Top Pair Production Cross Sections and Differential Cross Sections in the Semi-Leptonic channel using the CMS detector at $\sqrt{s} = 7$ and 8 TeV

Philip Symonds

School of Engineering and Design
Brunel University

A thesis submitted for the degree of
Doctor of Philosophy

Monday 2nd March, 2015

Acknowledgements

The work in this thesis would not have been possible without the help and support of my friends, family and colleagues to whom I am very grateful. I would firstly like to thank Ivan Reid and Matt Barrett who helped me get up and running with analysis work when I was starting out in 2010. I must also thank Enrique Palencia and Martijn Mulders who were always at hand once I was out at CERN.

It was a real pleasure to then begin work on the differential cross section analysis with my friends: Sergey, Jeson and Luke from Bristol University. Sharing an office at CERN really encouraged us to work well together and form a great team. The guidance of Greg Heath and Joel Goldstein was also very much appreciated. Their help along with that of our analysis review committees (ARCs) really helped us get our 2011 and 2012 analyses through the CMS approval procedure.

Throughout my time at CERN, I thoroughly enjoyed working on the test beam experiments. Even the Saturday morning graveyard shift (12am-8am). I'd like to thank Sasha for getting me involved in this work and Marco for the analytical support. CERN can feel like a bit too much of an office environment at times and it's nice to see where the physics actually happens.

Many thanks also go to my supervisor, Akram Khan, who really put faith in me and drove me on to succeed. Although I didn't always appreciate being put in daunting situations at times, I have gained real value from it. I must thank Jo Cole, my second supervisor, who I'm particularly grateful to for reading the first draft of my thesis. It must have been a very painful experience with all of the typos.

This research would have been a bit of a struggle without the funding from the Science and Technology Facilities Council (STFC). Their financial support helped me not only achieve my physics goals but also my skiing goals during the weekends out at CERN.

Finally, a big thanks goes out to all of the friends I made out in Geneva. I met so many amazing people that I can't name them all. You all gave me a truly memorable experience which I will never forget.

Abstract

The top quark has been extensively studied since the Large Hadron Collider (LHC) began operating in 2010. The excellent performance of both the LHC and the Compact Muon Solenoid (CMS) detector has enabled complex analyses of many properties of the top quark. In this thesis inclusive and differential top pair ($t\bar{t}$) production cross sections have been measured. Inclusive $t\bar{t}$ cross sections of 145.6 $\pm$ 8.2 (stat.) $^{+38.1}_{-28.3}$ (syst.) pb and 237.4 $\pm$ 1.9 (stat.) $^{+20.4}_{-16.9}$ (syst.) pb were measured at 7 TeV and 8 TeV center-of-mass collision energies using luminosities of 1 fb^{-1}and 19.7 fb^{-1}, respectively. These measurements were performed in the semi-leptonic channel by means of a maximum likelihood fit of the lepton's pseudorapidity. The work in this thesis focuses specifically on the muon-plus-jets channel.

The methods used for measuring the inclusive cross sections were built upon to measure differential cross sections with respect to event level observables. These observables include the missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) as well as some other kinematic distributions involving the jets, lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$ in the decay. These results are unfolded to remove detector and selection effects and have uncertainties in the range of 3% to 15%. A low uncertainty is achieved by normalising the differential cross section using the total cross section. This leads to cancellations of some uncertainties. The results were compared with different Monte Carlo generators and with different input parameters. No significant deviations from predictions of the Standard Model were observed.

This thesis also contains test beam results on CMS ECAL Endcap Lead Tungstate ($PbWO_4$) crystals. These crystals had been damaged using various doses of proton irradiation. The damage for some crystals is expected to be roughly equivalent to 300 fb^{-1} of integrated luminosity at $\sqrt{s} = 14$ TeV. The energy resolution for these crystals was seen to reduce by close to a factor of 20.

Contents

Chapter 1

Introduction

Particle physics is the study of the fundamental constituents of matter and the forces that govern how these particles interact with one another. Although the idea that all matter is composed of fundamental particles dates back to the 6^{th} century BC, a revolution in the field took place in the 20^{th} century. This revolution has continued into the 21^{st} century. Breakthroughs in both experimental and theoretical work have led to the concise picture of the Standard Model (SM) of particle physics that is seen today.

In this thesis, measurements of the inclusive top pair ($t\bar{t}$) production cross section at both 7 and 8 TeV will be presented. The top quark is an important part of the SM. It has been studied extensively at both the Tevatron and since the start up of the Large Hadron Collider (LHC). The excellent performance of the LHC machine opened up the possibilities of differential cross section measurements due to the millions of top quark pairs being produced. Global event-level variables have been studied such as the missing transverse energy ($E_{\text{T}}^{\text{miss}}$); the jet transverse momentum sum (H_{T}); the total observed transverse momentum sum (S_{T}); the transverse mass (M_{T}^{W}) and transverse momentum (p_{T}^{W}) of the W boson. These measurements allow the comparison and verification of Monte Carlo generators which aim to model top quark pair events. Knowing these distributions precisely will also help for future observations of rare Standard Model (SM) processes possible such at $t\bar{t}H$. This chapter introduces our current understanding of particle physics

to set the scene for the work presented in this thesis.

1.1 The Standard Model of Particle Physics

The Standard Model (SM) of particle physics is the theoretical framework that attempts to bring together all of what we know about the fundamental building blocks of the universe. The SM is an evolving theory which has been put together in a collaboration between many theorists and experimentalists. The main components of the SM are the electroweak theory and quantum chromodynamics (QCD) which together describe the fundamental interactions at the quantum scale.

The SM of particles and forces consists of both the fermions and the gauge bosons. Fermions have half-integer ($\frac{1}{2}\hbar$, $\frac{3}{2}\hbar$, $\frac{5}{2}\hbar, \ldots$) spin where $\hbar$ is the reduced version ($= \frac{h}{2\pi}$) of Planck's constant which is used to quantise spin. Planck's constant is equal to $6.62 \times 10^{-34}\mathrm{J} \cdot s$. In this thesis, natural units will be used where $\hbar = c = 1$. Here, c is the speed of light $\approx 3{\times}10^{8}\mathrm{m}s^{-1}$.

Fermions obey the Pauli exclusion principle which states that no two particles can be in the same quantum state. This means that they obey Fermi-Dirac statistics [1]. The fermions in the Standard Model and some of their key properties are shown in Table 1.1. Fermions consist of quarks and leptons. These particles are all spin-$\frac{1}{2}$ and have electric charge defined in terms of the elementary charge constant, e. Elementary charge units are defined as the charge of a proton (or a positron) where $e = 1.602 \times 10^{-19}$ C.

Exchanges of gauge bosons between fermions is what allows matter particles to interact with one another. Bosons have integer spin ($0\hbar$, $1\hbar$, $2\hbar$, ...) and obey Bose-Einstein statistics [2]. The gauge bosons within the SM are presented in Table 1.2. The electromagnetic (EM) interaction is mediated by the photon (γ) which couples to particles with electric charge. The weak interaction is mediated by the massive W and Z vector bosons. These bosons can interact with one another as well as with the Higgs boson. The high masses of these particles mean that the force is short ranged. The strong force is mediated by the massless gluon (g) which carries colour. Since gluons possess colour they are able to self interact which gives the strong interaction

Flavor	Symbol	Charge (e)	Mass (GeV/c^2)
Quarks			
Up	u	2/3	$\approx 2.3 \times 10^{-3}$
Down	d	-1/3	$\approx 4.8 \times 10^{-3}$
Charm	c	2/3	≈ 1.275
Strange	s	-1/3	$\approx 95 \times 10^{-3}$
Top	t	2/3	≈ 173.20
Bottom	b	-1/3	≈ 4.18
Leptons			
Electron	e	-1	0.511×10^{-3}
Electron Neutrino	ν_e	0	$\leq 2.2 \times 10^{-6}$
Muon	μ	-1	105.7×10^{-3}
Muon Neutrino	ν_μ	0	$\leq 0.17 \times 10^{-3}$
Tau	τ	-1	1.777
Tau Neutrino	ν_τ	0	$\leq 15.5 \times 10^{-3}$

Table 1.1: The fermions within the Standard Model of particle physics as well as their electric charges and masses.

its very short range. All of these force mediators are of spin-1. The force of gravity is not currently included within the Standard Model as it is hard to mathematically describe at the quantum scale. Gravity has a negligible effect at atomic length scales so does not effect the theory.

1.1.1 The Electroweak Interaction

The concept of a SM of particle physics first came about in the 1960s through the unification of the electromagnetic and weak forces [3, 4, 5]. The theory of quantum electrodynamics (QED) already provided an excellent description of the EM force. Unifying the EM and weak forces however, provides a more complete theory. This unification comes in the form of the gauge group SU(2)×U(1) and ensures that the electromagnetic and weak coupling constants are correlated.

The photon and the massive vector bosons which mediate the interactions are first introduced together as four massless fields $(W^1_\mu, W^2_\mu, W^3_\mu, B_\mu)$. The field triplet, W_μ, transforms via the special unitary group, SU(2). The

Name	Symbol	Charge (e)	Spin	Mass (GeV/c^2)
Strong Force				
Gluon	g	0	1	0
Electroweak Force				
Photon	γ	0	1	0
Z boson	Z^0	0	1	91.2
W boson	$W^{\pm}$	± 1	1	80.4
Electroweak Symmetry Breaking				
Higgs	H	0	0	≈ 125

Table 1.2: The Bosons within the Standard Model.

weak field triplet interacts with the third component of isopsin (I_{W_z}) of the particles. Isospin is an abstract quantum property of a particle that is always conserved in weak interactions. The B_μ field transforms unitary group, U(1), and interacts with the weak hypercharge of the particles. Weak hypercharge is another conserved quantum property defined as: $Y_W = 2(Q - I_{W_z})$ where Q is the charge of the particle.

The physical fields that mediate the forces are composed of linear combinations of the massless fields $(W^1_\mu, W^2_\mu, W^3_\mu, B_\mu)$. In the the charged current (CC) weak interaction, $W^{\pm}$ consists of the superposition:

$$W^{\pm} = \frac{1}{\sqrt{2}}(W^1 \mp iW^2).$$

The field W_μ of SU(2) couples to the weak isospin current with strength proportional to the coupling constant, g. An additional coupling constant, g', is introduced for U(1) to describe the strength of the coupling of B_μ to the weak hypercharge current. The physical fields (Z^0 and A) which mediate the weak neutral current (NC) and electromagnetic interactions can be constructed as follows:

$$\begin{pmatrix} Z^0 \\ A \end{pmatrix} = \frac{1}{\sqrt{g^2 + g'^2}} \begin{pmatrix} g & -g' \\ g' & g \end{pmatrix} \begin{pmatrix} W^3 \\ B \end{pmatrix} = \begin{pmatrix} \cos\theta_W & -\sin\theta_W \\ \sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} W^3 \\ B \end{pmatrix}.$$

The angle θ_W (weak mixing angle) quantifies the relationship between the

two coupling constants:

$$\theta_W \equiv \tan^{-1} \frac{g'}{g} \tag{1.1}$$

The unification of the electromagnetic and the weak force can be seen in the expression:

$$g \sin \theta_W = \frac{e}{\sqrt{\epsilon_0 \hbar c}} = \sqrt{4\pi\alpha} \tag{1.2}$$

where ϵ_0 is the permittivity of free space and α is the fine structure constant which is the coupling coefficient of the EM force ($\sim 1/137$ at low energy scales). The unification of EM and weak forces allows these interactions to be expressed using just a few fundamental parameters. The model does not predict the value of θ_W. This has to be determined from experiment or from global SM fits. θ_W has been measured through a variety of direct and indirect methods at different energy scales. A value of $\sin^2 \theta_W = 0.23116$ was obtained through a global fit [6] of standard model parameters using experimental constraints.

The decays of W bosons maximally violate parity conservation in that left-handed particles are always produced. The fermions within the SM therefore come in the form of left-handed isospin doublets:

$$\begin{pmatrix} I_{W_z} = +1/2 \\ I_{W_z} = -1/2 \end{pmatrix} = \begin{pmatrix} \nu_{eL} \\ e_L^- \end{pmatrix}, \begin{pmatrix} \nu_{\mu L} \\ \mu_L^- \end{pmatrix}, \begin{pmatrix} \nu_{\tau L} \\ \tau_L^- \end{pmatrix}, \begin{pmatrix} u_L \\ d'_L \end{pmatrix}, \begin{pmatrix} c_L \\ s'_L \end{pmatrix}, \begin{pmatrix} t_L \\ b'_L \end{pmatrix}.$$

Each quark doublet has a colour which is the charge carried by the strong force. Colour comes in three flavours: red, green and blue (R, G, B) and so there are 12 doublets in total including the leptons. Particles within a left handed doublet have equal hypercharge. The lower members of the quark doublets are shown with a prime to represent a rotated quark state. This shows that mixing can occur between different quark generations through the CC weak interaction. Kobayashi and Maskawa introduced the third generation of quarks as a way of explaining observations of charge-parity (CP) violation [7]. Quark mixing between generations is defined by the

Cabibbo-Kobayashi-Maskawa (CKM) matrix [8] which transforms as:

$$\begin{pmatrix} u \\ c \\ t \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} 0.974 & 0.225 & 0.004 \\ 0.225 & 0.973 & 0.041 \\ 0.009 & 0.040 & 0.999 \end{pmatrix} \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix}.$$

On the left are the up-type quarks (u, c and t). The components of the matrix, V_{ij}, represent the probability of the transition to down-type quarks (d', s' and b') on the right. The CKM matrix is unitary such that:

$$\sum_k V_{ik} V_{jk}^* = 0 \tag{1.3}$$

and it can be parameterised [7] using three angles (θ_{12}, θ_{23} and θ_{13}) to quantify the quark mixing between the different generations as well as a phase factor (δ) to account for CP violation. Within the SM right-handed singlets also exist with $I_W = 0$ and can interact in NC interactions:

$$e_R^-, \mu_R^-, \tau_R^-, d_R, u_R, s_R, c_R, b_R, t_R.$$

Right-handed neutrinos do not exist since they would possess both zero isospin and zero hypercharge and so would not interact with any of the force carriers of the SM. Right-handed anti-particle doublets and left handed anti-particle singlets also interact via the weak interaction.

The spontaneous symmetry breaking [9] mechanism of the Higgs field gives mass to the $W^\pm$ and Z^0, whilst leaving the photon massless. This is what gives the weak interaction its finite range whilst the photons range is infinite. The discovery of the Higgs boson by both the CMS and ATLAS experiments in the summer of 2012 [10] was the SM's most recent success. Two of the key decay channels that lead to the discovery of the Higgs boson where the di-photon ($\gamma\gamma$) and ZZ channels. Distributions of the invariant mass of $\gamma\gamma$s and ZZs can be seen in Figure 1.1. A clear excess can be seen at a mass of around 125 GeV indicating the existence of the Higgs. Further studies are required to determine the properties of this new particle such as its branching ratios in all decay channels. Initial indications show that its

properties are consistent with SM predictions.

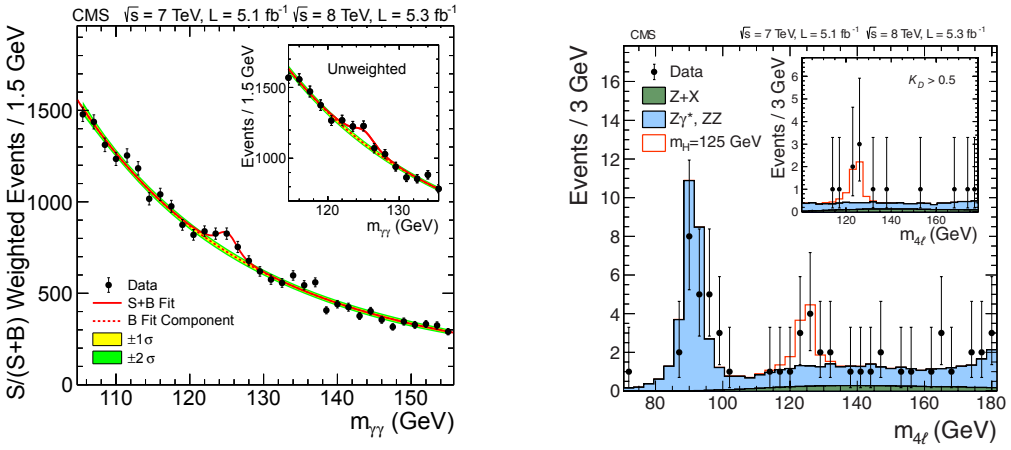

Figure 1.1: The invariant mass of two photons in the $H \to \gamma\gamma$ analysis (left). In the main figure the events are weighted by the signal/(signal+background) that the event is catergorised as. The colour bands represent the ± 1 and ± 2 standard deviation uncertainties on the background estimate. The invariant mass of the four leptons in the $H \to ZZ$ channel (right). The inset shows the distribution after selecting events with $K_D > 0.5$ where K_D is the S/(S+B) probability ratio of the events. Taken from the CMS Higgs discovery paper [10].

1.1.2 The Strong Interaction

The strong interaction is represented by the gauge symmetry group SU(3) and only effects particles that possess a colour charge [11, 12]. The underlying theory in this interaction is described by the quantum chromodynamic (QCD) field theory. The three colour charges (red (R), green (G) and blue (B)) come in both positive and negative varieties (e.g the red colour charge has a corresponding anti-red form). Quarks are the only fermions that possess a colour charge. Gluons also carry colour themselves which enables self interactions leading to the finite range of the strong force. This means that individual quarks can not exist on their own and must be confined to hadrons (hadronisation). In the strong interaction, colour is a conserved quantity whilst the force is independent of both the electric charge and quark flavour.

Hadrons are composed of bound colour-neutral states which interact within the fields of massless gluons. Mesons are composed of a coloured quark and an anti-quark which possesses the corresponding anti-colour. Baryons are a three quark system with each quark carrying a different colour. Together they are colour neutral. The SU(3) symmetry group can be represented by eight gluons which carry the colour charge:

$$g_1 = R\bar{G} \qquad g_2 = R\bar{B} \qquad g_3 = G\bar{R} \qquad g_4 = G\bar{B} \qquad g_5 = B\bar{R}$$

$$g_6 = B\bar{G} \qquad g_7 = \frac{1}{\sqrt{2}}(R\bar{R} - G\bar{G}) \qquad g_8 = \frac{1}{\sqrt{6}}(R\bar{R} + G\bar{G} - 2B\bar{B}).$$

Protons are composed of two up quarks and a down quark ("valence" quarks). The mass of these three quarks only constitutes roughly 1% of the proton's mass. The remainder comes from the potential energy contained within the gluon fields that hold the the quarks together. Gluons within a proton are continually producing $q\bar{q}$ pairs. These non-valence quarks and gluons within a proton are referred to as a "sea" of partons. The word parton was initially introduced by Feynman to describe the hard objects within nucleons [13]. Understanding the composition of the proton is crucial when calculating event production rates at the LHC as will be described in the next chapter.

1.2 Summary

The SM of particle physics provides a description of the fundamental forces which are at play in the quantum world. The theory has survived many of the tests it has faced over the past half century. It is able to describe to a high precision what has been observed by many particle physics experiments. This includes the prediction of the existence of the Higgs Boson. Since the Standard Model Higgs Boson couples to fermions with strength proportional to the fermion's mass, the Higgs coupling to the top quark is large. The top quark, being the most massive particle discovered so far will therefore play an important role in future Higgs measurements.

Several signatures of new physics accessible at the LHC either suffer from top-quark production as a significant background or contain top quarks themselves. Knowing the top quark pair production cross section to a high precision is therefore not only important as a test of the Standard Model but also in knowing how it contributes as a background to rarer processes.

Chapter 2

The Top Quark

The top quark is a fundamental building block of the Standard Model and was first observed in $p\bar{p}$ collisions at the Tevatron in 1995 using the CDF [14] and DØ [15] detectors. The top quark is the most massive of the known Standard Model particles with a current best measured value from a LHC and Tevatron combination of $m_t = 173.30 \pm 0.76$ [16]. Due to its large mass, it decays rapidly with a lifetime (τ_t) of $\sim 5 \times 10^{-24}$ seconds. This is before it can form a bound-state hadron ($\tau_{\text{hadronisation}} \approx 10^{-23}s$). This means that the study of the top quark's decay products allow direct access to various properties such as its mass, spin and charge in a way that is inaccessible for any of the other known quarks.

2.1 Top Quark Production

At the LHC, top pair production proceeds primarily through gluon-gluon (gg) fusion ($\sim 90\,\%$), with a small contribution from quark-antiquark ($q\bar{q}$) annihilation ($\sim 10\,\%$). There is also a small (almost negligible) contribution from quark (qg) and anti-quark ($\bar{q}g$) scattering with gluons. Leading order Feynman diagrams for top pair production are shown in Figure 2.1. Top quarks can also be produced singly which serves as an irreducible background to $t\bar{t}$ due to the similar decay products being produced. Feynman diagrams for leading order single-top production are shown in Figure 2.2.

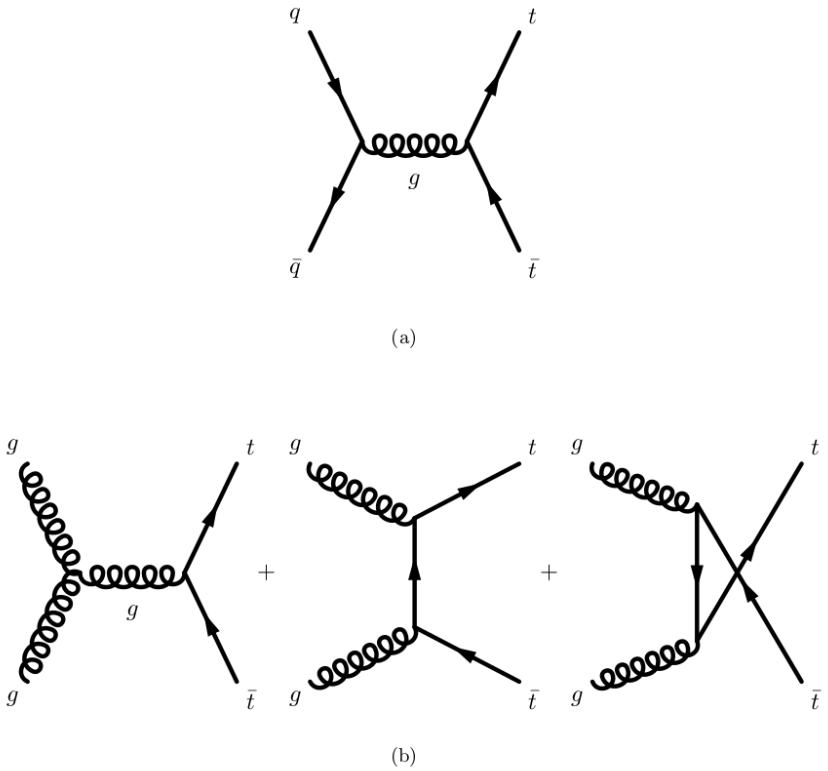

Figure 2.1: Schematic of leading order (LO) top pair production from quark-antiquark annihilation (a) and gluon-gluon fusion (b) in the s (space-like), t (time-like) and u (time-like but with a switch in the out going particles) channels.

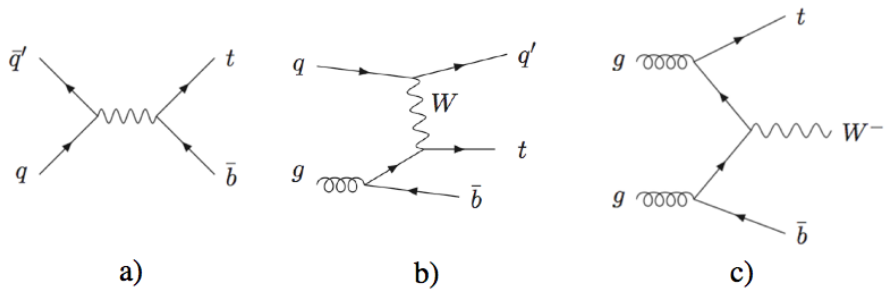

Figure 2.2: The three main LO single-top production mechanisms. The image shows (a) s-channel, (b) t-channel and (c) associated W single-top production (tW).

In top pair production the following hard scattering process is considered

$$i(p_1) + j(p_2) \rightarrow t(p_3) + \bar{t}(p_4) + X$$

where i and j represent the incoming partons $(q, \bar{q}, g)$ and X is additional radiation produced in the final state. The differential cross section is a sum over all possible spin and colour states of the incoming partons and can be calculated using:

$$d\sigma = \sum_{i,j \in q,\bar{q},g} \int dx_1 dx_2 f_1(x_1, \mu_F) f_2(x_2, \mu_F) d\hat{\sigma}_{ij}(x_1, x_2, \mu_F, \mu_R, \alpha_s(\mu_R)),$$

where μ_R and μ_F is the renormalisation and factorisation scales respectively, $\alpha_s(\mu_R)$ the strong coupling coefficient and $f_i(x_i, \mu_F)$ the parton distribution function of the parton, i. In this form, x_i, is the fraction of the momentum that a parton, i, carries with respect to the momentum of the proton.

The factorisation scale characterises the boundary between short (parton-level) and long (hadronisation) distances and is in place such that theoretical predictions are infra-red safe. Infra-red safe means that the calculation is stable against divergences due to soft gluon emissions which arise in QCD perturbation theory. The theoretical calculation can therefore be comparable to experimental measurements which are independent of μ_F. The renormalisation scale is in place as a consequence of ultra violet divergences that can occur in QCD perturbation theory due to the running of the coupling constant. The strong coupling constant decreases as a function of the momentum transfer in the collision so the correct energy scale needs to be chosen. In the case of top quark production both μ_R and μ_F are chosen as the hard-scattering scale squared, Q^2, and is set to the invariant mass of the produced particle, m_{T}.

The partonic cross section, $d\hat{\sigma}_{ij}$, can be expanded to a certain order within perturbative QCD

$$d\hat{\sigma}_{ij} = \alpha_s^2 \left[d\hat{\sigma}_{ij}^{(0)} + \frac{\alpha_s}{\pi} d\hat{\sigma}_{ij}^{(1)} + \frac{\alpha_s^2}{\pi^2} d\hat{\sigma}_{ij}^{(2)} \ldots \right].$$

The first term in the square bracket is the leading-order (LO) term and the second and third terms the next-to-leading-order (NLO) and next-to-next-to-leading-order (NNLO) terms respectively.

An important factor in the calculation of the production cross section for top pairs is the partonic composition of the incoming protons. This is characterised by the parton distribution functions (PDFs, $f_i(x_i, \mu_F)$) and is represented as a the number density of a given parton, i, with a certain momentum fraction within a proton. PDFs are determined using both theoretical and experimental input from deep-inelastic scattering experiments such as HERA [17]. The PDFs for a (squared) hard-scattering scale, $Q^2 = 10000$ GeV^2, are shown in Figure 2.3. This momentum transfer is of the order required for top quark or W and Z boson production. At high Q^2 the sea of quarks (xS) and gluons (xg) become resolvable and it is these gluons that are mostly responsible for producing $t\bar{t}$ pairs. CTEQ [18] PDF sets are commonly used in the calculations of the hard scattering process although others are available.

As a consequence of the large mass of the top quark, partons with a relatively high momentum fraction are required for its production. The momentum fractions of the two incoming partons must satisfy the following:

$$\sqrt{x_1 x_2 s} = 2m_{\mathrm{T}}$$

At a center-of-mass energy, $\sqrt{s} = 7$ TeV, back-to-back top pair production in the lab frame would take place when $x_1 = x_2 = 0.05$. The tops are produced at smaller angles with respect to one another when one of the momentum fractions is significantly larger than the other. The minimal angle at which tops can be produced with respect to one another is about 15° which would happen if $x_1 = 1$ and $x_2 = 0.002$ or vice versa.

Top quarks are usually produced quite centrally within the detector (i.e. with a large angle with respect to the beam line). This is the result of them being spin-half particles and hence their angular distribution follows $(1+\cos^2\theta)$. Due to its rapid decay, the angular distribution of the top quarks is passed on to the decay products. This results in $t\bar{t}$ events predominantly

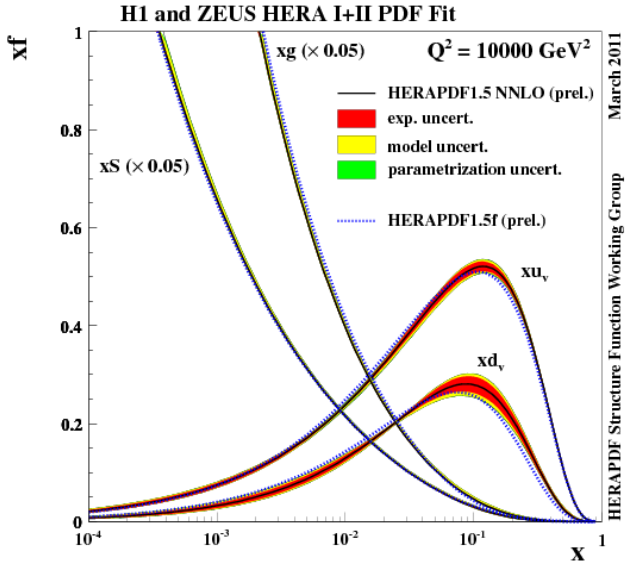

Figure 2.3: Parton distribution function (PDFs) as a function of parton momentum fraction for $Q^2 = 10000$. The number density, xf, is plotted against the momentum fraction, x, of the partons. At low x the sea of quarks (xS) and gluons (xg) dominate. It's only at higher x that the valence quarks (xd_{v} and xu_{v}) come into play. Image taken from [19].

occurring quite centrally within the detector. Lower mass particles such as the W and Z bosons can be produced by a broader range of x values. Therefore these particles are less centrally produced. This is an important factor and is used later on in the process to separate $t\bar{t}$ from the main background processes.

The top quark pair ($t\bar{t}$) production cross section has been precisely measured at three different centre of mass energies. The CDF and DØ collaborations made the measurement at $\sqrt{s}$ = 1.96 TeV in $p\bar{p}$ collisions at the Tevatron [20]. In addition to this, measurements by ATLAS and CMS at $\sqrt{s}$ of 7 and 8 TeV in proton-proton collisions have been performed [21, 22]. A summary of these results and a comparison to NNLO (Next-to-next-to-leading-order) theoretical predictions are shown in Figure 2.4. Recent numerical values of theoretical predictions and experimental measurements of $t\bar{t}$ and single-top production cross sections at both 7 and 8 TeV are summarised in Table 2.1.

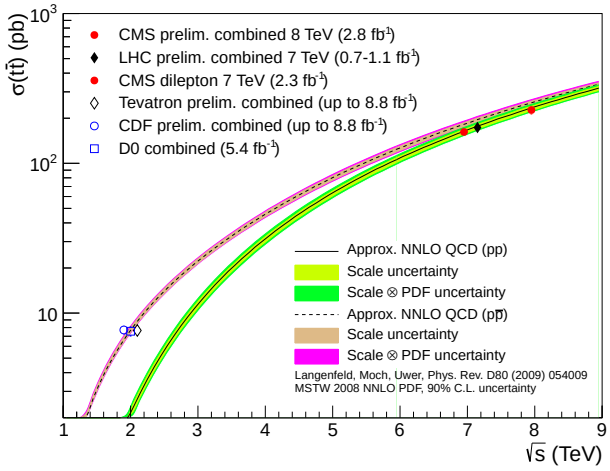

Figure 2.4: A summary of $t\bar{t}$ cross section results measured as a function of $\sqrt{s}$. Measurements of the CMS experiment at 7 and 8 TeV are shown along with the preliminary combined result with the ATLAS measurement. The CDF, DØ and the combined Tevatron results at $\sqrt{s}$= 1.96 TeV are also shown. Taken from [23].

Measured and theoretical top production cross sections (pb)				
Production	7 TeV th.	7 TeV meas.	8 TeV th.	8 TeV meas.
$t\bar{t}$	$163^{+7}_{-5} \pm 9$	$161.9 \pm 2.5^{+6.2}_{-6.1}$[21]	$234^{+10}_{-7} \pm 12$	$239 \pm 2 \pm 13$ [24]
Single-t (t-chan)	$65.9^{+2.1+1.5}_{-0.7-1.7}$	$67.2 \pm 3.7 \pm 4.8$[25]	$87.2^{+2.8+2.0}_{-1.0-2.2}$	$85 \pm 4 \pm 11$ [26]
Single-t (tW)	$15.6 \pm 0.4^{+0.5}_{-0.6}$	N/A	$22.2 \pm 0.6 \pm 1.4$	23.4 ± 5.5 [27]
Single-t (s-chan)	$4.56 \pm 0.07^{+0.18}_{-0.17}$	< 26.5@95% [28]	$5.55 \pm 0.08 \pm 2.21$	N/A

Table 2.1: This table displays the measured and theoretical top production cross sections. The 7 and 8 TeV theoretical values have been obtained from approximate NNLO calculations [29]. The first uncertainty on theoretical values is due to the choice of scale (renormalisation and factorisation) and the second due to the PDF uncertainty. The measured values have a statistical and a systematic error associated with them respectively.

2.2 Top Quark Decays

Within the Standard Model, the top quark is expected to decay via the weak process $t \to Wb$ approximately 100 % of the time. The decay amplitude is proportional to the $|V_{tb}|^2$ component of the CKM matrix. A value of $|V_{tb}| = 0.999146^{+0.000021}_{-0.000046}$ [30] was determined by performing a global fit of the SM using various experimental constraints and requiring unitarity of the CKM matrix. CMS has made a recent measurement of $|V_{tb}| = 0.998 \pm 0.041$ [31] using the single-top (t-channel) decay channel. It is also possible to make a direct measurement without CKM unitarity using single-top quark production in the s-channel. CDF measured a value of $|V_{tb}| = 0.91 \pm 0.11(\text{sys.} + \text{stat.}) \pm 0.7(\text{theory})$ [32] and set a lower limit of $|V_{tb}| > 0.71$ at the 95% confidence level.

The various $t\bar{t}$ decay topologies are therefore characterized by the decays of the two W bosons in the event. W bosons decay both hadronically (qq') and leptonically ($l\bar{\nu}_l$) giving yield to three distinct decay channels:

- Fully-leptonic: both Ws in the event decay leptonically

- Semi-leptonic: one W boson decays leptonically and the other hadronically
- Fully-hadronic: both Ws in the event decay hadronically

The various branching ratios (BRs) for these processes with the different lepton contributions are shown in Table 2.2. The hadronic decay BRs contain a factor of three coming from the different quark colours.

		W^+			
		$c\bar{s}, u\bar{d}$ ($\approx$6/9)	$e^+\nu_e$ ($\approx$1/9)	$\mu^+\nu_\mu$ ($\approx$1/9)	$\tau^+\nu_\tau$ ($\approx$1/9)
W^-	$\bar{c}s, \bar{u}d$ ($\approx$6/9)	45.7%	7.3%	7.3%	7.3%
	$e^-\bar{\nu}_e$ ($\approx$1/9)	7.3%	1.2%	1.2%	1.2%
	$\mu^-\bar{\nu}_\mu$ ($\approx$1/9)	7.3%	1.2%	1.2%	1.2%
	$\tau^-\bar{\nu}_\tau$ ($\approx$1/9)	7.3%	1.2%	1.2%	1.2%

Table 2.2: Measured branching ratios [6] of the various $t\bar{t}$ decay channels as a result of the different W boson decays.

2.2.1 The Fully-Leptonic Decay Channel

This decay channel is characterised by the presence of two opposite sign leptons which give a very clean event signature. The leptons which are usually used are electrons and muons and so there are three distinct sub channels: $ee, \mu\mu$ and $e\mu$. Tau particles can also be used but measurements using these leptons are usually less precise. This is because they are heavy enough to decay hadronically* via the weak interaction making it difficult to separate from the fully-hadronic and semi-leptonic decays. Large amounts of missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) are expected in these events due to the presence of the two undetected neutrinos from the W decays. At least two jets are expected to be present from the b quarks.

*Via $\tau \to \nu_\tau + W(\to u\bar{d})$, this accounts for roughly 65% of τ decays

2.2.2 The Semi-Leptonic Decay Channel

The final state is defined by the presence of a single, isolated, high transverse momentum (p_{T}) lepton. As with the di-lepton channel, $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected from the presence of the neutrino in the event. At least four central and energetic jets are expected, two of which originate from b quarks. This channel is sometimes referred to as a golden channel since it is relatively clean and has a higher branching fraction than in the di-lepton case. Figure 2.5 shows the LO Feynman diagram for a $t\bar{t}$ decay in the muon-plus-jets channel. The work in this thesis was performed using the muon-plus-jets decay channel.

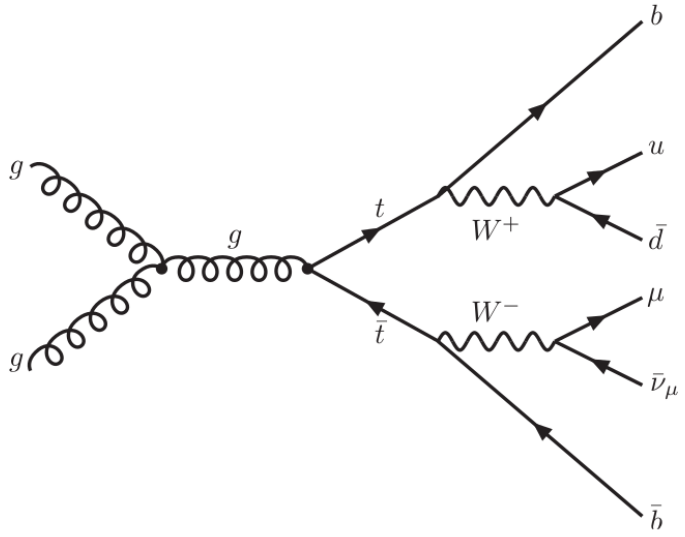

Figure 2.5: Feynman diagram illustrating the $t\bar{t}$ decay topology in the semi-leptonic (μ+jets) decay channel.

2.2.3 The Fully-Hadronic Decay Channel

In this channel both of the W bosons decay hadronically and so no isolated leptons are expected to be present in the decay. All decay products are likely to be jets of which there should be at least six. This channel suffers badly from QCD multi-jet background contamination. It is hard to separate this large background from the $t\bar{t}$ signal.

2.3 Event Simulation

Physics processes are simulated such that theoretical predictions can be compared to the data. Monte Carlo (MC) generators which utilise random numbers are used to do this. There are various MC generators available of which several are used by CMS to model specific physics processes. MCs are responsible for the production of the hard scattering processes via matrix element (ME) calculations and the resulting parton showering (PS). Parton matching is required to merge these two processes. Hadronisation also needs to be modelled as the showered partons form colourless bound states.

The MC generator gives a list of all the particles produced in an event and the corresponding kinematic properties of the particles. This includes the production of the underlying event (UE) and additional interactions from pile-up (PU). The underlying event is defined as anything in the event that does not originate from the hard scattering process. This includes interactions coming from the beam remnants and also initial and final state radiation. Pile-up is a result of multiple interactions taking place during a bunch crossing. Each interaction has it's own distinguishable primary vertex. Figure 2.6 shows a schematical overview of the MC production process. Geant4 [33] is used to simulate the CMS detector such that the particles produced in the MC are reconstructed in the same way as for the recorded data.

A list of the main MC generators is shown in Table 2.3. The default MC that is used for most CMS analyses is MADGRAPH [34]. POWHEG [35] and MC@NLO [36] are the alternative generators available on the market. All of these MCs are inclusive of all $t\bar{t}$ decays described in Section 2.2. The differences highlighted in the table will be summarised in the following section. One of the main reasons for performing differential cross sections is to try and verify which generator gives a better description of the data. The vast amounts of data available from the 2012 run makes it possible to test these models.

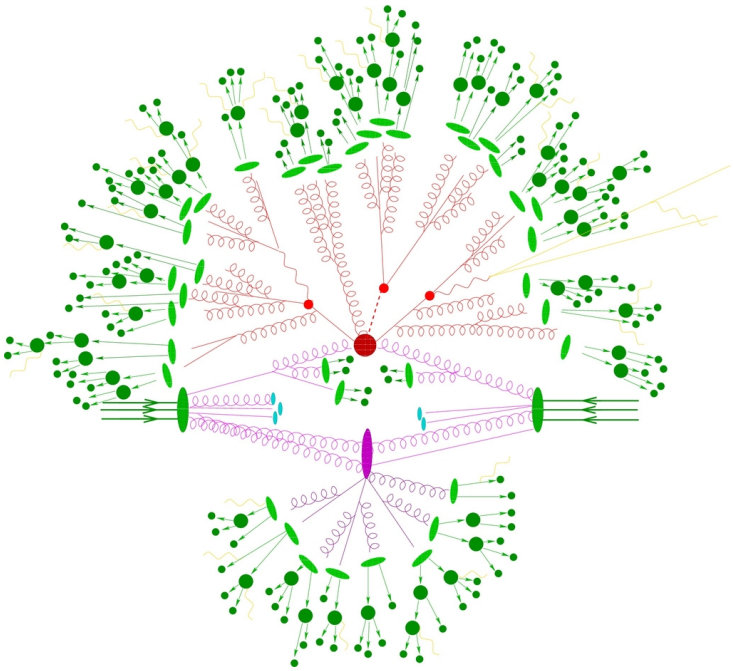

Figure 2.6: This image depicts how MC generation is performed. Coming in from the left and right are the two protons. Two gluons then enter the hard scattering interaction which is shown as the large red circle. This step represents the matrix element calculation. After the hard scatter prompt decays and parton showering then takes place as shown by the smaller red circles. Finally, hadronisation ensues which is shown in the diagram as the green ovals. We are then left with stable particles that are observed in the detector. The image also shows the underlying event which is initiated at the purple oval.

process	ME	PS	Method	PDFs
$t\bar{t}$ + jets	Madgraph v5.x	Pythia v6.42x	ME+PS	CTEQ6L1
$t\bar{t}$	POWHEG-box 1.0	Pythia v6.42x	NLO+PS	CTEQ6M
$t\bar{t}$	MC@NLO v3.41	Herwig v6.520	NLO+PS	CTEQ6M

Table 2.3: Table of the main Monte Carlo event generators.

2.3.1 Matrix Element (ME) Calculation

The general procedure of event simulation starts by determining all of the Feynman diagrams that contribute to a certain process. The incoming partons are generated using their parton distribution functions. The four-momenta of the outgoing partons is then predicted using fixed order matrix element calculations as was described in Section 2.1. Matrix element calculations are computationally expensive and are best suited to processes with just a few well separated hard partons. Due to this long computational time the highest order that MCs are calculated to is NLO (for MC@NLO and POWHEG). It is envisioned that NNLO generators will become available in the next few years. MADGRAPH uses a slightly different technique where LO calculations are performed where up to three additional partons can be included as either initial or final state radiation. This is represented as $t\bar{t}$ + jets in Table 2.3, where jets are the result of showered partons.

2.3.2 Parton Showering (PS)

Typically a different generator, better suited to soft QCD scattering interactions is used for the showering of final state quarks and gluons. PS is computationally cheap, has no limit on particle multiplicity and is valid for soft and collinear partons. Parton showering typically follows the following procedure:

1. A parton shower is initiated with a virtual mass scale, t_0, at the mass of the decaying particle. This parton has a momentum fraction, $z_0 = 1$.

2. At a later stage during the evolution of the shower the virtual mass scale becomes t_i and the momentum fraction z_i due to the emission of

radiation from a parton. A uniform random number between 0 and 1 is generated to replicate the probability of the parton splitting.

3. If $t_i < t_{\text{req.}}$, it signifies the end of the propagation of the shower. $t_{\text{req.}}$ is the minimum energy radiated by a parton and is usually defined to be ~ 1 GeV. This is the energy where hadronisation begins to dominate over perturbative QCD effects.

4. If this requirement is not reached a parton is radiated with $z = z_i/z_{i+1}$ with a distribution proportional to $(\alpha_s/2\pi)\hat{P}(z)$ where $\hat{P}(z)$ is the probability of the specific type of splitting (e.g qq, gq, qg, gg). The shower then continues to propagate repeating the process from step 1.

Parton showering is provided by two multi-purpose event generators. These being PYTHIA and HERWIG. These have a different ordering of the shower propagation. PYTHIA is transverse-momentum-ordered (starting with the highest p_{T}), whilst HERWIG uses angular ordering.

The physics of electromagnetic showers is modelled in a similar way as hadron showering. Electromagentic corrections are applied where soft photon emission becomes important. Electromagnetic radiaition from quarks is cut off at the typical hadronic scale.

2.3.3 Parton Matching Prescriptions

The matching (merging) between the initial hard partons produced at the matrix element and the evolution of the softer partons produced in the shower is a critical part of the MC. It is important that there is no double counting within the two schemes. To avoid the double counting of initial or final state radiation, phase space requirements are used between the two schemes. This usually includes a selection requirement on the p_{T} of the parton or a positional requirement. Jets are only matched to harder partons. The requirement on the energy of a parton that is matched to a jet is referred to as the matching threshold. The MADGRAPH generator uses a nominal requirement of 20 GeV such that only partons with an energy above this value initiate showers

producing jets. The MLM [37] approach to merging is the adopted method by the main generators mentioned in Table 2.3.

2.3.4 Hadronisation

After partons within a shower have evolved such that their energy is ~ 1 GeV, hadronisation into colourless baryons and mesons begins. As with PS, PYTHIA and HERWIG provide hadronisation modelling. These MC generators use two completely different models.

The Lund string model is adopted by PYTHIA. This is based on the assumption that the colour potential between two partons increases linearly with their separation for distances greater than about 1 fm. When the energy within the colour "string" holding the two partons together reaches a certain threshold it is energetically beneficial for a $q\bar{q}$ pair to be created out of the string. After the energy within the string has diminished, on-shell hadrons form out of the remaining colour connected hadrons.

HERWIG uses the cluster model where first, all gluons split into $q\bar{q}$ pairs. The remaining quarks that are in similar regions of phase space are then grouped into colourless clusters. These clusters are then sub-divided into smaller colourless clusters from which hadrons are formed. The decays of short-lived hadrons are modeled in both PYTHIA and HERWIG.

2.4 Summary

This chapter has introduced the top quark and has briefly summarised the theoretical tools used to make predictions about the top pair production rate. The MC generators that aim to replicate the data have also been outlined. These generators contain the matrix element calculation, parton showering and the hadronisation which together produces events as are expected to be observed in the detector.

Chapter 3

Experimental Apparatus

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is the highest energy particle accelerator ever built. It straddles the French/Swiss border and has a circumference of about 26.7 km. It was designed to run at a centre of mass energy of $\sqrt{s} =$ 14 TeV. A technical setback meant running finally began in 2010 at $\sqrt{s} = 7$ TeV through to the end of 2011. This was then increased up to 8 TeV for 2012 running. A scheduled upgrade is now in progress to bring the machine up to a centre of mass energy of 13 TeV which is due to be completed by early 2015.

3.1.1 The Path of the Protons

The layout of the LHC and all of its sub accelerators are shown in Figure 3.1. The protons are first extracted by injecting Hydrogen (H_2) gas into a Duoplasmatron cylinder that is surrounded by an electric field. This strips the electrons from the hydrogen atoms leaving the protons. The protons are then linearly accelerated by radio frequency (RF) cavities in LINAC 2 producing bunches of protons of energy 150 MeV. The bunches of protons then enter the first synchrotron called the booster. This hardware dates back to 1972 and accelerates the protons up to an energy of 1.4 GeV in 1.2 seconds. It also squeezes the bunches such that they have a smaller cross sectional

area. The protons are then injected into the Proton Synchrotron (PS) ring which has a radius 100 m. The protons leave this ring with an energy of 26 GeV within 3.2 seconds. From there, the protons make an underground trip into France where they enter the Super Proton Synchrotron (SPS). The SPS ramps the protons up to an energy of 450 GeV in 20 seconds. When not being used to fill the LHC, protons from the SPS can be sent off to various other experiments. The bunches of protons are then injected into the main LHC ring in two locations where they begin to circulate in opposite directions. Each beam was designed to contain 2808 bunches. The LHC still has some work to do to get the protons up to collision energy. RF cavities are again used to accelerate the protons up to higher energies whilst dipole magnets are used to bend the protons around the rings' 8 curved sections. Quadrupole magnets are used to keep the bunches tightly packed together. It takes around half an hour for both beams to become stabilised at full beam energy. A final shaping and cleaning of the beam takes place before collisions can begin. Table 3.1 contains a list of design parameters for the LHC. A more detailed description of the LHC can be found elsewhere [38].

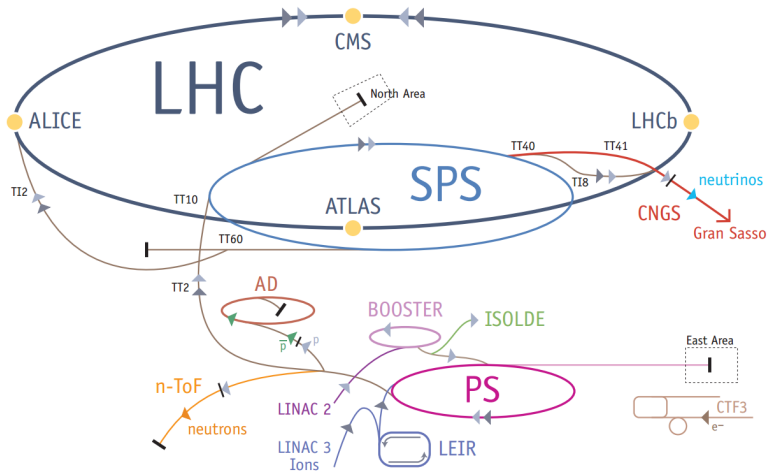

Figure 3.1: Sketch of the layout of the LHC. Protons start at LINAC 2 and Pb ions at LINAC 3 and then follow the arrows to the collision in CMS. Notice that the protons are also used in other experiments at the LHC which branch of the smaller accelerator rings.

LHC vital statistics	
circumference	26659 m
depth	50 - 175 m
total number of magnets	9600
number of main dipoles	1232
number of quadrupoles	392
temperature	1.9 K (-271.3C)
beam vacuum pressure	10^{-13} atm
nominal p energy	7 TeV
center-of-mass energy	14 TeV
design luminosity	10^{34} cm^{-2}s^{-1}
protons per bunch	1.1×10^{11}
bunch spacing	50 ns
bunches per beam	2808
turns per second	11245
collisions per second	600 million
length of each dipole	15 m
weight of each dipole	$\sim$ 35 tonnes
dipole field	8.33 T

Table 3.1: Design parameters for LHC running.

3.1.2 Luminosity at the LHC

In line with the LHC's extensive physics goal, vast amounts of collision data needed to be obtained. This means accumulating a large integrated luminosity: $\mathcal{L}^{\text{int}} = \int \mathcal{L} dt$. Achieving these goals required the LHC to maintain a high instantaneous luminosity, $\mathcal{L}$. The machine luminosity depends on the following beam parameters:

$$\mathcal{L} = \mathrm{F}\frac{\mathrm{N}_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi\epsilon_n \beta*}$$

where F is the geometric luminosity reduction factor due to the crossing angle at the interaction point (IP), N_b is the number of particles per bunch, n_b is the number of bunches per beam, f_{rev} the revolution frequency, γ_r the relativistic gamma factor, ϵ_n the normalised transverse beam emittance [39] and $\beta*$ the amplitude function at the collision point. ϵ_n and $\beta*$ are used to describe the shape of the beam at the interaction point and are related by:

$$\epsilon_n = \pi\frac{\sigma^2}{\beta*}$$

where σ is the cross sectional area of the bunch. Emittance is defined as the parallelism of the beam and so for a low emittance beam the particles are confined to a small area with approximately the same momentum. $\beta*$ is determined by the focusing strength of the quadrupole magnets. The LHC operates at low $\beta*$ where the beam is narrower or "squeezed". This allows for more interactions per bunch crossing.

The instantaneous luminosity of the LHC has been rapidly improving since running began in 2010. This was a result of increasing the number of bunches and optimising the optics to minimize transverse spread of the beam. In 2010 the LHC delivered a total of 44.22 pb^{-1} of which 40.76 pb^{-1} was recorded by the CMS detector. Since then the 2011 and 2012 runs have yielded total integrated luminosities of 5.55 fb^{-1} and 21.79 fb^{-1} respectively, thus making the 2010 data set irrelevant. Table 3.2 shows some luminosity information for the first 3 years of LHC running and Figure 3.2 presents the

accumulation of the total luminosity for these runs.

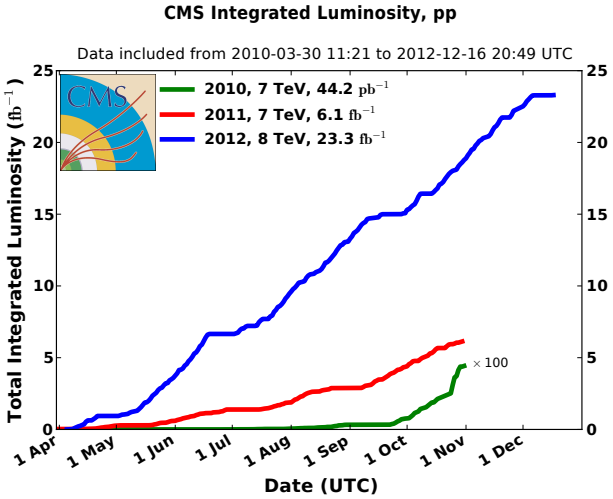

Figure 3.2: The accumulation of the integrated luminosity produced at the LHC vs time in 2010, 2011 and 2012. The 2010 integrated luminosity is multiplied by a factor of 100 so that it is visible on the plot. Image from [40].

3.2 The Compact Muon Solenoid Detector

The Compact Muon Solenoid (CMS) [41] is a multi-purpose detector located 100 meters underground at point 5 near the town of Cessy, France. Along with ATLAS (A Large Toroidal ApparatuS) it is designed to cover a broad range of physics goals including stringent tests of the Standard Model as well as searches for new physics. CMS has an overall length of 28.7 meters and a diameter of 15 meters. It weighs in at roughly 14000 tonnes.

3.2.1 Detector Components

CMS (shown in Figure 3.3) is characterised by a large superconducting solenoid which provides a 3.8 T magnetic field. The solenoid is about 14 meters long with a radius of 3 meters and is located between the Hadron Calorimeter (HCAL) and the muon detection system. At the design instantaneous luminosity of 10^{34} cm^{-2}s^{-1}, an average of 20 interactions are

2010	
Peak Inst. lumi.	205.14 $\times 10^{30}$ cm^{-2}s^{-1}
Total Integrated lumi. delivered	44.22 pb^{-1}
Total Integrated lumi. recorded	40.76 pb^{-1}
2011	
Peak Inst. lumi.	3535.08 $\times 10^{30}$ cm^{-2}s^{-1}
Total Integrated lumi. delivered	6.13 fb^{-1}
Total Integrated lumi. recorded	5.55 fb^{-1}
2012	
Peak Inst. lumi.	7670.19 $\times 10^{30}$ cm^{-2}s^{-1}
Total Integrated lumi. delivered	23.30 fb^{-1}
Total Integrated lumi. recorded	21.79 fb^{-1}

Table 3.2: CMS luminosity information

expected per bunch crossing. This can lead to the production of more than 1000 particles and so very high detector granularities are required to be able to decipher interesting events from the copious backgrounds. The innermost subdetector is comprised of the Tracker which is used to pin point the location of the primary vertex of an event and to track the trajectory of charged particles. The CMS tracking system is the largest tracking detector ever built for a high energy physics experiment with dimensions of 5.4 meters in length and 1.1 meter in radius. Outside of the Tracker is the Electromagnetic Calorimeter (ECAL), which is comprised of 61200 lead tungstate ($PbWO_4$) crystals within the barrel and 7324 in each of the two endcaps. It has an internal radius of 1.29 meters. The HCAL is located outside of the ECAL and is also made up of barrel and endcap sections built of brass layers. It has a radial width of 1.18 m. The Tracker, ECAL and HCAL are all located within the field volume of the solenoid. Outside the solenoid are the muon chambers which make up the outermost layer of the detector. Muons were pinpointed as a key ingredient for the discovery of the Higgs boson and so their detection was of high priority in the design of the detector.

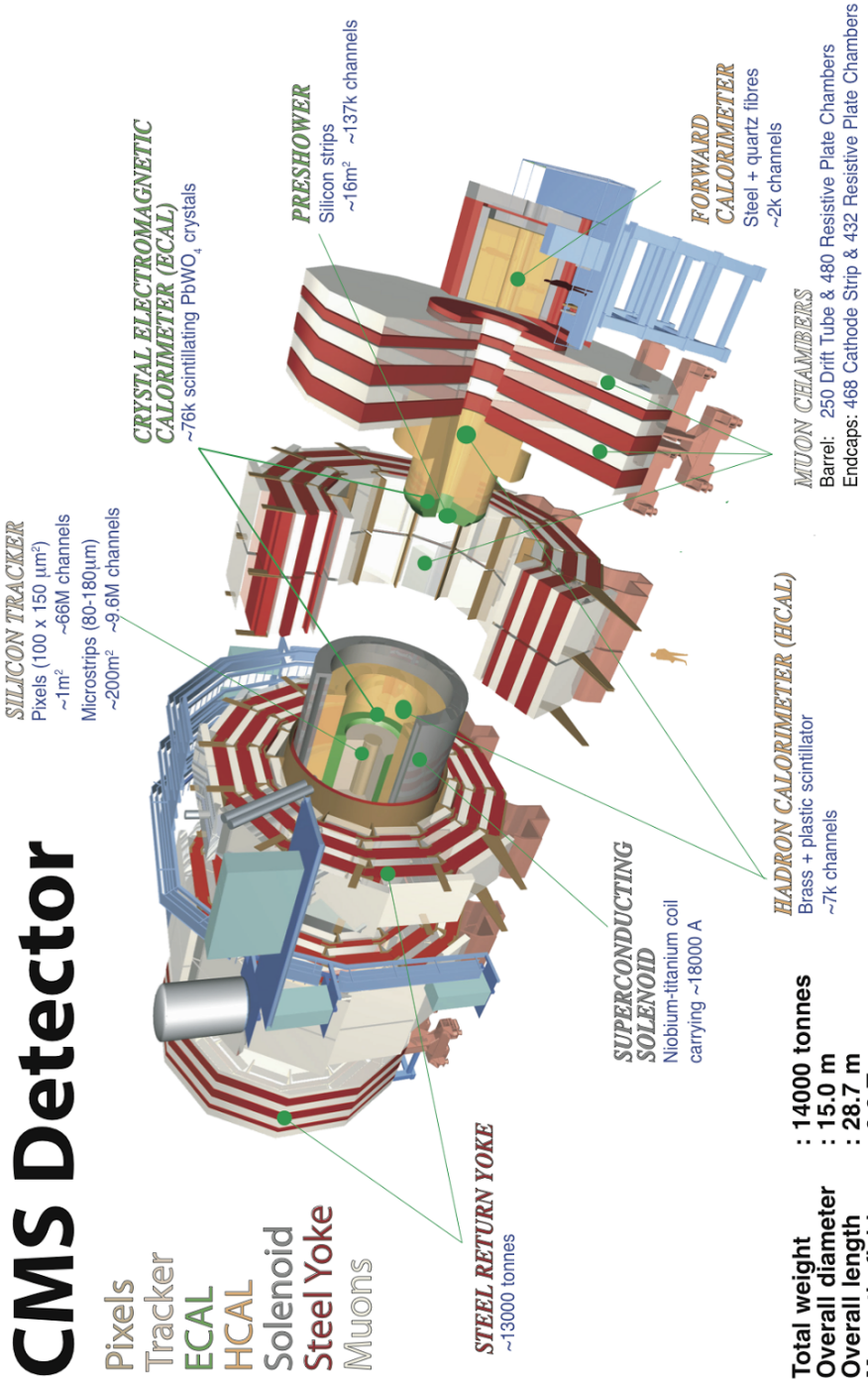

Figure 3.3: Schematic of the CMS detector with all the key components labelled.

3.2.2 Coordinate System

The standard CMS coordinate system is a right handed system with the z-axis along the beam direction. The x-axis points towards the centre of the LHC ring and the y-axis points vertically upwards. The momentum components of an individual particle in the x-y plane are p_x and p_y, and its transverse momentum is $p_{\mathrm{T}} = \sqrt{p_x^2 + p_y^2}$. The polar and azimuthal angles of a particle are $\theta = \tan^{-1}(p_{\mathrm{T}}/p_z)$ and $\phi = \tan^{-1}(p_y/p_x)$, respectively. In the polar coordinate system the radial position, r, is used. The pseudorapidity is defined as $\eta = -\ln\tan(\theta/2)$ and the transverse energy of an object with energy E is $E_{\mathrm{T}} = E\sin\theta$.

3.2.3 The Inner Tracking System

The CMS tracking system is shown in Figure 3.4 and covers a pseudorapidity (η) range up to 2.5. Its job is to measure the trajectories of charged particles such that the four-momentum of the particles and the position of the primary vertex of an event can be determined. Tracking the path that a charged particle makes as it traverses the Tracker enables the momentum of the particle to be calculated. The Tracker is comprised of two parts, the pixel detector and the strip Tracker. Both are made of silicon, which was chosen since it can be finely-segmented into sensors that provide a better spatial resolution and a faster readout than other technologies such as cloud and wire Trackers. Silicon is, however, expensive and requires sophisticated electronics and cooling systems. This means that the Tracker has quite a large material budget. In particular, in the region between the barrel and the endcap ($|\eta| \sim 1.5$) where the material budget reaches 1.8 radiation lengths (X_0) [42], where one radiation length corresponds to the distance over which an electron loses $1/e$ of its initial energy. As large fluxes of high energy particles are produced close to the interaction point the Tracker has to be built radiation hard.

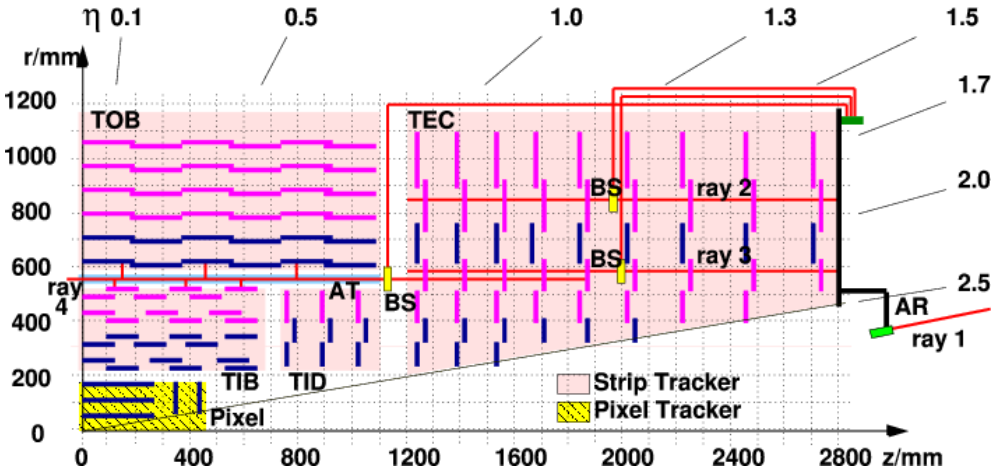

Figure 3.4: A cross section view of a quarter of the CMS inner tracking system in the r, z plane. The 0,0 position is the nominal interaction point. Taken from [43].

Pixel Detector

The pixel detector is the subdetector closest to the interaction point. It is primarily for vertexing and it also plays a crucial role in identifying decays from bottom quarks and tauons by observing displaced vertices. It is also the region of the detector where there will be the highest particle density. At full design luminosity, around 10 million particles are expected to pass through each cm^2 every second. It therefore needs to have a high granularity to be able to separate the tracks of different particles. To do this, silicon pixels of rectangular shape $100(r\phi) \times 150(z)$ μm^2 are used. A position resolution of approximately 15 μm can be obtained when charge is shared between neighbouring electrodes. The pixel detector is arranged into three layers at radii of 4.4, 7.3 and 10.7 cm, respectively. Two end cap disks that extend the pseudorapidity coverage are located at 34.5 cm and 46.5 cm, respectively, from the nominal interaction point.

Silicon Strip Tracker

After the inner pixel detector, the next part of the Tracker is the silicon strip Tracker. As there are lower particle fluxes whilst moving radially outwards

from the interaction point, strips can be used instead of pixels to save on cost. This part of the detector is composed of more than 15000 strips and covers an area of more than 200 m^2. The strip Tracker is comprised of four sub detectors as shown in Figure 3.4 and is divided into inner and outer segments.

In the inner part the strips are arranged into the Tracker Inner Barrel (TIB) and Tracker Inner Disk (TID). Strips in the TIB/TID come in a variety of geometries which range in width from 80 μm to 141 μm, have a length of 10 cm and are 320 μm thick. These two inner layers of the TIB achieve a position resolution in $r\phi$ of 16 to 27 μm in the barrel [44].

The outer part of the tracking system consists of the Tracker Outer Barrel (TOB) and the Tracker endcaps (TECs). Strips in these parts of the detector range in width from 97 μm to 184 μm, are 10 cm in length and have a thickness of 320 μm. The TOB is made of 6 layers and is able to resolve a position in $r\phi$ of 25 μm to 41 μm [44] in the barrel.

3.2.4 The Electromagnetic Calorimeter (ECAL)

The ECAL part of the detector as shown in Figure 3.5 is the part of the detector that is responsible for stopping particles that shower electromagnetically, these being electrons and photons. An electron initiates a shower via bremsstrahlung and a photon via pair production. Radiation losses due to bremsstrahlung scale with particle mass, m^{-4} (m^{-6}) when a charged particle travels perpendicular (parallel) to an electric field and hence heavier leptons are much less likely to shower. The ECAL is a hermetic, homogeneous detector which consists of 61200 lead tungstate ($PbWO_4$) crystals in the barrel and 7324 in each of the endcaps. Pre-shower detectors are also located at both ends. The design of this part of the detector is focused towards the discovery of the Higgs boson and hence the position and energy resolution of photons and electrons are required to a very high precision.

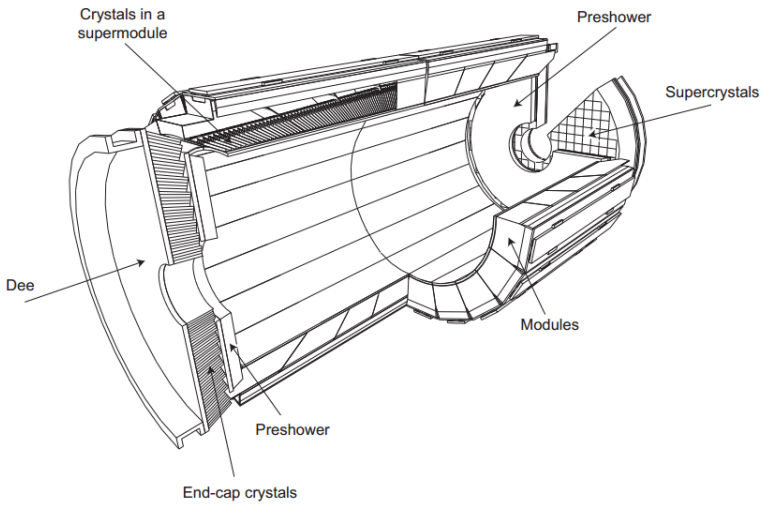

Figure 3.5: Illustration of the CMS ECAL. Taken from [45].

Lead Tungstate ($PbWO_4$) Crystals

The choice of $PbWO_4$ as the material to be used for this part of the detector is motivated by its high density (8.28 gcm^{-3}) which means it has a very short radiation length (0.89 cm) and a low Moliere radius (2.2 cm). This means that EM showers are contained within a small volume and, as such, a better position resolution can be obtained. These properties also mean that around 80% of the shower is initiated within 25 ns. It is important that a fast response is achieved when dealing with high pile up conditions.

ECAL Barrel (EB)

The barrel section of the ECAL covers a pseudo rapidity range up to 1.479. In this section the $PbWO_4$ crystals are tapered such that the smaller end ($2.2{\times}2.2$ cm^2) faces the beam line at a radial distance of 1.29 m. Crystals are divided into 36 super modules which each give 10° coverage in ϕ. Crystals in the barrel are 23.0 cm in length which equates to 25.8 X_0. The barrel crystals are read out by Avalanche Photodiodes (APDs) which have an area of $5{\times}5$ mm^2. The energy resolution is sensitive to both the voltage ($\frac{1}{E}\frac{dE}{dV} \sim 3.1\%V^{-1}$) and the temperature ($\frac{1}{E}\frac{dE}{dT} \sim -2.4\%°C^{-1}$) and so these both need

to be monitored closely such that the resolution is kept to within 0.1%.

ECAL Endcap (EE)

The ECAL endcaps (EE) are located at either end of the barrel at a distance of 3.15 m from the nominal interaction point. The endcap covers a range of $1.479 < |\eta| < 3.0$. Each EE is divided into 2 dees with each dee providing 180° ϕ coverage. The dees are arranged into an x-y grid of 5×5 $PbWO_4$ supercrystals. The EE crystals have a area of 2.86×2.86 cm^2 at the end facing the interaction point and a length of 22.0 cm (24.7 X_0). The light yield in the EE crystals is read out by Vacuum Phototriodes (VPTs) these having an area of 280 mm^2. The read out is done via a copper mesh anode at a pitch of 10 μm such that it is operational within the high magnetic field.

ECAL Pre-shower

A pre-shower (ES) detector is also in place just in front of each endcap covering a range of $1.653 < |\eta| < 2.5$. The ES is comprised of 137,000 Pb-Si strips and its high granularity helps discriminate between isolated photons and photons from neutral meson decays.

Energy Resolution of ECAL

The energy resolution, $\frac{\sigma_E}{E}$, of the ECAL crystals can be parameterised by the following equation:

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{A}{\sqrt{E}}\right)^2 + \left(\frac{B}{E}\right)^2 + C^2 \tag{3.1}$$

where A is the stochastic term due to the probabilistic nature of scintillation showers. B is the noise term due to the PMTs and the readout electronics. C is the constant term which is a direct measure of the performance of the crystals. The performance of the ECAL can be affected by non-uniformity in longitudinal light collection, intercalibration errors and energy leakage at the back of the crystal. The resolution has been measured in test beams using electrons of energies between 20 and 250 GeV. The coefficients were

determined to be $A = 2.8\%$, $B = 0.12\%$ and $C = 0.3\%$ [46]. This means that the constant term becomes dominant at electron energies of around 100 GeV. In-situ measurements of the energy resolution have also been performed using $Z \to e^+e^-$ decays with the 7 TeV data set [47]. An energy resolution for electrons from Z decays is determined to be 2% in the barrel for $|\eta| < 0.8$ and ranging from 2-5% elsewhere. The transmission of the $PbWO_4$ crystals is reduced by high radiation doses as is described in detail in Chapter 4.

3.2.5 The Hadronic Calorimeter (HCAL)

The Hadronic Calorimeter, shown in Figure 3.6, is used to measure the energy and position of hadrons in an event, which are usually produced in jets. The HCAL is also important in the measuring of the missing transverse energy in an event as the measurement of jet energy resolution is key. This part of the detector must therefore be hermetic in ϕ.

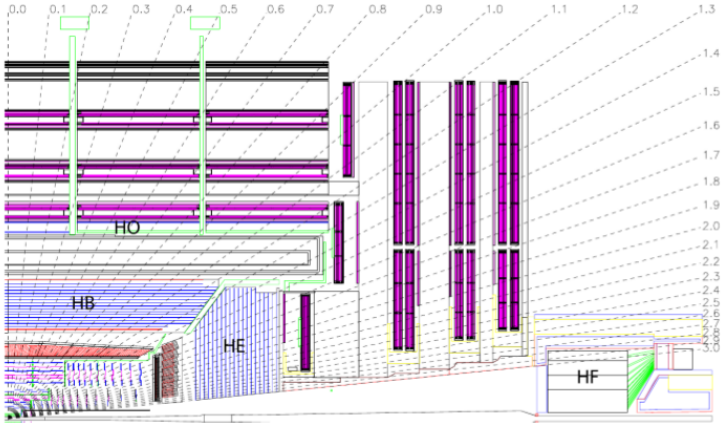

Figure 3.6: Illustration of the CMS HCAL. Taken from [48].

Components of the HCAL

The HCAL is divided in to four main segments. The Barrel Calorimeter (HB) is sandwiched between the ECAL and the solenoid and so is limited to a radial width of 1.18 m providing coverage up to $|\eta| < 1.4$. An additional

part of the HCAL extends outside (HO) the region of the solenoid using the magnet for additional absorption. The HO catches the tails of hadronic showers and extends the HCAL's width to 11.8 interaction lengths (λ_I). λ_I corresponds to the mean free path of a charged hadron within a medium. The Endcap Calorimeter (HE) is within the field volume of the solenoid and extends the $|\eta|$ coverage up to 3.0. Finally, the Forward Calorimeter (HF) is positioned at a distance of 11.2 m from the nominal interaction point and gives additional coverage in the region $3.0 < |\eta| < 5.2$.

The HCAL is a sampling calorimeter comprised of wedges made up of 14 azimuthal Brass (70% Cu and 30% Zn) plates with a steel plate either side for structural stability. The plates absorb particles from hadronic showers which then produce light in scintillating tiles which are placed between the absorbers. This light is then read out via optical fibres which are linked to Hybrid Photodiodes (HPDs).

Energy Resolution of HCAL

The resolution of the HCAL is limited by the nature of the interaction of the particles within it. The energy from neutral mesons decaying to two photons can be measured directly in the scintillators whilst charged particles produce scintillating photons via ionisation, excitation and also nuclear interactions. The resolution of the HB, HE and HO is described by:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{A}{\sqrt{E}}\right)^2 + C^2$$

where the stochastic term, A, is 90% and the constant term is 4.5%. For the HF, $A = 172\%$ and $C = 4.5\%$ [49]. The results were obtained from HCAL test beam measurements. The resolution can be substantially improved for charged hadrons by additionally using tracking information and reconstructing using the particle flow algorithm (see Section 5.2).

3.2.6 The Muon Detection System

Since "muon" is in the name of the detector the identification and the momentum resolution of this particle were clearly of high importance in the detectors design. Muons decay with a lifetime of 2.2 μs and so they typically travel well beyond the outer edge of the detector before decaying. Higher p_{T} muons travel in an almost straight line through the Tracker and hence a larger detector is required to obtain the desired momentum resolution. In total, the muon detector chambers are made up of a gaseous volume of approximately 25,000 m^3. The detection chambers are located outside the solenoid within the steel yoke. The yoke is in place to homogenise the magnetic field and returns the magnetic flux to the solenoid. It also absorbs hadrons that have managed to get this far. A large volume is required for this part of the detector. Gas chambers were used as thay are a cheap, robust and reliable option. Muons are detected as they ionise the gas as they pass through. A more detailed overview of the muon detection system can be found in the technical design report [50].

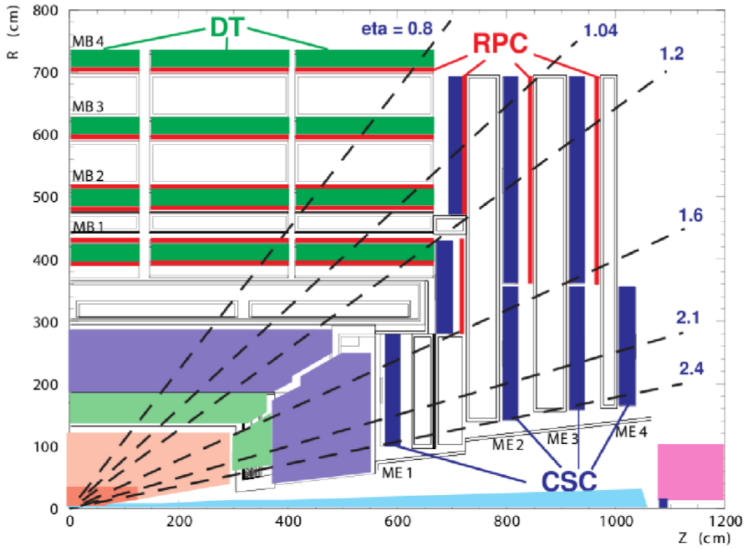

Figure 3.7: Illustration of the CMS muon detection system. Taken from [50].

Barrel Muon System

In the barrel, where there is a low muon flux and a uniform magnetic field is present, Drift Tubes (DT) are used. Four stations are embedded within the yoke as can be seen in Figure 3.7. The three innermost stations consist of four pairs of chambers which measure muon position in the $r\phi$ and z planes. The outer chamber is used primarily for the z position measurement. The gas in the chambers is a mixture of 85% Ar and 15% CO_2. A 50 μm anode wire runs down the centre of the chamber and the walls are grounded such that they act as a cathode. Each drift cell is offset by half a cell width with respect to its neighbours to help eliminate dead spots in efficiency. The barrel region of the muon detector covers the region up to $|\eta| < 1.2$.

Endcap Muon System

In the endcap region the muon fluxes are higher and there are also increased, non-uniform magnetic fields present. Cathode Strip (CS) Detectors are used in this part of the muon system. These have a faster response time and are more radiation hard than the DTs. The CS are comprised of 468 trapezoidal gas chambers each covering 10° or 20° in ϕ and provide coverage in the range $0.9 < |\eta| < 2.4$. The chambers contain 6 planes of anode wires within 7 cathode panels providing excellent spatial resolution. A gas mixture of 40% Ar and 50% CO_2 is used with the addition of 10% CF_4 to prevent polymerisation of the wires.

Resistive Plate Chambers

In the $|\eta| < 1.6$ region Resistive Plate Chambers (RPCs) are installed along side the DTs and CSs to provide an independent, efficient triggering system. They give a coarser position measurement than the other chamber types but give a fast read out and a good time resolution ($\sim$ 1 ns) which is important for triggering.

Resolution of the Muon System

The momentum resolution was initially determined prior to LHC start up using cosmic ray data [51]. In these tests the cosmic ray muons traverse the full detector and so comparisons of the momentum measurement in the upper and lower segments can be used to determine the resolution. When using both the muon detection system and the Tracker, the muon momentum resolution was found to be around 1% or 2% for lower p_{T} muons (10 GeV/c) and no more than about 10% for high p_{T} muons (1 TeV/c). This observation was supported using $Z \rightarrow \mu\mu$ events in early collision data [52].

3.2.7 Trigger and Data Acquisition (DAQ) System

When running at full luminosity an event rate of 40 MHz is expected. With each event corresponding to roughly 0.5-1 MB of data it is impossible to store all of them. The purpose of the trigger is to filter the more interesting events, such that only these events are saved for further analysis. Figure 3.8 shows the cross sections of various physics processes and the event rates associated with them at a luminosity of 10^{34} $\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The trigger operates in two layers. The first is the level 1 (L1) trigger which is based on hardware built into the CMS detector and in the underground control room. The L1 trigger's purpose is to reduce the event rate from 40 MHz to a maximum output of 100 kHz. If an event passes the L1 it is then passed on to the high level trigger (HLT) which is software based and executes algorithms similar to those used offline on a computer farm located at Point 5. An array of triggers are used to select events appropriate for a variety of analyses. This is known as the trigger menu. Overlap between similar triggers is minimised to no more than 10% of events that pass. Events that pass a given trigger go on to make up the data set needed for a specific physics analysis.

L1 Trigger

The L1 trigger uses coarse information from the muon and calorimetry systems to assess whether an event should be passed onto the HLT. The architec-

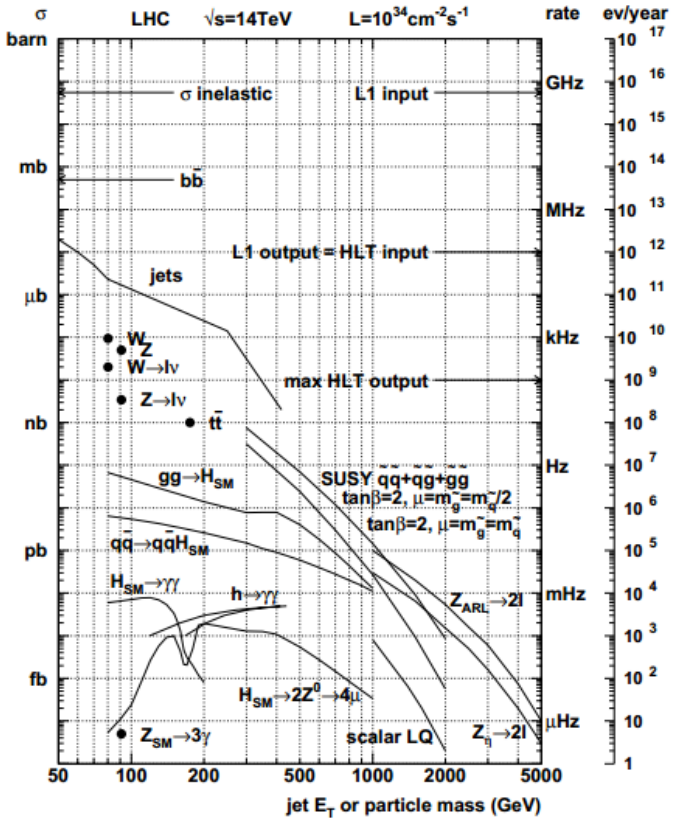

Figure 3.8: Production rates for various physics processes. Event rates for a luminosity of 10^{34} cm^{-2}s^{-1} is on the right hand scale. Figure from reference [53].

ture of the L1 trigger is depicted in Figure 3.9. The L1 trigger analyses every bunch crossing and has 3.2 μs to pass events on to the front-end electronics. The information from the detector must therefore be pipelined within the electronics. The first step at L1 are the Trigger Primitive Generators (TPG) which look for hit patterns in the muon system or energy depositions in the calorimeters. Regional triggers are then employed which match hits/clusters within local regions of the sub detectors to form tracks. These are designated primitive muon, electron, photon or jet candidates. This regional information is then ranked in terms of p_{T} or E_{T} and the quality of the deposition within the detector. The Global Calorimeter Trigger (GCT) and Global Muon Trigger (GMT) determine the highest ranked of these across the whole detector. Finally this information is passed to the Global Trigger (GT) which makes the decision on whether or not to pass the event onto the HLT.

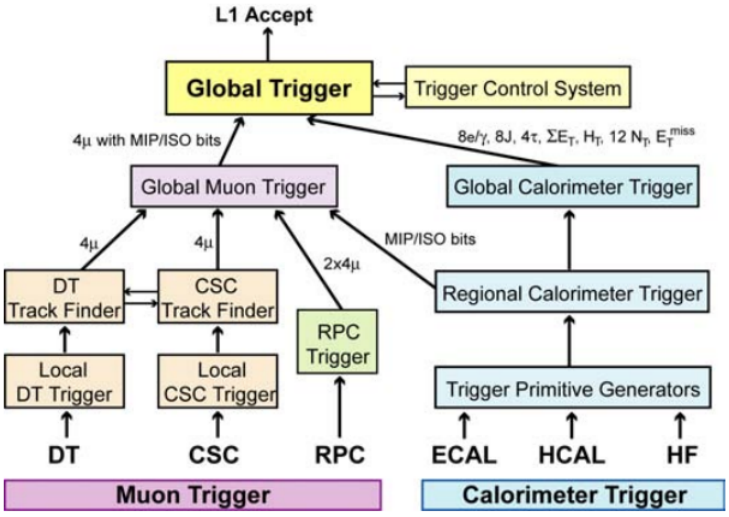

Figure 3.9: Image showing the structural flow of the data as it goes through the L1 trigger. Image from reference [41].

The High Level Trigger (HLT)

The HLT is composed of programmable software that has access to all of the information from the CMS detector. In the early stages of LHC running

the HLT selection criteria and trigger menus changed regularly to deal with the rapidly increasing instantaneous luminosity. The HLT has the following functions:

1. Collection of all the information from the various CMS sub-detectors;
2. Online reconstruction of interesting physics parameters such as p_{T}, E_{T}, η and the isolation of leptons and jets. Other event variables such as H_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$ are also reconstructed;
3. Selection (triggering) based on these reconstructed quantities.

The HLT takes roughly 50 ms to carry out these functions and reduces the recorded event rate by a factor of 10^3, resulting in an output rate of 300 Hz. The data accepted by the HLT is then transferred to the Tier 0 computer system located at the main CERN site where full event reconstruction takes place (See Chapter 5 for more details).

3.3 Summary

This chapter was dedicated to the LHC and the CMS detector. The design specifications for these experimental apparatus have been described. The LHC machine performed exceptionally during the first three years of running and the instantaneous luminosity increased by several orders of magnitude. In 2011 and 2012, total integrated luminosities of 6.1 fb^{-1} and 23.3 fb^{-1} were produced, respectively. This has meant that CMS has had to be very adaptable to these changing conditions in terms of the trigger and DAQ system. The detector has performed very well in terms of energy resolution and identification of physics objects.

Chapter 4

Hadron Induced Radiation Damage in $PbWO_4$ ECAL Endcap (EE) Crystals

The CMS EE (see Section 3.2.4), is designed to withstand radiation doses of up to 100 kGy* or charged hadron fluences of up to 6×10^{13} cm^{-2} [46]. In phase II of LHC running, a total integrated luminosity of around 300 fb^{-1} is expected to be accumulated. This is equivalent to the doses that have been used to irradiate some of the EE crystals studied in this chapter. This level of dose can lead to crystal light output losses of up to 50-60% for EE crystals at $|\eta|$ >1.7 [54]. The residual damage to $PbWO_4$ crystals is defined as the reduction in light output due to losses in light transmission. Transmission ($T = I/I_0$) is defined as the loss in the intensity of light traversing a crystal, where I_0 is the intensity prior to traversing the crystal and I is the intensity after traversing. There is almost no recovery from hadron induced damage at the ECAL operating temperature [55, 56, 57, 58]. It is therefore important to understand the effect of the damage and model its effects.

As was previously stated in Section 3.2.4, the performance of the ECAL can be parameterised using Equation 3.1. In this chapter the effect of radiation damage on the energy resolution and linearity will be examined using

*Gy (gray) is the SI unit for a radiation dose and is defined as the amount of energy (in Joules) absorbed per kg of matter.

the results from electron test beam experiments of energy in the range 10 to 150 GeV. The results are used to validate a phenomenological damage model [59] that has been developed to describe the observed effects. *Energy linearity* for a crystal matrix is defined as:

$$\text{Linearity } (E_{\text{beam}}) = \frac{E_{\text{reco}}}{E_{\text{beam}}} \tag{4.1}$$

where E^{reco} is the reconstructed energy and E_{beam} is the test beam energy (known to a high precision, $\delta p/p \sim 0.3\%$).

The energy resolution and linearity of crystals are both expected to be affected to some degree by hadron irradiation. In the EE, a high precision light monitoring system is designed to correct for transmission losses in order to maintain the contribution of the constant term to the energy resolution to less than 0.3% [46]. The effect of the transmission loss within the crystal will, however, degrade several aspects of the crystal response such as the pulse shape and the uniformity of the light collection efficiency which affects the energy linearity and resolution.

The work in this chapter was performed in collaboration with people based at CERN. The author was involved in the data taking (electron test beam shifts) as well as the reconstruction and analysis of the test beam data during the 2011 and 2012 test beam data taking periods.

4.1 Crystal Irradiation and Damage

The following subsections outline the proton irradation procedure and the method for measuring the damage caused to each crystal. $PbWO_4$ crystals with dimensions identical to the ones used in the EE ($30\times30\times220$ mm^3) have been proton-irradiated at the PS irradiation facility at CERN [60]. A spectrophotometer was then used to measure the damage and check if this damage is uniformly distributed accross the crystals.

4.1.1 Proton Irradiation Procedure

The protons that were used to irradiate the crystals were of energy 24 GeV. A proton beam of area 30×30 mm^2 was used with a proton average flux of 10^9 $s^{-1}cm^{-2}$. The transverse proton beam profile was tuned before each irradiation to be as uniform as possible over the entire (x-y) area of the crystal. At the start of the 2012 irradiation campaign a new beam detector with higher sensitivity and finer granularity was installed, which allowed better beam profile monitoring. This monitoring showed a slightly reduced flux of protons towards the edge of the beam with rates of about 75% of the central flux 20 mm from the centre in the x-y plane. A full list of the crystals that have been irradiated is shown in Appendix A.

4.1.2 Crystal Damage and Non-Uniformity

The damage to each crystal is defined by the induced absorption coefficient (μ_{ind}). This quantity is related to the transmission loss at 420 nm. This wavelength corresponds to the maximum transmission of the $PbWO_4$ emission spectrum. μ_{ind} is calculated using:

$$\mu_{\text{ind}} = \frac{1}{L} \ln \frac{T_{\text{before}}^{420\text{nm}}}{T_{\text{after}}^{420\text{nm}}} \tag{4.2}$$

where L is the crystal length equal to 0.22 m, T_{before} and T_{after} correspond to the crystal light transmission at 420 nm before and after irradiation. Damage profiles were measured at wavelengths in the 300-700 nm range. Results obtained for crystal 11135 ($\mu_{\text{ind}} = 7.4$ m^{-1}) are shown in Figure 4.1. Damage causes the transmission band edge to be shifted by several tens of nanometers to higher values causing it to start to overlap with the emission peak.

The transverse non-uniformity of the damage as a result of the non-uniformity of the proton beam was visible particularly for the crystals irradiated to high integrated fluences, Φ, of $\sim 10^{14}$ p/cm^2. The transverse damage non-uniformity was measured using a dedicated spectrophotometer (optical beam spot of 2 mm diameter). Comparisons of the light transmission at 420 nm for different positions along the x and y axes indicates a strong

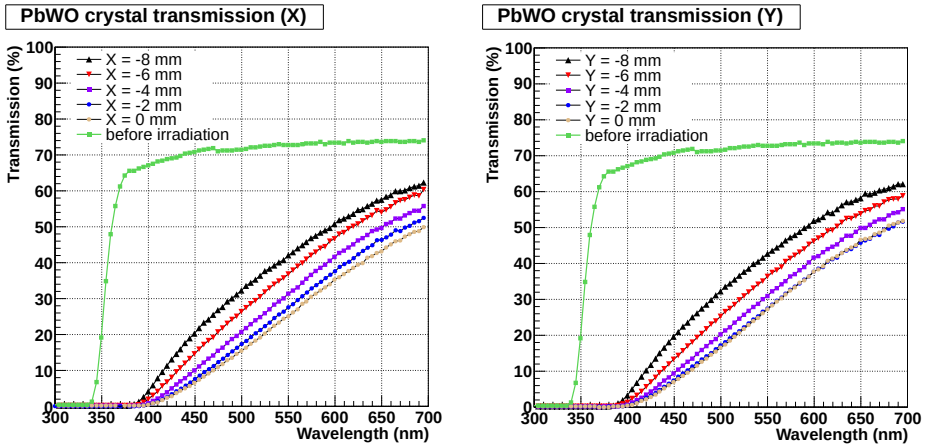

Figure 4.1: Transmission curves for crystal 11135 ($\mu_{\text{ind}} = 7.4$) at different positions on the x (left) and y (right) axis of the crystal. The $x = y = 0$ position corresponds to the centre of the crystal. The transmission curve for a standard non-irradiated crystal is shown in green.

non-uniformity within the crystal as shown in Figure 4.2.

The μ_{ind} for crystals irradiated in 2012 were averaged over each crystal's surface using the profiles shown in Figure 4.2. A polynomial fit to model the non-uniformity of the crystals. This fit gives a very good agreement to the data as seen with the low χ^2 value. The error bars on the μ_{ind} are propagated from the uncertainty on the spectrophotometry measurement. It is necessary to use the average μ_{ind} over each crystal for two main reasons:

- The value of accumulated fluence is integrated over the crystal surface;
- The development of an electromagnetic shower inside $PbWO_4$ crystals interacts with the whole crystal volume.

Only negligible non-uniformity was observed in crystals irradiated prior to 2012 and hence the average μ_{ind} does not need any correction. A 15% relative error on the estimation of μ_{ind} is assumed due to a 2% systematic error in the transmission measurement which propagates to $\sigma_{\mu_{\text{ind}}}$ via Equation 4.2. A 20% uncertainty is assigned for crystals irradiated in 2012 to represent the spread of μ_{ind} along the x axis, in the range (-7.5,7.5) mm.

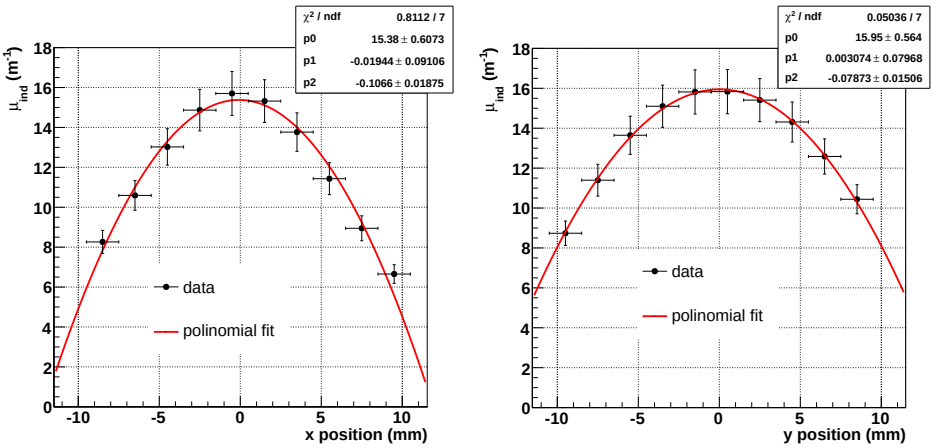

Figure 4.2: Transverse scan of the μ_{ind} for crystal 11135 calculated from transmission curves at 420 nm using a spectrophotometer with beam spot 4 mm^2. Similar profiles are observed for both the x (left) and y axis (right). Data points are fitted using a 2$^{\text{nd}}$ order polynomial function.

4.2 Test Beam Setup

4.2.1 The Test Matrix

The crystal configuration (Figure 4.3) used in test beam runs aims to reproduce the setup in the actual EE as much as possible [46]. A 5×5 EE supercrystal was placed in the standard supercrystal alveolar. The plastic inserts in front of the crystals were the same as the ones used in the EE, while the Al inserts at the photo detector side were modified to house Hamamatsu Photonics PMTs, which were used for the light readout instead of the EE-standard Vacuum Photo Triodes (VPT). Crystals were coupled to the PMT with optical grease of refractive index of 1.45 to enhance the light yields. The alveolar was surrounded by 1 cm thick Cu plates with pipes for water circulation used for the thermal stabilisation of the matrix. Two thermal sensors were installed on each side of the matrix. The precision of the temperature readout was better than 0.1 °C. The temperature variation was found to be lower than 0.2 °C and so its effect on the system stability was negligible.

An LED-based light monitoring system consisting of a blue LED light

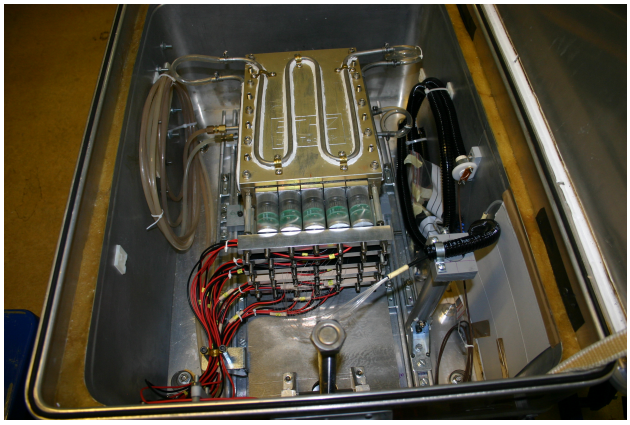

Figure 4.3: Image of the test beam matrix setup. The crystals are housed within a carbon fibre alveolar inside cooling plates (yellow). PMTs with HV dividers (green) and cables are behind the crystals readout by the red wiring. The LED light source and fibres for light monitoring are attached to the right wall of the box.

source (with a 455 nm emission peak) and a bundle of quartz optical fibres were mounted next to the crystal matrix. This was used to deliver the light to the back end of each crystal . The entire matrix was mounted inside an Aluminum box sealed to provide optical and thermal isolation. The box was installed on a remotely controlled x-y table with a displacement range of $\pm$ 20 cm and positioning precision of approximately 1 mm.

Crystal Configurations

The test matrices were configured to provide similar values of μ_{ind} in the central 3×3 matrix, surrounded by non-irradiated or slightly irradiated crystals. The matrices tested in 2011-2012 are shown in Figure 4.4. The values of μ_{ind} quoted in the Figures represent the average μ_{ind} measured prior to entering the test beam.

Average induced absorption for a 3×3 matrix

Since the matrices of crystals used for this analysis are characterized by different μ_{ind}, we define an average $\mu_{\text{ind}}^{3\times3}$ of the matrix by re-weighting the

$\mu_{\rm ind}$ of each crystal with its relative contribution to the reconstruction of an electromagnetic shower centred on the 3×3 matrix.

$$\mu_{\rm ind}^{3\times3} = \sum_{i=1}^{9} \mu_{\rm ind}^{i} \times w_i \ , \ \sum_{i=1}^{9} w_i = 1 \tag{4.3}$$

where for the central crystal w_5 is equal to 0.84, for the crystals adjacent to the central one $w_2 = w_4 = w_6 = w_8 = 0.032$ and for the crystals in the corners of the matrix $w_1 = w_3 = w_7 = w_9 = 0.008$. These weights have been determined by considering the relative average energy deposited in each crystal of the 3×3 matrix at 50 GeV using a 10×10 mm^2 beam spot.

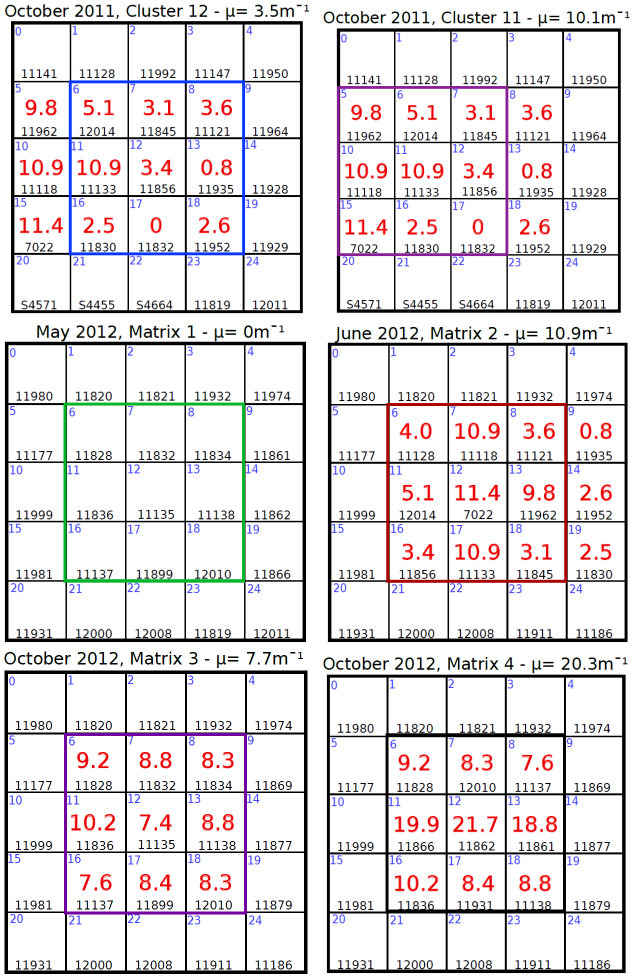

Figure 4.4: The crystal matrices from 2011 and 2012 test beam runs are shown in the above figure. The μ_{ind} of the irradiated crystals in m^{-1} are shown in red in the centre of the crystal, the crystal identification number (black, bottom) and the position number of each crystal (blue, top left) is also shown. The 3×3 clusters used for the following analysis are surrounded by a coloured thick line. The average value of the μ_{ind} for the selected cluster is shown above each matrix as was calculated using Equation 4.3 in Section 4.2.1.

4.2.2 H4 Experimental Setup

Beam counters

A sketch of the CMS H4 test beam zone is shown in Figure 4.5. The incoming beam was detected by a set of four scintillation counters: S1 (60×60 mm^2), S3 (40×40 mm^2), S4 (10×10 mm^2), and S6 (60×60 mm^2), installed at 5525 mm, 5340 mm, 5210 mm, and 2820 mm from the crystal matrix respectively. Two settings were used: "Wide beam" - coincidence of S1 and S6, and "Narrow beam" - coincidence of all four scintillators, S1, S3, S4 and S6. A narrow beam was usually used to maintain centralised electron impact. Wide beam was only used when the electron statistics were low. The impact position of the beam particles on the matrix was measured using two sets of beam hodoscopes. Each set was composed of two planes of 64 scintillating fibres of square cross section 0.5×0.5 mm^2, readout by a multi-anode PMT. The planes FH1 and FH3 measured the horizontal coordinate, and FH2 and FH4 the vertical one. These planes were installed at 6015 mm, 3490 mm, 5650 mm and 3125 mm from the matrix respectively.

Data Acquisition System (DAQ)

An event trigger could be one of three signals:

- "Electron" - coincidence of beam counting scintillators (four for "narrow" or two for "wide" beam).
- "LED" - generated by the LED pulsing at about 30 Hz.
- "Pedestal" - generated by the random gate generation at a rate of about 10 Hz.

An additional "spill" trigger, indicating the start and stop of particle spills from SPS was utilised. The SPS spill structure varied slightly between runs depending on the LHC operation. The average spill cycle was 38-45 seconds and consisted of a slow extraction to the fixed target ($\sim$10 sec), several fast extractions for the CNGS neutrino beam and also LHC injection. The data was taken during the slow extraction, recorded to disk and monitored during

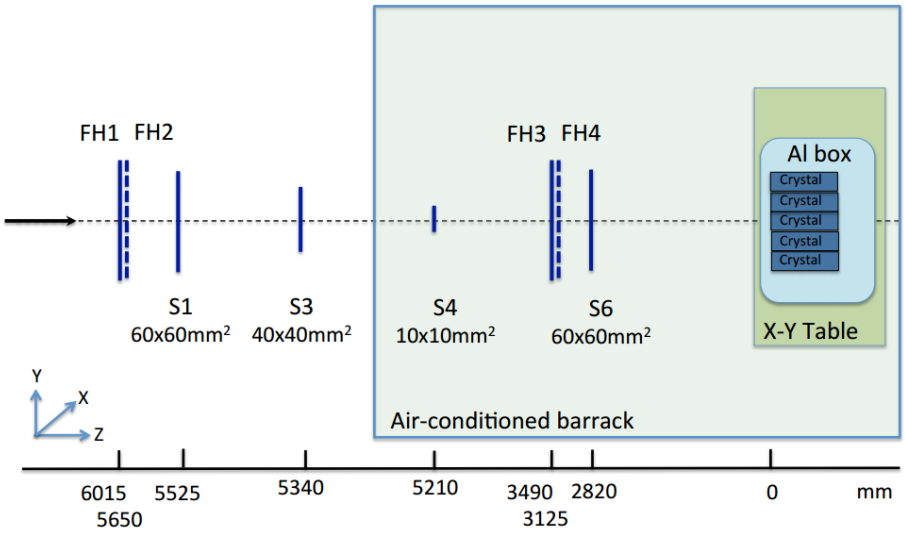

Figure 4.5: A birds eye view of the CERN H4 test beam zone showing the position of the fibre hodoscopes (FH1, FH2, FH3, FH4) and the scintillation counters (S1, S6, S3, S4) with respect to the box which contains the crystals. The black arrow indicates the beam direction.

the remaining 25-30 seconds of the SPS cycle. A standard data taking rate of approximately 10000 triggers per spill gave roughly 3000 electron events.

4.2.3 Data Taking Procedure and System Monitoring

There were two data taking modes in operation: Short calibration runs where a 50 GeV beam was positioned on each crystal and longer scans at a specified energy usually focusing on the central crystal in the matrix.

Calibration Runs

Calibration of the crystal matrix readout chain was performed by positioning the crystals, one-by-one, into the path of a 50 GeV electron beam by moving the table supporting the matrix. The particle rate was around 1 kHz. All three types of event: “electron”, “LED” and “pedestal” were recorded. The calibration run finished once 20000 “electron” events per crystal had been collected. The reproducibility of the table positioning was measured

to be better than 1 mm. The table movement was controlled by the DAQ computer and the calibration procedure was fully automated to run over all crystals. During standard beam conditions the full calibration run took about 60 minutes.

Energy Scans

Energy scans were performed with the beam incident on a specific crystal in the matrix. The beam energy was defined by magnet elements within the H4 beam line. The configuration of the beam line permitted an adequate quality of electron beam in the 10 to 250 GeV range. The quality of the beam was defined in terms of low amounts of hadron and muon contamination, a sufficient rate and a low angular divergence. Energy scans at 10, 20, 50, 100, 120, 150 and 250 GeV were performed (although only 10-150 GeV results are shown in this thesis). Typically a few scans were done for each energy requiring 50k events per scan.

System Monitoring

The gain of the PMTs change with anode current variation. LED light was injected in to all 25 crystals of the supercrystal to monitor the gain stability of individual PMTs. The purpose of this was to verify that the gain of the PMTs was constant during and between spills. LED stability ensures the stability of the crystal, PMT and electronic readout chain. On larger time scales (30 hours and upwards), a difference of a few percent in the recorded LED signal was observed. As a precaution, approximately every 20-30 hours, a repeat of the calibration run using the 50 GeV electron beam was performed.

4.3 Analysis Methods

The procedure used to calibrate a 3×3 crystal matrix involved beam spot selection, calibrating each individual crystal and finally obtaining a calibration coefficient for the entire 3×3 matrix. Calibration coefficients were used

to reconstruct the electron energy for all energy scans where the LED value remains stable. The energy reconstruction was then used to measure the energy linearity and resolution of the irradiated crystals. The noise term was fixed for each matrix by extracting it from the pedestal as is described later in this section.

4.3.1 Beam Spot Selection

The beam hodoscopes ($FH_{1,2}$ and $FH_{3,4}$) were positioned before the crystal matrix in the beam line ($z_{FH_{1,2}}$, $z_{FH_{3,4}}$), as seen in Figure 4.5. The hit position of the electron on the matrix (x_m, y_m) was reconstructed by extrapolating the particle trajectory using:

$$x_m = x_{FH_1} + (x_{FH_3} - x_{FH_1}) \times \frac{z_{FH_1}}{z_{FH_1 - FH_3}}. \tag{4.4}$$

The information about the electron position permits the study of crystal response as a function of electron impact point. These hit profiles enable the crystal center with respect to the matrix coordinate system (x_m, y_m) to be found. The coordinates (x_c^i, y_c^i) of the centre are obtained for each crystal, i, by fitting these distributions with a 2^{nd} order polynomial function in the range of ± 10 mm:

$$\text{Normalised Response} = p_0 + p_1 x_m + p_2 x_m^2 \tag{4.5}$$

This procedure allows the position of the centre of each crystal in the matrix to be found with a precision of approximately 1 mm.

A narrow beam spot selection of $4{\times}4$ mm^2 around the crystal centre was chosen for calibrating individual crystals. This selection was chosen in order to reduce fluctuations in amplitude. This improved the resolution of the electron peak and consequently the precision of the intercalibration coefficients from 11% to 9%. Beam spot selection reduces the statistics in proportion to the selected beam spot area. The beam hodoscope efficiency (70-90% per plane) used to record both x and y coordinates is an additional

limiting factor. A beam spot selection of 4×4 mm^2 reduced event statistics by up to 90%. The advantages of a narrower beam spot, however, outweigh these losses.

The energy spectra reconstructed using a 3×3 cluster does not improve when reducing the beam spot size from 10×10 mm^2 to 4×4 mm^2 because the shower is not confined to just one crystal. More statistics were also required in outer crystals. A 10×10 mm^2 beam spot area is therefore used when calibrating and reconstructing the energy within a 3×3 cluster of crystals.

4.3.2 Calibration of Individual Crystals

Individual crystals are calibrated by converting the recorded ADC (Analog-to-Digital-Converter) count to the beam energy (GeV) using the following procedure:

1. The ADC value corresponding to the electron peak, $\mathrm{ADC_i^{ele}}$, is obtained by fitting the amplitude distribution after the 4×4 mm^2 position section (see Figure 4.6 for example). A Crystal Ball function was used to extract this value:

$$f(x;\alpha,n,\bar{x},\sigma) = N.\begin{array}{ll} exp(-\frac{(x-\bar{x})^2}{2\sigma^2}), & \mathrm{for}\frac{(\mathrm{x}-\bar{\mathrm{x}})^2}{\sigma} > -\alpha \\ A.(B-\frac{x-\bar{x}}{\sigma})^{-n}, & \mathrm{for}\frac{(\mathrm{x}-\bar{\mathrm{x}})^2}{\sigma} \leq -\alpha \end{array}$$

   where N is a normalisation factor, and α, n and σ are parameters which are fitted to the distribution (x). The mean value of the peak of the electron distribution is given by $\bar{x}$. A Crystal Ball function was chosen as the amplitude distributions have asymmetric tails. A low energy tail is observed for non-irradiated crystals while a high energy tail appears for highly irradiated ones. A Gaussian function is unable to account for these tails.

2. The subtraction of the offset between the PMT output and the ADC input is known as "pedestal subtraction". The "pedestal" peak is fitted with a Gaussian function to extract the peak ADC position, ADC_i^{ped}.

The signal amplitude, $A_i = \mathrm{ADC_i^{ele}} - \mathrm{ADC_i^{ped}}$, can then be obtained for each crystal, i.

3. The ADC-to-energy conversion coefficient, G_i, was then obtained for each crystal using the following relationship: $E_{\mathrm{beam}} = G_i \times A_i$. The calibration beam energy, E_{beam}, was 50 GeV.

The uncertainties on this procedure were assessed using two methods. One was to vary the beam spot size between 2×2 mm^2 to 8×8 mm^2. This yielded an error on G_i of about 0.2%. The central position was also varied by ± 1 mm in both the x and y planes. This also gave approximately a 0.2% error on value of G_i.

Calibration of 3×3 Crystal Matrix

Once the individual crystal calibrations had been obtained, the energy deposited by a shower within a 3×3 cluster of crystals was calibrated by summing up all the contributions from the 9 crystals. A similar procedure was performed as for individual crystals but with a sum over 9 crystals:

$$E_{\mathrm{beam}} = G_{3\times3} \times \sum_{i=1}^{9} G_i \times A_i \tag{4.6}$$

As with the individual crystal calibrations a Crystal Ball function is used to find the peak position in the electron distribution. A 10×10 mm^2 beam spot selection was required. The energy calibration coefficient $G_{3\times3}$ in Equation 4.6 is defined by setting the peak of the energy reconstructed in the 3×3 matrix to be equal to the beam energy (50 GeV).

4.3.3 Energy Resolution Measurements

The energy resolution has been estimated by calculating σ_{eff}. σ_{eff} is defined as half of the interval containing 68% (1σ) of the 3×3 electron ($E_{3\times3}$) distribution. The energy resolution is then defined as:

$$\frac{\sigma_{\mathrm{E}}}{\mathrm{E}} = \frac{\sigma_{\mathrm{eff}}}{\mathrm{E}} \tag{4.7}$$

where E is the peak value obtained from the Crystal Ball fit using the $E_{3\times3}$ distribution. This method was preferred to using the width from the Crystal Ball function, σ_{CB}, which doesn't take the tails of distributions into account, hence biasing σ_E towards lower values. A slightly worse resolution is found using σ_{eff} rather than σ_{CB}. A comparison between the two methods using a non-irradiated matrix is shown in Figure 4.6.

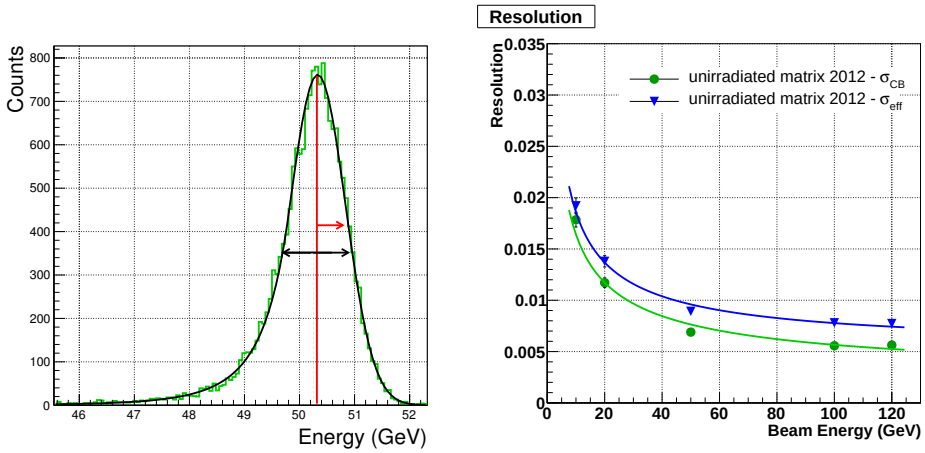

Figure 4.6: Left: The amplitude distribution for a single non-irradiated crystal using a 50 GeV electron beam. The red line shows the position of the peak estimated using a Crystal Ball fit and the red arrow represents the value σ_{CB}. The black arrow represents the smallest interval containing 68% of the distribution equivalent to $2\sigma_{\text{eff}}$. Right: a comparison of the energy resolution estimated using the two different methods is shown using a non-irradiated matrix from the 2012 test beam.

Noise Term Estimation

In the energy resolution measurements the contribution from the noise term (B in Equation 3.1) was determined using the pedestal and is fixed for each crystal matrix (see Table 4.1). The pedestal of each channel was continuously monitored. The mean pedestal value and the pedestal noise, σ^i_{ped}, were estimated by fitting the pedestal peak using a Gaussian function for each channel. The width of the Gaussian is referred to as σ_{ped} and has been used to estimate the noise contribution to the energy resolution.

The noise term has been calculated by adding in quadrature the contribution of the pedestal, σ_{ped}, for each channel involved in the reconstruction:

$$B = \sqrt{\sum_{i=1}^{9} (\sigma_{\text{ped}}^{i} \times G_i)^2} \tag{4.8}$$

where G_i is the ADC to GeV conversion coefficient for a specific channel i, which allows the different gain of each PMT to be taken into account (see Section 4.3.2).
The pedestal noise of a single channel varied between 40 and 150 MeV. The pedestal noise term, B, was 0.22 GeV for the central 3×3 cluster of the non-irradiated 2012 matrix. Table 4.1 shows the values of B obtained for each matrix configuration. The difference in noise term between matrices is mostly due to the intrinsic noise of the PMTs used in a certain test beam.

Year	Matrix #	B (GeV)
2011	cluster 12	0.17
2011	cluster 11	0.17
2012	1	0.22
2012	2	0.34
2012	3	0.19
2012	4	0.25

Table 4.1: Contribution of the noise term to the energy resolution for each configuration of the matrices used during test beam running.

4.4 Results

The results obtained during test beam periods from 2011 to 2012 are summarised in this section. The energy resolution and linearity reconstructed using 3×3 crystal matrices are shown for crystals with a range of μ_{ind} values. Test beam energies in the range 10 to 150 GeV were used.

The energy linearity of crystal matrices used in the 2011 and 2012 test beam set ups are displayed in Figure 4.7. The linearity curves have been fit

using the following parameterisation:

$$L(E) = 1 + S_{NL}(\mu_{\text{ind}}) \times log_{10}\left(\frac{E}{50}\right) \tag{4.9}$$

where S_{NL} is a non linearity parameter related to the μ_{ind} and E is the beam energy. S_{NL} is an 8^{th} order polynomial function of the μ_{ind} variable only [61, 62]. This is the reason perfect agreement with the data is not observed. This parametrization assumes that $L(50) = 1$ by construction.

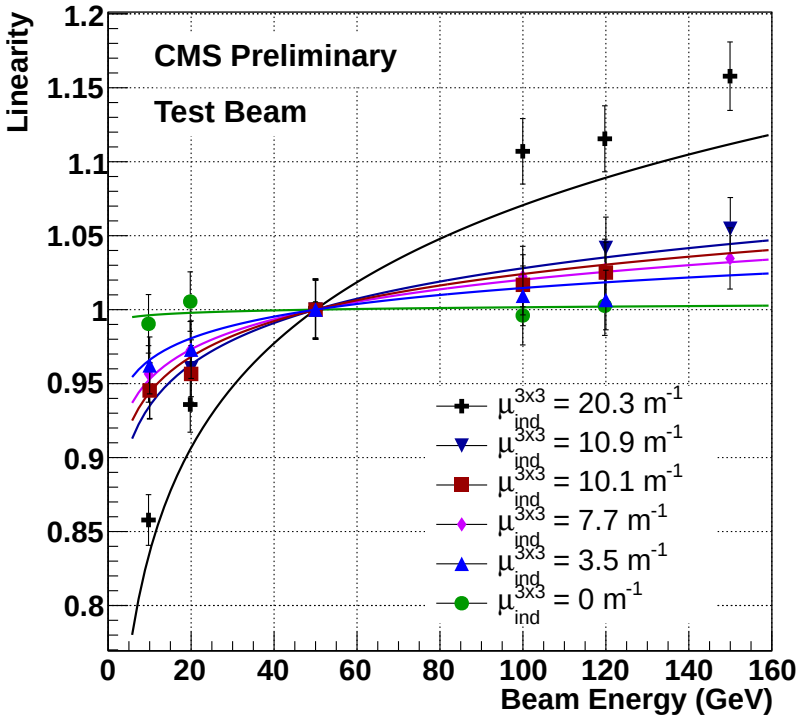

Figure 4.7: Fit of linearity for 2012 and 2011 irradiated matrices using the parametrization of Equation 4.9.

The energy resolution measurements using the method described in Section 4.3.3 are shown in Figure 4.8. The data has been fit using the parameterisation in Equation 3.1. The stochastic and constant term contributions to the energy resolution are shown in Table 4.2.

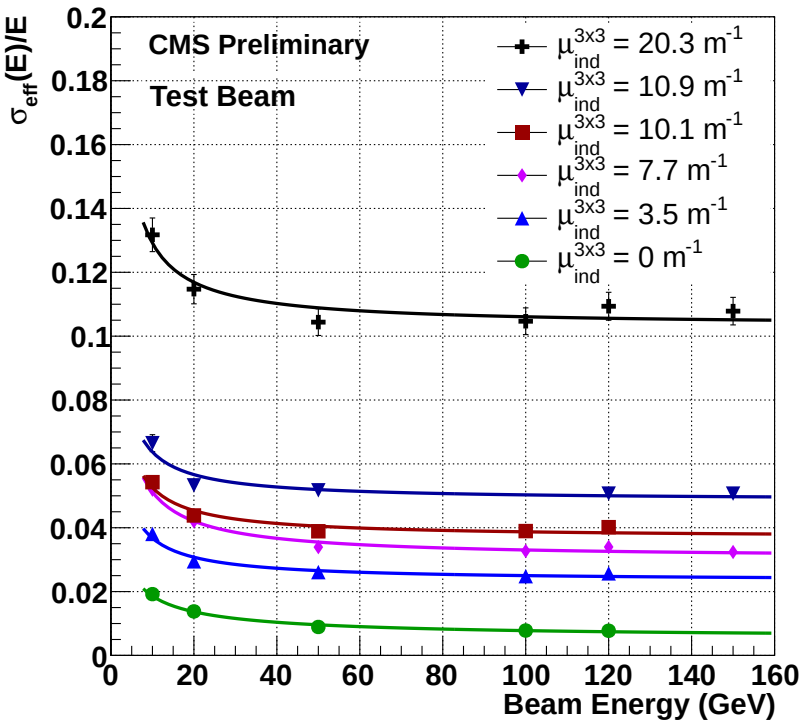

Figure 4.8: The energy resolutions as a function of beam energy are compared for different crystal matrices of varying radiation damage. The resolution is fit using Equation 3.1.

4.5 Discussion

The increasing non-linearity corresponding to higher values of μ_{ind} can be interpreted as the effect of transmittance loss leading to less efficient light collection from the crystal. Since the maximum of the shower, t_{max}, moves toward the rear face of the crystal with higher beam energy according to:

$$t_{max} \propto ln(E) \tag{4.10}$$

it is expected that light produced by low energy showers will suffer more from transmittance losses since it has to travel a longer path before reaching the PMT. This light will be more attenuated along the crystal introducing a non-

Matrix	$\mu_{\text{ind}}^{central}$ (m^{-1})	$\mu_{\text{ind}}^{3\times3}$ (m^{-1})	A (% $\text{GeV}^{1/2}$)	C (%)
2012 - 1	0.0	0.0	5.6 ± 0.2	0.5 ± 0.1
2011 - 1	3.4	3.5	8.9 ± 0.6	2.3 ± 0.2
2012 - 3	7.4	7.7	12.7 ± 1.0	3.1 ± 0.2
2011 - 2	10.9	10.1	11.8 ± 0.8	3.7 ± 0.2
2012 - 2	11.4	10.9	12.5 ± 1.4	4.9 ± 0.3
2012 - 4	21.7	20.3	24.5 ± 3.0	10.3 ± 0.5

Table 4.2: Stochastic and constant terms affecting the energy resolution as extracted from the fits.

linear effect on the energy reconstructed. This suggests a parametrization of the linearity as in Equation 4.9. The matrices of non-irradiated crystals show a good energy linearity in the range 10-120 GeV whereas the irradiated matrices show an increasing non-linearity, as can been seen in Figure 4.7.

The energy resolution decreases with increasing average μ_{ind} of the crystal matrices. Using the chosen parmeterisation of the energy resolution, the constant term evolves from 0.5% for non-irradiated matrices to 10.3% for the most damaged matrix of October 2012 (matrix 4). The constant term is expected to degrade with increasing average μ_{ind} of the crystals within a matrix. In this test beam configuration, there are several contributions affecting the constant term:

1. Different μ_{ind} of the crystals within a 3x3 cluster.
2. Partial containment of the shower inside a limited-size matrix.
3. Non uniformity of light collection efficiency along individual crystals.

4.6 Summary

The results presented, show the evolution of the performance of EE matrices constructed from 3×3 hadron-irradiated $PbWO_4$ crystals in terms of energy linearity and resolution. The set of matrices consists of $PbWO_4$ crystals with different levels of proton radiation damage corresponding to an induced

absorption μ_{ind} in the range 0-20 m^{-1}. Assuming no dose rate dependency and no annealing of the hadron damage, such induced absorptions coefficients are expected to largely cover the effect of the doses predicted for EE after an integrated luminosity of 300 fb^{-1} at $\eta = 2.3 - 2.5$.

Irradiated matrices show non-linearities in the reconstructed energy which increase with the loss of transparency of the crystals. The constant term of the energy resolution changes from 0.54% to 10.3%. These results suggest that the main effects of the proton radiation damage in $PbWO_4$ crystals can be described by a loss of the light transmission along the crystal. The results presented in this chapter provide important information on the effect of radiation in $PbWO_4$ crystals and provide information about how to validate a model for the damage within the CMS reconstruction software.

Chapter 5

Event Reconstruction and Weighting

In this chapter the transformation of the "raw" information from the CMS detector into the construction of physics "objects" such as particles, jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ will be described. In order to achieve this, a number of reconstruction algorithms have been implemented that help improve the resolution of measurable physics observables. This chapter outlines the variables used to identify specific particles such that mis-identified particles (fakes) can be minimised. The event weighting procedure is also described, which aims to correct for known differences between the Monte Carlo and the data.

5.1 Track and Primary Vertex Reconstruction

Tracking is the measurement of the path that charged particles take as they traverse the Tracking detector. It is vitally important for the measurement of the p_{T} of charged particles. The force, F, that a charged particle experiences is given by

$$\vec{F} = q(\vec{v} \times \vec{B}) \tag{5.1}$$

where q is the charge of the particle, B is the magnetic field strength and v is the velocity of the particle. Equation 5.1 can be rearranged such that the momentum, p, of the particle is related to the radius of curvature, r, of the particle within the detector

$$\vec{p} = q\vec{B}r \tag{5.2}$$

The tracking of particles is also important in the determination of the primary vertex, which is the location of the interaction point where the hard scattering was initiated. The reconstruction of a good primary vertex is key in separating the interesting physics events from soft QCD interactions (pileup) and also from the underlying event.

5.1.1 Track Reconstruction

The first step in track formation with CMS [63] is to cluster energy depositions from the inner pixel detector and tracker into "hits". So-called "seeds" are then formed using minimal tracking information. A particle trajectory is defined if at least three hits are present in various layers of the tracker. Hits are chosen based on their position and the associated uncertainty. A combination of two hits and a beam spot constraint can also be used.

Tracks are then constructed from the seeds using a Kalman Filter approach [64]. This involves propagating outwards through the tracker layers and forming potential tracks from hits on the outer layers. In some circumstances multiple hits can be compatible; in this case an additional track is formed. After this process, there can be multiple seeds assigned to tracks and multiple tracks assigned to seeds so some cleaning up is required; If more than 50% of the hits on a trajectory are multiply-assigned the track with the lowest number of hits is discarded. If the number of hits for both tracks are equal the track with the lowest χ^2 divided by the number of degrees of freedom is used where χ^2 is defined as:

$$\chi^2 = \sum_i \frac{(x_i - \mu_i)^2}{\sigma_i} \tag{5.3}$$

where x_i is a measured track position, μ_i is the expected position from ex-

trapolating a particles' trajectory at layers, i, in the tracker and σ_i is the uncertainty on the x_i measurement.

After the initial stringent fit criteria, the final particle track is reconstructed by refitting the Kalman Filter method. Hits with a large residual from the trajectory are removed. A track quality selection is then applied using the number of hits, normalised χ^2 of the track and the vertex compatibility. This helps to reject fake tracks in the case of high track multiplicity.

5.1.2 Primary Vertex Reconstruction

The reconstruction of the Primary Vertex (PV) is the determination of the interaction point from which the majority of tracks originate for a specific event. More than one primary vertex can be present in a bunch crossing due to pileup. A PV is reconstructed using tracks that fulfill certain quality criteria such as the χ^2, the number of hits in the inner tracking system and the impact parameter with respect to the beam spot position. These tracks are input into a vertex-fitting procedure [65]. This uses a Kalman Filter approach that calculates a weight related to the likelihood that a given track is associated with a particular vertex position. The sum of these weights relates to the number of degrees of freedom (ndof) associated with a PV. More than four dofs are required for a PV to be selected. The $|z|$ position of the PV along the beam line must be within 24 cm of the center of the detector and within a distance, ρ, of less than 2 cm from the nominal interaction point (beam spot) in the x-y plane.

5.2 The Particle Flow Algorithm

The particle flow (PF) reconstruction [66] algorithm aims to reconstruct all stable (long lived) particles within an event. These particles include charged and neutral hadrons, photons, electrons and muons. The information from all sub-detectors is combined for optimal identification of individual particles. The algorithm produces a list of all particles in the event, which can then be input into a clustering algorithm to reconstruct higher order objects such

as jets, hadronically decaying taus and $E_{\mathrm{T}}^{\mathrm{miss}}$. The CMS detector is ideally suited to use this algorithm with its large silicon tracker immersed within a high magnetic field. The algorithm combines information from the tracker and the calorimeters to form "blocks" in the detector. These blocks are then interpreted as particles.

5.2.1 Iterative Tracking

The iterative tracking algorithm is used to identify charged particle tracks. This algorithm starts by first seeding and then reconstructing tracks as described in Section 5.1 with a tight selection criteria. Once a well defined track has been formed its hits are removed from the next iteration of track forming. The next track is identified using a slightly looser selection requirement. The PV requirement is relaxed after more than three iterations, which helps to reconstruct charged tracks coming from photons or neutral hadron decays. Some longer lived neutral hadrons and some photons may not leave a track at all and their presence can only be inferred from energy depositions in the calorimeters.

5.2.2 Calorimeter Clustering

The hadronic and electromagnetic calorimeters are important in the identification and energy measurement of neutral hadrons and photons. Calorimetry is also used to help improve the energy/momentum measurement for particles and is particularly useful for hadrons that are associated with low quality tracks. The algorithm is used separately for each sub-calorimeter. The first step of the algorithm is to form "seeds" within cells where the energy deposited is greater than a certain threshold. This threshold is defined by the noise within the calorimetry electronics: 80 MeV in the ECAL barrel, 300 MeV in the ECAL endcap and 800 MeV in the HCAL. Next, topological clusters are formed by aggregating adjacent cells where the adjacent cell has an energy of more than 2σ above the average noise expected for a cell within that sub-detector. Finally, the granularity of the calorimeter is exploited to distinguish between particle flow clusters within topological clusters such

that the number of particle flow clusters match the number of initial seeds. This process helps distinguish particles having a low spatial separation.

5.2.3 Linking Algorithm

A linking algorithm is next performed to match tracks with the clusters. This allows the path of particles to be mapped as they traverse the detector and helps to remove double counting of tracks. A connection is made between a track and a cluster by extrapolating the trajectory of the track outwards from the tracker. This extrapolation is performed out to the expected shower maximum for a typical electron in the ECAL or the typical interaction length of a hadron shower in the HCAL. The quality of a match is defined by the distance in the $\eta\phi$ plane between a track and a matched cluster. A link is formed if the extrapolated track is within the boundary of a calorimetric cluster. Links can also be formed between pairs of clusters in the ECAL and HCAL or the ECAL and preshower sub-detectors in the case where the cluster in the more granular detector is within that of the less granular.

5.2.4 Particle Identification

The final part of the PF algorithm is the identification of particles types within an event. The algorithm starts by matching linked "blocks" to the most easily identifiable particles. As particles are matched to "blocks" by passing a set of identification criteria (outlined in the following sections), their tracks and clusters are removed from further processing. The algorithm then moves on to particles where the identification is more ambiguous. The algorithm starts by identifying electrons. Muons that are not contained within a jet are then identified by linking the inner tracks with those in the muon chambers. Muons within a jet are required to pass tighter identification requirements as charged pions can quite easily be misidentified.

Charged hadrons, neutral hadrons and photons are then identified. This begins with the identification and removal of tracks from charged hadrons. This is done by comparing the energy of matched clusters in the ECAL with those in the HCAL and also with the momentum measured from tracking. If

the charged hadron hypothesis is satisfied these clusters and tracks are then removed. In the case where multiple tracks match a cluster, the sum of the track momenta is required to be compatible with the calorimetric energy. In the opposing case where multiple clusters are linked to a single track, the momentum of the track is compared to the energy of the closest linked cluster. If the values are compatible, the track and corresponding cluster are removed. Otherwise, the sum of the energy in neighbouring clusters is used until there is a reasonable match to the track p_{T}.

If there are still clusters present at this stage, they are considered to be photons, unless there is an excess energy in the HCAL compared to that in the ECAL, in which case a neutral hadron is identified. The assignment is done in this order because photons are expected to carry around 25% of jet energy compared to only 3% from neutral hadrons.

5.3 Electron Reconstruction and Identification

Electrons are reconstructed [67] using information from both the ECAL and the tracker. The reconstruction starts with the deposition in the ECAL. This is then matched to the seeds within the tracker. The trajectory, and therefore the momentum can then be calculated using the particle flow algorithm. Bremsstrahlung associated to the electron must be included in its reconstruction.

The deposition in the ECAL is measured by clustering the hits within cells into clusters. In the EE, the cluster is propagated back to the preshower (PS) detector and the deposition here is also used in constructing the trajectory. As a result of CMS's large magnetic field, bremsstrahlung photons can be spread quite widely in ϕ from the primary electron. The deposition within an ECAL supercluster (5×5) is therefore used. Superclusters (SCs) with an energy deposition of greater than 4 GeV are selected and then matched to a tracker seed. The electrons' trajectory, and hence its p_{T} are then obtained using the particle flow algorithm.

Identity variable	tight	loose
Electron Identification	Particle Flow: MVA > 0.5	MVA > 0.5
Exclusion of EE-EB transition region ($1.4442 < \|\eta\| < 1.566$)	applied	not applied
Transverse impact parameter w.r.t primary vertex	< 0.02	< 0.04
Photon conversion rejection	applied	applied
Missing inner tracker hits	0	0
PF Relative Isolation	< 0.1, cone 0.3	< 0.15, cone 0.3

Table 5.1: Table containing electron identification requirements for both tight and loose selection as prescribed by the CMS top group.

These electron candidates must then pass a set of selection and identification criteria [68]. These aim to minimise mis-identification and are outlined in Table 5.1. A simple selection-based [69] criteria was used for the 2011 data with a more sophisticated multivariate (MVA) identification [70] requirement applied in 2012. The efficiencies of the MVA identification as a function of p_{T} in both the barrel and the endcap can be seen in Figure 5.1. Similar variables are used in both the selection-based and MVA identification. These include tracking variables such as the χ^2 of the track divided by the number of degrees of freedom and the impact parameter with respect to the primary vertex. Geometric properties of the candidate electron are also used including the tracks spread in both η and ϕ. Shower shape variables and energy matching terms are also included such as H/E, the ratio of the energy deposited in the HCAL to that in the ECAL, and E/p which is the energy measured in a ECAL supercluster divided by the momentum measured using the particle flow algorithm.

5.3.1 Photon Conversion Veto

An additional bit of identification is required to make sure the electron does not originate from a converted photon. Electrons can be misintrepreted in an event when a photon produces a pair of electrons. To minimise this occurrence, photon conversion and missing inner track vetos are applied. Photon conversions are identified by combining opposite sign pairs from a secondary vertex in the tracker. Photons do not leave hits in the inner pixel detector and hence missing hits would indicate that a photon produced the detected electron. Events containing a pair of electrons from a pair of

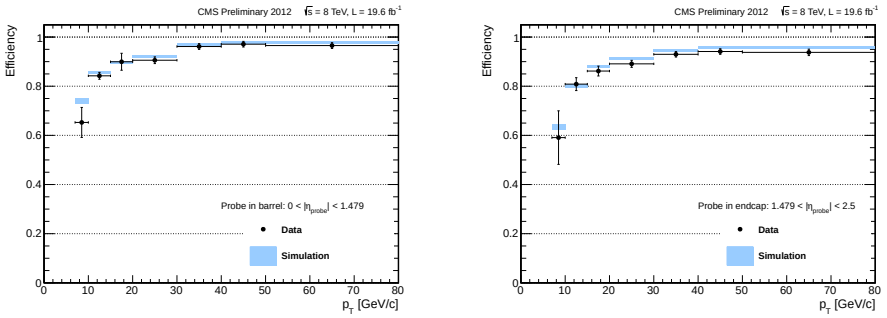

Figure 5.1: Identification efficiencies for both the data and the MC extracted from the the tag and probe method (see Section 5.7.1) using the MVA identification criteria. The left plot shows the efficiencies in the barrel and the right for electrons detected in the endcap. Taken from [71].

converted photon are rejected.

5.3.2 Relative Isolation

In top pair events, leptons are produced from the prompt decay of a W boson and are expected to be produced away from other event activity. This feature of an electron in events is characterised as relative isolation ($R^{\text{iso}}_{\text{lep}}$). It can be defined using all particles within an event that are reconstructed using the particle flow algorithm by the following:

$$R^{\text{iso}}_{\text{lep}} = \frac{\sum p_T^{\text{charged}} + \sum E_T^{\gamma} + \sum E_T^{\text{neutral}}}{p_T^{\text{lep}}} \tag{5.4}$$

where $\sum p_T^{\text{charged}}$, $\sum E_T^{\gamma}$ and $\sum E_T^{\text{neutral}}$ are the sums of the p_T or E_T of all the charged hadrons, photons and neutral hadrons, respectively, within a cone around the lepton in question. The cone radius, ΔR, is defined as:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2} \tag{5.5}$$

where $\Delta\eta$ and $\Delta\phi$ are radial distances in pseudorapidity and ϕ from the electrons' path. A cone size of $\Delta R = 0.3$ is used for electrons. Corrections are also applied to the isolation to remove the effect of neutral hadrons and

photons arising from pileup being included in the isolation cone. The pileup energy is simply subtracted in the calculation of the relative isolation:

$$R_{\text{lep}}^{\text{iso}} = \frac{\sum p_T^{\text{charged}} + \max(\sum E_T^{\gamma+\text{neutral}} - \rho \cdot A_{\text{eff}}^{\gamma+\text{neutral}}, 0.)}{p_T^{\text{e}}} \tag{5.6}$$

Here, ρ is a measure of the pileup energy density in an event and is calculated using jet activity from other interaction vertices. $A_{\text{eff}}^{\gamma+\text{neutral}}$ is defined as the effective area and is the probability of finding a photon or neutral hadron within a certain isolation cone as a function of pseudorapidity. The effective rather than the geometric area is used to take detector effects into account. The A_{eff} values used were obtained from 12.1 fb^{-1}of 2012 $Z \to ee$ data [72].

5.4 Muon Reconstruction and Identification

Muons have a long lifetime of $\tau \sim 2.2$ μs. This means that high p_{T} ($p_{\text{T}} \gtrsim 20$ GeV) muons can escape the entire CMS detector without losing all of their energy through Coulomb scattering. Low p_{T} muons ($p_{\text{T}} \lesssim 5$ GeV), however, are usually stopped in the detector and decay via the electroweak interaction to an electron and two neutrinos. Muon reconstruction is performed using the particle flow algorithm as has been previously described in Section 5.2. There are some slight deviations to the standard particle flow algorithm for muons which will be described. Muons produced in $t\bar{t}$ events are from the decay of a W boson and so are produced close to the primary vertex and isolated from other event activity. Reconstruction starts by building tracks separately in the tracker and muon detection system. The information from these two sub-detectors is reconstructed in two different ways:

- Global muon reconstruction - “from outside-in”
- Tracker muon reconstruction - “from inside-out”

5.4.1 Global Muon Reconstruction

Global muon reconstruction first reconstructs tracks in the muon detection system. The track is then propagated inwards using the Kalman Filter technique [64] starting with local segment seeds. If other hits have a compatible χ^2/ndof then these are included and the trajectory is recalculated. After the track has been reconstructed in the muon detector, it is propagated inwards taking into account the magnetic field and multiple scatterings effects. This continues until the closest approach to the beam line is reached. Since there is a much higher track density in the inner tracker, there could be multiple tracks compatible with the track propagated from the muon detector. Initially tracks are selected that are within an $\eta \times \phi$ region of the propagated track. Then more stringent p_{T} and spatial requirements are requested of the inner tracks. Selected inner tracks are then propagated outwards until it meets the inwardly propagated track from the muon system at a common surface between the two subdetectors. A quality test is done on the compatibility of the two tracks. The hits in both the inner tracker and the muon detector are then used to form a global track. If there is more than one compatible track the one with the lowest χ^2/ndof is used.

5.4.2 Tracker Muon Reconstruction

Tracker muon reconstruction relies more on the tracker information and only uses minimal information from the muon detection system. Tracks with p_{T} > 0.5 GeV/c and a total momentum greater than 2.5 GeV/c are considered as possible muon candidates. As in the global muon case the track is extrapolated outwards to the muon chambers. As long as the track can be matched to within a radial distance of 3 cm of a segment (few hits in the muon DT or CS detector) or has a pull of less than 4 then it is selected as a tracker muon. The pull is defined as:

$$\mathrm{pull} = \frac{(\mathrm{X_{ex}} - \mathrm{X_{hit)}}}{\sigma_{\mathrm{X}}} \tag{5.7}$$

where X_{ex} is the extrapolated position, X_{hit} is the hit position in the muon detector and σ_X is the combined error on both positions.

The tracker algorithm obtains a better efficiency for low p_{T} muons ($\text{p} \lesssim 5$ GeV) since lower p_{T} muons don't always pass through many of the outer muon detectors before losing all of their energy. The global muon algorithm improves the momentum resolution for higher p_{T} muons that escape the detector before losing all of their energy.

5.4.3 Identification

Global and tracker muon reconstruction algorithms provide the basis for tight and loose muon identification. A "tight" muon is required to be reconstructed globally whereas a "loose" muon may be reconstructed using either the global or tracker algorithms. Tight muons are then also required to pass some additional requirements that are outlined in Table 5.2. These requirements are related to the quality of the track, the impact parameter with respect to the PV and the z axis and the number of hits in various parts of the detector.

Identity variable	tight	loose
PF Muon Reconstruction	true	true
Muon Reconstruction Algorithm	Global muon	Global or Tracker muon
norm χ^2/dof	< 10	not applied
Tracker layers with measurement	> 5	not applied
Number of valid muon chamber hits	> 0	not applied
Transverse impact parameter w.r.t primary vertex	< 0.2 cm	not applied
Vertex$_\mu(z)$-PV(z)	< 0.5	not applied
Number of hits in inner pixel tracker	> 0	not applied
Matched hits in muon stations	> 1	not applied
PF Relative Isolation	< 0.12, $\Delta R = 0.4$	< 0.20, $\Delta R = 0.4$

Table 5.2: Muon identification requirements are shown for both "tight" and "loose" selections as prescribed by the CMS Muon object group [73].

5.4.4 Isolation

As with the electrons there is an isolation requirement for the muons. Again, particle flow isolation is used in a similar fashion as for electrons. There are, however, two differences compared to the electron case. The first difference is

that a slightly different method is used to remove the PU from the isolation cone as seen in the following equation:

$$R_{\text{lep}}^{\text{iso}} = \frac{\sum p_T^{\text{charged}} + \max(\sum E_T^{\gamma+\text{neutral}} - 0.5 \cdot \sum p_T^{\text{PU}}, 0.)}{p_T^{\mu}} \tag{5.8}$$

where $\sum p_T^{\text{PU}}$ is the sum of the p_{T} of the charged pileup within the cone of interest. The factor of 0.5 corresponds to the naive estimation of the neutral to charged hadron ratio in jets [74]. The second difference from the electron channel is that a cone of $\Delta R = 0.4$ rather than 0.3 is used.

5.5 Jet Reconstruction and Identification

Jets are an important feature in top pair events as two b quarks are produced. In the semi-leptonic channel two additional quarks are also produced from the decay of one of the W bosons. Jets are produced in the detector from the hadronisation of these quarks. Jets are reconstructed by combining particles reconstructed and identified using the PF algorithm described in Section 5.2. PF reconstruction has been shown to improve energy resolution substantially over using calorimetric reconstruction as seen in Figure 5.2 and gives closer values with respect to generator/truth level. There are various algorithms that can be used to reconstruct jets. The two main requirements are that the reconstructed jets should be:

- Infra-red safe: should not be affected by soft gluon emissions;
- Collinear safe: should not be affected by collinear parton splitting.

Along with these two requirements, an algorithm that has a low sensitivity to pileup and the underlying event is preferred.

5.5.1 Reconstruction

CMS currently uses the "anti-k_t" algorithm [75], which sequentially clusters particles into jets. The algorithm starts with the highest p_{T} particle, i,

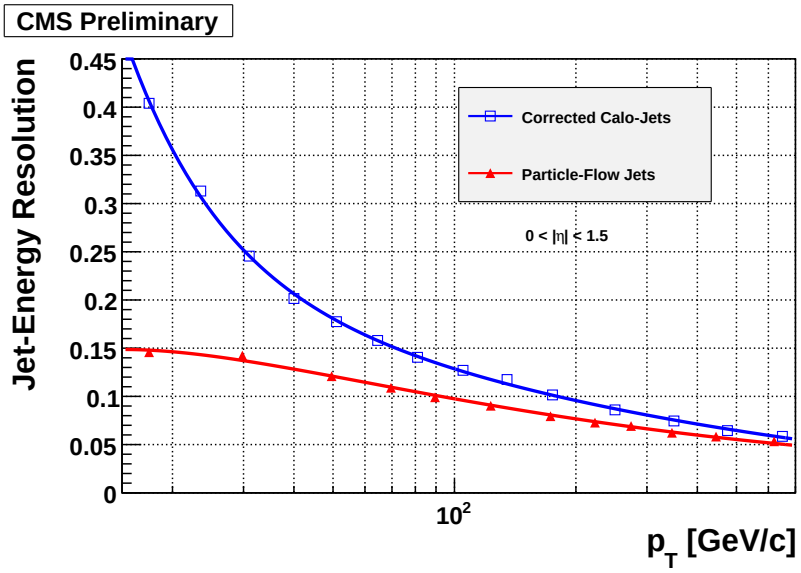

Figure 5.2: Jet-Energy resolution as a function of jet p_{T} for calorimetic and particle flow jets from simulation. An improved resolution is observed particularly for low p_{T} jets. Taken from [66].

and iterates over softer nearby particles calculating effective distances, d_{ij}, between neighbouring particles and the beam, d_{iB}:

$$d_{ij} = \min(k_{ti}^{2p}, k_{tj}^{2p}) \frac{\Delta_{ij}^2}{\Delta R^2} \tag{5.9}$$

$$d_{iB} = \frac{1}{k_{ti}^2} \tag{5.10}$$

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 azimuthal angle, respectively, of the particle. ΔR is the cone radius (0.5) as defined in Equation 5.5 and p is the parameter that defines the choice of clustering algorithm [76]. The anti-k_t algorithm uses a p value of -1. If the value of d_{ij} is less than d_{iB} then the softer particle is included in the jet, otherwise it will form part of another jet. The result of this approach is that soft particles surrounding a hard particle form conical jets.

5.5.2 Identification

PF reconstructed jets are required to pass a set of fairly loose quality requirements which are outlined in Table 5.3. In addition to this jets that are within ΔR of less than 0.5 to an electron or muon which pass the tight selection (defined in Tables 5.1 and 5.2) are rejected. This removes jets that were wrongly identified as electrons or muons.

Identity variable	selection requirement
Number of particles in jet	> 1
Neutral hadron energy fraction	< 0.99
Photon energy fraction	< 0.99
Charged EM energy fraction ($\|\eta\| < 2.4$)	< 0.99
Charged hadron energy fraction ($\|\eta\| < 2.4$)	> 0
Charged hadron multiplicity ($\|\eta\| < 2.4$)	> 0

Table 5.3: The loose jet identification requirements as prescribed by the CMS Jet object group [77].

5.5.3 B-tagging

The presence of two b quarks makes the identification of top pair events more recognisable. A method called b-tagging aims to identify jets produced by b quarks as opposed to lighter quarks or gluons (u, d, c, s, g). The properties of the bottom quark are used to make this distinction. A b quark decays via the weak interaction to either a charm or an up quark. These decays are suppressed by the CKM matrix and so the b quark has a relatively long lifetime ($\tau_b \sim 1.5$ ps). A relativistically travelling b quark would have a decay length of ≈ 450 μm. This is an observable length for the high resolution CMS pixel tracker. A displaced secondary vertex can therefore be used (as shown in Figure 5.3) as a b-tagging technique (Simple Secondary Vertex, SSV algorithm [78]).

The b quark has a large mass ($m_b = 4.18$ GeV/c^2) and so a large multiplicity of charged particles are produced during hadronisation. The b hadron carries a large fraction of the jet energy. In its decay the off-shell W boson

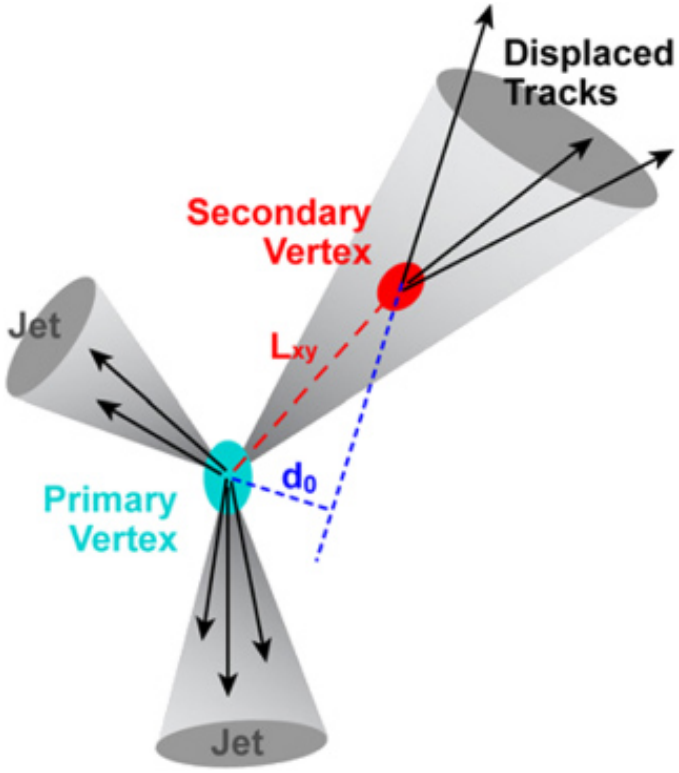

Figure 5.3: Schematic of the displacement of secondary vertex (SV) due to the b decay from the PV. Here the PV is associated with the tracks from two jets. Displaced tracks are used to reconstruct a SV. Each track has an impact parameter, d_0, which is its displacement with respect to the PV along the z-axis. L_{xy} is the transverse distance of the SV from the PV in the x-y plane.

decays with 20% probability to an electron or a muon so b-jets can often be associated with one of these leptons. Particle tracking information combined with the properties of the secondary vertex can therefore be used as input to a b-tagging algorithm. The algorithm that is currently used is called the Combined Secondary Vertex (CSV) algorithm [79].

All jets within an event are possible b candidates. The CSV algorithm starts by looking for a secondary vertex. All reconstructed particle tracks within a jet are input [80]. A similar Kalman Filter Vertex Finder [81] to that used for the PV is implemented to reconstruct the SV. The vertex must then pass a set of requirements to be classified as a true reconstructed vertex:

- The distance, L_{xy}, in the transverse plane between the primary and secondary vertex must be greater than 100 μm and be less than 2.5 cm.
- $L_{xy}/\sigma_{L_{xy}} > 3$, where $\sigma_{L_{xy}}$ is the error on L_{xy}.
- The invariant mass of the charged particles associated to the secondary vertex must not exceed 6.5 GeV/c^2 as this would exceed the the mass of a bottom quark, even taking into account the uncertainties associated with the value.
- A veto is applied for K_s^0 decays ($K_s^0 \to \pi^+\pi^-$), where an oppositely signed pair of particle tracks originate from a vertex with an invariant mass within 50 MeV/c^2 of the K_s^0 mass (≈ 0.5 GeV/c^2).

If a secondary vertex is not found a pseudo-secondary vertex is identified using tracks that are not compatible with the primary vertex. Otherwise no secondary vertex is found.

A set of discriminating variables is then combined into a single discriminator that is used to separate the charm and the lighter quark jets from the b-jets. The set of discriminating variables includes the invariant mass of the charged particles associated to the vertex, the number of charged tracks associated to the vertex, $L_{xy}/\sigma_{L_{xy}}$ and the energy of the charged particles associated to the vertex divided by the energy of the charged particles associated to the jet. B-tagging can be performed at various working points using

this discriminator. The partons that initiated candidate b-jets in $t\bar{t}$ enriched events is presented in Figure 5.4. Tight, medium and loose working points can be used depending on the purity requirements of a particular analysis. The loose, medium and tight working points are at 0.244, 0.679 and 0.898, respectively. The medium working point has a b-tagging efficiency of roughly 65% and a mis-tag (wrongly tags u, d, s, g as b quarks) rate of 1% [82].

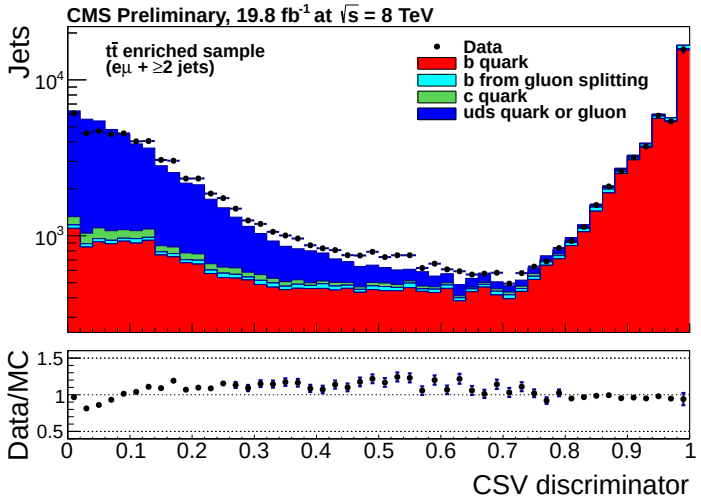

Figure 5.4: Logarithmic distribution of the partons as a function of the CSV b-tagging discriminant. Taken from [83].

5.6 Missing Transverse Energy Reconstruction

The missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) in an event is defined as the imbalance of transverse energy in an event. This arises naturally in semi-leptonic $t\bar{t}$ events due to the energy carried away by the undetected neutrino produced in the W decay. $E_{\mathrm{T}}^{\mathrm{miss}}$ is also used in a lot of searches for new physics since new exotic particles may be undetectable. $E_{\mathrm{T}}^{\mathrm{miss}}$ is calculated by obtaining the vectorial sum of the transverse momenta of all particles, i, identified

using the PF algorithm in an event:

$$E_{\mathrm{T}}^{\mathrm{miss}} = -\sum_{i}(p_x^2 + p_y^2)^{\frac{1}{2}}. \tag{5.11}$$

$E_{\mathrm{T}}^{\mathrm{miss}}$ can also arise unnaturally as a result of random electrical noise in a certain part of the detector. Noise filters are applied to remove events where there is excess noise in certain parts of the detector. The resolution of the $E_{\mathrm{T}}^{\mathrm{miss}}$ plotted against the sum of the transverse energy can be seen in Figure 5.5. Roughly a factor of two improvement is seen when using PF reconstruction over purely calorimetric reconstruction.

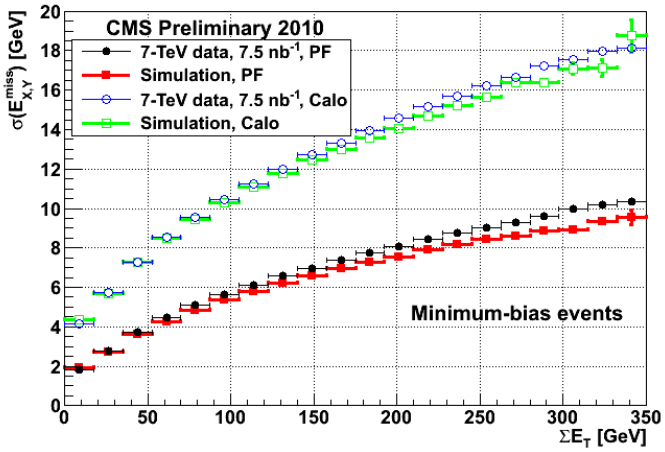

Figure 5.5: Resolution of $E_{\mathrm{T}}^{\mathrm{miss}}$ vs. the scalar sum of the E_{T} of all particles in an event using 7.5 nb^{-1} of 2010 data compared to simulation. Both PF and calorimeter reconstructed $E_{\mathrm{T}}^{\mathrm{miss}}$ are shown. Taken from [84].

5.7 Event Reweighting

The MC simulation does not always perfectly describe the data. An example of this is that in the 2011 data taking the amount of pileup changed as the instantaneous luminosity increased. The simulation included pileup assuming that it was flat up to 10 pileup events and had a Poissonian tail. As such,

reweighting had to be performed to ensure that the MC matched the data.

To correct for mis-modelling in the MC, correction factors are derived from the data. Well known physics processes such as Drell-Yan (Z boson decays to lepton pairs) are usually used to extract the event weights. Weights have been applied to account for the following data/MC mismatches:

- Trigger efficiency (SF_{Trig_μ});
- Lepton Identification and Isolation ($SF_{\mathrm{Id-Iso}_\mu}$);
- pileup Scale Factors (SF_{PU});
- B-tagging Scale Factors ($SF_{\mathrm{b-tag}}$).

A weight must also be applied for each MC sample to reweight the total number of events produced by the MC generator, $N_{\mathrm{processed}}$, to the integrated luminosity, $\mathcal{L}$, of the data and the theoretical cross section, σ, of the simulated process. The following scaling factor, SF_{MC}, is obtained:

$$SF_{\mathrm{MC}} = \frac{\mathcal{L} \times \sigma}{N_{\mathrm{processed}}} \tag{5.12}$$

All of these scale factors are applied on an event-by-event basis. The overall scaling factor, SF, is given by their product:

$$SF = SF_{\mathrm{MC}} \times SF_{\mathrm{Trig}_\mu} \times SF_{\mathrm{Id-Iso}_\mu} \times SF_{\mathrm{PU}} \times SF_{\mathrm{btag}} \tag{5.13}$$

An additional weight was applied to $t\bar{t}$ MC events to account for an observed difference between the data and MC top p_{T} distributions [85].

5.7.1 Muon Scale Factors

A trigger is applied to the data during the data taking process. The trigger being applied in the MC attempts to replicate the same implementation but it is not always able to mirror the exact details of the actual trigger. The "tag & probe" method [86] is used to measure the trigger efficiency as a function

of muon p_{T} and η for both data (ϵ_{data}) and MC (ϵ_{MC}). The scale factor is then given by:

$$SF = \frac{\epsilon_{\mathrm{data}}}{\epsilon_{\mathrm{MC}}} \tag{5.14}$$

In the tag and probe method the di-muon resonance from the decay of a Z boson is used. A tag is defined as a muon that passes a tight selection. The probe is a muon which passes only a loose muon requirement and also has an opposite charge to the tight muon. A passing probe is then a probe muon that passes a desired selection requirement, so in the case of the trigger it would have to pass the isolated muon trigger. The efficiency of this desired selection requirement is then:

$$\epsilon = \frac{N^{\mathrm{pass}}_{\mathrm{probe}}}{N^{\mathrm{all}}_{\mathrm{probe}}} \tag{5.15}$$

where $N^{\mathrm{pass}}_{\mathrm{probe}}$ is the number of passing probes and $N^{\mathrm{all}}_{\mathrm{probe}}$ is the total number of probes within the Z mass resonance. Drell-Yan enriched data is used to calculate the efficiency for data with Drell-Yan Monte Carlo being used for the MC.

The trigger scale factors used in the analysis were centrally produced by the CMS muon group for the various data taking periods [87]. The 2011 efficiencies for the muon η distribution are plotted in Figure 5.6.

Scale factors also need to be applied to account for muon identification and isolation differences between the data and MC. These SFs are again obtained using the tag and probe method, except now the probe must pass the tight identification or isolation requirement rather than the trigger. Figure 5.7 shows the identification and isolation scale factors for the 2011 data set. These muon SFs are dependent on the p_{T} and η of the "tight" muon in the event.

5.7.2 Pileup Reweighting

Due to the rapidly changing pileup conditions, particularly during the 2011 data taking period, it is impossible for the Monte Carlo to precisely model

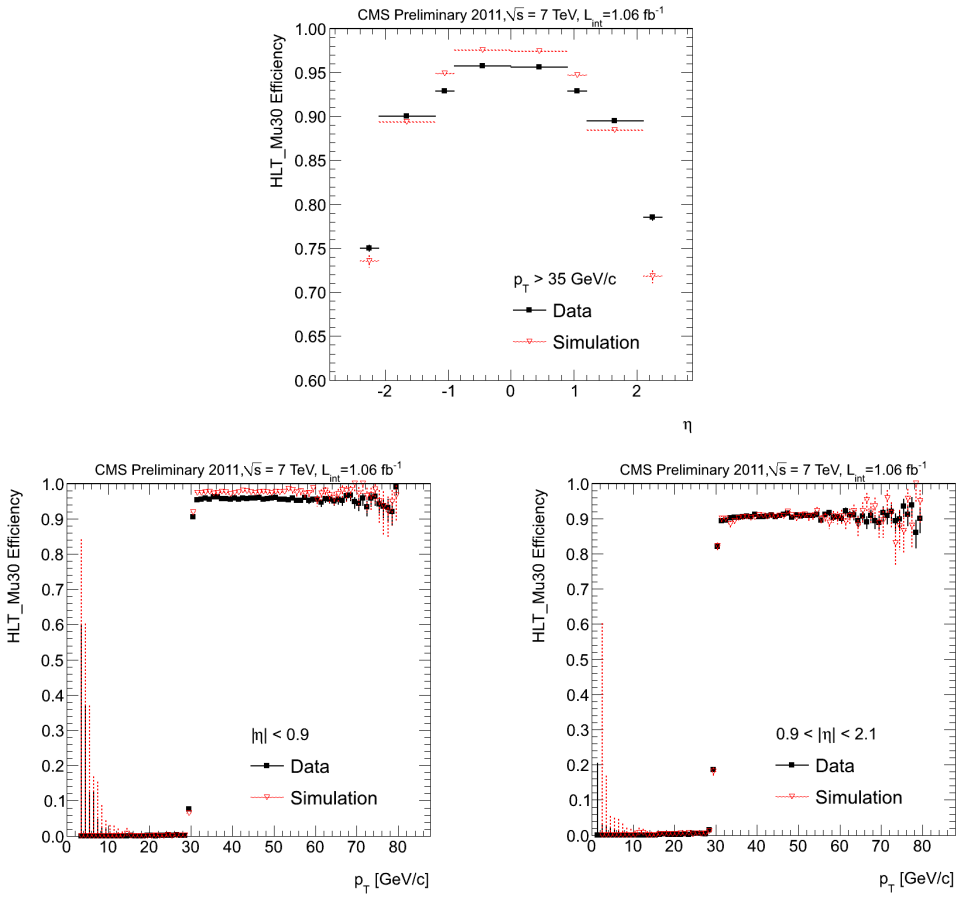

Figure 5.6: Efficiencies for data and MC passing the HLT_Mu30 trigger as a function of muon η (top) and p_{T} in the barrel (bottom left) and the endcaps (bottom right) for 1.1 fb^{-1}of 2011 data. The scale factors are obtained by dividing the data by MC. A 35 GeV p_{T} cut was made on the muon in the η study to ensure that the muon is well clear of the trigger turn on at 30 GeV. Images taken from [87].

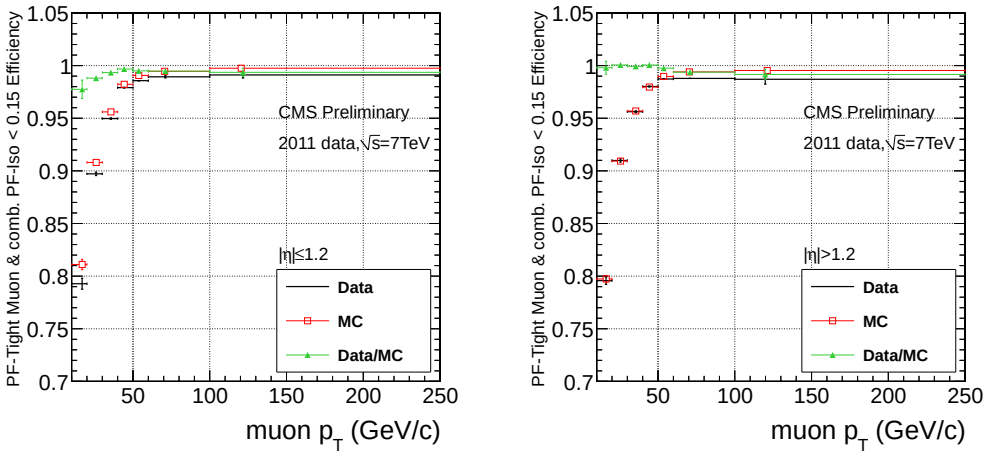

Figure 5.7: Efficiencies for data and MC along with $SF_{\mathrm{Id-Iso}_{\mu}}$ (green) as a function of muon p_{T} for 1.1 fb^{-1}of 2011 data. The left plot shows the barrel ($|\eta| < 1.2$) and the right plot the endcap ($|\eta| > 1.2$) region. Since the tag and probe is a data driven method the error bars are statistical only. Images taken from [87].

pileup. In the MC produced in the summer of 2011 for example, events were simulated with a flat pileup distribution up to 10 interactions with a Poissonian tail. The situation improved for the 2012 MC production somewhat. It still however isn't perfect as can be seen in Figure 5.8. This MC distribution clearly doesn't match the PU distribution of the data which is more or less Poissonian distributed with a mean of about 14 interaction vertices per bunch crossing.

To perform the reweighting, normalised PU distributions in data and the MC are used. The PU distribution for MC is taken directly from a centrally provided set of normalised values for the required MC production cycle. The PU distribution for the data is estimated using the number of in-time interactions per luminosity block. The estimated instantaneous luminosity for the runs included in the data sample and the total inelastic proton-proton cross section are used ($N_{\mathrm{PU}} = \mathcal{L} \times \sigma_{\mathrm{pp}}$) to produce a distribution of the expected number of pileup interactions. Figure 5.9 shows the true number of interactions per bunch crossing for different pp inelastic cross sections for

both the 7 and 8 TeV data. The event weights are then calculated using:

$$SF_{\mathrm{PU}} = \frac{N_{\mathrm{Normalised\ data}}}{N_{\mathrm{Normalised\ MC}}} \tag{5.16}$$

These weights are normalised such that their mean is 1, leaving the total number of simulated events unaffected.

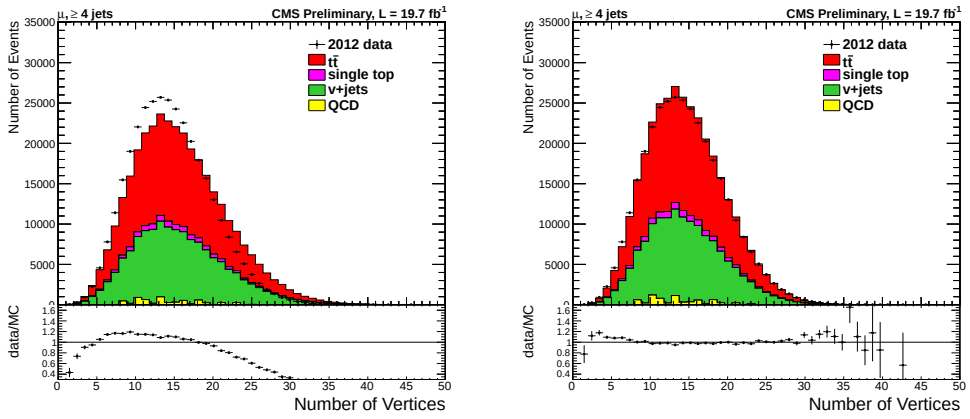

Figure 5.8: Number of reconstructed vertices in the 2012 data set. Events with at least four jets are used. The left shows the number of vertices before, and the right plot after, pileup reweighting. A vast improvement in the modelling of the data is observed.

5.7.3 B-tagging Scale Factors

In the analysis performed on the 8 TeV data b-tagging was implemented. A p_{T} dependent b-tag scale factor is applied to each b-tag for simulated events to account for the differences between data and simulation [88]. These scale-factors are derived using a combination of b-enriched QCD and $t\bar{t}$ data sets [89]. Simulated events passing the b-tagging requirements are given a weight (following the prescription from the b-tagging physics object group [90]) to ensure that the probability of selecting a simulated event matches the probability of a data event with the same jet flavour composition being selected.

The results of the reweighting can be seen in Figure 5.10. The number of MC events in the 0 and 1 b-tag bins are scaled up while the number of

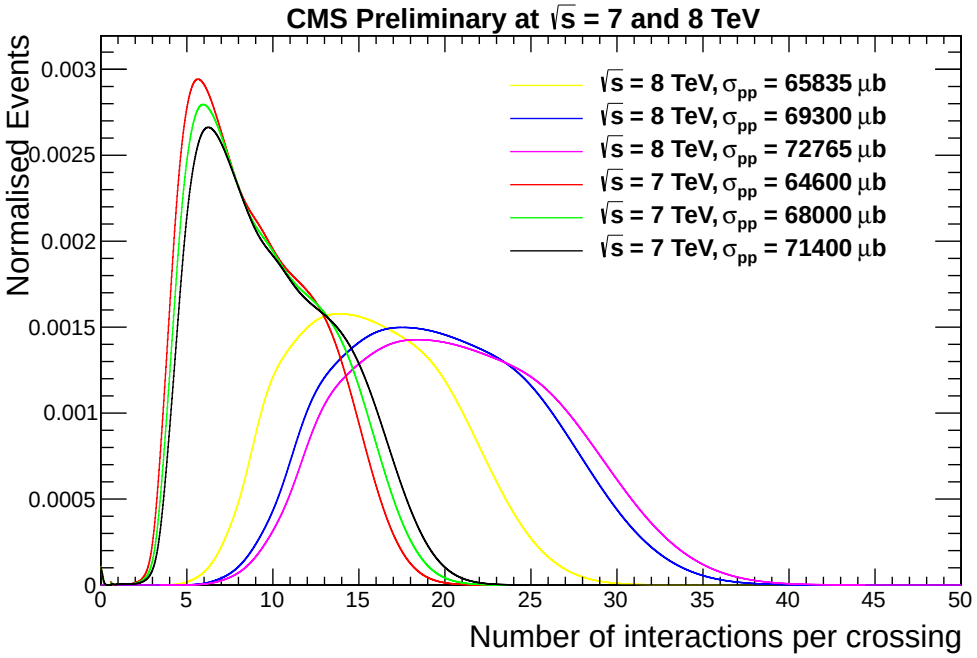

Figure 5.9: The normalised number of expected vertices per bunch crossing for various inelastic proton-proton cross sections for the 2011 and 2012 data.

MC events with b-tag multiplicity greater than 1 are scaled down. There is a slight mismatch between the data and MC for higher b-tag multiplicities. This can be put down to the fact that MADGRAPH MC is only simulated with up to three additional partons.

5.7.4 Jet Energy and Resolution Corrections

Particles produced in the underlying event (UE) and PU have the effect of smearing the energy of jets that are produced in a hard scattering process. Jet energy corrections are applied to remove the effect of this. A similar procedure as with the electron relative isolation (see Section 5.3) is used to correct for the effect of PU and the UE on the jets. Again, this uses the energy density (ρ) for a particular event and the effective area, A_{eff}, which is determined as a function of η within the detector.

Corrections are also applied to account for the non-linearity of the detector as a function of the η and p_{T} of jets. Di-jet data and Drell-Yan simulation

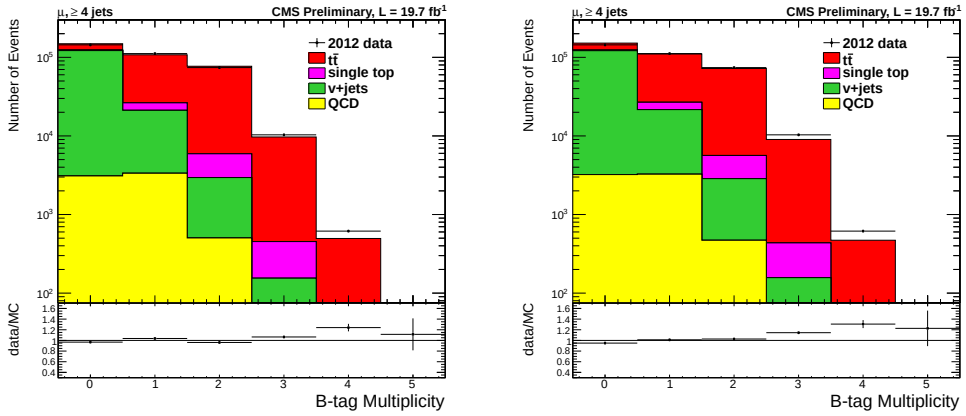

Figure 5.10: Logarithmic distribution of b-jet multiplicity before (left) and after (right) applying b-tag scale factors for the 8 TeV data.

are used to derive these corrections [91]. These corrections are applied not only to the MC but also to the data. The reason for this is that energy leakage from jets is dependent on where in the detector it is measured. In some cases, it may not be possible to fully reconstruct a jet, for example, if it punches through the HCAL.

Jet energy resolution corrections are applied to account for the fact that jet energy resolution is worse in data than in simulation. All PF jets in MC are smeared by the prescribed scale factors that were derived using 0.8 fb^{-1} of 2011 di-jet data [92]. Measurements with 2012 data and more statistics showed consistent results.

5.7.5 Missing Transverse Energy Corrections

The $E_{\text{T}}^{\text{miss}}$ is corrected by propagating the correction from the jet energy and resolution corrections described in the previous section to the $E_{\text{T}}^{\text{miss}}$. These are known as "Type I" $E_{\text{T}}^{\text{miss}}$ corrections. The $E_{\text{T}}^{\text{miss}}$ and $E_{\text{T}}^{\text{miss}}$ ϕ distributions are shown in Figure 5.11. These distributions show the small effect that these corrections have.

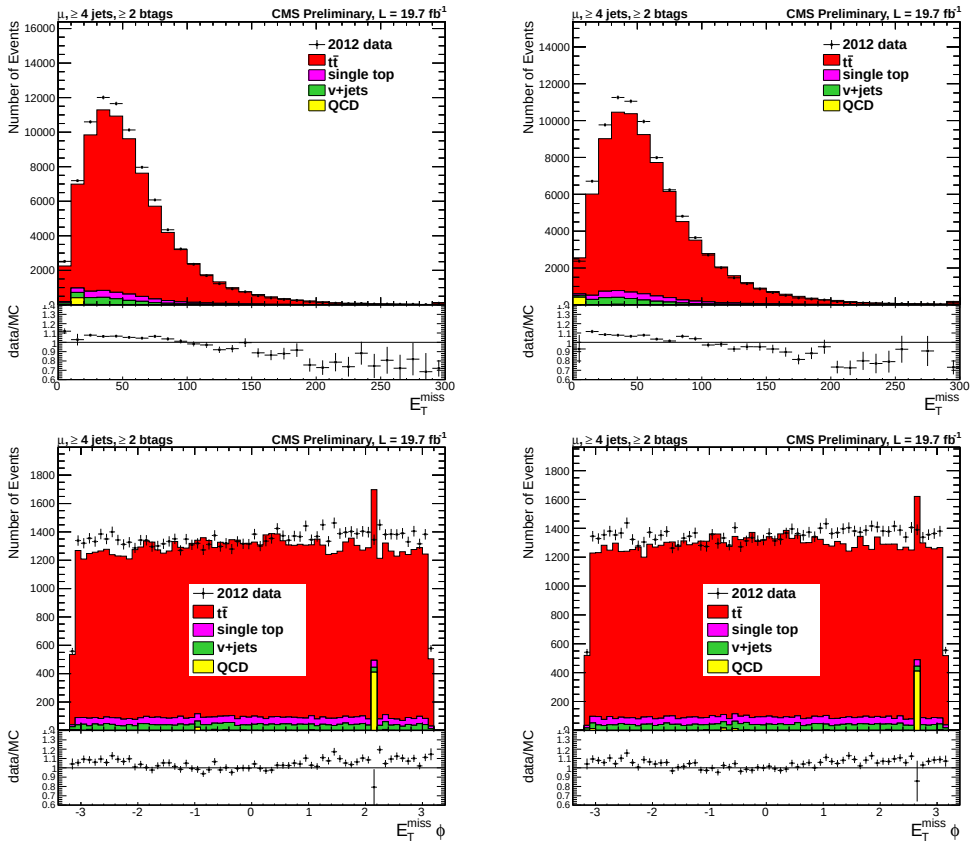

Figure 5.11: Distributions of raw $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}$ ϕ (left) and Type I corrected $E_{\mathrm{T}}^{\mathrm{miss}}$ and $E_{\mathrm{T}}^{\mathrm{miss}}$ ϕ (right) for the 8 TeV data. In the $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions the final bin is as an overflow, containing events beyond the range of the x-axis.

5.7.6 Top p_{T} Reweighting

The differential cross section measurement of the top quark p_{T} showed a discrepancy between the data and the prediction of MADGRAPH and the other MC generators. The data was observed to be softer than the predictions from the various generators [93, 94, 95]. This was first observed for the 7 TeV top p_{T} distribution as can be seen for the semi-lepton channel in Figure 5.12. This mis-modelling in MC can also have a slight effect on other distributions including the $E_{\mathrm{T}}^{\mathrm{miss}}$, H_{T} and S_{T}. The nature of this discrepancy is still not very well understood. Top p_{T} reweighting is applied to correct for it.

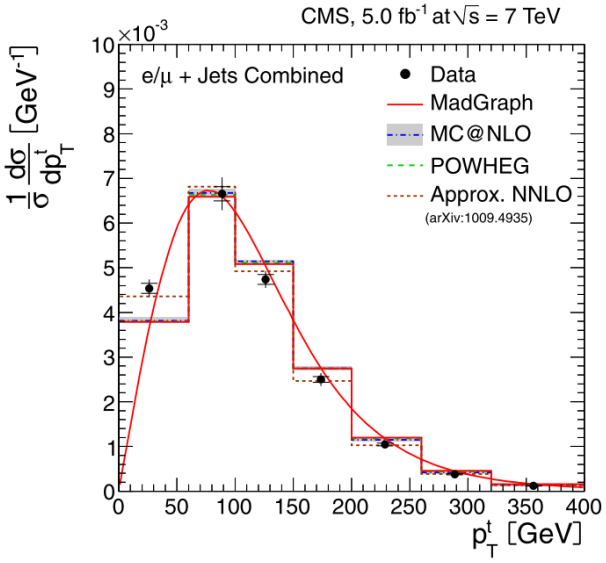

Figure 5.12: The top p_{T} distribution above shows the discrepancy between the data and the MADGRAPH prediction. The agreement is better for the approximate NNLO prediction. Image from [93].

A weight is applied to each event dependent on the generated p_{T} of the top and anti-top. The weight, w, is given by:

$$w = SF(\mathrm{top}) \times \mathrm{SF}(\mathrm{anti}-\mathrm{top}) \tag{5.17}$$

where $SF(\mathrm{top})$ is calculated for both the top and the anti-top quark and is

given as a exponential function of the top p_{T}:

$$SF(\mathrm{top}) = \exp(\mathrm{A} + \mathrm{Bx}) \tag{5.18}$$

where A and B are coeficients determined from a fit to remove the discrepancy [96]. x is the p_{T} of the top/anti-top. This correction is normalised such that the total events in MC is conserved.

5.8 Summary

In this chapter, all of the physics objects which are used to identify top-pair decays have been reviewed. The reconstruction and identification of the various particles that are present in $t\bar{t}$ decays has been described. Details on the corrections applied to account for differences between data and MC have also been given.

Chapter 6

Inclusive $t\bar{t}$ Cross Section Measurement at $\sqrt{s} = 7$ and 8 TeV

Top quark pair decays are capable of producing all particles and so all parts of the detector may be required. To ensure that the CMS detector was behaving as expected, $t\bar{t}$ was therefore one of the first things that had to be studied in detail when the LHC was first turned on. It was important to check that the particles, jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ are well understood such that the inclusive cross section can be measured precisely. This would give confidence going on to more complicated studies involving the top quark such as differential cross sections. This chapter will describe the methods used for measuring the inclusive cross section starting with the definition of the cross section measurement. The reader is reminded that the work done for this thesis was performed in the muon-plus-jets decay channel. The electron-plus-jets channel was worked on in parallel in a collaboration with people at CERN and Cornell University.

The general analysis strategy to extract the $t\bar{t}$ cross section is by means of a maximum likelihood fit of the muon $|\eta|$ distribution. This distribution is chosen since its shape is sensitive to differences between the signal and the background as was explained in Section 2.1. Other analyses have chosen to

use different variables such as $M3$ and m_{lb} [97]. $M3$ and m_{lb} are the invariant masses of the three jets with the highest reconstructed p_{T} and of the selected lepton and a b quark, respectively. The fit is performed after the final event selection described in Section 6.2. The event selection is an important part of the analysis as this helps remove some of the vast backgrounds arrising from other physics processes. Most of these backgrounds can be estimated using data driven techniques as will be described in detail in Section 6.3. The 7 and 8 TeV analyses have been performed using similar methods. There are, however, some differences between the analyses in the event selection and the background estimation. These differences will be outlined in this chapter, where appropriate.

6.1 Data Samples and Triggers

During the technical stop between the 2011 and 2012 LHC runs, the centre of mass energy was increased from 7 to 8 TeV. This provided the opportunity to measure the $t\bar{t}$ cross section at different centre of mass energies. The next section outlines the data samples and triggers used for both the 7 and 8 TeV analyses.

6.1.1 Dataset used in 7 TeV Analysis

The 7 TeV analysis was performed using the first 987.2 pb^{-1} collected by the CMS detector during the 2011 LHC run. The full data set was not used because the uncertainty on the measurement was already becoming systematically dominated. Additional statistics would not have drastically improved the sensitivity of the result. The single muon trigger with a p_{T} requirement of 30 GeV was used. There was no muon isolation requirement for this trigger, which meant that the QCD background could be studied in detail. This trigger was only in operation for run A (April-June of LHC running) of 2011 collecting an integrated luminosity of 1.1 $\mathrm{fb}^{-1}$of data. It was not used after the instantaneous luminosity reached $3{\times}10^{33}$ $\mathrm{cm}^{-2}\mathrm{s}^{-1}$ due to the event rate being too high. The CMS trigger rate had to be kept

under control in the face of rapidly increasing luminosities. The datasets that have been used along with the corresponding run ranges and integrated luminosities are shown in Table 6.1.

Dataset	**Run period**	$\mathcal{L}_{\mathbf{int}}$ / pb^{-1}	**Run Range**
Single Muon ($p_{\mathrm{T}} > 30$ GeV)	Run 2011 A	987.2	160404–167284

Table 6.1: Data samples used for the 7 TeV analysis.

6.1.2 Dataset used in 8 TeV Analysis

In the 8 TeV analysis, the full 2012 dataset was used (Runs A: April, B: May-June, C: July-September and D: October-December). Table 6.2 shows a list of the data samples and their corresponding integrated luminosities and run ranges. The total integrated luminosity equates to 19.7 fb^{-1}.

Dataset	**Run period**	$\mathcal{L}_{\mathbf{int}}$ / pb^{-1}	**Run Range**
Single Muon ($p_{\mathrm{T}} > 24$ GeV, $R_{\mu}^{\mathrm{iso}} < 0.15$)	Run 2012 A	889.4	190456–193621
Single Muon ($p_{\mathrm{T}} > 24$ GeV, $R_{\mu}^{\mathrm{iso}} < 0.15$)	Run 2012 B	4424.0	193834–196531
Single Muon ($p_{\mathrm{T}} > 24$ GeV, $R_{\mu}^{\mathrm{iso}} < 0.15$)	Run 2012 C	7152.0	198022–203742
Single Muon ($p_{\mathrm{T}} > 24$ GeV, $R_{\mu}^{\mathrm{iso}} < 0.15$)	Run 2012 D	7280.0	203777–208686

Table 6.2: Data sets by run period with corresponding integrated luminosities ($\mathcal{L}_{int}$) and run ranges.

The single muon trigger was also used for this analysis but this time with an isolation requirement and a reduced p_{T} requirement (24 GeV). The isolation requirement within the trigger uses detector-based rather than particle flow isolation (see Equation 5.8). Detector-based relative isolation is calculated using:

$$R_{\mu}^{\mathrm{iso}} = \frac{\sum p_T^{\mathrm{tracker}} + \sum E_T^{\mathrm{ECAL}} + \sum p_T^{\mathrm{HCAL}}}{p_T^{\mu}} \tag{6.1}$$

where $\sum p_T^{\mathrm{tracker}}$ is the sum of the p_T of the tracks reconstructed in the tracker, $\sum E_T^{\mathrm{ECAL}}$ and $\sum p_T^{\mathrm{HCAL}}$ are the sum of the energy deposits in the ECAL and

HCAL, respectively, surrounding the muon within a cone of size ΔR in the detector. Muons only pass the trigger if they have $R_{\mu}^{\text{iso}} < 0.15$ within a cone of $\Delta R = 0.3$. This trigger was in operation throughout the whole of the 2012 data taking period.

6.2 Event Selection

The 7 and 8 TeV analyses both use the muon-plus-jets $t\bar{t}$ decay channel and so the event selection is reflected by this. The semi-leptonic event topology was outlined in Section 2.2 with an isolated lepton and four jets expected in the final state. The purpose of the event selection is to maximise the efficiency (see Equation 6.16) of semi-leptonic $t\bar{t}$ events whilst removing as much of the background as possible. The systematic uncertainty associated with the backgrounds can be minimised by selecting a very pure $t\bar{t}$ sample, where the purity of the selection is defined as:

$$\text{purity}(t\bar{t}) \ = \frac{N_{t\bar{t}}^{sel}}{N_{t\bar{t}}^{sel} + N_{\text{BG}}^{sel}} \tag{6.2}$$

where $N_{t\bar{t}}^{sel}$ and N_{BG}^{sel} are the number of $t\bar{t}$ and background events passing a certain selection step.

The general selection requirements are applied in the following order:

1. Trigger and event cleaning. Event cleaning consists of removing events where parts of the detector are not performing as expected. For example, if there is noise in a part of the ECAL or HCAL.

2. The event must have a good primary vertex selection as defined in section 5.1.2.

3. Exactly one high p_{T} (see Table 6.3), isolated muon is then selected. This muon must satisfy the tight selection requirements in Table 5.2.

4. Events with additional looser, lower p_{T} muons are vetoed in order to remove contamination from the Drell-Yan (DY) process.

5. Events with additional low p_{T} electrons are also vetoed in order to remove contamination from di-lepton $t\bar{t}$ events.

6. Each event is further required to have at least three (7 TeV analysis) or four (8 TeV analysis) high p_{T} jets. These jets are reconstructed and identified as described in section 5.5. A more stringent selection was used in the 8 TeV analysis to reduce the systematic uncertainties related to the backgrounds.

7. The tagging of b quarks (see section 5.5.3) is implemented for the 8 TeV analysis but not in the 7 TeV analysis. The looser b-tagging and jet requirements in the 7 TeV analysis are due to the lower statistics data sample. A high signal purity is obtained in the 8 TeV analysis by adding this b-tagging requirement.

The main selection requirements are highlighted in the Table 6.3. The key differences between the 7 and 8 TeV analyses are presented. It was possible to reduce the muon p_{T} selection requirement because the trigger p_{T} requirement was reduced. The muon pseudorapidity range is extended from 2.1 to 2.5 for loose muons inorder to remove more dileptonic backgrounds. Tight muons are only triggered within $|\eta| < 2.1$. The Events that pass the selection described above are weighted according to the correction factors described in Section 5.7 and using the predicted cross sections in Table 6.4.

6.2.1 Event Yields and Key Kinematic Distributions

It is important to understand the effect that each selection requirement has on the number of signal and background events. Checks are done by looking at the number of events surviving each requirement and by studying the key kinematic distributions in the selected events. Comparisons have been made with other analysis groups to clarify that the selection is being performed correctly.

The contribution from the various physics processes after each selection step for the 8 TeV analysis can be seen in Figure 6.1. In the 8 TeV analysis there was an initial muon-plus-three-jet pre-selection applied. The purpose

Selection parameter	$\sqrt{s} = 7$ TeV	$\sqrt{s} = 8$ TeV
Trigger		
Trigger	Single Muon	Isolated Muon
p_{T}	30 GeV	24 GeV
Reliso	none	< 0.15 (PF $\Delta R = 0.4$)
Tight Selected Muon		
p_{T}	> 35 GeV	> 26 GeV
$\lvert\eta\rvert$	< 2.1	< 2.1
Reliso, $\Delta R = 0.4$	< 0.125 (PF)	< 0.12 (PF $\Delta\beta$ corr)
Loose Vetoed Muon		
p_{T}	> 10 GeV	> 10 GeV
$\lvert\eta\rvert$	< 2.5	< 2.5
Reliso, $\Delta R = 0.4$	< 0.2 (PF)	< 0.2 (PF $\Delta\beta$ corr)
Loose Vetoed Electron		
p_{T}	> 15 GeV	> 20 GeV
$\lvert\eta\rvert$	< 2.5	< 2.5
Reliso, $\Delta R = 0.5$	< 0.2 (PF)	< 0.15 (PF rho corr)
Selected Jets		
multiplicity	≥ 3	≥ 4
p_{T}	> 30 GeV	> 30 GeV
$\lvert\eta\rvert$	< 2.5	< 2.5
B-tagging		
multiplicity	≥ 0	≥ 2

Table 6.3: Selection requirments applied in the 7 and 8 TeV analyses.

Process	7 TeV (pb)	8 TeV (pb)
$t\bar{t}$	163	234
Single-top	86	115
$W \to l\nu$	31314	36257
$Z/\gamma^* \to l^+l^-$ ($m_{ll} > 50$ GeV)	3048	3504
QCD multi-jet (μ enriched)	≈ 84679	≈ 97381

Table 6.4: Cross sections (pb) of the major contributing physics processes at 7 and 8 TeV. The $t\bar{t}$ and single-top cross sections are provided by approximate NNLO calculations [29]. The W and Z boson production cross sections are calculated at NNLO using FEWZ [98]. Pythia [99] was used to estimate the QCD multi-jet LO cross section. The QCD sample only contains multi-jet events containing muons (muon enriched). This is because simulating QCD events is computationally taxing and it would take too long to simulate the inclusive sample.

of this pre-selection was to reduce the size of the files and hence the computer processing time of the next step of the analysis. As can be seen from Figure 6.1 it is clear that QCD dominates the event composition prior to any selection. After all the selection requirements have been applied the data becomes $t\bar{t}$ dominated. At least two b-tags are required for the final selection step such that systematic uncertainties relating to backgrounds are reduced. There is also a lot more MC than data in the first bin of the distribution because the trigger has been used to select the required data but is not applied to the MC until the second step "cleaning and HLT". The isolated muon requirement removes a lot of the QCD since multi-jet events rarely produce isolated muons. Quite a lot of V+jets (V stands for vector boson and is W and Z production) events are still present though as these are capable of producing genuinely isolated muons. The contribution from V+jets events slowly drops off with the requirement of additional jets and b-tagging.

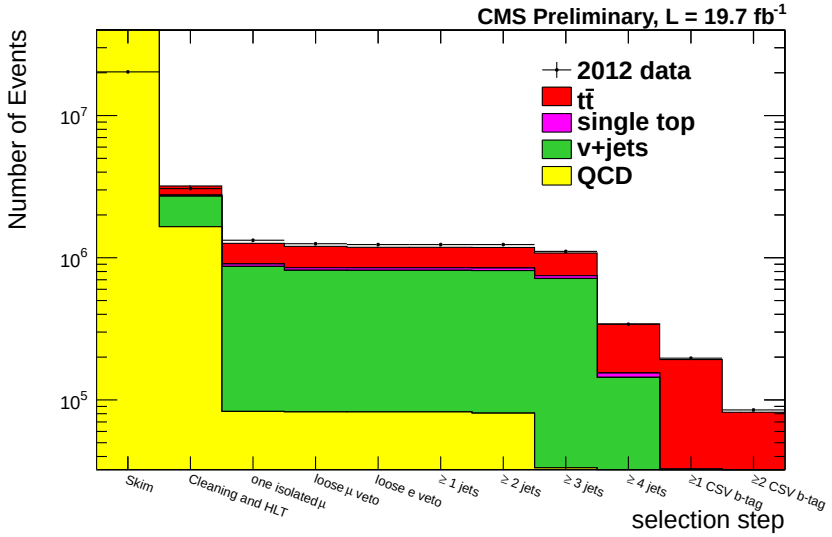

Figure 6.1: The number of events surviving each stage of the selection.

The event yields after each selection step for the 7 and 8 TeV analysis are presented in Tables 6.5 and 6.6, respectively. These tables show the number of events surviving each successive requirement. The number in the brackets shows the efficiency as a percentage for an individual requirement, i, with

respect to the previous requirement $(i-1)$:

$$\epsilon_{\text{req.}} = 100 \times (N^{\text{i}}_{\text{sel}}/N^{\text{i}-1}_{\text{sel}})\% \tag{6.3}$$

Selection step	$t\bar{t}$	W+Jets	Z+Jets	QCD	Single-t	Total MC	Data
Initial	160913	30913180	3008985	83595108	84925	1.17763×10^8	46976284
Cleaning and HLT	24496	3825206	601544	3985165	6991	8.4434e+06 (7.17%)	14817498 (31.5%)
one isolated μ	14910	2209137	263201	17812	4476	2.50954e+06 (29.7%)	2558884 (17.3%)
loose μ veto	14240	2209105	110793	17706	4401	2.35624e+06 (93.9%)	2394181 (93.6%)
loose e veto	12794	2208666	110186	17529	4246	2.35342e+06 (99.9%)	2390848 (99.9%)
$\geq$ 3 jets	9620	17889	2221	317	1047	31094 (1.32%)	31955 (1.33%)

Table 6.5: Numbers of $t\bar{t}$, single-top, W and Z+Jets and QCD events surviving after each selection step in the 7 TeV analysis. MC is scaled to a luminosity of 987.2 pb^{-1} using the 7 TeV theoretical cross sections. The number in brackets in the last two columns is the percentage of events surviving from the previous selection requirement.

Selection step	$t\bar{t}$	W+jets	Z+jets	QCD	Single-t	Total MC	Data
Initial	4609800	714262900	69028800	1.9184×10^9	2265500	3.4228×10^9	20284215
Cleaning and HLT	440179.2	960480.1	220771.8	1653207.6	44436.7	3196233.7 (0.1%)	3063569.0 (15.1%)
one isolated μ	358553.7	706122.6	131565.5	82992.3	35315.5	1262892.4 (39.5%)	1327738.0 (43.3%)
loose μ veto	351350.3	706050.6	79376.2	82324.6	35031.4	1202482.7 (95.2%)	1254896.0 (94.5%)
loose e veto	334289.5	705833.2	78887.9	82302.9	34408.8	1184094.7 (98.5%)	1237495.0 (98.6%)
$\geq$ 3 jets	332300.6	633859.1	70777.6	33287.1	32931.1	1079969.3 (91.4%)	1108272.0 (89.6%)
$\geq$ 4 jets	187768.9	124062.4	15118.2	6960.8	10487.7	342347.2 (31.7%)	340786.0 (30.7%)
$\geq$1 CSV b-tag	159651.0	18391.8	2519.3	3739.2	8268.4	192291.6 (56.2%)	196667.0 (57.7%)
$\geq$2 CSV b-tag	75827.0	2184.4	370.4	478.7	3076.8	81900.9 (42.6%)	85028.0 (43.2%)

Table 6.6: Expected and observed event yields after each event selection step for the 8 TeV analysis. The MC event yields are scaled to a luminosity of 19.7 fb^{-1} using the 8 TeV theoretical cross sections. The number in brackets in the last two columns is the percentage of events surviving from the previous selection requirement.

Figures 6.2 and 6.3 show a comparison between data and MC for some of the key kinematic variables after the final selection for the 7 and 8 TeV data, respectively. The muon p_{T}, η and relative isolation are shown along with the jet multiplicity, p_{T} and η. The individual MC samples are all stacked on top of one another with QCD at the bottom and $t\bar{t}$ at the top. It is possible to use a finer binning for the 8 TeV data due to the increased statistics. As shown in Tables 6.5 and 6.6, there is a slight excess of data with respect to the MC expectations for both 7 and 8 TeV. The excess seems to be more

pronounced at lower energies. It could be a similar affect as seen in the top p_{T} distribution. The normalisation factors for the various processes can be determined in the MLE fit which resolves any excess.

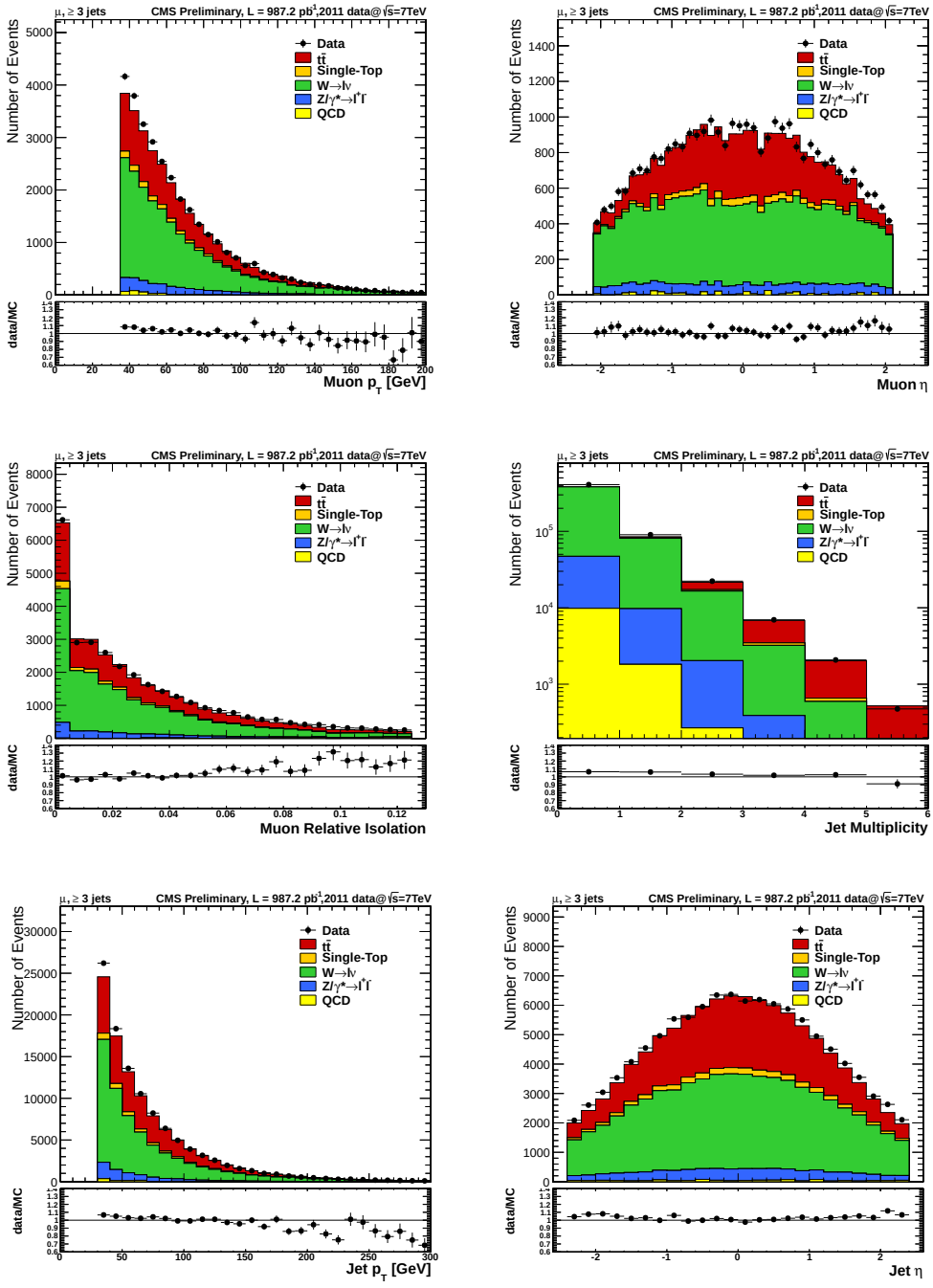

Figure 6.2: Data-MC comparison for some key kinematic distributions at 7 TeV for events passing the full selection. MC distributions are normalized to the number of events expected from MC for the amount of data analysed. The distributions underneath show the data/MC ratio.

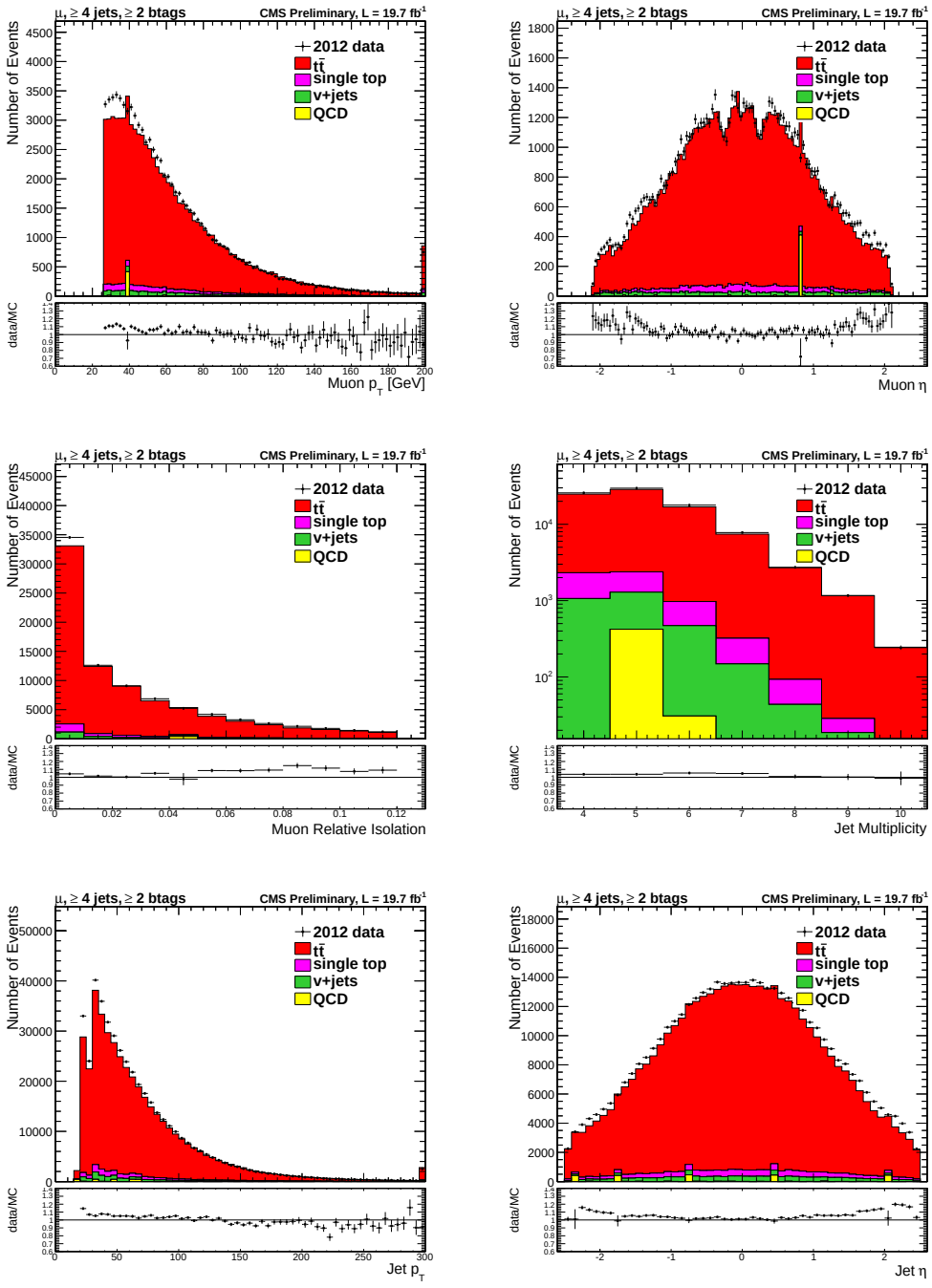

Figure 6.3: Kinematic distributions of some key variables comparing data and MC at 8 TeV. Events passing the full selection are used with MC normalised to the theoretically predicted cross section and a luminosity of 19.7 fb^{-1}. In the muon and jet p_{T} distributions the final bin is as an overflow, containing events beyond the range of the x-axis. The distributions underneath show the data/MC ratio.

6.3 Background Estimation Methods

As was introduced earlier in this chapter, process dependent normalised templates are used in the MLE fit. These templates can be either input using MC or they can be estimated using the data. Processes that are more difficult to model such as QCD multijet background can be better estimated using the data.

The increase in integrated luminosity between the 7 and 8 TeV datasets meant a more stringent selection could be applied to the 8 TeV data. Different techniques were therefore implemented in order to estimate the various background templates. Stricter selection requirements remove most of the background before the fit is even performed in the 8 TeV analysis. In the 7 TeV analysis, however, a lower statistics sample is used and so looser selection requirements meant that there was a higher contamination from reducible background processes such as W+Jets, Z+Jets and QCD multijet events. It is good practice to use data to estimate the backgrounds to check that the MC provides a good description. This allows MC to be used in the 8 TeV analysis. It was therefore necessary to develop a better understanding of the backgrounds, particularly for the muon pseudorapidity shapes that enter the MLE fit.

In following subsections, the template shapes for the three main reducible background processes are studied in detail using data driven techniques. These techniques usually involve inverting or changing part of the event selection criteria such that the background enriched data is selected. Corrections are then used to correct back to the signal region where the MLE is performed.

6.3.1 W+Jets

The template to be used in the MLE for the W background is constructed using the fact that the ratio of the W production cross sections

$$\frac{\sigma(\mathrm{pp} \to W^+)}{\sigma(\mathrm{pp} \to W^-)} \tag{6.4}$$

is larger than unity, i.e. the W background process produces more μ^+ than μ^-. This can be explained by referring back to Figure 2.3. At the LHC, W bosons can be produced with at low x of about 0.01 per parton where the sea quarks and gluons dominate. In W production with additional jets, however, higher x is required and this is where the contribution from valence quarks start to become important. The more energy (jets) that there is in the final state the more important the valence quarks become. As there are two up and one down quark within a proton this yields slightly more W^+ production than W^-.

As an example, the W+2 jet production process is shown in Figure 6.4 where an up-type quark interacts with a gluon. Other processes, excluding t-

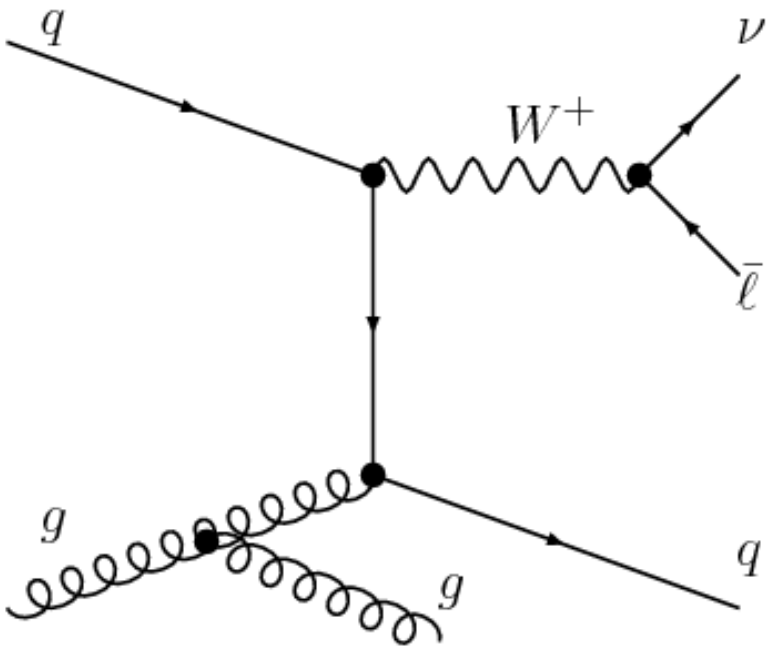

Figure 6.4: Feynman diagram showing the W+Jets production process with two additional jets.

and s-channel single-top production are expected to produce leptons of both charge at more or less the same rate.

Subtracting the $|\eta_{\mu^-}|$ distribution from the $|\eta_{\mu^+}|$ distribution produces a W enriched data sample. The charge subtracted distribution, Δ^i_μ, is defined

as:

$$\Delta_{\mu}^{i} = \mathrm{N}(|\eta_{\mu^+}|)^i - \mathrm{N}(|\eta_{\mu^-}|)^i \tag{6.5}$$

where $\mathrm{N}(|\eta_{\mu^+}|)^i$ and $\mathrm{N}(|\eta_{\mu^-}|)^i$ denote the number of events in an $|\eta|$ bin, i, for positively and negatively charged muons, respectively. The charge sum distribution, Σ_{μ}^{i}, corresponding to the desired W template shape is given by:

$$\Sigma_{\mu}^{i} = \mathrm{N}(|\eta_{\mu^+}|)^i + \mathrm{N}(|\eta_{\mu^-}|)^i \tag{6.6}$$

Obtaining Σ_{μ}^{i} from Δ_{μ}^{i} requires corrections to be applied to account for the differences in the shapes of the W^+ and the W^- muon pseudorapidity spectra. This correction factor, c_W^i, for each bin can be defined such that Σ_{μ}^{i} is obtained when c_W^i is multiplied with Δ_{μ}^{i}:

$$\Sigma_{\mu}^{i} = c_W^i \times \Delta_{\mu}^{i}. \tag{6.7}$$

where c_W^i can be derived such that:

$$c_W^i = \left(1 + \frac{2}{R_i - 1}\right) \tag{6.8}$$

The ratio factor, R_i, is defined as the ratio of the expected W^+ and W^- rates in bin i:

$$R_i = \mathrm{N}(|\eta_{\mu^+}|)^i / \mathrm{N}(|\eta_{\mu^-}|)^i \tag{6.9}$$

R_i can be expressed using normalized differential distributions, $d\tilde{\sigma}/d\eta_{\mu}$, the overall cross section ratio, $R = \sigma_{W^+}/\sigma_{W^-}$, and the ratio of the efficiencies, $\rho_{\varepsilon} = \varepsilon_{W^+}/\varepsilon_{W^-}$ such that

$$R_i = R\frac{\varepsilon_{W^+}}{\varepsilon_{W^-}} \left(\frac{d\tilde{\sigma}_{W^+}}{d|\eta_{\mu^+}|}\right)_i \left(\frac{d\tilde{\sigma}_{W^-}}{d|\eta_{\mu^-}|}\right)_i^{-1} = R\,\rho_{\varepsilon}\,\rho_i, \tag{6.10}$$

where the ratio of the normalized differential cross sections is defined as:

$$\rho_i = \left(\frac{d\tilde{\sigma}_{W^+}}{d|\eta_{\mu^+}|}\right)_i \left(\frac{d\tilde{\sigma}_{W^-}}{d|\eta_{\mu^-}|}\right)_i^{-1} \tag{6.11}$$

Equation 6.8 then becomes:

$$c_W^i = \left(1 + \frac{2}{R\rho_\varepsilon\rho_i - 1}\right) \tag{6.12}$$

R changes as a function of jet multiplicity. The objective is to extract the template shape for ≥ 3 jets, since this is where the signal region is defined. The charge subtraction is, however, done using 1 or 2 jets to enhance statistics. A χ^2 test was performed on a closure test using the R values measured with ≥ 1 and ≥ 3 jets. A lower χ^2 was observed when using the R value from the ≥ 3 jets measurement. This value of $R = 1.6 \pm 0.3$ as measured by CMS [100].

To estimate ρ_i and ρ_ε values, the parton-level NLO Monte Carlo program MCFM [101] was used. MCFM has no hadronising/showering effects but these effects are expected to cancel in the ratio calculation. Events are simulated using the same muon p_T and η requirements as in the 7 TeV event selection described in section 6.2.

The distributions of ρ_i and the resulting corrections factors, c_W^i, with uncertainties due to the error propagated from the R value are shown in Figure 6.5. The correction factors are relatively flat using an R value of 1.6.

The distributions of positive and negative muons for jet-multiplicities of $N_{\text{jet}} = 1||2$ (one or two jets) and $N_{\text{jet}} \geq 3$ are shown in Figures 6.6 and 6.7 respectively. There is a slight excess in data events with respect to MC. This could be a result of poor statistics in the QCD multijet MC sample. The corresponding charge-subtracted distributions, Δ_μ^i, for the same jet-multiplicities are shown in Figure 6.8.

The W template is finally obtained by taking the product of the charge-difference (Δ_μ) and the correction factor (c_W) distributions. Ideally it would be preferable to use the actual signal region, i.e. the $N_{\text{jet}} \geq 3$ region to extract the template shape. The statistics are, however, to limited to do that. There is also more contamination from other processes (mostly single-top) in the $N_{\text{jet}} \geq 3$ multiplicity region when using this region. The contamination is very small for $N_{\text{jet}} = 1||2$ and so no contamination subtraction is required.

The first test of the procedure is to perform it using MC rather than the

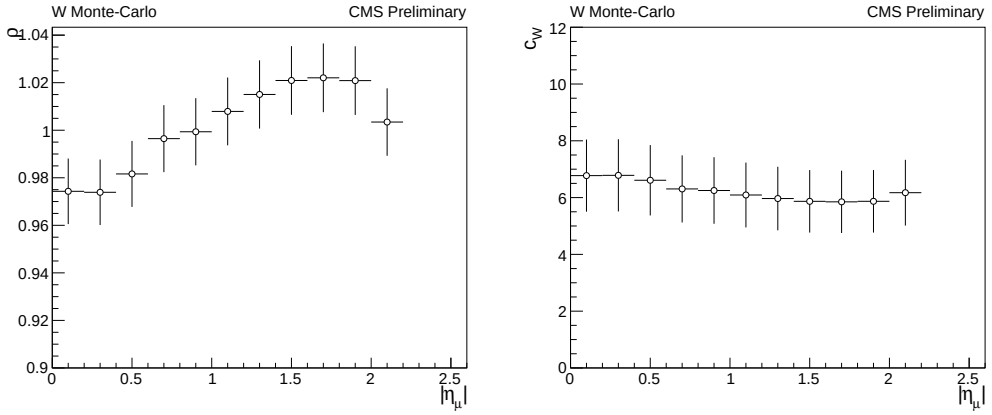

Figure 6.5: Ratio of normalized muon pseudorapidity distributions for W^+ and W^- (left) and muon pseudorapidity dependent correction factors (right) calculated using an R value of 1.6. These correction factors are almost flat within uncertainties. The uncertainties for the ρ values are set to 5% to include any miss-modelling effects in the MC used. The uncertainties on c_W also include the uncertainty on the measured value of R.

data. In this closure test, the correction factors, c_W, are applied to the W MC prediction in the $N_{\text{jet}} = 1||2$ charge-subtracted region. The results are then compared to the actual MC prediction in the signal region. Figure 6.9 shows the comparison of the "W MC template" in the $N_{\text{jet}} = 1||2$ muon charge-subtracted region and the actual MC prediction in the $N_{\text{jet}} = 1||2$ signal region on the left and in the $N_{\text{jet}} \geq 3$ signal region on the right. As expected, the W MC template describes the W+Jets MADGRAPH MC well as shown by the ratio plots. The uncertainties on the templates are calculated by propagating the error on c_W in addition to the statistical error.

The final W+Jets template, θ_W, can now be constructed using the charge subtracted data. Figure 6.10 shows the comparison between the W template extracted from data in the $N_{\text{jet}} = 1||2$ charge-subtracted region and the MADGRAPH predictions for the W process in the $N_{\text{jet}} = 1||2$ (left) and the $N_{\text{jet}} \geq 3$ (right) regions. The uncertainties associated with the normalised W template range from 5% at low $|\eta|$ to around 10% for the last $|\eta|$ bin where there are less statistics. The agreement between data and MC is good.

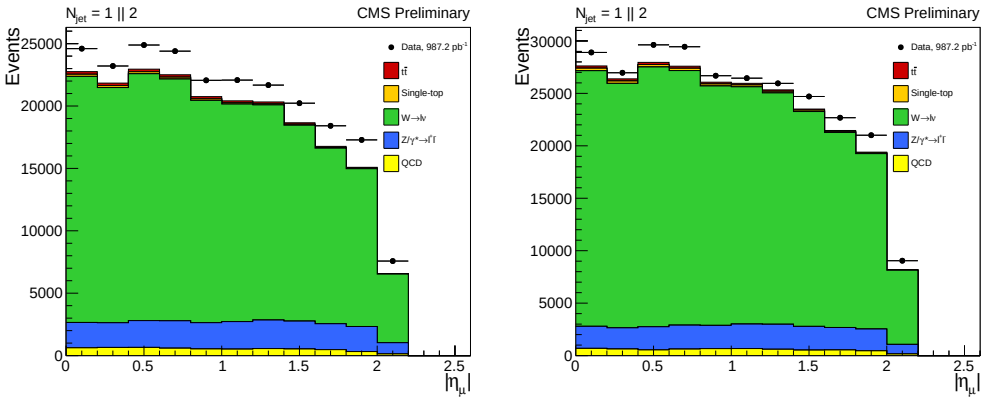

Figure 6.6: The absolute value of the pseudorapidity for negatively (left) and positively (right) charged muons in the $N_{\text{jet}} = 1||2$ jet-multiplicity bin.

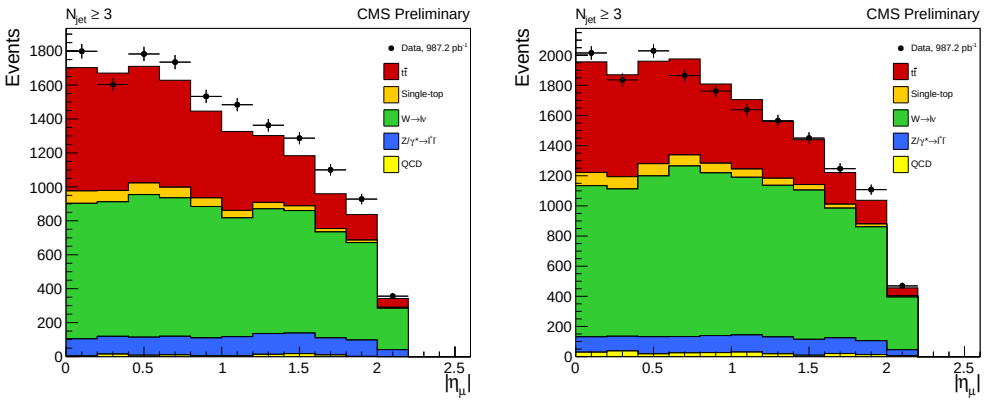

Figure 6.7: Absolute value of the pseudorapidity for negatively (left) and positively (right) charged muons in the $N_{\text{jet}} \geq 3$ jet-multiplicity bin.

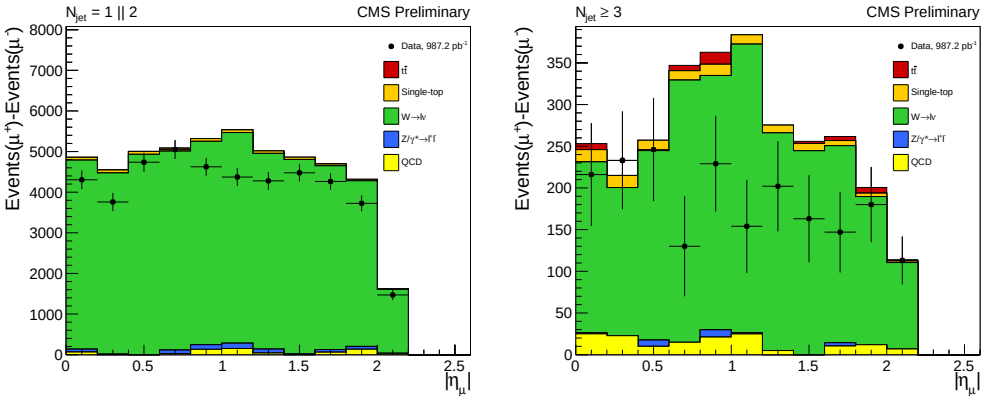

Figure 6.8: Charge-subtracted muon pseudo rapidity distributions for $N_{\text{jet}} = 1||2$ (left) and $N_{\text{jet}} \geq 3$ (right).

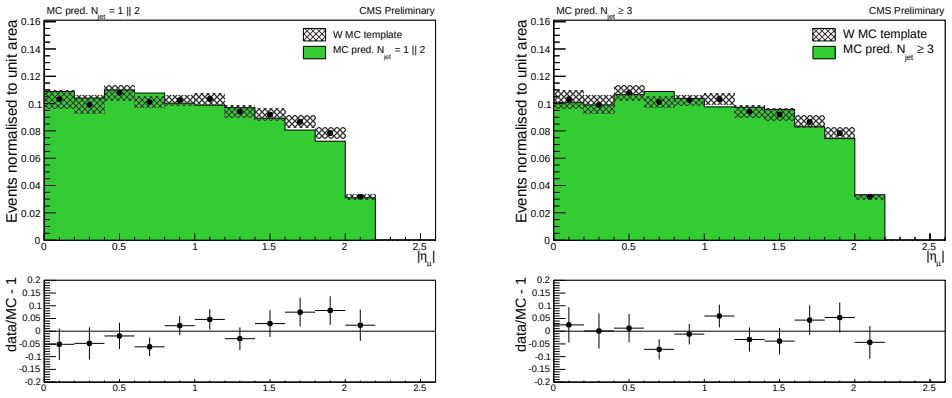

Figure 6.9: The results of the closure test where the W background template is extracted in the $N_{\text{jet}} = 1||2$ region using MC only. The template is compared to the MADGRAPH predictions in the $N_{\text{jet}} = 1||2$ (left) and $N_{\text{jet}} \geq 3$ (right) multiplicity bins.

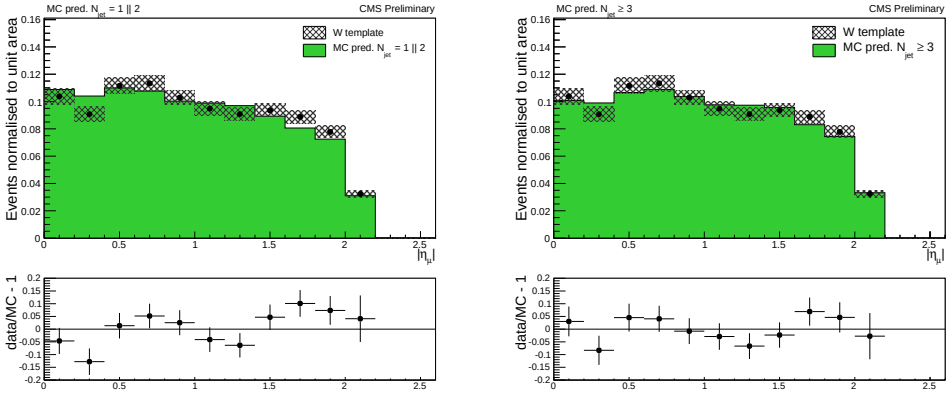

Figure 6.10: The W background template estimated from data compared to MADGRAPH predictions in the $N_{\text{jet}} = 1||2$ (left) and $N_{\text{jet}} \geq 3$ (right) jet multiplicity bins.

6.3.2 Drell-Yan

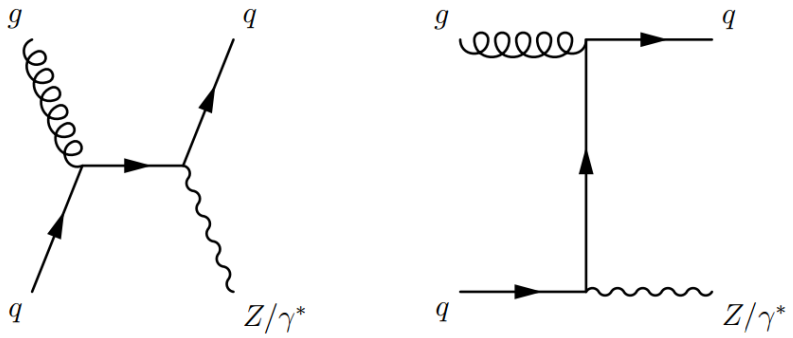

Figure 6.11: Leading order Feynman diagrams showing the Z+Jets production process.

The Drell-Yan (DY) process as shown in Figure 6.11 can pass the signal selection when one of the muons coming from the Z^0/γ^* decay is not detected. This can happen when the second muon is

- outside the muon selection pseudorapidity region $|\eta_\mu| > 2.1$;
- softer than the loose muon definition $p_{\text{T}} < 10$ GeV;

- fails the loose muon isolation requirement, i.e $R_{\text{iso}} > 0.2$;
- when one of the muons is mis-reconstructed as another particle such as a $\pi^{\pm}$.

As with the W+Jets template estimation, a region needs to be defined where the DY processes is enhanced. Events with two opposite sign muons passing the tight muon selection in Table 6.3 and an invariant mass in the Z^0-window defined as $|m_{\mu\mu} - m_Z^0| < 15$ GeV (DY background region) are therefore selected. The pseudorapidity distribution (Figure 6.12) of the harder of the two muons is then studied. This selection contains a small amount of contamination from other processes. The relative contamination from other processes in each bin is shown in Table 6.7 and originates almost entirely from $t\bar{t}$ events. Since the contamination in each bin is well below 1% it can be ignored.

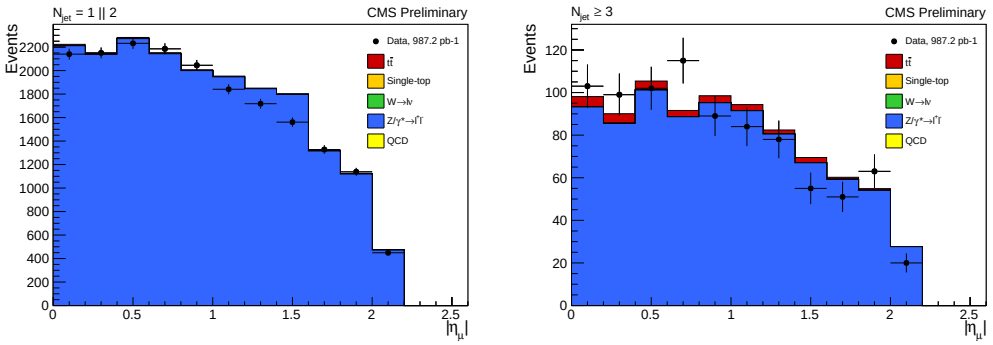

Figure 6.12: Absolute value of the muon pseudorapidity for $N_{\text{jet}} = 1||2$ (left) and $N_{\text{jet}} \geq 3$ (right) of the harder muon in the Drell-Yan selection.

$\lvert\eta_\mu\rvert$	0.0-0.2	0.2-0.4	0.4-0.6	0.6-0.8	0.8-1.0	1.0-1.2	1.2-1.4	1.4-1.6	1.6-1.8	1.8-2.0	2.0-2.2	≥ 2.2
cont.[%]	0.56	0.54	0.48	0.45	0.4	0.36	0.28	0.29	0.26	0.23	0.22	0

Table 6.7: Relative contamination of the Drell-Yan template region with other processes for jet multiplicities of one or two.

Bin-by-bin correction factors are then applied to extrapolate back to the signal region from the DY control region. These correction factors are ob-

tained using particle level DY Monte-Carlo with the MC@NLO [102] event generator. These $|\eta_\mu|$ dependent correction factors are calculated using:

$$c_F = \frac{N^{\text{Events}}_{\text{1 tight muon}}}{N^{\text{Events}}_{\text{2 tight muons}}} \tag{6.13}$$

which is the ratio of the number of events for a Z boson to produce one or two tight muons within the measurable η range ($|\eta| < 2.1$) for all jet multiplicities. The uncertainties on the correction factors resulting from limited MC statistics are small. A pessimistic uniform 5 % uncertainty is applied to account for any MC mis-modelling. The distribution of these correction factors is shown in Figure 6.13. The c_F values are reasonably large due to the requirement imposed by the single muon trigger.

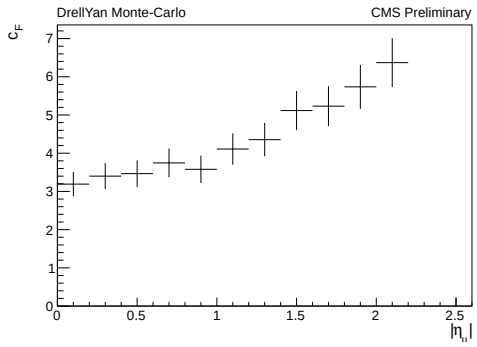

Figure 6.13: Correction factors to extrapolate from the DY background region to the signal region using MC@NLO.

The final Drell-Yan template is obtained by applying the correction factors to the muon pseudorapidity distribution in the $N_{\text{jet}} = 1||2$ DY-enriched region. To test this procedure a closure test is again performed using the MADGRAPH DY+Jets MC. The resulting template (DY MC template) is compared with the MC predictions for the Drell-Yan background in the signal region as shown in Figure 6.14. Good agreement is observed for both $N_{\text{jet}} = 1||2$ and $N_{\text{jet}} \geq 3$.

A comparison of the final DY template extracted from data is made to the MC predictions in both jet multiplicities and is shown in Figure 6.15. The

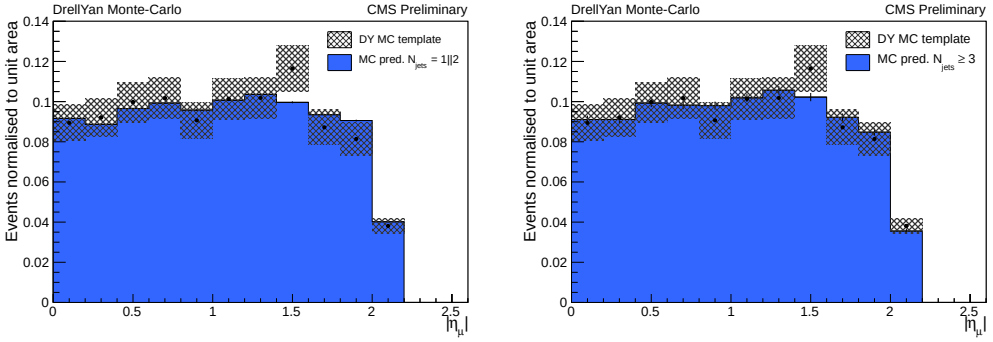

Figure 6.14: Comparison of the corrected MC prediction from the DY background region (DY MC template) to the MADGRAPH prediction in the signal region for $N_{\text{jet}} = 1||2$ (left) and $N_{\text{jet}} \geq 3$ (right).

good agreement indicates an accurate description of the Drell-Yan data by the MC. The uncertainties are the combination of the statistical uncertainty from the DY background region distribution and the modelling uncertainties on the correction factors. The total uncertainties are typically about 10% of the normalised values in each bin.

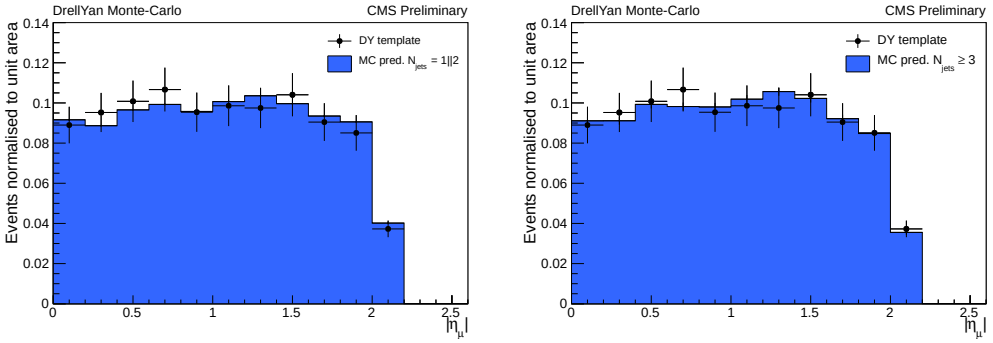

Figure 6.15: Comparison of the DY template obtained from data with the DY MADGRAPH predictions in the signal region for $N_{\text{jet}} = 1||2$ (left) and $N_{\text{jet}} \geq 3$ (right).

6.3.3 QCD Multi-Jet

The QCD multi-jet background can mimic the signal process when muons are produced in the decays of hadrons. The majority of the time, the muons are produced within jets. Occasionally, muons are produced that are isolated from jet activity. This could happen in charged pion decays or in the decay of a b quark to a virtual W boson. This background is difficult to model and so a data-driven method is required for both the 7 and 8 TeV analyses. Poor statistics in the QCD MC sample in the selected signal region means that the description suffers from wild statistical fluctuations producing an unrealistic representation.

When constructing the shape of the QCD template an inversion of the relative isolation, R_{μ}^{iso}, criteria on the selected muon is performed. Instead of requiring $R_{\mu}^{\text{iso}} < 0.12(5)$, a requirement of $R_{\mu}^{\text{iso}} > 0.2$ is used instead. Relaxed jet requirements are also used: ≥ 2 jets for the 7 TeV analysis and ≥ 3 jets for the 8 TeV analysis with an additional 0 b-tag requirement. These requirements are chosen such that the QCD background dominates as can be seen in Figure 6.16. This figure shows the muon relative isolation distribution with all other selection requirements applied. Muons in QCD events are expected to be non-isolated since they are produced close to the jet activity. After applying the inverted R_{μ}^{iso} selection requirement, the muon pseudorapidity is as shown in Figure 6.17. There are clearly a lot more events in the QCD enriched region in the 7 TeV analysis. This is because there was no isolation requirement in the trigger for the 7 TeV analysis.

The contamination from other processes is small in the 7 TeV analysis ($1 - 4\,\%$ per $|\eta_{\mu}|$ bin) and was therefore neglected. There was a much larger contamination ($6 - 10\,\%$ per $|\eta_{\mu}|$ bin) in the 8 TeV analysis. This was a result of the isolation requirement in the trigger and the use of ≥ 3 jets rather than ≥ 2 jets. The contamination was therefore subtracted using the MC prediction. A conservative 50% uncertainty on this subtracted value in each bin was assumed when producing the template.

When comparing the QCD Monte Carlo prediction for the $|\eta_{\mu}|$ distribution in the signal region and QCD enriched region, a clear difference in shapes

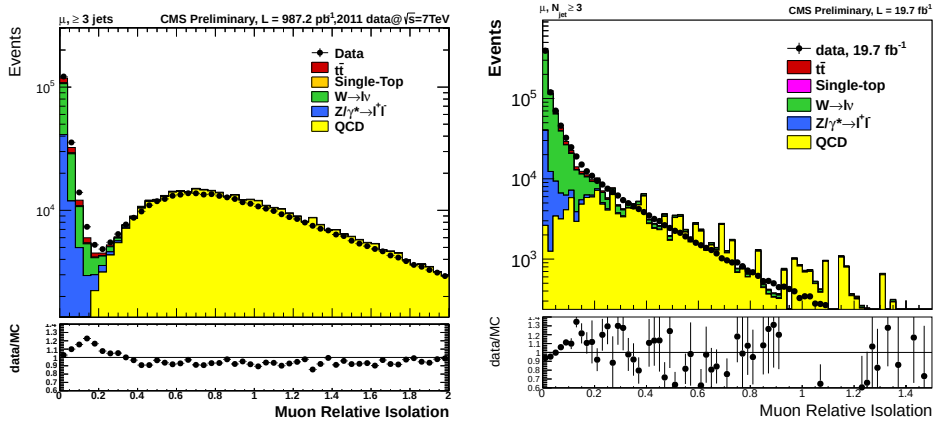

Figure 6.16: Plot showing R_{μ}^{iso} with no isolation selection applied for the 7 TeV (left) and 8 TeV analyses (right). This enables the choice of 0.2 for the inverted R_{μ}^{iso} selection requirement in order to obtain a QCD enriched data sample.

was observed in the 7 TeV data. The $|\eta_{\mu}|$ distribution in the signal region is a lot flatter than in the QCD control region. It was therefore necessary to introduce a correction factor derived from QCD MC to extrapolate back to the signal region. The correction factors are given by

$$C_f = \frac{N^{\text{Events}}_{\text{relIso}<0.12}}{N^{\text{Events}}_{\text{relIso}>0.2}} \tag{6.14}$$

where $N^{\text{Events}}_{\text{relIso}<0.12}$ is the number of events passing the baseline event selection described in section 6.2 with ≥ 2 jets, and $N^{\text{Events}}_{\text{relIso}} > 0.2$ is the number of events passing the same selection but with $R_{\mu}^{\text{iso}} > 0.2$. The correction factors have been normalised such that $\sum C_f = 1$ and are shown in Figure 6.18. This effect was not seen in the 8 TeV analysis and so no correction factors were applied in this case.

After applying the correction factors obtained for $N_{\text{jet}} \geq 2$(3 for 8 TeV) to the QCD enriched data, the QCD data driven template is obtained. A closure test was first performed where the correction factors were applied to MC from the QCD enriched region (Figure 6.19). The full data-driven QCD template is seen in Figure 6.20. These is a slight discrepacny between MC and the

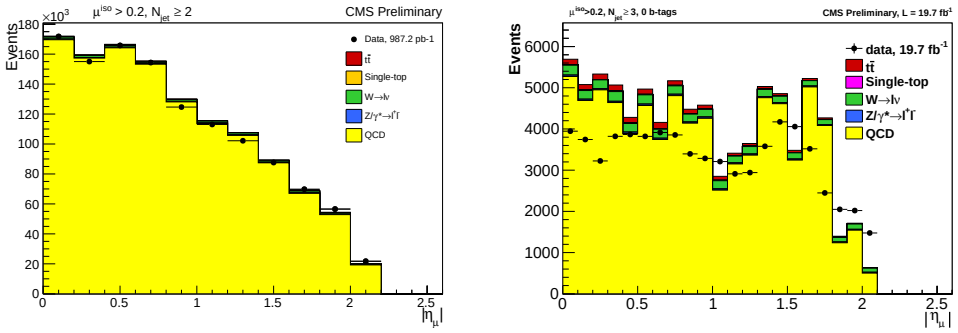

Figure 6.17: Absolute value of the pseudorapidity for $N_{\text{jet}} \geq 2$ for the 7 TeV analysis (left) and $N_{\text{jet}} \geq 3$ (0 b-tag) for the 8 TeV analysis (right) after the inversion of the relative isolation selection requirement.

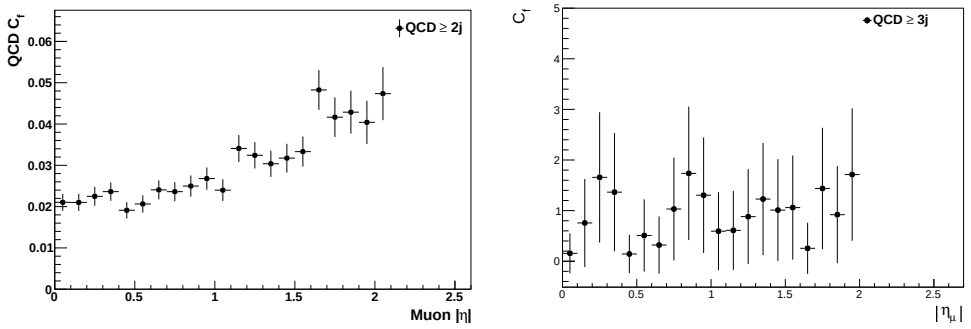

Figure 6.18: This plot shows the normalised correction factors which are applied to the QCD $|\eta_\mu|$ shape extracted from the control region in the 7 TeV analysis (left). These have been normalised to unit area. On the right the same correction factors for the 8 TeV analysis are shown. These are normalised to the number of bins.

data. This is a result of the poor MC statistics and is justification for using a data-driven approach. The uncertainties from the combined statistical limitations of both the data and the MC samples are around 8% and slightly higher for the last few bins where the number of events is low.

6.4 Inclusive Cross Section Definition

The importance of accurately measuring the $t\bar{t}$ production cross section was outlined in Chapter 2. A general formula for the calculation of the $t\bar{t}$ cross

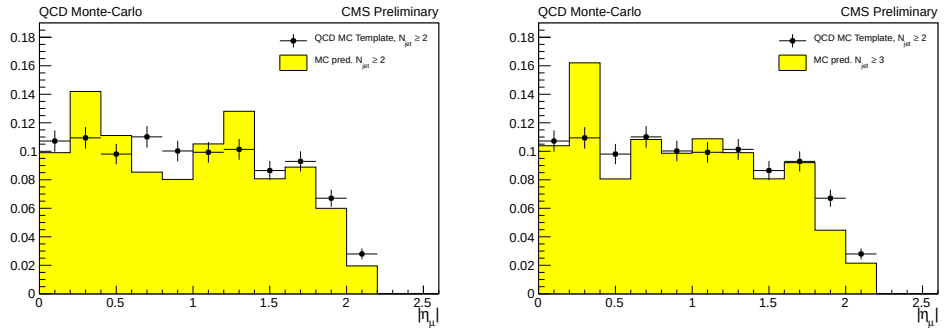

Figure 6.19: Closure test performed on the QCD MC for $N_{\text{jet}} \geq 2$ (left) and $N_{\text{jet}} \geq 3$ (right). The resulting MC template has been normalised. These closure tests were performed for the 7 TeV analysis.

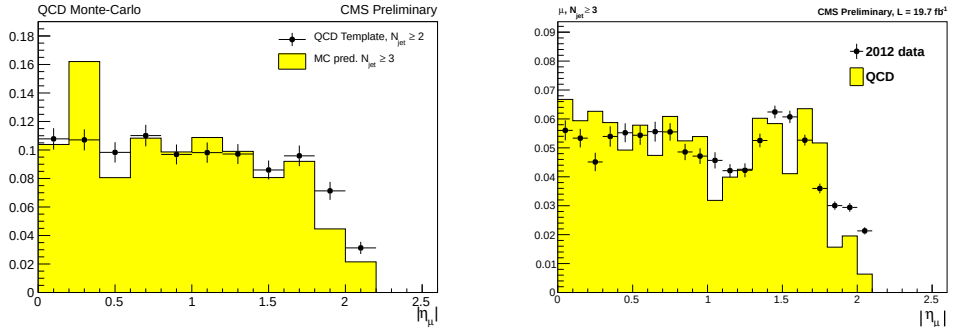

Figure 6.20: Normalised QCD data driven template compared with the MC prediction for $N_{\text{jet}} \geq 2$ (7 TeV analysis left) and $N_{\text{jet}} \geq 3$ (8 TeV analysis right).

section ($\sigma_{\text{t}\bar{\text{t}}}$) in experiment is given by:

$$\sigma_{\text{t}\bar{\text{t}}} = \frac{N_{t\bar{t}}}{\mathcal{L}} \frac{1}{\epsilon_{t\bar{t}}} \tag{6.15}$$

where $N_{t\bar{t}}$ is the number of observed $t\bar{t}$ events, $\epsilon_{t\bar{t}}$ is the overall efficiency of the $t\bar{t}$ events that pass the final selection within the measured phase space and $\mathcal{L}$ is the integrated luminosity of the data being used for the measurement. Cross sections in particle physics are measured in units of barn (b^*).

*$1b = 10^{28} m^2$ is approximately the area of a Uranium nucleus.

Calculations of cross section are made easier by converting the units of $\mathcal{L}$ into fb^{-1}.

The $t\bar{t}$ cross section is found by determining the contribution of the signal and background processes using a maximum likelihood (ML) fitting technique as will be described in the next section. The dominant reducible background processes include:

- W boson production with additional jets (W+Jets);
- Z boson production with additional jets (Z+Jets or Drell-Yan);
- QCD: multi-jet events from soft QCD interactions, in particular where a muon is produced in the decay of a hadron (i.e. a charged pion).

The contributions from these processes can be estimated using the ML fit. Single-top production is the largest irreducible background and must be subtracted using the MC prediction. Diboson backgrounds (WW/WZ/ZZ) have a negligable contribution to the selected data sample.

The total $t\bar{t}$ efficiency, $\epsilon_{t\bar{t}}$, is the number of selected $t\bar{t}$ events reconstructed by the detector, $N^{\text{sel}}_{\text{Reco}}$, divided by the total number of generated $t\bar{t}$ events, $N^{\text{tot}}_{\text{Gen}}$:

$$\epsilon_{t\bar{t}} = \frac{N^{\text{sel}}_{\text{Reco}}}{N^{\text{tot}}_{\text{Gen}}} \tag{6.16}$$

This efficiency correction can be considered as a conversion between the phase space in which the measurement is performed (i.e. accounting for detector and selection effects) and the full theoretical phase space. The value, $\epsilon_{t\bar{t}}$, is obtained from the prediction of the MADGRAPH event generator. The model dependence of this value is assessed in the analysis of the systematic uncertainties (see Chapter 8). Although the measurement is performed in the muon-plus-jets channel, other $t\bar{t}$ decay processes may be selected, such as tauons decaying to muons or fully-leptonic decays where one of the leptons is not detected. The correction above takes this into account when calculating the full inclusive cross section.

6.4.1 Maximum-Likelihood Estimation (MLE)

The number of observed $t\bar{t}$ events, $N_{t\bar{t}}$, and hence the cross section (from Equation 6.15) is extracted using a MLE. The Minuit implementation within the ROOT framework is used to perform this procedure [103]. It is used to estimate the fit parameters (e.g $N_{t\bar{t}}$) such that the underlying model (from the MC) best describes the observed CMS data. The method is equivalent to minimising the χ^2 between the model and the data. The number of events in each $|\eta_\mu|$ bin, i, is expected to be Poissonian distributed. This is because the number of proton-proton collisions tends to infinity whilst the $t\bar{t}$ cross section is low. This leads to a small but finite number of $t\bar{t}$ events being measured in each pseudorapidity bin. The Poissonian function is used in the Likelihood function that describes the probability of the model (the Standard Model) being consistent with the observed data:

$$L(\{\lambda_i, d_i\}) = \prod_i \frac{\lambda_i^{d_i} \cdot e^{-\lambda_i}}{d_i!} \tag{6.17}$$

where λ_i is the expected, and d_i the observed, number of events in each bin, i, of a chosen distribution. It is more convenient to work with the Logarithm of the Likelihood function (LL) such that the product over the bins becomes a sum:

$$LL(\{\lambda_i, d_i\}) = -2 \log\left(\prod_i \frac{\lambda_i^{d_i} \cdot e^{-\lambda_i}}{d_i!}\right) = -2\sum_i \log\left(\frac{\lambda_i^{d_i} \cdot e^{-\lambda_i}}{d_i!}\right) \tag{6.18}$$

The expected pseudorapidity distribution, λ_i, is modeled by process dependent (normalised to unit area) templates (θ_j^i) and normalisation factors (N_j), such that:

$$\lambda_i = \sum_j N_j \theta_j^i \text{ with } \sum_i \theta_j^i = 1 \tag{6.19}$$

for each physics process, j.

The normalization factors can be defined for the various contributing processes, i.e:

- N_{top}: for top processes ($t\bar{t}$ and single-top). Since the $t\bar{t}$ and single-top

$|\eta_\mu|$ distributions are very similar these are combined into one template and hence have one normalisation factor

- N_W: W+Jets background;
- N_{DY}: Drell-Yan (DY) background;
- N_{QCD}: QCD multijet background;
- N_{BG}: W+Jets, DY and QCD can be combined into one background template as was done for the 8 TeV analysis.

The shape of the signal template, θ_{top}, is taken from the MC prediction. In the 7 TeV analysis all of the background templates are constructed using data-driven techniques as described in section 6.3. In the 8 TeV analysis only the QCD was estimated from data due to the low statistics of the QCD MC in the signal region. The 7 TeV data-driven estimates for Z+Jets and W+Jets showed that MC does a good job of describing these backgrounds. The more stringent selection criteria used for the 8 TeV analysis also meant that these data-driven methods would be more difficult to use due to contamination from signal processes. MC was therefore used for these processes in the 8 TeV analysis.

The MLE fit is then performed using the Minuit framework [103]. This fit allows the normalisation factors to vary so as to maximise the value given by the likelihood function. The output of this fit provides the contribution of the signal and background normalisation factors used in the calculation of the $t\bar{t}$ cross section.

In order to calculate the $t\bar{t}$ cross section, the contribution from single-top production must be subtracted from the fitted number of top-like events using the MC prediction:

$$N_{t\bar{t}} = N_{\mathrm{top}}^{\mathrm{fit}} - N_{\mathrm{single-top}}^{\mathrm{MC}} \tag{6.20}$$

Single-top production contributes around 20% of the signal events. Once corrected, $N_{t\bar{t}}$ can be inserted into Equation 6.15 to obtain the final result.

Constraining the Fitted Normalisation Factors

In the likelihood fit it is possible to constrain the allowed values of the background normalisation factors. This can be done by adding additional terms to the LL function. Without constraints there can be quite large statistical fluctuations in the fitted results. It is therefore useful to add constraints to the fitted background normalisation values. These constraints are not absolute limits to the fitted values but are Gaussian in nature. The fit can be constrained by placing a constraint on the ratio of the number of Z+Jets events to W+Jets. For example, the fitted results can be constrained to within 5% of the expected MC value by adding the following term to the LL:

$$\frac{\left(N_Z/N_W - N_Z^{MC}/N_W^{MC}\right)^2}{(0.05 \cdot N_Z^{MC}/N_W^{MC})^2} \tag{6.21}$$

A constraint can also be applied to limit the number of events to within a certain percentage of its initial MC predicted value. This can be done, for example, by adding a 100% constraint on N_{QCD}. Again, an additional term must be added to the LL:

$$\frac{\left(N_{\mathrm{QCD}} - N_{QCD}^{MC}\right)^2}{(1 \cdot N_{QCD}^{MC})^2} \tag{6.22}$$

Different constraints were used in the 7 and 8 TeV analyses. The above ratio constraint on the Z+Jets and W+Jets normalisation factors was applied for the 7 TeV analysis whilst a 50% constraint was applied on the N_{BG} value (since the Z+Jets, Z+Jets and QCD templates were combined) for the 8 TeV analysis. The Z+Jets and W+Jets were combined for the 8 TeV analysis because more stringent selection criteria were applied, and so more MC statistics were required. The size of the normalisation constraints are justified, since they are large in comparison to the uncertainties predicted by theory [104].

6.5 Results

As has been described previously, the MLE method predicts the total number of signal and background events after selection. Figure 6.21 shows the normalised templates used as input into the binned likelihood fit of the muon pseudorapidity for both the 7 and 8 TeV analyses (left plots). The right plots in figure 6.21 show the muon pseudorapidity with the number of signal and background events scaled to the results from the template fit for the MC. A good agreement is seen with the data.

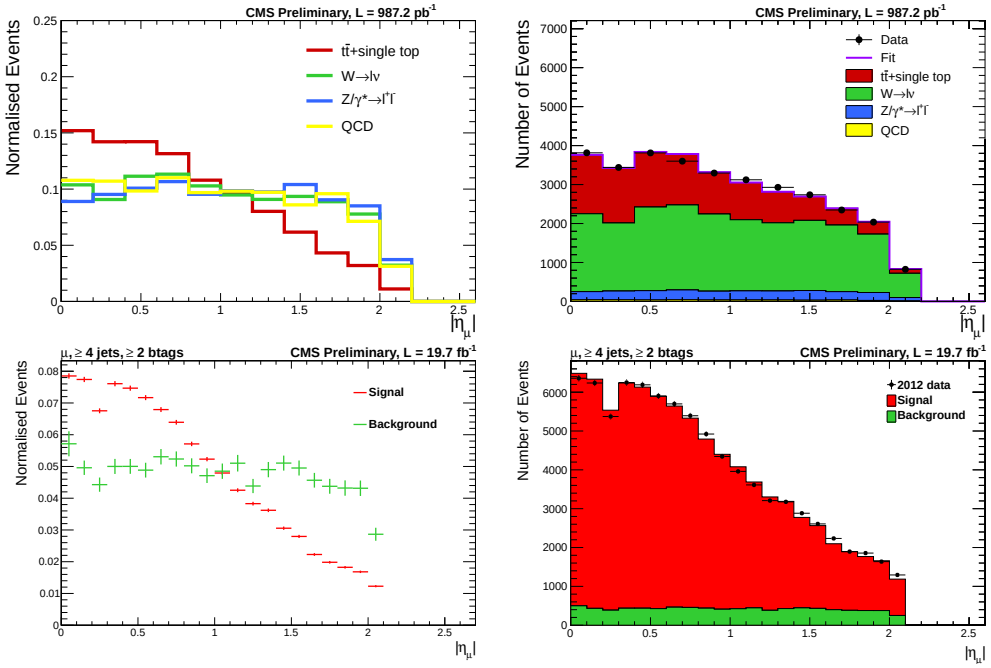

Figure 6.21: The left plots show the normalised templates used in the binned likelihood fit for the 7 TeV analysis (top) and the 8 TeV analysis (bottom). The background templates for the 7 TeV analysis are all estimated using data-driven techniques. Simulation is used to model the signal template and also the 8 TeV background template excluding the QCD component, which is always estimated using the data. The right plots show the muon pseudorapidity normalised to the number of events obtained from the fit for the 7 TeV analysis (top) and the 8 TeV analysis (bottom). It was possible to use a finer binning in the 8 TeV analysis as more statistics were available.

The number of events predicted by MC for each process is compared with

that obtained from the fit in Table 6.8. The number of top events from the fit and Equation 6.15 are used to calculate the $t\bar{t}$ production cross section. The measured cross section values are:

$$\sigma_{t\bar{t}}^{7\text{ TeV}} = 145.6 \pm 8.2\text{ (stat.)}\ ^{+38.1}_{-28.3}\text{ (syst.) pb}$$

$$\sigma_{t\bar{t}}^{8\text{ TeV}} = 237.4 \pm 1.9\text{ (stat.)}\ ^{+20.4}_{-16.9}\text{ (syst.) pb}$$

A discussion of the systematic uncertainties (syst.) and how these are estimated is described in Chapter 8.

7 TeV using 987.2 pb^{-1}					
Sample	Top	W+jets	Z+jets	QCD	Total
MC estimation	11170.8 ± 21.4	17893.3 ± 105.0	2221.9 ± 14.2	317.4 ± 37.2	31603.3± 114.3
Fit result	9941.5±463.9	19253.0±510.2	2398.8±121.4	361.5±314.0	31954.7±767.4
8 TeV using 19.7 fb^{-1}					
Sample	Top	W+jets	Z+jets	QCD	Total
MC estimation	78903.8 ± 241.9	2184.4 ± 24.2	370.4 ± 4.6	478.7 ± 412.9	81937.3 ± 475.5
Fit results	76310.1 ± 600.8	6243.7 ± 611.4	1058.6 ± 134.9	1367.1 ± 152.1	84979.5 ± 804.3

Table 6.8: Expected number of events from MC (before fit) and number of events from the fit for each process.

6.6 Summary

In this section the measurement of the inclusive cross section at both 7 and 8 TeV has been presented. The differences between the data samples meant that slightly different selection methods and estimations of the backgrounds have been used. The final results are within agreement of the theoretical predictions made at NNLO of 163 pb and 234 pb for 7 and 8 TeV, respectively. The methods used in this chapter are built upon to perform differential cross sections which are shown in the next chapter.

Chapter 7

Global Event Level Observables in $t\bar{t}$ Events at $\sqrt{s} = 8$ TeV

Following on from the measurements of the inclusive cross sections at 7 and 8 TeV presented in the previous chapter, the next natural step is to measure differential production cross sections. These measurements have been made achievable by the vast quantities of data obtained by the CMS experiment in the 2011 and 2012 data taking periods. Differential measurements can be made with respect to interesting properties of the top quark. Measurements by both ATLAS [105] and CMS [93] at 7 TeV have previously been performed. These measurements are with respect to the invariant mass of the $t\bar{t}$ pair ($m_{t\bar{t}}$), the p_{T} of the top quarks, as well as various other properties of the top quark and its decay products. Some of these distributions require kinematic reconstruction of the top pair events. This reconstruction determines which decay products are associated with each top quark. In the analysis presented in this thesis, no kinematic reconstruction is required as all variables of interest can be reconstructed easily without assigning decay products to top quarks. The work presented in this chapter was done by the author in collaboration with the Bristol Top Physics group.

To extend previous studies, a measurement of the differential cross section with respect to $E_{\mathrm{T}}^{\mathrm{miss}}$ in top pair production was performed using the 7 TeV dataset [106]. The $E_{\mathrm{T}}^{\mathrm{miss}}$ distribution from this analysis is shown in Figure

7.1.

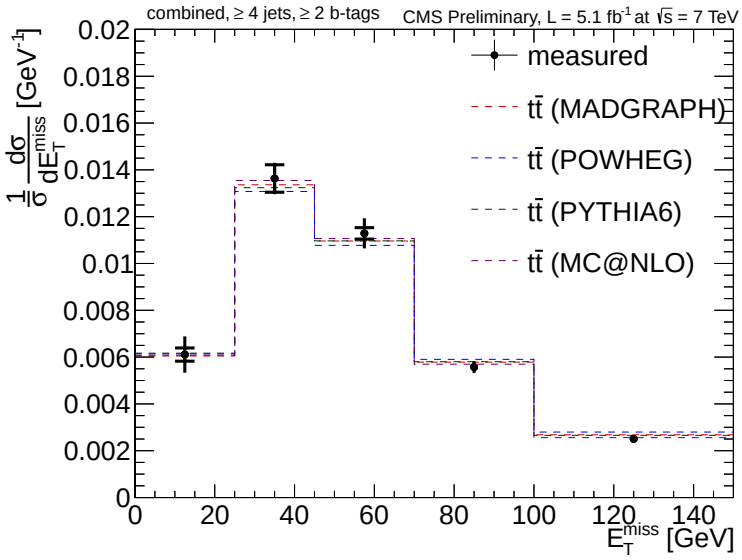

Figure 7.1: Comparison of measured normalised differential cross section with respect to $E_{\mathrm{T}}^{\mathrm{miss}}$ compared to different generators using the 7 TeV data.

Building on this measurement and using the full 2012 dataset, measurements of other event-level variables in the production of top-quark pairs have been studied. This includes variables such as the H_{T} which is defined as the the sum of the p_{T} of the jets measured in an event.

Differential cross section measurements of the global event variables will help verify the models that are embedded within the MC generators. This is done by comparing the result with the predictions of the generators listed in Table 2.3. Comparisons can also be made to the choice of renormalisation and factorisation (Q^2) scale and the matching threshold of the partons. This makes it possible to constrain the uncertainties due to these model inputs. Additional gluons are commonly produced in roughly 50% of $t\bar{t}$ decays. Differential measurements involving additional jets are a good test of perturbative QCD physics and the running of the strong coupling constant.

Measuring distributions such as $E_{\mathrm{T}}^{\mathrm{miss}}$ may help to uncover rare Standard Model processes such as top quark pair plus boson production. Additional $E_{\mathrm{T}}^{\mathrm{miss}}$ would appear in top pair decays with an associated vector boson ($t\bar{t}$+

$W \to l\nu$ or $t\bar{t}+ Z \to \nu\bar{\nu}$) from the presence of additional neutrinos. Top pair plus Higgs boson production is another example where additional $E_{\mathrm{T}}^{\mathrm{miss}}$ can be produced. This can happen when the Higgs boson decays to two Z bosons which in turn decay to two neutrinos. There is also a new physics scenario where the Higgs boson decays to massive non-interacting particles.

Top pair events are a major background to beyond the Standard Model (BSM) physics processes. BSM searches such as those for supersymmetry (SUSY) where the SUSY partner of the top ($\tilde{t}\bar{\tilde{t}} \to t\tilde{\chi_0}\bar{t}\bar{\tilde{\chi}}_0$) where $\tilde{\chi_0}$ is the lightest super-symmetric particle and is a possible candidate for dark matter. Exotic heavy partners of the top, e.g the T'^* have a very similar event signature to $t\bar{t}$. These searches often use variables such as $E_{\mathrm{T}}^{\mathrm{miss}}$ and H_{T}. It is therefore important that these distributions are well understood for the $t\bar{t}$ process. The tails of these distributions are particularly interesting as this is where signs of new physics would be likely to show up.

*A T' is a heavy partner that decays in a similar way to the top

7.1 Analysis Strategy

The differential cross section measurements have been performed using similar methods to those described in Chapter 6. The same selection criteria as were used in the 8 TeV inclusive cross section analysis have been used. A tight selection using at least two b-tags gives a high $t\bar{t}$ purity and hence the uncertainties on the backgrounds are reduced.

As with the inclusive cross section analyses, a maximum likelihood fit of the $|\eta_\mu|$ distribution was implemented as described in Section 6.4. The same dataset as for the inclusive cross section was used. However, this data has to be sub-divided into bins of each variable under study. The choice of the binning aims to reduce the migration of events between bins due to detector resolution effects and is described in Section 7.3.

Once the number of events has been extracted for a given bin in the selected (visible) phase space, the result has to be corrected back to the full theoretical phase space. This process is known as "unfolding" and is performed so that the result can be compared to different theoretical models as well as results from different experiments. In this analysis SVD unfolding has been implemented, which is described in detail in Section 7.5.

Finally, the result is combined with those from the corresponding electron-plus-jets channel and normalised using the total measured cross section. The final results are then presented as a normalised differential cross section in each bin of any given variable, X:

$$\frac{1}{\sigma_{\mathrm{t\bar{t}}}}\frac{d\sigma_{\mathrm{t\bar{t}}}}{dX} \tag{7.1}$$

$\frac{d\sigma_{\mathrm{t\bar{t}}}}{dX}$ is the differential cross section with respect to a variable, X and $\sigma_{\mathrm{t\bar{t}}}$ is the inclusive cross section measured for a particular variable given by:

$$\sigma_{\mathrm{t\bar{t}}} = \int_0^\infty \frac{d\sigma_{\mathrm{t\bar{t}}}}{dX} dX \tag{7.2}$$

The normalisation is performed because cancellations of the systematic uncertainties occur giving a more precise result. It is also useful in the presenta-

tion of the results when comparing to various models of the MC generators. Details about how the final results are calculated are given in Section 7.6.

7.2 Definitions of Event Level Observables

It is first of all important to formally define the variables that have been studied and to express mathematically how they are calculated experimentally. Sometimes variables are defined slightly differently at reconstruction and generator level. These differences will be accounted for in the unfolding to the generator level variable. It is also important to be clear about exactly what is being unfolded. It is possible to unfold the $E_{\mathrm{T}}^{\mathrm{miss}}$ to the generator $E_{\mathrm{T}}^{\mathrm{miss}}$ or the generated p_{T} of the neutrino. The generator $E_{\mathrm{T}}^{\mathrm{miss}}$ is different from the neutrino p_{T} as it includes missing energy from noise in parts of the detector. The variables under study are given in Table 7.1.

$E_{\mathrm{T}}^{\mathrm{miss}}$	Missing transverse energy
H_{T}	Scalar sum of all jet transverse momentum (p_{T}) in event
S_{T}	Scalar sum of jet p_{T}, lepton p_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$
p_{T}^{W}	p_{T} of W boson (that decays leptonically)
M_{T}^{W}	transverse mass of W boson (that decays leptonically)

Table 7.1: Variables under study.

The transverse momentum of a particle is defined as $p_{\mathrm{T}} = \left(p_x^2 + p_y^2\right)^{\frac{1}{2}}$. The overall missing transverse momentum in an event, or $E_{\mathrm{T}}^{\mathrm{miss}}$, is defined as

$$E_{\mathrm{T}}^{\mathrm{miss}} = -\left[\left(\sum_i p_{x,i}\right)^2 + \left(\sum_i p_{y,i}\right)^2\right]^{\frac{1}{2}} \tag{7.3}$$

where the sums extend over all measured particles in the event. It was decided to unfold to the transverse momentum of the neutrino as it is a more fundamental particle-level definition and therefore more useful to theorists.

The H_{T} in an event is defined as the scalar sum of the p_{T} of all jets in

the event:

$$H_{\mathrm{T}} = \sum_{\text{jets}} |p_{\mathrm{T}}| \tag{7.4}$$

At reconstruction level at least four jets are required to have a p_{T} of 30 GeV and additional jets must have a p_{T} of 20 GeV. Generator level jets are reconstructed from generated particles in an event using the clustering algorithm described in Section 5.5.

S_{T} is the scalar sum of the p_{T} of all jets, plus the $E_{\mathrm{T}}^{\mathrm{miss}}$ and the p_{T} of the single isolated lepton presumed to come from the decay of one of the W bosons in the event:

$$S_{\mathrm{T}} = H_{\mathrm{T}} + E_{\mathrm{T}}^{\mathrm{miss}} + |p_{\mathrm{T}}^{\mathrm{lepton}}| \tag{7.5}$$

The reconstructed lepton p_{T} could potentially come from a lepton that is not from the decay of a W boson. At generator level only the lepton from the W in semi-leptonic top decays are used in the calculation of S_{T}.

The transverse momentum of the W boson, p_{T}^{W}, is derived from the single isolated lepton and the $E_{\mathrm{T}}^{\mathrm{miss}}$ in the event:

$$p_{\mathrm{T}}^{W} = \sqrt{(p_x^{\mathrm{lepton}} + p_x^{\mathrm{miss}})^2 + (p_y^{\mathrm{lepton}} + p_y^{\mathrm{miss}})^2} \tag{7.6}$$

at reconstruction level. At generator level however the W boson p_{T} can be accessed directly.

Finally, M_{T}^{W} is the transverse mass of the W boson in the event, again using the single isolated lepton and $E_{\mathrm{T}}^{\mathrm{miss}}$

$$M_{\mathrm{T}}^{W} = \sqrt{(E_{\mathrm{T}}^{\mathrm{miss}} + \mathrm{E}_{T}^{\mathrm{lepton}})^2 - {p_{\mathrm{T}}^{W}}^2} \tag{7.7}$$

p_{T}^{W} and M_{T}^{W} are calculated in the same way at reconstruction and generator level.

7.3 Binning of Variables

In order to reduce statistical fluctuations when unfolding the differential cross section measurements, the binning of the variables must be chosen carefully

so as to minimise migration between neighbouring bins. The event statistics (N^k_{events}) within a bin after the final selection must also be substantial such that the likelihood fit can be performed (≥ 3000 events). In order to quantify the effect of bin migration, the purity, p^k, and stability, s^k, of a bins are defined:

$$p^k = \frac{N^k_{\text{rec\&gen}}}{N^k_{\text{rec}}} \tag{7.8}$$

$$s^k = \frac{N^k_{\text{rec\&gen}}}{N^k_{\text{gen}}} \tag{7.9}$$

where $N^k_{\text{rec\&gen}}$ is the number of events both generated and reconstructed in the same bin and N^k_{rec} and N^k_{gen} are the number of events reconstructed and generated, respectively, within a bin after the final event selection. The purity parameter is sensitive to migration into a bin while the stability parameter is sensitive to migration out of a bin. These values are calculated using $t\bar{t}$ Monte Carlo events.

Binning is chosen such that p^k and s^k values do not fall far below 0.5. This is done by incrementing over a range of values for each of the variables until threshold p^k, s^k and N^k_{events} are obtained. The boundaries of bins were chosen such that they are rounded to the nearest 5 GeV to improve aesthetics. The 2D, reconstructed vs. generated variable distributions are shown in Figure 7.2. Purity and stability values along with the number of events in each chosen bin for the five variables are presented in Tables 7.2 to 7.6. The H_{T} and S_{T} variables both have quite a large first bin. This is because of the kinematic requirements on the jets and the muon at reconstruction level. The values of purity and stability are consistent between the muon and electron-plus-jets channels.

	$0< E_{\text{T}}^{\text{miss}} <25$	$25\leq E_{\text{T}}^{\text{miss}} <45$	$45\leq E_{\text{T}}^{\text{miss}} <75$	$75\leq E_{\text{T}}^{\text{miss}} <100$	$100\leq E_{\text{T}}^{\text{miss}} <150$	$E_{\text{T}}^{\text{miss}} \geq 150$
events	11506	20348	22594	14413	8416.5	3274
p^k	0.52	0.49	0.45	0.42	0.52	0.69
s^k	0.43	0.42	0.47	0.5	0.66	0.87

Table 7.2: Stability and purity of the chosen $E_{\text{T}}^{\text{miss}}$ bins.

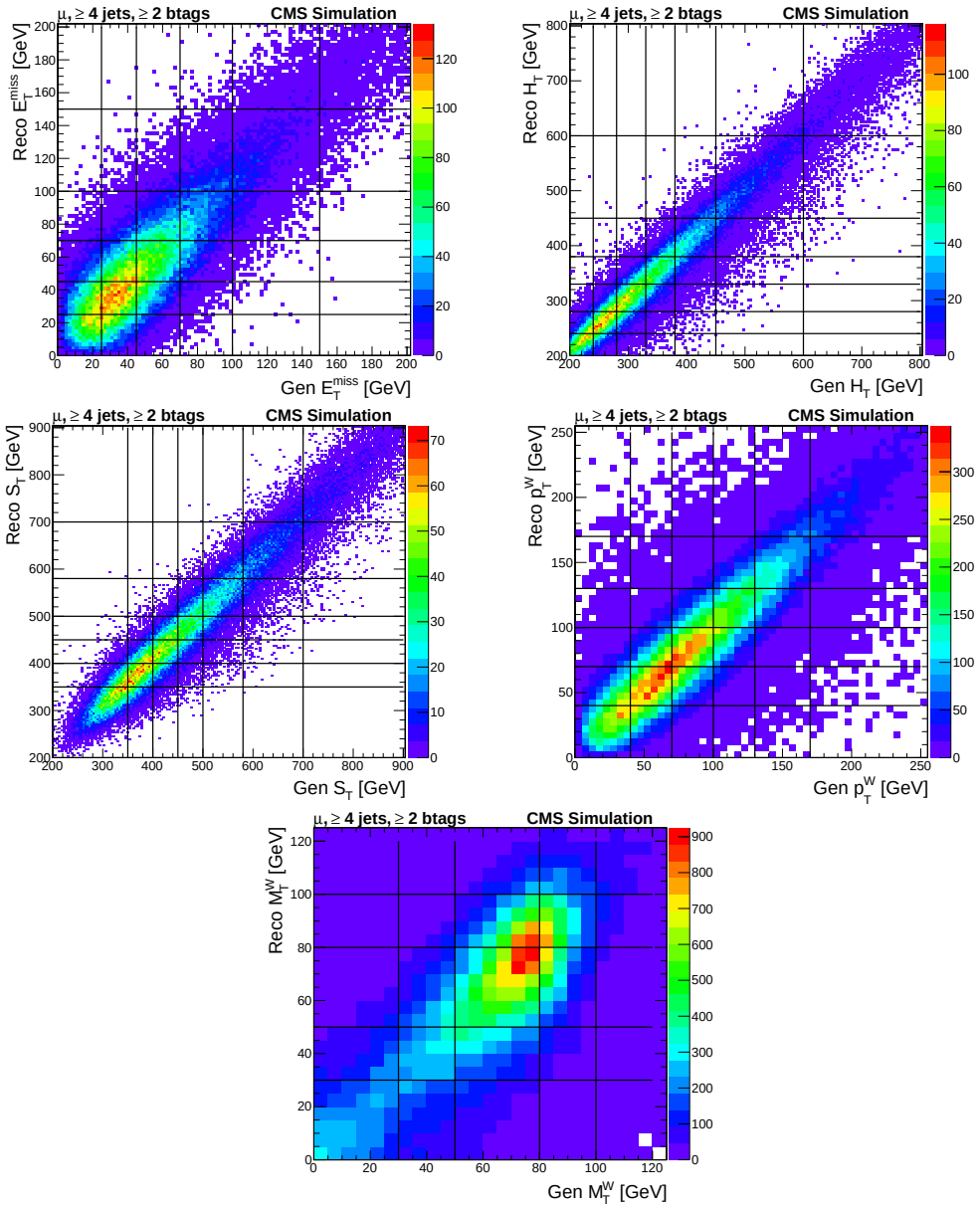

Figure 7.2: The above distributions show the reconstructed versus generated values for all of the variables. These 2D histograms are used to calculate the p^k and s^k values. The lines indicate the final choice of the binning. The z-axis represents the number of events in MC after event reweighting.

	$0< H_T<240$	$240\leq H_T<280$	$280\leq H_T<330$	$330\leq H_T<380$	$380\leq H_T<450$	$450\leq H_T<600$	$H_T\geq 600$
events	12337	14821	16747	14268	12258	12119	5856.7
p^k	0.82	0.63	0.64	0.63	0.67	0.81	0.9
s^k	0.76	0.64	0.65	0.65	0.68	0.81	0.89

Table 7.3: Stability and purity of the chosen H_T bins.

	$0< S_T<350$	$350\leq S_T<400$	$400\leq S_T<450$	$450\leq S_T<500$	$500\leq S_T<580$	$580\leq S_T<700$	$S_T\geq 700$
events	13593	15717	13878	12307	12732	10135	8874.8
p^k	0.83	0.61	0.54	0.54	0.64	0.71	0.87
s^k	0.74	0.61	0.55	0.56	0.66	0.74	0.9

Table 7.4: Stability and purity of the chosen S_T bins.

	$0< p_T^W<40$	$40\leq p_T^W<70$	$70\leq p_T^W<100$	$100\leq p_T^W<130$	$130\leq p_T^W<170$	$p_T^W\geq 170$
events	12379	18821	18330	13495	10260	8393.9
p^k	0.67	0.55	0.52	0.5	0.55	0.76
s^k	0.63	0.55	0.53	0.51	0.56	0.76

Table 7.5: Stability and purity of the chosen p_T^W bins.

	$0< M_T^W<30$	$30\leq M_T^W<50$	$50\leq M_T^W<80$	$80\leq M_T^W<100$	$M_T^W\geq 100$
events	13209	13636	30669	16974	9179.3
p^k	0.56	0.36	0.66	0.42	0.37
s^k	0.64	0.38	0.54	0.42	0.66

Table 7.6: Stability and purity of the chosen M_T^W bins.

7.4 Maximum Likelihood Fits

The same maximum likelihood estimation method was used as for the 8 TeV inclusive cross section measurement described in Section 6.4. The only difference being that instead of using the full data set that passed the selection, the data and the MC used in the fit are divided into the bins chosen in the previous section.

Three normalised templates are fit to the data: Signal, V+Jets and QCD. The same QCD template as was extracted from the data for the inclusive cross section was used. It was found that within statistical errors, the template shape is the same for all H_{T} bins. As usual, the Signal and V+Jets templates come from the MC prediction. The templates used for the $E_{\mathrm{T}}^{\mathrm{miss}}$ variable are shown in Figure 7.3, as an example.

After performing the fits for each variable, the total number events for each process has been used as the normalisation for the variables under study. These are shown in Figure 7.4. The correlations between the fitted input parameters for the $E_{\mathrm{T}}^{\mathrm{miss}}$ variable are shown in Appendix B.

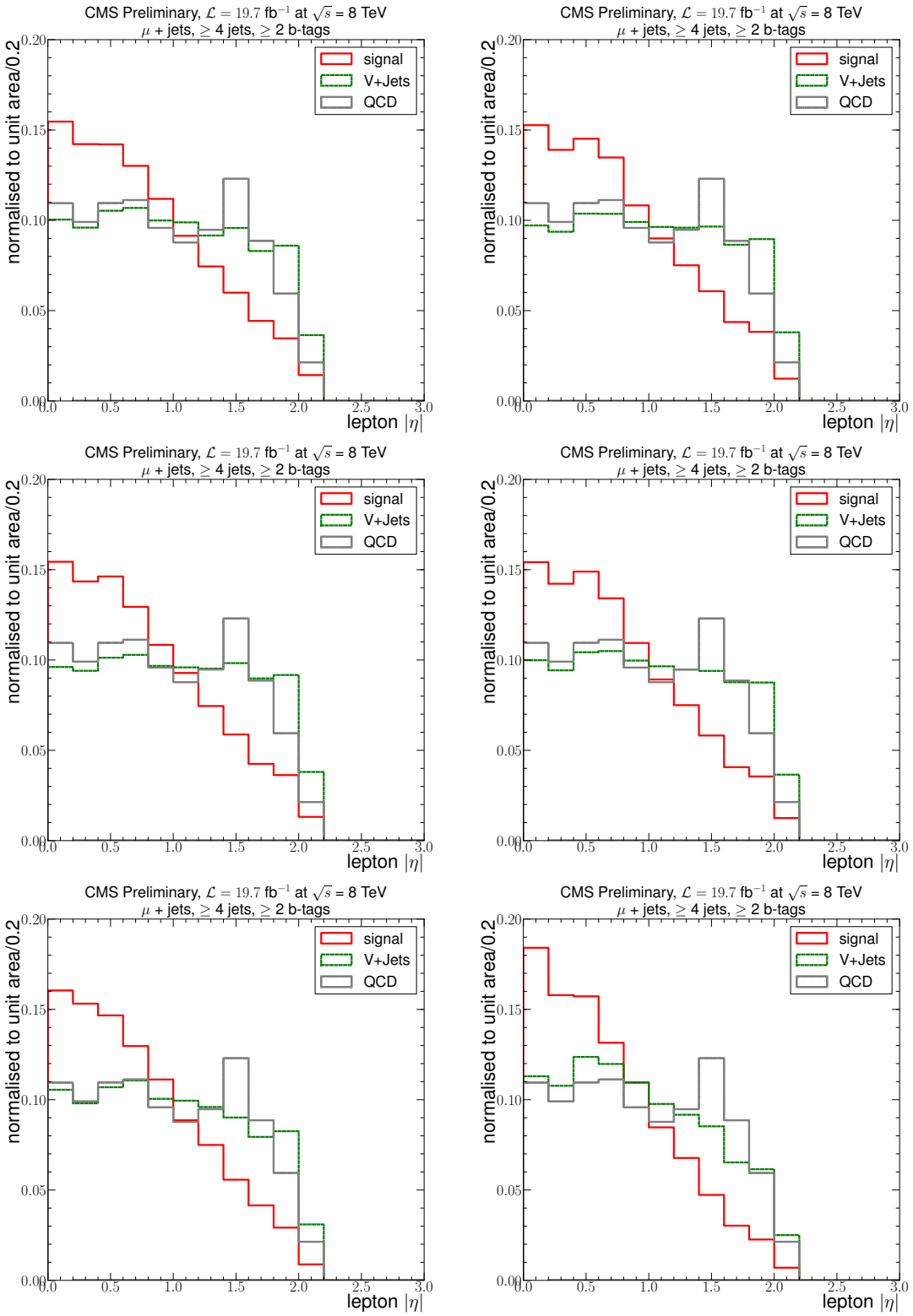

Figure 7.3: Muon $|\eta|$ templates for the fit in ≥ 2 b-tag events in different bins of $E_{\mathrm{T}}^{\mathrm{miss}}$: 0-25 GeV(top left), 25-45 GeV(top right), 45-70 GeV(middle left), 70-100 GeV(middle right), 100-150 GeV(bottom left) and ≥ 150 GeV(bottom right).

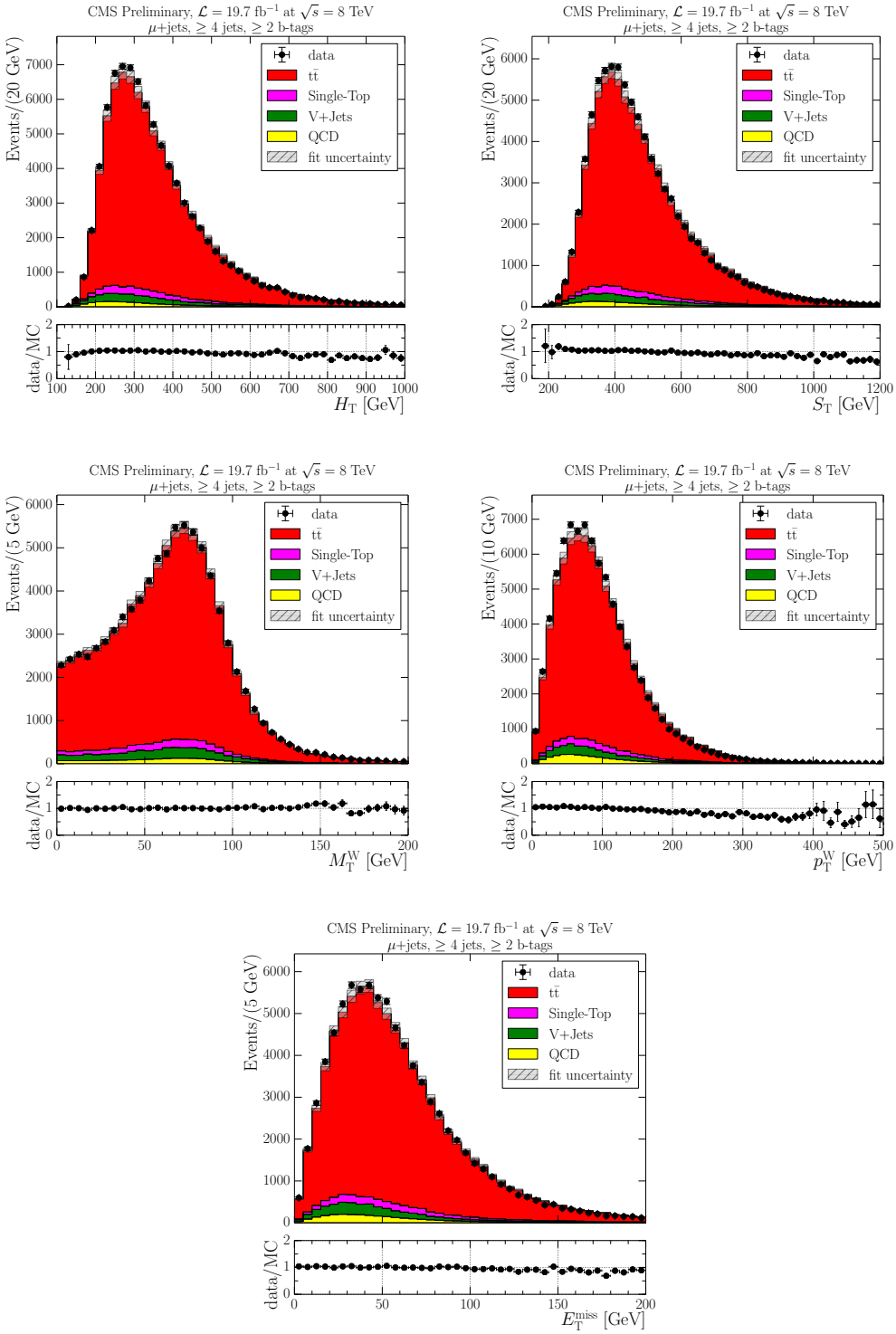

Figure 7.4: Data to MC comparison plots after the final event selection. The MC is normalised to the fit results obtained from each individual variable. The following variables are shown: H_T(top left), S_T (top right), M_T^W (middle left), p_T^W (middle right) and E_T^{miss} (bottom). The distributions below the varibales display the data/MC ratio.

7.5 Unfolding

Measured quantities are bound to certain restrictions. These arise as a result of finite detector resolution and acceptance constraints (for example, from applying the trigger and event selection). This can lead to $t\bar{t}$ events with a true $E_{\mathrm{T}}^{\mathrm{miss}}$ in a bin, i, ending up in an adjacent bin ($i \pm 1$) or not even passing the selection at all. It is of great interest to be able to interpret results independently of the detector and any selection requirements. Such results enable a direct comparison to be made between theoretical predictions as well as between different experiments.

A distribution of *true* x values will be measured as a distribution, y, with potentially different values. Here, x and y can be considered as vectors containing the contents of each bin, i:

$$x_n = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \tag{7.10}$$

and

$$y_m = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \tag{7.11}$$

where n is the number of bins for the true distribution and m for the measured distribution. In this analysis $m = n$ for all the considered variables. The migration of events and acceptance effects can be described by inserting an $m \times n$ response matrix $A_{m \times n}$ which represents the convolution between the measured the true distributions:

$$y_m = A_{m \times n} \times x_n \tag{7.12}$$

The response matrix is constructed using events from MC to form a probability density matrix of an event migrating from a bin m to n. By simply

inverting the matrix A, and applying it to the background subtracted data, the true "unfolded" distribution can be obtained:

$$u_m = A^{-1}_{m\times n} \times b_n \tag{7.13}$$

where b_n and u_m are the measured and unfolded vectors, respectively. These are given different letters to x_n and y_m to differentiate the use of data (b_n and u_m) from MC (x_n and y_m). The direct inversion of the response matrix is known as bin-by-bin unfolding and was used to obtain the final result for the 7 TeV $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement [106]. This method can produce large statistical fluctuations in the unfolded result. These spurious statistical fluctuations can arise from off diagonal elements of the response matrix with poor statistics. The problem is highlighted using a simple two binned example from reference [107]. In this simple case, the response matrix can be defined such that ϵ is introduced to account for inefficiencies in detection within a bin:

$$A_{2\times 2} = \frac{1}{2}\begin{pmatrix} 1+\epsilon & 1-\epsilon \\ 1-\epsilon & 1+\epsilon \end{pmatrix} \tag{7.14}$$

After the direct inversion of this response matrix the following unfolded vector, u, can be obtained by:

$$u = A^{-1}_{2\times 2} b = \frac{b_1 - b_2}{2\epsilon}\begin{pmatrix} 1 \\ -1 \end{pmatrix} + \frac{b_1 + b_2}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \tag{7.15}$$

It is clear that if ϵ is small the first term will dominate producing a nonsense value. This is why care was taken when choosing the binning inorder to minimise the migration of events, thus, keeping ϵ reasonably high (> 0.4). Another problem could arise if $(b_1 - b_2)^2 \leq (b_1 + b_2)$ the number of events in neighbouring bins is not statistically significant and the solution will be a random number. It was reasonable to perform bin-by-bin unfolding for the $E_{\mathrm{T}}^{\mathrm{miss}}$ variable as the purity and stability values are relatively high and there were quite large variations between bins and hence the values were statistically significant. A more accurate solution can be obtained by regularising

the unfolding which requires there to be certain constraints placed on the solution.

7.5.1 Singular Value Decomposition (SVD) Unfolding

Various regularised unfolding techniques and implementations exist. The method that has been adopted for this analysis is the singular value decomposition (SVD) approach [107] which uses a regularised inversion of the response matrix. Regularisation takes *a priori* knowledge of the true underlying distribution i.e. its curvature. Another constraint could be to impose that the result must have positive values. These constraints help to eliminate the spurious statistical fluctuations in the unfolded result. In SVD unfolding, only statistically significant contributions from the response matrix are used to derive the solution.

SVD unfolding is a simple implementation that decomposes the response matrix, A, into set of linear orthogonal equations such that exact solutions can be found for the values of u_i. This is done by applying rotation matrices (U and V^T) to the left and the right such that a diagonal matrix, S, is obtained:

$$A = USV^T \tag{7.16}$$

U and V^T are both orthogonal such that:

$$U^TU = UU^T = I \qquad V^TV = VV^T = I \tag{7.17}$$

The S matrix contains the *singular values*, s_{ij}, for the unfolding along the diagonal. The values within S are equal to 0 for $i \neq j$. A perfect detector has singular values that are equal to 1. In most cases however, some of the singular values will be low. These singular values correspond to components of the response matrix that are not statistically significant.

The problem of unfolding can be rewritten in such a form that the χ^2 between the unfolded result and the true distribution is minimised:

$$\chi^2 = (Au - b)^T \mathbf{cov}^{-1}(Au - b) + \tau^2 \cdot (Cu)^T(Cu) \tag{7.18}$$

The first term corresponds to the direct inversion of the response matrix with $\mathbf{cov}^{-1}$ being the covariance matrix for the measurement vector. The second term is the imposed *regularisation* condition. C is a matrix that contains an *a priori* curvature condition in order to minimise the curvature of the solution. τ is the regularisation strength parameter. A τ value set to zero means no regularisation and a high τ value can bias the result towards the regularisation condition. The regularisation causes the matrix containing the singular values to be reweighted emphasising the statistically significant terms. The RooUnfold [108] implementation of the the SVD algorithm has been used in this analysis. The regularisation term which is required as input in this implementation is a *k-value*. This *k-value* corresponds to the number of statistically significant values within the S matrix. It is also related to τ such that:

$$\tau = s_k^2 \tag{7.19}$$

where s_k is the k^{th} most statistically significant singular value. The *k-value* is chosen as to minimise the χ^2 of the solution to equation 7.18. By doing this the bias of the imposed curvature condition will be reduced whilst preventing large statistical fluctuations. The choice of *k-value* will now be discussed.

7.5.2 Closure Test and Choice of Regularisation Parameter

The first thing to check before justifying the choice of regularisation parameter is that the unfolding procedure works correctly using the RooUnfold framework. A closure test was performed using MC. Applying the unfolding to the MC distribution after event selection should return the MC truth distribution before any selection. This should be the case regardless of the choice of regularisation parameter. The result of the closure test can be seen in Figure 7.5.

A regularisation parameter is chosen to regulate the large statistical fluctuations that can arise from unfolding whilst not biasing the result to that of the MC input. There are various methods that can be used to determine the

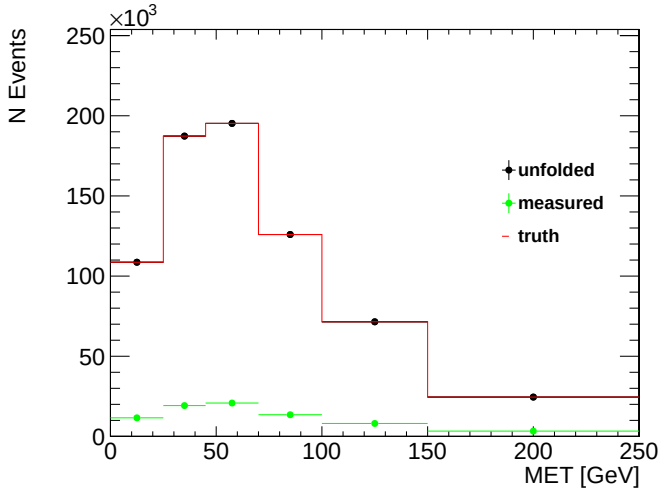

Figure 7.5: Closure test performed for the $E_{\mathrm{T}}^{\mathrm{miss}}$ variable. As expected, applying the response matrix to the MC measured returns the MC truth. The MADGRAPH MC generator was used for this closure test.

regularisation parameter, k or τ. In this analysis the *k-value* has been determined based on a suggested method in [107]. In this method, the closure test is again performed but this time using a different MC generator (MC@NLO) to produce the response matrix. The χ^2 is then calculated between the MC truth and the unfolded result using the MC@NLO response matrix. The χ^2 for all *k-values* is calculated within a range of 2 to the number of bins of the distribution. The *k-value* with the smallest χ^2 is that which is chosen to do the unfolding on the data. Table 7.7 presents the χ^2 values obtained for the different variables. A cross check of these values was performed using the POWHEG MC generator. This yielded similar results.

The H_{T}, S_{T} and M_{T}^{W} variables had a low optimised choice of *k-value* of two. The reasons for the low choice of *k-value* for these variables is that there is quite a low variation in event numbers between neighbouring bins. Therefore, there is a low statistical significance between bins as the truth distribution has a low curvature. There is more curvature between bins for the $E_{\mathrm{T}}^{\mathrm{miss}}$ and p_{T}^{W} variables and so optimised *k-values* of three and five are chosen, respectively. Due to the increased statistical significance between bins, less regularisation is required.

χ^2 values for different variables					
k-value	$E_{\mathrm{T}}^{\mathrm{miss}}$	H_{T}	S_{T}	p_{T}^{W}	M_{T}^{W}
2	215.2	2238.0	2293.0	576.4	268.8
3	79.6	5834.9	4291.4	284.5	803.0
4	298.3	7710.3	5319.8	266.9	1174.2
5	370.8	9126.6	6302.9	265.5	1693.8
6	418.5	10019.1	6866.0	266.4	NA

Table 7.7: χ^2 values obtained for the different variables in order to choose the *k-value*.

7.6 Normalised Differential Cross Section Calculation

After performing the unfolding on the fitted results, the number of $t\bar{t}$ events, $N_{\mathrm{t\bar{t}}}^{i}$, in the semi-leptonic decay channel is obtained for each bin, i, of a given variable. This value can then be converted into a partial cross section, $\Delta\sigma_{\mathrm{t\bar{t}}}^{i}$, for $t\bar{t}$ production in each bin by dividing by the total integrated luminosity:

$$\Delta\sigma_{\mathrm{t\bar{t}}}^{i} = \frac{N_{\mathrm{t\bar{t}}}^{i}}{\mathcal{L}} \tag{7.20}$$

Notice that this calculation no longer needs an efficiency correction. The unfolding takes care of this correction. The differential cross section for each variable, X, is then found after dividing the contents of each bin by the width of the bin, $\Delta\mathrm{X}^{i}$ (in GeV):

$$\frac{d\sigma_{\mathrm{t\bar{t}}}^{i}}{d\mathrm{X}} = \frac{\Delta\sigma_{\mathrm{t\bar{t}}}^{i}}{\Delta\mathrm{X}^{i}} \tag{7.21}$$

These are the widths that were determined in Section 7.3. Finally, the normalised differential cross sections are calculated using the sum of the partial differential cross sections:

$$\frac{1}{\sigma_{\mathrm{t\bar{t}}}^{\mathrm{tot}}}\frac{\mathrm{d}\sigma_{\mathrm{t\bar{t}}}^{i}}{\mathrm{dX}} = \frac{1}{\sum_i \frac{\mathrm{d}\sigma_{\mathrm{t\bar{t}}}^{i}}{\mathrm{dX}}}\frac{\mathrm{d}\sigma_{\mathrm{t\bar{t}}}^{i}}{\mathrm{dX}} \tag{7.22}$$

The electron and muon semi-leptonic channels are combined by first performing the fitting and unfolding procedures in the respective channels. The

distributions output by the individual channels are then summed. The final combined normalised differential cross section is determined using the equations above.

7.7 Results

The final normalised differential cross section results in the combined electron and muon-plus-jets channels are presented for each variable: $E_{\mathrm{T}}^{\mathrm{miss}}$, H_{T}, S_{T}, p_{T}^{W} and M_{T}^{W} in Figures 7.6 to 7.10. Systematic uncertainties will be discussed separately in Chapter 8. The results are compared to the predictions from MADGRAPH using different tunes of the Q^2 parameter and the parton to jet matching threshold (left). Comparisons are also made to the MC@NLO and POWHEG MC generators (right). The differences between these MC generators was outlined in Chapter 2.

The results show no significant deviations from the predictions made by the MC generators. No excess was observed in the tails of any of the distributions where signs of new physics would have been expected to show up. It has not been possible to make comparisons to approximate NNLO predictions as these are not available. The values in each bin for the presented variables along with the uncertainty associated with each value can be seen in Table 7.8. Results presented individually in the electron and muon+jets channels can be found in Appendix C.

It has been possible to differentiate the predictions of the various generators for some of the measurements. The $E_{\mathrm{T}}^{\mathrm{miss}}$ variable does not have very good separating power, as the different generators give similar predictions. The uncertainties on the measurement are not small enough to make any conclusions about which is best. In contrast, for the H_{T} and S_{T} variables, reasonably large distinctions can be made between the predictions. The measurements for both variables lie between the Q^2 up and down variations on the renormalisation and factorisation scale. This suggests that these variations are over estimated and that it is correct to use the top mass as the value of renormalisation. When comparing the different generators, quite a large difference is observed between MC@NLO and both MADGRAPH and

POWHEG. This is down to the fact that MC@NLO used a different hadronisation model (Herwig rather than Pythia). Pythia seems to give the better agreement at softer energy scales.

The W boson p_{T} distributions yields the largest discrepancy between the data and the MC. It is reasonable to suggest that this is caused by the same MADGRAPH mismodelling which resulted in the discrepancy in the top-p_{T} measurement [93]. The large uncertainties on the M_{T}^{W} measurement make it hard to draw any firm conclusions. This measurement will be improved in the future by reducing bin migrations.

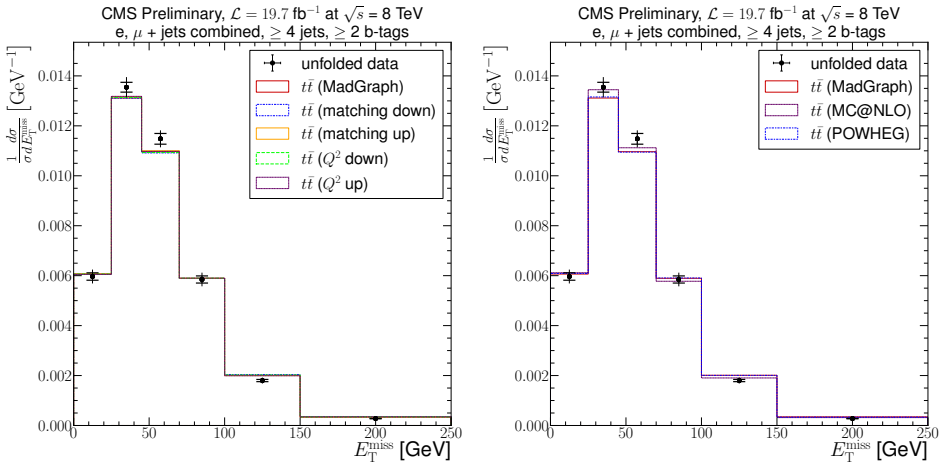

Figure 7.6: Comparison of measured normalised differential cross section with respect to $E_{\mathrm{T}}^{\mathrm{miss}}$ to different MC tunes (left) and different generators (right) for the combined channels. The error bars are divided into systematic (inner band) and statistical (outer band) uncertainties.

Measured normalised top-quark pair cross section for various bins in $E_{\mathrm{T}}^{\mathrm{miss}}$ (GeV^{-1})	
0–25 GeV	$(0.60 \pm 0.03) \cdot 10^{-2} (4.65\%)$
25–45 GeV	$(1.35 \pm 0.04) \cdot 10^{-2} (2.99\%)$
45–70 GeV	$(1.15 \pm 0.03) \cdot 10^{-2} (2.96\%)$
70–100 GeV	$(0.58 \pm 0.03) \cdot 10^{-2} (4.67\%)$
100–150 GeV	$(0.18 \pm 0.01) \cdot 10^{-2} (6.68\%)$
$\geq$ 150 GeV	$(0.03 \pm 0.00) \cdot 10^{-2} (7.66\%)$
Measured normalised top-quark pair cross section for various bins in H_{T} (GeV^{-1})	
0–240 GeV	$(0.24 \pm 0.02) \cdot 10^{-2} (7.78\%)$
240–280 GeV	$(0.41 \pm 0.02) \cdot 10^{-2} (4.61\%)$
280–330 GeV	$(0.31 \pm 0.01) \cdot 10^{-2} (4.38\%)$
330–380 GeV	$(0.21 \pm 0.01) \cdot 10^{-2} (6.55\%)$
380–450 GeV	$(0.13 \pm 0.01) \cdot 10^{-2} (8.75\%)$
450–600 GeV	$(0.05 \pm 0.01) \cdot 10^{-2} (11.74\%)$
$\geq$ 600 GeV	$(0.01 \pm 0.00) \cdot 10^{-2} (11.68\%)$
Measured normalised top-quark pair cross section for various bins in S_{T} (GeV^{-1})	
0–350 GeV	$(0.17 \pm 0.01) \cdot 10^{-2} (6.45\%)$
350–400 GeV	$(0.34 \pm 0.02) \cdot 10^{-2} (4.65\%)$
400–450 GeV	$(0.26 \pm 0.01) \cdot 10^{-2} (4.29\%)$
450–500 GeV	$(0.18 \pm 0.01) \cdot 10^{-2} (6.02\%)$
500–580 GeV	$(0.11 \pm 0.01) \cdot 10^{-2} (7.92\%)$
580–700 GeV	$(0.05 \pm 0.01) \cdot 10^{-2} (10.25\%)$
$\geq$ 700 GeV	$(0.02 \pm 0.00) \cdot 10^{-2} (10.95\%)$
Measured normalised top-quark pair cross section for various bins in p_{T}^{W} (GeV^{-1})	
0–40 GeV	$(0.49 \pm 0.02) \cdot 10^{-2} (3.62\%)$
40–70 GeV	$(0.91 \pm 0.02) \cdot 10^{-2} (2.53\%)$
70–100 GeV	$(0.77 \pm 0.02) \cdot 10^{-2} (2.05\%)$
100–130 GeV	$(0.49 \pm 0.02) \cdot 10^{-2} (4.30\%)$
130–170 GeV	$(0.23 \pm 0.01) \cdot 10^{-2} (3.38\%)$
$\geq$ 170 GeV	$(0.08 \pm 0.00) \cdot 10^{-2} (5.38\%)$
Measured normalised top-quark pair cross section for various bins in M_{T}^{W} (GeV^{-1})	
0–30 GeV	$(0.45 \pm 0.06) \cdot 10^{-2} (12.26\%)$
30–50 GeV	$(0.79 \pm 0.06) \cdot 10^{-2} (7.92\%)$
50–80 GeV	$(1.81 \pm 0.07) \cdot 10^{-2} (3.63\%)$
80–100 GeV	$(0.72 \pm 0.05) \cdot 10^{-2} (7.63\%)$
$\geq$ 100 GeV	$(0.02 \pm 0.00) \cdot 10^{-2} (14.40\%)$

Table 7.8: Measured normalised top-quark pair cross section for the different variables. The values in brackets show the total uncertainty (stat.+syst.) on the measurement as a percentage.

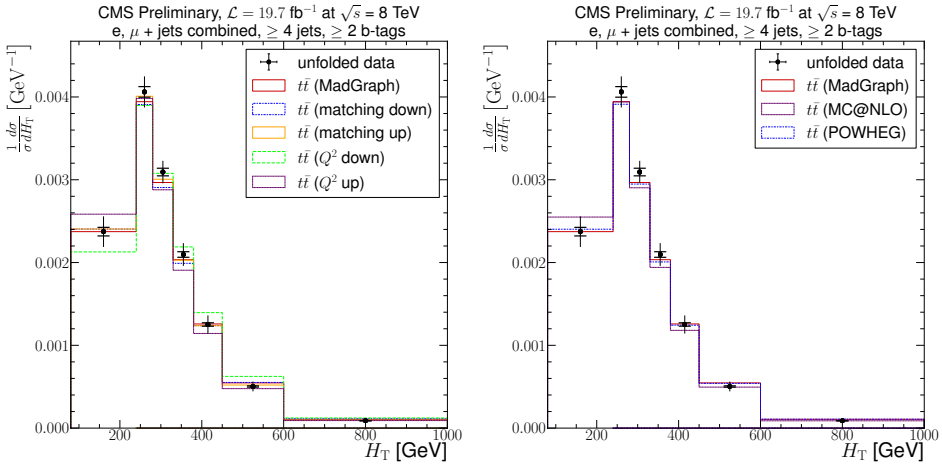

Figure 7.7: Comparison of measured normalised differential cross section with respect to H_T to different MC tunes (left) and different generators (right) for the combined channels. The error bars are divided into systematic (inner band) and statistical (outer band) uncertainties.

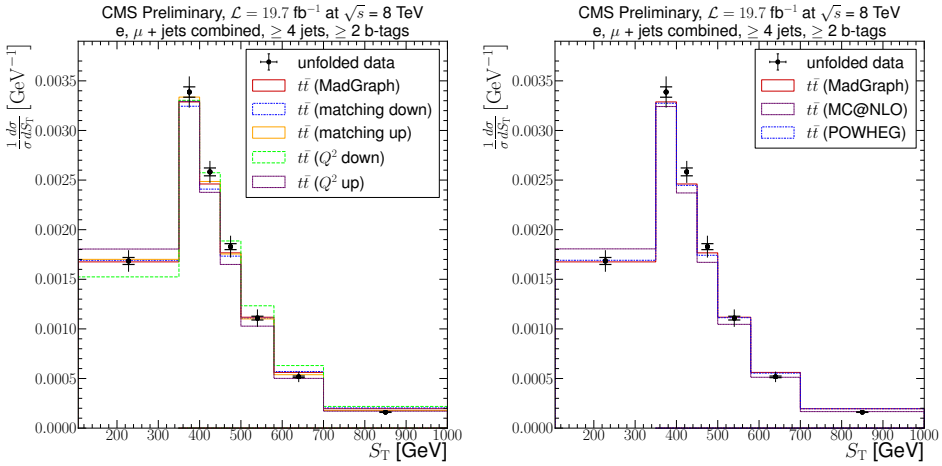

Figure 7.8: Comparison of measured normalised differential cross section with respect to S_T to different MC tunes (left) and different generators (right) for the combined channels. The error bars are divided into systematic (inner band) and statistical (outer band) uncertainties.

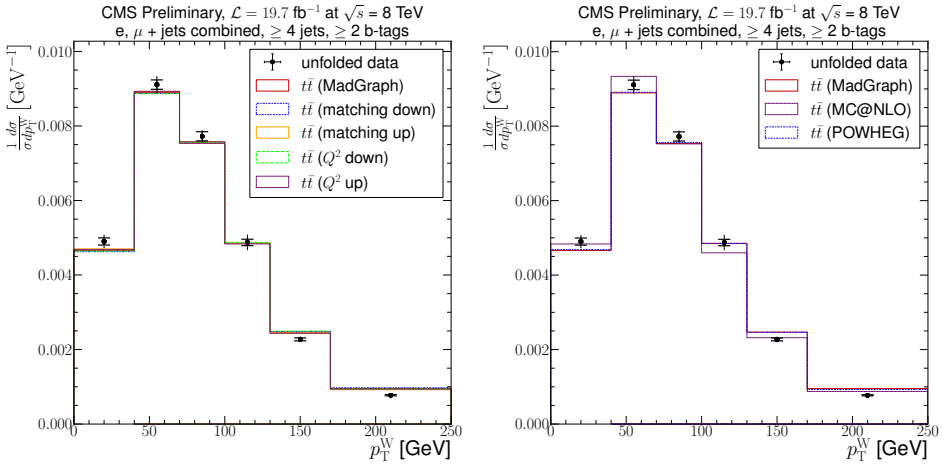

Figure 7.9: Comparison of measured normalised differential cross section with respect to p_{T}^{W} to different MC tunes (left) and different generators (right) for the combined channels. The error bars are divided into systematic (inner band) and statistical (outer band) uncertainties.

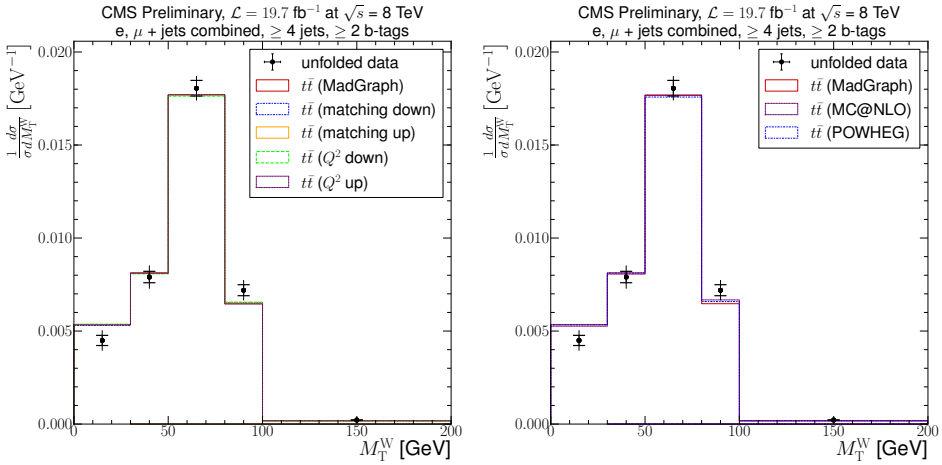

Figure 7.10: Comparison of measured normalised differential cross section with respect to M_{T}^{W} to different MC tunes (left) and different generators (right) for the combined channels. The error bars are divided into systematic (inner band) and statistical (outer band) uncertainties.

7.8 Summary

In this chapter the process of extracting differential cross sections with respect to global event-level variables has been described. The variables that have been studied are $E_{\mathrm{T}}^{\mathrm{miss}}$, H_{T}, S_{T}, p_{T}^{W} and M_{T}^{W}. The analysis involves optimising the binning for each variable, extracting the number of $t\bar{t}$ events in each bin using a maximum likelihood fit and then unfolding the true number of events in both the electron and muon-plus-jets channels to correct for selection and detector effects. The results in the two channels are then combined and the final results are presented as a normalised differential cross section. The uncertainties on the measured results will be detailed in the next chapter.

Chapter 8

Evaluation of Uncertainties on Measurements

There are various sources of uncertainties that can have an effect on the measurements. Uncertainties can be statistical by nature or can arise from systematic effects. Systematic uncertainties can be characterised as either experimental or theoretical. Experimental uncertainties originate from finite detector resolution in terms of both energy and momentum measurements of particles and jets, as well as efficiencies in identifying a particular particle type correctly. Theoretical systematic uncertainties are present due to theoretical parameters that are input into the Monte Carlo used to model the signal and background processes.

In the case of the inclusive cross section, uncertainties can affect the shapes of the $|\eta_\mu|$ distributions that are used in the fitting procedure and the efficiency of $t\bar{t}$ events, $\epsilon_{t\bar{t}}$, passing the final event selection. In order to assess a particular source of uncertainty, the $\pm 1\sigma$ systematics variations on both the template shape and $\epsilon_{t\bar{t}}$ are used as input. The effect that this variation has on the final result is then quantified as:

$$\Delta\sigma_{t\bar{t}}^{\pm} = \sigma_{t\bar{t}}^{\text{syst.}\pm} - \sigma_{t\bar{t}}^{\text{nominal}}. \tag{8.1}$$

Both $\sigma_{t\bar{t}}^{\text{syst.}\pm}$ and $\sigma_{t\bar{t}}^{\text{nominal}}$ are calculated using Equation 6.15, with $\sigma_{t\bar{t}}^{\text{nominal}}$ being the central result and $\sigma_{t\bar{t}}^{\text{syst.}\pm}$ the corresponding result obtained when

the appropriate variation is made. The $\Delta\sigma^{+}_{t\bar{t}}$ and $\Delta\sigma^{-}_{t\bar{t}}$ variations are not necessarily symmetric about $\sigma_{t\bar{t}}$ and so are assessed separately.

The uncertainties related to the differential cross section measurements are assessed by repeating the fitting and unfolding procedures using the $\pm 1\sigma$ variations as input. This involves changing the input $|\eta_{\mu}|$ templates and $\epsilon_{t\bar{t}}$ for the experimental uncertainties and the response matrix used for the unfolding for theoretical uncertainties. It is not necessary to change both the templates and the response matrix as this would result in double counting in the estimation of the uncertainty. Equation 7.22 is then used to calculate the result in the same way as for the nominal case. The difference from the nominal result is taken as the uncertainty. The normalisation of the differential cross section means that uncertainties only effecting normalisation, such as luminosity, cancel. Cancellations will also occur when uncertainties are positively correlated between the bins of a given distribution.

In this chapter, all of the sources of uncertainties will first be described. Each will be assessed individually and for some of the key systematics, the effect on $\epsilon_{t\bar{t}}$ and the shape of the $|\eta_{\mu}|$ templates will be investigated. To obtain the total uncertainty on a measurement, the correlations between all of the individual uncertainties must be taken into account. In the 7 TeV cross section analysis, this was done by performing pseudo-experiments using random numbers to simulate the variations on each individual uncertainty. This method was chosen for this analysis because it is able to take account of the uncertainties on the background templates by smearing the shapes within their uncertainties. These pseudo-experiments are used in a Neyman construction scheme which was first used in reference [109]. This removes any bias on the cross section measurement by using a range of alternate $t\bar{t}$ cross sections values as input. More details on the pseudo-experiments and the Neyman construction will be given in Section 8.4.

After the comprehensive study on the uncertainties in the 7 TeV cross section analysis, the uncertainties in the 8 TeV analysis are determined via a more conventional method. This involves passing each systematic variation of fit templates through the fitting procedure whilst using the systematic variation on $\epsilon_{t\bar{t}}$ in the calculation of $\sigma_{\mathrm{t\bar{t}}}$. The uncertainties from the different

sources are summed in quadrature to calculate the total systematic uncertainty on the result. Tables containing all of the systematic uncertainties for the various analyses are presented towards the end of this Chapter.

8.1 Statistical Uncertainties

The statistical error in each bin of the $|\eta_\mu|$ distribution that enters the fit is given by the Poisson standard deviation, i.e. the square root of the number of events in that bin. Statistical errors also arise from the statistical procedures that are used to extract the results such as in the maximum likelihood estimate and the unfolding. In both of these cases the statistical uncertainty is the square root of the $t\bar{t}$ component from the diagonal of the covariance matrix.

8.2 Experimental Systematic Uncertainties

8.2.1 Jet Energy Scale (JES) and Resolution (JER)

The uncertainty on the Jet Energy Scale (JES) potentially has an impact on the shapes of both the $t\bar{t}$ and single-top templates, as well as on the selection efficiencies for the two processes. To understand this effect, the four-momenta of each jet is varied simultaneously by one standard deviation, $\pm 1\sigma$. Jet energy uncertainties are related to the errors on the jet η- and p_T-dependent correction factors as was mentioned in Section 5.7.4.

The effect of varying the JES on the muon rapidity distribution is studied for both $t\bar{t}$ and single-top processes. The relative shape difference between the nominal JES and the varied values is shown in Figure 8.1. For both the $t\bar{t}$ and single-top the differences are well below 1%. The uncertainty is actually driven by the limited statistics of the Monte Carlo samples as can be seen from the fact that the points are all within the shaded region. The effect of the JES uncertainty on the signal template shapes can therefore be neglected.

When assessing the effect of the JES on the selection efficiencies, its effect can be parameterised using the strength-parameter, $\delta_{JES} \in [-1, +1]$, where a strength-parameter $\delta_{JES} = \pm 1$ corresponds to the $\pm 1\sigma$ variation of the jet four-momenta. The effects of varying the JES on the selection efficiencies for $t\bar{t}$ and single-top are shown in Figure 8.2. The three points show the efficiency ratios $(\epsilon(\delta)/\epsilon(0))$ to the nominal efficiency. This ratio is parameterised using a linear function that gives a good description of the efficiency ratio in this region. The resulting functional dependencies for $t\bar{t}$ and single-top are:

$$\frac{\epsilon(\delta^{t\bar{t}}_{JES})}{\epsilon^{t\bar{t}}(0)} = 0.060\delta_{JES} + 1 \tag{8.2}$$

$$\frac{\epsilon^{stop}(\delta_{JES})}{\epsilon^{stop}(0)} = 0.121\delta_{JES} + 1 \tag{8.3}$$

The linear parameterisation is required such that values of $\epsilon_{t\bar{t}}$ can be generated and entered into pseudo-experiments later. The JES has roughly double the effect on the single-top efficiency as it does for $t\bar{t}$. This could be because there are fewer jets produced in single-top production so it is closer to the three jet threshold required in this study.

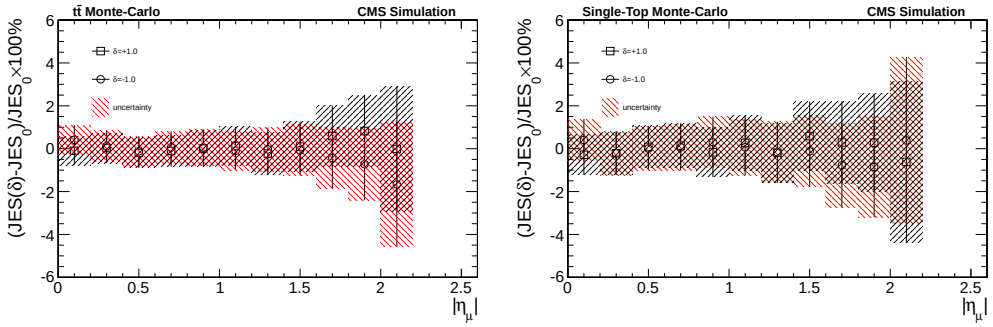

Figure 8.1: The relative shape uncertainty due to JES variations for $t\bar{t}$ (left) and single-top (right) from Monte Carlo predictions. The red hashed box represents the error on the down variation and the black hashed box on the up variation. The two shapes are within statistical error of one another. The errors on the points are due to the statistical error on the MC sample.

The JER uncertainty arises due to the uncertainty on the amount of

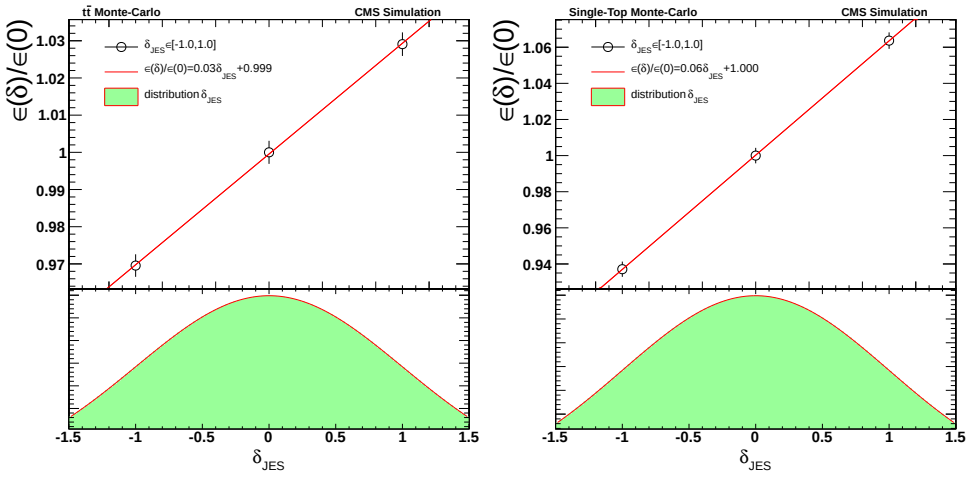

Figure 8.2: The relative uncertainty on the efficiency due to JES variations are shown for $t\bar{t}$(left) and single-top (right) using Monte Carlo predictions. In both cases the variation is parametrised using a linear fit. The Guassian at the bottom shows distribution of how $\epsilon_{t\bar{t}}$ will be generated in pseudo-experiments. The errors on the points are due to the statistical error on the MC sample.

smearing applied to the MC in order to make it consistent with the data. Jet asymmetry measurements [92] suggest that the jet-p_{T} resolution in data is about 10% worse than in MC. Smearing therefore needs to be applied to the jet four-momenta for the MC. This smearing effect has very little effect on the p_{T} of the jets and so the systematic uncertainty is very small.

8.2.2 $E_{\mathrm{T}}^{\mathrm{miss}}$

The $E_{\mathrm{T}}^{\mathrm{miss}}$ uncertainty is only considered for differential cross sections where $E_{\mathrm{T}}^{\mathrm{miss}}$ is used, that is, all of the differential variables except H_{T}. The uncertainties directly connected to the $E_{\mathrm{T}}^{\mathrm{miss}}$ measurement are evaluated by changing the energy and momentum of all objects which are used to calculate the $E_{\mathrm{T}}^{\mathrm{miss}}$. The main source of uncertainty is that propagated from the error on the JES and has roughly a 15% effect on the calculation of the $E_{\mathrm{T}}^{\mathrm{miss}}$. Other things that can effect the $E_{\mathrm{T}}^{\mathrm{miss}}$ include:

- Unclustered energy (energy that in an event from detector noise): $\pm 10\%$
- Tau energy (since semi-leptonic tau decays can sometimes be included these include additional $E_{\mathrm{T}}^{\mathrm{miss}}$ from the neutrinos): $\pm 3\%$
- Electron & photon energy: $\pm 0.6\%$ in EB and $\pm 1.5\%$ in EE
- Muon momentum: $\pm 0.2\%$

The percentages are the uncertainty propagated to the measurement of the $E_{\mathrm{T}}^{\mathrm{miss}}$ from the various sources. These values are provided by the dedicated physics object groups within CMS.

8.2.3 Luminosity

The luminosity has been measured for the both the 7 and 8 TeV physics runs. The luminosity enters the cross section calculation and has an effect on the global event rate. The uncertainty on the luminosity amounts to 2.2% [110]. The uncertainty completely cancels out for the normalised differential cross section measurements, but must be included for the inclusive cross section.

8.2.4 Pileup

Two sources of uncertainty need to be taken into account when assigning the pileup weights. These are the luminosity uncertainty, which is estimated to be 2.2% [110] and the uncertainty on the total inelastic cross section. An inelastic cross section, σ_{pp}, value of 68 mb was obtained using forward calorimetry with the 2011 data [111]. These values (luminosity + cross section) have a combined uncertainty of 3.6%. The PU weights were reproduced with $\pm 1\sigma$ variations on the pp inelastic cross sections and the luminosity. The impact of using the lower and upper bounds on the inelastic pp cross section and the luminosity can be seen in Figure 8.3. Pileup reweighting has little effect on the η distribution of the muon and so is not an major source of uncertainty in this analysis.

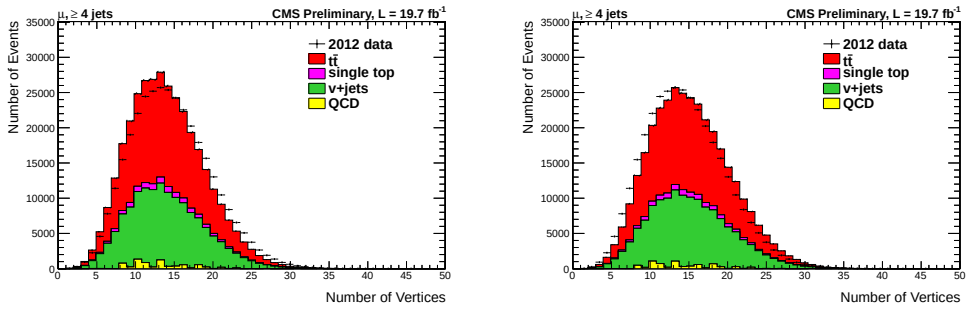

Figure 8.3: Number of reconstructed vertices for the -1σ (left) and the $+1\sigma$ (right) variations of the pileup reweighting procedure.

8.2.5 Lepton Identification, Isolation and Trigger Efficiencies

The uncertainties on the scale factors for the muon identification, isolation and trigger have been considered. These scale factors are provided by the CMS muon physics object group using the tag and probe method [86]. The scale factors are dependent on the η of the muon. They therefore have an effect on the template shapes that are used in the fitting procedure. Muon SFs (combined trigger + identification + isolation) have an overall uncertainty of 1%. The uncertainties on the electron scale factors (combined trigger +

identification + isolation) are also considered when combining in the two channels. The electron scale factors are flat in η and have uncertainties of typically 3%.

8.2.6 B-tagging

The b-tagging scale factors are applied as was described in Section 5.7.3. Since b-tagging was only implemented in the 8 TeV analysis the uncertainty is only considered for the measurements based on this data set. The uncertainties on the scale factors are dependent on the flavour of the parton that initiates the jet and its kinematics. The uncertainties are summarised as follows:

- b-flavoured jets with $p_T < 800$ GeV have p_{T}-dependent uncertainties in the range 1.6-8%;
- c-flavoured jets with $p_T < 800$ GeV have twice the uncertainty as that of b-flavoured jets;
- c- and b-flavoured jets with $p_T > 800$ GeV again have twice the uncertainty of those with $p_T < 800$ GeV. Uncertainties are fully correlated with b-tagging uncertainties);
- mis-tagging (or b-tagging of light-flavoured jets: u, d, s) occurs for roughly 1% of jets. These uncertainties are uncorrelated with b-tag uncertainties for c/b-flavoured jets and are again p_{T}-dependent in the range 6-10%.

8.3 Theoretical and Modelling Uncertainties

Some fundamental parameters are used as input in the MC generators. To take account of the theoretical uncertainties due to modelling effects, a different MC generator tune is used as input and the result from this tune is compared to the central result (using the nominal MC tune).

8.3.1 Renormalisation and Factorisation Scale

The renormalisation (μ_R) and factorisation (μ_F) scale which is set to be equal to the momentum transfer squared (Q^2) required for top pair production (see Chapter 2) is a key input parameter in MC generators. The nominal choice for Q^2 is to set it equal to the mass of the top ($m_t = 173$ GeV). In order to estimate the uncertainty on this choice of Q^2, MC samples have been simulated with $2Q^2$ (up) and $0.5Q^2$ (down). These samples are then used to evaluate the theoretical uncertainty. The renormalisation and factorisation scales are both varied together.

The effect of varying the scale Q^2 on the muon pseudorapidity shapes and selection efficiency is shown in Figure 8.4. The relative differences for each bin with respect to the default scale Q_0^2 are displayed and the uncertainty due to statistical limitations is represented by the coloured bands. Most values are consistent with each other within uncertainties. To assess the effect of the Q^2 scale on the selection efficiency, a linear parameterisation is performed as was done for the JES (again for later use in pseudo-experiments). A strength-parameter, δ_{Q^2}, is defined such that:

$$P(\delta_{Q^2}) = Q_0^2 \cdot e^{\delta_{Q^2}} \tag{8.4}$$

where $\delta_{Q^2} \in \{\log 0.5, \log 2\}$. This has the physical meaning of varying the Q^2 parameter in the range $[Q^2/2, 2Q^2]$. The effect on the selection efficiency is parameterised using a linear fit to give

$$\frac{\epsilon(\delta_{Q^2}^{t\bar{t}})}{\epsilon^{t\bar{t}}(0)} = -0.276\delta_{Q^2} + 0.998 \tag{8.5}$$

This systematic also had to be included in the 8 TeV analyses, where the W+jets and Z+jets backgrounds were taken from MC. In both of these cases the nominal Q^2 scale is set to the mass of the boson being produced. The same variations as for the $t\bar{t}$ case are used.

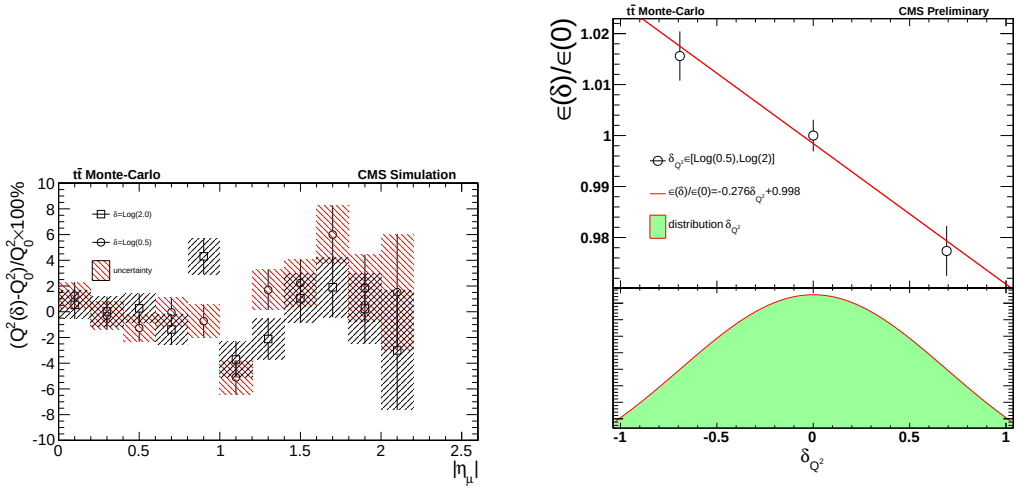

Figure 8.4: Relative uncertainty on the shape (left) and efficiency (right) due to variations of the Q^2 choice for the $t\bar{t}$ process. The red hashed box represents the error on the down variation and the black hashed box on the up variation. The errors on the points are due to the statistical error on the MC sample.

8.3.2 Jet-Parton Matching Threshold

Another fundamental input parameter into the MC is the matching threshold (th). This parameter is the parton p_{T} required to initiate a shower. The nominal choice of the matching threshold for the $t\bar{t}$ MC is 20 GeV. The "up" and "down" simulated samples have a matching threshold of 10 and 40 GeV, respectively.

To estimate the systematic uncertainties due to the choice of the matching threshold, the effect on the muon pseudorapidity shape and the selection efficiency of the $t\bar{t}$ process have again been investigated.

A similar procedure as for the Q^2 scale and JES was performed. The effect of varying the matching threshold on the $|\eta_\mu|$ distribution and the selection efficiency is shown in Figure 8.5. In contrast to the JES and Q^2 studies, a clear difference is observed between the $|\eta_\mu|$ shapes for the matching up and down samples. The right hand Figure shows that there is no linear representation of the selection efficiency with the matching threshold input.

Again, for the 8 TeV analyses the W+jets and Z+jets matching threshold must be assessed. The nominal matching threshold for these sample is 10 GeV and the "up" and "down" samples are simulated with thresholds of 5 and 20 GeV, respectively.

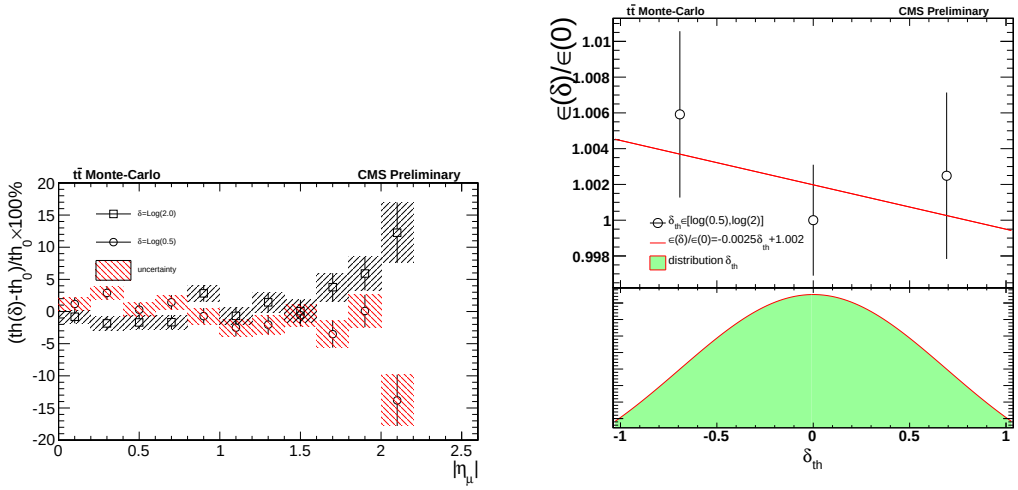

Figure 8.5: Relative uncertainty on the shape (left) and selection efficiency (right) due to variations of the matching threshold for the $t\bar{t}$ process using MADGRAPH MC. The red hashed box represents the error on the down variation and the black hashed box on the up variation. The errors on the points are due to the statistical error on the MC sample.

8.3.3 Hadronisation Modelling

A hadronisation uncertainty is present due to the choice of using MADGRAPH+PYTHIA to obtain the nominal result. PYTHIA uses a different hadronisation model to HERWIG as was described in Chapter 2. It is therefore necessary to account for this choice of modelling. To do this the results are obtained using both response matricies in the unfolding. The difference between the two results is taken as the uncertainty.

8.3.4 Single-Top Production

Since the $t\bar{t}$ and single-top templates are very similar they have to be fit together as a combined template (signal). It is therefore necessary to subtract the single-top contribution from the signal using the number of single-top events predicted by MC, normalised using the theoretical production cross sections given in Table 6.4. A pessimistic uncertainty of 30% on the single-top theoretical cross section is recommended by the CMS top quark group. In the differential cross section measurements this uncertainty largely cancels in the normalisation of the cross section.

8.4 Uncertainties on 7 TeV Inclusive Cross Section Measurement

The 7 TeV inclusive cross section was the first time this analysis had been performed and so it was important to make sure that there were no sources of bias. This was the reason for performing the full Neyman construction from pseudo-experiments.

In the 7 TeV analysis all of the fit templates and constraints on the normalization for the background processes are extracted from data and some theoretical predictions. Therefore, the only uncertainties to be taken into account on the background estimates are due to statistical limitations and uncertainties on the theoretical numbers used to extract the template. The backgrounds were extracted using different methods in various signal depleted phase spaces and so the uncertainties are completely uncorrelated.

The treatment of the background uncertainties for the 7 TeV analysis are summarised here for each process:

1. W+Jets background

    - Template shape: Each bin has an uncertainty defined by the statistical limitations of the charge subtraction method explained in Section 6.3.1 and the uncertainty on the correction factors used

to extrapolate from the charge subtraction region to the signal region. The relative uncertainties per $|\eta_\mu|$ bin typically range from $5\% - 10\%$.

2. Drell-Yan background

   - Template shape: The bin-wise uncertainties due to limited statistics and correction factors (again from the extrapolation into the signal region) are used. The uncertainty in each $|\eta_\mu|$ bin is typically around 10%.

3. QCD multi-jet background

   - Template shape: The only uncertainty is the statistical uncertainty for each bin. The typical range is $6\% - 14\%$.
   - Normalization: A 100% uncertainty on the number of events predicted by QCD MC is applied.

In the 8 TeV analyses the theoretical uncertainties on the Q^2 scale and the matching threshold described in the previous section have been considered.

8.4.1 Construction of Pseudo-Experiments

Pseudo-experiments using pseudo-data derived from MC is produced in order to assess uncertainties on the measurement. Distributions are generated for each sub-process (signal, W+jets, Z+jets and QCD) independently and then added to obtain the final observed pseudo-data. The construction of each background $|\eta_\mu|$ distribution follows the procedure below:

1. For each bin of the distribution a normally distributed random number is generated with a mean set to the expected number of events and width corresponding to the uncertainty in that $|\eta_\mu|$ bin.

2. In each $|\eta_\mu|$ bin a Poissonian random number is then generated about the mean from step 1. Step 2 represents the statistical uncertainty in each bin.

The construction of the signal pseudo-data is a little more involved since each source of systematic uncertainty has to be taken into account, all of which will have an effect on the fitted cross section result. Strength-parameters such as those defined for JES and the Q^2 uncertainties are used. These allow for the signal pseudo-data to be produced within $\pm 1\sigma$ bands for each uncertainty. This is performed for each source of systematic uncertainty defined previously in this chapter. The variation of each strength-parameter, δ_i, can have an impact on both the $t\bar{t}$ and single-top selection efficiencies.

The construction of the signal pseudo-data can now proceed as follows:

1. A strength-parameter, δ_i, is simulated for each source of systematic, i, using a random normal variable about the mean expected number of events with a width corresponding to its uncertainty.

2. These strength-parameters are combined to produce efficiencies for both the $t\bar{t}$ and single-top processes.

3. Normally distributed random numbers are simulated within uncertainties for systematic uncertainties such as the matching threshold where the $|\eta_\mu|$ shape is significantly different.

4. The expected number of $t\bar{t}$ and single-top events are also computed using
$$N = \mathcal{L} \times \sigma \times \epsilon \tag{8.6}$$

5. In each $|\eta_\mu|$ bin a Poisson distributed random number with mean constructed in the previous step is generated. This step again relates to the statistical error in each bin.

The final pseudo-data distributions are then the simple sum of the distributions of all sub-processes. These distributions are then fitted using the maximum likelihood method as explained in Section 6.4 to extract the result.

8.4.2 Full Uncertainty on the Cross Section: Neyman Construction

The fitted $t\bar{t}$ cross section is computed from the fitted number of events N^{fit} as was described in Section 6.4.1. The fitted result can be related to the true $t\bar{t}$ cross section via a Neyman construction scheme. In the process of this Neyman construction the following parameter is defined as

$$\beta_{t\bar{t}} = \frac{\sigma_{t\bar{t}}}{\sigma_{t\bar{t}}^{\text{theory}}}. \tag{8.7}$$

$\sigma_{t\bar{t}}$ is varied in a range of values such that $\beta_{t\bar{t}} \in [0.2, 0.5, 0.7, 0.9, 1.0, 1.2, 1.4, 1.8, 2.0]$. An ensemble of 100K pseudo-experiments are performed for each value of $\beta_{t\bar{t}}$. In each pseudo-experiment, pseudo-data is used to extract a fitted cross section result, σ^{fit}. A value or $\beta_{t\bar{t}}^{\text{fit}}$ is obtained from each pseudo-experiment, which is in turn defined as

$$\beta_{t\bar{t}}^{\text{fit}} = \frac{\sigma_{t\bar{t}}^{\text{fit}}}{\sigma_{t\bar{t}}^{\text{theory}}}. \tag{8.8}$$

These pseudo-experiments yield an approximately normal probability distribution, $P(\beta)$, related to the variable $\beta_{t\bar{t}}^{\text{fit}}$ for each $\beta_{t\bar{t}}$ value. The central value $\beta_{t\bar{t}}^{\text{fit}}(0)$ is defined as the mean of the probability distribution. The upper and lower 1σ (and 2σ) bands can be defined as $\beta_{t\bar{t}}^{\text{fit}}(\pm n\sigma)$, with $n = 1, 2$. With this in mind the following equations are satisfied:

$$1 - \frac{p_{n\sigma}}{2} = \int_{\beta_{t\bar{t}}^{\text{fit}}(+n\sigma)}^{\infty} P(\beta)d\beta = \int_{-\infty}^{\beta_{t\bar{t}}^{\text{fit}}(-n\sigma)} P(\beta)d\beta, \quad n = 1, 2 \tag{8.9}$$

with the 1σ and 2σ probabilities are defined where $p_{1\sigma} = 0.68$ and $p_{2\sigma} = 0.95$, respectively. The results of the pseudo-experiments are summarized in Table 8.1.

The relationship between $\beta_{t\bar{t}}^{\text{fit}}$ and $\beta_{t\bar{t}}$ follows a linear relationship (central values, as well as for the error bands). These dependencies can be seen in Figure 8.6. Linear functional dependencies are therefore constructed for the mapping of $\beta_{t\bar{t}}^{\text{fit}} \rightarrow \left(\beta_{t\bar{t}} + \Delta\sigma_{t\bar{t}}^{+} - \Delta\sigma_{t\bar{t}}^{-}\right)$ in order to obtain the cross section

		$\beta_{t\bar{t}}$								
		0.2	0.5	0.7	0.9	1.0	1.2	1.4	1.8	2.0
$\beta_{t\bar{t}}^{\text{fit}}$	-2σ	-0.045	0.207	0.369	0.515	0.591	0.735	0.881	1.16	1.29
	$-\sigma$	0.077	0.353	0.527	0.699	0.785	0.953	1.12	1.45	1.6
	central(0)	0.199	0.499	0.697	0.897	0.997	1.2	1.4	1.79	1.95
	$+\sigma$	0.319	0.645	0.869	1.1	1.21	1.45	1.68	2.13	2.27
	$+2\sigma$	0.441	0.791	1.03	1.29	1.41	1.67	1.93	2.39	2.46

Table 8.1: $\beta_{t\bar{t}}^{\text{fit}}$ values for several input values of $\beta_{t\bar{t}}$.

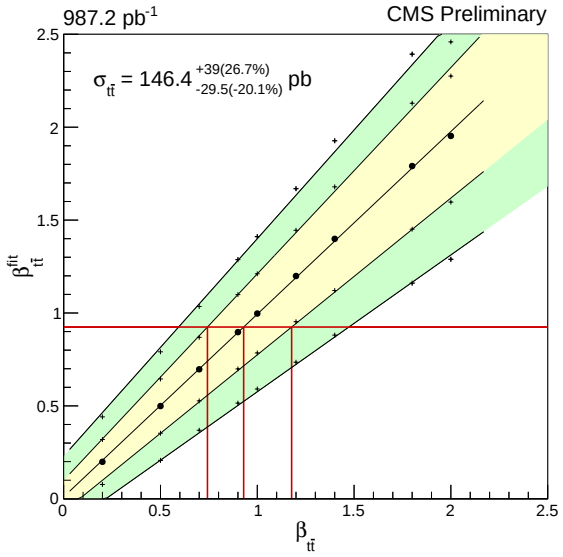

Figure 8.6: Functional dependence between $\beta_{t\bar{t}}^{\text{fit}}$ and $\beta_{t\bar{t}}$ with corresponding 1σ (yellow) and 2σ (green) uncertainty. The red lines correspond to the Neyman construction for the measured $t\bar{t}$ cross section.

and the uncertainties:

$$\sigma_{t\bar{t}} = \sigma_{t\bar{t}}^{\text{theory}} \times \left(\beta_{t\bar{t}} + \Delta\sigma_{t\bar{t}}^{+} - \Delta\sigma_{t\bar{t}}^{-}\right). \tag{8.10}$$

The fitted cross section $\sigma_{t\bar{t}}^{\text{fit}}$ using the nominal values is used to extrapolate across Figure 8.6 to obtain the final result. A value of $\beta_{t\bar{t}} = 0.93$ was found. This gives a final result for the $t\bar{t}$ cross section of $146.4^{+39.0(26.7\%)}_{-29.5(-20.1\%)}$ pb at $\sqrt{s} = 7$ TeV.

8.4.3 Breakdown of Uncertainties

Individual uncertainties can be assessed by only allowing one source of uncertainty at a time to influence the construction of the pseudo-data in experiments. The same mapping can then be performed as was described for the full uncertainty. The breakdown of the contributions from the individual uncertainties for the 7 TeV inclusive cross section are shown in Table 8.2. The dominant sources are those relating to the statistical uncertainties on the W+jets template and the Q^2 scale uncertainty. The systematic uncertainties on the background templates will improve with increased statistics but will eventually be limited by the uncertainties assigned to the correction factors. This limitation is almost reached for the 1 fb^{-1} of 7 TeV data set as was used in this analysis.

Source	$\beta_{t\bar{t}}(-\sigma)$ pb (%)	$\beta_{t\bar{t}}(0)$	$\beta_{t\bar{t}}(+\sigma)$ pb (%)
W+jets Temp stat.	-15.8 (-10.9%)	0.924	15.9 (11%)
DY Temp stat.	-2.9 (-2.01%)	0.924	2.99 (2.06%)
QCD Temp stat.	-0.4 (-0.247%)	0.924	0.315 (0.216%)
All backgrounds stat.	-16.2 (-11.1%)	0.924	16.2 (11.2%)
JES	-4.2 (-2.88%)	0.923	4.34 (2.98%)
Q^2 scale	-22.3 (-15.3%)	0.926	32 (22%)
matching threshold	-0.2 (-0.168%)	0.922	0.291 (0.2%)
luminosity	-3.5 (-2.41%)	0.923	3.69 (2.54%)
single top xsect.	-5.2 (-3.55%)	0.923	5.06 (3.48%)
stat.	-7.2 (-4.97%)	0.924	7.46 (5.12%)
syst. $\oplus$ stat.	-29.5 (-20.1%)	0.93	39 (26.7%)

Table 8.2: Breakdown of systematic and statistical uncertainties for the 7 TeV inclusive cross section analysis.

8.5 Uncertainties on 8 TeV Inclusive Cross Section Measurements

The comprehensive study of the uncertainties for the 7 TeV cross section measurement showed only a negligible bias on the final result in comparison

to the uncertainty on the measurement. It is therefore sufficient to simply repeat the analysis with the various systematic shifts as input for the 8 TeV analyses. This involves using systematically shifted event weights to fill histograms or using different systematic samples as has been described earlier in this Chapter. A breakdown of the uncertainties is shown in Table 8.3.

Source	$-\sigma$ pb (%)	$+\sigma$ pb (%)
JES	-7.4 (-3.1%)	10.1 (4.3%)
JER	-0.7 (-0.3%)	0.6 (0.2%)
PU	-1.5 (-0.6%)	2.1 (0.9%)
B-tag	-10.9 (-4.6%)	12.1 (5.1%)
lumi.	-5.1 (-2.2%)	5.3 (2.2%)
Q^2 scale (V+jets)	-6.4 (-2.7%)	10.4 (4.4%)
Matching threshold (V+jets)	-2.5 (-1.1%)	2.5 (1.1%)
QCD	-4.2 (-1.8%)	2.8 (1.2%)
Q^2 scale ($t\bar{t}$)	-0.4 (-0.2%)	0.7 (0.3%)
Matching threshold ($t\bar{t}$)	-2.5 (-1.0%)	0.1 (0.0%)
single top xsect.	-3.0 (-1.3%)	3.0 (1.3%)
stat.	-1.9 (-0.8%)	1.9 (0.8%)
syst. $\oplus$ stat.	-16.9 (7.1%)	20.4 (8.6%)

Table 8.3: Breakdown of systematic and statistical uncertainties for 8 TeV inclusive cross section measurement. The value in brackets is the percentage of the uncertainty on the result.

8.6 Uncertainties on Differential Cross Section Measurements

The systematic uncertainties for the differential cross section measurements are estimated in a similar way to those in the 8 TeV inclusive cross section. The main difference is that for the experimental uncertainties the error is also propagated through the unfolding procedure. The theoretical uncertainties are assessed by replacing the response matrix in the unfolding with that as-

sociated with the prescribed uncertainty. All of the systematic uncertainties for the $E_{\mathrm{T}}^{\mathrm{miss}}$ variable are summarised in Table 8.4 for the combined (e and μ+jets) measurement. Tables for the other variables can be seen in Appendix D.

Systematic Errors						
Systematic	$0< E_{\mathrm{T}}^{\mathrm{miss}} <25$	$25\leq E_{\mathrm{T}}^{\mathrm{miss}} <45$	$45\leq E_{\mathrm{T}}^{\mathrm{miss}} <75$	$75\leq E_{\mathrm{T}}^{\mathrm{miss}} <100$	$100\leq E_{\mathrm{T}}^{\mathrm{miss}} <150$	$E_{\mathrm{T}}^{\mathrm{miss}} \geq 150$
b-jets - (%)	0.12	0.03	-0.06	-0.05	-0.01	0.02
b-jets + (%)	-0.11	-0.03	0.06	0.04	-0.00	-0.02
JER - (%)	0.03	0.03	-0.04	-0.01	0.02	0.06
JER + (%)	-0.07	0.04	0.03	-0.03	-0.04	0.01
JES - (%)	1.80	1.22	0.10	-1.62	-2.81	-3.22
JES + (%)	-2.09	-1.21	0.20	1.67	2.38	2.83
Light jet - (%)	0.05	0.01	-0.04	-0.02	0.03	0.03
Light jet + (%)	-0.05	-0.01	0.04	0.01	-0.04	-0.03
Pile-up - (%)	-0.05	0.03	0.00	-0.01	0.01	-0.04
Pile-up + (%)	0.14	-0.04	-0.04	0.01	-0.00	0.04
QCD shape uncertainty (%)	-1.47	-0.24	0.53	0.52	0.36	0.33
hadronisation uncertainty (%)	1.16	0.01	-0.94	-0.27	1.08	1.65
$p_{\mathrm{T}}(\bar{t})$ reweight (%)	0.55	0.19	-0.10	-0.14	-0.48	-1.31
$t\bar{t}$ (matching down) (%)	-1.08	-0.16	0.07	0.62	1.16	-1.05
$t\bar{t}$ (matching up) (%)	0.67	-0.10	-1.04	-0.65	2.66	3.70
$t\bar{t}$ (Q^2 down) (%)	1.65	1.17	-1.35	-1.78	1.21	1.08
$t\bar{t}$ (Q^2 up) (%)	-0.09	-0.74	-0.87	1.30	1.94	2.20
V+jets (matching down) (%)	-0.07	0.49	0.15	-1.18	-0.19	2.11
V+jets (matching up) (%)	0.65	0.98	0.48	-1.51	-2.33	-1.00
V+jets (Q^2 down) (%)	0.65	-0.38	-0.26	-0.25	0.38	3.25
V+jets (Q^2 up) (%)	2.27	1.36	-0.49	-2.04	-2.55	0.81
Electron energy -1σ (%)	0.12	0.13	-0.16	-0.21	0.21	0.34
Electron energy $+1\sigma$ (%)	-0.40	0.08	0.28	-0.02	-0.33	-0.34
Muon energy -1σ (%)	0.08	-0.02	-0.08	0.01	0.10	0.21
Muon energy $+1\sigma$ (%)	-0.10	0.02	0.10	-0.02	-0.15	-0.10
Tau energy -1σ (%)	-0.06	0.02	0.03	-0.00	-0.03	0.00
Tau energy $+1\sigma$ (%)	0.13	-0.02	-0.09	0.02	0.08	0.03
Unclustered energy -1σ (%)	0.88	0.88	0.21	-1.25	-1.88	-1.49
Unclustered energy $+1\sigma$ (%)	-1.08	-0.93	0.02	1.17	1.76	1.64
Total (%)	4.65	2.99	2.96	4.67	6.68	7.66

Table 8.4: Systematic uncertainties for $E_{\mathrm{T}}^{\mathrm{miss}}$ variable after unfolding for the combination of channels. These uncertainties are presented as a percentage of the result for each uncertainty in each $E_{\mathrm{T}}^{\mathrm{miss}}$ bin.

It can be seen that for some systematics, the shift for the both the "up" and the "down" systematic is in the same direction. For this reason, the uncertainties for the final results shown in Chapter 7 are symetrised with the highest error taken from the $\pm 1\sigma$ variation. The most notable case was seen for the Q^2 systematic for the last $E_{\mathrm{T}}^{\mathrm{miss}}$ bin. These fluctuations in the same direction can be explained by poor statistics in the theory samples. The shapes of the templates that enter the fit are shown in the left plot of Figure 8.7. In this example the central $t\bar{t}$ shape is steeper than that of the two systematic samples. When the statistics were increased by going to the inclusive b-tag region (right plot of Figure 8.7) no significant deviations in

shape were observed. It was therefore decided to use this enhanced statistics region in determining the contribution to the uncertainty.

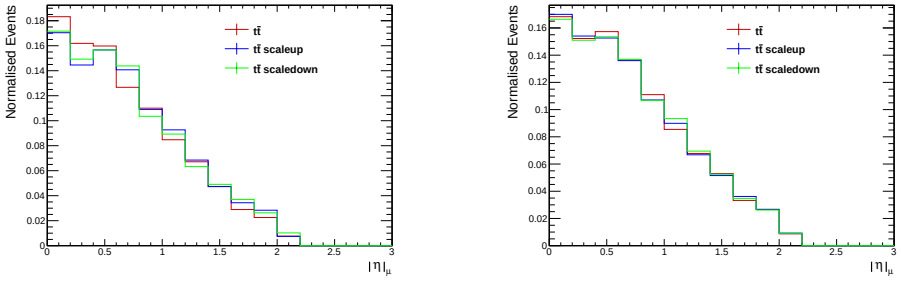

Figure 8.7: Comparison of Q^2 shapes after $\geq$ 2 b-tags (left) and $\geq$ 4 jet (right) selection.

8.7 Summary

In this chapter the systematic uncertainties that can have an effect on the inclusive and differential cross section have been described. The key sources of uncertainty are the JES and the modelling of the V+Jets background. The b-tagging uncertainty is also quite large where b-tagging is implemented. A discussion on how these might be improved will be made in the next chapter.

The systematics were assessed in more detail in the 7 TeV inclusive cross section analysis using a full Neyman construction to extract the final result and the uncertainties. This was done in order to check for any bias in the method. Since the bias was negligible in relation to the size of the uncertainties, for the 8 TeV analyses, a more standard way of assessing the uncertainties was used. One of the key problems in assessing the systematic uncertainties is the poor statistics, particularly after $\geq$ 2 b-tag selection for the differential cross section analysis. Higher statistics MC samples would allow these systematics to be more reliably assessed.

Chapter 9

Summary and Outlook

In this thesis, the inclusive and differential cross sections of the $t\bar{t}$ production process have been studied using both 7 and 8 TeV CMS data. The work in this thesis focuses on the muon-plus-jets decay channel. Collaborations have been made with others in order to combine the results with the electron-plus-jets channel for the final measurements.

The 7 TeV inclusive cross section measurement presented here was used to cross check the main result in reference [112]. This analysis used data-driven methods to extract information about the backgrounds. A maximum likelihood fit of the $|\eta_\mu|$ distribution was used to extract the result. This result is in agreement with results measured by ATLAS and CMS as can be seen in Table 9.1.

Measurement	Lumi. (fb^{-1})	Result (pb)
Thesis (μ+jets)	1.0	146 $\pm$ 8 (stat.) $^{+38}_{-28}$ (syst.)
CMS (e/μ+jets)	2.2-2.5	158 $\pm$ 2 (stat.) $\pm$ 10 (syst.) $\pm$ 4 (lumi.)
CMS ($ee, \mu\mu, e\mu$)	2.3	162 $\pm$ 3 (stat.) $\pm$ 5 (syst.) $\pm$ 4 (lumi.)
ATLAS (e/μ+jets)	2.05	186 $\pm$ 13 (stat.) $\pm$ 20 (syst.) $\pm$ 7 (lumi.)
ATLAS ($ee, \mu\mu, e\mu$)	0.7	176 $\pm$ 5 (stat.) $^{+14}_{-11}$ (syst.) $\pm$ 8 (lumi.)
Theory	N/A	163 $^{+7}_{-5}$ (μ_R and μ_F scale) $\pm$ 9 (PDF)

Table 9.1: A selection of CMS [112, 113] and ATLAS [114, 115] cross section measurements at $\sqrt{s} = 7$ TeV compared to the one presented in this thesis (blue). The approximate NNLO calculations [29] are also shown.

The 8 TeV inclusive cross section measurement was performed using the

same method as that for the 7 TeV. The result is compared to various other measurements performed by CMS and ATLAS and the theoretical prediction in Table 9.2.

Measurement	Lumi. (fb^{-1})	Result (pb)
Thesis (μ+jets)	19.7	237 ± 2 (stat.) $^{+20}_{-17}$ (syst.)
CMS (e/μ+jets)	2.8	228 ± 9 (stat.) $^{+29}_{-26}$ (syst.) ± 10 (lumi.)
CMS ($ee, \mu\mu, e\mu$)	2.4	227 ± 3 (stat.) ± 11 (syst.) ± 10 (lumi.)
ATLAS (e/μ+jets)	5.8	241 ± 2 (stat.) ± 31 (syst.) ± 9 (lumi.)
ATLAS ($e\mu$)	20.3	242 ± 2 (stat.) ± 6 (syst.) ± 8 (lumi.)
Theory	N/A	$234\ ^{+10}_{-7}$ (μ_R and μ_F scale) ± 12 (PDF)

Table 9.2: A selection of CMS [24, 23] and ATLAS [116, 117] cross section measurements at $\sqrt{s} = 8$ TeV compared to the one presented in this thesis (blue). The approximate NNLO calculations [29] are also shown.

Some enhancements were made when moving from the 7 TeV to the 8 TeV inclusive cross section measurement. Rather than estimating the V+jets background from data, MC was used. This is because it was possible to apply a much tighter event selection in the 8 TeV analysis. Using the same data-driven techniques would have yielded too much contamination from other processes, particularly $t\bar{t}$ and single-top production. It had also been proved in the 7 TeV analysis that MC models the data very well for the W and Z+jets processes. Tightening the event selection meant that the background is significantly reduced as well as the uncertainties associated with them. This can be seen when comparing tables 8.2 and 8.3, where the uncertainty due to the background modelling reduces from 11% to 5%. The largest experimental uncertainties in the 8 TeV analyses arise from b-tagging and the jet energy scale (JES). The uncertainty due to the JES will reduce as our understanding of the detector improves and more sophisticated methods of extracting the jet energy corrections are developed [118]. These uncertainties can also be reduced by adding nuisance parameters to the fitting procedure and performing fits on divided data sets as was done in reference [112]. The QCD multi-jet and W/Z+jets backgrounds can be distinguished better by using additional fitting variables such as M_{T}^{W}.

A measurement of the $t\bar{t}$ differential cross section with respect to the miss-

ing transverse energy in the event was first performed using the 7 TeV data [119]. This analysis was described in the conference proceedings for the Top 2012 workshop [120] (not presented in this thesis). Additional variables including the H_{T}, S_{T}, p_{T}^{W} and M_{T}^{W} were added for the 8 TeV analysis [121]. In general, these measurements showed a good agreement with the predictions from Monte Carlo generators. In the 8 TeV analysis, a regularised unfolding approach was adopted rather than bin-by-bin unfolding as was used in the 7 TeV $E_{\mathrm{T}}^{\mathrm{miss}}$ analysis. Regularisation helps to suppresses the large statistical fluctuations that can result from poor statistics in the response matrix. Improving the method of binning selection will help reduce migrations between bins and help improve the measurements. This could be done by allowing the bin widths to float as free parameters such that they can be optimised.

The study of inclusive and differential cross sections are important for the understanding of perturbative QCD physics in $t\bar{t}$ production. In particular constraints can be made on the PDFs. The H_{T} and S_{T} variables can give insight in to the hadronisation of the quarks produced in top pair decays. The p_{T}^{W} and M_{T}^{W} variables allow us to learn about the leptonically decaying W-boson produced by the top quark. All of these variables, in particular the $E_{\mathrm{T}}^{\mathrm{miss}}$, are used in BSM physics searches. This includes searches for SUSY and for heavy exotic top partners where top pair production is a large background. It is therefore important for these analyses to understand these variables well which could help with reducing the $t\bar{t}$ background contribution. Top pair production is also a major background to rare standard model processes such as $t\bar{t}V$ and $t\bar{t}H$ where neutrinos are produced in the decay of Vector boson or the Higgs.

The differential measurements are compared to the predictions of three different MC generators: MADGRAPH, MC@NLO and POWHEG. Comparisons are also made to different MC tunes using the MADGRAPH generator. These tunes include variations in the factorisation and renormalisation scale and the parton matching threshold. No significant deviations are observed between the data and these predictions. All theoretical tunes yield similar predictions for the $E_{\mathrm{T}}^{\mathrm{miss}}$ variable. No excess is observed in the tail of the distribution which could have been an indicator of new physics. The H_{T} and S_{T}

variables show more significant differences between the MADGRAPH tunes particularly for the Q^2 variations. This seems to suggest that it is correct to set the factorisation and renormalisation scales to the mass of the top quark. The POWHEG and MADGRAPH generators tend to make very similar predictions for the H_{T} and S_{T} variables. The explanation behind this is due to PYTHIA being used for the hadronisation for both of these generators. MADGRAPH tends to give a better description of softer transverse events whilst MC@NLO seems slightly better for harder events where more high p_{T} objects are produced. The p_{T}^{W} results show a slight discrepancy between the MC and data. It is reasonable to suggest that this is due to a similar effect as is seen in the top-p_{T} distribution [93] where the data is softer than the MC. It would be interesting to see if the NNLO prediction for this variable gives a better description as it does in the case of the top p_{T}. M_{T}^{W} is the least precisely measured variable. This is a result of significant bin-to-bin migrations. A refinement of the binning would help improve this measurement.

It is planned that all of the variables presented will be measured using the 7 TeV data and combined with the 8 TeV results in a publication. It will be interesting to look at a ratio of the differential cross sections between 7 and 8 TeV:

$$\frac{d\sigma^{7\mathrm{TeV}}/dX}{d\sigma^{8\mathrm{TeV}}/dX} \tag{9.1}$$

In these measurements, many systematic uncertainties would cancel giving very precise results. This could help when attempting to differentiate between MC generators. Another interesting future measurement could be to perform 2D differential cross sections, e.g:

$$\frac{d^2\sigma}{dE_{\mathrm{T}}^{\mathrm{miss}}/dS_{\mathrm{T}}} \tag{9.2}$$

These measurements could tell us a lot about top pair decays and would be very sensitive to new physics, particularly at high $E_{\mathrm{T}}^{\mathrm{miss}}$ and S_{T}. The presence of these sorts of events could indicate the presence of a massive undetected particle.

9.1 Summary

Measurements of top pair inclusive cross sections in the muon-plus-jets channel have been presented. The results agree with theoretical predictions and measurements from other analyses and experiments within uncertainties. Differential cross section measurements have also been made with respect to various global event variables. These include $E_{\mathrm{T}}^{\mathrm{miss}}$, H_{T}, S_{T}, p_{T}^{W} and M_{T}^{W}. The differential cross section results have been unfolded to remove selection and migration effects. This allows results to be compared with other experiments. Comparisons have been made to the available Monte Carlo simulation samples. A reasonable agreement is observed for these variables.

Appendices

Appendix A

Irradiated Test Beam Crystals

Table A.1 shows the proton fluence and the resulting damage (μ_{ind}) to crystals from proton irradiation between 2009 and 2012. The values shown are after the initial irradiation. Some of the crystals were re-irradiated or thermally recovered before being reused in test beam experiments.

Crystal ID	Year of irradiation	Accumulated fluence (p/cm^2)	μ_{ind} (m^{-1})
11124	2009	1.0×10^{13}	2.7
11992	2010	3.72×10^{13}	7.7
11952	2010	3.74×10^{13}	8.0
11935	2010	3.89×10^{13}	8.6
11121	2010	3.75×10^{13}	8.6
11128	2010	3.69×10^{13}	8.5
12014	2010	3.81×10^{13}	7.7
11962	2010	3.70×10^{13}	7.7
11118	2010	3.75×10^{13}	8.7
11133	2010	3.75×10^{13}	8.2
7022	2010	3.73×10^{13}	8.8
11856	2011	2.21×10^{13}	3.4
11830	2011	2.13×10^{13}	2.5
11845	2011	2.26×10^{13}	3.1
11133	2010+2011	5.94×10^{13}	10.9
11828	2012	5.60×10^{13}	9.2
11832	2012	5.33×10^{13}	8.8
11834	2012	5.21×10^{13}	8.3
11836	2012	6.65×10^{13}	10.2
11135	2012	5.36×10^{13}	7.4
11138	2012	5.87×10^{13}	8.8
11137	2012	5.34×10^{13}	7.6
11931	2012	5.90×10^{13}	8.4
12010	2012	5.75×10^{13}	8.3
11861	2012	1.14×10^{14}	18.8
11862	2012	1.34×10^{14}	21.7
11866	2012	1.17×10^{14}	19.9

Table A.1: List of the crystals irradiated at the PS Irradiation Facility from 2009-2012. The accumulated fluence and the induced absorption coefficient for each crystal are reported.

Appendix B

Correlations between Fitted Variables

The correlations between the fitted variables: N_{signal}, $N_{\text{V+jets}}$ and N_{QCD} for the $E_{\text{T}}^{\text{miss}}$ variable are shown in Figure B.1 for the muon+jets channel. It is seen that the correlation between signal and QCD remains very low for all $E_{\text{T}}^{\text{miss}}$ bins, while that between V+jets and QCD is negatively correlated at low $E_{\text{T}}^{\text{miss}}$ bins but not so at higher $E_{\text{T}}^{\text{miss}}$ bins as they contain almost no QCD multi-jet background.

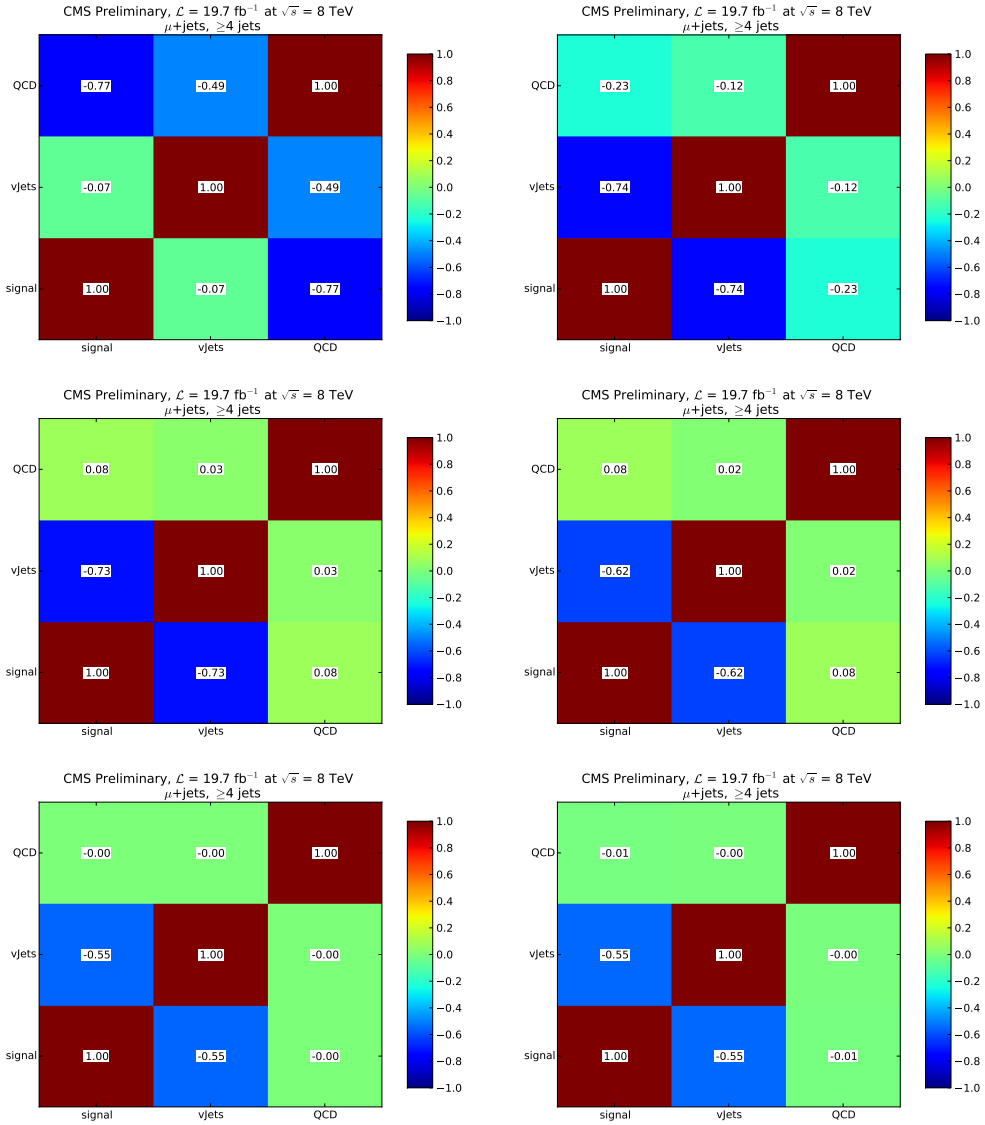

Figure B.1: Correlation between input fit processes for the $E_{\mathrm{T}}^{\mathrm{miss}}$ variable in the muon channel: 0-25 GeV(top left), 25-45 GeV(top right), 45-70 GeV(middle left), 70-100 GeV(middle right), 100-150 GeV(bottom left) and $\geq$ 150 GeV(bottom right).

Appendix C

Comparison of Differential Cross section Results in the Electron and Muon+Jets Channels

The measured differential cross section are shown for both the muon and electron+jets channels prior to the combination in Figures C.1-C.10 for all variables. A comparison is made to the Q^2 scale and matching threshold on the left and the different generators on the right.

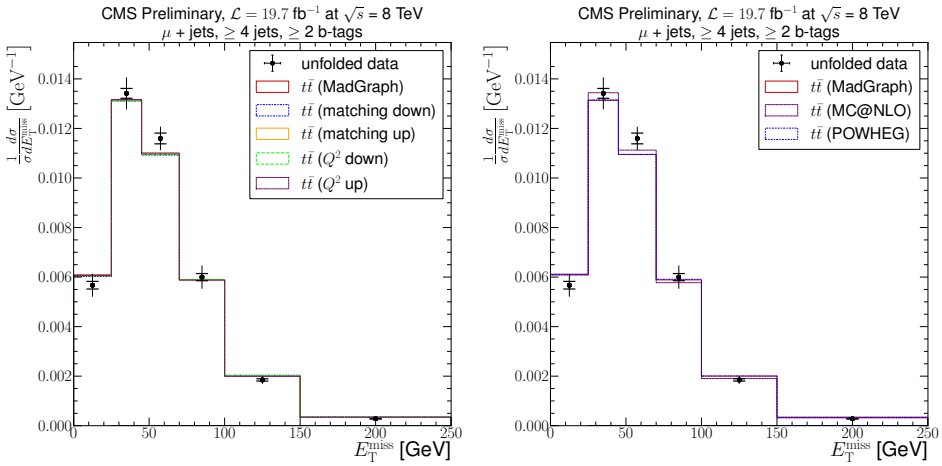

Figure C.1: Comparison of measured normalised differential cross section with respect to $E_{\mathrm{T}}^{\mathrm{miss}}$ to different MC tunes (left) and different generators (right) for the muon-plus-jets channel.

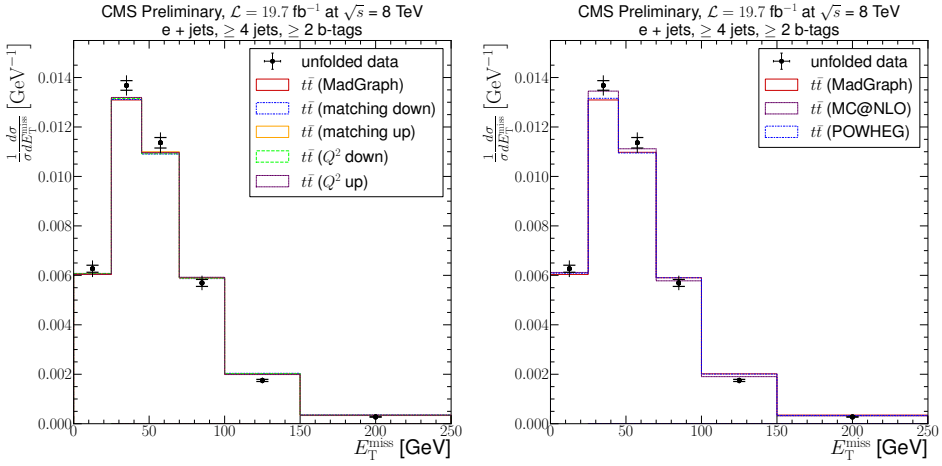

Figure C.2: Comparison of measured normalised differential cross section with respect to $E_{\mathrm{T}}^{\mathrm{miss}}$ to different MC tunes (left) and different generators (right) for the electron-plus-jets channel.

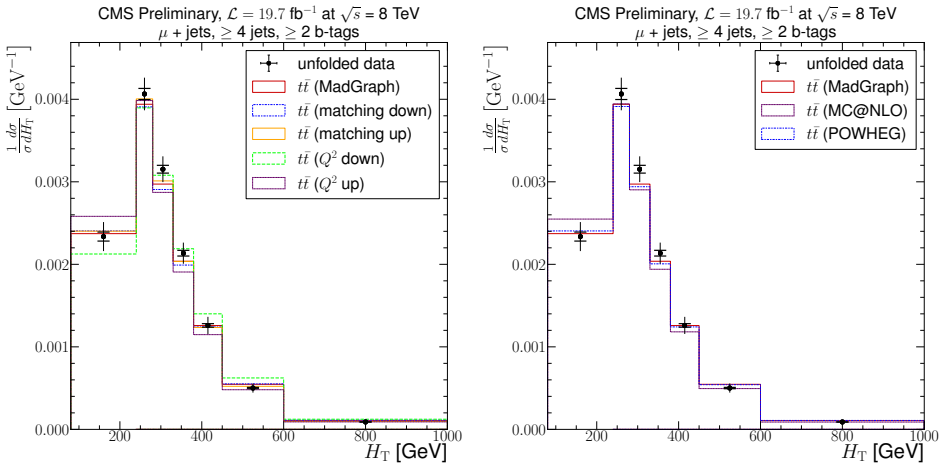

Figure C.3: Comparison of measured normalised differential cross section with respect to H_T to different MC tunes (left) and different generators (right) for the muon-plus-jets channel.

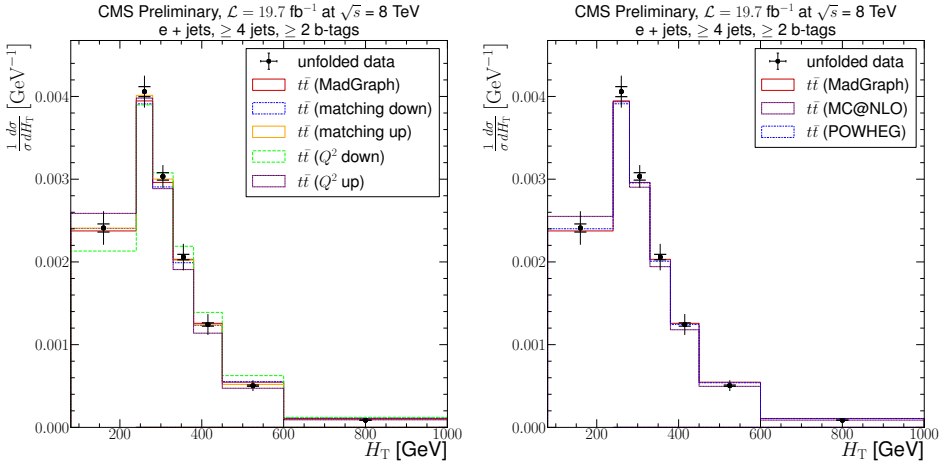

Figure C.4: Comparison of measured normalised differential cross section with respect to H_T to different MC tunes (left) and different generators (right) for the electron-plus-jets channel.

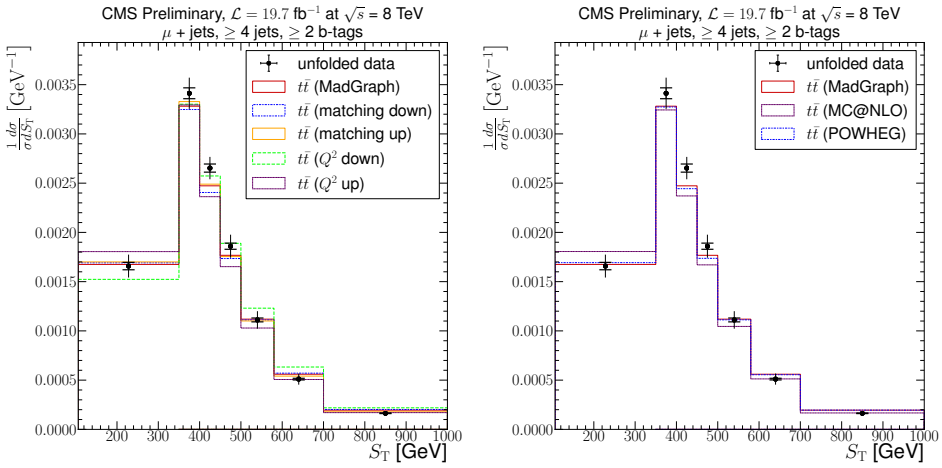

Figure C.5: Comparison of measured normalised differential cross section with respect to S_{T} to different MC tunes (left) and different generators (right) for the muon-plus-jets channel.

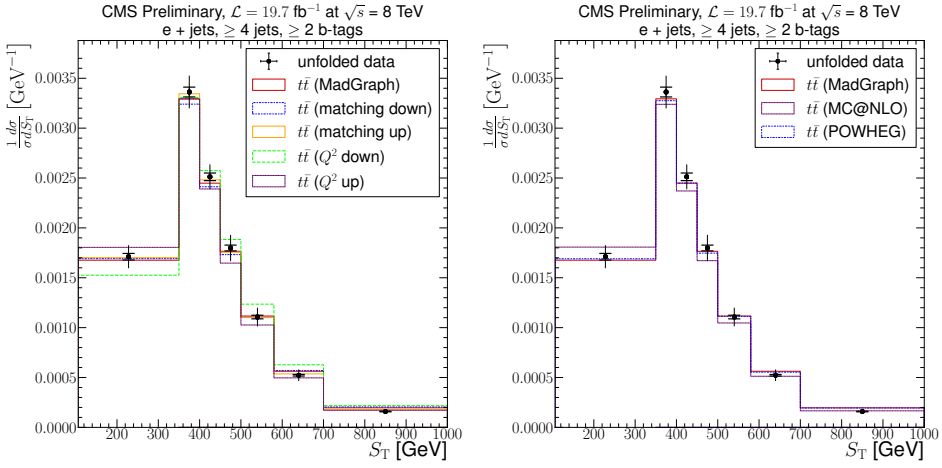

Figure C.6: Comparison of measured normalised differential cross section with respect to S_{T} to different MC tunes (left) and different generators (right) for the electron-plus-jets channel.

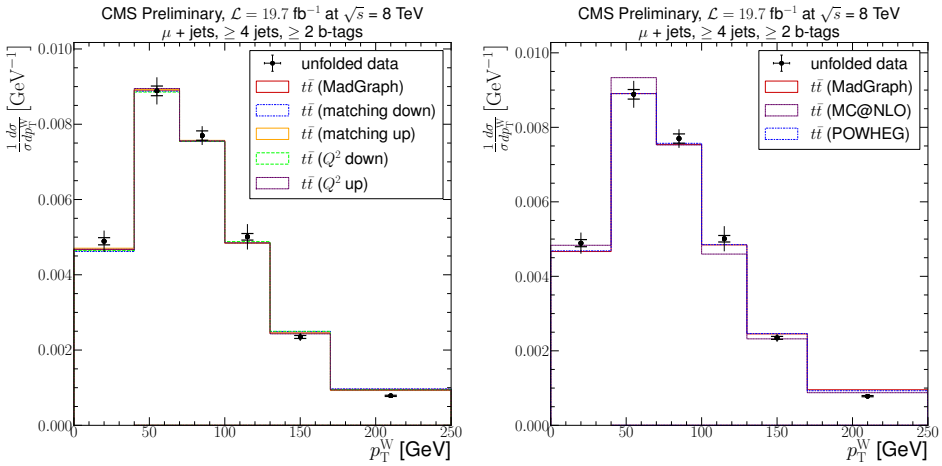

Figure C.7: Comparison of measured normalised differential cross section with respect to p_{T}^{W} to different MC tunes (left) and different generators (right) for the muon-plus-jets channel.

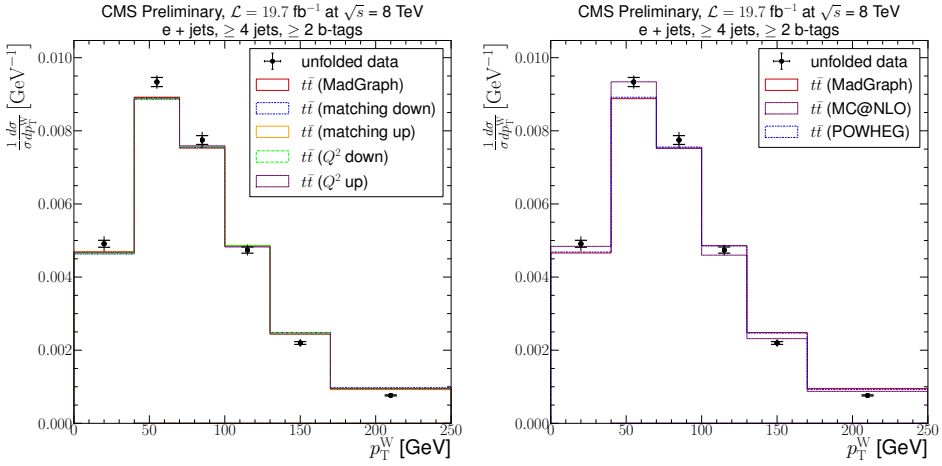

Figure C.8: Comparison of measured normalised differential cross section with respect to p_{T}^{W} to different MC tunes (left) and different generators (right) for the electron-plus-jets channel.

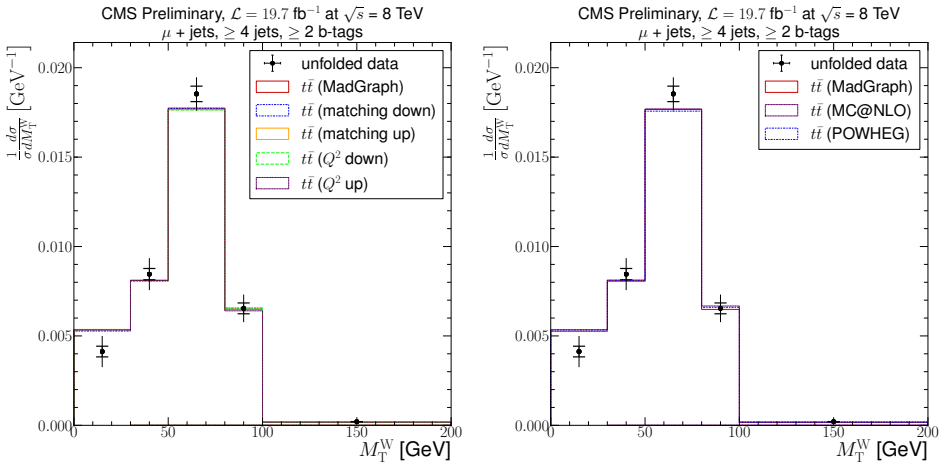

Figure C.9: Comparison of measured normalised differential cross section with respect to M_{T}^{W} to different MC tunes (left) and different generators (right) for the muon-plus-jets channel.

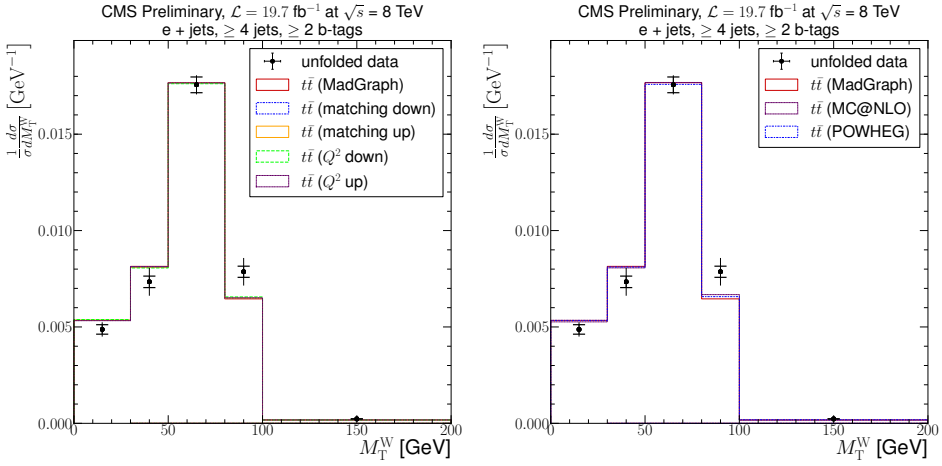

Figure C.10: Comparison of measured normalised differential cross section with respect to M_{T}^{W} to different MC tunes (left) and different generators (right) for the electron-plus-jets channel.

Appendix D

Systematics Uncertainties

Tables D.1 to D.9 show the exact values of the systematic uncertainties after unfolding the measurement for the muon-plus-jets channel and the combination of the muon and electron-plus jets channels for all of the considered variables.

Table D.1: Systematic uncertainties for $E_{\mathrm{T}}^{\mathrm{miss}}$ variable after unfolding in the μ-plus-jets channel.

Systematic Errors						
Systematic	$0< E_{\mathrm{T}}^{\mathrm{miss}} <25$	$25\leq E_{\mathrm{T}}^{\mathrm{miss}} <45$	$45\leq E_{\mathrm{T}}^{\mathrm{miss}} <75$	$75\leq E_{\mathrm{T}}^{\mathrm{miss}} <100$	$100\leq E_{\mathrm{T}}^{\mathrm{miss}} <150$	$E_{\mathrm{T}}^{\mathrm{miss}} \geq 150$
b-jets - (%)	0.17	0.07	-0.06	-0.11	-0.09	-0.01
b-jets + (%)	-0.15	-0.07	0.05	0.10	0.08	0.01
JER - (%)	0.04	0.04	-0.02	-0.02	-0.06	-0.04
JER + (%)	-0.00	0.06	0.01	-0.05	-0.09	-0.04
JES - (%)	2.38	1.35	-0.08	-1.85	-2.85	-2.95
JES + (%)	-3.03	-1.48	0.64	2.11	2.17	2.19
Light jet - (%)	0.05	0.03	-0.01	-0.05	-0.04	0.03
Light jet + (%)	-0.04	-0.03	0.01	0.05	0.03	-0.04
Pile-up - (%)	-0.14	-0.02	0.04	0.08	0.03	-0.13
Pile-up + (%)	0.34	-0.00	-0.12	-0.09	-0.02	0.12
QCD shape uncertainty (%)	0.00	0.00	0.00	0.00	0.00	0.00
hadronisation uncertainty (%)	2.40	0.73	-1.30	-1.67	0.62	3.06
$p_{\mathrm{T}}(\bar{t})$ reweight (%)	0.84	0.21	-0.18	-0.21	-0.54	-1.34
$t\bar{t}$ (matching down) (%)	-0.70	0.27	0.19	0.03	0.15	-1.66
$t\bar{t}$ (matching up) (%)	-0.96	0.96	-0.24	-1.95	2.01	4.01
$t\bar{t}$ (Q^2 down) (%)	3.99	2.24	-2.07	-3.97	0.58	3.35
$t\bar{t}$ (Q^2 up) (%)	1.84	-0.03	-1.96	-0.45	3.24	3.49
V+jets (matching down) (%)	-0.39	0.88	0.39	-1.91	-0.33	2.85
V+jets (matching up) (%)	3.89	1.62	-0.74	-2.63	-2.90	-1.13
V+jets (Q^2 down) (%)	2.21	-1.08	-1.62	0.63	2.65	3.15
V+jets (Q^2 up) (%)	2.47	2.91	0.23	-4.10	-5.60	2.07
Electron energy -1σ (%)	0.12	0.10	-0.03	-0.10	-0.12	-0.17
Electron energy $+1\sigma$ (%)	-0.11	-0.04	0.05	0.05	0.01	0.13
Muon energy -1σ (%)	0.20	-0.05	-0.17	0.04	0.16	0.37
Muon energy $+1\sigma$ (%)	-0.23	0.04	0.21	-0.01	-0.29	-0.29
Tau energy -1σ (%)	-0.10	0.07	0.08	-0.08	-0.15	-0.01
Tau energy $+1\sigma$ (%)	0.24	-0.06	-0.10	0.01	0.06	0.09
Unclustered energy -1σ (%)	0.88	0.81	0.36	-1.14	-2.20	-1.38
Unclustered energy $+1\sigma$ (%)	-0.90	-0.82	0.02	0.81	1.53	2.05
Total (%)	8.12	4.81	4.09	7.78	8.64	9.48

Table D.2: Systematic uncertainties for H_{T} variable after unfolding in the μ-plus-jets channel.

Systematic Errors							
Systematic	0–240 GeV	240–280 GeV	280–330 GeV	330–380 GeV	380–450 GeV	450–600 GeV	$\geq$ 600 GeV
b-jets - (%)	0.04	-0.02	-0.08	-0.07	-0.03	0.07	0.15
b-jets + (%)	-0.04	0.02	0.08	0.07	0.02	-0.07	-0.16
JER - (%)	-0.13	-0.04	0.05	0.19	0.18	0.23	-0.14
JER + (%)	0.10	-0.02	-0.19	0.02	0.03	-0.15	0.03
JES - (%)	4.83	0.79	-3.19	-4.62	-4.47	-4.73	-5.19
JES + (%)	-1.99	0.03	0.91	1.32	1.90	2.50	2.80
Light jet - (%)	0.02	-0.02	-0.07	-0.07	-0.01	0.09	0.14
Light jet + (%)	-0.02	0.02	0.07	0.07	0.01	-0.09	-0.15
Pile-up - (%)	-0.06	0.01	0.05	0.07	0.08	0.02	-0.04
Pile-up + (%)	0.19	-0.03	-0.16	-0.18	-0.18	-0.08	0.01
QCD shape uncertainty (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
hadronisation uncertainty (%)	-2.77	-3.61	-0.11	2.08	5.32	8.63	8.40
$p_{\mathrm{T}}(\bar{t})$ reweight (%)	0.25	0.00	-0.14	-0.14	-0.19	-0.28	-0.47
$t\bar{t}$ (matching down) (%)	2.82	0.14	-1.61	-1.63	-2.27	-3.64	-4.90
$t\bar{t}$ (matching up) (%)	0.47	-0.41	-1.48	-0.71	0.78	2.11	-0.80
$t\bar{t}$ (Q^2 down) (%)	-3.50	2.11	0.90	1.14	3.25	3.61	4.04
$t\bar{t}$ (Q^2 up) (%)	3.02	-1.40	-2.57	-1.43	-1.06	-1.91	-2.98
V+jets (matching down) (%)	-1.35	-0.67	0.92	1.01	0.25	2.63	3.91
V+jets (matching up) (%)	1.40	-0.29	-0.72	-1.16	-1.02	-0.95	-2.17
V+jets (Q^2 down) (%)	-1.77	-1.70	0.56	3.01	3.76	1.65	1.99
V+jets (Q^2 up) (%)	0.82	0.85	0.16	-0.98	-2.09	-2.07	-0.66
Electron energy -1σ (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Electron energy $+1\sigma$ (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Muon energy -1σ (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Muon energy $+1\sigma$ (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tau energy -1σ (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tau energy $+1\sigma$ (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unclustered energy -1σ (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unclustered energy $+1\sigma$ (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total (%)	7.51	4.86	4.94	5.99	8.23	11.05	11.87

Table D.3: Systematic uncertainties for H_{T} variable after unfolding in the combined channel.

Systematic Errors							
Systematic	0–240 GeV	240–280 GeV	280–330 GeV	330–380 GeV	380–450 GeV	450–600 GeV	$\geq$ 600 GeV
b-jets - (%)	0.02	-0.02	-0.04	-0.03	-0.02	0.03	0.11
b-jets + (%)	-0.02	0.02	0.04	0.03	0.02	-0.03	-0.11
JER - (%)	-0.05	0.00	0.02	0.06	0.06	0.09	-0.09
JER + (%)	0.10	-0.03	-0.16	-0.04	-0.00	-0.06	0.01
JES - (%)	4.37	0.57	-2.79	-4.00	-4.20	-4.68	-5.14
JES + (%)	-2.18	-0.24	1.21	1.91	2.13	2.46	3.03
Light jet - (%)	0.01	-0.01	-0.04	-0.03	-0.01	0.04	0.09
Light jet + (%)	-0.01	0.01	0.03	0.03	0.01	-0.05	-0.10
Pile-up - (%)	-0.04	-0.02	0.02	0.06	0.09	0.05	-0.01
Pile-up + (%)	0.11	0.01	-0.07	-0.10	-0.13	-0.10	-0.03
QCD shape uncertainty (%)	-0.50	-0.07	0.32	0.58	0.57	0.43	0.25
hadronisation uncertainty (%)	-4.14	-3.90	1.24	3.80	6.48	9.45	9.53
$p_{\mathrm{T}}(\bar{t})$ reweight (%)	0.19	0.00	-0.09	-0.10	-0.15	-0.25	-0.40
$t\bar{t}$ (matching down) (%)	2.67	0.03	-1.45	-1.69	-2.72	-3.61	-2.97
$t\bar{t}$ (matching up) (%)	0.73	0.17	-0.62	-1.28	-0.56	0.11	-1.02
$t\bar{t}$ (Q^2 down) (%)	-3.45	1.74	1.27	1.78	3.28	3.65	2.33
$t\bar{t}$ (Q^2 up) (%)	3.32	-1.21	-2.19	-2.04	-2.21	-3.24	-1.98
V+jets (matching down) (%)	-0.60	-0.19	0.47	0.28	-0.08	1.28	1.91
V+jets (matching up) (%)	0.73	-0.09	-0.49	-0.82	-0.48	-0.27	-1.01
V+jets (Q^2 down) (%)	-0.42	-0.65	-0.05	0.98	1.49	0.39	0.24
V+jets (Q^2 up) (%)	0.47	0.35	-0.09	-0.71	-1.17	-0.67	0.22
Electron energy -1σ (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Electron energy $+1\sigma$ (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Muon energy -1σ (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Muon energy $+1\sigma$ (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tau energy -1σ (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Tau energy $+1\sigma$ (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unclustered energy -1σ (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Unclustered energy $+1\sigma$ (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total (%)	7.78	4.61	4.38	6.55	8.75	11.74	11.68

Table D.4: Systematic uncertainties for S_{T} variable after unfolding in the μ-plus-jets channel.

Systematic Errors							
Systematic	0–350 GeV	350–400 GeV	400–450 GeV	450–500 GeV	500–580 GeV	580–700 GeV	$\geq$ 700 GeV
b-jets - (%)	0.05	-0.03	-0.09	-0.09	-0.05	0.04	0.12
b-jets + (%)	-0.04	0.02	0.08	0.08	0.04	-0.04	-0.12
JER - (%)	-0.17	0.15	0.14	0.04	0.09	0.12	0.10
JER + (%)	0.21	0.02	-0.24	-0.29	-0.22	-0.14	0.01
JES - (%)	3.55	0.37	-2.51	-3.76	-3.96	-4.11	-4.37
JES + (%)	-1.86	0.07	0.94	1.33	1.96	2.88	2.85
Light jet - (%)	0.03	-0.02	-0.08	-0.08	-0.03	0.08	0.12
Light jet + (%)	-0.03	0.01	0.08	0.08	0.02	-0.08	-0.12
Pile-up - (%)	-0.11	0.04	0.10	0.11	0.13	0.05	-0.03
Pile-up + (%)	0.26	-0.08	-0.24	-0.27	-0.26	-0.16	-0.03
QCD shape uncertainty (%)	0.00	0.00	0.00	0.00	0.00	0.00	0.00
hadronisation uncertainty (%)	-2.65	-3.36	0.32	3.08	4.94	7.95	7.99
$p_{\mathrm{T}}(\bar{t})$ reweight (%)	-0.02	0.22	0.15	-0.06	-0.18	-0.19	-0.34
$t\bar{t}$ (matching down) (%)	1.64	0.04	-1.74	-0.83	-0.91	-1.29	-4.11
$t\bar{t}$ (matching up) (%)	-1.54	-0.17	1.27	1.40	1.49	2.38	1.52
$t\bar{t}$ (Q^2 down) (%)	-4.07	2.16	1.75	2.42	3.85	4.99	3.55
$t\bar{t}$ (Q^2 up) (%)	1.50	-1.18	-1.19	-0.01	-0.77	-1.25	-2.04
V+jets (matching down) (%)	-0.63	0.14	0.52	-0.20	-0.11	1.37	2.18
V+jets (matching up) (%)	2.19	-0.50	-2.15	-2.20	-1.37	-1.36	-2.17
V+jets (Q^2 down) (%)	-1.33	-1.63	0.42	2.89	2.86	2.26	2.03
V+jets (Q^2 up) (%)	0.60	0.54	-0.03	-0.33	-1.17	-2.47	-0.89
Electron energy -1σ (%)	0.06	0.01	-0.05	-0.06	-0.07	-0.06	-0.02
Electron energy $+1\sigma$ (%)	-0.04	-0.03	0.02	0.08	0.09	0.03	-0.01
Muon energy -1σ (%)	0.01	-0.03	-0.05	0.00	0.04	0.02	0.05
Muon energy $+1\sigma$ (%)	0.08	-0.04	-0.06	-0.05	-0.10	-0.09	-0.03
Tau energy -1σ (%)	0.02	-0.00	-0.04	-0.02	-0.02	-0.02	0.01
Tau energy $+1\sigma$ (%)	-0.00	-0.01	0.01	0.00	-0.00	0.02	0.02
Unclustered energy -1σ (%)	0.32	0.05	-0.17	-0.35	-0.49	-0.39	-0.29
Unclustered energy $+1\sigma$ (%)	-0.36	-0.04	0.20	0.39	0.45	0.40	0.48
Total (%)	7.01	4.64	4.28	6.29	7.99	11.02	10.66

Table D.5: Systematic uncertainties for S_{T} variable after unfolding in the combined channel.

Systematic Errors							
Systematic	0–350 GeV	350–400 GeV	400–450 GeV	450–500 GeV	500–580 GeV	580–700 GeV	$\geq$ 700 GeV
b-jets - (%)	0.03	-0.01	-0.04	-0.04	-0.04	-0.00	0.06
b-jets + (%)	-0.02	0.01	0.03	0.03	0.04	0.00	-0.06
JER - (%)	-0.13	0.10	0.10	0.05	0.11	0.11	0.06
JER + (%)	0.13	0.03	-0.16	-0.21	-0.13	-0.09	-0.00
JES - (%)	3.46	0.28	-2.66	-3.66	-3.70	-4.07	-4.35
JES + (%)	-1.81	0.08	1.17	1.34	1.62	2.59	3.19
Light jet - (%)	0.02	-0.01	-0.03	-0.04	-0.03	0.01	0.05
Light jet + (%)	-0.02	0.01	0.03	0.04	0.03	-0.01	-0.05
Pile-up - (%)	-0.06	0.02	0.04	0.05	0.09	0.08	-0.00
Pile-up + (%)	0.14	-0.03	-0.11	-0.12	-0.15	-0.14	-0.05
QCD shape uncertainty (%)	-0.51	0.08	0.58	0.65	0.47	0.29	0.06
hadronisation uncertainty (%)	-3.23	-3.56	1.33	3.60	5.55	8.16	8.90
$p_{\mathrm{T}}(\bar{t})$ reweight (%)	0.06	0.16	0.07	-0.10	-0.20	-0.25	-0.39
$t\bar{t}$ (matching down) (%)	1.66	-0.06	-1.06	-1.21	-1.62	-1.96	-3.33
$t\bar{t}$ (matching up) (%)	-0.45	0.54	1.19	-0.23	-0.54	0.22	-0.14
$t\bar{t}$ (Q^2 down) (%)	-3.54	2.26	1.83	1.65	3.68	4.01	2.28
$t\bar{t}$ (Q^2 up) (%)	2.42	-1.32	-1.33	-1.76	-1.65	-3.03	-2.14
V+jets (matching down) (%)	-0.17	-0.00	0.06	-0.16	-0.10	0.62	0.97
V+jets (matching up) (%)	1.33	-0.22	-1.39	-1.51	-0.86	-0.82	-1.34
V+jets (Q^2 down) (%)	-0.01	-1.07	-0.66	0.75	1.18	0.94	0.80
V+jets (Q^2 up) (%)	0.57	0.18	-0.25	-0.39	-0.86	-1.54	-0.54
Electron energy -1σ (%)	0.06	-0.03	-0.10	-0.08	-0.01	0.05	-0.03
Electron energy $+1\sigma$ (%)	0.00	0.02	0.03	0.02	-0.01	-0.09	-0.08
Muon energy -1σ (%)	0.01	-0.02	-0.03	-0.00	0.03	0.02	0.03
Muon energy $+1\sigma$ (%)	0.03	-0.01	-0.02	-0.02	-0.05	-0.05	-0.01
Tau energy -1σ (%)	0.02	-0.00	-0.03	-0.02	-0.01	-0.00	0.02
Tau energy $+1\sigma$ (%)	0.01	-0.00	-0.01	-0.01	-0.01	0.01	0.01
Unclustered energy -1σ (%)	0.27	0.08	-0.09	-0.32	-0.46	-0.44	-0.30
Unclustered energy $+1\sigma$ (%)	-0.29	-0.05	0.10	0.31	0.45	0.46	0.36
Total (%)	6.45	4.65	4.29	6.02	7.92	10.25	10.95

Table D.6: Systematic uncertainties for p_{T}^{W} variable after unfolding in the μ-plus-jets channel.

Systematic Errors						
Systematic	0–40 GeV	40–70 GeV	70–100 GeV	100–130 GeV	130–170 GeV	$\geq$ 170 GeV
b-jets - (%)	0.00	0.11	0.03	-0.15	-0.18	0.03
b-jets + (%)	-0.00	-0.11	-0.03	0.14	0.18	-0.03
JER - (%)	-0.03	-0.06	-0.00	0.04	0.09	0.15
JER + (%)	0.09	-0.02	-0.05	-0.01	0.01	0.02
JES - (%)	1.35	1.07	-0.16	-1.22	-1.80	-2.57
JES + (%)	-1.18	-0.95	-0.05	1.26	1.71	2.31
Light jet - (%)	0.03	0.06	0.01	-0.11	-0.13	0.09
Light jet + (%)	-0.03	-0.07	-0.01	0.12	0.14	-0.08
Pile-up - (%)	-0.09	-0.09	0.07	0.09	0.07	0.10
Pile-up + (%)	0.10	0.17	-0.09	-0.17	-0.11	-0.12
QCD shape uncertainty (%)	0.00	0.00	0.00	0.00	0.00	0.00
hadronisation uncertainty (%)	-0.21	1.84	-0.35	-3.13	-0.57	2.45
$p_{\mathrm{T}}(\bar{t})$ reweight (%)	0.17	0.27	0.05	-0.20	-0.43	-0.77
$t\bar{t}$ (matching down) (%)	0.46	0.94	0.44	-1.04	-0.83	-3.35
$t\bar{t}$ (matching up) (%)	1.26	1.10	-0.10	-2.49	-1.26	-0.42
$t\bar{t}$ (Q^2 down) (%)	2.14	1.90	-1.29	-3.82	-0.99	0.64
$t\bar{t}$ (Q^2 up) (%)	0.90	1.21	-0.40	-2.25	-1.62	1.36
V+jets (matching down) (%)	4.40	-0.07	-2.33	-1.26	-0.60	-0.95
V+jets (matching up) (%)	0.20	0.53	-0.42	-0.33	0.70	-1.60
V+jets (Q^2 down) (%)	-1.09	-0.79	-0.50	0.65	2.46	3.36
V+jets (Q^2 up) (%)	-1.40	1.32	1.64	-1.28	-2.19	-0.93
Electron energy -1σ (%)	-0.03	0.07	0.09	-0.06	-0.16	-0.19
Electron energy $+1\sigma$ (%)	-0.09	-0.00	-0.00	-0.02	0.08	0.23
Muon energy -1σ (%)	-0.21	-0.09	0.04	0.04	0.18	0.52
Muon energy $+1\sigma$ (%)	0.15	0.17	0.06	-0.14	-0.31	-0.59
Tau energy -1σ (%)	0.12	-0.03	-0.05	0.01	-0.02	-0.08
Tau energy $+1\sigma$ (%)	0.03	-0.05	-0.03	0.02	0.05	0.11
Unclustered energy -1σ (%)	1.08	0.56	-0.17	-0.73	-0.99	-1.89
Unclustered energy $+1\sigma$ (%)	-1.05	-0.69	0.11	0.77	1.35	1.94
Total (%)	5.79	4.05	3.24	6.73	4.54	6.55

Table D.7: Systematic uncertainties for p_{T}^{W} variable after unfolding in the combined channel.

Systematic Errors						
Systematic	0–40 GeV	40–70 GeV	70–100 GeV	100–130 GeV	130–170 GeV	$\geq$ 170 GeV
b-jets - (%)	0.00	0.03	-0.01	-0.06	-0.04	0.08
b-jets + (%)	-0.00	-0.03	0.01	0.06	0.04	-0.07
JER - (%)	-0.01	-0.04	-0.02	0.01	0.08	0.17
JER + (%)	0.05	0.01	-0.04	-0.03	0.01	0.01
JES - (%)	1.08	0.83	0.10	-0.98	-1.82	-2.49
JES + (%)	-1.45	-0.84	0.21	1.27	1.69	2.05
Light jet - (%)	0.02	0.01	-0.02	-0.05	-0.02	0.08
Light jet + (%)	-0.02	-0.02	0.02	0.05	0.03	-0.08
Pile-up - (%)	-0.09	-0.03	0.06	0.02	0.01	0.12
Pile-up + (%)	0.10	0.07	-0.09	-0.06	-0.03	-0.12
QCD shape uncertainty (%)	-0.36	-0.39	0.07	0.58	0.58	0.38
hadronisation uncertainty (%)	-1.09	1.15	-0.23	-1.97	0.51	3.14
$p_{\mathrm{T}}(\bar{t})$ reweight (%)	0.18	0.22	0.06	-0.16	-0.44	-0.78
$t\bar{t}$ (matching down) (%)	-0.03	0.67	0.27	-0.91	-0.58	-0.82
$t\bar{t}$ (matching up) (%)	0.65	0.28	-0.18	-1.20	-0.22	0.53
$t\bar{t}$ (Q^2 down) (%)	1.18	1.00	-0.56	-2.20	-0.74	0.24
$t\bar{t}$ (Q^2 up) (%)	0.70	0.40	-0.09	-1.22	-1.01	0.70
V+jets (matching down) (%)	1.53	-0.06	-0.61	-0.16	-0.56	-1.09
V+jets (matching up) (%)	0.34	0.16	-0.06	-0.01	-0.10	-1.39
V+jets (Q^2 down) (%)	-1.20	0.02	0.75	0.51	-0.27	0.11
V+jets (Q^2 up) (%)	-0.81	0.61	0.78	-0.64	-0.89	-0.22
Electron energy -1σ (%)	-0.54	-0.09	0.18	0.18	0.34	0.51
Electron energy $+1\sigma$ (%)	0.33	0.15	-0.05	-0.13	-0.33	-0.73
Muon energy -1σ (%)	-0.11	-0.04	0.03	0.01	0.08	0.27
Muon energy $+1\sigma$ (%)	0.09	0.08	0.01	-0.08	-0.16	-0.28
Tau energy -1σ (%)	0.01	-0.01	0.01	0.02	-0.03	-0.02
Tau energy $+1\sigma$ (%)	0.02	-0.02	-0.03	0.00	0.02	0.08
Unclustered energy -1σ (%)	0.94	0.52	-0.09	-0.64	-1.04	-1.89
Unclustered energy $+1\sigma$ (%)	-1.08	-0.57	0.14	0.77	1.23	1.83
Total (%)	3.62	2.53	2.05	4.30	3.38	5.38

Table D.8: Systematic uncertainties for M_{T}^{W} variable after unfolding in the μ-plus-jets channel.

Systematic Errors					
Systematic	0–30 GeV	30–50 GeV	50–80 GeV	80–100 GeV	$\geq$ 100 GeV
b-jets - (%)	0.07	0.12	-0.01	-0.16	-0.19
b-jets + (%)	-0.03	-0.14	0.00	0.16	0.18
JER - (%)	-1.00	0.41	0.14	-0.22	0.22
JER + (%)	-0.47	0.13	0.16	-0.36	-0.16
JES - (%)	-3.91	0.83	0.66	-0.41	1.57
JES + (%)	2.55	0.08	-0.46	-0.33	-1.49
Light jet - (%)	-0.06	0.11	0.02	-0.16	-0.12
Light jet + (%)	0.07	-0.11	-0.02	0.16	0.12
Pile-up - (%)	-0.49	0.21	0.13	-0.38	0.23
Pile-up + (%)	1.03	-0.24	-0.21	0.28	-0.35
QCD shape uncertainty (%)	0.00	0.00	0.00	0.00	0.00
hadronisation uncertainty (%)	17.97	3.00	-3.68	-3.65	-10.37
$p_{\mathrm{T}}(\bar{t})$ reweight (%)	2.71	-0.25	-0.37	-0.58	-0.65
$t\bar{t}$ (matching down) (%)	-1.37	-1.53	-0.14	3.94	-0.36
$t\bar{t}$ (matching up) (%)	0.63	-4.68	0.19	4.76	-0.85
$t\bar{t}$ (Q^2 down) (%)	2.84	-0.93	0.39	-0.78	-15.18
$t\bar{t}$ (Q^2 up) (%)	-0.98	-0.52	-0.54	3.03	5.58
V+jets (matching down) (%)	-0.87	-6.79	0.54	6.81	3.35
V+jets (matching up) (%)	5.34	-2.07	-0.46	-0.35	-0.44
V+jets (Q^2 down) (%)	4.10	3.87	-1.16	-3.84	-0.62
V+jets (Q^2 up) (%)	-2.38	-2.32	2.17	-2.94	-6.64
Electron energy -1σ (%)	0.22	-0.25	0.11	-0.32	-0.21
Electron energy $+1\sigma$ (%)	-0.39	0.14	0.12	-0.33	-0.12
Muon energy -1σ (%)	1.48	0.49	-0.12	-1.24	-1.87
Muon energy $+1\sigma$ (%)	-0.59	-1.10	0.15	1.18	1.19
Tau energy -1σ (%)	0.18	-0.17	-0.05	0.22	0.21
Tau energy $+1\sigma$ (%)	0.08	0.08	-0.06	0.12	-0.22
Unclustered energy -1σ (%)	-0.53	-2.20	0.06	2.85	1.47
Unclustered energy $+1\sigma$ (%)	0.35	2.68	-0.44	-1.73	-1.20
Total (%)	21.13	10.56	5.01	11.77	20.10

Table D.9: Systematic uncertainties for M_{T}^{W} variable after unfolding in the combined channel.

Systematic Errors					
Systematic	0–30 GeV	30–50 GeV	50–80 GeV	80–100 GeV	$\geq$ 100 GeV
b-jets - (%)	0.00	0.06	0.01	-0.08	-0.07
b-jets + (%)	0.02	-0.07	-0.01	0.08	0.07
JER - (%)	-0.43	0.19	0.10	-0.19	0.13
JER + (%)	-0.34	0.00	0.13	-0.17	-0.06
JES - (%)	-3.43	-0.42	0.80	0.36	2.17
JES + (%)	2.09	1.25	-0.67	-0.59	-1.53
Light jet - (%)	-0.03	0.03	0.02	-0.06	-0.06
Light jet + (%)	0.03	-0.03	-0.02	0.06	0.06
Pile-up - (%)	-0.00	0.09	0.02	-0.16	0.02
Pile-up + (%)	0.25	-0.14	-0.05	0.11	-0.10
QCD shape uncertainty (%)	-0.15	-0.49	-0.09	0.87	1.03
hadronisation uncertainty (%)	9.41	5.10	-2.32	-4.69	-6.70
$p_{\mathrm{T}}(\bar{t})$ reweight (%)	1.22	0.00	-0.13	-0.57	-0.53
$t\bar{t}$ (matching down) (%)	0.10	-0.16	0.15	0.01	-3.20
$t\bar{t}$ (matching up) (%)	0.89	-1.40	0.06	0.65	-1.11
$t\bar{t}$ (Q^2 down) (%)	1.80	-0.07	0.23	-0.89	-10.65
$t\bar{t}$ (Q^2 up) (%)	-0.18	-1.36	-0.36	2.06	6.32
V+jets (matching down) (%)	-0.54	-3.12	0.27	2.62	1.87
V+jets (matching up) (%)	2.67	-1.54	-0.50	1.11	-0.21
V+jets (Q^2 down) (%)	1.60	2.40	-0.79	-1.05	-0.78
V+jets (Q^2 up) (%)	-1.07	-1.17	1.17	-1.62	-3.29
Electron energy -1σ (%)	1.31	0.99	-0.05	-1.74	-2.58
Electron energy $+1\sigma$ (%)	-2.05	-0.65	0.20	1.42	3.24
Muon energy -1σ (%)	0.67	0.24	-0.05	-0.58	-0.86
Muon energy $+1\sigma$ (%)	-0.27	-0.61	0.07	0.58	0.61
Tau energy -1σ (%)	-0.19	0.11	0.03	-0.08	0.14
Tau energy $+1\sigma$ (%)	0.05	-0.01	-0.02	0.02	0.03
Unclustered energy -1σ (%)	-1.07	-1.71	0.27	1.72	1.03
Unclustered energy $+1\sigma$ (%)	0.84	1.68	-0.40	-1.05	-0.48
Total (%)	12.26	7.92	3.63	7.63	14.40

Bibliography

[1] Paul A.M. Dirac. On the Theory of quantum mechanics. *Proc.Roy.Soc.Lond.*, A112:661–677, 1926.

[2] S.N. Bose. Planck's law and light quantum hypothesis. *Z.Phys.*, 26:178–181, 1924.

[3] Sheldon L. Glashow. Partial-symmetries of weak interactions, 1961.

[4] Abdus Salam and John Clive Ward. Electromagnetic and weak interactions. *Phys.Lett.*, 13:168–171, 1964.

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

[6] J Beringer et al. Review of Particle Physics, 2012-2013. Review of Particle Properties. *Phys. Rev. D*, 86(1):010001, 2012.

[7] Makoto Kobayashi and Toshihide Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog.Theor.Phys.*, 49:652–657, 1973.

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

[9] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, Oct 1964.

[10] CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys.Lett.*, B716:30–61, 2012.

[11] M. Gell-Mann. A schematic model of baryons and mesons. *Physics Letters*, 8(3):214 – 215, 1964.

[12] G Zweig. An SU_3 model for strong interaction symmetry and its breaking; Version 1. Technical Report CERN-TH-401, CERN, Geneva, Jan 1964.

[13] Richard P. Feynman. Very high-energy collisions of hadrons. *Phys.Rev.Lett.*, 23:1415–1417, 1969.

[14] CDF Collaboration. Observation of top quark production in $\bar{p}p$ collisions. *Phys.Rev.Lett.*, 74:2626–2631, 1995.

[15] D0 Collaboration. Observation of the top quark. *Phys.Rev.Lett.*, 74:2632–2637, 1995.

[16] CMS ATLAS, CDF and D0 Collaborations. First combination of Tevatron and LHC measurements of the top-quark mass. 2014.

[17] F.D. Aaron et al. Combined Measurement and QCD Analysis of the Inclusive e+- p Scattering Cross Sections at HERA. *JHEP*, 1001:109, 2010.

[18] D. R. Stump. Proceedings to amsterdam 2002, ichep. *Proceedings to Amsterdam 2002, ICHEP*, page 265, 2002.

[19] Ringaile Placakyte. Parton Distribution Functions. 2011.

[20] Timo Antero Aaltonen et al. Combination of measurements of the top-quark pair production cross section from the Tevatron Collider. 2013.

[21] CMS Collaboration. Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 1211:067, 2012.

[22] ATLAS and CMS Collaborations. Combination of ATLAS and CMS top-quark pair cross section measurements using proton-proton collisions at $\sqrt{s} = 7$ TeV. Technical Report CMS-PAS-TOP-12-003, CERN, Geneva, 2013.

[23] CMS Collaboration. Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s}$ = 8 TeV. *J. High Energy Phys.*, 02(arXiv:1312.7582. CMS-TOP-12-007. CERN-PH-EP-2013-234):024. 29 p, Dec 2013. Comments: Replaced with published version. Added journal reference and DOI.

[24] CMS Collaboration. Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 8$ TeV. 2013.

[25] CMS Collaboration. Measurement of the single-top-quark t-channel cross section in pp collisions at $\sqrt{s}$ = 7 TeV. *J. High Energy Phys.*, 12(arXiv:1209.4533. CMS-TOP-11-021. CERN-PH-EP-2012-274):035. 41 p, Sep 2012.

[26] CMS Collaboration. Combination of single top-quark cross-sections measurements in the t-channel at $\sqrt{s} = 8$ TeV with the ATLAS and CMS experiments. Technical Report CMS-PAS-TOP-12-002. ATLAS-CONF-2013-098, CERN, Geneva, Sep 2013.

[27] CMS Collaboration. Observation of the associated production of a single top quark and a W boson in pp collisions at $\sqrt{s} = 8$ TeV. 2014.

[28] ATLAS Collaboration. Search for s-Channel Single Top-Quark Production in *pp* Collisions at $\sqrt{s}$ = 7 TeV. (ATLAS-CONF-2011-118), Aug 2011.

[29] Nikolaos Kidonakis. Single top and top pair production. 2012.

[30] Andreas Hocker, H. Lacker, S. Laplace, and F. Le Diberder. A New approach to a global fit of the CKM matrix. *Eur.Phys.J.*, C21:225–259, 2001.

[31] Vardan Khachatryan et al. Measurement of the t-channel single-top-quark production cross section and of the $| V_{tb} |$ CKM matrix element in pp collisions at $\sqrt{s}$= 8 TeV. *JHEP*, 1406:090, 2014.

[32] T. Aaltonen et al. Observation of Single Top Quark Production and Measurement of $|Vtb|$ with CDF. *Phys.Rev.*, D82:112005, 2010.

[33] S. Agostinelli et al. Geant4: a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250 – 303, 2003.

[34] Fabio Maltoni and Tim Stelzer. MadGraph 5: Going Beyond. *JHEP*, 1106:128, 2011.

[35] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 1006:043, 2010.

[36] Stafano Frixione and Bryan R. Webber. The MC@NLO 3.3 Event Generator. 2006.

[37] S. Hoche et al. Matching Parton Showers and Matrix Elements. *hep-ph/0602031*, 2006.

[38] Lyndon Evans and Philip Bryant. LHC Machine. *Journal of Instrumentation*, 3(08):S08001, August 2008.

[39] Helmut Wiedemann. *Particle Accelerator Physics*. Springer-Verlag, Berlin, third edition, May 2007.

[40] CMS Collaboration. Public CMS Luminosity Information. https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.

[41] CMS Collaboration. The CMS experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[42] CMS Collaboration. Studies of Tracker Material. Technical Report CMS-PAS-TRK-10-003, 2010.

[43] V Karimki, M Mannelli, P Siegrist, H Breuker, A Caner, R Castaldi, K Freudenreich, G Hall, R Horisberger, M Huhtinen, and A Cattai. *The CMS tracker system project: Technical Design Report*. Technical Design Report CMS. CERN, Geneva, 1997.

[44] CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the CMS tracker. 2014.

[45] Addendum to the CMS ECAL technical design report: Changes to the CMS ECAL electronics. 2002.

[46] CMS Collaboration. The Electromagnetic Calorimeter Project - Technical Design Report. *J. Phys. G*, CERN/LHCC 97-33:290, 1997.

[47] CMS Collaboration. Energy calibration and resolution of the CMS electromagnetic calorimeter in pp collisions at $\sqrt{s} = 7$ TeV. 2013.

[48] *The CMS hadron calorimeter project: Technical Design Report.* Technical Design Report CMS. CERN, Geneva, 1997.

[49] CMS Collaboration. Design, Performance, and Calibration of CMS Hadron-Barrel Calorimeter Wedges. 2007.

[50] CMS Collaboration. *The CMS muon project: Technical Design Report.* Technical Design Report CMS. CERN, Geneva, 1997.

[51] CMS Collaboration. Performance of CMS muon reconstruction in cosmic-ray events. *Journal of Instrumentation*, 5:3022, March 2010.

[52] CMS Collaboration. Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV. *Journal of Instrumentation*, 7:2P, October 2012.

[53] CMS Collaboration. *CMS TriDAS project: Technical Design Report, Volume 1: The Trigger Systems.* Technical Design Report CMS.

[54] M. Huhtinen, P. Lecomte, D. Luckey, F. Nessi-Tedaldi, and F. Pauss. High-energy proton induced damage in PbWO4 calorimeter crystals. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 545(12):63 – 87, 2005.

[55] F. E. Maas et al. Measurement of the Transverse Beam Spin Asymmetry in Elastic Electron-Proton Scattering and the Inelastic Contribution to the Imaginary Part of the Two-Photon Exchange Amplitude. *Phys. Rev. Lett.*, 94:082001, Mar 2005.

[56] A.A Annenkov, M.V Korzhik, and P Lecoq. Lead tungstate scintillation material. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 490(12):30 – 50, 2002.

[57] Peter Kozma, Robert Bajgar, and Petr Kozma Jr. Radiation resistivity of PbF2 crystals. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 484(13):149 – 152, 2002.

[58] Masaaki Kobayashi, Kenji Harada, Yoshiaki Hirose, Mitsuru Ishii, and Isao Yamaga. Large-size bismuth silicate (Bi4Si3O12) scintillating crystals of good quality. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 400(23):392 – 400, 1997.

[59] P. Adzic et al. Radiation hardness qualification of PbWO(4) scintillation crystals for the CMS Electromagnetic Calorimeter. *JINST*, 5:P03010, 2010.

[60] M. Glaser, F. Ravotti, and M. Moll. Dosimetry Assessments in the Irradiation Facilities at the CERN-PS Accelerator. In *Radiation and Its Effects on Components and Systems, 2005. RADECS 2005. 8th European Conference on*, pages PI5–1–PI5–7, 2005.

[61] A. Ledovskoy. Non linearity of energy measurements in EE.

[62] A. Ledovskoy. Simulation of radiation damage effects in CMS ECAL. (CMS DN-2013/008).

[63] Tracking and Primary Vertex Results in First 7 TeV Collisions. Technical Report CMS-PAS-TRK-10-005, CERN, 2010. Geneva, 2010.

[64] R Frhwirth. Application of Kalman filtering to track and vertex fitting. *Nucl. Instrum. Methods Phys. Res., A*, 262(HEPHY-PUB-503):444. 19 p, Jun 1987.

[65] R Frhwirth, Wolfgang Waltenberger, and Pascal Vanlaer. Adaptive Vertex Fitting. Technical Report CMS-NOTE-2007-008, CERN, Geneva, Mar 2007.

[66] CMS Collaboration. Particle-Flow Event Reconstruction in CMS and Performance for Jets, Taus, and MET. Technical Report CMS-PAS-PFT-09-001, CERN, 2009. Geneva, Apr 2009.

[67] CMS Collaboration. Electron reconstruction in CMS. Technical Report CMS-NOTE-2006-040, CERN, Geneva, Feb 2006.

[68] CMS Collaboration. Electron reconstruction and identification at $\sqrt{s}$ = 7 TeV. Technical Report CMS-PAS-EGM-10-004, CERN, Geneva, 2010.

[69] CMS Collaboration. Cut-based electron identification. https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideCutBasedElectronID.

[70] CMS Collaboration. Multivariate electron identification. https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideNeuralNetElectronID.

[71] CMS Collaboration. CMS 2013 Public Electron Performance Results. https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideCutBasedElectronID.

[72] CMS Collaboration. Effective area values. https://twiki.cern.ch/twiki/bin/viewauth/CMS/EgammaEARhoCorrection.

[73] CMS Collaboration. Muon identification. https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonPOGIDNote.

[74] CMS Collaboration. Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV. *CMS PAS*, 2010.

[75] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The anti-kt jet clustering algorithm. *JHEP*, 04:063, 2008.

[76] CMS Collaboration. Performance of Jet Algorithms in CMS. *CMS PAS*, JME-07-003, 2007.

[77] CMS Collaboration. Jet identification. https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID.

[78] CMS Collaboration. B-tagging in CMS. *EPJ Web Conf.*, 28:12055, 2012.

[79] Christian Weiser. A Combined Secondary Vertex Based B-Tagging Algorithm in CMS. Technical Report CMS-NOTE-2006-014, CERN, Geneva, Jan 2006.

[80] Wolfgang Adam, Boris Mangano, Thomas Speer, and Teddy Todorov. Track Reconstruction in the CMS tracker. Technical Report CMS-NOTE-2006-041, CERN, Geneva, Dec 2006.

[81] Thomas Speer, Kirill Prokofiev, R Frhwirth, Wolfgang Waltenberger, and Pascal Vanlaer. Vertex Fitting in the CMS Tracker. Technical Report CMS-NOTE-2006-032, CERN, Geneva, Feb 2006.

[82] CMS Collaboration. Results on b-tagging identification in 8 TeV pp collisions. Feb 2013.

[83] Performance of b tagging at sqrt(s)=8 TeV in multijet, ttbar and boosted topology events. Technical Report CMS-PAS-BTV-13-001, CERN, Geneva, 2013.

[84] CMS Collaboration. CMS Particle Flow and Tau Identification Results. *https://twiki.cern.ch/twiki/bin/view/CMS/MuonReferenceEffs*, 2010.

[85] CMS Collaboration. Measurement of differential top-quark pair production cross sections in pp colisions at $\sqrt{s} = 7$ TeV. *Eur.*

Phys. J. C, 73(arXiv:1211.2220. CMS-TOP-11-013. CERN-PH-EP-2012-322):2339. 47 p, Nov 2012. Comments: Replaced with published version. Added journal reference and DOI.

[86] CMS Collaboration. Generic Tag and Probe Tool for Measuring Efficiency at CMS with Early Data. *CMS-AN-2009-111*.

[87] CMS Collaboration. Reference muon trigger, id and isolation efficiencies. https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsMUO.

[88] CMS Collaboration. Results on b-tagging identification in 8 TeV pp collisions. *CMS report*, CMS-DP-2013-005, Feb 2013.

[89] CMS Collaboration. Combination of b-tagging efciency measurements in 2012 data at 8 TeV pp collision. *CMS-AN-2012-470*.

[90] CMS Collaboration. Re-weighting of events to account BTag efficiency and mistag scale factors. *https://twiki.cern.ch/twiki/bin/viewauth/CMS/BTagWeight*, 2012.

[91] CMS Collaboration. Jet Energy Corrections determination at 7 TeV. *CMS PAS*, CMS-PAS-JME-10-010, 2010.

[92] CMS Collaboration. Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS. *JINST*, 6:P11002, 2011.

[93] CMS Collaboration. Measurement of differential top-quark pair production cross sections in *pp* colisions at $\sqrt{s} = 7$ TeV. *Eur.Phys.J.*, C73:2339, 2013.

[94] CMS Collaboration. Measurement of differential top-quark pair production cross sections in the lepton+jets channel in pp collisions at 8 TeV. Technical Report CMS-PAS-TOP-12-027, CERN, Geneva, 2013.

[95] CMS Collaboration. Measurement of the differential ttbar cross section in the dilepton channel at 8 TeV. Technical Report CMS-PAS-TOP-12-028, CERN, Geneva, 2013.

[96] CMS Collaboration. pt(top-quark) based reweighting of ttbar MC. https://twiki.cern.ch/twiki/bin/viewauth/CMS/TopPtReweighting.

[97] CMS Collaboration. Top pair cross section in e/mu+jets at 8 TeV. Technical Report CMS-PAS-TOP-12-006, CERN, Geneva, 2012.

[98] Ryan Gavin, Ye Li, Frank Petriello, and Seth Quackenbush. FEWZ 2.0: A code for hadronic Z production at next-to- next-to-leading order. 2010.

[99] Torbjrn Sjstrand, Patrik Edn, Christer Friberg, Leif Lnnblad, Gabriela Miu, Stephen Mrenna, and Emanuel Norrbin. High-energy-physics event generation with Pythia6.1. *Computer Physics Communications*, 135(2):238 – 259, 2001.

[100] CMS Collaboration. Jet Production Rates in association with W and Z Bosons in pp Collisions at $\sqrt{s} = 7$ TeV.

[101] K. Ellis J. Campbell. MCFM - Monte Carlo for FeMtobarn processes, http://mcfm.fnal.gov/.

[102] Stefano Frixione, Fabian Stoeckli, Paolo Torrielli, Bryan R. Webber, and Chris D. White. The MCaNLO 4.0 Event Generator. 2010.

[103] F James and M Roos. MINUIT: a system for function minimization and analysis of the parameter errors and corrections. *Comput. Phys. Commun.*, 10(CERN-DD-75-20):343–367. 38 p, Jul 1975.

[104] Kirill Melnikov and Frank Petriello. Electroweak gauge boson production at hadron colliders through O(alpha(s)**2). *Phys. Rev.*, D74:114017, 2006.

[105] ATLAS Collaboration. Measurements of top quark pair relative differential cross-sections with ATLAS in *pp* collisions at $\sqrt{s} = 7$ TeV. *Eur.Phys.J.*, C73:2261, 2013.

[106] CMS Collaboration. Measurement of $E_{\mathrm{T}}^{\mathrm{miss}}$ in $t\bar{t}$ events at 7 TeV. *CMS PAS*, TOP-12-019, 2012.

[107] Andreas Hoecker and Vakhtang Kartvelishvili. SVD Approach to Data Unfolding. 1995. arxiv:hep-ph/9509307.

[108] Tim Adye. Unfolding algorithms and tests using RooUnfold. 2011. arxiv:1105.1160.

[109] CMS Collaboration. Particle-Flow based Measurement of the Top-Quark Pair Production Cross-Section in the Muon+Jets Channel using CMS 2010 Data, AN-10-368.

[110] CMS Collaboration. Absolute Calibration of the Luminosity Measurement at CMS: Winter 2012 Update. 2012.

[111] CMS Collaboration. Measurement of the inclusive production cross sections for forward jets and for dijet events with one forward and one central jet in pp collisions at $\sqrt{s} = 7$ TeV. *JHEP*, 1206:036, 2012.

[112] CMS Collaboration. Measurement of the $t\bar{t}$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV with lepton + jets final states. *Phys. Lett. B*, 720(arXiv:1212.6682. CMS-TOP-11-003. CERN-PH-EP-2012-285):83–104. 31 p, Dec 2012.

[113] CMS Collaboration. Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 7$ TeV. *J. High Energy Phys.*, 11(arXiv:1208.2671. CMS-TOP-11-005. CERN-PH-EP-2012-224):067. 39 p, Aug 2012. Comments: Submitted to the Journal of High Energy Physics.

[114] ATLAS Collaboration. Measurement of the top quark pair cross section with ATLAS in pp collisions at $\sqrt{(s)} = 7$ TeV using final states with an electron or a muon and a hadronically decaying τ lepton. *Phys.Lett.*, B717:89–108, 2012.

[115] ATLAS Collaboration. Measurement of the cross section for top-quark pair production in *pp* collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector using final states with two high-pt leptons. *JHEP*, 1205:059, 2012.

[116] ATLAS Collaboration. Measurement of the top quark pair production cross section in the single-lepton channel with ATLAS in proton-proton collisions at 8 TeV using kinematic fits with b-tagging. Technical Report ATLAS-CONF-2012-149, CERN, Geneva, Nov 2012.

[117] ATLAS Collaboration. Measurement of the $t\bar{t}$ production cross-section using $e\mu$ events with b-tagged jets in pp collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS detector. 2014.

[118] Ricardo Eusebi and on behalf of the CMS Collaboration). Jet energy corrections and uncertainties in CMS: reducing their impact on physics measurements. *Journal of Physics: Conference Series*, 404(1):012014, 2012.

[119] CMS Collaboration. Measurement of missing transverse energy in top pair events. Technical Report CMS-PAS-TOP-12-019, CERN, Geneva, 2012.

[120] P Symonds and the CMS Collaboration. Measurement of Missing Transverse Energy in Top Pair Events. *Journal of Physics: Conference Series*, 452(1):012032, 2013.

[121] CMS Collaboration. Measurement of $E_{\mathrm{T}}^{\mathrm{miss}}$ and other global distributions in top pair events. Technical Report CMS-PAS-TOP-12-042, CERN, Geneva, 2013.