Alma Mater Studiorum · Università di Bologna

Scuola di Scienze
Corso di Laurea Magistrale in Fisica

STUDY OF A NEW METHOD FOR HIGH ENERGY TOP TAGGING AT THE ATLAS EXPERIMENT

Relatore:
Prof. Nicola Semprini Cesari

Correlatore:
Dott. Roberto Spighi

Presentata da:
Camilla Vittori

Sessione II
Anno Accademico 2013/2014

Abstract

A partire dalla sua scoperta, il top ha sempre svolto un ruolo di notevole interesse nella fisica delle particelle. Lo scopo di questa tesi è la ricostruzione di top adronici con un alto impulso trasverso ("boosted") attraverso il Template Overlap Method (TOM). A causa dell'alta energia, i prodotti di decadimento dei boosted top sono parzialmente o totalmente sovrapposti e risultano contenuti in un singolo jet di grandi dimensioni ("fat-jet"). Il TOM confronta le distribuzioni di energia del fat-jet con campioni di top ottenuti con simulazioni Monte Carlo ("template"). L'algoritmo è basato sulla definizione di una funzione di overlap, che quantifica il livello di accordo tra il fat-jet e il template, consentendo un'efficiente discriminazione del segnale dai contributi di fondo. Per ottenere un'efficienza sul segnale attorno al 90% e una corrispondente reiezione dal fondo del 70%, è stato necessario stabilire un punto di lavoro. Le performance del TOM sono state testate su campioni MC nel canale muonico e confrontate con i metodi presenti in letteratura. Tali metodi saranno inseriti in un'analisi multivariata al fine di creare un metodo di tagging globale che sarà incluso nella misura della sezione d'urto differenziale della produzione di coppie $t\bar{t}$ sui dati acquisiti nel 2012 a $\sqrt{s}$=8 TeV, nella regione dello spazio delle fasi in cui potrebbero essere possibili processi di nuova fisica. A causa della sua caratteristica di aumentare l'efficienza di identificazione all'aumento del p_T, il Template Overlap Method giocherà un ruolo fondamentale durante la prossima presa dati a $\sqrt{s}$=13 TeV, dove quasi la totalità dei top sarà prodotta ad alta energia, rendendo impossibile l'identificazione con le tecniche standard.

Abstract

Since its discovery, top quark has represented one of the most investigated field in particle physics. The aim of this thesis is the reconstruction of hadronic top with high transverse momentum ("boosted") with the Template Overlap Method (TOM). Because of the high energy, the decay products of boosted tops are partially or totally overlapped and thus they are contained in a single large radius jet ("fat-jet"). TOM compares the internal energy distributions of the candidate fat-jet to a sample of tops obtained by a MC simulation ("template"). The algorithm is based on the definition of an overlap function, which quantifies the level of agreement between the fat-jet and the template, allowing an efficient discrimination of signal from the background contributions. A working point has been decided in order to obtain a signal efficiency close to 90% and a corresponding background rejection at 70%. TOM performances have been tested on MC samples in the muon channel and compared with the previous methods present in literature. All the methods will be merged in a multivariate analysis to give a global top tagging which will be included in the measurement of the $t\bar{t}$ production differential cross section performed on the data acquired in 2012 at $\sqrt{s}$=8 TeV in high phase space region, where new physics processes could be possible. Due to its peculiarity to increase the identification efficiency with respect the top p_T, the Template Overlap Method will play a crucial role in the next data taking at $\sqrt{s}$=13 TeV, where the almost totality of the tops will be produced at high energy, making the standard reconstruction methods inefficient.

Contents

List of Figures

List of Tables

Introduction

With unprecedented high center-of-mass energy and luminosity, LHC provided an important development in the study of top quark, allowing to perform high-statistic measurements. Since its discovery, the study of the top quark has represented one of the most investigated field in particles physics, because of its peculiar properties, as the largest mass and the smallest decay time, that offer the unique possibility to study a bare quark.

The analysis presented in this thesis is focused on the reconstruction of the hadronic top decay ($t \rightarrow Wb \rightarrow qq^{'}b$) at high momentum; the results of this study will improve the measurement of the $t\bar{t}$ production differential cross section, performed on data collected by the ATLAS detector on 2012.

At high energy (p_T>300 GeV), the decay products of hadronic top quarks are so collimated that the standard reconstruction techniques begin to fail because the separation among the three emitted jets becomes negligible and they tend to be superimposed in a single, energetic and large radius jet (*fat-jet*). The aim of this analysis is to reconstruct high energy top quarks with the *Template Overlap Method* (TOM), a new procedure, still not applied in the standard analysis, specially optimized for hadronic top decays.

TOM performances have been evaluated on MC samples in the muon channel; the method has provided similar performances compared with the published results of previous techniques. At the moment, a working point has been chosen in order to have a signal efficiency of about 90% and a background rejection of about 70%.

An important TOM feature is the increase of the efficiency with the top p_T,

that is crucial for two reasons: first, it permits to study top produced in high momentum phase space region with the data acquired at $\sqrt{s}$= 8 TeV, where eventual processes coming from new physics are expected; second, it allows to reconstruct the top acquired in the next data taking at $\sqrt{s}$= 13 TeV, where the almost totality of them will be produced at high energy and the standard reconstruction will be not efficient.

The structure of the thesis is the following. In Chapter I a description of the Standard Model with particular attention to top quark features is presented. A synthetic panorama of the ATLAS detector is given in Chapter II in order to have a better comparison of the following analysis. A detailed description of the Template Overlap Method is provided in Chapter III, associated to a comparison with the other boosted top tagging algorithms. In Chapter IV the obtained results of the Template Overlap Method together with its systematics and the comparison with previous results are presented. In the end, the Conclusions.

Chapter 1

Top quark physics

1.1 The Standard Model

Developed in the early 1970s, the *Standard Model* (SM) is the theory which successfully describes the fundamental particles and the interactions among them (see Fig.1.1) in the language of grand unification theory. The SM forces involving fundamental particles are the electromagnetic, the weak and the strong interactions, while, until now, it as not been possible to construct a consistent theory of the gravitational interaction.
According to the strong force, the ultimate constituents of matter are divided into leptons and quarks, all point-like fermions with spin 1/2 [1]. On the basis of the weak interaction, leptons and quarks are both divided into three weak isospin doublets (see Tab.1.1 and 1.2), each one consisting of a massive charged particle (e, μ and τ) and the corresponding neutrino (ν_e, ν_μ and ν_τ).

$$\begin{pmatrix} e \\ \nu_e \end{pmatrix} \begin{pmatrix} \mu \\ \nu_\mu \end{pmatrix} \begin{pmatrix} \tau \\ \nu_\tau \end{pmatrix}$$

While electron was known from the end of XIX century, the muon, considered an unstable "heavy electron", was the first particle not involved in the structure of ordinary matter to be discovered (observed in cosmic rays in 1937). Tau was first revealed in accelerator experiments in 1974 and neutrino, after been

predicted by Pauli's theory, was finally found in beta decay in the 1950s [1]. Particles are identified by quantum numbers thus, for instance, to leptons corresponds the *leptonic number* L conserved by all the interactions. Each weak doublet is described by leptonic number L_e, L_μ and L_τ, approximatively conserved by all the interactions. In Tab.1.1 a summary of all lepton quantum numbers have been listed together with a quotation of neutrino mass superior limits [2].

Table 1.1: Standard Model leptons

Lepton	Q ($\|e\|$)	L	L_e	L_μ	L_τ	Mass (MeV/c^2)
e	-1	+1	1	0	0	0,511
ν_e	0	+1	1	0	0	$< 2{,}2\ 10^{-6}$
μ	-1	+1	0	1	0	105,65
ν_μ	0	+1	0	1	0	< 0.19
τ	-1	+1	0	0	1	1777,82
ν_τ	0	+1	0	0	1	$< 18{,}2$

Quarks occur in six different *flavours*, represented by the assignment of quantum numbers labeled *u, d, c, s, t, b* (see Fig.1.1). Because of the similarity between up and down mass values, these two quarks are grouped in a strong isospin doublet ($I = 1/2$, with $I_3 = \pm 1/2$ as third component). While leptons carry an integer charge value (0 or $\pm 1|e|$), quarks carry fractional charge; each weak doublet of quarks contains a quark with charge $+2/3|e|$ and another one with charge $-1/3|e|$.

$$\begin{pmatrix} u \\ d \end{pmatrix} \begin{pmatrix} c \\ s \end{pmatrix} \begin{pmatrix} t \\ b \end{pmatrix}$$

Because of **confinement**, the property of the strong interaction which force quarks bound in hadrons, quarks cannot exist as free particles. All the observed hadrons are quark-antiquark (*mesons*) or quark-quark-quark (*baryons*) combinations. To all quarks an additional quantum number is associated, the

baryon number, conserved by all the interactions, whose value is 1/3 ($-1/3$ for anti-quarks and 0 for leptons). The quark quantum numbers are listed in Tab.1.2 [2].

Table 1.2: Standard Model quarks

Quark	Q ($\|e\|$)	I	I_3	C	S	T	B	Mass (GeV/c^2)
u	+2/3	1/2	+1/2	0	0	0	0	2,3 $\cdot 10^{-3}$
d	-1/3	1/2	-1/2	0	0	0	0	4,8 $\cdot 10^{-3}$
c	+2/3	0	0	1	0	0	0	1,275
s	-1/3	0	0	0	1	0	0	95 $\cdot 10^{-3}$
t	+2/3	0	0	0	0	1	0	173,07
b	-1/3	0	0	0	0	0	1	4,18

In order to understand certain properties of hadrons is necessary to introduce for each flavour the *colour charge*, which can assume three possible values: red, blue and green. Considering that to each particle corresponds an antiparticle, with opposite quantum numbers, the total number of the fundamental particles allowed in the Standard Model amounts to

$$[6(leptons) + 6(quarks) \times 3(colours)] \times 2 = 48 \quad .$$

Table 1.3: Standard Model gauge bosons

Force	Gauge boson	Q ($\|e\|$)	Mass (GeV/c^2)
Strong	gluon (g)	0	0
Electromagnetic	photon (γ)	0	0
Weak	$W^{\pm}$	± 1	80,385 $\pm$ 0,015
Weak	Z^0	0	91,1876 $\pm$ 0,0021

In the Standard Model, particles interact with each other by coupling with specific fields whose quanta are spin-1 particles (*bosons* [2]). The field quanta of electromagnetic, weak and strong forces are respectively the photon γ, three massive particles W^+, W^- and Z^0 and eight gluons (see Tab.1.3). Of these,

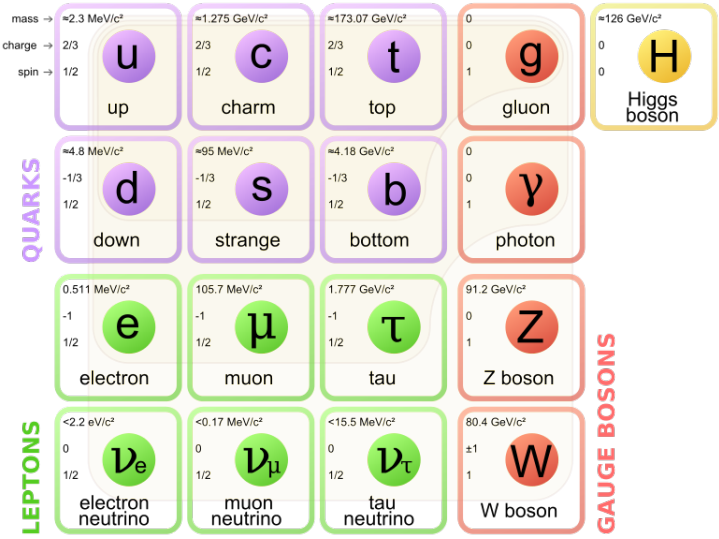


Figure 1.1: The fundamental fermions and bosons of the Standard Model.

only $W^{\pm}$ and Z^0 have mass because of the interaction with the Higgs field, a property which ensures the typical short range of the weak interaction. Charged leptons can interact through both the electromagnetic and the weak forces, while quarks, which are coloured particles, are affected by the strong interaction too, otherwise neutrino can interact only through the weak force. In order to show the relative magnitudes of the fundamental forces, the strong interaction amplitude has been fixed to 1 and all the other are refereed to it:

Strong	Electromagnetic	Weak	Gravitational
1	10^{-2}	10^{-7}	10^{-39} .

Fig.1.1 illustrates all particles allowed in the Standard Model, including gauge bosons.

1.1.1 The electromagnetic interaction

Described by quantum field theories, all the Standard Model interactions arise from the coupling between particles and fields. The intensity of the interactions is described by coupling constants, which enter in the matrix element of

each process [3]. In the SM, the structure of the different interactions is determined by a symmetry principle requiring that the corresponding Lagrangian is invariant under local gauge transformations. In this way, all the terms of the Lagrangian can be generated starting from the known term of the free material particle.

In the specific case of the electromagnetic interaction, the coupling of charged particles with the electromagnetic field is due to the electric charge. The quantum field theory describing this interaction is the Quantum Electrodynamics (QED) [4], symmetric with respect gauge rotation of U(1) group. The QED coupling constant is called fine-constant, a dimensionless quantity defined as

$$\alpha_e = \frac{e^2}{4\pi\epsilon_0\hbar c} = \frac{1}{137} \quad , \tag{1.1}$$

where e is the electric charge. The coupling constant is a function of energy and for this reason it is called "running". According to quantum field theory, in the vacuum medium photon emission, pair annihilation and pair creation phenomena happen continuously: this effect is called **vacuum polarization**. If a charged sphere is present, the e^+e^- pairs become oriented, forming a virtual cloud around the charged body. The net effect is a screening of the sphere and thus a gradually reduction of the power of its charge at increasing distance from it. In this ideal experiment, the distance of closest approach of a probe to the charge is a decreasing function of the energy of the probe: consequently, high-energy probes will see a larger charge on the sphere.

The QED Lagrangian can be obtained from the free Dirac Lagrangian:

$$\mathcal{L}_{free} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi \quad , \tag{1.2}$$

requiring the invariance under global and local gauge transformation in the electric charge space. The invariance under a global phase rotation, which is a continuous symmetry, through the Noether's theorem, leads to the conservation of the electric charge ($e = \sqrt{4\pi\alpha_e}$). Generalizing the global phase symmetry to a local one, allows to pass from a theory describing free particles to a theory in which particles experience electromagnetic interaction. In order

to preserve Lagrangian invariance under local gauge rotation, the introduction of the gauge covariant derivative is necessary

$$\partial_\mu \rightarrow \mathcal{D}_\mu = \partial_\mu + iqA_\mu(x) \quad , \tag{1.3}$$

where the quanta of the vector field A_μ is the photon. The free-particle Lagrangian of Eq.(1.2) is replaced by the locally gauge-invariant expression

$$\mathcal{L}_{QED} = \mathcal{L}_{free} - J^\mu A_\mu - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \quad , \tag{1.4}$$

which is indeed the QED Lagrangian. In Eq.(1.4) J^μ is the conserved electromagnetic current and the last term represents the propagation of free photons, in which $F_{\mu\nu}$ is the Maxwell's electromagnetic tensor ($F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$). A photon mass term with the form $\mathcal{L}_\gamma = \frac{1}{2} m^2 A^\mu A_\mu$ is not present in the expression of the QED Lagrangian because it would violate the local gauge invariance: this leads to the existence of massless photon.

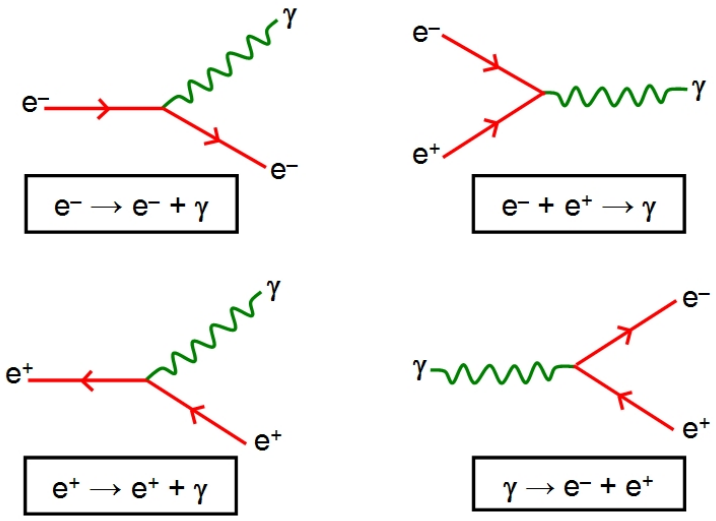


Figure 1.2: Feynman diagrams of the fundamental QED vertex (*top right*), the e^+e^- annihilation (*top left*), the emission of a photon by a positron (*bottom right*) and the couple creation by a photon (*top left*). All of these diagrams can be obtained through the fundamental vertex.

1.1.2 The weak interaction

The weak interaction takes place between all fundamental particles of the Standard Model. Because of the small strength of this force compared to electromagnetic and strong ones, weak interactions are observable only when the other forces cannot occur. The quantum field theory describing the weak interaction alone is often called Quantum Flavordynamics (QFD) [4], symmetric with respect gauge rotation of $SU(2)_L$ group ($SU(2)_L$ indicates that only left-handed particles can couple with the weak field). Three vector bosons mediate this interaction, two are electrically charged, W^+ and W^-, each the antiparticle of the other, and one is neutral, Z^0. In the weak interaction vertex two particles interact exchanging a vector boson: if it is a W, the charges of fermions in the final and initial states differ by a unit and the process is called "charge current interaction" (CC), if it is the Z, the two electric charges are equal and the process is labelled as "neutral current interaction" (NC). Fig.1.3 shows the fundamental vertexes of the weak interaction.

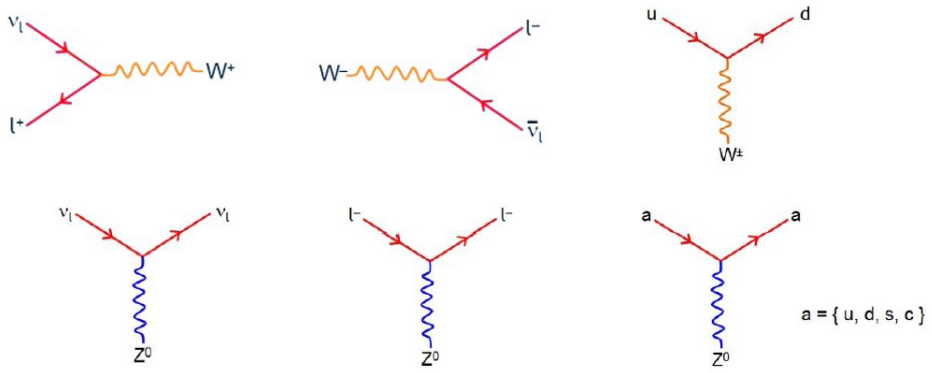


Figure 1.3: Fundamental vertexes of the weak interaction in both charged current CC (*top*) and neutral current (*bottom*).

The weak interaction was first observed in the process of β decay

$$n \rightarrow p + e^- + \overline{\nu_e} \tag{1.5}$$

and explained by Fermi through an effective pointlike theory. In this appro-

ximation, successful at low momentum transferred ($q^2 << M_W^2$), the virtual W boson exchanged was neglected and the decay was described by the Fermi coupling constant $G_F = g^2/M_W^2$, where g^2 is the weak charge. With the use of Fermi's approximation, it has been possible to measure the rates of lots of weak decays and to verify that they have the same coupling: this leads to the development of the *weak coupling universality* concept [1]. Fermi's phenomenological description of weak interaction was based on the similarity with the electromagnetic one; in order to improve the analogy it is important to leave the effective theory and to introduce the presence of a vector boson. The latter must carry charge $\pm 1|e|$ or 0, be enough massive to explain the short range of this force and with indefinite parity. In addiction, in order to involve the violation of parity and the coupling to left-handed particles $\Psi_L = \frac{1}{2}(1-\gamma_5)\Psi$ only, the structure of the weak interaction must be of type V-A (vector-axial vector).

Experimentally, decays into fermions belonging to the same weak doublet are found to be more frequent, although universality requires the corresponding matrix elements to be equal. In 1963 Cabibbo proposed the solution to this problem [5]. He assumed that d- and s-quark states participating in the weak interactions are rotated by a mixing angle θ_C, called "Cabibbo mixing angle"

$$\begin{pmatrix} u \\ d' \end{pmatrix} = \begin{pmatrix} u \\ d\, cos\theta_C + s\, sin\theta_C \end{pmatrix} \quad . \tag{1.6}$$

The same procedure can be applied to all the quark families. Therefore one can conclude that the eigenstates of the weak interaction do not coincide with the mass eigenstates, but are rotated by a unitary 3×3 matrix, called Cabibbo-Kobaiashi-Maskawa (**CKM**) matrix. The quark mixing transformation is:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad . \tag{1.7}$$

The squares of matrix elements give the decay amplitude probability and have

been experimentally determined [2]:

$$\begin{pmatrix} V_{ud} = 0,97427 \pm 0,00015 & V_{us} = 0,22534 \pm 0,00065 & V_{ub} = 0,00351^{+0.00015}_{-0.00014} \\ V_{cd} = 0,22520 \pm 0,00065 & V_{cs} = 0.97344 \pm 0.00016 & V_{cb} = 0.0412^{+0.0011}_{-0.0005} \\ V_{td} = 0,00867^{+0.00029}_{-0.00031} & V_{ts} = 0.0404^{+0.0011}_{-0.0005} & V_{tb} = 0.999146^{+0.000021}_{-0.000046} \end{pmatrix} \quad .$$

The off-diagonal values are small, therefore the corresponding mixing angles are small, while the diagonal elements are close to 1, meaning that the most favoured transitions are those happening among quarks that come from the same isospin doublet. Indeed the model predicts a specific sequence of decays: the top quark, for example, decays mostly $t \rightarrow W^+b$.

In 1967-1968 Weinberg and Salam proposed a gauge theory unifying weak and electromagnetic interactions, the so called **electroweak interaction** [6]. This force is based on the SU(2) group of weak isospin T and the U(1) group of weak ipercharge, with four generators and four massless gauge fields; both of them are connected with the electromagnetic charge Q by the relation

$$Q = T_3 + \frac{1}{2}Y \quad , \tag{1.8}$$

where T_3 is the third component of the weak isospin. The electroweak unification conserves the local gauge invariance, nevertheless it describes W and Z^0 as massless bosons, while they are massive particles, as proved by experiments. Through the introduction of the **Higgs mechanism** it is possible to preserve the local gauge invariance and to give mass to the vector boson of the weak interaction, keeping the photon massless. Predicted in 1960s as the main responsible of the electroweak symmetry breaking, the Higgs boson was finally discovered on 4 July 2012, with a mass around 125 GeV/c^2 [7].

1.1.3 The strong interaction

The strong interaction describes the interactions among quarks and gluons and how they bind together to form hadrons. Quantum Chromodynamics (QCD) [8] is the quantum field theory of the strong interactions, symmetric with respect gauge rotation of SU(3) group. The coupling magnitude can be estimated, from the decay probability of unstable hadrons: comparing for

example the lifetime of Σ^0 in the $\Sigma^0 \rightarrow \Lambda + \pi^0$ process ($\tau = 10^{-23}\, s$) with the electromagnetic decay $\Sigma^0 \rightarrow \Lambda + \gamma\, (\tau = 10^{-}19\, s)$, one can get the strong coupling constant α_s:

$$\frac{\alpha_s}{\alpha_e} = (\frac{10^{-19}}{10^{-23}})^{1/2} \simeq 10^2, \quad \alpha_s = \frac{g^2}{4\pi} \simeq 1 \quad , \tag{1.9}$$

where g is the value of the strong charge.

In order to explain the existence of hadrons made up of three quarks of the same flavour and quantum numbers, a new charge type has been inserted, the colour [3]. According to the Pauli's principle, the colour can assume three possible values called red, green and blue (R, G, B), making the total wavefunction of those hadrons antisymmetric. Only quarks carry colour charge, that means that only quarks are affected by the strong force. Moreover the interquark interactions are assumed to be invariant under colour interchange, meaning that the theory is described by the symmetry group SU(3). Colour symmetry is supposed to be exact, therefore the strong interaction is independent of the quark colours involved.

QCD invariance under global gauge transformations leads to the colour charge conservation. In order to guarantee the local gauge invariance, one should introduce a covariant derivative

$$\mathcal{D}_\alpha = \partial_\alpha + igt_A \cdot \mathcal{A}^A_\alpha(x) \quad , \tag{1.10}$$

where $\mathcal{A}^A_\alpha$ is the proper gauge field of the strong interaction, the gluon, and t_A is a matrix in the fundamental representation of SU(3). The field strength tensor $F^A_{\alpha\beta}$ can be expressed in function of the gluon field $\mathcal{A}^A_\alpha$:

$$F^A_{\alpha\beta} = [\partial_\alpha \mathcal{A}^A_\beta - \partial_\beta \mathcal{A}^A_\alpha - g f^{ABC} \mathcal{A}^B_\alpha \mathcal{A}^C_\beta] \quad , \tag{1.11}$$

where indices A, B, C run over the eight colour degrees of freedom of the gluon field. The third therm in Eq.(1.11) is a typical feature of a non-abelian theory: it gives rise to triplet and quartic gluon self-interactions (see Fig.1.5); f^{ABC} are the structure constants of the SU(3) colour group. The Lagrangian of the strong interaction is

$$\mathcal{L} = \sum_{flavours} \bar{q}_a (i\gamma_\mu \mathcal{D}^\mu - m)_{ab} q_b - \frac{1}{4} F^A_{\alpha\beta} F^{\alpha\beta}_A \quad . \tag{1.12}$$

It does not contain a $m^2\mathcal{A}^\alpha\mathcal{A}_\alpha$, that should represent the gluon mass but is not invariant under local gauge transformations.
In order to deeply understand the strong interaction, let me spend some words on its main features. Fist of all, the strong coupling constant is a function of energy, as the fine constant (see Fig.1.4). The quark-antiquark pairs coming out of vacuum shield the colour charge, reducing its value for increasing distance, or for increasing momentum transferred in the process. However the action of gluons is a smearing of the colour charge, which results in an opposite effect of that of quarks called **antiscreening** (Politzer, Gross and Wilezek, 1973 [9]).

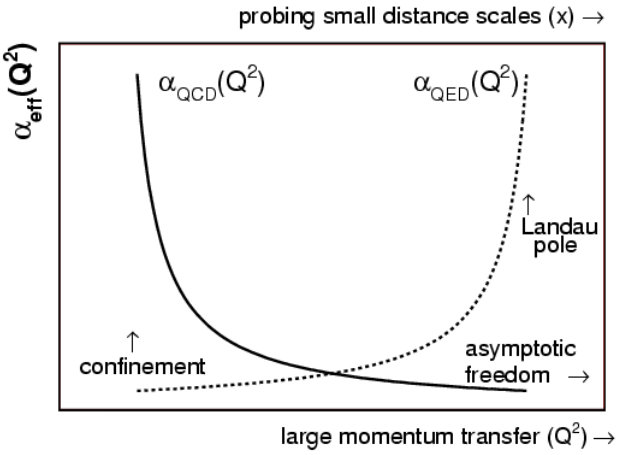


Figure 1.4: Comparison between electromagnetic and strong coupling constants.

What happens is that all around an isolate quark, all the vacuum pulsates; quark-antiquarks pairs create and then disappear, gluons appear and then fade away. This cloud of virtual particles antiscreens the central quark, making the colour charge grow with increasing distance from the quark. Nevertheless it would require an infinite energy. This divergence can be avoided if near a quark its antiquark is present, because they neutralise each other. Therefore neither quarks, nor antiquarks, nor pairs can exist alone.
Starting from the strong coupling constant, the Λ_{QCD} fundamental parameter

can be defined

$$\Lambda_{QCD}(n_f) = \mu^2 exp[-\frac{12\pi}{(33-2n_f)\alpha_s(\mu^2)}] \quad , \tag{1.13}$$

where μ^2 is the energy scale and n_f is the number of flavours which contribute to the strong process. Λ_{QCD} represents the scale at which the coupling would diverge: more qualitative, it indicates the order of magnitude at which the strong coupling constant becomes strong. This is an indication of **confinement**, the mechanism that keeps quarks and antiquaks together inside hadrons [9]. Confinement explains why the quark and the gluon degrees of freedom have never been observed as free particles, which is actually a consequence of the growth of the strong coupling constant at low energies. On the other hand, when the momentum transferred is large, namely when two quarks are really close, their interaction is feeble: this property is called **asymphtotic freedom**. Fig.1.5 illustrates the fundamental Feynman diagrams for the strong interactions.

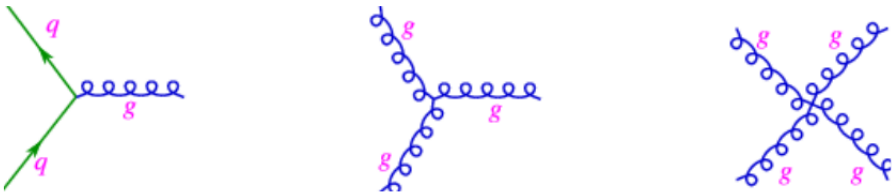


Figure 1.5: Feynman QCD diagrams: from the right the exchange of a gluon by two quarks and triplet and quartic gluon self-interactions are shown.

1.2 The top quark

The discovery of the top quark was made possible by the remarkable success of D0 and CDF experiments at Tevatron $p\bar{p}$ collider [10]-[11]. Based on data collected in 1994-1995 at a $67\,pb^{-1}$ and 44-56 pb^{-1} integral luminosity for CDF and D0 respectively, the observation of the top quark is the latest in a long

series of triumphs of the Standard Model. In fact the top quark is the last fundamental quark that has been discovered; with a charge $+2/3\,|e|$ and a weak isospin $+1/2$, it is the partner of the b-quark in the third weak doublet. Its existence was predicted many years before the experimental evidence, after the discovery of the b-quark in 1977. Its mass of $m_t = 173.2 \pm 0.9\,GeV$ [12], makes the top the heaviest of the six known quarks of the Standard Model. It is one of the fundamental parameters of the theory, because it appears in higher order loop diagrams of the electroweak theory. The large value of m_t also implies a large coupling with the Higgs boson: therefore the Yukawa coupling $y_t = m_t/v$, where $v = 246\,GeV$ is the vacuum expectation value, is of order of unity. Moreover the full decay width of the top quark is measured to be 1.33 GeV, implying a very short life time of about $\tau = 0.5 \cdot 10^{-24}\,s$, if compared to the hadronization timescale of $\tau_{had} = 3 \cdot 10^{-24} s$. Top is indeed the only quark of the SM with the property of decaying weakly ($t \to Wb$) before hadronizing and offers a unique opportunity to study the properties of a bare quark, including polarisation effects. For these reasons the top quark plays a special role in the Standard Model: an accurate knowledge of its features can be a key on the fundamental interactions at the electroweak breaking scale and beyond.

1.2.1 Top pair production

Because of its large mass, the top quark can only be observed directly in high energy experiments, where sufficiently high center-of-mass energies have been achieved. Significantly high energy has been reached at Tevatron ($\sqrt{s} = 1.8$ TeV) and LHC ($\sqrt{s} = 14$ TeV) hadron colliders.

According to the Standard Model, the dominant mechanism for the top pair production is governed by the strong interaction: since $m_t >> \Lambda_{QCD}$, the $t\bar{t}$ production can be successfully described by the perturbative QCD theory. The two main production channels at the leading order (LO) are quark-antiquark annihilation ($qq \to t\bar{t}$) and gluon-gluon fusion ($gg \to t\bar{t}$), while at next-to-leading order (NLO) there are also partonic sub-processes with $gq\,(g\bar{q})$ in the

initial state. Fig.1.6 shows the leading order diagrams for top pair production. Approximately 85% of the production cross section at the Tevatron is from $q\bar{q}$ annihilation [13], because the contribution of the valence quarks of the initial state favoured at that center-of-mass energy with respect to the gluon contribution. On the other hand at LHC about 90% of the production is from gluon-gluon fusion [14] because of the large gluon density in the proton at small x; the remainder is determined by the quark-antiquark annihilation. At both colliders the $gq\,(g\bar{q})$ processes contribute only at the percent level. At LHC the total $t\bar{t}$ cross section is $172.0^{+6.4}_{-7.5}$ pb at $\sqrt{s} = 7$ TeV and $254.8^{+8.8}_{-7.5}\,pb$ at $\sqrt{s} = 7$ TeV, which represents about 2/3 of all events containing top quarks.

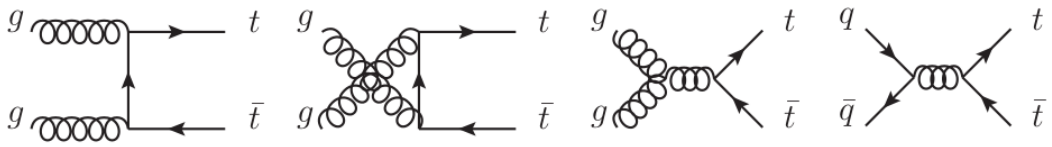


Figure 1.6: Gluon-gluon fusion and quark-antiquark annihilation Feynman diagrams for $t\bar{t}$ production at leading order QCD.

1.2.2 Single top production

The responsible for the single top production is the electroweak interaction through the vertex Wtb (about 100% of all cases since $|V_{tb}| >> |V_{td}|, |V_{ts}|$). The production cross section is predicted to be $\sigma_t = 20pb$ at $\sqrt{s} = 7$ TeV pp collisions, smaller than that for pair production [16]. The experimental signature of this process suffers from much more challenging background contamination; indeed the observation of single top quark production was only made in 2009 at D0 and CDF.

There are three different single top production processes distinguished by the virtuality of the W boson exchanged: the t-channel, the tW-channel and the s-channel, illustrated in Fig.1.8. At Tevatron, the significant channels were the t- and the s-channel, with a production cross section of about $2.2\,pb$ and

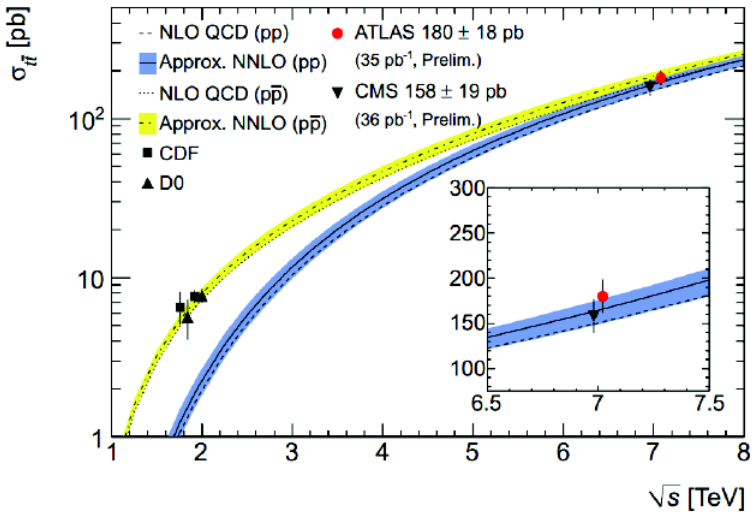


Figure 1.7: Theoretical Inclusive $t\bar{t}$ production cross section predicted for LHC and comparison between ATLAS, CMS, D0 and CDF measurements. LHC energy 4 times greater than that of the Tevatron corresponds to a top pair cross section 30 times greater [15].

$1\,pb$ respectively. Associated production with a W boson, although significant at the LHC, was negligible at Tevatron. A further analysis on the kinematics of different production processes al LHC follows (Tab.1.4). The s-channel process has the smallest cross section at LHC ($\sigma_t < 26.5(20.5)\,pb$, about five times larger than the SM expectations). In this production mode a time-like W boson is produced from two quarks belonging to an isospin doublet. Next in order of increasing cross section is the associated production of a top quark and a W boson, in which a initial state bottom quark emits W boson ($\sigma_t = 14.4^{+5.3}_{-5.1}(stat.)^{+9.7}_{-9.4}(syst.)\,pb$). The t-channel is the predominant single top production mode, accounting about 3/4 of single top quarks produced at LHC ($\sigma_t = 83 \pm 4(stat.)^{+20}_{-19}(syst.)\,pb$). In this process, a space-like W boson scatters with a b-quark, coming from the b-quark PDF of the proton or produced by gluon splitting $g \to b\bar{b}$. At proton-proton colliders the t-channel is a charge asymmetric process, due to the prevalence of u type valance quarks in the proton PDF.

All of the production modes are sensitive to the Wtb vertex in different ways: indeed non-standard couplings would indicate the presence of some new physical phenomena. In addiction the single top production allows to directly measure the CKM matrix element, without hypothesize the number of generations; deviations from the Standard Model expectations could be a signal for other generations of quarks.

Table 1.4: Expected single top quark production cross sections in different channels at a center-of-mass energy of 7 TeV and 8 TeV, given by approximate NNLO assuming $m_t = 172.5\,GeV$ [2].

Production mode	$\sigma_t[pb]$ 7 TeV	$\sigma_t[pb]$ 8 TeV
s-channel	4.6 ± 0.2	5.6 ± 0.2
t-channel	$64.6^{+2.7}_{-2.0}$	$87.8^{+3.4}_{-1.9}$
tW-channel	15.7 ± 1.1	22.4 ± 1.5

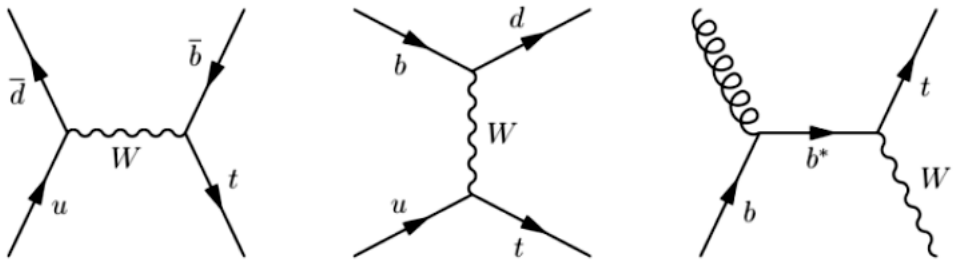


Figure 1.8: Leading-order Feynman diagrams for s-channel, t-channel and associated production with W boson.

1.2.3 Top decay

According to the the Standard Model, a vast majority of the top quarks decays into a W boson and a b-quark through the electroweak process. The width of such a decay is proportional to the square of the element in the Cabibbo-Kobayashi-Maskawa matrix (CKM). Since $|V_{tb}| >> |V_{td}|, |V_{ts}|$ (see Eq.1.1.2),

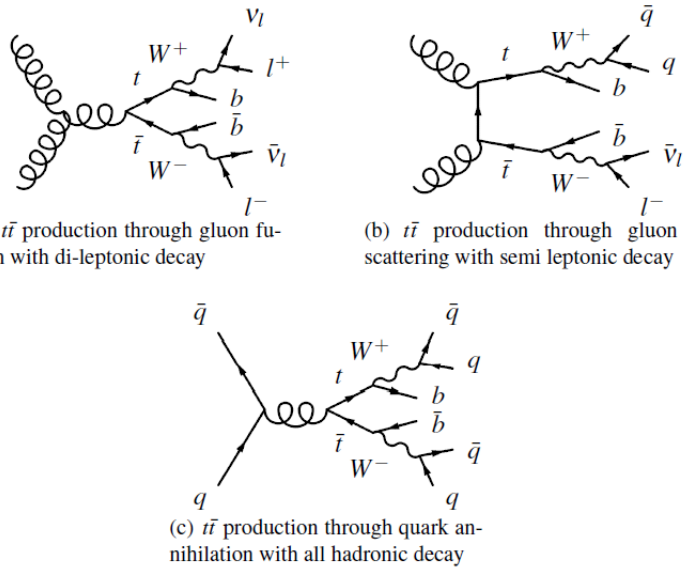
decays to the other down-type quarks, s and d, are suppressed. Neglecting the decays $t \to Wd(s)$, the total width of the top quark in the SM at NLO QCD is [17]:

$$\Gamma_t = \frac{G_F m_t^3}{8\pi\sqrt{2}}|V_{tb}|^2(1-\frac{m_W^2}{m_t^2})^2(1+2\frac{m_W^2}{m_t^2})[1-\frac{2\alpha_s}{3\pi}(\frac{2\pi^2}{3}-\frac{5}{2})] \quad , \qquad (1.14)$$

where G_F is the Fermi constant, m_W is the mass of the W boson, m_t is the mass of the top quark and α_s is the strong interaction coupling (here $\alpha_s(M_Z) = 0.118$). For a top mass of $172.5\,GeV$, the decay width of this vertex yields $\Gamma_t = 1.33\,GeV$, which corresponds to a very short lifetime $\tau_t = 1/\Gamma_t \sim 5 \cdot 10^{-25}\,s$. The fact that the top lifetime is one order of magnitude smaller than the typical formation time of hadrons means that top quark decays before hadronize. It is also an explanation of the absence of bound states containing top quarks (e.g. *toponium*). The top quark mass is even larger than the sum of the W boson (Tab.1.3) and b-quark (Tab.1.2); this implies that the W boson belonging from this decay is "on-shell". This is an important feature of $t\bar{t}$ events that makes the precision measurements of the top quark mass possible.

The top quark pairs decay modes are classified according to the decay of the W boson [16]: di-leptonic, lepton+jets and hadronic channels (see Fig.1.9 and Fig.1.10). The experimental signature varies in the different channels; the event topology and the background processes are summarized in Tab.1.5. In the di-leptonic channel both the W-bosons decay into lepton-neutrino pairs $t\bar{t} \to W^+bW^-\bar{b} \to \bar{l}\nu_l b l^{'}\bar{\nu}_{l'}\bar{b}$. The presence of two isolated high p_T leptons, a huge missing energy and at least two b-jets permits to easily identify this event, even if the two neutrinos make the reconstruction difficult. The branching ratio is small ($BR = 10.3\%$), but the backgrounds (mostly Z+jets), are also fairly small. This makes the di-leptonic topology a valid process to obtain a very clean sample of $t\bar{t}$ events. On the other hand, in the lepton+jets channel one W-boson decays into lepton and neutrino, while the other one into a quark-antiquark pair $t\bar{t} \to W^+bW^-\bar{b} \to q\bar{q}^{'}bl\bar{\nu}_l\bar{b}\,(or\,\bar{l}\nu_l bq\bar{q}^{'}\bar{b})$. Its signature is one high p_T isolated lepton, missing transverse energy and at least 4 jets: with a large branching ratio (43.5%) and not huge background (mostly W+jets), this

channel is often referred as the "golden channel". Finally, the hadronic channel is characterized by the decay of both the W bosons into quark-antiquark pairs $t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow q\bar{q}'bq''\bar{q}'''\bar{b}$. The typical signature is the presence of six jets, whose two belong to the b quark. Despite the large branching ratio (46.2%), the observation of this process is difficult by the presence of QCD multi-jets events not involving top quark.



(a) $t\bar{t}$ production through gluon fusion with di-leptonic decay

(b) $t\bar{t}$ production through gluon scattering with semi leptonic decay

(c) $t\bar{t}$ production through quark annihilation with all hadronic decay

Figure 1.9: Illustration of different top pair production and decay modes.

Table 1.5: Summary of main signatures and background of the three $t\bar{t}$ decay channels.

Channel	Event topology	Dominate background
Dileptonic	2 b-jets, 2 isolated leptons, E_T^{miss}	Z + jets
Lepton+jets	2 jets + 2 b-jets, 1 isolated leptons, E_T^{miss}	W + jets
Hadronic	4 jets + 2 b-jets, no isolated leptons, no E_T^{miss}	QCD multijets

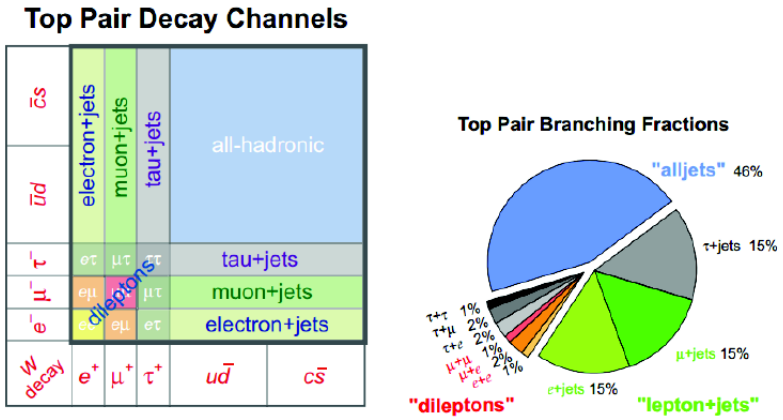


Figure 1.10: Top pair decay channels (*right*) and the corresponding branching-ratios (*left*).

1.2.4 Top quark mass

The top quark mass is a fundamental parameter of the Standard Model. A precise determination of this value induces large corrections in the theory predictions of many precision electroweak observables, including the mass of the Higgs boson. The top quark mass has been measured in the lepton+jets, dileptonic and hadronic channels by Tevatron and LHC experiments [16] (see Fig.1.11). The most precise measurement of the top quark mass was made by Tevatron [2]

$$m_t = 173.20 \pm 0.51(stat.) \pm 0.71(sist.)\, GeV/c^2 \quad ;$$

with a relative precision is 0.50%, it is a combination of Run I and Run II measurements based on data set corresponding to a luminosity of $8.7\, fb^{-1}$. Indirect constraints on m_t can be obtained from precision measurements of electroweak theory. In fact the mass of the W-boson can be expressed as a function of the QED coupling $\alpha(m_Z^2)$, the Fermi constant G_F and the electroweak mixing angle θ_W

$$m_W^2 = \frac{\pi\alpha(m_Z^2)/\sqrt{2}G_F}{sin^2\theta_W \cdot (1-\delta_r)} \quad . \tag{1.15}$$

The term δ_r contains contributions from higher order electroweak loop diagrams involving the square of the top quark mass m_t.

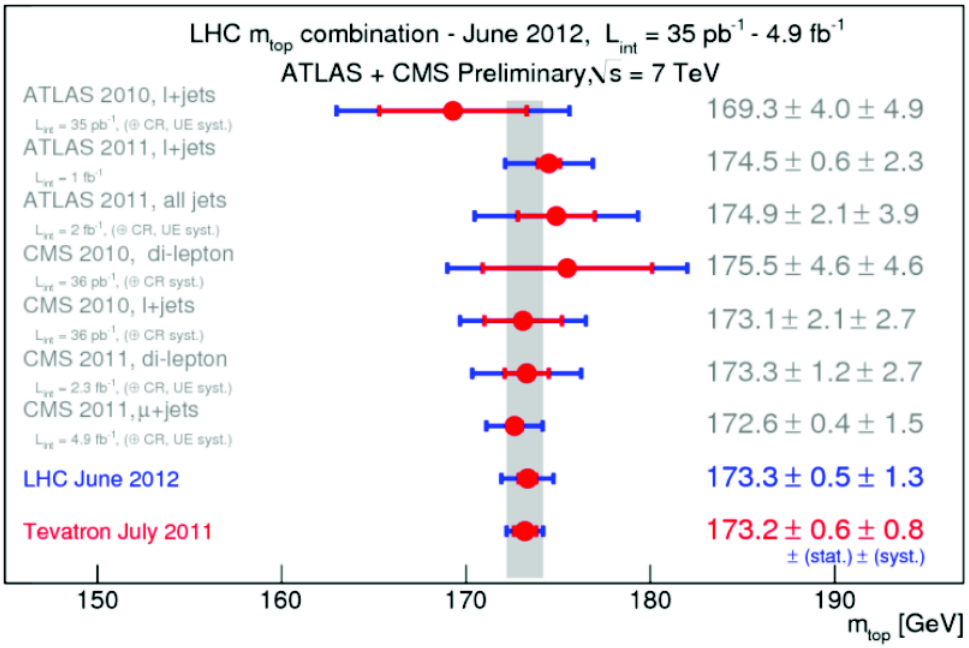


Figure 1.11: Overview of the top mass measurements from both ATLAS and CMS in the lepton+jets, dileptonic and hadronic channels [12].

1.3 Cross section measurements

In order to quantify particle production, it is important to evaluate the cross section of the process under consideration, that is a measure of the interaction probability. In high energy colliders such as LHC, protons can scatter and produce other particles; all those possible processes are described by the total *inclusive cross section*. On the other hand the *exclusive cross section* is the probability for a process to happen. The total cross section formula for a collision is given by

$$\sigma = \frac{N_{ev}}{\epsilon \mathcal{L}} \quad , \tag{1.16}$$

where N_{ev} is the number of scattered events, ϵ is the overall efficiency of the detector and $\mathcal{L}$ is the luminosity of the acquired data, meaning the luminosity obtained during the data acquisition. The instantaneous luminosity for a collider is given by

$$\mathcal{L} = f \cdot N \frac{n_1 n_2}{4\pi\sigma_x\sigma_y}, \tag{1.17}$$

where f is the collision frequency, n_1 and n_2 are the number of particles belonging to a bunch of the beam, N is the number of bunches and σ_x and σ_y are the transverse dimensions of the bunch along two orthogonal axis with respect to the beam direction.

At LHC energies, interactions happen among partons, the elementary particles (quarks and gluons) inside the proton which participate to the process almost independently. As a consequence, the available energy is the fraction of the center-of-mass energy carried by partons: $p_{q,g} = xP_p$, where x is called Bjorken variable. It varies between 0 and 1 and represents the fraction of the total momentum carried by the parton. The distribution of the momentum among all partons inside the proton is described by the Parton Distribution Functions (PDF), determined trough the combination of a large amount of experimental data on deep-inelastic scattering. The inclusive cross section of the process $pp \to t\bar{t}$ strongly depends on the center-of-mass energy of the collider and on the top mass; it can be expressed by means of the factorization theorem, which allows to convolute the parton distribution function and the partonic cross sections $\hat{s}$:

$$\sigma_{pp\to t\bar{t}}(s, m_t) = \sum_{i,j=q,\bar{q},g} \int dx_i dx_j f_i(x_i, \mu_f^2) f_j(x_j, \mu_f^2) \hat{\sigma}_{ij\to t\bar{t}}(\hat{s}, m_t, \mu_f \mu_r, \alpha_s) \quad . \tag{1.18}$$

The sum runs over all the quarks and gluons which contribute to the process, $x_{i,j}$ are the parton momentum fraction with respect to the proton momenta, $f_{i,j}(x_{i,j}, \mu_f^2)$ are the proton PDF, μ_f^2 and μ_r^2 are the factorization and renormalization scales, α_s is is the strong coupling constant and $\hat{s} \sim x_i x_j s$ is the partonic center-of-mass energy. The dependence from μ_r arises from the definition of the renormalized coupling α_s at a fixed order in a truncated perturbation

theory; on the other hand, μ_f indicates a transition between the perturbative and the non pertubative regime, thus arises from absorbing collinear initial state singularities in the PDF. The renormalization and factorization scale are usually set to the hard scale of the process: in the case of the total cross section, one usually sets $\mu_r = \mu_f = m_t$. However in the case of the differential cross sections, other scale choices are more appropriated (e.g. the transverse momentum of a jet $p_{T,jet}$ or the top pair invariant mass $M_{t\bar{t}}$).

1.3.1 $t\bar{t}$ total cross section

The top pair cross section was first measured in $p\bar{p}$ collisions at Tevatron with a center-of-mass energy of $\sqrt{s} = 1.96$ TeV. The most precise and recent measurements of D0 and CDF are

$$\begin{aligned} \sigma_{t\bar{t}}^{D0} &= 7.56^{+0.63}_{-0.56} \quad pb \\ \sigma_{t\bar{t}}^{CDF} &= 7.50 \pm 0.48 \quad pb \end{aligned}$$

in agreement with the Standard Model expected value of $\sigma_{t\bar{t}} = 7.16^{+0.20}_{-0.23} \quad pb$ at NNLO perturbation theory [16]. The strong dependence on the collision energy, explains why the theoretical production cross section at LHC is far greater than the Tevatron one. In particular at the center-of-mass energy of $\sqrt{7}$ TeV and $\sqrt{8}$ TeV, the SM predicted values are respectively

$$\begin{aligned} \sigma_{t\bar{t}}^{7TeV} &= 172.0^{+6.4}_{-7.5} \quad pb \\ \sigma_{t\bar{t}}^{8TeV} &= 254.8^{+8.8}_{-7.5} \quad pb \quad . \end{aligned}$$

ATLAS and CMS evaluated the top pair production cross section combining measurements performed in various channels. It follows a brief summary of the results in both ATLAS and CMS [2] at the center-of-mass energy of 7 TeV and 8 TeV respectively (see also Fig.1.12).

$$\begin{aligned} \sigma_{t\bar{t}}^{ATLAS(7TeV)} &= 173 \pm 3(stat.)^{+8}_{-4}(syst.) \pm 7(lumi.) \quad pb \\ \sigma_{t\bar{t}}^{CMS(7TeV)} &= 162 \pm 2(stat.) \pm 5(syst.) \pm 4(lumi.) \quad pb \\ \sigma_{t\bar{t}}^{ATLAS(8TeV)} &= 237.7 \pm 1.7(stat.) \pm 7.4(syst.) \pm 7.4(lumi.) \quad pb \\ \sigma_{t\bar{t}}^{CMS(8TeV)} &= 227 \pm 3(stat.) \pm 11(syst.) \pm 10(lumi.) \quad pb \quad . \end{aligned}$$

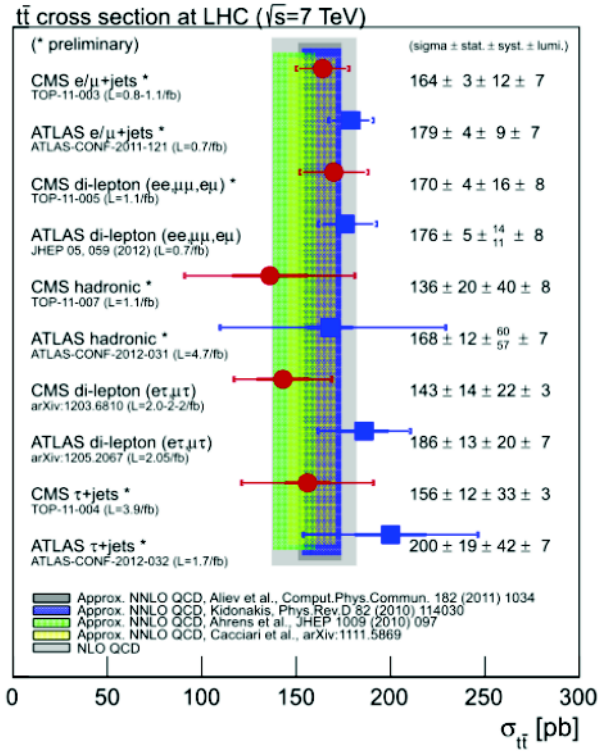


Figure 1.12: Summary of the ATLAS and CMS most precise measurements of top-antitop pair per decay mode, compared with several theory predictions at NLO and NNLO QCD [2].

These results are in agreement with NNLO Standard Model perturbation theory as it is shown in Fig1.7.

1.3.2 Differential cross section

The large abundance of top quark pair production at LHC allows not only to measure the total cross section $\sigma_{t\bar{t}}$, but also the differential cross section $d\sigma_{t\bar{t}}/dX$, where X is a relevant variable, such as the kinematic variables of the $t\bar{t}$ system. In fact cross section can be evaluated either after the extrapolation to the full phase space, as done in the case of the total cross section, or only within the kinematic range in which the decay products are measured. In particular a prominent role in the discovery of new physics have the invariant

mass distribution $M_{t\bar{t}}$ and the transverse momentum $p_{T,t\bar{t}}$, which could be significantly modified in presence of resonances decayed in top pairs.
In order to compare differential cross section measurements with theoretical predictions, it is important to clarify two different way to quote it: the particle level that considers only particles visible by the detector and easily comparable to Monte Carlo simulations, or the parton level which refers to particle before hadronisazion.
Thanks to the large available event samples, Tevatron and LHC performed first differential cross section measurements in top-antitop production [2]. Such measurements allow accurate tests on perturbative QCD, the extractions or the use of PDFs and enhance the sensitivity to possible new physics contributions. In particular both ATLAS and CMS performed several measurements with increasing statistic on different channels. Fig.1.13 shows a recent study of ATLAS collaboration in the lepton+jets channel; the result was obtained at an integrated luminosity of 2.6 fb^{-1} in the resolved channel. This was possible due to the fact that the analysis was focused the top pair with p_T only up to 1 GeV. The aim of this thesis instead is to extend the top pair production analysis to the highest energy regions, where new physics states may be found. In the reconstruction of high p_T top quarks ("boosted top"), standard techniques fail because of the partial or total overlap of its decay products, which form a huge signal (*fat-jet*) in the detector. In this context the Template Overlap Method (TOM) becomes necessary. TOM offers a new top tagging strategy based on the comparison of the fat-jet with a series of top decay states (*templates*) generated through Monte Carlo simulations. The comparison is based on a definition of an overlap function which quantifies the matching. Further details are provided in Chapter 3.

1.4 Beyond Standard Model

For the past few decades physicists have made measurements of particles and parameters of the Standard Model; provided the discovery of the Higgs boson

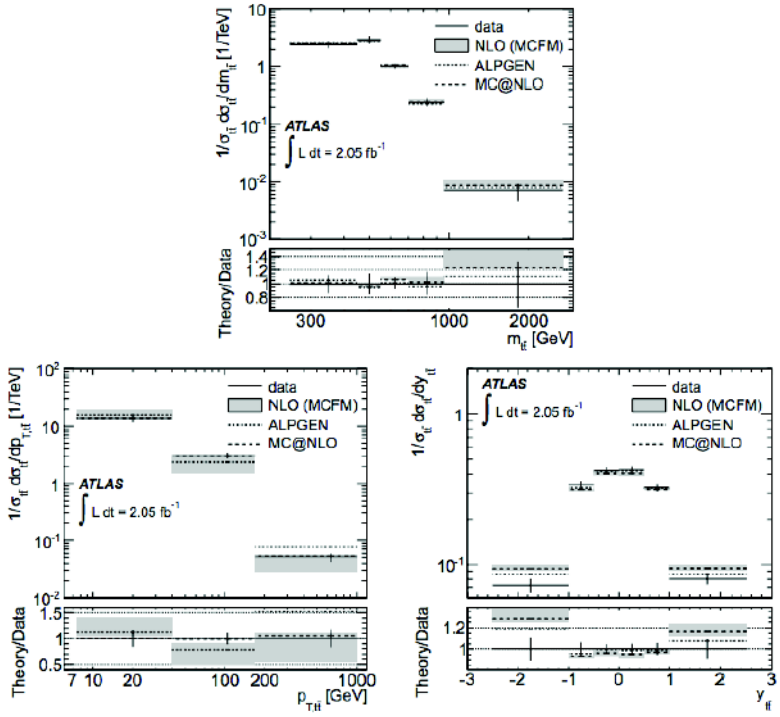


Figure 1.13: Lepton+jets channel normalised differential $t\bar{t}$ production cross section obtained by the ATLAS collaboration as a function of $M_{t\bar{t}}$, $p_{T,t\bar{t}}$ and $y_{t\bar{t}}$. The measurements is compared to the NLO prediction from MCFM [18].

and excluding the discovery of neutrino masses, no major deviations from the SM predictions have been found. Despite the remarkable success of the theory, there remain many unresolved questions that lead physicists to look beyond the Standard Model.

The structure of the SM is itself a deep mystery. The gauge groups that describe the various interactions seem to adequately describe nature, but why should nature choose these groups instead of others? Similarly, many parameters of the theory, as the coupling constants or the particle masses, are free parameters and have been evaluated from experiments. Theoretical problems can arise from the values of some of them: the hierarchy problem and the vacuum expectation value of the Higgs potential, which influences the W and Z masses, are some issues.

Another set of questions concerns major problems in physics that the SM does not address. First among these is the inability to integrate a theory of gravity in a consistent manner: this is a general problem of quantum field theories, as no renormalizable quantum field theory of gravity has been forthcoming. Moreover the expansion of the universe due to the dark energy phenomenon and the presence in the universe of the dark matter, have no explanation in the Standard Model. Furthermore the Standard Model is not sufficient to explain the observed asymmetry between matter and antimatter in nature: the CKM matrix predicts some CP violations which lead to this imbalance, but the known sources of CP violations are insufficient to account for this large discrepancy.

There are many scenarios of physics beyond the SM which involve top quarks [16]. One of them predicts new interactions with enhanced coupling to the top quark, resulting in new particles that would decay into $t\bar{t}$ pairs and may show up as resonances in the top-quark pair invariant mass distributions. New interactions imply the possible presence of new gauge bosons, as W' and Z', the heavier counterpart of the W and Z of the electroweak theory. At the moment no sign from new physics has been found in the $M_{t\bar{t}}$ distribution and new heavy resonances decaying into $t\bar{t}$ pairs have been excluded for masses up around 1.5 TeV.

Extended models add two charged Higgs bosons to the SM, which may be heavier or lighter than the top quark. In the case of a charged Higgs heavier than the top quark, an additional diagram would be added to t-channel production through the replacement of the W boson by a charged Higgs boson; although the impact on the t-channel cross section would be small, making this an unpromising discovery channel for charged Higgs. A charged Higgs lighter than the top quark would introduce additional diagrams contributing to t-channel production, but since the Higgs couples preferentially to massive particles the effect would be suppressed by the small coupling between the charged Higgs and the light quark in the initial state. The principal experimental clue for such a particle would be the introduction of a new decay mode

for the top quark.
In addiction, model independent studies can be carried out to search for non standard model interactions; they may be parametrized via effective field theories that allow the existence of flavour changing neutral current (FCNC) processes such as $cq \rightarrow tq$ or $qq \rightarrow Z \rightarrow tc$. Researches for FCNC top quark decay and production, same sign top quark production, fourth generation of quarks, charged Higgs and W' so far have turned out negative, but can already provided more stringent exclusion limits.
Finally an other problem concerning the Standard Model is the non unification of the electroweak and strong interactions: the theory that unifies these three forces is called "Grand Unification Theory" (GUT). One of the most popular extensions to the SM is Supersymmetry (SUSY) based on the Poincaré group $U(1) \otimes SU(2) \otimes SU(3)$. The basic idea behind SUSY is a symmetry between fermions and bosons, in such a way that every SM fermion should have a super partner boson and vice-verse. SUSY may provide particles that can solve the problem of the grand unification: it predicts an energy scale where all the interaction coupling constants meet. At the moment there are no experimental confirmations of supersymmetric particles; the symmetry between particles and their superpartner must be broken. This leads to the prediction of a superparticles mass much larger than the SM masses. Operating from March 2015 at the center-of-mass energy of $\sqrt{14}$ TeV, LHC Run II will offer a great possibility to explore energy regimes never reached so far, in the hope to observe new physics.

Chapter 2

LHC and ATLAS

2.1 LHC

The Large Hadron Collider (LHC) is the largest particle accelerator in the world. Located at CERN, beneath Franco-Swiss border near Geneva, it has been built in the same tunnel which hosted the former Large Electron-Positron (LEP) collider. The tunnel has a circumference of 27 km [19] and it is situated between 50 and 175 m under ground. LHC is a proton-proton collider with a design center-of-mass energy of 14 TeV at a peak luminosity of $10^{34}\,cm^{-2}s^{-1}$. It can also work as a lead ion collider, accelerating fully ionized leads atom at a center-of-mass energy of 1150 TeV ($\sim$ 2.76 TeV/nucleon) and at a luminosity of $10^{27}cm^{-2}s^{-1}$. It started be operating in 2008 and during 2010 and 2011 it reached the center-of-mass energy of $\sqrt{s} = 7TeV$; in 2012, the center-of-mass energy has been increased until 8 TeV, with a maximum luminosity of $\mathcal{L} = 4 \cdot 10^{33}\,cm^{-2}s^{-1}$ [19]. After a technical shut-down period of two years, LHC will start operating again at energies approaching its design parameters in March 2015. The high energy and luminosity will offer the opportunity for both precision measurements and high energy frontier explorations.
In the LHC tunnel two proton beams circulate in opposite directions into two separate ultra-high vacuum chambers at a pressure of 10^{-10} Torr. In order to keep the beams into circular trajectories, 1232 NbTi superconducting

dipole magnets produce a magnetic field of 8.4 T; other 392 superconducting quadrupole magnets generate a field of 6.8 T necessary to focalize the beams. The superconducting magnets are cooled with superfluid helium below 2 K. The magnet systems use a twin bore design to bend particles in both beams simultaneously, which collide every 25 ns.
Beams are not continuous, but are divided into a maximum of 2808 bunches gathered in trains of 80; moreover each bunch contains 10^{11} protons which give rise to 10^9 collisions per second, assuming a total proton-proton cross section of $10^{-25} cm^2$ at the LHC energy [20]. The most important LHC parameters are reported in Tab.2.1.

LHC protons are originated from ionised hydrogen and passed through a chain

Table 2.1: LHC technical parameters for 2012.

Maximum collision energy	8 TeV
Maximum Luminosity	$2.3 \cdot 10^{34} cm^{-2} s^{-1}$
Number of particles per bunch	$1.67 \cdot 10^{11}$
Number of fill bunches	2808
Bunch separation	25 ns
Bunch length	7.7 cm
Bunch width (Atlas)	16.7 μm
Total number of particles	$4.7 \cdot 10^{14}$
Mean current	0.584 A
Number of collision per bunch	25

of accelerators of progressively grater energy before entering in the beam-pipe (Fig.2.1). The process begin with the LINAC2, a linear accelerator which increases the proton energy to 50 MeV. The following three stages are synchrotrons: the Proton Synchrotron Booster (PSB) leads protons to 1.4 GeV, the Proton Synchrotron (PS) to 25 GeV and the Super Proton Synchrotron (SPS) to 450 GeV. In the LHC ring beams are further accelerated by 16 radiofrequency cavities with a maximum electric field of 5.5 MV/m.

Four interaction regions along the tunnel host the following experiments, as shown in Fig.2.1:

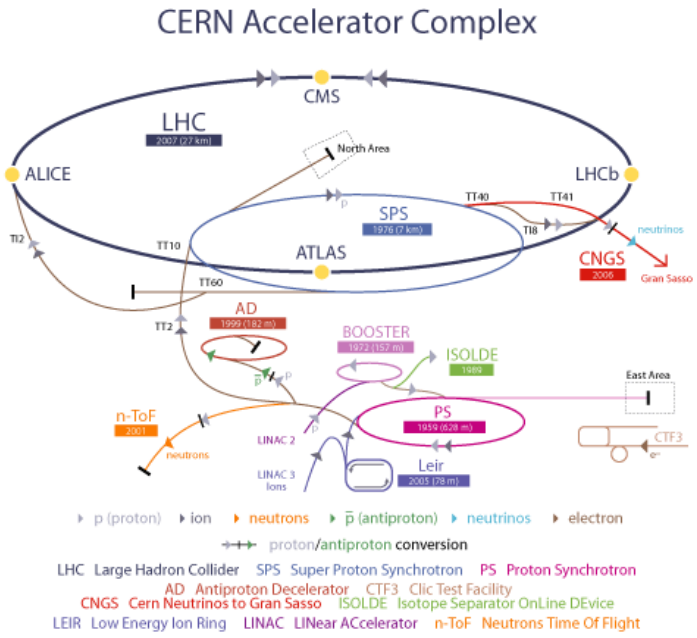


Figure 2.1: Schematic view of CERN accelerators.

- **ATLAS (A Toroidal LHC ApparatuS)** is a multi-purpose experiment which works at high luminosity ($\mathcal{L} = 10^{34} cm^{-2} s^{-1}$) to explore the Higgs boson and all the heavy particles, that may permit to solve the mass origin and the extradimension problems.

- **CMS (Compact Muon Solenoid)** is a multi-purpose experiment designed to work up to the same high luminosity of ATLAS, but implemented with different and complementary technologies.

- **LHCb** performs accurate measurements of the b-quark physics (e.g. CP violation of B mesons). It works at a luminosity lower than the one designed for the two previous experiments ($\mathcal{L} = 10^{32}\, cm^{-2} s^{-1}$), in order to better reconstruct the decay vertices of B-mesons, made difficult when there is more than one interaction per bunch crossing.

- **ALICE (A Large Ion Collider Experiment)** is dedicated to the study of heavy ion collisions, in order to reproduce the matter state

(called *Quark-Gluon Plasma*) for the first 30 μs of its life, that means early after the Big Bang. Due to the high nucleus-nucleus cross section, ALICE works up to luminosities of $\mathcal{L} = 10^{27} cm^{-2} s^{-1}$.

2.2 ATLAS

ATLAS is multi-purpose particle detector installed 100 m underground in the interaction Point 1, along the LHC tunnel. With a total length of 42 m, a diameter of 22 m and a weight of 7000 t, it is the most extended of the LHC experiments [21]. The detector is organized in a central barrel and two end-caps that close both ends. It has a cylindrical symmetry around the beam pipe: all of its subdetectors are arranged in concentric layers around the interaction point, each optimized to the detection of a specific type of particles. ATLAS is composed by six main subsystems:

- the **Magnetic System**, it is necessary in order to bend the trajectory of charged particles and to measure their momentum;
- the **Inner Detector**, that provides precise measurements of the trajectory of charged particles and reconstructs the interaction vertexes;
- the **Electromagnetic** and the **Hadronic Calorimeter**, optimized for the measurement of the photon and electron energy and jets of hadrons, respectively.
- the **Muon Spectometer**, specialized apparatus which identifies muons and measures their energy. Muons are indeed very penetrating particles which cross all the previous subdetectors without losing their energy, but leaving only an ionization signal;
- the forward detectors, among which **LUCID**, finalized to the luminosity measurement.

An overview of the ATLAS detector is shown in Fig.2.2. The standard ATLAS

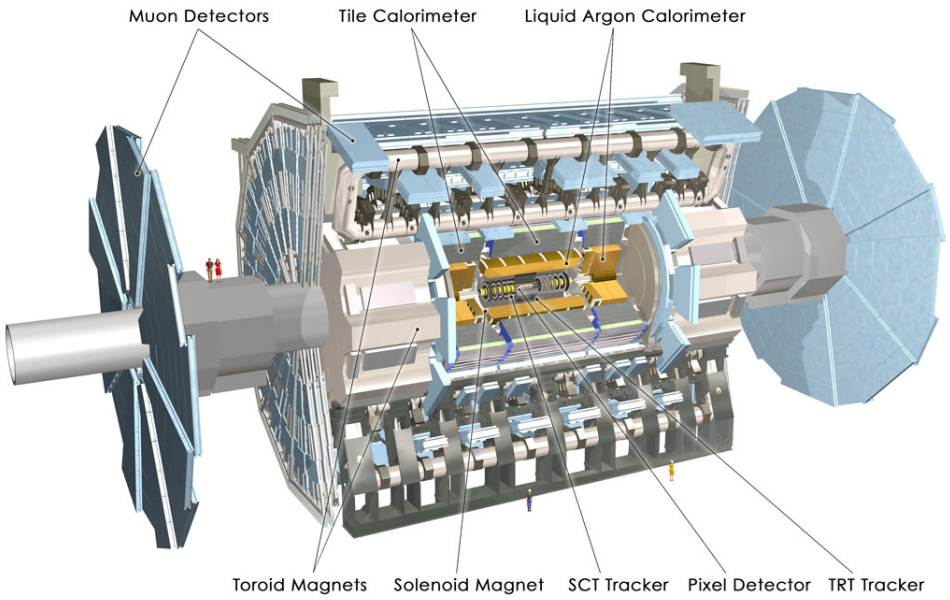


Figure 2.2: The ATLAS detector.

coordinate system is defined with respect to the interaction point, around which the detector is forward-backward symmetric [21]. The beam direction identifies the z-axis and the x-y transverse plane. According to the standard convention, the positive x-axis is defined as pointing to the center of the LHC ring, while the positive y-axis is defined as pointing upwards from the beam. Keeping in mind the cylindrical symmetry of ATLAS, the spherical (r, ϕ, θ) coordinate system is frequently employed, where r is the radius from the central axis, ϕ is the azimuthal angle measured around the beam direction and θ is the polar angle from the beam axis. As a function of θ, the *pseudorapidity* describes the angular position relative to the beam axis

$$\eta = -ln\, tan(\theta/2) \quad ; \tag{2.1}$$

in the non zero approximation for highly relativistic particles, this definition closely approximates the definition of *rapidity*

$$y = \frac{1}{2} ln(\frac{E + p_z}{E - p_z}) \quad . \tag{2.2}$$

The transverse momentum p_T, the transverse energy E_T and the missing transverse energy E_T^{miss} are defined in the x-y plane. The distance between two objects in the pseudorapidity-azimuthal plane (η, ϕ) is

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad . \tag{2.3}$$

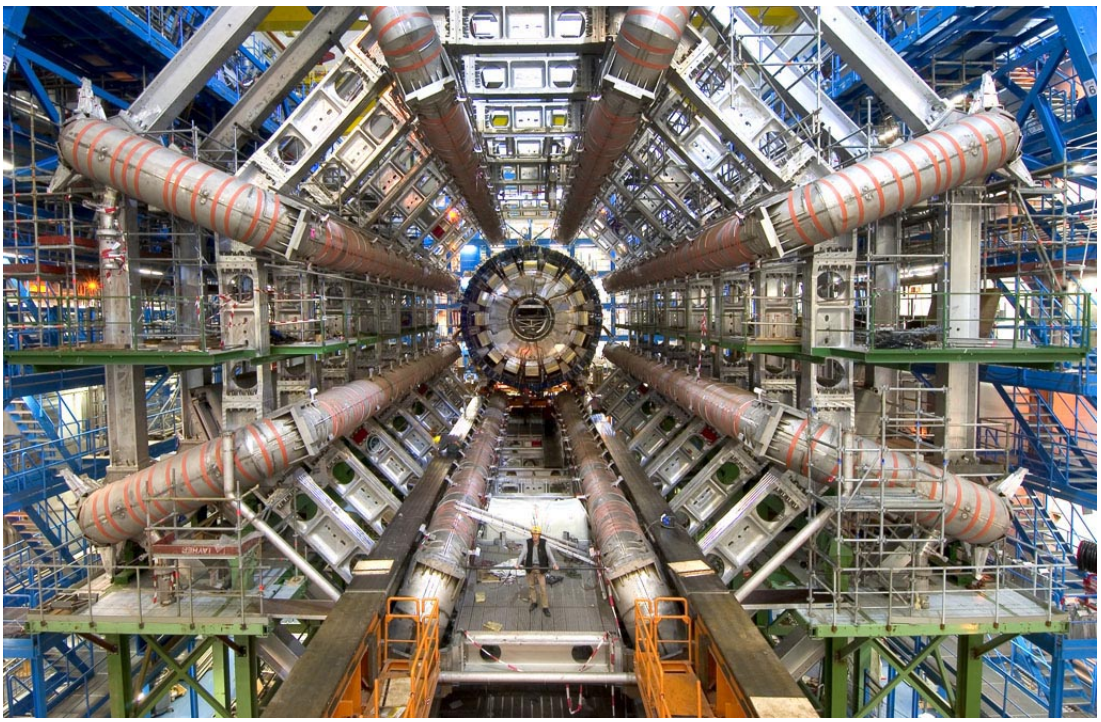


Figure 2.3: Section of ATLAS detector.

2.2.1 The magnets system

The magnets system is finalized to the evaluation of the charge q and the momentum $p\,[GeV/c]$ of particles, through the measurement of the curvature radius $\rho\,[m]$ of their trajectories, when they cross a region with a magnetic field $B[T]$:

$$p = 0.3\,\rho\,q\,B \quad . \tag{2.4}$$

ATLAS is characterized by three different superconductive magnetic systems:

- **Central Solenoid (CS)**: installed around the Inner Detector, it is a superconducting solenoid of a niobium-titanium (NbTi) alloy. With a radius of 1.2 m and a length of 5.3 m, it carries a 8 kA current to provide a magnetic field of 2 T. The coil is kept at a temperature of 4.5 K by a flux of liquid helium. CS is represented in blue in Fig.2.4.

- **Barrel Toroid (BT)**: it is composed by 8 rectangular superconducting coils arranged in a cylindrical configuration and kept at a temperature of 4.5 K. The total length is 25 m, the outer diameter is 20.1 m and the inner diameter is 9.4 m. Installed just outside the calorimeters, it bends particles with $\eta \leq 1$ and provides a magnetic field of 1.5 T.

- **End-Cap Toroid (ECT)**: it is composed by 8 rectangular coils arranged in a single cylindrical vessel. The total length is 5 m, the outer diameter is 10.7 m and the inner diameter is 1.65 m. The vessel is mounted at the end of ATLAS in order to close the magnetic field lines produced by the Barrel Toroid. The 2 T magnetic field is orthogonal to the beam axis and bends particles emitted at small polar angle ($1.4 < \eta < 2.7$). The entire Toroid Magnets (in red in Fig.2.4) reach a total magnetic field of 4 T mostly orthogonal to muon trajectories.

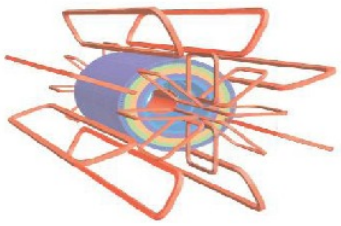


Figure 2.4: The Central Solenoid (*blue*), the Barrel Toroid and the End-Cap Toroids (*red*) of the magnetic system.

2.2.2 Inner Detector

The Inner Detector (ID) is the part of the apparatus closest to the beam pipe; it is placed in a cavity delimited by calorimeter cryostats around the beam pipe. It reconstructs the charged particle tracks by measuring the ionization energy they produce as they move through the detector medium. It has an inner radius of 45 mm, an outer radius of 115 cm and a length of 6.2 m. The ID is immersed in a 2 T magnetic field parallel to the beam axis. Its structure is composed by two silicon revelation systems, the Pixel Detector and the SemiConductor Tracker (SCT), and by a Transition Radiation Tracker (TRT), as presented in Fig.2.5 and Fig.2.6.
The two major goals of track reconstruction are the measurement of momentum and the reconstruction of interaction vertexes. The momentum can be inferred by measuring the curvature of the tracks produced as charged particles move through the field; on the other hand, interaction vertexes are reconstructed by extrapolating tracks recorded in the ID to their origin point in the beam pipe. This process is essential to reject pile-up and to tag jets produced by the decay of heavy particles. The Inner detector provides a transverse momentum resolution of about 4% for 100 GeV muons and a transverse impact parameter resolution of $\sim 35\,\mu m$ for $p_T = 100$ GeV and $\sim 10\,\mu m$ for $p_T = 5$ GeV pions. At designed luminosity, about 1000 charged particles are expected every 25 ns within the coverage of the tracking detectors, which extends out to $|\eta| < 2.5$ [20].

Pixel Detector

It is the innermost tracking detector made up of silicon pixel with high granularity. Therefore it is designed to measure the particle impact parameters, the production vertexes and the decay of short lived particles, as B mesons or τ leptons. The pixel detector consists of three concentric layers around the beam axis at average radii of 5 cm, 9 cm and 12 cm and five rings perpendicular to the beam axis on each side of the interaction point (11 cm inner radius and 20 cm outer radius) [22]. It is composed by 1744 modules, each consisting

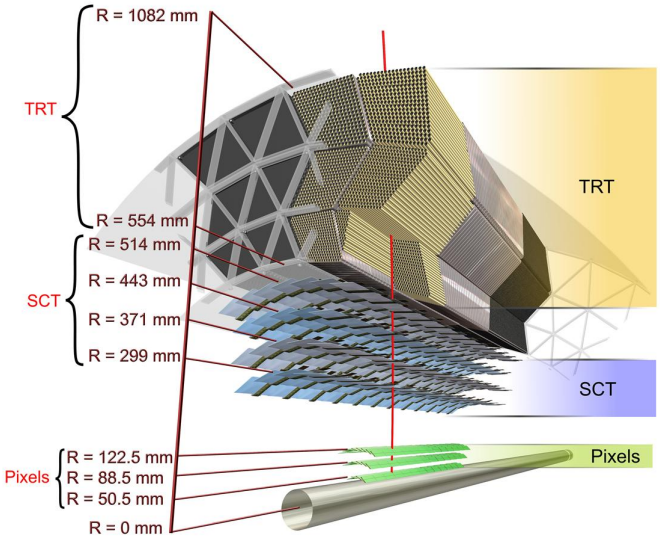


Figure 2.5: An illustration of the ATLAS Inner Detector. It highlights the major features of the design, showing the arrangement of modules in the barrel and end-caps of the Pixel Detector, the SCT and the TRT.

of a 250 μm layer of silicon pixels, adding up to a total of 80 million pixels. The system occupies a total area of 1.73 m^2 and has an intrinsic accuracy in the position determination of 10 μm in the $r - \phi$ plane and 115 μm in the z direction for the barrel detector, while an accuracy of 10 μm in $z - \phi$ and 115 μm in r for the end-caps.

SemiConductor Tracker (SCT)

Placed in the intermediate radial range of the ID, the SCT provides precise reconstruction of tracks and measurements of momentum, impact parameter and vertex position [23]. It employs the same semiconductor technology as the Pixel Detector, with the difference in the use of silicon microstrips instead of pixels systems. The SCT consists of four double layers in the barrel region and nine end-cap disks per side that cover up to $|\eta| < 2.5$: each of the 4088 detector

modules incorporates two layers with the strips rotated of 40 mrad one with respect to the other; information from both layers allows the reconstruction of a precise hit location. It occupies an area of 63 m^2 and has a spatial resolution of 17 μm along the $r - \phi$ plane and 580 μm in the z direction; moreover SCT provides a transverse momentum resolution of about 4% for 100 GeV pions. Its high granularity is important for the pattern recognition.

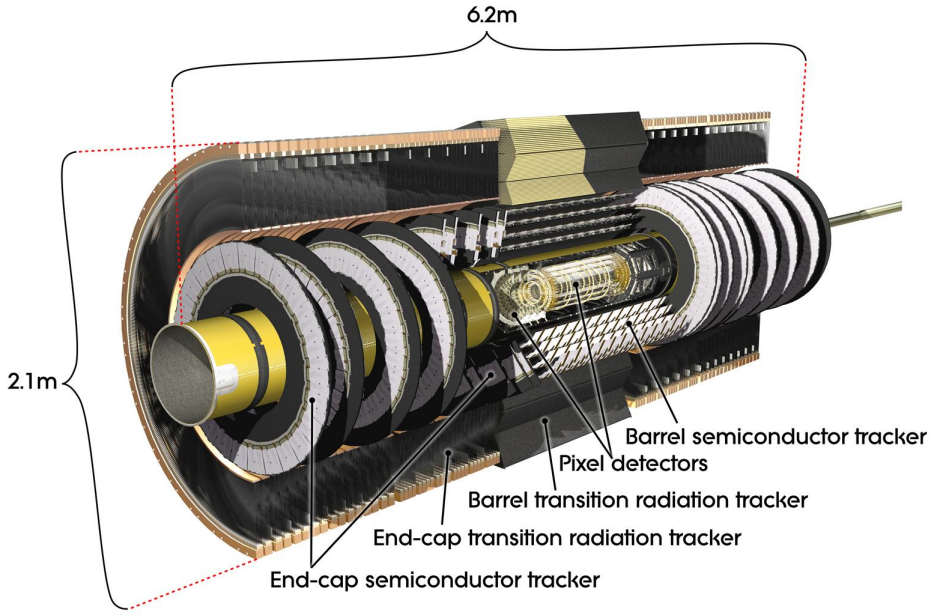


Figure 2.6: cross sectional view of the Inner Detector.

Transition Radiation Tracker (TRT)

It is the outer component of the Inner Detector and participates to the track reconstruction and to the momentum measurement in the $|\eta| < 2$ region. It is equipped with continuous tracking elements, based on the use of straw detectors. Each straw is 4 mm in diameter for a maximum straw length of 144 cm in the barrel [24]. A gold-plated tungsten wire in the middle of each tube collects the signal. Filled with a ionizing gas mixture of 70% Xe, 20% CO_2 and 10% CF_4, the tubes are arranged in 36 layers interleaved with transition

radiation material, in order to stimulate transition radiation from ultrarelativistic particles. The intensity of the transition radiation is proportional to the Lorentz factor of the traversing particle. It measures only the $r - \phi$ coordinate with a resolution of 130 μm.

2.2.3 Calorimetric System

The ATLAS calorimetric system surrounds the Inner Detector and provides a measure of the energy of the collision products. Covering a range $|\eta| < 4.9$, these detectors are known as sampling calorimeters, as they consist of alternating layers of dense absorber material (*passive*) and instrumented detector material (*active*). When they cross the calorimeter, particles interact with the dense medium, producing a cascade (or shower) of decay products, whose energy is measured in the instrumented region.
ATLAS calorimeters are of two basic types, differentiated by the kind of particles they are optimized to detect [25]. The electromagnetic Liquid-Argon (LAr) calorimeter is designed to identify and measure showers produced by charged particles, while the hadronic calorimeter is idealized to evaluate the larger showers produced by strongly interacting hadrons. A pictorial view of the whole system is shown in Fig.2.7.
In order to obtain informations about particles energies and directions, calorimeters must provide good containment for electromagnetic and hadronic showers and limit punch-through into the muon system. Applying energy and momentum conservation to the whole event, it is also possible to detect the eventual presence of a neutrino reconstructing the missing transverse energy (E_T^{miss}), that is the difference between the initial and final state energy of a weak process in the transverse plane of the detector.
Two important features describe electromagnetic and hadronic calorimeters: the *radiation length* and the *interaction length* respectively. The former is the mean distance over which a high energy electron will lose $1/e$ of its energy; the latter refers to the mean path length required in a particular material to reduce the number of relativistic charged particles by a factor $1/e$. The energy

resolution is determined by the relation

$$\frac{\Delta E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \quad , \tag{2.5}$$

where a is a sampling dependent parameter, b is related to the electronic noise of the readout technology and c is a constant term reflecting the calorimeter inhomogeneity. In Tab.2.2 electromagnetic and hadronic calorimeters resolution are presented.

Table 2.2: Nominal detector performances for the ATLAS calorimetric system [21].

Detector component	Δ_E/E
LAr	$10\%/\sqrt{E} \oplus 0.7\%$
HTC	$50\%/\sqrt{E} \oplus 3\%$
HEC	$50\%/\sqrt{E} \oplus 10\%$
FCAL	$100\%/\sqrt{E} \oplus 3\%$

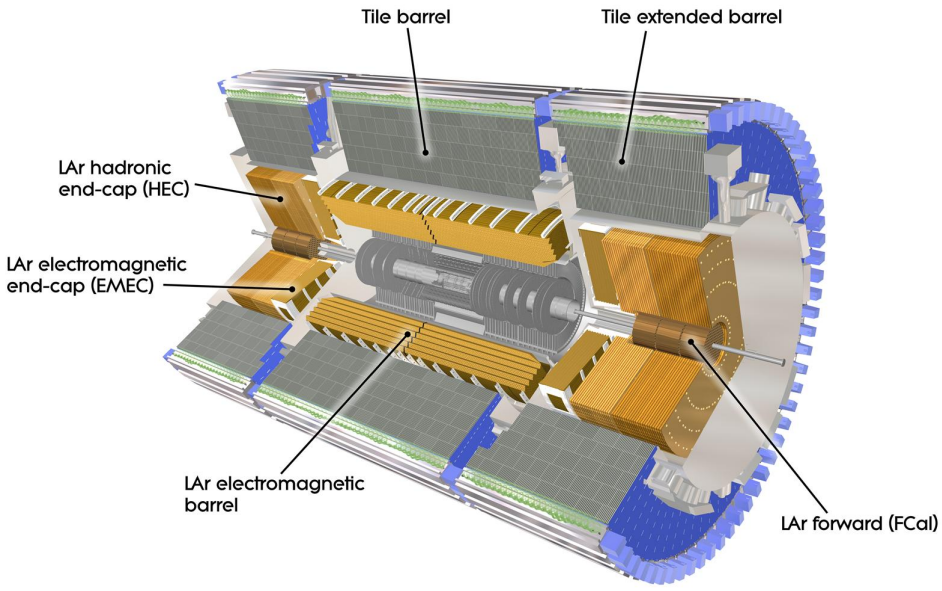


Figure 2.7: The ATLAS calorimetric system.

LAr electromagnetic calorimeter

The EM calorimeter is especially designed to absorb and measure the energy of photon and electrons. It is divided into a barrel part, covering the region $|\eta| < 1.475$, and two end-cap components. The barrel detector is composed by two half-barrels, each 3.2 m long with inner and outer radii of 2.8 and 4 m respectively. The two end-caps cover the range $1.375 < |\eta| < 3.2$; each is 63 cm thick with inner and outer radii of 0.33 and 2.098 m. The total thickness of the EM calorimeter is >24 radiation lengths in the barrel and >26 in the end caps.

The detector structure is composed by liquid argon (LAr), with accordion-shaped kapton electrodes, as active material and lead absorber plates, as passive medium. When a particle traverses the liquid argon gap, it ionizes the medium and the signal is then collected on readout electronics. LAr was chosen as active material due to its linear response to increasing energy deposition and good performance without degradation after long periods in a high-radiation environment. In the region $|\eta| < 1.8$, the calorimeter is complemented by pre-samples, an instrumented argon layer which provides a measurement of the energy losses in front of the electromagnetic calorimeter. The energy resolution has been measured to be less than 1.5%, on analysis performed with $p_T = 100$ GeV photons.

Hadronic calorimeters

The hadronic calorimeter provides measurements of the fully showers produced by hadrons interacting via the strong nuclear force. It consists in three different detectors due to the radiation level dependence on pseudorapidity. They consist of a Hadronic Tile Calorimeter (HTC) in the barrel region ($|\eta| < 1.7$), a liquid argon sampling calorimeter in the end-caps (Hadronic End-Caps Calorimeter, HEC) for $1.5< |\eta| < 3.2$ coverage, and a Forward Calorimeter (FCAL), very close to the beam pipe, that covers the region of $|\eta| < 5$ [25]. HTC is a sampling calorimeter composed of plastic scintillator tiles interspersed with steel absorber ones, arranged with the long axis orthog-

onal to the beam line into wedge shaped modules. HTC extends from 2.28 to 4.25 m. HEC consist of two wheels of 2.03 outer radius, composed by 25 mm and 50 mm copper plates as absorber material. The 8.5 mm gaps between the various plates are filled with liquid argon as active material. FCAL is located at a distance of 4.7 m from the interaction point and it is made of three sections, a first copper section and two tungsten ones.
The thickness of the calorimeter is 10 interaction lengths, tuned in order to contain all the hadronic shower, to minimize the punch-through into the muon system and to provide a good resolution for high energy jets. The energy resolution has been measured to be $\Delta E/E \sim 65\%/\sqrt{E} \oplus 2.5\% \oplus 5\%/E$.

2.2.4 Muon Spectrometer

The Muon Spectrometer (MS) is the outermost ATLAS sub-detector. With a length of 46 m and an outer diameter of 22 m, it is designed to detect muons and to measure their momentum from the bending caused by the toroidal magnets [26]. Muons are charged particles that interact in the tracking and in the calorimetry systems, but lose energy slowly and thus easily penetrate the calorimeters to decay well outside the detector. High p_T muons or isolated ones are common in interesting Standard Model events and can thus provide an important signature for the trigger system.
The muon system is instrumented with separate trigger and high-precision tracking chambers for excellent momentum resolution [27], as shown in Fig.2.8. Triggering in the central region is provided by the Resistive Plate Chambers (RPC), while in the forward region by the Thin Gap Chambers (TGC), covering the range $|\eta| < 2.4$. In order to tag on muons, a very rapid availability of tracking information is needed, within tens of nanoseconds after the passing of a muon and a timing signal with a narrow spread of 15-25 ns. The RPCs are positioned in three layers around the MDT. They are made of two resistive plates of phenolic-melaminic plastic laminate with an electric field of about 4.9 kV/mm. Ionizing tracks lead to avalanches as electrons accelerate across the potential difference toward the anode; the drift motion of these electrons

indices a signal on strips mounted on the faces of the resistive plates. The RPC signal width of about 5 ns is well within the performance envelope required for bunch crossing identification. The TGCs are arranged in multiple disks orthogonal to the beam pipe; they are thin multi-wire proportional counters composed of an array of wires between two cathodes made of metallic strips perpendicular to the wires. The small cathode-anode spacing allows a very short drift time of the electrons and thus an excellent response in time, less than 20 ns.

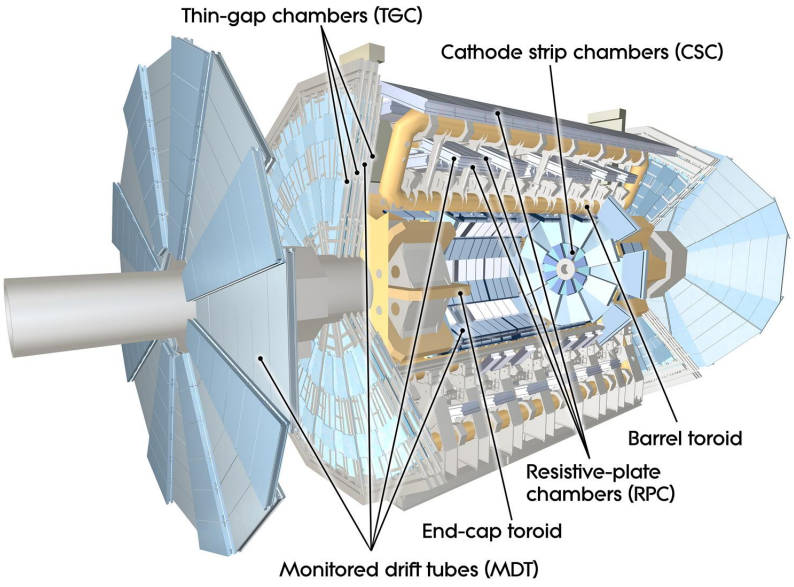


Figure 2.8: ATLAS Muon Spectrometer layout.

Two distinct technologies are also used for the precision tracking: Monitored Draft Tubes (MDT) in the central region and Cathode Strip Chambers (CSC) in the forward, measuring muon momentum in the pseudorapidity range $|\eta| < 2.7$. The MDT are gas-filled ionization detectors whose basic element is an aluminium drift tube about 30 mm in diameter with a tungsten-rhenium wire running down the centre. The tube is filled with an Ar/CO_2 gas mixture. A single MTD tube resolution is of 80 μm, while the all MDT system reach a

resolution of 35 μm along the ϕ direction. In the forward region where the flux of particles is the greatest, MDTs are replaced by the CSC detectors (counting rate of 1000 Hz/cm^2). The latter are multi-ware proportional chambers filled with an Ar/CO_2 mixture with cathodes segmented into strips; in particular one cathode is segmented parallel to the wires and the other in the perpendicular direction, in order to allow measurements in two coordinates. Therefore the spatial resolution is of 40 μm in the ϕ direction and of 5 mm in the η direction. The Muon System performance on the transverse momentum measurements has been evaluated using muons of both p_T = 10 GeV and p_T = 100 GeV, founding a resolution value of 3% and 12% respectively.

2.2.5 The Trigger and Acquisition System

The Trigger System is essential for reducing the enormous data flow and to select potentially interesting events from the large rate of inelastic collisions. These events are obtained as digital informations by the Data AcQuisition system (DAQ). The Trigger and AcQuisition system (TDAQ) works on 10^9 interactions for second at the LHC frequency of 40 MHz; although the rate must be reduced to about 100 Hz for permanent storage of collision records. As illustrated in Fig.2.9, the trigger system is based on a three levels of online selection: LVL1, LVL2 and LVL3, also called Event Filter [28].

The LVL1 trigger algorithms discard the major part of non interesting events, using reduced granularity data from only the TGC, RPC and calorimeters. This information is employed to make a decision to reject an event or pass it to the next level within 2.5 μs. During the LVL1 processing, the data from all subdetectors are held in pipeline memories in the front-end electronics until the decision is taken. In addiction, the LVL1 trigger identifies regions of interest (ROI) around parts of the detector containing the features that satisfy trigger requirements. The maximum rate of events passing the first level must be kept below 75 kHz. Events selected by LVL1 are held in Read Out-Buffers (ROB), until LVL2 trigger takes the decision to discard the event or accept it. LVL2 is a software-based system; it makes use of the ROI created by LVL1 to

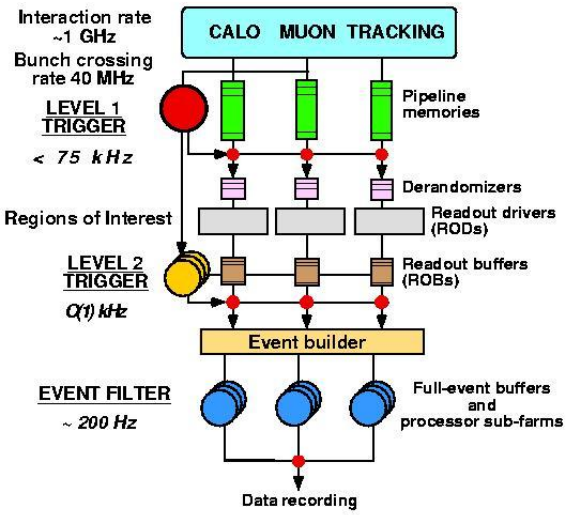


Figure 2.9: Schematic diagram of ATLAS trigger system.

read the minimum information from the ROBs necessary for the second level decision. The LVL2 trigger reduces the accepted rate from 100 kHz to 1 kHz, with a latency range between 1 ms and 10 ms depending on the event. After an event is accepted, the full data are sent to the Event Filter processor via the Event Builder. Event Filter use offline object reconstruction algorithms and the most recent calibration, to make a final decision to write or to reject an event, with a latency of 1 s. The Event Filter must achieve a data storage of 10-100 MB/s by reducing both the event rate and the event size.

ATLAS uses an advanced computing system to handle and distribute the large amount of data produced in the LHC collisions. The computing system is based on GRID technology, a way of sharing computing power and data storage capacity over the Internet, thanks to optical fibres capable to transfer bytes at 1 Gbyte/s. The GRID system is organized in three different levels: the first stage (TIER0) is located at CERN; 11 TIER1 are connected to it, which, in turn, are connected to 150 TIER2. Data are recorded according to three different formats: Raw Data Object (RDO), Event Summary Data (ESD) and Analysis Object Data (AOD).

2.2.6 LUCID

The luminosity accurate measurement is essential in the determination of the cross section of the physical process of interest. In ATLAS there are four detectors which can perform this measurement: LUCID, BMC, FCAL and HTC; although here only a brief description of LUCID is given.
The Luminosity Cherenkov Integrated Detector (LUCID) is the dedicated luminosity monitor of the ATLAS experiment. It measures inelastic proton-proton interactions in the forward direction, providing bunch-by-bunch luminosity with a precision of 2-4% and beam background information. It is composed by two identical modules placed around the beam-pipe on both forward ends of ATLAS, 17 m far from the interaction point, covering a range $5.6 < |\eta| < 6.0$. Each module is formed by 16 photomultipliers (PMTs) and 4 quartz fibers bundles read by PMTs themselves. PMTs detect Cherenkov light emitted by charge particles passing through the quartz window; the light is carried on 1.5 m long fibres to the front-end electronics. In every bunch crossing, PMT signals are discriminated by the acquisition system and converted in digital "hits" if over threshold (equivalent on average to 15 photoelectrons); the "event" is then defined from specific hits configurations. To measure the bunch-per-bunch luminosity, LUCID exploits event-algorithms or hit-algorithms [29]; with a view to II Run, another method has been developed based on the measurement of the charge gathered by the PMTs. The typical latency of this system is a few ns (maximum 10-15 ns), in order to ensure an accurate reading capability of each bunch crossing (separated from 25 ns).
Starting from March 2015, when the center-of-mass energy will reach the value of $\sqrt{s} = 14$ TeV, the average number of bunch crossings, the number of particles passing through the PMTs and thus the anodic current will increase. In order to be efficient at the new experimental conditions, LUCID will be supported by a new type of PMTs and a new readout electronics coupled with them (see Fig.2.10). As CERN Summer Student I personally calibrated and equalized the new PMTs before they were assembled in the detector.

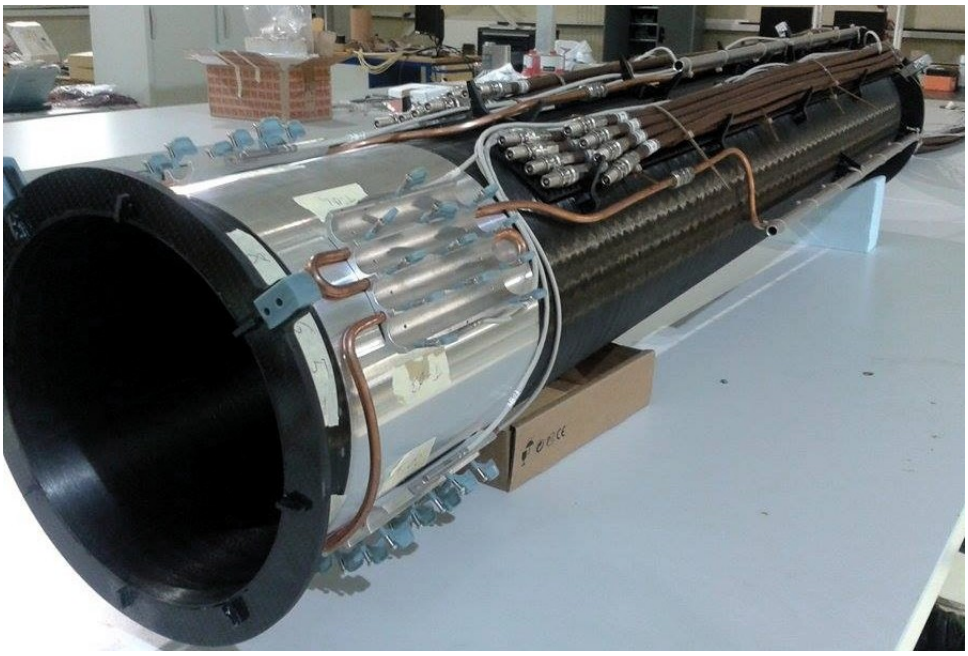


Figure 2.10: LUCID detector under construction in view of the II Run of LHC.

Chapter 3

Top Reconstruction and Selection

The analysis presented in this thesis has been performed on data collected by the ATLAS detector at $\sqrt{s}$= 8 TeV and is focused on a new method of reconstruction for the high momentum top quark. Top quark mainly decays weakly emitting a W boson and a b quark ($t \rightarrow Wb$), where the W boson can decay both in a couple of leptons or quarks. This new algorithm for the top reconstruction will be included inside the measurements of the $t\bar{t}$ differential cross section production in the semileptonic (lepton+jets) channel. In the lepton+jets channel one of the two W bosons decays in leptons and the other one in quarks, characterizing this channel by the presence of four jets, an isolated high p_T lepton and missing transverse energy indicating the presence of a neutrino. According to the energy of the emitted quarks, the events can be classified in two categories: if quarks have low p_T, the hadronic process can be identified by three well separated jets ("resolved"), otherwise if quarks have high p_T, the same decay is characterized by three partially or totally overlapped jets ("boosted").

3.1 Top Reconstruction

The standard analysis reconstructs the hadronic top as three separated jets (*resolved* [30]); this procedure has high efficiency till the top momentum is

not so high (about 300 GeV of p_T, see Sec.3.2), while it decreases significantly when the top momentum increases (in this analysis p_T >300 GeV) and the three jets result partially or totally overlapped (*boosted*). A description of the reconstruction of the main objects follows.

3.1.1 Jets

The majority of high energy physics processes involve quarks and gluons, that after the production fragment and hadronise, leading a collimated flux of energetic hadrons, called **jet**. Jets can be easily recognised as a flux of particles which reaches the hadronic calorimeter and by measuring its energy and momentum, it is possible to extract information of the "original" parton.
Due to the fact that the concept of partons, and thus of jets, is referred to distribution functions (PDF), a jet definition becomes necessary, meaning a set of rules how to group particles and how to assign a momentum to the jet. Identifying the jet as the product of a single parton could be ambiguous; in fact a high momentum particle could decay to various hadrons that generate jets completely or partially overlapped. Produced a small angle, they are indeed detected as a single big jet, although composed by more partons belonging to the decay of an energetic particle. In this contest, proper jet definition for specific analysis task becomes essential.
According to the "Snowmass accord" set out in 1990, a jet definition should meet several important properties [31]: simple to implement in experimental analysis and theoretical calculations, defined at any order perturbation theory and yielding finite cross section relatively insensitive to hadronisation. Moreover jet algorithms should be infrared and collinear safe, that means that modifying the event with the addition of a soft emission or a collinear splitting, the set of hard jets that are found in the event should remain unchanged.
Algorithms for the jet reconstruction can be divided into two broad categories: cone and sequential recombination algorithms. The former takes as initial direction the trajectory of a particle i and sum all the momenta of the particles j within a circle of a certain radius. In this thesis, only sequential recombination

algorithms will be considered, whose detailed explanation follows.

Sequential recombination algorithms

Sequential recombination algorithms repeatedly recombine the closest pair of particles according to some distance definition. The first simple one was introduced by the JADE collaboration [31] in the middle of the 1980's; for each pair of particles *i, j*, it evaluates the distance

$$y_{ij} = \frac{2E_iE_j(1 - cos\theta_{ij})}{Q^2} \quad , \tag{3.1}$$

where Q is the total visible energy of the event, $E_i\,(E_j)$ is the total energy of particle $i\,(j)$ and θ_{ij} is the angle between particles i and j. For massless particles this distance represents the normalized square invariant mass of the pair. The algorithm finds the minimum y_{min} among all the y_{ij} couples: if y_{min} is below a fixed threshold y_{cut}, it recombines i and j into a single new particle and repeats the steps from the beginning; otherwise, it declares all the remaining particles to be jets and terminates the interaction. This method is infrared and collinear safe, because any soft and collinear particle will be recombined at the beginning of the clustering. The presence of E_iE_j in this distance definition means that two very soft particles moving in opposite directions often get recombined into a single particle in the early stage of the clustering. This goes contrary to the intuitive idea that only particles emitted in a restricted angular region enter in the jet. The most common sequential recombination methods are k_t, Cambridge/Aachen and anti-k_t algorithms.

k_t algorithm

The idea of the k_t algorithm originated from the JADE method. Some improvements have been introduced to better describe hadron collisions, in which the energy of the interaction is no longer well defined. In addition, the divergences in the QCD branching probability are not just between pairs of outgoing particles, but also between an outgoing particle and the incoming beam direction. This leads to the additional idea of a *particle-beam* distance d_{iB}.

In pp collisions it is common to use variables that are invariant under longitudinal boosts; all these requirements lead to the k_t distance definition

$$\begin{aligned} d_{ij} &= min(p_{Ti}^2, p_{Tj}^2)\Delta R_{ij}^2 \,, \qquad \Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2 \\ d_{iB} &= p_{Ti}^2 \quad . \end{aligned} \tag{3.2}$$

where p_{Ti} is the transverse momentum, y_i the rapidity and ϕ_i the azimuth angle of the particle i (the same for the particle j). In the collinear limits, these measurements reduce to relative transverse momenta. The sequential recombination procedure with the distance measures of Eq.(3.2) is referred to the *exclusive* k_t algorithm [31], in which every particles is assigned either to a beam-jet or to a final-state jet. In the following, this method was slightly modified in the definition of the d_{ij} distance with the introduction of a new parameter R (inclusive k_t [32]):

$$d_{ij} = min(p_{Ti}^2, p_{Tj}^2)\frac{\Delta R_{ij}^2}{R^2} \quad , \tag{3.3}$$

remaining the same procedure and ΔR_{ij} and d_{iB} definitions of Eq.(3.2).
The algorithm first determined all the d_{ij} and the d_{iB} distances, then evaluates the minimum among them; if it is a d_{ij}, it recombines i and j into a single new particle, while if it is a d_{iB}, it declares i to be a final-state jet and removes it from the list of particles. The sequence stops when no particles remain. Here all the particles are included in final-state jets: if a particle i has no other particles within a distance R, then the d_{iB} will be smaller than the d_{ij} for any j and the particle will then become a jet. Consequently arbitrarily soft particles become jets and therefore a minimum transverse momentum should be specified. In collider experiments the k_t algorithm is always referred to the inclusive and the same terminology will be used in this thesis.

Cambridge/Aachen algorithm

Originally born for e^+e^- collisions, Cambridge/Aachen (C/A) algorithm [33] introduces two distance measures between pairs of particles: $v_{ij} = 2(1-cos\theta_{ij})$ and y_{ij} of Eq.(3.2). It proceeds as follows: if only one particle is left, it calls it

jet and stops, otherwise it finds the pair of particles with smallest v_{ij}. Then, if the corresponding $y_{ij} < y_{cut}$ (fixed parameter), it replaces i and j with the recombined one, or it takes the less energetic of i and j, removes it from the list of particles, calls it jet and restarts. The general idea is to combine the jet resolution of the k_t algorithm with a clustering sequence dictated by angular ordering. The most widely extension of the C/A algorithm to hadron collider uses longitudinally invariant variables, introduces an R parameter and applies the y_{ij} cut on jets. It recombines the pair of particles with the smallest ΔR_{ij}, and repeats the procedure until all objects are separated by a $\Delta R_{ij} > R$; final objects are jets.

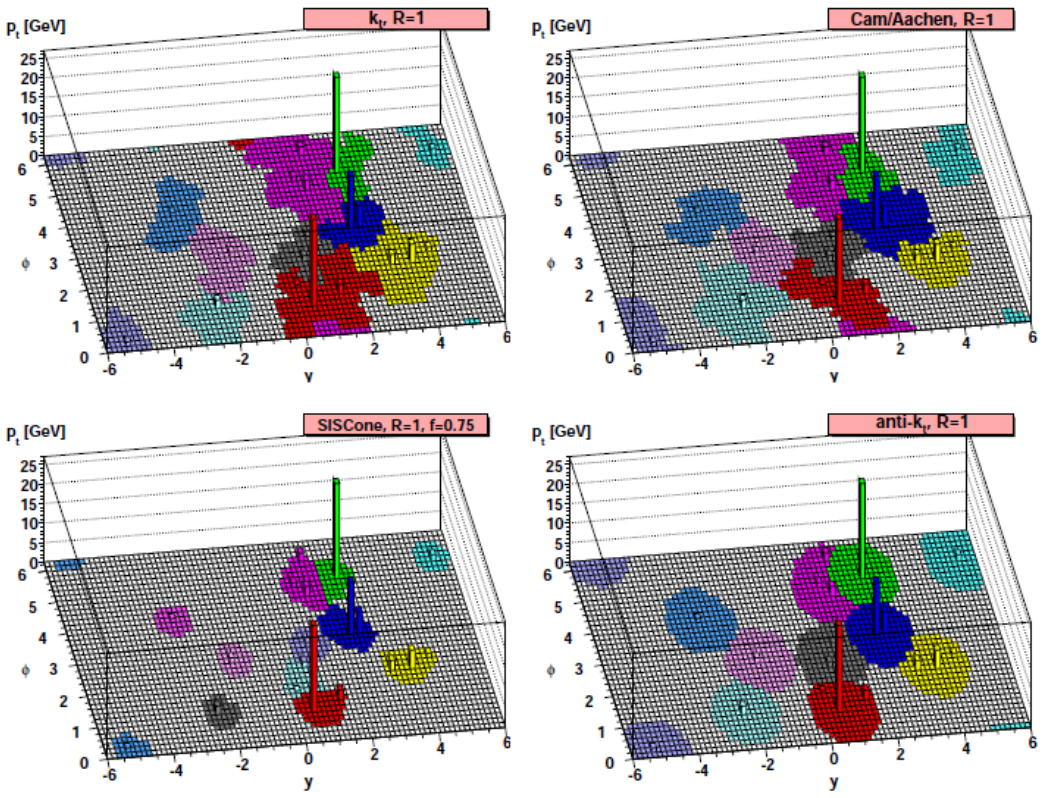


Figure 3.1: A sample parton level event, together with soft contributions, clustered with four different jet algorithms, illustrating the "active" catchment areas of the resulting hard jets [31].

anti-k_t algorithm

The k_t and C/A distance measures can be generalize as:

$$\begin{aligned} d_{ij} &= min(p_{Ti}^{2p}, p_{Tj}^{2p}) \frac{\Delta R_{ij}^2}{R^2} \\ d_{iB} &= p_{Ti}^{2p} \quad , \end{aligned} \tag{3.4}$$

where p is a parameter equal 1 for the k_t algorithm and 0 for C/A (ΔR_{ij} has the same definition as in Eq.(3.2)). In the the anti-k_t algorithm [34], the value of the p parameter is equal to -1; a feature of this method is to favour the clusterization of hard particles rather than soft ones (k_t algorithm) or energy-independent (C/A). A consequence of this is that isolated anti-k_t jets tend to be very close to circular in $\eta - \phi$ space (see Fig.3.1), because the axis of the jet is relatively fixed after the first few steps of recombination. The result is an infrared collinear safe algorithm that makes anti-k_t jets more robust than the k_t jets in high multiplicity environments. Standard ATLAS reconstruction employs anti-k_t algorithm with $R = 0.4$.

3.1.2 Leptons

The lepton+jets channel is characterised by the presence of a high energy p_T lepton, which is the main responsible of the event identification. In this thesis the analysis was based only on electrons and muons.

Electrons

The electron reconstruction performed in ATLAS is based on the matching between the Inner Detector tracks and the EM calorimeter clusters which give the angular direction and the energy respectively. The ATLAS electron reconstruction algorithm is therefore based on information coming from electromagnetic calorimeter layers, energy leakage in the hadronic calorimeter, the track quality criteria of the the ID objects and the cluster-track matching. This technique ensures a good discrimination from background objects, by mainly requiring electron isolation; it is based on the request of *Etcone*20 and

*Ptcone*30 variables, where *Etcone*20 is the total transverse energy (E_T) deposited in the calorimeter towers in a cone of radius $\Delta R = 0.2$ around the electron position, while *Ptcone*30 is the sum of the transverse momentum (p_T) of the tracks in the ID around a cone of $\Delta R = 0.3$.
The electron identification is based on three levels, called loose, medium and tight, each with progressively more stringent requirements. In this analysis the tight electrons have been used, which have passed all the following criteria:

- no error occurred in the LAr electromagnetic calorimeter during the data taking;
- the track should be identified as tight by a specific algorithm based on the shape of the shower deposited in the calorimeter, on the cluster matching with the associated track and on the number of hits in the inner part of the tracker;
- small distance between the track impact parameter and the primary vertex projection on the z-axis ($|z_0|$ <2 mm);
- the transverse energy should be above a fixed threshold $E_T > 25$ GeV;
- compatibility with the calorimeter acceptance 0< $|\eta_{cluster}|$ <1.37 and 1.52< $|\eta_{cluster}|$ <2.47;
- isolation cuts.

The tight electron identification efficiency is smaller than loose and medium one because of the requirements imposed in the selection; although its average value is around 80% as shown in Fig.3.2. In every identification level the efficiency distribution as a function of the number of primary vertexes is almost constant. In order to obtain a fine tune of the electromagnetic energy, a calibration based on the detection of the Z boson is used. The well-known mass of the Z boson and its decay in e^+e^- pairs are used to improve the knowledge of the electron energy scale and the linearity of the electromagnetic calorimeter.

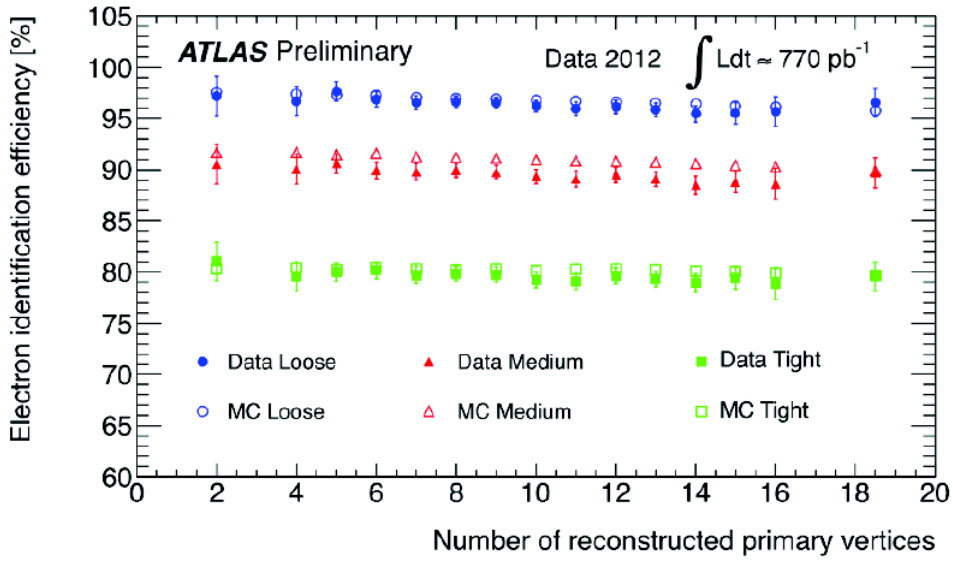


Figure 3.2: Electron identification efficiency with increasing number of primary vertices and pile-up, for different efficiency values [35].

Muons

Muon reconstruction employs information from ID tracks, calorimeter cells and Muon Spectrometer, even if the main reconstruction hit information is given by the MS. It reaches an identification efficiency better than 95%, as shown in Fig.3.3, and a relative momentum resolution of about 1-3%. According to the detector information used for their identification, four different muon candidates are identified, as shown in Fig.3.4:

- **combined**, reconstructed with information from all the three detectors;
- **standalone**, identified only through the muon spectrometer and the calorimeter cells;
- **segment-tagged**, reconstructed with the matching of the ID track with the corresponding segment in the MS region of poor coverage;
- **calo-tagged**, identified with the ID track combined with a minimal ionizing particle (MIP) signature in calorimeter cells.

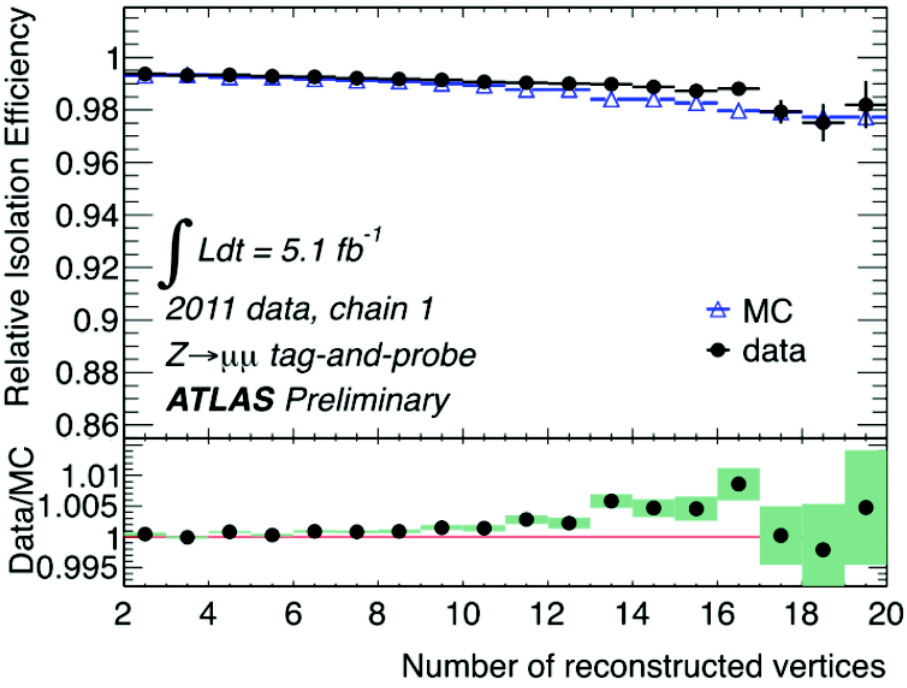


Figure 3.3: Stability of muon isolation efficiency with increasing number of primary vertices, for combined muons [38].

ATLAS uses two distinct muon reconstruction procedures, the MuID [36] and the STACO [37] methods, which differ in the combination scheme of ID ans MS tracks. MuID employs a refitting procedure starting from the ID tracks to the MS ones; on the other hand STACO combines all the track vectors. In this analysis the MuID muons have been used, with the requirements:

- to be identified as a tight muon, which means as a combined or standalone muon with at least three hits in both precision drift tube chambers and cathode strip chambers of the muon system;
- small distance between the track impact parameter and the primary vertex projection on the z-axis ($|z_0|$ <2 mm);
- to satisfy the isolation criteria Etcone20<4 GeV and Ptcone30<2.5 GeV.

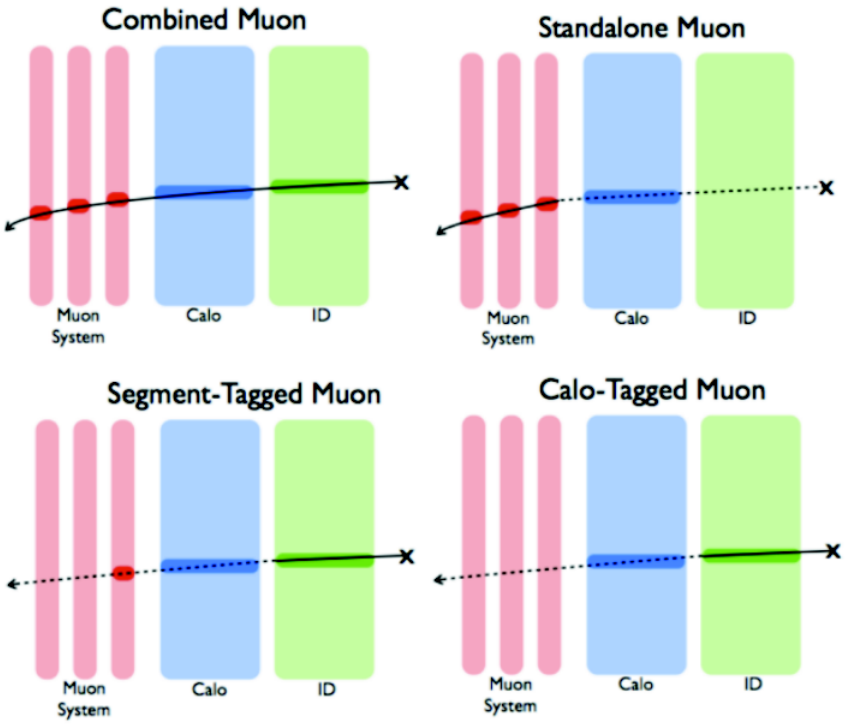


Figure 3.4: The four kinds of muon candidates in ATLAS: combined, standalone, segment-tagged and calo-tagged muons.

3.1.3 Neutrinos

Neutrinos arise from leptonic W boson decay and do not interact in the detector. Their presence can be inferred from the imbalance of the energy in the transverse plane (missing transverse energy). Because at hadron colliders the initial momentum of the colliding partons along the beam axis is not known a priori, the momentum conservation can only be applied in the transverse plane. In fact, the vector sum of the interacting partons transverse momenta is in good approximation zero, allowing the measurement of the *missing transverse energy* E_T^{miss}, defined as

$$E_T^{miss} = \sqrt{(E_x^{miss})^2 + (E_y^{miss})^2} \quad , \tag{3.5}$$

where x and y are the directions perpendicular to the beam pipe. This quantity is obtained from the vector sum of all calorimeter clusters in the transverse

plane. Cells not associated to muon, electron, photon, τ candidates and jets are included at the EM scale. The ATLAS algorithm [39] for the E_T^{miss} evaluation includes contributions from topo-clusters transverse energy $E_{x,calo}^{miss}$, corrections for energy losses in the cryostat system $E_{x,cryo}^{miss}$ and reconstructed muons $E_{x,\mu}^{miss}$

$$E_x^{miss} = E_{x,calo}^{miss} + E_{x,cryo}^{miss} + E_{x,\mu}^{miss} \quad , \tag{3.6}$$

where all the components are on the x axis (identical considerations for the y axis). The $E_{x,calo}^{miss}$ term is evaluated only from cells belonging to topological clusters and included in the range $|\eta|$ <4.9. The calorimetric cells are calibrated in different ways according to the reconstructed object they belong to; this implies that $E_{x,calo}^{miss}$ is the sum of different components evaluated as the negative sum of the energy deposits in the calorimeter cells associated to the correspondent object typology (muons, electrons, photons, taus and jets). The cryostat term $E_{x,cryo}^{miss}$ gets rid of the non-negligible loss of energy in hadronic showers due to the cryostat system installed between the LAr electromagnetic calorimeter and the Tile hadronic calorimeter. It is evaluated via the energy correlation between the last LAr layer and the first Tile one. The muon term $E_{x,\mu}^{miss}$ is evaluated from the ID and MS muon information.
The E_T^{miss} performance is established from differences between data and MC simulations distribution in $Z \rightarrow \mu\mu$; Fig.3.5 shows the resolution of $Z \rightarrow \mu\mu$ events as a function of the number of primary vertices before and after pile-up subtraction.

3.2 Boosted top reconstruction

LHC is exploring a completely new kinematic regime, where the Standard Model particles can be produced at a center-of-mass energy which largely exceeds their masses. At these energies, heavy particles as W and Z bosons and the top quark can be produced with a large momentum (*boosted* particles). At high enough p_T, their decay products will appear as collimated jets [40]; as a consequence the standard resolution techniques begin to fail, because of

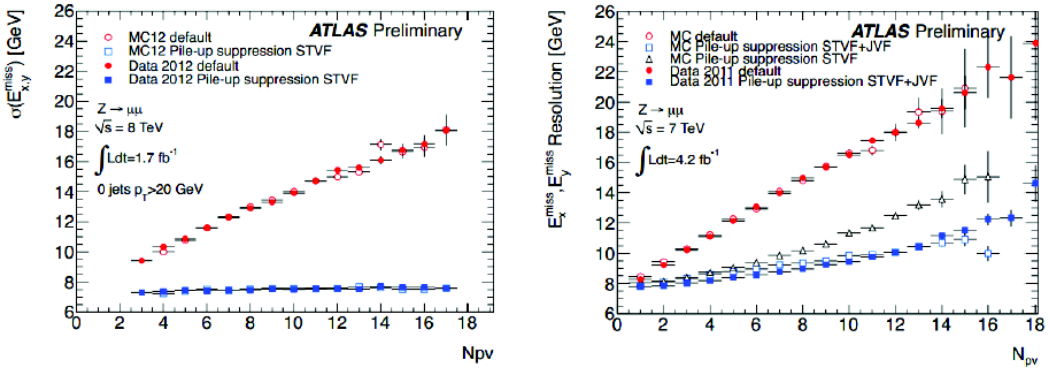


Figure 3.5: Comparison between data and simulation of E_T^{miss} (*letf*) and E_x^{miss}, E_y^{miss} (*right*) resolutions as a function of the number of primary vertices [39].

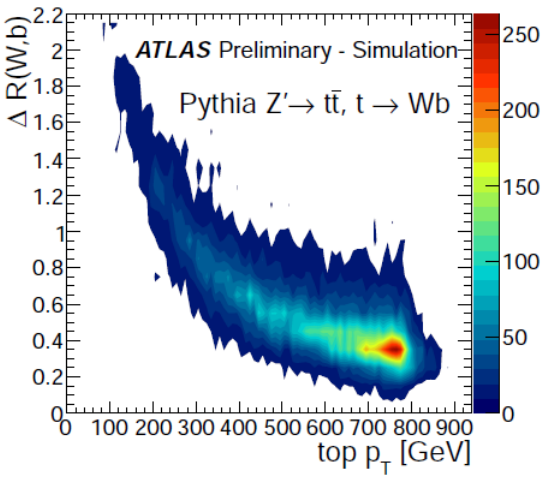
the partial or total overlap of the signals in the detector. The large integrated luminosity collected at $\sqrt{s}$ =8 TeV permitted to explore the top production at high energy with unprecedented sensitivity.

As the p_T of a top quark increases, its decay products are boosted into increasing narrower regions (Fig.3.6) complicating attempts to individually identify them. This implies that in the lepton decay the emerging charged lepton is no longer isolated from the nearby energy flow, increasing the background contamination. While in the hadronic top decay the emerging partons can be detected as a single, energetic and large radius jet "fat-jet". Fig.3.6a shows the distribution of the average distance ΔR between the b and W produced in a top decay process as a function of the top pT, as well as the separation between the light quarks of the subsequent hadronically-decaying W, as a function of the W transverse momentum. In each case, the angular separation of the decay products is approximately

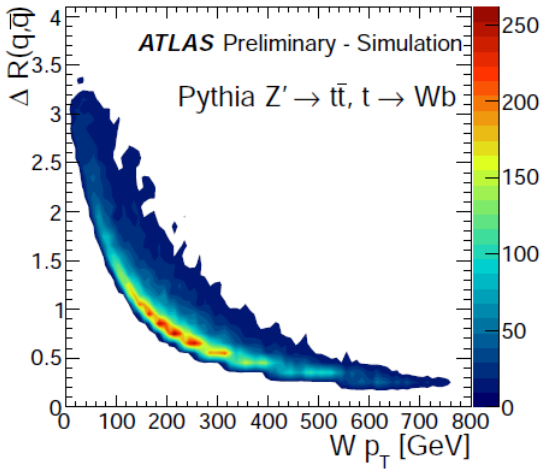
$$\Delta R \sim \frac{2m}{p_T} \quad , \tag{3.7}$$

where p_T and m are the transverse momentum and mass of the decaying particle respectively. For p_T^W ~200 GeV, the ability to resolve the individual hadronic decay products using standard narrow-cone jet algorithms begins to degrade, and above p_T^{top} ~350 GeV, the decay products of the top quark tend to have a separation ΔR <1, smaller than the double of the anti-k_t cone radius

usually fixed at 0.4 in standard jet reconstruction. Fig.3.7 shows the differences in the jet configuration for the resolved and the boosted tops, where the fat-jet contains all the decay products of the top quark.



(a) $t \to Wb$



(b) $W \to q\bar{q}$

Figure 3.6: (a) The opening angle between the W boson and b quark in top decays ($t \to Wb$) as a function of the top p_T in simulated PYTHIA events. (b) The opening angle of the $W \to q\bar{q}$ process from top decays as a function of the p_T^W. Both distribution are at the particle level [41].

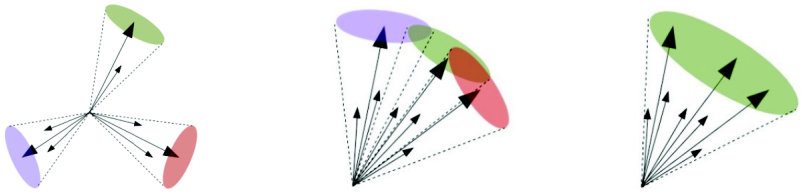


Figure 3.7: Graphical representation of jets produced in a top decay event in case of low (*left*) and high (*centre*) values of top p_T. On the right there is the high top p_T configuration reconstructed using a large-R jet.

The fat-jet has significantly different properties than a jet of the same p_T originating from a single light quark or gluon. The characteristic three-body decay results in a hard substructure that is absent from the light-quark or gluon jets. Recently, new techniques have been proposed in order to recognize fat-jets belonging to massive particle decays with the aim of increasing efficiency. The *tagging algorithms* involve the study of the substructures of the fat-jet; a detailed description of these methods is presented in Section 3.4. The first problem associated to the presence of a fat-jet is the increasing of the background contribution inside it, due to its angular extension. The rise of the background contributions are essentially two: pile-up caused by the fact that during the 2012 data taking ATLAS acquired on average 25 interactions per bunch crossing and the soft QCD radiation (underlying event, UE) due to other interactions of the other partons of the beam proton in the same event. For this reason particular studies have been implemented in order to recognize and decrease the background contributions (grooming methods) described in Section 3.3.

3.3 Jet Grooming Techniques

The jet grooming is the selective removal of soft radiation during the process of iterative recombination in jet reconstruction. Recently many jet grooming algorithms have been designed to remove soft energy deposit contributions

inside the fat-jet in order to improve resolution of the reconstruction of hard decay products from a boosted object. The differences among jets formed from light quarks or gluons and from hadronic particle decay are the basis of these tools. The general aims are to reduce the impact of fluctuations coming from parton shower and underlying event and to mitigate the influence of pile-up. Therefore jet grooming techniques can be a powerful method to discriminate between the dominant multi-jet background and the heavy particle decay. The main jet grooming techniques of the ATLAS experiment are introduced in the next sections.

3.3.1 Mass-drop filtering

This method is based on the attempt to isolate concentration of energy within a jet by identifying symmetric subjets, each with a smaller mass than that of their sum. The mass drop filtering is applied only to C/A jets because their structure provides an angular ordered description of substructures, which tends to be one of the most useful properties in the research of hard splittings within a jet. This procedure undoes the last step of the C/A clustering so that the jet divides into two subjets j_1 and j_2. The mass-drop criterion requires that there should be a difference between the original jet mass m_{jet} and the mass jet m_{j_1} after the splitting: $m_{j_1}/m_{jet} < \mu$, where μ is a parameter of the algorithm. The splitting is also required to be symmetric:

$$\frac{min[(p_T^{j_1})^2, (p_T^{j_1})^2]}{(m_{jet})^2} \times \Delta R^2_{j_1,j_2} > y_{cut} \quad , \tag{3.8}$$

where $\Delta R_{j_1,j_2}$ is the opening angle between the two jets and y_{cut} defines the energy sharing between the two highest p_T subjets within the original jet. The typical y_{cut} is 0.09, as obtained in previous studies. The next step of the algorithm is called filtering, a procedure in which the constituent of j_1 and j_2 are reclustered with the C/A algorithm with a smaller opening angle. The procedure ends when all constituents outside the three hardest subjets are discarded. This technique was developed in the search for a Higgs boson decaying in the channel $H \rightarrow b\bar{b}$ [42]. For this reason, the mass-drop filtering criterion

is not effective for three-body decay tagging, like the top quark; however it is used as starting point for other techniques.

3.3.2 Trimming

The trimming procedure [43] removes contamination from pile-up, multiple parton interactions (MPI) and initial-state radiation (IRS), taking advantage of the fact that they are much softer than the hard parton products. The selection criterion is in fact based on the ratio of the p_T of the constituents to that of the jet. The trimming method uses the k_t algorithm to create subjets with R_{sub}, usually set to 0.2, from the constituents of the fat-jet: any subjets with $p_{Ti}/p_T^{jet} < f_{cut}$ are removed, where p_{Ti} is the transverse momentum of the i^{th} subjet and the values of f_{cut} parameters are around a few percent. The remaining constituents form the trimmed jet; the selection procedure is illustrated in Fig.3.8.

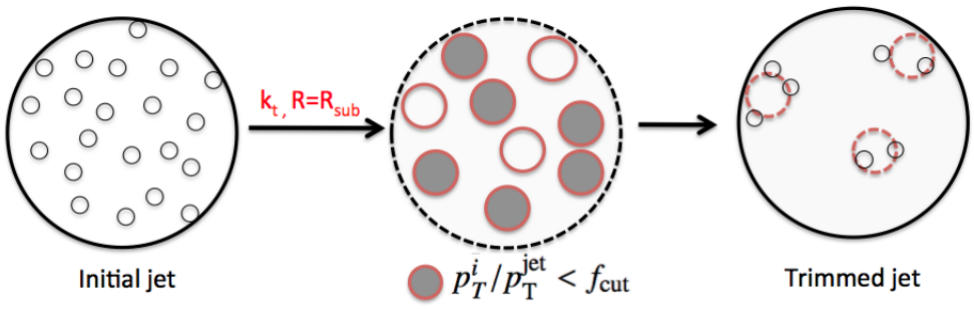


Figure 3.8: A representation of the trimming procedure.

Low-mass jets from a light quark or gluon usually lose 30-50% of their mass; on the other hand, jets containing the decay products of a boosted object will lose only a few percent of their mass, mostly due to pile-up and underlying events.

The trimming algorithm is the standard grooming procedure adopted by ATLAS for boosted top analysis, applying the method to the anti-k_t fat-jet with R=1.

3.3.3 Pruning

In addition to the soft component removal, the pruning procedure [44] employs a wide-angle radiation veto. The fat-jet constituents are utilized to reconstruct again the jet with the C/A or k_t algorithm; at each pseudo jet recombination step, the following cuts are placed:

$$\frac{p_T^{j_2}}{p_T^{j_1+j_2}} > z_{cut} \qquad and \qquad \Delta R_{j_1,j_2} < R_{cut} \times \frac{2m^{jet}}{p_T^{jet}} \quad , \tag{3.9}$$

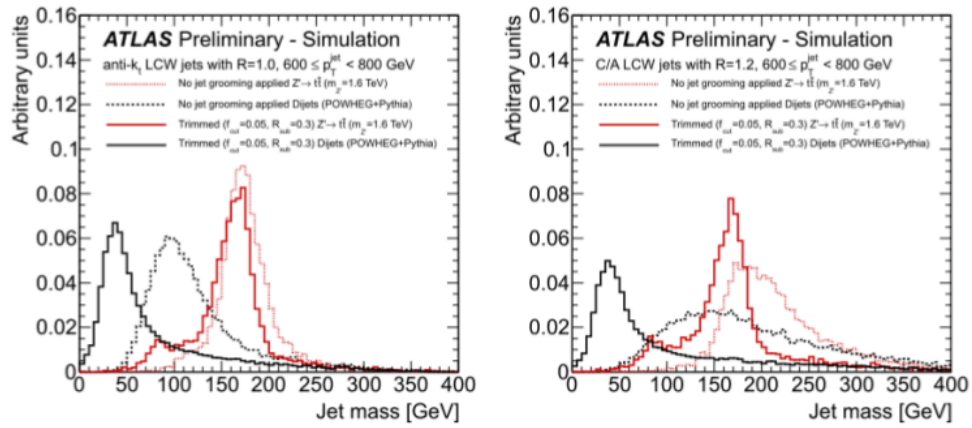
where j_1 and j_2 are the jet constituents ordered $p_T^{j_1} > p_T^{j_2}$ and z_{cut} and R_{cut} are the parameters of the tagger, which can assume a relatively wide range of possible configurations. If the previous criteria are met, j_1 and j_2 are merged, otherwise j_2 is discarded and the algorithm continues. It is important to remark that the requirement above are not directly related to the original fat-jet, but to the proto-jet formed in the new reconstruction process. In Fig.3.9 the effect of trimming algorithm on the distributions of jet mass and splitting scale variables is shown in the range 600$\leqslant p_T^{fat} \leqslant$800 GeV. For these studies a $Z' \rightarrow t\bar{t}$ Monte Carlo sample ($m_{Z'}$ =1.6 TeV) has been considered for signal-like events (red), compared to di-jets background (black). The grooming procedure increases the separation between signal and Monte Carlo distributions for all the substructure variables considered, helping the discrimination based on these quantities.

3.4 Top tagging techniques

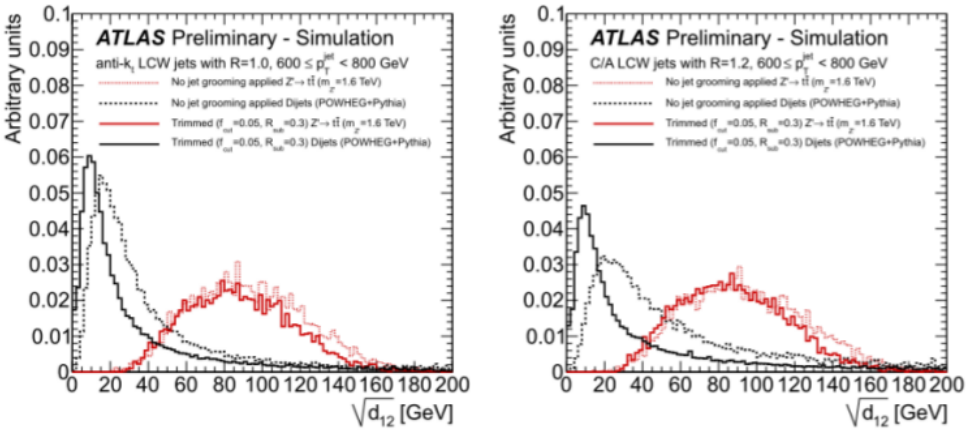
In this section the main tagging algorithms employed by the ATLAS experiment are presented [41].

3.4.1 Jet Mass

The jet mass is the result of the 4-vector addition of the jet components, thus it is interesting to discuss to what extent it is related to that of the original



(a) anti-k_t, R = 1.0 (b) C/A, R = 1.2



(a) anti-k_t, R = 1.0 (b) C/A, R = 1.2

Figure 3.9: Comparison of POWHEG $Z' \rightarrow t\bar{t}$ signal to multi-jet background as a function of jet mass and of splitting scale $\sqrt{d_{12}}$ in the range $600 \leqslant p_T^{jet} \leqslant 800$ GeV. The dotted lines show the ungroomed jet distribution, while the solid lines show the corresponding trimmed (f_{cut} =0.05 and R_{sub} =0.3) jets. The distributions are reconstructed both with the anti-k_t (*left*) and C/A (*right*) algorithms.

partons. The jet mass m_{jet} is calculated from the energy and momenta of its constituent

$$m_{jet} = \sqrt{(\sum_i E_i)^2 - (\sum_i p_i)^2} \quad , \tag{3.10}$$

where E_i and p_i are the energy and three-momentum of the i^{th} jet constituents. At the particle level jet constituents are the stable particles coming from parton hadronisation, while at the reconstruction level m_{jet} is calculated by summing over the topo-cluster or tracks belonging to the jet. The mass jet is a fundamental measurement in the research of boosted objects and can be used to discriminate signal from background.

3.4.2 Splitting Scale

The splitting scale is defined by reclustering the constituents of a fat-jet with the k_t recombination algorithm. The k_t-distance of an intermediate step combining two subjets into the final jet can be used to identify a splitting scale variable as

$$\sqrt{d_{ij}} = min(p_{Ti}, p_{Tj}) \times \Delta R_{ij} \quad , \tag{3.11}$$

in which ΔR_{ij} is the distance between the two subjets. According to this definition, the subjets determined at the last step of the k_t reclustering provide the $\sqrt{d_{12}}$ variable. Similarly, $\sqrt{d_{23}}$ represents the splitting scale in the second to last step of reclustering. The splitting scale definition is equivalent to the square of the distance parameter in Eq.(3.3), multiplied by the jet radius parameter R. The most useful parameters for boosted tagging are $\sqrt{d_{12}}$ and $\sqrt{d_{23}}$. Since the k_t algorithm combines the harder constituents last, they can distinguish heavy particle decays, which tend to be symmetric, from the largely asymmetric splittings in light quark and gluon jets.

3.4.3 N-Subjettiness

The N-Subjettiness or τ_N variable is related to the subjet multiplicity [45]. It is calculated by clustering the constituents of the jet with the k_t algorithm and requiring exactly N subjets. This is obtained using the exclusive k_t algorithm which stops when there are exactly N proto-jets remaining. These N final subjets define an axis within the jet, around which the jet constituent may be

concentrated. Thus τ_N is defined as the sum over all the k constituents:

$$\tau_N = \frac{1}{d_0} \sum_k p_{Tk} \times min(\delta R_{1k}, \delta R_{2k}, \ldots, \delta R_{Nk}) \qquad d_0 \equiv \sum_k p_{Tk} \times R \quad . \tag{3.12}$$

In the equation above, R is the jet radius parameter of the algorithm, p_{Tk} is the p_T of the k^{th} constituent and δR_{ik} is the distance from the subjet i to the constituent k. N-subjettiness variable specifies how well jets can be described as containing N or fewer k_t subjets by evaluating the distance between constituents from the axes of these subjets. The ratios τ_2/τ_1 (denoted τ_{21}) and τ_3/τ_2 (denoted τ_{32}) can provide discrimination between jets formed from the parton shower of light quarks or gluons and jet containing three hadronic products from boosted top quarks.

3.4.4 HepTop Tagger

An alternative class of tagging methods reverses the clustering sequence of the candidate top quark jet and checks the angular or energy scales involved at a given clustering step for consistency with top quark decays. The HEPTop Tagger [46] is one of such a method, optimized for moderately boosted top quarks that are sufficiently boosted for their decay products to lie inside a single fat-jet. In this method, extra radiation from the candidate jet are removed by filtering out soft constituents and the remaining energy deposits are reclustered into three hard subjets. Finally, if the invariant mass of these three subjets is consistent with the top quark mass, the jet is tagged as a top quark event. The performance of the HEPTop Tagger has been studied extensively using ATLAS pp collision data and simulated events.

The HEPTop Tagger method operates on a fat-jet that has been constructed using the C/A algorithm; C/A is employed to re-cluster the fat-jet constituents into subjets together with the mass-drop filtering criterion in order to utilize information about the recombination history of the jet. The HEPTop Tagger chain is shown in Fig.3.10.

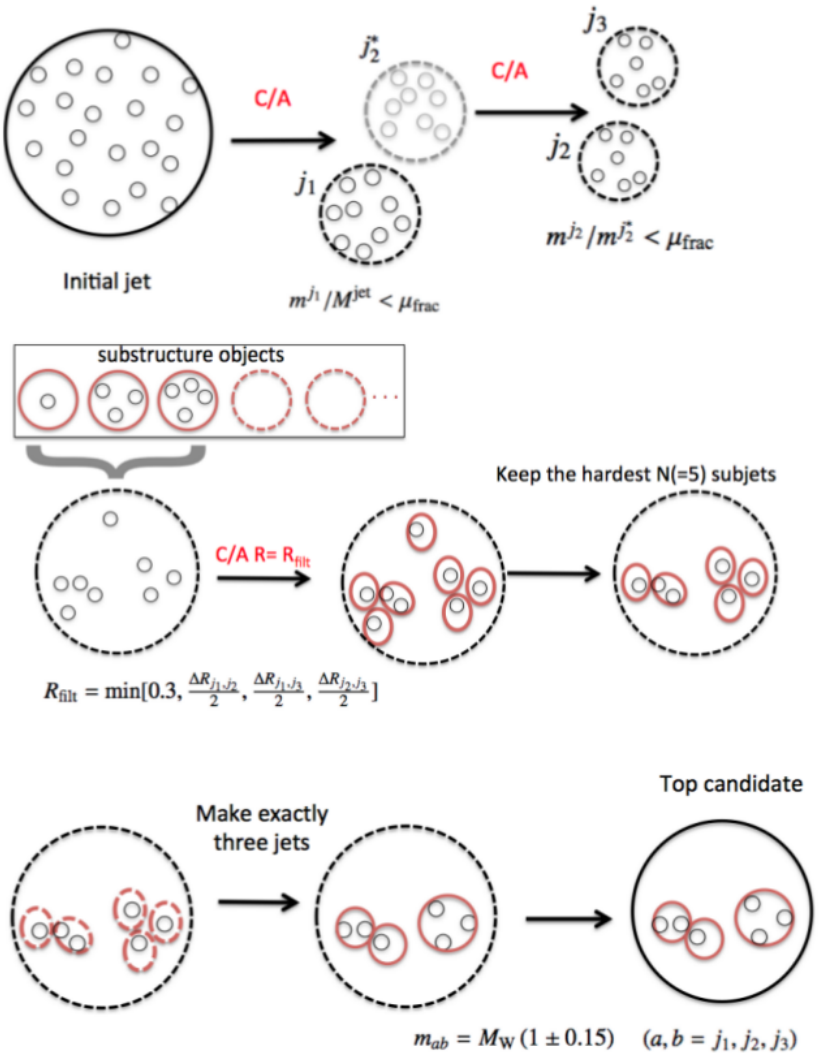


Figure 3.10: A representation of the HepTop Tagger algorithm chain.

The algorithm proceeds as follows:

- the mass-drop criterion is applied to a large-R C/A jet, in which j_1 and j_2 are the two subjets at the last stage of clustering. If the criterion $m_{j_i}/m_{jet} < \mu$ is satisfied, the procedure is iteratively applied to all the subjets passing the mass cut, until $m_i < m_{cut}$. When all the subjets have been de-clustered, at least three substructure objects must survive,

otherwise the jet is discarted;

- combinations of three substructure objects are filtered at a time, reconstructing their N_i constituents using the C/A algorithm with a distance parameter $R_{filt} = min[0.3, \Delta R_{j_1,j_2}/2]$, where $\Delta R_{j_1,j_2}$ is the minimum separation between all possible pairs in the current triplet;

- the invariant mass of the four-vector determined by the sum of the resulting proto-jet should be near to the mass of the top quark ($140 \leqslant m_{jet} \leqslant 200$ GeV), otherwise the triplet combination is ignored;

- from the N_i subjets formed from the chosen top triplet, only a number N_{subj} are selected, where $3 \leqslant N_{subj} \leqslant N_i$. Of these N_i subjets, exactly three jets are built by applying the C/A algorithm to the constituents of the N_{subj} subjets, which are then properly calibrated;

- the triplet is accepted as top candidate if one of the following criteria on the invariant mass of two (m_{ij}) and three (m_{123}) subjet combinations is satisfied:

$$R_{min} < \frac{m_{23}}{m_{123}} < R_{max} \qquad 0.2 < arctan\frac{m_{13}}{m_{12}} < 1.3 \quad , \tag{3.13}$$

where R_{min} and R_{max} are method parameters.

3.4.5 Template Overlap Method

In the analysis presented in this thesis, highly boosted top quarks are identified with the Template Overlap Method (TOM) [47]-[48], but this procedure is also applicable to other boosted particles (e.g. Higgs bosons decaying to $b\bar{b}$ pairs).

TOM employs a set of infrared-safe observables specifically designed to distinguish boosted heavy particle decays from QCD jets and others coming from background processes. TOM is a new method still not applied in the standard analysis, optimized for the hadronic top decay produced at high p_T when jets are partially or totally overlapped: for this reason the starting point of the

algorithm is the fat-jet reconstruction. It is based on a comparison between the p_T distribution in the $\eta - \phi$ plane (energy flow in the following) of the observed fat-jets, with the one from a simulated top decaying in the hadronic mode (the template). The template is a set of various millions of generated top with uniform momentum distribution, emitted along the x axis. Further details on template kinematics are given in the next paragraph. In order to evaluate the agreement between the detected fat-jet and a particular top of the template set, it has been implemented an overlap estimator defined between 0 and 1. For each fat-jet candidate, the overlap function is defined as

$$Ov = max_{\tau_n} exp\left[-\sum_{i=1}^{3}\frac{1}{\sigma_i^2}\left(p_{T,i} - \sum_j p_{T,j}F(\hat{n}_j, \hat{n}_i)\right)^2\right] \quad , \tag{3.14}$$

where τ_n is the set of templates, i is the index of partons in each template set that runs from 0 to 3, $p_{T,i}$ are the parton momenta of the top quark decay products for the given template, $p_{T,j}$ are the momenta of all the clusters belonging to the fat-jet close to the direction of a parton of the template and σ_i is the uncertainty of the reconstructed $p_{T,j}$ variable. The function $F(\hat{n}_j, \hat{n}_i)$ is defined to have nonzero value only in the regions around the directions of the template momenta:

$$F(\hat{n}_j, \hat{n}_i) = \begin{cases} 1 & if\, \Delta r < r_i \\ 0 & otherwise \end{cases} \quad , \tag{3.15}$$

where Δr is the Sub Cone Radius variable ($= \sqrt{\Delta\phi^2 + \Delta\eta^2}$) defined in analogy with R, which represents the angular distance between the template parton i and a fat-jet constituent j in the (η, ϕ) space. $\hat{n}_i$ and $\hat{n}_j$ are the directions of the parton and of cluster and the parameter r_i determines the angular scale of the template subjet. Typical value of is 0.2.

The first step of the method is the generation of millions of top quarks along the x axis with uniform momentum, obviously it will be rotated to the fat-jet axis before to perform the comparison. The produced hadronic top quarks decay to the three partons following the relativistic kinematics, without considering the parton hadronization. The p_T deposit of each cluster of the fat-jet

within a radius r_i from the direction of the parton of the template are summed and then compared with the p_T of each parton.
The overlap variable quantifies how well the template matches the p_T flow. For each template the overlap value is evaluated and, at the end, the maximum value is selected. The method provides an unique tool to match unequivocally the top decay products to the partonic partners. For construction the overlap function is a defined in the range [0,1], where values close to 1 (0) mean high (low) probability that the fat-jet has been produced by a top. In this analysis we have established to set ov=0.7 as the minimum value to define a fat-jet coming from a top.

Template Kinematic Distributions

The template set contains 4 millions top quarks generated with uniform momentum in the range [250;1000] GeV and along the x axis (ϕ =0 and η =0) of the ATLAS reference frame (see Fig.3.11).

The generated top quarks decay to three partons following the relativistic kinematics, without considering the hadronisation. In Fig.3.12 the kinematic quantities of the W coming from the top decay are plotted; as expected W bosons are emitted around top direction (especially the high momentum ones). In Fig.3.13 and Fig.3.14 are plotted respectively the momentum and the direction of the b quark coming directly from the top decay and the two quarks belonging to the W decays.

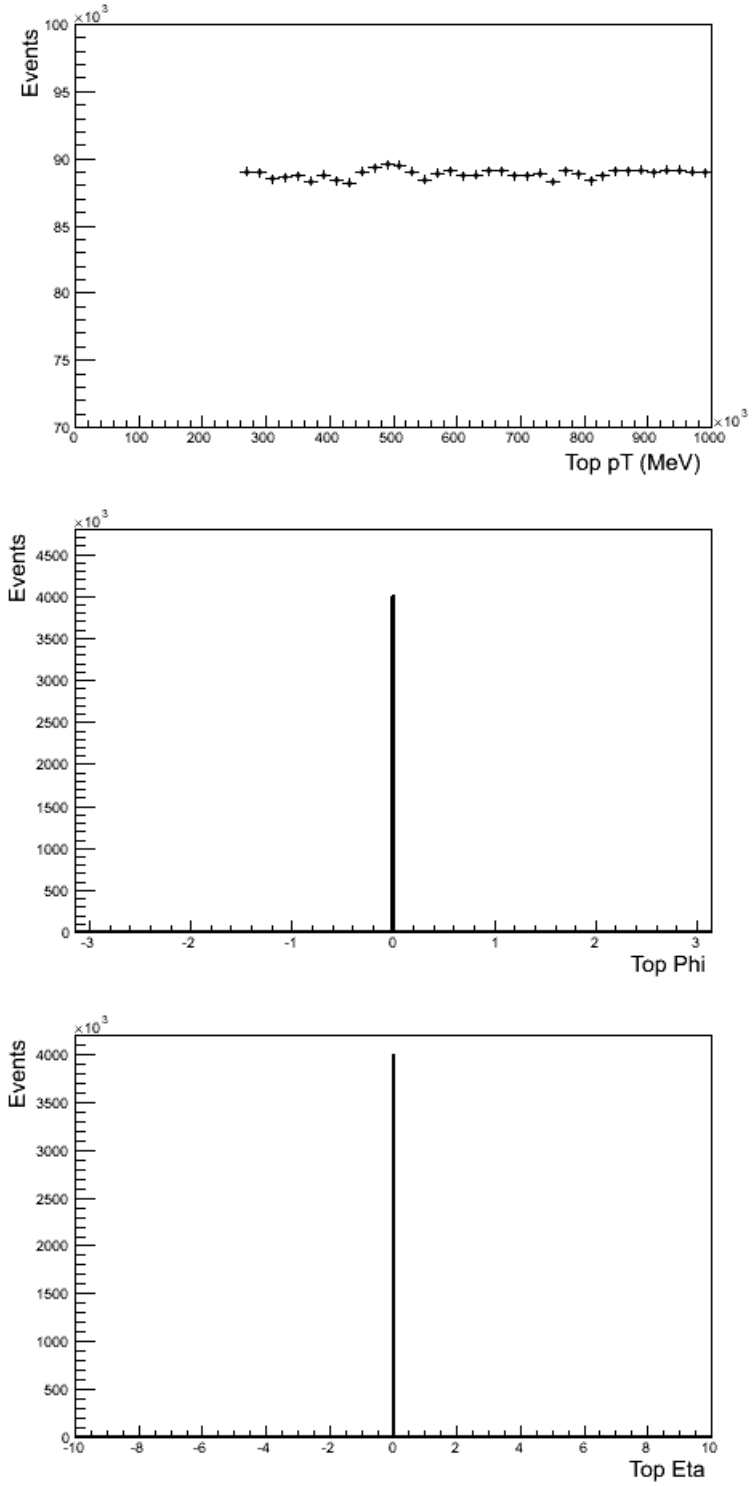


Figure 3.11: Distribution of the E, p_T, ϕ and η variables of the top template.

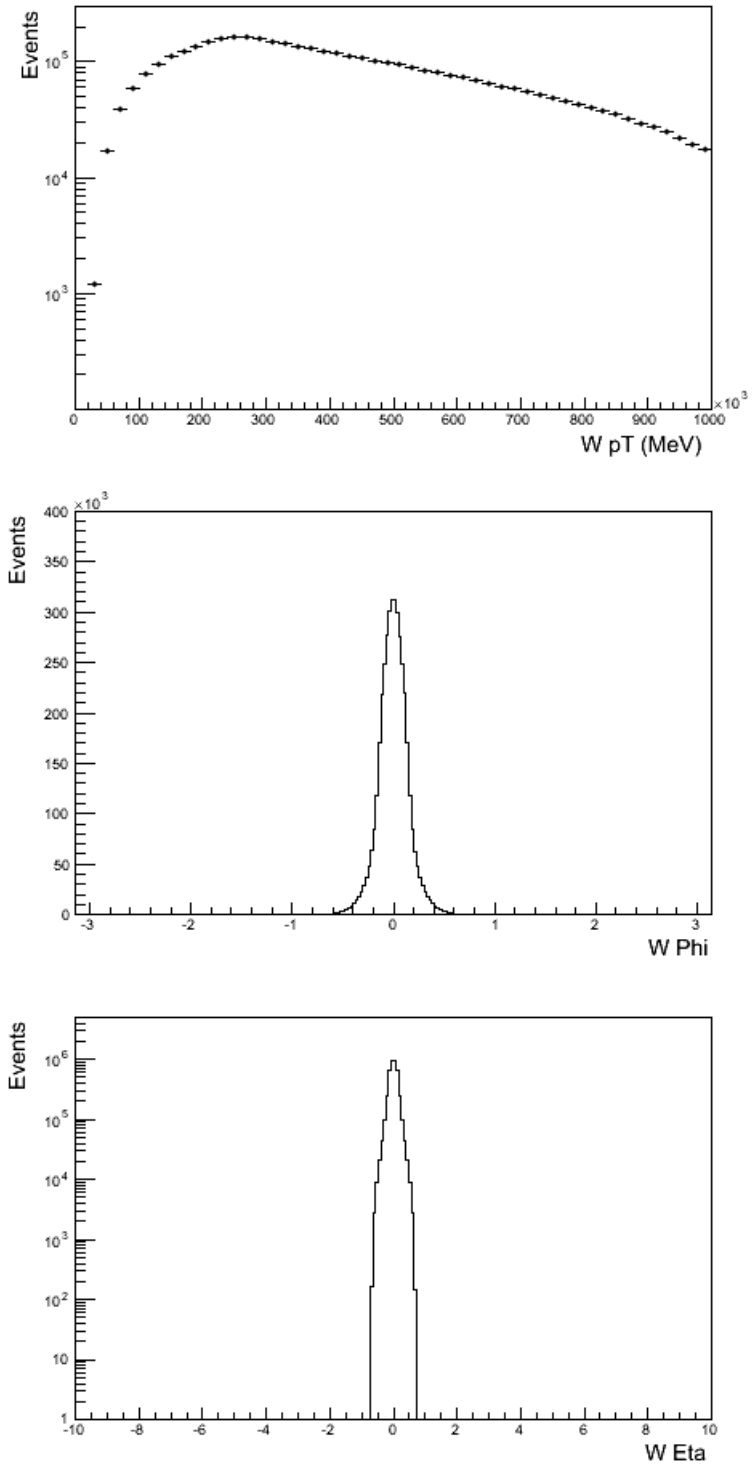


Figure 3.12: Distribution of E, p_T, ϕ and η variables of the W boson coming from the decay of the generated top. E, p_T and η are represented in a logarithmic scale.

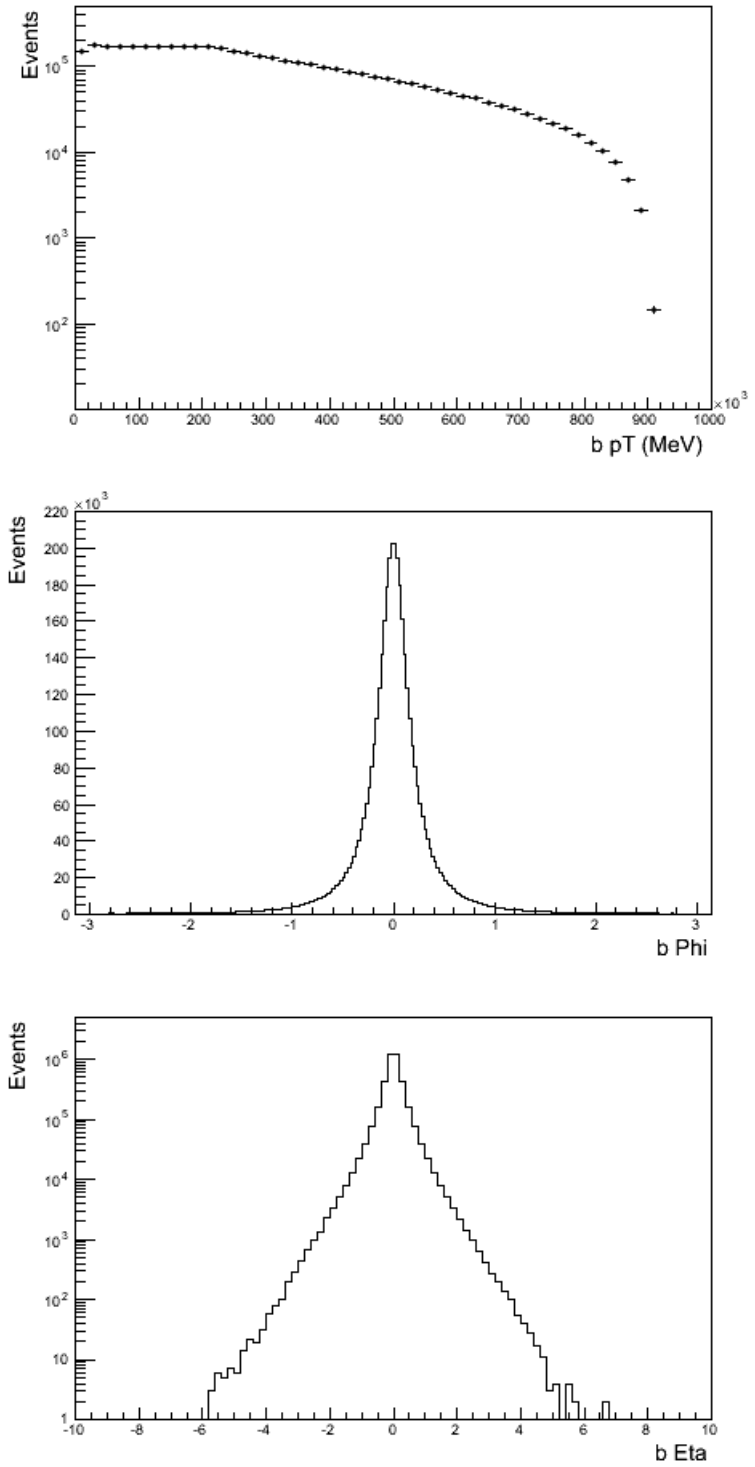


Figure 3.13: Distribution of E, p_T, ϕ and η variables of the b quark coming from the decay of the generated top. E, p_T and η are represented in a logarithmic scale.

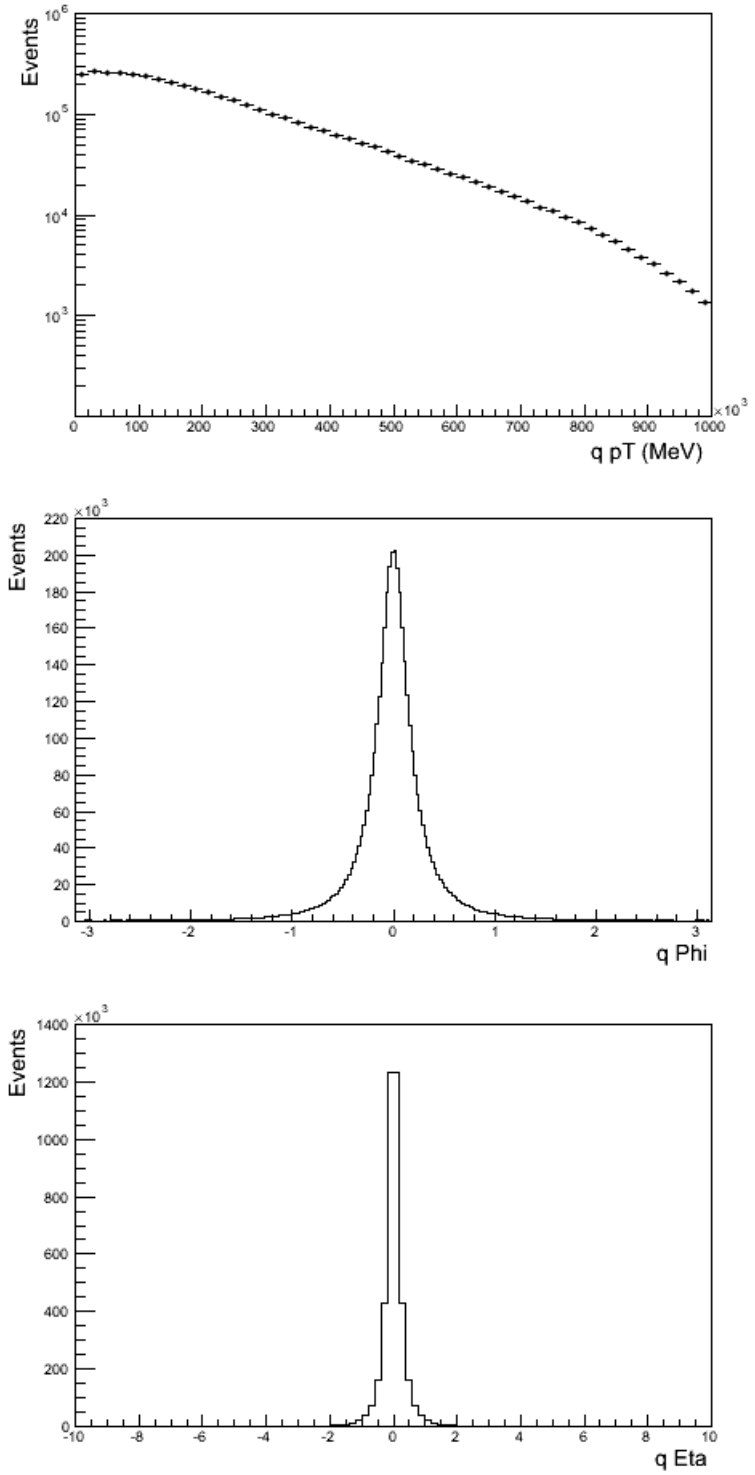


Figure 3.14: Distribution of E, p_T, ϕ and η variables of one of the two quarks produced from the W boson decay (analogue trend for the other quark). E, p_T and η are represented in a logarithmic scale.

3.5 Data and Monte Carlo Samples

3.5.1 Data sample

The data used for this analysis have been acquired during 2012 and correspond to pp collisions at $\sqrt{s}$ =8 TeV, with an integrated luminosity of 20.3 fb^{-1}. The uncertainty on the integrated luminosity is ±2.8%, derived from a calibration of the luminosity scale performed in dedicated Runs [49]. Only the events where the detector was fully functional have been used and the data are required to meet all the quality criteria. These requirements reject significant contamination from detector noise, non-collision beam backgrounds, and other spurious effects. The ATLAS data quality selection is based on individual assessments for each detector part, as well as for trigger and for each type of reconstructed physics object (jets, muons, electrons, neutrinos).

3.5.2 Monte Carlo simulation

The physical processes and the detector responses are simulated by Monte Carlo programs, which produce samples according to both theoretical prediction and phenomenological models. Monte Carlo simulations are used to verify our comprehension of the physics processes and the behaviour of the experimental apparatus.
The event generation is performed by theoretical calculation of the elementary processes from the pp interaction to the final state stable particles, through all the intermediate steps. First the Monte Carlo generators evaluate the matrix element of the hard process (parton-parton interaction) at a fixed perturbative order; then they calculate the cascades of radiation produced from QCD processes and interactions, called parton shower. Parton shower is a space-like process for the initial state partons and a time-like process if applied to the final state ones. In the initial state, the QCD radiation emission increases the virtuality of the initial state partons, allowing them to access the hard scale needed for the hard scattering process. After the scattering, the time-

like parton shower allows the high-virtuality partons emerging from the hard scattering to loose their virtuality towards the hadronization phase (which is a non-pertutbative process). The hadronisation is the mechanism which produce stable particles from quarks and gluons. These final particles are passed through the detector simulator, GEANT4 [50], that provides a model for the particle interaction through matter.
The Monte Carlo samples for signal and background are the ones used by the Top Working Group, mostly documented in [51]; in the following, only the samples used in this analysis are described.

Monte Carlo Signal

Two different signal MC samples have been taken into account in this analysis: direct $t\bar{t}$ production and $Z' \rightarrow t\bar{t}$. The $t\bar{t}$ sample is produced with the POWHEG [52] generator using the CT10 PDF set [53] in the matrix element; this is interfaced with PYTHIA6 [54] using CTEQ6L1 PDF and Perugia tune for the hadronization and underlying event models.
To study the boosted top quark tagging performance with higher p_T top distribution, a beyond Standard Model process has been supposed, generating an heavy Z' with mass $m_{Z'}$ =1750 GeV decaying to a $t\bar{t}$ couple. This process is produced using PYTHIA8 with the MSTW2008 LO PDF [58] and the ATLAS AU2 tune [57]. In Tab.3.1 the production cross section for the MC signal samples at the $\sqrt{s}$ =8 TeV center-of-mass energy are reported.

Monte Carlo Backgrounds

Background sources affecting the $t\bar{t}$ production channel are those processes leading to a signature similar to the $t\bar{t}$ one. This could be caused both to the similarity of the decay products or to the not negligible probability of object misidentification by the detector. The processes giving the larger background contribution to the $t\bar{t}$ pair signal in the lepton+jets channel are single top production, multijets events from QCD radiation, W+jets, Z+jets and diboson events. In Tab.3.1 the cross sections of the background processes at ATLAS

Table 3.1: Cross Section used in Monte Carlo production for signal and background samples. The cross section values reported involve only the semileptonic and dileptonic top decay channels. The number of QCD and dileptonic processes will be considerably reduced with the analysis cut application.

Signal	σ [pb]
$t\bar{t}$	114.49
$Z^{'} \rightarrow t\bar{t}$	3.93
Background	**σ [pb]**
single top	~ 72
multijets	$\sim 7 \cdot 10^9$
W+jets	$\sim 37 \cdot 10^3$
Z+jets	~ 33
diboson	~ 1148

at $\sqrt{s} = 8$ TeV are listed.

SINGLE TOP background arises from single top electroweak production which is about a factor of two smaller than the $t\bar{t}$ cross section. Because of the lower number of jets with respect to the top pair production, it contributes predominantly in low multiplicity events. This background is simulated using POWHEG (CT10 PDF) and PYTHIA8 (CTEQ6L1 PDF) with Perugia tune for what concerns the s- and Wt-channels, while AcerMC [59] and Pythia, with the addition of the CTEQ6L1 set of PDF used in the t-channel case.

MULTIJETS events is an important source of background in correspondence to a lepton misidentification by the detector that deceives the single lepton trigger. The objects much commonly identified as "fake leptons" are long living mesons, photons and hadronic jets. The misidentification rate is very small but due to the huge multi-jets cross section the contribution is not negligible. This source of background is usually called QCD or fake-leptons background. It is produced with PYTHIA8 using the CT10 PDF.

W+JETS events constitute the main background source for 2012 data taking analysis because of the high cross section and the signature very close to the $t\bar{t}$

one, especially in the high jet multiplicity case. W+jets background samples are produced with ALPGEN [60] and PYTHIA using CTEQ6L1 PDF set with Perugia tune.

Z+JETS events can be mismatched for $t\bar{t}$ processes in both electron and muon Z boson decays ($Z \to e^+e^-$ and $Z \to \mu^+\mu^-$), where one lepton is not detected, giving the needed fake E_T^{miss} contribution, and in the τ decay case, where one lepton decay leptonically and the other hadronically. Z+jets samples are generated using ALPGEN and PYTHIA with CTEQ6L1 PDF set with Perugia tune.

DIBOSON events $pp \to WW, WZ, ZZ$ provide a small background contribution, when the decay products have the same final state configuration as in $t\bar{t}$ events. Diboson samples are produces with HERWIG using with AUET2 tune [61] and CTEQ6L1 PDF set.

Chapter 4

Results

In this chapter the results of the analysis performed on the reconstruction of boosted top through the Template Overlap Method (TOM) in the hadronic decay channel are presented. This analysis is inserted in the measurement of the $t\bar{t}$ differential cross section in the lepton+jets channel performed on 2012 data sample collected with the ATLAS detector. At the moment, the TOM performances have been evaluated only on MC simulations, but in the next future they will be extended also to the real data.
In Section 4.1 the selection used to identify $t\bar{t}$ events is described. In Section 4.2 the study on the comparison between data and Monte Carlo is presented, in order to verify that the selection cuts and the reconstruction criteria are in agreement with data. Section 4.3 reports the TOM results obtained on the boosted top tagging and illustrates the comparison with other methods. Finally, Section 4.4 presents a study of the main systematics.

4.1 Selection Criteria

The event selection employs the object definitions presented in Section 3.1. It is applied in order reduce background contribution while keeping a high efficiency for $t\bar{t}$ events. The target event topology is illustrated in Fig.4.1 and described in further detail here.

The sample is collected using the logical OR of two electron triggers, with transverse momentum thresholds of 60 GeV lowered to 24 GeV in case of isolated electrons, and two single muon triggers, with transverse momentum thresholds of 36 GeV lowered to 24 GeV in case of isolated muons. The $t\bar{t}$ event candidates must succeed all the following criteria:

- the events must belong to the so called "good run list" (GRL) of events acquired when all the detector work properly;
- problematic events, such as containing LAr noise bursts or Tile calorimeter corrupted data, are rejected;
- at least two reconstructed primary vertexes, each with at least 4 tracks;
- exactly one good reconstructed lepton candidate per event matching with the triggered one (p_T>25 GeV and $|\eta|$ >2.5 for muons);
- longitudinal impact parameter with respect to the primary vertex having the highest $\sum p_T^2$ of constituents tracks smaller than 2 mm;
- isolated muon
- E_T^{miss} greater than 20 GeV;
- the scalar sum of E_T^{miss} and m_T^W, where $m_T^W = \sqrt{2p_T^{\mu}E_T^{miss}(1 - cos\,\Delta\phi)}$, greater than 60 GeV;
- at least one b-tagged anti-k_t R=0.4 jet, at 70% efficiency working point;
- at least one b-tagged anti-k_t R=0.4 jet per event is required to satisfy $\Delta R(jet, lepton)$ <1.5, since the leptonic top decay products are expected to be collimated and thus the b-quark is should be in the vicinity of the lepton;
- if there are more than one small-R jet with $\Delta R(jet, lepton)$ <1.5, the highest p_T one is taken as the leptonic side b-jet candidate;

- at least one large-R jet (anti-k_t R=1) candidate with transverse momentum p_T >300 GeV, reconstructed inside $|\eta|$ <2;

- the fat-jet should be tagged as top using the splitting scale method (see Section 3.4.2), in which it has been requested $\sqrt{d_{12}}$ >40 GeV and m_{top} >100 GeV;

- an isolated muon required suppress background leptons originating from jets. The variable "mini-isolation" is introduced and is defined as $I_{mini} = \sum_{track} p_T^{track}/p_T^l$, where p_T^l is the lepton transverse momentum and the sum is over all tracks (excluding the lepton track) that have $p_T > 0.4$ GeV, pass quality cuts and have $\Delta R < k_T/p_T^l$. The parameter k_T is set to 10 GeV and the isolation requirement I_{mini} <0.05 is applied for muons.

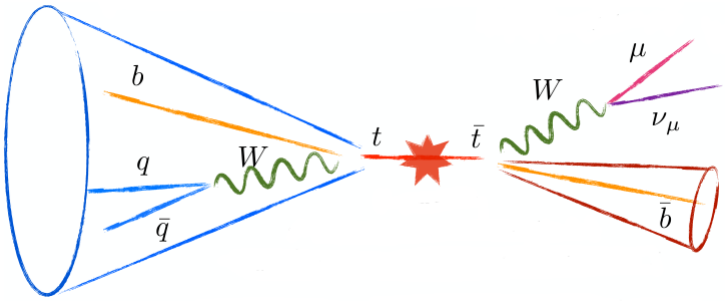


Figure 4.1: An illustration of the top pair event topology decaying to lepton+jets channel.

4.2 Data Monte Carlo Comparison

In order to check the agreement between Monte Carlo simulations and the real data, a collection of plots representing the principal quantities of physics interest have been compared in the muon and electron channels. In the following plots, real data are represented by black dots, while Monte Carlo distributions have different colours depending on the diverse physics process and have been

stacked keeping into account each cross section in order to be compared with data. The events where the generated top direction does not match with the reconstructed fat-jet have been considered as a source of background (labelled "untruth-matched" in the figures), exactly as the $t\bar{t}$ event decaying in the dilepton channel (labelled "dilepton" in the figures). Below each plot is reported the sum of all MC contributions with respect to the real data. The uncertainty band represents only statistic uncertainty.
The binning has been decided in order to fill each bin with about the same number of events; with this choice it has been possible to uniform statistical uncertainties among all the bins. In the boosted analysis a discrepancy on normalization of the order of 20% is observed between the MC and data in all the kinematic variable distributions considered. The origin of the discrepancy is under investigation: this can be due to the lack of the electroweak background corrections, or to a not perfect knowledge of the apparatus (reconstruction program, identification algorithm, trigger efficiency, ...) or also to some unknown processes, that could be index of new physics. Fig.4.2 shows the comparison between MC and data samples in transverse momentum p_t, pseudorapidity η and azimuthal angle ϕ kinematic distributions of the hadronic top in the muon channel. As anticipated, the disagreement between data and MC is around 20%, well visible in the plots of the η and ϕ variables and it increases with the rise of the top transverse momentum (see low part of the first plot in Fig.4.2).
In Fig.4.3 are represented the same kinematic quatities of the muon; the comparison between data and MC reflects exactly the conclusion obtained for the hadronic top distributions (Fig.4.2).

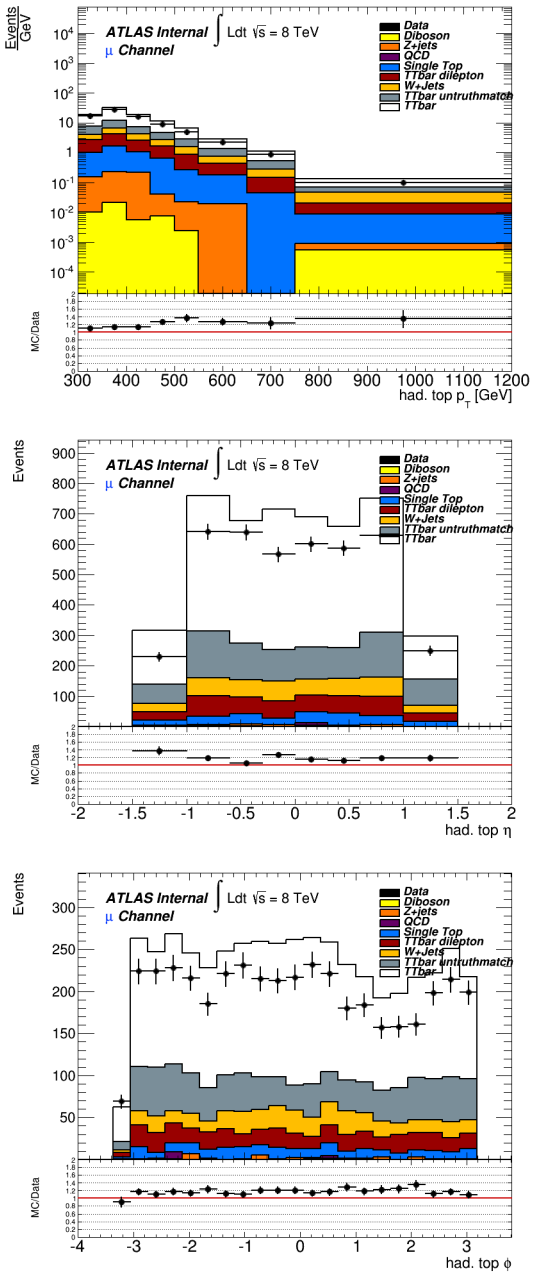


Figure 4.2: Comparison between Monte Carlo and data distribution in the analysis of the p_T, η and ϕ kinematic quantities of the hadronic top quark in the muon channel. Real data are represented with black dots, while Monte Carlo samples have different colours on the basis of their origin. Diboson, Z+jets, QCD, Single top, $t\bar{t}$ dilepton, W+jets, $t\bar{t}$ untruth-matched background samples have been considered. The same legend has been used in all the following plots.

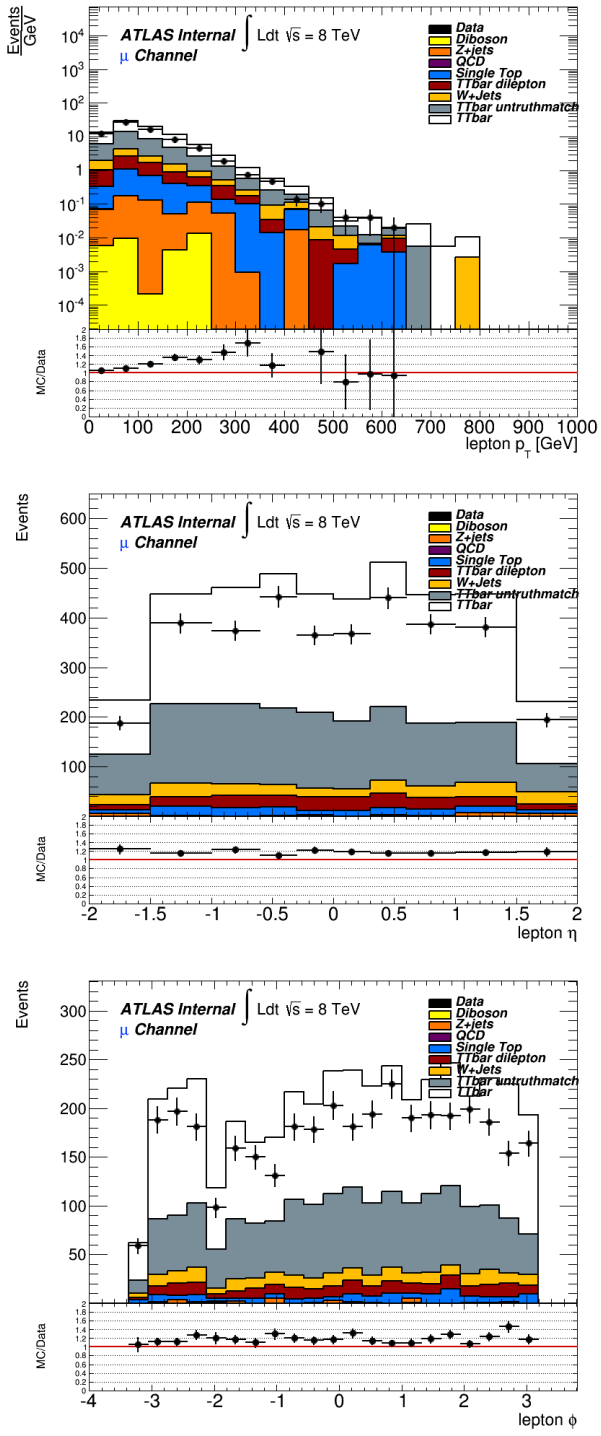


Figure 4.3: Comparison between Monte Carlo and data distribution in the analysis of the p_T, η and ϕ kinematic quantities of the muon.

To test the availability of the TOM method, in Fig.4.4 the real data MC comparison of the overlap variable has been represented, showing the same behaviour as the previous plots.

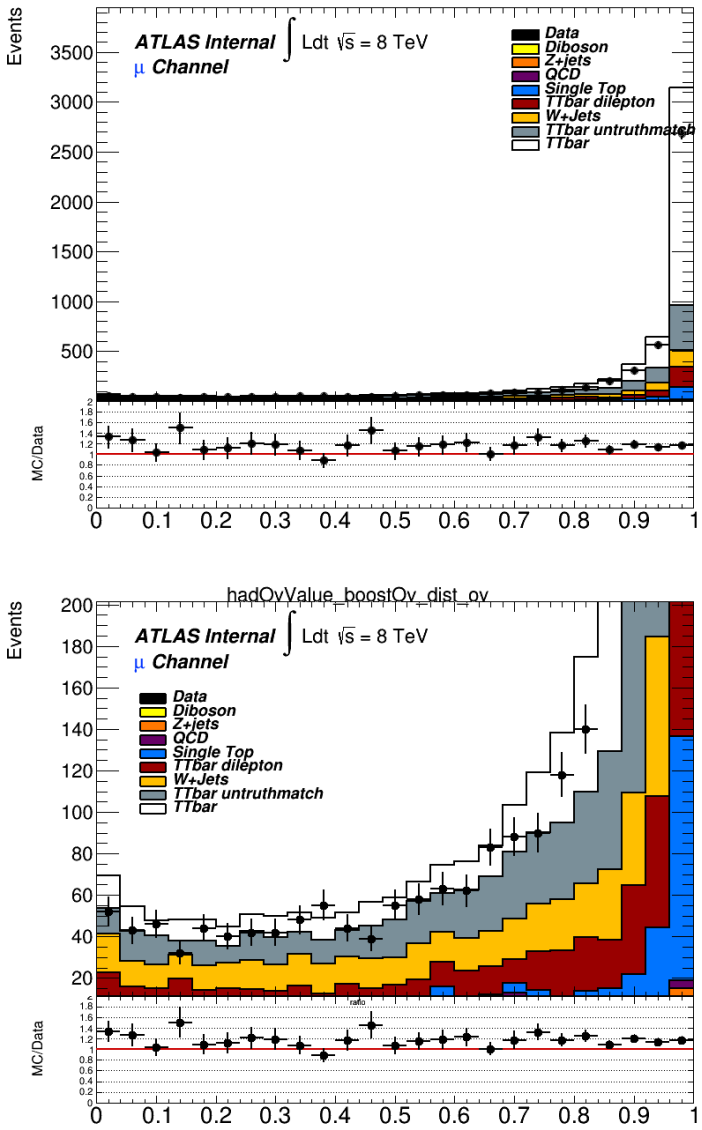


Figure 4.4: Comparison between Monte Carlo and data in the analysis of the overlap distribution in the muon channel. On the bottom there is an expansion of the lower part of the plot.

Other important quantities to be studied are the number of jets tagged as coming from a b quark and the E_T^{miss} (see Fig.4.5). Both quantities are fundamental in the event selection where the presence of at least one b-jet and a big amount of E_T^{miss}, revealing the presence of a neutrino, is explicitly requested.

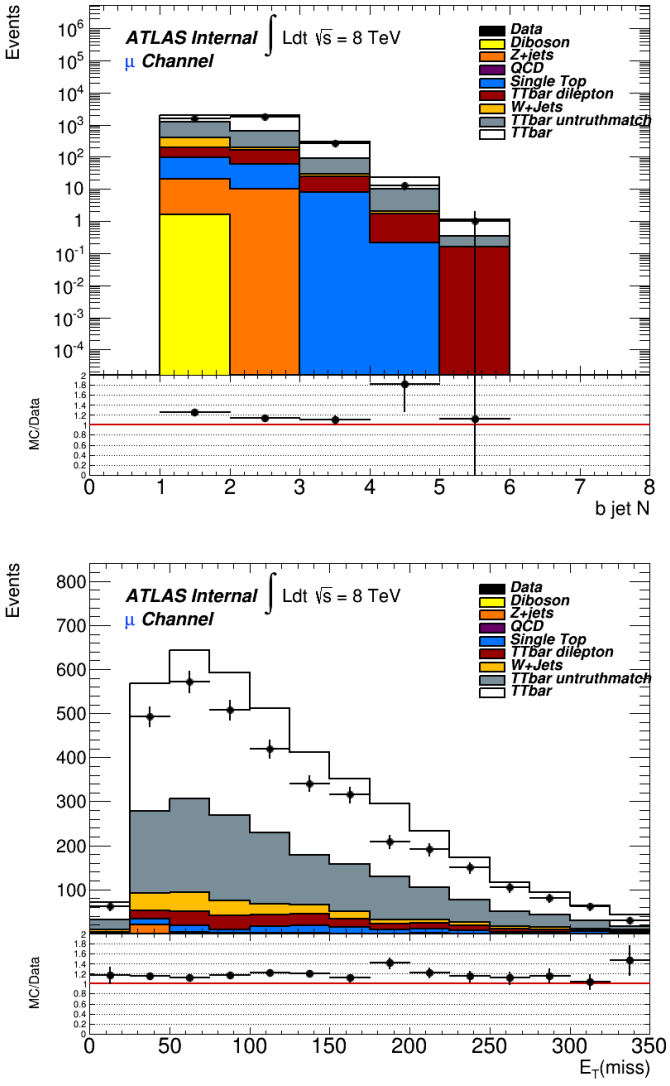


Figure 4.5: Comparison between Monte Carlo and data distribution in the analysis of the number of b-jets and E_t^{miss} in the muon channel.

Tab.4.1 summarises the different Monte Carlo contributions and the comparisons with data. These events are used for the analysis of the $t\bar{t}$ production differential cross section. As anticipated in the selection cuts, the request to tag as top the fat-jet using the splitting scale method has been included; for the next data taking, this request will be replaced by a multivariate analysis including all the other top tagging techniques.

Table 4.1: Summary of Monte Carlo and data number of event calculated in respect to the p_T distribution of the hadronic top with an overlap value greater than 0.7 in the muon channel.

Sample	Number of events	MC*100/data
data	4145	100
diboson	2.59567	0.0
Z+jets	33.2187	0.8
multijets	0.000595256	0.0
single top	222.606	5.4
$t\bar{t}$ dilepton	405.999	9.8
W+jets	409.055	9.9
$t\bar{t}$ untruth-matched	884.64	21.3
$t\bar{t}$	2895.16	69.8
Total MC	4853.275	117.09

The same comparisons have been performed also for the electron channel and the conclusions reflect the previous ones obtained for the muon channel. As an example, in Fig.4.6 the kinematic variable distributions for the electron are reported and in Fig.4.7 the overlap value distribution is presented.

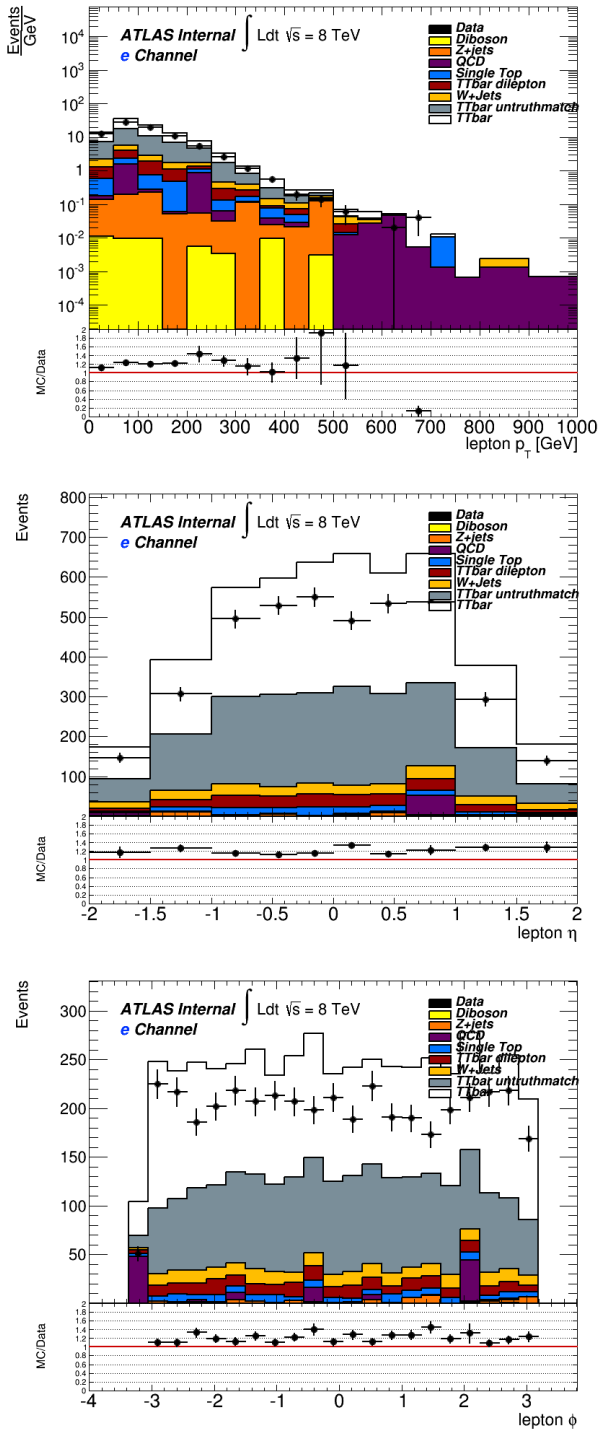


Figure 4.6: Comparison between Monte Carlo and data distribution in the analysis of the p_T, η and ϕ kinematic quantities of the electron.

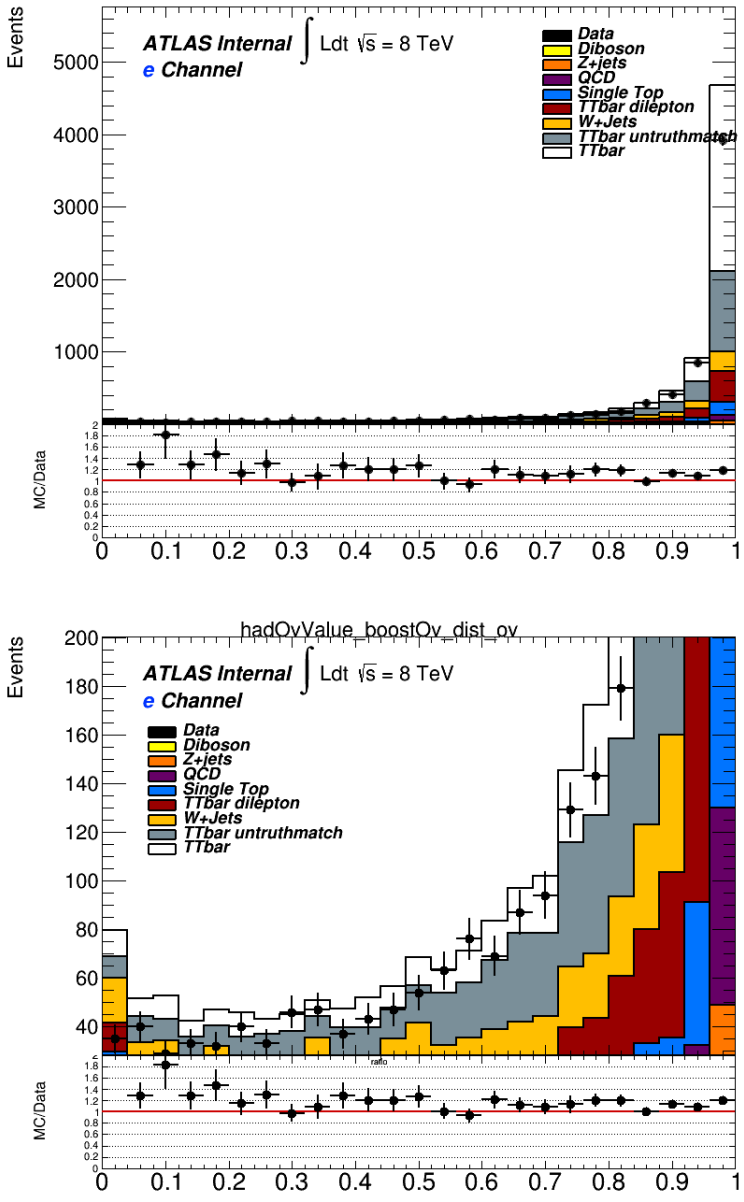


Figure 4.7: Comparison between Monte Carlo and data in the analysis of the overlap distribution in the electron channel. On the bottom there is an expansion of the lower part of the plot.

4.3 TOM Results

The events used to test the TOM method have passed the selection described in Section 4.1, with the exception of the last cut on the top tagging with the splitting scale technique. As a first step, the TOM algorithm has been applied to fat-jet where the grooming procedure was not employed; this choice is motivated by the necessity to distinguish the effect of TOM with respect to the grooming one. Obviously, after the application of the grooming method, better performances for the TOM techniques are expected.
At the moment, TOM performances have been evaluated on Monte Carlo, using the $t\bar{t}$ sample for the signal and the QCD as background. The reason to limit background only to QCD channel is motivated to test how TOM rejects jets mainly coming from gluons and light quarks. In the next future, TOM performances will be extended to all background contributions and also to real data.
The overlap distributions are represented in Fig.4.8 for $t\bar{t}$ and QCD events, respectively; as predicted the overlap distribution is closer to 1 for signal with respect to background, as confirmed also by the mean value of the histograms. The enhancement of the overlap distribution in the region close to 1 in the background sample is caused by an irreducible contribution due to those events which pass all the selection cuts and present a jet with a completely similar structure as coming from a top decay.
In Fig.4.9 the signal efficiency and the background rejection (defined in this case as $1 - \epsilon$) are plotted as a function of the overlap value. These plots have been used to determine the minimum overlap value to define when a fat-jet can be tagged as coming from a top. Since the selected cuts give a good signal background ratio, this value has been chosen to allow high efficiency on the signal and a not very high background rejection. At the end, the top tagging has been defined if the fat-jet gives an overlap value greater than 0.7, with at least one top generated in the template sample. This choice conduce to have a signal efficiency better than 90% and a background rejection around 70%.

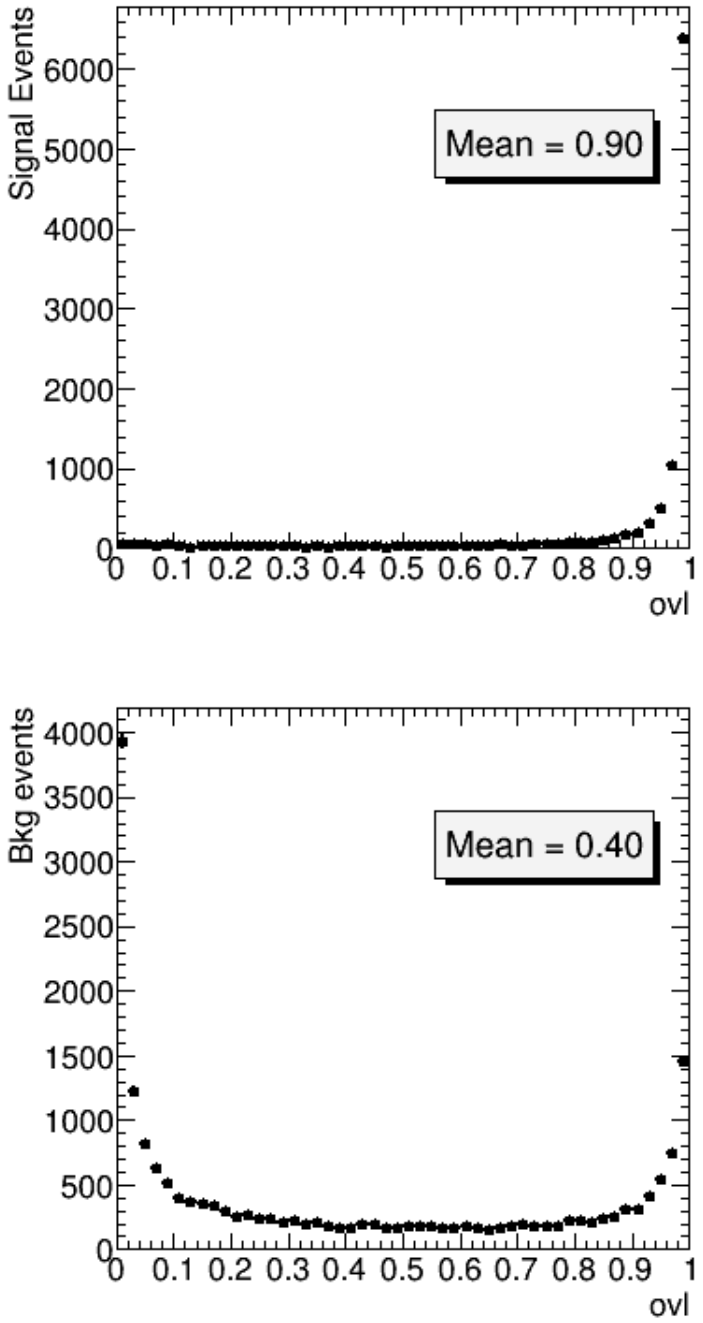


Figure 4.8: Overlap distribution for Monte Carlo $t\bar{t}$ signal (*top*) and QCD background (*bottom*).

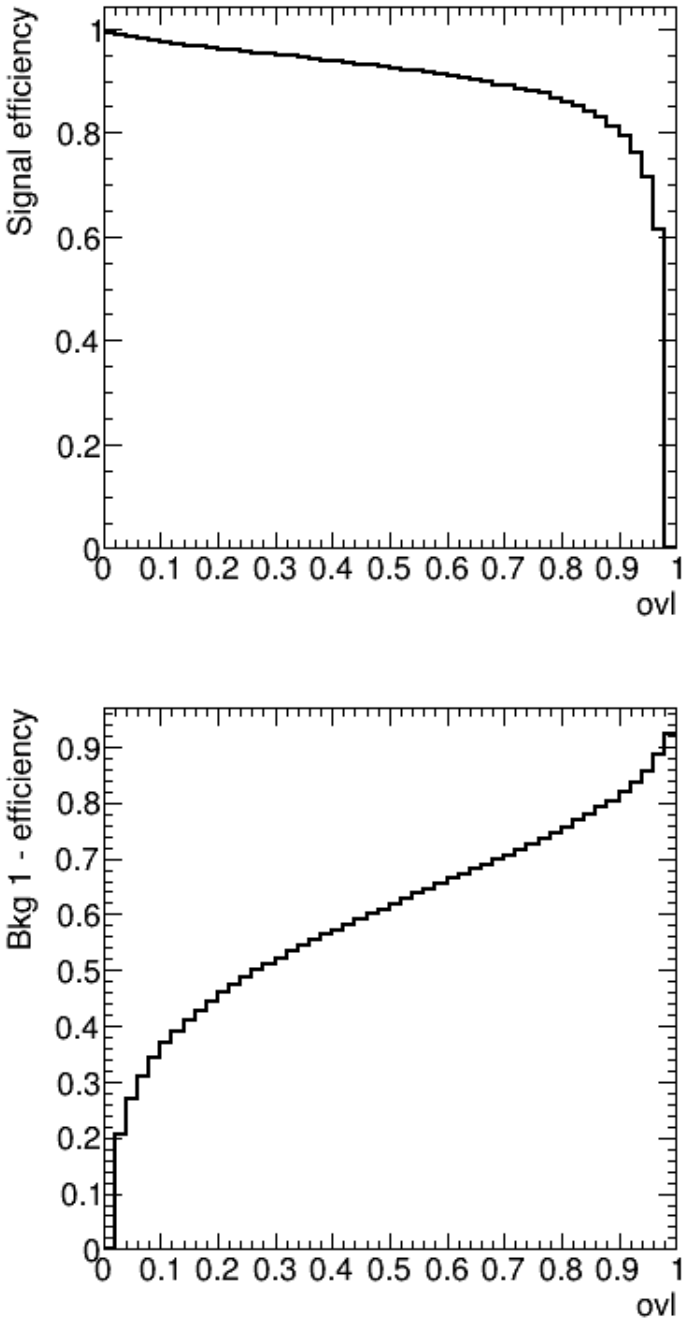


Figure 4.9: Distributions of $t\bar{t}$ signal (*top*) efficiency and QCD background rejection (*bottom*) as a function of the overlap.

The performances of all the methods presented in Section 3.4 have been considered in Fig.4.10 [62]. This illustration correlates the efficiency on the signal (x axis) with the background rejection (y axis), defined as $R = 1/\epsilon$ (different definition with respect to 4.9). As for the TOM procedure, these performances have been obtained on MC samples using direct $t\bar{t}$ production for signal and only QCD as background contribution. In order to compare TOM performances with the ones in literature, the same plot has been produced (see Fig.4.11 and Fig.4.12 for the comparison). TOM gives similar results with respect to the others techniques and this is a very interesting starting point considering that this procedure is innovative and never used before. In the next future all these methods will be merged in a multivariate analysis in order to increase the performance of each single method.

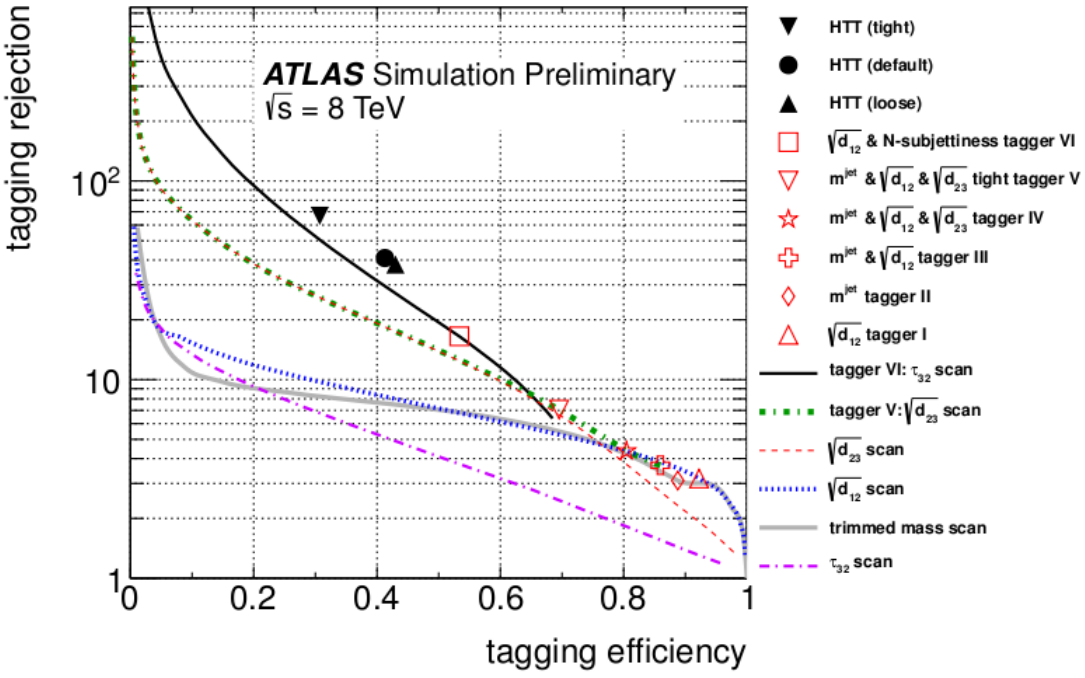


Figure 4.10: Comparison of the simulated fat-jet tagging efficiency and fat-jet light quark/gluon rejection [62].

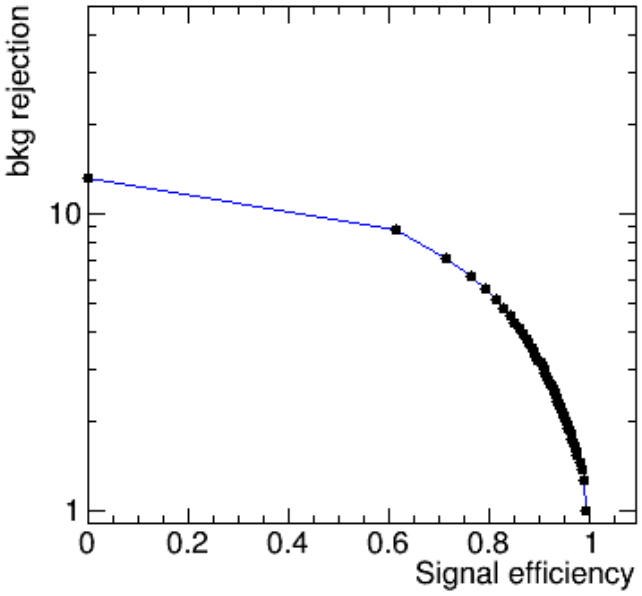


Figure 4.11: Distribution of the QCD rejection as a function of the $t\bar{t}$ efficiency. The rejection is represented in a logarithmic scale.

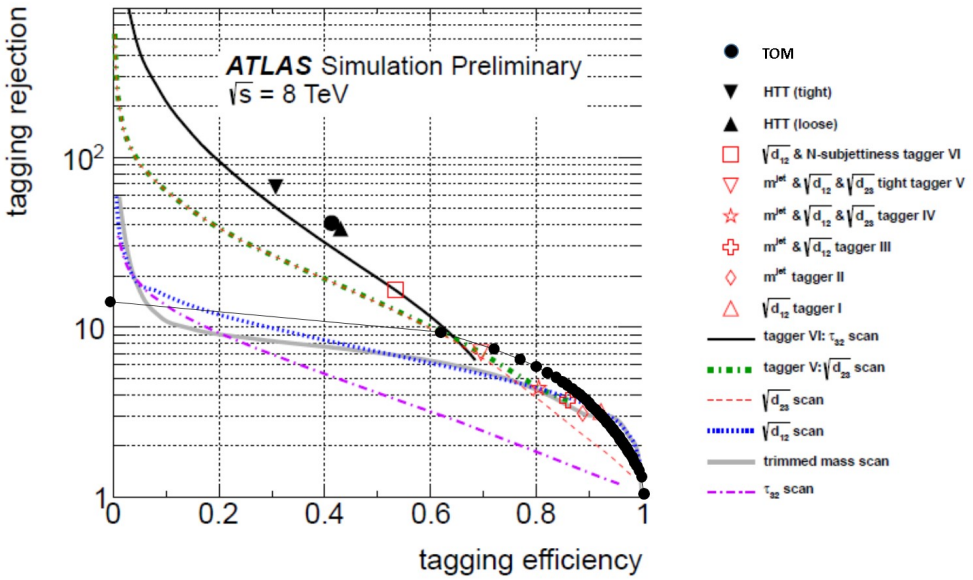


Figure 4.12: Comparison between the TOM performances with the other top tagging procedures presented in literature [62].

A crucial feature of the TOM is the improvement of the efficiency with respect the transverse momentum of the fat-jet (see Fig.4.13); this will be fundamental in the next data taking where the center-of-mass energy will increase to $\sqrt{s}$ =13 TeV and consequently also the top quark will have higher momentum.

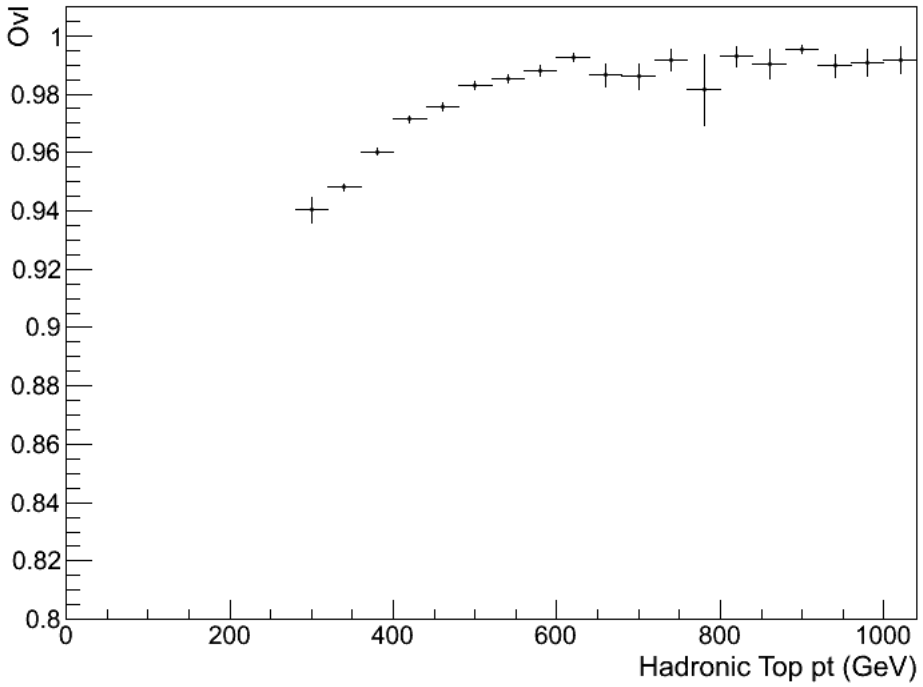


Figure 4.13: Overlap distribution as a function of the hadronic top momentum.

With the rise of center-of-mass energy designed for the LHC Run 2, the proton-proton cross section will increase and thus the average number of interactions for bunch crossing (pile-up). It becomes crucial to verify that TOM will not depend on the increase of the number of interactions, as shown in Fig.4.14, where the efficiency is stable. All these results make us confident that TOM could be an important method for the top tagging of the next data taking.

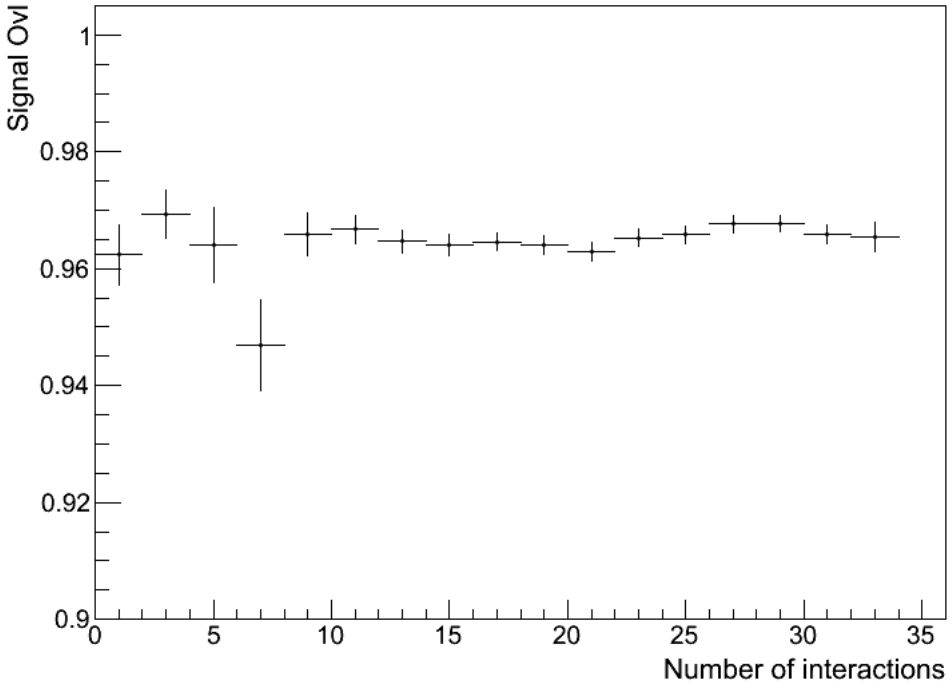


Figure 4.14: Overlap distribution as a function of pile-up for $t\bar{t}$ signal events.

4.4 TOM Systematics

The reproducibility of the results has been tested with a systematic study obtained varying all the parameters of the TOM procedure. In particular the TOM performances have been determined using:

- a different algorithm for the top template rotation over the fat-jet;
- the energy deposits of the cluster and of the template instead of the p_T;
- templates with momentum distribution equal to the top production;
- template with fixed momentum and top direction with uniform distribution in the η-ϕ plane;
- different values for the sigma parameter of Eq.(3.14);

- different values for the Sub Cone Radius;

- variable Sub Cone Radius depending on the energy of the cluster;

- the parton of the template with energy close to the one of the cluster.

The efficiency of the top tagging with the TOM method for each systematic is reported in Fig.4.15; the average value is (0.86 ± 0.12).

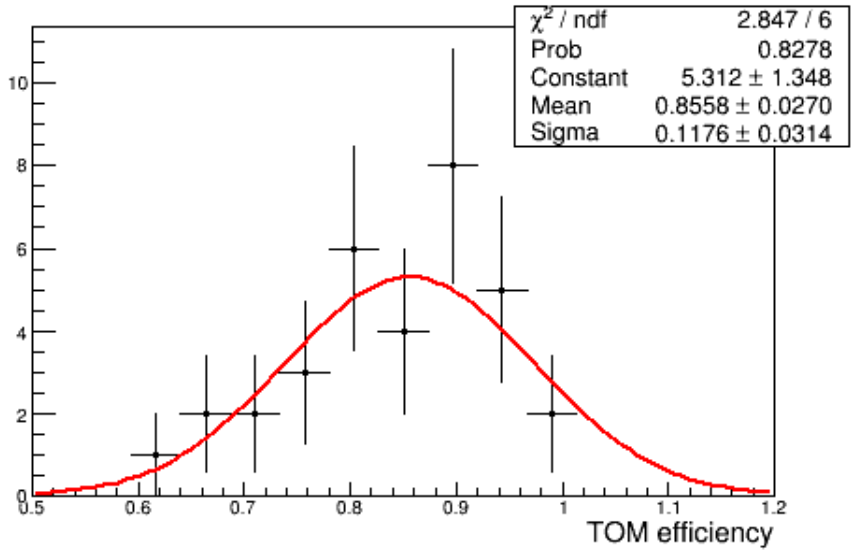


Figure 4.15: Distribution of the efficiency average value for each systematic.

Conclusions

The analysis presented in this thesis is focused on the study of a new method (Template Overlap Method) to reconstruct high momentum p_T >300 GeV top quark decaying in the hadronic channel. The decay products of these top quarks (usually referred to as "boosted") are emitted in a narrow region in the η-ϕ space, causing a partial or total overlap of the corresponding jets. The decay products are fully contained in a single large radius jet (fat-jet), which obviously absorbs many background contributions due to pile-up and soft QCD emissions.

The Template Overlap Method compares the energy distribution of the fat-jet with the one belonging to a sample of simulated top in which it has not taken into account the hadronization. The procedure is based on the definition of an overlap function which quantifies the similarity between the energy distribution of the fat-jet and the one of the template. The method has been optimised in order to obtain a signal efficiency of about 90% together with a background rejection at the level of 70%. At present, the analysis has been applied only to MC samples. To verify the effectiveness of the all analysis chain, a comparison between data and MC samples has been performed; it results a discrepancy of about 20% (present also in other analysis) that is under investigation. The obtained results are similar to the ones still present in literature and in the next future they will be merged in a multivariate analysis in order to provide a better top identification. All the developments reached in this analysis will be included in the measurement of the $t\bar{t}$ production differential cross section performed on the data acquired in 2012 at $\sqrt{s}$=8 TeV.

Important feature of the Template Overlap Method is that the efficiency is increasing with p_T of the fat-jet; this will be crucial in the next data taking at $\sqrt{s}$=13 TeV, where almost the totality of tops are produced with very high p_T and the standard reconstruction method will become totally inefficient.

Bibliography

[1] Donald H., *Perkins Introduction to High Energy Physics*, Addison-Wesley, 1986.

[2] K.A. Olive and al. *Particle Data Group*, Chin. Phys. C, 38, 090001, 2014.

[3] A. Bettini, *Elementary Particle Physics*, Cambridge, 2008.

[4] C. Quigg, *Gauge Theories of the Strong, Weak and Electromagnetic Interactions*, Westview, 1997.

[5] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys.Rev.Lett., 10:531–533, 1963.

[6] A. Salam, *Weak and Electromagnetic Interactions*, Conf.Proc.,C680519, 1968.

[7] The ATLAS Collaboration, *Observation of a New Particle in the Search for the Standard Model Higgs Boson with the ATLAS Detector at the LHC*, CERN-PH-EP-2012-218, 2012.

[8] M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*, Perseus Books, 1995.

[9] D. Gross and F. Wilezek, *Soft Strong Interaction - the asymptotic freedom of quarks*, Phys. Rev. Lett., B30 1343, 1973.

[10] CDF Collaboration, *Observation of Top Quark Production in pbar-p Collisions*, Phys. Rev. Lett. 74, 2626, 1995.

[11] D0 Collaboration, *Observation of the Top Quark*, Phys. Rev. Lett. 74, 2632, 1995.

[12] CDF and D0 collaborations, *Combination of CDF and D0 results on the mass of the top quark using up to 5.8 fb^{-1} of data*, Technical Report, FERMILAB-TM-2504-E, 2011.

[13] F. Abe et al., *Observation of the Top Quark Production in $p\bar{p}$ Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. 74:2626-2631, 1995.

[14] S.Moch and P.Uwer, *Theoretical status and prospects for top-quark pair production at hadron colliders*, Phys. Rev. D78, 034003, 2008.

[15] The ATLAS Collaboration, *A combined measurement of the top quark pair production using dilepton and single-lepton final states*, Tech. Rep. ATLAS-CONF-2011-040, 2011.

[16] F. P. Schilling, *Top Quark Physics at LHC: a review of the first two years*, Int. Jour. Mod. Phys., 2012.

[17] M. Jezabek and J. H. Kuhn, *Electroweak Corrections to the Top Quark Decay*, Nucl. Phys. B314, 1, 1989.

[18] The ATLAS collaboration, *Measurements of top quark pair relative differential cross sections with ATLAS in pp collisions at $\sqrt{s}$ = 7 TeV*, CERN-PH-EP-2012-165, 2012.

[19] CERN, *LHC Design Report Volume I: The LHC Main Ring*, CERN-2004-003-V-1, 2004.

[20] The ATLAS Collaboration, *The ATLAS Experiments at the CERN Large Hadron Collider*, JINST 3:S08003, 2008.

[21] The ATLAS Collaboration, *ATLAS Detector and Physics Performance Technical Design Report*, CERN/LHCC, Volume I-II, 1999.

[22] The ATLAS Collaboration, *ATLAS Pixel Detector Technical Report*, CERN/LHCC, 1998.

[23] J. N. Jackson, *The ATLAS Semiconductor Tracker (SCT)*, Nucl. Instrum. Meth. A541, 2005.

[24] V. A. Mitsou, *The ATLAS Transition Radiation Tracker*, ATL-CONF-2003-012., 2003.

[25] The ATLAS Collaboration, *ATLAS Calorimeter Performance Technical*, CERN/LHCC, 1997.

[26] The ATLAS Collaboration, *ATLAS Muon Spectometer Technical Design Report*, CERN/LHCC, 1997.

[27] G. Aad et al., *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics* CERN, 2009.

[28] The ATLAS Collaboration, ATLAS, High Level Trigger, Data Acquisition and Controls, CERN/LHCC/2003-022, 2003.

[29] The ATLAS collaboration, *Updated Luminosity Determination in pp collision at* $\sqrt{s}$ *=7 TeV using the ATLAS Detector*, ATLAS-CONF-2011-011, 2011.

[30] The ATLAS Collaboration, *Measurements of normalized differential cross sections for ttbar production in pp collisions at* $\sqrt{s}$ *= 7 TeV using the ATLAS detector*, Phys. Rev. D, 2014.

[31] G. P. Salam, *Towards Jetography*, JHEP 0906-1833, 2010.

[32] S. D. Ellis and D. E. Soper, *Successive Combination Jet Algorithm For Hadron Collisions*, Phys. Rev. D 48, 3160, 1993.

[33] Y. L. Dokshitzer, G. D. Leder, S. Moretti and B. R. Webber, *Better jet clustering algorithms*, JHEP 9708, 001, 1997.

[34] M. Cacciari, G. P. Salam and G. Soyez, *The anti-kt jet clustering algorithm*, HEP 0804 063, 2008.

[35] The ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, CERN-PH-EP-2013-05, 2013.

[36] T. Lagouri et al., *IEEE Transactions on Nuclear Science*, 51, 3030, 2004.

[37] S. Hassani et al., *Nuclear Instruments and Methods in Physics Research*, A 572-77, 2007.

[38] The ATLAS Collaboration, *Pile-up Depedence of the ATLAS Muon Performance*, ATL-COM-PHYS-2011-1640, 2011.

[39] The ATLAS Collaboration, *Performance of Missing Transverse Momentum Reconstruction in ATLAS with 2011 Proton-Proton Collisions at* $\sqrt{s}$*= 7 TeV*, ATLAS-CONF-2012-101, 2012.

[40] E. B. Kuutmann, *Physics with boosted top quarks*, ATL-PHYS-PROC-2014-108, 2014.

[41] The ATLAS Collaboration, *Performance of large-R jets and jet substructure reconstruction with the ATLAS detector*, ATLAS-CONF-2012-065, 2012.

[42] J. M. Butterworth et al., *Jet substructure as a new Higgs search channel at the LHC*, Phys. Rev. Lett. 100 242001, 2008.

[43] D. Krohn, J. Thaler, and L. T. Wang, *Jet trimming*, JHEP 2010 20, 2010.

[44] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, *Recombination Algorithms and Jet Substructure: Pruning as a Tool for Heavy Particle Searches*, Phys. Rev. D81 094023, 2010.

[45] J. Thaler and K. Van Tilburg, *Maximizing Boosted Top Identification by Minimizing N-subjettiness*, JHEP 1202 093, 2012.

[46] T. Plehn, M. Spannowsky, M. Takeuchi, and D. Zerwas, *Stop Reconstruction with Tagged Tops*, JHEP 1010 078, 2010.

[47] L. G. Almeida, S. J. Lee, G. Perez, G. Sterman and I. Sung, *Template Overlap Method for Massive Jets*, YITP-SB-10-21, 2010.

[48] M. Backovic, J. Junknevich and G. Perez, *Boosting the Standard Model Higgs with the Template Overlap Method*, CERN-PH-TH-2012-349, 2013.

[49] The ATLAS Collaboration *Luminosity Determination in pp Collisions at $\sqrt{s}$ = 7 TeV using the ATLAS Detector in 2011*, Technical Report ATLAS-CONF-2011-116, 2011.

[50] GEANT4 Collaboration, S. Agostinelli et al., *GEANT4: A simulation toolkit*, Nucl. Instrum. Meth. A506 250, 2003.

[51] The ATLAS Collaboration, *Measurement of the differential cross section of highly boosted top quarks as a function of their transverse momentum using the ATLAS detector in $\sqrt{s}$=8 TeV proton-proton collisions*, ATLAS-CONF-2014-057, 2014.

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

[53] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, et al., *New parton distributions for collider physics*, Phys. Rev. D82 074024, 2010.

[54] T. Sjostrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP 0605 026, 2006.

[55] J. Pumplin, D. Stump, J. Huston, H. Lai, P. M. Nadolsky, et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP 0207 012, 2002.

[56] P. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, Phys. Rev. D82 074018, 2010.

[57] The ATLAS Collaboration, *Further ATLAS tunes of PYTHIA6 and Pythia8*, ATL-PHYS-PUB-2011-014, 2011.

[58] A.D. Martin, W.J. Stirling, R.S. Thorne and G. Watt, *Parton distributions for the LHC*, Eur. Phys. J. C63, 2009.

[59] B. P. Kersevan and E. Richter-Was, *The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5*, TPJU-6, 2004.

[60] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP 07 001, 2003.

[61] G. Corcella et al., *HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes)*, JHEP 01 010, 2010.

[62] The ATLAS Collaboration, *Performance of boosted top quark identification in 2012 ATLAS data*, ATLAS-CONF-2013-084, 2013.