MEASUREMENT OF THE TOP QUARK PAIR PRODUCTION

CROSS SECTION WITH ATLAS DETECTOR

By

BABAK ABI

Bachelor of Science in Electrical Engineering
K.N.Toosi University of Technology
Tehran, Iran
1995

Master of Science in Physics
Shahid Beheshti University
Tehran, Iran
1998

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
December, 2012

MEASUREMENT OF THE TOP QUARK PAIR PRODUCTION

CROSS SECTION WITH ATLAS DETECTOR

Dissertation Approved:

Dr. Flera Rizatdinova

Dissertation Adviser

Dr. Satyanarayan Nandi

Dr. Alexander Khanov

Dr. Eric Benton

Dr. Joe Cecil

Acknowledgments

First and foremost, I must thank and express my sincere appreciation to my advisor Professor Flera Rizatdinova as she is a really kind and nice supervisor who became my first guide in High Energy Physics and gave me an opportunity to work in the ATLAS experiment with the OSU High Energy Physics group. I would never complete this dissertation without her support.

I would like to especially thank Professor Alexander Khanov to whom I am indebted. He was always there to help me patiently, carry out a profound scientific discussion, and offer a hand with his incredible skills in programming. I would also like to acknowledge my fellow analysis colleague Dr. M. Saleem, who taught me the basics of ATLAS analysis and discussed and worked with me every day for almost three years.

Acknowledgements reflect the views of the author and are not endorsed by committee members or Oklahoma State University.

Name: Babak Abi

Date of Degree: December, 2012

Title of Study: MEASUREMENT OF THE TOP QUARK PAIR PRODUCTION CROSS SECTION WITH ATLAS DETECTOR

Major Field: Physics

Abstract: This thesis describes two measurement of the $t\bar{t}$ production cross-section, the inclusive $\sigma_{t\bar{t}}$ and $\sigma_{t\bar{t}j}$ in association with at least one jet in lepton + jets channel with ATLAS detector. The measurements employ multivariate technique with a binned maximum likelihood template fit. Several kinematic variables of the reconstructed collision events are selected to discriminate $t\bar{t}$ signal events from the various background sources.

The inclusive $\sigma_{t\bar{t}}$ measurement uses approximately 35.3pb^{-1} of ATLAS data collected in 2010 and selecting events with one isolated high-p_T electron or muon, large missing transverse energy and two jet bins of 3 or at least 4 high-p_T jets. The measured $t\bar{t}$ cross section assuming the top quark mass of 172.5 GeV yields :

$$\sigma_{t\bar{t}} = 173 \pm 17(\text{stat.})^{+20}_{-17}(\text{syst.}) \pm 6(\text{lumi.})\,\text{pb}$$

The measured value is in a good agreement with the Standard Model prediction

The second measurement is performed on ATLAS data collected in 2011 corresponding to an integrated luminosity of 4.7 fb^{-1}. The same analysis technique and event selection employed as the first measurement but in 4th and 5th jet bins. The cross-section for $\sigma_{t\bar{t}}$ with at least one additional jet is found to be :

$$\sigma_{t\bar{t}j} = 102 \pm 2(\text{stat.}) \pm 6(\text{ISR/FSR})^{+22}_{-25}(\text{syst.})\,\text{pb}$$

Finally the results of study on PiN diodes radiation hardness are presented. Results were performed with several irradiation beam to investigate the proper optical components for ATLAS Pixel upgrade.

Contents

List of Tables

List of Figures

Chapter 1

Top quark Physics at the Large Hadron Collider

The top quark has been discovered in 1995 by the D0 and CDF experiments at the Tevatron.[3] Since that time, it remains a central part of physics programs at hadron colliders. It is the only known quark with the mass at the electroweak scale. Unlike other quarks, the top quark has very short lifetime and decays before it hadronizes. Because of its unique properties, the top quark is believed to play a special role in the Standard Model (SM). With the start of the Large Hadron Collider (LHC) at CERN, the top quark studies became an important part of the LHC experimental physics program. This is motivated by the fact that due to high beam energy and huge collision rate, the LHC is a top factory, with the top quark production rate orders higher than that at the Tevatron.

The interest in the top quark physics at the LHC is two-fold. On one hand, these studies promote the state-of-the-art theoretical calculations. These calculations are important for searches for new physics, since they allow to evaluate the SM backgrounds with high accuracy. On the other hand, studies of the top quark production and properties can reveal signs of new physics beyond the SM.

Family	Subfamily	Name	Symbol	Mass	Spin	Electric Charge
Fermions	Leptons	electron	e	0.511 MeV	1/2	-1
		muon	μ	105.7 MeV	1/2	-1
		tau	τ	1.717 GeV	1/2	-1
		electron neutrino	ν_e	< 2 eV	1/2	0
		muon neutrino	ν_μ	< 0.17 MeV	1/2	0
		tau neutrino	ν_τ	< 15.5 MeV	1/2	0
	Quarks	up	u	1.5 to 3.3 MeV	1/2	2/3
		charm	c	1.27 GeV	1/2	2/3
		top	t	172.2 GeV	1/2	2/3
		down	d	3.5 to 6.4 MeV	1/2	-1/2
		strange	s	104 MeV	1/2	-1/2
		bottom	b	4.20 GeV	1/2	-1/2
Bosons		photon	γ	$< 1 \cdot 10^{-18}$ eV	1	0
		Z	Z	91.188 MeV	1	0
		W	W	80.398 GeV	1	±1
		gluon	g	0 (< MeV)	1	$SU(3)_C$
		Higgs	H	> 114.4 GeV	1	0

Table 1.1: Properties of the fundamental Standard Model particles

1.1 Standard Model

The Standard Model describes the fundamental particles and their interactions. It consists of the electroweak theory which combines the electromagnetic and weak interactions, and the quantum chromodynamics (QCD) which addresses the strong interactions. The fourth fundamental interaction, gravity, is not part of the model. According to the SM, the building blocks of matter are fermions, leptons (electrons, muons, tau-leptons, and neutrinos) and quarks, and the forces between them are mediated by bosons, gluons (strong), photons (electromagnetic), and W and Z bosons (weak) as shown in Table 1.1.

The intrinsic symmetry between the electromagnetic and weak interactions in the SM is broken via Higgs mechanism which implies existence of a special particle called Higgs boson. It has not been observed until in Summer 2012 the two major LHC experiments ATLAS and CMS announced observation of a new particle at a mass of

125 GeV with properties consistent with those of the Higgs boson.

The current version of the SM was finalized in the mid 70's after the first evidence for quarks was obtained in deep inelastic scattering experiments at SLAC in 1968. Since that time, three generations of quarks have been discovered by experimentalists. The top quark is the last discovered quark in the SM.[4, 5, 6]

The SM has been extensively validated by multiple experiments with astonishing accuracy. Recent experimental results, however, suggest that the SM may be incomplete and has to be expanded. Experiments with neutrinos revealed that these fundamental particles have a mass, while in the SM they are massless. The dark matter which according to astrophysics constitutes a significant fraction of our Universe, is not part of the SM. The fact that our Universe consists of matter without traces of antimatter suggests a large violation of charge conjugation and parity (CP) symmetry at the level that cannot be described within the SM framework. Finally, the SM cannot explain why fundamental particles have different masses. By searching for new physics beyond the SM, the LHC experiments hope to solve these puzzles and provide the basis for the SM extension which will be able to explain all observed phenomena.

1.2 Top Quark Pair Production and Decay

At the LHC, the top quarks are produced predominantly in pairs. As opposed to the Tevatron where the dominant top production mechanism is the quark-anti quark annihilation $q\bar{q} \to \bar{t}$, at the LHC energies the leading source of top quarks pairs is the gluon fusion. The leading order Feynman diagrams for pair production are given in Figure 1.1.

For the analysis's in this thesis uses the $t\bar{t}$ that is mainly simulated with the event generator MC@NLO. Cross sections for different physics processes at various energies

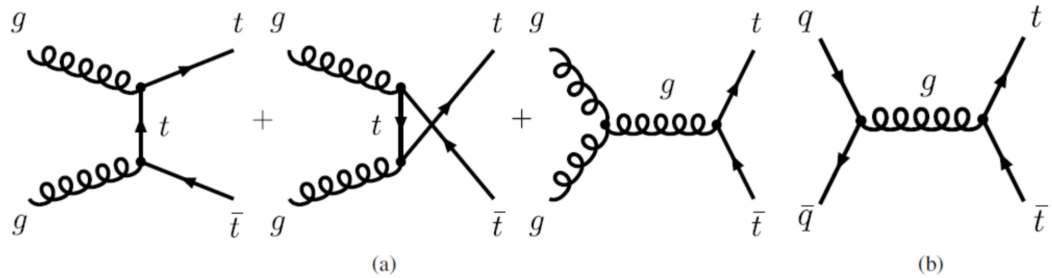


Figure 1.1: Feynman diagrams of the leading order processes for $t\bar{t}$ production: (a) gluon-gluon fusion and (b) quark-antiquark annihilation.

compared to $t\bar{t}$ are shown in Figure 1.2 for both Tevatron and LHC energies.

The $t\bar{t}$ cross section has been estimated with the HATHOR program at approximate next-to-next-to-leading order (NNLO) in QCD using the HATHOR program and CTEQ66 parton distribution functions. The result at a center-of-mass energy of $\sqrt{s} = 7\,TeV$ is [7]

$$\sigma(t\bar{t}) = 164.57^{+4.30}_{-9.27}(sacle) +^{+7.15}_{-6.51}(PDF)\,pb \tag{1.1}$$

In the SM, the top quark decays predominantly into a W boson and a down-type quark. The branching ratios are given by the Cabibbo-Kobayashi-Maskawa (CKM) unitary rotation matrix [8]. The diagonal elements are close to the unity and the off-diagonal are closer to zero, especially for heavy quarks. The leading top quark decay is $t \to Wb$, as the decays to light quarks (d and s) are suppressed by the square of CKM matrix elements V_{td} and V_{ts}, respectively.

$$\begin{bmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{bmatrix} = \begin{bmatrix} 0.97428 \pm 0.00015 & 0.2253 \pm 0.0007 & 0.00347^{+0.0016}_{-0.00012} \\ 0.2252 \pm 0.0007 & 0.97345^{+0.00015}_{-0.00016} & 0.0410^{+0.0011}_{-0.0007} \\ 0.00862^{+0.00026}_{-0.00020} & 0.0403^{+0:0011}_{-0.0007} & 0.999152^{+0:000030}_{+0:000045} \end{bmatrix} \tag{1.2}$$

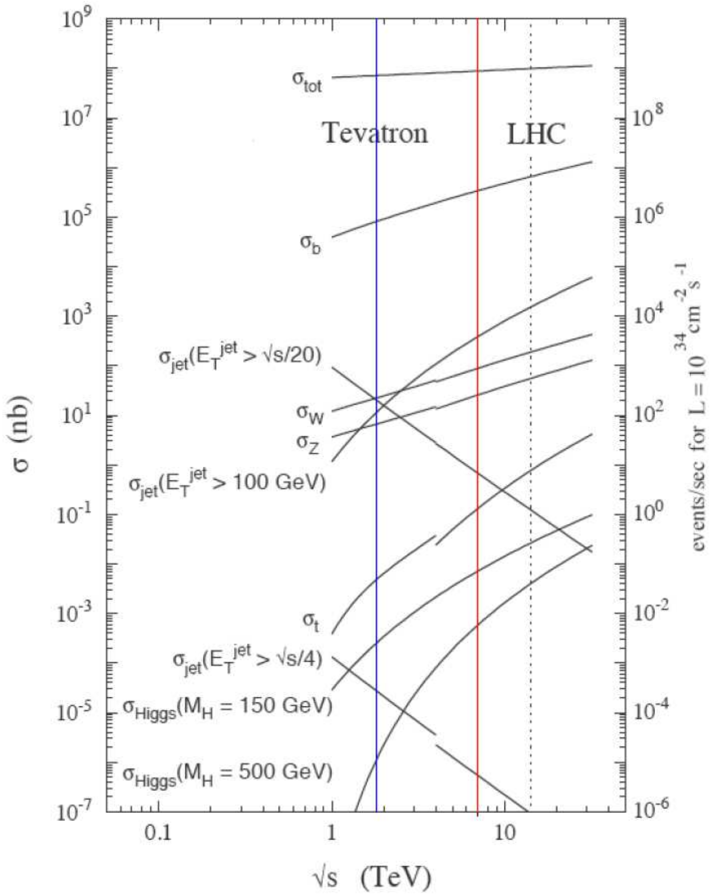


Figure 1.2: Cross sections for various physics processes at different center of mass energies. The blue line shows the Tevatron energy and the red lines shows the current and dashed and future LHC energies.[1]

The branching ratios of the decaying top quark to any of quarks are estimates as :

$$B(t \to bW^+) \cong 0.998 \quad (1.3)$$

$$B(t \to sW^+) \cong 1.63 \times 10^{-3} \quad (1.4)$$

$$B(t \to sW^+) \cong 1 \times 10^{-4} \quad (1.5)$$

The W boson in turn decays into a lepton and a corresponding neutrino, or into a pair of quarks. The quarks produced in the top and W decays hadronize giving rise to jets (groups of collimated particles). Depending on the W decay mode for each of the top quarks in the $t\bar{t}$ pair, the $t\bar{t}$ events are categorized into three main channels as shown in Figure 1.3:

1. **Dilepton channel:** $2\ \ell + 2\ \nu + b\bar{b}$ (two b quarks). This mode is quite clean and in principle can be easily identified by requiring two high p_T leptons and the presence ofE_T^{miss}.

2. **Semilepton channel:** $1\ \ell + 1\ \nu + b\bar{b}q\bar{q}'$ (i.e. being two of these four quarks two light quarks[1] and the two other b quarks). The presence of just one high p_T lepton and missing energy due to ν in the final state allows to suppress the QCD background. This channel is quite easy to trigger due to the presence of high P_T lepton.

3. **Fully hadronic channel:** $b\bar{b} + q\bar{q}' + q''\bar{q}'''$ (four of these six quarks light quarks and the other two b quarks). In this case, there are no leptons to trigger on, and the signal is not easily distinguishable from the abundant SM QCD

[1] The "light quark" term refers to u, d, c or s quarks.

multijets production, which is expected to be orders of magnitude higher than the signal (before any selection cuts). Furthermore, from the technical point of view, the presence of a high combinatorial background is another challenging problem.

The Branching ratio (BR) of each $t\bar{t}$ decay channel can be computed based on each W boson decay channel since the W decays into $e\nu$, $\mu\nu$, $\tau\nu$ and $q\bar{q}'$. Table 1.2 shows numerically the BRs for each $t\bar{t}$ decay channels, computed from their probabilities.

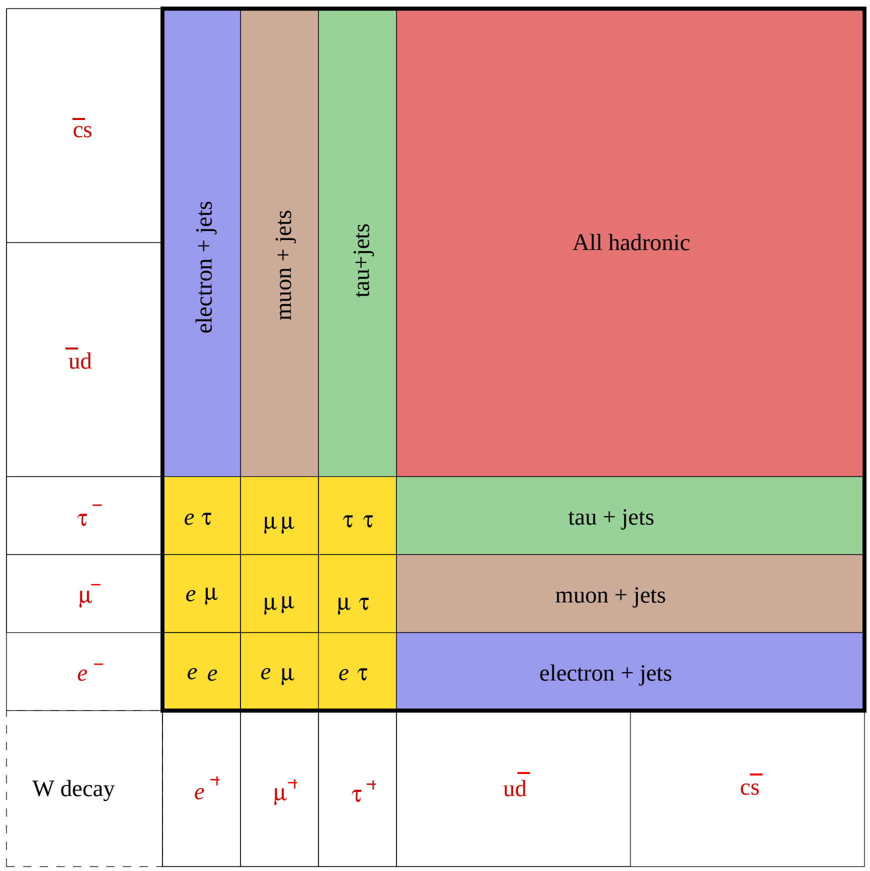


Figure 1.3: Graphical representation of the $t\bar{t}$ decay channels.

channel	Signature	BR(%)
Fully Hadronic	All Jets	45.7
Lepton + Jets	e + Jets	17.2
	μ + Jets	17.2
Dilepton	$e\mu$ + Jets	3.18
	$\mu\mu$ + Jets	1.59
	ee + Jets	1.50
Tau + Jets	τ + Jets	9.49
Lepton + Tau	$\tau + e/\mu$ + Jets	3.54
Tau + Tau	$\tau + \tau$ + Jets	0.49

Table 1.2: Top quark pair decay channels.

Chapter 2

LHC and Atlas Detector

The Large Hadron Collider (LHC) is the largest international particle accelerator in the world. It is located at European Organization for Nuclear Research (CERN) Conseil Europeen pour la Recherche Nucleaire which is the largest energy of collisions ever achieved by mankind.

The main purpose of building such a complex machine is to find out the answers on the most fundamental questions about the structure of our Universe and interactions between its building blocks. These questions include the origin of mass, the presence of matter and absence of antimatter, nature of the dark matter, the number of dimensions in our Universe, the state of matter right after the Big Bang and many others. No single detector can address all the questions, so six detectors were carefully designed to search for answers for specific questions.

LHC is a two-ring-superconducting hadron accelerator with a circumference of 27 km. The beams are crossed at four interaction points where the detectors are located. These detectors are ATLAS [9], ALICE [10], CMS [11], LHCb [12]. ATLAS and CMS are general purpose detectors designed to search for a new physics, like dark matter particles, extra dimensions, supersymmetric partners of the SM fermions and bosons, and for precise measurements of the SM predictions. LHCb detector was built to

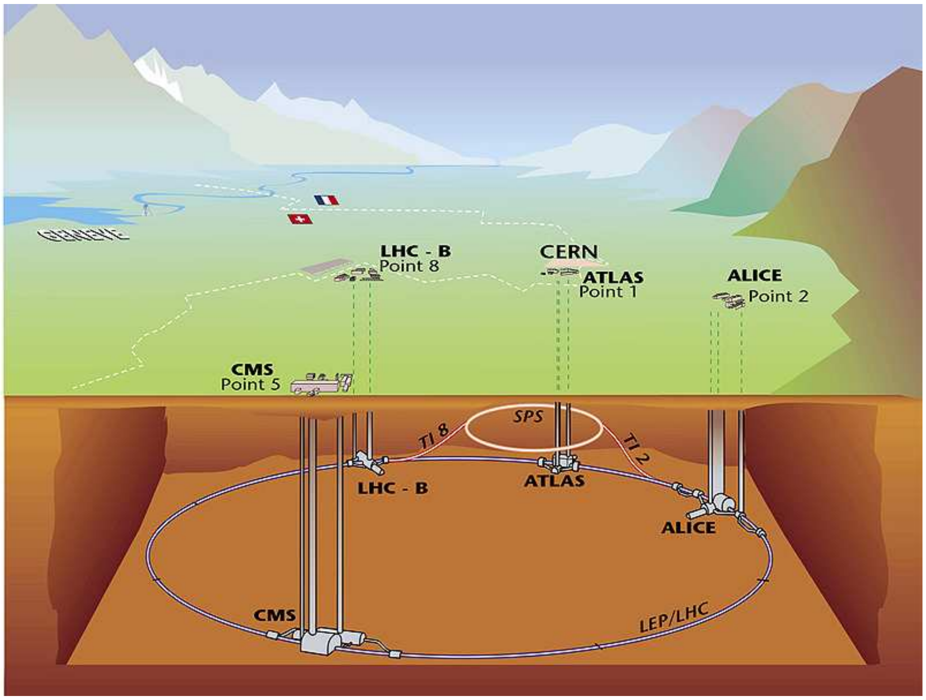


Figure 2.1: Representation of the LHC ring with its detectors and all its services.

investigate the CP violation, which led to the fact that our Universe is composed of matter. ALICE experiment collects heavy ion (lead) collisions to investigate the properties of quark-gluon plasma which probably existed just after Big Bang. The 2.1 shows an illustration of the LHC accelerator with its four main experiments. In this chapter a brief summary of LHC and ATLAS detector is provided.

2.1 Large Hadron Collider

The LHC particle collider can accelerate two intense proton beams to an energy up to 7 TeV each. LHC will start to work in this regime in 2014. Besides the center-of-mass energy, the other important parameter of a collider is instantaneous luminosity $\mathcal{L}$ which depends on the number of particles per bunch in the two colliding beams, N_1 and N_2, on the bunch crossing frequency f and on the bunch area. Thus, assuming a Gaussian transverse particle bunch profile, the $\mathcal{L}$ is defined as:

$$\mathcal{L} = \frac{1}{4\pi}\frac{N_1 N_2}{A} f N_b \tag{2.1}$$

where N_b is the number of bunches and A [1] is the beam profile at the interaction point.

The LHC's designed luminosity $\mathcal{L}$ is 10^{34} cm^{-2} s^{-1}. It can also collide heavy (Pb) ions with an energy of 2.8 TeV per nucleon with a peak luminosity of 10^{27} cm^{-2} s^{-1}. That leads us to collisions almost 30 times more intense and seven times more energetic than the Tevatron machine. The LHC is installed in the existing tunnel that was built in 80's for CERN LEP accelerator.(Figure 2.3). [13]

The LHC started for the first time in September 2008 by injecting and circulating the proton beams. But due to faulty electrical connection between two magnets that caused an explosion and a large helium leak into the tunnel, the operations were stopped shortly after its opening. After repair in electrical connection, LHC was set to go back online in November 2009 with energy of 450 GeV per beam. The first collision at a center of mass energy of $\sqrt{s} = 7$ TeV took place at the end of March 2010 with luminosity of 10^{27} cm^{-2} s^{-1}. The beam intensity raised in 2011 to reach the luminosity of 3.65×10^{33} cm^{-2} s^{-1}. An integrated luminosity of 45 pb^{-1} was delivered by end of 2010 and 5.25 fb^{-1} during the whole year of 2011. Figure2.2. shows the delivered integrated luminosity. At the end of the end of 2012, the LHC will shut down for maintenance for up to two years and then it will start to work at the design energy of 14 TeV.

[1] A defined as $A = \sigma_x \times \sigma_y$. σ_y and σ_x are the Gaussian widths in the horizontal and vertical plane of the bunch respectively.

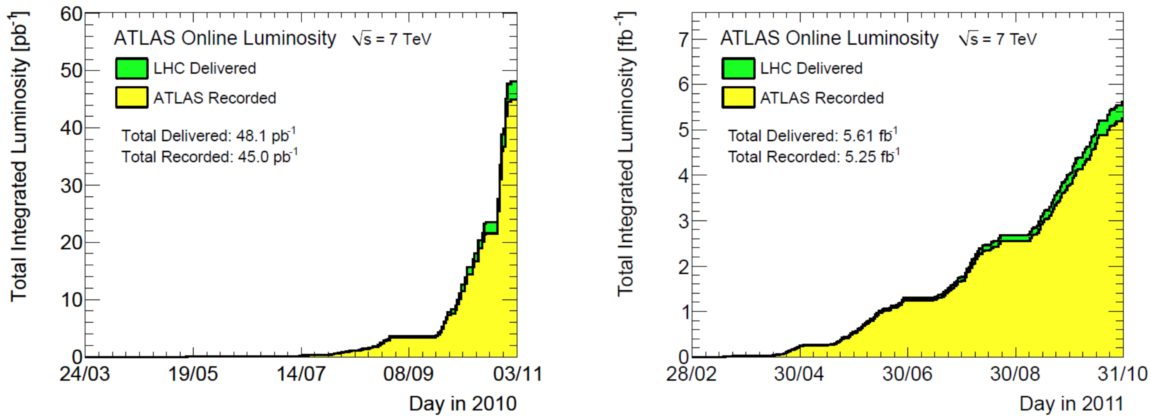


Figure 2.2: Cumulative luminosity versus day delivered to (green), and recorded by ATLAS (yellow) during 2010 (left) and 2011 (right) at $\sqrt{s} = 7$ TeV .

2.2 The Accelerator

The LHC is the name of main accelerator but there are series of accelerator before that, by which beam energy is raised step by step.

The starting point is a LINAC (Linear Accelerator) where protons are generated using hydrogen nuclei by separating the electron, and accelerated up to 50 MeV. These protons are injected to the Proton Synchrotron Booster (PSB, 157 m) and accelerated up to 1.4 GeV. Then, they are injected to the Proton Synchrotron (PS, 628 m), where protons are accelerated up to 26 GeV. Finally, they are injected to the Super Proton Synchrotron (SPS, 7 km) to increase its energy up to 450 GeV followed by LHC to achieve 7 TeV beam energy. At this energy, the relativistic of the proton reaches 7500, corresponding to 99.9999991% compared to the speed of light (c). The difference between the proton beams and the light after traveling a second become only 3 m. For the data-taking purpose, protons are bundled into 2808 bunches, each consist of 1.151011 protons. Those are controlled by 40 MHz RF (Radio Frequency) clock, which results in collision interval of 25 ns. RF clock is also distributed to each detector to achieve synchronized data-taking.

The existing CERN accelerator complex is used to accelerate these beams. The

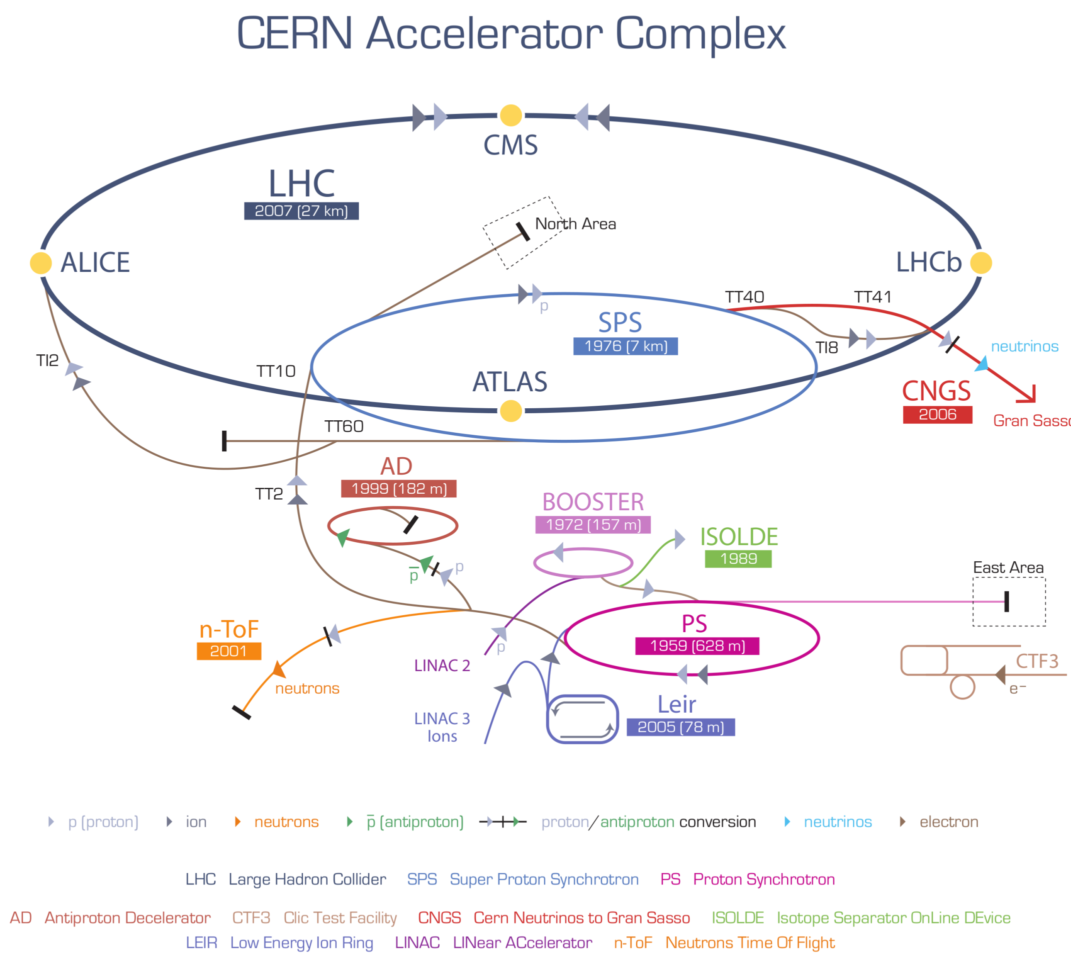


Figure 2.3: LHC accelerator complex at CERN

path that the beams follow is shown in Figure 2.3. Beam starts from the LINACs (50 MeV), then goes to the PS Booster (1.5 GeV) , after that continues on through the PS (26 GeV) and after that goes through the SPS which is 7 km length. The SPS accelerates the beams up to 450 GeV and finally they will be injected into the LHC ring that is a 27 km long tunnel and 100 m below ground.

Each beam has internal structure such arranged in bunches that condense up to $1.1 \cdot 10^{11}$ charged particles. These bunches are separated in space, thus collisions have a rate of 40 MHz (one collision every 25 ns). In table 2.1 one can find a summary of the main LHC parameters.

LHC Parameter	**p-p Collisions**	**Pb-Pb Collisions**
Beam energy	7 TeV	570 TeV (2.76 TeV/u)
Injection energy	450 GeV	1.774 TeV/u
Center of mass energy ($\sqrt{s}$)	14 TeV	1250 TeV
Luminosity ($\mathcal{L}$)	10^{34} cm^{-2} s^{-1}	10^{27} cm^{-2} s^{-1}
Frequency (f)	40 MHz	10 MHz
Bunch separation	25 ns	100 ns
Number of particles per bunch (N_b)	$1.1 \cdot 10^{11}$	$7 \cdot 10^{7}$
Average size of a bunch	7.5 cm	
Average radius of a beam	16 μm	
Beam current	0.56 A	
Collision regions	3 (2 for high $\mathcal{L}$)	1
Circumference length	26.66 km	
Radius	4.24 km	
Number of dipole magnets	1232	
Length of the dipole magnets	14.3 m	
Nominal magnetic field (B)	8.33 T	
Number of quadrupole magnets	392	
Total mass	27.5 tons	

Table 2.1: Main parameters of the LHC collider.

The LHC has more than 1200 superconducting dipole magnets that conduct and curve the beams through the LHC ring. The magnets uses twin bore magnets which consist of two sets of coils and beam channels within the same mechanical structure and cryostat. This design was chosen because the LHC needs to accelerate two beams of equally charged particles but in opposite directions and there is not enough space for two separate magnet arrays in the LHC tunnel. These magnets are able to generate bipolar magnetic field up to $8.33\,T$ due to their superconductivity capabilities. They fill more than 66% volume of the LHC ring. The coils design is based on niobium titanium (NbTi) that is a material that allow to reach the superconducting regime when it is at 1.9 K which is colder than the outer space. A comprehensive cross section of a dipole magnet is shown in Figure 2.4 with detail of where all its parts are precisely located.

In addition, there are almost 400 quadrupole magnets to perform beam focusing

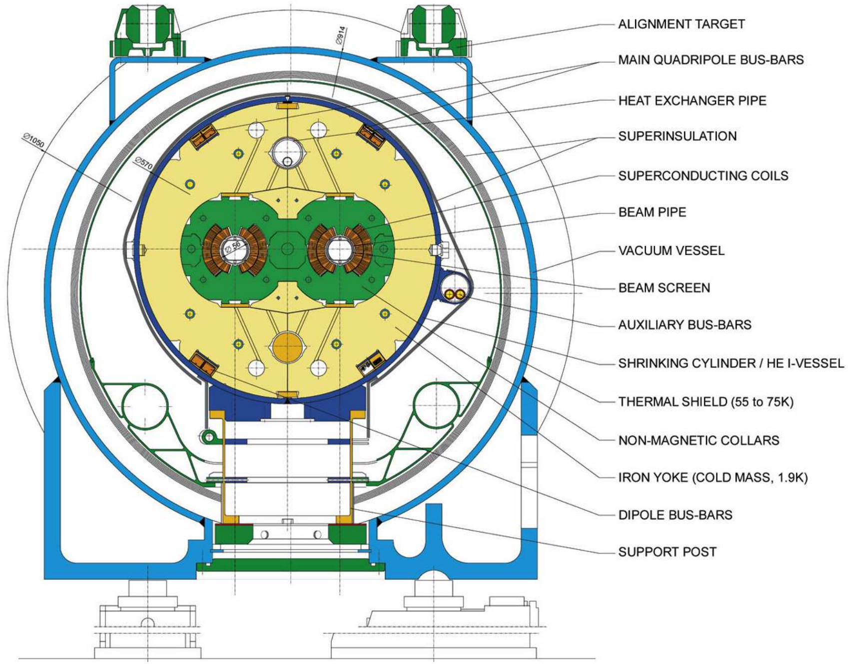


Figure 2.4: Cross section of a LHC dipole magnet design showing its components.

and beam corrections and also sextupole, octupole and decapole magnets mainly for compensating the systematic non-linearities.

2.3 The ATLAS Detector

ATLAS (A Toroidal LHC ApparatuS) is one of two experiments designed to realize the main physics goals of the LHC, namely searches for the Higgs boson and for supersymmetric particles. The formal proposal for ATLAS was introduced in 1994 [14] and ten years later the detector installation in the cavern began [9] . The ATLAS collaboration is quite large with about 3000 physicists scattered around the world. As its collaboration, the ATLAS detector is also quite large as seen in Figure 2.5 compared with the human figures which are included for a scale reference. Thus, with its 44 m long and 25 m high is the biggest LHC detector although not the heaviest with a weight of 7000 tons.

The ATLAS general layout is fairly typical general purpose high energy physics

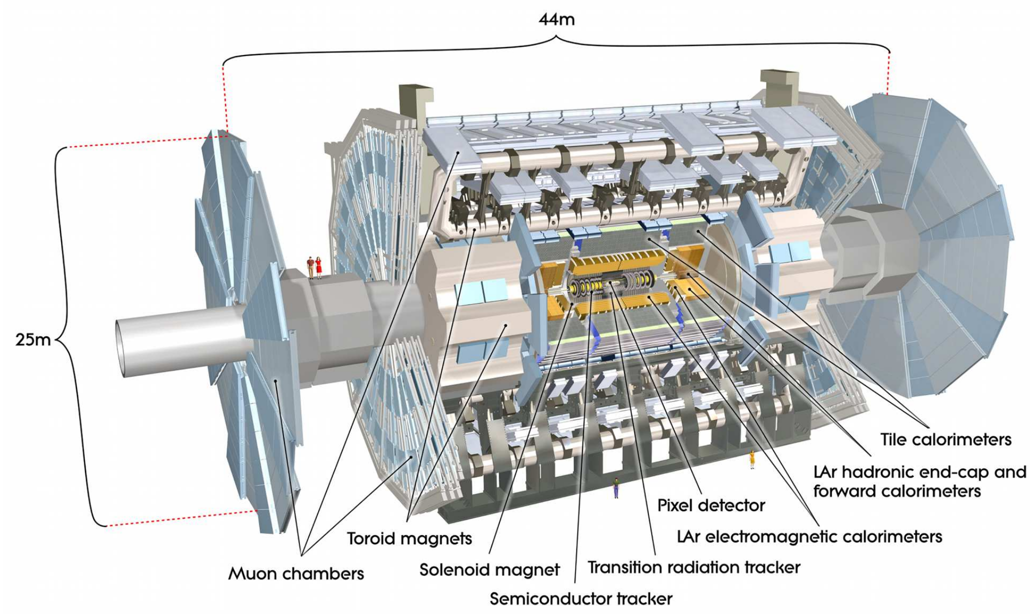


Figure 2.5: The ATLAS detector with all its subdetectors labeled. Moreover, several human figures, barely visible, are included for a scale reference.

collider detector with cylindrical shape and onion style subdetectors. The individual subsystems are arranged in cylindrical layers around the beam pipe which form the detector barrel. On each side of the barrel the subsystems are completed by wheels perpendicular to the beam pipe, the end-caps.

The innermost part of detector is the tracker that the closest part of detector to the beam pipe. It operates inside of a 2 T magnetic field and includes several of layers of subdetectors that are high resolution discrete silicon detectors and on drift tubes which is a continuous gaseous detector. This subdetector provide a robust pattern recognition, momentum and charge determination, precise primary and secondary vertex reconstruction as well as particle identification capability with pion separation. Surrounding the solenoid tracker are the calorimeters. It contains two major parts, the electromagnetic and hadronic calorimeter. The electromagnetic calorimeter uses liquid argon as an ionization medium. It allows the identification and measurement of photons and electrons. The electromagnetic calorimeter is surrounded by

detector component	η coverage	design resolution	readout channels
Inner Detector	± 2.5	$\sigma_{p_T}/p_T = 0.05\% \times p_T \oplus 1\%$	
Pixel	± 2.5		$\approx 80 \times 10^6$
SCT	± 2.5		$\approx 6.3 \times 10^6$
TRT	± 2.0		$\approx 351 \times 10^3$
EM calorimeter	± 3.2	$\sigma_E/E = 11\%/\sqrt{E} \oplus 0.7\%$	$\approx 173 \times 10^3$
Hadronic calorimeter			
barrel + end-cap	± 3.2	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	$\approx 15 \times 10^3$
forward	$3.1 < \lvert\eta\rvert < 4.1$	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$\approx 3.5 \times 10^3$
Muon Spectrometer	± 2.7	$\sigma_{p_T}/p_T = 10\%$	
MDT	± 2.7		$\approx 354 \times 10^3$
RPC	± 1.05		$\approx 373 \times 10^3$
TGC	$1.05 < \lvert\eta\rvert < 2.7$		$\approx 318 \times 10^3$
CSC	$2.0 < \lvert\eta\rvert < 2.7$		$\approx 31 \times 10^3$

Table 2.2: The designed coverage and resolution are listed for the individual subsystems of ATLAS.

the hadronic calorimeter which uses a scintillating tile technology that allows to measure hadronic jets and help to estimate the missing energy. Outside the calorimeters, ATLAS detector has a large and important muon spectrometer that provides the measurements of muon momenta. It also includes fast response trigger chambers. The muon spectrometer rests inside an air core toroidal magnetic field. The designed resolution and coverage of the individual systems are summarized in Table 2.2.

2.3.1 The Coordinate System

The coordinate system of the Atlas detector is right handed, and its origin is in the middle of the detector at the interaction point. The z-axis is defined from the beam direction, and the $x-y$ plane is the plane transverse to the beam direction. The side A of the detector is defined as the side with positive z, and the side C is the side with $z < 0$. The positive x-axis points from the interaction point to the center of the LHC ring, and the positive y-axis is pointing upwards. In experimental particle physics, pseudorapidity (η) is related with the azimuthal angle θ (i.e. it is related with the angle of a particle relative to the beam axis) as follows:

$$\eta = -ln\left[tan\left(\frac{\theta}{2}\right)\right] \tag{2.2}$$

The transverse momentum, p_T , and transverse energy, E_T , are defined in the $x - y$ plane and are given by the equations:

$$p_T = Psin(\theta),\ E_T = Esin(\theta) \tag{2.3}$$

The other common quantity is distance in the pseudorapidity-azimuthal angle space ΔR between two objects define as

$$\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \tag{2.4}$$

2.3.2 Inner Detector

The ATLAS inner detector (ID) is designed to measure the tracks of charged particles produced in the proton-proton interactions. It has high granularity and a large number of layers of tracking systems to provide an excellent pattern recognition and track reconstruction efficiency in the extremely large track density environment from the collisions.

The ID consists of the three tracking subsystems. These are pixel detector, semiconductor tracker (SCT) and transition radiation tracker (TRT). Pixel detector and SCT have silicon sensors. TRT is a set of drift tubes. The ID operates inside a 2 T axial magnetic field generated by a solenoid [15]. Combination of discrete high resolution semiconductor pixel and strip detectors in the inner part of the tracking volume and straw-tube tracking detectors with the capability detect transition radiation in its outer part makes possible to accurately reconstruct the primary vertex of

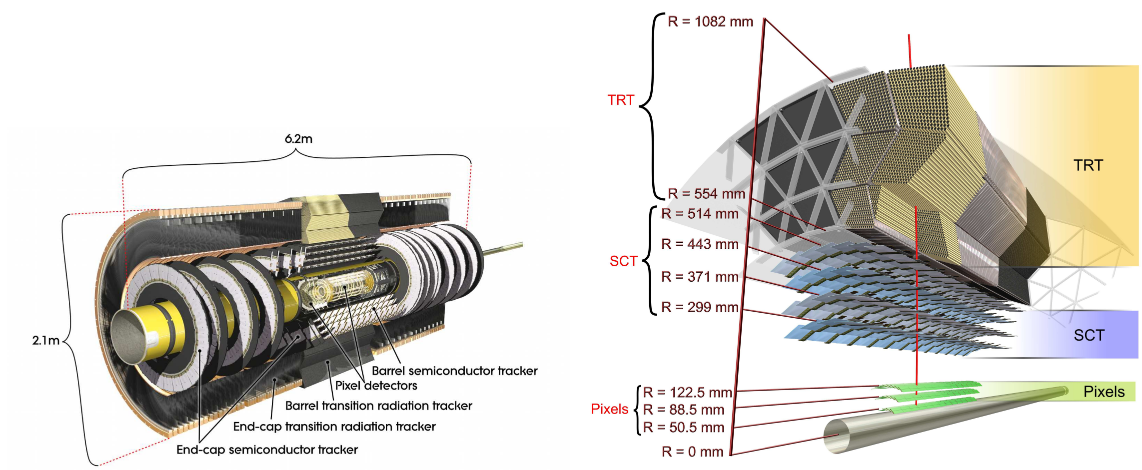


Figure 2.6: A sketch of the ATLAS inner detector.(at left) The pixel detector has three barrel layers and three end-cap discs on each side. The SCT has four barrel layers and nine end-cap discs each. The TRT has three barrel layers and two end-cap wheels on each side. Position of the structural elements of the ID.(at right) The red line indicates a track passing the barrel region.

the proton-proton collision of interest, tracks of the charged particles with transverse momenta $p_T > 0.5$ GeV, and decays of long-lived particles, which fly some distance away from the primary vertex before their decay. TRT allows one to identify electrons based on the relative fraction of hits with high energy with respect to the total number of hits in this tracker. Three-dimensional view of the ID is shown in Figure 2.6.

The ID has 5832 individual silicon modules with about 86 million of readout channels providing a a very fine resolution to measure the particle trajectories and allowing to reduce the influence of overlapping events. Each of these modules measure the position of particles with a precision of about 20 μm. The ID covers a pseudorapidity range up to $|\eta| < 2.5$.

The silicon pixel detector is the innermost subdetector of the ID. It provides the highest space-point resolution in the vertex region. It consists of three layers and three disks of pixel modules as shown in Figure 2.6. A module is an active device approximately 6 cm by 2 cm and contains 46,080 pixels. Each pixel is 50 μm in

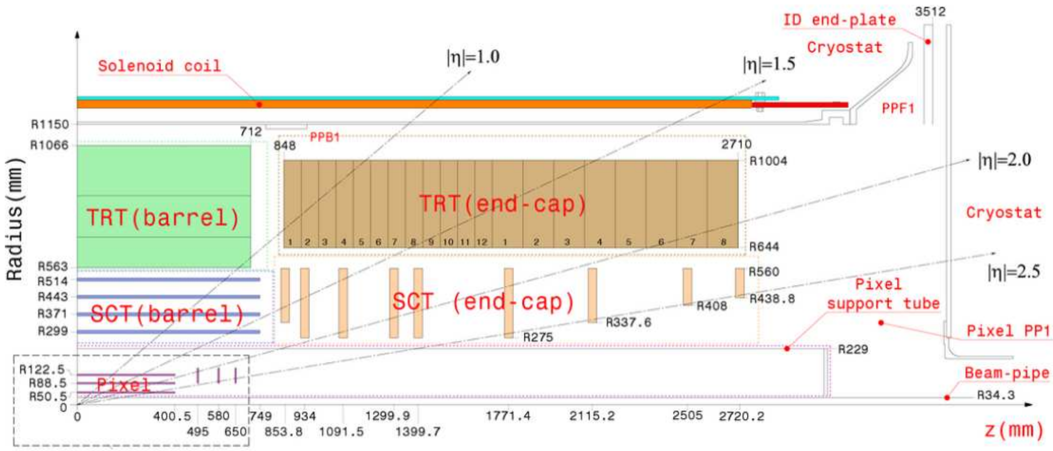


Figure 2.7: The η coverage of inner detector

azimuth by 400 μm along the beam. All modules are identical.

The designed expectation resolutions of the Pixel detector is $10\,\mu m$ by $115\,\mu m$. There are in total $174419 \times 63\,mm^2$ sensors in the Pixel detector which operate at the temperature $-5 \div -10°$ C. The design of the Pixel sensors are challenged and restricted by the specification in the high energy hadron collider experiment such as the radiation hardness, occupancy and resolution in the innermost layers. Accordingly, the sensors are using oxygenated n-type wafers with readout pixels on the n^+ implanted side of the detector with double-side processing involved so as to guarantee good charge-collection efficiency and radiation tolerance. The B-layer which is the innermost layer of the Pixel detector is essentially important in b-jet identification as well as helping to identify electrons from converted photons.

The SCT is a silicon microstrip detector that is placed in just after the pixel detector. The SCT modules are lined up on four barrel layers and nine end-cap discs on each side. The design is specified to perform eight precision measurements via 4 layers of back to back silicon microstrip detector modules with a relative 40 mrad stereo angle. There are five sensor parts, one for the barrel which is parallel strips with 80 μ m pitch and four for the end caps with fan-out structure with pitches varying between 54.53 and 90.34 μ m. The SCT module with 80 μ m strip pitch (on average) provides a 17 μ m precision in $r-\phi$ and 580 μ m precision in z-direction. The SCT in overall has more than 4000 modules (2112 barrel and 1976 end-cap modules), covering 61 m^2 by silicon sensors with 6.3 million channels.

The TRT is the outermost part of the ID and its design is based on straw detectors. It covers the radial range from 56 to 107 cm. Each straw has a diameter of 4 mm and is 144 cm long in the barrel region and 37 cm long in the end-caps. The straw tubes are filled with a gas mixture of Xe (70%) , CO2 (27%) and O2 (3%) . The anode wires placed in the center of the straw has a diameter 31 μm and are made from gold plated tungsten. There are approximately 351000 TRT read-out channels. The TRT system can provide 36 hits per track with an designed precision of 130μm per tube that covers a range of $\eta < 2.0$. The tubes, in the barrel region ($|\eta| < 1.0.$), are placed parallel to the beam axis, while in the end-cap regions ($0.8 < \eta < 2.0$) they are arranged radially on wheels. In principle the mechanism of operation is very similar to the drift tubes used in the muon spectrometer. The straws are enclosed by polypropylene fibers that is a key feature to help differentiate electrons from pions. The discriminating mechanism is based on effect that is called transition radiation (TR), since the charged particles traversing fibers emit photons proportional to the Lorentz factor $\gamma = E/mc^2$. Typically, for electrons the factor γ is greater than for other charged particles. The η coverage of inner detector and cross-section of one quadrant of the Inner Detector shown in Figure 2.7.[16]

2.3.3 Calorimetry

The calorimeters are designed to perform energy measurements and particle identification of all incoming particles by absorbing them. Muons and neutrinos are the only known particles that can traverse through the calorimeters. ATLAS detector has two

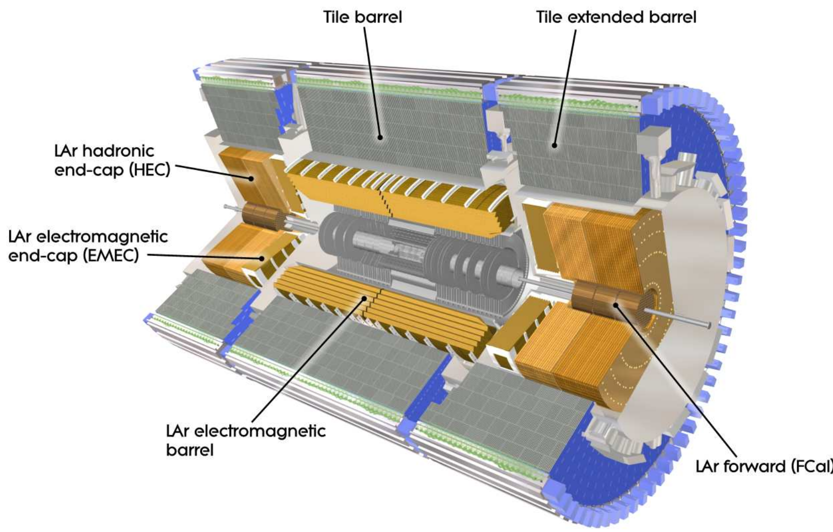


Figure 2.8: Cut-away view of the ATLAS calorimeter system with its subdetctors labeled.

types of calorimeters, electromagnetic and hadronic ones, as shown in Figure 2.8. The results of combining the ID and calorimeter measurements provides an identification of electrons, photons and gives probabilities for a jet to result from a b or τ decay.

The Electromagnetic Calorimeter (ECAL) is situated outside the solenoidal magnet that surrounds Inner Detector. The principle of work of this detector is based on the fact that some particles interact with the material primarily via electromagnetic force. Its main purpose is to identify photons and electrons that reach this calorimeter. ECAL uses liquid argon as an ionization medium, with the lead absorbers arranged in an accordion geometry [17, 18] . This kind of geometry provides complete ϕ symmetry without azimuthal cracks and the lead thickness in the absorber plates is optimized as a function of η in terms of performance in energy resolution. Electrons and photons release all their energy in the ECAL and are absorbed by its medium. Thus, ECAL provides an excellent performance in terms of energy and position resolution as well as in the identification of electrons and photons providing coverage up to $|\eta| < 3.2$. It is surrounded by cryostat as it needs very low temperatures to work properly.

Surrounding the ECAL is the Hadronic Calorimeter (HCAL), which absorbs energy of particles passed through the ECAL, and can interact strongly with the HCAL medium. HCAL provides a coverage up $|\eta| < 4.9$. The information from HCAL is used to reconstruct parameters of hadronic jets and to calculate the transverse missing energy in the event. A sampling technique with plastic scintillator plates (called tiles) embedded in an iron absorber is used for the hadronic barrel tile calorimeter (also known as TileCal)[19, 20, 21]. HCAL is separated into a large barrel and two smaller extended barrel cylinders, one on either side of the central barrel. In the end-caps ($|\eta| < 1.6$), LAr technology is also used for the hadronic calorimeters, matching the outer $|\eta|$ limits of end-cap electromagnetic calorimeters. The LAr forward calorimeters provide both electromagnetic and hadronic energy measurements, and extend the pseudorapidity coverage to $|\eta| < 4.9$.

In general, energy resolution of calorimeters can be parametrized as:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus \frac{b}{E} \oplus c \tag{2.5}$$

where a is the stochastic term (accounting for the statistical fluctuation in the number of primary and independent signal generating processes), b is the noise term (accounting for the electronics noise and fluctuations in energy carried by particles, other than the ones of interest) and c is the constant term (reflecting global non-uniformities in the response of the calorimeter as well as energy leakage and mechanical imperfections). While the sampling term dominates the energy resolution, the noise term limits the performance at low energies and the constant term sets the limit at high energies. Then, the tolerable size of the three terms depends on the energy range involved in the experiment.

Testbeam results have shown that the ECAL's energy resolution can be parametrized

$$\frac{\sigma_E}{E} = \frac{11.5\%}{\sqrt{E}} \oplus 0.5\% \quad (E \text{in GeV}) \tag{2.6}$$

For HCAL an energy resolution parametrization is following:

$$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\% \quad (E \text{in GeV}) \tag{2.7}$$

Information from calorimeters is further used in combination with the information from ID for identification of electrons, photons, τ-leptons and other physical objects.m.

2.3.4 The Muon Spectrometer

The outermost detector is the muon spectrometer which defines the overall dimensions of the ATLAS detector [15]. Muon spectrometer is designed specifically to identify muons and measure their momenta. Muons do not leave significant amount of energy in any other subsystem, so they reach the Muon Spectrometer. Its layout can be seen in figure 2.9. It consists of four different types of detectors . These are :

- Monitor Drift Tube Chambers (MDT), which cover 99.5% of the area and are used for precision measurements up to $|\eta| < 2.7$, except from the innermost end-cap layer which stops at $|\eta| < 2.0$.

- Cathode Strip Chambers (CSC), which are located on the innermost wheel of the endcap covering the region $2.0 < \eta < 2.7$ and have the same objectives as the MDTs. They are also used as a part of muon trigger system.

- Resistive Plate Chambers (RPC), which are utilized for triggering in the barrel region;

- Thin Gap Chambers (TGC), which are the trigger elements in the end-cap regions.

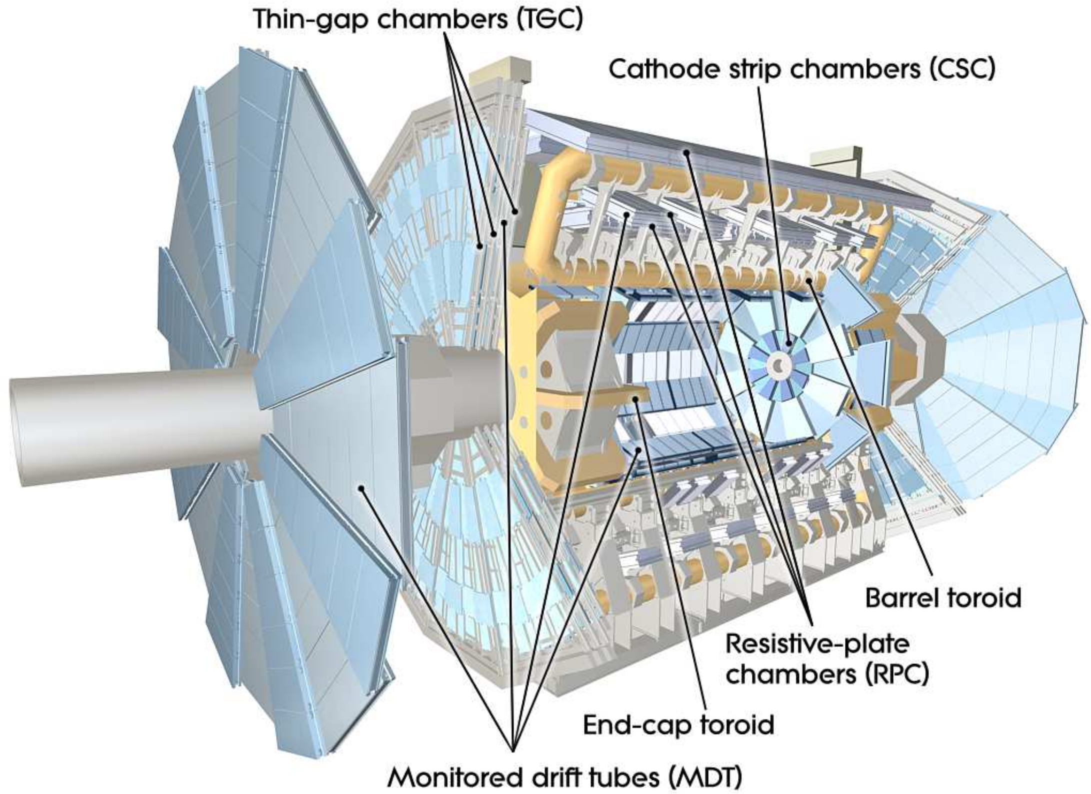


Figure 2.9: Cut-away view of the ATLAS muon system.

The muon spectrometer is designed to achieve a transverse momentum resolution of $\Delta p_T/p_T < 10^{-4}$ for p_T >300 GeV. At smaller momenta, the resolution is limited to a few per cent by Multiple Coulomb Scattering (MCS) effects in the magnet and detector structures, and by energy loss fluctuations in the calorimeters

An air-core toroid system generates strong bending power in a large volume within a light and open structure. Thus, MCS effects are thereby minimized, and excellent muon momentum resolution is achieved with three layers of MDT chambers achieving a precision of ~ 50 μm in muon position measurements. This magnetic system as can be seen in Figure 2.10 has a barrel (25 m long, with an inner bore of 9.4 m and an outer diameter of 20.1 m) and two inserted end-cap magnets (with a length of 5.0 m, an inner bore of 1.65 m and an outer diameter of 10.7 m).

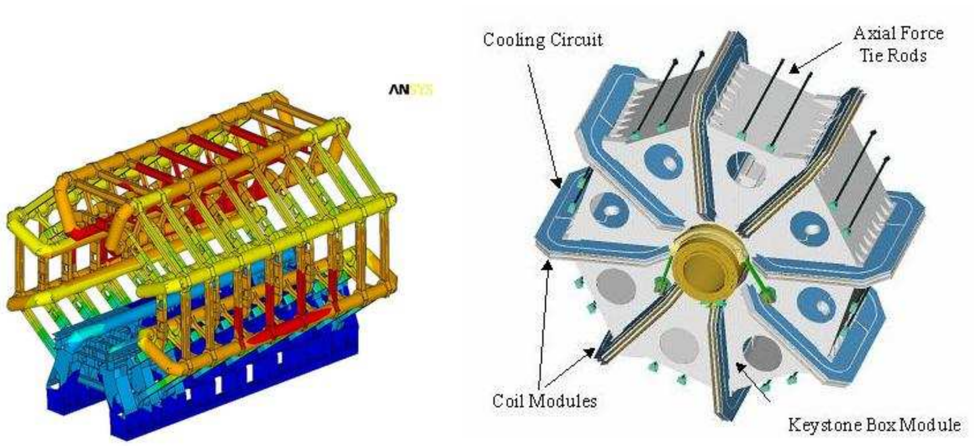


Figure 2.10: Barrel toroid and end-cap magnets.

Each barrel toroid consists of eight flat coils assembled radially and symmetrically around the beam axis and the magnetic field provides for typical bending powers of 3 Tm in the barrel and 6 Tm in the end-caps . The end-cap toroid coils are rotated in azimuth by an angle of 22.5^0 with respect to the barrel toroid coils to provide radial overlap, and to optimize the bending power in the transition region between the two toroids ($1.0 \leq |\eta| \leq 1.4$) where the magnetic field is a superposition of the barrel and end-cap fields.

2.3.5 Trigger

The LHC proton bunches will collide at a frequency of 40 MHz, i.e. every 25 ns and at the design luminosity of 10^{34} cm^{-2} s^{-1} on average about 23 inelastic proton-proton collisions will be produced at each bunch crossing. Thus, the trigger system needs to efficiently reject a large rate of background events and still select potentially interesting ones with high efficiency. To deal with this amount of data that these collisions will generate the ATLAS trigger is based on three levels of online event selection as it is shown in Figure 2.11. Each trigger level refines the event selection done by the previous level, applying new criteria.

The level-1 trigger (LVL1)[1, 22], that is hardware-based and it is implemented in

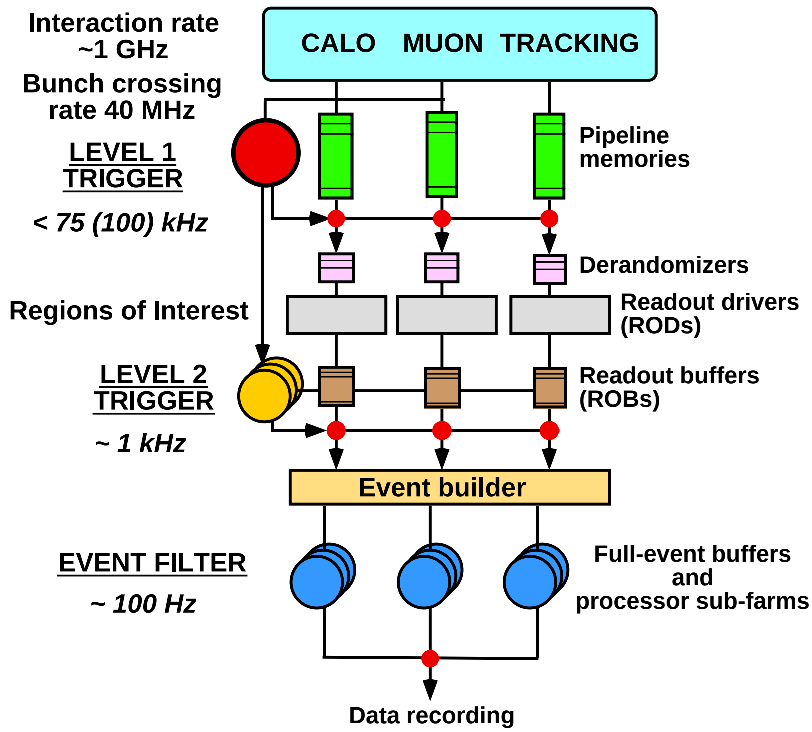


Figure 2.11: Diagram of the ATLAS trigger/DAQ system showing the 3 trigger levels.

custom electronics, is responsible for the first level of event selection. It reduces the initial event rate to less than 70 kHz which is limited by the bandwidth of the readout system. It uses information from the calorimeters and muon trigger chambers to make a decision on whether or not to continue processing an event in about 2.5 μs. The subsequent two levels are software-based and are collectively known as the High Level Trigger (HLT) . On the other side, the level-2 trigger (LVL2) decides in about 10 ms if the event should be rejected making use of the Regions of Interest (RoI) provided by the LVL1 data with full granularity from all detectors. On the other hand, the Event Filter (EF) uses offline algorithms to perform its refine selection in about 1 s. Thus, the HLT provides the reduction to a final data taking rate of approximately 200 Hz where each selected event is estimated to have a total size of ~1.6 MB. This reduction is possible because the HLT uses step-wise and fast selection algorithms that is based on the reconstruction of potentially interesting physical objects like electrons, muons, jets, etc and that can provide the earliest possible rejection of background events.

Chapter 3

Object Identification and Physics Backgrounds

3.1 Object Identification

The performed measurements are relying on reconstructed objects, which are jets, missing transverse energy, electrons, muons and tracks of charged hadrons. Brief description of each reconstructed object is given below.

Tracks: The tracks are the important inputs to the particle identification algorithms. They are formed from hits either in the silicon sensors of the pixel detector or SCT, or in the drift tubes of the TRT. Each hit defines a space point where particle has traversed the detector. The reconstruction algorithms that convert space-points into charged particle trajectories rely on several pattern recognition and track fitting methods. Track finding (i.e. associating hits to track candidates) and track fitting (i.e. estimating the track parameters and covariance matrix) are done simultaneously together with extrapolating the track through the B-field and estimating the energy loss in material. Three track reconstruction algorithms are performed as a result: [23]

- Inside-out algorithm: The reconstruction starts with silicon hits close to the interaction point and extend to TRT.

- Outside-in algorithm: Using the hits not used in the first algorithm, this algorithm start with TRT and extrapolate to silicon.

- Low p_T track reconstruction Similar as the inside-out algorithm, using the remaining hits and loose pattern recognition requirement. It mainly aims to recuperate the $p_T > 150$ MeV tracks.

Parameters of reconstructed tracks are used further for identification of leptons, jets originating from the main interaction point and many other purposes.

Electrons: The electron identification and reconstruction begins with finding of an energy cluster in the electromagnetic calorimeter. The special algorithm makes use of the difference of the shape of the cluster from an electron and any other charged particle. To distinguish between photons and electrons, the reconstructed electromagnetic cluster is required to have a track matched to it.[24] [25]

Electrons are required to have a transverse energy above 25 GeV and $\eta_{\mathrm{cluster}} < 2.47$, where η_{cluster} defines the electromagnetic cluster position. Electrons are vetoed in the calorimeter barrel-endcap overlap region,$1.37 < \eta_{\mathrm{cluster}} < 1.52$, where the calorimeter has limited capability of electron identification. Electrons found in the small regions with readout in-efficiencies in the liquid argon Calorimeter are removed in both data and MC.

Electrons are required to be isolated within the calorimeter and the inner detector. The calorimeter cells within a cone of $\Delta R = 0.2$ of the electron are summed. The energy deposit associated with the electron is subtracted from this sum. The calorimeter isolation threshold applied to the ratio of this cone energy to the electron energy is chosen such that electrons are reconstructed with 90% efficiency, including the effects of pileup and leakage. Isolation within the inner detector was performed in a similar manner, using a cone of $\Delta R = 0.3$ around the direction of the electron. The transverse momentum sum within the cone, after the exclusion of the electron track, was required to be within a 90% efficiency threshold. Electrons within $\Delta R = 0.4$ of any jet with $p_{\mathrm{T}} > 25$ GeV after the electron-jet overlap removal were rejected. Electron identification efficiencies for selected electrons have been measured on data using W and Z samples with tag and probe methods . The correction factors as functions of electron's pseudorapidity η and its transverse momentum p_{T} have been derived from these studies and applied to all MC events, to take into account for the differences between simulated and real data events.[26, 27]

Muons: Muons were reconstructed using the algorithm which combines muon spectrometer and inner detector track information. A global fit was performed to muon candidates, and results of this fit together with the information of all hits on the fitted track were used in the selection. Muon candidates were required to have $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. The background from leptonic b-hadron decays was reduced by requiring a hit in the b-layer of the inner most detector (pixel detector) if the track traversed a working module. Combined muon tracks were required to be associated with at least one hit in the pixel detector, at least five hits in the SCT, no more than three missing hits in the SCT and pixel detectors. Muon tracks within $0.1 < |\eta| < 1.9$ were required to have at least five TRT hits and outliers, with an outliers fraction less than 0.9. For muons with $|\eta| > 1.9$ or $|\eta| < 0.1$, the outliers fraction was selected to be less than 0.9 if the sum of TRT hits and outliers was larger than five. Isolation criteria were imposed within the calorimeter and the inner detector. The transverse energy in the calorimeter within a cone of $\Delta R = 0.2$ around muon was required to be less than 4 GeV, and the sum of transverse momenta of all tracks (excluding the muon) in a cone of $\Delta R = 0.3$ was constrained to be less than 2.5 GeV. Muons within $\Delta R(\mathrm{jet}, \mathrm{muon}) < 0.4$ of a jet with $|JVF| > 0.5$ and $p_{\mathrm{T,jet}} > 25$ GeV were rejected.[28] [29]

Jets: Jets were reconstructed using the anti-k_{t} algorithm with $\Delta R = 0.4$. The response of the calorimeter is corrected through a p_{T} and η-dependent factor, called jet energy scale (JES), which is derived from simulated events and validated with data. The jet energy scale allows to relate the energy and transverse momentum of an object as measured in the calorimeter to the hadron level. It includes corrections for a variety of instrumental and detector effects, like the intrinsically different response of the calorimeter to electromagnetic and hadronic deposits, reconstruction effects, dead material, noise and pile-up. The baseline for the jet energy scale determination is the electromagnetic (EM) scale, as determined from the response of electrons in the Liquid Argon and Tile calorimeters in test beam data. After that jets are corrected on two levels, via the properties of the clusters or cells and using the jet kinematics. More details can be found in [30, 31].

Jets from additional proton-proton interactions were suppressed by requiring the jet vertex fraction (JVF) to be larger than 0.5. The JVF is defined for tracks matched to the jet as the ratio of the scalar sum of transverse momenta of tracks from the primary vertex to that of all tracks. Jets overlapping with identified electron candidates within a cone of $\Delta R < 0.2$ were removed.

Missing transverse energy: The $E_{\mathrm{T}}^{\mathrm{miss}}$ is a measure of the momentum of the escaping neutrinos. It also includes energy loss due to detector in-efficiencies and its resolution, leading to the mismeasurement of the true E_{T} of the final interacting objects. The $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated as the vector sum over all energy clusters in the event, and is further refined by applying object level corrections for the contributions which arise from identified electrons, muons and jets. First, muons, which are not primarily measured by the calorimeter, are included using their momentum as measured from their track. Next, the topological clusters that are associated to electrons are calibrated at the EM-scale. Clusters belonging to jets (but not to electrons) are divided into those that belong to high p_T jets ($p_T > 20\,GeV$) and soft jets ($p_T < 20\,GeV$). High p_T jets are calibrated at the JES, whereas low p_T jets (SoftJets) are calibrated at the EM-scale. The remaining clusters not associated to any object are included at the EM-scale in what is called the CellOut term. The low p_T jet terms and the unassigned clusters are calibrated in the same way. [32]

3.2 Physics Backgrounds

There are some physics channels that have exactly the same final topology as the semi-leptonic signal and therefore they can not be distinguished. Moreover, there are other physics channels that in some cases can lead, for different reasons, to the same final topology as the requested signal although originally they do not have this final state objects. All these physics channels constitute the physics background. For the for presented analysis in this thesis, they can be classified in five groups as following (corresponding cross sections are shown in Table 3.1 [7]) :

- Single top production: Wt, s and t channels. These channels will be important backgrounds, especially the t channel at the LHC (contrary to the Tevatron) as they have high cross sections [33]. In case of $t\bar{t}$, the production rate of the single top processes is also be larger by a factor around 100 compared to the Tevatron. The main contribution comes from the t-channel and the s-channel is the most suppressed one. As can be seen in Figure 3.1 in case of additional gluon radiation and the leptonic decay of the associated W boson final topologies of these three channels are similar to $t\bar{t}$ channel. The contribution in comparison to W + jets is smaller.

- An important physics process at LHC is the production of $W+$ jets with $W \rightarrow \ell\nu$. The production of a W in association with at least four jets directly results the same final states as the semi-leptonic top quark pair decay. This process turns out to be one of the main backgrounds for top quark pairs which requires

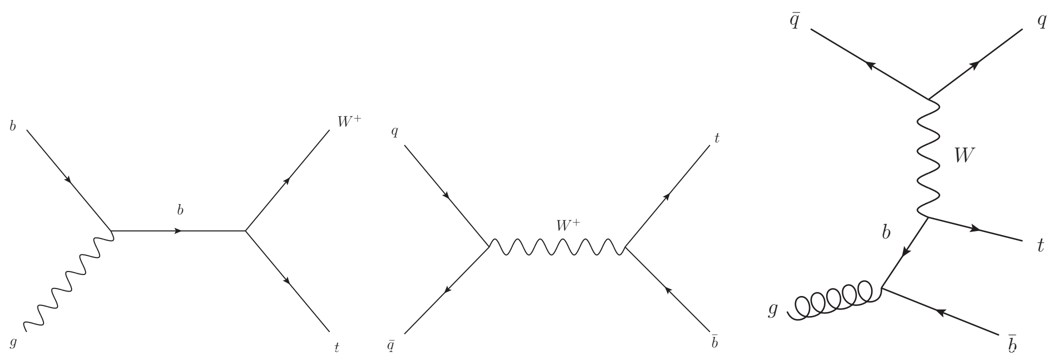


Figure 3.1: Feynman diagrams for single top production.

a thorough understanding of such W + n jets events (from now on W + jets). Figure 3.2 shows Feynman diagrams of this channel and taking into account the 3 type of lepton and the number of jets in the final state this background can be classified into 12 sub-channels: $e\nu_e + X$ jets, $\mu\nu_\mu + X$ jets and $\tau\nu_\tau + X$ jets, with $X = 2$, 3, 4 or 5. In the this analysis, we consider only final states with e or μ. In cases of $W \to \tau\nu$, contribution comes only τ decays to e or μ.

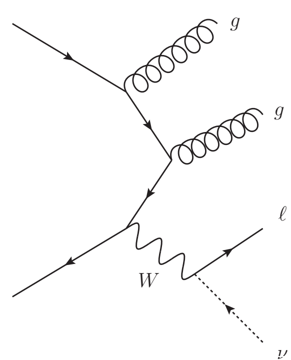


Figure 3.2: Feynman diagram of W+partons production with $W \to \ell\nu$.

- Gauge boson pair production: WW, WZ and ZZ (Figure 3.3 for Feynman diagrams). These channels are also classified as expected backgrounds. The largest contribution corresponds to W^+W^- production. this channel can fake one W boson decays leptonically and the other hadronically, together with ISR or FSR jets to produce the same final state as $t\bar{t}$.

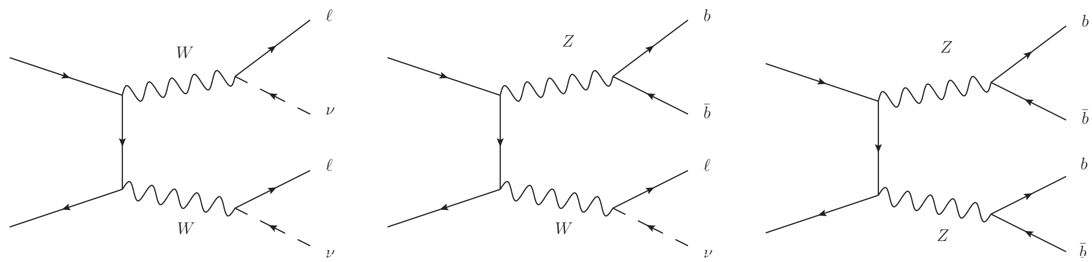


Figure 3.3: Feynman diagrams for the gauge boson pair production (WW, WZ and ZZ).

- Z+ jets events with $Z \rightarrow \ell\ell$ (Figure 3.4 for its Feynman diagram). The production of a Z boson in association with jets has a contribution almost 10 smaller than W boson just due to the branching ratio for the decay of the boson into leptons.

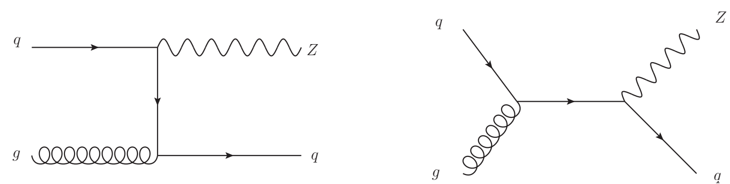


Figure 3.4: Feynman diagram of Z+partons events with $Z \rightarrow \ell\ell$.

-

- Multijets : The production of multijet events (QCD) which have an extremely large cross sections are another important source of background . A misreconstructioned jet can be misidentified as charged electron or muon and a neutrino, that also leads to an artificial missing transverse energy. This leads to a similar final state as the semi-leptonic top quark decay. Since there is no accurate theoretical prediction of production cross section, a large efforts needs to estimate and suppress the contribution of this background. The QCD would be measured directly in data.

Process	Cross section (pb)
$t\bar{t}(full-hadronic)$	90
$t-channel \rightarrow e\nu_e$	6.9
$t-channel \rightarrow \mu\nu_\mu$	6.8
$t-channel \rightarrow \tau\nu_\tau$	7.3
$s-channel \rightarrow e\nu_e$	0.5
$s-channel \rightarrow \mu\nu_\mu$	0.49
$s-channel \rightarrow \tau\nu_\tau$	0.52
Wt	16
$W+0$ jet	8300
$W+1$ jet	1600
$W+2$ jet	460
$W+3$ jet	120
$W+4$ jet	31
$W+5$ jet	8
$W+bb+0$ jet	57
$W+bb+1$ jet	43
$W+bb+2$ jet	21
$W+bb+3$ jet	8
$W+cc+0$ jet	153
$W+cc+1$ jet	126
$W+cc+2$ jet	62
$W+cc+3$ jet	20
$W+c+0$ jet	980
$W+c+1$ jet	312
$W+c+2$ jet	77
$W+c+3$ jet	17
$W+c+4$ jet	4
WW	17
WZ	6
ZZ	1
$Z+0$ jet	853
$Z+1$ jet	168
$Z+2$ jet	51
$Z+3$ jet	14
$Z+4$ jet	4
$Z+5$ jet	1

Table 3.1: Cross-sections for various processes including branching ratios and k-factors

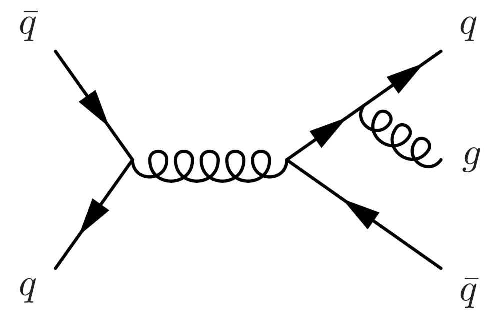


Figure 3.5: Feynman diagram of a dijet event with a gluon radiated in the final state.

3.3 Data driven QCD Methods

This section describes concepts of two methods to estimate multijet background and more detail of usage of QCD for shape and normalization uncertainty come in corresponding analysis. These methods are aimed not only to obtain a proper normalization of the QCD background but also to provide a model which can be used in multivariate analysis techniques to describe QCD shape.

3.3.1 Matrix Method for QCD Background Estimate

The matrix method (MM) exploits differences in lepton identification-related properties between prompt isolated leptons from W and Z decays (referred to as "real leptons" below) and those where the leptons are either non-isolated or result form misidentification of photons/jets (referred to as "fake leptons" below). For this purpose, two samples are defined after requiring the final kinematic selection criteria, differing only in the lepton identification criteria: a "tight" sample and a "loose" sample, the former being a subset of the latter. The tight selection typically employs the final lepton identification criteria used in the analysis, whereas the loose selection is adjusted in order to satisfy basic requirements for the method to work, which are outline below.[34]

The method assumes that the number of selected events in each sample (N^{loose}

and N^{tight}) can be expressed as a linear combination of the numbers of events with real and fake leptons, in such a way that the following system of equations can be defined:

$$N^{loose} = N^{loose}_{real} + N^{loose}_{fake} \tag{3.1}$$

$$N^{tight} = \varepsilon_{real} N^{loose}_{real} + \varepsilon_{fake} N^{loose}_{fake} \tag{3.2}$$

where ε_{real} (ε_{fake}) represents the probability for a real(fake) lepton that satisfies the loose criteria, to also satisfy the tight ones. Therefore, these efficiencies are defined as:

$$\varepsilon_{real} = N^{tight}_{real} / N^{loose}_{real} \tag{3.3}$$

$$\varepsilon_{fake} = N^{tight}_{fake} / N^{loose}_{fake} \tag{3.4}$$

Assuming ε_{real} and ε_{fake} can be estimated in control samples, the system of equations in Equation. 3.1 can be solved yielding the estimated number of events with a fake lepton in the tight sample:

$$N^{tight}_{fake} = \frac{\varepsilon_{fake}}{\varepsilon_{real} + \varepsilon_{fake}} (\varepsilon_{real} N^{loose} + N^{tight}) \tag{3.5}$$

The control sample used to measure ε_{real} is inclusive $Z \rightarrow \ell^+\ell^-$ events using tag and probe. The measurement of ε_{fake} requires selecting a sample enriched in QCD multijet events. Control region for the fake rate estimate is obtained by selecting events with one lepton, at least one jet and low missing transverse energy ($E_T^{miss} <$ 10 GeV).

3.3.2 Jet-electron method

The estimation of QCD background by matrix method for the e+jets channel is more difficult than for the μ+jets channel. The fake muon backgrounds are dominated by heavy flavour jets, whereas photon conversions and jets with high electromagnetic fraction also contribute in the electron channel. Since the composition of multijet backgrounds may be different in the control and signal regions in e+jets, the extrapolation of the fake efficiencies measured in the control samples to the signal one required by the matrix method may introduce a bias. Therefore an alternative method is used for the electron channel.

The Jet-electron method establishes a shape for the QCD background by selecting events with similar kinematics to the signal selection but with a jet used in place of the electron. The normalization is performed by fitting in the E_T^{miss} distribution. The model is using a generic jet sample, selected with a Jet trigger. Since a jet has to fake an electron, the ET threshold has to be lower than the E_T of the signal leptons. The jet, that resembles an electron has to have $E_T > 25$ GeV and the same coverage in h as the signal electrons. The fraction of the electromagnetic energy of the jet f_{EM} has to be between 0.8 and 0.95. By demanding that at least four tracks have been found in the jet, the amount of converted photons is reduced. In Table 3.2 the applied cuts are summarized. As jet-electrons are jet objects but the fakes that are modeled are leptons the missing transverse energy in wrongly calibrated. This is corrected for by scaling the jet-electron to the electromagnetic scale in the E_T^{miss} calculation. To estimate the number of expected QCD events both approaches utilize a binned likelihood fit. Several kinematic variables sensitive to QCD contribution are investigated and finally have been chosen to use the E_T^{miss} distribution. The fit

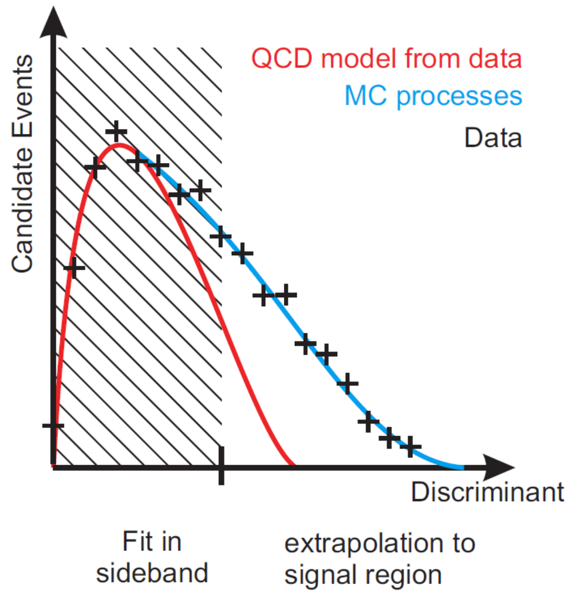


Figure 3.6: This plot shows the general concept how to determine the rate of the QCD-multijet background. A binned likelihood fit is performed, using the shapes of the data-driven fake-lepton model and the shapes of the real W events from MC simulation. Both shapes are fitted in a sideband to the data distribution of a discriminating variable, e.g the E^T_{miss}. [2]

itself is performed only in the side-band region $E^T_{miss} < 20$ GeV in order to be able to perform in the later stage of the analysis a fit in the signal region to extract the signal. The fit is performed before the triangle cut in order to allow for sufficient statistics in the sideband. The cut is then applied considering the cut efficiencies for each template separately. they are called anti-electrons.[2]

Variable	Cut
Trigger	EF Jet20
Transverse energy of jet	$E_T > 20$ GeV
η of jet	$\|\eta\| < 2.47$, no CRACK region
EM fraction	$0.8 < f_{EM} < 0.95$
Number of tracks within the jet	$n_{tracks} > 3$

Table 3.2: Cuts applied to define a jet-electron sample

Chapter 4

Multivariate Likelihood analysis

It is common to all analysis approaches that aim at measuring the $t\bar{t}$ cross section in the single-lepton channel to start with imposing some kinematic cuts. These preselection cuts assist to get a signal enriched sample.

In the Standard Model, top quarks decay to W bosons and b-quarks with probability close to 100%. If one of the W bosons decays leptonically (either to electron or muon with corresponding neutrino) and the other - hadronically, one can naively expect the final state that is characterized by the presence of a high p_T lepton, large missing transverse energy taken away by neutrino and four jets, two of which are from hadronic decay of W boson and two b-quark jets from top quarks decays. The number of jets can actually vary significantly due to several reasons. First, additional jets can arise in the event due to initial and final state gluon radiation. Initial state radiation is the process when the initial partons emit gluons before interaction, which leads to the $t\bar{t}$ production. Final state radiation is the process when either one of the produced top quarks or both of them emit gluons. If these emitted gluons have sufficient energy, they will appear as a jets after hadronization, leading to the increased number of jets in the event with respect to naive expectation of four jets from $t\bar{t}$ decays. The number of reconstructed jets can also be smaller than 4 because of different limitations in the detector coverage and reconstruction algorithms.

Preselection cuts are designed to select events with a high-p_T lepton, substantial amount of missing transverse energy in the final stage and at least three jets.

Events selected in this way originate not only from $t\bar{t}$ production, but from several other processes. These are vector boson production in association with jets, vector boson pair production in association with jets, single top production, and multijet production due to different detector related misidentifications.

To extract the signal cross section, one needs to separate the signal from all backgrounds in the preselected dataset. The simplest approach, called "cut and count method", is based on subtracting the expected background from the preselected sample. The expected background is calculated using theoretical cross sections for all physics processes except multijet background. The latter is estimated using data driven methods. However cut-and-count method has limited accuracy, in particular due to large systematic uncertainties in QCD and W+jets normalization. These uncertainties propagate to the final systematic uncertainty on the inclusive $t\bar{t}$ cross section measurement, making it quite large.

To overcome this limitation and perform more precise measurements of the $t\bar{t}$ fraction in data, multivariate likelihood approach is used in this thesis. This approach exploits the differences in the kinematics of $t\bar{t}$ and background events. Events with exactly 3 jets have larger background compared to the events with 4 jets, so it is practical to construct the likelihood discriminants in different jet multiplicity bins. After performing the measurements in separate bins, one can perform a global fit to all jet bins to get the final value of the $t\bar{t}$ cross section. This combined measurement in the different jet bins reduces statistical and systematic uncertainties significantly.

4.1 Choice of Discrimination Variables

None of the topological variables alone has enough discrimination power allowing them in order to separate the $t\bar{t}$ signal from the background. Therefore the information contained in a set of topological variables is combined into a single likelihood discriminant. More than 85 of these topological variables were studied for their discrimination power, their correlations and their sensitivity against systematic uncertainties. The criteria for the choice of the topological variables are as following:

- Variables should have a good separation power between signal and background.
- Variables should be uncorrelated or have pretty small correlations. Correlated variables only add new information if the correlation is different for signal and background. Since the variables are subject to statistical uncertainty, they increase only the statistical uncertainties if they add no new information.

To evaluate the separation power of discriminating variables, the parameter $< S_k^2 >$ of a variable k, defined by the equation 4.1, is calculated.

$$< S_k^2 >= \frac{1}{2} \int \frac{(p_{S,k}(x_k) - p_{B,k}(x_k))^2}{(p_{S,k}(x_k) + p_{B,k}(x_k))} dx_k \tag{4.1}$$

Here $p_{S,k}(x_k)$ and $p_{B,k}(x_k)$ are the signal and background probability density functions (PDFs) of variable k. For identical signal and background shapes the separation parameter $< S_k^2 >= 0$,while for the case without overlap of PDFs, $< S_k^2 >= 1$.

For a projective likelihood discriminant it is important that input variables are uncorrelated. Correlation between variables x_i and x_j is usually measured by the correlation coefficient, defined by 4.2

$$\rho(x_i, x_j) = \frac{cov(x_i, x_j)}{\sigma_{xi}\sigma_{xj}} \tag{4.2}$$

Here $cov(x_i, x_j)$ is the covariance between variable x_i and x_j, $\sigma_{xi(j)}$ is the standard deviation of variable $x_{i(j)}$. The correlation coefficient is symmetric regarding to x_i and x_j and lies within the interval $[-1, 1]$. The correlation coefficient quantities by definition have a linear relationship. Thus $\rho(x_i, x_j) = 0$ holds for independent variables while ± 1 expresses high correlation or anti-correlation between two variables.

4.2 Definition of Topological Variables

The topological variables can be divided into two categories:

- Variables which are proportional to the energy present in the event. These variables have a very good separation power. However these variables are also very sensitive to uncertainties in the jet energy scale, that translates into a large systematic uncertainty in the cross section measurement.

- The other class of variables is characterized by the property of having reduced sensitivity to JES uncertainties. These variables are either angular quantities or different types of ratios of energy dependent variables where the JES systematic uncertainty largely cancels out.

The topological variables that have been studied as inputs to the likelihood discriminant are constructed to include only up to four leading jets. This procedure reduces the dependence on systematic effects on the modeling of soft radiation such as initial and final state radiation via the parton shower and pile-up jets. The list of best 24 variables are given in Table 4.1. The first column shows the variables name as used in samples, second and third columns shows separation in e+jets and μ+jets channel in four inclusive jet multiplicity bin (events with at least four jets in the final state). The last column gives the definitions of the variables as following:

- The jets are ordered in the samples by their transverse momenta p_T. The jet with the highest p_T is defined as leading jet and its p_T denoted as $p_T(J1)$ and p_T of second leading jet as $p_T(J2)$ and so on.

- M_T^{Jets}, transverse mass of up to four leading jets. Calculated as the total transverse mass of the good jet in the event according to formula 4.3. To reduce the dependence on systematic effects that come from the modeling of soft radiation (e.g. underlying event, initial state radiation via parton shower, etc) only up to four of the leading jets are used to determine the topological variables.

$$M_T^{Jets} = \sqrt{(\sum_{j=0}^{NJets} E_j)^2 + (\sum_{j=0}^{NJets} p_{T_j})^2} \tag{4.3}$$

- H_T, scalar sum of transverse energies of up to four leading jets $\sum(p_{T,i})$.
- $\Delta R_{min}(jet, jet)$, the minimum distance between any two jets in $\eta - \phi$ space.
- $K_T^{min} = \Delta R_{min} \times E_T^{min}/E_W^T$, distance in $\eta - \phi$ space between the closest pair of jets, multiplied by transverse energy E_T^{min} of the less energetic jet of that pair, divided by the scalar sum of the transverse energies of the lepton and the event missing transverse energy $E_T^W = E_T^{lepton} + E_T^{miss}$.
- $p_T(Ji/Jj)$, Ratio of transverse momenta p_T of Jets i over j
- $|\eta_{jet}|^{max}$, largest pseudorapidity.
- $\Delta R_{min}(leading - jet, lepton)$, the distance between pseudorapidity of leading and lepton pseudorapidity in $\eta - \phi$ space
- $\sum(\eta_{jet})^2$, sum of the squared pseudorapidities of up to four jets.
- E_T^{miss}, the event missing transverse energy.
- η_ℓ, where η_ℓ is pseudorapidity of lepton.

- charge of the lepton

- Event aplanarity, $A = \frac{3}{2} \times \lambda_3$, where λ_3 is the smallest eigenvalue of the normalized momentum tensor. The normalized momentum tensor M is defined as

$$M_{ij} = \frac{\sum_n p_i^n p_j^n}{\sum_n |\overrightarrow{p^n}|^2}, \tag{4.4}$$

  where $|\overrightarrow{p^n}|$ is the momentum-vector of a reconstructed object of n, and i and j are Cartesian coordinates. By standard diagonalization of M_{ij} one may find three eigenvalues $\lambda_1 > \lambda_2 > \lambda_3$, with $\lambda_1 + \lambda_2 + \lambda_3$=1. The aplanarity A is a measure of the flatness of the event. The objects included in the sum are the jets and the lepton from the W decay. Large values of A are indicative of spherical events, whereas small values correspond to more planar events. The transformed aplanarity $exp(-8A)$ is used further since it is less sensitive to statistical fluctuations.

- The event planarity $= \frac{3}{2} \times (\lambda_1 + \lambda_2)$

- The event sphericity $= \frac{3}{2} \times (\lambda_3 + \lambda_2)$

- The event centrality $= \frac{H_T}{H}$, where H_T is the scalar sum of the transverse energy of the jets and H is the scalar sum of the energy of the jets.

The distributions for these variables overlaid for $t\bar{t}$ signal and W+jets background are shown in Fig. 4.1 and 4.2 for e+jets channel and Figures 4.3 and 4.4 for μ+jets channel. Correlation matrices are shown in Fig. 4.5 for $t\bar{t}$ signal and in Fig. 4.6 for W+jets background in e+jets channel. The same distributions in the muon channel are shown in Fig. 4.7 and Fig. 4.8.

Variable name	e+jets$< S^2 >$	μ+jets$< S^2 >$	Definition
LeptonEta	5.961e-02	5.871e-02	Lepton pseudorapidity, η_ℓ
LeptonCharge	1.025e-02	1.274e-02	Charge of the lepton
J1Pt	5.043e-02	4.730e-02	$p_T(J1)$
J2Pt	6.381e-02	6.296e-02	$p_T(J2)$
J3Pt	7.790e-02	8.532e-02	$p_T(J3)$
J4Pt	6.889e-02	8.310e-02	$p_T(J4)$
ExpAplanarity	6.550e-02	6.114e-02	$exp\,(-8A)$
planarity	1.054e-02	1.176e-02	The event planarity
centrality	2.821e-02	1.996e-02	The event centrality
sphericity	3.687e-02	3.961e-02	The event sphericity
MET	1.772e-02	2.036e-02	E_T^{miss}
JetMt	2.909e-02	2.632e-02	M_T^{Jets}
Ht	5.707e-02	6.070e-02	H_T
Ptratio_J3J1	4.254e-02	3.349e-02	$p_T(J3/J1)$
Ptratio_J4J1	3.889e-02	3.190e-02	$p_T(J4/J1)$
Ptratio_J2J1	1.992e-02	1.486e-02	$p_T(J2/J1)$
Ptratio_J1_J2toJ4	3.393e-02	2.678e-02	$p_T(J1+J2/J1)$
dRJ1J2	3.742e-02	3.190e-02	$\Delta R(leading-jet, 2^{th} jet)$
dRJ1J4	3.742e-02	3.190e-02	$\Delta R(leading-jet, 4^{th} jet)$
DRMin	3.051e-02	2.625e-02	$\Delta R_{min}(jet, jet)$
dRLeptonJ1	2.018e-02	1.633e-02	$\Delta R(leading-jet, lepton)$
EtaSqSum	2.813e-02	2.017e-02	$\sum(\eta_{jet})^2$
EtaMax	2.716e-02	1.768e-02	$\lvert\eta_{jet}\rvert^{max}$
KtMinP	3.731e-02	3.918e-02	K_T^{min}

Table 4.1: Separation power of best 24 variables that can be used for the likelihood discriminant.

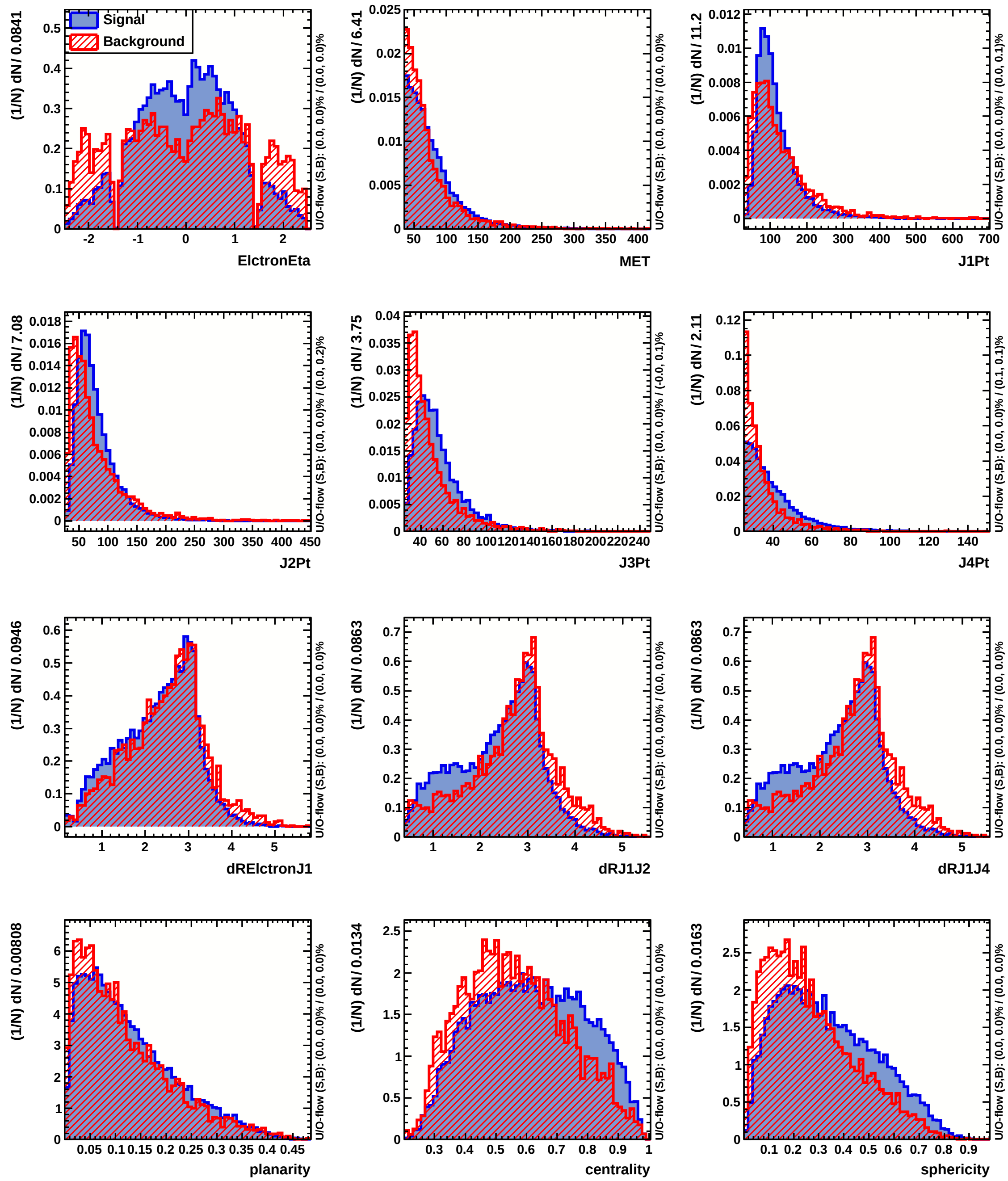


Figure 4.1: Distributions of the discriminant variables in the $e+ \geq 4$ jets channel. The $t\bar{t}$ signal is shown by blue solid line, and W+jets backgrounds - by red dashed line.

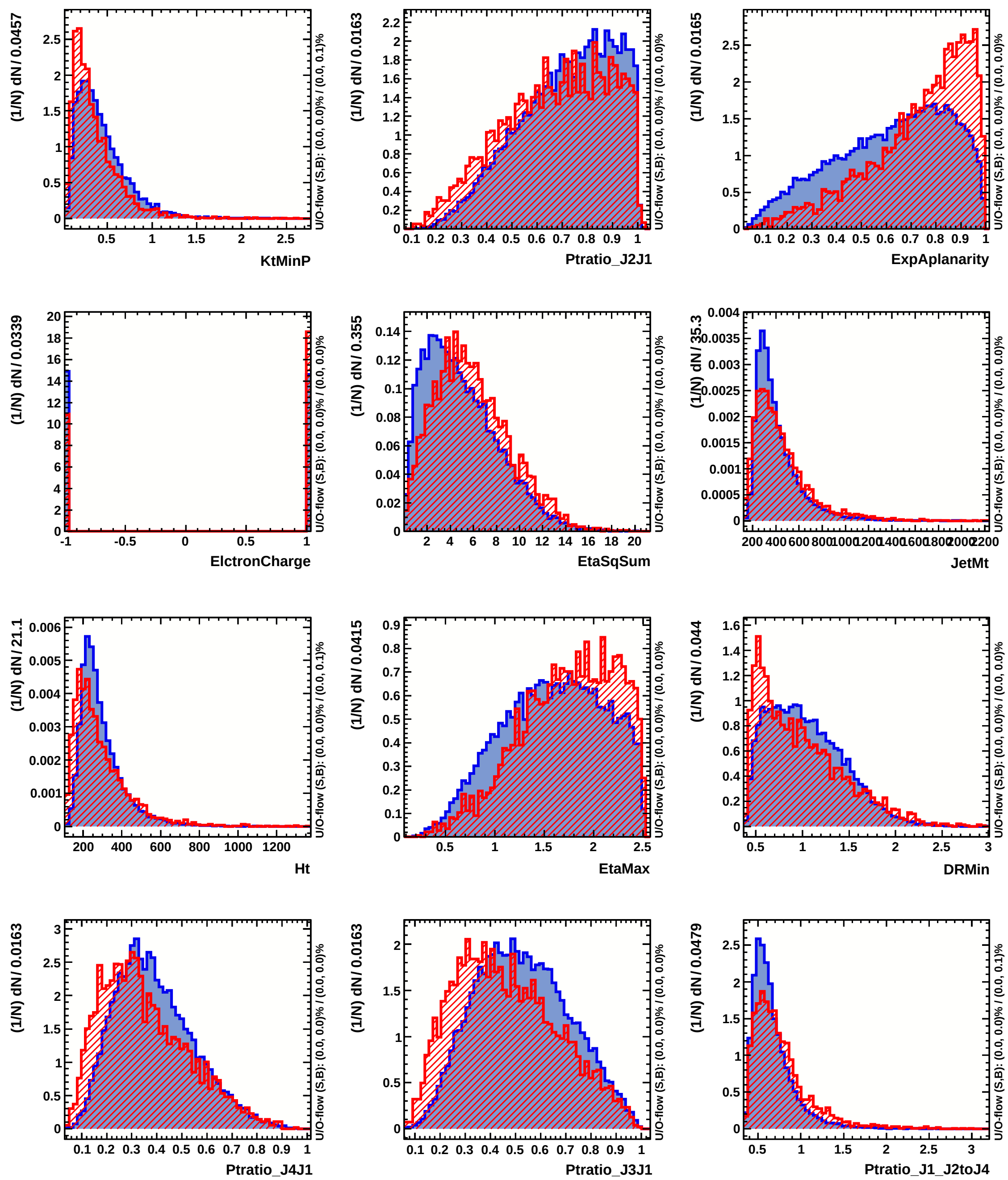


Figure 4.2: Distributions of the discriminant variables in the $e+ \geq 4$ jets channel. The $t\bar{t}$ signal is shown by blue solid line and combined W+jets background by red dashed line

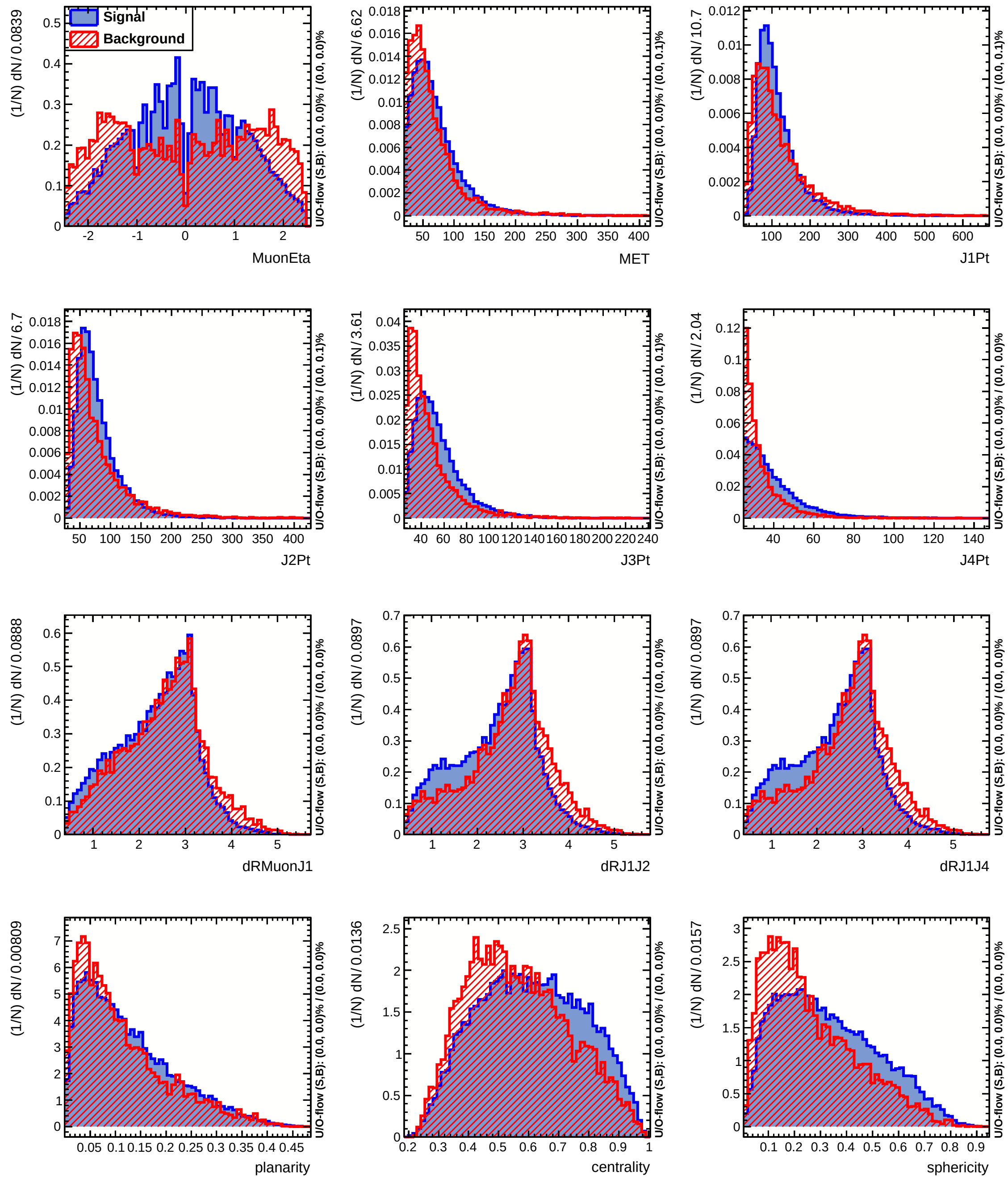


Figure 4.3: Distributions of the discriminant variables for events with four or more jets μ+jets channel. The $t\bar{t}$ signal is shown by blue solid line and combined W+jets background - by red dashed line.

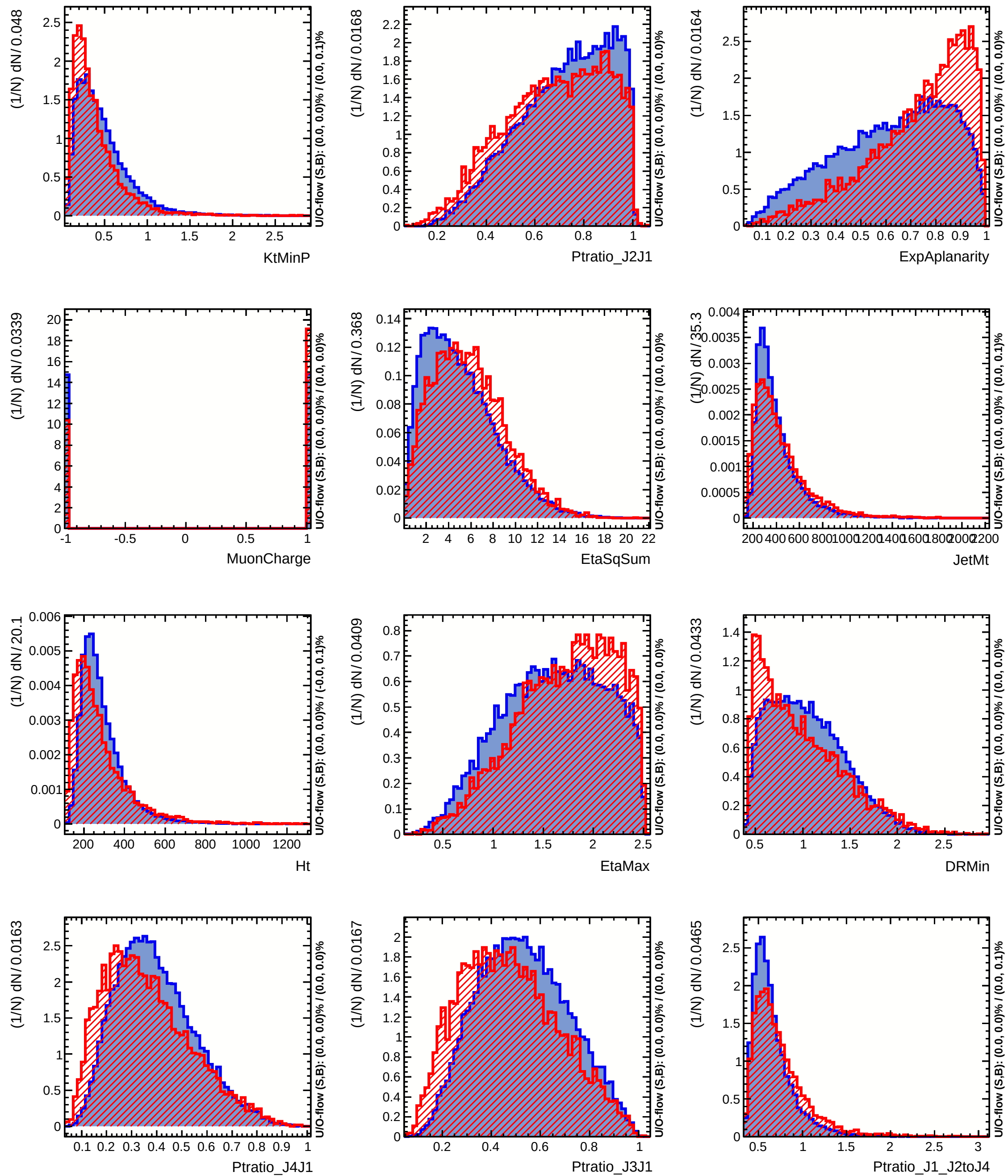


Figure 4.4: Distributions of the discriminant variables for events with four jets or more μ+jets channel. The $t\bar{t}$ signal is shown by blue solid line and combined W+jets background - by red dashed line.

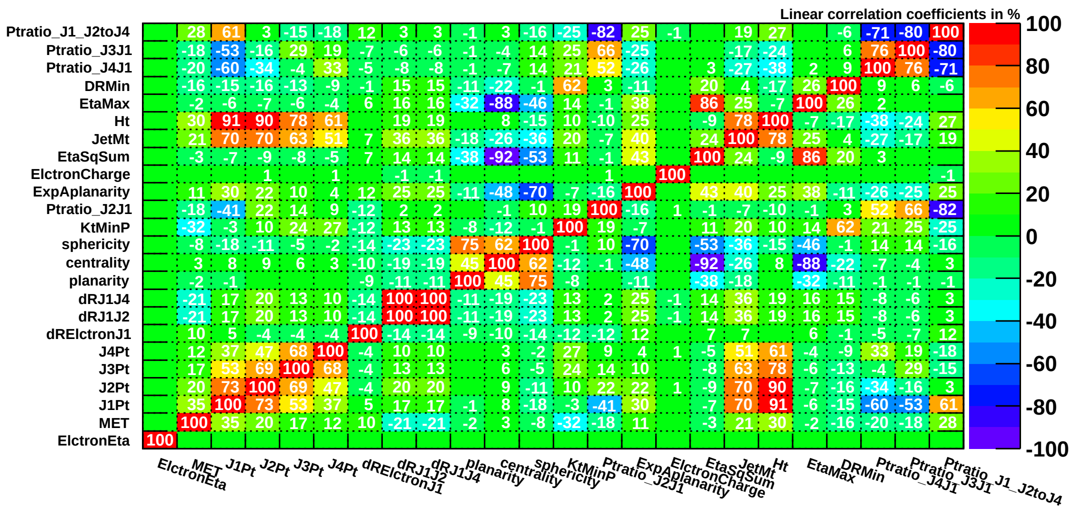


Figure 4.5: Correlation matrix of the discriminant variables for $t\bar{t}$ signal events with four or more jets in the electron channel.

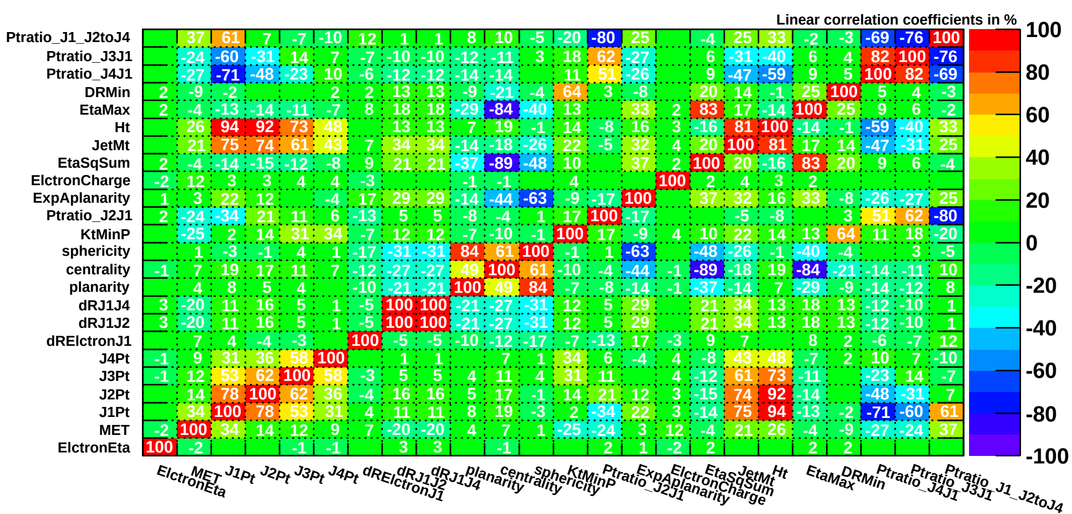


Figure 4.6: Correlation matrix of the discriminant variables for W+jets events with four or more jets in the electron channel.

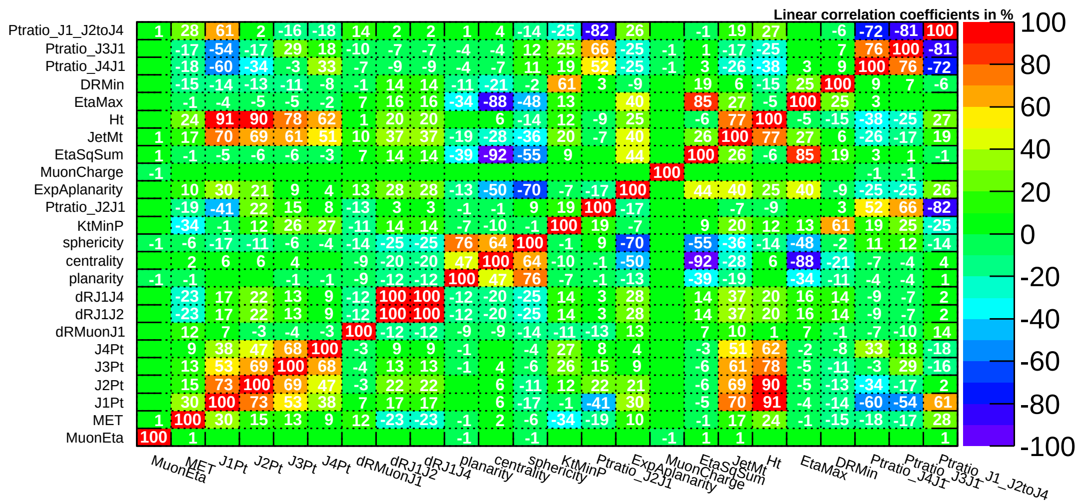


Figure 4.7: Correlation matrix of the discriminant variables for W+jets events with four or more jets in the muon channel.

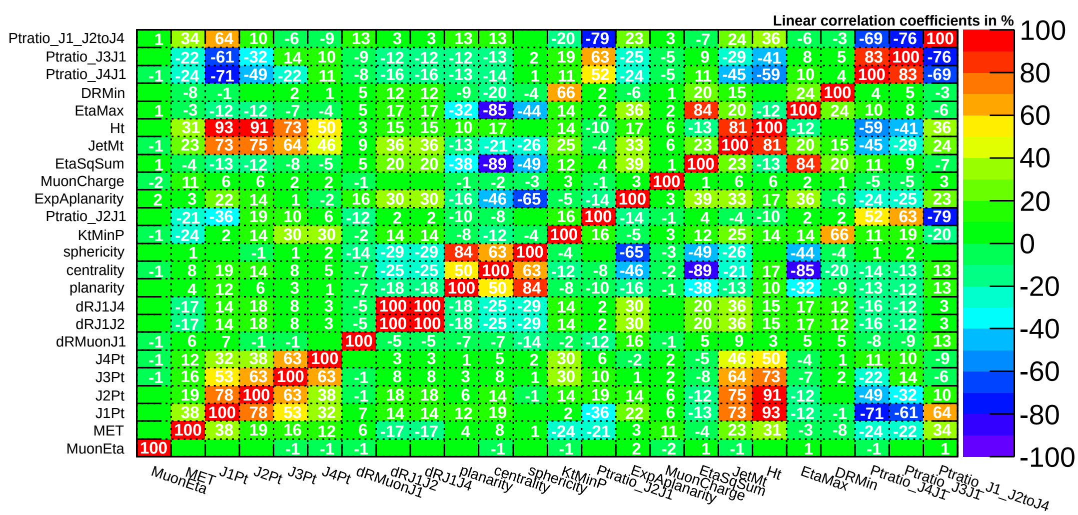


Figure 4.8: Correlation matrix of the discriminant variables for $t\bar{t}$ signal events with four or more jets in the muon channel.

4.3 Multivariate Likelihood Discriminant

The likelihood discriminant is created from input variables by following the projective likelihood approach defined in the TMVA package [35].

The likelihood discriminant D_i for an event i is defined as the ratio of the signal $\mathcal{L}_S(i)$ to the signal plus background $\mathcal{L}_S(i) + \mathcal{L}_B(i)$ likelihoods:

$$D_i = \frac{\mathcal{L}_S(i)}{\mathcal{L}_S(i) + \mathcal{L}_B(i)} \tag{4.5}$$

The individual likelihoods are products of the corresponding probability density functions of the discriminating input variables x_k defined above. For example for the signal likelihood:

$$\mathcal{L}_S(i) = \prod_{k=1}^{n_{variable}} p_{S,k}(x_k(i)) \tag{4.6}$$

where $p_{S,k}(x_k(i))$ is the signal (background) PDF for the kth input variable x_k. PDFs are normalized to unity. This simple multivariate approach assumes that the variables x_k are uncorrelated.

During the training, the $t\bar{t}$ events are used as signal, and W+jets events are treated as background, since W+jets process is the dominant background and the choice of likelihood variables is driven by the separation power between $t\bar{t}$ and W+jets. Adding other background processes at the training step can decrease the discrimination power of the final likelihood especially if the backgrounds behave similar to $t\bar{t}$. Example of such a background is single top quark production.

4.4 Cross Section Extraction Using Likelihood Fitting

After the likelihood discriminant function LHW is constructed, binned distributions (templates) are made for the $t\bar{t}$ signal and each of the backgrounds. These templates are then fitted to the LHW distribution in the data using the negative log likelihood method. The number of $t\bar{t}$ events in each jet multiplicity bin j is obtained by minimizing the quantity

$$-2ln(\mathcal{L}) = 2\sum_{i,j} \nu_{ij} - n_{ij} + n_{ij}\ln n_{ij} - n_{ij}\ln \nu_{ij} \tag{4.7}$$

where the sum is taken over all channels j and all template histogram bins i, n_{ij} is the observed number of events, and ν_{ij} is the expected number of events. The latter is defined as

$$\nu_{ij} = p_0 s_{ij} + p_j b_{ij} + m_{ij}, \tag{4.8}$$

where s_{ij} ,b_{ij} , and m_{ij} are the predicted numbers of events for the $t\bar{t}$ signal, W+jets, and small backgrounds (Z+jets, single top, dibosons, and QCD), respectively. and p_0, p_j are the parameters of the fit. The resulting value of the fit parameter p_0 characterizes the ratio of the measured to the theoretical $t\bar{t}$ production cross section.

4.5 Ensemble Tests

Both expected statistical and systematic uncertainties for a given likelihood discriminant are determined from ensemble tests, which consist of a series of pseudo-experiments (PEs). Each PE dataset consists of drawing random numbers of events from signal and background distributions assuming that the templates follow a multinomial distribution and the total number of events is Poisson-distributed, or equivalently that the content of each bin of the templates is independently Poisson-distributed. The PEs are then fitted as if they were the real data. In each pseudo experiment the measurement is repeated on a simulated data sample using the same statistics as in the real dataset.

The distribution of the fit results is an estimator of the statistical uncertainty, their shift with respect to the nominal result are used to quantify systematic uncertainties. In general a systematic variation will "shift the templates", i.e. change both shape and normalization of the templates. In practice there are two ways of running ensemble tests to quantify systematic uncertainties. The two main strategies are to either draw PEs from the nominal samples and run the fit using the shifted templates or to draw PEs from the shifted samples and fit with the nominal templates. We have compared the two strategies and found that the resulting systematic uncertainties are compatible.

The event sample used for deriving the templates and the event sample used for deriving the PE dataset should be independent from each other. In current analysis a half of $t\bar{t}$ and W+jets simulated events was used for the construction of the likelihood templates and the other half events used for making PEs.

The common approach used in this thesis is to perform $N = 1000$ pseudo data ensemble tests, when the fit is repeated for each ensemble. The average number of p_0 over all ensembles is taken to be the ratio of measured to injected $t\bar{t}$ cross section, and the RMS of the p_0 distribution gives the statistical uncertainty of the method. The systematic uncertainties are evaluated by replacing the nominal predicted values of s_{ij} ,b_{ij} , and m_{ij} with their modified versions and repeating the fit. The relative systematic uncertainty is defined as the difference between the p_0 values in two cases.

4.6 Optimization of the likelihood Discriminant

Likelihood discriminant constructed with several variables has better separation power than any given variable. It can also allow one to reduce some systematic uncertainties, like JES, leading to a better overall accuracy of the resulting measurement. The discrimination variables selected in section 4.2 were studied to determine the best combination of variables for the final likelihood discriminant.

The optimization procedure for the choice of topological variables is an estimation of the expected statistical and dominant systematic uncertainties using pseudo-experiments as described in previous section. The goal of the optimization is to find the combination of variables which gives the lowest combined statistical and systematic error on the $t\bar{t}$ cross section measurement. In addition, it is desired to use a relatively small number of variables to keep the analysis simple. Usually separation power and systematic error of constructed likelihood discriminant with n variable cannot be predicted by study of each single variable which likelihood function would include. The number of possible likelihood functions that can be constructed with n input variables, is 2^{n-1}. Due to computing time and resource limitation it is thus important to constrain the number of input variables to a manageable number and pick only the most promising variables out of all possible ones.

The first study has been done by randomly constructed likelihood with n variables and ranking the variables by their expected statistical uncertainty and dominant systematic uncertainty (JES). Results of this study has been reflected in selecting the discrimination variables to reduce the initial list of variable to table 4.1. The figures 4.9 and 4.10 show scattered plots of variables versus expected statistical uncertainty and dominant JES uncertainty for $e + jets$ channel.

Other optimization performed is the incremental search. It begins with discriminators made of a single variable than pick discriminators with the best expected statistical uncertainty than it creates all possible (n+1) variable discriminators by adding a new variable preselected n-variable discriminators. Figure 4.11 shows the total expected error as a function of number of variables in the likelihood. The average uncertainty decreases with an increasing number of variables used in likelihood and becomes flat quickly for likelihood functions after 3 or 4 and more input variables

In the final step of the optimization all possible three-variable likelihoods have been constructed. For each likelihood, the σ_{tot} defined by Equation. 4.9 was calculated as the figure of merit. It takes into account the statistical uncertainty (σ_{stat}) plus major source of systematic uncertainty due to JES, (σ_{JES}).

Tables 4.2 and 4.3 show five selected likelihoods with variables used for their construction and related σ_{tot}. We selected a likelihood discriminate with three variables, lepton charge, lepton pseudo-rapidity (η_ℓ) and transformed aplanarity($exp(-8A)$), which they have a good Data/MC agreement and minimal correlation. This likelihood of choice serves the analysis with the best uncertainty performance as well as simplicity.

$$\sigma_{tot} = \sqrt{\sigma_{stat}^2 + \sigma_{JES}^2} \tag{4.9}$$

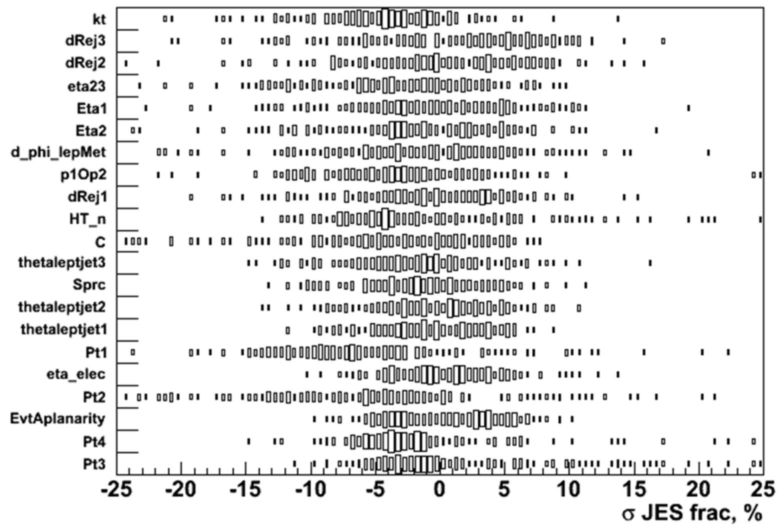


Figure 4.9: Expected JES systematic uncertainty versus input variables used in the likelihood discriminant.

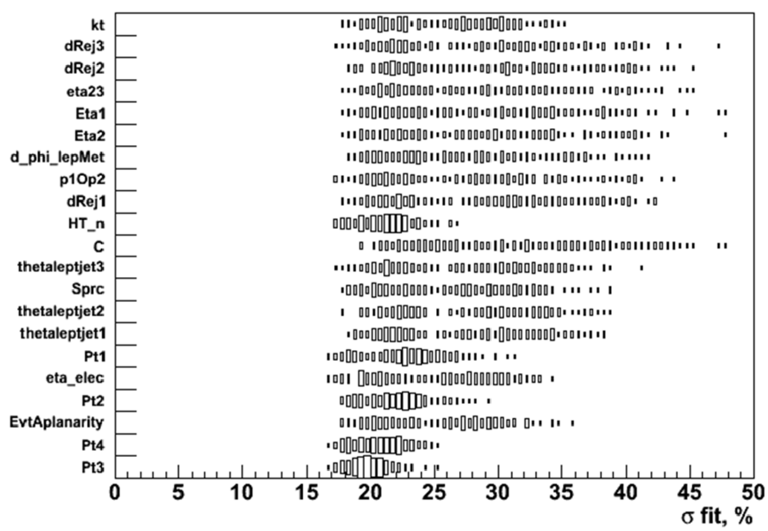


Figure 4.10: Total expected statistical uncertainty versus input variables used in the likelihood discriminant.

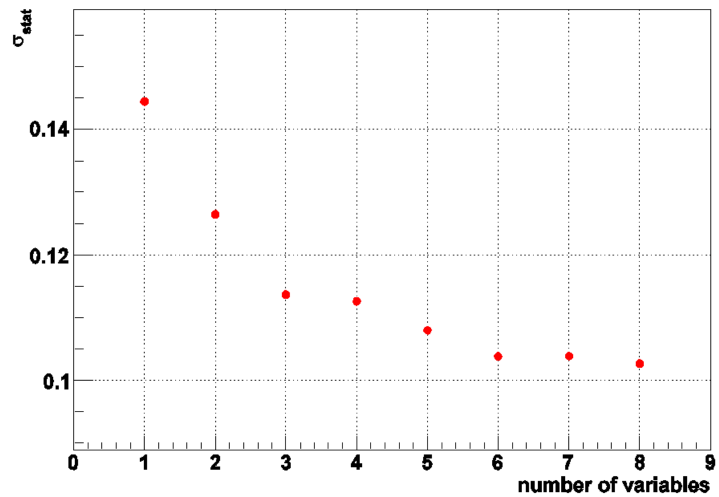


Figure 4.11: Total expected uncertainty versus the number of input variables used in the likelihood discriminant.

$\sigma_{tot}(\%)$	Variables in Likelihood $\mu + jets$ channel
13.83	sphericity, $p_T(J1+J2/J1)$, MuonCharge
14.55	$exp\,(-8A)$, sphericity, $p_T(J1+J4/J1)$
15.86	MuonCharge, $exp\,(-8A)$, η_{muon}
21.8	H_T, sphericity, K_T^{min}
21.2	$p_T(J4/J1)$, E_T^{miss}, MuonCharge

Table 4.2: Overview over the five selected likelihood functions found by the optimization procedure for $\mu + jets$ channel .

$\sigma_{tot}(\%)$	Variables in Likelihood $e + jets$ channel
12.93	$\eta_{electron}$, $p_T(J3)$, $p_T(J1+J4/J1)$
13.85	H_T, $p_T(J4)$, $exp\,(-8A)$
13.43	ElectronCharge, $exp\,(-8A)$, $\eta_{electron}$
20.13	$p_T(J2)$, $p_T(J4)$, ElectronCharge
24.2	$p_T(J2)$, $\sum(\eta_{jet})^2$, $\Delta R(leading-jet, electron)$

Table 4.3: Overview over the five selected likelihood functions found by the optimization procedure for $e + jets$ channel .

4.7 Performance of the Likelihood Discriminant

After the choice of the likelihood discriminant, it is important to check its performance in terms of the expected statistical error and correctness of the fit in the wide region of the possible values of $t\bar{t}$ cross section.

To evaluate the performance of method there are two checks that are usually performed. The first one the check of the pull distribution which is defined as $(p_0 - 1)/\sigma(p_0)$. Here $\sigma(p_0)$ is an uncertainty of an individual fit. The mean of the pull distribution in the adopted definition is expected to be zero. The pull distribution width is a measure of the accuracy of the estimated statistical uncertainty. A width of 1.0 confirms the correct error estimation of the single fit results. The second quality criterion for a method is called linearity test which investigates if the maximum likelihood fit is unbiased. The procedure is using ensemble tests with PEs datasets generated with different "values" of the $t\bar{t}$ cross section. The input $t\bar{t}$ cross section is varied between 50% and 150% of the theoretical $t\bar{t}$ cross section. After that, the cross section from the likelihood fit is compared to the input cross section. In case of the absence of bias in the fitting procedure one can expect the slope to be equal to 1.

Chapter 5

Measurements of $t\bar{t}$ Inclusive Cross Section

This chapter describes a measurement of the inclusive $t\bar{t}$ production cross-section in lepton + jets channel based on the analysis strategy summarized in chapter 4. The measurement is done using 35 pb^{-1} of ATLAS data collected in 2010.

The measurement employs a binned maximum likelihood template fit. Several kinematic variables of the reconstructed collision events are selected to distinguish $t\bar{t}$ signal events from the various background sources. Results of this analysis are published in [36] and presented on all international conferences in 2011.

5.1 Monte Carlo Simulation

The template fitting analysis relies on Monte Carlo (MC) samples for the description of most of the backgrounds and the $t\bar{t}$ signal. The MC $t\bar{t}$ signal sample was generated with the MC@NLO generator [37], assuming a top quark mass of $m_t = 172.5$ GeV, HERWIG [38] fragmentation and the JIMMY [39] underlying event model. Of the dominant backgrounds, vector boson production with associated jets ($W + jets$ and

$Z + jets$) was generated using the Alpgen generator separately for each channel with n additional massless partons, n varying from 0 to 5. The vector boson production in association with heavy flavors, i.e. c or b-quarks was done separately. $Wc\bar{c}$ an $Wb\bar{b}$ processes were generated with n = 0, 1, 2 additional massless partons. Wc was generated with n additional partons, n varying from to 4. The multijet background was evaluated by constructing the QCD templates from data using cut reversal or matrix methods. Events with single top quarks in t, s, and Wt channels were simulated with MC@NLO assuming the top quark mass of $m_t = 172.5$ GeV. Finally, diboson states WW, WZ and ZZ were generated with PYTHIA.

All samples were scaled such that the total number of events corresponds to recent theoretical predictions of the production cross sections at NLO or NNLO [40]. Bunch-train pileup MC samples with an average number of overlaid minimum bias events of two are used. All MC samples were passed through the Geant4 detector simulation. Data sets used for this measurement are listed in the appendix B.

5.2 Event Selection

The events are selected with one isolated high-p_T electron or muon, large missing transverse energy, and at least 3 high-p_T jets. Because of the different contribution of the multijet background to the muon (μ+jets) and electron (e+jets) channels, the requirements for selected events are slightly different. In the electron channel the selected events must have

- an isolated electron with transverse momentum $p_T > 25$ GeV, passed quality cuts and matched to an electron fired the trigger;
- missing transverse energy corrected for the presence of leptons and jet energy scale, $E^T_{miss} > 35$ GeV,

- Transverse mass of W-boson, reconstructed from electron and E_{miss}^T, $m_T(W)>$ 25 GeV,

- At least three well reconstructed and calibrated jets with $p_T > 25$ GeV.

Similarly, events in the muon channel are required to have:

- an isolated muon with $p_T > 20$ GeV, passed all muon quality cuts and matched to a muon trigger object;

- missing transverse energy, corrected for the presence of leptons and jet energy scale, $E_{miss}^T > 20$ GeV;

- Cut E_{miss}^T+$m_T(W)>$ 60GeV;

- At least three well reconstructed and calibrated jets with $p_T > 25$ GeV.

The background from QCD multijet production is determined in a data-driven way. This background enters the analysis through jets that are mis-identified as charged leptons. There are two alternative approaches to determine the multijet contribution to the selected samples. The matrix method (MM) determines the QCD background from the difference in efficiency to reconstruct leptons with loose and tight lepton identification criteria. The QCD fitting method is based on a data-derived model for the QCD background. The shapes of kinematic distributions are determined from the distributions of leptons for which one or more lepton identification criterion is inverted. A fit to a sideband region, typically low E_{miss}^T, fixes the normalization. Unless noted otherwise, the MM is used in the muon channel, while the fitting method is employed in the electron channel. More details on these methods can be found in section 3.3.1 and [41].

5.3 Construction of the Likelihood Discriminant

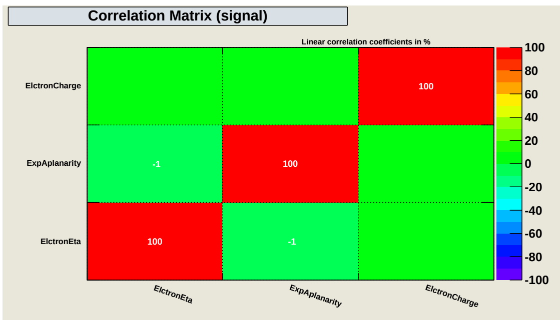
Multivariate analysis is using the maximum likelihood method to measure the $t\bar{t}$ cross section in data. The discriminating functions (likelihood functions, LH) for the signal and background events will be referred to as the signal and background templates, respectively. Data represents the sum of signal and background. The idea is to determine the fraction of the signal-like events by fitting the likelihood distribution in data to the sum of two templates, signal and background. Knowing the initial number of events, fraction of signal, efficiency of the signal selection and integrated luminosity, one can determine the production cross-section of the $t\bar{t}$ events.

The $t\bar{t}$ and $W + jets$ MC samples are used to construct signal and background templates, and use them to fit both real and simulated data. The analysis is done in 3 exclusive and 4 inclusive jet multiplicity bins in e+jets and μ+jets channels. The results of measurements in 3 and 4 jet multiplicity bins are combined together at the later stage.

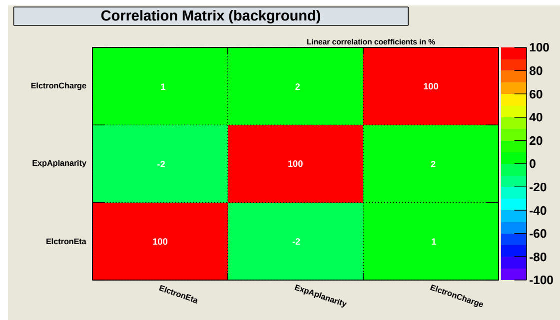
5.3.1 Input Variables

Optimization of the variables was done by minimization of combined statistical and JES error as described in Chapter 5. The best expected performance is achieved using η_{lepton}, charge of the lepton, and $exp(-8 \times A)$, where A is the aplanarity of the event. The correlation matrices for signal and background in e+jets and μ+jets are shown in Fig. 5.1. One can see that these variables are not correlated.

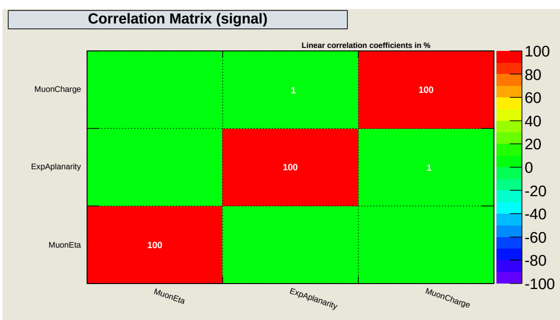
Agreement between data and MC for these three variables is shown in Fig. 5.2 for the electron channel and in Fig. 5.3 for the muon channel. Good agreement between data and sum of all MC backgrounds, QCD data and expected $t\bar{t}$ signal is observed for all three variables. Contribution of $t\bar{t}$ in these plots is evaluated using its theoretical cross section.



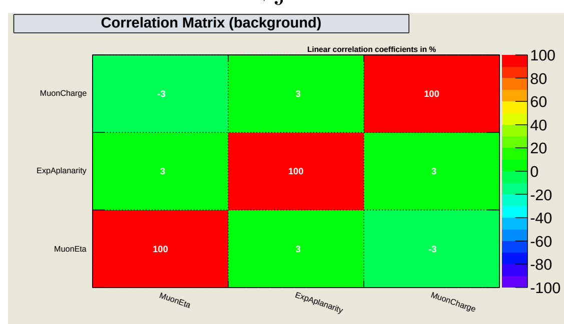
(a) Linear correlations between chosen variables in e+3 jets channel for signal



(b) Linear correlations between chosen variables in $e+ \geq 4$ jets channel for W+jets channel

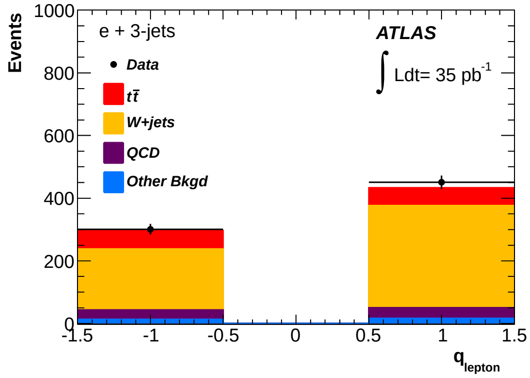


(c) Linear correlations between chosen variables in μ+3 jets channel for signal

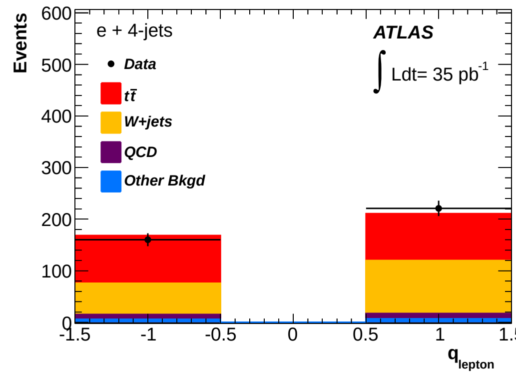


(d) Linear correlations between chosen variables in $\mu+ \geq 4$ jets channel for W+jets channel

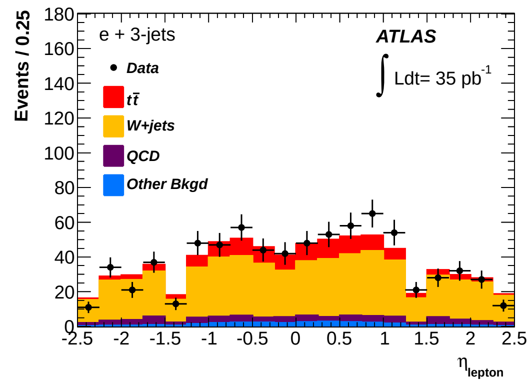
Figure 5.1: Linear correlations between chosen variables.



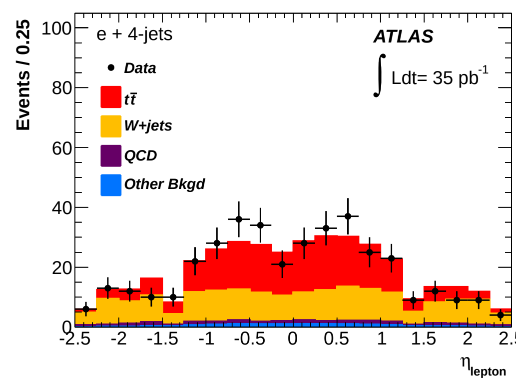
(a) Electron charge in $e + 3$ jets channel



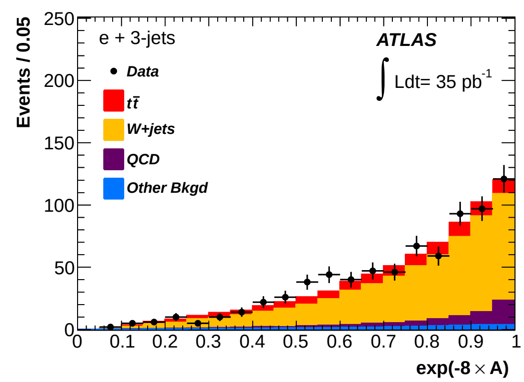
(b) Electron charge in $e+ \geq 4$ jets channel



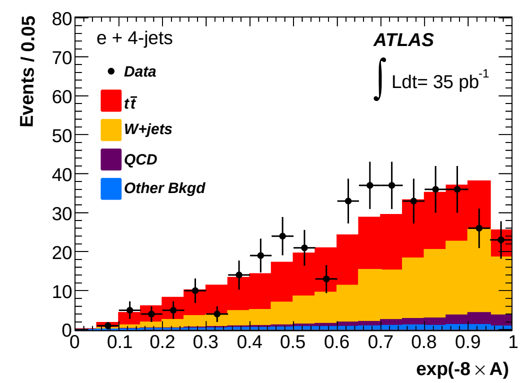
(c) Electron pseudorapidity in $e + 3$ jets channel



(d) Electron pseudorapidity in $e+ \geq 4$ jets channel

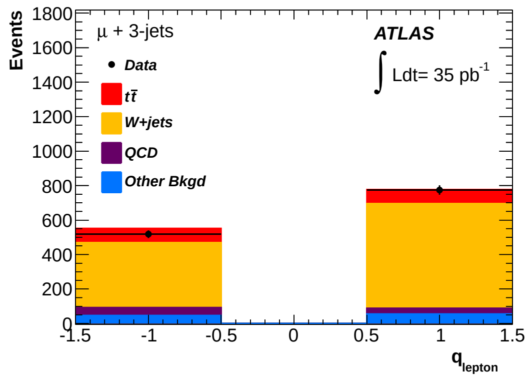


(e) Transformed aplanarity in $e + 3$ jets channel

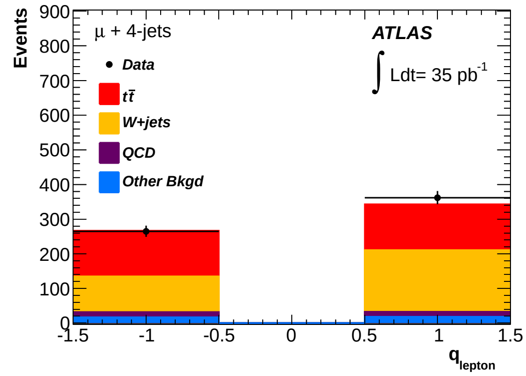


(f) Transformed aplanarity in $e+ \geq 4$ jets channel

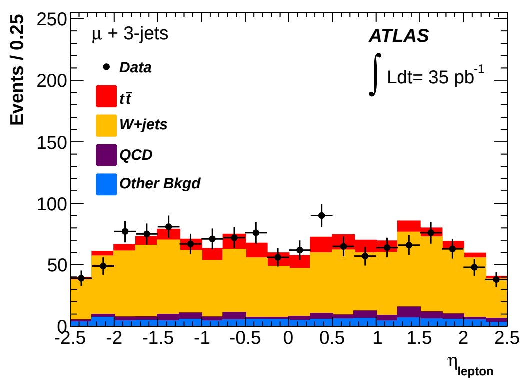
Figure 5.2: Input variables for the likelihood discriminant in e+jets channel.



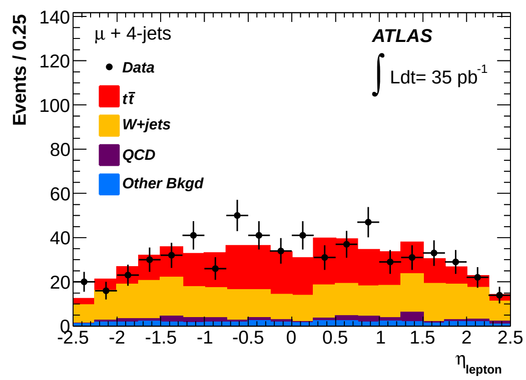
(a) Muon charge in $\mu + 3$ jets channel



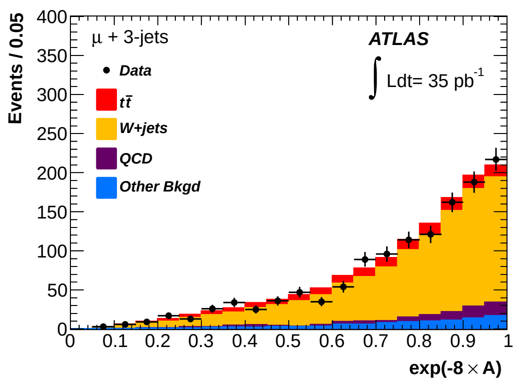
(b) Muon charge in $\mu+ \geq 4$ jets channel



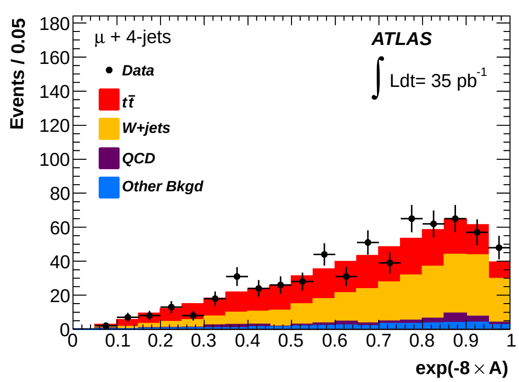
(c)Muon pseudorapidity in $\mu + 3$ jets channel



(d) Muon pseudorapidity in $\mu+ \geq 4$ jets channel



(e) Transformed aplanarity in $\mu + 3$ jets channel



(f) Transformed aplanarity in $\mu+ \geq 4$ jets channel

Figure 5.3: Input variables for the likelihood discriminant in μ+jets channel.

jet multiplicity	Channel	Statistical (%)
3 jet bin	e+jets	39.2
	μ+jets	40.7
4 jet bin	e+jets	17.7
	μ+jets	14.7

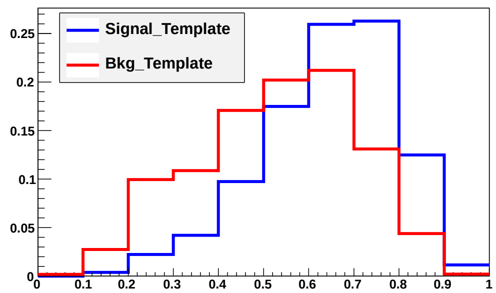
Table 5.1: Expected statistical errors in 3 and 4 jet multiplicity bins

5.3.2 Performance of the Likelihood Discriminant

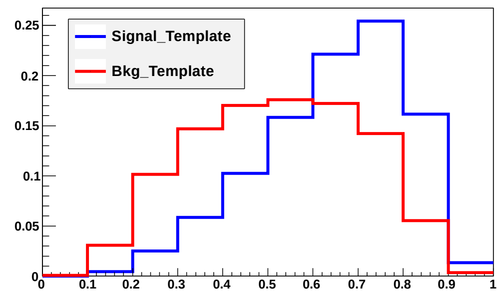
The LH distributions in 3 and 4 jet multiplicity bins are shown in Fig. 5.4 for electron and muon channel, respectively. Performance of the TMVA LH was studied by checking the fitting fraction of $t\bar{t}$ events as a function of input fraction as described in Chapter 5. The results are shown in Fig. 5.5 for both channels. No shifts are observed in fitted fractions with respect to input fractions.

The expected statistical uncertainty is evaluated with ensemble tests on the MC samples. When preparing the ensemble tests, total number of events are fixed to the number observed in data less the number of events due to small backgrounds (all except $W + jets$). The obtained number of events is fixed in each ensemble. The number of $t\bar{t}$ events in the sample is fluctuated according to Poisson distribution around the expected value of top quark events in data. The maximum likelihood fit is then performed on the simulated data (pseudo data) resulting in $N_{t\bar{t}}$. For each test, 1000 pseudo data samples are generated.

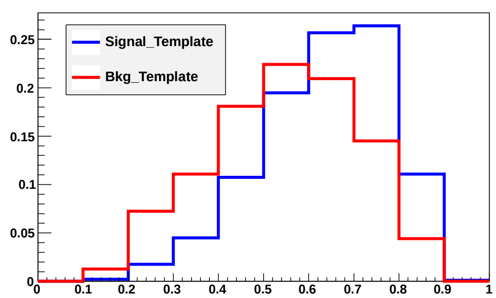
The statistical error is found from the width of the $N_{t\bar{t}}$ distribution found resulting from the fit when applied to pseudo data. Results of ensemble tests for statistical error are shown in Fig. 5.6 for both electron (right) and muon (left) channels. Here the ratio of fitted to input fraction is plotted. The validity of the pseudo data tests can be found by examining the pull distributions in Fig. 5.7. Summary of the expected statistical errors obtained with likelihood ensemble tests is given in Table 5.1.



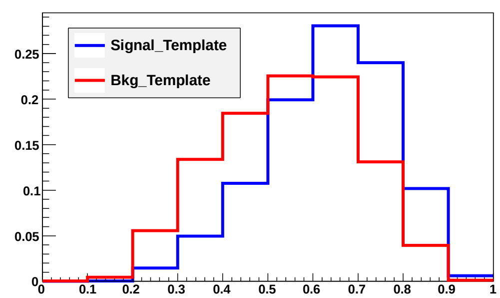
(a) LH templates for $e + 4\,jets$



(b) LH templates for $\mu + 4\,jets$

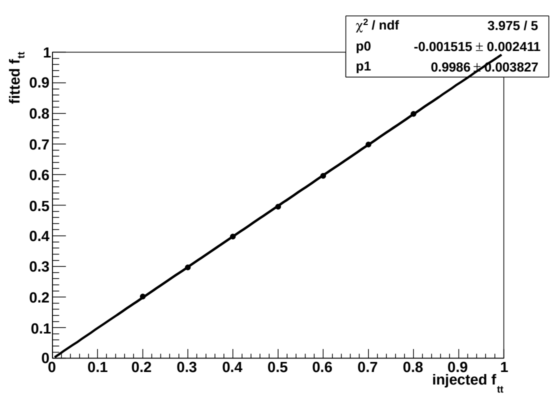


(c) LH templates for $e + 3\,jets$

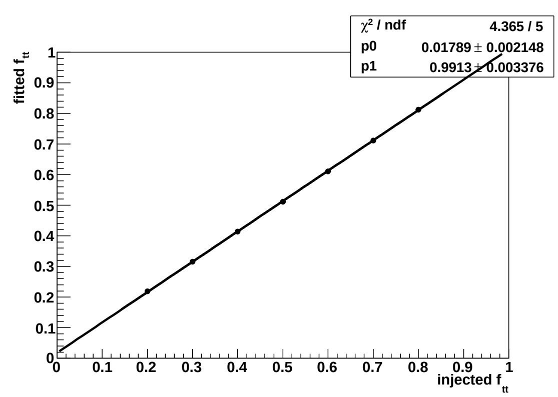


(d) LH templates for $\mu + 3\,jets$

Figure 5.4: LH templates for muon and electron channels.

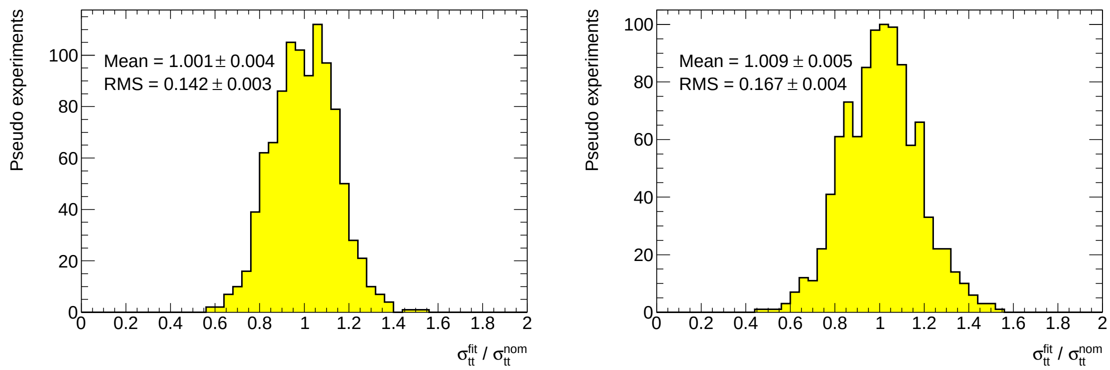


(a) Linearity test in $\mu + 4\,jets$



(b) Linearity test in $e + 4\,jets$

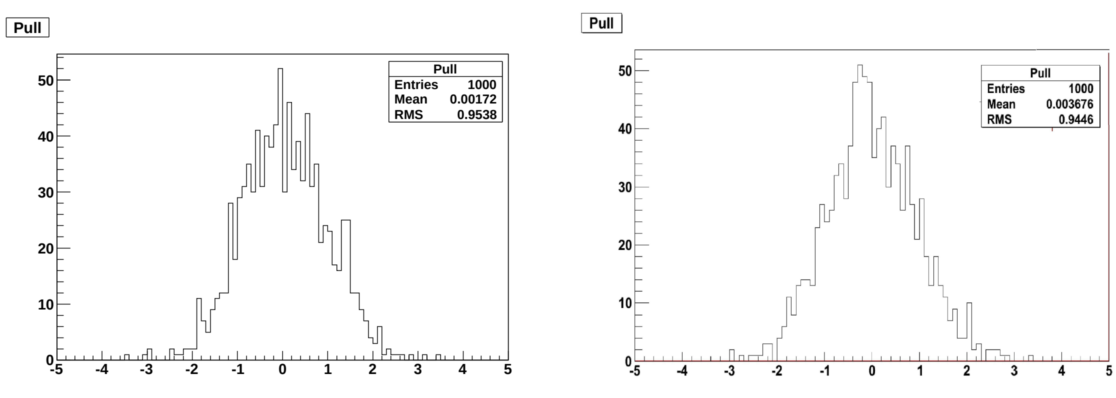
Figure 5.5: Linearity tests in $\mu + 4jets$and $e + 4jets$ channels.



(a) Ensemble tests for statistical error $\mu + 4\,jets$

(b) Ensemble tests for statistical error $e + 4\,jets$

Figure 5.6: Ensemble tests for statistical error in $\mu + 4jets$ and $e + 4jets$ channels.



(a) Pull distribution $\mu + 4\,jets$

(b) Pull distribution $e + 4\,jets$

Figure 5.7: Pull distributions in $\mu + 4jets$ and $e + 4jets$channels.

5.3.3 Evaluation of Systematic Errors

The systematic uncertainties in the inclusive $t\bar{t}$ cross section measurement arise from uncertainties associated with JES, ISR/FSR, W+jets modeling, signal modeling, parton shower modeling, uncertainty in QCD normalization and shape, jet reconstruction efficiency, lepton momentum scale and resolution, difference in pileup in MC and data, PDF, out-of-cell uncertainties and soft jet uncertainties in MET calculations.

Factors contributing to the systematic uncertainty on the $t\bar{t}$ cross-section measurement can be separated into two categories: uncertainties on the event preselection efficiency $\Delta\epsilon$ and uncertainties on the LH template shapes Δf. The resulting uncertainty is calculated as

$$\frac{\Delta\sigma_{t\bar{t}}}{\sigma_{t\bar{t}}} = \frac{\Delta f/f - \Delta\epsilon/\epsilon}{1 + \Delta\epsilon/\epsilon} \tag{5.1}$$

Uncertainties due to the jet energy scale, ISR/FSR, parton shower modeling and signal modeling affect both preselection efficiency and LH template shapes. Uncertainties due to W+jets modeling affect the LH shape only. Uncertainties on both QCD background normalization and shape are resulting in the change of the number of fitted $t\bar{t}$ events. Difference in the pileup in data and MC leads to a change of the LH templates and of preselection efficiency.

Systematic uncertainty due to JES uncertainty was calculated by varying jet momenta in the Monte Carlo W+jets and $t\bar{t}$ samples by $1\sigma_{JES}$ up and $1\sigma_{JES}$ down. This variation leads to lowering or increasing jet momenta with respect to their nominal values. After variation of jet momenta some jets will fall below or become above the selection threshold. This leads to the change of numbers of jets in the final state. New $t\bar{t}$ preselection efficiencies were obtained on JES up and JES down samples, and the difference between modified and nominal efficiencies divided by the nominal efficiency, $\Delta\epsilon/\epsilon$ is calculated. Then the shapes of the signal and background templates are re-derived, and 1000 ensemble tests are performed on pseudo data, picked from the nominal samples. The difference between new fraction of $t\bar{t}$ events, obtained with modified templates, and nominal fraction, divided by nominal fraction, $\Delta f/f$ is calculated. These two uncertainties are completely correlated and are combined into one total error according to the equation 5.1.

Systematic uncertainty related to the jet energy resolution uncertainty is evaluated by smearing the reconstructed jet p_T in Monte Carlo according to the resolution expected in data. The effect on the cross section obtained with smeared and nominal MC samples is evaluated.

Systematic uncertainty related to the jet reconstruction efficiency uncertainty is evaluated by random dropping jets. The rate of dropping was obtained by comparison of jet reconstruction efficiency in data and MC simulated events.

Systematic uncertainty related to MET calculation due out-of-cell uncertainty and soft-jet uncertainty were calculated. The changes in preselection efficiency and in fitted numbers of expected $t\bar{t}$ events are evaluated using ensemble tests with modified templates.

Systematic uncertainties related to the modeling of lepton momentum scale and resolution are calculated. The smearing of both values was done. Using the recalculated lepton momentum, the changes in preselection efficiency and in fitted numbers of expected $t\bar{t}$ events were evaluated using ensemble tests with modified templates.

Systematic uncertainties due to scale factors for lepton reconstruction efficiency, trigger efficiency were evaluated. In case of electron, the electron identification efficiency was included in this systematic uncertainty. After smearing of both values, changes in preselection efficiency and in fitted numbers of expected $t\bar{t}$ events using ensemble tests with modified templates were calculated.

The effect of parton shower modeling was estimated by comparison of the results obtained with default MC sample simulated with POWHEG generator interfaced to JIMMY for the parton showering with POWHEG interfaced to PYTHIA as an alternative shower generator. The difference in number of expected $t\bar{t}$ events due to modified preselection efficiency and LH shape, divided by the number of expected$t\bar{t}$ events, evaluated with the nominal sample, is interpreted as a total systematic error due to the parton shower modeling.

Two different generators, MC@NLO and POWHEG, are used to estimate the effect of the difference in $t\bar{t}$ kinematics simulation on the cross section measurement. Both modified preselection efficiency and expected number of $t\bar{t}$ events are calculated using modified templates for ensemble tests. The difference in the cross section measurements obtained with these two generators is assigned as systematic uncertainty due to the signal modeling.

To find the error due to ISR/FSR modeling, MC samples with modified ISR/FSR are used. Combined systematic errors due to change of the preselection efficiency and change of the LH shape were evaluated for each sample, and the largest positive and negative deviations are used as systematic uncertainties.

The effect of the $W+$ jets modeling are estimated on results of the LH fit. The spectra of jets depend on the MC generator, factorization scale, functional form of the factorization scale and other parameters. To evaluate these effects, the parameters used for simulation of $W+$ jets events, i.e. functional form of factorization scale and minimum ALPGEN parton p_T, were modified. New samples were used to create new background templates. These templates were used to fit the nominal pseudo data. Difference between fitted and nominal $t\bar{t}$ fractions divided by nominal fraction is interpreted as systematic error due to the W+jets modeling.

Expected systematic uncertainty on measured $t\bar{t}$ cross section due to pileup is evaluated by reweighting the MC samples according to the number of reconstructed primary vertices observed in data. The new LH templates are obtained and used to perform ensemble tests fitting nominal pseudo data data.

Systematic error on $t\bar{t}$ cross section due to the parton distribution function (PDF) uncertainty was done by using $t\bar{t}$ and W+jets MC samples with modified PDF parameters. PDFs describe the momentum distributions of partons (gluons and quarks) inside the colliding protons. Uncertainties due to change in preselection efficiency and due to change of likelihood shapes, were obtained and recalculated into the uncertainty on the $t\bar{t}$ cross section measurement.

Contribution of small backgrounds, like single top, diboson production and $Z + jets$ production, is taken into account based on their theoretical cross sections which have particular uncertainties. Systematic uncertainties due to normalization on theoretical cross sections were estimated by modifying the normalization of the corresponding background and repeating the fit. The difference between the nominal fitted number of $t\bar{t}$ events and the number of $t\bar{t}$ events using modified normalization is propagated into the error on measured cross section. The breakdown of the systematic uncertainties for electron channel and muon channel in 3th and 4th jet bins are given in Table 5.2 and 5.3 respectively.

Table 5.4 shows systematic errors due to small backgrounds normalization which is evaluated for the combined result e+jets and μ+jets in 3rd and 4th jet multiplicity bins.

The systematic error related to the limited Monte Carlo statistics available for the template construction is also evaluated. The Monte Carlo templates have been fluctuated according to the number of events in each bin using Poisson statistics and calculated the $t\bar{t}$ cross section on pseudo data using modified templates. This procedure was repeated 1000 times, and the RMS of fitted $t\bar{t}$ cross section distribution was taken to be the systematic uncertainty due to Monte Carlo statistics for a given pseudo data ensemble. This number was averaged over 1000 pseudo data ensembles. For the combined cross section the uncertainty was found to be ± 1.8 %.

Channel	$e + 3$jets			$e + 4$jets		
Source	$\triangle_{preselection}$	$\triangle_{shape}$	$\triangle_{total}$	$\triangle_{preselection}$	$\triangle_{shape}$	$\triangle_{total}$
JES up/dn	-1.1 $+1.3$	$+3.1$ -5.1	$+4.2$ -6.3	$+8.4$ -8.9	$+0.6$ -2.1	-7.3 $+7.4$
parton shower modeling	$+5.6$	-2.4	±1.5	-4.9	$+0.1$	±2.6
Monte Carlo generator *	-9.6	$+7.3$	±9.3	$+7.0$	$+2.1$	±2.3
ISR/FSR			$+2.0 - 11.8$			$+8.7 - 5.7$
W-shape modeling			±4.5			±1.1
pileup modeling	-3.9	-4.1	-0.2	-3.9	-0.6	$+3.5$
jet reconstruction efficiency *	$+0.4$	-0.9	±1.3	$+0.0$	-0.5	±0.5
MET corrections up/dn	< 0.1	-4.8 $+9.6$	-4.8 $+9.6$	< 0.1	-0.6 $+2.5$	-0.6 $+2.5$
jet energy resolution *	-0.02	-10.9	±5.4	$+0.3$	$+3.2$	±1.4
QCD background normalization			$+0.4$			$+0.8$
QCD background shape			$+2.0$			$+0.2$
PDF	-3.3	-4.6	-1.4	$+0.5$	-3.8	-4.3
lepton						
momentum scale and resolution (up/dn)	< 0.1	$-4.1 - 4.4$	$-4.1 - 4.4$	< 0.1	$-0.4 - 0.5$	$-0.3 - 0.4$
trigger and reconstruction efficiency	±3.9	±0.3	±4.3	±3.9	±0.2	±4.1

Table 5.2: List of systematic uncertainties for electron channel. *(symmetrized)

Channel	$\mu + 3$jets			$\mu + 4$jets		
Source	$\triangle_{preselection}$	$\triangle_{shape}$	$\triangle_{total}$	$\triangle_{preselection}$	$\triangle_{shape}$	$\triangle_{total}$
JES up/dn	-4.1 $+3.9$	$+9.7$ -9.3	$+14.4$ -12.7	$+6.7$ -6.6	-1.2 -1.9	-7.4 $+5.1$
parton shower modeling	$+3.6$	-2.7	±3.0	-3.7	$+0.9$	±2.4
Monte Carlo generator *	-5.4	$+0.2$	±2.9	$+5.2$	-0.9	±2.0
ISR/FSR			$+5.9-6.9$			$+10.6-7.7$
W-shape modeling			±8.5			±1.1
pileup modeling	-0.1	$+0.4$	$+0.5$	-0.2	-0.9	-0.7
jet reconstruction efficiency *	$+0.5$	$+1.2$	±0.7	$+0.6$	$+0.3$	±0.2
MET corrections up/dn	< 0.1	$+8.2$ -1.2	$+8.2$ -1.2	< 0.1	-2.5 -1.3	-2.5 -1.3
jet energy resolution *	-0.05	$+0.7$	±0.4	-0.2	$+1.1$	±0.7
QCD background normalization			$+7.8$			$+5.8$
QCD background shape			$+15.6$			$+5.3$
PDF	-3.1	-4.3	-1.3	$+0.5$	-3.8	-4.3
lepton momentum scale and resolution (up/dn) trigger and reconstruction efficiency	 $+1.0+1.0$ ±0.8	 $+1.3+1.4$ ±0.2	 $+2.3+2.5$ ±0.6	 $+1.0+1.0$ ±0.5	 $-2.4-2.3$ ±1.0	 $-1.3-1.3$ ±0.8

Table 5.3: List of systematic uncertainties for Muon channel. *(symmetrized)

Source	$\triangle_{theory}$variation	$\triangle$up, down
Single top	$\pm 10\%$	± 0.50
Dibosons	$\pm 5\%$	± 0.05
Z+jets	+0.8	± 0.08

Table 5.4: Systematic uncertainties due to variation of the theoretical cross sections of small backgrounds.

Channel	$t\bar{t}$	$W + jets$	$Z + jets$	Single top	dibosons	QCD	Sum	Data
$e + 3$jets	116	521	60	21	9	64	790	752
$e + 4$jets	183	164	29	10	2.3	20	408.3	381
$\mu + 3$jets	165	985	54	31	15	80	1300	1292
$\mu + 4$jets	264	281	17	14	3.5	29	608.5	627

Table 5.5: Number of preselected events.

5.3.4 Results

Table 5.5 summarizes the number of expected background and top quark pair events, number of QCD events and number of observed events in data.

In the final step, a maximum likelihood fit has been performed to the data after the preselection cuts in the e+jets and μ+jets channels. The fit is performed in the same manner as on pseudo data samples. The results of the fit are shown in Fig. 5.8.

The numbers of fitted $t\bar{t}$ events and corresponding cross sections are shown in Table 5.6. Quoted statistical error come from the fit result.

5.3.5 Combination of Measurements in all four channels

The negative log likelihood method is used to fit the data as described at chapter 4. The number of $t\bar{t}$ events in the sample is obtained by minimizing the quantity of

Channel	$N_{t\bar{t}}$	$\sigma_{t\bar{t}}$ (pb)
$e + 3$jets	146 ± 50	207 ± 71 (stat.)
$e + 4$jets	205 ± 32	184 ± 28 (stat.)
$\mu + 3$jets	170 ± 66	170 ± 65 (stat.)
$\mu + 4$jets	257 ± 40	160 ± 25 (stat.)

Table 5.6: Results of the likelihood fit to data

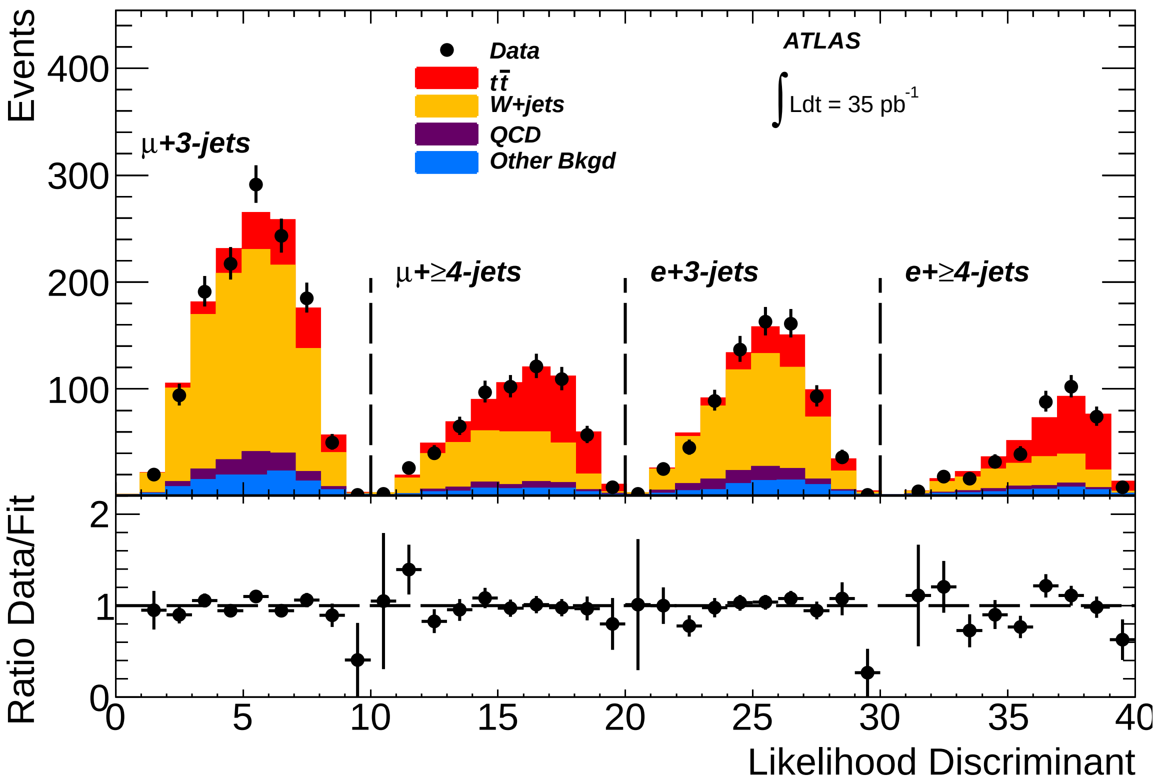
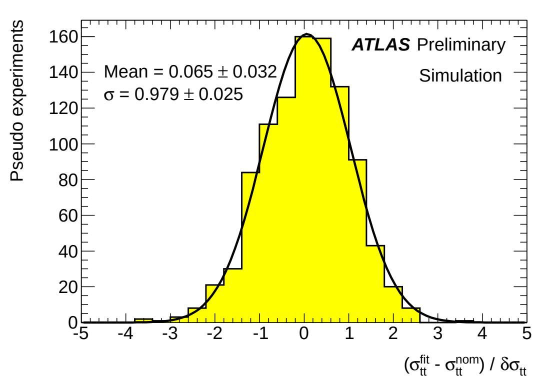
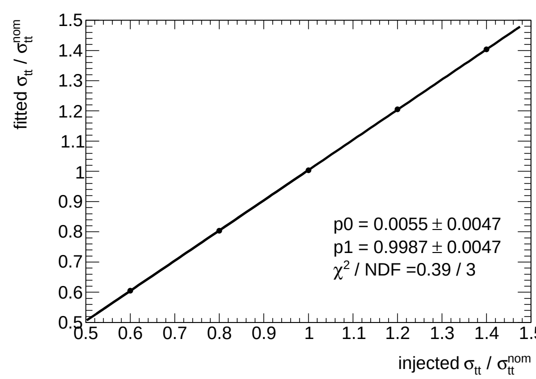


Figure 5.8: Fit to data.



(a) Pull of the combination fit



(b) Linearity of combination fit

Figure 5.9: Test of the correctness of the combination fit

NLL:

$$NLL = 2\sum_{i,j} \nu_{ij} - n_{ij} + n_{ij}\ln n_{ij} - n_{ij}\ln \nu_{ij} \tag{5.2}$$

where the sum is taken over all channels j ($\mu + 4$ jets, $\mu + 3$ jets, $e + 4$ jets, $e + 3$ jets) and all template histogram bins i; n_{ij} is the observed number of events, and ν_{ij} is the expected number of events. The ν_{ij} is defined as

$$\nu_{ij} = p_0 s_{ij} + p_j b_{ij} + m_{ij} \tag{5.3}$$

where s_{ij}, b_{ij}, and m_{ij} are the predicted numbers of events for the $t\bar{t}$ signal, W+jets, and small backgrounds (Z+jets, single top, dibosons, and QCD), respectively); p_0, p_j are the parameters of the fit. The resulting value of the fit parameterp_0 characterizes the ratio of the measured to the theoretical $t\bar{t}$ production cross section.

In case of Monte Carlo, the real data sample is replaced with N = 1000 pseudo data ensembles, and the fit is repeated for each ensemble. The average number of p_0 over all ensembles is taken to be the ratio of measured to injected $t\bar{t}$ cross section (and expected to be 1 for the nominal pseudo data), and the RMS of the p_0 distribution gives the statistical uncertainty of the method. It has been verify that the pull distribution $(p_0^k - 1)/\sigma_{p0}^k$ where σ_{p0}^k is uncertainty of an individual fit, has the RMS of 1 (see Fig. 5.9, left). It has been checked that it gets the same number of $t\bar{t}$ events after fit as it is injected - linearity of the fit is shown in Fig. 5.9, right.

The systematic uncertainties are evaluated by replacing the nominal predicted values of s_{ij}, b_{ij}, m_{ij} with their modified versions and repeating the fit. The relative systematic uncertainty is defined as the difference between the p_0 values in two cases. A combined likelihood fit has been performed in four channels, $e + 3$ jets, $e + 4$ jets, $\mu + 3$ jets and $\mu + 4$ jets, to obtain the combined cross section:

Source of Error	$\Delta\sigma_{t\bar{t}}/\sigma_{t\bar{t}}[\%]$
JES	$-5.2 + 3.8$
Jet reco. efficiency	< 0.1
Jet energy resolution	± 1.1
ISR/FSR	$-2.1 + 6.3$
MC generator	± 3.3
Parton shower	± 2.1
W-shape	± 4.0
PDF	± 1.8
QCD normalization μ	± 1.0
QCD normalization e	± 0.5
QCD shape μ	± 3.5
QCD shape e	± 0.2
Single top normalization	± 0.5
Diboson normalization	< 0.1
Z+jets normalization	± 0.4
Pile-up reweighting	± 1.2
Lepton momentum/energy scale and resolution	< 0.1
Lepton reco. efficiency and trigger SF	$-1.6 + 1.7$
Total Syst.	$+10.5 - 9.4$
Expected Stat.	± 9.7
Expected Stat. (+) Syst	$+15.0 - 13.9$
Measured Stat.	± 10.1
Measured Stat. (+) Syst	$+15.3 - 14.2$

Table 5.7: List of separate errors contributing to the combined cross-section over all 4 channels.

$$\sigma_{t\bar{t}} = 173 \pm 17(stat.)^{+20}_{-17}(syst.) \pm 6(lumi.)pb \qquad (5.4)$$

For luminosity error, 3.4% error is used [42, 43]. The breakdown of the combined error on the measurement is given in Table 5.7.

5.4 Conclusions

Measurement of the $t\bar{t}$ production cross-section in the single-lepton channel on data corresponding to 35.3pb^{-1} of integrated luminosity has been done.For the first time in the ATLAS experiment the multivariate technique was applied for this type of measurements. This allowed us to make the most precise measurement by that time at LHC. The measured value is in a good agreement with the Standard Model prediction.

Chapter 6

Measurement of the Cross Section in $t\bar{t}$ + Associated Jets

6.1 Introduction

The large statistics of top-anti-top quark pair events produced at the LHC provides an opportunity to study production of $t\bar{t}$ in association with at least one additional jet (further referred to as $tt\,j$). This process is an ideal precision test of the standard model. Due to the gluon-gluon fusion being the dominant production mechanism of $t\bar{t}$ pairs at the LHC, the majority of $t\bar{t}$ events is produced with additional jet(s). Precise description of hard interactions at LHC energies is an important theoretical task. The $tt\,j$ process represents a major background to $t\bar{t}H$ production which is crucial for Higgs Yukawa coupling studies, so one has to measure the $tt\,j$ production as accurately as possible in order to reduce the uncertainty in this measurement and verify available state of the art NLO calculations. The $tt\,j$ production is also a dominant background to new physics searches, in particular the heavy charged Higgs boson production. Lastly, a detailed analysis of the $tt\,j$ production characteristics helps to tune the Monte Carlo models which are used to simulate the physics processes at the LHC.

The measurement described in the chapter was carried out in the single-lepton final state using a multivariate technique. The CDF collaboration performed a similar measurement [44] at a centre-of-mass energy of 1.96 TeV.

The dataset used for this measurement corresponds to an integrated luminosity of $4.7\,fb^{-1}$. Events were selected by requiring the presence of a high-p_T electron or muon, large missing transverse energy (MET), and at least four jets. A likelihood discriminant was used to separate the $t\bar{t}$ events from the background (dominated by W+jets). The $tt\,j$ contribution was determined using reconstruction efficiencies in different jet multiplicity bins obtained with Monte Carlo. As a cross check, the $tt\,j$ production cross-section was also evaluated using the kinematic properties of $tt\,j$ events.

This chapter is organized as follows. Section 6.3 discusses the classification algorithm used to select additional particle jets in $t\bar{t}$ events. Section 6.4 lists the data samples used and details of the event selection, including the trigger selection. In Section 6.2 the Monte Carlo samples used are described. A summary of background sources and background estimation methods, including fake lepton estimates, is presented in Section 6.5. In Section 6.6 the technique of the cross-section measurement is discussed. Sources of systematic uncertainties are analyzed in Section 6.7. The comparison of different Monte Carlo generators is described in Section 6.7.2. Since initial and final state radiations can change the additional particle jets in the events for our analysis, this study is described in Section 6.7.1. Finally, the results and conclusions are presented in Section 6.10.

6.2 Monte Carlo Samples

The Monte Carlo samples intended for the analysis of the data collected until the end of 2011 are provided by the MC11 production campaign of the ATLAS Production

Group [45].

For the pile-up simulation the minimum bias events generated with PYTHIA6 [46] generator using the ATLAS AMBT2B(CTEQ6L1) tune [47] are used by default.

The template fitting analysis described in this chapter rely on Monte Carlo (MC) samples for the description of most of the backgrounds and the $t\bar{t}$ signal. The central sample, for this analysis, of the $t\bar{t}$ is generated using MC@NLO [48] generator v4.01 with the CT10 [49] parton distribution function. The parton shower and the underlying event were added using HERWIG v6.520 [50] and JIMMY 4.31 [39] generators using the CT10 HERWIG and JIMMY ATLAS AUET2 tune [51]. The $t\bar{t}$ cross section is normalized to the approximate next-to-next-to leading order prediction value of 166.78 pb, obtained using the HATHOR tool [52]. For the $t\bar{t}$ the samples simulated using POWHEG [53] interfaced with PYTHIA as well as HERWIG and JIMMY enable the comparisons of two different parton shower and hadronization models and (with the central samples) also a comparison of the POWHEG and MC@NLO engines. The $t\bar{t}$ multi-jet samples have been simulated using ALPGEN v2.13 with HERWIG and JIMMY [54] and SHERPA v1.3.1 [55] generators. For the systematics studies samples with more and less parton shower activity generated with AcerMC and PYTHIA generators is used.

Of the dominant backgrounds, vector boson production with associated jets ($W+jets$ and $Z+jets$) was simulated using ALPGEN interfaced to HERWIG and JIMMY [39] generators. The heavy flavour $Wb\bar{b}$ and $Wc\bar{c}$ processes were also generated from zero to two additional massless partons, whereas Wc was generated from zero to four additional partons. This analysis constructed QCD templates from data using Jet-electron or matrix methods.

The single top quark production t, s, and the Wt channels were generated with MC@NLO generator assuming a top quark mass of $m_{\mathrm{t}} = 172.5$ GeV. Finally the diboson events (WW, WZ and ZZ) were produced using ALPGEN and HERWIG.

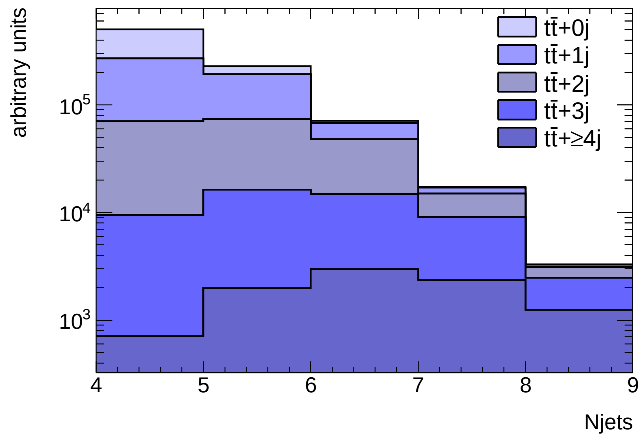


Figure 6.1: The multiplicity of reconstructed jets in $t\bar{t}$ events as predicted by MC@NLO simulation. Shown with various colors are the contributions of $t\bar{t}$ events with various number of additional jets calculated according to the procedure described in the text.

The pile-up re-weighting for these samples was also done to match the amount of pileup observed in the data using the pileup re-weighting tool. Scale factors like Lepton trigger, ID, reconstruction, and so on are used in compliance with [56] and [57].

6.3 Definition of Additional Jets

The additional jets in $t\bar{t}$ events are defined in a way that allows us to disentangle experimental effects and allow direct comparison of our results with theoretical models. The measurement begins with the jets observed in the detector and relate them to the jets of stable particles at the Monte Carlo generator level as illustrated by Fig. 6.1. In order to do that, the jets in Monte Carlo are classified according to the procedure described below (called definition 1).

For each generated sample of MC, stable particles were clustered into jets using the anti-k_t algorithm with a radius parameter of 0.4. These particle jets were then classified using $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$ matching between the jets and partons from top quark decays. Particle jets which were not matched to a parton from the top quark were flagged as additional jets. The two $t\bar{t}$ templates used in the present analysis were generated by requiring zero and one or more additional particle jets respectively.

The classification algorithm selects top quarks with status code 3 and 155 for PYTHIA and HERWIG6 respectively. The status code 155 is the HERWIG6 code for a heavy quark before decay. Immediate daughter partons and W boson indices are then stored. The algorithm then loops over the collection of selected particle jets. For each particle jet, each parton with $p_T > 5$ GeV is used to calculate a ΔR match. If the value of ΔR is less than 0.4, the algorithm checks to see if the parton is in the list of top daughter partons. If it is not found in this list, its parents are used to check against the list of top daughter partons. The recursive query finishes when there are no more parents. Jets which are not matched with partons that originate from the top quark are flagged as additional jets. All partons except top quarks are required to have $p_T > 5$ GeV .

The stability of the resulting number of additional jets as a function of the parton p_T cut was studied using POWHEG and ALPGEN generators. The term $f_1^{default}$ is defined as a fraction of events with at least 1 additional jet out of total number of $t\bar{t}$ events calculated with parton $p_T > 5$ cut. Then f_1^{cut} is calculated using several parton p_T cuts varying from 2 to 15 GeV. Fig. 6.2 shows the ratio $f_1^{cut}/f_1^{default}$ as a function of parton p_T cut. No dependence is observed for POWHEG and slight variation (within 2%) is observed for ALPGEN in the wide region of parton p_T cut values.

Partons from additional in-time and out-of-time interactions while typically present within analysis object data (AODs), were excluded from the data format used for this analysis. The classification of additional particle jets is therefore not affected by the presence of additional interactions within the MC record, but does include partons from the underlying event.

For definition 2 an event is classified as $tt\,j$ if it has at least 5 selected parton

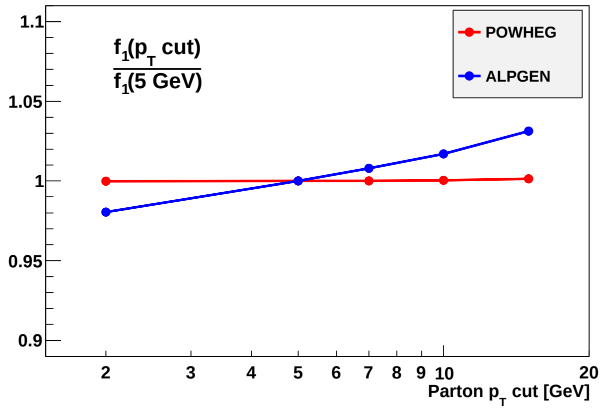


Figure 6.2: The ratio of fraction of events with at least one additional jet vs parton p_T cut to its value at 5 GeV.

jets. The $tt\,j$ production cross sections in fiducial volume defined in this way are denoted $\sigma^{fiducial}_{ttX\rightarrow e+\geq 5jets}$, $\sigma^{fiducial}_{ttX\rightarrow \mu+\geq 5jets}$. These cross sections can be further unfolded to the parton level using relevant Monte Carlo. This is not done in the present analysis.

6.4 Event Selection

6.4.1 Data sample

Data were selected from the 2011 LHC running period, corresponding to a centre-of-mass energy of 7 TeV. Basic data quality requirements were applied through the application of the `top_allchannel` good runs list, resulting in an integrated luminosity of 4.7 fb^{-1}. Events were selected using unprescaled single lepton (electron or muon) triggers. Due to the increasing instantaneous luminosity, the threshold and isolation requirements of these triggers were tightened during the 2011 running period.

6.4.2 Event Selection Criteria

Data were selected for the electron (muon) channel following,

- the single lepton trigger requirement according to the data period,
- a primary vertex reconstructed from at least five tracks,
- exactly one good selected electron (muon),
- no selected muons (electrons),
- a $\triangle R$ match between the selected lepton and the trigger used to select the event,
- at least four reconstructed jets passing the object selection,
- $E_{miss}^{T} > 30$ GeV (electron channel) or $E_{miss}^{T} > 20$ GeV (muon channel),
- a transverse W mass[1] $m_T(W) > 30$ GeV (electron channel) or $m_T(W) + E_{miss}^{T} > 60$ GeV (muon channel),

Events with noise bursts within the LAr calorimeter were rejected. The fraction of data affected by LAr front-end board failures was excluded within the data and MC samples.

6.5 Background Estimation

The dominant backgrounds for the $t\bar{t}$ + jets measurement are expected to be the vector boson production with associated jet and/or heavy quark production (W+jets γ/Z+jets, $W + c$+jets, $W + c\bar{c}$+jets, $W + b\bar{b}$+jets), the single top production, the diboson (WW) production and the QCD backgrounds. The Monte Carlo samples are used to estimate the background contributions from single top, diboson production and Z+jets. The data-driven estimates are used for QCD as described in section 6.5.1.

[1] $m_T(W)$ is defined as $\sqrt{2p_T^l p_T^{\nu}(1 - \cos(\phi^l - \phi^{\nu}))}$, where l and ν refer to the lepton and neutrino terms respectively. In the event selection the neutrino terms were replaced with the E_{miss}^{T} and its azimuthal angle.

For $W+$ jets the normalization is taken from final likelihood fit to the data and likelihood template shape is constructed from the Alpgen Monte Carlo sample using the selection cuts listed in section 6.4.

6.5.1 Fakes Lepton Estimates

The QCD fake contribution in muon channel was estimated using a data driven Matrix Method as described at section 3.3.1. In muon channel the signal and fake efficiencies were determined from data in a low transverse W mass region with an inverted triangular cut: $M_T(W) < 20$ GeV and $MET + M_T(W) < 60$ GeV. The loose selection of muons is identical to the one that is used for the analysis, except for the muon isolation. The contribution from $W+$ jets and $Z+$ jets events was subtracted to obtain as pure QCD estimate as possible. Both the signal and fake efficiencies were parametrized in muon η and leading jet p_T in order to account for dependencies on muon detector acceptance and hadronic activity from hard jets affecting the muon isolation. The signal efficiencies were determined with a Tag & Probe method by selecting prompt muons from the Z decay, while the fake efficiencies were determined in the described control region.

The QCD contribution to electron channel was calculated by jet-electron method . QCD in electron channel is mostly due to jets mimicking electrons. The jet-electron model is using a sample, selected with a Jet trigger. Jets with $E_T > 25$ GeV in the η range of real electrons, that posses a relatively high fraction of electromagnetic energy (80%-95%) compared to their total energy, are considered as fake electrons in this model. In order to reduce the fraction of real conversions that might leave a large signal in EM calorimeter, jets require to have at least 4 tracks associated with them. Events selected in this way are used to obtain the QCD shapes for all distributions.

An alternative method of QCD evaluation in the electron channel is Matrix Method. For this method, loose electrons were defined as those that pass the Medium++

electron identification criteria and satisfy conversion veto and isolation criteria. The first isolation cut was set on the transverse energy in the calorimeter around the electron candidate in a cone of $\Delta R = 0.2$ around the electron so that E_T^{max} value corresponds to 98% efficiency for isolated electrons. The second cut makes use of the transverse momenta of the electron candidate and of all reconstructed tracks around it in the cone of $\Delta R = 0.3$. The sum of p_T's of all tracks excluding electron should be less than p_T^{max}, where the chosen p_T^{max} value corresponds to 98% efficiency for isolated electrons. Since contamination from signal leptons from W and Z decays is still present in the low MET region, corrections of the number of observed loose and tight events in data based on MC were applied.

6.6 Overview of the Analysis Strategy

The present analysis was carried out in the four jet exclusive and five jet inclusive multiplicity bins in the e + jets and μ+jets channels. Using the particle-level jet classifier described in Section 6.3, the sample of MC $t\bar{t}$ events was split in two sets: (i) events without an additional jet (further denoted as $tt0$) and (ii) events with one or more additional jets (ttj). For a given classification, one can evaluate the reconstruction efficiencies ε_{ij} defined as probabilities to observe a $tt0$ ($i = 0$) or ttj ($i = 1$) event in the 4th exclusive ($j = 4$) or 5th inclusive ($j = 5$) jet multiplicity bin. These efficiencies were calculated using full detector simulation and take into account the data/MC correction factors and pile-up effects.

Two different methods were used to determine the ttj production cross-section. In 1-Dimensional Likelihood, the number $N_j^{t\bar{t}}$ of inclusive $t\bar{t}$ events in data first are evaluated using the method described below. The $tt0$ (ttj) production cross-sections

σ_{tt0} (σ_{ttj}) were then obtained from a system of equations shown at 6.1

$$\begin{aligned} N_4^{t\bar{t}} &= \mathcal{L}\sigma_{tt0}\varepsilon_{04} + \mathcal{L}\sigma_{ttj}\varepsilon_{14} \\ N_5^{t\bar{t}} &= \mathcal{L}\sigma_{tt0}\varepsilon_{05} + \mathcal{L}\sigma_{ttj}\varepsilon_{15} \end{aligned} \tag{6.1}$$

where $\mathcal{L}$ is total integrated luminosity. For a single channel (e or μ) this system consists of two equations and has two unknowns (σ_{tt0} and σ_{ttj}), so the solution is straightforward. For a combination of the two channels there are four equations, so the system is over constrained. In this case, the cross sections are determined from a log likelihood fit to the number of events observed in each channel. For convenience, the calculations are performed in terms of k_0 and k_1 which are the ratios of the observed to the predicted number of $tt0$ and ttj events, respectively:

$$\begin{aligned} N_4^{t\bar{t}} &= k_0 N_4^0 + k_1 N_4^1 \\ N_5^{t\bar{t}} &= k_0 N_5^0 + k_1 N_5^1, \end{aligned} \tag{6.2}$$

where $N_j^i = \mathcal{L}\sigma_{t\bar{t}} f_i \varepsilon_{ij}$ is the predicted number of $tt0$ ($i = 0$) or ttj ($i = 1$) events, $\sigma_{t\bar{t}}$ is the theoretical inclusive $t\bar{t}$ production cross-section, and f_0 (f_1) are the fractions of $tt0$ (ttj) events in the original MC sample ($f_0 + f_1 = 1$). The ttj production cross-section is then found as

$$\sigma_{ttj} = k_1 \sigma_{t\bar{t}} f_1. \tag{6.3}$$

(The result depends on the reconstruction efficiencies but not on $\sigma_{t\bar{t}}$ nor f_1, as shown in Appendix A).

In in second, 2-Dimensional Likelihood, the kinematic difference between the $tt0$ and ttj events is used in addition to efficiencies. The inclusive $t\bar{t}$ measurement step can be skipped and evaluate directly the numbers of events due to $tt0$, ttj, and non-$t\bar{t}$

backgrounds simultaneously using a 2D likelihood fit. The details of 2-Dimensional Likelihood method are outlined in section 6.9. The 1-Dimensional Likelihood method is used for the main result and 2-Dimensional Likelihood method as a cross check. In the rest of the current section, the method 1 will be discussed.

6.6.1 Variables used in the Likelihood

The first step of the analysis is the separation of $t\bar{t}$ events from the SM backgrounds. The latter include W/Z+jets production, diboson production, single top production, and QCD multijet events. The W+jets background is by far dominant, so only W+jets events were used to construct the likelihood. The variables used in the likelihood are lepton pseudorapidity η_l and transformed aplanarity $A' = \exp(-8A)$, where the (regular) aplanarity A is defined as $A = \frac{3}{2} \times \lambda_3$, λ_3 being the smallest eigenvalue of the normalized momentum tensor. The $A \to A'$ transformation is done for uniformity as discussed in [58] and in chapter 4.

Agreement between data and MC for the LHW variables is shown in Figs. 6.3-6.6 before and after the LHW fit. For plots made before the fit, the sum of all MC events and QCD is normalized to the number of observed events in data to compare the shapes of these distributions.

We construct the likelihood discriminant function LHW (Likelihood Discriminator against W+jets) using the projective likelihood method defined in the TMVA [35]. The analysis strategy for cross section extraction using likelihood fitting and ensemble testing and performance of method performed as described at chapter 4.

6.6.2 Performance of the Method

The LHW distributions in 4th and 5th jet multiplicity bins are shown in Fig. 6.7 for electron and muon channel, respectively.

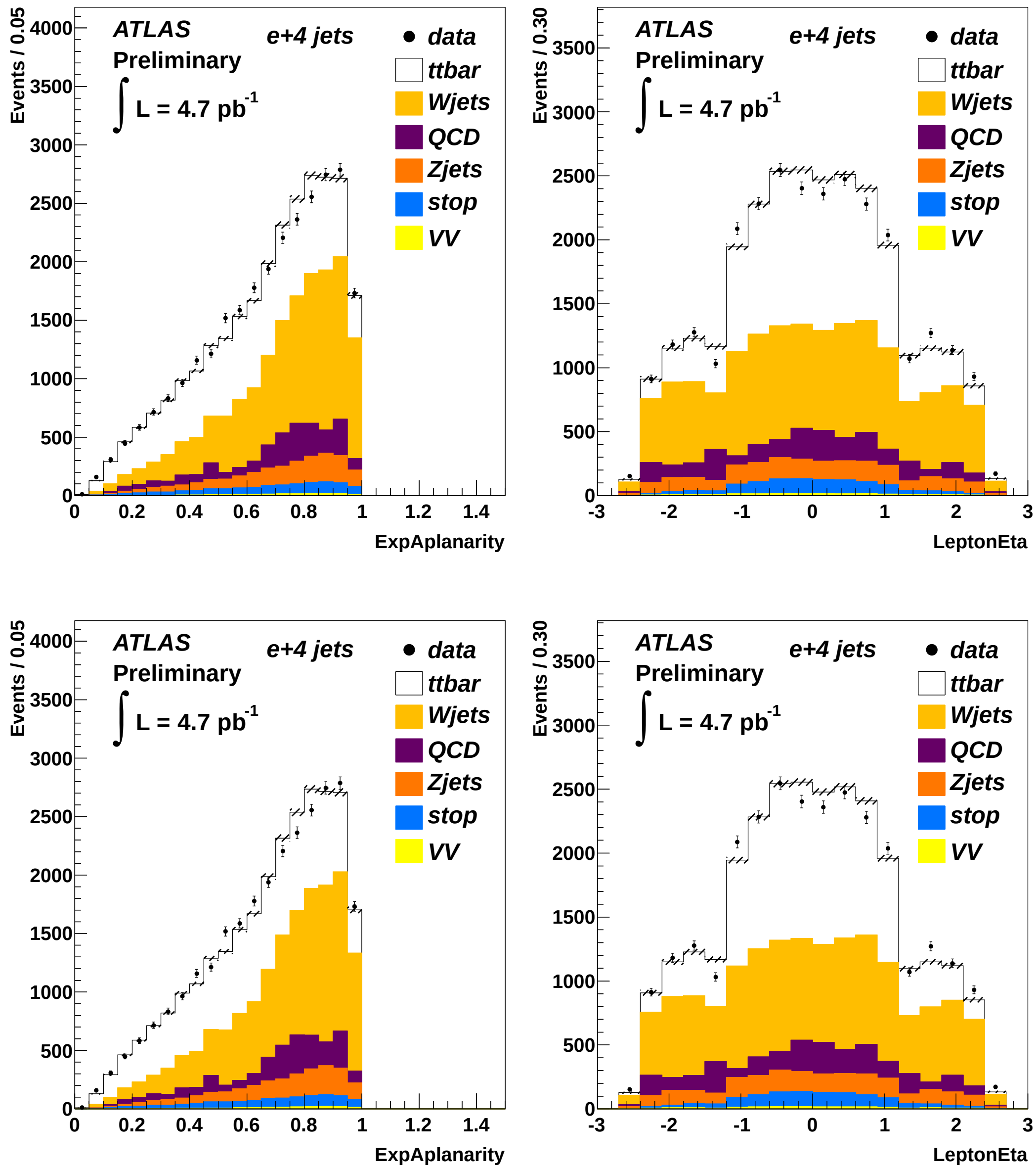


Figure 6.3: Control plots in the e+4 jets channel: transformed aplanarity (left) and electron pseudo rapidity (right), normalized to the total number of events in data (top) and after the LHW fit (bottom).

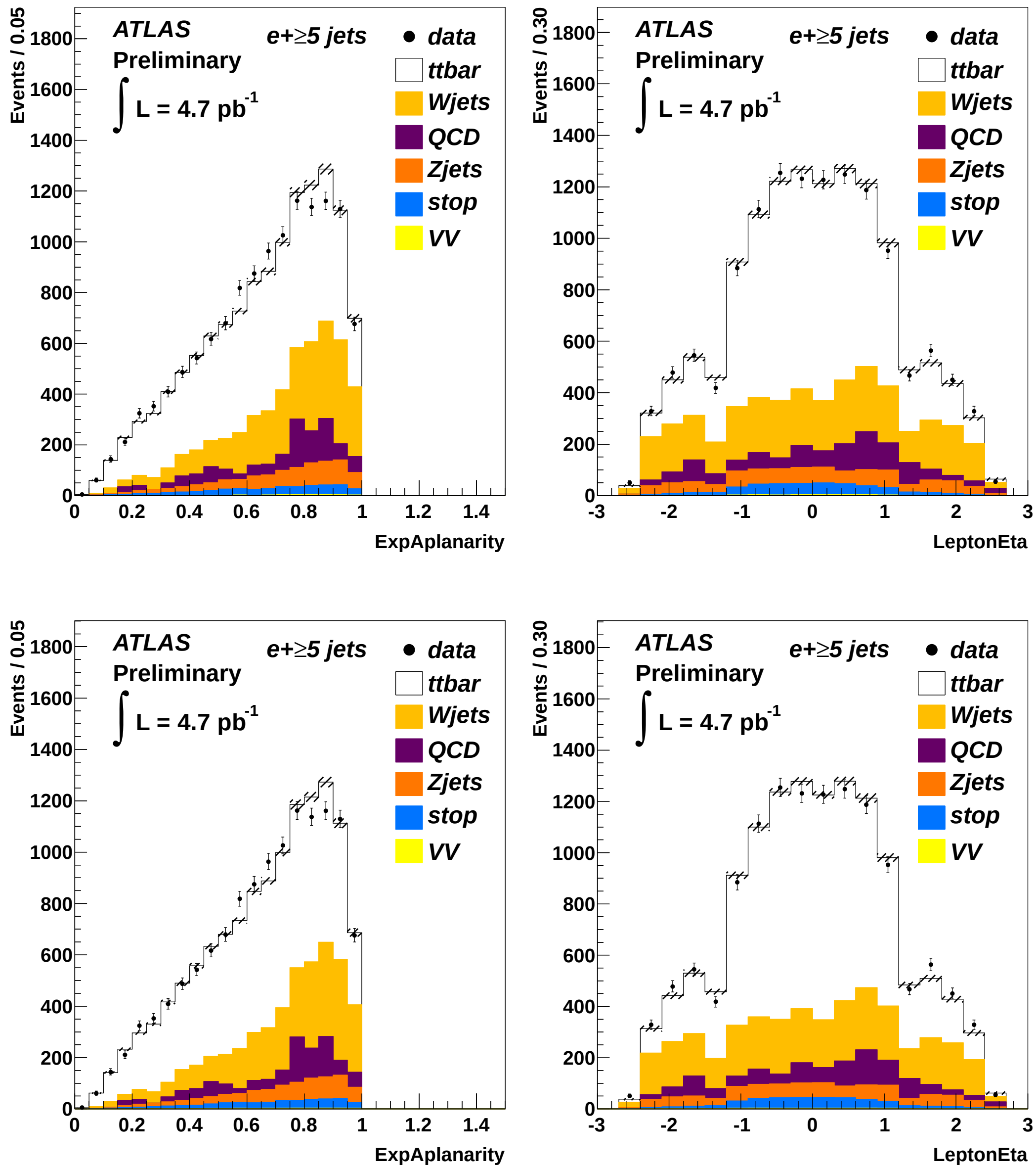


Figure 6.4: Control plots in the $e+\geq 5$ jets channel: transformed aplanarity (left) and electron pseudo rapidity (right), normalized to the total number of events in data (top) and after the LHW fit (bottom).

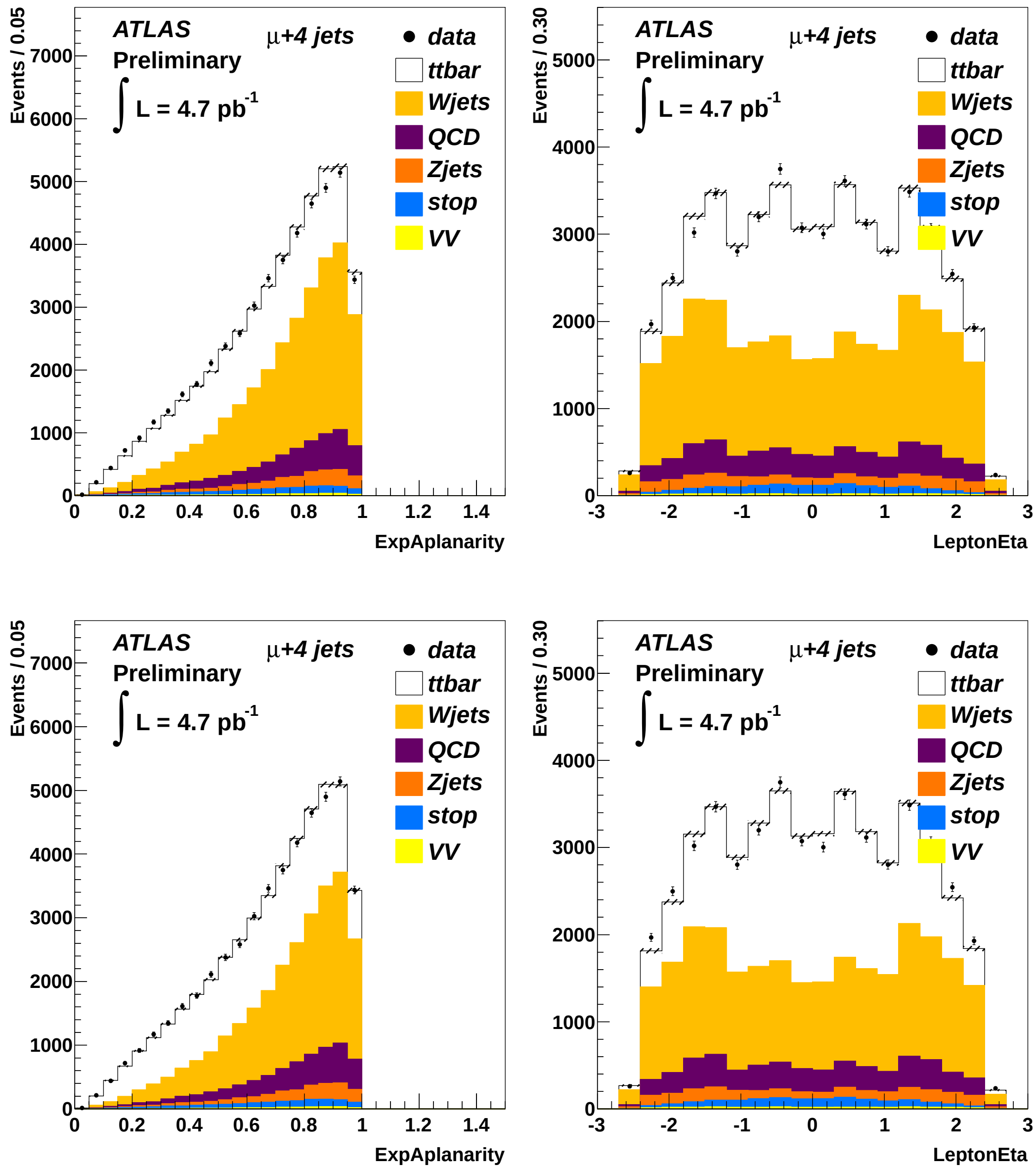


Figure 6.5: Control plots in the μ+4 jets channel: transformed aplanarity (left) and muon pseudo rapidity (right), normalized to the total number of events in data (top) and after the LHW fit (bottom).

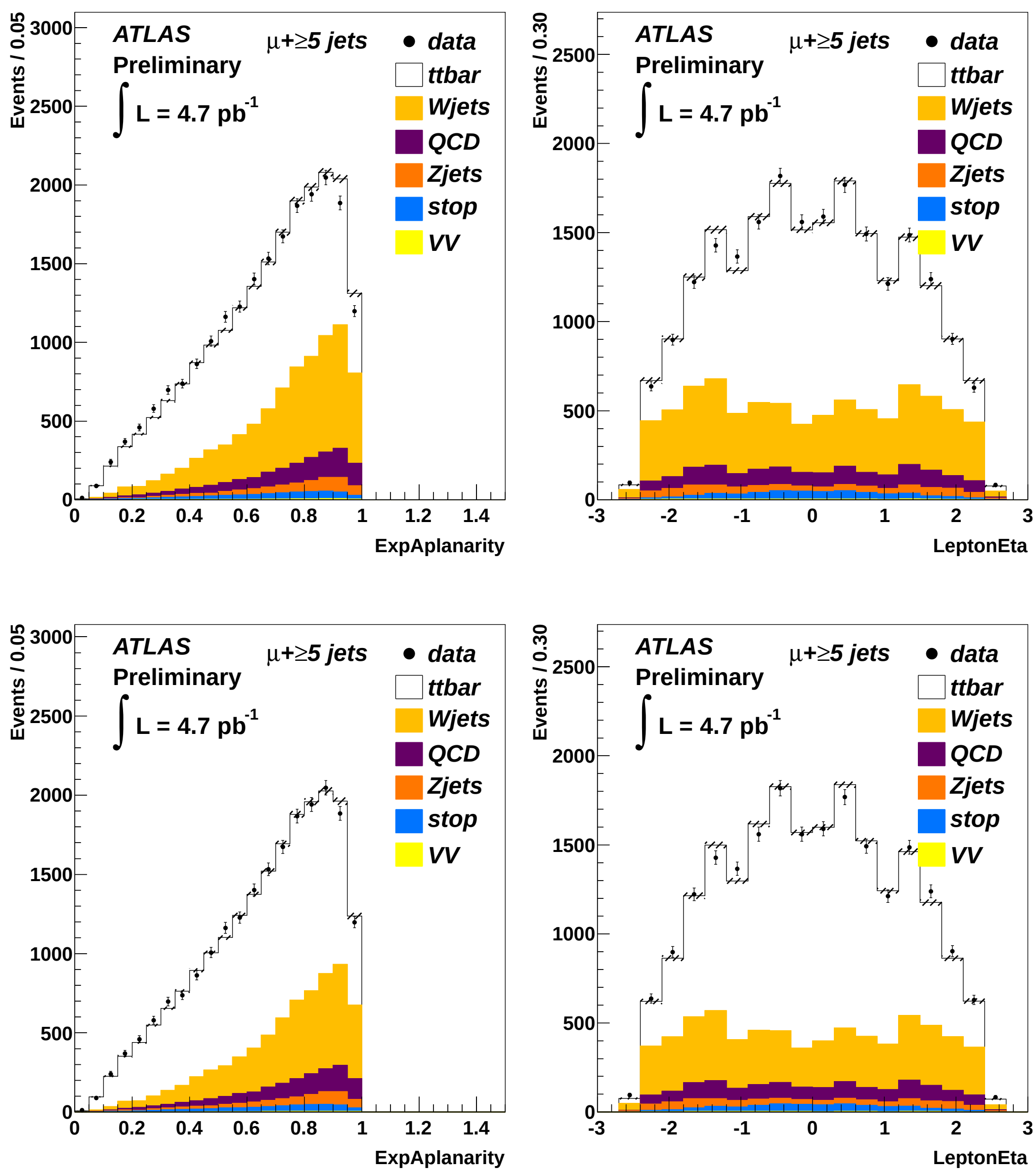


Figure 6.6: Control plots in the $\mu+\geq 5$ jets channel: transformed aplanarity (left) and muon pseudo rapidity (right), normalized to the total number of events in data (top) and after the LHW fit (bottom).

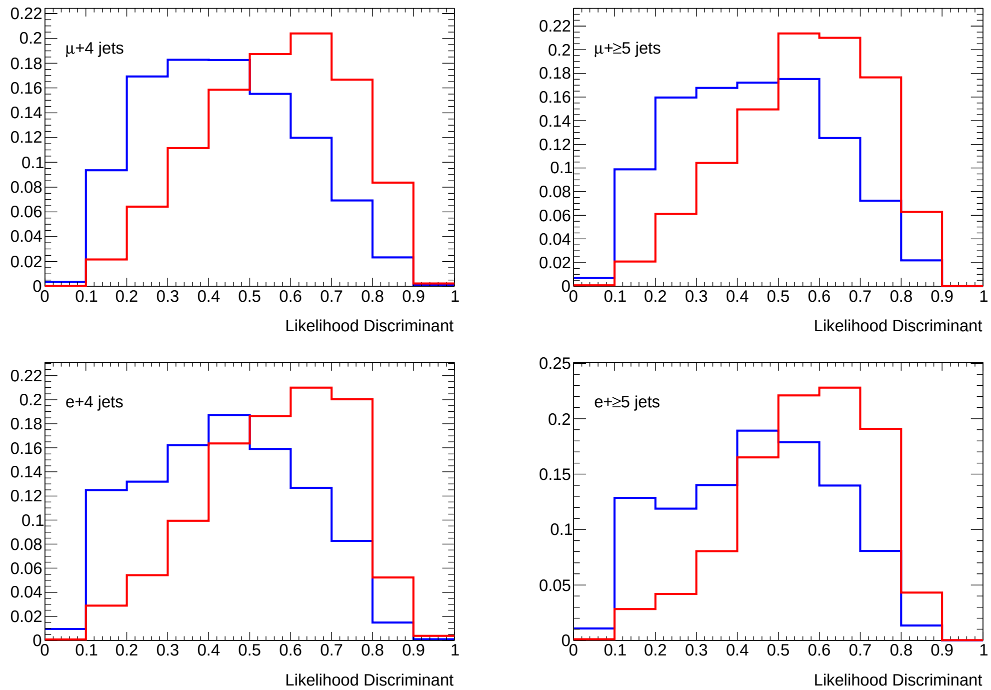


Figure 6.7: LHW templates for muon and electron channels, where the red line represents the $t\bar{t}$ and the blue line represents the background templates.

Performance of the LHW was studied by checking the fitting fraction of $t\bar{t}$ events as a function of input fraction. The results are shown in Fig. 6.8 for both channels. No shifts are observed in fitted fractions.

The expected statistical uncertainty is evaluated with ensemble tests on the MC samples. When preparing the ensemble tests, the total number of events is fixed to the number of events observed in data. This number is composed of the sum of $t\bar{t}$ events and expected numbers of events from each background and is fixed in each ensemble. The maximum likelihood fit is performed on the pseudo data resulting in $N_{t\bar{t}}$. For each test, 1000 pseudo data samples are generated.

The expected statistical error is found from the width of the $N_{t\bar{t}}$ distribution found resulting from the fit when applied to pseudo data and turned out to be of the order of 3%.

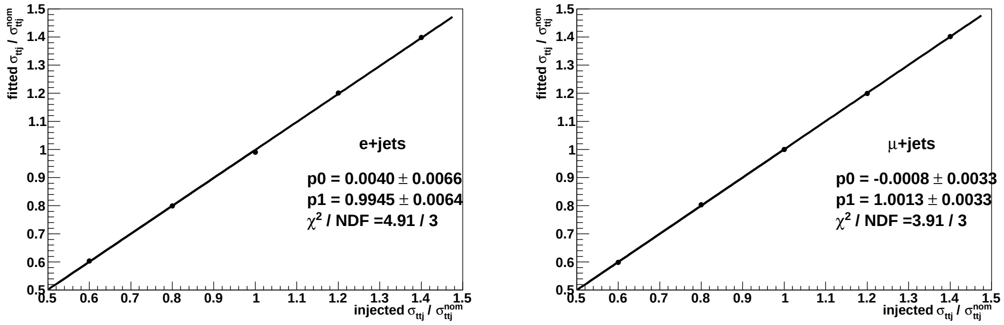


Figure 6.8: Linearity tests for the LHW in e+jets and $\mu+$ jets channels.

6.6.3 Likelihood Fit to Data

Event yields in data and expected yields from MC based on the theoretical cross-sections in electron and muon channels are shown in Table 6.1.

Channel	$t\bar{t}$	W+jets	Z+jets	Dibosons	Single top	QCD	Total expected	Data
e+4 jets	10920 $^{+1080}_{-1165}$	11527 ±5530	2043 ±612	188 ±9	1005 ±40	2467 ±1234	28151 ±5817	27599 ±166
e +≥5 jets	6856 $^{+678}_{-731}$	2900 ±1556	741 ±285	37 ±2	356 ±14	960 ±480	11851 ±1808	12778 ±113
μ+4 jets	17672 $^{+1080}_{-1165}$	21363 ±10249	1900 ±570	314 ±15	1169 ±47	4486 ±2243	46903 ±10571	47827 ±219
$\mu+$ ≥5 jets	11250 $^{+1113}_{-1200}$	5426 ±2912	586 ±225	56 ±3	414 ±17	1232 ±616	18963 ±3217	20985 ±144

Table 6.1: Expected and observed event yields in different jet multiplicity bins in electron and muon channels. Errors on the Monte Carlo predictions were calculated according to the variations of theoretical cross-sections for $t\bar{t}$, single top and diboson production. For W+jets and Z+jets, the errors expected from Berends scaling are shown. QCD background has 50% uncertainty.

Results of the likelihood fits to data in all channels are shown in Fig. 6.9. The resulting numbers of $t\bar{t}$ events from these fits are summarized in Table 6.6.3. While performing the likelihood fit in each sub-channel, the number of $W+$ jets events in the fit is allowed to float. The resulting numbers of $W+$ jets events are also listed in Table 6.6.3.

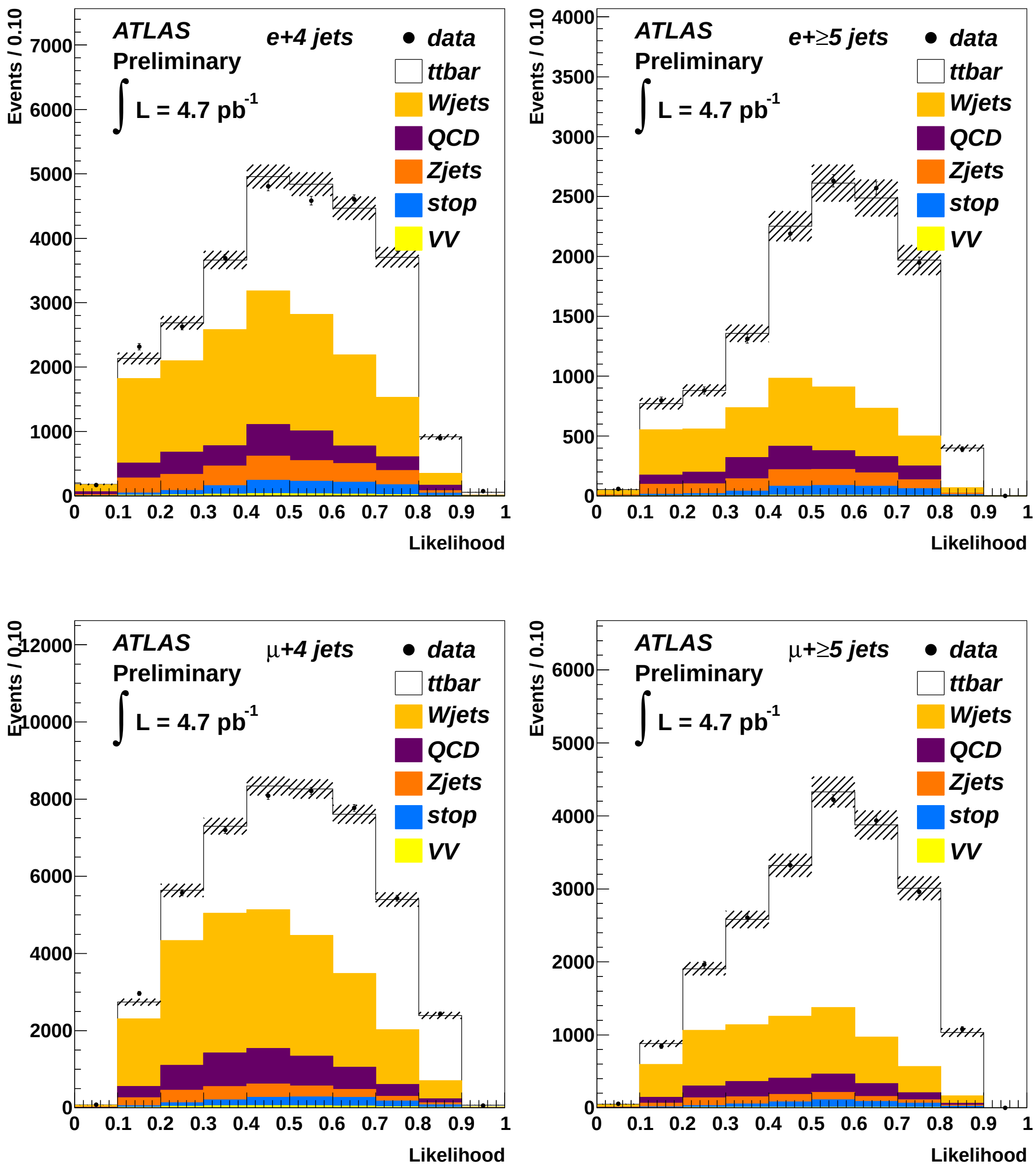


Figure 6.9: Likelihood fits in the e+jets and μ+jets channel.

Channel	N^{tt}	N^{tt}/N^{tt}_{pred}	N^{W+jets}
e+4 jets	10848	0.99 ± 0.02	11062
e+5 jets	7700	1.12 ± 0.03	2990
μ+4 jets	20257	1.14 ± 0.02	19714
μ+5 jets	13821	1.23 ± 0.02	4876

Table 6.2: Fitted numbers of W+jets events N^{W+jets}, $t\bar{t}$ events $N^{t\bar{t}}$ and the observed to predicted event ratios $N^{t\bar{t}}/N^{t\bar{t}}_{pred}$ in electron and muon channels in 4 and 5 jet multiplicity bins. The errors are statistical only.

6.6.4 Measurement of ttj Production Cross Section

The numbers of inclusive $t\bar{t}$ events obtained in each channel are combined according to the procedure described in the section 6.6. The efficiencies ε_{ij} are summarized in Table 6.3. The 5th inclusive jet multiplicity bin is dominated by the ttj events, as illustrated by Fig. 6.10.

The fitted k_1 values are shown in Table 6.4 for e+jets and μ+jets channels as well as for the combination. The errors in the table are statistical.

	ε_{04}	ε_{05}	ε_{14}	ε_{15}
e+jets	2.40	0.40	2.81	2.87
μ+jets	4.21	0.72	4.82	5.03

Table 6.3: $tt0$ and ttj reconstruction efficiencies (%).

The fitted k_1 values are shown in Table 6.4 for e+jets and μ+jets channels and for their combination. The errors in the table are statistical.

Channel	Fitted k_1
e+ jets	1.17 ± 0.04
μ+ jets	1.26 ± 0.03
Lepton + jets	1.22 ± 0.02

Table 6.4: Fitted k_1 values in electron and muon channels separately and in the combination. The errors are statistical only.

The ttj production cross-section is found to be

$$\sigma_{ttj} = 102 \pm 2(\text{stat.})\ \text{pb}. \tag{6.4}$$

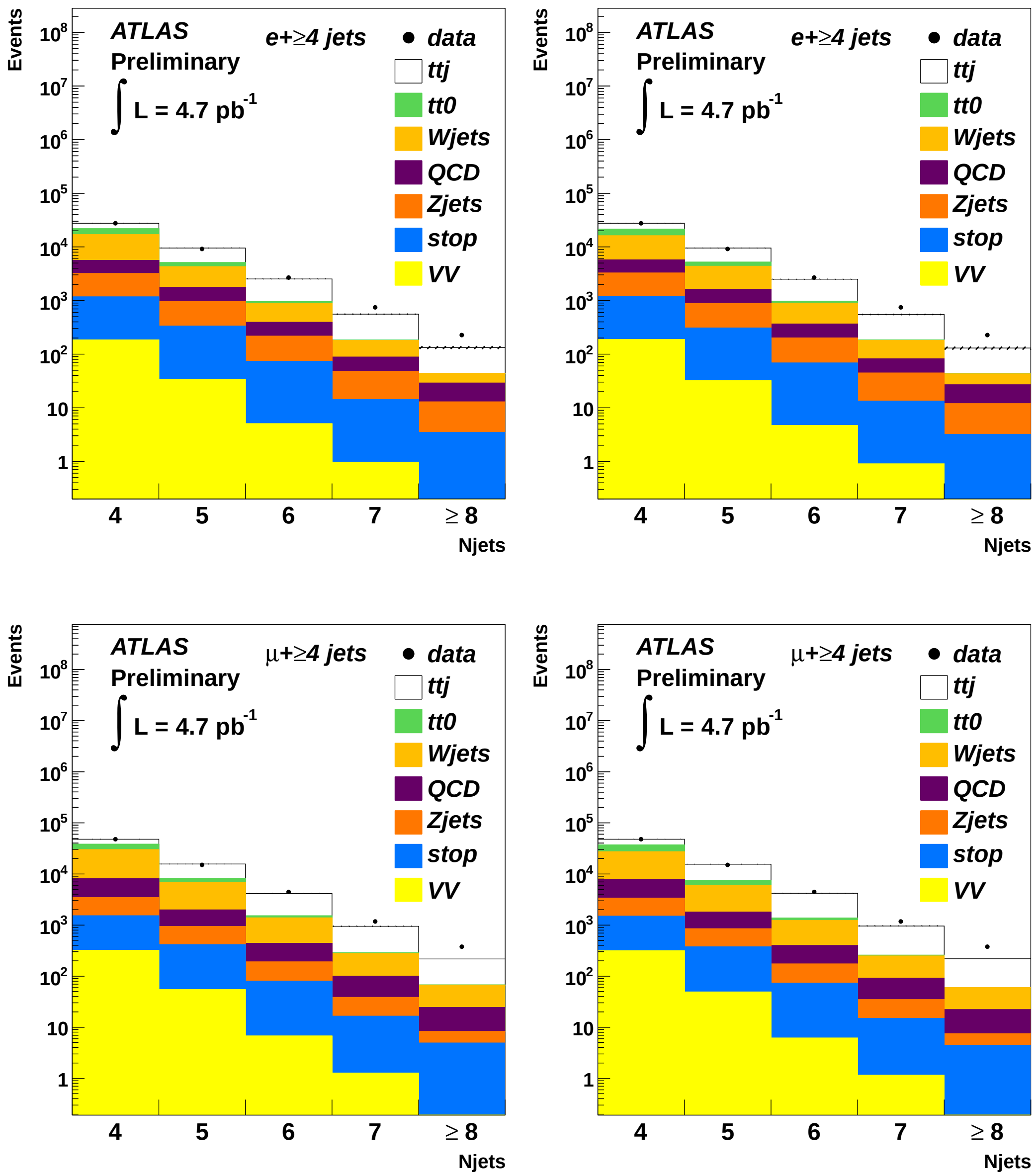


Figure 6.10: The distribution of events in jet multiplicity bins normalized to the total number of events in data (left) and from the fit to data (right) in the e+jets and μ+jets channels.

The numbers of fitted $t\bar{t}$ events in e+jets and μ+jets channels in 4-th and 5-th jet multiplicity bins can be used for calculation of the inclusive $t\bar{t}$ cross-section, which is found to be $\sigma_{t\bar{t}}^{incl} = 189 \pm 4$ pb. The ratio of ttj to inclusive $t\bar{t}$ production cross-section is

$$f_1^{\text{meas}} = 0.54 \pm 0.01(\text{stat.}). \tag{6.5}$$

6.7 Evaluation of Systematic Uncertainties on $\sigma_{t\bar{t}+jets}$

The estimated systematic uncertainties are estimated, Jet energy scale uncertainty, W+jets modeling, signal modeling, parton shower modeling, uncertainty in QCD normalization and shape, small backgrounds normalization, jet reconstruction efficiency, lepton momentum scale and resolution, difference in pile-up in MC and data, PDF, cell-out uncertainties, soft jet uncertainties in MET calculations and systematic uncertainty arising from the removal of events with jets in the LAr hole region.

Factors contributing to the systematic uncertainty of the $t\bar{t}+\geq 1$ particle-jet cross-section measurement can be separated into two categories: uncertainties on the event preselection efficiency and uncertainties on the likelihood template shapes. Systematic uncertainties except for small backgrounds and QCD normalization and QCD shape were evaluated using ensemble testing. Systematic uncertainties related to the small backgrounds were obtained from fit to data by variation of their cross-sections within theoretical uncertainties.

Uncertainties due to the jet energy scale, jet resolution, Monte Carlo model used for the signal simulation, and parton distribution functions (PDFs) affect both preselection efficiency of $t\bar{t}$ events and likelihood template shapes. The detailed description of these sources of systematic uncertainty is given below. The expected number of $t\bar{t}$ events referred to after all preselection cuts used in the analysis as to a nominal number. Similarly, the nominal signal and background templates are referred to templates obtained after application of all preselection cuts. Modified numbers of $t\bar{t}$ events and modified templates are defined as those obtained after the variation of sources of systematic uncertainties. Following there are more detail on the procedures used to calculate various types of systematic uncertainties.

Jet Energy Scale (JES): The uncertainty on the $t\bar{t}+\geq 1$ particle-jet cross-section due to jet energy scale uncertainty was measured by variation of JES by $\pm 1\sigma_{JES}$.[ref to JES plot for Jet Identification section] After applying the JES corrections to all reconstructed jets, the E^{T}_{miss} is recalculated. Change in the number of selected $t\bar{t}$ events due to variation of JES propagates as a systematics uncertainty to the cross-section of $t\bar{t}+\geq 1$ particle jets.

The likelihood templates are derived new to estimate the contribution of JES variation to the likelihood shape. Fig. 6.11 shows three sets of likelihood templates for signal and background used for evaluation of JES related uncertainty: nominal ones (solid lines), templates obtained by changing JES by 1σ up (dotted lines), and those obtained by changing JES by 1σ down (dashed lines). Modified templates were used in ensemble tests to derive modified numbers of the $t\bar{t}$ events in each jet multiplicity bin. The difference between the number of $t\bar{t}$ events obtained with nominal and modified likelihood templates is interpreted as an uncertainty due to the likelihood shape and is further propagated to calculate the JES uncertainty on the $t\bar{t}+\geq 1$ particle jets production cross-section. The total uncertainty due to JES variation is a combination of the two effects (change in the preselection efficiency and in the likelihood template shape). Results of ensemble tests (fitted to predicted number of events) with nominal and JES shifted LHW templates are shown in Fig. 6.12.

Jet Energy Resolution (JER): The uncertainty due to different jet energy resolution in data and MC was evaluated by smearing the MC jet transverse momentum depending on jet p_T and η. After smearing, all selection cuts were applied, and the number of expected $t\bar{t}$ events was calculated. The difference between the nominal

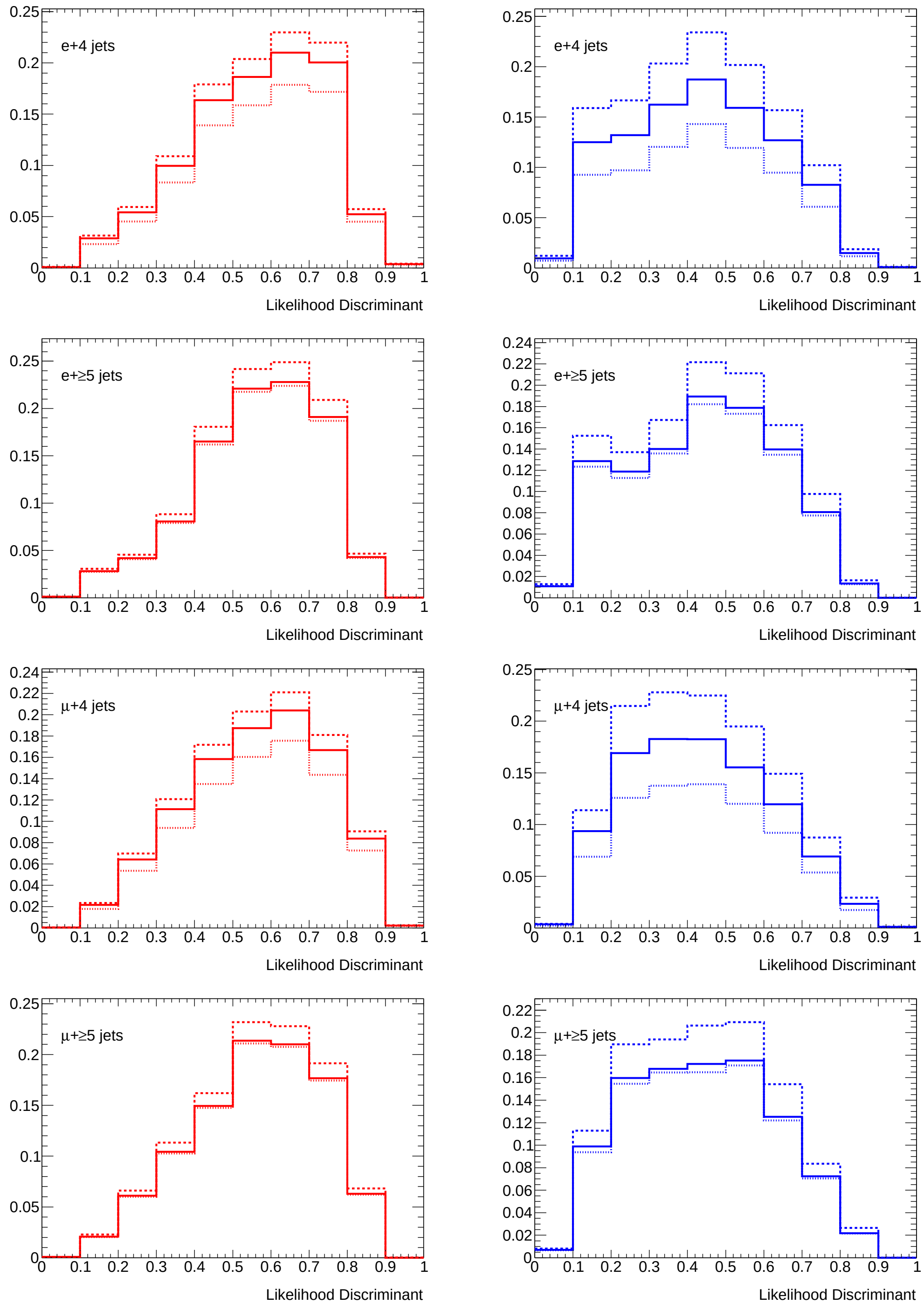


Figure 6.11: Nominal (solid line) and JES shifted (dotted line is due to JES-UP and dashed line is due to JES-DN) LHW templates for the $t\bar{t}$ signal (left) and W+jet background (right) events in the muon and electron channel.

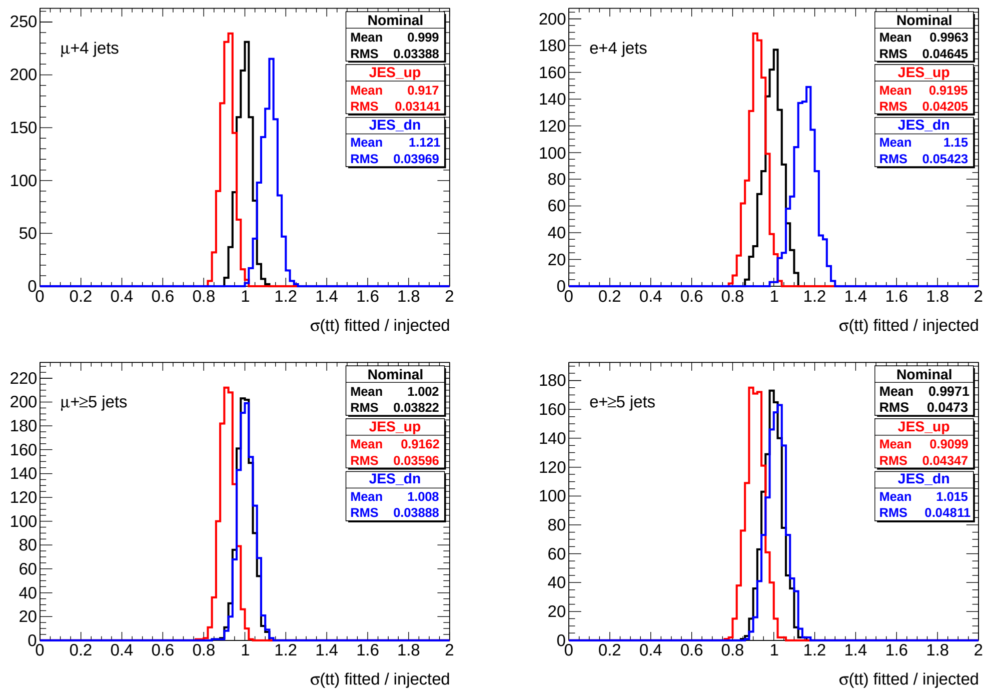


Figure 6.12: The ratios of fitted to predicted number of events obtained with nominal and JES shifted LHW templates in the muon and electron channel.

and modified numbers of $t\bar{t}$ events in each jet multiplicity bin was interpreted as a change in the preselection efficiency due to JER. Change in the fitted number of $t\bar{t}$ events due to variation of JER was obtained as a difference between the numbers of $t\bar{t}$ events obtained with nominal and modified templates. Sets of nominal and modified templates are shown for signal and background in Fig. 6.13. Variation in the number of $t\bar{t}$ events due to modified selection efficiency and template shapes was propagated to the $t\bar{t}+\geq 1$ particle jets cross-section uncertainty. Results of ensemble tests (fitted to predicted number of events) with nominal and JER shifted LHW templates are shown in Fig. 6.14.

Jet Reconstruction Efficiency: The uncertainty due to jet reconstruction efficiency was estimated by randomly dropping a fraction of jets according to the difference in jet reconstruction efficiency on data and MC. It has been calculated in the same way as uncertainty due to JER. Changes in the expected number of $t\bar{t}$ due to se-

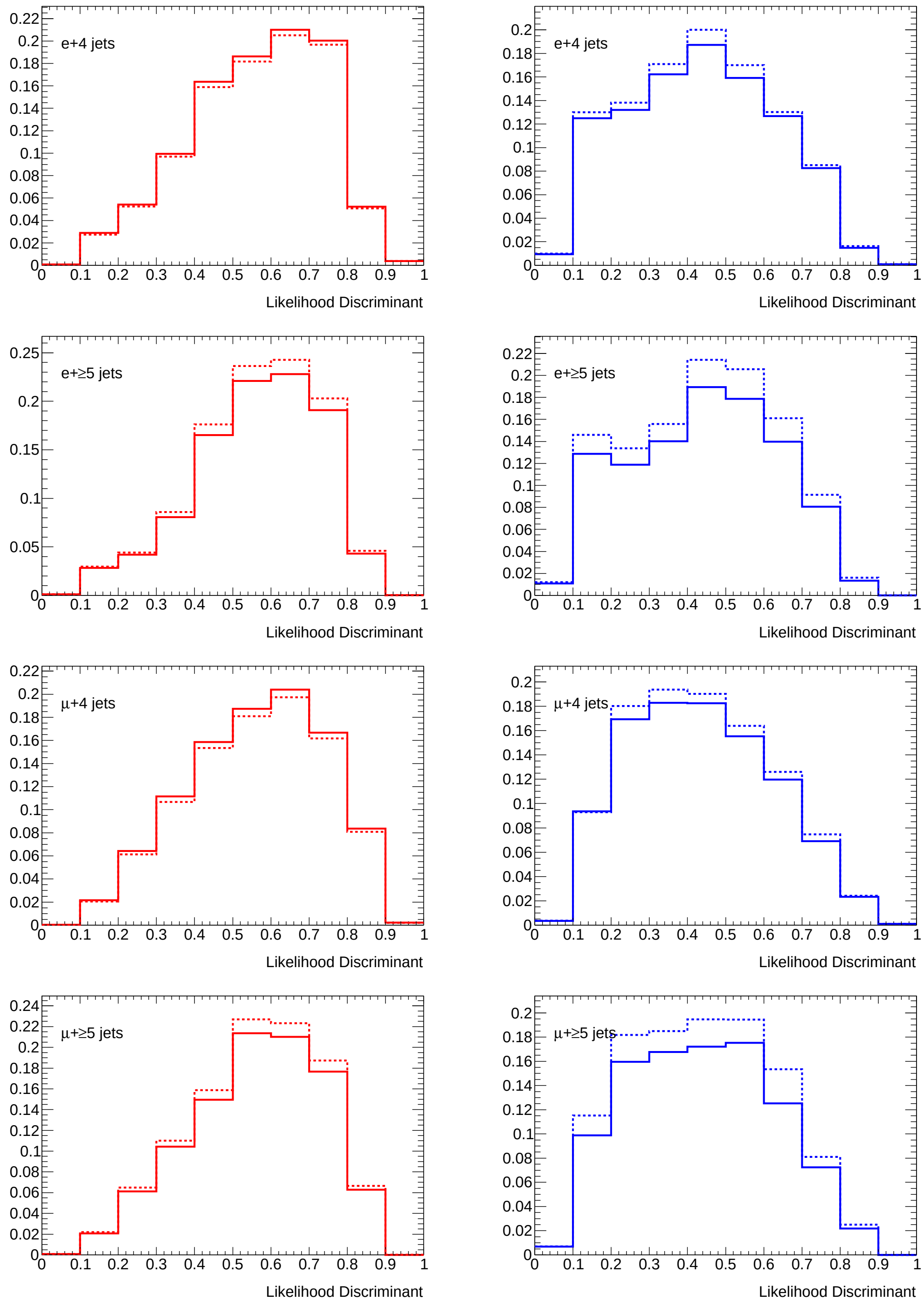


Figure 6.13: Nominal (solid line) and JER shifted LHW templates for the $t\bar{t}$ signal (left) and W+jet background (right) events in the muon and electron channel.

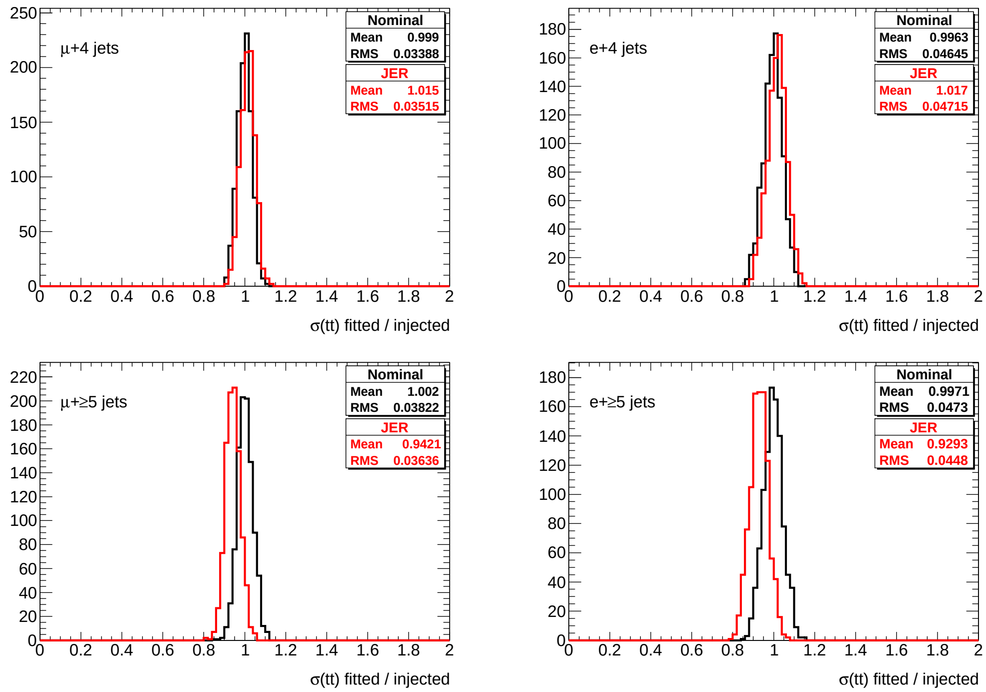


Figure 6.14: The ratios of fitted to predicted number of events obtained with nominal and JER shifted LHW templates in the muon and electron channel.

lection efficiency and likelihood fit were calculated and propagated to the uncertainty on the measured cross-section.

W + jets shape: This source of the uncertainty does not affect the number of $t\bar{t}$ events expected after applying the selection cuts, but it does change the shape of the background template, and therefore, the number of fitted top quark events. The uncertainty due to the shape of W+jets events generated with ALPGEN was estimated using the reweighting tool provided by the TopReco group. Reweighting functions based on the leading jet p_T spectrum were provided for the renormalization scale (iqopt3) and for the functional form of the factorization scale. Reweighting functons were obtained as follows. Truth-level samples with parameters varied with respect to the ATLAS nominal parameter set have been generated and event kinematics were compared.

After reweigting the W+jets MC events, the modified background templates are used to obtain the number of $t\bar{t}$ events. The difference between the nominal number of $t\bar{t}$ events and the new one obtained with modified templates was propagated to the uncertainty on the measured cross-section.

Lepton Identification, Reconstruction Efficiency and Energy/Momentum Resolution: Uncertainties due to lepton identification, energy or momentum resolution and reconstruction efficiencies were estimated by variation of the corresponding parameters by $\pm 1\sigma$ of their uncertainties. A procedure similar to JES uncertainty evaluation was repeated for each of those uncertainties, and variations in fitted numbers of $t\bar{t}$ events were propagated to the cross-section uncertainty.

Small backgrounds: Contributions from small backgrounds, like single top quark production, diboson production and Z+jets production were estimated using their theoretical cross-sections. Systematic uncertainties related to their normalization were evaluated by varying the relevant cross-sections by $\pm 1\sigma_{theor.}$. Fit to data was performed, and difference between nominal and new numbers of $t\bar{t}$ events for each type of background was propagated to the cross-section uncertainty.

QCD background: Systematic uncertainty due to the shape and normalization of QCD background in μ+jets and e+jets was estimated in different ways. QCD contribution in the muon channel was estimated using Matrix Method. The tool used to determine the fraction of QCD events, has also an option to evaluate the systematic uncertainty due to shape and normalization variation by $\pm 1\sigma$. This option is used to obtain new QCD templates in the muon channel and repeated the data fit with modified templates. The difference between nominal and new numbers of $t\bar{t}$ events was propagated to the final systematic uncertainty on the measured cross-section.

There are two methods available for QCD background evaluation in data for the electron channel. These are jet-electron (used as a base method) and matrix method. Uncertainty related to the overall QCD normalization was obtained by variation of the number of QCD events predicted by jet-electron method by $\pm 50\%$. The matrix

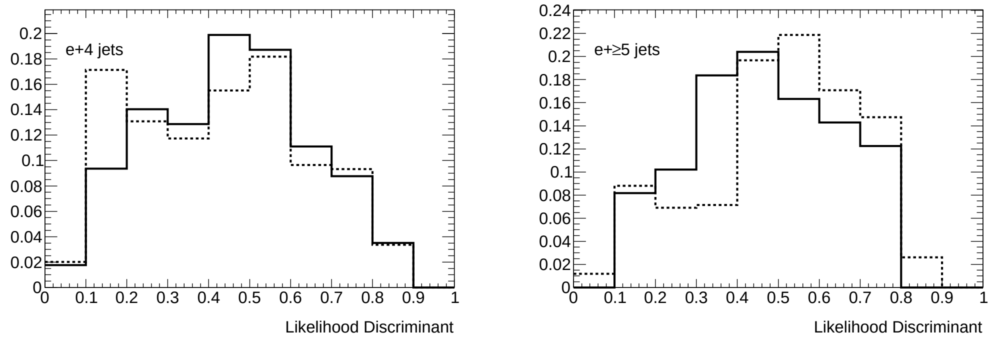


Figure 6.15: QCD templates obtained with jet-electron method (solid histogram) and with matrix method (dashed histogram) in the electron channel.

method is used to derive an alternative QCD templates for e+4 and e+5 inclusive jets bins. Normalization for the total number of expected QCD events in new "Matrix Method" templates was fixed to the numbers obtained with jet-electron method. Comparison of the QCD templates obtained with jet-electron and matrix method is shown in Fig. 6.15. The fit to data with modified templates is repeated and interpreted the difference between resulting cross-sections as a systematic uncertainty due to QCD background.

Pile-up: To evaluate the systematic uncertainty due to pile-up, the average number of interactions per bunch crossing (μ) is scaled up and down by 9% in simulated samples. The reweighted Monte Carlo samples is used to get new templates and selection efficiencies and used them to measure the ttj cross-section. The difference between nominal and recalculated cross-sections was quoted as systematic uncertainty due to pile-up.

6.7.1 Initial and Final State Modeling

Initial and final state radiation (ISR/FSR) changes the number of additional particle jets in the event. Since the kinematics of these events can be different from the nominal, the ISR/FSR variation may also affect the efficiencies used to calculate the

ttj cross-section according to Equation. 6.1. The variations of efficiencies due to ISR/FSR changing up and down are summarized in Table 6.5.

	ε_{04}	ε_{05}	ε_{14}	ε_{15}
e+jets, ISR/FSR up	2.35	0.29	2.89	3.12
e+jets, ISR/FSR down	2.57	0.29	3.02	2.79
μ+jets, ISR/FSR up	3.82	0.47	4.68	5.10
μ+jets, ISR/FSR down	4.19	0.48	4.73	4.44

Table 6.5: $tt0$ and ttj reconstruction efficiencies for samples with ISR/FSR shifted up and down (%).

The variation of measured ttj production cross-section due to ISR/FSR is shown in Table 6.7.1, and the variation of measured ratio of ttj to inclusive $t\bar{t}$ cross-section is shown in Table 6.7.1. The variations are defined as

$$\Delta\sigma_{ttj} = \frac{\sigma_{ttj}^{up} - \sigma_{ttj}^{down}}{\sigma_{ttj}^{up} + \sigma_{ttj}^{down}}, \qquad \Delta f_1 = \frac{f_1^{up} - f_1^{down}}{f_1^{up} + f_1^{down}} \tag{6.6}$$

where indices up and $down$ indicate the ISR/FSR modified up and down, respectively. Part of the variation is due to the difference in the value of f_1 in up and down samples, denoted as Δf_1^{MC} (higher ISR/FSR leads to increased extra jet production). If the reconstruction efficiencies ε_{ij} were the same in both samples, the whole change in measured ttj cross-section and ttj fraction would be due to Δf_1^{MC}. Since it is not the case, in addition to Δf_1^{MC} there is a second component due to efficiency change. As it is small compared to the change in the ttj fraction between the samples, our analysis is potentially capable of probing the ISR/FSR production model, but its sensitivity is diluted by other systematic uncertainties, particularly jet energy scale and jet energy resolution.

The ttj cross-section measurement is performed using different MC generators interfaced to the different parton showering models, and compared obtained results in the next section.

channel	σ_{ttj} variation due to efficiency	Δf_1^{MC}	total σ_{ttj} variation
e+jets	−5.2		−14.2
μ+jets	−7.1	+9.1	−16.1
total	−6.2		−15.2

Table 6.6: Variation of the ttj production cross-section in the current MC ISR/FSR model (%).

channel	f_1 variation due to efficiency	Δf_1^{MC}	total f_1 variation
e+jets	−2.2		−11.2
μ+jets	−3.4	+9.1	−12.4
total	−2.8		−11.8

Table 6.7: Variation of the ratio of ttj to inclusive $t\bar{t}$ production cross-section in the current MC ISR/FSR model (%).

6.7.2 Comparison to other Monte Carlo Generators

The presented measurement of the ttj cross-section is model dependent. It depends on (a) kinematics of inclusive $t\bar{t}$ events in a given generator and (b) on the efficiencies ϵ_{ij} obtained with this generator. As it was shown in previous sections, the detailed comparison of kinematic distributions is done in data and in MC with MC@NLO predictions for $t\bar{t}$ and found a good agreement between data and MC. Nevertheless it is important to check how stable is obtained result against the different MC generators and parton shower schemes. This section contains the results of measurement of the ttj cross-section using ALPGEN interfaced to JIMMY and POWHEG interfaced to PYTHIA for simulation of inclusive $t\bar{t}$ production.

In POWHEG the hardest radiation is generated first, and then the results of generation are interfaced to PYTHIA generator for subsequent, soft radiation. The details on the POWHEG generator [59] and on the comparison of POWHEG with MC@NLO can be found in [60]. The $t\bar{t}$ events generated with POWHEG are used and repeated the analysis excluding the evaluation of systematic uncertainties. The main steps of the analysis with POWHEG are documented in [61] , which contains the comparison of kinematic distributions, likelihood shapes, likelihood fits and the measured ttj cross-section.

ALPGEN is a MC generator based on Leading Order (LO) matrix element calculations [54]. ALPGEN generator has been interfaced to JIMMY for the following parton showering. ALPGEN interfaced to JIMMY has been used to produce the $t\bar{t} + n$ hard partons samples, where n changes from 0 to 3. ALPGEN generator describes modified aplanarity distributions slightly better compared to POWHEG and MC@NLO, but has problems with describing jet p_T distributions. It needs to be mention that the ALPGEN samples of $t\bar{t}$ have not been produced in association with heavy flavor jets, assuming that their kinematics is pretty close to the $t\bar{t}$ plus light parton jets.

As an additional cross check of the results obtained with different MC generators, the inclusive cross-section is calculated with signal templates from MC@NLO, POWHEG and ALPGEN. The results are presented in Table 6.8.

Generator	$N_{tt}^{data}/N_{tt}^{expected}$	$\sigma_{t\bar{t}}$ [pb]
MC@NLO	1.13	189 ± 4
POWHEG	1.13	189 ± 4
ALPGEN	0.97	163 ± 3

Table 6.8: Inclusive $t\bar{t}$ production cross-section $\sigma_{t\bar{t}}$ measured with $t\bar{t}$ templates obtained with different generators. The second column shows the ratio of observed to expected number of $t\bar{t}$ events. The quoted errors are statistical only.

These values can be compared to the theoretical value of $\sigma_{tt}^{theor.} = 166.8^{+16.5}_{-17.8}$ pb. One can see that with NLO generators the measured cross-section is 13% higher than the theoretical one, while using ALPGEN generator results in pretty much the same cross-section as the theoretical. It should be mentioned that the systematic uncertainty of these measurements is of the order of 13% (as it was evaluated for ttj cross-section) and the systematic error on the theoretical cross-section is of the order of 10 %, so all of them agree with the SM calculations.

The ttj cross-section has been calculated using Equation. 6.2 for each generator.

Generator	k_1	f_1	σ_{ttj} [pb]	Measured f_1
MC@NLO	1.21	0.503	102 ± 2	0.54 ± 0.01
POWHEG	1.11	0.568	105 ± 2	0.56 ± 0.01
ALPGEN	0.92	0.524	81 ± 2	0.50 ± 0.01

Table 6.9: The ttj production cross-section σ_{ttj} (the fourth column) and the ratio f_1 of the ttj to inclusive $t\bar{t}$ production cross-section (the fifth column) measured with $t\bar{t}$ templates obtained with different generators. The second column shows the ratio of observed to expected number of ttj events, and the third column shows the value of f_1 predicted by simulation. The quoted errors are statistical only.

Results obtained with MC@NLO and POWHEG are in a good agreement with each other. Although the fraction of the ttj events in ALPGEN is in between MC@NLO and POWHEG, the observed fraction of the ttj events is smaller compared to the expected one.

Another difference between MC samples generated with ALPGEN and MC@NLO is that ALPGEN sample was produced using leading order PDF, while MC@NLO events were produces with next-to-leading order PDF. This might potentially affect the kinematics of produced $t\bar{t}$ events, and as a result, the measured $t\bar{t}+$ jets cross-section.

Detailed studies of all three generators are presented in [61].

The largest difference between MC@NLO, POWHEG, and ALPGEN ($\pm 21\%$) id quoted as the systematic uncertainty due to choice of MC generator.

6.8 Cross Section in Fiducial Volume

The ttj cross-section has estimated in fiducial volume for direct comparison of our results with theoretical calculations. The definition of fiducial volume is based on particle-level objects as described in Section 6.3. Muons are required to have $p_T >$ 20 GeV, $|\eta| < 2.5$, and electrons are required to have $p_T > 25$ GeV, $|\eta| < 2.5$. The "dressed" electrons (four-momenta of photons within $\Delta R < 0.1$ from the electron

	e+jets		μ+jets	
generator	$\varepsilon_{acc}^{fiducial}$	$\sigma_{ttj}^{fiducial}$ [pb]	$\varepsilon_{acc}^{fiducial}$	$\sigma_{ttj}^{fiducial}$ [pb]
MC@NLO	0.216	2.59 ± 0.08	0.261	3.47 ± 0.08
POWHEG	0.189	2.39 ± 0.07	0.223	3.14 ± 0.07
ALPGEN	0.247	2.57 ± 0.08	0.301	3.17 ± 0.08

Table 6.10: The ttj acceptance efficiencies and cross sections in fiducial volumes in e+jets and μ+jets channels. Fiducial volumes in electron and muon channels are different (described in the text). Errors on measured cross sections are statistical only.

direction are added to the electron four-momentum) is considered. This procedure is not applied to muons. In both e+jets and μ+jets channel, exactly one lepton and at least five jets with $p_T > 25$ GeV, $|\eta| < 2.5$ are required. In addition in e+jets channel it is required $|MET| > 30$ GeV, $MTW > 30$ GeV , and $|MET| > 20$ GeV, $|MET| + MTW > 60$ GeV in μ+jets channel. Here MET is defined as the vector sum of the transverse energy of all neutrinos originating from W bosons from top quarks, and MTW is the invariant transverse mass of MET and the lepton. Events passing these cuts are considered to be in fiducial volume.

The acceptance efficiencies is evaluated for the $ttj \rightarrow e +$ jets and $ttj \rightarrow \mu +$ jets events in fiducial volume $\varepsilon_{acc}^{fiducial}$ for three Monte Carlo generators (MC@NLO, POWHEG, and ALPGEN). Using these efficiencies and fitted values of k_1, the cross sections is obtained in fiducial volume $\sigma_{ttj}^{fiducial}$. The results are summarized in Table 6.8. For the systematic uncertainty due to different Monte Carlo generators, we quote the largest difference between the three results (±10% in the μ+jets channel and ±8% in the e+jets channel).

LO and NLO calculations of σ_{ttj} in fiducial volume for muon channel have been done by Melnikov, Schulze and Scharf [62]. Fiducial volume was defined using cuts applied in our analysis. They found $\sigma_{ttj}^{LO} = 2.20^{+1.34}_{-0.77}$ pb and $\sigma_{ttj}^{NLO} = 2.35^{+0.05}_{-0.26}$ pb.

6.9 Measuring σ_{ttj} with 2-Dimensional Likelihood Method

In 2-Dimensional Likelihood, the kinematic difference between the $tt0$ and ttj events used in addition to efficiencies. A 2D likelihood is constructed out of two likelihoods: one capable of separating the $t\bar{t}$ events from the W+jets background (denoted LHW, the same likelihood which is used in method 1), and another one which distinguishes between the $tt0$ and ttj events (denoted LHD). The formula for expected number of events (Equation. 4.8) is modified as

$$\nu_{ij} = p_0 s_{ij}^{tt0} + p_1 s_{ij}^{ttj} + p_j b_{ij} + m_{ij}, \tag{6.7}$$

where s_{ij}^{tt0} and s_{ij}^{ttj} are predicted numbers of $tt0$ and ttj events, respectively. Since this method constraints the fraction of ttj events in each jet multiplicity bin, it is expected to be less sensitive to systematic effects resulting in migration of events between the bins (e.g. the jet energy scale uncertainty).

6.9.1 Variables used in the 2-Dimensional Likelihood

The variables used to separate the ttj from $tt0$ are $\Delta\eta(j1, j3)$ between first and third leading jets, and jet M_T, transverse mass of all jets in the event as defined at section 4.2. These variables are uncorrelated for both signal and background.

Agreement between data and MC for all variables is shown in Figure. 6.9.1 for the muon channel and in Figure 6.9.1 for the electron channel. Sum of all MC events and QCD is normalized to the number of observed events in data to compare the shapes of these distributions.

The LHD likelihood distributions in 4th and 5th jet multiplicity bins are shown in Figure 6.9.1 for electron and muon channel, respectively.

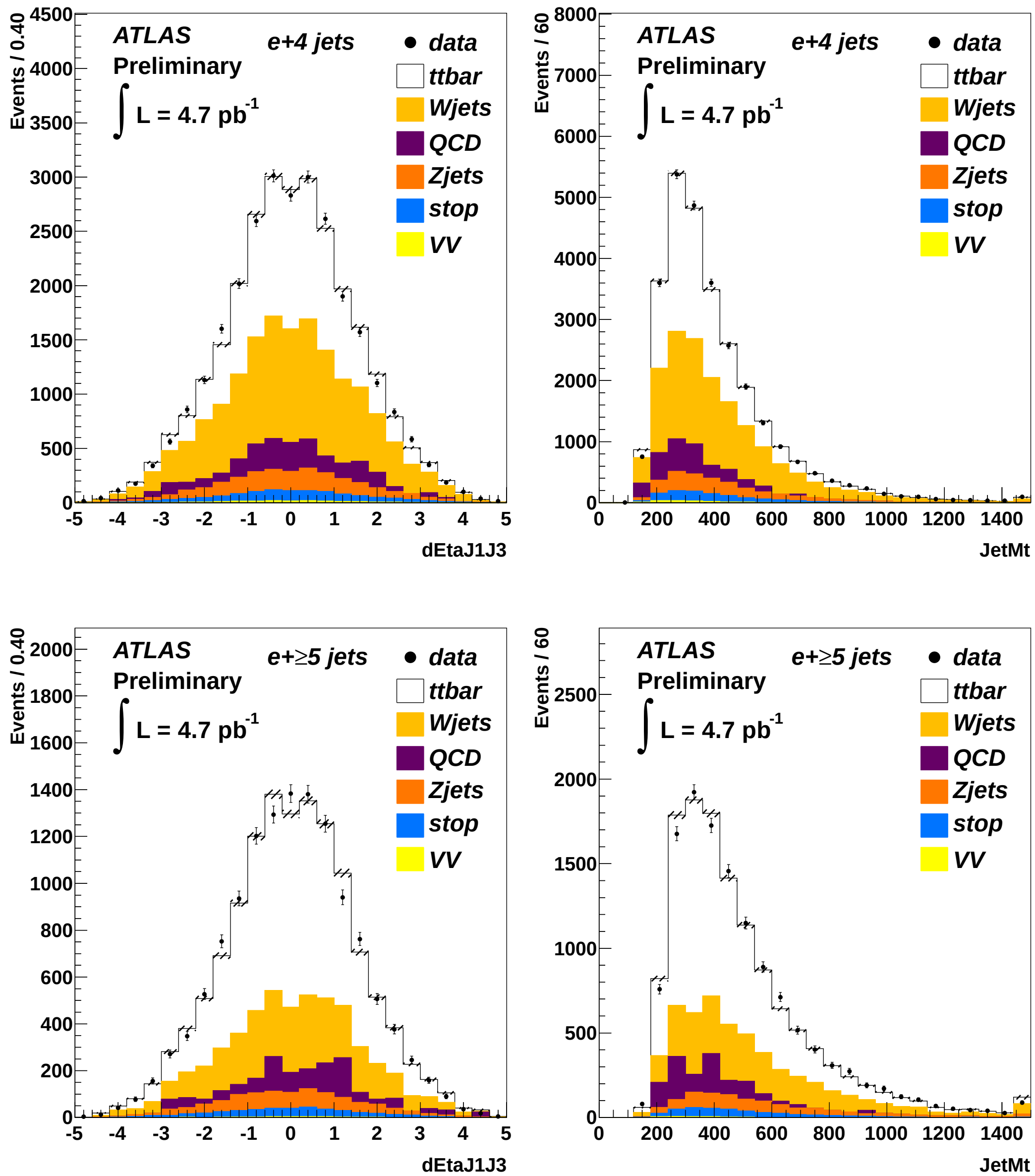


Figure 6.16: Control plots for chosen variables in electron channel.

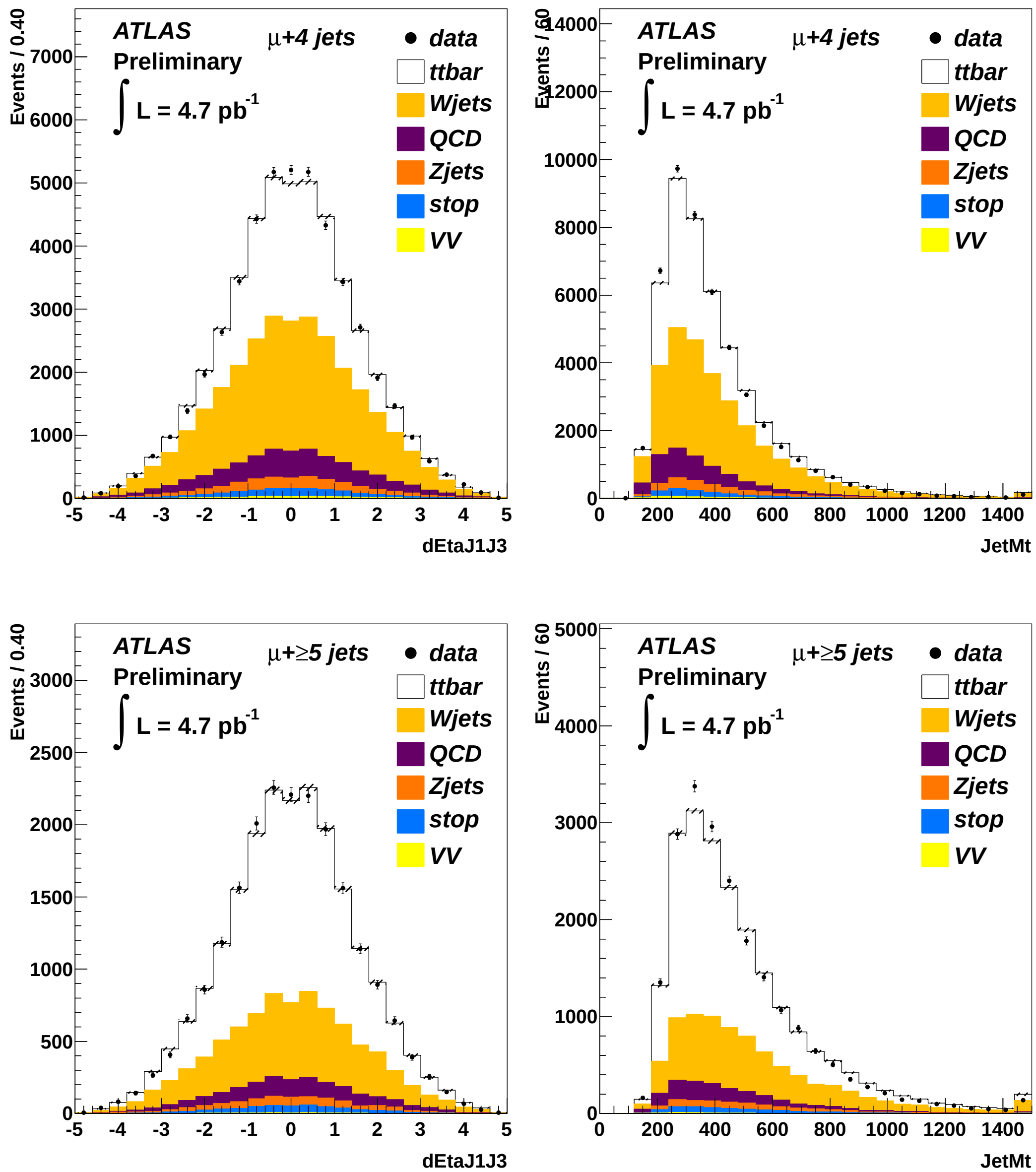


Figure 6.17: Control plots for chosen variables in muon channel.

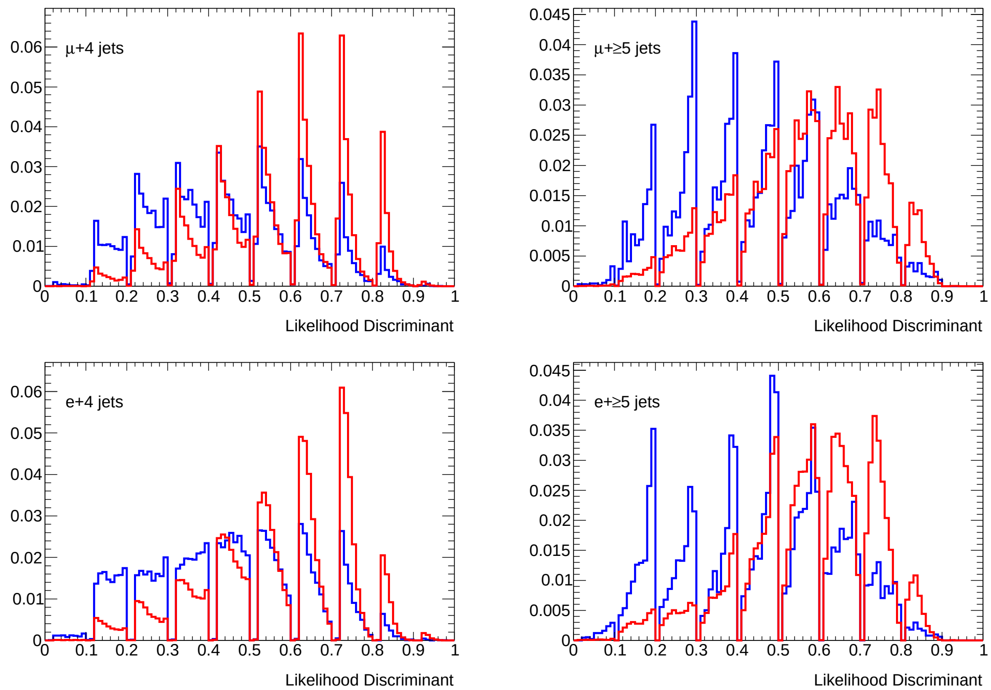


Figure 6.18: LHD templates for muon and electron channels.

6.9.2 Results of Fit to Data

Results of the likelihood fits to data in all channels are shown in Figure. 6.9.2. The resulting numbers of $tt0$, ttj events from these fits are summarized in Table 6.11.

Channel	N^{ttj}/N^{ttj}_{pred}
$e+$ jets	1.11 ± 0.03
$\mu+$ jets	1.20 ± 0.02
Lepton + jets	1.16 ± 0.02

Table 6.11: Observed to predicted event ratios N^{ttj}/N^{ttj}_{pred} in electron and muon channels and in the combined lepton+ jets channel. The errors are statistical only.

The ttj production cross section is found to be

$$\sigma_{ttj} = 114 \pm 2(\text{stat.})\,\text{pb}. \tag{6.8}$$

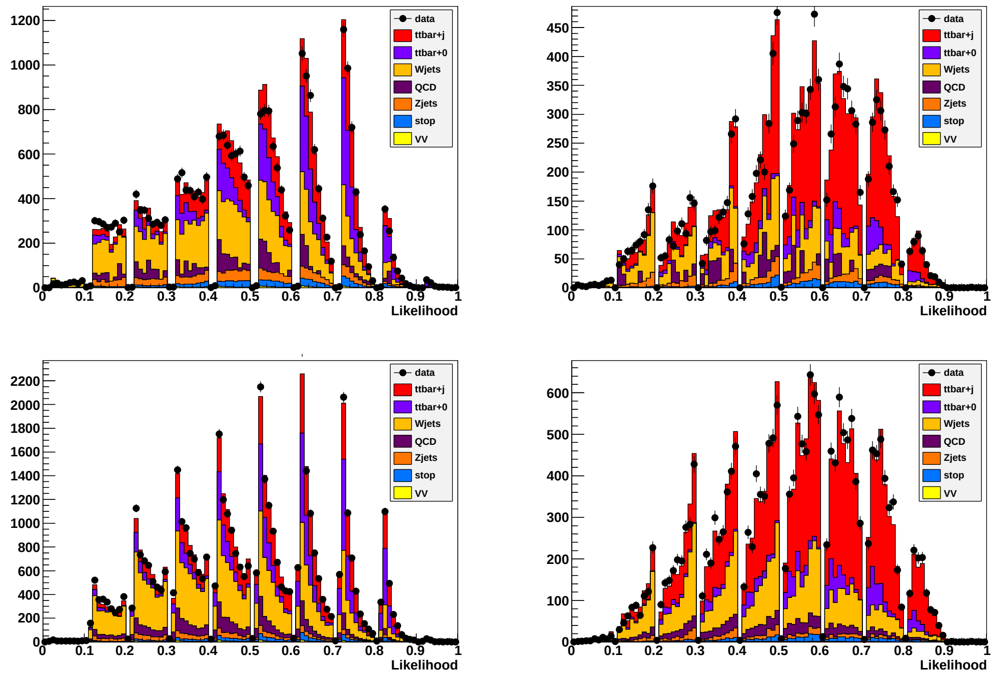


Figure 6.19: Likelihood fits in e+jets and $\mu+$ jets channels.

6.9.3 Systematic Uncertainties

Systematic uncertainties were estimated in the same manner as for method 1 (see 6.7).

Summary of the systematical errors are presented in Table 6.9.3.

6.9.4 Results 2-Dimensional Likelihood method

The measured cross section $t\bar{t}+\geq 1$ hard parton jet is found to be

$$\sigma_{ttj} = 114 \pm 2(\text{stat.}) \pm 15(\text{syst.})\,\text{pb} \tag{6.9}$$

The largest systematic uncertainties in this measurement are due to JES and Jet Energy Resolution. Since the method 2 is only for cross check of central value, the MC Generators, MC parton cut and ISR/FSR uncertainties are not done and included in total systematic. This result is in a good agreement with the one obtained with method 1.

Source	electron +jets	muon +jets	combined
JES Up, Down	-6.4 -8.8	-6.5 -9.9	-6.5 - 9.4
Jet Energy Resolution	-8.8	-8.5	-9.0
Jet Reconstruction Efficiency	-0.2	-0.3	-0.1
MET Cell out	+0.9 +0.8	-0.1 +0.4	+0.3 +0.6
Electron Energy Smearing	-0.4 -0.3	0	-0.1 -0.1
Electron Energy Scale	-0.5 -0.4	0	-0.2 -0.2
Electron Trigger Scale Factor	-1.4 +1.5		-0.6 +0.6
Electron Reco Eff. SF	-1.2 +1.2		-0.5 +0.5
Electron ID	-1.5 +1.5		-0.6 +0.6
Muon Momentum Smearing		-0.2 -0.1	-0.1
Muon Momentum Scale		-0.1	-0.2
Muon Trigger Scale Factor		-2.8 +2.9	-1.7 +1.7
Muon Reco Eff. SF		-0.3 +0.3	-0.2 +0.2
Muon ID		-0.1 -0.4	-0.1 -0.2
Muon ID SF		-0.7 +0.7	-0.4 +0.4
QCD electrons normalization	$+0.7 - 1.0$		$+0.2 - 0.3$
QCD muons normalization		± 0.8	± 0.5
QCD electrons shape	± 6.0		± 2.2
QCD muons shape		< 0.1	< 0.1
Z Background	± 1.0	± 0.1	± 0.3
Single top	± 0.6	± 0.3	± 0.4
Dibosons	< 0.1	< 0.1	< 0.1
W shape	-1.1 -0.8	-1.0 -1.0	-1.0 -0.9
PDF	± 0.9	± 0.9	± 0.9
Pileup	+1.1 +0.2	+2.5 +1.0	+1.9 +0.7
Monte Carlo Generator and parton shower	-3.5	+1.2	-1.0

Table 6.12: Summary of systematic errors in electron, muon channels and their combination (%).

6.10 Summary

The complete list of systematic uncertainties is presented in Table 6.10. The total uncorrelated systematic uncertainty which includes lepton ID, triggers, lepton reconstruction efficiency, and QCD estimation, is 4.9% in e+jets channel and 3.3% in μ+jets channel. The total uncorrelated uncertainty in two channels is 7%, which is of the order of the difference of the measurement in the two channels as shown in Table 6.4. Table 6.10 presents the complete list of systematic uncertainties calculated for the measured ratio of ttj to $t\bar{t}$ inclusive production cross section.

This chapter presents the measurement of the cross-section of $t\bar{t}+\geq 1$ jet production on the ATLAS data corresponding to the integrated luminosity of f 4.7 fb^{-1}. The measurement is done in a single lepton final state, requiring a presence of high-p_T isolated muons or electrons. The measured cross-section of $t\bar{t}+\geq 1$ jet is found to be

$$\sigma_{ttj} = 102 \pm 2(\text{stat.}) \pm 6(\text{ISR/FSR})^{+22}_{-25}(\text{syst.})\,\text{pb} \qquad (6.10)$$

The largest systematic uncertainties in this measurement are due to JES, Jet Energy Resolution and MC generators.

The cross sections in the fiducial region in e+jets and μ+jets channels are found to be

$$\sigma^{fiducial}_{ttj\to e+jets} = 2.59 \pm 0.09(\text{stat.}) \pm 0.15(\text{ISR/FSR})^{+0.21}_{-0.44}(\text{syst.})\,\text{pb} \qquad (6.11)$$

$$\sigma^{fiducial}_{ttj\to \mu+jets} = 3.48 \pm 0.08(\text{stat.}) \pm 0.21(\text{ISR/FSR})^{+0.38}_{-0.58}(\text{syst.})\,\text{pb} \qquad (6.12)$$

$$\sigma^{fiducial}_{ttX\to e+\geq 5jets} = 4.09 \pm 0.18(\text{stat.}) \pm 0.24(\text{ISR/FSR})^{+0.57}_{-0.82}(\text{syst.})\,\text{pb} \qquad (6.13)$$

$$\sigma^{fiducial}_{ttX\to \mu+\geq 5jets} = 5.27 \pm 0.16(\text{stat.}) \pm 0.31(\text{ISR/FSR})^{+1.00}_{-1.16}(\text{syst.})\,\text{pb} \qquad (6.14)$$

The cross section in muon channel in fiducial volume is within two standard deviations from theoretical NLO value, $\sigma^{N}_{ttj}LO = 2.35$ pb.

Source	electron +jets (%)	muon +jets (%)	combined (%)
JES Up, Down	-9.1 -3.2	-8.7 -3.5	-8.9 -3.3
Jet Energy Resolution	-10.0	-8.7	-9.2
Jet Reconstruction Efficiency	-0.2	-0.1	-0.1
MET Cell out	+0.4 +0.5	-0.1 +0.4	+0.2 +0.4
Electron Energy Smearing	-0.4 -0.3		-0.2 -0.2
Electron Energy Scale	-0.3 -0.3		-0.2 -0.2
Electron Trigger Scale Factor	-1.4 +1.5		-0.6 +0.6
Electron Reco Eff. SF	-1.1 +1.1		-0.5 +0.5
Electron ID	-1.6 +1.7		-0.7 +0.7
Muon Momentum Smearing		-0.1 +0	-0.1
Muon Momentum Scale		<0.1	0.1
Muon Trigger Scale Factor		-2.6 +2.7	-1.5 +1.5
Muon Reco Eff. SF		-0.3 +0.3	-0.2 +0.2
Muon ID		-0.1 -0.4	-0.1 -0.2
Muon ID SF		-0.7 +0.7	-0.4 +0.4
QCD electrons normalization	$-1.7 + 1.6$		±0.7
QCD muons normalization		±0.8	±0.5
QCD electrons shape	-6.0		-2.2
QCD muons shape		< 0.1	< 0.1
Z Background	±0.3	±0.1	±0.1
Single top	±0.3	±0.2	±0.3
Dibosons	< 0.1	< 0.1	< 0.1
W shape	+0.1 -0.2	+0.2 -0.3	+0.1 -0.3
PDF	±0.9	±0.9	±0.9
Pileup	+0.8 +0.5	+2.6 +0.5	+1.9 +0.5
MC parton cut	+1.2-1.6	+1.3-1.7	+1.2-1.7
MC Generator	±19	±23	±21
Total systematic error	+19 -24	+23 -26	+22 -25

Table 6.13: Summary of relative systematic errors on ttj production cross-section in electron and muon channels and their combination (%).

Source	electron +jets (%)	muon +jets (%)	combined (%)
JES Up, Down	-1.0 -10.5	-0.3 -9.0	-0.6 -9.5
Jet Energy Resolution	-7.8	-6.6	-7.1
Jet Reconstruction Efficiency	-0.1	-0.1	-0.1
MET Cell out	-0.3 + 0.1	+0.1 +0.0	< 0.1
Electron Energy Smearing	-0.1 -0.0		-0.1 -0.1
Electron Energy Scale	+0.3 -0.1		+0.1 -0.1
Electron Trigger Scale Factor	< 0.1		< 0.1
Electron Reco Eff. SF	+0.1 -0.1		< 0.1
Electron ID	-0.1 +0.1		-0.1 +0.1
Muon Momentum Smearing		-0.1 +0	-0.1
Muon Momentum Scale		< 0.1	+0.1 +0
Muon Trigger Scale Factor		+0.1 -0.1	+0.2 -0.2
Muon Reco Eff. SF		< 0.1	< 0.1
Muon ID		-0.1 -0.0	< 0.1
Muon ID SF		< 0.1	< 0.1
QCD electrons normalization	$-0.4 + 0.3$		±0.2
QCD muons normalization		±0.2	±0.2
QCD electrons shape	-6.5		-2.5
QCD muons shape		< 0.1	< 0.1
Z Background	±0.3	±0.1	±0.1
Single top	±0.2	±0.1	±0.2
Dibosons	< 0.1	< 0.1	< 0.1
W shape	-0.6 -0.5	-0.5 -0.5	-0.5 -0.5
PDF	±0.1	±0.1	±0.1
Pileup	-0.7 +0.1	+0.9 +0.5	+0.3 +0.3
MC parton cut	+1.2-1.6	+1.3-1.7	+1.2-1.7
MC Generator	±8	±9	±8
Total systematic error	+8 -17	+9 -14	+9 -14

Table 6.14: Summary of relative systematic errors on the ratio of ttj to inclusive $t\bar{t}$ production cross-section in electron and muon channels and their combination (%).

The ratio $f_1^{\text{meas.}}$ of σ_{ttj} to $\sigma_{t\bar{t}}^{incl}$ is determined to be

$$f_1^{\text{meas.}} = 0.54 \pm 0.01(\text{stat.}) \pm 0.02(\text{ISR/FSR})^{+0.05}_{-0.08}(\text{syst.}) \tag{6.15}$$

Chapter 7

Study of Radiation Hardness of PiN Diodes for the Pixel Detector Upgrade

The LHC is designed to deliver a nominal luminosity of 10^{34} $m^{-2}s^{-1}$ during the first ten years of its operation. As an extension of the LHC program, the project called Super Large Hadron Collider (SLHC) was proposed [63]. Its goal is to increase the luminosity of the machine to 10^{35} $m^{-2}s^{-1}$. The increase of luminosity leads to a significantly higher collision rate resulting in a larger data sets that can be collected by LHC experiments. Essential increase in the statistics will improve the sensitivity of the LHC experiments to new physics and allow them to perform more precise measurements. Recently new initiative, High Luminosity LHC (HL-LHC) program, was proposed as an intermediate upgrade program. It aims at providing luminosity of $5 \times 10^{34} m^{-2}s^{-1}$.

At SLHC the beam luminosity will be increased by a factor of ten compared to the LHC. Accordingly the radiation level is expected to increase by a similar factor. Since the pixel detector is the innermost part of the ATLAS detector, it will suffer most

	Total designed Luminosity after 5 years	readout link speed
LHC	300 fb^{-1}	40 Mb/s
SLHC	3000 fb^{-1}	624Mb/s $\rightarrow$ 3.24Gb/s

Table 7.1: Expected total integrated luminosity after 5 years of operation and designed optical link speed.

among the all ATLAS subsystems from the radiation damage. It has to be upgraded to address the higher radiation environment of the SLHC. The SLHC upgrade design not only requires higher radiation hardness of all parts of the pixel detector, but also much faster readout system. The requirements to the readout system of the pixel detector are shown in Table 7.1.

There are two dominant mechanisms that lead to performance degradation of semiconductor devices. First mechanism is due to the damage of the surface layer. In the surface layers, such as SiO_2 in silicon based devices, the electrons produced due to ionization quickly carry away from the oxide layer. But they leave behind the holes which have low mobility. These holes can be trapped at the SiO_2-Si interface, resulting in an increase of the oxide positive charge. The accumulated positive charge leads to attraction of free electrons from the bulk, and results in an increase of the voltage necessary to deplete the pixels completely. This effect leads to decreasing of the charge collection efficiency.

The second effect is displacement damage or bulk damage which is caused by the incident radiation displacing atoms from their lattice sites. The bulk damage depends on the type and energy of the incident radiation. It results in increasing of leakage current, change of the effective impurity concentration and loss in the charge collection efficiency. Therefore, it is important to scale measurements on radiation damage for different kinds of radiation.

The damage effects by energetic particles in the bulk of material can be described as being proportional to the so called displacement damage cross section. This quan-

	1-MeV $[n_{eq}/cm^2]$	200 MeV $[p/cm^2]$	24 GeV $[p/cm^2]$
Si	1.5×10^{15}	1.4×10^{15}	2.6×10^{15}
GaAs	8.2×10^{15}	1.2×10^{15}	1.6×10^{15}

Table 7.2: Beam fluences for Si and GaAs based diodes

tity is equivalent to the Non Ionizing Energy Loss (NIEL) and hence the proportionality between the NIEL value and the resulting damage effects is referred to as the NIEL-scaling hypothesis (more detail at [64]). The NIEL scaling hypothesis is used to estimate the SLHC fluences at the present optical link location patch panel zero of the pixel detector of the ATLAS [65] [66]. After five years of operation the SLHC is expected to achieve $3000\,fb^{-1}$ of integrated luminosity. A a result the pixel detector is expected to be exposed to a maximum total fluence of 1.5×10^{15} 1-MeV n_{eq}/cm^2 for Si components and corresponding fluence for a GaAs component (VCSEL) is 8.2×10^{15} 1-MeV n_{eq}/cm^2. [67]

The fluences of protons with momenta of 200 MeV and 24 GeV corresponding to the maximum expected fluences at SLHC are presented in Table 7.2. The results of Table 7.2 can be translated into radiation dosage of 69 and 34 Mead, for Si and GaAs based devices, respectively.

The purpose of experiment described in this thesis is to identify the PiN diode candidates which can withstand the SLHC dose and will have sufficient speed for the optical readouts to be used in tracking detectors at the SLHC. In order to accomplish this goal, special test stands are designed.

7.1 Pixel Detector and Opto-link

The pixel detector described in 2.3.2 is the innermost element of the ID and plays an important role in track reconstruction of charged particles. Figure 7.1 presents block diagram of the pixel detector module. Each pixel module contains 16 Front End

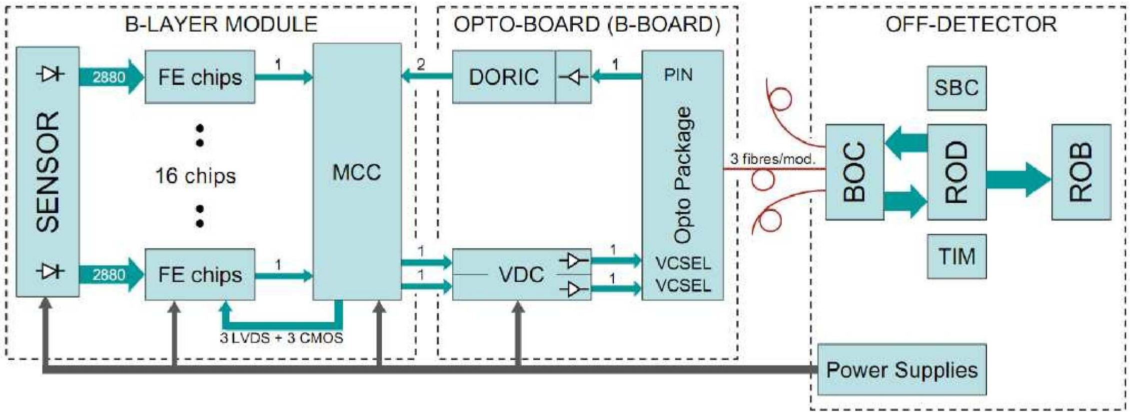


Figure 7.1: Block diagram of the pixel detector module.

(FE) chips which are controlled by a Module Control Chip (MCC) [9]. A FE chip accommodates the readout for 2,880 pixels and transmits signals from them to MCC. Each module is linked to the off-detector readout system through the opto-link. The opto-link is the best option for the pixel detector readout, since optical fibers have very small footprint and are capable of handling the high speed data and clock rates.

Opto-boards are located at the Service Quarter Panel (SQP) which are located on both sides of the detector and connected pixel modules MCC by to micro-twisted-pair wires (MTP).

Electrical signals come in a group of 7 or 14 MTP wires from MMC of pixel modules to Vertical Cavity Surface Emitting Lasers (VCSEL) controlled by the VCSEL Driver Chip (VDC) on an opto-board. The VCSEL diode converts the data signals to optical signals and transmits them to an RX plug-in mounted on the Back of Crate card (BOC) via optical fibers. Similarly Bi-Phase Modulated (BPM) clock and data of control signals are received by an array of PIN diodes on the opto-board and are separated by a Digital Optical Receiver IC (DORIC). The interconnection process is illustrated in Figure 7.2. The optical components at off-detector side are not subject to irradiation expose. Each optical link is responsible to transfer data from almost 45000 pixels. Therefore the reliability and radiation hardness of optical components

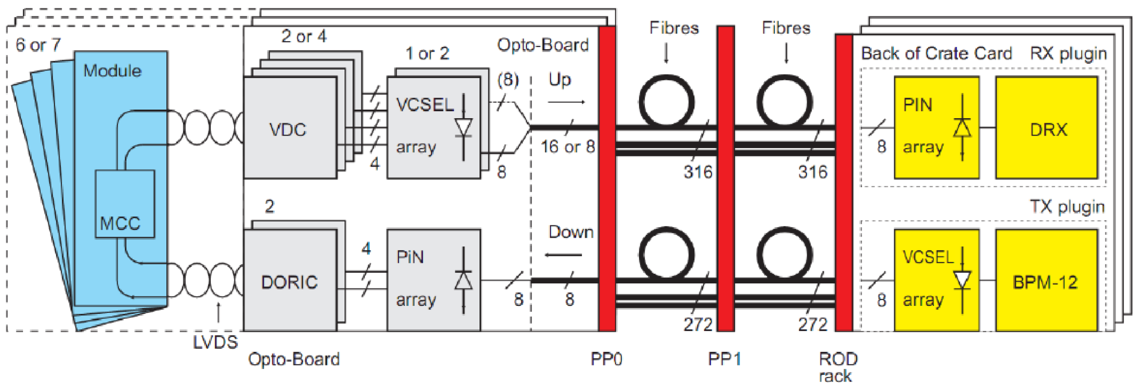


Figure 7.2: Block diagram of interconnections of Opto-boards to Pixel modules and off detector readout system via optical fibers from PP0 to BOC cards.

is an important matter to be investigated. In this thesis the focus is of the study of radiation hardness of PiN diodes.

7.2 PiN Diodes Selected for Tests

Following by study of the market for the available PiNs, six types of PiN diodes have been chosen for radiation hardness test:

- Si PiN diodes S9055-01 and S5973-01 (single devices) are silicon based PiN photodiodes. These diodes have a flat responsivity up to 1 GHz (responsivity is defined as the ratio of generated photo-current to incident light power. The responsivity may also be expressed as the ratio of the number of photogenerated carriers to incident photons). The PiN diodes need low bias voltage ($V_R = 2$ V). They come in 3-pin TO-18 package as shown in Figure 7.3.

- GaAs PiN diode G8522-01/3 (single device). This family includes three types with different optical active area size and cut-off frequency but the same physical structure, which provides an excellent opportunity to study the radiation hardness vs PiN diode frequency. They are packaged in TO-18 format too.

- GaAs PiN diode G8921-01 (diode die is available in 4 and 12 channel configura-

PiN diode	S9055-01	S5973-01	G8921-01	G8522-01	G8522-02	G8522-03
Peak Wavelength [nm]	700	760	850	850	850	850
Photo Sensitivity @850nm [A/W]	0.25/0.37	0.52	0.5	0.5	0.5	0.5
Dark Current [pA]	1	100	2	1	8	20
Cut-off Frequency [GHz]	2/1.5	1.5	2.5	3	1.9	1.5
Terminal Capacitance [pF]	0.5/1.6	1.5	0.5	0.3	0.45	0.8
Active area (µm)	100/200	400	60	40	80	120
Voltage bias (V)	5	2	2	2	2	2

Table 7.3: Electrical and optical characteristics of PiN diodes

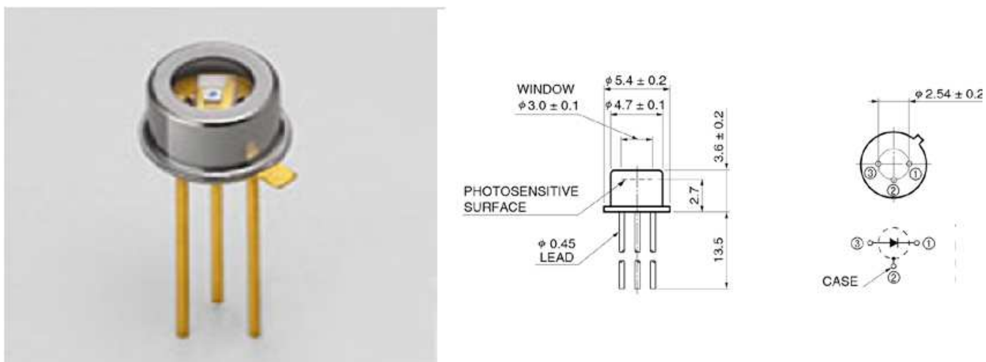


Figure 7.3: Dimensional outline for TO-18 packages (units are mm) that is common for all diodes tested in this thesis

tions Figure 7.4). This diode is a potential candidate for the high speed parallel optical transceiver. It has a flat response up to 2 GHz.

The electrical and optical characteristics of PiN diodes are shown in Table 7.3.

7.3 Performed Tests

Three major tests were performed to study radiation hardness characteristics of the PiN diodes. First, the total ionization damage (TID) test was performed at Solid State Gamma-Ray Irradiation Facility at Brookhaven National Laboratory (BNL) using gamma rays. PiN diodes were exposed to the total dose of 10 Mrad. The

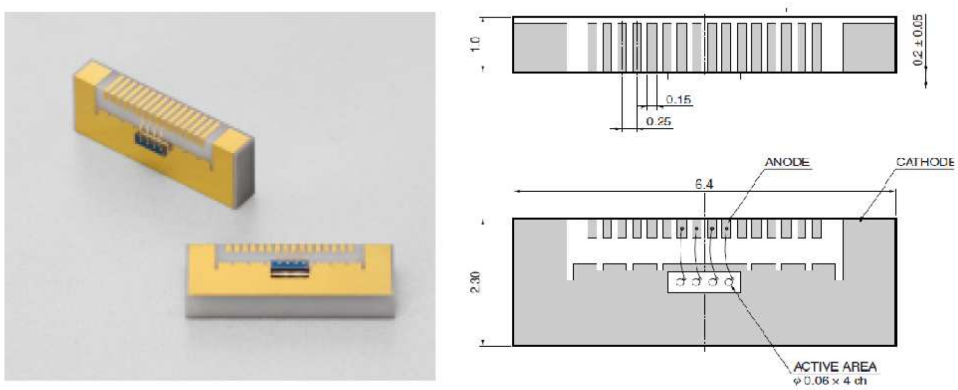


Figure 7.4: Dimensional outline for G8921 in 4 channel die (units are mm)

purpose of this test was to confirm that the performance of PiN diodes is not affected by the ionizing radiation which is one of the major irradiation sources at SLHC.

Two sets of tests have been done with proton beams to investigate the effect of the bulk damage in PiN diodes. The first set of tests was done at Indiana University Cyclotron Facility (IUCF) with 200 MeV protons. The tests were done in two phases with two weeks interval between them. This break allowed us to observe an annealing effect. Finally, the last set of irradiation tests was performed at CERN T7 test beam facility using 24 GeV proton beam.

The main purpose of the tests was to measure the PiN diode responsivity as a function of accumulated dose. In addition, it was investigated if the degradation of the responsivity depends on the active area and cut-off frequency of PiN diodes. For this purpose three PiN photodiodes from the G8522 family were used.

7.4 TID test at BNL

The gamma ray source at BNL is a cylindrical cobalt-60 source with maximum irradiation rate of $200 kRad/h$ positioned at 15 cm from the center of the source as shown in Figure 7.5. Four diodes chosen for the test were biased and measurement is done offline as responsivity v.s accumulated dose in three points. For the measurement

Figure 7.5: Gamma ray source and experimental setup

of responsivity a homogenous infrared (IR) source is designed which its current was controlled by feedback from photodiode installed at the close point to PiN diodes to keep the IR power constant. Figure 7.6 shows the schematics of the setup for optical calibration and responsivity measurement. Samples have been installed at the irradiation point at 15 cm from the source to get the maximal dose. The accumulated dose is known with 20% accuracy. Responsivity was measured offline at 0 Mrad, 5.6 Mrad and 9.6 Mrad. Figure 7.7 shows responsivity versus total dose. No change in responsivity has been observed up to 9.6 Mrad.

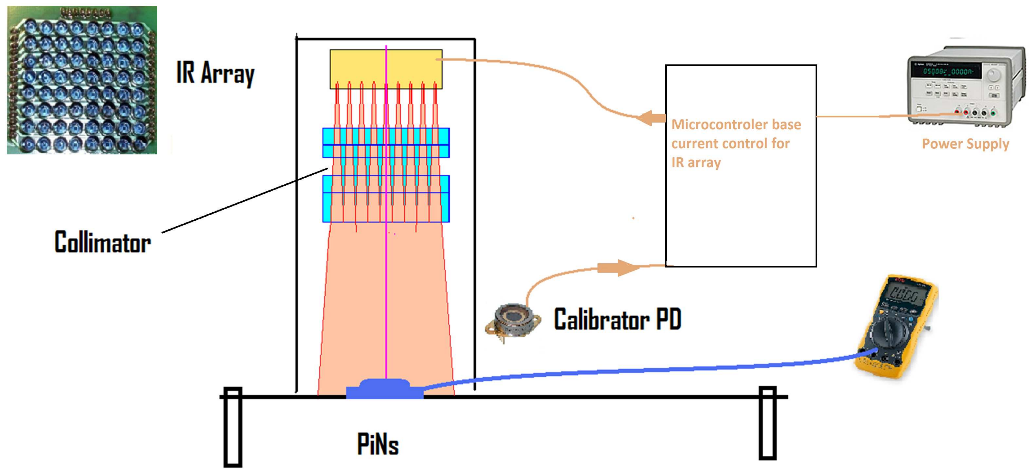


Figure 7.6: Simplified schematics of the optical setup for BNL tests.

7.5 Open-Air Optical Path Test Stand

For the second set of tests with protons irradiation, a new concept of a test stand was proposed for optical components like PiN diodes. It is named Open-air Optical Path (OAOP) and has several advantages over conventional optical fiber base test stands:

- It allows to avoid any parasitic effects and random mechanical attenuation from optical circuits and connectors. With this concept one can remove the effect of optical packages in final measurement and simplify the test setup.
- Ability to test a large number of PiN die samples in small area of test beam.
- The test stand design needs simpler electrical and DAQ circuitry and provides a simple monitoring and control of the optical power of the sources with high accuracy.
- Same apparatus can be used for accelerated life time test just by adding heat source to PiN diodes and isolating them from environment and IR source.

Figure 7.8 shows the general schematic of the OAOP and its IR source and monitoring photodiodes. The actual motherboard that shown in Figure 7.9 includes a high power

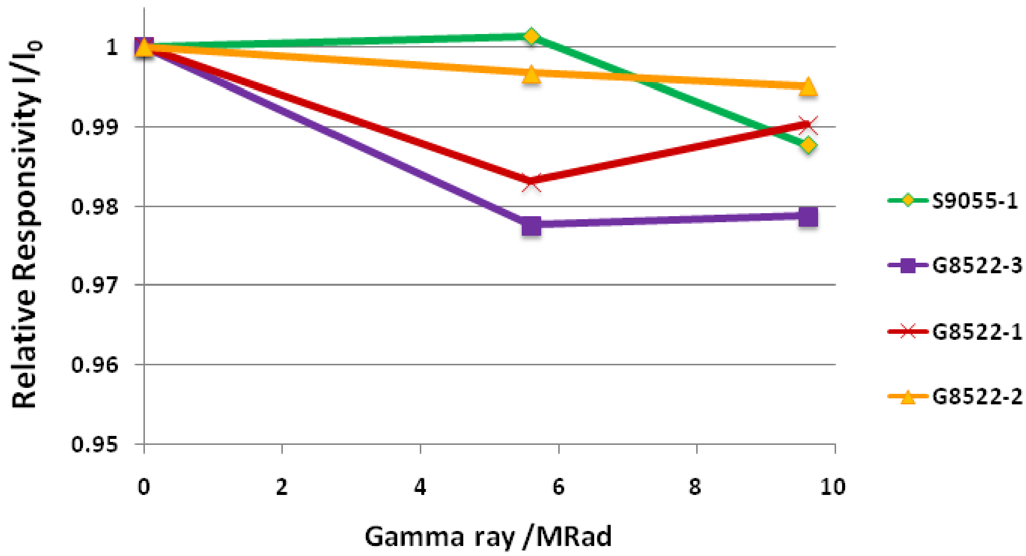


Figure 7.7: Responsivity vs. accumulated dose obtained at BNL tests.

IR power source with a wavelength of 850 nm focused precisely at the center of the daughter-board that carries the PiN diodes. The motherboard is positioned in such a way that the beam crosses the PiN diodes but does not touch the optical sources which are installed as a ring in top of board. The daughter-board is detachable from the motherboard to allow quick and easy measurement of multiple samples or exchange of samples. The IR source rings are supplied with a constant current to keep the optical power stable. Two photodiodes are installed away from beam but focused at center of the daughter-board monitor the optical power. Their output are measured during the test and are recorded to compensate for any optical power fluctuation later in analysis of data. A USB powered National Instruments data acquisition with 16 differential input used to measure the current of PiN diodes during irradiation. A program for recording of data is written in LabView environment. This program also monitors two photodiodes on board to compensate automatically for any optical power fluctuation during the experiment. The control panel of program shown at Figure 7.10. The OAOP setup at a test run is shown in Figure 7.11 at left and during experiment in front of the beam pipe at IUCF at right.

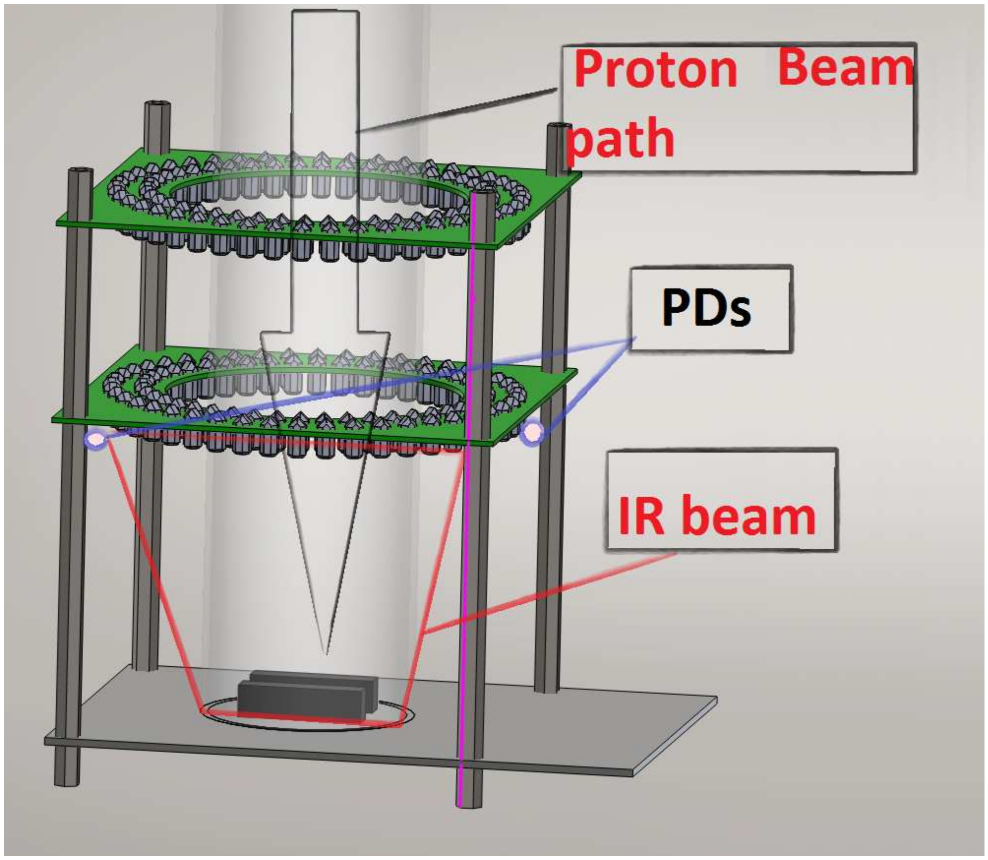


Figure 7.8: Schematic of IR source of OAOP test stand

7.6 Performed Test by Proton

7.6.1 Results of Irradiation Test at IUCF

The second set of tests was performed with the Open-air Optical Path test stand at Indiana University Cyclotron Facility (IUCF) using 200 MeV proton beam. The IUCF's beam is 5 cm in diameter and uniformity better than 30% variation over area. It has a maximum irradiation rate of 5 Mrad/h.

The test of selected PiN diodes has been performed in two separate runs. The accumulated dose in each run was $0.7 \times 10^{15} p/cm^2$, which corresponds to 40 MRad. Since the IR source was losing power during the run due to secondary radiation, its optical power has been monitored, and the results of the PiN diode responsivity

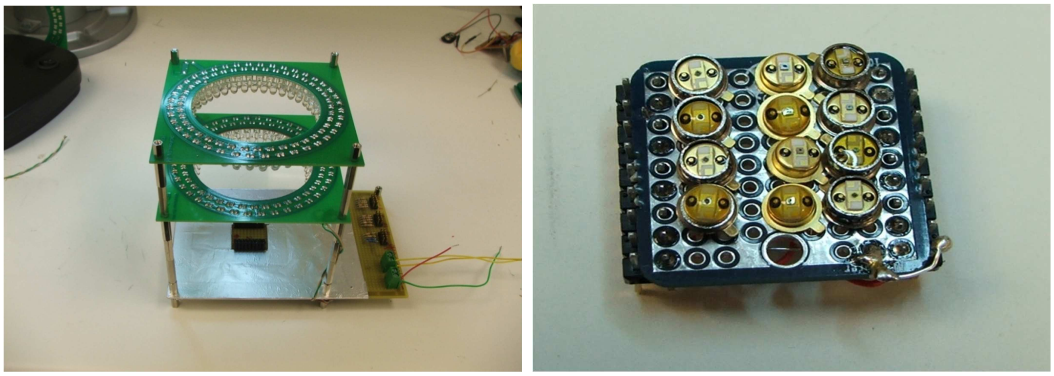


Figure 7.9: Open-air Optical Path motherboard (at left) and daughter-board with PiN diodes installed on (at right).

PiN	Total degradation due to 80Mrad [%]
S9055-01	12
S5973-01	33
G8522-01	34
G8522-02	29
G8522-03	55

Table 7.4: Total responsivity degradation after 80Mrad irradiation with 200 MeV proton for selected PiNs. Optical power fluctuation due IR source has been compensated

measurements have been corrected for this loss. Figure 7.12 shows the results of the first irradiation phase with normalized responsivity but without optical power variation compensation.

Figure 7.13 shows the results of the responsivity measurement for two types of PiN diodes (S8522-1 and 3) with and without compensation. The lines marked "ON" take the compensation into account. The small plot on top shows the normalized optical power of the source.

Table 7.4 shows total responsivity degradation after 80 Mrad of accumulated dose using 200 MeV protons. The results are corrected for the optical power fluctuation. The results for the G8522-0X diode family show no correlation between responsivity and size of active optical area.

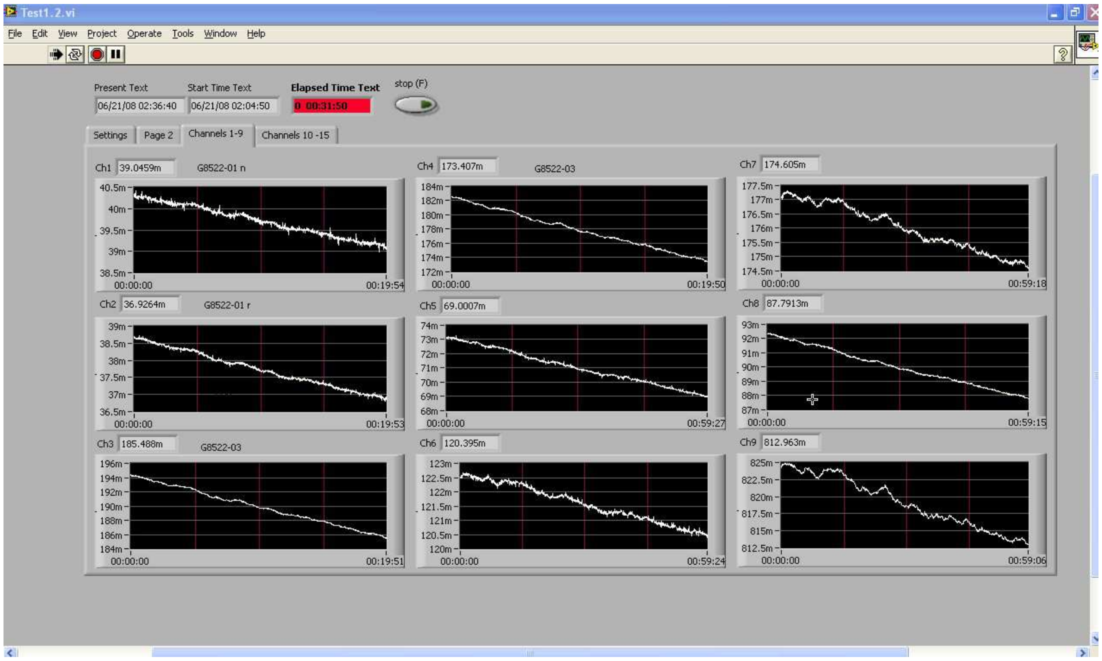


Figure 7.10: Panel of program written in LabView to record the output of PiN diodes

7.6.2 Tests at CERN T7 Facility

The last test was performed at CERN T7 facility with 24 GeV proton beam. The beam profile is approximately a square 2.2 cm by 2 cm wide and the flux is about $1.5 - 4.5 \times 10^9 p\, cm^{-2} s^{-1}$ [68].

Figure 7.14 shows the test stand designed to be used at T7. The stand has 32 optical channels, but can be expanded up to 256 channels. It has a compact, portable

Figure 7.11: Open-Air Optical Path Setup (at right) and position of test setup in front of beam at IUCF

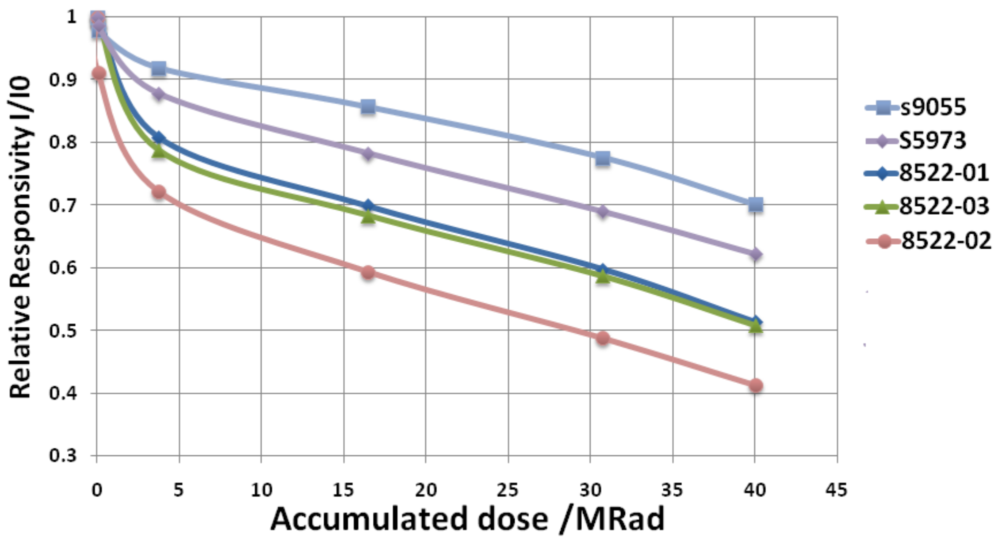


Figure 7.12: Responsivity vs. accumulated dose obtained in the first IUCF run (no optical variation compensation).

footprint. The stand provides a fine control of the optical power for each individual channel. It can be also modified into a bit error rate test stand. The test stand is controlled by a PC running LabView. At T7 facility, the proton beam is about 20 m away from the control room and samples have to be installed in a shuttle to reach the beam. Samples in a shuttle are connected via fiber ribbons to the VCSEL modules at the test stand. Every 30 seconds, the test stand reads out the optical power sent to the PiN diodes and the currents of each individual channel of PiN diodes as a measure of responsivity. Figures 8-10 show degradation of the PiN diode responsivity due to irradiation with 24 GeV protons. This degradation is in a good agreement with results obtained at 200 MeV at IUCF. Figure 7.15 shows that the S5793-01 diodes lost about 30% of their initial responsivity, to be compared with 33% at IUCF. Figure 7.16 shows that with total fluence of 2.6×1015 p/cm2 the S9055-01 diodes lost about 10% of their initial responsivity which is in agreement with 12% at IUCF. Finally, Figure 7.17 shows responsivity plots for the GaAs PiN array with an average of 50% loss of responsivity. The responsivity shown in these plots is calculated taking into

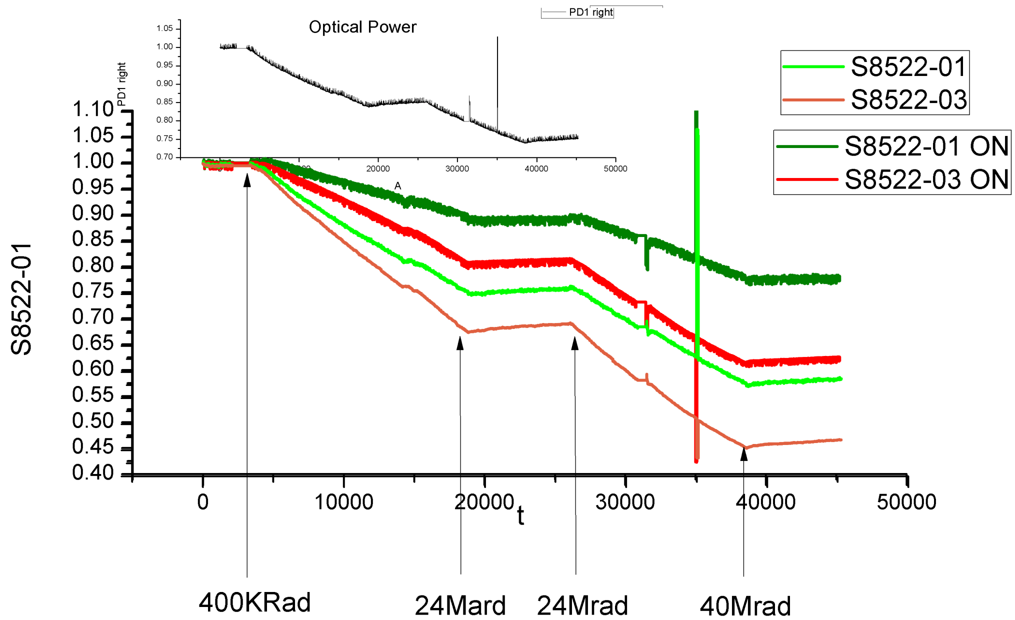


Figure 7.13: Responsivity vs. accumulated dose obtained at IUCF at second run 40 to 80Mrad with normalized responsivity (optical power degradation compensation)

account the optical power fluctuations of the optical sources.

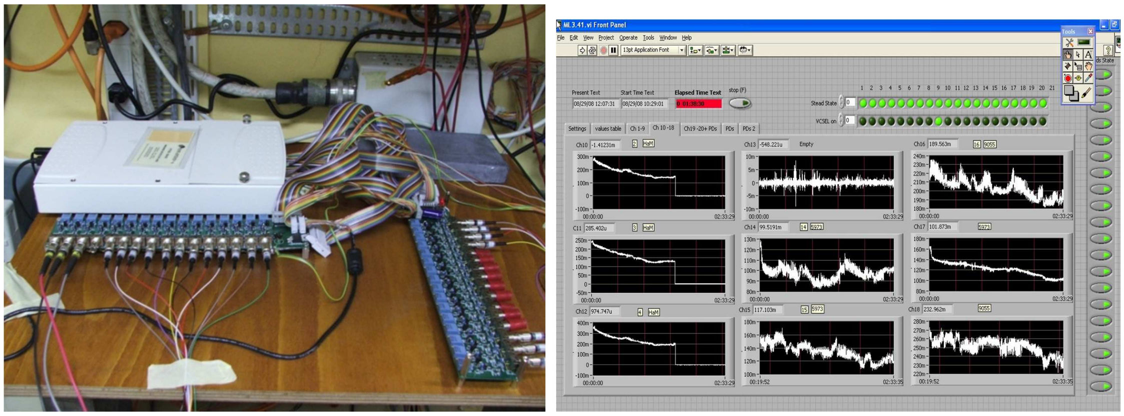


Figure 7.14: Test stand designed for includes 32 controlled VCSEL source (at left) control panel of program that control VCSELs and records the PiN currents during experiment (at right)

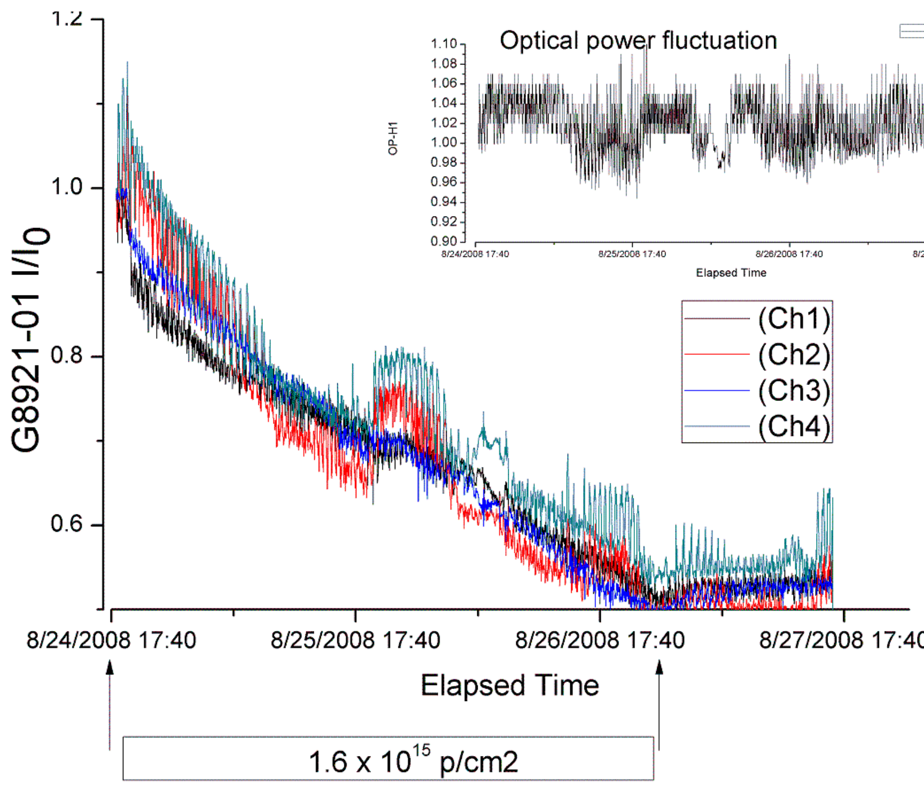


Figure 7.17: Responsivity vs. accumulated dose obtained at CERN T7 (diode G8921-01, 4 channels)

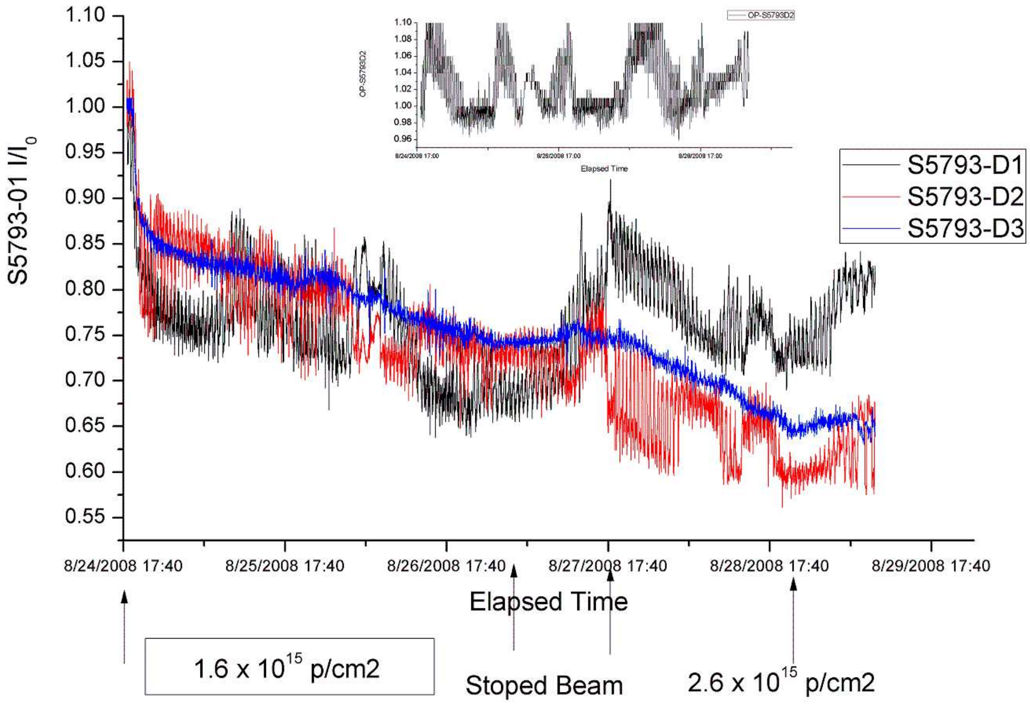


Figure 7.15: Responsivity vs. accumulated dose obtained at CERN T7 (three diodes S5793-01).

7.7 Summary

Three different tests of the radiation hardness of selected PiN diodes were done. For this purpose. the test stands were developed and built. The results demonstrate that radiation hardness does not depend on the active area of PiN diodes from the same family. Results obtained at two different proton beam energies 200 MeV and 24 GeV are found to be in a good agreement with each other. Based on results from IUCF and CERN irradiation runs, one can identify the following PiN diode candidates:

- GaAs array G8921-01 with total responsivity degradation less than 50% (initial photosensitivity is 0.5 A/W). Its responsivity after irradiation is better than for S9055-01. It is available in different configurations (4 to 16 channels per array).

- Si PiN 9055-01 with total degradation less than 10% (initial photosensitivity is 0.25 A/W).

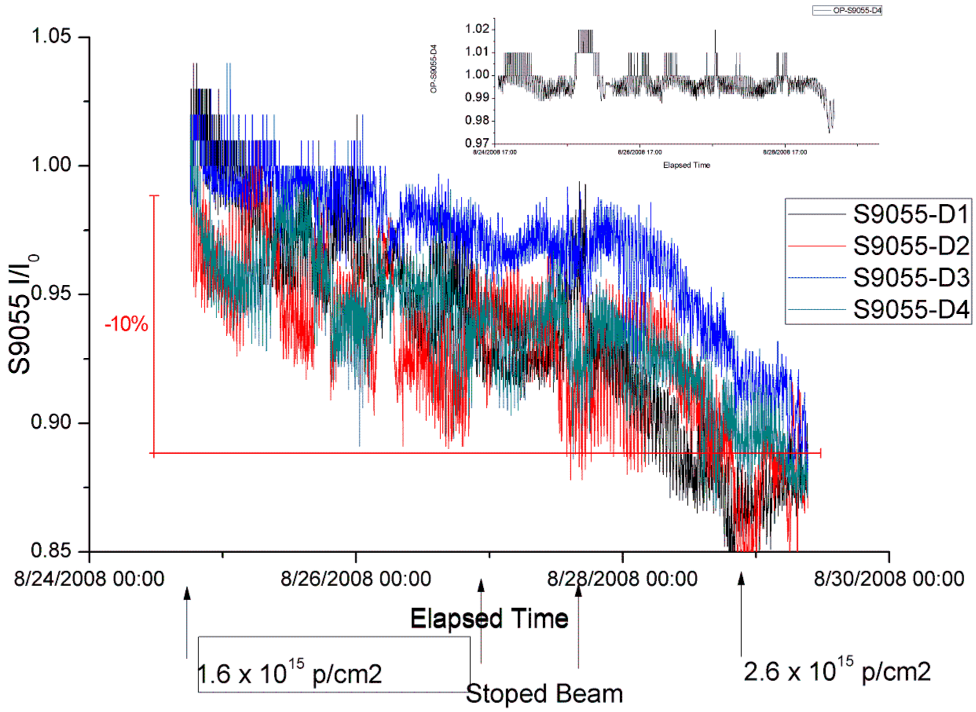


Figure 7.16: Responsivity vs. accumulated dose obtained at CERN T7 (four diodes S9055-01).

These candidates can be used for other applications in high radiation areas at the SLHC and HL-LHC.

Chapter 8

Conclusion

The LHC is known to be a top pair factory that produces $t\bar{t}$ with an order of magnitude more top quarks than at the Tevatron where the top quark was discovered. LHC will provide us precision measurements properties and improving of understanding of the heaviest known particle. Studying top physics is powerful probe for physics beyond the Standard Model related to EWSB. The observables (cross-section, polarization) are sensitive to different kinds of new physics. Also $t\bar{t}$ is important dominant source of background in searches for the Higgs boson and new physics beyond the SM.

This thesis describes two measurement of the $t\bar{t}$ production cross-section in lepton + jets channel, one inclusive $\sigma_{t\bar{t}}$ and the other $\sigma_{t\bar{t}j}$ in association with at least one jet. The measurements employ multivariate technique with a binned maximum likelihood template fit. Several kinematic variables of the reconstructed collision events are selected to distinguish $t\bar{t}$ signal events from the various background sources.

The first measurement is done by using approximately 35.3pb^{-1} of ATLAS data collected in 2010 and selecting events with one isolated high-p_T electron or muon, large missing transverse energy and two jet bins of 3 or at least 4 high-p_T jets. The measured $t\bar{t}$ cross section assuming the top quark mass of 172.5 GeV yields :

$$\sigma_{t\bar{t}} = 173 \pm 17(\text{stat.})^{+20}_{-17}(\text{syst.}) \pm 6(\text{lumi.})\,\text{pb} \tag{8.1}$$

The measured value is in a good agreement with the Standard Model prediction

The second measurement is performed on ATLAS for the first time with data 2011 corresponding to an integrated luminosity of 4.7 fb^{-1}. The same analysis technique and event selection employed as the first measurement with slight changes but in 4th and 5th jet bins. The cross-section for $\sigma_{t\bar{t}}$ with at least one additional jet is found to be :

$$\sigma_{t\bar{t}j} = 102 \pm 2(\text{stat.}) \pm 6(\text{ISR/FSR})^{+22}_{-25}(\text{syst.})\,\text{pb} \tag{8.2}$$

Also a study on PiN diodes radiation hardness is performed to investigate the proper optical components for ATLAS Pixel upgrade. Along side developing three the test-stands, several tests are performed with different beam and different energies. The results have been presented and two promising candidates are proposed.

Appendix A

Calculation of σ_{ttj} in method 1

Solving Eq. 6.2, we have

$$k_1 = \frac{N_4^0 N_5^{t\bar{t}} - N_5^0 N_4^{t\bar{t}}}{N_4^0 N_5^1 - N_4^1 N_5^0} \tag{A.1}$$

Substituting

$$N_4^0 = \mathcal{L}\sigma_{t\bar{t}} f_0 \varepsilon_{04}, \quad N_5^0 = \mathcal{L}\sigma_{t\bar{t}} f_0 \varepsilon_{05}, \quad N_4^1 = \mathcal{L}\sigma_{t\bar{t}} f_1 \varepsilon_{14}, \quad N_5^1 = \mathcal{L}\sigma_{t\bar{t}} f_1 \varepsilon_{15} \tag{A.2}$$

we get

$$k_1 = \frac{(f_0\varepsilon_{04} N_5^{t\bar{t}} - f_0\varepsilon_{05} N_4^{t\bar{t}})(\mathcal{L}\sigma_{t\bar{t}})}{(f_0\varepsilon_{04} f_1\varepsilon_{15} - f_1\varepsilon_{14} f_0\varepsilon_{05})(\mathcal{L}\sigma_{t\bar{t}})^2} = \frac{f_0(\varepsilon_{04} N_5^{t\bar{t}} - \varepsilon_{05} N_4^{t\bar{t}})}{f_0 f_1(\varepsilon_{04}\varepsilon_{15} - \varepsilon_{14}\varepsilon_{05})\mathcal{L}\sigma_{t\bar{t}}} \tag{A.3}$$

$$k_1 = \frac{\varepsilon_{04} N_5^{t\bar{t}} - \varepsilon_{05} N_4^{t\bar{t}}}{f_1(\varepsilon_{04}\varepsilon_{15} - \varepsilon_{14}\varepsilon_{05})\mathcal{L}\sigma_{t\bar{t}}}$$

so that

$$\sigma_{ttj} = k_1 \sigma_{t\bar{t}} f_1 = \frac{\varepsilon_{04} N_5^{t\bar{t}} - \varepsilon_{05} N_4^{t\bar{t}}}{(\varepsilon_{04}\varepsilon_{15} - \varepsilon_{14}\varepsilon_{05})\mathcal{L}} \tag{A.4}$$

does not depend on either f_1 or $\sigma_{t\bar{t}}$.

Bibliography

[1] P. Jenni, M. Nessi, M. Nordberg, and K. Smith. Atlas high-level trigger, data-acquisition and controls: Technical design report. *Technical Design Report ATLAS. CERN, Geneva*, 42, 2003.

[2] K Becker, A Cortes Gonzalez, , et al. Mis-identified lepton backgrounds in top quark pair production studies for eps 2011 analyses. (ATL-COM-PHYS-2011-768), Jun 2011.

[3] F. Abe, H. Akimoto, Akopian, et al. Observation of top quark production in p [over¯] p collisions with the collider detector at fermilab. *Physical review letters*, 74(14):2626–2631, 1995.

[4] S. L. Glashow. Partial Symmetries of Weak Interactions. *Nucl. Phys.*, 22:579–588, 1961.

[5] Abdus Salam. Weak and Electromagnetic Interactions. Originally printed in Svartholm: Elementary Particle Theory, Proceedings Of The Nobel Symposium Held 1968 At Lerum, Sweden, Stockholm, 1968.

[6] Steven Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.

[7] S Allwood-Spires, M Barisonzi, et al. Monte carlo samples used for top physics. (ATL-PHYS-INT-2010-132), Dec 2010.

[8] K. Nakamura *et al.* (Particle Data Group). The Review of Particle Physics. J. Phys. G 37, 075021 (2010).

[9] G. Aad, M. Ackers, FA Alberti, M. Aleppo, G. Alimonti, J. Alonso, EC Anderssen, A. Andreani, A. Andreazza, JF Arguin, et al. Atlas pixel detector electronics and sensors. *Journal of Instrumentation*, 3(07):P07007, 2008.

[10] K. Aamodt, A.A. Quintana, R. Achenbach, S. Acounis, D. Adamová, C. Adler, M. Aggarwal, F. Agnese, G.A. Rinella, Z. Ahammed, et al. The alice experiment at the cern lhc. *Journal of Instrumentation*, 3:S08002, 2008.

[11] S. Chatrchyan, G. Hmayakyan, V. Khachatryan, AM Sirunyan, W. Adam, T. Bauer, T. Bergauer, H. Bergauer, M. Dragicevic, J. Erö, et al. The cms experiment at the cern lhc. *Journal of Instrumentation*, 3:S08004, 2008.

[12] L.H.C. Collaboration, R. Aaij, B. Adeva, M. Adinolfi, C. Adrover, A. Affolder, Z. Ajaltouni, J. Albrecht, F. Alessio, M. Alexander, et al. Search for the rare decays and b0–>[mu]+[mu]-. *Physics Letters B*, 2011.

[13] L. Evans and P. Bryant. *JINST*, 3:S08001, 2008.

[14] WW Armstrong, W. Burris, DM Gingrich, P. Green, LG Greeniaus, JC Hewlett, L. Holm, JW McDonald, S. Mullin, WC Olsen, et al. Atlas: technical proposal for a general-purpose pp experiment at the large hadron collider at cern. 1994.

[15] ATLAS collaboration et al. Atlas magnet system technical design report. *CERN/LHCC*, pages 97–18, 1997.

[16] The ATLAS Collaboration. Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics, pp. 870-883. pages 870–883, 2009.

[17] ATLAS HEC Collaboration. The atlas liquid argon hadronic end-cap calorimeter: contruction and selected beam test results. 2004. arXiv:physics/04070226v1.

[18] The ATLAS Collaboration, G. Aad et al. Drift time measurement in the atlas liquid argon electromagnetic calorimeter using cosmic muons. *Eur. Phys. J.*, C70:755, 2010.

[19] The ATLAS Collaboration, G. Aad et al. Readiness of the atlas tile calorimeter for lhc collisions. *Eur. Phys. J.*, C70:1193, 2010.

[20] TileCal collaboration. Cesium monitoring system for atlas tile hadron calorimeter. *NIM A*, 494:381–384, 2002.

[21] ATLAS Tile Calorimeter Collaboration. Atlas tile calorimer technical design report. Technical report, CERN, 1996. CERN/LHCC 96/42.

[22] Giulio Usai. *Identification and triggering of soft muons in the Atlas detector.* PhD thesis, Universit?degli studi di Pisa, Dipartimento di fisica, June 2004.

[23] Tracking results and comparison to monte carlo simulation at $\sqrt{s}$ = 900 gev. (ATLAS-CONF-2010-011), Jul 2010.

[24] Electron and photon reconstruction and identification in atlas: expected performance at high energy and results at 900 gev. (ATLAS-CONF-2010-005), Jun 2010.

[25] Expected electron performance in the atlas experiment. (ATL-PHYS-PUB-2011-006), Apr 2011.

[26] The e/gamma combined performance group. Electron efficiency measurements using atlas 2010 data at sqrt(s) = 7 tev and https://twiki.cern.ch/twiki/bin/view/atlasprotected/efficiencymeasurements. ATL-COM-PHYS-2011-322.

[27] M. Agustoni *et al.* Electron energy scale in-situ calibration and performance. ATL-COM-PHYS-2011-263 (2011).

[28] S. Hassani, L. Chevalier, E. Lancon, J.F. Laporte, R. Nicolaidou, and A. Ouraou. A muon identification and combined reconstruction procedure for the atlas detector at the lhc using the (muonboy, staco, mutag) reconstruction packages. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 572(1):77–79, 2007.

[29] ATLAS Muon Collaboration et al. Atlas muon spectrometer technical design report. *CERN/LHCC*, 97:22, 1997.

[30] ATLAS Top Group. M2: Jets, missing transverse energy and taus for top physics analyses in release 16 with the 2010 dataset. *ATL-COM-PHYS-2011-132*, 2011.

[31] M. Cacciari, G.P. Salam and G. Soyez. The anti-k(t) jet clustering algorithm. *JHEP*, 04:63, 2008.

[32] G. Aad, B. Abbott, J. Abdallah, AA Abdelalim, A. Abdesselam, O. Abdinov, B. Abi, M. Abolins, H. Abramowicz, H. Abreu, et al. Performance of missing transverse momentum reconstruction in proton-proton collisions at with atlas. *The European Physical Journal C-Particles and Fields*, 72(1):1–35, 2012.

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

[34] B Abi, J Adelman, J Almond, et al. Mis-identified lepton backgrounds to top quark pair production. (ATL-PHYS-INT-2010-139), Dec 2010.

[35] A. Höcker et al. TMVA: Toolkit for multivariate data analysis. *PoS*, ACAT:040, 2007.

[36] Georges Aad et al. Measurement of the top quark pair production cross-section with ATLAS in the single lepton channel. *Phys.Lett.*, B711:244–263, 2012.

[37] Stefano Frixione and Bryan R. Webber. Matching NLO QCD computations and parton shower simulations. *JHEP*, 0206:029, 2002.

[38] G. Corcella, I.G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, et al. HERWIG 6: An Event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *JHEP*, 0101:010, 2001.

[39] J. M. Butterworth, Jeffrey R. Forshaw, and M. H. Seymour. Multiparton interactions in photoproduction at HERA. *Z. Phys.*, C72:637–646, 1996.

[40] L. Mijovic et al.

[41] E. Shabalina, J. Sjölin, et al. Estimation of the fake lepton background to ttbar events with l = 35 ipb-1. Note M4, In Preparation, 2010.

[42] ATLAS Collaboration. Updated luminosity determination in pp collisions at $\sqrt{s} = 7$ tev using the atlas detector. *ATLAS-CONF-2011-011*, 2011.

[43] The ATLAS Collaboration. Luminosity determination in pp collisions at $\sqrt{s} = 7$ tev using the atlas detector in 2011. 2011.

[44] The CDF Collaboration. Measurement of the t-tbar + jet cross section with 4.1 fb^{-1}. CDF Note 9850, 2009. http://www-cdf.fnal.gov/physics/new/top/2009/xsection/ttj_4.1invfb/.

[45] Twiki page: Production Group MC11c campaign page. `https://twiki.cern.ch/twiki/bin/viewauth/Atlas/MC11c`.

[46] T. Sjostrand and S. Mrenna and P.Z. Skands. Pythia 6.4 physics and manual. *JHEP*, 05:026, 2006.

[47] The ATLAS Collaboration. Atlas tunes of pythia6 and pythia8 for mc11. ATL-PHYS-PUB-2011-009, 2011.

[48] S. Frixione and B.R. Webber. Matching nlo qcd computations and parton shower simulations. *JHEP*, 0206:029, 2002.

[49] H.-L. Lai et al. New parton distributions for collider physics. *Phys.Rev.*, D82:074024, 2010.

[50] G. Corcella et al. Herwig 6: An event generator for hadron emission reactions with interfering gluons (including supersymmetric processes). *JHEP*, 0101:010, 2001.

[51] The ATLAS Collaboration. New atlas event generator tunes to 2010 data. ATL-PHYS-PUB-2011-008, 2010.

[52] M. Aliev et al. HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR. *Comput. Phys. Commun.*, 182:1034–1046, 2011.

[53] S. Frixione, P. Nason and C. Oleari. Matching nlo qcd computations with parton shower simulations: the powheg method. *JHEP*, 0711:070, 2007.

[54] M.L. Mangano et al. Alpgen, a generator for hard multiparton processes in hadronic collisions. *JHEP*, 0307:001, 2003.

[55] T. Gleisberg et al. Event generation with sherpa 1.1. *JHEP*, 0902:007, 2009.

[56] Twiki page: TopCommonObjects. `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopCommonObjects`.

[57] Twiki page: TopSystematicUncertainties. `https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TopSystematicUncertainties`.

[58] The ATLAS Collaboration. Top quark pair production cross-section measurement in atlas in the single lepton+jets channel without b-tagging. 2011.

[59] Stefano Frixione, Paolo Nason, and Carlo Oleari. Matching NLO QCD computations with Parton Shower simulations: the POWHEG method. *JHEP*, 11:070, 2007.

[60] Paolo Nason and Bryan Webber. Next-to-Leading-Order Event Generators. 2012.

[61] B Abi, M Saleem, A Khanov, and F Rizatdinova. Measurement of the cross section in ttbar + jets using kinematics fit method with the atlas detector. (ATL-COM-PHYS-2011-1718, CERN,Geneva), Dec 2011.

[62] K. Melnikov. private communications.

[63] F. Gianotti, M.L. Mangano, T. Virdee, S. Abdullin, G. Azuelos, A. Ball, D. Barberis, A. Belyaev, P. Bloch, M. Bosman, et al. Physics potential and experimental challenges of the lhc luminosity upgrade. *The European Physical Journal C-Particles and Fields*, 39(3):293–333, 2005.

[64] S. Baranov, M. Bosman, I. Dawson, V. Hedberg, A. Nisati, and M. Shupe. Estimation of radiation background, impact on detectors, activation and shielding optimization in atlas. *ATLAS Note,ATL-GEN-2005-001*, 2005.

[65] A. Van Ginneken. Non-ionizing energy deposition in silicon for radiation damage studies. *FERMILAB-FN-0522*, page 8pp., Oct 1989.

[66] I. Gregor. *Opticallinks for the ATLASpixeldetector*. PhD thesis, Ph. D. Thesis University of Wuppertal, 2001.

[67] KK Gan, K. Gill, J. Troska, F. Vasey, T. Huffman, C. Issever, T. Weidberg, J. Ye, P. Skubic, R. Boyd, et al. Joint atlas-cms working group on optoelectronics for slhc.

[68] M. Glaser, L. Durieu, F. Lemeilleur, M. Tavlet, C. Leroy, and P. Roy. New irradiation zones at the cern-ps. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 426(1):72–77, 1999.