University of BRISTOL
VRIJE UNIVERSITEIT BRUSSEL
School of Physics

The Search for the Standard Model Production of Four Top Quarks

Lana Beck

A dissertation submitted to the University of Bristol and Vrije Universiteit Brussel in accordance with the requirements of the degree of PhD in the Faculty of Science, School of Physics, Bristol and Faculteit Wetenschappen en Bio-ingenieurswetenschappen, Vakgroep Fysica, Brussel.

Dec 2016

University of Bristol, Vrije Universiteit Brussel

Abstract

The Standard Model (SM) of Particle Physics has been incredibly successful and accurate in describing the fundamental particles that make up the world around us and the way they behave. However, we know it is not the ultimate theory of nature as some phenomena remain unexplained. Outstanding questions include what is dark matter, this mysterious material which can be inferred from observations of the universe but has not been directly detected; and where does gravity fit into the picture? These questions motivate our search for physics Beyond the Standard Model (BSM) at the Large Hadron Collider at the CERN research facility. Here we accelerate protons around a 27 km ring and collide them at four experiments located underground. When we collide protons we effectively collide the more fundamental particles within the protons, quarks and gluons. This thesis focuses on research undertaken at the CMS experiment studying the heaviest quarks, top quarks, which are not found in nature but instead are produced in high-energy experiments. Top quarks are most often produced in pairs, however, this thesis focuses on the search for the simultaneous production of four top quarks, which is an incredibly rare process in comparison. A precision measurement of this rare process would be a stringent test on the SM and may give hints of physics beyond the standard model. Untangling the signal of four-top-quark production from the overwhelming background of top-quark-pair production in the output of the detector is incredibly difficult. Algorithms, which are often used in developing artificial intelligence, are therefore employed to exploit subtle differences in signatures, greatly increasing the sensitivity. Results are presented which place tight limits on the rate of four-top-quark production and projections of the future sensitivity are made. This includes an estimate of when CMS will have sufficient data to definitively observe this process at SM rates. The results also allow us to place constraints on properties of hypothesised BSM particles. Here we interpret the results to place constraints on the mass and top quark-coupling of one such particle, the sgluon.

Samenvatting

Het Standaard Model (SM) van deeltjesfysica is ongelofelijk succesvol en nauwkeurig in het beschrijven van de fundamentele deeltjes in de wereld om ons heen en de manier waarop ze zich gedragen. Maar we weten dat het niet de ultieme natuurkundige theorie is, omdat een aantal verschijnselen onverklaard blijven in het SM. Openstaande vragen zijn: wat is donkere materie, het bestaan van deze mysterieuze substantie kan worden afgeleid uit waarnemingen van het heelal, maar is nog niet rechtstreeks waargenomen; en waar past de zwaartekracht in het plaatje? Deze vragen motiveren onze zoektocht naar de fysica voorbij het Standaard Model (BSM) bij de Large Hadron Collider op het CERN. Hier versnellen we protonen rond een ring van 27 km omtrek. Op vier plaatsen aan de ring liggen ondergrondse experimenten waar de protonen worden gebotst. Wanneer protonen botsen, bestuderen we effectief de werkelijk fundamentele deeltjes in het proton, quarks en gluonen. Dit proefschrift richt zich op onderzoek gebruikmakend van de proton-proton botsingen waargenomen door het CMS-experiment. Het zwaarste bekende elementaire deeltje, de top-quark, is niet te vinden in de natuur, maar in plaats daarvan kan worden geproduceerd in de botsingen bij de LHC. Top quarks worden meestal geproduceerd in paren, maar dit proefschrift richt zich op de zoektocht naar de productie van vier top-quark tegelijkertijd, dat is een ongelooflijk zeldzaam proces in vergelijking met paarproductie. Een nauwkeurige meting van deze zeldzame proces zou een strenge test zijn van het SM en kan hints geven of er nieuwe deeltjes worden gemaakt samen die in vier top quarks uiteenvallen. Het identificeren van het signaal van vier top-quark productie in de overweldigende achtergrond van top-quark-paarproductie in de data is een ongelooflijk moeilijke wetenschappelijke uitdaging. Hiervoor worden machine-learning algoritmen toegepast. Op deze wijze kunnen subtiele verschillen tussen de productie van top quark paren en vier top quarks in de botsing worden benut, en dit leidt tot een aanzienlijke verhoging van de gevoeligheid van de data-analyse. De resultaten van dit onderzoek plaatsen de meest strakke grenzen aan de werkzame doorsnede voor productie van vier top quarks. De resultaten zijn ook in staat om beperkingen te geven op de eigenschappen van eventuele hypothetische BSM deeltjes. In dit onderzoek worden daarom de resultaten ook geïnterpreteerd als een functie van de massa en top quark-koppeling van zo’n hypothetisch nieuw deeltje, het zogenaamde sgluon.

Acknowledgements

First of all I would like to thank my supervisors, Freya Blekman and Joel Goldstein, for being incredibly supportive, helpful and great fun to work with over the past four years. I would also like to thank James Keaveney for teaching me everything possible about being a four-top expert in the first two years of my PhD, for continuing to discuss physics problems with me since then and for being great craic. It was also a pleasure to work with Jesse and Steve from UCR on combining our four-top searches together and I wish Denys and Long the best in taking the analysis forward from here. I would like to thank my collaborators on our phenomenology project, Didar, Benjamin and Kentarou, for a fruitful and interesting collaboration. I would like to thank both the Bristol and Brussels research group for being great colleagues and fun to work with. I would particularly like to thank Robin for all the advice, memes and the exchange of swearing at pieces of code, and Maarten for housing and feeding me at the very start of my PhD while I found somewhere to live. Many thanks go to Martisse, you are sunshine on a cloudy day. Thanks to Leonor for being a supportive flatmate over this past two years and helping me out in these last few weeks of writing. Thanks to Mark for being my rock, for encouraging me and helping me through this writing period and for always being there to laugh with me. I am incredibly appreciative of the never-ending support from my parents who brought me up to believe that I could be whatever I wanted to be. Lastly, of course, this thesis would not be possible without the work from everyone in the CMS collaboration who work together for the common goal of pushing the boundaries of physics.

The cover of this thesis was kindly designed by the talented Siphiwe Manda.

Author's Declaration

I declare that the work in this dissertation was carried out in accordance with the Regulations of the University of Bristol and Vrije Universiteit Brussel. The work is original except where indicated by special reference in the text and no part of the dissertation has been submitted for any other degree. Any views expressed in the dissertation are those of the author and in no way represent those of the University of Bristol or Vrije Universiteit Brussel. The dissertation has only been presented to the University of Bristol and Vrije Universiteit Brussel for attainment of a joint degree and has not been presented to any other university for examination either in the United Kingdom or overseas.

Signed

Date

Contents

6 **Analysis strategy and techniques** **55**

8 Phenomenological study of Run 1 four-top-quark production cross section limits 111

List of Figures

List of Tables

1 | Introduction

The standard model (SM) of particle physics is the currently accepted theory for describing the known fundamental building blocks of the universe. It includes the six quarks and six leptons (and their anti-particles) and the four fundamental forces, electromagnetism, weak force, strong force and gravity. It also includes the Higgs boson that arises from the Brout-Englert-Higgs mechanism which describes the origin of mass of the quarks, leptons and weak gauge bosons. It has stood up to rigorous testing at experiments such as at the Deutsches Elektronen-Synchrotron (DESY), the Large Electron Positron collider (LEP) at CERN and the Tevatron at Fermilab. However there are still many unanswered questions about the universe such as how does gravity fit in when it is not described in the SM? What is the dark matter in the universe that affects galaxy rotation curves and causes gravitational lensing where no baryonic matter is present? The Large Hadron Collider (LHC) at CERN aims to answers these question by studying the possible signatures of new physics that may arise around the electroweak scale. One of the main goals of the LHC was to find the Higgs boson which was confirmed in July 2012. The precision measurement of SM processes is also still very important as deviations from the SM expectation can give hints about new physics and they are also key background processes which should be well understood so that new physics signals can be found.

The SM process of the production of four top quarks is studied in this thesis. Although this process is predicted by the SM it is extremely rare and hence it has not yet been possible to measure it. A precision measurement of four top quark production would be an exceptional test of the SM. In addition to this, many models of new physics predict final states which contain four top quarks including supersymmetry, models with extra dimensions, models where the top quark/Higgs boson is a composite particle, or models which contain pair-produced scalar particle which decay into top quark pairs. Therefore the SM production of four top quarks is both an important background to these new models and the production rate of four top quarks may be enhanced by new physics models.

The thesis is structured as follows: First the background theory of particle physics and top quark physics is discussed in Chapter 2. The LHC and particularly the CMS detector are described in Chapter 3. The reconstruction of physics objects from the detector read-out are given in Chapter 4. In Chapter 6 the general strategy for searching for four top quarks and the analytical and statistical techniques required are described. The search for four top quarks in the single lepton channel in the 2012 dataset collected by the CMS experiment at $\sqrt{s} = 8$ TeV is described in Chapter 7. The phenomenological interpretation of the former analysis and another CMS same-sign dilepton analysis in the context of a model of new physics where a sgluon particle is predicted is given in Chapter 8. The search for four top quarks is continued in the 2015 dataset collected by the CMS experiment at $\sqrt{s} = 13$ TeV, described in Chapter 9, where enhancements are made to the analysis and additional search channels, opposite-sign and same-sign dilepton are combined with the single lepton channel to produce a more sensitive final result. Finally a summary of the thesis can be found in Chapter 10, including a discussion of the relevance of the analysis in the field of high energy physics and a look towards the future.

The author's personal contributions to the analyses in this thesis include:

- In Chapter 7, on the search for four top quarks at $\sqrt{s} = 8$ TeV, Sections 7.7, 7.10.1 and 7.12.
- All of the work in Chapter 8, except on Figure 8.3, the result from the single lepton channel was not the author's work.
- In Chapter 9, on the search for four top quarks at $\sqrt{s} = 13$ TeV, all sections except for the training of the hadronic top BDT in Section 9.7.1 are the author's work.

In this thesis the convention of using natural units, $\hbar = c = 1$, is adopted.

2 Theory

In this chapter the standard model (SM) of particle physics is introduced, including QED, weak and QCD interactions. A brief description of the theory of proton-proton collisions is given before a discussion of the physics of the top quark, including four-top-quark production. Some of the shortcomings of the SM are described, after which a discussion of some of the models beyond the standard model (BSM) which can produce four-top-quark signatures is presented. The concepts of 'effective field theories' and 'simplified models' are introduced.

2.1 Standard Model

The SM [1–4] is a theory which describes the elementary particles and their interactions. Matter consists of six quarks and six leptons, each of which has an anti-particle with the same mass and opposite-sign charge. They are organised into three generations, each of which contains heavier particles than the last, as seen in Table 2.1. All of the normal matter on Earth is made up of particles from the first generation, i.e. up and down quarks make up protons and neutrons, and combined with electrons they form atoms.

The leptonic sector consists of charged leptons, which can interact via the electromagnetic and weak forces, and neutrinos, which interact via the weak force only. The neutrinos are assumed to be massless in the SM, however, observations of neutrino oscillations have revealed that neutrinos have mass [5].

Quarks interact via the electromagnetic, weak or strong forces. Each quark has an electric charge, as seen in Table 2.1, and carries a colour charge of red, green or blue, where all three colours combined can form a colour-singlet state. A combination of quarks with a colour and its anti-colour can also form a colour-singlet state.

Generation	Quarks			Leptons		
	Flavour	Electric Charge	Mass (MeV)	Flavour	Electric Charge	Mass (MeV)
I	u	2/3	$2.2^{+0.6}_{-0.4}$	e	-1	0.511
	d	-1/3	$4.7^{+0.5}_{-0.4}$	ν_e	0	$< 2 \times 10^{-6}$
II	c	2/3	$(1.27 \pm 0.03) \times 10^3$	μ	-1	105.66
	s	-1/3	96^{+8}_{-4}	ν_μ	0	< 0.19
III	t	2/3	$(173.21 \pm 0.87) \times 10^3$	τ	-1	1776.86 ± 0.12
	b	-1/3	$(4.18^{+0.04}_{-0.3}) \times 10^3$	ν_τ	0	< 18.2

Table 2.1: The quarks and leptons in the SM [6].

The phenomenon of colour confinement means that quarks can only be found in colour-singlet states such as in baryons or mesons.

Finally, the force carriers consist of gauge bosons of integer spin, as seen in Table 2.2. Photons and Z bosons mediate neutral electroweak interactions whereas the W bosons mediate charged electroweak interactions. The gluons mediate the strong interaction and occur with 8 different types of colour charge which will be described in Section 2.1.3.

Gauge boson	Force	Electric Charge	Mass (GeV)	Spin	Range (m)
Photon (γ)	electromagnetic	0	0	1	∞
$W^\pm$	weak	± 1	80.385 ± 0.015	1	10^{-18}
Z	weak	0	91.1876 ± 0.0021	1	10^{-18}
gluon (g)	strong	0	0	1	10^{-15}

Table 2.2: The gauge bosons of the SM [6].

The discovery of the Higgs boson in 2012 [7–9] completed the SM with an explanation of how the fundamental particles acquire mass via the electroweak symmetry breaking mechanism.

The SM is a gauge theory based on the symmetry group SU(3)×SU(2)×U(1), which is a combination of the electroweak interactions based on the SU(2)×U(1) symmetry group and the strong interaction which is based on the SU(3) group.

In the following, the SM will be described in more detail.

2.1.1 The gauge principle

Emmy Noether showed that conservation laws result from underlying continuous symmetries of a physical system [10]. This principle had a profound impact in the development of quantum field theories. A quantum field possesses gauge symmetries, where the gauge regulates the degrees of freedom of the Lagrangian. For the action of a system to be conserved, a Lagrangian should be invariant under continuous local gauge transformations (i.e. transformations that depend on the space-time point). Each group of transformations has one or more generators, each with an associated gauge (vector) field. Gauge bosons are quantisations of these gauge fields. Global transformations are a subset of local transformations where the transformation is not dependent on the space-time point.

2.1.2 Electroweak theory

2.1.2.1 Quantum Electrodynamics

Quantum Electrodynamics is an Abelian gauge theory based on the $U(1)_{\mathrm{EM}}$ symmetry group, which has the electric charge, q, as its generator. Figure 2.2 shows the elementary QED vertex where a charged particle (charged lepton or quark) and its anti-particle interact with a photon. The convention used in this thesis is that time flows from left to right. However, these diagrams may be rotated as long as they conserve energy. Hence, this vertex can describe particle-anti-particle annihilation into a photon, a photon pair-producing a particle-anti-particle pair, a particle emitting a photon or an anti-particle emitting a photon, depending on the orientation of the diagram. QED interactions conserve lepton or quark flavour.

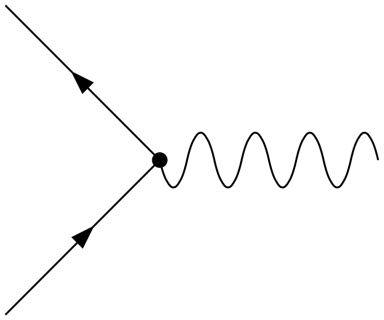


Figure 2.1: Elementary QED vertex.

Feynman diagrams are visual representations of the terms in a perturbative series expansion with respect to the interaction strength around the non-interacting

Lagrangian. Incoming particles are considered to be free particles, the interaction term then turns on and the outgoing particle are again considered to be free particles. Terms involving three or more fields represent interactions between the fields, and are referred to as vertex factors. The propagators can be described from the free particle Lagrangian and they represent intermediate states which connect vertices. Feynman diagrams are said to be at “tree level” if they are the representation of the leading order term in the expansion.
The QED Lagrangian can be found in Eq. 2.1 with ψ representing a relativistic spin-1/2 field and D_μ representing the covariant derivative, which is defined as $D_\mu = \partial_\mu + iqA_\mu$. Here q represents the charge of the particle and A^μ is the massless field of the electromagnetic four-potential. The mass of the particle is represented by m and $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field (or Faraday) tensor.

$$\mathcal{L} = \overline{\psi}\left(i\gamma^\mu D_\mu - m\right)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{2.1}$$

The Dirac Lagrangian is manifestly invariant to global phase transformations $\psi \to e^{i\theta}\psi$. The supplementation by iqA_μ is required to make the Lagrangian invariant to local phase transformations $\psi \to e^{i\theta(x)}\psi$ where A_μ transforms as $A_\mu \to A_\mu - \frac{1}{q}\partial_\mu\theta(x)$. Hence, the QED Lagrangian is gauge invariant under U(1) phase transformations, where U(1) is the unitary group of complex numbers.

Application of the Euler-Lagrange equations to Eq. 2.1 leads to the derivation of the Dirac Equation shown in Eq. 2.2.

$$\left(i\gamma^\mu D_\mu - m\right)\psi = 0 \tag{2.2}$$

The Dirac equation describes the motion for spin-1/2 particles with mass, ie. quarks and charged leptons.

In QED the vacuum acts like a dielectric medium which produces electron-positron pairs where the virtual electron is attracted to positive charges and the virtual positron is repelled (and vice versa for negative charges). This vacuum polarisation partially screens the charged particle and effectively reduces its field. However at short distances, the effective charge increases as the screening reduces.

2.1.2.2 Weak interactions

The charged weak interaction is the only interaction where a flavour changing process can occur. There are two W bosons, one of positive charge and one of negative charge, $W^{\pm}$. The diagram of weak nuclear decay in Fig. 2.2 illustrates the W boson's interaction with different flavour quarks and leptons.

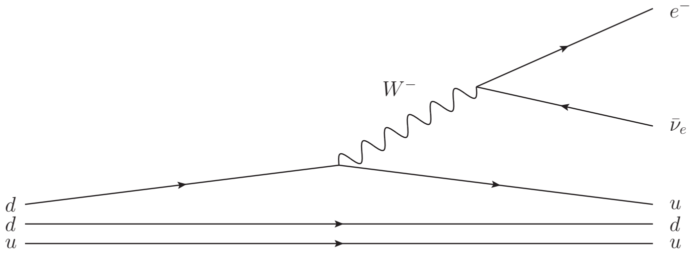


Figure 2.2: Weak nuclear decay of neutron to proton.

The neutral weak interaction is mediated by Z bosons which can interact with any quark or lepton as long as the flavour is conserved at the vertex.

The weak force only interacts with chirality left-handed (right-handed) particles (anti-particles). The charge for weak interaction is *weak isospin* (I_3) which is $\pm 1/2$ for left-handed fermions, ± 1 for $W^{\pm}$, and zero otherwise.

Weak interactions are the only interactions known to violate parity conservation, which was hypothesised by Yang and Lee [11] and experimentally validated by Wu in 1957 [12]. It was later discovered that weak interactions also violate the combined charge-parity (CP) symmetry [13, 14].

2.1.2.3 Electroweak Unification

Glashow [1], Weinberg [2] and Salam [3] formulated the unification of the weak and electromagnetic forces combined in the $SU(2)_L$ x $U(1)_Y$ gauge group. They are hypothesised to merge into one electroweak force above the unification energy of

$\approx$ 100 GeV. Weak hypercharge is defined as $Y_W = 2(q - I_3)$. Y_W is the generator of the $\mathrm{U(1)_Y}$ component of the electroweak gauge group, which is mediated by the B boson. In the $\mathrm{SU(2)_L}$ component of the electroweak theory, the left-handed fields transform as doublets and the right-handed fields transform as singlets. It is only in the weak interaction that left-handed and right-handed particles are treated differently. The generators of the $\mathrm{SU(2)_L}$ component are $\mathbf{T} = \boldsymbol{\sigma}/2$ where $\boldsymbol{\sigma}$ are the Pauli matrices. The left-handed doublet, ψ_L, can consist of either a left-handed up-type and a left-handed down-type quark, or a left-handed neutrino and left-handed charged lepton. The right-handed field, ψ_R, can be either a right-handed up-type quark, a right-handed down-type quark, or a right-handed charged lepton. Right-handed neutrinos have not been observed. The possible left-handed and right-handed fields are:

$$\psi_L = \begin{pmatrix} u_L \\ d_L \end{pmatrix}, \begin{pmatrix} \nu_L \\ \ell_L \end{pmatrix} \qquad \psi_R = \ u_R,\ d_R,\ \ell_R$$

where u_L and d_L represent all up-type and down-type quarks, respectively.

The projection operators, $\mathrm{P}_{L/R} = \frac{1}{2}(1 \pm \gamma^5)$, are used to obtain the left-handed and right-handed components of fermionic fields via:

$$\psi = \mathrm{P}_L\psi + \mathrm{P}_R\psi = \psi_L + \psi_R \tag{2.3}$$

where γ_5 is a combination of the gamma matrices:

$$\gamma_5 = i\gamma_0\gamma_1\gamma_2\gamma_3 = \begin{pmatrix} 0 & I_2 \\ I_2 & 0 \end{pmatrix} \tag{2.4}$$

To ensure local gauge invariance, the covariant derivative is defined as:

$$D_\mu = \partial_\mu + ig\mathbf{T}\cdot\mathbf{W}_\mu + i\frac{g'}{2}Y_W B_\mu \tag{2.5}$$

Here four new gauge fields are introduced. The triplet of fields $\mathbf{W} = W^a$ ($a = 1,\ 2,\ 3$) are required for $\mathrm{SU(2)_L}$ gauge invariance with coupling constant g, and the gauge field B_μ is required for $\mathrm{U(1)_Y}$ gauge invariance with coupling constant g'. For right-handed singlet particles, $\mathbf{T} = \mathbf{0}$ and the second term in Eq. 2.5 vanishes. Equation 2.6 shows the Lagrangian for electroweak interactions, where $\mathbf{W}^a_{\mu\nu}$ and $B_{\mu\nu}$ are the field strength tensors for the W^a_μ and B_μ fields, respectively.

$$\mathcal{L}_{\mathrm{EWK}} = \bar{\psi} i \gamma^\mu D_\mu \psi - \frac{1}{4}\mathbf{W}^a_{\mu\nu}\mathbf{W}^{\mu\nu}_a - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{2.6}$$

The fields of the four more familiar physical gauge bosons from Table 2.2, $W^{\mu\pm}$, Z^μ and the photon (A^μ), are linear combinations of the gauge fields W^a_μ and B_μ. This is represented in Eqs. 2.7 & 2.8, where θ_W is the weak mixing angle.

$$\begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B^\mu \\ W^\mu_3 \end{pmatrix} \tag{2.7}$$

$$\mathrm{W}^{\mu\pm} = \frac{1}{\sqrt{2}}\left(\mathrm{W}^\mu_1 \mp i\mathrm{W}^\mu_2\right) \tag{2.8}$$

The spontaneous symmetry breaking mechanism proposed by Brout, Englert [15] and Higgs [16] results in the Higgs field getting a vacuum expectation value (VEV) and the electroweak Lagrangian changing form as described in Ref [17]. It now has a kinetic term, a term for the neutral interactions and one for the charged interactions. Also introduced is a term for the Higgs three and four point self-interactions and another term for the Higgs interactions with vector bosons. There is a term for both the three and four point interactions of the vector bosons and another term for the Yukawa couplings between the Higgs field and the fermions. It is through the spontaneous symmetry breaking mechanism that the $\mathrm{W}^\pm$ and Z bosons gain mass. Fermions also acquire mass due to Yukawa terms as shown in Eq. 2.9. Here M_f is the mass of the fermion, Y_f is the Yukawa coupling and v is the vacuum expectation value.

$$M_f = Y_f \frac{v}{\sqrt{2}} \tag{2.9}$$

The charged current interaction is particularly interesting for the W decays from top quarks. It is described by the term in Eq. 2.10, where ψ_L is now rotated from flavour eigenstate to the weak eigenstate,.

$$\mathcal{L}_C = \frac{g}{\sqrt{2}} i\bar{\psi}_L \gamma^\mu \partial_\mu \psi_L \tag{2.10}$$

$$\psi_L = \begin{pmatrix} u_L \\ d'_L \end{pmatrix}, \begin{pmatrix} \nu_L \\ l_L \end{pmatrix}$$

The weak eigenstates are related to the flavour eigenstates through the CKM matrix, V_{CKM}, via $d'_L = V_{CKM}\ d_L$, where d'_L is a superposition of the flavour eigenstates. The CKM matrix, also known as the *quark mixing matrix*, is a unitary matrix which describes the strength of the couplings for weak decays, as shown in Eq. 2.11.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.11}$$

The amplitudes for the up-type quarks to transition to down-type quarks are given in Eq. 2.12 [6].

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97417 \pm 0.00021 & 0.2248 \pm 0.0006 & 0.00409 \pm 0.00039 \\ 0.220 \pm 0.005 & 0.995 \pm 0.016 & 0.0405 \pm 0.0015 \\ 0.0082 \pm 0.0006 & 0.040 \pm 0.0027 & 1.009 \pm 0.031 \end{pmatrix} \tag{2.12}$$

2.1.3 Quantum chromodynamics

Quantum chromodynamics (QCD) is a non-Abelian gauge theory based on the $\mathrm{SU}(3)_C$ symmetry group that describes the strong interactions between quarks and gluons. Quarks and gluons carry colour charge, C. Each (anti-)quark will carry one of (anti-) red, (anti-) green or (anti-) blue colour charge whilst there are 8 types of gluon which exist in a superposition of colour-anti-colour states. One of the elementary QCD vertices is shown in Fig 2.3 where two quarks couple to a gluon. There are also three and four-point interactions between gluons.

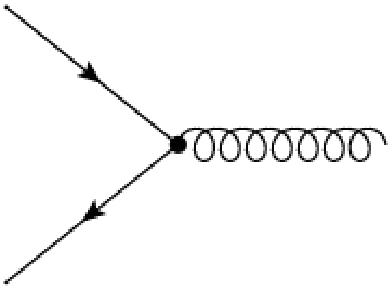


Figure 2.3: An elementary QCD vertex.

The Gell-Mann matrices, λ_α, are the generators of the SU(3)$_C$ group. In QCD, the covariant derivative is $D_\mu = \partial_\mu - igA^\alpha_\mu \lambda_\alpha$ where g represents the strong coupling constant and A^α_μ represents the gluon field. Equation 2.13 gives the QCD lagrangian, $\mathcal{L}_{\mathrm{QCD}}$, where $G^a_{\mu\nu} = \partial_\mu \mathcal{A}^a_\nu - \partial_\nu \mathcal{A}^a_\mu + gf^{abc}\mathcal{A}^b_\mu \mathcal{A}^c_\nu$ is the gluon field strength tensor and f^{abc} are the structure constants of SU(3).

$$\mathcal{L}_{\mathrm{QCD}} = \bar{\psi}_i \left(i(\gamma^\mu D_\mu)_{ij} - m\,\delta_{ij} \right) \psi_j - \frac{1}{4} G^a_{\mu\nu} G^{\mu\nu}_a \tag{2.13}$$

Asymptotic freedom and colour confinement

Quark-anti-quark loops lead to screening of the quark colour charge, however gluon loops contribute the opposite by ‘anti-screening’. It was found that in any theory with $11n > 2f$, where n is the number of “colours” and f is the number of “quark flavours”, the coupling constant, $\alpha_S\left(|q^2|\right)$, will decrease with increasing energy, q^2 [18, 19]. This is known as *asymptotic freedom* as the quarks inside hadrons effectively act like free particles. This is in contrast to QED where there is no ‘anti-screening’ effect.

With increasing spatial separation, the strong force increases. Energy which has gone into separating two quarks reaches a critical point (at a distance of ≈ 1 fm) where it is transferred into producing more quarks which accompany the separated quarks to form hadrons. This cascade of separated quarks, or parton shower, into hadrons is called *hadronisation*. This is the principle of *confinement* and it is the reason that colour doublets or octets are never found in nature, only colour singlet states such as mesons and baryons. This is in contrast, again, to QED where free particles can carry electric charge.

2.2 Proton-proton collisions

At the Large Hadron Collider (LHC) the particles involved in the collisions are protons[1], which are complex composite particles consisting of three valence quarks (two up quarks and one down quark) and gluons which exchange the strong force. The proton also contains 'sea quarks' which are quark-anti-quark pairs that come into and out of existence rapidly and continuously within the proton due to gluon colour field splitting.

Figure 2.4 shows the parton distribution functions for the proton. These are interpreted as the probability for a quark to be carrying a fraction, x, of the proton's momentum in the longitudinal direction.

Protons in the LHC may a) not interact at all and continue to be accelerated around the ring, b) interact via a soft scatter where the products mostly travel along the direction of the beam, c) participate in a hard interaction where two partons within the protons have a high energy collision in which the products travel transverse to the beam. In the latter case, the remaining partons which have not participated in the hard interaction hadronise and form what is known as the *underlying event* (UE).

---

[1] There are also collisions with lead ions but they are not considered in this thesis.

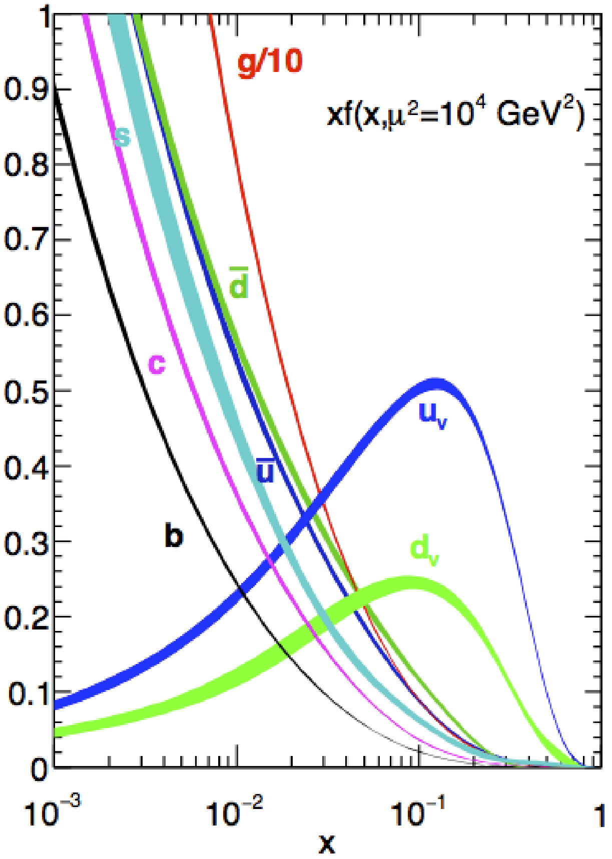


Figure 2.4: Proton parton distribution functions $xf(x)$ ($f = u_v,\ d_v,\ \overline{u},\ \overline{d},\ s \approx \overline{s},\ c \approx \overline{c},\ b \approx \overline{b},\ g$) for a given momentum fraction, x. The fermions are considered to be sea quarks except in the case of u_v and d_v, which are valence quarks. The gluon contribution has been scaled by a factor of ten for visibility. Obtained from NNLO NNPDF3.0 [20].

2.3 Top physics

It had been shown previously that CP-violation was not possible in a model which contained only two generations of quarks but that it required at least three generations [21]. When the bottom quark was discovered it was the only lone quark not contained in a weak isospin doublet and hence it was hypothesised that there must exist a partner to it, the top quark. The top quark mass was initially assumed to be much lighter than it is known to be currently. However, the ARGUS collboration found that $\mathrm{B}^0 - \overline{\mathrm{B}^0}$ mixing was much larger than expected, which implied that the top quark mass ($\mathrm{m_t}$), was larger than 50 GeV [22, 23]. The top quark was discovered in 1995 at the Tevatron by the CDF [24] and D0 [25] collaborations. It is the heaviest quark with a mass of $173.21 \pm 0.51 \pm 0.71$ GeV [6], approximately equivalent to the mass of a rhenium atom[2]. The top quark is the only quark which predominantly decays before it can form any bound states, due to its short lifetime of 5×10^{-25} seconds [6] and hence it is the only quark that can be studied for its spin and polarisation properties. The main decay mode for top quarks is to a bottom quark and a W boson, as shown in Fig. 2.5, which has a 99.8 ± 3.8 (exp.) ± 1.6(theo.)% [26] probability of occurring.

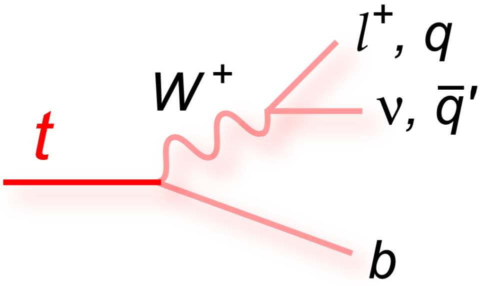


Figure 2.5: Top quark decay to a W boson and b-quark with subsequent decay of the W boson either leptonically or hadronically [27].

The top quark has the largest Yukawa coupling to the Higgs boson which is of the order of unity. The value of the top quark Yukawa coupling is important in calculations of the stability of the universe and of the energy scales where new physics may arise [28].

2.3.1 Top quark pair production

The first observations of top quarks were made on analyses of top pair production ($t\bar{t}$) as this is the dominant mechanism for producing top quarks at colliders. Figure 2.6 shows the leading order tree-level production mechanisms via gluon fusion and quark-anti-quark annihilation.

There are three possible decay modes depending on how each top quark decays, as shown in Fig. 2.5: *hadronic* where both W bosons from the top decay to a quark and anti-quark, *semi-leptonic* where one W boson decays to $q\bar{q}$' and one W boson

---

[2] Atomic number, $Z = 75$

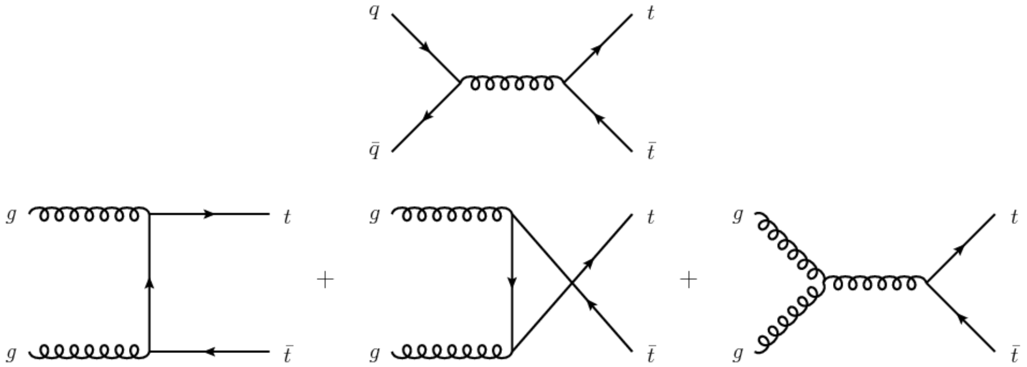


Figure 2.6: Representative diagrams of top quark pair production in the SM by quark-anti-quark annihilation (top) and via gluon fusion (bottom) at leading order [29].

decays to a lepton and a neutrino, and *dileptonic* where both W bosons decay to a lepton and a neutrino each.

2.3.2 Single top quark production

Single top quark production is much rarer than $\mathrm{t\bar{t}}$ production in the SM. It can occur via $\mathrm{q\bar{q}}$ annihilation, gq fusion or gluon fusion as shown in Fig. 2.7. In this figure the s-channel (left), t-channel (middle) and associated production with a W boson, tW-channel, (right) are shown.

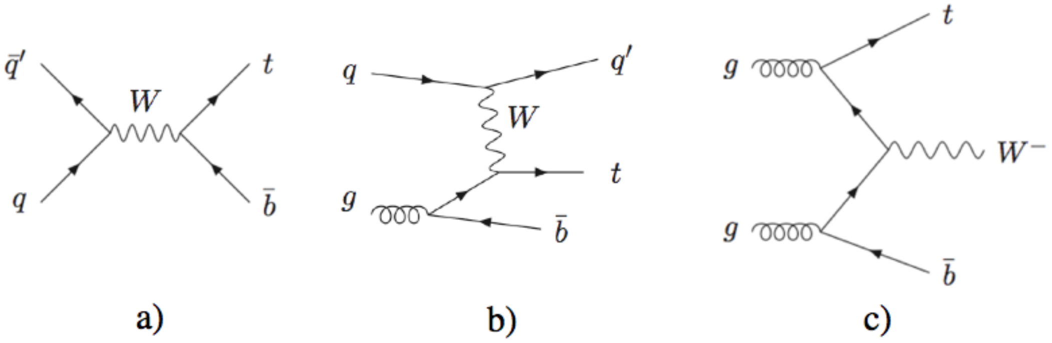


Figure 2.7: Representative diagrams of single top production at leading order in the a) s-channel , b) t-channel and c) tW-channel [30].

The CKM element $|V_{tb}|$ from Eq. 2.12 can be extracted from single top quark decays and the spin of the top quark can be ascertained by studying the leptonic decay of the single top as the charged lepton will point along the direction of the top spin [31].

2.3.3 Four top quark production

The production of four top quarks (t$\bar{\text{t}}$t$\bar{\text{t}}$) occurs predominantly via gluon fusion, as seen at leading order in Fig. 2.8, with a 10% contribution from quark-anti-quark annihilation. The production mechanism occurs predominantly via QCD, with smaller contributions from electroweak and Higgs boson mediated terms.

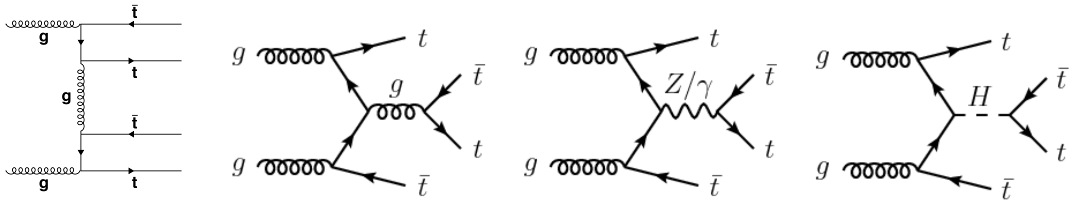


Figure 2.8: Representative diagrams of t$\bar{\text{t}}$t$\bar{\text{t}}$ production in the SM at LO [32].

Higgs boson mediated four top production, as seen in Fig. 2.8 (right), is interesting as the cross section is proportional to the fourth power of the top quark Yukawa coupling. Hence, limits can be set on the top quark Yukawa coupling using limits placed on four-top-quark production. The Higgs mediated process can have notable interference with the other terms of O(10%) [32]. Final states are determined upon whether the weak decay of the W bosons occurs leptonically or hadronically.

2.4 Shortcomings in the standard model

The SM has been resilient to many tests at the LHC and previous collider experiments. However there are many questions about the universe which the SM cannot answer, for instance:

Gravity : How can gravity be integrated into the SM and is there an associated boson for the gravitational force [33, 34]?

Matter-anti-matter asymmetry : How did the matter-anti-matter asymmetry in the universe arise when the amount of CP-violation in the SM is not sufficient to account for it [35]?

Hierarchy problem : The SM does not provide a solution as to why the gravitational force is so much weaker than the other forces. The gravitational coupling constant is $O(10^{42})$ smaller than the fine structure constant. Alternatively, one can consider how the masses of weak gauge bosons are $O(10^{16})$ smaller than the Planck Mass. The masses of the weak gauge bosons are determined by the Higgs VEV which lies at the relatively small but non-zero value of 246 GeV. The quadratic divergences in the SM calculation of the Higgs VEV suggest that the Higgs field is unstable and should be either zero or at the Planck Energy, ie. $O(10^{16})$ larger than it is. The Higgs VEV is considered the be *unnatural* and requires “fine tuning” [36].

Neutrino mass : The Sudbury Neutrino Observatory (SNO) observed that the flux of electron neutrinos from the sun was $\approx 1/3$ what it was expected to be [5]. This can be explained by neutrino oscillations - $\frac{2}{3}$ of the neutrinos which originated from the nuclear reaction in the sun oscillated into muon and tau neutrinos before reaching the Earth. This means that their mass basis is rotated from their weak-flavour basis and hence neutrinos must have mass, which is not part of the SM.

Dark matter : Observations of the universe show there is non-baryonic, non-luminous matter which is not accounted for within the SM [37]. Studies of galaxy rotation curves show that the angular velocity is relatively constant as a function radius when it should decrease, meaning that there is additional dark matter providing a contribution to the mass of galaxies [38–40]. The presence of dark matter can also be inferred by gravitational lensing. The light from distant stars in the background is curved around the strong gravitational presence of a dark matter cloud, which itself cannot be seen [41, 42]. This can cause arcs of light and repeated patterns of the same galaxy as seen in Fig 2.9.

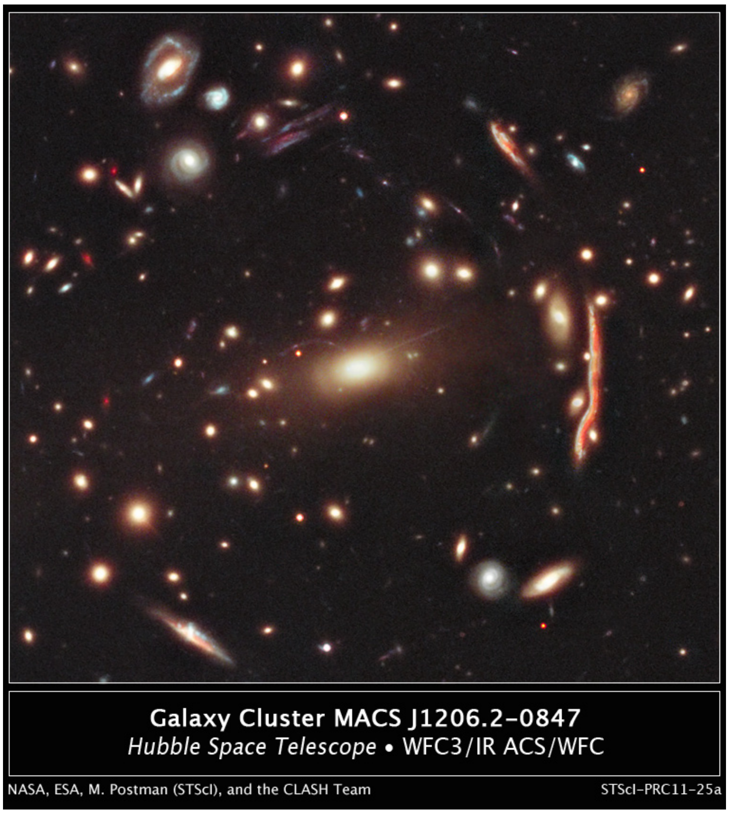


Figure 2.9: Gravitational lensing around MACS 1206 as captured by the Hubble Space Telescope [43].

2.5 BSM models with four top quark signatures

There are many theories which try to solve some or all of the problems listed in Section 2.4. Some theories take a simplified approach rather than using a full Ultraviolet Complete model. These approaches include using an Effective Field Theory (EFT) or Simplified Model, as discussed below.

2.5.1 Effective field theories

An effective field theory can be used in the low-energy limit of a more complete underlying theory. It is used to make observations on final state particles without assumptions on the intermediate particles involved. One example of this is Fermi's theory of beta decay, which was studied before the W boson was discovered. The intermediate W boson particle is replaced with a four-point interaction, as shown in Fig. 2.10. Effective field theories work best when there is a large separation between the energy scale being probed and the energy scale of new physics. In the case of the Fermi interaction, the separation between these two scales was three orders of magnitude when it was first studied, and hence it was a valid approximation.

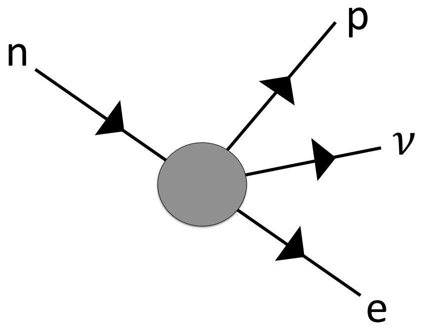


Figure 2.10: Fermi interaction for beta decay in an EFT.

2.5.2 Simplified models

Simplified models work by building a TeV-scale effective Lagrangian which describes a minimal number of new particles and their interactions. Hence, variables which can be directly observed by the detector can be studied, for example, particle masses, cross sections and branching ratios for different decay modes. Simplified models can sometimes be considered to be a subset of more general physics models where only a few of the particles are considered. Therefore, a simplified model cannot be considered to be model independent but they can help to identify the bounds of sensitivity for searches [44]. An example of where a simplified model can be used is in the case of $\mathrm{t\bar{t}}+\mathrm{X}$, as shown in Fig. 2.12, where the phenomenology of the resulting particles in the detector can be studied without knowledge of which underlying theory the new scalar particle, X, is coming from. In the limit of a large mediator the simplified model is equivalent to an EFT [45].

Some BSM theories which have final states containing four top quarks are discussed below.

Top quark Compositeness

It has been hypothesised that the top quark could be a composite particle made up of subparticles named preons, which are bound by a new confining force. Phenomenological studies have been performed in which an EFT is proposed where only the right-handed top quark is considered to be composite. It is argued that if only t_R is composite and no other SM component is, then the four top operator (shown in Fig 2.11) will be the most significant component of the EFT lagrangian. This can lead to an enhancement of $\approx 10^3$ to the production of $\mathrm{t\bar{t}t\bar{t}}$ compared to the SM rate [46, 47].

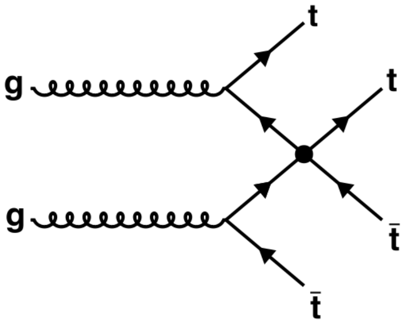


Figure 2.11: Four top interaction in an EFT [46].

$\mathbf{t\bar{t}} + \mathrm{X},\ \mathrm{X} \to \mathbf{t\bar{t}}$

There are several models which contain $\mathrm{t\bar{t}}$ plus an extra scalar particle which then decays to $\mathrm{t\bar{t}}$ as seen in Fig. 2.12. The mediator could be a dark matter mediator [48], a heavy Higgs boson [49], or a member of a scalar colour sextet [50], for instance.

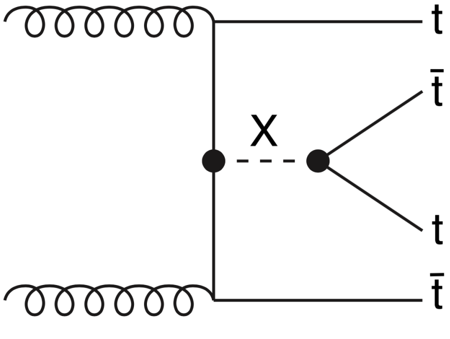


Figure 2.12: Four top production with an intermediate scalar [48].

Scalar pair production to $t\bar{t}t\bar{t}$

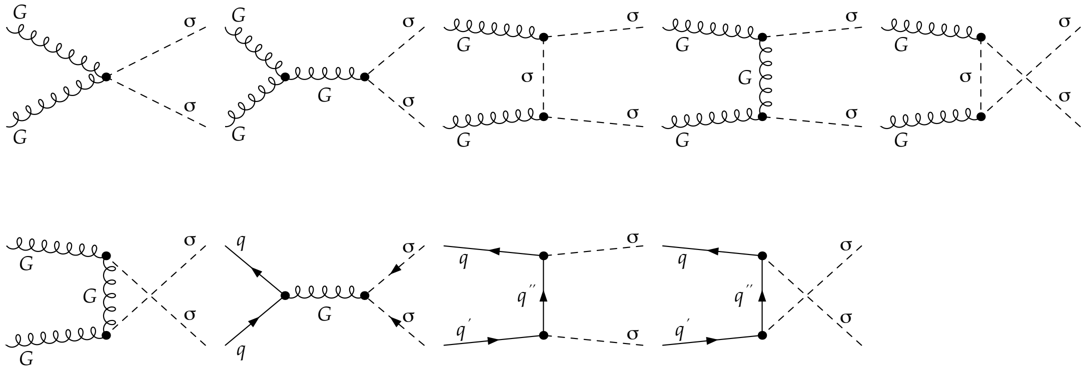


Figure 2.13: Sgluon pair production [51] where G represents gluons and σ represents sgluons.

An additional scalar gluon (*sgluon*) has been theorised in several models of physics beyond the SM. In N=1/N=2 hybrid [52, 53] and R-symmetric [54–56] versions of non-minimal supersymmetric models, the minimal supersymmetric model (MSSM) is supplemented by an additional chiral multiplet which lies in the adjoint representation of the QCD gauge group. This supermultiplet contains a two-component fermionic part which mixes with the Dirac gluino and a colour-octet complex scalar particle which is the sgluon field. This is particularly interesting because coloured particles will couple directly to gluons and hence should be produced in proton-proton collisions. Figure 2.13 shows the possible tree level production modes for sgluon pair production [51], whilst Fig. 2.14 shows each sgluon coupling to a $t\bar{t}$ pair resulting in a final state with four top quarks. Sgluons also arise in vector-like

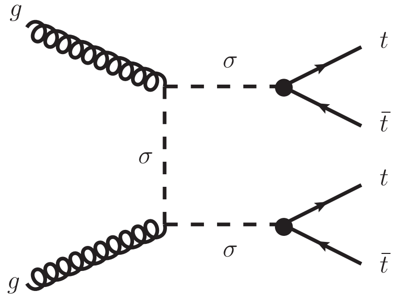


Figure 2.14: Representative diagram of sgluon pair production to four-top-quark final state [57].

confining theories [58] and extra-dimensional models [59]. Therefore it is viable to use a simplified model approach because the final state signatures are reasonably model-independent. Colour octets and sextets can also be pair produced in theories where the Higgs boson is a composite particle [50].

3 The CMS detector and the Large Hadron Collider

This chapter discusses the Large Hadron Collider (LHC) which is located on the Franco-Swiss border near Geneva, approximately 100 m underground at the site of the Conseil Européen pour la Recherche Nucléaire (CERN)[1]. It is a 26.7 km long synchrotron particle accelerator with four interaction points where four experiments are located. This thesis focuses on results from the Compact Muon Solenoid detector described in Section 3.2. The other experiments include ATLAS which is a multi-purpose experiment like CMS, the LHCb detector which focuses on the study of the physics of B hadrons, and the ALICE detector which is used to study the quark-gluon plasma. Details of the function and purpose of each of the CMS sub-detectors are given in this chapter. The algorithms which are used to reconstruct particles from the information given from the sub-detectors are given in Chapter 4.

3.1 LHC

The LHC accelerates two beams of protons which circulate in opposite directions. This is achieved using a system of superconducting dipole magnets which have an aperture for each beam direction. Quadrapoles are used to squeeze the beam. The protons are sourced from a bottle of hydrogen where a strong electric field is used to excite the electrons from the hydrogen atoms leaving protons behind. The protons are then accelerated through the linear accelerator LINAC2, where they reach an energy of 50 MeV. The LINAC2 uses radio frequency cavities to accelerate the protons. From LINAC2, the protons are accelerated by the Proton Synchrotron Booster (PSB) to 1.4 GeV, then they are accelerated by the Proton Synchrotron (PS) to 25 GeV. This is followed by a boost to 450 GeV in the Super

---

[1] This chapter is largely adapted from Ref. [60]

Proton Synchrotron (SPS), which is the final accelerator before the protons are injected into the LHC ring where their final collision energy can be achieved.

The protons are accelerated in bunches of O(10^{11}) protons, which are collided at the interaction points of each experiment. This results in many collisions per bunch crossing despite the fact that many protons will miss each other and continue to be accelerated around the LHC. The background of particles coming from the frequent less interesting collisions is called *pileup* (PU). The multiple interactions per bunch mean that PU can come from the same bunch crossing (in-time PU) or overlap from a previous bunch crossing (out-of-time PU).

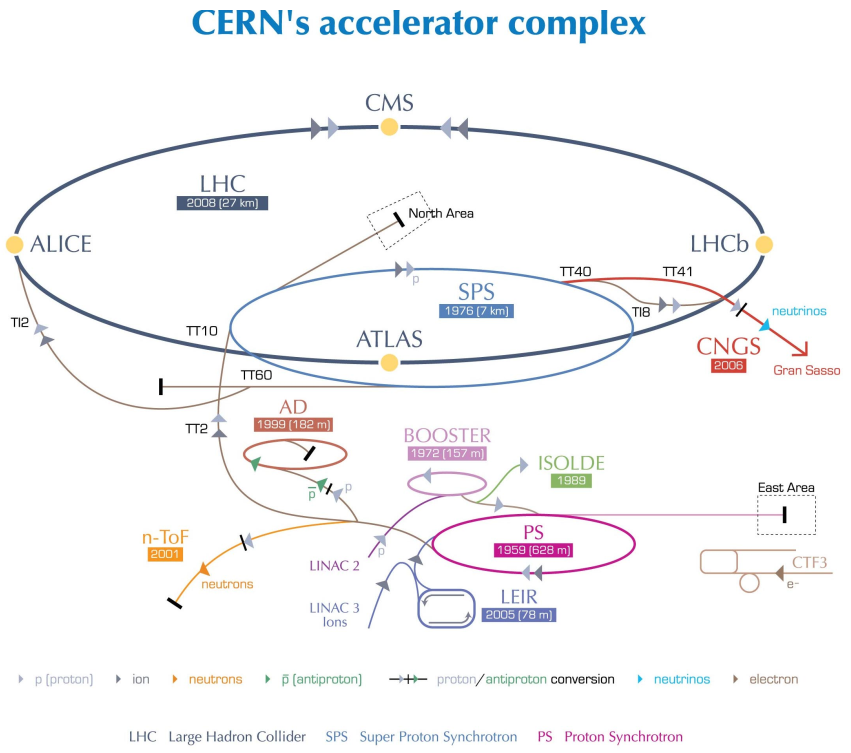


Figure 3.1: The LHC accelerator complex at CERN. Protons are accelerated from LINAC 2 into the BOOSTER synchrotron. From there they are further accelerated in the proton synchrotron (PS) and super proton synchrotron (SPS) before finally being injected in two counter-rotating beams in the large hadron collider (LHC). The beams are crossed at the four experiments: CMS, LHCb, ATLAS and ALICE [61].

The *Luminosity* ($\mathcal{L}$) is a measure of the instantaneous collision rate and can be calculated using Eq. 3.1, where f is the bunch frequency, and N_1 and N_2 are the numbers of particles in each bunch. The effective collision area for a gaussian distributed beam is $4\pi\sigma_x\sigma_y$. The widths σ_x and σ_y are for the x and y components of the beam, transverse to the beam direction. It is assumed that each beam has the same cross section.

$$\mathcal{L} = \frac{f N_1 N_2}{4\pi\sigma_x\sigma_y} \tag{3.1}$$

The *integrated luminosity* ($\mathcal{L}_{\text{int}}$) which is shown in Fig 3.2 for the CMS experiment from 2010 until 2016 for proton-proton collisions is the delivered luminosity integrated over time. The number of events, N_{events}, produced for a particular particle physics process can be calculated from the cross section, σ, using Eq. 3.2.

$$N_{\text{events}} = \mathcal{L}_{\text{int}} \times \sigma \tag{3.2}$$

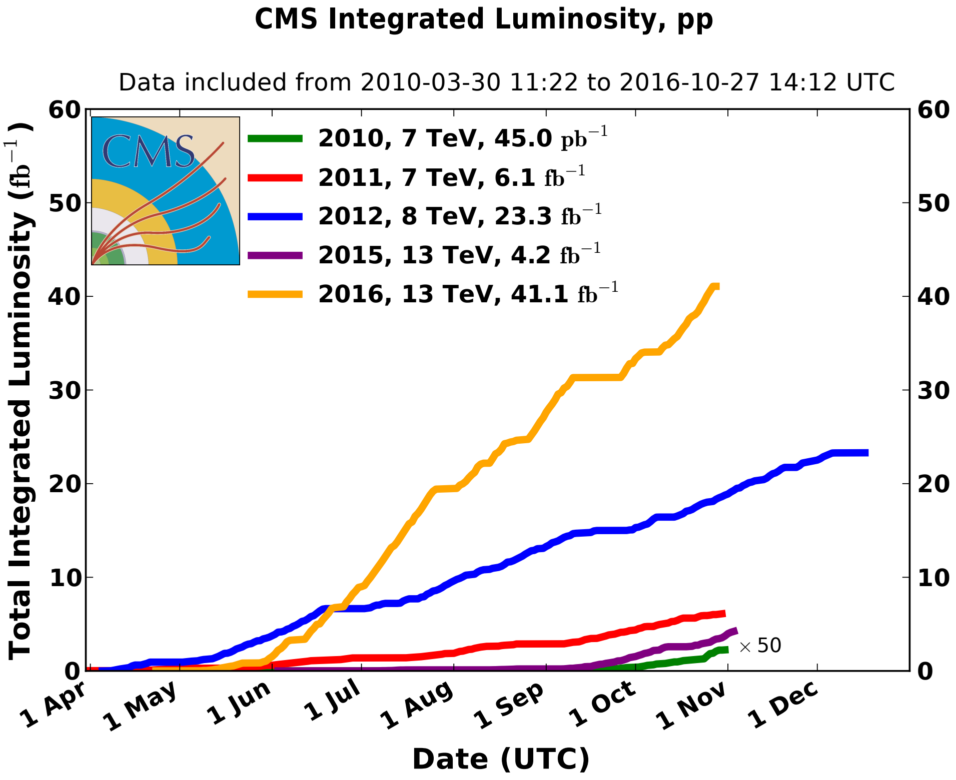


Figure 3.2: The integrated luminosity in fb^{-1} for proton-proton collision at the CMS experiment from 2010 to 2016 [62].

The LHC was designed to have a centre of mass collision energy ($\sqrt{s}$) of 14 TeV. The intention was to start the machine with a lower energy in September 2008 and to obtain $\sqrt{s}$ = 10 TeV by the end of 2008. However, an electrical fault 10 days into operation in October 2008 caused damage to over 50 of the superconducting magnets. The LHC was shut down for repairs until November 2009 when LHC achieved the record breaking 1.18 TeV per beam. The centre of mass energy was then ramped up to $\sqrt{s} = 7$ TeV in 2011. In 2012 this was increased to $\sqrt{s} = 8$ TeV, at which time the bunches of protons were crossed every 50 ns. A dataset with an integrated luminosity of ≈ 20 fb^{-1} was recorded. This dataset from the run phase known as ‘Run 1’ was used for the analysis in Chapter 7. In 2013 after Run 1, Long Shutdown 1 (LS1) commenced to make upgrades to the LHC and the detectors to allow the machine to run at $\sqrt{s} = 13$ TeV with a bunch crossing time of 25 ns for ‘Run 2’. Run 2 began in March 2015 and results from Run 2 are the focus of the analysis in Chapter 9.

Run 1 saw the great success of the discovery of the Higgs boson, one of the main objectives of the LHC [63, 64]. In Run 2, the search for new physics continues where precision measurements will test the predictions of the SM. The CMS and ATLAS detectors are considered to be general-purpose detectors which can be used to test many areas of physics within and beyond the standard model. This includes searches for dark matter, supersymmetric particles, vector-like-quarks, lepton-flavour-violating processes, light Higgs and charged Higgs bosons, and studies of the properties of the Higgs boson such as the couplings and rare decay modes.

3.2 CMS detector

The CMS detector is a hermetic detector with a large magnetic solenoid which causes charged particles to follow a curved trajectory as they traverse the detector. Closest to the beam line is the silicon tracker which makes the most accurate position measurements. Next are the calorimetry systems for electromagnetic and separately for hadronic particles. All of these detectors are contained within the magnetic solenoid. The muon chambers are outside the solenoid where they detect muons, which are much more penetrating than other particles. The cylindrical coordinate system of the detector is defined as follows: the x-axis points towards the centre of the LHC ring, the y-axis points upwards and the z-axis points along the beamline in the anti-clockwise direction. The azimuthal angle ϕ is measured in the (x, y) plane clockwise from the x axis. The polar angle θ is measured clockwise from the z-axis. More commonly the pseudorapidity, defined in Eq. 3.3, is used instead of θ.

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right] \tag{3.3}$$

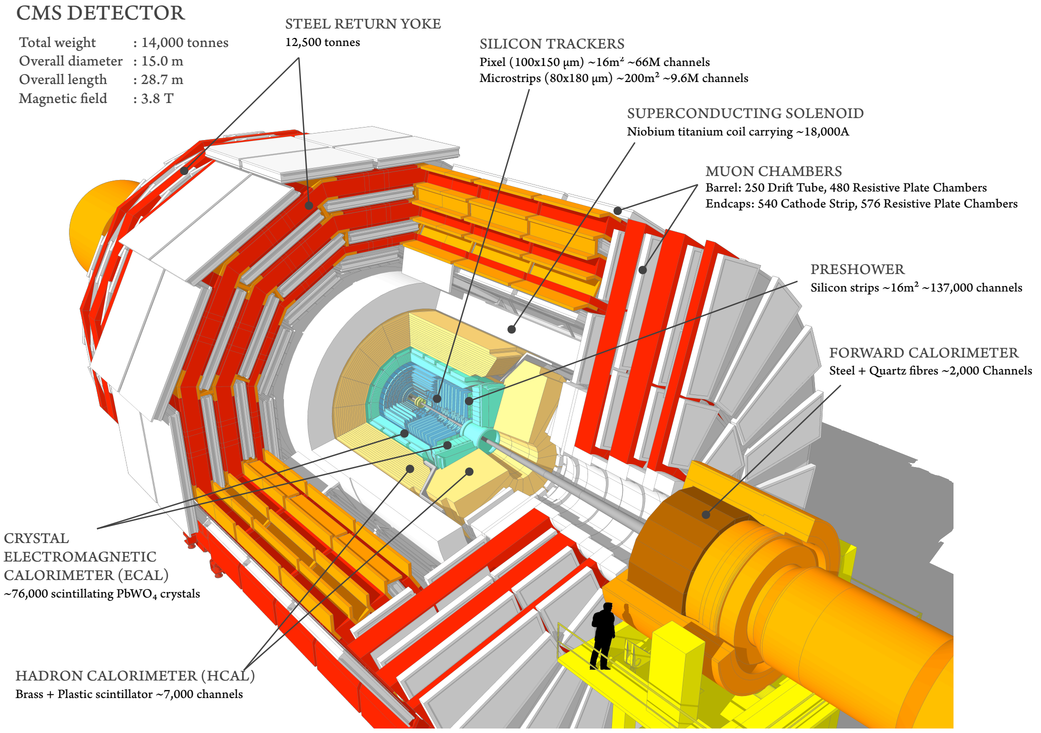


Figure 3.3: The CMS detector [65].

3.2.1 Magnetic Solenoid

The superconducting magnetic solenoid at the core of CMS was designed to have a magnetic field of 4 T. The free bore magnet has a diameter of 6.3 m and length of 12.5 m. It uses a 4-layer winding of NbTi superconductor which is required to generate this high magnetic field. The magnet is cooled within a cryostat using liquid helium to a temperature of 4.5 K. The magnetic field is returned via an iron yoke consisting of five barrel wheels and two endcaps which are made of three discs each. The outer dimension of the iron flats is 14 m. The tracker, electromagnetic calorimeter and hadronic calorimeter are constrained to be within the inner dimensions of the solenoid as can be seen in Fig. 3.3. The solenoid provides a homogeneous magnetic field, which bends particle trajectories transverse to the beam direction, over the silicon tracker region shown in Fig. 3.4.

3.2.2 Tracker

The tracking system of CMS lies inside the superconducting solenoid and surrounds the interaction point. It is 5.8 m long with a diameter of 2.5 m. Silicon detectors are used as they can provide the high granularity and fast response required to reliably reconstruct the trajectories of particles coming from the collision vertex. Reconstructing secondary vertices is also particularly important for identifying jets originating from heavy flavour quarks such as bottom quarks. This is integral for distinguishing final states involving top quarks.

It is estimated that at a PU of 20 there are 1000 particles traversing the tracker at every bunch crossing. The silicon detectors have been designed to have the radiation hardness to last for the design time of ten years. The minimum material possible was used in order to reduce the amount of multiple scattering, photon conversion and bremsstrahlung. Cooling the tracker helps to prevent thermal runaway from leakage current, hence it was cooled to 0°C in Run 1 and -20°C in Run 2. The tracking system has a nominal momentum resolution of 0.7 (5.0)% for a particle with a momentum of 1 (1000) GeV in the central region. The impact parameter resolution is around 10 μm for high momentum tracks [66].

The tracking system consists of two main sections: the pixel tracker makes up the innermost section and the strip tracker surrounds the pixel tracker as seen in Fig. 3.4.

Pixel tracker

As the pixel detector is closest to the interaction vertex, it experiences the highest flux of particles at ≈ 1 MHz per mm^2. The fine granularity of the pixel detector

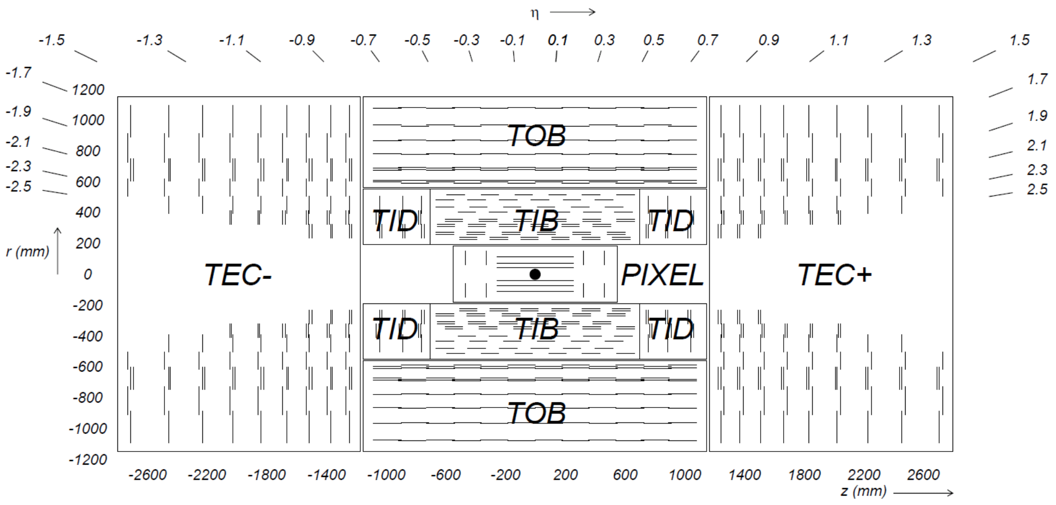


Figure 3.4: The tracking system [67].

(100 x 150 μm in $(r-\phi)$x$(x-z)$) is required in order to keep the occupancy below 1%. It consists of three barrel layers which range between 4.0 cm and 10.2 cm from the interaction point and 2 disks which are transverse to the beamline as seen in Fig. 3.4.

Strip tracker

Two types of silicon strip tracker are used. Closest to the interaction point (20 - 50 cm) are the tracker inner barrel detectors (TIB) which contain silicon microstrips (10 cm x 80 μm). An occupancy of $\approx$ 2-3% is achieved for a fluence of $\approx$ 60 kHz per mm^2. An increased strip pitch of 180 μm can be used in the tracker outer barrel (TOB) due to the lower fluence of 3 kHz per mm^2. To cover the larger surface area, the effective strip length is increased to 25 cm. However, increased strip length increases the noise. To combat this the strips are made thicker to 500 μm compared to 320 μm in the TIB. The tracker inner disk (TID) and tracker end cap (TEC) have strips which are aligned radially to the beamline with a strip pitch of 80 μm and 200 μm, respectively. The TID and TEC extend the tracker acceptance to $|\eta| < 2.5$.

3.2.3 Electromagnetic Calorimeter

The electromagnetic calorimeter (ECAL) is a homogeneous, hermetic detector made up of 61200 (7324) lead tungstate, $PbWO_4$, crystals in the barrel (endcap) region. In the barrel region the light from these scintillating crystals is collected in avalanche photodiodes and in the endcap region by vacuum phototriodes. The barrel region covers $|\eta| < 1.479$ whilst the two semi-circular 'Dees' which make up the endcaps extend the range to $|\eta| < 3.0$.

Good resolution can be obtained from $PbW0_4$ crystals and they are fast and radiation resistant. Equation 3.4 shows the dependence of the resolution on the energy of the particle, E. The stochastic term for the statistical fluctuations on the number of secondary particles produced is represented as S. The noise from the electronics and digitisation is given by N and the constant term C arises from calibration errors and leakage of the shower outside of the calorimeter.

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C \tag{3.4}$$

For an electron test beam with no magnetic field and no material between the beam and the ECAL, the parameters in Eq. 3.4 were measured to be $S = 0.028$, $N = 0.12$ GeV and $C = 0.0003$ [68]. This gives an energy resolution of $\approx$ 0.5% for a 100 GeV particle.

The crystals have a high density and small radiation length ($X_0 = 0.89$ cm) so that the ECAL can be compact. Another important quality is that they are optically clear allowing all of the light to be collected. The scintillation decay time is comparable to the time between bunch crossings with 80% of light being emitted within 25 ns. The crystals have a tapered shape in the barrel region and lie parallel in the endcap region. The crystals are 1.29 m from the beam line in the barrel region and 3.15 m from the the interaction point in the longitudinal direction in the endcap region. The crystals are contained in thin-walled aluminium structures which make up submodules.

In the endcap region there is a preshower detector, shown in Fig. 3.5, which is a two-layer sampling calorimeter. There are two layers of lead used as a radiator material to initiate the electromagnetic shower and silicon strip detectors after each layer (orthogonal in each plane) to measure the energy deposited. The preshower detector helps to identify neutral pions and to distinguish electrons from minimum ionising particles.

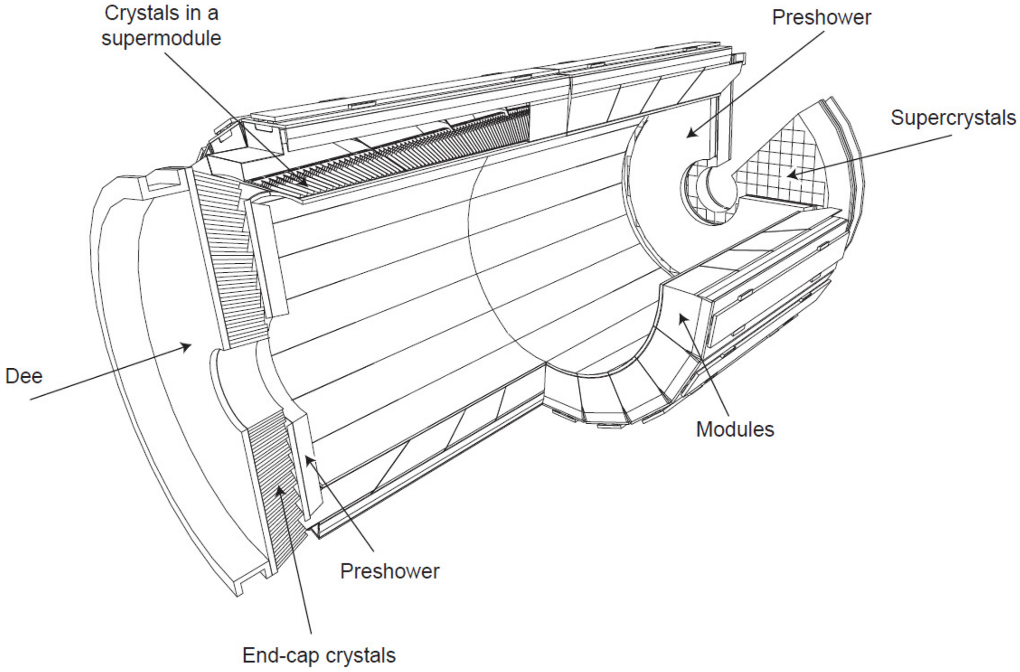


Figure 3.5: The ECAL system contains 61200 $PbWO_4$ crystals which are contained within modules. Each of the 36 supermodules shown are made up of four modules. The supercrystals in the endcap are made up of groups of 5 x 5 crystals, with a total of 7324 crystals [69].

3.2.4 Hadron Calorimeter

The hadron calorimeter (HCAL) is used for identifying hadron jets. It has barrel (HB) and endcap (HE) regions made up of sampling calorimeters which have coverage of $|\eta| < 1.3$ and $1.3 < |\eta| < 3.0$, respectively. The HB and HE are placed between the ECAL and solenoid magnet, and therefore are restricted to the radial dimensions, R, of $1.77 \text{ m} < R < 2.95 \text{ m}$. The scintillators in both the HB and HE have a granularity of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ (0.17×0.17) for $|\eta| < 1.6$ (≥ 1.6). The barrel consists of two halves, HB+ and HB-, which are composed of 36 wedges made up of 14 flat brass absorber plates parallel to the beam axis, alternated with plastic scintillator. Stainless steel plates are used for the innermost and outermost plates to provide structural support. The HE has a similar system of alternating absorber and plastic scintillator. Due to their restricted dimensions the HCAL and ECAL do not always contain all of the energy from the particle showers. Between five and ten interaction lengths are contained within the HB, depending on the pseudorapidity. Therefore another detector known as the outer hadronic calorimeter (HO) is embedded in the muon system, outside of the solenoid magnet, to measure the energy leakage from the HCAL and ECAL. This extends the combined thickness to around twelve interaction lengths. The HE contains around ten interaction lengths. The forward HCAL extends the range from $|\eta| < 2.3$ to $|\eta| < 5.2$ such that very forward jets can be detected. The hermeticity of the detector ensures good coverage on detecting the total hadronic energy and hence good resolution can be obtained on missing transverse energy which could come from neutrinos or BSM particles.

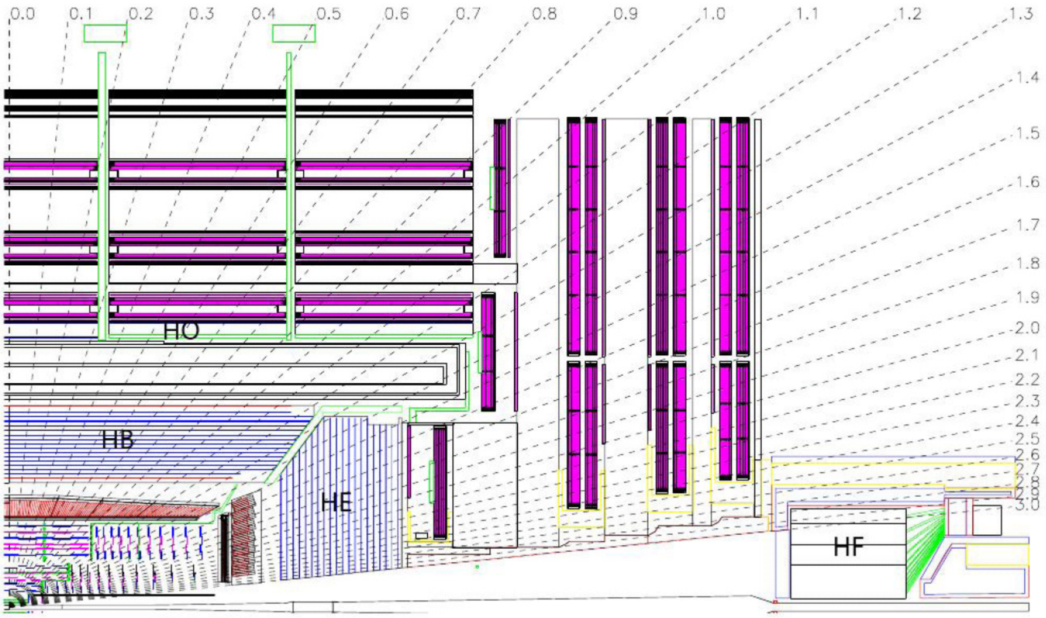


Figure 3.6: The HCAL system [69].

The energy resolution of the HCAL is $120\%/\sqrt{E(GeV)}$ for neutral hadronic particles [70]. Neutral hadronic interactions are the most important consideration for the energy resolution of the HCAL as, unlike charged particles, no additional information can be obtained from the tracker to combine with the calorimetry measurement to reduce the resolution.

3.2.5 Muon Chambers

As the name CMS suggests, muon identification and measurement was a central focus in the design of the detector. The muon chambers are interspersed within the iron flux return yoke. These thick layers of iron act as a hadron absorber. Muons are much less affected by radiative losses through the detector material than electrons, therefore they are able to penetrate through to the outermost layers of the detector. Figure 3.7 shows the layout of the muon chambers around the detector. The muon chambers consist of a cylindrical barrel section and 2 planar endcaps. The magnetic field is uniform in the barrel region where drift tubes (DT) are arranged into chambers, some of which make measurements in $r-\phi$ and some of which make measurements in z. This provides a high efficiency for matching individual hits in different stations to one single muon track. In the endcaps, where

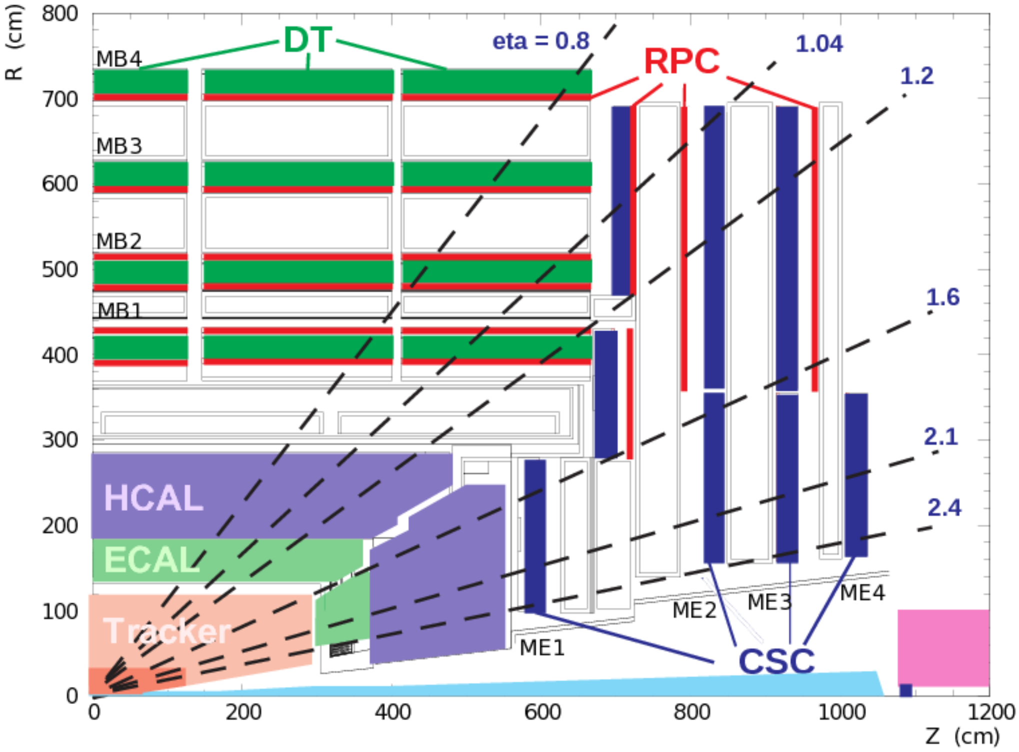


Figure 3.7: The muon chamber system [71].

the muon flux is high and the magnetic field is non-uniform, cathode strip chambers (CSC) are used due to their fine segmentation, fast response and radiation resistance. The CSC stations are aligned perpendicular to the beam line and are positioned between the flux return plates. The cathode strips are positioned as radial lines and the anode wires run perpendicular to the cathode strips. Together they can measure the position in $r - \phi$ and η. The CSCs provide robust pattern recognition for matching hits to other stations and to the tracker as well as for rejecting non-muon backgrounds. Together the barrel and end-cap region provide uninterrupted coverage up to $|\eta| < 2.4$.

Resistive plate chambers (RPC) are added in the barrel and endcap sections of the muon chambers. They consist of two parallel plate chambers which sandwich readout strips, known together as a double-gap module. The sum of the two signals from each gap creates one total induced signal. There are six RPCs in the barrel section and there are three layers of RPCs in the endcap. Where ambiguous tracks exist due to multiple hits in the muon chambers, the RPCs can help to distinguish the correct track. They have a fast response and can time-tag an ionising event in much less than the 25 ns bunch spacing time and hence they can assign candidate tracks to the relevant bunch crossing. The resolution of the RPCs is courser than the DTs and CSCs. This courser information from the RPCs is used in the trigger which is described in the following section.

The muon chamber measurements provide the dominant contribution to the energy resolution for high-momentum muons. For muons with low momentum the tracker provides the dominant contribution to the energy resolution when the tracks are more curved. At around 1 TeV both systems provide a momentum resolution of $\approx 5\%$. Figure 3.8 shows the muon transverse momentum resolution gained from using only the muon system, only the tracking system, and the combination of the two. It can be seen that above 200 GeV the combination of the information from the muon system with the tracking system improves the overall resolution compared to using the tracking system only.

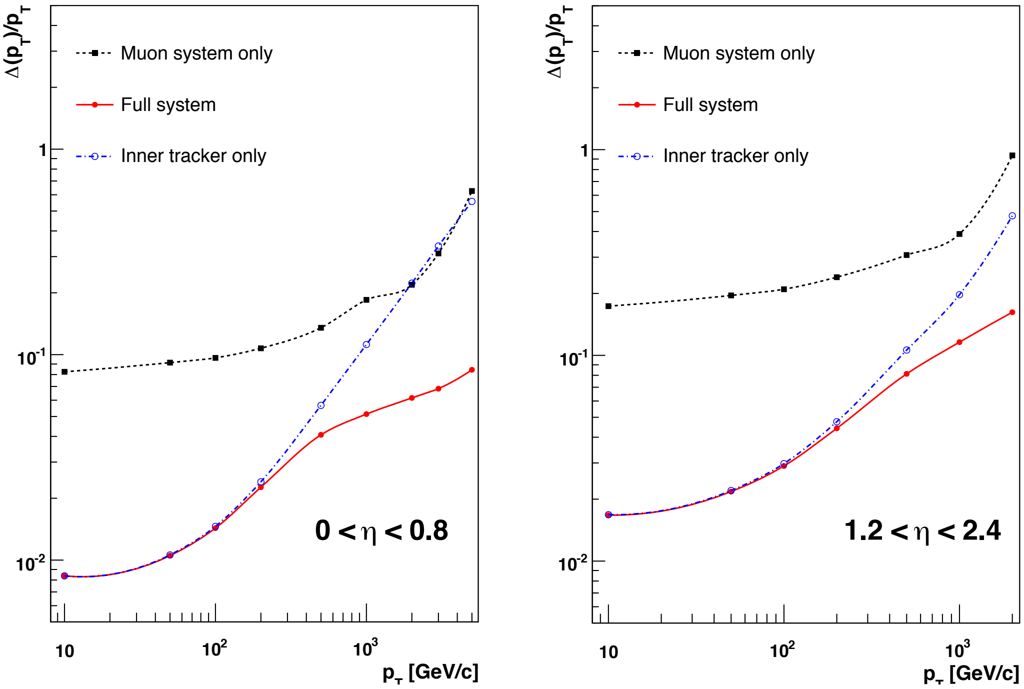


Figure 3.8: The muon transverse momentum resolution as a function of transverse momentum (p_{T}) using the muon system only (black), the inner tracking only (blue), and both (red), in regions of $|\eta| < 0.8$ (left) and $1.2 < |\eta| < 2.4$ (right) [67].

3.2.6 Trigger

The amount of information which can be stored is much less than the amount produced within the subdetectors. Collisions occur at a rate of 40 MHz (beam crossing interval of 25 ns) in Run 2. The rate reduction capability of the trigger system was designed to reduce the rate by at least a factor of 10^5. This occurs in two main stages, the Level-1 (L1) trigger and High-Level Trigger (HLT). The L1 trigger is composed of highly programmable custom electronics, mainly FPGAs. ASIC and programmable memory lookup tables are used where higher speed and radiation hardness is required closer to the beam spot. Full high resolution data are held in the front-end electronics whilst the L1 trigger decides whether or not to keep the event based on courser information from the calorimeters and muon chambers. The HLT is a software based filter farm ($\approx$ 1000 processors) which has access to all of the readout data from the subdetectors. The HLT can assess more complex information in order to filter out less interesting events and categorise the most interesting ones.

3.2.7 Upgrades for Run 2

In LS1, repairs were made to the LHC magnet splices to allow safe operation at the design energy of $\sqrt{s} = 14$ TeV. All of the detectors on the LHC ring were able to make essential repairs and upgrades. For CMS this included repairing damaged silicon pixels and strips in the tracker and inserting the tracker back into CMS with better centering around the beam line. The temperature of the tracker was lowered to -20°C to mitigate against radiation damage. In the ECAL the EE and ES subsystems underwent minor repairs. New photodetectors were added to the HO in the HCAL to improve the signal to noise ratio. The main change to the muon systems has been to the RPCs with the addition of the fourth disk (RE4). For the muon system CSCs, 72 chambers have been added to the existing 468 chambers. In addition to the detector upgrades, improvements have been made to the trigger and DAQ.

These upgrades have allowed CMS to operate at $\sqrt{s} = 13$ TeV in 2015 with the smaller bunch spacing of 25 ns compared to the previous 50 ns bunch spacing in Run 1.

3.2.8 Data collections

The data quality monitoring (DQM) group monitor incoming data online by reconstructing a small subset of the data immediately to ascertain whether all sub-detectors in CMS are operational. Offline they go through several stages of DQM as the data are reprocessed and reconstructed. Data which have been collected while all sub-detectors were working, the magnet was at full field and the beam conditions were stable are certified by the DQM group for use in physics analyses. Hence the amount of data used for analyses is less than the total collected, as in Fig. 3.2. Different run ranges are usually separated by short shutdowns for maintenance or they may have different run conditions such as different instantaneous luminosity [72].

The proton-proton collision datasets recorded by CMS and used in this thesis are given in Table 3.1 for the Run 1 data that are used in Chapter 7 and in Table 3.2 for the Run 2 data that are used in Chapter 9.

Dataset	Recorded	$\mathcal{L}$ pb^{-1}
Single Muon Run A	2012	888
Single Muon Run B	2012	4436
Single Muon Run C	2012	7125
Single Muon Run D	2012	7426
Total		19695
Single Electron Run A	2012	876
Single Electron Run B	2012	4420
Single Electron Run C	2012	7132
Single Electron Run D	2012	7294
Total		19721

Table 3.1: Run 1 datasets at 8 TeV, when they were recorded and how much data were recorded.

Dataset	Recorded	$\mathcal{L}$ pb^{-1}
Single Muon Run C	2015	17.2
Single Muon Run D	2015	2611.5
Total		2628.7
Single Electron Run C	2015	17.2
Single Electron Run D	2015	2611.5
Total		2628.7

Table 3.2: Run 2 datasets at 13 TeV, when they were recorded and how much data were recorded.

4 Event Reconstruction

In Chapter 3, each of the sub-detectors in CMS have been described; how particles interact with them and how electrical signals are read out. The next step is to combine the readouts from each detector in order to reconstruct the resulting particles from an interesting proton-proton collision. This snapshot of the collision output is known as an *event*. An event will also contain PU from reconstructed particles from other simultaneous uninteresting collisions from the same or previous bunch crossing. Algorithms are used in order to subtract PU particles from the stored event. As a particle will usually traverse more than one sub-detector, it is advantageous to combine these outputs in order to reconstruct and identify the particle. This is achieved using the *particle-flow* (PF) algorithm described in Section 4.3. The objects which can be reconstructed using the PF algorithm such as muons, electrons, and jets are discussed in Sections 4.5, 4.6, 4.7 respectively. Further information can be obtained from these reconstructed objects such as how likely a jet is to have originated from a b-quark (Section 4.8) and how the presence of neutrinos can be inferred by the imbalance of energy in the transverse plane of the detector (Section 4.9). The information from the detector is processed using a distributed computing infrastructure with custom software made by CMS, CMSSW.

4.1 Track reconstruction

Approximately 1000 charged particles are expected to traverse the CMS tracker at each bunch crossing at a PU of ≈ 20 concurrent collisions. Each charged particle will interact with the silicon tracker as it continues through its trajectory from the collision point. Algorithms are designed to match hits in the tracker along each particle's trajectory in order to reconstruct its path so that information about the charge and momentum of the particle can be obtained. Not only is the tracker information used in offline reconstruction but it is used in the HLT, therefore it must have a fast response. Reconstructed paths from random particle hits in the tracker are considered to be *fake tracks* [73].

Knowledge of particle trajectories can help to pinpoint the collision vertex of interest which caused the trigger to fire. This is known as the *primary vertex* and is described further in Section 4.2. Accuracy in reconstructing tracks is essential for b-tagging as described in Section 4.8. Electrons lose energy through the tracker material in a non-gaussian way such that their tracks can not be fitted using the standard Kalman Filter. A Gaussian-Sum-Filter refit [74], which uses a sum of gaussians to estimate the energy loss, takes into account the interaction of electrons through the tracker material.

4.2 Primary vertices

Primary vertices are the point at which the collision occurred, as opposed to secondary vertices which originate at the decay of subsequent particles coming from the collision. The first step in reconstructing primary vertices is to consider tracks which are consistent with the beam spot and cluster them into candidate vertices, separated along the z direction. Next a 3D fit is made and candidates which are compatible with originating from the beamline are kept [75]. Primary vertices are ranked according to the sum of the momentum squared of all the tracks considered to have originated from that vertex. The vertex with the largest sum is regarded as the signal vertex, ie. the most interesting event, with higher momentum objects, that is most likely to have fired the trigger.

4.3 Particle-flow algorithm

The particle-flow (PF) algorithm combines information from all sub-detectors described in Section 3, in order to improve the reconstruction of ‘final state’ particles such as electrons, muons, photons, neutral hadrons and charged hadrons. Using this information more complicated higher-level objects, such as jets, can be reconstructed as described in the subsequent sections. Collections of different subdetector objects, such as tracks or ECAL/HCAL hits, are created and each object is subsequently removed from each collection as they are identified as belonging to a final state particle within the algorithm.

The particle-flow algorithm reconstructs objects in an order starting from the easiest to reconstruct unambiguously. The hardest objects to reconstruct such as neutral hadrons are one of the last to be reconstructed because their properties can be constrained from the previously reconstructed objects. The first objects to be reconstructed are PF muons. Each muon identified from the muon chambers is associated to compatible hits in the tracker. This associated track and muon chamber hits are then removed from their respective collection. The muon behaves like a minimum-ionising particle (MIP) in the ECAL and HCAL. With this assumption, energy can be subtracted from the ECAL and HCAL deposits (of the order of a few GeV) where the muon track is interpolated through these subdetectors. Next, a Gaussian-Sum-filter refit is used to extrapolate electron candidate trajectories to the ECAL. On average, electrons have shorter trajectories than muons due to losing much of their energy through interactions with the tracker material. Tracker and ECAL variables are combined for the final identification of a PF electron after which the track and ECAL clusters are removed from their respective collections.

Charged hadrons are reconstructed from the remaining tracker, ECAL and HCAL deposits where the calorimeter hits are compatible with the tracker hits. Again, these hits are removed from the respective collections. Neutral hadrons leave no tracks in the tracker but have deposits in the ECAL and HCAL. Photons leave deposits in the ECAL but not the HCAL.

4.4 Isolation

Relative isolation (RelIso) is a measure of how isolated the muons or electrons are from surrounding hits in the detector from charged hadrons, neutral hadrons and photon energy which could contribute to a mis-measurement of their momentum. Only charged hadrons which are consistent with the signal primary vertex are considered in the calculation. As it is not possible to determine whether neutral hadrons are consistent with the signal primary vertex, instead the fact that the ratio of neutral hadronic to charged hadronic energy has been measured to be ≈ 0.5 can be used[1]. Hence the neutral hadronic energy in the transverse plane of the detector coming from the primary vertex, $E_{T,PV}^{NH}$, can be calculated as seen in Eq. 4.1, where $E_{T,sub}^{CH}$ is the transverse energy from charged hadrons which are associated to a sub-leading primary vertex and $E_{T,Tot}^{NH}$ is the total transverse neutral hadronic energy. The $max()$ function ensures that the corrected neutral hadronic energy is never defined as negative.

$$\Sigma E_{T,PV}^{NH} = max(0, \Sigma E_{T,Tot}^{NH} - 0.5 \times \Sigma E_{T,sub}^{CH}) \tag{4.1}$$

The RelIso formula for muons, with the correction to neutral hadronic energy, can be found in Eq. 4.2. The energy from photons is denoted as ΣE_T^{γ}. It is defined in a cone of radius R = 0.4 and scaled by 1 / p_{T}^{μ} (where p_{T} is the momentum in the transverse plane of the detector) so that lower momentum muons are required to have less energy from hadrons and photons in the cone to be considered isolated.

$$RelIso = \left(\Sigma E_T^{CH} + \Sigma E_{T,PV}^{NH} + \Sigma E_T^{\gamma}\right) \ / \ p_{\mathrm{T}}^{\mu} \tag{4.2}$$

---

[1] This is consistent with the π^0, π^+, and π^- mesons (which constitute the majority of particles produced from p-p collisions) existing in an isospin triplet such that they have equal probability of being produced in a parton shower

For electrons RelIso is similarly defined, however, the neutral hadronic energy is estimated using a different method. The relative isolation for electrons is defined in Eq. 4.3 where EA denotes the Effective Area used in the analyses in this thesis. The EA is calculated by CMS from the average PU energy density per unit area in the ρ ($\rho = \phi - \eta$) plane and the effective area based on shower shapes that has been measured by CMS (which depends on the η value in the supercluster) for each event.

Equation 4.3 gives the RelIso formula for electrons with neutral hadronic energy correction where ρ stands for the median density of pile-up contamination.

$$RelIso = \Big(\Sigma E_T^{CH} + \Sigma E_T^{NH} + \Sigma E_T^{\gamma} - (\rho \times \mathrm{EA})\Big) \ / \ p_{\mathrm{T}}^{e} \tag{4.3}$$

4.5 Muons

It is important to be able to identify isolated muons coming from the signal process rather than from further decays from within jets or from mismatched tracks. Applying the identification criteria in Table 4.1 can help to ensure a high purity of real muons is selected from the PF candidates described in Section 4.3. Two working points (WP) are defined: tight and loose. Tight muons have tighter requirements on various quantities including p_{T}, as lower momentum muons are harder to distinguish from other particles. Tight muons are used when making selection requirements on how many muons should be in the event from the signal process. Loose muons are used to veto additional objects which are still likely to be muons but could be a misidentified object such as a pion. The loose criteria will capture more real muons in its selection but with a lower purity. The cut values for each working point are given in Table 4.1 for the 8 TeV analysis and the 13 TeV analysis.

The transverse impact parameter in the ϕ-plane with respect to the leading primary vertex is denoted as d0 in the table. The distance between the leading primary vertex and the muon track in the z-direction is denoted dz. These two variables can be used to establish how consistent the muon track is with the leading primary vertex. A *Global Muon* is a muon which has been identified from both hits in the muon chamber and hits in the tracker whereas a *Tracker Muon* has only been identified from tracker hits.

	Tight WP		Loose WP	
Requirements	8 TeV	13 TeV	8 TeV	13 TeV
Is a Global Muon and a Tracker Muon	yes	yes	yes	yes
p_{T} (GeV)	> 30	> 26	> 10	> 10
$\vert\eta\vert$	< 2.1	< 2.1	< 2.5	< 2.5
RelIso	< 0.12	< 0.15	< 0.2	< 0.25
Number of valid hits in the tracker	> 5	> 5	-	-
Number of hits in the muon stations	> 0	> 0	-	-
d0 (cm)	< 0.2	< 0.2	-	-
dz (cm)	< 0.5	< 0.5	-	-
Number of hits in the pixel tracker	> 0	> 0	-	-
Normalised χ^2 of track	< 10	< 10	-	-
Number of muon stations matched to track	> 1	> 1	-	-

Table 4.1: The cuts used for the tight and loose muon identification at 8 TeV [76] and 13 TeV [77].

4.6 Electrons

In this thesis the word electron is mostly used to include both charges, electron and positron. Electrons typically lose a large fraction of their energy in the tracker material via Bremsstrahlung. This is one of the biggest challenges in reconstructing electrons as the Bremsstrahlung radiation needs to be taken into account to accurately measure their momentum. These Bremsstrahlung photons can also convert into electron-positron pairs in the tracker material creating secondary electrons which must be distinguished from the signal electrons coming from the hard process.

Similar to the muon reconstruction, two working points are defined for electrons from the PF candidates. Tight electrons are used when requiring electrons as part of the signal process and a looser set of criteria are used in order to veto on extra electrons in the event ensuring a strict selection. The loose selection will contain more electrons but with a lower purity.

There are multiple ways to identify electrons, two of which are used in this thesis. At 8 TeV a multivariate technique to identify electrons was used. At 13 TeV a *cuts based identification* was used as the electron tools from CMS were not as advanced by the time the analysis was performed.

The tight and veto working points at $\sqrt{s} = 8$ TeV can be found in Table 4.2. The multivariate algorithm used for electron identification assigns a discriminator value which is closer to one for candidate particles which are more consistent with being a real electron and closer to zero if not. This discriminator is shown in Fig. 4.1. A conversion veto is applied for tight electrons which mitigates against identifying electrons which have come from photons converting into an electron-positron pair in the detector [78].

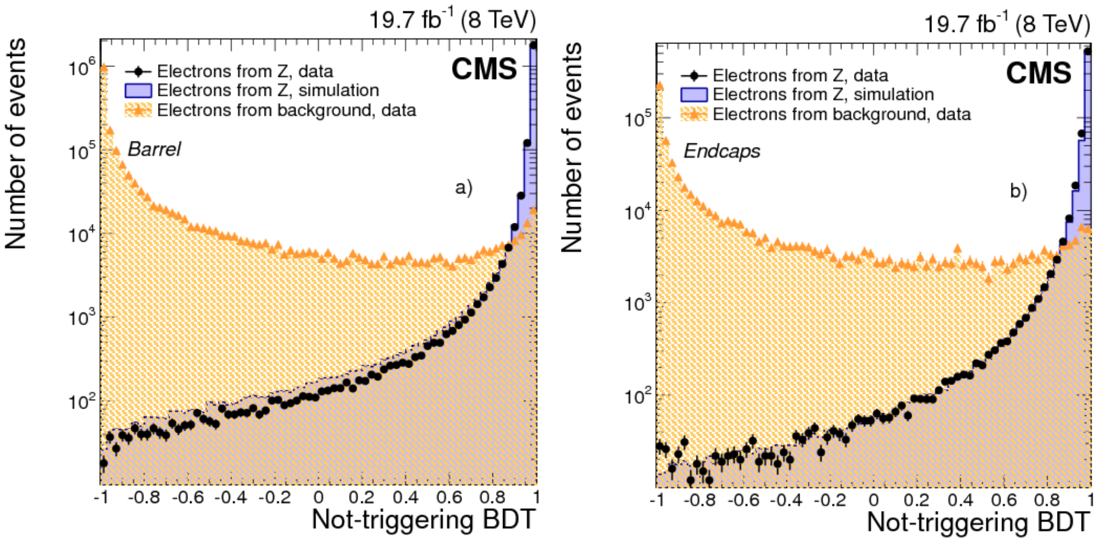


Figure 4.1: Output of the electron-identification boosted decision tree (BDT) multivariate algorithm for non-triggering electrons from Z $\rightarrow$ e$^+$e$^-$ data (black dots) and simulated (solid histograms) events, and from background-enriched events in data (triangles), in the ECAL a) barrel (left) and b) endcaps (right). For triggering electrons, to mimic the requirements applied at the HLT loose identification and isolation requirements are applied as a preselection [78].

The tight and veto working points are given for the barrel and endcap in Table 4.3 for the $\sqrt{s} = 13$ TeV analysis. The ECAL crystal that contains the largest energy

Requirements	Tight	Veto
E_{T} (GeV)	> 30	> 20
$\lvert\eta\rvert$	< 2.5	< 2.5
$\lvert$d0$\rvert$ (cm)	< 0.02	-
ConversionVeto	yes	-
MVA ID	> 0.9	> 0
RelIso	< 0.1	< 0.2

Table 4.2: The cuts used for the tight and veto electron identification at $\sqrt{s}$ = 8 TeV [79].

deposit is considered to be the *seed crystal.* Strips of ECAL crystals (in ϕ) are clustered together to create a "supercluster" (SC) with minimum requirements on initial seed crystal energy, as well as minimum energy of each strip. The distances in η and ϕ between the supercluster and the point in the ECAL where the track from the PV is extrapolated to are defined as $\Delta\eta_{In}$ and $\Delta\phi_{In}$, respectively. The ratio of the hadronic to electromagnetic energy around the seed cluster is denoted $\frac{h}{E}$. Photons which come from conversions often have tracks which do not pass through the innermost layer of the tracker, therefore limits are set on the number of expected missing hits, N_{miss}. The effect of the momentum loss through the tracker material can be quantified in the variable $\frac{1}{E_{SC}} - \frac{1}{p}$, where E_{SC} is the energy of the supercluster and p is the momentum of the track at closest approach to the vertex. The shower shape variable, $\sigma_{I_\eta I_\eta}$, describes the lateral extension of the shower along the η direction.

	Tight		Veto	
Requirements	Barrel	Endcap	Barrel	Endcap
$\sigma_{I_\eta I_\eta}$	< 0.0101	< 0.0279	< 0.0114	< 0.0352
$\lvert\Delta\eta_{In}\rvert$	< 0.00926	< 0.00724	< 0.0152	< 0.0113
$\lvert\Delta\phi_{In}\rvert$	< 0.0336	< 0.0918	< 0.216	< 0.237
$\frac{h}{E}$	< 0.0597	< 0.0615	< 0.181	< 0.116
RelIso	≤ 0.0354	≤ 0.0646	≤ 0.126	≤ 0.144
$\frac{1}{E_{SC}} - \frac{1}{p}$ (GeV)$^{-}$1	< 0.012	< 0.00999	< 0.207	< 0.174
$\lvert$d0$\rvert$ <	< 0.0111	< 0.0351	< 0.0564	< 0.222
$\lvert$dz$\rvert$	< 0.0466	< 0.417	< 0.472	< 0.921
N_{miss}	≤ 2	≤ 1	≤ 2	≤ 3
pass conversion veto	yes	yes	yes	yes

Table 4.3: The cuts used for the tight and veto electron identification at $\sqrt{s}$ = 13 TeV [80] where barrel is $\lvert\eta_{SC}\rvert < 1.4442$ and endcap is $(1.5660 < \lvert\eta_{SC}\rvert < 2.5)$.

Electron reconstruction cannot be performed accurately in the transition region (TR) between the ECAL barrel and endcap, $1.4442 < TR < 1.5660$ and hence they are excluded from physics analyses.

4.7 Jets

When partons such as quarks and gluons hadronise, they form a number of charged and neutral hadrons travelling in approximately the same direction of travel as the original parton. These final state particles can be clustered into what is known as a *jet* using the anti-κ_{T} reconstruction algorithm [81]. This is an infrared and collinear safe algorithm which starts with a high p_{T} 'seed' deposit in the calorimeter and uses the distance measure in Eq. 4.4 to find the nearest deposit to merge with. If the distance between the seed particle and the beam, d_{iB} in Eq. 4.5, is smaller than the distance to another hit, d_{ij}, the particle is merged with the beam [82]. Otherwise, the particle is merged with the nearest hit, according to d_{ij}. The algorithm terminates when $d_{iB} < d_{ij}$ and the merged particles are considered to be a jet. In this thesis, a distance parameter of $R = 0.5$ (0.4) is used to reconstruct jets in the Run 1 (Run 2) analysis in Chapter 7 (9). The distance parameter was changed for Run 2 to be consistent with the ATLAS experiment and to mitigate against PU contamination.

$$d_{ij} = min\left(p_{\mathrm{T}i}^{-2}, p_{\mathrm{T}j}^{-2}\right)\frac{\Delta R_{ij}^2}{R^2}\text{ , where }\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2 \tag{4.4}$$

$$d_{iB} = p_{\mathrm{T}i}^2 \tag{4.5}$$

Corrections are applied to the jet energy to account for the non-uniform response of the detector in p_{T} and η. The first correction is the *L1FastJet* correction which is applied to both data and simulation to remove the energy coming from PU events. The *L2Relative* and *L3Absolute* corrections respectively correct for the non-uniform response in η and p_{T} for both data and simulation. The *L2L3Residual* corrections are applied to simulation only and correct the remaining small differences in jet response between data and simulation such as correcting the absolute jet energy scale (JES). The jet energy resolution (JER) is also smeared by 10% as the resolution is worse in data than in simulation. Together these corrections are called the *jet energy corrections* (JEC) [83]. Jet identification criteria are applied to suppress fake jets arising from electrons showering in the ECAL due to Bremsstrahlung. This includes requiring $|\eta| < 2.5$, $p_{\mathrm{T}} > 30$ GeV and a separation from the nearest loose muon or electron of $\Delta R > 0.4$.

The biggest gains in using the PF algorithm come from performing the jet reconstruction on PF particles. The jet-matching efficiency, jet energy resolution and the reconstruction of the jet p_{T} are improved compared to using calorimeter information alone [84].

4.8 b-tagging

The presence of jets in an event indicates that the particles emerging from the collision include quarks and gluons. Being able to identify or *tag* which flavour of quark hadronised in the detector is extremely useful for a wide range of analyses. Particularly for searches for final states containing four top quarks, the ability to identify b-quarks originating from the decay of top quarks is incredibly beneficial in allowing us to discriminate between the signal and backgrounds. The particle shower coming from the hadronisation of b-quarks will contain B mesons and Λ_B (and other) baryons. These particles travel further in the detector due to having longer decay times than light flavour (u, d, s) mesons and baryons, resulting in a typical flight distance of up to a few centimetres [85]. The *impact parameter* (IP), defined as the distance between the primary vertex and the extrapolated point of closest approach of a track, will be larger for tracks coming from the decay of a hadron containing a b-quark. The tracks emerging from this decay can form a secondary vertex. This information is exploited in the Combined Secondary Vertex (CSV) algorithm [86]. The CSV algorithm is used to identify or *tag* jets which originate from b-quarks by assigning a discriminator value between 0 and 1, where larger values are more consistent with b-quark jets. Loose (CSVL), medium (CSVM) and tight (CSVT) working points are defined at values of the discriminator for a given mis-identification rate. For analyses at $\sqrt{s} = 13$ TeV the algorithm was improved and is called Combined Secondary Vertex version 2 algorithm (CSVv2) [87], the output of which is shown in Fig. 4.2. Working points, selection efficiencies for b-quarks and mis-identification rates are given in Table 4.4.

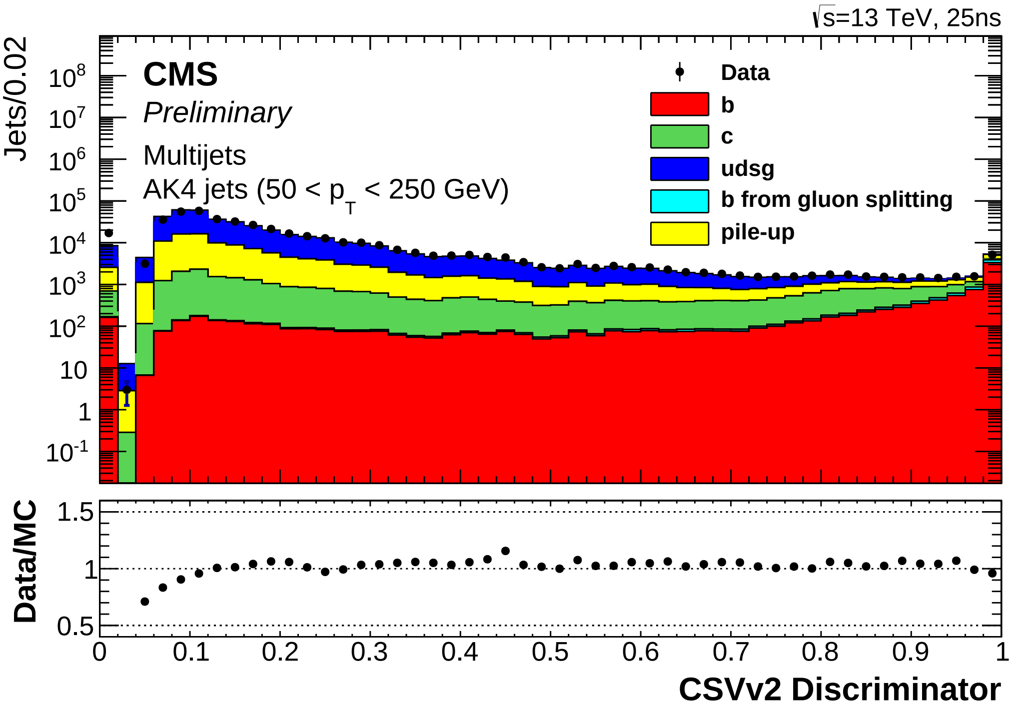


Figure 4.2: CSVv2 discriminator distribution at $\sqrt{s} = 13$ TeV using a multi-jets sample. Jets are reconstructed using the anti-k_t algorithm with $R = 0.4$ [87].

$\sqrt{s}$ (TeV)	Name	WorkingPoint	Selection Efficiency (%)	Mis-identification (%)
8	CSVL	0.244	< 80	0.1
	CSVM	0.679	< 62	0.01
	CSVT	0.898	< 35	0.001
13	CSVv2L	0.46	82	11.5
	CSVv2M	0.8	67	1.4
	CSVv2T	0.935	47	0.15

Table 4.4: b-tagging working points and their selection and mistagging efficiencies for PF jets.

4.9 Missing transverse energy

As it is not possible to detect neutrinos and potentially some BSM particles because they interact incredibly weakly with matter, their existence can be inferred by examining the sum of the momentum of particles in the transverse plane of the detector. The transverse plane is defined to be transverse to the beamline. Starting with the assumption that the total momentum in the transverse plane is zero, an imbalance in the sum of the momentum of detectable particles is considered to be missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$), as defined in Eq. 4.6. The p_{T} of jets is used after the JEC have been applied [88].

$$E_{\mathrm{T}}^{\mathrm{miss}} = \; - \sum_{\text{all particles},i} p_{\hat{\mathrm{T}},i} \tag{4.6}$$

5 Simulation

The search for new physics at the LHC requires comprehensive modelling of the overwhelmingly large SM backgrounds. Particle physics events are simulated using Monte Carlo (MC) simulation that can be used, for example, to measure properties of known particles by comparison with data or to predict and optimize searches for BSM physics. Hence, simulation is generated for SM backgrounds using current theoretical knowledge, and BSM models are simulated using theoretical models. There are four main stages; generation (GEN), simulation (SIM), digitisation (DIGI) and reconstruction (RECO). The GEN stage consists of producing the hard scattering between the partons from the protons and the outgoing particles. The SIM stage continues from the GEN stage simulating the paths of the outgoing particles through the detector after which the response of the detector is generated in the DIGI step. The RECO stage then uses the algorithms previously discussed in this chapter to produce collections of high-level physics objects which are the same as what is reconstructed in the detector from real data events.

Event generators are used to simulate the signal and background processes at GEN-level. Using the proton PDFs and calculations of the Matrix Element (ME) associated to the Feynman diagrams for a particular process, a proton-proton collision can be replicated in simulation so that theory can be compared to experimental results. Due to the high-dimensional phase space associated to the multi-particle events produced at the LHC, Monte Carlo simulation is the only viable option. The GEANT4 [89] program is used to perform the SIM-level simulation through the CMS detector.

Figure 5.1, shows the chain of processes which occur at GEN-level including:

- Incoming protons
- Hard interaction of partons from inside proton
- Parton shower from outgoing partons

- Hadronisation of partons
- Underlying event from the proton remnants

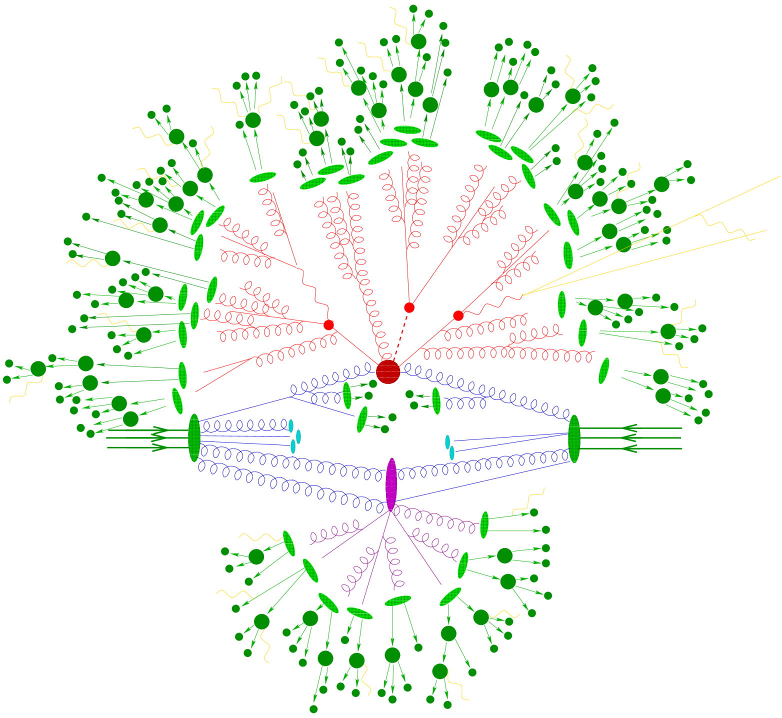


Figure 5.1: Depiction of hadron-hadron collision as simulated using an MC event generator. The red circle in the center indicates the hard collision. The surrounding tree-like structure represents Bremsstrahlung as simulated by parton showers. The purple circles represent a secondary hard scattering event. Light green ellipses represent parton to hadron transitions through hadronisation, darker green ellipses indicate hadron decays, and the yellow lines are soft photon radiation [90].

There are many MC event generators which specialise is one or more parts of the GEN-level simulation. Some generators are more specialised towards ME calculations and producing the hard interaction. Others are more optimised for simulating the parton shower, which is treated as a Markov process where four-momentum and probabilities are conserved with the creation of each new particle [90]. Using generators above LO is a necessity when precision measurements are required and when many high-p_{T} and well-separated jets are present in the signature for the signal process [91]. The main event generators used in the thesis are as follows.

MadGraph

MadGraph is a leading-order (LO) event generator [92] which calculates the ME at tree-level with a number of additional partons (which is process dependent and limited by computer memory contraints) for processes such as decays and $2 \rightarrow n$ scattering. It takes PDF sets as input, for example NNPDF3.0 as shown in Fig. 2.4, which describes the kinematics of the incoming partons from the proton. It generates all Feynman diagrams for a particular process and evaluates each ME for a given phase space point. The number and type of partons and the kinematics of the event are generated.

aMC@NLO

The aMC@NLO package [91] can simulate events at next-to-leading order (NLO) in perturbative QCD as it uses both tree-level and one-loop perturbations. These additional corrections from higher-order Feynman diagrams make the simulation more accurate than its LO counterparts. This package includes initial and final state radiation. Initial state radiation (ISR) refers to any particle which is radiated off of an incoming particle to the collision whereas final state radiation (FSR) refers to a particle radiated off the final state outgoing products of a collision.

Negative event weights By including higher order perturbations to the cross section calculated in aMC@NLO, it is necessary to consider terms which interfere destructively. This is achieved by assigning negative weights to some events within the generator so that the differential cross section is simulated correctly. The effective number of events, N_{eff}, produced by the generator corresponds to $N^{pos} - N^{neg}$ (this equates to the number of events that would be produced for the given cross section) whereas the total number of events produced correspond to $N^{pos} + N^{neg}$. Therefore these samples can be scaled using the negative event weight, W^{neg}_{event}, according to Eq. (5.1).

$$W^{neg}_{event} = \frac{N^{pos} + N^{neg}}{N^{pos} - N^{neg}} = \frac{N^{Total}}{N^{Total} - 2 \times N^{neg}} \tag{5.1}$$

POWHEG

The 'Positive Weight Hardest Emission Generator', known as POWHEG [93] generates the hardest process in the event first. This means that when it is interfaced with a shower MC (SMC) program for hadronisation which has emissions p_T-ordered, the double counting of the low-p_T radiation emitted can be avoided by using a p_T-veto. As the double counting of low-p_T emission is the cause for negative weight events, these can be avoided using POWHEG with an SMC with p_T-ordering but negative weight events are still necessary for angular-ordered SMCs [94].

PYTHIA

The PYTHIA program [95] can take the parton-level event generated by another generator and add soft emissions from the initial and final state particles, as well as performing the fragmentation and hadronisation of quarks to produce the parton shower (PS). It can also simulate everything standalone including the initial protons fragmentation, multi-particle production, hadronisation, beam remnants and the underlying event. PYTHIA is one of the most widely used generators amongst LHC experimentalists as it has been observed to produce particularly good agreement with data compared to other generators which produce parton showers.

Matching

“Matching” or “merging” refers to the method of combining the output of the hard scatter, which has well separated particles, with parton showers which have much softer low-p_T particles. The p_T threshold at which partons from the ME are matched to the PS is known as the ME-PS threshold. There are two types of matching used in MadGraph depending on the order at which the process is generated [91]. At LO, MLM-merging is used to combine multiple LO + PS samples which are produced with different final-state multiplicities. FxFx-merging is similarly defined however NLO matrix elements are used.

6 Analysis strategy and techniques

The LHC has been described as a “top quark factory” due to the large cross sections for $\mathrm{t\bar{t}}$ production at $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV, which were calculated at NNLO to be 253 pb and 831 pb [96, 97], respectively. Most analyses within the CMS collaboration which work on top quark physics study $\mathrm{t\bar{t}}$ production. There are also a number of analyses working on the rarer process of single top quark production. Four-top-quark production is rarer still at $\mathrm{O}(10^5)$ smaller than $\mathrm{t\bar{t}}$ production. In this chapter some of the common ideas and algorithms shared between both CMS $\mathrm{t\bar{t}t\bar{t}}$ analyses in this thesis are introduced. More analysis-specific details are given in Chapters 7 and 9 for the studies at $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV, respectively. In Section 6.1 the general strategy for searching for $\mathrm{t\bar{t}t\bar{t}}$ production is outlined. In Section 6.2, the signal $\mathrm{t\bar{t}t\bar{t}}$ process and background processes are each discussed. The various corrections which are made to the simulation to improve the modelling are given in Section 6.3. Analysis of the potential multi-jet background is given in Section 7.7. Next the multivariate techniques employed in this thesis are introduced in Section 6.5, specifically Boosted Decision Tree (BDT) algorithms (Section 6.7). BDTs are used both to reconstruct top quarks (Section 6.6) and then to separate the signal $\mathrm{t\bar{t}t\bar{t}}$ process from the main background of $\mathrm{t\bar{t}}$ production. The systematic uncertainties affecting the analyses in this thesis are described in Section 6.8. Finally, in Section 6.9 the statistical procedure used to set a limit on the cross section of $\mathrm{t\bar{t}t\bar{t}}$ production is discussed.

6.1 Strategy for searching for four top quarks

The strategy for selecting $\mathrm{t\bar{t}t\bar{t}}$ events while suppressing background processes is similar to the $\mathrm{t\bar{t}}$ selection but with the requirement of an additional two jets. The small cross section for $\mathrm{t\bar{t}t\bar{t}}$ production in the SM, 1.3 fb at $\sqrt{s} = 8$ TeV and 9.2 fb at $\sqrt{s} = 13$ TeV [98, 99] dictates the selection. For the single lepton channel, which is the focus of the studies in Chapters 7 and 9, the common selection between both analyses is as follows:

- Exactly 1 lepton (muon or electron)
- No additional leptons
- $\geq$6 jets
- $\geq$2 b-tagged jets

It would be preferential to require four b-tagged jets in the selection to obtain the highest signal-to-background ratio. However, this would have a detrimental effect on the number of $\mathrm{t\bar{t}t\bar{t}}$ events selected due to the fact that some b-jets will not be identified correctly or will not fall within the acceptance of the detector. Similarly, it is not advantageous to require a jet per quark hadronising in the detector in a $\mathrm{t\bar{t}t\bar{t}}$ final state as some of the jets may be merged or may not fall within the acceptance of the detector. However, it will be discussed in Section 6.9.2 how a looser selection can be used to an advantage to constrain the main background process. This thesis will focus on the single lepton channel where only single muon and single electron final states are considered. The criteria that each channel are required to pass are called the *Baseline Event Selection*.

6.2 Signal and background processes

6.2.1 Four-top-quark production signal process

It can be seen from Fig. 6.1 that the single lepton final state has the largest branching ratio which makes it a favourable place to study. This is the final state considered most in this thesis. The dilepton final state also has a large branching ratio and has particularly low backgrounds when same-sign lepton final states are

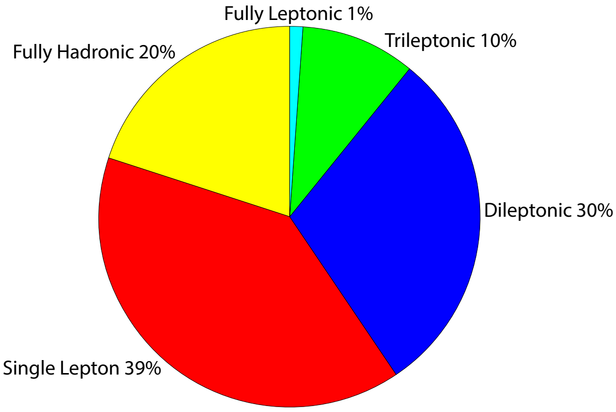


Figure 6.1: The possible decay channels for $t\bar{t}t\bar{t}$ production.

considered. The combination of studies on the dilepton final state will be discussed in Chapter 9.

Figure 6.2 shows the cross section for $t\bar{t}t\bar{t}$ production versus centre of mass energy in the proton-proton collision. It can be seen that the cross section rises with energy and indeed it increases faster than the main background of $t\bar{t}$ production, which means that the signal to background ratio increases with $\sqrt{s}$. Table 6.1 shows the cross sections for $t\bar{t}t\bar{t}$ production and $t\bar{t}$ production where it can be seen that the $t\bar{t}t\bar{t}$ cross section increases by a factor of $\approx$ 7 from $\sqrt{s} = 8$ TeV to $\sqrt{s} = 13$ TeV whereas the $t\bar{t}$ production increases by a factor of $\approx$ 3.5. This is due the fact that the parton luminosity for gluon-gluon fusion (the dominant production process for $t\bar{t}$ and $t\bar{t}t\bar{t}$) increases more rapidly for higher invariant mass processes between $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV.

Cross section	8 TeV (pb)	13 TeV (pb)
$t\bar{t}t\bar{t}$	0.0013	0.0092
$t\bar{t}$	245	831

Table 6.1: The cross sections for $t\bar{t}t\bar{t}$ (computed at NLO) and $t\bar{t}$ (computed at NNLO) production at 8 TeV and 13 TeV [96–99].

6.2.2 $t\bar{t}$ background

The main background to $t\bar{t}t\bar{t}$ production is $t\bar{t}$ production with additional jets. The final state of the $t\bar{t}$ process in the detector, shown in Fig. 6.3, is defined by whether each W boson decays leptonically or hadronically.

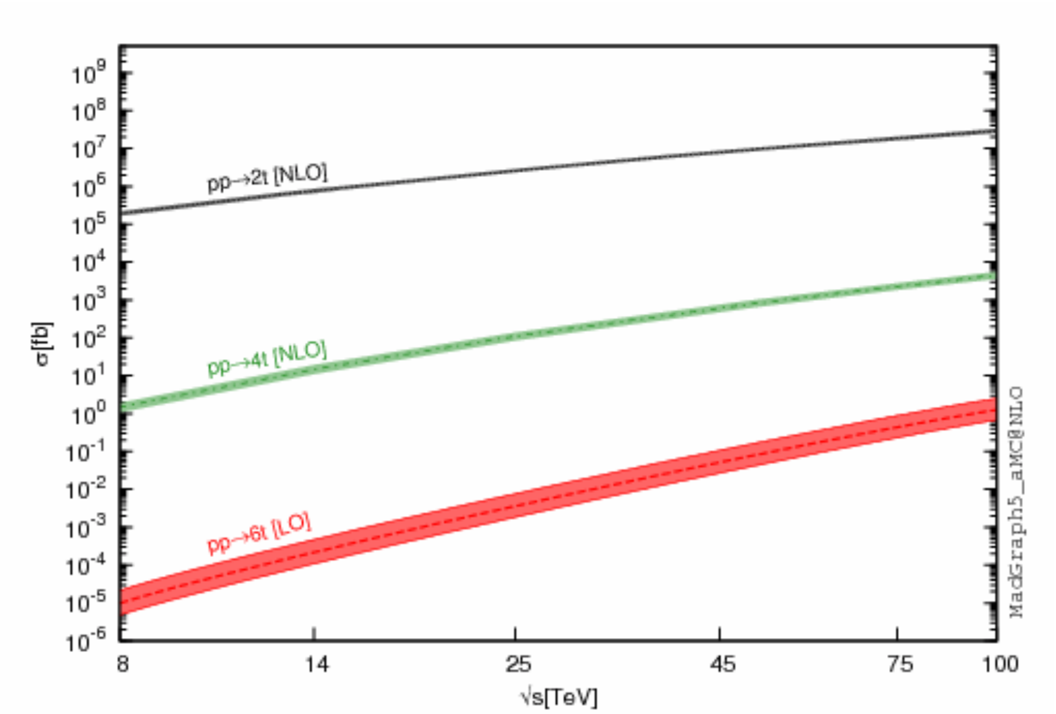


Figure 6.2: The $t\bar{t}$ (2t), four top quark (4t) and six top quark (6t) production cross sections at a range of centre of mass energies, $\sqrt{s}$ [100].

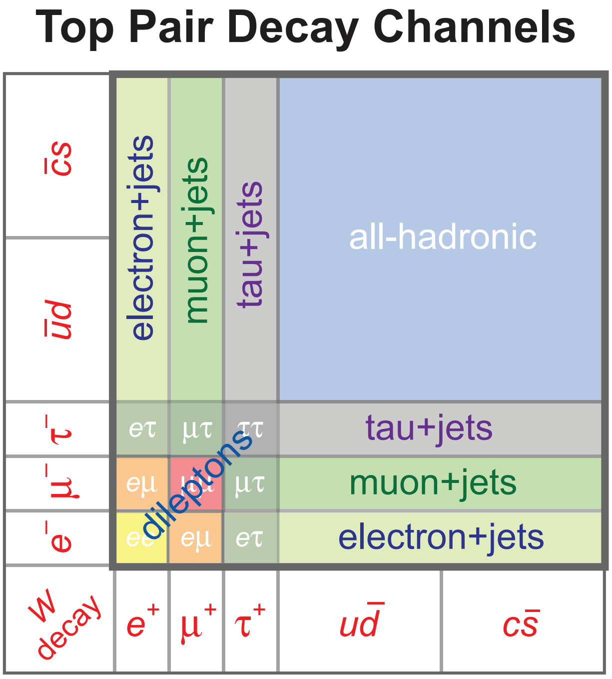


Figure 6.3: The possible decay channels for $t\bar{t}$ production where the area of each final state is proportional to its branching ratio [27].

The standard strategy is to require two b-jets to be present in the event and 0, 1 or 2 leptons depending on whether the final state is all-hadronic, semi-leptonic or dileptonic respectively, where 6, 4 or 2 total number of jets are required. Not all $\mathrm{t\bar{t}}$ events will be captured within the selection due to inefficiencies in tagging b-jets, identifying leptons and reconstructing jets. As the rate of $\mathrm{t\bar{t}}$ production is very high at the LHC, this selection still provides a large enough sample of events to perform studies of $\mathrm{t\bar{t}}$ production that are not limited by the size of the signal sample.

In the case of the single lepton channel, semi-leptonic $\mathrm{t\bar{t}}$ is the main final state which contributes to the background. Figure 6.4 shows semi-leptonic $\mathrm{t\bar{t}}$ production with ISR and FSR.

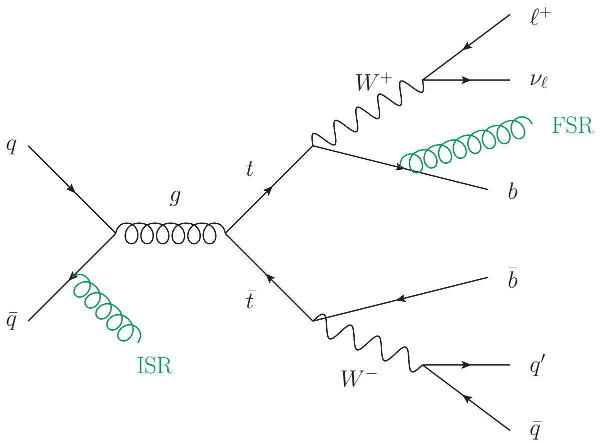


Figure 6.4: Semi-leptonic $\mathrm{t\bar{t}}$ production with demonstration of extra jets arising from initial and final state radiation.

The semi-leptonic $\mathrm{t\bar{t}}$ process can produce final states which look very much like a $\mathrm{t\bar{t}t\bar{t}}$ final state in the detector. There are two top quarks which will produce hard jets including two b-quark jets and one lepton. There are also the additional jets from ISR and FSR which increase the jet multiplicity in the event. Final state jets may arise that do not exceed the p_{T} threshold of the trigger, which is described in Section 3.2.6. Jets may also fall outside of the acceptance of the detector or jets may merge together, both of which can lower the $\mathrm{t\bar{t}t\bar{t}}$ multiplity to look more like $\mathrm{t\bar{t}}$. The spectrum of ‘additional jets’ which are not associated with coming from the decay of a top quark are shown in Fig. 6.5 for a CMS analysis at $\sqrt{s} = 13$ TeV. Events which lie in the tail of the additional jets distribution are more likely to fall within the requirements for a $t\bar{t}t\bar{t}$ event. However, in general, jets which come from ISR or FSR tend to be *softer*, ie. lower in p_T. The mis-modelling of the jet multiplicity spectrum, as observed in the lower panel of Fig. 6.5, is understood to be due to a suboptimal value of α_S being used to produce the $t\bar{t}$ simulation. The treatment of this mis-modelling is discussed further in Section 6.3.6.

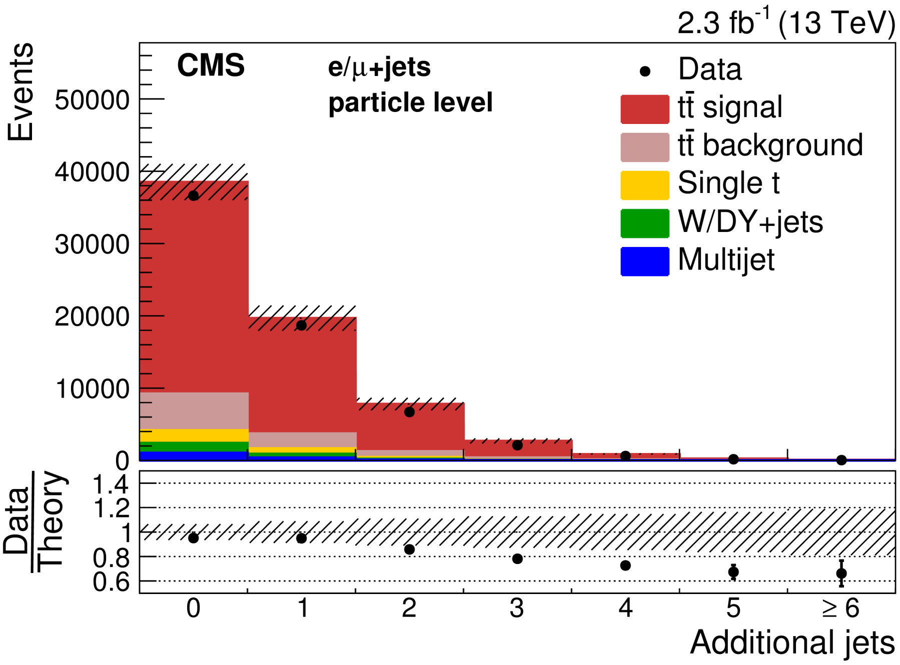


Figure 6.5: Number of jets which are not associated with a top quark decay in the $t\bar{t}$ semi-leptonic (lepton + jets) channel [101].

6.2.3 Electroweak backgrounds

The electroweak backgrounds include W and Z boson production in association with jets, which in their leptonic decays provide one and two leptons, respectively, and no jets. Additional jets come from ISR or FSR. A small fraction of electroweak events will have a large number of additional jets. However due to their large cross section, as can be seen in Fig. 6.6, there will be a small number of events which may pass the requirements for selecting $t\bar{t}t\bar{t}$ events.

The W and Z boson may also decay hadronically where each will provide two jets. Along with the additional jets from ISR/FSR, one jet may be misidentified as a

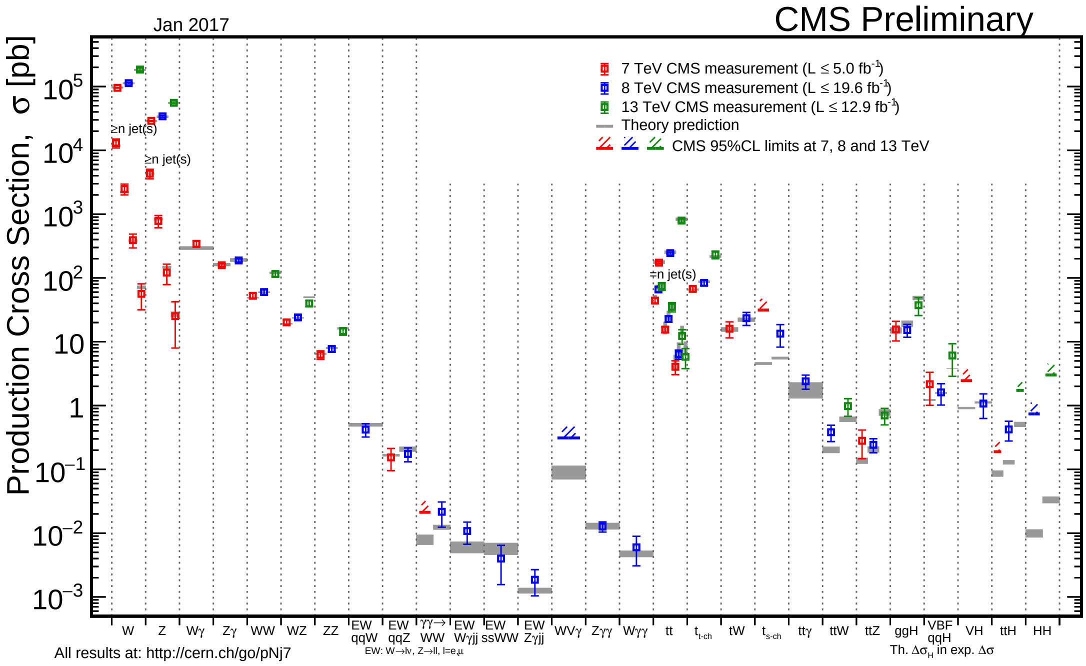


Figure 6.6: CMS standard model production cross sections [102].

lepton in the detector and hence may also pass the selection requirements for $t\bar{t}t\bar{t}$ events.

6.2.4 Rarer backgrounds

Single top production both has a smaller cross section than W and Z boson production and is less similar to $t\bar{t}t\bar{t}$ production than $t\bar{t}$ production by virtue of having one less top. However the cross section is still many times larger than $t\bar{t}t\bar{t}$ production. Particularly in the tW-channel there can be enough additional jets (including two b-jets) such that it may mimic $t\bar{t}t\bar{t}$ in the detector.

Other rarer backgrounds include $t\bar{t}H$, $t\bar{t}Z$ and $t\bar{t}W$ which resemble $t\bar{t}t\bar{t}$ production much more closely than $t\bar{t}$ by having extra bosons produced in the process which can decay to higher p_T jets and leptons than can occur from ISR/FSR. However, they are much more rare than $t\bar{t}$ production and do not produce a significant background. Diboson production (WW, WZ, ZZ) can also possibly mimic $t\bar{t}t\bar{t}$ production. However, they can be considered to be small perturbations on W + 2 jets and Z + 2 jets processes as they are at least an order of magnitude smaller in cross section. Simulated diboson events were processed through the analysis framework where a negligble (sometimes zero) number of events passed the baseline event selection. Hence, it is too rare to be considered in the analyses in this thesis.

6.3 Corrections to the simulation

Many of the parameters which go into simulating each particle physics process are not precisely known and some can not even be calculated. Therefore these parameters are tuned to produce the simulation that best matches the observed data. This simulation will still have residual discrepancies from data that can be measured and accounted for by producing *scale factors* (SF). These SFs are usually dependent on the p_T and η of the particles produced and may be dependent on other factors such as jet flavour. They are used to produce a per-event weight for simulated events such that the overall distributions more closely match data. SFs are measured using control samples. For example lepton SFs are measured using the "tag-and-probe" method using the Z→ee or Z→ $\mu\mu$, as described in Ref. [103].

Distributions of some of the SFs used in the $\sqrt{s} = 13$ TeV analysis can be found in Appendix C.1.

6.3.1 Pileup modelling

The PU SF (SF_{PU}) is derived using *minimum bias* data and simulation. This means that the data have been collected using a much looser trigger than those used to detect interesting physics events.

The distribution of the number of p-p interactions (n_{int}) in collisions varies between data and simulation for a given luminosity. The distribution of n_{int} in data can be estimated from the measured luminosity per bunch crossing (L_{bx}) and the total inelastic cross section (σ_{in}) using:

$$n_{int} = L_{bx} \times \sigma_{in}$$

The SF_{PU} can be derived by comparing the number of data events (N_{int}^{Data}) to the number of simulated events (N_{int}^{sim}) for a given n_{int} as in Eq. 6.1.

$$SF_{PU}\left(n_{int}\right) = \frac{N_{int}^{Data}\left(n_{int}\right)}{N_{int}^{sim}\left(n_{int}\right)} \tag{6.1}$$

6.3.2 b-tag modelling

There are residual differences between the b-tagging efficiencies measured by CMS in data and the efficiencies as measured in simulation as it is difficult to simulate the fragmentation and hadronisation of b-quarks. There are two methods used for correcting the b-tag modelling in this thesis. In Method 1, scale factors are derived to correct the distributions of the number of b-tags for each working point. In Method 2, it is the CSV(v2) distribution of all jets that is corrected by applying a weight to each jet.

6.3.2.1 Method 1

The b-tagging efficiencies are measured for each of the CSVL, CSVM, and CSVT working points in bins of p_{T} and η. Depending on which working point was used in an analysis, a scale factor, $SF(\eta, P_T)_i$, can be applied to the simulation for each jet. This is dependent on the p_{T}, η and flavour, i, of the jet, shown in Eq. 6.2, where $\epsilon(\eta, p_{\mathrm{T}})_i^{data}$ is the efficiency for a jet to be tagged as a b-jet in data and $\epsilon(\eta, p_{\mathrm{T}})_i^{sim}$ is the efficiency for identifying a jet as a b-jet in simulation.

$$SF(\eta, P_T)_i = \frac{\epsilon(\eta, p_{\mathrm{T}})_i^{data}}{\epsilon(\eta, p_{\mathrm{T}})_i^{sim}} \tag{6.2}$$

Separate scale factors are defined for b and light (u, d, s, g) jets. Scale factors for c jets are taken to be the same as for b jets. A weight, ω_{btag} can be applied to simulated events. The method proceeds by defining the probability of an event in simulation producing a given number of tagged and untagged jets, *P(MC)*, defined as

$$P(MC) = \prod_{tagged\ jets} \epsilon_i^{sim} \times \prod_{untagged\ jets} (1 - \epsilon_i^{sim})$$

where ϵ_i^{sim} is the efficiency of tagging a jet of flavour i with the CSV criterion in simulation. While the probability of an event in data producing a given number of tagged and untagged jets, P(DATA), is defined as

$$P(DATA) = \prod_{tagged\ jets} SF_i \cdot \epsilon_i^{sim} \times \prod_{untagged\ jets} (1 - SF_i \cdot \epsilon_i^{sim})$$

where SF_i is the appropriate scale factor for a jet of flavour i.

An overall event weight to be applied depending on the jet content can be derived from these scale factors using Eqn. 6.3. This weight, ω_{btag}, must be applied to the events in simulation in order to predict the correct event yield.

$$\omega_{btag} = \frac{P(DATA)}{P(MC)} \tag{6.3}$$

6.3.2.2 Method 2

Alternatively, the measurements at each working point can be used to fit the shape of the CSV distribution and provide scale factors in bins of p_{T} and η for each jet flavour. Therefore the scale factors for each jet can be derived from the p_{T}, η, and CSV discriminator value and for each jet flavour as seen in Eq. 6.4. For this method, the jet flavours are defined as heavy for bottom quarks and light for u, s, d, g whilst c-quarks are given $SF = 1$. The scale factor for a heavy (light) flavour jet is defined as SF^H (SF^L), the number of jets in data are defined as N_{jets}^{Data}, whilst the number of heavy (light) flavour jets in simulation are defined as N_{sim}^H (N_{sim}^L). SF^H and SF^L are derived in bins of p_{T}, η and CSV value.

$$\begin{aligned} SF^H(\text{CSV}, p_{\mathrm{T}}, \eta) &= \frac{N_{\text{jets}}^{Data} - N_{sim}^L}{N_{sim}^H} \\ SF^L(\text{CSV}, p_{\mathrm{T}}, \eta) &= \frac{N_{\text{jets}}^{Data} - N_{sim}^H}{N_{sim}^L} \end{aligned} \tag{6.4}$$

An event weight can be derived by taking the product of the per-jet scale factors for all heavy and light flavour jets in the event, as seen in Eq. 6.5.

$$SF_{\text{total}} = \prod_{i}^{N_L} SF_i^L \prod_{j}^{N^H} SF_j^H \tag{6.5}$$

Further details of the CSV distribution reshaping can be found in Ref. [104].

6.3.3 Heavy flavour jet modelling

The extra jets in $\mathrm{t\bar{t}}$ events can come from processes such as gluon splitting which pair produces $\mathrm{b\bar{b}}$. These $\mathrm{t\bar{t}b\bar{b}}$ events are most likely to resemble the features of the $\mathrm{t\bar{t}t\bar{t}}$ signal events and so it is essential that the proportion of them in simulation is correctly modelled. Table 6.2 shows the heavy flavour ratio $R = \sigma_{\mathrm{t\bar{t}b\bar{b}}} \ / \ \sigma_{\mathrm{t\bar{t}jj}}$ as measured by CMS at $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV in data and in simulation. To incorporate this ratio into the analysis, the MC truth information of the $\mathrm{t\bar{t}}$ + jets (MC) sample is used to split the sample into $\mathrm{t\bar{t}b\bar{b}}$, $\mathrm{t\bar{t}c\bar{c}}$ and $\mathrm{t\bar{t}ll}$, where l denotes light quarks and gluons (u, d, s, g). A scale factor SF = R(Data) / R(Sim) is derived from the information in Table 6.2 which is applied to the $\mathrm{t\bar{t}b\bar{b}}$ events. Another SF is applied to $\mathrm{t\bar{t}ll}$ to preserve the total number of $\mathrm{t\bar{t}}$ events.

$\sqrt{s}$ (TeV)	R(Data) (%)	R(Sim) (%)	SF
8	$2.2 \pm 0.4\,(\text{stat.}) \pm 0.5\,(\text{sys.})$	1.6 ± 0.2	1.35
13	$2.2 \pm 0.3\,(\text{stat.}) \pm 0.6\,(\text{sys.})$	1.2 ± 0.1	1.83

Table 6.2: Ratio of $R = \sigma_{\mathrm{t\bar{t}b\bar{b}}} \ / \ \sigma_{\mathrm{t\bar{t}jj}}$ for Data and Simulation at $\sqrt{s} = 8$ TeV [105] and $\sqrt{s} = 13$ TeV [106] alongside the scale factor derived for R(Data) / R(Sim).

6.3.4 Lepton modelling

A weight is applied to events which is dependent on the η, p_{T} and flavour of the selected leptons. The scale factors for each source of efficiency (isolation, identification, reconstruction and trigger) are designed to be multiplicative. The final event weight is:

$$\omega_{lepton} = SF_{iso} \times SF_{id} \times SF_{reco} \times SF_{trig} \tag{6.6}$$

6.3.5 Top p_T modelling

The top quarks which are reconstructed in $\mathrm{t\bar{t}}$ enriched regions of data tend to have a softer p_T spectrum than in $\mathrm{t\bar{t}}$ simulation, which is believed to be due to the NLO approximation used in the event generator. A scale factor can be derived as a function of the p_T of the top at generator level of the simulation (before it is reconstructed by the detector). The event weight can be derived by multiplying the independent SFs for each (anti-)top in the event and is shown in Eq. 6.7. This SF is only applied to $\mathrm{t\bar{t}}$ simulation samples.

$$\omega_{topPt} = SF(\text{top } p_\mathrm{T}) \times SF(\text{anti-top } p_\mathrm{T}) \tag{6.7}$$

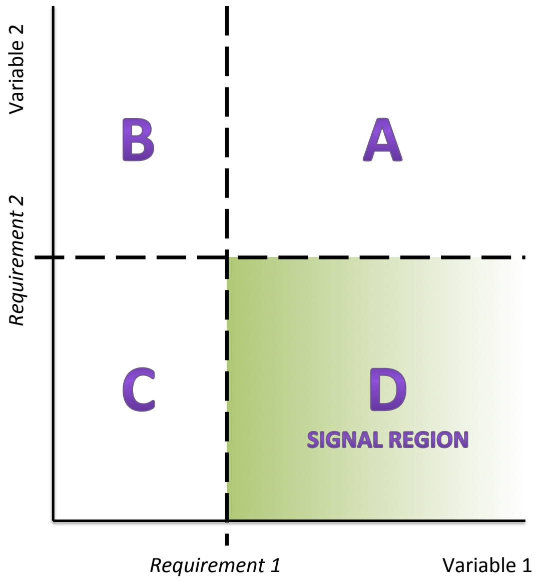
6.3.6 Jet multiplicity modelling

Good modelling of the jet multiplicity distribution up to a large number of jets in the main $\mathrm{t\bar{t}}$ background is very important for this analysis due to the fact that as the number of jets increases, the signal to background ratio increases. Hence, the higher jet multiplicities have the highest sensitivity in separating the signal $\mathrm{t\bar{t}t\bar{t}}$ process and the background $\mathrm{t\bar{t}}$ process. It is essential that $\mathrm{t\bar{t}}$ is well modelled in this high jet multiplicity region. Particularly at $\sqrt{s} = 13$ TeV the simulation was larger than data in the higher $\mathrm{N_{jets}}$ bins. The value of α_S in the $\mathrm{t\bar{t}}$ sample produced with POWHEG +PYTHIA 8 used in the analysis in Chapter 9 is 0.137, however the best tune was observed to have a value of $\alpha_S = 0.113^{+0.012}_{-0.010}$. A scale factor was therefore calculated by CMS in order to improve the modelling of the $\mathrm{N_{jets}}$ distribution [107].

6.4 Multi-jet background estimation

The presence of multi-jet events within the signal region defined by the baseline selection is investigated in this section. It is rare for multi-jet events to have a highly energetic undetectable particle as QCD processes do not produce neutrinos in the hard process. Therefore, the $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions for multi-jet events typically peak at low values. The peak can be above zero as some jets may be outside of the acceptance of the detector or mismeasured. It is difficult to simulate QCD events for the analyses in this thesis as QCD production has an incredibly large cross section with a small efficiency to produce events which pass the baseline selection for the $\mathrm{t\bar{t}t\bar{t}}$ analyses. Events which pass the selection will exist in the extreme tails of the $\mathrm{N_{jets}}$ and $\mathrm{N_{btags}}$ distributions of QCD events where there are more subtle effects in the modelling and more statistical fluctuations. Hence, it is not possible to use multi-jet simulation. In this case, a data-driven method known as the "ABCD method" may be used. This method proceeds by selecting two uncorrelated variables from the object or baseline selection and defining three control regions (A,B,C) and one signal region, D, in the 2-dimensional phase space of these variables, as shown in Fig. 6.7. A selection is made in each variable which defines one quadrant of the phase space as the signal region. The event variable $E_{\mathrm{T}}^{\mathrm{miss}}$ and the lepton variable RelIso were selected as they are uncorrelated, as shown in Fig 3 in Appendix .1. The signal region is defined in a low RelIso and higher $E_{\mathrm{T}}^{\mathrm{miss}}$ region.

For the background processes that are well modelled using simulation, their yields in each region are subtracted from the data. This should in theory leave only QCD multijet events remaining and so the number of events in the signal region can be estimated from the equation in Fig. 6.7. This method does have some dependence on the simulation and assumes that the other backgrounds are well modelled. The uncertainty on this assumption can be taken into account by repeating the procedure with different versions of the $\mathrm{t\bar{t}}$ simulation in which the ME and PS scale uncertainties (the main uncertainties on $\mathrm{t\bar{t}}$) are varied up and down.



$$\frac{N^{B}_{\text{multi-jet}}}{N^{A}_{\text{multi-jet}}} = \frac{N^{C}_{\text{multi-jet}}}{N^{D}_{\text{multi-jet}}}$$

Figure 6.7: Illustration of the ABCD method.

6.5 Multivariate analysis techniques

Four top quark production events are very rare at the energies of $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV studied in this thesis. Typically, within the selection defined in Section 6.1 the main background is $t\bar{t}$ production which is several orders of magnitude larger than the $t\bar{t}t\bar{t}$ signal. In this type of analysis, where the background and signal can be similar in many distributions and the signal is so rare in comparison to the backgrond, it can be advantageous to use multivariate analysis (MVA) methods to increase the signal to background separation. The main MVA method used in this thesis is *Boosted Decision Trees* (BDT). More details of this algorithm and its specific use in the analysis are given in the subsequent sections.

6.5.1 Boosted Decision Trees

To start with, the simpler concept of a single Decision Tree will be considered. Decision trees are used to maximise the separation between a signal (S) sample and a background (B) sample by looking at a set of distributions in which there is some initial separation between the two; a training sample for each is provided to the algorithm. A simple decision tree is shown in Fig. 6.8 where each decision is defined at a *node* and splits the dataset in two by placing a requirement in the most discriminating variable in order to separate as many background events down one side and signal events down the other side. One can define the purity at each node as $P = \frac{S}{S+B}$. Another useful measure is the *gain* at each node which gives a more symmetric measure of the node having high purity of either signal or background. The particular metric of gain used in this analysis is the commonly used *Gini Index* which is defined as $Gini = P \cdot (1 - P)$. A high Gini Index suggests the node contains a relatively equal amount of signal and background whereas a low Gini index shows that the node contains more of either signal or background. The goal is to scan across each distribution and find the requirement in the associated variable which maximises the *separation gain* (S_G) as defined in Eq. 6.8. This process is iterated upon until a predefined end condition such as the maximum depth of the tree or the minimum number of events in a node.

$$S_G = Gini(\text{parent}) - Gini(\text{child 1}) - Gini(\text{child 2}) \tag{6.8}$$

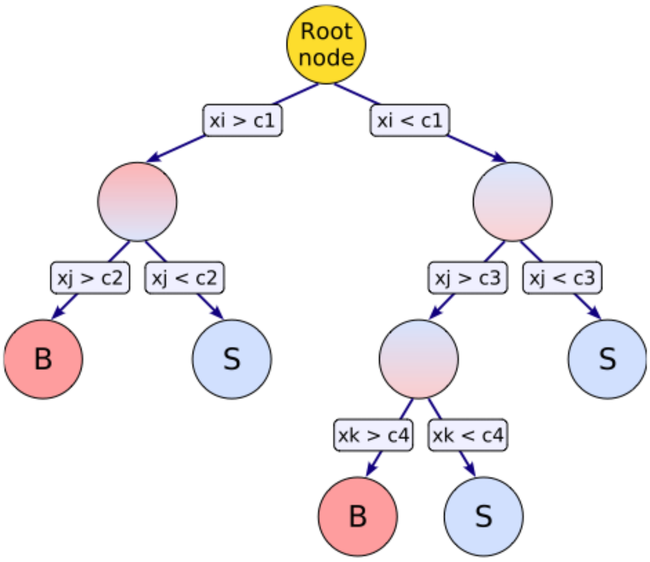


Figure 6.8: Illustration of a single decision tree of depth = 3 [108].

It has been found that using an ensemble (usually several hundred) of smaller decision trees of depth = 2 or depth = 3 is preferable to using one large decision tree as it can create a much stronger "learner" and minimise *overtraining*, ie. the propensity to train to the specific features of the training sample rather than the general trend.

The first tree begins as a normal decision tree as described above and it terminates when it reaches the maximum depth defined. The error rate, R_{err}, of the tree is defined as the number of events incorrectly classified divided by the total number of events. The events which were incorrectly categorised are weighted according to Eq. 6.9 where α is the *boost weight* which is dependent on R_{err}. The idea is that in the next iteration of a tree the incorrectly classified events are considered with higher importance. The degree of boosting can be adjusted by raising α to the power of β, the learning rate.

$$\alpha = \left(\frac{1 - R_{err}}{R_{err}} \right) \tag{6.9}$$

This process is repeated until the predefined maximum number of trees has been reached. This is known as the "Adaptive Boost" or "AdaBoost" algorithm A discriminator value, y_{BDT}, is defined by summing the response from each tree, t, where the boost weight for each tree is α_t. Each response is weighted by $\ln(\alpha_t)$ such that the trees with low error rates are considered more than trees with high error rates. The response from each individual tree is defined as $h(\mathbf{x})$ where $\mathbf{x}$ is the set of input variables. Events categorised in signal nodes are assigned $h(\mathbf{x}) = +1$ and events categorised in background nodes are assigned $h(\mathbf{x}) = -1$.

$$y_{BDT} = \frac{1}{Ntrees} \cdot \sum_{t}^{Ntrees} \ln(\alpha_t) \cdot h_t(\mathbf{x}) \tag{6.10}$$

Values of y_{BDT} which are closer to +1 (-1) are considered to be more signal-like (background-like).

Variable importance is a measure of how often a variable is chosen as the variable to split on at a node in all of the trees. Variables can be ranked in order of variable importance and this can be used to motivate the final selection of variables used in the BDTs.

6.6 Reconstruction of hadronic top quarks

In the $\mathrm{t\bar{t}t\bar{t}}$ single lepton channel there are three hadronically decaying top quarks while there is only one hadronically decaying top quark in the semi-leptonic $\mathrm{t\bar{t}}$ final state. Equivalently in the dilepton channel there are two (zero) hadronically decaying top quarks in $\mathrm{t\bar{t}t\bar{t}}$ ($\mathrm{t\bar{t}}$). Therefore reconstructing top quarks from their hadronic decay products should be a powerful way to separate the signal and background processes. The jets which are in addition to the hard process in $\mathrm{t\bar{t}}$ are likely to come from initial or final state radiation (ISR or FSR) or from pileup, hence it should be unlikely that a hadronic top quark could be reconstructed from these additional jets. Due to the large number of jets within the selection, it is a challenge to find the right combination of jets which originated from a top quark. This motivates using multivariate analysis in order to rank each combination of three jets (tri-jet) according to which is most likely to be the combination that originated from a real hadronically decaying top quark. A training sample from $\mathrm{t\bar{t}}$ events is provided to the BDT where the Monte Carlo truth information is used to classify whether a tri-jet combination was a *good* combination which originated from a top quark or a *bad* combination which is formed from random jets. The following variables were selected for use within the BDT:

Tri-jet invariant mass - Good tri-jet combinations should have an invariant mass distribution which peaks around the top mass. Bad tri-jet combinations will have a much broader distribution.

Di-jet invariant mass - The di-jet combination is formed from the two jets with the smallest ΔR separation. The invariant mass distribution should peak around the W mass for good tri-jet combinations.

$\mathbf{p_T^{Rat}}$ - This is the ratio of the modulus of the vectorial sum of the p_T to the scalar sum of the p_T of the jets in the tri-jet combination. This is likely to be smaller in a random combination of jets where the vectorial p_{T} of each jet will cancel out more than in a good tri-jet combination.

$\Delta\phi_{\text{T-W}}$ - This is the $\Delta\phi$ between the tri-jet and di-jet system which should be smaller in good tri-jet combinations.

$\Delta\phi_{\text{T-b}}$ - This is the $\Delta\phi$ between the tri-jet and remaining jet not included in the di-jet system which should be smaller for good tri-jet combinations.

CSV$_{\mathrm{j}}$ - For the jet not used in the di-jet system, the CSV b-tagging discriminator value is used. If the di-jet system correctly identifies the quark jets coming from a W boson decay then the remaining jet in a good tri-jet combination should be a b-jet and hence will have a higher CSV b-tagging discriminator than a typical randomly selected jet.

The AdaBoost BDT algorithm is trained on the six variables above, which are shown in 6.9 for the Run 1 analysis. Combinations of tri-jets which come from a top quark in simulation are considered to be the signal and tri-jets which come from random combinations of jets and do not originate from a top quark are considered as the background sample. It produces a discriminator value for each tri-jet combination. Higher BDT discriminator values are associated with tri-jet combinations which are more likely to have come from a top quark. Each tri-jet combination is ranked according to the BDT discriminator value. In the dilepton channel the value for the highest ranked tri-jet, known as BDT$_{tri-jet1}$, can be used as a discriminating variable between $\mathrm{t\bar{t}t\bar{t}}$ and $\mathrm{t\bar{t}}$. In the single lepton channel the three jets which make up the highest ranked tri-jet combination are removed from the collection of jets and the process is repeated. The discriminator value of the next highest ranked tri-jet, BDT$_{tri-jet2}$, in this reduced jet collection can be used to distinguish between $\mathrm{t\bar{t}t\bar{t}}$ and $\mathrm{t\bar{t}}$. BDT$_{tri-jet2}$ is shown in Fig. 7.8.

6.6.1 Reduced Variables

As mentioned above, BDT$_{tri-jet2}$ is calculated from a reduced jet collection where the three jets from the highest-ranked tri-jet are removed from the jet collection.

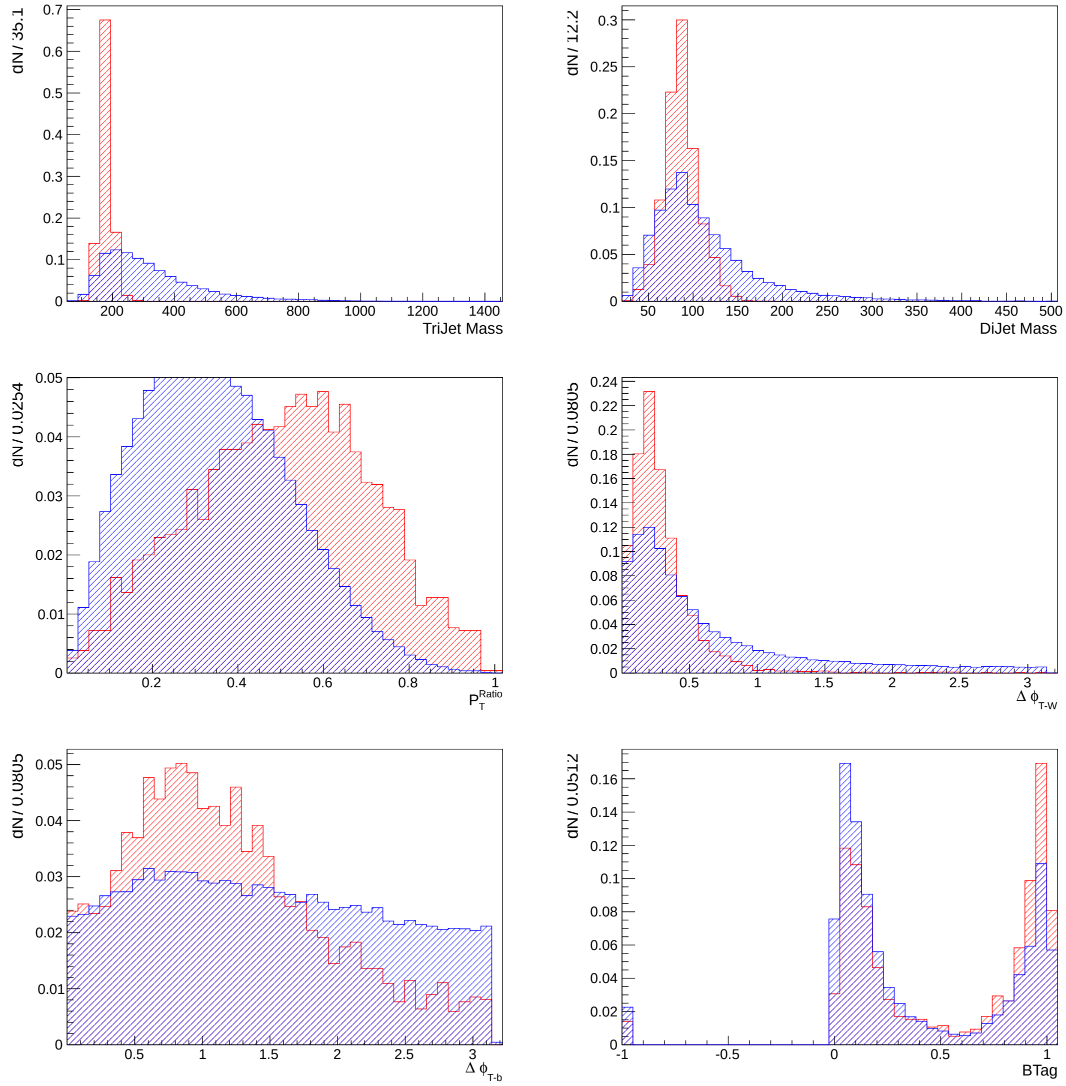


Figure 6.9: Input variables into hadronic top quark reconstruction BDT including: the invariant mass of trijets (top-left), the invariant mass of dijets (top-right), the p_{T}^{Ratio} (middle-left), $\Delta\phi_{T-W}$ between the top quark and dijet (middle-right), $\Delta\phi_{T-b}$ between the top quark and bottom jet (bottom-left) and BTag, which is the CSV value for the jet not in the di-jet (bottom-right).The blue distributions denotes the background consisting of random combinations of tri-jets. The red distributions are the signal, which consists of tri-jets which originate from top quarks in simulation.

In $\mathrm{t\bar{t}}$ events, the removal of the jets most likely to form a true top quark leave softer jets remaining in the reduced jet collection, usually from ISR, FSR or PU rather than the hard process of $\mathrm{t\bar{t}}$ production. In $\mathrm{t\bar{t}t\bar{t}}$ events the removal of the leading hadronic top quark candidate potentially leaves behind two additional hadronic top quarks, where some of the jets may not be reconstructed leaving harder jets from the hard process than what remains in $\mathrm{t\bar{t}}$ events. Therefore some discriminating variables can be formed from the reduced jet collection, as follows:

HT$_X$ - This is the H_T of the reduced event which should be higher for signal $\mathrm{t\bar{t}t\bar{t}}$ events, see Fig. 9.13.

SumJetMass$_\mathbf{X}$ - Invariant mass of all jets contained in the reduced event which should be higher for signal $\mathrm{t\bar{t}t\bar{t}}$ events, see Fig. 7.10.

6.7 Event-level BDT

A second BDT is employed to increase the separation between the $\mathrm{t\bar{t}}$ background and $\mathrm{t\bar{t}t\bar{t}}$ signal beyond what can be achieved with a simpler variable. Several discriminating variables have already been formed from the hadronic top reconstruction; BDT$_{tri-jet1}$, BDT$_{tri-jet2}$, HT$_X$, SumJetMass$_\mathrm{X}$. More discriminating variables can be formed from the event information, as follows:

b-jet content

As the branching ratio of top quarks to a b quark and a W boson is $\approx$ 100%, the main background process, $\mathrm{t\bar{t}}$ + $\mathrm{l\bar{l}}$, typically produces two b-quarks while $\mathrm{t\bar{t}t\bar{t}}$, typically produces four. Hence, the presence of more than two b-tagged jets is a potentially important source of discriminating power. At 8 TeV this was exploited by looking at the number of CSVM b-tags in an event, $\mathrm{N^{M}_{tags}}$ (see Fig. 7.2). At 13 TeV the CSV values for the *third-highest CSV* and *fourth-highest CSV* (see Figs. 9.2 and 9.2, respectively) are used to discriminate between signal and background and it is expected that these values will be lower for $\mathrm{t\bar{t}}$ + $\mathrm{l\bar{l}}$ events where the third-highest CSV and fourth-highest CSV are not likely to be b-jets. However for $\mathrm{t\bar{t}}$ + $\mathrm{b\bar{b}}$ events these variables do not provide much discrimination power. Another variable used is the scalar sum of the H_T of all CSVM b-jets in the event which should be higher for $\mathrm{t\bar{t}t\bar{t}}$ events, $\mathrm{H^{b}_{T}}$ (see Fig. 7.11).

Event-Activity

One of the most obvious variables to choose to distinguish between $\mathrm{t\bar{t}}$ and $\mathrm{t\bar{t}t\bar{t}}$ is the number of jets ($\mathrm{N_{jets}}$) as on average $\mathrm{t\bar{t}t\bar{t}}$ events have a higher number of jets than $\mathrm{t\bar{t}}$ (see Fig. 7.1). In semi-leptonic $\mathrm{t\bar{t}}$ events there are up to four hard jets from the hard process compared to up to ten from $t\bar{t}t\bar{t}$. Therefore, the fifth and sixth jet p_{T} can be used to distinguish between the event types (see Figs. 7.14). The same idea is used to form the variable $\mathrm{H_T^{Rat}}$ which is the ratio of the scalar sum of the H_{T} from the four highest p_{T} jets to the scalar sum of the H_{T} of the other jets in the events (see Fig. 7.13). This variable should be smaller for $t\bar{t}t\bar{t}$ events where there are more than four jets coming from the hard process. The *centrality* of the event is defined as the ratio of the H_{T} in the event to the H in the event, where H is the scalar sum of the total momentum (P) in the event (see Fig. 7.12).

The lepton p_{T} is used as a training variable as it is observed to be slightly larger in $t\bar{t}t\bar{t}$ events than in $t\bar{t}$ events (see Fig. 9.22).

The *Weighted jet multiplicity* ($\mathrm{N_j^W}$) takes into account a combination of the p_{T} spectra of the jets and the differences in the jet multiplicity (see Fig. 9.22). It is sensitive to the differences between the p_{T} spectra of the jets from a top quark decay and those originating from ISR/FSR. The $\mathrm{N_j^W}$ definition is given in Eq. 6.11 where p_T^{th} is the p_{T} threshold above which a jet is counted, $N_j\left(p_T > p_T^{th}\right)$ is the number of jets above the p_{T} threshold, $p_T^{low(up),i}$ takes values from the set [30 GeV; $p_{T,1}$; $p_{T,2}\cdots$; p_{T,N_j}] ([$p_{T,1}$; $p_{T,2}\cdots$; p_{T,N_j}; 125 GeV]), given that $p_{T,i}$ is the p_{T} of each jet in ascending order. The lower kinematic threshold is determined by the jet p_{T} requirement and the upper threshold is determined by the median jet p_{T}. The prefactor arises from the integral in the denominator combined with constant factors in the numerator.

$$N_j^W = \frac{\int_{30}^{125} N_j\left(p_T > p_T^{th}\right) \cdot p_T^{th}\, dp_T^{th}}{\int_{30}^{125} p_T^{th}, dp_T^{th}} = \frac{1}{14725} \sum_{i=1}^{N_{jets}} N_j\left(p_T > p_T^{low,i}\right) \cdot \left(p_T^{th}\right)^2 \Big|_{p_T^{low,i}}^{p_T^{up,i}}, \tag{6.11}$$

6.7.1 Training

Each analysis from the single-lepton channel at 8 TeV, the single-lepton channel at 13 TeV and the dilepton channel at 13 TeV, uses different subsets of the variables described above. The variables which are only used in the dilepton channel have not been described here and can be found in Ref [109]. The optimised set of variables for each training are provided to the TMVA package where the AdaBoost boosting algorithm is employed. The $\mathrm{t\bar{t}t\bar{t}}$ signal is trained against the main $\mathrm{t\bar{t}}$ background only as the other backgrounds are comparatively small. The *event-level* BDT will return a discriminator value which will be closer to +1 (-1) for signal-like (background-like) events. The treatment of negative weights in the $\sqrt{s} = 13$ TeV simulation will be discussed in Section 9.7.3.

6.8 Systematic uncertainties

Both the modelling of the simulation and the efficiency and resolution of the detector are not perfect. Hence, they bring some uncertainty to the measurement in addition to the theoretical uncertainties. This needs to be taken into account when fitting the simulation to the data as described in Section 6.9. There are two categories of systematic uncertainties; (i) Uncertainties which affect the normalisation of the distributions and (ii) uncertainties which affect the shape of the distributions. The quantity causing the uncertainty in each case can be varied up and down to produce alternative histogram templates.

6.8.1 Normalisation uncertainties

Normalisation uncertainties include the following:

- **Luminosity**

  There is uncertainty on the luminosity as measured by the CMS Luminosity group which affects the normalisation of all simulation samples as they are scaled to the measured luminosity.

- **Monte Carlo simulation cross sections**

  There is uncertainty on the cross sections given for each of the background processes as derived from theory. As $\mathrm{t\bar{t}}$ is the main background to $\mathrm{t\bar{t}t\bar{t}}$, the uncertainty on its MC cross-section is expected to be dominant over the other background processes.

- **Lepton ID, Iso and trigger SF**

  An uncertainty arises from the choice of lepton triggers and identification criteria in the baseline selection which is applied to all simulation data sets.

6.8.2 Shape uncertainties

Shape uncertainties arise from the following quantities.

- **Factorisation and renormalisation scales** This factorisation and renormalisation scale (μ_f, μ_s) can be shifted up and down at matrix element (ME) level and at parton shower (PS) level within the simulation. A shift in the Q^2 scale at PS level is equivalent to changing the value of α_S. Therefore the uncertainty from the jet multiplicity modelling in Section 6.3.6 can be included by inflating the PS scale systematic.

- **Matching threshold**

  The uncertainty due to the choice of matching threshold between the matrix element calculation and the parton shower is evaluated by producing alternative simulation samples where the matching threshold is varied between 20 GeV and 40 GeV.

- **Generator choice**

  The uncertainty on the choice of generator for the main $\mathrm{t\bar{t}}$ background is evaluated by considering an alternative $\mathrm{t\bar{t}}$ MC generator. The difference between the number of events in each bin of the output BDT distributions is converted into a symmetric uncertainty around the template created by the chosen MC generator.

- **Jet Energy Scale**

  The uncertainty on the Jet Energy Corrections, described in Section 4.7, applied to the simulation is evaluated by varying the Jet Energy Scale (JES) by $\pm 1\sigma$.

- **JER**

  The energies of the jets in simulation are smeared to match the observed discrepancy between the jet energy resolution, JER, in data and simulation.

- **b-tagging**

  There is a significant uncertainty on the measurement of b-tagging scale factors which is particularly important as variables derived from b-tagging jets are used in the event-level BDT. The method of accounting for this uncertainty varies depending on the b-tag modelling used from Section 6.3.2.

- **Pile up**

  The minimum bias cross section used in the simulation which is used to derive the pile up scale factors is varied by ±5% to account for the uncertainty on the minimum bias cross section.

- $\sigma_{\mathrm{t\bar{t}b\bar{b}}}$ / $\sigma_{\mathrm{t\bar{t}jj}}$ **modelling**

  The uncertainty on the measurement of the ratio of $\sigma_{\mathrm{t\bar{t}b\bar{b}}}$ / $\sigma_{\mathrm{t\bar{t}jj}}$ is translated into an uncertainty on the scale factors applied to heavy flavour jets. An anti-correlated uncertainty is applied to light flavour jets.

The subsets of this list of uncertainties used for the Run 1 and Run 2 analyses and which data sets they are applied to are described in sections 7 and 9.

6.9 Limit setting

The nominal histogram templates in the BDT output discriminator disctribution for each background and signal, the normalisation systematic uncertainties for each source, and the alternative histograms for the up and down shape systematic uncertainties have now been described. All of these quantities are provided to the Higgs Combine Tool, which is based on the Roostats package [110], where a Maximum Likelihood Fit (MLF) is performed. In reality the MLF actually minimises the negative log likelihood ($-\ln \mathcal{L}$) as it is computationally simpler to work with because it becomes a sum of log likelihoods rather than a product of likelihoods.

The expected number of events in bin i, ν_i, is given in Eq. 6.12 where L is the luminosity, σ_j is the cross section for the source of events from process j, and ϵ_{ij} is the efficiency for source j to be in bin i derived from simulation.

$$\nu_i = \sum_{j=1}^{n_{source}} L\sigma_j\epsilon_{ij} \tag{6.12}$$

Equation 6.12 can be written in terms of the signal and backgrounds separately. The expected number of signal (background) events, which is dependent on the nuisance parameters represented by θ, is denoted as $s_i = s_i(\theta)$ ($b_i = b_i(\theta)$) for each bin i. It is defined as $s_i = L\sigma_S\epsilon_{is}$, where the signal cross section is σ_s and the efficiency is ϵ_{is}. The signal strength modifier is denoted as μ. In Eq. 6.13, a number of background sources (n_{bkgd}) are summed using the index k. The total number of background events is summed over a n_{bkgd} sources $b_i = L\sum_{k=1}^{n_{bkgd}} \sigma_b^k \epsilon_{ib}^k$ where the cross section for each of the k backgrounds is σ_b^k with efficiency ϵ_{ib}^k.

$$\nu_i = \mu s_i(\theta) + b_i(\theta) \tag{6.13}$$

It is assumed that the number of observed events in bin i, n_i, will be poisson-distributed ($\mathcal{P}$) and hence the likelihood for the entire histogram is given in Eq. 6.14.

$$\mathcal{L} = \prod_{i=1}^{N} \mathcal{P}(n_i|\nu_i) = \prod_{i=1}^{N} \mathcal{P}(n_i|\mu s_i(\theta) + b_i(\theta)) \tag{6.14}$$

Each systematic uncertainty is modelled as a nuisance parameter in the fit. Normalisation uncertainties are modelled as lognormal nuisance parameters which is generally preferable to using a gaussian nuisance parameter. This is because lognormal nuisance parameters are better at modelling multiplicative uncertainties and do not allow the parameter to become negative. The shape uncertainties are modelled using vertical morphing [111, 112] between the three templates; the nominal template which has efficiency ϵ_{ij}^{0}, systematic shifted up with efficiency ϵ_{ij}^{+} and systematic shifted down with efficiency ϵ_{ij}^{-}. Quadratic interpolation is used between the two systematic templates and linear extrapolation beyond that range. The quadratic interpolation is shown in Eq. 6.15 where morphing parameter f has an Gaussian uncertainty with $\sigma_f = 1$.

$$\epsilon_{ij} = \frac{f\left(f-1\right)}{2}\epsilon_{ij}^{-} - \left(f-1\right)\left(f+1\right) + \epsilon_{ij}^{0}\frac{f\left(f+1\right)}{2}\epsilon_{ij}^{+} \tag{6.15}$$

From the MLF, the number of background events from each source and the number of signal events can be extracted from the *post-fit* distributions

6.9.1 CL$_\text{S}$ method

The procedure used for setting limits on the strength of the signal process is the modified frequentist method known as the *CL$_S$ method*. A test statistic, $\tilde{q}$, can be defined to distinguish between the background only (b-only) hypothesis and signal+background (s+b) hypothesis which includes the $\mathrm{t\bar{t}t\bar{t}}$ signal in the case of this analysis. The most common test statistic, and the one used by CMS and ATLAS, is the log likelihood ratio defined in Eq. 6.16. The nominal signal expectation can be scaled by the signal strength modifier μ. The test statistic $\tilde{q}_\mu$ is defined for an underlying model with signal strength modifier μ. The global maximum likelihood estimators for the signal strength modifier and the vector of nuisance parameters are $\hat{\mu}$ and $\hat{\theta}$. The maximum likelihood estimators of θ for a given μ are denoted $\hat{\theta}_\mu$. Equation 6.16 has the constraint of $0 \leq \hat{\mu} \leq \mu$ to ensure that the μ is positive and to guarantee a one-sided confidence interval. In Eq. 6.16, “data” can mean either the *real observed data* or *pseudodata*.

$$\tilde{q}_\mu = -2\ln\frac{\mathcal{L}\left(data|\mu s + b, \hat{\theta}_\mu\right)}{\mathcal{L}\left(data|\hat{\mu} s + b, \hat{\theta}\right)} \tag{6.16}$$

Larger values of $\tilde{q}_\mu$ are more“background-like". Using real data $\tilde{q}_\mu^{obs}$ can be derived for a particular value of μ using Eq. 6.16. Pseudodata is generated using toy MC and can be b-only ($\mu = 0$) or s+b for various signal strength modifier values. From this one can obtain the probability density function for $\tilde{q}_\mu$ given pseudodata containing a signal with strength μ in the s+b hypothesis, $f\left(\tilde{q}_\mu|\mu s + b, \theta_\mu^{obs}\right)$. For the b-only hypothesis, b-only pseudodata is generated to produce the probability density function is $f\left(\tilde{q}_\mu|b, \theta_0^{obs}\right)$.

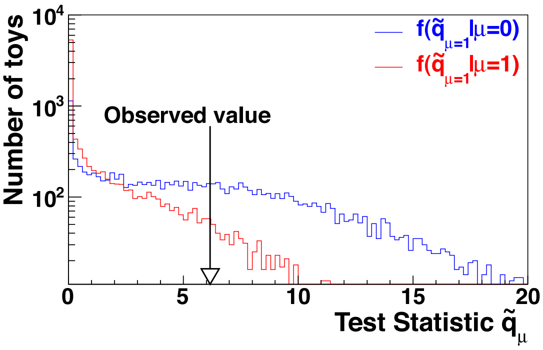


Figure 6.10: “Test statistic distributions for ensembles of pseudo-data generated for signal+background and background-only hypotheses.” [113].

The variable $p_{s+b}(\mu)$ can be defined for s+b hypothesis for a given signal strength modifier, as shown in Eq. 6.17. The value of $p_{s+b}(\mu)$ denotes the probability of obtaining a value of $\tilde{q}_\mu$ which is equal or less likely to happen relative to the q_{obs} given the s+b hypothesis.

$$p_{s+b}(\mu) = \int_{\tilde{q}_\mu^{obs}}^{\infty} f\left(\tilde{q}_\mu|\mu, \theta_\mu^{obs}\right) d\tilde{q}_\mu \tag{6.17}$$

One can similarly define p_b, however it is better to look at the definition of $1 - p_b$ in Eq. 6.18 as this is used in the $\mathrm{CL_S}$ definition in Eq. 6.19.

$$1 - p_b = \int_{q_0^{obs}}^{\infty} f\left(\tilde{q}_\mu|0, \theta_0^{obs}\right) d\tilde{q}_\mu \tag{6.18}$$

$$\mathrm{CL_S}\left(\mu\right) = \frac{p_{s+b}(\mu)}{1 - p_b} \tag{6.19}$$

It can be said that this model is excluded with a $(1-\alpha)$ confidence level (C.L.) if $\mathrm{CL_S} < \alpha$ for a given signal strength modifier, μ. In this thesis 95% Confidence Level upper limits are set on four top quark production within the standard model which corresponds to signal strengths which have $\mathrm{CL_S} < 0.05$ being excluded.

6.9.1.1 Asymptotic approximation to $\mathrm{CL_S}$ method

Often the "Asymptotic $\mathrm{CL_S}$ method" is used instead when there is a large enough number of expected events. In this method, the use of an ensemble of toy MC samples can be avoided and instead replaced by one representative dataset known as the "Asimov dataset". In this dataset the expected number of events in each bin exactly matches the observed data. Using this method can significantly reduce computing time [114].

6.9.2 Categorisation

Categorising events according to the numbers of jets and b-tags allows higher purity categories to be formed in which the signal-to-background ratio is much larger in some bins than it was prior to categorisation The signal sensitivity will be increased in signal rich and background-depleted regions. The main $\mathrm{t\bar{t}}$ background can be constrained in the signal-depleted regions which effectively act as control regions as the background-to-signal ratio is so large.

7 Search for standard model $\mathrm{t\bar{t}t\bar{t}}$ production in Run 1 at $\sqrt{s} = 8$ TeV

7.1 Introduction

In this chapter, the SM production of four top quarks ($\mathrm{t\bar{t}t\bar{t}}$) is sought. An analysis of the full 2012 CMS data set of proton-proton collisions at $\sqrt{s} = 8$ TeV, which corresponds to an integrated luminosity of 19.6 fb^{-1}, of data is presented. SM $\mathrm{t\bar{t}t\bar{t}}$ production has a cross section of $\sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}} \approx 1.3$ fb at NLO in QCD with NNLO corrections [99, 115].

In Section 7.2 the data and simulated samples of the signal and background are described. Section 7.3 details the initial requirements imposed on the signal region while Section 7.4 describes the corrections made the simulation. The effects of the selection requirements are given in Section 7.5 and control distributions comparing data with simulated events are given in Section 7.6. The multi-jet background estimation is described in Section 7.7 and the multivariate techniques used to increase the discrimination power between signal and background are contained in Section 7.8. Discussion of the systematic uncertainties and extraction of the limit on the $\mathrm{t\bar{t}t\bar{t}}$ cross section are in Sections 7.9 and 7.10, respectively. The final result can be found in Section 7.11 whilst validation of the analysis can be found in Section 7.12. Finally, a candidate four-top-quark event is shown in Section 7.12 and a summary and discussion of the ATLAS $\mathrm{t\bar{t}t\bar{t}}$ cross section limit can be found in Sections 7.14 and 7.15, respectively. Sections 7.7, 7.10.1 and 7.12 are the author's personal contribution to the analysis.

7.2 Data and Simulation

This analysis uses data from proton-proton collision at the CMS experiment in 2012 at $\sqrt{s} = 8$ TeV. For the muon (electron) channel, the data were collected using a trigger based on the presence of at least one muon (electron) candidate with $p_{\mathrm{T}} > 24$ (27) GeV and correspond to an integrated luminosity of 19.6 fb^{-1}. The signal SM $t\bar{t}t\bar{t}$ Monte Carlo (MC) samples and the background MC samples are given in Table 7.1, along with the MC generators used to produce these samples, the perturbative accuracy in QCD at which they were produced and the number of events produced. Simulated samples were produced for the scale and matching systematic uncertainties. These can be found in Table 7.2. In this analysis the scale uncertainty includes the ME scale and PS scale.

Dataset	Events	Generator	Order
$t\bar{t}t\bar{t}$	100K	MadGraph	LO
$t\bar{t}$ Semi leptonic	25M	MadGraph	LO
$t\bar{t}$ Hadronic	31M	MadGraph	LO
$t\bar{t}$ Dileptonic	12M	MadGraph	LO
W + 4 Jets $\rightarrow$ lν	13M	MadGraph	LO
$\bar{t}$ tW-channel	500K	POWHEG	NLO
t tW-channel	500K	POWHEG	NLO
t t-channel	3.8M	POWHEG	NLO
$\bar{t}$ s-channel	260K	POWHEG	NLO
t s-channel	139K	POWHEG	NLO
$\bar{t}$ t-channel	2M	POWHEG	NLO
DYJets $\rightarrow$ ll	6.7M	MadGraph	LO
$t\bar{t}$ Z	200K	MadGraph	LO
$t\bar{t}$ W	200K	MadGraph	LO
$t\bar{t}$ H$\rightarrow$ HToBB	1M	PYTHIA 6	LO
ZZ	10M	PYTHIA 6	LO
WZ	10M	PYTHIA 6	LO
WW	10M	PYTHIA 6	LO

Table 7.1: Dataset name, total number of events, MC generator and order of the simulated samples. PYTHIA 6 was used to hadronise all samples in this table.

Dataset	Events	Generator	Order
$t\bar{t}$ scale down	5M	MadGraph	LO
$t\bar{t}$ scale up	5M	MadGraph	LO
$t\bar{t}$ matching down	5M	MadGraph	LO
$t\bar{t}$ matching up	5M	MadGraph	LO

Table 7.2: Dataset name, total number of events, MC generator and order of the simulated systematic samples. PYTHIA 6 was used to hadronise all samples in this table.

7.3 Baseline Event Selection

To select $t\bar{t}t\bar{t}$ events and suppress background events, a set of criteria are applied to the reconstructed objects in events which are triggered by the single muon or single electron triggers. For the muon channel these are:

- Exactly one tight muon
- Exactly zero additional loose muons
- Exactly zero loose electrons
- At least 6 jets with $p_{\mathrm{T}} > 30$ GeV
- At least 2 CSVM tagged b-jets
- $H_{\mathrm{T}} > 400$ GeV
- $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV

For the electron channel these are:

- Exactly one tight electron
- Exactly zero additional loose electrons
- Exactly zero loose muons
- At least 6 jets with $p_{\mathrm{T}} > 30$ GeV
- At least 2 CSVM tagged b-jets
- $H_{\mathrm{T}} > 400$ GeV
- $E_{\mathrm{T}}^{\mathrm{miss}} > 30$ GeV

7.4 Corrections to the simulation

All corrections described in section 6.3 are applied to the simulation samples. The events for all background and signal samples which pass the baseline event selection are given a weight which is the product of the weights for the pile up corrections, a weight to correct the b-tag modelling from the method in 6.3.2.1 and a weight for lepton modelling corrections. The heavy flavour jet modelling correction from section 6.3.3 is applied to the $\mathrm{t\bar{t}}$ background.

7.5 Effect of selection requirements

The event counts, after weighting, are given for the muon (electron) channel in Table 7.3 (Table 7.4). It is quite evident that the main background is from $\mathrm{t\bar{t}}$ production, with smaller contributions coming from W+jets, Z+jets and single top (ST) as well as very small contributions from the diboson (WW, WZ, ZZ), $\mathrm{t\bar{t}}H$ and $\mathrm{t\bar{t}}$V background (V = Z, W).

	Data	$\mathrm{t\bar{t}t\bar{t}}$	*ttH*	*Wjets*	*Zjets*	*ST*	Diboson	$\mathrm{t\bar{t}}V$	$\mathrm{t\bar{t}}$
Trigger and PV	121 000	4.50	172	16 100	3920	3260	505	497	84 500
1 iso. μ	95723	3.50	142	13 700	2640	2750	405	398	70 000
Loose μ veto	93 400	3.20	139	13 700	1840	2740	381	376	69 500
Loose e veto	91 500	2.50	133	13 600	1810	2690	375	350	68 000
$\mathrm{N_{jets}} \geq 6$	24 800	2.30	59.1	2430	351	592	66.6	166	21 200
$\mathrm{N^{M}_{tags}} \geq 2$	9260	1.70	46.0	68.8	15.2	216	5.10	74.5	9140
HT $\geq$ 400 GeV	6340	1.60	37.7	49.3	10.4	157	4.00	62.4	6540
$E_{\mathrm{T}}^{\mathrm{miss}} \geq 30$ GeV	5220	1.50	31.7	41.5	7.10	132	3.20	52.7	5420

Table 7.3: Number of observed events in data and expected events in simulation after successive selection requirements in the μ + jets channel ($\mathcal{L} = 19.6$ fb).

	Data	$\mathrm{t\bar{t}t\bar{t}}$	*ttH*	*Wjets*	*Zjets*	*ST*	Diboson	$\mathrm{t\bar{t}}V$	$\mathrm{t\bar{t}}$
Trigger and PV	1 130 000	8.60	308	28 800	25 100	6490	1380	986	157 000
1 iso. e	104 000	3.20	118	9630	7180	2210	415	364	78 900
Loose e veto	101 000	3.00	116	9610	5350	2200	382	347	63 800
Loose mu veto	99 400	2.30	111	9610	5340	2170	415	321	62 600
$\mathrm{N_{jets}} \geq 6$	26 500	2.10	55.2	1660	1110	491	59.8	152	19 800
$\mathrm{N^{M}_{tags}} \geq 2$	8950	1.50	43.2	45.4	19.8	179	4.10	62.1	8360
HT $\geq$ 400 GeV	6280	1.50	35.6	33.2	13.9	134	2.80	52.2	6060
$E_{\mathrm{T}}^{\mathrm{miss}} \geq 30$ GeV	5070	1.30	30.0	25.4	7.00	114	2.10	46.0	4970

Table 7.4: Number of observed events in data and expected events in simulation after successive selection requirements in the e + jets channel ($\mathcal{L} = 19.6$ fb).

7.6 Control distributions between data and simulation

The distributions which show the agreement between data and simulation for $\mathrm{N_{jets}}$, $\mathrm{N^{M}_{tags}}$, H_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$ after the baseline event selection and corrections are applied are in Figs 7.1, 7.2, 7.3, 7.4. The background distributions are stacked and hence the agreement between the data and the b-only hypothesis is shown. The signal distribution is overlaid and multiplied by a factor of 100 for visibility. The scale uncertainty is the largest uncertainty on the background and is shown as a hatched error band. The signal has a different distribution in these variables from the background processes, particularly in the $\mathrm{N_{jets}}$, $\mathrm{N^{M}_{tags}}$, H_{T} distributions. This makes them viable candidates to be used in the event-level BDT.

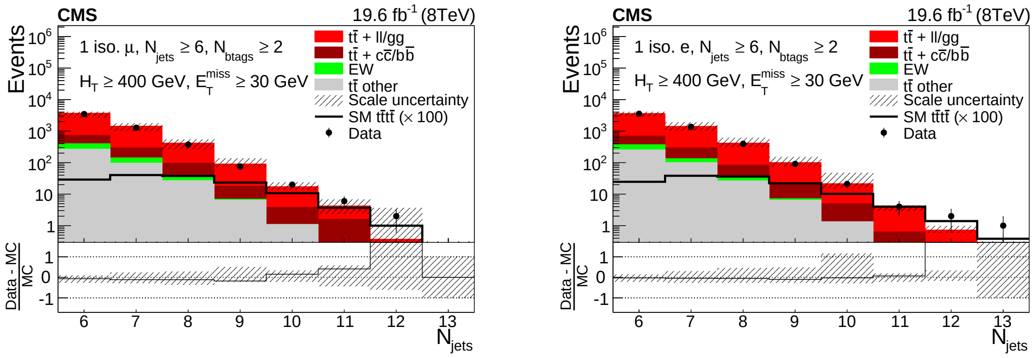


Figure 7.1: Distribution of $\mathrm{N_{jets}}$ after selection for μ + jets (left) and e + jets (right).

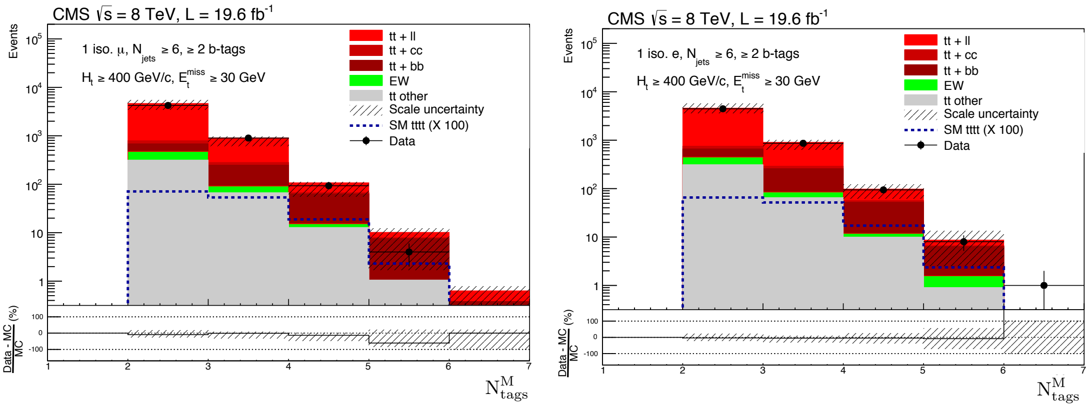


Figure 7.2: Distribution of $\mathrm{N^{M}_{tags}}$ after selection for μ + jets (left) and e + jets (right).

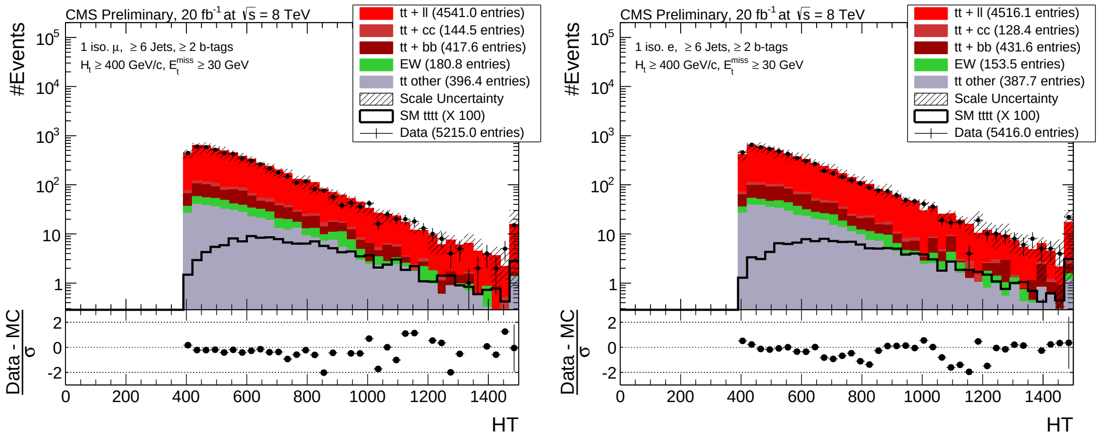


Figure 7.3: Distribution of H_{T} after selection for μ + jets (left) and e + jets (right).

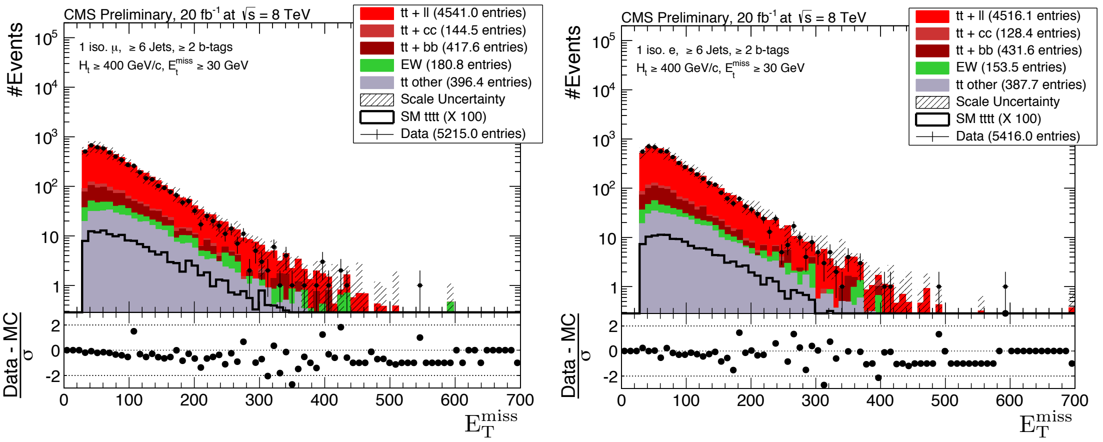


Figure 7.4: Distribution of $E_{\mathrm{T}}^{\mathrm{miss}}$ after selection for μ + jets (left) and e + jets (right).

There is a small discrepancy between the data and simulation in which the simulation overshoots data in the muon channel.

7.7 Multi-jet background estimation

The presence of multi-jet events within the signal region defined by the baseline selection has been investigated. It can be seen from the $E_{\mathrm{T}}^{\mathrm{miss}}$ distributions in Fig. 7.4 that the data agrees well with the simulation at low values. This suggests that there are very few multi-jet events which pass the tight requirements in the baseline selection. The data-driven ABCD method, which is described in Section 7.7, is used to estimated the multi-jet background. The defined regions in the uncorrelated variables of $E_{\mathrm{T}}^{\mathrm{miss}}$ and $RelIso$ are shown below. The upper bound in $RelIso$ is restricted by the minimum $RelIso$ values required by the HLT in the single muon and single electron data sets.

For the muon channel the bounds are:

- A : $30 < E_{\mathrm{T}}^{\mathrm{miss}} < 500$, $0.1 < \mathrm{RelIso} < 0.15$
- B : $0 < E_{\mathrm{T}}^{\mathrm{miss}} < 30$, $0.1 < \mathrm{RelIso} < 0.15$
- C : $0 < E_{\mathrm{T}}^{\mathrm{miss}} < 30$, $0 < \mathrm{RelIso} < 0.1$
- D : $30 < E_{\mathrm{T}}^{\mathrm{miss}} < 500$, $0 < \mathrm{RelIso} < 0.1$

and for the electron channel:

- A : $30 < E_{\mathrm{T}}^{\mathrm{miss}} < 500$, $0.12 < \mathrm{RelIso} < 0.2$
- B : $0 < E_{\mathrm{T}}^{\mathrm{miss}} < 30$, $0.12 < \mathrm{RelIso} < 0.2$
- C : $0 < E_{\mathrm{T}}^{\mathrm{miss}} < 30$, $0 < \mathrm{RelIso} < 0.12$
- D : $30 < E_{\mathrm{T}}^{\mathrm{miss}} < 500$, $0 < \mathrm{RelIso} < 0.12$

The results of the background subtraction from data and the defined ABCD regions are shown in Fig. 7.5. The number of multi-jet events, $\mathrm{N}_{\text{multi-jet}}$, in each of the control regions are used to predict the number of multi-jet events in region D using Eq. 7.1, the results of which are shown in Table 7.5.

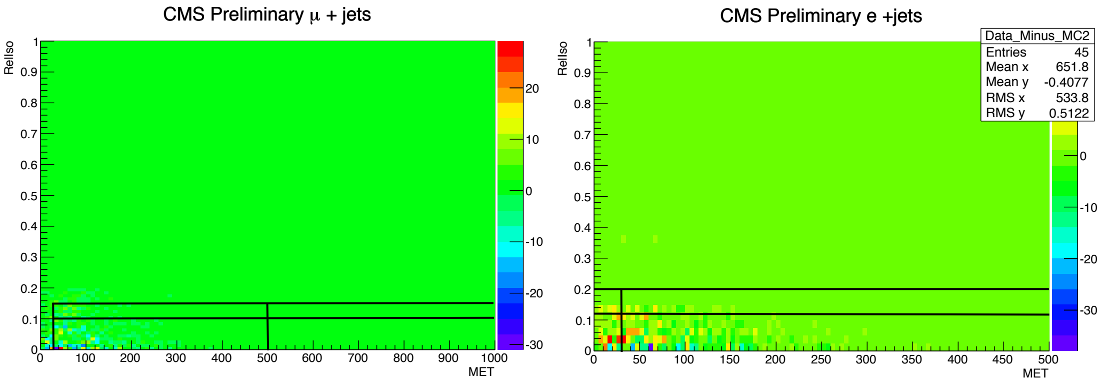


Figure 7.5: Number of events after selection requirements and background subtraction for μ + jets (left) and e + jets (right). The black lines shows the regions used in the ABCD method, as defined in the text.

$$\frac{\mathrm{N}^{\mathrm{B}}_{\text{multi-jet}}}{\mathrm{N}^{\mathrm{A}}_{\text{multi-jet}}} = \frac{\mathrm{N}^{\mathrm{C}}_{\text{multi-jet}}}{\mathrm{N}^{\mathrm{D}}_{\text{multi-jet}}} \tag{7.1}$$

Table 7.5: Multi-jet estimation

Channel	$\mathrm{N}^{\mathrm{A}}_{\text{multi-jet}}$	$\mathrm{N}^{\mathrm{B}}_{\text{multi-jet}}$	$\mathrm{N}^{\mathrm{C}}_{\text{multi-jet}}$	Prediction for $\mathrm{N}^{\mathrm{D}}_{\text{multi-jet}}$
μ + jets	19.1	16.1	-16.8	-20
e + jets	36.8	50.8	62.7	45.5

As the number of $\mathrm{t\bar{t}}$ events in simulation has fluctuated to be greater than the data in region C in the muon channel, the prediction for the signal region D is negative. As this prediction is unphysical, the number of events is estimated to be zero in the muon channel. In the electron channel, 45.5 multi-jet events are predicted which is considered negligible at $< 1\%$ of the large $\mathrm{t\bar{t}}$ background. Hence, the multi-jet background is not considered further.

7.8 Discriminating between signal and background

As seen from Tables 7.3 and 7.4, the dominant background process after the baseline selection is $\mathrm{t\bar{t}}$ production, which is approximately three orders of magnitude greater than $\mathrm{t\bar{t}t\bar{t}}$ production in the signal region. In this section, details of the hadronic top quark reconstruction (described in section 6.6) and the variables that can be defined from the output of this MVA will be discussed as well as the variables which are used in the event-level BDT.

There are three categories of observables which are used to discriminate: the number of top quarks which can be reconstructed in the event, the number of b-jets found in each event, and event activity such as H_T.

7.8.1 Hadronic top quark content

The anti-k_t algorithm can only distinguish separate jets if they are separated by $\Delta R > 0.5$, hence it may not always be possible to reconstruct all hadronic top quarks. To ascertain whether it is feasible to use the hadronic top reconstruction, described in Section 6.6, to aid in distinguishing $\mathrm{t\bar{t}t\bar{t}}$ and $\mathrm{t\bar{t}}$ given this criteria, the number of reconstructible tops in $\mathrm{t\bar{t}t\bar{t}}$ in $\mathrm{t\bar{t}}$ events is investigated. Hadronic top quarks are considered reconstructible at parton level if they decayed into partons with $\Delta R > 0.5$. Both $\mathrm{t\bar{t}t\bar{t}}$ and $\mathrm{t\bar{t}}$ events were analysed where there was one muon with $p_\mathrm{T} > 26$ GeV and exactly 6 jets with $p_\mathrm{T} > 30$ and $|\eta| < 2.5$. The number of reconstructible hadronic top quarks for $\mathrm{t\bar{t}t\bar{t}}$ and $\mathrm{t\bar{t}}$ are shown in Fig. 7.6.

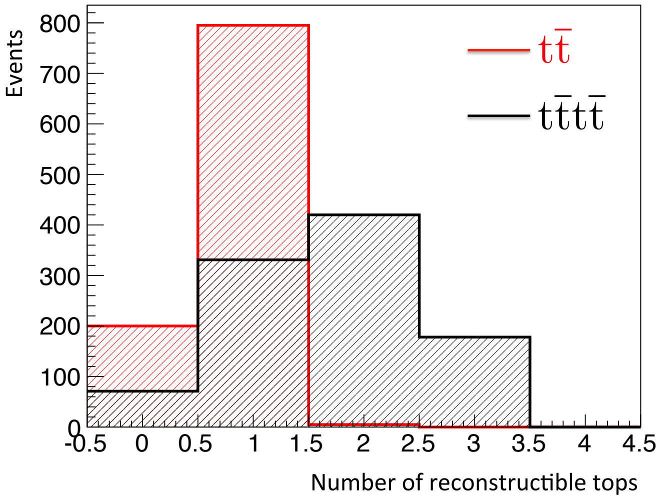


Figure 7.6: The number of reconstructible hadronic top quarks in semileptonic $\mathrm{t\bar{t}}$ and $\mathrm{t\bar{t}t\bar{t}}$ in the single lepton channel.

For $\mathrm{t\bar{t}t\bar{t}}$ production at parton level it is possible to reconstruct more than one hadronic top quark ≈ 63 % of the time compared to a negligible number of times in $\mathrm{t\bar{t}}$, as is evident in the Fig. 7.6. Hence, using the number of hadronic tops to distinguish between $\mathrm{t\bar{t}t\bar{t}}$ and $\mathrm{t\bar{t}}$ is worth investigating.

Hadronic top quark BDT

The hadronic top quark reconstruction BDT is trained on $\mathrm{t\bar{t}}$ events. The variables used in the BDT are shown in Fig. 6.9 where it can be seen that the tri-jet and dijet invariant masses provide the most powerful separation. This visualisation can help in finding discriminating variables but it should be considered that the shapes of these distributions will change as the events progress through the BDT due to correlations between variables. For instance the CSV_j variable appears to not have much separation power initially but the distribution may look quite different in certain areas of phase space or after boosting weights have been applied, and the inclusion of this variable improved the overall separation of the BDT output discriminator.

The distribution of the discriminator values output from the BDT is shown in Fig. 7.7, along with the distribution of the tri-jet vs BDT value. In the latter distribution, the features in the background shape are caused by the large discriminating power of the tri-jet and di-jet invariant mass. It can be seen in the tri-jet vs BDT output value distribution that tri-jet candidates with an invariant mass close to the top mass are associated with larger BDT output values and, in particular, there is a minimum in the BDT discriminator value at ≈ -0.5 which can be understood as the division between the two peaks in the BDT discriminator plot. Tri-jets with an invariant mass $\gtrsim 300$ GeV tend to have BDT discriminator values < -0.5. As signal tri-jets tend to have an invariant tri-jet mass close to the top mass and invariant di-jet mass close to the W boson, these entries are clustered to the right-hand side of Fig. 7.7, with values typically > 0.

BDT$_{tri-jet2}$

The distributions of the BDT$_{tri-jet2}$ variable, described in Section 6.6, are shown in Fig. 7.8 for the muon and electron channel. It can be seen that the data and simulation agree well. There is some difference between the shape of the signal and background distributions which can be exploited by the event-level BDT.

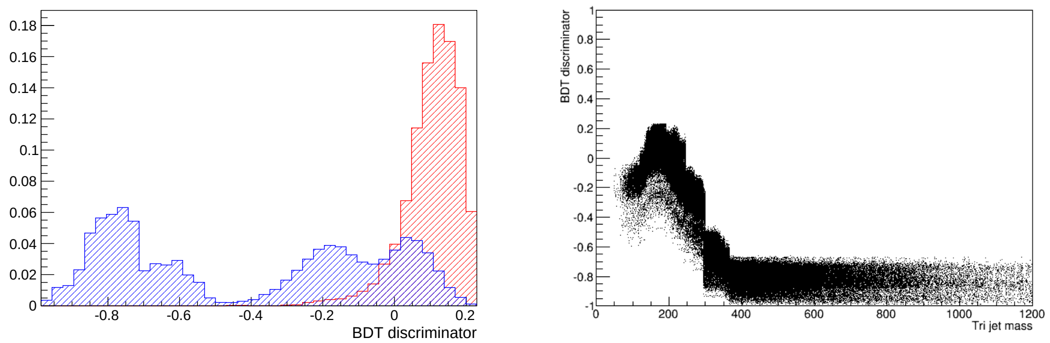


Figure 7.7: Output discriminator variable from top quark reconstruction BDT (left) and tri-jet mass vs BDT score (right). The blue distribution denotes the background consisting of random combinations of tri-jets. The red distribution is the signal, which consists of tri-jets which originate from top quarks in simulation.

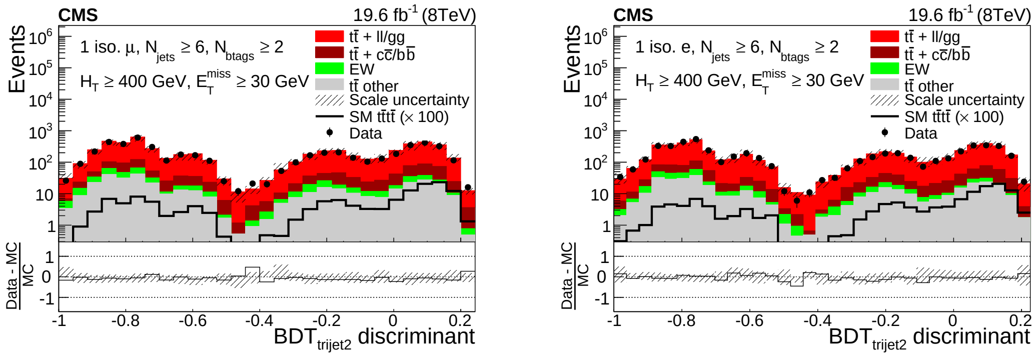


Figure 7.8: BDT$_{tri-jet2}$ in data and simulation for μ + jets (left) and e + jets (right).

Reduced Event Variables

The distributions for the HT$_\mathrm{X}$ and SumJetMass$_\mathrm{X}$ reduced variables, which are derived from the reduced event where jets from the highest ranked hadronic top quark candidate are removed, can be seen in Figs 9.13, 7.10. These variables have excellent agreement between data and simulation and it is apparent, particularly in the HT$_\mathrm{X}$ distribution, that these variables will provide some discrimination power in the event-level BDT.

7.8.2 Event activity and b-jet content variables chosen for the event-level BDT

The variables chosen for use in the event-level BDT due to their discrimination power include:

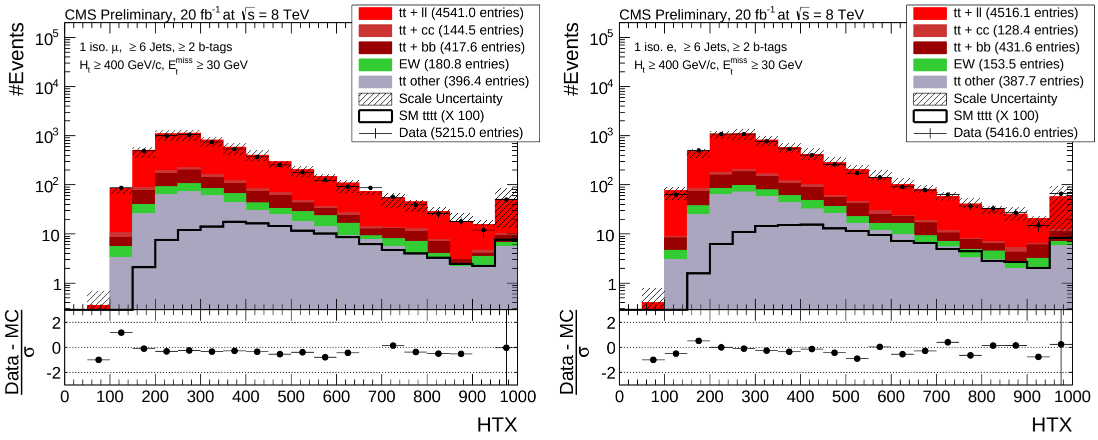


Figure 7.9: $\mathrm{HT_X}$ for μ + jets (left) and e + jets (right).

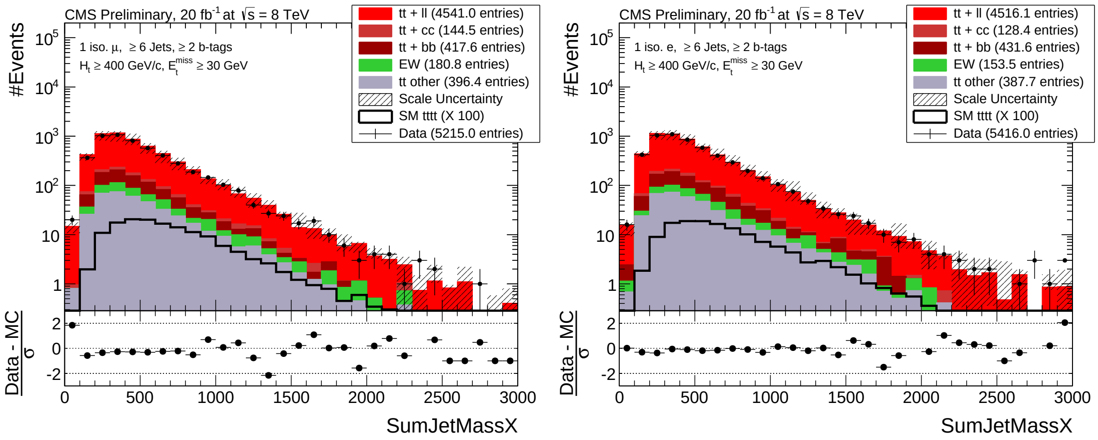


Figure 7.10: $\mathrm{SumJetMass_X}$ for μ + jets (left) and e + jets (right).

- $\mathrm{N_{jets}}$
- $\mathrm{N^M_{tags}}$
- $\mathrm{HT^b}$
- Centrality
- $\mathrm{H^{rat}_T}$
- 5th jet p_T
- 6th jet p_T
- $\mathrm{BDT}_{tri-jet2}$
- $\mathrm{SumJetMass_X}$
- $\mathrm{HT_X}$

These variables are fully described in section 6.5. All of the variables in Figures 7.11- 7.15 show good agreement between data and simulation and the discriminating power is quite apparent in the distributions of Centrality and $\mathrm{H^{rat}_T}$ in Fig. 7.12 and Fig. 7.13, respectively. Figure 7.2 shows the distribution of $\mathrm{N^M_{tags}}$ which again shows good agreement between data and simulation and it can be seen that the $t\bar{t}t\bar{t}$ signal is more likely to produce events with a higher number of N^{M}_{tags} than the background due to having four b-quarks in the final state.

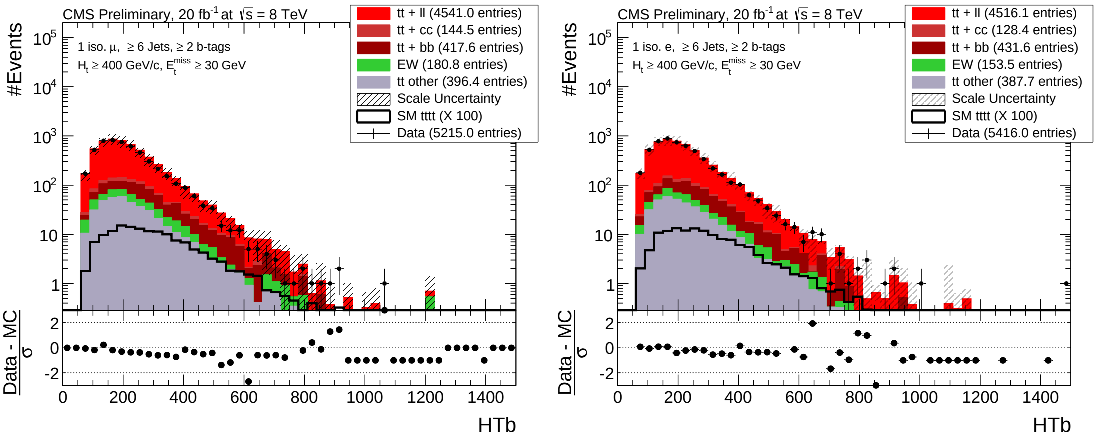


Figure 7.11: HT^{b} for μ + jets (left) and e + jets (right).

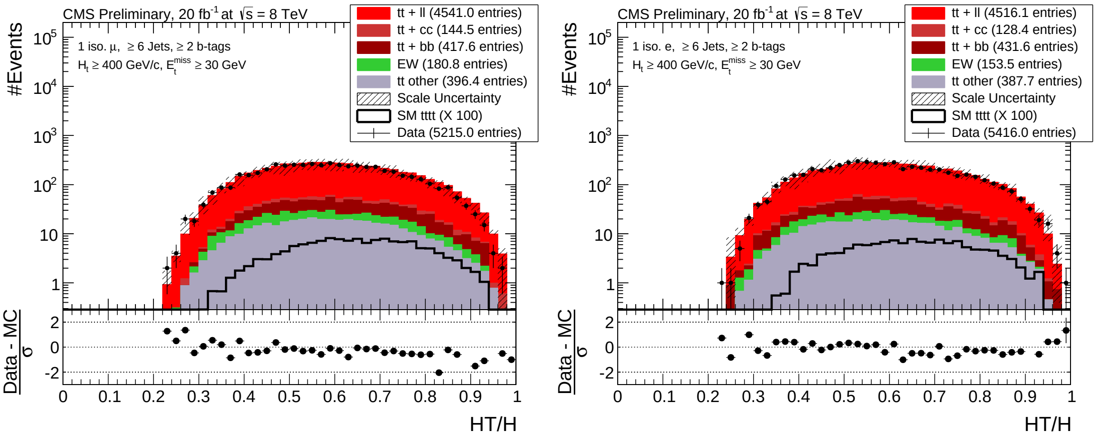


Figure 7.12: Centrality for μ + jets (left) and e + jets (right).

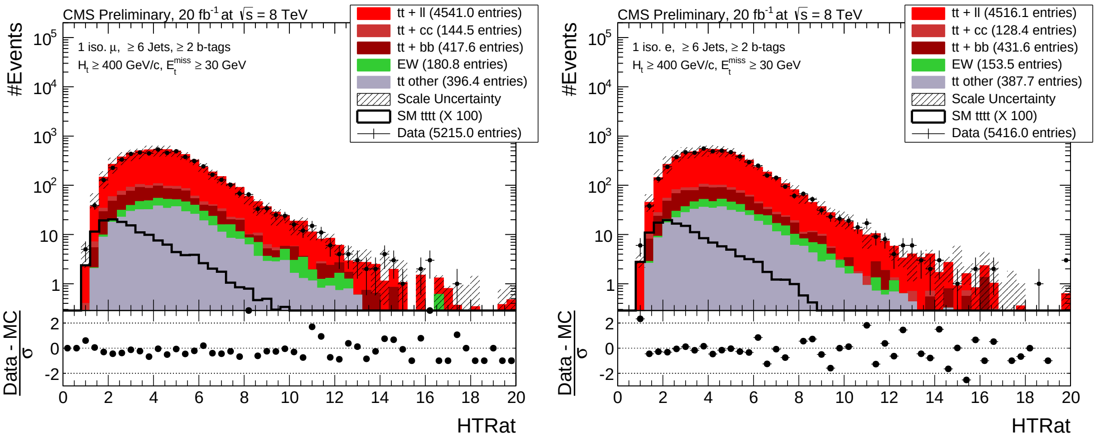


Figure 7.13: H_{T}^{rat} for μ + jets (left) and e + jets (right).

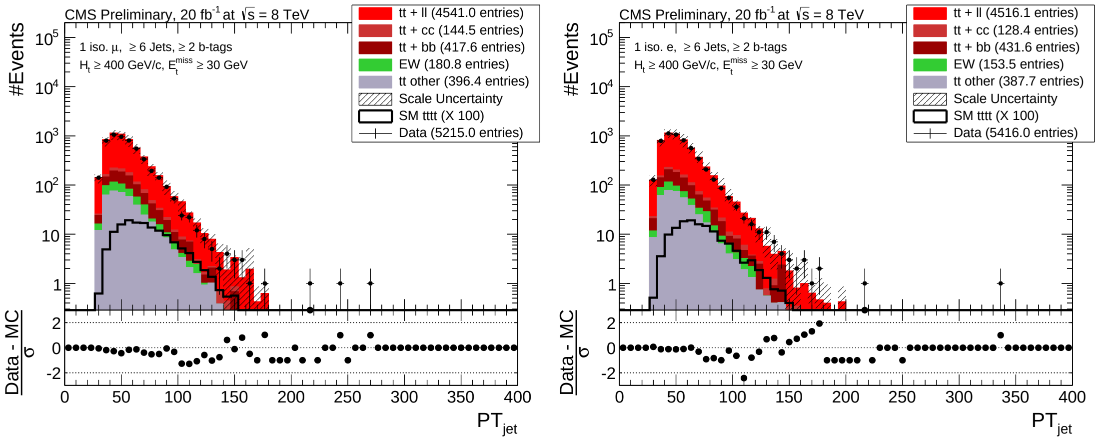


Figure 7.14: 5th jet p_T for μ + jets (left) and e + jets (right).

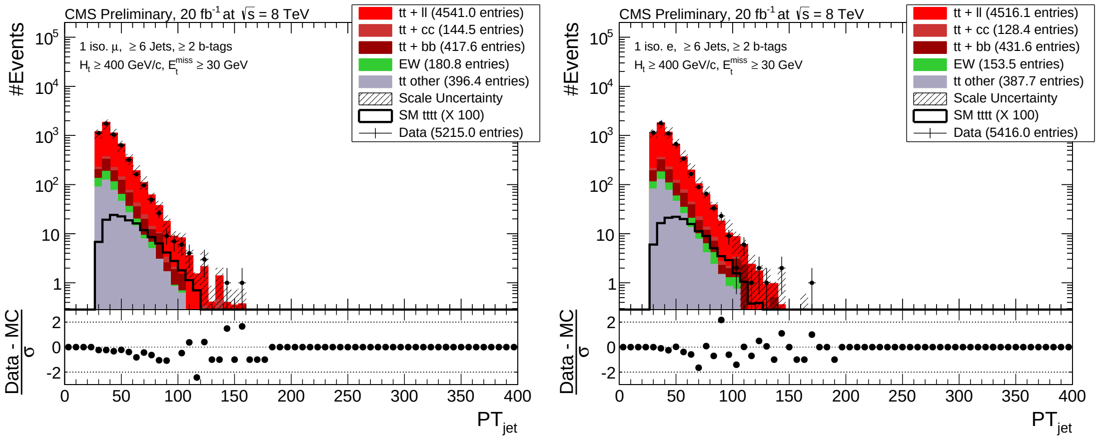


Figure 7.15: 6th jet p_T for μ + jets (left) and e + jets (right).

7.8.3 Event-level BDT training and output

A sample of events which were not used to train the kinematic reconstruction of top quarks are used to train the event-level BDT so that an orthogonal training sample can be provided. The output distribution of the BDT discriminator can be seen in Fig. 7.16. Again, there is good agreement between the data and simulation and it can be seen that there is an improved separation between the background distributions and the $t\bar{t}t\bar{t}$ distribution compared to the input variables to the BDT.

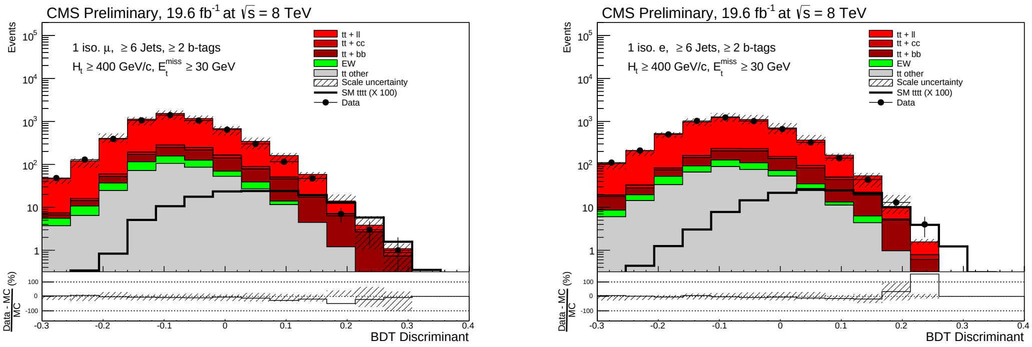


Figure 7.16: BDT discriminator variable in data and simulation for μ + jets (right) and e + jets (left).

7.9 Systematic uncertainties

All systematic uncertainties are described in section 6.8 and some further details which are specific to this Run 1 analysis are given below.

- **Luminosity**

  The CMS Luminosity Group gave a recommendation of 2.6% uncertainty on the luminosity [116].

- **Monte Carlo cross sections**

  The uncertainty on the $t\bar{t}$ cross section is $^{+2.5\%}_{-3.4\%}$ (renormalisation and factorisation scale) and $^{+2.5\%}_{-2.6\%}$ (PDF) [96]. The cross section uncertainties for the other background processes are modelled by assigning a 4% uncertainty and a 10% uncertainty is assigned to the signal process.

- **Factorisation and renormalisation scales**

  The alternative $t\bar{t}$ samples with ME and PS scale (u) varied by 2u, 1/2u were used as the alternative shapes.

- **Matching threshold**

  The alternative $t\bar{t}$ samples with the matching threshold changed between 20 GeV and 40 GeV were used as the alternative shapes.

- **JES**

  The JES uncertainty is derived by varying the JES by $\pm 1\sigma$.

- **JER**

  The JER uncertainty is derived by varying the smearing by $\pm 1\sigma$.

- **b tagging**

  The b tagging uncertainty is quantified by varying the scale factor by $\pm 1\sigma$.

- **Pile up**

  The PU systematic uncertainty is found by varying the MinBias cross section by $\pm 5\%$.

- $\sigma_{\mathrm{t\bar{t}b\bar{b}}}$ / $\sigma_{\mathrm{t\bar{t}jj}}$ **modelling**

  The uncertainty on the measurement of $\sigma_{\mathrm{t\bar{t}b\bar{b}}}$ / $\sigma_{\mathrm{t\bar{t}jj}}$ by CMS [105] is $\pm 0.4\,(\mathrm{stat.}) \pm 0.5\,(\mathrm{sys.})$. Alternative event weights are derived for $\sigma_{\mathrm{t\bar{t}b\bar{b}}}$ / $\sigma_{\mathrm{t\bar{t}jj}}$ which are used to provide the alternative systematic up and down templates.

The effects of the systematics are studied in Section 7.12.1.

Additional systematic uncertainties were studied which ultimately had a negligible impact on the analysis so were not included in the final result. For completeness they are discussed in Appendix C.

7.10 Template fit and upper limit

No excess was observed in the data. An upper limit was calculated on the signal strength, $\sigma_{\mathrm{t\bar{t}t\bar{t}}}$ / $\sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}$. Firstly, a binned maximum likelihood fit of the BDT distributions is made, as described in Section 6.9. This is performed using the Higgs Combine Tool [113] which is based on the statistical package RooStats [110, 114, 117, 118]. The background processes are grouped into the following templates for the fit: ‘$\mathrm{t\bar{t}}$’ for semi-leptonic $\mathrm{t\bar{t}}$, ‘electroweak’ for W + jets and Z + jets, and ‘$\mathrm{t\bar{t}}$ _other’ for all other $\mathrm{t\bar{t}}$ processes. Using the $\mathrm{CL_S}$ method [119, 120] with the asymptotic approximation, the best fit values of the nuisance parameters can be obtained along with the corresponding limit on $\sigma_{\mathrm{t\bar{t}t\bar{t}}}$ / $\sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}$. The cross section limit, $\sigma_{\mathrm{t\bar{t}t\bar{t}}}$, can be derived by multiplying $\sigma_{\mathrm{t\bar{t}t\bar{t}}}$ / $\sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}$ by $\sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}$ =1.30 fb.

7.10.1 Splitting into N_{jets} categories

To increase the sensitivity of the analysis, the BDT templates were split into three N_{jets} categories ($N_{\mathrm{jets}} = 6$, $N_{\mathrm{jets}} = 7$ & $N_{\mathrm{jets}} \geq 8$) for each of the muon and electron channels, as shown in Fig. 7.17. The higher N_{jets} categories benefit from a better ratio of signal to background events as $\mathrm{t\bar{t}t\bar{t}}$ events are likely to have a larger number of jets. As the opposite is true in the $N_{\mathrm{jets}} = 6$ category it helps to constrain the main background of $\mathrm{t\bar{t}}$.

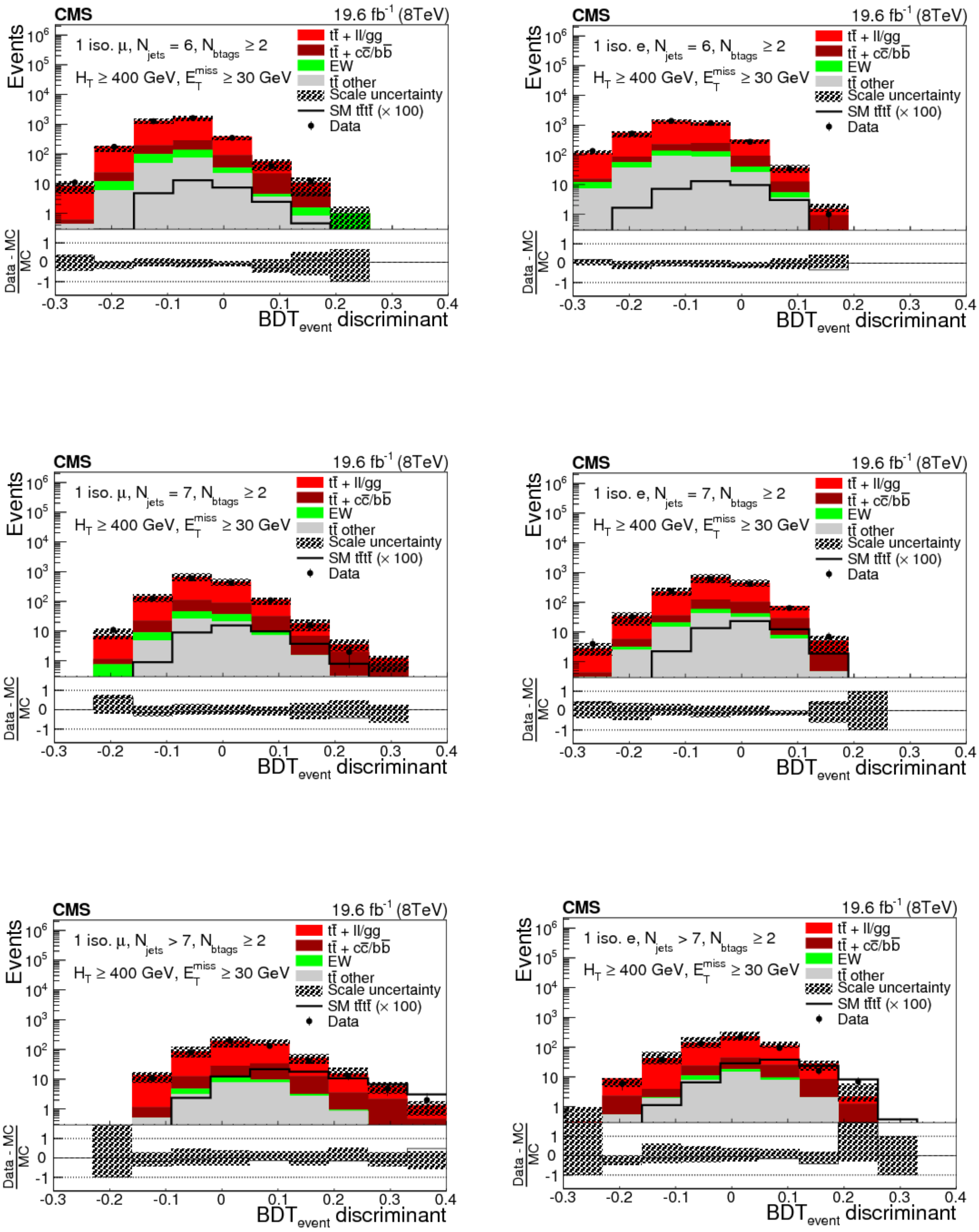


Figure 7.17: The BDT discriminator distributions in exclusive jet bins of $N_{\text{jets}} = 6$, $N_{\text{jets}} = 7$ and $N_{\text{jets}} \geq 8$, are shown for the μ + jets channel (left) and e + jets channel (right).

7.11 Results

A simultaneous fit of the three N_{jets} categories is made for the muon channel, the electron channel and a combination of the two. The limits extracted using the CL_S method are shown in Table 7.6.

Channel	Exp.	Obs.
μ + jets	35.6 ± 18	34
e + jets	36.1 ± 16	36
combined	24.6 ± 13	25

Channel	Exp. (fb)	Obs. (fb)
μ + jets	46.3 ± 23	44
e + jets	47.0 ± 20	47
combined	32.0 ± 17	32

Table 7.6: CLS limits on ($\sigma_{\mathrm{t\bar{t}t\bar{t}}}$ / $\sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}$) (left) and $\sigma_{\mathrm{t\bar{t}t\bar{t}}}$ (right).

In comparison, the corresponding limit for a simultaneous fit of the electron and muon channels in an inclusive N_{jets} category had an expected limit of 42^{+18}_{-13} fb [121]. Therefore, it can be seen that splitting into N_{jets} categories provides a significant improvement in the sensitivity of the analysis.

7.12 Cross checks on the analysis

7.12.1 Individual effects of systematics

To deduce the effect each shape systematic has on the results, the expected limits were recalculated multiple times with one of the systematic effect removed. These limits are denoted lim._1. In addition, another set of limits are calculated with only one of the systematic uncertainties included. These limits are denoted lim._2. The values of lim._1 and lim._2 for each systematic uncertainty are detailed in Table 7.7.

Figure 7.18 shows the post-fit values and errors for the lognormal and gaussian uncertainties. The deviation of the fitted values in terms of both the error on the fitted value and the prior uncertainty on the parameter from the nominal values are shown for fits in the μ + jets channel, e + jets channel, and both channels combined. It can be seen that the post-fit uncertainty on the scale and matching systematics have decreased with respect to their input values and hence were overestimated.

Sys.	exp. lim.$_1$	exp. lim.$_2$
None	32.0^{+17}_{-17}	-
JER	$33.15^{+13.84}_{-10.35}$	$30.19^{+14.21}_{-9.66}$
JES	$31.62^{+15.05}_{-9.16}$	$23.31^{+12.25}_{-8.05}$
Matching	$32.79^{+14.37}_{-10.21}$	$32.79^{+14.37}_{-10.21}$
PU	$32.42^{+13.91}_{-10.00}$	$28.05^{+13.32}_{-8.93}$
Scale	$32.46^{+14.56}_{-10.35}$	$24.80^{+13.21}_{-6.73}$
bTag	$33.40^{+14.27}_{-10.48}$	$23.32^{+12.71}_{-8.35}$
leptonSF	$33.36^{+14.24}_{-10.39}$	$28.20^{+13.32}_{-8.96}$
misTag	$33.24^{+14.14}_{-10.34}$	$28.62^{+13.46}_{-9.13}$
ttbb	$33.23^{+13.87}_{-10.17}$	$27.97^{+13.26}_{-8.91}$

Table 7.7: Effects of systematic uncertainties.

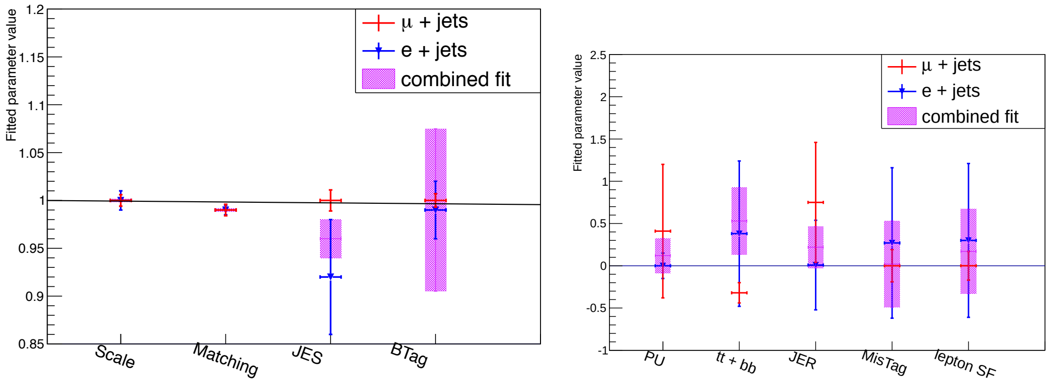


Figure 7.18: The fitted values and approximate errors for the lognormal (left) and gaussian (right) nuisance parameters describing the shape systematics.

7.12.2 Signal Injection Test

In order to validate the limit-setting method, the following closure test was performed. A range of signal strengths were injected into the real data to simulate what the data collected by CMS look like with an enhanced t̄tt̄t cross section and the limit-setting procedure was repeated. Figure 7.19 shows the BDT discriminator distributions with injected signal strengths of 20fb, 40fb and 60fb. With the increasing signal strength, there is an increase for pseudo-data in the bins with higher BDT discriminator values as would be expected for signal events.

As can be seen in Fig. 7.20, the fitted value for the signal strength and observed limit increase with increasing injected singal strength as expected. This shows that the analysis is capable of detecting an enhanced signal if it were present. However, there appears to be some bias in the analysis towards selecting background as opposed to signal as can be seen from the gradient of the line, which is <1.

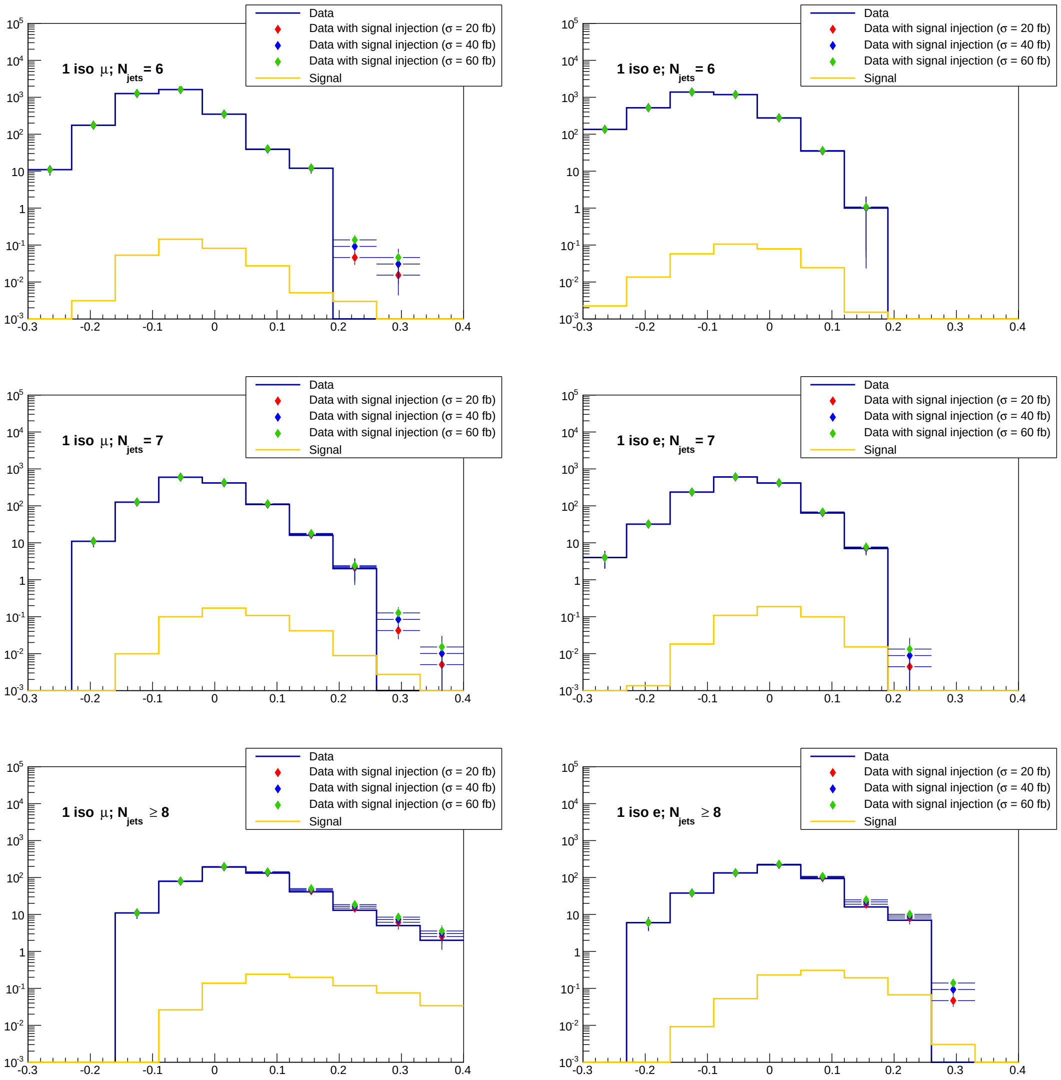


Figure 7.19: The BDT discriminator distributions of data and the mixtures of data and injected signal are compared for μ + jets channel (left) and e + jets channel (right).

7.12.3 Comparisons with the Theta package and the fully-frequentist approach using the Higgs Combine Tool

A cross-check was performed using an alternative framework, the Theta package [122], to set a limit using $\mathrm{CL_S}$ with the asymptotic approximation. As another cross-check, the fully frequentist $\mathrm{CL_S}$ limit setting procedure is performed, using Combine, in place of the asymptotic limit setting used for the final results. Due to the extremely large CPU demands of the fully frequentist technique, the range of values for the parameter of interest, $\frac{\sigma_{tttt}}{\sigma^{SM}_{tttt}}$, and the number of points investigated is reduced. The results, shown in Table 7.8, show that the observed differences

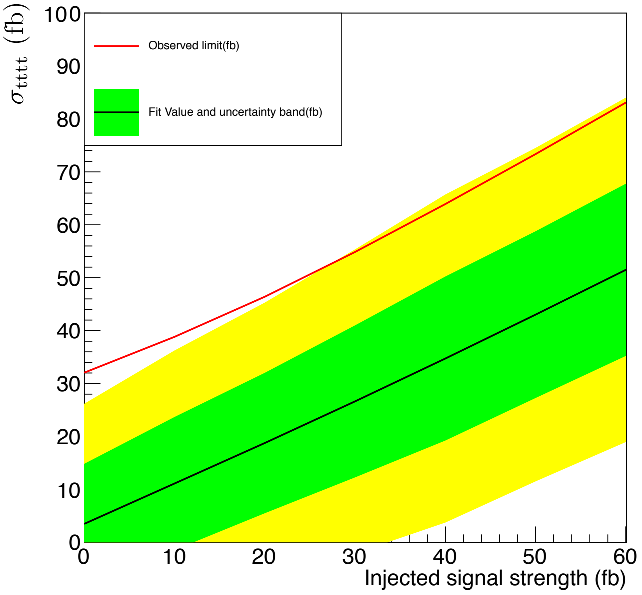


Figure 7.20: Fitted values and observed for various injected signal strength in units of fb.

between the limit setting packages is below the precision of the individual expected limits. Thus, and differences can be ignored.

are consistent with the results obtained using the asymptotic approximation in Combine.

Setup	Exp. $\frac{\sigma_{signal}}{\sigma^{tttt}_{SM}}$	Obs. $\frac{\sigma_{signal}}{\sigma^{tttt}_{SM}}$	Exp.(fb)	Obs.(fb)
Combine (asymptotic.)	24.6 ± 13	24.7	32 ± 17	32
Theta (asymptotic)	27.4^{+12}_{-8}	24.66	35.6^{+16}_{-10}	32
Combine (ful. freq.)	-	23.78	-	31

Table 7.8: Comparison of CL_S limits on $\sigma_{t\bar{t}t\bar{t}}$ from Higgs Combine Tool (asymptotic + fully frequentist) and Theta packages.

7.12.4 Fitted nuisance parameters and uncertainties

In Table 7.9 the fitted values, Δx, and the post-fit uncertainties, σ_{out}, in units of the input uncertainty, σ_{in}, are given for the background only (b-only) hypothesis and the signal + background (s+b) hypothesis. If the associated uncertainty is well modelled, the value of $\frac{\Delta x}{\sigma_{\mathrm{in}}}$ should be close to zero and the value of $\frac{\sigma_{\mathrm{out}}}{\sigma_{\mathrm{in}}}$ should be close to one. The largest deviations in $\frac{\sigma_{\mathrm{out}}}{\sigma_{\mathrm{in}}}$ arise from the scale and matching uncertainties, which is expected as these uncertainties were assigned ad hoc variations. Table 7.9 shows that the results do not vary significantly between the b-only and s+b hypotheses which is consistent with no signal being present in the data.

	b-only fit		$s+b$ fit	
name	$\Delta x/\sigma_{\mathrm{in}}$	$\sigma_{\mathrm{out}}/\sigma_{\mathrm{in}}$	$\Delta x/\sigma_{\mathrm{in}}$	$\sigma_{\mathrm{out}}/\sigma_{\mathrm{in}}$
btag	+0.16	1.15	+0.16	1.24
ew_norm	-0.01	0.99	-0.00	0.99
JER	+0.64	0.48	+0.63	0.49
JES	+0.30	0.37	+0.29	0.37
lepton sf	-0.02	1.19	-0.02	1.29
lumi	-0.92	0.81	-0.93	0.82
matching	-0.30	0.17	-0.30	0.17
mistag	+0.22	1.04	+0.27	1.05
pu	+0.27	0.82	+0.28	0.86
scale	+0.12	0.25	+0.11	0.27
tt_norm	-1.34	0.61	-1.35	0.62
ttbb	+0.18	0.84	+0.20	0.87
ttother_norm	-0.06	0.99	-0.06	0.99

Table 7.9: Fitted values and post-fit uncertainties in units of input uncertainties (σ_{in}) for nuisance parameters in b-only and $s+b$ fits.

7.12.5 Alternative parameterisations

As the combined ME and PS scale systematic uncertainty may vary in shape across the $\mathrm{N_{jets}}$ categories, another cross-check on the analysis was to allow the scale systematic to vary independently in each $\mathrm{N_{jets}}$ category, as in [123, 124]. In this approach, both channels have three nuisance parameters for the scale systematic compared with one nuisance parameter used for all $\mathrm{N_{jets}}$ categories. The results, shown in Table 7.10, show no significant change in the limit compared with the original approach. The associated nuisance parameters and their uncertainties are show in Table 7.11.

$N_{scale\ params.}$	Exp. (fb)	Obs. (fb)
1	32 ± 17	32
3	31.8 ± 12	32.5

Table 7.10: $\mathrm{CL_S}$ limits on $\sigma_{t\bar{t}t\bar{t}}$ for the nominal result (top row) and with scale uncertainty fit independently in each $\mathrm{N_{jets}}$ bin (bottom row).

	b-only fit		$s+b$ fit	
name	$\Delta x/\sigma_{\text{in}}$	$\sigma_{\text{out}}/\sigma_{\text{in}}$	$\Delta x/\sigma_{\text{in}}$	$\sigma_{\text{out}}/\sigma_{\text{in}}$
btag	+0.04	1.35	-0.04	1.51
jer	+0.43	0.53	+0.42	0.54
jes	+0.30	0.37	+0.28	0.39
leptonsf	-0.01	1.36	-0.02	1.56
lumi	-0.92	0.81	-0.94	0.82
matching	-0.27	0.19	-0.26	0.19
mistag	+0.18	1.17	+0.27	1.13
pu	+0.29	0.87	+0.28	0.95
scale6j	+0.31	0.23	+0.31	0.23
scale7j	+0.31	0.25	+0.31	0.25
scale8j	-0.71	0.34	-0.70	0.36
tt_norm	-1.33	0.61	-1.36	0.62
ttbb	+0.40	0.71	+0.44	0.75

Table 7.11: Fitted values and post-fit uncertainties in units of input uncertainties (σ_{in}) for all nuisance parameters in b-only and s+b fits, for three N_scale parameters.

A separate check was made where the normalisation uncertainty on the $t\bar{t}$ component was allowed to vary independently in each N_{jets} category. In each channel, three nuisance parameters are associated to the normalisation uncertainty on the $t\bar{t}$ for each N_{jets} category compared to the one nuisance parameter across all N_{jets} categories in the original approach. The results, shown in Table 7.12, also show no significant change to the original limit.

	Exp. (fb)	Obs. (fb)
nominal	32 ± 17	32
varying $t\bar{t}$ in N_{jets} bins	36 ±17	41

Table 7.12: CL_S limits on $\sigma_{t\bar{t}t\bar{t}}$ for the nominal result (top row) and with $t\bar{t}$ normalisation uncertainty fit independently in each N_{jets} bin (bottom row).

7.12.6 Using N_{jets} ≥8jet category only

The greatest separation of signal and background occurs in the $N_{\text{jets}} \geq 8$ jet category as $t\bar{t}t\bar{t}$ events typically have higher jet multiplicities than $t\bar{t}$. In order to ascertain whether the analysis could benefit from a higher jet multiplicity cut, the limit setting procedure is repeated using only the $N_{\text{jets}} \geq 8$ jet category. The results are shown in Table 7.13, where it can be seen that the limit becomes weaker by removing the $N_{\text{jets}} = 6$ and $N_{\text{jets}} = 7$ categories. This implies that including the $N_{\text{jets}} = 6$ category aids in constraining the $t\bar{t}$ in the fit. The $N_{\text{jets}} = 7$ category also has some discrimination between signal and background and simultaneously fitting the $\mathrm{N_{jets}} = 6$, $\mathrm{N_{jets}} = 7$ and $\mathrm{N_{jets}} \geq 8$ categories provides a better limit.

	Exp. (fb)	Obs. (fb)
nominal result	32 ± 17	32
8 jet bin only	45 ± 17	47

Table 7.13: $\mathrm{CL_S}$ limits on $\sigma_{t\bar{t}t\bar{t}}$ for nominal scenario and 8 jet bin only.

7.12.7 Investigation of binning in BDT distributions

In order to ascertain what the optimum binning should be within each BDT distribution, a variety a values are shown in Fig. 7.21. Both the inclusive result before splitting the templates into $\mathrm{N_{jets}}$ categories and exclusive $\mathrm{N_{jets}}$ categories results are shown, where the significant improvement in the limit by the latter method can be seen. The limit does not appear to improve significantly when increasing the number of bins in the BDT distributions. In order to ensure that most bins were well populated, the number of bins was chosen to be ten in the exclusive analysis.

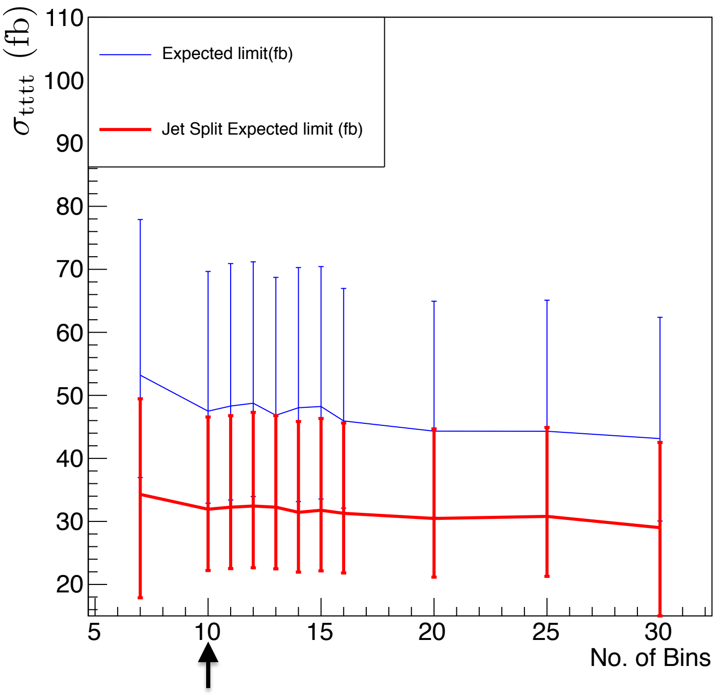


Figure 7.21: Expected limit (fb) with various numbers of fixed width bins in both the inclusive analysis and when split into $\mathrm{N_{jets}}$ categories. The arrow indicates the chosen binning.

7.13 Candidate four-top-quark event

Although it is not possible to know which specific events are produced from the decays of four top quarks, candidate events can be found which are more likely to have come from $t\bar{t}t\bar{t}$ production. Figure 7.22 shows an event from data which has 11 jets, three of which are b-tagged. It has a relatively high event-level BDT score and also a high score from the hadronic top quark BDT, where the *Multitopness* variable on the figure equates to $\mathrm{BDT}_{tri-jet2}$. Both the hadronic transverse energy from the reduced event, HTX, and from b-quarks, HTb, is large, as expected for $t\bar{t}t\bar{t}$ events.

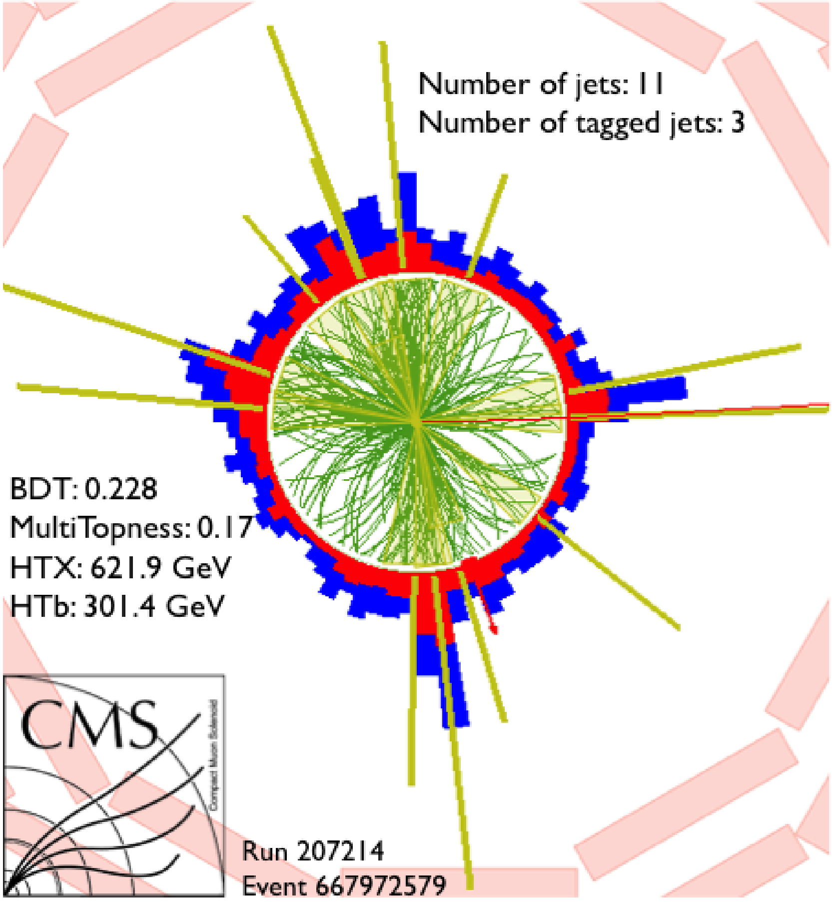


Figure 7.22: A candidate four-top-quark event from the Run 1 2012 CMS dataset.

7.14 Summary and conclusion

The full 2012 data set of proton-proton collisions at $\sqrt{s} = 8$ TeV with 19.6 fb^{-1} of data was used to set limits on the production of $t\bar{t}t\bar{t}$ in the single muon and single electron final states in the absence of a signal. After a baseline selection the dominant background is $\mathrm{t\bar{t}}$ production. The simulation is corrected with scale factors. Two BDTs are used in the analysis, one for reconstructing hadronically decaying top quarks and one event-level BDT to separate the signal and background using discriminators derived from the hadronic top quark reconstruction BDT. Analysis of the systematic uncertainties showed the factorisation and renormalisation scale on the $\mathrm{t\bar{t}}$ background to be the largest uncertainty. A template fit was performed on the event-level BDT discriminator output which was separated into $\mathrm{N_{jets}}$ categories. A simultaneous fit of the templates for the single electron and single muon histograms, each split into $\mathrm{N_{jets}}$ categories, was performed and a $\mathrm{CL_S}$ limit at 95% C.L. on a cross section ratio of $\sigma_{\mathrm{t\bar{t}t\bar{t}}}$ / $\sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}$ of 25 observed and 24.6 ± 13 expected was found. This corresponds to 32 fb observed with 32.0 ± 17 fb expected.

7.15 Discussion of other searches for $\mathrm{t\bar{t}t\bar{t}}$ production studies at $\sqrt{s} = 8$ TeV

The ATLAS collaboration have also published limits on the $\mathrm{t\bar{t}t\bar{t}}$ cross section in the single lepton + jets channel at $\sqrt{s} = 8$ TeV [57] using 20.3 fb^{-1} of data. This search takes a more simple approach by making an initial selection, splitting events into categories and then making a fit to the H_{T} distributions. The selection is similar to the selection in this chapter: 1 isolated lepton (e or μ), ≥ 5 jets, ≥ 2 b-tagged jets, $E_{\mathrm{T}}^{\mathrm{miss}}$ >20 GeV and ($E_{\mathrm{T}}^{\mathrm{miss}} + M_T^W$) > 60 GeV (where M_T^W is the transverse mass of the lepton). The electron and muon channels are considered as one analysis. Events are categorised according to 5 or ≥ 6 jets and 2, 3 or ≥ 4 b-tagged jets. The categories of ≥ 6 jets and 3 or ≥ 4 b-tags are further split into two categories each with either $M_{bb}^{min\Delta R} < 100$ or $M_{bb}^{min\Delta R} > 100$, where $M_{bb}^{min\Delta R}$ represents the invariant mass of the two least separated b-tagged jets. A fit is made to the H_{T} distributions in these categories which results in a 95% CL upper limit of 23 fb (32 fb) observed (expected). Therefore, this result is consistent with the studies of four-top-quark production in the single lepton channel by CMS in this chapter. Where the CMS analysis has benefitted from fitting to the BDT discriminator value distribution with respect to using a simpler variable such as H_{T}, the ATLAS analysis has gained sensitivity by further categorisation into b-tagged jet categories.

There are also same-sign dilepton analyses performed by both CMS [125] and ATLAS [126] at $\sqrt{s} = 8$ TeV, which use 19.5 fb^{-1} and 20.3 fb^{-1} of data, respectively. These analyses benefit from relatively small SM backgrounds compared to the single lepton analyses, however they suffer more from uncertainties due to charge misidentification. Both analyses select two same sign leptons, ≥ 2 jets, ≥ 2 b-tagged jets and have requirements on H_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$. They both use a number of search regions categorised according to the number of jets, b-tagged jets, H_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$. ATLAS observe a 2.5σ excess on the $\mathrm{t\bar{t}t\bar{t}}$ cross section, however this is not significant and therefore an upper limit is set of 70 fb (27 fb) observed (expected). In comparison, CMS set a limit of 45 fb (36^{+16}_{-9} fb) observed (expected). These results are consistent with each other. The excess observed by ATLAS is likely to be a statistical fluctuation.

It can be seen that the CMS single lepton + jets channel discussed in this chapter is very competitive with the other results produced by both CMS and ATLAS.

8 | Phenomenological study of Run 1 four-top-quark production cross section limits

The 95% CL upper limit placed on the production of four top quarks at $\sqrt{s} = 8$ TeV can be used to place constraints on new physics models which predict an enhancement in the $t\bar{t}t\bar{t}$ production cross section. One particular model, which is described further in Chapter 2, is a simplified model of NMSSM in which a new particle, the sgluon, arises. The results of the search in the single lepton channel in Chapter 7 are used to place constraints on the mass and coupling of the sgluon alongside another complementary CMS search for new physics in the same-sign dilepton channel which also places a limit on SM four-quark-production using the same dataset at $\sqrt{s} = 8$ TeV [125]. The author personally made contributions to reinterpreting that latter search and hence it will be primarily discussed in this thesis. Full details of the combined analysis and particularly the reinterpretation of the single lepton analysis can be found in [127].

The chapter is organised as follows: in Section 8.1 the simplified model which is considered for study is discussed while in Section 8.2 the methodology of reinterpreting the CMS same-sign dilepton search is discussed. The simulation of sgluon events is discussed in Section 8.3 and the analysis of the simulation in the reinterpretation is discussed in Section 8.4. The results are given in Section 8.5 and conclusions in Section 8.6.

8.1 A simplified model for describing top-philic sgluons

In this section the simplified model of sgluon production and decay, which is a minimal extension of the SM, is described. The SM is supplemented by the Lagrangian in Eq. 8.1 by adding a single real colour-octet field for the sgluon, S^a. The gauge interactions of a sgluon pair to gluons is represented by the QCD covariant derivative, $D_\mu S^a = \partial_\mu S^a + g_s {f_{bc}}^a G^b_\mu S^c$.

$$\mathcal{L} = \frac{1}{2} D_\mu S^a D^\mu S_a - \frac{1}{2} m_S^2 S^a S_a \tag{8.1}$$

The strong coupling constant is represented by g_s, ${f_{bc}}^a$ are the structure constants of the SU(3) gauge group and G^a_μ denotes the gluon field. The effective Lagrangian in Eq. 8.2 represents the sgluon coupling to the SM degrees of freedom, where the first term represents the dimension-four coupling of the sgluon to a top-anti-top pair and the second term represents the sgluon's interaction with gluons.

$$\mathcal{L}_{\text{eff}} = \bar{t}\, T_a (a_t^L P_L + a_t^R P_R)\, t\, S^a + \frac{a_g}{\Lambda} {d_a}^{bc} S^a F_{\mu\nu,b} F^{\mu\nu,c} + \text{h.c.} \tag{8.2}$$

The dimensionless coupling of the sgluons to gluons is represented by a_g, which is suppressed by the theory cutoff scale Λ. The coupling of the sgluon to the top quark is denoted by a_t^L and a_t^R for the left-handed and right-handed couplings. In the fundamental representation, ${d_a}^{bc}$ is the symmetric structure constant and P_L (P_R) is the left (right) parity projection operator. The generators of SU(3) are T_a. Although sgluons can, in theory, couple via flavour-changing-neutral currents to different quark flavours, we only consider the interactions with top quarks which conserve flavour. This is motivated by R-symmetric supersymmetry with minimal flavour violation where only interactions of sgluons with top quarks and gluons are non-negligible.

In this model sgluons are top-philic, ie. they are only allowed to decay to $\mathrm{t\bar{t}}$ or to two gluons. This is consistent with the class II scenarios in Ref. [51], which gives the following benchmark values for the couplings and theory cut-off scale:

$$a_t^L = a_t^R = a_t = 1.5 \times 10^{-3}\ , \quad a_g = 1.5 \times 10^{-3}, \quad \text{and} \qquad \Lambda = 1\ \text{TeV}$$

The a_t coupling is varied in the range $[0.5, 5] \times 10^{-3}$ and a_g is varied in the range $[1.35, 1.65] \times 10^{-3}$. The mass of sgluons studied is in the range $[350, 900]$ GeV. These values are consistent with ultraviolet-complete theories of supersymmetry which have coloured superpartners with mass in the range $[1, 2]$ TeV.

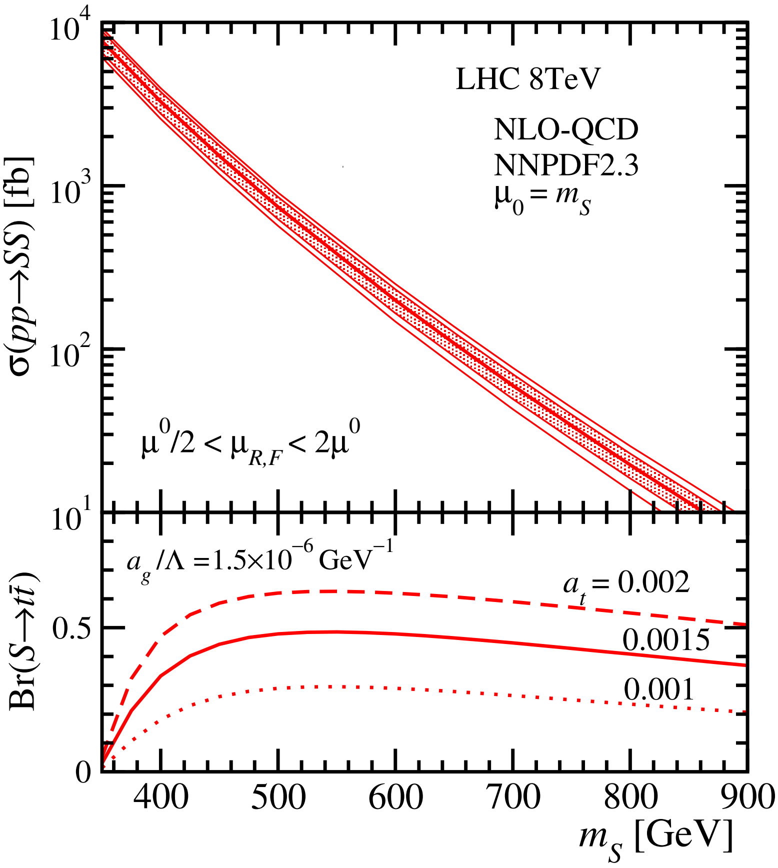


Figure 8.1: The top panel shows the m_s dependence of the production of sgluon pairs from proton-proton collision at NLO with variations to the renormalisaton and factorisation scale by a factor of 1/2 and 2. The bottom panel shows the m_s dependence of the branching fraction of the sgluon to a top-anti-top pair, $\mathrm{Br}(S \to \mathrm{t\bar{t}})$, for different values of the coupling a_t and a fixed value of a_g [127].

This simplified model is consistent with an enhancement in the production of $\mathrm{t\bar{t}t\bar{t}}$ final states in proton-proton collisions, depending on the mass and couplings of the sgluon.

In Fig. 8.1 the upper panel shows the dependence of the sgluon pair production cross section on the mass of the sgluon itself calculated at NLO accuracy in QCD [91, 128]. If sgluons were to have higher masses they would be less likely to be produced as it requires the quarks or gluons from the proton to have a high momentum fraction from the proton. The lower panel in Fig. 8.1 shows the sgluon mass dependence of the branching fraction of a sgluon to a $\mathrm{t\bar{t}}$ pair for a range of a_t coupling values. It can be seen that $\mathrm{Br}(S \to \mathrm{t\bar{t}})$ increases from zero at twice the top quark mass.

8.2 Reinterpretation of same-sign dilepton channel

The analysis by CMS of the same-sign dilepton channel at $\sqrt{s} = 8$ TeV [125] utilises the entire 2012 dataset of proton-proton collision corresponding to an integrated luminosity of 19.6 fb^{-1}. The analysis is a collection of counting experiments that contain at least two same sign charged leptons and a number of jets. There are 28 signal regions defined in Ref. [125] (see Appendix B.1) which are categorised according to $\mathrm{N_{jets}}$, $\mathrm{N_{btags}}$, H_{T}, $E_{\mathrm{T}}^{\mathrm{miss}}$ and require the leptons to satisfy $p_{\mathrm{T}} > 20$ GeV and$|\eta| < 2.4$. Parameterisations are provided in this CMS paper which give the efficiency for reconstructing an object in simulation given generator-level information. These include:

- The reconstruction of jets with respect to generator-level jet p_{T} (ε_{jet})
- The b-tagging of b-quark jets by the CSV algorithm with respect to generator-level jet p_{T} (ε_b)
- The reconstruction of H_{T} with respect to generator-level H_{T} (ε_{H_T})
- The reconstruction of muons and electrons with respect to generator-level lepton p_{T} (ε_μ, ε_e)
- The reconstruction of $E_{\mathrm{T}}^{\mathrm{miss}}$ with respect to generator-level $E_{\mathrm{T}}^{\mathrm{miss}}$ ($\varepsilon_{E_{\mathrm{T}}^{\mathrm{miss}}}$)

as shown in Fig. 8.2.

Signal region 28 (SR28) was chosen to be reinterpreted in the context of the sgluon model as it closely corresponds to the $\mathrm{t\bar{t}t\bar{t}}$ signature and it can be fully emulated

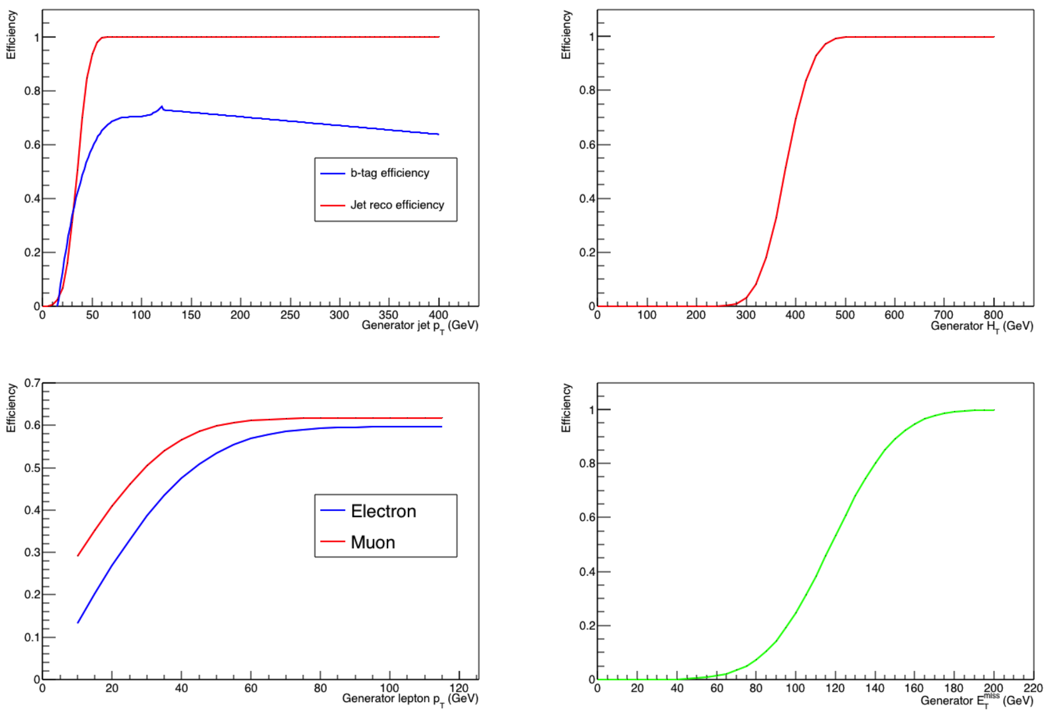


Figure 8.2: Selection efficiencies for the reconstruction of jets and for b-tagging b-quark jets (top-left), for reconstructing H_T (top-right), for reconstructing muons and electrons (bottom-left) and for reconstructing $E_\mathrm{T}^\mathrm{miss}$ (bottom-right). Adapted from [125]. The small bump in the top-left panel is discussed in Appendix B.2.

using the parameterisations of selection efficiency given in the paper. The requirements of SR28 are $\mathrm{N_{jets}} \geq 4$, $\mathrm{N_{btags}} \geq 2$, $H_\mathrm{T} > 400$ GeV, $E_\mathrm{T}^\mathrm{miss} > 120$ GeV and two same-sign leptons with $p_\mathrm{T} > 20$ GeV and $|\eta| < 2.4$. Events which contain any additional leptons which form an opposite-sign-same-flavour (OSSF) pair with one of the signal leptons are vetoed if the invariant mass of the OSSF ($\mathrm{M}_{\ell\ell}$) satisfies either $\mathrm{M}_{\ell\ell} < 12$ GeV or $76 < \mathrm{M}_{\ell\ell} < 106$ GeV. This is to replicate the veto in the CMS search for events which may contain a low mass bound state or a Z boson. It is not possible to combine the results from multiple signal regions without knowledge of the correlation of the uncertainties on the background.

The efficiency for obtaining at least 2 b-tagged jets, $\varepsilon_{\geq\text{2b-tags}}$, can be calculated using Eq. 8.3, where P_i denotes the probability for reconstructing a b-jet, in this case. P_i is the product of the probability to reconstruct a jet$_i$ and to tag it as b-jet, $\varepsilon_{jet} \times \varepsilon_b$ which can be determined from the parameterisations of the curves in Fig. 8.2 (top-left). A uniform efficiency for light quarks to be tagged as b-jet of 1% is included to replicate the same-sign dilepton study.

$$w(\geq 2\ b\ tags) = 1 - w(0) - w(1) \tag{8.3}$$

$$w(0) = \prod_i (1 - P_i) \tag{8.4}$$

$$w(1) = \sum_i \left[P_i \prod_j (1 - P_j) \right] \ ;\ j \neq\ i \tag{8.5}$$

The efficiency for obtaining one same-sign dilepton pair in an event, $\varepsilon_{\geq 1\mathrm{SS2L}}$, can be obtained by using Eq. 8.6 where the weight for obtaining ≥ 2 same-sign positive or negative leptons, $w(\geq 2\ l^+)$ or $w(\geq 2\ l^-)$, can be calculated using 8.3 along with the probabilities from the parameterisation of the muon and electron efficiency curves, (ε_μ & ε_e), in Fig. 8.2 (bottom-left).

$$\varepsilon_{\geq 1\mathrm{SS2L}} = 1 - \left[1 - w(\geq 2\ l^+)\right] \left[1 - w(\geq 2\ l^-)\right] \tag{8.6}$$

From the derived efficiencies for obtaining a specific H_T and $E_\mathrm{T}^\mathrm{miss}$ per event as well as ≥ 2 b tags and ≥ 1 SS dilepton pair, an overall event weight can be calculated using equation 8.7. These parameterisations were designed to be multiplicative.

$$w_\mathrm{event} = \varepsilon_{H_T} \times \varepsilon_{E_\mathrm{T}^\mathrm{miss}} \times \varepsilon_{\geq 2\text{b-tags}} \times \varepsilon_{\geq 1\mathrm{SS2L}} \tag{8.7}$$

8.3 Simulation of sgluon events

The MADGRAPH5_aMC@NLO [98] software package was used to simulate sgluon pair production with exclusive decays to top quarks at $\sqrt{s} = 8$ TeV. PYTHIA6 [95] was used to perform the hadronisation. The jet clustering was performed using the anti-k_{t} clustering algorithm in the FastJet package with a distance parameter of R $= 0.5$ and $p_{\mathrm{T}} > 5$ GeV, which is consistent with the same-sign analysis.

A sample of standard model four top quark events was also simulated using MadGraph5_aMC@NLO and PYTHIA6. A signal efficiency for these events passing the SR28 selection was 0.6% which can be compared to the 0.49% signal acceptance listed in the same-sign analysis paper. These two numbers agree within the 30% uncertainty quoted for the analysis.

Sgluon events were simulated for a number of sgluon mass points, m_S, between $350 \geq m_S \geq 1000$ GeV and in the range of coupling to the top quark, α_t, of $0.5 \times 10^{-3} \geq \alpha_t \geq 5 \times 10^{-3}$.

8.4 Analysis

The MadAnalysis framework [129] is used to perform the analysis, including event selection and the application of the efficiency model using Eq. 8.7. As the efficiency functions are provided in Ref. [125] which map GEN-level quantities, such as H_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$, to RECO-level, it is not necessary to perform a detector simulation. Using Eq. 8.7 one can obtain the fraction of events which pass the selection and hence the signal efficiency for each sgluon (m_S, α_t) point. The number of sgluon events, n^S_{events}, which would be found in the same-sign analysis is obtained using Eq. 8.8, where $\varepsilon\left(m_S, \alpha_t\right)$ is the signal efficiency, $\sigma\left(m_S, \alpha_t\right)$ is the cross section and $K_{NLO}\left(m_S, \alpha_t\right)$ is a factor for the ratio of the NLO to LO cross section for that mass and coupling point. The luminosity, $\mathcal{L}$, is taken from the analysis to be 19.6 fb^{-1}. Rather than running the analysis in MadAnalysis for every coupling point, α_t, a nominal coupling can be provided and the program will produce weights, $F^{\alpha_t}\left(m_S, \alpha_t\right)$, for each alternative coupling, where $F^{\alpha_t}\left(m_S, \alpha_t\right) = 1$ for the nominal case. This is possible as the cross section will scale with α_t^4 and it reduces computing time.

$$n^S_{events}\left(m_S, \alpha_t\right) = \varepsilon\left(m_S, \alpha_t\right) \times \sigma\left(m_S, \alpha_t\right) \times K_{NLO}\left(m_S, \alpha_t\right) \times \mathcal{L} \times F^{\alpha_t}\left(m_S, \alpha_t\right) \quad (8.8)$$

This gives the number of sgluon events for each mass and coupling point in the chosen range, which then needs to be compared to the CMS analysis. CMS obtained 2 (2.2) events observed (expected) in SR28. This is converted to an 95% C.L. upper limit of 4.68 (4.89) events observed (expected) by using the asymptotic $\mathrm{CL_S}$ method implemented in the RooStat package. An uncertainty of 30% on the signal yield was assumed, which is consistent with the CMS analysis result. Hence regions of the (m_S, α_t) plane where n^S_{events} is expected to be larger than 4.68 events can be excluded.

8.5 Results

Figure 8.3 shows the exclusion boundary for the dilepton analysis, where the number of events exceeds the CMS 95% C.L. limit in the exclusion zone. The CMS single lepton analysis from Refs. [125, 127] is also included in the figure. The single lepton analysis uses a simplified Matrix Element Method to quantify how SM-like the sgluon event kinematics are for each mass and coupling point. The solid lines show the exclusion boundary for a fixed value of $a_g/\Lambda = 1.5 \times 10^{-6}$ GeV^{-1} and the dashed lines show the change in the limit due to a variation of a_g by ±10%. In the mass region from 400 GeV to 500 GeV sgluon models couplings down to 6×10^{-4}, where the cross section of $\sigma(pp \to SS \to \mathrm{t\bar{t}t\bar{t}})$ is maximal, are excluded. The sensitivity worsens at low mass values due to the decreasing branching ratio, $\mathcal{B}(S \to \mathrm{t\bar{t}})$, and at higher mass values due to the decreasing cross section $\sigma(pp \to SS)$. The dilepton analysis is more constraining as it does not depend on event kinematics. It constrains up to $m_S = 750$ GeV at $a_t = 2 \times 10^{-3}$ compared to around $m_S = 650$ GeV for the single lepton channel.

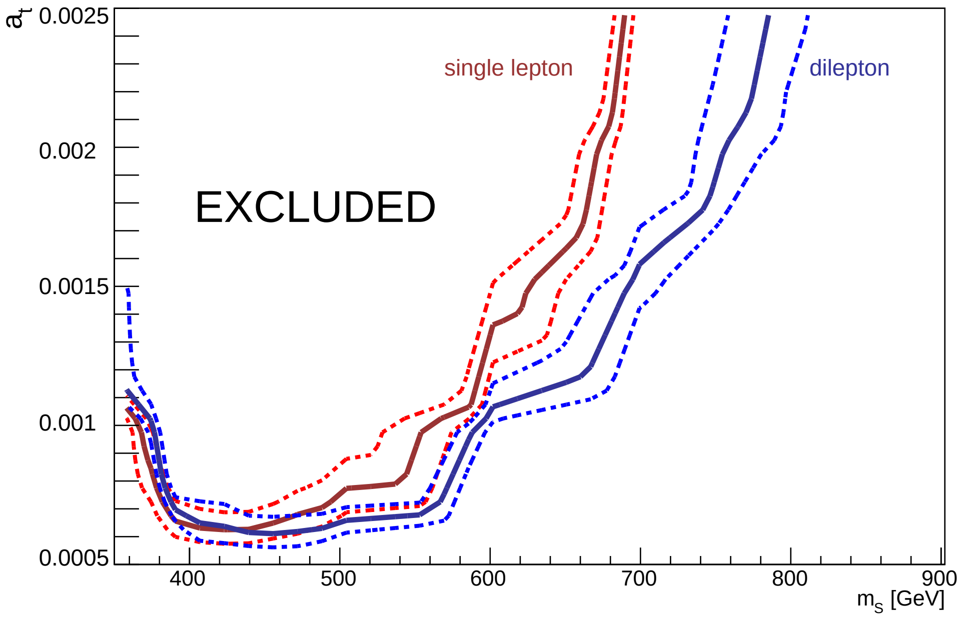


Figure 8.3: Solid lines show the exclusion boundary for a sgluon mass, m_S, and coupling to the top quark, α_t, at $a_g/\Lambda = 1.5 \times 10^{-6}$ GeV^{-1} and dashed lines show the results with a ±10% variation in a_g [127].

8.6 Conclusion

A simplified model approach has been used to describe the dynamics of top-philic sgluons where the parameter space of $(m_S, a_t, a_g/\Lambda)$ has been explored. Two analyses, which set limits on the four-top-quark cross section, are used to place constraints on the phase space regions where sgluons may exist. The dilepton analysis is more powerful at excluding m_S up to 700 GeV for the benchmark scenario of $a_t = 1.5 \times 10^{-3}$, which is a typical value in the NMSSM for superpartners that have masses around 1-2 TeV. For $a_t < 0.75 \times 10^{-3}$ the sgluon mass range of $400 < m_S < 580$ GeV is excluded. Varying a_g by 10% from its benchmark value of $a_g/\Lambda = 1.5 \times 10^{-6}$ GeV^{-1} does not have a significant impact on the result.

9 Search for standard model $\mathrm{t\bar{t}t\bar{t}}$ production in Run 2 at $\sqrt{s} = 13$ TeV

9.1 Introduction

In this chapter, an analysis of the 2015 CMS data set with 2.6 fb^{-1} of data is presented where the SM production of four top quarks ($\mathrm{t\bar{t}t\bar{t}}$) is sought in pp collisions at $\sqrt{s} = 13$ TeV. The SM production of $\mathrm{t\bar{t}t\bar{t}}$ has a cross section at $\sqrt{s} = 13$ TeV of $\sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}} \approx 9.2$ fb at NLO with NNLO corrections [98, 99]. The focus of this chapter is on the single lepton channel where only the μ + jets and e + jets final states are considered. The combination of this analysis with an opposite-sign dilepton analysis and a same-sign dilepton analysis is discussed in Section 9.11.

All sections apart from the training of the hadronic top quark reconstruction in Section 9.7.1 are the author's personal contribution to the analysis.

9.2 Data and Simulation

This analysis uses data from proton-proton collisions at the CMS experiment in 2015 at $\sqrt{s} = 13$ TeV. Data were collected using a trigger based on the presence of at least one muon (electron) candidate with $p_{\mathrm{T}} > 18$ (23) GeV for the muon (electron) channel. These data sets have an integrated luminosity of 2.6 fb^{-1}. The signal SM $\mathrm{t\bar{t}t\bar{t}}$ MC samples and the background MC samples are given in Table 9.1, along with the MC generator used to produce these samples, the order at which they were produced and the number of events produced. MC samples were produced for some systematic uncertainties, which can be found in Table 9.2. In this analysis the ME scale and PS scale are treated as separate uncertainties.

Dataset	Events	Generator	Order
$\mathrm{t\bar{t}t\bar{t}}$	960K	MadGraph aMC@NLO	NLO
$\mathrm{t\bar{t}}$	97M	POWHEG	NLO
W+Jets→ $l\nu$	47M	MadGraph_MLM	LO
Tbar_tW-channel	1M	POWHEG	NLO
T_tW-channel	1M	POWHEG	NLO
DYJetsToLL	9M	MadGraph_MLM	LO
TTZ	400K	MadGraph aMC@NLO	NLO
TTW	250K	MadGraph aMC@NLO FxFx	NLO
TTH_HToBB	4M	POWHEG	NLO

Table 9.1: Dataset name, total number of events, MC generator and order of the simulated samples. PYTHIA 8 was used to hadronise all samples in this table.

Dataset	Events	Generator	Order
TTJets_scaledown	10M	POWHEG	NLO
TTJets_scaleup	10M	POWHEG	NLO
TTJets	5M	MadGraph_MLM	LO
TTJets	5M	MadGraph aMC@NLO FxFx	NLO

Table 9.2: Dataset name, total number of events, MC generator and order of the simulated systematic samples. PYTHIA 8 was used to hadronise all samples in this table.

Comparisons of the alternative $\mathrm{t\bar{t}}$ samples used for the $\mathrm{t\bar{t}}$ generator systematics can be found in Appendix D.1. Studies of the $\mathrm{t\bar{t}W}$, $\mathrm{t\bar{t}Z}$ and $\mathrm{t\bar{t}H}$ backgrounds can be found in Appendix D.2, where merging these backgrounds into the main $\mathrm{t\bar{t}}$ sample is motivated (indicated on the figures as $\mathrm{t\bar{t}}$ +X).

9.3 Baseline Event Selection

The set of criteria applied to the reconstructed objects in events to preferentially select $\mathrm{t\bar{t}t\bar{t}}$ events and suppress background events is detailed below. The exact definition of these objects is given in Chapter 4.

For the muon channel these are:

- Exactly one tight muon
- Exactly zero additional loose muons
- Exactly zero loose electrons

- At least 6 jets with $p_{\mathrm{T}} > 30$ GeV
- At least 2 CSVM tagged b-jets

For the electron channel these are:

- Exactly one tight electron
- Exactly zero additional loose electrons
- Exactly zero loose muons
- At least 6 jets with $p_{\mathrm{T}} > 30$ GeV
- At least 2 CSVM tagged b-jets

9.4 Corrections to the simulation

All corrections are described in Section 6.3. The PU corrections are applied, producing a good agreement in the distribution of the number of vertices as seen in Fig. 9.1. Muon scale factors [77] and electron scale factors [130] are applied. By comparing the efficiencies in data with the efficiencies in simulation, a value was obtained of 1.0001 ± 0.0001 for the electron trigger scale factor which was taken to be 1 in the analysis.

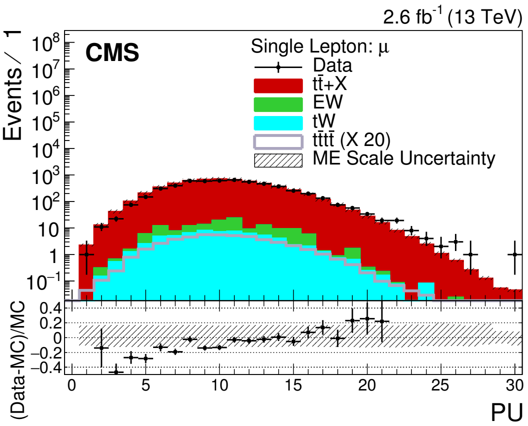


Figure 9.1: The number of primary vertices for data and simulation after application of PU corrections for μ + jets.

The method in Section 6.3.2.2 was used to derive the b tagging scale factors on a per-event basis. Figures 9.2 and 9.3 show the effect of applying the b-tagging scale factor to correct the CSV discriminator distributions. It can be seen that the agreement between data and simulation has been improved.

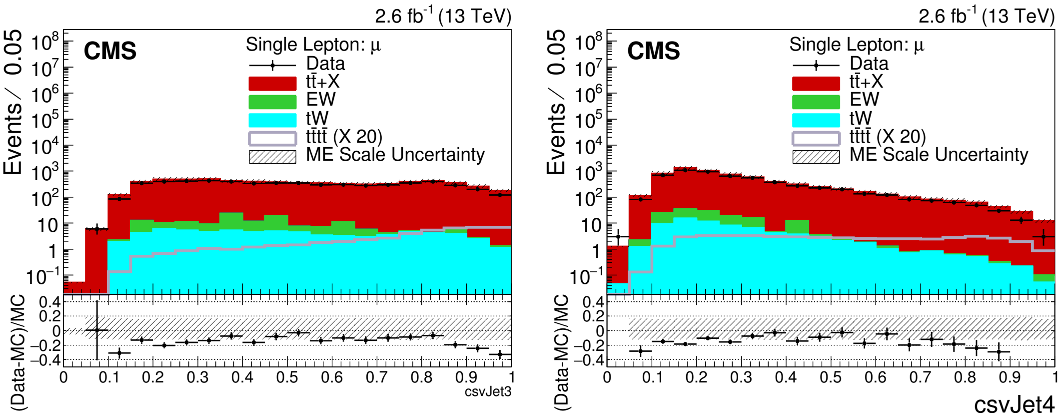


Figure 9.2: The third-highest (left) and fourth-highest (right) ranked CSV jet distributions for data and simulation in the μ + jets channel before b-tagging corrections.

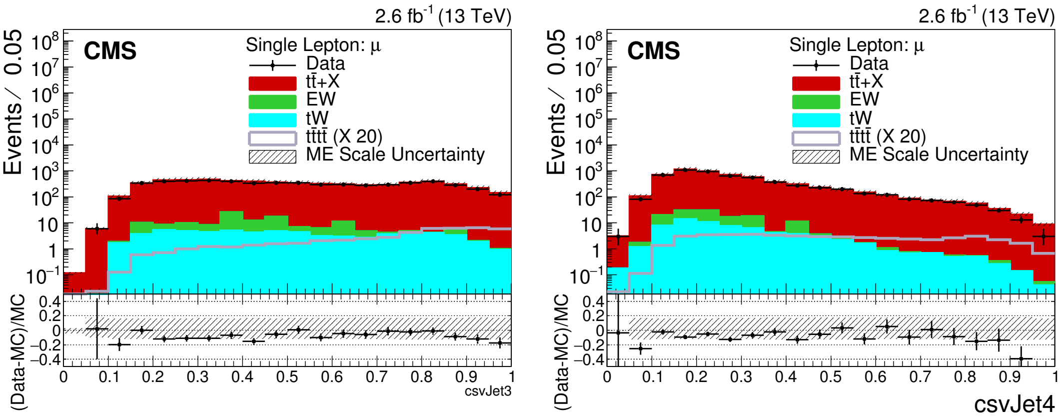


Figure 9.3: The third-highest (left) and fourth-highest (right) ranked CSV jet distributions for data and simulation in the μ + jets channel after b-tagging corrections.

The jet multiplicity modelling from Section 6.3.6, which corrects the MC to correspond to the best tune of α_S in simulation, was applied. It can be seen from Figs. 9.5 and 9.4 that the jet multiplicity modelling is greatly improved by applying this correction.

The scale factors applied for the heavy flavour modelling are described in Section 6.3.3. The distributions for the N^{M}_{tags} are shown with and without the heavy

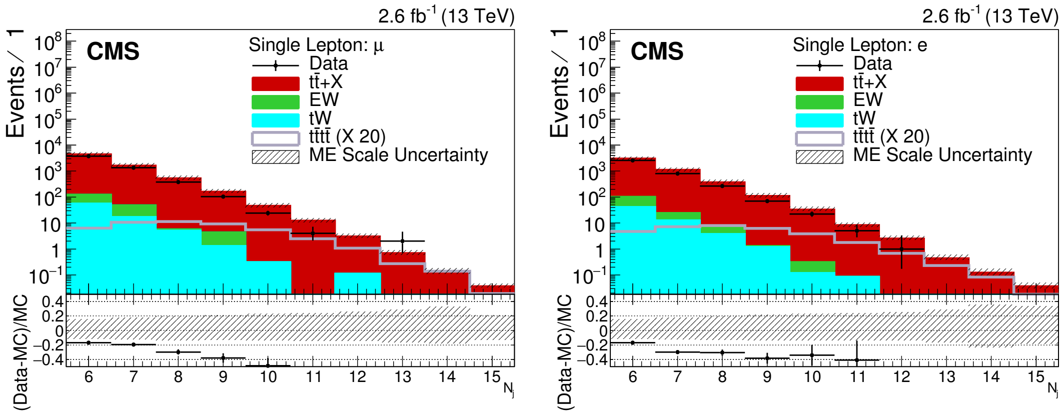


Figure 9.4: The N_{jets} distributions for data and simulation in the μ + jets channel (left) and e + jets channel (left) without jet multiplicity modelling scale factors applied.

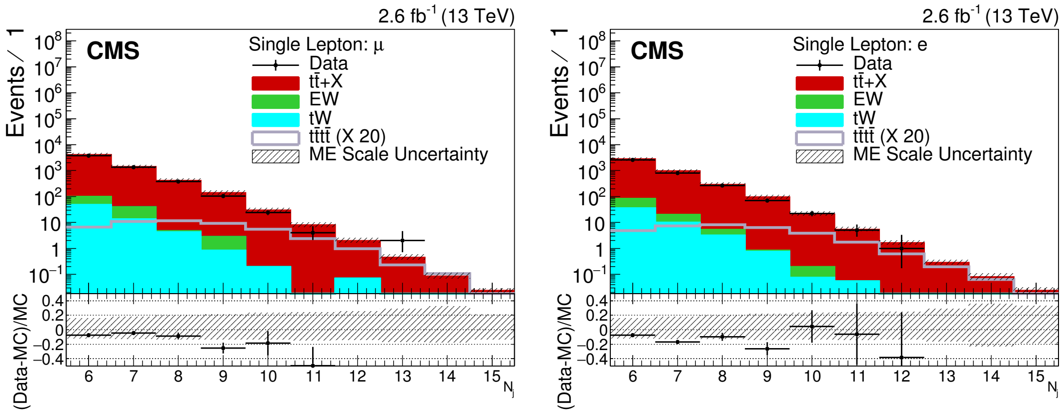


Figure 9.5: The N_{jets} distributions for data and simulation in the μ + jets channel (left) and e + jets channel (left) with jet multiplicity modelling scale factors applied.

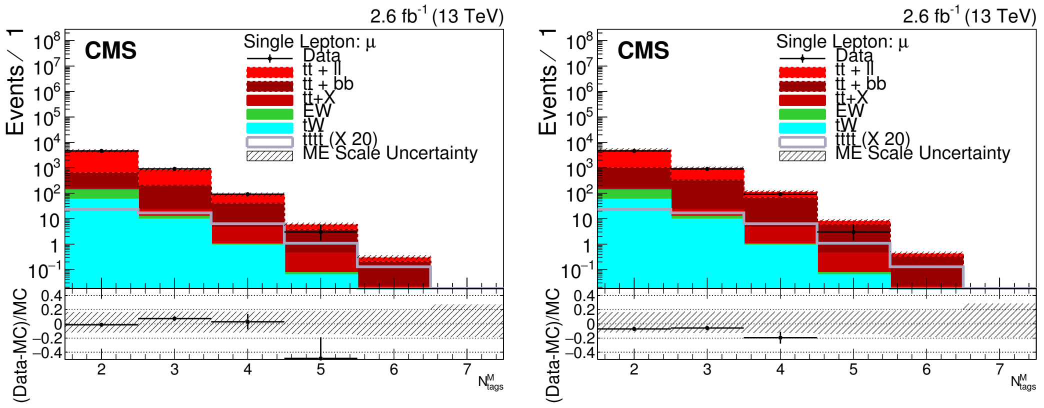


Figure 9.6: $N^{\text{M}}_{\text{tags}}$ are shown for the muon channel with heavy flavour reweighting (right) and without(right).

flavour modelling scale factors applied. It is not obvious that there is a significant improvement in the N^{M}_{tags} distribution after the scale factors have been applied. However, the heavy flavour fraction is allowed to float as a shape nuisance parameter in the template fit.

9.5 Effect of selection requirements

The event counts, which are weighted by each correction factor in simulation, are given after each selection requirement for μ + jets (e + jets) in Table 9.3 (Table 9.4). Small discrepancies between the initial simulated events in each table are due to different lepton scale factors being applied in each case. After the baseline selection has been applied the $t\bar{t}$ component represents ≈ 97% of the number of events combined across all background samples. The ttbb and ttll/ttcc components of the main $t\bar{t}$ sample are also given in the tables. The processes $t\bar{t}Z$ and $t\bar{t}W$ have been merged into $t\bar{t}V$ for the tables. [1]

	Data	$t\bar{t}t\bar{t}$	Single top	DY	W	$t\bar{t}H$	$t\bar{t}V$	$t\bar{t}$	ttbb	ttll&ttcc
initial	85 500 000	24.8	90 700	16 000 000	165 000 000	751	1310	2 130 000	104 000	2 030 000
Trigger	55 400 000	6.66	14 600	3 880 000	27 300 000	113	371	327 000	15 000	312 000
Exactly 1 mu	21 100 000	4.57	11 200	1 570 000	18 300 000	82.7	248	244 000	11 200	233 000
lepton Veto	20 200 000	3.16	9720	890 000	18 300 000	72.0	162	212 000	9860	202 000
≥ 1 Jets	4 620 000	3.16	9440	266 000	3 060 000	72	161	210 000	9840	200 000
≥ 6 Jets	13 600	2.84	171	195	1530	24.8	29.9	11 600	1320	10 300
≥ 2 CSVM bs	5130	2.14	62.7	10.4	63.8	20.0	12.6	5170	657	4510

Table 9.3: Number of observed events in data and expected events in simulation after successive selection requirements in the μ + jets channel ($\mathcal{L} = 2.6$ fb).

	Data	$t\bar{t}t\bar{t}$	Single top	DY	W	$t\bar{t}H$	$t\bar{t}V$	$t\bar{t}$	ttbb	ttll&ttcc
initial	125 000 000	24.9	90 700	16 100 000	165 000 000	752	1310	2 130 000	104 000	2 030 000
Trigger	110 000 000	5.72	12 000	3 200 000	20 000 000	93.6	327	266 000	12 300	254 000
Exactly 1 e	13 300 000	3.45	8050	1 410 000	10 100 000	59.1	196	170 000	7810	163 000
lepton Veto	12 500 000	2.24	6810	692 000	10 100 000	50.4	118	145 000	6750	138 000
≥ 1 Jets	3 800 000	2.24	6620	374 000	1 930 000	111	56.3	144 000	6740	137 000
≥ 6 Jets	9930	2.01	129	214	1070	17.5	22.6	8090	923	7160
≥ 2 CSVM bs	3580	1.51	48.8	12.5	42.9	14.2	9.40	3590	462	3130

Table 9.4: Number of observed events in data and expected events in simulation after successive selection requirements in the e + jets channel ($\mathcal{L} = 2.6$ fb).

---

[1] These studies were performed when method 1 from Section 6.3.2.1 was used for b-tagging before the analysis was updated to using method 2 from Section 6.3.2.2.

9.6 Control distributions

Distributions are shown for HTb, $\mathrm{H_T^{rat}}$, $p_{\mathrm{T\ trijet1}}$, $\mathrm{N_{tags}^{L}}$, $\mathrm{N_{tags}^{T}}$, $\mathrm{M_{RE}^{H}}$, $\mathrm{HT_X}$ and lepton isolation. Good agreement is seen between data and simulation in all distributions with almost all data points within the ME scale uncertainty, the dominant systematic uncertainty, represented by the hatched band.

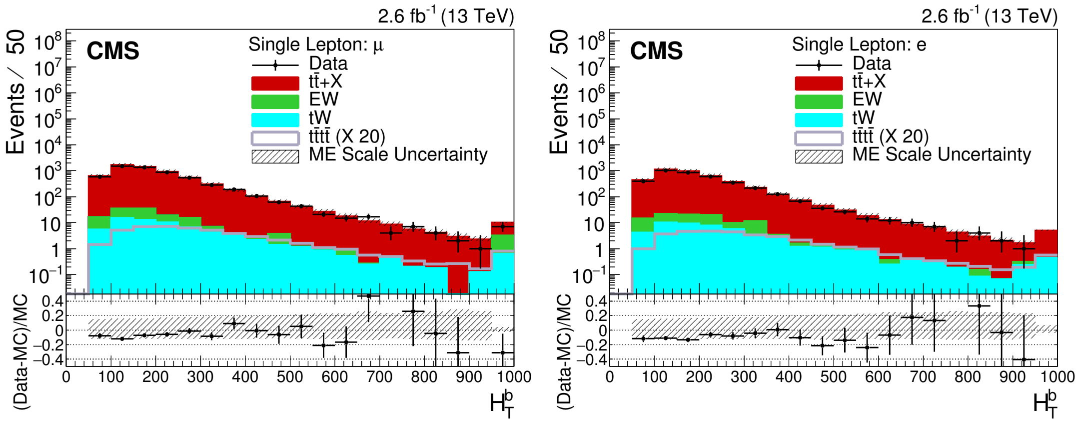


Figure 9.7: The HTb distributions for data and simulation in the μ + jets channel (left) and e + jets channel (right).

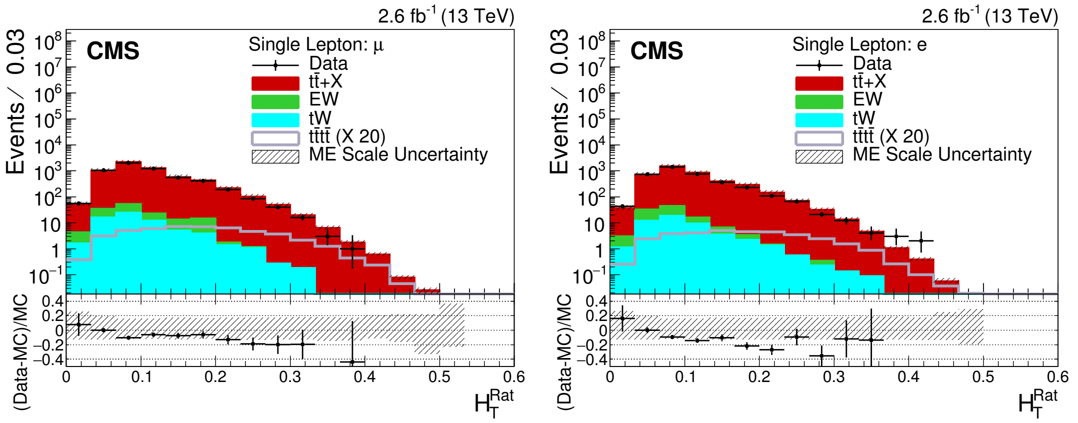


Figure 9.8: The $\mathrm{H_T^{rat}}$ distributions for data and simulation in the μ + jets channel (left) and e + jets channel (right).

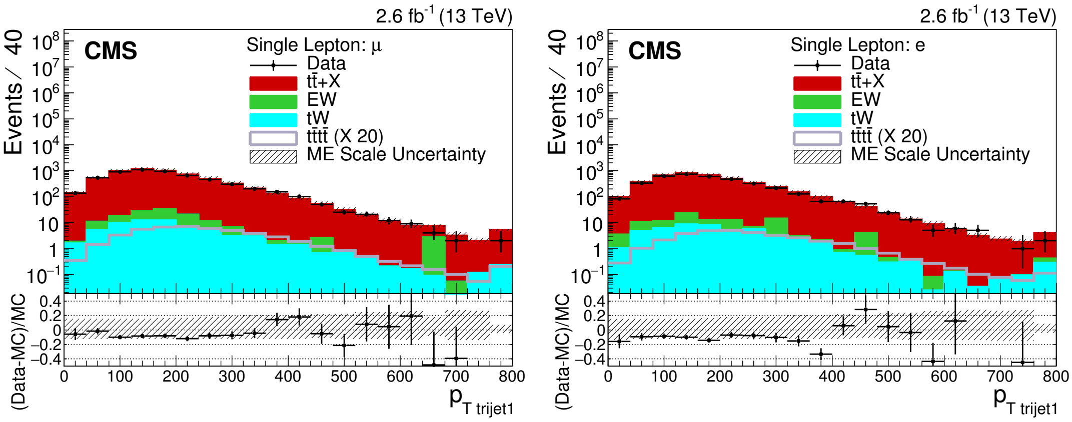


Figure 9.9: The $p_{\mathrm{T\ trijet1}}$ distributions for data and simulation in the μ + jets channel (left) and e + jets channel (right).

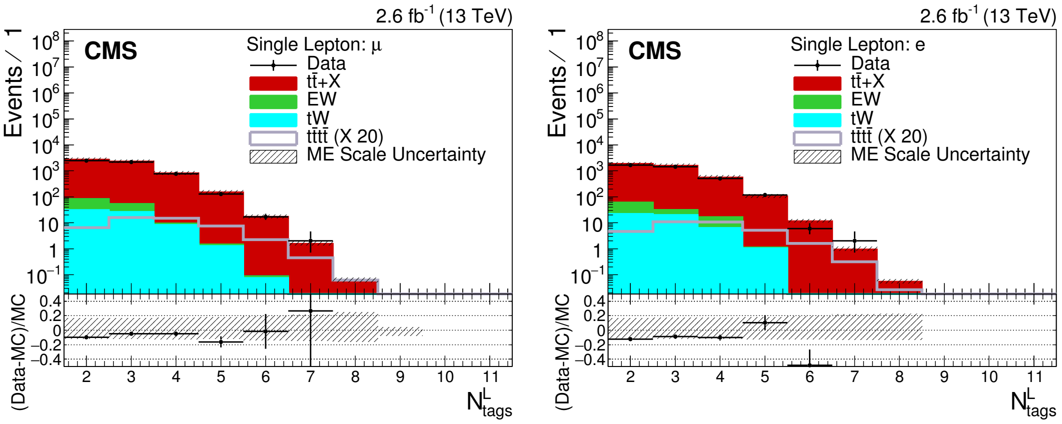


Figure 9.10: The $\mathrm{N}^{\mathrm{L}}_{\mathrm{tags}}$ distributions for data and simulation in the μ + jets channel (left) and e + jets channel (right).

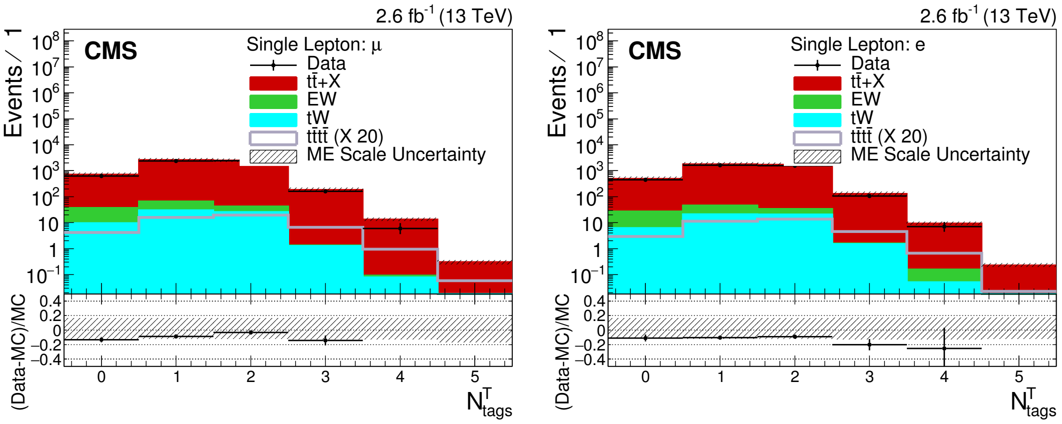


Figure 9.11: The $\mathrm{N}^{\mathrm{T}}_{\mathrm{tags}}$ distributions for data and simulation in the μ + jets channel (left) and e + jets channel (right).

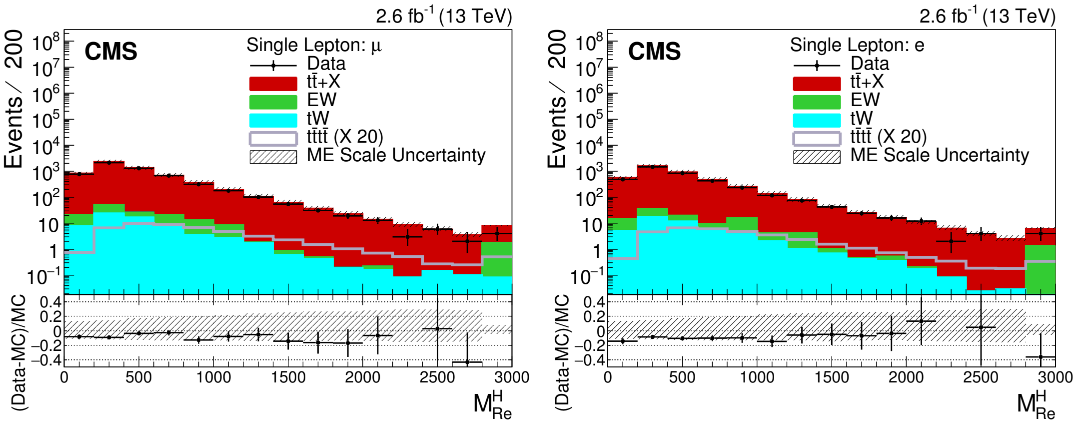


Figure 9.12: The M^{H}_{RE} distributions for data and simulation in the μ + jets channel (left) and e + jets channel (right).

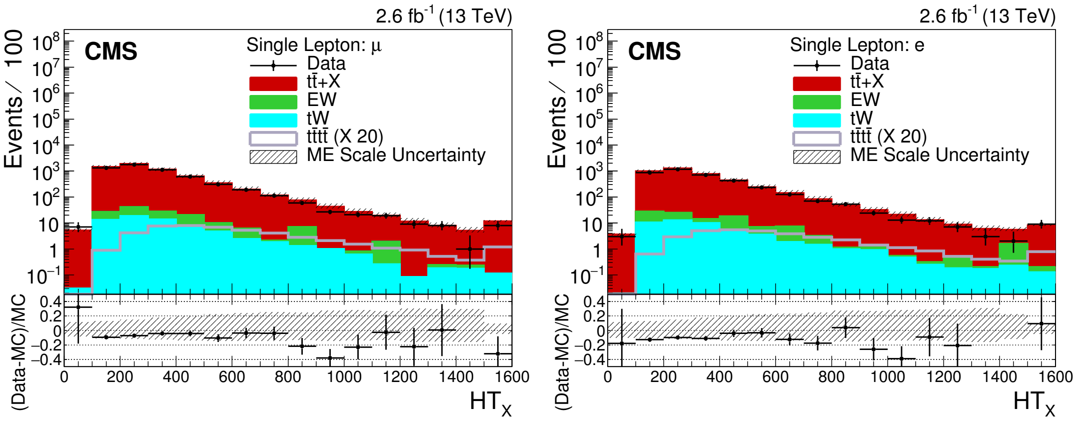


Figure 9.13: The HT_X distributions for data and simulation in the μ + jets channel (left) and e + jets channel (right).

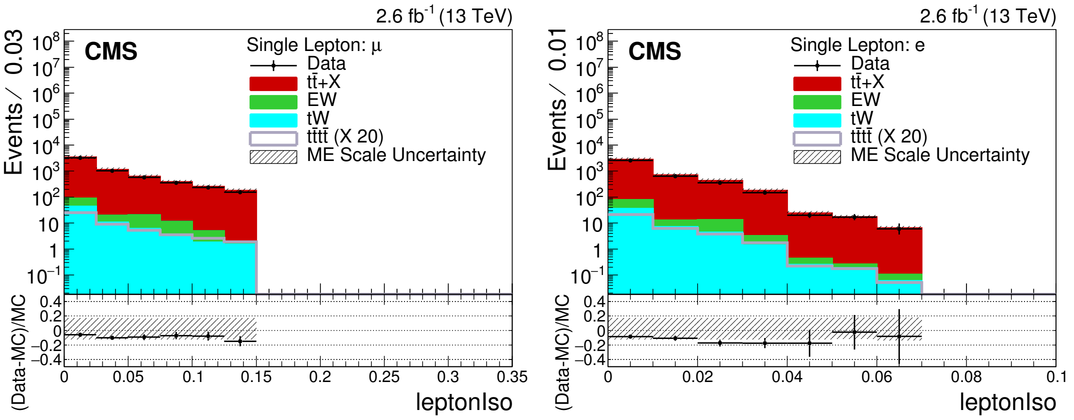


Figure 9.14: The lepton isolation distributions for data and simulation in the μ + jets channel (left) and e + jets channel (left) where the selection requirements on lepton isolation are evident.

9.7 Discriminating between signal and background

It can be seen in Section 9.5 that the $\mathrm{t\bar{t}}$ background is three orders of magnitude larger than the $\mathrm{t\bar{t}t\bar{t}}$ signal in the signal region. The variables used to discriminate between $\mathrm{t\bar{t}}$ and $\mathrm{t\bar{t}t\bar{t}}$ are described below.

9.7.1 Hadronic top quark content

The hadronic top quark reconstruction is fully described in Section 6.6.
As the anti-k_t algorithm cannot resolve jets which have $\Delta R = \sqrt{\eta^2 + \phi^2} < 0.4$, a hadronically decaying top quark can only be deemed *reconstructible* if the minimal ΔR between all three jets is > 0.4. This happens $> 98\%$ of the time in $\mathrm{t\bar{t}}$ and $\mathrm{t\bar{t}t\bar{t}}$ simulation when studying the decay products of the top quarks.
The BDT training was performed on 273,000 $\mathrm{t\bar{t}}$ events. The input variables to the hadronic top quark reconstruction BDT are shown in Fig. 9.15. The separation power for each of these variables and for the output BDT discriminator distribution in Fig. 9.16 is eviden.

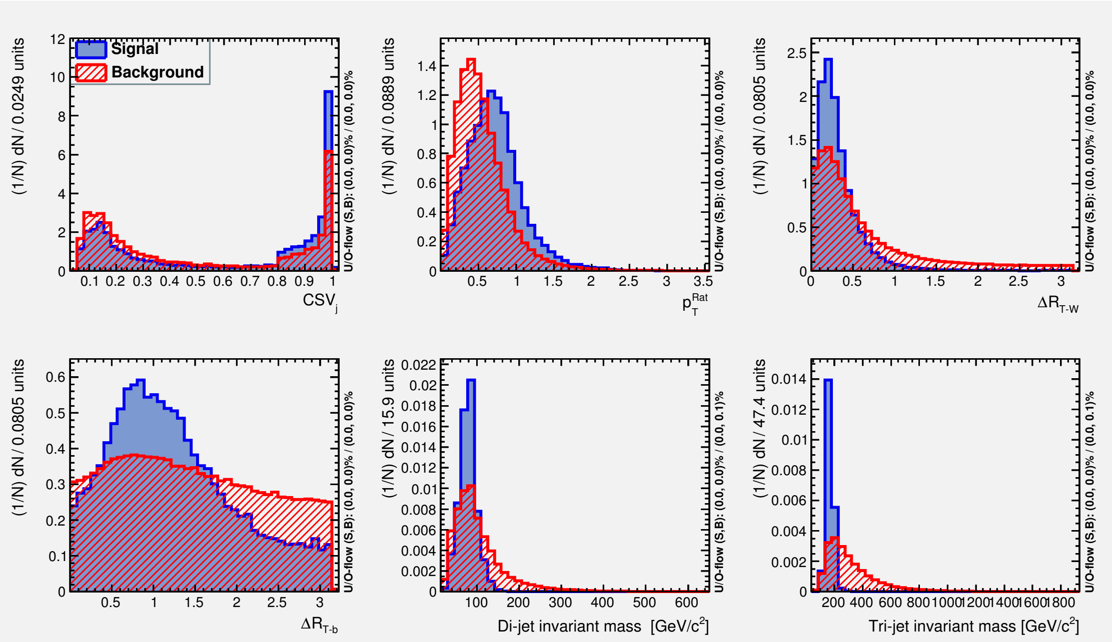


Figure 9.15: Normalised distributions of the six variables used the MVA hadronic Top kinematic reconstruction are shown for good (hatched-red histograms) and bad (solid-blue histograms) tri-jets.

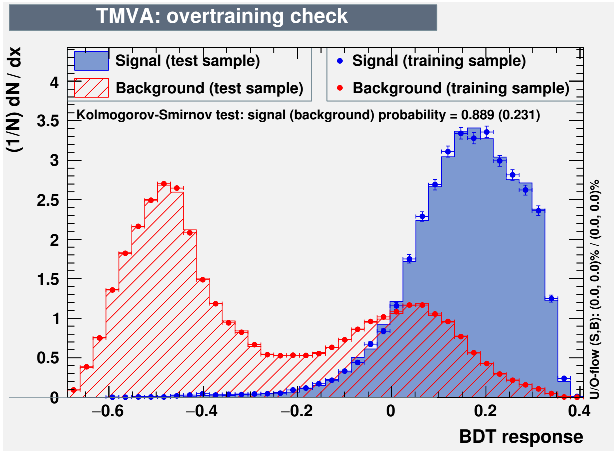


Figure 9.16: The discriminator distributions for the BDT classifier for good (solid blue) and bad (hatched-red) tri-jets in training and validation samples.

The effect of the tri-jet invariant mass variable in the BDT is shown in Fig. 9.17 where it can be clearly seen that this variable contributes to the strong splitting of the BDT output distribution at a value of ≈ -0.2.

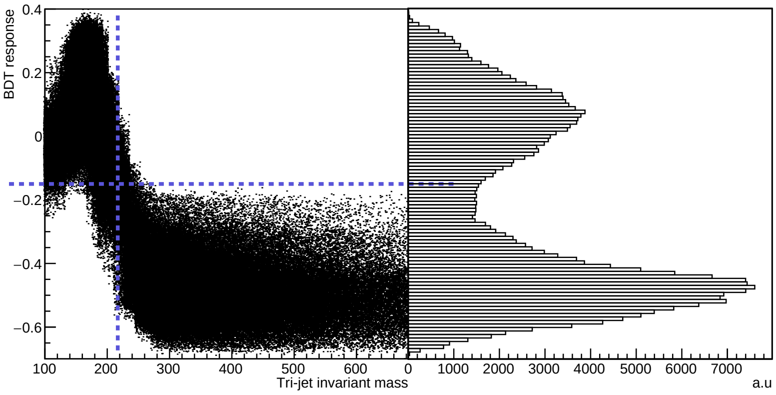


Figure 9.17: The discriminator distributions for the BDT classifier versus tri-jet invariant mass and the projection on the vertical axis. The vertical dashed line indicates the approximate cut value on tri-jet invariant mass at the BDT root node.

Figure 9.18 (left) shows the distribution of good and bad tri-jet combinations in the phase space of tri-jet and di-jet invariant mass. It can be seen from Fig. 9.18 (right) that high BDT discriminator values are found in the region where the good tri-jet combinations are clustered at the top mass.

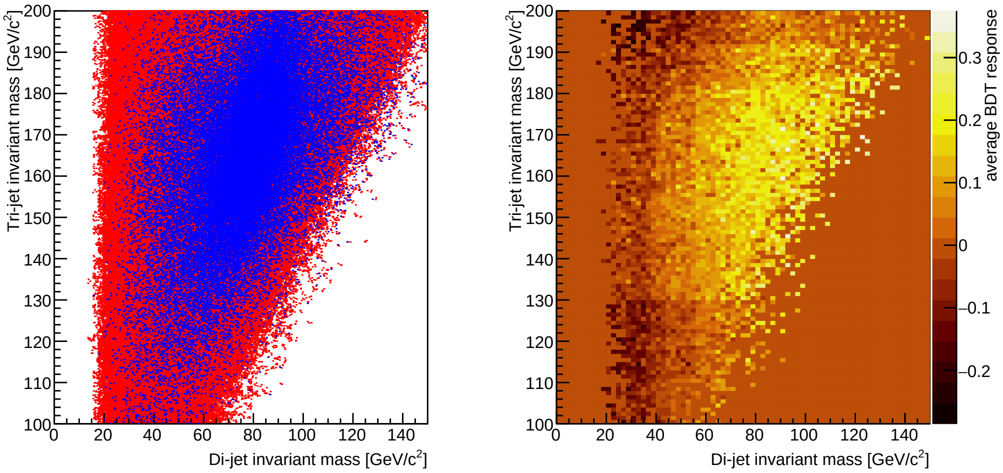


Figure 9.18: (Left) Di-jet versus tri-jet invariant mass distribution for good (blue) and bad (red) tri-jet combination (Right) The average BDT response as a function of Di-jet versus tri-jet invariant mass input variables.

BDT$_{tri-jet2}$

The BDT score of the second highest ranked tri-jet combination, BDT$_{tri-jet2}$, as discussed in Section 6.6, is shown in Fig. 9.19. There is good agreement between the data and simulation and sufficient discrimination power to be used in the event-level BDT.

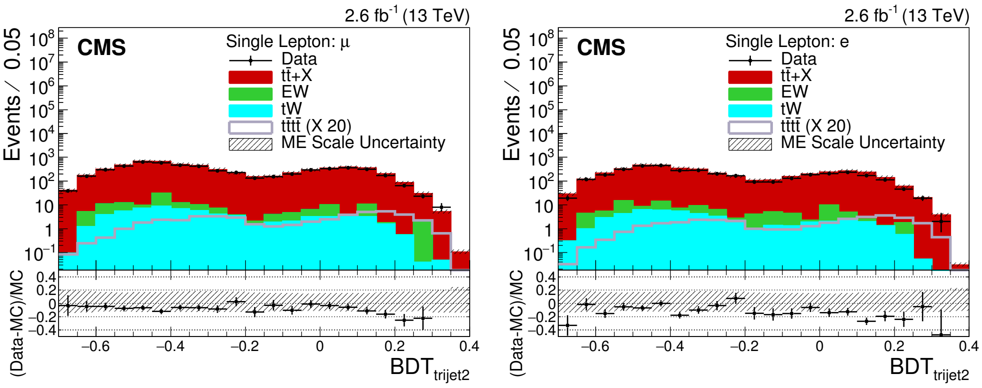


Figure 9.19: The BDT$_{trijet2}$ distributions for data and simulation event in the μ + jets channel (left) and e + jets channel (right).

Reduced Event Variables

The reduced variables formed from the reduced event, where the jets from the highest-ranked hadronic top quark have been removed from the collection of jets, are shown in Figs. 9.20 and 9.21. Again, good agreement is observed between the data and simulation and both variables were found to have good discrimination power in the event-level BDT.

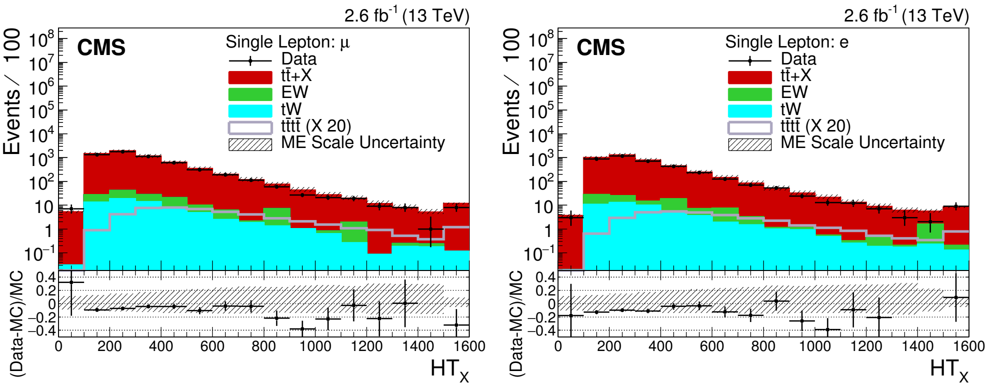


Figure 9.20: The $\mathrm{HT_X}$ distributions for data and simulation event in the μ + jets channel (left) and e + jets channel (right).

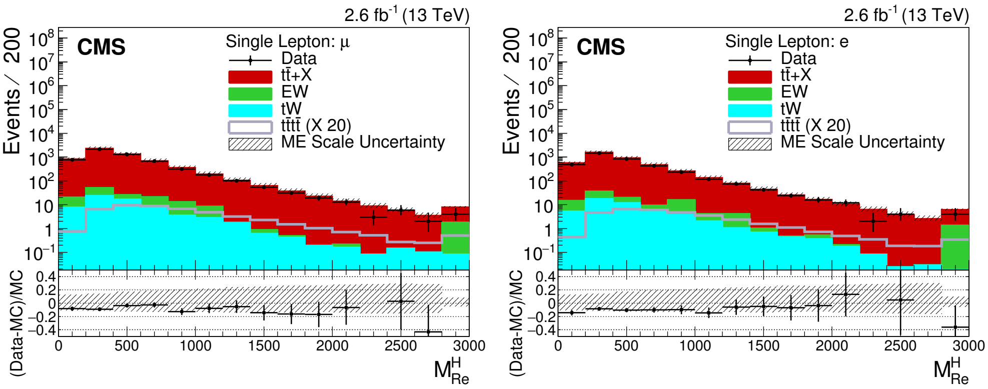


Figure 9.21: The $\mathrm{M^H_{RE}}$ distributions for data and simulation event in the μ + jets channel (left) and e + jets channel (right).

9.7.2 Event activity and b-jet content variables chosen for the event-level BDT

The following variables were chosen for their discrimination power within the event-level BDT. The details are described in Section 6.1. It should be noted that the third-highest CSV and fourth-highest CSV values can be used in the $\sqrt{s}$ = 13 TeV analysis as the CSV distributions have been corrected by the modelling in Section 6.3.2.2.

- $\mathrm{H_T^b}$
- $\mathrm{H_T^{Rat}}$
- $\mathrm{N_{jets}}$
- lepton p_T, $\mathrm{p_T^{ll}}$
- $\mathrm{N_j^W}$
- third-highest CSV
- fourth-highest CSV

9.7.3 Event-level BDT

A $\mathrm{t\bar{t}}$ sample and a $\mathrm{t\bar{t}t\bar{t}}$ sample are provided to the TMVA package to train and test the performance of the event-level BDT using the AdaBoost boosting algorithm [131]. The MadGraph aMC@NLO $\mathrm{t\bar{t}t\bar{t}}$ sample was used with all negative weights set to one in the training. The gradient boosting algorithm [132] can be used with negative weights, hence it was used to verify that the inclusion of negative weights had negligible impact on the final limit compared to setting all weights to unity (See Appendix D.3). Ultimately the AdaBoost algorithm produced a stronger expected limit on the $\mathrm{t\bar{t}t\bar{t}}$ cross section than the gradient boosting algorithm with the negative weights set to one. Therefore, it was the algorithm of choice for this analysis. The jet modelling scale factor weight from Section 6.3.6 is supplied to the BDT as it is important to correct the mismodelling of the most powerful variable input into the BDT.

The separation of the input variables, before any boost weights are applied, is shown in Fig. 9.22 and the ranking of the variables in terms of variable importance are shown for the muon channel in Table 9.5. The variables $\mathrm{N_{jets}}$, third-highest CSV and $\mathrm{H_T^{Rat}}$ are the highest ranked variables and their discriminating power is evident in their initial separation as seen in Fig. 9.22. The lowest ranked variable is $\mathrm{p_T^{ll}}$, which is also seen in Fig. 9.22 to have poor separation power initially. However it still enhances the discrimination power of the BDT to include the $\mathrm{p_T^{ll}}$ variable and it is preferable to have at least one leptonic variable in the list of input variables rather than all hadronic variables.

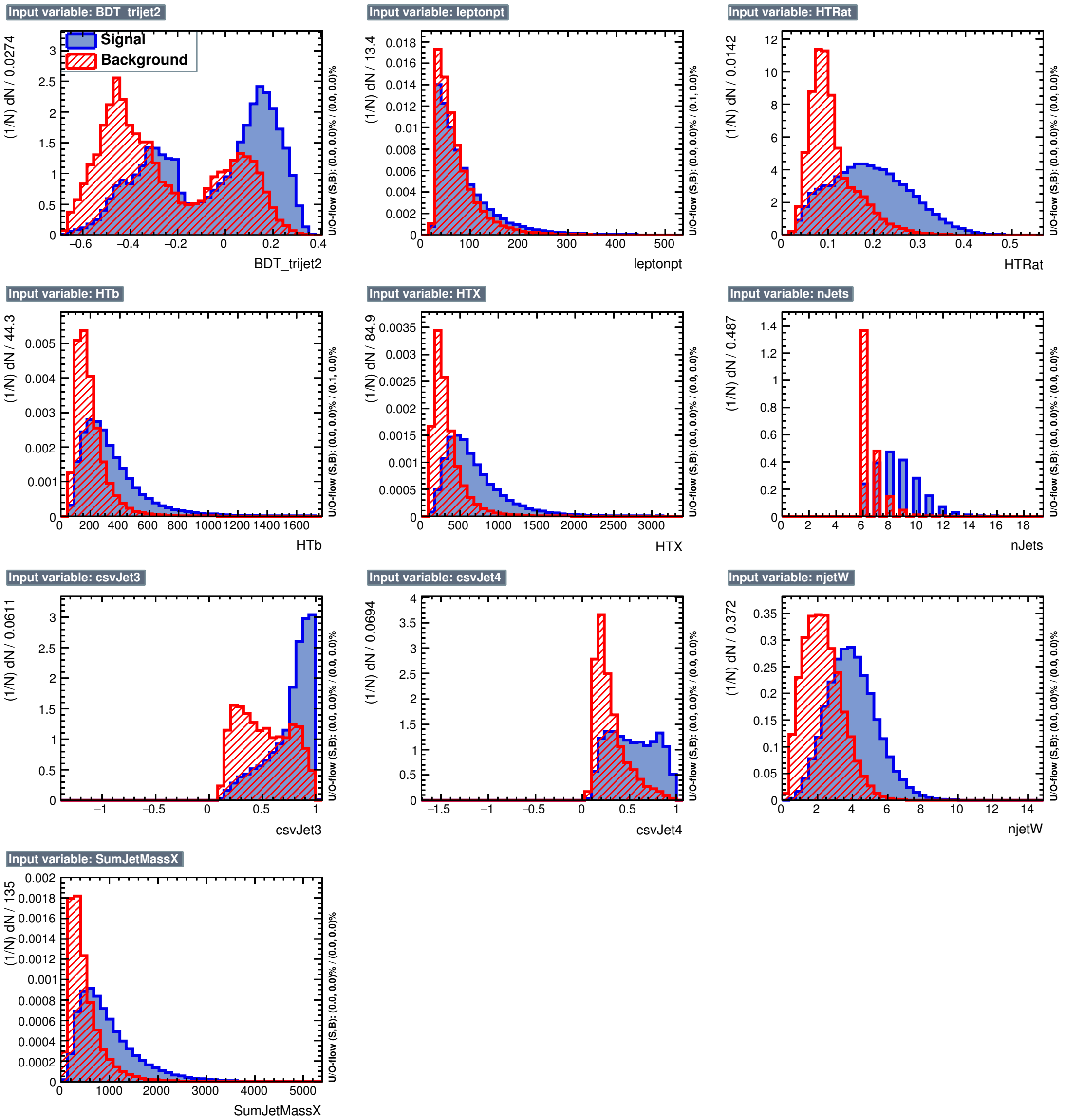


Figure 9.22: Normalised distributions of the input variables in the muon channel taken from TMVA.

Rank	Variable	Importance
1	$\mathrm{N_{jets}}$	1.340e-01
2	third-highest CSV	1.180e-01
3	$\mathrm{H_T^{Rat}}$	1.133e-01
4	$\mathrm{BDT}_{trijet2}$	1.091e-01
5	$\mathrm{N_j^W}$	1.082e-01
6	$\mathrm{M_{RE}^H}$	1.026e-01
7	$\mathrm{HT_X}$	9.867e-02
8	$\mathrm{H_T^b}$	8.650e-02
9	fourth-highest CSV	7.630e-02
10	$\mathrm{p_T^{l1}}$	5.334e-02

Table 9.5: Ranking of variables in order of discrimination power within the BDT.

The output discriminator value for the event level BDT is split into $\mathrm{N_{jets}}$ categories of 6, 7, 8 and ≥9 jets . Further to the Run 1 analysis in Chapter 7, the distributions are also split into $\mathrm{N^M_{tags}}$ categories of 2, 3 and ≥4 b-tags which further categorises the distributions into regions which are more sensitive to the signal and regions which are better for constraining the background. The output BDT distributions are shown for the $N_{\mathrm{jets}} = 6$ and $N^{\mathrm{M}}_{\mathrm{tags}} = 2$ category in Fig. 9.23, which has a large ratio of background to signal. Comparatively in the $N_{\mathrm{jets}} = 6$ and $N^{\mathrm{M}}_{\mathrm{tags}} = 3$ category and $N_{\mathrm{jets}} = 6$ and $N^{\mathrm{M}}_{\mathrm{tags}} \geq 4$ category, in Figs. 9.24 and 9.25 respectively, it can be seen that the background contribution is two orders of magnitude smaller and there is better separation between signal and background. The other BDT categories can be found in Appendix D.4. The lower $\mathrm{N_{jets}}$ and $\mathrm{N^M_{tags}}$ categories are still included as they can be used to help to constrain the $t\bar{t}$ background. These categories act like control regions due to the very large background to signal ratio.

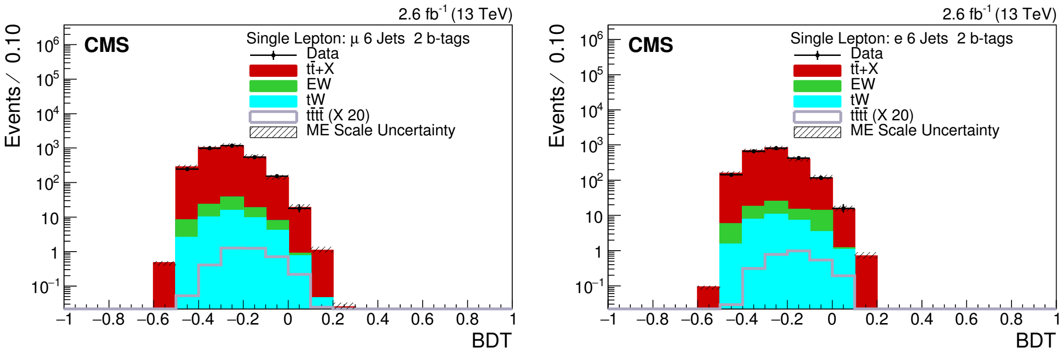


Figure 9.23: The BDT output distributions for the Event-level BDT for data and simulation in the μ + jets channel (left) and e + jets channel (left) are shown for the 6 $\mathrm{N_{jets}}$ and $2\mathrm{N^M_{tags}}$ category.

9.7.3.1 Stability of the event-level BDT

The BDT was seen to be very stable with respect to changes in the number of trees used and to the minimum number of events required at a node for further splitting to occur. The impact on the final expected limit was negligible when changing these BDT hyperparameters. Training was also performed in jet categories of 6-7 jets and ≥8 jets separately to study whether a separate training in the ≥ 8 jets category would improve the power of the BDT to separate $t\bar{t}$ and $t\bar{t}t\bar{t}$ events

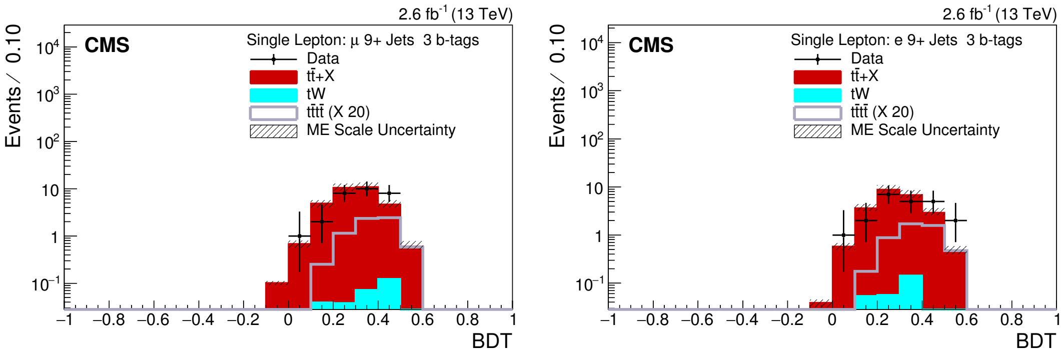


Figure 9.24: The BDT output distributions for the Event-level BDT for data and simulation in the μ + jets channel (left) and e + jets channel (left) are shown for the ≥ 9 $\mathrm{N_{jets}}$ and 3 $\mathrm{N^{M}_{tags}}$ category.

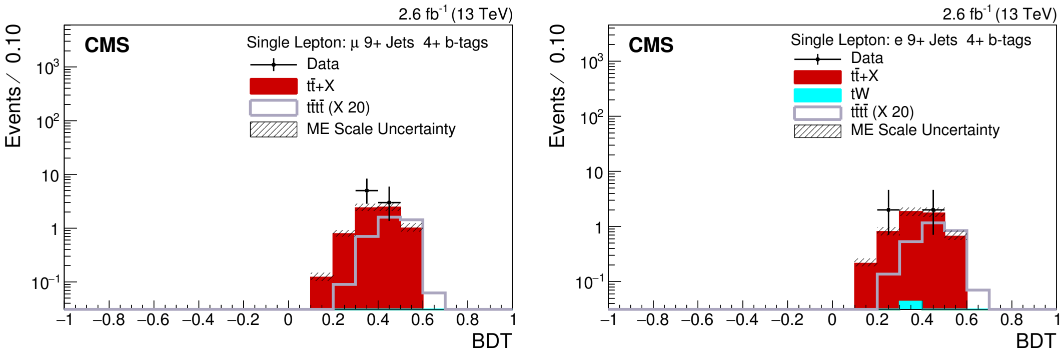


Figure 9.25: The BDT output distributions for the Event-level BDT for data and simulation in the μ + jets channel (left) and e + jets channel (right) are shown for the ≥ 9 $\mathrm{N_{jets}}$ and ≥ 4 $\mathrm{N^{M}_{tags}}$ category.

in a signal-rich region. However, the expected limit derived from this alternative training was not enhanced with respect to the inclusive training in all jet categories. The most significant improvement in the expected limit comes from optimising the $\mathrm{N_{jets}}$ and $\mathrm{N^{M}_{tags}}$ categories. Further categorisation into higher $\mathrm{N_{jets}}$ and $\mathrm{N^{M}_{tags}}$ categories improves the expected limit. However, these categories become very statistically limited with 2.6 fb^{-1} of data.

9.7.3.2 Correlation matrices for BDT input variables

Figure 9.26 shows the correlation matrices for the input BDT variables for the background $\mathrm{t\bar{t}}$ (left) and signal $\mathrm{t\bar{t}t\bar{t}}$ (right). It can be seen that the most highly correlated variables are $\mathrm{N^{W}_{j}}$ and $\mathrm{M^{H}_{RE}}$. This correlation arises from the fact that the $\mathrm{N^{W}_{j}}$ variable is weighted with respect to the p_{T} of the jets in the event and more additional higher-p_{T} jets will lead to a larger value of $\mathrm{M^{H}_{RE}}$. BDTs are effective at handling correlated variables and this was one reason for the choice of using this multivariate algorithm.

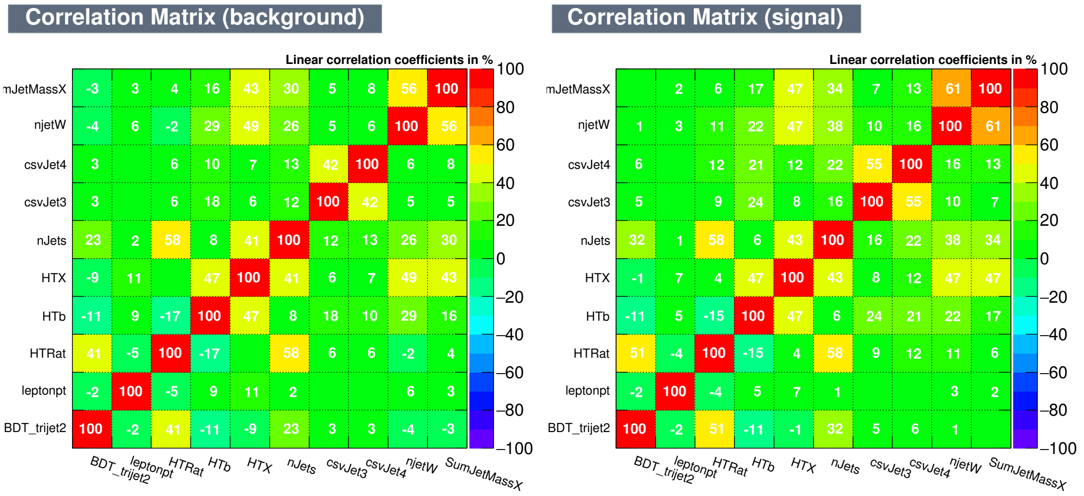


Figure 9.26: The correlation matrices for background (left) and signal (right).

Correlation matrices were used when choosing which variables to include in the final set used to train the BDT for this analysis. Variables which were highly correlated variables in both signal and background were removed as they are largely superfluous.

9.7.3.3 Overtraining tests

Figure 9.27 shows the BDT distribution from training and testing (left) and the associated ROC[2] curve for this training. The training sample is shown as a filled histogram and the testing sample is shown as overlaid data points for the signal $t\bar{t}t\bar{t}$ (blue) and background $t\bar{t}$ (red). It can be seen that there is good agreement between the training and testing samples with Kolmogorov-Smirnov values of 0.159 for background and 0.998 for signal. This suggests that the BDT has not been overtrained.

---

[2] Receiver Operating Characteristic curve: Illustrates the performance of a binary classifier as its discrimination threshold is varied.

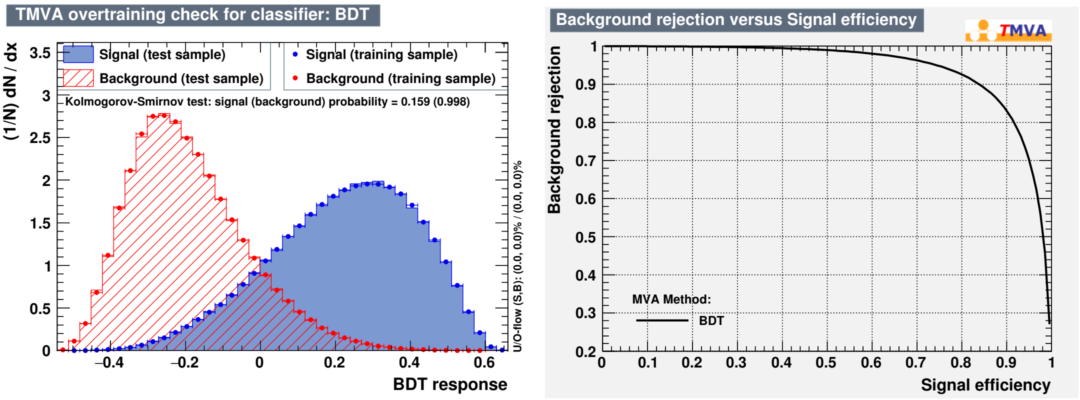


Figure 9.27: The over-training test for the BDT (left) and the rejection of background $t\bar{t}$ (red) vs signal $t\bar{t}t\bar{t}$ (blue) (right).

9.8 Systematic uncertainties

All systematic uncertainties are described in Section 6.8 and some further details about them are given below. The systematic shape templates can be found in Appendix D.5.

- **Luminosity**

  The CMS Luminosity Group gave a recommendation of 2.7% uncertainty on the luminosity [133] which is applied to all simulated backgrounds.

- **Monte Carlo cross sections**

  The uncertainty on the main background of $t\bar{t}$ is $^{+2.5\%}_{-3.4\%}$ renormalisation and factorisation scale and $^{+6.2\%}_{-6.4\%}$ (PDF) [96]. The cross section uncertainties for the other background processes are modelled by assigning a 4% uncertainty and a 10% uncertainty is assigned to the signal process.

- **Lepton SF**

  The lepton SF is applied to all backgrounds. The uncertainty on these SFs is 1.3% in muon channel and 3.6% in electron channel

- **Matrix Element Factorisation and renormalisation scales**

  Weights are available in the $t\bar{t}$ and $t\bar{t}t\bar{t}$ samples which correspond to the factorisation and renormalisation scale (μ_f, μ_s) being individually varied through 1/2u, u and 2u, where u represents the central value. This gives nine weights. However, the extreme unphysical values of (1/2u, 1/2u) and (2u, 2u) are not included. Applying these weights individually gives six alternative event-level BDT histograms to the central histogram, which uses (u,u).

- **Parton Shower Factorisation and Renormalisation Scales**

  The effect of the parton showering (PS) scale in the PYTHIA 8 generator is evaluated by using alternative $t\bar{t}$ samples with the PS scale (u_{PS}) varied by $2u_{PS}$, $1/2u_{PS}$. This shift in the PS scale is equivalent to modifying the value of α_S, hence the alternative parton shower histogram shapes have been inflated by a factor of 1.5 relative to the nominal template to take into account the uncertainty on the jet multiplicity modelling.

- **$t\bar{t}$ generator**

  The POWHEG generator was used to produce the nominal $t\bar{t}$ sample for this analysis. The dependence of the generator can be estimated by running the analysis with the MadGraph_MLM and MadGraph aMC@NLO FxFx $t\bar{t}$ samples. A symmetric envelope is formed around the nominal template by symmetrising the difference in event counts in each bin between the alternative and nominal samples. The MadGraph_MLM sample was found to produce the most conservative uncertainty for this effect and so it was used to produce the symmetric up and down histograms.

- **JES**

  The JES uncertainty is derived by varying the JES by $\pm 1\sigma$ for the $t\bar{t}$ and $t\bar{t}t\bar{t}$ samples.

- **JER**

  The JER uncertainty is derived by varying the smearing by $\pm 1\sigma$ for the $t\bar{t}$ and $t\bar{t}t\bar{t}$ samples.

- **b tagging**

  As detailed in Section 9.4, the difference between b-tagging efficiency in data and simulation is accounted for by the application of scale factors to simulated events via an event weighting procedure. This event weighting procedure was developed for Ref. [134], details of which are documented here [104]. Given the significant uncertainty on these scale factors and that the CSV distributions are input variables to the BDT algorithm, a significant systematic effect is expected. Light flavour contamination, ‘lf’, and linear statistical and quadratic statistical fluctuations, ‘hfstats1’ and ‘hfstats2’, are applied to heavy flavour jets. Heavy flavour contamination, ‘hf’, and linear statistical and quadratic statistical fluctuations, ‘lfstats1’ and ‘lfstats2’, are applied to light flavour jets. Linear and quadratic uncertainties, ‘cferr1’ and ‘cferr2’, are applied to charm flavour jets. A b-tagging JES systematic is applied to light and heavy flavour jets when the standard jet energy scale systematics are applied and hence it is incorporated into the alternative JES systematic shapes. The b-tagging systematic is studied for the $\mathrm{t\bar{t}}$ and $\mathrm{t\bar{t}t\bar{t}}$ samples.

- **Pile up**

  The PU systematic uncertainty is found by varying the MinBias cross section by $\pm5\%$ and is applied to the $\mathrm{t\bar{t}}$ and $\mathrm{t\bar{t}t\bar{t}}$ samples.

- $\sigma_{\mathrm{t\bar{t}b\bar{b}}}$ / $\sigma_{\mathrm{t\bar{t}jj}}$ **modelling** The uncertainty on the measurement of $\sigma_{\mathrm{t\bar{t}b\bar{b}}}$ / $\sigma_{\mathrm{t\bar{t}jj}}$ by CMS [105] is $\pm 0.3\,(\mathrm{stat.}) \pm 0.6\,(\mathrm{sys.})$. Alternative event weights are derived for $\sigma_{\mathrm{t\bar{t}b\bar{b}}}$ / $\sigma_{\mathrm{t\bar{t}jj}}$ which are used to provide the alternative systematic up and down histograms.

9.9 Template fit and upper limit

As no excess of events over the background expectation consistent with SM $\mathrm{t\bar{t}t\bar{t}}$ production was observed, upper limits on $\sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}$ are calculated. The limit setting proceeds by simultaneously fitting the BDT output distributions of signal and backgrounds to the BDT distribution of data in both the μ + jets and e + jets channels in all 12 $\mathrm{N_{jets}}$ and $\mathrm{N^{M}_{tags}}$ categories. The Higgs Combine Tool is used to perform the fit, assigning lognormal uncertainties to normalisation systematic uncertainties and the vertical morphing technique described in Section 6.9 for the shape systematic uncertainties. The fit produces a fitted shape and normalisation and best-fit values for all nuisance parameters and the parameter of interest. To avoid prohibitively large computing times, the approximate *asymptotic* approach is used to calculate the $\mathrm{CL_S}$ limits, which can be found in Table 9.6 in units of σ_{SM}.

Channel	Expected limit	Uncertainty	Observed limit
μ	20.6	$+12.9 - 7.2$	20.8
e	26.4	$+16.6 - 9.3$	33.5
Combined	16.0	$+9.8 - 5.5$	16.8

Table 9.6: Extracted expected limits for $\mathrm{N_{jets}}$ and $\mathrm{N^{M}_{tags}}$ categorized templates in multiples of σ_{SM}.

The combined expected limit is 147.2^{+90}_{-51} fb and the observed limit is 154.6 fb for the single lepton + jets channel.

9.9.0.4 Nuisance parameters

Figure 9.28 shows the variation of the post-fit nuisance parameters, θ, with respect to their pre-fit values. As all of the parameters have not shifted outside of the pre-fit uncertainty, the number of uncertainties and their modelling is deduced to be appropriate for modelling the data. It can be seen that both the $\mathrm{t\bar{t}}$ ME scale uncertainty and the PS scale, denoted *ttMEScale* and *scaleH* respectively on the figure, have much smaller post-fit uncertainties with respect to their pre-fit uncertainties which suggests that they were larger than necessary.

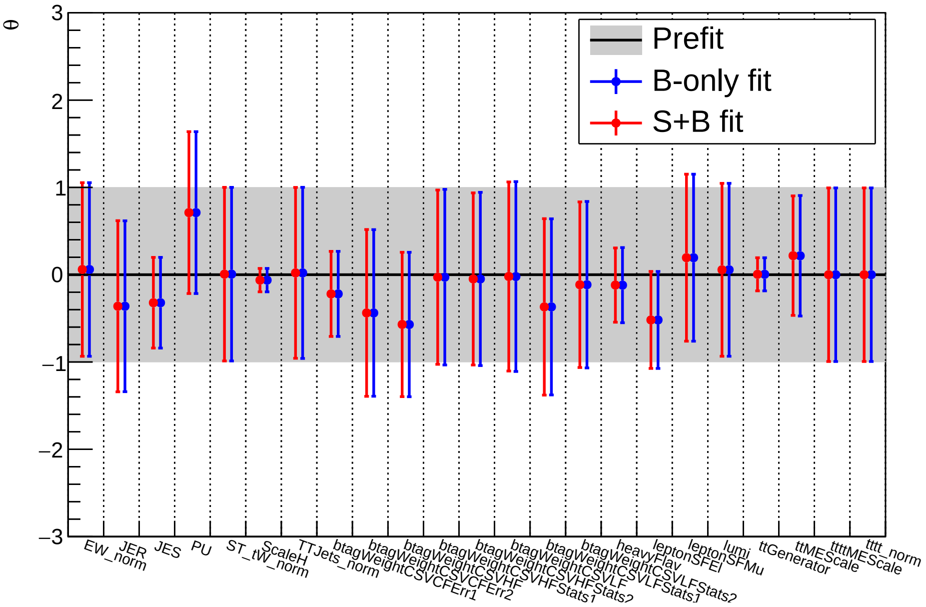


Figure 9.28: Post-fit nuisance parameters for data.

9.10 Alternative limit setting using H_{T} distributions for template fitting

The limit setting procedure was repeated using the H_{T} distributions to perform the template fit. This is to compare a simple discriminating variable between $\mathrm{t\bar{t}}$ and $\mathrm{t\bar{t}t\bar{t}}$ to the output BDT discriminator variable to estimate the gain of using a BDT.

Channel	Categorized	Uncertainty
μ	26.4	$+16.2 - 9.2$
e	30.6	$+18.8 - 10.7$
Combined	19.6	$+11.7 - 6.7$

Table 9.7: Extracted expected limits for $\mathrm{N_{jets}}$ and $\mathrm{N^{M}_{tags}}$ categorized templates of H_{T} in multiples of σ_{SM}.

The results in Table 9.7 correspond to a combined expected limit of 180.3^{+108}_{-62} fb using H_{T} to make the template fit compared to an expected limit of 147.2^{+90}_{-51} fb when using the event level BDT to make the template fit. Hence, using the BDT the expected limit is improved by ≈ 20%.

Further studies of the expected limit with each systematic removed can be found in Appendix D.5.1. The correlation matrix for the fit nuisance parameters can be found in Appendix D.6.

9.11 Combination with OS dilepton channel and SS dilepton channel

The sensitivity of the search for standard model four top quark production can be improved by combining with other search channels. An opposite-sign (OS) search was developed in parallel with the single lepton channel study described in this chapter [109]. The analysis selects events which contain any combination of $\mu^+\mu^-$, $\mu^\pm e^\mp$, e^+e^-. It uses the same hadronic top quark reconstruction as in Section 9.7.1 to identify the BDT value for highest-ranked top quark candidate, $BDT_{trijet1}$. This variable is fed into the event-level BDT along with other variables based on the event-topology, event activity and b-jet content. A simultaneous fit was performed using the BDT histogram templates described above for the single lepton channel and the BDT histogram templates (which are split only in $\mathrm{N_{jets}}$ categories due to statistical limitations) from the dilepton channel. All systematic uncertainties apart from the lepton scale factors were treated as correlated. The results of this fit can be found in Table 9.8 in the row labelled *Combined (single lepton and OS dilepton)*. It is clear that the OS dilepton channel alone is not as sensitive as the single lepton channel, which is due in part to it having a smaller branching ratio. However, its combination with the single lepton channel improves the overall sensitivity.

The analysis was then further combined with a search for new physics in events with same-sign (SS) dileptons which places limits on the SM production of four top quarks [135]. This search benefits from very low numbers of events from background processes which gives rise to its good signal sensitivity. The luminosity, JES and PU systematic uncertainties were treated as correlated between the SS dilepton channel and the other two channels. The uncertainty in response of the CMS trigger system to events containing dileptons is also treated as correlated between the two dilepton analyses, whilst all other systematic uncertainties were treated as fully uncorrelated between the SS dilepton analysis and the other two search channels. The combination of all channels is listed in Table 9.8. It can be seen that this gives a significant improvement in the expected limit compared to any individual channel.

Table 9.8: Expected and observed 95% CL upper limits on the SM $\mathrm{t\bar{t}t\bar{t}}$ production as a multiple of $\sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}$ and in fb. The values quoted on the expected limits are the 1 standard deviation uncertainties and include all statistical and systematic uncertainties.

Channel	Expected Limit (x $\sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}$)	Observed Limit (x $\sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}$)	Expected limit (fb)	Observed Limit (fb)
Single lepton	$16.0^{+9.8}_{-5.5}$	16.8	147^{+90}_{-51}	155
Dilepton (opposite sign)	$24.0^{+16.3}_{-8.9}$	14.1	221^{+150}_{-82}	130
Combined (single lep -ton and OS dilepton)	$12.6^{+7.9}_{-4.4}$	9.9	116^{+73}_{-41}	91
Dilepton (same sign)	$11.0^{+6.2}_{-3.8}$	12.9	101^{+57}_{-35}	119
Combined (all channels)	$7.5^{+4.1}_{-2.5}$	7.2	69^{+37}_{-23}	66

9.12 Summary and conclusion

The SM production of four top quarks at $\sqrt{s} = 13$ TeV has been studied with 2.6 fb^{-1} of data from the 2015 CMS dataset. In the absence of an excess, limits were placed on the SM cross section, which is 9.2 fb. The single lepton channel was primarily studied in the μ + jets and e + jets final states. Baseline selection requirements were implemented to suppress backgrounds and select the signal $\mathrm{t\bar{t}t\bar{t}}$ process. Good agreement was observed between the simulation and data in many variables after corrections were applied to the events. BDTs were employed to reconstruct hadronically decaying top quarks and then to increase the separation of the signal and background processes using several discriminating variables including those formed from the hadronic top quark BDT. Many tests were performed to check the performance of the BDTs, that it wasn't overtrained and was stable with respect to changing the hyperparameters. Categorisation of the histograms used in the template fit was optimised to enhance the sensitivity of the analysis and hence reduce the expected limit. The limit set on the SM production of four top quarks is $16.0^{+9.8}_{-5.5} \times \sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}$ expected and $16.8 \times \sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}$ observed which equates to 147.2^{+90}_{-51} fb expected and 154.6 fb observed.

9.13 Discussion of other searches for $\mathrm{t\bar{t}t\bar{t}}$ production studies at $\sqrt{s} = 13$ TeV

There are a number of searches for the production of four top quarks at $\sqrt{s} =$ 13 TeV in the single lepton, opposite-sign dilepton and same-sign dilepton channels, at both the CMS and ATLAS experiments. A summary of the results is given in Table 9.9 including three CMS analyses in three different channels using the 2015 CMS dataset and a combination of those results as discussed in Section 9.11. The table also includes two different ATLAS searches, ATLAS-CONF-2016-013 and ATLAS-CONF-2016-020, which place limits on four-top-quark production in the single lepton channel using the 2015 ATLAS dataset of 3.2 fb^{-1}. ATLAS-CONF-2016-104 is a progression from ATLAS-CONF-2016-013 that uses the 2015 and 2016 ATLAS datasets using a total of 13.2 fb^{-1}. Finally there is a same-sign dilepton search by ATLAS which uses the 2015 ATLAS dataset with 3.2 fb^{-1}.

Of all the analyses the same-sign dilepton searches are the most sensitive. ATLAS-CONF-2016-020 is more sensitive than ATLAS-CONF-2016-013 as it categorises into more jet and b-tag categories. There is a significant reduction in the limit between ATLAS-CONF-2016-013 and ATLAS-CONF-2016-140 of $\approx 40\%$ by using four times more data. The results, between searches in the same channels in CMS and ATLAS are compatible with each other within the systematic errors. Overall the CMS combination has the strongest limit on four-top-quark production so far of 66 fb (69^{+37}_{-23} fb) observed (expected).

Analysis	Channel	$\mathcal{L}$ (fb^{-1})	Expected limit (fb)	Observed limit (fb)
CMS (This analysis)	Single lepton	2.6	147^{+90}_{-51}	155
CMS	OS dilepton	2.6	221^{+150}_{-82}	130
CMS [135]	SS dilepton	2.6	101^{+57}_{-35}	119
CMS combination	Above combined	2.6	69^{+37}_{-23}	66
ATLAS-CONF-2016-013 [136]	Single lepton	3.2	180	370
ATLAS-CONF-2016-020 [137]	Single lepton	3.2	143	190
ATLAS-CONF-2016-032 [138]	SS dilepton	3.2	107	95
ATLAS-CONF-2016-104 [139]	Single lepton	13.2	110	130

Table 9.9: Limits of four-top-quark production by a variety of searches in CMS and ATLAS at $\sqrt{s} = 13$ TeV.

10 Conclusion

In this thesis, searches for the production of four top quarks at $\sqrt{s} = 8$ TeV and $\sqrt{s} = 13$ TeV at the CMS experiment at CERN were presented. In addition to this, a phenomenological interpretation of the results at $\sqrt{s} = 8$ TeV in the context of a simplified model with sgluon particles was presented.

10.1 Summary of results

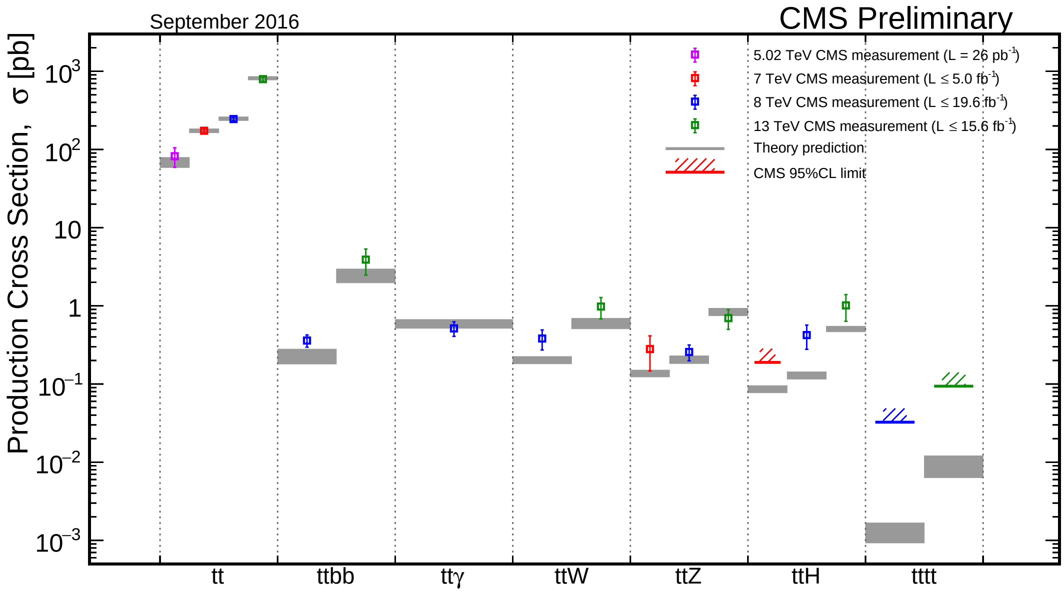
In Run 1 and Run 2, the main background after the baseline selection requirement was $\mathrm{t\bar{t}}$ production where the extra jets required to meet the selection requirements tend to come from ISR and FSR. Smaller contributions to the background came from single top production, W and Z boson production, and even less so from $\mathrm{t\bar{t}H}$, $\mathrm{t\bar{t}W}$, $\mathrm{t\bar{t}Z}$.

One BDT was used to identify hadronically decaying top quarks and a second BDT was used to separate signal and background using variables which described the event activity, b-jet content and variables derived from the hadronic top BDT.

At $\sqrt{s} = 8$ TeV, the BDT templates were split into $\mathrm{N_{jets}}$ categories only and a simultaneous fit was made in the electron and muon channels. This resulted in a 95% CL on $\sigma_{\mathrm{t\bar{t}t\bar{t}}}$ / $\sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}$ of 25 (24.6 ± 13) observed (expected) which corresponds to a cross section of 32 fb (32.0 ± 17 fb) observed (expected). This result was both competitive and consistent with the other limits on $\mathrm{t\bar{t}t\bar{t}}$ production in Run 1 from both CMS and ATLAS.

At $\sqrt{s} = 13$ TeV, the analysis saw improvements from using larger samples of simulated events and better modelling of $\mathrm{t\bar{t}t\bar{t}}$ by using a NLO sample rather than LO. Improvements were also made to the modelling of b-tagging by using scale factors which correct the CSV discriminator distributions. The variables input into the event-level BDT were reoptimised and the BDT was shown to be very stable to variations in its hyperparameters. Further to the categorisation of the

BDT templates in N_{jets} categories in the Run 1 analysis, the templates were also categorised by $N^{\text{M}}_{\text{tags}}$ categories. This was motivated by the increase in sensitivity observed by other searches for $\text{t}\bar{\text{t}}\text{t}\bar{\text{t}}$ production. A simultaneous fit was performed in the electron and muon channels in all N_{jets} and $N^{\text{M}}_{\text{tags}}$ categories which resulted in a 95% CL limit of 16.8 $(16.0^{+9.8}_{-5.5}) \times \sigma^{SM}_{\text{t}\bar{\text{t}}\text{t}\bar{\text{t}}}$ observed (expected) in the single lepton channel. This result was further combined with an OS-dilepton and a SS-dilepton search for four top quark production in CMS, which resulted in an expected limit of 7.2 $(7.5^{+4.1}_{-2.5}) \times \sigma^{SM}_{\text{t}\bar{\text{t}}\text{t}\bar{\text{t}}}$ observed (expected) which equates to a cross section of 66 fb $(69^{+37}_{-23}$ fb) observed (expected) - the world's tightest limit on $\text{t}\bar{\text{t}}\text{t}\bar{\text{t}}$ production at present.



All results at: https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsTOP

Figure 10.1: Top quark production cross section summary plot showing measurement using square points and 95% upper limits by a hatched band. Theory predictions are shown where the grey band represents the uncertainty on the prediction [140].

Figure 10.2 shows the CMS top quark cross section measurements as of September 2016. This figure highlights how much larger the main background of $\text{t}\bar{\text{t}}$ production is compared to $\text{t}\bar{\text{t}}\text{t}\bar{\text{t}}$ production. Evidence has been seen for some of the rarer top physics processes, $\text{t}\bar{\text{t}}\text{b}\bar{\text{b}}$, $\text{t}\bar{\text{t}}\gamma$, $\text{t}\bar{\text{t}}$W, $\text{t}\bar{\text{t}}$Z and $\text{t}\bar{\text{t}}$H, shown in Fig. 10.2. It can be seen that the production of $\text{t}\bar{\text{t}}\text{t}\bar{\text{t}}$ is at least an order of magnitude smaller than the other rare top physics processes being studied. The limit at $\sqrt{s}$ = 8 TeV is from the analysis in this thesis and the limit shown for $\sqrt{s} = 13$ TeV on the figure comes from the preliminary result [109] which was released before changing the b-tagging scale factors for the final result presented in Chapter 9.

10.2 Future prospects

Figure 10.2 shows the expected 95% upper limit on $\sigma_{t\bar{t}t\bar{t}}$ / $\sigma^{SM}_{t\bar{t}t\bar{t}}$ against integrated luminosity, where the integrated luminosity has been artificially enhanced within the Higgs Combine Tool. The dashed black line indicates the Run 2 result from Chapter 9. It can be seen that, in the absence of a signal, without any further enhancements or optimisations to the analysis workflow, the expected limit should decrease down the level of the SM expectation simply by increasing the amount of data up to 30-100 fb^{-1}.

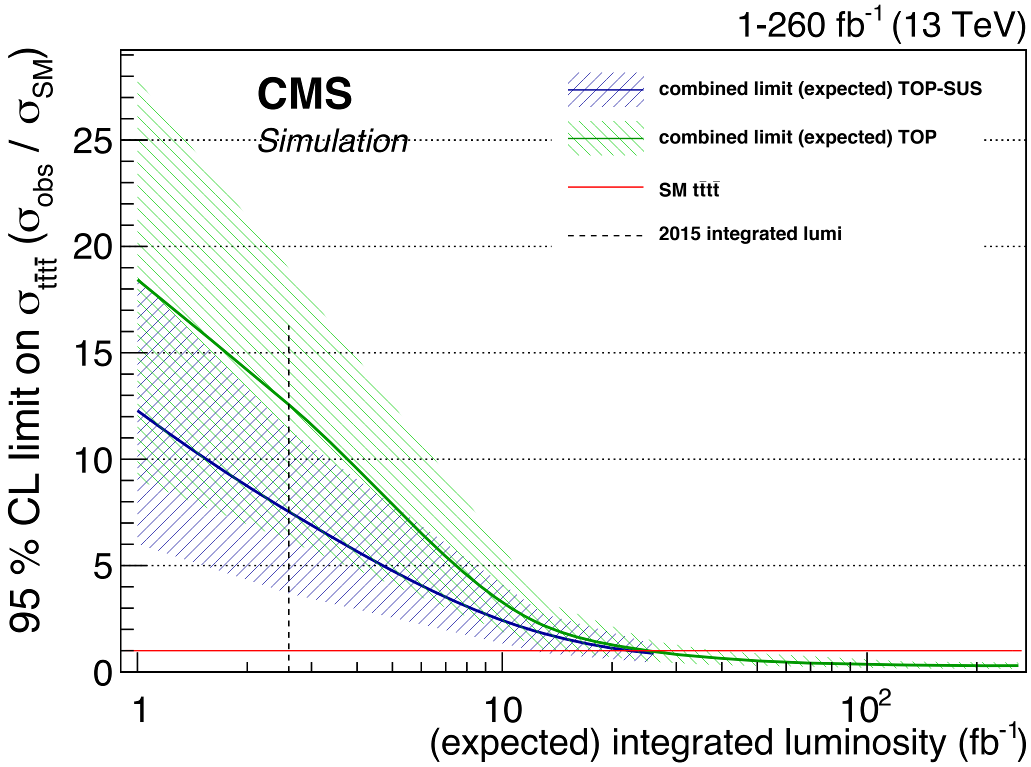


Figure 10.2: Extrapolated limits on four-top-quark production using the single lepton and opposite-sign dilepton analyses (denoted TOP and in green) and using the single lepton, opposite-sign dilepton and the same-sign dilepton analyses (denoted SUS and in blue). The red line indicate the SM production rate. See Table 9.9 for details of these analyses using 2.6 fb^{-1}.

This suggests that it may be possible to produce a limit consistent with $1 \times \sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}$ using the dataset which has been collected by CMS in 2016 which has 41.1 fb^{-1} of data as shown in Figure 3.2.

It may be possible to optimise the analysis further by including the $N_{\mathrm{jets}} = 5$ category, as in Ref. [137], to further constrain $\mathrm{t\bar{t}}$. Many four-top-quark searches have also categorised in H_{T} and $E_{\mathrm{T}}^{\mathrm{miss}}$ [125, 137]. So far the hadronic top quark BDT has been trained on top quarks from $\mathrm{t\bar{t}}$ events, but it may be beneficial to also train on hadronic tops from $\mathrm{t\bar{t}t\bar{t}}$ events. It may be possible to find more discriminating variables for use in the event-level BDT. Lastly, the systematic uncertainties used in Chapter 9 were potentially larger than necessary to sufficiently model the simulation. There were several shape systematic uncertainties which could potentially be substituted into normalisation uncertainties.

In summary, the analyses in this thesis have produced the most stringent limit on four-top-quark production to date. This result constrains which BSM scenarios are possible, for instance by reducing the allowable phase-space of the sgluon particle, in mass and coupling to the top quark, as seen in Chapter 8. It is possible that, with the enhancements to the analysis mentioned and the much larger dataset collected by CMS in 2016, direct evidence of the production of four top quarks is within our reach.

Appendix A | Cross check on Multi-jet background estimation

Figure 3 shows that there is no significant correlation between $E_{\mathrm{T}}^{\mathrm{miss}}$ and Reliso in $\mathrm{t\bar{t}}$ events.

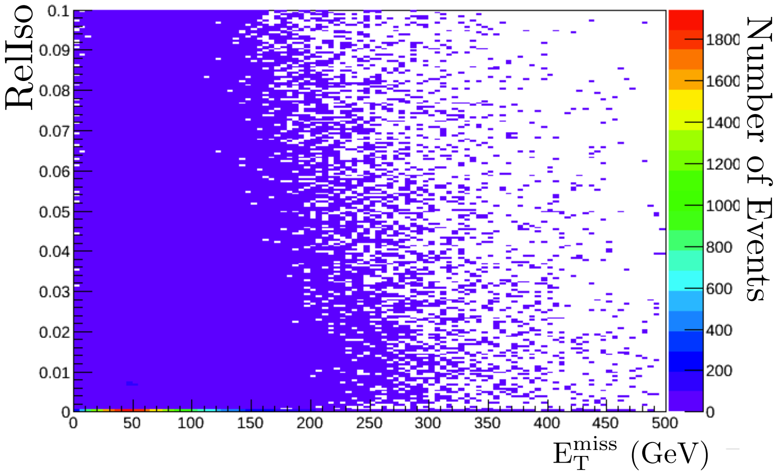


Figure A.1: $E_{\mathrm{T}}^{\mathrm{miss}}$ versus RelIso in $\mathrm{t\bar{t}}$ events.

Appendix B | Cross checks on Run 1 $t\bar{t}t\bar{t}$ analysis at $\sqrt{s} = 8$ TeV

This study was undertaken before the decision was made to split the BDT templates by the N_{jets} categories.

In the 2012 CMS differential cross-section $t\bar{t}$ analysis [141] the p_T spectrum of top quarks tends towards higher values in the MadGraph simulation than it is in data. Scale factors were derived by CMS to compensate for this effect. The analysis was performed with and without the application of the top p_T scale factors and the observed effect was negligible, as seen in Fig. A.1, and hence these scale factors were not applied for the final result and no systematic uncertainty was included.

The uncertainty on the parton distribution functions (PDFs) are a potential source of systematic uncertainty. The method used by CMS to model this effect is given in [142]. The BDT distributions which correspond to the maximal downward and upward fluctuation due to the uncertainty on the PDFs have a small effect on the shape of the BDT, as seen in Fig. A.1, and are not considered further.

The uncertainty due to the choice of PYTHIA tune used in the hadronisation of $t\bar{t}$ events is considered. The nominal tune used is the $Z2^*$ tune which is compared to the alternative $P11$ tune [143, 144]. Again, there is a very small effect on the shape of the BDT, as seen in Fig. A.1, so this uncertainty is not included in the final fit.

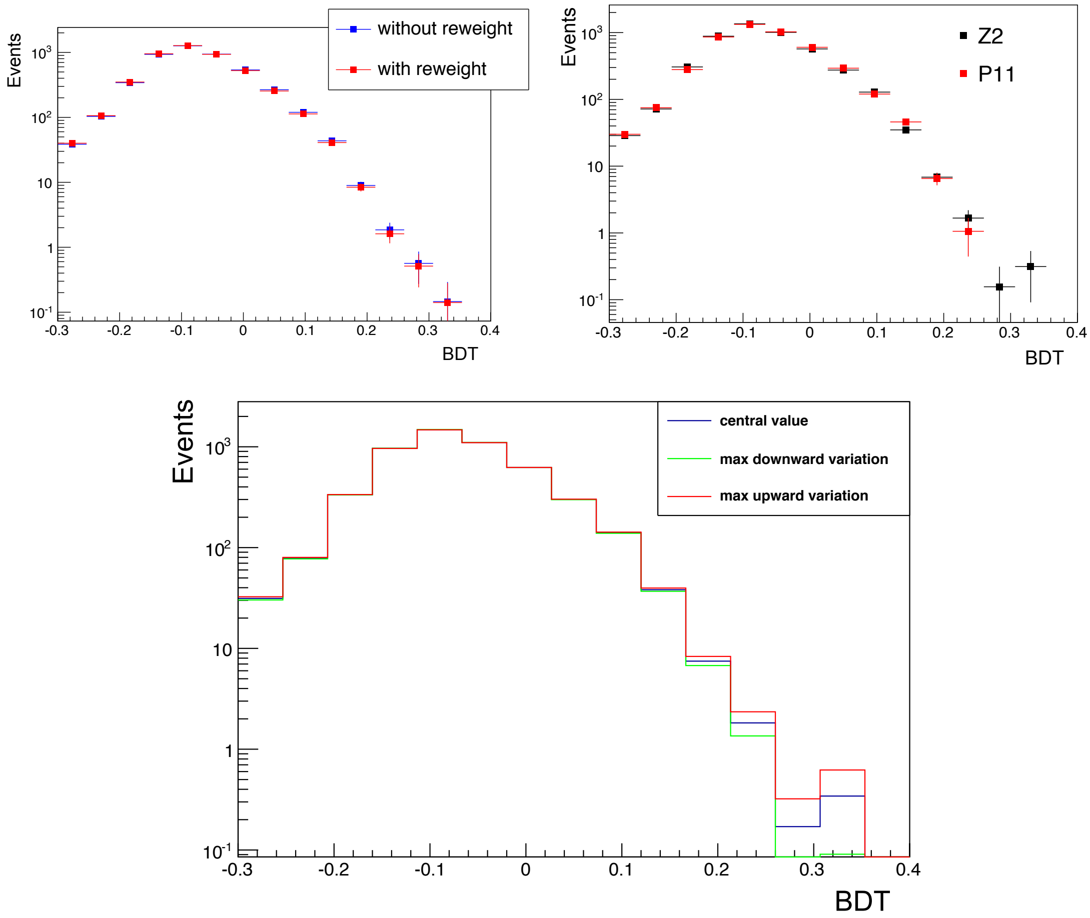


Figure B.1: The BDT discriminator distributions of $t\bar{t}$ simulation with and without the top quark p_{T} reweighting (top left), PYTHIA tunes (top right) and PDF uncertainty (bottom).

B.1 Cross-checks on the BDT

Rankings of the input variables in terms of importance in the BDT are provided in tables A.1 and A.2, respectively.

Rank	Variable	Variable Importance
1	Jet6Pt	1.109e-01
2	HTX	1.092e-01
3	MultiTopness	1.074e-01
4	HTRat	1.058e-01
5	HTb	1.003e-01
6	Jet5Pt	9.848e-02
7	HTH	9.838e-02
8	nJets	9.664e-02
9	SumJetMassX	9.266e-02
10	nTags	7.732e-02

Table B.1: The rankings of the input variables in terms of importance in the BDT for the μ + jets channel are provided.

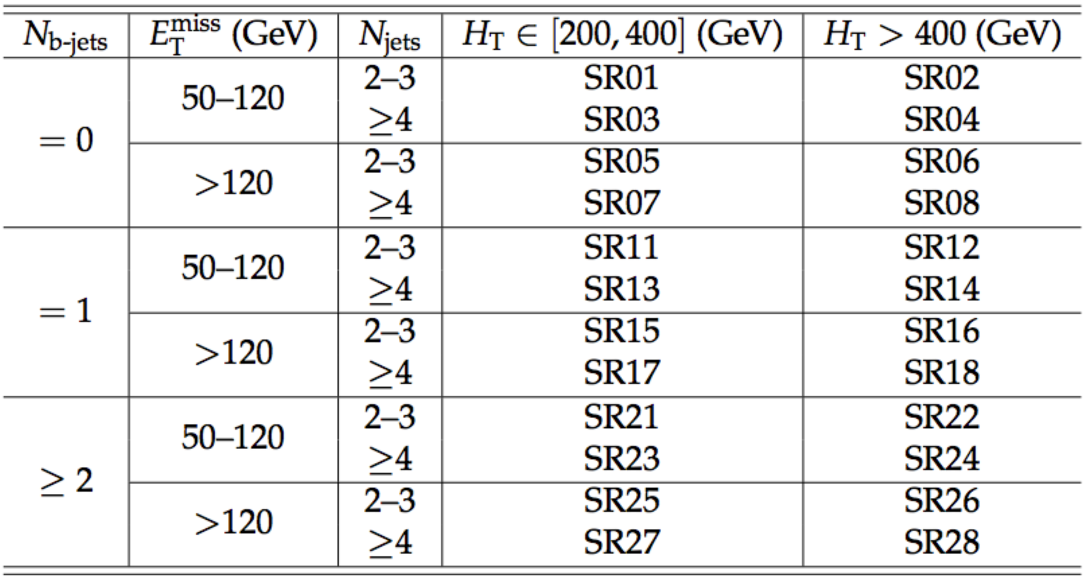
Rank	Variable	Variable Importance
1	MultiTopness	1.201e-01
2	HTRat	1.186e-01
3	HTX	1.175e-01
4	HTb	1.148e-01
5	HTH	1.142e-01
6	Jet5Pt	9.730e-02
7	SumJetMassX	9.671e-02
8	Jet6Pt	8.913e-02
9	nJets	7.319e-02
10	nTags	5.850e-02

Table B.2: The rankings of the input variables in terms of importance in the BDT for the e + jets channel are provided.

Appendix C | Further detail on the phenomenological study in Chapter 8

C.1 Signal regions

Figure B.1 shows the signal regions defined in Ref. [125]. These search regions are defined for high-p_T (lepton p_T >20 GeV) analyses and low-p_T analyses lepton p_T >10 GeV), the former of which is used in this thesis for SR28.



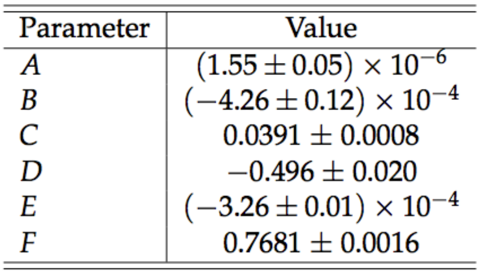
$N_{\text{b-jets}}$	E_T^{miss} (GeV)	N_{jets}	$H_T \in [200, 400]$ (GeV)	$H_T > 400$ (GeV)
$= 0$	50–120	2–3	SR01	SR02
		≥4	SR03	SR04
	>120	2–3	SR05	SR06
		≥4	SR07	SR08
$= 1$	50–120	2–3	SR11	SR12
		≥4	SR13	SR14
	>120	2–3	SR15	SR16
		≥4	SR17	SR18
≥ 2	50–120	2–3	SR21	SR22
		≥4	SR23	SR24
	>120	2–3	SR25	SR26
		≥4	SR27	SR28

Figure C.1: Definition of the signal regions for the high-p_T analysis [125].

Different combinations of these signal regions are used to provide “broad coverage of strongly produced SUSY particles, including signatures with low hadronic activity as well as signatures involving third-generation squarks” [125].

C.2 Parameterisation of the b-tagging of b-quark jets

There is a small bump in the efficiency curve, shown in Fig 8.2 (top-left). This is due to the parameterisation of the curve fitted from data. A polynomial of the form $Ax^3 + Bx^2 + Cx + D$ is used for $p_{\mathrm{T}} < 120$ GeV and a linear fit $Ex + F$ is used above that p_{T} threshold. The matching of these two function is the cause of the small bump in the curve and can also be seen in Ref. [125].



Parameter	Value
A	$(1.55 \pm 0.05) \times 10^{-6}$
B	$(-4.26 \pm 0.12) \times 10^{-4}$
C	0.0391 ± 0.0008
D	-0.496 ± 0.020
E	$(-3.26 \pm 0.01) \times 10^{-4}$
F	0.7681 ± 0.0016

Figure C.2: b-tagging efficiency parameters.

The efficiency functions for all efficiencies shown in Fig 8.2 can be found in Ref. [125].

Appendix D | Cross checks on Run 2 $t\bar{t}t\bar{t}$ analysis at $\sqrt{s} = 13$ TeV

D.1 Scale factors

The distributions of the lepton SF, b-tagging CSV SF, PU SF and jet multiplicity modelling scale factor are given for the $\sqrt{s} = 13$ TeV analysis in Chapter 9.

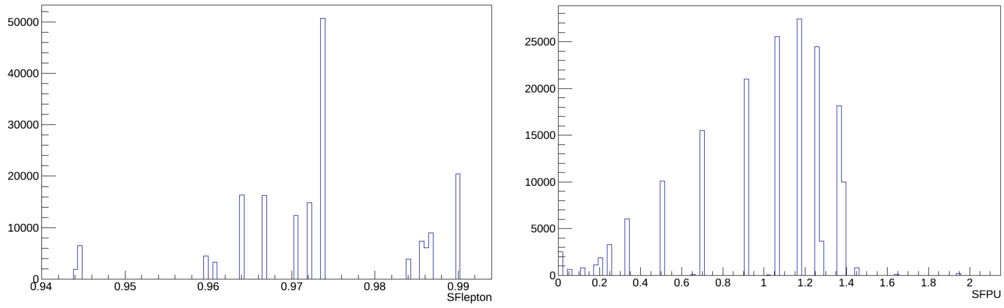


Figure D.1: Lepton SF (left) and PU SF (right).

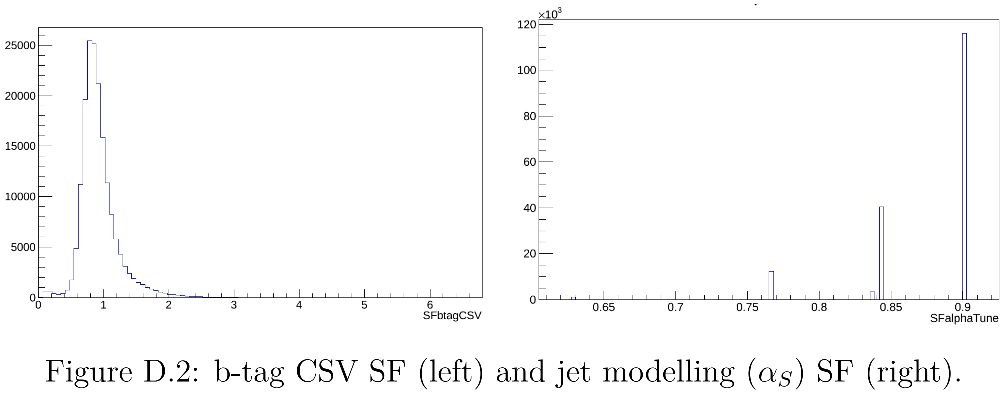


Figure D.2: b-tag CSV SF (left) and jet modelling (α_S) SF (right).

Appendix E | Cross checks on Run 2 $\mathrm{t\bar{t}t\bar{t}}$ analysis at $\sqrt{s} = 13$ TeV

E.1 Comparison of alternative $\mathrm{t\bar{t}}$ generators

Figure D.1 shows that the uncertainty from the MadGraph aMC@NLO generator is contained within the uncertainty from the MadGraph_MLM generator, therefore it is conservative to use the MadGraph_MLM generator as the systematic shape for differences in the BDT distribution due to generator choice.

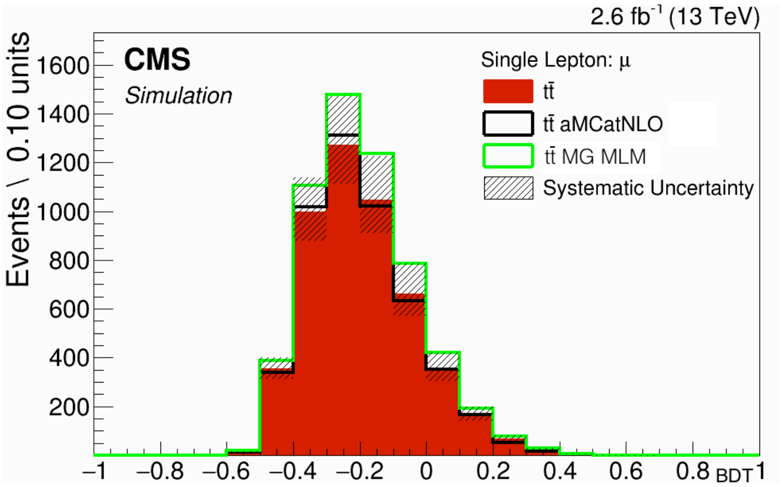


Figure E.1: Inclusive BDT distribution for $\mathrm{t\bar{t}}$ generators POWHEG +PYTHIA, MadGraph_MLM and MadGraph aMC@NLO FxFx.

E.2 TTZ, TTW, TTH MC backgrounds

The contributions from $\mathrm{t\bar{t}}$ + B, where B = W, Z or H, were added to the predicted $\mathrm{t\bar{t}}$ yields to give a prediction for the net $\mathrm{t\bar{t}}$ + B background. The event-level BDT discriminant shapes for these contributions closely follow those of the $\mathrm{t\bar{t}}$

contribution and are very different from those predicted for the $\mathrm{t\bar{t}t\bar{t}}$ signal as a function of both the number of jets and the number of b-tagged jets. The differences are small and they over covered by the $\mathrm{t\bar{t}}$ scale uncertainties. Therefore, no additional systematic uncertainties were considered necessary to cover these backgrounds.

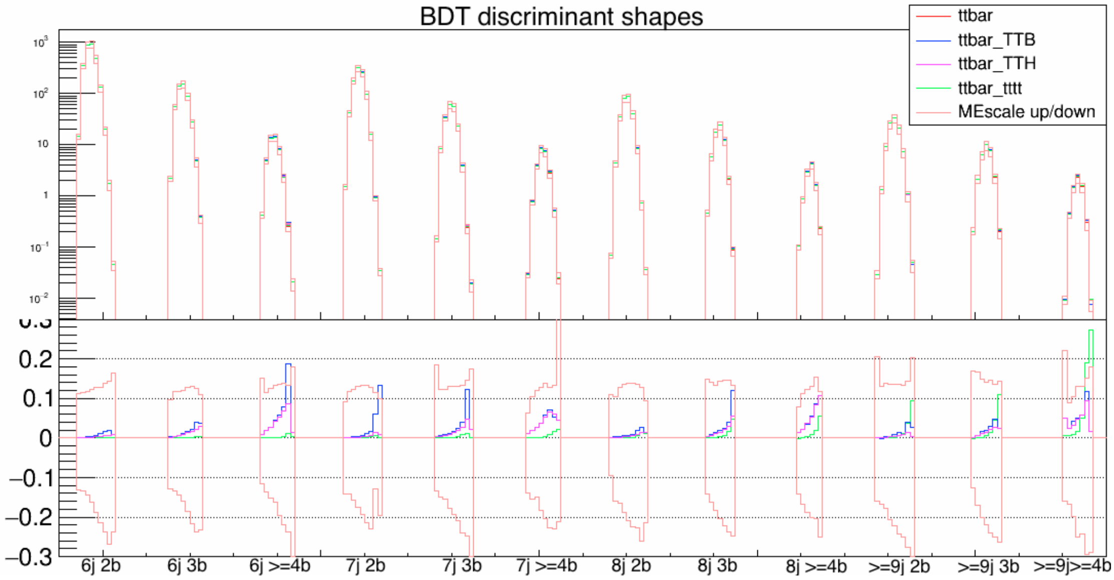


Figure E.2: BDT discriminator shapes for all categories, as indicated along the x axis. The ratio plot shows the difference between each distribution and the nominal $\mathrm{t\bar{t}}$ distribution divided by the $\mathrm{t\bar{t}}$ distribution.

E.3 Comparison of the Gradient Boost and AdaBoost boosting algorithms within the BDT

For this study, the following three BDTs were trained:

1. *GradNeg* - Gradient boosting taking into account negative weighting information in training and testing
2. *GradBoost* - Gradient boosting ignoring negative weighting information in training and testing

3. *AdaBoost* - AdaBoost boosting ignoring negative weighting information in training and testing

Each BDT was trained with the same set of input features and using the same sample of events to train and test the BDTs. The expected limits and uncertainties are shown for each strategy in Table D.1 for the μ + jets and e + jets final states. Note that this study was performed at an earlier stage in the analysis so the results to do not correspond exactly to the final expected limit given in Section 9.9. The BDT output discriminator distribution was only split into $\mathrm{N_{jets}}$ categories of of 6, 7, 8, 9+ jets at this stage rather than $\mathrm{N_{jets}}$ and $\mathrm{N^{M}_{tags}}$ categories.

Table E.1: Expected limits using jet categories of 6, 7, 8, 9+ jets for different BDT boosting algorithms.

Algorithm	μ + jets	uncertainty	e + jets	uncertainty
GradNeg	18.1	+8.0, -5.3	27.6	+12.9, -8.3
GradBoost	18.7	+8.3, -5.5	28.8	+12.9, -8.3
AdaBoost	10.7	+6.4, -4.0	21.6	+10.9, -7.0

It can be seen from Table D.1 that the difference between including negative weight information in the GradNeg strategy and not including it in the GradBoost strategy have a negligible effect on the expected limit within the uncertainties. Using negative weights may slightly optimise the modelling for training but not significantly hence the AdaBoost strategy can be used without negative weights as it has a significant benefit in lowering the expected limit.

E.4 Event-level BDT templates

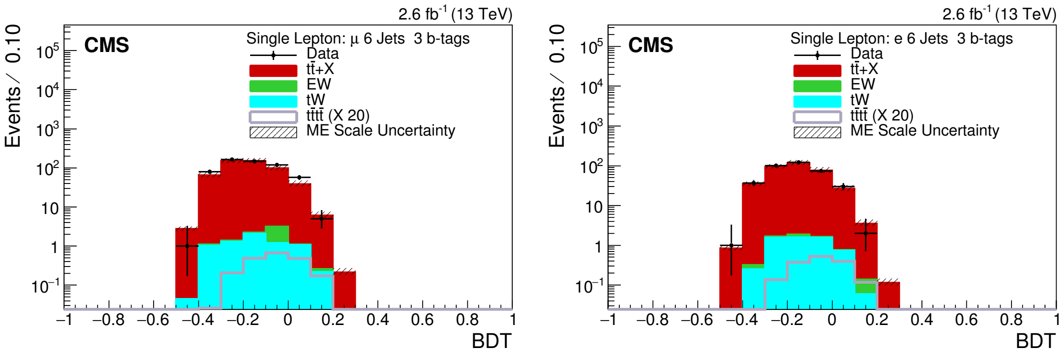


Figure E.3: The BDT output distributions for the Event-level BDT for data and simulation in the μ + jets channel (left) and e + jets channel (left) are shown for the 6 $\mathrm{N_{jets}}$ and $3\mathrm{N^{M}_{tags}}$ category.

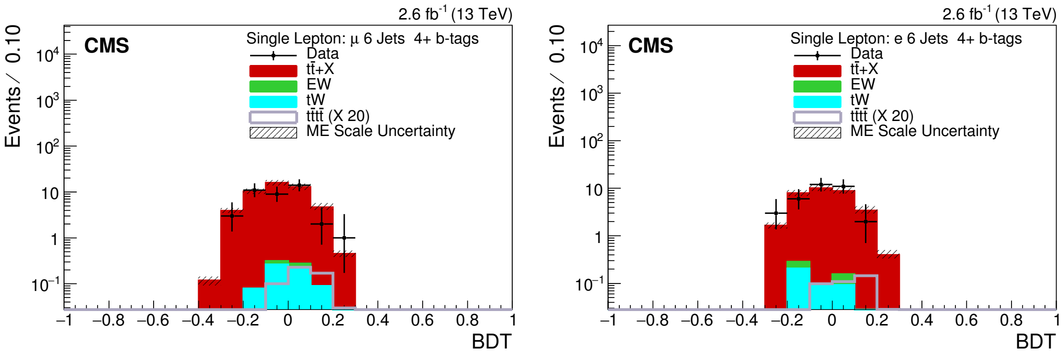


Figure E.4: The BDT output distributions for the Event-level BDT for data and simulation in the μ + jets channel (left) and e + jets channel (right) are shown for the 6 $\mathrm{N_{jets}}$ and ≥ 4 $\mathrm{N^{M}_{tags}}$ category.

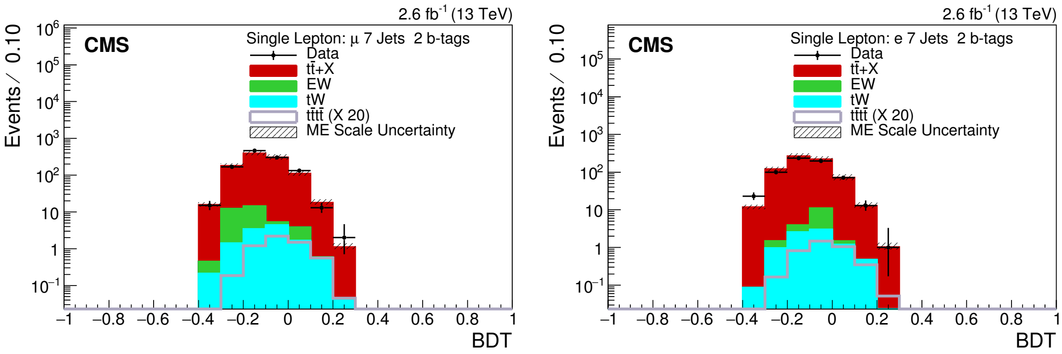


Figure E.5: The BDT output distributions for the Event-level BDT for data and simulation in the μ + jets channel (left) and e + jets channel (left) are shown for the 7 $\mathrm{N_{jets}}$ and 2 $\mathrm{N^{M}_{tags}}$ category.

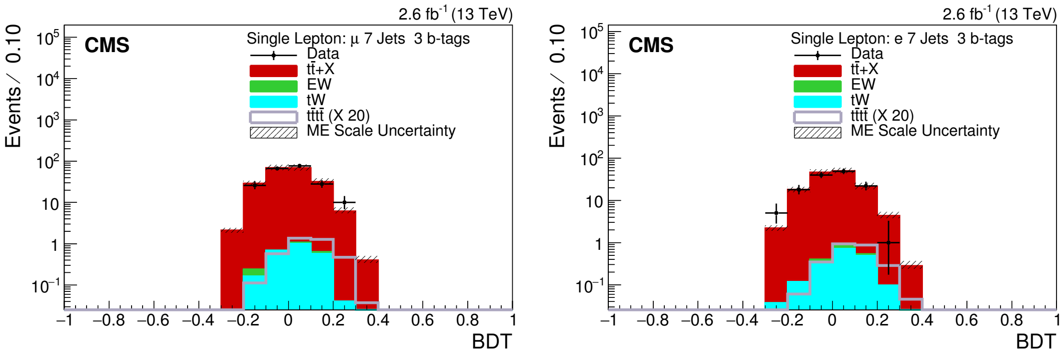


Figure E.6: The BDT output distributions for the Event-level BDT for data and simulation in the μ + jets channel (left) and e + jets channel (left) are shown for the 7 $\mathrm{N_{jets}}$ and 3 $\mathrm{N^{M}_{tags}}$ category.

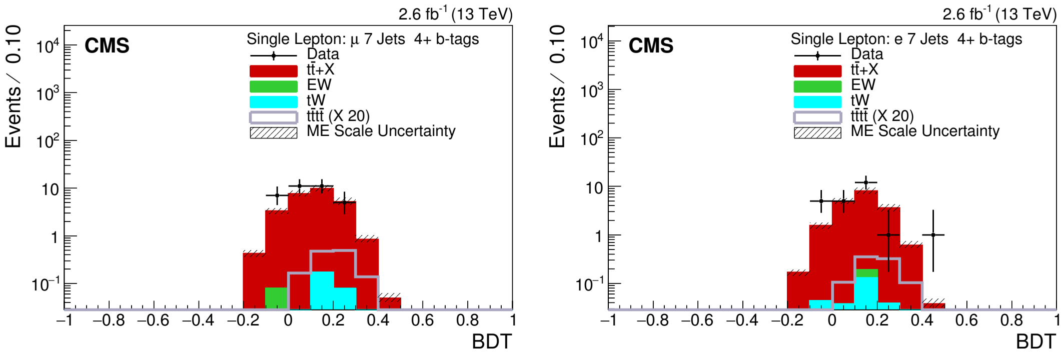


Figure E.7: The BDT output distributions for the Event-level BDT for data and simulation in the μ + jets channel (left) and e + jets channel (right) are shown for the 7 $\mathrm{N_{jets}}$ and $\geq$ 4 $\mathrm{N^{M}_{tags}}$ category.

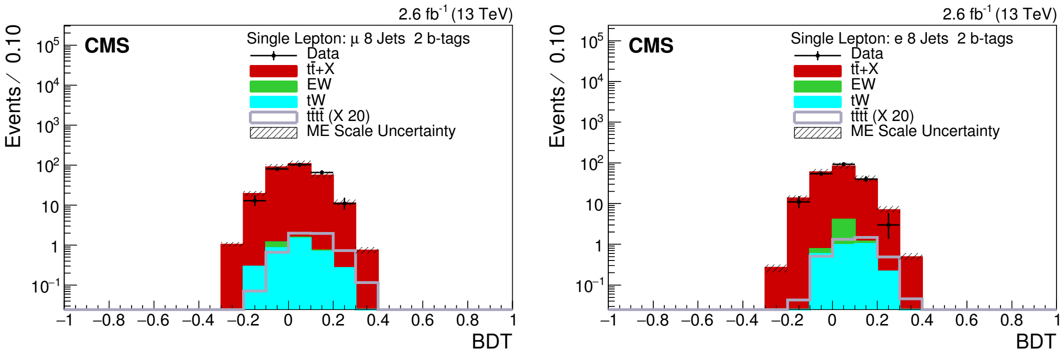


Figure E.8: The BDT output distributions for the Event-level BDT for data and simulation in the μ + jets channel (left) and e + jets channel (left) are shown for the 8 $\mathrm{N_{jets}}$ and 2 $\mathrm{N^{M}_{tags}}$ category.

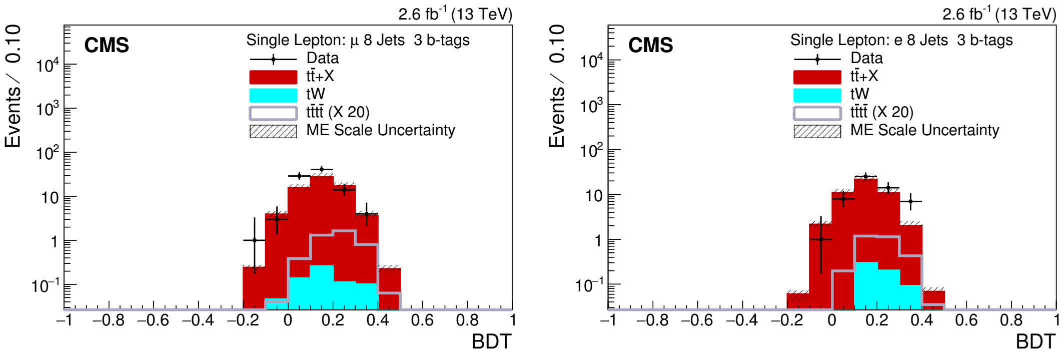


Figure E.9: The BDT output distributions for the Event-level BDT for data and simulation in the μ + jets channel (left) and e + jets channel (left) are shown for the 8 $\mathrm{N_{jets}}$ category and 3 $\mathrm{N^{M}_{tags}}$ category.

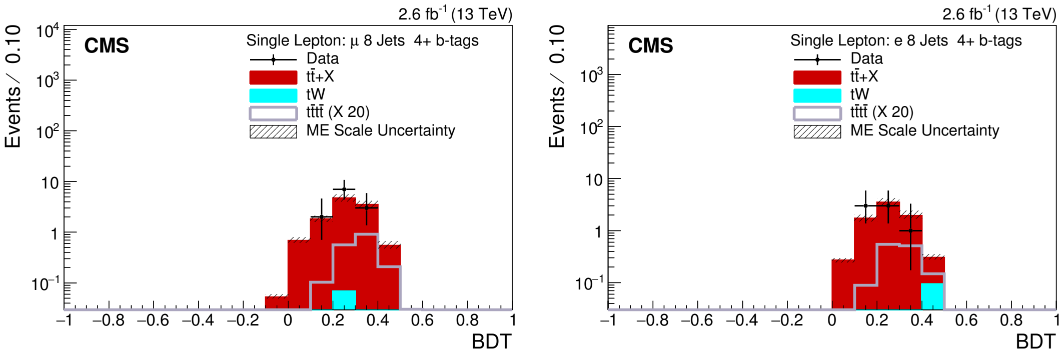


Figure E.10: The BDT output distributions for the Event-level BDT for data and simulation in the μ + jets channel (left) and e + jets channel (right) are shown for the 8 $\mathrm{N_{jets}}$ and ≥ 4 $\mathrm{N^{M}_{tags}}$ category.

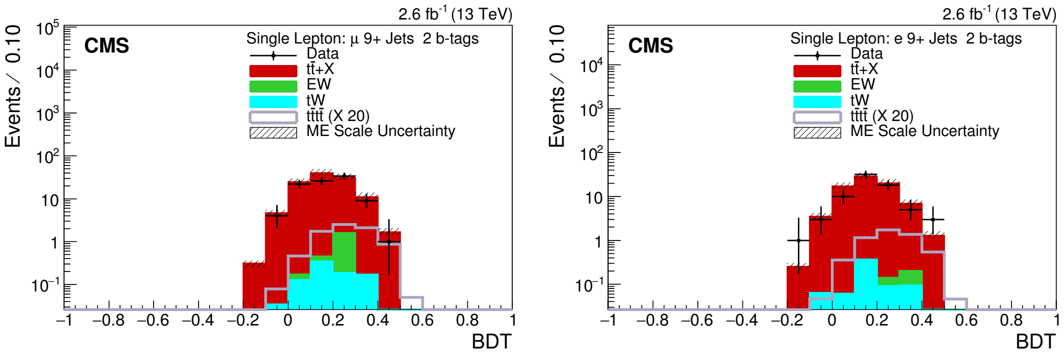


Figure E.11: The BDT output distributions for the Event-level BDT for data and simulation in the μ + jets channel (left) and e + jets channel (left) are shown for the ≥ 9 $\mathrm{N_{jets}}$ and 2 $\mathrm{N^{M}_{tags}}$ category.

E.5 Systematic shape studies

In this section the alternative BDT distribution shapes are examined for a few of the shape systematic described in Section 9.8. The largest systematic uncertainties are the JES, ME scale systematics and the $t\bar{t}$ generator choice. The JER and PU up/down (red/cyan) shapes deviate very little from the nominal distributions (blue). In $t\bar{t}$ there are several distributions including JER and PU systematics show relatively flat behaviour with respect to the nominal distribution and in future analyses could be considered for incorporation into a normalisation systematic uncertainty.

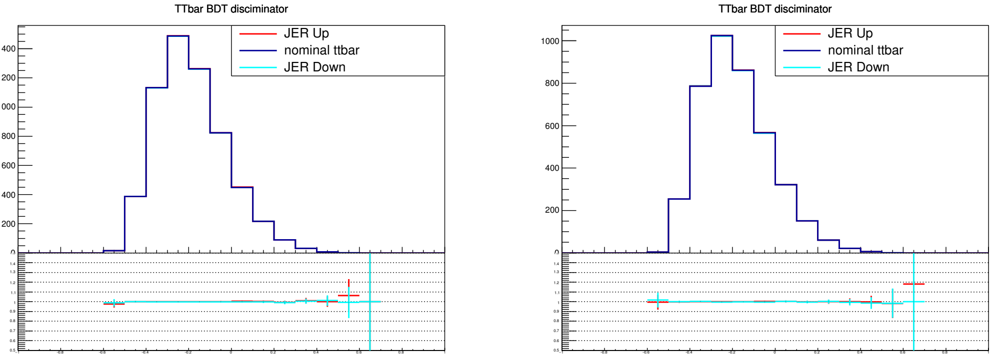


Figure E.12: The BDT shapes for JER systematic in $t\bar{t}$ for the μ + jets channel (left) and e + jets channel (right).

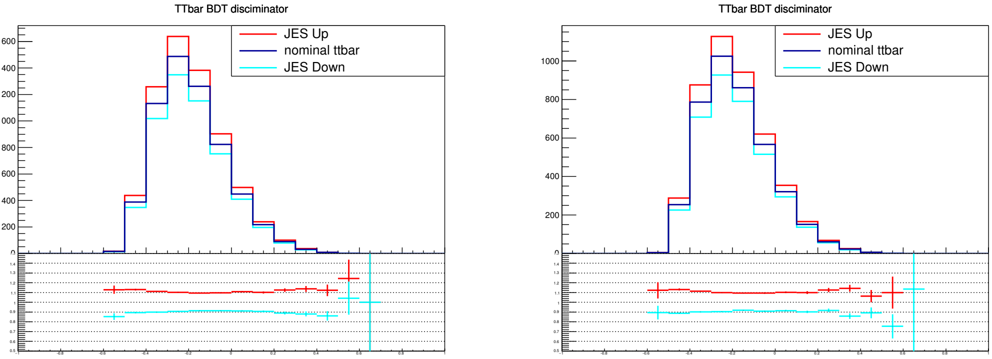


Figure E.13: The BDT shapes for JES systematic in $t\bar{t}$ for the μ + jets channel (left) and e + jets channel (right).

E.5.1 Studies of impact of systematic uncertainties

The impact of each systematic uncertainty on the expected limit is shown in Table D.2 by removing each systematic from the fit and recalculating the expected

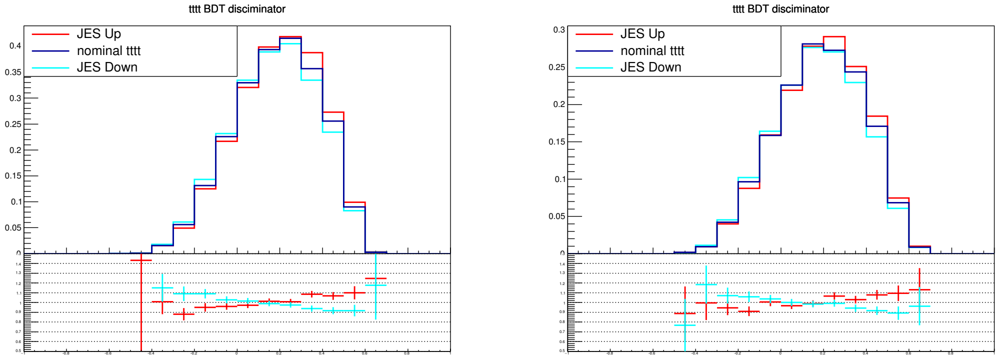


Figure E.14: The BDT shapes for JES systematic in $t\bar{t}t\bar{t}$ for the μ + jets channel (left) and e + jets channel (right).

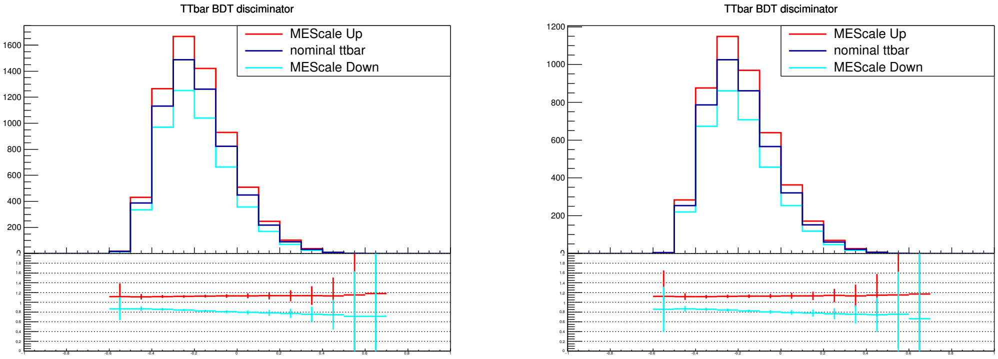


Figure E.15: The BDT shapes for ME scale systematic in $t\bar{t}$ for the μ + jets channel (left) and e + jets channel (right).

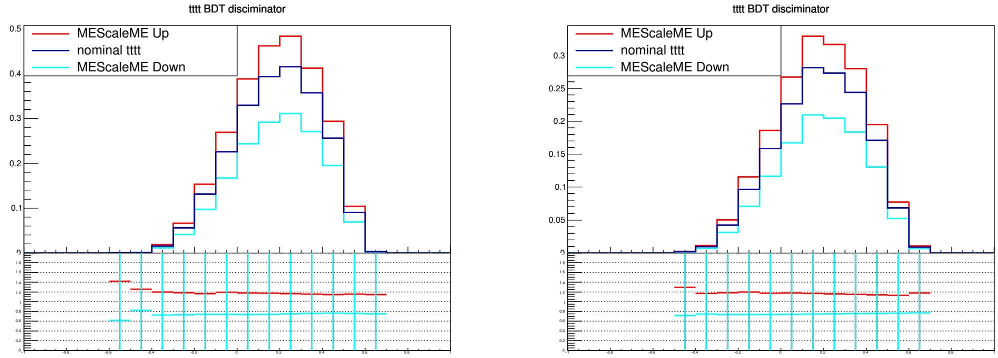


Figure E.16: The BDT shapes for ME scale systematic in $t\bar{t}t\bar{t}$ for the μ + jets channel (left) and e + jets channel (right).

limit. The systematic uncertainties which have the largest impact are the $t\bar{t}$ ME scale, $t\bar{t}$ PS scale and the JES scale, which is applied to both $t\bar{t}$ and $t\bar{t}t\bar{t}$. The $t\bar{t}$ ME scale has a large impact on the modelling of the signal process whereas the $t\bar{t}$ PS scale has a big effect on the modelling of the additional jets produced in high jet multiplicity $t\bar{t}$ events which pass the baseline event selection.

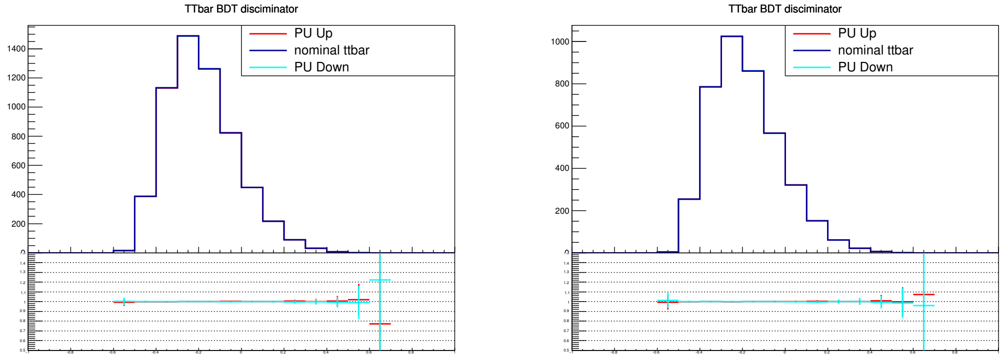


Figure E.17: The BDT shapes for PU systematic in $t\bar{t}$ for the μ + jets channel (left) and e + jets channel (right).

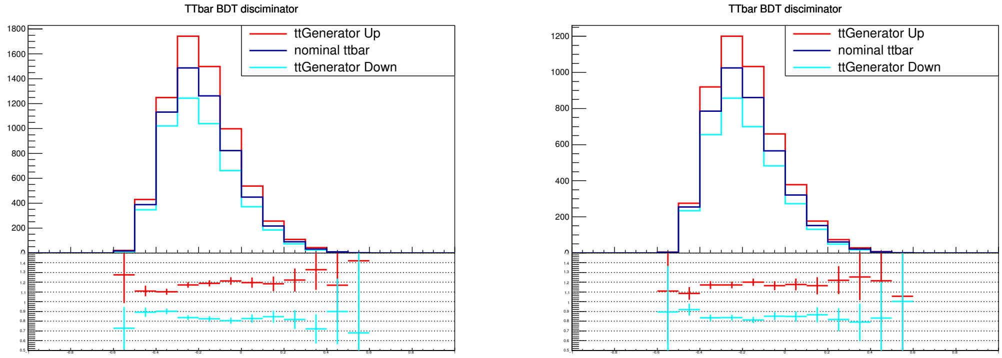


Figure E.18: The BDT shapes for ttGenerator choice systematic in $t\bar{t}$ for the μ + jets channel (left) and e + jets channel (right).

Table E.2: Expected limits on $\mathrm{t\bar{t}t\bar{t}}$ production which each systematic removed in turn

Systematic uncertainty removed	Expected limit ($\times\sigma^{SM}_{\mathrm{t\bar{t}t\bar{t}}}$)
None	16.0000
$\mathrm{t\bar{t}}$ ME scale	16.0625
$\mathrm{t\bar{t}t\bar{t}}$ ME scale	14.4375
JER	15.9375
JES	15.3125
PS scale	15.0625
PU	16.0625
Generator uncertainty	15.9531
$\mathrm{t\bar{t}}$ heavy flav	15.5625
Luminosity	16.0625
Lepton SF Mu	16.0625
Lepton SF El	15.9062
$\mathrm{t\bar{t}}$ norm	16.0000
$\mathrm{t\bar{t}t\bar{t}}$ norm	15.9062
EW norm	16.0000
Single top norm	16.0000
btagWeightCSVCFErr1	16.0625
btagWeightCSVCFErr2	16.0625
btagWeightCSVHF	15.5625
btagWeightCSVHFStats1	15.9375
btagWeightCSVHFStats2	15.9531
btagWeightCSVLF	15.9375
btagWeightCSVLFStats1	16.0625
btagWeightCSVLFStats2	15.9062

E.6 Correlation matrices for fit nuisance parameters

The correlation matrix for the fit nuisance parameters in the background only scenario can be see in Fig. D.19. There is some correlation between the various b-tagging scale factors and also a correlation between the heavy flavour $\sigma_{\mathrm{t\bar{t}b\bar{b}}}$ / $\sigma_{\mathrm{t\bar{t}jj}}$ modelling and the $\mathrm{t\bar{t}}$ ME scale, where the $\sigma_{\mathrm{t\bar{t}b\bar{b}}}$ / $\sigma_{\mathrm{t\bar{t}jj}}$ is expected to be related to the choice of ME scale.

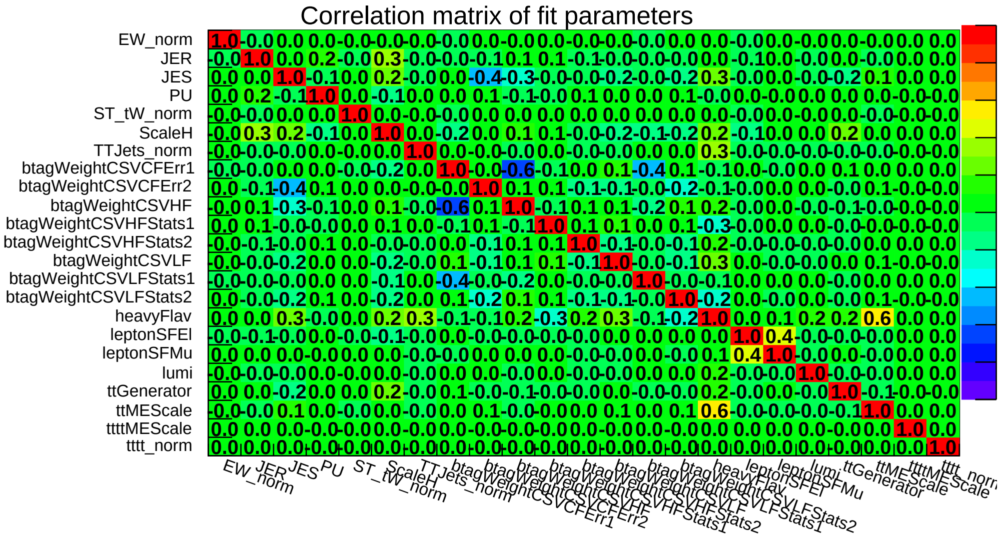


Figure E.19: The correlation matrices for background only for the fit parameters.

References

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

[2] Steven Weinberg. A model of leptons. *Phys. Rev. Lett.*, 19:1264–1266, Nov 1967.

[3] Abdus Salam. Weak and Electromagnetic Interactions. *Conf. Proc.*, C680519:367–377, 1968.

[4] D. Griffiths. *Introduction to elementary particles.* Wiley-VCH, second edition, 2008.

[5] B. et al. Aharmim. Combined analysis of all three phases of solar neutrino data from the sudbury neutrino observatory. *Phys. Rev. C*, 88:025501, Aug 2013.

[6] Particle Data Group, C. Patrignani, et al. Review of particle physics. *Chin. Phys. C*, 40:100001, 2016.

[7] The CMS Collaboration. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”, 2012. arxiv:1207.7235.

[8] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012.

[9] Georges Aad et al. Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments. *Phys. Rev. Lett.*, 114:191803, 2015.

[10] E. Noether. Invariant variation problems. *Transport Theory and Statistical Physics*, 1:186–207, January 1971.

[11] T. D. Lee and C. N. Yang. Question of parity conservation in weak interactions. *Phys. Rev.*, 104:254–258, Oct 1956.

[12] C. S. Wu, E. Ambler, R. W. Hayward, D. D. Hoppes, and R. P. Hudson. Experimental test of parity conservation in beta decay. *Phys. Rev.*, 105:1413–1415, Feb 1957.

[13] J. W. Cronin. The discovery of CP violation. *The European Physical Journal H*, 36(4):487–508, 2012.

[14] J. H. Christenson, J. W. Cronin, V. L. Fitch, and R. Turlay. Evidence for the 2π decay of the k_2^0 meson. *Phys. Rev. Lett.*, 13:138–140, Jul 1964.

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

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

[17] A. Pich. The Standard model of electroweak interactions. In *2004 European School of High-Energy Physics, Sant Feliu de Guixols, Spain, 30 May - 12 June 2004*, pages 1–48, 2005.

[18] David J. Gross and Frank Wilczek. Ultraviolet behavior of non-abelian gauge theories. *Phys. Rev. Lett.*, 30:1343–1346, Jun 1973.

[19] H. David Politzer. Reliable perturbative results for strong interactions? *Phys. Rev. Lett.*, 30:1346–1349, Jun 1973.

[20] Richard D. et al. Ball. Parton distributions for the LHC run II. *Journal of High Energy Physics*, 2015(4):40, 2015.

[21] Makoto Kobayashi and Toshihide Maskawa. CP Violation in the Renormalizable Theory of Weak Interaction. *Prog. Theor. Phys.*, 49:652–657, 1973.

[22] H. Albrecht et al. Observation of B0 - $\bar{B}$0 Mixing. *Phys. Lett.*, B192:245–252, 1987. [,51(1987)].

[23] F. Lehner and S. Faverot-Spengler, editors. *ARGUS Fest, 20 years of B meson mixing 1987-2007. Proceedings, ARGUS-symposium, DESY, Hamburg, Germany, November 9, 2007*, 2007.

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

[25] S. Abachi et al. Observation of the top quark. *Phys. Rev. Lett.*, 74:2632–2637, 1995.

[26] CMS Collaboration. Measurement of the t-channel single-top-quark production cross section and of the |Vtb| CKM matrix element in pp collisions at sqrt(s) = 8 TeV. *ArXiv e-prints*, March 2014.

[27] https://www-d0.fnal.gov/Run2Physics/top/top_public_web_pages/top_feynman_diagrams.html. Top physics diagrams by d0. Accessed: 16-12-14.

[28] Fedor Bezrukov and Mikhail Shaposhnikov. Why should we care about the top quark Yukawa coupling? *J. Exp. Theor. Phys.*, 120:335–343, 2015. [Zh. Eksp. Teor. Fiz.147,389(2015)].

[29] Fabian Kohn. *Measurement of the charge asymmetry in top quark pair production in pp collision data at $\sqrt{s} = 7$ TeV using the ATLAS detector.* PhD thesis, U. Gottingen (main), 2012.

[30] Kevin Lannon, Fabrizio Margaroli, and Chris Neu. Measurements of the Production, Decay and Properties of the Top Quark: A Review. *Eur. Phys. J.*, C72:2120, 2012.

[31] E. Boos and L. Dudko. The Single Top Quark Physics. *Int. J. Mod. Phys.*, A27:1230026, 2012.

[32] Qing-Hong Cao, Shao-Long Chen, and Yandong Liu. Probing Higgs Width and Top Quark Yukawa Coupling from $t\bar{t}H$ and $t\bar{t}t\bar{t}$ Productions. 2016.

[33] A. O. Sushkov, W. J. Kim, D. A. R. Dalvit, and S. K. Lamoreaux. New experimental limits on non-newtonian forces in the micrometer range. *Phys. Rev. Lett.*, 107:171101, Oct 2011.

[34] D. Keppel and P. Ajith. Constraining the mass of the graviton using coalescing black-hole binaries. *Phys. Rev. D*, 82:122001, Dec 2010.

[35] Michael Dine and Alexander Kusenko. Origin of the matter-antimatter asymmetry. *Rev. Mod. Phys.*, 76:1–30, Dec 2003.

[36] Leonard Susskind. Dynamics of spontaneous symmetry breaking in the weinberg-salam theory. *Phys. Rev. D*, 20:2619–2625, Nov 1979.

[37] V. C. Rubin and W. K. Ford, Jr. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *The Astrophysical Journal*, 159:379, feb 1970.

[38] Louise Volders. Neutral hydrogen in M 33 and M 101. *Bulletin of the Astronomical Institutes of the Netherlands*, 14:323, sep 1959.

[39] Chanda J. Jog. Large scale asymmetry of rotation curves in lopsided spiral galaxies. *Astron. Astrophys.*, 391:471, 2002.

[40] Massimo Persic, Paolo Salucci, and Fulvio Stel. The Universal rotation curve of spiral galaxies: 1. The Dark matter connection. *Mon. Not. Roy. Astron. Soc.*, 281:27, 1996.

[41] A. Einstein. Lens-Like Action of a Star by the Deviation of Light in the Gravitational Field. *Science*, 84:506–507, dec 1936.

[42] Richard S. Ellis. Gravitational lensing: a unique probe of dark matter and dark energy. *Philos Trans A Math Phys Eng Sci*, Mar 2010. 20123743[pmid].

[43] NASA, ESA, M. Postman (STScI) and the CLASH team. Galaxy Cluster MACS J1206.2-0847.

[44] Daniele Alves et al. and the LHC New Physics Working Group. Simplified models for lhc new physics searches. *Journal of Physics G: Nuclear and Particle Physics*, 39(10):105005, 2012.

[45] Andrea De Simone and Thomas Jacques. Simplified models vs. effective field theory approaches in dark matter searches. *Eur. Phys. J.*, C76(7):367, 2016.

[46] Kunal Kumar, Tim M.P. Tait, and Roberto Vega-Morales. Manifestations of top compositeness at colliders. *Journal of High Energy Physics*, 2009(05):022, 2009.

[47] Ben Lillie, Jing Shu, and Timothy M. P. Tait. Top Compositeness at the Tevatron and LHC. *JHEP*, 04:087, 2008.

[48] Chiara et al. Arina. A comprehensive approach to dark matter studies: exploration of simplified top-philic models. *Journal of High Energy Physics*, 2016(11):111, 2016.

[49] W. Bernreuther, P. Galler, C. Mellein, Z. G. Si, and P. Uwer. Production of heavy Higgs bosons and decay into top quarks at the LHC. *Phys. Rev.*, D93(3):034032, 2016.

[50] Giacomo Cacciapaglia, Haiying Cai, Aldo Deandrea, Thomas Flacke, Seung J. Lee, and Alberto Parolini. Composite scalars at the lhc: the higgs, the sextet and the octet. *Journal of High Energy Physics*, 2015(11):201, 2015.

[51] Samuel Calvet, Benjamin Fuks, Philippe Gris, and Loic Valery. Searching for sgluons in multitop events at a center-of-mass energy of 8 TeV. *JHEP*, 04:043, 2013.

[52] Pierre Fayet. Fermi-Bose Hypersymmetry. *Nucl.Phys.*, B113:135, 1976.

[53] Luis Alvarez-Gaume and S.F. Hassan. Introduction to S duality in N=2 supersymmetric gauge theories: A Pedagogical review of the work of Seiberg and Witten. *Fortsch.Phys.*, 45:159–236, 1997.

[54] Abdus Salam and J.A. Strathdee. Supersymmetry and Fermion Number Conservation. *Nucl.Phys.*, B87:85, 1975.

[55] Pierre Fayet. Supergauge Invariant Extension of the Higgs Mechanism and a Model for the electron and Its Neutrino. *Nucl.Phys.*, B90:104–124, 1975.

[56] Graham D. Kribs, Erich Poppitz, and Neal Weiner. Flavor in supersymmetry with an extended R-symmetry. *Phys.Rev.*, D78:055010, 2008.

[57] Georges Aad et al. Search for production of vector-like quark pairs and of four top quarks in the lepton-plus-jets final state in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *JHEP*, 08:105, 2015.

[58] Can Kilic, Takemichi Okui, and Raman Sundrum. Vectorlike Confinement at the LHC. *JHEP*, 02:018, 2010.

[59] Gustavo Burdman, Bogdan A. Dobrescu, and Eduardo Ponton. Resonances from two universal extra dimensions. *Phys. Rev.*, D74:075008, 2006.

[60] The CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004, 2008.

[61] Fabienne Marcastel. CERN's Accelerator Complex. La chaîne des accélérateurs du CERN. Oct 2013. General Photo.

[62] CMS Luminosity - Public Results.

[63] G. Aad et al. Observation of a new particle in the search for the standard model higgs boson with the ATLAS detector at the LHC. *Physics Letters B*, 716(1):1 – 29, 2012.

[64] S. Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B*, 716(1):30 – 61, 2012.

[65] Tai Sakuma and Thomas McCauley. Detector and Event Visualization with SketchUp at the CMS Experiment. *Journal of Physics: Conference Series*, 513(2):022032, 2014.

[66] V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, T. Bergauer, M. Dragicevic, J. Erö, and etal. CMS tracking performance results from early LHC operation. *The European Physical Journal C*, 70(4):1165–1192, 2010.

[67] The CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation*, 3(08):S08004, 2008.

[68] P Adzic. Energy resolution of the barrel of the CMS Electromagnetic Calorimeter. *Journal of Instrumentation*, 2(04):P04004, 2007.

[69] Bora Isildak. *Measurement of the differential dijet production cross section in proton-proton collisions at* $\sqrt{s} = 7$ *tev*. PhD thesis, Bogazici U., 2011.

[70] Florian Beaudette. The CMS Particle Flow Algorithm. In *Proceedings, International Conference on Calorimetry for the High Energy Frontier (CHEF 2013): Paris, France, April 22-25, 2013*, pages 295–304, 2013.

[71] Min Suk Kim et al. CMS reconstruction improvement for the muon tracking by the RPC chambers. *PoS*, RPC2012:045, 2012. [JINST8,T03001(2013)].

[72] L Tuura, A Meyer, I Segoni, and G Della Ricca. CMS data quality monitoring: Systems and experiences. *Journal of Physics: Conference Series*, 219(7):072020, 2010.

[73] The CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the CMS tracker. *Journal of Instrumentation*, 9(10):P10009, 2014.

[74] W Adam, R Frühwirth, A Strandlie, and T Todorov. “Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC”. *Journal of Physics G: Nuclear and Particle Physics*, 31(9):N9, 2005.

[75] A. Caner, S. Banerjee, A. Khanov, and N. Stepanov. Track and vertex finding performance with the CMS inner tracker. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 435(1âĂŞ2):118 – 143, 1999.

[76] https://twiki.cern.ch/twiki/bin/view/CMSPublic/SWGuideMuonId. Muon scale factor twiki for run one. Accessed: 2016-12-01.

[77] https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffsRun2. Muon scale factor twiki for run two. Accessed: 2016-08-12.

[78] Vardan Khachatryan et al. Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV. *JINST*, 10(06):P06005, 2015.

[79] S. Khalil and T. J. Kim. Electron object in top analysis in run 1. Accessed: 2016-12-01.

[80] https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2. Cut based electron id for run 2. Accessed: 2016-12-01.

[81] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. The Anti-k(t) jet clustering algorithm. *JHEP*, 04:063, 2008.

[82] Gavin P. Salam. Towards jetography. *The European Physical Journal C*, 67(3):637–686, 2010.

[83] Vardan Khachatryan et al. Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV. *Submitted to: JINST*, 2016.

[84] CMS Collaboration. Commissioning of the particle-flow event reconstruction with the first LHC collisions recorded in the CMS detector. CMS Physics Analysis Summary CMS-PAS-PFT-10-001, 2010.

[85] C. M. S. Collaboration and LHCb Collaboration. Observation of the rare $b_s^0 \to \mu^+\mu^-$ decay from the combined analysis of cms and lhcb data. *Nature*, 522(7554):68–72, Jun 2015. Letter.

[86] Serguei Chatrchyan et al. Identification of b-quark jets with the CMS experiment. *JINST*, 8:P04013, 2013.

[87] Identification of b quark jets at the CMS Experiment in the LHC Run 2. Technical Report CMS-PAS-BTV-15-001, CERN, Geneva, 2016.

[88] CMS Collaboration. Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector. 2016.

[89] 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.

[90] Stefan HÃűche. Introduction to parton-shower event generators. In *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014) Boulder, Colorado, June 2-27, 2014*, 2014.

[91] Céline Degrande, Benjamin Fuks, Valentin Hirschi, Josselin Proudom, and Hua-Sheng Shao. Automated next-to-leading order predictions for new physics at the LHC: the case of colored scalar pair production. *Phys.Rev.*, D91(9):094005, 2015.

[92] Johan Alwall, Michel Herquet, Fabio Maltoni, Olivier Mattelaer, and Tim Stelzer. Madgraph 5: going beyond. *Journal of High Energy Physics*, 2011(6):128, 2011.

[93] Stefano Frixione, Paolo Nason, and Carlo Oleari. "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", 2007. arxiv:0709.2092.

[94] Carlo Oleari. The POWHEG-BOX. *Nucl. Phys. Proc. Suppl.*, 205-206:36–41, 2010.

[95] Torbjörn Sjöstrand, Patrik Edén, Christer Friberg, Leif Lönnblad, Gabriela Miu, Stephen Mrenna, and Emanuel Norrbin. "High-energy-physics event generation with Pythia 6.1", 2001.

[96] Michał Czakon, Paul Fiedler, and Alexander Mitov. Total top-quark pair-production cross section at hadron colliders through $\mathcal{O}(\alpha_S^4)$. *Phys. Rev. Lett.*, 110:252004, Jun 2013.

[97] MichałCzakon and Alexander Mitov. Top++: A program for the calculation of the top-pair cross-section at hadron colliders. *Computer Physics Communications*, 185(11):2930 – 2938, 2014.

[98] J. Alwall, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, O. Mattelaer, H.-S. Shao, T. Stelzer, P. Torrielli, and M. Zaro. The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations. *Journal of High Energy Physics*, 2014(7):79, 2014.

[99] G. Bevilacqua and M. Worek. Constraining bsm physics at the lhc: four top final states with nlo accuracy in perturbative qcd. *Journal of High Energy Physics*, 2012(7), 2012.

[100] Aldo Deandrea and Nicolas Deutschmann. Multi-tops at the LHC. *JHEP*, 08:134, 2014.

[101] Measurement of the inclusive and differential $\mathrm{t\bar{t}}$ production cross sections in lepton + jets final states at 13 TeV. Technical Report CMS-PAS-TOP-16-008, CERN, Geneva, 2016.

[102] https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsCombined. Summaries of CMS cross section measurements. Accessed: 2016-12-03.

[103] Measuring Electron Efficiencies at CMS with Early Data. 2008.

[104] The CMS Collaboration. Calibration of the Combined Secondary Vertex b-Tagging discriminant using dileptonic $\mathrm{t\bar{t}}$ and Drell-Yan events. Technical Report CMS-NOTE-2013-130, CERN, Geneva, 2013.

[105] Vardan Khachatryan et al. Measurement of the cross section ratio $\sigma_{\mathrm{t\bar{t}b\bar{b}}}/\sigma_{\mathrm{t\bar{t}jj}}$ in pp collisions at $\sqrt{s} = 8$ TeV. *Phys. Lett.*, B746:132–153, 2015.

[106] Measurement of the cross section ratio $\mathrm{t\bar{t}}$ $+\mathrm{b\bar{b}}/\mathrm{t\bar{t}}$ $+\mathrm{jj}$ using dilepton final states in pp collisions at 13 TeV. Technical Report CMS-PAS-TOP-16-010, CERN, Geneva, 2016.

[107] Vardan Khachatryan et al. Measurement of $\mathrm{t\bar{t}}$ production with additional jet activity, including b quark jets, in the dilepton decay channel using pp collisions at $\sqrt{s} = 8$ TeV. *Eur. Phys. J.*, C76(7):379, 2016.

[108] A. et al. Hoecker. TMVA - Toolkit for Multivariate Data Analysis. *ArXiv Physics e-prints*, March 2007.

[109] Search for standard model production of four top quarks in proton-proton collisions at 13 TeV. Technical Report CMS-PAS-TOP-16-016, CERN, Geneva, 2016.

[110] Lorenzo Moneta, Kevin Belasco, Kyle S. Cranmer, S. Kreiss, Alfio Lazzaro, Danilo Piparo, Gregory Schott, Wouter Verkerke, and Matthias Wolf. The RooStats Project. *PoS*, ACAT2010:057, 2010.

[111] Kyle Cranmer. Practical Statistics for the LHC. In *Proceedings, 2011 European School of High-Energy Physics (ESHEP 2011): Cheile Gradistei, Romania, September 7-20, 2011*, pages 267–308, 2015. [247(2015)].

[112] J. S. Conway. Incorporating Nuisance Parameters in Likelihoods for Multisource Spectra. In *Proceedings, PHYSTAT 2011 Workshop on Statistical*

Issues Related to Discovery Claims in Search Experiments and Unfolding, CERN,Geneva, Switzerland 17-20 January 2011, 2011.

[113] The LHC Higgs Combination Group The ATLAS Collaboration, The CMS Collaboration. Procedure for the LHC Higgs boson search combination in Summer 2011. Technical Report CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, Aug 2011.

[114] G. Cowan, K. Cranmer, E. Gross, and O. Vitells. Asymptotic formulae for likelihood-based tests of new physics. *Eur. Phys. J. C.*, 71:1554, 2011.

[115] Vernon Barger, Wai-Yee Keung, and Brian Yencho. Triple-top signal of new physics at the LHC. *Physics Letters B*, 687(1):70 – 74, 2010.

[116] CMS Collaboration. CMS Luminosity Based on Pixel Cluster Counting - Summer 2012 Update. Technical Report CMS-PAS-LUM-12-001, CERN, Geneva, 2012.

[117] A.L. Read. Presentation of search results: the CL_s technique. *J. Phys. G.*, 28:2693, 2002.

[118] Junk T. Confidence level computation for combining searches with small statistics. *Nucl. Instrum. Meth. A.*, 434:435, 1999.

[119] Thomas Junk. Confidence level computation for combining searches with small statistics. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 434(2âĂŞ3):435 – 443, 1999.

[120] A L Read. Presentation of search results: the cl s technique. *Journal of Physics G: Nuclear and Particle Physics*, 28(10):2693, 2002.

[121] Search for standard model four top quark production at 8 TeV in the lepton + jets channel. Technical Report CMS-PAS-TOP-13-012, CERN, Geneva, 2013.

[122] Thomas Müller, Jochen Ott, and Jeannine Wagner-Kuhr. "Theta - a framework for template-based statistical modeling and inference", 2010. http://www-ekp.physik.uni-karlsruhe.de/õtt/theta/html/index.html.

[123] S. et al. Chatrchyan. Search for heavy quarks decaying into a top quark and a W or Z boson using lepton + jets events in pp collisions at $\sqrt{s} = 7$ TeV. *Journal of High Energy Physics*, 2013(1):154, 2013.

[124] Serguei et al. Chatrchyan. Search for heavy quarks decaying into a top quark and a W or Z boson using lepton + jets events in pp collisions at $\sqrt{s} = 7$ TeV. *J. High Energy Phys.*, 01(arXiv:1210.7471. CMS-B2G-12-004. CERN-PH-EP-2012-309):154. 29 p, Oct 2012. Comments: Submitted to the Journal of High Energy Physics.

[125] Serguei Chatrchyan et al. Search for new physics in events with same-sign dileptons and jets in pp collisions at $\sqrt{s} = 8$ TeV. *JHEP*, 1401:163, 2014.

[126] Georges Aad et al. Analysis of events with b-jets and a pair of leptons of the same charge in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector. *JHEP*, 10:150, 2015.

[127] Lana Beck, Freya Blekman, Didar Dobur, Benjamin Fuks, James Keaveney, and Kentarou Mawatari. Probing top-philic sgluons with LHC Run I data. *Physics Letters B*, 746:48 – 52, 2015.

[128] Dorival Goncalves-Netto, David Lopez-Val, Kentarou Mawatari, Tilman Plehn, and Ioan Wigmore. Sgluon Pair Production to Next-to-Leading Order. *Phys.Rev.*, D85:114024, 2012.

[129] Eric Conte, Benjamin Fuks, and Guillaume Serret. Madanalysis 5, a user-friendly framework for collider phenomenology. *Computer Physics Communications*, 184(1):222 – 256, 2013.

[130] https://twiki.cern.ch/twiki/bin/view/CMS/EgammaIDRecipesRun2# Electron_efficiencies_and_scale. Electron scale factor twiki. Accessed: 2016-08-12.

[131] Yoav Freund and Robert E Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1):119 – 139, 1997.

[132] Llew Mason, Jonathan Baxter, Peter Bartlett, and Marcus Frean. Boosting algorithms as gradient descent in function space. NIPS, 1999.

[133] CMS Luminosity Measurement for the 2015 Data Taking Period. Technical Report CMS-PAS-LUM-15-001, CERN, Geneva, 2016.

[134] Search for $\mathrm{t\bar{t}H}$ production in the $\mathrm{H} \to \mathrm{b\bar{b}}$ decay channel with $\sqrt{s} = 13$ TeV pp collisions at the CMS experiment. Technical Report CMS-PAS-HIG-16-004, CERN, Geneva, 2016.

[135] Vardan Khachatryan et al. Search for new physics in same-sign dilepton events in proton-proton collisions at $\sqrt{s} = 13$ TeV. 2016.

[136] Search for production of vector-like top quark pairs and of four top quarks in the lepton-plus-jets final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Technical Report ATLAS-CONF-2016-013, CERN, Geneva, Mar 2016.

[137] Search for four-top-quark production in final states with one charged lepton and multiple jets using 3.2 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC. Technical Report ATLAS-CONF-2016-020, CERN, Geneva, Apr 2016.

[138] Search for new physics using events with b-jets and a pair of same charge leptons in 3.2 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Technical Report ATLAS-CONF-2016-032, CERN, Geneva, Jun 2016.

[139] Search for new phenomena in $t\bar{t}$ final states with additional heavy-flavour jets in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector. Technical Report ATLAS-CONF-2016-104, CERN, Geneva, Sep 2016.

[140] https://twiki.cern.ch/twiki/bin/view/CMSPublic/PhysicsResultsTOP. Top physics results twiki. Accessed: 2016-12-17.

[141] Martin Goerner. Top p_{T} reweighing wiki. Technical Report , 2013.

[142] C. Hof A Meyer H. Pieta P. Biallass, T. Hebbeker. Parton distribution uncertainty determination within cmssw. Technical Report AN-2009-48, CMS, 2009.

[143] V. Khachatryan, A. M. Sirunyan, A. Tumasyan, W. Adam, T. Bergauer, M. Dragicevic, and etal. Charged particle multiplicities in pp interactions at $\sqrt{s} = 0.9$, 2.36, and 7 TeV. *Journal of High Energy Physics*, 2011(1):79, 2011.

[144] Rick Field. Min-Bias and the Underlying Event at the LHC. *Acta Phys. Polon.*, B42:2631–2656, 2011.