be a factor of about 1.5 higher than the other scintillator plates.
- the **Hadron Endcap (HE)** covers the region $1.3 < |\eta| < 3.0$ with 14 towers. The separation for the 5 outermost towers, at smaller η is of 5° and 0.087 in η, while for the 8 innermost towers is 10° in ϕ and varies from 0.09 to 0.35 (for the highest values of η) in η. The total number of HE towers is 2304.
- the **Hadron Forward (HF)** covers the range $3.0 < |\eta| < 5.0$, and is made by steel/quartz fiber; the neutral component of hadron shower is preferentially sampled in the HF technology. The front face is located at 11.2 m from the interaction point; the depth of the absorber is 1.65 m. The signal originates from Cerenkov light emitted in the quartz fibers, which is then channeled by the fibers to photomultipliers. The quartz fibers run parallel to the beam line and are inserted into grooves in order to create 2 effective longitudinal samplings. There are 13 towers in η with a size given by $\Delta\eta = 0.175$ (except for the lowest η tower and the highest η tower with $\Delta\eta = 0.1$ an $\Delta\eta = 0.3$, respectively) and $\Delta\phi = 10°$ (except for the highest η tower that has $\Delta\phi = 20°$, for a total of 900 towers.

The gap between the barrel and the endcap HCAL, through which the services of the ECAL and the inner tracker pass, is inclined at 53° and points away from the centre of the detector.

The Hadron Outer

Outside the magnet coil there is another hadron calorimeter, the **Hadron Outer (HO)**, which contains 0.1 cm thick scintillators lining the outside of the outer vacuum tank of the coil and covers the region $|\eta| < 1.26$.
The tiles are grouped in 30° sectors, matching the segmentation of the Drift Tube Chamber; in fact, HO is physically located inside the barrel muon system, hence it is constrained by the geometry and construction of that system (see section 2.3.6).
The HO samples the energy from penetrating hadron showers leaking through the rear of the calorimeters and so serves as a tail-catcher after the magnet coil and increases the effective thickness of the hadron calorimetry to over $10\lambda_I$, therefore reducing the tails in the energy resolution function. An improvement in the Missing Transverse Energy (MET) resolution is provided by the HO.It is divided in 5 sections along η, called "rings" with enumeration from -2 to 2. Each ring covers 2.5 m in the z coordinate. HO scintillators follow the HCAL barrel tower geometry in η and ϕ.

Parameter	Value
Field	3.8 T
Inner Bore	5.9 m
Length	12.9 m
Current	19.5 kA
Stored Energy	2.7 GJ
Number Of Turns	2168

Table 2.2: Parameters of CMS superconducting solenoid.

2.3.5 The Superconducting Solenoid

As mentioned before, one of the first requirement for the CMS experiment design was to measure with high precision the momentum of the muons. To do it, a large bending power is needed, hence an intense magnetic field. In fact, the CMS magnet enables the muon system to reach the required performance, as the momentum resolution of $\Delta p/p \approx 10\%$ at $p = 1\,\mathrm{TeV/c}$. The CMS detector contains a large superconducting solenoid, whose parameters are listed in Tab. 2.2. The main features of the solenoid are:

- the use of a high-purity and aluminum-stabilised conductor;
- indirect cooling, by a thermosyphon;
- full epoxy impregnation.

The Rutherford-type cable is co-extruded with pure aluminum, which acts as a thermal stabilizer. The flux is returned through a 10^8kg iron yoke comprising 5 wheels and 2 endcaps, composed of three disks each. The iron return yoke allows for a constant 1.8 T field also in the region outside the magnet.

2.3.6 The Muon System

The detectors used for the muon system are all gaseous, a choice driven by the large surface to be covered and by the different particle flux.

Muon detectors in the Barrel region

In the barrel region ($|\eta| < 1.2$), where the muon rate is low, **Drift Tube Chambers (DTCs)** are used. In particular, 250 DTCs are arranged in 4 layers at different distances from the beam axis (4.0 m, 4.9 m, 5.9 m and 7 m), alternated with iron yokes. It is divided in 5 wheels, and each wheel is divided into 12 sectors; each sector covers 30° azimuthal angle. There are 12 chambers in each of the 3 inner layers. In the 4th layer, the top and bottom sectors host 2 chambers each, thus leading to a total of 14 chambers per wheel in this outermost layer.

The chambers belonging to different station are staggered so that high-p_T muon produced near a sector boundary crosses at least 3 out of the 4 stations.

Each DTC has 1 or 2 **Resistive Plate Chambers (RPCs)** coupled to it, so that high-p_T muons cross up to 6 RPCs and up to 4 DTCs, producing up to 44 measured points in the DT system from which a muon-track candidate can be built.

Muon detectors in the Endcap region

The Muon Endcap (ME) system comprises 468 CSCs in the 2 endcaps. Each CSC is trapezoidal in shape and consists of 6 gas gaps, each gap having a plane of radial cathode strips and a plane of anode wires running almost perpendicularly to the strips. Most of the CSCs are overlapped in ϕ in order to avoid gaps in the muon acceptance.

The gas ionization and subsequent electron avalanche caused by a charged particle traversing each plane of a chamber produces a charge on the anode wire and an image charge on a group of cathode strips.
The signal on the wires is fast and is used in the Level-1 Trigger. However, it leads to a coarser position resolution. Each CSC measures up to 6 space coordinates (r, ϕ, z). The spatial resolution provided by each chamber from the strips is typically about 200 µm.
In the endcap regions there are RPCs too, to help resolve ambiguities in the CSCs.

Muon momentum resolution

The centrally produced muons are measured in the inner tracker, after the coil and in the return flux (where the curvature has changed).

A momentum measurement performed by the muon system only is essentially determined by the bending angle of the muons when they come out the magnet (taking as origin the interaction point, which is known with precision of 20 µm); as it is shown in Fig. 2.9, the resolution is dominated by the multiple scattering contribution, for muon transverse momentum values up to 200 GeV/c; then the chamber spatial resolution starts to dominate.

For low momentum muons, the best momentum resolution is obtained in the tracker. The muon trajectory beyond the return yoke extrapolates back to the beamline due to the compensation of the bend before and after the coil when multiple scattering and energy loss can be neglected. This fact can be used to improve the muon momentum resolution at high momentum when combining the inner tracker and muon detector measurements (full system in Fig. 2.9).
Figure 2.9 shows the muon momentum resolution dependency from the momentum itself, in two η ranges.

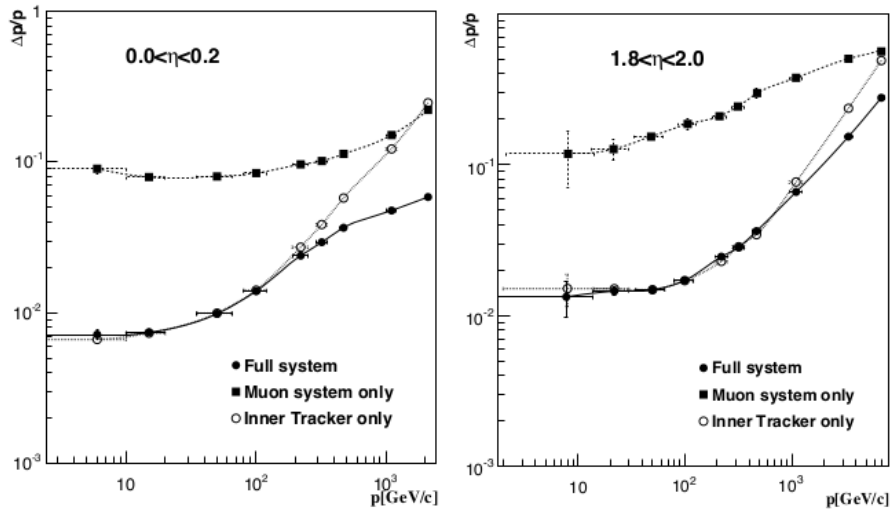


Figure 2.9: he muon momentum resolution versus p using the muon system only, the inner tracker only, or both ("full system"), in the barrel with $|\eta < 0.2|$ (left panel) and in the endcap with $1.8 < |\eta| < 2.0$ (left panel).

2.4. The CMS Trigger

The Trigger and Data Acquisition (DAQ) system of an experiment at a hadron collider plays an essential role because both the collision and the overall data rates are much higher than the rate at which one can write data to mass storage. At the LHC, the proton beams cross each other at a frequency of 40 MHz.
The online event selection process (trigger) must reduce the approximately 1 billion interactions per second to no more than about $\mathcal{O}(100)$ events/s for storage and subsequent analysis. The trigger system consists of two levels designed to select events of potential physics interest. The required rejection power of $\mathcal{O}(10^5)$ is too large to be achieved in a single processing step. For this reason, the full selection task is split into two steps.
The first level of the trigger, the **L1-Trigger** is implemented in hardware, and selects events containing detector signals consistent with a particle candidate [44].

The second level, called the **High Level Trigger (HLT)**, implemented in software, further refines the purity of the output stream, selecting an average rate of 400 Hz for offline event storage [45].

The L1 Trigger

The first level (**L1**) of the CMS trigger is implemented in custom hardware, and selects events containing candidate objects, e.g., ionization deposits consistent with a muon, or energy clusters consistent with an electron, photon, τ lepton, missing transverse energy (MET), or jet. Within 4 µs from a collision, the system must decide if an event should be accepted

or rejected by using information from the calorimeter and muon detectors. The trigger thresholds are adjusted to the LHC instantaneous luminosity during data taking in order to restrict the output rate to 100 kHz, the upper limit imposed by the CMS readout electronics.

A more detailed scheme of the L1 Trigger architecture is shown in Fig. 2.10: data from the HF and the HCAL, and from the ECAL, are processed first regionally (RCT) and then globally (GCT). Energy deposits (hits) from the RPCs, CSCs, and DTs are processed either via a pattern comparator or via a system of segment- and track-finders and sent to a global muon trigger (GMT).

The information from the GCT and GMT is combined in a global trigger (GT), which makes the final trigger decision. This decision is sent to the tracker (TRK), ECAL, HCAL or muon systems (MU) via the trigger, timing and control (TTC) system. The data acquisition system (DAQ) reads data from various subsystems for offline storage [46].

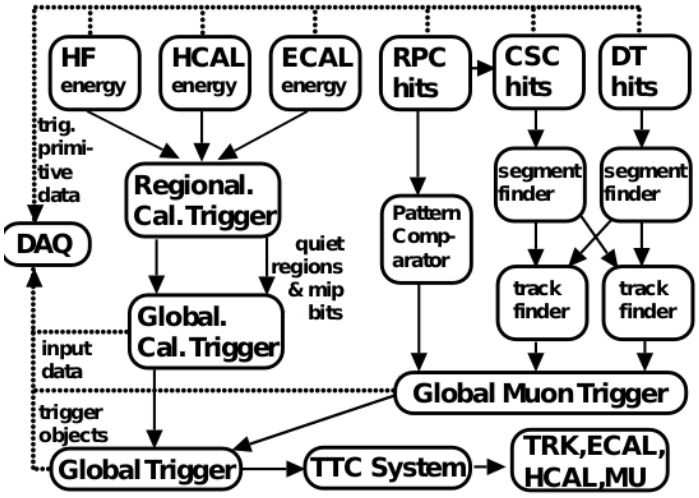


Figure 2.10: Overview of the CMS L1 trigger system.

The High Level Trigger (HLT)

The HLT hardware consists of a single processor farm composed of commodity computers, the event filter farm (EVF). The event filter farm consists of filter-builder units. In the builder units, individual event fragments from the detector are assembled to form complete events. For each event, objects such as electrons, muons, and jets are reconstructed and identification criteria are applied in order to select only those events which are of possible interest for data analysis. The filtering process uses the full precision of the data from the detector, and the selection is based on offline-quality reconstruction algorithms. The High-Level Trigger is designed to reduce this maximum Level-1 accept rate of 100 kHz to a final output rate of 100 Hz.

3 Data sample and event selection

3.1. Data and simulated samples

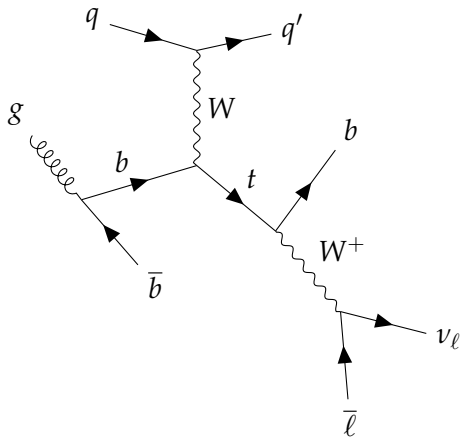
In order to predict the signal and background process, simulated samples with the Monte Carlo (MC) technique have been used, collected with the requirement of one muon in the final state. State-of the art simulation softwares are used to reproduce the hard-scattering interaction of the standard model, either at tree-level, i.e. Leading-Order (LO) or up to the first perturbative expansion, or Next-to-Leading-Order (NLO). The simulation of parton shower, hadronization and the underlying event is performed by PYTHIA 8.180 [47] with tune CUETP8M1 [48], except for the $t\bar{t}$ background, where the CUETP8M2T4 tune is used [49]. The default parametrisation of the parton distribution functions (PDFs) used in all simulations is NNPDF30 nlo [50]. All the generated events undergo a full simulation of the detector response according to the implementation of the CMS detector within GEANT4 [51]. Additional proton-proton interactions within the same or nearby bunch crossing (pileup) are included in the simulation with the same distribution as observed in data [52].

The final states of interest are constituted by one lepton (muon or electron[1]), some missing transverse energy that brings neutrino information, and a different jet and b-jet multiplicity depending on the process; a more detailed description of the selection of those events can be found in Sec. 3.3. In the following are briefly described the simulated signal and background samples.

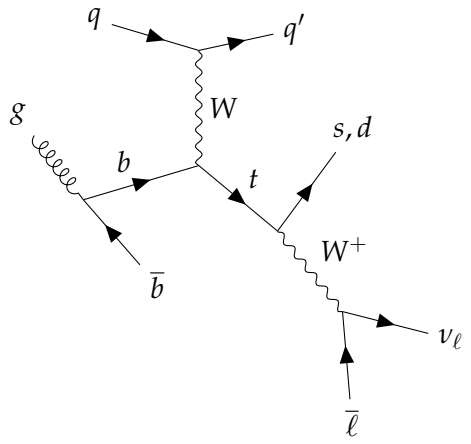
Signal events: single top t-channel processes. Their Feynman diagrams are represented in Fig. 3.1. In principle, they can have a different final state topology from the considered one, where the W decays hadronically. This topology is not considered in this analysis, since it is way more challenging to separate from the background sources, and the lepton

[1] In this analysis the considered lepton in the final state is the muon. Adding the samples with an electron in the final state will be a future developement.

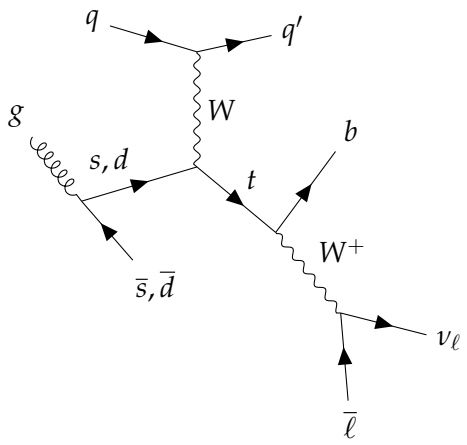
signature is more clear. The MC simulated sample is generated next-to-leading order (NLO) with POWHEG 2.0 in the four-flavour-scheme (4FS) for the $|V_{tb}|$-$|V_{tb}|$ events and in the five-flavour-scheme(5FS) for the $|V_{tb}|$-$|V_{tq}|$ or $|V_{tq}|$-$|V_{tb}|$ events [53]; the top-quark decays (both in $|V_{tb}|$ channel and in $|V_{ts}|$ a $|V_{td}|$ channels) are simulated with MADSPIN [54].

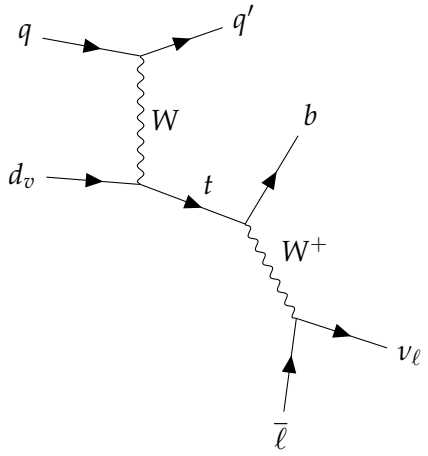


(a) S1: $|V_{tb}|$ in production and in decay.



(b) S2: $|V_{tb}|$ in production and $|V_{ts}|$ or $|V_{td}|$ in decay.





(c) S3: $|V_{td}|$ in production and $|V_{tb}|$ in decay.

Figure 3.1: Feynman diagrams of the single top productions and decays in the *t*-channel, corresponding to the signals.

Top-antitop pair production events This is the main background, since it contains two top quarks. It can simulate the signal in the semi-leptonic decays, since there is a lepton-neutrino pair and some jets in the final state, as shown in Fig. 3.2. The simulated sample is generated with POWHEG. Only the tWb decay vertices are considered for the analysis, whereas the $t\bar{t}_{b,q}$ is not considered in the analysis.

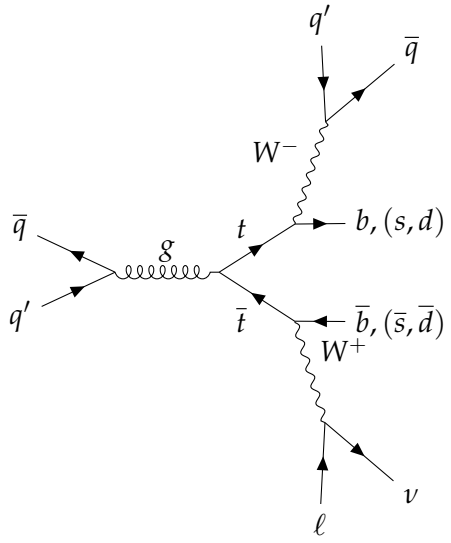


Figure 3.2: Feynman diagram at leading order of the $t\bar{t}$ pair production.

W or Z boson associated with jets events This is another main background source, when the jet multiplicity is the same as in the signal events and the bosons decay leptonically; the Feynman diagrams at the leading order of those processes are represented in Fig. 3.3. This sample is simulated using aMC@NLO [55] and the FxFx merging scheme [56], where up to two additional partons are generated at matrix-element level.

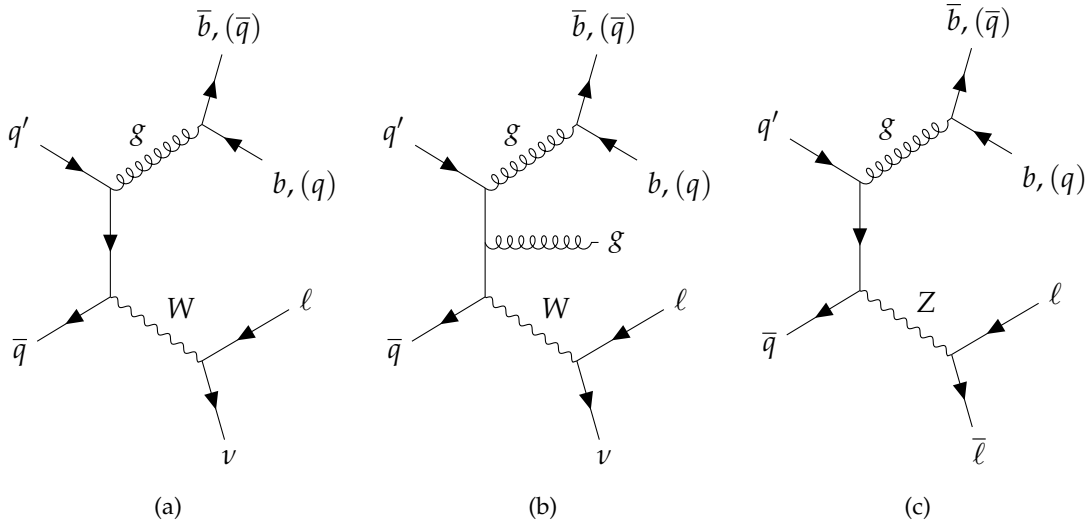


Figure 3.3: Feynman diagrams of W/Z bosons production associated with jets

Single top *s*-channel and tW associated production These events involve the single top-quark production, but in this analysis are considered as background sources. The Feynman

diagrams of those processes are shown in Fig. 3.4. The tW and $\bar{t}W$ samples are simulated with POWHEG, in the five-flavour-scheme (5FS) [57], whereas the s-channel events are generated using aMC@NLO in the four-flavour-scheme (4FS). Those processes are minor background sources, due to the smaller cross section at LHC, that are reported in Tab. 1.5 and Tab. 1.6.

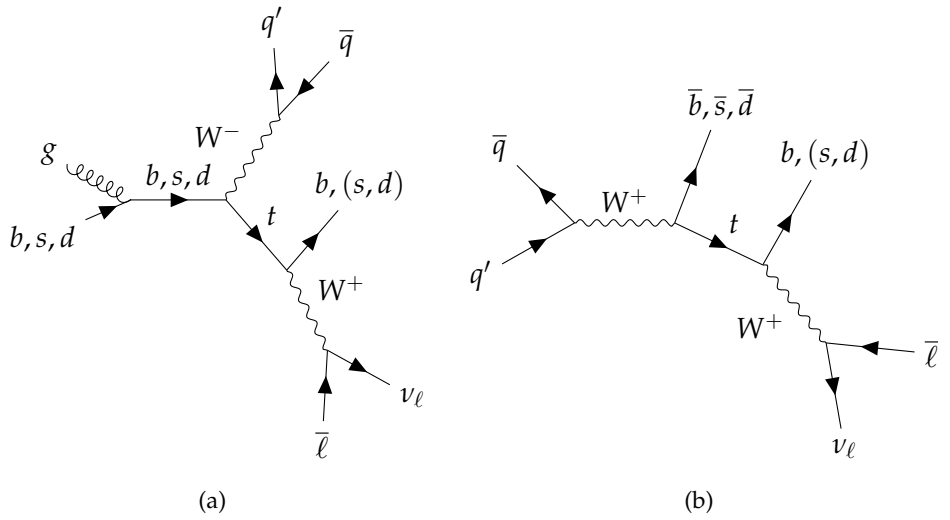


Figure 3.4: Feynman diagram at leading order of single top s-channel and tW associated production backgrounds.

Drell-Yan events This background sample is generated with M_50 aMC@NLO and the decays are simulated with MADSPIN. For the analysis purposes only Drell-Yan events with the couple lepton-antilepton in the final state have been considered, since they are more similar to the signal. The Feynman diagram of this process is shown in figure Fig. 3.5.

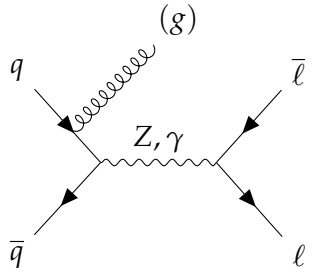


Figure 3.5: Feynman diagram at leading order of Drell-Yan background.

Double vector boson events Those samples are generated with aMC@NLO or POWHEG and the decays are simulated with MADSPIN [58, 59]. The Feynman diagrams of some of those processes are represented in Fig. 3.6.

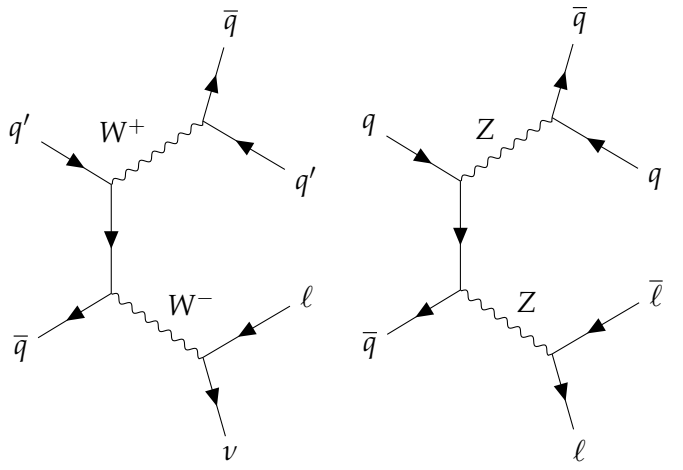


Figure 3.6: diboson background

QCD multi-jet background This is an important source of background. the Feynman diagrams at leading order of this QCD multi-jet events are shown in Fig. 3.7. The large cross section and the large rejection power of the final event selection for the QCD multi-jet processes makes it impossible to generate samples of simulated events in reasonable size. To overcome this difficulty, this background contribution is suppressed as much as possible and the remaining contamination is estimated directly from data. As will be described Sec. 3.4, some cuts are applied to the data sample to suppress the QCD multijet production.

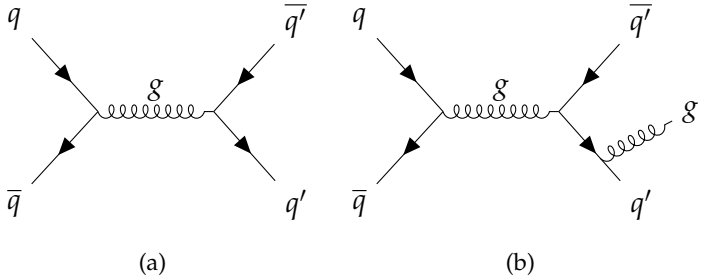


Figure 3.7: qcd multi-jet background

In table Tab. 3.1 are listed the cross sections used in the analysis for each simulated sample.

Signals	$\sigma(\cdot BR)$ (pb)
ST_{bb}	136.02
$S\overline{T}_{bb}$	80.95
ST_{bb} +$S\overline{T}_{bb}$	216.99
ST_{bq}	0.27
$S\overline{T}_{bq}$	0.17
ST_{qb}	0.27
$S\overline{T}_{qb}$	0.17
Backgrounds	$\sigma(\cdot BR)$ (pb)
$(\bar{t}+t)W$ associated production	71.1
Single $(\bar{t}+t)$ s-channel	10.32
$t\bar{t}$	831.76
$Z/\gamma(\rightarrow \ell\ell)+$ jets	5765.4
$W(\rightarrow \ell\nu)+0$ jets	49670.0
$W(\rightarrow \ell\nu)+1$ jet	8264.0
$W(\rightarrow \ell\nu)+2$ jets	2544.0
$WW(\rightarrow 2\ell 2\nu)+$ jets	12.178
$WW(\rightarrow 1\ell 1\nu)+$ jets	45.85
$WZ(\rightarrow 2\ell 2\nu)+$ jets	5.595
$ZZ(\rightarrow 2\ell)+$ jets	3.22
$WZ(\rightarrow 1\ell 1\nu)+$ jets	10.71

Table 3.1: Monte Carlo samples used in the analysis. For the samples restricted to specific decay channels the branching ratio (BR) is included in the cross section value quoted. The notation ST_{ij} ($S\bar{T}_{ij}$) stands for the single top (anti-top) t-channel with the i^{th} quark in production and the j^{th} quark in decay.

Data sample The data sample used in this analysis consists of 2016 collision data at $\sqrt{s} = 13\,\text{TeV}$, corresponding to an integrated luminosity of $L = 35.9\,\text{fb}^{-1}$.

3.2. Event reconstruction

The reconstruction algorithm used in CMS is the so-called Particle Flow (PF) algorithm [60], which aims to identify and reconstruct the single particles from pp collisions, combining optimally information from the sub-detectors.
The CMS detector design ad layout is well suited to particle-flow (PF) reconstruction, since it has a highly-segmented tracker, a fine-grained electromagnetic calorimeter, a hermetic hadron calorimeter, a strong magnetic field, and an excellent muon spectrometer. Hence, the event reconstruction is performed by a PF algorithm, tuned to the CMS detector.

The result leads to improved performances in the jet and τ leptons reconstruction, MET (missing transverse energy) determination and muon and electron identification.
This approach also allows to identify particles from pile-up interactions and enables efficient pileup mitigation methods.

3.2.1 The Particle Flow Algorithm

The PF algorithm is designed to [60]:

- maximize the separation between charged and neutral hadrons, exploiting a large field integral and the good electromagnetc calorimeter granularity achieved in CMS;
- get an efficient pure track reconstruction;
- perform a clustering algorithm able to disentangle overlapping showers and an efficient link procedure to connect together the deposits of each particle in the sub-detectors.

Tracks falling within the clusters' boundaries are extrapolated through the calorimeters, and are associated to the clusters. The set (tracks, cluster) is identified to be a *charged hadron* and it is not considered anymore in the algorithm. Before of hadrons, muons are identified so that their tracks is not associated to a charged hadron; for a more detailed description about muon identification and reconstruction, see in Sec. 3.3.3.
Electrons are way harder to deal with, because of the bremsstrahlung emissions. A more specific description of electrons reconstruction can be found in [61].
After all charged particles have been identified, the other calorimetric deposits are identified in photons (neutral hadrons) if their deposits are in the ECAL (HCAL).

Once associated the calorimetric deposits to the particles, their four momenta are determined. If the cluster energy (a linear combination of the ECAL and HCAL energies) associated to the track exceeds the track momentum of more than 1 sigma, a neutral particle is considered to be responsible of this excess. The energy associated to the neutral particle is given by the difference of the two measurements. The resulting particle list containing charged and neutral hadrons, photons, electrons and muons, is used to reconstruct more complex quantities as jets, missing transverse energy (MET) and to identify and reconstruct τ leptons.

3.2.2 Particle isolation

Particle isolation defined as the ratio of p_T of all the objects within a cone around the particle trajectory, expressed in the $\eta - \phi$ plane as $\Delta R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}$, and the p_T of the isolated object. For the particle x, the isolation is so computed:

$$I^x = \frac{\sum_{i=0}^{N_{obj}} p_T(i|\Delta R)}{p_T(x)}$$

Such variable is quantifies the energy deposited around the isolated object, as the muon or electron. In this analysis it has been used the *particle flow relative isolation*, defined as follows:

$$I^x_{rel} = \frac{S^{ch.h.} + max((S^\gamma + S^{n.h.} - S^{P.U.}), 0)}{p_T(x)} \tag{3.1}$$

where $S^{ch.h.}$, $S^{n.h.}$ and S^γ are the sum of the transverse energies deposited by charged hadrons, neutral hadrons, and photons respectively. In the muon case, $S^{P.U.} \equiv \Delta\beta \times \sum p_T^{P.U.}$, is the sum of pile-up transverse momenta, corresponding to tracks associated to non-leading vertices; it is used to estimate the contribution of neutral particles from pile-up events by applying the $\Delta\beta$ multiplying factor, that is $\Delta\beta = 0.5$, since it takes into account the neutral-to-charged particles ratio expected from isospin invariance; this factor maps the expected neutral contribution in the isolation cone from the observed PU charged one. The data-to-MC correction factors for muon are obtained using a "Tag and Probe" method and are applied to simulated events to account for differences in efficiencies between data and simulation. For the electron the definition $S^{P.U.} \equiv \rho \times A$, where rho is the average energy density not clustered in jets, and A is the effective area of the electron, i.e. the detector area hit by the calorimetric deposit is fitted in the calorimeter in order to ge the electron energy[2]. For this analysis, the working point chosen for the isolation cone is $\Delta R = 0.4$, and a particle is defined *isolated* if $I^x_{rel} < 0.06$ or *anti-isolated* if $I^x_{rel} > 0.15$.

3.2.3 Missing Transverse Energy

The vector momentum imbalance in the plane perpendicular to the beam direction is particularly useful in pp and $p\overline{p}$ colliders, and is known as missing transverse momentum, $\vec{\not{p}}_T$. Its magnitude is called missing transverse energy, $\not{E}_T$.

Missing transverse energy is one of the most important observables for discriminating leptonic decays of W bosons and top quarks from background events which do not contain neutrinos, such as multi-jet and DrellYan events, as will be explained later.

The reconstruction of $\not{E}_T$ is very sensitive to particle momentum mismeasurements, particle misidentification, detector malfunctions, particles impinging on poorly instrumented regions of the detector, cosmic-ray particles, and beam-halo particles, which may result in artificial $\not{E}_T$ [62].

The missing transverse momentum is computed by the PF algorithm as the opposite of the vectorial sum of the transverse momenta of the particles and the missing transverse energy is its module.

[2]Since the ECAL is not homogeneus the shape of the cluster may vary with position; hence an effective area is defined, from which is subtracted the Pile Up contribution that is assumed to be uniform.

3.2.4 W boson transverse mass

To further suppress contributions from processes where the muon does not come from a leptonically decaying W boson, a selection based the reconstructed transverse W-boson mass, m_T^W, can be applied. It is defined as:

$$m_T^W = \sqrt{(p_{T_L} + p_{T_\nu})^2 - (p_{x_L} + p_{x_\nu})^2 - (p_{y_L} + p_{y_\nu})^2} \tag{3.2}$$

where the transverse momentum of the neutrino is approximated by the missing transverse energy vector $\vec{\not{p}}_T$.

This quantity plays an important role in the determination of the data-driven QCD contribution.

3.3. Offline selection

A baseline selection on the events has been applied in order to have signal enriched samples, from the HLT to lepton and jets quality requirements.

3.3.1 High Level Trigger selection

Muon channel events are required to pass the `HLT_IsoMu24` OR the `HLT_IsoTkMu24`, that correspond to the presence of an isolated online muon candidate with $p_T > 24\,\text{GeV}/c$. Efficiencies for this trigger in simulation are corrected using data-to-MC scaling factors. Muon efficiencies are calculated using the official "Tag and Probe" tool provided by the Muon Physics Object Group [63] which utilizes the known resonances, for example J/ψ, Z, to measure object efficiencies. In the tag-probe method, "tagged" muon is required to pass a very tight selection where as the "probe" muon is required to pass a loose selection as compared to the tagged muon. The "tag-and-probe" pair is selected such that the invariant mass of tag-and-probe falls within the selected mass resonance window. The efficiency ($\epsilon = N_{passing\ probe}/N_{all}$) both for the data and MC samples is extracted by fitting the resonance peak. The data-to-MC simulation scale factors (SF) are than derived as $SFs = \epsilon^{Data}/\epsilon^{MC}$. As is shown in Fig. 3.8 the trigger efficiency for p_T around the threshold is $\sim 95\%$, and for $p_T \gtrsim 40\,\text{GeV}/c$ it is stabilized around the 99%.

Electron channel events must fulfill the `HLT_Ele32_eta2p1_WPTight_Gsf_vX` trigger, which requires the online electron to have $p_T > 32\,\text{GeV}/c$ and $|\eta| < 2.1$. Electron trigger efficiencies are calculated using the official "Tag and Probe" tool provided by the E/gamma Physics Object Group. The efficiency to select events with electrons having an HLT match was estimated using a well known resonance decaying in the dileptonic channel, for example the decay of a Z boson to two oppositely charged electrons. To select events with this topology, an electron called "tag" passing tight selection criteria is required and an oppositely

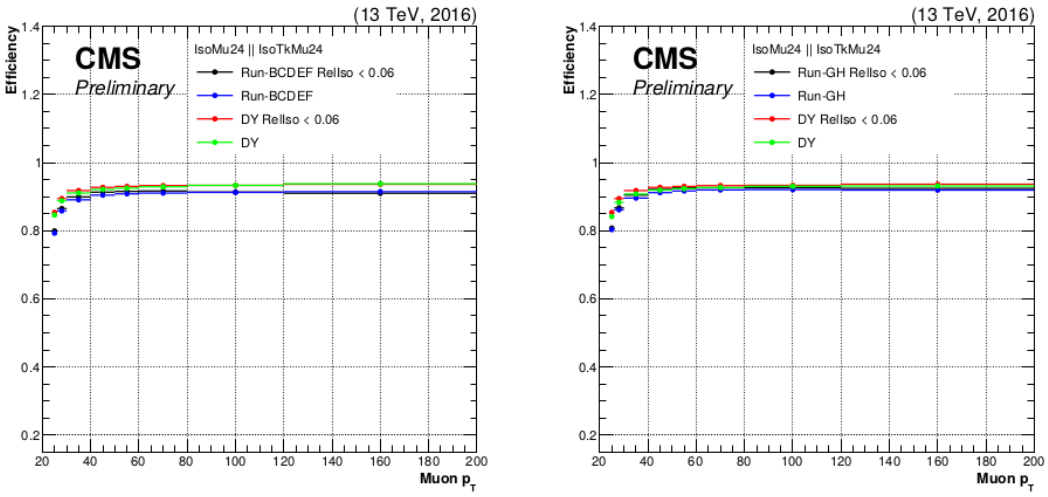


Figure 3.8: Muon trigger efficiency in bins in bins of p_T two subsamples of the Run 2 (Run-BCDEF and Run-GH) for $I_{rel} < 0.06$ super imposed with trigger efficiency with $I_{rel} < 0.15$ from Muon POG.

charged electron called "probe" passing looser criteria is necessary. The invariant mass of the dilepton-pair is restricted to be between $60 < m_{e^+e^-} < 120$ GeV$/c^2$. Splitting the probes into collections, where one contains electrons with an HLT match, i.e. the "passing probes", and one containing electrons without an HLT match, i.e. the "failing probes", allows to estimate the efficiency as in the muon case ($\epsilon = N_{passing\ probe}/N_{all}$). The scale factors were hence calculated by dividing the efficiency estimated in data to the simulation one: $SFs = \epsilon^{Data}/\epsilon^{MC}$. As is shown in Fig. 3.9 the trigger efficiency for p_T around the threshold is $\sim 85\%$ and the SF is around the unit.

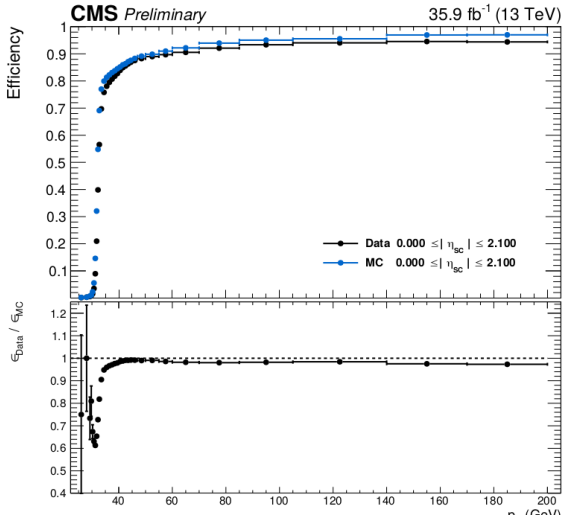


Figure 3.9: Distribution of the efficiencies in data and MC and the corresponding SF in dependence of the transverse momentum of the probe.

3.3.2 Primary vertex requirements

From the sample of triggered events, only those with at least one primary vertex reconstructed in the region $|d_z| < 24\,\text{cm}$, $|d_{xy}| < 2\,\text{cm}$ with respect to the nominal interaction point, are considered. Furthermore, a "good" primary vertex candidate has to be fitted with $ndof > 4$. A more detailed explanation of reconstruction of tracks and primary vertices can be found in [64].

3.3.3 Muon identification and reconstruction

Muon tracks are reconstructed independently both in the inner tracker (tracker tracks) and in the muon system (standalone muon tracks). Two approaches can be exploited for the reconstruction:

- The *Global* muon reconstruction (outside-in). A tracker track is associated to every standalone track by comparing the parameters of the tracks, and a *global muon track* is fitted by combining tracker hits and muon system ones and using the Kalman Filter method. For high p_T the global muon fit improves the momentum resolution, compared to the only-tracker one. This technique is designed to have an high efficiency for muons crossing several stations.
- The *Tracker* muon reconstruction (inside-out). Tracks with $p_T > 0.5\,\text{GeV}/c$ and $p > 2.5\,\text{GeV}/c$ are considered muon candidates and are extrapolated in muon system considering the magnetic field, the average energy losses and the multiple scattering in the detector. If at least a muon "segment" (i.e. a brief track whose hits are in one or more DT/CSC) correspond to the extrapolated track, it is associated to a muon candidate; this method is known as *track-to-segment matching*. This technique is more efficient for low momenta since it requires just a segment in the muon system.

Thanks to the high reconstruction efficiency both in the tracker and in the muon system, the $\sim 99\%$ of muons which fall within the muon system detectors geometric acceptance is reconstructed. The two techniques are often used together, since the combination of different reconstruction algorithms provides a more efficient and robust reconstruction.
For the identification of muons there are several identification criteria with different selection efficiency and fake or non-prompt muon rejection. The ones used in the analysis, are here presented.

Tight muon selection

The tight muon selection requires:

- a global reconstructed muon with a $\chi^2/ndof < 10$, where the chi-squared refers to the global track fit;
- at least one hit of a muon chamber considered in this fit;
- to have the Particle-Flow muon ID (see Sec. 3.2.1);
- muon segments must be in at least two muon stations;
- there must be at least 10 hits in the inner tracker (one of which must be in the pixel tracker),
- there is a cut on number of tracker layers with $hits > 5$;
- a transverse (longitudinal) impact parameter to be $|d_{xy}| < 2\,\text{mm}\,(d_z < 5\,\text{mm})$ with respect to the vertex;
- transverse momentum $p_T > 26\,\text{GeV}/c$;
- relative isolation $I^{\mu}_{rel} < 0.06$ (see section 3.2.2). This cut is applied in order to discriminate prompt muons (i.e. from hard interaction as Drell-Yan events or from W or Z bosons) from hadronic decays in jets;
- Pseudorapidity range $|\eta| \leq 2.4$.

Loose muon selection

The loose muon selection requires:

- global reconstructed or a tracker reconstructed muon;
- to have Particle-Flow muon ID (see section 3.2.1);
- to fall within $|\eta| < 2.4$;
- to have a transverse momentum $p_T > 10\,\text{GeV}/c$.

3.3.4 Electron selection

A specific electron track reconstruction is required and a specific treatment for the bremsstrahlung photons in order to associate them to the electrons and to not double count them; in particular, in CMS the reconstruction of electrons is done with the Gaussian-Sum Filter (gsf). Some quality criteria are applied based on nine variables through an optimized cut-based approach [65]. The cuts for these working points depend on several factors, such as running conditions, and are derived separately for each distinct data sample.

Tight electron selection

Tight electron selection requires:

- a "GsfElectron", i.e. an electron track fitted with the Gaussian-Sum-Filter that is matched to an active ECAL cell by the PF algorithm, with transverse energy $E_T > 35\,\text{GeV}$, $|\eta| < 2.1$, while excluding the ECAL barrel-endcap transition region of $1.44. < |\eta| < 1.57$;
- transverse momentum $p_T > 32\,\text{GeV}/c$;
- $I^e_{rel} < 0.06$ (see section 3.2.2). This cut is applied in order to discriminate prompt muons (i.e. from hard interaction as Drell-Yan events or from W or Z bosons) from hadronic decays in jets;
- Pseudorapidity range $|\eta| \leq 2.1$.

. The "tight" working point yields an efficiency of $80 - 90\%$ for electrons above $E_T > 45\,\text{GeV}$.

Veto electron selection

Veto electron candidate have to pass a cut-based quality selection similar to the tight electron selection defined above whereas the kinematic requirements are $E_T > 15\,\text{GeV}$, $|\eta| < 2.5$.

This analysis then uses the cut-based "electron veto" working point yielding an efficiency of 98% at a plateau which is reached around $E_T > 45\,\text{GeV}$.

3.3.5 Leptons multiplicity selection

An event is selected if there is one lepton passing the tight selection (muon OR electron) and no other leptons passing neither the tight or loose selections.

3.3.6 Jet selection

The jets are required:

- to be **PF-CHS jets**: jet clusterd from the PF candidates, with the pile-up reduction technique of Charge Hadron Subtraction [66], where PF candidates from non primary vertices are removed before clustering; the clustering algorithm used is the **Anti-Kt algorithm**, with the requirement on the cone-size of $R = 0.4$. For more detailed information on this clustering algorithm, see [67];
- transverse momentum $p_T > 40\,GeV$;
- jet distance from the lepton $\Delta R > 0.4$.

3.3.7 The *b*-tagged jets

The *b*-tagging consists in the identification of jets originating from *b* quarks; *b*-tagging algorithms exploit displaced tracks, the presence of secondary vertices or a combination of the two kinds of information.

In this analysis a **combined multivariate analysis (cMVA) tagger** is used as the b-tagging algorithm. It exploits the combined secondary vertex algorithm, matching the track-based lifetime information with secondary vertices inside the jet to provide a MVA discriminator for b-jet identification.

3.3.8 Event selection

In this analysis, final state topologies are selected in order to have a signal enriched sample, where a single top-quark is produced and decays in a d, s, or b quark and a W boson that decays in a lepton-neutrino pair.
The Feynman diagram of signal events are shown in Fig. 3.1. Three different signals are considered:

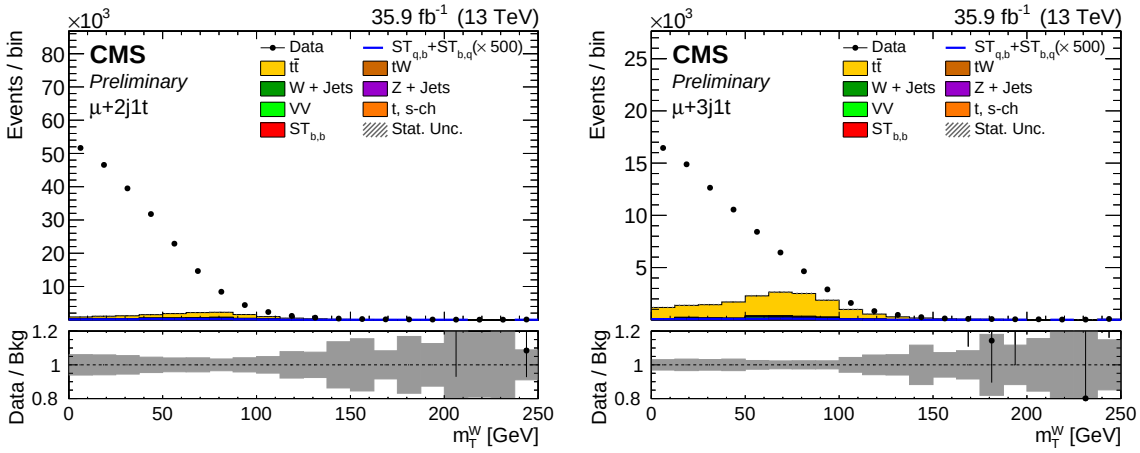
- **S1** or $\mathbf{ST_{bb}}$ for the top quark, and $\mathbf{S\overline{T}_{bb}}$ for the anti-top anti-quark. The top quark is produced via the exchange of a W boson between the b quark (from a sea gluon producing a $b\bar{b}$ couple) and the spectator quark, and decays in a couple bW: therefore the b quark is both in the production and in the decay vertices, as shown in Fig. 3.1(a),;
- **S2** or $\mathbf{ST_{bq}}$ for the top, and $\mathbf{S\overline{T}_{bq}}$ for the anti-top anti-quark. In this case the top is produced via the same interaction of S1, but decays in a d or s quark and a W boson. The Feynman diagram of this process is represented in Fig. 3.1(b);
- **S3** or $\mathbf{ST_{qb}}$ in the top quark case, and $\mathbf{S\overline{T}_{qb}}$ for the anti-top anti-quark. This process has a d or s quark in the top-quark production vertex, and then the top quark decays in a bW pair. Since S3 is not distinguishable from S1 when sea quarks are considered, it is restricted to the interaction of the d quark (in particular the valence one) with a spectator quark, which is usually another valence quark. This signal, defined $\mathbf{ST_{d_vb}}$ ($\mathbf{S\overline{T}_{d_vb}}$ in the anti-top anti-quark case), has different kinematic features from the ST_{bb} . The ST_{qb} and ST_{d_vb} Feynman diagrams are represented in the left and right panels of Fig. 3.1(c), respectively.

Single-top-quark via s-channel and via tW associated production are present in the sample, but they are not considered as signals in this analysis.
Events with one muon in the final state and with different jet multiplicity of jets and b-jet multiplicity are selected: **3 jet - 1 tag** (3j1t) events, where there are three jets passing the jet selection, one of which passes the b-tagging requirements too; the **2 jet - 1 tag** (2j1t) events, where there are two jets passing the jet selection, one of which passes the b-tagging requirements too.

In addition to this baseline selection, other requirements have been performed in order to reduce the background contributions.

3.4. Estimation of QCD multi-jet background from data

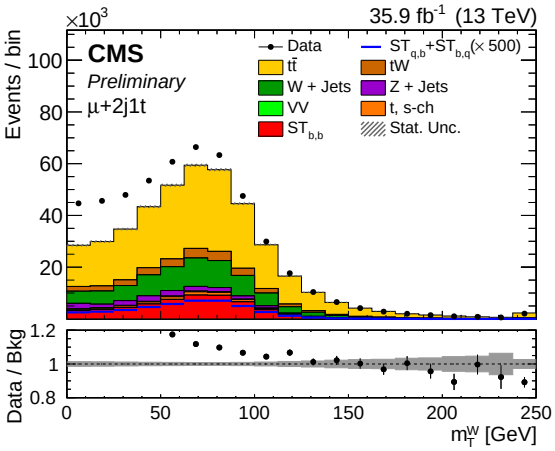
The large cross section and rejection power of the final event selection for the QCD multi-jet processes make it impossible to generate samples of simulated events in reasonable size. To overcome this problem, this background contribution is suppressed as much as possible and the remaining contamination is estimated directly with a data-driven technique; furthermore, selected events usually populate the tail of kinematic distributions, thus it is important to control this background from data. The estimation is performed in the anti-isolated lepton region ($I_{rel} > 0.15$), since it is QCD enriched, as is shown in Fig. 3.10. Furthermore, this region is orthogonal to the selected one for the analysis ($I_{rel} < 0.06$).



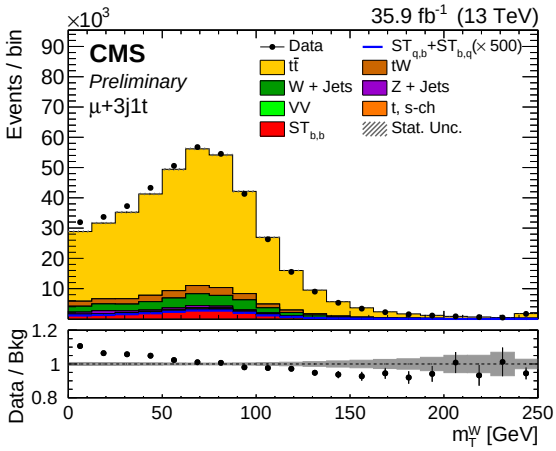
(a) 2j1t (b) 3j1t

Figure 3.10: Distribution of data and Monte Carlo samples in the muon-antiisolated region.

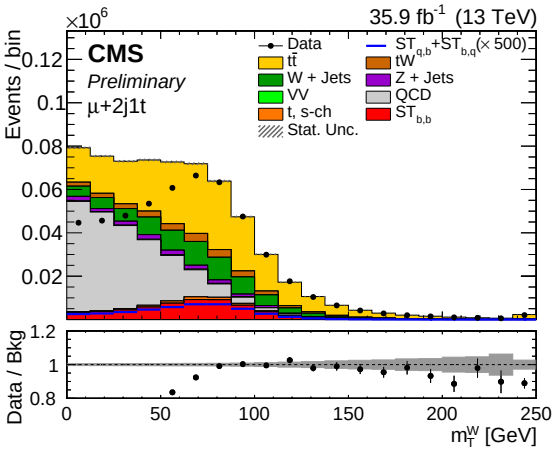
Figure 3.11 shows the distributions of the data with all the MC samples (the so-called *stacked plot*) without the QCD contribution, in 2j1t and 3j1t channels, where the lepton is a muon. The variable m_W^T provides the highest separation power between QCD multi-jet events and other processes including the t-channel single-top-quark process. The m_T^W variable is also used to estimate the remaining contribution by fitting their distributions over the entire range, i.e., without applying the described requirement. A maximum likelihood fit is performed on the distribution using two probability distribution functions, one for the QCD multi-jet process and one for all other processes. The latter distribution is obtained by adding the different non-QCD contributions from simulation, including the t-channel signal, according to their theory predictions, while the former is derived from a side-band region enriched in QCD multi-jet events. The resulting distributions before and after the fit for the 3j1t and 2j1t channels are shown in Fig. 3.13. The distributions of m_T^W after the fit are shown in Fig. 3.13. Good agreement between the fit results and the data is found. In addition to the baseline selection discussed in Sec. 3.3.8, a cut on m_T^W is performed: $m_T^W > 50\text{GeV}/c^2$.



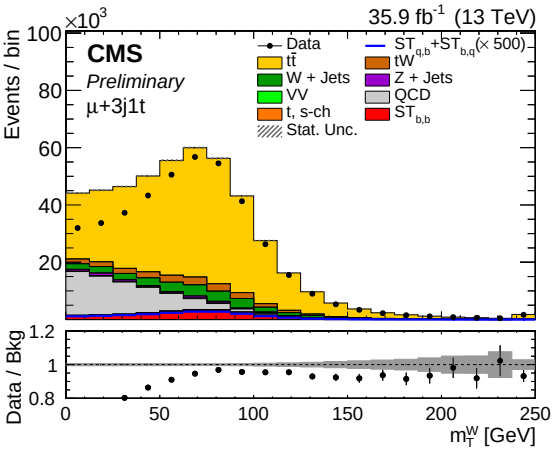
(a) 2j1t without QCD contribution



(b) 3j1t without QCD contribution

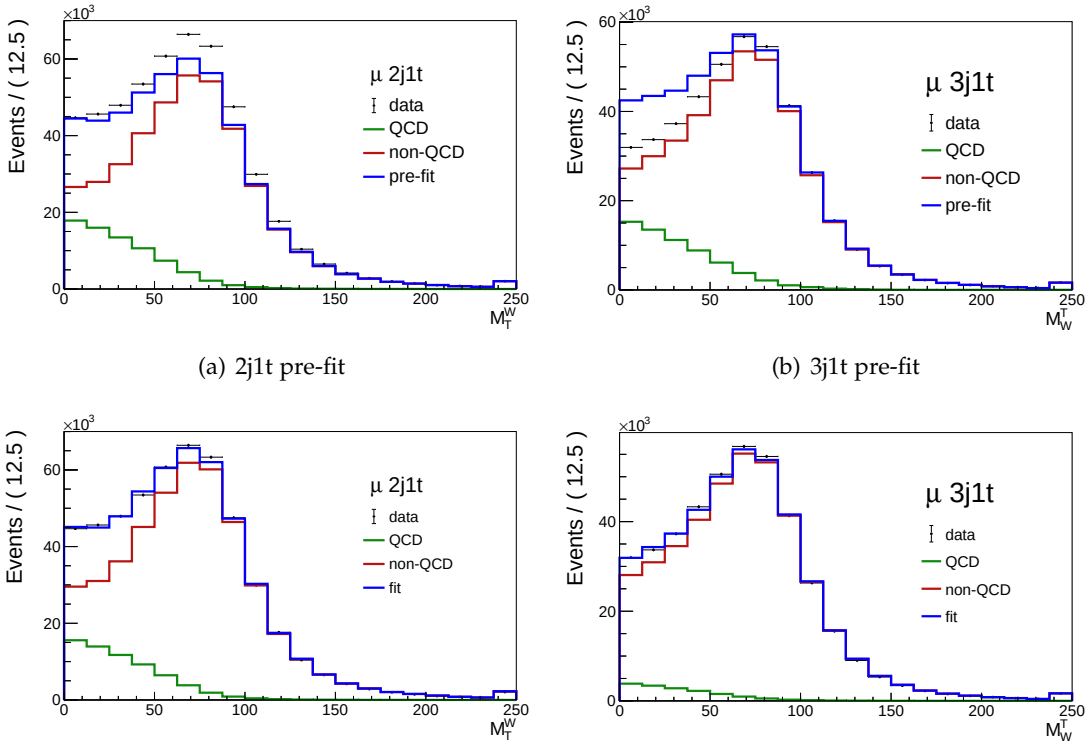


(c) 2j1t with QCD contribution pre-fit



(d) 3j1t with QCD contribution pre-fit

Figure 3.11: Estimation of QCD background from the anti-isolated muon region. The upper panels show the comparison between data and simulation without the QCD contribution, and the lower ones contain the QCD estimation before applying the fit to the signal region.



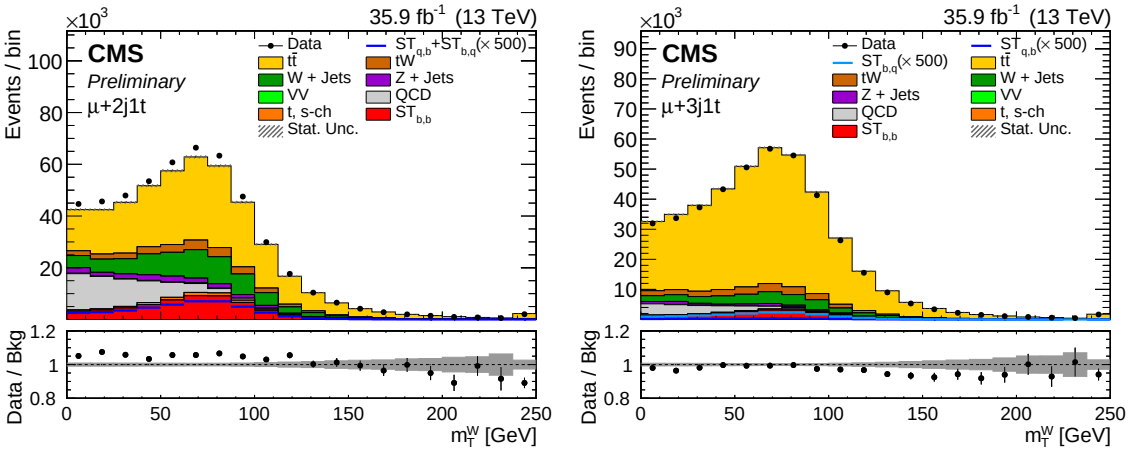
(a) 2j1t pre-fit

(b) 3j1t pre-fit

(c) 2j1t post-fit

(d) 3j1t post-fit

Figure 3.12: Fit to the m_T^W distribution for events with muons in the 2j1t (left), 3j1t (right) before and after the fit, respectevely in the upper and lower panels. The QCD template is extracted from a sideband region in data. For the fit, only statistical uncertainties are considered.



(a) 2j1t with QCD contribution post-fit

(b) 3j1t with QCD contribution post-fit

Figure 3.13: Distribution of m_T^W after the fit.

4 The top-quark reconstruction

This chapter focuses on the selection of the lepton-jet pair coming from the top quarks. Figure 3.1 shows the Feynman diagrams of the events of interest: single-top-quark produced by the interaction of a quark with a sea (or valence) quark. The quark that recoils against the top quark is called "spectator" quark. The decay processes are $t \to Wq$, and then $W \to \ell\nu$. So the final states of interest contain: one lepton-neutrino pair coming from the W decay and at least two jets, coming from the spectator quark and from the tWq vertex).
Leptons and jets are reconstructed in the detector, whereas the neutrino's observables can be inferred from the Missing Transverse Energy.
There are two different sources of background that simulate the processes of interest:

- *Physical background*, due to events that have the same final state configuration with respect to the signal, but they are produced by a different physical mechanism: the main sources of this background are the $t\bar{t}$ pair production, the W/Z production associated with jets;
- *Combinatorial background*: this is related to the choice of the lepton-jet pair in the events. Although there is just one selected lepton, the jet multiplicity varies from 2 to 3, so a wrong selection of the jet for the top-quark reconstruction leads a contribution to this background.

Two different approaches to reduce the combinatorial background will be presented and compared in this chapter.

4.1. Combinatorial background

Section 3.3 describes the final state topologies selected to get a signal-enriched sample. In the **3 Jets - 1 Tag** (3j1t) final state configuration, the selected jets are: the *b-jet*, that passes *b*-tagging requirements; the *forward jet*, the one that is supposed to come from the spectator quark, that goes preferentially along the direction of the incoming proton; finally, the *extrajet*, the third, less energetic jet, that, in signal events, is supposed to come from the gluon-splitting.
In the **2 Jets - 1 Tag** (2j1t) final state configuration, the selected jets are the same of the 3j1t ones. In this topology, there is an *extrajet* if another jet passes looser requirements with respect to the ones discussed in Sec. 3.3.6, but it will not be considered in this analysis. For each jet in both the categories there is an associated reconstructed top-quark candidate.
The combinatorial background arises from the wrong lepton-jet pair selection. Figure 4.1(a) shows the top-quark reconstruction with the correct jet, i.e. the one coming from the top-quark decay; Fig. 4.1(b) and Fig. 4.1(c) show the wrong selection of the jets, that are coming from the gluon-splitting (GS) and the spectator quark, respectively. To deal with this background source the basic idea is to select variables with high discrimination power in order to separate the correct lepton-jet pair from the wrong ones.

4.2. Variables of interest

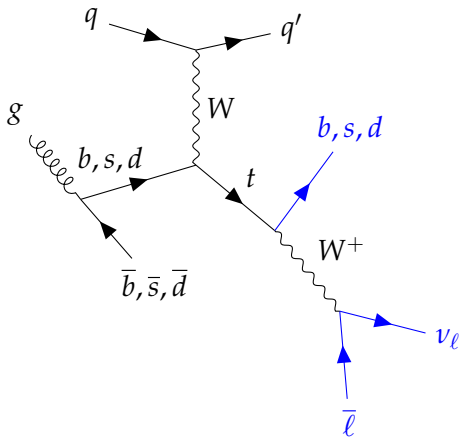
In order to perform an unbiased selection, the reconstructed top quark is selected without looking at its kinematic variables or at the jets *b*-tagging, but considering only:

- **jet kinematic variables**: transverse momentum $p_{T_{jet}}$, pseudorapidity η_{jet};
- **lepton variables**: charge Q_{lep}, transverse momentum $p_{T_{lep}}$, pseudorapidity η_{lep};
- **lepton-jet pair variables**: distance $\Delta R_{lep-jet}$, invariant mass $m_{lep-jet}$;
- **global variables**: W reconstructed transverse mass (defined in Sec. 3.2.4) m_T^W, missing transverse energy.

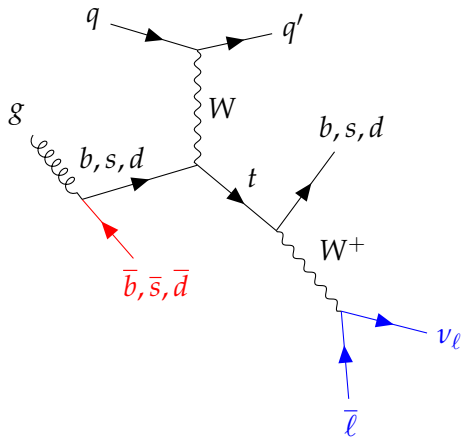
The idea is to re-organize the jets for each event by the probability of the jet to be originated from a top, the *jet score*. This has been done in two different ways:

- taking a 2D distribution from Monte Carlo (MC) samples, using the two observables that achieve the best separation between signals and combinatorial background, and then evaluating the likelihood ratio in the bakground and signal hypotheses. This is exploited to order the jets, and to compute the ROC curve, as it is explained in Sec. 4.3;
- with a multivariate discriminant, based on the variables listed above, obtained with different machine learning techniques (a Boosted Decsion Tree and an Artificial Neural Network), discussed in Sec. 4.4.

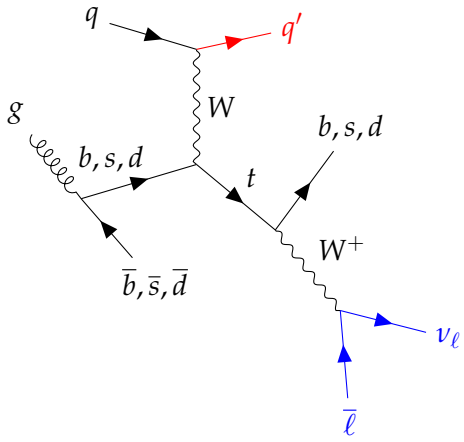
The two approaches are then compared and the jet score is computed with the method that achieves the best performances.



(a) Jet associated to the tWq vertex quark.



(b) Jet associated the gluonsplitting quark.



(c) Jet associated to the spectator quark.

Figure 4.1: Possible jet combinations for the top-quark reconstruction in signal events. In Fig. 4.1(a) it is shown the correct lepton-jet combination, whereas in Fig. 4.1(b) and Fig. 4.1(c) the chosen (red lines) jets are not associated to the top quark, but they are from the gluon-splitting process and the spectator quark associated jet, respectively.

4.3. Jet classification with the likelihood ratio

The first step in order to separate the signal from the combinatorial background is to identify the variables that operate the best discrimination, in order to build a 2D likelihood, the distribution of the events in the 2D plane in those variables. The variables among those mentioned in Sec. 4.2 that achieve the best separation are shown in Fig. 4.2, and are found to be ΔR and $m_{lep-jet}$. The background contributions are divided in two components: the jet coming from the GS quark and all the other jets, that are mainly the forward jet other jets that are not part of the event.

After the variables selection, a 2D likelihood in ΔR and $m_{lep-jet}$ has been built both in the signal hypothesis and in the background one, as it is shown in Fig. 4.3. In Sec. A there are the 2D likelihood distribution under the two different background hypotheses too.

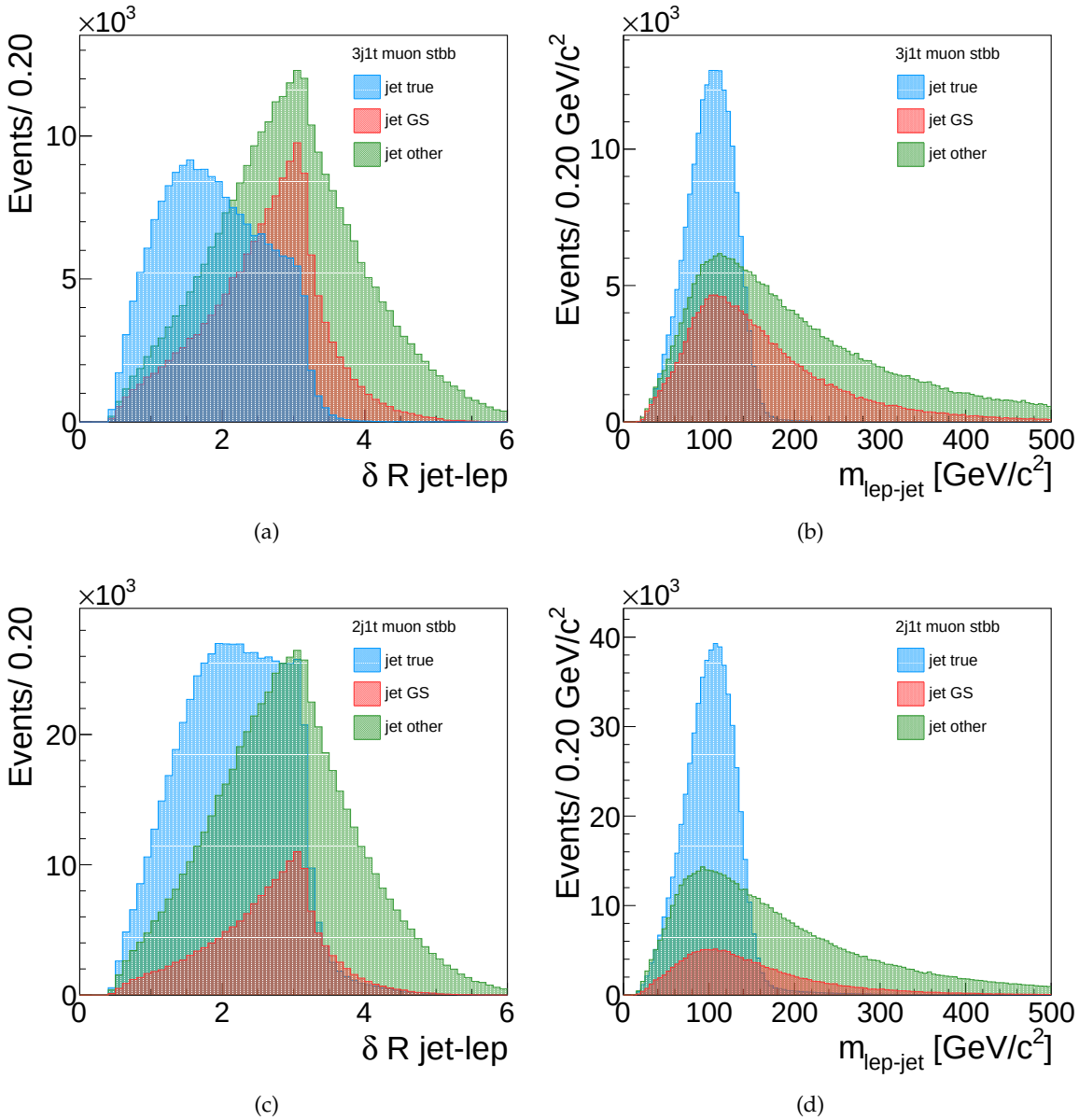


Figure 4.2: Distributions of the variables used to build the 2D likelihood: $\Delta R_{jet-lep}$ in the left panels, lepton-jet invariant mass in the right panels, in the 3j1t (upper panels) and 2jt (lower panels) final state configurations, for the ST_{bb} sample.

When the top quark is produced by the interaction of a sea quark (d, s, or b) with a valence one, the top-quark decay products are supposed to have the same kinematic features, up to minor differences in the sea quarks parton distribution functions (PDFs). This assumption does not apply in the ST_{db} case where the d quark is one of the valence quarks, since they have a very different Bjorken x spectrum with respect to sea quarks, thus the kinematic distributions of final state particles are different compared to the cases when a sea quark is involved. Moreover, the ST_{db} component is less relevant with respect to the ST_{bb} one. In this analysis kinematic differences of the two signals are thus neglected and the 2D likelihood is chosen to be the distribution in the $\Delta R_{jet-lep}$ - $m_{lep-jet}$ plane for the ST_{bb} sample.

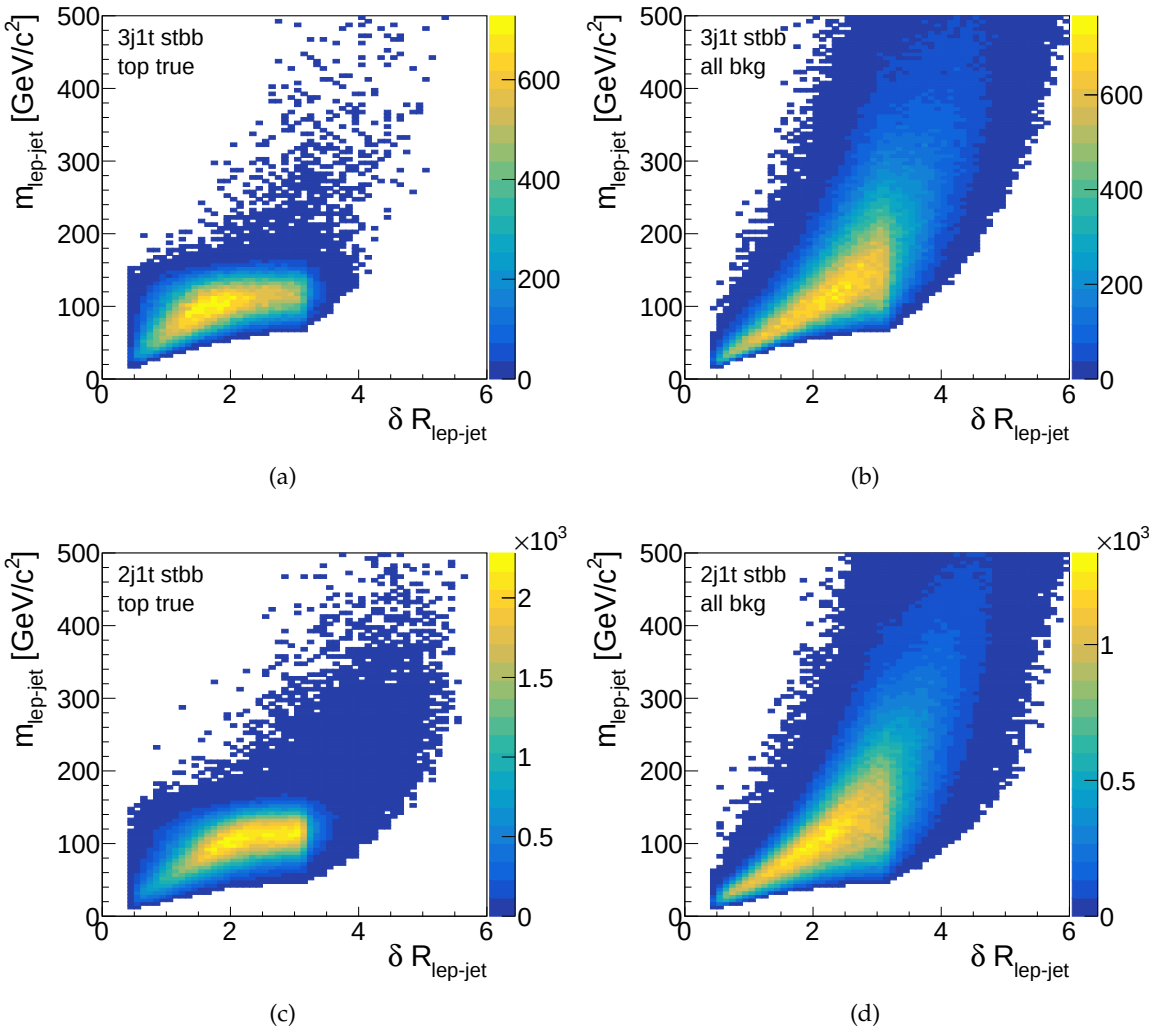


Figure 4.3: Likelihood 2D under the H_1 hypothesis (left panels) and H_0 hypotesis (right panels) for the 3j1t final state topology (upper panels) and for the 2j1t final state topology (lower panels).

The performance of a selection criterion can be considered optimal if it achieves the smallest mis-identification probability for a fixed value of the selection efficiency [68]. According to the Neyman-Pearson lemma [69], test statistic that achieves the best separation between two hypotheses, H_0 and H_1, is given by the ratio of the likelihood functions $\mathcal{L}(\Delta R_{lep-jet}, m_{lep-jet}|H_1)$ and $\mathcal{L}(\Delta R_{lep-jet}, m_{lep-jet}|H_0)$.

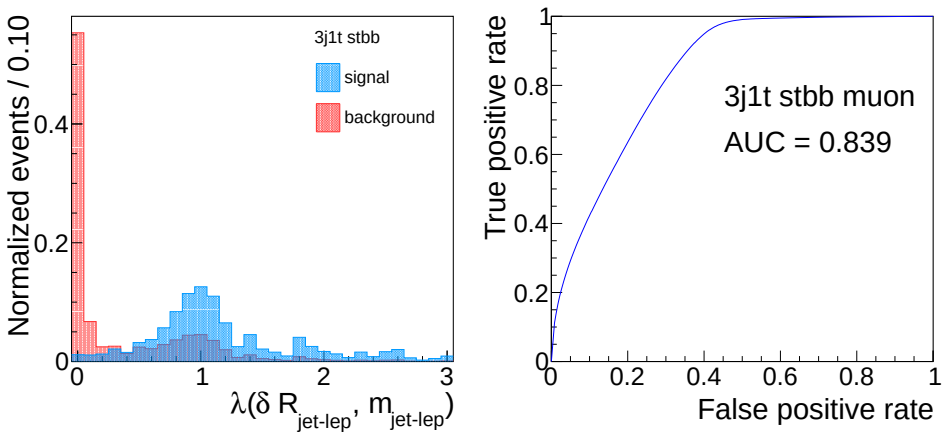
$$\lambda_{ST_{bb}}(\Delta R_{lep-jet}, m_{lep-jet}) = \frac{\mathcal{L}(\Delta R_{lep-jet}, m_{lep-jet}|H_1)}{\mathcal{L}(\Delta R_{lep-jet}, m_{lep-jet}|H_0)},$$

where $\mathcal{L}$ is the 2D likelihood for the ST_{bb} channel, H_1 denotes the signal hypothesis, i.e. the jet is a top-quark decay product, and in H_0 denotes the background hypothesis, corresponding to a jet that is not a top-quark decay product.
For each jet in the events and for each sample, $\lambda_{ST_{bb}}(\Delta R_{lep-jet}, m_{lep-jet})$ has been computed: it corresponds to the probability for a jet to be originated from a top.
A Receiver Operating Characteristic (ROC) curve has been computed: it represents the true positive rate (the signal efficiency), versus the false positive rate (the mis-identification probability). The ROC curve has been plotted in order to eventually select the working point and to compare it with machine-learning output based selections (see Sec. 4.4). Figure 4.4 shows the likelihood ratio and the ROC curve for the ST_{bb} sample with a muon in the final state. The likelihood ratio distributions of some other samples involving top-quark production are in App. A.

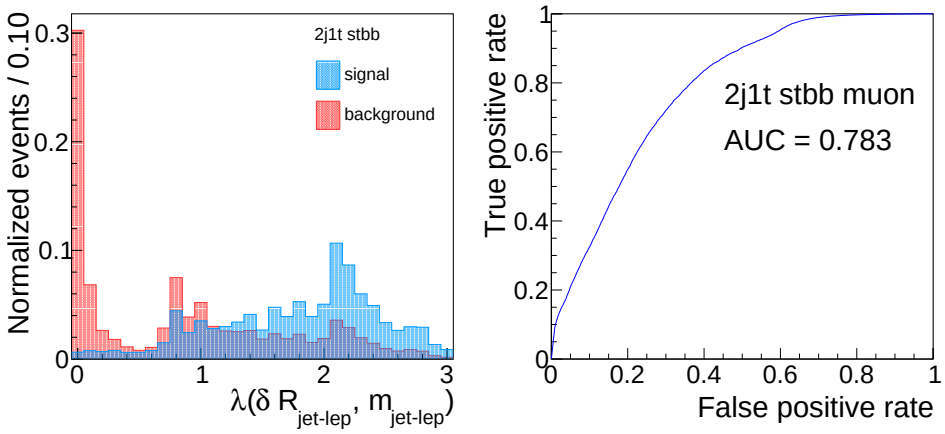
4.4. Multi-variate jets classification

The jet selection with the likelihood ratio could be in line of principle further optimized by evaluating the likelihood in a more variables. This is called *multivariate analysis* (MVA). Nevertheless, this requires to build multi-dimensional histograms and the number of bins for each one explodes exponentially with the increase of the dimensions. Therefore, the number of events for each bin becomes too small to give an accurate estimation of the likelihood. This phenomenon is known as the *curse of dimensionality*.

The Neyman-Pearson lemma sets an upper limit to the performance of any possible multivariate discriminant: the exact evaluation of the likelihood function in the two hypotheses H_0 and H_1 is not always possible, and approximate methods have been developed in order to approach the performance of an ideal selection based on the multi-dimensional likelihood ratio, the limit set by the Neyman-Pearson lemma. Among those approximate methods there are **machine-learning algorithms**, such as Artificial Neural Networks and Boosted Decision Trees [68].



(a) 3j1t configuration



(b) 2j1t configuration

Figure 4.4: Likelihood ratio distribution for the signal and the background events (left panels) and corresponding ROC curves by varying the cut (right panels) for the ST_{bb} data sample, in 3j1t and 2j1t configurations (upper and lower panels, respectively).

4.4.1 Machine-Learning algorithms

Machine learning (ML) is a sub-field of artificial intelligence that involves the development of self-learning algorithms, whose goal is to gain knowledge from the data sample and to make predictions [70]. There are three types of machine learning algorithms:

- *Supervised learning*: it learns a model from labeled training data that allows to make predictions about unseen data. The term "supervised" refers to a set of samples where the desired output signals (labels) are already known. Supervised learning is well suited in *classification* or *regression* problems, where the predictions are discrete or continuous, respectively. The *model* in supervised learning usually refers to the

mathematical structure of by which the prediction y_i is made from the input x_i.
- *Reinforcement learning*, where the goal is to develop a system (named *agent*) that improves its performance based on interactions with the *environment*. Through the interaction with the environment, an agent can then use reinforcement learning to learn a series of actions that maximizes this reward via an exploratory trial-and-error approach or deliberative planning.
- *Unsupervised learning*: it is used with unlabeled data or data of unknown structure, in order to extract meaningful information without the guidance of a known outcome variable or reward function.

The prediction value can have different interpretations, depending on the task, i.e., regression or classification. The discriminant output value can be taken as a test statistic too, and in this case it can be adopted to perform the signal selection.
In this part of the analysis, classification supervised learning is used to classify jets coming from a top quark in the data sample knowing the MC truth for the signal samples.

Classification It is a subcategory of supervised learning, where the goal is to predict the labels of new samples based on past observations, among discrete possible labels that identify different classes. The algorithm receives as input a set of discriminating variables, each of which individually does not allow to reach an optimal selection power, then it computes an output that combines the input variables.

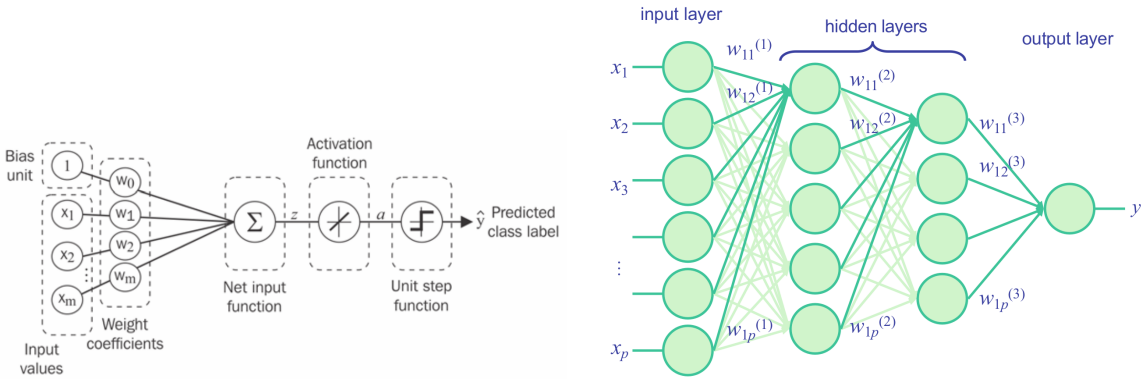
The objective function The task of training the model consists in finding the parameters set $\vec{\theta}$ that best fits the training data $\vec{x}_i$ and known labels y_i. In order to train the model, an objective function that measures how well the model fits the training data needs to be defined.
The objective function is divided in two parts, training loss and regularization term:

$$\text{obj}(\theta) = L(\theta) + \Omega(\theta)$$

where L is the training loss function and Ω is the regularization term. The first measures how predictive the model is with respect to the training data, whereas the regularization term controls the complexity of the model, in order to avoid overfitting, that occurs when the ML algorithms learns some specific features of the training sample that are not general and therefore is not performing in the prediction as well as in the training. The trade-off between the two is also referred as *bias-variance trade-off*.

4.4.2 Artificial Neural Networks

Artificial neural networks (ANN, or simply NN) are supervised machine learning algorithms vaguely inspired by the biological neural networks that constitute animal brains. There are single-layer and multi-layers NN. A schematic view of the principle of working of a NN is shown in Fig. 4.5.



(a) single-layer (b) multi-layer

Figure 4.5: Scheme of neural networks with one (left panel) or more (right panel) hidden layers.

The basic example of a NN machine learning algorithm is the single layer NN (Fig. 4.5(a)), but its performances substantially increase if some hidden layer is added (Fig. 4.5(b)); although the price to pay is in the complexity of the algorithm. In order to avoid overtraining, while building a NN, a trade-off between the parameters of a NN (such as the number of hidden layers) and the efficiency of the algorithm must be found and chosen as working point.

One of the main Python-based libraries that implement a great variety of tools to build a NN is `keras` [71].

Single layer neural network

A schematic view of a single layer neural network can be found in Fig. 4.5(a): here, a linear combination of the features x_i with the coefficients w_i, $z = w_0 + \sum_{i=1}^{m} w_i x_i$ is taken as input for a so-called *activation function* $f(z)$. The weights are initialized to certain values, the output value $\hat{y}$ is computed through the *activation function* and it is then compared to the known y value. This is done in order to update the weights $w_j \rightarrow w_j + \Delta w_j$, with $\Delta w_j = \eta(y^{(i)} - \hat{y}^{(i)})x_j$, where η is the learning rate, a constant parameter in the range $[0, 1]$, $y^{(i)}$ is the known outcome of the i^{th} training sample and $\hat{y}^{(i)}$ is the algorithm prediction.

The main activation functions used to build a NN are:

- *step function* : $f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \geq 0 \end{cases}$
- *linear* $f(x) = x$. This is often coupled with the the step function since its range is $\mathbb{R}$, especially when the label is binary;
- *sigmoid or logistic* function: $f(x) = \frac{1}{1+e^{-x}}$, in the range $[0, 1]$. It is used often for models that aim to predict the probability as an output;
- *hyperbolic tangent*: $f(x) = tanh(x)$. The shape is similar to the sigmoid, but its range is

$[-1, 1]$. It is often used for binary classification;
- *Rectified Linear Unit (ReLU)*: $f(x) = max(0, x)$. It is the most used activation function; its range is $[0, \infty)$.

The objective function is divided into the training loss term and the regularization term. The loss function varies with the machine learning algorithms. For the neural networks, the most common loss function used are:

- the sum of squared errors (SSE): $J(w) = \frac{1}{2} \sum_{i=1}^{n} (y^{(i)} - \hat{y}^{(i)})$;
- the logistic regression loss function: $J(f(z), y; w) = \sum_{i=1}^{n} -log(f(z^{(i)})) - (1 - y^{(i)})log(1 - (f(z^{(i)})))$. This can be minimized through the gradient descendent method [70].

Multi layer neural network

To build a multi layer neural network one has to choose:

- the *model*: in this analysis the *sequential model* is chosen: it is a linear stack of layers.
- the model needs to know what *input shape* it should expect. For this reason, the first layer in a sequential model (and only the first, because following layers can do automatic shape inference) needs to receive information about its input shape;
- the *single-layer parameters*: for this analysis *dense layers* are chosen. This means that each neuron in a layer receives an input from all the neurons present in the previous layer; thus, theyre densely connected. The values of the layers parameters used in this analysis are listed in Tab. 4.1;
- the *learning rate*, chosen to be 0.001;
- the *loss function*: for this analysis the loss function is the *binary cross entropy*;
- the *metric*: the chosen one is the *accuracy metric*.

Other adjustable single-layer parameters are set to their default values in this analysis.

Parameter	Type	Layer	Chosen value
units	int	1	12
		2 , 3	8
		4	1
activation function	string	1	ReLU
		2 , 3	ReLU
		4	sigmoid

Table 4.1: List of some ANN parameters' choices for each layer.

4.4.3 Boosted Decision Tree in XGBoost

Decision trees

The decision tree model learns from a series of questions the features of the training set, and exploits those information to the infer the class labels.

The decision algorithm starts at the tree root and splits the data on the most informative feature. In an iterative process, this splitting procedure is repeated until all the samples in each node belong to the same class. In practice, this can result in a very deep tree with many nodes, which can easily lead to overfitting. Thus, one typically wants to prune the tree by setting a limit for the maximal depth of the tree that achieves a good separation without the risk of overfitting.

In the decision trees algorithms, the objective function is the *information gain*:

$$IG(D_p, f) = I(D_p) - \sum_{j=1}^{m} \frac{N_j}{N_p} I(D_j),$$

where f is the feature that performs the split, $D_{p,j}$ are the dataset and $N_{p,j}$ are the samples of the parent and the j^{th} child node, respectively, I is the impurity measure. It is the difference between the impurity of the parent node and the sum of the child node impurities (the lower the impurity of the child nodes, the larger the information gain) and it has to be maximized at each split. For the sake of simplicity and to reduce the combinatorial search space, most libraries implement binary decision trees.

XGBoost Classifier

The model choice of XGBoost (XGB) Classifier is the *decision tree ensemble*, a set of **classification and regression trees (CART)** [72]. In decision trees the leaf only contains decision values, whereas in CART a real score is associated with each of the leaves, which gives richer interpretations that go beyond classification. Usually, a single tree is not strong enough to be used in practice. What is actually used is the ensemble model, which sums the prediction of multiple trees together. The prediction scores of each individual tree are summed up to get the final score. The model can be written as follows:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i)\,,\ f_k \in \mathcal{F},$$

where K is the number of trees, f is a function in $\mathcal{F}$, the *functional space*, that is the set of all the possible CARTs.

The objective function to be optimized is given by:

$$\text{obj}(\theta) = \sum_{i=0}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{4.1}$$

One of the most used loss function is the mean squared error (MSE) function. Possible ways to define the regularization functions are given by parameters such as the number of notes in the tree or the depth and the L2 norm of leaf weights.

Additive training: boosting Equation 4.1 does not answer the problem to how train the trees. To do it, one must know the parameters of the trees, and therefore needs to learn the functions f_i in Eq. **??**, each containing the structure of the tree and the leaf scores. Learning tree structure is much harder than traditional optimization problem where you can simply take the gradient; an additive strategy is used: fix what one has learned, and add one new tree at a time. The prediction value at step t as $\hat{y}_i^{(t)}$ is given by:

$$\begin{aligned}
\hat{y}_i^{(0)} &= 0 \\
\hat{y}_i^{(1)} &= f_1(x_i) = \hat{y}_i^{(0)} + f_1(x_i) \\
\hat{y}_i^{(2)} &= f_1(x_i) + f_2(x_i) = \hat{y}_i^{(1)} + f_2(x_i) \\
&\dots \\
\hat{y}_i^{(t)} &= \sum_{k=1}^{t} f_k(x_i) = \hat{y}^{(t-1)} + f_t(x_i).
\end{aligned}$$

For each step the chosen tree optimizes the objective function:

$$\begin{aligned}
\text{obj}^{(t)} &= \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t)}) + \sum_{i=0}^{t} \Omega(f_i) \\
&= \sum_{i=1}^{n} l(y_i, (\hat{y}_i^{(t-1)} + f_t(x_i))) + \Omega(f_t) + const\ .
\end{aligned}$$

The goal is thus to find f_t to minimize the objective function. By developing in Taylor series the objective function one gets:

$$\text{obj}^{(t)} = \sum_{i=1}^{n} \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t),$$

where $g_i = \partial_{\hat{y}^{t-1}} l(y_i, \hat{y}^{(t-1)})$, $h_i = \partial^2_{\hat{y}^{t-1}} l(y_i, \hat{y}^{(t-1)})$. Even if it seems to be way more complicated then just grow trees, this method gives more control on what we are learning and it ensures the convergence.
For what concerns the regularization term, a good definition is:

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^{T} w_j^2, \tag{4.2}$$

where the first term is the number of leaves multiplied for a certain parameter γ that is adjustable, and the second part is the L2 norm[1] for the leaf scores.

[1] For more detailed information on the L2 norm regularization see [70].

XGBoost Classifier parameters

The XGBoost Classifier is trained on the ST_{bb} sample and returns the output for the data and all the samples. It has a large variety of parameters that can be adjusted in order to achieve the best performances. In Tab. 4.2 are reported the main ones, their meaning and the choices for them.

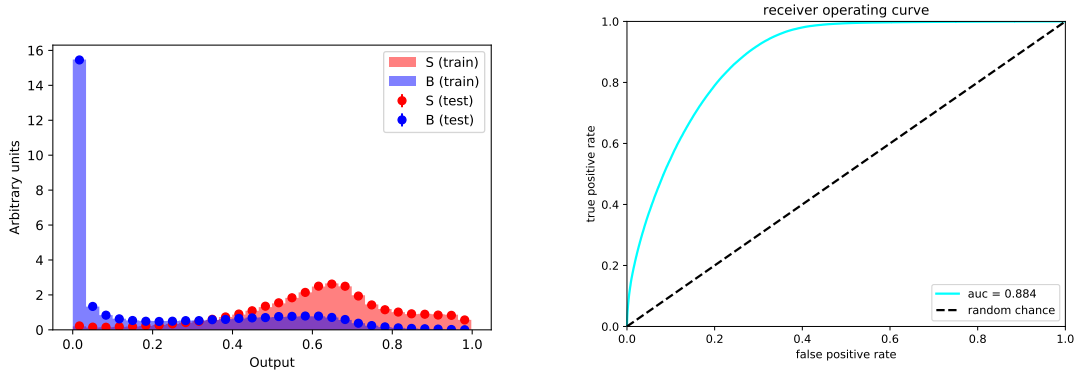
Parameter	Type	Chosen value	
max depth	integer	4	maximum tree depth for base learners
learning rate	floating in $[0,1]$	0.32	reweighting for trees correction[2]
number of estimators	integer	100	number of trees to fit
objective	string	binary logistic	objective function to be used [3]
γ	floating	0	the parameter γ in Eq. 4.2
minimum child weight	integer	0.4	minimum sum of instance weight(hessian) in a child[4]
reg α	floating	0	L1 regularization term coefficient
reg λ	floating	0	L2 regularization term coefficient

Table 4.2: List of the main parameters of the XGB Classifier algorithm with the values adopted in this analysis.

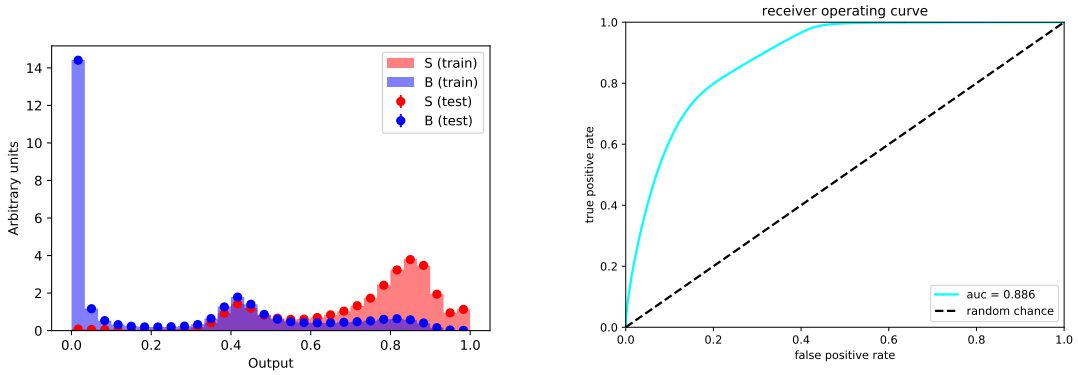
[2] A large value of the learning rate can produce overfitting.
[3] It can be callable too, if one wants a custom objective function.
[4] The greater is, the more conservative the algorithm is.

4.4.4 Machine learning selection results

In Fig. 4.6 and Fig. 4.7 are shown the XGB Classifier and the NN outputs and their respective ROC curves. In both the cases the training has been performed on the ST_{bb} and $S\overline{T}_{bb}$ samples, as well as the evaluation of the likelihood ratio. The machine learning algorithms output distributions and their respective ROC curves for some other samples involving top-quark production are in App. A.

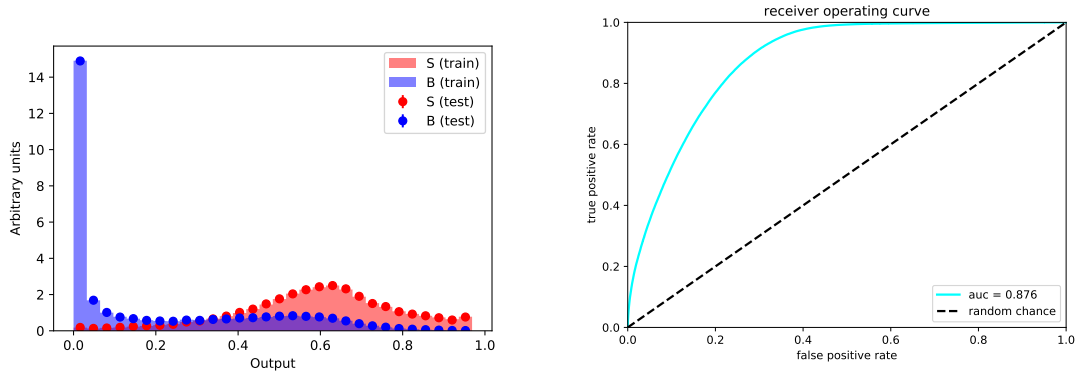


(a) 3j1t configuration

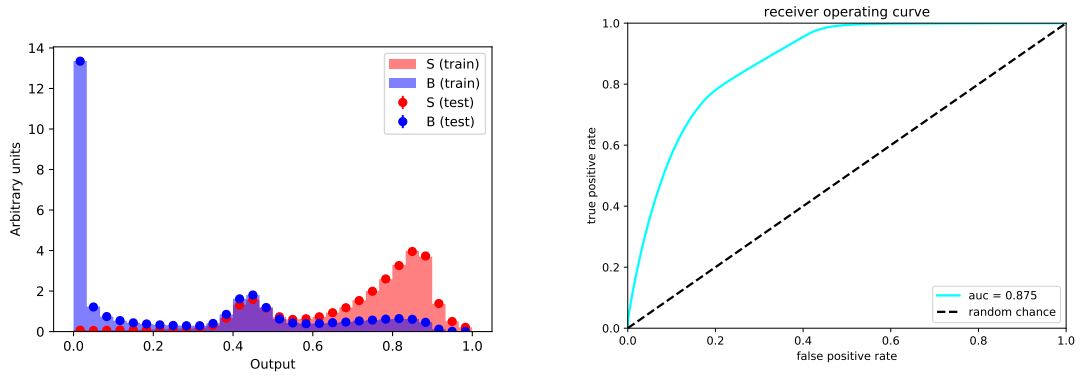


(b) 2j1t configuration

Figure 4.6: XGB Classifier output distribution for the signal and the background events (left panels) and corresponding ROC curves (right panels) for the ST_{bb} data sample, in 3j1t (upper panels) and 2j1t (lower panels) final states.



(a) 3j1t configuration



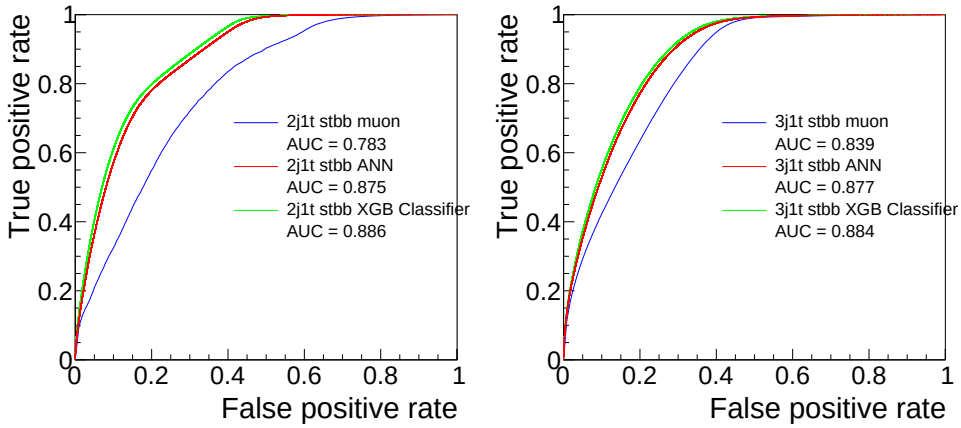
(b) 2j1t configuration

Figure 4.7: NN output distribution for the signal and the background events (left panels) and corresponding ROC curves (right panels) for the ST_{bb} data sample, in 3j1t (upper panels) and 2j1t (lower panels) final states.

4.5. Likelihood ratio vs ML algorithms

Figure 4.8 shows the comparisons between the 2D likelihood ratio based selection and the machine learning ones. Since the XGB Classifier has the highest ROC curve (or, equivalently, the largest Area Under the Curve), the final choice for the jet ordering is the XGB Classifier output. Table 4.3 summarizes the importance of the training variables set for the XGB Classifier.

The top-quark production processes are considered to be equivalent in this analysis, for what concerns the variables used in the training as well as while computing the likelihood ratio in Sec. 4.3. For this reason the only samples considered for the training are ST_{bb} and $S\overline{T}_{bb}$. Adding all the samples involving the top-quark and the anti-top anti-quark could improve the jet selection in both the approaches.



(a) 3j1t muon selection

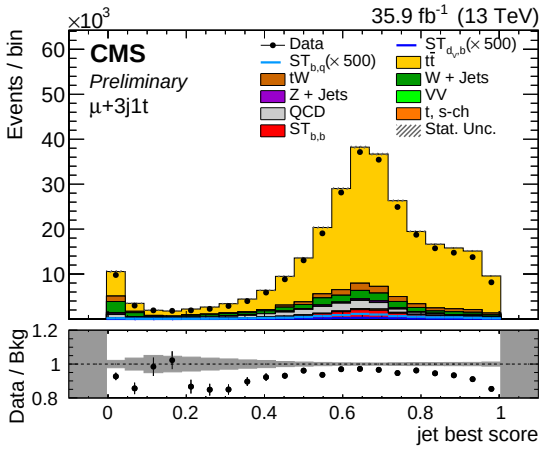
(b) 2j1t muon selection

Figure 4.8: ROC curves for the 2D likelihood ratio based selection (blue line), the XGB Classifier (green line) and the ANN (red line) machine learning algorithms.

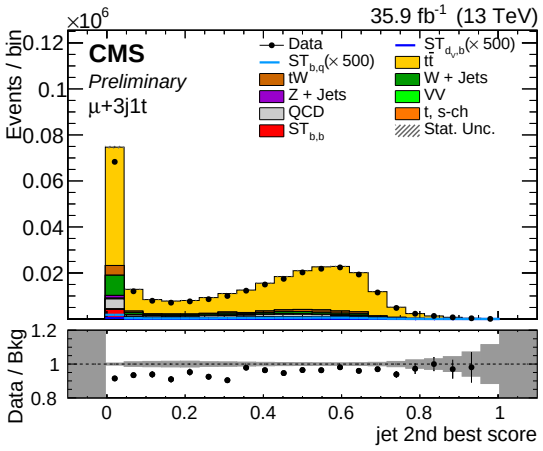
variable	importance
$m_{lep-jet}$	59.5%
Y_{lep}	9.3%
$p_{T_{lep}}$	7.9%
η_{jet}	7.9%
$p_{T_{jet}}$	6.4%
Y_{lep}	4.4%
$\Delta R_{lep-jet}$	2.5%
η_{lep}	1.9%
lepton Charge	0.2%

Table 4.3: List of the variables used in the training and their importance for the XGB Classifier.

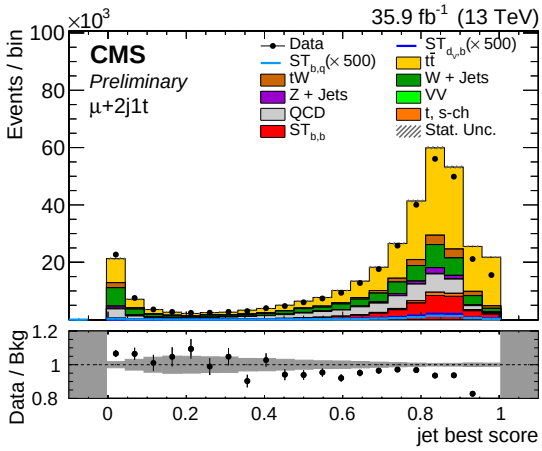
The jet classification with XGB Classifier has been performed on both MC and data samples. After having obtained the jet score, that is the output of the XGB Classifier, jets have been sorted from the jet with highest score (jet best) to the jet with lowest score (jet second or, when there is, third best). The jet score represents the probability for a jet to come from a top-quark decay. Figure 4.9 shows the jet scores for the best and the second best jet.



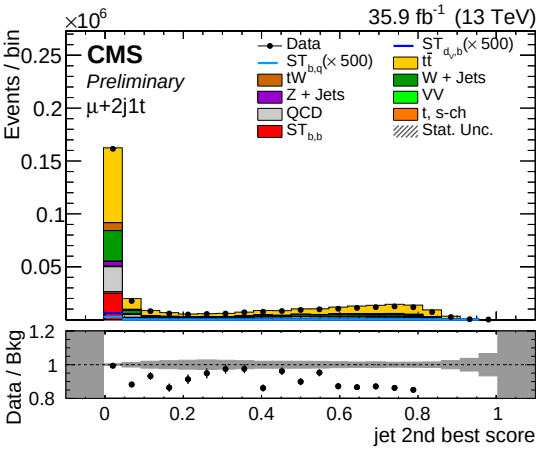
(a) 3j1t jet best score



(b) 3j1t jet 2nd best score



(c) 2j1t jet best score



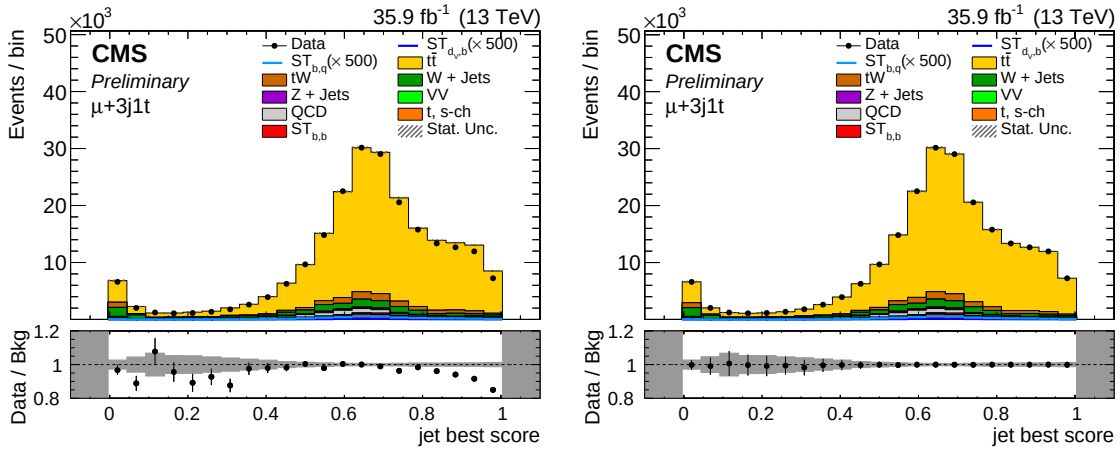
(d) 2j1t jet 2nd best score

Figure 4.9: Jet score for the best and second best jet in the 3j1t (upper panels) and 2j1t (lower panels) final state configurations.

4.5.1 Monte Carlo Reweighing

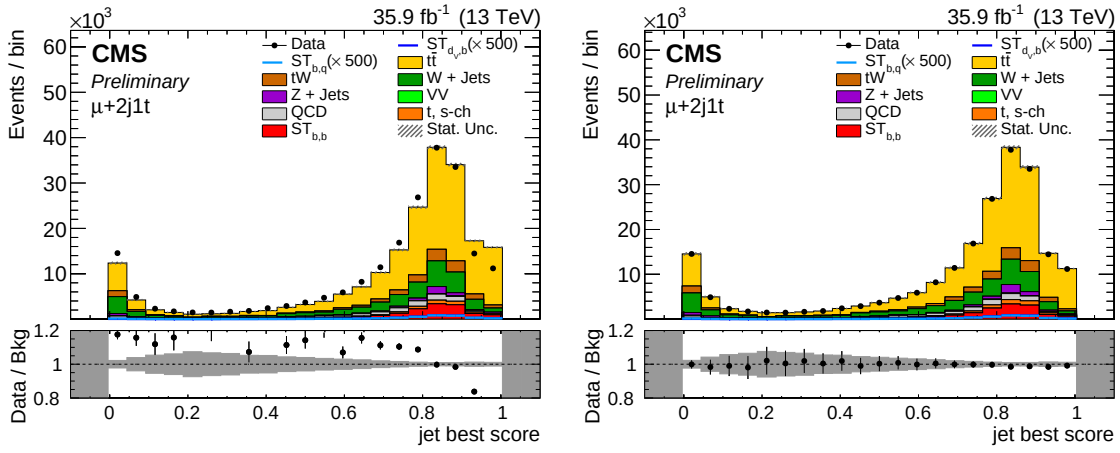
There is a difference between data and MC samples in the machine-learning output, mainly due to variables correlations, and to the top-quark reconstruction, that is difficult to precisely reproduce in simulations. To account for this effect, a reweighing procedure has been performed: the idea is to define a signal-depleted region and to evaluate the ratio between data and simulation as a function of the jet best score, then to apply them as weight in the signal-enriched region. To define these orthogonal regions the variable $|\eta_{forward}|$, that is the pseudo-rapidity of the most forward jet in each event, has been chosen, since it should not be correlated with the jet score.

The distributions of $\eta_{forward}$ vs jet best score in the 2j1t and 3j1t channels for the signals and for the main backgrounds are shown in Fig. 4.12 and Fig. 4.13, respectively. As it was expected, there is no appreciable correlation between the two variables. Therefore in the 2j1t (3j1t) channel the control region[5] (CR), is determined by $|\eta_{forward}| < 2$ ($|\eta_{forward}| < 2.5$), where the weights were computed in the jet best score variable.



(a) 3j1t before reweighting

(b) 3j1t after reweighting



(c) 2j1t before reweighting

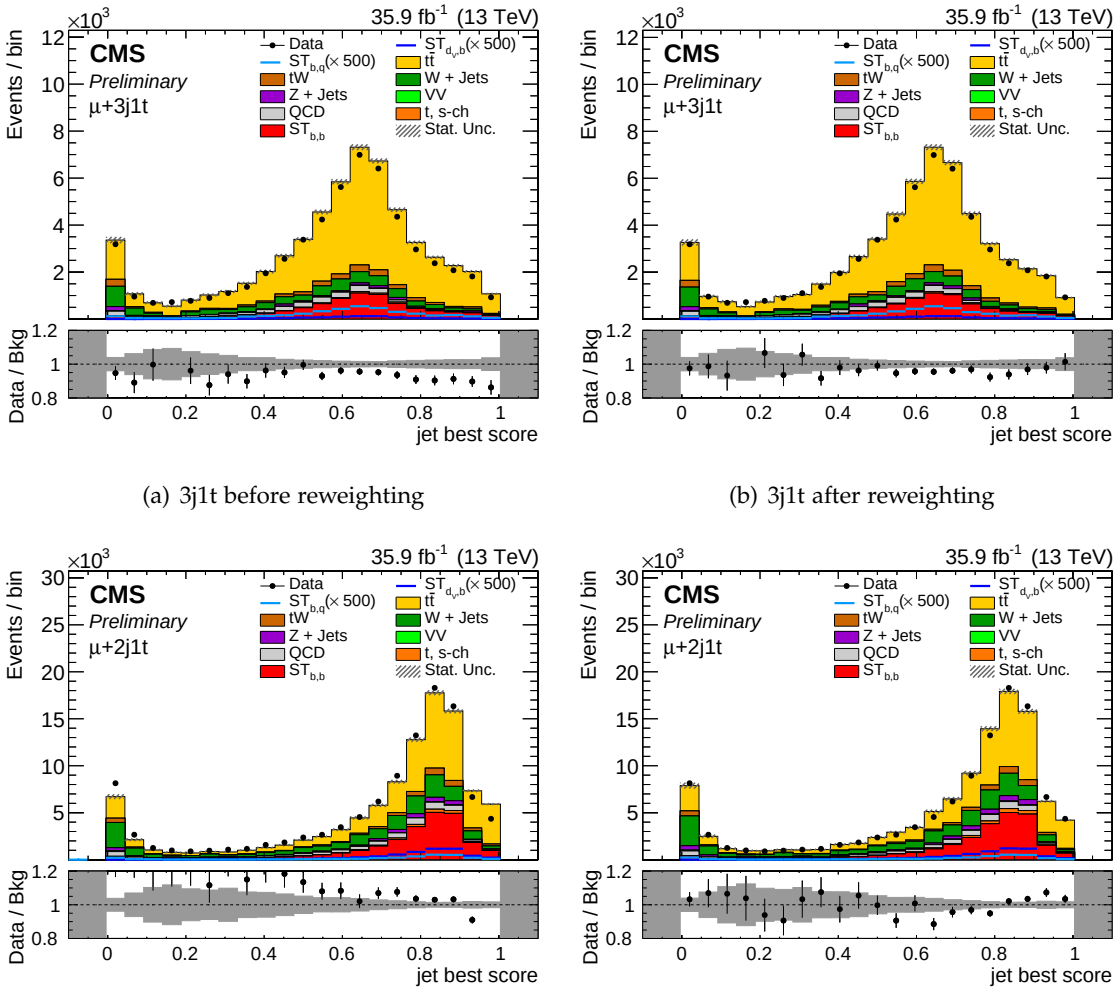
(d) 2j1t after reweighting

Figure 4.10: Jet best score distribution before (left) and after (right) reweighting in the CR, for 3j1t (upper panels) and 2j1t (lower panels) categories.

After having evaluated the ratio of Data/MC in the CR, the weights have been applied both in the CR itself, as a check, and in the signal region (SR), as shown in Fig. 4.10 and Fig. 4.11, respectively.
In the CR the agreement is perfect for construction; in the SR it is good too. Further

[5]It corresponds to a signal depleted region.

improvements can be object of study, since this is a rough reweighing; in addition, the systematic errors are not yet accounted for.



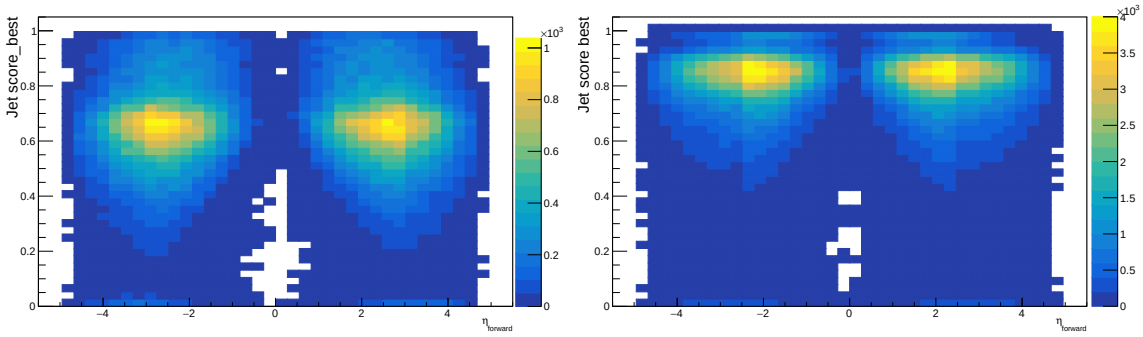
(a) 3j1t before reweighting

(b) 3j1t after reweighting

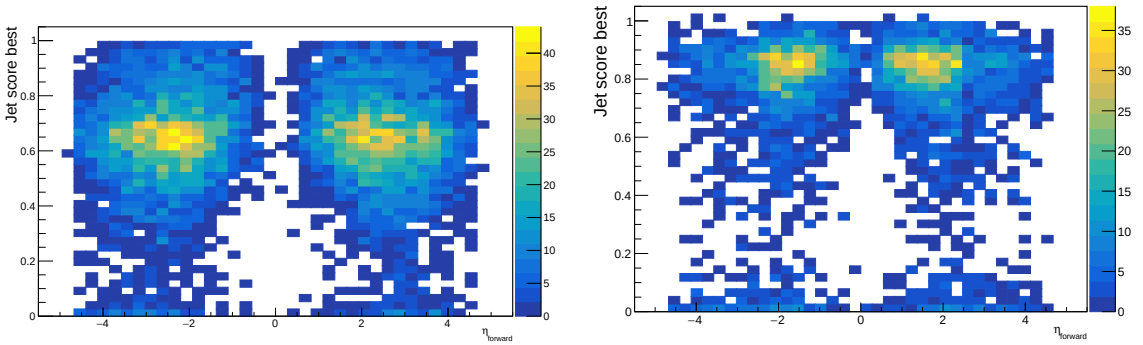
(c) 2j1t before reweighting

(d) 2j1t after reweighting

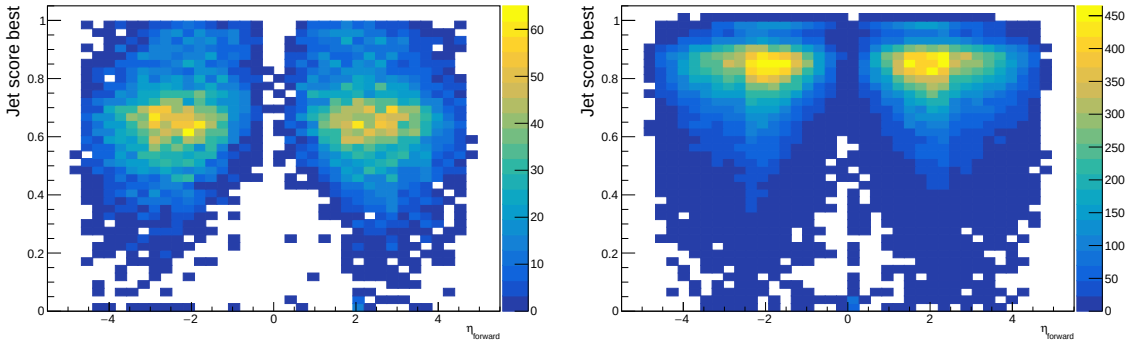
Figure 4.11: Jet best score distribution before (left) and after (right) reweighting in the SR, for 3j1t (upper panels) and 2j1t (lower panels) categories.



(a) 3j1t ST_{bb} (b) 2j1t ST_{bb}

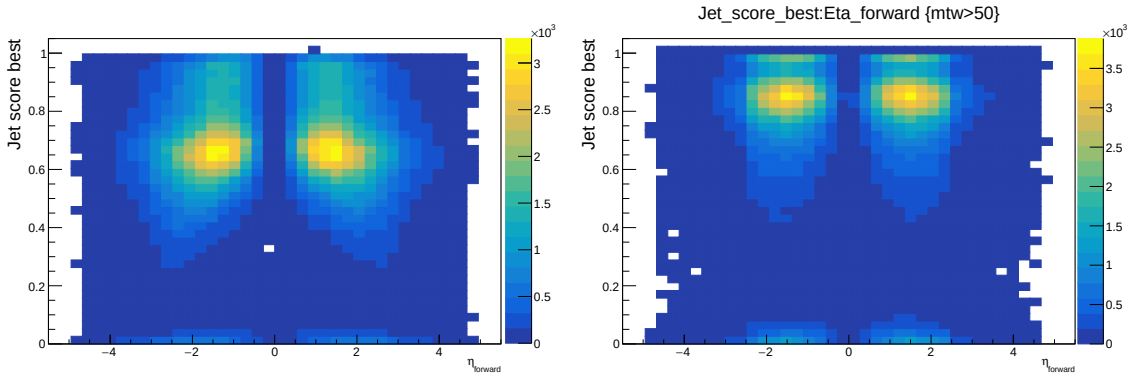


(c) 3j1t ST_{bq} (d) 2j1t ST_{bq}



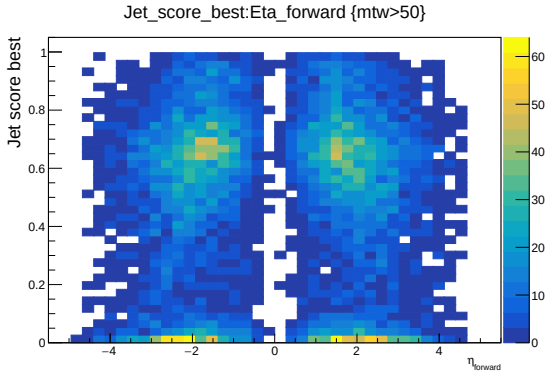
(e) 3j1t ST_{db} (f) 2j1t ST_{db}

Figure 4.12: Distributions of jet best score vs $\eta_{forward}$ for the signal samples: from top to bottom there are ST_{bb} , ST_{bq} and ST_{db} in 3j1t (left) and 2j1t (right) final state configurations.

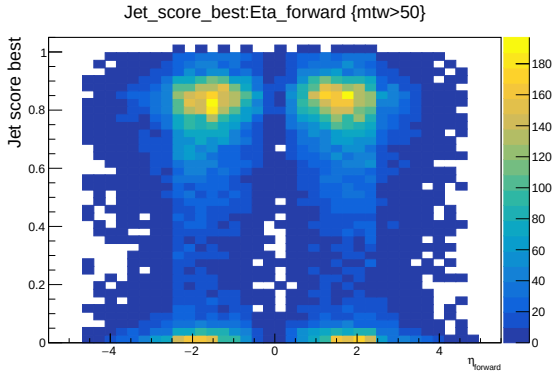


(a) 3j1t $t\bar{t}$

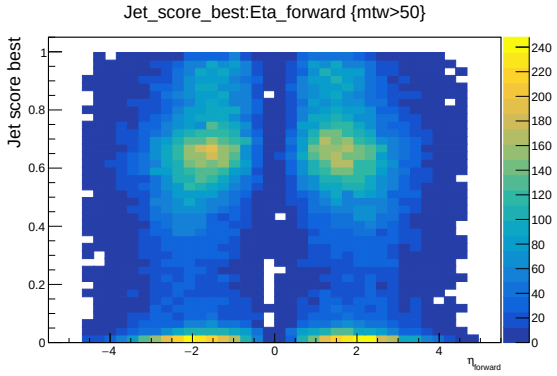
(b) 2j1t $t\bar{t}$



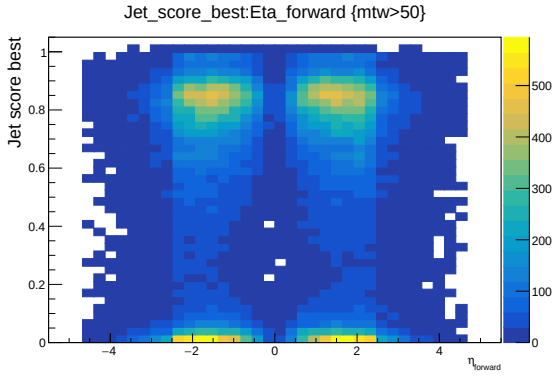
(c) 3j1t W+Jets



(d) 2j1t W+Jets



(e) 3j1t Diboson



(f) 2j1t Diboson

Figure 4.13: Distributions of jet best score vs $\eta_{forward}$ for the main background samples: from top to bottom there are $t\bar{t}$, W+Jets and diboson production in 3j1t (left) and 2j1t (right) final state configurations.

5 Signal extraction

In order to extract the value for $|V_{tb}|$, $|V_{ts}|$ and $|V_{td}|$ parameters, three different signal regions have been defined, allowing to fit the main background contributions and differentiate among the signals. A simultaneous maximum likelihood fit is performed, in order to extract $|V_{tb}|^4$, the sum $|V_{ts}|^2$+$|V_{td}|^2$ and the ratio R_{sd} =$|V_{ts}|^2/|V_{td}|^2$. An upper limit has been set on the sum $|V_{ts}|^2$+$|V_{td}|^2$.

5.1. Signal features

The predicted branching fractions of top quark to d, s, and b quarks can be written as a function of the overall magnitude of the following quantity:

$$\mathcal{B}(t \to Wq) = \frac{|V_{tq}|^2}{|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2}.$$

The decay is assumed to differ only by the strength of the coupling between the top quark and the down-type quark, whereas the kinematic part is the same for all the down-type quarks. The theoretical input for $|V_{tq}|$ and $\mathcal{B}(t \to Wq)$ are shown in Tab. 5.1.

Quark	b	s	d
$\lvert V_{tq}\rvert$	$0.999119^{+0.000024}_{-0.000012}$	$0.04108^{+0.00030}_{-0.00057}$	$0.008575^{+0.000076}_{-0.000098}$
$\mathcal{B}(t \to Wq)$	$0.998239^{+0.000048}_{-0.000024}$	$0.0016876^{+0.0000025}_{-0.0000047}$	$0.000074^{+0.000013}_{-0.000017}$

Table 5.1: Values of the third-row elements of the CKM matrix inferred from low-energy measurements [1], with the respective values of the top quark decay branching fractions.

The channels considered as signal and the corresponding cross sections are reported in Tab. 5.2. The cross sections are evaluated at NLO in the five-flavour scheme using POWHEG

2.0, for $\sigma_{tch.,d/s}$, and using HATHOR for $\sigma_{tch.,b}$.

Process	Cross section $\times$ branching fraction (pb)
$\sigma_{t-ch.,b}\mathcal{B}(t \to Wb)$	216.99 ± 8.37 (scale $\oplus$ pdf)
$\sigma_{t-ch.,b}(\mathcal{B}(t \to Ws) + \mathcal{B}(t \to Wd))$	0.41 ± 0.05 (scale $\oplus$ pdf $\oplus$ exp)
$\sigma_{t-ch.,d}\mathcal{B}(t \to Wb)$	0.102 ± 0.015 (scale $\oplus$ pdf $\oplus$ exp)

Table 5.2: Production cross section of signal processes: in the uncertainty, the "scale" component refers to renormalization and factorization scale uncertainties, the "pdf" part contains the contribution to the uncertainty due to the PDFs, and "exp" denotes the experimental component of the uncertainty from low-energy measurements.

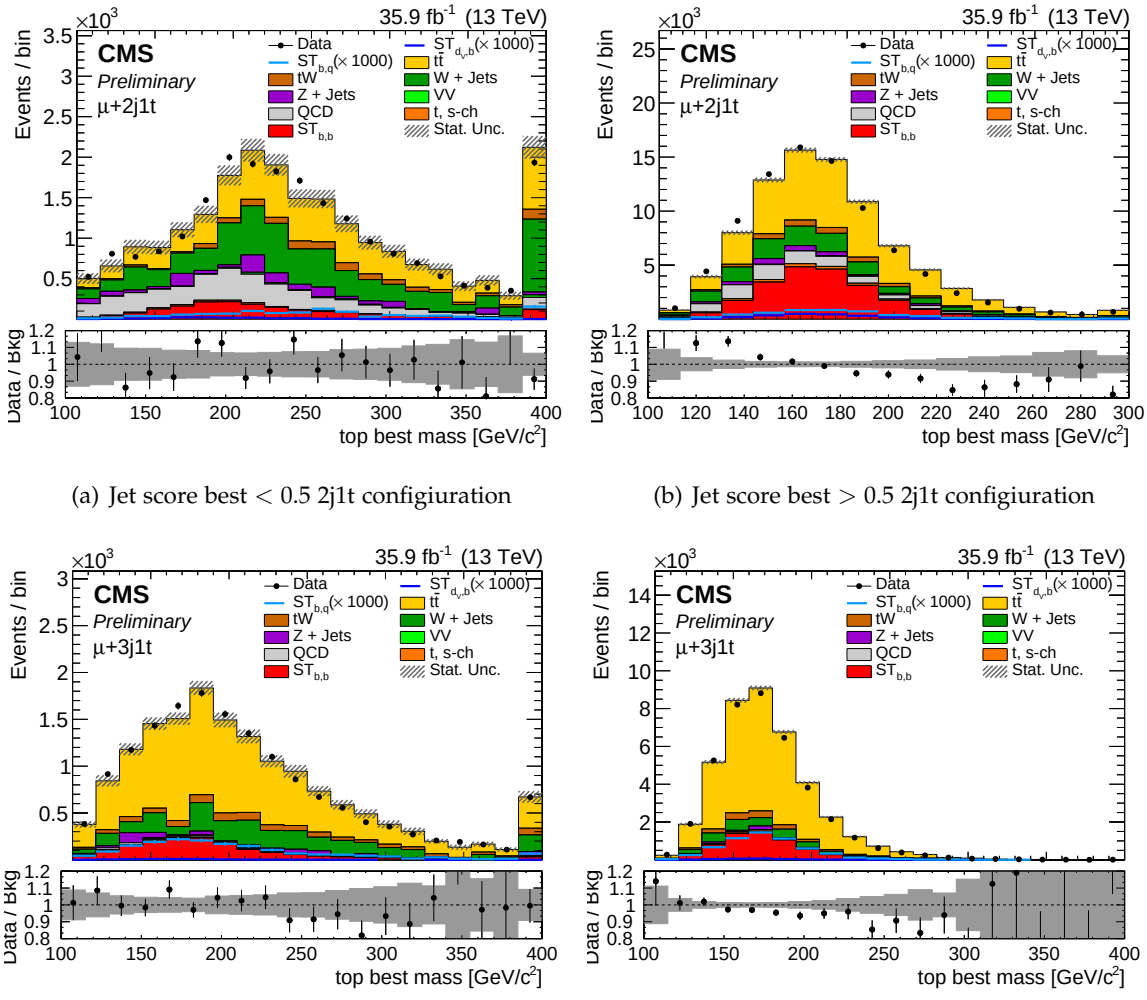
The signatures for single-top-quark t-channel processes involving $|V_{tb}|$, $|V_{ts}|$ and $|V_{td}|$ either in production or decay differ in three aspects: the number of b-tagged jets produced, the features of the jet involved in the reconstruction of the top-quark candidate, and the kinematic features of the events as a result of different PDFs involved for b, s, or d incoming quarks or in the case where the d is the valence one.

The discrimination between the three signals ST_{bq} , ST_{qb} , and ST_{bb} is based on the following considerations:

- ST_{bb} events should have two b quarks in the final state: the samples with 2 b-tagged jets in the final state are enriched in this signal. This final state configuration is not considered in the analysis, as the regions with 2 b-jets have a less favorable signal-to-background ratio. In the final state configurations with 1 b-tagged jet in the final state, the b-tagged jet is typically the one associated to the top quark. Nevertheless, there is a fraction of events that have the jet associated to the reconstructed top quark that is not b-tagged, where the b-tagged jet is the one from the quark coming from the gluon splitting, but this constitutes a small fraction of the total events. In the 2j1t the fraction of b-jets coming from the top is $\approx 85\%$, whereas in 3j1t this fraction is much lower ($\approx 50\%$), since more stringent requirements are applied on the extrajet too. So, both the 2j1t and 3j1t samples are enriched in this process, but in this analysis the 2j1t samples are chosen to fit the ST_{bb} signal;
- the ST_{bq} channel does not distinguish $|V_{td}|$ and $|V_{ts}|$. This sample should have a low energy b-tagged jet, coming from gluon splitting, while the jet associated to the right top quark should be neither the forward jet or the b-tagged jet. Hence, the 2j1t sample will be depleted, whereas the 3j1t should be enriched in this signal;
- the ST_{sb} events have a very similar kinematics to the ST_{bb} ones and are not distinguishable from the ST_{bb} channel when one of the two b-jets is not tagged, as in 3j1t and 2j1t final state configurations. For those reasons they are not considered in the analysis. The ST_{db} events with the valence d quark have different kinematic features, and hence are the only ST_{qb} events considered;
- both ST_{qb} and ST_{bq} processes therefore differ from ST_{bb} by having a single b quark in

the final state, as opposed to two for the latter process. However, this feature can only be exploited when the jet from gluon splitting is energetic enough to be reconstructed, and this feature is not exploited in this analysis since the final states considered have just one b-tagged jet in the final state. This feature can be object of further studies.

The kinematic and angular properties of the top-quark decay products will exhibit some differences depending on whether the correct jet is chosen, or if the jet that originated from the gluon splitting-quark is used instead. This effect is clear looking at the differences between distributions of the jets with higher or lower jet score, as shown in Fig. 5.1. Furthermore, those distributions show that the events with jet score best > 0.5 are signal enriched.



(a) Jet score best < 0.5 2j1t configiuration

(b) Jet score best > 0.5 2j1t configiuration

(c) Jet score best < 0.5 3j1t configiuration

(d) Jet score best > 0.5 3j1t configiuration

Figure 5.1: Reconstructed top-quark mass distributions with the requirements on the jet score for the best jet to be less (left panels) or larger (right panels) than 0.5.

5.2. Signal region and fit variables

Data and MC samples have been divided in three regions based on the final state topologies and other requirements, in order to get a better discrimination between the signals and the backgrounds. Each region is optimized for a particular signal (ST_{bb} , ST_{bq} or ST_{db}).
All the regions have three common requirements:

- the $m_T^W > 50\,\mathrm{GeV}/c^2$ condition, that allows to select a QCD-depleted sample;
- $|\eta_{forward}| > 2.5\,(2)$ in the 3j1t (2j1t) final state configuration: this condition is applied to account for the MC reweighing (see Sec. 4.5.1);
- jet score best > 0.5 is required in order to get a single-top enriched sample.

In order to get a $\mathbf{ST_{bb}}$ **enriched region**, the 2j1t final state configuration is chosen, and the fit variable is $|\eta_{forward}|$. Both the main backgrounds ($t\bar{t}$ and W+Jets) and the ST_{bb} signal contribution are evaluated in this region.

A $\mathbf{ST_{db}}$ **enriched region** is found in the events with the 3j1t final state configuration, with the requirements on the best jet to be b-tagged. Since this region is enriched in the ST_{bb} events too, the fit variable is the output of the XGB Classifier trained to discriminate the ST_{db} from the ST_{bb} events.

Finally, to identify a $\mathbf{ST_{bq}}$ **enriched region**, the 3j1t final state configuration has been chosen, requiring the best jet not to be b-tagged and the second best jet to be b-tagged. This region is orthogonal to the previous one and achieves a optimization for this particular signal, but contains the fraction of ST_{bb} events where the b-tagged jet does not come from the top-quark decay, that is $\approx 50\%$. Therefore, the fit variable is the output of the XGB Classifier trained to discriminate the ST_{bq} from the ST_{bb} events.
The variables used for the XGB Classifier trainings are:

- the top-related kinematic variables: p_T, η, $\cos\theta_{hel}$[1], $\cos\theta_{pol}$[2], Y, m and M_T;
- the missing transverse energy;
- the reconstructed W boson transverse mass defined in Sec. 3.2.4;
- the jet score;
- the jet value of the CMVA b-tagging algorithm.

Since the reconstructed top quarks are associated to the jets, they are ordered from the best (associated to the highest score jet) to the 3rd best (associated with the lowest score jet). All the variables used to train the XGB Classifier and their respective importance in the two different trainings are listed in Tab. 5.3.

The distributions of the $|\eta_{forward}|$ and XGB Classifier outputs in the in the signal regions are shown in Fig. 5.2 and Fig. 5.3, respectively. The signal regions and the variables where the fit procedure has been applied are summarized in Tab. 5.4.

[1]This is the cosine in the top rest frame of the angle between the lepton 3-momenta in the W boson rest frame and the W boson 3-momenta

[2]The cosine of the angle between the lepton and W boson in the top rest frame

variable	importance in ST_{db} discrimination (%)	importance in ST_{bq} discrimination (%)
m_T^W	1.46	0.27
MET	3.6	1.34
η_{lep}	3.28	0.47
$p_{T_{lep}}$	1.75	3.82
$\eta_{forward}$	2.37	4.6
p_T(best top)	3.07	1.61
$\cos\theta_{hel}$(best top)	2.83	3.44
$\cos\theta_{pol}$(best top)	2.39	2.96
Y(best top)	1.44	1.09
m(best top)	3.16	5.13
M_T(best top)	4.68	2.96
p_T(2nd best top)	1.78	3.53
$\cos\theta_{hel}$(2nd best top)	4.17	1.49
$\cos\theta_{pol}$(2nd best top)	3.17	5.24
η(2nd best top)	1.79	2.77
Y(2nd best top)	2.12	1.69
m(2nd best top)	3.98	3.33
M_T(2nd best top)	4.5	2.78
p_T(3rd best top)	3.04	2.66
$\cos\theta_{hel}$(3rd best top)	1.7	2.29
$\cos\theta_{pol}$(3rd best top)	4.33	3.66
η(3rd best top)	2.28	2.84
Y(3rd best top)	5.07	3.61
m(3rd best top)	1.91	2.62
M_T(3rd best top)	3.46	3.09
jet best CMVA	3.33	19.94
jet best score	3.48	4.6
jet 2nd best CMVA	10.19	0.69
jet 2nd best score	3.31	3.56
jet 3rd best CMVA	3.45	0.58
jet 3rd best score	2.88	1.30

Table 5.3: List of the input variables for the XGB Classifier trained for the ST_{db} vs the ST_{bb} channel and for the ST_{bq} vs the ST_{bb} channel, with their respective importances.

Variable	Region	Requirements
$\lvert\eta_{forward}\rvert$	2j1t ST_{bb} enriched	$m_W^t > 50\,\mathrm{GeV}/c^2$ jet score > 0.5 $\lvert\eta_{forward}\rvert > 2$
XGB Classifier output vs ST_{db}	3j1t ST_{db} enriched	$m_W^t > 50\,\mathrm{GeV}/c^2$ jet score > 0.5 $\lvert\eta_{forward}\rvert > 2.5$ jet best b-tagged
XGB Classifier output vs ST_{bq}	3j1t ST_{bq} enriched	$m_W^t > 50\,\mathrm{GeV}/c^2$ jet score > 0.5 $\lvert\eta_{forward}\rvert > 2.5$ jet best not b-tagged jet 2nd best b-tagged

Table 5.4: Signal regions and requirements on the variables.

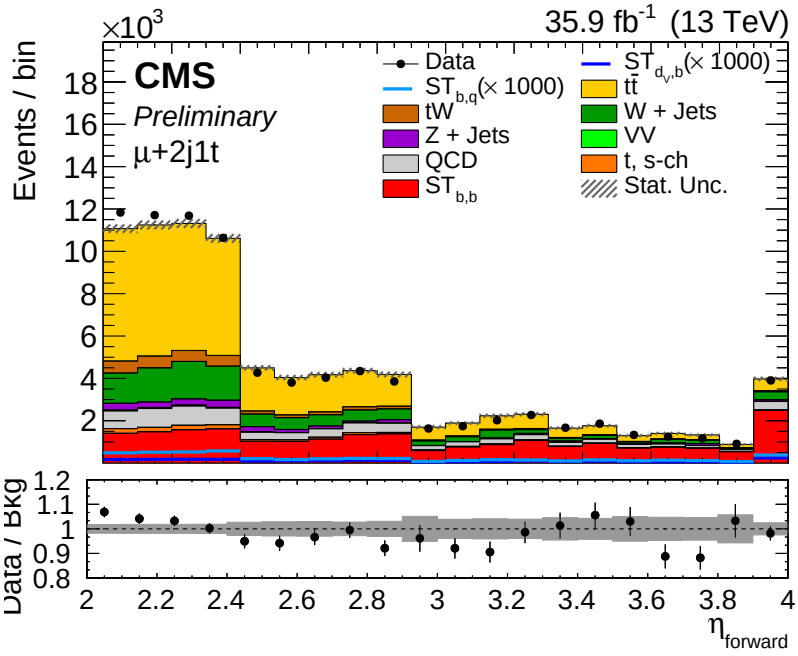
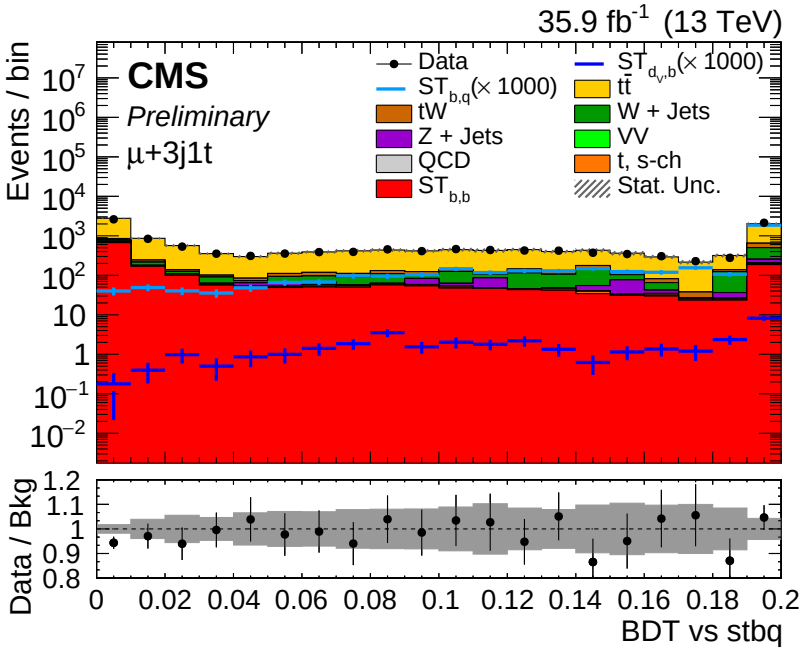
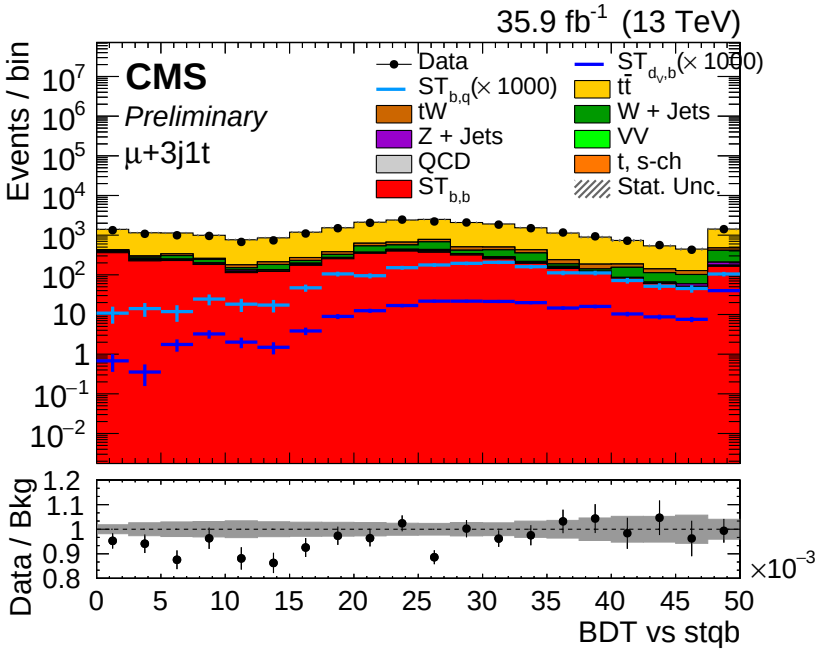


Figure 5.2: Distribution of $\lvert\eta_{forward}\rvert$, i.e. the η of the most forward jet, in 2j1t ST_{bb} enriched region.



(a)



(b)

Figure 5.3: Distributions of XGB Classifier output vs ST_{bq} in 3j1t ST_{bq} enriched region (upper panel) and XGB Classifier output vs ST_{db} in 3j1t ST_{db} enriched region (lower panel)

5.3. Fit procedure

The CKM matrix elements are extracted from the measurement of the production cross sections and branching fractions of single-top-quark t-channel processes, drawing contributions from $|V_{tb}|$, $|V_{ts}|$, and $|V_{td}|$ in production and decay and the $|V_{ts}|^2/|V_{td}|^2$ ratio deviations from the one used in the analysis. The fit procedure is divided into two steps. In the first step, a maximum likelihood (ML) fit to the m_T^W distribution is performed separately for the 2j1t and 3j1t regions in order to extract the QCD multijet contribution, as explained in Sec. 3.4. The QCD multijet normalization and the relative uncertainty are extrapolated to the QCD multijet-depleted regions and used as an input to the next part of the fit procedure. In the second step, a ML fit is performed simultaneously in the three regions listed in Tab. 5.4, in order to discriminate between ST_{bb} and (ST_{db} +ST_{bq}). The extended likelihood function can be written as the superposition of signal and background components, weighted by the expected signal and background fractions, respectively:

$$\mathcal{L}_{s+b}(\vec{x}_1,\ldots,\vec{x}_N;\mu,\vec{\theta}) = \frac{e^{-\lambda}}{N!}\prod_{i=1}^{N}(\mu s f_s(\vec{x}_i;\vec{\theta}) + b f_b(\vec{x}_i;\vec{\theta})),$$

where $\lambda = \mu s(\vec{\theta}) + b(\vec{\theta})$, s and b are the expected signal and background yields from theory, f_s and f_b are the probability distribution functions (p.d.f.) for the signal and for the background, respectively, $\vec{x}_i$ is the i^{th} variable set, and N is the total number of measurements. The parameter μ is called the *signal strength*. As will be explained in Sec. 5.4, the systematic uncertainties are introduced as nuisance parameters. In this notation they are taken into account in the parameters set $\vec{\theta}$.

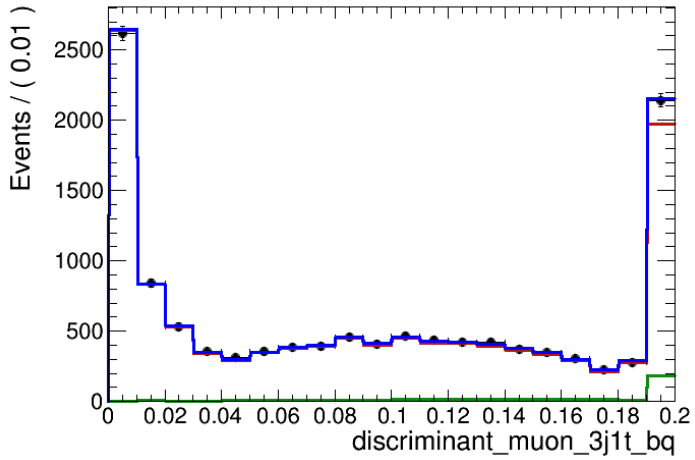


Figure 5.4: Distribution of XGB Classifier output vs ST_{bq} after the ML fit. The red and lines represent the background and signal contributions, respectively, whereas the blu line is the global fit.

Parameters of interest They are the signal strength parameters of the processes of interest: μ_b, that refers to the ST_{bb} process, and μ_{sd}, that is the signal strength for the (ST_{db} +ST_{bq}) processes. Since there are two signal strength parameters, it is possible to separate the ST_{bb} and (ST_{db} +ST_{bq}) contributions in the likelihood function, thus obtaining:

$$\mathcal{L}_{s+b}(\vec{x}_1,..,\vec{x}_N;\mu_{sd},\mu_b,\vec{\theta}) = \frac{e^{-\lambda}}{N!}\prod_{i=1}^{N}(\mu_{sd}s_{sd}f_{s_{sd}}(\vec{x}_i;\vec{\theta}) + \mu_b s_b f_{s_b}(\vec{x}_i;\vec{\theta}) + b f_b(\vec{x}_i;\vec{\theta}))\prod_{j=1}^{k}\pi_j(\vec{\theta}), \tag{5.1}$$

where $\lambda = \mu_{sd}s_{sd}(\vec{\theta}) + \mu_b s_b(\vec{\theta}) + b(\vec{\theta})$ and $\pi_j(\vec{\theta})$ are penalty functions for the k nuisance parameters to take into account their experimental constraints.

Nuisance parameters They are additional parameters in the likelihood function, and are usually not the main focus of the measurement. In the notation adopted in Eq. 5.1, nuisance parameters are the $\vec{\theta}$ set. They are:

- QCD-multijet contribution, with a prior uncertainty and central value taken from the QCD fit in the first step of the analysis;
- $t\bar{t}$ and W+jets normalization parameters are left free to float;
- profiled systematic uncertainties, described in Sec. 5.4 .

In order to determine the uncertainty on maximum likelihood estimate, a scan of the negative log likelihood, $-2\ln\mathcal{L}$ is performed around its minimum value $\mathcal{L}_{max} = \mathcal{L}(\hat{\hat{x}};\hat{\mu}_{sd},\hat{\mu}_b,\hat{\hat{\theta}})$, where the $(\hat{\mu}_{sd},\hat{\mu}_b,\hat{\hat{\theta}})$ parameters allow the likelihood to reach the maximum value. An interval corresponding to an increase of $-2\ln\mathcal{L}$ by one unit with respect to the likelihood maximum value can be determined. The signal strength, μ, is a multiplying factor for the amount of signal, s, in the likelihood [68]. The negative log likelihood dependence on the signal strength, μ_{sd} is shown in Fig. 5.5. The fit result in the 3j1t region of the XGB Classifier trained with the ST_{bq} sample can be found in Fig. 5.4.

5.4. Systematic uncertainties

Systematic uncertainties are not directly due to the statistics of the data sample. The sources of a systematic uncertainty in HEP are related to:

- the detector simulation, as calibrations (jet energy scale), efficiencies, particle ID, reconstruction (electron or muon detection), resolutions (jet energy, muon momentum);
- theoretical uncertainties, as factorization or normalization scale of MC generators, the choice of MC generator itself;
- statistical uncertainties due to the limited number of simulated events.

Systematic uncertainties are treated with the profiling approach, that takes them into account as nuisance parameters, considering all the templates corresponding to the variations of

each systematic scenario, and performing an interpolation between them in order to define a smooth dependence of the likelihood function from each parameter. Several sources of systematic uncertainties have been considered as profiled uncertainties in the fit of the XGB Classifier output distributions and of the $|\eta_{forward}|$ in the control region. These include the uncertainty sources related to the modeling of the signal process, and jet energy scale uncertainties, which play a major role for events featuring hadronic activity in the high-pseudorapidity region of the detector. The different uncertainty sources that are considered in the analysis are:

- **limited size of simulated event samples**: the statistical uncertainty due to the limited size of the simulated event samples is evaluated for each bin and each process with the Barlow-Beeston "light" method [73, 74];
- **lepton trigger and reconstruction efficiencies**, estimated with a "tag-and-probe" method from Drell-Yan events with the dilepton invariant in the Z boson mass peak;
- the uncertainty in the average expected number of **pileup interactions** is propagated as a source of systematic uncertainty in this measurement by varying the total proton-proton inelastic cross section by $\pm$4.6%.
- the **QCD multijet background normalization** is introduced as a nuisance parameter of initial value the result of the m_W^T fit performed in the previous section;
- **$t\bar{t}$ normalization and uncertainties** in the two regions: those uncertainties cover potential mismodelling of this process;
- **W+jets normalization**: a separate uncertainty is dedicated to the fraction of W+jets events that covers potential mismodelling of those processes;
- **renormalization and factorization scale uncertainties**: those uncertainties are considered for the signal process, tW, $t\bar{t}$, and W + jets. The effect is estimated for each process separately;
- uncertainties on the **b-tagging and mistagging** efficiency measurements are split into two different components;
- **jet energy scale (JES)**: All reconstructed jet four-momenta in simulated events are simultaneously varied according to the η and p_T dependent uncertainties in the JES [75]. This variation in jet four-momenta is also propagated to the missing transverse energy;
- **$|\mathbf{V_{ts}}|^2/|\mathbf{V_{td}}|^2$ ratio separation**: it is introduced with a gaussian penalty function π_{sd} with standard deviation $\sigma_{R_{sd}} = 0.5$ that allows to quantify the measured $|V_{ts}|^2/|V_{td}|^2$ ratio with respect to the value used for the simulations;
- **luminosity**: the integrated luminosity is known with a relative uncertainty of $\sigma_{lumi} =$ 2.7% [76].

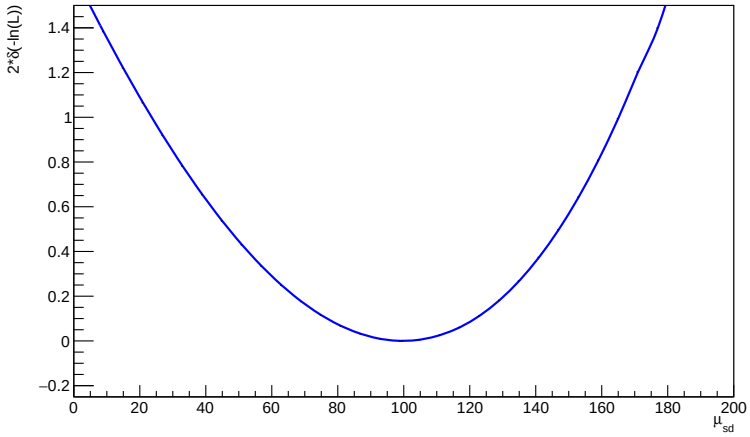


Figure 5.5: Negative log likelihood as a function of the signal strength, μ_{sd} parameter.

5.5. Fit results and interpretation

The fit has been performed by considering the (ST_{bq} + ST_{db}) contributions and by introducing a gaussian parameter that quantifies the difference in terms of standard deviations of the $|V_{ts}|^2/|V_{td}|^2$ ratio with respect to the one used to do the fit. The fit results are then converted in terms of multiplying factors for the $|V_{ts}|$ and $|V_{td}|$ matrix elements.

The contributions of each CKM matrix element come from two different signal strength parameters: μ_b, that refers to the ST_{bb} process, and the second one, μ_{sd} , refers to and (ST_{bq} + ST_{db}) processes. In the SM, top quarks only decay to W bosons associated with b, s, or d quarks, and their branching fractions are proportional to the respective matrix element.

By neglecting terms proportional to $|V_{td}|^4$ and $|V_{ts}|^4$, the signal strength parameters can be written:

$$\mu_b = \frac{\sigma^{obs}_{t-ch,b}\mathcal{B}(t \to Wb)^{obs}}{\sigma_{t-ch,b}\mathcal{B}(t \to Wb)},$$
$$\mu_q = \frac{\sigma^{obs}_{t-ch,q}\mathcal{B}(t \to Wb)^{obs} + \sigma^{obs}_{t-ch,b}\mathcal{B}(t \to Wq)^{obs}}{\sigma_{t-ch,q}\mathcal{B}(t \to Wb) + \sigma_{t-ch,b}\mathcal{B}(t \to Wq)},$$

where q stands for s or d. In the signal strength μ_q both production and decay in s or d quarks are considered; it is possible to re-write the following quantities in terms of $|V_{tb}|$, $|V_{ts}|$ and $|V_{td}|$:

$$\mu_b = \frac{|V_{tb}|^4_{obs}\left(\Gamma_b/\Gamma_{top}\right)^{obs}}{|V_{tb}|^4\left(\Gamma_b/\Gamma_{top}\right)},$$
$$\mu_q = \frac{|V_{tb}|^2_{obs}(|V_{ts}|^2_{obs}+|V_{td}|^2_{obs})\left(\Gamma_q/\Gamma_{top}\right)^{obs}}{|V_{tb}|^2(|V_{ts}|^2+|V_{td}|^2)\left(\Gamma_q/\Gamma_{top}\right)}.$$

In addition, it is possible to write those quantities in function of the $|V_{ts}|^2/|V_{td}|^2$ ratio:

$$\mu_b = \frac{|V_{tb}|^4_{obs}\left(\Gamma_b/\Gamma_{top}\right)^{obs}}{|V_{tb}|^4\left(\Gamma_b/\Gamma_{top}\right)},$$
$$\mu_q = \frac{|V_{tb}|^2_{obs}|V_{td}|^2_{obs}(1+R^{obs}_{sd})\left(\Gamma_q/\Gamma_{top}\right)^{obs}}{|V_{tb}|^2|V_{td}|^2(1+R_{sd})\left(\Gamma_q/\Gamma_{top}\right)}.$$

where $R_{sd} = |V_{ts}|^2/|V_{td}|^2$.
The fit has been performed introducing a gaussian parameter that quantifies the smearing of R_{sd} from the nominal value, in terms of standard deviations. The results of the unconstrained fit, in the SM scenario, are:

$$|V_{tb}|^4 = 1.13 \pm 0.05,$$
$$|V_{tb}| = 1.031 \pm 0.013,$$
$$|V_{td}|^2 + |V_{ts}|^2 = 0.20 \pm 0.12,$$
$$\frac{1}{\sigma_{R_{sd}}}\left(\frac{1}{1+R^{obs}_{sd}} - \frac{1}{1+R_{sd}}\right) = 0.04 \pm 0.99.$$

All the results are compatible within two standard deviations with the SM predictions. The above results do not consider the ratio Γ_q/Γ_{top} being different from the one observed. Moreover, these results do not include the unitary constraint on $|V_{tb}|$, $|V_{ts}|$ or $|V_{td}|$.

Upper limit on $|V_{td}|^2 + |V_{ts}|^2$

Since the value obtained for $|V_{td}|^2 + |V_{ts}|^2$ is near to the physical limit, an upper limit on their squared sum has been put. The test statistic adopted to compute this limit is the modified profile likelihood, defined as follows:

$$\widetilde{q}_\mu = \begin{cases} -2\ln\dfrac{\mathcal{L}(\vec{x};\mu,\hat{\hat{\vec{\theta}}}(\mu))}{\mathcal{L}(\vec{x};0,\hat{\vec{\theta}})} & \hat{\mu} < 0, \\ -2\ln\dfrac{\mathcal{L}(\vec{x};\mu,\hat{\hat{\vec{\theta}}}(\mu))}{\mathcal{L}(\vec{x};\mu,\hat{\vec{\theta}})} & \hat{\mu} \geq 0. \end{cases}$$

The reason this test statistic is chosen is that $\widetilde{q}_\mu$ accounts for nuisance parameters: the $\vec{\theta}$ parameters set is treated as nuisance parameters, whereas the signal strength μ is the parameter of interest. Additionally, since it is a likelihood ratio based test statistic, for the Neyman-Pearson lemma [69], it achieves optimal performances with respect to the numerator-vs-denominator hypotheses separation. In this case, the μ null value is replaced with 0 in the denominator of the profile likelihood if the observed signal strength $\hat{\mu}$ is negative, to avoid an un-physical values of it.

This test statistic is used to compute the p-values in the different hypotheses on μ. To do it, larges Monte Carlo pseudo-samples are generated. Nevertheless, since this approach requires high computation power, asymptotic approximations are performed [77]. The upper limit is computed with the CLs method, where confidence level is defined as follows:

$$CL_s(\mu) = \frac{p_{s+b}(\mu)}{1 - p_b(\mu)},$$

where:

$$p_{s+b}(\mu) = P_{s+b}(t(\mu) \geq t_{\hat{\mu}}),$$
$$p_b(\mu) = P_b(t(\mu) \leq t_{\hat{\mu}}).$$

In this notation, p_{s+b} and p_b are the two p-values corresponding to both the signal+background ($s+b$) hypothesis and the only background (b) one, $P_{s+b}(t(\mu) \geq t_{\hat{\mu}})$ is the probability that the test statistic t_μ is greater or equal to the observed value of test statistic $t_{\hat{\mu}}$ in the signal+background hypothesis, and $P_b(t(\mu) \leq t_{\hat{\mu}})$ is the probability that the test statistic t_μ is less or equal to the observed value of test statistic $t_{\hat{\mu}}$ in the only background hypothesis. This means to compute the ratio of the right integral of the test statistic, that represents the probability to get a greater value than the one obtained, in the signal+background hypothesis, and the left integral of the test statistic,that represents the probability to get a lower value than the one obtained in the only background hypothesis. The upper limit obtained with this method is:

$$|V_{td}|^2 + |V_{ts}|^2 < 0.41 \text{ at } 84\% \ C.L.$$

Conclusions

In this thesis a direct measurement of the Cabibbo-Kobayashi-Maskawa matrix elements $|V_{tb}|$, $|V_{ts}|$ and $|V_{td}|$ with single-top t-channel events is presented. The t-channel is the dominant electroweak production mechanism for single top quarks and it contains a tWq vertex, where q stands for a b, s or d quark, both in production and in decay. The cross sections and branching fractions of these processes are therefore sensitive to the strength of the electroweak coupling, making them suitable channels for direct measurements of the magnitude of CKM matrix elements $|V_{tb}|$, $|V_{ts}|$ and $|V_{td}|$.

Events in proton-proton collision data with a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$, collected with the CMS experiment at the LHC, have been used for this analysis. The subset of analyzed data corresponds to an integrated luminosity of $35.92\,\text{fb}^{-1}$ and contains one isolated muon and two or three jets, one of which has to pass the b-tagging requirement. The events of both the simulated samples and the data are divided in two different categories based on the final state jets multiplicity: two or three jets, one of which has to be identified as coming from a b-quark.

The signal is composed by three different single-top t-channel production processes: the ones that have a coupling to a b quark both in decay and in production, and the processes with a b quark either in decay or production, while the other quark is a d or s quark. Events with d or s quark both in production and in decay are highly suppressed because of the magnitude of $|V_{td}|$ and $|V_{ts}|$, therefore they are not considered in this analysis. The main background processes that can mimic the final state configuration of signal events are the $t\bar{t}$ pair production, and the W/Z + jets events. Both background and signal samples are simulated with the state-of the art simulation softwares. QCD multijet events are estimated from data with a data-driven procedure.

The analysis has been divided in two main parts. In the first part, jets have been ordered according to a certain score that is assigned to each jet, that represents the probability to be originated from a top-quark decay. Different ways to define it have been studied and the

method that achieves the best performances has been chosen. This approach is based on a machine learning algorithm, the XGB Classifier: the output of the XGB Classifier is taken as the jet score. This is used both to individuate the jets coming from top and to order jets in the event. A reweighing in a signal depleted region has been performed in order to account for differences in data and Monte Carlo samples in the score due to variables correlations that are difficult to accurately reproduce and modify the output of the classifier. This is first time a top-tagging approach based on a multi-variate classifier is chosen to select the events.

In the second part of the analysis, three regions characterized by different final state topologies and requirements on the jet variables have been identified, in order to get a better discrimination between the signals and the backgrounds; each region is optimized for a particular signal. Several requirements are applied in order to define those regions, selecting a QCD-depleted sample and events that contain at least one top-quark according to the jet score. For the first time, a multivariate analysis with a machine learning algorithm, the XGB Classifier, is performed in order to recognize the most abundant single-top-quark events, with a b-quark both in production and in decay, from the rarest signals that have a s or d quark either in production or in decay. The output of the XGB Classifier is evaluated in two out of the three regions, in order to fully exploit the most important kinematic variables, and it is adopted as fit variable. In the region where the XGB Classifier output has not been performed, the $|\eta_{forward}|$ variable is chosen for the fit. A simultaneous maximum likelihood fit is performed, in order to extract $|V_{tb}|^4$, $|V_{ts}|^2$+$|V_{td}|^2$ and the ratio $R_{sd} = |V_{ts}|^2/|V_{td}|^2$.

The results obtained are:

$$
\begin{aligned}
&|V_{tb}|^4 = 1.13 \pm 0.05\,, \\
&|V_{tb}| = 1.031 \pm 0.013\,, \\
&|V_{td}|^2 + |V_{ts}|^2 = 0.20 \pm 0.12\,, \\
&\frac{1}{1 - R_{sd}^{obs}} - \frac{1}{1 - R_{sd}} = 0.04 \pm 0.99\,.
\end{aligned}
$$

Since the sum $|V_{td}|^2 + |V_{ts}|^2$ is found to be near to the physical limit (it is compatible with zero within two sigma), an upper limit computed with the CLs method has been set:

$$|V_{td}|^2 + |V_{ts}|^2 < 0.41 \text{ at } 84\%\ C.L.$$

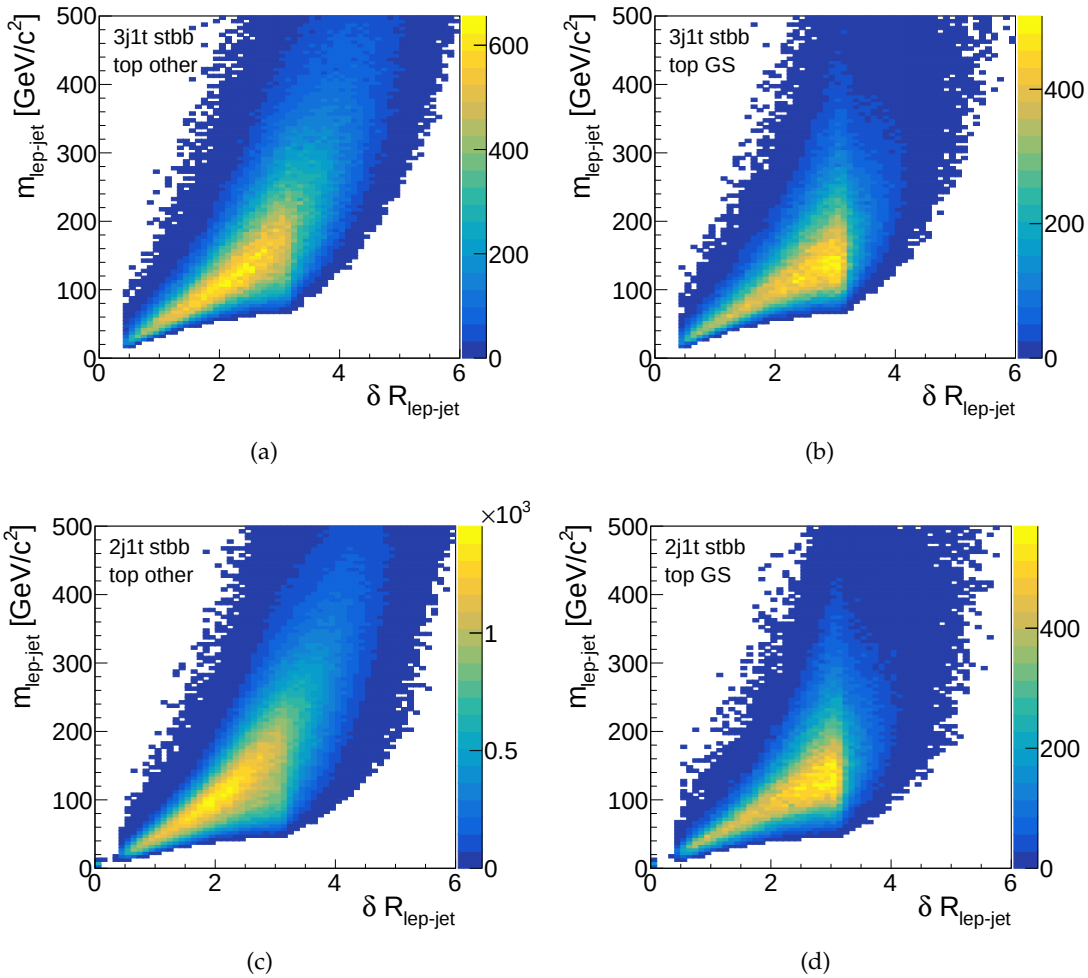
There is no significant tension from with the existing measurements.

This measurement could be improved by including the full Run 2 data collected by CMS, both with an isolated muon and with an isolated electron in the final state, and by adding all the systematic uncertainties. Moreover, with the inclusion of the samples wth 3 jets, 2 of which are b-tagged, in the final state, a better discrimination of the events that have a b quark both in production and in decay can be achieved. The differences between the kinematic features of the events where the top quark is produced by a valence d quark (and has a b quark in its decay), and the ones of the events with a b quark both in production and in decay, can be further exploited in the 2j1t configuration to optimize their separation.

Finally, another improvement could be inclusion in the fit of direct constraints among the different CKM matrix elements $|V_{tb}|$, $|V_{ts}|$ and $|V_{td}|$ to allow for a wider variety of possible hypotheses to be tested.

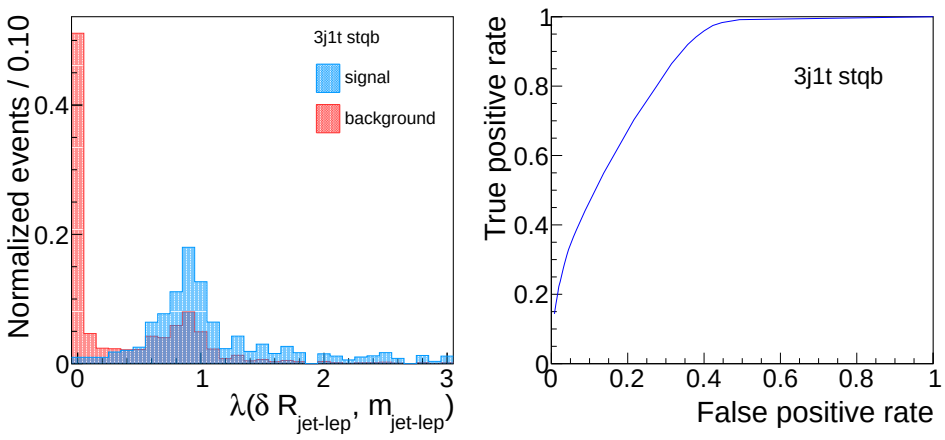
A

Classifiers performances

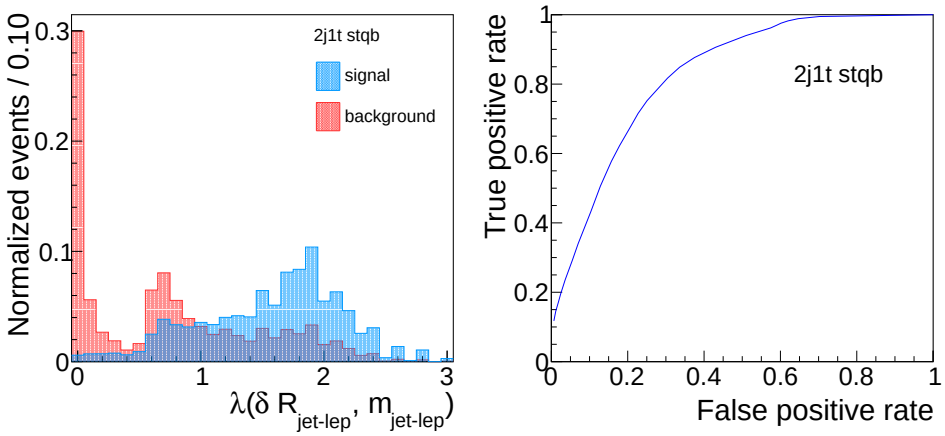


(a) (b)

(c) (d)

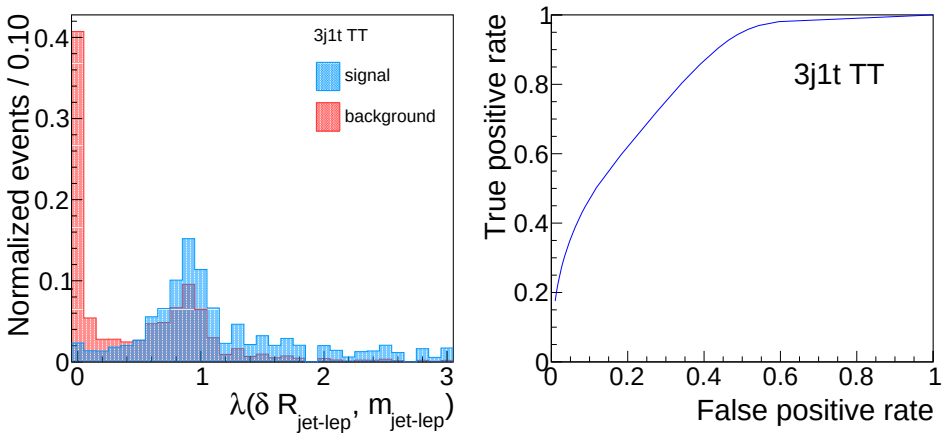


(e) 3j1t configuration

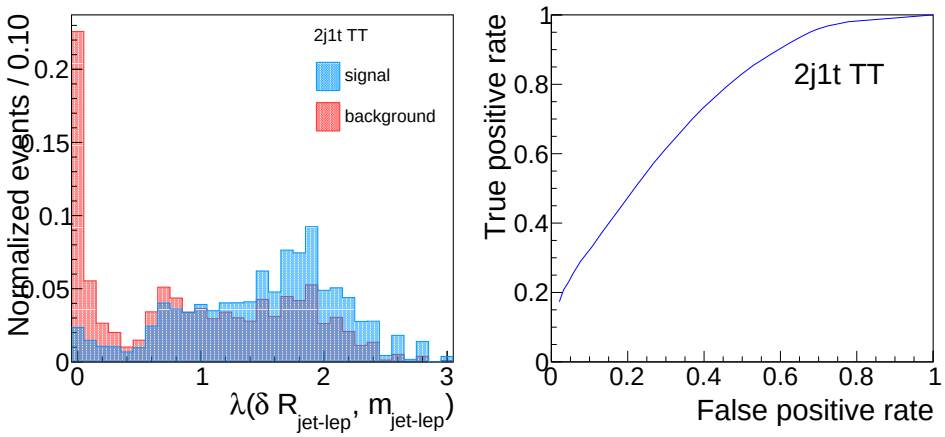


(f) 2j1t configuration

Figure A.1: Likelihood ratio distribution for the signal and the background events (left panels) and corresponding ROC curves by varying the cut (right panels) for the ST_{qb} data sample, in 3j1t and 2j1t configurations.

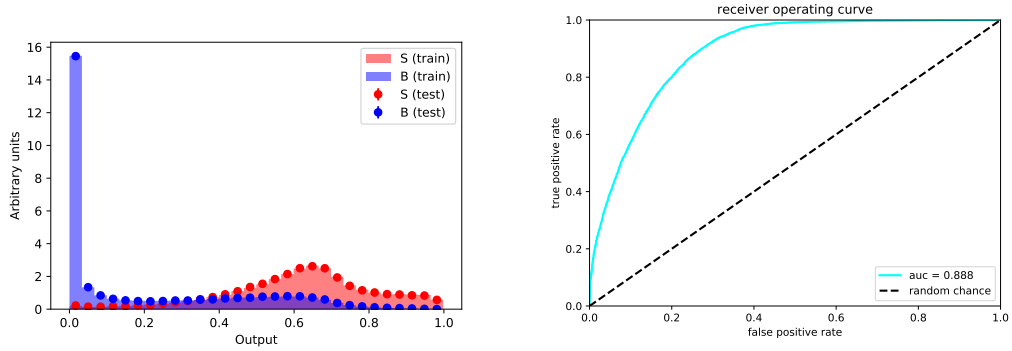


(a) 3j1t configuration

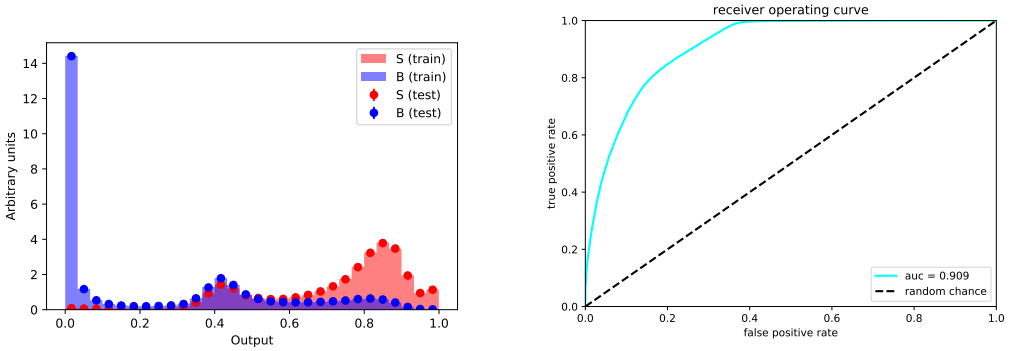


(b) 2j1t configuration

Figure A.2: Likelihood ratio distribution for the signal and the background events (left panels) and corresponding ROC curves by varying the cut (right panels) for the $t\bar{t}$ data sample, in 3j1t and 2j1t configurations.

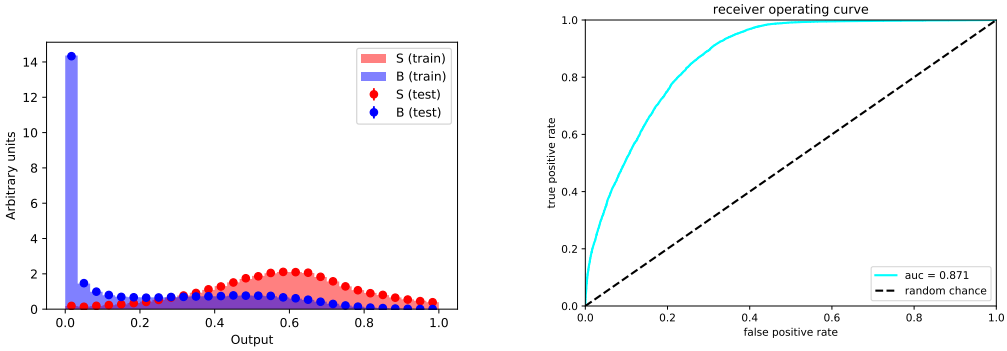


(a) 3j1t configuration

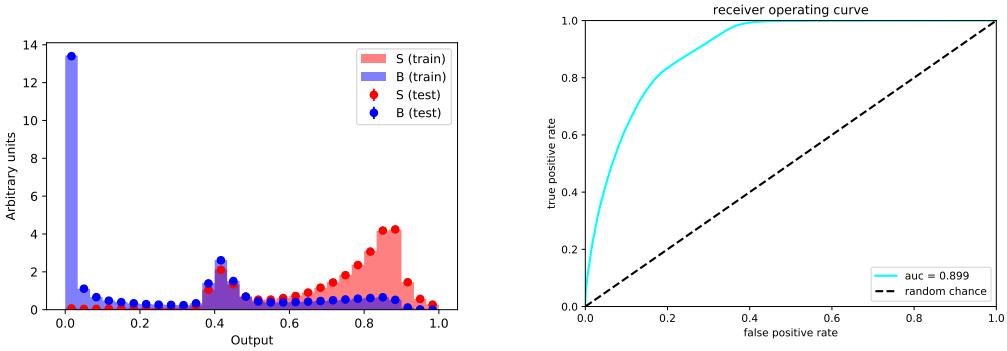


(b) 2j1t configuration

Figure A.3: XGB Classifier output distribution for the signal and the background events (left panels) and corresponding ROC curves (right panels) for the ST_{qb} data sample, in 3j1t and 2j1t configurations.

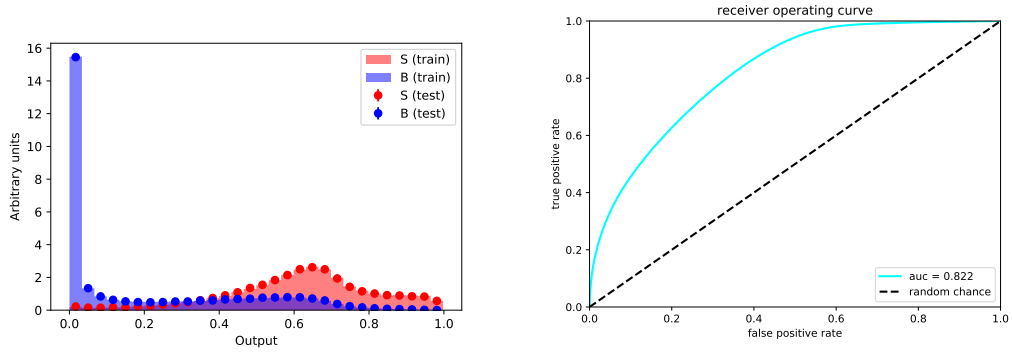


(a) 3j1t configuration

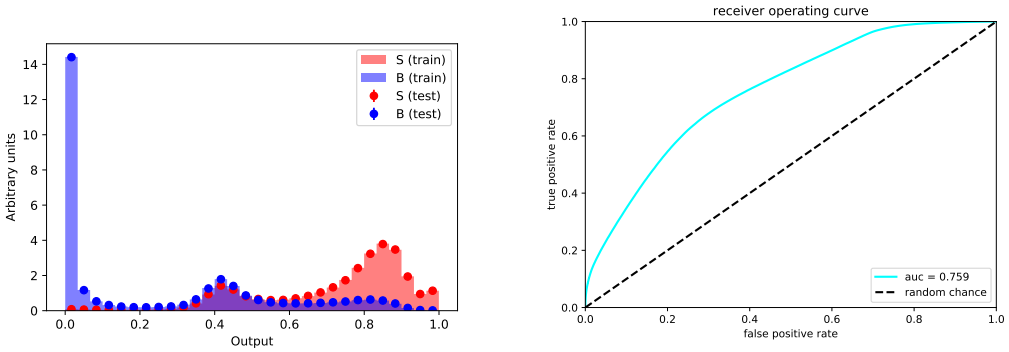


(b) 2j1t configuration

Figure A.4: NN output distribution for the signal and the background events (left panels) and corresponding ROC curves (right panels) for the ST_{qb} data sample, in 3j1t and 2j1t configurations.

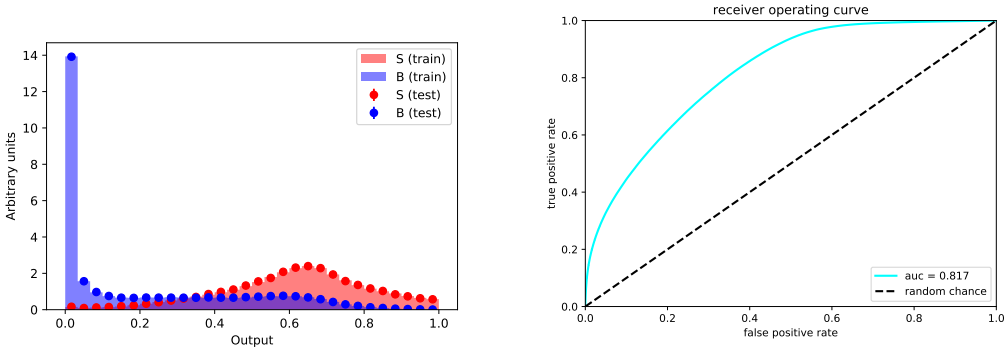


(a) 3j1t configuration

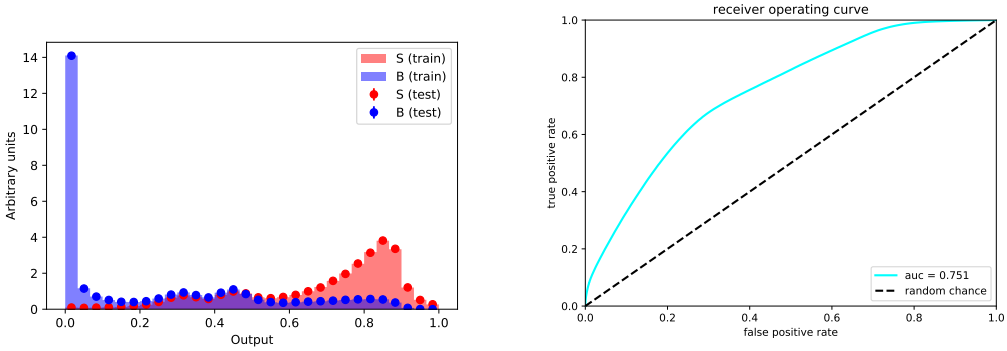


(b) 2j1t configuration

Figure A.5: XGB Classifier output distribution for the signal and the background events (left panels) and corresponding ROC curves (right panels) for the ST_{qb} data sample, in 3j1t and 2j1t configurations.



(a) 3j1t configuration



(b) 2j1t configuration

Figure A.6: NN output distribution for the signal and the background events (left panels) and corresponding ROC curves (right panels) for the $t\bar{t}$ data sample, in 3j1t and 2j1t configurations.

Bibliography

[1] Particle Data Group, M. Tanabashi *et al.*, "*Review of Particle Physics*", Phys. Rev. **D98**, 030001 (2018).

[2] S. L. Glashow, "*Partial Symmetries of Weak Interactions*", Nucl. Phys. **22**, 579 (1961).

[3] S. Weinberg, "*A Model of Leptons*", Phys. Rev. Lett. **19**, 1264 (1967).

[4] A. Salam, *Elementary Particle Theory* (, 1969).

[5] N. Cabibbo, "*Unitary Symmetry and Leptonic Decays*", Phys. Rev. Lett. **10**, 531 (1963).

[6] M. Kobayashi and T. Maskawa, "*CP-Violation in the Renormalizable Theory of Weak Interaction*", Progress of Theoretical Physics **49**, 652 (1973), http://oup.prod.sis.lan/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf.

[7] Z. Maki, M. Nakagawa, Y. Ohnuki, and S. Sakata, "*A Unified Model for Elementary Particles:* ", Progress of Theoretical Physics **23**, 1174 (1960), http://oup.prod.sis.lan/ptp/article-pdf/23/6/1174/5461148/23-6-1174.pdf.

[8] L.-L. Chau and W.-Y. Keung, "*Comments on the Parametrization of the Kobayashi-Maskawa Matrix*", Phys. Rev. Lett. **53**, 1802 (1984).

[9] L. Wolfenstein, "*Parametrization of the Kobayashi-Maskawa Matrix*", Phys. Rev. Lett. **51**, 1945 (1983).

[10] CKMfitter Group, J. Charles *et al.*, "*CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories*", Eur. Phys. J. **C41**, 1 (2005), hep-ph/0406184, updated results and plots available at: `http://ckmfitter.in2p3.fr`.

[11] CDF, F. Abe *et al.*, "*Observation of top quark production in $\bar{p}p$ collisions*", Phys. Rev. Lett. **74**, 2626 (1995), hep-ex/9503002.

[12] D. Heymes, Differential top-quark pair production at NNLO, in *Proceedings, 51st Rencontres de Moriond on QCD and High Energy Interactions: La Thuile, Italy, March 19-26, 2016*, pp. 23–26, ARISF, ARISF, 2016.

[13] M. Grazzini, S. Kallweit, and M. Wiesemann, "*Fully differential NNLO computations with MATRIX*", Eur. Phys. J. **C78**, 537 (2018), 1711.06631.

[14] CMS, A. M. Sirunyan *et al.*, "*Observation of Single Top Quark Production in Association with a Z Boson in Proton-Proton Collisions at* $\sqrt{s}$ *=13 TeV*", Phys. Rev. Lett. **122**, 132003 (2019), 1812.05900.

[15] CMS, A. M. Sirunyan *et al.*, "*Evidence for the associated production of a single top quark and a photon in proton-proton collisions at* $\sqrt{s} = 13$ *TeV*", Phys. Rev. Lett. **121**, 221802 (2018), 1808.02913.

[16] M. Jezabek and J. H. Kuhn, "*QCD Corrections to Semileptonic Decays of Heavy Quarks*", Nucl. Phys. **B314**, 1 (1989).

[17] I. I. Y. Bigi, "*The QCD perspective on lifetimes of heavy flavor hadrons*", (1995), hep-ph/9508408.

[18] G. Kane, G. Ladinsky, and C. Yuan, "*Using the top quark for testing standard-model polarization and CP predictions*", Physical review D: Particles and fields **45**, 124 (1992).

[19] A. Buckley *et al.*, "*Constraining top quark effective theory in the LHC Run II era*", JHEP **04**, 015 (2016), 1512.03360.

[20] CDF, T. Aaltonen *et al.*, "*Forward-Backward Asymmetry in Top Quark Production in* $p\bar{p}$ *Collisions at sqrts* $= 1.96$ *TeV*", Phys. Rev. Lett. **101**, 202001 (2008), 0806.2472.

[21] CMS, G. K. Krintiras, Results on TOP physics from CMS, in *International Workshop on Future Linear Collider (LCWS2017) Strasbourg, France, October 23-27, 2017*, 2018, 1801.08398.

[22] ATLAS, T. A. collaboration, "*Measurement of the top-quark decay width in top-quark pair events in the dilepton channel at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*", (2019).

[23] CDF, T. Aaltonen *et al.*, "*Exclusion of exotic top-like quarks with -4/3 electric charge using jet-charge tagging in single-lepton ttbar events at CDF*", Phys. Rev. **D88**, 032003 (2013), 1304.4141.

[24] D0, V. M. Abazov *et al.*, "*Measurement of the Electric Charge of the Top Quark in* $t\bar{t}$ *Events*", Phys. Rev. **D90**, 051101 (2014), 1407.4837, [Erratum: Phys. Rev.D90,no.7,079904(2014)].

[25] CMS, C. Collaboration, "*Constraints on the Top-Quark Charge from Top-Pair Events*", (2012).

[26] ATLAS, G. Aad *et al.*, "*Measurement of the top quark charge in pp collisions at* $\sqrt{s} = 7$ *TeV with the ATLAS detector*", JHEP **11**, 031 (2013), 1307.4568.

[27] U. Baur, A. Juste, L. H. Orr, and D. Rainwater, "*Probing electroweak top quark couplings at hadron colliders*", Phys. Rev. **D71**, 054013 (2005), hep-ph/0412021.

[28] O. Bessidskaia Bylund, F. Maltoni, I. Tsinikos, E. Vryonidou, and C. Zhang, "*Probing top quark neutral couplings in the Standard Model Effective Field Theory at NLO in QCD*", JHEP **05**, 052 (2016), 1601.08193.

[29] M. Schulze and Y. Soreq, "*Pinning down electroweak dipole operators of the top quark*", Eur. Phys. J. **C76**, 466 (2016), 1603.08911.

[30] CDF, T. Aaltonen *et al.*, "*Evidence for $t\bar{t}\gamma$ Production and Measurement of $\sigma_{t\bar{t}\gamma}/\sigma_{t\bar{t}}$*", Phys. Rev. **D84**, 031104 (2011), 1106.3970.

[31] ATLAS, G. Aad *et al.*, "*Observation of top-quark pair production in association with a photon and measurement of the $t\bar{t}\gamma$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector*", Phys. Rev. **D91**, 072007 (2015), 1502.00586.

[32] ATLAS, M. Aaboud *et al.*, "*Measurement of the $t\bar{t}\gamma$ production cross section in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*", JHEP **11**, 086 (2017), 1706.03046.

[33] G. Mahlon and S. J. Parke, "*Maximizing spin correlations in top quark pair production at the Tevatron*", Phys. Lett. **B411**, 173 (1997), hep-ph/9706304.

[34] G. Mahlon and S. J. Parke, "*Single top quark production at the LHC: Understanding spin*", Phys. Lett. **B476**, 323 (2000), hep-ph/9912458.

[35] M. Jezabek and J. H. Kuhn, "*V-A tests through leptons from polarized top quarks*", Phys. Lett. **B329**, 317 (1994), hep-ph/9403366.

[36] J. A. Aguilar-Saavedra and J. Bernabeu, "*W polarisation beyond helicity fractions in top quark decays*", Nucl. Phys. **B840**, 349 (2010), 1005.5382.

[37] J. A. Aguilar-Saavedra, "*Single top quark production at LHC with anomalous Wtb couplings*", Nucl. Phys. **B804**, 160 (2008), 0803.3810.

[38] J. A. Aguilar-Saavedra, "*A Minimal set of top anomalous couplings*", Nucl. Phys. **B812**, 181 (2009), 0811.3842.

[39] F. Bach and T. Ohl, "*Anomalous Top Couplings at Hadron Colliders Revisited*", Phys. Rev. **D86**, 114026 (2012), 1209.4564.

[40] CMS, V. Khachatryan *et al.*, "*Measurement of Top Quark Polarisation in T-Channel Single Top Quark Production*", JHEP **04**, 073 (2016), 1511.02138.

[41] L. Evans and P. Bryant, "*LHC Machine*", Journal of Instrumentation **3**, S08001 (2008).

[42] CMS, G. L. Bayatian *et al.*, "*CMS Physics*", (2006).

[43] CMS, S. Chatrchyan *et al.*, "*Energy Calibration and Resolution of the CMS Electromagnetic Calorimeter in pp Collisions at $\sqrt{s} = 7$ TeV*", JINST **8**, P09009 (2013), 1306.2016, [JINST8,9009(2013)].

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

[45] CMS, P. Sphicas, "*CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger*", (2002).

[46] CMS, V. Khachatryan *et al.*, "*The CMS trigger system*", JINST **12**, P01020 (2017), 1609.02366.

[47] T. Sjostrand, S. Mrenna, and P. Z. Skands, "*A Brief Introduction to PYTHIA 8.1*", Comput. Phys. Commun. **178**, 852 (2008), 0710.3820.

[48] CMS, V. Khachatryan *et al.*, "*Event generator tunes obtained from underlying event and multiparton scattering measurements*", Eur. Phys. J. **C76**, 155 (2016), 1512.00815.

[49] CMS, C. Collaboration, "*Investigations of the impact of the parton shower tuning in Pythia 8 in the modelling of $\mathrm{t\bar{t}}$ at $\sqrt{s} = 8$ and 13 TeV*", (2016).

[50] NNPDF, R. D. Ball *et al.*, "*Parton distributions for the LHC Run II*", JHEP **04**, 040 (2015), 1410.8849.

[51] J. Allison *et al.*, "*Geant4 developments and applications*", IEEE Trans. Nucl. Sci. **53**, 270 (2006).

[52] CMS, C. Collaboration, "*Extraction of CKM matrix elements in single top quark t-channel events in proton-proton collisions at $\sqrt{s} = 13$ TeV*", (2019).

[53] S. Frixione, P. Nason, and C. Oleari, "*Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*", JHEP **11**, 070 (2007), 0709.2092.

[54] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, "*Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*", JHEP **03**, 015 (2013), 1212.3460.

[55] J. Alwall *et al.*, "*The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*", JHEP **07**, 079 (2014), 1405.0301.

[56] R. Frederix and S. Frixione, "*Merging meets matching in MC@NLO*", JHEP **12**, 061 (2012), 1209.6215.

[57] E. Re, "*Single-top Wt-channel production matched with parton showers using the POWHEG method*", Eur. Phys. J. **C71**, 1547 (2011), 1009.2450.

[58] T. Melia, P. Nason, R. Rontsch, and G. Zanderighi, "*W+W-, WZ and ZZ production in the POWHEG BOX*", JHEP **11**, 078 (2011), 1107.5051.

[59] P. Nason and G. Zanderighi, "W^+W^- *, WZ and ZZ production in the POWHEG-BOX-V2*", Eur. Phys. J. **C74**, 2702 (2014), 1311.1365.

[60] CMS, F. Beaudette, The CMS Particle Flow Algorithm, in *Proceedings, International Conference on Calorimetry for the High Energy Frontier (CHEF 2013): Paris, France, April 22-25, 2013*, pp. 295–304, 2013, 1401.8155.

[61] W. Adam, R. Frühwirth, A. Strandlie, and T. Todor, "*Reconstruction of Electrons with the Gaussian-Sum Filter in the CMS Tracker at the LHC*", (2005).

[62] CMS, S. Chatrchyan *et al.*, "*Missing transverse energy performance of the CMS detector*", JINST **6**, P09001 (2011), 1106.5048.

[63] `https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonTagAndProbeTreesRun2.`

[64] CMS, S. Chatrchyan *et al.*, "*Description and performance of track and primary-vertex reconstruction with the CMS tracker*", JINST **9**, P10009 (2014), 1405.6569.

[65] `https://twiki.cern.ch/twiki/bin/viewauth/CMS/CutBasedElectronIdentificationRun2#Electron_ID_Working_Points_WP_de`.

[66] CMS Collaboration, CERN Report No. CMS-PAS-JME-14-001, 2014 (unpublished).

[67] M. Cacciari, G. P. Salam, and G. Soyez, "*The anti-k_t jet clustering algorithm*", JHEP **04**, 063 (2008), 0802.1189.

[68] L. Lista, *Statistical Methods for Data Analysis in Particle Physics*Lecture Notes in Physics (Springer International Publishing, 2015).

[69] J. Neyman and E. S. Pearson, "*On the Problem of the Most Efficient Tests of Statistical Hypotheses*", Phil. Trans. Roy. Soc. Lond. **A231**, 289 (1933).

[70] S. Raschka, *Python Machine Learning* (Packt Publishing, Birmingham, UK, 2015).

[71] F. Chollet *et al.*, Keras, `https://keras.io`, 2015.

[72] T. Chen and C. Guestrin, XGBoost: A scalable tree boosting system, in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '16, pp. 785–794, New York, NY, USA, 2016, ACM.

[73] R. J. Barlow and C. Beeston, "*Fitting using finite Monte Carlo samples*", Comput. Phys. Commun. **77**, 219 (1993).

[74] J. Conway, "*Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra*", p. 115 (2011), Comments: Presented at PHYSTAT 2011, CERN, Geneva, Switzerland, January 2011, to be published in a CERN Yellow Report.

[75] T. C. collaboration, "*Determination of jet energy calibration and transverse momentum resolution in CMS*", Journal of Instrumentation **6**, P11002 (2011).

[76] CMS, C. Collaboration, "*CMS Luminosity Based on Pixel Cluster Counting - Summer 2013 Update*", (2013).

[77] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "*Asymptotic formulae for likelihood-based tests of new physics*", Eur. Phys. J. **C71**, 1554 (2011), 1007.1727, [Erratum: Eur. Phys. J.C73,2501(2013)].

Acknowledgments

Questo percorso di tesi per me ha rappresentato il coronamento di una crescita accademica e personale, in cui ho potuto osservare e sentirmi un po' parte del mondo della ricerca. Ho imparato ad affrontare la paura di "non essere abbastanza", ad accettare i miei limiti, e a prendere consapevolezza dei miei punti di forza.

Il merito di tutto ciò va in primis al dottor Orso Iorio per essere stato una guida sempre disponibile e competente in questo percorso di tesi, rendendomi fiera del mio lavoro forse per la prima volta, e per avermi dato la possibilità di approcciarmi al mondo della ricerca. Per me è stato un modello, professionalmente, per tutto quello che mi ha insegnato in questo percorso, ed umanamente, per la passione che mi ha trasmesso in questo lavoro e per l'analisi dati. Grazie, davvero.

Ringrazio il prof. Ambrosino, il mio correlatore, per i preziosi consigli che mi ha dato in questo percorso di tesi.

Ringrazio la mia mamma, Isabella, perchè mi è *sempre* stata vicina anche quando l'ho data per scontata, e perchè mi ha sempre incoraggiata in questo percorso: grazie a lei ho compreso quanto sia speciale l'amore di una mamma. Spesso in questo periodo mi ha detto: "ricordo il giorno in cui tu andasti in prima elementare, ed ora hai finito". Ora che ho finito non so esattamente che strada intraprenderò, ma so che lei è e sarà sempre il modello di donna che voglio diventare. Ringrazio il mio papà, Cosmo, perchè la sua dolcezza e la sua stima sono state determinanti in questo percorso. Grazie a lui ho sempre trovato una ragione per rialzarmi anche quando ero a terra. Non ci sono parole per esprimere quanto io vi sia grata perchè mi avete sempre insegnato ad inseguire i miei sogni, e se li ho realizzati è soprattutto grazie a voi.

Ringrazio Roberto, per tutto. Per quello che c'è fra di noi, perchè fin dal primo momento in cui ci siamo conosciuti c'è stata una grande complicità che è la stessa che giorno per giorno ci porta a migliorarci l'un l'altra e ad amarci sempre di più, perchè ha sempre creduto in me anche quando io avevo smesso di farlo. Non c'è dimostrazione più bella d'amore.

Questa esperienza, così come quella in America, non sarebbe stata così bella se non ci fosse stato con me Francesco. Dopo un paio di ore che ci conoscevamo già ci sembrava di conoscerci da una vita. E' davvero un amico prezioso, ha sempre saputo darmi il consiglio giusto e mi è sempre stato vicino. Non so davvero come avrei fatto senza di lui. Grazie anche perchè lui e Daniele sono stati i migliori compagni di studio che potessi avere, siamo stati un fantastico "terzetto". Le giornate di studio, in cui invadevamo casa di uno di noi, erano sempre stimolanti e passavano in un baleno.

In questo percorso ho conosciuto tante persone, alcune sono andate via, altre sono rimaste anche nei momenti più difficili. Ringrazio Ludovica e Sara, perchè sono entrate nella mia vita e in men che non si dica il legame tra noi è diventato indissolubile. Ci sono state in tanti momenti importanti, abbiamo vissuto insieme tanti cambiamenti e tante nuove esperienze, che hanno reso la nostra amicizia solida come se fossimo amiche da una vita. Grazie Ludovica, perchè siamo così simili e ci *triggeriamo* a vicenda. Grazie Sara, perchè siamo così diverse, e ci compensiamo. Non ci sono frasi, parole, che possano racchiudere quanto io sia grata della vostra presenza nella mia vita, siete davvero rare.

Ringrazio Mauro, perchè in questi anni è sempre stato un grande amico e confidente. Con la tua saggezza hai sempre trovato le parole giuste, con la tua simpatia hai sempre saputo dirmele nel modo giusto. Con il nostro legame quasi fraterno non ci siamo mai persi. Ringrazio Gege perchè ha sempre creduto in me, è sempre stato un amico leale e disponibile, e mi ha sempre spronato a superare i miei limiti. È una persona da cui ho molto da imparare. Grazie a Damiano, Mattia, Roberto e a tutti i ragazzi del *Secondo Tavolo a sinistra*, anche le giornate più buie sono state illuminate dalla loro dolcezza e dai loro consigli, ma soprattutto mi hanno ricordato *quante cose al mondo posso fare*.

Durante il percorso accademico, una persona in un certo senso delinea la propria identità. Lo studio, la passione, l'impegno, le esperienze, le relazioni umane, sono ingredienti determinanti alla formazione della persona, permettendole di sviluppare una maturità tale da essere in grado di scegliere chi vuole diventare al termine degli studi. Ma non è solo tra le mura di un ateneo che avviene questa crescita. Ringrazio Ester, perchè siamo così diverse, ed è grazie alla nostra diversità che siamo così amiche. Da quando è arrivata in terza media e siamo diventate compagne di banco ed io parlavo senza fermarmi un attimo e lei invece era così tranquilla, le nostre differenze e il nostro affetto hanno reso la nostra amicizia unica. Il nostro essere diametralmente opposte ci porta ad avere continuamente pareri differenti su tutto, ed è per questo che è sempre stimolante parlare con lei. Grazie per tutti questi anni di amicizia. Ringrazio Giorgia, la mia *amica gemella*, una delle persone che stimo di più. Ci conosciamo da più di metà delle nostre vite ormai, di avventure insieme ne abbiamo vissute tante. A qualsiasi ora del giorno e della notte mi ha sempre tranquillizzata sulle mie ipocondrie (ma non solo) ed è sempre stata disponibile ad ascoltarmi. In quello che fa, così come nelle relazioni umane, mette tanta passione, e la trasmette a tutti quelli che le sono intorno, e questa è una dote rara. Di persone come lei non se ne trovano tante in giro.

Ringrazio Alice, perchè mi ha sempre sostenuta ed è sempre stata una grande amica, leale e presente in ogni momento.

Ringrazio Veronica perchè c'è sempre stata nei momenti più importanti della mia vita.

Tra le esperienze più belle in questi anni c'è stata la scuola estiva al Fermilab, che mi ha permesso di imparare tante cose, e mi ha fatto conoscere tante splendide persone. Tra queste primeggia Cecilia, *la Zezi*, la mia compagna di stanza e di viaggi, sempre *componente*. La nostra amicizia va oltre le prime impressioni e la distanza. Ci siamo sempre sostenute e spronate in questo percorso di tesi che abbiamo portato avanti contemporaneamente. Sono davvero contenta di aver trovato in lei una grande amica.

Per coltivare un'amicizia non è necessaria la quotidianità quando c'è un profondo legame alla base. A prova di ciò c'è Mattia: tra una chiamata su Skype e un caffè quando torna a Napoli, il nostro legame non si è mai indebolito nonostante siano anni che lui non vive più in Italia.

Ringrazio mia cugina Federica, che mi ha insegnato ad avere il coraggio di crescere. Il nostro rapporto ha subito tante variazioni di paradigma nel corso degli anni, ma il bene profondo che ci vogliamo è sempre rimasto immutato.

Ringrazio inoltre tutte le persone che ho incontrato in questi anni che abbiano anche solo in minima parte contribuito a questo traguardo.

Ringrazio infine me stessa per aver sempre saputo cogliere le occasioni, per l'impegno e la dedizione che ho messo in questo lavoro. La mia passione ha sempre vinto su tutto.

Valeria