

A search for massive resonances decaying to top quark pairs and jet trigger performance studies with the ATLAS detector at the Large Hadron Collider

DISSERTATION

zur Erlangung des akademischen Grades

Dr. rer. nat.
im Fach Physik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät I
Humboldt-Universität zu Berlin

von
M.Sc. Luz Stella Gomez Fajardo

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Jan-Hendrik Olbert

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät I:
Prof. Stefan Hecht, Ph.D.

Gutachter:

1. Prof. Dr. Klaus Mönig
2. Prof. Dr. Jörgen Sjölin
3. Prof. Dr. Thomas Lohse

Tag der mündlichen Prüfung: 8.05.2014

Abstract

This thesis presents the search for new particles that decay into top quark pairs ($t\bar{t}$). The analysis is performed with the ATLAS experiment at the LHC, using an integrated luminosity of 2.05 fb^{-1} of proton–proton collision data, collected at a center-of-mass energy of $\sqrt{s} = 7 \text{ TeV}$. The lepton plus jets final state is used in the $t\bar{t} \rightarrow WbWb$ decay, where one W boson decays leptonically and the other hadronically. The $t\bar{t}$ system is reconstructed using both resolved and boosted topologies of the top-quark decay. For the first time, correlations between the two search channels have been employed by creating a third channel with the events selected by both analyses. The sensitivity to new physics phenomena is thereby improved. Upper limits are derived on the production cross-section times branching ratio for narrow and wide massive states, at the 95 % confidence level. These are extracted by combining the two approaches of the $t\bar{t}$ reconstruction. For a narrow Z' resonance, the observed (expected) upper limits range from 4.85 (4.81) pb for a mass of 0.6 TeV, to 0.21 (0.13) pb for a mass of 2 TeV. A narrow leptophobic topcolor Z' resonance with a mass below 1.3 TeV is excluded. Observed (expected) limits are also derived for a broad color-octet resonance. They vary between 2.52 (2.59) pb and 0.37 (0.27) pb for a mass of 0.7 TeV and 2 TeV, respectively. The wide Kaluza-Klein gluon with a mass below 1.65 TeV is excluded.

Another aspect of this thesis are performance studies of the level-1 jet trigger. Trigger efficiencies have been measured, using data collected by the ATLAS detector in 2010 at $\sqrt{s} = 7 \text{ TeV}$. The turn-on curves obtained for a variety of jet triggers, showed good agreement between data and simulation in the plateau region. The efficiency results were used at the first stage of analyses for multi-jet cross-section measurements.

Zusammenfassung

Diese Arbeit behandelt die Suche nach neuen Teilchen, die in Top-Quark-Paare zerfallen ($t\bar{t}$). Die Analyse beruht auf Daten des ATLAS-Experiments von Proton-Proton-Kollisionen am LHC bei einer Schwerpunktsenergie von $\sqrt{s} = 7$ TeV und einer Gesamtluminosität von 2.05 fb^{-1} . Hierzu wird der Lepton+Jets Endzustand im $t\bar{t} \rightarrow WbWb$ Zerfallskanal verwendet, worin ein W -Boson leptonisch und das andere hadronisch zerfällt. Das $t\bar{t}$ -Ereignis wird sowohl in aufgelösten als auch geboosteten Zerfallstopologien rekonstruiert. Zum ersten Mal werden die Korrelationen beider Kanäle in Form einer dritten Kategorie nutzbar gemacht, welche aus Ereignissen besteht, die in beiden Topologien selektiert wurden. Die Sensitivität der Analyse wird hierdurch erhöht.

Obere Schranken bei 95% Vertrauensniveau auf den Wirkungsquerschnitt multipliziert mit der Zerfallsbreite für massive Zustände großer und kleiner Zerfallsbreite werden berechnet. Diese werden aus der Kombination der beiden Ansätze der $t\bar{t}$ -Rekonstruktion gewonnen. Für die Z' -Resonanz kleiner Breite reicht die beobachtete (erwartete) obere Grenze auf den Wirkungsquerschnitt von 4.85 (4.81) pb, für eine Masse von 0.6 TeV, bis 0.21 (0.13) pb, für eine Masse von 2 TeV. Eine schmale leptophobic Topcolor- Z' -Resonanzen mit einer Masse unterhalb von 1.3 TeV kann ausgeschlossen werden. Weiterhin konnten beobachtete (erwartete) obere Grenzen auch für eine breite Farboktett-Resonanz berechnet werden. Diese liegen zwischen 2.52 (2.59) pb und 0.37 (0.27) pb für Massen von 0.7 TeV bzw. 2 TeV. Breite Kaluza-Klein-Gluon-Resonanzen mit einer Masse unter 1.65 TeV können ausgeschlossen werden.

Ein weiterer Teil dieser Arbeit untersucht die Leistung des Level-1-Jet-Trigger. Die Trigger-Effizienzen wurden mit Daten bestimmt, die vom Atlas-Detektor bei einer Schwerpunktsenergie von $\sqrt{s} = 7$ TeV im Jahr 2010 gemessen wurden. Die Effizienzkurven für verschiedene Jet-Trigger zeigen eine gute Übereinstimmung zwischen Daten und Simulation in der Plateau-Region. Die Effizienzen wurden für erste Analysen von Messungen der Multi-Jet Wirkungsquerschnitte verwendet.

Contents

1	Introduction	1
2	Theoretical overview of the physics at the LHC	3
2.1	Introduction to the Standard Model	3
2.1.1	Electroweak theory	5
2.1.2	The Strong Interaction	8
2.2	Development of a proton-proton collision	9
2.3	The top quark in the Standard Model	10
2.3.1	Top quark pair production at the LHC	11
2.3.2	Single top production	12
2.3.3	Top quark decay channels	13
2.4	Searches beyond the Standard Model	14
2.4.1	The top quark beyond the Standard Model	15
2.5	Modeling a proton-proton collision	18
2.5.1	Simulation of the hard subprocess	20
2.5.2	Hadronization and decays	21
2.5.3	Underlying and pile-up events	22
2.5.4	Monte Carlo generators	22
3	The ATLAS Detector and simulation	25
3.1	The Large Hadron Collider	25
3.2	The ATLAS Detector	27
3.2.1	The ATLAS Coordinate System	28
3.2.2	Magnet system	29
3.2.3	Inner detector	30
3.2.4	Calorimeter system	32
3.2.5	Muon System	36
3.2.6	Forward detectors	37
3.3	Experiment performance	38
3.3.1	Pile-up	40
3.4	Trigger and data acquisition	40
3.4.1	Level-1 trigger	41
3.4.2	Level-2 Trigger	42
3.4.3	Event Filter	43
3.4.4	Trigger Menu	43
3.5	Data Quality	44

3.6	ATLAS simulation and data management	44
3.6.1	The LAr failure	45
3.6.2	Pile-up re-weighting	46
4	Object reconstruction and identification	47
4.1	Track reconstruction	47
4.2	Electrons	48
4.2.1	Reconstruction, identification and isolation	48
4.2.2	Electron trigger and resolution	51
4.3	Muons	51
4.3.1	Reconstruction, identification and isolation	51
4.3.2	Muon trigger and resolution	54
4.4	Jets	54
4.4.1	Reconstruction and identification	55
4.4.2	Jet calibration	57
4.4.3	Jet energy resolution	59
4.4.4	Jet selection and efficiency	60
4.5	b-jets	60
4.5.1	b-jets identification	61
4.5.2	b-tagging efficiency and scale factors	62
4.6	Missing transverse momentum	62
4.6.1	Neutrino reconstruction	64
5	Level-1 jet trigger efficiencies	67
5.1	Level-1 calorimeter trigger	67
5.2	Jet triggers in the 2010 physics menu	69
5.3	Jet trigger efficiency	70
5.3.1	Estimation methods	71
5.4	Results	72
5.4.1	Single-jet trigger efficiencies	72
5.4.2	Multi-jet trigger efficiencies	74
5.4.3	Total jet E_T trigger efficiencies	79
5.5	Conclusions	80
6	Top-quark reconstruction in $t\bar{t}$ resonances searches	81
6.1	Analysis strategy	81
6.2	Simulation of signal and background processes	82
6.2.1	Background processes	83
6.2.2	Benchmark signal processes	85
6.3	Data sample	85
6.4	Selection of events in $t\bar{t}$ topologies	86
6.4.1	Jets and event topology in the resolved reconstruction	88
6.4.2	Jets and event topology in the boosted reconstruction	90
6.4.3	Selection efficiencies and overlap channel	92

6.5	Reconstruction of the $t\bar{t}$ invariant mass	94
6.5.1	The $t\bar{t}$ system in the resolved approach	94
6.5.2	The $t\bar{t}$ system in the boosted approach	97
6.6	Data-driven estimation of background processes	99
6.6.1	W +jets background	99
6.6.2	QCD multi-jet background	102
6.7	Systematic uncertainties	105
6.7.1	Event level effects	105
6.7.2	Systematic uncertainties related to the physics objects	106
6.7.3	Systematic effects on the minor backgrounds	112
6.7.4	Systematic uncertainties affecting the SM $t\bar{t}$ background only . . .	112
6.7.5	Shape systematics affecting the W +jets background only	113
6.7.6	Impact of the systematic uncertainties on the analysis	114
6.8	Observed data and Standard Model background prediction	114
6.8.1	Observed and predicted $m_{t\bar{t}}$ distribution	120
6.9	Summary and conclusions	121
7	Search for heavy $t\bar{t}$ resonances in ATLAS	123
7.1	Existing limits	123
7.2	Determination of upper limits on the signal cross-section	124
7.3	Inputs to the limits setting procedure	127
7.4	Results	128
7.4.1	Combination I: Inclusive boosted	130
7.4.2	Combination II: Exclusive boosted + Overlap $m_{t\bar{t}}^{\text{boosted}}$	132
7.4.3	Combination III: Exclusive boosted + Overlap $m_{t\bar{t}}^{\text{boosted}}$ + Exclu- sive resolved	133
7.4.4	Combination IV: Exclusive boosted + Overlap $m_{t\bar{t}}^{\text{resolved}}$ + Exclu- sive resolved	134
7.5	Comparison of the sensitivity for the different combinations	135
7.6	Comparisons to other $t\bar{t}$ resonances searches	138
7.7	Summary and conclusions	139
8	Summary and conclusions	141

1 Introduction

Colliding particles at very high energies is the most prolific access for understanding the fundamental building blocks of matter. The Large Hadron Collider (LHC) located at CERN¹, produces the world's most powerful high-energy beams of protons to collide at four main experiments. ATLAS (A Toroidal LHC Apparatus) is one of the two general purpose experiments, and is designed to investigate the fundamental interactions within and outside of the Standard Model framework.

The vast majority of the collision events contain well-known physics. To collect a considerable number of interesting events with new processes, very high collision rates are needed. This produces an enormous data output, which exceeds any storage capacity. The ATLAS trigger system is dedicated to reduce the event rates, by deciding in real-time whether an event is interesting and thus stored, or it is discarded and lost forever. A part of this thesis aims to quantify the selection quality of the ATLAS trigger system.

The theoretical framework of elementary particle physics is the Standard Model (SM). It is an effective quantum field theory, which encompasses all known fundamental particles and their interactions. There are twelve fundamental constituents of matter: six quarks and six leptons. The top quark is the heaviest fundamental particle known so far, at a level of two orders of magnitude larger than the second heaviest quark. It was expected for being the electroweak partner of the bottom quark. The existence of the top quark was confirmed in 1995 at the Fermi National Accelerator Laboratory (Fermilab).

Despite the predictive power of the SM, certain questions still remain unanswered. The SM fails to provide a complete description of the universe, since it does not incorporate a quantum theory of the gravity interaction. In addition, the SM does not give an explanation to cosmological observations like the matter-antimatter asymmetry. Another unsatisfactory aspect of the SM is that it only covers a few percent of the energy and matter of the universe. Measurements of velocities of galaxies have shown that there must be much more matter (dark matter) than the visible one. Theoretical shortcomings come also upon the SM theory. The hierarchy problem is one of them. It requires a precise cancellation of the order of 10^{16} times the Higgs mass, in order to protect it from quadratic divergences.

Beyond the standard model (BSM) theories have been proposed to solve these imperfections. As a result of the top quark mass, several BSM theories predict the existence of heavy particles decaying primarily to top quark pairs. These new states are expected to be visible as resonances in the invariant mass distribution of the top anti-top ($t\bar{t}$) system.

¹European Organization for Nuclear Research

1 Introduction

This thesis presents the results of a search for production of heavy particles decaying to top quark pairs. The analysis is performed using collision data collected with the ATLAS experiment at the LHC, at a center-of-mass energy of 7 TeV. The search uses a combination of *resolved* and *boosted* approaches to reconstruct the $t\bar{t}$ system. These are defined by the topology of the top-quark decay products in the detector, which depends on the transverse momentum of the parent particle. In the resolved topology, the top quarks produced nearly at rest will have well-separated decay products to be individually reconstructed. As the transverse momentum increases, these decay products become increasingly more and more collimated until they occupy the same region of the detector. In such a case, the boosted approach of the $t\bar{t}$ reconstruction is needed. For the first time, the strategy of combining the two approaches is applied to construct an overlapping channel. The aim is to improve the sensitivity to the resonances in a broad range of transverse momenta, from top quarks at rest up to the TeV-regime.

The outline of this thesis is the following. Chapter 2 gives an overview of the Standard Model and top-quark physics. The plausibility of certain BSM theories is also summarized. Two benchmark models used to guide the $t\bar{t}$ resonances analysis are in addition introduced. The chapter ends by characterizing the proton-proton collisions and their modeling by Monte Carlo techniques.

Chapter 3 describes the experimental setup of the LHC collider and the ATLAS detector, which were used to produce and record the collision data of this thesis. The data quality and the trigger system of the ATLAS experiment are also introduced.

In Chapter 4 the event and object reconstruction in ATLAS are discussed. The description focuses on the physics objects that make up the final state of the $t\bar{t}$ system.

Chapter 5 comprises the performance studies whose emphasis lies on the trigger efficiencies to select calorimeter-based signatures.

In the final Chapters 6 and 7, the search for $t\bar{t}$ resonances is presented. Chapter 6 describes the selection of events required by the resolved and boosted approaches. After applying the selections, the procedure for reconstructing the $t\bar{t}$ system and the invariant mass is detailed. A comparison of the $t\bar{t}$ invariant mass distributions in data and the SM background prediction is shown. These distributions are the inputs to perform the statistical analysis outlined in Chapter 7. Upper limits on the production cross-section² times branching ratio for narrow and wide massive states are set. Finally, the sensitivity of the search to the $t\bar{t}$ resonances is evaluated. This is done by comparing the limits for a range of resonance masses, obtained from different combinations of the $t\bar{t}$ reconstruction approaches.

²In this work, the cross-section is given in units of picobarn (pb), with $1 \text{ pb} = 10^{-36} \text{ cm}^2$.

2 Theoretical overview of the physics at the LHC

The fundamental quantum fields of matter and their interactions are defined in the Standard Model (SM) of elementary particles. This theory has been developed over the past century and involves our current understanding of the experimental results. In addition to describing the nature of the fundamental particles, the SM provides a complete scenario of their decay modes and production mechanisms via hadronic collisions. Thus, the SM is the starting point for investigating new signatures at the TeV scale. Several theoretical formalisms, known as Beyond the Standard Model (BSM) theories, have been proposed in order to extend the SM. Their aim is to cover the limitations of the SM when explaining certain experimental observations. This chapter presents a brief review of the SM, together with some of its limitations. BSM models used in the developments of this thesis are also outlined. The final part of the chapter is dedicated to the description of the modeling of proton-proton collisions.

2.1 Introduction to the Standard Model

The SM describes three out of four fundamental interactions found in nature: the electromagnetism, the weak and the strong. Gravity is neglected in the SM because its strength is about 43 orders of magnitude weaker than the strong interaction. The SM is a quantum field theory [1] that classifies the constituents of matter into two categories: quarks (q) and leptons (ℓ). Quarks and leptons belong to the type of particles called fermions, which carry spin 1/2 and obey the Fermi-Dirac statistics. Each has its corresponding antiparticle with identical properties except for the reversal of their quantum numbers. Interactions between fermions are represented as fundamental forces. These forces are described by fields whose mediators are gauge bosons, which carry spin 1 and obey the Bose-Einstein statistics [2]. The electromagnetic force is mediated by the massless photon (γ), the strong force is mediated by eight massless gluons (g), and the weak force is mediated by the massive W^\pm and Z^0 gauge bosons.

There are six leptons (electron, electron neutrino, muon, muon neutrino, tau and tau neutrino) and six quarks (up, charm, top, down, strange and bottom) [3]. They are grouped into three generations as shown in Table 2.1. Charged leptons interact through electromagnetic and weak forces, whereas neutral leptons interact only weakly. None of them interact via the strong force. Neutrinos are electrically neutral. They do not decay and rarely interact with matter, making them difficult to detect. In addition, they oscillate between generations and have very small but non-zero mass [4, 5]. Quarks whose charge is $+2/3$ are referred to as up-type quarks, whereas down-type quarks have

2 Theoretical overview of the physics at the LHC

Fermion	Particle	Generation	Mass	Charge [e]	Interaction
leptons	e^\pm	1	511 keV	± 1	electromagnetic, weak
	ν_e	1	< 2 eV	0	weak
	μ^\pm	2	105.7 MeV	± 1	electromagnetic, weak
	ν_μ	2	< 2 eV	0	weak
	τ^\pm	3	1.78 GeV	± 1	electromagnetic, weak
	ν_τ	3	< 2 eV	0	weak
quarks	u, \bar{u}	1	1.7-3.3 MeV	$2/3, -2/3$	electromagnetic, weak, strong
	d, \bar{d}	1	4.1-5.8 MeV	$-1/3, +1/3$	electromagnetic, weak, strong
	c, \bar{c}	2	1.27 GeV	$2/3, -2/3$	electromagnetic, weak, strong
	s, \bar{s}	2	101 MeV	$-1/3, +1/3$	electromagnetic, weak, strong
	t, \bar{t}	3	173.2 GeV	$2/3, +2/3$	electromagnetic, weak, strong
	b, \bar{b}	3	4.19 GeV	$-1/3, +1/3$	electromagnetic, weak, strong

Table 2.1: Fermions of the Standard Model and their properties. The charge is given in units of the electron charge, e . For leptons, the sign (+) labels the particles and their corresponding antiparticles are labeled with the sign (-). The six quarks up (u), down (d), charm (c), strange (s), top (t) and bottom (b), are also listed. The barred quarks are the antiquarks which have opposite electric charge to the quarks [3].

charge $-1/3$. Quarks couple to all force mediators. Charged particles belonging to the first generation build up the visible matter, while those belonging to second and third generation are unstable. They can only be observed in high energy interactions, decaying into first generation particles.

For all interactions, particles are perturbations of fermion fields. Since the SM is a quantum field theory, the interactions between particles are described by a Lagrangian density and represented in Feynman diagrams [1, 2].

The theory of the SM is based on combinations of gauge symmetries associated to conservation laws. In this context, the SM is a gauge theory that contains the fermion fields, one Higgs field and the gauge fields of the SM gauge group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. Here, $SU(3)_C$ is the symmetry group of the strong interaction, while the electromagnetic and the weak interactions are unified under the $SU(2)_L \otimes U(1)_Y$ group. The generators of the SM gauge group represent the mediators of the interactions, i.e. the gauge bosons. If the Lagrangian density, which describes the dynamics of the system, is invariant under local transformations of the symmetry group, a corresponding quantum number is conserved [6]. The guiding principle in a gauge theory is this local gauge invariance. However, the massive gauge bosons are obtained by introducing a mass term in the Lagrangian. Such a term is not gauge invariant, whereby a new scalar field must be introduced. The introduced scalar field is the Higgs field. The non-zero vacuum expectation value of the Higgs field breaks spontaneously the electroweak symmetry, giving rise to the massive gauge bosons. This is known as Spontaneous Symmetry Breaking (SSB) mechanism or Higgs mechanism [7–11]. The interactions involved in the SM gauge group and the Higgs mechanism are described below.

2.1.1 Electroweak theory

The electromagnetic interaction is responsible for binding electrons to atomic nuclei forming atoms. Historically, quantum electrodynamics (QED) was the first formulated gauge theory to describe this interaction. The QED is an abelian group theory [8] based on the gauge group $U(1)$. The QED Lagrangian is constructed such that by applying the Euler-Lagrange equation, it results in the Dirac-equation of a free fermion. The local gauge invariance of this Lagrangian implies that the electrical charge is conserved locally. To ensure the gauge invariance, the mediator of the interaction is required to be massless. The scale-dependent coupling of the electromagnetic interaction α_{QED} , is related to the elemental electrical charge as $e = \sqrt{4\pi\alpha_{\text{QED}}}$. It describes how the effective charge depends on the separation of the two charged particles. As the energy scale (Q^2) increases, the coupling also grows from $1/137$ at $Q^2 = 0$ to $1/127$ at energies corresponding to the mass of the Z boson ($m_Z = 91.18 \pm 0.0021$ GeV [12]).

The weak interaction is responsible for beta decays. It is several orders of magnitude smaller than the electromagnetic one. The weak interaction is known for being the only interaction that changes the flavor of fermions via charged currents. Strictly speaking, the W^\pm gauge bosons mediate the charged weak currents between fermions, while the Z^0 mediates the neutral currents.

The weak and electromagnetic interactions are successfully described as different forms of the same interaction: the electroweak. This was originally proposed in the 60s by Glashow [13], Weinberg [14] and Salam [15]. In the electroweak description of the SM, fields associated to the electromagnetic interaction are combined with those describing weak couplings. The electroweak quantum numbers are the weak-isospin, I , and the hypercharge Y [13, 14]. The hypercharge is expressed in terms of the third component of the weak-isospin and the electric charge Q , as $Q = I_3 + \frac{Y}{2}$. For neutrinos and up-type quarks the third component of the weak-isospin is $I_3 = +1/2$. For charged leptons and down-type quarks, $I_3 = -1/2$.

Fermions can be grouped into doublets or singlets. Those belonging to a doublet have the same hypercharge: $Y = -1$ for leptons and $Y = 1/3$ for quarks. The left-handed fermions¹ have a weak-isospin $I = 1/2$ and form weak-isospin doublets. The right-handed fermions have $I = 0$ and form singlets. The electroweak quantum numbers for the SM fields are displayed in Table 2.2.

Since left-handed fermions have $I \neq 0$, only these transform under the $SU(2)_L$ symmetry. However, requiring gauge invariance under a $SU(2)_L$ transformation, results in two charged and one neutral gauge bosons. The latter does not agree with the observation, as the observed weak neutral current has a right-handed component. This component is derived from the fact that both, left- and right-handed fermions transform under the group $U(1)_Y$. Thus, the Lagrangian of the electroweak theory must be invariant under the $SU(2)_L \otimes U(1)_Y$ symmetry group. If so, Y and I are conserved under $U(1)_Y$ and

¹The helicity massless fermions is right-handed if the direction of its momentum vector is the same as the direction of its spin, whereas if the directions of spin and motion are opposite, the helicity is left-handed. Thus, the subscript L in the electroweak gauge group indicates that only the left-handed fermions interact weakly via vector minus axial currents, $(V - A)$ [3].

				Y	I	I_3	Q
$Q_L^i =$	$\begin{pmatrix} u \\ d \end{pmatrix}_L$	$\begin{pmatrix} c \\ s \end{pmatrix}_L$	$\begin{pmatrix} t \\ b \end{pmatrix}_L$	1/6	1/2	+1/2 -1/2	+2/3 -1/3
$u_R^i =$	u_R	c_R	t_R	2/3	0	0	+2/3
$d_R^i =$	d_R	s_R	b_R	-1/3	0	0	-1/3
$L_L^i =$	$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$	$\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$	$\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$	-1/2	1/2	+1/2 -1/2	+1 -1
$e_R^i =$	e_R	μ_R	τ_R	-1	0	0	-1
$\nu_R^i =$	ν_R^e	ν_R^μ	ν_R^τ	0	0	0	0
$\phi =$	$\begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$			1/2	1/2	+1/2 -1/2	+1 0

Table 2.2: The electroweak fields and their quantum numbers: I the isospin and Y the hypercharge. The subscript L (R) identifies left-(right-)handed fermions [16].

$SU(2)_L$ transformations, respectively. The unification of the $SU(2)_L$ and $U(1)_Y$ groups enforces the definition of two interaction couplings: g associated with $SU(2)_L$ and g' associated with $U(1)_Y$.

The local gauge invariance of the electroweak interaction leads to four gauge bosons $\mathbf{W}^\mu = (W_1^\mu, W_2^\mu, W_3^\mu)$ associated to the $SU(2)_L$ group, and B^μ in the case of $U(1)_Y$. The spin-1 fields of electroweak bosons are mixtures of these gauge boson fields:

$$\begin{aligned}
 (W^\pm)^\mu &= \frac{1}{\sqrt{2}}(W_1^\mu \mp iW_2^\mu) \\
 Z^\mu &= \frac{-g'B^\mu + gW_3^\mu}{\sqrt{g^2 + g'^2}} \\
 &= -B^\mu \sin \theta_W + W_3^\mu \cos \theta_W \\
 A^\mu &= \frac{gB^\mu + g'W_3^\mu}{\sqrt{g^2 + g'^2}} \\
 &= B^\mu \cos \theta_W + W_3^\mu \sin \theta_W,
 \end{aligned}$$

where W_μ^\pm and Z_μ are the fields of the electroweak bosons and A^μ is the photon field. The Weinberg angle, $\sin \theta_W$, describes the mixing between the gauge groups $SU(2)$ and $U(1)$. The electromagnetic coupling and the Weinberg angle fix the gauge couplings as $\alpha_{\text{QED}} = g \sin \theta_W = g' \cos \theta_W$.

Although experimental measurements indicate otherwise, the SM describes massless particles. To generate the masses of the electroweak bosons, the Higgs mechanism introduces

an additional $SU(2)_L$ isospin doublet of four complex-scalar fields $\phi(x)$, with $Y = 1$:

$$\phi(x) \equiv \begin{pmatrix} \phi^{(+)}(x) \\ \phi^{(0)}(x) \end{pmatrix}, \quad \text{with} \quad \begin{aligned} \phi^{(+)}(x) &\equiv \frac{1}{\sqrt{2}} (\phi_1(x) + i\phi_2(x)) \\ \phi^{(0)}(x) &\equiv \frac{1}{\sqrt{2}} (\phi_3(x) + i\phi_4(x)) \end{aligned}. \quad (2.1)$$

The potential of the Higgs field, which must be added to the electroweak Lagrangian is

$$V(\phi) = \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2, \quad \text{with} \quad (\lambda > 0, \mu^2 < 0). \quad (2.2)$$

Here, the mass is given by $m = \sqrt{-\mu^2}$ and the second term corresponds to the gauge self-interactions. The parameters of the Higgs potential are chosen in such way that its vacuum expectation value ν , is given by

$$\nu^2 = -\frac{\mu^2}{\lambda}. \quad (2.3)$$

As $\mu^2 < 0$ and all the terms are introduced into a forced-invariant Lagrangian, the field $\phi(x)$ is expanded around a minimum. This causes a spontaneous breaking of the $SU(2)_L$ symmetry, by the non-zero vacuum expectation value of the scalar field [17]. The result is:

$$\phi(x) = \sqrt{\frac{1}{2}} \begin{pmatrix} 0 \\ \nu + h(x) \end{pmatrix}. \quad (2.4)$$

Of the four scalar fields only one Higgs field $h(x)$ remains. The others are absorbed in longitudinal degrees of freedom of the massive gauge bosons. The gauge bosons acquire mass by their coupling to the Higgs field:

$$m_W = \frac{1}{2}\nu g, \quad m_Z = \frac{1}{2}\nu\sqrt{g^2 + g'^2}, \quad m_\gamma = 0. \quad (2.5)$$

The SM predicts a neutral spin-zero Higgs boson whose vacuum expectation value is $\nu = 246$ GeV. This value can be determined from the Fermi coupling constant G_F . On July 4th 2012 the LHC experiments ATLAS and CMS announced the observation of a new boson with a mass of ~ 125 GeV [18, 19]. The particle is compatible with the production and decay of the SM Higgs boson. The datasets used for these analyses were collected at a center-of-mass energy of $\sqrt{s} = 7$ TeV in 2011 and $\sqrt{s} = 8$ TeV in 2012. In ATLAS, the excess of events above the expected background in the region between 122 GeV and 131 GeV, was observed with a local significance of 5.9 standard deviations (5.9σ). The search has been performed mainly in the $H \rightarrow \gamma\gamma$, $H \rightarrow ZZ^{(*)}$, and $H \rightarrow WW^{(*)}$ channels. Further studies of the properties of this particle such as its spin, parity and couplings are needed to confirm if this is the SM Higgs boson.

The mass of each fermion is also generated via its specific coupling to the Higgs field. This *Yukawa coupling* [20] is defined by

$$m_f = y_i \frac{\nu}{\sqrt{2}}, \quad (2.6)$$

where m_f is the mass of the fermion and y_i is the Yukawa coupling constant for the quark i . Thus, the fermion mass depends on the strength of the interaction with the Higgs field.

As mentioned above, the weak interaction changes the flavor of the fermions via charged currents. However, the mass eigenstates of fermions are not identical to the charged-current weak eigenstates. The transformation from one base into the other is carried out via a 3×3 matrix: the Cabibbo-Kobayashi-Maskawa, \mathbf{V}_{CKM} matrix [21, 22]. Under the assumption of unitarity of the \mathbf{V}_{CKM} matrix, the change of basis is applicable on the fermion mass eigenstates, resulting in the weak eigenstates as

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \mathbf{V}_{\text{CKM}} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}_L. \quad (2.7)$$

Because this matrix has off-diagonal terms, the W boson mixes the three generations. The probability for a quark of flavor i to be transformed to a quark of flavor j by exchange of a W boson is proportional to $|V_{ij}|^2$. The values have been measured accurately and are listed as follows [23, 24]:

$$\mathbf{V}_{\text{CKM}} = \begin{pmatrix} 0.97428 \pm 0.00015 & 0.2253 \pm 0.0007 & 0.00347 \pm \begin{smallmatrix} 0.00016 \\ -0.00012 \end{smallmatrix} \\ 0.2252 \pm 0.0007 & 0.97345 \pm \begin{smallmatrix} 0.00015 \\ -0.00016 \end{smallmatrix} & 0.0410 \pm \begin{smallmatrix} 0.0011 \\ -0.0007 \end{smallmatrix} \\ 0.00862 \pm \begin{smallmatrix} 0.00026 \\ -0.00020 \end{smallmatrix} & 0.0403 \pm \begin{smallmatrix} 0.0011 \\ -0.0007 \end{smallmatrix} & 0.999152 \pm \begin{smallmatrix} 0.000030 \\ -0.000045 \end{smallmatrix} \end{pmatrix}. \quad (2.8)$$

The most likely transitions occur between same-generation quarks with couplings ≈ 1 , while the off-diagonal couplings are smaller.

2.1.2 The Strong Interaction

Quantum Chromodynamics (QCD) is a non-abelian theory [25] that describes the strong interaction of quarks mediated by gluon exchange. The strong interaction is responsible for the ensemble of quarks to form hadrons². Hadrons composed of three quarks are called baryons, which have half-integer spin. Hadrons composed of a quark-antiquark pair are called mesons and are bosons. The QCD quantum number is the color³. Each (anti)quark can have only one of three types of color charge: (anti)red, (anti)green or (anti)blue. Eight different combinations of color (anti-color) for gluons are possible, forming a color octet in $\text{SU}(3)_C$. The fact that the theory is non-abelian, i.e. the operators of the symmetry group are non-commutative, makes it possible for the color-carrying gluons to self-interact. Thus, the QCD Lagrangian describes the free propagation of quarks and gluons, the quark-gluon interaction and gluon self-interactions.

The coupling of the strong interaction α_s , is the fundamental parameter of QCD. For energies approaching zero, α_s tends to infinity. This means that if two quarks are pulled to be far away from each other, the interaction gets stronger. As a consequence,

²All quarks, but the top, in their final states form hadrons.

³The subscript C in $\text{SU}(3)_C$ refers to the color quantum number.

neither gluons nor quarks are observed as isolated particles. This property of the QCD dynamics is called color *confinement* [26]. Hadrons are then bound color-singlet states (colorless), formed by confined quarks. The generation of colorless hadrons is known as *hadronization* or *fragmentation*.

In the context of perturbative QCD (pQCD), observables — usually cross-sections — are expanded in a power series of α_s . The power of the expansion is the order at which the observable is perturbatively calculated. From the QCD Lagrangian, Feynman diagrams contributing to each perturbative order are obtained. Higher order diagrams may contain gluon and fermion loops. The momentum of virtual particles in such loops is not defined, and for large momenta, the observables undergo ultraviolet divergences. To yield finite and calculable results, the divergences are absorbed by renormalizing the theory. The effect of renormalization is the introduction of an unphysical renormalization scale, μ_R . The observables have to be scale-independent, therefore the dependence is included into the QCD coupling as $\alpha_s(\mu_R^2)$ [2]. For a determined process, if the scale μ_R approaches the scale of the momentum transfer Q , $\alpha_s(\mu_R^2 \approx Q^2)$ indicates the effective strength of the interaction in that process. At the leading order, i.e. $\mathcal{O}(\alpha_s^2)$ in the perturbative expansion, $\alpha_s(Q^2)$ satisfies:

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu_R^2)}{1 + \alpha_s(\mu_R^2)\beta_1 \ln(Q^2/\Lambda)}, \quad (2.9)$$

where β_1 is the 1-loop β -function coefficient of the renormalization group equation [26]. This equation controls the scale dependence of the renormalized coupling. For i -loops the β_i coefficients depend on the number of quark flavors and can be negative due to gluon self-interactions. For sixteen or less flavors, the coupling decreases for hard processes, i.e. those involving large momentum transfer. This means that at high energies and short distances, particles behave as free ones, being governed by *asymptotic freedom* [27]. When the energy scale approaches zero, the coupling blows up, indicating a scale $\Lambda = \Lambda_{\text{QCD}}$, where pQCD breaks down. This is the color-confinement. Since hadronization arises from confinement, interactions between bound states of hadrons lie outside the area of applicability of the pQCD. Therefore, phenomenological models are required to describe low-energy processes.

2.2 Development of a proton-proton collision

The theoretical description of the scattering between two colliding particles is given by transition probabilities. The probability of an initial state transforming into a final state is obtained by the absolute square of the elements of the S-matrix (scattering matrix). The expected interaction rate between any colliding particles is determined from the cross-section σ . This depends on the coupling strength, and the energy and momentum of the particles [28, 29]. In high-energy pp collisions however, the structure of the proton must be accounted for. Namely, the colliding protons are hadrons whose constituents are *partons*. QCD is the underlying theory for all processes, i.e. it describes the dynamics between the partons inside the colliding protons, and the interaction between

the hadronized states. In the parton model⁴, partons are quasi-free quarks and gluons sharing the longitudinal hadron's momentum. The actual pp scattering process takes place between the partons that form the protons. This hard process among two partons i, j , is described by pQDC (short-distance process). On the other side, the hadronic soft processes (long-distance) occur at low momentum transfer scales (≈ 1 GeV), and have a non-perturbative nature.

The separation of the whole pp collision into the hard scattering cross-section and the non-perturbative contribution is known as *factorization*. The hard-scattering cross-section $\hat{\sigma}_{ij}$, is given by the S-matrix elements with high accuracy over two or three orders of perturbation. Each matrix element is calculated by using Feynman diagrams extracted from the QCD Lagrangian. The non-perturbative dynamics of the proton is accounted into the Parton Distribution Functions (PDFs), $f_i(x_i, \mu_F^2)$. These give the probability of locating a parton i carrying a fraction x_i of its parent proton's longitudinal momentum, when probed at μ_F^2 scale. μ_F^2 is an unphysical energy scale called *factorization scale*. This is introduced to separate what is counted as part of the hard-scattering, and what is taken as part of the internal dynamics of the proton. PDFs are universal as they do not depend on the scattering process. They are extracted from data and evolved to other kinematic scales using the DGLAP equation [30], which is analog to the renormalization group equation.

When colliding two protons A and B , the production cross-section for a process X , is calculated as a convolution of the PDFs for the protons $f_{i,A}(x_i, \mu_F^2)$ and $f_{j,B}(x_j, \mu_F^2)$, and the hard parton-parton cross-section:

$$\sigma(AB \rightarrow X) = \sum_{i,j} \int dx_i dx_j f_{i,A}(x_i, \mu_F^2) f_{j,B}(x_j, \mu_F^2) \hat{\sigma}_{ij}(ij \rightarrow X; \hat{s}, \mu_F, \mu_R, \alpha_s(\mu_R^2)), \quad (2.10)$$

where $\hat{\sigma}_{ij}$, includes all possible combinations of the two partons i and j . $\hat{s} = x_i x_j s$ is the effective center-of-mass energy squared for the partonic process, which depends on the total energy squared s , available in the center-of-mass of the collision.

Several particles can be created in the scattering process. Depending on their properties, these can decay further into other particles. The total decay width Γ , is a measure of the probability of a particle decaying into any state. The branching ratio BR , is the width of a particle decaying to a specific final state, divided by the width of this particle decaying into all possible states.

2.3 The top quark in the Standard Model

The top quark is the heaviest fundamental particle known in nature. Its existence as the weak-isospin partner of the bottom quark (b -quark) was predicted by the SM. The Yukawa coupling of the top quark to the Higgs boson (~ 1), is the largest for any quark,

⁴The parton model defines partons explicitly as valence quarks and a sea of gluons and quark-antiquark pairs of all flavors coming from the conversion of the gluons, which bind the valence quarks.

and it is the only fermion with a mass of order the electroweak symmetry breaking scale (EWSB): $= \nu \approx 246$ GeV. The first direct observation of the top quark took place in 1995 by the CDF and DØ experiments [31, 32] at FERMILAB, using the proton-antiproton ($p\bar{p}$) collider TEVATRON. Indirect constraints on the top quark mass (m_t) have been inferred from electroweak precision measurements, performed mainly at LEP [33]. The current value stands at $m_t = 173.5 \pm 0.6 \pm 0.8$ GeV [34]. In addition, the decay width of the top quark, obtained from TEVATRON data is $\Gamma_t = 2.0^{+0.7}_{-0.6}$ GeV [33]. Since this value is far from $\Lambda_{\text{QCD}} = 261$ MeV, the top quark decays can be described by pQCD.

2.3.1 Top quark pair production at the LHC

Pairs of top and anti-top quarks ($t\bar{t}$) can be produced simultaneously at the LHC via the strong interaction. There are two possible processes contributing to the $t\bar{t}$ production: quark-antiquark annihilation and gluon-gluon fusion. Their corresponding Feynman diagrams at leading order of perturbation are shown in Figure 2.1.

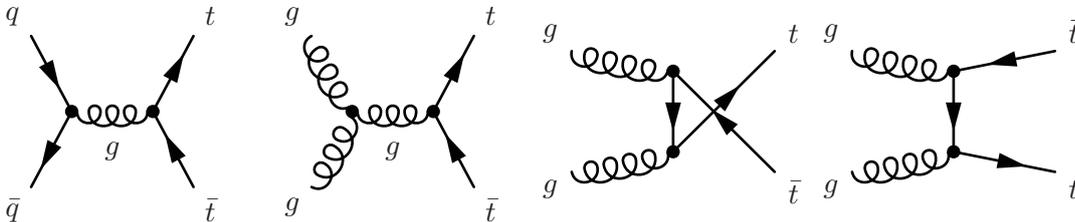


Figure 2.1: Feynman diagrams at leading order contributing to $t\bar{t}$ production. The first on the left corresponds to quark-antiquark annihilation ($q\bar{q} \rightarrow t\bar{t}$) processes, while the three on the right correspond to gluon-gluon fusion ($gg \rightarrow t\bar{t}$).

The dominant process depends on the PDFs and the center-of-mass of the pp collisions. In the process $pp \rightarrow t\bar{t}$ where $\hat{\sigma}_{ij \rightarrow t\bar{t}}$ is calculated in pQCD, the energy of the partons has to be at least enough to produce $t\bar{t}$ pairs at rest $\hat{s} \gtrsim 4m_t^2$. Therefore, in the simplified scenario describing partons with equal momentum fraction $x_i = x_j$, the proton's momentum fraction required to produce top quark pairs is $x \approx 2m_t/\sqrt{s}$. Clearly, the parton momentum fraction decreases as the total center-of-mass energy of colliding protons grows large. In particular for $\sqrt{s} = 7$ TeV, this kinematic threshold for $t\bar{t}$ production of the two colliding partons is $x \approx 0.06$ [35]. Figure 2.2 shows the PDFs for partons inside the proton in the MSTW2008 NLO [36] parametrization. The PDFs are determined by global fits to hard-scattering data at the next-to-leading order (NLO) that is, including terms of order $\mathcal{O}(\alpha_s^3)$. Notably for large x , the valence quark momentum density distributions are much larger than those of the other partons. Furthermore, the probability of finding gluons with $x \approx 0.06$ dominates over the probability of finding quarks required for $q\bar{q}$ annihilation (\bar{u} or \bar{d}). Thus, the dominant $t\bar{t}$ production mechanism at the LHC is gluon-gluon fusion with about 80% contribution (at $\sqrt{s} = 7$ TeV). The $t\bar{t}$ production cross-section $\sigma_{t\bar{t}} = 165$ pb, is calculated at next-to-next-to-leading order by using HATHOR, the HAdronic Top and Heavy quarks crOSS section calculatoR [37].

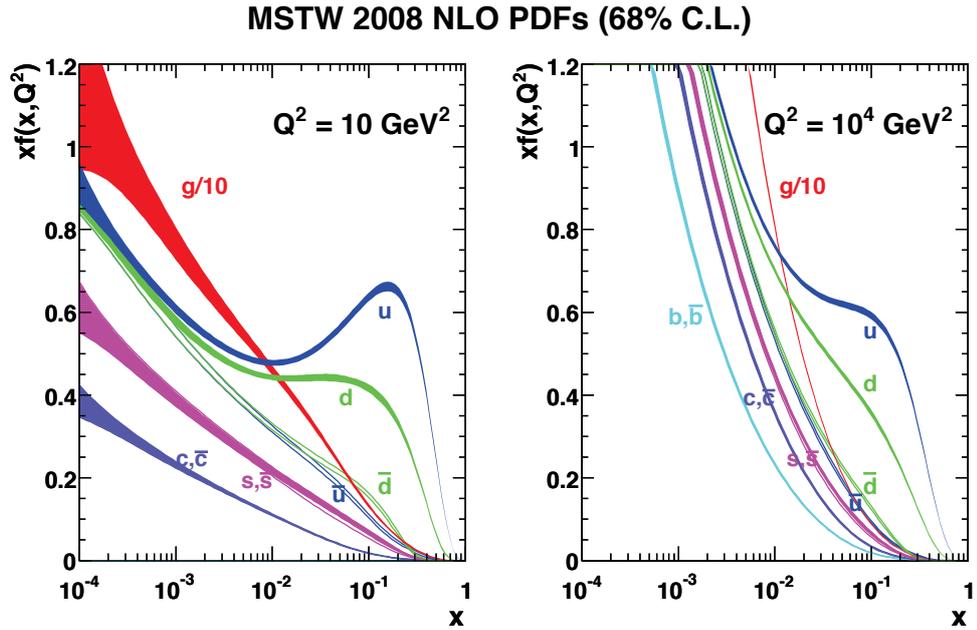


Figure 2.2: MSTW next-to-leading order PDFs at $Q^2 = 10 \text{ GeV}^2$ and $Q^2 = 10^4 \text{ GeV}^2$. The momentum densities of the partons are shown as a function of the proton momentum fraction x [36]. The probability of finding a quark of momentum fraction x decreases with increasing x .

2.3.2 Single top production

Top quarks are also produced individually via the electroweak interaction. The production mechanisms for single top are separated according to the process mode into: W -gluon fusion or t -channel process, Wt associated production and quark-antiquark annihilation or s -channel. The contributions of these processes are proportional to the CKM matrix element $|V_{tb}|^2$. Figure 2.3 shows exemplary some Feynman diagrams at LO for single top production.

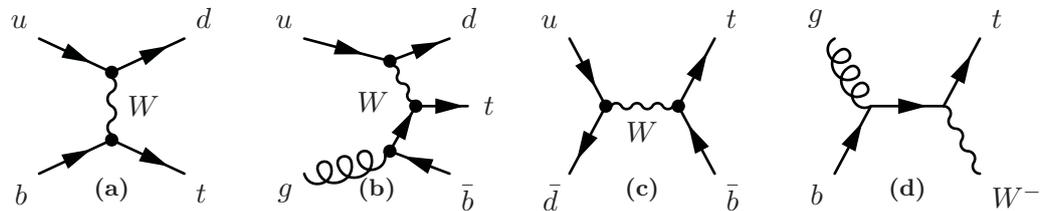


Figure 2.3: Example Feynman diagrams of the leading order processes for single top production: t -channel production as (a) flavor excitation and (b) W -gluon fusion, (c) s -channel production and (d) associated Wt production.

At the LHC energies, the dominant contribution is the t -channel. The cross-section of these processes σ_{st} , is expected to be half of the cross-section for $t\bar{t}$ production, $\sigma_{st} = 90_{-22}^{+32}$ pb [38–41].

2.3.3 Top quark decay channels

At NLO the top-quark decay width is given by [42]:

$$\Gamma(t \rightarrow qW) = \frac{G_F m_t^3}{8\pi\sqrt{2}} \left(1 - \frac{m_W^2}{m_t^2}\right)^2 \left(1 + 2\frac{m_W^2}{m_t^2}\right) \left[1 - \frac{3\alpha_s}{3\pi} \left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right] \times |V_{tq}|^2,$$

where m_W is the mass of the W boson ($m_W \approx 80.4$ GeV). For $m_t \approx 175$ GeV, $\Gamma(t \rightarrow qW) \approx 1.40$ GeV, with a very small theoretical uncertainty. Since Λ_{QCD} is smaller than Γ_t , the top quark lifetime ($\tau_t = \frac{1}{\Gamma_t} \approx 5 \cdot 10^{-25}$ s) is one order of magnitude shorter than the hadronization time: $\tau_{\text{had}} = \Lambda_{\text{QCD}}^{-1} \approx 3 \cdot 10^{-24}$ s. Therefore, the top quark decays before hadronizing.

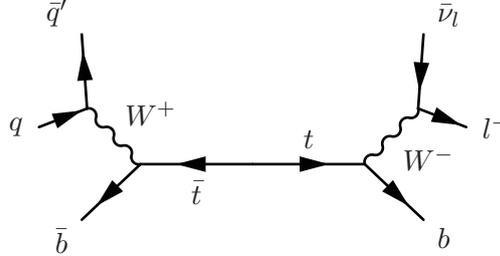
The top quark decays through the weak interaction in a W boson and a down-type quark. The decay rates are proportional to the CKM matrix elements $|V_{tq}|^2$, where q are the down-type quarks only. By assuming unitarity of the CKM matrix, Equation 2.8 shows that $|V_{tb}|$ is close to one. Thus, the decay rate $t \rightarrow Wb$ is nearly 100%, i.e. $BR(t \rightarrow Wb) \approx 1$. The W boson produced in the decay is real because the top quark mass exceeds the combined masses of the b -quark and the W boson. Hence, the SM top quark decays predominantly into a b -quark and a W boson. The former subsequently hadronizes into a jet⁵ of hadrons, while the latter further decays into either a quark-antiquark — *hadronically* — or a charged lepton-neutrino pair — *leptonically* —. These decay modes define the final state of $t\bar{t}$ events. Three decay channels for the $t\bar{t}$ final states can be characterized:

$$\begin{aligned} \text{dilepton channel:} & \quad t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow \ell\bar{\nu}_\ell b\ell'\bar{\nu}_{\ell'}\bar{b} \quad , \\ \text{semi-leptonic channel:} & \quad t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow q\bar{q}'b\ell\bar{\nu}_\ell\bar{b} + \bar{\ell}\nu_\ell b\bar{q}'\bar{b} \quad , \\ \text{fully-hadronic channel:} & \quad t\bar{t} \rightarrow W^+bW^-\bar{b} \rightarrow q\bar{q}'b\bar{q}''\bar{q}'''\bar{b} \quad . \end{aligned}$$

Table 2.3 displays the branching ratios for the W -boson decays. The fully hadronic and lepton+jets channels dominate over the dilepton channel due to the color charge in the hadronic final states. In this thesis, only events decaying into the semi-leptonic channel are selected. Figure 2.4 shows the final-state topology of these events. Events in which the charged lepton is a τ -lepton are not considered. Because of its large mass, τ -leptons will subsequently decay into a final state containing hadrons or a lighter lepton and a neutrino. The reconstruction of the event is more challenging and can be ambiguous due to the escape of the additional neutrino.

⁵When collimated particles generated by hadronization are grouped within a certain distance measure, a “jet” of hadrons is observed.

W Decay Mode	Branching Ratio (BR)
$e + \nu_e$	$(10.75 \pm 0.13)\%$
$\mu + \nu_\mu$	$(10.57 \pm 0.15)\%$
$\tau + \nu_\tau$	$(11.25 \pm 0.20)\%$
$q\bar{q}$	$(67.60 \pm 0.27)\%$

Table 2.3: Branching ratio for the W boson decays [3].**Figure 2.4:** Topology of the semi-leptonic top quark pair decay. The experimental topology consists of one W boson decaying hadronically into a jet initiated from a b -quark (b -jet) and two additional jets. The other W boson decays leptonically into a charged lepton, missing energy originating from the escaping neutrino and a b -jet.

2.4 Searches beyond the Standard Model

The SM is the most successful description of elementary particles. However, theoretical shortcomings and unanswered questions still remain. Some of the main limitations of the SM are listed in the following.

- Neutrinos in the SM are massless particles. This contradicts the experimental observations of neutrino oscillations which require massive neutrinos [43]. The existence of massive neutrinos will require an extension of the SM.
- Even though elementary particles can interact via gravity, the gravitational field is not considered into the SM framework. Despite the success of the SM in linking the other forces, a grand unified theory of all fundamental interactions is not feasible without gravity.
- The SM does not explain the *hierarchy problem*. The problem sets the question why the electroweak strength is $\sim 10^{32}$ times stronger than gravity. The question of the large difference in orders of magnitude between the Higgs mass ($\sim 10^2$ GeV) and the the Planck mass ($\sim 10^{19}$ GeV) is still unsolved.

The formulation of the hierarchy problem is closely related to the “fine-tuning”: the Higgs field is constrained by the known masses of the weak gauge bosons. However cancellations of the order of 10^{16} times the Higgs mass must be applied to reconcile its value with the radiative corrections from fermion and boson loops predicted by the SM.

- The SM does not explain the number of fermion generations, nor does it explain why the electric charges of electrons and protons cancel each other, and why the quark charge is quantized in thirds of the elementary charge of the electron.
- Measurements related to the rotation curves of galaxies indicate that the visible matter constitutes $\sim 17\%$ of the mass in the universe [44]. The remaining mass is defined as “dark matter” composed by stable particles. The interaction of these particles is weaker than the weak force. When visible and dark matter particles are combined only $\sim 30\%$ of the energy in the universe is accounted for. The rest of the energy is “dark energy” responsible for the acceleration of the universe [45]. Dark matter particles are not included within the SM model description.

New theoretical formalisms Beyond the SM (BSM) have been introduced in order to correct the weaknesses of the SM. None of the BSM theories has overcome all its limitations. However, testing its validity is one of the physics goals of the LHC program.

2.4.1 The top quark beyond the Standard Model

The fact that the top quark is the only fermion with mass very close to the EWSB has many implications. Its production is potentially sensitive to unknown intermediate-heavy states, whose properties can appear as an extension of the electroweak theory. These new states would primarily decay into a $t\bar{t}$ pair, showing itself as a resonance in the $t\bar{t}$ invariant mass spectrum. In many BSM schemes this resonance is predicted as the Z' boson, which is heavy, neutral, colorless and self-adjoint [46]. The $t\bar{t}$ resonance can also be a color octet predicted in some models as either a vector or an axial-vector⁶ particle. Figure 2.5 shows exemplary the resonances in the invariant mass spectrum from $t\bar{t}$ production including both, the color singlet and the color octet. The precise width and height of the peak depends on the parameters for the resonance given by the particular model. As benchmark in this illustration, the Z' vector boson and the color octet couple to quarks with the same strength as the SM particles. Interference effects of the Z' with the SM $t\bar{t}$ production can be neglected, so the peak is nearly independent of the couplings. The interference of the axial vector with the QCD $t\bar{t}$ production does not change the shape of the distribution. Hence, the effects of the color-octet axial vector and the color singlet are very similar, apart from the size of the resonance due to the different coupling constant. The interference of the color octet vector particle with the QCD $t\bar{t}$ production does affect the $t\bar{t}$ invariant mass distribution: there is negative interference in the invariant mass region below the resonance mass and positive interference for heavier invariant masses [46].

The analysis of $t\bar{t}$ resonances presented in this thesis strives to be independent of the underlying model and thus widely applicable. However, two benchmark models predicting the color- singlet and -octet states are used to quantify the sensitivity to the resonances. These models are presented as follows.

⁶The axial vector is a pseudo-vector as it transforms like a vector under a proper rotation, but gains an additional sign flip under an improper rotation such as a reflection.

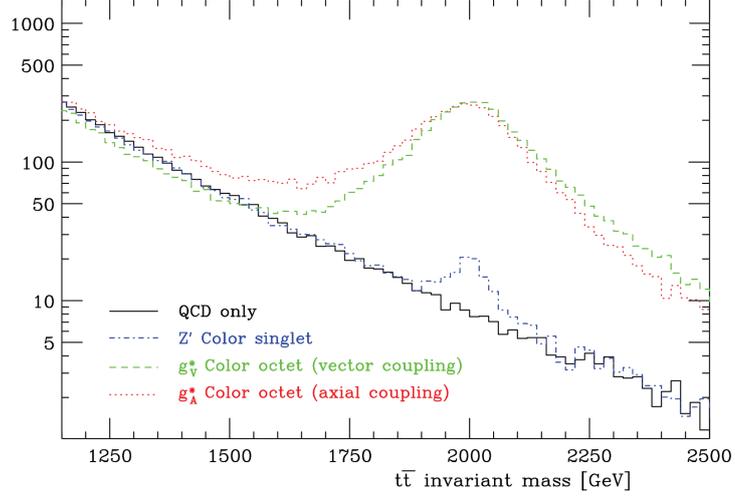


Figure 2.5: Invariant $t\bar{t}$ mass distribution for $pp \rightarrow t\bar{t}$, for a generic color-singlet and for a color octet with a masses of 2 TeV. All plots were produced using the CTEQ6L1 PDF set [47] with $\mu_R = \mu_f = 2$ TeV. The couplings to the SM fermions are implemented using MADGRAPH [48]. Taken from [46].

2.4.1.1 The leptophobic topcolor Z' boson

In a typical Topcolor model the $SU(3)_C$ group is inserted into a larger arrangement, e.g. $SU(3)_1 \otimes SU(3)_2$ [49]. If the $SU(3)_1 \otimes SU(3)_2$ is broken, massive gauge bosons called colorons are generated. The interaction field of the colorons is called the “topcolor” field and its strength becomes stronger near 1 TeV. Perturbations to the topcolor field generate heavy top quarks through the formation of $t\bar{t}$ dynamical condensates coupling to the Higgs field. The topcolor models are an extension of a more general set of theories called Technicolor (TC) [50, 51]. They provide alternative solutions to the hierarchy problem by introducing new gauge interactions known as technicolor forces.

The Topcolor Assisted Technicolor models (TC2) [50] bring together the $t\bar{t}$ condensation arising from the topcolor interactions, and the so-called technicolor (ETC) symmetry breaking [51]. At the ETC symmetry breaking scale (≈ 100 GeV), the masses of the fermions and gauge bosons are generated.

To ensure that the $t\bar{t}$ condensates generated by the topcolor field are heavier than $b\bar{b}$ condensates, the TC2 models add to the $SU(3)_1 \otimes SU(3)_2$ symmetry an extended hypercharge group, $U(1)_Y$. If the $SU(3)_1 \otimes SU(3)_2 \otimes U(1)_Y$ symmetry is broken, a massive Z' gauge boson with enhanced coupling to the first and third generation of quarks is produced. Some models describe this resonance as alike to the SM Z boson, since both couple with the same strength to fermions. However, unlike in the SM, the Z' belongs to a heavier mass scale, namely $\mathcal{O}(\text{TeV})$. In contrast, a Z' resonance with suppressed couplings to leptons and substantial couplings to SM quarks could be lighter. This topcolor resonance, called “leptophobic” Z' boson, is predicted to exist within the TC2 models by introducing a “leptophobic interaction”. The Lagrangian, which is dominated

by $t\bar{t}$, $u\bar{u}$, and $d\bar{d}$ interactions, is configured in such way that the couplings to leptons are minimized. From it, the cross-section of the process $pp \rightarrow Z' \rightarrow t\bar{t}$ is:

$$\frac{d\sigma_{Z' \rightarrow t\bar{t}}}{dm_t} = \frac{2}{m_t} \int_{-\ln(\sqrt{s}/m_t)}^{\ln(\sqrt{s}/m_t)} dy_b \tau \mathcal{L}(x_1, x_2) \hat{\sigma}(q\bar{q} \rightarrow Z' \rightarrow t\bar{t}), \quad (2.11)$$

where $\hat{\sigma}(q\bar{q} \rightarrow Z' \rightarrow t\bar{t})$ is the partonic cross-section, which depends on the width of the resonance $\Gamma_{Z'}$ [51]. The fraction's momentum of the partons x_1 and x_2 is related by $\tau = x_1 x_2 = m_t^2/s$, where s is the center-of-mass energy of the protons. The boost of the partonic system y_b , is given by $y_b = (1/2) \ln(x_1/x_2)$ [49]. The term $\mathcal{L}(x_1, x_2)$ is the product of PDFs for the colliding hadrons. Using this in the simplest scenario in which the formation of $t\bar{t}$ states is enhanced, the total width for the leptophobic Z' resonance decaying into a $t\bar{t}$ pair can be calculated as

$$\Gamma_{Z'} \approx C_L m_{Z'} \left[\sqrt{1 - \frac{4m_t^2}{m_{Z'}^2} \left(2 + 4 \frac{m_t^2}{m_{Z'}^2} \right) + 4} \right]. \quad (2.12)$$

Here $m_{Z'}$ the mass of the Z' and the factor C_L contains fine-tuning terms that minimize the couplings to leptons, along with spin and color parameters.

Although the topcolor Z' boson arises from an alternative EWSM mechanism, its validity as benchmark model continues regardless the recent evidence for the Higgs boson. As shown in Figure 2.5, the topcolor Z' is a very narrow resonance. Therefore, the upper limits which are determined on its production cross-section times BR are valid for any resonance whose width is narrower than the experimental resolution (typically 10% of the resonance mass). The specific leptophobic topcolor Z' resonance used in this thesis has a width of $\Gamma_{Z'}/m_{Z'} = 1.2\%$ of the Z' mass.

2.4.1.2 The Kaluza-Klein gluon

The color octet Kaluza-Klein gluon (KK-gluon or g_{KK}) is predicted in theories of extra dimensions [52]. They are proposed to solve non-perturbatively the hierarchy problem, by suggesting that gravity effects could be perceivable at the TeV-scale. These models were first explored by Kaluza [53] and Klein [54], who introduced a unified scenario that includes electromagnetic interactions and relativistic gravity fields in a five dimensional theory. Recently, this idea has been exploited in theories BSM. For instance in the Randall-Sundrum (RS) model, one new extra dimension is projected on a topological space or orbit-manifold [55]. This manifold is known as ‘‘warped’’ space-time because it is warped along the new dimension. The non-factorizable metric of the fifth-dimensional warped space is written as [56, 57]

$$ds^2 = e^{-2\sigma} \eta_{\mu\nu} dx^\mu dx^\nu - dy^2, \quad (2.13)$$

where y is the coordinate of the new fifth dimension and x^μ are the coordinates of the familiar fourth-dimensional space. The warp factor is given by $e^{-2\sigma}$ in which $\sigma = k|y|$ and k determines the space-time curvature. The term $\eta_{\mu\nu} = \text{diag}(-1, 1, 1, 1)$ represents the four-dimensional Minkowski metric.

2 Theoretical overview of the physics at the LHC

Two branes are defined in the warped space-time: the Planck-brane and the TeV-brane. The boundary at $y = 0$ is the Planck-brane and the limit at $y = \pm\pi r_c$ defines the TeV-brane. The gravitational field resides in the Planck-brane while the TeV-brane contains the SM fields. Choosing the radius of the space r_c such that $kr_c = 11.27$, the ratio of the Planck to TeV scales is set as [56]

$$\frac{e^{-k|y(\text{Planck})|}}{e^{-k|y(\text{TeV})|}} = e^{-k\pi r_c} \approx 10^{-15}. \quad (2.14)$$

As seen, the Planck scale is quite low in the Planck-brane and gravity is actually strong. Nonetheless, in the TeV-brane the Planck scale lays in a weaker scale ($\sim\text{TeV}$) and gravity is a weak interaction, thus solving the hierarchy problem. SM particles excited in the extra dimension acquire an effective mass in the forth-dimensional space-time. These excited states are called Kaluza-Klein (KK) excitations whose couplings rise as the warp factor increases. The KK-gluon is one of these excited states. The RS model predicts excitations of top quarks and KK-gluons likely located towards the TeV-brane. Since the dominant $t\bar{t}$ production mechanism at the LHC is gluon-gluon fusion, one can expect high rates for KK-gluons that decay primarily into $t\bar{t}$ pairs. The mass of the KK-gluon is modulated by electroweak precision constraints with values ranging between 2 and 3 TeV. Thereby, the resonance signature is given by decays into high energy $t\bar{t}$ final-states. The LHC energy is large enough to produce these KK-gluons states. However, the detection is difficult as the color octet is a broad resonance, namely wider than the experimental resolution. The particular KK-gluon chosen in this thesis has large width $\Gamma_{g_{\text{KK}}}/m_{g_{\text{KK}}} = 15.3\%$ of the resonance mass. The rationale for using this benchmark model is driven by previous searches of $t\bar{t}$ resonances in ATLAS, e.g. [58, 59]. The identification of the the g_{KK} is also challenging due to its color-octet nature. As already seen in Figure 2.5, the fact of being a colored resonance results in interference effects in the $t\bar{t}$ invariant mass. This somewhat diminishes the claim of model-independence and additional properties of the resonance have to be considered. Quantities such as spin, the charge asymmetry between the top and the anti-top quarks and couplings to quarks could be more sensitive to the resonance [60].

2.5 Modeling a proton-proton collision

The Monte Carlo (MC) simulation is an assembling of numerical methods intended to mimic the processes involved in high energy interactions. It is relevant when extracting BSM signals from SM background processes, estimating the experimental feasibility of different researches, and when deriving correction factors for data analysis. Three processes take place during the simulation: the event generation, the simulation of interactions in the detector and digitalization of detector signals into voltages and currents for readout electronics. This section is dedicated to the simulation of events and immediate decays in proton-proton collisions. Different MC event generators used in this thesis are also introduced. The detector simulation and digitalization processes are discussed in the next chapter.

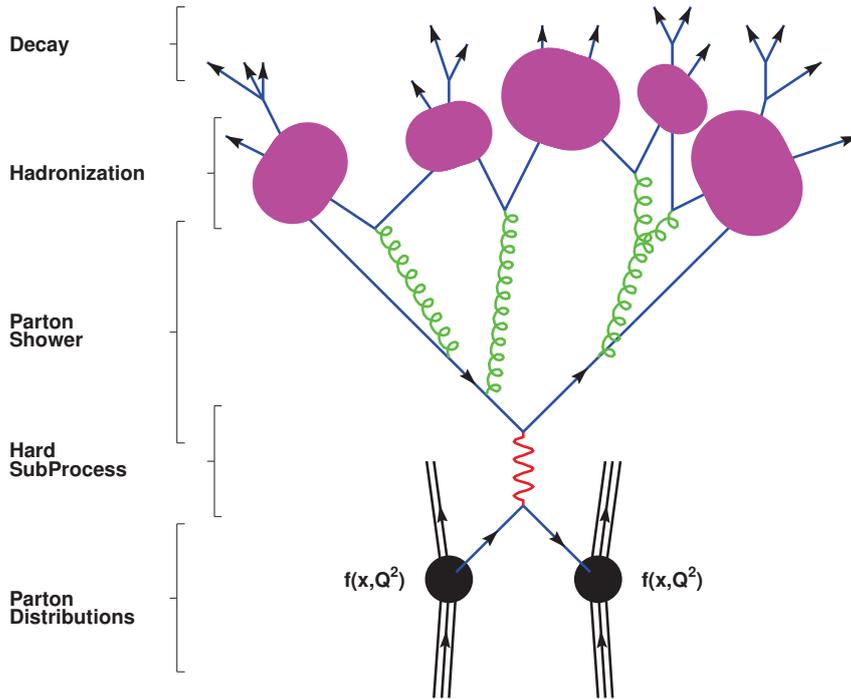


Figure 2.6: Steps of Monte Carlo event generation [61].

The event generation is related to the scattering process from the initial state towards the observed final-state. It includes the production mechanisms and the decay products of the pp collision. The complexity of the pp scattering poses theoretical challenges when modeling the dynamics of the whole interaction. Namely, the number of particles produced during scattering increases with increasing the energy of the colliding protons. Since the scale of momentum transfer varies over many orders of magnitude, the simulation must involve pQCD together with phenomenological models; the latter to account for the transition from partons to hadrons. MC generators distort the physics of the pp collision into sequential steps, depicted in Figure. 2.6 and listed in the following.

1. Two protons collide with a large momentum transfer. The collision between two partons (*incoming partons*, one from each proton) occurs, leaving behind other partons (*outgoing proton remnants*). This parton-level scattering is the hard process of interest, which is described using pQCD (see section 2.2). As a result of the hard interaction, the two partons produce short-lived resonances or massive SM particles, e.g. top quarks, W^\pm bosons or maybe some heavy state like the Z' . The decays of these resonances are viewed as part of the hard process itself.
2. Since electromagnetic and color charges interact during the hard process, the partons are accelerated, and thereby bremsstrahlung can occur. The accelerated incoming partons radiate (e.g a quark can radiate a gluon). Emissions that are

associated with the two colliding partons are called Initial-State Radiation (ISR). Emissions associated with the outgoing partons created by the hard scattering are called Final-State Radiation (FSR). Further emissions can emerge from the emitted gluons⁷, generating a parton shower which is simulated by using pQCD.

3. Given the multi-partonic nature of the proton, further parton pairs might interact, generating multiple interactions (MI). MI predominantly occur at scales much lower than the hard process scale, thus contributing to the soft part of the interaction. Each MI may be associated with its ISR and FSR.
4. Partons generating the hard process take a fraction of its parent proton's momentum. However, much of the energy remains in the outgoing proton remnants.
5. When partons are sufficiently "far" from each other, the coupling becomes strong and confinement forces are significant. Since this is a non-perturbative process, the behavior of the color interaction must be modeled using phenomenological schemes. At this point the hadronization process starts. Nearby partons whose color charge is opposite will recombine and create a hadron. This hadron is observed in the final state. The hadronization is assumed to be independent of the initial hard process.

2.5.1 Simulation of the hard subprocess

The reliable simulation of the hard process must include the modeling of low-momentum scale subprocesses associated to it. As described in Section 2.2, this is carried out via factorization. Thereby, the MC simulation incorporates the convolution with a PDF of choice. The default scales μ_F^2 and μ_R^2 are typically set to the scale of the hard process [62]. All event generators provide access to the most used PDF sets by means of the LHAPDF interface [63]. However, each generator uses a default PDF set with an ansatz for the initial distribution at some low μ_F^2 . As already mentioned, the evolution toward larger scales is obtained from the DGLAP equations. These give the probability for a parton to branch following a Markov process [64]. In the MC generator algorithms, these probabilities are ordered in forward evolution of the outgoing partons (FSR) or backward evolution of incoming partons (ISR) downwards from the scale set by the hard process [62]. The splits are iterated up to a cut-off scale (usually 1 GeV) at which the hadronization starts. Multiple partons in the final states are handled by using two methods: matrix elements calculation and parton showers modeling. The description is as follows.

2.5.1.1 Matrix elements calculation

In the simulation of the hard scattering the generators run across the phase-space, Φ_n , over all X final-states of processes like $2 \rightarrow X$. The fully parton cross-section (Equation 2.10) is given by the product of the square modulus of the corresponding matrix

⁷For example, a gluon can radiate other gluon or even produce quark-antiquark pairs.

element $|\mathcal{M}_{ij \rightarrow X}|^2(\Phi_n; \mu_F^2, \mu_R^2)$, averaged over the spin initial-states. The matrix elements are evaluated as the sum over the Feynman diagrams, $\mathcal{F}^{(k)}$, as

$$\mathcal{M}_{ij \rightarrow X} = \sum_k \mathcal{F}_{ij \rightarrow X}^{(k)}. \quad (2.15)$$

If further degrees of freedom are considered, the MC generator can integrate not only over the phase-space, but also over helicities and color configurations [62]. The use of the matrix-elements calculation provides good description of the hard-scattering cross-section, mainly at high energy scales.

2.5.1.2 Parton showers

The factorial increase of Feynman diagrams with the number of outgoing partons leads to subtleties when calculating the matrix elements. These calculated in fixed-order suffer from divergences in the cases of soft emissions (at very low momentum) and collinear splitting (at very small angles), limiting their usability for predicting cross-sections. Parton-shower models are used to describe the interaction in these phase-space regions. They are still based on pQCD and allow to add an arbitrary number of branchings of one parton into two (or more) partons. Parton showers provide good description of soft and collinear emissions, however hard emissions are poorly approximated.

2.5.1.3 Matching matrix elements with parton showers

The most accurate predictions are often obtained by combining the two methods outlined above. However, care must be taken to avoid double-counting. Strictly speaking, a radiated parton may come either from the matrix elements or the parton showers. The double-counting is evaded by matching the parton shower to the matrix element calculation. The MC generators used in this thesis implement the MLM algorithm [65], which is fixed to perform the matching between a parton (from the matrix elements) and a jet (from the parton shower).

2.5.2 Hadronization and decays

Once a cut-off defined during the parton showering is reached, the hadronization takes place. Colored partons are grouped into color-singlets, following a hadronization model. This is based on phenomenological approaches with parameters tuned to data [66]. The most frequently used models are:

- The Lund Model [67] assumes a linear confinement between partons. The color force among two partons is represented like the tension force in a string. The two extremities are opposite color charges, and the energy increases linearly with the separation between the charges. The string breaks when the separation is very large and two additional color charges ($q\bar{q}$ pair) are produced at the endpoints. The quark and antiquark from two adjacent broken strings can combine to form a meson.

- The cluster Model [68] is based on a pre-confinement property of the angular-ordered parton shower. After the parton showering, all outgoing gluons are split into light $q\bar{q}$ pairs. In the limit of a large number of colors, all quarks and antiquarks can form color-singlet clusters. The high mass clusters split into low mass clusters using string-like mechanism. Cluster splitting remains until all hadrons are built.

2.5.3 Underlying and pile-up events

The MC simulation also includes the so-called *underlying events* (UE). An UE is defined as everything resulting from a pp collision except the hard scattering of two partons. Examples of underlying events are the MI and the beam remnants. Regardless their origin, UE are included in the hadronization process as they might be color-connected to the hard process. The simulation of *pile-up* events is in addition carried out. These events arise when more than one proton-pair collides during the same protons's bunch crossing.

2.5.4 Monte Carlo generators

The Monte Carlo generators can be multi-purpose if they follow all the steps listed previously, or specialized when they develop individual steps. Different MC generators are listed below. The interfaces and matchings are specified, focusing on their usage in the $t\bar{t}$ resonances analysis of this thesis.

PYTHIA [66]: is a multi-purpose generator. It uses $2 \rightarrow X (X \leq 3)$ matrix elements at leading-order, with the rest of the phase-space filled by parton showers. The parton showers are modeled by using a leading-logarithmic approximation with momentum ordering [69]. PYTHIA follows the Lund string model for hadronization and includes its own model for generating UE. The parametrization used for the PDFs is the MRST [36] at leading-order.

HERWIG [70]: is a multi-purpose generator. It uses $2 \rightarrow 2$ matrix elements at leading-order and parton showers with angular ordering. The cluster model is used for hadronization, and in contrast to PYTHIA, UE are generated by implementing the UE-dedicated simulation program JIMMY [71]. As in PYTHIA the MRST PDF set is used for HERWIG.

ALPGEN [72]: is a leading-order matrix elements generator. Unlike PYTHIA, ALPGEN offers multi-parton final-states. The matrix elements go from $2 \rightarrow 2$ up to six partons in the final state. The matrix elements and parton showers are matched by using the MLM scheme. Since parton showers and hadronization are not implemented in the ALPGEN code, it is interfaced to HERWIG. UE are simulated with JIMMY and the PDFs are included from the CTEQ6L1 parametrization [47].

MC@NLO [73, 74]: is a generator specialized in the calculation of the hard process. MC@NLO provides the calculation of the next-to-leading order matrix elements and has to be consistently matched to HERWIG and JIMMY to cover parton showers and hadronization.

POWHEG [75]. In the same fashion as MC@NLO, POWHEG is a next-to-leading order generator. Unlike MC@NLO, POWHEG does not depend on a showering interface. This means that POWHEG can be interfaced to any general-purpose generator, thus enabling hadronization and UE. POWHEG does not generate events with negative weights.

MADGRAPH [48]: is a specialized generator for the production of parton-level events. It calculates the leading-order matrix elements for processes $2 \rightarrow X$ up to six partons in the final state. For the production of the parton showers MADGRAPH is interfaced to PYTHIA.

ACERMC [76]: is a generator specialized in the production of the hard process at leading-order. It has been specifically developed to simulate the SM events of the LHC collisions. ACERMC uses $2 \rightarrow 2$ matrix elements, including spin corrections, and is interfaced with PYTHIA or HERWIG for hadronization and UE. Since ACERMC allows a convenient variation of initial- and final-state radiation, it is commonly used in ATLAS to evaluate systematic uncertainties due to these effects.

3 The ATLAS Detector and simulation

The proton-proton (pp) collision data used in this thesis have been recorded by the ATLAS experiment located at the Large Hadron Collider (LHC), at a center of mass energy of $\sqrt{s} = 7$ TeV. This chapter presents a description of the ATLAS detector and its sub-systems. The experimental setup, design and performance to acquire and process data will be introduced. A brief description of the detector simulation is also outlined.

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [77, 78] is a two-ring-superconducting hadron accelerator with a circumference of 26.65 km. It is located at CERN near Geneva, crossing the border between Switzerland and France. Its depth varies between 50 m and 175 m underground, positioned at the former tunnel of the large electron-positron (e^-e^+) Machine, LEP [79]. The LHC is designed to accelerate protons and produce pp collisions at four locations around the accelerator. The ATLAS detector is based at one of them. The nominal center-of-mass energy for pp collisions is 14 TeV, i.e. 7 TeV per proton beam. The data analyzed in this thesis were collected at a beam energy of 3.5 TeV.

The protons of the LHC are created from hydrogen atoms that are ionized in an electric field. Their acceleration is sequentially conducted in the CERN complex of accelerators shown in Figure 3.1. The linear accelerator LINAC2 provides pulsed 1 Hz beams of protons up to 175 mA at 50 MeV, with pulse lengths varying between 20 and 150 μ s, depending on the required number of protons. The beams are boosted into the Proton Synchrotron Booster ring accelerator (PSB) to increase their energy up to 1.4 GeV. They are then sent to the Proton Synchrotron (PS), where they are further accelerated to 26 GeV [81]. The protons are thereupon injected to the Super Proton Synchrotron (SPS) to reach the minimum energy at which the LHC can maintain a stable beam: 450 GeV [82]. By using two transfer lines, the protons are injected into the two beam pipes of the LHC in clockwise and anti-clockwise directions. The protons arrive at the LHC in bunches. The bunches circulating in the LHC ring contain around 10^{11} protons. They measure a few centimeters long and a millimeter wide when they are far apart from a collision point. However, when they are about to collide, their size is squeezed by quadrupole magnets to 16 μ m in order to increase the probability of pp collisions at high momentum transfer. At nominal conditions, the two beams consist of 2808 bunches. The bunch interval is designed to be ~ 25 ns, although this was held at a minimum of 50 ns for the data collected for this thesis, with a total number of bunches per beam of 1380.

3 The ATLAS Detector and simulation

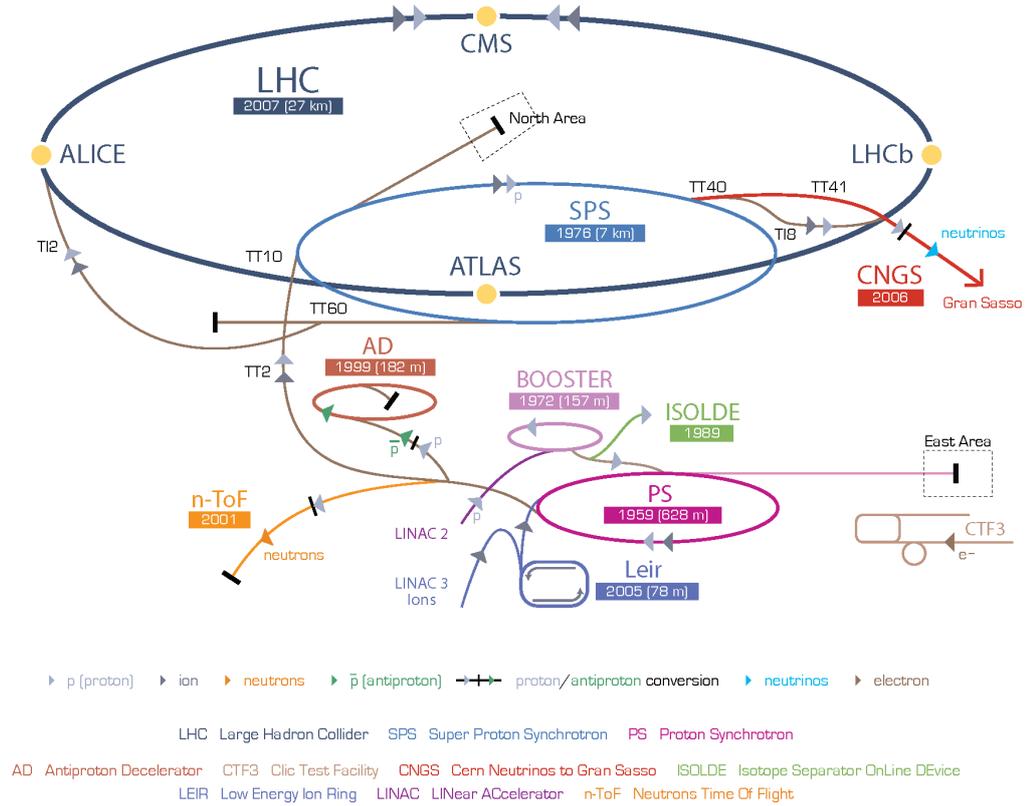


Figure 3.1: The CERN accelerator complex [80].

At each interaction point the proton beams collide at a small crossing angle. The number of collisions per unit time per area is known as the instantaneous luminosity \mathcal{L} . This depends on the LHC running parameters: the accelerator frequency, the number of colliding bunches per beam, the number of protons per bunch and the transverse cross-section of the overlap between beams. The design goal with a bunch spacing of ~ 25 ns implies a frequency of 40 MHz, and an instantaneous luminosity of $10^{34} \text{cm}^{-2} \text{s}^{-1}$. Integrating the instantaneous luminosity over time gives the integrated luminosity \mathcal{L}_{int} , which is the amount of collision data collected. The number of events N of a physics process is calculated as $N = \sigma \cdot \mathcal{L}_{\text{int}}$, where σ is the probability of the process occurring. The head-on collision between two bunches circulating in opposite directions is called bunch crossing. In each bunch crossing one or more pp collisions may simultaneously take place. Multiple proton collisions in the same bunch crossing are referred to as *in-time* pile-up. The number of collisions per bunch crossing, averaged over all bunch crossings in a discrete time is denoted by $\langle \mu \rangle$. The maximum $\langle \mu \rangle$ is a measure of the amount of pile-up in a particular event.

The LHC provides pp collision data to four experiments located at the four interaction points: ATLAS (A Toroidal LHC ApparatuS [83]), CMS (Compact Muon Solenoid [84]), LHCb [85] and ALICE (A Large Ion Collider Experiment [86]).

3.2 The ATLAS Detector

The ATLAS detector is one of the two general-purpose experiments recording LHC collision data. With the high luminosity and center-of-mass energy of the LHC, the ATLAS detector is able to perform precision tests of the Standard Model (SM) as well as searches for physics Beyond the Standard Model. Figure 3.2 is a cut-away view of the ATLAS detector. The detector is about 44 m long, 25 m high and weighs ~ 7000 tonnes. ATLAS aims to observe as much as possible of the many particles resulting from the pp collisions. For this purpose, it has a cylindrical and forward-backward symmetry, covering most of the solid angle surrounding the interaction point. The detector is equipped with successive sub-detectors intended to different purposes: tracking detectors, calorimeters and muon detectors.

Particles produced in the pp collision traverse first the tracking sub-detectors. These form the Inner Detector (ID) system, which is located closest to the beam axis. It comprises three complementary components, which successively enclose each other: the Pixel Detector, the Semiconductor Tracker and the Transition Radiation Tracker. The ID is immersed in a 2 Tesla magnetic field generated by a superconducting solenoid. It allows to measure the momenta of charged particles, by reconstructing the radius of curvature of their tracks. Outside of the solenoid, the next detector is the electromagnetic calorimeter. It is designed to stop electrons and photons as well as to measure their positions and energy deposits. This is carried out by using liquid Argon as active sampling material and lead as the absorber. Partons hadronize in collimated showers of particles known as

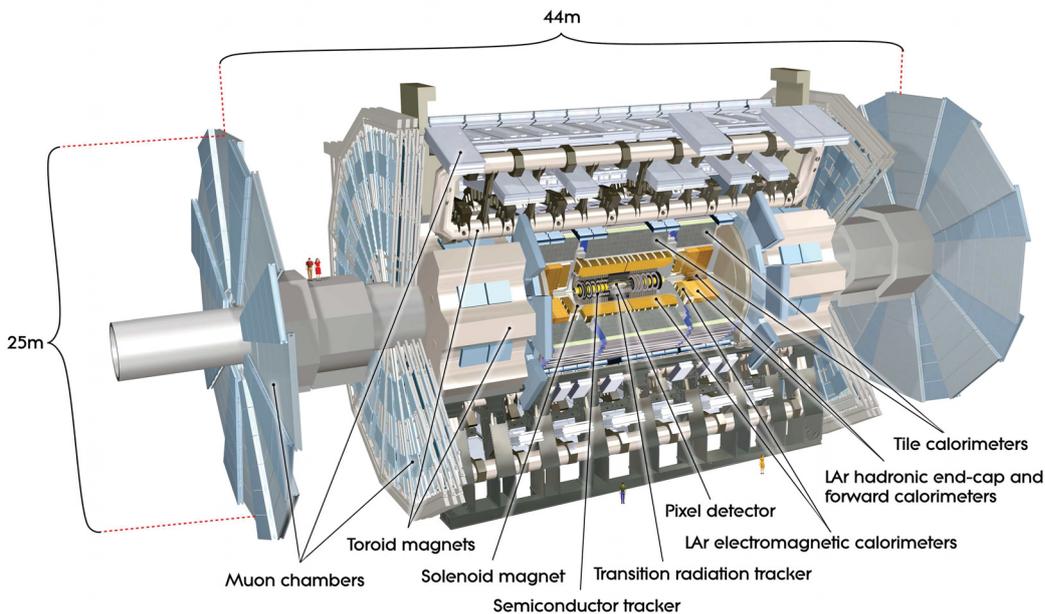


Figure 3.2: The ATLAS detector. The main sub-detectors are identified as well as the overall dimensions [83].

3 The ATLAS Detector and simulation

jets. Jets and other particles with enough energy to not be completely stopped in the electromagnetic calorimeter deposit their remaining energy in the hadronic calorimeter. Its location is outside of the electromagnetic calorimeter. The hadronic calorimeter is made of plastic scintillating tiles and steel in the central region, and liquid Argon in the forward region. Muons penetrate matter more deeply than other particles; their energy loss produces only little deposits in the calorimeters. Since they are not stopped, an additional sub-system beyond the calorimeters is required. The muon chambers are the outermost of the ATLAS sub-detectors. They are placed within a toroidal magnetic field and form the muon spectrometer. This is used for tagging muons and measuring their momenta before they escape the volume of the ATLAS. There are also a set of forward detectors, placed close to the beam pipe and away from the interaction point.

The basic requirements considered when constructing the detector were:

- High detector granularity to reduce the influence of pile-up.
- Good charged-particle momentum resolution and track identification.
- Good electromagnetic calorimetry for electron and photon identification.
- Full-coverage hadronic calorimetry for jet and missing energy measurements.
- High-precision muon momentum measurements.
- Efficient tracking at high luminosity.

The performance of the components of ATLAS have to fulfill the conditions listed in Table 3.1. The ATLAS geometry and sub-detectors are briefly described in the following sections.

3.2.1 The ATLAS Coordinate System

In ATLAS a right-handed coordinate system is used. The origin is located at the center of the detector in the nominal collision point. The z -axis is defined along the anti-clockwise beam line and the x - y plane is transverse to the beam. The positive y -axis points upwards while the positive x -axis points to the center of the LHC ring.

The ATLAS geometry is defined in cylindrical (r, ϕ) coordinates, ϕ being the azimuthal angle measured around the beam axis, and r the distance in the transverse plane defined as

$$r = \sqrt{x^2 + y^2}. \quad (3.1)$$

In the spherical coordinate system the inclination measured from the positive z -axis is given by the polar angle, θ . The azimuth is given by the angular distance ϕ in the transverse plane from the positive x -axis. The polar angle is usually replaced by the *pseudorapidity* (η) defined as

$$\eta = -\ln \left(\tan \frac{\theta}{2} \right). \quad (3.2)$$

Detector component	Required resolution	η coverage	
		Measurement	Trigger
Tracking	$\frac{\sigma_{p_T}}{p_T} = 0.05\% p_T \oplus 1\%$	± 2.5	–
Electromagnetic calorimetry	$\frac{\sigma_E}{E} = 10\%/\sqrt{E} \oplus 0.7\%$	± 3.2	± 2.5
Hadronic calorimetry:			
barrel and end-cap	$\frac{\sigma_E}{E} = 50\%/\sqrt{E} \oplus 3\%$	± 3.2	± 3.2
forward	$\frac{\sigma_E}{E} = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \eta < 4.9$	$3.1 < \eta < 4.9$
Muon spectrometer	$\frac{\sigma_{p_T}}{p_T} = 10\%$ at $p_T = 1\text{TeV}$	± 2.7	± 2.4

Table 3.1: Requirements for the ATLAS detector components. The term \oplus symbolizes addition in quadrature. The units for the energy and momenta are given in GeV [83].

In the transverse xy -plane $\theta = 90^\circ$ and $\eta = 0$. In the positive (negative) z -direction, $|\eta|$ diverges at $\cos\theta = 0$ (180°). The distance in these coordinates is defined as

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \quad (3.3)$$

The transverse momentum and the transverse energy are measured in terms of these coordinates. They are given by: $p_T = p_x^2 + p_y^2$ and $E_T = E \cdot \sin\theta$, respectively. E denotes the energy and is usually combined with the p_z component of the momentum to define the *rapidity* (y):

$$y = 0.5 \times \ln[(E + p_z)/(E - p_z)]. \quad (3.4)$$

For massless particles, the rapidity and pseudorapidity are equal. Since the overall boost of the collision is unknown, the difference Δy is commonly used for the description as it is longitudinal boost-invariant.

The cylindrical geometry of ATLAS leads to the description of its sub-systems as being either part of the *barrel* when they are in the central region, or part of the *end-caps* when they are in the forward regions. The central part of the detector corresponds to small $|\eta|$ values, while the forward direction are those regions close to the beam pipe, at high $|\eta|$ values. The $|\eta|$ coverages for the ATLAS sub-detectors are listed in Table 3.1.

3.2.2 Magnet system

A strong magnetic field bends the tracks of charged particles to measure their momentum and identify the charge. Two different fields are generated by a combined system of four superconducting NbTi magnets. These consists of one solenoid magnet and three toroid

3 The ATLAS Detector and simulation

magnets. The solenoid magnet generates the 2 T axial field immersing the ID. This extends over a length of 5.3 m with a diameter of 2.5 m. The toroid system provides the magnetic field for the muon system. Figure 3.3 illustrates the ATLAS magnet system.

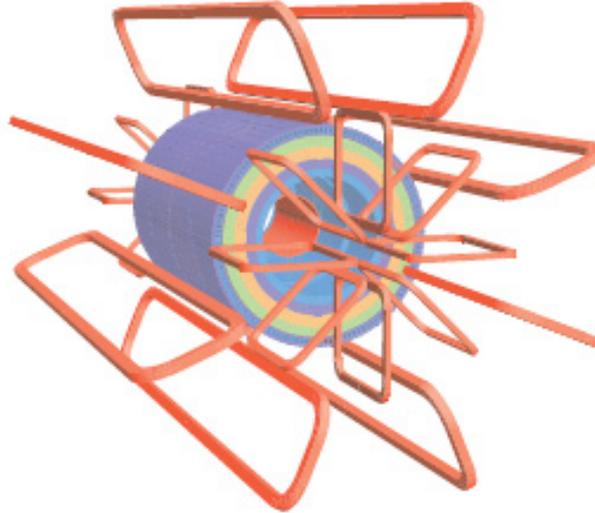


Figure 3.3: The magnetic system of the ATLAS detector. The solenoid is embedded within the calorimeters. The eight barrel toroid coils with the interleaved end-cap coils are visible [83].

In addition, the end-caps of the muon spectrometer are equipped with smaller toroids, whose total magnetic field is non-uniform. They are arranged such that the magnetic field is mostly orthogonal to the particle trajectory. The strengths are approximately 0.5 T in the barrel, and 1 T in the end-caps. An accuracy of 1 mT of the magnetic field strength and 3 mrad in direction is necessary for the required momentum resolution in the muon system.

3.2.3 Inner detector

The inner detector is designed to record the tracks and vertices left by charged particles with $p_T > 0.5$ GeV. Figure 3.4 shows a cut-away view of the ATLAS Inner Detector. To achieve high-precision measurements in the innermost pixel layer, the ID is characterized by its fine granularity. It is crucial for a precise identification of B -hadrons (hadrons containing a bottom quark) and enables the reconstruction of primary and secondary vertices a few millimeter distant from the interaction point.

At small radii, the ATLAS Pixel Detector [83, 87] has the highest granularity of all of the sub-detectors. Designed to operate in the high particle multiplicity of LHC, it is located closest to the beam pipe and therefore closest to the decay vertices of the short-lived

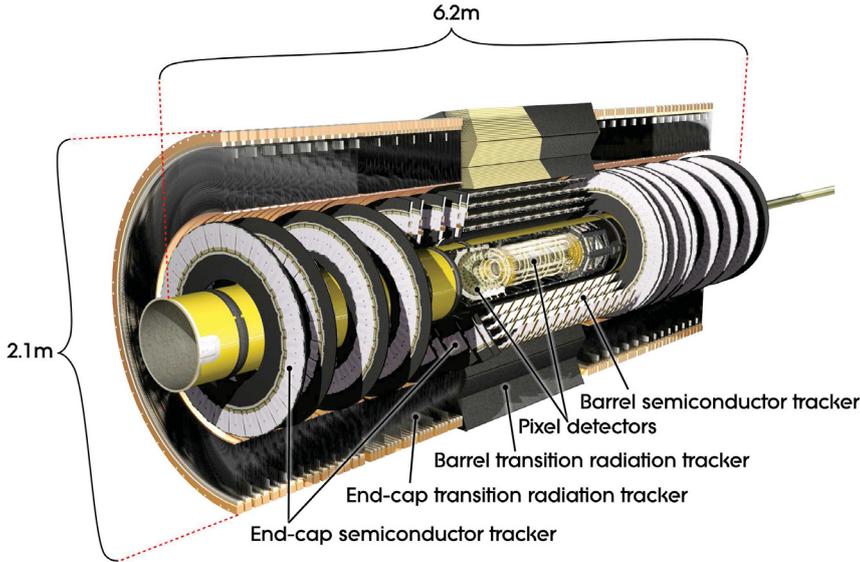


Figure 3.4: Cut-away view of the ATLAS Inner Detector. The layers of sub-detectors are arranged on concentric cylinders around the beam axis [83] .

particles ($|\eta| < 2.5$). The pixel detector consists of three silicon pixel layers in the barrel region and three silicon pixel discs in each of the end-caps. The barrel layers are cylindrical and contain 1456 modules while in the two end-caps there is a total of 288 modules. The barrel and disk modules are identical. Their main components of a module are: a silicon sensor which represents the active part of the module, 16 electronic read-out chips (FEs 18×160 pixels each) whose primary function is to amplify and discriminate the charge collected from the sensor, a module controller chip (MCC) that distributes clock and sends the data collected by the FEs to off-detector read-out drivers, and an interconnection foil (flex) [88]. The sensitive silicon detector area is connected via bump bonds with the front-end chips. The typical pixel size is $50 \times 400 \mu\text{m}^2$ and is defined by the read-out pitch of the front-end chips. To enable full coverage in the regions between front-end chips, approximately 10% of the sensor pixels have a size of $600 \times 50 \mu\text{m}^2$. The pixel layers are segmented in $r - \phi$ and z with typically three pixel layers crossed by each track. The intrinsic accuracies in the barrel are $10 \mu\text{m}$ ($r - \phi$) and $115 \mu\text{m}$ (z), and in the disks are $10 \mu\text{m}$ ($r - \phi$) and $115 \mu\text{m}$ (r). The time resolution is less than the 25 ns collision rate of the LHC and the total number of read-out channels is approximately 80.4 million.

The pixel detector is surrounded by the Semi Conductor Tracker (SCT), which also uses silicon technology. It is built out of four double layers of silicon microstrip detectors with radii between 299 mm and 514 mm and a full length of 1492 mm. The detector uses small-angle (40 mrad) stereo strips to measure both coordinates, with one set of strips in each layer parallel to the beam direction, measuring $r - \phi$. There is a total of 4088 modules typically providing a total of eight strip measurements (four space-points)

3 The ATLAS Detector and simulation

for particles originating in the beam-interaction region. The hit resolution per module is $17\ \mu\text{m}$ ($r - \phi$) and $580\ \mu\text{m}$ (z) in the barrel, and of $17\ \mu\text{m}$ ($r - \phi$) and $580\ \mu\text{m}$ (r) in the end-caps. The total number of read-out channels in the SCT is approximately 6.3 million.

The outermost part of the Inner Detector is the Transition Radiation Tracker (TRT). With a large number of hits, typically 36 hits per track, the TRT provides the most continuous tracking of charged particles in a radial covering up to 1066 mm. It is composed of polyimide straw tubes (4 mm diameter each), which allow electron identification by the transition radiation. The straw tubes are aligned with the beam axis in the barrel section and radially in the two end-caps. In the barrel region, the straws are parallel to the beam axis and are 144 cm long, with their wires divided into two halves, approximately at $\eta = 0$. In the end-cap region, long straws (37 cm) are arranged radially in wheels. The straw tubes are filled with a gas mixture of Xe (70%), CO₂ (27%) and O₂ (3%). A gold-plated tungsten anode wire of $31\ \mu\text{m}$ diameter runs along the tube axis. When a high voltage is applied between the tube and the wire, an electric field is induced. The passage of a charged particle ionizes the gas-filled area, forming pairs of electrons and ions. The electrons drift towards the wire, ionizing other atoms on their way and produce an avalanche multiplication that amplifies the signal and allows its read-out. The charge collected on the wire is proportional to the energy loss of the primary ionizing particle and the arrival time of the signal defines the drift radius of the electrons and thus the track position of the primary particle in the $r - \phi$ plane. The tubes together with the gas mixture give 48 ns maximum collection time of the signal. The TRT has an intrinsic accuracy of $130\ \mu\text{m}$ per straw. The total number of TRT readout channels is approximately 351,000. Figure 3.5 shows the three components of the ID traversed by a charged track of $p_T = 10\ \text{GeV}$, at $\eta = 0.3$. The sensors and structures of each component are also illustrated.

3.2.4 Calorimeter system

The ATLAS Calorimeter system measures the energy of particles interacting with the detector via the electromagnetic and/or strong interactions. The calorimeters are composed by sampling detectors with full ϕ -symmetry and coverage around the beam axis ($|\eta| < 4.9$). The calorimeters closest to the beam-line are housed in three cryostats: one barrel and two end-caps. The barrel cryostat contains the electromagnetic barrel calorimeter. Each of the two end-cap cryostats contain an electromagnetic end-cap calorimeter (EMEC), a hadronic end-cap calorimeter (HEC), located behind the EMEC, and a forward calorimeter (FCal) to cover the region closest to the beam line. All the calorimeters employ liquid Argon as active detector material. This was chosen because its stability, read-out speed and its intrinsic radiation-hardness. Lead is used as absorber for the electromagnetic calorimeter, Copper for the hadronic end-cap calorimeters and Copper-tungsten for the forward calorimeter. The outer hadronic calorimeter uses scintillator tiles as the sampling medium and steel as the absorber medium. A schematic view of the components is shown in Figure 3.6.

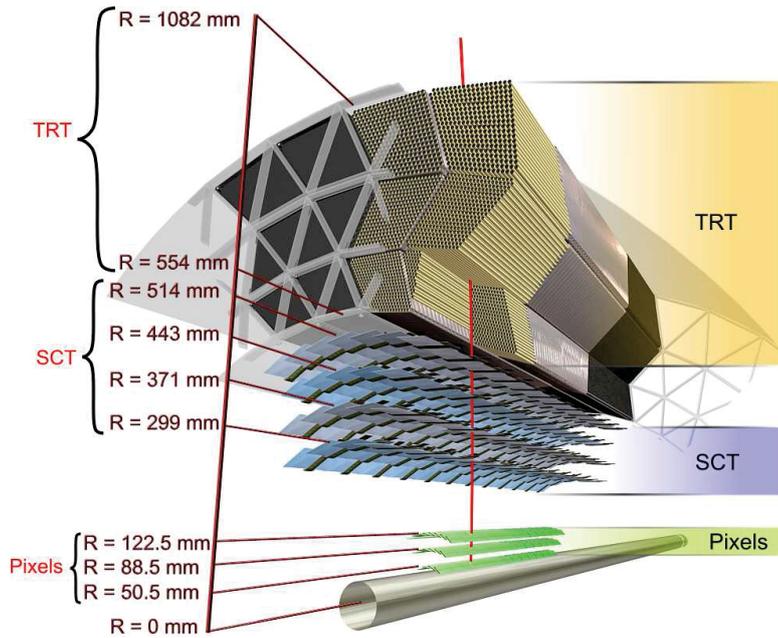


Figure 3.5: Layers-structure of the ATLAS Inner Detector. The red line is a charged track of $p_T = 10$ GeV crossing the ID layers: the three cylindrical silicon-pixel layers, the four cylindrical double of the barrel silicon-microstrip sensors (SCT) and the transition-radiation tracker modules (TRT) [83].

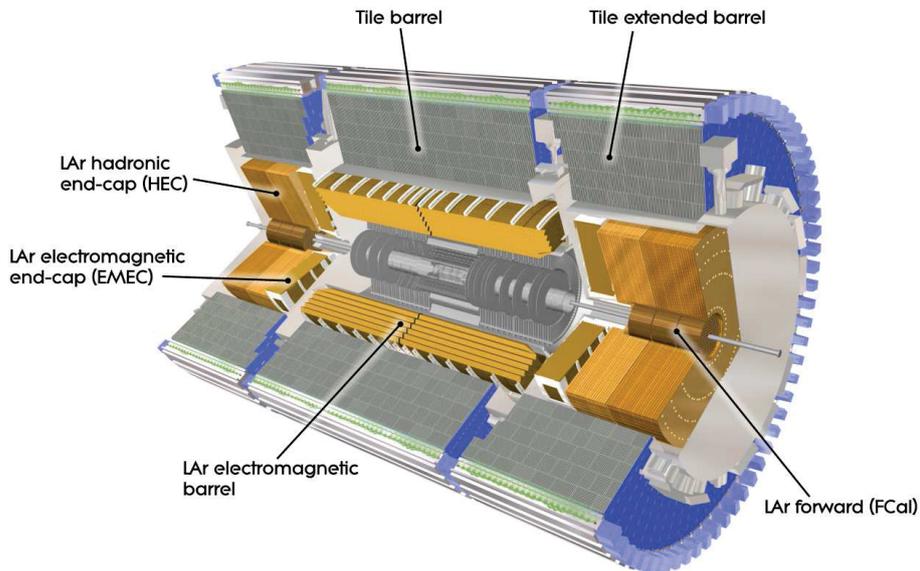


Figure 3.6: Cut-away view of the calorimeter system [83].

3 The ATLAS Detector and simulation

An accordion geometry for the absorbers and the electrodes allows the liquid Argon calorimeters to have several active layers in depth, three in the central region $0 < |\eta| < 2.5$ and two in the forward region $2.5 < |\eta| < 3.2$ and in the overlap region $1.35 < |\eta| < 1.475$ between the barrel and the end-caps. Such a geometry provides naturally a full coverage in ϕ without gaps a fast extraction of the signal at the electrodes.

The maximum granularity of the electromagnetic calorimeter is $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ at the depth at which the largest number of secondary particles are produced. This allows for very precise measurements of the EM shower position. The barrel electromagnetic calorimeter consists of two identical half barrels, separated by a small gap of 4 mm at $z = 0$. The length of each half-barrel is 3.2 m, and their inner and outer diameters are 2.8 m and 4 m, respectively.

The LAr alternating layers have different functionality. The first layer is finely segmented along η in order to measure the direction of particles. The second layer collects the largest fraction of the energy of the electromagnetic shower, and the third layer collects only the tail of the EM shower and is therefore less segmented in η . The third layer also performs trigger on energy deposits from the EM showers [83]. Figure 3.7 shows the specific layers of the electromagnetic calorimeter, together with their $\eta - \phi$ granularity. In the figure, X_0 is the radiation length. It defines the average distance over which the energy of an electron is reduced by a factor of $1/e$ due to radiation losses.

In the central region, $0 < |\eta| < 1.8$, the electromagnetic calorimeter is complemented by a presampler detector. It is an instrumented Argon layer with 11 mm in depth, placed in front of the barrel. The presampler detector is used to measure the energy lost by charged particles in front of the electromagnetic calorimeters. The EMEC calorimeter consists of two wheels, one on each side of the electromagnetic barrel. Each wheel is 63 cm thick and weighs 27 tones, has an external and internal radii of 2098 mm and 330 mm, respectively. It covers the region $1.375 < |\eta| < 3.2$. In the transition region between the barrel and the end-cap calorimeters where the material in front of the calorimeter amounts to several interaction lengths, a LAr presampler detector is implemented. It covers the region $1.5 < |\eta| < 1.8$. The total thickness of the electromagnetic calorimeter is more than 24 (26) radiation lengths in the barrel (end-caps). Front-end boards (FEBs) contain the electronics for amplifying, shaping and digitizing the detector signals. The energy resolution of the electromagnetic calorimeter is $\sigma_E/E = 10\%/\sqrt{E} [\text{GeV}] \oplus 0.7\%$.

There is a “crack” region located at $1.375 < |\eta| < 1.52$. This is a gap between the barrel and the end-cap calorimeters, filled with cables and services for the inner detector and the electromagnetic calorimeter. Since this region can be only partially instrumented, this analysis in this thesis discards events with electrons falling in the crack region.

The hadronic tile calorimeter is a sampling calorimeter located in the region $|\eta| < 1.7$, surrounding the electromagnetic calorimeter. It is divided into a central barrel, 5.8 m in length, and two extended barrels, 2.6 m in length, each having an inner radius of 2.28 m and an outer radius of 4.25 m. The radial depth of the tile calorimeter is approximately 7.4 interaction lengths in the barrel and 10 interaction lengths in the

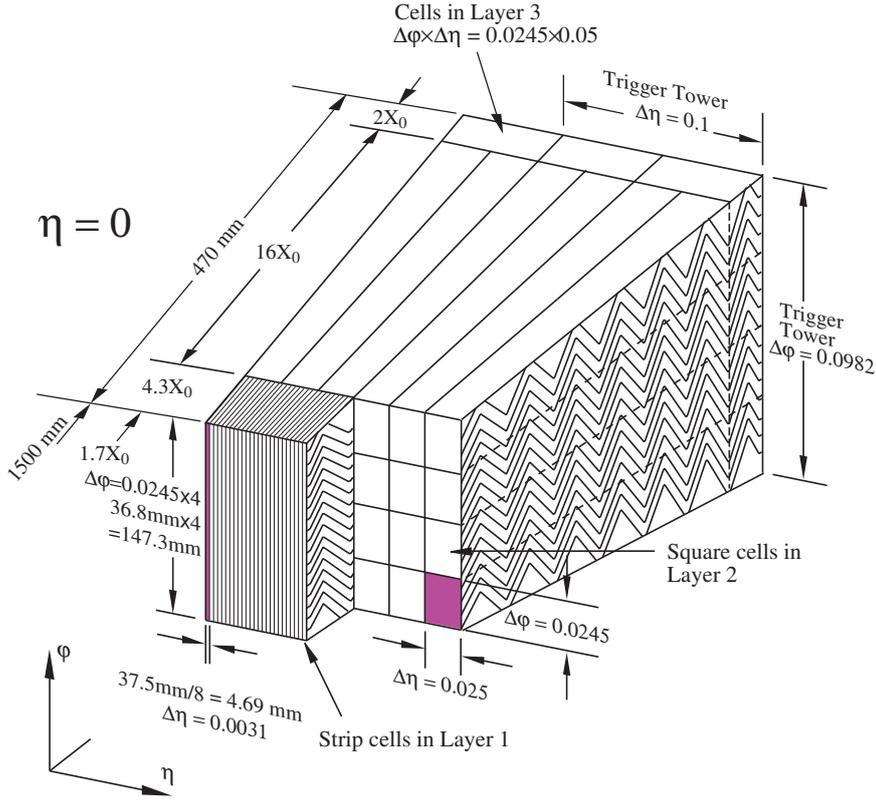


Figure 3.7: The three LAr sampling layers of the ATLAS electromagnetic calorimeter. An electromagnetic shower spreads from layer 1 up to $4.3X_0$, most of which then propagates through layer 2 up to $16X_0$. Only a small fraction of the shower arrives up to layer 3 [83].

end-caps¹. Each barrel comprises 64 modules or wedges of size $\Delta\phi = 5.625^\circ$, made of steel plates and scintillating tiles. The assembled module forms an almost-periodic steel-scintillator structure with a ratio by volume of approximately 4.7:1. The hadronic end-cap calorimeter covers the range $1.5 < |\eta| < 3.2$, overlapping with the tile and forward calorimeters to reduce the drop in material density in the crack region. The hadronic end-cap calorimeter is a copper/liquid-Argon sampling calorimeter with a flat-plate design, which covers the range $1.5 < |\eta| < 3.2$. It consists of two wheels per end-cap and 32 identical modules per wheel. Each wheel is divided into two segments in depth, for a total of four layers per end-cap. The wheels closest to the interaction point are built from 25 mm parallel copper plates, while those further away use 50 mm copper plates. The former are interspersed with active material made of 8.5 mm LAr. The outer radius of the copper plates is 2.03 m, while the inner radius is 0.475 m. The energy resolution for the barrel and end-cap calorimeter is $\sigma_E/E = 50\%/\sqrt{E [\text{GeV}]} \oplus 3\%$.

¹In general, the hadronic interaction length defines the average distance that a proton travels without having a hadronic interaction.

3 The ATLAS Detector and simulation

The forward calorimeter is placed in the region $3.1 < |\eta| < 4.9$. It consists of three modules in each end-cap. The first module is made of Copper and optimized for electromagnetic measurements. The other two are made of tungsten and measure mainly the energy of hadronic interactions. The energy resolution of the forward calorimeter is $\sigma_E/E = 100\%/\sqrt{E} [\text{GeV}] \oplus 10\%$.

In general, the granularity of the hadronic calorimeter is coarser than in the EM calorimeter. For $|\eta| < 2.5$, $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$, and in the sampling calorimeter region $\Delta\eta \times \Delta\phi = 0.2 \times 0.1$. The tiles are read out at two sides using wavelength shifting fibres into separate photomultiplier tubes.

3.2.5 Muon System

The Muon Spectrometer (MS) shown in Figure 3.8 is the outermost part of the ATLAS detector, and hence covers the largest surface. It is designed to provide a measurement of muon momentum, with a resolution of around 4 (12)% for muons with a p_T of 100 (1000) GeV. The momentum is reconstructed by measuring muon tracks, which are deflected by a magnetic field described in Section 3.2.2. The barrel toroid provides the magnetic field in the region $0 < |\eta| < 1.4$ and the end-cap toroids in the range of $1.6 < |\eta| < 2.7$. In the transition region where the two systems overlap, the bending power is lower. The barrel contains eight racetrack-shaped coils encased in individual stainless-steel vacuum vessels, and supported by rings of struts. The overall size is 25.3 m in length, with inner and outer diameters of 9.4 m and 20.1 m, respectively. Each end-

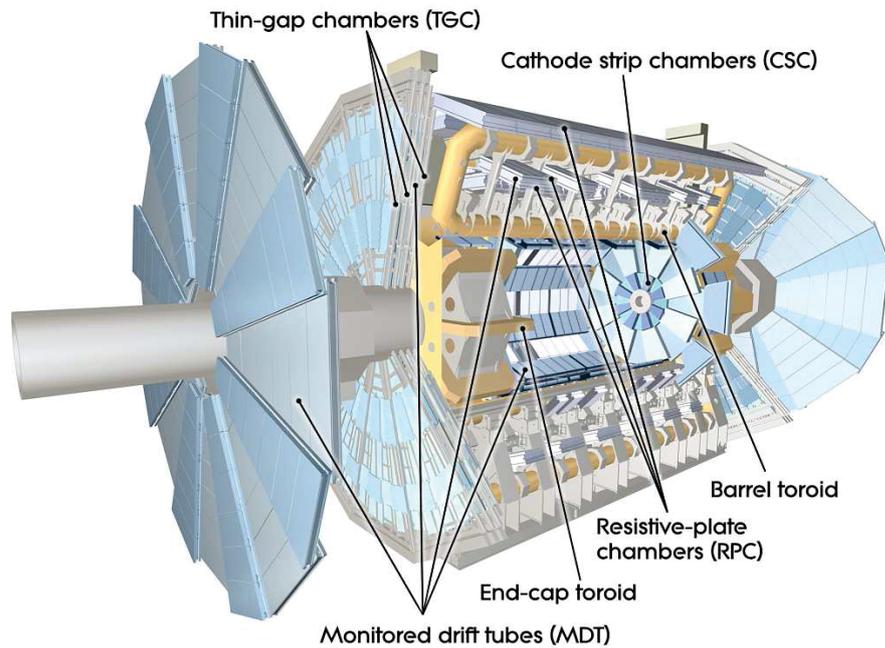


Figure 3.8: Cut-away view of the ATLAS muon system [83].

cap toroid is made of a single cold mass build up from eight flat, square coil units and eight keystone wedges. The length of end-cap toroids is 5.0 m, the inner diameter is 1.65 m and the outer diameter is 10.7 m.

The muon trajectory is measured in the η -direction at three stations. Over most of η range the track measurement is performed by the Monitored Drift Tubes (MDT), which have a precision higher than $100\ \mu\text{m}$. The MDT are multilayers made of aluminum pressurized drift tubes. The tubes contain a tungsten-rhenium wire of 3 cm diameter and are filled with a mixture of CO_2 (7%) and Ar (93%) acting as a drift gas. A high voltage is applied between the wire and the tube and an electric field is created between them. The muons pass through the tubes and ionize the gas. The stripped electrons drift to the wire generating an avalanche multiplication proportional to the energy deposited by the muon. The maximum drift time is about 700 ns and the single wire resolution is about $80\ \mu\text{m}$.

Due to a high particle flux in the transition region and in the end-caps, the cathode strip chambers (CSC) have been installed in the region $2 < |\eta| < 2.7$. The CSCs have a higher rate capability, higher granularity and time resolution than the MDTs. The CSCs are multi-wire proportional chambers with segmented cathode strips and a wire spacing of 2.5 mm. The wires can undergo voltages up to 1900 V and are interspersed with strip cathodes filled with a gas mixture of CO_2 (20%) and Ar (80%). The maximum drift time is approximately 30 ns. The measurement of the charge induced on the cathode by the avalanche formed on the anode wire gives the coordinates. The θ -coordinate is obtained by arranging the strips of one cathode perpendicularly to the radial anode wires. A second set of cathode strips parallel to the wires provide the second spatial coordinate ϕ . The resolution in the measure of the coordinates reaches 1 cm.

Since the MDT and CSC have large drift times to be used in the trigger (see next Section), the muon system also has a dedicated trigger system covering a region up to $|\eta| = 2.4$. The trigger system is build up of resistive plate chambers (RPCs) in the barrel thin gap chambers (TGCs) in the end-caps. It measures the muon tracks in two orthogonal projections with a spatial resolution of around 1 cm and temporal resolution of 1 ns. The RPCs are gaseous parallel electrode plate detectors forming three concentric cylindrical layers around the beam axis at radii of about 5, 7.5, and 10 m. The large distance between the inner and outer RPC is used for the selection of high momentum tracks of (up to 35 GeV), while the two inner chambers provide low p_{T} -values (below 9 GeV). The TGC are multi-wire proportional chambers. They are arranged in four disks at distances of 7, 10, 14, and 21-23 m from the interaction point, concentric with the beam axis. The spatial coverage of the trigger system is about 99% in the end-cap regions and 80% in the barrel region.

3.2.6 Forward detectors

There are complementary detectors that provide a good coverage in the very forward region. These are used for additional physics measurements or monitoring purposes, and

3 The ATLAS Detector and simulation

have significantly different acceptance and response to pile-up. The description of the forward detectors relevant in this work is given in the following.

Closest to the ATLAS detector, at the distance of ± 17 m from the interaction point, a Cherenkov detector called LUCID (Luminosity measurement using Cherenkov Integrating Detector [89]) is located. Inelastic pp scattering is detected in the forward direction, using the number of particles passing as a measure of the luminosity. The detector is made of 20 aluminum tubes with a length of 1.5 m and 15 mm diameter each, surrounding the beam axis and pointing to the interaction point. Each tube is filled with gas C_4F_{10} under constant pressure, inducing particles passing through the tubes to emit a cone of Cherenkov light. The Cherenkov radiation is then led to the other end of the tubes where a photo-multiplier is located to read out and collect the signal.

The Beam Condition Monitor (BCM) [90] consists of two sets of diamonds sensors located at a distance of ± 184 cm from the interaction point and at $r = 5.5$ cm. It provides a bunch-by-bunch luminosity measurement by counting single-sided and coincidence rates as a function of bunch crossing. This detector is designed to monitor the beam background level to issue a beam abort in case of beam losses.

The Zero-Degree Calorimeter (ZDC) is located in the very forward region just after the LHC beam-pipe splits in two, at a distance of ± 140 m from the collision point. This detector offers other luminosity measurements but it is mainly aimed to determine the centrality and forward neutrons in heavy-ion collisions. It consists of interspersed quartz rods and tungsten plates for pseudorapidity measurements up to $|\eta| = 8.2$.

The Minimum Bias Trigger Scintillators (MBTS) [91] consist of two scintillator wheels with 32 counters mounted in front of the calorimeter end-caps, covering the region $2.1 < |\eta| < 3.8$. In a collision event during a bunch crossing, wavelength-shifting fibers collect light produced in each scintillator and transport it to the photomultiplier tubes for signal amplification. The signals out of the photomultiplier tubes are processed using the tile calorimeter read-out electronics. They are shaped and amplified, and a threshold of 50 mV is used to define a hit. The MBTS has been designed to detect collision activity in the early operation of ATLAS. It is also used for the luminosity determination and in this thesis is a reference to calculate trigger efficiencies.

3.3 Experiment performance

The LHC ran at $\sqrt{s} = 7$ TeV throughout the year 2010 delivering a total integrated luminosity of $\mathcal{L} = 48 \text{ pb}^{-1}$, of which ATLAS recorded $\mathcal{L} = 45 \text{ pb}^{-1}$ [91, 92] (see Figure 3.9a). During 2011, a maximum of 1380 bunch pairs were collided with a spacing of 50 ns. The maximum peak of luminosity reached was $3.65 \cdot 10^{33} \text{ cm}^{-2}\text{s}^{-1}$ (see Figure 3.9c). At the end of 2011 a total of $\mathcal{L} = 5.25 \text{ fb}^{-1}$ was recorded by ATLAS, as shown in Figure 3.9b. This corresponds to a data taking efficiency of 95% of the detector [94].

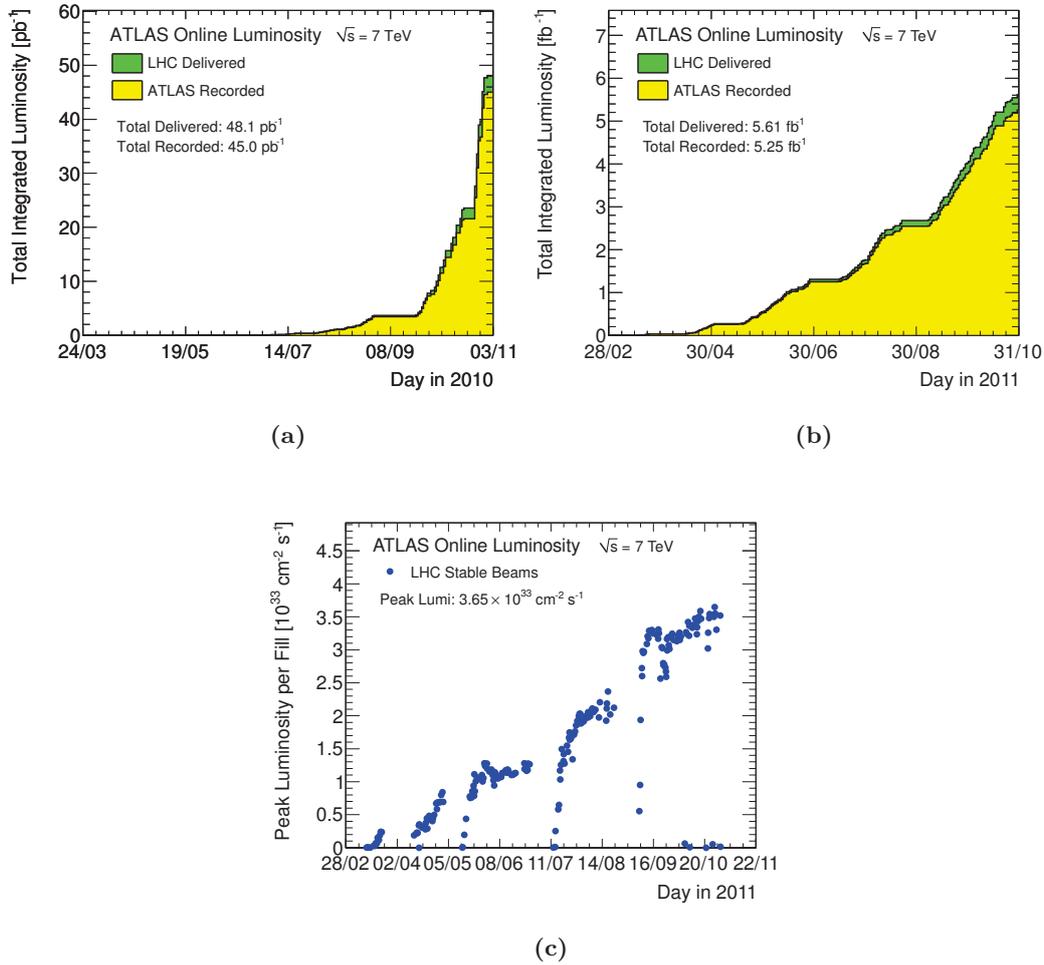


Figure 3.9: Cumulative luminosity versus day delivered by the LHC (green), and recorded by ATLAS (yellow) during stable beams and for pp collisions at $\sqrt{s} = 7$ TeV in (a) 2010 and (b) 2011, and (c) the maximum instantaneous luminosity versus day delivered to ATLAS [93].

The record luminosity of $7.73 \cdot 10^{33} \text{ cm}^{-2} \text{ s}^{-1}$ was achieved during August 2012. This contributed to a total integrated luminosity recorded by ATLAS of $\mathcal{L} = 21.7 \text{ fb}^{-1}$ at $\sqrt{s} = 8$ TeV. The luminosity is determined by measuring the observed interaction rate per bunch crossing. The calibration of the luminosity detectors is obtained by beam-separation techniques, using van-der-Meer scans [95]. The beam profile is measured by moving the beams horizontally and vertically against each other. This profile gives all the running parameters needed for calculating the luminosity mentioned in Section 3.1.

3.3.1 Pile-up

Pile-up interactions can be “in-time” pile-up when the collisions occur within the same bunch crossing, or “out-of-time” pile-up when they come from the previous bunch crossing due to the short bunch spacing. The probability of having inelastic pp interactions per bunch crossing μ increases proportionally with the luminosity. As shown in Figures 3.9a and 3.9b, the excellent performance of the LHC in 2010 and 2011 produced more than $\mathcal{L} = 5 \text{ fb}^{-1}$ of integrated luminosity. As a consequence, the number of pile-up events have, on average, exceeded 3 pp interactions at the end of 2010 and 18 in 2011. It can be seen in Figure 3.10. The pile-up identification through primary vertex reconstruction is complementary to the corrections applied when hadronic jets are identify.

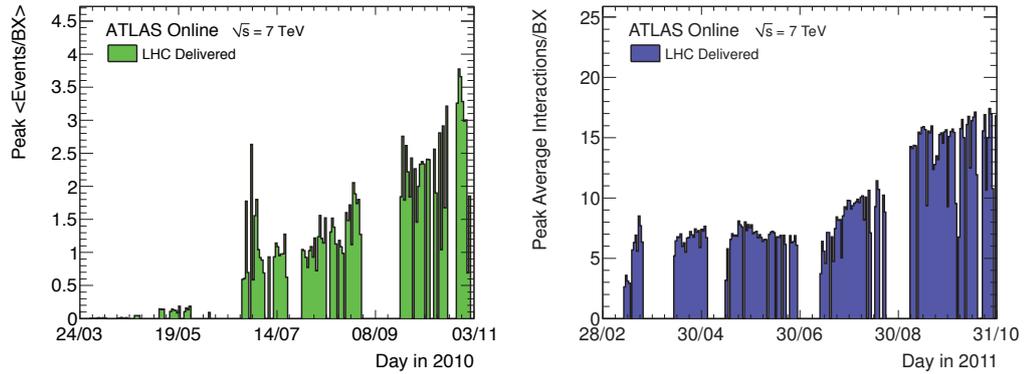


Figure 3.10: Mean number of interactions per bunch crossing at the peak of the fill for each day in 2010 (left) and 2011 (right). The number of events per bunch crossing is averaged over a short time period. [96]

3.4 Trigger and data acquisition

The incoming interaction rate at the design luminosity is about 1 GHz and the event size is, on average $\sim 1.3 \text{ MB}$ [97]. With the current technology, the input to be stored by ATLAS exceeds any acquisition capacity. Moreover, the vast majority of events product of the LHC pp collisions are uninteresting since they are known physics. Figure 3.11 shows the production cross-section and rates of some processes at the LHC’s design energy and luminosity. The primary purpose of the ATLAS trigger system is to select those interesting events, reducing the input from the order of GHz to a rate of $\sim 200 \text{ Hz}$; this is the rate at which events can be written to disk. This limit corresponds to an average data rate of $\approx 300 - 500 \text{ MB/s}$ and acted upon the computing resources available for storage and processing of the data. To reduce the rates, an overall rejection factor of 5×10^6 is required. Such a reduction is achieved by the trigger using a set of hardware/software implementations. It consists of a three-levels chain (level-1 , level-2 and Event Filter),

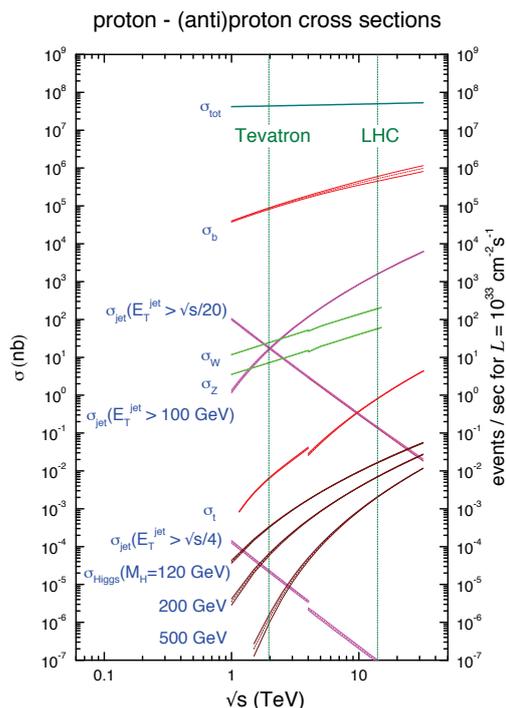


Figure 3.11: Some of the relevant cross sections as a function of center-of-mass energy in proton-antiproton (left) and proton-proton (right) colliders [80]. The dashed lines correspond to the design center-of-mass energy at the TEVATRON and the LHC. The cross-section is measured in nb and events per second at the LHC design instantaneous luminosity (beam intensity).

that selects events by rapidly identifying trigger signatures above a certain threshold. The signatures are physics-object candidates, e.g. electrons, photons, muons, jets, jets with b -flavor tagging, and so on. In addition, there are triggers for inelastic pp collisions (MinBias) and triggers based on global event properties such as missing transverse energy (E_T^{miss}) and summed transverse energy ($\sum E_T$). Only events passing established ATLAS trigger chains are stored to subsequently be analyzed.

3.4.1 Level-1 trigger

The level-1 (L1) system receives all the collision data and takes trigger decision within $2.5 \mu\text{s}$. The aim is to reduce the initial rate to a maximum of 75 KHz. The trigger decision is based on the multiplicities and thresholds for the physics objects reconstructed by the level-1 trigger algorithms. The collection of total energy and transverse momentum thresholds (called trigger items) are defined in a trigger menu, which is located in the central trigger processor (CTP). The trigger items have conditions that must be

satisfied by the trigger signatures when firing the trigger². To achieve quick response, the level-1 is implemented in fast custom electronics allowing to accept no more than 75000 events per second. The level-1 selection is based on reduced-granularity information from the calorimeters and the muon sub-detectors. Because of its complexity and processing speed, no ID information is used. Besides applying multiplicity requirements, the CTP implements pre-scale factors configured through the trigger menu. The pre-scales determine how often an event that has passed a particular trigger item should be saved. For instance, a pre-scale of 100 indicates that events passing a trigger condition will only be accepted one out every hundred times. A trigger item is pre-scaled in order to control the rate of stored events passing this particular trigger.

The level-1 trigger decision is distributed according to the timing and control signal regions through the ATLAS sub-detectors. The L1-calorimeter trigger obtains inputs from the read-outs of the electromagnetic and hadronic calorimeters in a region $|\eta| < 4.9$. Thus, the localized signatures to be triggered by the L1-calorimeter trigger are electrons, photons, τ -leptons and jets. The L1-muon trigger system processes information read-out from the muon detectors and identifies the bunch crossing in which they were produced. It mainly selects muon candidates passing trigger items with imposed thresholds in the transverse momentum above 6 GeV. The whole level-1 system identifies *Regions of interest* (RoIs) within the (η, ϕ) space of the detector. The RoIs are selected according to an energy threshold. Based on a required multiplicity of RoIs, the level-1 decides whether the event passes the trigger item or not. If so, the event is then processed and sent to the High Level Trigger (HLT) composed by the level-2 trigger and the Event Filter

3.4.2 Level-2 Trigger

The software-based level-2 trigger system (L2) is designed to reduce the trigger event rates from 75 kHz to 3.5 KHz in an average latency of 40 ms. It consists of farms of commodity processors connected by fast dedicated networks. Events accepted by the level-1 trigger decision activate data sent from each sub-detector to the specific read-out buffers (ROB) in which the event is partially stored while the level-2 decision is taken. One or more ROBs are connected to read-out systems (ROS) sending signals to the level-2 networks. The level-2 trigger receives the position of the RoIs to determine their read-outs servers and continues the analysis of the event using more complex algorithms. Explicitly, the level-2 processors receive data from ROS corresponding to detector elements inside each RoI, and take a decision which is sent to data-flow managers. These distribute the data from the read-out servers to the Event Builder in which each data is processed until reaching a full event structure. The event is then sent to the Event Filter for the final trigger analysis.

²If the event passes an item, it is said to have fired the trigger.

3.4.3 Event Filter

The software-based Event Filter (EF) realizes the final trigger selection, reducing the output rates from 3.5 KHz to 200-300 Hz. The EF receives full event information from the ROBs to be processed using offline algorithms, i.e. those run after the data taken. Each event is on average processed by the EF in ~ 4 s. Any event with an EF processing time exceeding 180 s is recorded as a *timeout* event. This is separately written to be analyzed and in some cases, recovered.

The event filter classifies selected events based on the trigger items they have been accepted by. Events are thus written to inclusive data *streams* designed to classify them in different datasets. There are four primary streams for physics analyses: **Egamma** (for photon and electron triggers), **Muons** (for muon triggers), **JetTauEtmiss** (for jet, tau-lepton and missing transverse energy triggers) and **MinBias** (for minimum-bias and random triggers). If one event fires, for example, both a jet and electron triggers, the event will be written to both, the **JetTauEtmiss** and the **Egamma** streams. In addition, there are several calibration streams in which only partial information from one or more sub-detectors is written. For instance, the debug stream that stores events that are processed in the event filter during more than 5 s and do not reach a physics stream (e.g. *timeout* events). They are kept to possibly be recovered for offline analysis (reprocessing).

3.4.4 Trigger Menu

The trigger system is configured via the trigger menu. It defines the sequential trigger chain, starting from a level-1 towards the HLT. The trigger menu specifies the trigger items composing each trigger chain, as well as their parameters, i.e. the total energy or transverse momentum thresholds. Most of the trigger items define the object multiplicity required in the event final-states. Thus, different classes of triggers can be found in the trigger menu:

Single triggers: used for final states with at least one object of the same type, e.g. single-jet triggers.

Multiple object triggers: used for final states with two or more objects of the same trigger-type. Multi-jet triggers for instance, are designed to two or more jets in the final state.

Combined triggers: composed of two or more signatures of different trigger-types.

The trigger system can take input from the luminosity detectors described in Section 3.2.6. In particular, the CTP receives signals from the MBTS and if one event is recorded by the level-1, it becomes available for further processing and analysis. Trigger chains dedicated to select these events are included in the trigger menu as MBTS triggers.

In addition to the pre-scale factors, the trigger rates can be controlled by modifying the thresholds at any trigger level, or applying different sets of selection cuts at HLT. Applied to a given trigger object, these cuts can be **loose**, **medium**, and **tight**, depending on their severity level. In the trigger menu, they are suffixed to the trigger name, e.g. as **e20_tight**. Here, **e** labels one **egamma** trigger object with a 20 GeV threshold in

transverse energy. Object-isolation requirements can also be imposed to reduce the trigger rates. They are indicated in the trigger menu by an *i* appended to the trigger name, e.g. `e20i_tight`.

3.5 Data Quality

The data recorded by the ATLAS detector are organized in run periods, runs and luminosity blocks. A luminosity block is defined as the accumulated data in a short period of time in which the detector is stable. A new luminosity block starts when the trigger pre-scales change. A run is a collection of luminosity blocks corresponding to continuous data taken. It starts after the LHC beams have been accelerated to the nominal center-of-mass energy and declared stable. The run's end is given hours later, shortly before a scheduled beam dump. However, problems with the detector or the data-taking infrastructure can force the run to be aborted. When runs are brought together according to constant conditions in the detector, accelerator and trigger systems, they are collected in run periods. The declaration of the a new data-taking period indicates that there has been a major change with non-negligible impact on the detector performance.

The performance of each ATLAS sub-detector is individually monitored. To guarantee that the data are taken under full detector functionality and quality, suitable runs for analysis are contained in the so-called *GoodRunsLists* (GRL). Runs included in the GRL are collision data originated only from stable beams at the LHC.

The mechanism to inspect the performance of each sub-system in ATLAS is through Data Quality (DQ) flags. They are issued for every location where the ATLAS sub-detectors exist, usually segmented per barrel ($|\eta| < 1.5$), end-caps (for the calorimeter, $1.5 < |\eta| < 3.2$), and forward (for the calorimeter, $3.2 < |\eta| < 4.9$). DQ flags are also assigned to trigger streams and to each physics object to be reconstructed. The outputs of the DQ flags are saved per luminosity block (LB). The DQ flag per LB is part of the information used to classify lists of good runs/luminosity blocks.

3.6 ATLAS simulation and data management

The detector simulation is the simulation of interactions between the final-state particles generated in the collision, and the detector. The digitization is the simulation of the detector read-out, i.e. the conversion of energy deposited in the detector to times, currents and voltages for read-out electronics (see Section 2.5). The entire ATLAS simulation system is built within the ATHENA framework [98], as an integration of many processes comprising the simulation stages: event generation (see Section 2.5), detector simulation and digitization. At the first, ATHENA is interfaced with more than 40 Monte Carlo generators. The events generated are then taken as input to simulate the propagation of the products of the interaction through the detector layers. The detector simulation follows the time-history of the particles, and is performed within ATHENA using GEANT4 [99]. The detector description is taken from a geometry database that is common to simulation and reconstruction (described in the next chapter). The database includes variants

of the detector, e.g. extra material and sub-detector misalignments that are regularly updated. All the tunable parameters in GEANT4 have been fixed according to the results from test-beam analyses. At the final stage, the digitization software of ATLAS transforms the hits into detector response, taking into account all the sub-detectors, their inherent inefficiencies and resolutions. The electronic noise and channel-dependent variations are also considered, using a database of run-dependent conditions. In addition, the simulation of the signal events is overlaid with pile-up events, cavern background and cosmic rays.

The output of the simulation is provided at three levels. The parton level refers to the partons at the matrix-element level and fragmentation modeling. The truth (or particle) level are the physics objects reconstructed after the fragmentation and before any detector simulation. The reconstruction level refers to the physics objects reconstructed from the detector signals.

The output format of the simulation is identical to the real detector output format, and is run through the same trigger and reconstruction algorithms. The data and Monte Carlo samples used in this analysis are constructed with ATHENA release 15.6 (16.6), implemented in the data taking of the year 2010 (2011). The official data format of Athena is the AOD (Analysis Object Data). Each AOD stores thousands of variables and can reach the size of several gigabytes. Since a physics analysis needs to use hundreds of AODs, many ATLAS Working Groups have decided to conveniently derive smaller samples (D3PD) containing only the relevant variables of filtered events passing a basic selection. The performance studies presented in Chapter 5 have been developed using both data formats.

The D3PDs used in Chapter 6 were privately produced, and also centrally provided by the ATLAS Top Working Group. The data were analyzed with a software framework SFrame³.

Monte Carlo samples are produced before or during a data-taking period. When the simulated conditions do not reproduce completely the data it is necessary to correct the generated events at the analysis level. Two examples relevant in this thesis are described in the following.

3.6.1 The LAr failure

Between 30 April and 13 July 2011 several FEBs of the electromagnetic calorimeter were inactive in the region $0 < \eta < 1.5$ and $-0.8 < \phi < 0.6$. There were interruptions to localized regions of the high-voltage supply across the active medium. This is known as the “LAr hole” problem. Since this failure occurred after the production of the Monte Carlo samples, its effect had to be reproduced at the analysis level.

³<http://sourceforge.net/projects/sframe>

3.6.2 Pile-up re-weighting

The average number of interactions per bunch crossing is a parameter of the simulation. The effects of pile-up are simulated by overlaying various numbers of minimum-bias events onto the hard process of interest. The value chosen for the 2010 production was found to be smaller than the average number of interactions occurring at the LHC in early 2011; this of course caused a discrepancy between the data and the simulation. Additionally, in 2011 the proton bunch spacing was decreased from 75 ns to 50 ns producing also an out-of-time component of pile-up, coming from overlapping signals in the detector from neighboring bunch crossings. A tool centrally provided by the ATLAS collaboration has been used to correct the simulation, according to the pile-up conditions observed in data. The tool contains a set of weights to be applied on an event-by-event basis.

4 Object reconstruction and identification

Particles produced in proton-proton collisions traverse the detector material and leave signals. These are recorded, digitalized and read-out by the detector electronics. The electronic signals are used to reconstruct physics objects. Following the object reconstruction, identification criteria are imposed. The reconstruction and identification of different object-types are described in the chapter. In particular, those involved in the final state of the $t\bar{t}$ decay: hadrons visible as jets, neutrinos, electrons and muons with large transverse momentum (p_T). The performance studies presented in Chapter 5 are also based on objects identified as jets. The reconstruction, identification and selection of jets is described in detail in this chapter.

4.1 Track reconstruction

Tracks display the reconstructed trajectories of charged particles when traversing the detector. These are reconstructed in the Inner Detector (ID) within an acceptance range of $|\eta| < 2.5$. The reconstruction is divided into three stages. At the first, the hits in the pixel and SCT detectors are converted into clusters and transformed into three-dimensional space points [96]. At the same time, the timing information in the TRT is converted into calibrated drift circles. At the second stage, a track-finding is performed as follows. The space points in the three layers of the pixel detector and the first layer of the SCT form track seeds. Track candidates are obtained by extending the seeds across the outer SCT layers, and fitted excluding outlier clusters. If a layer has no hits, the track has a so-called “hole”. A track candidate is required to have a limited number of holes, as well as clusters shared with other tracks. Finally, selected tracks are extended into the TRT and refitted using the three sub-detectors of the ID. Hits on the extended tracks that result in a bad fit are labelled as outliers [100]. Since the tracks are contained in a magnetic field, they are described by a helix whose parameters are: the signed transverse (d_0) and longitudinal (z_0) impact parameters, the azimuthal and polar angles θ and ϕ of the track tangent at perigee, and q/p_T , where q is the signed charge of the track. Primary and secondary vertices are also reconstructed using exclusively ID information. The primary event vertex is defined as the vertex with the highest sum of the squared p_T values of the associated tracks ($\sum p_{T,\text{track}}^2$) in the event. This is reconstructed from at least five tracks, each with $p_T > 150$ MeV.

4.2 Electrons

4.2.1 Reconstruction, identification and isolation

The electron reconstruction uses a sliding window algorithm, which builds fixed size clusters from selected seeds. The seed clusters have a size of 3×5 cells, where one cell corresponds to $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$. The algorithm finds a seed cluster with an energy deposit of more than 2.5 GeV, and then reconstructs the electron candidate if it matches an ID track with $p_T > 0.5$ GeV. If more than one track is matched to the cluster, the track with the smallest ΔR is taken.

The energy cluster is determined by summing four different contributions [101]:

1. The estimated energy deposited in material in front of the EM calorimeter.
2. The measured energy deposited in the cluster.
3. The estimated energy deposited outside of the cluster (lateral leakage).
4. the estimated energy deposited beyond the EM calorimeter (longitudinal leakage).

The EM-scale is the scale at which the energy of electrons is correctly measured in the calorimeters. In order to obtain this scale, several corrections are applied to the final clusters. These corrections involve the built four-momentum of the electron, the energy measurement of the cluster and the η and ϕ coordinates of the ID track. The transverse energy of an electromagnetic cluster is defined in terms of the calorimeter cluster energy and the direction of the track pointing to it as [102]

$$E_{T,\text{cluster}} = \frac{E_{\text{cluster}}}{\cosh \eta_{\text{track}}}. \quad (4.1)$$

The largest background to the isolated electrons arises from hadronic jets that pass the electron identification. These are referred to as “fake” electrons. For example, the process $\pi^0 \rightarrow \gamma\gamma$ produces photons which may be reconstructed as one or more electrons if there is at least one nearby reconstructed track. The accurate reconstruction of electrons is also affected by sources producing non-prompt electrons, i.e. those that do not result directly from the decay of a W or Z boson¹. For instance, the semi-leptonic heavy-flavor decays (of B - or D -hadrons), the decay of charged pions (when the pions are “in-flight”) and photon conversions to electron-positron pairs. Identification criteria are imposed to obtain high efficiency for selecting prompt electrons and to suppress fake electrons [104]. These cuts involve variables measured either separately in the calorimeter and the ID, or in both combined. There are three classifications, **loose**, **medium** and **tight**, meaning an increasingly background-rejection power of 500, 5000 and 50,000 respectively [105]. The **loose** set uses shower shape variables of the EM calorimeter middle layer and hadronic leakage variables. The **medium** set adds requirements on the EM calorimeter strip layer variables, track quality and track-cluster matching. The

¹The bosons are prompt sources if they were produced in the proton-proton hard-scatter or, for the W boson, in the decay of a top quark [103].

Type	Description	Name
Loose selection		
Acceptance	$ \eta < 2.47$	
Hadronic leakage	Ratio of E_T in the first layer of the hadronic calorimeter to E_T of the EM cluster (used over the range $ \eta < 0.8$ and $ \eta > 1.37$)	R_{had1}
	Ratio of E_T in the hadronic calorimeter to E_T of the EM cluster (used over the range $ \eta > 0.8$ and $ \eta < 1.37$)	R_{had}
Middle layer of EM calorimeter	Ratio of the energy in 3×7 cells over the energy in 7×7 cells centered at the electron cluster position	R_η
	Lateral shower width, $\sqrt{(\sum E_i \eta_i^2)/(\sum E_i) - ((\sum E_i \eta_i)/(\sum E_i))^2}$, where E_i is the energy and η_i is the pseudorapidity of cell i and the sum is calculated within a window of 3×5 cells	$w_{\eta 2}$
Medium selection (includes loose)		
Strip layer of EM calorimeter	Shower width, $\sqrt{(\sum E_i (i - i_{\text{max}})^2)(\sum E_i)}$, where i runs over all strips in a window of $\Delta\eta \times \Delta\phi \approx 0.0625 \times 0.2$, corresponding typically to 20 strips in η , and i_{max} is the index of the highest-energy strip	w_{stot}
	Ratio of the energy difference between the largest and second largest energy deposits in the cluster over the sum of these energies	E_{ratio}
Track quality	Number of hits in the pixel detector (≥ 1)	n_{pixel}
	Number of total hits in the pixel and SCT detectors (≥ 7)	n_{Si}
	Transverse impact parameter ($ d_0 < 5$ mm)	d_0
Track-cluster matching	$\Delta\eta$ between the cluster position in the strip layer and the extrapolated track ($ \Delta\eta < 0.01$)	$\Delta\eta$
Tight selection (includes medium)		
Track-cluster matching	$\Delta\phi$ between the cluster position in the middle layer and the extrapolated track ($ \Delta\phi < 0.02$)	$\Delta\phi$
	Ratio of the cluster energy to the track momentum	E/p
	Tighter $\Delta\eta$ requirement ($ \Delta\eta < 0.005$)	$\Delta\eta$
Track quality TRT	Tighter transverse impact parameter requirement ($ d_0 < 1$ mm)	d_0
	Total number of hits in the TRT	n_{TRT}
Conversions	Ratio of the number of high-threshold hits to the total number of hits in the TRT	f_{HT}
	Number of hits in the b-layer (≥ 1) (if expected)	n_{BL}
	Veto electron candidates matched to reconstructed photon conversions	

Table 4.1: Definition of variables used for loose, medium and tight electron identification cuts. The cuts are optimized in 10 bins of cluster η and 11 bins of cluster E_T from 5 GeV to above 80 GeV [101].

tight cut set applies cuts on E/p , the ratio between the energy reconstructed in the EM calorimeter and the track momentum. These criteria are detailed in Table 4.1.

In addition, an isolation requirement is imposed to reduce the contribution from non-prompt electrons. The isolation is requested either in the calorimeter or in the ID. The

4 Object reconstruction and identification

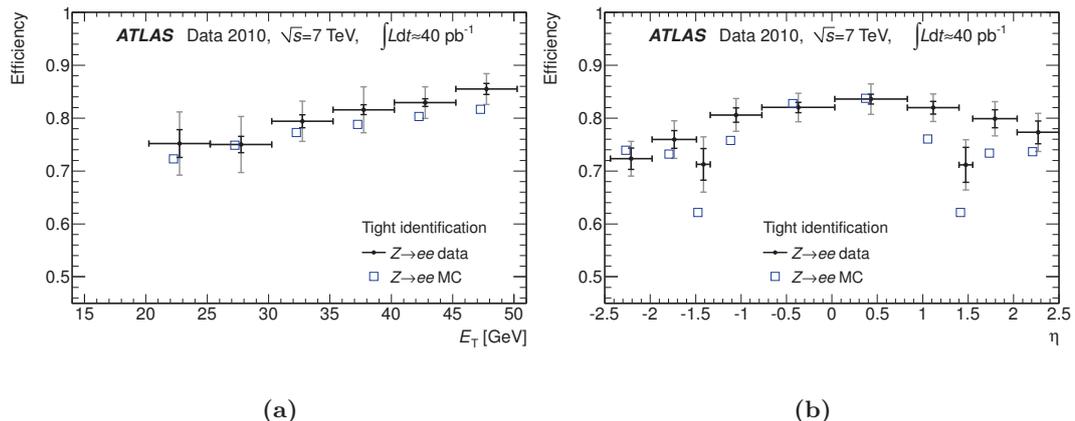


Figure 4.1: Identification efficiencies of tight electrons, measured from $Z \rightarrow ee$ events in data (black dots) and predicted by MC (blue squares), as a function of (a) E_T and (b) η . The results for the data are shown with their statistical (inner error bars) and total (outer error bars) uncertainties. The statistical error on the MC efficiencies plotted as open squares is negligible [101].

calorimeter isolation is measured using the variable EtCone20 . This is the sum of the E_T recorded in all EM and hadronic calorimeter cells within a cone of radius $R = 0.2$, centered on the electron. The energy of the electron itself must be subtracted from the cone. A p_T -dependent correction is applied accounting for the leakage of the electron energy out of the central core, and another correction to account for soft energy deposits from pileup events. The track isolation on the other hand, is required by summing the transverse momenta of all tracks in the ID that fall inside a cone of size $\Delta R = 0.3$. The assigned variable is PtCone30 .

The performance of the reconstruction, identification and isolation is determined through efficiency (ε) measurements. The efficiencies and their uncertainties are calculated in bins of kinematical variables, using the *tag-and-probe* (T&P) method [101]. The T&P method is commonly utilized in $Z \rightarrow \ell\ell$ events, in which one of the leptons satisfy tight selection criteria (the “tag”), and the other has looser criteria applied (the “probe”). In this case, the invariant mass of the two electrons has to be close to the Z boson mass. This technique is equally implemented in Monte Carlo (MC) simulation and data, resulting in different efficiency values. Figure 4.1 shows the efficiencies of the tight electron identification cuts measured as a function of the electron E_T and $|\eta|$.

The ratio $\varepsilon_{\text{data}}/\varepsilon_{\text{MC}}$ is defined as the scale factor (SF). This is a multiplicative factor applied to the weight of the individual simulated events, and is parameterized in terms of kinematic variables. SFs are applied to correct the MC efficiency, adjusting its value with that measured in data. The combined SFs for the electron identification and isolation are binned in eighteen η_{cluster} and five E_T bins to account for regions with different efficiencies, while retaining enough statistics in each bin. They are not measured beyond $|\eta_{\text{cluster}}| = 2.47$. Discrepancies between the efficiencies measured in the simulation and in data are mostly due to mis-modeling in the TRT-related variables [101]. In re-

gions at $|\eta_{\text{cluster}}| < 1.37$ the SFs are ≈ 1 and for regions at $|\eta_{\text{cluster}}| > 1.37$ they drop slightly to ≈ 0.97 . The electron reconstruction SFs are obtained in three bins of $|\eta_{\text{cluster}}|$. They are listed, together with their combined statistical and systematic uncertainties as follows [106]:

$$\begin{aligned} 0 < |\eta_{\text{cluster}}| < 0.8 & \quad \text{SF}_{\text{reco}} = 0.9984 \pm 0.0066, \\ 0.8 < |\eta_{\text{cluster}}| < 2.37 & \quad \text{SF}_{\text{reco}} = 1.0091 \pm 0.0070, \\ 2.37 < |\eta_{\text{cluster}}| < 2.47 & \quad \text{SF}_{\text{reco}} = 0.9759 \pm 0.0184. \end{aligned}$$

4.2.2 Electron trigger and resolution

For the analysis presented in this thesis two single-electron triggers are used. Firstly, a trigger requiring an EM cluster with $E_{\text{T}} > 20$ GeV was implemented. As the instantaneous luminosity increased, the rates became higher and this trigger was pre-scaled. The next higher-threshold and un-pre-scaled trigger was then chosen: $E_{\text{T}} > 22$ GeV. Triggers without pre-scale whose thresholds are moderately higher are requested in order to guarantee no further loss of events to be analyzed. The thresholds indicate the regions where the trigger is fully efficient selecting events.

The trigger efficiency is measured in data and MC simulation using the T&P method on samples with $W \rightarrow e\nu$ and $Z \rightarrow ee$ events. The differences are accounted in flat SF of 0.995 ± 0.01 [103].

The EM-scale was derived from test-beam measurements with an uncertainty of 3%. The uncertainty is reduced using the masses of the Z boson and the J/ψ particle. Further reduction is achieved by comparing the ratio E/p in $W \rightarrow e\nu$ events. From these results, the energy scale in the Monte Carlo simulation is smeared to match the di-electron mass distribution. Specifically, the width of this reconstructed mass peak must match that in data. Figure 4.2a shows the ratio E/p for electrons emitted in $W \rightarrow e\nu$ decays selected in the barrel: E/p is nearly unity with a significant tail at large values due to Bremsstrahlung occurring in the ID. Figure 4.2b shows good agreement between data and simulation for the di-electron mass distribution of electrons from $J/\psi \rightarrow ee$ events. The width of this latter distribution is an estimate of the calorimeter electron energy resolution. The energy resolution of calibrated electrons with $E_{\text{T}} > 25$ GeV is better than 2%.

4.3 Muons

4.3.1 Reconstruction, identification and isolation

The analysis presented in this thesis involves combined (CB) muons. These are retrieved from the combination of tracks reconstructed separately in the MS (stand-alone) and in the ID [107]. A MS track must be associated to an ID track within $|\eta| < 2.5$, given by the coverage of the ID. This matching is performed by constructing a χ^2 fit whose parameters account for the traversed material, the non-uniformity of the magnetic field and the muon trajectory. Since the procedure is done for all good stand-alone tracks, the

4 Object reconstruction and identification

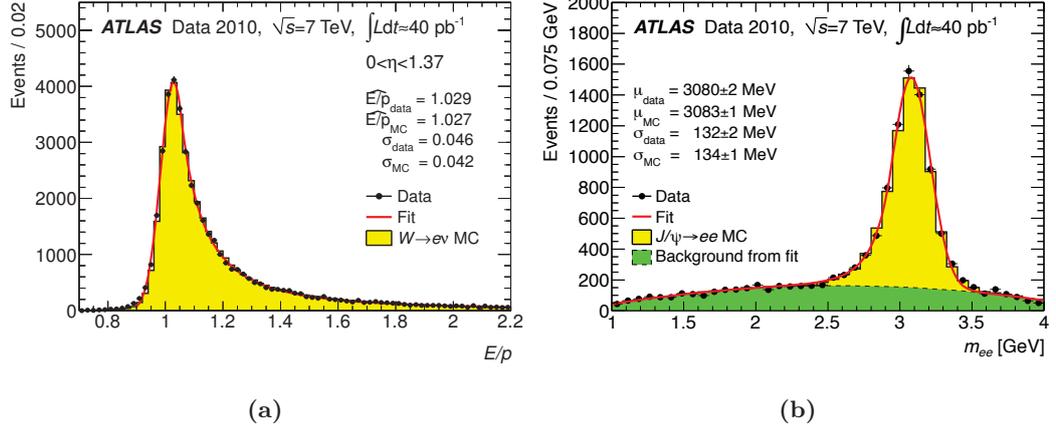


Figure 4.2: Electron energy calibration distributions: (a) E/p distributions of electrons from $W \rightarrow e\nu$ decays for $0 < |\eta| < 1.37$ in data (full circles with statistical error bars) and simulation (filled histogram). (b) Reconstructed di-electron mass distribution for $J/\psi \rightarrow ee$ decays, as measured after applying the baseline $Z \rightarrow ee$ calibration: the energy of all electrons with $E_T > 10$ GeV, is corrected so that the reconstructed Z boson mass peak is centered on its known mass. The data (full circles with statistical error bars) are compared to the sum of the MC signal (light filled histogram) and the background contribution (darker filled histogram) [101].

MS-ID track combination leading to the best χ^2 is chosen to form the track of the muon candidate. A set of offline cuts are needed for the ID segments of the tracks associated to combined muons [108]. These cuts are applied in order to reject muons that do not come from the process of interest in this thesis, i.e. $t \rightarrow W \rightarrow \mu\nu$. They are listed as follows:

- The ID track must have at least one hit in the innermost layer of the pixel detector. If the muon has traversed an area which is known to be not functioning or not instrumented this requirement is dropped. In addition, the total number of pixel hits plus the total number of crossed dead pixel sensors must be greater than one.
- The number of SCT hits plus the number of crossed dead SCT sensors must not be smaller than six.
- No more than two holes (see Section 4.1) in the pixel and SCT detectors are allowed.
- For $|\eta| < 1.9$: $n > 5$ and $n_{\text{outliers}}/n < 0.9$, where n denotes the number of TRT hits plus the number of TRT outliers (see also Section 4.1).
- for $|\eta| \geq 1.9$: if $n > 5$ require $n_{\text{outliers}}/n < 0.9$, using the same notation as above.

The momentum of the reconstructed combined muon is computed by averaging the separate measurements of the stand-alone and the good-quality ID tracks, weighted with a factor extracted from their covariance matrix. These muons are classified as **tight CB**.

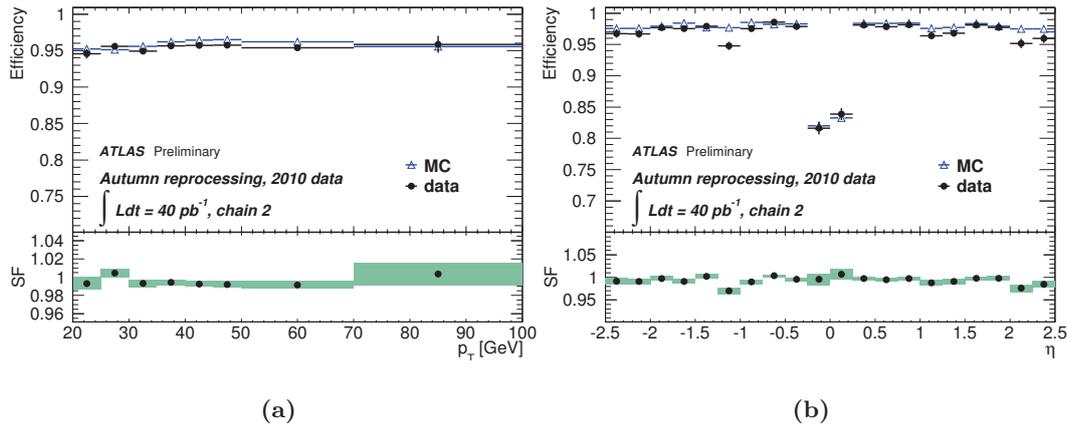


Figure 4.3: Muon reconstruction efficiency as a function of the muon (a) p_T and (b) η , measured in both data and MC simulation and integrated over $20 < p_T < 100$ GeV. The efficiencies retrieved from data (dots) without background correction and Monte Carlo simulation (open triangles) including backgrounds, are depicted in the upper part of each figure [109].

The muon isolation is implemented to suppress non-prompt muons. As defined for electrons, the isolation is requested in the calorimeters and in the tracking system. The scalar sum of the p_T of all tracks and the E_T measured in the calorimeters, in a cone of radius 0.3 surrounding the muon candidate: $PtCone30$. The calorimeter isolation requirement is assigned to the variable $EtCone30$. The muon track and the energy deposited by the muon itself are subtracted from the sums.

The T&P technique is used on $Z \rightarrow \mu\mu$ decays to measure the muon reconstruction and identification efficiencies [109]. The di-muon invariant mass has to be close to the mass of the Z boson, and the muons oppositely charged. The total efficiency is the product of the efficiency derived from the ID, the efficiency from the MS and the matching efficiency between the ID and the MS systems. The individual efficiencies are calculated as follows: one muon is required to be reconstructed (identified) in both systems whereas the other muon is identified by just one of the systems in order to probe the efficiency of the other [103]. Figure 4.3 shows the reconstruction efficiency for muons measured in both data and simulation, as a function of the muon p_T and η . The average SF across all regions is 0.995 ± 0.002 . The dominating systematic uncertainty in the derivation of this SF is the one associated with the background estimation [110].

The identification efficiencies are parametrized as a function of the muon p_T , η , the ΔR with its closest jet and the number of primary vertices (N_{PV}) in the event. Flat SFs are applied as they do not exhibit a significant dependence on the variables: 1.008 ± 0.0003 (stat.) ± 0.0003 (syst.) for the first 1.5 fb^{-1} of analyzed data and 1.0034 ± 0.0003 (stat.) ± 0.0002 (syst.) thereafter up to 2.05 fb^{-1} [103]. The same source of systematic uncertainty is considered as for the reconstruction efficiency.

4.3.2 Muon trigger and resolution

To avoid the usage of pre-scaled triggers, two single triggers requiring a muon with $p_T > 18$ GeV are used in this thesis. The first requires only two layers of the muon trigger system to have fired, and the other requires all the three layers. The T&P method is again implemented to obtain the trigger efficiencies [103]. These are measured by requiring a match of a reconstructed muon with a trigger object from the muon triggers. The matching condition is given within a distance of $\Delta R < 0.15$. The SFs for the two triggers are derived as function of η and ϕ in three p_T bins: [20-60] GeV, [60-120] GeV and [120-150] GeV. A dependence of the SFs on the muon p_T was found in the barrel region. This stems from a misconfiguration of the level-2 trigger causing an underestimation of the MC efficiency for muons with $p_T > 40$ GeV. This effect increased linearly with the muon transverse momentum and caused a 30% muon loss at $p_T = 500$ GeV [111]. Therefore, the muon trigger requirement is not applied to simulated events. Instead, these events are weighted by the muon trigger efficiency measured in data rather than by a SF. The procedure is described in [110]. A systematic uncertainty due to this effect is estimated. The systematic uncertainties on the SFs derivation arise from the di-muon invariant mass threshold chosen in the T&P implementation. The choice of the trigger-matching threshold and the isolation criteria also propagate systematic uncertainties.

The muon momentum calibration and resolution are obtained from $Z \rightarrow \mu\mu$ and $W \rightarrow \mu\nu$ events [112]. These events are sensitive to the resolution since the width of the reconstructed di-muon invariant mass peak (at the Z pole) is a convolution of the natural width of the Z boson and the muon p_T resolution. $W \rightarrow \mu\nu_\mu$ events are used to estimate the difference of the separate p_T measurements in the ID and the MS. The differences between data and MC simulation are corrected by shifting the muon momentum scale, and smearing the resolution.

4.4 Jets

Partons emerging from pp collisions hadronize; their final states manifest as sprays of collimated particles in the detector. Jets are defined to capture the hadronization products into a measurable object whose kinematics is extracted in the reconstruction. Jets confine colorless bound states which are generated by colored partons interacting strongly. As partons undergo showering before hadronization, they can radiate additional partons while the strong coupling increases (splitting). Thus, the emission probabilities can diverge. As mentioned in Section 2.5.1.3, the divergences show up when soft partons are emitted (infrared divergences) and when the radiation occurs at small angles with respect to the original parton (collinear divergences). The challenge when reconstructing jets is to connect the theoretically predicted colored states with the observed confined-states produced after hadronization. This implies to characterize the partonic structure of the jet, protecting its definition against infrared and collinear divergences (infrared and collinear safety).

4.4.1 Reconstruction and identification

Jets can produce electromagnetic and hadronic showers, leaving energy deposits in the calorimeters. They also can produce tracks in the ID. Jets reconstructed from calorimeter read-outs are called calorimeter jets, and those reconstructed from the ID are track jets. Jets are defined by the jet algorithms. They can be classified in two mayor categories: cone algorithms [113] and sequential recombination algorithms [114, 115]. The latter are infrared and collinear safe and are therefore implemented in ATLAS.

The sequential recombination algorithms build jets step-by-step from their initial constituents. The k_T algorithms belong to this category. These are cluster type jet-finders and are specifically based on a sequential pair-wise recombination. The algorithm starts from the calorimeter clusters or towers referred to as objects. Let ϕ_i , η_i and $p_{T,i}$ be, respectively, the azimuthal angle, pseudorapidity, and transverse momentum of object i . For each pair of objects ij calculate

$$d_{ij} = \frac{\min(p_{T,i}^{2s}, p_{T,j}^{2s})}{R^2} \times \Delta R_{ij}^2, \quad (4.2)$$

where,

$$\Delta R_{ij} = \sqrt{(\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2}, \quad (4.3)$$

is the distance in the $\eta - \phi$ plane between the two objects, and R is an adjustable distance parameter used to control the width of the jet: the larger R , the wider the jet. The parameter s is an integer number among 1, 0, -1 [116].

Also, define the distance between the object i and the beam:

$$d_{i,B} = p_{T,i}^{2s} \quad (4.4)$$

Find the minimum of all d_{ij} and $d_{i,B}$. If it is a d_{ij} , merge the two objects into one by adding up their four-momenta. If it is a $d_{i,B}$, declare the object a jet and remove it from the list of objects. This process is repeated until all objects have become part of jets, i.e. until no objects are left. The d values are calculated for all possible combinations.

There are three different cluster algorithms, depending on the value of variable s :

- $s = 1$: The k_T algorithm [117] is used because of the particular order in which it performs the clustering: it iterates across soft particles first, leaving hard objects for the final iterative process. This is designed to somehow reverse the parton splitting before hadronization.
- $s = 0$: The Cambridge-Aachen algorithm [114] in which the p_T of the objects is irrelevant and the clustering is carried out according to the angular distance R between two objects, starting with the closest one.
- $s = -1$: The anti- k_T algorithm [115, 116] has the effect that soft objects cluster with the hard ones, before they cluster among themselves. This somehow reverts the effect of radiation. If the soft objects are within a radius less than $2R$ from

4 Object reconstruction and identification

the hard object, the resulting jet tends to be circular in the $\eta - \phi$ plane. Two hard objects within $R < \Delta R < 2R$ share the energy between them, depending on the relative p_T and the d_{ij} . The hard objects within $\Delta R < R$ will be merged in one single jet.

Objects making up an anti- k_T jet can be given to the k_T algorithm for recombination. If a jet contains the products of a two-body decay of a heavy particle, the final clustering step of the k_T algorithm will usually be to combine these two decay products. There is a variable associated with this final step: first splitting scale, $\sqrt{d_{12}}$. This variable is used to distinguish between different types of particles initiating the jet, and enables its decomposition into sub-jets. It is defined as

$$\sqrt{d_{12}} = \min(p_{T,1}, p_{T,2}) \times \Delta R_{12}, \quad (4.5)$$

where 1 and 2 are the two sub-jets that are present just before the final recombination into a single jet. The mentioned features of the first splitting scale are useful for studying the structure of the jet. Methods using jet substructure have been developed for improving the identification of boosted top quarks [118–121]. This technique is implemented in this thesis to determine whether the jet contains the decay products of a top quark, or those come from a light quark or gluon.

Jets reconstructed with the anti- k_T algorithm are used in this thesis. It has been demonstrated that anti- k_T jets are less affected by underlying events and pile-up effects, and manifest high reconstruction efficiency for all jet flavors, fragmentation and showering models [122].

4.4.1.1 Calorimeter jets

Calorimeter signals are taken by the jet algorithm to reconstruct jets. The ATLAS calorimeters contain about 200 000 cells. The cells where signal is found are grouped into larger signal objects, whose four-momenta are the inputs of the algorithm. As mentioned above, there are two possible definitions of these groups: topological clusters (topo-clusters) and calorimeter towers. The analyses involving top quarks at ATLAS are recommended to use jets formed from topo-clusters. A brief description is given as follows.

Topo-clusters are designed to follow the shower development, exploiting the fine segmentation of the ATLAS calorimeters. The clustering starts with seed cells whose ratio between signal and noise² is greater than 4. All direct neighbors are grouped first. Neighbors of neighbors are added iteratively to the cluster if their signal-to-noise ratio is > 2 . When no further adjacent cells are found with signal-to-noise ratio ≥ 2 , a ring comprising all adjacent cells is included in the topo-cluster. All cells of the reconstructed topo-cluster are successively searched for local maxima of deposited energy. These maxima are then used as seed cells for a new iteration of the topological clustering, which

²The noise factor accounts for electronic noise and interferences caused by pile-up events interacting with the calorimeter. It is defined as the RMS of the cell noise distribution.

will split the original cluster into more topo-clusters. A topo-cluster is defined to have an energy equal to the energy sum of all the included cells. Its reconstructed direction is that of a vector with its origin in the origin of the ATLAS coordinate system, and pointing to the energy-weighted topo-cluster barycenter [123]. Jets are built from well-defined topo-clusters using the `FastJet` program [124]. It computes the total four-momentum of the jet by adding the four-momenta of its constituents, while ensuring energy and momentum conservation. The transverse momentum threshold for the `FastJet` program to reconstruct a jet is $p_T > 7$ GeV.

4.4.2 Jet calibration

The ATLAS calorimeters are non-compensating. This means that an incident hadron with equivalent energy to an electron will have a lower reconstructed energy. Similar to the electrons, jets are reconstructed at the EM-scale. This scale is set using test-beam measurements for electron deposits in the barrel and the end-caps calorimeters [125]. The EM-scale accounts correctly for the energy deposited in the calorimeters by electromagnetic showers. Since jets also contain hadrons, it is necessary to calibrate the jet energy to the scale of the hadrons. This procedure is known as hadronic scale calibration. It relates the energy of the jet measured in the calorimeters to the energy of the incident particles which initiate the jet. The hadronic jet energy scale (JES) is on average restored using data-derived corrections and calibration constants. These are obtained by comparing the kinematics of the reconstructed jet to its corresponding Monte Carlo truth-level jet³. Further effects besides the non-compensation are corrected via jet calibration:

- Dead material: energy losses in inactive regions of the detector.
- Leakage: energy deposits from particles reaching outside the calorimeters.
- Out of calorimeter jet cone: energy deposits from particles that are not included in the reconstructed jet.
- Signal losses in the clustering procedure and thus in the reconstructed jet.

The calibration scheme used for top quark analyses in ATLAS is called EM+JES. This is implemented because it allows a simple evaluation of the systematic uncertainties [126]. The EM+JES calibration applies jet-by-jet corrections as a function of the jet energy and pseudorapidity to jets reconstructed at the EM-scale. There are three subsequent stages carried out during this calibration:

1. **Pile-up corrections or offset correction:** The average additional energy due to in-time pile-up is subtracted from the energy measured in the calorimeters. The correction factors are extracted from in-situ measurements, using minimum bias data as a function of the N_{PV} , the jet p_T and pseudorapidity [126]. Since this

³The truth jet recreates the hadronization before the detector level and maps the kinematical conditions of the measured jet by using the same reconstruction algorithm (see Section 3.6).

4 Object reconstruction and identification

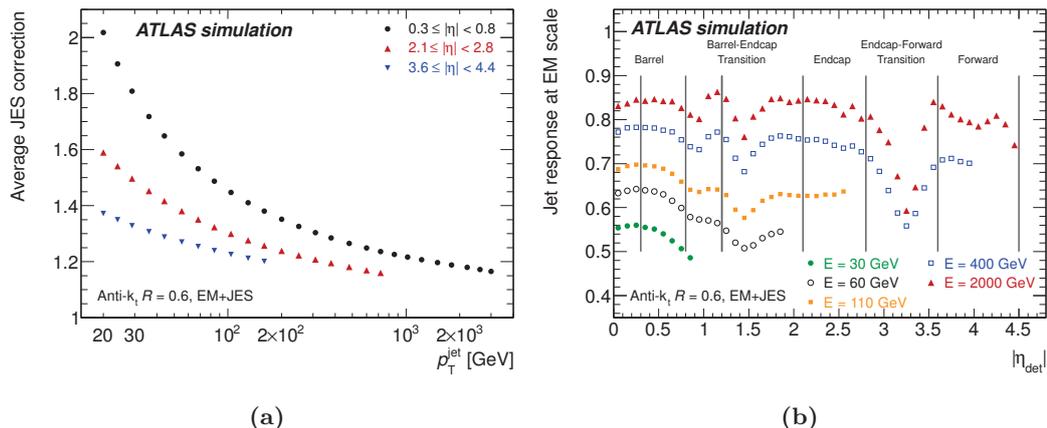


Figure 4.4: EM+JES jet calibration: (a) average jet energy scale correction as a function of the calibrated jet transverse momentum in three detector pseudorapidity bins. (b) simulated jet energy jet response at the EM scale as a function of the detector pseudorapidity. The response is shown for different EM+JES anti- k_T calibrated jets [126].

correction is calculated with respect to the p_T of the jet instead of its E_T , the derivation can result in reconstructed jets with negative energy. These jets are not considered as part of events to be analyzed.

2. **Jet origin correction:** Rather than towards the geometrical center of the ATLAS detector, the direction of each input topo-cluster is corrected to point back to the primary vertex in the event. This is valid as long as the jet originates from that vertex and not from pile-up.
3. **Residual JES correction:** The energy response function $\mathcal{R}_{\text{EM}}^{\text{jet}} = E_{\text{EM}}^{\text{jet}}/E_{\text{truth}}^{\text{jet}}$, corrects the energy of reconstructed jets $E_{\text{EM}}^{\text{jet}}$ to the energy of their corresponding Monte Carlo truth jet $E_{\text{truth}}^{\text{jet}}$. The correction is applied in data and MC simulated events. Figure 4.4 shows the average energy correction as a function of the calibrated jet p_T and the average simulated jet response at EM scale.

The p_T correction ranges from 1.2 for high p_T jets to 2.1 for soft jets, in the central region. Low p_T jets need a larger correction because they have a smaller fraction of electromagnetically-interacting particles than harder jets. The jet pseudorapidity is also corrected for a small effects caused by poorly instrumented regions of the calorimeter.

The EM+JES scheme is used in this thesis to calibrate anti- k_T jets whose radius parameter is $R = 0.4$ or $R = 0.6$. However, anti- k_T jets with $R = 1.0$ are also used. A procedure called Local Cluster Weighting (LCW or LC) [127, 128] is used to calibrate these larger-radius jets. In this calibration, the topo-clusters are first classified as either electromagnetic or hadronic. A weight is then applied to each cell in the cluster according

to: energy and the cell energy density, depth in the calorimeter, fractional energy deposited in the calorimeter layer and energy measured around it. The weights are derived from Monte Carlo simulation of charged and neutral pions, and are applied individually before the reconstruction. Following jet reconstruction from the LC-scale topo-clusters, the JES calibration then proceeds as described above for EM-scale topo-clusters. The whole calibration scheme is thus called LC+JES.

The benefit of the LC+JES calibration is that individual topo-clusters are properly calibrated according to their type, before the jets are reconstructed. It is particularly suitable for analyses involving the jet substructure. In this thesis, the mass jet is a substructure variable used to identify anti- k_T $R = 1.0$ jets. A correction for the mass of anti- k_T $R = 1.0$ jets is derived, along with the LC+JES calibration. The mass and energy are corrected by 10 - 20% on average.

4.4.3 Jet energy resolution

The jet energy resolution is determined in di-jet events using two methods [129]. The di-jet balance method is based on momentum conservation in the transverse plane. The asymmetry distribution of the two-leading-jets p_T is fitted with a Gaussian, from which the jet p_T resolution is obtained. This method applies corrections for soft jets undetected in the calorimeter. The bi-sector method is based on an imbalance vector, given by the vector sum of the two leading jets in the di-jet event. The vector is zero by construction for perfectly balanced jets, and thus any fluctuation can be investigated [129]. Figure 4.5 shows the data and MC agreement obtained from the two methods. The simulation describes the jet energy resolution within 14% for jets with $20 \text{ GeV} < p_T < 80 \text{ GeV}$. The same uncertainty is assumed for $p_T \geq 80 \text{ GeV}$. An additional smearing in the jet p_T is applied in order to enhance the agreement.

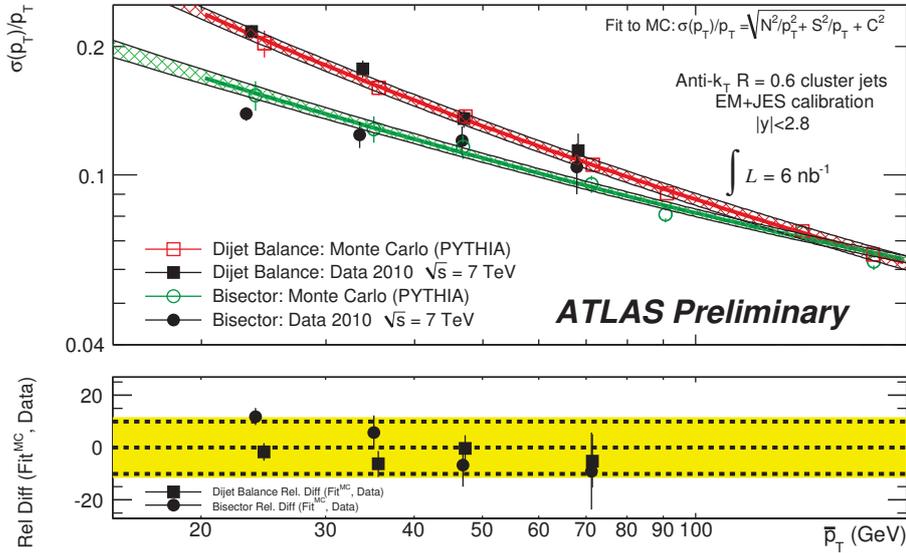


Figure 4.5: Jet energy resolution for the di-jet balance and bi-sector techniques as a function of the average jet transverse momenta. The lower plot shows the relative difference between the Monte Carlo fit and the data results. The yellow band indicates the relative uncertainty [129].

4.4.4 Jet selection and efficiency

Mis-reconstructed jets can arise from sources such as hardware problems and cosmic-ray contributions. Background processes resulting in jets can include interactions between protons and residual gas in the beam pipe and noise in the calorimeters. Jets that in some way are problematic are flagged as “bad” quality. A jet is classified as bad if any of the following criteria is not fulfilled [126, 130, 131]:

- The 5 energy-ordered cells with highest energy deposition in the jet must contain less than 90% of the energy of the jet.
- The fraction of the jet energy deposited in the hadronic end-cap calorimeters (f_{HEC}) is less than or equal to 0.8.
- The absolute value of the jet quality variable, which quantifies how closely the measured calorimeter pulses match a reference pulse is less than $1 - f_{\text{HEC}}$.
- The fraction of energy of the jet deposited in the electromagnetic calorimeter is less than or equal to 0.95.
- The cell-weighted time of the jet is less than 50 ns different from the average event time.
- The fraction of the jet energy coming from masked cells, whose energy is extrapolated using the energy density of neighboring cells, must be less than or equal to 0.5.

Bad jets are vetoed in any selection of this thesis. Jet candidates which are not in-time with the collision events are a signal for cosmic rays or non-collision background. Hence, these are not selected. In addition, data events are removed if they contain a jet with $p_{\text{T}} > 20$ GeV closer than 0.1 in ΔR to the LAr problematic region (see Section 3.6.1). The same procedure has been applied to the fraction of the simulated sample proportional to the integrated luminosity of the affected data.

The selection efficiency is measured using the T&P method in di-jet events with a back-to-back topology [129]. A track jet is used as tag objects, and a second track jet is balancing the event in ϕ ; this will be the probe candidate. The efficiency is derived by matching a reconstructed calorimeter jet to the probe jet. Figure 4.6 shows the data and Monte Carlo efficiencies obtained after the T&P implementation. Since the jet reconstruction is almost fully efficient for high- p_{T} jets, namely above 20 GeV, only jets located in this region are used in the $t\bar{t}$ analysis of this thesis.

4.5 b-jets

As a consequence of the branching ratio $BR(t \rightarrow Wb) \approx 1$, one bottom quark (b -quark) is expected per top quark decay. A b -jet originates from a b -quark, which produces a bottom-hadron (B -hadron) in the hadronization. The properties of the B -hadrons are used to identify them from hadrons containing light quarks:

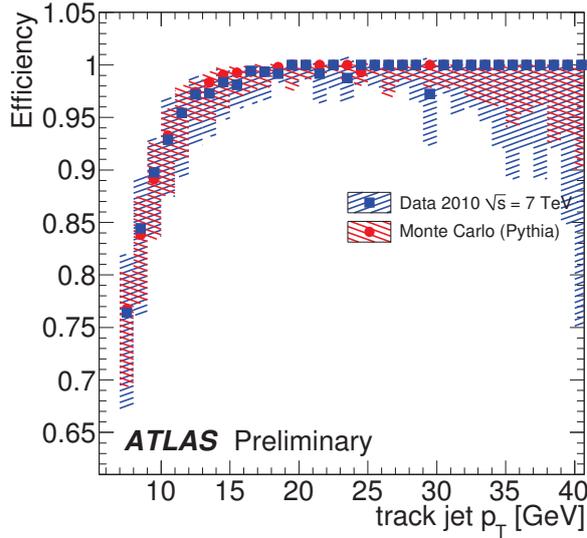


Figure 4.6: Jet Selection efficiency relative to track jets as a function of the probe track jet p_T in data and MC simulation. The hatched areas indicate the statistical and systematic uncertainties added in quadrature [129].

- B -hadrons preserve most of the original b -quark momentum (about 70%) and exhibit large masses above 5 GeV. The decay products carry large transverse momenta with respect to the jet axis. Furthermore, the opening angle of the decay products is large enough to recognize their spatial separation.
- Hadrons containing b -quarks are characterized by a long lifetime of about 1.5 ps. So, they typically travel about 3 mm in the transverse plane before decaying.

4.5.1 b -jets identification

The identification of b -jets is referred to as b -tagging. Several b -tagging techniques are implemented in ATLAS [132], all relying on a likelihood-ratio approach to build a discriminating variable, called jet weight [133]. Jets having a jet weight above a certain cut value are tagged as b -jets. Each b -tagging method exploits the two facts: 1) a certain number of tracks point to a secondary vertex instead of pointing to the reconstructed primary vertices, 2) the impact parameters of these tracks are large.

The weights obtained from different taggers can be combined to enhance the discrimination power. For instance, the `JetFitterComNN` [134, 135] algorithm used in this thesis, combines the `IP3D` and `JetFitter` [132, 135] algorithms. The `IP3D` algorithm is an impact parameter (IP) based tagger. It uses the IP significance to derive the likelihood corresponding to the b -tag probability. For each track contained in the respective jet, the IP significance S is obtained from its transverse ($S_{d_0} = d_0/\sigma_{d_0}$) and longitudinal ($S_{z_0} = z_0/\sigma_{z_0}$) components (see Section 4.1). If the probabilities of finding a track from a b -jet or any other jet, e.g. u -jet are P_b and P_u , respectively, the track weight is determined by w_t : $w_t = P_b/P_u$. From all tracks i in the jet passing a quality cut, the jet

4 Object reconstruction and identification

weight is defined as

$$w_{\text{jet}} = \sum_{i \in \text{jet}} \ln w_t^i. \quad (4.6)$$

The secondary vertex-based `JetFitter` tagger exploits the topology of weak B - and D -hadron decays inside the jet. The algorithm identifies both decays by applying a Kalman filter [136] on a common flight direction, along with the position of additional vertices on it. The jet weight is calculated for different decay topologies, by deriving a likelihood ratio for each of the following discriminating variables:

- invariant mass of all the tracks associated to the secondary vertex, which is higher for b -jets than for light jets,
- energy fraction of the tracks fitted to the vertex to all tracks in the jet, due to in general, b -jet tracks contain large fraction of the jet energy,
- number of tracks pairs contained in the secondary vertex, which is larger in the case of b -jets,
- angle between the directions of the jet and the B -hadron flight.

The `JetFitterComNN` tagger uses an artificial neural network to combine the weights of `JetFitter` and `IP3D` algorithms, together with variables describing the topology of the hadron decay chain. The neural network is trained with Monte Carlo simulated samples [135].

The b -tag efficiency is defined as the fraction of reconstructed jets originating from b -quarks that are tagged by the b -tagging algorithm. For a $w_{\text{jet}} > 0.6017$ derived from the MC simulation, the b -jet identification efficiency is of 70%, which sets the working point of the `JetFitterComNN` algorithm [135, 137]. It is also used to measure the mistag rate, i.e., the rate at which light-jets are misidentified and tagged as b -jets. The inverse of the of the mistag rate is then the efficiency for rejecting light-jets.

4.5.2 b -tagging efficiency and scale factors

The b -tag efficiency is measured from kinematic variables such as jet p_T and η . Quantities accounting for the fraction of jets coming from gluons are also used. Figure 4.7 shows the tagging rates for an enriched sample in heavy-flavor jets, as obtained in data and as measured in a simulated sample of $t\bar{t}$ events. The tagging rate predicted by the simulation agrees with experimental data to within 20%.

4.6 Missing transverse momentum

Missing transverse momentum is the signature of undetected objects. For an event in which all its particles are detected, the missing transverse momentum is zero. However, particles like neutrinos do not interact with the detector material, and the missing transverse momentum arises to balance the total momentum in the event. The transverse momenta computed from the transverse energies of all the detected objects are summed

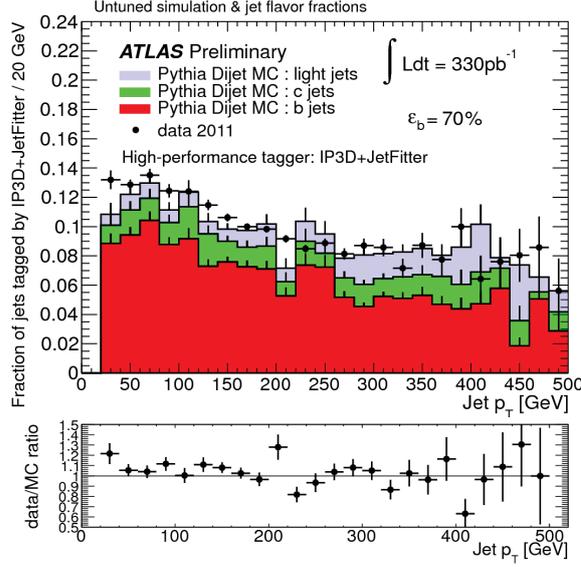


Figure 4.7: Distribution of the tagging rate for the IP3D+JetFitter tagging algorithm at an operating point leading to 70% b -jet efficiency for experimental data (solid black points) and for simulated data (filled histograms for the various flavors) as a function of the jet transverse momentum, using $t\bar{t}$ events [135].

vectorially. The negative value of this sum is the missing transverse momentum denoted by E_T^{miss} .

The E_T^{miss} reconstruction proceeds in two stages. Firstly, the calorimeter cells whose deposits have been associated with reconstructed objects are used to form the *calorimeter terms*. All cells have to be calibrated according to the object they are associated with:

1. Identified electrons with $p_T > 10$ GeV and calibrated to the EM-scale.
2. Anti- k_T $R = 0.4$ jets with $p_T > 20$ GeV and calibrated to the EM+JES scale (**HardJets**).
3. Anti- k_T $R = 0.4$ jets with $7 \text{ GeV} < p_T < 20$ GeV and calibrated to the EM-scale (**SoftJets**).

Each calorimeter term is calculated from the negative sum of the cell energies for the corresponding objects as [138]

$$E_x^{\text{miss}} = - \sum_{i=1}^{N_{\text{cell}}} E_i \sin \theta_i \cos \phi_i \quad , \quad E_y^{\text{miss}} = - \sum_{i=1}^{N_{\text{cell}}} E_i \sin \theta_i \sin \phi_i \quad ,$$

where E_i , θ_i and ϕ_i are the energy, the polar angle and the azimuthal angle, respectively, of the cells associated to energy clusters. The sum runs over all cells in the pseudorapidity range $|\eta| < 4.5$. The ordering of the objects indicates the order of association of the

4 Object reconstruction and identification

cluster to the objects. If a cell belongs to more than one object, only the first association is taken. The second stage of the E_T^{miss} reconstruction corresponds to a *muon term*. This is defined by the track's momenta of reconstructed muons, within $|\eta| < 2.7$. The contribution is calculated as

$$E_{x,y}^\mu = - \sum_{\text{muons}} p_{x,y}^\mu. \quad (4.7)$$

Only good-quality combined muons with a matched track in the ID are considered.

The remaining energy from cells not associated with any object is included as an additional EM-scale term called `CellOut`.

The E_T^{miss} components are calculated as

$$-E_{x,y}^{\text{miss}} = E_{x,y}^{\text{Electron}} + E_{x,y}^{\text{HardJet}} + E_{x,y}^{\text{SoftJet}} + E_{x,y}^\mu + E_{x,y}^{\text{CellOut}}, \quad (4.8)$$

and the E_T^{miss} value is then calculated as

$$E_T^{\text{miss}} = \sqrt{(E_x^{\text{miss}})^2 + (E_y^{\text{miss}})^2}. \quad (4.9)$$

The resolution of the E_T^{miss} is determined by the precision and the energy scales of the objects which are used in the reconstruction. It is expected to be proportional to the square root of the total E_T deposited in one bunch crossing. By using $t\bar{t}$ simulated events, the resolution of the x and y components ranged from about 2 GeV (for $E_T^{\text{miss}} = 20$ GeV) to 10 GeV (for $E_T^{\text{miss}} = 400$ GeV) [139].

4.6.1 Neutrino reconstruction

Neutrinos from the W boson decay are used in this thesis. As mentioned above, the neutrino cannot be detected. Only its transverse momentum is reconstructed as E_T^{miss} . The procedure to reconstruct neutrino's longitudinal momentum ($p_{z,\nu}$) is described as follows.

The sum of the lepton and neutrino four-vectors $\vec{\mathbf{P}}_\ell$ and $\vec{\mathbf{P}}_\nu$, respectively, is equal to the four-vector of the W boson, $\vec{\mathbf{P}}_W$:

$$\vec{\mathbf{P}}_W = \vec{\mathbf{P}}_\ell + \vec{\mathbf{P}}_\nu. \quad (4.10)$$

This equation is used assuming an on-shell W boson production. Thus, its pole mass $m_W = 80.4$ GeV is employed as constraint to obtain a $p_{z,\nu}$ equation. By neglecting the invariant masses of the lepton and the neutrino, and after squaring the terms in Equation 4.10, the $p_{z,\nu}$ equation is written as

$$p_{z,\nu}^2 - 2 \cdot \frac{\mu p_{z,\ell}}{E_\ell^2 - p_{z,\ell}^2} \cdot p_{z,\nu} + \frac{E_\ell^2 p_{T,\nu}^2 - \mu^2}{E_\ell^2 - p_{z,\ell}^2} = 0, \quad \text{with} \quad \mu = \frac{1}{2} m_W^2 + p_{T,\ell} p_{T,\nu} \cos \Delta\phi, \quad (4.11)$$

4.6 Missing transverse momentum

where ϕ is the azimuthal angle between the E_T^{miss} and the lepton. $p_{z,\ell}$ is the longitudinal component of the lepton momentum and $p_{T,\nu}$ is the transverse component of the neutrino momentum. The energy of the lepton is denoted by E_ℓ . The solution of Equation 4.11 is given by:

$$p_{z,\nu}^\pm = \frac{\mu p_{z,\ell}}{p_{T,\ell}} \pm \sqrt{\frac{\mu^2 p_{z,\ell}^2}{p_{T,\ell}^4} - \frac{E_\ell^2 p_{T,\nu}^2 - \mu^2}{p_{T,\ell}^2}}. \quad (4.12)$$

If the radicand of Equation 4.12 is positive, two real solutions are obtained. The solution with the smallest $|p_z|$ is chosen since for about 70% of the $t\bar{t}$ events, it produces values close to the true values. Complex solutions arise if the radicand is negative. This happens when the reconstructed transverse mass of the W boson is larger than its pole mass. Likely, that is due to an imperfect reconstruction of the E_T^{miss} [140]. In such a case, the E_T^{miss} is adjusted by rotating the x and y components, until a real solution is found [59, 140].

5 Level-1 jet trigger efficiencies

The production of QCD jets is the dominant hard process at the LHC. Jet triggers are the primary means for selecting events containing jets with high-transverse momentum (p_T). These events are used in ATLAS for physics analyses ranging from Standard Model QCD to Beyond the Standard Model searches. The performance of jet triggers is evaluated in terms of the trigger efficiency. This must be precisely measured since it enters in the calculation of cross-sections. This chapter presents an overview of the jet trigger efficiencies using data collected by the ATLAS detector in 2010, at $\sqrt{s} = 7$ TeV. Since in these data only the level-1 was used to actively select events¹, the studies described are exclusively aimed to the level-1 triggers.

5.1 Level-1 calorimeter trigger

The jet triggers discriminate events on basis of E_T and jet multiplicity. This is carried out in the ATLAS calorimeter system. The level-1 calorimeter trigger is a pipelined digital system built of custom electronics [141]. The calorimeters are segmented into 7168 trigger towers, with granularity of approximately 0.1×0.1 in $\Delta\eta \times \Delta\phi$. These represent 7168 analogue inputs that must first be shaped and amplified, before they are digitized and synchronized in a pre-processing phase. At the pre-processing stage, 10-bit flash ADCs are used as analog-to-digital converters, while noise thresholds and energy calibration are set by a Pre-Processor (PPr) system [83]. Hardware-based lookup tables are used for pedestal subtraction to apply noise thresholds and for final energy calibration. The electromagnetic tower response is calibrated at the electromagnetic scale, whereas the hadronic tower response is calibrated for jets. The final energies of the trigger towers are expressed as 8-bit integers with accuracy of 1 GeV. In the last step of the pre-processing, signals are associated with a particular LHC bunch crossing. Once digital transverse energies per LHC bunch crossing are formed, two separate processor systems working in parallel, run the trigger algorithms for further trigger processing [141]. One of these systems is the jet/energy sum Processor (JEP). It extracts information from data provided by the PPr using 32 jet/energy sum processor modules (JEMs).

The JEM combines digitally the electromagnetic and hadronic trigger towers, yielding an effective calorimeter-tower granularity of $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$. The resulting signal are called jet elements. These are used to find jet trigger candidates when running the trigger algorithm. The level-1 jet trigger is based on a sliding window algorithm that identifies jet candidates, which have E_T values above a given threshold. Specifically, the

¹The HLT algorithms were executed in *pass-through* mode. This means that their output was recorded, but was not considered in the actual trigger decision.

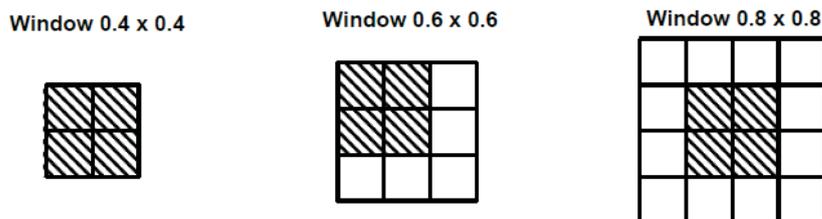


Figure 5.1: The three jet sizes with 2×2 , 3×3 or 4×4 jet trigger elements, respectively. The shaded regions represent jet RoIs. In the 0.6×0.6 size, there are four possible windows containing a given RoI. In the 0.8×0.8 case the RoI is required to be in the center position, in order to avoid the possibility of two jets per window [83].

algorithm identifies E_T sums within overlapping windows consisting of 2×2 , 3×3 or 4×4 jet elements, corresponding to window sizes of $\Delta\eta \times \Delta\phi = 0.4 \times 0.4$, $\Delta\eta \times \Delta\phi = 0.6 \times 0.6$ or $\Delta\eta \times \Delta\phi = 0.8 \times 0.8$, respectively. This is shown in Figure 5.1. These sums are then compared to pre-defined jet energy thresholds.

The trigger algorithm windows overlap. To avoid double counting of jet candidates, each jet window must surround a 2×2 jet-element area that is a local maximum in E_T . Figure 5.2 shows the procedure to determine the local maximum. The location of the maximum also defines the coordinates of the jet RoI (see Section 3.4.1). Eight independent combinations of jet E_T thresholds and window sizes are available for trigger menus.

An energy-summation algorithm performs global transverse energy summations: missing transverse energy (E_T^{miss}), total transverse energy ($\sum E_T$) and jet-sum transverse energy ($\sum E_T^{\text{jet}}$). Each JEM transmits the resulting information to four Common Merg-

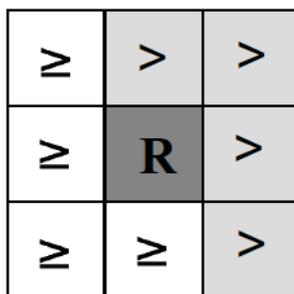


Figure 5.2: The local E_T maximum test for level-1 jet finding. R represents the 2×2 jet element area (RoI) under testing. The RoI must be more energetic ($>$) than its neighbors along two connected edges, and at least as energetic (\cong) as its neighbors along the opposite two edges. The η axis runs from left to right and the ϕ from bottom to top [83].

ing Modules (CMM). They merge the results from all the modules, counting objects passing each jet trigger threshold. In the end, the CMMs send the total multiplicities to the CTP. The CTP then makes the overall acceptance decision based on objects found in the JEP and the multiplicities counted by the CMMs. The final level-1 trigger decision is made by the CTP, applying the multiplicity requirements and pre-scale factors, which are previously set in the level-1 trigger menu.

5.2 Jet triggers in the 2010 physics menu

The trigger menu in 2010 consisted of more than 470 trigger items, the majority tailored to physics analysis [97]. As the LHC luminosity increased and the trigger rates approached the bandwidth limits, pre-scale factors were imposed to level-1 triggers (see Sections 3.4 and 3.4.4). Table 5.1 shows exemplary the lowest un-pre-scaled (UP) single-jet triggers, together with their corresponding integrated luminosities. The listed triggers belong to the `JetTauEtmis` physics stream defined in Section 3.4.3.

Period	Dates	$\int \mathcal{L} dt$ [pb^{-1}]	Max. \mathcal{L} [$\text{cm}^{-2}\text{s}^{-1}$]	Lowest UP
B	23/4 - 17/5	9.0×10^{-3}	6.8×10^{28}	L1_J5
C	18/5 - 23/6	9.5×10^{-3}	2.4×10^{29}	L1_J5
D	24/6 - 28/7	0.3	1.6×10^{30}	L1_J15
E	29/7 - 18/8	1.4	3.9×10^{30}	L1_J30
F	19/8 - 21/9	2.0	1.0×10^{31}	L1_J55

Table 5.1: Single-jet triggers for each period in 2010 data-taking. Lowest UP is the lowest un-pre-scaled trigger for each period. \mathcal{L} is the instantaneous luminosity. Two separated data samples corresponding to a integrated luminosities of 17 nb^{-1} and 2.42 nb^{-1} are used in this study.

Level-1 jet triggers used in 2010 are separated into four categories: single or *inclusive jets* (J), *forward jets* (FJ), *multi-jets* (nJ, n=2,3..), and *total jet* E_T (JE). The triggers of interest in this chapter are organized, together with their corresponding pre-scale factors in Table 5.2. Single and multi-jet triggers are inclusive, meaning that an n -trigger fires if n or more jets are encountered. Consider for instance, an event containing three jets. If the jets have energies between 5 GeV and 10 GeV, the event is passing only L1_J5 because the level-1 requirement is $E_T > \text{threshold}$, in this case 5 GeV. If one of these jets has $E_T > 10$ GeV, the event passes also the L1_J10 trigger. Thus, L1_JX is the short-hand notation for requiring at least one jet, identified by the level-1 with E_T above the threshold X.

The total jet transverse energy $\sum E_T^{\text{jet}}$ (JE), is a trigger specialized in global event selection (global jet trigger). A standard jet trigger will require that the event has e.g. at least one jet $> X$ GeV, or at least three jets $> Y$ GeV. Unlike, the JE trigger will tend to select both events with a few high- E_T jets or a larger number of lower- E_T jets. The selection is performed without putting fixed limits on the actual jet E_T . JE does

5 Level-1 jet trigger efficiencies

Category	L1 Threshold	Pre-scale	Motivation
<i>Inclusive</i> or single jets	J5	$\sim 10^5$	jets in central region ($ \eta < 3.2$)
	J10	$\sim 10^4$	
	J15	$\sim 10^4$	
	J30	$\sim 10^3$	
<i>Multi-jets</i>	2J10	~ 30	two or more central jets above threshold
	2J15	~ 200	
	2J30	~ 5	
	3J10	1	
<i>Total jet E_T</i> $\sum E_T^{\text{jet}}$	JE60	~ 70	total E_T of all jets above threshold, in the central part of the calorimeter
	JE100	~ 70	
	JE120	1	

Table 5.2: Level-1 jet triggers in 2010 used in this study. Triggers are shown in categories with their L1 threshold and approximate pre-scale factor for an instantaneous luminosity of $\sim 10^{32} \text{ cm}^{-2}\text{s}^{-1}$. Pre-scale value of 1 means no pre-scale. Since jet triggers are calorimeter-based, the thresholds are applied on the transverse energy (E_T) in units of GeV. The number before the letters represents the minimal multiplicity required, while that after the letters the jet threshold [142].

require that at least one jet threshold is passed, but after that every jet contributes to the trigger. Thus, the JE trigger accounts for the total jet multiplicity, rather than just for sum of the E_T of the jets passing a trigger threshold. To obtain the $\sum E_T^{\text{jet}}$, the resulting multiplicity of jets per threshold is weighted by a factor. The calculation is performed using a simple multiplicity-based estimator implemented in the CTP. The inputs to the algorithm are the counts of the numbers of jets passing each of the main jet thresholds (inclusive jet trigger thresholds) and the values of those thresholds. To understand the estimator, consider an event with n jets passing the threshold $E_T > X$ GeV and m jets passing $E_T > Y$ GeV, with $n > m$ and $Y > X$. There are $n - m$ jets in the E_T range $X \text{ GeV} \leq E_T \leq Y \text{ GeV}$. These jets then have a total E_T of at least $(n - m) X$ GeV and at most $(n - m) Y$ GeV.

5.3 Jet trigger efficiency

Level-1 jet triggers must select interesting events in $2.5 \mu\text{s}$, within a limited trigger bandwidth². The online identification and selection of physical signals implies:

- run the reconstruction algorithms,
- calculate jet variables which can be used for identification,
- use these quantities to decide whether the event is selected or not.

²With a luminosity of $10^{32} \text{ cm}^{-2}\text{s}^{-1}$ the total bandwidth allowed for the single, forward and multi-jet triggers is about 20 Hz, and for the E_T^{miss} , $\sum E_T$ and $\sum E_T^{\text{jet}}$ triggers is 10 Hz.

The ability of the trigger system in the listed processes is evaluated by measuring the trigger efficiency, ε . It is defined as the ratio between the number of surviving events after the trigger selection and the total number of events without trigger applied. In the case of the jet triggers, the jet calorimeter clustering and calibration lead to a finite resolution in the energy measurements. Thereby, the online measurement is not accurately determined and the reconstructed offline jet is taken as the best available estimate. The trigger efficiency can accordingly be redefined as the probability of an event with a property as reconstructed by the offline software, to pass a trigger based on the online reconstruction of the trigger.

The trigger efficiency can be calculated either per event (event-level) or per object (jet-level). When measured per event, the efficiency is the probability of an offline event having a certain set of properties to fire the jet trigger. When measured per object, it is the probability that a specific jet in an event satisfies the jet trigger. The difference is that while in the first case all jets in an event that passed the trigger are positively contributing to the efficiency, in the second only the offline jets matched in $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} = 0.3$ to a RoI jet (online jet) are contributing. This ΔR value has been found highly efficient for identifying whether a level-1 jet corresponds to an offline jet [143].

For either calculation, the trigger efficiency is represented as a function of quantities reconstructed offline, e.g. p_T or rapidity (y). The trigger efficiency *turn-on* curve is shaped by the ratio of the jet distribution in events that passed the trigger, with respect to the offline-observable distribution. The characteristic shape of the curve has a *turn-on* region where the efficiency increases quickly, and a “plateau” region where the maximum efficiency is reached and no longer varies. Thus, the jet trigger under study is fully efficient selecting jet-events when these fall in the plateau region.

5.3.1 Estimation methods

The efficiency of a jet trigger cannot blindly be estimated using a sample of events selected online by the same trigger. An unbiased or independent measurement of the trigger efficiency is always desired. To achieve it, two methods are implemented:

- Orthogonal triggers: The event is triggered by a different and unbiased trigger with respect to the other for which the efficiency is being determined.
- Bootstrapping: The events triggered using a lower threshold are used to determine the efficiency of a trigger using a higher one.

In this study, the level-1 minimum bias trigger (L1_MBTS_1) is used as orthogonal trigger. This trigger selects events by requiring at least one hit in the Minimum Bias Trigger Scintillators (MBTS), on either side of the interaction point (see Section 3.2.6). Events passing this trigger are stored in the the MinBias stream defined in Section 3.4.3. The orthogonality of the L1_MBTS_1 trigger with respect to the jet triggers was verified by comparing its efficiency to that obtained by instead using the Zero-Degree Calorimeter (ZDC). A systematic uncertainty accounting for the limited statistics of the ZDC sample has been propagated. This is summed in quadrature to the statistical errors for the jet

5 Level-1 jet trigger efficiencies

trigger efficiencies derived from the L1_MBTS_1 trigger. This systematic uncertainty is 5% assigned to the efficiency points less than 80% and 1% for points above this value [143]. If the orthogonal triggers method is employed, the ratio

$$\varepsilon_{\text{L1_nJX}} = \frac{N(\text{Sel} \wedge \text{L1_nJX} \wedge \text{L1_MBTS_1})}{N(\text{Sel} \wedge \text{L1_MBTS_1})}, \quad (5.1)$$

is the efficiency of the L1_nJX trigger and Sel is the offline selection described in Section 4.4.4.

When using the bootstrapping method the trigger efficiency is calculated as

$$\varepsilon_{\text{L1_nJY}} = \frac{N(\text{Sel} \wedge \text{L1_nJX} \wedge \text{L1_nJY})}{N(\text{Sel} \wedge \text{L1_nJX})} \cdot \varepsilon_{\text{L1_nJX}}, \quad (5.2)$$

with $\text{L1_nJY} \subseteq \text{L1_nJX}$. L1_nJX and L1_nJY correspond to the lower and higher threshold-triggers, respectively.

5.4 Results

The trigger efficiency turn-on curves shown in this section are obtained by parametrizing Equations 5.1 and 5.2 as a function of the offline jet distributions. The errors in all figures are calculated — as is conventional for trigger efficiency errors — using the Bayesian method assuming binomial statistics. Since the efficiency cannot be higher than 1, this method implements a prior distribution going from 0 to 1 [144]. In all the studies a PYTHIA sample has been used for the Monte Carlo simulation (see Section 2.5.4). The efficiencies are calculated using jets reconstructed with the anti- k_T algorithm with $R = 0.4$ and $R = 0.6$. The scheme to calibrate these jets is described in Section 4.4.2.

5.4.1 Single-jet trigger efficiencies

Single-jet triggers are applicable for instance, to bootstrap events to be triggered by multi-jet triggers. The latter are used in analyses requiring more than one jet in the final states, e.g. in measurements of multi-jet cross-sections [145–148]. For a total integrated luminosity of 17 nb^{-1} , the single jet trigger with a E_T threshold of 5 GeV (L1_J5) was implemented in the analysis [145]. The orthogonal triggers method has been used to calculate the efficiency of this trigger. By using anti- k_T $R = 0.6$ jets, the L1_J5 trigger is 99% efficient at 60 GeV, as seen in Figure 5.3. This figure shows the event-level trigger efficiency as a function of the reconstructed offline leading jet p_T , integrating over $|y| < 2.8$. Data and simulation are in good agreement. According to Table 5.1, the instantaneous luminosity was low enough for the L1_J5 trigger to stay un-pre-scaled. Thus, there was no need to employ multi-jet triggers. Events whose leading jet in p_T fell in the plateau region $p_T^{\text{Lead}} > 60 \text{ GeV}$ were therefore used for the further analysis [145].

For an integrated luminosity of 2.4 pb^{-1} , the orthogonal triggers method has also been employed to determine single-jet trigger efficiencies. Figure 5.4 shows the efficiency of

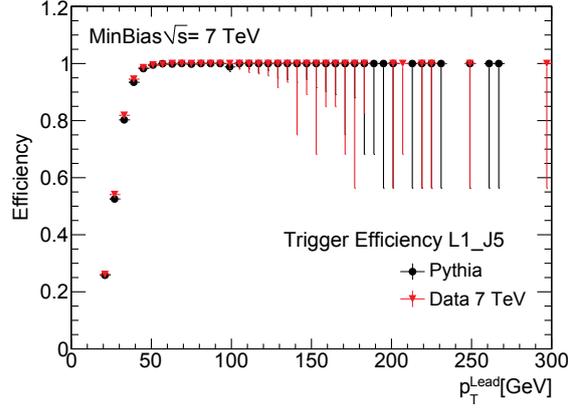


Figure 5.3: Event-level efficiency of the L1_J5 trigger as a function of the offline leading jet p_T for events containing anti- k_T $R = 0.6$ jets. The efficiency is measured relative to the L1_MBTS_1 trigger. Data and PYTHIA Monte Carlo simulation are compared.

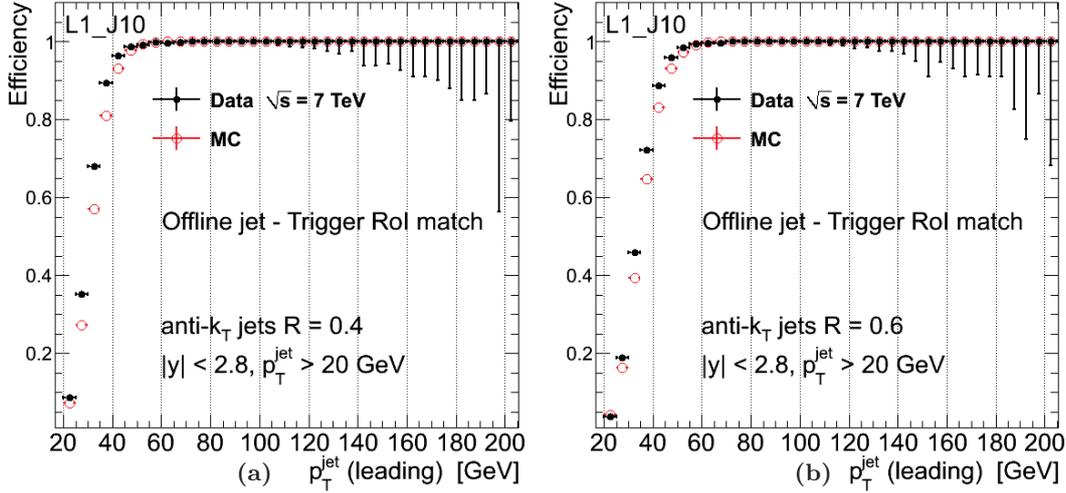


Figure 5.4: Efficiency of the L1_J10 trigger as a function of the leading jet p_T , integrated over rapidity $|y| < 2.8$, for anti- k_T jets with (a) $R = 0.4$ and (b) $R = 0.6$. The efficiency is shown as obtained in the Monte Carlo simulation and as calculated using the orthogonal triggers method in data. Events accepted by the unbiased L1_MBTS_1 trigger are used for calculating the efficiency. Only offline-selected jets with $p_T > 20$ GeV are accepted.

the L1_J10 trigger to select events as a function of the reconstructed offline leading jet p_T . In this case, anti- k_T $R = 0.4$ and anti- k_T $R = 0.6$ are required within $|y| < 2.8$. Good agreement between data and simulation is observed in the plateau region for both sizes of jets. The trigger reaches the maximum efficiency at $p_T \approx 50$ GeV and $p_T \approx 60$ GeV

5 Level-1 jet trigger efficiencies

for jets reconstructed with $R = 0.4$ and $R = 0.6$, respectively. This is expected as larger jets have more spread-out energy, and that the level-1 RoIs are the same in both cases. Events falling in the plateau of this trigger are used to test the efficiency as a function of the leading jet rapidity. Figure 5.5 shows the efficiency of the L1_J10 trigger, using events in which the p_T of the leading jet is larger than 60 GeV.

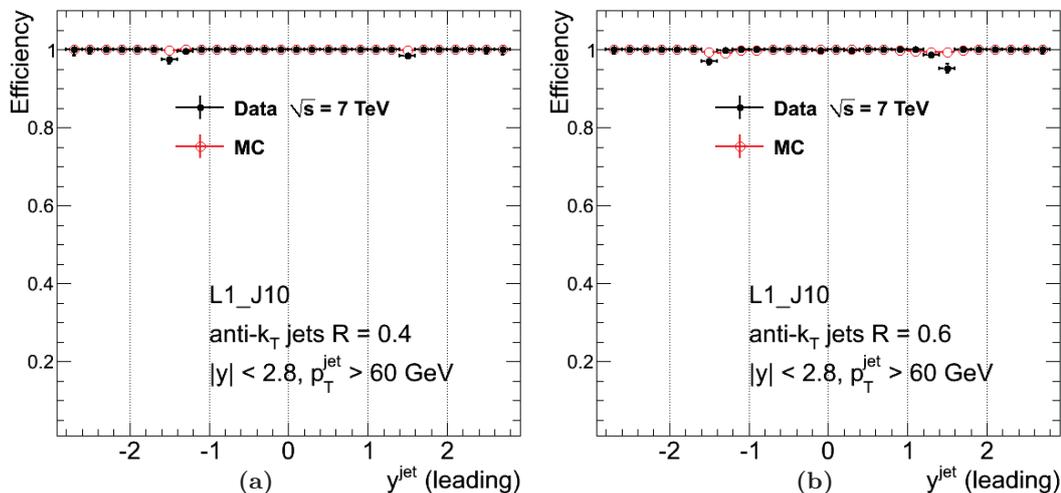


Figure 5.5: Efficiency of the L1_J10 trigger as a function of the leading jet rapidity for events whose the leading jet has $p_T > 60$ GeV, for anti- k_T jets with (a) $R = 0.4$ and (b) $R = 0.6$. The efficiency is shown as obtained in the Monte Carlo simulation and as calculated using the orthogonal L1_MBTS_1 trigger.

The trigger is fully efficient over the full rapidity range, except in the transition region between the barrel and end-cap calorimeters around $|y| = 1.5$ in data. The reason is that from periods B to F, the level-1 jet trigger-tower timing was commissioned. Thereby, only part of the calorimeter in that region was included in the trigger.

A similar study is carried out to calculate the efficiency of the L1_J15 trigger. Figure 5.6 shows the L1_J15 trigger efficiency parametrized as a function of the leading jet p_T and rapidity. The turn-on curve shows that the L1_J15 trigger is $\sim 100\%$ efficient for jets with p_T leading above 80 GeV, regardless the jet size. Therefore events where the leading jet has $p_T > 80$ GeV, form the sample to calculate the trigger efficiency as function of the leading jet $|y|$. The inefficiencies in data are produced by the same inefficiencies in the transition region between the barrel and end-cap calorimeters described above.

The L1_J30 trigger efficiency is shown in Figure 5.7. In this case, the trigger is 100% efficient for events containing jets with p_T above ≈ 110 GeV.

5.4.2 Multi-jet trigger efficiencies

Efficiencies of multi-jet triggers have been calculated via bootstrapping for data periods from B to F. This corresponds to an integrated luminosity of 2.43 pb^{-1} . During these periods, only the L1_3J10 trigger ran without pre-scale. The efficiencies are determined

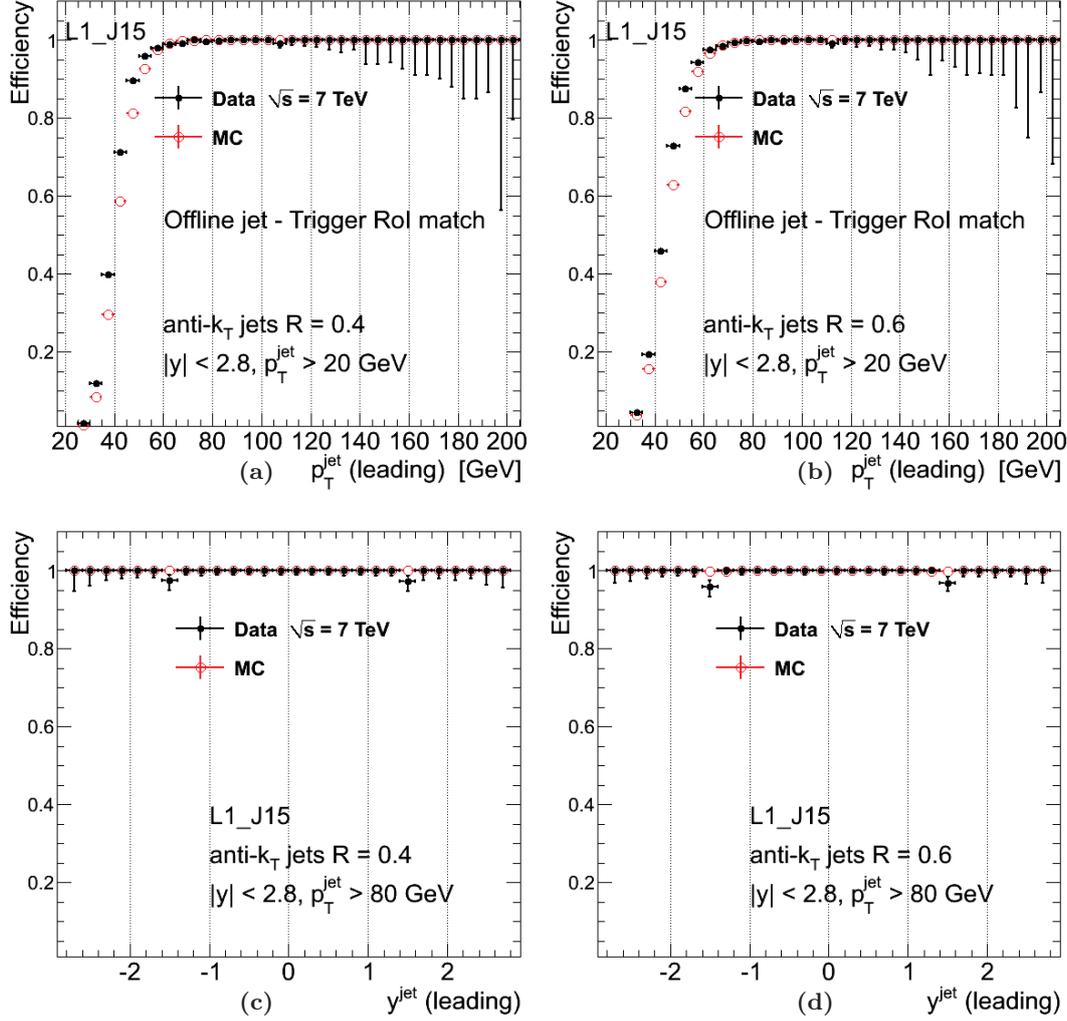


Figure 5.6: Efficiency of the L1_J15 trigger: efficiency as a function of the leading jet p_T for anti- k_T jets with (a) $R = 0.4$ and (b) $R = 0.6$, and efficiency as a function of the leading jet rapidity for events in which the leading jet has $p_T > 80$ GeV, for anti- k_T jets with (c) $R = 0.4$ and (d) $R = 0.6$. The efficiencies are shown as obtained in the Monte Carlo simulation and as calculated using the orthogonal L1_MBTS_1 trigger in data.

starting from the low multiplicity trigger L1_2J10, and using the plateaus obtained for the single-jet triggers. Namely, a two-jet event whose leading jet falls in the plateau region of the single-jet trigger, will be 100% efficient to select two-jet events, regardless of the second jet p_T . Therefore, the efficiency of the L1_2J10 trigger to fire on the second leading jet is measured by requiring that the leading jet belongs to the plateau of the L1_J10 trigger (i.e. leading jet with $p_T > 60$ GeV). In addition, the leading jet has

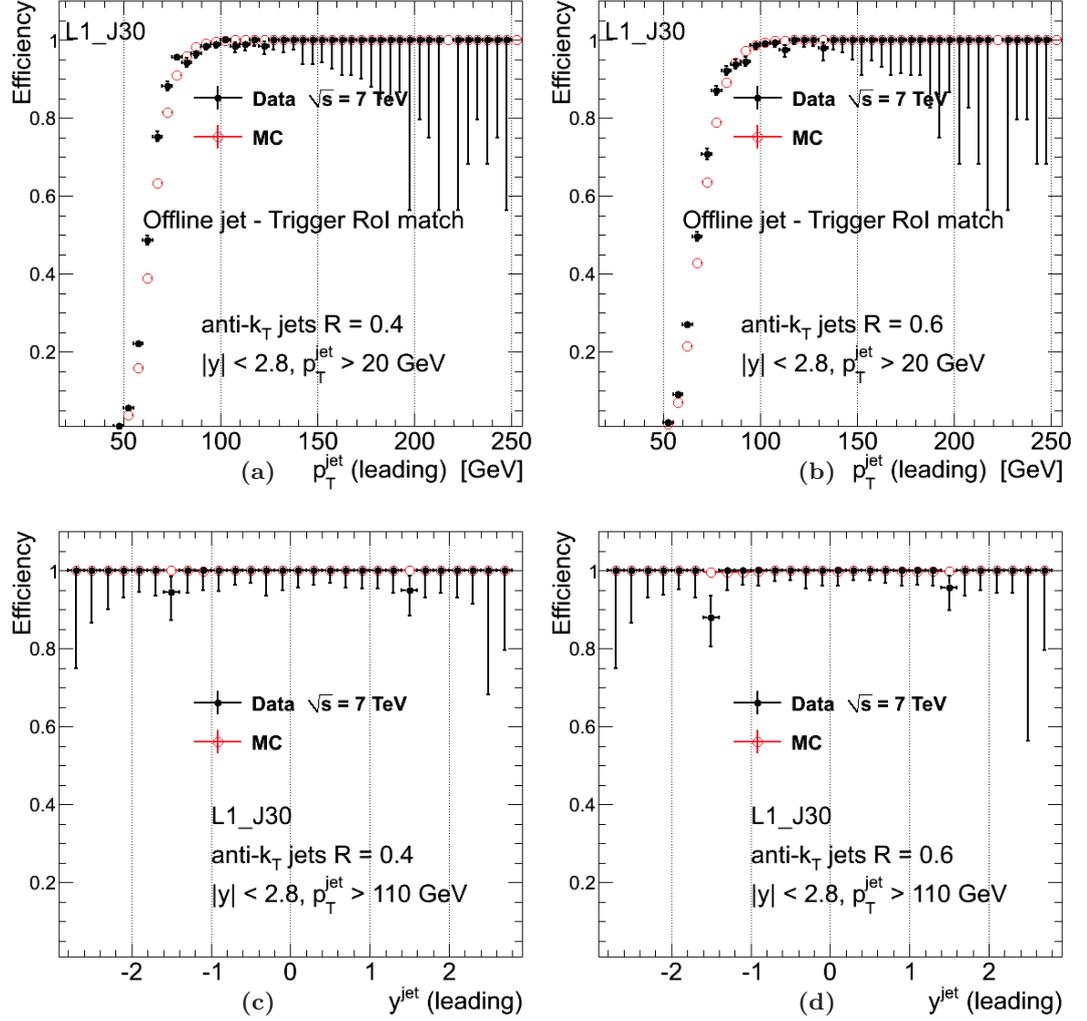


Figure 5.7: Efficiency of the L1_J30 trigger: efficiency as a function of the leading jet p_T for anti- k_T jets with (a) $R = 0.4$ and (b) $R = 0.6$, and efficiency as a function of the leading jet rapidity for events in which the leading jet has $p_T > 110$ GeV, for anti- k_T jets with (c) $R = 0.4$ and (d) $R = 0.6$. The efficiencies are shown as obtained in the Monte Carlo simulation and as calculated using the orthogonal L1_MBTS_1 trigger in data.

to be matched to a level-1 RoI as defined for jet-level efficiencies. Figure 5.8 shows the jet-level efficiency of the L1_2J10 trigger as a function of the p_T of the second leading jet, within $|y| < 2.8$. An offline cut on the jet p_T at 20 GeV is applied. Notably, the L1_J210 trigger selects two-jet events efficiently if the second jet p_T is above 60 GeV. The efficiencies determined in the simulation give a good description of the data in the plateau region. Once the plateau for the second leading jet is known, the same strategy

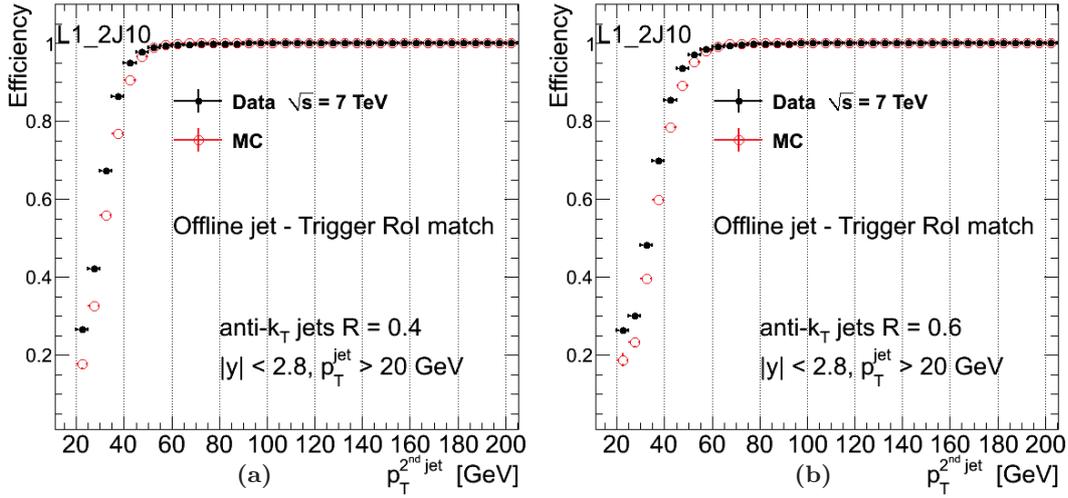


Figure 5.8: Jet-level efficiency of the L1_2J10 trigger as a function of the reconstructed second leading jet p_T within $|y| < 2.8$, for anti- k_T jets with (a) $R = 0.4$ and (b) $R = 0.6$. The efficiencies are shown as calculated in data using the unbiased bootstrap method with the L1_J10 trigger as seed, and as obtained in Monte Carlo simulation.

is used to measure the efficiency of the L1_3J10 trigger. In this case, the two leading jets in the event are required to be in the plateau of the L1_2J10 trigger, besides to be matched to the level-1 RoI. Figure 5.9 shows the efficiency of the L1_3J10 trigger to select three-jet events, as a function of the p_T of the third leading jet. As for the

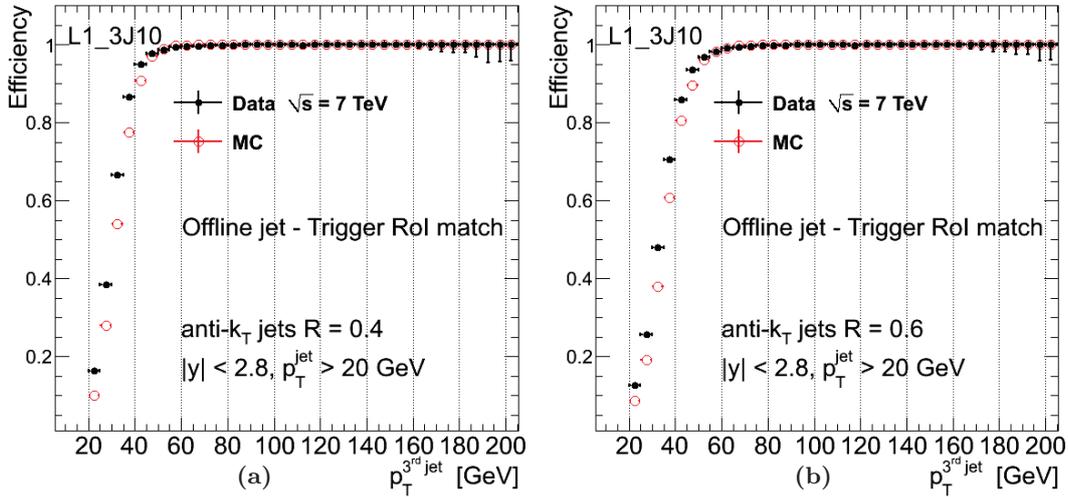


Figure 5.9: Jet-level efficiency of the L1_3J10 trigger as a function of the third leading jet p_T within $|y| < 2.8$, for jets with (a) $R = 0.4$ and (b) $R = 0.6$. The efficiencies are shown as calculated in data using the unbiased bootstrap method with the L1_2J10 trigger as seed, and as obtained in Monte Carlo simulation.

5 Level-1 jet trigger efficiencies

L1_2J10 trigger, data are well described by the Monte Carlo simulation, showing that the three-jet trigger L1_3J10, reaches the plateau region when the third leading p_T is above 60 GeV. Further multi-jet trigger efficiencies have been calculated by using the same bootstrapping strategy. Figure 5.10 shows the efficiencies of the L1_2J15 and L1_2J30 triggers, bootstrapped from the L1_J15 and L1_J30 triggers, respectively. The efficiencies are shown as a function of the p_T of the second leading jet in the event. For $R = 0.4$ jets, the L1_2J15 trigger reaches its maximum efficiency if the second leading jet p_T is above 70 GeV. For $R = 0.6$, the plateau is reached slower, at p_T values of the second leading jet above 80 GeV. The similar comparison is done for the L1_2J30 trigger: the efficiency reaches faster the plateau, at $p_T^{2^{\text{nd jet}}} > 110$ GeV for $R = 0.4$

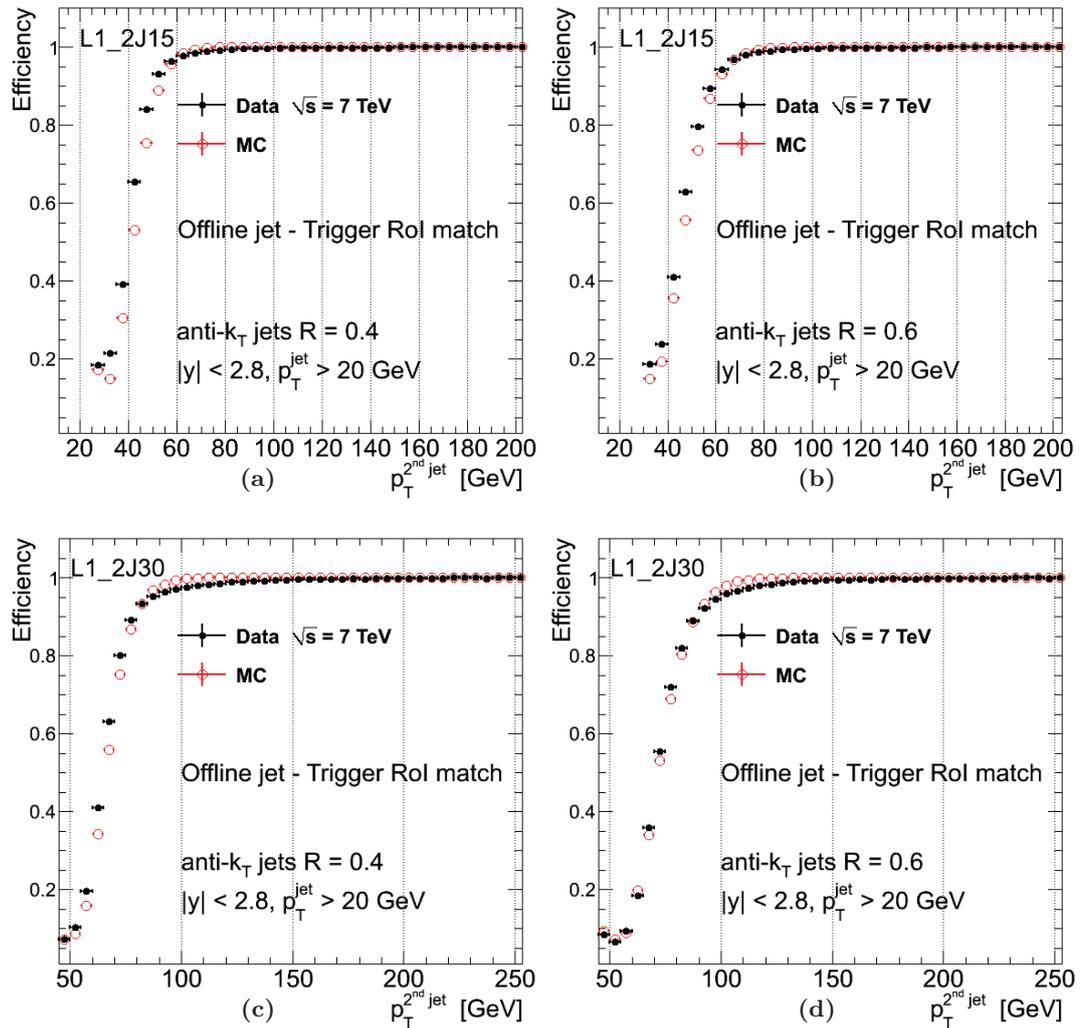


Figure 5.10: Multi-jet jet-level trigger efficiencies as a function of the p_T of the second leading jet: efficiency of the L1_2J15 trigger for (a) $R = 0.4$ jets and (b) $R = 0.6$ jets, and efficiency of the L1_2J30, for jets with (c) $R = 0.4$ and (d) $R = 0.6$. The efficiencies are shown as obtained in the Monte Carlo simulation and as calculated using the bootstrapping method in data.

jets, whereas it is reached at $p_T^{2^{\text{nd jet}}} > 120$ GeV for $R = 0.6$ jets. The inefficiencies in the plateau region of the data efficiencies for the L1_2J30 trigger are caused by the differences in the transition region.

5.4.3 Total jet E_T trigger efficiencies

JE triggers are useful for topologies characterized by high jet multiplicity with large transverse energies. They were tested in analyses aimed to measurements of multi-jet cross-sections, using data corresponding to an integrated luminosity of 17 nb^{-1} . Since JE are global triggers, the obvious way to parametrize their efficiencies is in terms of an offline global-variable: $H_T = \sum p_T^{\text{jet}}$. Figure 5.11 shows the H_T distributions for all the events selected offline (described in Section 4.4.4), and for those that pass three different JE trigger thresholds. Figure 5.12 shows the JE trigger efficiencies as function of the offline-reconstructed H_T , integrated over $|\eta| < 2.8$. The efficiencies determined in simulation provide a reasonable description of the data, for jets for both sizes of jets. As observed for single-jets, the plateau region is reached faster for $R = 0.4$ jets.

The level-1 $\sum E_T^{\text{jet}}$ estimator is sensitive to the size of the window RoI. It works in such a way that e.g. a jet passing L1_J5 but not L1_J10 would contribute 8 GeV to the sum (assume E_T is half way between 5 GeV and 10 GeV, and round up), while a jet passing L1_J10 but not L1_J15 would give an E_T of 13 GeV, etc. This implies larger E_T contributions per jet threshold to the sum, in comparison to the actual values of the jet- E_T . Thus, the sum will exceed the threshold and the efficiencies evaluated in terms of the offline-reconstructed H_T are not accurate. However, this feature seems not to affect anti- k_T with $R = 0.4$ jets in which the plateau is well defined in data and simulation.

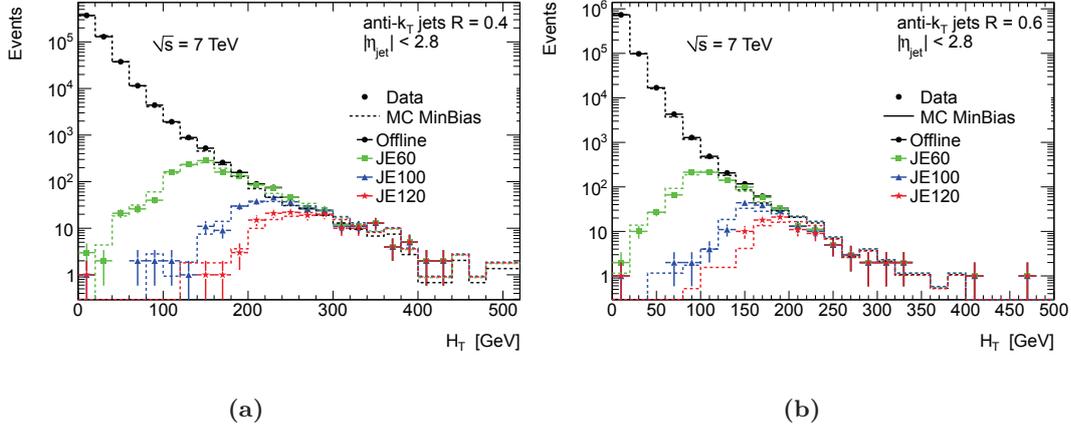


Figure 5.11: H_T distributions for events unbiased by the L1_MBTS_1 trigger and for events that satisfy the JE trigger thresholds, using anti- k_T jets with (a) $R = 0.4$ and (b) $R = 0.6$.

5 Level-1 jet trigger efficiencies

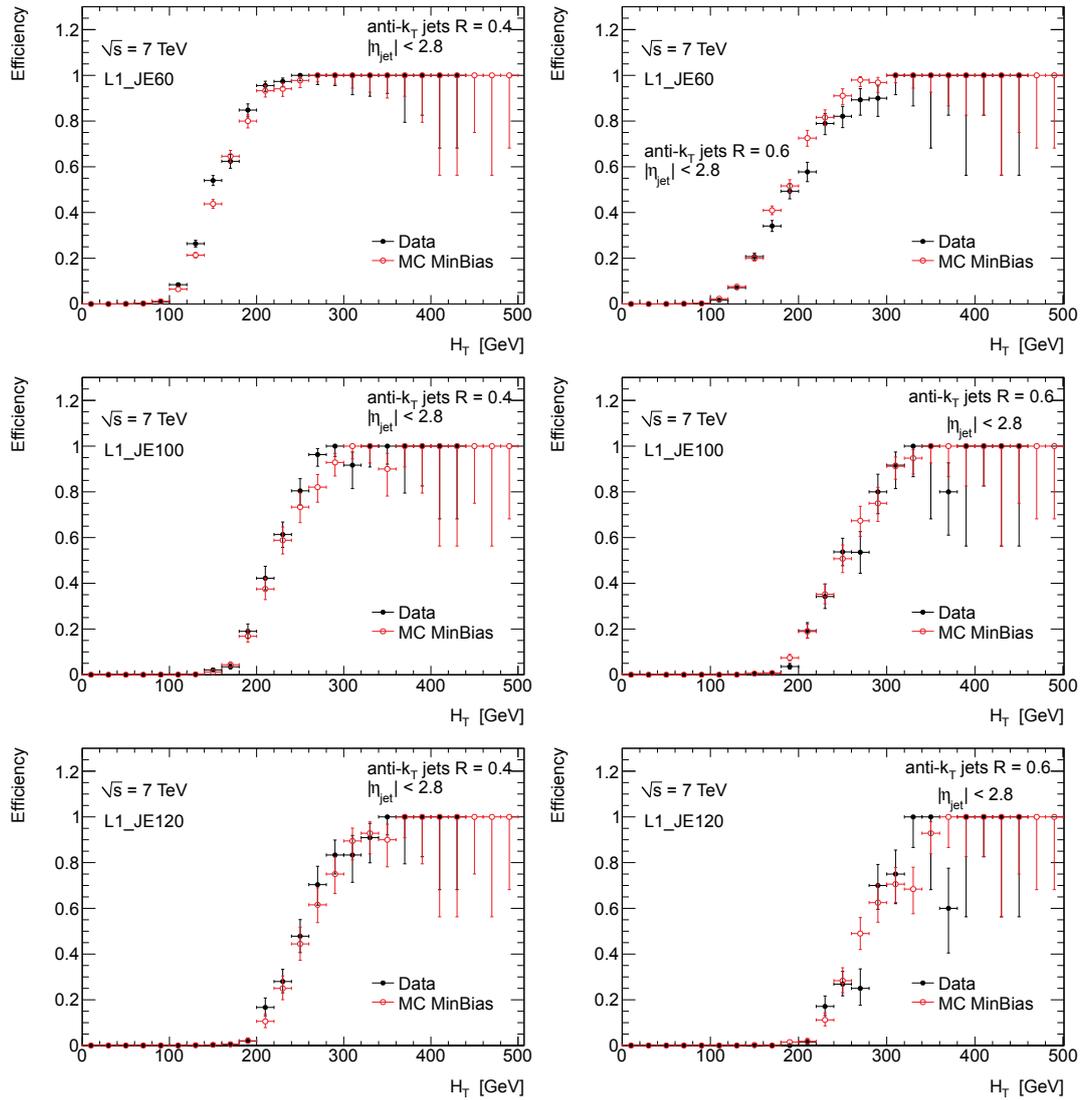


Figure 5.12: Event-level efficiencies of the total transverse jet energy triggers parametrized as a function of the offline-reconstructed H_T .

5.5 Conclusions

The performance of the ATLAS level-1 jet trigger has been shown for the early data at $\sqrt{s} = 7$ TeV. Efficiency turn-on curves were obtained for a variety of jet triggers, showing good agreement between data and simulation in the plateau region. The trigger efficiencies calculated for inclusive, multi-jet and global jet triggers for the level-1 have been used at the first stage of analyses for multi-jet cross-section measurements.

6 Top-quark reconstruction in $t\bar{t}$ resonances searches

The analysis described in the next two chapters involves the search for massive states coupling to top quarks. The leading hypothesis is that $t\bar{t}$ pairs can be produced from the decay of heavy resonances predicted by theories Beyond the Standard Model (BSM). This study is not predisposed to any particular model or subsequent decay signature. However, the two benchmark models introduced in Section 2.4.1 (the leptophobic top-color Z' boson and the KK-gluon, g_{KK}) are used to guide the analysis. The production cross-section for the Z' and g_{KK} signal is assessed across a wide range of masses of the $t\bar{t}$ system. The assessment is carried out by analyzing data events collected by the ATLAS detector at the LHC, with a center-of-mass energy $\sqrt{s} = 7$ TeV.

Top quark pairs can be produced starting at the $t\bar{t}$ production threshold up to a few TeV. Different approaches have to be applied to reconstruct top quarks in a broad range of top quark transverse momenta. These are the *resolved* and *boosted* approaches, and are defined by the final configuration of the top-quark decay products in the detector. Each approach, together with their strategies to select $t\bar{t}$ events are thoroughly described in this chapter. The methods to reconstruct the invariant mass of the $t\bar{t}$ system and the estimation of systematic uncertainties are also discussed.

6.1 Analysis strategy

Searching for new states implies the selection of events produced by the possible signal. The aim is to select from all the events recorded by ATLAS, only those whose characteristics resemble the signal, while rejecting as much background as possible. To achieve this, all events are required to have the desirable $t\bar{t}$ topology (Figure 2.4) of the signal processes. Each selected event enters into a sample hereafter called signal region. Any resonance manifests itself through a local deviation in the reconstructed $t\bar{t}$ invariant mass spectrum ($m_{t\bar{t}}$) from the Standard Model (SM) prediction. The deviation can be identified by counting the number data and predicted background events that lie in the signal region. However, when searching for $t\bar{t}$ signal, there is no way to separate the BSM processes from those produced by the SM $t\bar{t}$ background. Namely, the SM $t\bar{t}$ processes are irreducible “continuum” background, meaning that its rejection leads to the rejection of signal events. On top of that, the uncertainty on the $t\bar{t}$ production cross-section is larger than the number of the expected signal events. Since such a resonance would create a bump in the invariant $t\bar{t}$ mass spectrum¹, one can determine $m_{t\bar{t}}$ in data, and

¹The position of the bump in the $m_{t\bar{t}}$ spectrum depends on the mass of the resonance.

compare it to the roughly exponentially dropping spectrum as predicted by the SM.

The absence of bumps implies that the analyzed dataset is compatible with the SM prediction. In such case, the $m_{t\bar{t}}$ distribution serves to impose upper limits on the cross-section for signal production processes. Specifically, upper limits on the cross-section, times branching ratio for decay to a $t\bar{t}$ pair ($\sigma \times BR(\rightarrow t\bar{t})$) are estimated for different signal masses. Once upper limits on the production cross-section are set, exclusion regions of the benchmark models are inferred within certain mass range.

The resolved and boosted approaches are used to reconstruct the $m_{t\bar{t}}$ distribution. The resolved approach is used at low $t\bar{t}$ masses. Its main feature is that the products from the top quark decay are well-separated in the detector such that they can be reconstructed as individual jets. The boosted reconstruction is advantageous at high $t\bar{t}$ masses where the angles between the decay products become smaller and smaller with the increasing transverse momenta. In this case, the decay products of the hadronically decaying top quark overlap in the detector and are reconstructed in one larger-size jet. The intermediate stage between the two extrema above, is the possibility that only two products from the hadronic top quark decay are reconstructed in a single jet.

This analysis covers a broad range of transverse momenta at which the top quarks can be produced: starting by the resolved approach at low p_T regions, and then passing the intermediate regime up to high p_T regions where the boosted approach becomes more efficient.

The resolved and boosted reconstructions lead to different requirements to select $t\bar{t}$ events. These are two different mechanisms to enhance the Z' and g_{KK} signal over the background fraction. A combination of both approaches would put forward further strength to reject background contributions. Thus, the objective of this analysis is to combine these independent techniques in order to enhance the sensitivity of the search to $t\bar{t}$ resonances. An overlap channel composed by the events selected in the two approaches is created. Different combinations of the resolved, boosted and overlap selection channels are statistically analyzed and for each one, the upper limits on the signal production cross-section are derived.

6.2 Simulation of signal and background processes

The SM processes can be misidentified as the signal of interest at rates large enough to affect the $t\bar{t}$ reconstruction. For example, QCD multi-jet events can mimic the signal signature if a lepton from a b -quark decay or a jet is mistakenly identified as an isolated lepton. To evaluate the sensibility of the analysis to the resonances, signal and SM background samples have been simulated using Monte Carlo (MC) techniques. The generated samples are run through the detector simulation based on GEANT4. Subsequently, simulated events are reconstructed and analyzed using the same software used for data. The generated processes employed in this analysis are listed below.

6.2.1 Background processes

- The irreducible SM $t\bar{t}$ background is the main source of background for this analysis. It has been generated using the MC@NLO v4.01 with the CTEQ6.6 PDF set at NLO accuracy. The showering is generated with HERWIG v6.5 in association with JIMMY (see Section 2.5.4). The top quark mass is set to $m_t = 172.5$ GeV. Only events in which at least one of the W bosons decays leptonically are produced. A K-factor of 1.117 is applied to normalize the samples to the NNLO cross-section of 165 pb [37].
- The electroweak single top-quark production is simulated using the same generators as for the SM $t\bar{t}$ background. For the s -channel and t -channel production, the W boson from the top quark decays leptonically only. The inclusive NNLO cross-sections are: $64.57^{+2.63}_{-1.74}$ pb (t -channel) [38], $4.63^{+0.20}_{-0.18}$ pb (s -channel) [39] and $15.74^{+1.17}_{-1.21}$ pb (Wt -channel) [38].
- Separate samples of W +jets events are considered as background sources. The samples generated for top-quark analyses in ATLAS are:
 1. W +light-jets: The jets arising from the matrix elements are initiated from gluons and from massless u , d , s and c quarks. The b -jets originate from massless b -quarks, and from the parton shower.
 2. $W + b\bar{b}$ +jets: The samples are generated for events with massive b -quarks, produced through the matrix elements.
 3. $W + c(\bar{c})$ +jets, and $W + c\bar{c}$ +jets: Samples with massive c -quarks are produced separately using the matrix elements.

The numerals 2. and 3. are collectively called W +heavy-flavor jets. For the generation of events, ALPGEN v2.13 interfaced as described in Section 2.5.4 is used. The samples are split in bins of parton multiplicity in the final state: five samples with an exclusive number of partons that goes from zero to four, and an inclusive sample with more than four partons. The MLM matching method is used to remove overlaps between samples with the same parton multiplicity from the matrix elements and parton shower. Events are simulated twice: 1) in the inclusive $W + n$ parton samples and 2) in the W with associated heavy quark samples. To remove double-counted events, an overlap removal (HFOR) based on a ΔR matching is used. The application of the HFOR removes all W +light-jets events with c and b quarks from the matrix elements. To reproduce the NNLO cross-section, a K-factor of 1.20 is applied. Table 6.1 shows the cross-sections for the W +jets samples.

- The Z +jets samples are simulated using the same generators and interfaces as for the W +jets samples. Drell-Yan contributions from $\gamma^* \rightarrow \ell\ell$ and Z/γ^* photon interference are included. For the production of Z +jets events, it is required to have a di-lepton invariant mass $40 < m_{\ell\ell} < 2000$ GeV. The corresponding cross-sections are listed in Table 6.2.

6 Top-quark reconstruction in $t\bar{t}$ resonances searches

Sub-sample	Cross-section [pb]	Sub-sample	Cross-section [pb]
<i>W</i> + light-jets		<i>W</i> + heavy-jets	
$W \rightarrow e\nu + \text{Np0}$	6921.60	$W \rightarrow \ell\nu + c + \text{Np0}$	644.40
$W \rightarrow e\nu + \text{Np1}$	1304.30	$W \rightarrow \ell\nu + c + \text{Np1}$	205.01
$W \rightarrow e\nu + \text{Np2}$	378.29	$W \rightarrow \ell\nu + c + \text{Np2}$	50.80
$W \rightarrow e\nu + \text{Np3}$	101.43	$W \rightarrow \ell\nu + c + \text{Np3}$	11.42
$W \rightarrow e\nu + \text{Np4}$	25.87	$W \rightarrow \ell\nu + c + \text{Np4}$	2.80
$W \rightarrow e\nu + \text{Np5}$	7.00	$W \rightarrow \ell\nu + c\bar{c} + \text{Np0}$	127.53
$W \rightarrow \mu\nu + \text{Np0}$	6919.60	$W \rightarrow \ell\nu + c\bar{c} + \text{Np1}$	104.68
$W \rightarrow \mu\nu + \text{Np1}$	1304.20	$W \rightarrow \ell\nu + c\bar{c} + \text{Np2}$	52.08
$W \rightarrow \mu\nu + \text{Np2}$	377.83	$W \rightarrow \ell\nu + c\bar{c} + \text{Np3}$	16.96
$W \rightarrow \mu\nu + \text{Np3}$	101.88	$W \rightarrow \ell\nu + b\bar{b} + \text{Np0}$	47.35
$W \rightarrow \mu\nu + \text{Np4}$	25.75	$W \rightarrow \ell\nu + b\bar{b} + \text{Np1}$	35.76
$W \rightarrow \mu\nu + \text{Np5}$	6.92	$W \rightarrow \ell\nu + b\bar{b} + \text{Np2}$	17.33
$W \rightarrow \tau\nu + \text{Np0}$	6918.60	$W \rightarrow \ell\nu + b\bar{b} + \text{Np3}$	7.61
$W \rightarrow \tau\nu + \text{Np1}$	1303.20		
$W \rightarrow \tau\nu + \text{Np2}$	378.18		
$W \rightarrow \tau\nu + \text{Np3}$	101.51		
$W \rightarrow \tau\nu + \text{Np4}$	25.64		
$W \rightarrow \tau\nu + \text{Np5}$	7.04		

Table 6.1: Leading order cross-sections for $W \rightarrow \ell\ell$ +jets sub-samples generated using ALPGEN+HERWIG+JIMMY. “Np” refers to the number of partons. A K-factor equals to 1.20 is applied to each sample in order to reproduce the NNLO cross-section.

Sub-sample	Cross-section [pb]
$Z \rightarrow ee + 0\text{Np}$	668.32
$Z \rightarrow ee + 1\text{Np}$	134.36
$Z \rightarrow ee + 2\text{Np}$	40.54
$Z \rightarrow ee + 3\text{Np}$	11.16
$Z \rightarrow ee + 4\text{Np}$	2.88
$Z \rightarrow ee + 5\text{Np}$	0.83
$Z \rightarrow \mu\mu + 0\text{Np}$	668.68
$Z \rightarrow \mu\mu + 1\text{Np}$	134.14
$Z \rightarrow \mu\mu + 2\text{Np}$	40.33
$Z \rightarrow \mu\mu + 3\text{Np}$	11.19
$Z \rightarrow \mu\mu + 4\text{Np}$	2.75
$Z \rightarrow \mu\mu + 5\text{Np}$	0.77
$Z \rightarrow \tau\tau + 0\text{Np}$	668.40
$Z \rightarrow \tau\tau + 1\text{Np}$	134.81
$Z \rightarrow \tau\tau + 2\text{Np}$	40.36
$Z \rightarrow \tau\tau + 3\text{Np}$	11.25
$Z \rightarrow \tau\tau + 4\text{Np}$	2.79
$Z \rightarrow \tau\tau + 5\text{Np}$	0.77

Table 6.2: Leading order cross-sections for $Z \rightarrow \ell\ell$ + jets sub-samples generated using ALPGEN+HERWIG+JIMMY. A K-factor equals to 1.25 is applied to each generated sample in order to reproduce the NNLO cross-section.

Subsample Topcolor Z'	Cross section $\times \text{BR}(Z' \rightarrow t\bar{t})$ (pb)	Subsample g_{KK}	Cross section $\times \text{BR}(g_{KK} \rightarrow t\bar{t})$ (pb)
$m_{Z'} = 500$ GeV	19.60	$m_{g_{KK}} = 700$ GeV	20.80
$m_{Z'} = 600$ GeV	10.30	$m_{g_{KK}} = 800$ GeV	11.60
$m_{Z'} = 700$ GeV	5.60	$m_{g_{KK}} = 900$ GeV	6.80
$m_{Z'} = 800$ GeV	3.20	$m_{g_{KK}} = 1000$ GeV	4.10
$m_{Z'} = 1000$ GeV	1.20	$m_{g_{KK}} = 1150$ GeV	2.58
$m_{Z'} = 1300$ GeV	0.29	$m_{g_{KK}} = 1300$ GeV	1.09
$m_{Z'} = 1600$ GeV	0.09	$m_{g_{KK}} = 1600$ GeV	0.35
$m_{Z'} = 2000$ GeV	0.02	$m_{g_{KK}} = 1800$ GeV	0.18
$m_{Z'} = 3000$ GeV	0.01	$m_{g_{KK}} = 2000$ GeV	0.09

Table 6.3: Production cross sections times branching ratio (BR) for the resonant signal processes $pp \rightarrow Z' \rightarrow t\bar{t}$ in the topcolor model and $pp \rightarrow g_{KK} \rightarrow t\bar{t}$ for the KK gluon in Randall–Sundrum models with warped extra dimensions.

- The simultaneous production of two vector bosons (diboson) has been generated using HERWIG v6.5 associated with JIMMY and the MRST2007LO* [149] PDF set. Both bosons can decay leptonically and hadronically. Thus, the final configuration includes a charged lepton, missing transverse momentum and jets. A filter selecting events with at least one lepton with $p_T > 10$ GeV and $|\eta| < 2.8$ is applied. The corresponding cross-sections (K-factors at NLO) are: 11.50 pb (1.48) for WW production, 3.46 pb (1.60) for WZ production, and 0.97 pb (1.30) for ZZ production.

6.2.2 Benchmark signal processes

For the simulation of processes involving the leptophobic topcolor Z' boson, PYTHIA v6.421 [66] with the CTEQ6L1 PDF set is used. A K-factor equals to 1.3 is applied to each sample in order to account for NLO corrections.

The samples for the KK-gluon are generated with MADGRAPH v4.4.51 [48] and showered with PYTHIA. Both use the CTEQ6L1 PDF set. Since the background and signal samples were produced separately, the effect of interferences with the SM processes is ignored. No K-factor is applied as there is no published cross-sections at higher orders. The cross-section times branching ratio ($\sigma \times BR$) for the signal subsamples is shown in Table 6.3.

6.3 Data sample

The data used were collected by the ATLAS detector between 21 March and 4 August 2011. This corresponds to an integral luminosity of $2.05 \pm 0.08 \text{ fb}^{-1}$ [91]. Of this, 0.87 fb^{-1} were affected by the LAr failure described in Section 3.6.1. These data were processed with the release 16 of the ATLAS software and split up into periods from B to K (see Sections 3.5 and 3.6). The center-of-mass energy of the pp collisions producing the data was 7 TeV. Only data taken under full-operational conditions of the ATLAS sub-detectors were analyzed. In this way, the data were first required passing the DQ

criteria as defined in Section 3.5, within luminosity blocks accepted by the *Good Runs List* (GRL). The GRL was provided by the Top Physics Group of ATLAS [150].

6.4 Selection of events in $t\bar{t}$ topologies

This analysis is carried out in the semi-leptonic ℓ +jets channel of the $t\bar{t}$ decay. As mentioned in Section 2.3.3, the top quarks decay to a W boson and a b -quark. One W boson from one top quark decays to a lepton and a neutrino, and the other to a pair of quarks which subsequently hadronize (see Figure 2.4). The leptons can be either an electron (e +jets channel) or a muon (μ +jets channel). τ -leptons are not considered because they are difficult to identify from the electron or muon decays, and thus contribute to the signal². The reconstructed final-state objects of the selected events are therefore leptons, the $E_{\text{T}}^{\text{miss}}$ due the escaping neutrino and jets. All objects are required to pass the identification criteria specified in Chapter 4. In addition, they should have high transverse momentum (p_{T}), since the masses of the Z' and the g_{KK} resonances are of the order of several top quark masses.

Events are first filtered either by a single-electron trigger (electron channel) or a single-muon trigger (muon channel). These triggers varied according to the data periods as shown in Table 6.4. The plateau region of the efficiency turn-on curve for the triggers indicates the E_{T} or p_{T} thresholds to select the reconstructed leptons (See Chapter 5).

Data Period	$\int \mathcal{L}$ (pb ⁻¹)	Electron Trigger	Muon Trigger
B-I	1340.3	EF_e20_medium	EF_mu18
J	212.2	EF_e20_medium	EF_mu18_medium
K	500.0	EF_e22_medium	EF_mu18_medium

Table 6.4: Single-lepton triggers used in this analysis. Triggers are shown as applied according to the data periods in 2011 for an integral luminosity of $2.05 \pm 0.08 \text{ fb}^{-1}$.

After passing a single-lepton trigger, the reconstructed primary event vertex is required to have at least five tracks with $p_{\text{T}} > 0.4 \text{ GeV}$. This is applied to avoid contributions from noise in the calorimeters, cosmic rays or beam background. Following the selection, additional criteria on the reconstructed objects have been imposed.

Events in the e +jets channel must contain one **tight** electron with $E_{\text{T}} > 25 \text{ GeV}$ and within the volume of the inner detector: $\eta < 2.47$. Each electron is required to be isolated with **EtCone20** $< 3.5 \text{ GeV}$ and **PtCone30** $< 4.0 \text{ GeV}$ (Section 4.2.1). To avoid a decrease in the selection efficiency and fakes, events with electrons in the transition region $1.37 < |\eta| < 1.52$ or near the LAr problematic region are excluded.

² τ -leptons which decay solely to hadrons may be misidentified as electrons, as the τ -jets are similarly narrow and contain low track multiplicity.

In the μ +jets channel events must have one **tight** combined muon with $p_T > 20$ GeV and $|\eta| < 2.5$. The muon has to be isolated with $\text{EtCone30} < 4$ GeV and $\text{PtCone30} < 4$ GeV (see Section 4.3.1). To reduce the amount of non-prompt muons, e.g. those produced in jets coming from B -hadrons, reconstructed muons close to jets within a cone of $\Delta R = 0.4$ are rejected. These cut is imposed only when the jets have a p_T larger than 20 GeV, and allows the suppression of multi-jet background contributions, which have no real isolated leptons.

To ensure only one lepton in the final states, events in the electron channel must not contain muons with $p_T > 20$ GeV. Similarly in the muon channel, events containing electrons with $E_T > 25$ GeV are vetoed for further selection. Events are also removed if the accepted electron shares an inner detector track with a reconstructed muon. This condition guarantees that a unique energy deposit in the calorimeter has not been used to reconstruct both leptons.

Additional constraints on the events are enforced to reduce multi-jet events with a jet faking an electron. Specifically, thresholds are applied on the missing transverse momentum E_T^{miss} , and on the transverse mass of the W boson, defined as

$$m_T(W) = \sqrt{2p_T^\ell E_T^{\text{miss}}(1 - \cos(\Delta\phi))}, \quad (6.1)$$

where p_T^ℓ is the p_T of the selected lepton and $\Delta\phi$ is the angle between the lepton and the neutrino measured in the transverse plane (the neutrino is reconstructed as described in Section 4.6.1). The thresholds are:

- $E_T^{\text{miss}} > 35$ GeV, $m_T(W) > 25$ GeV in the electron channel and
- $E_T^{\text{miss}} > 20$ GeV, $E_T^{\text{miss}} + m_T(W) > 60$ GeV in the muon channel.

The cut in the muon channel excludes a triangular region in the E_T^{miss} - $m_T(W)$ plane. This is motivated by the fact that $W \rightarrow \ell\nu$ events with large E_T^{miss} also have large $m_T(W)$. This is true when the W boson arises directly from top quark decay, but is not reproduced in events with non-prompt leptons like multi-jet events. In the electron channel a more stringent cut on $m_T(W)$ is required because of the more important QCD background.

Jets³ differ for the cases of resolved or boosted selection. For the resolved case only anti- k_T $R = 0.4$ jets are selected. In the boosted approach two jets are required: one anti- k_T $R = 0.4$ jet to complete the final configuration of the leptonic top quark, and one anti- k_T $R = 1.0$ jet which encompasses the majority of the decay products of the hadronic top quark. The configuration of the jets in the final state defines the topology of a $t\bar{t}$ event, i.e. as resolved or boosted. The selection of jets for the reconstruction of each topology is given individually in the next sections.

³The reconstruction and calibration of the jets used in this analysis is discussed in Section 4.4.1.

6.4.1 Jets and event topology in the resolved reconstruction

The resolved selection requires four or three anti- k_T $R = 0.4$ jets in the final state, each having $p_T > 25$ GeV and $|\eta| < 2.5$. Three of the jets can be assumed to result from the hadronic decay of the top quark, while the fourth is from the b -jet of the leptonically decaying W boson (Section 4.5.1). To avoid double counting of electrons as jets, the closest jet within $\Delta R < 0.2$ from a selected electron is removed. Further contamination from W/Z +jets and multi-jet backgrounds usually contain low p_T jets in their final states. This contribution is reduced by requiring events whose jet with the highest p_T (leading p_T jet) has $p_T > 60$ GeV.

The resolved approach relies on the fact that the jets from the hadronic decay are sufficiently separated in space to allow their individual reconstruction. Figure 6.1a is a sketch of each parton from the hadronic top quark decay, surrounded by a cone of radius $R = 0.4$. The hadronic top is not specifically reconstructed, so this is just a basic representation. As the p_T of the top quark increases, two of the partons resulting from top quark decay cannot be reconstructed as individual jets and hence, are merged into a single jet. Figure 6.1b illustrates this situation. The p_T of the top quark and the radius of the cone containing its decay products are related by [151]:

$$\Delta R \approx \frac{2m_t}{p_T}, \quad (6.2)$$

where m_t is the mass of the top quark and ΔR is the angular separation of the its decay products in $(\eta - \phi)$ -space. As three jets are involved in the hadronic decay, $\Delta R = 0.4 \times 3 = 1.2$, leading to $2m_t/1.2 \approx 300$ GeV. Therefore, the resolved approach is adequate for reconstructing top quarks decaying with p_T up to ≈ 300 GeV.

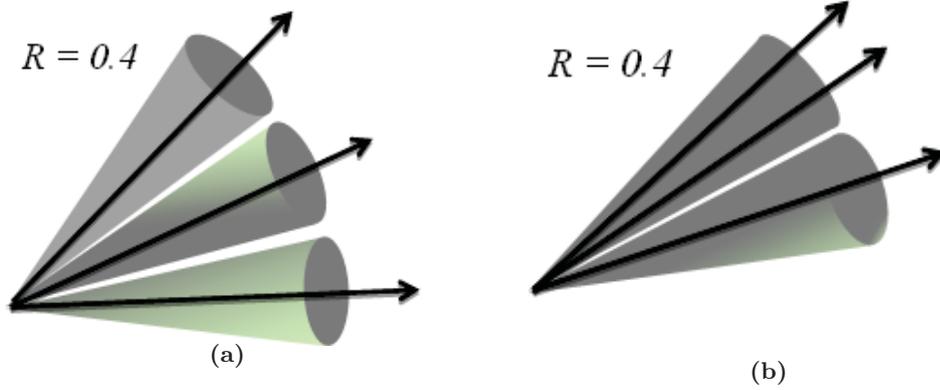


Figure 6.1: Representation of the jets in the: (a) resolved approach and (b) intermediate stage between the resolved and boosted approaches of the $t\bar{t}$ reconstruction. The representation of the jets in the boosted reconstruction is shown in Figure 6.3. Each arrow represents a hard parton arising from the hadronic decay of the top quark. The cones represent the jets used during the reconstruction. The topology of the jets becomes more collimated as the p_T of the top quark increases from 6.1a to 6.1b.

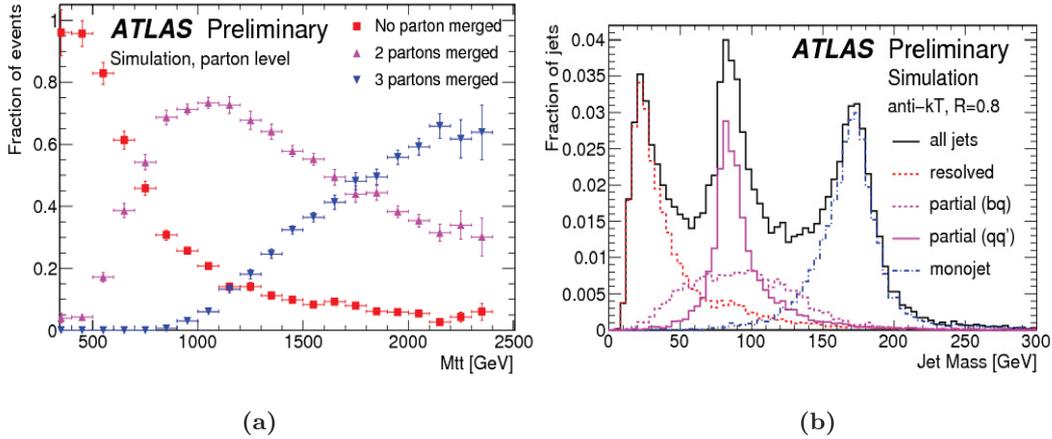


Figure 6.2: (a) Fraction of events with a certain reconstructed $t\bar{t}$ mass in which three partons resulting from the hadronically decaying top quark are found within a cone of radius $\Delta R = 0.8$, or merged. The pink triangles show events in which two partons are merged, but the third remains well-separated. The blue triangles are events in which all three partons are merged. (b) Reconstructed invariant mass of the leading anti- k_T $R = 0.8$ jet in $pp \rightarrow X \rightarrow t\bar{t} \rightarrow \ell + \text{jets}$ events. The resolved topology with no partons merged is represented by the red dashed line. The partially merged topology is represented by the pink lines and the blue lines represent events where all three partons have been merged. Each event is classified on basis of the ΔR matching of the quarks to the reconstructed jet [152].

For a wide range of $t\bar{t}$ masses, Figure 6.2a shows the fraction of events in which three partons resulting from the hadronically decaying top quark are found within a cone of radius $\Delta R = 0.4 \times 2 = 0.8$, or merged. Clearly, the resolved approach in which no partons are merged dominates at $m_{t\bar{t}} = 2m_t$. The decay products of the hadronic top quark are contained in a more collimated configuration when moving from $m_{t\bar{t}} = 2m_t$ to the TeV-regime. Events falling in the region $800 < m_{t\bar{t}} < 1200$ GeV have probabilities up to 70% that two of the three partons from the decay will be merged.

The jet invariant mass⁴ (m^{jet}) provides sensitive information about the parton merging [151, 152]. Namely, for partially (fully) merged topologies, quarks from the W boson (top quark) decay will be merged in one jet whose mass is distributed around the W boson (top quark) mass. Figure 6.2b displays different degrees of parton merging using anti- k_T $R = 0.8$ jets. Notably, events in which the leading jet has $m^{\text{jet}} \gtrsim 60$ GeV, most of the jets are merging at least two partons. This corresponds to the illustration shown in Figure 6.1b. Thus, when the leading jet mass is above 60 GeV, the minimum number of required jets is reduced from four to three. The selected four or three jets, the lepton and the neutrino identified as E_T^{miss} make up the $t\bar{t}$ events of the resolved

⁴At the reconstruction level, the invariant jet mass is calculated from the four-vectors of the topological-calorimeter clusters.

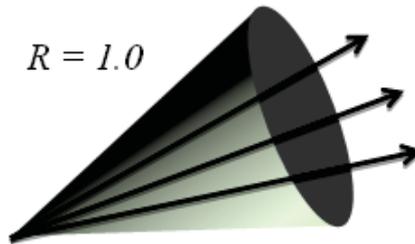


Figure 6.3: Configuration of partons in the boosted reconstruction. A further increase in the p_T of the top quark will yield an overlapping of the jets represented in Figure 6.1b. That leads to $\Delta R = 0.8$, which corresponds to top quarks with $p_T \gtrsim 450$ GeV.

reconstruction. The whole $t\bar{t}$ system is reconstructed without attempts to reconstruct individual top quarks.

6.4.2 Jets and event topology in the boosted reconstruction

If the hadronic top quark is produced with a large p_T compared to its mass, its decay products tend to be emitted in a *back-to-back* topology in ϕ to the leptonically decaying top quark⁵. The decay products are closer to each other than $\Delta R = 0.4 \times 3 = 1.2$, i.e. the radius of the cone that contains them should be smaller than 1.2. The boosted reconstruction uses a “fat” jet of radius $R = 1.0$. This size is good enough to overcome the partially merged topology represented in Figure 6.1b, while not too large such that the mass resolution is affected due to pile-up effects inside the jet. The merging scheme is illustrated in Figure 6.3.

Since the $R = 1.0$ jet contains most of the products resulting from the hadronic top decay, it is identified as the hadronic top quark candidate of the $t\bar{t}$ pair: the “top jet”. The selection of the top jet is carried out in a “top-tagging” procedure [153]. The idea behind the method is to explore the substructure of the jet to determine whether it is a top jet or a QCD-jet, i.e. initiated by gluons and light quarks. Contributions from multi-jet backgrounds can thus be discriminated and reduced. The substructure variables are the jet mass m_j , and the first k_T -splitting scale $\sqrt{d_{12}}$, defined in Section 4.4.1. Figure 6.4 shows their corresponding distributions for separated samples of top- and QCD-jets. In both samples, the jets are required having $m_j > 100$ GeV and $p_T > 200$ GeV. If the decay products of the top quark are fully contained in the $R = 1.0$ jet, its invariant mass will peak around the top-quark mass. This behavior is evident in the simulated signal containing $Z' \rightarrow t\bar{t}$ events but not in the QCD-jets. The separation between top- and QCD-jets occurs when $m_j \gtrsim 150$ GeV and $\sqrt{d_{12}} \gtrsim 80$ GeV. However, Figure 6.5a shows the m_j and $\sqrt{d_{12}}$ distributions for all reconstructed $R = 1.0$ jets

⁵The two top quarks are emitted in opposite directions in the transverse plane, and the top quark boost ensures that their decay products retain the approximate direction of the top quarks. Top quarks undergoing such a high Lorentz boost are known as boosted tops.

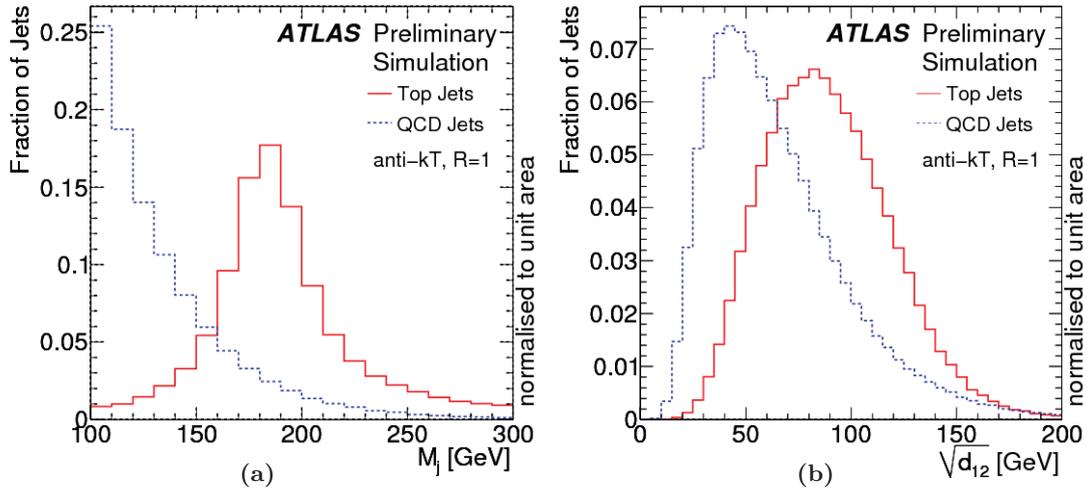


Figure 6.4: Discriminant variables in the boosted approach of the top-quark reconstruction: (a) jet mass m_j and (b) first k_T -splitting scale $\sqrt{d_{12}}$, for $R = 1.0$ jets containing top quark decay products, i.e. top-jets (red continuous line), and QCD-jets (blue dashed line). The signal sample of $Z' \rightarrow t\bar{t}$ events is simulated with $m_{Z'} = 1$ TeV and $m_{Z'} = 2$ TeV. QCD-jets are obtained from di-jet events simulated with PYTHIA [152].

having $p_T > 150$ GeV. Only $\approx 30\%$ of the fat jets from signal events would pass these cuts. To ensure the retaining of signal statistics while the background contributions are depleted, the thresholds are reduced to $m_j > 100$ GeV and $\sqrt{d_{12}} > 40$ GeV.

Although a $R = 1.0$ cone size would correspond to top quarks with $p_T \gtrsim 350$ GeV (see Equation 6.2), the p_T of the top jet is required to be larger than 250 GeV. The justification for lowering the threshold will be detailed later.

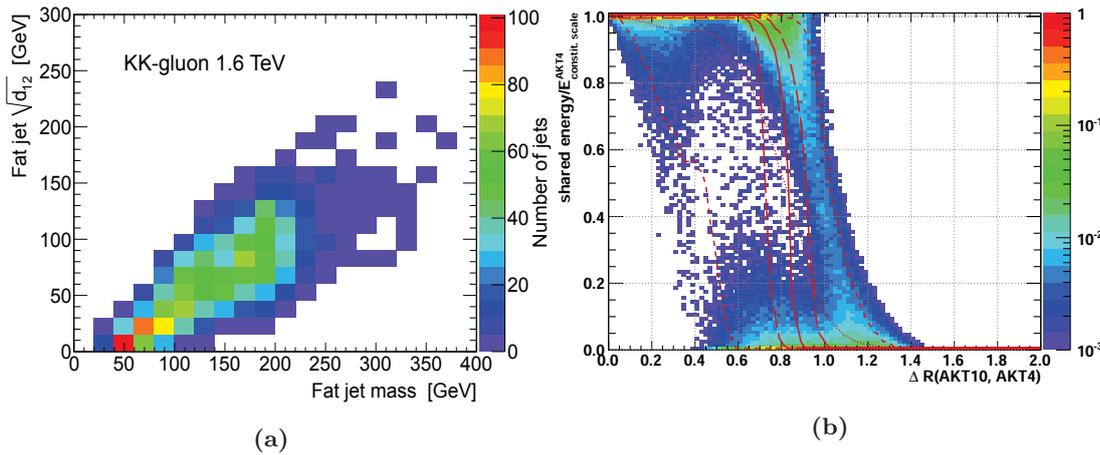


Figure 6.5: Jet substructure variables and jet energy-sharing in the boosted selection: (a) Fat jet mass and first k_T -splitting scale distributions for $R = 1.0$ jets having $p_T > 150$ GeV, in a signal sample of $g_{KK} \rightarrow t\bar{t}$ events with $m_{g_{KK}} = 1.6$ TeV. (b) Energy sharing between hadronic and leptonic jets in a sample of $Z' \rightarrow t\bar{t}$ simulated events with $m_{Z'} = 2$ TeV. 91

Events are selected if they contain at least one top jet satisfying the above thresholds. The hardest of these jets is the hadronic top jet of the decay: the hadronically decaying top quark.

The highly-collimated topology of boosted approach would ideally suggest an overlap of the leptonic decay products, similar to the hadronic side of the decay. However, the reconstruction of the top quark requires isolated leptons, so they cannot be merged within a jet⁶. Therefore, the resolved strategy of employing separately the selected lepton, one $R = 0.4$ jet, and the E_T^{miss} , is used in the boosted case to reconstruct the leptonically decaying top quark. The boosted selection aims to preserve the event topology as much collimated as possible, though. The specific criteria are listed as follows:

- The lepton and the $R = 0.4$ jet must be close to each other without overlapping. Events are accepted only if they contain at least one jet close to the lepton, within a distance $0.4 < \Delta R$ (jet-lepton) < 1.5 . If the event has more than one $R = 0.4$ jet, the nearest to the lepton is taken to be from the leptonic top quark decay.
- Fat jets overlapping with the jet from the leptonic top decay are discarded. This guarantees that the two jets have not been reconstructed from the same topoclusters. Figure 6.5b exhibits the relative energy sharing between the jets from the hadronic and leptonic decays in the boosted reconstruction. Since, no clusters are shared for distances $\Delta R > 1.5$, any jet from the hadronic top decay is removed from the selection if ΔR ($R = 0.4$ jet, $R = 1.0$ jet) < 1.5 .

6.4.3 Selection efficiencies and overlap channel

The described selection cuts applied on the physics objects are summarized in Table 6.5. The specific criteria for each approach introduced in the previous section can be identified. Figure 6.6 shows the efficiencies for selecting $Z' \rightarrow t\bar{t}$ events. They are obtained separately from the fraction of events passing the resolved and boosted selections, out of the total number of events in each Z' sample. At masses $m_{Z'} < 1$ TeV the resolved selection is dominating, whereas the boosted selection becomes the most efficient at masses $m_{Z'} > 1.3$ TeV. This is because the probability for selecting events increases as the boost of the top quark increases: the merging of the decay products in the top jet, leads to have higher p_T and mass, large enough to pass the top-tagging criteria. An increasing efficiency from low to higher masses is beneficial to enhance sensitivity to TeV-scale resonances. However, this behavior is not fully reached and the efficiency starts to fall at $m_{Z'} \approx 1.6$ TeV. The decrease is related to the lepton isolation criteria. Likely in the resolved selection, events fail the requirement when the lepton and jet resulting from the leptonically decaying W boson are close-by. In the boosted selection, the lepton isolation itself restricts significantly the event acceptance, as a highly collimated topology is required.

⁶The isolation requirement reduces the selection efficiency for the boosted events, but is needed to reject overwhelming multi-jet backgrounds. Recent analysis in ATLAS redefine the lepton isolation to increase the selection efficiency [154].

Physics object	Resolved selection	Boosted selection
Electrons	tight $E_T > 25$ GeV $ \eta < 1.37$ or $1.52 < \eta < 2.47$ EtCone20 < 3.5 GeV PtCone30 < 4.0 GeV	tight $E_T > 25$ GeV $ \eta < 1.37$ or $1.52 < \eta < 2.47$ EtCone20 < 3.5 GeV PtCone30 < 4.0 GeV
Muons	tight combined $p_T > 20$ GeV $ \eta < 2.5$ EtCone30, PtCone30 < 4 GeV $\Delta R(\text{muon}, \text{jet}) < 0.4$ if $p_T^{\text{jet}} > 20$ GeV	tight combined $p_T > 20$ GeV $ \eta < 2.5$ EtCone30, PtCone30 < 4 GeV $\Delta R(\text{muon}, \text{jet}) < 0.4$ if $p_T^{\text{jet}} > 20$ GeV
E_T^{miss} El. chn. E_T^{miss} Mu. chn. QCD cut El. chn. QCD cut Mu. chn.	$E_T^{\text{miss}} > 35$ GeV $E_T^{\text{miss}} > 20$ GeV $m_T(W) > 25$ GeV $E_T^{\text{miss}} + m_T(W) > 60$ GeV	$E_T^{\text{miss}} > 35$ GeV $E_T^{\text{miss}} > 20$ GeV $m_T(W) > 25$ GeV $E_T^{\text{miss}} + m_T(W) > 60$ GeV
anti- k_T $R = 0.4$ jets	If $m^{\text{jet}} > 60$ GeV $N = 3$ else $N = 4$ 1 tagged as b -jet leading jet in p_T with $p_T > 60$ GeV remove jet if $\Delta R(\text{jet}, e^{\text{sel}}) < 0.2$ —	$N = 1$ — — remove jet if $\Delta R(\text{jet}, e^{\text{sel}}) < 0.2$ $0.4 < \Delta R(\text{jet-lepton}) < 1.5$
anti- k_T $R = 1.0$ jets	— — —	$m_j > 100$ GeV, $\sqrt{d_{12}} > 40$ GeV $p_T > 250$ GeV remove if $\Delta R(\text{jet}^{0.4}, \text{jet}^{1.0}) < 1.5$

Table 6.5: Summary of the object selection cuts for the resolved and boosted approaches of the $t\bar{t}$ reconstruction. N indicates the jet multiplicity required in each selection. The specific cuts in the electron (El.) and muon (Mu.) channels (chn.) are listed. These object cuts are added to the initial requirements related to GRL, vertex and trigger. Events whose physics objects pass the criteria in each selection, are used to reconstruct the $t\bar{t}$ system of the corresponding approach. Events passing all the listed criteria are used to create the overlap channel.

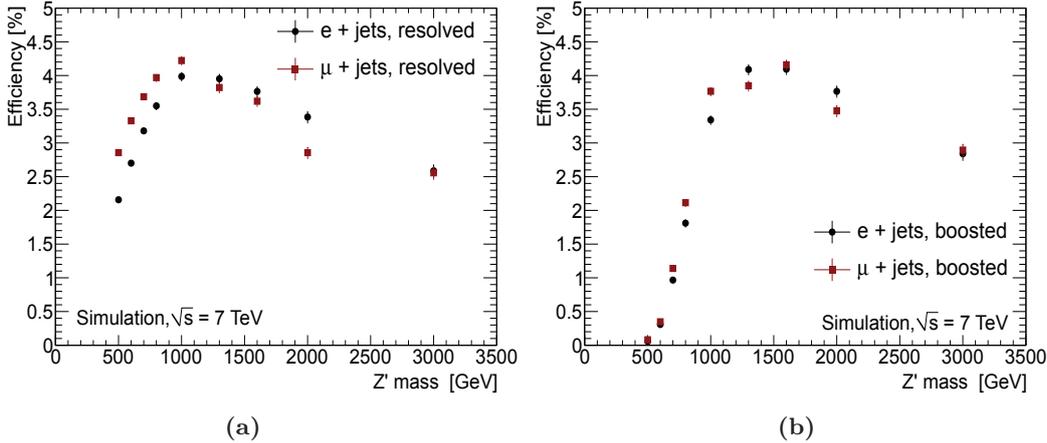


Figure 6.6: Selection efficiencies for the (a) resolved and (b) boosted approaches, as a function of the resonance mass. Only events from the leptophobic Z' benchmark model are considered. The error bars indicate the Monte Carlo statistical uncertainty.

6 Top-quark reconstruction in $t\bar{t}$ resonances searches

Channel	Resolved selection			Overlap selection			Boosted selection		
	Efficiency [%]			Efficiency [%]			Efficiency [%]		
	$t\bar{t}$	Bkg.	g_{KK}	$t\bar{t}$	Bkg.	g_{KK}	$t\bar{t}$	Bkg.	g_{KK}
e +jets	4.18	0.03	3.62	0.27	0.10×10^{-3}	2.24	0.28	0.03	3.85
μ +jets	5.35	0.03	3.99	0.30	1.44×10^{-3}	2.95	0.33	0.04	4.45

Table 6.6: Selection efficiencies in the resolved, overlap and boosted channels, for selecting SM $t\bar{t}$ background ($t\bar{t}$), electroweak backgrounds (Bkg, by adding the W +jets, Z +jets, single-top and diboson backgrounds) and g_{KK} signal with $m_{g_{KK}} = 1.3$ TeV.

As mentioned in Section 6.1, the resolved and boosted selections are combined as further mechanism to reject SM background processes. The “overlap” channel contains only events passing *both* the resolved and boosted criteria. It means that they are conditioned by the two different selections listed in Table 6.5. The overlap events are acquired separately for each sample of data, background and signal, and in the two decay channels, i.e. e +jets and μ +jets. Table 6.6 shows the overlap selection efficiency, compared to that obtained in the original resolved and boosted channels. The results are shown for all the backgrounds comprising the analysis, as well as for a representative signal sample.

The overlap selection of course results in a decrease of the background efficiencies, and is not highly detrimental for selecting signal events.

6.5 Reconstruction of the $t\bar{t}$ invariant mass

The selected objects are brought together to reconstruct the $t\bar{t}$ system and from it, the $t\bar{t}$ invariant mass is determined. Two methods to reconstruct the $t\bar{t}$ invariant mass, depending on the event topology are discussed in this section. In both methods, the smallest $|p_z|$ solution is taken for the neutrino reconstruction (see Section 4.6.1). The invariant mass distribution and the resolution for narrow and broad resonances are also introduced.

6.5.1 The $t\bar{t}$ system in the resolved approach

The $t\bar{t}$ system is reconstructed in the resolved approach by using the so-called dR_{\min} method [59, 155]. It is a refinement of a primary procedure called the four-hardest jets method [155]. To reconstruct the $t\bar{t}$ pairs, the four selected highest p_T jets in each event, are combined with the lepton and the neutrino. If one of the jets has mass greater than 60 GeV, only three hardest jets are used, since two jets from the top quark decay are already merged. The method is affected by long, non-gaussian tails in the $m_{t\bar{t}}$ distribution [59]. This is because one or more jets used in the reconstruction can arise from initial state radiation (ISR), instead of being produced from the top quark decay. The four-hardest jets method is also sensitive to pile-up effects. As pile-up increases, the probability of using a pile-up jet instead of a jet from the top quark increases as well [59, 140].

6.5 Reconstruction of the $t\bar{t}$ invariant mass

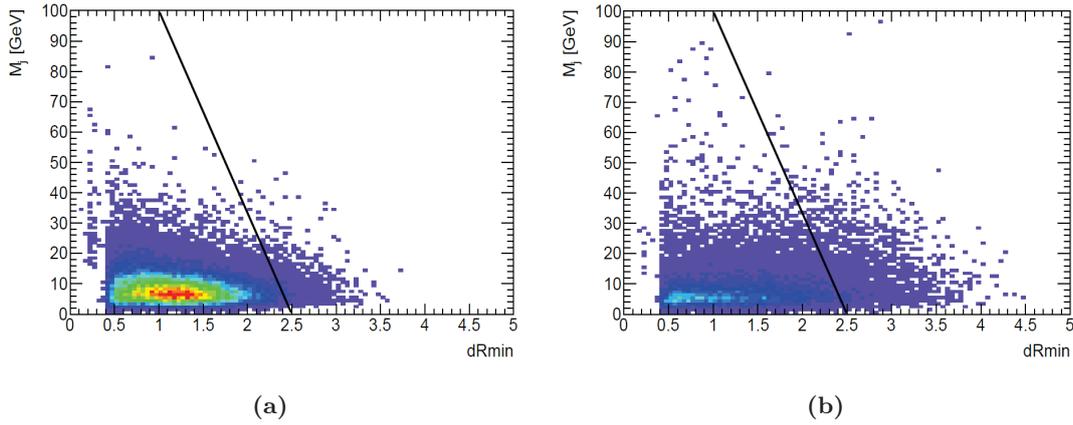


Figure 6.7: Correlation between the angular separation to the closest jet and jet mass in the resolved approach of the $t\bar{t}$ reconstruction, for jets (a) matched and (b) not matched to the top-quark decay products. Jets to the right of the black line are rejected for further consideration. The absolute color scale is the same in the two Figures. Only SM $t\bar{t}$ events are used [59, 156].

The dR_{\min} method aims to reduce the contamination from ISR and pile-up jets. The method explores the correlation between the jet mass and the distance to the closest object. The dR_{\min} algorithm considers the four highest p_T jets in the event, and removes any jet if its angular distance to the lepton or closest jet satisfies $\Delta R_{\min} > 2.5 - 0.015 \times m^{\text{jet}}$. Here, m^{jet} is the mass of the jet in units of GeV. If more than one jet fulfills this condition, the jet with the largest ΔR_{\min} is excluded. In addition, if a jet is discarded and more than three jets remain, the procedure is iterated [155]. Figure 6.7 shows the correlation between the ΔR to the closest jet, and m^{jet} for both, matched and not matched jets to top-quark decay products. Only SM $t\bar{t}$ events have been used. The $t\bar{t}$ system is then reconstructed as a whole with the selected lepton, the reconstructed neutrino and the four or three leading p_T jets, if only three jets remain. The latter happens when one of the jets of the top quark decay is outside of the detector acceptance or merged with another jet.

Although the resolved approach does not reconstruct individual top quarks, the dR_{\min} method is slightly redefined allowing such possibility. Specifically, the strategy to reconstruct the $t\bar{t}$ system is adapted when only three jets survive after the final selection. If one of the jets in the event has a mass $m^{\text{jet}} > 60$ GeV, it is combined with the jet closest to it in ΔR . The reconstructed object is the hadronic top quark candidate. The leptonic top quark will be reconstructed by combining the reconstructed leptonic W boson candidate with the jet closest to it in ΔR . This reconstruction procedure is needed to account for the partially merged jets.

The invariant mass of the $t\bar{t}$ system is reconstructed in the resolved approach by adding the four-vectors of the lepton, the neutrino and the four or three leading p_T jets. The $m_{t\bar{t}}$ distributions for a variety of resonance masses at the reconstruction level are shown in Figure 6.8a. They are obtained from simulated signal events that pass the resolved

6 Top-quark reconstruction in $t\bar{t}$ resonances searches

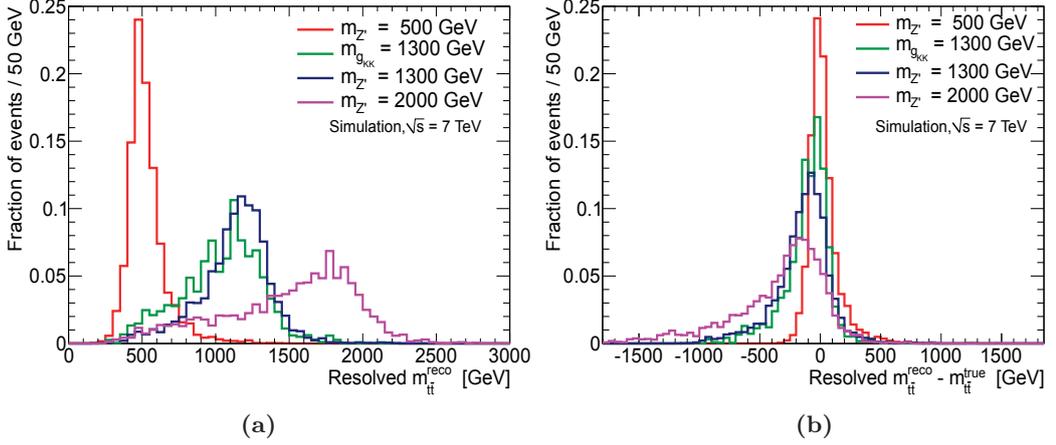


Figure 6.8: (a) Reconstructed invariant mass of the $t\bar{t}$ system $m_{t\bar{t}}^{\text{reco}}$, and (b) difference between the reconstructed and the true $m_{t\bar{t}}^{\text{true}}$ invariant mass. The distributions are obtained for the resolved selection, at a variety of simulated Z' masses ($m_{Z'}$). The broad g_{KK} resonance at $m_{g_{KK}} = 1.3$ TeV is also shown for comparison. All distributions are normalized to the unit area.

selection. The chosen signal masses vary from $m = 0.5$ TeV up to $m = 2$ TeV. For $m = 1.3$ TeV the KK-gluon distribution is notably wider than the Z' distribution. The difference is quantified by the full width at half maximum (FWHM) of the distributions. Table 6.7 shows the FWHM values for a set of resonance masses. The difference is not larger than 100 GeV for resonances with mass up to 1 TeV, although it increases for higher masses. Thus, the broader nature of the g_{KK} resonance in comparison to the leptophobic Z' boson is validated.

The mass resolution shown in Figure 6.8b is determined by subtracting the $m_{t\bar{t}}$ distributions at the truth level from those at the reconstruction level. Under ideal conditions the resolution would behave similar to a δ -function. Since this is not the case, the broadening of the distributions reveals certain effects taking place during the reconstruction. Firstly, it is possible that the solution chosen for the neutrino $|p_z|$ is not the correct. Secondly, additional jets from radiated gluons are not taken into account during the $t\bar{t}$ reconstruction, thus the invariant mass could be underestimated. As the production

Mass [TeV]	FWHM		$m_{t\bar{t}}^{\text{reco}}$	
	Resolved	g_{KK}	Z'	g_{KK}
0.7	150	250	100	150
1.0	200	250	200	250
1.3	300	350	250	300
1.6	350	750	300	700

Table 6.7: The full width at half maximum FWHM for different resonance masses. The FWHM values are given in GeV.

of this radiation increases with the p_T of the top quark, the zero-divergent behavior is most evident for higher masses. The extended tail in the lower half of the high-mass distributions are mainly due to the lack of reconstruction of some decay products. In such cases, the invariant mass of the $t\bar{t}$ system gets smaller. The limited resolution of the detector also affects the mass resolution. Namely, each sub-detector has a finite resolution for the measurements it takes, resulting in a finite resolution on the variables associated with the reconstructed objects.

6.5.2 The $t\bar{t}$ system in the boosted approach

The reconstruction of the $t\bar{t}$ system in the boosted approach is simpler. The hadronic top quark is the tagged top jet, and the leptonic top quark is obtained by adding the four-vectors of the lepton, the neutrino p_z and the $R = 0.4$ jet from the leptonic decay. Figure 6.9a shows the mass distributions for the hadronic top jet after the event selection. These are obtained for a signal sample of g_{KK} with $m_{g_{KK}} = 1.3$ TeV and for the SM $t\bar{t}$ background. The peak around the top quark mass is visible in the signal distribution. However, there is no peak in the distribution from the SM $t\bar{t}$ background. This is because the hadronic top is reconstructed using a fat jet, which is required having $p_T > 250$ GeV. Only top quarks with $p_T \gtrsim 350$ GeV are expected to be boosted enough, such that their decay products will be enclosed within the $R = 1.0$ cone. Therefore, the SM $t\bar{t}$ sample of tagged top jets will actually include a significant fraction that do not contain all the top quark decay products. The peak in the SM $t\bar{t}$ background distribution is apparent if the threshold is increased to $p_T > 350$ GeV, as noted in Figure 6.9b.

Figure 6.10 shows the mass distribution of the leptonic top quark. The peak around the top quark mass can be seen in both signal and SM $t\bar{t}$ distributions. Unlike the

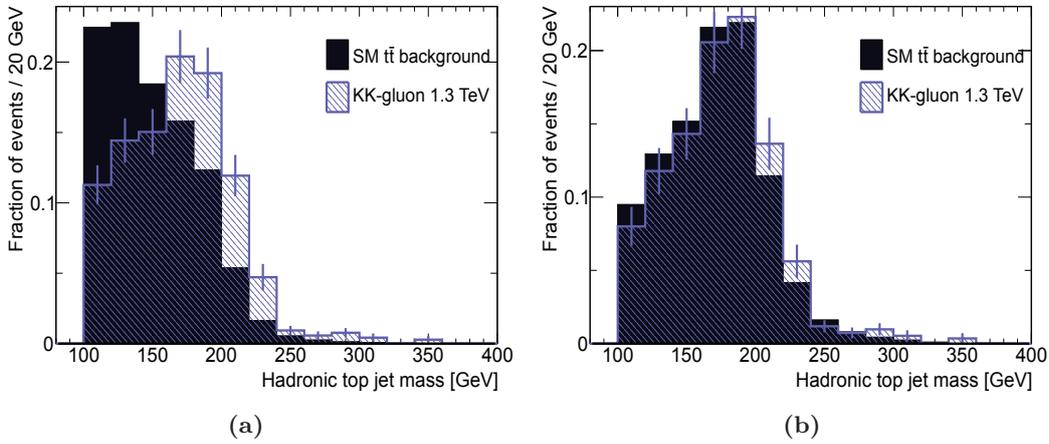


Figure 6.9: Mass distribution of the hadronic top quark in the boosted approach of the $t\bar{t}$ reconstruction after the top-tagging applied on signal and SM $t\bar{t}$ background events, for jets having (a) $p_T > 250$ GeV and (b) $p_T > 350$ GeV. The distributions are normalized to the unit area.

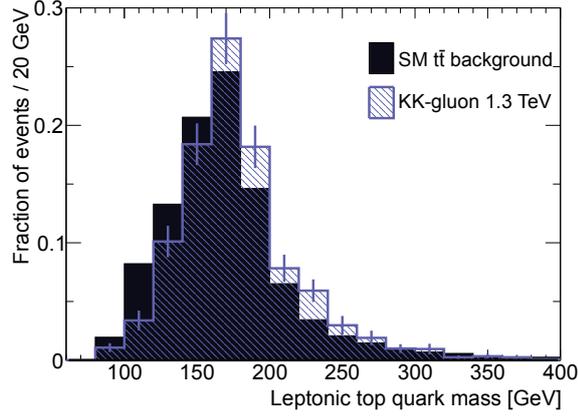


Figure 6.10: Leptonic top quark mass distributions in the boosted approach for signal and SM $t\bar{t}$ events. The signal corresponds to a g_{KK} with $m_{g_{KK}} = 1.3$ TeV. The distributions are normalized to the unit area.

hadronic top, the reconstruction of the leptonically decaying top quark includes all its decay products, regardless the p_T of the reconstructed top quark. If the decay is too soft that the $0.4 < \Delta R$ (jet-lepton) < 1.5 condition is not satisfied, the reconstruction does not proceed and the event is simply discarded.

The invariant mass of the $t\bar{t}$ system is reconstructed in the boosted approach by adding the four-vectors of the leptonic and hadronic top quarks. Figure 6.11 illustrates the rationale for the choice of the $p_T > 250$ GeV criterium, imposed on the hadronic top jet. The signal sample corresponds to g_{KK} with $m_{g_{KK}} = 1$ TeV. At this mass, the boosted reconstruction becomes suitable in relation to the resolved one. The signal and SM $t\bar{t}$ background distributions are well separated if the p_T threshold on the hadronic top jet is 250 GeV. This facilitates the identification of the resonances at the TeV scale. However if the p_T threshold is raised to 350 GeV, the peaks in the distributions are

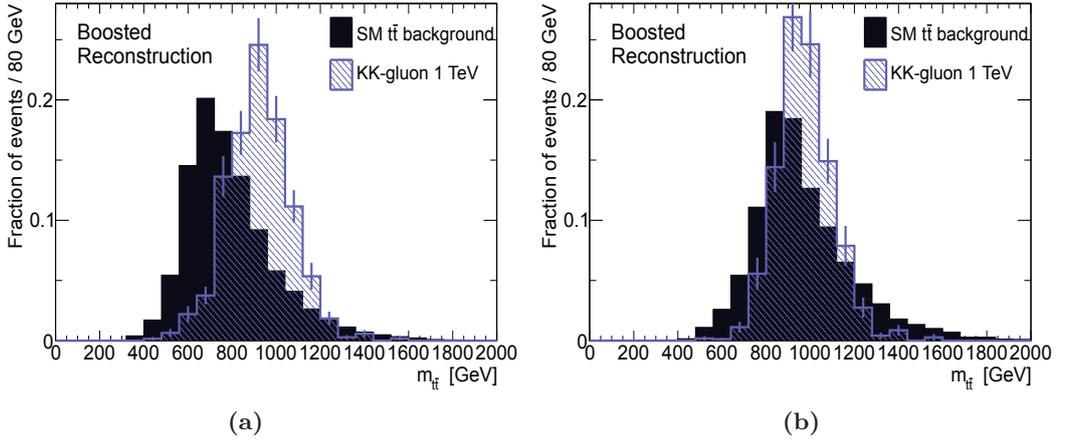


Figure 6.11: Invariant mass distribution of the $t\bar{t}$ system in the boosted approach for signal and SM $t\bar{t}$ background events. Jets are required having (a) $p_T > 250$ GeV and (b) $p_T > 350$ GeV.

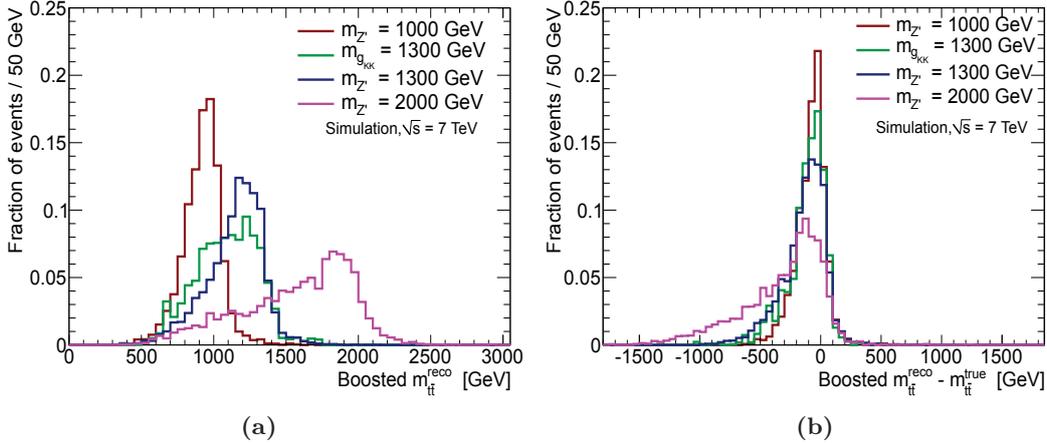


Figure 6.12: (a) $t\bar{t}$ invariant mass and (b) mass resolution in the boosted approach of the $t\bar{t}$ reconstruction, at different simulated masses. All distributions are normalized to the unit area.

almost coincident, restricting such identification.

Figure 6.12 shows the distributions of the reconstructed invariant mass and mass resolution in the boosted approach. Some FWHM values obtained from the boosted distributions also listed in Table 6.7. At high resonance masses, the residuals in the boosted distributions are narrower than those from the resolved distributions. Thus, the boosted reconstruction is most suitable for resonances at high masses. The effects which broaden the resolved distributions also affect the boosted reconstruction, generating the residual and tails in the distributions.

6.6 Data-driven estimation of background processes

The Monte Carlo (MC) simulation of the multi-jet background is difficult due to the production of jets that can be mis-reconstructed as leptons (fakes). The W +jets background is easy to simulate but it can introduce large systematic uncertainties. Therefore, these background processes are estimated from data, where possible. An entirely data-driven estimation is used for the QCD multi-jet background. For the W +jets background, only the normalization is estimated from data. The estimation mechanisms, together with the associated systematic uncertainties are described in this section.

6.6.1 W +jets background

The leptonic decays of W bosons and top quarks have the same final-state signature in the detector. At the LHC, W +jets processes are produced at rates large enough to hinder the identification of signal. The contribution is estimated using the ALPGEN generator. ALPGEN provides an appropriate description of the final-state kinematics in

6 Top-quark reconstruction in $t\bar{t}$ resonances searches

low H_T regions⁷. However, its description in the signal regions ($H_T > 700$ GeV) is not accurately known [157]. Hence, scale factors (SFs) are derived from data to normalize the ALPGEN yield in regions where the signal can be observed. For the resolved approach, default normalization SFs as defined in the ATLAS Top Working group are used [150]. For the boosted approach, the correction factors are carefully computed in the relevant kinematic regions of the boosted topology. Both approaches use a common procedure called “charge asymmetry method”. The method exploits the asymmetry of the cross-section distributions for positive and negative charged leptons from W boson decays. Since u quarks dominate over d quarks inside the proton, the rate of W^+ + jets is higher than that of W^- + jets production. The ratio $r_{MC} = \sigma(pp \rightarrow W^+)/\sigma(pp \rightarrow W^-)$ is the rate of the yields from W^+ and W^- production derived from MC simulation, and is well understood in theory. This is predicted within a few percent at LHC energies and its main theoretical uncertainty is associated with the PDFs [36, 72].

The formula used to extract the W +jets background is

$$N_{W^+} + N_{W^-} = \left(\frac{r_{MC} + 1}{r_{MC} - 1} \right) (D_+ - D_-), \quad (6.3)$$

where $D_{+(-)}$ is the event yield in data after the $t\bar{t}$ selection with a lepton charged positively (negatively). $N_{W^+} + N_{W^-}$ is the total yield for W +jets production derived from data. Except the SM $t\bar{t}$ production which is charge symmetric, the non- W +jets contributions in data have a degree of charge asymmetry. These are derived from MC simulation and subtracted from Equation 6.3. The SFs are defined as the ratio between the number of W +jets events estimated with the charge asymmetry method and the MC prediction. They are calculated separately in the electron and in the muon channels.

6.6.1.1 W +jets scale factors for the resolved approach

Two sets of SFs are calculated for the resolved analysis: overall normalization SFs and the flavor component SFs. For the former, the difference $N_{W^+} + N_{W^-}$ and the ratio r_{MC} are extracted as a function of the final jet and b -jet multiplicities. The overall normalization SFs obtained in the inclusive four-jet bin with at least one b -tag are 0.906 and 0.814 for the electron and muon channels, respectively. The overall normalization uncertainty on the W +jets background is 48% [158].

The flavor-component SFs are derived to match each flavor component of the ALPGEN yield, i.e. $W + b\bar{b}$ +jets, $W + c\bar{c}$ +jets, $W + c$ +jets and W +light jets, to the yield obtained in data. The data yield is estimated from a “tag-counting”, together with the b -tagging efficiency described in Section 4.5.2 [156, 158]. A system of equations is solved in order to determine the flavor fractions in n -jets events, before and after b -tagging. The ratio of events containing two b -quarks to events with two c -quarks is taken from MC simulation; it is used to reduce the number of unknown variables and is applicable in the 2-jets bin only. For each jet bin, the sum of all flavor fractions is constrained to unity [158]. Once the flavor fraction is determined in the two-jets bin, it is extrapolated to a higher

⁷ H_T is the scalar sum of the p_T of all partons and of the lepton and neutrino from the W boson decay.

multiplicity bin. The final SFs are 1.63 to scale the $W + bb$ and $W + cc$ components, 1.11 to the $W + c$ fraction and the light component is scaled by a factor of 0.83. The latter is estimated in such a way that the normalization in the 2-jets bin is unaffected. The systematic uncertainty due to the application of these SFs is obtained by varying the flavor SFs within their statistical uncertainties [156, 158]. They are 13% for $W + bb$ and $W + cc$ and 9% for $W + c$. An additional 25% is added in quadrature to account for the extrapolation to higher multiplicity. Uncertainties related to jet energy scale, b -tagging efficiency, E_T^{miss} , etc, are considered besides the W +jets flavor-composition uncertainties. These are of course correlated with the object-uncertainties and must be properly included. This is achieved by re-calculating the flavor-composition SFs for each object-uncertainty scenario of the resolved selection.

6.6.1.2 W +jets scale factors for the boosted approach

The total data yield for W +jets production is significantly reduced after the boosted selection, affecting the evaluation of the SFs. To increase the statistics for a reliable W +jets normalization estimate, the boosted selection cuts are loosened. Specifically, the top-tagging criteria related to the jet mass and $\sqrt{d_{12}}$ are not applied. Moreover, the p_T threshold on the hadronic top jet is lowered from 250 GeV to 150 GeV. Once the selection is done, the data yields $D_{+(-)}$, and the ratio r_{MC} are inserted into Equation 6.3. This gives the data-driven W +jets yield for the “normalization” region (NR) of the boosted selection. To obtain the final normalization SFs, this yield is divided by the total yield for W +jets production estimated from MC simulation. The resulting SFs are 0.75 and 0.79 for the e +jets and μ +jets channel, respectively. These values are applied to normalize the W +jets yield in the signal region, despite the derivation is carried out in the NR. The validation of this procedure is detailed in [58]. The contributions of non- W +jets backgrounds which are subtracted are: single top with a degree of asymmetry $\approx 40\%$, diboson with $\approx 10\%$ and QCD multi-jet with $\approx 5\%$.

Several sources of systematic uncertainties are associated with the derivation of the SFs. A systematic uncertainty of 5.2% (2.2%) in the electron (muon) channel is assigned to account for all background subtractions described above. It is derived by re-calculating the SFs after assuming a conservative 100% uncertainty to each background. An additional uncertainty of 5.2% in both channels is considered due to the choice of the PDFs. The values of the jet p_T thresholds are effected by uncertainties in the jet energy scale. The thresholds are varied according to these uncertainties and the SFs are re-calculated. The resulting uncertainty is found to be 7.9% (10.1%) in the electron (muon) channel. Finally, the systematic uncertainty associated with the ALPGEN modeling of the W +jets processes has an average effect of 7%. The above uncertainties are assumed as uncorrelated. Therefore, they are added in quadrature to obtain a first part of the uncertainty associated with the NR SFs. The second part accounts for the applicability of SFs in the signal region. The uncertainty is estimated by varying the fat jet p_T and top-tagging variables within the levels of their uncertainties. The SFs are found to vary by $\pm 0.27(\pm 0.18)$ in the electron (muon) channel. The first and second parts of the

uncertainty are assumed to be uncorrelated and thus added in quadrature. The final estimate of the SFs in the NR is $0.75 \pm 0.06 \pm 0.11$ (17% total uncertainty) in the e +jets channel, and $0.79 \pm 0.05 \pm 0.11$ (15% total uncertainty) in the μ +jets channel.

6.6.2 QCD multi-jet background

Multi-jet events from QCD production processes can fake leptons resulting from non-prompt sources, e.g. heavy-flavor quark decays. Even though they are not produced from the W -boson decay, non-prompt leptons can occasionally enter the selected sample, thereby constituting a background source. Since the production cross-section for the multi-jet events is high⁸, but the probability for jets to fake leptons is low, samples with very large statistics are required for a proper modeling of this background. Data-driven methods are used instead to estimate the multi-jet contribution to the total background.

6.6.2.1 Multi-jet contribution to the resolved selection using the jet-electron method

Data events are used in the jet-electron method to model jets mis-identified as electrons. This is achieved by mimicking the resolved selection, but instead of selecting one electron, one jet with certain properties is required. Jet candidates must have a high fraction of electromagnetic energy (f_{EM}). Events containing these jets have to be triggered by a single-jet trigger with a threshold $0.8 < f_{EM} < 0.95$. To avoid converted photons, jets are required having at least four tracks [159]. A veto on good electrons of medium quality is applied. Events containing exactly one of the surviving jets are subject for the further selection. However, the b -tagging requirement is removed in order to prevent statistical fluctuations [156]. The remaining events form the jet-electron template, used to model the kinematic distributions of the multi-jet processes. To determine the normalization of the multi-jets background, the jet-electron template together with the simulated SM $t\bar{t}$, single top, W +jets and Z +jets samples is fitted to the E_T^{miss} distribution obtained in data⁹. The fit is performed in a side-band region, which is defined by the events passing the resolved selection cuts, excluding those applied on the E_T^{miss} . The simulated samples mentioned above are allowed to vary in each bin of the E_T^{miss} distribution according to a Gaussian distribution. The Gaussian is centered at the bin height, with 10% RMS given by the modeling uncertainties.

The multi-jets contribution is estimated in the muon channel using the same templates as derived for the electron channel. The strategy relies on the fact that in a tagged analysis, the QCD background is dominated by events with true leptons from heavy flavor decays [156].

In general, normalization uncertainties modify the total number of events that make up a distribution, and shape uncertainties modify its shape. The systematic uncertainty on

⁸ 10^6 times larger than the cross-section for SM $t\bar{t}$ production

⁹The reason why the E_T^{miss} distribution is used to perform the fit is that the E_T^{miss} shapes in multi-jet and signal samples are different enough so that, fitting the multi-jet contribution to part of the signal region will not mask a potential signal [156].

the multi-jet background is separated out into the two components. The normalization uncertainty is derived by comparing the default E_T^{miss} fit above, to a similar fit using the $m_T(W)$ distribution. To estimate the impact of pile-up on the determination of the multi-jet background, the fit is carried out under two different pile-up conditions: the primary vertex multiplicity is required to be above and below six vertices. The found difference of 17% (48%) in the electron (muon) channel is taken as the normalization uncertainty due to pile-up. The same procedure is used to estimate the shape uncertainty affecting the $m_{t\bar{t}}$ distribution, due to different pile-up conditions. In addition, the dependence of the multi-jet estimate on the scalar sum of all momenta in the event is tested. For the different pile-up conditions, two samples with roughly the same number of events are separated. In the electron channel the separation is set at $H_T = 350$ GeV and in the muon channel at $H_T = 310$ GeV. After all samples are normalized to unity, the differences in the shape are summed in quadrature.

6.6.2.2 Multi-jet contribution to the boosted selection using the matrix method

The nominal contribution of the multi-jet background is estimated in the boosted approach using the matrix method. The strategy is based on a counting of events that fulfill two sets of selection criteria. The first corresponds to the lepton selection (tight selection) outlined in Section 6.4. To enhance the fraction of multi-jet events in the data sample, a looser selection is defined by removing some of the lepton requirements. In the muon channel the track-based and the calorimeter-based muon isolation cuts are dropped. Thus, muons are allowed to overlap with jets. For the loose electron selection, the `tight` electron quality is replaced by the `medium` one, with the additional requirement of a hit in the innermost layer of the pixel detector. The electron isolation requirement is also loosened: the value of `EtCone20` is required to be less than 6 GeV instead of 4 GeV [159]. Leptons accepted after the loose selection but failing the tight requirements are considered as fake leptons¹⁰. The loose selection is implemented in data, together with the remaining criteria of the boosted selection. In each channel, the number of events which contain a lepton satisfying the loose criteria above is

$$N_L = N_{\text{prompt}} + N_{\text{fake}} . \quad (6.4)$$

Here, N_{fake} is the number of events containing a fake lepton, and N_{prompt} otherwise. Among them, the number of events containing a `tight` lepton is given by:

$$N_T = \varepsilon \times N_{\text{prompt}} + f \times N_{\text{fake}} , \quad (6.5)$$

where the efficiency ε (f) is the probability that a prompt (fake) lepton passing the loose criteria, will also pass the tight selection. These probabilities are derived from data and will be discussed later. Solving the equations for N_{prompt} and N_{fake} , the QCD

¹⁰Fake leptons arise from semi-leptonic b -quark jet decays, long-lived weakly decaying states e.g. π^\pm , mis-reconstruction of π^0 showers, etc.

6 Top-quark reconstruction in $t\bar{t}$ resonances searches

contribution to the signal region is given by:

$$f \times N_{\text{fake}} = \left[\frac{\varepsilon - 1}{\varepsilon - f} \right] N_T + \left[\frac{\varepsilon f}{\varepsilon - f} \right] (N_L - N_T). \quad (6.6)$$

Since **loose** leptons from multi-jet processes have a small probability to pass the tight selection, f must be estimated using a control sample CR0, enriched of multi-jet events. CR0 region is defined similarly to the signal region of the boosted configuration. Therefore, the ΔR between the leptonic and the hadronic jets, and some of the kinematic requirements on the final-state objects are retained. However, the p_T cut on the hadronic top jet is lowered from 250 GeV to 150 GeV, and only events with a **loose** lepton are accepted. To suppress contributions from the SM $t\bar{t}$ background, the top-tagging conditions are reversed: the hadronic jet is required with $m_j < 100$ GeV and $\sqrt{d_{12}} < 40$ GeV. The same logic is used to eliminate W +jets events, but the E_T^{miss} and $m_T(W)$ cuts are instead reversed: $E_T^{\text{miss}} < 35$ GeV and $m_T(W) < 25$ GeV. Finally, the Z +jets contribution is suppressed by vetoing any event containing a Z boson candidate. To identify this candidate, the invariant mass of the system formed by two **loose** leptons is expected to be close to the mass of the Z boson.

All the CR0 conditions aim to collect data events with **loose** leptons. Some of these leptons do not pass the tight criteria, whereby those events are dominated by multi-jet processes, as desired. Nevertheless, events whose leptons also pass the tight selection will have a fraction stemming from prompt sources. This fraction is estimated from the MC simulation to be subtracted from the multi-jet background. The final values calculated for f in the electron and muon channels are 0.066 ± 0.010 and 0.00395 ± 0.00161 , respectively. The associated systematic uncertainties are estimated by evaluating the precision of the MC background samples subtracted from CR0. This is done by defining several control regions parallel to CR0. These are listed below:

1. **Top-CR:** The thresholds on the E_T^{miss} and the $m_T(W)$ are reversed back to be the same as in the signal region. Additionally, the b -jet requirement defined in the resolved selection is required. The remaining CR0 criteria are retained.
2. **W +jets-CR:** Equivalent to the top-CR above, but the b -jet is now vetoed.
3. **Z +jets CR:** Equivalent to the CR0 but the Z +jets veto is inverted.

An overall 20% uncertainty is assumed on all the simulated subtracted samples.

Uncertainties associated with the variables that define CR0 are also considered. The top-tagging thresholds, the E_T^{miss} and the $m_T(W)$ cuts are varied exclusively, and for each variation f is recalculated. For these variables, the difference of the updated f with respect to the CR0 f is taken as the uncertainty of the cut. The final systematic uncertainty is then given by the combination of all the differences.

The efficiency ε is determined using the tag-and-probe method in $Z \rightarrow \ell\ell$ data events. Each event must have exactly two **loose** leptons of the same flavor, but at least one of them must also be **tight**. The invariant mass of the two leptons is required to be

between 86 GeV and 96 GeV, thus the sample is ensured to be dominated by $Z \rightarrow \ell\ell$ production. The efficiency is defined as

$$\varepsilon = \frac{1}{\frac{N_{TA}/2}{N_{TT}} + 1}, \quad (6.7)$$

where N_{TA} is the number of events with one **tight** (T) lepton and one anti-**tight** (A) lepton, and N_{TT} is the number of events in which both leptons are **tight**. The value obtained is 0.853 (0.969) in the electron (muon) channel, with an associated uncertainty less than 0.3%.

The total yields (N_{fake}) due to multi-jet production processes are calculated from Equation 6.6, together with f and ε . The final values are $49.1 \pm 2.3(\text{stat.}) \pm 7.1(\text{sys.})$ in the electron channel and $29.7 \pm 0.3(\text{stat.}) \pm 7.9(\text{sys.})$ in the muon channel. The statistical uncertainty in these values is estimated by propagating the statistical fluctuations on the number of events with a **tight** lepton, and the number of events with a **loose**, but not **tight** lepton. The total systematic uncertainty is given by the propagation of the f and ε uncertainties. Both vary the shape of the $m_{t\bar{t}}$ distribution. In addition, a total uncertainty of 50% is taken as the overall normalization uncertainty affecting the yield of the multi-jet background.

6.7 Systematic uncertainties

Sources of systematic uncertainty affecting the reconstructed $m_{t\bar{t}}$ distribution can be divided into two categories: those affecting the physics objects, which are related to the MC modeling and limited knowledge of the detector response, and those associated with the MC simulation of the physics processes. The majority of the uncertainties alter the normalization and the shape of the $m_{t\bar{t}}$ distribution. All are discussed in this section. Since this analysis incorporates two scenarios to reconstruct the $t\bar{t}$ system, the estimation of the systematic uncertainties is carried out separately in the resolved and boosted selections. Unless otherwise specified, the estimation method is common for both approaches.

6.7.1 Event level effects

Luminosity: Normalization-changing uncertainty of 3.7% applied as a constant shift to each MC-based background [91].

PDF uncertainty: Systematic uncertainties due to the use of different PDF sets in MC simulation are estimated. To determine this effect, the $m_{t\bar{t}}$ distribution obtained using the original PDF set, is compared to that obtained with other sets. In the resolved approach the NLO CTEQ66 [160] and MSTW2008-NLO68CL [36] sets are compared. In the boosted case the NNPDF20 [161] set is in addition compared. The MC truth information about the partons involved in the hard interaction is used to re-weight the events. The probability for an event with a particular kinematic configuration (e.g. the momentum transfer scale Q^2 , or the

fraction of the hadron's momentum introduced in Section 2.2) to be produced, is evaluated and re-weighted according to the new PDF set within 1σ variation of the original PDF set. The recommended methods to extract and combine the PDF uncertainties in ATLAS are described in detail in [162].

The total PDF uncertainty is absorbed by a total normalization uncertainty described in Section 6.7.4. The $t\bar{t}$ background yield is affected by 1.7% (3%) after the resolved (boosted) selection.

Since the W +jets background is normalized by data-driven SFs, the PDF uncertainty must be assigned only to the shape of the distribution given by ALPGEN. To ensure that the yield remains constant, the PDF variations are applied to the $m_{t\bar{t}}$ distribution, which is then normalized to the nominal yield.

6.7.2 Systematic uncertainties related to the physics objects

Systematic uncertainties associated with physics objects are evaluated varying the nominal value by ± 1 standard deviation (i.e. up/down) from the expected value. The whole analysis including the event selection and the $m_{t\bar{t}}$ reconstruction is re-run on the varied signal and background samples.

Lepton-related and E_T^{miss} uncertainties are estimated in the resolved and boosted selections using the same methods. Although the reconstruction of the $R = 0.4$ and $R = 1.0$ jets is based on the same anti- k_T technique, their calibration scheme differs. Therefore, the estimation of the jet uncertainties varies between the both approaches. Regardless the radius-parameter, uncertainties associated with jets have the largest impact on the analysis. b -jets are required by the resolved selection, whereby their uncertainties are evaluated in the resolved analysis only. Leptons and jets have uncertainties associated which change their energy. Any systematic scaling or smearing in the energy must be propagated accordingly to the E_T^{miss} measurement. This implies to re-calculate the E_T^{miss} after applying the systematic variation.

The estimation of these systematic uncertainties are presented in this section. Firstly, the uncertainties associated with leptons are described. The specific methods to derive the uncertainties for each jet-size are also discussed. Finally, the uncertainties associated with b -jets, the E_T^{miss} and the LAr failure are introduced.

6.7.2.1 Electron-related uncertainties

The electron reconstruction, identification and isolation systematic uncertainties are the uncertainties on their corresponding scale factors. Since they are derived in E_T and η_{cluster} bins, the systematic effects change the normalization and the shape of the $m_{t\bar{t}}$ distribution. The uncertainties related to the energy scale, resolution and trigger also modify the normalization and shape of the distribution. All the procedures and their uncertainties were presented individually in Section 4.2. A total impact of about 1.3% (1.2%) on the background distributions is found after running the resolved (boosted) selection. This uncertainty is obtained from the combination of all systematic uncertainties mentioned above.

6.7.2.2 Muon-related uncertainties

Muon reconstruction, identification and trigger SFs are applied before the muon selection on event-by-event basis. For the resolved selection their uncertainties have impacts of 1%, 1.2% and 1.9% respectively. The effects for the boosted selection are larger with about 2.6%, 2.2% and 2.1%. In both selections, the normalization and shape of the $m_{t\bar{t}}$ distribution are modified. Since the trigger matching requirement is not applied in simulation, an additional uncertainty of 1.5% (1.8%) is applied in the resolved (boosted) uncertainty estimation. This uncertainty is added in quadrature to the systematic effect associated with the muon trigger. All muon-related uncertainties are combined to obtain a total effect of 1.9% (2.1%) in the resolved (boosted) selection. Systematic uncertainties associated with the muon momentum scaling and resolution have a minor effect in both approaches.

6.7.2.3 Jet-related uncertainties: anti- k_T $R = 0.4$ jets

Small-radius jets are reconstructed from topo-clusters and calibrated to the jet energy scale, as described in Section 4.4.1. The EM+JES calibration employs full-detector simulation tests and data-MC comparisons to derive the calibration constants. The components of the JES uncertainty associated with the EM+JES calibration are discussed in this section. The resolution of the jet energy can be mis-estimated in the simulation, when compared to the resolution measured in data. The systematic effect due to the jet energy resolution (JER) is also outlined here.

Jet energy scale: A first component of the JES uncertainty is associated with the general calorimeter response to hadrons. The uncertainty is obtained from single-pion test-beam measurements and from single particle-response estimation. This JES component is parametrized as a function of the jet η and p_T , and varies between 1.5% and 4% for high and low p_T jets, respectively. An additional component is given by the mis-modeling of the signal-to-noise ratio in the calorimeter cells. Discrepancies between the simulated noise and the real noise in data can alter the cluster shape and produce fake topo-clusters. The effect is studied by reconstructing topo-clusters and thereafter jets in MC, using the noise measured from data. This component of the uncertainty, in addition to the uncertainty in the detector material description is below 3% [163].

Systematic effects due to the fragmentation and the underlying event modeling are obtained by comparing the nominal PYTHIA sample (used to derive the calibration constants) to ALPGEN+HERWIG + JIMMY. The IFS/FSR effects are estimated by varying the parameters in PYTHIA responsible for production of less or more parton shower [163]. The fragmentation and showering modeling depends on the jet flavor and has an influence on the calorimeter response. An additional uncertainty due to different calorimeter response to gluon-initiated and light-quark initiated jets (light-jets) is estimated. This is performed by varying the jet flavor with respect to the nominal MC sample. It has been found that for low p_T jets, light-quark jets

have on average, 5.5% higher response than gluon-jets, whereas this is about 2% in high p_T regions [164].

Since the calorimeter response to jets is η -dependent, an uncertainty due to the relative calibration of the jet energy scale in the end-cap and forward regions is also propagated. This component of the uncertainty is determined by exploiting the p_T balance of a central and a forward jets in di-jet events [163].

The realistic scenario for $t\bar{t}$ topologies is characterized by multi-jet events. A source of systematic uncertainty arises in events containing close-by jets. This is because the calorimeter response might be degraded due to overlapping energy deposits. The uncertainty is estimated by comparing calorimeter and track jets in MC and data events, parametrized in terms of the jet p_T and the ΔR between the jets. It is found to vary between 1.7% and 2.8% [165].

A systematic uncertainty accounting for the calorimeter response to heavy-quark initiated jets is in addition considered. The estimation of this effect is based on MC studies of the calorimeter response to b -jets, parametrized as a function of the b -jet p_T . The resulting values vary as shown in Table 6.8.

p_T range [GeV]	b -JES uncertainty [%]
20 — 40	2.5
40 — 80	2.0
80 — 210	1.7
210 — 600	1.1
> 600	0.76

Table 6.8: Jet energy scale uncertainty due to different calorimeter response to b -quark initiated jets [163].

The last component of the JES uncertainty is due to soft interactions in the same bunch crossing, i.e. in-time pile-up. Since pile-up contributions affect the energy of the reconstructed jet, an E_T offset correction to account for additional energy deposits is applied [166]. The uncertainty on the offset correction is parametrized as a function of the jet η and p_T and is found within a range between $\approx 2\%$ for high p_T jets and $\approx 5\%$ for low p_T jets.

The total JES uncertainty is derived by considering all the sources described above. The JES components were estimated from data collected in 2010, except those due to pile-up and b -jet scaling which were derived from 2011 data and MC, respectively. The final fractional JES systematic uncertainty for the sources evaluated in 2010 is shown in Figure 6.13, together with the individual contributions. The uncertainty components derived from pile-up and b -quark initiated jets are treated as uncorrelated to it, and therefore are added in quadrature. Up and down templates for the JES uncertainty are produced by adding and subtracting the total JES calibration uncertainty from the E_T and p_T of all the $R = 0.4$ jets, before the resolved and boosted selections.

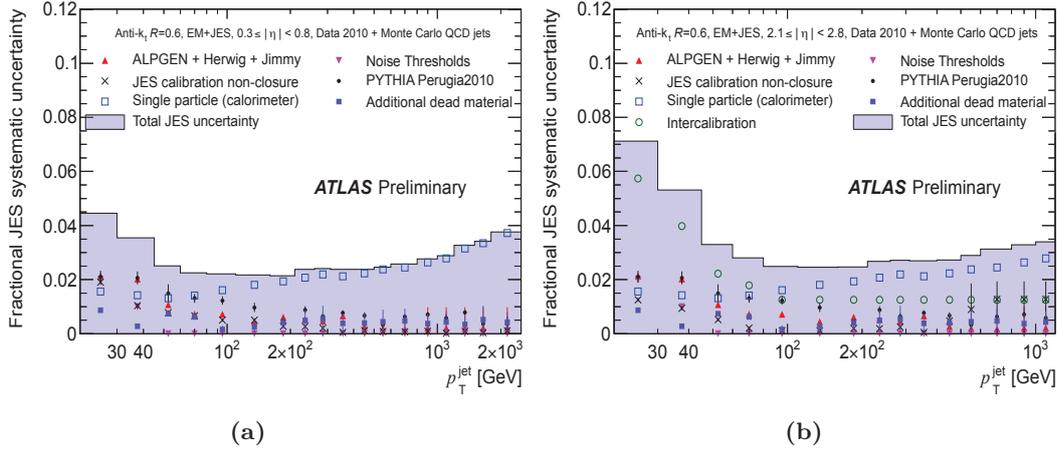


Figure 6.13: Fractional jet energy scale systematic uncertainty as a function of the jet p_T (p_T^{jet}), for $R = 0.4$ jets in (a) the calorimeter barrel and (b) the end-cap. The total uncertainty is shown as the solid light blue area [163].

Jet energy resolution: The systematic uncertainty associated with the jet energy resolution accounts for data-MC discrepancies in measurements of the jet energy. This was measured using the first 1 fb^{-1} of data, from which a data-MC agreement was estimated within $\approx 10\%$ for jets with $|\eta| < 2.8$ [129]. To propagate this effect on the $m_{t\bar{t}}$ distribution, all the $R=0.4$ jets with $p_T > 15$ GeV are smeared to match the JER measured in data, which is ≈ 25 (12) % for jets with $p_T = 20$ (100) GeV. The set of smearing y - and p_T -dependent SFs are provided centrally by the Top Working Group in ATLAS [167]. These are applied before the resolved and boosted selections. After the symmetrization of the resulting distributions, the up and down JER templates are obtained.

6.7.2.4 Jet-related uncertainties: anti- k_T $R = 1.0$ jets

Large-size jets are reconstructed from locally calibrated topo-clusters. During the energy calibration, the mass of each jet is corrected back to the jet energy scale as described in Section 4.4.2. Systematic uncertainties associated with the jet energy scale and the jet mass scale (JMS) of anti- k_T $R = 1.0$ jets are estimated for the boosted selection only. The estimation mechanisms take advantage of the fact that the ID and calorimeter have largely uncorrelated systematic effects. Comparisons of variables such as the mass and energy of the jet in the two sub-detectors, allows for some separation of the systematic uncertainties due to the physics modeling and due to the detector effects [168]. The strategy begins by matching track jets to calorimeter jets if they are within a distance of $\Delta R < 0.3$ of each other. For a substructure property X (e.g. p_T , mass, $\sqrt{d_{12}}$), the ratio between a calorimeter jet and its matched track jet is given by:

$$r^X = \frac{X_{\text{calorimeter-jet}}}{X_{\text{track-jet}}}. \quad (6.8)$$

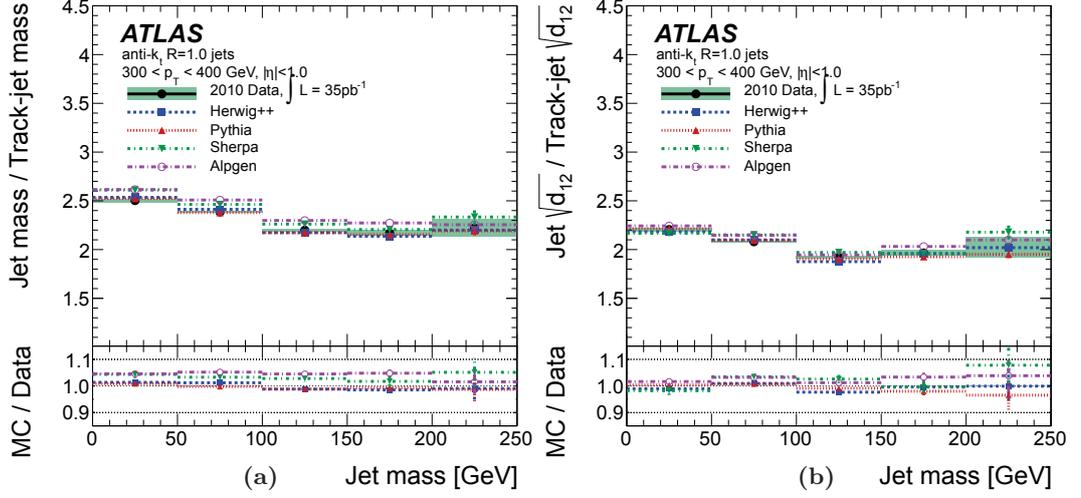


Figure 6.14: The ratio of (a) jet mass and (b) first k_t -splitting scale determined using the calorimeter to that determined using the ID tracks, as a function of the calorimeter jet mass for anti- k_T $R = 1.0$ jets. All jets are required having $300 < p_T < 400$ GeV [168].

Double ratios of r^X measured in data and MC simulation are defined in order to quantify the level of agreement:

$$\rho^{p_T} = \frac{r_{\text{data}}^{p_T}}{r_{\text{MC}}^{p_T}}, \quad \rho^m = \frac{r_{\text{data}}^m}{r_{\text{MC}}^m}, \quad \rho^{\sqrt{d_{12}}} = \frac{r_{\text{data}}^{\sqrt{d_{12}}}}{r_{\text{MC}}^{\sqrt{d_{12}}}}, \quad (6.9)$$

where the substructure variables of the top-tagging have been brought here. The MC generator used for these ratios is PYTHIA v6.423 [66], with the PERUGIA2010 PYTHIA tune [169]. It has been found to describe more accurately the substructure variables in ATLAS [170]. The substructure distributions of the variables $X_{\text{calorimeter-jet}}$ themselves are not necessarily expected to be correctly modeled by MC simulation. However, if the simulation correctly models the effect of the detector on these variables, the double ratios ρ^X , are expected to be consistent with unity. Thus, deviations from unity can represent e.g. that the jet calibration in MC simulation is different from data. Figure 6.14 shows the double ratios for jet mass and $\sqrt{d_{12}}$, using different generators.

The double ratio calculated using PYTHIA is within 1% of unity, meaning that the detector effects are well-modeled by the simulation. This variation is added in quadrature to the 3-5% uncertainty obtained from the ID measurements, in order to estimate the final uncertainties on the scale of energy, mass and $\sqrt{d_{12}}$ listed in Table 6.9. The MC generators HERWIG++ [171], SHERPA 1.2.3 [172] and ALPGEN 2.13 are used to estimate systematic uncertainties associated with the modeling of the fragmentation and hadronization processes. For each generator, the double ratios are calculated resulting in deviations from unity below 5% in all the cases.

Table 6.9 also displays systematic uncertainties associated with the jet energy (JER),

Uncertainty	200 - 300 GeV	300 - 400 GeV	400 - 500 GeV	500 - 600 GeV
JES	4.0%	5.2%	6.0%	3.9%
JMS	4.5%	4.5%	6.0%	6.0%
JER	20.0%	20.0%	20.0%	20.0%
JMR	20.0%	20.0%	20.0%	20.0%
$\sqrt{d_{12}}$ scale	4.4%	3.8%	6.0%	6.8%
$\sqrt{d_{12}}$ resolution	21.0%	22.0%	28.0%	31.0%

Table 6.9: Final systematic uncertainties associated with kinematic quantities for anti- k_T $R = 1.0$ jets. The percentage is expressed with respect to the nominal value [168].

jet mass (JMR) and $\sqrt{d_{12}}$ resolution. These are derived from MC simulation only, using simulated samples from the different generators listed above. Parameters in the samples are varied to emulate some of the uncertainty sources described in Section 6.7.2.3. In particular, samples including variations in the following properties have been generated:

- the detector material,
- the fragmentation and hadronic shower model,
- the underlying event model.

The maximum deviation across all the samples is 20%, with a 30% deviation when the ALPGEN samples are included.

Similar to $R = 0.4$ jets, the JER for $R = 1.0$ jets measured in data and MC simulation are assumed to vary within 10%. The systematic uncertainties due to the variations in the simulations, and the mis-matching between data and MC simulation, are added in quadrature to obtain the total uncertainty on the resolution. It has been found to be $\approx 20\%$.

Large-size jets are more susceptible to in-time pile-up than smaller-radius jets. The mass of the $R = 1.0$ jet is corrected to account for the additional energy deposits due to pile-up. The systematic uncertainty associated with this correction is applied to the JMS. Its estimation is carried out by varying the average number of pp interactions as $\langle \mu \rangle + 1$ and $\langle \mu \rangle - 1$ in the SM $t\bar{t}$ sample. The nominal value of $\langle \mu \rangle$ has been adjusted to match that in data, using the pile-up re-weighting procedure introduced in Section 3.6.2. The effect on the yield of the SM $t\bar{t}$ sample is found to be $\approx 5\%$, and is included by increasing the total JMS uncertainty by 5% [58].

6.7.2.5 b -tagging Efficiency

Jet p_T - and η -dependent SFs are applied on MC simulated events to correct to the b -tagging efficiency and mis-tag rates determined from data (see Section 4.5.2). The systematic uncertainty associated with the SFs has been found ranging from 8% to 16% for the tagging efficiency of b -quark initiated jets. For the mis-tagging rate of light-quark initiated jets it ranges between 23% up to 45%. The c -tagging efficiency is assumed to be fully correlated with the b -tagging efficiency.

6.7.2.6 Missing transverse momentum uncertainty

The E_T^{miss} is a function of the energy measured in the detector, that corresponds to the physics objects of each event. It is added to the energy of clusters that, after event reconstruction, are not associated with a physics object. The systematic uncertainties on the E_T^{miss} stem from different sources listed as follows:

- Uncertainties on the energy scale of the topo-clusters associated to the `CellOut` and `SoftJet` terms (see Section 4.6). These effects come from discrepancies between the processes occurring in the detector and the detector simulation. They have been found to be $\approx 13.2\%$ for `CellOut` and $\approx 10.5\%$ for `SoftJets`. The details of the estimation are developed in [139]. Since the source of these effects is the same, the uncertainties are assumed to be fully correlated, thus added linearly.
- Uncertainty due to additional energy deposits left in the calorimeters by pile-up events. This uncertainty is assumed as 100% correlated with the `CellOut` and `SoftJets` uncertainties. For the two terms, the x and y components of the E_T^{miss} are shifted by 10%, to account for the pile-up mis-modeling on the E_T^{miss} .

6.7.2.7 LAr failure

The systematic uncertainty due to the LAr problem (see Section 3.6.1) is propagated varying by 4 GeV the energy threshold for all jets falling in the hole vicinity. This systematic is normalization and shape changing uncertainty.

6.7.3 Systematic effects on the minor backgrounds

Minor sources of background are the simulated single top, Z +jets and diboson processes. These sources contribute to the total background yield with less than 20% after both selections. Since their systematic effects are entirely due to cross-section uncertainties, they change the normalization of the $m_{t\bar{t}}$ only. The systematic uncertainties associated with the theoretical prediction of the cross-sections are accounted by re-scaling the single top, Z +jets and diboson nominal backgrounds by $\pm 10\%$ [173], $\pm 48\%$ [174] and $\pm 5\%$ [174], respectively.

6.7.4 Systematic uncertainties affecting the SM $t\bar{t}$ background only

There are several sources of systematic uncertainties in the MC modeling of the SM $t\bar{t}$ background. As mentioned previously, the uncertainty due to the choice of the PDFs to simulate $t\bar{t}$ events is absorbed by an overall normalization uncertainty of ${}_{-9.6}^{+7.0}\%$ [175]. Other contributions arise from the uncertainty on the renormalization μ_R and factorization μ_F scales (Section 2.1.2 and Section 2.2). This effect is calculated by independently varying each scale between $0.5m_t$ and $2m_t$, with the constraint $0.5 < \mu_F/\mu_R < 2$. The envelope of the obtained cross-section is taken as the total uncertainty. The resulting impact on the SM $t\bar{t}$ yield is 6.6% (4.9%) after the resolved (boosted) selection. No uncertainty on the top quark mass is considered in this normalization uncertainty [176].

Additional systematics uncertainties affecting the SM $t\bar{t}$ background change the normalization and the shape of the $m_{t\bar{t}}$ distribution. They are described in the following.

Fragmentation and parton shower: A systematic uncertainty due to the use of fragmentation and parton shower models (provided by HERWIG and JIMMY) is estimated. Since the nominal generator of the hard process MC@NLO can only be interfaced with HERWIG/JIMMY, POWHEG, samples interfaced with PYTHIA are used to generate the up and down variations. These have to be scaled to the difference between MC@NLO and POWHEG when both interfaced with HERWIG/JIMMY. The variation in the $m_{t\bar{t}}$ distribution with respect to the nominal one is quoted as systematic uncertainty: 3.3% (3.4%) after the resolved (boosted) selection.

Monte Carlo generator: The systematic uncertainty due to the modeling of the hard process is estimated by comparing the $t\bar{t}$ mass spectra using the default MC@NLO generator, to the POWHEG generator. POWHEG is interfaced, like MC@NLO, with HERWIG/JIMMY. The resulting effects due to this systematic uncertainty are 0.5% and 2.8% in the resolved and boosted approaches, respectively.

ISR-FSR: The uncertainty due to the modeling of initial- and final-state radiation is estimated by using SM $t\bar{t}$ samples generated with ACERMC. ACERMC is interfaced with PYTHIA for the hadronization. The PYTHIA parameters that control the probability to produce ISR/FSR are varied. Specifically, those related to the scale of the hard-scatter and to the strong coupling constant [177]. The impact of this systematic uncertainty is found to be 5.9% and 6.4% on the resolved and boosted SM $t\bar{t}$ yields, respectively. For the resolved selection only, this uncertainty is also evaluated in signal samples; the resulting effect is about 5%.

6.7.5 Shape systematics affecting the W +jets background only

Normalization uncertainties on the data-derived W +jets SFs were given individually in Sections 6.6.1.1 and 6.6.1.2. As mentioned above, the shape of the reconstructed invariant mass of the $t\bar{t}$ system is determined by the ALPGEN generator. The systematic effect on the $m_{t\bar{t}}$ distribution shape is estimated by mimicking the variation of key parameters in ALPGEN. Specifically, variations of the relative contributions of the W +light-jets simulated sub-samples listed in Table 6.1 are considered. This method provides the correspondence between the parameter variations and the sub-samples cross-sections. The ALPGEN varied parameters are enumerated as follows [72]:

1. **iqopt:** Is a functional form of the factorization scale. The default form is $m_W^2 + \sum_{\text{partons}}(m^2 + p_T^2)$, where the sum is over all final state partons, including heavy quarks. The decay products of the W boson are excluded from the sum. The default form is varied up to $m_W^2 + p_{T_W}^2$ and down to m_W^2 .
2. **qfac:** Is a multiplicative factor applied to the factorization scale. The default value is 1 and it is varied to 0.5 and to 2.

3. **ktfact**: Is a multiplicative factor applied to the renormalization scale. The default and varied values are the same as for qfac.
4. **ptmin**: Is the minimum p_T of partons generated in the simulation of the matrix element (see Section 2.5.2). The p_T threshold is required so that the phase space region of jets generated by the matrix element does not overlap with those generated in the parton shower. The default value is 15 GeV, and it is varied up to 20 GeV and down to 10 GeV.

6.7.6 Impact of the systematic uncertainties on the analysis

The impact of the systematic uncertainties on the background and signal yields are listed in Table 6.10. The chosen signal is the Z' with a mass of $m_{Z'} = 1.3$ TeV. The shifts are given in percent with respect to the nominal values as obtained for the resolved and boosted selections individually. The impact of each systematic uncertainty is separated into three categories: inclusive, overlap and exclusive. The “inclusive” column indicates the impact estimated using all selected events of each approach. The “overlap” column corresponds to the impact of each uncertainty derived on the overlap channel, i.e. from those events that pass both the resolved and the boosted selections. Considering that these impacts can change accordingly to the selection, the uncertainties in the overlap channel are estimated for the resolved and boosted categories separately. The “exclusive” column in the Table contains the impact of each uncertainty, estimated using events that pass the resolved selection but not the boosted selection and vice-versa. These definitions are being used hereafter.

Even if a systematic e.g. b -tagging is not applied for one selection (in this case boosted), this systematic can cause event migration between the overlap and the exclusive categories, but the total number of events in the inclusive category remains constant. Thus, the \pm (up/down) variation in the overlap channel \mp (down/up) the variation in the exclusive overlap channel retrieves the zero variation of the inclusive boosted channel. This explains why the b -tagging uncertainty in the exclusive boosted sample is non-zero, even though no b -tagging condition is required in this channel. The impact of the systematic uncertainties treated in this way is specified in the Table with the \pm and \mp operators.

6.8 Observed data and Standard Model background prediction

The selections described in Section 6.4 yield a total of 22326 and 1837 data events in the inclusive channels of the resolved and boosted approaches, respectively. It is found to be in good agreement with the corresponding background prediction of 22263 ± 4186 and 1833 ± 234 events. The individual contribution from data and from each background expectation is displayed in Table 6.11. The signal yields for certain mass points are also summarized in Table 6.12. All yields are listed for the electron and muon channels separately, and for both combined. In addition, the event yields are given for the inclusive, overlap and exclusive categories, specified in Section 6.7.6 above. The total systematic uncertainties in the background yields are also listed.

Systematic source	Resolved selection yield impact [%]						Boosted selection yield impact [%]					
	Inclusive		Overlap		Exclusive		Inclusive		Overlap		Exclusive	
	Bkg.	Z'	Bkg.	Z'	Bkg.	Z'	Bkg.	Z'	Bkg.	Z'	Bkg.	Z'
Luminosity*	3.1	3.7	3.4	3.7	3.1	3.7	2.5	3.7	3.5	3.7	1.7	3.7
$t\bar{t}$ PDF uncertainty	1.7	--	2.9	--	0.4	--	3.0	--	3.0	--	1.1	--
Electron-related	1.3	1.3	1.3	1.3	1.3	1.3	1.2	1.3	1.3	1.3	1.2	1.3
Muon-related	1.9	2.0	2.1	2.1	1.9	1.8	2.1	2.2	2.1	2.1	1.9	2.2
Jet energy scaling	12.8	4.1	6.5	2.0	10.2	1.1	6.7	1.9	6.3	1.8	7.0	3.5
Jet energy resolution	2.8	1.7	12.5	9.7	3.5	17.3	4.6	4.0	5.5	2.8	3.7	6.9
b -tag efficiency	11.7	7.5	10.1	9.2	7.5	9.7	--	--	± 7.6	± 5.6	∓ 6.9	∓ 13.1
E_T^{miss} -related	1.1	0.2	0.1	0.3	1.2	0.1	0.7	0.3	0.1	0.3	1.2	0.2
LAr failure	0.5	0.6	0.4	0.5	0.4	0.1	0.5	0.5	0.4	0.5	0.3	0.4
W +jets-normalization*	5.6	--	3.1	--	5.7	--	4.3	--	0.6	--	7.7	--
W +jets bb/cc -normalization*	2.3	--	1.6	--	2.3	--	--	--	± 1.1	--	∓ 5.3	--
W +jets c -normalization*	0.4	--	0.1	--	0.4	--	--	--	± 2.7	--	∓ 6.3	--
Multi-jet normalization*	1.8	--	0.9	--	1.9	--	2.1	--	1.3	--	2.9	--
Z +jets-normalization*	0.4	--	0.3	--	0.4	--	2.0	--	0.3	--	3.5	--
Single-top normalization*	0.4	--	0.3	--	0.5	--	0.3	--	0.3	--	0.2	--
$t\bar{t}$ normalization*	6.6	--	7.2	--	6.6	--	4.9	--	7.2	--	2.9	--
$t\bar{t}$ Parton shower & frag.	3.3	--	8.2	--	2.1	--	3.4	--	8.5	--	3.3	--
$t\bar{t}$ MC generator	0.5	--	2.5	--	0.4	--	2.8	--	2.6	--	1.0	--
$t\bar{t}$ & signal ISR/FSR	5.9	4.5	7.3	7.7	8.0	7.7	6.4	--	7.5	± 7.7	3.0	∓ 7.5

Table 6.10: Systematic uncertainties and their impact on the total background yield and on the yield of the Z' signal sample with $m_{Z'} = 1.3$ TeV. The e +jets and μ +jets spectra are added. Certain systematic effects are not applicable (i.e. zero-variation) to all the samples, which is indicated with the bars (---). JES(JER) uncertainties in the boosted columns include the JES(JER), JMS(JMR) and $\sqrt{d_{12}}$ scale (resolution) uncertainties. Sources of uncertainty labelled with “*” affect the normalization of the $m_{t\bar{t}}$ distribution only. The non-labelled sources affect both its normalization and shape.

Type	Resolved selection			Boosted selection		
	$e+\text{jets}$	$\mu+\text{jets}$	Sum	$e+\text{jets}$	$\mu+\text{jets}$	Sum
SM $t\bar{t}$	7752 \pm 1675	9906 \pm 2066	17658 \pm 3741	513 \pm 97	615 \pm 108	1128 \pm 205
$W+\text{jets}$	1130 \pm 786	1452 \pm 1026	2582 \pm 1812	202 \pm 41	299 \pm 56	501 \pm 97
Multi-jet	344 \pm 172	471 \pm 236	815 \pm 408	45 \pm 23	30 \pm 15	75 \pm 38
$Z+\text{jets}$	85 \pm 60	88 \pm 54	173 \pm 114	41 \pm 20	32 \pm 16	73 \pm 36
Single top	449 \pm 123	550 \pm 137	999 \pm 260	21 \pm 4	27 \pm 5	48 \pm 9
Diboson	18 \pm 7	18 \pm 6	36 \pm 13	3.4 \pm 1.2	4.2 \pm 0.8	7.6 \pm 2
Total	9778 \pm 1863	12485 \pm 2323	22263 \pm 4186	826 \pm 110	1007 \pm 124	1833 \pm 234
Data	9622	12704	22326	803	1034	1837

Type	Resolved selection			Overlap selection			Boosted selection		
	$e+\text{jets}$	$\mu+\text{jets}$	Sum	$e+\text{jets}$	$\mu+\text{jets}$	Sum	$e+\text{jets}$	$\mu+\text{jets}$	Sum
SM $t\bar{t}$	9491 \pm 1753	16875 \pm 3192	368 \pm 74	415 \pm 84	783 \pm 158	145 \pm 37	200 \pm 51	345 \pm 88	
$W+\text{jets}$	1433 \pm 808	2547 \pm 1437	16 \pm 16	19 \pm 11	35 \pm 27	186 \pm 75	280 \pm 77	466 \pm 152	
Multi-jet	460 \pm 231	792 \pm 400	12 \pm 8	11 \pm 8	22 \pm 16	33 \pm 18	19 \pm 11	52 \pm 29	
$Z+\text{jets}$	85 \pm 51	167 \pm 109	3 \pm 1.5	2.7 \pm 1.4	5.7 \pm 2.9	38 \pm 19	30 \pm 15	68 \pm 34	
Single-top	437 \pm 101	973 \pm 218	12 \pm 2	14 \pm 3	27 \pm 5	9 \pm 2	13 \pm 2	22 \pm 4	
Diboson	17 \pm 6	34 \pm 13	0.21 \pm 0.12	0.11 \pm 0.02	0.32 \pm 0.14	3.2 \pm 1.1	4.1 \pm 0.8	7.3 \pm 1.9	
Total	9366 \pm 1584	12022 \pm 1948	21388 \pm 3532	411 \pm 76	461 \pm 86	872 \pm 161	415 \pm 87	961 \pm 182	
Data	9232	12227	21459	390	477	867	413	557	970

Table 6.11: Number of observed and expected events after applying the resolved and boosted selections. For each approach, the inclusive channel on the top is separated into the exclusive and overlap categories on the bottom. The total systematic uncertainty in the expected background yields is listed.

Signal m [TeV]	Resolved selection			Boosted selection		
	e+jets	μ +jets	Sum	e+jets	μ +jets	Sum
$m_{Z'} = 0.7$	89 \pm 13	103 \pm 14	192 \pm 27	27 \pm 3	32 \pm 4	59 \pm 7
$m_{Z'} = 1$	26 \pm 3	27 \pm 3	53 \pm 6	19 \pm 1	22 \pm 2	41 \pm 3
$m_{Z'} = 1.3$	6 \pm 1	6 \pm 1	12 \pm 2	6 \pm 0.1	5 \pm 0.1	11 \pm 0.2
$m_{Z'} = 1.6$	2 \pm 0.3	1.9 \pm 0.2	3.9 \pm 0.5	1.9 \pm 0.1	1.9 \pm 0.1	3.8 \pm 0.2
$m_{Z'} = 2$	0.4 \pm 0.05	0.35 \pm 0.04	0.7 \pm 0.10	0.40 \pm 0.02	0.37 \pm 0.03	0.77 \pm 0.05
$m_{g_{KK}} = 0.7$	1238 \pm 242	1639 \pm 233	2877 \pm 475	485 \pm 43	472 \pm 41	957 \pm 84
$m_{g_{KK}} = 1$	345 \pm 48	422 \pm 56	767 \pm 104	242 \pm 19	334 \pm 30	576 \pm 49
$m_{g_{KK}} = 1.3$	97 \pm 11	106 \pm 11	203 \pm 22	76 \pm 5	92 \pm 7	168 \pm 12
$m_{g_{KK}} = 1.6$	34 \pm 5	31 \pm 4	65 \pm 9	30 \pm 2	28 \pm 2	58 \pm 4
$m_{g_{KK}} = 2$	5 \pm 1	5 \pm 1	10 \pm 2	4.3 \pm 0.3	4.3 \pm 0.3	8.6 \pm 0.6

Signal m [TeV]	Resolved selection			Overlap selection			Boosted selection		
	e+jets	μ +jets	Sum	e+jets	μ +jets	Sum	e+jets	μ +jets	Sum
$m_{Z'} = 0.7$	70 \pm 11	82 \pm 12	152 \pm 23	19 \pm 3	21 \pm 4	40 \pm 7	8 \pm 1	11 \pm 3	19 \pm 4
$m_{Z'} = 1$	12 \pm 3	12 \pm 2	24 \pm 5	14 \pm 2	15 \pm 2	29 \pm 4	5 \pm 1	7 \pm 1	12 \pm 2
$m_{Z'} = 1.3$	2.4 \pm 0.7	2.6 \pm 0.5	5 \pm 1.2	4.2 \pm 0.5	3.8 \pm 0.4	8.0 \pm 0.9	1.7 \pm 0.3	1.7 \pm 0.3	3.4 \pm 0.6
$m_{Z'} = 1.6$	0.72 \pm 0.21	0.68 \pm 0.17	1.4 \pm 0.38	1.3 \pm 0.2	1.3 \pm 0.2	2.6 \pm 0.4	0.55 \pm 0.11	0.63 \pm 0.13	1.18 \pm 0.24
$m_{Z'} = 2$	0.14 \pm 0.04	0.13 \pm 0.03	0.27 \pm 0.07	0.3 \pm 0.04	0.23 \pm 0.04	0.52 \pm 0.08	0.12 \pm 0.02	0.13 \pm 0.03	0.25 \pm 0.05
$m_{g_{KK}} = 0.7$	966 \pm 195	1358 \pm 202	2324 \pm 397	272 \pm 41	281 \pm 47	553 \pm 88	213 \pm 42	191 \pm 39	404 \pm 81
$m_{g_{KK}} = 1$	180 \pm 38	197 \pm 35	377 \pm 73	165 \pm 22	225 \pm 28	390 \pm 50	77 \pm 12	109 \pm 26	186 \pm 38
$m_{g_{KK}} = 1.3$	41 \pm 12	41 \pm 7	82 \pm 19	56 \pm 6	65 \pm 7	121 \pm 13	20 \pm 4	27 \pm 4	47 \pm 8
$m_{g_{KK}} = 1.6$	12 \pm 4	13 \pm 4	25 \pm 8	22 \pm 4	18 \pm 3	40 \pm 7	8 \pm 2	10 \pm 2	18 \pm 4
$m_{g_{KK}} = 2$	2.2 \pm 0.7	2.4 \pm 0.6	4.6 \pm 1.2	3 \pm 0.4	2.9 \pm 0.4	5.9 \pm 0.8	1.3 \pm 0.2	1.4 \pm 0.2	2.7 \pm 0.4

Table 6.12: Expected number of signal events for the two benchmark models, at a variety of mass points. For each approach, the inclusive channel on the top is separated into the exclusive and overlap categories on the bottom. The total systematic uncertainty in the expected signal yields is listed.

6 Top-quark reconstruction in $t\bar{t}$ resonances searches

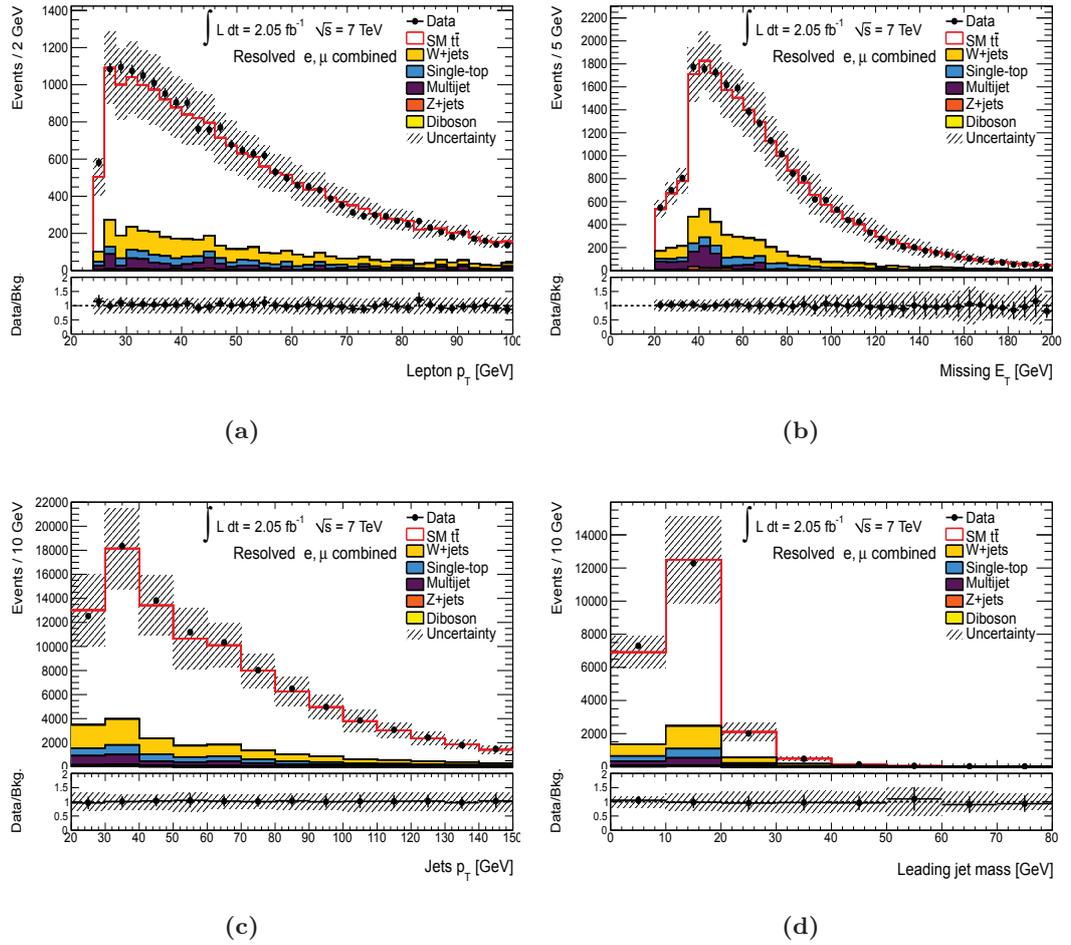


Figure 6.15: Comparison of the observed data and expected background for distributions of final-state objects in the resolved selection: (a) lepton p_T , (b) missing transverse momentum, (c) jets p_T and (d) mass of the leading jet in p_T .

The SM $t\bar{t}$ production is the dominant background in all selections. The second main background is W +jets production. The overlap channel contains approximately 50% of events of the inclusive boosted selection, whereas just over $\approx 4\%$ of the inclusive resolved events. This can be turned as an advantage to inspect the sensitivity of resonances at the TeV-scale.

Figure 6.15 shows a set of distributions obtained after the resolved selection. The lepton p_T distribution is shown in Figure 6.15a and the E_T^{miss} distribution in 6.15b. Figures 6.15c and 6.15d show the p_T of the anti- k_T $R = 0.4$ jets and the mass of the leading of these jets, respectively. Good agreement is observed between the data and the expected background. Similar shapes are observed in the distributions related to the leptonic jet

6.8 Observed data and Standard Model background prediction

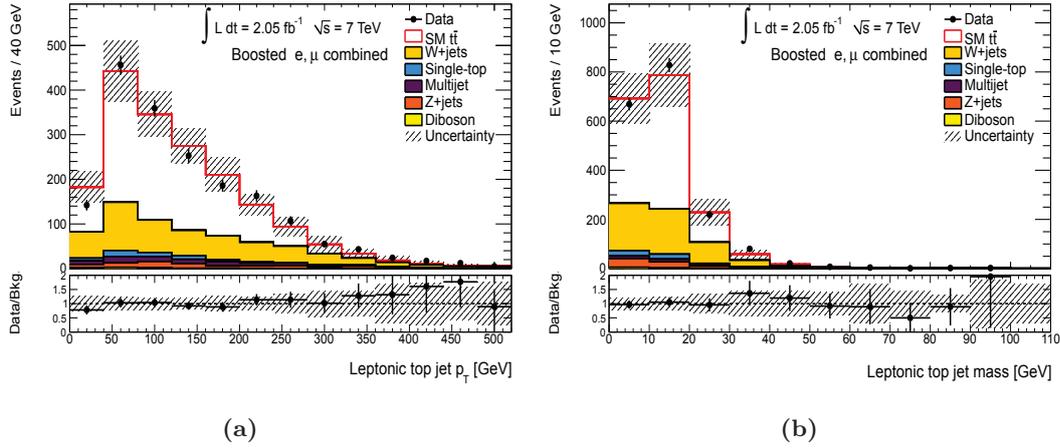


Figure 6.16: Comparison of the observed data and expected background for the leptonic top jet (a) p_T and (b) mass of the boosted reconstruction.

in the boosted selection, displayed in Figure 6.16. Most of the jets fall in the p_T range 50-100 GeV and a very small fraction of events contain leptonic jets with mass greater than 60 GeV. The W transverse mass distributions after both selections are depicted in Figures 6.17a and 6.17b. An expected Jacobian peak with an upper edge at the W mass $m_W \approx 80$ GeV is obtained in the two distributions. This occurs when the lepton lies in the transverse plane. The significant fraction of events lying below the upper edge is primarily due to leptons having some momentum in the longitudinal direction. There is also a long tail above m_W resulting from the finite width of the W boson [178]. Figure 6.18 shows the distributions related to the boosted reconstruction of the $t\bar{t}$ system. The minimal thresholds allowed on the top-tagging variables are clearly apparent in the hadronic top distributions. The mass distribution of the leptonic top quark is shown in Figure 6.18d. The peak around the top quark mass for the SM $t\bar{t}$ background is visible.

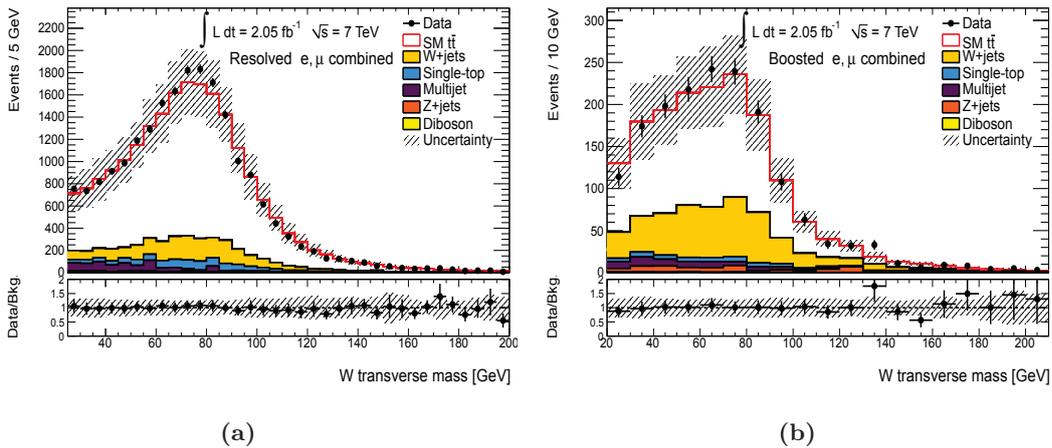


Figure 6.17: Comparison of the observed data and expected background distributions for the W transverse mass in the (a) resolved and (b) boosted selections.

6 Top-quark reconstruction in $t\bar{t}$ resonances searches

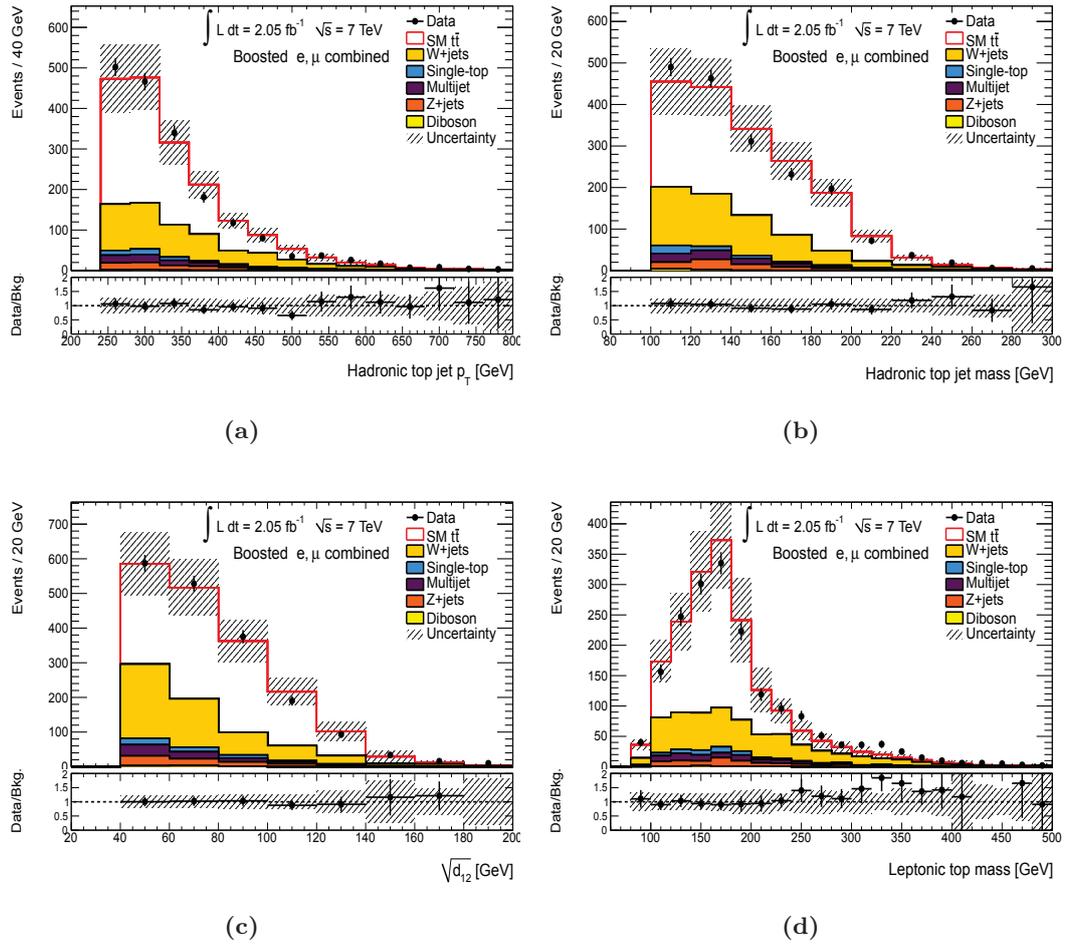


Figure 6.18: Comparison of the observed data and expected background for distributions related to boosted reconstruction: hadronic top jet (a) p_T , (b) mass, (c) first k_T -splitting scale and (d) mass of the leptopnically decaying top quark.

6.8.1 Observed and predicted $m_{t\bar{t}}$ distribution

The $m_{t\bar{t}}$ distribution is used to evaluate the sensitivity to $t\bar{t}$ resonances. Figure 6.19 shows the $m_{t\bar{t}}$ spectra for the two approaches studied in this chapter. These are formed from the spectrum of the observed data and each background in the signal region. The distributions are depicted in linear and logarithmic scales for the y -axis and in a variable bin width for the x -axis. This bin variation is set in such a way that, for the low mass region, the width does not exceed half of the experimental resolution [58]. For the high mass region, the widths of the bins are increased in order to reduce the sensitivity to statistical fluctuations. The data agree with the SM prediction within the uncertainties.

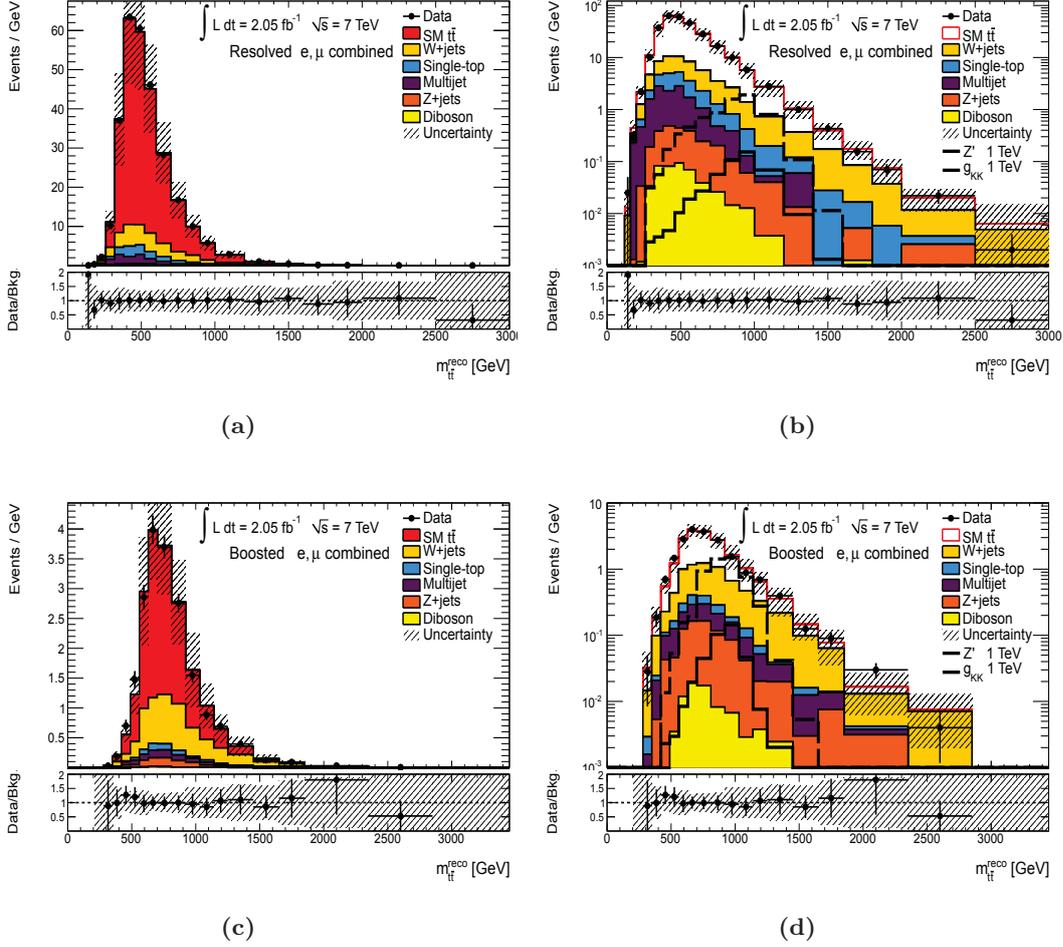


Figure 6.19: Invariant mass of the reconstructed $t\bar{t}$ system after signal selection in the (a), (b) resolved and (c), (d) boosted approaches of this analysis. The shaded area indicates the total systematic uncertainties. The two benchmark signals at $m = 1$ TeV are shown on top of the background in the logarithmic distributions. The two lepton channels are combined.

6.9 Summary and conclusions

Two approaches for $t\bar{t}$ resonances searches in the lepton+jets decay channel have been developed. They differ in the technique to assign the final-state objects to the $t\bar{t}$ system, which must be adjusted as the p_T of the hadronically decaying top quark increases. In the resolved approach of the $t\bar{t}$ reconstruction, the hadronic top-quark decay is reconstructed using two or three $R = 0.4$ jets, whereas in the boosted approach it is reconstructed as one $R = 1.0$ jet. Each approach involves its own set of criteria to select signal events. Since the final-state topology varies between the approaches, two methodologies to derive background contributions from data are stressed.

The selected objects are combined to reconstruct the invariant mass of the $t\bar{t}$ system, $m_{t\bar{t}}$. Comparisons from the selection efficiency and the mass resolution have confirmed that the boosted reconstruction is the most suitable for high resonance masses. The estimation of the systematic uncertainties affecting the two selections was also provided. From the selected events of each approach, the observed data are compared to the SM background prediction. The shapes and yields of distributions for relevant variables in the analysis are found to be in good agreement, all within the estimated uncertainties.

The development of two approaches to reconstruct the $t\bar{t}$ system can increase the sensitivity to $t\bar{t}$ resonances. The combined strategy explored here relies on the ability to discriminate between signal and background. In order to enhance the discrimination power, the total selected events in each approach (inclusive channels) are separated into two statistically independent channels: exclusive and overlap. The overlap channel contains only events that satisfy both the resolved and boosted selections. If the overlap events are subtracted from the inclusive samples, the exclusive resolved and boosted channels are created. The aim of this analysis is to quantify the benefit of this channel-separation in terms of the signal sensitivity. This is performed by means of a statistical analysis, which determines upper limits on the cross-section for signal production. Different combinations of the created channels are defined, giving rise to the results discussed in the next chapter.

7 Search for heavy $t\bar{t}$ resonances in ATLAS

The benchmark models topcolor Z' boson and KK-gluon (g_{KK}) serve to quantify the sensitivity of the analysis to $t\bar{t}$ resonances. The sensitivity is evaluated using the resolved and boosted approaches of the $t\bar{t}$ reconstruction. By combining the two selections, upper limits on the cross-section for the narrow and wide resonant states are determined. These limits are set on the cross-section times branching ratio $\sigma \times BR(Z' \rightarrow t\bar{t})$ and $\sigma \times BR(g_{\text{KK}} \rightarrow t\bar{t})$, at the 95 % confidence level (CL). Their derivation is given on the basis that no significant excess of signal events in data is found. The combination of theoretical predictions and the obtained limits, allows to constrain experimentally where possible, certain mass regions of the two models.

7.1 Existing limits

The distribution of the reconstructed $t\bar{t}$ invariant mass is used to test the compatibility of the data with the Standard Model (SM) prediction. For the same dataset as used in this analysis, the results shown in [59, 179] demonstrated that there is no significant excess of signal events over the background prediction. Therefore, observed and expected upper limits on the signal cross-section have been set, at the 95% CL. The observed limits are determined using the observed data, whereas the expected are derived using the SM prediction. The obtained limits can be transformed into exclusion regions of the resonance mass. Published upper limits on the cross-section for the Z' boson and the g_{KK} models are outlined in turn below.

Although the center-of-mass energy at the Tevatron is relatively low compared to the LHC, the sensitivity to narrow $t\bar{t}$ resonances is comparable. This is motivated by the fact that the leptophobic topcolor Z' boson is produced through $q\bar{q}$ fusion. The CDF [180–182] and DØ [183] collaborations have performed searches for $t\bar{t}$ resonances using a maximal luminosity of 5.3 fb^{-1} . The most stringent limit found excluded the Z' boson with $m_{Z'} < 0.915 \text{ TeV}$. Neither of the Tevatron experiments set limits on the KK-gluon.

For an integrated luminosity of 4.7 fb^{-1} , at $\sqrt{s} = 7 \text{ TeV}$, the CMS collaboration performed searches in the all-hadronic, dileptonic and lepton+jets decay channels [184–186]. In the lepton+jets channel, the KK-gluon with $m_{g_{\text{KK}}} < 1.4 \text{ TeV}$ and the leptophobic Z' boson with $m_{Z'} < 1.3 \text{ TeV}$ have been excluded.

At the time that this thesis was written, ATLAS published the results of the search using an integrated luminosity of 4.7 fb^{-1} . In this work, the leptophobic Z' resonance with $m_{Z'} < 1.74 \text{ TeV}$ and the KK-gluon with $m_{g_{\text{KK}}} < 2.07 \text{ TeV}$ are excluded [154].

The same 2.05 fb^{-1} of data as analyzed in this thesis were used by two separated analyses at ATLAS. The first looked at resolved $t\bar{t}$ reconstruction [59]. The observed (expected) upper limit on $\sigma \times BR(Z' \rightarrow t\bar{t})$ ranges from 9.3 (8.5) pb at $m_{Z'} = 0.5 \text{ TeV}$ to 0.95 (0.62) pb at $m_{Z'} = 1.3 \text{ TeV}$. This led to the exclusion of the leptophobic Z' boson with $0.5 \text{ TeV} < m_{Z'} < 0.88 \text{ TeV}$. The expected mass exclusion is $0.5 \text{ TeV} < m_{Z'} < 1.01 \text{ TeV}$. The observed (expected) upper limit on $\sigma \times BR(g_{KK} \rightarrow t\bar{t})$ ranges from 10.01 (10.3) pb at $m_{g_{KK}} = 0.5 \text{ TeV}$ to 1.6 (0.9) pb at $m_{g_{KK}} = 1.3 \text{ TeV}$. g_{KK} resonances with masses between 0.5 TeV and 1.13 TeV are excluded, while the expected mass exclusion is $0.5 \text{ TeV} < m_{g_{KK}} < 1.36 \text{ TeV}$.

The second search used the boosted reconstruction [179]. The observed (expected) upper limit for the Z' boson ranges from 7.7 (10.04) pb for a mass of 0.6 TeV to 0.27 (0.27) pb for a mass of 3 TeV. The leptophobic Z' is excluded for $0.6 \text{ TeV} < m_{Z'} < 1.15 \text{ TeV}$. The observed (expected) upper limit on $\sigma \times BR(g_{KK} \rightarrow t\bar{t})$ on the other hand, ranges from 2.8 (2.9) pb at $m_{g_{KK}} = 0.7 \text{ TeV}$ to 0.61 (0.38) pb at $m_{g_{KK}} = 2 \text{ TeV}$. The KK-gluon is ruled out for $0.6 \text{ TeV} < m_{g_{KK}} < 1.5 \text{ TeV}$.

7.2 Determination of upper limits on the signal cross-section

The reconstructed $t\bar{t}$ invariant mass is evaluated to determine upper limits on the signal cross-section. In general, a binned likelihood method is applied on each $m_{t\bar{t}}$ distribution formed from the selected events of each approach. A Bayesian technique [187] is used in the limits setting procedure. This is addressed to find from the analyzed dataset the so-called *posterior probability*. To determine this probability, the Bayesian interpretation requires an initial degree of belief that the signal hypothesis is true; this is the *prior probability*, $P(\text{theory})$. The prior is then updated by the observed data to infer the degree of belief that the signal hypothesis is true, given the observed data; this is the posterior probability, $P(\text{theory} | \text{data})$. The posterior probability is contrasted with the likelihood function, which is the probability of obtaining the data, given the model-hypothesis: $P(\text{data} | \text{theory})$. The posterior probability, the likelihood function and the prior probability can be related by the Bayes theorem in the way:

$$P(\text{theory} | \text{data}) \propto P(\text{data} | \text{theory})P(\text{theory}) . \quad (7.1)$$

The choice of the initial prior implies a level of subjectivity. However, rather than accounting for the degree of belief towards a given hypothesis, the prior is commonly chosen such that it maximizes the amount of missing information, while minimizing the degree of belief. This fact characterizes the *reference priors* [188]. Since a “flat” prior is a good approximation to the reference prior, this is implemented in the current analysis [189]. Given a signal model and a specific mass point, the prior for the signal cross-section is defined by [187]:

$$\pi(\sigma) = \begin{cases} \frac{1}{\sigma_{\max}} & \text{if } 0 \leq \sigma \leq \sigma_{\max} ; \\ 0 & \text{otherwise .} \end{cases} \quad (7.2)$$

7.2 Determination of upper limits on the signal cross-section

where σ is the cross-section for signal production and σ_{\max} is a numerical value chosen as will be described below. The next step to construct the posterior probability given the prior of Equation 7.2, is to obtain the likelihood function L . For counting experiments like this one, it has been proven that the data follow the Poisson statistics. The form of L is the Poisson distribution with expectation value μ :

$$L(D | \mu(I)) = \frac{e^{-\mu} \mu^D}{D!}, \quad (7.3)$$

where D is the count of data events resulting from the experiment and I encloses all the information used to obtain μ , as well as the assumption of the Poisson distribution [187]. Thus, Equation 7.3 is the probability to observe a count D in a given bin, if the mean count μ is provided. μ is the sum of the predicted contributions from background and signal. The total background yield b , is the sum over all background sources displayed in Table 6.11. For a particular resonance mass ν , the signal yield is the product of the cross-section for the resonance σ_ν , and the effective luminosity a_ν . This luminosity is defined as the product of the signal acceptance and the integrated luminosity. It should be noted that the prior knowledge of the parameters related to the experiment a and b , is independent of the signal cross-section. This fact enables to factorize the prior probability as $\pi(\sigma, a, b) = \pi(a, b)\pi(\sigma)$.

Since the probability to observe a count in a given bin (or channel) is independent of the counts observed in other bins, the total likelihood function can be expressed as the product of all single-bin likelihoods:

$$L(D | \sigma_\nu, a_\nu, b) = \prod_{i=1}^N \frac{e^{-(a_{\nu,i}\sigma_\nu + b_i)} (a_{\nu,i}\sigma_\nu + b_i)^{D_i}}{D_i!}. \quad (7.4)$$

The index i runs over all the bins of all $m_{t\bar{t}}$ spectra in all the channels. The Bayes theorem, together with the prior probability 7.2 and the likelihood function 7.4 are used to obtain the posterior probability density function (PDF):

$$P(\sigma | D) = \frac{1}{N} \int_0^\infty da_\nu \int_0^\infty db L(D | \sigma_\nu, a_\nu, b) \pi(\sigma). \quad (7.5)$$

Here, the variable of interest is σ . Any dependence on a_ν and b is removed by integrating over all possible signal and background yields. These yields can change according to statistical and systematical variations. The overall normalization constant N is determined from the condition that the probability over the full sample space must be one:

$$\int_0^\infty d\sigma \int_0^{\sigma_{\max}} da_\nu \int_0^\infty db P(\sigma | D) = 1. \quad (7.6)$$

The σ_{\max} boundary is chosen to be large enough such that the likelihood function for $\sigma > \sigma_{\max}$ is negligible, i.e. the posterior probability should be close to zero. The lower bound is set knowing that the production cross-section for signal is non-negative.

The posterior obtained is the final result, i.e. the probability density for the signal cross-section. The Bayesian upper limit on the cross-section (σ_{UL}), specified at certain level

of confidence (CL) of $(100 \times \beta)\%$, is obtained by integrating the posterior probability such that [187]

$$\beta = \int_0^{\sigma_{\text{UL}}} d\sigma P(\sigma | D). \quad (7.7)$$

The limit is calculated in different ways by using the *top_statistics* package [190].

The method to propagate the systematic uncertainties in the likelihood is based on a direct sampling approach. Each systematic uncertainty is sampled from a Gaussian distribution and then the same shift is applied to each bin of each source in each channel. The procedure is described in the following.

For an individual systematic uncertainty *isys*, a shift factor $g(0, 1)_{isys}$ is first computed by sampling a random number from a Gaussian distribution with mean zero and width one. The nominal yield y in each bin of the $m_{t\bar{t}}$ distribution is then shifted by a factor Δy_{isys} corresponding to this Gaussian and determined as

$$\Delta y_{isys} = \begin{cases} s_{tot}^+ \times g(0, 1)_{isys} \times (y_{isys}^+ - y) & \text{if } g(0, 1)_{isys} > 0; \\ s_{tot}^- \times g(0, 1)_{isys} \times (y - y_{isys}^-) & \text{otherwise.} \end{cases} \quad (7.8)$$

Here $y_{isys}^{+(-)}$ is the yield in the bin after applying the systematic *isys* at the level of one standard deviation (1σ), whose factor $g(0, 1)_{isys}$ is positive (negative). The multiplicative factor $s_{tot}^{+(-)}$ is one if the effect of the systematic uncertainty modifies both, the shape and normalization of the $m_{t\bar{t}}$ distribution. If it affects the shape rather than the normalization of the distribution, then $s_{tot}^{+(-)}$ is set so the overall normalization can be corrected. The new background or signal yield in the bin j is obtained from the sum of all individual shifts with respect to the nominal value, plus the nominal yield:

$$y'_j = y_j + \sum_{isys} \Delta y_{isys}, \quad (7.9)$$

where y'_j is used to calculate the new likelihood. Integration over the sampled Gaussian is performed by drawing random numbers from its distribution n -times and computing the likelihood each time. The integral of each posterior is preserved, which means that those individual likelihoods with a large integral will be the main contributions to the final posterior. Uncertainties whose effect is too large can truncate the Gaussian, leading to a negative yield in a particular bin. To reduce this possibility, systematic effects larger than 20% are sampled using a log-normal distribution instead a Gaussian [190].

To validate the method and estimate the sensitivity, pseudo-data from the input data distributions have been generated. A count in each bin is sampled from a Poisson distribution, with mean the same as the nominal distribution and within one standard deviation. The procedure is repeated in N pseudo-experiments and for each one, the expected upper limit is calculated¹. Figure 7.1a shows an example of the 95% CL upper limit distribution. The median defined as the 50% quantile of this distribution is the expected limit. The corresponding posterior PDF is shown in Figure 7.1b. The 84%

¹To calculate the expected upper limit the description outlined up to Equation 7.7 is repeated, with D being the bin count of the expected background yield.

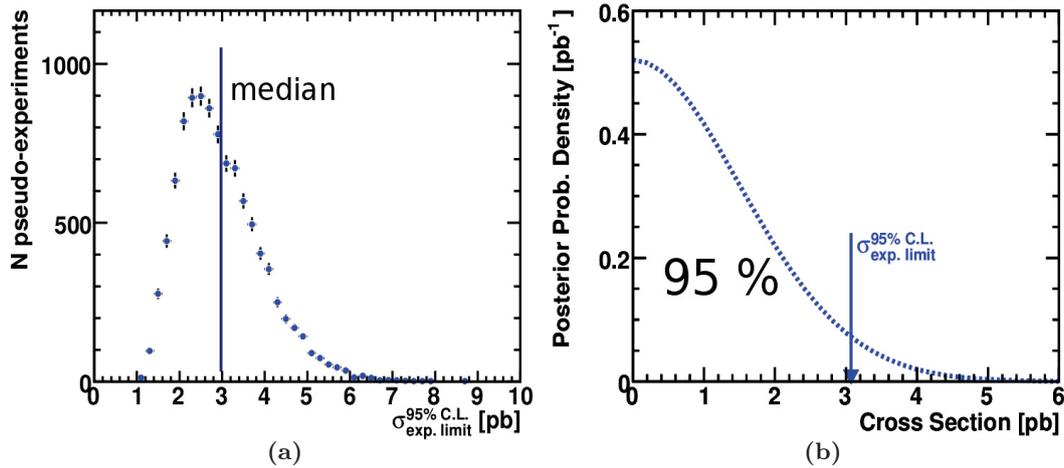
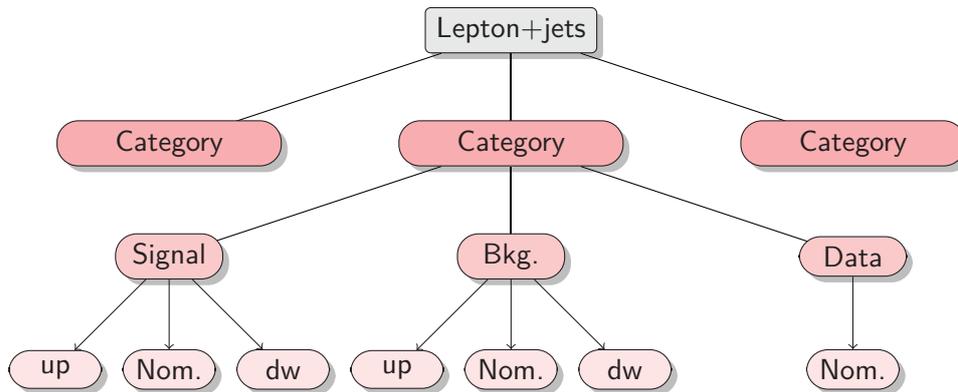


Figure 7.1: (a) Example of the upper limits distribution for 10000 pseudo-datasets. The expected cross-section is defined as a median of the distribution. (b) Posterior probability density function. The upper limit on the signal cross-section is calculated at 95% confidence level [191].

and 16% of the integral of the latter distribution are respectively, the $+1\sigma$ and -1σ uncertainties on the cross-section. The $+2\sigma$ and -2σ uncertainties are the 97.5% and 2.5%, respectively [191].

7.3 Inputs to the limits setting procedure

The reconstructed $t\bar{t}$ invariant mass distributions from data, signal and from the individual backgrounds are the inputs to the limits setting procedure. Each input is obtained from the events selected separately in the e +jets and μ +jets channels. The nominal (Nom.) distributions are provided as reconstructed in data, signal and each background (Bkg.). The background and signal distributions must in addition be given with the variations accounting for the systematic uncertainties. The whole chain of inputs required is illustrated in the following diagram:



with each ‘‘Category’’ being the channel in which the selected events have been separated, i.e. inclusive, overlap and exclusive (see Section 6.7.6). The distributions after applying the up and down (dw) variations are provided individually for each systematic uncertainty. All the distributions are created with a variable bin widths, lying between 80 to 600 GeV. These values are set such that for masses below $m_{t\bar{t}} = 1.4$ TeV the bin widths are no wider than half of the experimental resolution². Beyond this region, the bin width is increased to reduce statistical fluctuations relevant for the background distributions. The stability of the limits due to the changing binning is discussed in [179]. Specifically, the limits obtained from variable-binning inputs were compared to those resulting from a constant binning of 80 GeV. The variation was found to be between 5% and 15%, which is less than half of the 1σ shift due to statistical fluctuations (30% - 50%) [58].

7.4 Results

Four combinations of the categories are established for the limits calculation. Each combination contains its own inputs as described in Section 7.3 above. These combinations are listed as follows:

- I. Inclusive boosted:** Is not exactly a combination of categories but is considered as such for simplicity in the description below. It is used as reference for comparisons of the results. The corresponding distributions are shown in Figures 7.2a and 7.2b.
- II. Exclusive boosted + Overlap $m_{t\bar{t}}^{\text{boosted}}$:** As already defined in Section 6.7.6, the exclusive categories are created from the selected events in each approach, with the overlap events subtracted from the inclusive channels. The exclusive boosted distributions are shown in Figures 7.2c and 7.2d. The overlap $m_{t\bar{t}}$ distributions are obtained from the overlap events, using the boosted reconstruction (overlap $m_{t\bar{t}}^{\text{boosted}}$). These are shown in Figures 7.2e and 7.2f.
- III. Exclusive boosted + Overlap $m_{t\bar{t}}^{\text{boosted}}$ + Exclusive resolved:** Figures 7.3a and 7.3b show the $m_{t\bar{t}}$ distributions corresponding to the exclusive resolved category. The $m_{t\bar{t}}$ distributions shown in Figures 7.3c and 7.3d are created in the overlap channel using the resolved reconstruction (overlap $m_{t\bar{t}}^{\text{resolved}}$).
- IV. Exclusive boosted + Overlap $m_{t\bar{t}}^{\text{resolved}}$ + Exclusive resolved.**

The sensitivity of the analysis to $t\bar{t}$ resonances is assessed in terms of the upper limits on the production cross-section for signal. The limits have been set for each of the four combinations above, following the statistical method described in Section 7.2. A simultaneous statistical evaluation of all the channels for each combination is performed. Any combination is executed independently from the others, resulting in four sets of upper limits on the signal cross-section.

²Considered to be approximately 6-12% of the resonance mass.

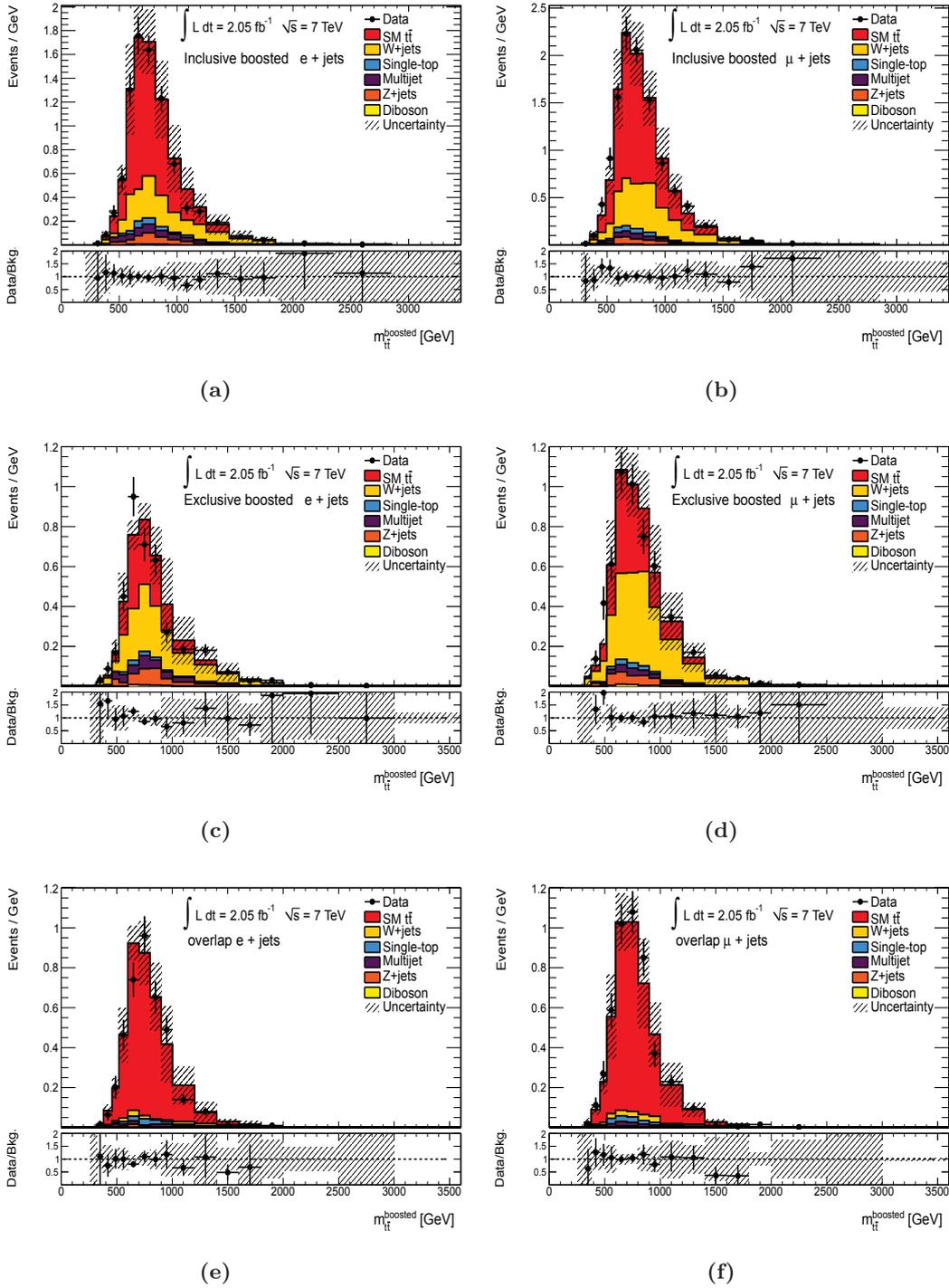


Figure 7.2: Invariant mass of the reconstructed $t\bar{t}$ system for all events in the categories of the boosted reconstruction: inclusive boosted in the (a) $e + \text{jets}$ and (b) $\mu + \text{jets}$ channels, exclusive boosted in the (c) $e + \text{jets}$ and (d) $\mu + \text{jets}$ channels and overlap boosted in the (e) $e + \text{jets}$ and (f) $\mu + \text{jets}$ channels. The distributions derived from data and SM background prediction are shown. The shaded band indicates the total systematic uncertainty.

7 Search for heavy $t\bar{t}$ resonances in ATLAS

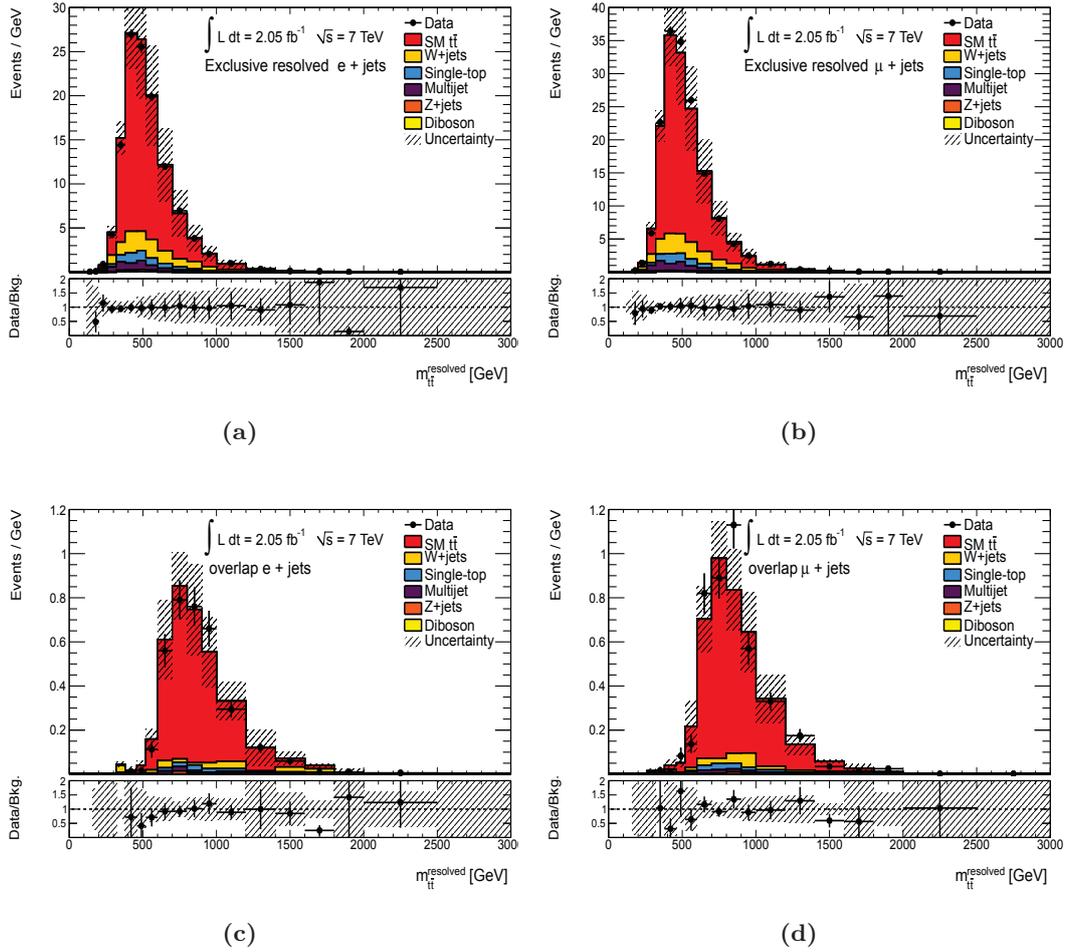


Figure 7.3: Invariant mass of the reconstructed $t\bar{t}$ system for all events in the exclusive and overlap categories of the resolved reconstruction: (a) exclusive resolved in the e +jets channel, (b) exclusive resolved in the μ +jets channel, (c) overlap resolved in the e +jets channel and (d) overlap resolved in the μ +jets channel. The distributions obtained from data events are compared to those from the SM background prediction. The shaded band indicates the total systematic uncertainty.

7.4.1 Combination I: Inclusive boosted

Table 7.1 shows the resulting upper limits from the inclusive boosted category. The expected upper limits for the topcolor Z' boson range from 10.71 pb for $m_{Z'} = 0.6$ TeV to 0.27 pb for $m_{Z'} = 3$ TeV. For the wide g_{KK} resonance they are set to 2.77 pb and 0.37 pb for $m_{g_{KK}} = 0.7$ TeV and $m_{g_{KK}} = 2$ TeV, respectively. For the two benchmark models, the same limits are shown in Figure 7.4, together with the theoretical values of $\sigma \times BR(\rightarrow t\bar{t})$. These are represented by the red band whose finite width depicts the effect of the PDF uncertainty.

$Z' \rightarrow t\bar{t}$ limits					
Z' Mass [GeV]	Observed [pb]	Expected [pb]	-1σ [pb]	$+1\sigma$ [pb]	
600	7.58	10.71	7.0	15.98	
700	2.20	2.63	1.76	3.93	
800	1.13	1.49	0.99	2.30	
1000	0.49	0.69	0.47	1.05	
1300	0.56	0.38	0.26	0.56	
1600	0.22	0.25	0.17	0.36	
2000	0.35	0.18	0.12	0.26	
3000	0.24	0.27	0.18	0.40	
$g_{KK} \rightarrow t\bar{t}$ limits					
g_{KK} Mass [GeV]	Observed [pb]	Expected [pb]	-1σ [pb]	$+1\sigma$ [pb]	
700	2.55	2.77	1.91	4.14	
800	2.20	2.14	1.41	3.16	
900	1.06	1.50	0.99	2.26	
1000	0.60	0.92	0.63	1.43	
1150	0.50	0.63	0.43	0.94	
1300	0.78	0.57	0.39	0.86	
1600	0.39	0.39	0.27	0.57	
1800	0.49	0.38	0.26	0.56	
2000	0.60	0.37	0.25	0.54	

Table 7.1: Observed and expected upper limits on the production cross-section times $t\bar{t}$ branching ratio for the leptophobic Z' boson and the KK-gluon g_{KK} . The limits are obtained at the 95% confidence level from the inclusive boosted category. Systematic and statistical uncertainties have been included. The $\pm 1\sigma$ variation in the expected limits is also given.

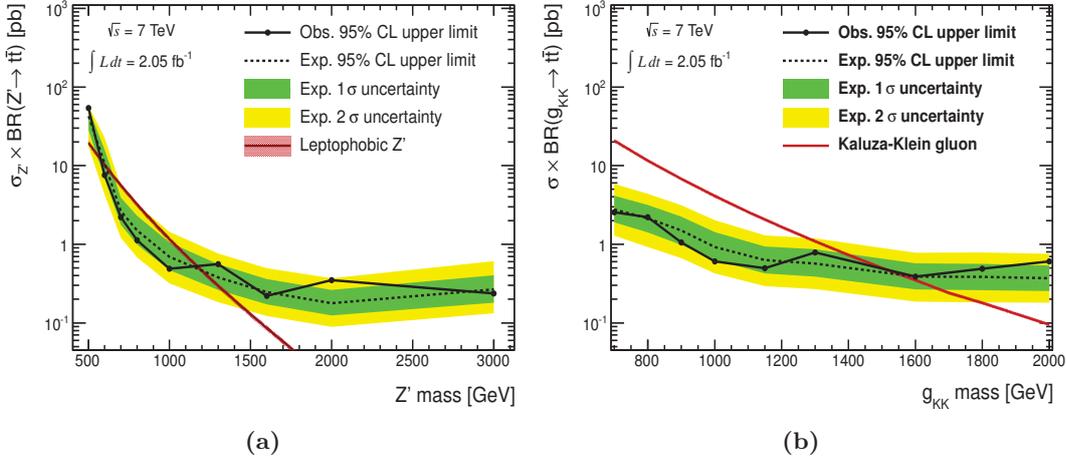


Figure 7.4: Upper limits on the production cross-section times $t\bar{t}$ branching ratio for (a) the Z' boson and (b) the KK-gluon g_{KK} , in the inclusive boosted category. The observed and expected limits are represented by the solid and dashed lines, respectively. The green and yellow bands show the range in which the limit is expected to lie in 68% and 95% of pseudo-experiments, respectively, and the smooth red lines correspond to the predicted production cross-section times branching ratio for the signal model in question. The band around the cross-section curve is based on the effect due to the systematic uncertainty associated with the determination of the PDFs.

7.4.2 Combination II: Exclusive boosted + Overlap $m_{t\bar{t}}^{\text{boosted}}$

The events in the exclusive boosted plus overlap $m_{t\bar{t}}^{\text{boosted}}$ channels are the same as those in the inclusive boosted category, i.e. combination I. Table 7.2 and Figure 7.5 show the limits resulting from the combination II. For the full range of resonances, the expected limits are stronger after the separation of the inclusive boosted channel into the exclusive boosted and overlap.

$Z' \rightarrow t\bar{t}$ limits					
Z' Mass [GeV]	Observed [pb]	Expected [pb]	-1σ [pb]	$+1\sigma$ [pb]	
600	7.11	7.05	4.45	10.77	
700	1.92	2.53	1.57	3.82	
800	1.15	1.26	0.80	2.00	
1000	0.63	0.56	0.37	0.89	
1300	0.35	0.31	0.20	0.47	
1600	0.13	0.19	0.13	0.29	
2000	0.23	0.13	0.09	0.19	
3000	0.29	0.17	0.12	0.26	
$g_{\text{KK}} \rightarrow t\bar{t}$ limits					
g_{KK} Mass [GeV]	Observed [pb]	Expected [pb]	-1σ [pb]	$+1\sigma$ [pb]	
700	3.07	2.53	1.64	3.87	
800	2.03	1.71	1.06	2.51	
900	1.10	1.21	0.77	1.83	
1000	0.69	0.74	0.48	1.13	
1150	0.42	0.51	0.32	0.81	
1300	0.41	0.44	0.28	0.67	
1600	0.24	0.29	0.19	0.43	
1800	0.33	0.28	0.19	0.41	
2000	0.39	0.26	0.17	0.39	

Table 7.2: Observed and expected upper limits on the production cross-section times $t\bar{t}$ branching ratio for the leptophobic Z' boson and the KK-gluon g_{KK} , at the 95% CL, combining the exclusive boosted and overlap $m_{t\bar{t}}^{\text{boosted}}$ categories.

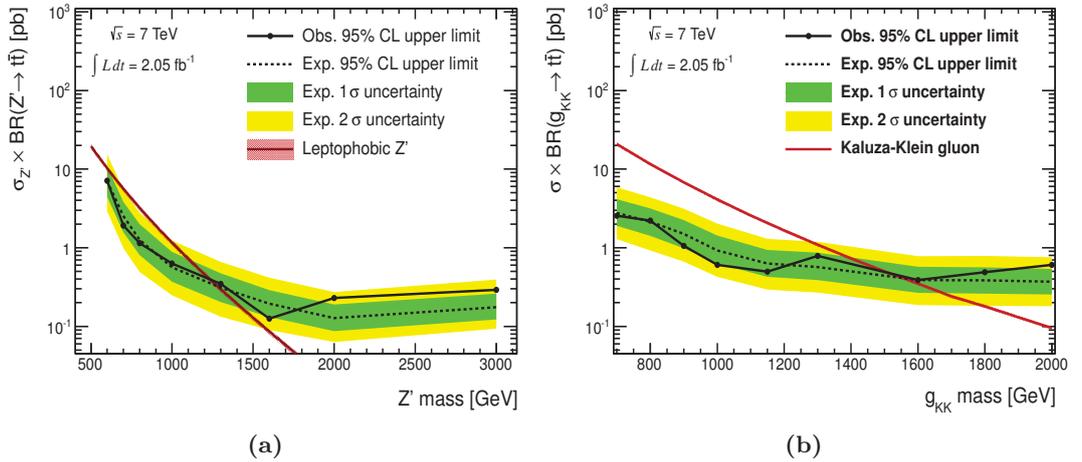


Figure 7.5: Observed and expected upper limits on the production cross-section times $t\bar{t}$ branching ratio for (a) the leptophobic Z' boson and (b) the KK-gluon g_{KK} , combining the exclusive boosted and overlap $m_{t\bar{t}}^{\text{boosted}}$ categories. Systematic and statistical uncertainties are included.

7.4.3 Combination III: Exclusive boosted + Overlap $m_{t\bar{t}}^{\text{boosted}}$ + Exclusive resolved

Table 7.3 and Figure 7.6 show the upper limits obtained from the combination III. For the topcolor Z' boson the expected limits range from 4.81 pb for $m_{Z'} = 0.6$ TeV to 0.18 pb for $m_{Z'} = 3$ TeV. For the g_{KK} they vary between 2.59 pb and 0.27 pb for $m_{g_{\text{KK}}} = 0.7$ TeV and $m_{g_{\text{KK}}} = 2$ TeV, respectively.

$Z' \rightarrow t\bar{t}$ limits				
Z' Mass [GeV]	Observed [pb]	Expected [pb]	-1σ [pb]	$+1\sigma$ [pb]
600	4.85	4.81	2.27	8.33
700	1.71	2.34	1.05	3.64
800	0.85	1.18	0.56	2.02
1000	0.35	0.51	0.27	0.85
1300	0.30	0.30	0.17	0.49
1600	0.15	0.19	0.12	0.31
2000	0.21	0.13	0.09	0.22
3000	0.25	0.18	0.11	0.31
$g_{\text{KK}} \rightarrow t\bar{t}$ limits				
g_{KK} Mass [GeV]	Observed [pb]	Expected [pb]	-1σ [pb]	$+1\sigma$ [pb]
700	2.52	2.59	1.11	4.31
800	1.09	1.52	0.74	2.50
900	0.77	1.05	0.51	1.70
1000	0.45	0.67	0.34	1.17
1150	0.42	0.48	0.24	0.80
1300	0.39	0.41	0.22	0.71
1600	0.23	0.28	0.16	0.45
1800	0.34	0.28	0.17	0.46
2000	0.37	0.27	0.16	0.44

Table 7.3: Observed and expected upper limits on the production cross-section times $t\bar{t}$ branching ratio for the leptophobic Z' boson and the KK-gluon g_{KK} , at the 95% confidence level, combining the exclusive boosted, overlap $m_{t\bar{t}}^{\text{boosted}}$ and exclusive resolved categories.

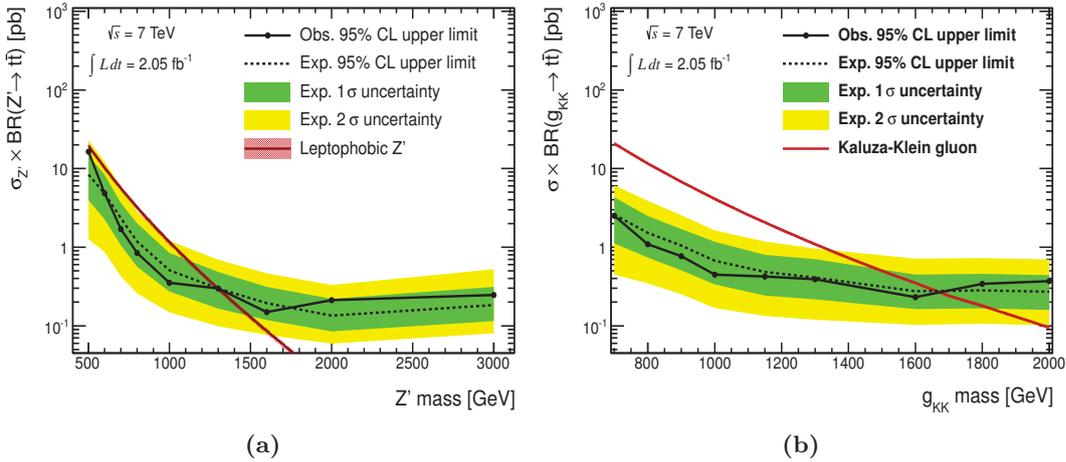


Figure 7.6: Observed and expected upper limits on the production cross-section times $t\bar{t}$ branching ratio for (a) the Z' boson, (b) the KK-gluon, combining the exclusive boosted, overlap $m_{t\bar{t}}^{\text{boosted}}$ and exclusive resolved categories.

7.4.4 Combination IV: Exclusive boosted + Overlap $m_{t\bar{t}}^{\text{resolved}}$ + Exclusive resolved

Table 7.4 and Figure 7.7 show the upper limits on the signal cross-section derived from the combination IV. The expected limits for the topcolor Z' boson vary from 4.75 pb for $m_{Z'} = 0.6$ TeV to 0.22 pb for $m_{Z'} = 3$ TeV. For the wide g_{KK} resonance they range between 2.57 pb and 0.35 pb for $m_{g_{\text{KK}}} = 0.7$ TeV and $m_{g_{\text{KK}}} = 2$ TeV, respectively.

$Z' \rightarrow t\bar{t}$ limits				
Z' Mass [GeV]	Observed [pb]	Expected [pb]	-1σ [pb]	$+1\sigma$ [pb]
600	5.13	4.75	2.17	8.35
700	1.85	2.35	1.00	3.64
800	0.91	1.34	0.59	2.22
1000	0.74	0.64	0.32	1.04
1300	0.61	0.37	0.20	0.59
1600	0.29	0.26	0.15	0.42
2000	0.27	0.19	0.12	0.30
3000	0.26	0.22	0.13	0.37
$g_{\text{KK}} \rightarrow t\bar{t}$ limits				
g_{KK} Mass [GeV]	Observed [pb]	Expected [pb]	-1σ [pb]	$+1\sigma$ [pb]
700	3.97	2.57	1.11	4.34
800	1.06	1.63	0.75	2.75
900	1.20	1.19	0.54	1.93
1000	0.85	0.81	0.37	1.37
1150	0.57	0.57	0.28	0.92
1300	0.57	0.50	0.27	0.84
1600	0.42	0.38	0.21	0.64
1800	0.39	0.36	0.20	0.61
2000	0.56	0.35	0.20	0.59

Table 7.4: Observed and expected upper limits on the production cross-section times $t\bar{t}$ branching ratio for the leptophobic Z' boson and the KK-gluon g_{KK} , at the 95% confidence level, combining the exclusive boosted, overlap $m_{t\bar{t}}^{\text{resolved}}$ and exclusive resolved categories.

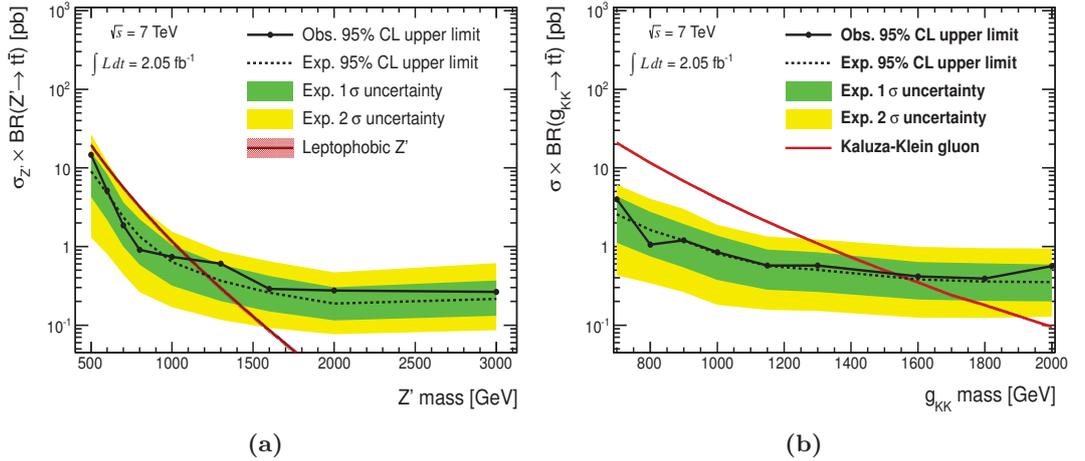


Figure 7.7: Observed and expected upper limits on the production cross-section times $t\bar{t}$ branching ratio for (a) the leptophobic Z' boson, (b) the KK-gluon g_{KK} , combining the exclusive boosted, overlap $m_{t\bar{t}}^{\text{resolved}}$ and exclusive resolved categories. Systematic and statistical uncertainties are included.

7.5 Comparison of the sensitivity for the different combinations

The sensitivity on the exclusion for $t\bar{t}$ resonances is evaluated in this section. This is achieved by comparing the upper limits on the production cross-section for the two benchmark models. The four combinations of categories defined previously are listed again for convenience:

- Combination I** : Inclusive boosted ,
- Combination II** : Exclusive boosted + Overlap $m_{t\bar{t}}^{\text{boosted}}$,
- Combination III** : Exclusive boosted + Overlap $m_{t\bar{t}}^{\text{boosted}}$ + Exclusive resolved ,
- Combination IV** : Exclusive boosted + Overlap $m_{t\bar{t}}^{\text{resolved}}$ + Exclusive resolved .

Figures 7.8 — 7.12 show the expected and observed limits for resonance masses between 0.8 TeV and 2.0 TeV. Regardless the resonance mass, the expected limits derived from the combination II are always stronger than those from the combination I. Such a gain is due to the splitting of the inclusive boosted category in the exclusive and overlap. The overlap channel contains the events that pass the resolved and boosted selections. If the background contribution is small in the overlap channel, it has to be large in the exclusive channel. This guarantees that when the two contributions are added, the total number of background events in the inclusive channel remains constant. The sensitivity is only identical if the background level in the exclusive and overlap channels is the same. Unlike, having much less background for one channel and much more for the other yields stronger limits, thereby better sensitivity. Therefore, the sensitivity is better in the combination II than in the combination I, despite they have the same events.

The sensitivity is further improved in the combination III, demonstrating the benefit of adding categories with additional events. The combination III is defined by adding the

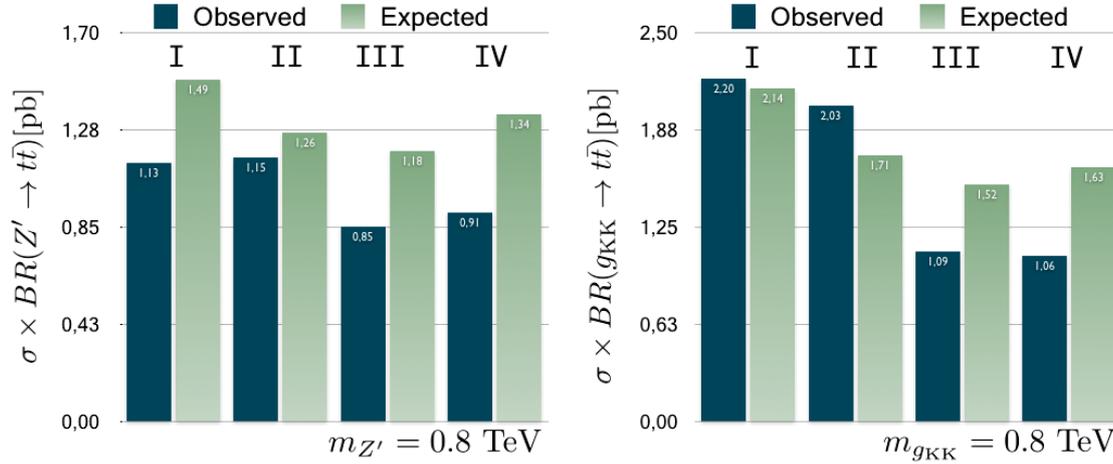


Figure 7.8: Observed and expected upper limits on the production cross-section times $t\bar{t}$ branching ratio for the narrow (left) and wide (right) resonant states with a mass of $m = 0.8 \text{ TeV}$, for distinct combinations of selection categories: (I) inclusive boosted, (II) exclusive boosted + overlap $m_{t\bar{t}}^{\text{boosted}}$, (III) exclusive boosted + overlap $m_{t\bar{t}}^{\text{boosted}}$ + exclusive resolved and (IV) exclusive boosted + overlap $m_{t\bar{t}}^{\text{resolved}}$ + exclusive resolved.

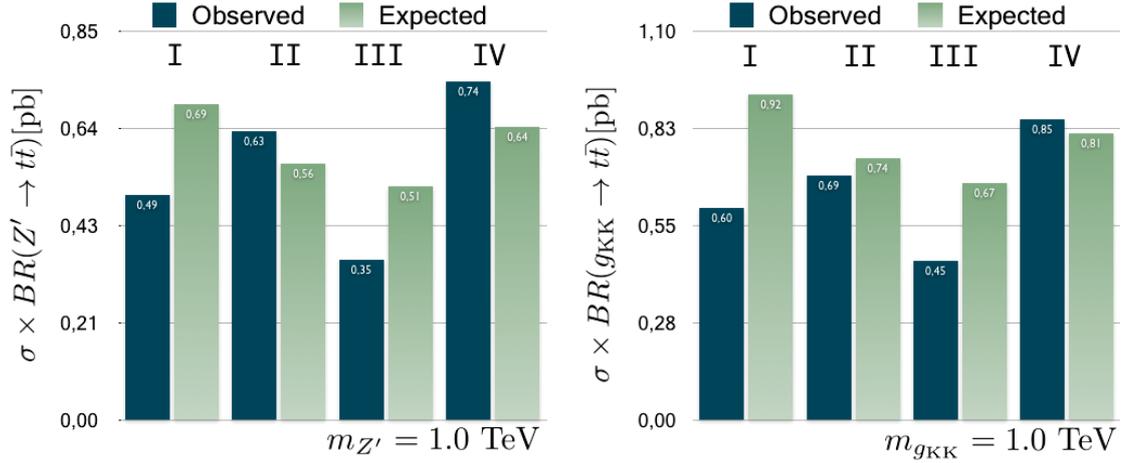


Figure 7.9: Observed and expected upper limits on the production cross-section times $t\bar{t}$ branching ratio for the narrow (left) and wide (right) resonant states with a mass of $m = 1.0$ TeV. Other details as for Figure 7.8.

resolved exclusive category to the boosted information. The addition of this category improves the sensitivity obtained from the combination II, specifically at low masses where the resolved reconstruction is dominant. By contrast, no gain is visible for high masses: the expected limits fluctuate within a 5% uncertainty due to the limited number of pseudo-experiments. The drop in the selection efficiency of the resolved approach is responsible for the loss on sensitivity with increasing the resonance mass.

The combination IV is created by adding the exclusive boosted category to the resolved information. For all masses, the limits obtained from this combination are weaker than those from the combination III, i.e. the sensitivity is lower.

The gain in sensitivity is achieved by introducing the overlap category. This category, together with the exclusive information from each approach, enables to enhance the

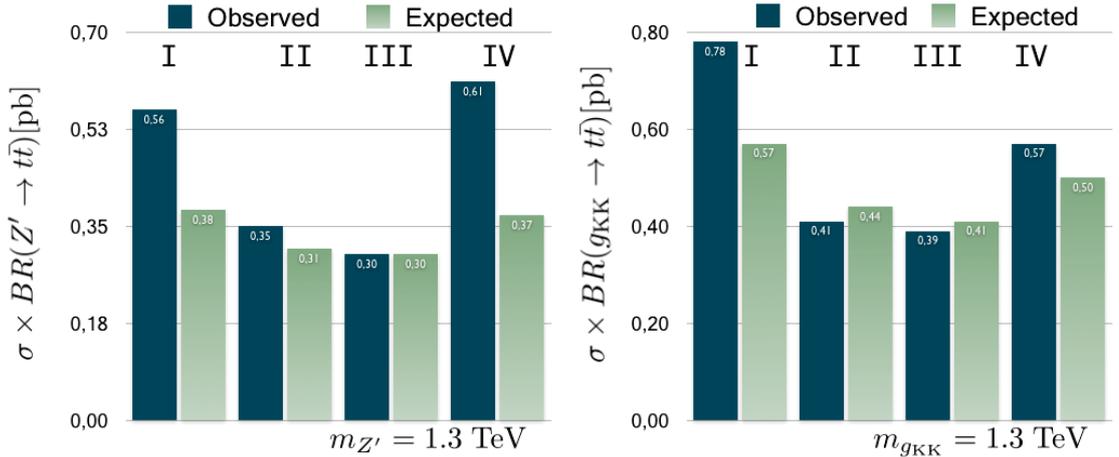


Figure 7.10: Observed and expected upper limits on the production cross-section times $t\bar{t}$ branching ratio for the narrow (left) and wide (right) resonant states with a mass of $m = 1.3$ TeV. Other details as for Figure 7.8.

7.5 Comparison of the sensitivity for the different combinations

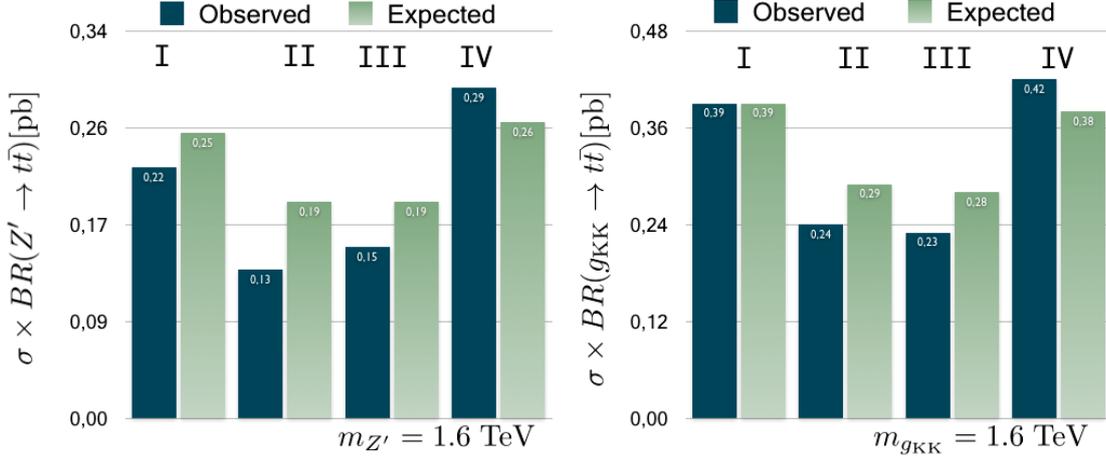


Figure 7.11: Observed and expected upper limits on the production cross-section times $t\bar{t}$ branching ratio for the narrow (left) and wide (right) resonant states with a mass of $m = 1.6 \text{ TeV}$, for distinct combinations of selection channels. Other details as for Figure 7.8.

sensitivity as more categories are added: starting from the combination I up to the combination III. Note that the invariant mass in overlap category of these combinations is reconstructed by using the boosted definition.

The previous comparisons show that the boosted reconstruction in the overlap channel provides better sensibility in a) regions where the resolved reconstruction is most suitable, b) in intermediate regions where the resolved and boosted reconstructions are valid and c) in regions where the boosted reconstruction dominates. The effect is stronger at the TeV-scale where the sensitivity improves by a maximum factor of 1.4, with respect to the combination I.

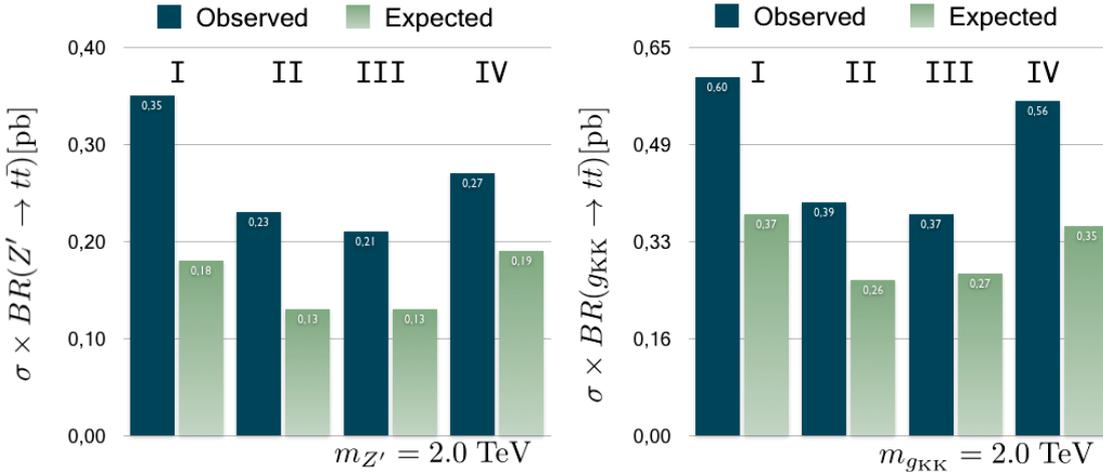


Figure 7.12: Observed and expected upper limits on the production cross-section times $t\bar{t}$ branching ratio for the narrow (left) and wide (right) resonant states with a mass of $m = 2.0 \text{ TeV}$. Other details as for Figure 7.8.

The highest sensitivity given by the boosted reconstruction in the overlap channel is a consequence of the better reconstructed mass resolution in the boosted approach. Namely, if the resolution is better, the background interference in the signal region is less relevant and the statistic power is higher. Shape-changing effects due to the systematic uncertainties also affect the sensitivity. Strictly speaking, larger systematic effects in the resolved overlap channel produce weaker limits, thus loss on sensibility. In the same way, if the effects in the overlap boosted channel are small, the limits tend to be strong.

The increasing sensitivity from the combination I to the combination III is also notable in the observed limits. In particular for high resonance masses, these limits are improved by a maximum factor of 2, i.e. 50% improvement. Statistical fluctuations cause the weaker observed limits for resonances at $m = 1$ TeV, shown in Figure 7.9.

The width of the resonances is reflected by the obtained results: the upper limits on the signal cross-section derived for broad resonances are weaker than those for narrow resonances. All the limits set in this analysis are applicable to any resonance whose width is narrow in comparison to the experimental resolution. This condition is valid as long as the width does not contribute significantly to the reconstructed $t\bar{t}$ invariant mass. The validity is in particular tailored for the topcolor Z' boson. For the KK-gluon however, further considerations related to the parameters of the $t\bar{t}$ decay must be taken into account (see Section 2.4.1.2).

The upper limits obtained are transformed into mass exclusion regions of the two predicted resonances, at the 95% CL. From the combination III and based on Figure 7.6a, the leptophobic topcolor Z' boson is excluded in the mass region $0.6 < m_{Z'} < 1.3$ TeV. From Figure 7.6b, the g_{KK} with a mass below 1.65 TeV is also excluded.

7.6 Comparisons to other $t\bar{t}$ resonances searches

The upper limits obtained from the inclusive boosted category are compared to the results presented in [179]. The published analysis implemented the boosted reconstruction on the same dataset as used in this thesis. For the full range of masses, the observed and expected limits on the signal cross-section were found to be in very good agreement with the results obtained from the combination I. Any slight difference lies within the $\pm 1\sigma$ uncertainties on the cross-section. Such a level of agreement enabled the use of the combination I as a reliable reference for the comparisons shown in this chapter³.

The search [59] used the same dataset as in this thesis, but implemented the resolved reconstruction only. The published upper limits in this analysis are listed in Table 7.5. Notably, the limits obtained from the boosted inclusive category of this thesis (Table 7.1) are stronger for masses greater than 0.8 TeV. The boosted reconstruction is more sensitive to the $t\bar{t}$ resonances at the TeV-scale, as expected.

³The analysis [179] is exactly the same as that developed in the combination I.

Mass [GeV]	Z' Obs. [pb]	Z' Exp. [pb]	g_{KK} Obs. [pb]	g_{KK} Exp. [pb]
500	9.3	8.5	10.1	10.3
600	4.8	6.0	5.0	6.0
700	2.5	3.1	3.1	4.2
800	1.9	2.1	2.2	2.7
1000	2.4	1.1	2.9	1.4
1300	0.95	0.62	1.6	0.90
1600	0.76	0.46	1.4	0.68

Table 7.5: Observed (Obs.) and expected (Exp.) 95% CL upper limits on the production cross-section times $t\bar{t}$ branching ratio for the leptophobic Z' boson and the KK-gluon g_{KK} in the resolved analysis published in [59]. These results are compared to those obtained in this thesis from the inclusive boosted category shown in Table 7.1.

The upper limits displayed in Table 7.3 can be compared to the results published in [154], in which 4.7 fb^{-1} of data were analyzed. This analysis combines the resolved and boosted reconstructions but, unlike the strategy implemented in this thesis, only two orthogonal categories are created. The results in the cited analysis are as follows. For a leptophobic topcolor Z' the expected upper limits are 4.2, 2.3, 1.45 and 0.49 pb, for $m_{Z'} = 0.6, 0.7, 0.8$ and 1 TeV, respectively. For the g_{KK} resonance the expected limits are 3.5, 1.86 and 0.76 pb, for $m_{g_{KK}} = 0.7, 0.8$ and 1 TeV, respectively. The comparison demonstrates the reason behind the strategy followed, with three categories instead only two. The results derived from the combination III are highly competitive with the results outlined in the publication, taking into account that the latter used more than double of the integrated luminosity.

The competence of the results of this thesis is confirmed by contemporary results obtained by the CMS collaboration [192]. These were achieved from searches in the lepton+jets channel using 4.7 fb^{-1} of data. They excluded the leptophobic topcolor Z' for masses below 1.3 TeV, and the KK-gluon with a mass below 1.4 TeV. Again, it is worth saying that the data is more than twice the amount of data used in the $t\bar{t}$ resonances analysis of this thesis.

7.7 Summary and conclusions

Different combinations of three channels have been defined in order to investigate the sensitivity of searches to $t\bar{t}$ resonances. Each channel is composed of $t\bar{t}$ events selected by the resolved and boosted approaches of the $t\bar{t}$ reconstruction, and the overlap between them. For each combination, 95% CL upper limits on the production cross-section are determined by using a Bayesian technique. The limits are set on the cross-section times branching ratio of the Z' and KK-gluon resonances, across a wide range of masses of the $t\bar{t}$ system.

The most stringent upper limits are achieved by combining the exclusive categories of the boosted and resolved selections, together with the overlap channel. In the latter, the $t\bar{t}$ invariant mass is obtained by using the boosted reconstruction. For the Z' resonance, the observed (expected) upper limits range from 4.85 (4.81) pb for $m_{Z'} = 0.6$ TeV, up to 0.21 (0.13) pb for $m_{Z'} = 2$ TeV. For the KK-gluon, these vary between 2.52 (2.59) pb and 0.37 (0.27) pb for $m_{g_{\text{KK}}} = 0.7$ TeV and $m_{g_{\text{KK}}} = 2$ TeV, respectively. As expected, the resolved reconstruction only improves the sensitivity at low resonance masses where it dominates. Unlike, the boosted reconstruction provides better sensitivity in the low mass regions, as well as in intermediate regions and at the TeV-scale. The upper limits obtained are used to exclude the topcolor Z' boson with a mass below 1.3 TeV and the KK-gluon with mass below 1.65 TeV.

The strategy of creating an overlap channel, allowing for the addition of different selection categories, represents a significant improvement in the sensitivity of the searches to $t\bar{t}$ resonances. This is ratified through comparisons to previous analyses in ATLAS, which used the same dataset analyzed in this thesis. The potential of the strategy is also highlighted by means of comparisons to searches using higher luminosities.

8 Summary and conclusions

In this thesis, the efficiencies of inclusive, multi and global jet triggers were studied. The focus has been on the level-1 triggers, which were used in analyses requiring multi-jets in the final states. The data employed for these performance studies were collected by the ATLAS detector at $\sqrt{s} = 7$ TeV. Trigger efficiency turn-on curves have been shown as calculated in data using the unbiased bootstrap method, and as obtained in Monte Carlo simulation. Orthogonal triggers were also implemented as baseline to derive the efficiencies of triggers with high thresholds. The plateau of the efficiency curves for the studied triggers was used to extract the multi-jet cross-section in analyses using data collected in 2010.

Beyond the Standard Model theories predict the existence of new heavy particles decaying to a top and an anti-top quark ($t\bar{t}$) pair. The main part of the thesis concentrates on the search for such $t\bar{t}$ resonances. The search has been performed in the lepton+jets channel of the $t\bar{t}$ decay. The data used were collected with the ATLAS detector, at a center-of-mass energy of 7 TeV. This data sample corresponds to an integrated luminosity of 2.05 fb^{-1} of proton-proton collisions. The analysis was carried out by combining two different approaches to reconstruct the $t\bar{t}$ system. In the resolved approach, the top-quark decay products are individually reconstructed as they are well-separated in the detector. In particular, the hadronic top-quark decay is reconstructed as two or three anti- k_T $R = 0.4$ jets. With increasing transverse momentum of the top quark, its decay products are more collimated. In this situation, the boosted approach is the most efficient, and the hadronic top-quark decay is reconstructed as one anti- k_T $R = 1.0$ jet. The mass and the substructure of this jet have been used to select events satisfying the boosted topology.

Each approach involves its own set of criteria to select $t\bar{t}$ events. From these events, the invariant mass of the $t\bar{t}$ system has been reconstructed. The mass distribution from the selected data was compared with the Standard Model prediction. Good agreement within the uncertainties was found between the data and the expected background.

An overlap channel comprising those events that satisfy both the resolved and the boosted selections has been created. Exclusive categories for each approach were defined by subtracting the overlap channel from the resolved and boosted selected events. Different combinations of the selection categories have been defined in order to improve the sensitivity of searches to $t\bar{t}$ resonances. Each combination was used as input of the statistical analysis based on a Bayesian technique. From this analysis, upper limits were derived on the production cross-section times branching ratio for narrow and wide resonant states, at the 95% confidence level.

8 Summary and conclusions

A significant gain in sensitivity has been achieved by introducing the overlap channel. For the full range of resonance masses, this channel, added to the exclusive information of the boosted approach, provides stronger limits than those obtained when the channel is not defined. Such an improvement was made possible by using the $t\bar{t}$ invariant mass in the overlap channel. Furthermore, it has been proved that the exclusive resolved information has impact on the sensitivity only at low masses. Therefore, a maximum sensitivity in the whole mass range was reached by combining the resolved and boosted exclusive channels, together with the overlap boosted category.

The results obtained from this combination are outlined in the following. The observed (expected) upper limits for the narrow resonance range from 4.85 (4.81) pb at a resonance mass of 0.6 TeV to 0.21 (0.13) pb at 2 TeV. The observed (expected) upper limits on the cross-section determined for the broad resonance range from 2.52 (2.59) pb at 0.7 TeV to 0.37 (0.27) pb at 2 TeV. These results represent a considerable improvement on the ATLAS search for $t\bar{t}$ resonances, using the same dataset. The obtained upper limits enabled the exclusion of the topcolor leptophobic Z' boson with a mass below 1.3 TeV, and the KK-gluon with mass below 1.65 TeV. Both exclusion regions are determined at 95% confidence level.

In summary, an alternative strategy for searching new massive states decaying to $t\bar{t}$ pairs has been developed in this thesis. It consists of creating an overlap channel that relates the two standard search approaches: resolved and boosted. The definition of the overlap channel allows for the characterization of statistically independent categories whose combination improves significantly the sensitivity to $t\bar{t}$ resonances. Given the competence of the obtained results, this strategy is proposed for future $t\bar{t}$ resonances searches.

Bibliography

- [1] F. Halzen & A. Martin. *Quarks & Leptons: An introductory course in modern particle physics*. John Wiley & Sons, USA, 1984.
- [2] M. E. Peskin & D. V. Schroeder. *An introduction to quantum field theory*. Perseus Books, USA, 2004.
- [3] K. Nakamura et al. Particle Data Group. Review of particle physics. *J. Phys.* **G37** (2010) 075021.
- [4] Super-Kamiokande Collaboration. Evidence for oscillation of atmospheric neutrinos. *Phys. Rev. Lett.* **81** (1998) 1562-1567.
- [5] SNO Collaboration. Direct evidence for neutrino flavor transformation from neutral-current interactions in the Sudbury Neutrino Observatory. *Phys. Rev. Lett.* **89** (2002) 011301.
- [6] E. Noether. Invariante Variationsprobleme. *Nachr. v. d. Ges. d. Wiss. zu Göttingen* (1918) 235-257.
- [7] F. Englert & R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.* **13** (1964) 321-323.
- [8] G. Guralnik, C. Hagen & T. Kibble. Global conservation laws and massless particles. *Phys. Rev. Lett.* **13** (1964) 585-587.
- [9] P. W. Higgs. Broken symmetries, massless particles and gauge fields. *Phys. Lett.* **12** (1964) 132-133.
- [10] P. W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.* **13** (1964) 508-509.
- [11] P. W. Higgs. Spontaneous symmetry breakdown without massless bosons. *Phys. Rev.* **145** (1966) 1156-1163.
- [12] W. A. Bardeen et al. Deep inelastic scattering beyond the leading order in asymptotically free gauge theories. *Phys. Rev.* **D18** (1978) 3998.
- [13] S. L. Glashow. Partial symmetries of weak interactions. *Nucl. Phys.* **22** (1961) 579-588.
- [14] S. Weinberg. A Model of Leptons. *Phys. Rev. Lett.* **19** (1967) 1264-1266.

Bibliography

- [15] A. Salam. Gauge unification of fundamental forces. *Rev. Mod. Phys.* **52** (1980) 525-538.
- [16] A. Quadt. Top quark physics at hadron colliders. *Eur. Phys. J.* **C48** (2006) 835-1000.
- [17] J. Goldstone, A. Salam & S. Weinberg. Broken symmetries. *Phys. Rev.* **127** (1962) 965-970.
- [18] ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.* **B716** (2012) 1-29, .
- [19] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.* **B716** (2012) 30.
- [20] H. Yukawa. On the interaction of elementary particles. *Proc. Phys. Math. Soc. Jap.* **17** (1935) 48-57.
- [21] N. Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.* **10** (1963) 531.
- [22] M. Kobayashi & T. Maskawa. CP violation in the renormalizable theory of weak interaction. *Prog. Theor. Phys.* **49** (1973) 652-657.
- [23] A. Höcker et al. A new approach to a global fit of the CKM matrix. *Eur. Phys. J.* **C21** (2001) 225-259.
- [24] CKMfitter Group. CP violation and the CKM matrix: assessing the impact of the asymmetric B factories. *Eur. Phys. J.* **C41** (2005) 1-131.
- [25] T. Kibble. Symmetry breaking in non-Abelian gauge theories. *Phys. Rev.* **155** (1967) 1554-1561.
- [26] K. G. Wilson. Confinement of Quarks. *Phys. Rev.* **D10** (1974) 2445-2459.
- [27] G. Altarelli & G. Parisi. Asymptotic freedom in parton language. *Nucl. Phys.* **B126** (1977) 298.
- [28] E. L. Berger & H. Contopanagos. Threshold resummation of the total cross section for heavy quark production in hadronic collisions. *Phys. Rev.* **D57** (1998) 253-264.
- [29] B. Kersevan & I. Hinchliffe. A consistent prescription for the production involving massive quarks in hadron collisions. *JHEP* **09** (2006) 033.
- [30] V.N. Gribov & L. Lipatov. Deep inelastic ep scattering in perturbation theory. *Sov.J.Nucl.Phys.* **15** (1972) 438-450.
- [31] CDF Collaboration. Observation of top quark production in $\bar{p}p$ collisions. *Phys. Rev. Lett.* **74** (1995) 2626-2631.

- [32] D0 Collaboration. Observation of the top quark. *Phys. Rev. Lett.* **74** (1995) 2632-2637.
- [33] ALEPH Collaboration, DELPHI Collaboration, L3 Collaboration, OPAL Collaboration & LEP Electroweak Working Group. Combined preliminary data on Z parameters from the LEP experiments and constraints on the Standard Model. Technical Report CERN-PPE-94-187-CERN-L3-078, CERN, Geneva, Nov. 1994.
- [34] J. Beringer et al. Particle Data Group. Review of particle physics. *Phys. Rev.* **D86** (2012) 010001.
- [35] MSTW PDFs key plots: LHC Parton Kinematics. URL <http://mstwpdf.hepforge.org/plots/plots.html>.
- [36] A. D. Martin et al. Parton distributions for the LHC. *Eur. Phys. J.* **C63** (2009) 189-285.
- [37] M. Aliev et al. HATHOR: HAdronic Top and Heavy quarks crOss section calculator. *Comput. Phys. Commun.* **182** (2011) 1034-1046.
- [38] N. Kidonakis. Next-to-next-to-leading-order collinear and soft gluon corrections for t -channel single top quark production. *Phys. Rev.* **D83** (2011) 091503.
- [39] N. Kidonakis. NNLL resummation for s -channel single top quark production. *Phys. Rev.* **D81** (2010) 054028.
- [40] H. X. Zhu et al. Factorization and resummation of s -channel single top quark production. *JHEP* **1102** (2011) 099.
- [41] N. Kidonakis. Two-loop soft anomalous dimensions for single top quark associated production with a W - or H -. *Phys. Rev.* **D82** (2010) 054018.
- [42] J. H. Kühn. Theory of top quark production and decay. arXiv:hep-ph/9707321v1.
- [43] Y. Ashie et al. Evidence for an oscillatory signature in atmospheric neutrino oscillations. *Phys. Rev. Lett.* **93** (2004) 101801.
- [44] G Bertone et al. Particle dark matter: Evidence, candidates and constraints. *Phys. Rept.* **405** (2005) 279-390.
- [45] E. Komatsu et al. Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *Astrophys. J. Suppl.* **180** (2009) 330-376.
- [46] R. Frederix & F. Maltoni. Top pair invariant mass distribution: A window on new physics. *JHEP* **0901** (2009) 047.
- [47] J. Pumplin et al. New generation of parton distributions with uncertainties from global QCD analysis. *JHEP*, 07:012, 2002.

Bibliography

- [48] Fabio Maltoni and Tim Stelzer. MadEvent: Automatic event generation with MadGraph. *JHEP*, 0302:027, 2003.
- [49] R. M. Harris et al. Cross-section for topcolor Z' (t) decaying to top anti-top: Version 2.6. arXiv:hep-ph/9911288v1.
- [50] C. T. Hill. Topcolor assisted technicolor. *Phys. Lett.* **B345** (1995) 483-489.
- [51] R. Harris, & J. Supriya. Cross Sections for leptophobic topcolor Z' decaying to top-antitop. *Eur. Phys. J.* **C72** (2012) 2072.
- [52] L. Randall & R. Sundrum. A Large mass hierarchy from a small extra dimension. *Phys. Rev. Lett.* **83** (1999) 3370-3373.
- [53] T. Kaluza. On the problem of unity in physics. *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* **1921** (1921) 966-972.
- [54] O. Klein. Quantum theory and five-dimensional theory of relativity. *Z. Phys.* **37** (1926) 895-906.
- [55] H. Davoudiasl et al. Phenomenology of the Randall-Sundrum gauge hierarchy model. *Phys. Rev. Lett.* **84** (2000) 2080.
- [56] B. Lillie, L. Randall & L. Wang. The bulk RS KK-gluon at the LHC. *JHEP* **0709** (2007) 074.
- [57] K. Agashe et al. LHC signals for coset electroweak gauge bosons in warped/composite PGB Higgs models. *Phys. Rev.* **D81** (2010) 096002.
- [58] ATLAS Collaboration. A Search for $t\bar{t}$ resonances in lepton+jets events with highly boosted top quarks in 2 fb^{-1} of pp collisions at $\sqrt{s} = 7\text{ TeV}$. Technical Report ATL-COM-PHYS-2011-1692, CERN, Geneva, Jun. 2012.
- [59] ATLAS Collaboration. A search for $t\bar{t}$ resonances with the ATLAS detector in 2.05 fb^{-1} of proton-proton collisions at $\sqrt{s} = 7\text{ TeV}$. *Eur. Phys. J.* **C72** (2012) 2083.
- [60] A. Djouadi, G. Moreau & R. K. Singh. Kaluza-Klein excitations of gauge bosons at the LHC. *Nucl. Phys.* **B797** (2008) 1-26.
- [61] M. A. Dobbs et al. Les Houches guidebook to Monte Carlo generators for hadron collider physics. 2004.
- [62] Andy Buckley, Butterworth, et al. General-purpose event generators for LHC physics. *Phys.Rept.*, 504:145–233, 2011.
- [63] M.R. Whalley, D. Bourilkov, and R.C. Group. The Les Houches accord PDFs (LHAPDF) and LHAGLUE. 2005.

- [64] A.A. Markov. An example of statistical investigation of the text Eugene Onegin concerning the connection of samples in chains. *Science in Context*, 19(4):591–600, 2006.
- [65] J. Alwall et al. Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *Eur. Phys. J.*, C53:473–500, 2008.
- [66] T. Sjostrand, S. Mrenna & P. Z. Skands. PYTHIA 6.4 physics and manual. *JHEP* **05** (2006) 026.
- [67] Bo Andersson, G. Gustafson, and B. Soderberg. A PROBABILITY MEASURE ON PARTON AND STRING STATES. *Nucl.Phys.*, B264:29, 1986.
- [68] Bryan R Webber. A qcd model for jet fragmentation including soft gluon interference. oai:cds.cern.ch:147786. *Nucl. Phys. B*, 238(CERN-TH-3713):492–528. 53 p, Sep 1983.
- [69] T. Sjostrand and Peter Z. Skands. Transverse-momentum-ordered showers and interleaved multiple interactions. *Eur.Phys.J.*, C39:129–154, 2005.
- [70] G. Corcella et al. HERWIG 6.5: an event generator for Hadron Emission Reactions With Interfering Gluons (including supersymmetric processes). *JHEP*, 01:010, 2001.
- [71] J. M. Butterworth, Jeffrey R. Forshaw, and M. H. Seymour. Multiparton interactions in photoproduction at HERA. *Z. Phys.*, C72:637–646, 1996.
- [72] M. L. Mangano et al. ALPGEN, a generator for hard multiparton processes in hadronic collisions. *JHEP* **07** (2003) 001.
- [73] S. Frixione & B.R.Webber. Matching NLO QCD computations and parton shower simulations. *JHEP* **06** (2002) 029.
- [74] S.Frixione, P. Nason & B.R. Webber. Matching NLO QCD computations and parton shower in heavy flavor production. *JHEP* **08** (2003) 007.
- [75] Paolo Nason. A New method for combining NLO QCD with shower Monte Carlo algorithms. *JHEP*, 0411:040, 2004.
- [76] Borut Paul Kersevan and Elzbieta Richter-Was. The Monte Carlo event generator AcerMC version 2.0 with interfaces to PYTHIA 6.2 and HERWIG 6.5. 2004.
- [77] Brüning et al. LHC design report. Vol. I: The LHC main ring. 2004. CERN-2004-003-V-1.
- [78] J. Evans. LHC Machine. *JINST*, 3:S08001, 2008.
- [79] CERN. LEP Design Report: Vol. 2. The LEP Main Ring. 1984. CERN-LEP-84-01.

Bibliography

- [80] A. D. Martin et al. Parton Distributions for the LHC., 2006. URL <http://cdsmedia.cern.ch/collection/CERN%20PhotoLab?ln=en>.
- [81] K. Schindl. The PS booster as preinjector for LHC. *Part.Accel.*, 58:63–78, 1997.
- [82] P. Collier, B. Goddard, R. Jung, K.H. Kissler, T. Linnecar, et al. The SPS as injector for LHC: Conceptual design. 1997.
- [83] ATLAS Collaboration. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST* **3** (2008) S08003.
- [84] R. Adolphi et al. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.
- [85] A. Augusto Alves et al. The LHCb Detector at the LHC. *JINST*, 3:S08005, 2008.
- [86] K. Aamodt et al. The ALICE experiment at the CERN LHC. *JINST*, 3:S08002, 2008.
- [87] M.S. Alam et al. ATLAS pixel detector: Technical design report. 1998.
- [88] G. Aad, M. Ackers, F.A. Alberti, M. Aleppo, G. Alimonti, et al. ATLAS pixel detector electronics and sensors. *JINST*, 3:P07007, 2008.
- [89] P. Jenni et al. *ATLAS Forward Detectors for Measurement of Elastic Scattering and Luminosity*. Number CERN-LHCC-2008-004 in Technical Design Report. CERN, Geneva, 2008.
- [90] V. Cindro, D. Dobos, I. Dolenc, H. Fraiss-Kolbl, H. Fraiss-Koelbl, et al. The ATLAS beam conditions monitor. *JINST*, 3:P02004, 2008.
- [91] ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector at the LHC. *Eur. Phys. J.* **C71** (2011) 1630.
- [92] ATLAS Collaboration. Updated luminosity determination in pp collisions at $\sqrt{s}=7$ tev using the atlas detector. Technical Report ATLAS-CONF-2011-011, CERN, Geneva, Mar 2011.
- [93] ATLAS Luminosity Working Group. URL <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResults>.
- [94] ATLAS Collaboration. Luminosity Determination in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS Detector in 2011. Technical Report ATLAS-CONF-2011-116, CERN, Aug 2011.
- [95] S van der Meer. Calibration of the effective beam height in the ISR. Technical Report CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, 1968.

- [96] ATLAS Collaboration. Alignment of the ATLAS Inner Detector Tracking System with 2010 LHC proton-proton collisions at $\sqrt{s} = 7$ TeV. Technical Report ATLAS-CONF-2011-012, CERN, Geneva, Mar. 2011.
- [97] ATLAS Collaboration. Performance of the ATLAS Trigger System in 2010. Eur. Phys. J. **C72** (2012) 1849, .
- [98] G. Barrand et al. GAUDI - A software architecture and framework for building HEP data processing applications. Comput. Phys. Commun. **140** (2001) 45.
- [99] S. Agostinelli et al. GEANT4: A simulation toolkit. Nucl. Instrum. Meth. **A506** (2003) 250-303.
- [100] ATLAS Collaboration. Performance of the ATLAS Inner Detector track and vertex Reconstruction in the high pile-up LHC environment. Technical Report ATLAS-CONF-2012-042, CERN, Geneva, Mar. 2012.
- [101] ATLAS Collaboration. Electron performance measurements with the ATLAS detector using the 2010 LHC proton-proton collision data. Eur. Phys. J. **C72** (2012) 1909, .
- [102] ATLAS Collaboration. $pp \rightarrow \gamma/Z \rightarrow \mu^+\mu^-$ and $pp \rightarrow W \rightarrow \mu\nu$ inclusive cross-section measurement in early data with the ATLAS detector. Technical Report ATL-COM-PHYS-2010-124, CERN, Geneva, Mar. 2010.
- [103] ATLAS Collaboration. Lepton trigger and identification for the Winter 2011 top quark analyses. Technical Report ATL-COM-PHYS-2011-123, CERN, Geneva, Feb. 2011.
- [104] ATLAS Collaboration. Performance studies for e/γ calorimeter isolation. Technical Report ATL-COM-PHYS-2011-1186, CERN, Geneva, Sep. 2011.
- [105] ATLAS Collaboration. Expected electron performance in the ATLAS experiment. Technical Report ATL-PHYS-PUB-2011-006, CERN, Geneva, Apr. 2011.
- [106] ATLAS Collaboration. Electron Gamma Performance Group, Feb. 2012. URL <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ElectronGammaPublicCollisionResults>.
- [107] Th. Lagouri et al. A muon identification and combined reconstruction procedure for the ATLAS detector at the LHC at CERN. IEEE Trans. Nucl. Sci. **51** (2004) 3030-3033.
- [108] ATLAS Collaboration. Identification of muon candidates in pp collisions at $\sqrt{s} = 900$ GeV with the ATLAS detector. Technical Report ATLAS-CONF-2010-015, CERN, Geneva, Jul. 2010.

Bibliography

- [109] ATLAS Collaboration. Muon reconstruction efficiency in reprocessed 2010 LHC proton-proton collision data recorded with the ATLAS detector. Technical Report ATLAS-CONF-2011-063, CERN, Geneva, Apr. 2011.
- [110] ATLAS Collaboration. Muon Performance Group, Apr. 2012. URL <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/MuonPerformancePublicPlots>.
- [111] ATLAS Collaboration. A measurement of the ATLAS muon reconstruction and trigger efficiency using J/ψ decays. Technical Report ATLAS-CONF-2011-021, CERN, Geneva, Mar. 2011.
- [112] ATLAS Collaboration. Muon momentum resolution in first pass reconstruction of pp collision data recorded by ATLAS in 2010. Technical Report ATLAS-CONF-2011-046, CERN, Geneva, Mar. 2011.
- [113] G. P. Salam & G. Soyez. A practical seedless infrared-safe cone jet algorithm. *JHEP* **05** (2007) 086.
- [114] Y. L. Dokshitzer et al. Better jet clustering algorithms. *JHEP* **08** (1997) 001.
- [115] M. Cacciari & G. P. Salam. Dispelling the N^3 myth for the k_T jet-finder. *Phys. Lett.* **B641** (2006) 57-61.
- [116] M. Cacciari, G. P. Salam & G. Soyez. The anti- k_T jet clustering algorithm. *JHEP* **0804** (2008) 063, .
- [117] S. Catani et al. The K^- perpendicular clustering algorithm for jets in deep inelastic scattering and hadron collisions. *Phys. Lett.* **B285** (1992) 291-299.
- [118] K. Agashe et al. LHC signals from warped extra dimensions. *Phys. Lett.* **D77** (2008) 015003.
- [119] J. M. Butterworth. Reconstructing sparticle mass spectra using hadronic decays. *JHEP* **0705** (2007) 033.
- [120] J. Thaler & L. T. Wang. Strategies to identify boosted tops. *JHEP* **07** (2008) 092.
- [121] U. Baur & L. H. Orr. High p_T top quarks at the CERN Large Hadron Collider. *Phys. Rev.* **D76** (2007) 094012.
- [122] ATLAS Collaboration. Performance of jet algorithms in the ATLAS detector. Technical report, CERN, Geneva, Dec. 2009.
- [123] ATLAS Collaboration. Calorimeter clustering algorithms: description and performance. Technical Report ATL-COM-LARG-2008-003, CERN, Geneva, Apr. 2008.
- [124] M. Cacciari, G. P. Salam & G. Soyez. FastJet user manual. *Eur. Phys. J.* **C72** (2012) 1896, .

- [125] ATLAS Collaboration. Performance of the electron and photon trigger in pp collisions at $\sqrt{s} = 7$ TeV. Technical Report ATLAS-CONF-2011-114, CERN, Geneva, Aug. 2011.
- [126] ATLAS Collaboration. Jet energy measurement with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J.* **C73** (2013) 2304, .
- [127] C. Issever, K. Borrás & D. Wegener. An improved weighting algorithm to achieve software compensation in a fine grained LAr calorimeter. *Nucl. Instrum. Meth.* **A545** (2005) 803.
- [128] ATLAS Collaboration. Local hadronic calibration. Technical Report ATL-LARG-PUB-2009-001, CERN, Geneva, Jan. 2009.
- [129] ATLAS Collaboration. Jet energy resolution and selection efficiency relative to track jets from in-situ techniques with the ATLAS Detector using proton-proton collisions, at a center of mass energy $\sqrt{s} = 7$ TeV. Technical Report ATLAS-CONF-2010-054, CERN, Geneva, Jul. 2010.
- [130] ATLAS Collaboration. Data-Quality requirements and event cleaning for jets and missing transverse energy reconstruction with the ATLAS detector in proton-proton collisions at a center-of-mass energy of $\sqrt{s} = 7$ TeV. Technical Report ATLAS-CONF-2010-038, CERN, Geneva, Jul. 2010.
- [131] ATLAS Collaboration. How to clean jets, Nov. 2011. URL <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/HowToCleanJets>.
- [132] ATLAS Collaboration. Expected performance of the ATLAS experiment - detector, trigger and physics, . arXiv:0901.0512v4 [hep-ex].
- [133] M. Lehmacher. b -tagging algorithms and their performance at ATLAS. arXiv:0809.4896v3 [hep-ex].
- [134] G. Piacquadio & C. Weiser. An improved weighting algorithm to achieve software compensation in a fine grained LAr calorimeter. *J. Phys. Conf. Ser.* **119** (2009) 032032.
- [135] ATLAS Collaboration. Commissioning of the ATLAS high-performance b -tagging algorithms in the 7 TeV collision data. Technical Report ATLAS-CONF-2011-102, CERN, Geneva, Jul. 2011.
- [136] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME-Journal of Basic Engineering.* **D82** (1960) 35-45.
- [137] ATLAS Collaboration. Calibrating the b -tag efficiency and mistag rate in 35 pb^{-1} of data with the ATLAS Detector. Technical Report ATLAS-CONF-2011-089, CERN, Geneva, Jul. 2011.

Bibliography

- [138] ATLAS Collaboration. Performance of missing transverse momentum reconstruction in proton-proton collisions at 7 TeV with ATLAS. *Eur. Phys. J.* **C72** (2012) 1844, .
- [139] ATLAS Collaboration. Missing transverse energy for top analyses in release 16.6.5.5.1 with the 2011 dataset. Technical Report ATL-PHYS-INT-2011-092, CERN, Geneva, Nov. 2011.
- [140] T. Chwalek. Ph.D. Thesis (pp. 81-85). KIT Karlsruhe, Feb. 2010. IEKP-KA/2010-5.
- [141] ATLAS Collaboration. ATLAS level-1 trigger: Technical Design Report. Technical Report CERN-LHCC-98-014, CERN, Geneva, Nov. 1998.
- [142] ATLAS Collaboration. Inclusive jet trigger efficiencies for the early 2011 data. Technical Report ATL-COM-DAQ-2011-031, CERN, Geneva, May. 2011.
- [143] ATLAS Collaboration. Performance of the ATLAS Jet Trigger in the early $\sqrt{s} = 7$ TeV data. Technical Report ATLAS-CONF-2010-094, CERN, Geneva, Oct. 2010.
- [144] M. Paterno. Calculating efficiencies and their uncertainties. FERMILAB-TM-2286-CD, Fermilab, May. 2004.
- [145] ATLAS Collaboration. Measurement of multi-jet cross-sections in proton-proton collisions at 7 TeV center-of-mass energy with the ATLAS detector. Technical Report ATLAS-CONF-2010-084, CERN, Geneva, Oct. 2010.
- [146] ATLAS Collaboration. Measurement of multi-jet cross-sections in proton-proton collisions at 7 TeV center-of-mass energy. Technical Report ATLAS-CONF-2011-043, CERN, Geneva, Mar. 2011.
- [147] ATLAS Collaboration. Measurement of multi-jet cross-sections in proton-proton collisions at 7 TeV center-of-mass energy, updated. Technical Report ATL-PHYS-INT-2012-029, CERN, Geneva, Apr. 2012.
- [148] ATLAS Collaboration. Measurement of multi-jet cross-sections in proton-proton collisions at 7 TeV center-of-mass energy. *Eur. Phys. J.* **C71** (2011) 1763, .
- [149] A. Sherstnev and R.S. Thorne. Parton Distributions for LO Generators. *Eur.Phys.J.*, C55:553–575, 2008.
- [150] ATLAS Collaboration. Top working group, Nov. 2012. URL <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopWorkingGroup>.
- [151] ATLAS Collaboration. Performance of large-R jets and jet substructure reconstruction with the ATLAS detector. Technical Report ATLAS-CONF-2012-065, CERN, Geneva, Jul. 2012.

- [152] ATLAS Collaboration. Prospects for top anti-top resonance searches using early ATLAS data. Technical Report ATL-PHYS-PUB-2010-008, CERN, Geneva, Jul. 2010.
- [153] T. Plehn et al. Stop reconstruction with tagged tops. *JHEP* **10** (2010) 078.
- [154] ATLAS Collaboration. A search for $t\bar{t}$ resonances in the lepton plus jets final state with ATLAS using 4.7 fb^{-1} of pp collisions at $\sqrt{s} = 7\text{ TeV}$. *Phys. Rev.* **D88** (2013) 012004.
- [155] ATLAS Collaboration. A Search for $t\bar{t}$ resonances in the lepton plus jets channel in 200 pb^{-1} of pp collisions at $\sqrt{s} = 7\text{ TeV}$. Technical Report ATLAS-CONF-2011-087, CERN, Geneva, Jun. 2011.
- [156] ATLAS Collaboration. A Search for $t\bar{t}$ resonances in the lepton plus jets channel in 2.05 fb^{-1} of pp collisions at $\sqrt{s} = 7\text{ TeV}$. Technical Report ATLAS-CONF-2012-029, CERN, Geneva, Mar. 2012.
- [157] ATLAS Collaboration. Study of jets produced in association with a W boson in pp collisions at $\sqrt{s} = 7\text{ TeV}$ with the ATLAS detector. *Phys. Rev.* **D85** (2012) 092002.
- [158] ATLAS Collaboration. A search for $t\bar{t}$ resonances in the lepton plus jets final state using 4.66 fb^{-1} of pp collisions at $\sqrt{s} = 7\text{ TeV}$. Technical Report ATLAS-CONF-2012-136, CERN, Geneva, Sep. 2012.
- [159] ATLAS Collaboration. Mis-identified lepton backgrounds in top quark pair production studies for EPS 2011 analyses. Technical Report ATL-COM-PHYS-2011-768, CERN, Geneva, Jun. 2011.
- [160] P. M. Nadolsky et al. Implications of CTEQ global analysis for collider observables. *Phys. Rev.* **D78** (2008) 013004.
- [161] NNPDF Collaboration. A first unbiased global NLO determination of parton distributions and their uncertainties. *Nucl. Phys.* **B838** (2010) 136-206.
- [162] ATLAS Collaboration. Top PDF uncertainty, Aug. 2012. URL <https://twiki.cern.ch/twiki/bin/viewauth/AtlasProtected/TopPdfUncertainty>.
- [163] ATLAS Collaboration. Jet energy scale and its systematic uncertainty in proton-proton collisions at $\sqrt{s} = 7\text{ TeV}$ in ATLAS 2010 data. Technical Report ATLAS-CONF-2011-032, CERN, Geneva, Mar. 2011.
- [164] ATLAS Collaboration. Light-quark and gluon jets: calorimeter response, jet energy scale systematics and properties. Technical Report ATLAS-CONF-2012-138, CERN, Geneva, Sep. 2012.

Bibliography

- [165] ATLAS Collaboration. Close-by jet Effects on jet energy scale calibration in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS Detector. Technical Report ATLAS-CONF-2011-062, CERN, Geneva, Apr. 2011.
- [166] ATLAS Collaboration. In-situ jet energy scale and jet shape corrections for multiple interactions in the first ATLAS data at the LHC. Technical Report ATLAS-CONF-2011-030, CERN, Geneva, Mar. 2011.
- [167] ATLAS Collaboration. Jet energy resolution from in-situ techniques with the ATLAS detector using proton-proton collisions, at a center of mass energy $\sqrt{s} = 7$ TeV. Technical Report ATL-COM-PHYS-2011-240, CERN, Geneva, Mar. 2011.
- [168] ATLAS Collaboration. Jet mass and substructure of inclusive jets in $\sqrt{s} = 7$ TeV pp collisions with the ATLAS experiment. *JHEP* **05** (2012) 128.
- [169] P. Z. Skands. Tuning Monte Carlo generators: the PERUGIA tunes. *Phys. Rev.* **D82** (2010) 074018.
- [170] ATLAS Collaboration. Study of jet shapes in inclusive jet production in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector. *Phys. Rev.* **D83** (2011) 052003.
- [171] M. Bahr et. al. HERWIG++ Physics and manual. *Eur. Phys. J.* **C58** (2008) 639.
- [172] T. Gleisberg et al. Event generation with SHERPA 1.1. *JHEP* **02** (2009) 007.
- [173] ATLAS Collaboration. Measurement of the top quark-pair production cross-section with ATLAS in pp collisions at $\sqrt{s} = 7$ TeV. *Eur. Phys. J.* **C71** (2011) 1577.
- [174] J. M. Butterworth et al. Single boson and diboson production cross-sections in pp collisions at $\sqrt{s} = 7$ TeV. Technical Report ATL-COM-PHYS-2010-695, CERN, Geneva, Oct. 2010.
- [175] S. Moch & P. Uwer. Theoretical status and prospects for top-quark pair production at hadron colliders. *Phys. Rev.* **D78** (2008) 034003.
- [176] F. Demartin et al. The impact of PDF and α_s uncertainties on Higgs production in gluon fusion at hadron colliders. *Phys. Rev.* **D82** (2010) 014002.
- [177] ATLAS Collaboration. Technical Report ATL-COM-PHYS-2011-768, CERN, Geneva, Oct. 2010.
- [178] J. S. Miller. Top and W mass measurements at the Tevatron. SSI04-THT001, SLAC, Aug. 2004.
- [179] ATLAS Collaboration. A search for $t\bar{t}$ resonances in lepton+jets events with highly boosted top quarks collected in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector. *JHEP* **09** (2012) 041.

- [180] CDF Collaboration. A search for resonant production of $t\bar{t}$ pairs in 4.8 fb^{-1} of integrated luminosity of $p\bar{p}$ collisions at $\sqrt{s} = 1.96 \text{ TeV}$. *Phys. Rev.* **D84** (2011) 072004.
- [181] CDF Collaboration. A search for resonant production of $t\bar{t}$ decaying to jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.96 \text{ TeV}$. *Phys. Rev.* **D84** (2011) 072003.
- [182] CDF Collaboration. A search for resonant $t\bar{t}$ production in the semi-leptonic decay mode using the full CDF data set. *Phys. Rev. Lett.* **110** (2013) 121802.
- [183] DØ Collaboration. Search for a narrow $t\bar{t}$ resonance in $p\bar{p}$ collisions at $\sqrt{s} = 1.96 \text{ TeV}$. *Phys. Rev.* **D85** (2012) 051101.
- [184] CMS Collaboration. Search for anomalous $t\bar{t}$ production in the highly-boosted all-hadronic final state. *JHEP* **09** (2012) 029.
- [185] CMS Collaboration. Search for resonant $t\bar{t}$ production in lepton+jets events in pp collisions at $\sqrt{s} = 7 \text{ TeV}$. *JHEP* **12** (2012) 015.
- [186] CMS Collaboration. Search for Z' resonances decaying to $t\bar{t}$ in dilepton+jets final states in pp collisions at $\sqrt{s} = 7 \text{ TeV}$. *Phys. Rev.* **D87** (2013) 072002.
- [187] DØ Collaboration. A recipe for the construction of confidence limits. FERMILAB-TM-2104, Fermilab, Apr. 2000.
- [188] J. O. Berger, J. M. Bernardo & D. Sun. The formal definition of reference priors. *Ann. Statist.* **37** (2009) 905-938.
- [189] D Casadei. Reference analysis of the signal + background model in counting experiments. *JINST.* **7** (2012) 01012.
- [190] DØ Collaboration. Statistical methods implemented in the package top statistics. Technical report, Feb. 2011.
- [191] T. Lenz. Private communication. 14 Oct. 2013.
- [192] CMS Collaboration. A search for resonances in semileptonic top pair production at $\sqrt{s} = 7 \text{ TeV}$. Technical Report CMS-PAS-TOP-11-009, CERN, Geneva, Mar. 2012.

List of Figures

2.1	Feynman diagrams at leading order contributing to $t\bar{t}$ production.	11
2.2	MSTW2008 next-to-leading order Parton Distribution Functions.	12
2.3	Feynman diagrams of the leading order processes for single top production.	12
2.4	Topology of the semi-leptonic top quark pair decay.	14
2.5	Resonances in the $t\bar{t}$ invariant mass spectrum.	16
2.6	Steps of Monte Carlo event generation.	19
3.1	The CERN accelerator complex.	26
3.2	The ATLAS detector.	27
3.3	The magnetic system of the ATLAS detector.	30
3.4	Cut-away view of the ATLAS Inner Detector.	31
3.5	Layers-structure of the ATLAS Inner Detector.	33
3.6	The ATLAS calorimeter system.	33
3.7	The three LAr sampling layers of the ATLAS electromagnetic calorimeter.	35
3.8	Cut-away view of the ATLAS muon system.	36
3.9	Luminosity in 2010 and 2011.	39
3.10	Mean number of interactions per bunch crossing.	40
3.11	LHC relevant cross-sections as a function of center-of-mass energy.	41
4.1	Identification efficiencies of tight electrons.	50
4.2	Electron energy calibration distributions.	52
4.3	Muons reconstruction efficiencies.	53
4.4	EM+JES jet calibration.	58
4.5	Jet energy resolution.	59
4.6	Jet reconstruction efficiency.	61
4.7	Fraction of jets tagged by the <code>JetFitterComNN</code> algorithm using $t\bar{t}$ events.	63
5.1	Jet windows for the level-1 trigger algorithms.	68
5.2	The local maximum test for level-1 jet finding.	68
5.3	Efficiency of the L1_J5 trigger as a function of the offline leading jet p_T	73
5.4	Efficiency of the L1_J10 trigger as a function of the offline reconstructed leading jet p_T	73
5.5	Efficiency of the L1_J10 trigger as a function of the offline reconstructed leading jet rapidity.	74
5.6	Efficiency of the L1_J15 trigger as a function of the offline reconstructed leading jet p_T and rapidity.	75

List of Figures

5.7	Efficiency of the L1_J30 trigger as a function of the offline reconstructed leading jet p_T and rapidity.	76
5.8	Efficiency of the L1_2J10 trigger as a function of the reconstructed second leading jet p_T	77
5.9	Efficiency of the L1_3J10 trigger as a function of the second leading jet p_T	77
5.10	Multi-jet trigger efficiencies as a function of the second leading jet p_T	78
5.11	H_T distributions for events that satisfy the JE trigger thresholds.	79
5.12	Efficiencies of the total transverse jet energy triggers.	80
6.1	Representation of the jets in the resolved $t\bar{t}$ reconstruction.	88
6.2	Merging probability and jet mass distributions for partons resulting from the hadronically decaying top quark.	89
6.3	Configuration of partons in the boosted reconstruction.	90
6.4	Discriminant variables in the boosted approach of the top-quark reconstruction.	91
6.5	Jet substructure variables and jet energy-sharing in the boosted approach of the top quark reconstruction.	91
6.6	Selection efficiencies for the resolved and boosted approaches of the top quark reconstruction.	93
6.7	Correlation between the angular separation to the closest in the resolved approach of the $t\bar{t}$ reconstruction.	95
6.8	Reconstructed invariant mass of the $t\bar{t}$ system and mass resolution in the resolved approach.	96
6.9	Mass distribution of the hadronic top quark in the boosted approach of the $t\bar{t}$ reconstruction.	97
6.10	Leptonic top quark mass distribution in the boosted approach.	98
6.11	Invariant mass distribution of the $t\bar{t}$ system in the boosted approach.	98
6.12	Reconstructed invariant mass of the $t\bar{t}$ system and mass resolution in the boosted approach.	99
6.13	Jet energy scale systematic uncertainty for anti- k_T $R = 0.4$ jets.	109
6.14	Ratios of jet mass and first k_t -splitting scale defined to estimate systematic uncertainties on anti- k_T $R = 1.0$ jets.	110
6.15	Comparison of the observed data and expected background for distributions of final-state objects in the resolved selection.	118
6.16	Comparison of the observed data and expected background for distributions related to the leptonic top jet in the boosted approach.	119
6.17	Comparison of the observed data and expected background distributions for the W transverse mass in the resolved and boosted selections.	119
6.18	Comparison of the observed data and expected background for distributions related to the boosted reconstruction.	120
6.19	Invariant mass of the reconstructed $t\bar{t}$ system after signal selection in the resolved and boosted approaches.	121
7.1	Upper limits distribution and posterior probability density function.	127

7.2 Invariant mass of the reconstructed $t\bar{t}$ system for all events in the inclusive, exclusive and overlap categories of the boosted reconstruction. 129

7.3 Invariant mass of the reconstructed $t\bar{t}$ system for all events in the exclusive and overlap categories of the resolved reconstruction. 130

7.4 Upper limits on the production cross-section times $t\bar{t}$ branching ratio, for the benchmark signal models in the inclusive boosted category. 131

7.5 Upper limits on the production cross-section times $t\bar{t}$ branching ratio, for the benchmark signal models, combining the exclusive boosted and overlap boosted categories. 132

7.6 Upper limits on the production cross-section times $t\bar{t}$ branching ratio for the benchmark signal models, combining the exclusive boosted, overlap boosted and exclusive resolved categories. 133

7.7 Upper limits on the production cross-section times $t\bar{t}$ branching ratio for the benchmark signal models, combining the exclusive boosted, overlap resolved and exclusive resolved categories. 134

7.8 Observed and expected upper limits on the production cross-section times $t\bar{t}$ branching ratio for the narrow and wide resonant states with a mass of $m = 0.8$ TeV, for distinct combinations of selection categories. 135

7.9 Observed and expected upper limits on the production cross-section times $t\bar{t}$ branching ratio for the narrow and wide resonant states with a mass of $m = 1.0$ TeV, for distinct combinations of selection categories. 136

7.10 Observed and expected upper limits on the production cross-section times $t\bar{t}$ branching ratio for the narrow and wide resonant states with a mass of $m = 1.3$ TeV, for distinct combinations of selection categories. 136

7.11 Observed and expected upper limits on the production cross-section times $t\bar{t}$ branching ratio for the narrow and wide resonant states with a mass of $m = 1.6$ TeV, for distinct combinations of selection categories. 137

7.12 Observed and expected upper limits on the production cross-section times $t\bar{t}$ branching ratio for the narrow and wide resonant states with a mass of $m = 2.0$ TeV, for distinct combinations of selection categories. 137

List of Tables

2.1	Fermions of the Standard Model and their properties.	4
2.2	The electroweak fields and their quantum numbers.	6
2.3	Branching ratios for the W boson decays.	14
3.1	Requirements for the ATLAS detector components in order to reach the physics goals.	29
4.1	Definition of variables used for loose , medium and tight electron identification cuts.	49
5.1	Data-taking periods in 2010 running.	69
5.2	Level-1 jet triggers in 2010.	70
6.1	Leading order cross-sections for W +jets sub-samples.	84
6.2	Leading order cross-sections for Z +jets sub-samples.	84
6.3	Production cross-section for the resonant signal processes.	85
6.4	Triggers used in the $t\bar{t}$ resonances analysis during 2011.	86
6.5	Summary of the object selection cuts for the resolved and boosted approaches of the $t\bar{t}$ reconstruction.	93
6.6	Selection efficiencies in the resolved, overlap and boosted channels of the $t\bar{t}$ reconstruction.	94
6.7	The full width at half maximum and the position of the maximum for different resonance masses.	96
6.8	b -jet energy scale uncertainty.	108
6.9	Systematic uncertainties associated with anti- k_T $R = 1.0$ jets.	111
6.10	Systematic uncertainties and their impact on the background and signal yields.	115
6.11	Number of observed and expected events after applying the resolved and boosted selections.	116
6.12	Expected number of signal events for the two benchmark models.	117
7.1	Upper limits on the production cross-section times $t\bar{t}$ branching ratio for the benchmark signal models in the inclusive boosted category.	131
7.2	Upper limits on the production cross-section times $t\bar{t}$ branching ratio for the benchmark signal models, combining the exclusive boosted and overlap boosted categories.	132

List of Tables

7.3	Upper limits on the production cross-section times $t\bar{t}$ branching ratio for the benchmark signal models, combining the exclusive boosted, overlap boosted and exclusive resolved categories.	133
7.4	Upper limits on the production cross-section times $t\bar{t}$ branching ratio for the benchmark signal models, combining the exclusive boosted, overlap resolved and exclusive resolved categories.	134
7.5	Observed and expected upper limits in the published resolved analysis. . .	139