Measurement of the top quark pair production cross section in the muon + jets channel with the CMS detector at the LHC

Sinéad Walsh

faculty of sciences
Ghent University

Promoter: Prof. Dr. Martin Grünewald
Department of Physics and Astronomy

2012

Contents

Chapter 1

Introduction

Over the past century, the Standard Model of particle physics has been successful in predicting the existence and interactions of fundamental particles. The start up of the Large Hadron Collider (LHC) in 2009 heralded a new era in precision physics measurements at previously unavailable collision energies. Precision studies of the top quark, the highest mass fundamental particle, serve to benefit immensely from the LHC collision energies.

More than a million top quarks have been produced by the LHC in 2011 from the collision of protons at a center of mass energy of $\sqrt{s} = 7$ TeV. The subsequent decay of the top quark results in a variety of particles which are detected by the multi-purpose Compact Muon Solenoid (CMS) detector. The physics processes resulting from the collision, and the response of the CMS detector, are simulated to allow for comparison of the observed physics processes with SM expectations. Algorithms designed to take maximum advantage of the detector topology are then used to reconstruct the particles in the event, allowing for the identification of events originating from the decay of a top quark.

This thesis details a method developed to measure the production cross section of top-anti-top pairs at a center of mass energy of $\sqrt{s} = 7$ TeV. The measurement is performed with data collected by the CMS detector through the identification of $t\bar{t}$ events decaying to the muon + jets final state. The cross section is determined with the equation

$$\sigma_{t\bar{t}} = \frac{N_{t\bar{t}}}{\int \mathcal{L} \cdot \varepsilon_{t\bar{t}}}. \tag{1.1}$$

The determination of each component of Equation 1.1 is described in this thesis. The integrated luminosity, $\int \mathcal{L}$, of the collected data is measured offline with a pixel cluster counting method. The efficiency to select $t\bar{t}$ events in any decay channel, $\varepsilon_{t\bar{t}}$, is a combination of the detector efficiency and the efficiency of the selection criteria applied. It is provided by simulation with corrections derived from data applied. The selection criteria are optimised to select muon + jets events. $\varepsilon_{t\bar{t}}$ is then determined by combining $\varepsilon_{t\bar{t},\,\mu}$ and $\varepsilon_{t\bar{t},\,\text{other}}$ according to their respective branching fractions. The overall number of $t\bar{t}$ events, $N_{t\bar{t}}$, is extracted from data with a template fitting method.

This method avoids making assumptions on the normalisation of the signal or major background physics processes by performing a multi-parameter maximum likelihood fit to data. The fit makes use of the template shapes for the processes, estimated using simulation and data, to determine the contribution of each process to the selected data sample.

The systematic uncertainty on the measured $t\bar{t}$ cross section is evaluated by repeating the cross section extraction with systematic variations applied to the template shapes and the $t\bar{t}$ selection efficiency. The predicted $t\bar{t}$ cross section, calculated at NLO with NNLL corrections, is found to be in agreement with the measured $t\bar{t}$ cross section, within the statistical and systematic uncertainties of the measurement.

The SM is outlined in Chapter 2. The current understanding of top quark production and decay is described in 3. The $t\bar{t}$ final states, and the final states of background physics processes which affect the cross section measurement, are described in Chapter 4.

Chapters 5 and 6 introduce the LHC and CMS and detail the estimation of the integrated luminosity of collected data. The CMS trigger system, which determines the data to be stored or discarded, is described in Chapter 7. The simulation of proton collisions and the detector response to the particles produced are detailed in Chapter 8. The algorithms which reconstruct particles from the recorded or simulated detector signal are described in Chapter 9.

The requirements which are applied to select muon + jets events are given in Chapter 10. Chapter 11 describes the measurement of muon identification and trigger efficiencies in data and simulation. The method to extract the number of $t\bar{t}$ events in the selected data sample is described in Chapter 12. The systematic uncertainties on the measured cross section are outlined and quantified in Chapter 13.

Chapter 14 describes a method which is used to determine the top quark mass based on the measured $t\bar{t}$ cross section. The measured cross section is compared with results published by the CMS and ATLAS collaborations in Chapter 15. Finally, the analysis and results are summarised in Chapter 16.

Part I

Theory

Chapter 2

The Standard Model

Generation	Name	Spin	Charge	Mass
	Quarks (q)			
I	up (u)	1/2	2/3	2.3 MeV
	down (d)	1/2	-1/3	4.8 MeV
II	charm (c)	1/2	2/3	1.28 GeV
	strange (s)	1/2	-1/3	95 MeV
III	top (t)	1/2	2/3	173.5 GeV
	bottom (b)	1/2	-1/3	4.7 GeV
	Leptons (ℓ)			
I	electron neutrino (ν_e)	1/2	0	< 2 eV
	electron (e)	1/2	-1	0.511 MeV
II	muon neutrino (ν_μ)	1/2	0	< 0.19 MeV
	muon (μ)	1/2	-1	105.7 MeV
III	tau neutrino (ν_τ)	1/2	0	< 18.2 MeV
	tau (τ)	1/2	-1	1.777 GeV
	Gauge Bosons			
	photon (γ)	1	0	$< 1 \times 10^{-18}$ eV
	gluon (g)	1	0	0
	Z boson (Z^0)	1	0	91.2 GeV
	W boson ($W^\pm$)	1	± 1	80.4 GeV

Table 2.1: Overview of the Standard Model of particle physics. Properties provided by the 2012 review of particle physics [1].

The Standard Model (SM) [2, 3, 4] of particle physics is a theory explaining the fundamental particles and how they interact. Developed in the late 20th century, the SM unifies the quantum theories of electromagnetic, weak and strong interactions into one model. However, it is not a complete theory of the fundamental interactions. Attempts to incorporate the gravitational force remain unsuccessful. The Higgs mechanism aims to explain the observed masses of SM particles.

The properties of each fundamental particle predicted by the SM, with the exception of the SM Higgs, are summarised in Table 2.1. For each particle there is an anti-particle

with the same properties but opposite quantum number. Anti-particles are denoted with a line over the symbol of the corresponding particle, such as $\bar{t}$ for an anti-top quark.

The experimental detection of fundamental particles predicted by the SM has lead to general acceptance of the validity of the theory. The Higgs boson is the final SM particle to be observed experimentally. Recently, searches for the Higgs boson by the CMS and ATLAS experiments at the LHC uncovered a scalar boson with a mass in the region of 125 GeV [5, 6].

In this thesis the Heaviside-Lorentz system of units is employed with $\hbar = c = 1$.

2.1 Fermions

Fermions are particles with half integer spin which obey the Pauli exclusion principle. The Pauli exclusion principle states that no two fermions can exist simultaneously in the same quantum state. Consequently, fermions are the constituents of matter.

There are two types of fundamental fermions: quarks and leptons. The SM has three generations of each fermion type with two flavours, labelled up and down, per generation resulting in six quarks and six leptons. Fermions from the first generation are less massive and more stable than fermions from the second and third generations. Therefore, the higher generation fermions decay into first generation fermions.

2.1.1 Quarks

Quarks have both electric and colour charge, allowing them to interact via the electroweak and strong nuclear force. Up-flavour quarks have $+2/3$ electric charge while down-flavour quarks have $-1/3$ electric charge. A quark can have one of three colour charges: red, green or blue.

Due to colour confinement, quarks do not exist in isolation. Therefore, it is not possible to directly observe an individual quark. Instead, quarks hadronise to form composite particles called hadrons in which the quarks are held together via the strong force. The top quark is an exception as it decays before it can hadronise, as such there are no bound states with top quarks. Baryons are hadrons which consist of three quarks or three anti-quarks. The most commonly known baryons are protons, with two up and one down quark, and neutrons, with one up and two down quarks. Mesons are hadrons which consist of one quark and one anti-quark.

2.1.2 Leptons

Leptons have electric charge but no colour charge, therefore they participate in electroweak interactions but not strong interactions. The electron is a stable lepton and is found in nature orbiting the nucleus in atoms. The muon and tau leptons are created in high energy particle collisions and decay rapidly into a stable electron and neutrinos.

The tau lepton, with the largest mass, is the least stable lepton. The electron or muon leptonic decay modes of the tau lepton have equal probability.

Each charged lepton has a corresponding neutrino. Neutrinos have a small mass and are difficult to detect as they only interact via the weak force. In high energy colliders the presence of a neutrino in an interaction is generally inferred from a momentum imbalance, or missing energy, in the detector.

2.2 Forces

The forces which determine particle interactions in the SM are mediated by spin 1 gauge bosons.

2.2.1 The Strong Force

Gluons are the mediators for the strong force, affecting particles with colour charge. Quarks interact through the eight coloured gluons. Hadrons and mesons are colour neutral. Quantum chromodynamics (QCD) is the theory describing how quarks and gluons interact. QCD describes the processes related to the strong interactions observed at hadron colliders such as jet and heavy quark production. In high energy hadron collisions, perturbative QCD is used to predict the production cross sections for particles produced via the strong interaction. Quarks and gluons are also referred to as partons.

The strong coupling constant, α_s, increases as energy decreases or distance increases. Consequently, the attraction between quarks due to the strong force increases as the distance between the quarks increases. A quark-anti-quark pair is created from vacuum once the energy required to create the $q\bar{q}$ pair is less than then energy required to maintain the large separation between the original particles. This process, referred to as hadronisation, continues until the energy of the original $q\bar{q}$ pair is sufficiently reduced. With the exception of the top quark, an individual quark created in an interaction will result in a shower of particles created via hadronisation. This shower of particles is referred to as a jet.

Quarks in hadrons are quasi-free due to the small α_s arising from the short distances involved. The QCD mass scale, $\Lambda_{\mathrm{QCD}} \approx 200$ MeV, is the energy scale at which α_s approaches infinity.

2.2.2 The Electroweak Theory

The electroweak (EWK) theory is a unified description of electromagnetism and the weak force. Photons are the mediators of electromagnetic interactions. Quantum electrodynamics (QED) [7] is the theory describing how photons and fermions interact via the absorption and emission of photons. The W^+, W^- and Z bosons are the mediators for the weak force .

There are two types of EWK interactions, charged current and neutral current. Neutral current weak interactions are mediated by the Z boson. Neutral current electromagnetic interactions are propagated by the photon. The term neutral is used since the flavour of the fermion is not changed by the interaction. The colour and charge of the fermion are also conserved in such interactions.

Charged current interactions are mediated by the W^+ and W^- bosons. Charged current interactions mediate flavour change for quarks and leptons. The initial and final state quarks or leptons differ by one unit of electric charge.

The coupling of the charged current interactions to the up and down type quarks are described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix [1]. The CKM matrix is

$$V_{CKM} = V_{ij} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}. \tag{2.1}$$

The probability of quark i decaying to quark j is proportional to $|V_{ij}|^2$. The elements of the CKM matrix have been independently measured. The magnitude of the individual elements are then determined with improved accuracy using a global fit [1]. The fit result for the magnitude of the elements is

$$V_{CKM} = \begin{pmatrix} 0.97428 \pm 0.00015 & 0.2253 \pm 0.0007 & 0.00347^{+0.00016}_{-0.00012} \\ 0.2252 \pm 0.0007 & 0.97345^{+0.00015}_{-0.00016} & 0.0410^{+0.0011}_{-0.0007} \\ 0.00862^{+0.00026}_{-0.00020} & 0.0403^{+0.0011}_{-0.0007} & 0.999152^{+0.000030}_{-0.000045} \end{pmatrix} \tag{2.2}$$

2.2.2.1 The Higgs Mechanism

The Higgs mechanism [8, 9, 10] refers to the spontaneous breaking of EWK symmetry and explains the observed masses of fundamental particles, both fermions and exchange bosons. The SM predicts 'symmetric' zero mass for all of the EWK force carriers. However, while the photon has zero mass, the W and Z bosons break the EWK symmetry of boson masses by having non-zero masses.

The Higgs mechanism postulates the existence of a self-interacting complex doublet of scalar fields, referred to as the Higgs doublet. EWK symmetry breaking occurs when the neutral component of the Higgs doublet acquires a vacuum expectation value. The W and Z bosons gain mass through the absorption of massless Goldstone bosons generated in the symmetry breaking. The remaining component of the complex doublet becomes the Higgs boson. Fermion masses arise from the coupling of fermions to the Higgs doublet through Yukawa interactions.

Chapter 3

The Top Quark

The top quark is a third generation fermion. It was initially postulated as the up type partner to the third generation bottom quark discovered in 1977 [11]. The first observation of the top quark was made in 1995 by the CDF and D0 collaborations at the Tevatron [12, 13]. The top quark mass has since been measured precisely as $173.5 \pm 0.6 \pm 0.8$ GeV [1], marking it as the heaviest of the fundamental particles.

3.1 Significance

The top quark is expected to couple strongly to forces which break EWK symmetry due to its significant mass, placing it in a unique position to study the mechanism which results in quark masses. The top quark mass is also used, along with the W boson mass, to constrain the mass of the Higgs boson within the SM theory.

The measurement of top quark properties, in addition to its mass, is a useful test of the SM. In hadron colliders top quarks are produced predominantly in pairs via the strong interaction, providing validation of the QCD theory. The production of single top quarks is an EWK process, as is the decay of the top quark. Measurement of the decay modes validate the EWK theory or may suggest physics beyond the SM.

The large top quark mass results in a short lifetime of $\tau_t = 1/\Gamma_t \approx 0.5 \cdot 10^{-24}$ s, where Γ_t is the full decay width $2.0^{+0.7}_{-0.6}$ GeV. Consequently, the top quark decays before it can hadronise. This allows for study of the bare quark properties, such as spin and charge.

The measurement of top quark properties is facilitated by an accurate measurement of the top quark production cross section. The top quark mass can also be derived from the production cross section. Top quark production is often a predominant background in new physics searches. In such cases an accurately measured cross section is important to constrain the top quark contribution.

Since the early 1990's the Tevatron has been producing top quarks at a centre of mass energy of $\sqrt{s} = 1.8$ TeV and 1.96 TeV. The Large Hadron Collider (LHC) began operating at a centre of mass energy of $\sqrt{s} = 7$ GeV in 2010, resulting in a top quark production rate an order of magnitude above that of the Tevatron, and heralding a new era in the study of top quark physics. The expected $t\bar{t}$ production rates at the LHC and Tevatron energies are shown in Figure 3.1.

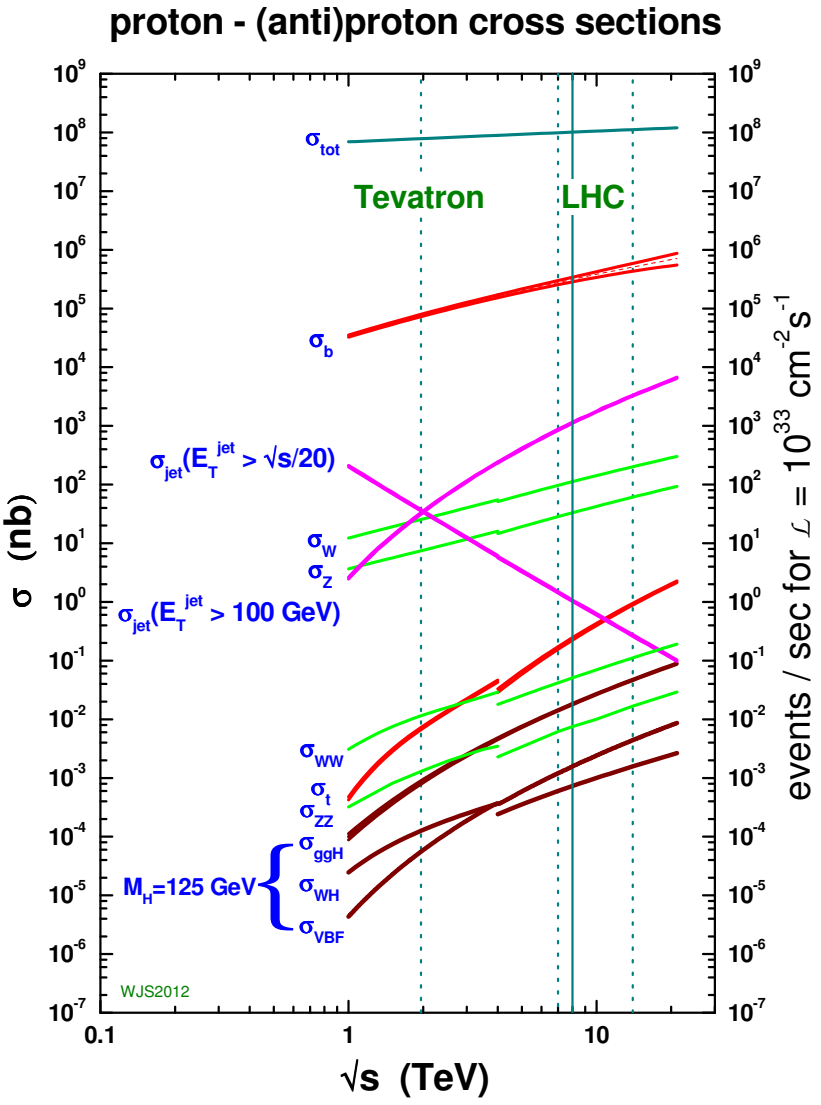


Figure 3.1: Calculated production cross section as a function of center of mass energy [14]. The vertical lines indicate the Tevatron center of mass energy, as well as the LHC energy in 2011 (7 TeV), 2012 (8 TeV) and the nominal LHC center of mass energy (14 TeV). The $t\bar{t}$ production cross section is labelled as σ_t.

3.2 Pair Production

Top quark pairs ($t\bar{t}$) are produced in the collision of two high energy leptons or hadrons, protons in the case of the LHC. Protons are made up of three valence quarks, two up quarks and a down quark, which make up most of the proton momentum. The proton also contains "sea" quarks and gluons which contribute to the total momentum of the proton. There are two mechanisms for $t\bar{t}$ production, $q\bar{q}$ annihilation and gluon fusion, shown in Figure 3.2. Gluon fusion is the dominant $t\bar{t}$ production mechanism at the LHC, responsible for approximately 85% of $t\bar{t}$ pairs produced at $\sqrt{s} = 7$ TeV.

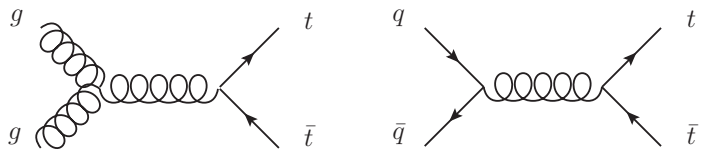


Figure 3.2: Feynman diagram for the production of $t\bar{t}$ pairs via gluon fusion, left, and $q\bar{q}$ annihilation, right [15].

The Leading Order (LO) approximation for the $t\bar{t}$ production cross section, Equation 3.1, is a convolution of the Parton Distribution Functions (PDF) [16] of the proton and the cross section for the partonic processes $q\bar{q} \to t\bar{t}$ and $gg \to t\bar{t}$. The total $t\bar{t}$ production cross section for hard scattering processes in pp collisions can be calculated with Equation 3.1 from [17].

$$\sigma(pp \to t\bar{t}X) = \sum_{ij} \int dx_i dx_j f_i^{p_i}(x_1, \mu^2) f_j^{p_j}(x_2, \mu^2) \hat{\sigma_{ij}}(ij \to t\bar{t}X, Q^2, \mu^2) \tag{3.1}$$

The interacting partons, quarks or gluons, are denoted by i and j. $f_i^p(x)$ are the PDFs where x is the fraction of proton momentum carried by parton i. $\hat{\sigma}(ij \to t\bar{t}X)$ is the partonic cross section. The factorisation and renormalisation scale is denoted by μ. The hard scattering scale, $Q^2 = x_1 x_2 s$, is the squared center of mass energy of the colliding partons in the protons.

The X in $ij \to t\bar{t}X$ denotes particles produced in addition to the $t\bar{t}$ pair, such as gluons from initial-state radiation (ISR) or final-state radiation (FSR). ISR occurs when one of the colliding quarks radiates a gluon which hadronises to form a jet. FSR occurs when a quark from the hard process emits a gluon, resulting in two jets which share the energy of the original quark. Both processes result in additional jets in the final state which are difficult to distinguish from jets from the hard scattering process. An inclusive cross section calculation takes such additional particles into account.

The factorisation scale μ_F [18] defines the separation between the short distance partonic cross section and the long distance aspects factored into the PDFs. As such, it determines the splitting of the perturbative and non-perturbative elements of the interaction. The partonic cross section also depends on the renormalisation scale μ_R [19] of the running coupling constant $\alpha_s(\mu^2)$. μ_R is introduced to regulate divergent terms in higher order calculations of the partonic cross section.

If all orders were included, the $t\bar{t}$ cross section would not depend on μ_F or μ_R. However, fixed order calculations have a scale dependence which becomes less significant with higher order calculations. The μ_F and μ_R scales are generally chosen to be equal and proportional to the mass of the particle generated in the interaction.

3.2.1 The Partonic Cross Section

The partonic cross section is calculated in perturbative QCD. Parton collisions occur at a hard energy scale characterised by large momenta and short distance interactions. The mass of the top quark, m_t, is much larger than the scale of QCD confinement, $\Lambda_{\mathrm{QCD}}(200\ \mathrm{MeV})$. The strong coupling constant becomes small at large energies allowing for the calculation of the partonic cross section with QCD perturbative theory.

3.2.2 Parton Distribution Functions

Soft interactions between the remaining partons are considered separately to the hard interaction. Partons within the proton interact at a low energy scale, $\Lambda_{\mathrm{QCD}} < 1$ GeV. The small momentum transfer indicates large $\alpha_s(Q^2)$ coupling, therefore, soft interactions cannot be described by perturbative QCD.

PDFs describe the probability density for a parton within the proton to carry a momentum fraction x at a squared energy scale Q^2. PDFs are determined experimentally with deep-inelastic scattering data in electron-proton collisions, such as in HERA [20] experiments. The PDFs are evolved to the relevant energy scale with the Dokshitzer-Gribov-Lipatov-Altarelli-Parisi Equation [21].

The squared center of mass energy at the parton level must be least $(2\mathrm{m}_t)^2$ to produce a top quark pair. The momentum fraction threshold for $t\bar{t}$ production is given by $< x >= \frac{2\mathrm{m}_t}{\sqrt{s}}$, which is 0.05 for a center of mass energy of $\sqrt{s} = 7$ TeV. The inverse relationship between the center of mass energy and the $t\bar{t}$ production threshold allows for gluon fusion to dominate at higher energies. For high momentum fraction thresholds $q\bar{q}$ annihilation dominates the $t\bar{t}$ production rate. However, as the center of mass energy increases, and consequently the momentum fraction threshold decreases, gluons in the proton are increasingly likely to have enough of the proton momentum to produce $t\bar{t}$ pairs. Since there are more gluons than quarks in the proton, gluon fusion will eventually dominate the $t\bar{t}$ production rate. The low threshold also means $t\bar{t}$ pairs tend to be produced away from threshold at $\sqrt{s} = 7$ TeV.

3.2.3 Cross Section Predictions

The LO formalisation for the calculation of the $t\bar{t}$ production cross section is given by Equation 3.1. To increase the accuracy and stability of the cross section prediction, higher order effects should be included in the calculation. Next-to-leading order (NLO) corrections include real and virtual calculations of soft gluon emissions and one loop virtual corrections.

Next-to-leading logarithm (NLL) resummation of higher order corrections due to soft gluon radiation improves the accuracy of the calculation. Resummation takes soft gluon emission contributions to the $t\bar{t}$ production cross section from all orders of α_s into account, leading to an increase in the calculated $t\bar{t}$ cross section. Resummation also improves the stability of the cross section predictions with respect to changes in μ_F and μ_R.

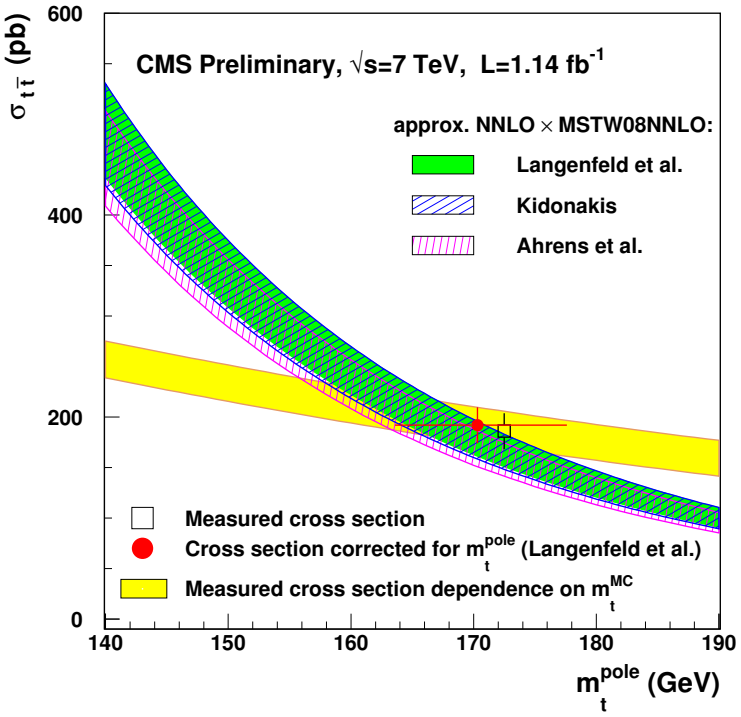


Figure 3.3: Predicted $t\bar{t}$ cross section as a function of the assumed top mass [22]. The $t\bar{t}$ cross section is measured in the dilepton channel.

Next-to-next-to-leading logarithm (NNLL) resummation requires the calculation of soft anomalous dimensions at two loops. Approximate next-to-next-to leading order

(NNLO) calculations based on NNLL resummation give the current best estimate for the $t\bar{t}$ cross section. At $\sqrt{s} = 7$ TeV, assuming $m_t = 173$ GeV, the approximate NNLO $t\bar{t}$ pair production cross section is calculated as 165 ± 10 pb [23]. The dependence of the calculated $t\bar{t}$ cross section on the assumed top mass is shown in Figure 3.3.

3.3 Decay

The short lifetime of the top quark relative to the hadronisation timescale results in the decay of the top quark before hadronisation. As m_t is larger than the W boson mass (80.385 GeV), $t \to Wq$ dominates the decay width, where q is one of the down type quarks d, s, or b. The probability for t to decay to a quark q is proportional to the CKM matrix element $|V_{tq}|^2$, given in Equations 2.1 and 2.2. Since the $t \to Wb$ decay dominates, with almost 100% probability, the contribution from $t \to Wd$ and $t \to Ws$ are negligible.

W^+ decay modes	Branching Fraction (%)
$e^+\nu$	10.75 ± 0.13
$\mu^+\nu$	10.57 ± 0.15
$\tau^+\nu$	11.25 ± 0.20
$\ell^+\nu$	32.57 ± 0.28
hadrons	67.60 ± 0.27

Table 3.1: Decay modes and branching ratios for the W^+ boson [1]. The W^- modes are charge conjugates to the W^+ modes. The symbol ℓ indicates the types of lepton e, μ and τ.

With $t \to Wb$ dominating the top quark decay, the $t\bar{t}$ decay channels are driven by the decay modes of the W boson. The W decay modes and their branching fractions are given in Table 3.1. The final states for the $t\bar{t} \to W^+bW^-\bar{b}$ process are described in Section 4.1.

Chapter 4

Physics Processes

Figure 3.1 shows that the top quark does not dominate the total production rate for particles in proton-proton interactions. Selection requirements are applied to the reconstructed final state particles to isolate event signatures similar to the muon + jets $t\bar{t}$ final state. This increases the relative fraction of muon + jets events in the data sample, however, background physics processes with event topologies similar to the muon + jets final state are still present. This chapter presents the physics processes which are expected to have a non-negligible contribution to the data sample after the application of selection requirements.

The expected cross sections for each of the processes are shown in Table 4.1.

Process (approx. NNLO)	Cross Section (pb)
$t\bar{t}$	165 ± 10
W + jets	31314 ± 1558
Z + jets	3048 ± 132
t-channel single top	$64.6^{+2.7}_{-2.0}$
tW single top	15.7 ± 1.1
s-channel single top	$4.6^{+0.14}_{-0.12}$
muon enriched multijet (LO)	85000

Table 4.1: Expected cross sections for $t\bar{t}$ and dominant background processes.

The $t\bar{t}$ cross section is calculated at approximate NNLO with NNLL resummations [23]. The W + jets cross section is calculated at approximate NNLO with FEWZ [24] for inclusive $W \to \ell\nu$ production, where $\ell = e, \mu$ or τ. The Z + jets cross section is calculated at approximate NNLO with FEWZ for inclusive $Z \to \ell\ell$ production for dilepton masses greater than 50 GeV. The t-channel, tW and s-channel cross sections are calculated at approximate NNLO with NNLL resummations [25, 26, 27]. An accurate estimate for the QCD multijet production cross section is not available. The LO PYTHIA [28] estimate for the production cross section of multijet events is for events where the p_{T} of outgoing muons on matrix element level is greater than 20 GeV. The

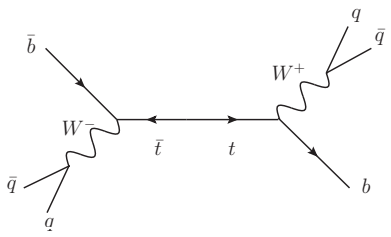
multijet cross section estimate includes muons from the decay of b and c quarks as well as muons from decays in flight of pions, kaons and K-longs.

4.1 The Signal Process

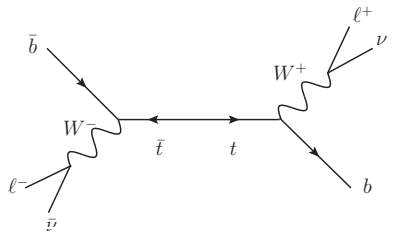
The final states for the $t\bar{t} \to W^+bW^-\bar{b}$ process can be divided into three classes, given in Table 4.2. The Feynman diagrams for the $t\bar{t}$ fully hadronic, dilepton and lepton + jets final states are shown in Figure 4.1.

Channel Name	Final State Particles	Branching Fraction (%)
fully hadronic	$q\bar{q}'\,b\,q''\bar{q}'''\bar{b}$	45.7
lepton+jets	$q\bar{q}'\,b\,\ell^-\bar{\nu}_\ell\bar{b} + \ell^+\nu_\ell b\,q''\bar{q}'''\bar{b}$	43.8
dilepton	$\bar{\ell}\nu_\ell b\,\ell'\,\bar{\nu}_{\ell'}\bar{b}$	10.5

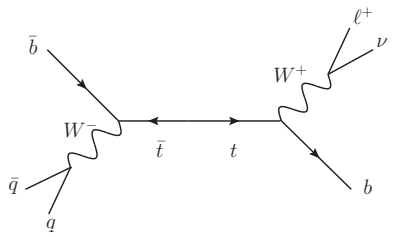
Table 4.2: Final state $t\bar{t}$ decay modes [1]. The symbol ℓ indicates each type of lepton e, μ and τ.



(a) fully hadronic



(b) dilepton



(c) lepton + jets

Figure 4.1: Feynman diagrams for $t\bar{t}$ fully hadronic, dilepton and lepton + jets final states [15].

The quarks in the final state hadronise to form jets, as described in Section 2.2.1.

Initial and final state quarks can radiate hard gluons which may be detected as additional jets.

The fully hadronic final state is the most prevalent $t\bar{t}$ final state, however, the presence of only jets in the event signature makes it difficult to distinguish these events from QCD multijet events. The dilepton final state has the cleanest event signature, with two well isolated high p_{T} leptons and a large amount of missing energy due to the two undetected neutrinos. However, the dilepton final state also has the lowest branching fraction.

The lepton+jets final state provides a compromise between the advantages and disadvantages of the other decay channels. The presence of one well isolated high p_{T} lepton and four jets, including two b quark jets, is an event signature which can be disentangled from background events with relatively simple selection requirements. The large branching fraction means that these events are produced with a large rate.

The lepton + jets final state where the lepton is a muon is the signal event signature studied in this analysis, also referred to as the muon + jets channel. The remaining lepton + jets final states, and the fully hadronic final state, may be misidentified as signal if a muon is reconstructed in the event. Dilepton events containing muons may be misidentified as signal if additional jets appear in the event.

The event selection criteria applied to the data sample are designed to accept muon + jets events while rejecting all other events. However, $t\bar{t}$ events from all decay channels are combined in the extraction of the $t\bar{t}$ cross section from the selected data sample. This is due to the difficulty in distinguishing between event kinematic distributions in different decay channels, which is relevant for the method used to extract the cross section in Chapter 12. The decay channels are combined with the muon + jets branching fraction determined by the W branching fractions given in Table 3.1.

4.2 The Background Processes

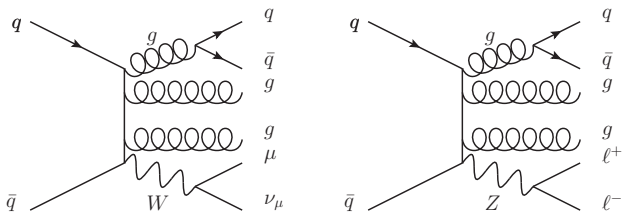
A background final state containing jets but no muon may be misidentified as a muon + jets event if a muon is spuriously reconstructed in the event. The decay of W or Z bosons is the dominant source of isolated muons in pp collisions. Less isolated muons occur from the decay of heavy flavour particles, such as b and c quarks or the τ lepton. Muons may also occur from the decay in flight of light flavour hadrons, such as pions and kaons, or the decay of particles produced in nuclear interactions in the detector material. A fake muon may be reconstructed if muon detector signals are produced by a non-muon particle, as is the case for hadronic punch-through where hadron shower remnants pass through the calorimeters and reach the muon system.

A background final state containing a muon but no or few jets may be misidentified as a muon + jets event if additional jets are reconstructed in the event. Additional jets arise due to ISR and FSR, described in Section 3.2.

4.2.1 W/Z + jets

The Feynman diagrams for a W + jets event with the W boson decaying to a lepton and a neutrino and a Z + jets event with the Z boson decaying to two leptons are shown in Figure 4.2. W + jets events where the final state lepton is a muon are the dominant source of background after the selection requirements are applied.

The inclusive cross section for $W/Z + \geq 0$ jets, shown in Table 4.1, is many times larger than the $t\bar{t}$ cross section. However, as a background to $t\bar{t}$ events only the production rate for $W/Z + \geq 4$ jets events is relevant.



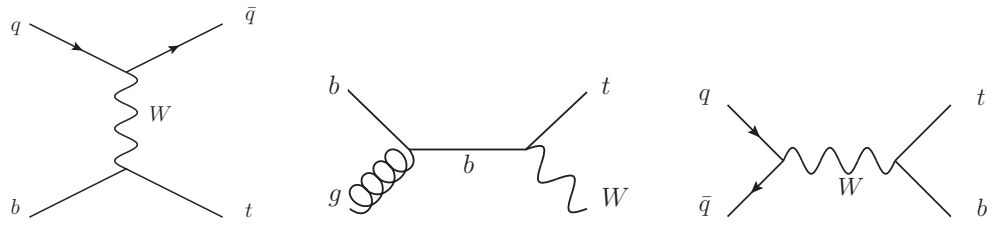
(a) W + jets (b) Z + jets

Figure 4.2: Feynman diagrams for the production of W + 4 jets and Z + 4 jets events [15].

4.2.2 Single Top

Top quarks are produced individually via the EWK interaction. There are three distinct processes for single top production: the t-channel exchange of a W boson, the associated production of a W boson and a top quark (tW), and the s-channel production and decay of a virtual W boson. The LO Feynman diagram for each of the single top processes is shown in Figure 4.3.

The tW process has the largest potential to be misidentified as a $t\bar{t}$ event as only one additional jet in the final state would result in a $t\bar{t}$-like event signature.



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

Figure 4.3: Feynman diagrams for the production of single top quarks via the t-channel, tW and s-channel production mechanisms [15].

4.2.3 QCD multijet

QCD multijet background is the most difficult background to predict. The overwhelming production rate for multijet events makes it a non-negligible background to many physics processes. QCD multijet events consist mainly of jets, however fake muons or muons from jets may result in a signal-like event signature.

No assumption is made on the production cross section for multijet events in this analysis. The kinematic distributions for multijet events are estimated from data by selecting a sample of events which is expected to be dominated by multijet events. Section 12.2 describes the method employed to extract multijet kinematic distributions from data.

4.2.4 Underlying Event

The composite nature of hadrons means more than one parton from the colliding proton can undergo scattering, referred to as multiple parton interactions. Particles from the breakup of protons involved in the hard process or initial-state radiation are referred to as beam remnants. In both cases particles will be detected in the detector which are not from the hard interaction. The processes which produce the extra particles are collectively referred to as the underlying event. The simulation of collision events, described in Chapter 8, includes the simulation of the underlying event.

4.2.5 Pile-up

Pile-up involves collisions between other protons in the same bunch crossing as the hard interaction. This results in extra particles in the detector, distorting the event signature of the hard interaction. Particles from pile-up interactions are removed by identifying the collision vertex from which each particle in the detector originated. Pile-up interactions are included in the simulation of collision events, described in Chapter 8.

Part II

Detection

Chapter 5

The Large Hadron Collider

The LHC [29] is an underground proton-proton accelerator and collider transversing the Franco-Swiss border near Geneva, Switzerland. It is the largest, highest energy particle accelerator currently available, with a circumference of 27.6 km and a nominal center of mass energy of $\sqrt{s} = 14$ TeV. The LHC began operating at a center of mass energy of $\sqrt{s} = 7$ TeV in 2010, allowing for examination of the SM at energy scales not previously attainable. Most notably, the LHC offers the potential for discovery of the Higgs boson and exploration of physics beyond the SM.

5.1 The Proton Beams

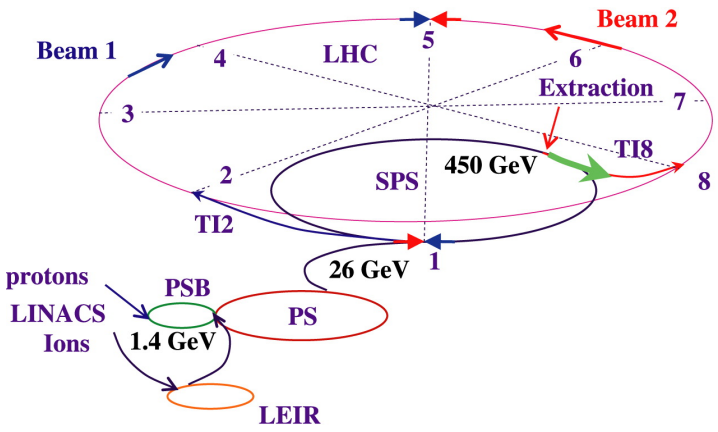


Figure 5.1: Schematic view of the LHC and its injection chain [29]. Points 1, 2, 5 and 8 on the LHC beam line correspond to regions in which the two proton beams interact.

In order to reach the desired proton beam energy the protons first travel through a series of pre-accelerators, shown in Figure 5.1. The protons are produced in the Linac2 duoplasmatron source [30]. The proton beams are first accelerated by the linear accelerator Linac2 before further acceleration by the PSB, PS and SPS synchrotron accelerators. The accelerators and their beam momenta at injection are listed in Table 5.1. Finally,

the two proton beams are injected into the LHC, in opposite directions, through transfer lines TI1 and TI8.

Machine	L (m)	ρ (m)	beam momentum (GeV)	bunches
LINAC	30	-	10^{-4}	4×2
PSB	157	8.3	0.05	4×2
PS	628.318	70.676	1.4	72
SPS	6911.56	741.257	26	4×72
LHC	26658.883	2803.98	450	2×2808

Table 5.1: Length (circumference), bending radius ρ and beam momentum at injection of main accelerators in LHC injection chain [29].

The proton beams are accelerated in the LHC up the desired beam energy. The beam energy was 3.5 TeV in 2011 and will increase to 7 TeV later this decade. The proton beam energy limits the particles which may be produced in collisions according to the particle mass. Synchrotron radiation is a source of energy loss when charged particles are accelerated radially and is proportional to $1/\mathrm{m}^4$, where m is the mass of the accelerated particle. The LHC beam consists of protons to reduce beam energy losses due to synchrotron radiation. The proton is significantly more massive than alternative candidates, such as the electron, therefore the energy loss is greatly diminished.

Dipole magnets control the bending of the proton beam in the LHC. Maximising the integrated dipole field increases the bending power and therefore the beam energy achievable. The LHC is built with superconducting Nb-Ti magnets, conducting electricity without resistance or loss of energy. Fields of up to B = 8.33 T, and consequently proton beam energies of up to 7 TeV, can be achieved by operating at a temperature of 1.9 K. At this temperature helium is superfluid with a large thermal conductivity.

The proton beams are accelerated in two separate beam pipes which cross at four interaction points, depicted as the interaction regions (IR) 1, 2, 5 and 8 in Figure 5.1. Separation and recombination dipole magnets left and right of the interaction regions adjust the separation of the beams at the interaction points in preparation for collisions.

The acceleration of two counter rotating proton beams necessitates the use of separate magnet and vacuum systems for each beam, since opposite magnet dipole fields are needed. To save space and reduce costs a two-in-one magnet system was designed, with twin bore magnets consisting of two sets of magnet coils and beam channels within the same mechanical structure and cryostat.

The four interaction points correspond to six detectors installed on the beam line. Two general purpose detectors, CMS and ATLAS, are installed at IR5 and IR1. LHCb, at IR8, is designed for the study of b-quark physics. ALICE, at IR2, is designed to study heavy ion (Pb-Pb) and proton-ion (p-Pb) collisions, for this reason the LHC also accel-

erates and collides Pb beams. There are two additional special purpose experiments. LHCf is installed close to ATLAS and is intended for the study of cosmic rays. TOTEM is located close to CMS and intends to study total cross section, elastic scattering and diffractive processes.

5.2 Instantaneous Luminosity

The number of events generated per second depends on the LHC machine luminosity. The LHC is designed for proton-proton collisions, instead of proton-anti-proton, to avoid luminosity limitations from the rate of anti-proton production. The instantaneous luminosity for a Gaussian beam distribution can be written as

$$\mathcal{L} = \frac{\mathrm{N}_b^2 n_b f_{\mathrm{rev}}}{A}. \tag{5.1}$$

N_b is the number of particles per bunch, n_b is the number of bunches colliding at the interaction point, f_{rev} is the revolution frequency and A is the effective beam overlap cross section at the interaction point. For beams colliding head on with horizontal and vertical r.m.s beam sizes σ_x and σ_y respectively

$$A = 4\pi\sigma_x\sigma_y. \tag{5.2}$$

Increasing the number of proton bunches results in a linear increase on the luminosity, while increasing the number of protons per bunch results in a quadratic luminosity increase. The luminosity also depends on the transverse beam size and beam overlap at interaction. Quadrapole magnets are installed around the interactions regions to focus the beams. The beam overlap is determined by the separation and recombination dipole magnets.

5.3 LHC in 2011

During its 2010 operation the LHC collected almost $50\,\mathrm{pb}^{-1}$ of data at a center of mass collision energy of $\sqrt{s} = 7$ TeV. In 2011 it was decided to continue with the same collision energy with an aim to collect $1\,\mathrm{fb}^{-1}$ of data, a goal which was surpassed early in the year. Figure 5.2 shows the evolution of luminosity delivered by the LHC to the CMS detector in 2011. Table 5.2 compares the design parameters of the LHC with the peak values reached in the 2011 data taking period.

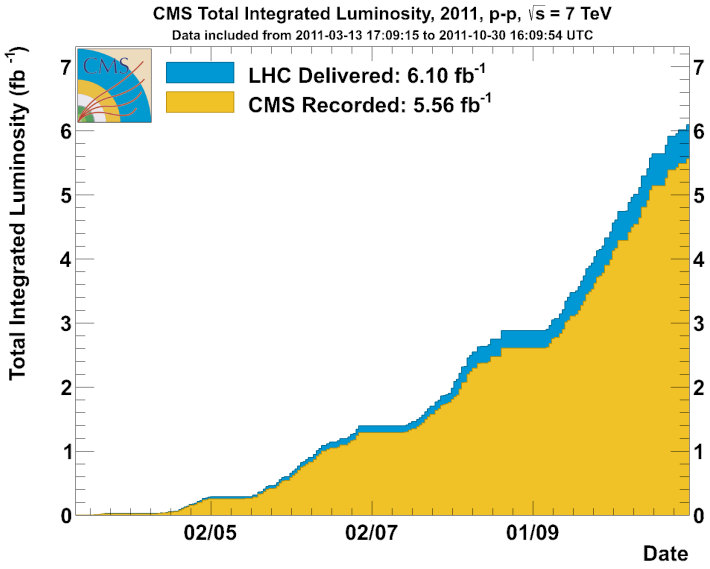


Figure 5.2: Luminosity delivered by the LHC and recorded by the CMS detector in 2011 [31].

	Design	2011
Beam energy, TeV	7	3.5
Peak Luminosity, $\mathrm{cm}^{-2}\mathrm{s}^{-1}$	10^{34}	3.65×10^{33}
Number of bunches per beam	2808	1380
Number of protons per bunch	1.15×10^{11}	1.4×10^{11}
Bunch spacing, ns	25	50

Table 5.2: LHC design parameters compared with the peak values reached during the 2011 period of operation [29, 32].

Chapter 6

The CMS Experiment

The Compact Muon Solenoid (CMS) [33, 34] detector is a general purpose particle detector installed 100 m underground on the LHC beam line at interaction region five. A schematic of the CMS detector is shown in Figure 6.1. It is 21.6 m in length, 14.6 m in diameter and weighs 12,500 tonnes. From the inside out, the interaction point is surrounded by the silicon tracking system, the electromagnetic and hadronic calorimeters and the muon system. A 3.8 T magnetic field is provided by a superconducting solenoid magnet.

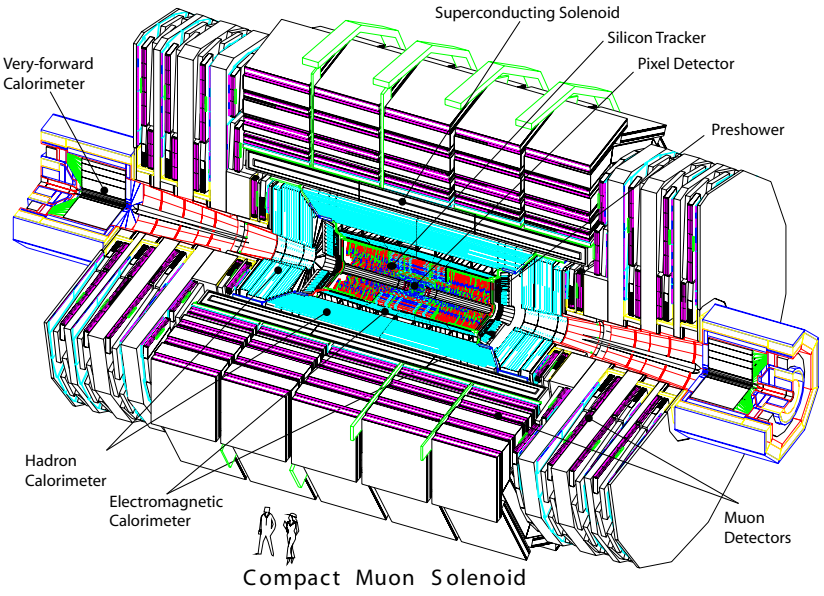


Figure 6.1: Schematic overview of the CMS detector [33].

The CMS design is driven by the need to identify interesting physics events with optimal efficiency while maintaining high background rejection. A fast detector response is also required as it is a high luminosity detector, expecting a luminosity of up to $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The detector granularity must be sufficient to allow for identification of the interaction under study with up to 23 inelastic collisions per bunch crossing

expected at design luminosity. Finally, the large flux of particles implies high radiation levels, therefore the detectors and electronics must be radiation hard.

6.1 Coordinate Conventions

The CMS coordinate system places the origin at the nominal collision point. The x-axis is perpendicular to the beam line, in the direction of the center of the LHC ring. The y-axis is vertical pointing upwards, and the z-axis is defined anti-clockwise along the beamline. The azimuthal angle ϕ is measured from the x-axis on the xy plane. The radial coordinate in the xy plane is denoted r. The polar angle θ is defined in the rz plane. Rapidity is defined as $y = \frac{1}{2}\ln\frac{E+p_z}{E-p_z}$ where E is energy and p_z is momentum in the z direction. Pseudorapidity is defined as $\eta = -\ln\tan(\theta/2)$. The momentum and energy components transverse to the beam direction, p_T and E_T respectively, are computed from the x and y components. The distance between two objects in the detector is given by $\Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2}$. The distance parameter is used to define the size of a cone around a particle track in the detector where ΔR is the radius of the cone.

6.2 Solenoid

The bending power of the magnet is defined by the requirement for a muon momentum resolution of $\frac{\Delta p}{p} \approx 10\%$ at $p = 1$ TeV. A superconducting solenoid magnet is installed in the CMS detector to provide a magnetic field of up to 4 T. During the 2011 running period the magnetic field remained at 3.8 T as a safety measure.

The solenoid magnet is 13 m in length with an inner diameter of 6 m. It is positioned in CMS such that the tracker and calorimeters are within its bore. The magnetic flux is returned through a 10,000 ton iron yoke made up of five wheels in the barrel, and two endcaps with three disks each. The return field saturates 1.5 m of iron, allowing for the integration of the muon detectors to ensure full geometric coverage.

6.3 Tracker

The tracker is intended for the precise and efficient measurement of charged particle trajectories and effective, high quality vertex reconstruction. For this purpose the tracker design includes fine granularity to resolve nearby tracks, high momentum resolution and high impact parameter resolution.

The tracker, shown in Figure 6.2, is made entirely of silicon with a total active area of 200 m^2 and coverage up to $|\eta| < 2.5$. Two types of detectors make up the tracker: a pixel detector, 4.4 cm $< r <$ 20 cm, and a strip detector, 20 cm $< r <$ 120 cm. To

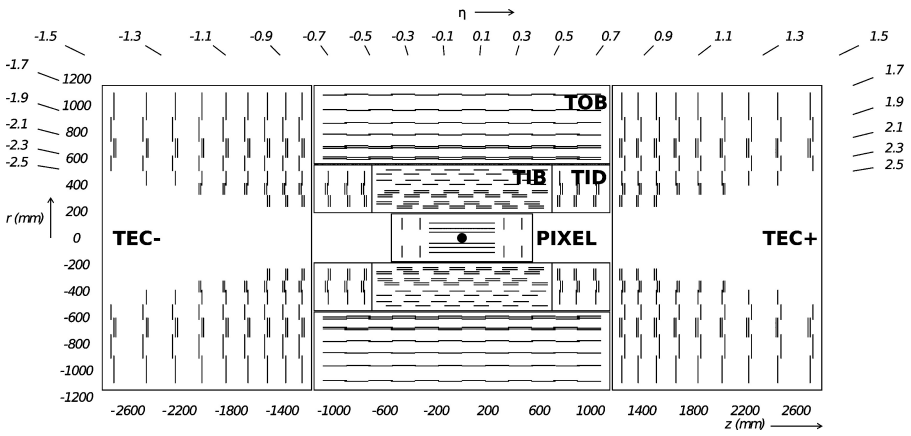


Figure 6.2: Tracker slice in r-z [35]. The tracker inner and outer barrel are labelled TIB and TOB respectively. TID and TEC denote the tracker inner disks and endcap.

mitigate radiation damage the tracker is designed to run at temperatures in the region of $-20°C$. Temperature stability is maintained by a mono-phase cooling system which uses C_6F_{14} as the cooling fluid.

The pixel detector is comprised of three barrel layers and two endcap disks, providing three hit coverage for $|\eta| < 2.2$ and two hit coverage for $|\eta| < 2.5$. The pixel size is $100 \times 150\,\mu\text{m}^2$ allowing for an occupancy of less than 10^{-4} at design luminosity, where occupancy is the fraction of detector channels with a hit. Accurate position resolutions are achieved by estimating the position of pixel clusters independently in both dimensions, based on the relative charges of the pixels at the edges of the cluster and the associated reconstructed track angle. The expected width of the charge distribution collected on the sensor surface is taken as input to the algorithm which determines the hit position. At design luminosity, the position resolution is expected to be better than $15\,\mu$m in the barrel transverse direction and between 15 and 30 μm in the barrel longitudinal direction and the endcap disks. The corresponding position resolutions achieved at $\sqrt{s} = 7$ TeV are shown in Figure 6.3.

The strip detector is comprised of four subsystems, providing up to 14 high precision measurements of track impact points in the region $|\eta| < 2.4$. The barrel section of the strip detector consists of a four-layer tracker inner barrel and a six-layer tracker outer barrel. At design luminosity, the tracker inner barrel is expected to have a position resolution of 23 to 24 μm in the r-ϕ direction and 230 μm in the z direction. The tracker outer barrel is expected to have a position resolution of 35 to 52 μm in the r-ϕ direction and 530 μm in the z direction. The corresponding position resolutions achieved at $\sqrt{s} = 7$ TeV are shown in Figure 6.3. Each endcap has three tracker inner disks and nine disks of tracker endcaps.

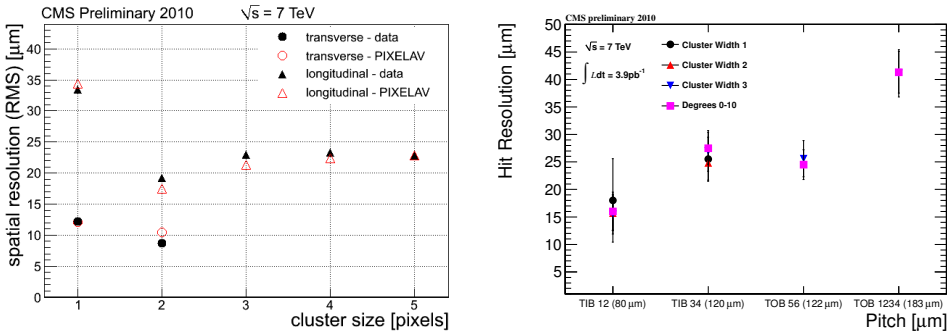


Figure 6.3: Transverse and longitudinal hit resolutions measured at $\sqrt{s} = 7$ TeV in the pixel tracker as a function of the cluster size, left. Hit resolution in the strip tracker measured at $\sqrt{s} = 7$ TeV as a function of strip pitch, right [36].

The particle density decreases with distance from the interaction point, therefore the strip detectors do not require the fine granularity of the pixel detectors. The tracker inner barrel and disks have sensors with length 10 cm and pitch 80 μm, where pitch is the inter strip distance. The pitch is larger in the outer barrel and endcap due to lower occupancy with increasing distance from the interaction region. The tracker outer barrel and endcap have sensors with length 25 cm and pitch 180 μm. This leads to an overall tracker occupancy of less than 3%. At design luminosity, the transverse momentum resolution, $\frac{\Delta p_{\mathrm{T}}}{p_{\mathrm{T}}}$, of the overall tracker is expected to be less than 3.5% for muons with $|\eta|$ less than 2.1 and a p_{T} of 100 GeV. Studies of the tracker performance in early LHC operation are described in [35].

6.4 Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL), Figure 6.4, is intended for the energy measurement of electromagnetically interacting particles, with a focus on electrons and photons. It is designed to achieve an energy resolution, $\frac{\Delta E}{E}$, of between 1.4% and 0.4% for electrons and photons with energy greater than 10 GeV. The energy resolution measured at $\sqrt{s} = 7$ TeV is shown in Figure 6.5. Further results for the comissioning of the ECAL at $\sqrt{s} = 7$ TeV can be found in [37].

The ECAL is made up of scintillation lead tungstate $PbWO_4$ crystals providing coverage up to $|\eta| < 3.0$. The barrel covers 25.8 radiation lengths while the endcap covers 24.7 radiation lengths. An additional preshower detector is placed in front of the endcap at $1.653 < |\eta| < 2.6$ to facilitate the rejection of photons produced in pairs in neutral pion decay.

$PbWO_4$ is chosen as the active material due to its fast response time and radiation

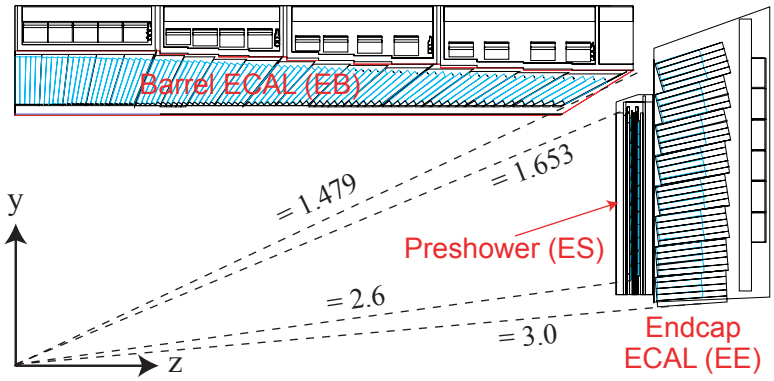


Figure 6.4: Transverse section through the ECAL [33].

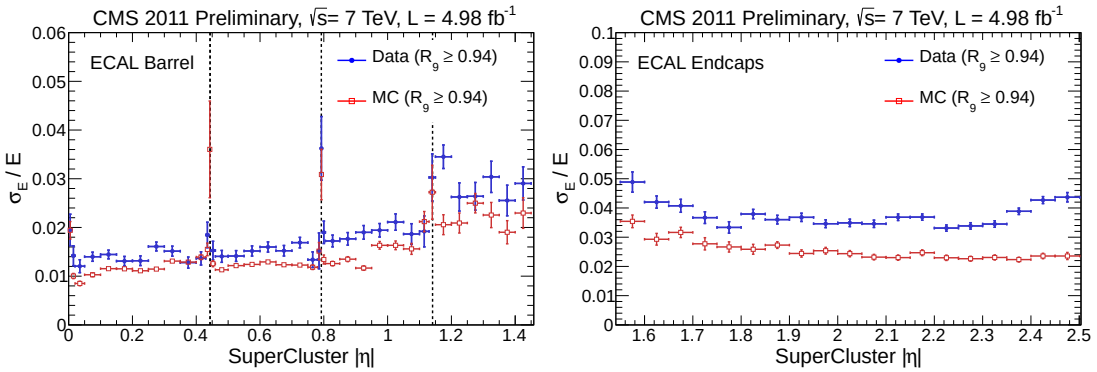


Figure 6.5: Relative electron energy resolution measured at $\sqrt{s} = 7$ TeV with respect to η in the barrel, left, and endcap, right [38].

hardness, with 80% of light emitted within 25 ns and a radiation hardness of up to 10 Mrad. The crystals also have a short radiation length of 0.89 cm and a Moliere radius of 2.2 cm. However, it also has a low light yield, 30 γ/MeV, and is temperature sensitive. To compensate for the low light yield the scintillation light produced in the crystals is read out by silicon avalanche photodiodes in the barrel and by vacuum phototriodes in the endcap. To avoid temperature dependent effects a nominal operating temperature of $18°\text{C} \pm 0.05°\text{C}$ is maintained by a water cooling system.

6.5 Hadronic Calorimeter

The Hadronic Calorimeter (HCAL), Figure 6.6, is intended for energy measurement with a focus on the measurement of the energy and direction of jets and missing transverse

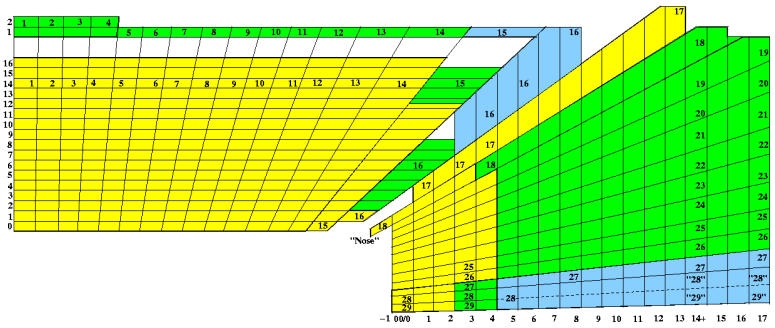


Figure 6.6: Schematic view of the tower mapping in r-z of the HCAL barrel and endcap regions [33]. The colours indicate the depth segmentation of the HCAL towers.

energy. To achieve these requirements the HCAL is designed to be fully hermetic with sufficient depth, greater than 10 interaction lengths, to contain hadronic showers and an energy resolution, $\frac{\Delta E}{E}$, between 0.10 and 0.12 for 100 GeV pions. The performance of the HCAL at $\sqrt{s} = 7$ TeV is outlined in [39].

The HCAL is a sampling calorimeter with alternating layers of passive absorbing brass plates and active plastic scintillator tiles arranged in trays. The HCAL covers $|\eta| < 3.0$. The first scintillator plate is placed before the first absorber plate to sample showers developing in the material between the ECAL and the HCAL. The last scintillator plate is positioned after the last absorber plates to correct for late developing showers. The majority of the HCAL is contained within the bore of the solenoid magnet, with an additional layer of scintillator material positioned outside the magnet. The magnet provides extra absorbing material giving the HCAL a thickness of 11 interaction lengths, allowing sufficient containment for most high energy hadronic showers.

A forward calorimeter is installed to cover $2.9 < |\eta| < 5.2$. The forward calorimeter is made from steel absorber plates with quartz fibers inserted as the active medium. It detects Cherenkov radiation emitted by charged particles in a shower, therefore it is mainly sensitive to the electromagnetic component of the shower.

6.6 Muon System

The muon system is designed for muon identification and momentum measurement. It is also intended for the effective triggering of events containing muons, requiring fast response times.

The muon system layout is shown in Figure 6.7. The muon system is composed of three separate gaseous detector technologies integrated into the magnet return yoke

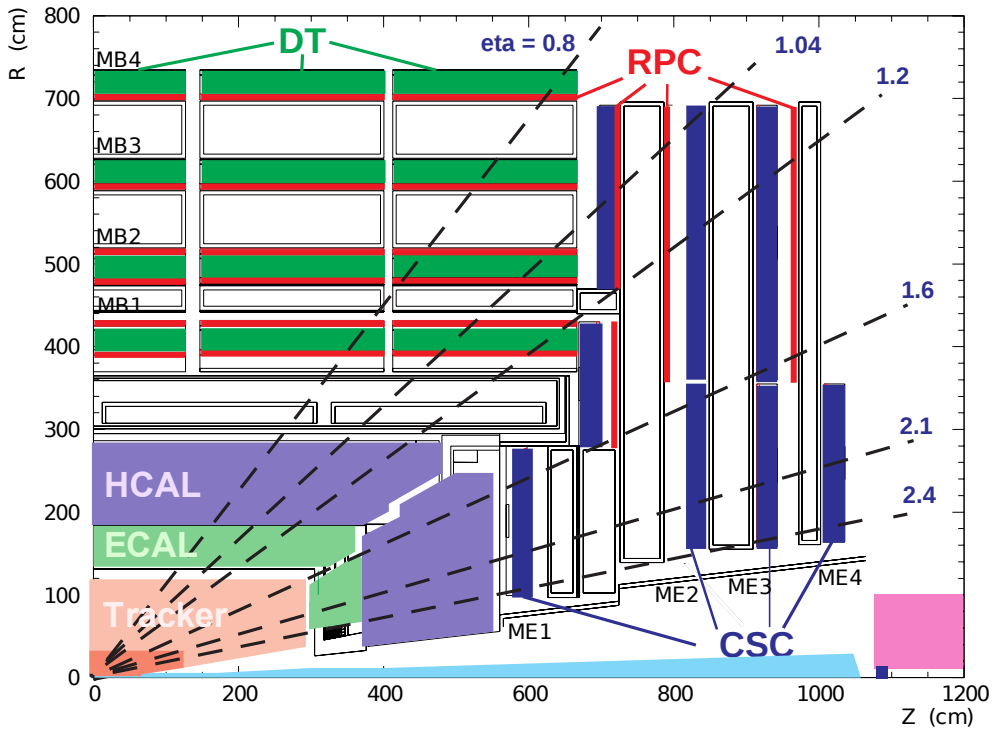


Figure 6.7: Layout of one quadrant of the CMS muon system [40].

and providing coverage up to $|\eta| < 2.4$. The choice of technology is driven by the magnetic field and neutron induced background in the different detector regions. The barrel component is contained in five separate iron wheels which are segmented into 12 sectors in ϕ. There are four concentric layers of muon stations. The endcap muon system is contained in four disks, with the muon detectors arranged in concentric rings, three in the innermost disk and two in the remaining disks.

6.6.1 Drift Tubes

The aluminium drift tubes (DT) cover the barrel, up to $|\eta| < 1.2$. In this region the residual magnetic field in the chambers, the muon rate and the neutron induced background rate are low. A DT cell consists of a gas tube with a positively charged stretched wire inside. There are four layers of DT chambers. In the inner three layers, DT chambers consist of twelve planes of DTs organised into three superlayers with four planes each, as shown in Figure 6.8. The central superlayer measures the z coordinate of the muon tracks. The two outermost superlayers are separated by about 20 cm and

measure the r-ϕ coordinate of the muon tracks. In the outer layer of DT chambers, the chambers consist of eight planes of DTs which measure the r-ϕ coordinate of the muon track. The point resolution is approximately 200 μm. Each station is designed to give a muon vector with a ϕ precision better than 100 μm in position and approximately 1 mrad in direction.

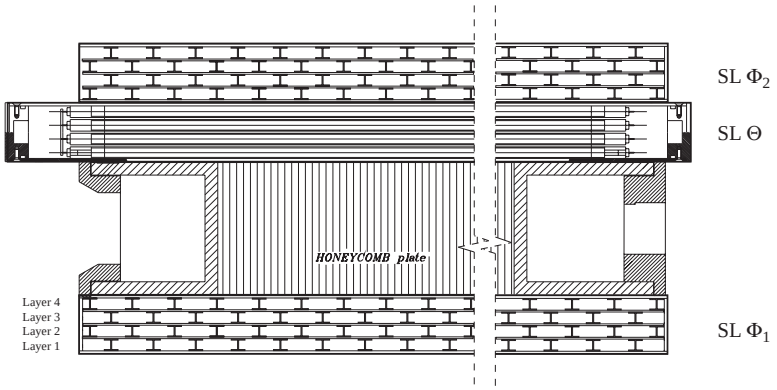


Figure 6.8: Layout of a DT chamber with three superlayers of DT cells [33].

6.6.2 Cathode Strip Chambers

The cathode strip chambers (CSC) provide coverage in the endcap, for $0.9 < |\eta| < 2.4$. In this region the magnetic field, muon rate and the neutron induced background rate are high. CSCs are used due to their fast response, high granularity and radiation resistance. A CSC station contains trapezoid shaped chambers consisting of six 9.5 mm gas gaps, each with a plane of anode wires placed perpendicular to a plane of radial cathode strips, shown in Figure 6.9. The strips are separated by about 0.5 mm while the wires are separated by about 3.2 mm. The copper strips provide a position measurement in the r-ϕ plane, the anode wires provide measurements of η. The spatial resolution provided by each CSC is approximately 200 μm, with an angular resolution in ϕ on the order of 10 mrad.

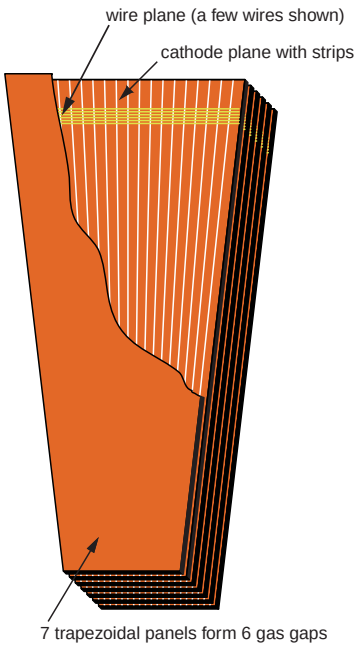


Figure 6.9: Schematic view of a CSC chamber [33].

6.6.3 Resistive Plate Chambers

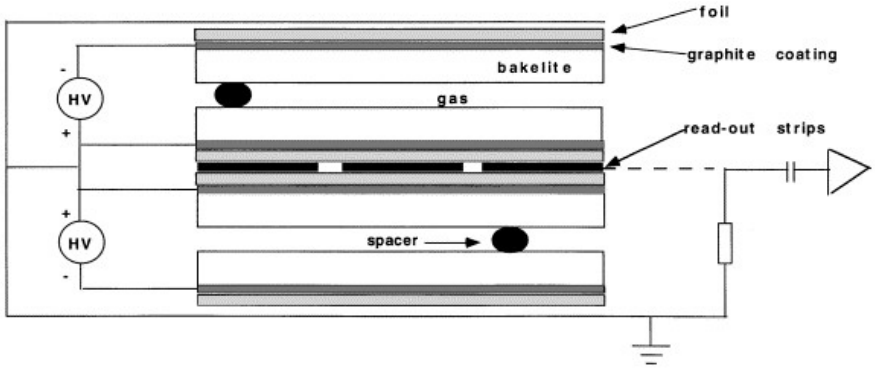


Figure 6.10: Schematic view of the RPC double gap structure [33].

The resistive plate chambers (RPC) provide coverage in the barrel and endcap, up to $|\eta| < 1.6$. They are highly segmented with a time resolution of 3 ns, allowing for accurate bunch crossing identification and prompt muon triggering. Each RPC detector consists of two 2 mm gaps filled with gas. The gaps are sandwiched by resistive bakelite

plates, as illustrated in Figure 6.10.

6.7 Luminosity Determination

The precise measurement of the $t\bar{t}$ cross section requires an accurate determination of the luminosity recorded by the CMS detector. The pixel cluster counting [41] method is used offline to determine the integrated luminosity recorded in 2011.

The pixel cluster counting method begins by determining the pixel cluster cross section, σ_{pixel}, with a Van der Meer scan [42], where σ_{pixel} is the interaction cross section multiplied by the average number of clusters per interaction. A Van der Meer scan determines the effective overlap area of colliding beams by scanning the beams through one another in the transverse direction. The absolute luminosity is then calculated from Equation 5.1. The value of σ_{pixel} is determined in the Van der Meer scan using

$$\sigma_{\text{pixel}} = \langle \text{N}_{\text{cluster}} \rangle f(\mathcal{L})^{-1} \tag{6.1}$$

where $\mathcal{L}$ is the instantaneous luminosity from Equation 5.1 estimated with the results of the Van der Meer scan, $f = 11,246$ Hz is the LHC orbital frequency and $\langle \text{N}_{\text{cluster}} \rangle$ is the mean number of pixel clusters per zero-bias trigger. The only requirement on a zero-bias trigger is that the proton bunches from each beam pass through each other.

Once σ_{pixel} has been calculated, the integrated luminosity for a given luminosity section is given by

$$\int \mathcal{L} = \frac{\langle \text{N}_{\text{cluster}} \rangle n_{\text{bx}} t_{\text{ls}} f}{\sigma_{\text{pixel}}}. \tag{6.2}$$

The number of active bunch crossings, n_{bx}, is determined from data by selecting crossings where at least 20% of zero-bias triggered events contain at least one well defined primary vertex. The duration of a single luminosity section is given by $t_{\text{ls}} = 2^{18}$ orbits ≈ 23.31 s.

The pixel data is corrected for afterglow effects from late arriving particles and energy originating from activated detector material. The afterglow correction corresponds to a subtraction of $\approx 2.8\%$ of the integrated luminosity per luminosity section for a typical 2011 fill with 1380 bunches.

The dominant systematic uncertainty on the measured luminosity is from scan-to-scan variations in σ_{pixel}, resulting in an uncertainty of 1.5%. The uncertainty on the afterglow correction is also a significant source of systematic uncertainty at 1%.

Length-to-scale corrections are applied to the beam widths measured in the Van der Meer scans. The uncertainty on these corrections results in a 0.5% uncertainty on the luminosity. The beam widths are extrapolated in time over the duration of the scale, resulting in an uncertainty of 0.5% due to beam width evolution.

The effect of pixel detector gain and pedestal changes on the cluster counts is determined by comparing the fraction of cluster counts as a function of time throughout 2011, resulting in a systematic uncertainty of 0.5%. The effect of front-end buffer overflow at high instantaneous luminosity, referred to as dynamic inefficiencies, is estimated to be 0.4%. Small variations in the fractional cluster counts across pixel detector regions translate to a 0.3% systematic uncertainty.

Three contributions to the uncertainty on the beam intensity are considered. An uncertainty of 0.3% originates from the absolute calibration of each individual bunch while an uncertainty of 0.5% is related to the current contributions of each individual bunch. This affects the measured luminosity as the Van der Meer scan is analysed for individual bunches. An uncertainty of 0.2% relates to beam charge not in the nominally-filled bunch slots.

The total uncertainty on the measured luminosity is obtained by adding the uncertainty from each source in quadrature. The integrated luminosity of certified data collected in 2011 by the muon trigger used in this analysis is estimated with the pixel cluster counting method as $\int \mathcal{L} = 4.76$ fb^{-1} $\pm 2.2\%$. This is less than the total luminosity of 5.56 fb^{-1} recorded by the CMS detector as some of the CMS subsystems were not operational at certain times during data taking. Certified data only includes luminosity sections recorded with a fully operational CMS detector.

Chapter 7

CMS Trigger System

At nominal operating conditions, with a bunch spacing of 25 ns, the CMS detector expects a bunch crossing rate of 40 MHz. Due to the limit of the archive rate of the online computer farm, only a few hundred Hz of these events may be stored for further study. The decision to keep an event is made by the CMS trigger system [43, 44].

The trigger is the first step in the selection of interesting physics events. If an event is discarded by the trigger it can not be retrieved. Therefore, the trigger must be highly efficient in selecting events which would be selected after the full offline reconstruction. Events which are interesting for new physics searches should also be stored. In addition, CMS needs certain events for calibration and monitoring of the detector.

In order to remain within the bandwidth of the computer farm, these events must be selected with high purity. Furthermore, the decision to keep an event must be made before the next bunch crossing, every 50 ns in 2011. This limits the complexity of the object reconstruction algorithms which can be employed, thus the requirement of a high purity, high efficiency selection is a challenge.

Computing time is optimised by discarding uninteresting events as quickly as possible. In the CMS trigger system the decision to store or discard an event is made in two stages, at Level-1 (L1) and in the High Level Trigger (HLT).

7.1 Level-1

The Level-1 trigger system is hardware based and, at the 2011 bunch spacing of 50 ns, reduces the bunch crossing rate of 20 MHz to 100 kHz for processing by the HLT. This rate is set by the speed of the detector electronics readout and the rate at which data can be accepted by the HLT.

The data available for the L1 decision can be stored for no more than $3.2\,\mu$s due to constraints on signal propagation from electronics technology. In this time, data is collected from the front end electronics, the L1 decision is made, and data is propagated to the readout electronics front end buffers. The L1 trigger calculations must be done in less that 1 ms.

The time restriction means preshower and tracker data can not be used in the L1 decision. Extensive processing or corrections to the data are also not feasible. Instead, only a subset of the available detector information is used at L1.

The L1 trigger is made up of three subsystems: the L1 calorimeter trigger, the L1 muon trigger and the L1 global trigger. The L1 calorimeter and muon triggers reconstruct coarsely segmented data from the calorimeter and muon detectors respectively. The global trigger decides to store or reject an event based on trigger data from the calorimeter and muon triggers. Coordinate information in η-ϕ space is available at L1, allowing for the variation of thresholds based on the location of the trigger object. Thresholds applied at L1 are more relaxed than HLT thresholds to allow for low L1 momentum resolution due to the limited detector information used in L1 reconstruction.

The L1 decision is transmitted through the trigger throttle system (TTS) to the timing, trigger and control (TTC) system. If the bandwidth of a particular trigger needs to be reduced, a prescale is applied in the TTS. A prescale limits the amount of events passing the L1 trigger which are sent to the HLT by assigning a probability, the inverse of the prescale value, that the L1 triggered event is forwarded to the HLT. The TTS can also shut off the L1 accept signal in case the detector readout or DAQ buffers are at risk of overflow. The TTC transmits the decision to all detector subsystem front end and readout systems. All trigger objects found at L1 are sent to the HLT if any of the objects pass the trigger decision. The triggered object is referred to as the L1 seed.

7.2 High Level Trigger

The High Level Trigger system reduces the event rate from 100 kHz to a few hundred Hz for writing to mass storage systems. The HLT runs an optimised version of the full CMS reconstruction software. Since the HLT is software based, it is possible to modify and improve the system during data taking. Reconstructed HLT objects should be as close as possible to objects reconstructed offline to allow for tight selection criteria without removing interesting events.

The real time nature of the selection constrains the resources available for object reconstruction, limiting the performance of the reconstruction algorithms. The selection is optimised by rejecting events as quickly as possible. First, the minimum detector information necessary for background rejection is reconstructed. Then, a partial event reconstruction is performed using a limited region of the detector. The region to be reconstructed is based on the L1 seed.

The HLT complex processes all events accepted by the L1 trigger in a single processor farm. However, the reconstruction and selection of events takes place in three software based steps. Level-2 (L2) uses information from the calorimeter and muon detectors. Level-2.5 uses partial tracker information. Level-3 includes reconstruction of full tracks in the tracker. Track reconstruction is performed in the final step due to its time consuming nature. Having the three software based steps on a single processor farm adds flexibility to the HLT system, allowing for selection improvements or adjustments to unforeseen circumstances.

7.3 Muon Trigger System

7.3.1 L1 Muon Trigger

The L1 muon trigger system is designed to identify muons, determine their location and transverse momentum, and assign them to a beam crossing. It is made up of three subsystems, one for each of the muon detectors: the DT trigger in the barrel, the CSC trigger in the endcap and the RPC trigger in the barrel and endcap.

The DT and CSC electronics first process information from each station locally. Information on the position, direction, quality and assigned bunch crossing for each muon in each station is sent to the track finder. The DT and CSC track finders build tracks and assign p_{T}, exchanging information between the subsystems in the region of overlap. Up to four muon candidates, selected based on p_{T} and quality, are forwarded from each subsystem to the L1 global muon trigger.

Hits from the RPC stations are collected by a pattern comparator trigger, which is based on the spatial and time coincidence of hits in the RPC muon system. If the hits are time coincident with patterns aligned along a possible muon track, a muon candidate is formed and p_{T} assigned. The muon candidates are sorted by p_{T} and quality. Up to eight candidates, four from the barrel and four from the endcap, are sent to the L1 global muon trigger.

The L1 global muon trigger combines information from the DT, CSC, RPC and calorimeter trigger systems. DT and CSC candidates are matched with RPC candidates based on their proximity in η-ϕ space. If matched, the candidate information is combined for additional precision, otherwise candidates are stored or discarded based on quality information. The muon tracks are extrapolated back to the calorimeter trigger towers to obtain isolation information. Muon candidates are sorted based on quality, subsystem, p_{T} and η. The top four muon candidates are sent to the L1 global trigger for the final L1 decision.

Quality code	Meaning
7	DT/RPC or CSC/RPC matched candidate
6	DT or CSC unconfirmed candidate
5	RPC unconfirmed candidate
4	very low quality type 3
3	very low quality type 2
2	very low quality type 1
1	halo muon
0	no track

Table 7.1: Quality codes for L1 muon candidates.

The three bit quality code defines the quality of a muon candidate. One of seven quality codes is assigned to each global muon trigger candidate, with a higher code indicating a higher quality candidate. The meaning of each code is described briefly in Table 7.1. A halo muon is a muon which is detected when the halo of particles around one of the proton beams interacts with the CMS detector material.

7.3.2 Muon HLT

L1 seeds forwarded by the L1 global trigger are examined by the HLT, regardless of which object was triggered. The muon track is reconstructed in the muon system at L2, refining the p_{T} measurement and confirming the L1 decision.

L3 reconstruction extends the muon trajectories to include hits from the silicon tracker system, further refining the p_{T} measurement. First, the L2 muon trajectory is extrapolated from the innermost muon station to the outer track surface. Then, a region of interest is defined for track reconstruction by determining which silicon layers are compatible with the muon trajectory. Finally, the reconstructed tracks are fit to the muon reconstructed at L2 with a Kalman-filter technique [45] to achieve a L3 global muon. The p_{T} and η requirements of the HLT are applied to the L3 muon.

Calorimeter isolation requirements are applied after L2 reconstruction. The calorimeter isolation is obtained by summing the calorimeter energy in a cone around the muon. The cone axis is defined as the muon direction at the impact point. The muon contribution is subtracted from the sum by removing the energy deposit in a small cone around the muon. This technique becomes less effective at high luminosity as more pile-up is included in the sum.

Isolation based on the pixel tracker information is determined after L3 reconstruction. This step is computationally time intensive as it requires full tracks to be reconstructed regionally. The tracker isolation selection is based on the sum of the p_{T} of tracks in a cone around the muon after removing the muon contribution. This definition of isolation is less sensitive to pile-up as only tracks from the same collision vertex as the muon are considered.

7.4 HLT_Mu40 in 2011

HLT_Mu40 is a single muon trigger with a p_{T} threshold of 40 GeV. The relatively high p_{T} threshold of 40 GeV is necessary to keep the trigger rate within the available bandwidth.

Trigger requirements evolve with LHC luminosity, remaining loose at low luminosities to maximise physics output. New trigger menus are introduced for significant increases in the LHC luminosity. A trigger menu defines all the available triggers, their requirements and prescales. To remain within bandwidth the higher luminosity menus

increase thresholds, tighten identification criteria, and, if nothing else will suffice, apply prescales.

The 40 GeV p_{T} threshold for HLT_Mu40 survived the menu changes in 2011. However, the identification criteria were tightened for high luminosity menus in order to maintain the p_{T} threshold while remaining within bandwidth. The trigger menus for 2011 are shown in Table 7.2, along with the amount of data collected by HLT_Mu40 with each menu.

Menu	$\int \mathcal{L}$ (pb^{-1})	First run number	Start date (dd/mm)
1e33	954	165088	15/05
2e33	831	170249	16/07
3e33	2092	173236	12/08
5e33	884	178420	13/10

Table 7.2: Luminosity collected with each of the trigger menus used to collect data for this analysis. HLT_Mu40 was used to collect data in menus 1e33 and 2e33 while HLT_Mu40_eta2p1 was used for the 3e33 and 5e33 menus.

Starting from the 1e33 menu, the L1 seeds for HLT_Mu40 are required to pass the L1_SingleMu16 trigger. The L1_SingleMu16 trigger has a p_{T} threshold of 16 GeV and requires muon candidates to have a quality code of at least four. The reconstructed trigger object must be detected by segments in at least two muon stations to obtain a meaningful estimate of the muon p_{T}.

The 2e33 menu sees the introduction of quality criteria at L2 to suppress an untenable trigger rate. Selected muon tracks in the region $0.9 < |\eta| < 1.5$ or $|\eta| > 2.1$ are required to have at least two muon stations with matched segments and at least one valid hit in the muon system. This suppresses triggering of unwanted muon candidates and also improves the behaviour of the muon trigger cross section with respect to pile-up.

The 3e33 menu is introduced after a technical stop of the CMS detector. Changes to the configuration of the detector made during the technical stop, such as the deactivation of some muon chambers, affect the performance of the trigger. A prescale is applied to HLT_Mu40 due to increasing trigger rates. A separate trigger, HLT_Mu40_eta2p1, is available which is a duplicate of the HLT_Mu40 trigger with the exception of an additional $|\eta| < 2.1$ requirement. The η restricted trigger remains unprescaled, therefore it is used by analyses studying centrally produced muons. The L1 seeds for HLT_Mu40_eta2p1 are required to pass the L1_SingleMu16_eta2p1 trigger.

Changes to the muon p_{T} assignment in the L1 global muon trigger are introduced in a revision of the 3e33 menu. When the muon candidate passes through several subsystems the GMT must make a decision on which p_{T} should be assigned. In earlier menus the

minimum p_{T} was chosen. However, choosing the p_{T} of the subsystem candidate which returns the highest muon quality is found to result in higher trigger efficiencies.

In the 5e33 menu, the p_{T} assignment of muons at HLT is based on a tracker fit instead of the global fit (tracker and muon system) previously used. This change reduces the trigger rate without noticeable efficiency loss. For muons with $p_{\mathrm{T}} < 200\,\mathrm{GeV}$ the tracker fit has better resolution than the global fit. Also, the p_{T} from a tracker fit shows better agreement with the offline reconstructed muon p_{T} .

The measurement of the trigger efficiency is an important ingredient in the analysis of the $t\bar{t}$ production cross section and is presented in Chapter 11. HLT_IsoMu24, an isolated muon trigger with a lower p_{T} threshold of 24 GeV, is described in Appendix B.

Part III

Simulation and Reconstruction

Chapter 8

Event Simulation

In order to validate SM predictions, it is necessary to be able to compare the collected data with the theoretical expectations. This is achieved through the simulation of collision events with Monte Carlo (MC) event generators. Event generators simulate everything from the initial collision to the final state particles.

The detector response to the final state particles is then simulated to allow for comparison of the observed data with what is expected from specific physics processes. Simulated events are used for detector design optimisation, calibration, object identification and physics analysis. Therefore, the accurate simulation of physics events and the detector response is crucial.

8.1 Event Generators

There are many event generators available to simulate physics events [46]. The optimal choice of generator depends on the relevant physics process. The event generators used in the cross section measurement are described in this section.

8.1.1 PYTHIA

PYTHIA [28] is a general purpose generator which can be used to simulate hadronic events in pp collisions. It provides full event simulation including the hard and soft interactions, parton distributions, initial and final state parton showers, hadronisation, decay and the underlying event. It may also be used for hadronisation of events which have been generated separately at parton level. QCD multijet background events are simulated by PYTHIA. For the simulation of the remaining physics samples a separate event generator is used in combination with PYTHIA.

The first step in event generation is the evaluation of the hard process. These processes are mainly $2 \to 2$ reactions at LO, with some $2 \to 1$ or $2 \to 3$ reactions.

The event then proceeds through parton shower (PS) simulation to include higher order effects. A PS occurs when a parton, which has been strongly accelerated by the hard subprocess, emits radiation in the form of gluons. The initial parton and the emitted gluons lose energy through the radiation of gluons. The strong coupling constant, α_s, increases as the partons lose energy, eventually leading to hadronisation. ISR and FSR are modelled by showering the initial and final state partons.

Partons are hadronised via the Lund string model [47]. In the Lund string model quarks or other colour triplets are considered to be located at the ends of string, with gluons as energy and momentum carrying kinks in the string. The string breaks in the production of quark-anti-quark pairs, with the quark from one break combining with an anti-quark from an adjacent break to form a colour singlet hadron. Particles with a short lifetime which are present at this point are decayed.

After full simulation the final state hadrons, leptons and photons are stored. The characteristics of intermediate particles such as quarks and bosons are also available to allow for the history of the final state particles to be traced back.

8.1.2 MadGraph

MadGraph [48] is a matrix element (ME) generator which computes tree level matrix elements with a fixed number of partons in the final state. The final state partons consist of bare quarks and leptons which are then delivered to PYTHIA for PS simulation. Top pair production is simulated with MadGraph then delivered to PYTHIA. W/Z + jets events are simulated in the same way, with up to four additional partons from MadGraph.

MadGraph ME calculations are computationally expensive and do not include virtual loops. However, MadGraph simulation is valid when partons are hard and well separated whereas the PYTHIA PS model is valid when partons are collinear or soft. Therefore, MadGraph is used for physics processes with multiple jets in the final state.

When combining ME and PS calculations it is important to avoid double counting events. Even if the initial number of partons is orthogonal, additional partons can arise due to parton showering. This results in an overlap in the final jet multiplicity states produced by each method. Double counting is avoided via matching, whereby a decision to take higher orders from MadGraph or PYTHIA is made on an event by event basis.

For the simulated samples used in this analysis the ME to PS matching is performed with the k_t-MLM matching scheme [49]. This method matches partons from the ME calculation to jets reconstructed after the perturbative shower. The matching threshold which separates the phase-space for ME or PS modelling is based on the E_T of the parton. Partons above the threshold are modelled by ME calculations. After PS simulation, particles are clustered into jets using the threshold. The jets are then matched to ME partons. The event is rejected if each jet does not have one matching parton.

8.1.3 POWHEG

POWHEG [50] is a hard event generator for heavy quark production, accurate to NLO. POWHEG is interfaced with PYTHIA to implement NLO calculations while maintain-

ing both the leading log and NLO accuracy of the PS. Single top background events are simulated with POWHEG interfaced with PYTHIA.

8.2 Detector Simulation

The CMS detector simulation is based on the GEANT4 [51] toolkit. GEANT4 is a toolkit for the simulation of the passage of particles through matter. It models the interaction of particles with the detector material, decays of long-lived particles, and allows to describe the full CMS detector geometry and interfaces required to retrieve information from particles travelling through the detectors and the magnetic field.

PYTHIA output is taken as input for the detector simulation. First, the CMS detector response is simulated to produce hits in the sensitive detector elements. Then the hits are converted to digitisations corresponding to the electronic readout used to acquire data by the detector and DAQ systems.

8.3 Signal and Background Modelling

$t\bar{t}$ and W/Z + jets events are simulated as described in Section 8.1.2. The simulation of single top events is described in Section 8.1.3. The muon enriched sample of QCD multijet background events is produced with PYTHIA, with a minimum p_{T} of 20 GeV for outgoing muons on matrix element level. The multijet sample includes muons from the decay of b and c quarks as well as muons from decays in flight of pions, kaons and K-longs. A filter is applied to the simulated sample requiring a final state muon with p_{T} greater than 15 GeV. The technical details for each of the simulated samples are documented in Appendix A.

Event simulation requires input for parameters such as the quark and boson masses. The input parameters are summarised in Table 8.1. The trigger response is simulated for the 3e33 trigger menu. The factorisation and renormalisation scale is defined as $Q^2 = \sum m_t^2 + p_{\mathrm{T}}^2$ in MadGraph and $Q^2 = m_t^2$ in POWHEG.

CTEQ [16] PDF libraries are used in the simulation of events. The CTEQ6L1 PDFs are used in MadGraph event generation while POWHEG uses CTEQ6M PDFs. However, the CTEQ6L1 library does not include PDF uncertainties. In this analysis, the simulated events are reweighted from CTEQ6L1 or CTEQ6M to CTEQ6.6 to make use of the error sets provided by the latter.

Parameter	Value
t mass	172.5 GeV
W mass	80.398 GeV
W width	2.141 GeV
Z mass	91.1876 GeV
Z width	2.4952 GeV
b mass	4.8 GeV
c mass	1.27 GeV

Table 8.1: Parameters input in the simulation of physics processes at $\sqrt{s} = 7$ TeV.

Chapter 9

Event Reconstruction

Once an event has been triggered, the event information is available in the form of detector electronic readout and kept on mass storage for analysis. This information is available in the same format for generated events which have been passed through the detector response simulation. This chapter describes the reconstruction of the physics objects which produce the detector signal.

9.1 The Particle Flow Algorithm

The particle flow algorithm [52, 53] is used to identify and reconstruct stable particles arising from the pp collision. The algorithm combines information from all sub detectors for the optimal determination of the direction, energy and type of each particle.

Specifically, the particle flow algorithm aims to individually reconstruct all electrons, muons, photons, charged hadrons and neutral hadrons in the event. This is achieved with a combination of information from charged particle tracks in the silicon tracker, calorimeter clusters from the ECAL and HCAL, and muon tracks from the muon system. The individual reconstructed particles are then used to construct higher level objects such as jets.

Particle flow event reconstruction proceeds in three stages. First, the fundamental elements of the event are reconstructed. The charged particle tracks, calorimeter clusters and muon tracks must be delivered with high efficiency and low fake rate for the algorithm to be successful. Advanced tracking and clustering algorithms have been developed to achieve this.

Secondly, the fundamental elements are topologically linked in blocks. Finally, these blocks are interpreted in terms of particles. The first and second stages are covered in this section. The interpretation of the blocks in terms of particles is then described for each type of particle.

9.1.1 Iterative Tracking

The momentum of a charged particle is measured in the silicon tracker with high efficiency and low fake rates for a p_{T} range of 150 MeV up to several hundred GeV. Reconstructed tracker tracks provide a precise measurement of the charged particle direction at the production vertex, before any deviation by the magnetic field.

An iterative tracking procedure [54] with six iterations is employed for track reconstruction. In the first iteration, hits are required in three pixel layers to produce a seed for track reconstruction. This leads to a negligibly small fake rate with moderate tracking efficiency. In the second iteration, hits which have been assigned to high quality tracks in the previous step are removed. Seeds are selected from the remaining hits by requiring hits in two pixel layers, thereby increasing the tracking efficiency. The fake rate remains low due to the reduced combinatorics from the hit removal.

The majority of high p_{T} tracks from the primary production vertex are reconstructed in the first two iterations. The remaining iterations are designed to pick up lower p_{T} tracks. In the third and fourth iterations the p_{T} and beam spot constraints on track seeding are loosened. The final two iterations use seeds from the strip tracker layer, allowing for reconstruction of decay particles produced outside of the pixel tracker.

The iterative tracking strategy successfully reconstructs charged particles with as few as three tracker hits, p_{T} as low as 150 MeV, arising from a vertex as far as 50 cm from the beam axis [52]. The fake rate is on the order of 1%. The muon tracking efficiency measured at $\sqrt{s} = 7$ TeV is above 99%, as shown in Figure 9.1.

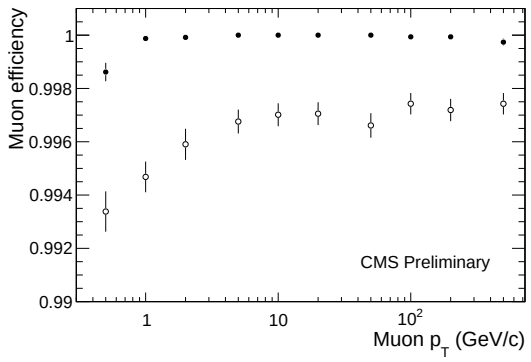


Figure 9.1: Muon tracking efficiency measured at $\sqrt{s} = 7$ TeV as a function of p_{T} for isolated muons in data (open circles) and simulation (solid circles) [55]. The tracking efficiency is above 99% for the full range of muon p_{T}. The efficiencies measured in data and simulation differ by less than 0.5%.

9.1.2 Calorimeter Clustering

Stable neutral particles, such as photons and neutral hadrons, are detected in the calorimeters. The calorimeters can also improve the measurement of charged hadrons for which track parameters are not determined accurately. Electrons, as well as all accompanying Bremsstrahlung photons, are detected in the calorimeters.

The calorimeter clustering algorithm aims to achieve high detection efficiency over a wide range of particle energy. Close energy deposits must be distinguished, as well as those from neutral hadrons or charged hadrons. Clustering is performed separately for the ECAL endcap, ECAL barrel, HCAL endcap, HCAL barrel, preshower first layer and preshower second layer. No clustering is performed in the forward HCAL, with each cell counted as one cluster.

Cluster seeds are individual calorimeter cells selected by finding cells with more energy than their adjacent cells and energy above a given threshold. Topological clusters are then constructed, starting from a cluster seed, by including all cells neighbouring a cell already in the cluster and with energy above a given threshold. The threshold is set to represent two standard deviations of electronic noise.

One topological cluster can result in multiple particle flow cluster seeds, allowing for overlap. The energy of each calorimeter cell is shared among all particle flow clusters according to the distance between the cell and the cluster, with an iterative determination of the cluster energies and positions. This makes maximal use of the calorimeter granularity.

9.1.3 Topological Linking

The particle flow elements from each sub detector are linked for the complete reconstruction of each individual particle. This also removes the risk of double counting from the different sub detectors. The linking algorithm produces blocks of linked elements. The linked elements are input to particle reconstruction and identification algorithms for muons, electrons and jets.

Silicon tracker tracks are linked to particle flow calorimeter clusters by extrapolating the track into the calorimeters. A link is established if the extrapolated track is within the cluster boundaries.

The silicon tracker induces significant Bremsstrahlung photon emission from electrons. Due to the magnet bending of the charged electron tracks, the electron and photon energy deposits in the calorimeters can be widely separated in ϕ, though not η. To connect energy from emitted photons to the electron track, tangents to the track are extrapolated to the ECAL from the intersection points between the track and each of the tracker layers. If the extrapolated tangent is within the boundaries of a cluster, the cluster is linked to the charged track.

Clusters from separate calorimeters are linked if the cluster from the more granular calorimeter (preshower or ECAL) is within the envelope of the cluster in the less granular calorimeter (ECAL or HCAL).

A global fit is performed to link muon tracks with tracker tracks. A link is established if the χ^2 of the fit is within a given limit. The corresponding muon candidate is referred to as a global muon.

9.2 Reconstructed Muons

Reconstructed muons [56] are categorised as either standalone, tracker or global muons. A reconstructed muon from any category is then filled to a collection of particle flow muons if it passes additional quality criteria. Standalone muons are reconstructed using only information from the muon system while tracker muons only use information from the tracker. Global muons combine information from the standalone muon tracks and silicon tracker tracks to improve the momentum resolution of the muon candidate.

Standalone muon reconstruction begins with the creation of track segments. Track segments are constructed with a linear fit the the position of hits in each layer of an individual DT or CSC muon chamber. Track segments from the DT chambers, track segment hits from the CSC chambers and hits from the RPC chambers are used to build muon trajectories with a fit based on the Kalman-filter technique [45].

Tracker muon reconstruction considers all tracker tracks with $p_{\mathrm{T}} > 0.5$ GeV and $p > 2.5$ GeV as possible muon candidates. Each tracker track is extrapolated to the muon system and counted as a tracker muon if at least one muon segment matches the extrapolated track position.

In preparation for global muon reconstruction, the standalone muon tracks are extrapolated from the inner muon station to the outer tracker surface. Each standalone muon is matched to a tracker track. The global muon trajectory is built by combining hits from the tracker tracks and the standalone muon track using the Kalman-filter technique.

9.2.1 Particle Flow Muons

Reconstructed muons include a significant amount of misidentified charged hadrons. The particle flow algorithm applies selection requirements to the reconstructed standalone, tracker and global muons to obtain a pure sample of muon candidates. The collection of particle flow muons consists of three subcategories of particle flow muons selected at the event reconstruction stage: isolated, pf-tight and pf-loose. The subcategories are defined here within the context of particle flow muon reconstruction and are not used later in the analysis.

In the isolated particle flow muon selection, the sum p_{T} of the tracks and transverse energy of the calorimeter hits within a cone size of 0.3 are required to be less than 10% of the muon p_{T}. The isolated muon is required to have been successfully reconstructed as a global muon in Section 9.2. The isolation requirement restricts the amount of neighbouring particles so to avoid loss of muon selection efficiency no further selection requirements are applied.

The pf-tight and pf-loose muon requirements are applied to reconstructed muons failing the isolated selection. Pf-tight muons are required to have a minimum number of hits in the muon track, and muon segments compatible with calorimeter deposits.

Pf-loose muons have a looser requirement on the number of hits and the tracker track is required to be compatible with hits in the muon stations.

At a center of mass energy of $\sqrt{s} = 7$ TeV, the fake rate for particle flow muons, including all three subcategories, with $p_{\mathrm{T}} > 20$ GeV is found to be on the order of 0.0004 [53]. The identification efficiency for particle flow muons with respect to reconstructed muons is greater than 99% for muons from W boson decay.

The particle flow muon tracks are removed from the block of linked particle flow elements for further processing.

9.3 Electrons

An ECAL-driven electron reconstruction algorithm is designed to form super-clusters of ECAL energy. ECAL energy deposits within an η-ϕ window are combined into a super-cluster and used to seed the reconstruction of the electron track. Electron track reconstruction is performed with the Gaussian-Sum Filter [57].

9.3.1 Particle Flow Electrons

Bremsstrahlung photons emitted from the electron can convert to an e^+e^- pair in the detector material. The first step in particle flow electron reconstruction is the identification of tracks due to electrons from photon conversion and recovery of the corresponding ECAL energy deposits. In the case of isolated electrons, the ECAL-driven super-cluster from Section 9.3 is then used to recover nearby clusters. The electron is identified and its momentum reconstructed by combining the track and recovered cluster observables.

Tracks and ECAL clusters which have been assigned to an electron are removed from the block of linked particle flow elements for further processing.

9.4 Particle Flow Hadrons and Photons

Charged hadrons, neutral hadrons, photons and, more rarely, additional muons are reconstructed from the remaining blocks of linked particle flow elements. Tracks are discarded if the relative uncertainty on the measured track p_{T} is greater than the relative calorimeter energy resolution expected for charged hadrons. 90% of tracks rejected with this requirement are fake tracks, i.e tracks which have been misreconstructed by the iterative tracking procedure described in Section 9.1.1. The energy of rejected tracks from real particles is measured independently, with better precision, in the calorimeters.

Neutral hadrons and photons are detected by comparing the momentum of tracker tracks and energy detected in the calorimeters in order to link tracks to ECAL or HCAL clusters.

A relaxed search for muons and fake tracks is performed if the total calorimeter energy is more than three standard deviations less than the particle flow linked total track momentum. Global muons, which have not already been identified, with a momentum precision better than 25% are considered to be particle flow muons. The remaining tracks are ordered according to p_{T} uncertainty and then progressively removed. The removal stops when the p_{T} uncertainty reaches 1 GeV or when the removal would make the total track momentum less than the calorimeter energy.

Any tracks remaining after the removal of fake tracks are considered as charged hadrons.

Further considerations for neutral hadrons and photons are necessary when the total energy of the particle flow calorimeter cluster linked to a track is significantly larger than the total associated charged particle momentum. If the relative energy excess is more than the expected calorimeter energy resolution, it is identified as a photon or neutral hadron. If the energy excess is more than the total ECAL energy, a photon is created with the ECAL energy and a neutral hadron with remaining part of excess. Otherwise only a photon is identified.

Any remaining ECAL or HCAL clusters are identified as photons or neutral hadrons respectively.

9.5 Jets

A quark or gluon in the event will hadronise, resulting in photons, hadrons, muons and/or electrons in the detector. These particles must be combined into a so-called jet to reproduce the initial hard quark or gluon. This particle reclustering is performed by the anti-k_t algorithm [58].

The anti-k_t algorithm takes the four-momentum of the reconstructed particle flow particles as input. The particles are clustered based on the distance parameters, d_{ij}, between particles i and j, and d_{iB}, between particle i and the beam (B). The distance parameters are defined as

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{R^2}, \tag{9.1}$$

$$d_{iB} = k_{ti}^{2p}, \tag{9.2}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and k_{ti}, y_i and ϕ_i are the transverse momentum, rapidity and azimuth of particle i respectively. R is a radius parameter while p governs the relative power of the energy versus geometrical (Δ_{ij}) scales. In the anti-k_t algorithm p is set to -1. The k_t algorithm [59] and the Cambridge/Aachen algorithm [60] are two

alternative jet reclustering algorithms. They define the distance parameter with $p = 1$ and $p = 0$ respectively.

The clustering begins by identifying if the smallest distance is between two particles or between a particle and the beam. If the smallest distance is between two particles i and j they are merged into a new particle with $E_{\mathrm{T}} = E_{\mathrm{T},i} + E_{\mathrm{T},j}$, $\eta = [E_{\mathrm{T},i}\eta_i + E_{\mathrm{T},j}\eta_j]/E_{\mathrm{T}}$ and $\phi = [E_{\mathrm{T},i}\phi_i + E_{\mathrm{T},j}\phi_j]/E_{\mathrm{T}}$. If the smallest distance is between a particle and the beam, the particle is called a jet and removed from the list of particles. The procedure is repeated until there are no particles remaining.

The anti-k_t algorithm is an infrared and collinear safe jet reconstruction algorithm. Collinear safety means the output of the algorithm remains the same if the energy of a particle is split among two collinear particles. The output of an infrared safe reconstruction algorithm remains stable with the addition of soft particles.

The behaviour of each of the jet algorithms in an event containing both hard and soft jets is illustrated in Figure 9.2. The k_t, Cambridge/Aachen and SISCone algorithms reconstruct varied shapes for the hard jets in green, blue and red. In contrast, the jets from hard partons are reconstructed as circular with the anti-k_t algorithm, while the soft jets have varied shapes.

9.5.1 Particle Flow Jets

All reconstructed particle flow particles are clustered into jets with the anti-k_t algorithm. However, for specific use cases particles may be removed from the jets. In this analysis, charged hadrons from pile-up interactions are subtracted from the jets. Each particle flow particle is assigned to a reconstructed primary vertex. Particles from pile-up interactions are identified by their association with a pile-up vertex. A muon is not clustered into a jet if it has a relative isolation less than 0.2 within a cone size of 0.4. An electron is not included in the reclustering if it has a relative isolation less than 0.2 within a cone size of 0.3. The relative isolations of the leptons are defined in Sections 10.2 and 10.3.

9.5.2 Jet Energy Scale Corrections

Translating the measured jet energy to the energy of the particle which produced the jet is complicated by the non-linear response of the calorimeters. Jet energy scale (JES) corrections are applied to map the measured jet energy deposition to the particle level.

JES corrections are factorised, with each level correcting for a different effect. The corrections are applied by scaling the jet four-momentum with a scale factor which depends on jet related quantities. Corrections are applied in four steps to jets in both simulation and data. An additional residual correction to the second and third steps is applied to jets in data. The methods for measuring the jet energy corrections are

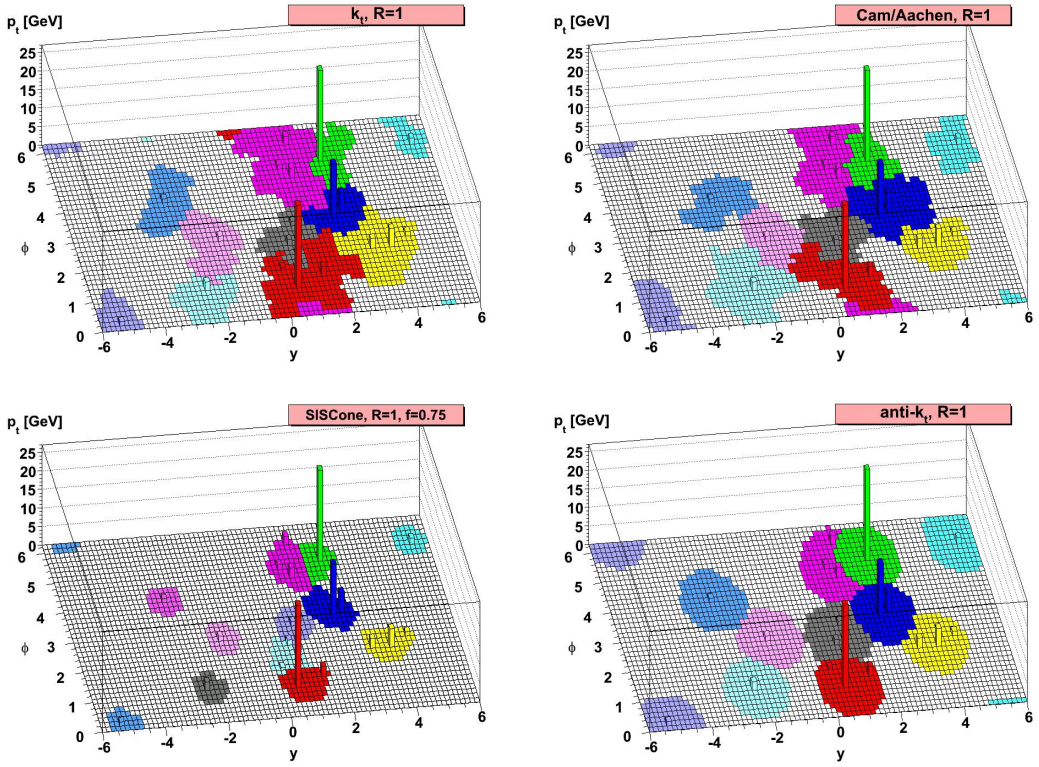


Figure 9.2: A parton-level event with many soft jets, clustered with four different jet reconstruction algorithms [58]. The SISCone algorithm [61] is based on the search for stable cones, with a split-merge step which disentangles overlapping cones.

described in [62]. The JES corrections are referred to as L1, L2 Relative, L3 Absolute and L2L3 Residual.

In the first JES step (L1) pile-up corrections remove energy due to neutral hadrons from pile-up interactions. This is done by means of a ρ correction, described in [63], where ρ is the mean amount of p_{T} per unit area that has been added to the event by pile-up. The relative corrections in the second step (L2 Relative) are derived to flatten the jet response with respect to η. To achieve this, jets in an arbitrary η region are corrected relative to jets in the central region of $|\eta| < 1.3$. The absolute corrections in the third step (L3 Absolute) are designed to make the jet response flat with respect to p_{T}. The residual corrections to the second and third steps (L2L3 Residual) are applied to data to correct for a small difference in jet energy response between data and simulation.

The overall JES uncertainty measured as a function of jet p_{T} with 36 pb^{-1} of data

is shown for particle flow jets in Figure 9.3. The JES uncertainties measured with 4.76 fb^{-1} of data are similar to or smaller than the uncertainties shown. The JES uncertainties for two alternative jet reconstruction algorithms are also shown. Particle flow jets have the smallest uncertainty for the jet p_{T} ranges which dominate in events selected in this analysis. The distributions of jet p_{T} for selected events are shown in Figures 12.16 and 12.17.

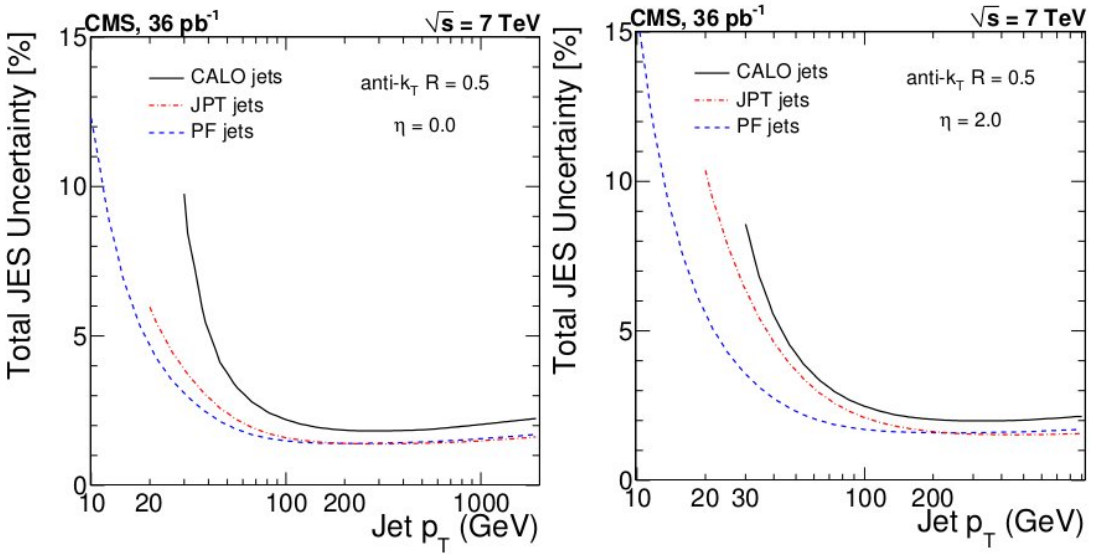


Figure 9.3: Overall JES uncertainty as a function of jet p_{T} for jets with $\eta = 0$, left, and jets with $\eta = 2$, right [62].

9.5.3 Jet Energy Resolution

The jet energy resolution (JER) is measured by examining the p_{T} imbalance in dijet and γ/Z + jets, as described in [62]. The JER in simulation is corrected by scaling the difference in the reconstructed jet p_{T} and the matched generator level jet p_{T} by the η dependent scale factors derived from the JER measurement. A generator level jet is reconstructed by using the anti-k_t algorithm to cluster the four-momenta of all stable particles generated in the simulation.

9.6 Beam Spot

The beam spot is the luminous region produced by the collision of the proton beams. The $d_0 - \phi_0$ algorithm [64] is a track based algorithm used to determine the beam spot

position with micron precision. d_0 is the signed impact parameter distance between the track and origin at minimun approach while ϕ_0 is the direction of the track at the point of minimum approach. The $d_0 - \phi_0$ algorithm is a simple iterative χ^2 fitter which uses the correlation between d_0 and ϕ_0 to extract the beam parameters.

Basic quality requirements are applied to tracks considered in the fit. The contribution from each track is weighted by its uncertainty. After the χ^2 is minimised, tracks with the largest contribution to the total $\chi^2$and tracks with the largest d_0 with respect to the beamline, which is a priori known, are removed. The complete set of initially selected tracks are re-evaluated at each iteration as the estimate of the beam position improves. With one thousand tracks a statistical precision of $2\,\mu$m is expected for the transverse beam position.

9.7 Primary Vertex

Primary vertex reconstruction is required for the precise determination of collision points in proton-proton collisions. Vertex reconstruction algorithms must be able to reconstruct multiple collisions occurring in one bunch-crossing, due to pile-up, and assign tracks to each collision. Vertex reconstruction uses information from the silicon tracker to find vertex candidates and then performs fitting to determine a best estimate of the vertex parameters for a given set of tracks.

Reconstructed tracks are selected based on their compatibility with the beam spot, the number of associated hits in the tracker, and the track fit quality. Tracks are then clustered into primary vertex candidates with the deterministic algorithm [65]. The tracks are clustered according to the z coordinate of the point of closest approach of a track to the z-axis.

A vertex fit is performed with tracks in each cluster using full track information and the adaptive vertex fitter algorithm [66]. This algorithms applies an iterative reweighted fit to down weight tracks according to their χ^2 distance from the vertex candidate. The weights vary per iteration until the fit converges. The sum of weights from the fit roughly corresponds to the effective number of tracks accepted by the adaptive fitter. The number of degrees of freedom is defined as $\mathrm{Ndof} = 2\sum w_i - 2$, where w_i is the weight of track i.

The primary vertex collection is sorted according to the sum of the p_{T}^2 of the tracks associated to each vertex. If no reconstructed vertex is found, a vertex based on the beam spot is put into the event. In this case no tracks are associated to the vertex, the χ^2 and Ndof are set to 0, and the vertex is flagged as fake.

The primary vertex resolution depends strongly on the number of tracks used by the fitter algorithm. The resolutions in x, y and z measured at $\sqrt{s} = 7$ TeV are less than $70\,\mu$m (x, y) and $80\,\mu$m (z) for primary vertices reconstructed with more than 10 tracks [67], which is the case for the majority of primary vertices in $t\bar{t}$ events.

Part IV

Selection

Chapter 10

Event Selection

Selection criteria are applied to reduce the amount of background events in the data sample. The same selection criteria are applied to the simulated samples to estimate how efficient the event selection is at retaining signal events while discarding background events.

The muon + jets event topology is described in Section 4.1. The muon + jets final state leads to four or more jets and one isolated muon in the detector. An example of a muon + jets-like event selected from the 2011 data sample is shown in Figure 10.1.

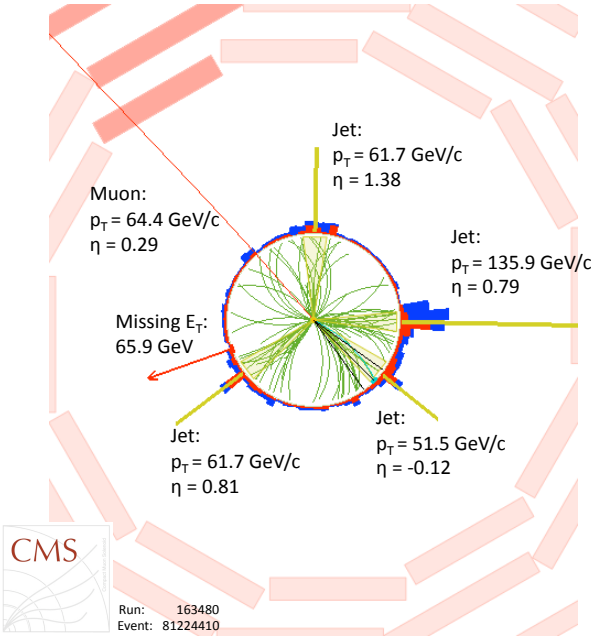


Figure 10.1: A muon + jets-like event detected by the CMS detector at $\sqrt{s} = 7$ TeV [68]. The view transverse to the beam axis is shown. The energy deposited in the ECAL and HCAL is represented in red and blue respectively. The DT muon chambers are shown at the edge of the image, with the chambers which detected the passage of the muon highlighted.

Four high p_{T} jets and one high p_{T} muon, all well separated from each other, are seen in the display. The event selection criteria applied in the $t\bar{t}$ cross section analysis is based on this event signature. Quality criteria are applied to the muon and jets to avoid misidentification.

10.1 Vertices

Selection criteria are applied to the first vertex in the reconstructed vertex collection, sorted according to the sum of the p_{T}^2 of the tracks associated to each vertex, to ensure the primary $t\bar{t}$ event vertex is well reconstructed. The requirements applied are summarised in Table 10.1. Requiring at least four degrees of freedom implies the selected vertices will have at least four tracks accepted by the adaptive vertex fitter, as explained in Section 9.7. The parameters $|z|$ and ρ represent the z-coordinate and radial coordinate of the primary vertex. Requiring the vertex to not be flagged as fake ensures the vertex was successfully reconstructed, rather than based on the beam spot.

Primary Vertex	
Parameter	Value
Ndof	≥ 4
$\lvert z\rvert$	< 24 cm
ρ	< 2 cm
is fake	`false`

Table 10.1: Selection criteria applied to reconstructed vertices to obtain a collection of good offline primary vertices.

10.2 Muons

Tight selection criteria are applied to reconstructed particle flow muons, from any of the particle flow muon subcategories, to obtain a collection of tight isolated muons. The selection is optimised to retain muons from W boson decay while rejecting muons from other sources. The sources of reconstructed muons are described in Section 4.2. Relaxed criteria are applied to the reconstructed particle flow muons to obtain a second collection of loose muons in the event.

10.2.1 Tight Isolated Muons

The requirements for a tight isolated muon are summarised in Table 10.2. The p_{T} value of 42 GeV is chosen to avoid the turn on of the HLT_Mu40 trigger efficiency, shown in Figure 11.8. This avoids efficiency dependencies due to incorrect modelling of the

efficiency turn on curve with respect to p_{T} in simulation. The muon $|\eta|$ is limited to 2.1 to correspond with the $|\eta|$ requirement in the HLT_Mu40_eta2p1 trigger used to collect data with the high luminosity trigger menus. The muon is required to be a global muon as the use of information from both the tracker and muon systems improves the estimates for the parameters of the reconstructed particle flow muon.

Tight Isolated Muon	
Parameter	Value
p_{T}	> 42 GeV
$\|\eta\|$	< 2.1
is global	`true`
Identification Criteria (ID)	
χ^2/Ndof	< 10
N muon hits	≥ 1
N matched stations	≥ 2
$\mathrm{d_{xy}}$	< 0.2 cm
N pixel hits	≥ 1
N tracker layers	≥ 9
$\Delta\mathrm{R}(\mu,\mathrm{jet})$	> 0.3
$\mathrm{Iso_{rel.}}(\mu)$	< 0.15

Table 10.2: Selection criteria applied to reconstructed particle flow muons to obtain a collection of tight isolated muons.

Quality criteria are applied to reject non-collision muons based on studies performed in [40] and [69]. The requirement on the normalised χ^2 of the global muon track fit and the number of muon chamber hits included in the fit suppresses hadronic punch-through and muons from decays in flight, described in Section 4.2. The majority of muons in the reconstructed particle flow muon collection have a normalised χ^2 less than five. The number of hits in the muon chambers ranges from 0 to 50 with the distribution peaking at 20.

The muon must be detected by segments in at least two muon stations to be consistent with the muon trigger requirements. This suppresses hadronic punch-through and accidental matches between tracks and muon segments. Reconstructed particle flow muons are detected by zero to five muon stations, with the distribution peaking at three.

Cosmic muons passing through the detector, and muons from decays in flight, are unlikely to have tracks close to the primary vertex. These are removed by the requirement on the transverse impact parameter, $\mathrm{d_{xy}}$, of the muon tracker track with respect to the primary vertex. The majority of reconstructed particle flow muons have a $\mathrm{d_{xy}}$ of

less than 0.02 cm.

Requiring hits in the pixel tracker and at least nine tracker layers provides further rejection for muons from decays in flight. The large number of tracker layers with hits guarantees an accurate muon p_{T} measurement. Reconstructed particle flow muons have from 0 to 4 hits in the pixel tracker and from 0 to 18 tracker layers with hits. The distribution for the number of pixel tracker hits peaks at 3 while the distribution for the number of tracker layers with hits peaks at 13.

The distance between the selected muon and the closest jet is given by $\Delta\mathrm{R}(\mu,\mathrm{jet})$. The $\Delta\mathrm{R}(\mu,\mathrm{jet})$ requirement is included for the data driven multijet estimation described in Section 12.2. The motivation for the $\Delta\mathrm{R}$ requirement is explained further in Section 12.2.1. The vast majority of muons which pass the other selection requirements have $\Delta\mathrm{R}(\mu,\mathrm{jet}) > 0.3$.

The isolation requirement is primarily intended to reject muons from within jets, the main source of muons in multijet events. The distribution of relative isolation for reconstructed particle flow muons peaks strongly close to zero, where muons from W or Z boson decay are expected to dominate. However, muons may have relative isolation values up to five or larger.

The relative isolation of a muon is defined by Equation 10.1 within a cone of radius 0.4 around the muon. The cone radius of 0.4 is chosen to improve the separation between muons from W or Z boson decay and muons from within jets.

$$\mathrm{Iso}_{\mathrm{rel.}}(\mu) = \frac{\sum p_{\mathrm{T}}(\mathrm{CH}) + \max(0., \sum p_{\mathrm{T}}(\mathrm{NH}) + \sum p_{\mathrm{T}}(\mathrm{Ph}) - 0.5 \cdot \sum p_{\mathrm{T}}(\mathrm{CP}))}{p_{\mathrm{T}}^{\mu}} \quad (10.1)$$

CH, NH and Ph denote charged hadrons, neutral hadrons and photons respectively. CP denotes charged particles within the cone of interest, where the particles are assigned to a pile-up vertex. The 0.5 factor corresponds to an estimate of the average neutral to charged particles ratio, measured in [70]. As such, $0.5 \cdot \sum p_{\mathrm{T}}(\mathrm{CP})$ is an estimate of the pile-up contribution from neutral particles after the charged particle pile-up contribution has been subtracted. The subtraction of this contribution reduces the pile-up dependence of the relative isolation of the muon.

10.2.2 Loose Muons

The p_{T}, η and relative isolation requirements are relaxed for the loose muon collection to allow for identification of secondary muons from Z boson decay and di-muon $t\bar{t}$ events.

Loose Muon	
Parameter	Value
p_{T}	> 15 GeV
$\lvert\eta\rvert$	< 2.5
is global	`true`
$\mathrm{Iso}_{\mathrm{rel.}}(\mu)$	< 0.2

Table 10.3: Selection criteria applied to reconstructed particle flow muons to obtain a collection of loose muons.

The loose muon selection criteria are summarised in Table 10.3. The muon $|\eta|$ is limited to 2.5 due to the acceptance of the pixel tracker. The muon is required to be a global muon to ensure the p_{T} estimate is accurate. The p_{T} and relative isolation requirements are chosen to increase the selection efficiency for prompt muons from W or Z boson decay while rejecting muons from within jets.

10.3 Electrons

A set of loose selection criteria, Table 10.4, are applied to reconstructed particle flow electrons to identify isolated electrons in the event. The $|\eta|$ requirement is based on the pixel tracker acceptance. The p_{T} and relative isolation requirements are chosen to increase the selection efficiency for prompt electrons from W or Z boson decay while rejecting electrons from within jets.

Loose Electron	
Parameter	Value
p_{T}	> 15 GeV
$\lvert\eta\rvert$	< 2.5
$\mathrm{Iso}_{\mathrm{rel.}}(\mathrm{e})$	< 0.2

Table 10.4: Selection criteria applied to reconstructed particle flow electrons to obtain a collection of loose electrons.

The relative isolation of an electron is defined by Equation 10.2 within a cone of radius 0.3 around the electron. The cone radius of 0.3 is chosen to improve the separation between electrons from W or Z boson decay and electrons from within jets.

$$\mathrm{Iso}_{\mathrm{rel.}}(\mathrm{e}) = \frac{\sum p_{\mathrm{T}}(\mathrm{CH}) + \max(0., \sum p_{\mathrm{T}}(\mathrm{NH}) + \sum p_{\mathrm{T}}(\mathrm{Ph}) - \rho \cdot \mathrm{EA})}{p_{\mathrm{T}}^{e}} \tag{10.2}$$

EA is the effective area of the isolation cone, defined as the ratio between the slope of the average isolation and ρ as a function of the number of primary vertices. The $\rho \cdot \mathrm{EA}$ correction is an estimate of the neutral hadron contribution from pile-up interactions and is subtracted to reduce pile-up dependence in the calculation of relative isolation. The optimisation of cone radius and pile-up subtraction studies were performed separately for electrons and muons, resulting in different cone radii and pile-up subtraction methods.

10.4 Jets

The requirements for a reconstructed particle flow jet to be considered in the analysis are summarised in Table 10.5. Jets arising from the $t\bar{t}$ hard scattering process tend to have larger p_{T} than jets from initial or final state radiation. The η requirement is based on the coverage of the strip tracker.

Jet	
Parameter	Value
p_{T}	> 30 GeV
$\lvert\eta\rvert$	< 2.4
Quality Criteria	
N constituents	> 1
Charged hadron fraction	> 0
Neutral hadron fraction	< 0.99
Charged EM fraction	< 0.99
Neutral EM fraction	< 0.99
N charged hadrons	> 0

Table 10.5: Selection criteria applied to reconstructed particle flow jets to obtain a collection of tight jets.

The fraction of electromagnetic energy deposited by charged constituents is required to be less than 0.99 to reject electrons. The fraction of energy from neutral hadronic particles and the fraction of electromagnetic energy deposited by neutral constituents is expected to be greater than 0.99 for jets reconstructed due to HCAL or ECAL noise respectively. Jets reconstructed from noise are also removed by requiring at least one charged hadron in the jet reconstruction, contributing to some fraction of the jet energy. Finally, selected jets are required to have at least two constituents, which is expected for real jets due to hadronisation and avoids the misidentification of muons or electrons as jets.

10.5 Event Selection

Data events considered for analysis are collected by the HLT_Mu40 trigger. Generated events are also required to pass a simulation of the trigger response. Exactly one tight isolated muon is required. The quality criteria are designed to remove non-collision muons and the isolation criteria are intended to reject multijet events. Events with an additional muon, from the tight or loose collections, are vetoed to reject Z + jets and $t\bar{t}$ di-muon events. Events with a loose electron are vetoed to reject $t\bar{t}$ dilepton events. The four jet requirement is the primary factor in the removal of W + jets events, which have the same lepton signature as the signal events. The event selection is summarised in Table 10.6.

Event Selection	
Object	Multiplicity
Primary Vertex	≥ 1
Tight Isolated Muon	1
Loose Muon	0
Loose Electron	0
Jets	≥ 4
Trigger	HLT_Mu40

Table 10.6: Number of objects from each of the object collections required to select a muon + jets-like event. Only events passing the HLT_Mu40 trigger are considered.

The result of the event selection applied to simulation and data is shown in Table 10.7. The contributions from simulation are normalised to the expected event yields for 4.76 fb^{-1} of data with the respective production cross sections in Table 4.1. Data driven corrections to the selection efficiency, derived in Chapter 11, are applied.

The efficiency to select $t\bar{t}$ events in any decay channel, $\varepsilon_{t\bar{t},\,\text{total}} = (N_{t\bar{t},\,\mu} + N_{t\bar{t},\,\text{other}})/(N_{t\bar{t},\,\mu\,\text{produced}} + N_{t\bar{t},\,\text{other produced}})$. The selection efficiencies for the muon + jets and other decay channels are derived separately then combined according to the branching fractions in Table 3.1. $\varepsilon_{t\bar{t},\,\text{total}}$ is used in the extraction of the $t\bar{t}$ cross section with Equation 1.1. The separate decay channels are combined due to the difficulty in distinguishing between kinematic distributions in different $t\bar{t}$ decay channels, which is relevant for the method used to extract the $t\bar{t}$ cross section in Chapter 12.

The efficiency to select $t\bar{t}$ muon + jets events, $\varepsilon_{t\bar{t},\,\mu} = N_{t\bar{t},\,\mu}/N_{t\bar{t},\,\mu\,\text{produced}}$, and the purity, $\pi = N_{t\bar{t},\,\mu}/(N_{t\bar{t},\,\mu} + N_{\text{background}})$, of the selection are indications of the effectiveness of the event selection in retaining signal events while rejecting background. In the purity calculation $N_{\text{background}}$ is the sum of all background events, including the contribution from $t\bar{t}$ events which do not result in the muon + jets final state.

Sample	N produced	N Selected
$t\bar{t}$ (muon + jets)	115357	17603 ± 14
$t\bar{t}$ (other)	670109	2188 ± 5
W + jets	149067197	12522 ± 101
Z + jets	14509702	1606 ± 17
multijet	403107424	328 ± 53
t, s-channel	21898	28 ± 1
t, t-channel	307522	258 ± 3
t, tW-channel	74738	880 ± 4
Data		33565 ± 183
$\sum$ Simulation		35413 ± 116
$\varepsilon_{t\bar{t},\,\mu}$		0.1526 ± 0.0001
π		0.49708 ± 0.00002
$\boldsymbol{\varepsilon_{t\bar{t},\,\text{total}}}$		$\mathbf{0.02520 \pm 0.00002}$

Table 10.7: Number of events selected when the event selection criteria are applied to data and to the simulated samples. N produced signifies the number of events expected for 4.76 fb^{-1} of data according to the cross sections calculated in Table 4.1. The selection efficiency for $t\bar{t}$ muon + jets events is denoted by $\varepsilon_{t\bar{t},\,\mu} = N_{t\bar{t},\,\mu}/N_{t\bar{t},\,\mu\,\text{produced}}$. The efficiency to select $t\bar{t}$ events in any channel is denoted by $\varepsilon_{t\bar{t},\,\text{total}} = (N_{t\bar{t},\,\mu} + N_{t\bar{t},\,\text{other}})/(N_{t\bar{t},\,\mu\,\text{produced}} + N_{t\bar{t},\,\text{other produced}})$. The purity, π, of the event selection is given by $N_{t\bar{t},\,\mu}/(N_{t\bar{t},\,\mu} + N_{\text{background}})$.

Chapter 11

Muon Efficiency

An estimate of the efficiency to select $t\bar{t}$ events is required to measure the $t\bar{t}$ cross section with Equation 1.1. In principle, the selection efficiency can be derived from the simulated $t\bar{t}$ sample. However, an imperfect description of the detector in the simulation can cause the selection efficiency in data to diverge from the simulated selection efficiency. To account for this, a correction factor is derived by measuring the selection efficiency in both data and simulation. The scale factor, defined in Equation 11.1, is then applied to the simulated muon selection efficiency to reproduce the efficiency measured in data.

$$\text{scale factor} = \frac{\varepsilon_{\text{Data}}}{\varepsilon_{\text{Simulation}}} \tag{11.1}$$

The scale factor is measured individually for both the muon identification and the muon trigger efficiency. No scale factor is applied to the simulated efficiency for the remaining event selection. Instead, sources of systematic uncertainty, such as JES, are considered explicitly as described in Chapter 13.

11.1 Tag and Probe Method

Muon selection efficiencies are measured directly in data using the tag and probe method. The tag and probe method is based on the selection of $Z \to \mu\mu$ events. The tag muon is selected with tight identification criteria while a second muon, the probe, is selected with loose selection criteria. The muons are required to have oppositely signed charge and a di-muon mass in the region of 81.2 to 101.2 GeV, bracketing the Z boson mass of 91.2 GeV [1]. This selection leads to a high purity sample of $Z \to \mu\mu$ events, as is apparent from the agreement between the reconstructed mass distributions in data and Z + jets simulation in Figure 11.1.

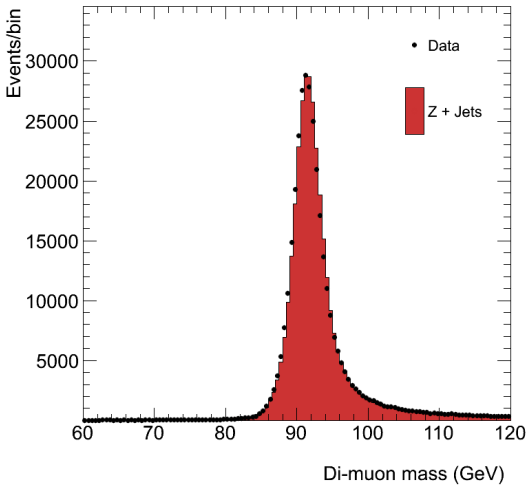


Figure 11.1: Di-muon mass reconstructed from two oppositely signed muons after the tag and probe selection. The number of Z + jets events is normalised to the number of data events. The distribution peaks at the nominal Z boson mass of 91.2 GeV. The similarity between the shape of the distribution in both simulation and data is an indication of the purity of the data sample after the tag and probe selection.

Since muons from Z boson decay are similar to muons from top decay, a fully efficient selection would result in all probe muons passing further selection. Therefore, the efficiency for specific muon selection criteria is measured by counting how many probe muons pass the selection using the equation

$$\varepsilon = \frac{\mathrm{N}_{\text{probes, passing}}}{\mathrm{N}_{\text{probes}}} = \frac{\mathrm{N_{TS}}}{\mathrm{N_{TS}} + \mathrm{N_{TP}}} \tag{11.2}$$

where $\mathrm{N_{TS}}$ is the number of events where one muon is tagged and the second passes the selection criteria being examined. The number of events where one muon is tagged and the second is selected as a probe but fails further selection is represented by $\mathrm{N_{TP}}$. Both of these quantities depend on the efficiency of the selection being examined according to the relation

$$\mathrm{N_{TS}} = \mathrm{N_{Data}} \cdot \varepsilon_{\mathrm{T}} \cdot \varepsilon_{\mathrm{S}}, \tag{11.3}$$

$$\mathrm{N_{TP}} = \mathrm{N_{Data}} \cdot \varepsilon_{\mathrm{T}} \cdot (1 - \varepsilon_{\mathrm{S}}), \tag{11.4}$$

where N_{Data} is the number of data events, ε_T is the efficiency of the tag selection and ε_S is the selection efficiency to be measured. By replacing the elements of Equation 11.2 with Equations 11.3 and 11.4 it can be shown that the selection efficiency is returned.

The tag muon is required to pass the tight isolated muon selection criteria in Table 10.2. In addition, the tag muon is required to match a trigger object passing the HLT_Mu40 trigger. A match is counted if a L3 trigger object which passed the HLT_Mu40 requirements is within $\Delta R < 0.2$ of the tag muon. This matching is necessary for the muon trigger efficiency measurement to avoid events where the trigger may have been prescaled, which would result in an artificially low measured trigger efficiency.

The probe muon definition depends on the selection criteria being examined. For the muon identification (ID) efficiency measurement the probe muon is identified as a global muon with $p_T > 42$ GeV and $\eta < 2.1$. The efficiency of this selection is accurately reproduced in simulation [40]. The probe muon is counted as a passing probe in Equation 11.2 if it passes the remaining ID selection criteria in Table 10.2.

The probe muon for the trigger efficiency measurement is required to fulfill the selection criteria in Table 10.2. A passing probe is counted in Equation 11.2 if the probe muon matches a HLT_Mu40 trigger object. As with the tag muon matching, a match is counted if a L3 trigger object which passed the HLT_Mu40 requirements is within $\Delta R < 0.2$ of the probe muon.

The overall muon selection efficiency is given by

$$\varepsilon_\mu = \varepsilon_{reco} \cdot \varepsilon_{ID/reco} \cdot \varepsilon_{trigger/ID}. \tag{11.5}$$

The efficiency to reconstruct a global particle flow muon with $p_T > 42$ GeV and $\eta < 2.1$, ε_{reco}, is provided by simulation. The efficiency for a reconstructed muon to pass the ID criteria, $\varepsilon_{ID/reco}$, and the efficiency for a muon passing the ID criteria to pass the HLT_Mu40 trigger, $\varepsilon_{trigger/ID}$, are measured with the tag and probe method.

11.2 Muon ID Efficiencies and Scale Factors

The muon ID efficiency is measured with respect to a number of kinematic quantities to examine possible dependencies. The muon ID scale factor is calculated from the measured efficiencies using Equation 11.1.

Figure 11.2 and Table 11.1 show the muon ID efficiency and scale factor in regions of η. The scale factor varies with respect to the η of the muon, therefore the η dependent scale factor is applied to the simulated ID efficiency before further studies. After reweighting the simulated ID efficiency, the efficiency measurement is repeated with respect to other variables to uncover further dependencies.

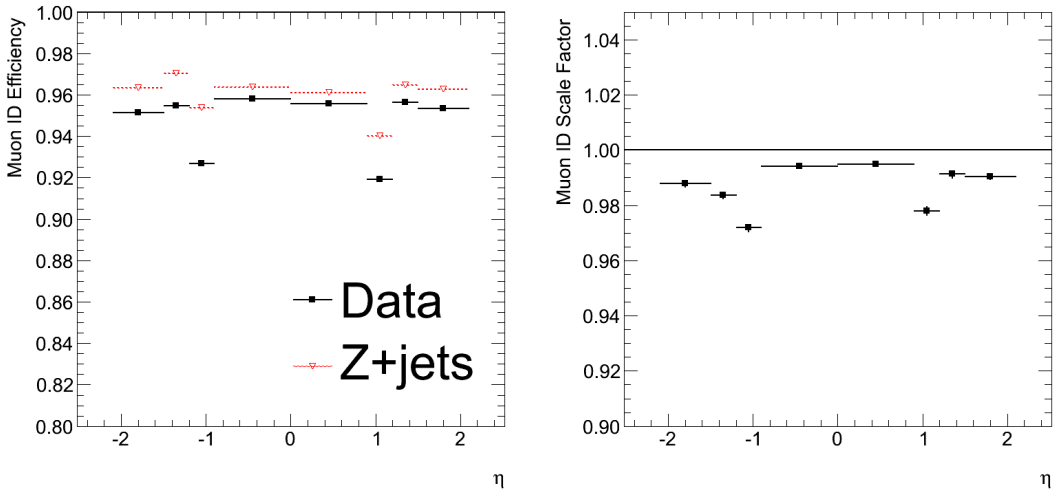


Figure 11.2: Muon ID efficiency and scale factor measured in regions of η. The muon η distribution after the $t\bar{t}$ signal selection is shown in Figure 12.15.

The efficiency and scale factor are presented as a function of p_{T} in Figures 11.3 and 11.4. The scale factor does not have an apparent dependence on the p_{T} of the muon above the muon p_{T} requirement of 42 GeV, varying within ± 0.005.

The $Z \to \mu\mu$ events used in the efficiency measurement tend to have low jet multiplicities whereas the signal selection requires at least four jets. Therefore, it is important to quantify any scale factor dependence on the jet multiplicity. Figure 11.5 shows the muon ID efficiency measured with respect to the jet multiplicity and number of primary vertices. Figure 11.6 uncovers a possible scale factor dependence on the jet multiplicity. The scale factors measured with respect to the jet multiplicity agree within ± 0.01, however it is not clear if this is a systematic effect or a statistical effect. In this analysis it is treated as a systematic effect.

The number of primary vertices reconstructed in an event corresponds to the amount of pile-up in the event. There is a downward trend in the scale factor with increasing number of primary vertices in Figure 11.6. This is because a steeper efficiency dependence is observed in data than expected in simulation, shown in Figure 11.5. However, considering the statistical uncertainties, the scale factors agree within ± 0.01.

The scale factors applied to the simulated ID efficiency are listed in Table 11.1, with the statistical uncertainty on the scale factor measured in each η region. An overall systematic uncertainty of ± 0.015 on the muon ID scale factor is obtained by adding in quadrature the systematic uncertainties on the scale factor measured with respect to the considered event and muon properties.

η Region	Scale Factor
-2.1 to -1.5	$0.9878^{+0.0006}_{-0.001}$
-1.5 to -1.2	$0.9837^{+0.0006}_{-0.001}$
-1.2 to -0.9	$0.9718^{+0.0007}_{-0.001}$
-0.9 to 0.0	$0.9942^{+0.0004}_{-0.0007}$
0.0 to 0.9	$0.9947^{+0.0004}_{-0.0007}$
0.9 to 1.2	$0.978^{+0.0008}_{-0.002}$
1.2 to 1.5	$0.9912^{+0.0007}_{-0.001}$
1.5 to 2.1	$0.9904^{+0.0006}_{-0.001}$

Table 11.1: Muon ID scale factor in regions of muon η. Uncertainties are statistical only.

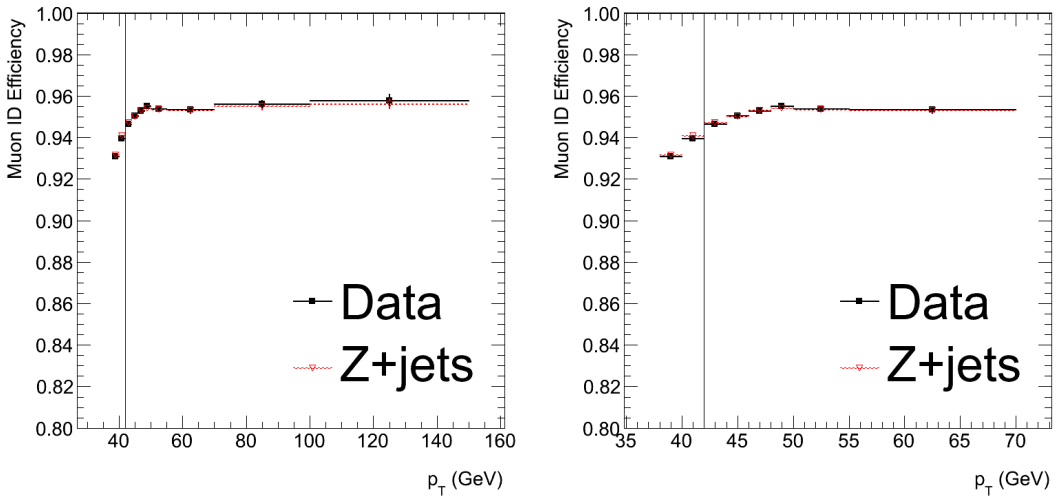


Figure 11.3: Muon ID efficiency measured in regions of p_{T}. The η dependent muon ID scale factor has been applied to the simulated efficiency. Muons selected in the cross section analysis are required to have a $p_{\mathrm{T}} > 42$ GeV.

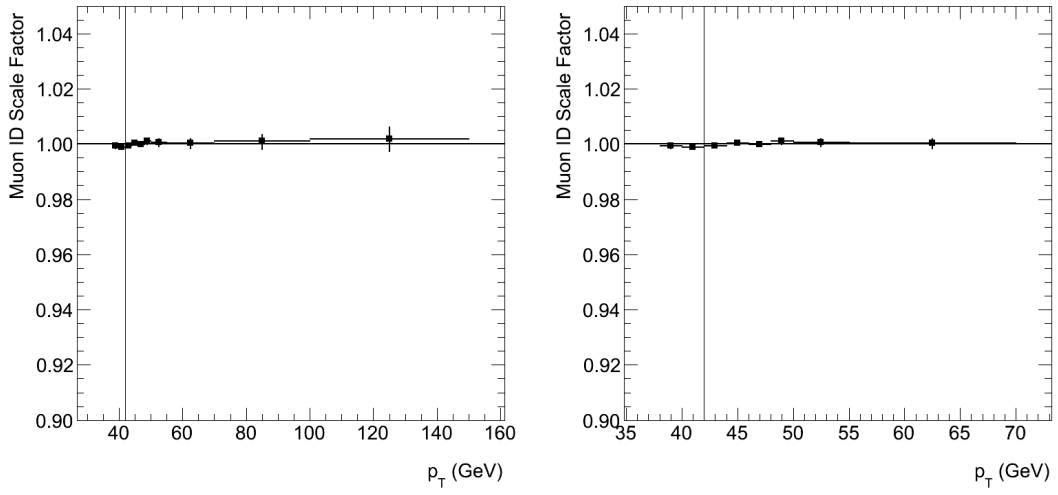


Figure 11.4: Muon ID scale factor measured in regions of p_{T}. Muons selected in the cross section analysis are required to have a $p_{\mathrm{T}} > 42$ GeV. The η dependent scale factor has been applied. The muon p_{T} distribution after the $t\bar{t}$ signal selection is shown in Figure 12.13.

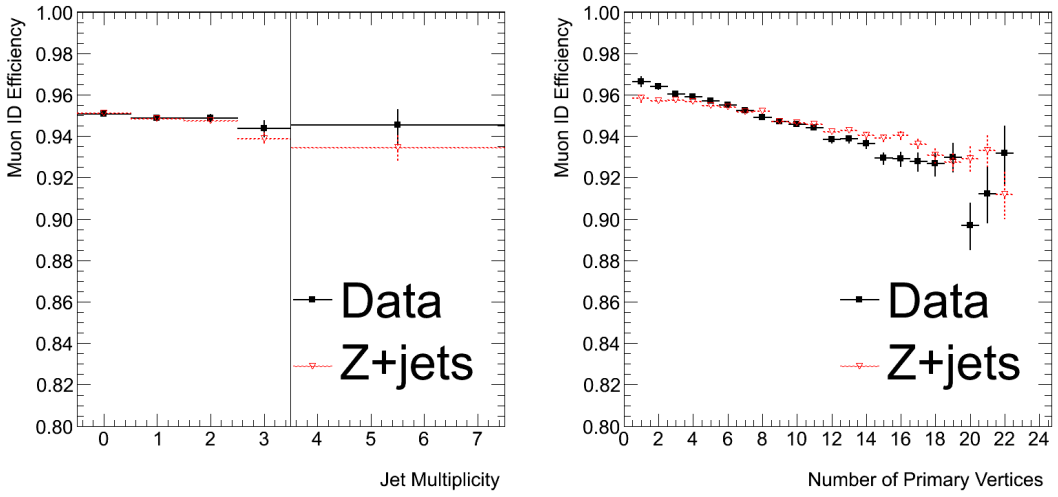


Figure 11.5: Muon ID efficiency measured with respect to jet multiplicity and number of primary vertices. The η dependent muon ID scale factor has been applied to the simulated efficiency. The $t\bar{t}$ signal selection requires at least four jets in selected events.

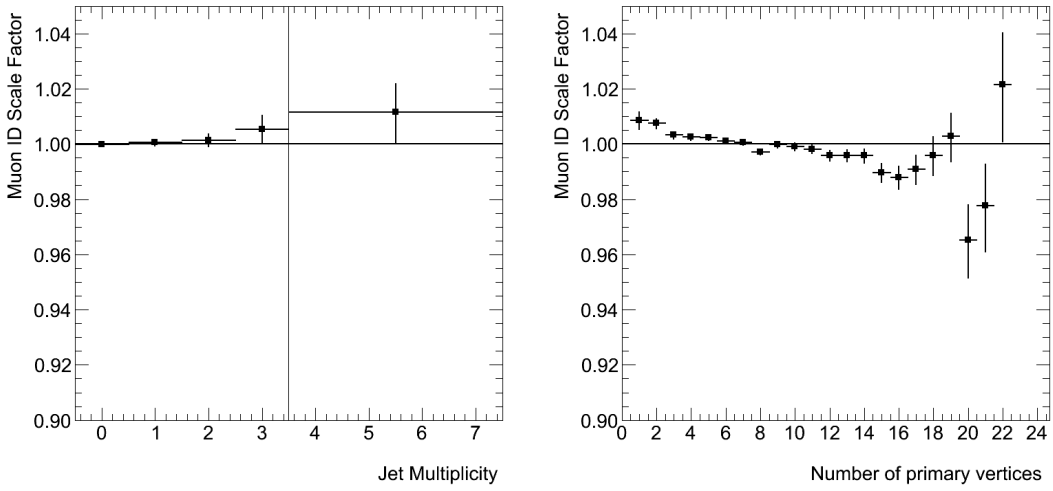


Figure 11.6: Muon ID scale factor measured with respect to jet multiplicity and number of primary vertices. The η dependent scale factor has been applied. The jet and primary vertex multiplicity distributions after the $t\bar{t}$ signal selection are shown in Figures 12.20 and 13.9.

11.3 Muon Trigger Efficiencies and Scale Factors

The efficiency of the HLT_Mu40 trigger in regions of η is shown in Figure 11.7 with the corresponding scale factors. The simulation overestimates the efficiency for muons with $|\eta|$ greater than 1.2 and underestimates the efficiency to trigger muons in regions of $|\eta|$ less than 1.2. This is due to η dependent adjustments to the HLT_Mu40 trigger in the different trigger menus used for data taking, described in Section 7.4. The trigger menu adjustments which affected the efficiency with respect to η are discussed in detail in Section 11.3.2.

The efficiency differences result in a significant scale factor dependence on the η of the muon. The η dependent efficiency observed in data is reproduced by applying an η dependent scale factor, given in Table 11.2, to the simulated trigger efficiency.

The turn on of the trigger efficiency with respect to muon p_{T} is shown in Figure 11.8, after the η dependent scale factor has been applied. The corresponding scale factors are shown in Figure 11.9. The efficiency reaches a plateau above 42 GeV while the scale factor varies within ± 0.005 above this threshold.

The trigger efficiency and scale factor for different jet multiplicities and number of primary vertices are shown in Figures 11.10 and 11.11. The efficiency and scale factor measured with respect to relative isolation and $\Delta\mathrm{R}(\mu,\mathrm{jet})$ are shown in Figures

11.12 and 11.13. In all four cases the scale factor varies within ±0.005, displaying no significant dependence.

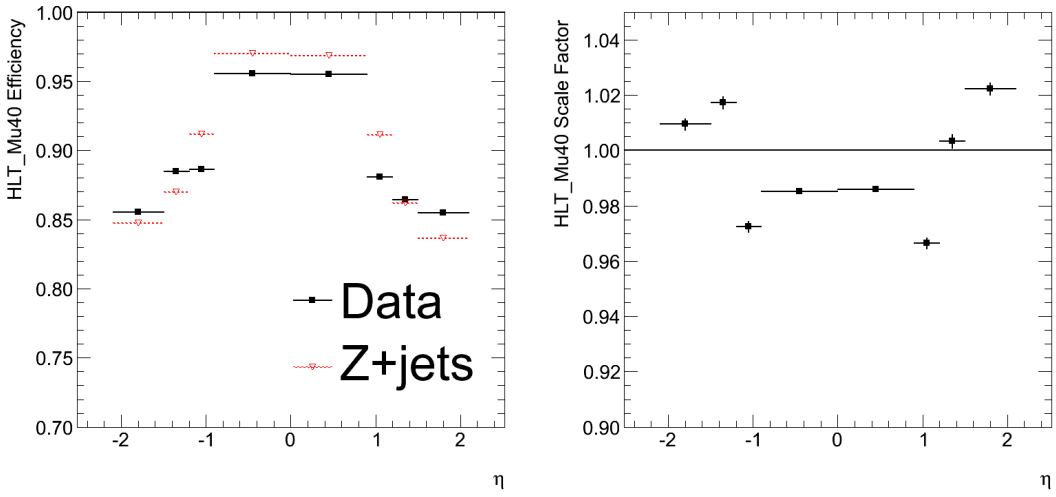


Figure 11.7: HLT_Mu40 efficiency and scale factor measured in regions of η. The muon η distribution after the $t\bar{t}$ signal selection is shown in Figures 12.15.

The scale factor correction applied to the simulated trigger efficiency is given in Table 11.2. The overall systematic uncertainty on the scale factor is taken as ±0.01 to account for the possibility of dependencies on the considered event and muon properties by adding the respective maximum scale factor variations in quadrature.

η Region	Scale Factor
-2.1 to -1.5	$1.0095^{+0.001}_{-0.002}$
-1.5 to -1.2	$1.0172^{+0.001}_{-0.002}$
-1.2 to -0.9	$0.9724^{+0.001}_{-0.002}$
-0.9 to 0.0	$0.9851^{+0.0003}_{-0.0007}$
0.0 to 0.9	$0.9858^{+0.0004}_{-0.0007}$
0.9 to 1.2	$0.9665^{+0.001}_{-0.002}$
1.2 to 1.5	$1.0033^{+0.001}_{-0.003}$
1.5 to 2.1	$1.0223^{+0.001}_{-0.002}$

Table 11.2: Muon trigger scale factors in regions of muon η. Uncertainties are statistical only.

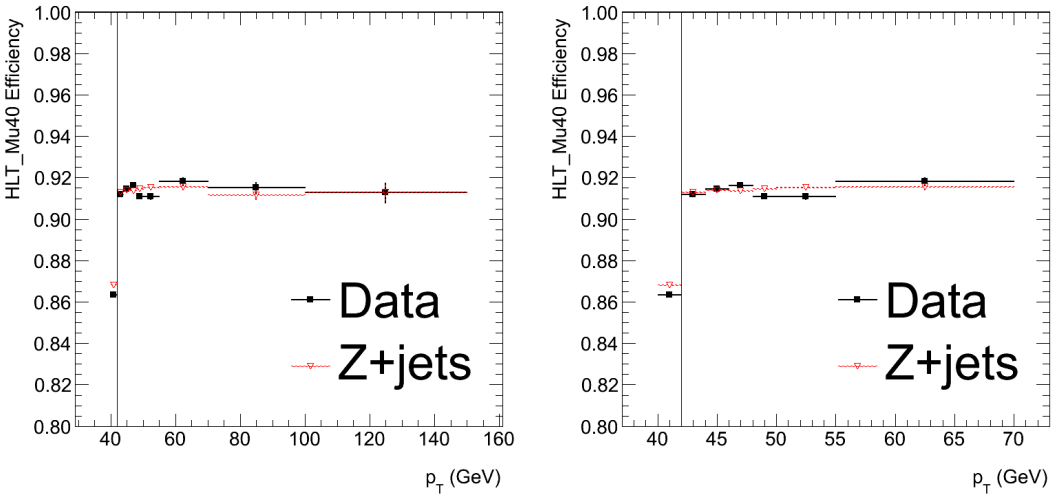


Figure 11.8: HLT_Mu40 efficiency measured in regions of p_{T}. Muons selected in the cross section analysis are required to have a $p_{\mathrm{T}} > 42$ GeV. The η dependent scale factor has been applied.

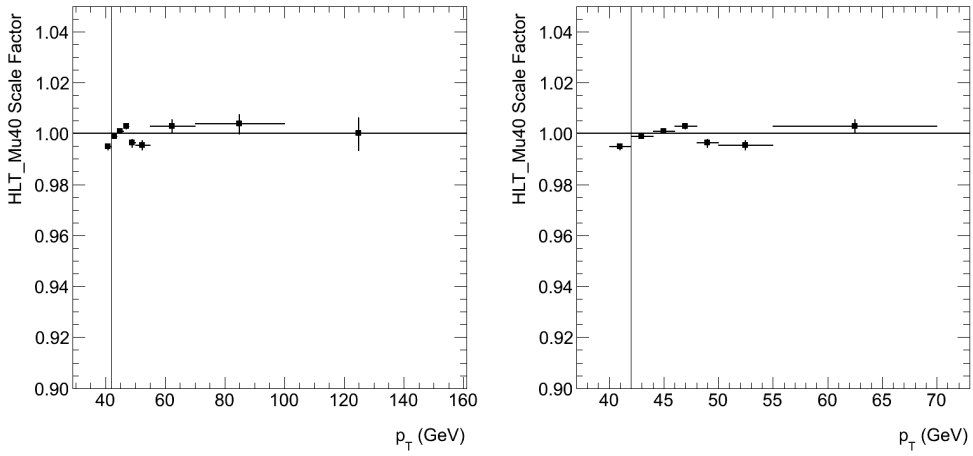


Figure 11.9: HLT_Mu40 scale factor measured in regions of p_{T}. Muons selected in the cross section analysis are required to have a $p_{\mathrm{T}} > 42$ GeV. The plot on the right shows a smaller range of muon p_{T} to examine the scale factor behaviour close to the trigger p_{T} threshold. The η dependent scale factor has been applied. The muon p_{T} distribution after the $t\bar{t}$ signal selection is shown in Figure 12.13.

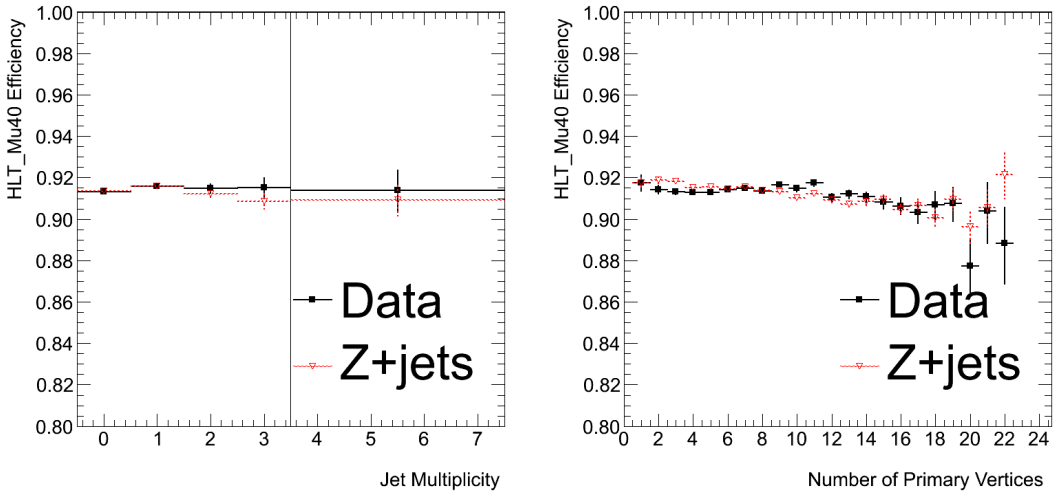


Figure 11.10: HLT_Mu40 efficiency measured with respect to jet multiplicity and number of primary vertices. The η dependent scale factor has been applied. The $t\bar{t}$ signal selection requires at least four jets in selected events.

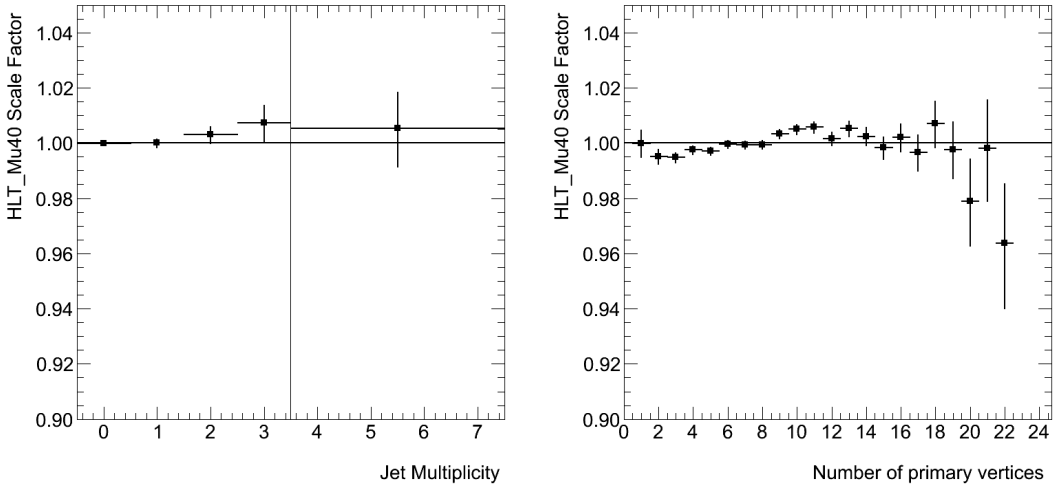


Figure 11.11: HLT_Mu40 scale factor measured with respect to jet multiplicity and number of primary vertices. The η dependent scale factor has been applied. The jet and primary vertex multiplicity distributions after the $t\bar{t}$ signal selection are shown in Figures 12.20 and 13.9.

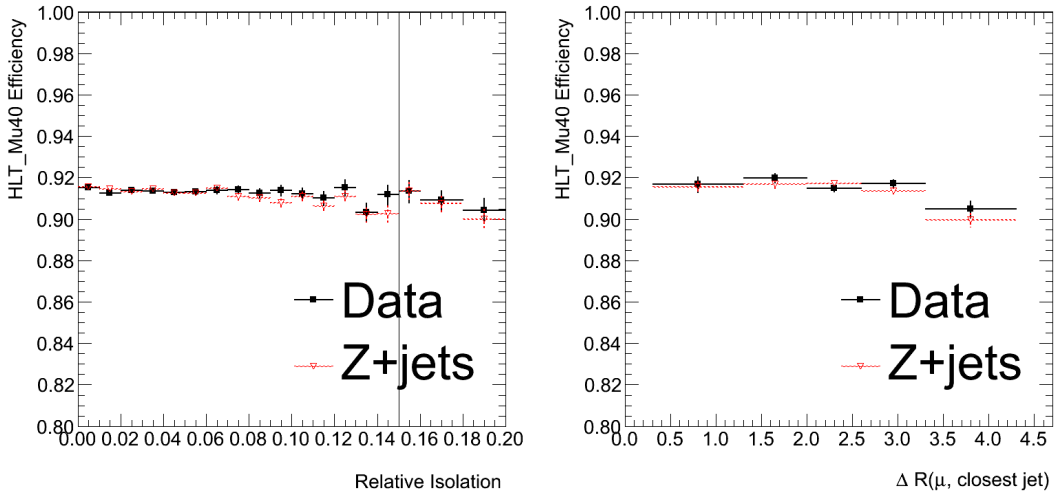


Figure 11.12: HLT_Mu40 efficiency measured with respect to relative isolation and ΔR(μ,jet). The η dependent scale factor has been applied to the simulated efficiency.

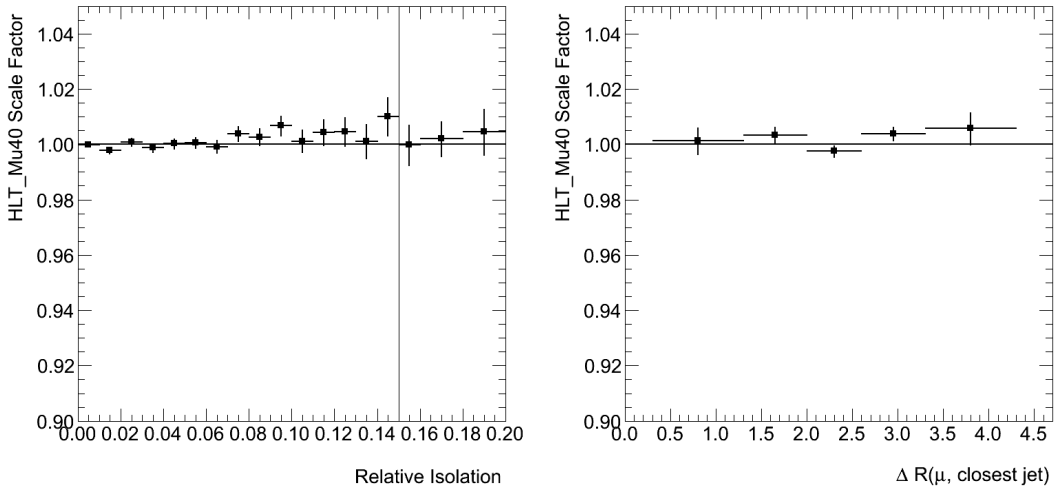


Figure 11.13: HLT_Mu40 scale factor measured with respect to relative isolation and ΔR(μ,jet). The η dependent scale factor has been applied. The relative isolation and ΔR(μ,jet) distributions after the $t\bar{t}$ signal selection are shown in Figures 12.4 and 12.20.

11.3.1 L1 and HLT Efficiencies

The efficiency of the HLT_Mu40 trigger is a combination of the efficiency of the L1 trigger and the HLT given by

$$\varepsilon_{\text{HLT\_Mu40}} = \varepsilon_{\text{L1}} \cdot \varepsilon_{\text{HLT}}. \tag{11.6}$$

The contribution from the L1 trigger can be measured separately by matching L1 trigger objects passing the L1 trigger to the probe muon. A match is counted if the L1 trigger object is within $\Delta\text{R} < 0.3$ of the probe muon. The L1 matching requirement is more relaxed than the L3 matching requirement as the L1 trigger object reconstruction is not as accurate as the L3 reconstruction. The efficiency of the HLT is extracted by measuring the HLT_Mu40 efficiency for probe muons which have passed the L1 trigger, $\varepsilon_{\text{HLT}} = \varepsilon_{\text{HLT\_Mu40}}/\varepsilon_{\text{L1}}$.

The L1 and HLT contributions to the HLT_Mu40 efficiency with respect to η are shown in Figure 11.14. The HLT_Mu40 efficiency differences between simulation and data are dominated by the L1 efficiency difference for $|\eta| > 1.5$, whereas the discrepancies in $|\eta| < 1.5$ region are mainly seen in the HLT efficiency. The discrepancies are due to the fact that the data was collected with a number of trigger menus while only one trigger menu was used in simulation. The trigger menu changes which affect the efficiency are discussed further in Section 11.3.2.

The L1 and HLT efficiencies in regions of muon p_{T} are shown in Figure 11.15. The HLT efficiency shows a sharp turn on at the trigger p_{T} threshold of 40 GeV, reaching an efficiency plateau at 42 GeV. The requirement of muon p_{T} greater than 42 GeV in the analysis is chosen to remain in the efficiency plateau.

The L1 and HLT efficiencies with respect to jet multiplicity and number of primary vertices are shown in Figure 11.16. The efficiencies show no significant dependence on the number of jets in an event. A flat efficiency response with respect to the number of primary vertices is measured in simulation for both the L1 and HLT. However, the efficiencies measured in data appear to show a dependence, with the L1 efficiency increasing as a function of the primary vertex multiplicity and the HLT efficiency decreasing. This is explained by the fact that the number of primary vertices in an event increased for data taken later in 2011. The trigger menu used to collect data also changed, resulting in an increase in the overall L1 efficiency and a decrease in the overall HLT efficiency. The low primary vertex multiplicities are dominated by events taken with early trigger menus while the high multiplicities are dominated by events taken with later trigger menus. Therefore the change in the overall trigger efficiency is reflected in the efficiencies measured with respect to the number of primary vertices. The L1 efficiency shows a stable response with respect the number of primary vertices when measured separately for each trigger menu, as shown in Figures 11.20. The HLT efficiency measured for each

trigger menu shows a reduced dependence on the number of primary vertices in Figure 11.21.

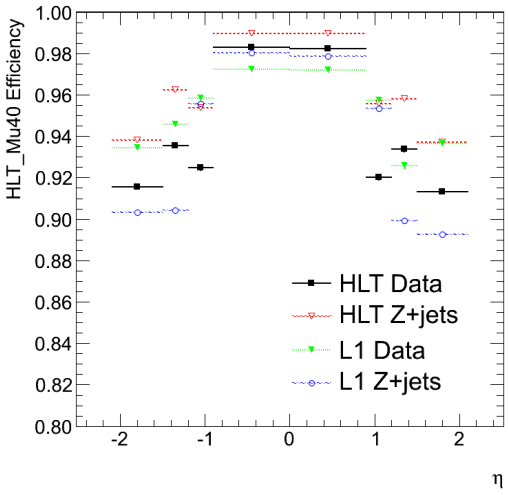


Figure 11.14: L1 and HLT efficiencies for HLT_Mu40 measured in regions of η.

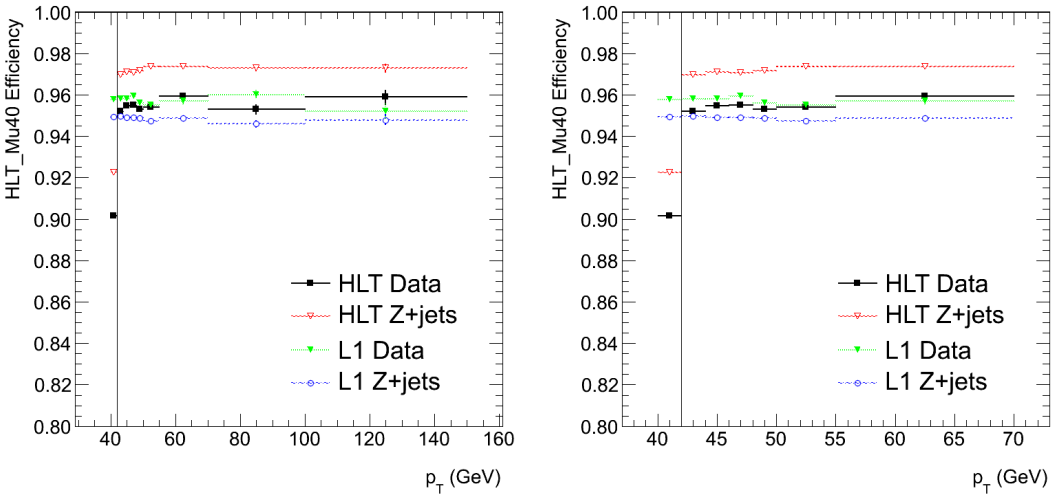


Figure 11.15: L1 and HLT efficiencies for HLT_Mu40 measured in regions of p_{T}. Muons selected in the cross section analysis are required to have $p_{\mathrm{T}} > 42$ GeV.

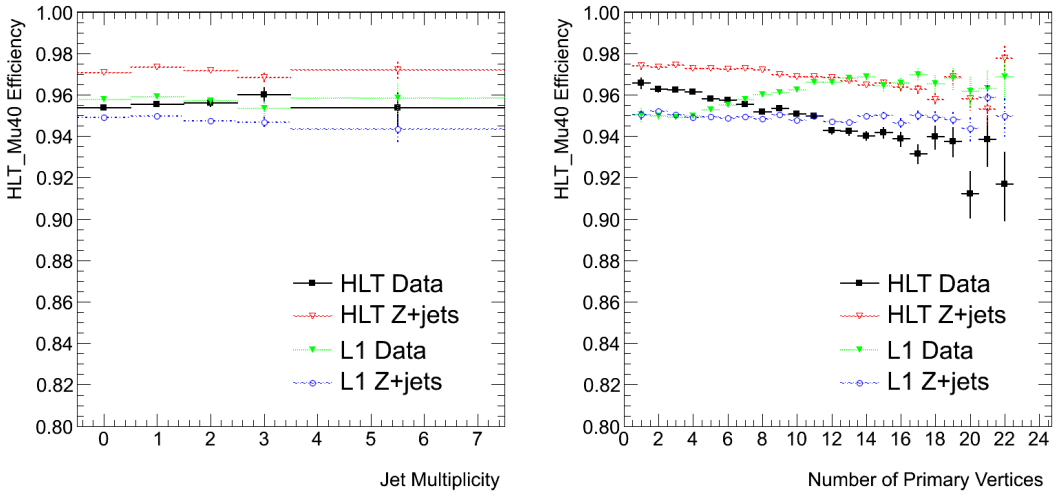


Figure 11.16: L1 and HLT efficiencies for HLT_Mu40 measured in regions of jet multiplicity and number of primary vertices.

11.3.2 Efficiency Evolution

The evolution of the L1 and HLT components of HLT_Mu40 for each trigger menu is examined to study variations in the trigger efficiency. The trigger menus and the changes made to the trigger are described in Section 7.4.

The evolution of the L1 trigger with respect to η is shown on the left in Figure 11.17. The change in the L1 muon p_{T} assignment in a revision of the 3e33 menu results in a significant increase in the L1 efficiency across the full η range.

The evolution of the HLT trigger with respect to η is shown on the right in Figure 11.17. The drop in efficiency in the $0.9 < |\eta| < 1.5$ region in the 2e33 menu is due to the addition of L2 quality criteria. Otherwise the HLT efficiency decreases with each new menu due the efficiency dependence on the number of primary vertices in the event, caused by the HLT quality criteria. The instantaneous luminosity of the data collected increases with each new menu, corresponding to an increase in pile-up and therefore number of primary vertices. Since the data sample is concentrated in higher primary vertex multiplicities, corresponding with lower trigger efficiencies in Figure 11.16, the result is a downward trend in the overall HLT efficiency.

The L1 and HLT efficiencies in regions of muon p_{T} are shown in Figures 11.18 and 11.19. While the overall efficiency varies, the efficiency dependence for muons with p_{T} greater than 42 GeV is stable for each trigger menu.

The L1 and HLT efficiency with respect to jet multiplicity and number of primary vertices in the event is shown in Figure 11.20 and 11.21. The overall efficiency varies,

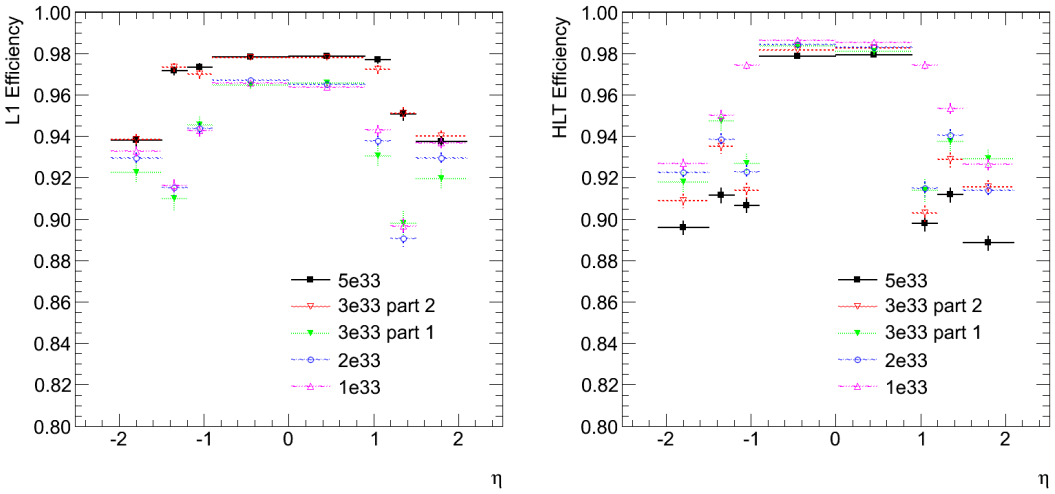


Figure 11.17: L1, left, and HLT, right, efficiency for HLT_Mu40 measured for each trigger menu in regions of η.

however there are no significant deviations in the behaviour of the trigger with respect to the jet multiplicity or number of primary vertices.

A summary of the overall efficiency for each trigger menu is shown in Figure 11.22.

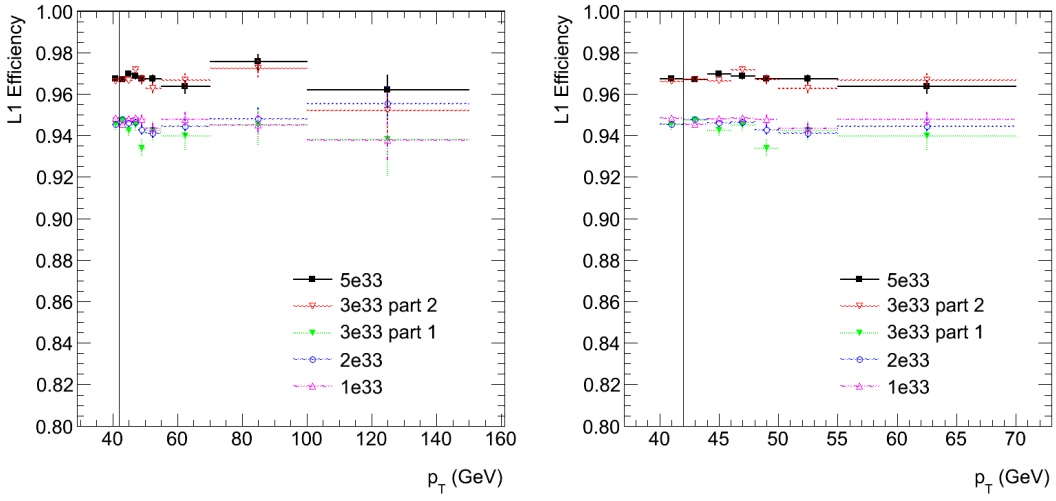


Figure 11.18: L1 efficiency for HLT_Mu40 measured for each trigger menu in regions of p_{T}. Muons selected in the cross section analysis are required to have $p_{\mathrm{T}} > 42$ GeV.

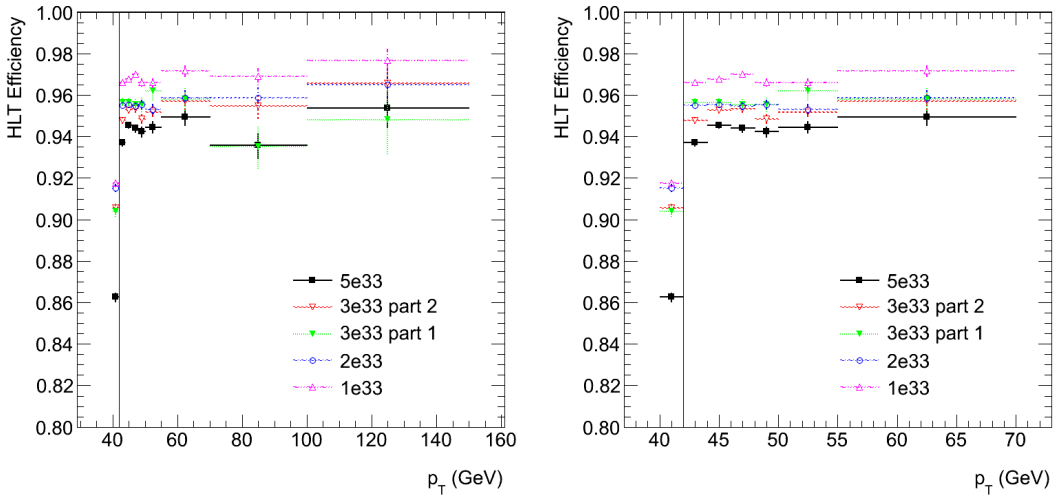


Figure 11.19: HLT efficiency for HLT_Mu40 measured for each trigger menu in regions of p_{T}. Muons selected in the cross section analysis are required to have $p_{\mathrm{T}} > 42$ GeV.

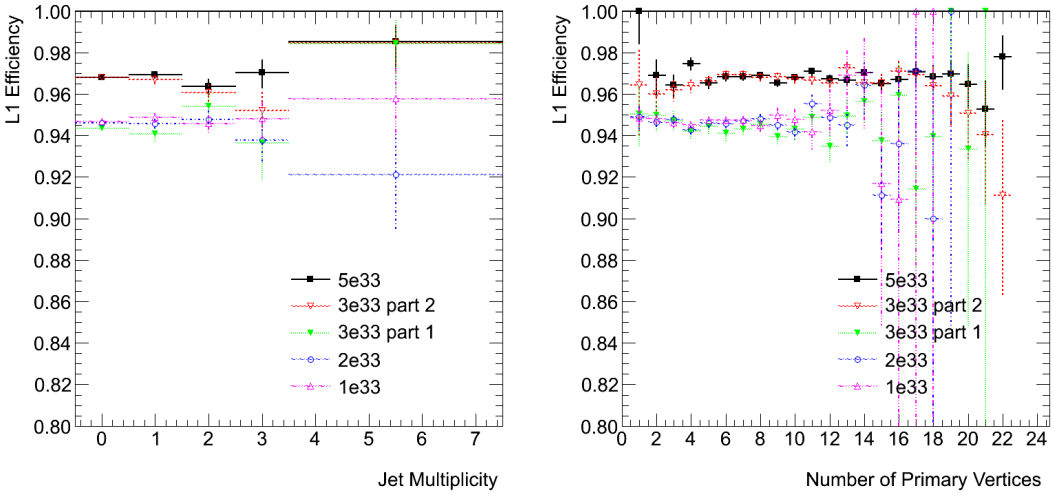


Figure 11.20: L1 efficiency for HLT_Mu40 measured for each trigger menu in regions of jet multiplicity and number of primary vertices.

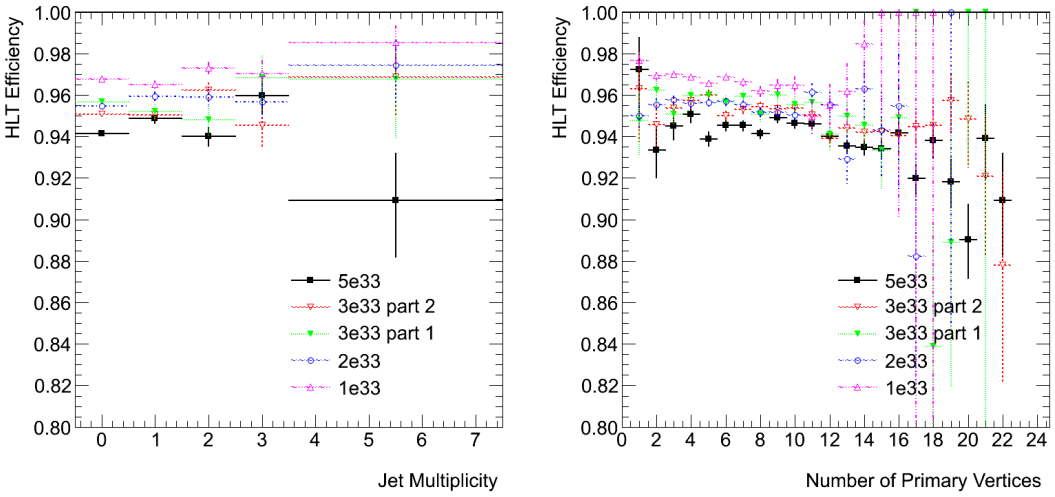


Figure 11.21: HLT efficiency for HLT_Mu40 measured for each trigger menu in regions of jet multiplicity and number of primary vertices.

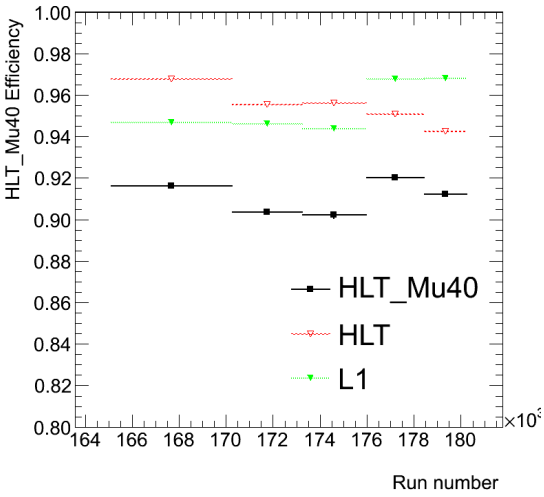


Figure 11.22: HLT_Mu40, L1 and HLT efficiencies measured for each trigger menu. The bin boundaries correspond with the boundaries of the trigger menus.

Part V

Cross Section Extraction

Chapter 12

Cross Section Extraction

The $t\bar{t}$ production cross section is measured with Equation 1.1. The selection efficiency for $t\bar{t}$ events in all decay channels, $\varepsilon_{t\bar{t}} = 0.02520 \pm 0.00002$ in Table 10.7, is estimated in simulation with corrections derived from data applied as described in Chapter 11. It is derived from the efficiency to select muon + jets events and the efficiency to select other $t\bar{t}$ events, combined according to their respective branching fractions. The calculation of the integrated luminosity of the collected data, $\int \mathcal{L} = 4.76$ fb^{-1} $\pm 2.2\%$, is described in Section 6.7. The only remaining information required for the cross section measurement is the number of selected $t\bar{t}$ events, $N_{t\bar{t}}$.

The shape of a kinematic distribution for the different physics processes is used to extract the number of signal events from the same distribution in data. If the signal and background shapes, also referred to as templates, are sufficiently different and well estimated, the relative contribution of each physics process to the selected data sample can be extracted with a binned maximum likelihood fit to the data distribution.

Templates are constructed from distributions where the $t\bar{t}$ shape is expected to differ significantly from the dominant W + jets background shape. Four templates are used in the fit to data. The signal $t\bar{t}$ template is used to estimate the number of events from muon + jets $t\bar{t}$ events and other $t\bar{t}$ events. The W/Z + jets template models the contribution from W + jets and Z + jets events. The single top template models the contribution from the single top t-channel, tW and s-channel processes. The $t\bar{t}$, W/Z + jets and single top template shapes are derived from simulation. The multijet template shape is estimated from data in a multijet dominated background region, as described in Section 12.2.

The signal and single top template shapes tend to be similar as both processes involve the decay of a top quark. The single top contribution to the total number of events is expected to be small and the single top cross section is well estimated. Therefore, the single top normalisation in the maximum likelihood fit is constrained by a Gaussian to within 30% of the expected value, where 30% is a conservative constraint at more than four times the uncertainty on the tW cross section prediction in Table 4.1. Events from the tW-channel are expected to dominate the single top contribution to the data sample after event selection, as shown in Table 10.7. Also, the relative uncertainty on the expected tW cross section is larger than for the other single top contributions. The normalisations of the signal, W/Z + jets, and multijet templates are unconstrained in the fit.

The likelihood function used to fit templates to a data distribution is

$$L(\mathrm{N}_k) = \left(\prod_{i=1}^{\mathrm{N_{bins}}} \frac{(\sum_k \mathrm{N}_{k,i})^{n_i} \cdot e^{-(\sum_k \mathrm{N}_{k,i})}}{n_i!} \right) \cdot \left(\frac{1}{\sqrt{2\pi\Delta_{\mathrm{st}}^2}} e^{-\frac{(\mathrm{N_{st}} - \mathrm{N_{Sim\,st}})^2}{2\Delta_{\mathrm{st}}^2}} \right). \quad (12.1)$$

The fitted number of events for template k is N_k, n_i is the observed number of data events in bin i and $\mathrm{N}_{k,i} = \mathrm{N}_k \cdot r_{k,i}$ where $r_{k,i}$ is the template shape value for process k in bin i. The sigma of the single top Gaussian constraint is denoted by Δ_{st}, while $\mathrm{N_{Sim\,st}}$ is the expected number of single top events derived from simulation.

The fit proceeds by varying the normalisation N_k of each template k. The fit parameters are the number of events for each process, $\mathrm{N}_{t\bar{t}}$, $\mathrm{N}_{W/Z+\mathrm{jets}}$, $\mathrm{N_{st}}$ and $\mathrm{N_{Multijet}}$, which scale the normalised template shapes to find the best fit to data. The best fit is found when the product of the per bin likelihood for data to Poisson fluctuate from the predication is maximised. This likelihood maximisation is performed with the `RooFit` toolkit [71], a library of C++ classes designed to facilitate physics analysis modelling in the `ROOT` [72] environment.

12.1 Choice of Template Distribution

The template distribution is chosen such that the signal template shape can be distinguished from the background template shapes. The candidate distributions studied are listed in Table 12.1.

M3 is the invariant mass of the three-jet system with the largest transverse momentum. M_{jjj} and $M_{j\nu\mu}$ are the hadronic and leptonic top masses reconstructed by optimising equation

$$\chi^2 = \frac{(M_W - M_{jj})^2}{\sigma_{M_W}^2} + \frac{M_{jjj} - M_{j\nu\mu}}{\sigma_\Delta^2}, \quad (12.2)$$

where M_W is the nominal mass of the W boson (80.4 GeV), M_{jj} and M_{jjj} denote the invariant mass of a pair and triplet of jets and $M_{j\nu\mu}$ is the invariant mass reconstructed from missing transverse energy, the muon and one jet.

The M3, M_{jjj} and $M_{j\nu\mu}$ distributions, Figures 12.1 and 12.2, are expected to peak close to the top quark mass of 173.5 GeV for top quark events. In contrast, the W/Z + jets and multijet templates are expected to have a broad distribution of events with no significant peak. The minimised χ^2 , Figure 12.1, is also considered as a discriminant as it should reach lower values in $t\bar{t}$ events than in the background events.

The muon η distribution, Figure 12.3, is a candidate as the top quark decay is expected to result in a higher concentration of centrally produced muons than the

background physics processes. Since there are two u quarks and one d quark in the colliding protons, the sample of W + jets events is expected to contain more positively charges muons than negatively charged muons. Therefore multiplying the $|\eta|$ of the muon by its charge should provide further discrimination between the $t\bar{t}$ and W + jets template shapes. The $|\eta_\mu| \cdot \mu$ charge template shapes are shown in Figure 12.3.

The validity and stability of the fit performed with each distribution is examined to determine if the distribution is suitable, as described in Sections 12.3 and 12.4. All of the candidate distributions pass the validity checks.

The distribution to be used in the cross section extraction is decided based on the expected correlation between the $t\bar{t}$ and background fit parameters. The correlations between the $t\bar{t}$ and unconstrained background fit parameters are listed in Table 12.1. The large correlation between the multijet and $t\bar{t}$ fit parameters measured with the $|\eta_\mu|$ and $|\eta_\mu| \cdot \mu$ charge templates indicates that the fit has difficulty distinguishing between the respective contributions. This difficulty is reflected in the large statistical uncertainty on the fit. The M3 distribution is chosen to extract the $t\bar{t}$ cross section as it results in the lowest correlations.

Fit distribution	Correlation with $N_{t\bar{t}}$		Stat. Uncertainty
	$N_{W/Z+\text{jets}}$	N_{Multijet}	(pb)
M3	0.01	0.3	3.9
M_{jjj}	0.3	0.5	4.1
$M_{j\nu\mu}$	0.6	0.5	4.9
χ^2	0.1	0.6	4.8
$\|\eta_\mu\|$	0.6	0.9	10.5
$\|\eta_\mu\| \cdot \mu$ charge	0.4	0.9	9.3

Table 12.1: Candidate distributions for the extraction of $N_{t\bar{t}}$ in data with template fitting. The absolute magnitude of the correlations between the $t\bar{t}$ fit parameter and the dominant backgrounds is quoted. The statistical uncertainty on the fit is also quoted. The correlations and statistical uncertainty are obtained from pseudo-experiments.

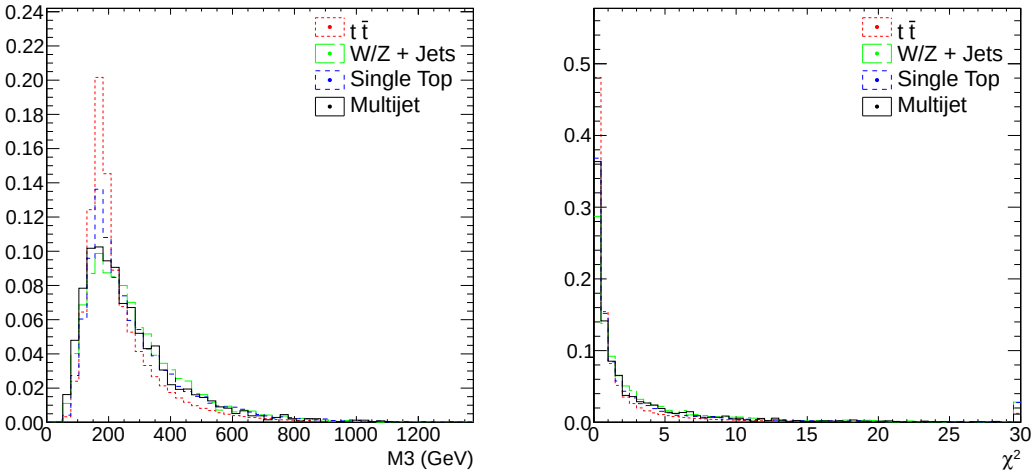


Figure 12.1: M3 and χ^2 template shapes, normalised to unit area.

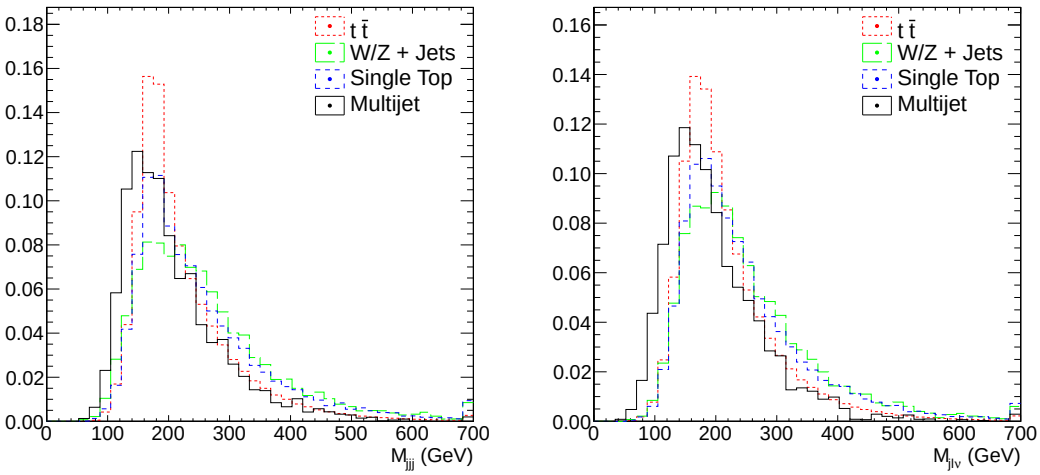


Figure 12.2: M_{jjj} and $M_{j\nu\mu}$ template shapes, normalised to unit area.

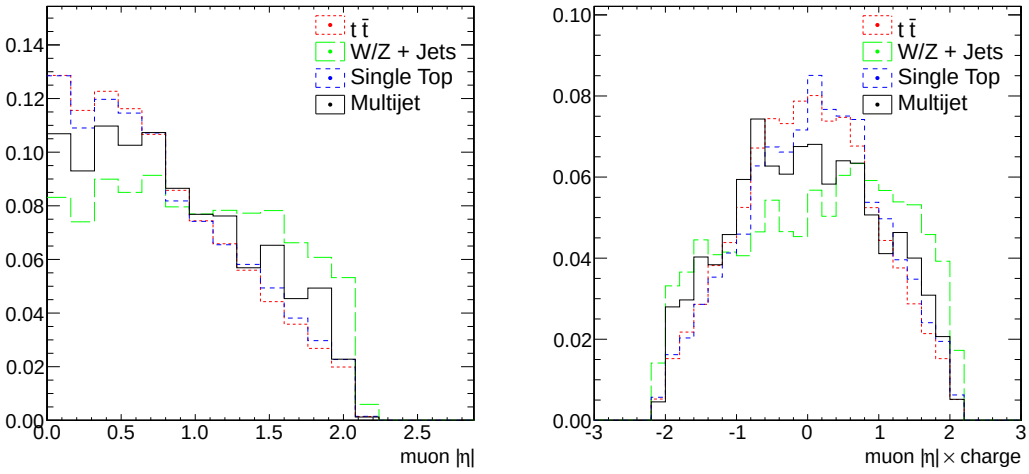


Figure 12.3: $|\eta_\mu|$ and $|\eta_\mu| \cdot \mu$ charge template shapes, normalised to unit area.

12.2 Multijet Template Shape Estimation

The $t\bar{t}$, W/Z + jets and single top templates in the maximum likelihood fit are constructed with simulation. The multijet simulation is statistically limited, therefore it can not be relied upon to provide an accurate estimate of the multijet template shape. Instead, the multijet template shape is extracted from data by selecting a background region, which does not overlap with the signal region, where the multijet contribution is expected to dominate. It is assumed that the template shape is the same in both regions. This assumption is studied later in this section.

The anti-relative isolation (ARI) method is used to estimate the multijet template shape. The background region is selected by inverting the relative isolation requirement on the muon in the event selection. The distributon of selected events with respect to the relative isolation of the muon is shown in Figure 12.4. The simulated samples are normalised according to the expected cross sections from Table 4.1. The multijet contribution is not included in the distribution. The discrepancy between the number of events in simulation and data for muons with relative isolation greater than 0.2 is indicative of the number of multijet events in that region.

The M3 distribution is constructed from the background sample selected with the anti-relative isolation requirement. Simulation is used to estimate the small contamination from the remaining $t\bar{t}$, W/Z + jets and single top events. The contamination is subtracted from the M3 distribution to achieve an estimate of the M3 template shape for multijet events.

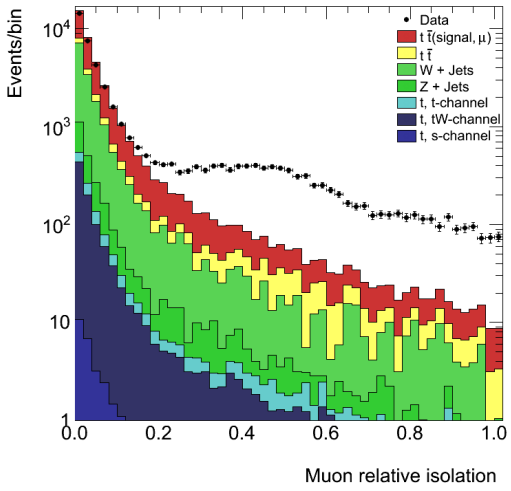


Figure 12.4: Distribution of relative isolation for muons passing the event selection without the $\mathrm{Iso}_{\mathrm{rel.}}(\mu) < 0.15$ requirement applied. The contribution from simulation is normalised according to the expected cross sections in Table 4.1. The prediction for the contribution from multijet events is not included.

Muons in the background region are required to have a relative isolation above 0.35. Events with an additional muon with a relative isolation of less than 0.35 are removed. The 0.35 threshold is chosen to reduce contamination while retaining enough events to avoid unphysical features in the template shape due to statistical fluctuations. Selecting less $t\bar{t}$, W/Z + jets and single top events in the background region reduces the dependence of the estimated template shape on the expected cross sections which are used to subtract the contamination. There is a gap between the > 0.35 threshold and the signal selection of < 0.15 as the intermediate region expects a large contribution from $t\bar{t}$, W/Z + jets and single top events.

As already mentioned, the ARI method is based on the premise that the template shape to be extracted does not depend on the relative isolation of the muon. This assumption is tested by extracting the template shape from non-overlapping regions of muon relative isolation. The template shape is said to be independent of relative isolation if the template shapes from the separate regions are compatible within statistical uncertainties and do not vary the measured $t\bar{t}$ cross section when used in the fit.

Three non-overlapping background regions are selected based on the requirements of reduced contamination and sufficient event multiplicity. Table 12.2 lists the contamination and number of events in each region.

$\mathrm{Iso}_{\mathrm{rel.}}$	N data	Total cont. (%)	$t\bar{t}$ (%)	W/Z + jets (%)	Single top (%)
$0.35 - 0.5$	2900	20.9	13.8	6.2	0.9
$0.5 - 0.75$	2694	17.0	11.9	4.2	0.8
$0.75 - 1.75$	2816	18.1	11.0	6.3	0.7

Table 12.2: Number of selected data events and relative contamination for different values of minimum relative isolation. The nominal multijet template is taken from the $0.35 < \mathrm{Iso}_{\mathrm{rel.}} < 0.5$ region. The signal selection requires muons with $\mathrm{Iso}_{\mathrm{rel.}} < 0.15$.

The M3 template shapes extracted from the three regions of relative isolation are shown in Figure 12.5. The M3 template shapes extracted from the relative isolation regions closest to the signal region agree within the statistical uncertainty. However, the peak in the template shape extracted from the $0.75 < \mathrm{Iso}_{\mathrm{rel.}} < 1.75$ region is noticeably subdued. This suggests a template shape dependence on the relative isolation of the muon. In order to reduce the effect of any dependence on the measured $t\bar{t}$ cross section, the M3 template shape extracted from the region closest to the signal region is selected as the nominal multijet template. The template shapes extracted from the remaining regions are used to estimate the systematic effect of a template shape dependence with respect to relative isolation on the measured $t\bar{t}$ cross section, as explained in Section 13.11.

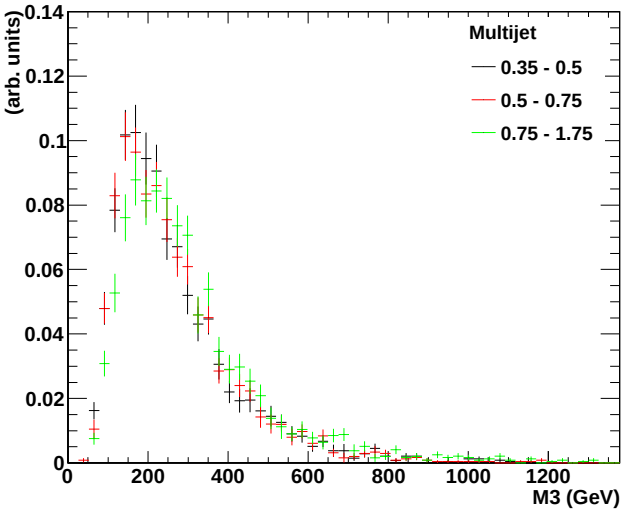


Figure 12.5: M3 multijet template shape estimated in non-overlapping regions of muon relative isolation.

12.2.1 Minimum Jet Muon Distance Requirement

The requirement on the minimum distance between the selected muon and the closest jet, $\Delta\mathrm{R}(\mu,\mathrm{jet}) > 0.3$, is introduced to avoid a systematic contribution to the modelling of the multijet template shape.

Figure 12.6 compares the $\Delta\mathrm{R}(\mu,\mathrm{jet})$ distribution for data events selected with the signal selection and with the ARI requirement. In both case the $\Delta R > 0.3$ requirement is not applied. There are relatively few events in the $\Delta R < 0.3$ region with the signal selection applied, however the events in this region dominate the ARI sample.

In Figure 12.7 a comparison is made of the M3 multijet distributions extracted using the ARI method with the requirements of $\Delta\mathrm{R} < 0.3$ and $\Delta\mathrm{R} > 0.3$. The M3 multijet distribution extracted from the $\Delta R < 0.3$ region appears to peak at a higher M3 value than the distribution extracted from the $\Delta R > 0.3$ region. Since the $\Delta\mathrm{R}(\mu,\mathrm{jet})$ of the selected muon appears to affect the M3 template shape, and there are few events with $\Delta R < 0.3$ in the signal region, the $\Delta R > 0.3$ requirement is applied when selecting muons with the ARI method. The $\Delta R > 0.3$ requirement is also applied in the signal selection for consistency.

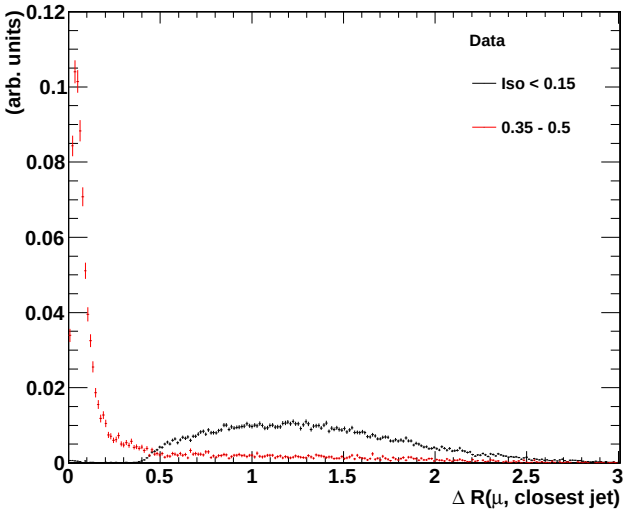


Figure 12.6: $\Delta\mathrm{R}(\mu,\mathrm{jet})$ for data events selected with the signal selection compared to data events selected with the ARI $0.35 < \mathrm{Iso}_{\mathrm{rel.}} < 0.5$ requirement. The number of events in each distribution is normalised to one.

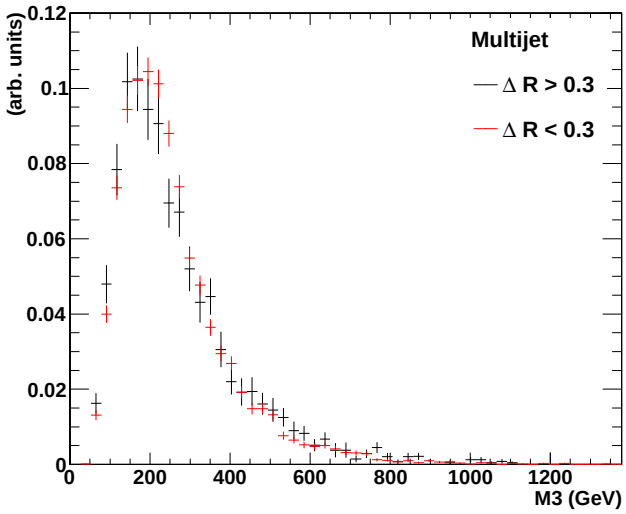


Figure 12.7: M3 multijet template shape estimated in the ARI $0.35 < \mathrm{Iso}_{\mathrm{rel.}} < 0.5$ region with $\Delta\mathrm{R}(\mu, \mathrm{jet}) < 0.3$ and $\Delta\mathrm{R}(\mu, \mathrm{jet}) > 0.3$. The number of events in each distribution is normalised to unit area.

12.3 Fit Stability

The maximum likelihood fit is binned, therefore the result of the fit should not vary according to the bin width chosen. The stability of the fit with respect to bin width is examined by repeating the fit for a range of bin widths. The fit uncertainty is expected to increase as the bin widths become so large than the features of the signal and background template shapes are no longer distinguishable. The fit is considered stable if the fit uncertainty and number of fitted events are consistent for a range of bin widths. The bin width for the cross section extraction is chosen to be within this range.

The stability of the fitted cross section and uncertainty are shown in Figure 12.8 for simulation based pseudo-experiments. As expected, the uncertainty on the measured cross section increases with respect to the bin width chosen. The uncertainty is stable for bin widths between 20 and 28 GeV. Therefore, a bin width of 26 GeV is chosen for the cross section measurement. The measured cross section is completely stable with respect to bin width chosen. This is because the pseudo-data is generated from the simulation, therefore there are no systematic shape differences between the template shapes and pseudo-data which may affect the rebinned results.

The stability measured with fits to data is shown in Figure 12.9. There is a noticeable trend in the measured cross section at the limits of the the bin widths studied. However,

the cross section is stable when measured with M3 bin widths between 20 and 30 GeV. The fit uncertainty is stable for bin widths between 18 and 28 GeV.

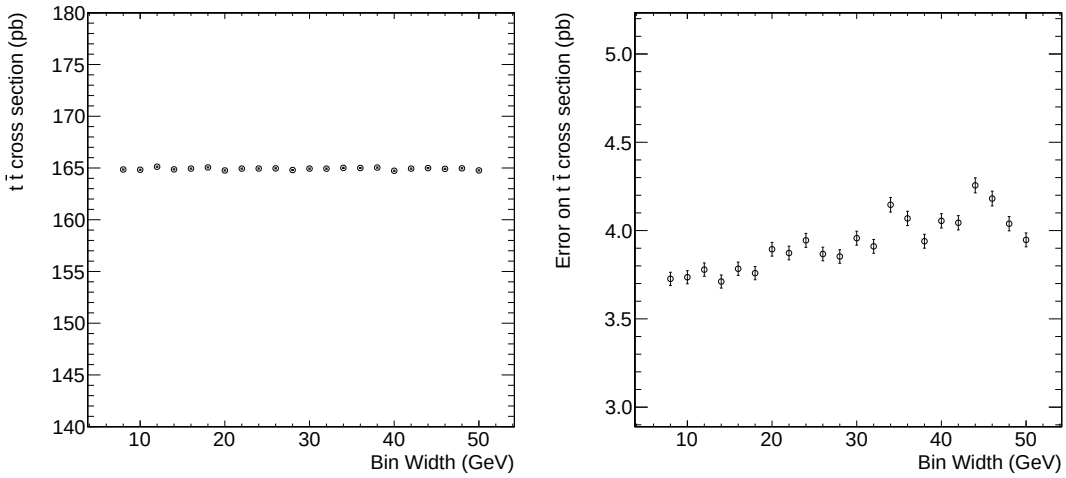


Figure 12.8: Stability of the template fit to pseudo-data with respect to bin width in pseudo-experiments. A bin width of 26 GeV is chosen for further studies. The statistical uncertainty on the $t\bar{t}$ cross section (error) is given by the uncertainty on the mean (sigma) determined by a Gaussian fit to the distribution for the $t\bar{t}$ cross section from the pseudo-experiments.

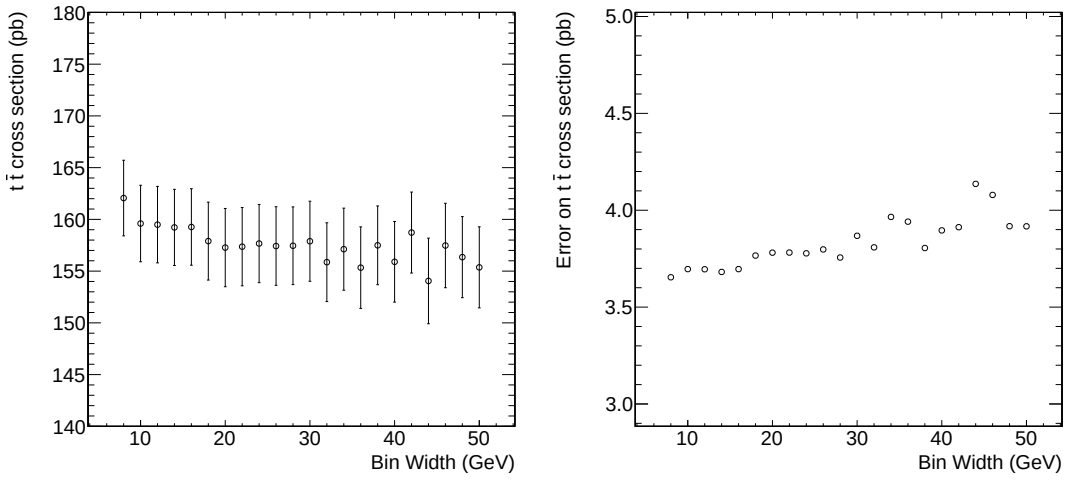


Figure 12.9: Stability of template fit to data with respect to bin width in data. A bin width of 26 GeV is chosen for further studies. The statistical uncertainty on the measured $t\bar{t}$ cross section is the fit error shown in the figure on the right. There is no statistical uncertainty shown for the fit error on the measured cross section.

12.4 Method Validation

Before applying the fit to the data, the validity of the template fitting method is examined by performing simulation based pseudo-experiments. In these experiments the templates described in the previous section are fit to pseudo-data generated from simulation. The pseudo-data is generated from a distribution where the templates for each process have been scaled to the expected number of events in Table 10.7. In each pseudo-experiment, event counts are Poisson fluctuated around the expectation, according to the statistical uncertainty for the expected number of events in each bin, to generate the pseudo-data for that pseudo-experiment. The unchanged template shapes are then fit to the Poisson fluctuated pseudo-data distribution.

Pseudo-experiments are performed to quantify the ability of the likelihood maximisation to separate the contributions from the different processes. The separation power of the fit is reflected in the correlation between the fit parameters. Pseudo-experiments are also used to test for sources of bias in the fit, or inaccurate estimates of the fit uncertainty. This is achieved by filling a distribution for the fitted number of events, the fit uncertainties, and the pull resulting from each pseudo-experiment. The pull is calculated as described in [73] as

$$\mathrm{pull_k} = \frac{\mathrm{N}_k - \mathrm{N}_{\mathrm{Sim}\,k}}{\sigma_{\mathrm{N}_k}} \tag{12.3}$$

where σ_{N_k} is the uncertainty on N_k, the fitted number of events for process k. $\mathrm{N}_{\mathrm{Sim}\,k}$ is the expected number of events for process k.

A fit of a Gaussian function is applied to the pull distribution. A Gaussian mean value of zero indicates the likelihood fit is unbiased. A Gaussian sigma of one indicates the uncertainty estimated by the likelihood fit is accurate.

Distributions for the fitted number of $t\bar{t}$ events, the uncertainty on the number of events, and the pull are filled for 5000 pseudo-experiments. The sigma of the fitted number of $t\bar{t}$ events in Figure 12.10, left, is consistent with the mean of the distribution for the fit uncertainty in Figure 12.10, right. This indicates that the fit uncertainty is well estimated. The pull distribution in Figure 12.11 provides further validation of the method as it is centered around zero with a sigma of one signifying an unbiased method with a well estimated fit uncertainty.

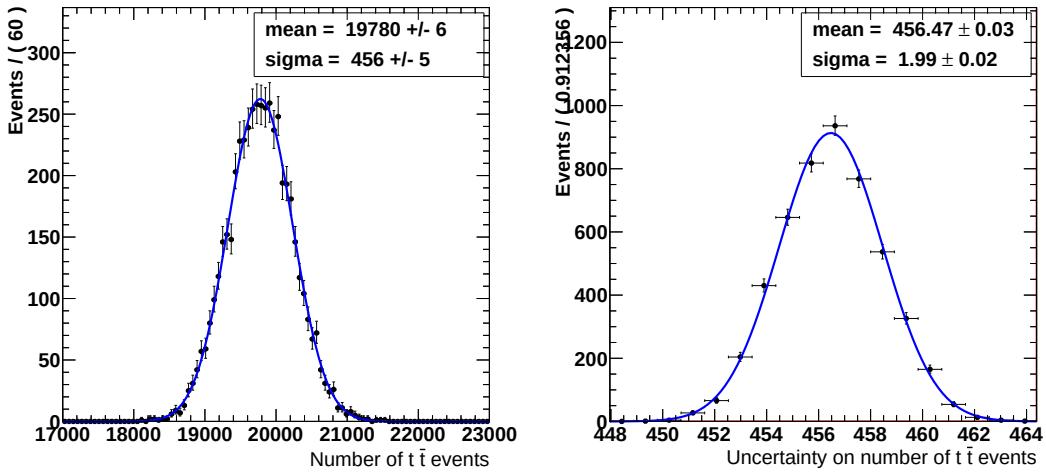


Figure 12.10: Distribution of the number of $t\bar{t}$ events, left, and the uncertainty on this number, right, from pseudo-experiments.

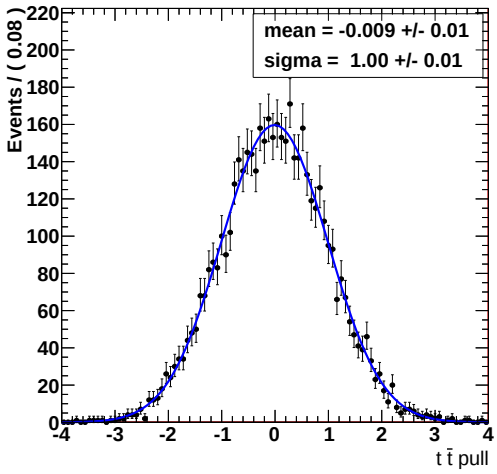


Figure 12.11: Distribution of the fit pull calculated in pseudo-experiments to validate the template fitting method.

12.5 Measured Cross Section

The $t\bar{t}$ cross section is extracted by fitting templates to the M3 distribution of data events passing the event selection. The correlations between the templates are shown in Table 12.3. The low correlation between the $t\bar{t}$ template and the background templates indicates the likelihood fit is well able to separate the signal contribution from the background contributions. The fit is less able to separate the W/Z + jets and multijet contributions due to the large anti-correlation between their fit parameters.

The measured $t\bar{t}$ cross section is given in Table 12.4. In addition, the expected and measured cross sections for each of the physics processes represented in the likelihood fit are shown. In Figure 12.12, good agreement is observed between the fit results and data after normalising the templates to the result of the template fit and overlaying with the M3 distribution in data.

The deviation from the expected cross sections for W/Z + jets and multijet events is explained by the large anti-correlation, given in Table 12.3, between their fit parameters. The fit has difficulty distinguishing between the contribution from W/Z + jets and multijet events due the similarity in their template shapes, shown in Figure 12.1. Therefore the combined number of W/Z + jets and multijet events estimated by the fit, 8820+4659 = 13479, should be considered. Assuming the simulated ratio of W/Z + jets to multijet events is accurate, it can be used to calculate the individual cross sections as $\sigma_{W/Z+\text{jets}} = 31991\,\text{pb}$ and $\sigma_{\text{QCD}} = 0.07\,\text{pb}$, which show agreement with the SM expec-

tation on the order of one sigma. The fit is repeated with the W/Z + jets normalisation constrained to the SM expectation. The resulting $t\bar{t}$ cross section is 156.7 pb, indicating that while the measured W/Z + jets and multijet cross sections can not be considered separately, the measured $t\bar{t}$ cross section is stable.

	$t\bar{t}$	W/Z + jets	Single Top	Multijet
$t\bar{t}$	1	-0.014 ± 0.012	-0.278 ± 0.004	-0.330 ± 0.007
W/Z + jets	-0.014 ± 0.012	1	-0.294 ± 0.008	-0.867 ± 0.005
Single Top	-0.278 ± 0.004	-0.294 ± 0.008	1	0.069 ± 0.012
Multijet	-0.330 ± 0.007	-0.867 ± 0.005	0.069 ± 0.012	1
N events	18875^{+456}_{-455}	8820^{+927}_{-921}	1197 ± 334	4659^{+933}_{-934}

Table 12.3: Correlations between the templates in the likelihood fit to data. N events is the number of events estimated by the fit for each template as represented by the fit parameters $N_{t\bar{t}}$, $N_{W/Z+\text{jets}}$, N_{st} and N_{Multijet}. The correlation values are from the fit to data. The uncertainty on the correlations is taken from the sigma of the Gaussian distribution of correlation values in the pseudo-experiments.

Template	Expected σ (pb)	Measured σ (pb)
$t\bar{t}$	165	157.4 ± 3.8
W/Z + jets	$31314 + 3048 = 34362$	21456^{+2255}_{-2241}
Single Top	$64.6 + 4.6 + 15.7 = 84.9$	87 ± 24
Multijet	0.07 ± 0.02	1.0 ± 0.2

Table 12.4: Expected and measured cross sections for the physics processes contributing to the selected data sample. The multijet cross section is the cross section for events passing the event selection.

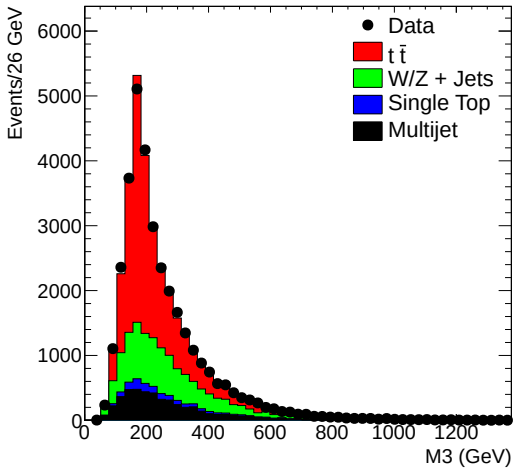


Figure 12.12: Result of the likelihood fit compared to the M3 distribution in data.

12.6 Kinematic Distributions

Kinematic distributions are produced with events passing the event selection and templates normalised to the fit results in Table 12.4 to examine the agreement between the fit results and the data with respect to various muon and jet quantities.

The distribution of events with respect to muon p_{T}, η and ϕ is shown in Figures 12.13 and 12.15. There is a discrepancy between the distribution of events with respect to muon p_{T} in the simulation scaled to the fit results and in data, especially in the low p_{T} region. This is explained by the difference between the muon p_{T} distribution for W/Z + jets and multijet events, and the high correlation between the two in the cross section extraction. Due to the high correlation between the fit parameters, the likelihood fit to the M3 distribution has difficulty distinguishing between W/Z + jets and multijet events. This results in a W/Z + jets normalisation which is significantly below the theoretical prediction and a multijet normalisation which is an order of magnitude above the LO expectation. The shape of the muon p_{T} distribution is significantly different for W + jets and multijet events, as shown in Figure 12.14, therefore the ambiguity between the W/Z + jets and multijet normalisations in the fit results in a discrepancy in the muon p_{T} distribution.

The p_{T} and η distributions for the four highest jets are shown in Figures 12.16 to 12.19. The jets are sorted in order of decreasing p_{T}. The jet multiplicity and $\Delta\mathrm{R}(\mu, \mathrm{jet})$ distributions are shown in Figure 12.20. The distributions show reasonable agreement between data and simulation.

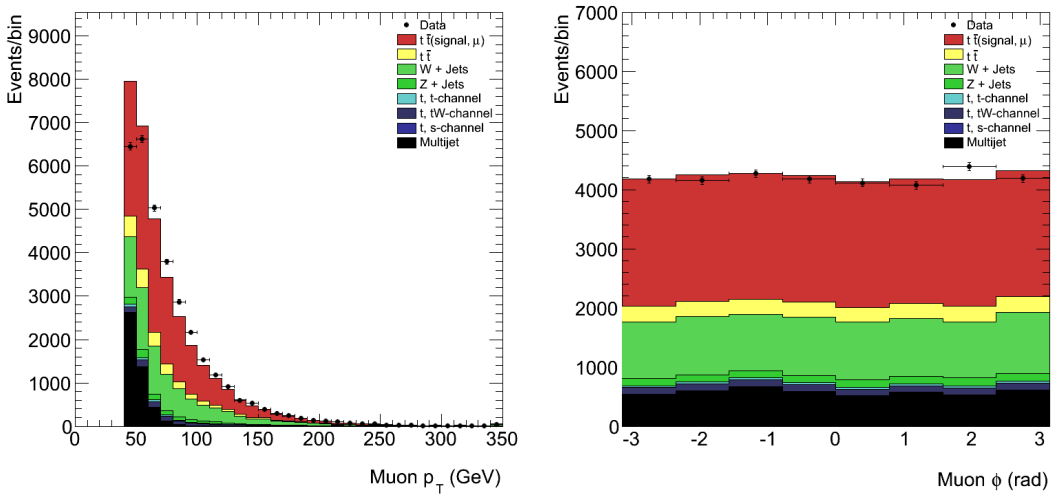


Figure 12.13: Muon p_{T}, left, and ϕ, right.

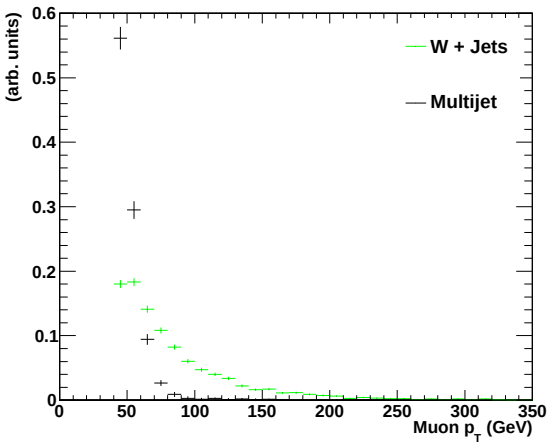


Figure 12.14: Muon p_{T} distribution for W + jets and multijet events normalised to unit area.

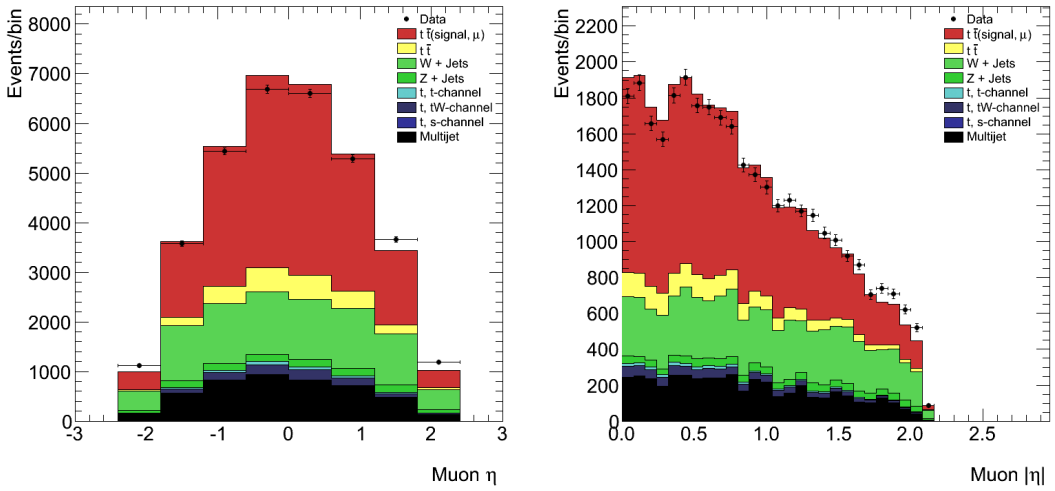


Figure 12.15: Muon η, left, and $|\eta|$, right.

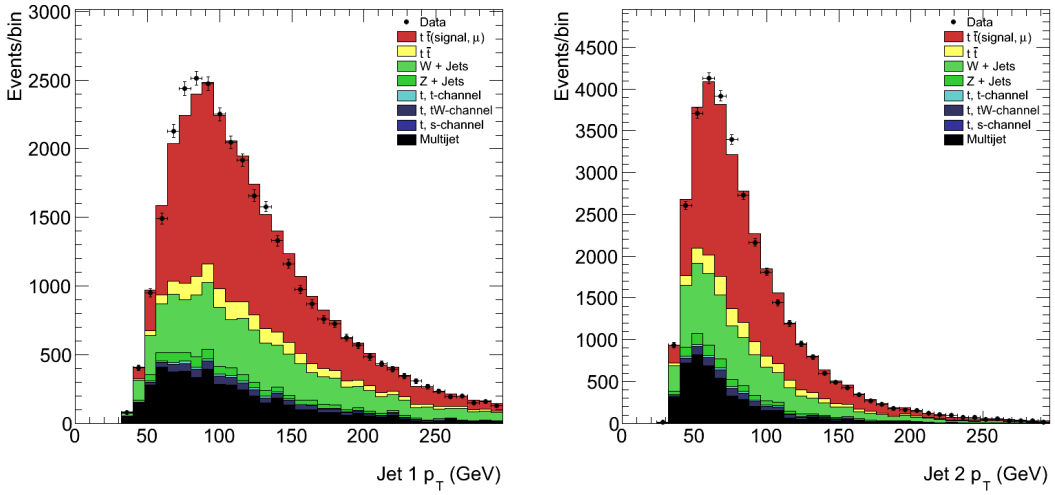


Figure 12.16: 1st and 2nd jet p_{T}.

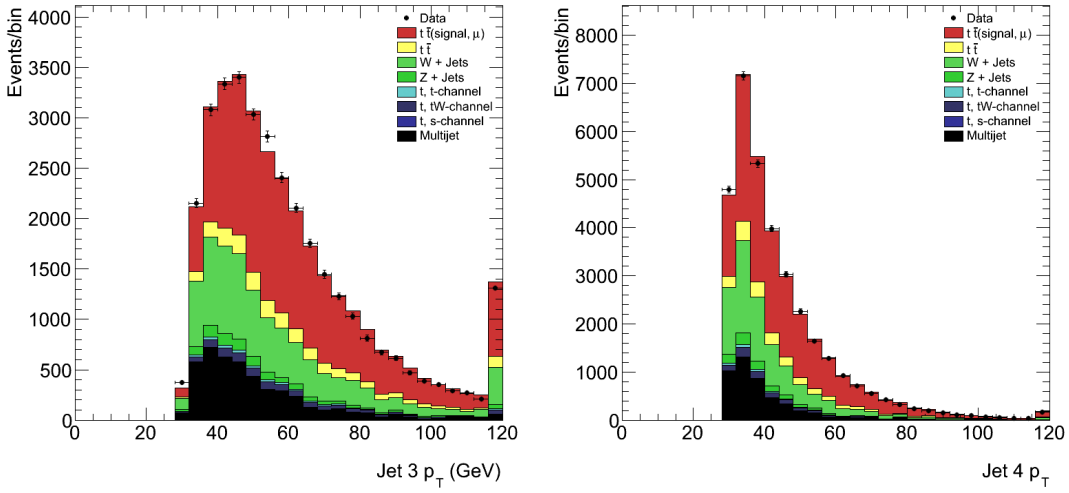


Figure 12.17: 3rd and 4th jet p_{T}.

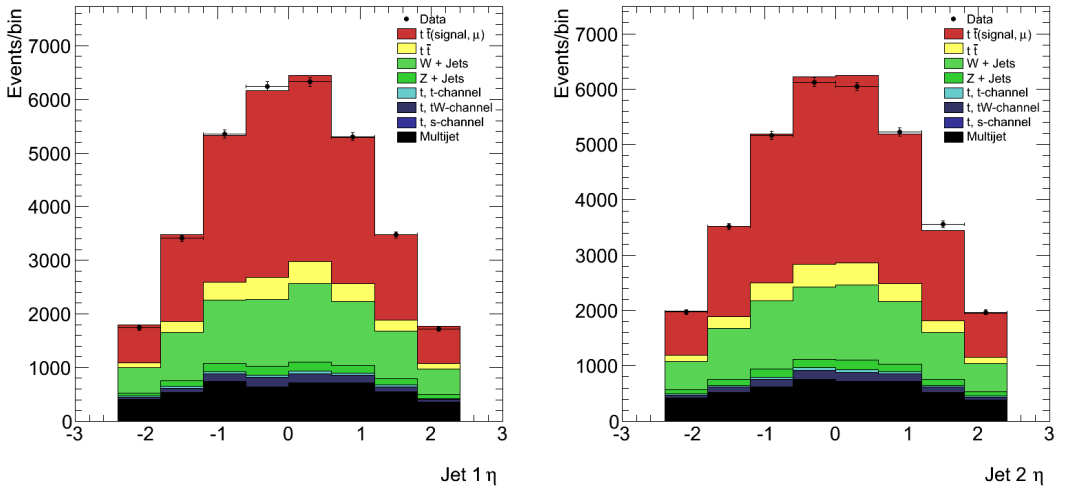


Figure 12.18: 1st and 2nd jet η.

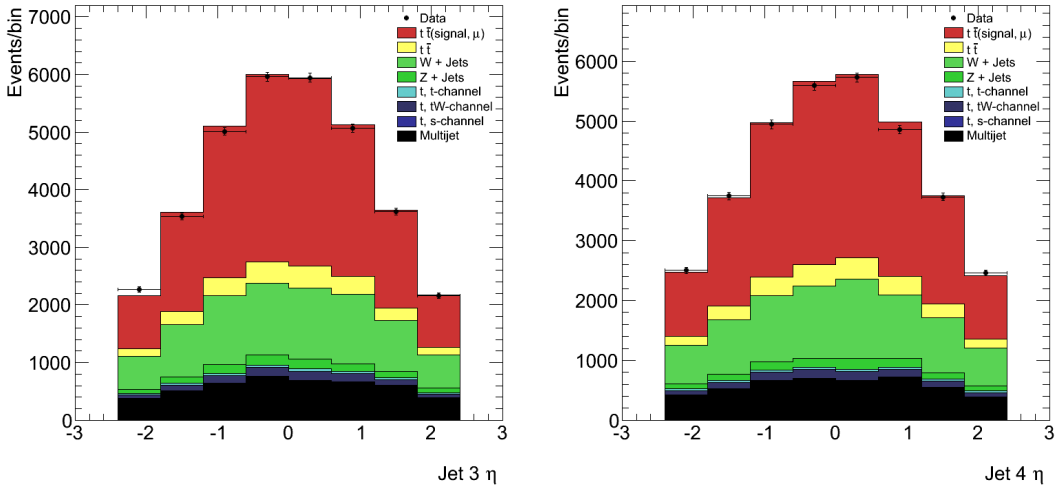


Figure 12.19: 3$^{\text{rd}}$ and 4$^{\text{th}}$ jet η.

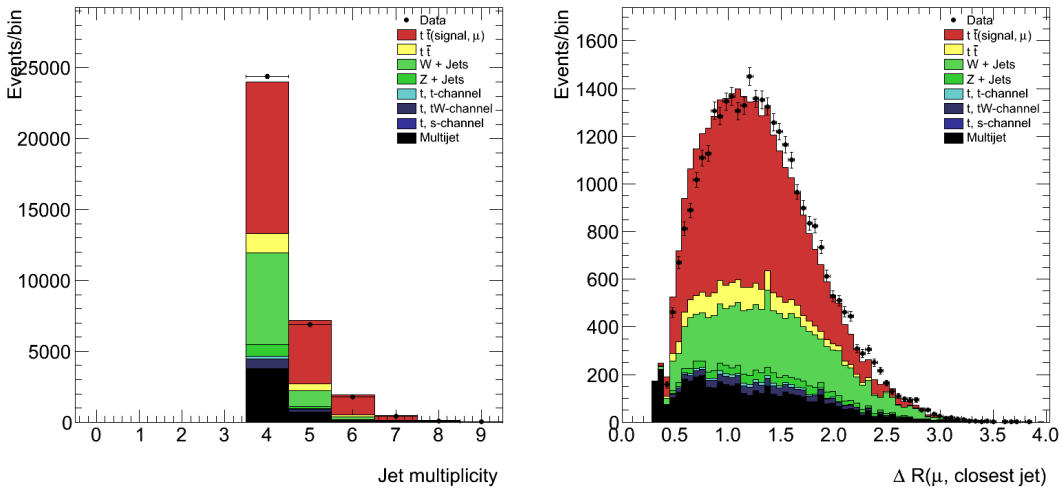


Figure 12.20: Jet multiplicity in selected events, left, and the distance between the selected muon and the closest jet, right.

Chapter 13

Systematic Uncertainties

In this chapter the sources of systematic uncertainty which are expected to have a non-negligible effect on the cross section measurement are introduced. The estimation of the systematic uncertainty on the measured cross section due to each source is described. The total systematic uncertainty on the measured cross section is then calculated by summing the measured systematic uncertainties for each source in quadrature.

The measured systematic uncertainties are summarised in Table 13.3. The $t\bar{t}$ selection efficiencies for each source of systematic systematic uncertainty are summarised in Table 13.4.

13.1 Estimating the Systematic Uncertainty

The effect of a systematic uncertainty on the cross section measurement is quantified by repeating the cross section measurement with the systematic variation applied to the simulated events. A systematic variation can affect the template shapes or the signal selection efficiency. A systematic variation which is applied to the $t\bar{t}$ sample will affect both.

If the systematic variation alters at least one of the template shapes, the number of $t\bar{t}$ events in data is re-estimated by fitting the altered template(s) to the unchanged data. The systematic $t\bar{t}$ cross section is then extracted with Equation 1.1 using the systematically altered values for the signal selection efficiency and the number of $t\bar{t}$ events as shown:

$$\sigma_{t\bar{t}}^{\text{systematic}} = \frac{N_{t\bar{t}}^{\text{systematic}}}{\int \mathcal{L} \cdot \varepsilon_{t\bar{t}}^{\text{systematic}}} \tag{13.1}$$

The systematic uncertainty on the nominal cross section is quoted as the difference between the nominal cross section and the cross section measured with the systematic variations applied:

$$\text{Systematic uncertainty} = \sigma_{t\bar{t}}^{\text{systematic}} - \sigma_{t\bar{t}} \tag{13.2}$$

The measured systematic uncertainty is cross checked by repeating the uncertainty estimation using only simulation. In this case the systematically varied template shapes are fit to pseudo-data as described in Section 12.4, with the pseudo-data generated using the nominal template shapes and normalisations. The simulated uncertainty is then extracted with Equation 13.2, where $\sigma_{t\bar{t}}$ is the predicted cross section of 165 pb^{-1}.

13.2 Factorisation and Renormalisation Q^2 Scale

The factorisation and renormalisation scale μ has been introduced in Section 3.2 and defined in Section 8.3. To account for ambiguity in the choice of values for the Q^2 scale, physics samples are generated with the Q^2 scale varied by a factor of two up and down in both the ME and PS simulation. These samples also account for variations to the initial and final state radiation. The Q^2 variation in $t\bar{t}$ events is considered uncorrelated to the variation in W/Z + jets events. The uncertainty on the measured $t\bar{t}$ cross section is estimated by re-evaluating the cross section using the samples with the varied Q^2 scales.

The scale values are shown in Table 13.1. The parameters `factscfact` and `renscfact` are the factorisation and renormalisation scale factors respectively. The `scalefact` parameter multiplies the factorisation scale with the renormalisation scale of QCD vertices that are not considered "extra jet" vertices. The `alpsfact` parameter multiplies the renormalisation scale for extra jets. For space-like PS evolution `PARP(64)` is multiplied by k^2, the squared transverse momentum evolution scale. This determines the production of additional parton branchings between the hard interaction and the initial pp state. For time-like PS evolution Λ_{QCD} is given by `PARP(72)`.

Variation	MadGraph		PYTHIA	
	`scalefact`/`factscfact`	`alpsfact`/`renscfact`	`PARP(64)`	`PARP(72)`
up	2.0	2.0	4.0	0.125
nominal	1.0	1.0	1.0	0.25
down	0.5	0.5	0.25	0.5

Table 13.1: Factorisation and renormalisation Q^2 scale applied to samples simulated at $\sqrt{s} = 7$ TeV.

The systematic templates obtained by applying the signal selection to the samples simulated with scale variations applied are shown in Figure 13.1. The templates are normalised to unit area. The W/Z + jets systematic templates show features consistent with statistical fluctuations due to a limited number of events in the simulated sample. Therefore the systematic uncertainty due to scale variations is quoted with the statistical uncertainty on the systematic template shape. This is obtained by repeating the fit

to data with the relevant template shape Poisson fluctuated within its statistical uncertainties. The procedure for the estimation of template shape statistical uncertainties is described in Section 13.12.

The systematic uncertainties due to the $t\bar{t}$ and W/Z + jets scale uncertainties have a large effect on the overall systematic uncertainty.

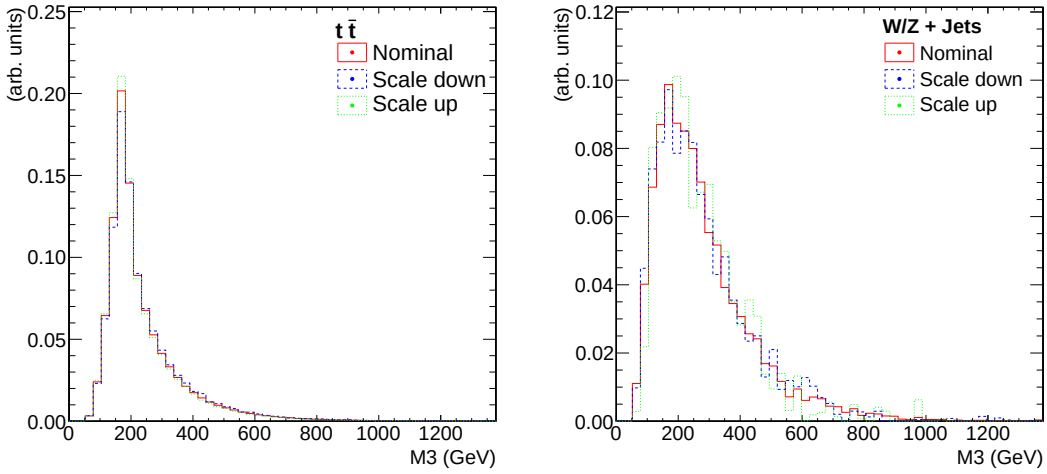


Figure 13.1: $t\bar{t}$ and W/Z + jets templates with scale variations applied.

13.3 Matrix Element to Parton Shower Matching

ME to PS matching is described in Section 8.1.2. In principle the thresholds applied in MadGraph and PYTHIA should be the same, however the jet p_{T} changes due to showering. Therefore the PS threshold is chosen to be 50% higher than the ME threshold.

In order to account for any ambiguity in the choice of matching threshold, the ME/PS samples are simulated with higher or lower matching thresholds applied. The thresholds applied are shown in Table 13.2. The matching variation in $t\bar{t}$ events is considered uncorrelated to the variation in W/Z + jets events. The uncertainty on the measured $t\bar{t}$ cross section is estimated by re-evaluating the cross section using the samples with the varied thresholds applied to construct the templates.

Variation	Threshold (GeV)			
	$t\bar{t}$		W/Z + jets	
	ME	PS	ME	PS
up	40	70	20	30
nominal	20	40	10	20
down	10	20	5	10

Table 13.2: Matching thresholds applied to samples simulated with MadGraph and PYTHIA at $\sqrt{s} = 7$ TeV.

The systematic templates obtained by applying the signal selection to the samples simulated with variations applied to the matching threshold are shown in Figure 13.2. The templates are normalised to unit area. The W/Z + jets systematic templates show features consistent with statistical fluctuations due to a limited number of events in the simulated sample. Therefore the systematic uncertainty due to matching variations is quoted with the statistical uncertainty on the systematic template shape. This is obtained by repeating the fit to data with the relevant template shape Poisson fluctuated within its statistical uncertainties. The procedure for the estimation of template shape statistical uncertainties is described in Section 13.12.

The contribution to the overall systematic uncertainty due to the $t\bar{t}$ matching uncertainty is large, while the W/Z + jets matching uncertainty has a relatively small effect.

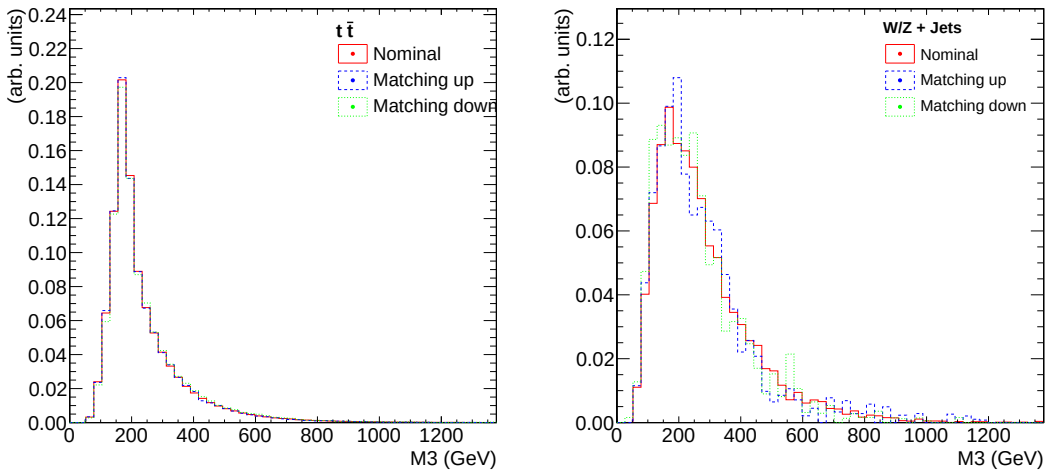


Figure 13.2: $t\bar{t}$ and W/Z + jets templates with matching variations applied.

13.4 Parton Distribution Functions

The significance of PDFs in hadron collisions is explained in Section 3.2.2. The PDF library used for the simulation of events in this analysis is given in Section 8.3. The PDF uncertainties in the CTEQ6.6 PDF set are quantified using the Hessian matrix method [74]. This method yields a PDF set containing 45 members, the central value plus $2 \cdot 22$ variations. The 22 eigenvectors each probe a direction in PDF parameter space. Each eigenvector direction is varied up and down within its tolerance parameter, resulting in $2 \cdot 22$ variations. The tolerance parameter is chosen to encompass variations in the quality of the fit between theoretical calculations and experimental measurements of PDFs when the fit parameters are moved away from the minimum which defines the best fit.

The effect of the PDF uncertainties on the measured $t\bar{t}$ cross section are propagated by repeating the measurement 44 times. Each time the events are weighted according to one of the elements in the error matrix. The uncertainty on the cross section measurement due to PDF uncertainties is then estimated through the use of a "Master Equation" 13.3, taken from [75], which considers maximal positive and negative variations of the physical observable separately.

$$\begin{aligned} \Delta X^{+}_{\mathrm{max}} &= \sqrt{\sum_{i=1}^{N} [\max(X_i^{+} - X_0, X_i^{-} - X_0, 0)]^2} \\ \Delta X^{-}_{\mathrm{max}} &= \sqrt{\sum_{i=1}^{N} [\max(X_0 - X_i^{+}, X_0 - X_i^{-}, 0)]^2} \end{aligned} \tag{13.3}$$

X_0 is the nominal cross section measured, $X_i^{\pm}$ are the cross section values obtained by applying the weight to the simulated events to correspond to the error set.

The 45 PDF templates are normalised to unit area and overlayed in Figures 13.3 and 13.4. The $t\bar{t}$ selection efficiency for each variation is shown on the right in Figure 13.3. The selection efficiency at zero on the x-axis is the nominal selection efficiency. The maximal positive and negative variations of the selection efficiency given by Equation 13.3 are listed in Table 13.4. The PDF uncertainties are a dominant source of systematic uncertainty in the measurement of the $t\bar{t}$ cross section.

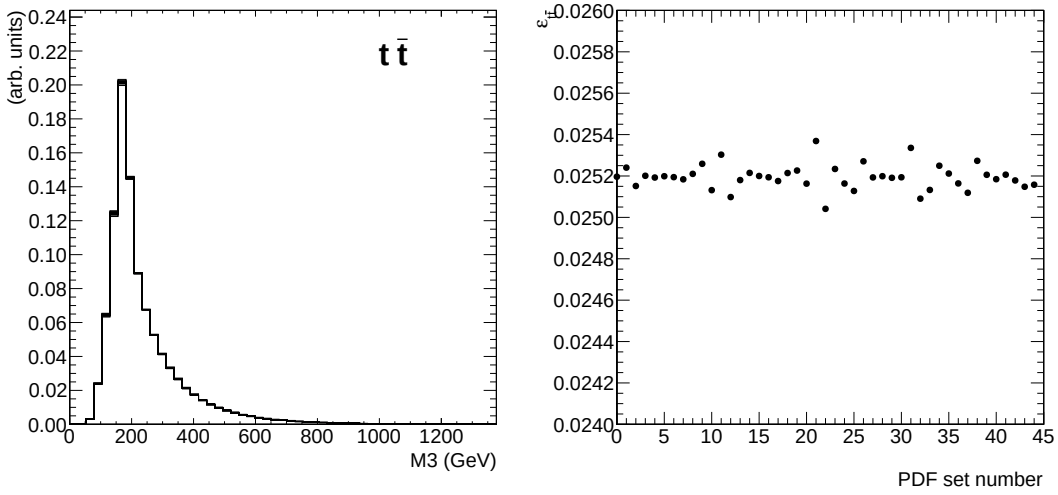


Figure 13.3: $t\bar{t}$ templates with PDF variations applied, left. Selection efficiencies obtained by weighting selected events for each member of the PDF set, right. The selection efficiency at zero on the x-axis is the nominal selection efficiency.

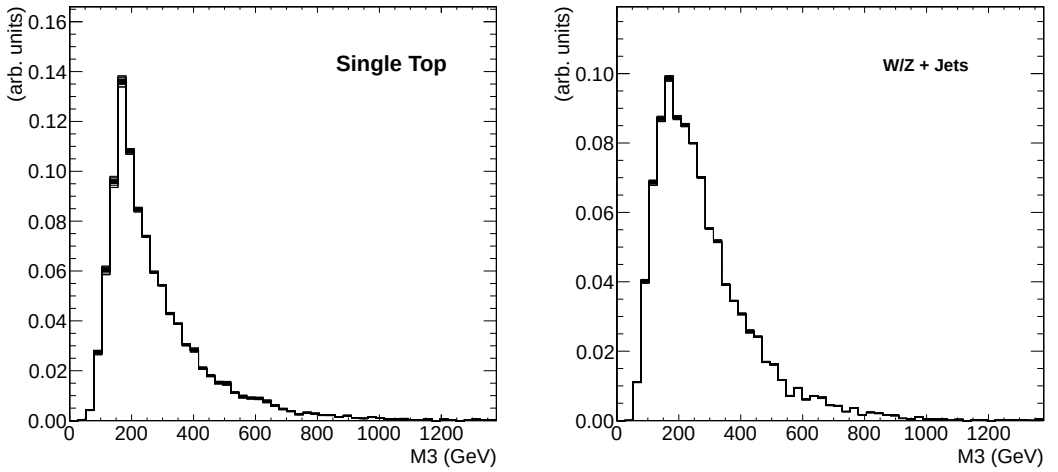


Figure 13.4: Single top and W/Z + jets templates with PDF variations applied.

13.5 Top Quark Mass

The top mass is an input parameter in the generation of $t\bar{t}$ events. However, the simulation does not take into account the uncertainty on this value. The current most precise measurement of the top mass is $173.5 \pm 0.6 \pm 0.8$ GeV [1]. In the $t\bar{t}$ event simulation a top mass of 172.5 GeV is assumed. Therefore, the effect on the measured cross section due to a +2 GeV uncertainty on the input mass must be accounted for.

Eight $t\bar{t}$ samples are simulated with the input top mass varied between 161.5 GeV and 184.5 GeV. The samples are simulated for eight masses, extending beyond the uncertainty on the top mass, to ensure the $t\bar{t}$ cross section can be estimated for any mass value within the simulated range of masses. The $t\bar{t}$ cross section measurement is repeated eight times with the selection efficiency and $t\bar{t}$ template given by a different $t\bar{t}$ mass sample for each measurement. The measured cross section values are plotted against the input top mass values, then a straight-line fit is performed between the measured cross sections. The systematically shifted cross section is extracted from the linear fit at the mass point 174.5 GeV, corresponding to the +2 GeV uncertainty on the simulated top mass.

The eight $t\bar{t}$ templates obtained by varying the top mass input to simulation are shown in Figure 13.5. The peak of the template shape moves to higher values as the top mass used for the simulation increases. The $t\bar{t}$ selection efficiency for each of the simulated samples is shown on the right in Figure 13.5. The efficiency shows a linear dependence on the top mass used to simulate events. The selection efficiency from the nominal mass sample of 172.5 GeV is slightly out of line as the systematic samples were generated separately with small changes to the generator configuration.

The cross sections measured with the eight mass samples and the nominal sample are shown in Figure 13.6 for fits to data, left, and pseudo-experiments, right. The behaviour of the cross section measured with higher and lower mass samples differs due to the behaviour in the $t\bar{t}$ template shape. As the top mass decreases the peak of the template shape shifts to lower values, with little effect on the measured cross section. As the top mass increases the peak of the template shape shifts to higher values, resulting in a decrease in the measured $t\bar{t}$ cross section. Therefore, separate linear fits are performed for each region. The uncertainty shown for the fits to data is the statistical uncertainty on the $t\bar{t}$ template shape, measured as described in Section 13.12 with the $t\bar{t}$ template Poisson fluctuated within its statistical uncertainties.

The systematic uncertainty on the measured cross section due to the uncertainty on the top quark mass is large. The systematic uncertainty is remeasured with only the $t\bar{t}$ template shape or $\varepsilon_{t\bar{t}}$ varied to determine which effect is dominant. Figure 13.7 shows the cross sections measured from fits to data with only the template shape varied, left, and only $\varepsilon_{t\bar{t}}$ varied, right. The slope of the linear fit is more significant when $\varepsilon_{t\bar{t}}$ is varied, therefore the overall uncertainty is mainly due the uncertainty on $\varepsilon_{t\bar{t}}$.

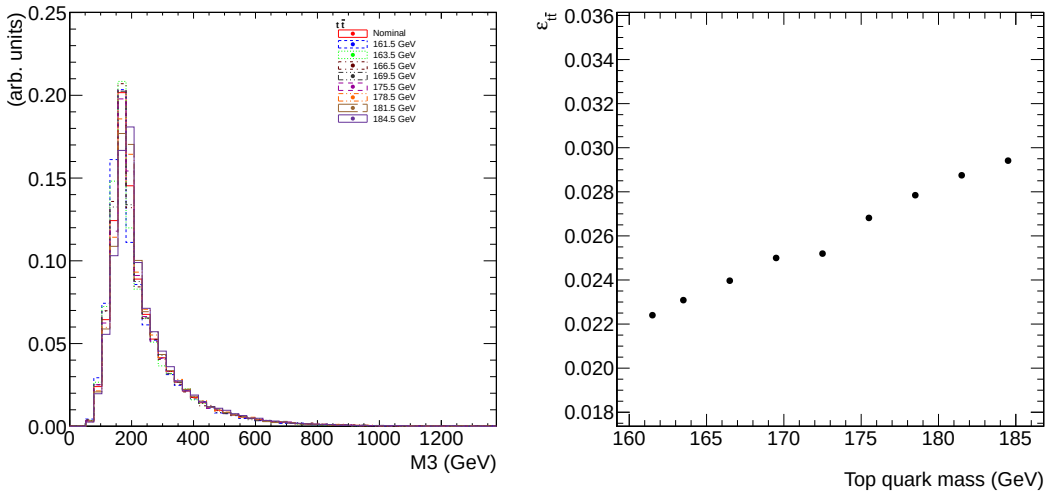


Figure 13.5: $t\bar{t}$ templates for samples simulated with a different top quark masses, left. $t\bar{t}$ selection efficiency obtained by applying the event selection to samples simulated with different top quark masses, right.

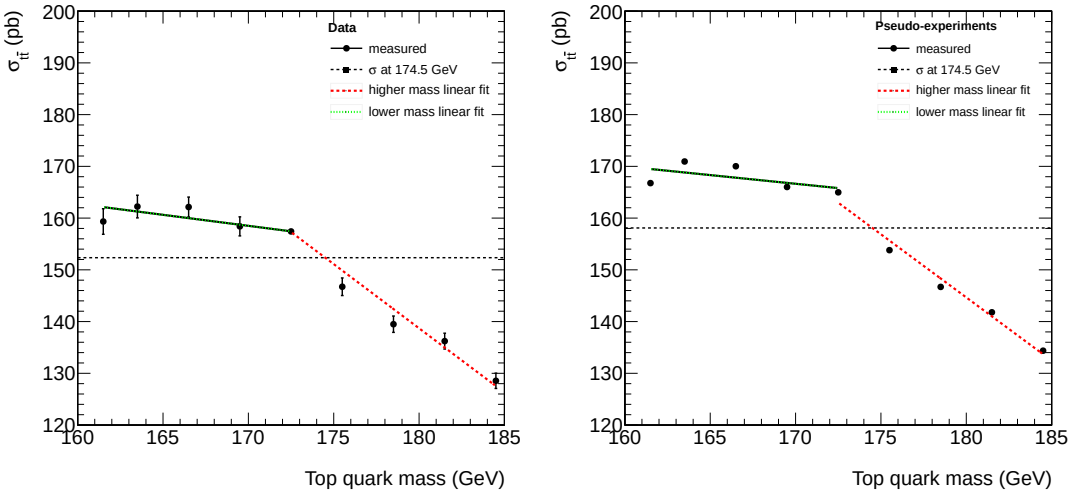


Figure 13.6: Cross sections measured with separate top mass samples in fits to data and pseudo-experiments. The uncertainty shown is the statistical uncertainty on the $t\bar{t}$ template shape. The horizontal line indicates the cross section estimated for a top mass of 174.5 GeV from the linear fit.

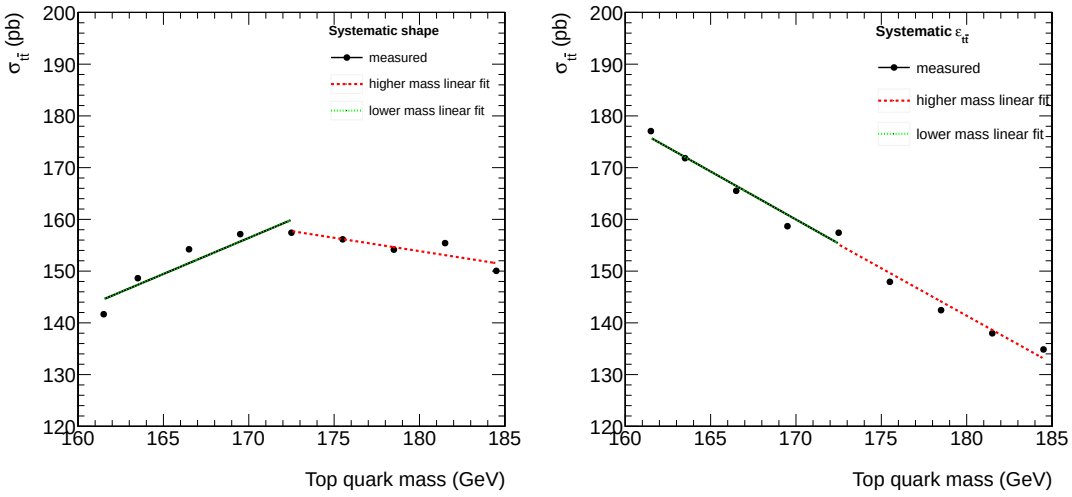


Figure 13.7: Cross sections measured with separate top mass samples in fits to data with only the $t\bar{t}$ template shape, left, or selection efficiency, right, varied.

13.6 Pile-up

An estimate of the number of interactions per event in data is included in the simulated events. A distribution representing the mean number of interactions expected in data is input to the event generation. For each event, the mean number of interactions per bunch crossing is chosen from the input distribution. This is referred to as the true number of interactions for this event and sets the instantaneous luminosity to be simulated for all bunch crossings in the event. The number of interactions which will be part of the event is then randomly sampled, for each bunch crossing, from a Poisson distribution with a mean set to the true number of interactions.

The pile-up distribution in data evolves continuously as the luminosity delivered by the LHC increases. Therefore, the pile-up distribution in simulation is reweighted to correspond to the distribution in the analysed data. The luminosity information for each luminosity section is used to calculate a pile-up distribution for the data. The distribution of the the true number of interactions in simulation is then reweighted to the distribution in data.

Figure 13.8 shows the true pile-up distributions in data and simulation, before pile-up reweighting is applied to the simulated sample. Pile-up reweighting reweights the simulated true pile-up distribution to agree with the data distribution. The number of reconstructed primary vertices after the final event selection is representative of the number of pile-up interactions in an events. Figure 13.9 shows the number of primary

vertices in events passing the signal selection before and after pile-up reweighting is applied. The distribution after pile-up reweighting shows good agreement between simulation and data. In addition to validating the results of the template fit for the $t\bar{t}$ cross section measurement, this indicates that the reweighting of the simulated pile-up distribution to reproduce the pile-up distribution in data has been successful.

The uncertainty on the number of interactions due to pile-up modelling is estimated to be $\pm 5\%$. This uncertainty is propagated to the cross section measurement by repeating the measurement with the variation applied to the true simulated distribution used for reweighting.

The systematic templates obtained by applying the signal selection to the samples simulated with pile-up variations applied are shown in Figures 13.10 and 13.11. The effect of the pile-up variations on the template shapes are small. The uncertainty on the measured $t\bar{t}$ cross section due to pile-up uncertainties is negligible.

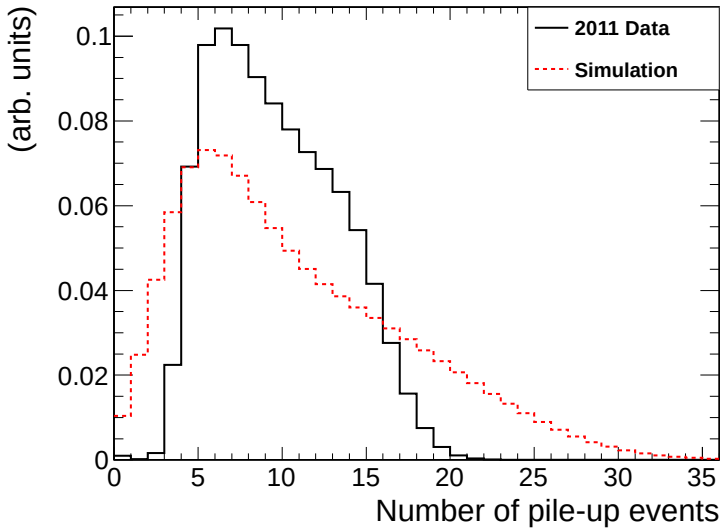


Figure 13.8: True pile-up distributions in simulation and data, before pile-up reweighting is applied to the simulated sample.

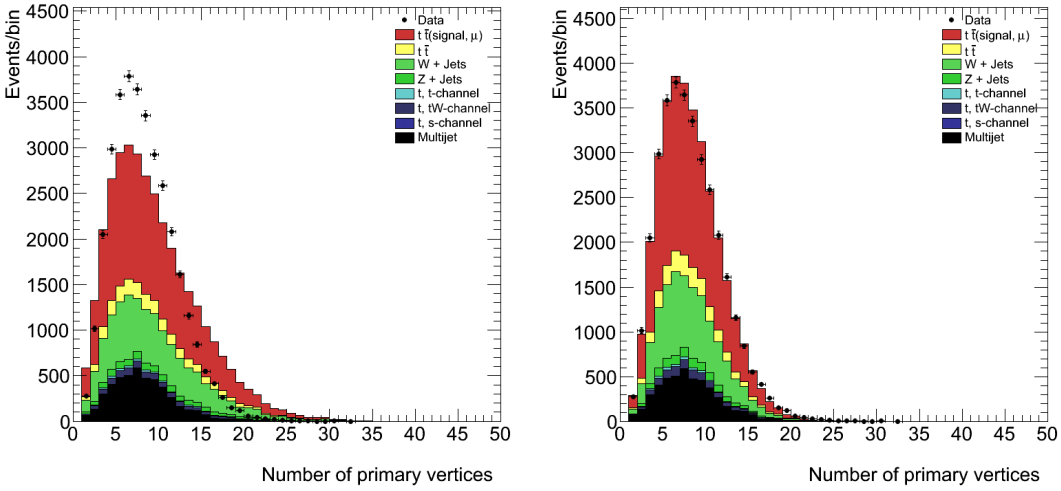


Figure 13.9: Number of primary vertices in selected events before, left, and after, right, pile-up reweighting is applied. The number of events in the simulated samples is normalised to the results of the template fit to data, given in Table12.4.

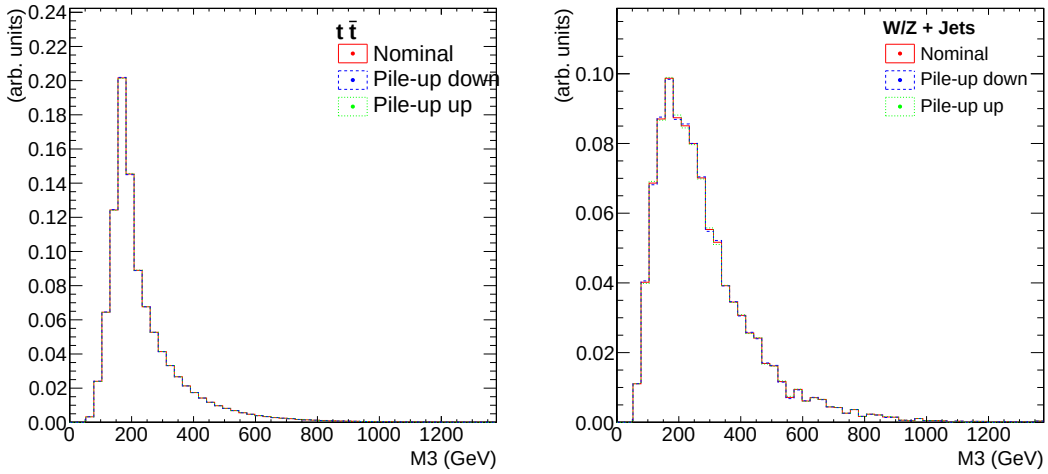


Figure 13.10: $t\bar{t}$ and W/Z + jets templates with pile-up variations applied.

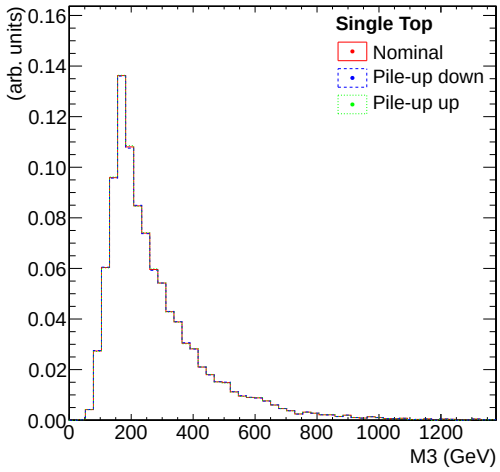


Figure 13.11: Single top templates with pile-up variations applied.

13.7 Jet Energy Scale and Resolution

The dominant contributions to the JES uncertainty are due to pile-up in the event and the flavour dependence of the JES, i.e. the difference in JES between jets from gluons, light and heavy flavour quarks. The overall JES uncertainties measured at $\sqrt{s} = 7$ TeV are shown in Figure 9.3.

The systematic uncertainties on the JES and JER are propagated to the analysis by repeating the cross section measurement with variations applied to the simulated jets. The energy of the jets in simulated events is varied according to the uncertainty on the JES corrections or the JER. The event selection is then applied to the simulated samples with the systematically varied jets.

The systematic JES and JER templates are shown in Figures 13.12, 13.13 and 13.14. The JES uncertainty is the dominant source of systematic uncertainty in the cross section measurement. The JER uncertainty has a negligible effect on the measured cross section.

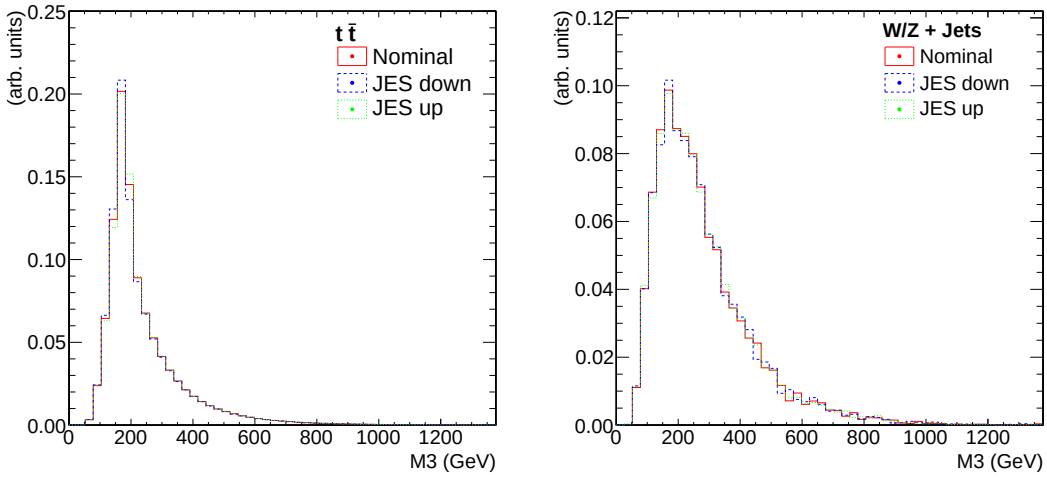


Figure 13.12: $t\bar{t}$ and W/Z + jets templates with JES variations applied.

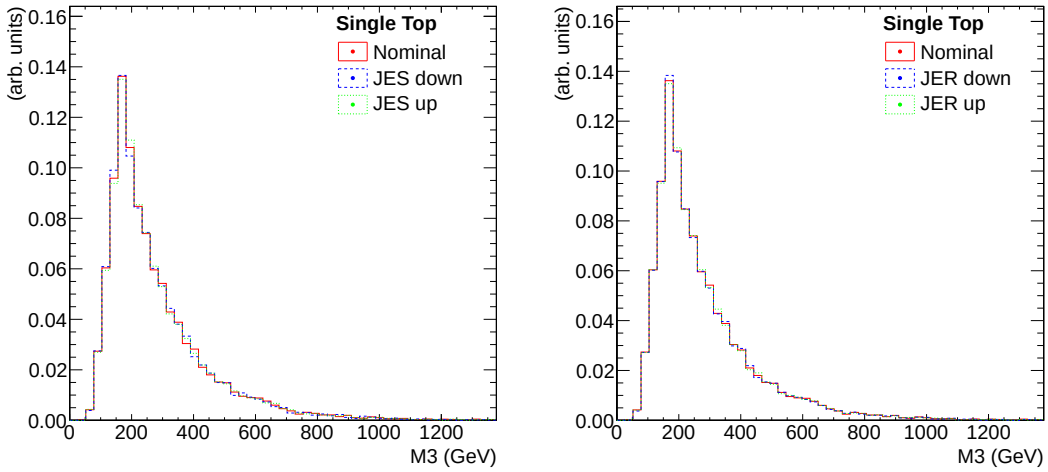


Figure 13.13: Single top template with JES, left, and JER, right, variations applied.

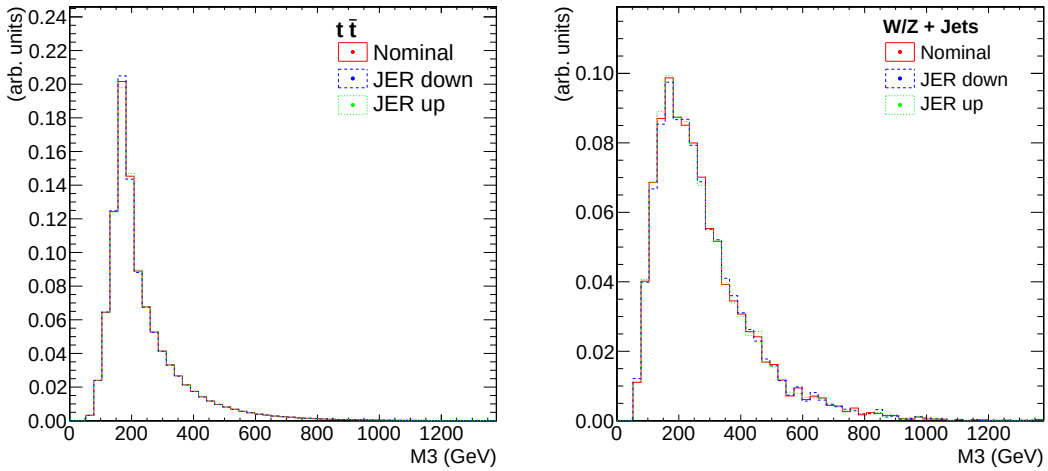


Figure 13.14: $t\bar{t}$ and W/Z + jets templates with JER variations applied.

13.8 Muon Selection Efficiency

Scale factors for the muon identification and selection efficiencies are measured in Chapter 11. The systematic uncertainty on the measured scale factors is estimated to be 0.015 for the muon ID scale factor and 0.01 for the muon trigger scale factor. The systematic deviation in the measured cross section due to these uncertainties is quantified by varying the selection efficiency in Equation 13.1 by the scale factor uncertainties.

The uncertainties on the muon ID and trigger scale factors have a small effect on the measured $t\bar{t}$ cross section.

13.9 Ratio of W + jets to Z + jets Events

The W/Z + jets template shape is constructed from the simulated W + jets and Z + jets samples. While the W/Z + jets normalisation is unconstrained in the template fit, an assumption is made on the relative contribution from W + jets and Z + jets events in the construction on the W/Z + jets template. The ratio of W + jets to Z + jets events is provided by the respective simulated selection efficiencies and the cross sections in Table 4.1.

The W + jets and Z + jets template shapes are expected to be comparable, therefore the assumption on the ratio is not expected to have a significant effect on the measured $t\bar{t}$ cross section. However, to account for any possible dependence the fit is repeated with the ratio of W + jets to Z + jets varied by $\pm 20\%$. This 20% is more than four times the uncertainty on the predicted W + jets and Z + jets cross sections. It is intended as a conservative estimate of the uncertainty to illustrate the stability of the measured $t\bar{t}$ cross section with respect to the W + jets to Z + jets ratio.

The systematic templates obtained by varying the ratio of W + jets to Z + jets events are shown in Figures 13.15. It is evident that varying the ratio has a very small effect on the template shapes. Varying the ratio of W + jets to Z + jets events is found to have a negligible effect on the measured $t\bar{t}$ cross section.

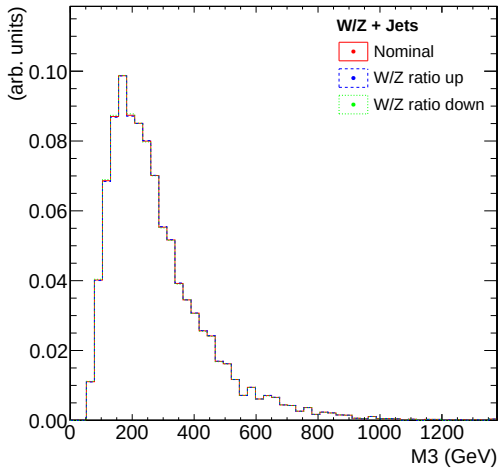


Figure 13.15: W/Z + jets templates with variations applied to the relative fraction of W + jets and Z + jets events.

13.10 Number of Single Top Events

The normalisation of the single top template is constrained by a Gaussian to within 30% of the expectation, as described in Chapter 12, due to the similarity between the single top and $t\bar{t}$ template shapes. However, due to the high correlation between the signal and single top template shapes, the pseudo-experiments result in a distribution for the number of single top events with $\sigma_{\mathrm{N\,st}}$ much less than 30%.

The effect of the 30% single top uncertainty is estimated by repeating the fit to data 5000 times. Before each fit the mean of the Gaussian constraint, $\mathrm{N}_{\mathrm{Sim\,st}}$ in Equation 12.1, is randomly sampled from a Gaussian distribution with a mean value at the expected number of single top events and a σ of 30%. The σ of the Gaussian constraint, Δ_{st}, remains at 30% for each fit to data. This method is also implemented in pseudo-experiments and is recommended to ensure the correct calculation of the single top pull distribution [73].

This method results in a Gaussian distribution for the number of fitted single top events with $\sigma_{\mathrm{N\,st}}$ consistent with 30%. The distribution for the number of $t\bar{t}$ events is also a Gaussian, with a spread which is a direct result of the variation in the number of single top events. The estimation of the single top systematic uncertainty uses Equation 13.2 to estimate $\sigma_{t\bar{t}}^{\mathrm{systematic}}$ for each pseudo-esperiment. The systematic uncertainty on the measured $t\bar{t}$ cross section is then given by the spread of the Gaussian distribution for $\sigma_{t\bar{t}}^{\mathrm{systematic}}$. The systematic uncertainty due to a 30% uncertainty on the number of

single top events has a small effect on the measured $t\bar{t}$ cross section.

13.11 Multijet Template Shape

13.11.1 Contamination Removal

Inaccurate contamination removal or a relative isolation dependence would cause a systematic shift in the estimate of the multijet template shape. The contamination removal assumes the calculated cross sections from Table 4.1. A conservative uncertainty of 50% is applied to the cross sections in the contamination removal. The uncertainty is propagated by varying the cross sections for contamination removal up and down then repeating the $t\bar{t}$ cross section measurement with the multijet template shapes resulting from each variation. The systematic uncertainty due to contamination removal is given by the change in the cross section measured.

The systematic multijet templates are shown on the left in Figure 13.16. The effect of the systematic variation on the template shape is small because the contamination in the selected background region is low. The uncertainty on the multijet contamination removal is a relatively small contributor to the overall systematic uncertainty on the measured $t\bar{t}$ cross section.

As a cross check the cross section measurement is repeated with the measured $t\bar{t}$ cross section of 157.4 pb assumed in the multijet contamination removal. This results in a measured cross section of 157.1 pb, a shift of -0.3 pb from the nominal cross section measured. The cross section measurement is repeated with 157.1 pb assumed in the multijet contamination removal, again resulting in a measured $t\bar{t}$ cross section of 157.1 pb. This variation of 0.3 pb is well within the systematic uncertainty due to contamination removal of ± 3.5 pb.

13.11.2 Isolation Dependence

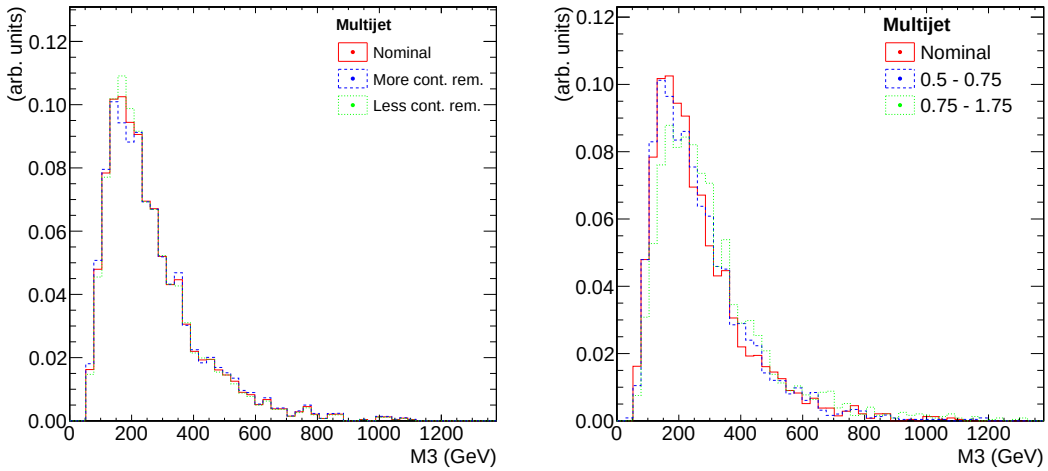


Figure 13.16: Multijet templates with varied contamination removal, left, and in non-overlapping regions of relative isolation, right.

The multijet template shape appears to show a dependence on the relative isolation of the muon in Figure 12.5. To account for this the $t\bar{t}$ cross section measurement is repeated using the multijet template shapes extracted from non-overlapping regions of relative isolation. The regions are chosen to have an equivalent number of events and are described in Section 12.2. The systematic multijet templates are shown on the right in Figure 13.16.

A linear fit is applied to the measured cross sections and extrapolated into the signal relative isolation region. The systematic uncertainty is the difference between the cross section measured with the nominal QCD slice and the cross section in the signal region estimated by the linear fit.

The cross sections measured with the separate multijet templates are shown in Figure 13.17 for fits to data, left, and pseudo-experiments, right. The measured cross section shows a dependence on the relative isolation region from which the multijet template is extracted. When extrapolated to the signal region the linear fit estimates a cross section of 154.3 pb for the fit to data. The difference between this value and the nominal cross section measured is considered as a systematic uncertainty on the cross section measured with the nominal template. This uncertainty is a relatively small contributor to the overall systematic uncertainty on the measured $t\bar{t}$ cross section.

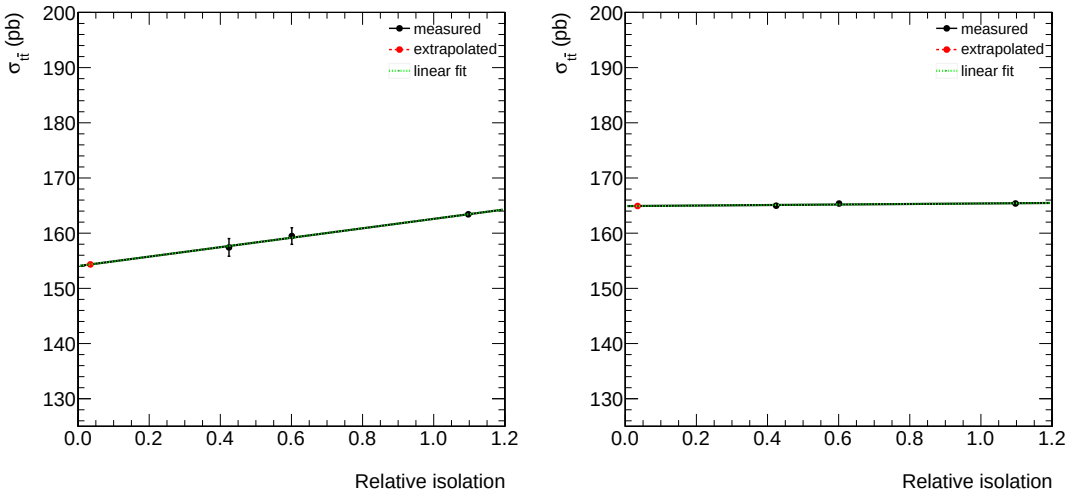


Figure 13.17: Cross sections measured with multijet templates estimated in separate regions of relative isolation in fits to data, left, and pseudo-experiments, right. The cross sections are placed at the weighted average of the relative isolation region. The red point denotes the cross section estimated in the signal relative isolation region by the linear fit. The errors are the statistical uncertainties on the multijet template shapes.

The pseudo-experiments do not reproduce the dependence seen in the fits to data. This is due to the fact that the simulation predicts 328 multijet events, Table 10.7, while the fit the data measures 4659 multijet events in Table 12.3. The expected number of multijet events is used to generate the pseudo-data in the pseudo-experiments. However, since the fit to data predicts significantly more multijet events it is more sensitive to the multijet template shape. The multijet template isolation dependence is remeasured in pseudo-experiments in Figure 13.18. In this case the pseudo-data is generated with an input multijet normalisation of 4659. The measured cross sections show a similar trend to that seen in the fits to data in Figure 13.17.

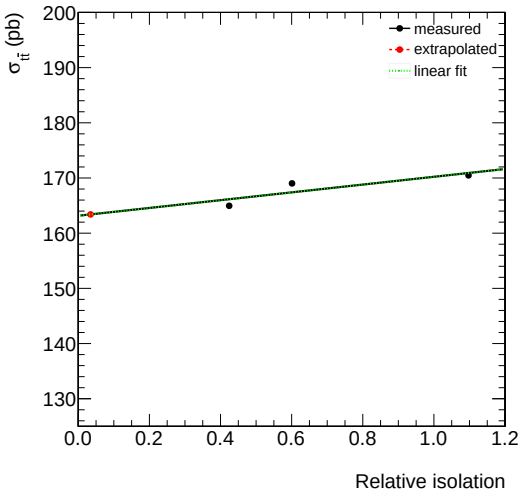


Figure 13.18: Cross sections measured with multijet templates estimated in separate regions of relative isolation in pseudo-experiments generated with increased multijet normalisation. The cross sections are placed at the weighted average of the relative isolation region. The red point denotes the cross section estimated in the signal relative isolation region by the linear fit.

13.12 Template Shape Statistical Uncertainties

The statistical uncertainty on the template shapes is not considered when performing the likelihood fit. To estimate the effect of the template shape statistical uncertainties on the measured cross section the fit to data is repeated 5000 times. Before each fit the template shapes are Poisson fluctuated according to their statistical uncertainties.

The distribution of the number of $t\bar{t}$ events from the 5000 fits is Gaussian. The spread of this distribution is taken as the systematic uncertainty on the measured cross section due to limited template statistics. The uncertainty on the measured $t\bar{t}$ cross section due to the statistical uncertainty in the templates shapes is relatively small.

13.13 Luminosity

There is a 2.2% uncertainty on the luminosity measured with the pixel counting method described in Section 6.7. This uncertainty is propagated to the cross section analysis by varying the integrated luminosity in Equation 1.1 by $\pm 2.2\%$. The uncertainty on the luminosity has a relatively small effect on the measured $t\bar{t}$ cross section.

13.14 Summary of Systematic Uncertainties

The results for the estimation of systematic uncertainties in the cross section measurement are summarised in Table 13.3. The W/Z + jets template shapes for the matching and scale systematic uncertainties, shown in Figures 13.2 and 13.1, show features consistent with a statistically limited simulated sample. The statistical uncertainty on the matching and scale template shapes is obtained by fluctuating the systematic template shapes as described in Section 13.12. The statistical uncertainties are not included in the calculation of the overall systematic uncertainty.

	Uncertainty (pb)			
Systematic	Fit to Data		Pseudo-Exp (Cross check)	
	up	down	up	down
$t\bar{t}$ Scale	$+1.1 \pm 1.2$	$+4.0 \pm 1.2$	$+0.8$	$+4.4$
W/Z + jets Scale	-1.7 ± 1.4	$+3.2 \pm 1.4$	-2.5	$+4.1$
$t\bar{t}$ Matching	-0.5 ± 1.1	-4.2 ± 1.8	-0.6	-2.7
W/Z + jets Matching	-2.8 ± 1.4	$+0.2 \pm 1.6$	-2.4	0
PDF	$+5.8$	-4.8	$+6.8$	-5.9
Top Mass	-5.1	—	-6.9	—
Pile-up	-0.3	$+0.1$	-0.3	$+0.2$
JES	-6.3	$+5.0$	-6.3	$+6.0$
JER	$+0.3$	$+0.6$	-0.4	$+1.8$
ID Efficiency	-2.3	$+2.4$	-2.4	$+2.5$
Trigger Efficiency	-1.6	$+1.6$	-1.6	$+1.7$
W/Z ratio	$+0.2$	-0.2	$+0.2$	-0.3
Single top	$+1.0$	-1.0	incl. in stat.	incl. in stat.
Multijet cont. rem.	$+3.5$	-3.5	$+0.3$	-0.2
Multijet iso. dep.	—	-3.1	—	-0.1
Shape stat.	$+2.5$	-2.5	—	—
Luminosity	-3.4	$+3.5$	-3.6	$+3.7$

Table 13.3: Summary of the systematic uncertainties on the measured cross section.

The systematic template shapes and selection efficiencies for each source of systematic uncertainty are given in Table 13.4. The top mass selection efficiency is determined by applying a linear fit to the selection efficiencies in Figure 13.5 to determine the selection efficiency at a top quark mass of 174.5 GeV. The PDF selection efficiencies are determined by applying Equation 13.3 to the selection efficiencies shown in Figure 13.3.

If the systematic variation affects both the template shape and $t\bar{t}$ selection efficiency, Table 13.4 states which variation dominated the measured systematic uncertainty. This

is determined by repeating the systematic cross section measurements with only the template shape or only the $t\bar{t}$ selection efficiency varied. The results of this study are given in Table 13.5.

The template shape and selection efficiency uncertainties measured for the top mass and PDF systematic variations do not correspond directly to the total systematic uncertainty measured. After measuring the cross section with only the template shape or selection efficiency varied, the top mass uncertainties are estimated by applying a linear fit to the systematic cross sections shown in Figure 13.7. The PDF uncertainties are determined by applying Equation 13.3 to the systematic cross sections.

The uncertainties on the JES corrections and the PDFs are the dominant sources of systematic uncertainty. The overall systematic uncertainty is determined by adding the individual uncertainties in quadrature with the positive and negative uncertainty from each source added separately. If the uncertainty for a particular source is asymmetric, the largest uncertainty is included in the overall calculation. The luminosity uncertainty is not included in the overall systematic uncertainty, instead it is quoted separately in the final result. The overall systematic uncertainty on the measured cross section is $+10.6 - 12.4$ pb. The pseudo-experiments predict an overall systematic uncertainty of $+11.7 - 12.6$ pb.

Therefore, the measured $t\bar{t}$ cross section is

$$157.4 \pm 3.8(\text{stat.})^{+10.6}_{-12.4}(\text{syst.}) \pm 3.5(\text{lumi.})\,\text{pb},$$

with relative uncertainties of $\pm 2.5(\text{stat.})^{+6.7}_{-7.9}(\text{syst.}) \pm 2.2(\text{lumi.})\,\%$. The pseudo-experiments predict a cross section of $165 \pm 3.9(\text{stat.})^{+11.7}_{-12.6}(\text{syst.}) \pm 3.7(\text{lumi.})$ pb, with relative uncertainties of $\pm 2.4(\text{stat.})^{+7.1}_{-7.6}(\text{syst.}) \pm 2.2(\text{lumi.})\,\%$. The expected cross section of 165 ± 10 pb [23] is consistent with the measured cross section within the uncertainties of the measurement.

Systematic	Shapes	$\varepsilon_{t\bar{t}}$		Dominant Unc.	
	(Figure)	up	down	up	down
Nominal	12.1	0.02520		—	—
$t\bar{t}$ Scale	13.1	0.02341	0.02647	ε	shape
W/Z + jets Scale	13.1	—	—	shape	shape
$t\bar{t}$ Matching	13.2	0.02512	0.02632	shape	ε
W/Z + jets Matching	13.2	—	—	shape	shape
PDF	13.3, 13.4	0.02549	0.02485	shape	shape
Top Mass	13.5	0.02622	—	ε	—
Pile-up	13.10, 13.11	0.02523	0.02517	ε	ε
JES	13.12, 13.13	0.02615	0.02409	ε	ε
JER	13.13, 13.14	0.02530	0.02511	shape	ε
ID Efficiency	—	0.02558	0.02482	ε	ε
Trigger Efficiency	—	0.02545	0.02495	ε	ε
W/Z ratio	13.15	—	—	shape	shape
Multijet cont. rem.	13.16	—	—	shape	shape
Multijet iso. dep.	13.16	—	—	shape	shape
Shape stat.	—	—	—	shape	shape

Table 13.4: Summary of the systematic template shape and selection efficiency variations. For each source of systematic uncertainty it is stated if the uncertainty is dominated by the template shape variation or the change in selection efficiency. The systematic uncertainties due to the single top constraint and the luminosity uncertainty are not included as they do not include any template shape or $\varepsilon_{t\bar{t}}$ variation. The top mass selection efficiency is determined by applying a linear fit to the selection efficiencies in Figure 13.5 to determine the selection efficiency at a top quark mass of 174.5 GeV. The PDF selection efficiencies are determined by applying Equation 13.3 to the selection efficiencies shown in Figure 13.3.

Systematic		Uncertainty		
		Total	Shape	$\varepsilon_{t\bar{t}}$
$t\bar{t}$ Scale	up	+1.1	−10.2	+12.1
	down	+4.0	+12.1	−7.6
$t\bar{t}$ Matching	up	−0.5	−1.0	+0.5
	down	−4.2	+2.6	−6.7
PDF	up	+5.8	+7.0	+1.7
	down	−4.8	−5.9	−1.8
Top Mass	up	−5.1	−0.7	−6.0
Pile-up	up	−0.3	−0.1	−0.2
	down	+0.1	−0.1	+0.2
JES	up	−6.3	−0.6	−5.7
	down	+5.0	−2.2	+7.2
JER	up	+0.3	+1.0	−0.6
	down	+0.6	+0.	+0.6

Table 13.5: Breakdown of the systematic uncertainty measured from a fit to data for each systematic variation which affected both the template shapes and the $t\bar{t}$ selection efficiency. The systematic $t\bar{t}$ cross section is remeasured for each source of uncertainty with only the template shapes or only the $t\bar{t}$ selection efficiency varied. The top mass systematic uncertainties are then estimated by applying a linear fit to the systematic cross sections, shown in Figure 13.7. The PDF uncertainties are determined by applying Equation 13.3 to the systematic cross sections.

Chapter 14

Top Quark Mass from Cross Section

In this chapter the pole mass of the top quark is extracted from the measured $t\bar{t}$ cross section. This is achieved by comparing the measured cross section to approximate NNLO calculations which use the pole quark mass definition.

Direct measurements of the top mass use information from simulation. Therefore the measured top mass depends on the top mass definition used in the simulation. The top mass depends on the renormalisation scheme, and the results of any direct measurement must be interpreted in terms of the renormalisation conventions [76]. The inability of event generators to fix the renormalisation scheme results in an uncertainty on the top mass definition used in simulation, which affects the measured top mass.

In this chapter the top mass is determined by comparing the measured $t\bar{t}$ cross section with approximated NNLO calculations which include an unambiguous definition of the top pole mass. This provides a measurement of the top mass which is complementary to direct measurements, due to the different sensitivity to systematic uncertainties. It is also a test of the mass scheme used in simulation.

The approximate NNLO calculations for the cross section are provided by Kidonakis [77] for the pole mass definition of the top mass. The uncertainties on the calculations are due to uncertainties in the determination of the PDFs, renormalisation and scale variations, and variations to the strong coupling constant.

The measured cross section dependence on the top mass is determined by repeating the cross section measurement with $t\bar{t}$ samples simulated with different top mass values, as described in Section 13.5. The uncertainty band on the measured cross section is given by the relative overall uncertainty on the nominal cross section measurement, excluding the top mass systematic uncertainty. This results in an absolute cross section uncertainty of $\pm 3.8(\text{stat.})^{+10.6}_{-11.3}(\text{syst.}) \pm 3.5(\text{lumi.})\,\text{pb}$, which corresponds to a relative uncertainty of $+7.5 - 7.9\%$.

The top pole mass and uncertainty are extracted by maximising the joint likelihood

$$L(\mathrm{m}_t) = \int f_{\text{exp}}(\sigma_{t\bar{t}}|\mathrm{m}_t) f_{\text{th}}(\sigma_{t\bar{t}}|\mathrm{m}_t) d\sigma_{t\bar{t}}. \tag{14.1}$$

The probability density functions $f_{\text{exp}}(\sigma_{t\bar{t}}|\mathrm{m}_t)$ and $f_{\text{th}}(\sigma_{t\bar{t}}|\mathrm{m}_t)$ are constructed from Gaussian distributions with a mean at the measured and predicted cross sections respec-

tively. The width of the distributions are given by the uncertainties on their respective means.

The mass dependence is parameterised with a third order polynomial

$$\sigma_{t\bar{t}}(\mathrm{m}_t) = \frac{1}{\mathrm{m}_t^4}(a + b \cdot \mathrm{m}_t + c \cdot \mathrm{m}_t^2 + d \cdot \mathrm{m}_t^3). \tag{14.2}$$

The top quark pole mass extracted from the fit to the cross section dependence measured in data is

$$178.2^{+6.5}_{-7.4}\,\mathrm{GeV},$$

at a $t\bar{t}$ cross section of 143.3 pb.

The results are shown compared to the measured and predicted cross section dependences in Figure 14.1. The results are compared with the results of similar studies performed by the D0, CMS and ATLAS collaborations [78, 22, 79] in Figure 14.2. The top pole mass extracted from the cross section measured in this thesis is in agreement with the results from the other experiments, within the uncertainties.

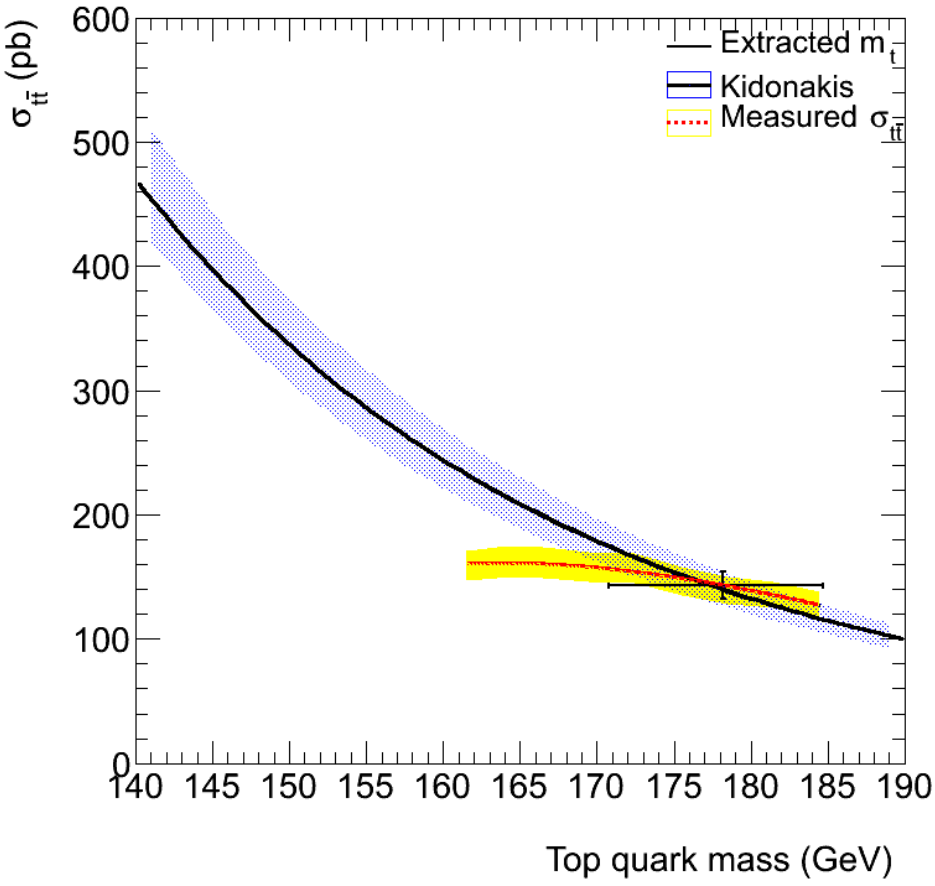


Figure 14.1: Measured and predicted cross section dependence on the top quark mass. The top quark pole mass extracted from the joint likelihood fit is shown at $\sigma_{t\bar{t}} = 143.3\,\text{pb}$. The blue uncertainty band on the theoretical cross section is due to uncertainties in the determination of the PDFs, renormalisation and scale variations, and variations to the strong coupling constant. The yellow uncertainty band on the measured cross section is given by the 7% overall uncertainty on the nominal cross section measured, excluding the top mass systematic uncertainty.

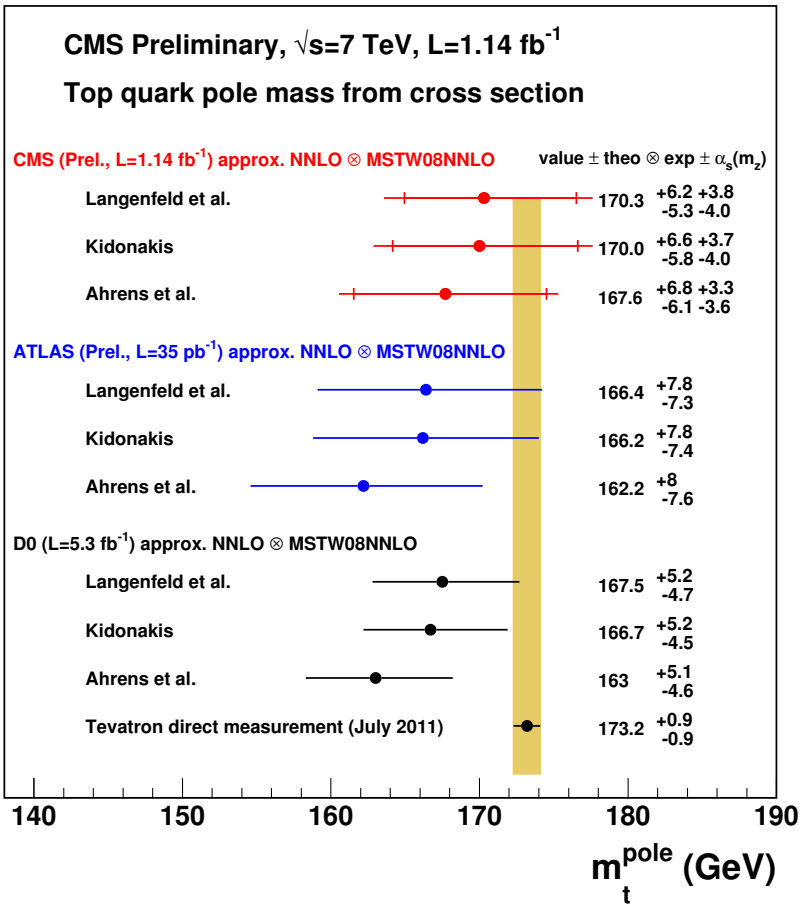


Figure 14.2: Top quark pole mass extracted from $t\bar{t}$ cross sections measured in D0, CMS and ATLAS [78, 22, 79]. The 2011 world average for the direct top quark mass measurement is also shown. The top mass extracted from the cross section measured in this thesis is $178.2^{+6.5}_{-7.4}$ GeV.

Chapter 15

Comparison with results from CMS and ATLAS

In this section the cross section measured in this thesis is compared with results for the measurement of the $t\bar{t}$ cross section published by the CMS and ATLAS collaborations. The results for the $t\bar{t}$ cross section measured in the lepton + jets channel, where the lepton is an electron or a muon, at $\sqrt{s} = 7$ TeV are shown in Table 15.1. The cross section measured in this thesis agrees within its uncertainties with the cross sections measured by CMS and ATLAS, and shows comparable sensitivity.

Analysis	Channel	Lumi (fb^{-1})	Cross Section (pb)
Thesis	μ+jets	4.8	$157.4 \pm 3.8(\text{stat.})^{+10.6}_{-12.4}(\text{syst.}) \pm 3.5(\text{lumi.})$
CMS 2011 [80]	μ+jets	1.1	$163.2 \pm 3.4(\text{stat.}) \pm 12.7(\text{syst.}) \pm 7.3(\text{lumi.})$
	e/μ+jets	1.1	$164.4 \pm 2.8(\text{stat.}) \pm 11.9(\text{syst.}) \pm 7.4(\text{lumi.})$
ATLAS 2011 [81]	e/μ+jets	0.7	$179 \pm 9.8(\text{stat} + \text{syst}) \pm 6.6(\text{lumi.})$
CMS 2010 [82]	e/μ+jets	0.04	$150 \pm 9(\text{stat.}) \pm 17(\text{syst.}) \pm 6(\text{lumi.})$
ATLAS 2010 [83]	e/μ+jets	0.04	$186 \pm 10(\text{stat.})^{+21}_{-20}(\text{syst.}) \pm 6(\text{lumi.})$
CMS 2011 [84]	Combi.	0.8-1.1	$165.8 \pm 2.2(\text{stat.}) \pm 10.6(\text{syst.}) \pm 7.8(\text{lumi.})$
ATLAS 2011 [85]	Combi.	0.7-1.0	$177 \pm 3(\text{stat.})^{+8}_{-7}(\text{syst.}) \pm 7(\text{lumi.})$

Table 15.1: Comparison of $t\bar{t}$ cross sections measured by the CMS and ATLAS collaborations. The combi. measurement is a combination of the $t\bar{t}$ cross section measurements in the fully hadronic, e/μ + jets, ee, $\mu\mu$ and $e\mu$ channels. The $\mu\tau$ channel is also included in the combined CMS cross section. In both CMS and ATLAS the combination is performed with a likelihood fit.

The CMS e+jets and μ+jets cross sections are extracted with a simultaneous binned likelihood fit to the secondary vertex mass distribution for different jet multiplicities and number of b-tagged jets. The 2011 CMS analysis has a similar event selection to

this thesis with two significant differences: a lower $p_{\mathrm{T}} > 35$ GeV threshold on the muon, as the non-isolated HLT_Mu30 trigger is used which remained unprescaled in the early 2011 data taking period, and a minimum requirement on the amount of missing energy in the event, $\not\!\!E_T > 20$ GeV.

The 2011 ATLAS analysis extracts the cross section with a simultaneous likelihood fit to a likelihood discriminant distribution for events with three, four and at least five jets. The 2010 ATLAS analysis makes use of b-tagging and extracts the cross section with a profile likelihood fit to a discriminant built from several kinematic variables. The analysis presented in this thesis achieves comparable or better sensitivity than previous cross section measurements without the use of b-tagging or multivariate techniques.

The cross sections measured with data collected in 2011 are compared with the predicted cross section in Figure 15.1. In all cases the expected cross section overlaps with the measured cross sections within the uncertainties.

The cross section measured in this thesis is compared with the cross sections measured by the CDF and D0 collaborations at $\sqrt{s} = 1.96$ at the Tevatron in Figure 15.2. The SM predicted cross sections remain consistent with the measured cross sections over a wide range of center of mass energies.

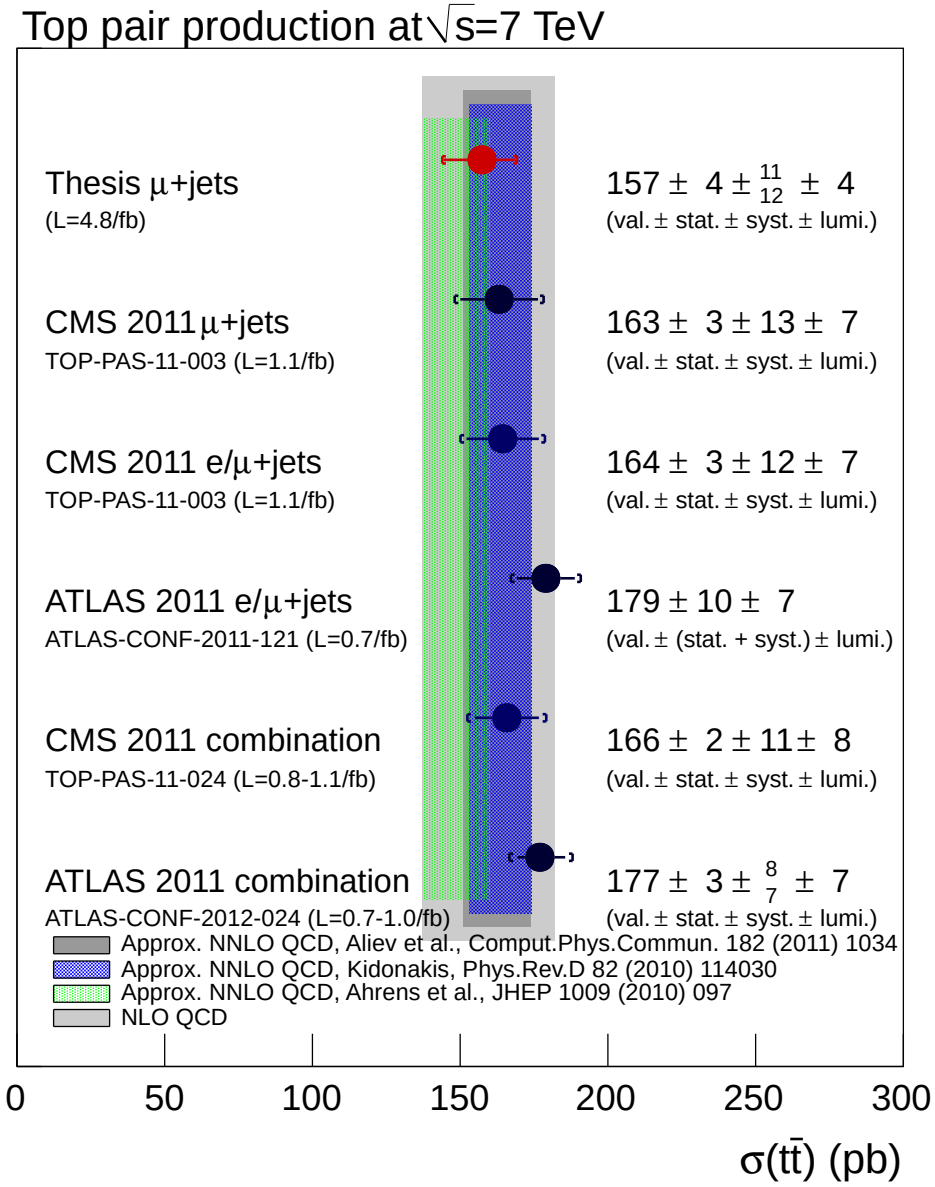


Figure 15.1: $t\bar{t}$ cross sections measured by the CMS and ATLAS collaborations compared with SM expectations.

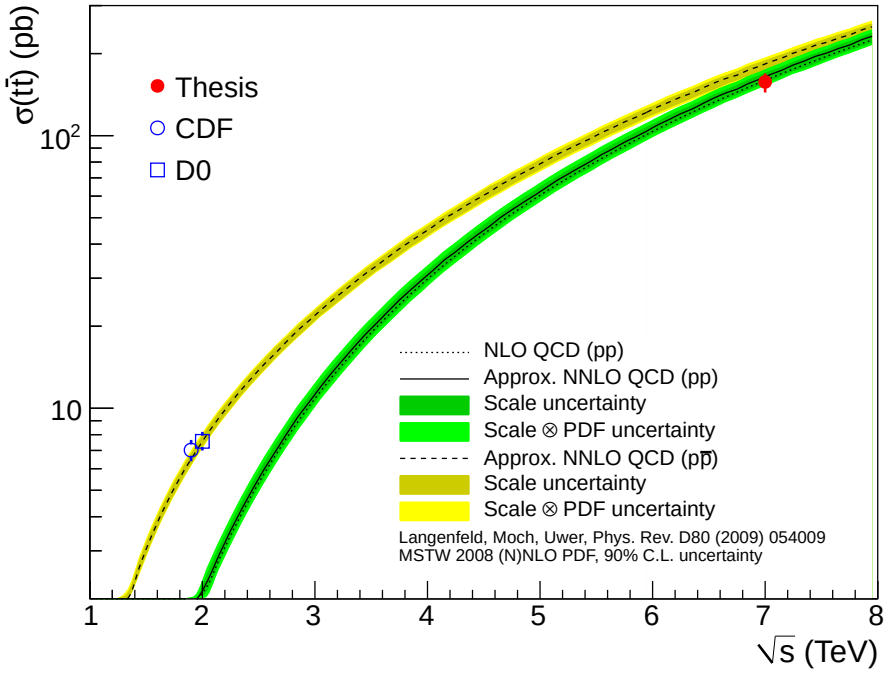


Figure 15.2: Measured $t\bar{t}$ cross sections as a function of center of mass energy. The CDF and D0 results are combinations of cross section measurements in the lepton + jets and dilepton channels [86, 87].

Chapter 16

Conclusion

The top quark pair production cross section has been measured at $\sqrt{s} = 7$ TeV with 4.76 fb^{-1} of data collected by the CMS detector at the LHC in 2011. The cross section was estimated with events in the muon + jets $t\bar{t}$ final state. The muon trigger and selection efficiencies were measured in data with the tag and probe method.

A template fitting method was used to extract the cross section. The method involves estimating template shapes for each of the physics processes expected to contribute to the selected data sample: $t\bar{t}$, W/Z + jets, multijet and single top. The multijet template shape is derived from data by selecting events in a background region dominated by multijet events with an inverted relative isolation requirement. The remaining template shapes are provided by simulation. The $t\bar{t}$ cross section is then extracted by performing a binned likelihood fit of the templates to the distribution in the selected data sample to determine the contribution from each physics process to the data sample. No assumption is made on the normalisation of the $t\bar{t}$, W/Z + jets or multijet templates in the fit.

The systematic uncertainty on the measured cross section is determined by repeating the fit to data with systematic variations applied to the template shapes or selection efficiencies for the expected sources of systematic uncertainty. The measurement is repeated with pseudo-experiments, using pseudo-data generated from simulated events, to ensure the behaviour of the fit is understood.

The measured cross section is

$$157.4 \pm 3.8(\text{stat.})^{+10.6}_{-12.4}(\text{syst.}) \pm 3.5(\text{lumi.})\,\text{pb}.$$

The expected cross section of 165 ± 10 pb [23] is consistent with the measured cross section within the uncertainties of the measurement, supporting the validity of the SM. The precisely measured $t\bar{t}$ cross section is useful for analyses measuring properties of the top quark. An accurate cross section estimate also facilitates the control of top quark background contributions in searches for new physics processes.

The uncertainty on the measured $t\bar{t}$ cross section is dominated by systematic uncertainties, with the most significant contribution coming from JES and PDF uncertainties. In future cross section measurements, improved understanding of the CMS detector should reduce the JES systematic uncertainty. Efforts to improve the precision of PDF estimates are ongoing and should further improve the cross section measurement.

The systematic uncertainty on the measured cross section would be further reduced by quoting the systematic uncertainties relative to the cross section measured at a top quark mass of 173.5 GeV, instead of 172.5 GeV, as $173.5 \pm 0.6 \pm 0.8$ GeV is the current most precise measurement for the top quark mass. The estimate of the systematic uncertainties due to the matching and scale variations in the simulation would be improved by increasing the number of events in the simulated samples, to avoid statistical features in the systematic template shapes.

Improvements can also be made to the method presented in this thesis to increase the accuracy of the measured $t\bar{t}$ cross section. The HLT_Mu40 trigger was chosen to avoid a trigger with an isolation requirement, however the consequent $p_{\mathrm{T}} > 42$ GeV requirement on muons selected in the analysis reduces the $t\bar{t}$ selection efficiency significantly. An alternative would be to use a trigger which requires at least one muon, without isolation requirements, and at least three jets. The additional jets requirement allows for a lower p_{T} threshold on the muon without reducing the number of events available for studies of the multijet template shape.

Alternative methods for the estimation of the multijet template shape should be studied to reduce or remove the systematic uncertainties due to contamination removal and isolation dependence. One such alternative is the matrix method, described in [88]. The isolation dependence of the multijet template shape estimation may also be reduced by using a different kinematic distribution in the cross section extraction.

A potential source of systematic uncertainty which has not been considered in this analysis is the uncertainty due to the choice of event generator. In this analysis the $t\bar{t}$ simulation is generated with MadGraph and PYTHIA. The cross section measurement should be repeated using $t\bar{t}$ events simulated with POWHEG and PYTHIA or MC@NLO [89] and HERWIG [90] to study the effect of the event generator choice on the measured cross section.

The LHC center of mass energy increased to $\sqrt{s} = 8$ TeV in 2012, providing a new opportunity to test the SM. With further development, the method presented in this thesis can continue to provide precision measurements of the top quark pair production cross section at the LHC.

Chapter 17

Samenvatting

De werkzame doorsnede voor productie van topquark paren werd gemeten bij $\sqrt{s} = 7$ TeV met 4.76 fb$_{-1}$ aan gegevens verzameld door de CMS detector aan de LHC versneller in 2011. De werkzame doorsnede werd geschat op basis van evenementen in de $t\bar{t}$ finale toestand met muonen en jets. De muon trigger en selectie efficiënties werden bepaald uit de gegevens aan de hand van de "tag and probe" methode.

Een fit methode op basis van een sjabloon werd gebruikt om de werkzame doorsnede af te leiden. Deze methode vereist het schatten van de vorm van een sjabloon voor elk van de fysische processen waarvan verwacht kan worden dat ze bijdragen tot de geselecteerde evenementen: $t\bar{t}$, W/Z+jets, multijet en "single top". De vorm van het multijet sjabloon werd afgeleid uit gegevens waarbij evenementen geselecteerd werden in een gebied waar de achtergrond gedomineerd wordt door multijet evenementen met omgekeerde vereisten wat betreft de relatieve isolatie. De overige sjabloonvormen werden uit simulaties bepaald. De $t\bar{t}$ werkzame doorsnede werd dan bepaald door een gebinde likelihood fit van de sjablonen aan de distributie in het geselecteerde sample uit te voeren, en zo de bijdrage van elk fysisch proces te bepalen. Er werd geen enkele vooronderstelling gemaakt over de normering van de $t\bar{t}$, W/Z+jets of multijet sjablonen in de fit.

De systematische fout op de gemeten werkzame doorsnede werd bepaald door de fit aan de gegevens te herhalen met systematische variaties op de sjabloon vormen of selectie efficiënties voor de verwachte bronnen van systematische fout. De meting wordt dan herhaald met pseudo-experimenten, gebruik makende van pseudo-gegevens gegenereerd met gesimuleerde evenementen, om zeker te zijn dat het gedrag van de fit goed begrepen wordt.

De gemeten werkzame doorsnede bedraagt:

$$157.4 \pm 3.8(\text{stat.})^{+10.6}_{-12.4}(\text{syst.}) \pm 3.5(\text{lumi.}) \text{ pb}$$

De onzekerheid op de gemeten $t\bar{t}$ werkzame doorsnede wordt gedomineerd door systematische fouten, waarbij de belangrijkste bijdrage komt van de energieschaal van de jets en onzekerheden op de PDF. De verwachte werkzame doorsnede van 165$\pm$10 pb is consistent met de gemeten waarde binnen de onzekerheden van het experiment, hetgeen de juistheid van het Standaard Model ondersteunt. Deze precieze waarde voor de $t\bar{t}$ werkzame doorsnede is belangrijk voor analyses van de eigenschappen van de top quark.

Een nauwkeurige schatting van de werkzame doorsnede vergemakkelijkt ook de controle van de top quark achtergrond in zoektochten naar processen te wijten aan nieuwe fysica.

Part VI

Appendices

Appendix A

Technical Information

This appendix provides technical information about the samples, software version and global tags used in this analysis. This information is relevant for CMS personnel who wish to understand the CMS specific details of the analysis.

The technical details of the samples used in this analysis may be accessed via the DAS website [91]. The DAS identifiers for the nominal simulated samples used for the cross section extraction are given in Table A.1. The identifiers for the simulated samples with matching and scale systematic variations applied are given in Table A.2. Table A.3 lists the samples simulated with varied top quark masses for the estimation of the top mass systematic uncertainty.

The DAS identifiers for the data samples are given in Table A.4. The corresponding JSON files, available at [92], are given in Table A.5. CMSSW version 4.2 is used in this analysis with the global tag START42_V17 for simulation and GR_R_42_V25 for data.

DAS identifier
/TTJets_TuneZ2_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v2/AODSIM
/WJetsToLNu_TuneZ2_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM
/DYJetsToLL_TuneZ2_M-50_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM
/T_TuneZ2_t-channel_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM
/Tbar_TuneZ2_t-channel_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM
/T_TuneZ2_s-channel_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM
/Tbar_TuneZ2_s-channel_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM
/T_TuneZ2_tW-channel-DR_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM
/Tbar_TuneZ2_tW-channel-DR_7TeV-powheg-tauola/Fall11-PU_S6_START42_V14B-v2/AODSIM
/QCD_Pt-20_MuEnrichedPt-15_TuneZ2_7TeV-pythia6/Fall11-PU_S6_START42_V14B-v1/AODSIM

Table A.1: Nominal simulated samples for physics process relevant to the measurement of the $t\bar{t}$ cross section in the muon + jets channel. The physics processes are described in Chapter 4. The simulated samples are described in Chapter 8.

DAS identifier
/TTjets_TuneZ2_matchingup_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v2/AODSIM
/TTjets_TuneZ2_matchingdown_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v2/AODSIM
/WJetsToLNu_TuneZ2_matchingup_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM
/WJetsToLNu_TuneZ2_matchingdown_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM
/ZJetsToLL_TuneZ2_matchingup_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM
/ZJetsToLL_TuneZ2_matchingdown_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM
/TTjets_TuneZ2_scaleup_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM
/TTjets_TuneZ2_scaledown_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v2/AODSIM
/WJetsToLNu_TuneZ2_scaleup_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM
/WJetsToLNu_TuneZ2_scaledown_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM
/ZJetsToLL_TuneZ2_scaleup_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM
/ZJetsToLL_TuneZ2_scaledown_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v1/AODSIM

Table A.2: Simulated samples for matching and scale systematic studies, described in Chapter 13.

DAS identifier
/TTJets_TuneZ2_mass161_5_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v3/AODSIM
/TTJets_TuneZ2_mass163_5_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v3/AODSIM
/TTJets_TuneZ2_mass166_5_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v3/AODSIM
/TTJets_TuneZ2_mass169_5_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v3/AODSIM
/TTJets_TuneZ2_mass175_5_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v3/AODSIM
/TTJets_TuneZ2_mass178_5_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v3/AODSIM
/TTJets_TuneZ2_mass181_5_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v3/AODSIM
/TTJets_TuneZ2_mass184_5_7TeV-madgraph-tauola/Fall11-PU_S6_START42_V14B-v3/AODSIM

Table A.3: Simulated samples for top quark mass systematic studies, described in Chapter 13.

DAS identifier
/SingleMu/Run2011A-05Aug2011-v1/AOD
/SingleMu/Run2011A-May10ReReco-v1/AOD
/SingleMu/Run2011A-PromptReco-v4/AOD
/SingleMu/Run2011A-PromptReco-v6/AOD
/SingleMu/Run2011B-PromptReco-v1/AOD

Table A.4: 2011 samples of data collected with a trigger requiring at least one muon.

JSON files
Cert_160404-180252_7TeV_PromptReco_Collisions11_JSON.txt
Cert_170249-172619_7TeV_ReReco5Aug_Collisions11_JSON_v2.txt
Cert_160404-163869_7TeV_May10ReReco_Collisions11_JSON_v3.txt

Table A.5: JSON files with the certified runs and luminosity sections used in this analysis.

Appendix B

HLT_IsoMu24

HLT_IsoMu24 is a single muon trigger with a p_{T} threshold of 24 GeV and an absolute isolation requirement which does not vary relative to the p_{T} of the muon. The isolation requirement ensured the trigger rate remained within the limitations of the HLT bandwidth during 2011 data taking, despite the rather low p_{T} threshold of the trigger. This trigger is suitable for analyses studying muons from W or Z boson decay. The low p_{T} threshold allows for increased signal event retention while the isolation requirement has little effect on the efficiency for such muons to trigger, since they are expected to be isolated. Despite this advantage, the HLT_IsoMu24 trigger is not used in this analysis. It was initially foreseen to use HLT_IsoMu24, however it turned out that the isolation requirement limits the number of events available in the multijet sideband region required for the extraction of the multijet template shape in Section 12.2. Thus the HLT_Mu40 trigger is used to improve the reliability of the multijet template shape extraction.

An overview of the HLT_IsoMu24 trigger in 2011 data taking is given in this appendix. A muon selection optimised for analyses using the HLT_IsoMu24 trigger is described, and the ID and trigger efficiencies and scale factors measured.

B.1 HLT_IsoMu24 in 2011

The HLT_IsoMu24 trigger is implemented in the same way as the HLT_Mu40 trigger, the only differences are the isolation and p_{T} requirements. Therefore the changes applied to HLT_Mu40, described in Section 7.4, which were not related to isolation or p_{T} , were also applied to the HLT_IsoMu24 trigger. The remaining changes are described in this section.

Starting from the 1e33 menu, the L1 seeds for HLT_IsoMu24 are required to pass the L1_SingleMu12 trigger, with no isolation requirement and a 12 GeV p_{T} threshold. At HLT, η dependent detector based isolation criteria are applied to the muon tracks. Selected muon tracks require a calorimeter isolation less than 4.0 within a cone of radius 0.24 and a tracker isolation less than 1.2 within a cone of radius 0.24.

In the 3e33 menu, a prescale is applied to the trigger. A duplicate trigger with an additional η requirement is introduced which remains unprescaled, as was done with HLT_Mu40. The HLT_IsoMu24_eta2p1 trigger is used to collect data starting

from this menu. The L1 seeds for HLT_IsoMu24_eta2p1 are required to pass the L1_SingleMu14_eta2p1 trigger.

B.2 Muon ID Efficiencies and Scale Factors

Muons selected for this study are required to fulfill the same selection criteria as in the cross section analysis except for the p_{T} and relative isolation criteria. Probe muons are required to have a p_{T} greater than 30 GeV to avoid the scale factor turn on of the HLT_IsoMu24 trigger, shown in Figure B.10. The ID selection requires a muon with relative isolation less than 0.14 due to the scale factor dependence of the HLT_IsoMu24 trigger, shown in Figure B.9.

The muon identification (ID) efficiency is measured with the tag and probe method, as described in Chapter 11. The denominator in Equation 11.2, the number of probe muons, is given by requiring a global muon with $p_{\mathrm{T}} > 30$ GeV and $\eta < 2.1$. The numerator is the number of probes passing the ID criteria.

The muon ID efficiency and scale factor measured in regions of η are shown in Figure B.1. The scale factor varies with respect to the η of the muon, therefore an η dependent scale factor should be applied to the simulated ID efficiency.

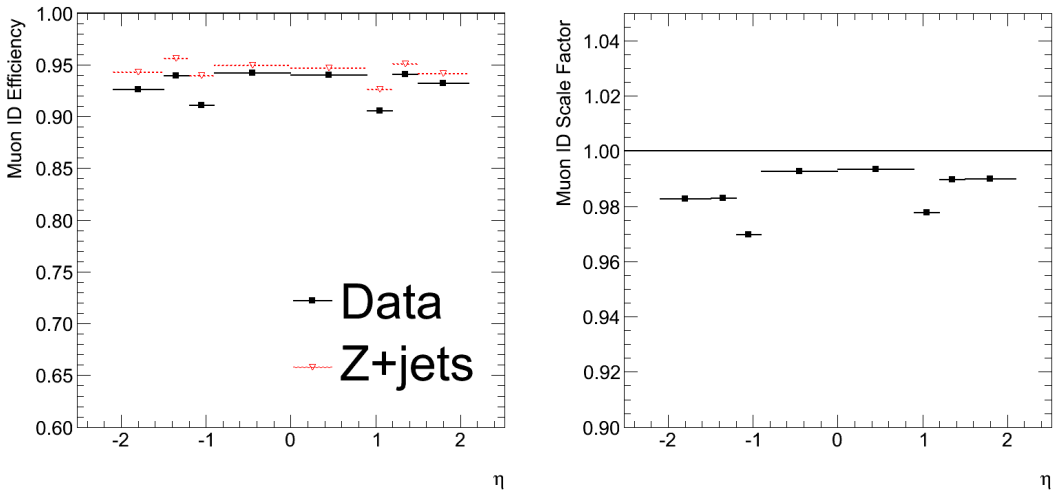


Figure B.1: Muon ID efficiency and scale factor measured in regions of η.

After reweighting the simulated ID efficiency as a function of the reconstructed muon η, the efficiency measurement is repeated in regions of other variables to uncover further scale factor dependencies.

The efficiency and scale factor are presented as a function of p_{T} in Figures B.2 and B.3. A gradual turn on in observed in the measured ID efficiencies, however, above the muon p_{T} requirement of 30 GeV the scale factor is stable within ± 0.005. The ± 0.005 range is considered to be a systematic uncertainty on the scale factor measured with respect to the muon p_{T}.

The $Z \rightarrow \mu\mu$ events used in the scale factor measurement tend to have low jet multiplicities whereas the signal selection requires at least four jets. Therefore, it is important to also quantify any scale factor dependence on the jet multiplicity. Figure B.5 shows that the scale factors measured with respect to the jet multiplicity agree within ± 0.005. The number of primary vertices reconstructed in an event corresponds to the amount of pile-up in the event. There is a downward trend in the scale factor with increasing number of primary vertices. However, considering the statistical uncertainties, the scale factors agree within ± 0.01 over the full range of number of primary vertices.

The η dependent scale factors applied to the simulated ID efficiency are listed in Table B.1 with the statistical uncertainty on the scale factor measured in each region. A systematic uncertainty of ± 0.01 accounts for scale factor deviations seen with respect to muon p_{T}, jet multiplicity and number of primary vertices.

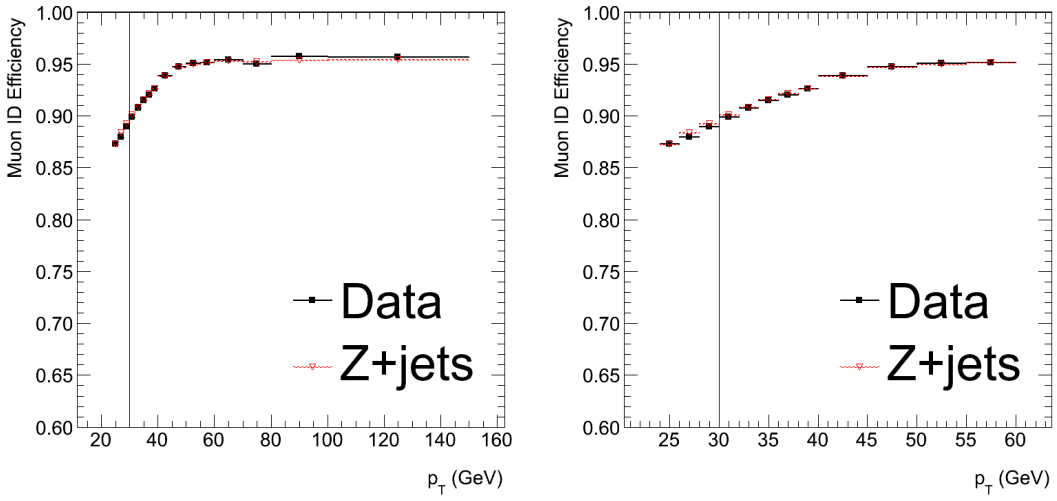


Figure B.2: Muon ID efficiency measured in regions of p_{T}. The η dependent muon ID scale factor has been applied to the simulated efficiency.

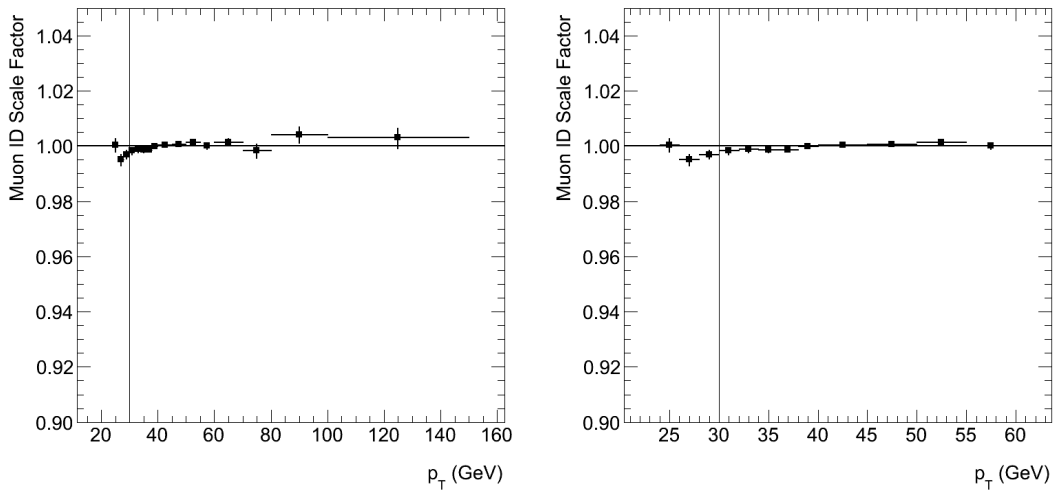


Figure B.3: Muon ID scale factor measured in regions of p_{T}. Probe muons are required to have a $p_{\mathrm{T}} > 30$ GeV. The same scale factors are shown in both plots, with the plot on the right shown for a smaller range of muon p_{T} to examine the scale factor behaviour close to the p_{T} threshold of 30 GeV. The η dependent scale factor has been applied.

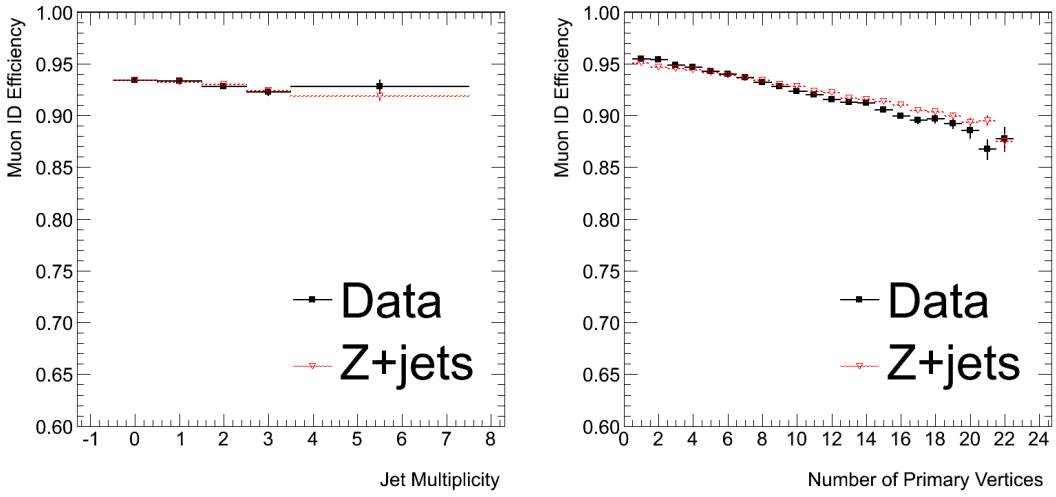


Figure B.4: Muon ID efficiency measured with respect to jet multiplicity and number of primary vertices. The η dependent muon ID scale factor has been applied to the simulated efficiency.

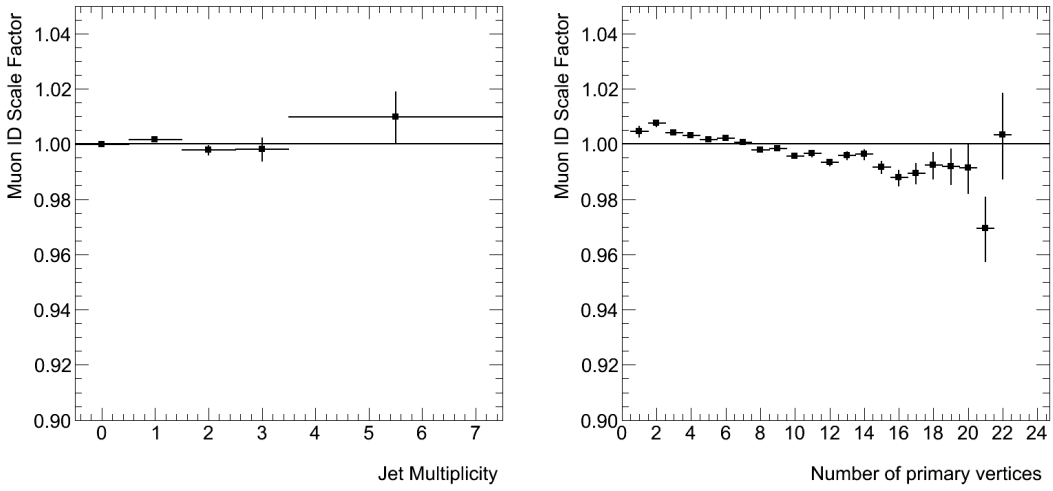


Figure B.5: Muon ID scale factor measured with respect to jet multiplicity and number of primary vertices. The η dependent scale factor has been applied.

η Region	Scale Factor
-2.1 to -1.5	$0.9826^{+0.0004}_{-0.0008}$
-1.5 to -1.2	$0.9828^{+0.0004}_{-0.0008}$
-1.2 to -0.9	$0.9697^{+0.0005}_{-0.0009}$
-0.9 to 0.0	$0.9926^{+0.0002}_{-0.0004}$
0.0 to 0.9	$0.9933^{+0.0002}_{-0.0004}$
0.9 to 1.2	$0.9776^{+0.0005}_{-0.001}$
1.2 to 1.5	$0.9896^{+0.0005}_{-0.0009}$
1.5 to 2.1	$0.9899^{+0.0004}_{-0.0008}$

Table B.1: Muon ID scale factors in regions of muon η. Uncertainties are statistical only.

B.3 Trigger Efficiency

The efficiency of the HLT_IsoMu24 trigger to select tight isolated muons is measured with the tag and probe method. The probe muon required to pass the tag selection criteria, except that it is not matched to a trigger object. The probe muon is considered to have passed the trigger if it matches to a HLT_IsoMu24 trigger object. A match is counted if a L3 trigger object which passed the HLT_IsoMu24 requirements is within

$\Delta R < 0.2$ of the probe muon.

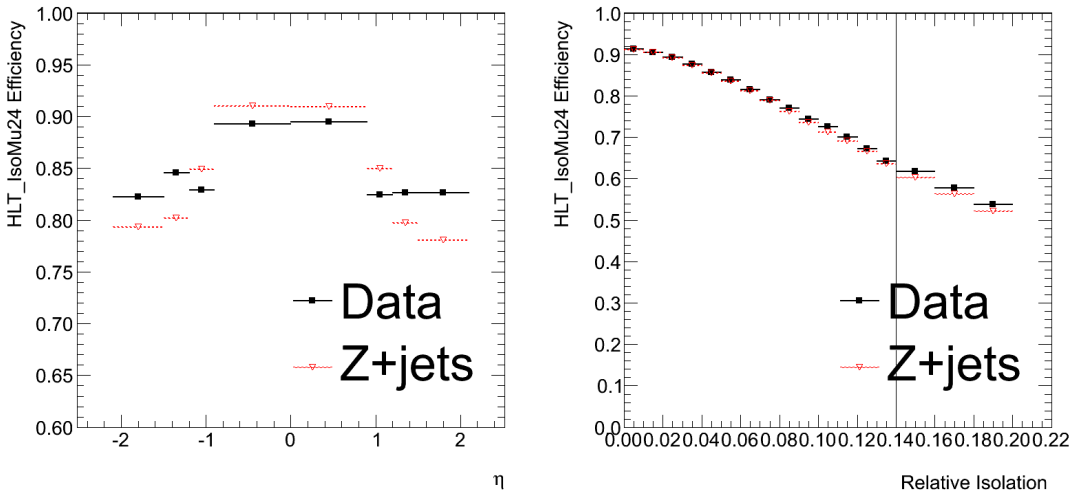


Figure B.6: HLT_IsoMu24 efficiency measured in regions of η and relative isolation. The probe ID selection requires muons to have a relative isolation less than 0.14.

The efficiency of the HLT_IsoMu24 trigger in regions of η and relative isolation is shown in Figure B.6. The simulation overestimates the efficiency for muons with $|\eta|$ greater than 1.2 and underestimates the efficiency for triggering muons with $|\eta|$ less than 1.2. This is due to η dependent adjustments to the HLT_IsoMu24 trigger in the trigger menu used for data taking. The trigger menu adjustments which affected the efficiency with respect to η are discussed further in Section B.4.2.

The HLT_IsoMu24 efficiency in simulation agrees with data for muons with relative isolation less than 0.08. However, for muons with relative isolation greater than 0.08 the trigger efficiency is underestimated in simulation. This underestimation becomes more significant for muons with relative isolation greater than 0.14, which is why the probe ID selection requires muons to have a relative isolation less than 0.14.

The turn on of the trigger efficiency with respect to muon p_{T} is shown in Figure B.7 after the simulated efficiency has been scaled to agree with the efficiency measured in data with respect to η and relative isolation.

The trigger efficiency for different jet multiplicities and number of primary vertices is shown in Figure B.8. The increase in the trigger efficiency in data for events with three jets appears to be a statistical effect. The trigger efficiency dependence on the number of primary vertices results in a drift in the overall trigger efficiency during data taking as the instantaneous luminosity, and consequently the event pile-up, increases. This dependence is discussed further in the Section B.4.2.

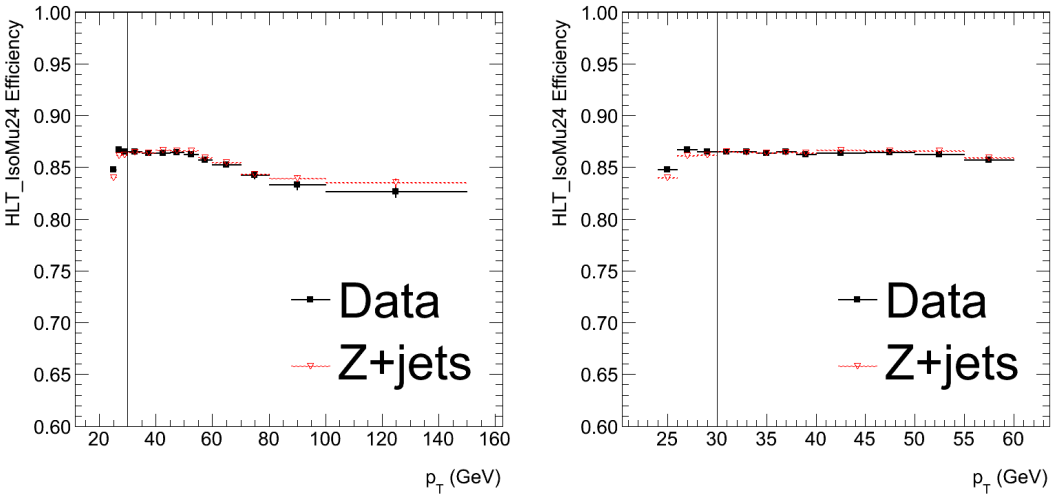


Figure B.7: HLT_IsoMu24 efficiency measured in regions of p_{T}. Probe muons are required to have a $p_{\mathrm{T}} > 30$ GeV. The simulated efficiency has been scaled to agree with the efficiency measured in data with respect to η and relative isolation.

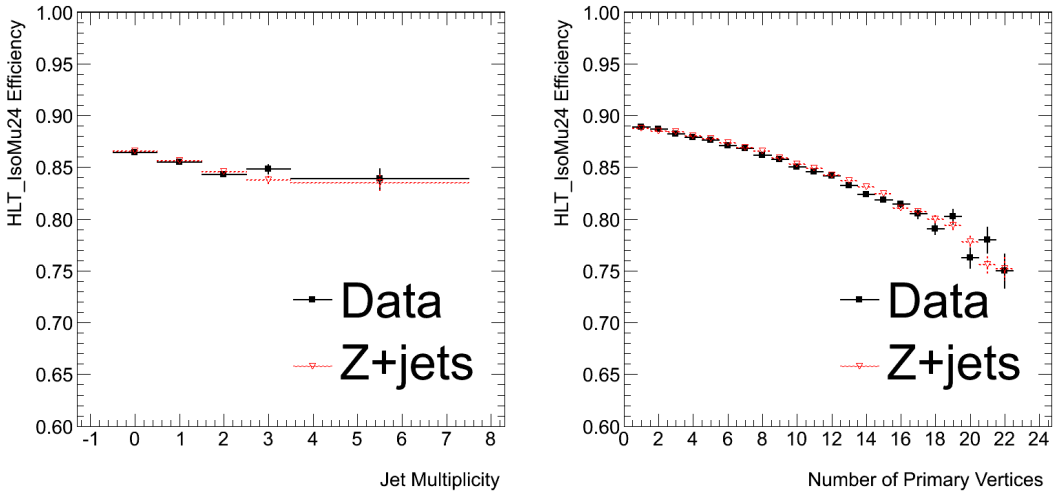


Figure B.8: HLT_IsoMu24 efficiency measured with respect to jet multiplicity and number of primary vertices. The simulated efficiency has been scaled to agree with the efficiency measured in data with respect to η and relative isolation.

B.4 Trigger Scale Factors

A significant scale factor dependence on the η of the muon is observed in Figure B.9. This is due to η dependent changes to the HLT_IsoMu24 trigger in different trigger menus, described in Section B.1. The effects of these trigger menu changes on the trigger efficiency are examined later in this appendix.

In Figure B.9 it is shown that the scale factor increases with increasing muon relative isolation. The selection requirement of 0.14 on the muon relative isolation was chosen to reduce the effect of this scale factor dependence. The choice of a muon relative isolation requirement of 0.14 instead of 0.08 was motivated by the retention of $t\bar{t}$ events.

The η and relative isolation efficiencies observed in data are reproduced by applying a two dimensional scale factor to the simulated trigger efficiency. This scale factor is derived in each η region for muon relative isolation less than 0.08 and between 0.08 and 0.14. The resultant scale factors are given in Table B.2.

Figures B.10 and B.11 display the trigger scale factor with respect to p_{T}, jet multiplicity and number of primary vertices after the two dimensional trigger scale factor is applied. Deviations from unity are accounted for by considering a systematic uncertainty of $\pm$ 0.006 on the applied scale factor.

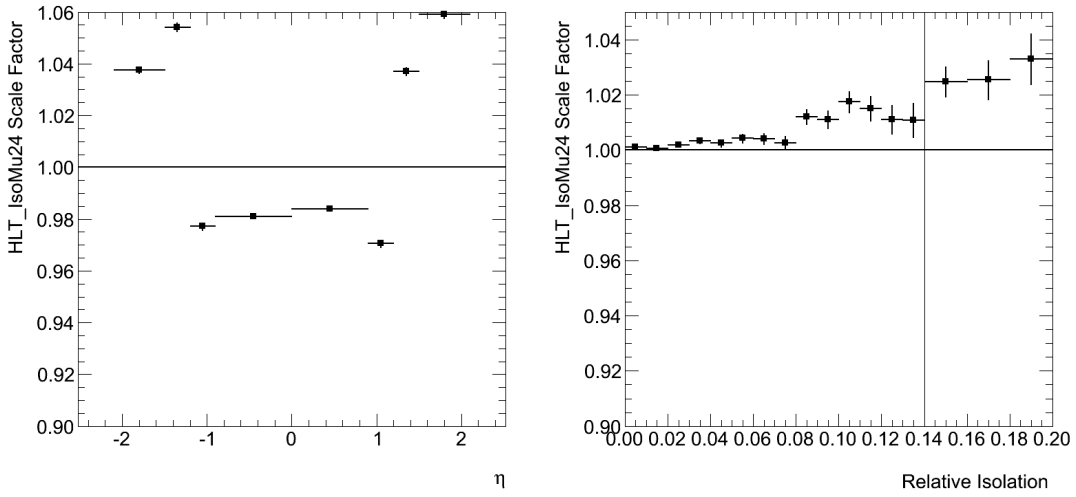


Figure B.9: HLT_IsoMu24 scale factor measured in regions of η and relative isolation. Probe muons are required to have a relative isolation less than 0.14.

	Relative Isolation	
η	$0.0 - 0.08$	$0.08 - 0.14$
-2.1 to -1.5	1.036 ± 0.001	1.062 ± 0.006
-1.5 to -1.2	1.054 ± 0.002	1.072 ± 0.008
-1.2 to -0.9	0.979 ± 0.001	0.984 ± 0.007
-0.9 to 0.0	0.983 ± 0.001	0.982 ± 0.003
0.0 to 0.9	0.985 ± 0.001	0.993 ± 0.003
0.9 to 1.2	0.972 ± 0.001	0.978 ± 0.007
1.2 to 1.5	1.038 ± 0.002	1.048 ± 0.008
1.5 to 2.1	1.058 ± 0.001	1.078 ± 0.006

Table B.2: Muon trigger scale factors in regions of muon η and relative isolation. Uncertainties are statistical only.

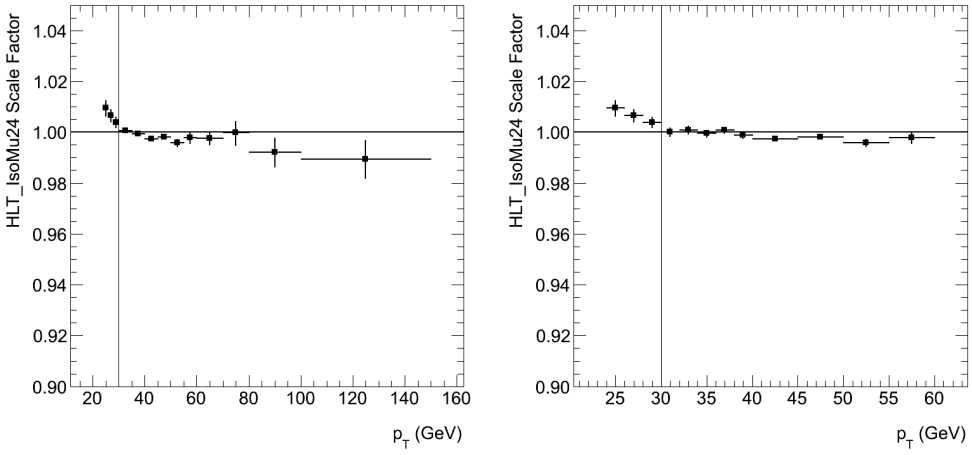


Figure B.10: HLT_IsoMu24 scale factor measured in regions of p_{T}. Probe muons are required to have a $p_{\mathrm{T}} > 30$ GeV. The same scale factors are shown in both plots, with the plot on the right shown for a smaller range of muon p_{T} to examine the scale factor behaviour close to the trigger p_{T} threshold. The two dimensional η-Iso$_{\mathrm{rel.}}$ scale factor has been applied.

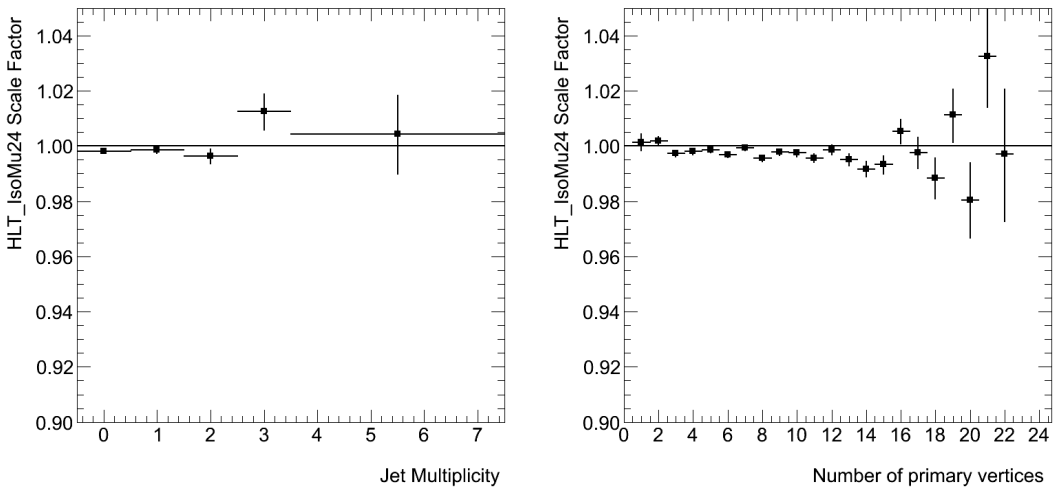


Figure B.11: HLT_IsoMu24 scale factor measured with respect to jet multiplicity and number of primary vertices. The two dimensional η-Iso$_{\mathrm{rel.}}$ scale factor has been applied.

B.4.1 L1 and HLT Efficiencies

The efficiency of the HLT_IsoMu24 trigger is a combination of the efficiency of the L1 trigger and of the HLT. The contribution from the L1 trigger can be measured separately by matching L1 trigger objects passing the L1 trigger to the probe muon. A match is counted if the L1 trigger object is within $\Delta R < 0.3$ of the probe muon. The L1 matching requirement is more relaxed than the L3 matching requirement as the L1 trigger object reconstruction is not as accurate as the L3 reconstruction. The efficiency of the HLT is extracted by measuring the HLT_IsoMu24 efficiency for probe muons which have passed the L1 trigger.

The L1 and HLT contributions to the HLT_IsoMu24 efficiency are shown in Figure B.12 with respect to muon η and relative isolation. The HLT_IsoMu24 efficiency differences between simulation and data are dominated by the L1 efficiency difference for $|\eta| > 1.5$, whereas the discrepancies in $|\eta| < 1.5$ region are mainly seen in the HLT efficiency. The discrepancies are due to the fact that the data was collected with a number of trigger menus while only one trigger menu was used in simulation. The trigger menu changes which affect the efficiency are discussed further in Section B.4.2.

Figure B.12 shows that the efficiency dependence with respect to relative isolation is only present in the HLT. This is due to the requirements on the tracker and calorimeter isolation in the HLT, which affect the relative isolation of the muon but are not equivalent to relative isolation. The L1 trigger has no isolation requirement therefore the L1 efficiency is stable with respect to relative isolation.

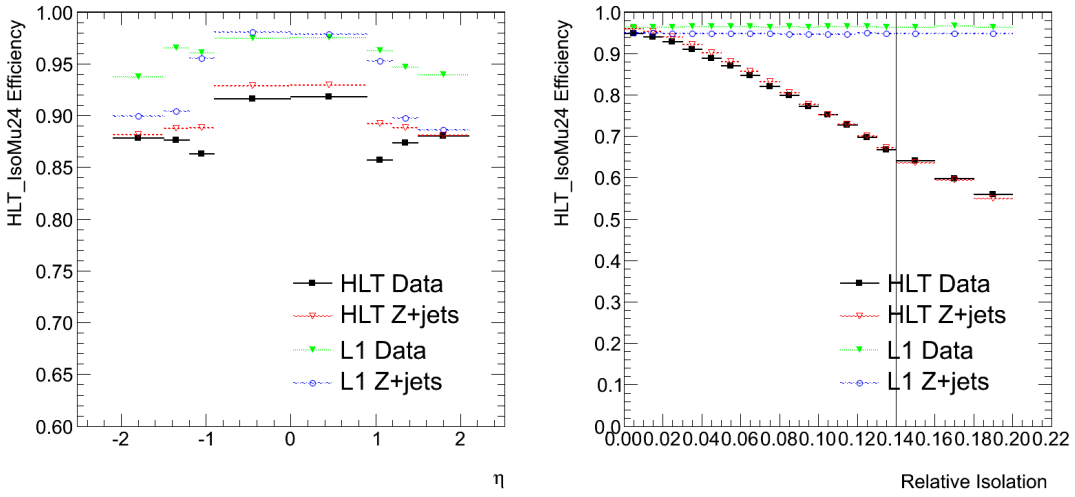


Figure B.12: L1 and HLT efficiencies for HLT_IsoMu24 measured in regions of η and relative isolation. Probe muons are required to have a relative isolation less than 0.14.

The L1 and HLT efficiencies with respect to muon p_{T} are shown in Figure B.13. The HLT efficiency shows a sharp turn on at the trigger p_{T} threshold of 24 GeV, reaching an efficiency plateau at 26 GeV. However, the more gradual L1 efficiency turn on in simulation results in scale factor dependencies with respect to muon p_{T}. The requirement of muon p_{T} greater than 30 GeV is chosen to minimise this dependence. The decrease in efficiency at high muon p_{T} is due to the absolute isolation requirements in the HLT, as high p_{T} muons will have more signal in the surrounding detector material and are therefore more likely to fail the absolute isolation criteria.

The L1 and HLT efficiency with respect to jet multiplicity and number of primary vertices is shown in Figure B.14. The decrease in the HLT efficiency as the number of jets in the event increases is due to the isolation requirements in the HLT. Muons in events with higher jet multiplicities are more likely to have detector hits close to the muon track, causing the muon to fail the HLT isolation requirements.

The HLT isolation requirements are also the cause of the HLT efficiency dependence with respect to the number of primary vertices. An increase in the number of primary vertices corresponds to an increase in pile-up. The resulting increase in detector hits causes an increase in the rate of muons failing the isolation requirements. The L1 trigger has no isolation requirement therefore the L1 efficiency is stable with respect to jet multiplicity and number of primary vertices.

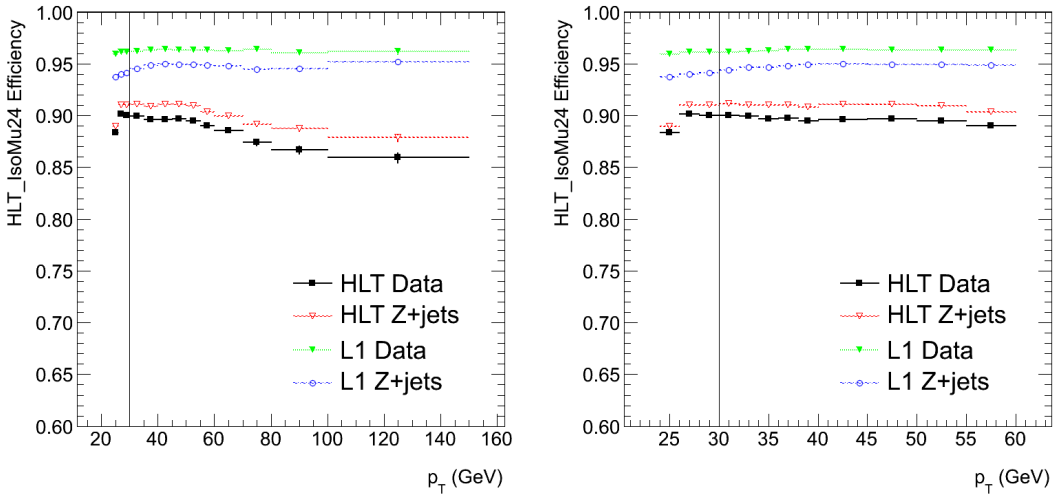


Figure B.13: L1 and HLT efficiencies for HLT_IsoMu24 measured in regions of p_{T}. Probe muons are required to have $p_{\mathrm{T}} > 30$ GeV.

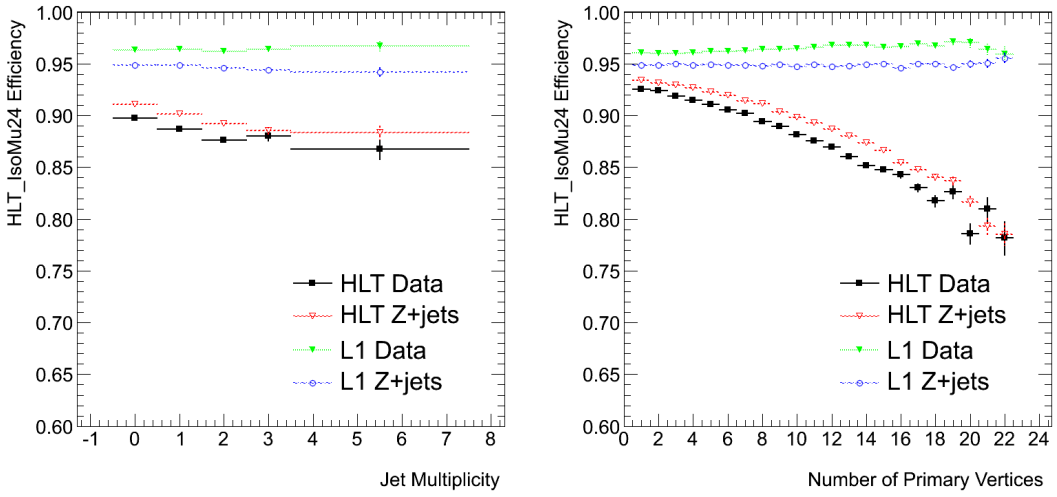


Figure B.14: L1 and HLT efficiencies for HLT_IsoMu24 measured in regions of jet multiplicity and number of primary vertices.

B.4.2 Efficiency Evolution

The evolution of the L1 and HLT components of HLT_IsoMu24 for each trigger menu is examined to study variations in the trigger efficiency. The trigger menus and the changes made to the trigger are described in Section B.1.

The evolution of the L1 trigger with respect to η and relative isolation is shown in Figure B.15. The 3e33 menu is introduced after a technical stop of the CMS detector. The decrease in L1 efficiency is due to changes to the detector configuration during the technical stop, such as an increase in the number of deactivated muon chambers. The change in the L1 muon p_{T} assignment in a revision of the 3e33 menu results in a significant increase in the L1 efficiency across the full η range.

The evolution of the HLT efficiency with respect to η and relative isolation is shown in Figure B.16. The drop in efficiency in the $0.9 < |\eta| < 1.5$ region in the 2e33 menu is due to the addition of L2 quality criteria. Otherwise the HLT efficiency decreases with each new menu due the efficiency dependence on the number of primary vertices in the event. The instantaneous luminosity of the data collected increases with each new menu, corresponding to an increase in pile-up and therefore number of primary vertices. Since the data sample is concentrated in higher primary vertex multiplicities, corresponding with lower trigger efficiencies in Figure B.20, this results in a downward trend in the overall HLT efficiency in Figure B.21.

The L1 and HLT efficiencies in regions of muon p_{T} are shown in Figures B.17 and B.18. While the overall efficiency varies, the efficiency dependence for muons with p_{T}

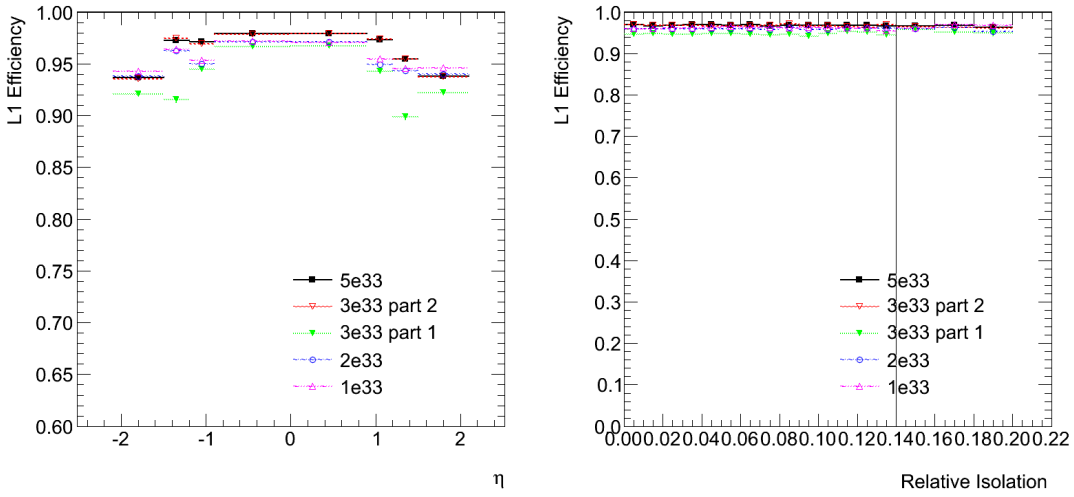


Figure B.15: L1 efficiency for HLT_IsoMu24 measured for each trigger menu in regions of η and relative isolation. Probe muons are required to have a relative isolation less than 0.14.

greater than 30 GeV is stable for each trigger menu. The change of L1 trigger from L1_SingleMu12 to L1_SingleMu16_eta2p1 in the 3e33 menu does not affect the L1 efficiency as both triggers reach a plateau before 30 GeV.

The L1 and HLT efficiency with respect to jet multiplicity and number of primary vertices in the event is shown in Figure B.19 and B.20. While the overall efficiency varies, there are no significant deviations in the behaviour of the trigger with respect to the jet multiplicity or number of primary vertices.

A summary of the overall efficiency for each trigger menu is shown in Figure B.21.

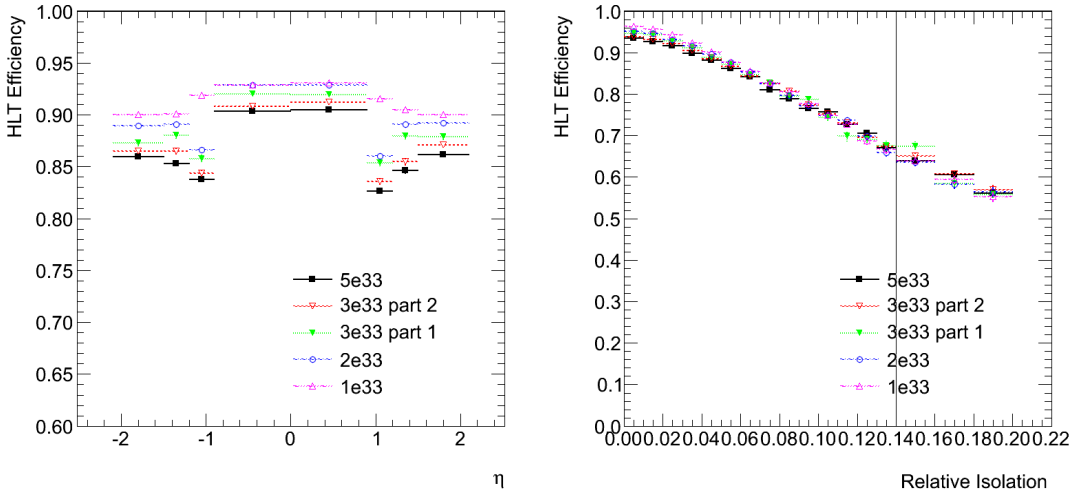


Figure B.16: HLT efficiency for HLT_IsoMu24 measured for each trigger menu in regions of η and relative isolation. Probe muons are required to have a relative isolation less than 0.14.

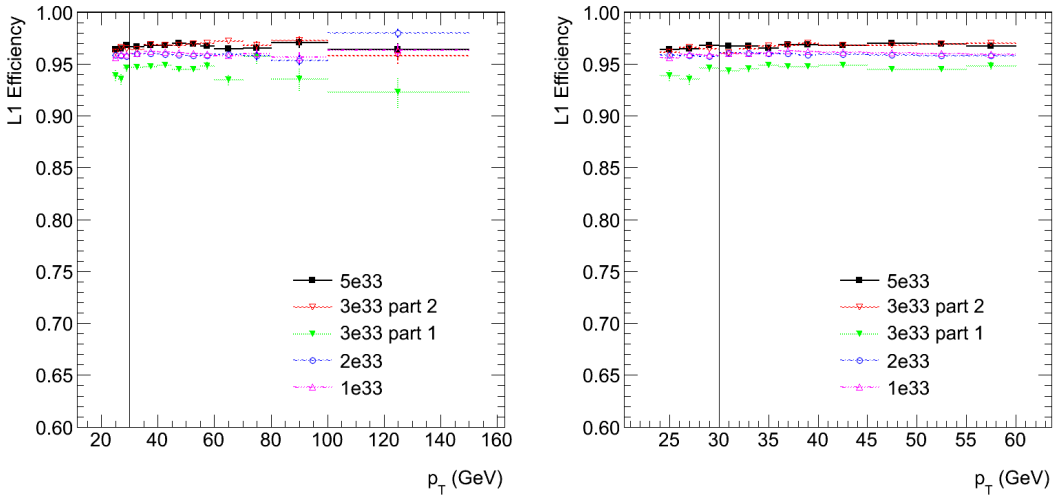


Figure B.17: L1 efficiency for HLT_IsoMu24 measured for each trigger menu in regions of p_{T}. Probe muons are required to have $p_{\mathrm{T}} > 30$ GeV.

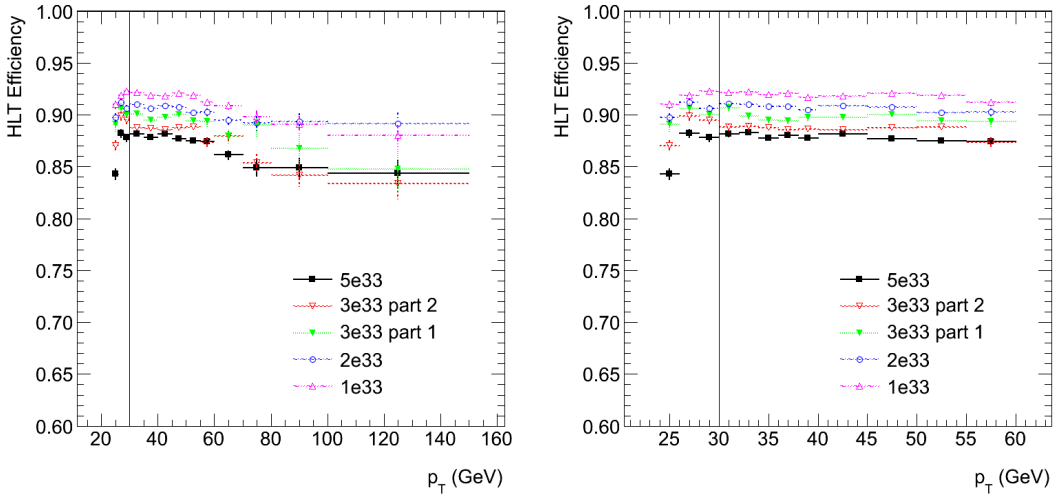


Figure B.18: HLT efficiency for HLT_IsoMu24 measured for each trigger menu in regions of p_{T}. Probe muons are required to have $p_{\mathrm{T}} > 30$ GeV.

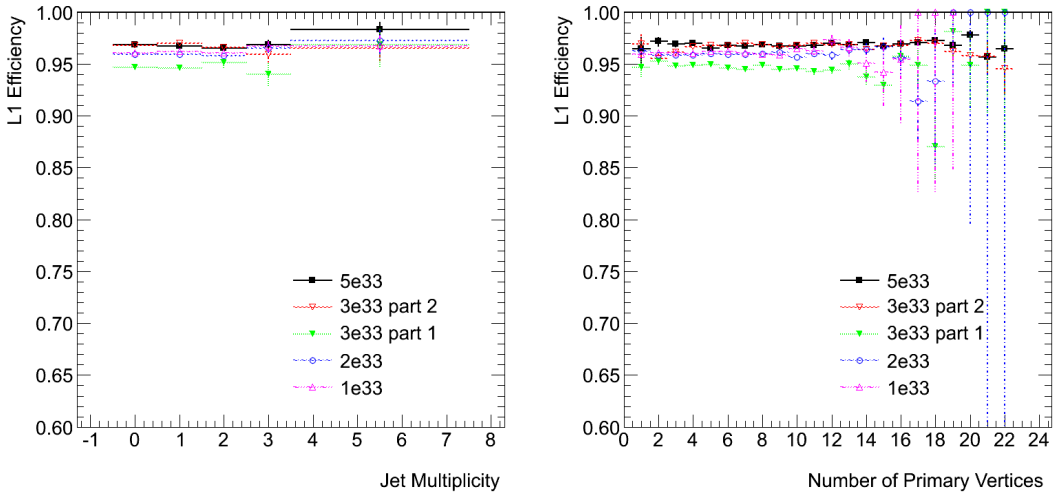


Figure B.19: L1 efficiency for HLT_IsoMu24 measured for each trigger menu in regions of jet multiplicity and number of primary vertices.

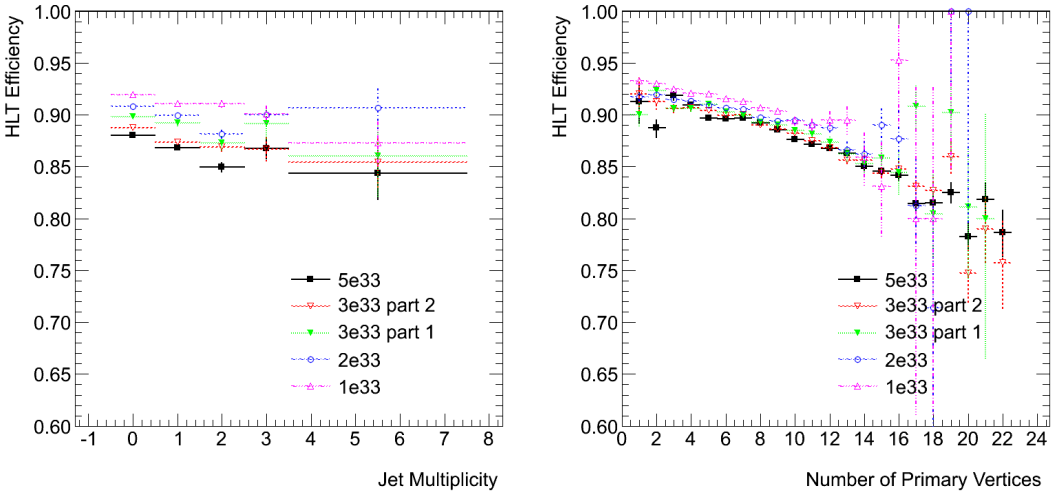


Figure B.20: HLT efficiency for HLT_IsoMu24 measured for each trigger menu in regions of jet multiplicity and number of primary vertices.

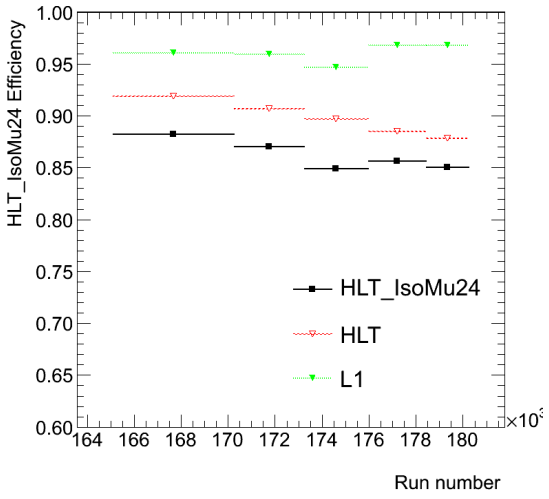


Figure B.21: HLT_IsoMu24 L1 and HLT efficiency measured for each trigger menu. The bin boundaries correspond with the boundaries of the trigger menus.

List of Abbreviations

ARI	anti-relative isolation
CKM	Cabibbo-Kobayashi-Maskawa
CMS	Compact Muon Solenoid
CSC	Cathode Strip Chamber
DT	Drift Tube
ECAL	Electromagnetic Calorimeter
EWK	Electroweak
FSR	Final State Radiation
HCAL	Hadronic Calorimeter
HLT	High Level Trigger
IR	Interaction Region
ISR	Initial-State Radiation
JER	Jet Energy Resolution
JES	Jet Energy Scale
L1	Level-1
L2	Level-2
L3	Level-3
LHC	Large Hadron Collider
MC	Monte Carlo
ME	Matrix Element
NLL	Next-to-leading logarithm
NLO	Next-to-leading order
NNLL	Next-to-next-to-leading logarithm

NNLO	Approximate next-to-next-to-leading order
PDF	Parton Distribution Function
PS	Parton Shower
QCD	Quantum chromodynamics
QED	Quantum electrodynamics
RPC	Resistive Plate Chamber
SM	Standard Model
TTC	Timing, Trigger and Control
TTS	Trigger Throttle System

List of Figures

List of Tables

Bibliography

[1] Particle Data Group, "Review of Particle Physics," *Phys. Rev. D* **86** (2012) 010001. pdg.lbl.gov. (Cited on pages 5, 8, 9, 14, 16, 71 and 117.)

[2] S. L. Glashow, "Towards a unified theory: Threads in a tapestry," *Rev. Mod. Phys.* **52** (1980) 539–543. doi:10.1103/RevModPhys.52.539. (Cited on page 5.)

[3] A. Salam, "Gauge unification of fundamental forces," *Rev. Mod. Phys.* **52** (1980) 525–538. doi:10.1103/RevModPhys.52.525. (Cited on page 5.)

[4] S. Weinberg, "Conceptual foundations of the unified theory of weak and electromagnetic interactions," *Rev. Mod. Phys.* **52** (1980) 515–523. doi:10.1103/RevModPhys.52.515. (Cited on page 5.)

[5] **CMS** Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Submitted to Phys. Lett. B* (2012) arXiv:1207.7235. CERN-PH-EP-2012-220. (Cited on page 6.)

[6] **ATLAS** Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC," *Submitted to Phys. Lett. B* (2012) arXiv:1207.7214. CERN-PH-EP-2012-218. (Cited on page 6.)

[7] S. G. Karshenboim, "Precision physics of simple atoms: QED tests, nuclear structure and fundamental constants," *Phys. Reports* **422** (2005) 1–63. doi:10.1016/j.bbr.2011.03.031. (Cited on page 7.)

[8] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys. Rev. Lett.* **13** (1964) 321–323. doi:10.1103/PhysRevLett.13.321. (Cited on page 8.)

[9] P. W. Higgs, "Broken Symmetries and the Masses of Gauge Bosons," *Phys. Rev. Lett.* **13** (1964) 508–509. doi:10.1103/PhysRevLett.13.508. (Cited on page 8.)

[10] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, "Global Conservation Laws and Massless Particles," *Phys. Rev. Lett.* **13** (1964) 585–587. doi:10.1103/PhysRevLett.13.585. (Cited on page 8.)

[11] S. W. Herb et al., "Observation of a Dimuon Resonance at 9.5 GeV in 400-GeV Proton-Nucleus Collisions," *Phys. Rev. Lett.* **39** (1977) 252–255. doi:10.1103/PhysRevLett.39.252. (Cited on page 9.)

[12] **CDF** Collaboration, "Observation of Top Quark Production in Pbar-P Collisions," *Phys. Rev. Lett.* **74** (1995) 2626–2631, arXiv:hep-ex/9503002. (Cited on page 9.)

[13] **D0** Collaboration, "Observation of Top Quark," *Phys. Rev. Lett.* **74** (1995) 2632–2637, arXiv:hep-ex/9503003. (Cited on page 9.)

[14] mstwpdf.hepforge.org/plots/plots.html. (Cited on pages 10 and 173.)

[15] B. Binosi and L. Theußl, "JaxoDraw: A graphical user interface for drawing Feynman diagrams ," *Comp. Phys. Comm.* **161** (2004) 76–86. doi:10.1016/j.cpc.2004.05.001. (Cited on pages 11, 16, 18, 19 and 173.)

[16] J. Pumplin et al., "New generation of parton distributions with uncertainties from global QCD analysis," *JHEP* **07** (2002) 012, arXiv:hep-ph/0201195. (Cited on pages 11 and 49.)

[17] W. Wagner, "Top quark physics at hadron colliders," *Rep. Prog. Phys.* **68** (2005) 2409–2494, arXiv:hep-ph/0507207. (Cited on page 11.)

[18] F. Maltoni et al., "Choosing the Factorization Scale in Perturbative QCD," 2007, arXiv:hep-ph/0703156. (Cited on page 12.)

[19] S. J. Brodsky and L. Di Giustino, "Setting the Renormalization Scale in QCD: The Principle of Maximum Conformality," 2011, arXiv:1107.0338. (Cited on page 12.)

[20] P. Schmueser, "The electron-proton colliding beam facility HERA," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **235** (1985), 2, 201 – 208. doi:10.1016/0168-9002(85)90554-6. (Cited on page 12.)

[21] A. Cafarella, C. Coriano, and M. Guzzi, "NNLO logarithmic expansions and exact solutions of the DGLAP equations from x-space: New algorithms for precision studies at the LHC," *Nucl. Phys. B* **748** (2006) 253–308, arXiv:hep-ph/0512358. (Cited on page 12.)

[22] **CMS** Collaboration, "Determination of the Top Quark Mass from the ttbar Cross Section at sqrt(s) = 7 TeV," 2011. CMS-PAS-TOP-11-008. cdsweb.cern.ch/record/1387001. (Cited on pages 13, 136, 138, 173 and 176.)

[23] N. Kidonakis, "Higher-order corrections to top-antitop pair and single top quark production," 2009, arXiv:0909.0037. (Cited on pages 14, 15, 132 and 143.)

[24] R. Gavin, Y. Li, F. Petriello, and S. Quackenbush, "FEWZ 2.0: A code for hadronic Z production at next-to-next-to-leading order," *Comput. Phys. Commun.* **82** (2011) 2388–2403, arXiv:1011.3540. (Cited on page 15.)

[25] N. Kidonakis, "Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production," *Phys. Rev. D* **83** (2011), 091503, arXiv:1103.2972. (Cited on page 15.)

[26] N. Kidonakis, "Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-," *Phys. Rev. D* **82** (2010), 054018, arXiv:1005.4451. (Cited on page 15.)

[27] N. Kidonakis, "NNLL resummation for s-channel single top quark production," *Phys. Rev. D* **81** (2010), 054028, arXiv:1001.5034. (Cited on page 15.)

[28] T. Sjöstrand, S. Mrenna, and P. Z. Skands, "PYTHIA 6.4 Physics and Manual," *JHEP* **05** (2006) 026, arXiv:hep-ph/0603175. (Cited on pages 15 and 47.)

[29] O. Brüning, H. Burkhardt, and S. Myers, "The large hadron collider," *Prog. Part. Nucl. Phys.* **67** (2012) 705–734. doi:10.1016/j.ppnp.2012.03.001. (Cited on pages 23, 24, 26, 173 and 179.)

[30] R. Scrivens et al., "Overview of the status and developments on primary ion sources at CERN," 2011. CERN-ATS-2011-172. cdsweb.cern.ch/record/1382102. (Cited on page 23.)

[31] twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults2011. (Cited on pages 26 and 173.)

[32] lhc-commissioning.web.cern.ch/lhc-commissioning/. (Cited on pages 26 and 179.)

[33] **CMS** Collaboration, "CMS Physics Technical Design Report Volume I : Detector Performance and Software". Technical Design Report CMS. CERN, Geneva, 2006. CMS-TDR-008-1. (Cited on pages 27, 31, 32, 34, 35 and 173.)

[34] **CMS** Collaboration, "CMS physics Technical Design Report, Volume II: Physics Performance ", volume 34. 2006. CERN-LHCC-2006-021. CMS-TDR-008-2. (Cited on page 27.)

[35] **CMS** Collaboration, "CMS Tracking Performance Results from early LHC Operation," *Eur. Phys. J.* **C70** (2010) 1165 – 1192, arXiv:1007.1988. (Cited on pages 29, 30 and 173.)

[36] twiki.cern.ch/twiki/bin/view/CMSPublic/DPGResultsTRK. (Cited on page 30.)

[37] **CMS** Collaboration, "Electromagnetic calorimeter commissioning and first results with 7 TeV data," 2010. CMS-NOTE-2010-012. cdsweb.cern.ch/record/1278160. (Cited on page 30.)

[38] twiki.cern.ch/twiki/bin/view/CMSPublic/EcalDPGResults. (Cited on pages 31 and 173.)

[39] **CMS** Collaboration, "HCAL performance from first collisions data," 2010. CMS-DP-2010-025. cdsweb.cern.ch/record/1364222. (Cited on page 32.)

[40] **CMS** Collaboration, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV," 2012, arXiv:1206.4071. CMS-MUO-10-004. (Cited on pages 33, 65, 73 and 173.)

[41] **CMS** Collaboration, "Absolute Calibration of the Luminosity Measurement at CMS: Winter 2012 Update," 2012. CMS-PAS-SMP-12-008. cdsweb.cern.ch/record/1434360. (Cited on page 36.)

[42] S. Van der Meer, "Calibration of the effective beam height in the ISR," 1968. CERN-ISR-PO-68-31. ISR-PO-68-31. cdsweb.cern.ch/record/296752/. (Cited on page 36.)

[43] **CMS** Collaboration, "CMS TriDAS project: Technical Design Report; 1, the trigger systems". Technical Design Report CMS. 2000. CMS-TDR-006-1. (Cited on page 39.)

[44] S. Cittolin, A. Rácz, and P. Sphicas, "CMS trigger and data-acquisition project: Technical Design Report". Technical Design Report CMS. CERN, Geneva, 2002. CERN-LHCC-2002-026. (Cited on page 39.)

[45] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," *Transactions of the ASME - Journal of Basic Engineering* (1960), 82, 35–45. www.citeulike.org/user/alexv/article/347166. (Cited on pages 42 and 54.)

[46] M. A. Dobbs et al., "Les Houches Guidebook to Monte Carlo Generators for Hadron Collider Physics," 2004, arXiv:hep-ph/0403045. (Cited on page 47.)

[47] B. Andersson, G. Gustafson, G. Ingelman, and T. Sjöstrand, "Parton fragmentation and string dynamics," *Phys. Rept.* **97** (1983) 31–145. doi:10.1016/0370-1573(83)90080-7. (Cited on page 48.)

[48] T. Stelzer and W. F. Long, "Automatic Generation of Tree Level Helicity Amplitudes," *Comput.Phys.Commun.* **81** (1994) 357–371, arXiv:hep-ph/9401258. (Cited on page 48.)

[49] S. Hoeche et al., "Matching Parton Showers and Matrix Elements," 2006, arXiv:hep-ph/0602031v1. (Cited on page 48.)

[50] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method," 2007, arXiv:0709.2092. GEF-TH-21/2007. (Cited on page 48.)

[51] J. Allison et al., "Geant4 developments and applications," *IEEE Trans. Nucl. Sci.* **53** (2006) 270–278. doi:10.1109/TNS.2006.869826. (Cited on page 49.)

[52] **CMS** Collaboration, "Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET," 2009. CMS-PAS-PFT-09-001. cdsweb.cern.ch/record/1194487. (Cited on pages 51 and 52.)

[53] **CMS** Collaboration, "Particle-flow commissioning with muons and electrons from J/Psi and W events at 7 TeV," 2010. CMS-PAS-PFT-10-003. cdsweb.cern.ch/record/1279347. (Cited on pages 51 and 55.)

[54] **CMS** Collaboration, "CMS Tracking Performance Results from Early LHC Operation," *Eur. Phys. J. C* **70** (2010) 1165. 29 p, arXiv:1007.1988. (Cited on page 52.)

[55] **CMS** Collaboration, "Measurement of Tracking Efficiency," 2010. CMS-PAS-TRK-10-002. cdsweb.cern.ch/record/1279139. (Cited on pages 52 and 173.)

[56] **CMS** Collaboration, "Performance of muon identification in pp collisions at $\sqrt{s} =$ 7 TeV," 2010. CMS-PAS-MUO-10-002. cdsweb.cern.ch/record/1279140. (Cited on page 54.)

[57] W. Adam, R. Frühwirth, A. Strandlie, and T. Todorov, "Reconstruction of Electron Tracks With the Gaussian-Sum Filter," 2003. CMS-RN-2003-001. cdsweb.cern.ch/record/687491. (Cited on page 55.)

[58] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm," *JHEP* **04** (2008) 063, arXiv:0802.1189. (Cited on pages 56, 58 and 173.)

[59] S. D. Ellis and D. E. Soper, "Successive Combination Jet Algorithm For Hadron Collisions," *Phys. Rev. D* **48** (1993) 3160–3166, arXiv:hep-ph/9305266. (Cited on page 56.)

[60] Y. L. Dokshitzer, G. D. Leder, S. Moretti, and B. R. Webber, "Better Jet Clustering Algorithms," *JHEP* **08** (1997) 001, arXiv:hep-ph/9707323. (Cited on page 56.)

[61] G. P. Salam and G. Soyez, "A practical Seedless Infrared-Safe Cone jet algorithm," *JHEP* **0705** (2007) 086, arXiv:0704.0292. (Cited on page 58.)

[62] **CMS** Collaboration, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS," *JINST* **6** (2011) 11002, arXiv:1107.4277. (Cited on pages 58, 59 and 174.)

[63] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas," *Phys.Lett.* **B659** (2008) 119–126, arXiv:0707.1378. doi:10.1016/j.physletb.2007.09.077. (Cited on page 58.)

[64] T. Miao, N. Leioatts, H. Wenzel, and F. Yumiceva, "Beam Position Determination using Tracks," 2007. CMS-NOTE-2007-021. cdsweb.cern.ch/record/1061285. (Cited on page 59.)

[65] E. Chabanat and N. Estre, "Deterministic Annealing for Vertex Finding at CMS," 2005. cdsweb.cern.ch/record/865587. (Cited on page 60.)

[66] R. Frühwirth, W. Waltenberger, and P. Vanlaer, "Adaptive Vertex Fitting," 2007. CMS-NOTE-2007-008. cdsweb.cern.ch/record/1027031/. (Cited on page 60.)

[67] **CMS** Collaboration, "Tracking and Primary Vertex Results in First 7 TeV Collisions," 2010. CMS-PAS-TRK-10-005. cdsweb.cern.ch/record/1279383. (Cited on page 60.)

[68] twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsTOP. (Cited on pages 63 and 174.)

[69] **CMS** Collaboration, "Performance of muon identification in pp collisions at $\sqrt{s} = 7$ TeV," 2010. CMS-PAS-MUO-10-002. cdsweb.cern.ch/record/1279140. (Cited on page 65.)

[70] **CMS** Collaboration, "Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV," 2010. CMS-PAS-PFT-10-002. cdsweb.cern.ch/record/1279341. (Cited on page 66.)

[71] W. Verkerke and D. Kirkby, "The RooFit toolkit for data modeling," in *Talk from the 2003 Computing in High Energy and Nuclear Physics*. 2003. arXiv:physics/0306116. See also http://roofit.sourceforge.net. (Cited on page 92.)

[72] R. Brun and F. Rademakers, "ROOT - An Object Oriented Data Analysis Framework," in *Proceedings AIHENP'96 Workshop, Lausanne.*, Nucl. Inst. & Meth. in Phys. Res. A, volume 389, pp. 81–86. 1997. See also http://root.cern.ch/. (Cited on page 92.)

[73] L. Demortier and L. Lyons, "Everything you always wanted to know about pulls," technical report, CDF, 2002. CDF/ANAL/PUBLIC/5776. (Cited on pages 101 and 126.)

[74] J. Pumplin et al., "Uncertainties of predictions from parton distribution functions. 2. The Hessian method," *Phys. Rev. D* **65** (2001), 014013, `arXiv:hep-ph/0101032`. `doi:10.1103/PhysRevD.65.014013`. (Cited on page 115.)

[75] D. Bourilkov and M. R. Whalley, "LHAPDF : PDF Use from the Tevatron to the LHC," May, 2006, `arXiv:hep-ph/0605240`. (Cited on page 115.)

[76] H. H. Hoang et al., "Infrared Renormalization Group Flow for Heavy Quark Masses," *Phys. Rev. Lett.* **101** (2008), 151602, `arXiv:0803.4214`. (Cited on page 135.)

[77] N. Kidonakis, "Next-to-next-to-leading soft-gluon corrections for the top quark cross section and transverse momentum distribution," *Phys. Rev. D* **82** (2010), 114030, `arXiv:1009.4935`. (Cited on page 135.)

[78] **D0** Collaboration, "Determination of the pole and MSbar masses of the top quark from the ttbar cross section," *Phys. Lett. B* **703** (2011) 422 – 427, `arXiv:1104.2887`. (Cited on pages 136, 138 and 176.)

[79] **ATLAS** Collaboration, "Determination of the Top-Quark Mass from the ttbar Cross Section Measurement in pp Collisions at sqrt(s)=7 TeV with the ATLAS detector," 2011. ATLAS-CONF-2011-054. `cdsweb.cern.ch/record/1342551`. (Cited on pages 136, 138 and 176.)

[80] **CMS** Collaboration, "Measurement of ttbar Pair Production Cross Section at sqrt(s)=7 TeV using b-quark Jet Identification Techniques in Lepton + Jet Events," 2011. CMS-PAS-TOP-11-003. `cdsweb.cern.ch/record/1386709`. (Cited on page 139.)

[81] **ATLAS** Collaboration, "Measurement of the ttbar production cross-section in pp collisions at sqrts = 7 TeV using kinematic information of lepton+jets events," 2011. ATLAS-CONF-2011-121. `cdsweb.cern.ch/record/1376413`. (Cited on page 139.)

[82] **CMS** Collaboration, "Measurement of the $t\bar{t}$ Production Cross Section in pp Collisions at 7 TeV in Lepton + Jets Events Using b-quark Jet Identification," *Phys. Rev. D* **84** (2011) 092004. 39 p, arXiv:1108.3773. CMS-TOP-10-003. cdsweb.cern.ch/record/1376068. (Cited on page 139.)

[83] **ATLAS** Collaboration, "Measurement of the top quark-pair cross-section with ATLAS in pp collisions at sqrt(s) = 7 TeV in the single-lepton channel using b-tagging," 2011. ATLAS-CONF-2011-035. cdsweb.cern.ch/record/1337785. (Cited on page 139.)

[84] **CMS** Collaboration, "Combination of top pair production cross section measurements," 2011. CMS-PAS-TOP-11-024. cdsweb.cern.ch/record/1401250. (Cited on page 139.)

[85] **ATLAS** Collaboration, "Statistical combination of top quark pair production cross-section measurements using dilepton, single-lepton, and all-hadronic final states at sqrt(s) = 7 TeV with the ATLAS detector," 2012. ATLAS-CONF-2012-024. cdsweb.cern.ch/record/1430733. (Cited on page 139.)

[86] **CDF** Collaboration, "Combination of CDF top quark pair production cross section measurements with 2.8/fb," 2008. CDF note 9448. www-cdf.fnal.gov/physics/new/top/confNotes/. (Cited on page 142.)

[87] **D0** Collaboration, "Measurement of the ttbar production cross section using dilepton events in ppbar collisions," *Phys. Lett. B* **704** (2011) 403–410. doi:10.1016/j.physletb.2011.09.046. (Cited on page 142.)

[88] V. Abazov et al., "Measurement of the t[overline t] production cross section in p[overline p] collisions at sqrt(s)=1.96 TeV using kinematic characteristics of lepton+jets events," *Physical Review D* **76** (2007) 1550–7998. doi:10.1103/PhysRevD.76.092007. (Cited on page 144.)

[89] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations," *JHEP* **06** (2002). doi:10.1088/1126-6708/2002/06/029. (Cited on page 144.)

[90] G. Corcella et al., "HERWIG 6.5," *JHEP* **0101** (2001) arXiv:hep-ph/0011363. (Cited on page 144.)

[91] cmsweb.cern.ch/das/. (Cited on page 149.)

[92] cms-service-dqm.web.cern.ch/cms-service-dqm/CAF/. (Cited on page 149.)

Acknowledgements

It has been a privilege to work as part of Ghent University and the CMS Collaboration. Over the past four years there have been many people whose expertise and discussions were integral to the development of this research. In the following paragraphs I will attempt to express my gratitude to the colleagues who contributed to the successful completion of this thesis.

I would like to extend my sincerest thanks to Prof. Martin Grünewald for the opportunity to work under his supervision. I am very appreciative of his expertise, advise and discussions.

I would also like to thank Prof. Dirk Ryckbosch for the chance to work in the Department of Physics and Astronomy at Ghent University and for creating a collaborative group environment.

I would like to thank the members of my thesis defense committee for taking the time to scrutinise my thesis and for providing valuable comments and feedback.

I would like to thank Dr. Ivan Mikulec, Dr. Silvia Goy Lopez and Dr. Slava Valuev for providing insight into the operation of the muon trigger system at CMS. It has been a pleasure working with them in monitoring and understanding the behaviour of the triggers.

I would like to thank the administrators of the computing facilities at the IIHE.

Finally, I would like express my gratitude to the members of the Ghent top group who have contributed to this analysis.