Measurement of single top t-channel production cross-section and top quark mass in single top events using CMS detector

A Thesis
Submitted to the
Tata Institute of Fundamental Research, Mumbai
for the degree of Doctor of Philosophy
in High Energy Physics

by
Soureek Mitra

School of Natural Sciences
Tata Institute of Fundamental Research,
Mumbai
September, 2017
Final Version Submitted in December, 2017

Declaration

This thesis is a presentation of my original research work. Wherever contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature, and acknowledgement of collaborative research and discussions.

The work was done under the guidance of Professor Tariq Aziz, at the Tata Institute of Fundamental Research, Mumbai.

Soureek Mitra

Soureek Mitra

In my capacity as supervisor of the candidates thesis, I certify that the above statements are true to the best of my knowledge.

Tariq Aziz

Date: 18/12/2017

Dedication

To my family and all tax-payers of India

Acknowledgement

Science in general and especially physics has intrigued me right from my childhood. The opportunity to choose research in physics as a career has been a "dream-come-true" for me. On the verge of submitting my Ph.D. dissertation marking the end of apprentice period, my mind gazes back in time to acknowledge the crucial efforts of the people surrounding me which has helped/guided me in the meandering paths of life till now. My heartiest apologies if I have missed anyone here.

First and foremost, I would like to thank my parents for giving highest priority to my choice in every step of my career, not weighing me with their own unfulfilled aspirations, for sacrificing plenty in their life so that I could succeed. I also cordially thank my elder brother, who has been more like a friend, always giving me right suggestions whenever I have been in doubt.

I am deeply thankful to my supervisor Prof. Tariq Aziz for his continuous support, positive attitude and friendly behaviour. His constructive criticism has helped me to improve and blossom in research. I enjoyed engrossing discussions with him, covering a larger dimension beyond my thesis topic on various aspects of physics and experimental particle physics, economics, the poetry of Mirza Ghalib etc. He has always been there offering his help in different situations regardless of academic and non-academic nature of them. His presence has been highly enjoyable and I would definitely keep in touch with him in future. I thank my co-supervisor Prof. Gagan Mohanty for his constant enthusiasm, enlightening suggestions on the physics analyses and his help on various academic and non-academic matters. He has always tried to keep my spirits high for doing research by sharing his own unique experiences. I thank my collaborators Dr. Alberto Maria Orso Iorio, Dr. Thorsten Chwalek, Dr. Nadijeh Jafari, Prof. Andrea Giammanco for healthy physics discussions and their guidance to handle different technical issues. I would also like to thank Prof. Gobinda Majumder, Prof. Maonoranjan Guchait and other faculty members of Department of High Energy Physics (DHEP) for many insightful discussions. I would also like to thank Brij, Puneet, Prashant for being the backbone behind the remarkable performance of the Tier-2 computing cluster without which physics analysis would not have been possible. I cordially thank our departmental secretaries Mrs. Minal Rane and Divekarji for their help in many logistic issues.

I joined TIFR as an integrated PhD. student after finishing my college and was enrolled in DHEP. I take this opportunity to recall some of the excellent teachings during my graduate course work by Prof. Shiraz Minawala, Prof. Amol Dighe, Prof. Sunil Mukhi, Prof. Sreerup Raychaudhuri, Prof. Sandip Trivedi, Prof. Mustansir Burma, Prof. Mandar Deshmukh, Prof. Vaibhav Prabhudesai. Their insights and way of teaching have helped me immensely in the transition towards research from the regular curriculum.

Research in High Energy Physics demands collaborative work. I thank our colleagues at the CERN accelerator departments for ensuring the outstanding performance of the LHC as well as the technical and administrative staff at CERN, also the other CMS institutes for their contributions to the successful data taking with the CMS detector.

Friends have been very special in my life at TIFR. I take this opportunity to cordially thank Nairit, Atreyee, Debjyoti, Ritam, Shubhadeep, Khadiza, Rickmoy without whom life at TIFR would not have been so charming. From whole-night *adda* to watching late night movie shows in and around Colaba; from organizing sports tournaments, especially football, Saraswati puja (even with a living idol), terrace party and even an election, to the late night snacks discussion-turned-into crazy debates on atheism, sex vs gender, Federer's 19th, world politics; from relishing *Parsi* delicacies to dipping into the ice-cream from "Natural's", from crazy nights in "Gokul" to late night coffee table banters at McDonald's or CCD; from *diwali* celebrations in the campus to pandal hopping across Mumbai during *Durga Puja*; we have done it all together. Everything happened so naturally and spontaneously. You guys hold a special place in my heart. I also thank my seniors: Sayan Chakraborti, Shamasish Sengupta, Sanmay Ganguly, Gourab Chatterjee, Nilay Kundu, Saikat Nandi, Gouranga Kole for the sheer fun in their company, for enriching our taste buds with cuisines of different varieties, for generously explaining tough assignments, for all the chocolates and sweets. I thank my flatmates in Wadala residence, Shubhrangshu, Ramkrishna for all the fun we had there together. I thank my fellow departmental students both senior and junior to me namely Dipan Sengupta, Kolahal Bhattacharya, Saranya Ghosh, Rajdeep Mohan Chatterjee, Saurabh Shandilya, Sanjeev, Suman, Soham, Pallabi, Aravind (popularly known as *Holur*), Bibhu, Jacky, Muzamil, Varghese, Debashshis, Ravindra (the "student"), Mintu, Uttiya and post-docs Sandhya, Depanwita, Ravindra (the "post-doc"), Vipin for all the chit-chats during afternoon tea sessions, for engrossing discussions on the subject of particle physics specially on the statistical aspect of it. I have to specially mention Tamali Bhattacharya Sikder, a doctor by profession, whose presence kept us in a jovial mood besides in good health, outside our office/lab hours. I thank the elite group of gossip-mongers comprising of Shamik, Manibrata, Moniruzzaman, Toushik, Sudip, Indranil (popularly known as *Pocha*), for keeping me updated about munching campus gossips. Football has been a good part of my life outside the office at TIFR. I thank Chandradoy, Sumeru, Krishnendu Gope, Vivekda, Abhijit for all those wonderful memories on the field during the (departmental) tournaments through your tricks and skills and for all the related fun from outside. I would love to see this continue with the same enthusiasm in future and would love my departmental team snatch back the trophy from the current champions.

I did spend a significant amount of time during my Ph.D at CERN, Geneva. There I met people not only from different ethnicity and culture but also from different backgrounds. Interactions and experiences with them have instilled a strong global citizen entity within me turning myself more accepting towards cultural, linguistic and social differences.

I also would like to acknowledge persons outside TIFR, a few of them whom I have known from childhood, namely Sutapa, Saptarshi, Srimanta, Sudipto, Arpan, Debapratim for their constant inspiration. They have been key figures for shaping me into the person I have become.

Finally, I am very much thankful to TIFR for providing such an excellent atmosphere within its campus and opportunities for research in a multi-national collaboration like CMS. In both cases, I got ample scope to improve on technical skill-set and knowledge as well as on a personal and social level.

Synopsis

Introduction

The top quark is the heaviest among all known elementary particles and has unique features in many ways. It decays ($\approx 10^{-25}$s) faster via electroweak interaction than typical hadronization time scale ($\approx 10^{-24}$s) implying that it can exist as a free quark before its decay. Due to this unique property, the top quark provides an excellent probe to signatures of physics beyond the Standard Model (SM) through comparison of measurements with precise SM predictions. Top quarks are produced copiously in the proton-proton (pp) collision at the Large Hadron Collider (LHC). Single top quarks are produced in pp collisions through charged-current interaction via exchange of a virtual W boson, unlike the dominant top pair ($t\bar{t}$) production process which occurs through strong interaction. Single top production at the LHC can be characterized via 3 modes: t-channel, tW-channel and s-channel, ordered according to decreasing cross section, which are shown in Figure 1. It is one of the important electroweak processes occurring at the LHC. It allows measurement of fundamental parameters of SM such as $|V_{tb}|$, top quark mass etc. It provides a direct probe to the tWb coupling structure for any beyond SM contribution such as presence of anomalous tensor couplings [1], flavor changing neutral current, fourth generation, new resonances like W′ [2], charged Higgs [3] etc. Differential measurements can also provide powerful inputs to constrain the b-quark parton distribution function (PDF) at high momentum fractions [4].

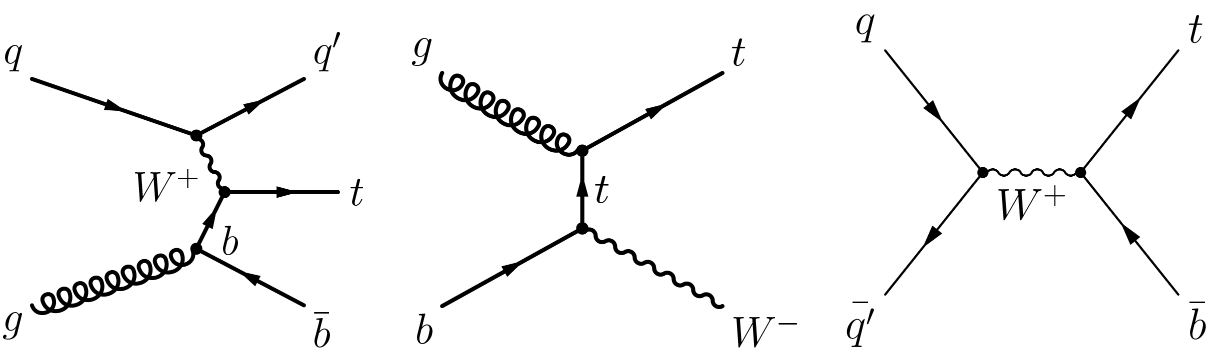


Figure 1: Feynman diagrams of electroweak single top production in the standard model: t-channel (left) , tW-channel (middle) and s-channel (right).

The t-channel is the most abundant single top production mode in pp collisions at the LHC, amounting to almost 80% of the total cross section. The striking feature of this process is the presence of a light quark recoiling against the top, which decays to a b-quark and a W boson. This study has focused on the muon final state arising from direct and cascade (via τ leptons) decays of the W boson originating from the top decay. So the final state consists of an isolated high transverse momentum (p_T) muon, two jets due to hadronization of the light quark and b-quark and missing energy due to the escaping neutrino. The next-to-leading-order (NLO) calculation with HATHORv2.1 [5,6] predicts

$$\sigma_{t-ch,t} = 136.0^{+4.1}_{-2.9}(\text{scale}) \pm 3.5(\text{PDF} + \alpha_s)\text{pb}, \tag{1}$$

$$\sigma_{t-ch,\bar{t}} = 81.0^{+2.5}_{-1.7}(\text{scale}) \pm 3.2(\text{PDF} + \alpha_s)\text{pb}, \tag{2}$$

$$\sigma_{t-ch,t+\bar{t}} = 217.0^{+6.6}_{-4.6}(\text{scale}) \pm 6.2(\text{PDF} + \alpha_s)\text{pb} \tag{3}$$

for the t-channel production at $\sqrt{s} = 13$ TeV of a top quark, top anti-quark and the sum, respectively.

The top quark mass is an important parameter of the SM. Its precise measurement is of profound importance, both in theory and in experiment. It constitutes a major input to the global electroweak fit and radiative corrections to the mass of the Higgs boson, which is related to the stability of electroweak vacuum [7]. In experiment, it provides an essential benchmark to determine the calibration and performance of the detector and/or reconstruction algorithm. The latest world average of the top mass, based on measurements performed in $t\bar{t}$ events, from ATLAS, CMS , CDF & D0 is $m_{top} = 173.34 \pm 0.27(\text{stat}) \pm 0.71(\text{syst})$ GeV [8]. The single top process offers an independent statistical sample for this measurement. Single top events have different final state particle multiplicities compared to $t\bar{t}$ events, leading to different amount of uncertainties in the measurement.

CMS detector

The central feature of the Compact Muon Solenoid (CMS) detector is a superconducting solenoid of 6 m internal diameter, providing a constant magnetic field of 3.8 T inside. Within the solenoid volume are a silicon (Si) pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass (absorber) and scintillator (sensor) hadron calorimeter (HCAL). The Si tracking detectors measure tracks within $|\eta| \leq 2.5$. The ECAL and HCAL, each composed of a barrel and two endcap sections, covers upto $|\eta| \leq 3.0$. Forward calorimeters on either side of the interaction point extend the η-coverage provided by the barrel and endcap detectors and covers the region defined by $3.0 \leq |\eta| \leq 5.0$. Muons are measured in the range $|\eta| \leq 2.4$ using gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid. The detector provides almost 4π solid angle coverage allowing an accurate estimate of missing transverse energy. Detailed description of the CMS detector along with the coordinate system convention and relevant kinematic variables is presented in [9]. Information from all sub-detectors has been used in the measurements presented in this thesis.

Event reconstruction

The particle flow (PF) algorithm [10] is at the heart of reconstruction of all the particles produced in the final state of the event. This algorithm uses an optimal combination of information from various sub-detectors in CMS to reconstruct and identify individual particles. The PF candidates are classified as: electrons, muons, photons, charged and neutral hadrons. During the collisions of two proton bunches, multiple pp interactions occur, leading to multiple collision vertices. The primary vertex of the event is considered to be the one reconstructed with the largest sum of charged-track p_T^2 values and is required to lie within 24 cm in the direction along the beam axis and 2 cm in the direction transverse to the beam axis about the center of the detector. All other vertices except the primary vertex are considered as pile-up interactions. The muon candidates are reconstructed combining the information from both the silicon tracker and the muon spectrometer in a global fit. An identification is performed using the quality of the geometrical matching between the tracker and the muon system measurements. The transverse momentum of muons is obtained from the curvature of the corresponding tracks. The energy of photons is obtained from the ECAL measurement directly, corrected for zero-suppression effects [11]. The energy of electrons is estimated from a combination of the electron momentum at the primary interaction vertex determined by the tracker, the energy of the corresponding ECAL cluster, and the energy sum of all bremsstrahlung photons spatially compatible with the track originating from the electron. The energy of charged hadrons is determined from a combination of their momenta measured in the tracker and the matching ECAL and HCAL energy deposits, corrected for zero-suppression effects and for the response function of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is obtained from the corresponding corrected ECAL and HCAL energy. Jets are reconstructed by clustering PF particle candidates using the anti-k_T clustering algorithm [12] with a distance parameter of 0.4 in the $\eta - \phi$ space. A correction to account for pileup interactions is estimated on an event-by-event basis using the jet area method described in Ref. [13], and is applied to the reconstructed jet p_T . Further corrections to the jet energy are also applied as a function of jet p_T and η to account for any residual effects due to non-uniform responses of the detector. The missing transverse momentum vector $\overrightarrow{E\!\!\!/_T}$ is defined as the projection onto the plane perpendicular to the beam axis of the negative vector sum of the momenta of all reconstructed PF objects in an event. Its magnitude is referred to as $E\!\!\!/_T$.

Measurement of t-channel single top cross section at 13 TeV

Event selection and categorization

Events are selected using a dedicated trigger requiring an isolated muon with high p_T in the final state. A few additional criteria are also imposed in order to reject backgrounds.

- Events must have exactly one muon having $p_T > 22$ GeV within $|\eta| < 2.1$ and passing tight identification criteria [14].
- Relative isolation (I_{rel}) of the selected muon, defined in Eq.(4), where $I^{\text{charged hadron}}$, I^{γ}, $I^{\text{neutral hadron}}$ and $I^{\text{Pile-up}}$ are the scalar p_T sums of the charged hadrons, photons, neutral hadrons and particles associated with pile-up vertices respectively,

computed in a cone of $\Delta R = 0.4$ should be $< 6\%$, i.e., $\mathrm{I_{rel}} < 0.06$.

$$\mathrm{I_{rel}} = \frac{\mathrm{I^{charged\ hadron}} + \max((\mathrm{I}^{\gamma} + \mathrm{I^{neutral\ hadron}} - \mathrm{I^{Pile-up}}), 0)}{\mathrm{p}_{\mathrm{T},\mu}} \tag{4}$$

- Events with additional muons (electrons) with $\mathrm{p_T} > 10(20)$ GeV and within $|\eta| < 2.5$ passing loose identification and isolation criteria are rejected in order to reduce Z+jets and $\mathrm{t\bar{t}}$ di-lepton backgrounds

- Events are required to consist of exactly 2 jets having $\mathrm{p_T} > 40$ GeV within $|\eta| < 4.7$ which pass loose identification criteria [15] depending on the jet constituents.

- One of the 2 selected jets is required to pass a b-tagging discriminator tight criterion, which corresponds to 0.1% mistag rate and 40% average efficiency in the selected $\mathrm{p_T}$ range, to be identified as b-tagged within $|\eta| < 2.4$.

- The transverse W mass ($\mathrm{m_T^W}$) defined in Eq.(5) of the (μ, $\not{E}_\mathrm{T}$) system in the final state is required to be more than 50 GeV, i.e., $\mathrm{m_T^W} > 50$ GeV to reduce the QCD multijet background.

$$m_\mathrm{T}^\mathrm{W} = \sqrt{(p_{T,\mu} + \not{E}_\mathrm{T})^2 - (p_{x,\mu} + \not{p}_x)^2 - (p_{y,\mu} + \not{p}_y)^2} \tag{5}$$

The events are categorized in the bins of number of jets and number of b-tagged jets and the generic nomenclature "NjMt" is attributed to the events having N jets and M b-tagged jets. The 2j1t event category is the most signal enriched region. The 3j2t category is used to control the $\mathrm{t\bar{t}}$ background. The 2j0t region is used to validate the QCD multijet background estimation method.

Top quark reconstruction

The 4-momentum of the top quark is reconstructed from the available kinematic information in the event. The top-quark decays to a b-tagged jet, a charged lepton and a neutrino, whose transverse momentum can be inferred from $\not{E}_\mathrm{T}$. The longitudinal momentum of the neutrino, $\mathrm{p}_{z,\nu}$, is determined from other kinematic constraints such as the W boson mass, $\mathrm{m_W} = 80.4$ GeV [16]. Assuming energy-momentum conservation at the W$\rightarrow \mu\nu$ vertex, one can obtain Eq.(4.4) for $\mathrm{p}_{z,\nu}$. Two cases can arise for a solution, namely,

- A negative discriminant leads to complex solutions for $\mathrm{p}_{z,\nu}$. In this case, the imaginary part is eliminated by setting $\mathrm{m_T^W} = \mathrm{m_W}$, while still respecting the $\mathrm{m_W}$ constraint. This leads to the fact that the discriminant i.e., the square root term in Eq.(6) is null. This condition gives a quadratic relation between $\mathrm{p}_{x,\nu}$ and $\mathrm{p}_{y,\nu}$ with two possible solutions, and one remaining degree of freedom. The solution corresponding to the minimal distance between $\mathrm{p}_{\mathrm{T},\nu}$ and $\not{E}_\mathrm{T}$ is chosen.

- For events with a positive discriminant, the solution with the smallest absolute value is chosen [17, 18].

$$p_{z,\nu} = \frac{\Lambda p_{z,\mu}}{p_{\mathrm{T},\mu}^2} \pm \frac{1}{p_{\mathrm{T},\mu}^2}\sqrt{\Lambda^2 p_{z,\mu}^2 - p_{\mathrm{T},\mu}^2(\mathrm{E}_\mu^2 \not{E}_\mathrm{T}^2 - \Lambda^2)}, \text{where } \Lambda = \frac{\mathrm{m_W^2}}{2} + \overrightarrow{p_{\mathrm{T},\mu}} \cdot \overrightarrow{\not{E}_\mathrm{T}} \tag{6}$$

This implies that the 4-momentum of the W boson candidate can be completely determined.

In the 2j1t and 3j1t regions, the b-tagged jet is assumed to come from the top quark decay. In 3j1t region, the most forward jet is considered to be the jet stemming from the light quark recoiling against top. In 3j2t region, the jet with the largest b-tagging discriminant value is attributed to the b quark from the top decay while the non-b-tagged jet is identified to originate from the light quark. In 2j0t region, three cases are considered, namely,

- If both jets pass b-tagging algorithm but do not satisfy the b-tagging discriminator tight criterion, then the jet with higher b-tagging discriminant value is attributed to b-quark originating from top.
- If only one of the two jets passes b-tagging algorithm but does not satisfy the b-tagging discriminator tight criterion, then that jet is identified as the b-jet.
- In the case where both jets fail to pass the b-tagging algorithm, the more central jet is assigned as the b-jet.

The 4-momentum of the W boson is then added to the 4-momentum of the b-jet candidate to obtain the 4-momentum of the decaying top quark.

Measurement with early data (42 pb^{-1}) collected in 2015

Analysis strategy

The 2j1t region is divided into two segments by putting requirements on the reconstructed top quark mass. The window $130 < m_{top} < 225$ GeV is the signal region (SR) whereas the regions $m_{top} < 130$ GeV and $m_{top} > 225$ GeV provide side-bands (SB) to validate the W+jets background. The 3j2t region is used as control sample for $t\bar{t}$ background. The cross-section is extracted by performing a simultaneous fit to the light-quark jet pseudorapidity $|\eta_{j'}|$ distribution in the 2j1t SR and 3j2t regions [19].

Background estimation

- **QCD:** The production of multijets through QCD processes has a huge cross section. On the other hand, only a small fraction of these events mimic the lepton+jets final state of the applied event selection. The selection efficiency for QCD multijet events is therefore tiny. The combination of huge cross section and tiny selection efficiency would require generation of extremely large MC samples for this process in order to produce sufficient events that survive the event selection to guarantee a proper QCD modeling with decent statistics in the signal region. An alternative way of modeling is to define control regions in data that are enriched in QCD events (purity $\approx 95\%$) by inverting the muon isolation requirement ($I_{rel} > 0.12$) and estimate QCD contribution to the signal region by means of a likelihood fit to the full m_T^W distribution in data using QCD templates derived from the control region. The QCD yield in the $m_T^W > 50$ GeV is obtained by calculating the integral of the m_T^W -distribution from QCD, normalized to the fitr esult, for $m_T^W > 50$ GeV. The

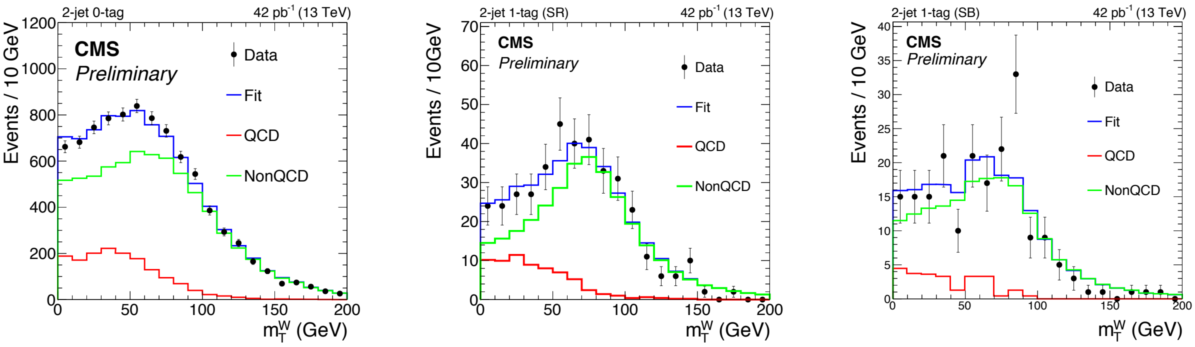


Figure 2: QCD estimation in 2j0t region (left), 2j1t SR (middle) and 2j1t SB (right).

method is validated in the 2j0t region, as the number of available events is larger compared to signal region and the final estimate is taken from the 2j1t region as shown in Figure 2. The QCD yield thus obtained is kept fixed during $|\eta_{j'}|$-fit.

- **W+jets:** Processes with W bosons produced in association with heavy flavor (b or c quark) jets constitute an important background to the 2j1t sample. The contribution from this background is independent of the reconstructed top-quark mass which indicates that the events inside and outside the m_{top} window have a similar W+heavy quark content. The $|\eta_{j'}|$ distribution in the side band is therefore used to determine the contribution from W+heavy quark events in the signal sample.

- **$t\bar{t}$:** The modeling of the $t\bar{t}$ background in simulation is validated using the 3j2t region. A simultaneous fit is performed including the 3j2t region to control $t\bar{t}$ normalization.

Results

The normalization scale factor (S_i) defined in Eq.(7) of each process category is allowed to float during fit,

$$S_i = \frac{N_i^{obs}}{N_i^{pred.}}, \tag{7}$$

where N_i^{obs} and $N_i^{pred.}$ are the number of events observed and predicted for the i-th process category, respectively.
Normalization scale factor for t-channel is allowed to float freely whereas top ($t\bar{t}$ + tW) and electroweak (W/Z + Jets) background normalizations are constrained by Gaussian priors with 20% and 30% uncertainties, respectively. The QCD multijet yield is kept fixed to the estimated value from m_T^W -fit. Uncertainty in the fit results due to QCD multijet background estimation is calculated separately by varying the QCD multijet normalization conservatively by 50% to account for effects due to difference of m_T^W -templates, trigger efficiency etc. in the signal and control regions. Figure 3 shows the post-fit distribution of the light quark jet $|\eta|$ in signal and control regions. The cross section of t-channel single top process is determined using Eq.(8),

$$\sigma_{t-ch} = \frac{N_s}{\epsilon \times \mathcal{B}(t \to b\ell\nu) \times \mathcal{L}} \tag{8}$$

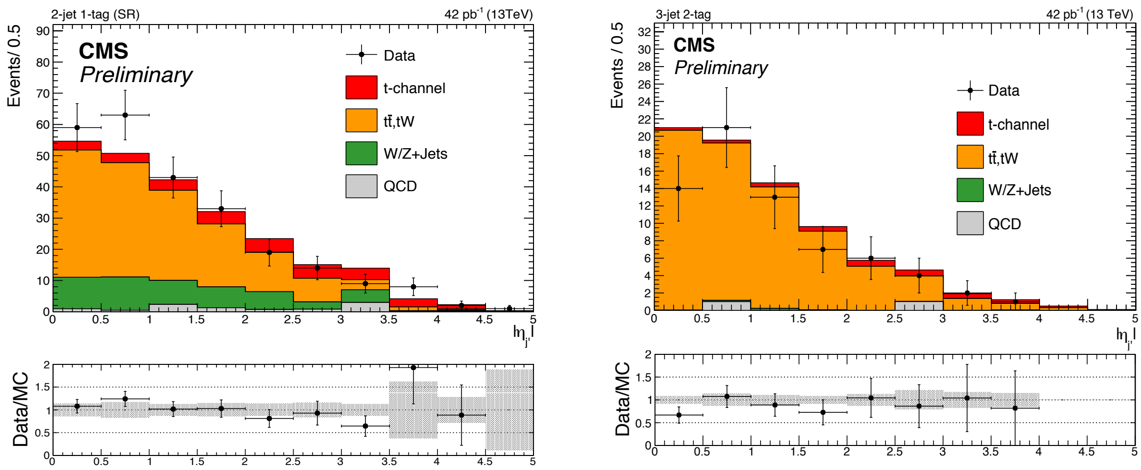


Figure 3: Post-fit $|\eta_{j'}|$ distribution for 2-jets-1-tag signal region (left) and 3-jets-2-tag region (right).

where N_s is the t-channel event yield after fit, ϵ is the selection efficiency, $\mathcal{L} = 42$ pb^{-1} is the integrated luminosity and $\mathcal{B}(\mathrm{t} \to \mathrm{b}\ell\nu) = 0.324$ is the total branching fraction of top decay into a b-quark and a lepton in the final state. The quantitative impact on the fit results from each of the uncertainty sources is evaluated by performing pseudo experiments with the corresponding variation. The individual uncertainty is taken from the difference between the nominal result of the fit and that from the varied pseudo experiment. The uncertainty in the measurement is mostly dominated by statistics (36%) since it is performed on such a small data sample; uncertainties due to jet energy scale (17%), b-tagging efficiency (5.6%), PDF (4.5%) are the dominant ones among the systematic sources.

The cross section of t-channel single top process using 42 pb^{-1} of pp collision data at 13 TeV is measured to be

$$\sigma_{\mathrm{t-ch}} = 274 \pm 98(\mathrm{stat.}) \pm 52(\mathrm{syst.}) \pm 33(\mathrm{lumi.})\,\mathrm{pb}, \tag{9}$$

in agreement with the SM prediction. In the approximation, $|V_{td}|, |V_{ts}| << |V_{tb}|$, the top quark decay branching fraction into Wb is considered to be close to 1 and the magnitude of the Cabibbo-Kobayashi-Maskawa (CKM) matrix element V_{tb} can be extracted as $|V_{tb}| = \sqrt{\frac{\sigma^{\mathrm{meas.}}_{\mathrm{t-ch}}}{\sigma^{\mathrm{theo.}}_{\mathrm{t-ch}}}}$, yielding

$$|V_{tb}| = 1.12 \pm 0.24(\mathrm{expt.}) \pm 0.02(\mathrm{theo.}) \tag{10}$$

Measurement with full data (2.3 fb^{-1}) collected in 2015

Analysis strategy

An advanced analysis technique compared to the early measurement has been used to exploit the full potential of available data (2.3 fb^{-1}) in order to make more precise measurement. The ratio ($R_{\mathrm{t-ch.}}$) of top to anti-top production cross sections is also determined directly from data in this analysis. A neural network based multivariate (MVA) discriminator is developed using the input variables listed in Table 1 to separate signal from background. The t-channel signal is trained against $t\bar{t}$ and W+jets backgrounds in the 2j1t region. The MVA response is also evaluated for events belonging to 3j1t and 3j2t category. Each event category is sub-divided into two segments based on the charge of the muon in the final state to distinguish between top and anti-top production. The cross-section is estimated by performing a simultaneous fit to the MVA response distribution in the 2j1t, 3j1t and 3j2t region [20].

Table 1: Input variables used in the neural network ranked according to their importance.

Rank	Description
1	Absolute value of the pseudorapidity of the light-quark jet
2	Invariant mass of the top quark reconstructed from muon, neutrino, and b-tagged jet
3	Invariant mass of the two selected jets
4	Transverse mass of the W boson
5	Scalar sum of the transverse momenta of the two jets
6	Cosine of the angle between the muon and the light-quark jet in the rest frame of the top quark
7	Invariant mass of the jet with the largest transverse momentum
8	ΔR between the momentum vectors of the light-quark jet and the b-tagged jet.
9	Transverse momentum of the light-quark jet
10	Invariant mass of the light-quark jet
11	Absolute value of the pseudorapidity of the reconstructed W boson

Background estimation

- **QCD:** QCD has been estimated using the same method as described in Section from control regions defined by inverting the muon isolation. Separate estimation of QCD yield has been made depending on the charge of the muon in the 2j1t signal region whereas 2j0t control region is used for validation of the estimation method as shown in Figure 4.

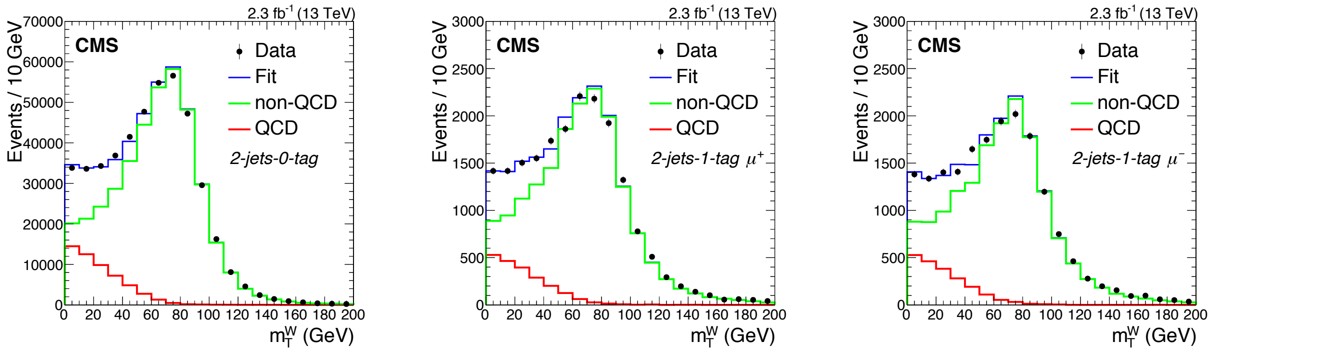


Figure 4: QCD estimation in 2j0t region (left) and in 2j1t region for μ^+ (middle) and μ^- (right) final states.

- **W+jets:** W+jets background is validated using 3j1t region, which has similar heavy flavor content. W+jets normalization is controlled by simultaneous fit to the MVA response including the 3j1t region.
- **$\mathrm{t\bar{t}}$:** The $\mathrm{t\bar{t}}$ normalization is controlled by including both 3j1t and 3j2t regions in the fit.

Results

A maximum likelihood fit is performed to the MVA response distribution as shown in Figure 5 with 2j1t, 3j1t and 3j2t events simultaneously to extract the cross section and $\mathrm{R_{t-ch.}}$. The free parameters of the fit are the normalization scale factors of the single top quark production and background processes as defined in Eq.(7) and $\mathrm{R_{t-ch.}}$. The background scale factors are constrained by log-normal priors with an uncertainty of 10% for the top quark background, 30% for the electroweak background, and 50% for the QCD multijet background. The t-channel normalization is allowed to float unconstrained. The fit is performed using the Barlow-Beeston [21] method which correctly accounts for limited-size simulation samples. The cross section is then determined from Eq.(8) as described in Section . The measurement of the cross section is affected by various sources of systematic uncertainties, which can be grouped into two categories, experimental uncertainties and theoretical uncertainties. Several of the former category such as uncertainties due to jet energy scale (JES), jet energy resolution (JER), b-tagging and mis-tagging efficiency, muon reconstruction etc. are considered as nuisance parameters in the fit to the MVA discriminator distribution and are thus included in the total uncertainty of the fit. Pseudo experiments are performed to determine the impact of the sources of the remaining uncertainties coming from theory such as signal modeling, factorization/renormalization ($\mu_\mathrm{F}/\mu_\mathrm{R}$) scale of signal and background events, PDF etc. The uncertainty in this measurement are dominated by the theoretical uncertainties such as generator modeling of signal (±8.2%) and $\mathrm{t\bar{t}}$ (±4.5%), $\mu_\mathrm{F}/\mu_\mathrm{R}$ scale of t-channel signal (-7.2/+5.1%) and $\mathrm{t\bar{t}}$ (-4.7/+3.1%) etc.
The cross section for the production of single top quarks and the top quark to top antiquark cross section ratio as a result of the fit are

$$\sigma_{\mathrm{t-ch,t}} = 150 \pm 8(\mathrm{stat.}) \pm 9(\mathrm{expt.}) \pm 18(\mathrm{theo.}) \pm 4(\mathrm{lumi.})\ \mathrm{pb} = 150 \pm 22\ \mathrm{pb}, \qquad (11)$$

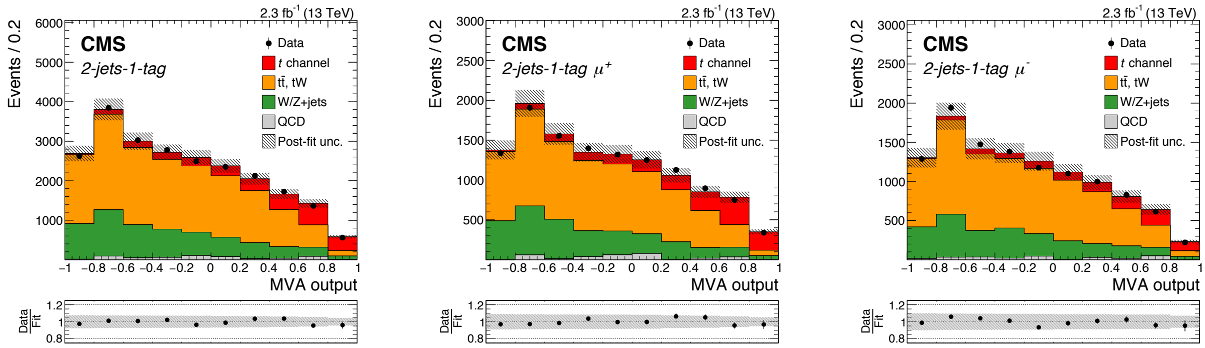


Figure 5: Neural network distributions for all (left), μ^+ (middle), and μ^- (right) normalized to the yields obtained from the simultaneous fit in the 2j1t region.

$$R_{\mathrm{t-ch.}} = 1.81 \pm 0.18(\mathrm{stat.}) \pm 0.15(\mathrm{syst.}) \tag{12}$$

A comparison between the measured ratio and the prediction of different PDF sets is shown in the left plot of Figure 6. Using the $\sigma_{\mathrm{t-ch.,t}}$ and $\mathrm{R}_{\mathrm{t-ch.}}$ measurements, the cross section of the top anti-quark production is computed as

$$\sigma_{\mathrm{t-ch,\bar{t}}} = 82 \pm 10(\mathrm{stat.}) \pm 4(\mathrm{expt.}) \pm 11(\mathrm{theo.}) \pm 2(\mathrm{lumi.})\ \mathrm{pb} = 82 \pm 16\ \mathrm{pb}, \tag{13}$$

where the uncertainties are evaluated using the correlation matrix of the simultaneous fit. This leads to the total cross section,

$$\sigma_{\mathrm{t-ch,t+\bar{t}}} = 232 \pm 13(\mathrm{stat.}) \pm 12(\mathrm{expt.}) \pm 26(\mathrm{theo.}) \pm 6(\mathrm{lumi.})\ \mathrm{pb} = 232 \pm 31\ \mathrm{pb} \tag{14}$$

The comparison of this measurement with the SM expectation and measurements of the single top quark t-channel cross section at other centre-of-mass energies is shown in the right plot of Figure 6. The total cross section is used to determine the absolute value of the CKM matrix element $|V_{tb}|$ defined as $|f_{LV}V_{tb}| = \sqrt{\frac{\sigma^{meas.}_{\mathrm{t-ch,t+\bar{t}}}}{\sigma^{theo.}_{\mathrm{t-ch,t+\bar{t}}}}}$ with the assumptions described in Section .

$$|f_{LV}V_{tb}| = 1.03 \pm 0.07(\mathrm{expt.}) \pm 0.02(\mathrm{theo.}) \tag{15}$$

The possible presence of an anomalous Wtb coupling is taken into account by an anomalous form factor f_{LV} [1], which is 1 in the SM and deviates from 1 for physics beyond the SM. The $\sigma^{\mathrm{theo.}}_{\mathrm{t-ch,t+\bar{t}}}$ is the SM prediction from Eq.(3) which assumes $|V_{tb}| = 1$.

Measurement of top quark mass in single top events at 13 TeV

Single top events provide an independent statistical sample to measure the mass of top quark. The top quark mass measurement in single top topology enhances the range of available measurements, with systematic uncertainties partially uncorrelated from those considered in $\mathrm{t\bar{t}}$ production. Due to low multiplicity of particles in the final state compared to $\mathrm{t\bar{t}}$, the chances of high uncertainty in top quark 4-momentum reconstruction

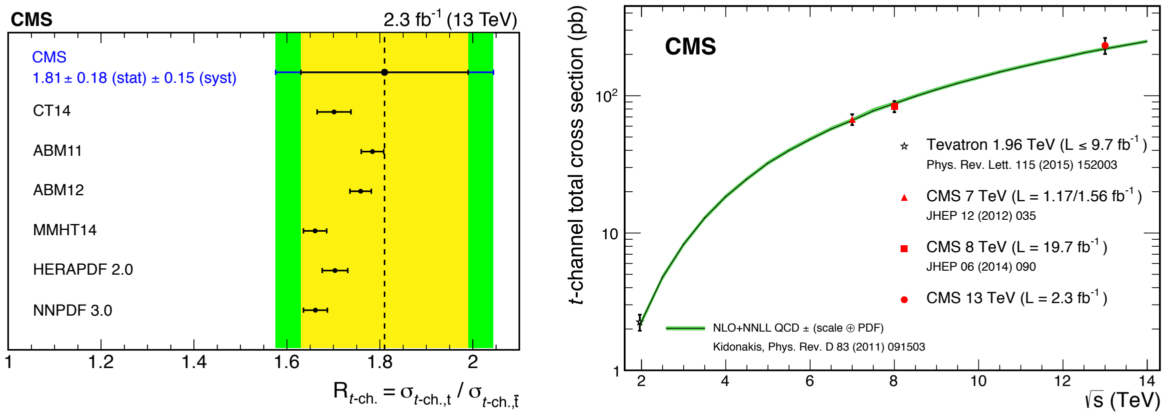


Figure 6: Comparison of measured value of $R_{t-ch.}$ with predictions from different PDF sets [22], [23], [24], [25], [26] (left) and summary of the most precise CMS measurements [27], [28] along with Tevatron combination [29] for the total t-channel single top quark cross section, in comparison with NLO+NNLL QCD calculations [30] (right).

due to cross-talk among the final state particles are relatively low. This essentially lowers the uncertainty occuring due to colour reconnection effect, which is one of the dominant source of systematics in top mass measurements using $t\bar{t}$ events in the lepton+jets final state [31,32]. An event selection designed for t-channel single top enriched events is considered for this analysis since t-channel is the most abundant single top production mode. Precise measurement of the top-quark-mass in single top events demands a high purity selection. An MVA analysis is performed on top of the baseline event selection criteria as described in Section to ensure high signal purity. The top mass (m_{top}) is determined from the reconstructed top quark 4-momentum using the method described in Section . The analysis is ongoing in a blinded mode in order to avoid any bias. A detail study of robustness of the analysis technique is currently in progress using pseudo experiments before "opening the box".

Multivariate analysis

An MVA discriminator is designed in order to discriminate signal from backgrounds. The signal and background distributions of the variables listed in Table 2 are given as inputs to a boosted decision tree (BDT). The input variables are chosen such that the BDT response has low correlation with the reconstructed top mass. The t-channel single top process is trained against $t\bar{t}$ and W+jets backgrounds with additionally QCD multijet events mixed in. Relative weights are assigned to signal and backgrounds based on their purity in the training sample. The simulated events in the 2j1t region are divided randomly into two independent samples, one to train the BDT and the other to test its response. The BDT response distribution from training and test samples are shown on the left plot in Figure 7. Probabilities for signal-like and background-like events estimated from training are then applied to events in data.

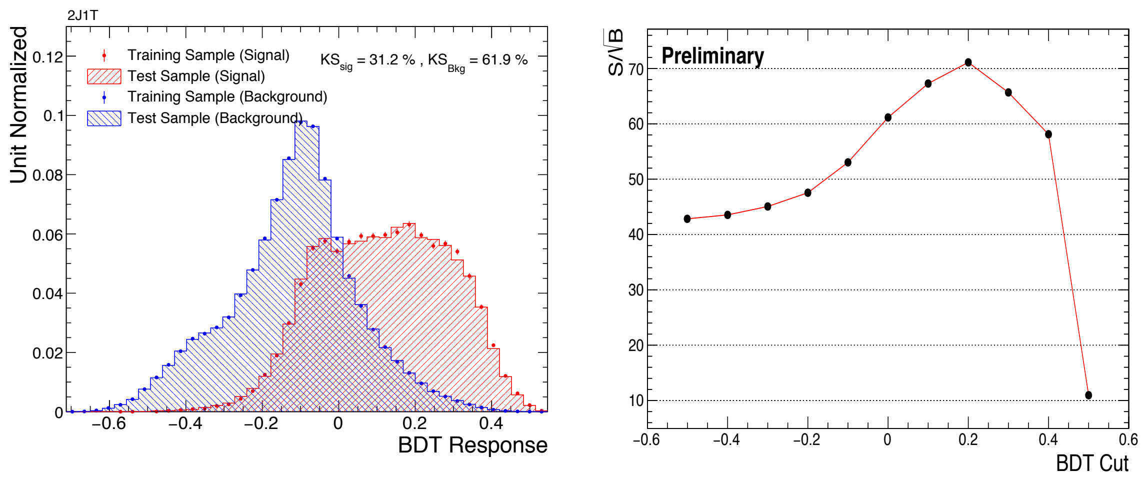


Figure 7: BDT response distribution in training and test sample (left) and significance study with BDT response (right).

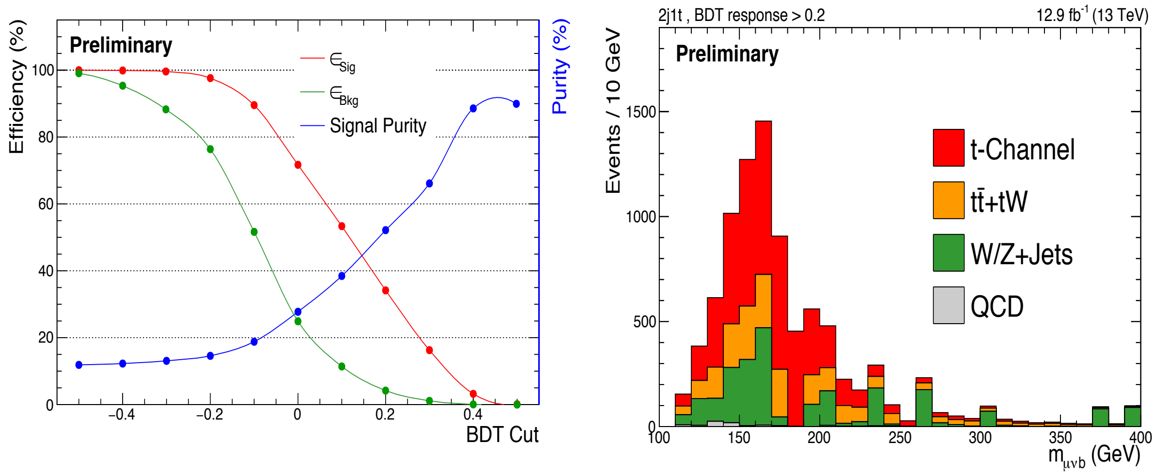


Figure 8: Study of Signal and background efficiencies and signal purity with cut on BDT response (left) and $m_{b\mu\nu}$ distribution after requiring BDT response > 0.2 (right).

Table 2: Input variables used in the BDT ranked according to their separation power.

Rank	Description
1	ΔR between the momentum vectors of the light-quark jet and the b-tagged jet.
2	Absolute value of the pseudorapidity of the light-quark jet
3	Invariant mass of the two selected jets
4	Cosine of the angle between the muon and the light-quark jet in the rest frame of the top quark
5	Transverse mass of the W boson
6	ΔR between the momentum vectors of the muon and b-tagged jet
7	Scalar sum of the transverse momenta of the two jets
8	Absolute value of the pseudorapidity of the muon

Optimization of the cut on MVA discriminator

An optimized cut on the BDT response is applied in order to reject backgrounds and obtain a sample with high signal purity to make a precise measurement. The optimization of the BDT response cut is performed by studying the signal significance shown on the right plot of Figure 7, along with signal and background efficiency and signal purity shown on the left plot of Figure 8. A cut of BDT response > 0.2 is chosen to ensure maximum signal significance, relatively high signal purity and efficiency and very low background efficiency. This cut implies the ratio (S/B) of signal to background yield to be 1.09. The composition of signal and background events in the sample after applying the cut on the BDT response corresponding to an integrated luminosity of 12.9 fb^{-1} is shown on the right plot in Figure 8.

Ongoing studies

Modeling of parametric shapes for signal and background m_{top} distribution for the final fit is in progress. The parametric model will be tested with different top quark mass hypothesis for calibration and offset correction. Preliminary uncertainties are going to be obtained from pseudo experiments in order to establish the robustness of the analysis technique. The final measurement in data will follow immediately after the method is approved by the collaboration.

Summary

Measurement of single top t-channel production cross section has been presented with 13 TeV pp collision data recorded by the CMS detector. The measurement is in agreement with SM prediction. The measurement with 2.3 fb^{-1} data is the most precise result available at $\sqrt{s}$ = 13 TeV. The uncertainties in the cross sections are dominated by uncertainties due to theory. The ratio of top to anti-top production in t-channel is also determined from data and compared with the predictions from different PDF sets. The value of CKM matrix element V_{tb} is also estimated by comparing the cross section from measurement with theoretical prediction. Lastly, some glimpses of the ongoing study on the top quark mass measurement in single top events with 13 TeV pp collision data are shown. Current studies hold the promise for an independent top mass measurement in single top events with good precision in near future.

List of Publications

Journals

- Cross section measurement of t-channel single top quark production in pp collisions at $\sqrt{s} = 13$ TeV, arXiv:1610.00678, Phys. Lett. B 72 (2017) 752

CMS Physics Analysis Summary (Public Results)

- Measurement of the inclusive cross section of single top quark production in the t channel at $\sqrt{s} = 13$ TeV, CMS-PAS-TOP-16-003.
- Measurement of the t-channel single top-quark cross section at $\sqrt{s} = 13$ TeV, CMS-PAS-TOP-15-004.

Conference Proceedings

- Single top measurements with CMS, PoS(DIS2016)155, Soureek Mitra on behalf of CMS collaboration, presented in "DIS 2016: XXIV International Workshop on Deep-Inelastic Scattering and Related Subjects" during 11-15 April, 2016, held at DESY, Hamburg, Germany.
- Single top quark and V_{tb} measurements, PoS(CKM2016)119, Soureek Mitra on behalf of ATLAS and CMS collaborations, presented in "CKM2016: 9th International Workshop on the CKM Unitarity Triangle" during 28 Nov - 3 Dec, 2016, held at TIFR, Mumbai, India

CMS Internal Notes

- Measurement of the top quark mass in single top events at 13 TeV, CMS-AN-16-443 (In preparation).
- Measurement Measurement of the single top t-channel inclusive cross section and charge ratio at 13 TeV, CMS-AN-15-329.
- Measurement of the single top t-channel inclusive cross section at 13 TeV, CMS-AN-15-114.
- Measurement of the top-quark mass in leptonic decays of t-channel single top events at 8 TeV, CMS-AN-14-194.
- Measurement of WW+WZ cross section in semi-leptonic decays in pp collisions at $\sqrt{s} = 8$ TeV, CMS-AN-12-464.

Contents

List of Figures

List of Tables

Abbreviation

SU(N)	Special Unitary group of degree N
SM	Standard Model
QCD	Quantum Chromodynamics
GeV	Giga electron Volt
TeV	Tera electron Volt
CM	Centre-of-Mass
CERN	European Organisation for Nuclear Research
LHC	Large Hadron Collider
LEP	Large Electron-Positron Collider
DESY	Deutsches Elektronen-Synchrotron
SLAC	Stanford Linear Accelerator Center
ATLAS	A Toroidal LHC ApparatuS
CMS	Compact Muon Solenoid
ALICE	A Large Ion Collider Experiment
LHCb	Large Hadron Collider beauty
PU	Pile-up
LO	Leading Order
NLO	Next-to-Leading Order
NNLO	Next-to-Next-to-Leading Order

NNLL	Next-to-Next-to-Log Order
HLT	High Level Trigger
PF	Particle Flow
JES	Jet Energy Scale
JEC	Jet Energy Calibration
JER	Jet Energy Resolution
DY	Drell Yan
MC	Monte Carlo

Symbols

Symbol	**Physical Quantity**	**Unit**
G^a_μ	Gauge field	
γ^μ	Dirac's gamma matrix	
$\not{A}$	$\gamma^\mu A_\mu$	
$\mathrm{p_T}$	Transverse momentum	GeV
y	Rapidity	–
η	Pseudo-rapidity	–
$\sqrt{s}$	CM energy	GeV or TeV
$\mathcal{L}$	Luminosity	fb^{-1}
$\mathcal{L}_{SM}$	Standard Model Lagrangian	
iM	Scattering amplitude	–
$\mathcal{B}$	Branching fraction	–
σ	Cross section	pb
$\not{E}_\mathrm{T}$	Missing transverse momentum	GeV
$\mathrm{I_{rel}}$	Relative Isolation	–
$\mathrm{m_{top}}$	Mass of the top quark	GeV
$\mathrm{m_T^W}$	Transverse W-boson mass	GeV

Chapter 1

Introduction

The human mind has always been intrigued to know what constitutes matter. In ancient civilizations such as those in Greece and India, the idea was already prevalent that matter is made up of discrete, indivisible units called atoms. In fact, the word "atom" is derived from the Greek word *atomos* which means indivisible. The works of John Dalton and Robert Brown in the 19th century established the concept of atoms. Later on Ernest Rutherford and Niels Bohr successfully showed in the early part of the 20th century that atoms perceived by Dalton are not indivisible, rather they have an internal structure consisting of a massive nucleus at the centre with electrons revolving around it. It was only after the discovery of neutron by Chadwick in 1932 the nuclear model of the atom was complete. Protons and neutrons compose the nucleus inside an atom. As time passed on, a plethora of new particles was discovered and it was Gell-Mann, Zweig and Nishijima who had put forward the quark model to explain the symmetries of hadrons. They proposed that quarks make up hadrons such as protons and neutrons, which were previously thought to be the elementary constituents of matter. Figure 1.1 shows how our knowledge about the fundamental constituents of matter has evolved over time. The elementary constituents such as quarks and leptons interact among themselves via four fundamental interactions, namely strong, electromagnetic, weak and gravitational. The strong force which bind the quarks into protons is the strongest of all. The strong and weak forces, being short-ranged, act only in the sub-nuclear distances. On the other hand, electromagnetic and gravitational forces have an infinite range. The theoretical structure of the standard model (SM) unifies the strong, weak and electromagnetic interactions among elementary particles, while it does not include gravity due to its minuscule effect on elementary particles. However, this assumption breaks down at the Planck Scale where gravity starts to dominate.

The top quark is the heaviest among the known elementary particle within the SM. It decays ($\approx 10^{-25}$s) faster via electroweak interaction than typical hadronization time scale ($\approx 10^{-24}$s) implying that it can exist as a free quark before its decay. Due to this unique property, the top quark provides an excellent probe to signatures of physics beyond the standard model (BSM) through comparison of measurements with precise SM predictions. It plays a key role in testing the SM as it couples most readily to the Higgs boson due to its large mass and is responsible for the largest corrections to the Higgs mass among the SM particles.

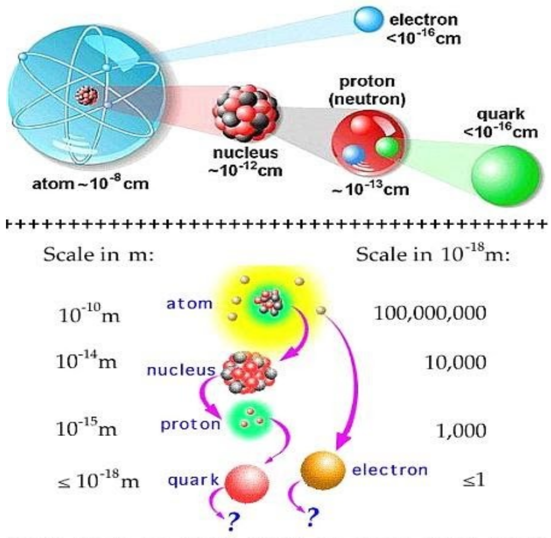


Figure 1.1: Looking deep down into matter: from atoms to quarks; taken from Ref. [33]

For these reasons, ever since its discovery at TeVatron in 1995, the top quark persists to be in the focus of various particle physics programmes. Precise measurements of its properties and cross sections characterizing several production mechanisms have been a primary interest at the Large Hadron Collider (LHC) where top quarks are produced at a significant rate due to higher collision energy and large instantaneous luminosity. This thesis presents a piece of a wide research puzzle, aimed to characterize electroweak production of the top quark via the dominant t-channel production mode and to measure the top quark mass using the t-channel single top events at centre-of-mass (CM) energy, $\sqrt{s}$ =13 TeV.

The theoretical structure is laid out in Chapter 2. This is followed by the details of the experimental set-up and physics objects in Chapter 3 and 4, respectively. An early measurement of the t-channel cross section at $\sqrt{s}$ = 13 TeV is presented in Chapter 5, followed by a more precise measurement with 2.2 fb^{-1} data collected in 2015 at 13 TeV in Chapter 6. Chapter 7 presents an ongoing measurement of the top quark mass using single top events based on 36 fb^{-1} data recorded in 2016.

Chapter 2

Standard Model and Top Quark

2.1 Standard Model

The goal of particle physics is to identify and understand the fundamental constituents of matter and their interactions. Towards this end, the standard model (SM) [34] [35] [36] of particle physics has been supremely successful in describing all known elementary particles and their interactions except for gravity. It is a low-energy effective quantum field theory that treats, in a very elegant way, both matter and force fields with the same formalism, since interactions are themselves regarded as mediated by particles.
The matter content of the SM is categorized into two classes: the quarks, which participate in the strong interaction, and the leptons, which do not. Both quarks and leptons are grouped into three generations (families) each. All quark generations contain a particle with charge $+2/3$ and a particle with charge $-1/3$, while lepton generations include a particle with charge -1 (e.g. the electron) and a neutral particle, named neutrino. The quantum numbers of all quarks and leptons are listed in Table 2.1 while Figure 2.1 shows all matter particles along with the mediators of their interactions.
The SM unifies three out of four types of fundamental interactions, namely, weak, electromagnetic and strong; through the gauge group $\mathcal{G}_{SM}$ defined as follows:

$$\mathcal{G}_{SM} = SU(3)_C \otimes SU(2)_L \otimes U(1)_Y \tag{2.1}$$

Table 2.1: Matter content of the standard model

Particle				Spin	Charge
Leptons	ν_e	ν_μ	ν_τ	$\frac{1}{2}$	0
	e	μ	τ		-1
Quarks	u	c	t	$\frac{1}{2}$	$+\frac{2}{3}$
	d	s	b		$-\frac{1}{3}$

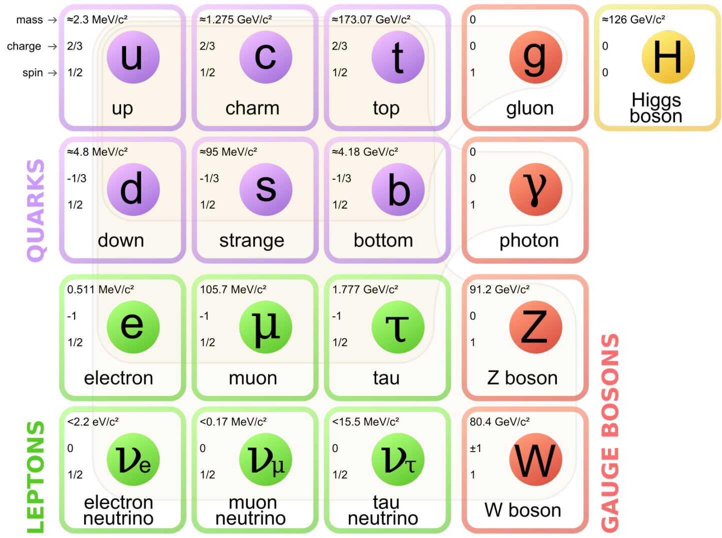


Figure 2.1: SM particles and force mediators; taken from Ref. [37]

where C,L and Y stands for colour quantum number, left-handedness and hypercharge. The core of this model relies on the symmetries represented by three Lie groups whose parameters are continuous variables. Such continuous symmetries can be associated to conservation laws in physics via the Noether's theorem [38]. According to Yang-Mill theory [39] , the renormalizable Lagrangian describing the elementary fermions and their interactions has to be invariant under a set of local gauge transformations that are represented by a Lie group; this entails, inevitably, the existence of gauge bosons as mediators of the interactions themselves. The full $\mathcal{G}_{SM}$ gauge-invariant SM Lagrangian is given by:

$$\mathcal{L}_{SM} = \mathcal{L}_{Gauge} + \mathcal{L}_{Fermions} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa} \tag{2.2}$$

Each term in Eq.(2.2) is explained as follows.
$\mathcal{L}_{Gauge}$ represents the kinetic energy term as well as the self-interaction between the gauge fields and is expressed by:

$$\mathcal{L}_{Gauge} = -\frac{1}{4}\left(B_{\mu\nu}B^{\mu\nu} + \sum_{a=1}^{3} W^a_{\mu\nu}W^{a\mu\nu} + \sum_{a=1}^{8} G^a_{\mu\nu}G^{a\mu\nu}\right) \tag{2.3}$$

with the gauge field tensors defined as:

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu \tag{2.4}$$

$$W^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g \in^{abc} W^b_\mu W^c_\nu \tag{2.5}$$

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu + g_s f^{abc} G^b_\mu G^c_\nu \tag{2.6}$$

The quantity $\in^{abc}$ is the usual anti-symmetric tensor of $SU(2)_L$ group, while f^{abc} represents the structure constants of the $SU(3)_C$ gauge group. The coupling constants g, g' and g_s correspond to the $SU(2)_L$, $U(1)_Y$ and $SU(3)_C$ gauge groups, respectively.
These fields can be exploited to construct a covariant derivative (D_μ) replacing the usual derivative (∂_μ) with the aim of expressing a Lagrangian invariant under local $\mathcal{G}_{SM}$ gauge transformations also for the fermions.

$$D_\mu = \partial_\mu + ig'\frac{Y}{2}B_\mu + ig\sum_{j=1}^{3}\tau^j W^j_\mu + ig_s\sum_{a=1}^{8}\lambda^a G^a_\mu \tag{2.7}$$

where τ^j and λ^a are the generators of the $SU(2)_L$ and $SU(3)_C$ groups and Y is the hypercharge. The second term in Eq.(2.2) corresponds to the kinetic energy for the fermions as well as the interactions between these fermions and the gauge fields, can be symbolically written as:

$$\mathcal{L}_{Fermions} = \bar{\psi_L} \not{D} \psi_L + \bar{\psi_R} \not{D} \psi_R \tag{2.8}$$

where ψ_L stands for fermion $SU(2)_L$ doublets and ψ_R denotes $SU(2)_L$ singlet states and $\not{D} = \gamma^\mu D_\mu$, with γ^μ being the Dirac matrices.
The third term in Eq.(2.2) corresponds to the dynamics of the Higgs field Φ, predicted by Brout, Englert [40] and Higgs [41], responsible for spontaneous breaking of electroweak symmetry within the SM. Φ is an all-pervading complex scalar field represented by the Eq.(2.9), which is a doublet in the $SU(2)$ space but carries non-zero $U(1)$ hypercharge.

$$\Phi = \begin{bmatrix} \phi^+ \\ \phi^0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{bmatrix} \tag{2.9}$$

The Lagrangian for the Higgs field is given by:

$$\mathcal{L}_{Higgs} = (D^\mu \Phi)^\dagger (D_\mu \Phi) - \mu^2 \Phi^\dagger \Phi - \lambda (\Phi^\dagger \Phi)^2 \tag{2.10}$$

Firstly, we note that λ is required to be positive so that the Higgs potential $V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda(\Phi^\dagger \Phi)^2$ is bounded from below corresponding to a stable vacuum. Now $\mu^2 > 0$ corresponds to the unique possibility of potential minimum at $\langle 0|\Phi|0\rangle = 0$ and there is no scope for breaking any symmetry spontaneously. However, in the case of $\mu^2 < 0$, the potential has the famous "Mexican hat" shape as shown in Figure 2.2 such that the minimum of the potential lies on a circle pointing towards a degenerate family of vacuum states, which is an additional degree of freedom. The consequent arbitrary choice of

$$\langle 0|\Phi|0\rangle = \left\langle 0 \left| \begin{bmatrix} \phi^+ \\ \phi^0 \end{bmatrix} \right| 0 \right\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ v \end{bmatrix}, v = -\sqrt{\frac{-\mu^2}{\lambda}} \tag{2.11}$$

implies spontaneous breaking of the initial symmetry of the system. A perturbation around the chosen minimum leads to couplings of the Higgs field with the gauge bosons, and are interpreted as mass terms. Also, an additional massive scalar particle, the Higgs boson H, supervenes naturally. The discovery of the Higgs boson was finally achieved in 2012 at the LHC [42, 43] and marked another silverware in the success story of the SM so far.

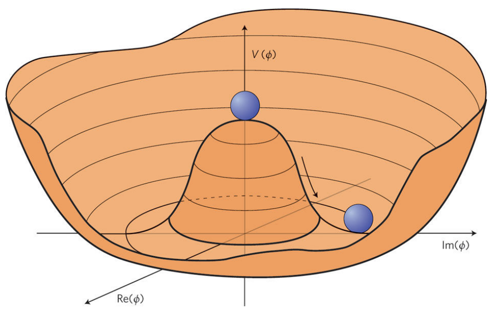


Figure 2.2: Higgs Potential; taken from Ref. [44]

The last piece in the SM Lagrangian in Eq.(2.2) is the $SU(2)_L$-invariant Yukawa interaction of the Higgs field with the fermion Dirac fields, responsible for the fermion masses, is given by

$$\mathcal{L}_{Yukawa} = \sum_{i=1}^{3}\sum_{j=1}^{3} \Gamma_u^{ij}\bar{Q}_L^i i\tau^2\Phi^* u_R^j + \Gamma_d^{ij}\bar{Q}_L^i\Phi d_R^j + \Gamma_\ell^{ij}\bar{L}_L^i\Phi\ell_R^j + h.c \tag{2.12}$$

where the coefficients Γ_u, Γ_d and Γ_ℓ represent the 3×3 Yukawa complex matrices in the space of the three fermion generations with the superscripts u, d and ℓ denote the up- and down-type quarks and leptons, respectively. The mass matrix corresponding to the leptons is given by, $m_\ell^{ij} = \frac{\Gamma_\ell^{ij} v}{\sqrt{2}}$, when the Higgs field has a vacuum expectation value of v. As the SM does not contain any right-handed neutrino field, the neutrinos remain massless in the model. Therefore, the mass matrix m_ℓ^{ij} turns out to be diagonal in the flavour space. However, for quarks, a mixing between the mass eigenstates of different flavours can happen, which is the origin of the Cabibbo-Kobayashi-Maskawa (CKM) mixing matrix [45] [46]. The mass matrices for quarks are given by $m_u^{ij} = \frac{\Gamma_u^{ij} v}{\sqrt{2}}$, $m_d^{ij} = \frac{\Gamma_d^{ij} v}{\sqrt{2}}$. Such matrices can be diagonalized via a bi-unitary transformation (e.g., $\Gamma_u \rightarrow U_u^{L\dagger}\Gamma_u U_d^L$) in order to obtain the physical mass eigenstates of the particles. The choice of a new basis to express the fermion fields has no effect on the leptons, on the other hand, a new unitary matrix as shown in Eq.(2.13) is associated with the quark sector, which in effect performs a rotation linking the eigenstates of electroweak interaction with those of mass.

$$V_{CKM} = U_u^{L\dagger}U_d^L = \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{td} \end{bmatrix} \tag{2.13}$$

The unitarity condition $\sum_{k=1}^{3} V_{ik}V_{jk}^* = \delta_{ij}$ limits the number of parameters to three mixing angles (θ_{ij}) and one irreducible phase (δ). The standard CKM matrix representation, in terms of sines and cosines of the mixing angles ($c_{ij} = \cos\theta_{ij}$; $s_{ij} = \sin\theta_{ij}$) and the phase δ is given by:

$$V_{CKM} = \begin{bmatrix} c_{12}c_{13} & s_{12}c_{13} & s_{13}e^{-i\delta} \\ -s_{12}c_{23} - c_{12}s_{23}s_{13}e^{i\delta} & c_{12}c_{23} - s_{12}s_{23}s_{13}e^{i\delta} & s_{23}c_{13} \\ s_{12}s_{23} - c_{12}c_{23}s_{13}e^{i\delta} & -c_{12}s_{23} - s_{12}c_{23}s_{13}e^{i\delta} & c_{23}c_{13} \end{bmatrix} \tag{2.14}$$

Currently best known values for the CKM matrix parameters are [16]:

$$\theta_{12} = 13.04 \pm 0.05^0,\ \theta_{13} = 0.201 \pm 0.011^0,\ \theta_{23} = 2.38 \pm 0.06^0,\ \delta = 1.20 \pm 0.08\,\text{rad} \tag{2.15}$$

It is known experimentally that $s_{13} << s_{23} << s_{12} << 1$, so it is convenient to exhibit this hierarchy using the Wolfenstein parameterization [47] written in a series expansion of $\lambda = \sin\theta_{12}$. Defining

$$s_{12} = \lambda = \frac{|V_{us}|}{\sqrt{|V_{ud}|^2 + |V_{us}|^2}},\ s_{23} = A\lambda^2 = \lambda|\frac{V_{cb}}{V_{us}}| \tag{2.16}$$

$$s_{13}e^{i\delta} = A\lambda^3(\rho + i\eta) = \frac{A\lambda^3(\bar{\rho} + i\bar{\eta})\sqrt{1 - A^2\lambda^4}}{\sqrt{1-\lambda^2}[1 - A^2\lambda^4(\bar{\rho} + i\bar{\eta})]} = V_{ub}^* \tag{2.17}$$

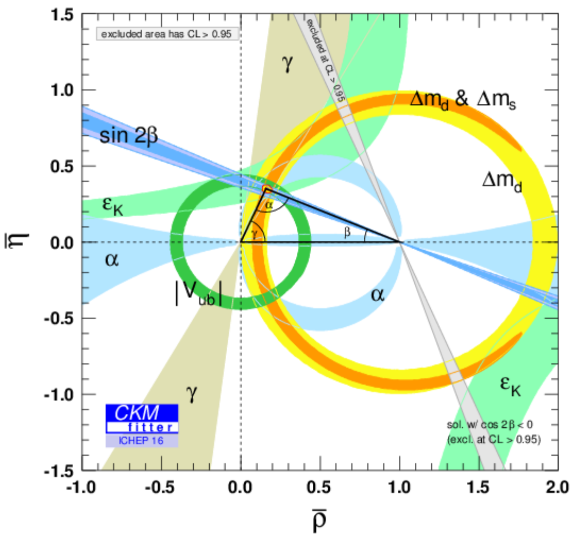


Figure 2.3: Constraints in the $\bar{\rho}$-$\bar{\eta}$ plane are shown as the result of multiple experimental measurements of the CKM matrix elements; the shaded areas correspond to a confidence level of 95% [48]

the CKM matrix becomes:

$$V_{CKM} = \begin{bmatrix} 1-\frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1-\frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1-\rho-i\eta) & -A\lambda^2 & 1 \end{bmatrix} + \mathcal{O}(\lambda^4) \tag{2.18}$$

The new variables can then be exploited, together with the unitarity condition, to draw a unitary triangle in the $(\bar{\rho}, \bar{\eta})$ plane. Its three sides, renormalized to the values $\frac{V_{ud}V_{ub}^*}{V_{cd}V_{cb}^*}$, 1, $\frac{V_{td}V_{tb}^*}{V_{cd}V_{cb}^*}$, and internal angles, are an explicit visualization of some fundamental parameters of the SM, so that their precise determination is crucial. One of the goals of flavour physics is to experimentally overconstrain the CKM unitarity triangle, as depicted in Figure 2.3. The individual measurements of different processes involving quark mixing are reported with distinct color bands which should converge on the angle labeled α, and are used to perform a global fit with the CKMFitter [48]. Such a fit, which relies as well on the SM constraints, gives the most precise estimate of the nine elements of the CKM matrix as in Eq.(2.19) [48].

$$\begin{bmatrix} 0.974334^{+0.000064}_{-0.000068} & 0.22508^{+0.00030}_{-0.00028} & 0.003715 \pm 0.000060 \\ 0.22494^{+0.00029}_{-0.00028} & 0.973471 \pm 0.000067 & 0.04181^{+0.00028}_{-0.00060} \\ 0.008575^{+0.000076}_{-0.000098} & 0.04108^{+0.00030}_{-0.00057} & 0.999119^{+0.000024}_{-0.000012} \end{bmatrix} \tag{2.19}$$

2.2 Top quark production at hadron colliders

Hadrons such as protons are composite particles made up by valence quarks, sea quarks and gluons arising due to vacuum polarization, which are generically termed as "partons".

Due to heavier mass of the (anti) protons, hadron colliders can provide high centre-of-mass (CM) energies due to very low radiation loss in head-on collisions. At LHC two counter-rotating proton beams collide head on. According to the QCD factorization theorem [49] any high energy hadronic reaction with a large momentum transfer can be factorized into two distinct contribution: the short-distance interaction or "hard scattering" between the participating partons of type i and j, $\hat{\sigma}^{ij}$, convoluted with the long-distance effect encapsulated in the "parton distribution functions" f_i and f_j [50]. The $\hat{\sigma}^{ij}$ contribution can be calculated in the perturbative QCD regime. The parton distribution functions (PDFs) describe the probability densities of partons i, j carrying $x_{i,j}$ fractions of total momenta of the incoming protons. They also depend on the factorization scale μ_F; i.e., $f_{i,j}{=}f_{i,j}(x_{i,j};\mu_F^2)$. The cross section of any final state X of interest in pp collisions can be formalized under the factorization approach as shown in Figure 2.4, as :

$$\sigma(pp \rightarrow X) = \sum_{i,j=g,q,\bar{q}} \int dx_i dx_j \underbrace{f_i(x_i,\mu_F^2)f_j(x_j,\mu_F^2)}_{\text{long}-\text{distance part}} \times \underbrace{\hat{\sigma}^{ij}(x_i,x_j)}_{\text{short}-\text{distance part}} \tag{2.20}$$

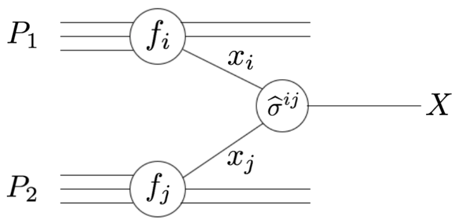


Figure 2.4: Hard scattering in the parton model at hadron colliders.

A precise knowledge of the Parton Distribution Functions (PDFs) of the (anti-) proton is essential in order to make predictions for the SM and beyond the SM processes at hadron colliders. As QCD does not predict the parton content of the (anti-)proton, the shapes of the PDFs are constrained by global fits on a large number of cross section data in various processes from experiemnts performed at SLAC, TeVatron, CERN and deep inelastic e-p scattering at HERA collider. The recent precise data at LHC has the potential to improve the constraints further. However, the evolution of the PDFs with the scale is predicted via the DGLAP evolution equation [51,52,53] by pertubative QCD. The extensive analyses resulted in the development of many parametrizations for the description of the PDFs at Leading Order (LO), Next-to-Leading-Order (NLO) and Next-to-Next-to-Leading-Order (NNLO): CTEQ [54], GJR [55], NNPDF [26], HERAPDF [25], ABM [23], MMHT [24], MSTW [56] etc. Such PDF sets differ essentially in the dataset used to constrain the fit, the value of the strong coupling α_s and of the quark masses, the way in which higher order corrections are implemented and the way the systematic uncertainties are treated. Figure 2.5 shows the MMHT14 PDFsets [24] at two different scales: the gluon distribution functions in the proton are peaked at low x values, while valence quark functions are enhanced at high x values.

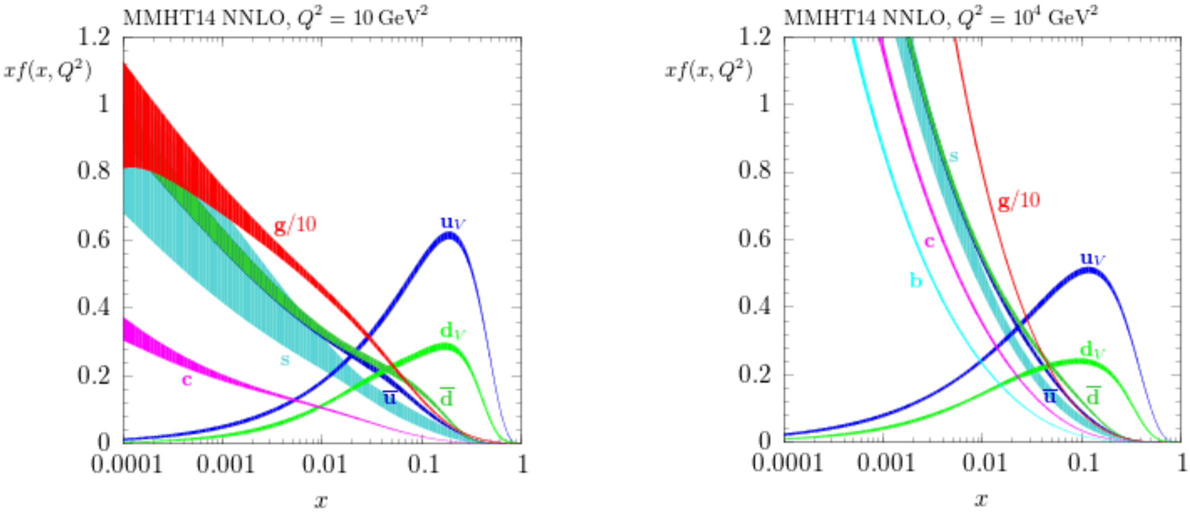
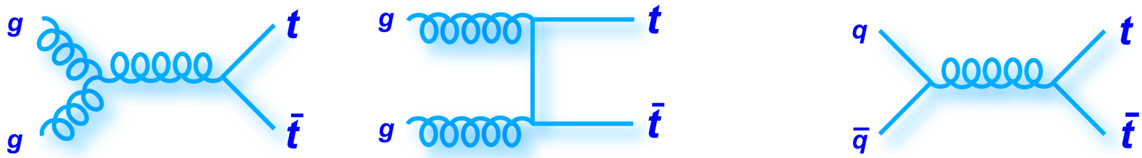


Figure 2.5: MMHT14 PDFset [24] at different scales (Q^2): $Q^2 = 10$ GeV^2 (left) and $Q^2 = 10^4$ GeV^2 (right). The associated 1σ (68% C.L.) uncertainty bands are also represented.

2.2.1 Top-quark pair production

Top quarks are mostly produced in pairs via gluon-gluon fusion or quark-anti-quark annihilation under the factorization approach, at pp collisions of the LHC. The corresponding Feynman diagrams for top pair productions are shown in Figure 2.6. The process is ini-



(a) gluon-gluon fusion (b) $q\bar{q}$ annihilation

Figure 2.6: Leading order (LO) Feynman diagrams for top-quark pair productions via strong interactions at hadron colliders [57].

tiated if energy shared by the partons at least corresponds to the total rest mass of the top quark pair; that is $\hat{s} \geq 4m_t^2$, and therefore:

$$x_i x_j = \frac{\hat{s}}{s} \geq \frac{4m_t^2}{s} \tag{2.21}$$

If we consider the distribution of PDFs, the apparent feature is a steep fall with increasing momentum fractions. This implies that the partons will typically have energy available for close to the $t\bar{t}$ production threshold. We can also expect, with a further degree of simplification, that the interacting partons would share roughly the same $\langle x \rangle$, given by

$$\langle x \rangle \sim \sqrt{x_i x_j} \sim \frac{2m_t}{\sqrt{s}} \tag{2.22}$$

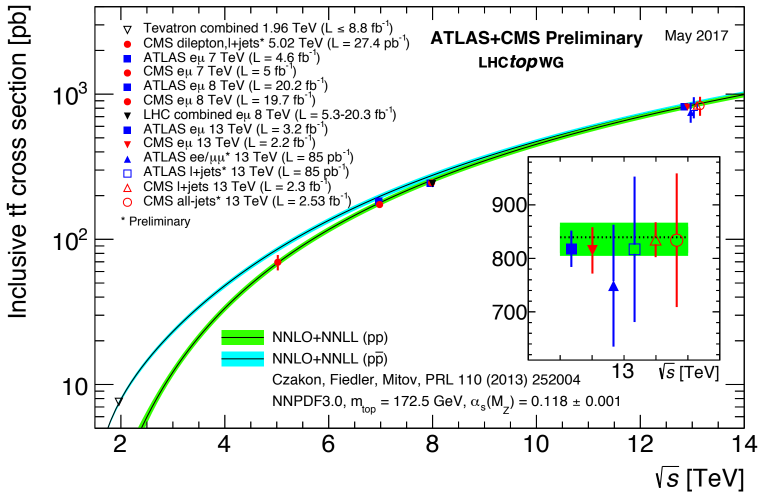


Figure 2.7: $t\bar{t}$ cross-sections at different CM energies [61]

Thus it is evident that higher the CM energy, lower the average momentum fraction carried by each parton for $t\bar{t}$ production. As gluon PDFs typically have larger values compared to the ones from valence or sea quarks at lower momentum fractions, the production of top pair proceeds almost exclusively via gluon-gluon fusion at the LHC. The cross-sections of $t\bar{t}$ production at different CM energies for the LHC are calculated with an accuracy up to next-to-next-to-leading-order (NNLO) in QCD, including resummation of soft gluon terms at next-to-next-to-leading-log (NNLL) accuracy [58].These calculations have been performed assuming the top quark mass to be 172.5 GeV.
Several measurements of $t\bar{t}$ cross-sections have been performed at TeVatron and LHC with higher sensitivity following the first measurement by CDF [59] and DØ [60] Collaborations in 1995. Figure 2.7 summarizes the latest results from CDF, DØ, ATLAS and CMS as function of the CM energy. All experimental results are found to be consistent with each other and with the theoretical predictions within uncertainties.

2.2.2 Single top-quark production

Top quarks are also produced singly via charged-current electroweak interaction in pp collisions taking place at the LHC. Such productions proceed via three channels: t-channel, tW-channel and s-channel. The LO diagrams of the three different production modes are shown in Figure 2.8. The Lagrangian at the tWb production vertex ($\mathcal{L}_{tWb}$) can be expressed as

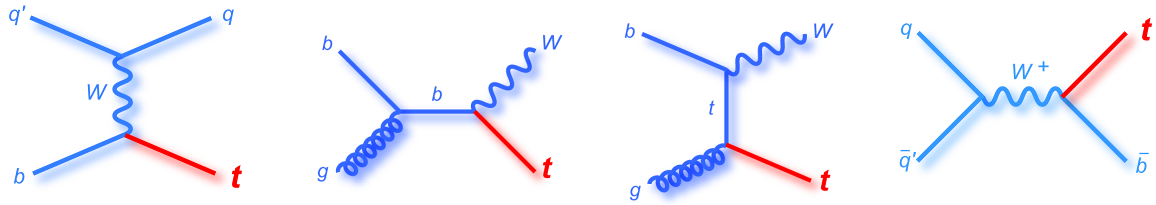
$$\mathcal{L}_{tWb} = -\frac{g}{2\sqrt{2}} V_{tb} \bar{b} \gamma^{\mu} (1-\gamma^5) t W^-_{\mu} + h.c \tag{2.23}$$

with the scattering matrix element ($i\mathcal{M}$) given by

$$|i\mathcal{M}| \propto \frac{g}{2\sqrt{2}} |V_{tb}| \tag{2.24}$$

This implies that the single top production, which is proportional to the absolute square of the scattering matrix element, is directly proportional to $|V_{tb}|^2$, i.e.,

$$\sigma \propto |i\mathcal{M}|^2 \propto |V_{tb}|^2 \tag{2.25}$$

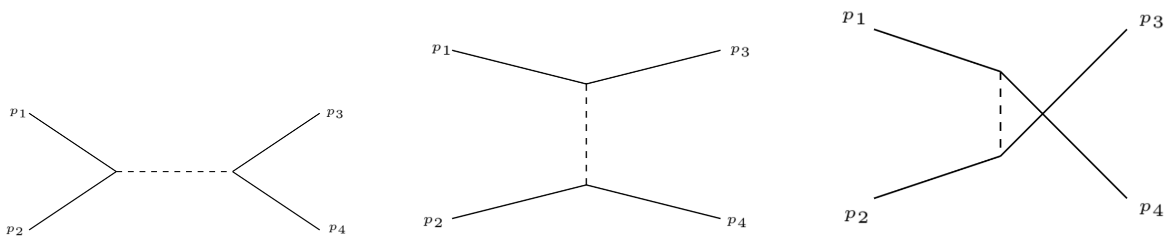


(a) t-channel (b) tW-channel (c) s-channel

Figure 2.8: LO Feynman diagrams for single top-quark production via electroweak interactions at hadron colliders [57]

It is quite evident from Figure 2.8 that the three production modes are distinguished by the virtuality of the W boson participating at the production vertex, which can be elegantly formulated in terms of the Mandelstam variables describing a generic 2→2 scattering process (Figure 2.9). The variables are defined as:

$$s = (p_1 + p_2)^2 = (p_3 + p_4)^2 \tag{2.26}$$
$$t = (p_1 - p_3)^2 = (p_2 - p_4)^2 \tag{2.27}$$
$$u = (p_1 - p_4)^2 = (p_2 - p_3)^2 \tag{2.28}$$



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

Figure 2.9: Visualization of Mandelstam variables in 2→2 scattering

t-channel

The t-channel production, which is characterized by the 4-momentum transfer squared, proceed via the exchange of a virtual, space-like ($Q^2 < 0$ GeV^2) W boson between the two initial partons. The parton-level cross section at LO as shown in Figure 2.8a is given by

$$\hat{\sigma}^{t-ch.}_{LO} \propto \frac{g^4}{4}|V_{tb}|^2|V_{qq'}|^2\frac{s^2}{Q^2 - M_W^2}(1 - \frac{m_t^2}{s}) \tag{2.29}$$

where Q(=$\sqrt{t}$) denotes the 4-momentum exchange in the reaction and $V_{qq'}$ is the CKM matrix element that couples the other initial- and final-state quarks q' and q, respectively. Eq.(2.29) clearly indicates that cross section increases considerably with the increase of

CM energy ($\sqrt{s}$). In fact, the t-channel has the largest cross-section among the three production processes at LHC energies as indicated in Table 2.2. As the probability of finding a b-quark is less inside a proton, the next-to-leading-order (NLO) diagram (2→3) as shown in Figure 2.10, where the b-quark appears due to gluon splitting, is dominant for t-channel production in pp collisions at the LHC.

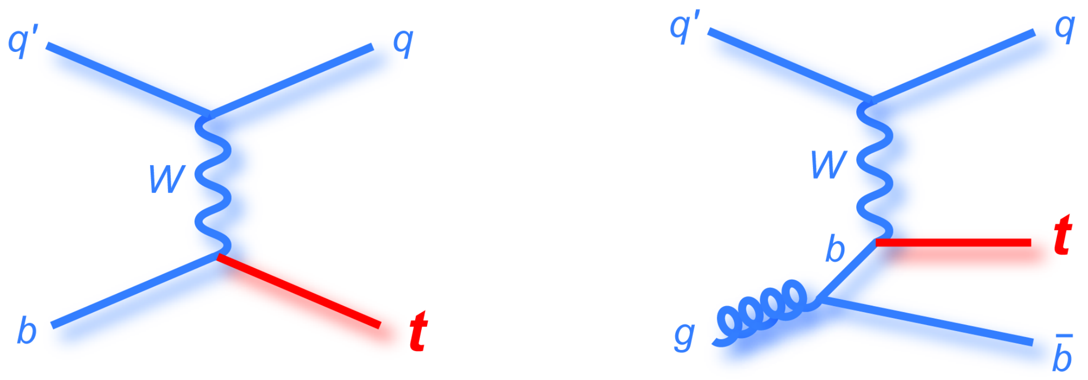


Figure 2.10: t-channel diagrams: leading order (left) and next-to-leading-order (right) [57]

tW-channel

The top quark can also be singly produced in association with a real W boson via the so called tW-channel. The main partonic diagram contributing to this channel is $gb \to tW$ as shown in Figure 2.8b, as the contributions from other initial states gd and gs are suppressed by the corresponding CKM matrix elements V_{td} and V_{ts}. Since two different Feynman diagrams of s- and u-channel type, participate at LO, the computation of total scattering amplitude is more complex than other production mechanisms. The total scattering amplitude also do not show explicit dependence on the CM energy. Further, the tW-channel has large interference from $t\bar{t}$ at NLO where the initial b-quark comes from another gluon, which makes it very sensitive to model uncertainties. The presence of a gluon in the initial state makes this mode interesting to measure at LHC. In fact, this is the second most dominant single top production mode at LHC.

s-channel

The s-channel process, characterized by the total CM energy squared, proceeds via annihilation of the two incident partons into an intermediate virtual, time-like $[Q^2 \geqslant (m_t + m_b)^2]$ W-boson, which then decays into a top and an anti-bottom quarks. This process is therefore particularly sensitive to resonances or stable unknown particles that may replace the standard model mediator. The partonic cross section from the LO diagram shown in Figure 2.8c (Q=$\sqrt{s}$) is :

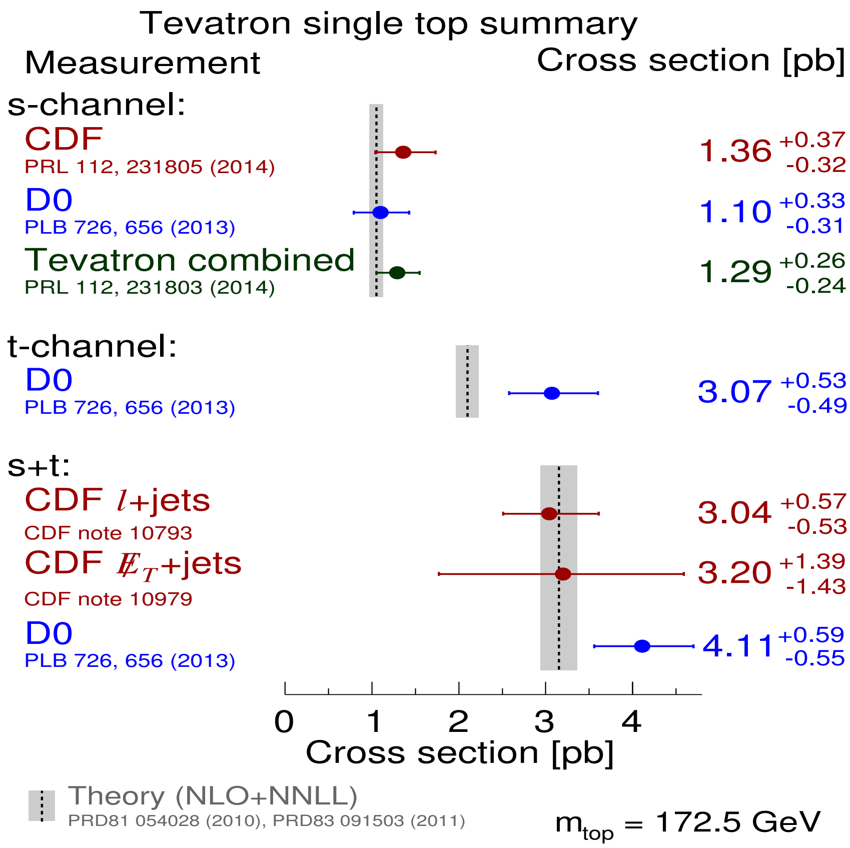
$$\hat{\sigma}_{LO}^{s-ch} \propto \frac{g^4}{4}|V_{tb}|^2|V_{qq'}|^2\frac{t^2}{Q^2 - M_W^2}(1 - \frac{m_t^2}{t}) \tag{2.30}$$

which can be easily derived from Eq.(2.29) using crossing symmetry. Eq.(2.30) underlines that the single top s-channel cross section depends slightly on the collider CM energy, and consequently no enhancement is favoured at LHC with increasing CM energy. Moreover, the presence of an anti-quark in the initial state suggests further suppression due to PDF in pp collisions. In fact, this is the least dominant single top production mode at LHC. Backgrounds are overwhelming for this channel even at 7 TeV whose contribution increase more steeply with $\sqrt{s}$ reducing the signal purity further, and thus presents a very difficult task to make precise measurements in this mode.

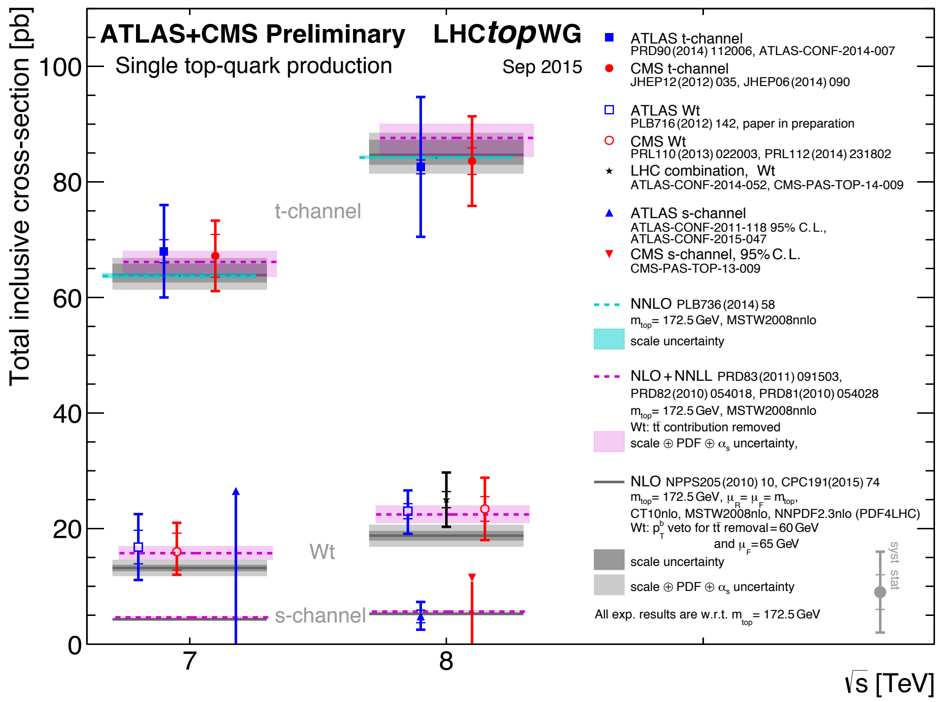
Table 2.2 summarizes predictions of single top cross section at different $\sqrt{s}$ at the LHC and Figure 2.11 lays out previous measurements of single top cross-sections at TeVatron and during Run-I of LHC.

Table 2.2: Single top cross-section prediction at different CM energy at the LHC [62]

Single top channel	$\sigma\,(\sqrt{\mathrm{s}} = 7\mathrm{TeV})$	$\sigma\,(\sqrt{\mathrm{s}} = 8\ \mathrm{TeV})$	$\sigma\,(\sqrt{\mathrm{s}} = 13\ \mathrm{TeV})$
t-channel	$63.89^{+2.91}_{-2.52}$ pb	$84.69^{+3.76}_{-3.23}$ pb	$216.99^{+9.04}_{-7.71}$ pb
tW-channel	$15.74^{+1.17}_{-1.21}$ pb	$22.37{\pm}1.52$ pb	$71.7{\pm}3.85$ pb
s-channel	$4.29^{+0.19}_{-0.17}$ pb	$5.24^{+0.22}_{-0.20}$ pb	$10.32^{+0.40}_{-0.36}$ pb



(a) Summary of TeVatron measurements of single top cross-sections [63]



(b) Summary of LHC measurements of single top cross-sections in Run-I [61]

Figure 2.11: Summary of single top cross-section measurements at TeVatron and LHC Run-I

2.2.3 Interest in single top measurements

Extraction of V_{tb}

Determination of various elements of the CKM matrix governing the electroweak interaction of quarks are of profound importance. Single top measurements are extremely valuable on this regard, as they can be used to extract the CKM matrix element V_{tb}. Since the production cross-section is proportional to $|V_{tb}|^2$ as shown in Eq.(2.25), V_{tb} can be extracted from the ratio of to measured cross-section to predicted cross section, under the assumptions $|V_{tb}| \gg |V_{td}|, |V_{ts}|$ and left-handed electroweak coupling. No hypothesis on the number of fermion generations is needed.

$$|V_{tb}| = \sqrt{\frac{\sigma^{\text{meas.}}_{\text{single top}}}{\sigma^{\text{theory}}_{\text{single top}}}} \tag{2.31}$$

In the case where $\sigma^{\text{meas.}}_{\text{singletop}} > \sigma^{\text{theory}}_{\text{singletop}}$, bounds on $|V_{tb}|$ are set.

Search for new couplings

The SM Lagrangian for the tWb vertex can be minimally extended to a more generic form as given in Eq.(2.32) by introducing anomalous vector and tensor couplings [1] in addition to the (V-A) coupling structure from the SM:

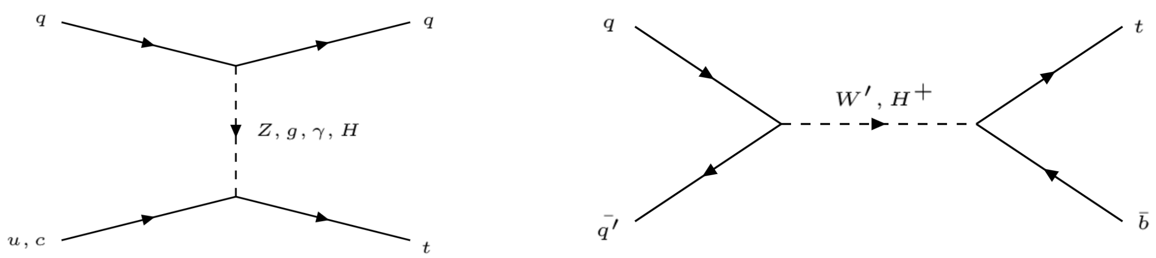
$$\mathcal{L}^{\text{anom.}}_{tWb} = -\frac{g}{\sqrt{2}}\bar{b}\gamma^{\mu}(f^V_L P_L + f^V_R P_R)tW^-_{\mu} - \frac{g}{\sqrt{2}}\bar{b}\frac{-i\sigma^{\mu\nu}q_{\nu}}{\text{m}_W}(g_L P_L + g_R P_R)tW^-_{\mu} + \text{h.c} \tag{2.32}$$

where q_{ν} and m_W are the 4-momentum and mass of the W boson, $P_{L,R} = \frac{(1\mp\gamma^5)}{2}$ are the left- and right-handed projection operators, $\sigma^{\mu\nu} = \frac{i[\gamma^{\mu},\gamma^{\nu}]}{2}$, $f^V_{L,R}$ and $g_{L,R}$ are the complex left- and right-handed vector and tensor couplings, respectively. At tree level in the SM, $f^V_L = V_{tb}$ and $f^V_R = g_{L,R} = 0$. The additional tensor couplings can alter the polarization of the top quark as well as the helicity fractions of the W boson from its decay. The t-channel production being the dominant of the single top process is an ideal hunting ground for such new couplings via measurement of the spin-polarization of the top decay products. The tW-channel is also very sensitive to the anomalous coupling, as the contamination from other beyond SM sources such as flavour-changing neutral currents (FCNC) and coupling with new heavier gauge bosons are negligible at the production vertex thanks to the presence of an on-shell W boson.

In the SM, the flavor quantum number of fermions can be changed by charged currents, i.e., via weak interactions mediated by a $W^{\pm}$ boson. FCNC processes such as qX$\rightarrow$t, where X is an electrically neutral boson e.g., photon, Z, gluon or H, are absent at tree level. The same can happen in higher-order loop diagrams with the help of a virtual W-boson; however this kind of processes is highly suppressed through the GIM mechanism [64] with a typical branching ratio in the range $10^{-10} - 10^{-14}$. Several new physics models suggest that this FCNC branching ratio can be enhanced by 7-10 orders of magnitude [65]. The t-channel is most sensitive to the FCNC processes as shown in Figure 2.12a. As the required u or c-flux at the reaction initial state is higher than the b-flux in a pp collision as indicated in Figure 2.5, it compensates for the smaller FCNC coupling. Also, the FCNC operators involve a different set of spectator quarks (denoted by q in Figure 2.12a) in the reaction and thus do not interfere with the SM t-channel process. Hence, the presence of FCNC beyond the SM can be seen as an enhancement of the t-channel cross section. FCNC may also be seen in the s-channel through reactions such as $q\bar{q}$ →Z→tc; however, these new physics signals may be difficult to extract in practice as it calls for the charm-quark identification along with that of the top quark. As the Wt-channel requires the $W^{\pm}$ to be on-shell, it receives no FCNC contribution.

Search for new bosons

In a minimal extension of the SM, it is possible to postulate the existence of a larger gauge group that somehow reduces down to the $\mathcal{G}_{SM}$ in Eq.(2.1) at low energy. Such theoretical models predict additional massive gauge bosons; top quark being the most massive among SM particles is expected to strongly couple with these heavy particles. The s-channel production cross section can be increased or decreased owing to the interference of the diagrams which are mediated by heavy gauge bosons, such as W′ [2] or charged Higgs bosons [3], as illustrated in Figure 2.12b with those allowed within the SM. If enough energy is available, these massive particles can be produced on-shell resulting in a resonant behavior. Similar interaction between heavy gauge bosons and the top quark is possible in the t-channel, but they will be suppressed by $\frac{1}{M^2}$, where M is the mass of the new gauge boson, due to the space-like nature of the propagator. Exotic production of such gauge bosons on-shell in the tW-channel is also possible, but would result in a very different final states with two top quarks and one b-quark, since the heavy gauge boson dominantly decays to bt assuming minimal flavour violation.



(a) FCNC contribution to t-channel (b) New resonance in s-channel

Figure 2.12: Potential new physics contributions to various single top processes

2.3 Top quark mass

2.3.1 Significance

The top quark is the heaviest known elementary particle. Its mass, m_{top}, is a fundamental parameter of the SM and is of significant importance due its strikingly large value (as heavy as a tungsten atom). The high mass of the top quark indicate its Yukawa coupling to the Higgs field of $\mathcal{O}(1)$. Further the Yukawa structure of electroweak Lagrangian implies that the top quark has large effect on radiative corrections to the electroweak processes as well as to the masses of the W,Z and Higgs bosons. Precision elctroweak measurements of the masses of W,Z and top can therefore be used to test the self-consistency of the SM.

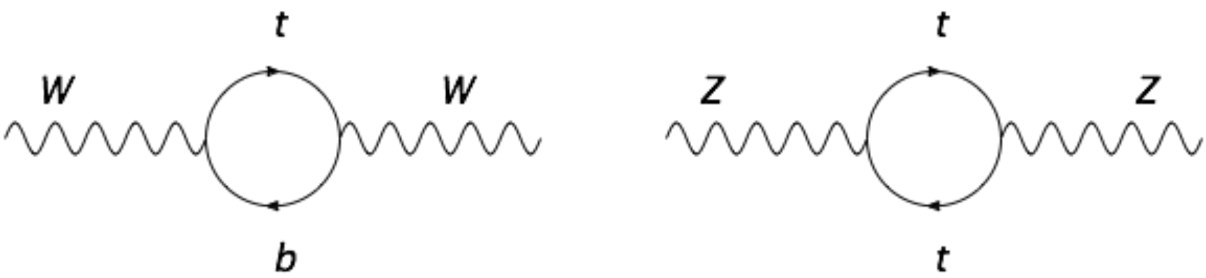
All electroweak quantities such as mass, width and couplings of the W- and Z-boson depend on five parameters in the SM. At leading order this dependence boils down to only three parameters: two gauge couplings and the Higgs-field vacuum expectation value. The three best-measured electroweak quantities can be used to determine these three parameters: The electromagnetic coupling constant α, measured in low-energy experiments [66], the Fermi constant, $\mathrm{G_F}$ determined from the μ lifetime measurement [67], and the mass of the Z-boson precisely measured in $\mathrm{e^+e^-}$ colliders [68]. By defining the electroweak mixing angle θ_{W} through $sin^2\theta_{\mathrm{W}} = 1 - \frac{m^2_{\mathrm{W}}}{m^2_{\mathrm{Z}}}$, the W-boson mass can be expressed as [69]:

$$m^2_{\mathrm{W}} = \frac{\pi\alpha}{\sqrt{2}\mathrm{G_F} sin^2\theta_{\mathrm{W}}(1-\Delta r)} \tag{2.33}$$

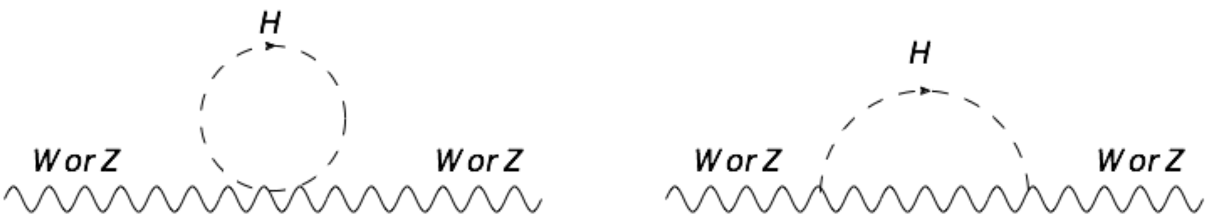
where Δr contains all one-loop corrections. Contributions to Δr due to the top quark and Higgs boson, at one-loop level as shown in Figure 2.13, are given by:

$$(\Delta r)_{\mathrm{top}} \simeq -\frac{3\mathrm{G_F}}{8\sqrt{2}\pi^2\tan^2\theta_{\mathrm{W}}} m^2_t \tag{2.34}$$

$$(\Delta r)_{\mathrm{Higgs}} \simeq -\frac{3\mathrm{G_F} m^2_{\mathrm{W}}}{8\sqrt{2}\pi^2}(2\ln\frac{m_{\mathrm{H}}}{m_{\mathrm{Z}}} - \frac{5}{6}) \tag{2.35}$$



(a) One loop corrections due to the top quark



(b) One loop corrections due to the Higgs boson

Figure 2.13: Radiative corrections to W and Z due to the top quark and Higgs boson

From the above equations, the leading $\mathrm{m_t}$ dependence is quadratic, i.e. very strong, whereas leading $\mathrm{m_H}$ dependence is only logarithmic, i.e. rather weak. The inferred constraints on $\mathrm{m_H}$ are thus much weaker than those on $\mathrm{m_t}$ from precision $\mathrm{m_W}$ measurements only. In fact, electroweak precision data were used to successfully predict the mass of the top quark before its discovery in 1995 [59, 60]. Post its discovery, direct $\mathrm{m_t}$ measurements together with precision $\mathrm{m_W}$ data can indirectly constrain the mass of the newly observed Higgs boson and perform a stringent test of the structure of the spontaneous symmetry breaking in the SM. Figure 2.14 shows 68% and 95% confidence-level (CL) contours obtained from a scan of fits with fixed variable pairs $\mathrm{m_W}$ vs. $\mathrm{m_t}$. The narrower blue and larger grey allowed regions are the results of the fit including and excluding the $\mathrm{m_H}$ measurements, respectively. The horizontal bands indicate the 1σ regions of the $\mathrm{m_W}$ and $\mathrm{m_t}$ measurements [70].

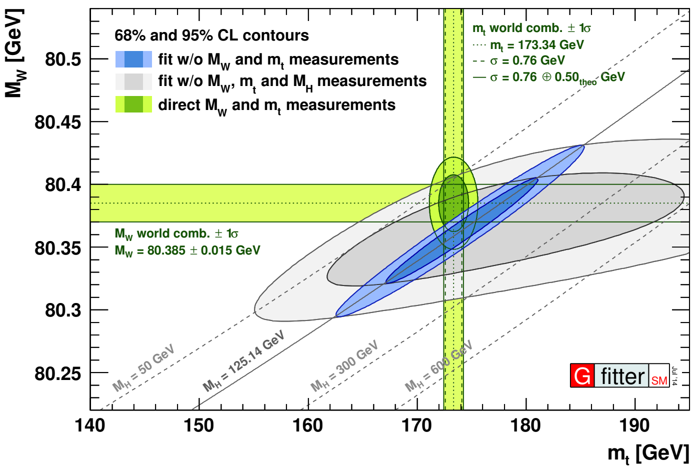


Figure 2.14: Electroweak fits performed with precision measurements of $\mathrm{m_W}$ and $\mathrm{m_t}$ [70]

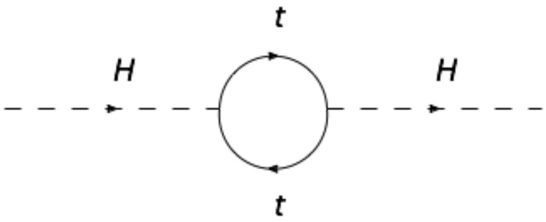
The stability of electroweak vacuum depends on the value of Higgs self-coupling λ. If $\lambda < 0$, then the scalar potential $V(\Phi)$ (Section 2.1) is unbounded from below, which leads to instability even under a tiny perturbation. The Higgs self-coupling λ gets radiative corrections ($\Delta\lambda$) from all the SM particles. The leading contribution comes from the top quark due to diagrams like Figure 2.15a since the top quark has the largest coupling to the Higgs field and corrections scale as:

$$(\Delta\lambda)_{\mathrm{top}} \propto \mathrm{m_t^4} \tag{2.36}$$

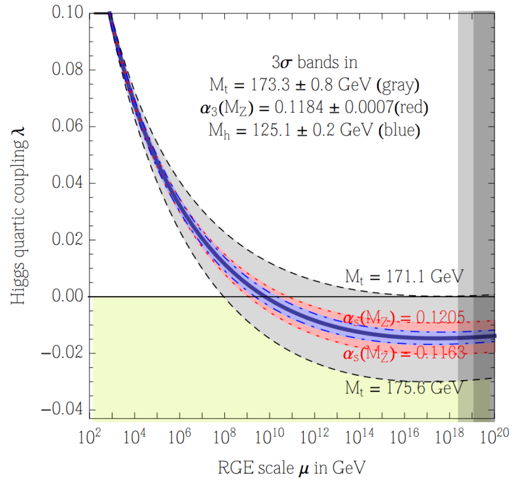
If we evolve $\Delta\lambda$ all the way up to the Planck scale ($\sim 10^{19} GeV$), the value of $\mathrm{m_t}$ will decide whether λ would remain positive at all energies or not (Figure 2.15b) [7, 71]. Therefore, a precise knowledge of the value of the top-quark mass might just reveal the fate of our universe.

2.3.2 Definition

Quark masses are fundamental parameters of the gauge theory of strong interaction, namely, QCD. In quantum field theory, the propagator of a massive fermion has a pole at m_0 called the pole mass, which also corresponds to the invariant mass calculated from its decay products. However, in the case of quarks a special situation arises owing to their confinement of the quarks due to nonzero colour charge. This applies to the top quark, although it decays via weak interaction to a b-quark and a W boson long before



(a) One loop corrections to the Higgs boson due to the top quark



(b) Evolution of Higgs self coupling λ [71]

Figure 2.15: Dependence of λ on m_t

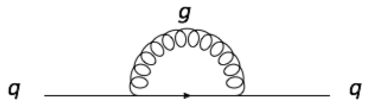


Figure 2.16: Radiative corrections to the quark propagator due to strong interaction

it can hadronize. The b-quark arising from the decay carries the colour charge of the top quark before hadronizing to form a colour-neutral hadron, which requires an additional quark coming from "somewhere else", i.e., not from the top decay. Such long-range effects, which are softer than the threshold of the perturbative calculations, lead to an irreducible ambiguity of the order of $\lambda_{QCD} \approx 0.2$ GeV for quark pole masses. It is possible to define the so-called short-distance masses that could in principle be measured with an arbitrary accuracy and are at the same time well-behaved parameters from a theorists point of view. At short distances and, equivalently, at high energies, particle masses can be calculated within the perturbation series. However, the propagator receives radiative corrections as shown in Figure 2.16 due to strong interaction, leading to divergences. These divergences are then removed using several renormalization procedures, such as $\overline{\text{MS}}$ [72], PS [73], 1S [74, 75], MSR [76] etc.
Simulated events generated using Monte Carlo (MC) techniques constiutte important inputs for many top quark studies. The parameters corresponding to the top quark mass in these MC generators are similar (not necessarily identical) to the pole mass rather than to the renormalized mass. Experimental techniques used so far in the determination of the top quark mass typically involve some MC-based calibration of the measured mass. Furthermore, most measurements are based on the invariant mass of the top quark decay products. For these reasons, the measured top quark mass has been interpreted as the pole mass by the Particle Data Group [16]. For all other quarks, in contrast, the renormalized masses are used.
Based on the measurements from LHC and TeVatron, the latest world average of the top mass is, m_t=173.34±0.27(stat.)±0.71(syst.) GeV [8]. Figure 2.17 shows a comparison of

various top mass results from LHC and TeVatron along with the combination.

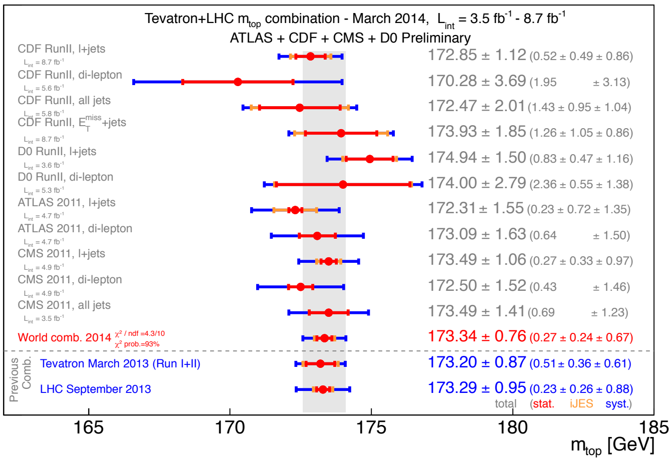


Figure 2.17: Summary of top mass measurements by the ATLAS, CDF, CMS and DØ Collaborations as of March 2014 [8]

Room for top mass measurements in single top events

Most measurements of top quark mass till date have been obtained using $t\bar{t}$ events. Higher particle multiplicity in the final state in case of $t\bar{t}$ events makes it difficult to model colour-flow in the event. The top decay products carrying the colour information of the parent top quark mostly hadronize with additional quark coming from somewhere else. Modeling of colour connection becomes one of the dominant source of systematics in top mass measurements using such events in the lepton+jets final state [31,32] . The distinct production mechanism of the t-channel single top process dictates only a color connection between the top quark and the proton, from which the initial b-quark is coming and not to the whole event like in $t\bar{t}$ production. Such measurements are therefore a useful check for any large unknown systematic effects arising due to the modeling of non-perturbative QCD effects.

2.4 Event signature and background processes

The analyses presented in this thesis focus on the leptonic decay of the top quark singly produced in the t-channel. The top quark almost exclusively decays to a b quark and W boson, which in turn, decays into a charged lepton and the associated neutrino. Therefore, in the final state we have a high-transverse-momentum (p_T), isolated lepton, two or more jets and missing transverse energy due to the escaping neutrino. The b-quark from top decay has high-p_T and hadronizes to form a b-tagged jet, while the light-flavour quark in the final state recoils against the top and acts as a spectator. Considering the dominant 2→3 diagram (see Figure 2.18) relevant for the t-channel single top process at the LHC, the other b-quark in the final state has lower p_T and broad pseudo-rapidity

(η); thus it is mostly either outside detector acceptance or fails selection criteria. Processes that mimic the lepton+jets final state as seen in the t-channel single top signal are considered as backgrounds. The dominant background diagrams are shown in Figure 2.19 and a comparison of cross sections of various SM processes as a function of CM energy ($\sqrt{s}$) at hadron colliders is illustrated in Figure 2.20.

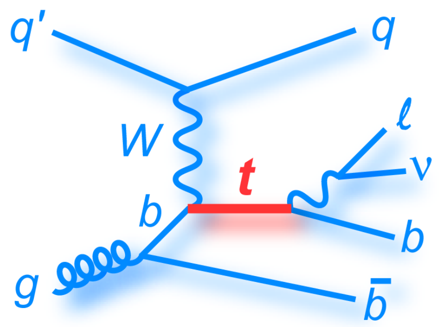


Figure 2.18: Production of top quark in the t-channel and its decay [57]

- **t$\bar{\text{t}}$:** Top-quark pair production, where one top undergoes a leptonic decay and the other decays hadronically, is the most significant background. As discussed in Section 2.2.1, t$\bar{\text{t}}$ production via strong interaction has a large cross section at the LHC, out of which about 45% decay to a semi-leptonic final state. Thus, t$\bar{\text{t}}$ contamination to signal cannot be reduced significantly making it very challenging to separate out single top from t$\bar{\text{t}}$.

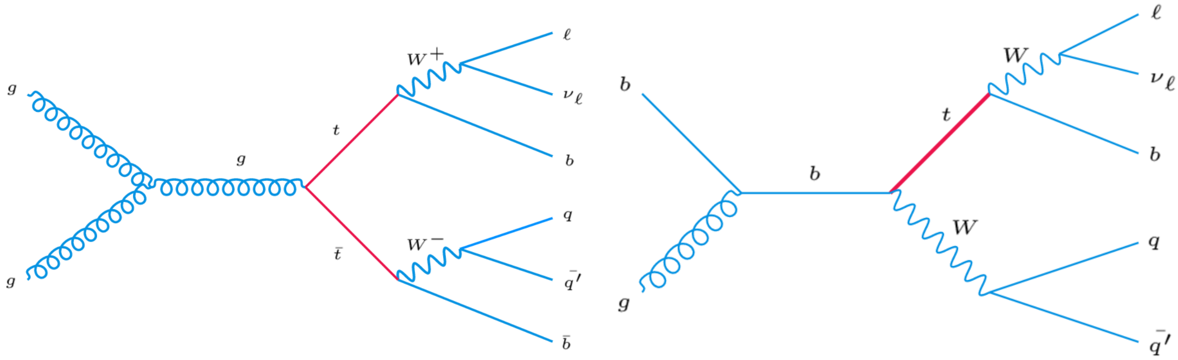
- **tW and s-channel single top:** The tW-channel has the second largest production cross section among the single top processes, of which about 44% undergoes a semi-leptonic decay to produce the same final state as the signal. The s-channel production has the least cross section and has an extra b-jet in the final state in addition to the one from top quark decay which has to either fail selection criteria or be mis-identified, so that the final state mimics that of signal. Therefore, the s-channel background contribution is rather tiny and often neglected.

- **W/Z+jets:** Production of a W or Z bosons in association with jets where the former decays leptonically, constitutes one of the dominant backgrounds. The W produced with either one or more heavy flavour jets or light flavour jets misidentified as b-jets can fake the signal. In case of Z+jets, one of the leptons originating from the Z being not identified or beyond the detector acceptance, results in a similar final state as signal. It is not always possible to significantly reduce W+jets background due to its large production cross section which is several orders of magnitude higher than that of signal. The effect of Z+jets events is slightly less due to their relatively smaller cross section.

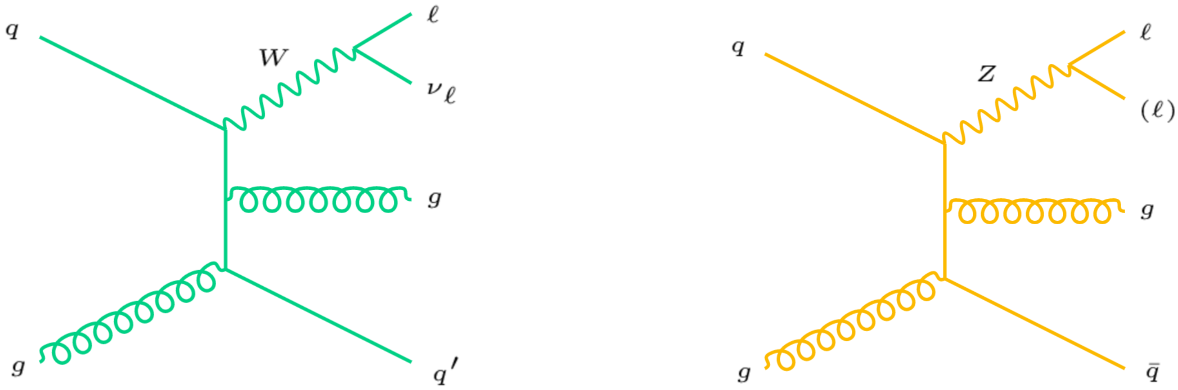
- **QCD multijets:** At the LHC, pp collisions produce with the largest probability multijet events that should hardly pass our selection criteria. Nonetheless, several causes can induce the mis-identification of a hadronic jet as an electron or a muon, giving rise to a "fake lepton": b-jets or long lived mesons like $\pi^{\pm}$ or $\text{K}^{\pm}$

can decay producing genuine leptons that appear to be isolated. Also showers due to π^0 , as well as converted or direct photons, can be reconstructed as electrons. These circumstances may lead to a semileptonic final state analogous to signal. As the multijet production cross section is several order of magnitude higher than signal and the aforementioned mis-reconstruction mechanisms above described are detector-dependent, data-driven methods are the most appropriate way to estimate this background contribution.

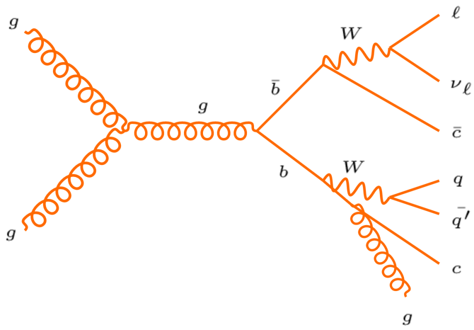
- **Di-boson:** Di-bosons such as WW, WZ and ZZ can also result in "signal-like" final states if one of them undergoes a leptonic decay while the other decays into hadronic final states. Due to lower production cross sections such processes do not constitute as a major background and are often neglected for not having significant contributions.



(a) $t\bar{t}$ (left) and tW-channel (right)



(b) W+jets (left) and Z+jets (right)



(c) QCD multijet

Figure 2.19: Feynman diagrams of the different background processes.

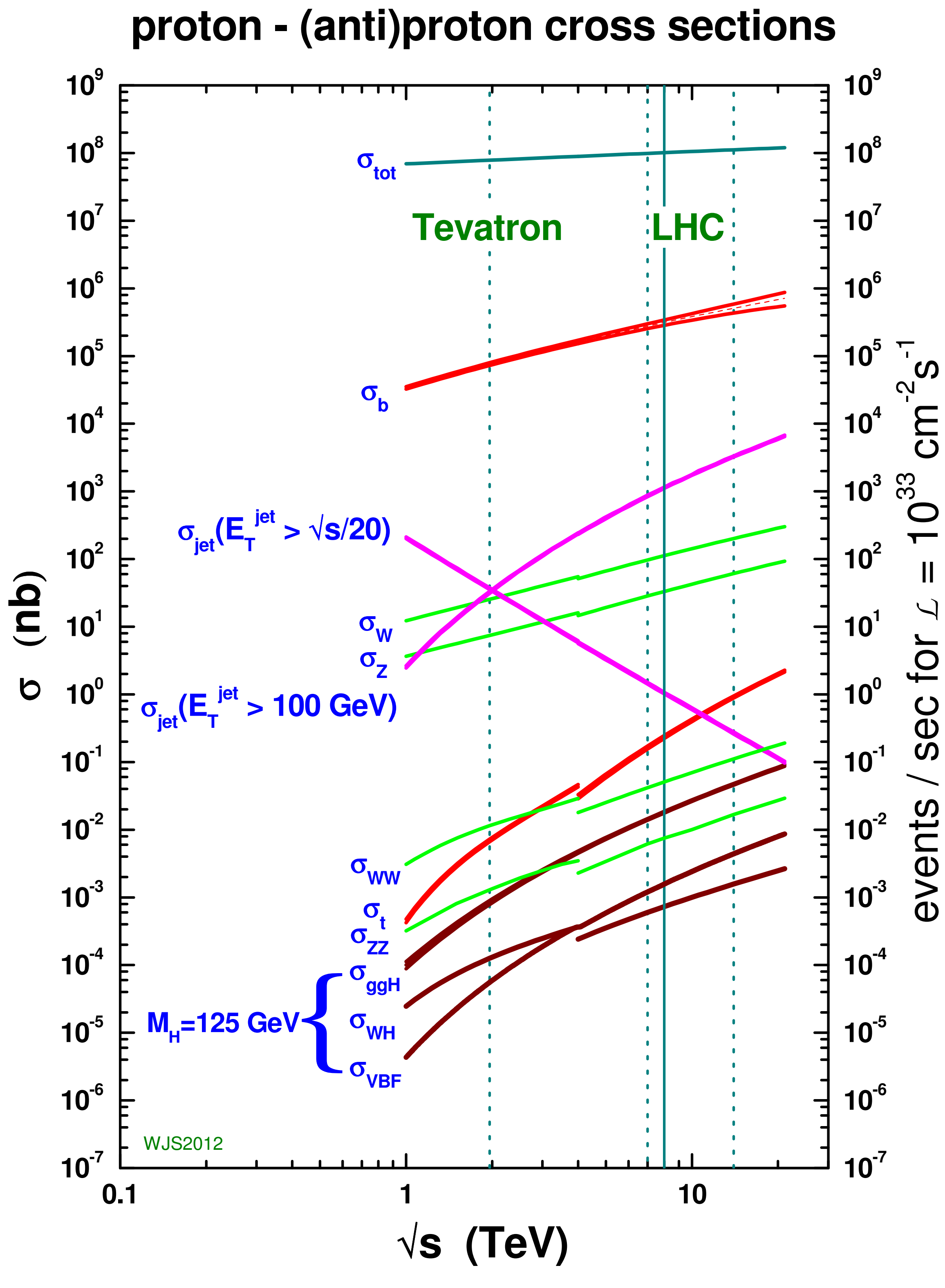


Figure 2.20: Cross sections and expected event yields for various SM processes at hadron colliders [77]

Chapter 3

LHC and CMS

Validity of the SM has been extensively tested for energies up to the electroweak scale, i.e., $\mathcal{O}(100$ GeV), by previous experiments such as Large Electron Positron Collider (LEP), TeVatron and in b-factories like Belle and BaBar. However, the current knowledge of particle physics indicates that the most likely energy frontier for the validity of the SM and hints of new physics is located at the energy scale of $\mathcal{O}(1$ TeV). Though the TeVatron collided protons with anti-protons at an CM energy of 1.96 TeV , the energy available for the elementary partonic collisions is much lower and is determined by the PDFs of the proton and anti-proton. Therefore, a new machine was required which would reach such energies in the elementary collision process. Also, the new machine was required to have an instantaneous luminosity high enough to make the hunt for the higgs boson and eventually for new physics possible in few years of data taking, so to quickly surpass the discovery phase and to proceed to the precision studies of the discovery. An hadronic collider would allow to search in a wide spectrum of energy, automatically scanning the range made available by the PDFs. Keeping these guidelines in mind, the idea of the Large Hadron Collider (LHC) was conceived. At the same time, the experimental conditions at the LHC would require a careful design for the detectors which are to perform physics measurements. Any experiment taking LHC data would have to cope with extremely high rates of events due to the high luminosity conditions and the high CM energy available for collisions. This causes several important consequences on the design of the experiment: detectors, especially the ones closer to the interaction points, have to be robust against extreme radiation conditions. To select interesting events and to reduce the information produced by detectors to a manageable size, a sophisticated system for *triggering* had to be developed. The data acquisition and transfer systems is required to sustain a constant stream of data and transfer it to the storage facilities, ensuring the highest bandwidth available to reduce loss of data. Finally, all reconstruction algorithms for the physics objects present an event is required to be robust against contamination from the *underlying events* that does not directly stem from the hard interaction and from pile-up i.e., effect of the “piling up” of several pp collisions in the same bunch crossing.
This chapter focuses on the design and main characteristics of the LHC machine and the details of the Compact Muon Solenoid (CMS) detector, showing how it has been carefully designed in order to cope with the extreme experimental conditions and to achieve maximum performance in the physics measurements of interest at the LHC.

3.1 The LHC

The LHC is a circular proton-proton (pp) collider operated by the European Centre for Nuclear Research (CERN), capable of accelerating particles at the highest energy ever achieved in a laboratory. It occupies the same 26.7 km circumference tunnel that previously hosted the LEP collider. The tunnel is located very near to the city of Geneva across the French-Swiss border, below the surface of the Earth at a depth between 45 m and 170 m. The LHC is designed to accelerate two counter-circulating proton beams up to $\sqrt{s} = 14$ TeV. It can also handle beams of heavy ions, such as the lead (Pb) nuclei, accelerating them up to $\sqrt{s} = 5.5$ TeV. The beams collide at four interaction points located at underground caverns hosting the four major particle detectors of the LHC program: A Toroidal LHC Apparatus (ATLAS) [78], Compact Muon Solenoid (CMS) [9], Large Hadron Collider beauty (LHCb) [79] and A Large Ion Collider Experiment (ALICE) [80]. The LHC-forward (LHCf) and TOTal, Elastic and diffractive cross-section Measurement (TOTEM) are two comparatively smaller experimental facilities that are located few hundreds of meters away from the ATLAS and CMS interaction points, respectively. In 2010, the Large Hadron Collider (LHC) approved the Monopole and Exotics Detector at the LHC (MoEDAL) to search directly for the hypothetical magnetic monopole. Figure 3.1 shows an aerial view of the underground LHC complex. The major experiments at the LHC are:

- ATLAS and CMS are two general-purpose detectors conceived to explore a wide range of physics, such as precision measurements of the SM including the Higgs and top quarks, search for super-symmetric particles, extra dimensions, dark matter candidates etc.

- LHCb aims mainly at studying matter-antimatter asymmetry through the measurements related to B hadrons.

- ALICE is designed primarily to investigate on the physics of strongly interacting matter at extreme energy densities using the heavy ion collisions.

- TOTEM is built to measure the total elastic and diffractive cross section in pp collisions.

- LHCf is devised to detect particles emitted in the forward direction as a source to understand particle cascades generated by cosmic rays reaching the ground level; by studying how collisions inside the LHC can cause similar cascades of particles in laboratory conditions.

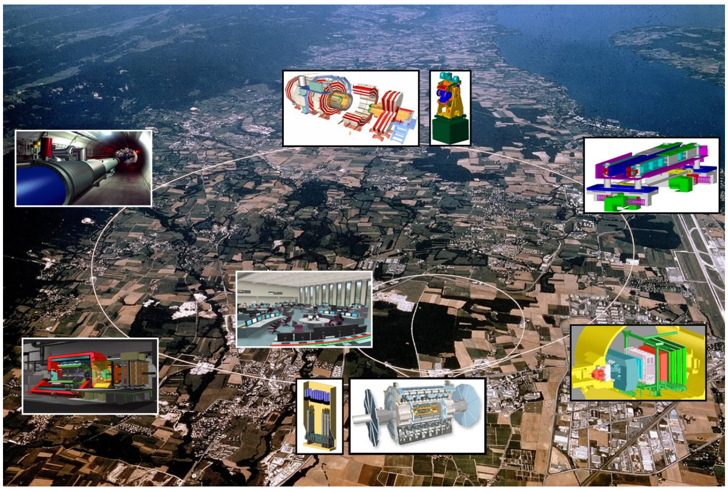


Figure 3.1: An aerial view of the LHC. Starting from the top, anti-clockwise, different LHC experiments: CMS, LHC RF cavity, ALICE, ATLAS, LHCb and TOTEM [81]

The LHC started operation towards the end of 2009. During 2010 and 2011 it produced samples of pp (PbPb) collisions at 7 (2.76)TeV. The so-called “Run-I” data taking period, started in 2009, has ended during early 2013 colliding proton beams even at $\sqrt{s}$=8 TeV. Post Run-I, the machine went into the Long Shutdown 1 (LS1) phase, which was mainly dedicated to the repair and upgrade of certain aspects of the LHC machine and the particle detectors in order to safely accelerate proton beams, close to the design energy. It has restarted again in 2015 and currently is running at $\sqrt{s} = 13$ TeV for the second data taking period marked as “Run-II”. The next target is to gradually move to its design CM energy of 14 TeV. This thesis deals with the data collected at $\sqrt{s} = 13$ TeV in the years 2015 and 2016 during Run-II. Here we describe few main features of the LHC machine [82].

3.1.1 The accelerator

The protons within the beams have to be grouped into bunches and accelerated upto collision energy in multiple stages, with each stage boosting the protons to higher energy before they are actually injected into the LHC ring. Thus the LHC operational cycle for pp collisions follows a long and tedious path. The acceleration procedure is carried out in the following steps.

- Hydrogen (H_2) atoms from a gas bottle are stripped of their electrons via an electrical discharge. The resulting protons are accelerated up to 50MeV in the Linear Accelerator 2 (LINAC2).
- Protons are passed into the Proton Synchrotron Booster (PSB), which accelerates them up to 1.4 GeV and then are introduced to the Proton Synchrotron (PS).
- The PS accelerates the protons up to 25 GeV. Protons come out of the PS packed in bunches separated by 25 nanoseconds each. This posed an unprecedented challenge for the PS.

- The protons then arrive at the Super Proton Synchrotron (SPS) where they are accelerated further to 450 GeV, the LHC injection energy. Injection kicker rise time of SPS separates two consecutive bunches. Such time is 225 ns and is equivalent to 8 bunches.

- SPS injects protons in the LHC ring in groups of 3 and 4 trains of 72 bunches each. Each bunch contains about 1.6×10^{11} protons. The LHC accelerates the bunches up to the energy required for physics (14 TeV by design). LHC injection kicker rise time separates each group of 3/4 train coming from SPS and the following. Such time is 0.975 μs (38 bunches), and rises to 1μs (39 bunches) every three injections. The LHC beam dump kicker rise time at the end of the train is 3 μs, corresponding to 119 bunches. Therefore an overall LHC train has a length corresponding to 3654, while actual proton bunches are 2808. The different stages of the LHC accelerator are illustrated in Figure 3.2.

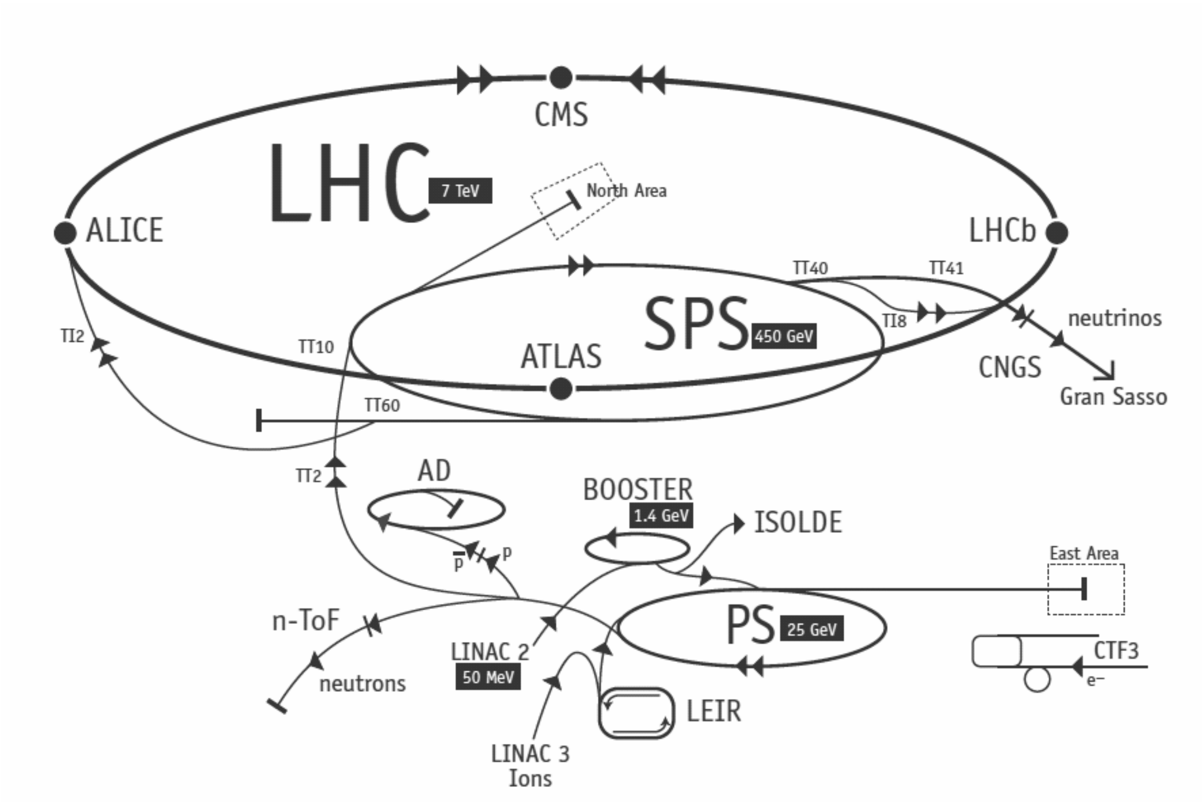


Figure 3.2: The LHC accelarator complex [83]

Radio Frequency Cavities

To accelerate the particles upto desired CM energy and then to maintain them circulating around the ring at the same energy, 16 radio frequency (RF) cavities housed in 4 cylindrical refrigerators are used. The RF modules are molded to a specific size and shape, so that electromagnetic waves initially set up by a power generator become resonant and build up inside the cavity. The protons and lead-ions passing through feel the overall force and direction of the resulting electromagnetic field, and are therefore pushed forwards along the accelerator. More precisely, since the field in a RF cavity oscillates at a given frequency, particles can be accelerated or decelerated depending on their arrival time. In this way the beam is sorted into discrete bunches.

3.1.2 Magnets and Cryogenics

Since the LHC collides protons with protons, two separate beam pipes with opposite magnetic fields were needed. To keep the 7 TeV proton beams onto their orbit, the required magnetic field of 8.33 T is supplied using the niobium-titanium dipole magnet system that operate at a temperature of 1.9 K. Also the whole magnet system had to fit into the limited inner diameter of the tunnel, 3.7 m, which led to the choice of a twin-bore magnet design as can be seen in Figure 3.3a. In total 1232 of these ~15 m long dipole magnets are used, each weighing about 27.5 tons. The focusing of the bunches to a small size is mainly provided by 392 quadrupoles (Figure 3.3b), made up by four magnetic poles arranged symmetrically to squeeze the beams either vertically or horizontally; their shape will also be corrected by approximately 3700 additional multipole magnets. Finally, the most complex insertion magnets are the 8 inner triplets that will further “compress” the bunches and bring them into collision at the four interaction points (IP).

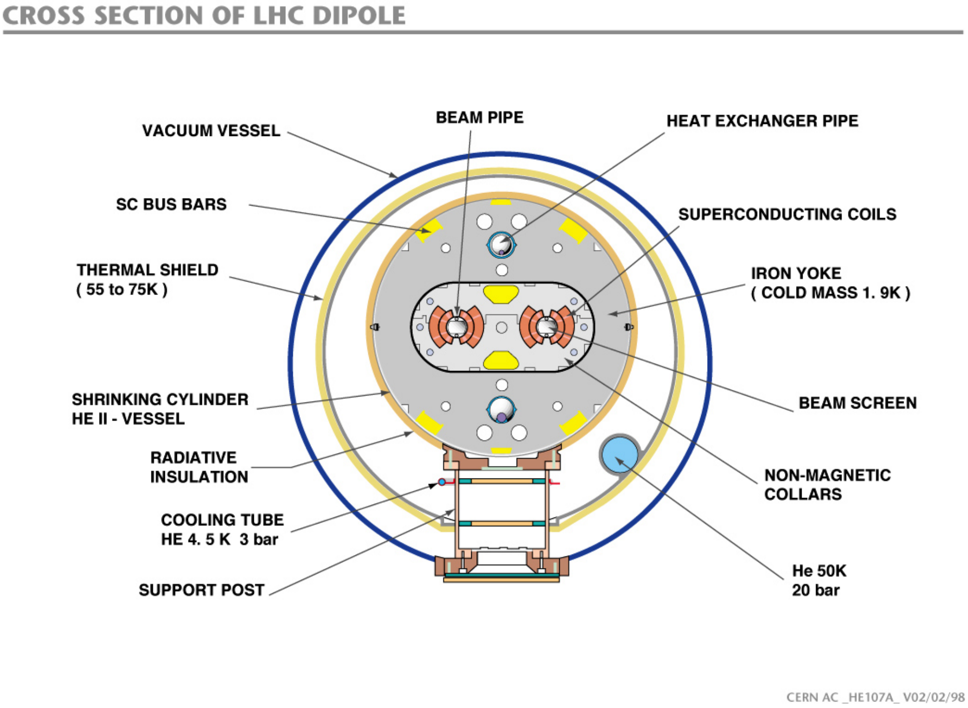
The LHC magnets are cooled with pressurized superfluid helium (He), which has some interesting properties that make it a unique material. Best known is the very low viscosity which allows it to permeate the smallest cracks and in particular enabling the fluid to be in contact with the strands of the superconductor. The large specific heat (typically 105 times that of the superconductor per unit mass), combined with the enormous heat conductivity at moderate flux (3000 times that of cryogenic-grade OFHC copper, peaking at 1.9 K) can have a powerful stabilizing action on thermal disturbances. The machine is cooled using 8 cryogenics plants located in pairs at the even points. At the tunnel level the conventional refrigerators are supplemented by cold compressors. These multi-stage axial centrifugal compressors pump the cold He gas producing the 15 mbar pressure in the linear heat-exchangers inside the magnets in order to produce the primary superfluid. The connection to the magnets is made through a cryogenic distribution line running in the tunnel parallel to the machine.

3.1.3 Luminosity

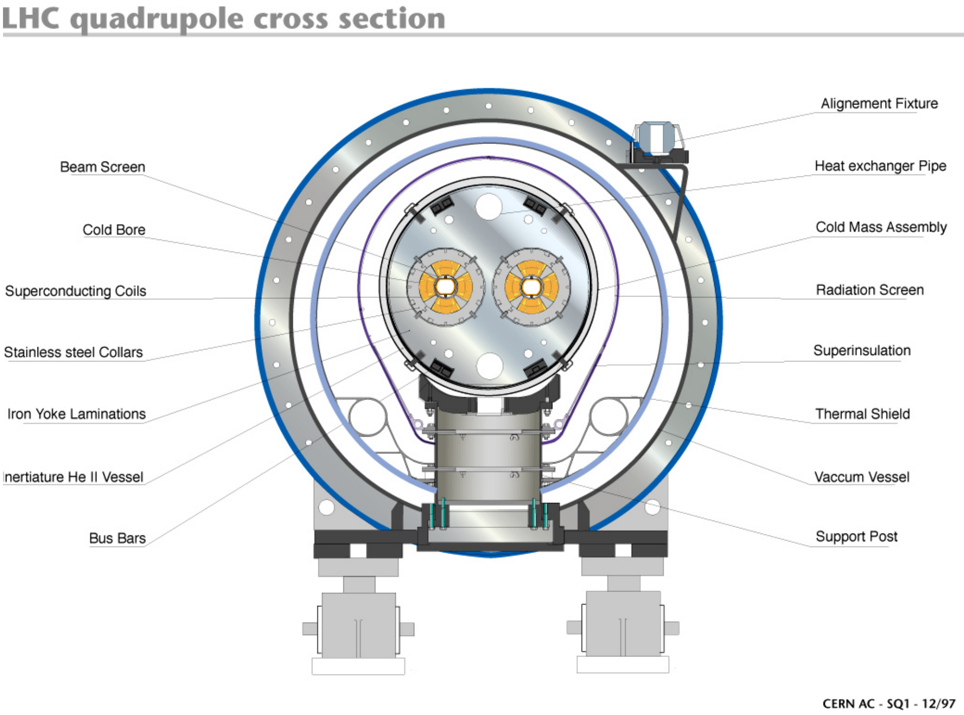
The main goal of the LHC is to search for physics beyond the SM with highest possible energies in the laboratory (14 TeV). As described in Eq. (4.1), the number of interesting events produced per second in an LHC collision is the product of the cross section and the luminosity. The luminosity, which can be controlled unlike the cross section, depends on various beam parameters and can be expressed as:

$$\mathcal{L} = \frac{\mathrm{N_b^2 n_b f_{rev}} \gamma_\mathrm{r}}{4\pi \epsilon_n \beta^*} \mathrm{F} \tag{3.1}$$

where $\mathrm{N_b}$ is the number of particles per bunch, $\mathrm{n_b}$ is the number of bunches per beam, $\mathrm{f_{rev}}$ is the revolution frequency of the beam (11.245 kHz),γ_r is the relativistic boost factor, ϵ_n is the normalized transverse beam emittance, β^* is the beta function at the collision

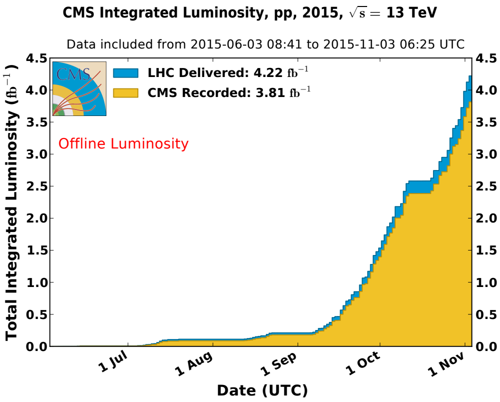


(a) Dipole magnets [84]

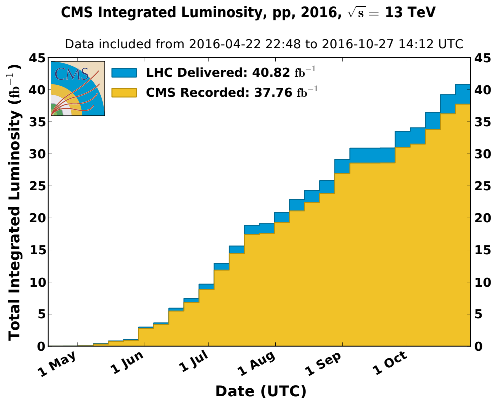


(b) Quadrupole magnets [85]

Figure 3.3: Magnets used in LHC for beam maneuvering.



(a) 2015



(b) 2016

Figure 3.4: Integrated Luminosities delivered by LHC and recorded by CMS in Run-II [88]

point, and F is the luminosity reduction factor due to the finite crossing angle at the IP. The F value is given by

$$F = [1 + (\frac{\theta_c \sigma_z}{2\sigma^*})^2]^{-\frac{1}{2}} \tag{3.2}$$

where with θ_c is the crossing angle at the IP, σ_z is the RMS bunch length, and σ^* is the transverse RMS beam size at the IP. The above expression assumes circular beams, with $\sigma_z << \beta^*$, and with same parameters for both beams. To explore the rare events at LHC, both high energy (to produce new resonances on shell) and high luminosity (to produce in good numbers) are required. By tuning various machine parameters, the LHC luminosity was steadily increased during the 2015-2016 as shown in Figure 3.4. The corresponding luminosities recorded by the CMS detector were $\sim$3.8 fb^{-1} and $\sim$38 fb^{-1} respectively in 2015 and 2016.
For measuring the luminosity at its IP, the CMS uses two methods: the Hadron Forward (HF) method is based on the determination of particle flux in the hadron forward calorimeter while the second method relies on the number of vertices reconstructed in the CMS pixel tracker. In the vast majority of run periods, the two measurements agree. The absolute luminosity measurement uses Van-Der-Meer scans in separate runs [86]. The relaive uncertainty on the measured luminosity is 2.3% for the 2015 dataset [87].

3.2 CMS detector

The Compact Muon Solenoid (CMS) detector is one of the two general purpose detectors installed at LHC. It is located at IP5, near the village of Cessy in France, between Lake Geneva and the Jura mountains. CMS allows both pp and heavy-ion collisions. The three striking features of CMS, as suggested by the name, are:

- **C**ompact design; all the detector material is contained within 21 m in length and 15 m is height.
- Highly Accurate **M**uon detection system

- The most powerful **S**olenoid magnet ever made which produces a homogeneous magnetic field of 3.8 T over the entire volume of inner detector.

Identification and precise measurements of electrons, photons, muons, electromagnetic and hadronic showers and estimation of missing energy are necessary to cover a broad variety of physics phenomena and CMS boasts of several unique features, such as:

- A silicon (Si) based inner tracking system with unique capabilities to measure particle momenta over a wide range
- Measurement of electromagnetic shower with optimal resolution for high energy electrons and photons
- Hermetic design of the hadronic calorimeter ensuring better reconstruction of hadronic jets and missing energy
- Strong magnetic field and efficient muon system for the best possible momentum measurement of high-energy muons

Figure 3.5 shows an inside view of the various sub-detectors while Figure 3.6 illustrates the arrangement of the sub-detectors in a transverse section. The sub-detectors are organized in the following order when go radially outward from the centre: Si-based inner tracker, electromagnetic calorimeter, hadron calorimeter, superconducting solenoid and muon system. In this section, all these various sub-detector components of the CMS detector will be discussed one by one in some details.

3.2.1 CMS Coordinate System

Before proceeding any further, let's introduce the coordinate system followed by the LHC and all of its experiments. For any given experiment at the LHC, the centre of the reference frame corresponds to the IP and the $\hat{z}$ axis corresponds to the tangent to the beam at the IP, in the anticlockwise direction. The $\hat{x}$ axis is perpendicular to the $\hat{z}$ axis and virtually joins the IP with the centre of the LHC ring, pointing towards the centre. The $\hat{y}$ axis is perpendicular to both $\hat{z}$ and $\hat{x}$ axes, and points towards the ground surface. These coordinates can then be used to define the following quantities:

1. $r = \sqrt{x^2 + y^2}$, the distance relative to the beam line
2. $\phi = \tan^{-1}(\frac{y}{x})$, the azimuthal angle
3. $\eta = -ln \tan(\frac{\theta}{2})$, the pseudo-rapidity

Each point in space is described by the (r, ϕ, η) triplet of coordinates and in particular, the (η, ϕ) coordinate set describes the direction of a particle produced at the IP.
In a hadron collider such as the LHC, the momentum of the initial state of the interacting partons is not precisely known. However, it can be well approximated to zero along the transverse direction with respect to the beam, as the transverse components of the momentum of the partons are negligible with respect to the longitudinal ones. As a consequence, physics analyses take into account a balance only in the transverse component of the interesting kinematic quantities, e.g., the transverse momentum defined as

$$\mathrm{p_T} = \mathrm{p} \sin\theta \tag{3.3}$$

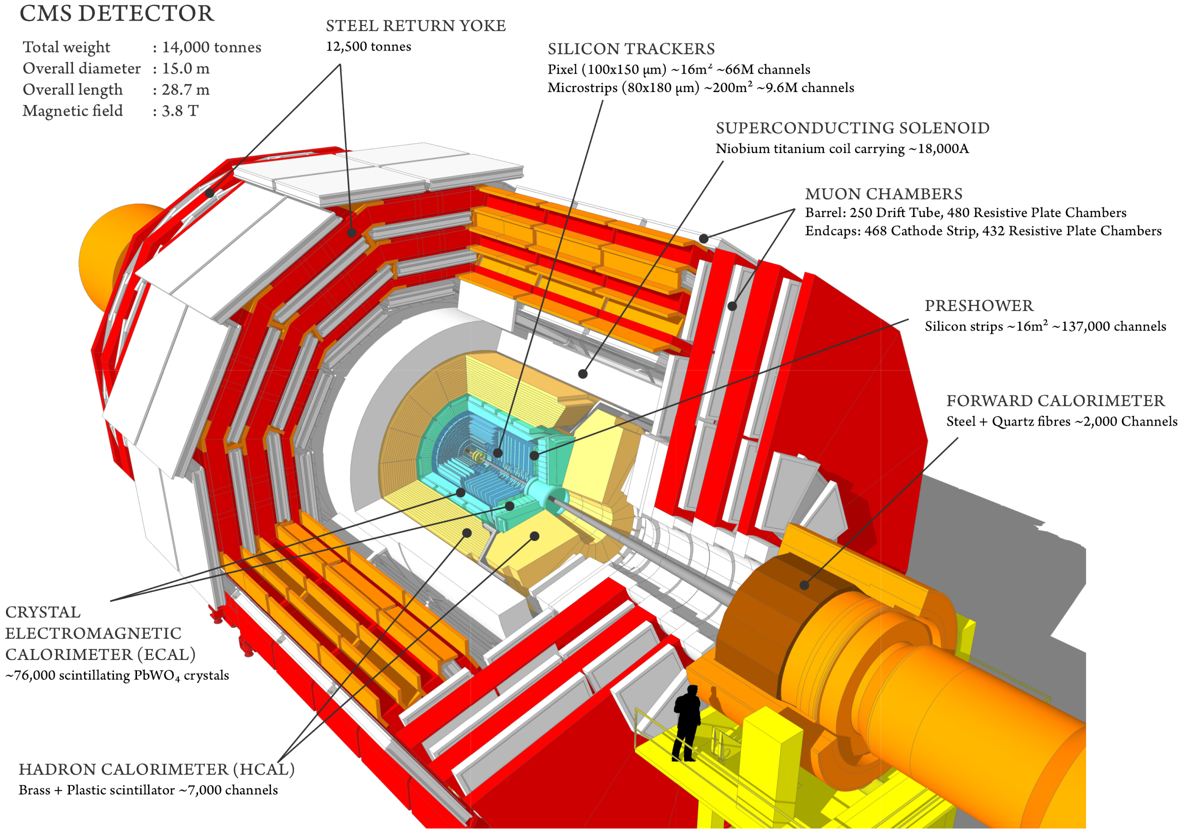


Figure 3.5: Inside view of the CMS detector [89]

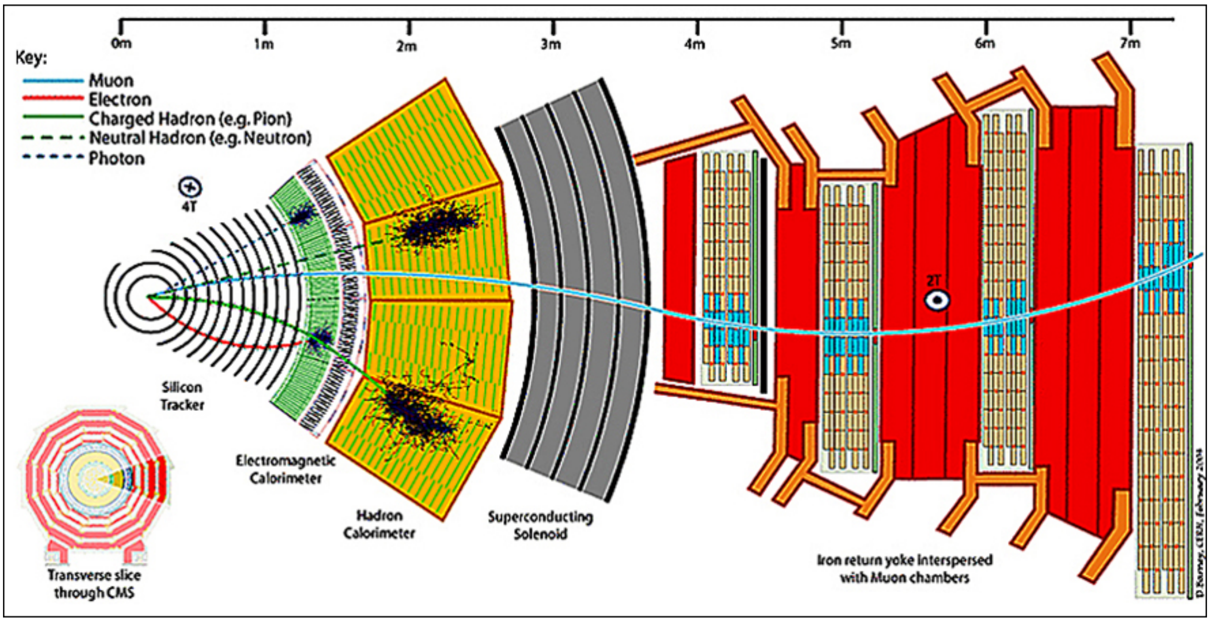


Figure 3.6: Transverse view of the CMS detector [90]

In hadron-collider experiments, we use pseudo-rapidity (η) most of the time as the difference between rapidities (Δy) is invariant under the longitudinal Lorentz boost and η

itself is independent of the mass of the particle. In the massless limit, both η and y are same. The angular distance between to particles is defined as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{3.4}$$

The 4-momenta of a particle is often referred by the components $(E, \mathrm{p_T}, \eta, \phi)$ or $(\mathrm{p_T}, \eta, \phi, m)$ instead of the standard Cartesian components (E, p_x, p_y, p_z).

3.2.2 Superconducting Solenoid

In order to keep the muon system and the detector itself reasonably compact, CMS has been designed around a superconducting solenoid that is capable of producing a 3.8T magnetic field, providing the necessary bending power required to achieve a good momentum resolution for final-state particles. As our main goal is to get an accurate measurement of momentum, the magnet was a driving element in the design of the detector. The magnet is a 12.9 m long solenoid made of a high-purity aluminum-stabilized conductor, which is maintained to be cooled indirectly. The conductor has a cross-section of 64×22 mm^2, through which 20 kA of current flows. The solenoid is located outside the HCAL, and provides a magnetic field of up to 3.8T inside the cylindrical volume.

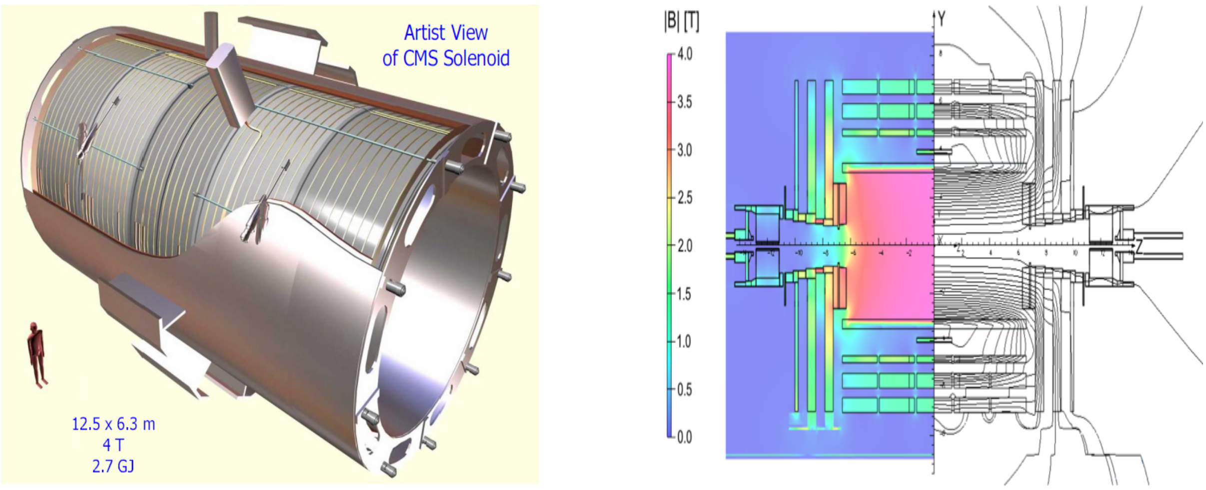


Figure 3.7: An artistic view of CMS solenoid magnet (left) and magnetic field of CMS experiment (right) [91, 92]

3.2.3 Inner Tracker

As shown in Figure 3.6, tracking system [93] [94] is the innermost part of the CMS detector surrounding the thin beryllium beam pipe. It is 5.8 m in length and 2.5 m in diameter centered around the IP. It provides an efficient and precise reconstruction of charged particles tracks (and, hence their momenta) coming from the IP, with a $\mathrm{p_T} > 1$ GeV and within $|\eta| < 2.5$, radiation heavy region of the detector. Additionally, the tracking system precisely reconstructs secondary decay vertices for an efficient identification of heavy flavour objects such as b-jet, τ etc. At LHC design luminosity of $10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$

there are more than 20 pp interactions per each bunch crossing which results in $\mathcal{O}(1000)$ particles crossing the tracker region every 25ns. This sets the requirements for the tracking system:

- High granularity in order to discriminate even adjacent trajectories so that ambiguities can be easily removed at reconstruction level.
- Fast response time to cope with the high bunch crossings frequencies.
- Radiation robustness: the intense particle flux causes damage to the detectors over time, so it was compelling to design a tracking system which could operate under LHC conditions and stay efficient over the entire lifetime of the experiment.

As a consequence of the first two requirements high power density of on-detector electronics is needed, and therefore an efficient cooling system is also necessary. The detector design had to cope with those issues while at the same time keeping at the minimum possible the amount of material present in the tracking volume to limit the effect of particle interactions (bremstrahlung, multiple scattering, photon conversion, nuclear interactions). To deal with these complexities, an all-silicon-based tracker was built, *first of its kind in the history of experimental particle physics.*
The inner tracking detector constitutes two subsystems: Pixel detector and Strip tracker as shown in Figure 3.8. The detailed description of design and functionality of these subsystems are as follows.

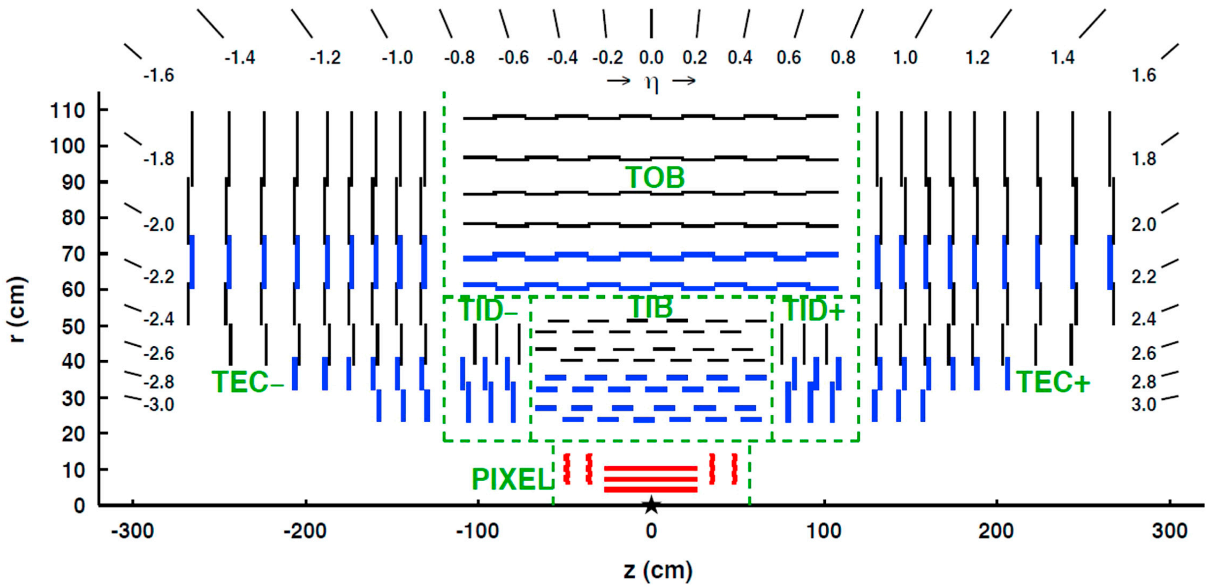


Figure 3.8: A schematic diagram of the CMS tracking detector along r-z plane [95]

Pixel Detector

The pixel detector (Figure 3.9) occupies the innermost region of the system, up to a maximum distance of 10 (15) cm in barrel (endcap) from the IP. The best possible spatial resolution in the pixel detectors is needed to cope with a flux of particles as high as $10^7 \text{cm}^{-2}\text{s}^{-1}$ as well as to discriminate against pileup. Each pixel of the detector has a size of 100 μm × 150 μm. The central part of the detector,i.e., barrel pixel (BPix) has a cylindrical shape with its axis coinciding with the beam, and consists of three layers of pixels. In the forward and backward regions, there are two endcaps (FPix) each comprising two layers of pixels. The pixel detector is made of 768 modules in the barrel and 672 modules in the endcaps, summing up to a total of 66 million pixels. It is capable of detecting particles having $|\eta| \leq 2.4$ with an average spatial resolution of 10 μm along the r-ϕ plane and of 20 μm along the z direction. Such high spatial resolution is crucial for a precise determination of vertex positions and to achieve the required momentum resolution. The latter is as good as 1% for a track momentum of 10 GeV. After the end of 2016, a new layer of pixel detectors in barrel as well as in endcaps has been added in order to cope with increasing instantaneous luminosity scenario in near future.

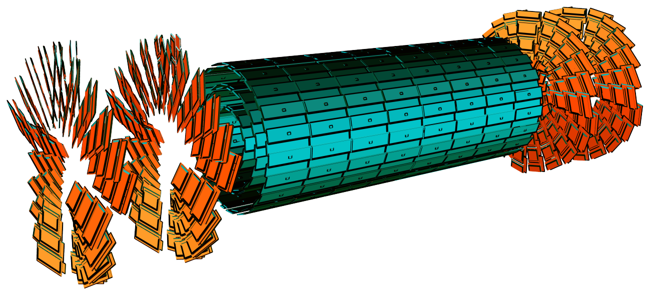


Figure 3.9: CMS Pixel detector [96]

Strip Detector

The strip tracker (Figure 3.10) measures the track trajectories with a little reduced resolution ($\leq$ 23 μm), reflecting the smaller particle flux at larger distances from the IP. It consists of 10 barrel layers (4 tracker inner barrel (TIB) and 6 tracker outer barrel (TOB) layers), 3 tracker inner discs (TID) and 9 tracker endcap (TEC) disks. Layers 1 and 2 (counted from inside) of TIB and TOB as well as TID rings 1 and 2 and TEC rings 1, 2 and 5 have back-to-back mounted modules with 100 μrad tilt, providing a stereo coordinate measurement (both r-ϕ and z). Modules of TIB, TID and the four innermost TEC rings are made of 300μm thick sensors. The entire TOB and the three outermost rings of TEC consist of two sensors of 500 μm thickness, bonded together. The one-dimensional strip layout allows the tracker to be constructed at a lower cost. In total, 24240 silicon micro-strip sensors cover an area of 200 m^2, making it the largest Si-based detector ever built. The tracker operates at a temperature of around -20^0C. Based on simulation studies, the material budget for tracker is estimated as a function of η in terms of the radiation length X=X_0 and interaction length λ/λ_0. The distributions of these two quantities, enlightening the contributions coming from different parts of the tracker, are presented in Figure 3.11.

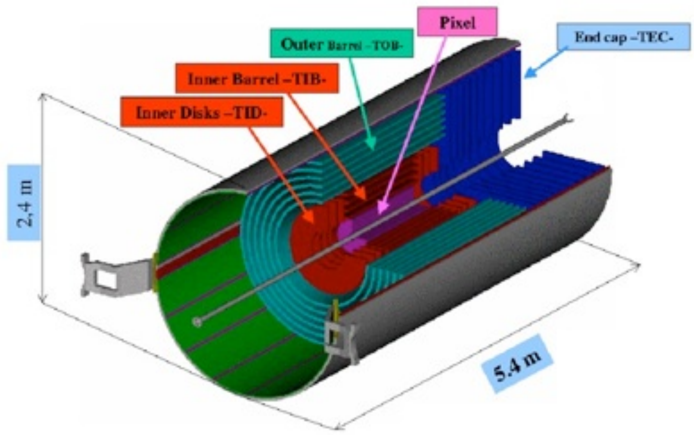


Figure 3.10: CMS Strip detectors [97]

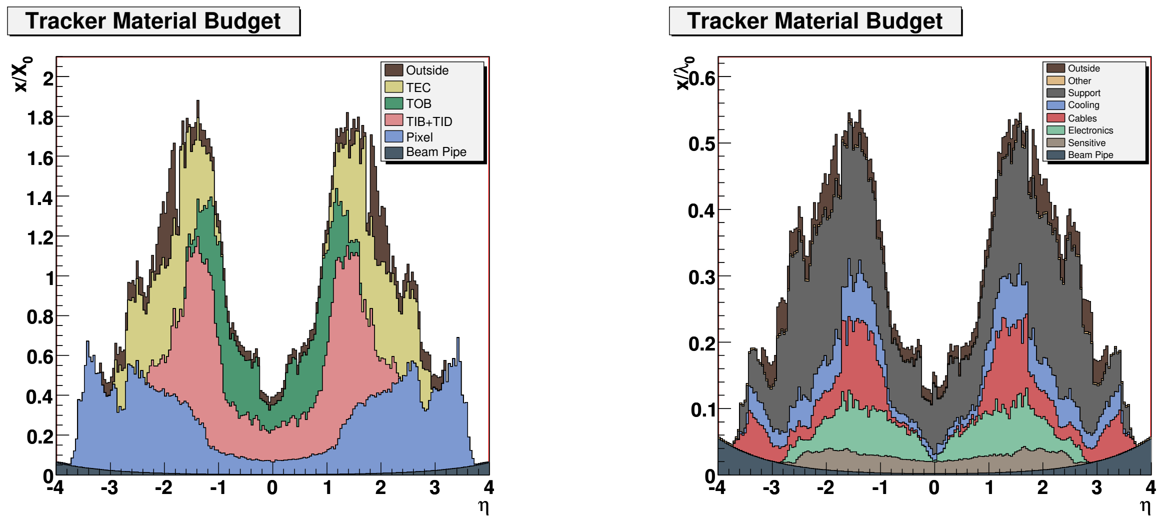
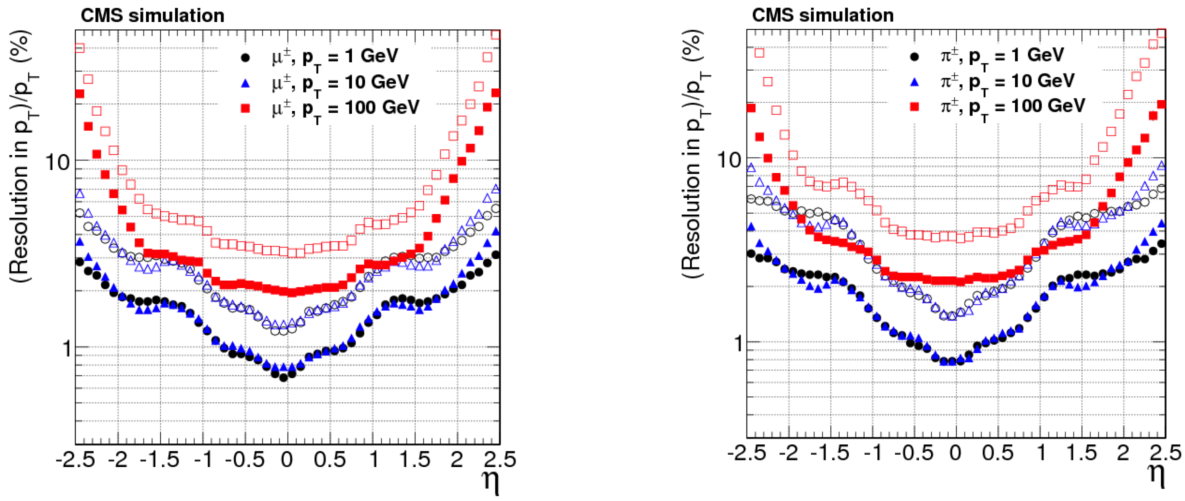


Figure 3.11: Radiation length (left) and Interraction length (right) of the tracker as a function of η [98]

Tracker Performance

For $t\bar{t}$ events under typical 2011 pileup conditions, the average track-reconstruction efficiency for promptly-produced charged particles with transverse momenta of $p_T > 0.9$ GeV is 94% for pseudorapidities of $|\eta| < 0.9$ and 85% for $0.9 < |\eta| < 2.5$ [99]. The inefficiency is caused mainly by hadrons that undergo nuclear interactions in the tracker material. The transverse momentum resolution δ p_T / p_T for muons and pions were measured for $p_T = 1$, 10, and 100 GeV as a function of η and are shown in Figure 3.12.

The position resolution achieved for reconstructed primary vertices that correspond to interesting pp collisions is 10-12 microns in each of the three spatial dimensions.



(a) μ (b) π

Figure 3.12: p_T resolution as a function of η [99]

3.2.4 Electromagnetic Calorimeter

The CMS Electromagnetic calorimeter (ECAL) is a homogeneous calorimeter made out of 75848 lead tungstate $PbWO_4$ crystals, providing excellent energy and position resolution for electron and photons. When a particle, such as a photon or an electron, enters the ECAL, starts to lose its energy by producing electromagnetic shower, caused by the successive processes of pair production and bremsstrahlung which results in the production of large number of secondary particles. At the end, the energy of the particle is deposited in the calorimeter material mostly via ionization and photoelectric effect. The main features of $PbWO_4$ crystals are as follows.

- High density, $\rho = 8.3$ g/cm^3 to sustain high levels of radiation
- Short radiation length, $X_0 = 0.89$ cm, so that the 99% of the electromagnetic shower is contained within its depth $\sim 25\ X_0$.
- Small Moliere radius, $R_M = 2.2$ cm, so that transverse profile of electromagnetic shower is contained within a single crystal.
- Emits scintillation light in blue-green region of the spectra with maximum at $\lambda \approx$ 440 nm.
- Fast response due to very small scintillation light decay time constant ≈ 10 ns
- Ability to sustain high magnetic field (3.8 T)
- Mechanically stable to fit into a compact calorimeter with fine granularity

The ECAL comprises of two parts: the barrel (EB) up to $|\eta| = 1.479$ and the two endcaps (EE), covering the $|\eta|$ within 1.479 to 3.0. The scintillation light from the $PbWO_4$ crystals are read by avalanche photo-diodes (APD) in the barrel region while vacuum photo-triodes (VPT) are used in endcap to be more radiation tolerant. In the EB, crystals are grouped into modules each containing 400 -500 crystals.4 such modules along η makes a "Supermodule" which subtends 20^0 angle in ϕ. There are 36 such Supermodules in EB. EE is made up of 4 half-discs or "Dees". 268 identical "Supercrystals" (each consisting of 36 PbWO4 crystals) are used to cover each endcap (2 Dee-s) with a further 64 used to complete the inner and outer perimeter. For a better discrimination of photons against neutral pions, a preshower device (ES) is mounted in front of the EE. Silicon strip sensors with a resolution of 2 mm, placed behind two planes of lead, are able to distinguish single photons from π^0-s decaying into photon pairs that would not be possible with a typical crystal size of 30 mm. Figure 3.13 shows different components of the ECAL.

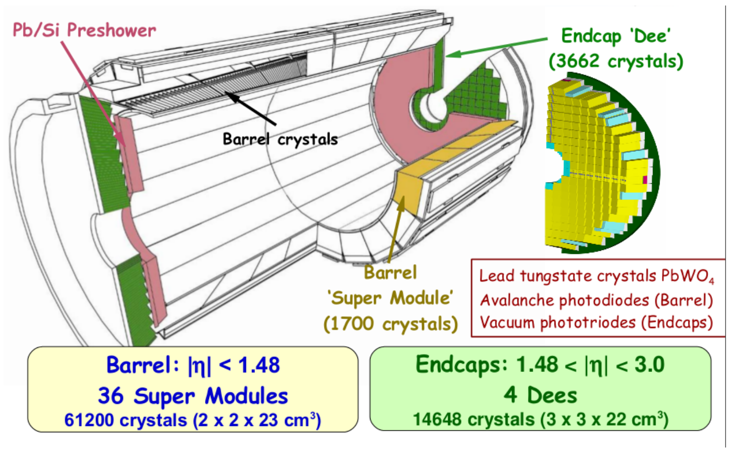


Figure 3.13: Layout of several components of CMS ECAL [100]

ECAL performance

The energy resolution of the CMS ECAL can be parametrized as

$$\frac{\sigma_E}{E} = \frac{S}{\sqrt{E}} \oplus \frac{N}{E} \oplus C \tag{3.5}$$

where the energy E is measured in GeV, S is the stochastic term, N is the noise term and C is the constant term. The stochastic term accounts for the combination of event-to-event fluctuations in the lateral shower containment, photo-statistics contribution and fluctuations in the energy deposited in the preshower absorber (where present) with respect to what is measured in the preshower silicon detector. The noise term conatins contribution due to electronics noise, digitization noise and event pile-up. The constant term arises due to : non-uniformity of the longitudinal light collection, inter-calibration errors, leakage of energy from the back of the crystal, imperfections in mechanical structure, temperature gradient etc. From test beam measurements [101] with 20-250 GeV electrons where the incident electron energy is reconstructed by summing 3×3 crystals,

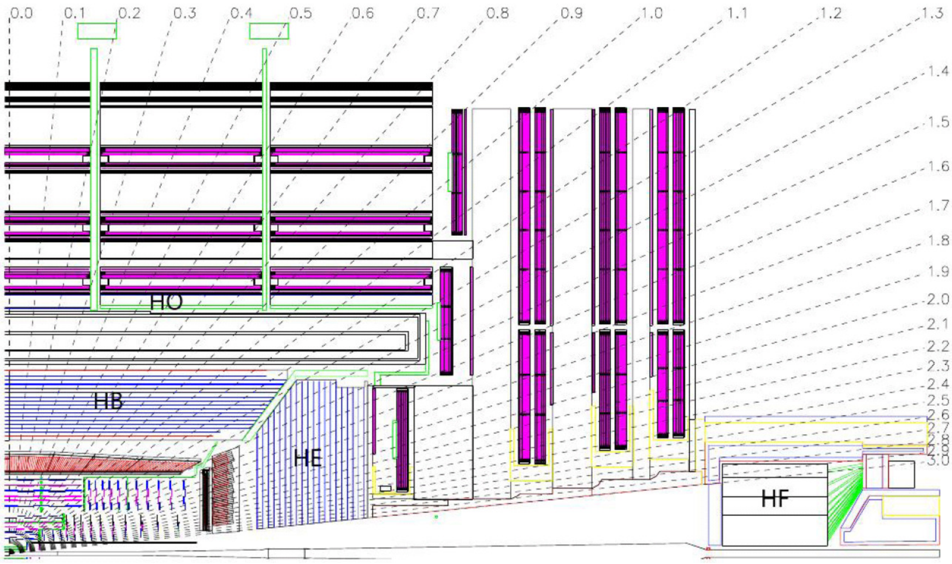


Figure 3.14: A schematic diagram of CMS HCAL showing its components [103]

a typical energy resolution was found to be

$$\frac{\sigma_{\mathrm{E}}}{\mathrm{E}} = \frac{2.8\%}{\sqrt{\mathrm{E(GeV)}}} \oplus \frac{12\%}{\mathrm{E(GeV)}} \oplus 0.3\% \tag{3.6}$$

3.2.5 Hadron Calorimeter

The CMS detector is designed to study a wide range of high-energy processes, resulting in a variety of final state topology involving hadronic jets, neutrinos or also exotic particles manifesting through missing energy in the transverse plane of the detector. It is therefore crucial to have a precise measurement of jets energy and ensure the hermeticity of the detector to reduce the uncertainty on transverse energy lost due to cracks in the detector. The CMS hadronic calorimeter (HCAL) [102] is a sampling calorimeter surrounding the aforementioned parts of the detector. It reconstructs both charged and neutral hadrons produced during pp collisions. The compact design is a consequence of the overall detector layout that permits the placement of the main parts of both calorimeters (ECAL and HCAL) inside the solenoid. Brass (70%Cu and 30%Zn) has been chosen as the absorber material for the HCAL, as it is non-magnetic and has a relatively short interaction length of $\lambda_{\mathrm{I}} = 16$ cm. The large fraction of passive material is cost effective. On the other hand, it requires detailed in-situ calibrations and the energy resolution is reduced compared to what is possible with homogeneous calorimeters. The entire subsystem is divided into four parts, as shown in Figure 3.14, explained as follows.

Hadron Barrel (HB)

The Hadronic Calorimeter Barrel (HB) covers the pseudo-rapidity range $|\eta| < 1.3$. The HB is divided in two half-barrel sections (HB+ and HB-) consisting of 36 identical azimuthal wedges made out of flat brass absorber plates. The plates geometrical configuration that contains no projective dead material for the full radial extent of a wedge.

The transverse granularity of the wedges $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. The total absorber thickness at $|\eta| = 0$ is $5.82\lambda_{\mathrm{I}}$ but increases with $|\eta|$up to $10.6\lambda_{\mathrm{I}}$ at $|\eta| = 1.3$. Plastic scintillator of 3.7 mm thickness is used as active material for its long term stability and moderate radiation hardness. The light output from the scintillator tiles are guided using wavelength shifting fibers to the hybrid photo-diodes (HPD) for detecting the hadronic showers.

Hadron Endcap (HE)

The HE has 14 additional calorimeter towers, covering the pseudo-rapidity region 1.3 $< |\eta| < 3.0$. Its ϕ-segmentation is 5^0 near HB that increases to 10^0 near the outermost towers. The η-segmentation increases to 0.35 at the highest $|\eta|$ value. The HE being closer to the beam pipe is a high-radiation region containing about 34% of the particles produced in the final state. High radiation tolerance, non magnetic properties, sufficiently high radiation length, good mechanical properties at a reasonable cost prompted the choice for C26000 cartridge brass as the absorber material. The brass plates, arranged in a staggered geometry with no projective dead material, are 79 mm thick with 9 mm gaps for scintillators accommodation and result in about $10\lambda_{\mathrm{I}}$ (including the ECAL crystals). The light output from the scintillator is guided using wavelength shifting fibers to multipixel HPDs.

Hadron Outer (HO)

As the energetic hadrons are unlikely to be fully contained within the limited range of the central part of the HCAL, a small extension is placed outside the solenoid magnet, called the outer hadron calorimeter (HO). The HO is used as a tail catcher to measure the energy of hadron showers penetrating the magnetic coil and thereby covering for very high energy particles and late starting showers. It stretches up to a pseudo-rapidity range of $|\eta| = 1.26$, which is comparable to the $|\eta|$ range of HB. The magnetic field, outside the vacuum tank of the solenoid, is returned through an iron yoke organized in five rings (the tail catcher iron), in which the HO is placed as the first sensitive layer essentially using the iron yoke as the absorber material.

Hadron Forward (HF)

To cover even higher pseudorapidity values in the range $2.8 \leq |\eta| \leq 5.2$, the HF is located at $|z| = 11.2$m from the IP close to the beam pipe. Jets in the very forward (backward) direction as well as the hadronization products of beam remnants are detected with this system. A square grid of quartz fibers is inserted into grooves in the steel plates that serve as an absorber material. The Cherenkov light, emitted when charged particles traverse through the quartz fibers, produces signal in the photo multipliers. The HF is 1.65m long and divided into 900 towers with an average granularity of $\Delta\eta \times \Delta\phi = 0.175 \times 0.175$. There are two types of quartz fibers, namely long (1.65m) and short (1.43m), in each HF tower in order to identify electromagnetic (signal in long fiber only) and hadronic showers (signal in long and short fibers) in HF.

HCAL performance

The hadronic energy resolution for the barrel HCAL and ECAL combination can be parametrized as

$$\frac{\sigma_{\mathrm{E}}}{\mathrm{E}} = \frac{\mathrm{a}}{\sqrt{\mathrm{E}}} \oplus b \tag{3.7}$$

where, "a" is the stochastic term and "b" is the constant term. From LHC test beam measurements [104], a = (0.847 ± 0.016) $\sqrt{\mathrm{GeV}}$ and b = 0.074 ± 0.008. For HF, the measured values are: a = 1.98 $\sqrt{\mathrm{GeV}}$ and b = 0.09.

3.2.6 Muon System

Muon detection is a very important tool to recognize signatures of interesting processes over the very high background rate expected at the LHC with full luminosity. Precise and robust muon measurement has been a central theme of CMS. Muons characteristics make them optimal candidates for precision measurements: muons of energies below the scale of $\mathcal{O}$(few TeVs) lose energy mainly due to ionization through the detectors, so that a sufficiently energetic muon ($\sim$ 5GeV) can reach the outside of the CMS superconducting coil. Also, since muons are charged particles, precision measurement of their momentum can be performed in the tracking systems. Those features make muon detection a powerful tool for recognizing signatures of interesting processes over the very high background rate expected at the LHC.

The muon system provides muon identification, momentum measurement and triggering. Good momentum resolution and trigger capability are achievable with the high solenoidal magnetic field and the flux return yoke (this serving also as hadron absorber for muon identification), while the identification of muons is provided through three different types of gaseous particle detectors installed depending on their radiation hardness and triggering performance, namely, drift tubes (DT), cathode strip chambers (CSC) and resistive plate chambers (RPC), as shown in Figure 3.15. Due to the shape of the solenoid magnet, the muon system was naturally driven to have a cylindrical barrel section and 2 planar endcap regions. Various components of muon system are described below. Figure 3.16 shows the individual components.

Drift Tubes

The muon barrel (MB) consists of four layers of DT chambers. Each chamber is 4 cm wide tube that contains a stretched wire within a gas volume. Muons as they pass through the volume knock out electrons off the atoms of the gas. These electrons follow the electric field ending up at the positively-charged wire. By keeping track of the wire electrons hit as well as by calculating the muons' original distance away from the wire, the DTs give two coordinates for the position of each muon. Each DT, on average 2.0m×2.5m in size, consists of 12 aluminum layers, arranged in three groups of four, each with up to 60 tubes. The middle group measures the coordinate along the direction parallel to the beam while the two outside groups measure the perpendicular coordinate.

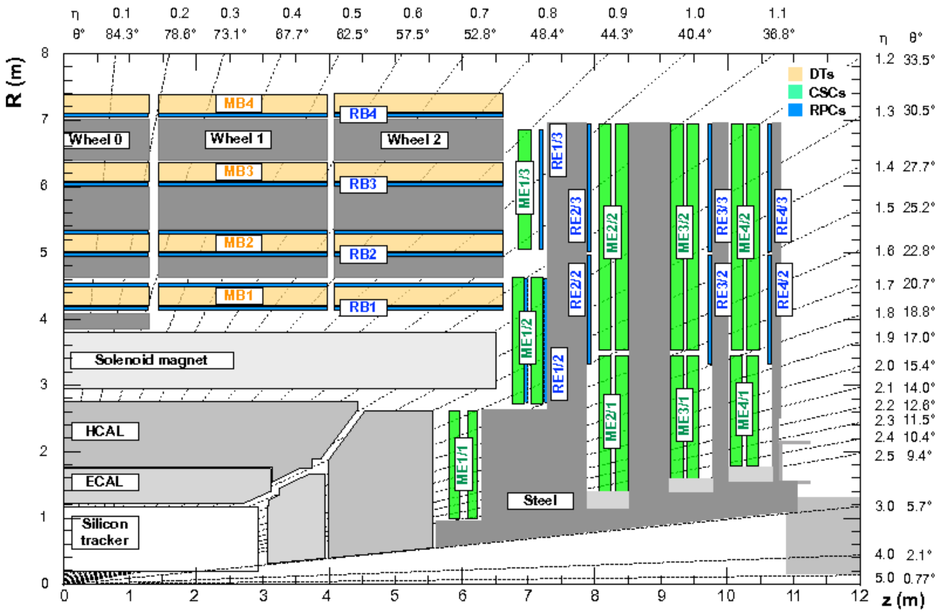


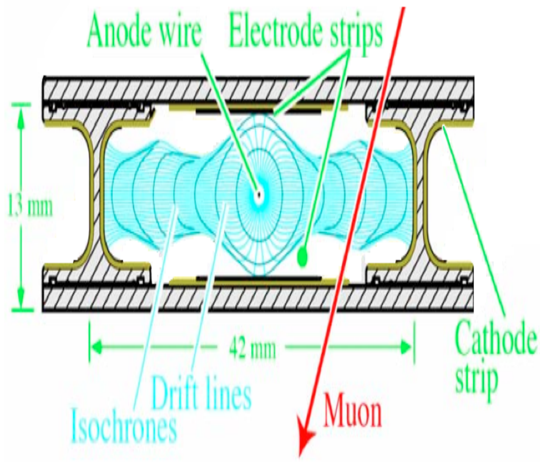
Figure 3.15: Muon detectors in CMS [105]

Cathode Strip Chambers

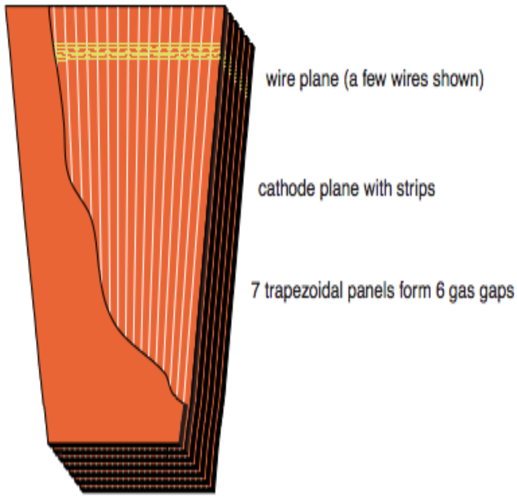
The muon endcap (ME) system consists of 468 Cathode Strip Chambers (CSCs). The chambers are trapezoidal and through overlapping provide contiguous ϕ-coverage. A muon in the range $1.2 < |\eta| < 2.4$ crosses 3 or 4 CSCs, in the range $0.9 < |\eta| < 1.2$ is detected by both the barrel drift tubes and endcaps CSCs. The CSCs are multi-wire proportional chambers comprised of 6 anode wire planes interleaved among 7 cathode strips running in radial direction within a gas volume. When muons pass through, they knock out electrons off the gas atoms, which travel to the anode wires creating an avalanche of secondary electrons. Positive ions move away from the wire and towards the cathode, also inducing a charge pulse in the strips, at right angles to the wire direction. As the strips and the wires are perpendicular, we get two position coordinates for each passing particle. In addition, to provide precise space and time information, the closely spaced wires make the CSCs fast detectors suitable for triggering.

Resistive Plate Chambers

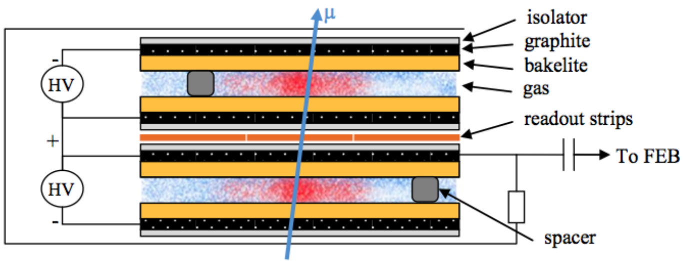
The resistive plate chambers (RPCs) cover the pseudorapidity region up to $|\eta| < 2.1$. They act as an additional muon trigger system and allow to resolve ambiguities in the CSCs. A total of 36 chambers are arranged in each of the two rings in each endcap. The RPCs are fast gaseous detectors that provide a trigger system parallel with those of the DTs and CSCs. RPCs are gaseous parallel-plate detectors able to combine adequate spatial resolution with a time resolution comparable to that of the scintillators. An RPC is capable of tagging the time of an ionizing event in a much shorter time than the 25 ns between 2 consecutive LHC bunch crossings (BX), reaching a time resolution of 1.5 ns. When a muon passes through the chamber, electrons are knocked out of the gas atoms. These electrons in turn hit other atoms causing an avalanche of electrons. The electrodes are transparent to the signal (the electrons), which are instead picked up by external metallic strips after a small but precise time delay. The pattern of hit strips gives a quick measure of the muon momentum, which is then used by the trigger to make an immediate decision about whether the data are worth keeping or not.



(a) Drift Tube



(b) Cathode Strip Chamber

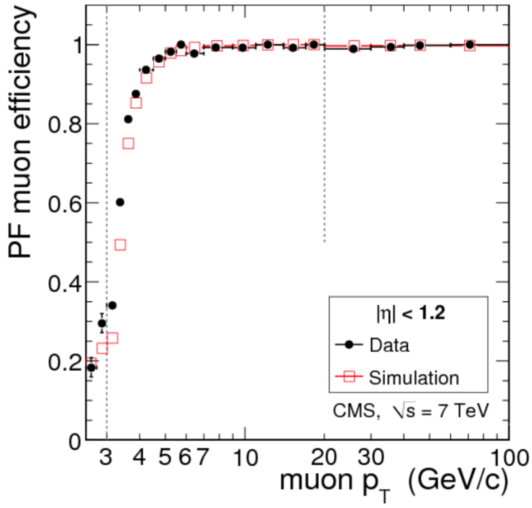


(c) Resistive Plate Chamber

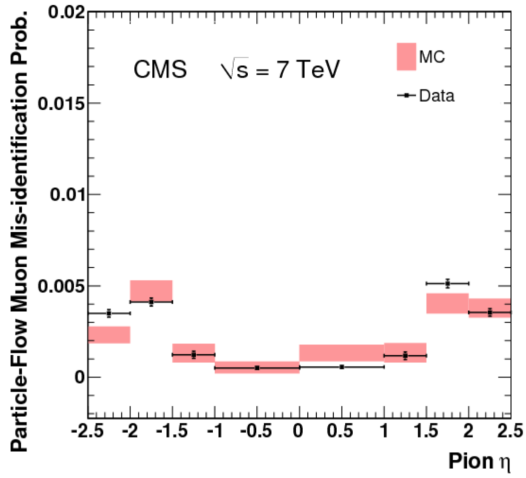
Figure 3.16: Various components of muon system in CMS [106]

Muon System Performance

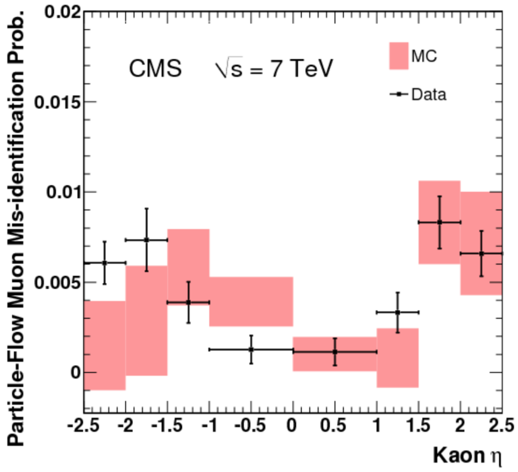
The muon momentum is measured in the inner tracker as well as in the return yoke flux. When only used the muon system, a momentum resolution of 10% (20%) is achieved for 40 GeV muons in the barrel (endcaps). However, the resolution of a global muon that combines information from the tracker with the muon system, is about 1% (2%) for muons with $p_T < 100$ GeV in the barrel (endcaps). For high p_T muons (above 100 GeV), the resolution degrades to about 8% for the barrel and 10% for the endcaps. The muon reconstruction and identification efficiency as function of p_T [107] is measured in $J/\psi \to \mu\mu$ and Z$\to \mu\mu$ events using the tag-and-probe [108] method with data collected at 7 TeV as shown in Figure 3.17.. The probability of π-K misidentified as muons is also estimated as a function of η as shown in Figure 3.17.



(a) μ efficiency vs p_T



(b) π mis-ID probability vs η

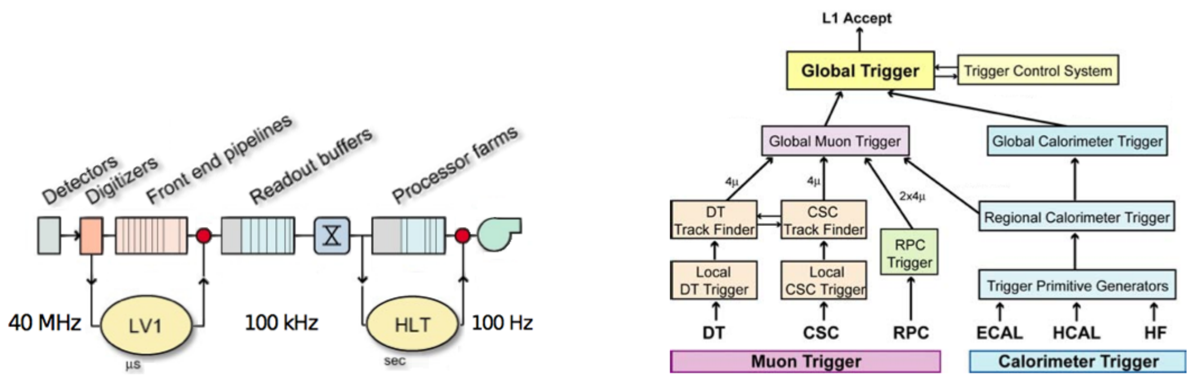


(c) Kaon mis-ID probability vs η

Figure 3.17: Muon performance [107]

3.2.7 Triggers

One of the biggest challenge LHC poses from the experimental point of view is to deal with a large volume of data to be stored in a limited storage capacity. The design BX of 25 ns corresponds to a frequency of 40 MHz. At nominal LHC luminosity, the expected rate of events is about 109 Hz, with an average size for a raw event of about 1 MB. Given this high rate, it is clearly not feasible to record all the collision information into a physical support, which would have a capability of around 100 Hz. On the other hand, most of the interactions are soft comprising low momentum particles, that are not relevant for the CMS physics purposes. Therefore, it becomes necessary to filter available information and record only the physics-wise interesting part of the collisions, and hence lowering the rate enough but still retaining all the physics potentials. The CMS trigger system is designed to exactly do this job by reducing the event rate down via a two-tier scheme: the Level-1 trigger (L1), mainly hardware-based, and the High Level Trigger (HLT), that exploits also software systems, as shown in Figure 3.18a.



(a) Two-tier trigger system

(b) Level-1 trigger

Figure 3.18: Schematic diagram for triggering system in CMS [106]

Level-1 Trigger

The Level-1 (L1) trigger [109] is designed to reduce the rate of events down to a maximum of 100 kHz from the BX frequency of 40 MHz. It takes a decision to whether keep or discard the data from a particular BX within 3.2 μs; if the L1 accepts the event, the data are moved to be processed by the next stage, i.e., HLT. As the available decision time is too short, instead of using the whole detector information, the L1 trigger takes advantage of the fast response coming from the calorimeter and muon systems only. The information coming from the calorimetric and muon triggers are then passed to a Global Trigger that takes the final decision (see Figure 3.18b). The calorimetric trigger is based on the information from the trigger towers, a 5×5 matrix of ECAL crystals, grouped in 4×4 squares. This trigger is able to distinguish classes of candidates: electrons, photons, jets, tau jets and missing energy by analyzing the energy deposits. The muon trigger runs in parallel taking information from the muon system. The final information is then merged and passed on to the Global Trigger, which analyzes the characteristics of the single object before passing information to the HLT.

High Level Trigger

The role of the High Level Trigger (HLT) [109] is to reduce the rate further down upto a maximum total bandwidth of ~ 1 kHz. The HLT is based on sophisticated and dedicated software tools called HLT paths that are designed to separate specific events topologies. It exploits a system, which connects information coming from all the sub-detectors of CMS and is controlled by an event manager system. These triggers have access to all the information used in L1 stored locally in the L1 trigger crates. Consequently, the HLT can make further combinations on a digital list of objects transmitted from L1, as much information is not available on the time scale of the L1 trigger decision. The tracking information as well as the full granularity calorimeter information is taken at the HLT stage, and all the CMS information are combined to make the final decision about whether the given event should be keep or rejected. The last stage of HLT processing is the reconstruction and event filtering with an aim of making dataset for different physics processes on easily accessed media.
This thesis deals with events which pass the lowest unprescaled single muon triggers. The typical rate of such triggers at an instantaneous luminosity 10^{33} cm^{-2} s^{-1} is $\mathcal{O}(50$ Hz$)$.

Chapter 4

Physics Objects and Top quark reconstruction

This chapter focuses on the description of different high level physics objects relevant for the analyses presented in this thesis are reconstructed from detector signals and the typical identification criteria applied to them.

4.1 Particle Flow (PF) algorithm

Most of the objects used in the analysis are reconstructed using the information from all the sub-detectors of CMS based on the so called Particle-flow (PF) algorithm [10]. First the particle candidates, i.e., charged hadrons, neutral hadrons, photons, electrons and muons are constructed from the tracks and calorimeter clusters as shown in Figure 4.1. These particle objects are then in turn used infer about higher level information, such as lepton isolation, tau lepton, jets after applying specific jet algorithms on the particle candidates and finally the missing energy.

4.2 Primary Vertex

The goal of primary vertex reconstruction is to determine the hardest vertex i.e. the precise position of pp interaction point. This task is accomplished with the so called Deterministic Annealing [111] clustering of tracks. The reconstructed vertex is required to pass the following criteria: ndf > 4; $|z| < 24$ cm; and $|d_{xy}| < 2$ cm. Due to high instantaneous luminosity at the LHC, there are additional pp interactions (pile-up) in the same bunch-crossing leading to a large number of reconstructed primary vertices in each event. From a set of reconstructed vertices, the one corresponding to the maximum $\sum \mathrm{p}_{\mathrm{T}}^2$ from all the tracks is chosen as the primary vertex or the collision point. As the luminosity varies over a given data taking period, the number of pp interactions per bunch-crossing is not well reproduced in simulations. As a consequence, there is a difference in the distribution of number of reconstructed primary vertices (N_{PV}) between data and MC

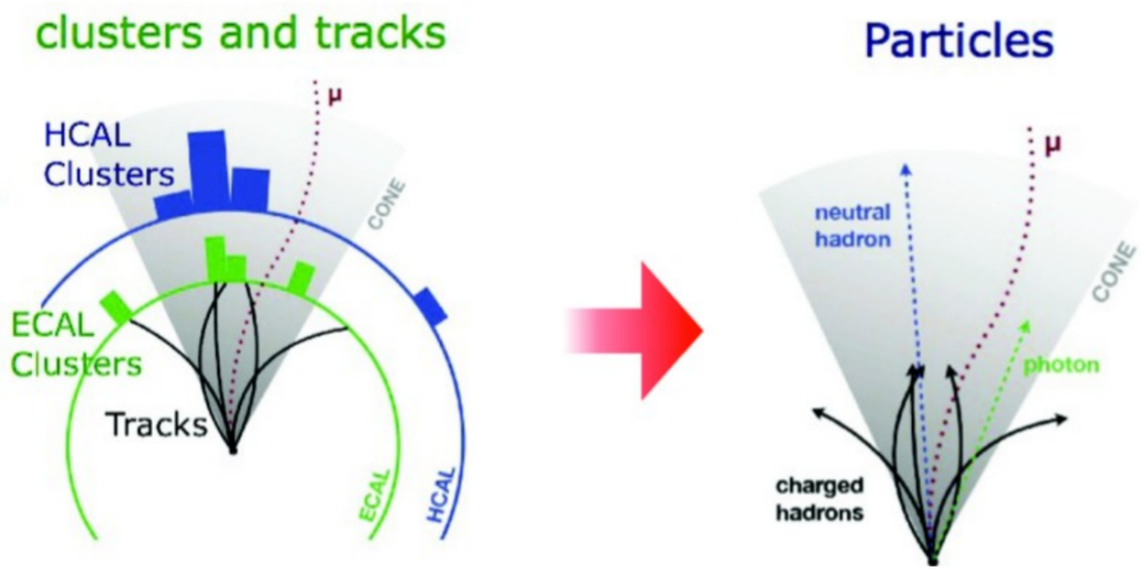
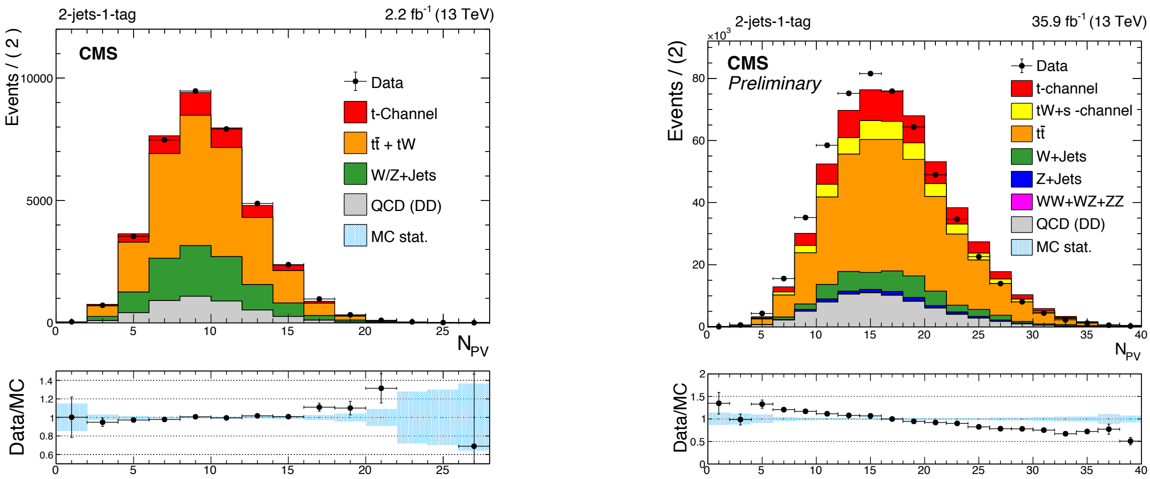


Figure 4.1: A schematic diagram of the PF algorithm [110]

simulated samples. To bring them to consistency, MC events are re-weighted to match pile -up distributions as observed in the data.



(a) 2015

(b) 2016

Figure 4.2: Data-MC comparison of number of reconstructed primary vertices (N_{PV}) distribution after pile-up re-weighting

4.3 Trigger

Events are required to pass the HLT criteria which demands presence of an isolated muon candidate above the p_T threshold for lowest unprescaled trigger. The value of the trigger p_T threshold changes for different analyses presented in later chapters depending on the data recording condition. Efficiency for the trigger(s) in simulation are corrected using data-to-MC correction factors obtained from a tag-and-probe method [108].

4.4 Muons

The muon reconstruction [112] chain starts with the "local reconstruction" using the hits in muon chambers and tracker separately and then matching the candidates to complete "global reconstruction". Based on this, muon objects are typically identified as one of the following.

- **standalone muons**: The segments reconstructed in the muon chambers are used to generate seeds consisting of position and direction vectors and an estimate of the muon transverse momentum. These initial estimates are used as seeds for the track fits in the muon system, which are performed using segments and hits from DTs, CSCs and RPCs and are based on the Kalman filter technique. The resulting collection is called "standalone muons"
- **tracker muons**: A complementary approach consists in considering all tracker tracks to be potential muon candidates and in checking this hypothesis by looking for compatible signatures in the calorimeters and in the muon system. Tracker tracks identified as muons by this method are referred to as "tracker muons". This algorithm is particularly useful for the identification of low-p_T muons (with p_T of the order of a few GeV), which may not leave enough hits in the muon stations for a "standalone muon" to be reconstructed.
- **global muons**: For each standalone muon track, a search for tracks matching it among those reconstructed in the inner tracking system is performed, and the best-matching tracker track is selected. For each such pair the track fit using all hits in both tracks is performed, again based on the Kalman filter technique which results in a collection of objects referred to as "global muons".

4.4.1 Tight Muons

These muons are reconstructed as global muons. Events with exactly one muon above a given high-p_T threshold and within $|\eta| < 2.4$ are selected for the analyses presented in this thesis. The selected high-p_T muon are required to meet additional quality requirements referred to as "tight muon ID" [14] which are as follows.

- Muon tracks should have $\chi^2/ndf < 10$
- at least one valid hit in the muon chamber
- at least two segments in two muon stations must match the global muon object, since the drift tube requires two segments to trigger.
- number of valid tracker hits > 5 out of which at least one in pixel detector so that at least a minimal number of measurement points are available in the tracker for good p_T measurement.

- Impact parameter in the transverse plane $|d_{xy}| < 2$mm w.r.t the centre of estimated beam spot position and distance along the beam-line w.r.t primary vertex should be $dz < 5$mm to reject cosmic muon tracks and tracks from pile-up vertices.

The first 3 criteria mostly helps to suppress hadronic punch-through effect and from muons which decay in flight. The third criterion is also required for triggering purpose. Further, the muon must be well isolated in terms of PF-based "DeltaBeta" [113] corrected relative isolation described below.

$$I_{rel} = \frac{I^{ch} + \max[(I^{\gamma} + I^{nh} - 0.5 \times I^{PU}), 0]}{p_{\mathrm{T}}^{\mu}}, \tag{4.1}$$

where I^{ch}, I^{γ} and I^{nh} are the sum of the transverse energies of charged hadrons, photons and neutral hadrons, respectively, in a cone size of $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4$ around the muon direction. I^{PU} is the $\mathrm{p_T}$ sum of charged hadrons associated to the primary vertices other than the leading primary vertex. It is used to estimate the contribution of neutral particles from pileup events.The factor 0.5 is used to take into account the neutral to charged particle ratio. For the tight muon, our selection requires $I_{rel} < 0.06$.The data-to-MC correction factors for muon are obtained using a tag-and-probe [108] method and are applied to simulated events to account for differences in efficiency between data and simulation.

4.4.2 Loose Muons

Events with additional charged muon are vetoed. The selection requirement for additional muon are loosened. Events having another muon(s) with $\mathrm{p_T} > 10$ GeV within $|\eta| < 2.4$ reconstructed as either the "global" or the "tracker muon" objects and relative isolation $\mathrm{I}_{rel} < 0.2$, are rejected.

4.5 Veto Electrons

Electrons are reconstructed using information from the pixel detector, strip tracker and ECAL clusters. First the ECAL clusters are used to determine the energy and position of the electrons which is then tightly matched with corresponding tracks in tracker. The energy loss due to bremsstrahlung for electrons propagating in matter introduces non-Gaussian contributions to the event-by-event fluctuations of the tracking measurements. A Gaussian-sum filter (Gsf) algorithm for electron track reconstruction in the CMS tracker has therefore been developed to model the bremsstrahlung energy loss distribution by a Gaussian mixture rather than a single Gaussian, as opposed to the standard Kalman filter which relies solely on Gaussian probability density functions [114,115]. The electrons used in the this analysis are thus reconstructed as "GsfElectron" objects.
Events with a charged electron along with the a tight isolated muon described in section 4.4.1 are vetoed. Any event which contains an additional electron with $\mathrm{E_T} > 15$ GeV within $|\eta| < 2.5$ and passes the cut-based "veto" identification criteria (efficiency 98% at plateau reached around 45 GeV),is rejected.

4.6 Jets

The jets used in the analysis are reconstructed using the anti-k_T algorithm [12] with a cone size of 0.4, using PF candidates as inputs and after rejecting charged hadrons associated to pile-up vertices. Once the jets are reconstructed, several levels of corrections [13] are applied depending on the p_T and $|\eta|$ of the jets. Corrections are applied sequentially in a factorized approach on both data and simulation to get corrected energy and other parameters of jets. For data we further apply residual corrections in order to reduce remaining differences with simulation.
This analysis requires at least 2 jets within $|\eta| < 4.7$ and corrected $p_T > 40$ GeV. To reduce the effect of misidentified jets from pile-up vertices or detector noise, $|\eta|$-dependent loose "Jet ID" [15] is applied. The $|\eta|$-dependency of the "Jet ID" criteria can be attributed to the limited tracker acceptance ($|\eta| < 2.4$) and presence of HF which has worse granularity compared to HB and HE in the high-$|\eta|$ region of the detector.
For Jets within $|\eta| < 2.7$,

- Neutral Hadron Energy Fraction < 0.99
- Neutral EM Energy Fraction < 0.99
- Number of constituents > 1

Additionally for PF jets with $|\eta| < 2.4$,

- Charged Hadron Energy Fraction > 0
- Charged multiplicity > 0
- Charged EM Energy Fraction < 0.99

For Jets within $2.7 < |\eta| < 3.0$,

- Neutral EM Energy Fraction > 0.01
- Neutral Hadron Energy Fraction < 0.98
- Neutral multiplicity > 2

For Jets with $|\eta| > 3.0$,

- Neutral EM Energy Fraction < 0.90
- Neutral multiplicity > 10

Jets are also required to have a distance more than 0.4 in the $\eta - \phi$ plane w.r.t the selected tight muon described in Section 4.4.1.

4.7 b-Tagging

As the life time of b-flavoured meson is large, it travels some distance before it decays. CMS has explored various technique [116] to identify jets that originate from b-quark using the secondary vertex and lifetime information. Figure 4.3 shows the tagging efficiency vs mis-identification probability corresponding different tagging algorithm. For the analyses presented in this thesis, mainly two different algorithms has been used for b-tagging, which are described as follows.

- **Combined Secondary Vertex version 2 (CSVv2):** This algorithm is based on secondary vertex and track-based lifetime information, it is an updated version of the CSV algorithm used in Run-I combining the variables such as vertex mass, number of tracks at vertex and inside a jet, 2D flight distance significance of the b-flavoured meson, ratio of energy carried by the tracks at the secondary vertex vs all tracks inside the jet etc. into a neural network instead of a likelihood ratio and the secondary vertex information is obtained with the inclusive vertex finder algorithm.
- **combined MultiVariate Algorithm version 2 (cMVAv2):** This algorithm uses a Boosted Decision Tree (BDT) taking as input the different algorithm outputs of CSVv2, a variant of CSVv2 using alternate vertex reconstruction, Jet Probability (JP), Jet B Probability (JBP), Soft Electron (SE) and Soft Muon (SM) taggers. As seen in Figure 4.3, this algorithm performs better compared to CSVv2 in discriminating jets originating from b-quark from jets due to c-quark, in the final state.

A tight threshold is applied on the b-tagging discriminator value which corresponds to mis-tag rate of 0.1%.

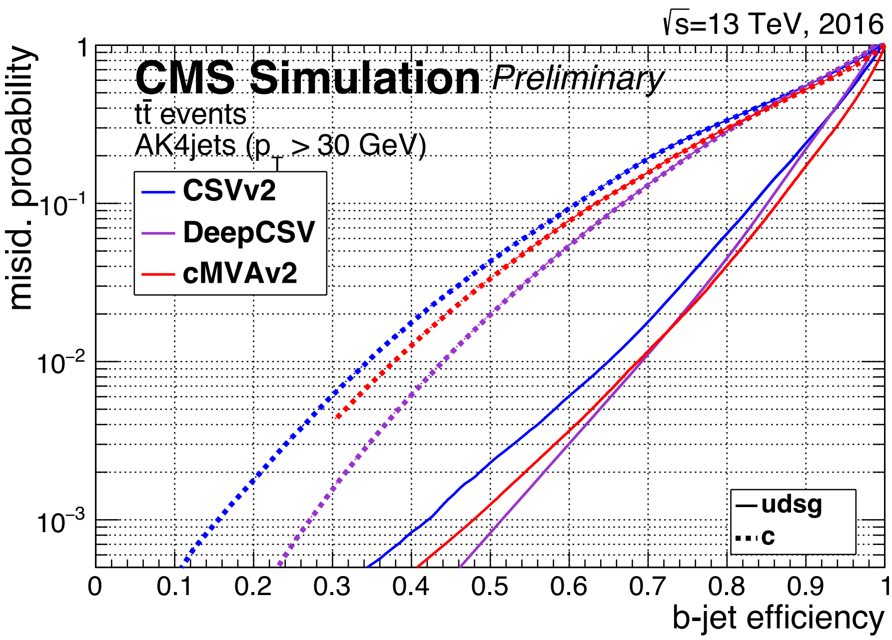


Figure 4.3: Performance of different b-tagging algorithms [117]

4.8 Missing Transverse Momentum

Missing transverse momentum (MET) is a measure of the imbalance in the p_T of all "visible" final state particles that interact with the detector material as described in Eq.4.2. In the analyses present in this thesis, the source of MET can be attributed to the p_T carried away by the neutrino, an "invisible" final state particle, originating from W boson, which in turn, comes from top quark decay. Corrections are applied on MET to mitigate the contamination due to pile-up and to account for jet energy calibration effects.

$$\not{E}_{\mathrm{T}} = |\overrightarrow{\not{E}_{\mathrm{T}}}| = |-\sum_{i\in visible} \overrightarrow{p_{\mathrm{T,i}}}| \tag{4.2}$$

4.9 Event Categorization

The t-channel single top production has one light quark recoiling against the virtual W boson and one b quark from top-quark decay and a second b quark from initial gluon splitting. The p_T -spectrum of the second b quark is softer compared to the one coming from top quark decay. It also has large $|\eta|$ compared to more central b-quark originating from top. As a result, the jet due to the second b-quark in the final state is often rejected by jet p_T -threshold. If selected, it is mostly found to be out of tracker acceptance ($|\eta| < 2.4$) to be tagged as a b-jet. Hence, the region having 2 jets, with one of them being b-tagged, has the largest fraction of signal events. Depending on the number of jets and b-tagged jets we define few categories in order to validate the normalization and modeling of main background processes. The generic nomenclature "nJmT" is attributed to the events having n jets and m b-tagged jets. In addition to the signal-enriched 2J1T region, 2J0T (enriched in W+light flavour jets and QCD), 3J1T (enriched in $t\bar{t}$ and W+ heavy flavour jets) and 3J2T (enriched in $t\bar{t}$) regions have been studied as control regions.

4.10 Transverse W-boson Mass

The transverse W-boson mass (m_T^W) is defined as

$$m_{\mathrm{T}}^{\mathrm{W}} = \sqrt{(p_{T,\mu} + \not{E}_{\mathrm{T}})^2 - (p_{x,\mu} + \not{p}_x)^2 - (p_{y,\mu} + \not{p}_y)^2} \tag{4.3}$$

where, $p_{x,\mu}$, $p_{y,\mu}$ are the x and y-components of muon 3-momentum and $\not{p}_x$, $\not{p}_y$) are the x and y-components of $\overrightarrow{\not{E}_{\mathrm{T}}}$.

m_T^W is very sensitive to the processes with prompt muons from leptonically decaying W boson. The m_T^W distributions for QCD multijet backgrounds and other processes is shown in the left plot of Figure 4.4. Beyond 50 GeV the QCD probability goes down monotonically whereas the probability for other processes starts to increase. So by requiring m_T^W > 50 GeV most of the QCD multijet background can be rejected.

4.11 Top Quark Reconstruction

The 4-momentum of the top quark is reconstructed from the available kinematic information in the event. The top-quark decays to a b-tagged jet, a charged lepton and a neutrino, whose transverse momentum can be inferred from $\not\!\!E_T$. The longitudinal momentum of the neutrino, $p_{z,\nu}$, is determined from other kinematic constraints such as the W boson mass, $m_W = 80.4$ GeV [16]. Assuming energy-momentum conservation at the W→ $\mu\nu$ vertex, one can obtain Eq.(4.4) for $p_{z,\nu}$. Two cases can arise for a solution, namely,

- A negative discriminant leads to complex solutions for $p_{z,\nu}$. In this case, the imaginary part is eliminated by setting $m_T^W = m_W$, while still respecting the m_W constraint. This leads to the fact that the discriminant i.e., the square root term in Eq.(4.4) is null. This condition gives a quadratic relation between $p_{x,\nu}$ and $p_{y,\nu}$ with two possible solutions, and one remaining degree of freedom. The solution corresponding to the minimal distance between $p_{T,\nu}$ and $\not\!\!E_T$ is chosen.
- For events with a positive discriminant, the solution with the smallest absolute value is chosen [17, 18].

$$p_{z,\nu} = \frac{\Lambda p_{z,\mu}}{p_{T,\mu}^2} \pm \frac{1}{p_{T,\mu}^2}\sqrt{\Lambda^2 p_{z,\mu}^2 - p_{T,\mu}^2(E_\mu^2 \not\!\!E_T^2 - \Lambda^2)}, \text{where } \Lambda = \frac{m_W^2}{2} + \overrightarrow{p_{T,\mu}} \cdot \overrightarrow{\not\!\!E_T} \quad (4.4)$$

The $|p_{z,\nu}|$ distributions for t-channel single top signal and other background processes are shown in Figure 4.4. Apart from QCD, other processes having a real W boson in the final state, $|p_{z,\nu}|$ peaks at zero, as expected, since the neutrino from W-decay carries away enough energy giving rise to high $\not\!\!E_T$. Using the $p_{z,\nu}$ information, the 4-momentum of the W boson candidate can be completely determined.

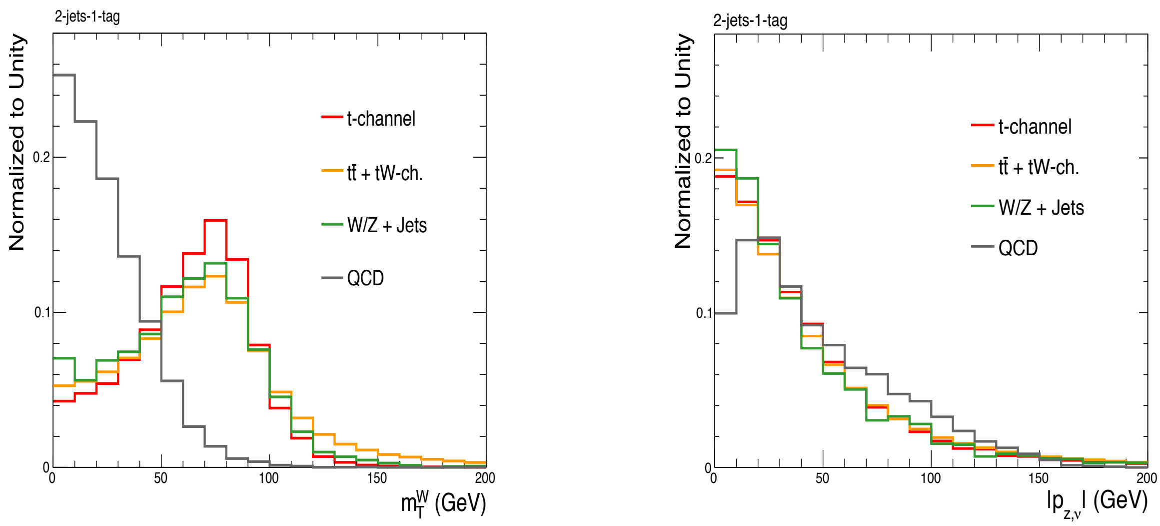


Figure 4.4: m_T^W (left) and $|p_{z,\nu}|$ distribution of signal and background processes.

In the 2J1T and 3J1T regions, the b-tagged jet is assumed to come from the top quark decay. In 3J1T region, the most forward jet is considered to be the jet stemming from the light quark recoiling against top. In 3J2T region, the jet with the largest b-tagging discriminant value is attributed to the b quark from the top decay while the non-b-tagged jet is identified to originate from the light quark. In 2J0T region, three cases are considered, namely,

- If both jets pass b-tagging algorithm but do not satisfy the b-tagging discriminator tight criterion, then the jet with higher b-tagging discriminant value is attributed to b-quark originating from top.
- If only one of the two jets passes b-tagging algorithm but does not satisfy the b-tagging discriminator tight criterion, then that jet is identified as the b-jet.
- In the case where both jets fail to pass the b-tagging algorithm, the more central jet is assigned as the b-jet.

The 4-momentum of the W boson is then added to the 4-momentum of the b-jet candidate to obtain the 4-momentum of the decaying top quark.
Figure 4.5 shows the signature of a typical t-channel single top event during the pp collision at 13 TeV.

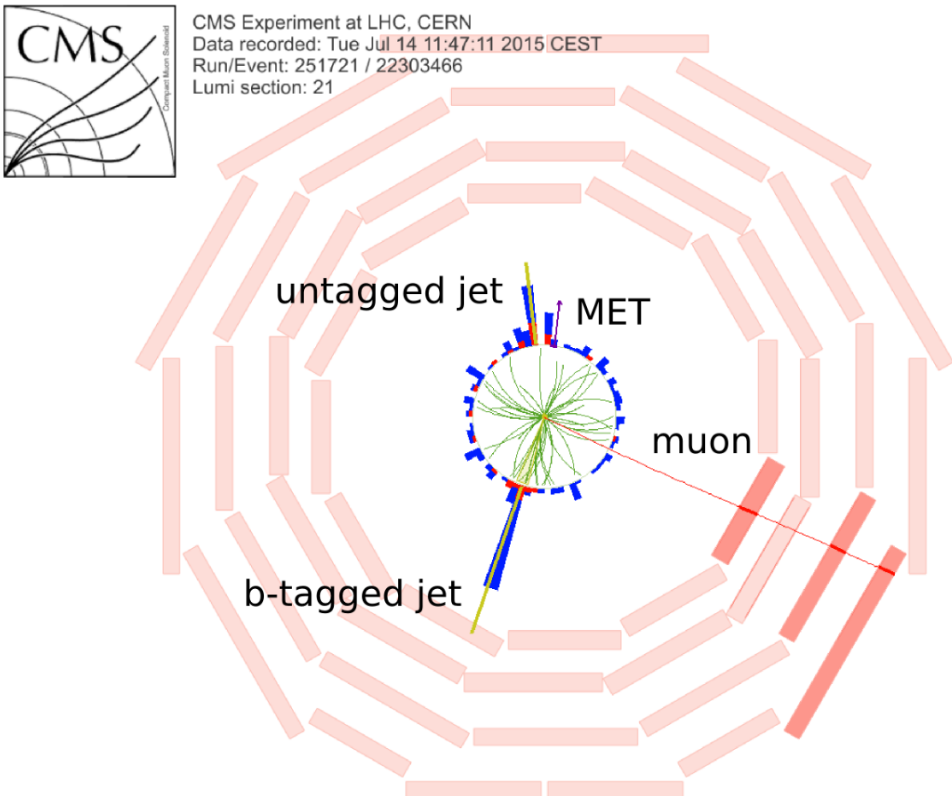


Figure 4.5: Event display for a t-channel single top event during pp collision at 13 TeV [118]

Chapter 5

Cross section measurement with early 42 pb^{-1} data

Since the discovery at Tevatron [17,18], single top-quark processes have gradually become a unique testing ground for the study of QCD and electroweak processes, specifically at the tWb vertex, as well as the measurement of the CKM matrix element V_{tb}. It has already been studied extensively at the LHC during Run-I at CM energies of 7 and 8 TeV respectively. Run-II of LHC marks the period for precision measurements of the single top processes.
A preliminary measurement of the t-channel cross-section was performed with the very first pp collision data at $\sqrt{s} = 13$ TeV recorded by the CMS detector with 50ns bunch spacing, as the LHC restarted after shutdown period. Measurement is performed in the final state containing one high-p_T isolated muon, two hadronic jets with one of them identified to have originated from a b-quark in the final state and large $\not{E}_T$ due to the escaping neutrino. Backgrounds are extracted using side-band regions. The signal is extracted from a fit to the pseudo-rapidity distribution of the recoiling non-tagged jet.

5.1 Data and MC samples

The MC and data samples used for this analysis are listed below in Table 5.1 and Table 5.2.

Table 5.1: List of MC samples

Process	$\sigma(\cdot BR)$[pb]	Generator	Nevents
t-channel single top, top + anti-top, leptonic decays	217($\cdot$ 0.324) (NLO) [6, 119]	aMC@NLO-pythia8	4M
tW-channel single top, top, inclusive decays	35.6 (NNLL) [120]	powheg-pythia8	1M
tW-channel single top, anti-top, inclusive decays	35.6 (NNLL) [120]	powheg-pythia8	1M
$t\bar{t}$, inclusive decays	831.76 (NNLO) [58, 121]	powheg-pythia8	20M
$W(\rightarrow \ell\nu)$ + Jets	61526.0 (NLO)	aMC@NLO-pythia8	24M
$Z/\gamma(\rightarrow \ell\ell)$ + Jets, $m_{\ell\ell} > 50$ GeV	6025.2 (NLO)	aMC@NLO-pythia8	20M
μ -enriched QCD, $\hat{p}_T > 20$ GeV, $p_T^{\mu} > 15$ GeV	240680 (LO)	pythia8	13M

Table 5.2: List of Data samples

Dataset	Run Range	Integrated Luminosity (pb^{-1})
/SingleMuon/Run2015B-17Jul2015-v1/MINIAOD	251162-251562	42
/SingleMuon/Run2015B-PromptReco-v1/MINIAOD	251585-251883	

5.2 Event Selection

The following event selection criteria was applied to reduce backgrounds.

- Events are required to pass high level trigger criteria which require at least one prompt global muon or tracker muon with p_T above 20 GeV within $|\eta| < 2.1$ in the final state ("HLT_IsoMu20_eta2p1" OR "HLT_IsoTkMu20_eta2p1"). The two HLT bits combined through the logical "OR" operator maximizes the efficiency.

- Events with exactly one muon having $p_T > 22$ GeV within $|\eta| < 2.1$ passing "tight ID" criteria and $I_{rel} < 0.6$ is selected.

- Events with another muon above p_T threshold of 10 GeV within $|\eta| < 2.5$ are rejected in order to reduce Z+Jets and $t\bar{t}$ di-muon backgrounds.

- Events with another electron in addition to selected high-p_T muon described above with p_T above 20 GeV within $|\eta| < 2.4$ are rejected in order to reduce $t\bar{t}$ di-lepton background.

- Event having exactly 2 anti-k_T jets with cone size 0.4 in η-ϕ space made out of PF candidates are required to have p_T above 40 GeV within $|\eta| < 4.7$ passing the loose identification criteria based on their constituents and having $\Delta R > 0.4$ with respect to the high-p_T muon candidate.

- Out of the two selected jets one jet is required to pass the tight working point of the neural network based CSVv2 algorithm (as explained in Section 4.7) within $|\eta| < 2.4$

to be identified as a b-tagged jet originating from the b-quark in final state due to the decay of the top quark. Thus we obtain the signal enriched region defined as 2J1T

- To reduce QCD background further, as explained in Section 4.10, events are required to have $m_T^W > 50$ GeV.
- In the signal region (2J1T) an additional cut on the invariant mass of the reconstructed top quark is applied. Events inside the mass window corresponding to $m_{\mu\nu b} \in [130, 225]$ GeV are kept while the events outside this region are used to model the background from W+Jets.

5.3 Estimation of backgrounds

5.3.1 Estimation of QCD background

The production of multijets through QCD processes has a huge cross section. On the other hand, only a small fraction of these events mimic the muon+jets final state of the applied event selection. The selection efficiency for QCD multijet events is therefore tiny. The combination of huge cross section and tiny selection efficiency would require generation of extremely large MC samples for this process in order to produce sufficient events that survive the event selection to guarantee a proper QCD modeling with decent statistics in the signal region. An alternative way of modeling is to define control regions in data that are enriched in QCD events (purity ≈ 95%) by inverting the muon isolation requirement ($I_{rel} > 0.12$) and estimate QCD contribution to the signal region by means of a likelihood fit to the full m_T^W distribution in data using QCD templates derived from the control region. The QCD yield in the $m_T^W > 50$ GeV is obtained by calculating the integral of the m_T^W -distribution from QCD, normalized to the fit result, for $m_T^W > 50$ GeV. The method is validated in the 2J0T region, as the number of available events is larger compared to signal region and the final estimate is taken from the 2J1T region.

The m_T^W -distribution in data is modeled according to Eq.(5.1). $Q(m_T^W)$ and N_{QCD} denote the QCD m_T^W -template and QCD yield parameter respectively. $W(m_T^W)$ denotes the combined MC template of all other processes with a prompt muon in the final state scaled according to their respective cross-sections and N_{nonQCD} is the yield parameter corresponding to $W(m_T^W)$. An extended maximum likelihood fit with 2 parameters, namely N_{QCD} and N_{nonQCD} , is performed to the m_T^W -distribution in data.

$$F(m_T^W) = N_{\text{QCD}}\,Q(m_T^W) + N_{\text{nonQCD}}\,W(m_T^W) \tag{5.1}$$

Estimation in 2J0T region

The 2J0T was used to validate the QCD estimation method with MC and data-driven templates for $Q(m_T^W)$ from the anti-isolated side-band. A comparison of $Q(m_T^W)$ templates from isolated region with the ones obtained from anti-isolated side-band is shown in Figure 5.1 and the templates are observed to agree with each other within the uncertainties. In addition, an unbinned parametric template was also tested. All fits are shown in Figure 5.2 and Table 5.3 lists the fit results using different templates. The difference in the respective final QCD yields can be attributed to the difference in the corresponding QCD templates. The data-driven QCD m_T^W -template is susceptible to the statistical fluctuations of the m_T^W distributions of the non-QCD processes in $I_{rel} > 0.12$ region, since combined non-QCD distribution is subtracted from the data distribution in order to obtain the data-driven QCD template. The fluctuation mostly arise due to the differences in trigger , muon isolation and identification efficiencies for different processes in the $I_{rel} > 0.12$ region. In order to account for the differences, 50% conservative uncertainty on the QCD normalization, obtained using the data-driven template, is assigned as systematic.

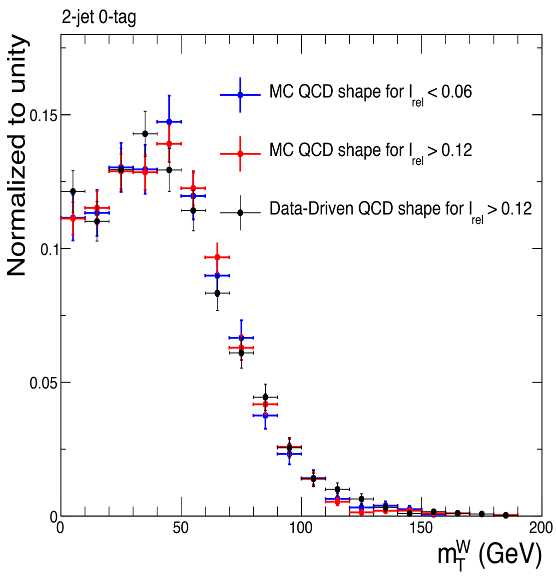
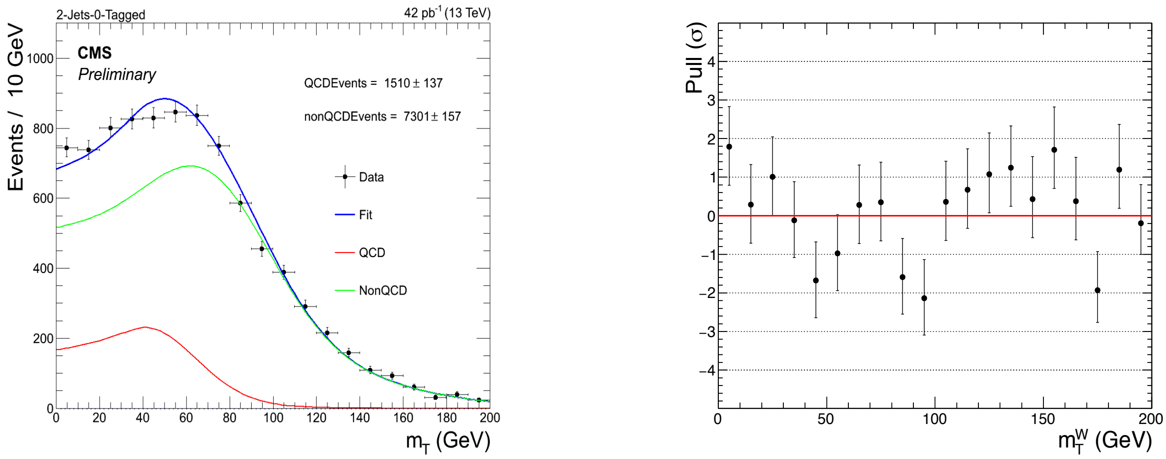


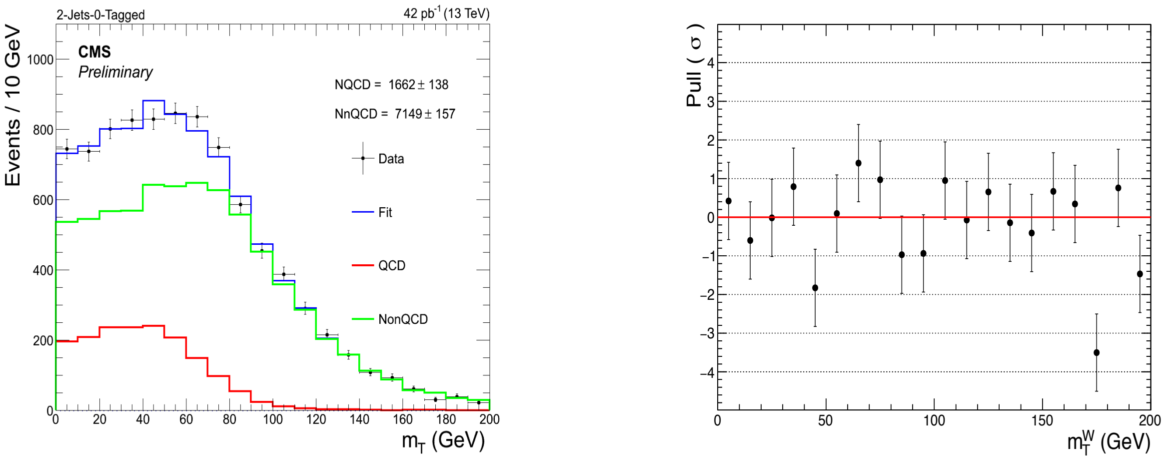
Figure 5.1: Comparison of QCD m_T^W -templates from isolated region and anti-isolated side-band in 2J0T

Table 5.3: QCD estimation in the 2J0T region

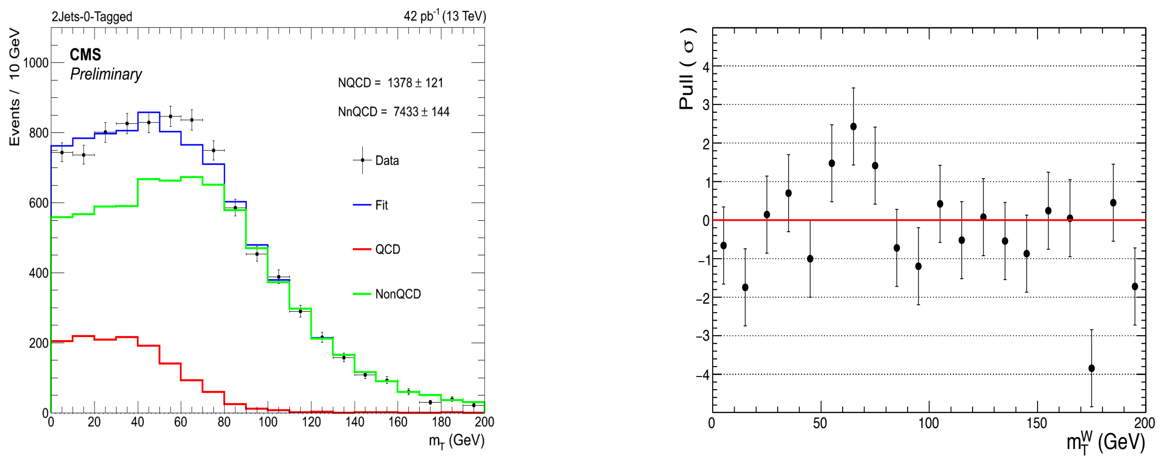
Variable	Fit Type	QCD shape	Process	Fitted Yield (full range)	Extrapolated Yield ($m_T^W > 50$GeV)
m_T^W	Unbinned	Parametric	QCD	1510±137	501±46
			nonQCD	7301±157	4407±95
	Binned	From AntiIso MC	QCD	1662±138	547±46
			nonQCD	7149±157	4291±94
		Data Driven	QCD	1378±121	342±30
			nonQCD	7433±144	4462±86



(a) Parametric template



(b) QCD template from MC ($I_{rel} > 0.12$)



(c) Data-driven QCD template from $I_{rel} > 0.12$

Figure 5.2: QCD fit in 2J0T region using different templates

Estimation in 2J1T region

The same binned likelihood fit as described in the previous section is applied to the m_T^W -distribution in the 2J1T region. The template for the non-QCD processes is again derived from MC simulation by summing up the individual contributions from the different processes weighted with the corresponding cross sections, while the QCD-template is again derived from the anti-isolated side-band in data ($I_{rel} > 0.12$). Again, the entire m_T^W -distribution is fitted and the QCD contribution in the signal region ($m_T^W > 50$ GeV) is extrapolated from the post-fit m_T^W -distribution, normalized to the fit results. The estimation is performed separately inside and outside the top mass window in 2J1T. Table 5.4 summarizes the fitted QCD and non-QCD yields for the two regions based on top mass window and the fits are shown in Figure 5.3.

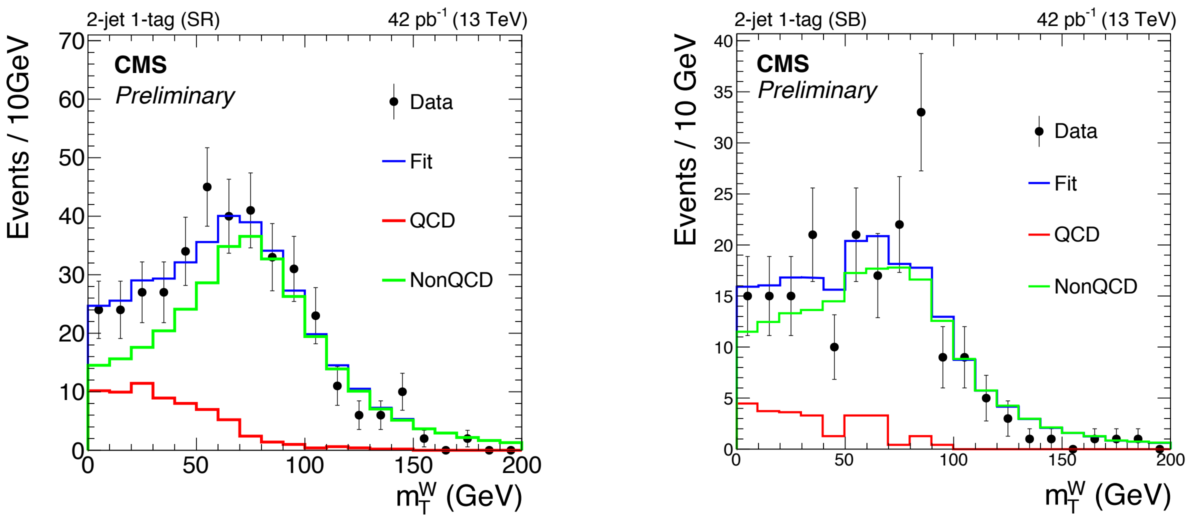


Figure 5.3: QCD fit in 2J1T region: inside top mass window (left) and outside top mass window (right)

Table 5.4: QCD estimation in the 2J1T region

$m_{\mu\nu b}$ range	Process	Fitted Yield (full m_T^W range)	Extrapolated Yield ($m_T^W > 50$ GeV)
$m_{\mu\nu b} < 130$ & $m_{\mu\nu b} > 225$	QCD	20±13	2±1
	nonQCD	181±16	116±10
$130 < m_{\mu\nu b} < 225$	QCD	65±17	10±5
	nonQCD	330±22	232±15

5.3.2 Estimation of W+Jets background

W+jets events are one of the main backgrounds surviving the selection cuts of the 2J1T signal region. Among them, those with jets stemming from heavy flavour partons, are not modeled well. Therefore, we model the shapes of relevant kinematic variables from W+jets events based on data distributions and simulation of well known and well modeled other processes.The m variable is used to define a signal region populated mainly with signal events, to be $130 < m_{\mu\nu b} < 225$ GeV, and a control region (side-band region) outside this window, where the W+jets fraction is significantly larger (about 45% of the events in the side-band region are W+jets events). The lower and upper edges of the window is chosen such that a signal efficiency of about 82% is obtained. The idea is that the shape of the distributions of kinematic variables from W+jets events in the side-band region is similar to that of W+jets events within the mass window. The variable which is later used to determine the contributions of the signal and various background processes is the pseudo-rapidity of the light quark jet, $|\eta_{j'}|$. The similarity of the distribution of this variable from W+jets events contributing to the signal and side-band regions, as shown in Figure 5.4 is checked using a Kolmogorov test, which yields a p-value of more than 90%. The side-band in the 2J1T region consists of about 45% W+jets events. Thus, the $|\eta_{j'}|$ from the side-band data has to be corrected for the non-W+jets processes in order to be used as W+jets template in the signal region. The contributions from all other processes are subtracted from the data in the side-band region using the shape and normalization from MC simulations, except for the QCD-template, which is obtained directly from data as described in Section 5.3.1. The available dataset being very tiny, the number of data events in the side-band region is rather small leading to large statistical fluctuations in the template from data. So for the validation of W+jets modeling, a scale factor is extracted from the ratio of W+jets templates from the side-band region to that in the signal region, which is found to be compatible with one.

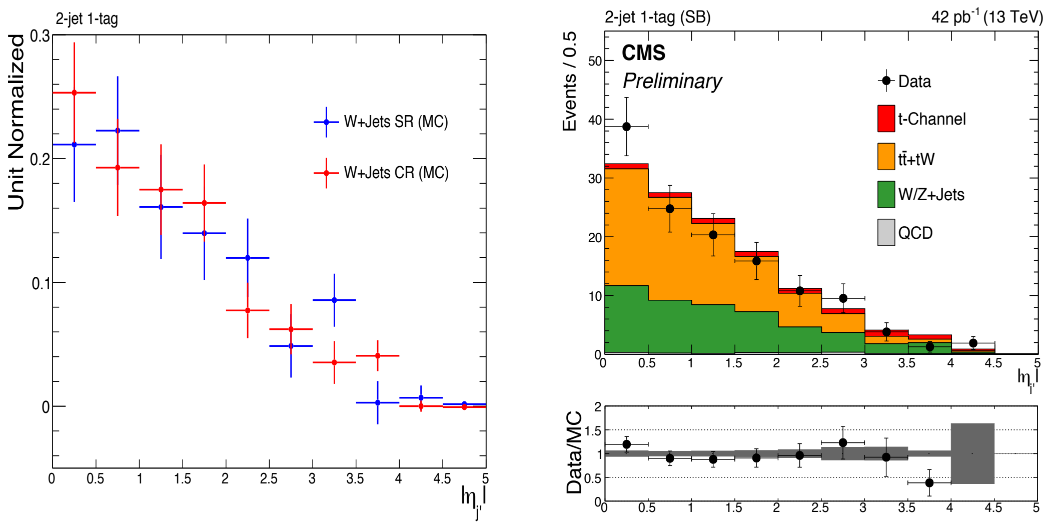


Figure 5.4: Comparison of W+Jets templates from SR defined by 130 GeV $< m_{\mu\nu b} <$ 225 GeV and SB defined by $m_{\mu\nu b} <$ 130 GeV and $m_{\mu\nu b} >$ 225 GeV (left) and Data-MC comparison of $|\eta_{j'}|$ in SB (right)

5.3.3 Estimation of $t\bar{t}$ background

The modeling of the $t\bar{t}$ background in simulation is validated using the 3J2T region. the 3J2T region is mostly dominated by $t\bar{t}$ background as shown in Figure 5.5. A simultaneous fit is performed including the 3J2T region to control $t\bar{t}$ normalization.

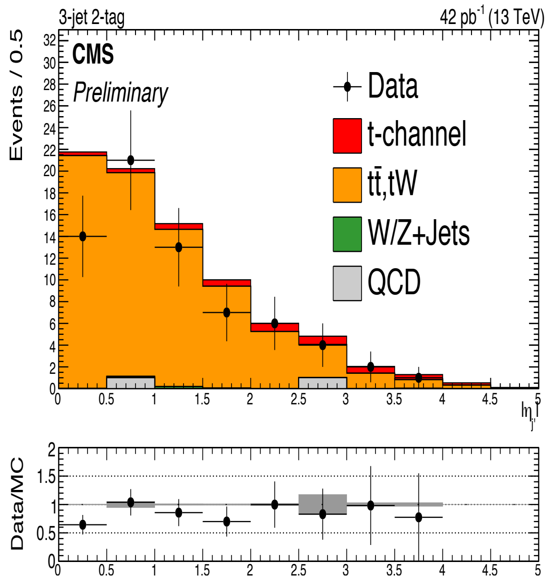


Figure 5.5: Data-MC comparison of $|\eta_{j'}|$ in 3J2T region

5.4 Event yield

The final event yield in data and MC after applying full selection criteria and background estimation methods are listed in Table 5.5. Except for the number of QCD multijet events all yields are taken from simulation and their uncertainties correspond to the size of the simulated samples. The yield of QCD multijet events is determined from the data (see Section. 5.3.1). Data-MC comparisons control plots for muon p_T , $\not{E}_T$, b-jet $|\eta|$, $|\eta_{j'}|$ in the 2J1T SR are shown in Figures 5.6 - 5.7.

5.5 Cross section extraction

A binned maximum likelihood fit is performed on the $|\eta_{j'}|$ distribution to determine the signal cross section [19]. Different background processes with the same behavior are grouped together in order to reduce the impact of yield uncertainties of the individual process. Background contributions from top quarks, namely $t\bar{t}$ and tW-channel single top quark production (s-channel single top quark production is neglected), are grouped together in the Top background category, whereas W+jets and Z+jets defines the EWK background category. A third background category consists of QCD events, its contribution is derived from the procedure described in Section 5.3.1 and kept fixed. The fit is

Table 5.5: Final event yields of the signal and main background processes in the 2-jet-1-tag sample inside (SR) and outside (SB) the $m_{\mu\nu b}$ window

Process	SR	SB
$t\bar{t}$ & tW	157 ± 1	71.7 ± 0.4
W/Z+jets	40 ± 4	47 ± 4
QCD	10 ± 5	2 ± 1
t-channel	33 ± 1	7.2 ± 0.3
Total expected	240 ± 6	128 ± 4
Data	252	127

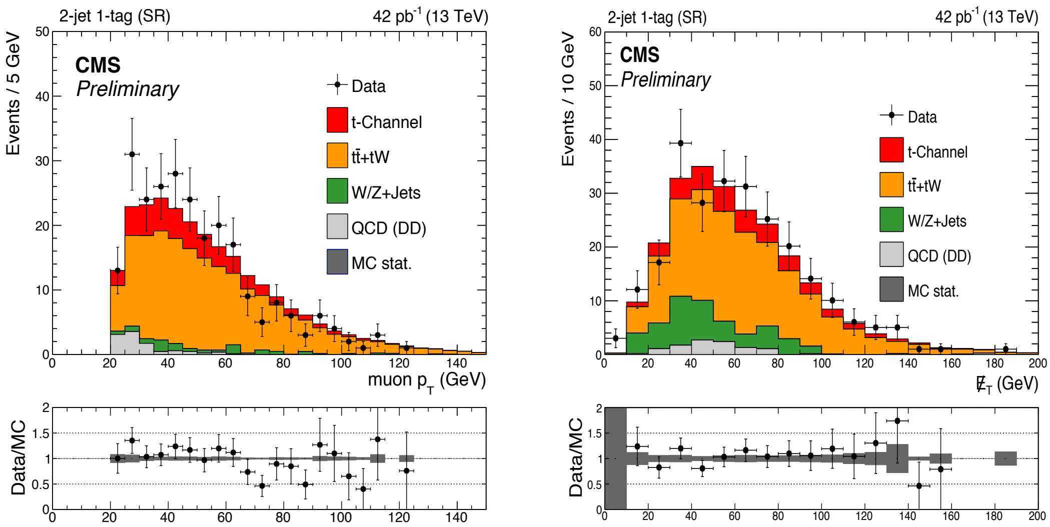


Figure 5.6: Data-MC comparison of muon p_T (left) and $\not{E}_T$ (right) in SR

done simultaneously in the signal enriched 2J1T region within the $m_{\mu\nu b}$ window between 130 GeV and 225 GeV and the 3J2T region. The latter one is dominated by $t\bar{t}$ events, which helps to reduce the uncertainty of the top background component.
The normalization scale factor (S_i) defined in Eq.(5.2) of each process category is allowed to float during fit,

$$S_i = \frac{N_i^{obs}}{N_i^{pred.}}, \tag{5.2}$$

where N_i^{obs} and $N_i^{pred.}$ are the number of events observed and predicted for the i-th process category, respectively.
Normalization scale factor for t-channel is allowed to float freely whereas top ($t\bar{t}$ + tW) and electroweak (W/Z + Jets) background normalizations are constrained by Gaussian priors with 20% and 30% uncertainties, respectively to account for theoretical and

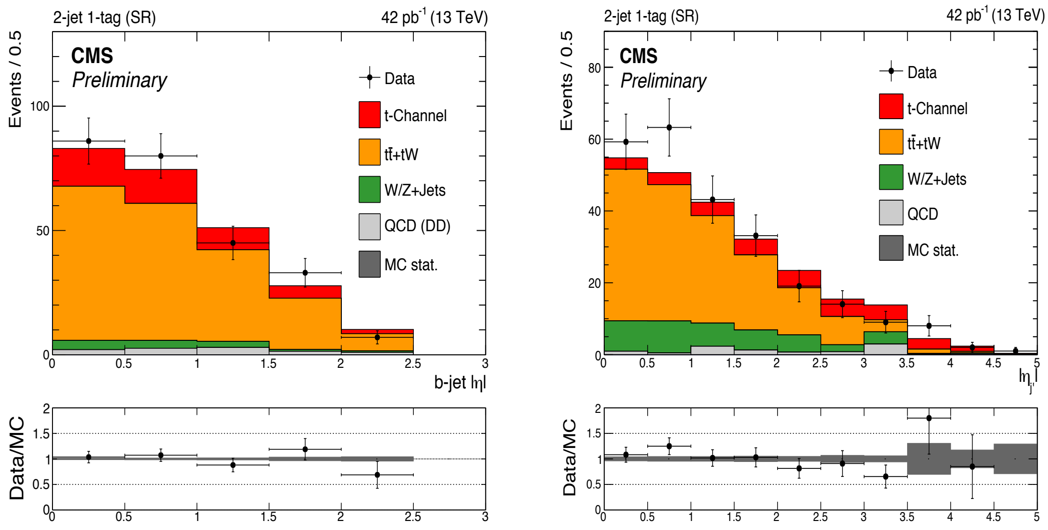


Figure 5.7: Data-MC comparison of b-jet $|\eta|$ (left) $|\eta_{j'}|$ (right) in SR

experimental uncertainties. The fit results are listed in Table 5.6.

Table 5.6: Fit Scale factors

Process Category	Scale factor
t-channel	1.26 ± 0.45
Top	1.03 ± 0.10
EWK	1.02 ± 0.27

Figure 5.8 shows the post-fit distribution of the $|\eta_{j'}|$ in signal and control regions. The cross section of t-channel single top process is determined using Eq.(5.3),

$$\sigma_{\mathrm{t-ch}} = \frac{\mathrm{N_s}}{\epsilon \times \mathcal{B}(\mathrm{t} \to \mathrm{b}\ell\nu) \times \mathcal{L}} \tag{5.3}$$

where N_s is the t-channel event yield after fit, ϵ is the selection efficiency, $\mathcal{L} = 42\ \mathrm{pb}^{-1}$ is the integrated luminosity and $\mathcal{B}(\mathrm{t} \to \mathrm{b}\ell\nu) = 0.324$ is the total branching fraction of top decay into a b-quark and a lepton in the final state.

5.6 Uncertainties

The measurement of the cross section is affected by various sources of systematic uncertainty. The quantitative impact on the fit results from each of the sources is evaluated by performing pseudo-experiments with the corresponding variation. The individual uncertainty is taken from the difference between the nominal result of the fit and that

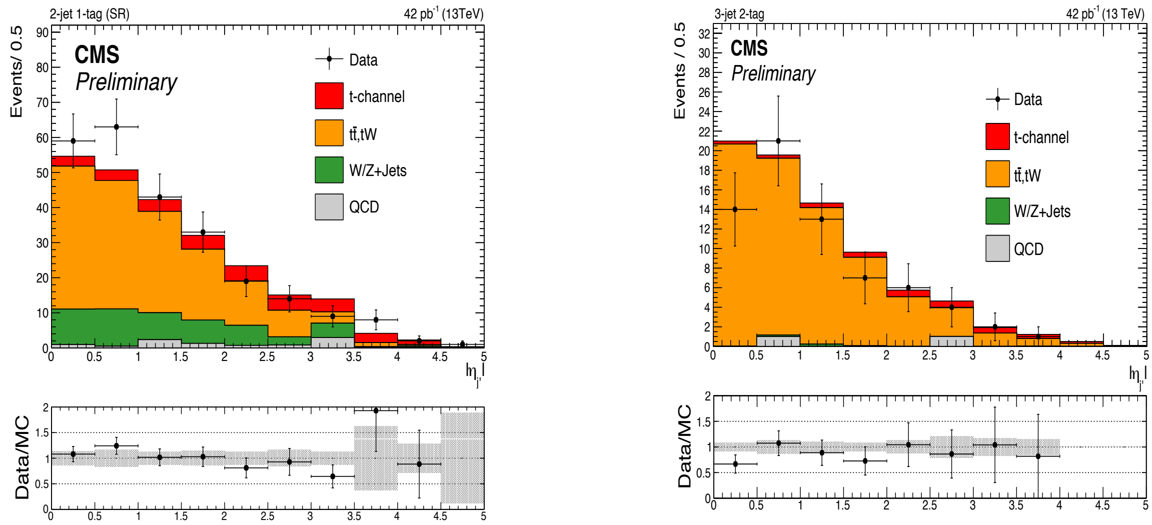


Figure 5.8: Post-fit $|\eta_{j'}|$ distribution in 2J1T region inside top mass window (left) and 3J2T region (right)

from the varied pseudo-experiment. The following sources of systematic uncertainties are considered.

- **Jet energy scale (JES):** All reconstructed jet four-momenta in the simulation are simultaneously varied according to the η and p_T -dependent uncertainties on the jet energy scale [32]. This variation in jet four-momenta is also propagated to $\not{E}_T$.

- **Jet energy resolution (JER):** A smearing is applied in the simulation to account for the known difference in jet energy resolution with respect to that in the data [32], increasing or decreasing the extra resolution contribution by its uncertainty.

- **b tagging:** b-tagging and mis-identification efficiencies are estimated from control samples in 13 TeV data. Scale factors (determined according to the method of Ref.[33]) are applied to the simulated samples to reproduce efficiencies in data and the corresponding uncertainties are propagated as systematic uncertainties.

- **Muon trigger and reconstruction:** Single-muon trigger and reconstruction efficiencies are estimated with a tag-and-probe method using DrellYan (DY) events. The uncertainty includes the difference in kinematics between events in DY data and single-top-quark data samples, as well as the dependence of the efficiencies from pileup events.

- **QCD extraction:** the normalization of the QCD multijet background is determined from a fit to the data as described in Section 5.3.1. In the fit to extract the single-top-quark cross section, a normalization uncertainty of 50% is assigned to account for the variations in the QCD background template.

- **Signal generator:** the results obtained using the nominal aMC@NLO signal samples are compared with the result obtained using signal samples generated with powheg to account for effects due to different NLO subtraction schemes.

- **Factorization and renormalization scales (Q^2):** The uncertainties due to variations in the renormalization and factorization scales are studied for the signal process, $t\bar{t}$ and W+jets by re-weighting the distributions with different combinations of halved/doubled factorization and renormalization scales and the envelope of the observed variations is taken as an estimate for the uncertainty.

- **PDF:** The uncertainty due to the choice of the parton distribution functions (PDF) is estimated using re-weighted histograms derived from the NNPDF 3.0 [34] replicas.

- **MET:** The effect on $\not{E}_T$ of a 10% uncertainty on the unclustered energy deposits in the calorimeters is estimated after subtracting from $\not{E}_T$ all jets and leptons.

- **Pile-up:** The uncertainty due to the modeling of the pile-up contributions has been estimated in a conservative way by comparing the event yields with and without applying pile-up weights to the simulation.

- **Luminosity:** the luminosity is known with a relative uncertainty of ±12%, utilizing the methods from [35].

Table 5.7 summarizes the contributions of the individual sources to the uncertainty on the measured cross section.

5.7 Results

Using an integrated luminosity of 42 pb^{-1} of pp collision data at 13 TeV, the cross section of the t-channel single top-quark production is measured to be

$$\sigma_{\text{t-ch}} = 274 \pm 98(\text{stat.}) \pm 52(\text{syst.}) \pm 33(\text{lumi.})\,\text{pb}, \tag{5.4}$$

which is in agreement with the standard model prediction. The measurement has an observed (expected) significance of 3.5 (2.7) standard deviations respectively. These values are evaluated based on the statistical uncertainty, which is the most dominant.
In the approximation, $|V_{td}|, |V_{ts}| << |V_{tb}|$, the top quark decay branching fraction into Wb is considered to be close to 1 and the magnitude of the CKM matrix element V_{tb} can be extracted as $|V_{tb}| = \sqrt{\frac{\sigma^{\text{meas.}}_{\text{t-ch}}}{\sigma^{\text{theo.}}_{\text{t-ch}}}}$, yielding

$$|V_{tb}| = 1.12 \pm 0.24(\text{expt.}) \pm 0.02(\text{theo.}) \tag{5.5}$$

From this result, the confidence interval for $|V_{tb}|$, assuming $|V_{tb}| < 1.0$, is determined in the Feldman-Cousins unified approach [122] which gives a lower bound of $|V_{tb}| > 0.7$ at the 95% confidence level.

Table 5.7: Relative impact of uncertainty sources with respect to the observed cross section value $\sigma_{t-\mathrm{ch}}^{\mathrm{obs}}$, given in percent.

Uncertainty source	$\Delta\sigma_{t-\mathrm{ch}}/\sigma_{t-\mathrm{ch}}^{\mathrm{obs}}$
Statistical uncertainty	36%
JES	17%
JER	1.1%
b-tagging	5.6%
Muon trigger/reconstruction	3.4%
QCD extraction	1.1%
Signal generator	1.9%
Factorization and renormalization scales (Q^2)	3.3%
PDF	4.5%
MET	1.2%
Pile-up	1.4%
Total systematic uncertainty	19%
Luminosity	12%
Total uncertainty	42%

Chapter 6

Cross section measurement with full 2015 data

The early analysis was followed by a precision measurement with full 2.2 fb^{-1} data collected by CMS during 2015 with 25ns bunch spacing. The full potential of ≈55 times more data compared to the early analysis presented in Chapter 5 is utilized by performing a multivariate analysis for better separation of signal from backgrounds. Further, the sample is divided based on the charge of the lepton in the final state and the ratio of top and anti-top production cross-sections is determined from data. A precise estimate of the CKM matrix element V_{tb} is determined from the results. The results are compared with SM prediction(s) and are found to agree with the prediction(s) within uncertainties. This analysis results in the most precise t-channel single top cross section measurement at $\sqrt{s}$ = 13 TeV till date.

6.1 Data and MC samples

The MC and data samples used for this analysis are listed below in Table 6.1 and Table 6.2.

Table 6.1: List of MC samples

Process	$\sigma(\cdot BR)$[pb]	Generator	Nevents
t-channel single top, top + anti-top, leptonic decays	217($\cdot$ 0.324) (NLO) [6, 119]	aMC@NLO-pythia8	20M
tW-channel single top, top, inclusive decays	35.6 (NNLL) [120]	powheg-pythia8	1M
tW-channel single top, anti-top, inclusive decays	35.6 (NNLL) [120]	powheg-pythia8	1M
$t\bar{t}$, inclusive decays	831.8 (NNLO) [58, 121]	powheg-pythia8	98M
W($\rightarrow \ell\nu$) + Jets	61526.7 (NLO)	aMC@NLO-pythia8	238M
Z/γ($\rightarrow \ell\ell$) + Jets, $m_{\ell\ell} > 50$ GeV	6025.2 (NLO)	aMC@NLO-pythia8	121M
μ -enriched QCD, $\hat{p}_T > 20$ GeV, $p_T^{\mu} > 15$ GeV	381304 (LO)	pythia8	22M

Table 6.2: List of Data samples

Dataset	Run Range	Integrated Luminosity (fb^{-1})
/SingleMuon/Run2015D-16Dec2015-v1/MINIAOD	256630-260627	2.2

6.2 Event Selection and Analysis Strategy

The following event selection criteria was applied to reduce backgrounds.

- Events are required to pass the high level trigger criteria which requires at least one prompt muon with p_T above 20 GeV ("HLT_IsoMu20_v") in the final state.
- Events with exactly one muon having $p_T > 22$ GeV within $|\eta| < 2.1$ passing "tight ID" criteria and $I_{rel} < 0.06$ is selected.
- Events with another muon above p_T threshold of 10 GeV within $|\eta| < 2.5$ are rejected in order to reduce Z+Jets and $t\bar{t}$ di-muon backgrounds.
- Events with another electron in addition to selected high-p_T muon described above with p_T above 20 GeV within $|\eta| < 2.4$ are rejected in order to reduce $t\bar{t}$ di-lepton background.
- Event having exactly 2 anti-k_T jets with cone size 0.4 made out of PF candidates are required which have p_T above 40 GeV within $|\eta| < 4.7$ passing the loose identification criteria based on their constituents and having $\Delta R > 0.4$ (in $\eta - \phi$ space) with respect to the high-p_T muon candidate.
- Out of the two selected jets one jet is required to pass the tight working point of the neural network based CSVv2 algorithm within $|\eta| < 2.4$ to be identified as a b-tagged jet originating from the b quark in final state due to the decay of the top quark. Thus we obtain the signal enriched region defined as 2J1T
- To reduce QCD background further, events are required to have $m_T^W > 50$ GeV (see Section 4.10).

An MVA discriminator is designed in order to distinguish signal from backgrounds. The MVA discriminant distribution is used as the fit variable in order to extract the cross section from data. W+Jets and $t\bar{t}$ background normalizations are constrained using the 3J1T and 3J2T regions respectively [20].

6.3 QCD Estimation

QCD estimation is performed using the isolation side-band ($I_{rel} > 0.12$) in the similar way as described in Section 5.3.1. Data is modeled using Eq.(5.1). The method is first validated in 2J0T and applied in the signal-rich 2J1T region.

6.3.1 Estimation in 2J0T region

The m_T^W QCD templates from isolated region and anti-isolated side-band is compared in order to check their agreement. Figure 6.1 shows the comparison and the QCD templates for agree with each other within the uncertainties for each of m_T^W and $\not{E}_T$. The data-

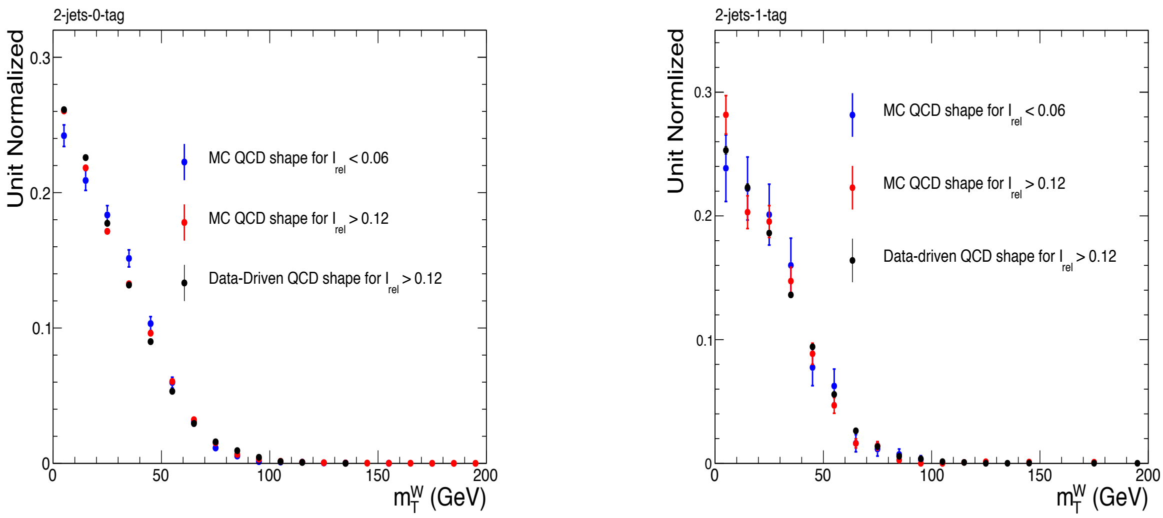


Figure 6.1: Comparison of QCD m_T^W templates in 2J0T region (left) and 2J1T region (right)

driven QCD templates are derived by subtracting contributions from all other processes, except QCD, scaled to their respective cross-sections from data in the anti-isolated side-band. The data-driven templates for QCD are then used to fit the data distributions of m_T^W in the isolated region. The nonQCD templates are obtained from MC simulation by summing up the individual contributions from the different processes scaled to the respective cross sections. The final QCD normalization is obtained by taking the integral of post-fit QCD templates scaled to fit result above 50 GeV for m_T^W .
In addition, the fit is performed separately for μ^+ and μ^- final states with the data-driven m_T^W -templates for QCD as shown in Figure 6.3. All fit results are listed in Tables 6.3 and 6.4. A conservative uncertainty of 50% is assigned to the QCD normalization obtained from the fit results corresponding to data-driven m_T^W template for QCD to cover for the yield differences, as explained in Section 5.3.1.

6.3.2 Estimation in 2J1T region

The same binned likelihood fit as in the 2J0T control region is applied to the m_T^W -distribution in the 2J1T region. The comparison of different QCD m_T^W templates can be found in Figure 6.1. The template for the non-QCD processes is again derived from MC simulation through the combination of the individual contributions from the different processes scaled according to their corresponding cross sections, while the QCD-template is again derived from anti-isolated side-band in data as described in Section 6.3.1. The fit is performed inclusively as well as separately for μ^+ and μ^- final states as shown in Figure 6.4. The fit results are quoted in Table 6.5.

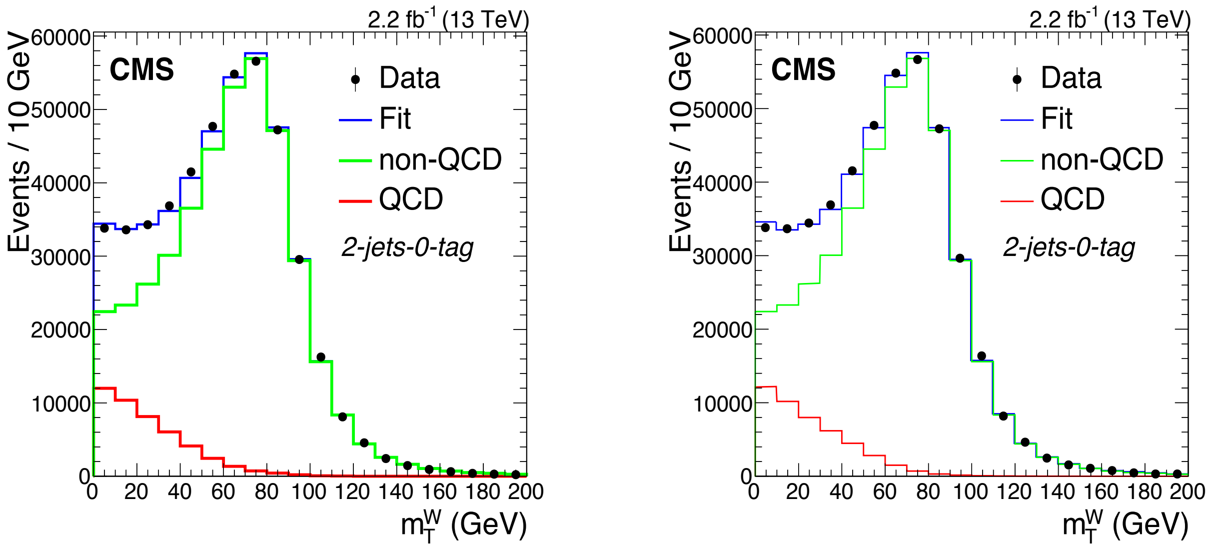


Figure 6.2: Fit to m_T^W in 2J0T region with QCD templates from anti-isolated region: data-driven (left) and MC (right)

Table 6.3: QCD estimation in the 2J0T region with different templates

Variable	Fit Type	QCD shape	Process	Fitted Yield (full range)	Extrapolated Yield ($m_T^W > 50$GeV)
m_T^W	Binned	From AntiIso MC	QCD	46619±538	5646±65
			nonQCD	404485±805	266108±529
		Data Driven	QCD	45884±529	5223±60
			nonQCD	405220±800	266592±526

Table 6.4: QCD estimation in the 2J0T region separately for final states with positively and negatively charged muon

μ charge	Process	Fitted Yield (full m_T^W range)	Extrapolated Yield ($m_T^W > 50$ GeV)
μ^+	QCD	23177±388	2717±46
	nonQCD	220913±590	145388±389
μ^-	QCD	22713±360	2507±40
	nonQCD	184300±539	121199±355

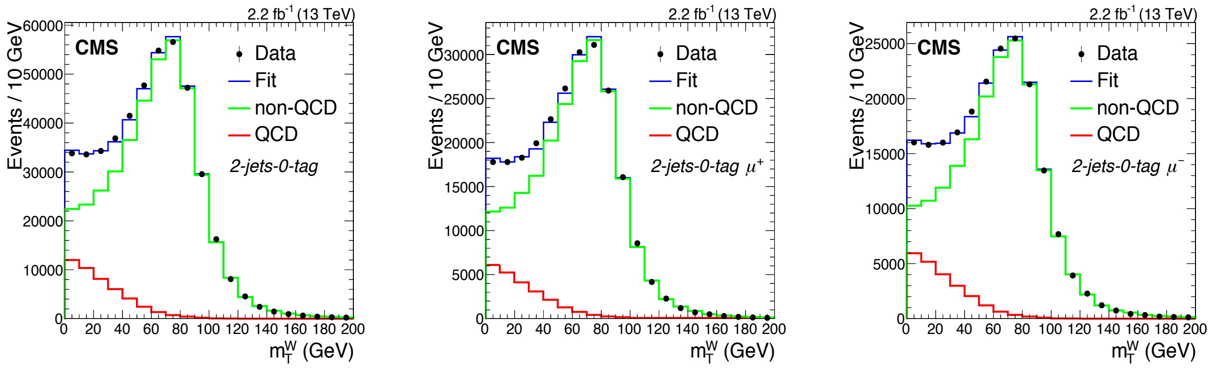


Figure 6.3: m_T^W -Fit in 2J0T region with data-driven QCD templates from anti-isolated region for inclusive (left), μ^+ (middle) and μ^- (right) final states

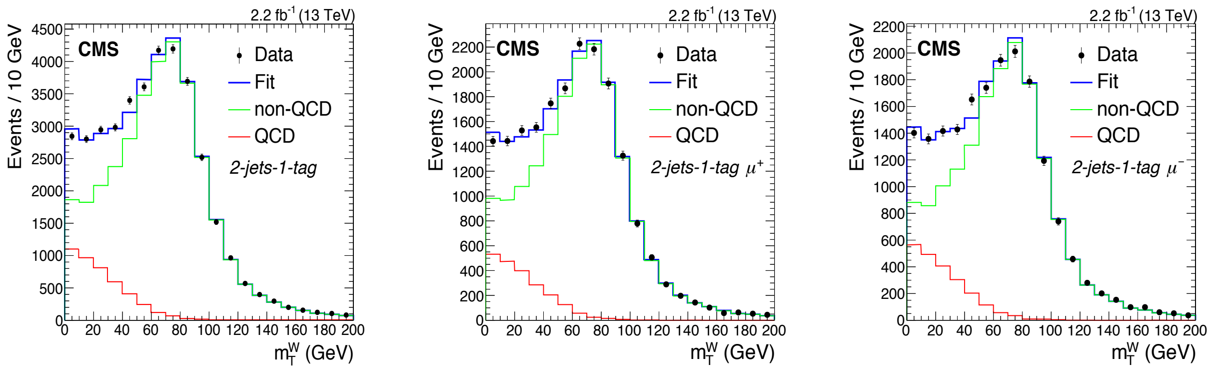


Figure 6.4: m_T^W -Fit in 2J1T region with data-driven QCD templates from anti-isolated region for inclusive (left), μ^+ (middle) and μ^- (right) final states

Table 6.5: QCD estimation in the 2J1T region separately for final states with positively and negatively charged muon

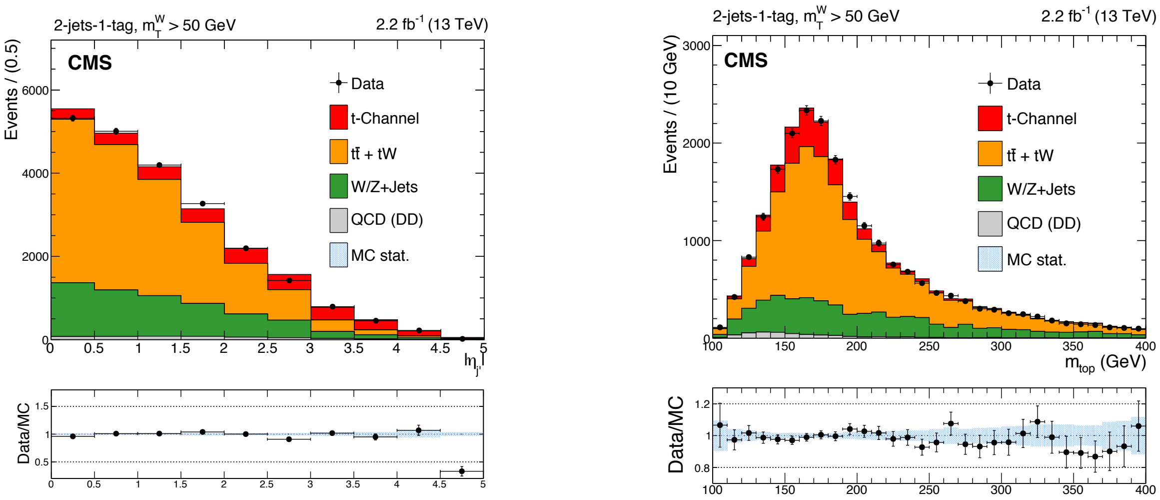
μ charge	Process	Fitted Yield (full m_T^W range)	Extrapolated Yield (m_T^W > 50 GeV)
inclusive	QCD	4317±2158	574±287
	nonQCD	33244±228	21361±153
μ^+	QCD	2126±1063	308±154
	nonQCD	17326±165	11082±84
μ^-	QCD	2200±1100	266±133
	nonQCD	15908±157	10279±78

6.4 Multivariate Discriminator

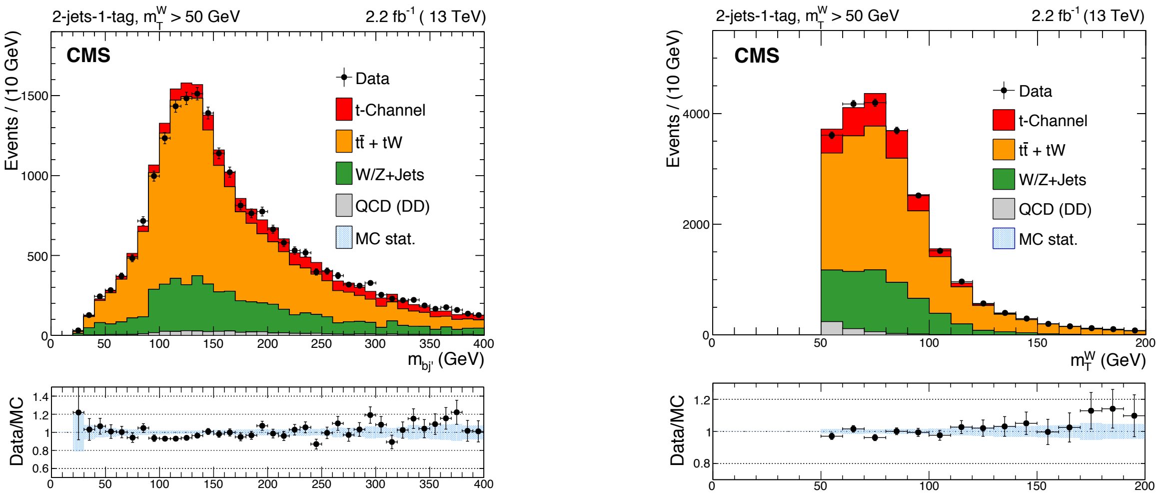
Multivariate analysis technique have been used to optimally discriminate signal from backgrounds. Several input variables are used to design the neural network. The input variables are ranked according to their importance in Table 6.6. The importance is defined as the loss of significance when removing this variable from the list. The variable with maximum importance is the pseudo-rapidity of the light-quark jet, i.e., $|\eta_{j'}|$. This is not surprising since the presence of a light-quark jet which is recoiling against the produced top quark, in high-$|\eta|$ region, is a typical feature of the signal topology. The second most important variable is the invariant mass of the reconstructed top quark, which discriminates processes with top quarks, from background processes without any produced top quark. All input variables are validated by comparing the distributions in data with those in the simulations as shown in Figures 6.5 and 6.6. The artificial neural network available (multi-layer perceptron with bayesian regulator) in the standard TMVA package [123] is used. The correlations of the input variables (Figure 6.7) are taken into account by the neural network.
Simulated t-channel single top quark events are trained against simulated $t\bar{t}$ and W+ jets events, as well as QCD multijet events from a side-band region in data, with all processes weighted according to their predicted relative contribution. The neural network is trained on a subset of the simulated samples. Application on the remaining sample shows similar performance and no sign of over-training is observed as shown in Figure 6.8. The area under the receiver-operator-characteristics (ROC) curve is found to be 85%. The neural network is trained using the inclusive 2J1T sample for events with positively and negatively charged muons, and afterwards applied to the 2J1T, 3J1T, and 3J2T data samples, each split in two, depending on the charge of the muon.

Table 6.6: Input variables to MVA ordered according to their importance

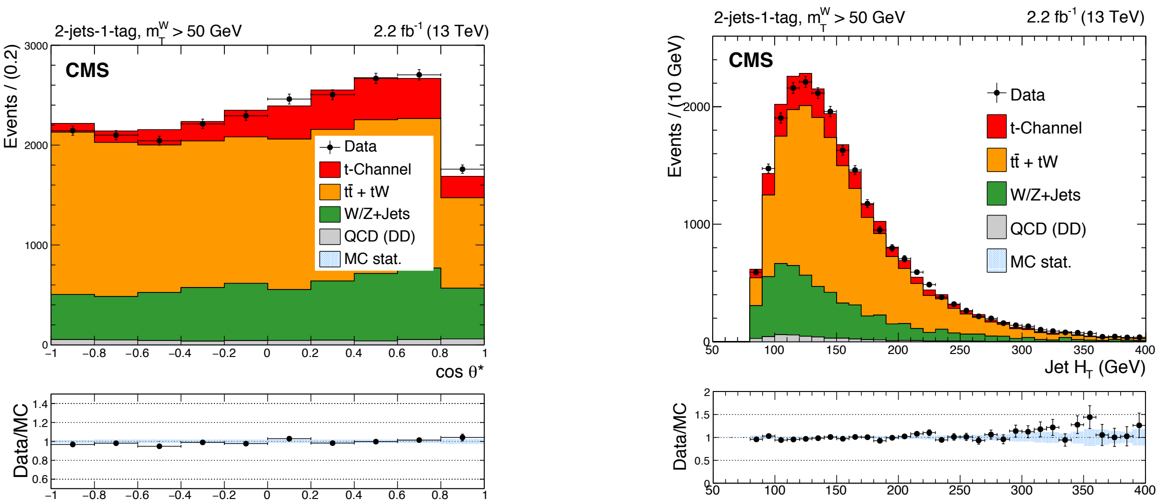
Rank	Variable	Description	Importance
1	$\|\eta_{j'}\|$	Light-quark jet $\|\eta\|$	59.07
2	m_{top}	Reconstructed top quark mass	43.04
3	$m_{bj'}$	Di-jet invariant mass	29.08
4	m_T^W	Transverse W boson mass	26.49
5	$\cos\theta^*$	Cosine of the angle between lepton and light-quark in the rest frame of the top quark	23.29
6	Jet H_T	Scalar sum of the p_T of b-tagged jet and light-quark jet	14.54
7	$m_j^{hardest}$	Mass of the jet with largest p_T	10.35
8	$\Delta R_{bj'}$	ΔR between b-tagged jet and light-quark jet in $\eta - \phi$ space	7.83
9	$p_T^{j'}$	p_T of light-quark jet	5.48
10	$m_{j'}$	mass of the light-quark jet	4.95
11	$\|\eta\|$ of W	$\|\eta\|$ of reconstructed W boson	4.49



(a) $|\eta_{j'}|$ (left) and $\mathrm{m_{top}}$ (right)

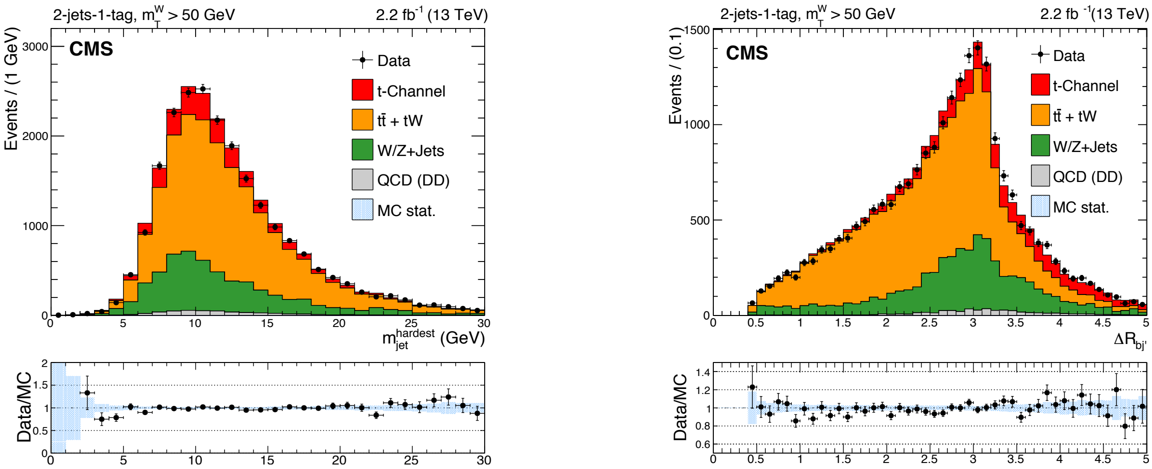


(b) $\mathrm{m}_{bj'}$ (left) and $\mathrm{m_T^W}$ (right)

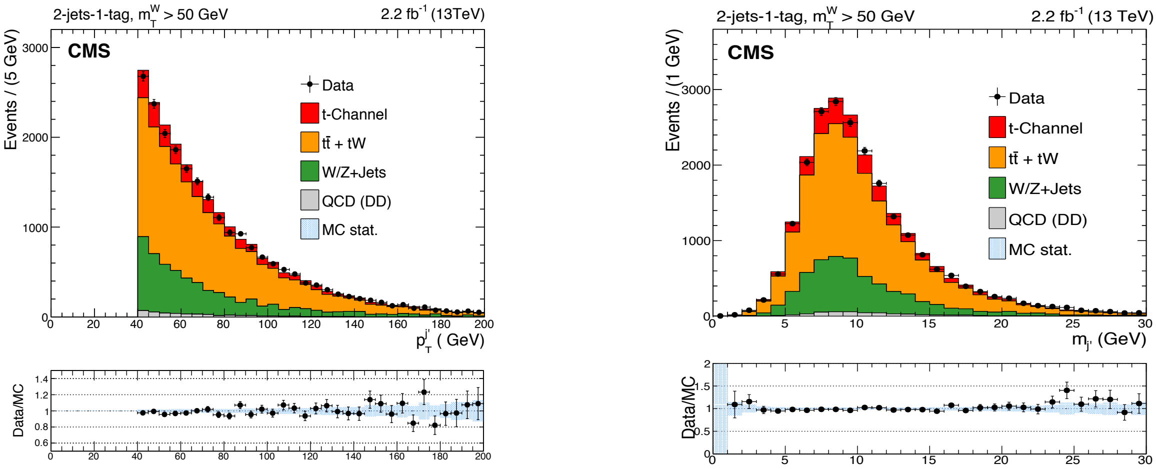


(c) $\cos\theta^*$ (left) and jet $\mathrm{H_T}$ (right)

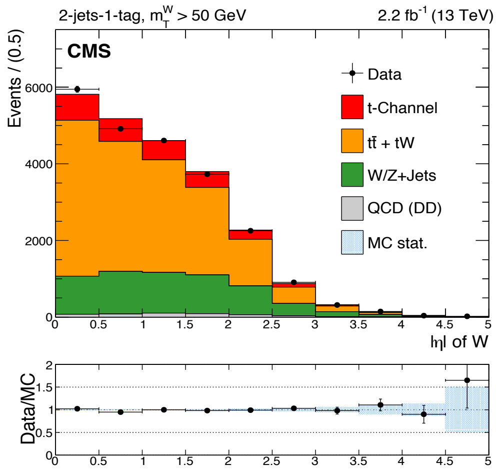
Figure 6.5: Data-MC comparison of MVA input variables in 2J1T region



(a) $m_j^{hardest}$ (left) and $\Delta R_{bj'}$ (right)

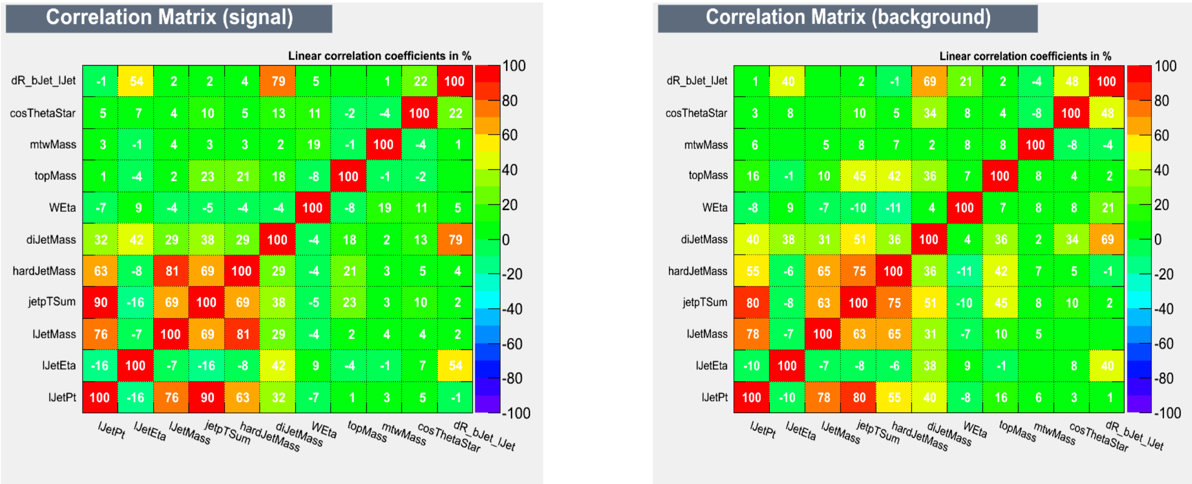


(b) $p_T^{j'}$ (left) and $m_{j'}$ (right)



(c) $|\eta|$ of W boson

Figure 6.6: Data-MC comparison of MVA input variables in 2J1T region



(a) t-channel single top signal

(b) Backgrounds

Figure 6.7: Correlation of input variables in the training sample for signal and backgrounds

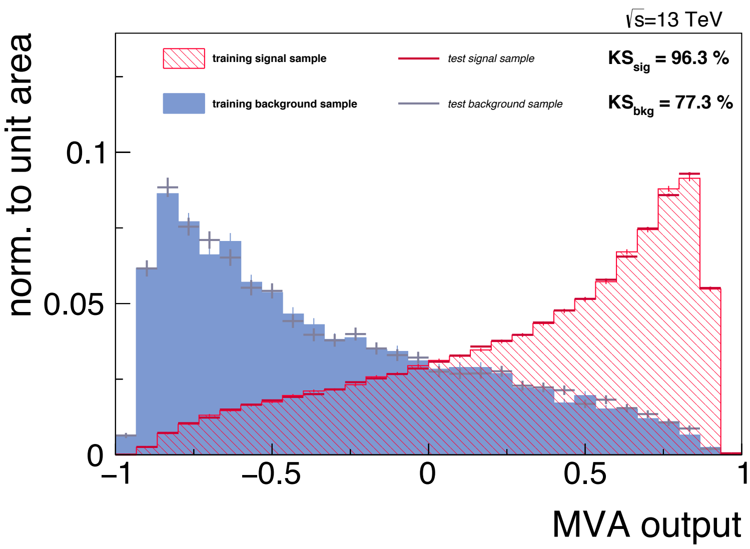


Figure 6.8: Comparison of neural network output from training samples against test samples for over-training check

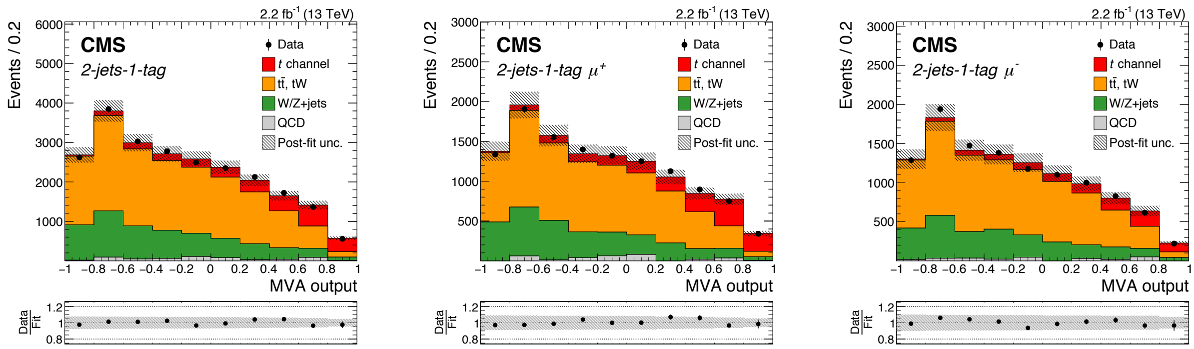
6.5 Cross section extraction

Binned likelihood fits are performed on the distributions of the MVA discriminators to extract signal cross section. The background contributions are made up of three templates to account for: i) top quark production including $t\bar{t}$ and tW, ii) electroweak production including W+jets and Z+jets processes, and iii) QCD multijet production. The fit is performed using the BarlowBeeston method [21] which correctly accounts for

Table 6.7: Post-fit values of normalization scale factors and R_{t-ch} from fit

Process Category	Post-fit value
Signal, t-channel	1.13±0.08
Top ($t\bar{t}$ and tW-channel)	1.00±0.02
EWK (W+Jets and Z+Jets)	1.11±0.09
QCD multijet	0.86±0.29
R_{t-ch}	1.81±0.19

limited-size simulation samples. The distributions of the MVA discriminators in the signal region (2J1T) and the two control regions (3J1T and 3J2T) are fitted simultaneously. As the latter are dominated by $t\bar{t}$ events, including these control regions improves the precision of the $t\bar{t}$ contribution determination. The free parameters of the fit are the scale factor for the normalization of the single top quark production, the scale factors for the normalization of the background processes, and the ratio of single top quark to top anti-quark production $R_{t-ch.}$. The background scale factors are constrained by log-normal priors with an uncertainty of 10% for the top quark background, 30% for the electroweak background, and 50% for the QCD multijet background. The latter is motivated by the uncertainties in the QCD estimation from data, while the other two are determined by the uncertainty on the theoretical cross sections. The scale factors are defined as before in Eq. 5.2. where N_i is the number of events after the fit, $N_i^{pred.}$ the predicted number of events and i the process category. Table 6.7 shows the results obtained from the fit for events with a positively charged muon. The fitted distributions are shown in Figure 6.9.



(a) Inclusive (b) μ^+ (c) μ^-

Figure 6.9: Post fit distribution of the MVA discriminant in signal region

6.6 Uncertainties

The measurement of the cross section is affected by various sources of systematic uncertainties, which can be grouped into two categories, experimental uncertainties and theoretical uncertainties. Several of the former category of uncertainties are considered as nuisance parameters in the fit to the MVA discriminator distribution and are thus included in the total uncertainty of the fit. To determine the impact of the sources of the remaining uncertainties, pseudo-experiments are performed. Pseudo-data are drawn from the nominal samples. Fits to the discriminator distributions are performed with templates, including the variations in the shapes that correspond to systematic variations of one standard deviation. The difference between the mean values of the results from these fits, and from fits using the nominal shapes as fit templates, is taken as an estimation for the corresponding uncertainty. The contributions from different sources are summed together with the method in Ref. [124]: the asymmetric components of each uncertainty are treated as the standard deviations of two halved Gaussian functions, and thus the convolution of the resulting distributions for all uncertainties is performed by making use of Thie′le′s semi-invariants [124].

Experimental Uncertainties - included in the fit

The following sources of systematic uncertainty are included in the fit either through the applied Barlow-Beeston method or by using nuisance parameters in the fit (profiled uncertainties). By variations of the default samples, two dedicated variation templates corresponding to ± 1 standard deviations of the respective uncertainty source are created. These variation templates are interpolated using the piece-wise linear "template morphing" model (see page 48-50 of [125]) and the corresponding nuisance parameter is determined from the fit.

- **Limited size of samples of simulated events:** To account for the limited number of available simulated events the fit is performed using the BarlowBeeston method [21], and the effect is therefore included in the total uncertainty of the fit. To estimate the impact of the sample size the nominal central value is compared with the central value obtained without the BarlowBeeston method. The latter effectively corresponds to assuming an infinite size of the samples of simulated events.

- **Jet energy scale (JES):** All reconstructed jet four-momenta in simulated events are simultaneously varied according to the η- and p_T -dependent uncertainties in the JES [13]. This variation in jet four-momenta is also propagated to $\not{E}_T$.

- **Jet energy resolution (JER):** A smearing is applied to account for the difference in the JER between simulation and data [13], increasing or decreasing the resolutions by their uncertainties.

- **b-tagging:** b tagging and mis-identification efficiencies are estimated from control samples in 13 TeV data [116]. Scale factors are applied to the simulated samples to reproduce efficiencies observed in data and the corresponding uncertainties are propagated as systematic uncertainties.

- **Muon trigger and reconstruction:** Single-muon trigger efficiency and reconstruction efficiency are estimated with a tag-and-probe method [108] from DY events in the Z boson mass peak. To take the difference in kinematic properties between DY and the single top quark process into account, an additional systematic uncertainty depending on the number of jets in an event is applied.

Experimental Uncertainties - not included in the fit

- **Pile-up:** The uncertainty in the average expected number of pileup interactions is propagated as a source of systematic uncertainty to this measurement by varying the minimum bias cross section by ±5%. The effect on the result is found to be negligible and is therefore not considered further.
- **Luminosity:** The integrated luminosity is known with a relative uncertainty of ±2.3% [87].

Modelling uncertainties

- **Signal modelling:** To estimate the influence of possible mis-modelling of the signal process, the default sample (MG5 aMC@NLO) is compared to a sample generated with powheg, another NLO matrix-element generator. The effect of different PS models is estimated by comparing the default sample (MG5 aMC@NLO interfaced with pythia) with a sample using a different PS description (MG5 aMC@NLO interfaced to HERWIG++).
- **$t\bar{t}$ modelling:** For the estimation of the uncertainty due to possible mis-modelling of the $t\bar{t}$ background, the same procedure as for the signal modelling is applied. The default sample, generated with POWHEG, is compared to a sample generated with MG5 aMC@NLO to estimate the impact of the choice of the matrix-element generator, and the two PS models implemented in PYTHIA and HERWIG++ [126] are compared to estimate the influence of the PS modelling.
- **W+jets modelling:** The impact of incorrectly modelled relative fractions of W boson production in association with heavy flavour jets in the W+jets sample is estimated by varying the fractions of W+b and W+c events independently by ±30%.
- **Modelling of the top quark p_T :** Differential measurements of the top quark p_T in $t\bar{t}$ events [127] have shown that a harder spectrum is predicted than observed. Therefore, the results derived using the default simulation for $t\bar{t}$ are compared to the results using simulated $t\bar{t}$ events that are re-weighted according to the observed difference between data and simulation in Ref. [127].
- **Renormalization and factorization scale uncertainty (μ_R/μ_F):** The uncertainties due to variations in the renormalization and factorization scales are studied for the signal process, tW, $t\bar{t}$, and W + jets by reweighting the distributions with different combinations of halved/doubled factorization and renormalization scales. The effect is estimated for each process separately.

Table 6.8: Relative Impact of different uncertainty sources on the measurement

Uncertainty Source	$(\frac{\Delta\sigma}{\sigma})_{\mathrm{t-ch,t+\bar{t}}}$	$(\frac{\Delta\sigma}{\sigma})_{\mathrm{t-ch,t}}$	$(\frac{\Delta\sigma}{\sigma})_{\mathrm{t-ch,\bar{t}}}$	$\Delta\mathrm{R}_{\mathrm{t-ch}}/\mathrm{R}_{\mathrm{t-ch}}$
Statistical unc.	$\pm5.5\%$	$\pm5.3\%$	$\pm11.5\%$	$\pm9.7\%$
Profiled exp. unc.	$\pm5.2\%$	$\pm5.7\%$	$\pm4.9\%$	$\pm3.3\%$
Total fit unc.	$\pm7.6\%$	$\pm7.8\%$	$\pm12.5\%$	$\pm10.3\%$
Luminosity	$\pm2.3\%$	$\pm2.3\%$	$\pm2.3\%$	–
Signal modelling	$\pm6.9\%$	$\pm8.2\%$	$\pm8.5\%$	$\pm5.3\%$
$\mathrm{t\bar{t}}$ modelling	$\pm3.9\%$	$\pm4.3\%$	$\pm4.5\%$	$\pm4.0\%$
W+jets modelling	$^{+2.1}_{-1.8}\%$	$^{+2.3}_{-1.6}\%$	$^{+2.3}_{-2.5}\%$	$^{+2.0}_{-1.7}\%$
t-channel $\mu_{\mathrm{R}}/\mu_{\mathrm{F}}$ scale	$^{+6.1}_{-4.6}\%$	$^{+5.2}_{-5.7}\%$	$^{+5.1}_{-7.2}\%$	$^{+1.2}_{-0.7}\%$
$\mathrm{t\bar{t}}$ $\mu_{\mathrm{R}}/\mu_{\mathrm{F}}$ scale	$^{+2.9}_{-3.5}\%$	$^{+4.1}_{-3.5}\%$	$^{+3.1}_{-4.7}\%$	$^{+1.0}_{-1.1}\%$
tW-channel $\mu_{\mathrm{R}}/\mu_{\mathrm{F}}$ scale	$^{+0.5}_{-0.3}\%$	$^{+0.8}_{-0.6}\%$	$^{+0.7}_{-1.1}\%$	$^{+0.1}_{-0.2}\%$
W+jets $\mu_{\mathrm{R}}/\mu_{\mathrm{F}}$ scale	$^{+3.7}_{-2.9}\%$	$^{+3.0}_{-3.5}\%$	$^{+3.8}_{-4.9}\%$	$^{+0.9}_{-1.2}\%$
PDF	$^{+1.9}_{-1.5}\%$	$^{+1.6}_{-2.1}\%$	$^{+2.1}_{-1.8}\%$	$^{+2.5}_{-2.2}\%$
Top quark $\mathrm{p_T}$ modelling	$\pm0.1\%$	$\pm0.2\%$	$\pm0.2\%$	$\pm0.1\%$
Total Theory unc.	$^{+11.1}_{-10.7}\%$	$^{+12.1}_{-12.2}\%$	$^{+12.9}_{-13.6}\%$	$\pm7.5\%$
Total unc.	$^{+13.7}_{-13.4}\%$	$\pm14.7\%$	$^{+18.2}_{-18.7}\%$	$\pm12.7\%$

- **PDF:** The uncertainty due to the choice of PDFs is estimated using re-weighted histograms derived from all PDF sets of NNPDF 3.0 [128].

The impact of the individual uncertainty sources are summarized in Table 6.8.

6.7 Results

The cross section for the production of single top quarks and the top quark to top anti-quark cross section ratio as a result of the fit are as follows.

$$\sigma_{\mathrm{t-ch,t}} = 154 \pm 8(\mathrm{stat.}) \pm 9(\mathrm{expt.}) \pm 19(\mathrm{theo.}) \pm 4(\mathrm{lumi.})\ \mathrm{pb} = 154 \pm 22\ \mathrm{pb} \quad (6.1)$$

$$R_{\mathrm{t-ch.}} = 1.81 \pm 0.18(\mathrm{stat.}) \pm 0.15(\mathrm{syst.}) \quad (6.2)$$

A comparison between the measured ratio and the prediction of different PDF sets is shown in Figure 6.10.
Using the $\sigma_{\mathrm{t-ch,t}}$ and $\mathrm{R}_{\mathrm{t-ch}}$ measurements, the cross section of the top anti-quark production is computed as

$$\sigma_{\mathrm{t-ch,\bar{t}}} = 85 \pm 10(\mathrm{stat.}) \pm 4(\mathrm{expt.}) \pm 11(\mathrm{theo.}) \pm 2(\mathrm{lumi.})\ \mathrm{pb} = 85 \pm 16\ \mathrm{pb}, \quad (6.3)$$

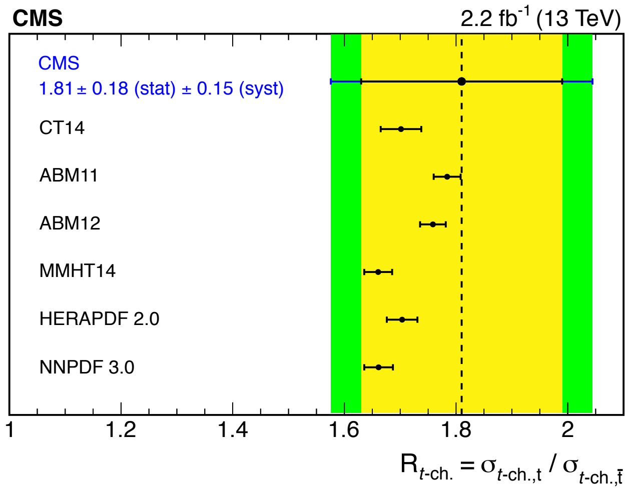


Figure 6.10: Comparison of measured $\mathrm{R_{t-ch.}}$ value with different PDF predictions

where the uncertainties are evaluated using the correlation matrix of the simultaneous fit. This leads to the total cross section,

$$\sigma_{\mathrm{t-ch,t+\bar{t}}} = 238 \pm 13(\mathrm{stat.}) \pm 12(\mathrm{expt.}) \pm 26(\mathrm{theo.}) \pm 5(\mathrm{lumi.})\ \mathrm{pb} = 238 \pm 32\ \mathrm{pb} \quad (6.4)$$

Figure 6.11 shows the comparison of this measurement with the SM expectation and measurements of the single top quark t-channel cross section at other CM energies. The total cross section is used to determine the absolute value of the CKM matrix element $|V_{tb}|$ defined as $|f_L^V V_{tb}| = \sqrt{\frac{\sigma^{\mathrm{meas.}}_{\mathrm{t-ch,t+\bar{t}}}}{\sigma^{\mathrm{theo.}}_{\mathrm{t-ch,t+\bar{t}}}}}$ with the assumption that $|V_{tb}| >> |V_{ts}|, |V_{td}|$ and $\sigma^{\mathrm{theo.}}_{\mathrm{t-ch,t+\bar{t}}}$ taken from Table 2.2.

$$|f_L^V V_{tb}| = 1.05 \pm 0.07(\mathrm{expt.}) \pm 0.02(\mathrm{theo.}) \quad (6.5)$$

The possible presence of an anomalous tWb coupling is taken into account by an anomalous form factor f_L^V [1], which is 1 in the SM and deviates from 1 for physics beyond the SM as discussed in Section 2.2.3. The $\sigma^{\mathrm{theo.}}_{\mathrm{t-ch,t+\bar{t}}}$ is the SM prediction from which assumes $|V_{tb}| = 1$. The first uncertainty on the r.h.s of Eq.(6.5) contains all uncertainties on the cross section measurement, and the second uncertainty is the uncertainty on the theoretical SM prediction.
The measured total cross section of t-channel single top production is currently the most precise result at $\sqrt{s}$ =13 TeV. The unique feature was to extract $\mathrm{R_{t-ch.}}$ from data without making a priori assumption on the correlation of the uncertainties for top quark and anti-quark production cross-section measurements. With future data, this observable is expected to be sensitive to different PDF descriptions.

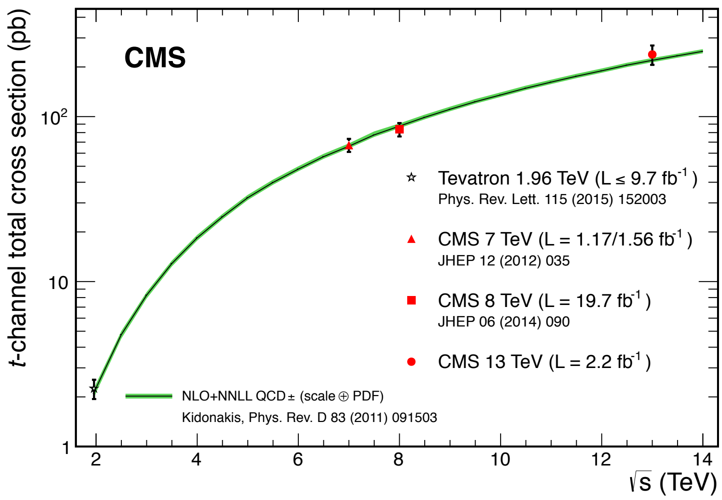


Figure 6.11: Comparison of t-channel single top production cross section measurements by CMS and TeVatron with SM predictions at different $\sqrt{s}$

Chapter 7

Top quark mass measurement in single top events with 36 fb^{-1} data collected in 2016

The latest world average of the top mass, based on measurements performed in top-antitop pair ($t\bar{t}$) events, from ATLAS, CMS , CDF & D0 is $m_{top} = 173.34 \pm 0.27(stat) \pm 0.71(syst)$ GeV [8]. As discussed in Section 2.3.2, measurments in the single top topology has distinct advantage in dealing with modelling uncertainties of the colour-flow in the event. Also, single top events provide an independent statistical sample to measure the mass of top quark. It occurs at a lower energy scale compared to $t\bar{t}$ production. It enriches the range of available measurements with systematics partially uncorrelated from those considered in $t\bar{t}$ events. Previous measurements of top quark mass in single top events from CMS [129] and ATLAS [130] with data collected at $\sqrt{s} = 8$ TeV are quoted in Eqs. 7.1 and 7.2 respectively.

$$m_t = 172.95 \pm 0.77\,(stat.)^{+0.97}_{-0.93}\,(syst.)\,GeV \tag{7.1}$$

$$m_t = 172.2 \pm 0.7\,(stat.)\,\pm 2.0\,(syst.)\,GeV \tag{7.2}$$

We report herein a measurement of the top-quark mass in t-channel single top events using 35.9 fb^{-1} of data collected by the CMS detector in 2016[1] at a CM energy of 13 TeV. The striking feature of this process is the presence of a light quark recoiling against the top, which decays to a b-quark and a W boson. This analysis is focused on the muon final state arising from direct and cascade (via τ leptons) decays of the W boson originating from the top decay. The outgoing b quark hadronizes to form a jet that has characteristic signature inside the detector. The energy and momenta of the top-quark decay products including the neutrino are measured or estimated to calculate the

[1]The analysis has been slightly modified to use the full potential of almost 3 times more available data compared to that presented in Synopsis.

invariant mass of the $\mu\nu$b system. An event selection tailored to obtain high purity t-channel single top sample is used to perform the measurement and thereby ensured to have a result as independent as possible from those from $t\bar{t}$. We also attempt to measure the masses of top and anti-top quarks separately by selecting events with positively and negatively charged muons respectively and determine the mass difference.

7.1 Data and MC samples

All the processes consisting of one muon (μ) and two or more jets are considered in this analysis. Signal samples are generated with powheg. Other background processes considered are tW-channel single top, s-channel single top, $t\bar{t}$, W+Jets, Z+Jets, QCD and dibosons(WW,WZ and ZZ). A total of 35.9 fb^{-1} data collected with the CMS experiment during 2016 at a center-of-mass energy of 13 TeV is used in the analysis. Good runs in data are selected for the analysis based on the performance of different sub-detectors and physics object reconstruction at the online and offline level. Table 7.2 and Table 7.1 summarizes MC simulation and data samples used in this analysis.

Table 7.1: List of Data samples

Run Period	Run Range	Dataset Name	Integrated Luminosity (fb^{-1})
Run B	272007 - 275376	/SingleMuon/Run2016B-23Sep2016-v3/MINIAOD	5.8
Run C	275657 - 276283	/SingleMuon/Run2016C-23Sep2016-v1/MINIAOD	2.6
Run D	276315 - 276811	/SingleMuon/Run2016D-23Sep2016-v1/MINIAOD	4.2
Run E	276831 - 277420	/SingleMuon/Run2016E-23Sep2016-v1/MINIAOD	4.0
Run F	277772 - 278808	/SingleMuon/Run2016F-23Sep2016-v1/MINIAOD	3.1
Run G	278820 - 280385	/SingleMuon/Run2016G-23Sep2016-v1/MINIAOD	7.5
Run H	280919 - 284044	/SingleMuon/Run2016H-PromptReco-v2/MINIAOD	8.4
		/SingleMuon/Run2016H-PromptReco-v3/MINIAOD	0.2
Total			35.9

7.2 Event Selection

The following event selection criteria was applied to reduce backgrounds.

- Events are required to pass the high level trigger criteria which requires at least one prompt muon with p_T above 24 GeV in the final state.
- Events with exactly one muon having $p_T > 26$ GeV within $|\eta| < 2.1$ passing "tight ID" criteria and $I_{rel} < 0.06$ is selected.
- Events with another muon above p_T threshold of 10 GeV within $|\eta| < 2.4$ passing "loose ID" criteria and having $I_{rel} < 0.2$ are rejected in order to reduce Z+Jets and $t\bar{t}$ di-muon backgrounds.

Table 7.2: List of MC samples

Process	$\sigma(\cdot BR)$[pb]	Generator	Nevents
t-channel,top, inclusive decays	136.02 (NLO) [6, 119]	powheg-pythia8	67240808
t-channel, anti-top, inclusive decays	80.95 (NLO) [6, 119]	powheg-pythia8	38811017
s-channel, top+anti-top, leptonic decays	10.32(0.324) (NNLL) [119]	aMC@NLO-pythia8	1000000
tW-channel, top, inclusive decays	35.6 (NNLL) [120]	powheg-pythia8	992024
tW-channel, anti-top, inclusive decays	35.6 (NNLL) [120]	powheg-pythia8	998276
$t\bar{t}$, inclusive decays	831.76 (NNLO+NNLL) [58, 121]	powheg-pythia8	77229341
W($\rightarrow \ell\nu$)+Jets, inclusive	61466.1 (NNLO)	aMC@NLO-pythia8	23948810
Z/γ^*($\rightarrow \ell^+\ell^-$)+Jets, inclusive($M_{\ell\ell} > 50$ GeV)	5765.4 (NNLO)	aMC@NLO-pythia8	28968252
WW$\rightarrow 1\ell1\nu2$Q	45.85	aMC@NLO-pythia8	5176114
WW$\rightarrow 2\ell2\nu$	12.178	powheg-pythia8	1999000
WZ$\rightarrow 1\ell1\nu2$Q	10.71	aMC@NLO-pythia8	24221923
WZ$\rightarrow 2\ell2$Q	5.595	aMC@NLO-pythia8	26517272
ZZ$\rightarrow 2\ell2$Q	3.22	aMC@NLO-pythia8	15345572
μ-enriched QCD	720648000*0.00042 (LO)	pythia8	22094081

- Events with another electron in addition to selected high-p_T muon described above with p_T above 15 GeV within $|\eta| < 2.4$ are rejected in order to reduce $t\bar{t}$ di-lepton background.
- Event having exactly 2 anti-k_T jets with cone size 0.4 made out of PF candidates are required which have p_T above 40 GeV within $|\eta| < 4.7$ passing the loose identification criteria based on their constituents and having $\Delta R > 0.4$ (in $\eta - \phi$ space) with respect to the high-p_T muon candidate.
- Out of the two selected jets one jet is required to pass the tight working point of the neural network based cMVAv2 algorithm within $|\eta| <2.4$ to be identified as a b-tagged jet originating from the b-quark in the final state due to the decay of the top quark. Thus we obtain the signal enriched region defined as 2J1T
- To reduce QCD background further, events are required to have $m_T^W > 50$ GeV as explained in Section 4.10.

An MVA discriminator is designed in order to distinguish signal from backgrounds (see Section 7.5). A cut on MVA discriminant is optimized and applied further to have a high-purity signal selection for the top quark mass measurement in single top events.

7.3 Cut Flow

Table 7.3 summarizes the step by step event yields for simulated signal and background events as well as in data, when the object selections are applied sequentially.

Table 7.3: Cut flow table

Cut	t-channel	tW+s -channel	$t\bar{t}$	W+Jets	Z+Jets	di-boson	QCD	Total MC	Data	Data/MC(%)
No Cut	7778510	2922066	2.98178e+07	2.20583e+09	2.06715e+08	2780995	1.08515e+010	13307344371	786809782	5.91
Trigger	522403	431879.5	4.03312e+06	3.09301e+08	4.46127e+07	576838.2	1.56223e+08	515700940.7	469761390	91.09
==1 Tight isolated μ	406617	331738	2.97164e+06	2.33916e+08	2.28107e+07	434859.5	2.47365e+07	285608054.5	258382120	90.47
Loose μ veto	404505	319712.8	2.84126e+06	2.33563e+08	1.29449e+07	406448.96	2.38424e+07	274322226.76	245109050	89.35
ele veto	402917	292559.1	2.56513e+06	2.33386e+08	1.28375e+07	375387.93	2.38121e+07	273671594.03	244545880	89.36
== 2 Jets	151988.9	105039.5	695845	6.60088e+06	590102	81312.9	1.46003e+06	9685198.3	8000147	82.60
== 1 tag	64472.5	39140.78	278100	103099	18926.7	2160.19	150628	656527.18	544798	82.98
$m_T^W > 50$ GeV	44946.6	26523.49	191354	74884.2	9781.65	1407.99	36526.1	385424.04	340453	88.33

7.4 Estimation of QCD background

The transverse mass (m_T^W) determined from the (μ, $\not{E}_T$) final state provides good discrimination power between QCD and processes with prompt muons. An extended maximum likelihood fit with 3 parameters is performed to the m_T^W distribution in the signal and control regions. We assume that the m_T^W -distribution in data $F(m_T^W)$ can be parameterised as

$$F(m_T^W) = N_{QCD}\, Q(m_T^W) + N_{EWK}\, E(m_T^W) + N_{Top}\, T(m_T^W) \quad (7.3)$$

where $Q(m_T^W)$ stands for the QCD m_T^W - template, $E(m_T^W)$ is the combined m_T^W - template of W+Jets, Z+Jets and dibosons (WW+WZ+ZZ) and $T(m_T^W)$ is the combined m_T^W -template of single top and $t\bar{t}$ processes. Both $E(m_T^W)$ and $T(m_T^W)$ are obtained by summing up the simulated contributions from the individual component processes in the isolated region according to their predicted cross section. The $Q(m_T^W)$ is typically derived from a side-band region defined by $I_{rel} > 0.2$, which is dominated by QCD. All templates,i.e., $Q(m_T^W)$, $E(m_T^W)$ and $T(m_T^W)$ are then normalized to an integral of 1.0. The fit parameter N_{QCD} denotes the number of QCD events and N_{EWK} represents the total number of W+jets, Z+jets and diboson events and N_{Top} is the total number of single top and top pair events. N_{QCD}, N_{EWK} and N_{Top} are allowed to float during the fit to the m_T^W distribution in data. The entire range of the m_T^W -distribution is fitted. From the resulting QCD yield we extrapolate the QCD-contribution to the signal region ($m_T^W > 50$ GeV) by calculating the integral of the m_T^W -distribution from QCD normalized to the fit-result from $m_T^W = 50$ GeV to infinity. The method is first validated in 2J0T region which has significantly larger QCD events and then applied in the signal-enriched 2J1T region.

7.4.1 Estimation of the QCD-contribution in 2J0T

The 2J0T control region is dominated by QCD and W+light flavour jets events. The fraction of QCD events can be significantly increased by inverting the isolation criterion for the muon ($I_{rel} > 0.2$) in the 2J0T region. Figure 7.1 shows for the transverse mass of the W boson a comparison between different QCD templates in the isolated ($I_{rel} < 0.06$) and anti-isolated ($I_{rel} > 0.2$) region in 2J0T. The data-driven template is determined by subtracting the total non-QCD contribution from the data distribution, where the non-QCD contribution is obtained by adding up individual contributions from all other processes weighted to their respective cross section, except QCD. Good agreement in the shapes of the distributions from the two orthogonal regions is found. Using samples of simulated events from all relevant signal and background processes the QCD-purity of the anti-isolated side-band region has been estimated to be $\approx 93\%$. The fitted (and extrapolated) yields from m_T^W -fit with different QCD templates are summarized in table 7.4. Overall a good agreement between the different fits can be observed. The maximum difference among the final yields above $m_T^W > 50$ GeV is taken as uncertainty on the QCD normalization.

Table 7.4: QCD estimation in the 2J0T region

Variable	Fit Type	QCD template	Process	Fitted Yield (full range)	Extrapolated Yield ($m_T^W > 50$ GeV)
m_T^W	Binned	From Anti-Iso MC	QCD	708531±2401	136445±463
			EWK	5959478±14311	4045740±9715
			Top	730883±14058	491138±9447
		Data Driven	QCD	746583±2509	153283±515
			EWK	5935018±13680	4029140±9172
			Top	717227±13427	481953±8914
		From Iso MC	QCD	748866±2510	151681±508
			EWK	5865560±12267	3981980±8328
			Top	784402±12003	527102±8066

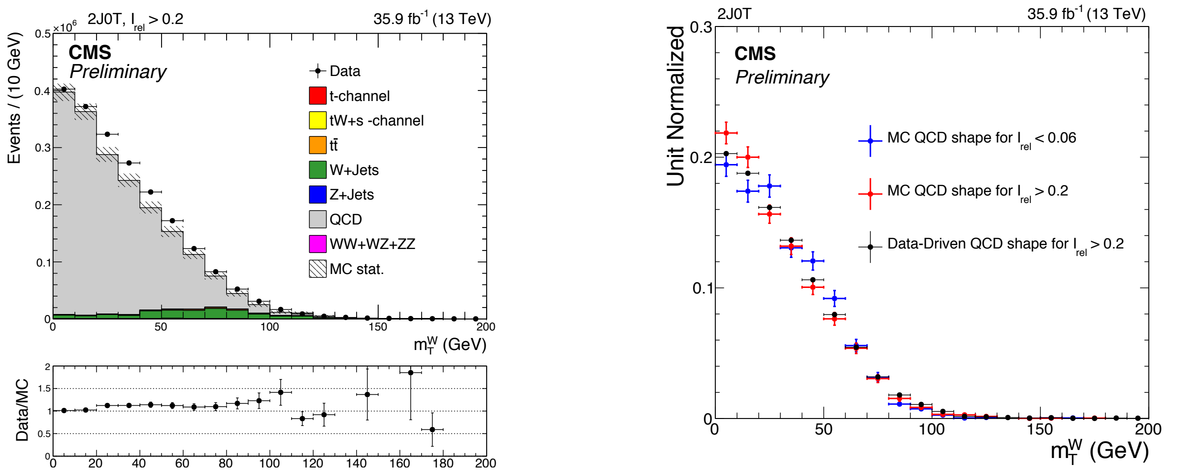
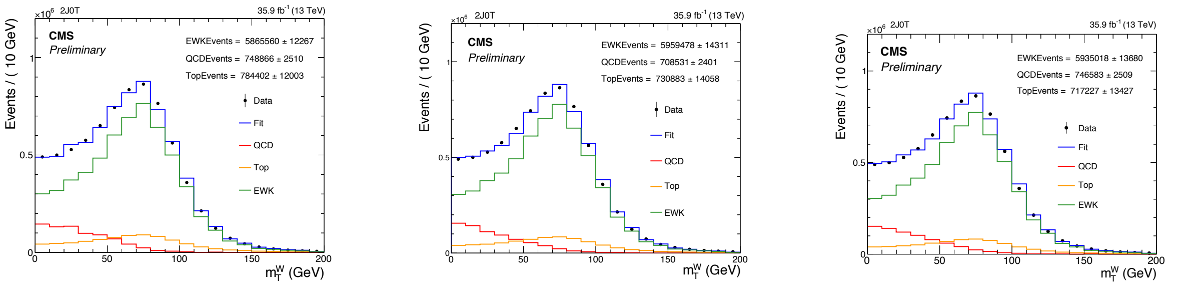


Figure 7.1: Data-MC comparison for m_T^W in the anti-isolated region (left) and Comparison of the QCD m_T^W templates between isolated and anti-isolated region(right) of 2J0T

7.4.2 Estimation of the QCD-contribution in 2J1T

Inclusive μ charge

The same binned likelihood fit as in the 2J0T control region is applied to the m_T^W -distribution in the 2J1T region. The templates $T(m_T^W)$ and $E(m_T^W)$ are again derived from MC simulation by summing up the individual contributions of different processes weighted with the corresponding cross sections, while the QCD-template is again derived from the anti-isolated side band in data ($I_{rel} > 0.2$). Figure 7.3 shows the data-MC



(a) MC($I_{rel} < 0.06$) (b) MC ($I_{rel} > 0.2$) (c) Data-driven ($I_{rel} > 0.2$)

Figure 7.2: Post fit distribution of m_T^W for binned fit with different QCD templates in 2J0T

comparison form m_T^W in the anti-isolated region as well as comparison of m_T^W -templates from isolated and anti-isolated region.

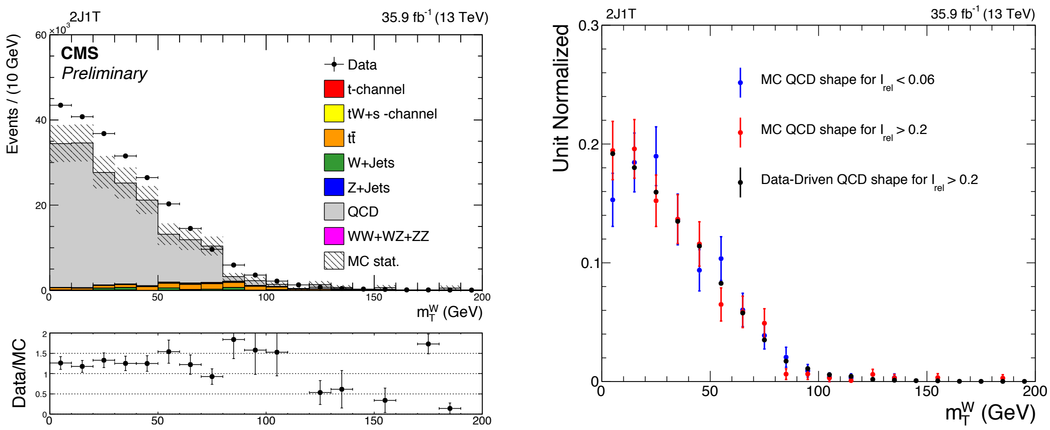


Figure 7.3: Data-MC comparison for m_T^W in the anti-isolated region (left) and Comparison of the QCD m_T^W -templates between isolated and anti-isolated region(right) of 2J1T

Again, the entire m_T^W -distribution is fitted and the QCD contribution in the signal region ($m_T^W > 50$ GeV) is extrapolated from the post-fit-m_T^W -distribution, normalized to the fit results. Figure 7.4 exhibits representative examples of the fitting process in the isolated region ($I_{rel} < 0.06$) where Table 7.5 reports the fitted QCD and Top and EWK yields.

Table 7.5: QCD estimation in the 2J1T region

Process	Fitted Yield (full m_T^W range)	Extrapolated Yield ($m_T^W > 50$ GeV)
QCD	72326±708	15869±155
EWK	53903±2897	37340±2010
Top	416962±2957	285399±2021

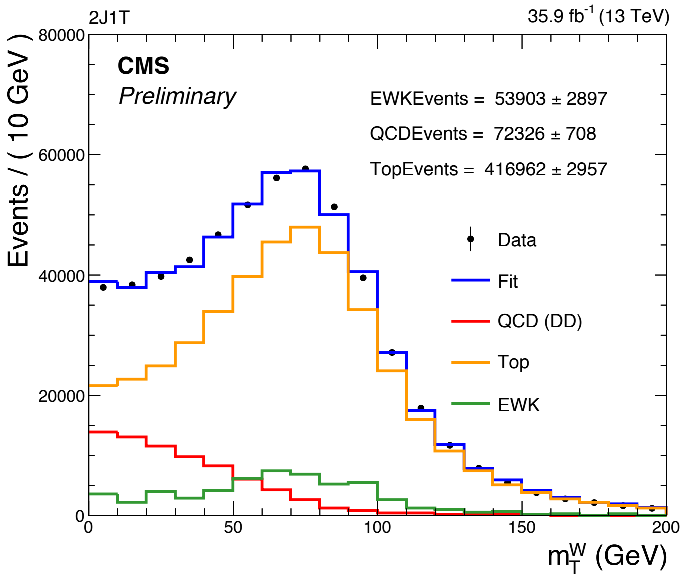


Figure 7.4: Fit in the isolated region ($I_{rel} < 0.06$) using data-driven QCD m_T^W template from anti-isolated ($I_{rel} > 0.2$) region

Separation on the μ charge

The same binned likelihood fit as previously is applied to the m_T^W distribution in the 2J1T region, but for μ^+ and μ^-, i.e. separately for positive and negative charge of the muon. This is necessary as the final mass measurement is separately performed for positive and negative muon charge. The templates for the Top and EWK process categories are again derived from MC simulation by summing up the individual contributions from the different processes weighted with the corresponding cross sections, while the QCD-template is again derived from the anti-isolated side band in data ($I_{rel} > 0.2$). The entire m_T^W -distribution is fitted and the QCD contribution in the signal region ($m_T^W > 50$ GeV) is extrapolated from the post-fit-m_T^W -distribution, normalized to the fit results. Table 7.6 reports the fitted QCD and non-QCD yields. The post-fit distributions are shown in Figure 7.5.

Table 7.6: QCD estimation in the 2J1T region separated by charge

	Process	Fitted Yield (full m_T^W range)	Extrapolated Yield ($m_T^W > 50$ GeV)
μ^+	QCD	35978±512	7823±111
	EWK	20834±1728	14102±1170
	Top	226409±1783	155208±1222
μ^-	QCD	36588±490	8099±108
	EWK	22526±1596	15986±1133
	Top	200855±1693	137234±1157

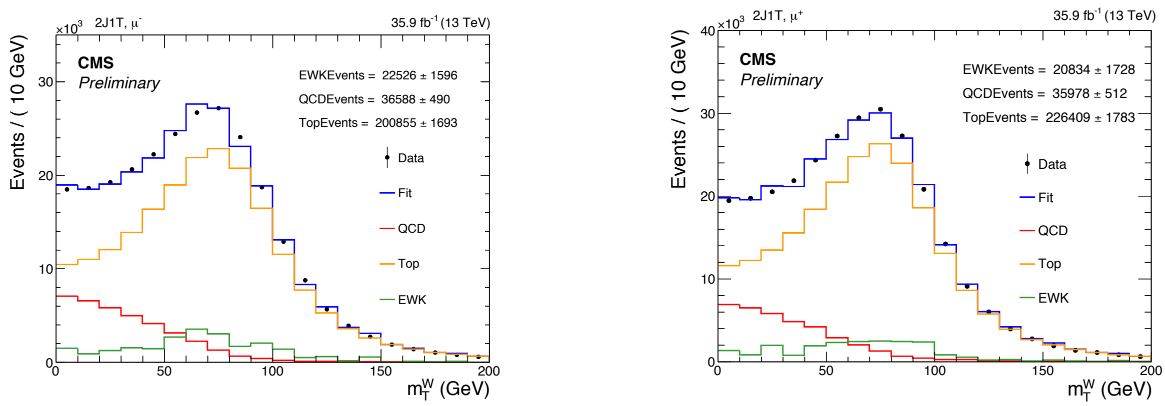


Figure 7.5: Fit in the isolated region ($I_{rel} < 0.06$), separately for μ^- (left) and μ^+ (right).

7.5 Multivariate Analysis

Several input variables as listed in Table 7.7 are combined into a multivariate (MVA) discriminator to have better separation between t-channel single top signal and backgrounds. The correlations among the input variables are also taken into account by the MVA discriminator. In this analysis, a boosted decision tree (BDT) is developed using the input variables listed in Table 7.7. TMVA package [123] built into ROOT has been used in the analysis.

Table 7.7: Input variables to BDT ranked according to their separation power

Rank	Variable	Description	Separation Power
1	$\Delta R_{bj'}$	ΔR between b-tagged jet and light-quark jet in η-ϕ space	2.090e-01
2	$\lvert\eta_{j'}\rvert$	Light-quark jet $\lvert\eta\rvert$	2.056e-01
3	$m_{bj'}$	Di-jet invariant mass	1.656e-01
4	$\cos\theta^*$	Cosine of the angle between lepton and light-quark in the rest frame of the top quark	6.769e-02
5	m_T^W	Transverse W boson mass	4.497e-02
6	$\Delta\eta_{\mu b}$	$\Delta\eta$ between muon and b-tagged jet	1.343e-02
7	Jet H_T	Scalar sum of the p_T -s of b-tagged jet and light-quark jet	7.352e-03
8	$\lvert\eta_\mu\rvert$	Muon $\lvert\eta\rvert$	3.584e-04

All input variables are validated by comparing data and MC distributions. The t-channel single top signal is trained against $t\bar{t}$ and EWK (V+jets and VV, V=W/Z) processes in 2J1T region with additionally QCD mixed in, after application of the m_T^W cut. All signal and background processes are weighted according to their purity in 2J1T region during training. The BDT training is checked against over-training by dividing the MC samples in 2J1T region into two independent sections, one for training the BDT and the other to evaluate similar performance. The result is shown in Figure 7.6 and no over-training is observed. The input variables are so chosen that the correlation between the reconstructed top mass and BDT response is as low as possible. This ensures that any cut on BDT output minimally affect the reconstructed top mass distribution. The distribution of the BDT output from data and MC in signal and control regions are shown in Figure 7.7.

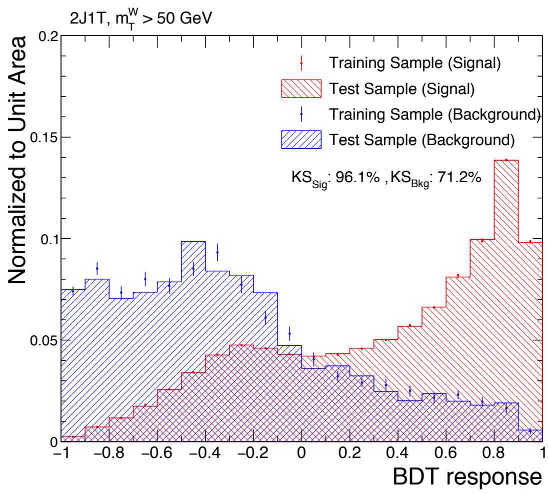


Figure 7.6: Overtraining check of BDT discriminator between test and training sample

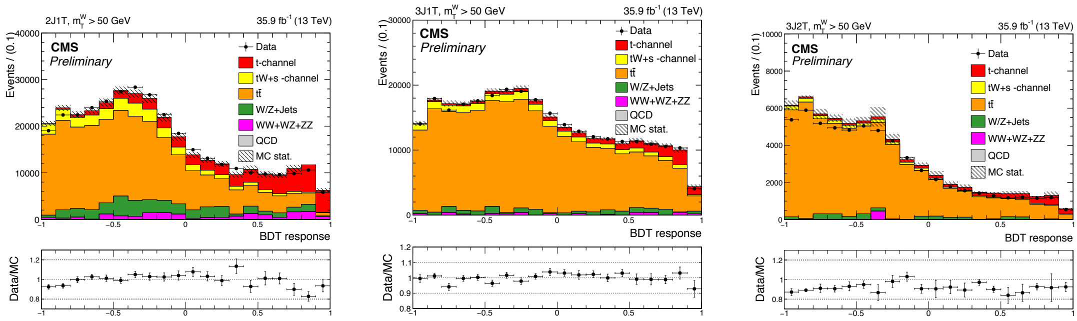


Figure 7.7: Data-MC comparison of BDT response in 2J1T (left), 3J1T (middle) and 3J2T (right) regions

A cut on the BDT response distribution is then applied in order to select a sample enriched in t-channel single top events. The cut on the BDT response is optimized by studying the signal and background efficiencies and signal purity of the resulting sample after cut, as shown in Figure 7.8. In order to make a good measurement, ideally, high signal purity with high signal statistics (efficiency) in the resulting sample after the cut is desirable. However, in reality, the signal efficiency and signal purity have opposite trends as shown in the left plot of Figure 7.8 with increasing cut value. So a trade-off between the two has to be met so that the resulting sample has high enough signal purity and enough statistics. Keeping this is mind, BDT response > 0.8 is applied so that the resulting sample has enough signal statistics, and more importantly, corresponds to 62.2 % signal purity and signal over background ratio of 1.65, for t-channel single top process. The resulting m_{top} distribution after BDT selection is shown in Figure 7.8.

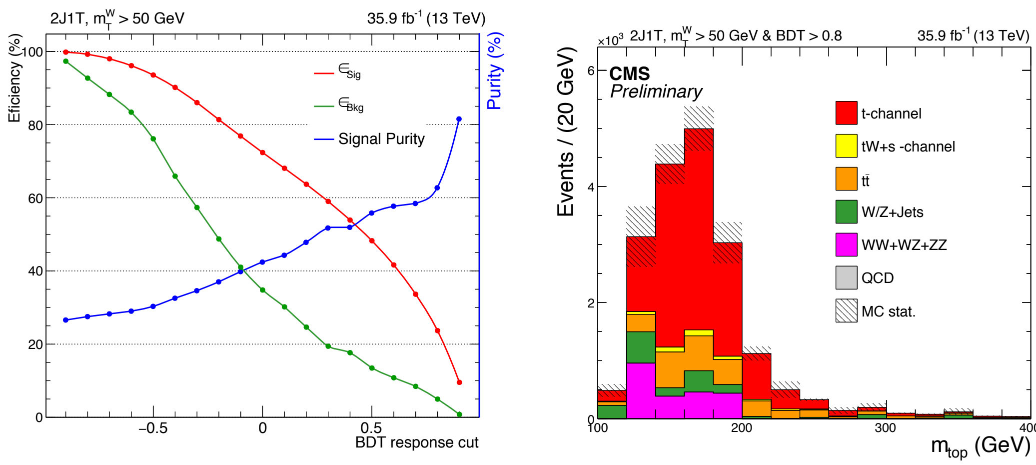


Figure 7.8: Study of signal and background efficiencies and signal purity as a function of cut on BDT response (left) and m_{top} distribution in 2J1T from MC after applying BDT response > 0.8 (right)

7.6 Top Mass extraction

Parametric templates for modeling signal and background component are used to model m_{top} distribution in MC. High skewness of the m_{top} distributions from signal and background processes and lower background rate after MVA selection poses considerable challenge to obtain parametric templates for all signal and background processes. Instead a way around is followed by considering the natural logarithm of the m_{top} . It is typically observed that for a distribution of a positive random variable such as m_{top} , which is skewed to the right, the skewness is reduced by taking logarithm of the random variable. This happens because taking logarithm "pulls in" more extreme values on the right relative to the mode of the original distribution, whereas the extreme values left of the mode are "stretched back" farther away from the mode, thus reducing the skewness. Moreover, logarithm being a monotonic function, transformed of probability distribution functions are well behaved.

Henceforth, $\ln m_{top}$ distributions from signal and background processes are used for parametric modeling. The top quark mass is determined by taking the exponential of the peak position of $\ln m_{top}$ distribution obtained from maximum likelihood fit. Figure 7.9 shows the distribution of $\ln m_{top}$ from signal and background MC in the 2J1T region after applying the BDT selection.

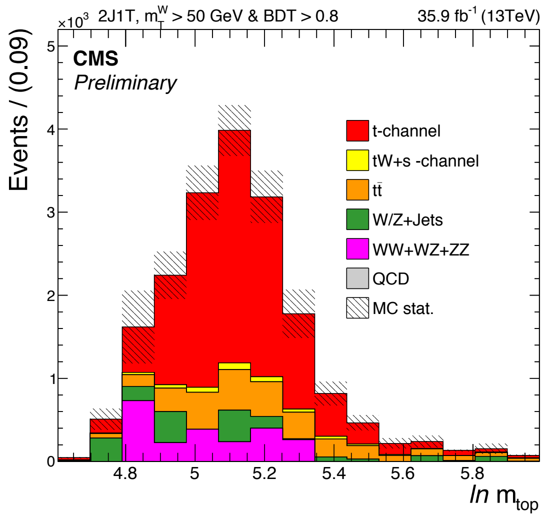


Figure 7.9: Distribution of $\ln m_{top}$ after BDT selection in 2J1T for signal and backgrounds from MC

The parametric 1D template that has been used model the $\ln m_{top}$ distribution in MC as shown in Figure 7.10 can be described as follows.

$$F(\ln m_{top}) = f_{t-ch}F_{t-ch} + f_{t\bar{t}}F_{t\bar{t}} + (1 - f_{t-ch} - f_{t\bar{t}})F_{EWK} \tag{7.4}$$

where F_{t-ch}, $F_{t\bar{t}}$ and F_{EWK} are modeled respectively as sum of an asymmetric gaussian and landau distribution with different mean, the crystal ball function and the Novosibirsk function. The peak position of the combined t-channel and $t\bar{t}$ templates termed as "mean" and fraction of t-channel (f_{t-ch}) and Top background category ($f_{t\bar{t}}$) yields are kept free to float during fit. Individual shape parameters of the templates comprising the model are tested on MC distributions and each individual template is found lie within relatively good accuracy. The shape parameters of the templates are kept fixed to the values obtained during best fit to the respective MC distribution.

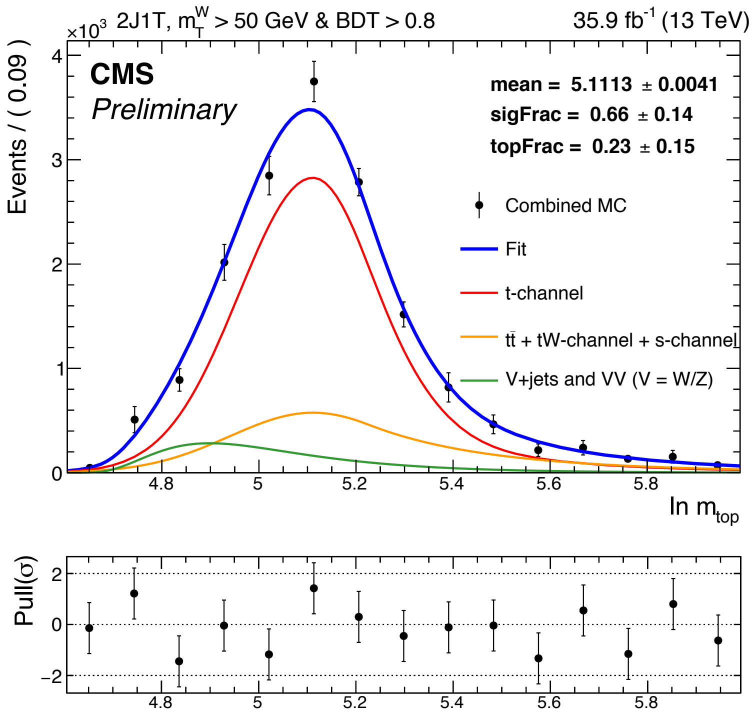


Figure 7.10: Parametric model along with its components for ln m_{top} distribution in MC.

The linearity of the fitter output is checked with different top mass hypothesis as shown in Figure 7.11 and the extracted mass is calibrated with respect to the true mass. The deviation of the slope (p1) of the red line shown in Figure 7.11 from unity can be attributed mostly due to the energy resolution effects of the b-tagged jet and missing energy. Very little dependence is observed for the yields of t-channel signal and $t\bar{t}$ background with different top mass hypothesis. The bias of fit parameters are checked using pseudo-experiments. The respective pull distributions are observed to follow a gaussian distribution centered at 0 with a standard deviation of 1 as shown in Figure 7.12. The expected statistical uncertainty on m_{top} is obtained to be ± 0.36 GeV from 10000 pseudo-experiments.

This analysis is still in progress in a blinded mode. Evaluation of the systematic uncertainties are being finalized. A final result is expected soon.

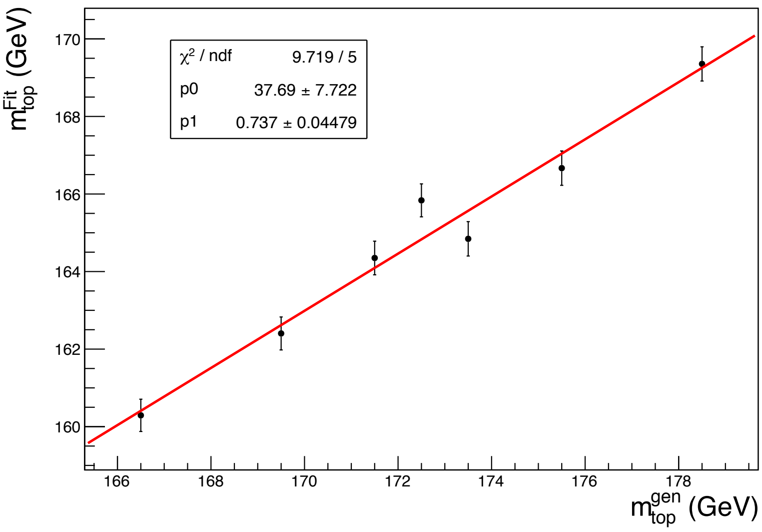
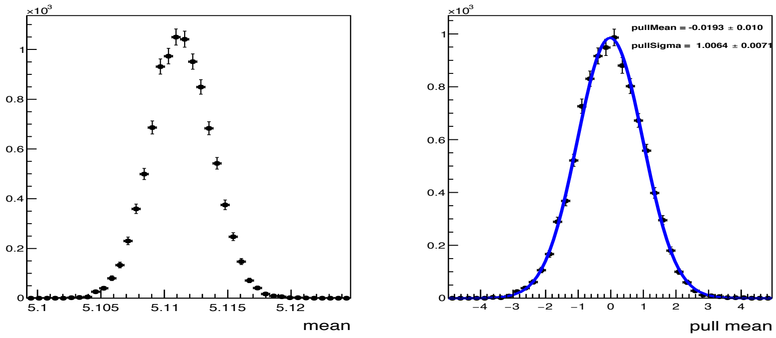
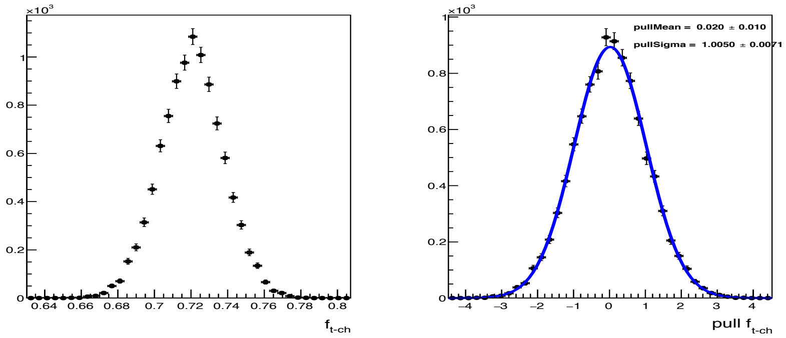


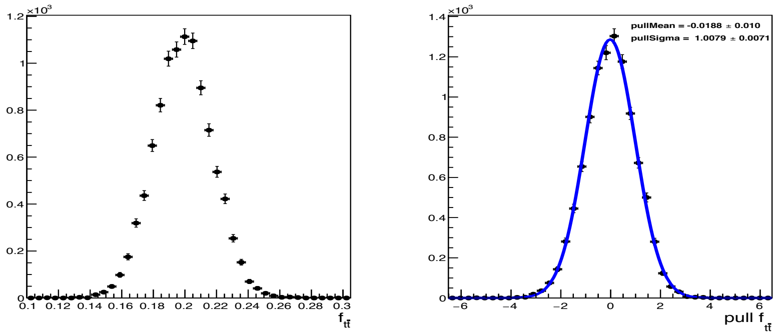
Figure 7.11: Linearity check of the m_{top} extracted from fit relative to the true m_{top} .



(a) Distribution of the post-fit “mean” and its pull in pseudo-experiments



(b) Distribution of the post-fit $f_{\mathrm{t-ch}}$ and its pull in pseudo-experiments



(c) Distribution of the post-fit $f_{\mathrm{t\bar{t}}}$ and its pull in pseudo-experiments

Figure 7.12: Post-fit distribution of fit parameters and their pull in pseudo-experiments

Chapter 8

Conclusions and Future Prospects

In this thesis, a detailed study of the single top production in the t-channel through charged current interaction, has been presented. The cross section of the t-channel single top production has been measured very precisely. In fact, the measured value of the t-channel single top production cross section with 2.2 fb^{-1} data is the most accurate result at CM energy of 13 TeV [20]. The ratio, $\mathrm{R_{t-ch}}$, of the top quark and anti-quark production has also been extracted directly from data. The CKM matrix element V_{tb} is determined from the ratio of measured cross section to theory cross section. The measured and estimated values are:

$$\sigma_{\mathrm{t-ch,t}} = 154 \pm 8(\text{stat.}) \pm 9(\text{expt.}) \pm 19(\text{theo.}) \pm 4(\text{lumi.})\ \text{pb} = 154 \pm 22\ \text{pb} \quad (8.1)$$

$$R_{\mathrm{t-ch.}} = 1.81 \pm 0.18(\text{stat.}) \pm 0.15(\text{syst.}) \quad (8.2)$$

$$\sigma_{\mathrm{t-ch,\bar{t}}} = 85 \pm 10(\text{stat.}) \pm 4(\text{expt.}) \pm 11(\text{theo.}) \pm 2(\text{lumi.})\ \text{pb} = 85 \pm 16\ \text{pb} \quad (8.3)$$

$$\sigma_{\mathrm{t-ch,t+\bar{t}}} = 238 \pm 13(\text{stat.}) \pm 12(\text{expt.}) \pm 26(\text{theo.}) \pm 5(\text{lumi.})\ \text{pb} = 238 \pm 32\ \text{pb} \quad (8.4)$$

$$|f_L^V V_{tb}| = 1.05 \pm 0.07(\text{expt.}) \pm 0.02(\text{theo.}) \quad (8.5)$$

It is very evident that the uncertainties owing to the sources related to modeling of signal and background processes are the dominant ones. Better understanding and validation of MC generators is required to overcome such issues.
In this chapter, we discuss impact of the results presented in this thesis in a bigger context by presenting results from ATLAS collaboration and status of LHC combinations with new results from Run-II and look for future prospects and scope for improvement.

8.1 ATLAS Result

The ATLAS collaboration has performed a measurement [131] of t-channel single top cross section at 13 TeV with 3.2 fb^{-1} in the μ+jets and e+jets final states inclusively. A neural network (NN) has been trained in the 2J1T region to discriminate t-channel single top signal from dominant backgrounds such as $\mathrm{t\bar{t}}$, W+jets etc. The $\mathrm{t\bar{t}}$ events are validated using 3J2T region and W+jets events are validated using a control region with less stringent b-tagging requirement. The cross section is extracted from a maximum likelihood fit to the NN discriminant separately for positively and negatively charged lepton in the final state to be

$$\sigma(tq) = 156 \pm 5(\text{stat.}) \pm 27(\text{syst.}) \pm 3(\text{lumi.})\,\text{pb} \tag{8.6}$$
$$\sigma(\bar{t}q) = 91 \pm 4(\text{stat.}) \pm 18(\text{syst.}) \pm 2(\text{lumi.})\,\text{pb} \tag{8.7}$$

These results can be combined to determine the total cross section for the t-channel single top production to be

$$\sigma(tq + \bar{t}q) = 247 \pm 6(\text{stat.}) \pm 45(\text{syst.}) \pm 5(\text{lumi.})\,\text{pb} = 247 \pm 46\,\text{pb} \tag{8.8}$$

The ratio of production cross section of top quark and anti-quark, $\mathrm{R_{t-ch}}$ is given by

$$\mathrm{R_{t-ch}} = 1.72 \pm 0.09(\text{stat.}) \pm 0.18(\text{syst.}) \tag{8.9}$$

the measured $\mathrm{R_{t-ch}}$ is then compared with several PDF predictions and found compatible with each of them within uncertainty. The CKM matrix element V_{tb} is extracted from the total cross section given in Eq.(8.8) to be

$$f_L^V|V_{tb}| = 1.07 \pm 0.01(\text{stat.}) \pm 0.09(\text{syst.}) \pm 0.02(\text{theor.}) \pm 0.01(\text{lumi.}) \tag{8.10}$$

where f_L^V is the anomalous form factor taking into account the possible anomalous tWb coupling as in Eq.(6.5).
The dominant uncertainty sources to the measurements are parton shower modelling, b-tagging efficiency, jet energy resolution etc.

LHC Combination

Individual measurements of single top processes from CMS and ATLAS collaborations are compared and then combined by the LHC top working group. Together with TeVatron combinations these provide global picture regarding the precision of measurements achieved in the top quark sector. The latest summary of single top cross section measurements and corresponding extracted V_{tb} by CMS and ATLAS till today are shown in Figure 8.1 and Figure 8.2.

8.2 Prospect for single top measurements

Comparing with the ATLAS result, the measurement of total t-channel single top cross section by CMS presented in Chapter 6 of this thesis, has already achieved slightly more accuracy even with $\sim$1 fb^{-1} less data. The installation of a new layer for the inner pixel detector in CMS tracking system at the end of 2016, would improve the secondary vertex finding and thus heavy flavour jet identification capabilities in CMS. This is expected to reduce the uncertainties mainly due to the flavour JES and b-tagging etc. and therefore increase the scope for a measurement with high precision in the single top processes. Also the possibility of collecting a total of $\sim$ 100 fb^{-1} data during LHC Run-II, the statistical uncertainties will become lower and would also enrich our understanding of detector response, which in turn would reduce systematic effects. A precise measurement of $\mathrm{R_{t-ch}}$ in near future would not only provide an additional handle to constrain u/d-quark PDFs, but also extend the current knowledge of b-quark PDFs inside a proton. A precise measurement of V_{tb} with more data and improved detector would also help to reduce the

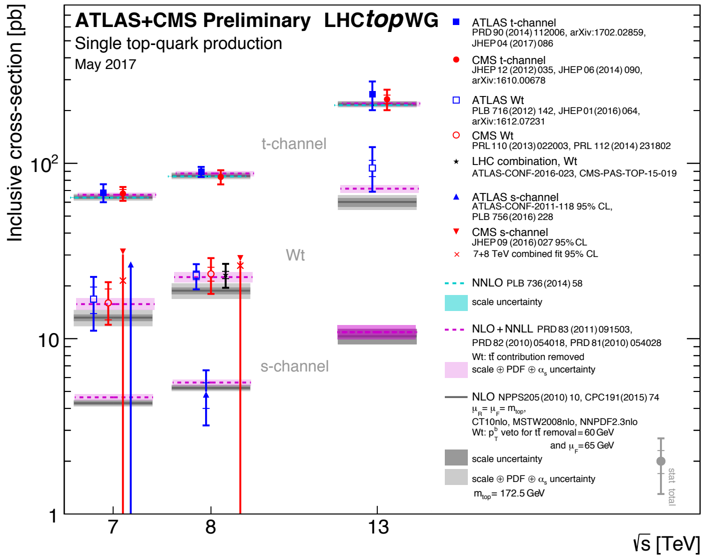


Figure 8.1: Summary single top cross section measurements at LHC [61]

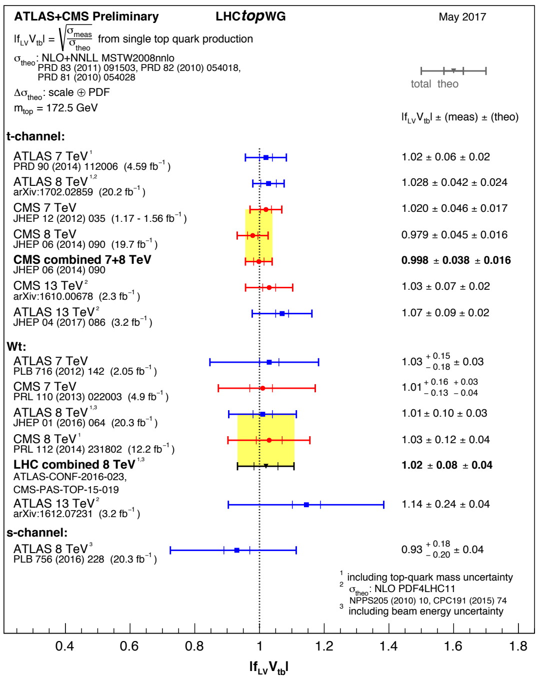


Figure 8.2: V_{tb} measurements in single top processes at LHC [61]

available phase space for anomalous couplings at the tWb vertex in near future. The addition of electron final states will also improve the overall precision.

The ongoing analysis to measure the top quark mass has the potential to surpass previous measurements using single top events in accuracy. Already the expected statistical uncertainty on the top mass is estimated to be ± 0.36 GeV which is lower compared to previous CMS measurement at 8 TeV (± 0.77 GeV) as discussed in Chapter 7. Regarding the systematic uncertainties, the following "guess-estimates" are made for some of the dominant sources.

- JES : ± 0.6 GeV
- b-jet modelling : ± 0.15 GeV
- Fit Calibration: ± 0.2 GeV
- Signal modelling: ±0.25 GeV
- Background modelling: ±0.35 GeV
- Underlying events: ±0.2 GeV

Based on these estimates, the total guess-estimated uncertainty δm_{top} on the measured top mass in single top events is given by $\delta m_{top} \sim \pm$ 0.9 GeV (stat. + syst.). However, these estimates of uncertainty are not final as the analysis is still in progress.
This techniques developed for this analysis can be further extended to directly measure the top decay width from the reconstructed top quark mass spectrum in data using single top events and thus constraining V_{tb} indirectly.

Bibliography

[1] J. A. Aguilar-Saavedra, "A Minimal set of top anomalous couplings," *Nucl. Phys.*, vol. B812, pp. 181–204, 2009.

[2] T. M. P. Tait and C. P. Yuan, "Single top quark production as a window to physics beyond the standard model," *Phys. Rev.*, vol. D63, p. 014018, 2000.

[3] M. Hashemi and G. Haghighat, "Observability of s-channel Heavy Charged Higgs at LHC Using Top Tagging Technique," *JHEP*, vol. 02, p. 040, 2016.

[4] A. Giammanco, "Single top quark production at the LHC," *Reviews in Physics*, vol. 1, pp. 1 – 12, 2016.

[5] M. Aliev, H. Lacker, U. Langenfeld, S. Moch, P. Uwer, and M. Wiedermann, "HATHOR: HAdronic Top and Heavy quarks crOss section calculatoR," *Comput. Phys. Commun.*, vol. 182, pp. 1034–1046, 2011.

[6] P. Kant, O. M. Kind, T. Kintscher, T. Lohse, T. Martini, S. Mlbitz, P. Rieck, and P. Uwer, "HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions," *Comput. Phys. Commun.*, vol. 191, pp. 74–89, 2015.

[7] S. Alekhin, A. Djouadi, and S. Moch, "The top quark and Higgs boson masses and the stability of the electroweak vacuum," *Phys. Lett.*, vol. B716, pp. 214–219, 2012.

[8] "First combination of Tevatron and LHC measurements of the top-quark mass," 2014.

[9] S. Chatrchyan *et al.*, "The CMS experiment at the CERN LHC," *JINST*, vol. 3, p. S08004, 2008.

[10] A. M. Sirunyan *et al.*, "Particle-flow reconstruction and global event description with the CMS detector," *JINST*, vol. 12, no. 10, p. P10003, 2017.

[11] V. Khachatryan *et al.*, "Performance of Photon Reconstruction and Identification with the CMS Detector in Proton-Proton Collisions at sqrt(s) = 8 TeV," *JINST*, vol. 10, no. 08, p. P08010, 2015.

[12] M. Cacciari, G. P. Salam, and G. Soyez, "The Anti-k(t) jet clustering algorithm," *JHEP*, vol. 04, p. 063, 2008.

[13] S. Chatrchyan *et al.*, "Determination of Jet Energy Calibration and Transverse Momentum Resolution in CMS," *JINST*, vol. 6, p. P11002, 2011.

[14] "Muon Reconstruction and Identification Improvements for Run-2 and First Results with 2015 Run Data," Tech. Rep. CMS-DP-2015-015, CERN, Geneva, Jul 2015.

[15] "Performance of the Particle-Flow jet identification criteria using proton-proton collisions at 13 TeV with the 2015 dataset," Tech. Rep. CMS-AN-15-269, CERN, Geneva.

[16] C. Patrignani *et al.*, "Review of Particle Physics," *Chin. Phys.*, vol. C40, no. 10, p. 100001, 2016.

[17] T. Aaltonen *et al.*, "First Observation of Electroweak Single Top Quark Production," *Phys. Rev. Lett.*, vol. 103, p. 092002, 2009.

[18] V. M. Abazov *et al.*, "Observation of Single Top Quark Production," *Phys. Rev. Lett.*, vol. 103, p. 092001, 2009.

[19] "Measurement of the t-channel single top-quark cross section at 13 TeV," Tech. Rep. CMS-PAS-TOP-15-004, CERN, Geneva, 2015.

[20] A. M. Sirunyan *et al.*, "Cross section measurement of t-channel single top quark production in pp collisions at $\sqrt{s} = 13$ TeV," *Phys. Lett.*, vol. B772, pp. 752–776, 2017.

[21] R. Barlow and C. Beeston, "Fitting using finite monte carlo samples," *Comput. Phys. Commun.*, vol. 77, no. 2, pp. 219 – 228, 1993.

[22] S. Dulat, T.-J. Hou, J. Gao, M. Guzzi, J. Huston, P. Nadolsky, J. Pumplin, C. Schmidt, D. Stump, and C. P. Yuan, "New parton distribution functions from a global analysis of quantum chromodynamics," *Phys. Rev.*, vol. D93, no. 3, p. 033006, 2016.

[23] S. Alekhin, J. Blumlein, and S. Moch, "The ABM parton distributions tuned to LHC data," *Phys. Rev.*, vol. D89, no. 5, p. 054028, 2014.

[24] L. A. Harland-Lang, A. D. Martin, P. Motylinski, and R. S. Thorne, "Parton distributions in the LHC era: MMHT 2014 PDFs," *Eur. Phys. J.*, vol. C75, no. 5, p. 204, 2015.

[25] F. D. Aaron *et al.*, "Combined Measurement and QCD Analysis of the Inclusive $e^{\pm}p$ Scattering Cross Sections at HERA," *JHEP*, vol. 01, p. 109, 2010.

[26] R. D. Ball, V. Bertone, F. Cerutti, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, J. Rojo, and M. Ubiali, "Reweighting and Unweighting of Parton Distributions and the LHC W lepton asymmetry data," *Nucl. Phys.*, vol. B855, pp. 608–638, 2012.

[27] S. Chatrchyan *et al.*, "Measurement of the single-top-quark t-channel cross section in pp collisions at $\sqrt{s} = 7$ TeV," *JHEP*, vol. 12, p. 035, 2012.

[28] V. Khachatryan *et al.*, "Measurement of the t-channel single-top-quark production cross section and of the $\mid V_{tb} \mid$ CKM matrix element in pp collisions at $\sqrt{s}$= 8 TeV," *JHEP*, vol. 06, p. 090, 2014.

[29] T. A. Aaltonen *et al.*, "Tevatron Combination of Single-Top-Quark Cross Sections and Determination of the Magnitude of the Cabibbo-Kobayashi-Maskawa Matrix Element $\mathbf{V_{tb}}$," *Phys. Rev. Lett.*, vol. 115, no. 15, p. 152003, 2015.

[30] N. Kidonakis, "Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production," *Phys. Rev.*, vol. D83, p. 091503, 2011.

[31] S. Argyropoulos and T. Sjstrand, "Effects of color reconnection on $t\bar{t}$ final states at the LHC," *JHEP*, vol. 11, p. 043, 2014.

[32] "Measurement of the top-quark mass in t t-bar events with lepton+jets final states in pp collisions at sqrt(s)=8 TeV," Tech. Rep. CMS-PAS-TOP-14-001, CERN, Geneva, 2014.

[33] `https://sites.google.com/site/csetstudyguidechemistry/home/1-2c---atomic-structure`.

[34] S. L. Glashow, "Partial Symmetries of Weak Interactions," *Nucl. Phys.*, vol. 22, pp. 579–588, 1961.

[35] S. Weinberg, "A Model of Leptons," *Phys. Rev. Lett.*, vol. 19, pp. 1264–1266, 1967.

[36] A. Salam, "Weak and Electromagnetic Interactions," *Conf. Proc.*, vol. C680519, pp. 367–377, 1968.

[37] `https://www.quantumdiaries.org/2014/03/14/the-standard-model-a-beautiful-but-fla`

[38] E. Noether, "Invariant Variation Problems," *Gott. Nachr.*, vol. 1918, pp. 235–257, 1918. [Transp. Theory Statist. Phys.1,186(1971)].

[39] C.-N. Yang and R. L. Mills, "Conservation of Isotopic Spin and Isotopic Gauge Invariance," *Phys. Rev.*, vol. 96, pp. 191–195, 1954.

[40] F. Englert and R. Brout, "Broken Symmetry and the Mass of Gauge Vector Mesons," *Phys. Rev. Lett.*, vol. 13, pp. 321–323, 1964.

[41] P. W. Higgs, "Broken symmetries, massless particles and gauge fields," *Phys. Lett.*, vol. 12, pp. 132–133, 1964.

[42] S. Chatrchyan *et al.*, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC," *Phys. Lett.*, vol. B716, pp. 30–61, 2012.

[43] 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," *Phys. Lett.*, vol. B716, pp. 1–29, 2012.

[44] W. de Boer, "The Discovery of the Higgs Boson with the CMS Detector and its Implications for Supersymmetry and Cosmology," in *Time and Matter 2013 (TAM2013) Venice, Italy*, 2013.

[45] N. Cabibbo, "Unitary symmetry and leptonic decays," *Phys. Rev. Lett.*, vol. 10, pp. 531–533, Jun 1963.

[46] M. Kobayashi and T. Maskawa, “CP Violation in the Renormalizable Theory of Weak Interaction,” *Prog. Theor. Phys.*, vol. 49, pp. 652–657, 1973.

[47] L. Wolfenstein, “Parametrization of the Kobayashi-Maskawa Matrix,” *Phys. Rev. Lett.*, vol. 51, p. 1945, 1983.

[48] J. Charles, A. Hocker, H. Lacker, S. Laplace, F. R. Le Diberder, J. Malcles, J. Ocariz, M. Pivk, and L. Roos, “CP violation and the CKM matrix: Assessing the impact of the asymmetric B factories,” *Eur. Phys. J.*, vol. C41, no. 1, pp. 1–131, 2005.

[49] J. C. Collins, D. E. Soper, and G. F. Sterman, “Factorization of Hard Processes in QCD,” *Adv. Ser. Direct. High Energy Phys.*, vol. 5, pp. 1–91, 1989.

[50] T. Han, “Collider phenomenology: Basic knowledge and techniques,” pp. 407–454, 2005.

[51] V. N. Gribov and L. N. Lipatov, “Deep inelastic e p scattering in perturbation theory,” *Sov. J. Nucl. Phys.*, vol. 15, pp. 438–450, 1972. [Yad. Fiz.15,781(1972)].

[52] G. Altarelli and G. Parisi, “Asymptotic Freedom in Parton Language,” *Nucl. Phys.*, vol. B126, pp. 298–318, 1977.

[53] Y. L. Dokshitzer, “Calculation of the Structure Functions for Deep Inelastic Scattering and e+ e- Annihilation by Perturbation Theory in Quantum Chromodynamics.,” *Sov. Phys. JETP*, vol. 46, pp. 641–653, 1977. [Zh. Eksp. Teor. Fiz.73,1216(1977)].

[54] H.-L. Lai, M. Guzzi, J. Huston, Z. Li, P. M. Nadolsky, J. Pumplin, and C. P. Yuan, “New parton distributions for collider physics,” *Phys. Rev.*, vol. D82, p. 074024, 2010.

[55] M. Gluck, P. Jimenez-Delgado, E. Reya, and C. Schuck, “On the role of heavy flavor parton distributions at high energy colliders,” *Phys. Lett.*, vol. B664, pp. 133–138, 2008.

[56] A. D. Martin, W. J. Stirling, R. S. Thorne, and G. Watt, “Parton distributions for the LHC,” *Eur. Phys. J.*, vol. C63, pp. 189–285, 2009.

[57] https://www-d0.fnal.gov/Run2Physics/top/top_public_web_pages/top_feynman_diagrams.html.

[58] https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO.

[59] “Observation of top quark production in $\overline{p}p$ collisions with the collider detector at fermilab,” *Phys. Rev. Lett.*, vol. 74, pp. 2626–2631, Apr 1995.

[60] “Observation of the top quark,” *Phys. Rev. Lett.*, vol. 74, pp. 2632–2637, Apr 1995.

[61] https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots.

[62] https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec.

[63] "Tevatron combination of single-top-quark cross sections and determination of the magnitude of the cabibbo-kobayashi-maskawa matrix element V_{tb}," *Phys. Rev. Lett.*, vol. 115, p. 152003, Oct 2015.

[64] S. L. Glashow, J. Iliopoulos, and L. Maiani, "Weak Interactions with Lepton-Hadron Symmetry," *Phys. Rev.*, vol. D2, pp. 1285–1292, 1970.

[65] W. Bernreuther, "Top quark physics at the LHC," *J. Phys.*, vol. G35, p. 083001, 2008.

[66] J. Z. Bai *et al.*, "Measurements of the cross-section for $e^+ e^- \to$ Hadrons at center-of-mass energies from 2-GeV to 5-GeV," *Phys. Rev. Lett.*, vol. 88, p. 101802, 2002.

[67] E. Sanchez, "Muon Lifetime Measurement with the FAST Detector at PSI," *Physics Procedia*, vol. 17, pp. 168 – 174, 2011. 2nd International Workshop on the Physics of fundamental Symmetries and Interactions - PSI2010.

[68] S. Schael *et al.*, "Precision electroweak measurements on the Z resonance," *Phys. Rept.*, vol. 427, pp. 257–454, 2006.

[69] J. R. Incandela, A. Quadt, W. Wagner, and D. Wicke, "Status and Prospects of Top-Quark Physics," *Prog. Part. Nucl. Phys.*, vol. 63, pp. 239–292, 2009.

[70] M. Baak, J. Cth, J. Haller, A. Hoecker, R. Kogler, K. Mnig, M. Schott, and J. Stelzer, "The global electroweak fit at NNLO and prospects for the LHC and ILC," *Eur. Phys. J.*, vol. C74, p. 3046, 2014.

[71] G. Degrassi, S. Di Vita, J. Elias-Miro, J. R. Espinosa, G. F. Giudice, G. Isidori, and A. Strumia, "Higgs mass and vacuum stability in the Standard Model at NNLO," *JHEP*, vol. 08, p. 098, 2012.

[72] F. Olness and R. Scalise, "Regularization, Renormalization, and Dimensional Analysis: Dimensional Regularization meets Freshman E&M," *Am. J. Phys.*, vol. 79, p. 306, 2011.

[73] M. Beneke, "A Quark mass definition adequate for threshold problems," *Phys. Lett.*, vol. B434, pp. 115–125, 1998.

[74] A. H. Hoang, Z. Ligeti, and A. V. Manohar, "B decays in the upsilon expansion," *Phys. Rev.*, vol. D59, p. 074017, 1999.

[75] A. H. Hoang and T. Teubner, "Top quark pair production close to threshold: Top mass, width and momentum distribution," *Phys. Rev.*, vol. D60, p. 114027, 1999.

[76] A. H. Hoang and I. W. Stewart, "Top Mass Measurements from Jets and the Tevatron Top-Quark Mass," *Nucl. Phys. Proc. Suppl.*, vol. 185, pp. 220–226, 2008.

[77] `http://www.hep.ph.ic.ac.uk/~wstirlin/plots/plots.html`.

[78] ATLAS Collaboration, "The ATLAS Experiment at the CERN Large Hadron Collider," *Journal of Instrumentation*, vol. 3, no. 08, p. S08003, 2008.

[79] LHCb Collaboration, "The LHCb Detector at the LHC," *Journal of Instrumentation*, vol. 3, no. 08, p. S08005, 2008.

[80] ALICE Collaboration, "The ALICE experiment at the CERN LHC," *Journal of Instrumentation*, vol. 3, no. 08, p. S08002, 2008.

[81] http://ts-dep.web.cern.ch/ts-dep/groups/lea/int/workshops/.

[82] L. Evans and P. Bryant, "LHC Machine," *Journal of Instrumentation*, vol. 3, no. 08, p. S08001, 2008.

[83] https://cds.cern.ch/record/2197559.

[84] https://cds.cern.ch/record/841539.

[85] http://cds.cern.ch/record/841485.

[86] S. van der Meer, "Calibration of the effective beam height in the ISR," Tech. Rep. CERN-ISR-PO-68-31. ISR-PO-68-31, CERN, Geneva, 1968.

[87] "CMS luminosity measurement for the 2015 data-taking period," Tech. Rep. CMS-PAS-LUM-15-001, CERN, Geneva, 2017.

[88] https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults.

[89] http://cms.web.cern.ch/news/cms-detector-design.

[90] J. Nielsen, "Fundamentals of LHC Experiments," in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 2010). String Theory and Its Applications: From meV to the Planck Scale: Boulder, Colorado, USA, June 1-25, 2010*, pp. 127–152, 2011.

[91] http://newsline.linearcollider.org/2011/05/05/one-hundred-years-of-superconductivity/cms-solenoid-magnet/.

[92] S. Chatrchyan *et al.*, "Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays," *JINST*, vol. 5, p. T03021, 2010.

[93] V. Karimki, M. Mannelli, P. Siegrist, H. Breuker, A. Caner, R. Castaldi, K. Freudenreich, G. Hall, R. Horisberger, M. Huhtinen, and A. Cattai, *The CMS tracker system project: Technical Design Report*. Technical Design Report CMS, Geneva: CERN, 1997.

[94] *The CMS tracker: addendum to the Technical Design Report*. Technical Design Report CMS, Geneva: CERN, 2000.

[95] L. Viliani, "CMS tracker performance and readiness for LHC Run II," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 824, pp. 67 – 69, 2016. Frontier Detectors for Frontier Physics: Proceedings of the 13th Pisa Meeting on Advanced Detectors.

[96] S. Chatrchyan *et al.*, "Commissioning and Performance of the CMS Pixel Tracker with Cosmic Ray Muons," *JINST*, vol. 5, p. T03007, 2010.

[97] http://accms04.physik.rwth-aachen.de/~schael/Stefan%20Schael,%20RWTH%20Aachen/CMS%20Tracker.html.

[98] http://hep.fi.infn.it/CMS/software/ResultsWebPage/CMSTrackerGeometry.html.

[99] S. Chatrchyan *et al.*, "Description and performance of track and primary-vertex reconstruction with the CMS tracker," *JINST*, vol. 9, no. 10, p. P10009, 2014.

[100] https://agenda.infn.it/getFile.py/access?contribId=35&sessionId=26&resId=0&materialId=slides&confId=352.

[101] "Performance and operation of the CMS electromagnetic calorimeter," *Journal of Instrumentation*, vol. 5, no. 03, p. T03010, 2010.

[102] *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS, Geneva: CERN, 1997. The following files are from ¡a href=.

[103] http://xdaq.web.cern.ch/xdaq/setup/images/HCAL.png.

[104] "Performance of the cms hadron calorimeter with cosmic ray muons and lhc beam data," *Journal of Instrumentation*, vol. 5, no. 03, p. T03012, 2010.

[105] G. Abbiendi, "The CMS muon system in Run2: preparation, status and first results," *PoS*, vol. EPS-HEP2015, p. 237, 2015.

[106] T. Lenzi, "Development and Study of Different Muon Track Reconstruction Algorithms for the Level-1 Trigger for the CMS Muon Upgrade with GEM Detectors," Master's thesis, U. Brussels (main), 2013.

[107] S. Chatrchyan *et al.*, "Performance of CMS muon reconstruction in pp collision events at $\sqrt{s} = 7$ TeV," *JINST*, vol. 7, p. P10002, 2012.

[108] V. Khachatryan *et al.*, "Measurements of Inclusive W and Z Cross Sections in pp Collisions at $\sqrt{s} = 7$ TeV," *JHEP*, vol. 01, p. 080, 2011.

[109] S. Cittolin, A. Rcz, and P. Sphicas, *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project*. Technical Design Report CMS, Geneva: CERN, 2002.

[110] G. Petrucciani, A. Rizzi, and C. Vuosalo, "Mini-AOD: A New Analysis Data Format for CMS," *J. Phys. Conf. Ser.*, vol. 664, no. 7, p. 072052, 2015.

[111] E. Chabanat and N. Estre, "Deterministic annealing for vertex finding at CMS," in *Computing in high energy physics and nuclear physics. Proceedings, Conference, CHEP'04, Interlaken, Switzerland, September 27-October 1, 2004*, pp. 287–290, 2005.

[112] G. Abbiendi *et al.*, "Muon Reconstruction in the CMS Detector," Tech. Rep. CMS-NOTE-2008-097, CERN, Geneva, July 2009.

[113] https://indico.cern.ch/event/190899/contributions/343838/attachments/271380/379779/BachtisMuonPOGDeltaBeta.pdf.

[114] S. Baffioni, C. Charlot, F. Ferri, D. Futyan, P. Meridiani, I. Puljak, C. Rovelli, R. Salerno, and Y. Sirois, "Electron reconstruction in CMS," Tech. Rep. CMS-NOTE-2006-040, CERN, Geneva, Feb 2006.

[115] W. Adam, R. Frhwirth, 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*, vol. 31, no. 9, p. N9, 2005.

[116] "Identification of b quark jets at the CMS Experiment in the LHC Run 2," tech. rep., CERN, 2016.

[117] "Heavy flavor identification at CMS with deep neural networks," Mar 2017.

[118] "First look at top-quark selections with 13 TeV data," Jul 2015.

[119] https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec.

[120] N. Kidonakis, "Differential and total cross sections for top pair and single top production," in *Proceedings, 20th International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS 2012): Bonn, Germany, March 26-30, 2012*, pp. 831–834, 2012. [,831(2012)].

[121] M. Czakon and A. Mitov, "Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders," *Comput. Phys. Commun.*, vol. 185, p. 2930, 2014.

[122] G. J. Feldman and R. D. Cousins, "A Unified approach to the classical statistical analysis of small signals," *Phys. Rev.*, vol. D57, pp. 3873–3889, 1998.

[123] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, "TMVA: Toolkit for Multivariate Data Analysis," *PoS*, vol. ACAT, p. 040, 2007.

[124] R. Barlow, "Asymmetric systematic errors," 2003.

[125] https://indico.in2p3.fr/event/9742/contributions/50419/attachments/40828/50594/sos2014_systprof_v38.pdf.

[126] M. Bahr *et al.*, "Herwig++ Physics and Manual," *Eur. Phys. J.*, vol. C58, pp. 639–707, 2008.

[127] V. Khachatryan *et al.*, "Measurement of the differential cross section for top quark pair production in pp collisions at $\sqrt{s} = 8\,\mathrm{TeV}$," *Eur. Phys. J.*, vol. C75, no. 11, p. 542, 2015.

[128] M. Botje *et al.*, "The PDF4LHC Working Group Interim Recommendations," 2011.

[129] A. M. Sirunyan *et al.*, "Measurement of the top quark mass using single top quark events in proton-proton collisions at $\sqrt{s} = 8$ TeV," *Eur. Phys. J.*, vol. C77, no. 5, p. 354, 2017.

[130] "Measurement of the top quark mass in topologies enhanced with single top-quarks produced in the t-channel in $\sqrt{s} = 8\,\mathrm{TeV}$ ATLAS data," Tech. Rep. ATLAS-CONF-2014-055, CERN, Geneva, Sep 2014.

[131] M. Aaboud *et al.*, "Measurement of the inclusive cross-sections of single top-quark and top-antiquark t-channel production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector," *JHEP*, vol. 04, p. 086, 2017.