Measurement of Top Pair Production Cross-Section in l + jets Channels in ATLAS at 7 TeV

Jie Yu

THÈSE

présentée par

Jie YU

pour obtenir le grade de

Docteur ès Sciences de L'Université Nanjing Chine et

l'Université Paris Sud 11 Orsay (Particles, Noyaux, Cosmos (ED 517))

Measure de la section efficace de production de paires de quarks, top-antitop à l'aide des premières donnees du détecteur ATLAS auprès du LHC

Sous la direction de:

Pr. Shenjian Chen

Dr. Bruno Mansoulié

Soutenue le 15 mai 2011 devant le jury composé de:

Pr. Shan Jin	Président du Jury
Pr. Xueyao Zhang	
Pr. Liang Han	
Pr. Etienne Augé	
Dr. Didier Vilanova	
Dr. Jérôme Schwindling	

Rapporteur externe: Dr. Arnaud Lucotte

Thèse préparée au Service de Physiques des Particules du CEA de Saclay

et département de physique de l'Université Nanjing

THESIS

presented by

Jie YU

to obtain the degree of

Doctor of Philosophy of Nanjing University and

University of Paris Sud 11 Orsay (Particle, Noyaux, Cosmos (ED 517))

Measurement of Top Pair Production in l + jets Channels in ATLAS at $\sqrt{s} = 7$ TeV

Under the direction of:

Pr. Shenjian Chen

Dr. Bruno Mansoulié

Defense on May 15, 2011. Committee members:

Pr. Shan Jin	Chair person
Pr. Xueyao Zhang	
Pr. Liang Han	
Pr. Etienne Augé	
Dr. Didier Vilanova	
Dr. Jérôme Schwindling	

With an external report by: Dr. Arnaud Lucotte

Thesis prepared at IRFU / SPP / CEA-Saclay and

the department of physics in Nanjing University

Abstract

Top quark was discovered by CDF and DØ collaborations in 1995. Since then, the measurement on the top production is a very interesting topic in the realm of high energy physics. This thesis describes a measurement of the $t\bar{t}$ production cross section at $\sqrt{s} = 7\,\text{TeV}$ using the first LHC data. This measurement depends on the cut and count method. The requirements of one good lepton, large $E_{\text{T}}^{\text{miss}}$, at least four jets are applied to suppress the backgrounds. It also requires in particular the knowledge of the W+Jets background, which will be measured using two methods, and a measurement of the b-jet tagging efficiency, which is measured using $t\bar{t}$ events. The QCD background is measured by applying a data driven method. Finally the systematics are estimated by considering the methods used in the measurement. The measurement gives a result of $t\bar{t}$ cross section

$$\sigma_{t\bar{t}} = 148.1 \pm 34.9\ pb$$

which is consistent with the results from different top pair decay channels and agrees with the theoretical prediction in the Standard Model.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

The author wishes to express his sincere appreciation to University of Nanjing, where he has the opportunity to work on the subject in High Energy Physics Group. Professor Chen Shenjian is really a kind and nice supervisor who gives him the first guide on HEP.

The author wishes to show his grateful thanks to the Atlas Group of CEA-Saclay, where he has spent half of the time during the three years and has done most of the analysis work. Bruno Mansoulié has been a good ATLAS group leader who expressed strong support for young students. Without him, the author cannot have so great experience of education in France. The author also would like very much to thank Jérôme Schwindling, who has been helping and guiding the author all through the journey of his analysis. Under his careful direction, the author has learned a lot of things on HEP, especially on physics analysis in hadron-hadron collisions. The author would like to express his same thankfulness to the University Paris-sud 11, which has been supporting his study in France with both finance and guidance.

The author wishes to thank also all the reporters, who have carefully read the thesis and given many helpful comments and suggestion. Dr. Arnaud Lucotte and Prof. Xueyao Zhang are the reporters who have written the feedback in Fench or in English, while Prof. Liang Han and Prof. Ming Qi have made it in Chinese. Besides, Dr. Traudl Hansl-Kozanecki has been reading it and given many comments and much suggestion without being officially a reporter. The author appreciates the effort very much.

After all, the author would like to thank his parents and all the relatives of the family who have been so supportive and show their sincere wish since the very beginning of the author's research career.

DEDICATION

To the universe, which is infinite and unknown

To the sun, which is warm and lightful

To the earth, which has oceans, fields and mountains

To the animals, who never lie

To the plants, who insist to stay

To the people, who survive

To the ground, where we can put our legs on

To everything that I know

To everyone who knows me

To anything that I don't know

To anyone who doesn't know me

Chapter 1

INTRODUCTION

This thesis presents the measurement of the $t\bar{t}$ cross section in the semi-leptonic decay channels, which means $t\bar{t} \to W^+\bar{b}W^-b \to l\nu bqqb$. It is using a cut and counting method expressed in the Formula 1.1:

$$\sigma_{t\bar{t}} = \frac{N_{data} - N_{bkg}}{L \cdot A \cdot \epsilon} = \frac{N_{sig}}{L \cdot A \cdot \epsilon} \tag{1.1}$$

The measurement needs to fulfill every member in the formula.

In Chapter 2, it first gives a short introduction to Standard Model (SM) of High Energy Physics. In the second part of this chapter, it revises the physics related to the top quark. The $t\bar{t}$ production cross section under different renormalization orders is calculated and used to check the validation of the theory.

In Chapter 3, the ATLAS detector is briefly introduced. The electromagnetic calorimeter and the muon spectrometer are emphasized, where the most important objects (electrons, muons and jets) are reconstructed. The reconstruction of $E_{\mathrm{T}}^{\mathrm{miss}}$ is also shown in this chapter after the other objects are built. The objects studied in this chapter can be used for the selection cuts. The luminosity measurement, which is used in Formula 1.1, is also described.

In Chapter 4, a series of selection cuts are applied to gain a high S/B ratio and at the same time to keep sufficient signal events. The selection efficiency (ϵ) is calculated for signal $t\bar{t}$ non-hadronic sample with MC events. Due to the imperfect of understanding the theory and detectors, the trigger, reconstruction and identification efficiencies of leptons and b-tagging efficiency of jets are different between MC and data. Therefore scale factors are introduced to correct the effects. The measurement of QCD background using a "Matrix Method" is also introduced in this chapter. QCD is one of the dominating backgrounds especially in the electron channel.

Chapter 5 presents two methods to measure the W+Jets background, the largest background in both electron and muon channels. The Berends-Giele scaling method is simple and mostly used for the first data. Charge asymmetry method is established for later stage, but limited by the lack of data. The two methods are applied and compared.

Chapter 6 introduces a method called "tag counting" using the number of b-tagged jets to extract both the efficiency of b-tagging and the number of signal events (N_{sig}). The method avoids adopting a scale factor to the ϵ_b, but obtains N_{sig} by using the information of b-tagging. After the N_{sig} is measured, a direct calculation gives the results of $t\bar{t}$ cross section. Finally the systematic uncertainties of the cross section are also considered in this chapter.

Finally, in Chapter 7, the expectation of the $t\bar{t}$ cross section measurement with ATLAS data in LHC is discussed. The results can be used to check the theoretical expectation of top pair cross section in Standard Model or beyond. Some improvements are expected in the future with increasing collision data in Atlas at 7 TeV.

Chapter 2

THE STANDARD MODEL AND TOP QUARK PHYSICS

The Standard Model (SM) in particle physics is a theory including electromagnetic, weak and strong interactions which rule the realm of the fundamental subatomic particles. The SM is established mainly in the early and middle of 20th century, while the electroweak theory is built in 1970s. The gauge transformation invariance of the fields is developed, which enhances the importance of the concept of symmetry. The spontenous symmetry breaking is introduced to generate the mass. The SM also has shortcomings, for example: it could not include the other fundamental force, the gravity. Anyway, the SM is so important that it plays an essential role in particle physics. It is a basis for the models beyond, like extra dimensions, supersymmetry, string theory and so on.

Top quark is one of the least well measured particles in SM. It is interesting to study the properties of the top quark because it gives insights not only to SM but also to the area beyond SM. An accurate measurement of top quark mass is a constrain to the precision of electroweak analysis. The single top processes could be used to directly measure the CKM matrix element V_{tb}. Because of its large mass, the Higgs boson couples most strongly to the top quark. Top decay events are important backgrounds to new physics, especially with $t\bar{t}$ production, which is also a path to investigate strong interaction.

In this chapter, the SM which describes the fundamental particles and their interactions is discussed. The theoretical calculation of the $t\bar{t}$ cross section is also described. Since now the Large Hadron Collider (LHC) is producing data at 7 TeV, it is a real top factory. An exciting journey to understand the top quark, the SM and even the theories beyond is now in front of us.

2.1 The Standard Model

Beginning with treating the particles also with a wave function, physicists built the monument of Quantum Mechanics in the first three decades of the 20th century. Afterwards, particle physics has experienced a glorious journey with not only theoretical progressing but also experimental proving. The so called "Standard Model" is almost established during the past 100 years.

At present, people believe that there are two classes of elementary particles, leptons and quarks, which combine to be matters we see in the normal world. They are connected by four fundamental interactions (or forces), gravity, electromagnetism, weak and strong nuclear force. The theories which describe these particles and interactions (except gravity) are now unified as "Standard Model".

It is well believed that symmetry plays a very important role in the history of physics. In the language of field theory, the Noether's theorem claims that symmetries imply conservation laws. In Quantum Electrodynamics (QED) the existence of the gauge field (the photon) follow from a principle of invariance under local gauge transformations of the U(1) group. The gauge principle which was introduced by Salam and Ward [4] for QED in 1961 is expected to be generalized to all the other interactions.

2.1.1 The Elementary Particles

All the quarks and leptons are fermions and have spin equal to $\frac{1}{2}$. Quarks and leptons are grouped into two families, who separately have three generations. For the quark family:

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

The name of the 6 flavors of particles are commonly up, down, charm, strange, top and bottom. The upper row in doublets has $+\frac{2}{3}$ electric charge and the lower row $-\frac{1}{3}$. The quarks also carry a baryon number $\frac{1}{3}$. The quarks encounter with both electroweak interaction and strong interaction.

Gauge Boson	charge	spin	Mediate	Mass (GeV)
γ	0	1	EM	0
$W^{\pm}$	± 1	1	Weak	80.398$\pm$0.025
Z^0	0	1	Weak	91.1876$\pm$0.0021
gluons	0	1	Strong	0
graviton	0	2	Gravity	0

Table 2.1: Summary of the gauge bosons mediating fundamental interactions.

Just like the quarks, the 6 flavors of the leptons also form a family with three generations:

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

The leptons are carrying -1 electric charge, while the neutrinos are chargeless. Each generation of the leptons have its own lepton number. Leptons do not have strong interaction since they have no colour charge.

In addition to the particles expressed above, every member in the family has its own partner, the anti-particle. The concept of anti-particle is introduced by Dirac in 1929. The anti-particles have the exact same properties as the corresponding particles except that they have opposite sign of quantum numbers like charge, lepton number, etc.

The interactions, in the fundamental gauge theory, are carried by the intermediate gauge bosons, photons, $W^{\pm}$, Z^0 and gluons.

The properties of these particles are shown in the Table 2.1 and Table 2.2.

2.1.2 The Electroweak Theory

In Standard Model, the electromagnetic and weak interactions are unified as "electroweak interaction" with the gauge symmetry $SU(2)_L \otimes U(1)_Y$.

The local gauge invariance is first introduced to quantum field theory which leads to the electromagnetic interaction. Afterwards, Yang and Mills developed the theory with the

Fermions	charge	spin	Mass
u	$+\frac{2}{3}$	$\frac{1}{2}$	$1.5 \sim 3.3$ MeV
d	$-\frac{1}{3}$	$\frac{1}{2}$	$3.5 \sim 6.0$ MeV
c	$+\frac{2}{3}$	$\frac{1}{2}$	$1.27^{+0.07}_{-0.11}$ GeV
s	$-\frac{1}{3}$	$\frac{1}{2}$	104^{+26}_{-34} MeV
t	$+\frac{2}{3}$	$\frac{1}{2}$	171.2 ± 2.1 GeV
b	$-\frac{1}{3}$	$\frac{1}{2}$	$4.20^{+0.17}_{-0.07}$ GeV
e	-1	$\frac{1}{2}$	0.510998910 ± 1.3E-8 MeV
ν_e	0	$\frac{1}{2}$	< 2 eV
μ	-1	$\frac{1}{2}$	105.658367 ± 4E-6 MeV
ν_μ	0	$\frac{1}{2}$	< 0.19 MeV
τ	-1	$\frac{1}{2}$	1776.84 ± 0.17 MeV
ν_τ	0	$\frac{1}{2}$	< 18.2 MeV

Table 2.2: Summary of the fermions

non-abelian gauge group [5]. In 1957, Schwinger [6], Lee and Yang [7] developed the idea of intermediate vector boson. They found that the weak current favors to be a $V-A$ structure because of the experimental results that the weak interaction requires charge changing and leptons are left-handed. The electroweak Lagrangian should also allow the violation of parity. Until 1958, Feynman and Gell-Mann constructed the universal $V-A$ weak interactions in the form [8]: $J^{+\mu}_{lepton} = \bar{\psi}_l \gamma_\mu (1-\gamma_5)\psi_\nu$.

With these understanding, the electroweak Lagrangian can be expressed,

$$\mathscr{L}_{e.m} = \bar{\psi} i\gamma^\mu D_\mu \psi - \frac{1}{4} F^i_{\mu\nu} F_i^{\mu\nu} - \frac{1}{4} G_{\mu\nu} G^{\mu\nu} \tag{2.1}$$

where, ψ indicates the spinors of fermions including the quarks, the leptons and the neutrinos. D_μ is the covariant derivative, with g and g' the gauge couplings,

$$D_\mu \equiv \quad \partial_\mu + i\frac{g}{2} T_i W^i_\mu + i\frac{g'}{2} Y B_\mu \tag{2.2}$$

And the field tensors with the gauge fields (W^i_μ with i=1,2,3 and B_μ) corresponding to the group generators are,

$$\begin{aligned} F^i_{\mu\nu} &\equiv \partial_\mu W^i_\nu - \partial_\nu W^i_\mu + g\epsilon^{ijk} W^j_\mu W^k_\nu \\ G_{\mu\nu} &\equiv \partial_\mu B_\nu - \partial_\nu B_\mu \end{aligned} \tag{2.3}$$

With the experimental observation that the strangeness changes in weak decays while such processes are strongly suppressed, Cabibbo first introduced a mixing angle between the quarks under the same generation which indicates that the $SU(2)$ gauge eigenstates are different from mass eigenstates [9]. With the development of Kobayashi and Maskawa [10] to expand the matrix to three dimensions in 1973, we can now express the quark mixing with the CKM matrix:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.4}$$

It is good to mention that they predicted a third generation of quarks in order to explain the CP violation in kaon decay. It is astonishing because even the fouth quark *charm* in the second generation hadn't been discovered at that time.

The consequence of the existence of the CKM matrix is that the charged weak currents are flavor non-diagonal. The best determined CKM matrix elements currently are listed [11],

$$\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 0.97419 \pm 0.00022 & 0.2257 \pm 0.0010 & 0.00359 \pm 0.00016 \\ 0.2256 \pm 0.0010 & 0.97334 \pm 0.00023 & 0.0415^{+0.0010}_{-0.0011} \\ 0.00874^{+0.00026}_{-0.00037} & 0.0407 \pm 0.0010 & 0.999133^{+0.000044}_{-0.000043} \end{pmatrix}$$

The square of the matrix element $\mid V_{ij} \mid^2$ represents the probability of a quark flavor i decaying to another quark flavor j. If taking the matrix elements related to the top quark, one finds that the top quark decays to bottom quark with a branching ratio close to 99.8%. The $\mid V_{tb} \mid$ could directly be measured through the single top production channels at LHC.

With the symmetric electroweak Lagrangian $\mathscr{L}_{e.m}$, no mass term is allowed either for fermions or for gauge bosons in order to preserve the gauge invariance. The $\mathscr{L}_{e.m}$ could be separated into three currents, the charged current, the electromagnetic current and the weak neutral current (Z boson), which was a prediction of the theory at that time and only confirmed more than five years later [12]. The Z vector and axial couplings to fermions, at tree level in Standard Model, are given in Table 2.3.

fermion	t_3	Q	g_v	g_a
ν_e, ν_μ, ν_τ	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$
e, μ, τ	-$\frac{1}{2}$	-1	-$\frac{1}{2} + 2sin^2\theta_w$	-$\frac{1}{2}$
u, c, t	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{1}{2} - \frac{4}{3}sin^2\theta_w$	$\frac{1}{2}$
d, s, b	-$\frac{1}{2}$	-$\frac{1}{3}$	-$\frac{1}{2} + \frac{2}{3}sin^2\theta_w$	-$\frac{1}{2}$

Table 2.3: Z vector, axial couplings to fermions

2.1.3 Quantum Chromodynamics

The strong interaction can be described by the Quantum Chromodynamics theory (QCD) with a $SU(3)_c$ gauge symmetry on quark colour fields. The hadrons are assumed to be combination of quarks, with mesons being $(q\bar{q})$ states and baryons (qqq) states. With the assumption of quarks carrying colours (red, blue and green), one can nicely classify the entire hadronic spectrum without violating Pauli's exclusion principle for fermions. Based on the knowledge that no free quarks are found in the natural world, one has to make the assumption that all the hadrons have to be colourless, which says that the same amount of red, blue and green components are combined. Then the quarks have to be confined within colour singlet bound states, which is known as confinement hypothesis. Another feature of QCD is the asymptotic freedom of quarks inside a hadron, which implies that the quarks behave as free particles at high energy.

Then, the $SU(3)_c$ gauge invariant Lagrangian can be written as,

$$\mathscr{L}_{QCD} = \quad \bar{q_f}(i\gamma^\mu D_\mu - m_f)q_f - \frac{1}{4}G_a^{\mu\nu}G^a_{\mu\nu} \tag{2.5}$$

where the covariant derivative and the gauge field intensity is,

$$\begin{aligned} D_\mu &\equiv \partial_\mu - ig_s\tfrac{\lambda_a}{2}G_a^\mu \\ G^a_{\mu\nu} &\equiv \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc}G^b_\mu G^c_\nu \end{aligned} \tag{2.6}$$

G^a_μ are the gluon fields and the coupling strength is characterized as g_s while f^{abc} are the structure constants. If the Lagrangian $\mathscr{L}_{QCD}$ is expanded, one gets the colour interaction

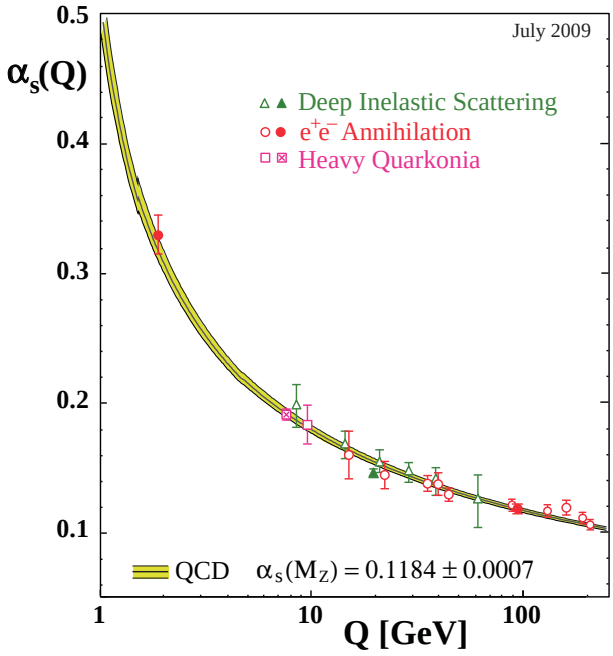


Figure 2.1: Measurements of α_s as a function of the energy scale Q

between quarks and gluons and the cubic and quartic gluon self-interaction owing to the non-abelian character of the colour group.

In spite of the simple $\mathscr{L}_{QCD}$ caused by the colour symmetry, the physics in it is quite rich. Not like the case in QED where there is only one gauge boson, the photon, there are eight gluons in QCD which are able to interact not only with quarks but also among themselves. Since all the interactions in $\mathscr{L}_{QCD}$ are given by one single universal coupling g_s, the coupling constant α_s should be expressed as a function of g_s. The problem is that α_s is close to 1 at low energy scale, which means that perturbation theory is no longer valid. Actually, α_s is not exactly a "constant". An energy scale Λ is generated using a β function,

$$\alpha_s(Q^2) = \frac{2\pi}{-\beta_1 ln(\frac{Q^2}{\Lambda^2})} \tag{2.7}$$

α_s goes to zero if the scale energy Q is much larger than Λ, which says that perturbation theory is recovered at high energy. The experimental evidence of α_s depending on the scale energy is presented in Figure 2.1.

2.1.4 The Higgs Mechanism

With the gauge symmetry groups, the matter field particles and the gauge bosons are all massless. P. W. Higgs introduced a method of spontaneous breaking of the symmetries to generate mass for gauge particles, $W^{\pm}$ bosons and Z^0 boson, while keeping the photon massless in 1964 [13] [14] [15].

In SM, when the temperature is high enough, the exact symmetries give rise to the exact conservation laws. All the elementary particles except the Higgs boson are massless. When it goes down to a critical temperature, the Higgs field is spontaneously broken down. The Lagrangian generates a small term that breaks the symmetry,

$$\mathscr{L} = \mathscr{L}_{symmetry} + \epsilon \mathscr{L}_{break} \tag{2.8}$$

In order to apply the Higgs mechanism to electroweak theory, the Higgs scalar field is introduced,

$$\mathscr{L}_{scalar} = \partial_\mu \Phi^\dagger \partial^\mu \Phi - V(\Phi^\dagger \Phi) \tag{2.9}$$

where the scalar doublet and the potential are,

$$\begin{aligned} &\Phi \equiv \binom{\phi^+}{\phi} \\ &V(\Phi^\dagger \Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \end{aligned} \tag{2.10}$$

The electroweak covariant derivative D_μ is applied in order to preserve the gauge invariance under $SU(2)_L \otimes U(1)_Y$ gauge transformation. If $\mu^2 < 0$, the symmetry breaks down spontaneously and the original symmetry group $SU(2)_L \otimes U(1)_Y$ goes into $U(1)_{em}$. The vacuum expectation of the the Higgs field could be chosen as,

$$< \Phi >_0 = \begin{pmatrix} 0 \\ \frac{v}{\sqrt{2}} \end{pmatrix}, \qquad with \qquad v = \sqrt{-\frac{\mu^2}{\lambda}} \tag{2.11}$$

Instead of writing the whole Lagrangian after symmetry breaking, only the terms containing the boson masses are presented,

$$\mathscr{L}_{mass} = \frac{g^2 v^2}{4} W_\mu^+ W^{-\mu} + \frac{g^2 v^2}{8 cos^2 \theta_w} Z_\mu Z^\mu - \frac{1}{2}(2v^2 \lambda) H^2 \tag{2.12}$$

With the $\mathscr{L}_{mass}$, we can verify that,

$$M_W = \frac{gv}{2}, \quad M_Z = \frac{gv}{2c_w} = \frac{M_w}{c_w}, \quad M_H = \sqrt{2v^2 \lambda} \tag{2.13}$$

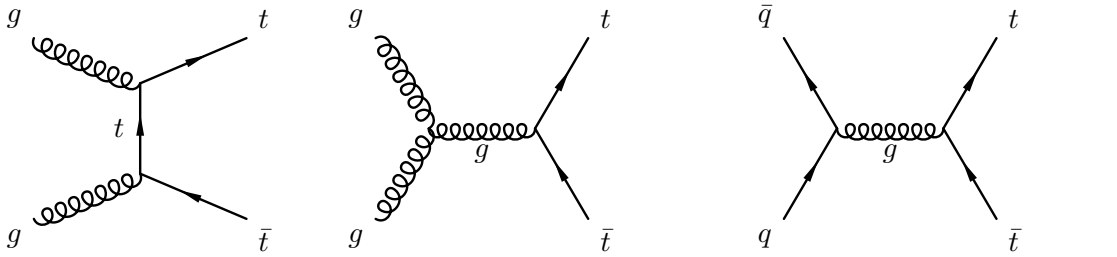
When the result from low energy phenomenology is taking into account, the vacuum expectation value is $v = (\sqrt{2}G_F)^{\frac{1}{2}} \simeq 246\,\text{GeV}$. Since the parameter λ is a priori unknown, the SM would not give a hint to the Higgs mass. One of the main purpose of LHC is to search for the Higgs boson and to measure its mass if it exists. The possible decay modes of Higgs boson is related to its mass. Final states of dibosons with W^+W^- or Z^0Z^0 are golden channels of Higgs decay. With the experiment of Large Electron-Positron (LEP) collider at CERN during 1990s, the SM Higgs boson with low mass $M_H < 114.4\,\text{GeV}$ is excluded at 95% confidence level (C.L.). The data collected from CDF and DØ experiments at Tevatron until January 2010 excludes the range between 162 GeV and 166 GeV at 95% C.L. [16]

2.2 Top Quark Physics

As shown in Section 2.1.2, a third generation of quarks was predicted in 1973. Not long after this remarkable prediction, the fourth quark *charm* was discovered at Brookhaven National Laboratory (BNL) [17] and the Stanford Linear Accelerator Center (SLAC) [18] in 1974 and the fifth quark *bottom* at Fermilab in 1977 [19]. This strongly suggested the existence of the sixth quark *top* to complete the doublets.

After the gauge bosons W and Z were found in the Super Proton Synchrotron (SPS) at CERN in early 1980s, a lower bound was also set for the top mass at 77 GeV. But the top quark was left undiscovered. Before the start of LHC, the Tevatron is the only place which has the ability to produce a heavy member of the quark family like top. The top quark was finally discovered by the experiments CDF and DØ at Tevatron in 1995 [20] [21] at a mass of 176 GeV and 180 GeV respectively. After starting to gather data of LHC at 7 TeV collision in 2010, the first top quark candidates are found in June and July. It is quite exciting to expect that the properties of the top quark will be carefully studied in the next few years.

The top quarks can be produced at LHC through top pair production ($t\bar{t}$) and single top production. The $t\bar{t}$ events are from strong interaction of quark anti-quark annihilation mode or gluon fusion mode. The leading order feynman diagrams are shown in Figure 2.2. By the proton-proton collision at LHC, the contribution of the two processes are approximately 82% and 18% at $\sqrt{s} = 7\,\text{TeV}$ [22].



(a) gluon fusion mode: $gg \to t\bar{t}$ (b) quark pair interaction: $q\bar{q} \to t\bar{t}$

Figure 2.2: The Feynman diagrams for $t\bar{t}$ production at leading order at LHC

2.2.1 The Cross Section for Top Pair Production

The total cross section for $t\bar{t}$ in pp collision at a center of mass energy $\sqrt{s}$ can be calculated using the formula,

$$\sigma_{pp\to t\bar{t}}(\sqrt{s}, m_t) = \sum_{i,j=q,\bar{q},g} \int dx_i dx_j \sigma_{ij\to t\bar{t}}(\tilde{s}, \alpha_s, \mu_R, m_t) f_i^p(x_i, \mu_F^2) f_j^p(x_j, \mu_F^2) \tag{2.14}$$

where i and j indicate the partons including quarks, anti-quarks and gluons in the collision. f_i^p is the i's parton distribution function (PDF) for the proton (p), which means the probability density of finding a parton of a given x_i, while p_i is the four-momentum of the protons and x_i is the four-momentum fraction of i parton to the proton. An example of the $PDFs$ (actually $xf(x, Q^2)$) as a function of x_i are shown in Figure 2.3. The square of the total energy of the collision in the parton level is represented by $\tilde{s}$,

$$\tilde{s} \equiv (x_1 p_1 + x_2 p_2)^2 \approx 2x_1 x_2 p_1 p_2 \approx x_1 x_2 (p_1 + p_2)^2 = x_1 x_2 s \tag{2.15}$$

The parton mass (m_u, m_d, ...) has been neglected for they are quite negligible comparing to the collision energy. It is obliged that the parton collision energy $\sqrt{\tilde{s}}$ is greater than $2m_t$ in order to produce a top pair. Thus, if $x_1 = x_2 = x$ is set, the typical value of x for $t\bar{t}$ production would be,

$$x = \sqrt{\frac{\tilde{s}}{s}} = \frac{2m_t}{7\,\text{TeV}} \approx 0.05 \tag{2.16}$$

By looking at the Figure 2.3, one find in the plot that the partonic density largely depends on the four-momentum fraction x. It is noticed that gluon dominates in the range of low x, which is the case of LHC. With the energy level of Tevatron ($x \approx \frac{2m_t}{2\,\text{TeV}} \approx 0.17$), u quark and d quark have the largest contribution. In the formula of cross section, the choice of the factorization scale μ_F will affect the structure function by entering the definition of PDF and is eventually related to the cross section $\sigma_{ij\to t\bar{t}}$. And, μ_R is the renormalization scale, which is in general identical to μ_F. It is convenient and conventional to set $\mu_R = \mu_F = \mu$ and in the production of top pair, $\mu = m_t$ is chosen.

The $PDFs$ are experimentally extracted from the structure functions as measured in deep inelastic scattering. In the real calculation the analysis has to expend until a finite number of orders of α_s. The integrated total cross section to the Next-to-Leading-Order

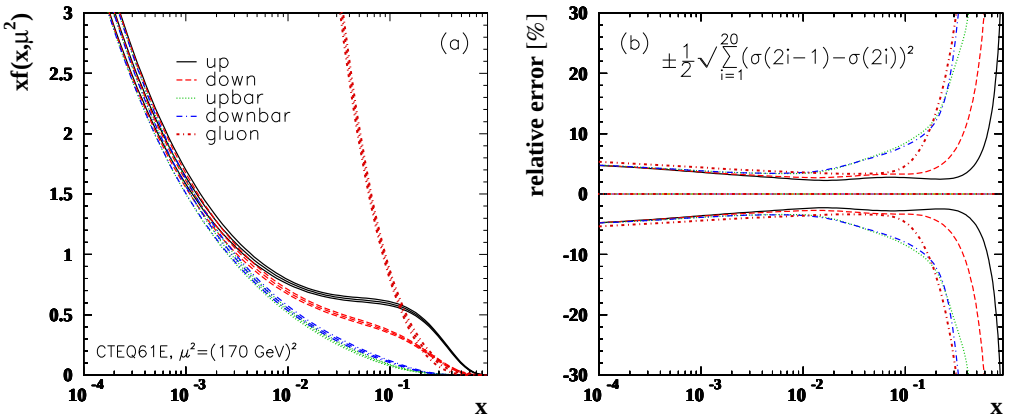


Figure 2.3: (left) CTEQ6.1 [23] PDFs with their uncertainty bands with $Q^2 = (170\,\mathrm{GeV})^2$. (right) Relative uncertainties on the PDFs[24].

(NLO) can be written in a simple form,

$$\sigma_{ij}(\tilde{s}, m_t^2, \mu^2) = \frac{\alpha_s^2(\mu^2)}{m_t^2}\Big[\sigma_{ij}^{(0)}(\rho) + 4\pi\alpha_s(\mu^2)\Big(\sigma_{ij}^{(1)}(\rho) + \bar{\sigma}_{ij}^{(1)}(\rho) ln\frac{\mu^2}{m_t^2}\Big)\Big] \tag{2.17}$$

where the definition for ρ is $\rho = \frac{4m_t^2}{\tilde{s}}$. The leading order terms for the total cross section are $gg \to t\bar{t}$ [25] and $q\bar{q} \to t\bar{t}$ [26] corresponding to the Feynman diagrams in Figure 2.2, which are well known in the following form with $\beta \equiv \sqrt{1-\rho}$,

$$\begin{aligned} \sigma_{gg}^{(0)}(\tilde{s}) &= \frac{\pi\alpha_s^2}{3\tilde{s}}\Big[(1+\rho+\frac{\rho^2}{16})ln\frac{1+\beta}{1-\beta} - \beta(\frac{7}{4}+\frac{31}{16}\rho)\Big] \\ \sigma_{q\bar{q}}^{(0)}(\tilde{s}) &= \frac{8\pi\alpha_s^2}{27\tilde{s}}\beta[1+\frac{\rho}{2}] \end{aligned} \tag{2.18}$$

It is argued in (ECFA LHC Workshop 1990 CERN 90-10) that the sub-dominant $2 \to 3$ processes soft gluon radiation like $gg \to t\bar{t}g$, $q\bar{q} \to t\bar{t}g$ etc are added up to contribute less than 10%. Physically, these α^3 terms have distinct components including the initial state gluon bremsstrahlung, the final state gluon bremsstrahlung, the gluon splitting and the flavor excitation. The contributions of the processes are more important for $\sqrt{\tilde{s}} \gg 2m_t$ which is more likely the situation of LHC rather than Tevatron. The NLO theoretical prediction is rather well known, and a more recent result can be found in reference [27].

A theoretical prediction of the total cross section to Next-to-Next-to-Leading-Order (NNLO) is investigated for Tevatron at $\sqrt{s} = 1.96\,\mathrm{TeV}$ as well as for LHC at $\sqrt{s} = 10\,\mathrm{TeV}$ and $14\,\mathrm{TeV}$ [28]. A recent study on the top pair production beyond NLO is presented in the

	Tevatron	LHC (7 TeV)	LHC (10 TeV)	LHC (14 TeV)
σ_{LO}	$6.66^{+2.95+0.34}_{-1.87-0.27}$	122^{+49+6}_{-32-7}	$305^{+112+14}_{-76\ -16}$	$681^{+228+26}_{-159-34}$
σ_{NLL}	$5.20^{+0.40+0.29}_{-0.36-0.19}$	103^{+17+5}_{-14-5}	253^{+44+10}_{-36-10}	$543^{+101+18}_{-88\ -19}$
σ_{NLO}	$6.72^{+0.36+0.37}_{-0.76-0.24}$	159^{+20+8}_{-21-9}	402^{+49+17}_{-51-18}	$889^{+107+31}_{-106-32}$
$\sigma_{\mathrm{NLO+NNLL}}$	$6.48^{+0.17+0.32}_{-0.21-0.25}$	146^{+7+8}_{-7-8}	368^{+20+19}_{-14-15}	813^{+50+30}_{-36-35}

Table 2.4: The total cross section in *pb* at different perturbation level.

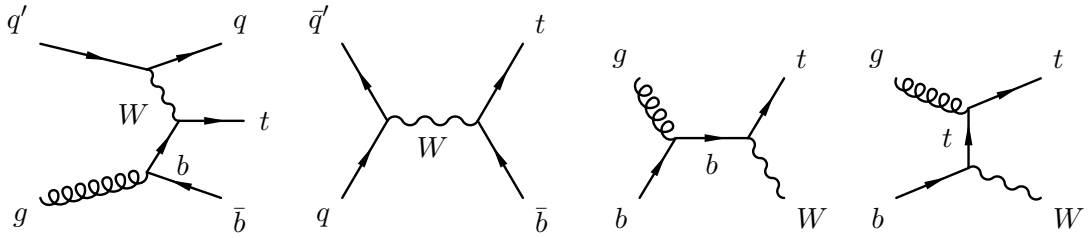
reference [29] with NLO plus a Next-to-Next-to-Leading logarithmic (NNLL) resummation scheme. A summary of the results at Tevatron with 1.96 TeV and at LHC with 7 TeV, 10 TeV and 14 TeV is listed in Table 2.4.

The results in Table 2.4 are obtained by using the choice of $\mu_F = m_t$. The first set of errors refers to perturbative uncertainties associated with scale μ variations, and the second to the PDF uncertainties. The PDFs used is MSTW2008NNLO along with the choice of $\alpha_s(M_Z) = 0.117$ and $m_t = 173.1$ GeV. The most advanced prediction is the NLO+NNLL expansion in the table. The result by changing only $\mu_F = 400$ GeV for LHC at 7 TeV would be $149 \pm 7 \pm 8$ in the same reference [29].

A different top process is the single top production via weak interaction. The significant importance of the channels is to directly measure the CKM matrix V_{tb}. The first evidence is found by DØ [30] and CDF [31] in 2009. The expected cross section for single top "s-channel" at 7 TeV is $3.99 \pm 0.05^{+0.14}_{-0.21}$ *pb*, and for "W + t" channel is $7.8 \pm 0.2^{+0.5}_{-0.6}$ *pb* [32], while the leading order diagrams of the processes can be found in Figure 2.4.

2.2.2 The Top Quark Decay

In the SM, about 99.8% of the top quarks decay in the mode: $t \to W + b$. It can also decay into Wd and Ws through Cabibbo-Kobayashi-Maskawa (CKM) mixing, which is so tiny and thus neglected in the analysis. Table 2.5 shows the branching fractions of the W boson decay. Eventually, the approximate top pair decay channels branching ratios are reassembled in Figure 2.5. What's so unique is that the top quark is extremely short lived



(a) t-channel (b) s-channel (c) Wt channel

Figure 2.4: The Feynman diagrams for single top production at leading order at LHC

because of its large mass. It decays before it is hadronized, which unveils the properties of a bare quark.

W Decay Mode		Branching Fraction
Leptonic Decay	$e\nu_e$	10.75 ± 0.13%
	$\mu\nu_\mu$	10.57 ± 0.15%
	$\tau\nu_\tau$	11.25 ± 0.20%
Hadronic Decay		67.60 ± 0.27%

Table 2.5: The branching fraction of W boson decay

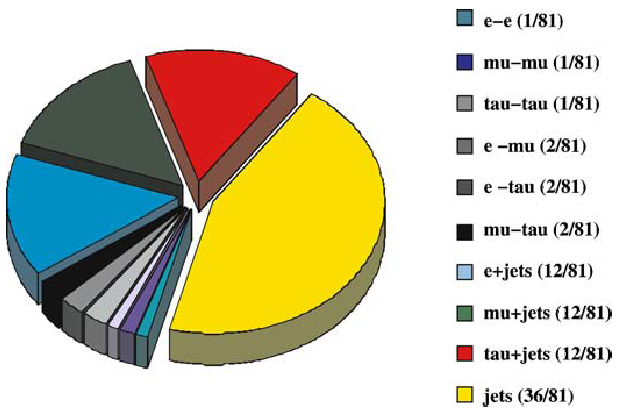


Figure 2.5: The branching ratios of the different $t\bar{t}$ decay channels at Born level[1].

Chapter 3

THE LHC AND ATLAS DETECTOR

In this chapter the world's large hadron collider, LHC, is first briefly described. Then, we concentrate to describe the detector ATLAS. The subdetectors, the inner detector, the electromagnetic calorimeter, the hadron calorimeter and muon spectrometer, are introduced. We then focus on the objects reconstruction which is critically important to event selection. Finally, the measurement of the luminosity is also presented.

3.1 The LHC

The Large Hadron Collider (LHC) is nowadays the world's largest and highest-energy instrument used for the experiments on high energy physics. The LHC sits in a circular tunnel 27 km in circumference, located from 50 to 175 metres beneath the boarder of France and Switzerland near Geneva. The LHC is an international project joining countries all over the world and is organized by the European Organization for Nuclear Research (CERN).

The LHC is built with the intention of investigating or answering some big questions in high energy physics, such as the beginning of the universe, the existence of the Higgs boson(s), the dark matter, a bunch of new particles predicted by supersymmetry, etc. Two beams of hadrons, either protons or heavy (lead) ions, are accelerated inside the circular in opposite directions and collide at very high energy. The designed proton-proton collision energy is at 7 TeV per beam. In November 2009, the proton beams were successfully circulated in the main ring of the LHC, and first data was recorded at 450 GeV per beam. On 30 March 2010, collisions took place between two beams, each of which is at 3.5 TeV. This is the new world record for the highest energy particle collisions.

The LHC planned to stay at the energy of 3.5 TeV and collect the first fb^{-1} data in 2010 - 2011 before going to the designed beam energy. In order to achieve such high energy of 7 TeV per beam, the LHC needs to provide a magnetic field of 8.33 Tesla (T) with the superconducting dipole magnets. Totally, over 1600 magnets are installed, each weighting more than 27 tonnes. Besides this, approximately 96 tonnes of liquid helium are supplied to keep the magnets at the operating temperature of 1.9K (-271.25°C). The direction of the beams in the circular path is kept by 1,232 dipole magnets, while some other 392 quadrupole magnets are used to focus the beam so that the chances of the collisions between the two beams are maximized.

Figure 3.1 shows the chain of accelerators at CERN. Protons obtained from hydrogen atoms by removing electrons are injected from the linear accelerator (LINAC2) generating 50 MeV. Protons are then fed into the Proton Synchrotron Booster (PSB) and accelerated to 1.4 GeV. Consequently, protons are injected into the Proton Synchrotron (PS), where they are boosted to 26 GeV. After that, the Super Proton Synchrotron (SPS) is used to

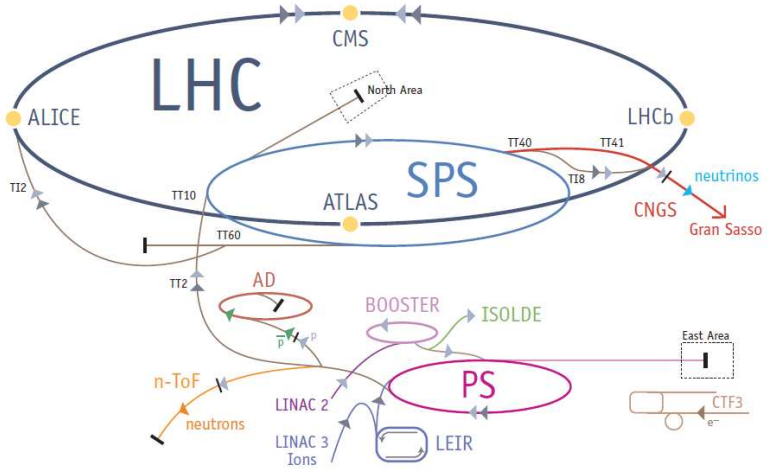


Figure 3.1: The accelerator chain at LHC

further increase the protons energy to 450 GeV before they are finally injected into the LHC. Protons are first accelerated about 20 minutes in the ring to reach the colliding energy, 7 TeV or 14 TeV. Protons in LHC are not with a continuous beam current. They are separated into 2808 bunches and the time interval between two bunch-bunch collisions will last at least 25 ns. They will stay in the main ring for several hours with a speed rather close to the speed of light while collisions are collected at the four interaction points. The LHC mainly focus on the proton-proton collision, while the heavy-ion collisions and light-ion collisions proceed for a shorter period of time. Lead ions are used and accelerated in a similar method. They finally reach the LHC ring with energy of 2.76 TeV per nucleon.

There are four main detectors, A Large Ion Collider Experiment (ALICE), A Toroidal LHC ApparatuS (ATLAS), Compact Muon Solenoid (CMS) and Large Hadron Collider beauty (LHCb), which are located at four colliding points in the ring. ATLAS and CMS are the so called general purpose detectors for proton-proton collision. Both of them are looking forward to hunt for the Higgs boson, to look for the sign of new physics, to search the clues about dark matter and so on. With the two separately designed detectors, they can cross confirm any new discovery of physics. The ALICE experiment with the heavy-ion program aims to investigate the "liquid" state of matter, so called the quark-gluon plasma, which is believed to exist only shortly after the big bang. LHCb is a detector specific for b-physics, which is much smaller than ATLAS and CMS. One of the main purpose is to measure the parameters of CP violation in the interactions of b-hadrons. In addition, the

other two detectors, TOTal Elastic and diffractive cross section Measurement (TOTEM) and Large Hadron Collider forward (LHCf), are much smaller and are built to do some very specialized research. TOTEM shares the intersection point IP5 with CMS. The detector aims at measuring the total cross section of LHC, the elastic scattering and the diffractive processes. LHCf is going to collect the particles in the forward region in line with the beam direction. The experiment is to complement other experimental high-energy cosmic ray measurements.

One of the most prominent issue for the LHC is to find out the origin of the spontaneous symmetry breaking generation masses. The existence of SM (or beyond SM) Higgs boson(s) would be a good manifestation. For example, the ATLAS detector is optimized to search for the Higgs particles, heavy W- and Z-like objects, supersymmetric particles etc, and study the top quark in detail. The LHC requires high luminosity with the primary goal of $10^{34} cm^{-2} s^{-1}$. The initial lower luminosity running put emphasis on the performance necessary signatures such as tau-lepton detection and heavy-flavors tags with secondary vertices.

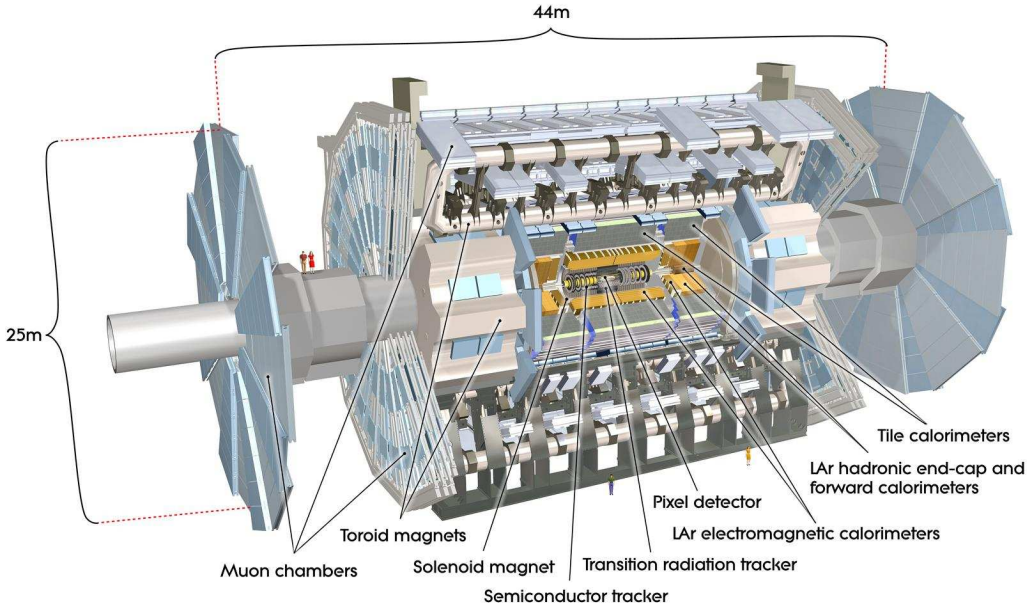


Figure 3.2: The construction of the ATLAS detector.

3.2 The ATLAS Experiment

The ATLAS detector is the largest particle detector ever constructed. ATLAS is 46 metres long, 25 metres high and 25 metres wide, weighing about 7000 tons. ATLAS consists of the barrel and the end caps leaving only small part of the space uncovered by the detectors. Some conventional parameters and concepts are made to define the performance of the detectors [33].

- z-axis: the beam direction
- x-y plane: transverse to the beam direction
- x-axis: pointing from the interaction point to the centre of the LHC ring
- y-axis: pointing upwards
- ϕ: the azimuthal angle in x-y plane, range $[-\pi, \pi]$
- θ: the polar angle from the z-axis
- η: the pseudorapidity defined as $\eta \equiv -ln[\tan(\frac{\theta}{2})]$
- ΔR: distance in the pseudorapidity-azimuthal angle space $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$

ATLAS is generally consisting of four main parts, the magnet system, the inner detector, the calorimeters and the muon spectrometer. An overview of ATLAS detector is presented in Figure 3.2.

3.2.1 Magnet System

The ATLAS superconducting magnet system is an arrangement of a central solenoid (CS) and a surrounding system of three large air-core toroids, two end-cap toroids (ECT) and one barrel toroid (BT). The CS provides the magnetic field to the inner detector and the toroids generate the magnetic field for the muon spectrometer. The overall magnet system measures 26 m in length and 20 m in diameter.

The CS provides a central field of about 2 T along the z-axis. BT and ECT generate a peak magnetic field at 3.9 and 4.1 T respectively. The bending power is lower in the transition region $1.3 < \mid \eta \mid < 1.6$ where the two magnets overlap.

The position of the CS in front of the EM calorimeter demands a careful minimization of the material in order to achieve the desired calorimeter performance. As a consequence, the CS and the liquid argon calorimeter share one common vacuum vessel, thereby eliminating two vacuum walls.

Each of the three toroids consists of eight coils assembled radially and symmetrically around the beam axis.

3.2.2 Inner Detector

The ATLAS inner detector (ID) will see approximately 1000 charged particle tracks for every beam crossing at the design luminosity of the LHC. High precision measurements of the momentum and vertex resolution in ID will be required to achieve the expected physics performance. Pixel, silicon microstrip tracker (semiconductor tracker: SCT) and the transition radiation tracker (TRT) are the three components of the ID.

The ID is approximately 7 meters long and 2.3 meters in diameter. It is located within the solenoidal magnetic field of about 2 Tesla. The layout of the detector is illustrated in Figure 3.3, while the basic overall parameters are summarized in Table 3.1.

The pixel detector is made up of 1744 modules arranged in three layers in the barrel and three disks in each endcap. The pixel cell size (R-ϕ $\times$ z) is 50 μm $\times$ 400 μm with a corresponding intrinsic resolution of 10 μm $\times$ 115 μm. A more precious measurement is in the ϕ direction, the bending direction of the magnetic field. There are in total 1744

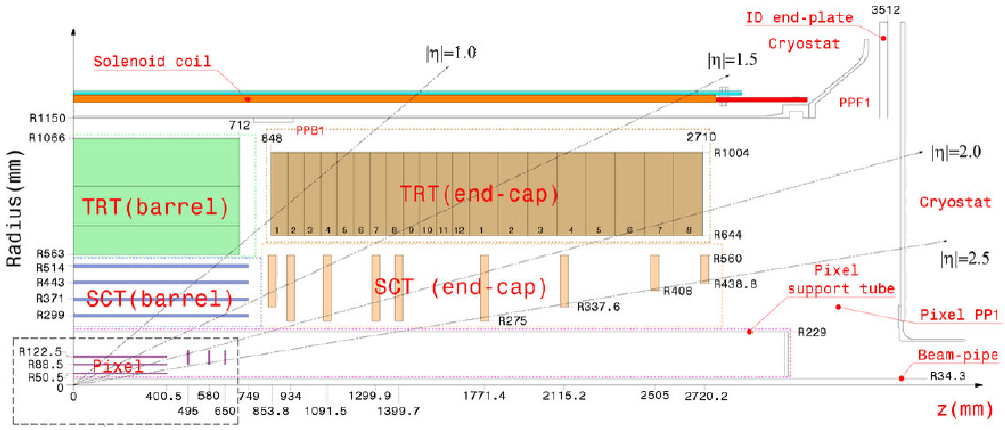


Figure 3.3: Plan view of a quarter-section of the inner detector showing each of the major elements with its active dimensions.

Item		Radial extension (mm)			Length (mm)		
Pixel		45.5	$<$ R $<$	242	0	$<\lvert z \rvert<$	3092
SCT	(barrel)	255	$<$ R $<$	549	0	$<\lvert z \rvert<$	805
	(end-cap)	251	$<$ R $<$	610	810	$<\lvert z \rvert<$	2797
TRT	(barrel)	554	$<$ R $<$	1082	0	$<\lvert z \rvert<$	780
	(end-cap)	617	$<$ R $<$	1106	827	$<\lvert z \rvert<$	2744

Table 3.1: Main parameters of the ID overall envelope

modules. The innermost layer is called the "vertexing layer" with a radius of 51 mm used to optimize the vertex measurement.

The SCT consists of four barrel layers and 9 disks in each endcap giving 4 space points for every track. There are 4088 modules in all. Each module is made of four silicon microstrip sensors glued back to back with 40 mrad stereo angle to give a 3D reconstruction. The strip pitch is about 80 μm with a resolution of 17 μm in the ϕ direction. A space point resolution in z (barrel) or R (end-cap) direction is about 580 μm.

The outer sub-detector TRT is made up of straw drift tubes with a diameter of 4 mm. On average 30 hits are provided for a track within $\mid \eta \mid < 2.0$. The barrel is segmented into 96 modules arranged in three rings while each endcap consists of 20 wheels. The TRT gives only R-ϕ information, for which it has an intrinsic resolution of 130 μm per straw.

3.2.3 Calorimeters

The total weight for the calorimeter system is about 4000 tons. Generally, the calorimetry consists of an electromagnetic (EM) calorimeter covering the pseudorapidity region $\mid \eta \mid < 3.2$, a hadronic calorimeter of barrel $\mid \eta \mid < 1.7$ and end-cap $1.5 < \mid \eta \mid < 3.2$ and forward calorimeters covering $3.1 < \mid \eta \mid < 4.9$.

The EM calorimeter is a lead liquid-argon (LAr) detector, where the LAr geometry view is shown in Figure 3.4. Over the range $\mid \eta \mid < 1.8$, it is preceded by a pre-sampler detector used to correct for the energy lost in the material ahead of the calorimeter.

The hadronic barrel calorimeter contains three sections: the central barrel and two identical extended barrels, which are based on a sampling technique with plastic scintillator plates (tiles) embedded in an iron absorber. The hadronic end-cap calorimeter is using the radiation-hard LAr technology, which is also used for a copper LAr detector with parallel-plate geometry, and the forward calorimeter, a dense LAr calorimeter with rod-shaped electrodes in a tungsten matrix.

The barrel cryostat surrounding the inner detector contains the barrel EM calorimeter. The solenoid is integrated into the vacuum of the barrel cryostat and is placed in front of the EM calorimeter. Two end-cap cryostats house the end-cap EM and hadronic calorimeters, as

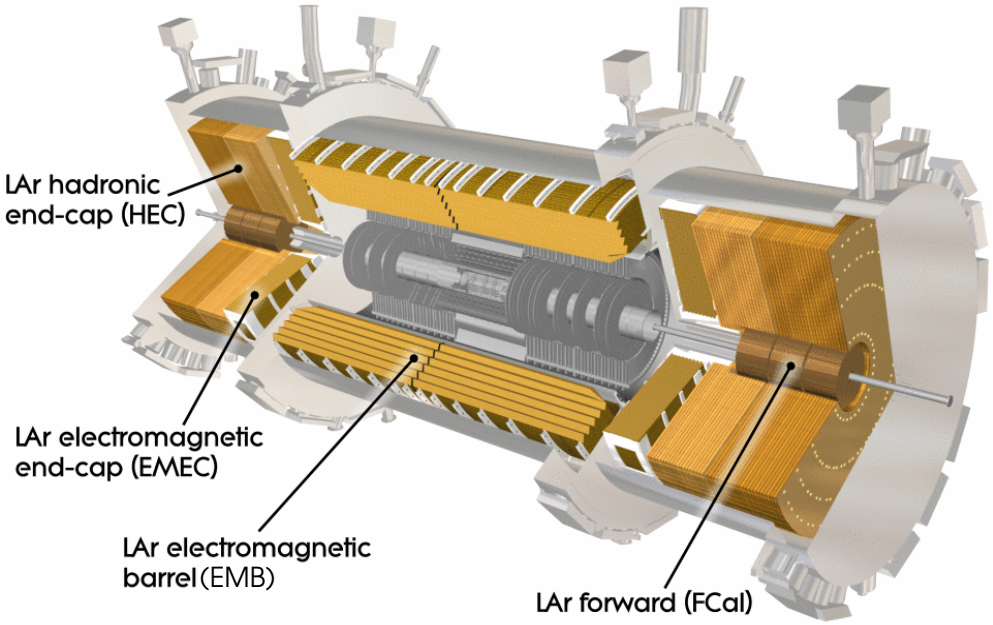


Figure 3.4: LAr calorimeter for EM, HEC and FCal[2]

well as the integrated forward calorimeter. The barrel and extended barrel tile calorimeters support the LAr cryostats. The signals leaving the cryostats go through the cold-to-warm feedthroughs located between the barrel and the extended barrel tile calorimeters and at the back of each end-cap.

Electromagnetic Calorimeter

The EM calorimeter is divided into a barrel part ($| \eta | < 1.475$) and two end-caps, each of which has an inner wheel ($2.5 < | \eta | < 3.2$) and an outer wheel ($1.375 < | \eta | < 2.5$). The barrel calorimeter consists of two identical half-barrels, separated by a small gap (6 mm) at $z = 0$. Over its full coverage, a lead LAr detector with accordion-shaped Kapton electrodes and lead (Pb) absorber plates is used. The lead thickness and LAr gap thickness are listed in Table 3.2 as a function of η.

The lead thickness in the absorber plates has been optimized for the EM calorimeter

	Rapidity		Pb (mm)	LAr gap (mm)
barrel		$\mid \eta \mid < 0.8$	1.5	2.1
	$0.8 <$	$\mid \eta \mid < 1.475$	1.1	2.1
end-cap	$1.375 <$	$\mid \eta \mid < 2.5$	1.7	2.8-0.9
	$2.5 <$	$\mid \eta \mid < 3.2$	2.2	3.1-1.8

Table 3.2: Lead and LAr gap thickness as a function of rapidity

performance in energy resolution. The LAr gap has a constant thickness in the barrel and increases with radius in the end-cap because of the increase of the amplitude of the accordion waves with radius.

The EM calorimeter is segmented into three longitudinal sections over the region of $\mid \eta \mid < 2.5$ devoted to precision physics. The front sample is the strip section, which has a constant thickness of about 6 radiation lengths (X_0). It acts as a "preshower" detector, enhancing particle identification of γ/π^0 and e/π and providing a precise position measurement in η. The middle section segmented into square towers of size $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ extends the thickness up to 24 X_0. The granularity of the back section is 0.05 in $\mid \eta \mid$ and the thickness varies between 2 X_0 and 12 X_0.

In the region of $\mid \eta \mid < 1.8$, a pre-sampler (PS) is used to correct the energy loss of electrons and photons, where the material exceeds about 2 X_0. The transitional part of barrel and end-cap with $1.37 < \mid \eta \mid < 1.52$ is not used for precision physics measurements because of the bad performance caused by the large amount of material in front of the EM calorimeter.

The sketches of a $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ region of the EM calorimeter with the different layers in depth is presented in Figure 3.5.

Hadronic Calorimeters

The hadronic calorimeters use different techniques to best meet the physics performance requirements over a large range $\mid \eta \mid < 4.9$. The iron scintillating-tile technique is used for

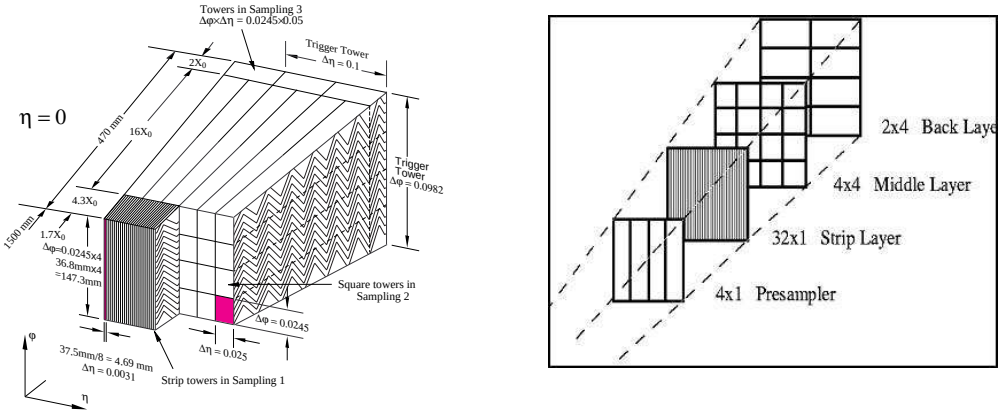


Figure 3.5: Sketches of the EM calorimeter with the different layers in depth.

the barrel and extended barrel tile calorimeters within $| \eta | < 1.7$. LAr calorimeters are chosen for the hadronic end-cap calorimeter (HEC, $| \eta | < 3.2$) and the high density forward calorimeter (FCAL, $3.1 < | \eta | < 4.9$). The HEC and the FCAL are housed together in the same cryostat as the EM end-caps. A relatively large thickness of the whole hadron calorimeter is required to obtain good containment for hadronic showers and reduce leakage into the muon system.

The hadronic barrel calorimeter, called the tile calorimeter, is a sampling calorimeter using iron as the absorber and scintillating tiles as the active material. Glued by the tiles of 3 mm thick, the iron plates in one period are 14 mm in thickness. The tile calorimeter shown in Figure 3.6 is composed of one barrel ($| \eta | < 1.0$) and two extended barrels ($0.8 < | \eta | < 1.7$) divided into 64 modules. It is radially extending from a radius of 2.28 m to 4.25 m and longitudinally segmented in three layers of approximately 1.4λ, 4.0λ and 1.8λ thick at $| \eta | = 0$, where λ is the interaction length. The granularity is $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ in the first two layers and 0.2×0.1 in the last layer. The tile calorimeter is placed behind the EM calorimeter and the solenoid coil with a distance of about 1.2λ.

The intermediate tile calorimeter (ITC) in the gap of barrel and extended barrel calorimeter contains a scintillator in $1.0 < | \eta | < 1.6$. The scintillator plays an important role in sampling the energy lost in the cryostat walls and dead material.

The HEC and FCAL are the hadronic liquid argon calorimeters. Each HEC consists of two equal diameter independent wheels, of outer radius 2.03 m. The first wheel is built out of 25 mm copper plates divided into two longitudinal readout segments with 8 layers and

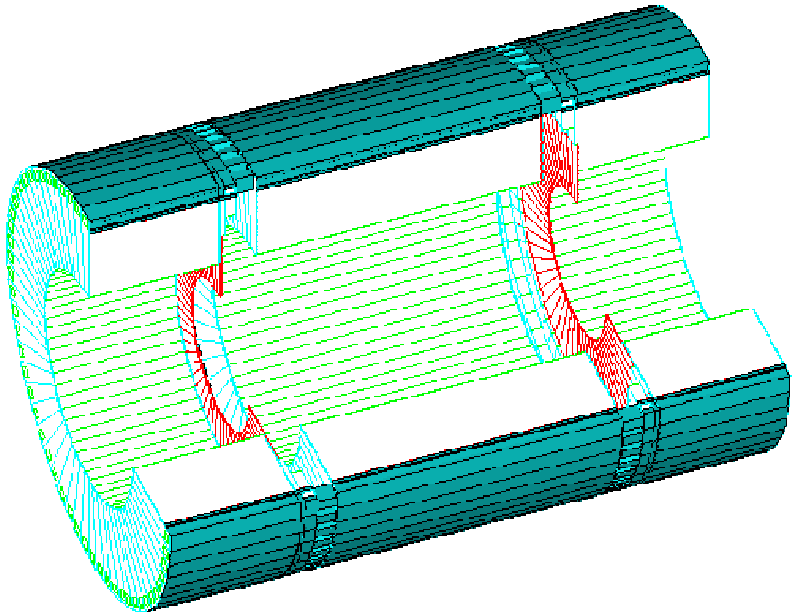


Figure 3.6: The tile calorimeter barrel and extended barrel

16 layers respectively. The second wheel uses 50 mm plates having one plate of 16 layers. The readout cells are pointing in ϕ and η directions.

The forward calorimeter is integrated into the end-cap cryostat, with a front face at about 4.7 m away from the interaction point. Thus, the FCAL is rather challenging due to the high level of radiation. On the other hand, the integrated FCAL provides a uniformity of the calorimetric coverage and reduces radiation background levels in the muon spectrometer.

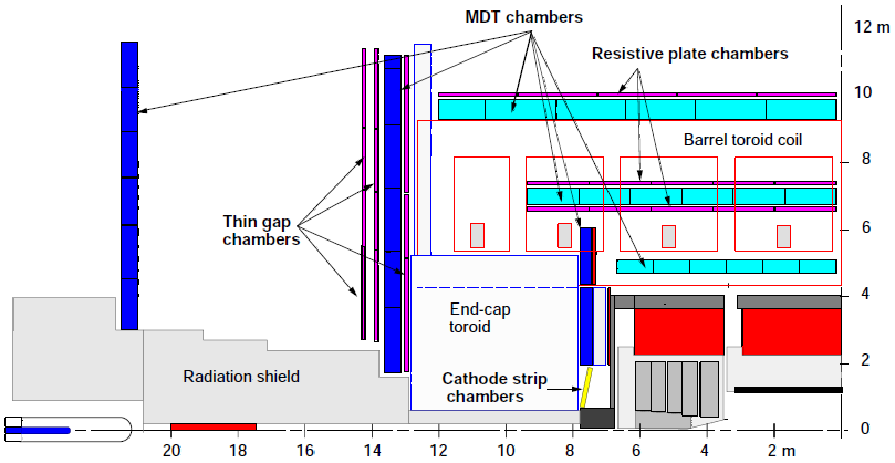


Figure 3.7: The structure of the muon spectrometer

3.2.4 Muon Spectrometer

The muon spectrometer consists of monitored drift tubes (MDT), resistive plate chambers (RPC), thin gap chambers (TGC) and cathode strip chambers (CSC), which can be seen in Figure 3.7, covering a region of $|\eta| < 1.0$ with barrel and $1.0 < |\eta| < 2.7$ with end-caps.

The magnet system providing approximately 1 tesla field, is constructed with a set of air-core toroids (8 coils) for the barrel and other two sets for the end-caps. The magnetic field provided is rather non-uniform, especially in the barrel endcap transition region near $1.0 < |\eta| < 1.4$.

A precision tracking measurement in the principal spectrometer bending plane is provided by MDTs, over most of the η range. However, at large pseudo-rapidities and close to the interaction point, MDTs would have occupancy problems, so CSCs with higher granularity are used in the innermost plane over $2 < |\eta| < 2.7$. The sense wires of the MDTs are mechanically isolated from each other. Thanks to this, the MDTs could provide an accuracy of $\sim$30 μm. RPCs used in the barrel and TGCs in the end-cap regions serve as the trigger system covering the pseudorapidity range $|\eta| < 2.4$. Table 3.3 gives a summary of the muon spectrometer detector components.

The MDT tubes are operated with a non-flammable mixture of Argon (93%) and CO_2 (7%) at 3 bar absolute pressure and the single-wire resolution is around 80 μm. To improve the resolution of a chamber beyond the single-wire limit , the MDT chambers are constructed

Type	Purpose	Location	Coverage	Chambers	Channels
MDT	Tracking	barrel, end-cap	$0.0 < \mid \eta \mid < 2.7$	1194	370k
CSC	Tracking	end-cap layer 1	$2.0 < \mid \eta \mid < 2.7$	32	67k
RPC	Trigger	barrel	$0.0 < \mid \eta \mid < 1.0$	596	355k
TGC	Trigger	end-cap	$1.0 < \mid \eta \mid < 2.4$	192	440k

Table 3.3: Muon spectrometer sub-detectors

from 2×4 monolayers of drift tubes for the inner station and 2×3 monolayers for the middle and outer stations. Each drift tube is read out at one end by a low-impedance current sensitive preamplifier.

The CSCs, arranged in 2×4 layers, are multiwire proportional chambers with cathode strip readout and with a symmetric cell. The baseline CSC gas is a non-flammable mixture of Ar (30%), CO_2 (50%) and CF_4 (20%). One obtains the coordinate precision by measuring the charge induced on the segmented cathode. Good spatial resolution is achieved by segmentation of the readout cathode and by charge interpolation between neighboring strips. It is sensitive to the inclination of tracks and the Lorentz angle.

The RPCs are gaseous detectors with the gas mixture based on $C_2H_2F_4$ and some small proportion of SF_6, which is non-flammable and allows for a relatively low operating voltage. The RPCs provide a typical space-time resolution of 1 cm × 1 ns with digital readout. Each chamber consists of two detector layers and four readout strip panels, which are held together by two support panels providing the required stiffness of the chambers. To preserve the excellent intrinsic time resolution of the RPCs, the readout strips are optimized for good transmission properties and are terminated at both ends to avoid signal reflections.

The TGCs are similar in design to the multiwire proportional chambers with a gas mixture of CO_2 (55%) and n-C_5H_{12} (45%). The gas mixture is highly flammable and requires adequate safety precautions. The TGCs are constructed in three stations with doublets and triplets of chambers. The inner station is used to measure the second coordinate, while the middle station provides the trigger and the second coordinate measurement. To form a

trigger signal, several anode wires are grouped together to feed a common readout channel.

3.3 Objects Reconstruction in ATLAS

Since $t\bar{t}$ decay channels produce leptons, missing transverse energy, jets (and in particular b-tagged jets), identifying and precisely measuring these decays require a good reconstruction of all these objects. Electrons are reconstructed by the combined information from inner detector and EM calorimeter. Jets will mainly leave their energy in the hadronic calorimeters, while algorithms used to reconstruct the jets are investigated. Muons fly a long distance to leave a track into the muon spectrometers. Eventually, $E_{\mathrm{T}}^{\mathrm{miss}}$ is obtained after all the objects in one event are considered. The b-tagging algorithm used in the thesis is called SV0 tagger, which needs the reconstruction of the secondary vertex. In the end of this section, the measurement and calibration of luminosity is introduced.

3.3.1 Reconstruction and Identification of Electrons

The offline algorithms for electron reconstruction and identification are used to reconstruct good electrons of high p_{T} and low p_{T} as well. An overlap removal procedure is applied to avoid one electron being reconstructed into two electrons or into one electron plus a jet. There are three algorithms for electron reconstruction. The first one is the standard algorithms dedicated for isolated high p_{T} electrons, which is seeded by a cluster in the EM calorimeter and extrapolated with the tracks in ID. This algorithms is the most important for the high p_{T} isolated electrons in $t\bar{t}$ decays, and will be described in more details below. The second one is mainly for low p_{T} electrons and electrons in jets, seeded by a track in the ID. The reconstruction of forward electrons in the region $2.5 <| \eta |< 4.9$ is performed by a third algorithm, where no track matching is required.

The standard electron reconstruction starts with a clusterisation algorithm called sliding window clustering. The clusterisation is basically separated into three steps:

- Tower Building: An electromagnetic tower approximately over 3 GeV is built by summing the energy of all cells across the longitudinal layers.

- Seeds Finding: If the transverse energy in a size fixed window (5×5) is above a threshold of 3 GeV, a seed is formed for this window. A smaller window size is used for position computation for the seeds, saying 3×3 cells. If two seeds are overlapped with a distance $\Delta\eta \leq 2$ or $\Delta\phi \leq 2$ in cell units, only the seed with larger transverse energy is kept while the other one is removed.

- Cluster Filling: The positions of the found seeds are used for the final clusters, which will use the cells around. The final clusters are in rectangle with some different size 3×5, 3×7 and 5×5. The e/γ candidates with a given cluster size above are kept after the e/γ identification is done.

After clustering, the reconstructed tracks in the ID, which do not belong to a photon-conversion pair, are searched and required to match the clusters. A track is regarded as matching the cluster if they are within a distance of $\Delta\eta < 0.05$ and $\Delta\phi < 0.1$. The second

sampling of the cluster is used for the distance computation between the track and the cluster.

The identification for standard electrons is based on the variables of tracks and clusters to distinguish the electrons from pions or jets. Since the electromagnetic showers deposit most of their energy in the second sampling of the electromagnetic calorimeter, the variables, hadronic leakage, the ratio of energy in 3×7 box to energy in 7×7 box ($R_\eta = \frac{E_{237}}{E_{277}}$) and the energy deposit width $w_{\eta 2}$ in the second sampling are efficient to reject jets of high energy and wide showers. Photon conversions are further reduced by requiring a good quality track pointing to an electromagnetic cluster with good energy-momentum match.

After the electron candidates are selected, an "isem" flag is used to define whether any of the electron identification cuts has passed or not. Basically there are three levels "loose", "medium" and "tight" of "isem", where "medium" is used on the first stage of our analysis. Once in the future, the variables which are used to define the "isem" flag are better understood, "tight" electrons are more favorable. Figure 3.8 displays the transverse energy and pseudo rapidity spectra for the "medium" electron candidates with comparison between 7 TeV data and Monte Carlo. Generally, shower shape variables of the second calorimeter layer and hadronic leakage variables are used in the "loose" electron selection. First calorimeter layer cuts, track quality requirements and track-cluster matching are added at the level of the "medium" electron selection. The "tight" electron selection adds E/p ratio, b-layer hit requirements and the particle identification potential of the TRT. All those variables used in the "loose", "medium" and "tight" selections are listed in Table 3.4

More details about the identification cuts applied on the selected electron candidates could be found in [34]. The Figure shows agreement between data and simulation, while the uncertainties would get larger if lower energy electrons are considered.

Since the MC is found to be mismodeling the R_η and $w_{\eta 2}$ distributions by investigating the data MC comparison, "isem" is redefined where R_η and $w_{\eta 2}$ cuts are loosened. The original definition of R_η and $w_{\eta 2}$ can be found in Table 3.4. Then, the electrons satisfy the loose cuts with the new R_η and $w_{\eta 2}$ cut values are called "RobustLoose". Consequently, there are "RobustMedium", which is used in our analysis, and "RobustTight" electrons.

An electron selected in our analysis is asked to pass the requirements of "RobustMedium".

A_{mc}	Description	Variable names
	Loose cuts	
Acceptance of the detector	• $\|\eta\| < 2.47$	
Hadronic leakage	• Ratio of ET in the first layer of the hadronic calorimeter to ET of the EM cluster (used over the range $\|\eta\| < 0.8$ and $\eta > 1.37$)	R_{had1}
	• Ratio of ET in the hadronic calorimeter to ET of the EM cluster (used over the range $\|\eta\| > 0.8$ and $\eta < 1.37$)	R_{had}
Second layer of EM calorimeter	• Ratio in η of cell energies in 3×7 versus 7×7 cells.	R_{η}
	• Lateral width of the shower.	$w_{\eta 2}$
	Medium cuts (includes Loose)	
First layer of EM calorimeter.	• Total shower width.	w_{stot}
	• Ratio of the energy difference associated with the largest and second largest energy deposit over the sum of these energies	E_{ratio}
Track quality	• Number of hits in the pixel detector (≥ 1).	
	• Number of hits in the pixels and SCT (≥ 7).	
	• Transverse impact parameter (<5 mm).	d0
Track matching	• $\Delta\eta$ between the cluster and the track (< 0.01).	$\Delta\eta_1$
	Tight cuts (includes Medium)	
b-layer	• Number of hits in the b-layer (≥ 1).	
Track matching	• $\Delta\phi$ between the cluster and the track (< 0.02).	$\Delta\phi_2$
	• Ratio of the cluster energy to the track momentum	E/p
	• Tighter $\Delta\eta$ cut (< 0.005)	$\Delta\eta_1$
Track quality	• Tighter transverse impact parameter cut (< 1 mm).	d_0
TRT	• Total number of hits in the TRT.	
	• Ratio of the number of high-threshold hits to the total number of hits in the TRT.	
Conversions	• Electron candidates matching to reconstructed photon conversions are rejected	

Table 3.4: Definition of variables used for “loose”, “medium” and “tight” electron identification cuts for the central region of the detector ($|\eta| < 2.47$)

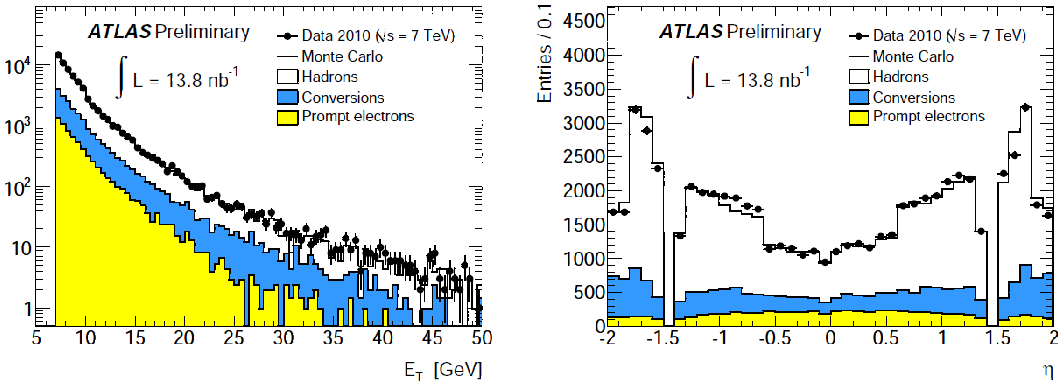


Figure 3.8: Distributions of cluster transverse energy E_T (left), and pseudorapidity η (right), for electron candidates passing the "medium" identification cuts.

Besides, it is also required to have an offline $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and be isolated by requiring the transverse energy in a cone of size ΔR =0.2 to be less than C1 + C2×p_{T} with C1=4 GeV, C2=0.023. Because of a worse electron reconstruction performance in the barrel endcap overlap region, the electrons in $1.37 < |\eta| < 1.52$ are excluded, and only the electrons in $|\eta| < 2.47$ are considered. In order to reject the electrons from photon-conversion, it requires at least one hit in the innermost Pixel detector layer (known as the b-layer) and the ratio of E/p to be close to unit, where E is the cluster energy and p is the track momentum. E and p are asked to be associated with the reconstructed electron. The electrons matching all the criteria above are stated as good electrons and are consequently selected in the analysis.

3.3.2 Reconstruction of Muons

A variety of strategies for muon identification and reconstruction are applied in ATLAS. The method of finding tracks in the muon spectrometer and extrapolating them to the beam line afterwards is called a "standalone" method. Combined muons are built after matching standalone muons to nearby inner detector tracks, where the measurements from the two systems are used. In a reverse order, "tagged" muons are found when tracks in the ID are extrapolated to the nearby hits in the spectrometer.

A standalone muon is found first by connecting the segments in each of the three muon stations to form tracks. The algorithm of extrapolating the spectrometer tracks to the beam line used by Staco is called "Muonboy", while the one used by "Muid" as the first stage is

"Moore". Multiple scattering and energy loss in the calorimeter should be considered when the extrapolation is performed. Standalone algorithms have slightly greater η coverage of $| \eta | < 2.7$, but have holes near $\eta = 0.0$ and $| \eta | = 1.2$.

The tagged muons are found by propagating the ID tracks with sufficient energy to the muon spectrometer looking for nearby hits. The tagging algorithms used for this purpose are MuTag and MuGirl. MuTag uses the difference between the predicted extrapolated track and its nearby segment to define a χ^2, while an artificial neural network multi-variable tool is used for MuGirl. Then the ID track is tagged as corresponding to a muon if it is close enough to a segment. A significant difference is that MuGirl consider all the segments around the ID track, while MuTag only uses the segments left by Staco.

A strategy to merge the standalone muons with the combined and/or tagged muons is useful to increase the muon reconstruction efficiency. An overlap removal between the algorithms is applied during the data recording. For instance, in the cases that a standalone muon is combined with more than one ID track, a best matched combination is selected among them. The muon reconstruction efficiency for "Staco" and "Muid" algorithms are presented in Figure 3.9. The plots show that the two algorithms are compatibly good with high muon reconstruction efficiencies. More details can be found in the Atlas muon reconstruction conference note [35].

Since the muon spectrometer is the outermost subdetector, muons traverse the ID and the calorimeters and leave an average amount of energy. The energy loss could be used to obtain better performance for muon reconstruction and identification.

Muons reconstructed by Muid with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $| \eta | < 2.5$ is used in this analysis, where a "tight" flag is also required ("tight" implies MuidCombined OR MuidStandalone $| \eta | > 2.5$ OR MuGirl with extended track). A muon also has to be isolated with both transverse momentum and transverse energy within a 0.3 cone (ptcone30 and etcone30) less than 4 GeV and not overlapped with any jet $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ at a distance $\Delta R < 0.4$. A muon selected matching all the requirements above is defined as a good muon and finally used in the analysis.

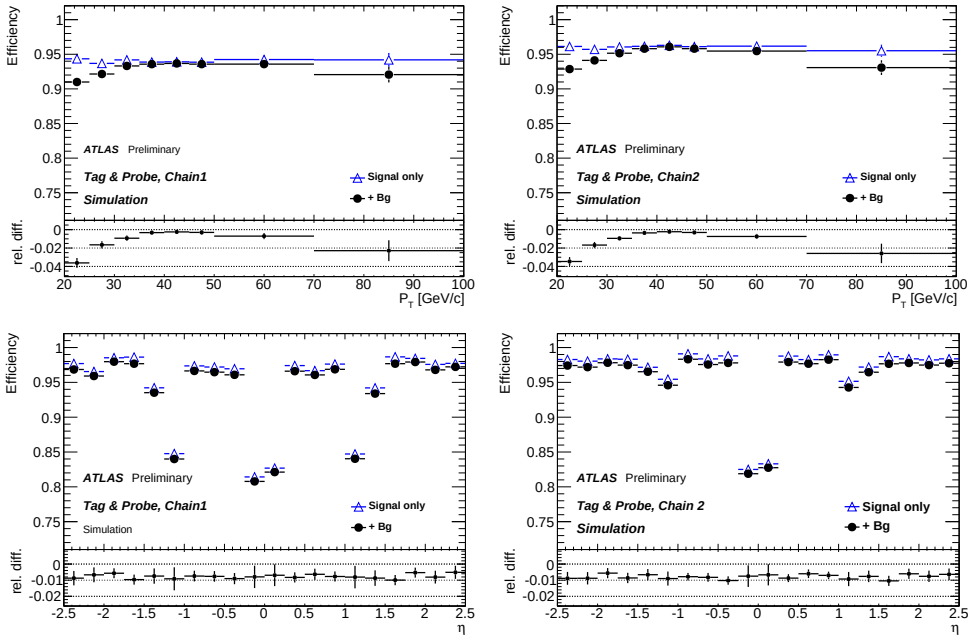


Figure 3.9: Measured muon reconstruction efficiencies for combined muons from Monte Carlo simulated data for "Staco" (left) and "Muid" (right) as a function of muons' p_{T} (up) and η (down). In the plots, triangles correspond to muon signal only and black dots correspond to signal plus all studied background contributions.

3.3.3 Reconstruction of Jets

The calorimeter system including the EM calorimeter and the hadronic calorimeters are used as the principal detectors for jet reconstruction. In principle, the pseudorapidity coverage of the detectors is near hermetic, $\mid \eta \mid < 4.9$. Several jet finding algorithms in ATLAS are provided to meet different purposes of physics analysis. This is in general due to the hadronic final states in all of the topologies interested, for example, jets in the inclusive QCD processes are wider spreads jets from the hard scattered parton, while W boson hadronic decay products are more likely to be narrow jets.

In general, the jet finding algorithms include fixed sized cone algorithms, sequential recombination algorithms and an algorithm based on event shape analysis. A common feature of the mentioned jet finder implementations is the full four-momentum recombination whenever the constituents of a jet change.

There is no universal way to reconstruct a jet, while the most commonly used jet finder implementations are the seeded fixed cone finder and the k_T algorithm which will be described below. The *anti-*k_T algorithm used in the analysis is based on the k_T algorithm.

The default implementation of a sequential recombination jet finder in ATLAS is the k_T algorithm. The p_{T}^2 of the objects (partons, particles, reconstructed detector objects with four-momentum representation) are defined as $d_i = p_{T,i}^2$. All pairs ij of input objects are defined to be their relative transverse momentum squared $d_{ij} = \min(p_{T,i}^2, p_{T,j}^2)\frac{\Delta R_{ij}^2}{R^2}$. The k_T algorithm parameter R gives a control of a jet size, where $R = 0.4$ for narrow jets and $R = 0.6$ for wide jets. The minimum p_{T}^2 are chosen from all the d_{ij} and d_i list as d_{min}. If d_{min} is a d_{ij}, the corresponding objects i and j are combined into a new object k using four-momentum recombination. The objects i and j are removed afterwards and k is added to the list. If d_{min} is a d_i, the object i is considered to be a jet and removed from the list. The objects list are finally becoming a list of all jets. The procedure is infrared safe and no object is shared by the jets formed.

The k_T algorithm is used and extended to a new jet finder tool named *anti-*k_T, which has a distance definition in the following formula[36].

$$d_{ij} = min(p_{T,i}^{2P}, p_{T,j}^{2P})\frac{\Delta R_{ij}^2}{R^2}, \quad d_i = p_{T,i}^{2P} \tag{3.1}$$

The inclusive k_T algorithm is recovered if one takes $P = 1$. It can be argued that the jet algorithm with respect to soft radiation is rather similar with any value of $P > 0$. It could be seen that the negative values of P is not pathological and the behavior is quite similar for all $P < 0$, where the *anti-k_T* jet-clustering algorithm is defined with P = -1. Generally, if we consider an event with several hard separated objects with $p_{T,1}$, $p_{T,2}$, ... It is easy to find that the d_{1i} of a hard particle 1 and a soft particle i is exclusively determined by the p_{T} of the hard particle, where $d_{1i} = \min(\frac{1}{p_{T,1}^2}, \frac{1}{p_{T,i}^2}) \frac{\Delta R_{1i}^2}{R^2}$. Then d_{1i} would be smaller than the d_{ij} from two similarly separated soft objects. As a consequence, the soft objects would be absorbed by the hard ones before they are combined themselves. The algorithm favors to have a hard jets well built conically with a radius R and the soft ones have more complex shapes.

Jet reconstruction from calorimeter towers starts with a re-summation step, where negative signal towers with $E_{tower} < 0$ are combined with the nearby positive ones $E_{tower} > 0$. This reduces the noise fluctuations and signal biases. The negative towers without any positive signals nearby are simply dropped. The remaining towers with physically valid four-momenta are the inputs to the jet finding algorithms. The outputs of the jet finder at this stage are then jets with energies on the *e.m.* energy scale. After the treatment of calibration, jets with $p_{\mathrm{T}} > 7\,\mathrm{GeV}$ are kept, otherwise they are discarded. Some refined corrections are applied to calibrate the tower jets to the particle level, which could be addressed in a specific physics analysis. The corrections of jet energy include the non-linearity in the jet response (such as missing energy from the jet or adding energy not belonging to the jet) and suppression of signal contributions from the underlying event and/or pile-up.

Besides the jet reconstruction in calorimeters, the jet finding in the ID tracks is also useful, which could improve the reconstruction efficiency and the quality of the jet parameters. The p_{T} fraction of the calorimeter jet and the corresponding track jet, $f_{trk} = \frac{p_{T,trk}}{p_{T,calo}}$, could be used to refine the jet energy measurement. ID tracks could also provide a reconstructed vertex associated with the jet characterized by a parameter named z_{vtx}.

The *anti-k_T* 0.4 TopoCluster jets with EM+JES calibration is produced by jet reclustering and is used in the analysis presented here. The jet close to a selected good electron ($\Delta R < 0.2$) is removed from the jet list. The JES uncertainty is estimated based

on Monte-Carlo studies [37].

In addition, the jets not associated to in-time real energy deposits in the calorimeters are tagged as "bad" jets. Several variables concerning the jet energy deposition fraction in different part of the detectors, jet timing and jet quality are used to identify if the jet is "bad". "Bad" jets arise from various sources, like hardware problems, LHC beam conditions and cosmic ray showers. An event containing any number of "bad" jets will be removed.

3.3.4 The Measurement of $E_{\mathrm{T}}^{\mathrm{miss}}$

A good measurement of $E_{\mathrm{T}}^{\mathrm{miss}}$ can highly improve the ability of some physics analysis measurement, such as the $t\bar{t}$ events with one or two top decaying leptonically. It is also important in Higgs hunting and new physics like supersymmetry and extra dimensions. Some aspects of the detectors would have impact on the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement, such as the limited detector coverage, the finite detector resolution, the presence of dead regions and different sources of noise. High energy particles escaping from the forward direction are limited by the large coverage of the calorimeters. Even so, the inactive transition regions between the calorimeters and those dead and noisy readout channels would produce some fake $E_{\mathrm{T}}^{\mathrm{miss}}$.

There are two $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction algorithms used in ATLAS, Cell-based and Object-based. The $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction is essentially done with energy deposits in the calorimeter (classified as jets and electrons, and calibrated separately) and reconstructed muon tracks.

The noise suppression in the calorimeters is a common step for both algorithms and is crucial for $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement. The standard noise suppression method is based on the cells whose energy is over a symmetric threshold, $\mid E_{cell} \mid > 2\times\sigma_{noise,cell}$. The second noise suppression method only uses cells in 3D topological calorimeter clusters (TopoClusters) combined with the cells called TopoCells, which are described in section 3.3.3. These chosen thresholds as S=4, N=2, P=0, are also optimized to suppress electronic noises as well as pile-up events.

The Cell-based $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction includes contributions from three terms, referred to as calorimeter, cryostat and muon terms:

$$E_{x,y}^{\mathrm{miss},Final} = E_{x,y}^{\mathrm{miss},Calo} + E_{x,y}^{\mathrm{miss},Cryo} + E_{x,y}^{\mathrm{miss},Muon} \tag{3.2}$$

The x and y components of the calorimeter term are calculated from the sum of the transverse energies in TopoCells, $E_{x,y}^{\mathrm{miss},Calo} = -\sum_{TopoCells} E_{x,y}$. The straightforward operation will give a large shift in the $E_{\mathrm{T}}^{\mathrm{miss}}$ scale of about 30% with respect to the true $E_{\mathrm{T}}^{\mathrm{miss}}$. Because of the truth that electromagnetic showers tend to have higher energy density compared to hadronic showers, energy deposits of TopoClusters are classified using further information related to shape and depth of the shower. Then some calibration schemes are applied to calibrate globally all the calorimeter cells. This can improve the linearity and the resolution of the measurement.

Hadronic showers will lose energy in the cryostat between the LAr barrel EM calorimeter and the tile hadronic calorimeter. The energy loss in the cryostat is used to $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction as a croystat term, $E_{x,y}^{\mathrm{miss},Cryo} = -\sum_{RecJets} E_{x,y}^{Cryo}$, where the jet energy in the third layer of the EM calorimeter and in the first layer of the hadronic calorimeter is used to calculate E^{Cryo}.

The muon term is calculated from the muons measured in a large range of $\mid \eta \mid < 2.7$, $E_{x,y}^{\mathrm{miss},Muon} = -\sum_{RecMuons} E_{x,y}$. High quality muons with a matching track in the ID ($\mid \eta \mid < 2.5$) can help to reduce the the $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution. Because of the acceptance of the muon spectrometer, a loss of muons will affect the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement.

Finally, the calibration is refined when the cells are associated with one or more good reconstructed objects. The association is done in a chosen order: electrons, photons, muons, hadronically decaying τ-leptons, b-jets and light jets. This improves the $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction as it is known that the objects are better calibrated to a higher accuracy. The performance of $E_{\mathrm{T}}^{\mathrm{miss}}$ with early 7 TeV data after the refined calibration based on energy corrections of physics objects is presented in Figure 3.10[38]. And Figure 3.11 displays the resolution as a function of the final $\sum E_T$.

The object based $E_{\mathrm{T}}^{\mathrm{miss}}$ reconstruction are designed to be sensitive to the low p_{T} objects like τ jets, soft jets, etc. The objects are first classified as high p_{T} and low p_{T} objects. The calorimeters are mostly used for the method to reconstruct $E_{\mathrm{T}}^{\mathrm{miss}}$ while track information in ID is also considered for objects like electrons, muons and soft jets. TopoClusters are used and all contributions of high p_{T} objects are first included in the calculation. TopoCells not part of any high p_{T} objects are classified as low p_{T} deposit.

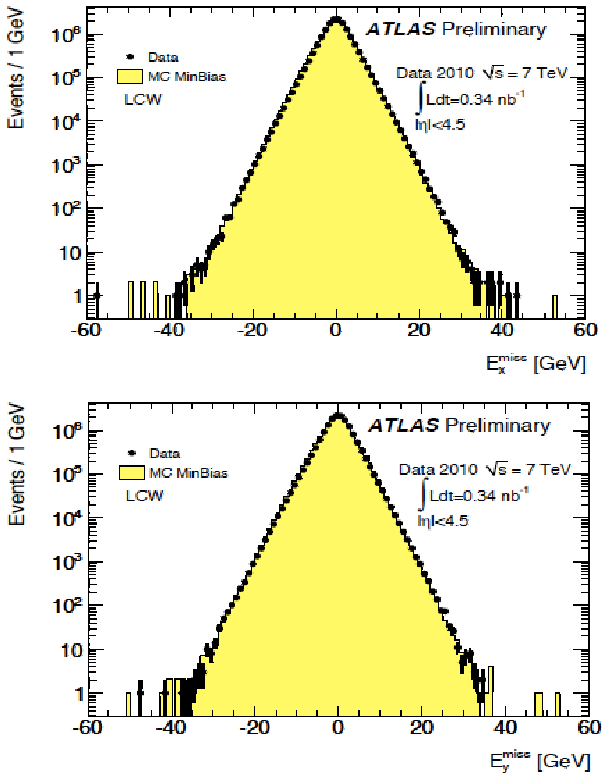


Figure 3.10: Comparison of minimum bias events for 7 TeV data and MC for distributions of E_x^{miss} (left) and E_y^{miss} (right)

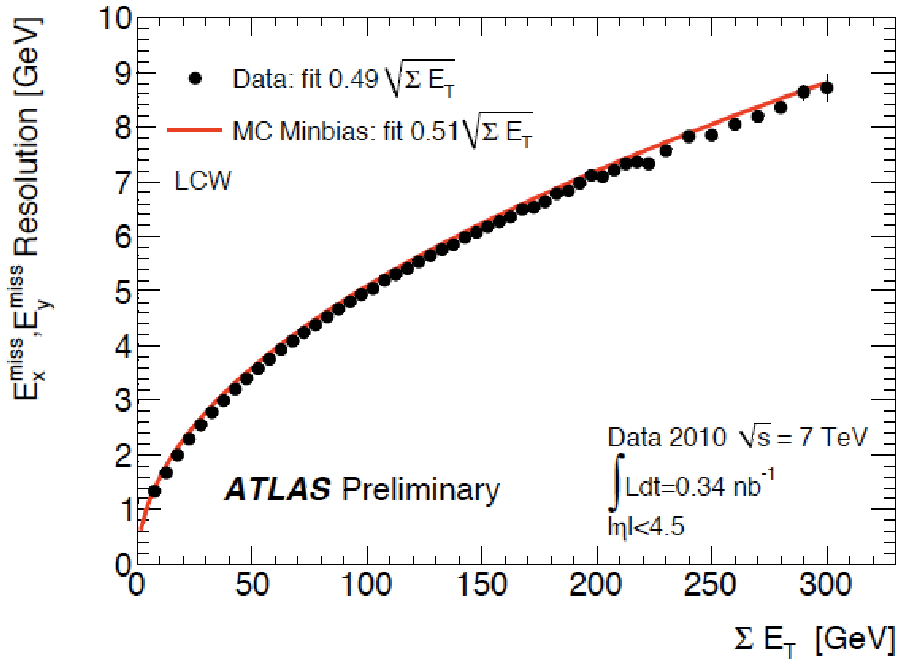


Figure 3.11: The $E_{\mathrm{T}}^{\mathrm{miss}}$ resolution distribution as a function $E_{\mathrm{T}}^{\mathrm{miss}}$ with 7 TeV data and simulation

3.3.5 The SV0 b-tagging algorithm

Since a top quark almost inclusively decays to a b-quark and a W boson, the identification of jets originating from b-quarks is very important in top quark related physics. The jet b-tagging is of great use to suppress the background processes containing large amount of light flavor jets.

The SV0 tagging algorithm is a lifetime-based b-tagger, which requires the reconstruction of the secondary vertices from tracks associated with a jet. The operation of this tagging algorithm involves placing a cut on the signed decay length significance, $L/\sigma(L)$, of the reconstructed secondary vertex, which is further referred as SV0Weight. The sign of $L/\sigma(L)$ is given by the sign of the projection of the decay length vector on the jet axis illustrated in Figure 3.12.

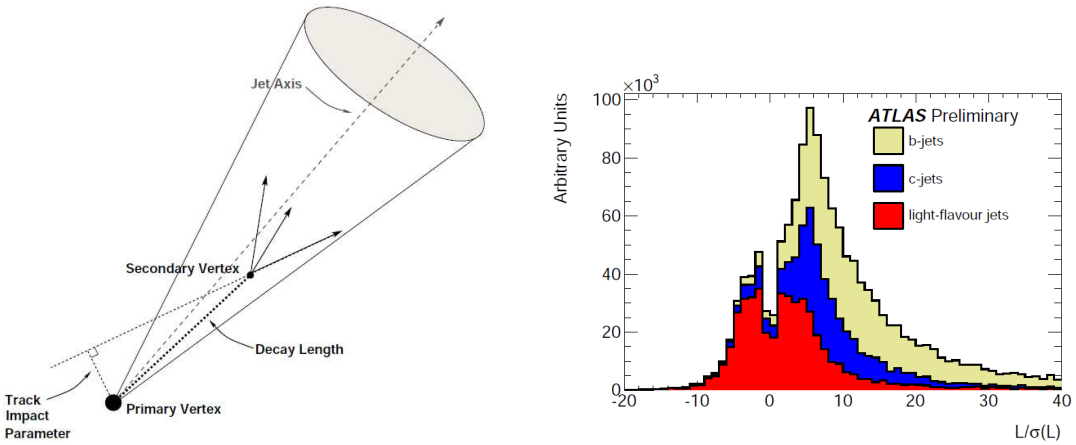


Figure 3.12: (left) A secondary vertex with a significant decay length indicates the presence of a long-lived particle in the jet. The secondary vertex is reconstructed from tracks with a large impact parameter significance with respect to the primary vertex. (right) The signed decay length significance $L/\sigma(L)$ for the SV0 b-tagging algorithm (known as SV0Weight) in simulation. The distribution extends to much larger values for jets originating from b-quarks compared to those originating from c-quarks, light quarks or gluons.

The SV0 tagging algorithm is first given a list of tracks associated to the calorimeter jet. Those tracks which fulfill certain quality criteria are used in the secondary vertex

fit. Secondary vertices reconstruction starts from a common vertex merged with two-track vertices. Tracks giving large χ^2 to the vertex are removed until the vertex survives certain quality criteria. A more detailed description can be found in [39] for the SV0 algorithm.

In this thesis, a jet is defined as b-tagged if $L/\sigma(L)$, therefore the SV0Weight, is over a threshold of 5.72, an operating point that yields a 50% b-tagging efficiency in simulated $t\bar{t}$ events with jets $p_{\mathrm{T}} > 15\,\mathrm{GeV}$. The SV0Weight distributions can be seen in the right plot of Figure 3.12 for b-jet, c-jet and light jets with simulated QCD multijet jet sample generated with Pythia.

3.3.6 The Measurement of Luminosity

A reliable luminosity measurement for high and low luminosity running is essential for cross section measurement in ATLAS. A number of complementary methods are used to determine the absolute and relative luminosity. The relative measurements are not normalized overall, and therefore need to be calibrated by measurements in the corresponding conditions. When the calibration is done, the relative methods can be used in different cases, where it is not suitable for the absolute methods. The precise absolute luminosity may be hard to obtain. It is good to notice that the relative measurements can always be re-normalized to a more precise absolute result at a later stage.

The instantaneous luminosity can be calculated in the following formula,

$$L = \frac{\mu \cdot n_b \cdot f_r}{\sigma_{inel}} = \frac{\mu^{meas} \cdot n_b \cdot f_r}{\epsilon \cdot \sigma_{inel}} = \frac{\mu^{meas} \cdot n_b \cdot f_r}{\sigma_{vis}} \tag{3.3}$$

Where,

- μ: number of inelastic proton-proton collisions per bunch crossing (BC)
- n_b: number of bunch pairs colliding at the interaction point (IP)
- f_r: LHC revolution frequency ($11245.5Hz$)
- σ_{inel}: total inelastic proton-proton collision cross section
- μ^{meas}: measured average number of events per BC
- ϵ: efficiency of the luminosity algorithm including the acceptance for a certain detector
- σ_{vis}: visible cross section as a detector calibration constant

The equation is only valid with the assumption of a linear response of the detector with respect to μ, otherwise the correction of the non linearity should be taken into account.

Some general principles of the absolute and relative luminosity measurements are reflected and summarized in the following items, the effects of which would be known in the LHC running data.

Absolute luminosity measurement methods:

- Roman Pots: One pair of Roman Pot stations are mounted at a distance of 240 m on each side of the IP and used to determine the absolute luminosity primarily from elastic proton scattering in the Coulomb interference region. This measurement will be alternative at low luminosity. The Roman Pot detectors are proposed to provide level 1 trigger information and data to the ReadOut System (ROS) for triggered events. The resolution of this method is expected to be 2∼3%.

- W/Z Counting: A precise knowledge of the cross-section for W/Z production in the leptonic decay channel is assumed and consequently used to calculate the absolute luminosity. The measurement will be determined to about 10% when the PDF of proton uncertainties and detector effects are considered. Thanks to the high rate of W/Z production, the statistical precision of online relative luminosity monitoring can reach 1% at high luminosity ($10^{34} cm^{-2} s^{-1}$).

- $\mu\mu/ee$ Counting:The cross section of the electromagnetic process of lepton pair production from two photon fusion could be calculable to a precise level. However, the cross-section is quite small. The method will require significant analysis work and backgrounds need to be well understood and controlled.

Relative luminosity measurement methods:

- LUCID: It is a Cerenkov detector which consists of 168 tubes filled with C_4F_{10} gas and mounted at a distance of 17m on each side of the IP. The detectors are designed

to have good time resolution to identify individual BCs. LUCID approximately covers $5.5 < | \eta | < 6.1$ and monitors luminosity by inelastic events at a bunch by bunch level. It provides level 1 trigger information and data to the ROS.

- Beam Condition Monitor (BCM): It consists of four 1 cm×1 cm×500μm diamond sensors mounted at a distance of 1.8 m on each side of the IP. The BCM covering $5.5 < | \eta | < 6.1$, provides a sufficient time resolution and good sensitivity to particle transversing. The luminosity monitoring is based on inelastic events and the system is well suited for the whole luminosity range at the LHC.

- The Minimum Bias Trigger Scintillator counters (MBTS): The MBTS counters consist of one plane with 2×8 scintillator segments mounted in front of the LAr end-cap. The MBTS will only be used during the commissioning phase of ATLAS to determine the luminosity by counting the minimum bias trigger rate. The η range covered is around $1.9 < | \eta | < 3.8$ and inelastic collisions are monitored. The MBTS will experience significant radiation damage during the LHC data taking, thus it is hard to provide a precise measurement of the luminosity.

- Tile Calorimeter: The hadronic tile calorimeter has a minimum bias monitor system which is best suited for the luminosity monitoring at high luminosity. It will provide relative luminosity information from the local monitor system outside the event stream.

- LAr Calorimeter: The LAr calorimeters have the possibility to measure the relative luminosity by monitoring the high voltage current in the LAr system.

Measurement of μ^{meas}

Beyond the short description above, a more detailed introduction on the measurement of μ^{meas} is presented in the following.

The MBTS is used to determine the luminosity item μ^{meas} both online and offline. Online measurements are based on the information available to the Central Trigger Pro-

cessor (CTP) before any L1 trigger decision is taken. The online calculation is done by central trigger scalars, which count the number of events firing any given L1 item. The counting time for each scalar is recorded by reporting the number of LHC turns. Every 10 seconds, the counters are read out and made available on the information service (IS), where they are retrieved by the online luminosity calculator (OLC). The MBTS trigger with the lowest background rate, L1_MBTS_1_1_paired, is most reliable for luminosity monitoring, because the online luminosity calculation does not perform a background subtraction. L1_MBTS_1_1_paired means at least one hit in both A side and C side of the MBTS in a paired bunch trigger.

The offline event selection (MBTS_1_timing) uses the timing information in the MBTS to remove background. Figure 3.13 shows the time difference between the MBTS side A and side C ($\Delta t_{A,C}$) for the 900 GeV data and simulation. Data and simulation are normalized to the same number of events within $\mid \Delta t_{A,C} \mid < 10$ ns. The peak at $\Delta t_{A,C} = 0$ is attributed to the collision at the center of ATLAS, which is well reproduced by simulation. The secondary peaks at ± 24 ns correspond to a round-trip distance of about 7.2 m. They are consistent with particles coming from upstream of ATLAS from beam-halo and beam-gas interactions, which are not included by the Monte Carlo. The estimated efficiency of the offline selection $\mid \Delta t_{A,C} \mid < 10$ns cut would be around 0.986. The variation in selection efficiency by changing the cut ± 2ns is about 1%. It would be smaller than 10^{-3} when the center of the distribution of $\Delta t_{A,C}$ is changed by ± 0.5ns. For the 7 TeV running, the backgrounds after the timing cuts are very small, which would be less than 10^{-4}. A further background subtraction is done by using the unpaired bunches, which produce similar peaks at ± 24 ns.

LUCID values for online luminosity measurements are obtained from a purpose built electronics card (LUMAT), which are programmed with different luminosity algorithms. At present there are four algorithms implemented in the LUMAT firmware:

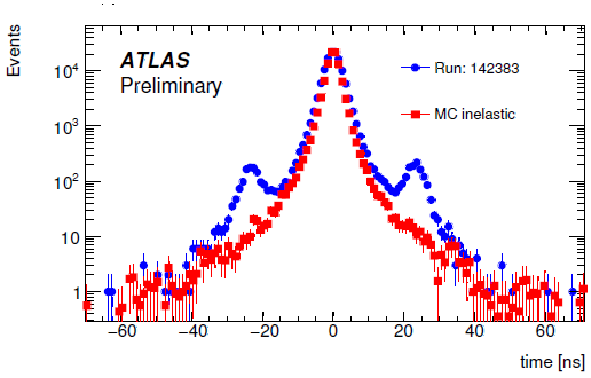


Figure 3.13: Time difference for L1_MBTS_1_1_paired selected events in 900 GeV run 142383 compared with MC simulation of inelastic processes.

- LUCID_Zero_AND: the number of events per BCID when no hits are found in either detector arm
- LUCID_Zero_OR: the number of events per BCID when one of the two or both detector arms have no hits
- LUCID_Hit_AND: the number of hits per BCID when at least one hit in each detector arm
- LUCID_Hit_OR: the number of hits per BCID when at least one hit in both detector arms

The counting of empty events was first used, while it was later converted to counts of non-empty events simply because this is conceptually easier to understand. It could be obtained by subtracting the number of empty events, which could be expressed in the following:

- LUCID_Event_AND: the number of events per BCID when at least one hit in each detector arm, the probability per beam crossing is $P^{LUCID\_Event\_AND} = 1$ - $P^{LUCID\_Zero\_OR}$

- LUCID_Event_OR: the number of events per BCID when at least one hit in both detector arms, the probability per beam crossing is $P^{LUCID\_Event\_OR} = 1$ - $P^{LUCID\_Zero\_AND}$

As long as the number of interactions per bunch crossing (μ^{meas}) is small, the event counting methods are more precise than hit counting methods. Therefore event counting has first

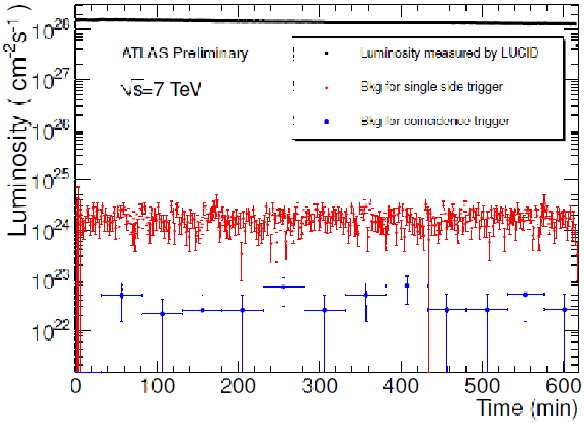


Figure 3.14: The luminosity measured by LUCID at 7 TeV obtained from paired bunches. The background has been calculated from unpaired bunches.

been used to calculate the luminosity. The hit counting methods will be more important when the luminosity increases. LUCID_Event_OR counting is preferred when background is small, since it gives a smaller statistical uncertainty and a less complicated dependence on the number of interactions. Similar to the case in MBTS measurement, the background is evaluated using the unpaired bunches in order to account for beam-related effects. Figure 3.14 presents the results of a typical run at 7 TeV, where the background level is for the single sided LUCID_Event_OR ($< 10^{-4}$) and the coincidence LUCID_Event_AND ($< 10^{-5}$) triggers. Uncertainty contributions from cosmic rays and electronic noise have been measured in the absence of beam and are negligible when only the colliding BCIDs are selected.

The efficiency for the LUCID_Event_OR method is used, since the average hit-multiplicity is high enough for the LUCID simulation to provide an accurate estimation. The result of the efficiency calibration is used to compare LUCID with measurements by other detectors. In addition, LUCID calibrations are done using MBTS, because it provides a higher average hit-multiplicity.

The LAr endcap calorimeters are also used for the measurement of luminosity which is performed by analyzing minimum bias data offline. Events are required to pass the L1_MBTS_1_paired trigger and the energy deposits in the EMEC Inner Wheels and the FCAL are measured, corresponding to the pseudo rapidity range $2.5 < | \eta | < 4.9$. Cells are required to have an energy 5σ above the noise level and two cells passing these requirements

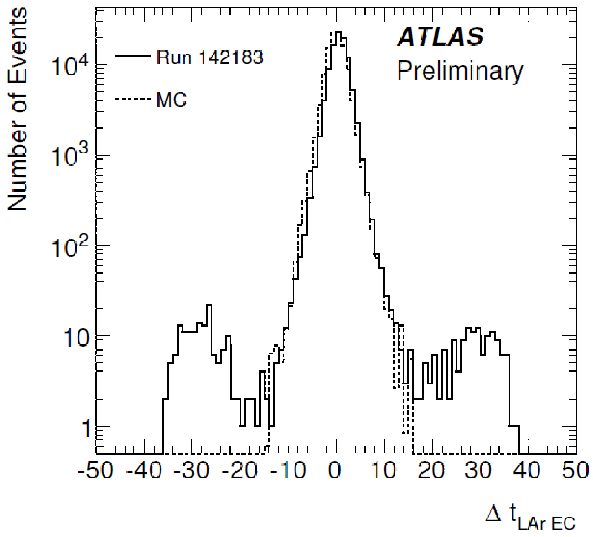


Figure 3.15: The difference in ns between the arrival time of the signal in the LAr endcaps A side and C side

are required on both the A side and the C side. An example of the distribution of the LAr endcap timing difference is shown in Figure 3.15. A clear peak at $\Delta t_{LArEC} = 0$ is from particles in the collision point, while the secondary peaks at ±30ns correspond to particles coming from upstream of ATLAS, a beam-gas interaction for example. A cut, $\Delta t_{LArEC} <$ 5ns is applied and the background remains after the cut can be estimated from the activity in the non-colliding bunches. At 7 TeV collision, the instantaneous luminosity is high and the background level is of order 10^{-4}.

Measurement of σ_{vis}

The total "visible" cross section are divided into individual inelastic processes contributions: non-diffractive (ND), single-diffractive (SD), and double-diffractive (DD). Therefore the luminosity and the cross section could be expressed:

$$L = \frac{\mu^{meas} \cdot n_b \cdot f_r}{\sigma_{vis}} = \frac{\mu^{meas} \cdot n_b \cdot f_r}{\epsilon_{ND} \cdot \sigma_{ND} + \epsilon_{SD} \cdot \sigma_{SD} + \epsilon_{DD} \cdot \sigma_{DD}} \tag{3.4}$$

Before the absolute luminosity measurement is available, for example the ALFA detector will provide an absolute luminosity calibration through the measurement of elastic pp-scattering at small angles, Monte Carlo based calibration relying on the modeling of the inelastic cross section will incur a significant systematic uncertainty. The uncertainty can be obtained

Process	PYTHIA	PHOJET
Non-Difractive	48.5	61.6
Single-Difractive	13.7	10.7
Double-Difractive	9.3	3.9
Total	71.5	76.2

Table 3.5: Predicted inelastic pp cross sections in mb at 7 TeV for PYTHIA and PHOJET

ϵ(%)	MBTS_1_timing		LUCID_Event_OR		LUCID_Event_AND		LAr Endcap	
Process	PYTHIA	PHOJET	PYTHIA	PHOJET	PYTHIA	PHOJET	PYTHIA	PHOJET
ND	97.4	97.9	79.2	74.2	30.8	25.5	96.0	94.3
SD	41.3	44.3	28.7	44.8	1.2	2.4	21.4	27.9
DD	50.8	68.1	39.4	62.0	4.4	14.8	25.9	53.6
σ_{vis} (mb)	57.6	67.8	46.1	52.9	15.5	16.4	51.9	63.2

Table 3.6: Efficiency at 7 TeV for the various luminosity methods

by comparing the cross sections and acceptances with two Monte Carlos: PYTHIA and PHOJET. Table 3.5 shows the PYTHIA and PHOJET predictions for the inelastic cross sections at 7 TeV, where the predicted cross section in PHOJET is higher than PYTHIA by 6.5%. Table 3.6 shows the predicted efficiencies for observing ND, SD and DD events for the different methods introduced in the previous section.

The instantaneous luminosity calculated with independent calibrations is shown in Figure 3.16 for one 7 TeV run. The multiplicity in LUCID is high enough to allow Monte Carlo calibration. The MBTS and Liquid Argon results are consistent to better than 1%. The results for LUCID_Event_OR are consistent with MBTS and Liquid Argon to 3%. The LAr instantaneous luminosity is corrected for the dead time in the data acquisition system, while the MBTS and the LUCID methods are not affected by data-acquisition dead time.

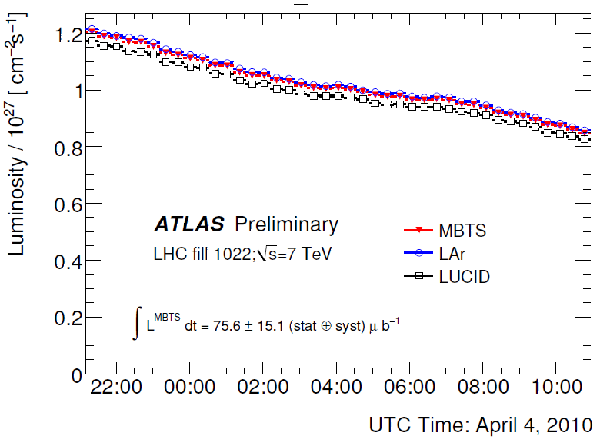


Figure 3.16: The instantaneous luminosity as determined with the LAr, the LUCID_Event_OR, and MBTS sub-detectors for run 152409.

As long as the visible cross section is determined by the Monte Carlo models of total pp inelastic processes, the systematic uncertainty on the luminosity will remain large. When much larger integrated luminosities are accumulated, it is possible to normalize cross-section measurements to electroweak processes for which precise NNLO calculations exist, for example W and Z production.

Absolute luminosity calibration with beam-separation scans

An alternative method to calibrate the absolute luminosity is to use the data collected with beam-separation scans (also called van der Meer (vdM) scans or luminosity scans). The beam scans has been carried out in Atlas, during which the collision rates are recorded by the luminosity detectors and beams are moved both horizontally and vertically in the transverse plane.

The luminosity in terms of colliding-beam parameters can be expressed as in Formula 3.5.

$$L = n_b \cdot f_r \cdot n_1 \cdot n_2 \cdot \int \hat{\rho}_1(x,y)\hat{\rho}_2(x,y)dxdy \tag{3.5}$$

where $\hat{\rho}_1(x,y)$ is the normalized particle density in the transverse plane of beam 1 at the IP as a function of (x and y), where n_1 is the number of particles per bunch. The parameters labeled 2 is the same for beam 2.

Under the assumption of no correlation between x and y direction, the particle densities can be re-written as $\hat{\rho}(x,y) = \rho(x)\rho(y)$. Therefore, the integration in the formula can be

replaced by Ω_x and Ω_y, defined as $\Omega_x = \int \rho_1(x)\rho_2(x)$ in x direction and an analogous definition in y direction. In order to measure these two parameters, one can introduce the method proposed by van der Meer [40] in 1968 under the formula 3.6:

$$\Omega_x(\rho_1(x), \rho_2(x)) = \frac{R_x(0)}{\int R_x(\delta)d\delta} \tag{3.6}$$

where $R_x(\delta)$ is the luminosity measured during a horizontal scan when the two beams are separated by a distance δ, which is, as a function of δ, then called as a luminosity curve. Defining a variable using unit of distance out of the luminosity curve, one have in x direction as:

$$\Sigma_x = \frac{1}{\sqrt{2\pi}} \frac{\int R_x(\delta)d\delta}{R_x(0)} \tag{3.7}$$

which leads to the luminosity formula expressed as machine parameters in Formula 3.8. The formula is quite general and Σ_x and Σ_y only depend on the area under the luminosity curve, which is in arbitrary units. With the scan data, Σ_x and Σ_y are measured and consequently gives a calibration results on luminosity.

$$L = \frac{n_b \cdot f_r \cdot n_1 \cdot n_2}{2\pi \cdot \Sigma_x \cdot \Sigma_y} \tag{3.8}$$

As well described in the first luminosity paper [41], three van der Meer scans have been performed at the Atlas interaction point. The procedure of obtaining the scan data is to span a range of $\pm 6\sigma_b$ in horizontal and vertical beam-separation, where σ_b is the nominal transverse size of either beam at the IP.

The full luminosity-calibration scan starts with a mini-scan to center the beams on each other in both directions on the transverse plane. It is activating closed orbit bumps, a local distortion of the beam orbit, to vary the IP positions of both beams by $\pm 1\sigma_b$ and then achieve best match of the two beams.

The full horizontal and vertical scans are carried out separately, where the orbit bumps are used to displace the two beams in opposite directions by $\pm 3\sigma_b$ (thus, $\pm 6\sigma_b$ relatively) first horizontally then vertically. Therefore each scan contains 27 steps, and each step lasts about 30 seconds to record the data. The instant luminosity and other relevant parameters are recorded and transmitted to Atlas.

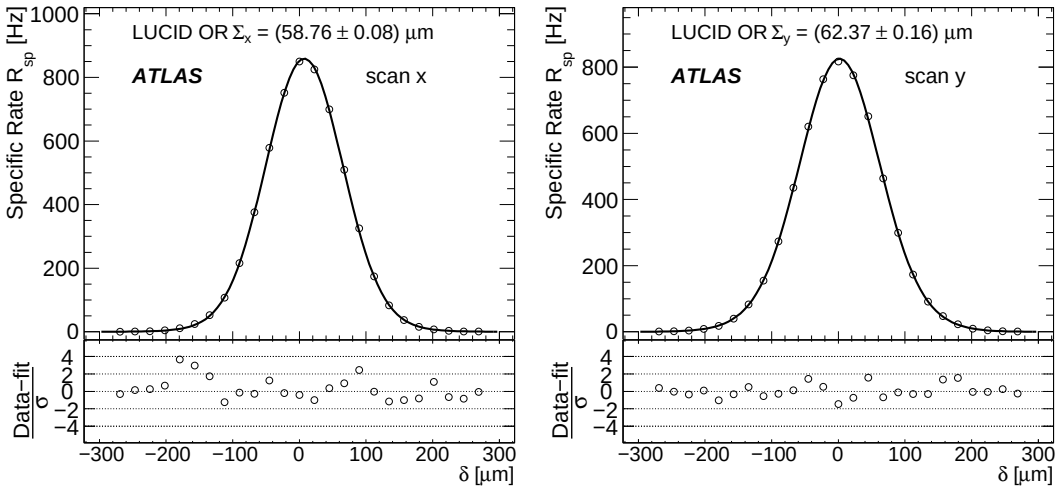


Figure 3.17: An example of the fitting results of a luminosity scan in the x (left) and y (right) direction for the LUCID_Event_OR algorithm. Other luminosity measurement algorithms give similar fitting results, which are not presented.

With the scan data, the Σ_x and Σ_y are extracted by fitting the relative variation of bunch luminosity as a function of the beam separation, as describe in 3.7. An example of the fitting using LUCID_Event_OR luminosity measurement algorithms is shown in Figure 3.17.

Combining with the measured bunch currents, the fitting results of Σ_x and Σ_y are used to determine the absolute luminosity expressed in Formula 3.8. The systematic uncertainties affecting the luminosity measurement arise from several sources. The main sources are the beam intensities (10%), the length-scale calibration (2%), the imperfectness of the beam centering (2%) and so on, which totally give an uncertainty of 11%. The luminosity used in this thesis is measured with this method of calibration and the uncertainties estimated is also applied correspondingly.

Chapter 4

TOP PAIR EVENT SELECTION

Top pair candidate events decaying to lepton+jets channels are selected using some common cuts in the ATLAS top group, which include exactly one isolated good lepton, large missing transverse energy and at least 4 jets. These cuts are designed with the help of the MC samples, which are first introduced in this chapter. The event generation, simulation and reconstruction algorithms are briefly described.

After that, the QCD multijet background measurement using data driven method is also included. The dominating background, W+jets, is measured with two data driven methods, which are introduced in Chapter 5. Other backgrounds are estimated from MC samples. The number of events yields after each selection cut is presented. Some control plots are drawn after the cuts to see performance and power.

This efficiency in MC is corrected for different lepton trigger, reconstruction and identification efficiencies between data and MC by using the scale factors (SF) measured in data. The effects of the correction are presented and final selection efficiencies are described in the last section of the chapter.

4.1 Top pair event selection cuts

The event topology for $t\bar{t}$ semi-leptonic decay channels is characterized by some strong features, for example, one well reconstructed and isolated high p_{T} lepton, at least 4 good jets and large missing transverse energy. The objects, electrons, muons, jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ used in this section are defined in Section 3.3. So, the event selection is done according to the characters of $t\bar{t}$ events, and the selection cuts are listed in the following:

1. Exactly one tight and isolated lepton (electron or muon) with $p_{\mathrm{T}} > 20\,\mathrm{GeV}$, no explicit veto is applied on events with leptons $10 < p_{\mathrm{T}} < 20\,\mathrm{GeV}$.

2. Passing trigger for electron or muon depending on the channel.

3. At least 5 good tracks from the primary vertex.

4. $E_{\mathrm{T}}^{\mathrm{miss}} > 20\,\mathrm{GeV}$.

5. Triangle cut, $E_{\mathrm{T}}^{\mathrm{miss}}$ plus the transverse mass (M_T) of the lepton and neutrino: $E_{\mathrm{T}}^{\mathrm{miss}} + M_T > 60\,\mathrm{GeV}$.

6. At least 4 good jets with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$.

Data Period	Electron	Data Period	Muon
A - D (To 159224)	L1_EM10	A - E3 (To 160879)	L1_MU10
E (To 161948)	EF_g17_etcut	E4 - F	EF_mu10_MSonly
F	EF_e10_medium	G1-G5	EF_mu13
G - I	EF_e15_medium	G6 - I	EF_mu13_tight

Table 4.1: Electron and muon triggers used in data periods

There should be one tight and isolated high p_{T} lepton, electron in EM Calorimeter or muon in the muon Spectrometer. Corresponding to the selected lepton, the event is required

to pass a trigger chain depending on the period of the collision. The details of the triggers chosen for the data are shown in Table 4.1.

The triggers used in data are the lowest unprescaled triggers. They depend on the data period because of the increase in instantaneous luminosity. In MC, the simulated triggers is chosen to be the one which is the closest to the trigger used for real data: EF_e15_loose in the electron channel and EF_mu13 in the muon channel. The effect of using slightly different trigger chains in MC and data is counted as a systematic, which is estimated to be less than 1%.

Table 4.2 and Table 4.3 list the numbers of events after each cut in the electron channel and in the muon channel. In both tables, the contributions from $t\bar{t}$, diboson, single top, W+jets and Z+jets are estimated in MC, which is normalized to 33.7 pb^{-1}. The QCD multijet background is measured using the "Matrix Method" detailed in Section 4.3. The W+jets number of events is also measured in data as shown in Chapter 5.

The number of events with different simulated top masses is listed in Table B.1 and Table B.2 in Appendix B. Since at a given collision energy the cross section of $t\bar{t}$ decreases with higher top mass, the measurement when compared to the prediction of $\sigma_{t\bar{t}}$ gives an idea of the cross check with top mass measurement.

The requirement of the number of tracks associated to the primary vertex is designed to reject the non-collision events like cosmic, elastic pp collision, radiation from the non ideal vacuum. The tight and isolated lepton cut as well as the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut is used to reduce the QCD multijet background. The QCD multijet events are estimated only after the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut, which is better measured with the triangle cut in one or more jets bins. The $E_{\mathrm{T}}^{\mathrm{miss}}$ cut and the triangle cut efficiently remove a large fraction of Z+jets background. There are close to 50% signal events surviving the number of jets cut, while only a few percent of W+jets background events pass the cut. The cut on at least 4 jets is applied to suppress W+jets background, where jets from initial or final state radiation have lower p_{T}. The detail about the performance of the selection cuts is shown in Section 4.4, while the MC samples used in the analysis is presented in Section 4.2.

channel	total	1 elec	trigger	nTrk≥5
$t\bar{t}$ NoFullHad (MC)	2947.1±0.6	585.3±0.3	580.2±0.3	579.0±0.3
diboson (MC)	824.4±0.2	200.3±0.1	198.5±0.1	197.9±0.1
single top (MC)	1262.3±1.2	197.0±0.5	195.4±0.5	194.5±0.5
W+jets (MC)	1072334.7±82.2	146419.8±30.3	144876.3±30.2	142179.0±29.9
Z+jets (MC)	104483.2±16.9	14782.7±6.4	14675.7±6.4	14505.1±6.4
Total Bkg (MC)	1178904.6±83.9	161599.8±31.0	159945.9±30.8	157076.5±30.6

Continue	$E_{\mathrm{T}}^{\mathrm{miss}} > 20$	$E_{\mathrm{T}}^{\mathrm{miss}} + M_T > 60$	No Bad Jet	$N_{jet} \geq 4$
$t\bar{t}$ NoFullHad (MC)	525.2±0.3	493.9±0.3	483.1±0.2	230.4±0.2
diboson (MC)	160.3±0.1	149.3±0.1	146.1±0.1	2.9±0.0
single top (MC)	173.7±0.5	163.8±0.4	160.9±0.4	12.5±0.1
W+jets (MC)	126514.4±28.2	125114.7±28.1	124182.4±28.0	189.9±0.7
Z+jets (MC)	2766.0±2.8	669.0±1.3	659.5±1.3	22.4±0.2
Total Bkg (MC)	129614.4±28.4	126096.8±28.1	125148.9±28.0	227.7±0.7
QCD multijet (DD)	-	24098.9±155.2	23547.1±153.5	249.4±15.8
data	-	152907	150447	709

Table 4.2: Number of MC events normalized to 33.7 pb^{-1} after successive selection cuts in the electron channel compared to the number of events found in data. The uncertainties shown in the table are statistical in MC. The uncertainties on the QCD multijet background are the square root of the estimated values. No scale factors have been applied to the number of events in $t\bar{t}$ signal.

channel	total	1 elec	trigger	nTrk≥5
$t\bar{t}$ NoFullHad (MC)	2947.1±0.6	717.5±0.3	601.5±0.3	600.9±0.3
diboson (MC)	824.4±0.2	260.8±0.1	218.8±0.1	218.1±0.1
single top (MC)	1262.3±1.2	248.1±0.5	205.3±0.5	204.9±0.5
W+jets (MC)	1072334.7±82.2	208926.6±36.4	175009.8±33.4	171697.7±33.0
Z+jets (MC)	104483.2±16.9	14150.0±6.2	12250.4±5.8	12104.1±5.7
Total Bkg (MC)	1178904.6±83.9	223585.5±37.0	187684.3±33.9	184224.8±33.5

Continue	$E_{\mathrm{T}}^{\mathrm{miss}} > 20$	$E_{\mathrm{T}}^{\mathrm{miss}} + M_T > 60$	No Bad Jet	N_{jet} ≥4
$t\bar{t}$ NoFullHad (MC)	549.4±0.3	520.6±0.3	507.6±0.3	240.3±0.2
diboson (MC)	185.0±0.1	176.3±0.1	171.8±0.1	3.1±0.0
single top (MC)	183.0±0.5	173.0±0.4	169.5±0.4	11.8±0.1
W+jets (MC)	159016.7±31.8	157732.3±31.7	155953.0±31.6	233.2±0.8
Z+jets (MC)	6281.1±4.1	5831.7±4.0	5735.4±4.0	16.3±0.2
Total Bkg (MC)	165665.8±32.1	163913.3±32.0	162029.7±31.8	264.4±0.8
QCD multijet (DD)	-	3095.1±55.6	2998.6±54.8	40.3±6.3
data	-	162561	159492	583

Table 4.3: Number of MC events normalized to 33.7 pb^{-1} after successive selection cuts in the muon channel compared to the number of events found in data. The uncertainties shown in the table are statistical in MC. The uncertainties on the QCD multijet background are the square root of the estimated values. No scale factors have been applied to the number of events in $t\bar{t}$ signal.

4.2 Phenomenology and Monte Carlo Samples

In order to obtain the event selection efficiency, we use the MC samples for the signal ($t\bar{t}$ non fully hadronic decay channels) and backgrounds, which are generated and simulated with ATLAS interface in Top Working Group. In general, a Geant4-based simulation suite for the ATLAS experiment is used for all MC samples with a geometry version labeled ATLAS-GEO-10-00-00. In our analysis, the default MC generator used for signal is MC@NLO (Matrix Element) + Herwig/Jimmy (Parton Shower). The cross section for the MC signal sample is 80.201 *pb* with a k-factor of 1.11, which has been calculated to the level NLO+NLL detailed in Section 2.2.1. The k-factor is used to correct the cross section for higher order effects. The reference top mass in the MC@NLO sample is 172.5 GeV. The MC samples for other top masses are also generated with MC@NLO + Herwig, and the cross sections and k-factors are listed in Appendix A Table A.1.

For the background MC samples, the single top process is generated by MC@NLO + Herwig, while the generators for other backgrounds are Alpgen + Herwig/Jimmy, including Z+jets, W+jets, W+bb+jets, Di-Boson. Concerning the PDF sets used to calculate the cross section, CTEQ is the default PDF. Specifically $t\bar{t}$, diboson and Z+jets are using CTEQ66, while single top is using CTEQ6M and W+jets is using CTEQ66M. The whole list of the sample details are listed in Appendix A.

The total no full hadronic $t\bar{t}$ decay channels are considered as signal, when an events passes the selection of one lepton, large missing transverse energy and at least 4 jets. The semi-leptonic decay channel ($t\bar{t} \to \tau\nu b, qqb$) is treated as signal when the τ lepton decays to electron or muon. It is also true for the di-leptonic decay channels ($t\bar{t} \to l\nu b, l\nu b$), when one of the leptons is lost or misidentified.

The single top channels including s-channel, t-channel and Wt channel shown in Figure 2.4, have also a top quark in the events, thus have similar final states, although they have smaller cross sections. The single top t-channel with the W boson decaying leptonically can fake a $t\bar{t}$ event when additional jets from radiation are associated. The t-channel has relatively larger cross section (about 22 *pb*) than the two others, where the s-channel has 1 *pb* and the Wt channel has 15 *pb* for inclusive decay channels of the two W bosons. Since

it does not have large cross section, it is estimated using MC samples.

The background of $Z \to l^+l^-$ is also considered when one lepton is missing in the final state. It has a total cross section of 3 nb and it diminishes to 10 pb if at least 4 partons are required. The Z boson background is largely suppressed by the selection cut of $E_{\mathrm{T}}^{\mathrm{miss}}$, because there is no neutrino in Z boson decays.

The processes of di-boson are also expected to be sources of backgrounds, where WW has 15 pb, WZ has 5 pb and ZZ has 1 pb. For WW background, where one W decays hadronically and another leptonically can fake a signal event. The W boson decays to lepton and neutrino in WZ channel or a Z boson decays to two leptons in ZZ channel associated with the other Z boson decaying to two quarks can produce a similar final state as the $t\bar{t}$ signal.

The backgrounds of single top, Z+jets and diboson are not expected to be large, therefore they are estimated using MC samples.

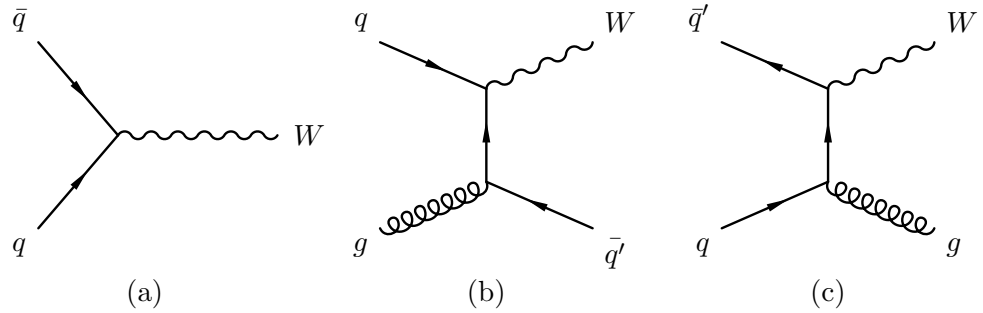


Figure 4.1: The leading order (a) and next to leading order (b & c) Feynman diagrams for W+jets process

Figure 4.1 shows the leading order and next to leading order Feynman diagrams for the W+jets background. A leptonic decay W boson produces a lepton and a neutrino. More partons can be produced when initial state radiation and final state radiation are included. The signature of one lepton, large $E_{\mathrm{T}}^{\mathrm{miss}}$ and several jets in W+jets is similar to the $t\bar{t}$ signal. Besides, the total cross section of W+jets with leptonic decay W boson is about 25 nb including inclusive number of partons and about 100 pb including at least 4 partons. With such large cross section, although it is efficiently reduced with the requirement of at

least four jets, the W+jets background still contributes the number of events competitive to the number of signal events after event selection.

However, using the information of b-tagging, e.g. the requirement of at least one b-tagged jet, it is more effectively eliminated, since the cross section of the W+bb+jets with at least 2 additional partons is much smaller, about 3 pb. It is found that c-jet can be misidentified as b-tagged jet. Therefore the W+cc+jets and W+c+jets backgrounds are also considered as contamination to b-tagged sample. The W+cc+jets and W+c+jets events have been included in the W+jets MC samples in Appendix A in Table A.3. The cross section of W+cc+jets and W+c+jets is 1.8 nb with inclusive number of partons and it is reduced to 90 pb with at least 2 additional partons. The content and cross section of W+cc+jets and W+c+jets are detailed in the note [42]. In Figure 4.2 the leading order Feynman diagrams of W+c and W+cc processes are shown. The W+bb process has the similar diagram, while the corresponding W+b is highly suppressed in the theory of SM. Inclusive W+jets, including contributions from W+jets with heavy flavors, is the dominant background to the signal. Therefore, it is measured using data driven methods, which is detailed in Chapter 5.

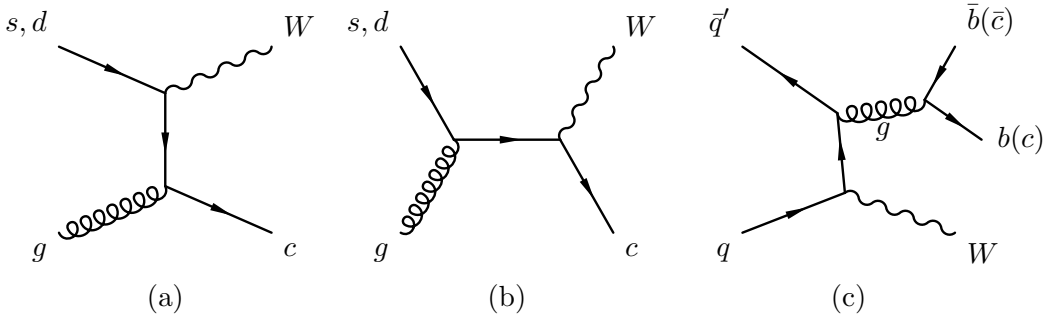


Figure 4.2: The leading order Feynman diagrams for dominent W+jets with heavy flavor, (a) and (b) W+c, (c) W+bb or W+cc

Finally, the QCD multijet is also an important background, because it has extremely large cross section from hadronic processes. A QCD jet has the possibility to fake a lepton. Consequently if the lepton is identified and isolated, it has the chance to finally pass all the selection requirements and contaminate the signal. QCD multijet background is difficult to be simulated in MC, because its theoretical cross section is not well known. Besides, it

has so large cross section that it is hard to produce enough QCD multijet events in MC simulation. Therefore, a data driven method is applied to determine the shape and the normalization of it, which will be introduced in Section 4.3.

4.3 Measurement of the QCD multijet background

The dominant background to the selected $t\bar{t}$ in semi-leptonic channels is W+jets, which has a similar final state signature. The method used to measure the W+jets background will be described in Chapter 5. Another potentially important background is QCD multijet background, which has a very large cross section. A QCD multijet event will contaminate the e+jets signal when a jet is mis-identified as an electron, or an electron originates from photon conversions. An electron can also come from a heavy flavor quark decay, which is also true for a muon. In the muon channel, $\pi^{\pm}/K$ decay can lead to final states similar to signal events. Besides the mis-identified isolated leptons, significant $E_{\mathrm{T}}^{\mathrm{miss}}$ can arise from the fluctuations and mismeasurements of the reconstructed object energy. All these small contributions cannot be reliably estimated from MC, and the QCD multijet background will be estimated directly with data driven methods. In this work, the QCD multijet background is directly measured using a "Matrix Method", which has been applied to the analysis in DØ at Tevatron [43].

In the "Matrix Method", two samples are defined, loose and tight corresponding to the selected lepton, where the tight is a sub-sample of the loose one in the electron channel and muon channel respectively. The tight electrons and muons are using the default object selection presented in Section 3.3.1 and 3.3.2. The loose electrons have the same definition except the requirements of E/p and b-layer hit. The loose muons are defined as the tight without the cuts on ptcone30 and etcone30. It is assumed that the selected numbers of events in loose or tight samples are linear combinations of real and fake lepton contributions, which is expressed in Equation 4.1.

$$\begin{aligned} N^{loose} &= N_{real}^{loose} + N_{fake}^{loose} \\ N^{tight} &= \epsilon_{real} N_{real}^{loose} + \epsilon_{fake} N_{fake}^{loose} \end{aligned} \tag{4.1}$$

N^{loose} and N^{tight} are the numbers of events we find in data. ϵ_{real} is the probability of a real loose lepton to also satisfy the tight criteria with a definition of $\epsilon_{real} = \frac{N_{real}^{tight}}{N_{real}^{loose}}$. The corresponding definition for ϵ_{fake} is $\epsilon_{fake} = \frac{N_{fake}^{tight}}{N_{fake}^{loose}}$. ϵ_{real} and ϵ_{fake} can be obtained

from control samples detailed below. Finally, what needs to be measured is N_{fake}^{tight} found in Equation 4.2.

$$N_{fake}^{tight} = \epsilon_{fake} N_{fake}^{loose} = \frac{\epsilon_{fake}}{\epsilon_{real} - \epsilon_{fake}} (\epsilon_{real} N^{loose} - N^{tight}) \tag{4.2}$$

The method is valid only if ϵ_{real} and ϵ_{fake} are sufficiently different and very much independent of the event topology. In general, the control sample used for obtaining ϵ_{real} is $Z \to l^+l^-$, while for ϵ_{fake} it is a QCD multijet sample obtained with same cuts but with upper threshold for $E_{\mathrm{T}}^{\mathrm{miss}}$ cut value as defined below. The ϵ_{real} and ϵ_{fake} values as a function of $|\eta|$ are listed in Table 4.4. The method description is common for both electron channel and muon channel. The same procedure is applied after b-tagging to get the estimation for tagged sample.

$\|\eta\|$	$\epsilon_{real}(e)$	$\epsilon_{fake}^{untagged}(e)$	$\epsilon_{fake}^{tagged}(e)$	$\|\eta\|$	$\epsilon_{real}(\mu)$	$\epsilon_{fake}^{untagged}(\mu)$	$\epsilon_{fake}^{tagged}(\mu)$
$[0.0-0.3]$	0.947	0.515	0.722	$[0.0-0.3]$	0.991	0.409	0.235
$[0.3-0.6]$	0.945	0.501	0.660	$[0.3-0.6]$	0.991	0.410	0.201
$[0.6-0.9]$	0.946	0.487	0.696	$[0.6-0.9]$	0.991	0.370	0.192
$[0.9-1.2]$	0.927	0.538	0.673	$[0.9-1.2]$	0.991	0.420	0.172
$[1.2-1.37]$	0.913	0.517	0.529	$[1.2-1.5]$	0.992	0.420	0.243
$[1.37-1.52]$	0	0	0	$[1.5-1.8]$	0.990	0.393	0.269
$[1.52-1.8]$	0.868	0.405	0.424	$[1.8-2.1]$	0.990	0.292	0.138
$[1.8-2.1]$	0.852	0.246	0.294	$[2.1-2.5]$	0.990	0.330	0.177
$[2.1-2.47]$	0.770	0.222	0.274				

Table 4.4: The efficiencies of a loose lepton to be identified also as a tight one for real and fake leptons are listed as a function of $|\eta|$ range. The ϵ_{fake} results are separated into the events without b-tagged jet or with at least one b-tagged jet. No uncertainties to the efficiencies are presented. The systematic uncertainties for QCD multijet background are detailed in the text.

In the muon channel, it is found that the ϵ_{real} measured in Z events in data and in MC samples with mixture of W+jets, Z+jets and $t\bar{t}$ are compatible [44]. No significant kinematic dependence on the jet multiplicity is observed. A result of ϵ_{real} measured with the inclusive $Z \to \mu^+\mu^-$ control sample is used as the input: $\epsilon_{real}(\mu) = 0.990 \pm 0.003$ [45].

The ϵ_{fake} is measured in data (period A-F 2.9 pb^{-1}) control sample $E_{\mathrm{T}}^{\mathrm{miss}} < 10$ GeV in each $\mid \eta \mid$ bins with a combined result of $\epsilon_{fake}(\mu) = 0.339 \pm 0.013$ (stat.) $\pm$ 0.061 (syst.) [45].

A 30% relative systematic uncertainty is estimated for muon channel, evaluated with A-F periods data [46], and applied in the top observation paper [45]. The systematic uncertainty is mainly due to the control region uncertainty. The selection of control region as $E_{\mathrm{T}}^{\mathrm{miss}} <$ 10 GeV compared to the one as $E_{\mathrm{T}}^{\mathrm{miss}} < 20$ GeV gives about 15%. The application of the measurement in $E_{\mathrm{T}}^{\mathrm{miss}} < 10$ GeV used in $E_{\mathrm{T}}^{\mathrm{miss}} > 20$ GeV is studied in MC, which gives an uncertainty of about 25%. The uncertainty on $\epsilon_{real}(\mu)$ is relatively smaller, which is about 3%.

To simplify the measurement, the "Matrix Method" is applied to both e+jets and μ+jets channels. The numbers of events and the distributions for QCD multijet background in the following parts are obtained from the "Matrix Method".

In the electron channel, the method is more difficult to be applied because photon conversions and jets also contribute a large fraction to the QCD multijet background. Therefore an alternative method is also applied. A sample is built if it has similar kinematics to the signal selection but with a different object used in place of the electron. This object can be a jet or an electromagnetic cluster constructed from jet-triggered events. Then, a binned likelihood fit to $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in $E_{\mathrm{T}}^{\mathrm{miss}} < 20$ region is applied to estimate the number of expected QCD multijet events.

No systematic uncertainties for the parameters associated to the "Matrix Method" are obtained in the electron channel from other references. Therefore, after estimating the QCD multijet background, the uncertainty is evaluated using the distribution of the variable M_T. Figure 4.3 shows the M_T distribution in 1 jet bin and 2 jets bin in the region of $30 < M_T < 60$, where QCD multijet is dominating. The plots are made after the cut of $E_{\mathrm{T}}^{\mathrm{miss}} >$ 20 GeV and the triangle cut, where the triangle cut means the $E_{\mathrm{T}}^{\mathrm{miss}}$ plus M_T is no more than 60 GeV.

The number of the QCD multijet events in the control region, $30 < M_T < 60$, is shown in Table 4.5, where all the other physics processes are estimated from MC. The data events subtracting all the non-QCD multijet events give an estimation of number of the QCD multijet events. The difference between this estimation and the measured QCD multijet

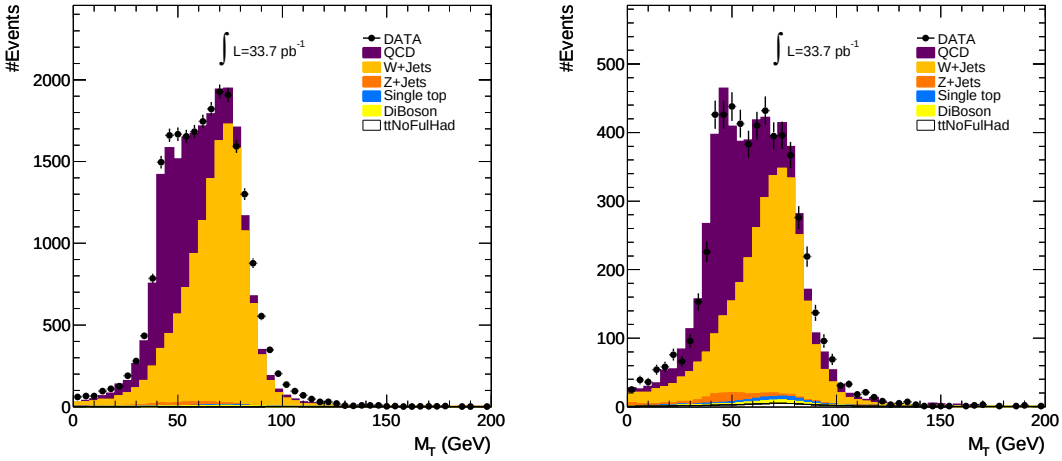


Figure 4.3: The $M_T(l,E_{\mathrm{T}}^{\mathrm{miss}})$ distribution in 1 jet sample (left) and in 2 jets sample (right) in the electron channel. The $E_{\mathrm{T}}^{\mathrm{miss}}$ cut and the triangle cut are applied. The normalizations and shapes of the M_T in both samples are correctly estimated with QCD multijet events from "Matrix Method".

events from the "Matrix Method" is taken to be the systematic uncertainty of the QCD multijet background. The uncertainties of the MC samples are taken into account, where the uncertainty on $t\bar{t}$ is estimated to be 100%, and 30% on all the other processes. A relatively larger uncertainty is assigned to $t\bar{t}$ events in order to avoid any bias on the $\sigma_{t\bar{t}}$ measurement. Therefore the 100% uncertainty is an overestimation. These uncertainties on all the samples give an error to the uncertainty on the QCD multijet, which is finally evaluated to 4.6±18.6% in 1 jet bin and 3.3±17.0% in 2 jets bin. The error bars on the relative uncertainties arise from the uncertainties of the items "data - MC" and "QCD multijet (DD)" shown in Table 4.5. Depending on this study, the uncertainty on the number of estimated QCD multijet background events in the electron channel is taken to be 30% as a conservative estimation.

Finally, after all the selection cuts presented in Section 4.4, the number of events found for QCD multijet background is 40.3±12.1 in the muon channel and 249.4±74.8 in the electron channel in the signal region with at least 4 jets. The background level in the

samples	1 jet	2 jets
data	9535.0±97.6	2514.0±50.1
$t\bar{t}$ NoFullHad (MC)	4.1±4.1	19.5±19.5
diboson (MC)	14.1±4.2	12.2±3.7
single top (MC)	14.4±4.3	15.7±4.7
W+jets (MC)	3473.9±1042.2	868.9±260.7
Z+jets (MC)	126.5±37.9	80.2±24.1
Total MC	3633.0±1042.9	996.6±262.6
data - MC	5902.0±1047.4	1517.4±267.3
QCD multijet (DD)	5642.4±75.1	1569.3±39.6
Uncertainty (absolute)	259.6±1050.1	51.9±270.2
Uncertainty (relative)	4.6±18.6%	3.3±17.0%

Table 4.5: The number of events in the control region, $30 < M_T < 60$ with 1 jet and 2 jets in the electron channel. QCD multijet is measured with "Matrix Method", where the uncertainty is the statistical uncertainty, the square root of the measured number of events. The other backgrounds are estimated from MC, where the uncertainty of $t\bar{t}$ is taken to be 100% and others are 30%. The difference between the measured QCD multijet and the estimation from data subtracting other backgrounds are taken as the uncertainty on the measurement of QCD multijet. The error on this uncertainty is taken from the square root of the quadratic sum of the uncertainty of "data - MC" and the statistical uncertainty of the measured QCD multijet.

electron channel can be reduced when a tighter lepton isolation requirement is applied. The measured QCD multijet background will be used in Chapter 6 to be subtracted from data.

4.4 Selected events distribution

The top pair events selection include 1 isolated lepton, large missing transverse energy and at least 4 good jets. The detailed selection cuts are presented in Section 4.1. The variables distributions are checked in this section, like $E_{\mathrm{T}}^{\mathrm{miss}}$, p_{T} of the selected lepton, leading p_{T} of the selected jet. It explains the power how each cut is used to suppress certain background.

The distribution of $E_{\mathrm{T}}^{\mathrm{miss}}$ in Figure 4.4 gives an idea that most of the QCD multijet background events are stacked in the region of low $E_{\mathrm{T}}^{\mathrm{miss}}$ region. QCD multijet events are not included in this plot, because they are not well estimated before the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut. The $Z \rightarrow l^{+}l^{-}$ background is also highly suppressed by the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut at 20 GeV, while the W+jets background remains a large fraction because of the leptonic decay of the W boson emitting a neutrino.

Figure 4.5 shows the p_{T} distribution of all the jets before the triangle cut without any good jet selection, where QCD multijet is not drawn as well for the same reason as the plot of $E_{\mathrm{T}}^{\mathrm{miss}}$. In $t\bar{t}$ not-fully-hadronic decay channels, jets from top decay and jets from other sources like initial state radiation (ISR), final state radiation (FSR), are separated. The jets from top decay tend to have higher p_{T}, while other jets have more probability to be low p_{T} jets. A large amount of jets in W+jets background are also found in the low p_{T} region. The p_{T} cut of the jets is chosen to be 25 GeV, which is optimized to get the higher signal over background ratio.

Figure 4.6 shows the distribution of the transverse mass reconstructed from the lepton and missing transverse energy before the triangle cut. No number of jets requirement is applied either. At this stage, the QCD multijet background is not well measured especially for no jet events, which can be emphasized by being compared to the Figure 4.3 of M_T in 1 jet bin and 2 jets bin. Therefore, a discrepancy is found between data and MC simulation in both channels, which is even larger in the electron channel. Despite of the small discrepancy, the plots show that the QCD multijet distribution describes reasonably well normalization and shape of the low M_T distribution in data. The discrepancy is less after the triangle cut and jet multiplicity cut, which can be seen from the control plots after the requirement of at least 4 jets, for example Figure 4.8 to Figure 4.12.

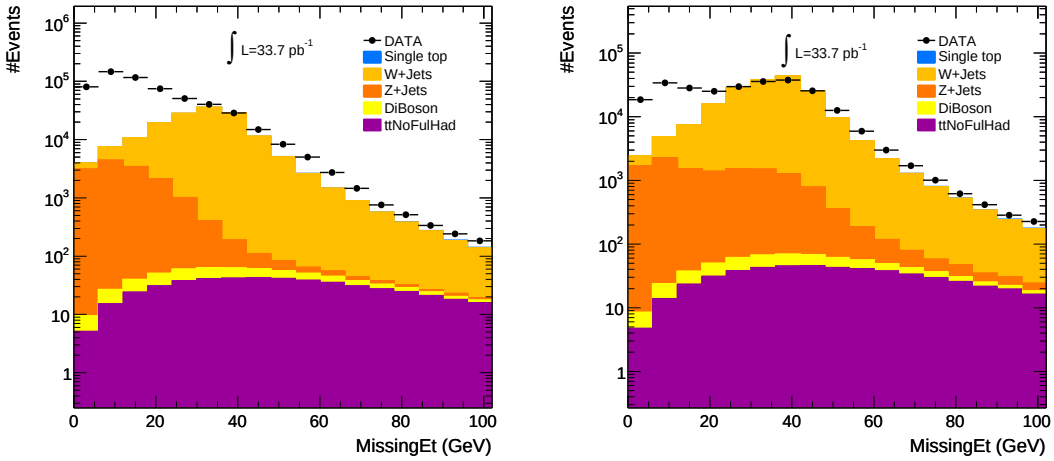


Figure 4.4: The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution in the electron channel (left) and in the muon channel (right) in log scale before the cut of $E_{\mathrm{T}}^{\mathrm{miss}}$. The triangle cut and the requirement of at least 4 jets are not applied as well. QCD multijet events are not presented, because of no estimation by the "Matrix Method" before the $E_{\mathrm{T}}^{\mathrm{miss}}$ cut. QCD multijet events are expected in the low $E_{\mathrm{T}}^{\mathrm{miss}}$ in data, as well as the Z+jets background shown in the plots.

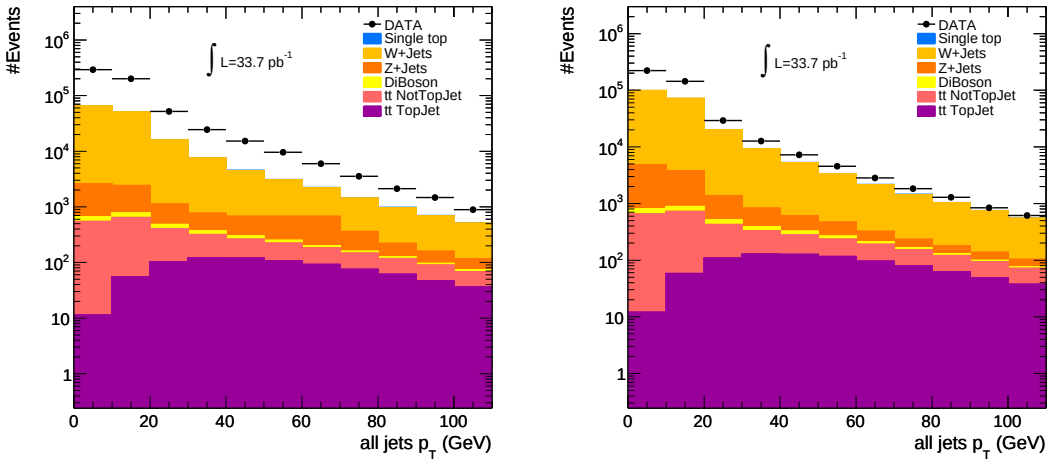


Figure 4.5: The distribution of p_{T} of all the selected jets in the electron channel (left) and in the muon channel (right) in log scale. The cut on $E_{\mathrm{T}}^{\mathrm{miss}}$ is applied, while the triangle cut and the requirement of at least 4 jets are not applied. The jets originated from top quark decay have large p_{T}, when the ones from ISR/FSR in $t\bar{t}$ have smaller p_{T}. The jets in W+jets and Z+jets also tend to stack in the low p_{T} region.

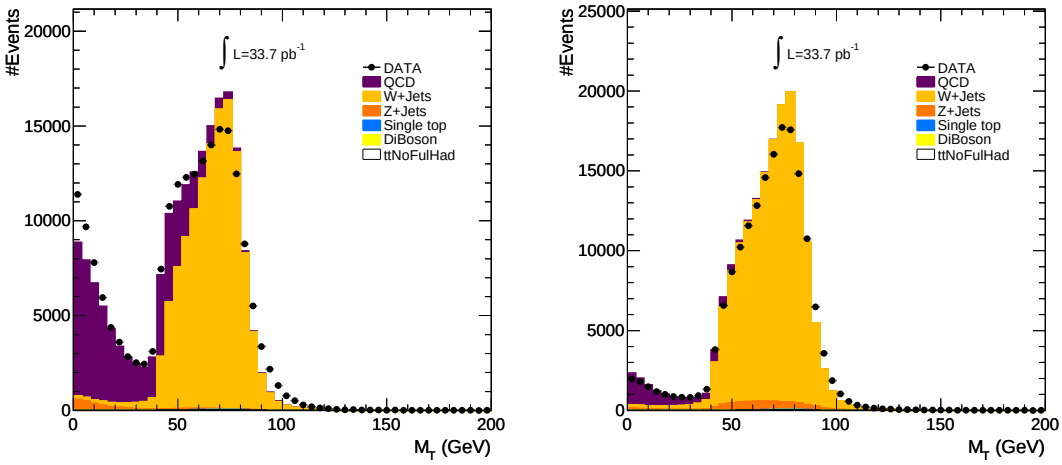


Figure 4.6: The $M_T(l,E_{\mathrm{T}}^{\mathrm{miss}})$ distribution before the triangle cut without jet multiplicity requirement in the electron channel (left) and in the muon channel (right). The shapes of the low M_T in both channels are correctly estimated by QCD multijet events from "Matrix Method". The discrepancy between data and MC mainly arises from no jet events.

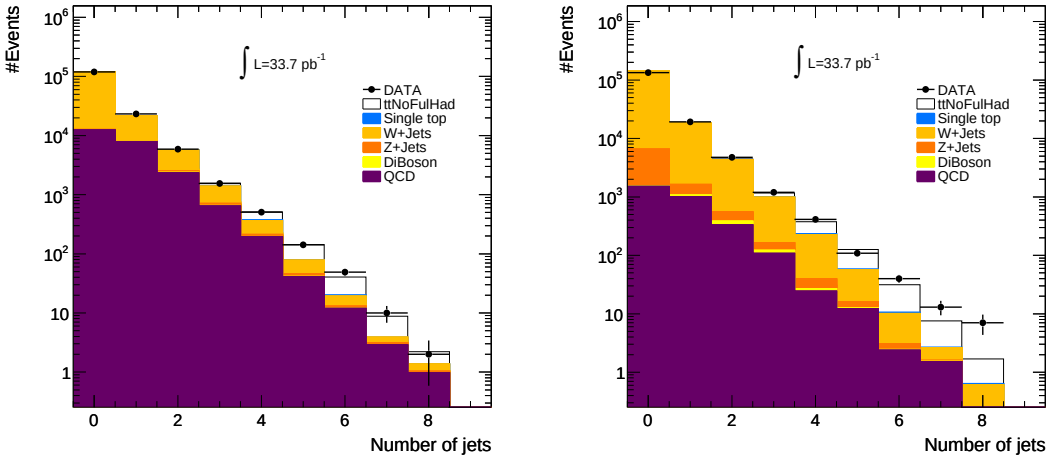


Figure 4.7: The number of jets in the electron channel (left) and in the muon channel (right) in log scale. The triangle cut is applied, while no jet multiplicity requirement is used. The plots show a rich W+jets contamination especially in low jet multiplicity region. It also contributes most in the signal region with at least 4 jets. The QCD multijet background is also large in the electron channel.

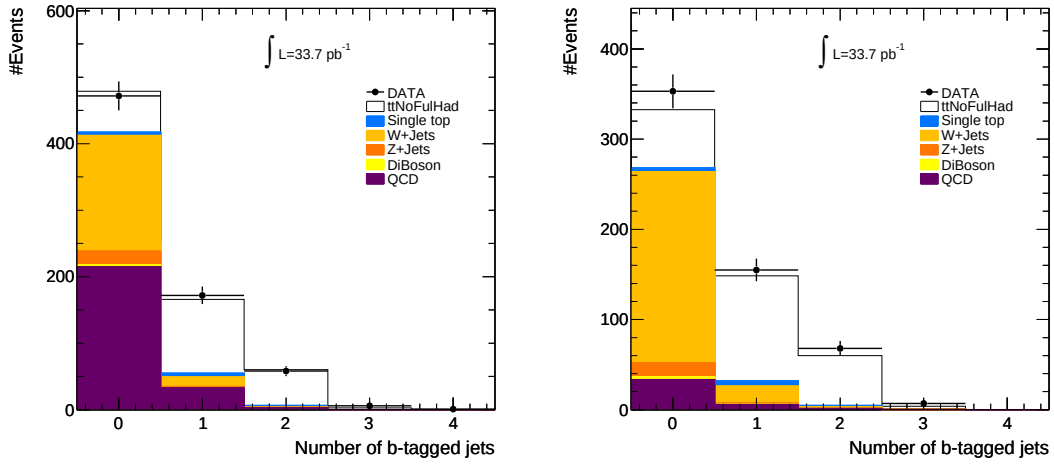


Figure 4.8: The number of b-tagged jets in the electron channel (left) and in the muon channel (right) in the signal region with at least 4 jets. Large background contamination is found in 0 b-tagged jet bin in both channels. With at least one b-tagged jet, the purity of signal events is largely improved.

The jet multiplicity distribution Figure 4.7 shows that the $t\bar{t}$ signal contributes more in a higher jet multiplicity region. With the number of jets larger than four, one finds large contamination of W+jets in both channels and large QCD multijet background in the electron channel. With the requirement of $N_{jets} > 4$, the distribution of number of b-tagged jets shows a lower background contamination, see Figure 4.8, especially in the bins with equal to or more than one b-tagged jet.

After all the selection cuts described in Section4.1 are applied, the left number of events in each physical process is listed in Table 4.6. In the electron channel, 709 data events are found. About one-third of the events are W+jets background, which shown in the table is measured with the Berends Scaling data driven method detailed in Section 5.3. QCD multijet events contribute about one-third of the total events found in data. Other backgrounds estimated in MC are relatively much smaller. The expected number of event for $t\bar{t}$ signal is 191.9 $\pm$ 112.2 events with an error bar mainly originated from W+jets and QCD multijet measurement. Therefore, the signal over background (S/B) ratio is about 0.4

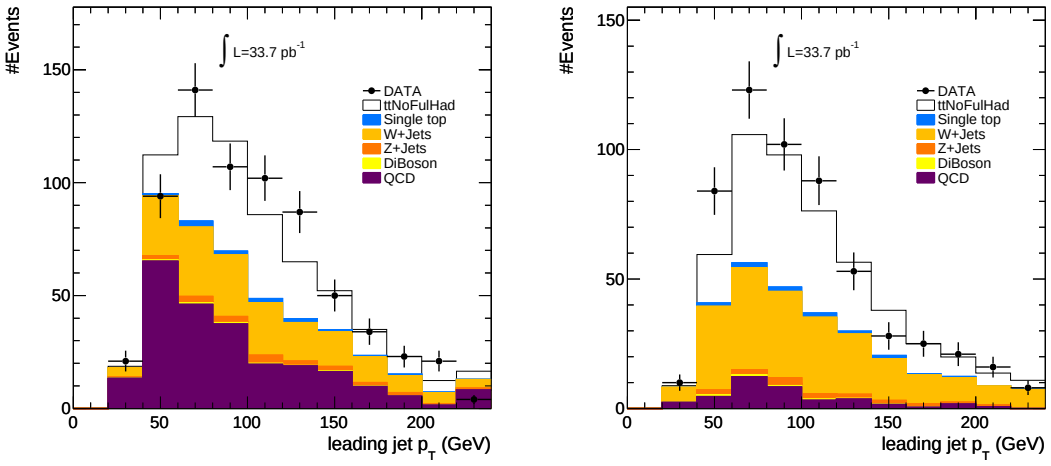


Figure 4.9: The p_{T} distribution of the leading jet in the electron channel (left) and in the muon channel (right) in the signal region with at least 4 jets. The signal events tend to have higher p_{T} for the leading jets, while the p_{T} of the leading jets in W+jets and QCD multijet are relatively lower.

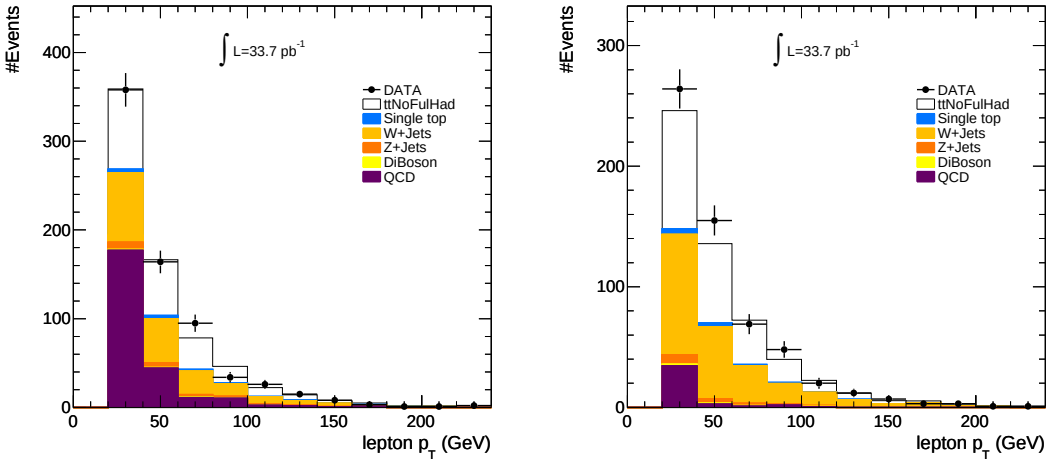


Figure 4.10: The p_{T} distribution of the lepton selected in the electron channel (left) and in the muon channel (right) in the signal region with at least 4 jets. The p_{T} of the leptons in QCD multijet events are lower than those in $t\bar{t}$ signal and W+jets background.

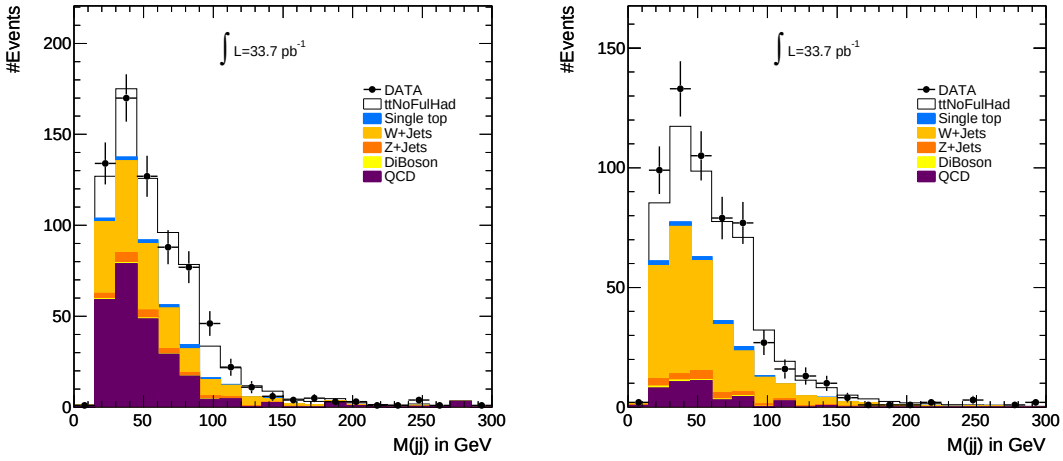


Figure 4.11: The mass distribution of jet-jet pair with smallest ΔR among all the jet-jet combinations in each event in the electron channel (left) and in the muon channel (right) in the signal region with at least 4 jets. The shapes of the signal and backgrounds are not well separated in both channels. With only the two jets from W boson decay in $t\bar{t}$ sample, the mass distribution has a mean value in agreement with the expected W boson mass with a reasonable resolution, which is detailed in the texts.

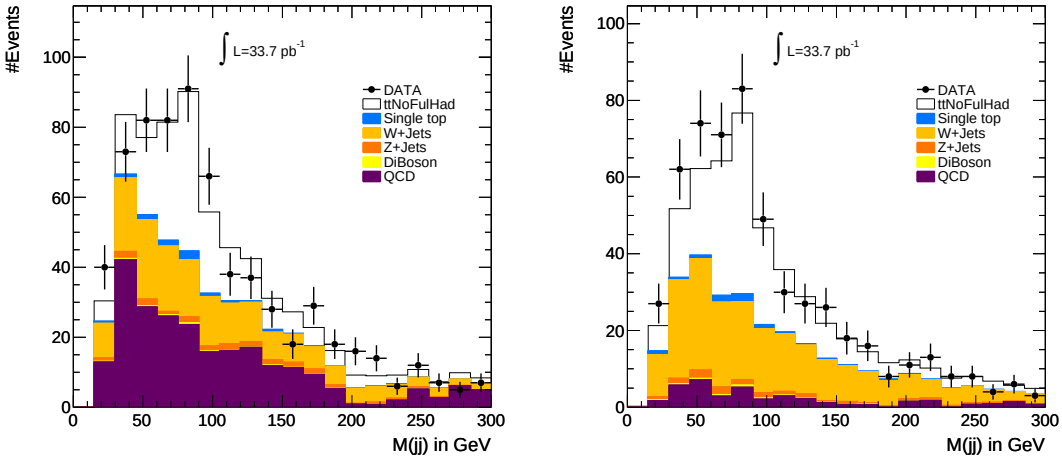


Figure 4.12: The mass distribution of jet-jet pair with largest p_{T} among all the jet-jet combinations in each event in the electron channel (left) and in the muon channel (right) in the signal region with at least 4 jets. The shapes of the signal and backgrounds are not well separated in both channels. With this jet-jet pair selection for those selected pairs from W boson decay in $t\bar{t}$ sample, the mass distribution has a mean value in agreement with the expected W boson mass with a reasonable resolution, which is detailed in the texts.

N_{evt} selected	e-chan	μ-chan
Data Obs.	709±26.6	583±24.1
diboson (MC)	2.9±0.9	3.1±1.0
singletop (MC)	12.5±3.9	11.8±3.7
Z+jets (MC)	22.4±7.0	16.3±5.1
W+jets (DD)	229.9±78.9	291.4±77.7
QCD multijet (DD)	249.4±74.8	40.3±12.1
$t\bar{t}$ Expect	191.9±112.2	220.1±82.5
$t\bar{t}$ (MC)	230.4±0.2	240.3±0.2

Table 4.6: Number of events after selection cuts in data, W+jets and QCD multijet from data driven methods and other samples from MC estimation. W+jets background is measured by the Berends Scaling method, while QCD multijet by the "Matrix Method" with about 30% uncertainties. Other backgrounds are from MC estimation with uncertainty levels assumed at 30%. The expected $t\bar{t}$ number of events is obtained using data subtracting total backgrounds with an uncertainty inherited from the processes. The statistical uncertainty is shown for the number of $t\bar{t}$ events in MC.

in the electron channel.

In the muon channel, 583 events are finally selected in data. A large W+jets background is found which contributes about one-half of the total number of events. The QCD multijet background is much smaller than that in the electron channel. Other backgrounds estimated in MC are also small compared to W+jets. After all, the number of signal events expected in the muon channel is 220.1 ± 82.5, with an error bar much smaller than that in the electron channel. The S/B is 0.6 in this channel.

Large backgrounds are found in both channels. The uncertainties on the backgrounds can lead to the uncertainties of the signal results. By looking at the Figure 4.8 of the number of b-tagged jets, one find the background contamination is much higher in no b-tagged jet bin than in other $\geq$1 b-tagged jet(s) bins. If at least one b-tagged jet is required, the S/B ratio can increase to 2.8 in the electron channel and 3.8 in the muon channel. The method using the b-tagging information is described in Chapter 6.

4.5 Control plots

After all of the selection cuts, the control plots containing the signal, backgrounds and data are compared. Figure 4.9 shows the p_{T} distribution of the leading jet. In the electron channel, the QCD multijet background contributes a large fraction, mostly in the lower p_{T} region, while the $t\bar{t}$ signal tends to have larger p_{T}. In the muon channel, the QCD multijet is much smaller, while W+jets has the largest contamination and has lower p_{T} compared to the signal. Figure 4.10 shows the distribution of the lepton p_{T}, where the contributions from QCD multijet populates the lower region.

In order to reconstruct the W boson mass and eventually the top quark mass through the hadronic decay, two ways of jet-jet pair combinations are applied to the non b-tagged jets. The first one is to combine the two jets with smallest ΔR between them. The mass of the jet-jet pair for MC and data is shown in Figure 4.11. Data and MC with QCD multijet background expectations are found to be matching, although large backgrounds are seen in the plots. Concerning the $t\bar{t}$ MC signal, 15.6% of the jet-jet pairs are matching the two quarks originated from the W boson decay. Those matched jet-jet pairs gives a W boson mass mean value of 80.5 GeV with an RMS of 10.9 GeV. This demonstrates a reasonable reconstruction of the W boson mass. The second method chooses the jet-jet pair with largest p_{T}, shown in Figure 4.12. This selects 15.0% of "true" W boson decay jets from $t\bar{t}$. Those matched pairs have a mean value of reconstructed mass of 81.7 GeV with an RMS of 11.0 GeV. Although 85% of combinatorial backgrounds in $t\bar{t}$ signal sample are found, the match of data and expectation validates the two methods to select a W boson decay jet-jet pair.

Another jet either non b-tagged jet or b-tagged jet is chosen with the largest p_{T} to combine with the jet-jet pair reconstructed above. If the third jet is a b-tagged jet, their invariant mass of the combination jet-jet-b is shown in Figure 4.13 and Figure 4.14. Among all selected events in MC $t\bar{t}$, the combination of jet-jet-b for a jet-jet pair having minimum ΔR, 11.5% of them are matching $t \to Wb \to qqb$ from top decay. The corresponding ratio is 10.6% for jet-jet pair with maximum p_{T}. The reconstructed masses of these jet-jet-b combinations in both cases are similar, 167.1±16.8 and 167.7±16.9, where the errors shown

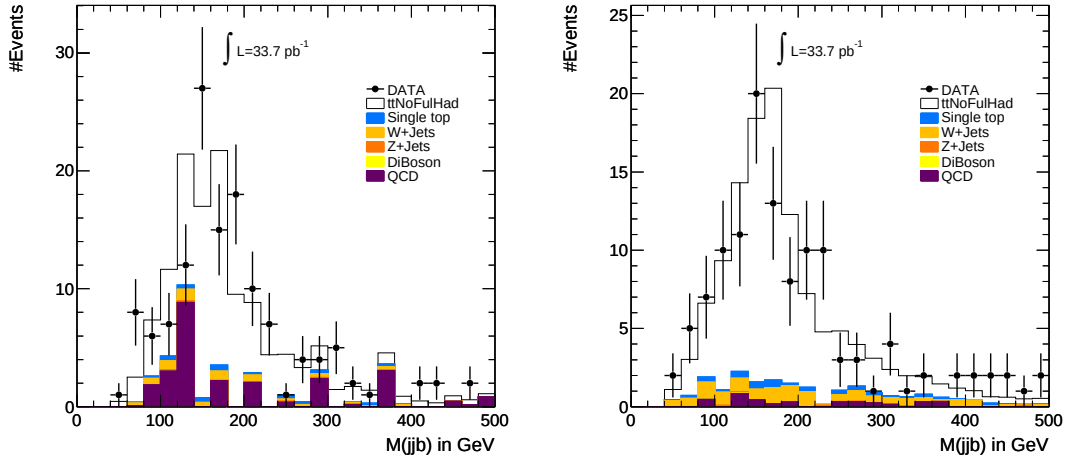


Figure 4.13: The mass distribution of jet-jet-b combination with largest p_{T} among all the jet-jet-b combinations in each event with at least one b-tagged jet in the electron channel (left) and in the muon channel (right) in the signal region with at least 4 jets. The jet-jet pair is first selected with smallest ΔR among all the jet-jet combinations. The jet-jet-b distribution has a mean value in agreement with the expected top mass with a reasonable resolution when it is combined with the two light jets from W boson decay and the b-tagged jet from top decay in $t\bar{t}$ MC sample, which is detailed in the texts.

are the RMS errors. The selected MC signal events give a reconstructed jet-jet-b mass close to the expectation, which demonstrate the selection of $t\bar{t}$ events in data using the topology of $t\bar{t}$ in SM prediction. There are also chances that 3 light jets combined to have a largest p_{T} even in an event there exits one or more b-tagged jets. However, the mass distribution is worse than the combination jet-jet-b because of more combinatorial background.

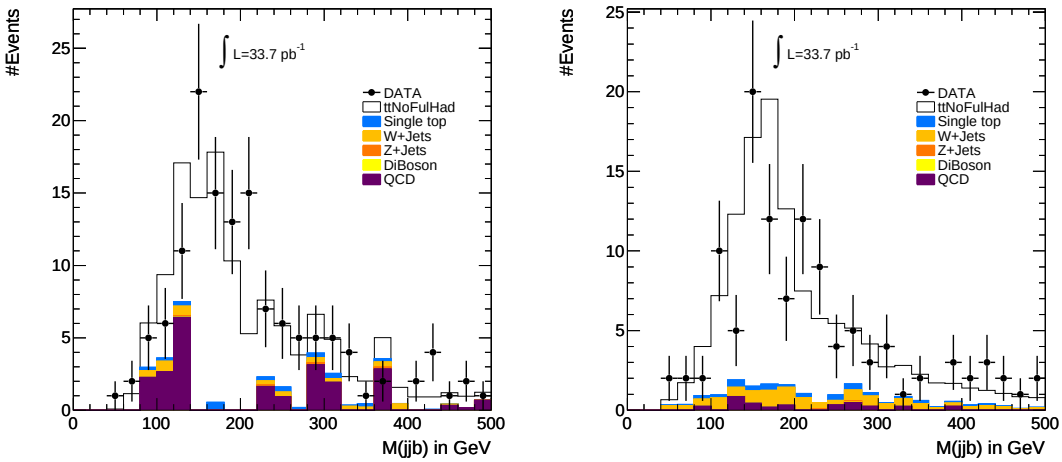


Figure 4.14: The mass distribution of jet-jet-b combination with largest p_{T} among all the jet-jet-b combinations in each event with at least one b-tagged jet in the electron channel (left) and in the muon channel (right) in the signal region with at least 4 jets. The jet-jet pair is first selected with largest p_{T} among all the jet-jet combinations. The jet-jet-b distribution has a mean value in agreement with the expected top mass with a reasonable resolution when it is combined with the two light jets from W boson decay and the b-tagged jet from top decay in $t\bar{t}$ MC sample, which is detailed in the texts.

4.6 Scale factors

The object reconstruction/identification/trigger scale factor is defined as SF = $\frac{\epsilon^{data}}{\epsilon^{mc}}$ for electron and muon reconstruction efficiency, trigger efficiency and identification efficiency, which account for mis-modeling of detector and lepton reconstruction and identification effects in MC with respect to data. The SFs will be used to rescale the acceptance term entering the cross section measurement. The SF for the resolution of lepton momentum (or energy) is also considered for electron and muon.

The SFs are obtained from $Z \to l^+l^-$ events using a Tag & Probe method, assuming that the SFs measured with Z events can be applied to $t\bar{t}$ events. The difference between the two samples is considered as systematic uncertainty. The Tag & Prob method has been used in the Tevatron experiments and the application in Atlas is detailed in [47]. The SFs are defined in Equation 4.3, where i stands for trigger, reconstruction or identification.

$$SF_i = \frac{\epsilon_i(Z, data)}{\epsilon_i(Z, mc)} \tag{4.3}$$

In general, the SFs depend on the kinematic variables of the lepton, e.g. η, ϕ and p_{T}. The detailed description for the SF usage and results is shown in [48]. The SFs enter the formula of measured cross section:

$$\sigma_{t\bar{t}} = \frac{N_{sig}}{L \cdot A \cdot \epsilon} = \frac{N_{sig}}{L \cdot A \cdot \prod_i \epsilon_i \cdot SF_i} \tag{4.4}$$

Concerning the muon trigger, Table 4.7 presents the trigger SFs for data periods A-E3, while for other periods, the SFs are measured with the reprocessed data and detailed in [49]. Table C.1 define the binning for the measurement on the latter periods E4-I and the values are reported in Table C.2 in Appendix C. By using the numbers of events found in the regions defined in MC signal sample, one can calculate the combined trigger scale factor for data periods E4-I is 1.005±0.008. This trigger SF is combined with the one for data A-E3 to get the overall muon trigger SF to be 1.003±0.008, which has much smaller error because period A-E3 contributes only 3% of the data.

Table 4.8 shows the reconstruction and identification SFs for muons, which are consistent with 1, showing a good data and MC agreement. The SFs in this table are measured with

Period (Trigger)	Muon SF (Barrel)	Muon SF (End-cap)
A-E3(L1 MU10)	$0.959^{+0.039}_{-0.041}(\text{stat})^{+0.013}_{-0.002}(\text{syst})$	$0.943^{+0.027}_{-0.029}(\text{stat})^{+0.006}_{-0.010}(\text{syst})$

Table 4.7: Muon trigger SFs from tag & probe method on $Z \to \mu^+\mu^-$ data periods A-E3 and MC events.

2.89 pb^{-1}. Unlike the trigger SF where the trigger changes from period to period and relies much on detector condition, the reconstruction and identification SFs are relatively stable. Therefore, those SFs are used for all data periods in this analysis.

Stage	Scale Factor (Muon)
Reconstruction	1.004±0.004(stat)±0.011(syst)
Identification	0.994±0.004(stat)±0.001(syst)
Reco+Id	1.000±0.006(stat)±0.011(syst)

Table 4.8: Muon reconstruction and identification SFs from tag and probe method on $Z \to \mu^+\mu^-$ data 2.89 pb^{-1} and MC events.

It has been observed that the width of the Z boson in $Z \to \mu^+\mu^-$ is larger in data than in MC. This is due to the momentum resolution in the muon spectrometer and the inner detector are not yet perfectly described in MC, because the actual material budget and the misalignment of various detectors in the reconstruction is not yet final. In order that the reconstructed Z boson in MC and data can match, we smear each muon in MC by the smearing function used in $W \to \mu\nu$ and $Z \to \mu^+\mu^-$ cross section paper [50]. The smearing function is:

$$p_T^s = \frac{C_1 \times p_{\mathrm{T}}}{1 + x \times C_2} \tag{4.5}$$

where p_T^s is the smeared muon p_{T} in MC, C_1 and C_2 are two parameters for barrel ($\mid \eta \mid < 1.05$) and end-cap ($1.05 < \mid \eta \mid < 2.5$) muons, which are listed in table 4.9, and x is a random number generated with a Gaussian distribution $Gaus(0, 1)$. The scale factor for the muon

selection efficiency cut with this muon momentum smearing is $0.987\pm0.004(\text{stat.})^{+0.008}_{-0.002}(\text{syst.})$. The systematic error is due to the uncertainty on the two parameters C_1 and C_2.

	C_1	C_2
Barrel	0.992±0.010	0.031±0.020
End-cap	0.980±0.012	0.063±0.031

Table 4.9: Parameters in muon smearing function for barrel and end-cap

The SFs for electrons shown in Table 4.10 are also close to 1, while the identification SF deviates by 2.4%, which still gives a good match for data and MC. The same measurement for electron trigger SF for data periods E4-I gives 0.995±0.005, which leads to a final trigger SF = 0.996±0.005. Identification and reconstruction scale factors measured with 2.89 pb^{-1} data are also applied to the whole data periods in electron channel.

Stage	Scale Factor (Electron)
Trigger	$1.007^{+0.002}_{-0.003}(\text{stat})\pm0.003(\text{syst})$
Identification	0.976±0.016(stat)±0.016(syst)
Reconstruction	1.000±0.022

Table 4.10: Electron trigger, reconstruction and identification SFs from the tag and probe method for $Z \to e^+e^-$ data 2.89 pb^{-1} and MC events.

Because of the deviation of the Z boson mass in the $Z \to e^+e^-$ channel, a smearing for electron energy is also applied according to the following function:

$$E^s = E \times [1 + Gaus(0, \frac{0.11}{\sqrt{E}}) + Gaus(0, 0.011)] \tag{4.6}$$

The results for electron momentum energy smearing is 1.000±0.004 with only statistical uncertainties.

The final SFs combining reconstruction, identification, trigger and resolution presented above for muon and electron are shown in Table 4.11, which is 0.971 ± 0.032 in the electron channel and 0.990 ± 0.018 in the muon channel. The corrections have been applied to MC via the SFs for both electron and muon channels. The dependence of lepton p_{T} and η is considered in the SFs. The measured SFs are finally applied to determine the top pair acceptance in the Section 4.7.

	electron	muon
Combined SF	0.971±0.032	0.990±0.018

Table 4.11: Combined SF for muon and electron, including trigger, reconstruction, identification and energy resolution.

4.7 Conclusion: acceptance for $t\bar{t}$ events

In Equation 1.1 the acceptance times the selection efficiency ($A \cdot \epsilon$) is finally used to calculate the $t\bar{t}$ cross section. The selection efficiencies as a function of top mass is found in Appendix D in Table D.1.

The scale factors discussed in Section 4.6 for electron and muon are included to obtain the final results for the $A \cdot \epsilon$, which is shown in Table 4.12. The uncertainties on the acceptances in the table include the statistical uncertainties and the uncertainties of the scale factors. Other systematics on the acceptance like the jet energy scale, ISR/FSR, pile up are considered in the method of tag counting detailed in Chapter 6 when the results for the number of signal events are obtained.

M_{Top} (GeV)	e+jets	μ+jets
160	6.68±0.23%	7.17±0.14%
165	7.00±0.24%	7.65±0.15%
167.5	7.19±0.25%	7.93±0.16%
170	7.27±0.25%	7.80±0.16%
172.5	7.59±0.25%	8.07±0.15%
175	7.78±0.27%	8.24±0.17%
177.5	7.91±0.27%	8.30±0.17%
180	8.11±0.28%	8.62±0.17%
190	8.78±0.30%	9.31±0.19%

Table 4.12: Acceptance times efficiency including the corrections from SFs $A \cdot \epsilon \cdot SFs$ as a function of top mass in the electron channel and in the muon channel. The statistical and systematic uncertainty are included. The results with top mass at 172.5 GeV are used as default to get the top pair cross section measurement.

The results of $A \cdot \epsilon$ with top mass at 172.5 GeV are used to get the results of top pair cross section. In the electron channel, it is 7.59% having a relative error at 3.3%. Similarly, it is 8.07% in the muon channel with a relative error at 1.9%. The uncertainties on the

results are mainly from the measurement of the scale factors.

In the Equation 1.1, the two items essential to cross section measurement is the parameter $A \cdot \epsilon$ and the number of background events N_{bkg}. QCD multijet background is already described in this chapter. Therefore, the largest background, W+jets background, is introduced in the next chapter.

Chapter 5

MEASUREMENT OF W+JETS IN DATA

The final state of W boson production in association with jets is similar to $t\bar{t}$ semi-leptonic decay channels. W+jets production has a significantly large cross section, which is more than O(10^2) larger than $t\bar{t}$ production. Therefore W+jets is the dominant background in the analysis. The W+jets normalization should be determined by data driven techniques, so we measure the background with three methods, the charge asymmetry method, the Berends scaling method and the combined method using part of both methods above.

A detailed study on the systematics of these methods is also presented. The charge asymmetry method measures directly the number of events in $\geq$4 jets region. The method is limited by statistics with the data collected in 2010 in Atlas, which has 70% in the electron channel and 25% in the muon channel. The total systematic uncertainty of the method is about 10%. The Berends Scaling method uses the large samples in 1 jet region and 2 jets region, where the statistical uncertainties are much smaller than in $\geq$4 jets region, about 10% in both channels. The method is largely limited by the systematic uncertainty of 24%. The combined method takes advantage of the number of W+jets events in 2 jets region with the charge asymmetry method and uses it with other parameters in Berends Scaling method. It reduces much the total uncertainty on the measured number of W+jets event, especially in the electron channel.

By using b-tagging in the top pair selection, the background of W+jets with heavy flavors will arise to have a larger contribution. The samples of W+jets with heavy flavors are introduced in Section 4.2, including W+bb+jets, W+cc+jets and W+c+jets. Therefore, The number of events for W+jets with at least 1 b-tagged jet is also measured in this chapter.

5.1 Charge asymmetry measurement in W+jets events

In proton-proton collisions, W boson is produced at the parton level through for example $u\bar{d} \to W^+$, $\bar{u}d \to W^-$. In a proton, the contributions of $\bar{d}$ quark and $\bar{u}$ are from gluon splitting, referred as "see quarks". A proton contains the "valence quarks" with combination uud, where a larger contribution from u "valence quark" is expected than d "valence quark". This indicates that the cross section of W^+ production is larger than W^- production, which leads to a charge asymmetry. The ratio of the cross sections for W^+ and W^- is defined as $r = \frac{\sigma(pp\to W^+)}{\sigma(pp\to W^-)}$.

Theoretically, this ratio is better understood than W+jets cross section [51] [52], while the main theoretical uncertainty comes from the uncertainties in PDFs. The uncertainty of the ratio is predicted to be a few percent in theory. Besides, most of the physical processes in SM like $t\bar{t}$, QCD multijet, Z+jets and single top Wt channel are charge symmetric, while single top s-channel and t-channel are charge asymmetric. The total number of single top events is relatively small, about 5% of $t\bar{t}$ signal after the top pair selection, among which 70% is from s-channel and t-channel and 30% is from Wt channel.

Therefore the different number of events for W^+ and W^- candidates seen in data subtracting single top are assumed to be a good prediction for W events, $N_{W^+} - N_{W^-} \approx N_{D^+} - N_{D^-}$. The single top events are quite negligible comparing to the W+jets events. Thus, all single top channels are subtracted, s-channel, t-channel and Wt channel as well. By subtracting the charge symmetric Wt channel, one hardly looses statistics and it is not biasing the number of $N_{D^+} - N_{D^-}$ events. After the single top background subtraction, the W+jets measurement with charge asymmetry method is based on the Formula 5.1.

The assumption of all the other physical samples, including $t\bar{t}$, Z+jets, QCD multijet, being charge symmetric may not be true and can cause some deviation during the application of the method. By looking at the number of events with charge positive and charge negative presented in Table E.1 and Table E.2 in Appendix E, one finds out that the effect is negligible comparing to the large statistical uncertainty.

$$N_{W+jets} = N_{W^+} + N_{W^-} = (\frac{r_{mc}+1}{r_{mc}-1})(N_{D^+} - N_{D^-}) \tag{5.1}$$

A_{mc}	W+light		Wbb		single top	
NJet	e-chan	μ-chan	e-chan	μ-chan		
0	0.217±0.001	0.239±0.001	0.305±0.053	0.272±0.047	0.206±0.058	0.263±0.060
1	0.178±0.004	0.197±0.003	0.310±0.040	0.223±0.038	0.235±0.027	0.242±0.025
2	0.187±0.006	0.214±0.005	0.222±0.037	0.318±0.034	0.197±0.024	0.171±0.025
3	0.221±0.011	0.242±0.010	0.223±0.052	0.298±0.046	0.134±0.038	0.075±0.035
≥4	0.264±0.020	0.249±0.019	0.326±0.066	0.304±0.061	0.021±0.055	0.162±0.053

Table 5.1: A_{mc} estimated in MC with W+light samples are $W \to e\nu$ and $W \to \tau\nu$ for electron channel and $W \to \mu\nu$ and $W \to \tau\nu$ for muon channel. A_{mc} in Wbb and single top channels are also listed to be compared. The uncertainty is due to the MC statistics only.

N_{D^+} and N_{D^-} are the number of W candidates in data after backgrounds subtraction. The asymmetry variable is defined as $A_{mc} = \frac{r_{mc}-1}{r_{mc}+1} = \frac{W^+_{mc}-W^-_{mc}}{W^+_{mc}+W^-_{mc}}$. The ratio r is obtained from MC simulation by including W($\to \tau\nu$)+jets in both the W($\to e\nu$)+jets and the W($\to \mu\nu$)+jets channels. The W+bb+jets background is not used to calculate the A_{mc} parameter, because it is much smaller contribution limited by the statistics. It is 2% of the W+light jets sample and will not affect the results to A_{mc}.

The charge asymmetry method can be applied with the requirement of at least one b-tagged jet. In that case, the W+bb+jets sample is increased to 14% of the W+light jets sample and should not be neglected. With the limitation of statistics at the present stage, the method is not in reality applied to data.

The asymmetry parameter A_{mc} is shown in Table 5.1 and Figure 5.1 as a function of the number of jets and the lepton pseudo-rapidity. By looking at the left plot in Figure 5.1, one finds that there is a significante difference in A_{mc} between no jet event and at least one jet event. This is because the asymmetry is expected to be dominated by quark-quark scattering in no jet event and quark-gluon scattering or quark-quark scattering associated with gluon splitting in at least one jet event. As the PDF distributions shown in 2.3, quark-quark requires high momentum fraction of the parton to the proton, known as x, in which region u and d quarks are mainly contributed from "valence quarks". Concerning quark-gluon scattering, it requires a rich content of gluons, thus lower x, where u and d

quarks have large contribution from "see quarks" and the distributions are closer to each other. Therefore A_{mc} in no jet event is higher than that in 1 jet event. When the jet multiplicity increases, larger x of the partons are expected, because more radiation partons are produced. This effect leads to a larger difference of u quark and d quark content so that A_{mc} increases with the number of jets.

Besides, A_{mc} also depends on the distribution of the lepton pseudo-rapidity. By looking at the PDF distributions, one finds that the u "valence quark" tends to have large x, while the $\bar{d}$ "sea quark" is much less energetic with lower x. Therefore largely boosted W^+ is produced, which will decay to leptons and the leptons are also boosted and can be found in high pseudo-rapidity region. The largely boosted W^- is produced in the same way with d "valence quark" in high x region. In high x region, large difference is found for u and d "valence quark", thus large A_{mc} is expected. The less boosted W^+ and W^- are produced with more contributions of u and d "sea quark" in the lower x region. Consequently, the leptons from W^+ and W^- decay also tend to be in low pseudo-rapidity region, which finally gives lower A_{mc}. The leptons are considered to be highly correlated with the mother W boson.

By using the MC simulation of W+jets, one finds the charge asymmetry distribution as a function of number of jets and lepton pseudo-rapidity. It matches the theoretical prediction in SM. The parameter A_{mc} estimated with at least 4 jets can be used to get the number of W+jets event using the Formula 5.1, when the number of events in data with charge positive and negative is applied. The measurement of W+jets background using the charge asymmetry method is presented in the next section. The systematic uncertainties in the method are also considered.

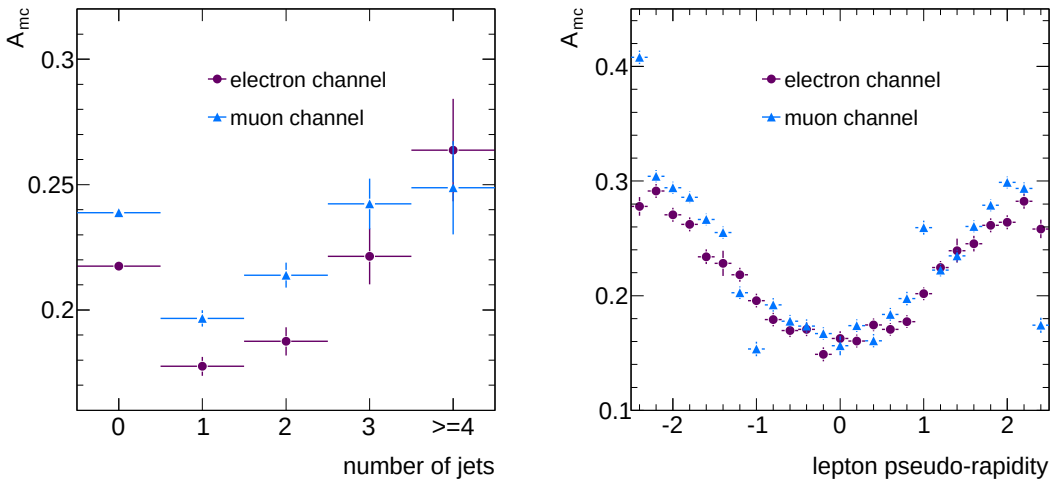


Figure 5.1: The parameter asymmetry A_{mc} obtained from MC W+jets as a function of the number of jets (left) and the lepton pseudo-rapidity (right). Overlap removal between leptons and jets are done at the stage of objects selection, where good electrons are first selected and any jets within $\Delta R < 0.4$ are removed, and muons are last selected without overlapping a jet with $\Delta R < 0.2$.

e+jets	1 jet	2 jets	3 jets	$\geq$4 jets
data(+)	12906.0±113.6	3302.0±57.5	858.0±29.3	373.0±19.3
single top (MC)(+)	32.1±9.6	36.6±11.0	14.6±4.4	6.5±1.9
D^+(subtract)	12873.9±114.0	3265.4±58.5	843.4±29.6	366.5±19.4
data(−)	10234.0±101.2	2572.0±50.7	687.0±26.2	336.0±18.3
single top (MC)(−)	19.5±5.9	24.3±7.3	10.3±3.1	6.1±1.8
D^-(subtract)	10214.5±101.3	2547.7±51.2	676.7±26.4	329.9±18.4
$D^+ - D^-$	2659.4±152.5	717.7±77.8	166.7±39.7	36.6±26.8
W+jets (Meas)	14979.6±915.4	3828.2±430.6	752.9±183.2	138.8±102.0
W+jets (MC)	13426.2±8.8	3092.3±3.1	672.1±1.3	185.9±0.7

Table 5.2: Number of selected events as a function of the number of jets in the electron channel. Events are separated into positive and negative charge lepton. In each bin, the contribution of W+jets is obtained as the difference between positively charged and negatively charged lepton events and compared with the MC expectations. Only statistical uncertainties are shown.

5.2 Results and systematic uncertainties to the charge asymmetry method

The charge asymmetry method is applied to the events after top pair selection. The number of events in data and single top MC sample is separated with the lepton charge positive and negative. After data events subtract single top events, the difference between positively charged and negatively charged lepton events is used to obtain the number of W+jets events. The measured results are listed in Table 5.2 for the electron channel and in Table 5.3 in the muon channel. The results measured are compared to W+jets MC expectations. Because of the limitation of statistics, the measured number of events for W+jets with at least four jets has a 72% uncertainty in the electron channel and 29% in the muon channel.

Figure 5.2 shows the fraction of number of W+jets events measured to MC expectation. The statistical uncertainties increase with the number of jets. The fractions are consistent with 1 in 1 jet, 2 jets and 3 jets bins with the given uncertainties. The fractions are quite

μ+jets	1 jet	2 jets	3 jets	$\geq$4 jets
data(+)	11303.0±106.3	2836.0±53.3	708.0±26.6	335.0±18.3
single top (MC)(+)	36.0±10.8	35.7±10.7	17.2±5.2	6.9±2.1
D^+(subtract)	11267.0±106.9	2800.3±54.3	690.8±27.1	328.1±18.4
data(−)	7889.0±88.8	1944.0±44.1	490.0±22.1	248.0±15.7
single top (MC)(−)	21.3±6.4	24.1±7.2	13.7±4.1	4.8±1.5
D^-(subtract)	7867.7±89.0	1919.9±44.7	476.3±22.5	243.2±15.8
$D^+ - D^-$	3399.3±139.1	880.4±70.3	214.6±35.2	84.9±24.3
W+jets (Meas)	17283.3±762.8	4115.9±342.6	885.5±150.1	341.4±100.9
W+jets (MC)	16562.1±9.5	3734.2±3.3	805.2±1.4	228.8±0.8

Table 5.3: Number of selected events as a function of the number of jets in the muon channel. Events are separated into positive and negative charge lepton. In each bin, the contribution of W+jets is obtained as the difference between positively charged and negatively charged lepton events and compared with the MC expectations. Only statistical uncertainties are shown.

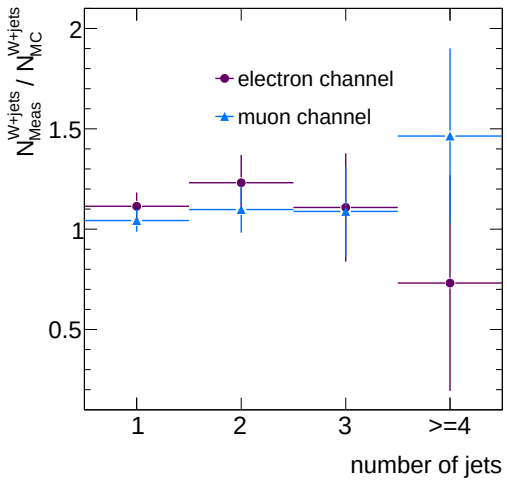


Figure 5.2: The number of W+jets events measured with asymmetry method is compared to the expectation in MC. The fraction of $\frac{N^{W+jets}_{Meas}}{N^{W+jets}_{MC}}$ is shown as a function of the number of jets in the electron and muon channels.

off 1 in the $\geq$4 jets bin in both the electron and the muon channels. The reason is that the difference between positively charged and negatively charged lepton events is smaller than MC expectation, which leads to a smaller number of events measured than that found in MC. In the muon channel, it is the same kind of reason but in the opposite direction. The difference between lepton charges is larger and the measured number of events is larger than that in MC. The uncertainty is greater in the electron channel than in the muon channel because of the smaller difference between positive and negative lepton charge events.

The asymmetry method can also be applied to the events with at least 1 b-tagged jet. The results are shown in Table F.1 and Table F.2 in Appendix F, which are very much limited by statistical uncertainty at the current stage. The method can be applied as soon as enough statistics is collected. Other approaches can be used to get the number of W+jets measured with tagged sample and are presented in Section 5.6.

The systematic uncertainties are considered to the measured number of W+jets background events without the requirement of b-tagged jet. The jet energy scale (JES) can change the distribution of the jet multiplicity. Since the value of asymmetry depends on

the number of jets, the JES is considered as a source of systematic uncertainty. The JES is taken to be ±10%, and the deviation caused for A_{mc} is shown in Table 5.4.

Re. Syst.	e+jets	μ+jets
0 jet	±0.5%	±0.3%
1 jet	±2.1%	±0.1%
2 jets	∓2.2%	∓0.6%
3 jets	∓0.5%	∓2.8%
≥4 jets	∓4.8%	∓1.4%

Table 5.4: A_{mc} systematic uncertainties due to the JES ±10% as a function of the number of jets in the electron and muon channels.

Since A_{mc} is determined from MC, the systematic due to MC sample is considered by using two different MC models, Sherpa and Alpgen. The A_{mc} values found with Alpgen MC are default. A_{mc} is also calculated as a function of jet multiplicity as well as lepton pseudo-rapidity with Sherpa MC and the difference to the default is taken as systematic uncertainty. The A_{mc} results from Sherpa MC are shown in Figure 5.3. The corresponding systematic uncertainty found in ≥4 jets is 7.7% in the electron channel and 3.1% in the muon channel.

The uncertainties due to the imprecise knowledge of PDFs can be evaluated through the comparison between using the default CTEQ6 PDF and other PDF sets. The difference found in A_{mc} can be taken as systematic uncertainty. Because of the lack of MC samples with other PDF sets, the systematic is not yet included. Since the charge asymmetry method is right now limited by statistic uncertainty, the missing of the systematic PDF is not expected to be crucial. The systematic should be added when the statistics increases.

The systematic uncertainties are combined and the final results measured with at least 4 jets in the electron and muon channel are summarized in the following Table 5.5. The large uncertainty is mainly caused by the statistics limitation of $N_{D^+} - N_{D^-}$. The measured number of W+jets background is 138.8 ± 104.6 (e+jets) and 341.4 ± 101.6 (μ+jets).

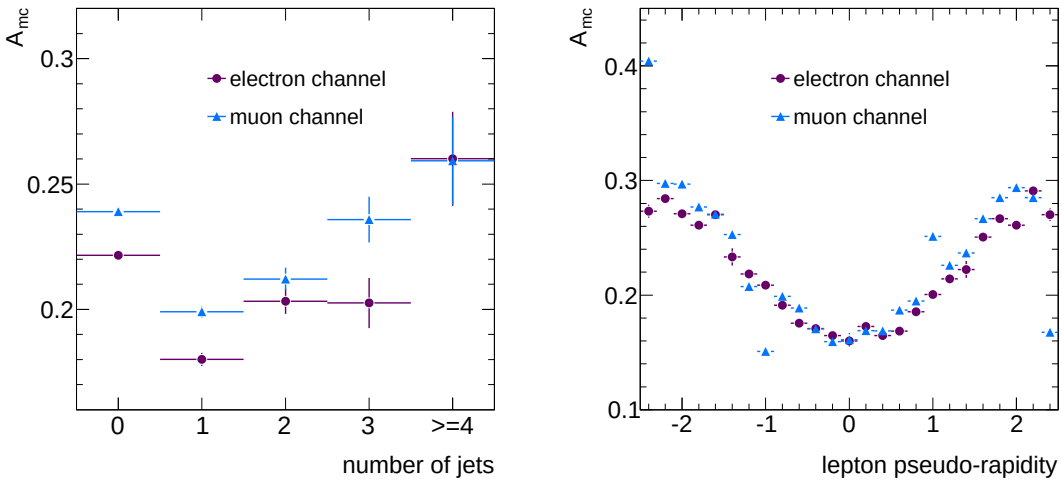


Figure 5.3: The estimated A_{mc} with Sherpa MC as a function of the number of jets (left) and lepton pseudo-rapidity (right) in the electron and muon channels.

	e+jets	μ+jets
A_{mc}	$0.264 \pm 0.020 \pm 0.013 \pm 0.020$	$0.249 \pm 0.019 \pm 0.004 \pm 0.08$
$N_{D^+} - N_{D^-}$	36.6 ± 26.8	84.9 ± 24.3
$W^{\geq 4j}$	138.8 ± 104.6	341.4 ± 101.6

Table 5.5: The asymmetry parameters from MC multiplying the data difference between charge positive and charge negative give measurements of number of W+jets events in electron and muon channel with at least four jets. The uncertainties on A_{mc} are statistical, JES systematic and MC generator Alpgen compared to Sherpa systematic respectively.

Another data driven method is applied to get the W+jets background events, which has smaller uncertainty at the current stage. It is called the Berends Scaling method and is introduced in the next Section 5.3.

5.3 Berends-Giele Scaling method

There is another method to measure the W+jets background with at least 4 jets using data driven approach, which is called Berends-Giele Scaling method, and in short Berends Scaling method. The method states that the ratio of number of events for W+n jets to that for W+$n+1$ jets is constant, which is described in [53] and [54]. Therefore the number of W+jets events with $\geq$4 jets is estimated:

$$W^{\geq 4j} = W^{2j} \cdot \sum_{n=2}^{\infty} (\frac{W^{2j}}{W^{1j}})^n \tag{5.2}$$

In this formula, W^{1j} and W^{2j} are estimated by data subtracting non-W events. Table 5.6 lists the number of events in 1 jet and 2 jets region in the electron and muon channels. The uncertainty assumed is 30% for QCD multijet background, 100% for $t\bar{t}$, 30% for single top, dibosons and Z+jets in both electron and muon channels in the table. The QCD multijet background is estimated with data driven method, while other samples are from MC estimation.

sample	1j (e)	2j (e)	1j (μ)	2j (μ)
observed	23140	5874	19192	4780
QCD multijet (DD)	7908.8±2372.7	2327.5±698.3	1003.9±301.2	332.1±99.6
Z+jet (MC)	225.3±67.6	137.0±41.1	557.8±167.3	163.6±49.1
$t\bar{t}$ (MC)	19.2±19.2	80.0±80.0	20.6±20.6	85.8±85.8
Single Top (MC)	51.6±15.5	60.9±18.3	57.2±17.2	59.7±17.9
DiBoson (MC)	55.7±16.7	44.5±13.4	64.9±19.5	51.2±15.4
W+jet (Meas)	14879.4±2378.7	3224.1±708.9	17487.6±375.8	4087.6±160.7
W+jet (MC)	13446.2±9.0	3109.0±3.1	16575.5±9.6	3749.1±3.4

Table 5.6: Number of events with 1 jet and 2 jets in the electron and muon channels. The uncertainties include statistical and systematic. It is assumed to be 100% for $t\bar{t}$, 30% for the QCD multijet background data driven method and 30% for other processes from MC estimation. The measured W+jets is obtained by data subtracting non-W events.

The fraction $\frac{W^{2j}}{W^{1j}}$ in MC shows no deviation between the electron and the muon channel. The fraction measured is compatible with that found in MC. Consequently, the sum of the fraction until infinite $\sum(\frac{W^{2j}}{W^{1j}})^n$ is calculated with the measured value of the fraction. The values of the fraction and the sum are listed in Table 5.7, where the uncertainties on the measured results include statistical and systematic.

	e+jets	μ+jets
$\frac{W^{2j}}{W^{1j}}$ (MC)	0.2312±0.0003	0.2262±0.0003
$\frac{W^{2j}}{W^{1j}}$ (Meas)	0.2167±0.0590	0.2337±0.0110
$\sum(\frac{W^{2j}}{W^{1j}})^n$ (Meas)	0.0599±0.0372	0.0713±0.0078

Table 5.7: The fraction $\frac{W^{2j}}{W^{1j}}$ found in MC is compared to the results in data. The difference between e+jets and μ+jets is around 2%. Therefore the results measured in muon channel are further used in W+jets measurement for both channels.

In the muon channel, less QCD multijet background is found and it leads to lower uncertainty than that in the electron channel. Since there is no bias found in MC study between the two channels, the result of $\sum(\frac{W^{2j}}{W^{1j}})^n = 0.0713 \pm 0.0078$ measured in the muon channel is applied to both electron and muon channels. It is further used in the Formula 5.2 to get the number of W+jets events with at least 4 jets, which is detailed in the next Section 5.4.

5.4 Results and systematic uncertainties to the Berends Scaling method

The measurement of W+jets with at least 4 jets using the Formula 5.2 needs the parameters, the number of events in 2 jets sample and $\sum(\frac{W^{2j}}{W^{1j}})^n$, which are calculated in the previous section. Concerning the systematics associated with $\sum(\frac{W^{2j}}{W^{1j}})^n$, one important systematic uncertainty is the purity of the W+jets sample in 1 jet and 2 jets region. This systematic has been counted when the non-W events are subtracted from data shown in Table 5.6. The leading systematic uncertainty to it is the assumption that the ratio $\frac{W+(n+1)jets}{W+njets}$ is constant. It is evaluated to 24% from the results in [55].

After the total systematic uncertainty is included, the parameter $\sum(\frac{W^{2j}}{W^{1j}})^n = 0.0713 \pm 0.0188$ in the muon channel. It is multiplied by the number of W+jets background events in 2 jets sample, where it is 3224.1 ± 708.9 in the electron channel and 4087.6 ± 160.7 in the muon channel. After these numbers are applied to the Formula 5.2, the results of W+jets background with at least 4 jets is obtained in the Berends Scaling method:

$$W^{\geq 4j}(e) = 229.9 \pm 78.9, \quad W^{\geq 4j}(\mu) = 291.4 \pm 77.7 \tag{5.3}$$

The uncertainty to the final result includes statistical and systematics. With the Berends scaling method, the total uncertainty is smaller than the charge asymmetry method at the current stage. However, it will be limited by 24% systematic uncertainty, while the charge asymmetry method is becoming more challenging with more statistics.

Finally, the number of W+jets background events is 229.9 ± 78.9 in the electron channel and 291.4 ± 77.7 in the muon channel. The measured number of events is higher than that in MC expectation, which is 21% higher in the electron channel and 40% in the muon channel. By looking at the Table 5.2 in Section 5.2, one finds that the number of events in 2 jets sample has lower uncertainty with the charge asymmetry method, which can be used to Formula 5.2 and get a better result in the electron channel. The so called combined method is presented in the next Section 5.5.

5.5 Combination of charge asymmetry and Berends scaling method

In the electron channel, W+jets background measurement using Berends Scaling method using the Formula 5.2 is largely restricted by the large uncertainty to the number of events in 2 jets sample, which is mainly from the contamination of QCD multijet events. In the charge asymmetry method, the number of W+jets in 2 jets region has lower uncertainty than that in Berends Scaling method, which is $W^{2jet} = 3828.2 \pm 430.6$ in the electron channel shown in the Table 5.2. Correspondingly, it is $W^{2jet} = 3224.1 \pm 708.9$ in Berends Scaling method.

When the W^{2jet} found in the charge asymmetry method multiplies the sum of the ratios measured in Berends Scaling method in the muon channel, $\sum(\frac{W^{2j}}{W^{1j}})^n = 0.0713 \pm 0.0188$, using the Formula 5.2, one gets the number of W+jets events with at least 4 jets as $W^{\geq 4j} = 273.0 \pm 78.2$ in the electron channel. This is the result using the so called combined method. It has an uncertainty of 29% including statistical and systematics, which is smaller than that in Berends Scaling method 34% and that in charge asymmetry method 75%.

In the muon channel, the combined method is equally applied like in the electron channel. The number of events in 2 jets region is $W^{2jet} = 4115.9 \pm 342.6$ in the charge asymmetry method, which is compatible with that in the Berends Scaling method 4087.6 ± 160.7. The uncertainty is larger in the charge asymmetry method 8% than that in the Berends Scaling method 4%. The total uncertainty of the W+jets background measured with the combined method is compatible with that in Berends Scaling method, because it is dominated by the uncertainty of $\sum(\frac{W^{2j}}{W^{1j}})^n$, 26%. It is the parameter used in both methods. Finally, the measured W+jets in the combined method is $W^{\geq 4j} = 293.5 \pm 81.1$ in the muon channel.

With the method combining the charge asymmetry and Berends Scaling, the uncertainty is reduced in the electron channel. The measured number of W+jets background events is used in Chapter 6. It is subtracted from data to get the measurement of the number of signal events. The number of W+jets background using the combined method is finally listed in the following:

$$W^{\geq 4j}(e) = 273.0 \pm 78.2, \quad W^{\geq 4j}(\mu) = 293.5 \pm 81.1 \tag{5.4}$$

5.6 W+jets in the b-tagged selection

In order to reduce the systematic uncertainty due to the knowledge of the background, one will need to use the events with at least one b-tagged jet. Therefore, the W+jets background with at least 1 b-tagged jet needs to be measured. As seen in Table F.1 and Table F.2 of the tagged results of the charge asymmetry method, the direct measurements are possible but greatly limited by the statistical uncertainty about 300% in the electron channel and 200% in the muon channel with the charge asymmetry method. The following method will use the b-tagging information in the 2 jets region and extrapolate it to $\geq$4 jets region. After the number of W+jets with at least 4 jets is measured, the following formula is applied to get the number of events among which at least one jet is b-tagged:

$$W_{tag}^{\geq 4j} = W^{\geq 4j} \cdot f_{tag}^{\geq 4j} \tag{5.5}$$

where $f_{tag}^{\geq 4j}$ is the fraction to be at least one b-tagged jet among the $\geq$ 4 jets sample. The factor can be calculated as:

$$f_{tag}^{\geq 4j} = f_{tag}^{2j} \cdot f_{corr}^{2\rightarrow\geq 4} \tag{5.6}$$

where f_{tag}^{2j} is the fraction with at least 1 b-tagged measured in 2 jets sample, and $f_{corr}^{2\rightarrow\geq 4}$ is the correction factor when the tagged fraction in 2 jets is applied to $\geq$4 jets sample. First, the fraction f_{tag}^{2j} is measured with an estimation of the number of W+jets events before and after requiring at least 1 b-tagged jet by subtracting other processes from the observed number of events in 2 jets sample. Table 5.8 shows the number of event in 2 jets sample with at least one b-tagged jet. By comparing the total uncertainties of number of events in 2 jets bin, one can tell that for the electron channel the charge asymmetry method has lower uncertainty. In the muon channel, the result with lower uncertainty is from Berends scaling method found in Table 5.6. The results for f_{tag}^{2j} in separated and combined channels are listed in Table 5.9, where the ones in muon channel is finally used because of its lower uncertainty. It is reasonable, because the factors f_{tag}^{2j} found in MC in both channels are only 2% different, which states that the factor f_{tag}^{2j} is not sensitive to the difference in the two

sample	2j (e)	2j (μ)
observed	339	304
QCD multijet (DD)	186.0±55.8	23.6±7.1
Z+jet (MC)	46.8±46.8	50.4±50.4
$t\bar{t}$ (MC)	3.3±3.3	4.4±4.4
Single Top (MC)	30.7±9.2	30.8±9.2
DiBoson (MC)	2.7±0.8	3.3±1.0
W+jet (Meas)	69.5±75.7	191.5±54.6
W+jet (MC)	100.9±0.5	119.6±0.6

Table 5.8: Number of events in 2 jets bin among which ≥1 b-tagged to compare with the MC expectation in both electron channel and muon channels. 100% uncertainty is assumed for $t\bar{t}$ and 30% for QCD multijet and all the other MC samples.

channels. The combined channel result can be used once it is not limited by the uncertainty on number of b-tagged events in electron channel.

The correction factor $f_{corr}^{2\to\geq 4}$ is measured in ALPGEN W+jets events and is defined as $f_{corr}^{2\to\geq 4} = f_{tag}^{\geq 4j}/f_{tag}^{2j}$, which is 2.8±0.8(syst.). The total systematic uncertainty arises from the uncertainties on the assumed flavor composition of the 2 jets sample before requiring b-tagging, the uncertainties on the b-tagging efficiency for b-jet, c-jet and light jet, and the uncertainty on the ratio of the fractions in the 2 jets and ≥4 jets samples for W+heavy flavor processes (W+bb+jets, W+cc+jets and W+c+jets).

By applying the results of f_{tag}^{2j} and $f_{corr}^{2\to\geq 4}$ to Equation 5.5, we get the estimated $f_{tag}^{\geq 4j}$ to be 0.131±0.053. The uncertainty includes the errors from f_{tag}^{2j} and $f_{corr}^{2\to\geq 4}$. The number of W+jets tagged events is shown in Table 5.10 for $W^{\geq 4j}$ measured both from charge asymmetry method and Berends Scaling method.

With a large uncertainty on the measured number of W+jets, the results from the three methods shown in Table 5.10 are compatible with the expectation in MC. The number of W+jets background events with at least one b-tagged is measured to be 35.8 ± 17.8 in

sample	2j (e)	2j (μ)	2j (combined)
W+jet (Tagged)	69.5±75.7	191.5±54.6	261.0±93.3
W+jet (PreTag)	3828.2±486.5	4087.6±160.7	7915.8±512.4
f_{tag}^{2j} (Meas)	0.0182±0.0199	0.0468±0.0135	0.0330±0.0120
f_{tag}^{2j} (MC)	0.0325±0.0002	0.0319±0.0002	0.0321±0.0002

Table 5.9: Number of events in 2 jets bin for W+jets ≥1 b-tagged and pretag to measure the b-tagging fraction in both electron channel and muon channels. The pretag in e+jets is measured with Asymmetry method, where total systematic uncertainties are included. The pretag in μ+jets is measured with data subtracting backgrounds. The fractions f_{tag}^{2j} found in MC is presented to see the consistency between channels and be compared with measured results.

$W_{tag}^{\geq 4j}$	e-chan	μ-chan
Asymmetry	18.2±15.6	44.7±22.5
Berends Scale	30.1±16.0	38.2±18.5
Combined	35.8±17.8	38.4±18.8
MC	15.7±0.2	20.7±0.2

Table 5.10: Number of events for W+jets measured with ≥4 jets and ≥1 b-tagged in both electron and muon channels, where $W^{\geq 4j}$ before b-tagging is measured with asymmetry method, Berends Scaling method and the method combined with both.

the electron channel and 38.4 ± 18.8 in the muon channel, where the number of W+jets events with at least 4 jets without b-tagging requirement is measured with the combined method. Finally, the b-tagged number of W+jets events is applied to the method described in Chapter 6 to measure the number of signal events. The $\sigma_{t\bar{t}}$ result using the W+jets originated from the combined method is finally chosen in this analysis.

5.7 W/Z ratio method

Another W+jets measurement method is W/Z ratio method which is more likely to be used in the future when the integrated luminosity is high enough. The W/Z ratio method uses the fact that the ratio of W and Z production is relatively better understood than W+jets itself, which can be found in [56] and [57]. Since in the signal region with at least 4 jets, the number of Z+jets events is better measured when there are enough statistics, the result of Z+jets can be extrapolated to W events. Therefore the W+jets background can be estimated using the Formula 5.7.

$$(\frac{W^{\geq 4j}}{W^{1jet}})_{data} = (\frac{Z^{\geq 4j}}{Z^{1jet}})_{data} \cdot C_{mc} = (\frac{Z^{\geq 4j}}{Z^{1jet}})_{data} \cdot (\frac{W^{\geq 4j}}{W^{1jet}})_{mc} \cdot (\frac{Z^{1jet}}{Z^{\geq 4j}})_{mc} \tag{5.7}$$

Here, 1 jet bin is used as a control region, for both W+jets and Z+jets events. The selection cuts for Z+jets is a little different, where it will require two opposite charged leptons instead of one, etc.

Concerning the systematic uncertainties, the consistency of C_{mc} based on the comparison of different Monte Carlo generators yields a systematic uncertainty of 12% [58]. And, for the time being, the method is limited by the statistical uncertainty of $Z^{\geq 4j}$ events from data. With a luminosity of 33.7 pb^{-1} data, a preliminary study shows the W+jets measured are compatible with charge asymmetry method and Berends Scaling method, which has an uncertainty 30% in the electron and muon channels. The error bars are competitive with the ones in Berends Scaling. With the accumulated statistics, one can expect the method can give a better measurement of W+jets background.

5.8 Summary

The W+jets measurement with the methods introduced in this chapter can give early results, which is quite important to the measurement of $t\bar{t}$ cross section. The W+jets background events are selected with the top pair selection cuts, including one isolated lepton, large missing transverse energy and at least 4 jets. A summary of the results is shown in Table 5.11. The uncertainties on the W+jets measurements finally lead to the uncertainties on the number of events found for signal. Concerning the uncertainties using the methods at this stage, the results are consistent with different methods.

W+jets	method	e+jets	$\frac{\Delta N}{N}$	μ+jets	$\frac{\Delta N}{N}$	e/μ+j	$\frac{\Delta N}{N}$
pretag	Asymmetry	138.8±104.6	75.4%	341.4±101.6	29.8%	480.2±145.8	30.4%
	Berends	229.9±78.9	34.3%	291.4±77.7	26.7%	521.3±110.7	21.2%
	Combine	273.0±78.2	28.6%	293.5±81.1	27.6%	566.5±112.7	19.9%
tagged	Asymmetry	18.2±15.6	85.7%	44.7±22.5	50.3%	62.9±27.4	43.5%
	Berends	30.1±16.0	53.2%	38.2±18.5	48.4%	68.3±24.5	35.8%
	Combine	35.8±17.8	49.7%	38.4±18.8	49.0%	74.2±25.9	34.9%

Table 5.11: Measured pretag W+jets using asymmetry method, Berends scaling or the combined. The tagged numbers of events are also presented according to the pretag methods. The relative systematics are listed beside, which stand for the total uncertainties statistical and systematical.

The measured number of W+jets events is consequently used to measure the top pair cross section using the Formula 1.1 to extract the number of $t\bar{t}$ signal events. All results from the three W+jets methods are used to be compared with each other. In Table 5.11, the result has lower uncertainty with the combined method in the electron and muon combined channel, which is 566.5 ± 112.7 for pretag and 74.2 ± 25.9 for tagged W+jets background in at least 4 jets region. The application using the measured W+jets events is detailed in Chapter 6 to measure N_{sig}.

Chapter 6

SIMULTANEOUS MEASUREMENT OF TOP PAIR CROSS SECTION MEASUREMENT AND B-TAGGING EFFICIENCY

In the Standard Model, a top quark decays into a b-quark and a W boson. The b-tagging is important select b-jet from all jets. Having b-tagged jets is a good signature for $t\bar{t}$ events. With the events selected as described in Chapter 4, there are several approaches to extract the $t\bar{t}$ cross section through cut and count method.

- With a limited statistics, one can subtract the number of background events without the requirement of b-tagged jet. With a S/B ratio of 1, this leads to an uncertainty limited to 30%, if the background is known to this level.

- With an additional requirement of at least one b-tagged jet, one can have a cleaner signal sample. The b-tagging efficiency can be measured in data using various methods. The measured ϵ_b needs to be transferred to $t\bar{t}$ through the weighting of each jet. This has been used in the first top observation paper [45].

- A method called tag counting uses the b-tagging information to simultaneously measure the number of signal events and the efficiency of b-tagging. With this method, no b-tagging cut is applied, but the information of the number of b-tagged jets multiplicity distribution is used.

In this chapter, the tag counting method is introduced. Afterwards, the σ_{ttbar} is measured with the number of signal events within this method.

A b-jet has its own properties which can be used to tag it, for example tracks inconsistent with the primary vertex, secondary vertex, lepton inside the jet from b-quark leptonic decay, etc. There are several taggers built by using these properties of b-jets. The one used in this

analysis is called SV0, which is introduced in Section 3.3.5. A jet having a SV0Weight $>$ 5.72 is regarded as a b-tagged jet.

6.1 The tag counting method

The tag counting method uses $t\bar{t}$ events to measure the b-tagging efficiency and the number of total $t\bar{t}$ events simultaneously in the lepton+jets decay channels.

The ϵ_b is defined to be the efficiency of a b-jet originated from a b-quark to be tagged as a b-tagged jet. Similarly, ϵ_c and ϵ_l are the efficiency of a c-jet (c-quark jet) and light jet (light quark jet or gluon jet) to be tagged as a b-tagged jet. Ideally, if top pair events produce exactly 2 b-jets and only b-jets can be tagged, then the expected number of events with 2 b-tagged jets is $N_{sig} \cdot \epsilon_b^2$, and with exactly one b-tagged jet is $2N_{sig} \cdot \epsilon_b(1-\epsilon_b)$, where N_{sig} is the number of $t\bar{t}$ signal events. Thus, the measurement of events with exactly one and exactly 2 b-tagged jets can be used to determine the b-tagging efficiency. At the same time, the corresponding number of selected events can be used to measure the top pair cross section.

In reality, the number of b-jets can be different from 2, since a b-jet from top decay cannot always be selected or additional b-jets can be produced through gluon radiation. Moreover, c-jets and light jets, which come from the hadronic W decay or ISR/FSR gluons, can also be tagged as b-jets, and consequently contribute to the number of tagged jets in the event. These effects are taken into account through the fractions (F_{ijk}) of number of events containing i b-jet(s), j c-jet(s) and k light jet(s) among total N_{sig} events. F_{ijk} is estimated from MC, among which the ones having non-zero fractions are shown in Table 6.1. The statistical uncertainties are listed in MC. It is interesting to notice that the contribution of 2 b-jets with 2 light jets only gives about one-fourth of the total events.

So, the expected number of events with n b-tagged jet(s) can be calculated with the Formula 6.1 combining all the possible contributions:

$$N^{exp}(\epsilon_b, N_{sig}) = N_{sig} \cdot \sum_{i,j,k} F_{ijk} \sum_{i'+j'+k'=n} A_i^{i'} \cdot {\epsilon_b}^{i'} \cdot (1-\epsilon_b)^{i-i'} \cdot A_j^{j'} \cdot {\epsilon_c}^{j'} \cdot (1-\epsilon_c)^{j-j'} \cdot A_k^{k'} \cdot {\epsilon_l}^{k'} \cdot (1-\epsilon_l)^{k-k'} \quad (6.1)$$

where $A_i^{i'}$ is the number of arrangements $\frac{i!}{i'! \cdot (i-i')!}$, and i' is the number of b-tagged jets of a given flavor while i is the number of jets before applying b-tagging. $N^{exp}(\epsilon_b, N_{sig})$ is the number of events expected in the n^{th} jet bin.

N_{jet}	F_{ijk}	e+jet	μ+jet	N_{jet}	F_{ijk}	e+jet	μ+jet
4	F_{004}	0.54±0.03%	0.50±0.03%	5	F_{221}	0.39±0.03%	0.38±0.03%
4	F_{013}	0.37±0.03%	0.38±0.03%	5	F_{302}	0.27±0.02%	0.26±0.02%
4	F_{103}	8.85±0.12%	8.89±0.12%	5	F_{311}	0.19±0.02%	0.18±0.02%
4	F_{112}	5.85±0.10%	5.79±0.10%	6	F_{105}	0.69±0.03%	0.68±0.03%
4	F_{121}	0.20±0.02%	0.22±0.02%	6	F_{114}	0.62±0.03%	0.57±0.03%
4	F_{202}	28.81±0.19%	28.77±0.19%	6	F_{204}	4.14±0.08%	4.08±0.08%
4	F_{211}	15.34±0.15%	15.09±0.15%	6	F_{213}	2.92±0.07%	2.85±0.07%
4	F_{220}	0.21±0.02%	0.16±0.02%	6	F_{222}	0.22±0.02%	0.22±0.02%
4	F_{301}	0.34±0.02%	0.30±0.02%	6	F_{303}	0.11±0.01%	0.11±0.01%
4	F_{310}	0.06±0.01%	0.08±0.01%	6	F_{312}	0.08±0.01%	0.10±0.01%
5	F_{005}	0.13±0.02%	0.14±0.02%	7	F_{106}	0.18±0.02%	0.14±0.02%
5	F_{014}	0.14±0.02%	0.11±0.01%	7	F_{115}	0.14±0.02%	0.10±0.01%
5	F_{104}	2.91±0.07%	3.04±0.07%	7	F_{205}	0.98±0.04%	0.89±0.04%
5	F_{113}	2.23±0.06%	2.20±0.06%	7	F_{214}	0.69±0.04%	0.72±0.04%
5	F_{122}	0.12±0.01%	0.13±0.02%	7	F_{223}	0.07±0.01%	0.07±0.01%
5	F_{203}	12.86±0.14%	13.15±0.14%	8	F_{206}	0.18±0.02%	0.19±0.02%
5	F_{212}	8.60±0.12%	8.79±0.12%	8	F_{215}	0.10±0.01%	0.12±0.01%

Table 6.1: The non-zero F_{ijk} from $t\bar{t}$ not-fully-hadronic MC sample in the electron channel and in the muon channel. The statistical uncertainties are listed.

A likelihood as in Formula 6.2 is used to fit, where N_n is the number of events found in the n jet(s) bin in data.

$$L = \Pi(Poisson(N_n, N^{exp})) \tag{6.2}$$

Figure 4.8 shows that the 0 b-tagged jet bin is largely contaminated by backgrounds and eventually more affected by the background level uncertainty. We thus compare the results of the fit using all bins from 0 b-tagged jet to 3 b-tagged jets (N0∼N3) and removing the 0 b-tagged bin (N1∼N3). Besides, only N_{sig} and ϵ_b are allowed to fluctuate and are fitted, where $\epsilon_b = 52.5 \pm 0.1\%$ in MC. ϵ_c and ϵ_l are fixed by the values found in MC, which is $\epsilon_c = 11.7 \pm 0.1\%$ and $\epsilon_l = 0.46 \pm 0.01\%$.

Before fitting the function, the method requires a background subtraction bin by bin of b-tagged jet multiplicity. So, it is important to measure the QCD multijet and W+jets backgrounds, especially the fraction containing heavy flavor, which have been considered separately as systematics. The QCD multijet and W+jets backgrounds are subtracted using the number of events measured with the data driven methods. Other backgrounds, which are much smaller than the dominant ones, are subtracted using the expectation from MC.

6.2 Likelihood fit results

In Chapter 5, the W+jets background is measured with asymmetry method, Berends Scaling method and the combined method, before and after requiring at least one b-tagged jet. In the coming sections, the fitted results are shown by using W+jets background measured with the combined method, because it measures W+jets events with lower stat.+syst. uncertainties. The corresponding results by using W+jets background from the other two methods are listed in Appendix G.

In the tag counting method, backgrounds are required to be subtracted bin by bin of b-tagged jet multiplicity. For the W+jets background, the discrepancy between the pretag and tagged number of events give an estimation of events without any b-tagged jet. The fractions of events with exactly 1, 2 and 3 b-tagged jets are computed from W+jets MC, which is shown in Table 6.2. The number of events for W+jets as function of b-tagged jet multiplicity is shown in Table 6.3.

channel	1 btag	2 btag	3 btag
e+jets	92.2±9.2%	7.5±1.7%	0.3±0.2%
μ+jets	91.8±7.9%	7.3±1.4%	0.9±0.6%

Table 6.2: Fraction of n^{th} b-tagged jet with respect to tagged sample from W+jets MC, only statistical uncertainty is presented.

After subtracting the number of background events from data, one can apply the fit to obtain the final measured number of signal events in electron, muon and the combined channel. The results for N_{sig} and ϵ_b are shown in Table 6.4, where only statistical uncertainties are listed. The results of N_{sig} by fitting N0∼N3 are quite different with W+jets background measured from asymmetry method or from the other two methods as shown in Table G.2 in Appendix G, since the number of W+jets events are varying from the other two methods especially in the no b-tagged jet bin. The fluctuation in the background measurement also reflects in the measurement of ϵ_b especially in the electron channel using N0∼N3. Therefore,

Combined	0 btag	1 btag	2 btag	3 btag
e+jets	237.2±78.3	33.0±16.7	2.7±1.5	0.1±0.1
μ+jets	255.1±83.3	35.7±17.5	2.8±1.5	0.3±0.3

Table 6.3: The number of W+jets events with b-tagged jet multiplicity. The pretag number of events is measured with the charge asymmetry and the Berends Scaling method. The number of no b-tagged events is measured by pretag events subtracting tagged events. The number of tagged events times the fractions from Table 6.2, which gives the number of events for exactly 1, 2 and 3 b-tagged events.

	Fit N0∼N3			Fit N1∼N3		
	e+jets	μ+jets	l+jets	e+jets	μ+jets	l+jets
N_{sig}	148.1±12.2	218.3±14.8	366.4±19.2	203.4±19.8	221.5±19.6	424.6±27.8
ϵ_b	80.9±3.2%	61.1±2.8%	69.0±2.1%	57.0±5.1%	60.2±4.7%	58.7±3.5%

Table 6.4: Fitted results of N_{sig} and ϵ_b results with or without the no b-tagged events in the electron, muon and the combined channel. Statistical uncertainties in data fitting are presented.

without using the no b-tagged jet bin, the results have lower systematic uncertainties.

In this section, N_{sig} and ϵ_b are simultaneously fitted. The result in the combined lepton+jets channel fitting N1∼N3 has lower statistical uncertainty, which is 424.6 ± 27.8 for the total number of signal events and 58.7 ± 3.5% for b-tagging efficiency. The systematic uncertainties fitting N1∼N3 are detailed in Section 6.3 in the electron, muon and combined channel.

6.3 Systematics to the simultaneous measurement of ϵ_b and N_{sig}

The systematic uncertainties of ϵ_b and N_{sig} of the counting method of fitting 1 b-tagged to 3 b-tagged are in Table 6.5 and Table 6.6. The corresponding results of fitting 0 b-tagged to 3 b-tagged are summarized in Table G.4 and Table G.5 in Appendix G. The results from fitting using the no b-tagged jet bin have larger total systematic uncertainties for both ϵ_b and N_{sig}, because the dominant systematic uncertainties, the W+jets and QCD multijet background level, are about 3 times larger than those for the results fitted without the no b-tagged jet bin. Therefore, the results from fitting 1 to 3 b-tagged jet bins are considered as default. The W+jets listed in the tables are measured from the combination of asymmetry method and Berends scaling method introduced in Section 5.5, which gives lower uncertainty especially in the electron channel.

The systematics related to the usage of different MC samples are considered separately, showing the effect of ISR/FSR, parton shower, MC generator and pile up effect. The ISR/FSR effect changes the jet multiplicity in each event. Therefore, the acceptance varies mainly due to the cut of at least 4 jets. With the method of tag counting, the fitted ϵ_b also changes correspondingly. The ISR/FSR MC samples have Matrix element and Parton Shower being ACerMC+pythia, which are compared to the nominal MC sample with ACerMC+pythia. Each of the ISR/FSR samples are separated into 40 pseudo experiments using F_{ijk} from the nominal MC. The results of N_{sig} and ϵ_b are fitted using the pseudo experiments. The largest positive and negative deviations with respect to the nominal sample results are used as the systematics. The systematics are symmetrized in the end.

The MC generator item reflects the effect by using different MC, where Powheg+Herwig is compared to MC@NLO+Herwig, both of which produce final states at NLO accuracy. The fitted results using pseudo experiments of Powheg sample are taken to be the uncertainties with respect to the corresponding results from the fitted results of MC@NLO sample, which is also used to quote the F_{ijk}. The effect is one sided, either positive or negative, and is symmetrized to the other side.

The parton shower reminds us the difference between MC simulation hadronization. Pythia is compared to Herwig accompanying the generator Powheg. The pile-up of events

are also considered using MC@NLO with a different tag. The systematics of parton shower and pile-up are also symmetrized before being put into the tables.

Some systematic uncertainties on N_{sig} are found to be largely different in the electron and muon channel, for example, ISR/FSR and pile up in Table 6.6. It is caused by the limitation of number of MC events, thus not enough pseudo experiments for those samples.

relative syst.	e-chan	μ-chan	combined
ISR/FSR	3.9%	4.2%	3.9%
Parton shower	1.4%	6.3%	3.9%
MC generator	1.4%	1.5%	0.3%
Pile up	0.6%	0.9%	0.5%
JES ($\pm$7%)	0.5%	0.5%	0.5%
ϵ_c ($\pm$100%)	7.3%	7.5%	7.4%
ϵ_l ($\pm$100%)	2.2%	2.3%	2.3%
QCD+HF total ($\pm$50%)	7.9%	0.8%	4.2%
QCD+HF $\geq$2 b ($\pm$100%)	2.9%	1.7%	2.2%
W+HF total (Combine)	7.7%	7.3%	5.4%
W+HF $\geq$2 b ($\pm$100%)	3.3%	4.1%	3.4%
Other Bkg ($\pm$30%)	0.5%	0.4%	0.5%
Other Bkg ($\epsilon_b\pm$10%)	0.1%	0.4%	0.3%
Total Systematic	14.8%	14.0%	12.4%
Statistical (33.7 pb^{-1})	9.3%	8.1%	6.2%

Table 6.5: Summary of systematic and statistical uncertainties for ϵ_b by fitting 1 to 3 b-tagged bins. The total systematic uncertainty is quadratically combined with the systematic uncertainties in the list.

The jet energy scale (JES) can change the fraction of the jet content in F_{ijk}. Besides, since the b-tagging efficiency also depends on the p_{T} of the jets, JES can cause a systematic on the fitted ϵ_b and N_{sig}, which will give about 1% for ϵ_b and 14% for N_{sig}. This effect

relative syst.	e-chan	μ-chan	combined
ISR/FSR	7.8%	7.0%	7.0%
Parton shower	1.3%	8.5%	5.0%
MC generator	8.9%	6.4%	9.5%
Pile up	6.8%	1.2%	5.4%
JES ($\pm$7%)	9.6%	9.6%	9.7%
ϵ_c ($\pm$100%)	2.9%	3.1%	2.9%
ϵ_l ($\pm$100%)	0.9%	0.8%	0.9%
QCD+HF total ($\pm$50%)	15.1%	2.5%	8.6%
QCD+HF $\geq$2 b ($\pm$100%)	1.7%	1.0%	1.3%
W+HF total (Combine)	15.7%	14.9%	11.0%
W+HF $\geq$2 b ($\pm$100%)	1.7%	2.2%	1.9%
Other Bkg ($\pm$30%)	1.9%	1.6%	1.9%
Other Bkg ($\epsilon_b\pm$10%)	0.3%	<0.1%	<0.1%
Total Systematic	27.8%	22.4%	22.3%
Statistical (33.7 pb^{-1})	9.6%	8.8%	6.5%

Table 6.6: Summary of systematic and statistical uncertainties for N_{sig} by fitting 1 to 3 b-tagged bins. The total systematic uncertainty is quadratically combined with the systematic uncertainties in the list.

is shown in Figure 6.1. The N_{sig} is more affected by JES, because the cut of the number of jets with the modified jet energies can change the selected number of events. The fitted number of events is consequently changed.

The c-jet and light jet tagging efficiency (ϵ_c, ϵ_l) are taken from $t\bar{t}$ MC as inputs during the fitting. Both of the efficiencies can be measured in other methods [59]. The errors on ϵ_c and ϵ_l are taken to be 100%, and the systematic caused by ϵ_c is about three times larger than that by ϵ_l, which are shown in Figure 6.2 and Figure 6.3.

A significant background is seen in the figure of number of b-tagged jets. The background

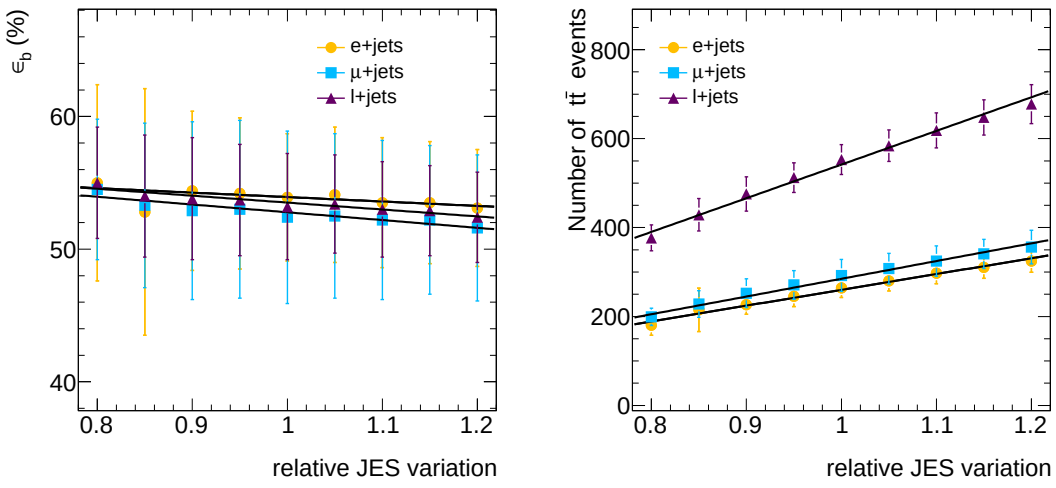


Figure 6.1: The jet energy scale systematic to ϵ_b and N_{sig} by fitting 1 to 3 b-tagged in e-chan, μ-chan and the combined channel of e-chan and μ-chan with the counting method

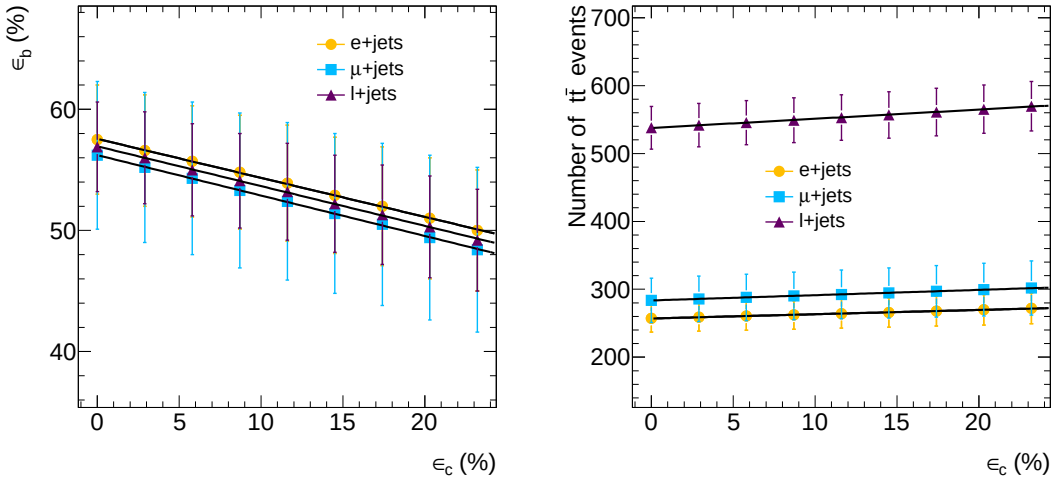


Figure 6.2: The c-jet tagging efficiency systematic to ϵ_b and N_{sig} by fitting 1 to 3 b-tagged in e-chan, μ-chan and the combined channel of e-chan and μ-chan with the counting method

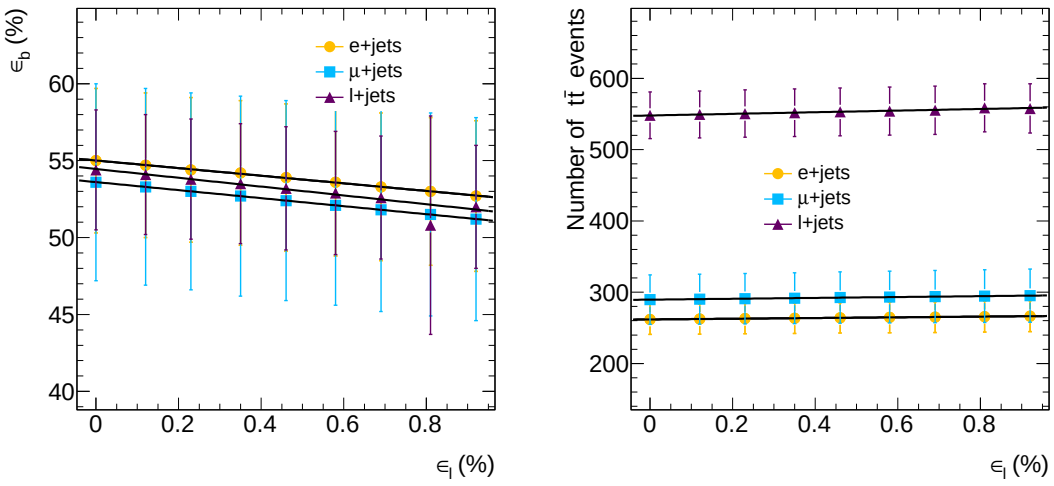


Figure 6.3: The light jet tagging efficiency systematic to ϵ_b and N_{sig} by fitting 1 to 3 b-tagged in e-chan, μ-chan and the combined channel of e-chan and μ-chan with the counting method

level is one of the main systematics. The backgrounds of QCD multijet, W+jets, single top, diboson can contaminate the number of jet bins, especially for the bins which have a large contribution. Therefore background levels of QCD multijet and W+jets are separately considered, while other backgrounds are combined to fluctuate a certain level.

QCD multijet is measured with "Matrix Method" in both electron channel and muon channel. As discussed in Section 4.3, a 30% uncertainty is estimated. The systematic on at least 1 b-tagged jet is estimated to be 50%, if the pretag QCD multijet measurement and the b-tagging efficiency are counted. An uncertainty of 100% for at least 2 b-tagged jets is assumed for QCD multijet background. For fitting N1∼N3, only QCD multijet background uncertainties with btag are considered. The Figure 6.4 shows the effect of QCD multijet heavy flavor level with at least 1 b-tagged jet by fitting 1 to 3 b-tagged.

W+jets is taken from the combined method which gives ±30% pretag events in each channel. A large uncertainty is found for fitting including 0 btag events. Concerning fitting N1∼N3, no 0 btag events is needed. Therefore, the W+jets events with at least 1 btag events contribute the systematic, which is about 50% uncertainty. For the fraction of the number

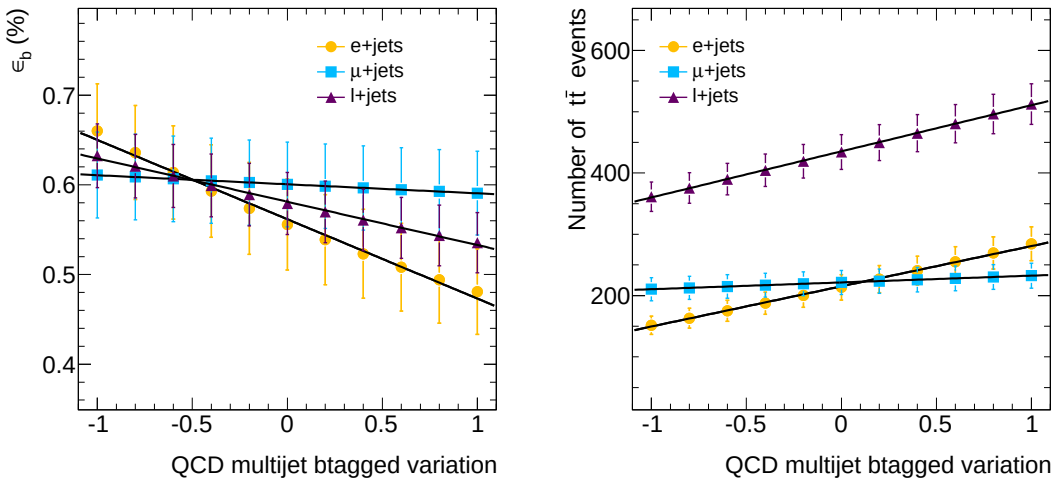


Figure 6.4: The QCD multijet background heavy flavor level systematic with at least one btag to ϵ_b and N_{sig} by fitting 1 to 3 b-tagged in e-chan, μ-chan and the combined channel of e-chan and μ-chan.

of events with no less than 2 btag jets, no measurement is accessible, an uncertainty of 100% is assumed. The W+jets level has the effect to the fitted results is shown in Figure 6.5. With W+jets measured in asymmetry or Berends scaling method, larger uncertainties are found especially in electron channel. Table 6.7 summarizes the systematics with all three methods.

Finally, other backgrounds are taken to fluctuate ±30%, which have much smaller effect, shown in Figure 6.6.

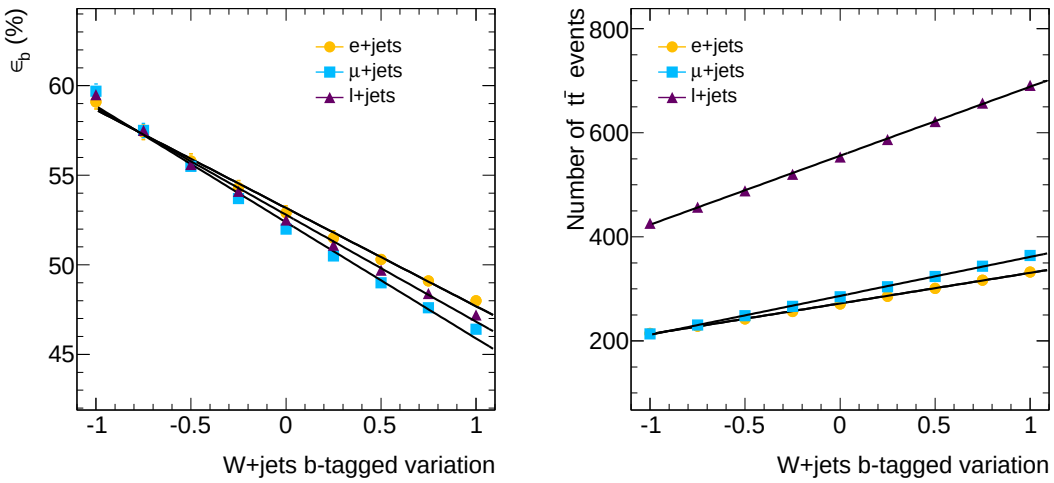


Figure 6.5: The W+jets background with at least 1 btag jet events systematic to ϵ_b and N_{sig} by fitting 1 to 3 b-tagged in e-chan, μ-chan and the combined channel of e-chan and μ-chan with the counting method. Tagged W+jets number of events is measured with the combined method.

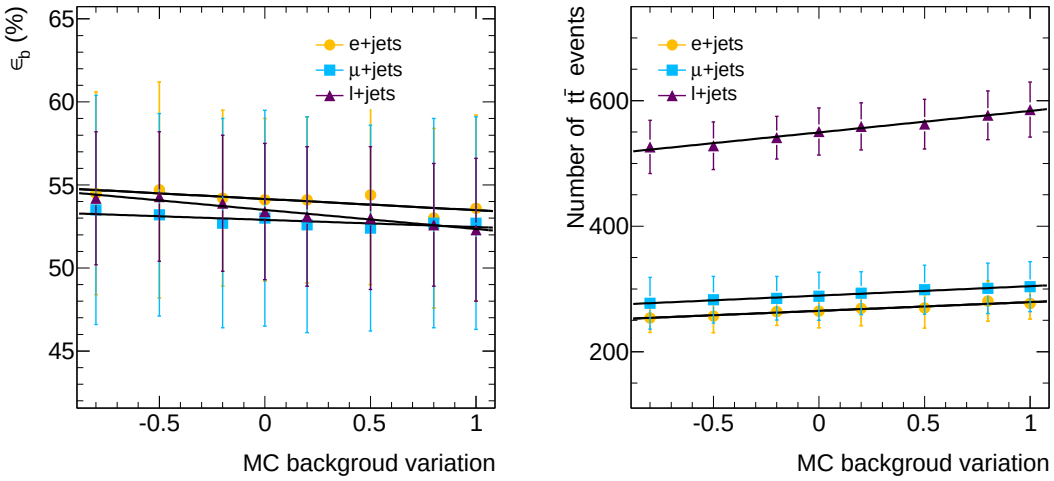


Figure 6.6: The combined of single top, diboson and ZJets background level systematic to ϵ_b and N_{sig} by fitting 1 to 3 b-tagged in e-chan, μ-chan and the combined channel of e-chan and μ-chan with the counting method

relative systematic		e-chan	μ-chan	combined
N0∼N3 (ϵ_b)	W+jets total (asymmetry)	26.7%	28.7%	19.3%
	W+HF total (asymmetry)	8.0%	13.0%	7.2%
	W+jets total (Berends)	31.8%	22.3%	18.9%
	W+HF total (Berends)	10.6%	9.4%	7.1%
	W+jets total (Combine)	27.7%	22.4%	16.5%
	W+HF total (Combine)	11.4%	9.4%	7.0%
N1∼N3 (ϵ_b)	W+HF total (asymmetry)	5.3%	8.9%	5.5%
	W+HF total (Berends)	8.2%	7.3%	5.5%
	W+HF total (Combine)	7.7%	7.3%	5.4%
N0∼N3 (N_{sig})	W+jets total (asymmetry)	36.1%	59.3%	31.4%
	W+HF total (asymmetry)	<0.1%	<0.1%	<0.1%
	W+jets total (Berends)	53.7%	39.3%	33.9%
	W+HF total (Berends)	<0.1%	<0.1%	<0.1%
	W+jets total (Combine)	54.5%	39.6%	30.6%
	W+HF total (Combine)	<0.1%	<0.1%	<0.1%
N1∼N3 (N_{sig})	W+HF total (asymmetry)	11.1%	17.8%	11.4%
	W+HF total (Berends)	16.8%	14.9%	11.2%
	W+HF total (Combine)	15.7%	14.9%	11.0%

Table 6.7: Summary of systematic uncertainties for the results of ϵ_b and N_{sig} due to the uncertainties on total W+jets events and W+HF with at least 1 btag.

6.4 Determination of the top pair cross section

The measured number of signal events determined in the previous section can be used to derive the cross section using the Formula 1.1. The $\sigma_{t\bar{t}}$ results fitting N1∼N3 are shown in Table 6.8 for the electron, muon and the combined channels. The results using different measures of W+jets background and fitting also N0∼N3 are listed in Table G.6 in Appendix G. Only the statistical uncertainties are presented related in the method measuring the N_{sig}. The systematic uncertainties only related to the cross section rather than the b-tagging efficiency are considered in the next section. The branching ratio for the not-fully-hadronic channels is 0.543, which is used for the default MC sample. The acceptances are given in 4.7 corresponding to the top mass at 172.5 GeV, which are 7.59±0.25% (e+jets) and 8.07±0.15% (μ+jets). By counting the numbers of events selected for both channels, the acceptance for combined l+jets channel is 15.66±0.29%

	e+jets	μ+jets	l+jets
$\sigma_{t\bar{t}}$	146.4±14.3	149.9±13.3	148.1±9.7

Table 6.8: Measured $\sigma_{t\bar{t}}$ results by subtracting W+jets measured from the charge asymmetry and Berends Scaling combined method fitting N1∼N3. Only the statistical uncertainties on fitting the data are included into the results.

The results of top pair cross section are calculated in this section, where the ones fitting N1∼N3 with W+jets background fitted with the combined method have lower systematic uncertainties. The result in the combined electron and muon channel has lower statistical uncertainty, and will be used as the final result in the thesis, which is 148.1 ± 9.7 with only the statistical uncertainty. The systematic uncertainties related to the measured $\sigma_{t\bar{t}}$ are considered in Section 6.5.

6.5 Systematic uncertainties to the top pair cross section $\sigma_{t\bar{t}}$

The main systematics, like W+jets background level, QCD multijet background level, jet energy scale and etc, are included into the systematics for N_{sig}. The systematics other than these are listed in Table 6.9, combined with the statistical uncertainties. The item "branching ratio" is the systematic due to the W boson decay and is negligible. The systematic of event selection efficiency, ϵ are from the uncertainties on the scale factors discussed in Section 4.6.

The uncertainty on the luminosity is quite reduced with a more careful calibration in October 2010, combined the earlier results in April and May 2010, whose results are published in the first Atlas luminosity paper [41]. The uncertainties on the luminosity are updated and in total it gives 3.4% instead of 11% previously, among which the main systematic uncertainties like "bunch charge product", "beam centering", "beam position jitter" and so on are better understood and quite reduced [60].

relative	e-chan	μ-chan	combined
luminosity	3.4%	3.4%	3.4%
branching ratio	1.0%	1.0%	1.0%
event select ϵ	3.3%	1.9%	2.1%
N_{sig} (fit N1∼N3)	27.8%	22.4%	22.3%
total systematic	28.0%	22.5%	22.4%
Statistical	9.6%	8.8%	6.5%
Lumi+Syst+Stat	29.8%	24.4%	23.6%

Table 6.9: Total uncertainties including luminosity, all the systematics and statistics to the $\sigma_{t\bar{t}}$ in e+jets, μ+jets and electron and muon combined channel. The uncertainties related to N_{sig} is from the tag counting method fitting N1∼N3, inside which the W+jets measurement is using the charge asymmetry and Berends Scaling combined method.

The table shows that the systematic in the e+jets and μ+jets combined channel with fitting N1∼N3 is lower, which gives a result of cross section:

$$\sigma_{t\bar{t}} = 148.1 \pm 9.7(\text{stat.}) \pm 33.2(\text{syst.}) \pm 5.0(\text{lumi.})\ pb = 148.1 \pm 34.9\ pb$$

The systematics in this method are contributed a lot by the number of signal events after fitting. Fitting with the 0 b-tagged bin is obviously more affected by the dominating backgrounds like W+jets. Even fitting without the 0 b-tagged bin, systematics arising from the N_{sig} are much larger than those from luminosity or statistical uncertainties. These can be improved when the background cross section is better understood, the heavy flavor fraction is better known or the jet energy is better measured, etc. It also includes the systematics concerning the MC modeling, which are also large. These can only be constrained by better understanding of the theory and well modeled in MC simulation.

A 2D contour plot with the fitting results of ϵ_b and $\sigma_{t\bar{t}}$ is shown in Figure 6.7. Only statistical uncertainties on the results are included in this figure. The fitted result is at 1.5σ of top pair cross section statistical uncertainty. This is acceptable, since the measurement is limited by the total systematic uncertainty, which is much larger than the statistical uncertainty. Both the fitted ϵ_b and $\sigma_{t\bar{t}}$ are compatible with those from other methods in ATLAS with the statistical and systematic uncertainty presented.

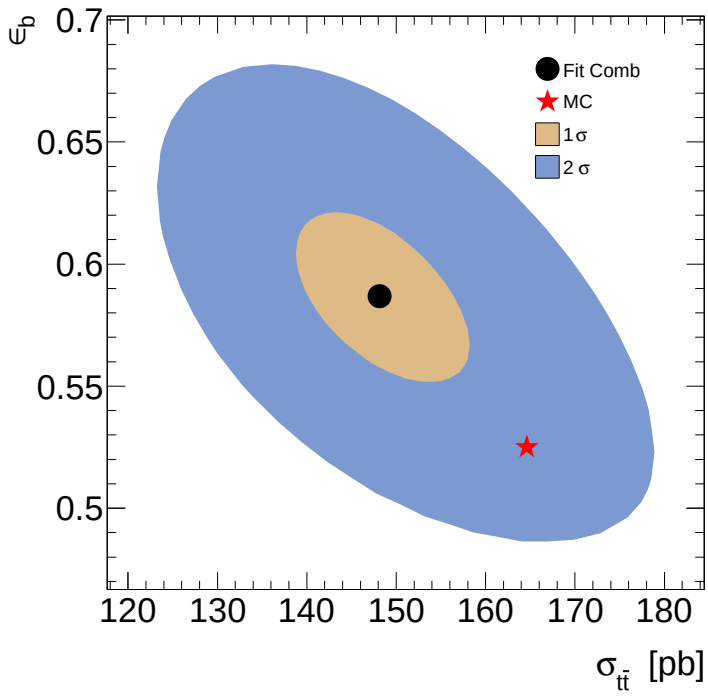


Figure 6.7: The 2D contour plot with the fitting results of ϵ_b and $\sigma_{t\bar{t}}$ corresponding to the result shown in this conclusion. Only statistical uncertainties on the results are included in the plot.

6.6 Prospectives

As some reprocessing procedure has been going on for data taken in 2010 since the beginning of 2011, some improvements have been made on the objects reconstruction. The corresponding measurements also gain a better measurement, for example, the W+jets and QCD multijet results. Besides, LHC has decided to keep running at 7 TeV in 2011. A lower statistical and systematic uncertainty can be expected. Some of the possible improvements related to the $\sigma_{t\bar{t}}$ measurement in this thesis are listed:

- Electron identification quality requirements are tightened by using "tight" for "isem" flag instead of "medium". This can reduce the fake electron background from multijet events, with an additional efficiency loss of about 10% [49]. Therefore, the QCD multijet background in e+jets channel can be suppressed to the level as in μ+jets channel.

- The reconstructed jets used in the analysis are *anti-k_T*-4 jets, where the JES uncertainty can be accessed with a tool named "MultijetJESUncertaintyProvider". Some details can be found in the note [61] for the usage of jets in top group. The tool incorporates all basic effects accounted for the aspects in JES calibration. The uncertainty on the JES is around 5% depending on the jet's η and p_{T} instead of a plain 7%, which is used in this thesis. The tool will further include systematics to account for flavor composition and close-by jets effects. An update on the JES of the b-tagged jets is also expected to be implemented into the tool.

- In the method of tag counting to measure the number of $t\bar{t}$ events, one can slightly change the Formula 6.1 into Formula 6.3 to directly measure the top pair cross section.

$$N^{exp}(\epsilon_b, \sigma_{t\bar{t}}) = \sum_{i,j,k} (\sigma_{t\bar{t}} \cdot BR \cdot A \cdot L \cdot F_{ijk}^{t\bar{t}} + N_{bkg} \cdot F_{ijk}^{bkg}) \times \sum_{i'+j'+k'=n} A_i^{i'} \cdot {\epsilon_b}^{i'} \cdot (1-\epsilon_b)^{i-i'} \cdot A_j^{j'} \cdot {\epsilon_c}^{j'} \cdot (1-\epsilon_c)^{j-j'} \cdot A_k^{k'} \cdot {\epsilon_l}^{k'} \cdot (1-\epsilon_l)^{k-k'} \quad (6.3)$$

where BR is branching ratio of the channel used to fit, A is the acceptance. Besides, by including the F_{ijk} contribution from backgrounds using corresponding MC samples,

one can avoid the uncertainty due to the difference of ϵ_b between MC and measurement when the backgrounds are subtracted.

- The tag counting method use MC samples to calculate the F_{ijk} parameters, which highly depends on MC to count the flavor composition. The top default MC uses MC@NLO generator with hadronisation by Herwig. The positive weight emission generator (Powheg) also produces final states at NLO accuracy and is hadronised by Herwig (or Pythia). These two MC with different generators are compared to denote the systematic on MC generator.

Chapter 7

CONCLUSION

Thanks to the successful running of LHC and ATLAS in 2010, more than 40 pb^{-1} data are recorded and around 33.7 pb^{-1} of them are used in top group before the reprocessing in 2011. The analysis on the very first data, 2.9 pb^{-1}, gives a preliminary result on the $t\bar{t}$ cross section, which is included into the first top observation paper [45]. In this thesis, the analysis is applied to the lepton+jets channels, which gives a final result of $t\bar{t}$ cross section with top mass at 172.5 GeV:

$$\sigma_{t\bar{t}} = 148.1 \pm 34.9\ pb$$

This is an early measurement of $t\bar{t}$ cross section with a mature QCD multijet data driven "Matrix Method", several W+jets background measurements and a first application of the simultaneous measurement of ϵ_b and number of signal events, which leads to $\sigma_{t\bar{t}}$. All the methods included in this analysis are making use of the reconstruction of electrons, muons, jets and missing transverse energy and the b-tagging algorithms, which span the whole usage of the detector.

The measured cross section in e+jets channel, μ+jets channel or the combined channel are consistent with each other. The kinematic distributions shown in the control plots are consistent with the prediction in SM $t\bar{t}$ production. The $\sigma_{t\bar{t}}$ measured is in good agreement with the prediction of NNLO top quark cross section calculation, which is presented in Figure 7.1.

With a larger statistics in the very near future and with the help of a better understanding of the detectors, a more precise measurement can be expected. The W+jets background can be measured with an uncertainty of 10% with the charge asymmetry method or W/Z ratio method, which can reduce the W+jets background level uncertainty on N_{sig} from 30%

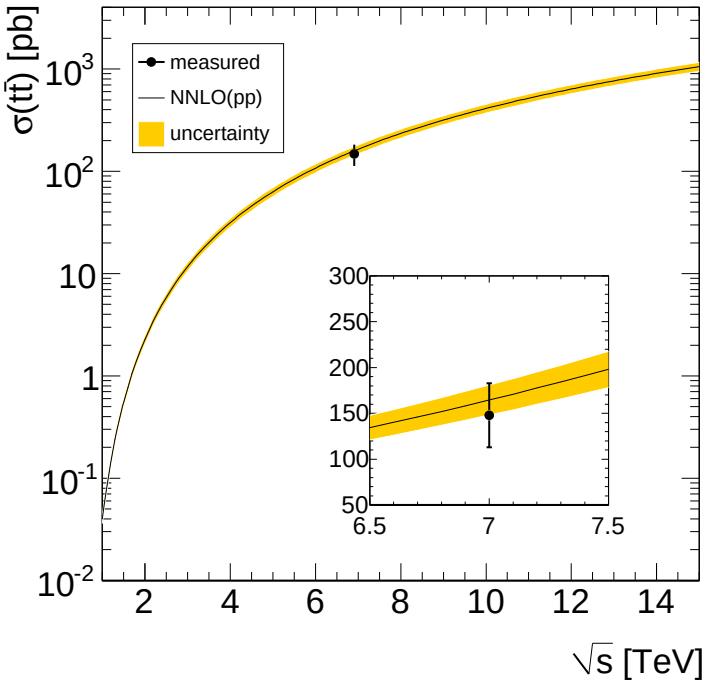


Figure 7.1: The cross section of $t\bar{t}$ measured in this thesis. The theoretical predictions for pp collision are obtained with HATHOR tool with the CTEQ6.6 PDFs [3] at a top mass of 172.5 GeV. The scale and PDF systematics are included, which give about 10% uncertainty in total.

to 10% by fitting N0∼N3 in the combined channel. The QCD multijet background can be better measured when the ϵ_{real} and ϵ_{fake} of the lepton in the "Matrix Method" is better known with large statistics. Correspondingly, the QCD multijet background level uncertainty can be reduced from 16% to 8%. The total uncertainty on the result fitting N0∼N3 can be reduced from 39% to 25% because of the improvement of the measurement of the two dominant backgrounds.

Concerning the default fitting used in the thesis with N1∼N3, the more important is the background level of the W+jets and the QCD multijet with at least 1 b-tagged jet. This can be improved consequently from the improvement of the measurement of the two backgrounds without the requirement of b-tagging. And the better understanding of b-tagging

efficiency, c-jet and light jet mistag rates can help to reduce the background level. With the charge asymmetry method and the W/Z ratio method, one can expect up to 15% on the W+jets background level with b-tagged jet. With the "Matrix Method" with b-tagging on QCD multijet background level, 15% is expected. The jet energy scale uncertainty can be reduced a factor of two in the coming data. With all those improvements from the dominant systematic uncertainties, the total uncertainty on N_{sig} can be reduced to about 12%, where the improvement on the MC modeling has not been counted. All those measurements with the method are expected to be challenge the SM prediction and also the models beyond SM.

Appendix A

MC SAMPLES FOR BACKGROUNDS

M_{Top} (GeV)	Run	Decay	σ (pb)	k-factor
160	6203	No Full Hadronic	117.4	1.12
	6223	Fully Hadronic	93.58	1.18
165	6208	No Full Hadronic	99.72	1.12
167.5	6205	No Full Hadronic	92.04	1.13
170	6201	No Full Hadronic	86.29	1.11
	6221	Fully Hadronic	69.97	1.16
172.5	5200	No Full Hadronic	80.201	1.11
	5204	Fully Hadronic	64.05	1.15
175	6206	No Full Hadronic	73.94	1.12
177.5	6207	No Full Hadronic	68.22	1.13
180	6202	No Full Hadronic	63.39	1.13
	6222	Fully Hadronic	52.50	1.15
190	6204	No Full Hadronic	49.03	1.10
	6224	Fully Hadronic	38.93	1.16

Table A.1: MC of $t\bar{t}$ samples at different top masses.

Run	Channel	σ (pb)	k-factor
8340	tchan $\rightarrow e$	7.152	1
8341	tchan $\rightarrow \mu$	7.176	1
8342	tchan $\rightarrow \tau$	7.128	1
8343	schan $\rightarrow e$	0.4685	1
8344	schan $\rightarrow \mu$	0.4684	1
8345	schan $\rightarrow \tau$	0.4700	1
8346	Wt $\rightarrow$ inclusive	14.581	1

Table A.2: MC of single top samples

Run	Channel	σ (pb)	k-factor
7680	W($\rightarrow e\nu$) + Np0	6913.3	1.22
7681	W($\rightarrow e\nu$) + Np1	1293.0	1.22
7682	W($\rightarrow e\nu$) + Np2	377.1	1.22
7683	W($\rightarrow e\nu$) + Np3	100.9	1.22
7684	W($\rightarrow e\nu$) + Np4	25.3	1.22
7685	W($\rightarrow e\nu$) + Np5	6.9	1.22
7690	W($\rightarrow \mu\nu$) + Np0	6935.4	1.22
7691	W($\rightarrow \mu\nu$) + Np1	1281.2	1.22
7692	W($\rightarrow \mu\nu$) + Np2	375.3	1.22
7693	W($\rightarrow \mu\nu$) + Np3	101.1	1.22
7694	W($\rightarrow \mu\nu$) + Np4	25.7	1.22
7695	W($\rightarrow \mu\nu$) + Np5	7.0	1.22
7700	W($\rightarrow \tau\nu$) + Np0	6835.8	1.22
7701	W($\rightarrow \tau\nu$) + Np1	1276.8	1.22
7702	W($\rightarrow \tau\nu$) + Np2	376.6	1.22
7703	W($\rightarrow \tau\nu$) + Np3	100.8	1.22
7704	W($\rightarrow \tau\nu$) + Np4	25.7	1.22
7705	W($\rightarrow \tau\nu$) + Np5	7.0	1.22

Table A.3: MC of W+Jets samples

Run	Channel	σ (pb)	k-factor
7650	Z($\rightarrow ee$) + Np0	661.9	1.22
7651	Z($\rightarrow ee$) + Np1	133.3	1.22
7652	Z($\rightarrow ee$) + Np2	40.3	1.22
7653	Z($\rightarrow ee$) + Np3	11.2	1.22
7654	Z($\rightarrow ee$) + Np4	2.7	1.22
7655	Z($\rightarrow ee$) + Np5	0.8	1.22
7660	Z($\rightarrow \mu\mu$) + Np0	657.7	1.22
7661	Z($\rightarrow \mu\mu$) + Np1	132.8	1.22
7662	Z($\rightarrow \mu\mu$) + Np2	39.6	1.22
7663	Z($\rightarrow \mu\mu$) + Np3	11.1	1.22
7664	Z($\rightarrow \mu\mu$) + Np4	2.8	1.22
7665	Z($\rightarrow \mu\mu$) + Np5	0.8	1.22
7670	Z($\rightarrow \tau\tau$) + Np0	657.4	1.22
7671	Z($\rightarrow \tau\tau$) + Np1	133.0	1.22
7672	Z($\rightarrow \tau\tau$) + Np2	40.4	1.22
7673	Z($\rightarrow \tau\tau$) + Np3	11.0	1.22
7674	Z($\rightarrow \tau\tau$) + Np4	2.9	1.22
7675	Z($\rightarrow \tau\tau$) + Np5	0.7	1.22

Table A.4: MC of Z+Jets samples

Run	Channel	σ (pb)	k-factor
6280	Wbb + Np0	3.2	1.22
6281	Wbb + Np1	2.6	1.22
6282	Wbb + Np2	1.4	1.22
6283	Wbb + Np3	0.6	1.22

Table A.5: MC of W+bb+Jets samples (overlap $\sim$10% with W+Np2 onwards)

Run	Channel	σ (pb)	k-factor
5985	WW (both→inclu)	11.75	1.52
5986	ZZ (both→inclu)	0.977	1.20
5987	WZ (both→inclu)	3.432	1.58

Table A.6: MC of W+cc+Jets samples

Appendix B

EVENT SELECTION FOR $t\bar{t}$ AT DIFFERENT TOP MASS

M_{Top} (GeV)	total	1 elec	trigger	nTrk≥5
160	4602.1±2.0	873.8±0.9	865.8±0.9	864.1±0.9
165	3909.0±1.7	739.1±0.7	732.1±0.7	730.4±0.7
167.5	3640.2±1.6	691.9±0.7	685.8±0.7	684.4±0.7
170	3352.4±1.6	635.0±0.7	629.3±0.7	627.8±0.7
175	2898.4±1.2	559.7±0.6	554.8±0.6	553.4±0.6
177.5	2698.1±1.2	522.1±0.5	517.7±0.5	516.7±0.5
180	2507.1±1.1	485.8±0.5	481.2±0.5	479.8±0.5
190	1887.7±0.8	371.3±0.4	367.9±0.4	367.0±0.4
Continue	$E_{\mathrm{T}}^{\mathrm{miss}}$ >20	$E_{\mathrm{T}}^{\mathrm{miss}} + M_T$ >60	No Bad J	N_{jet} ≥4
160	780.6±0.8	735.3±0.8	718.2±0.8	305.1±0.5
165	660.4±0.7	618.3±0.7	604.8±0.7	271.4±0.5
167.5	620.1±0.7	584.0±0.6	570.0±0.6	259.5±0.4
170	568.1±0.7	535.5±0.6	523.4±0.6	241.8±0.4
175	504.0±0.5	473.4±0.5	463.5±0.5	223.5±0.4
177.5	470.8±0.5	442.4±0.5	432.6±0.5	211.8±0.3
180	437.2±0.5	411.2±0.4	400.8±0.4	201.7±0.3
190	335.8±0.3	316.2±0.3	309.1±0.3	164.4±0.2

Table B.1: The number of events (33.7 pb^{-1}) after each selection cut in $t\bar{t}$ no full hadronic samples with different top mass in electron channel. The uncertainties shown in the table are statistical in MC. No scale factors have been applied to the number of events.

M_{Top} (GeV)	total	1 muon	trigger	nTrk≥5
160	4602.1±2.0	1073.3±1.0	898.9±0.9	898.2±0.9
165	3909.0±1.7	914.6±0.8	767.1±0.8	766.4±0.8
167.5	3640.2±1.6	860.6±0.8	722.5±0.7	722.2±0.7
170	3352.4±1.6	781.8±0.8	658.3±0.7	657.5±0.7
175	2898.4±1.2	685.1±0.6	572.8±0.6	572.2±0.6
177.5	2698.1±1.2	630.0±0.6	526.9±0.6	526.4±0.6
180	2507.1±1.1	588.7±0.5	493.4±0.5	492.8±0.5
190	1887.7±0.8	452.1±0.4	378.7±0.4	378.3±0.4
Continue	$E_{\mathrm{T}}^{\mathrm{miss}}$ >20	$E_{\mathrm{T}}^{\mathrm{miss}} + M_T$ >60	No Bad J	N_{jet} ≥4
160	818.1±0.9	774.6±0.8	754.7±0.8	318.2±0.5
165	698.1±0.7	660.2±0.7	643.6±0.7	290.5±0.5
167.5	659.6±0.7	623.3±0.7	606.8±0.7	280.7±0.4
170	600.7±0.7	569.7±0.7	554.0±0.6	254.4±0.4
175	522.4±0.5	493.9±0.5	481.2±0.5	232.3±0.4
177.5	480.8±0.5	456.0±0.5	445.5±0.5	217.8±0.4
180	450.9±0.5	428.2±0.5	417.2±0.4	210.3±0.3
190	349.0±0.4	330.5±0.3	321.8±0.3	170.9±0.2

Table B.2: The number of events (33.7 pb^{-1}) after each selection cut in $t\bar{t}$ no full hadronic samples with different top mass in muon channel. The uncertainties shown in the table are statistical in MC. No scale factors have been applied to the number of events.

Appendix C

SCALE FACTORS FOR MUON TRIGGER EFFICIENCY

Name	η range	ϕ range
EC	$\lvert\eta\rvert > 1.05$	any ϕ
B1P1	[-1.05, 0.6]	[-π, 5π/16] $\bigcup$ [11π/16, π]
B1P2	[-1.05, 0.6] $\bigcup$ [-0.6, -0.5] $\bigcup$ [-0.5, 0.2] $\bigcup$ [0.3, 0.6]	[5π/16, π/2]
B1P3	[-1.05, 0.6]	[π/2, 11π/16]
B2P1	[-0.6, 0.6]	[-π, 5π/16] $\bigcup$ [11π/16, π]
B2P2	[-0.5, -0.4] $\bigcup$ [0.2, 0.3]	[5π/16, π/2]
B2P3	[-0.6, 0.6]	[5π/16, 11π/16]
B3P1	[0.6, 1.05]	[-π, 5/16π] $\bigcup$ [11π/16, π]
B3P2	[0.6, 1.05]	[5π/16, π/2]
B3P3	[0.6, 1.05]	[π/2, 11π/16]

Table C.1: Definition of the $\eta - \phi$ bins of the muon for the expression of the trigger SF. The Barrel is divided into 9 regions according to a similar measured SF.

Region	Trigger SF
EC	0.987 ± 0.003 (stat) $^{+0.001}_{-0.001}$ (syst)
B1P1	1.026 ± 0.010 (stat) $^{+0.003}_{-0.002}$ (syst)
B1P2	0.919 ± 0.017 (stat) $^{+0.007}_{-0.000}$ (syst)
B1P3	0.952 ± 0.030 (stat) $^{+0.002}_{-0.003}$ (syst)
B2P1	1.009 ± 0.006 (stat) $^{+0.001}_{-0.002}$ (syst)
B2P2	0.657 ± 0.050 (stat) $^{+0.010}_{-0.000}$ (syst)
B2P3	0.906 ± 0.019 (stat) $^{+0.000}_{-0.004}$ (syst)
B3P1	1.005 ± 0.010 (stat) $^{+0.002}_{-0.003}$ (syst)
B3P2	0.843 ± 0.053 (stat) $^{+0.000}_{-0.013}$ (syst)
B3P3	1.046 ± 0.029 (stat) $^{+0.011}_{-0.009}$ (syst)

Table C.2: Muon trigger scale factors from $Z \to \mu\mu$ data and MC events, using the tag & probe method. The results correspond to data periods E4-I.

Appendix D

SELECTION EFFICIENCY WITH DIFFERENT TOP MASS

M_{Top} (GeV)	e+jets	μ+jets
160	6.88±0.07%	7.18±0.08%
165	7.21±0.08%	7.72±0.08%
167.5	7.40±0.08%	8.01±0.08%
170	7.49±0.08%	7.88±0.08%
172.5	7.82±0.04%	8.15±0.04%
175	8.01±0.08%	8.32±0.08%
177.5	8.15±0.08%	8.38±0.08%
180	8.35±0.08%	8.71±0.08%
190	9.04±0.08%	9.40±0.08%

Table D.1: Acceptance times efficiency $A \cdot \epsilon$ with the statistical uncertainty as a function of top mass in the electron channel and in the muon channel.

Appendix E

NUMBER OF EVENTS SEPARATED WITH CHARGE

e+jets	0 jet	1 jet	2 jets	3 jets	$\geq$4 jets
$t\bar{t}$ no full had (MC) (+)	0.8±0.2	9.3±2.8	39.4±11.8	75.3±22.6	113.8±34.1
$t\bar{t}$ no full had (MC) (−)	0.9±0.3	9.8±3.0	40.6±12.2	76.6±23.0	116.6±35.0
Z+Jets (MC) (+)	117.6±35.3	114.6±34.4	69.8±20.9	28.4±8.5	11.4±3.4
Z+Jets (MC) (−)	101.1±30.3	110.7±33.2	67.2±20.2	27.7±8.3	10.9±3.3
di-boson (MC) (+)	15.8±4.7	28.7±8.6	23.7±7.1	6.6±2.0	1.6±0.5
di-boson (MC) (−)	15.0±4.5	27.0±8.1	20.8±6.2	5.6±1.7	1.4±0.4
QCD (DD) (+)	6204.8±1861.4	4040.8±1212.2	1202.3±360.7	345.9±103.8	128.9±38.7
QCD (DD) (−)	6218.3±1865.5	3868.0±1160.4	1125.2±337.6	293.0±87.9	120.5±36.1

Table E.1: Events are separated by lepton charge for positive and negative as a function of number of jets in electron channel. The uncertainties to the number of events are assumed to be 30%. The results show a charge symmetry in these channels with the assumed uncertainties.

e+jets	0 jet	1 jet	2 jets	3 jets	$\geq$4 jets
$t\bar{t}$ no full had (MC) (+)	1.0±0.3	10.4±3.1	42.7±12.8	79.2±23.8	119.9±36.0
$t\bar{t}$ no full had (MC) (−)	1.0±0.3	10.2±3.1	43.1±12.9	79.7±23.9	120.4±36.1
Z+Jets (MC) (+)	2669.7±800.9	292.7±87.8	84.9±25.5	20.5±6.1	8.4±2.5
Z+Jets (MC) (−)	2287.1±686.1	265.1±79.5	78.8±23.6	20.4±6.1	7.8±2.4
di-boson (MC) (+)	19.5±5.9	33.5±10.0	27.3±8.2	7.6±2.3	1.6±0.5
di-boson (MC) (−)	18.7±5.6	31.5±9.4	23.9±7.2	6.6±2.0	1.5±0.5
QCD (DD) (+)	801.9±240.6	509.7±152.9	171.8±51.5	51.2±15.4	17.6±5.3
QCD (DD) (−)	712.7±213.8	494.3±148.3	160.3±48.1	56.3±16.9	22.7±6.8

Table E.2: Events are separated by lepton charge for positive and negative as a function of number of jets in muon channel. The uncertainties to the number of events are assumed to be 30%. The results show a charge symmetry in these channels with the assumed uncertainties.

Appendix F

W+JETS MEASUREMENT USING CHARGE ASYMMETRY METHOD WITH B-TAGGED JET

e+jets	1 jet	2 jets	3 jets	$\geq$4 jets
data(+)	247.0±15.7	183.0±13.5	144.0±12.0	121.0±11.0
single top (MC)(+)	11.9±3.6	18.6±5.6	8.7±2.6	3.8±1.1
D^+(subtract)	235.1±16.1	164.4±14.6	135.3±12.3	117.2±11.1
data(−)	216.0±14.7	156.0±12.5	105.0±10.2	116.0±10.8
single top (MC)(−)	6.6±2.0	12.1±3.6	6.2±1.9	4.0±1.2
D^-(subtract)	209.4±14.8	143.9±13.0	98.8±10.4	112.0±10.8
$D^+ - D^-$	25.6±21.9	20.5±19.6	36.5±16.1	5.3±15.5
Asymmetry	0.0397±0.0338	0.1314±0.0325	0.2171±0.0519	0.1707±0.0783
WJets (Meas)	646.0±779.3	155.8±153.9	167.9±84.3	30.9±91.8
WJets (MC)	166.3±1.0	93.5±0.5	31.0±0.3	13.2±0.2

Table F.1: The numbers of events with $\geq$1 tagged are presented as a function of number of jets. The description concerning the channels are common as in Table 5.2. The results shown are in electron channel with statistical uncertainty.

μ+jets	1 jet	2 jets	3 jets	$\geq$4 jets
data(+)	182.0±13.5	159.0±12.6	111.0±10.5	120.0±11.0
single top (MC)(+)	13.2±4.0	18.9±5.7	10.3±3.1	4.5±1.3
D^+(subtract)	168.8±14.1	140.1±13.8	100.7±11.0	115.5±11.0
data(−)	152.0±12.3	145.0±12.0	92.0±9.6	110.0±10.5
single top (MC)(−)	7.6±2.3	12.0±3.6	7.4±2.2	3.1±0.9
D^-(subtract)	144.4±12.5	133.0±12.6	84.6±9.8	106.9±10.5
$D^+ - D^-$	24.4±18.8	7.1±18.7	16.1±14.7	8.6±15.3
Asymmetry	0.1037±0.0286	0.1313±0.0294	0.0781±0.0471	0.2233±0.0663
WJets (Meas)	235.5±193.0	54.1±142.8	206.3±226.1	38.6±69.3
WJets (MC)	216.0±1.1	111.2±0.6	38.9±0.3	17.7±0.2

Table F.2: The numbers of events with $\geq$1 tagged are presented as a function of number of jets. The results shown are in muon channel with statistical uncertainty.

Appendix G

RESULTS AND SYSTEMATIC UNCERTAINTIES IN TAG COUNTING METHOD

Asymmetry	0 btag	1 btag	2 btag	3 btag
e+jets	120.6±105.8	16.8±14.5	1.4±1.2	0.1±0.1
μ+jets	296.7±104.1	41.5±21.0	3.2±1.7	0.4±0.3
Berends Scaling	0 btag	1 btag	2 btag	3 btag
e+jets	199.8±80.5	27.7±15.0	2.3±1.3	0.1±0.1
μ+jets	253.2±79.9	35.5±17.3	2.8±1.4	0.3±0.3

Table G.1: Number of W+jets events with b-tagged jet multiplicity. The pretag number of events is measured with the charge asymmetry the and the Berends Scaling method. The number of no b-tagged events is measured by pretag events subtracting tagged events. The number of tagged events times the fractions from Table 6.2, which gives the number of events for exactly 1, 2 and 3 b-tagged events.

N_{sig}	W+jet sub	e+jets	μ+jets	l+jets
Fit N0∼N3	Asymmetry	282.6±16.8	170.2±13.1	452.8±21.3
	Berends	191.3±13.9	220.4±14.9	411.7±20.3
	Combine	148.1±12.2	218.3±14.8	366.4±19.2
Fit N1∼N3	Asymmetry	236.8±22.9	210.4±18.7	445.1±29.1
	Berends	214.0±20.7	221.9±19.7	435.3±28.4
	Combine	203.4±19.8	221.5±19.6	424.6±27.8

Table G.2: Fitted N_{sig} results by subtracting W+jets measured from asymmetry method, Berends scaling and the combined method. Other backgrounds subtracted are identical in all the three cases, where QCD multijet is from "Matrix Method" and others from MC estimation. Statistical uncertainties in data fitting are presented.

ϵ_b	W+jet sub	e+jets	μ+jets	l+jets
Fit N0∼N3	Asymmetry	43.6±2.4%	77.9±3.0%	56.3±2.0%
	Berends	62.9±3.0%	60.6±2.8%	61.7±2.1%
	Combine	80.9±3.2%	61.1±2.8%	69.0±2.1%
Fit N1∼N3	Asymmetry	53.0±5.0%	61.7±4.8%	57.4±3.5%
	Berends	55.6±5.1%	60.1±4.7%	58.0±3.5%
	Combine	57.0±5.1%	60.2±4.7%	58.7±3.5%

Table G.3: Fitted ϵ_b results by subtracting W+jets measured from asymmetry method, Berends scaling or the combined method. Other backgrounds subtracted are identical in all the three cases, where QCD multijet is from "Matrix Method" and others from MC estimation.

relative syst.	e-chan	μ-chan	combined
ISR/FSR	3.1%	2.9%	3.0%
Parton shower	2.4%	1.3%	1.8%
MC generator	1.2%	0.7%	0.9%
Pile up	1.3%	<0.1%	0.6%
JES (±7%)	0.5%	0.5%	0.5%
ϵ_c (±100%)	4.2%	4.2%	4.2%
ϵ_l (±100%)	1.3%	1.3%	1.3%
QCD multijet level (±30%)	17.9%	3.6%	11.3%
QCD+HF total (±50%)	9.3%	1.9%	5.4%
QCD+HF ≥2 b (±100%)	0.9%	0.6%	0.7%
WJets (combine method)	27.7%	22.4%	16.5%
W+HF total (combine)	11.4%	9.4%	7.0%
W+HF ≥2 b (±100%)	1.0%	1.2%	1.1%
Other Bkg (±30%)	3.2%	2.4%	2.8%
Other Bkg (ϵ_b±10%)	1.3%	1.2%	0.9%
Total Systematic	36.8%	25.4%	22.9%
Statistical (33.7 pb^{-1})	5.0%	4.8%	3.5%

Table G.4: Summary of systematic and statistical uncertainties for ϵ_b by fitting 0 to 3 b-tagged bins. The total systematic uncertainty is quadratically combined with the systematic uncertainties in the list.

relative syst.	e-chan	μ-chan	combined
ISR/FSR	7.2%	6.5%	6.9%
Parton shower	2.0%	4.3%	3.2%
MC generator	8.9%	11.3%	10.1%
Pile up	6.2%	4.5%	5.4%
JES ($\pm$7%)	9.6%	9.6%	9.6%
ϵ_c ($\pm$100%)	<0.1%	<0.1%	<0.1%
ϵ_l ($\pm$100%)	<0.1%	<0.1%	<0.1%
QCD multijet level ($\pm$30%)	27.8%	4.5%	15.8%
QCD+HF total ($\pm$50%)	<0.1%	<0.1%	<0.1%
QCD+HF $\geq$2 b ($\pm$100%)	<0.1%	<0.1%	<0.1%
WJets total (Combine)	54.5%	39.6%	30.6%
W+HF total (Combine)	<0.1%	<0.1%	<0.1%
W+HF $\geq$2 b ($\pm$100%)	<0.1%	<0.1%	<0.1%
Other Bkg ($\pm$30%)	4.5%	3.5%	3.9%
Other Bkg ($\epsilon_b\pm$10%)	0.9%	0.7%	0.8%
Total Systematic	63.5%	43.6%	38.5%
Statistical (33.7 pb^{-1})	6.7%	6.1%	4.5%

Table G.5: Summary of systematic and statistical uncertainties for N_{sig} by fitting 0 to 3 b-tagged bins. The total systematic uncertainty is quadratically combined with the systematic uncertainties in the list.

$\sigma_{t\bar{t}}$	W+Jets	e+jets	μ+jets	e/μ+j
Fit N0∼N3	Asymmetry	203.4±12.1	115.2±8.9	158.0±7.4
	Berends	137.7±10.0	149.2±10.1	143.6±7.1
	Combine	106.6±8.8	147.8±10.0	127.8±6.7
Fit N1∼N3	Asymmetry	170.4±16.5	142.4±12.7	155.3±10.1
	Berends	154.0±14.9	150.2±13.3	151.8±9.9
	Combine	146.4±14.3	149.9±13.3	148.1±9.7

Table G.6: Measured $\sigma_{t\bar{t}}$ results by subtracting W+Jets measured from asymmetry, Berends scaling or the combined method. Only the statistical uncertainties on fitting the data are included into the results.

BIBLIOGRAPHY

[1] A. Quadt. Top quark physics at hadron colliders. *Euro. Phys. C*, 48:835–1000, 2006.

[2] ATLAS Collaboration. Readiness of the atlas liquid argon calorimeter for lhc collisions. *accepted for publication in EPJC*, page 31, 2009. arXiv:0912.2642v4.

[3] M. Aliev et al. Hathor hadronic top and heavy quarks cross section calculator. 2010. arXiv:1007.1327 [hep-ph].

[4] A. Salam and J. C. Ward. On a gauge theory of elementary interactions. *Nuovo Cimento*, 19:166–170, 1961.

[5] C. N. Yang and R. L. Mills. Conservation of isotopic spin and isotopic gauge invariance. *Phys. Rev.*, 96:191–195, 1954.

[6] J. Schwinger. A theory of the fundamental interactions. *Ann. Phys.*, 2:407–434, 1957.

[7] T. D. Lee and C. N. Yang. Possible nonlocal effects in μ decay. *Phys. Rev.*, 108:1611–1614, 1957.

[8] R. P. Feynman and M. Gell-Mann. Theory of the fermi interaction. *Phys. Rev.*, 109:193–198, 1958.

[9] N. Cabibbo. Unitary symmetry and leptonic decays. *Phys. Rev. Lett.*, 10:531–533, 1963.

[10] M. Kobayashi and T. Maskawa. Cp-violation in the renormalizable theory of weak interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[11] C. Amsler *et al*. Review of particle physics: The ckm quark-mixing matrix. *Phys. Lett. B*, 667:21, 2008.

[12] F. J. Hasert *et al*.. Search for elastic muon-neutrino electron scattering. *Phys. Lett. B*, 46:121–124, 1973.

[13] F. Englert and R. Brout. Broken symmetry and the mass of gauge vector mesons. *Phys. Rev. Lett.*, 13:321–323, 1964.

[14] P. W. Higgs. Broken symmetries and the masses of gauge bosons. *Phys. Rev. Lett.*, 13:508–509, 1964.

[15] P. W. Higgs. Spontaneous symmetry breakdown without massless bosons. *Phys. Rev.*, 145:1156–1163, 1966.

[16] T. Aaltonen *et al* (CDF and DØCollaborations). Combination of tevatron searches for the standard model higgs boson in the w^+w^- decay mode. *Phys. Rev. Lett.*, 104:11, 2010. arXiv:1001.4162.

[17] J. J. Aubert *et al*. Experimental observation of a heavy particle j. *Phys. Rev. Lett.*, 33:1404–1406, 1974.

[18] J. E. Augustin *et al*. Discovery of a narrow resonance in e^+e^- annihilation. *Phys. Rev. Lett.*, 33:1406–1408, 1974.

[19] S. W. Herb *et al*. Observation of a dimuon resonance at 9.5 GeV in 400- GeV proton-nucleus collisions. *Phys. Rev. Lett.*, 39:252–255, 1977.

[20] F. Abe *et al*. Observation of top quark production in $\bar{p}p$ collisions with the collider detector at fermilab. *Phys. Rev. Lett.*, 74:2626–2631, 1995.

[21] S. Abachi *et al*. Search for high mass top quark production in $\bar{p}p$ collision at $\sqrt{s} = 1.8$ TeV. *Phys. Rev. Lett.*, 74:2422–2426, 1995.

[22] Chris Quigg. Lhc physics potential vs. energy. 2009. arXiv:0908.3660v2 [hep-ph].

[23] F. I. Olness S. Kretzer, H. L. Lai and W. K. Tung. Cteq6 parton distributions with heavy quark mass effects. *Phys. Rev. D*, 69:11, 2004.

[24] M. Pleier. Review of properties of the top quark from measurements at the tevatron. *Int. J. Mod. Phys. A*, 24:2899–3037, 2009. arXiv:0810.5226v2.

[25] J. F. Owens M. Glück and E. Reya. Gluon contribution to hadronic j/ψ production. *Phys. Rev. D*, 17:2324–2331, 1978.

[26] B. L. Combridge. Associated production of heavy flavour states in pp and $p\bar{p}$ interactions: Some qcd estimates. *Nucl. Phys. B*, 151:429–456, 1979.

[27] Z. G. Si W. Bernreuther, A. Brandenburg and P. Uwer. Top quark pair production and decay at hadron colliders. *Nucl. Phys. B*, 690:81–137, 2004. arXiv:hep-ph/0403035v1.

[28] S. Moch and P. Uwer. Theoretical status and prospects for top-quark pair production at hadron colliders. *Phys. Rev. D*, 78:20, 2008. arXiv:0804.1476.

[29] A. Ferroglia V. Ahrens and et al. Renormalization-group improved predictions for top-quark pair production at hadron colliders. page 60, 2010. arXiv:1003.5827.

[30] V. M. Abazov *et al*. DØ Collaboration. Observation of single top-quark production. *Phys. Rev. Lett.*, 103:7, 2009. arXiv:0903.0850v1.

[31] T. Aaltonen *et al*. CDF Collaboration. Observation of electroweak single top-quark production. *Phys. Rev. Lett.*, 103:8, 2009. arXiv:0903.0885v1.

[32] N. Kidonakis. Single top quark production cross section at hadron colliders. page 60, 2010. arXiv:1005.3330v1.

[33] ATLAS Collaboration. Atlas detector and physics performance: Technical design report. 1:475, 1995.

[34] The ATLAS Collaboration. Observation of inclusive electrons in the atlas experiment at $\sqrt{s} = 7$ tev. *ATLAS CONF NOTE*, page 22, 2010. ATLAS-CONF-2010-073.

[35] The ATLAS collaboration. Determination of the muon reconstruction efficiency in atlas at the z resonance in proton-proton collisons at $\sqrt{s} = 7\,\text{TeV}$. 2011. ATLAS-CONF-2011-008.

[36] G. P. Salam M. Cacciari and G. Soyez. The anti-kt jet clustering algorithm. *Joural of High Energy Physics*, 04:15, 2008. arXiv:0802.1189.

[37] ATLAS Collaboration. Jet energy scale and its systematic uncertainty for jets produced in proton-proton collisions at $\sqrt{s} = 7$ TeV and measured with the atlas detector. 1:28, 2010. ATLAS-CONF-2010-056.

[38] A. Schwartzman. Performance of jets, missing transverse energy and tau reconstruction with atlas in pp collisions at 7 TeV. July 2010. 35th International Conference on High Energy Physics.

[39] The ATLAS collaboration. Performance of the atlas secondary vertex b-tagging algorithm in $7tev$ collision data. 2010. ATLAS-CONF-2010-042.

[40] S. van der Meer. Calibration of the effective beam height in the isr. *CERN-ISR-PO-68-31*, 1968.

[41] The ATLAS Collaboration. Determination in pp collisions at $\sqrt{s}$=7 tev using the atlas detector at the lhc. *EPJ*, page 61, 2011. arXiv:1101.2185v1 [hep-ex].

[42] The ATLAS collaboration. Monte carlo samples used for top physics. 2010. ATL-PHYS-INT-2010-132.

[43] V. M. Abazov *et al*. DØ Collaboration. Measurement of the $t\bar{t}$ production cross section in $\bar{p}p$ collisions at $\sqrt{s} = 1.96\,\text{TeV}$ using kinematic characteristics of lepton+jets events. *Phys. Rev. D*, 76:22, 2007.

[44] The ATLAS Collaboration. Observation of top quark pair production in the semileptonic decay channel at $\sqrt{s} = 7$ TeV with the atlas detector. *ATLAS NOTE*, page 36, 2010. ATL-COM-PHYS-2010-855.

[45] Measurement of top quark-pair production cross section with atlas in pp collisions at $\sqrt{s} = 7$ tev. submitted to EPJC.

[46] B. Abia *et. al.* Mis-identified lepton backgrounds to top 1 quark pair production. 1, 2010. ATL-COM-PHYS-2011-144.

[47] A. Baroncelli *et. al.* ATLAS Collaboration. $pp \to \gamma/z \to \mu\mu$ and $pp \to w \to \mu\nu$ inclusive cross-section measurement in early data with the atlas detector. 1, 2010. ATL-COM-PHYS-2010-124.

[48] N. Benekoss *et. al.* Lepton trigger and identification for the first top quark observation. 1, 2010. ATL-COM-PHYS-2010-826.

[49] N. Benekoss *et. al.* Lepton trigger and identification for the winter 2011 top quark analyses. 1, 2011. ATL-COM-PHYS-2011-123.

[50] J. Barreiro Guimaraes da Costa et al. $w \to \mu\nu$ and $z \to \mu\mu$ cross-sections measurements in proton-proton collisions at $\sqrt{s} = 7$ TeV with the atlas detector. *ATLAS NOTE*, 2010. ATL-COM-PHYS-2010-685.

[51] R. S. Thorne A. D. Martin, W. J. Stirling and G. Watt. Parton distributions for the lhc. *Euro. Phys. C*, 63:189–285, 2009. arXiv:0901.0002 [hep-ph].

[52] R. S. Thorne A. D. Martin, W. J. Stirling and G. Watt. Uncertainties on $\alpha(s)$ in global pdf analyses and implications for predicted hadronic cross sections. *Euro. Phys. C*, 64:653–680, 2009. arXiv:0905.3531 [hep-ph].

[53] H. Kuijf F. A. Berends and W. T. Giele B. Tausk. On the production of a w and jets at hadron colliders. *Nucl. Phys. B*, 351:32–64, 1991.

[54] Z. Bern C. F. Berger and et al. Precise predictions for w + 4 jet production at the large hadron collider. page 5, 2010. arXiv:1009.2338v1 [hep-ph].

[55] J. Alwall et al. Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions. *Eur. Phys. J. C*, 53:435, 1985.

[56] Erin Abouzaid and Henry Frisch. Ratio of w+n jets to z0/γ^*+n jets versus n as a precision test of the standard model. *Phys. Rev. D*, 68:10, 2003.

[57] W. T. Giele F. A. Berends and et al. Multijet production in w, z events at $p\bar{p}$ colliders. *Phys. Lett. B*, 224:237–242, 1989.

[58] B. Acharya et al. Prospects for measuring the top quark pair production cross-section in the single lepton channel at atlas in 10 tev p-p collisions. *ATL-PHYS-INT-2008-019*, 2009.

[59] The ATLAS collaboration. Calibrating the b-tag efficiency and mistag rate of the sv0 b-tagging algorithm in 3 pb^{-1} of data with the atlas detector. 2010. ATLAS-CONF-2010-099.

[60] The ATLAS Collaboration. Updated luminosity determination in pp collisions at root(s)=7 tev using the atlas detector. 2011. ATLAS-COM-CONF-2011-034.

[61] S. Cabrera Urbang *et. al.* Jets, missing transverse energy and taus for top physics analyses in release 16 with the 2010 dataset. 1, 2011. ATL-COM-PHYS-2011-132.