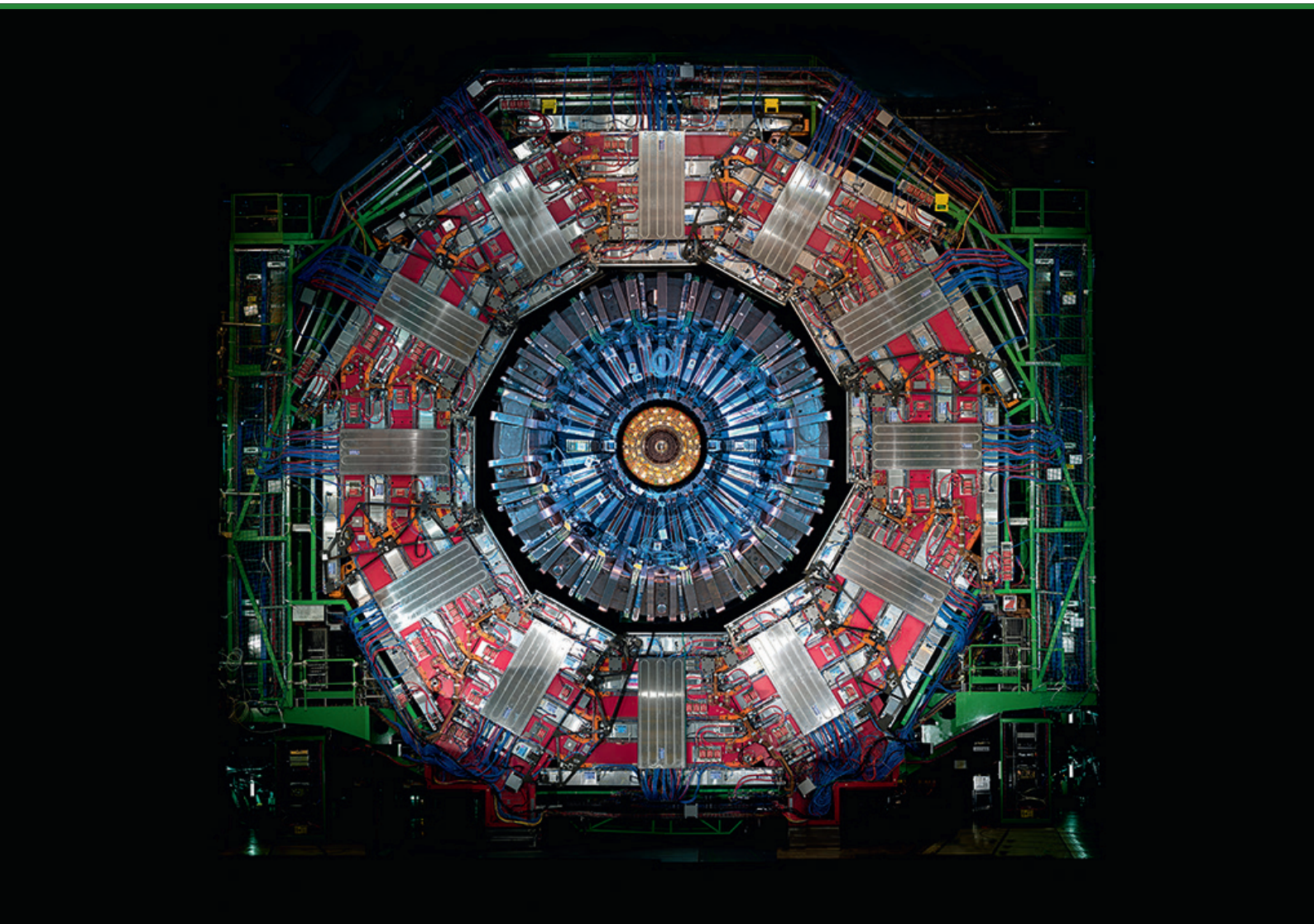
RĪGAS TEHNISKĀ UNIVERSITĀTE

Viesturs Veckalns

STUDIES OF COLOUR FLOW IN TOP QUARK PAIR DECAYS AT 13 TEV AT THE CMS EXPERIMENT OF THE CERN LHC

Summary of Doctoral Thesis



RTU Izdevniecība
Rīga 2019

RĪGAS TEHNISKĀ UNIVERSITĀTE
Faculty of Power and Electrical Engineering
Institute of Industrial Electronics and Electrical Engineering

Viesturs Veckalns
Doctoral Student of the Study Programme "Computerised Control of Electrical Engineering"

STUDIES OF COLOUR FLOW IN TOP QUARK PAIR DECAYS AT 13 TEV AT THE CMS EXPERIMENT OF THE CERN LHC

Summary of Doctoral Thesis

Scientific supervisor
Professor Dr. habil. sc. ing.
LEONĪDS RIBICKIS

RTU Izdevniecība
Rīga 2019

Veckalns, V. Studies of Colour Flow in Top Quark Pair Decays at 13 TeV at the CMS Experiment of the CERN LHC. Summary of the Doctoral Thesis. Rīga: RTU Izdevniecība, 2019. 73 pp.

Published in accordance with the decision of the Council for Conferral of Doctoral Degrees "RTU P-14" of 15 February 2019, Minutes No. 2019-1 (68)

Cover photo "Cross section of the CMS detector" by Hoch, M; Brice, M

ISBN 978-9934-22-270-2 (print)
978-9934-22-271-9 (pdf)

DOCTORAL THESIS PROPOSED TO RIGA TECHNICAL UNIVERSITY FOR THE PROMOTION TO THE SCIENTIFIC DEGREE OF DOCTOR OF ENGINEERING SCIENCES

To be granted the scientific degree of Doctor of Engineering Sciences, the present Doctoral Thesis has been submitted for defence at the open meeting of RTU Council for Conferral of Doctoral Degrees on May 28, 2019 at the Faculty of Power and Electrical Engineering of Rīgas Tehniskā universitāte, Āzenes iela 1, Rīga, room 212.

OFFICIAL REVIEWERS
Professor Dr. habil. sc. phys. Mārcis Auziņš
Latvijas Universitāte, Latvia

Professor Dr. Mikko Voutilainen
Helsinki Institute of Physics, Finland

Professor Dr. Roberto Tenchini
INFN Sezione di Pisa, Universita' e Scuola Normale Superiore, Italy

Professor Dr. Oskars Krievs
Rīgas Tehniskā universitāte

DECLARATION OF ACADEMIC INTEGRITY

I hereby declare that the Doctoral Thesis submitted for review at Riga Technical University for the award of a scientific degree of Doctor of Engineering Sciences is my own. I confirm that this Doctoral Thesis has not been submitted to any other university for an award of a scientific degree.

Viesturs Veckalns ..... (signature)

Date: .....

The Doctoral Thesis has been written in English. It consists of an Introduction; 9 Chapters; Conclusion; 122 figures; 32 tables; 1 appendix; the total number of pages is 144. The Bibliography contains 126 titles.

Abstract

The LHC operating at 13 TeV centre of mass energy is a factory of top quarks. The cross section of the production of the top quark pair at the LHC is 803 pb. The lifetime of the top quark is 3.3×10^{-25} s and it is so short that unlike other quarks the top quark decays before it hadronises. The top quark decays weakly emitting a W boson. In the case of the hadronic decay of the W boson, jets of particles are created through the strong nuclear interaction. This process is described by quantum chromodynamics and allows us to model the top quark decay process in terms of colour charge and colour strings. The jets from the hadronic decay of the W boson are interacting in the colour field (they are colour-connected). The colour connection leaves distinct experimental signatures that we are able to resolve in the CMS detector, particularly relying on its tracker, 4 T solenoid and calorimeters. Such a study is conducted for the first time at the CMS experiment. The colour connection between jets from the decay of top quark pairs is studied using final states with one lepton, two light jets and two b-tagged jets. Pull angles and projections of particle directions onto a plane formed by two jets are used. A colour octet W toy model is used to benchmark the performance of the methods.

Acknowledgements

The groundwork for this work was done in the scope of the Picosecond Siliconphotomultiplier-Electronics-Crystal research Marie-Curie Network project. I am thankful to Etiennette Auffray (CERN, Switzerland) for organising this project. I am also thankful to Michele Gallinaro (LIP, Portugal) for providing guidance for my initial work in the CMS experiment.

I am also grateful to my team members Martijn Mulders (CERN, Switzerland), Pedro Silva (CERN, Switzerland) and Markus Seidel (CERN, Switzerland) for providing guidance and sharing experience.

I acknowledge the relentless support of Rīgas Tehniskā universitāte, Latvia which allowed to continue my work at the CMS experiment.

Contents

OVERVIEW

I Physical background

The Large Hadron Collider (LHC) is a synchrotron of 27 km in circumference. It is located in Geneva area on the Franco-Swiss border. In the experimental insertions of the LHC proton-proton (pp) collisions take place. The majority of these collisions are inelastic and we analyse the debris with the help of detectors such as the CMS. In the debris we hope to find answers to fundamental questions in physics such as the existence and properties of the Higgs boson, Dark Matter and the properties of the top quark.

The centre of mass energy of pp collisions is 13 TeV. Such an energy is sufficient to create millions of $t\bar{t}$ pairs – the cross section of this process at $\sqrt{s} = 13$ TeV is 803 pb [1]. The LHC can be called a factory of the top quark. The top quark decays in the weak process emitting a W boson. The ensemble of the hadronic decay products of the W boson belong to the colour singlet. If the W boson decays into colourful products (quarks) then these products interact in the chromodynamic field – they are colour-connected.

The light quarks that are created in the decay of the W boson hadronise and can be observed in the detector as jets. The silicon tracker of CMS, the electromagnetic and hadronic calorimeter provide means to resolve jet constituents – the products of hadronisation (baryons and mesons). Additionally, the 4 T superconducting solenoid of CMS allows us to measure the transverse momentum of jet constituents with a high resolution. The particles are identified and their parameters measured by correlating measurements in various subdetectors [2]. In the case of colour-connected jets their constituents tend to fill the space between jets in the laboratory frame. This property underlies the methods used in this work. The task is to observe colour-connected jets in the detector by applying these methods.

We also investigate the decay of a hypothetical colour octet W boson. In this case the light jets are not colour-connected. We can use these results to benchmark the colour-connected case.

We select events according to the $tt \to bW(q_1q_2)bW(\ell\nu)$ topology. The events must have 2 light jets, 2 b-tagged jets as well as a charged lepton.

In this work we use 2016 data of the LHC CMS with the total integrated luminosity of 35.9 fb^{-1}. The experimental observations are compared to Monte Carlo (MC) simulations. The MC simulations provide means to assess background, determine the efficiency of the event selection and in our case also to assess the accuracy of hadronisation models. The hard process is simulated with POWHEG whilst the showering is simulated with PYTHIA. The showering simulated by PYTHIA is compared to showering simulated by HERWIG++. The detector is being modelled by GEANT4. To correct for differences in MC and experimental observations scale factors are applied to the former. The error due to various systematics is also evaluated.

Having regard to the poor track reconstruction efficiency of particles the transverse momentum of which is less than 1 GeV we include only particles the transverse momentum of which is greater than 1 GeV.

II Methods

We use the method of the pull angle [3]. According to this method a pull vector is constructed given the centre of the jet, and distance of jet constituents from it weighted by the transverse momentum p_{T} of the jet constituents. It is expected that the pull vector of a jet will point to another jet to which it is colour-connected. Hence it is expected that the distribution of the pull angle will have a peak centred at 0 rad.

We investigate the distribution of the pull angle between colour connect jets (both light jets) and compare these results to the distribution of pull angle between physical objects not connected in colour – b-tagged jets, light jet and the lepton. An interesting case is the pull angle between the jet and the beam.

We separate cases when ΔR between jets is greater than or less than 1. In the latter case the anti-k_{T} jet clustering algorithm induces a pull from the hard jet to the soft jet creating a significant bias on observations according to the pull angle method.

We assess the sensitivity of the pull angle method to various parameters – including only charged particles (only the charged particles are deflected in magnetic field), the transverse momentum of the W boson, the number of jet constituents, a threshold on the transverse momentum of jet constituents, the magnitude of the pull vector.

In order to revert the effects of the detector on the observation we carry out unfolding. This method provides means to estimate the expected true distribution of the observable albeit at a cost of a course granularity of the phase space. We evaluate the goodness of fit between the unfolded observations and the generated Monte Carlo observations. We also assess the effect of various systematics.

We also use an adaptation of a method used in the Large Electron Positron collider (LEP) (hereinafter referred to as the "LEP method") where the jet constituents are projected on inter-jet planes [4], [5], [6]. It is expected that a plane between jets connected in colour will be filled more densely with particles than a plane between jets not connected in colour.

The results are obtained using a CMSSW release CMSSW_8_0_26_patch1, initially also with RIVET [7].

Finally we carry out hypothesis testing. In this task we combine the $t\bar{t}$ signal with the colour octet W signal and assess the goodness of fit of this combination to data.

III Personal contribution

The results described in this thesis are my personal effort. A significant part of this effort has been the development of a sizable computer code. I owe to my colleagues for the development of the event selection code, especially having regard to its complexity and effort that must be put in.

IV Novelty

The method of the pull angle has been applied at the DØ experiment of the Fermilab Tevatron [8], in Run I at the ATLAS experiment [9], as well as in Run II at the ATLAS experiment [10]. At CMS this method has first been applied by Seidel, M. et al [11] however these

results have never been published. Compared to ATLAS the CMS detector has a better momentum resolution for tracks in the central region by roughly a factor of 2 (ATLAS has a much smaller 2 T solenoid with big toroid magnets on the outside [12]).

"LEP method" has not yet been applied at the LHC.

This work is the first contribution of Latvia to the experimental programme of the CERN LHC. When work referenced in this thesis was in full progress we in May, 2018 celebrated the adhesion of Rīgas Tehniskā universitāte to a full membership of the CMS experiment.

V Outreach

This thesis shows the results from a research activity undertaken by the Top Quark group of the CMS experiment. The results at various stages have been presented in the Top Modelling and Generator physics meetings on 19 January 2016, 29 March 2016, 7 June 2016, 30 August 2016, 13 February 2018 and 17 October 2018. They have also been presented at the CERN Science Week in Rīga from May 22–26 of 2017 and in the European School of High Energy Physics in Evora, Portugal from September 6–19 of 2017.

These results have not been approved yet according to the CMS procedures for approval and publication accepted by the CMS Collaboration Board hence they cannot be published in refereed journals or presented in official conferences.

SUMMARY OF CHAPTERS

1. Physical Background

The cross section of top quark pair production in proton-proton collisions at $\sqrt{s}$ =13 TeV is measured be 803 $\pm$ 2 (stat) $\pm$ 25 (syst) $\pm$ 20 (lumi) [1]. The cross section increases once the centre of mass energy is increased as illustrated in Fig. 1.1.

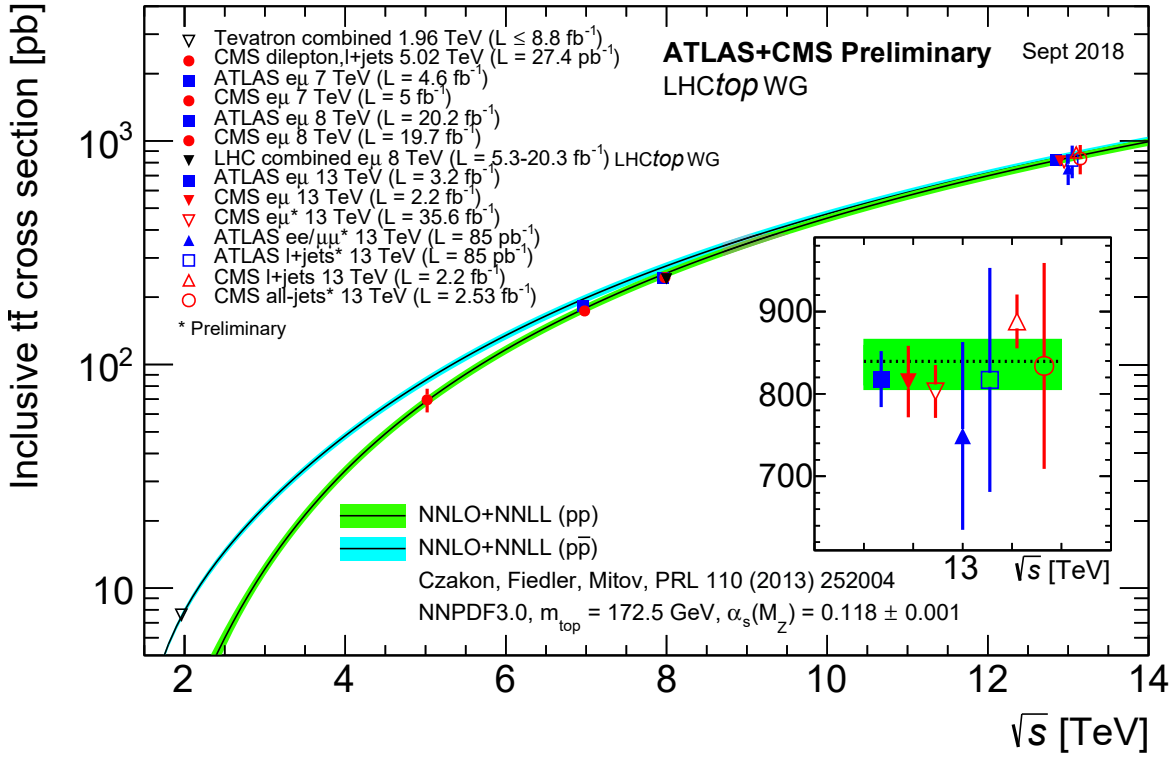


Figure 1.1. Inclusive cross section of the top pair at different centre of mass energies [13]. The result shows the cross section at pp and $p\overline{p}$ collisions. The results from CMS and ATLAS are given for different channels of the decay of the $t\bar{t}$ pair.

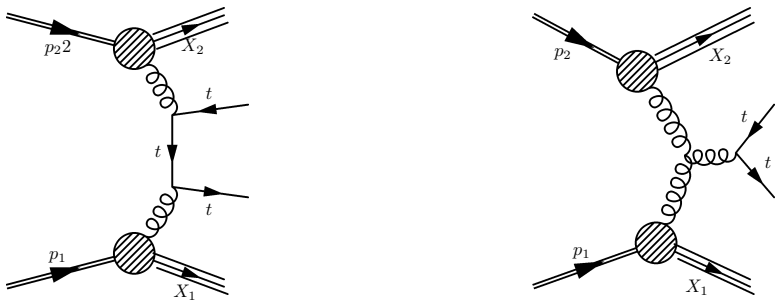
Using the relationship

$$N = \sigma \int L(t)dt, \tag{1.1}$$

where N is the number of $t\bar{t}$ pairs, σ is the $t\bar{t}$ cross section, L is the instantaneous luminosity, at 35.9 fb^{-1} integrated luminosity 26.7×10^6 of such pairs are expected to be created.

In the LHC 2 protons collide with an energy large enough to "squeeze" the protons so closely together that the quarks in one proton are able to interact with the quarks in the other proton. They interact by exchanging a gluon. By such an exchange the top quark-antiquark pair can be

created. Fig. 1.2 illustrates 2 such scenarios. The gluon exchanged is so energetic as to smash the proton into debris. Such a collision is called inelastic.



(a) Pair creation. (b) Gluon fusion.

Figure 1.2. Top quark pair production in a pp collision.

The top quark decays exclusively in the weak decay process (Fig. 1.3). In the weak decay W boson and a quark of different flavour and magnitude of electrical charge $^1/_3e$ is emitted.

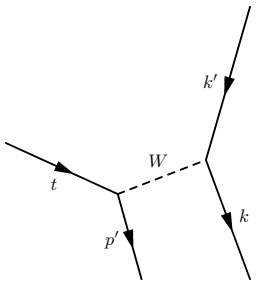


Figure 1.3. Weak decay of the top quark t. k and k' are fermions resulting from the decay of the W boson.

The average of measurements by CDF, DØ experiments of Tevatron [14], and ATLAS and CMS experiments of LHC [13] yield the result of the $|V_{tb}|$ term of the Cabibbo–Kobayashi–Maskawa matrix

$$|V_{tb}| = 1.009 \pm 0.031. \tag{1.2}$$

This would imply the top quark decays by emitting the b quark in at least $(0.98)^2$ of the cases. The other elements of the CKM matrix involving the top quark are very small [15]:

$$|V_{td}| = 8.4 \times 10^{-3}, \qquad |V_{ts}| = 40.0 \times 10^{-3}. \tag{1.3}$$

The width of the top quark as measured by the DØ collaboration [16] with 2.3 fb^{-1} of integrated luminosity is $\Gamma = 1.99^{+0.69}_{-0.55}$ GeV. This translates into a lifetime of $\tau = 3.3 \times 10^{-25}$ s.

This lifetime is smaller than the hadronisation timescale ($^1/_\Lambda \sim 10^{-24}$ s), where Λ^2 is the value of Q^2 of the exchanged gluon at which the strong coupling constant α_s becomes $\sim$1, close

to its asymptotic value at the confinement barrier. Thus the top quark decays before it hadronises and the experimentalist has a unique opportunity to observe a "bare" quark for a very short time.

The lifetime of the top quark is also smaller than the spin decorrelation of the top quark pair $M/\Lambda^2 = 3 \times 10^{-21}$ s. This means that the top quark pair maintain their spin states before they decay and transfer the spin states to their decay products [17].

The branching ratios of the decay of the top quark are essentially those of the decay of the W boson. The W boson can decay to any of the pair of leptons ($e\ \nu_e$, $\mu\ \nu_\mu$, $\tau\ \nu_\tau$) or the pairs of u, d' and c, s' quarks (where the apostrophe means flavour symmetry is not exactly conserved). However, the quark pairs can have 3 colours. Thus the total number of states is $3 + 2 \times 3 = 9$. A simple estimate and experimentally observed branching ratios from the decay of the W boson are given in Table 1.1.

Table 1.1

Branching ratios from the decay of the W boson.

Mode	Γ_j/Γ **simplified**	Γ_j/Γ **observed [15]**
$e\nu_e$	$1/9$	(10.71 ± 0.16) %
$\mu\nu_\mu$	$1/9$	(10.63 ± 0.15) %
$\tau\nu_\tau$	$1/9$	(11.38 ± 0.21) %
pair of quarks	$2/3$	(67.41 ± 0.27) %

colour-connected jets are emitted in the hadronic decay of the W boson (Fig. 1.5). The quarks originating these jets have opposite momenta in their COM frame. As the quarks try to separate, their kinetic energy is transferred to the colour field. The extra energy in the colour field equal to about m_W (80.4 GeV) is expended to create new particles. A simplified portrayal of the birth of new hadrons is given in Fig. 1.4, which is based on the Lund model [18]. An alternative portrayal based on Feynman diagrams is given in Fig. 1.6.

The following species of particles are created in the case of a hadronic decay of the W boson:

Table 1.2

New particles created in the colour field between energetic colour-connected quarks originating from a hadronic decay of the W boson.

Particle	**Mass [GeV]**	**Lifetime [s]**	**Observable signal**
π^0	135.0	8.5×10^{-27}	2 γ absorbed at ECAL
$\pi^\pm$	139.6	2.6×10^{-8}	tracker, ECAL, HCAL showers
K_S^0	497.6	8.95×10^{-11}	ECAL, HCAL showers
K_L^0	497.6	5.1×10^{-8}	ECAL, HCAL showers
$K^\pm$	493.7	1.2×10^{-8}	tracker, ECAL, HCAL showers
n	939.6	881.5	ECAL, HCAL showers
p	938.3	∞	tracker, ECAL, HCAL showers

The respective resonances are clearly discernible at the generation level (Fig. 1.7). Only the neutral pion decays before being directly observed in the detector.

The distribution of the number of particles that constitute the leading light jet is given in Fig. 1.8, while the distribution of the ratio of the number of electrically charged to the total

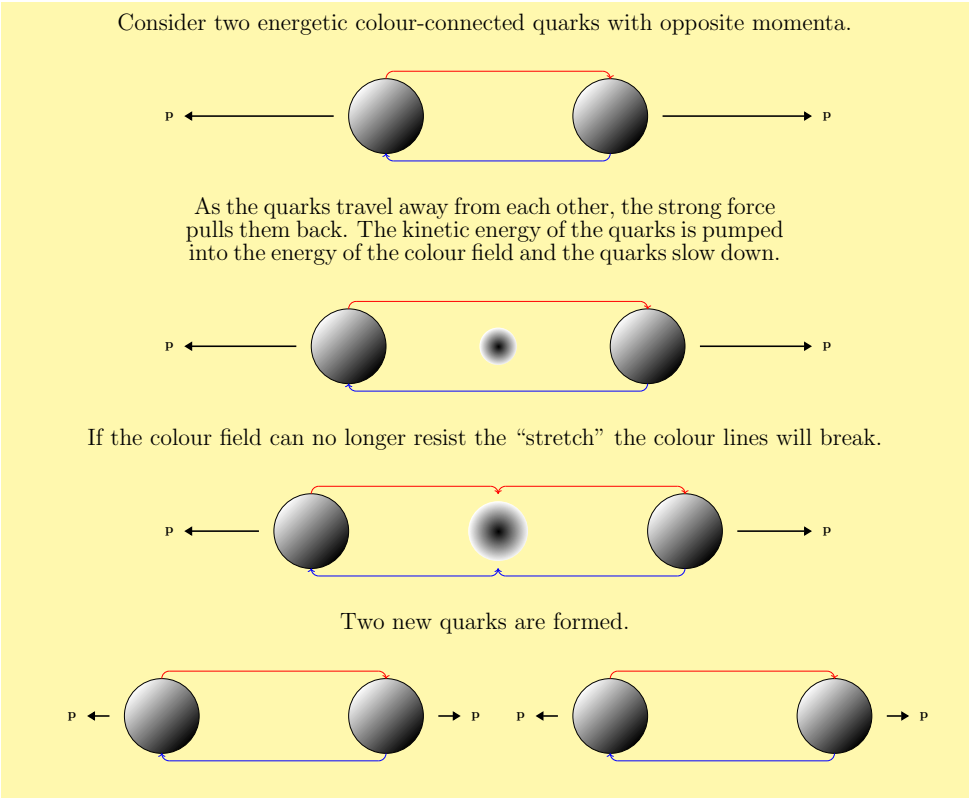


Figure 1.4. Creation of new hadrons by two energetic quarks.

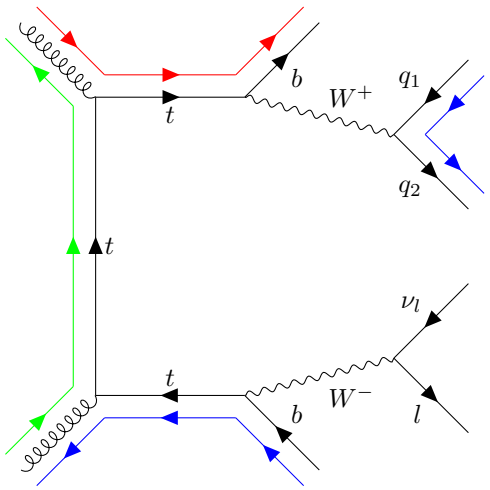


Figure 1.5. Colour flow in the decay of a top quark pair.

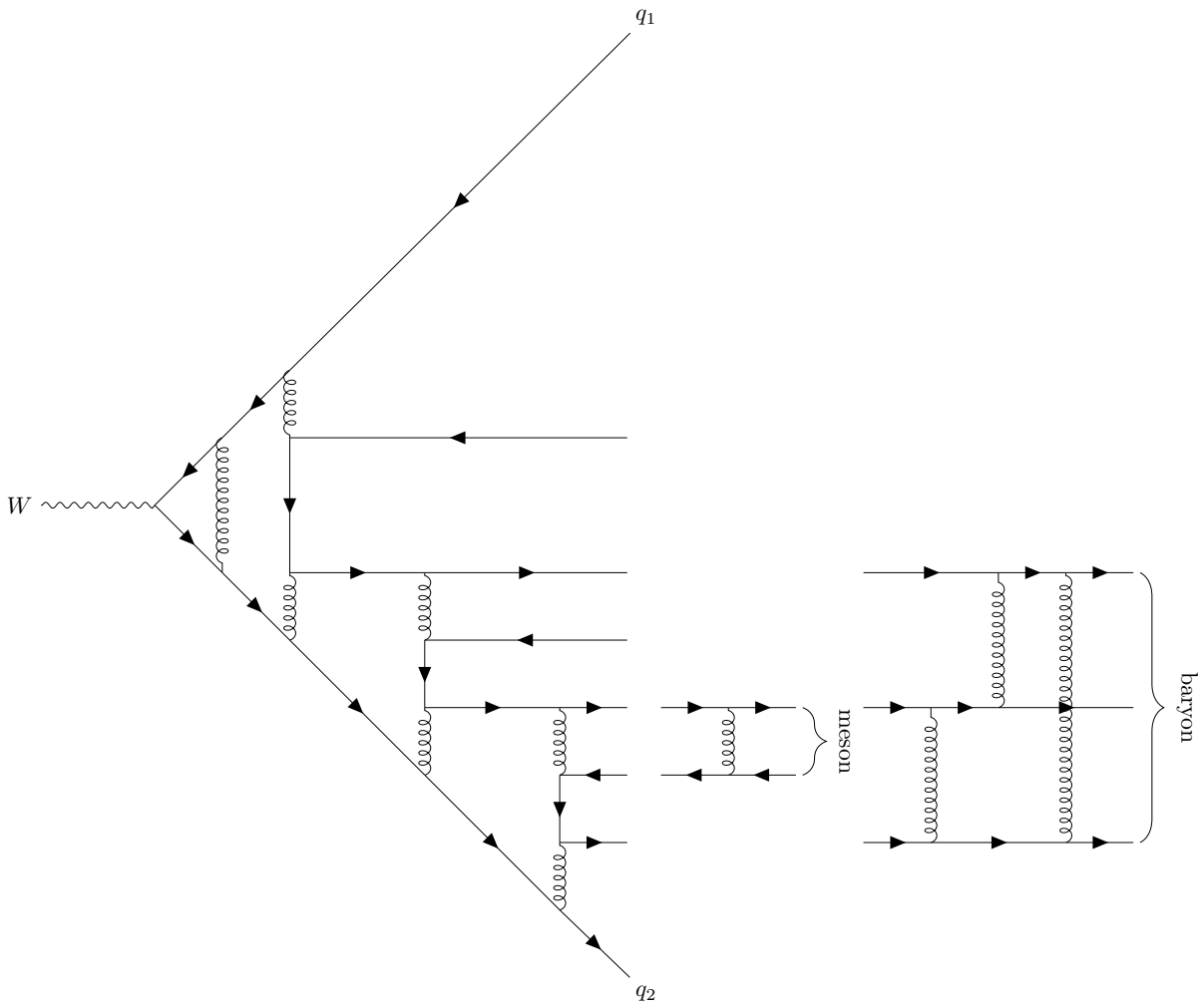


Figure 1.6. Creation of hadrons in the colour field of two quarks.

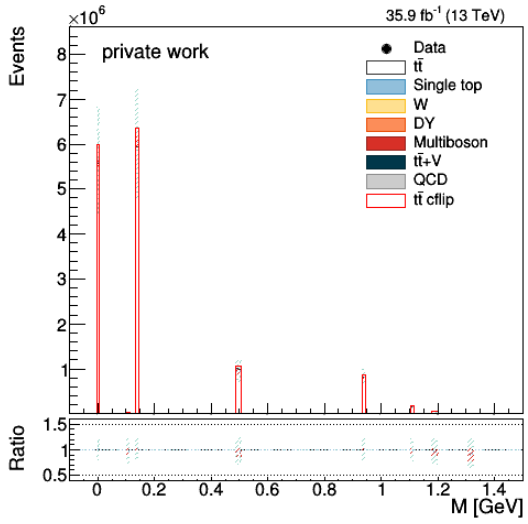


Figure 1.7. Resonances corresponding to particles listed in Table 1.2.
Note: This plot and a number of subsequent plots follow a format adopted at CMS to plot the value of an observable in a counting experiment. For an explanation of the format see Chap. 7.

number of particles is given in Fig. 1.9. The leading light jet is the jet from the decay of the W boson that has the highest transverse momentum p_{T}.

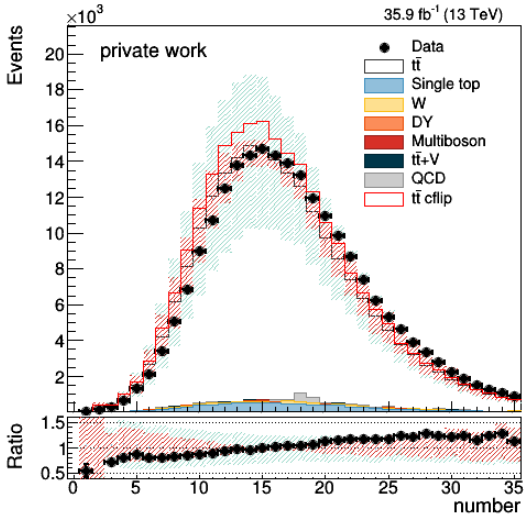


Figure 1.8. Distribution of the total number of particles constituting the leading light jet.

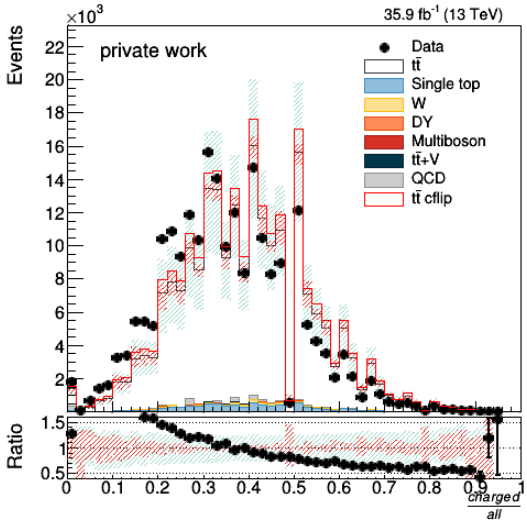


Figure 1.9. Distribution of the ratio of the number of charged particles to the total number of particles constituting the leading light jet.

Since we study light jets from the decay of the W boson it is interesting to ask why we need to concentrate on the $t\bar{t}$ process. The W production cross section is $> 20 \times$ larger than the $t\bar{t}$ cross section. In the study we need to use $W \to qq'$ events as the leptonic decays do not have colour flow. It is hard to trigger on resolved $W \to qq'$ events with sufficiently low p_{T} thresholds, so we use $t\bar{t}$ events where one of the W bosons decays leptonically and it is used to trigger the event while the other one decays hadronically and it is used to study colour flow.

In the case of a hadronic decay the W boson decays to the colour singlet:

$$\frac{1}{\sqrt{3}}\left(R\overline{R} + G\overline{G} + B\overline{B}\right), \tag{1.4}$$

where R, G and B are the three quantum states of the colour wave function. We mention this feature in light of our subsequent discussion of the colour octet W boson.

A W boson belonging to the colour octet is assumed. Its colour wavefunctions can take any of the 8 combinations:

$$R\overline{G},\quad R\overline{B},\quad G\overline{R},\quad G\overline{B},\quad B\overline{R},\quad B\overline{G},\quad \frac{1}{\sqrt{2}}\left(R\overline{R}-G\overline{G}\right),\quad \frac{1}{\sqrt{6}}\left(R\overline{R}+G\overline{G}-2B\overline{B}\right). \quad (1.5)$$

The only known particle in nature that belongs to the colour octet is the gluon. The colourful W boson is a purely hypothetical particle. The mass of the colour octet W boson is assumed to be equal to m_W. This boson would couple in colour field the light quarks to the hadronic b and the hadronic t, while the light quarks would become uncoupled from each other (Fig. 1.10).

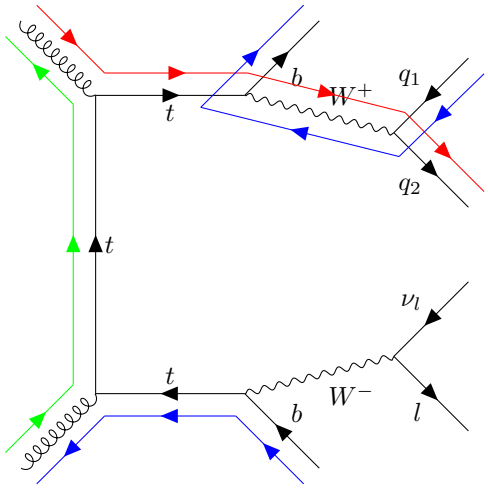


Figure 1.10. Colour flow in the decay of a top quark pair involving a hypothetical colour octet W boson.

2. Experimental Setup

The present study is conducted using arguably the most complex and largest experimental setup in the history of humanity, involving one of the most global collaborations in research. The LHC and its experiments were designed and built to answer some of the most fundamental questions in physics:

- Study electroweak symmetry breaking and search for the Higgs boson. Predicted in 1964 [19], [20] the Higgs boson had been the missing piece of the Standard Model. If discovered, it would confirm fundamental concepts of our understanding of the subatomic world. The relevant discovery was announced simultaneously by CMS and ATLAS in 2012 [21], [22] after almost 50 years of search.
- Study Standard Model physics to unprecedented detail with state-of-the-art detectors, high integrated luminosity and high centre-of-mass energy. One of the most interesting areas is studying the newly discovered top quark. Due to its high mass the top quark is predicted to couple well with the Higgs boson.
- Create the conditions for the primordial Quark-Gluon Plasma thus answering fundamental questions about the evolution of our Universe.
- Search for the Dark Matter, exotic particles, supersymmetric partners, extra dimensions and other puzzling and hypothetical topics beyond the Standard Model. These questions are still elusive and are motivations behind the High-Luminosity LHC, Future Circular Collider and other experimental concepts on a grand scale.

The CMS experiment is one of the flagship experiments of the Large Hadron Collider. Hence, in the present discussion the LHC will be presented first followed by a description of the CMS apparatus.

2.1 The LHC

The LHC is a two-ring superconducting hadron accelerator and collider installed in a 26.7 km tunnel 45–170 m underground traversing the Franco-Swiss border in Geneva area (Fig. 2.1). The hadrons circulate in the LHC with a constant radius but variable frequency. Hence, the LHC is a synchrotron. It reuses the tunnel and injection chain of the Large Electron-Positron collider (LEP).

Initially the LHC project faced severe competition from the more powerful Superconducting Super Collider in the USA. C. Rubbia argued that the luminosity higher by a factor of 10 at the

LHC would compensate its lower energy vis-à-vis the SSC. Eventually, the SSC project was cancelled in 1993. Cost overruns played a role. The CERN Council approved the LHC project in 1994. It started data taking in 2008.

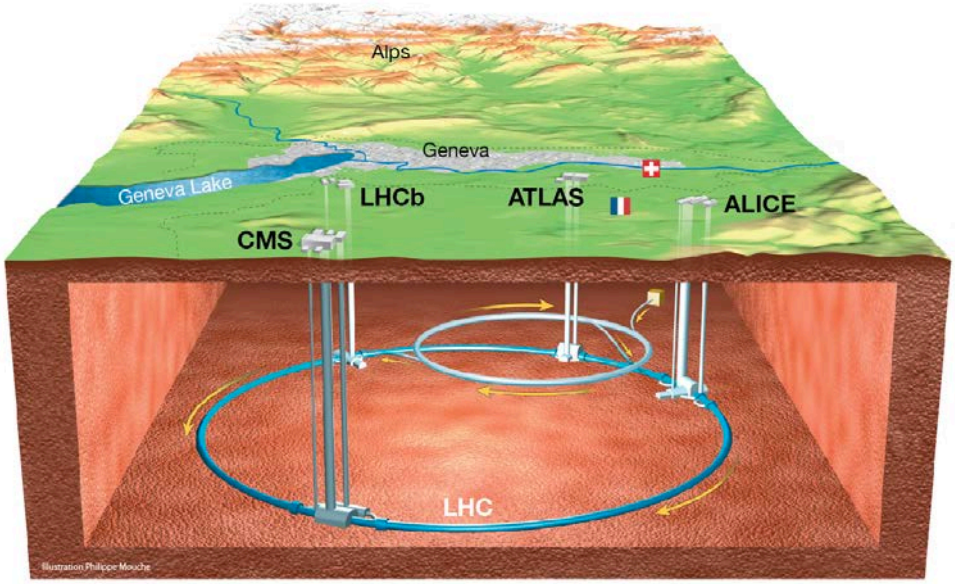


Figure 2.1. The Large Hadron Collider situated underground on the French-Swiss border in Geneva area [23].

The protons at the LHC circulate at nearly the speed of light. The per proton energy is 7 TeV, the γ factor being 7461. It is not practical to accelerate a proton from zero velocity to such an energy in one accelerator. Therefore before reaching this energy the protons undergo a sequence of accelerations in the CERN accelerator complex (Fig. 2.2):

- up to 50 MeV in Linac2,
- up to 1.4 GeV in PS Booster,
- up to 26 GeV in the Proton Synchrotron (PS),
- up to 450 GeV in the Superproton Synchrotron (SPS).

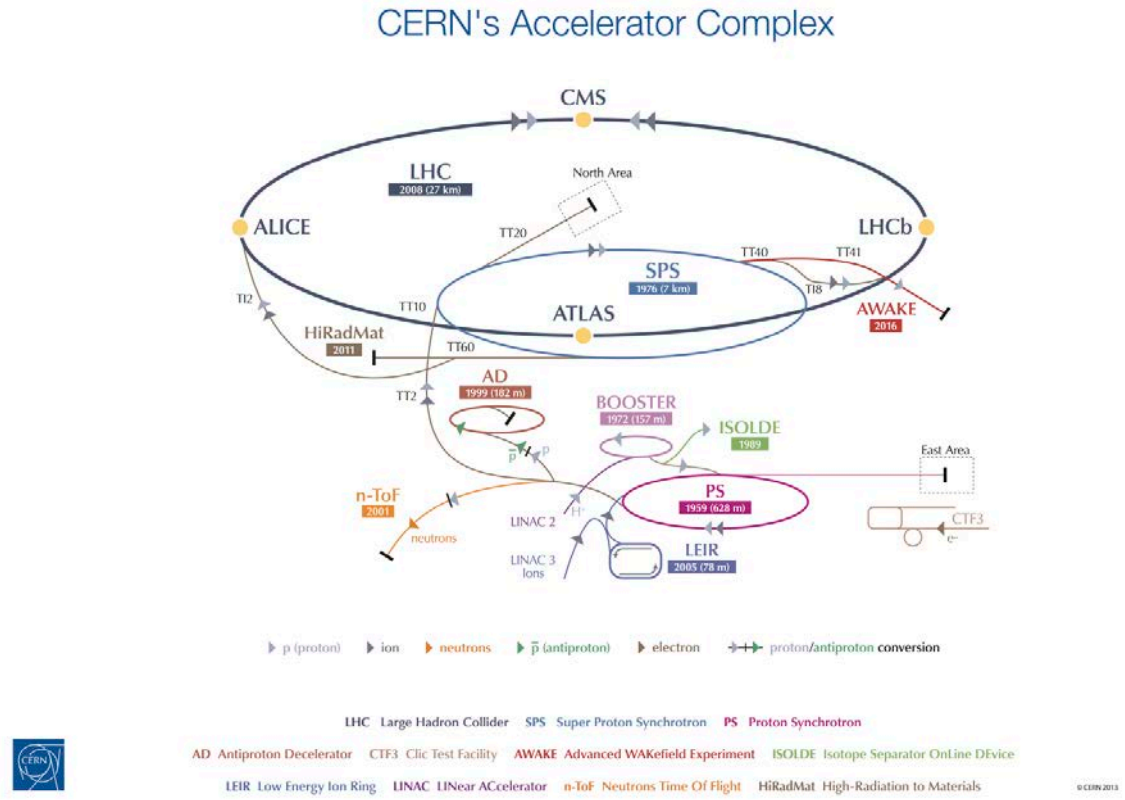


Figure 2.2. The CERN accelerator complex [24].

After the protons have been fully accelerated they are allowed to circulate in the LHC – the LHC is a storage ring. There are 1.15×10^{11} protons in each bunch and 2808 bunches in circulation. The revolution frequency is 11.245 kHz [25]. Each bunch crossing lasts 25 ns. There is an ultra-high vacuum maintained in the beam pipes.

The LHC uses superconducting magnet systems. Particularly, the dipole magnets bend the beam in a circular arc, and quadrupole magnets squeeze the beam near the collision points. Magnets of higher orders provide steering and correction to the beam. The magnet systems rely on the NbTi Rutherford cable, that is cooled by helium to below 2 K – below the lambda point of helium. Thus unlike other large accelerators that use NbTi but operate above the lambda point of helium (Tevatron-FNAL, HERA-DESY and RHIC-BNL) a much higher field of 8 T can be achieved in the dipole magnets at the LHC. A special two-in-one dipole magnet was designed for the LHC that uses the same yoke but fields of different polarities for the two proton beams circulating in opposite directions. Cooling the magnets requires the largest cryogenic system on Earth [26], [27].

The design COM of the LHC is 14 TeV. In its first data taking period from 2010–2013 it operated at $\sqrt{s}$ =7–8 TeV. This period is referred to as Run I. In its second data taking period from 2015–2018 referred to as Run II it operated at $\sqrt{s}$ =13–14 TeV. The present study is conducted with Run II data.

The LHC houses two high-luminosity experimental insertions – CMS and ATLAS each targeting a luminosity above $10^{1}/_{\text{pb}\cdot\text{s}}$, one b physics experiment LHCb targeting a luminosity of $0.1\ ^{1}/_{\text{pb}\cdot\text{s}}$ and one dedicated ion collision experiment – ALICE.

2.2 The CMS detector

The CMS detector is located at Point 5 of the LHC, close to the French village of Cessy, between Lake Geneva and the Jura mountains. It is placed in underground caverns about 100 m deep that were excavated to house the detector complex.

The CMS detector is designed to operate in diverse physics programmes in the TeV range. It is an onion-type detector covering 4π of solid angle around the collision point. The CMS detector is composed of the following layers starting from the beam axis – a silicon pixel and strip tracker, a lead tungstate electromagnetic calorimeter, a brass and a plastic scintillator hadron calorimeter, a superconducting magnet producing 3.8–4.0 T of magnetic field, and a gas-ionisation muon spectrometer [28]. The shape of the CMS detector is a cylinder. It has endcaps on both ends while the cental part is called the barrel. The length of the CMS detector is 21.6 m, diameter 14.6 m and total weight 12 500 t. A cut-away view of the CMS detector is presented in Fig. 2.3.

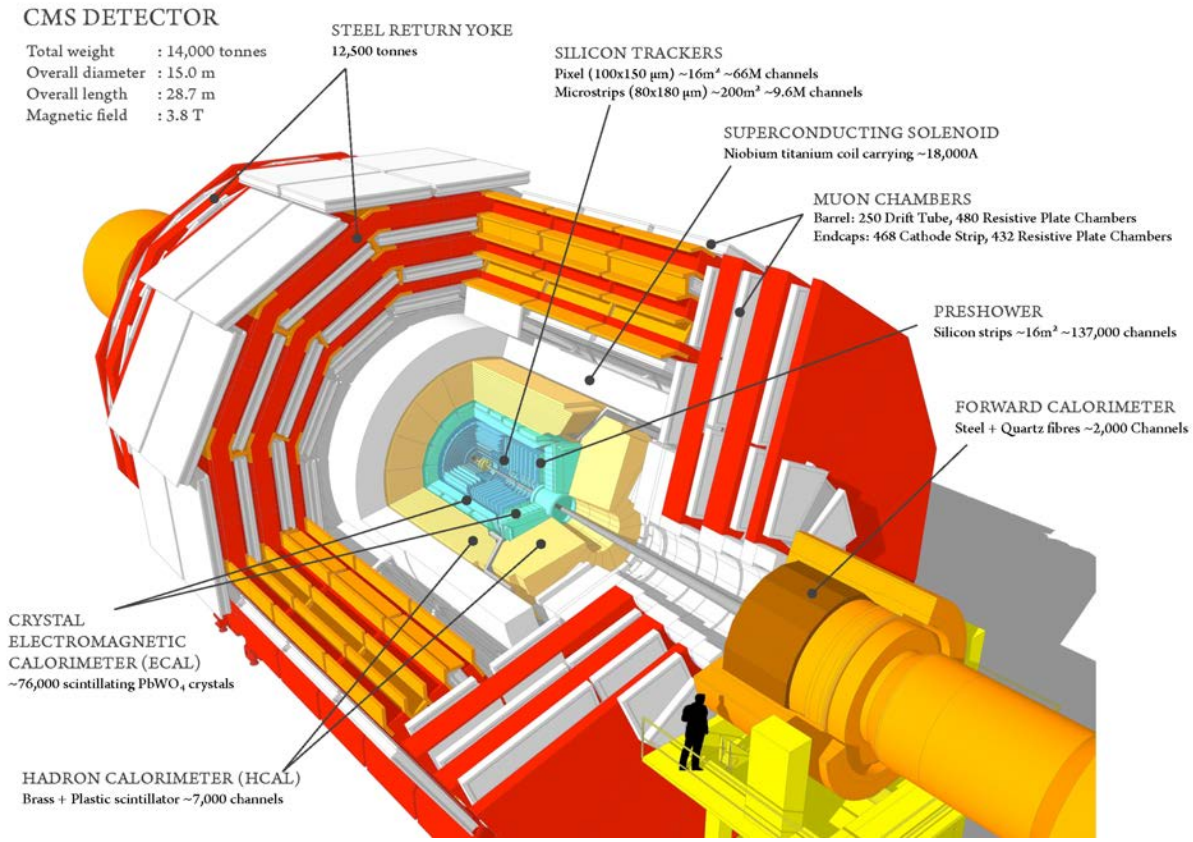


Figure 2.3. A cut-away view of the CMS detector [29].

Starting from the beam interaction region, particles first enter a tracker, in which charged-particle trajectories (tracks) and origins (vertices) are reconstructed from signals (hits) in the sensitive layers. The tracker is immersed in a magnetic field that bends the trajectories and allows the electric charges and momenta of charged particles to be measured. Electrons and photons are then absorbed in the electromagnetic calorimeter (ECAL). The corresponding electromagnetic showers are detected as clusters of energy recorded in neighbouring cells, from which the energy and direction of the particles can be determined. Charged and neutral hadrons may initiate a hadronic shower in the ECAL as well, which is subsequently fully absorbed in the

hadron calorimeter (HCAL). The corresponding clusters are used to estimate their energies and directions. Muons and neutrinos traverse the calorimeters with little or no interactions. While neutrinos escape undetected, muons produce hits in additional tracking layers called muon detectors, located outside the calorimeters. This simplified view is graphically summarised in Fig. 2.4, which displays a sketch of a transverse slice of the CMS detector.

A significantly improved event description can be achieved by correlating the basic elements from all detector layers (tracks and clusters) to identify each final-state particle, and by combining the corresponding measurements to reconstruct the particle properties on the basis of this identification. This holistic approach is called particle-flow (PF) reconstruction [2].

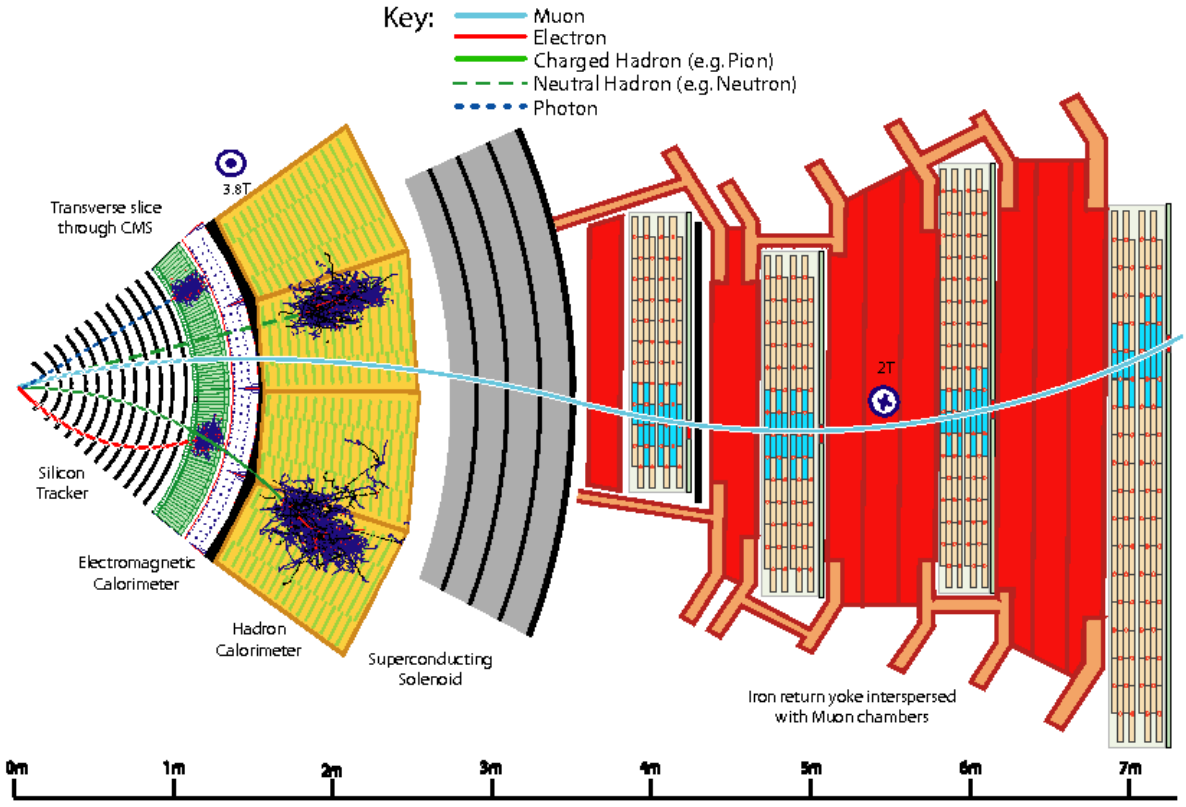


Figure 2.4. A sketch of the specific particle interactions in a transverse slice of the CMS detector, from the beam interaction region to the muon detector [2].

The fine-granularity and fast response tracker [30], [31] is an important segment in resolving the fine jet constituents. It is closely aligned to the beam axis and has a length of 5.8 m and radius of 2.5 m. The CMS solenoid provides a homogeneous and coaxial magnetic field of 3.8–4.0 T over the full volume of the tracker. At radius below 10 cm a hit rate at the order of $100\ {}^{\mathrm{kHz}}/_{\mathrm{mm}^2}$ is encountered. In order to achieve the desired resolution 100 µm × 150 µm pixel detectors are used. At a higher radius the reduced particle flux allows the use of silicon micro-strip detectors with a typical size of 10 cm × 80 µm to 25 cm × 150 µm, the size increasing with an increasing radius. There are 66 million pixels with 1 m^2 active area in the pixel detector and 9.3 million strips and 193 m^2 active area in the strip detector.

The electromagnetic calorimeter of CMS (ECAL) is a hermetic homogeneous calorimeter made of 61 200 lead tungstate ($PbWO_4$) crystals mounted in the central barrel part, closed

by 7324 crystals in each of the two endcaps. The barrel part covers the pseudorapidity range $|\eta| < 1.479$ while the endcaps cover the pseudorapidity range $1.479 < |\eta| < 3.0$. A preshower detector is placed in front of the endcap crystals. Avalanche photodiodes (APDs) are used as photodetectors in the barrel and vacuum phototriodes (VPTs) in the endcaps. The $PbWO_4$ crystals exhibit characteristics that make them an appropriate choice for an electromagnetic calorimeter at the LHC. The high density 8.28 $^{g}/_{cm^3}$, short radiation length (0.89 cm) and small Molière radius (2.2 cm) result in a fine granularity and a compact calorimeter. The scintillation decay time of $PbWO_4$ is of the same order of magnitude as the LHC bunch crossing time: about 80 % of the light is emitted in 25 ns. In the barrel the crystal cross section corresponds to approximately 0.0174 $\times$ 0.0174 in η-ϕ, corresponding to a front cross section 22 $\times$ 22 mm^2 and a rear cross section 26 $\times$ 26 mm^2. The crystal length is 230 mm, corresponding to 25.8 X_0. There are 61 200 crystals in the barrel. In the endcaps the crystals have a rear face cross section 30 $\times$ 30 mm^2, a front face cross section 28.62 $\times$ 28.62 mm^2 and a length of 220 mm (24.7 X_0). Additionally in the fiducial region $1.653 < |\eta| < 2.6$ there is a preshower detector whose principal aim is to identify neutral pions in the endcaps. The energy resolution of the barrel electromagnetic calorimeter depends on the incident energy and is measured from 0.94 % ($^{\sigma}/_{E}$) at 20 GeV to 0.34 % at 250 GeV [32]. The preshower detector consists of a lead radiator where electromagnetic showers from incoming electrons/photons are initiated. Behind the lead radiator there are silicon strips to measure the deposited energy and transverse shower profiles.

The hadronic calorimeter [33] consists of a barrel ($|\eta| < 1.3$) and two endcap disks ($1.3 < |\eta| < 3.0$). The space of the hadron calorimeter in the central pseudorapidity region is constrained. Therefore, an outer tail catcher layer behind the soleonoid is used. The solenoid is used as an additional absorber for the tail catcher. The absorber consists of a 40 mm thick front steel plate, followed by eight 50.5 mm thick brass plates, six 56.5 mm thick brass plates, and a 75 mm thick steel back plate. The total absorber thickness at 90 ° is 5.82 interaction lengths (λ_{I}). As the active material plastic scintillator arranged in tiles is used. Wavelength shifting fibres are used to bring out the light. The hadronic calorimeter is read out in individual towers with a cross section $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$ for $|\eta| < 1.6$ and 0.17×0.17 at larger pseudorapidities. The hadronic calorimeter at $|\eta|$ extending up to $\simeq 5.0$ where particle flux and radiation damage is highest is complemented by hadron forward calorimeters. The hadron forward calorimeter consists of a steel absorber composed of grooved plates. Radiation-hard quartz fibres are inserted in the grooves along the beam direction and are read out by photomultipliers. The signals are grouped so as to define calorimeter towers with a cross section $\Delta\eta \times \Delta\phi = 0.175 \times 0.175$ over most of the pseudorapidity range.

The magnet is located behind the calorimeters and the tracker to ensure that as less material as possible is situated between these subdetectors and the interaction point. The length of the magnet is 12.5 m and the free-bore radius is 3.15 m. The coil delivers a 3.8–4.0 T uniform and axial magnetic field to the tracker and the calorimeters. The magnet operates at 4.45 K and uses a NbTi superconducting coil. The magnet is characterised by a high stored-energy/mass ratio 11.6 $^{kJ}/_{kg}$.

The muon channel is a very powerful tool for studying interesting HEP processes and has been very important for CMS since the experiment's inception. This is because of the relative ease of detecting muons and because they are minimally affected by radiative losses in the tracker material. Four muon detector planes are located outside the solenoid coil interleaved with three

layers of steel yoke [34]. In the barrel region – $|\eta| < 1.2$ where the muon rate is low, and the 4 T magnetic field is uniform and mostly contained in the steel yoke, drift chambers are used. In the endcaps $0.9 < |\eta| < 2.4$ where the muon rates and background levels are high and the magnetic field is large and non-uniform, the muon system uses cathode strip chambers (CSC). Because of the uncertainty in the eventual background rates and in the ability of the muon system to measure the correct beam-crossing time when the LHC reaches full luminosity, a complementary, dedicated trigger system consisting of resistive plate chambers (RPC) is added in both the barrel and endcap regions. The RPCs provide a fast, independent, and highly segmented trigger with a sharp p_{T} threshold over a large portion of the rapidity range ($|\eta| < 1.6$) of the muon system. The particle flow reconstruction involves a global trajectory fit across the muon detectors and the inner tracker.

Jets are reconstructed using the anti-k_{T} algorithm [35] with radius parameter $R = 0.4$ as implemented by the FastJet [36] package. Distance d_{ij} between jets is determined by using $p = -1$ in the general formula:

$$d_{ij} = \min(k_{\mathrm{T}i}^{2p}, k_{\mathrm{T}j}^{2p}) \frac{\Delta_{ij}^2}{R^2}, \tag{2.1}$$

where $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$ and $k_{\mathrm{T}i}$, y_i, ϕ_i are respectively the transverse momentum, rapidity and azimuth of particle i.

The key feature of this algorithm is that soft particles do not modify the shape of the jet. Given separation between jets $\Delta_{ij} \leq 2R$ the jets have conical shapes.

3. Methodology

3.1 Pull angle

We adopt the methodology proposed by [3] to use the pull angle to reveal colour connection between two quark jets. The pull angle θ_{p} formed by the pull vector $\vec{v}_{\mathrm{p}}$ and difference between two jets $\vec{J}_2 - \vec{J}_1$ is shown in Fig. 3.1. The ϕ-y coordinate system is used.

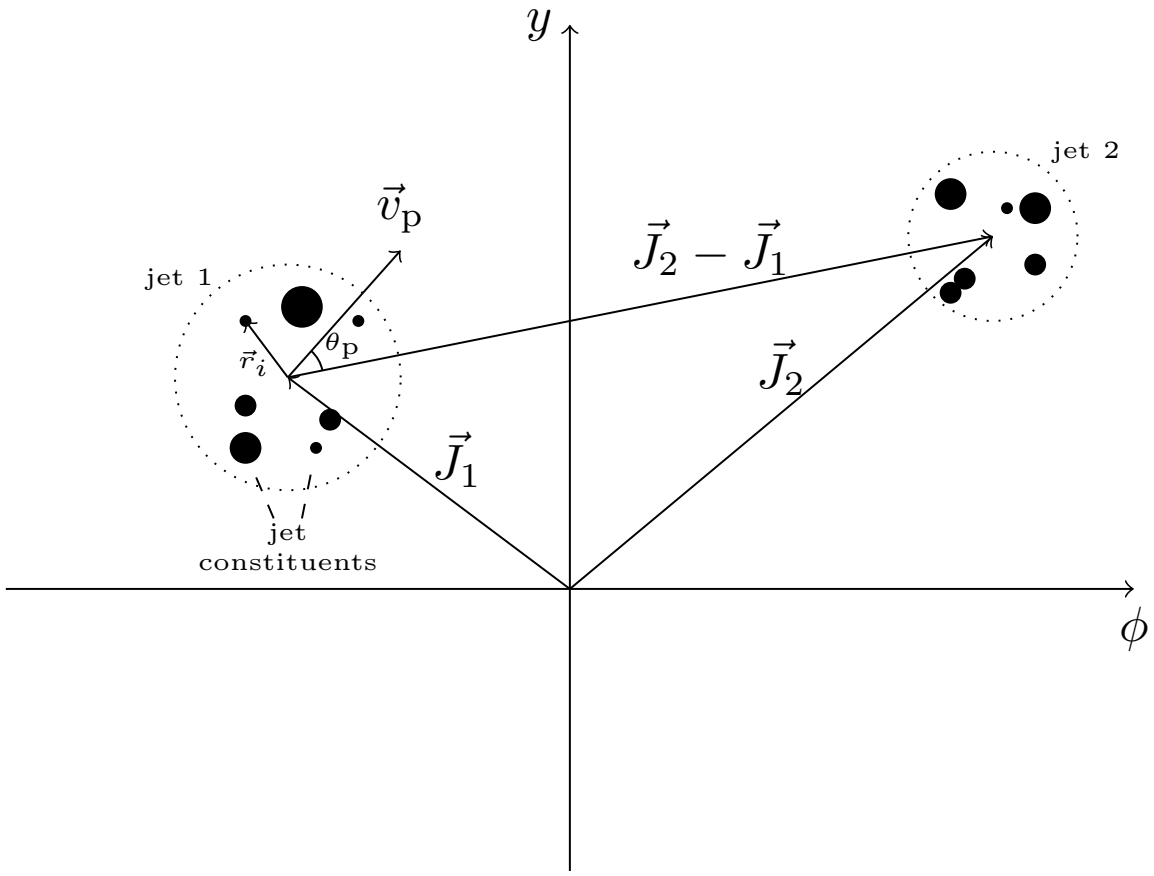


Figure 3.1. Pull angle θ_{p}, pull vector $\vec{v}_{\mathrm{p}}$ in a ϕ-y plane.

The pull vector is given by the formula

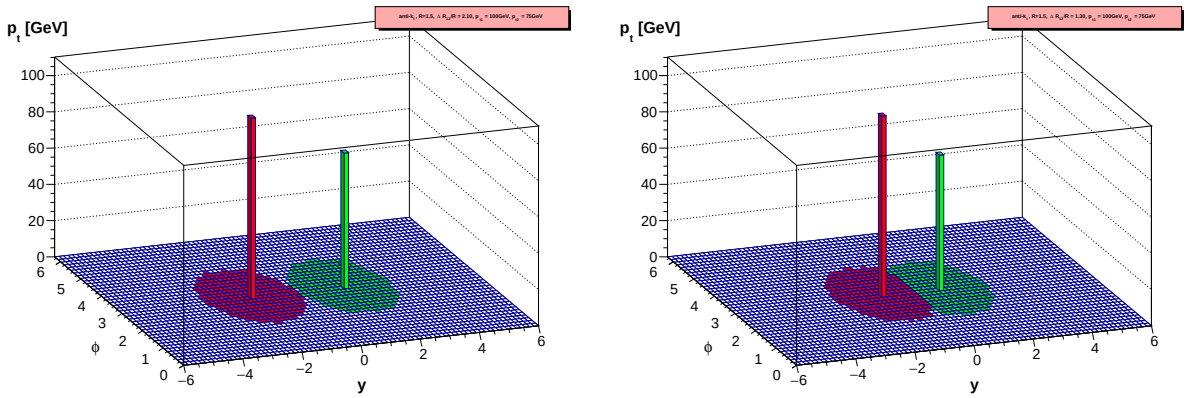
$$\vec{v}_{\mathrm{p}} = \sum_{i \in J} \frac{p_{\mathrm{T}}^i |\vec{r}_i|}{p_{\mathrm{T}}^J} \vec{r}_i, \tag{3.1}$$

where i is the index of the constituent of jet J, p_{T}^{i} is the transverse momentum of the jet constituent, $\vec{r}_i$ is the vectorial difference between the jet component and the jet, p_{T}^{J} is the transverse momentum of the jet.

Two jets that are colour connected are expected to have jet constituents dispersed in the area between the two jets. Hence the pull vector of J_1 would point towards J_2 and the pull angle would be narrow. For jets that are not colour connected the pull angle is expected to be distributed isotropically.

The methodology of the pull angle has been applied in the DØ experiment of Tevatron [8] and the ATLAS experiment at the LHC in Run I [9] and in Run II [37]. We hope to outperform all results with the methodology of the pull angle with the state-of-the-art tracker of the CMS detector.

The anti-k_{T} clustering algorithm ensures a conical jet shape in case the jet separation ΔR is more than double of the parameter R, which is set at 0.4 at CMS. This case is illustrated in Fig. 3.2(a). In case of separation between jets ΔR being less than double of the parameter R the hard jet will wean constituents from the soft jet. This is illustrated in Fig. 3.2(b). This latter effect will have consequences for the colour flow analysis with the pull angle as it will induce a pull from the involved jets to each other. This warrants a separation of the cases $\Delta R \leq 2R$ and $\Delta R > 2R$.



(a) $\Delta_{ij} = 3.15$. (b) $\Delta_{ij} = 1.95$.

Figure 3.2. Jet shapes obtained with the anti-k_{T} clustering. $R = 1.5$ is used. Two cases are shown – $\Delta_{ij} = 3.15$ and $\Delta_{ij} = 1.95$. The p_{T} of the hard jet is 100 GeV, the p_{T} of the soft jet is 75 GeV. Courtesy of Cacciari, Salam and Soyez [38].

Tracking efficiency of the detector is not perfect. It depends on the quality of the track finder algorithm and properties of the detector such as geometrical acceptance and material content. Fig. 3.3 shows the tracking efficiency of pions, a particle commonly resulting from quark hadronisation. Tracking efficiency is defined as the fraction of simulated charged particles that can be associated with corresponding reconstructed tracks. The tracking efficiency drops at low p_{T} of the particle. In our analysis we choose 1 GeV as the threshold and exclude particles whose p_{T} is below it from our analysis.

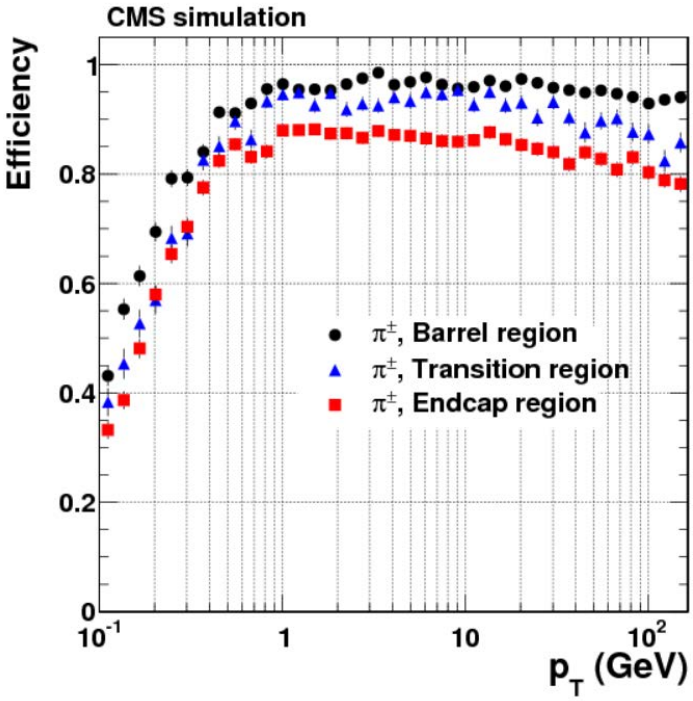


Figure 3.3. Track reconstruction efficiencies for pions passing the high-purity quality requirements. Results are shown as a function of p_{T}, for the barrel, transition, and endcap regions, which are defined by η intervals of 0–0.9, 0.9–1.4 and 1.4–2.5, respectively [39].

3.2 LEP method

Another methodology of studying colour-connected jets in the process $e^+e^- \to q\bar{q}q\bar{q}$ at $\sqrt{s}$ =189–207 GeV was used in various experiments of LEP [5], [4], [6]. Two inter-W planes formed by colour-connected quarks and two intra-W planes formed by quarks that are not colour-connected are introduced as shown in Fig. 3.4. Particles are projected onto these planes and the angle with the leftmost quark χ_1 is taken. If this angle is less than the angle χ_0 between the quarks forming the plane (which means the particle is projected between the respective quarks) then the normalised angle $\chi_{\mathrm{R}} = \chi_1/\chi_0$ is plotted in the region corresponding to the plane after a linear transformation

$$\chi = \chi_{\mathrm{R}} + n_{\mathrm{plane}} - 1 \tag{3.2}$$

has been performed on the normalised angle.

In the $t\bar{t}$ semileptonic decay an arrangement as shown in Fig. 3.4 is not possible. Therefore a modification as shown in Fig. 3.5 is proposed. There is one plane formed by colour-connected jets – the leading light jet j_1^W and the second leading light jet j_2^W from the hadronic decay of the W boson. Additionally there are 3 colour-free regions formed by 1) the furthest light jet j_{f}^W and the b jet from the hadronic decay of the W boson j_1^b, 2) the hadronic b jet and the closest light jet j_{c}^W, 3) the leading b jet j_1^b and the second leading b jet j_2^b. Whether a jet is close or far is determined with regard to the angle between jets in the Euclidian space. In the regions shown in Fig. 3.5(b) and Fig. 3.5(c) we may hope to observe colour reconnection effects.

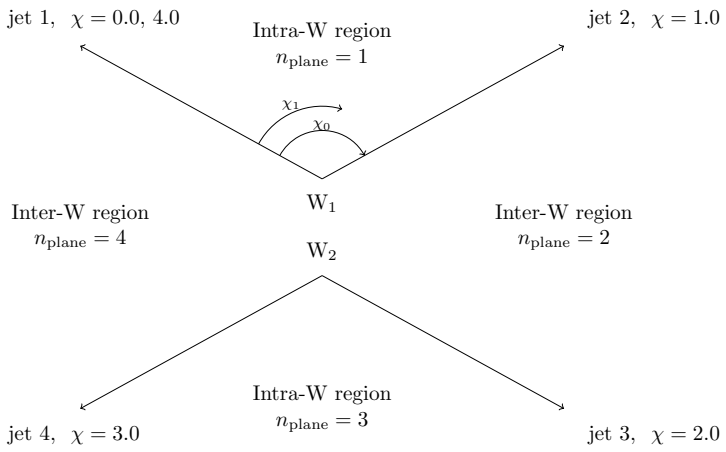
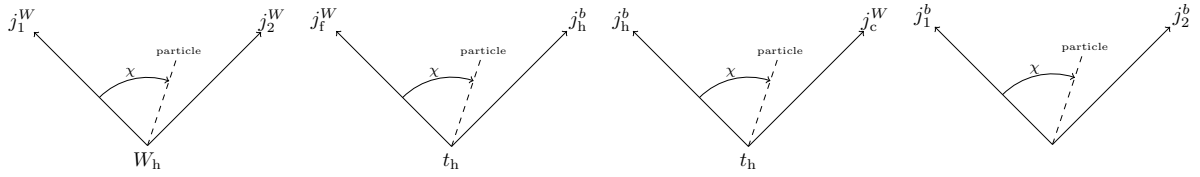


Figure 3.4. Inter-W and intra-W planes in the process $e^+e^- \to q\bar{q}q\bar{q}$ and the relative angle $\chi_{\text{R}} = \chi_1/\chi_0$.



(a) Colour-connected region j_1^W–j_2^W. (b) Colour-free region j_{f}^W–j_{h}^b. (c) Colour-free region j_{h}^b–j_{c}^W. (d) Colour-free region j_1^b–j_2^b.

Figure 3.5. Adaptation of the LEP method to $t\bar{t}$ semileptonic decay involving a colour-connected region and 3 colour-free regions.

The method calls for a separation of hadronic and leptonic b quarks. Each b quark is paired to each W boson and the invariant mass is compared to the mass of the t quark – 173.34 GeV. The b quark is assigned to the branch where the difference of the masses is the smallest.

4. Data and MC Samples

The discussion of this section is adapted from [40] and [41] as these studies use a similar set of data and MC samples.

The data analysed for the present study consist of the 2016B-H data taking periods for a total certified luminosity of 35.9 fb^{-1} for all the channels analysed. The luminosity has been computed with the BRILCALC tool [42] using the following command:

```
brilcalc lumi -b "STABLE BEAMS" --normtag /afs/cern.ch/user/l/
    lumipro/public/Normtags/normtag_DATACERT.json -i lumiSummary
    .json
```

All data used for this study are listed in Table 4.1.

Table 4.1
Primary datasets used in this analysis. PD is an abbreviation for SingleMuon or SingleElectron [41].

Primary dataset	Integrated luminosity
/PD/Run2016B-23Sep2016-v3/MINIAOD	35.9 fb^{-1}
/PD/Run2016C-23Sep2016-v1/MINIAOD	
/PD/Run2016D-23Sep2016-v1/MINIAOD	
/PD/Run2016E-23Sep2016-v1/MINIAOD	
/PD/Run2016F-23Sep2016-v1/MINIAOD	
/PD/Run2016G-23Sep2016-v1/MINIAOD	
/PD/Run2016H-PromptReco-v2/MINIAOD	
/PD/Run2016H-PromptReco-v3/MINIAOD	

The list of simulated samples from the RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6 production can be found in Table 4.2.

The cross sections we use are theoretical predictions. Practically, they are obtained from [43] and [44] except for $t\bar{t}$ for which the generator cross section is quoted according to [45]. At NNLO the expected $t\bar{t}$ cross section is 832^{+20}_{-29} (scale) ± 35(PDF + α_s) [46]. We use the NNLO reference to normalise all $t\bar{t}$ samples.

Table 4.2
List of simulation samples. We quote the cross section used to normalise the sample in the analysis. Adapted after [41].

Process	Dataset	σ [pb]
	Signal	
$t\bar{t}$	TT_TuneCUETP8M2T4_13TeV-powheg-pythia8	832
	Background	
$t\bar{t}$ + W	TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.20
	TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.41
$t\bar{t}$ + Z	TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia	0.53
	TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.25
WZ	WZTo3LNu_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	5.26
WW	WWToLNuQQ_13TeV-powheg	50.0
	WWTo2L2Nu_13TeV-powheg	12.2
ZZ	ZZTo2L2Nu_13TeV_powheg_pythia8	0.564
	ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.22
W + jets	WToLNu_0J_13TeV-amcatnloFXFX-pythia8	49 540
	WToLNu_1J_13TeV-amcatnloFXFX-pythia8	8041
	WToLNu_2J_13TeV-amcatnloFXFX-pythia8	3052
Drell–Yan	DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	18 610
	DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	6025
μ enriched QCD	QCD_Pt-30to50_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	1 652 471.46
	QCD_Pt-50to80_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	437 504.1
	QCD_Pt-80to120_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	106 033.66
	QCD_Pt-120to170_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	25 190.52
	QCD_Pt-170to300_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	8654.49
	QCD_Pt-300to470_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	797.35
	QCD_Pt-470to600_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	45.83
	QCD_Pt-600to800_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	25.1
	QCD_Pt-800to1000_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	4.71
	QCD_Pt-1000toInf_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	1.62
e enriched QCD	QCD_Pt-30to50_EMEnriched_TuneCUETP8M1_13TeV_pythia8	6 493 800.0
	QCD_Pt-50to80_EMEnriched_TuneCUETP8M1_13TeV_pythia8	2 025 400.0
	QCD_Pt-80to120_EMEnriched_TuneCUETP8M1_13TeV_pythia8	478 520.0
	QCD_Pt-120to170_EMEnriched_TuneCUETP8M1_13TeV_pythia8	68 592.0
	QCD_Pt-170to300_EMEnriched_TuneCUETP8M1_13TeV_pythia8	18 810.0
	QCD_Pt-300toInf_EMEnriched_TuneCUETP8M1_13TeV_pythia8	1350.0

The colour octet sample is listed in Table 4.3. We will occasionally refer to the colour octet W sample as the $t\bar{t}$ $cflip$ sample.

Table 4.3

Simulation samples for the colour octet W boson. We quote the cross section used to normalise the sample in the analysis.

Process	**Dataset**	σ **[pb]**
	Background	
Colour octet W boson	TT_TuneCUETP8M2T4_13TeV-powheg-colourFlip-pythia8	832

Based on differences between data and simulated events different sets of corrections are applied to the latter:

1) pile-up re-weighting,

2) lepton identification and isolation efficiency,

3) trigger efficiency,

4) generator level weights,

5) jet energy scale and resolution,

6) b tagging efficiency.

5. Event Selection

The goal of event selection is to separate signal from background. Separate selection is applied to detector level MC events and generator level MC events. Simulated events are tagged as passing only the reconstruction-based, only the particle-based or both selections. The selection for data is that of the detector level MC events.

The discussion of this section is adapted from [41], which uses a similar event selection.

5.1 Detector level

The event selection is based on the $t\bar{t} \rightarrow$ lepton + jets decay topology where one of the W bosons decays to a charged lepton ($\ell = e, \mu$) and a corresponding neutrino, while the other W boson decays to quarks yielding jets.

The particle flow PF algorithm is used for reconstruction of final state objects [2]. This algorithm combines signals from all sub-detectors to enhance the reconstruction performance and it allows to identify muons, electrons, photons, charged hadrons and neutral hadrons produced after a pp collision.

Data samples are collected using the single lepton trigger paths of the High Level Trigger summarised in Table 5.1.

Table 5.1

Trigger paths used for online selection in the analysis.

Final state	Path	Run range	Function	L1 seed
e + jets	HLT_Ele32_eta2p1_WPTight_Gsf_v	all	Select e with $\|\eta\| < 2.1$ and $p_{\mathrm{T}} > 32$ GeV with the tight working point and using the GSF to reconstruct tracks	L1_SingleEG40 OR L1_SingleIsoEG22er OR L1_SingleIsoEG24er OR L1_SingleIsoEG24 OR L1_SingleIsoEG26
μ + jets	HLT_IsoMu24_v	all	Select isolated μ with $p_{\mathrm{T}} > 20$ GeV using L3 tracker algorithm	L1_SingleMu18
	HLT_IsoTkMu24_v	all	Select isolated μ with $p_{\mathrm{T}} > 20$ GeV using HLT tracker muon algorithm	

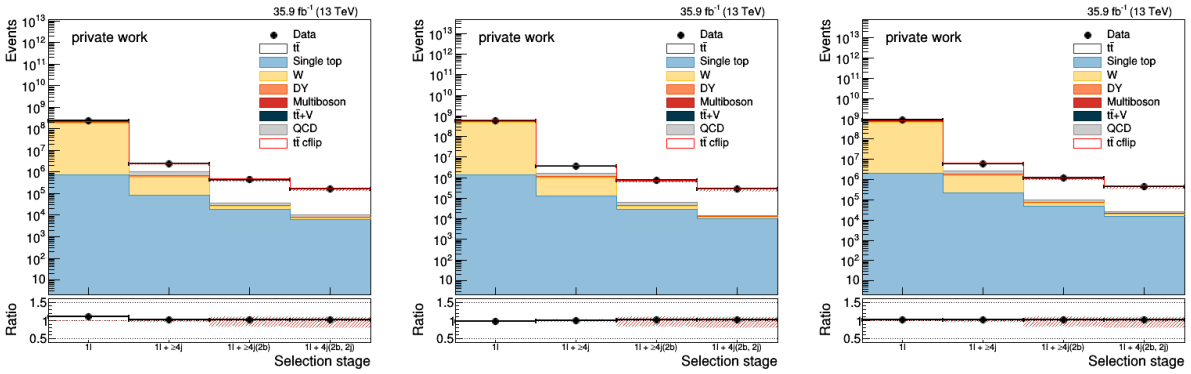
Offline, we require exactly one tight electron/muon with $p_{\mathrm{T}} > 34/26\,\mathrm{GeV}$ and $|\eta| < 2.1/2.4$. The tight working point allows to identify an electron/muon when it is really an electron/muon, important in a high background environment. The event is vetoed in the presence of a second loose lepton with $p_{\mathrm{T}} > 15\,\mathrm{GeV}$ and $|\eta| < 2.4$.

The events are required to have in addition four jets clustered with the anti-k_{T} algorithm with jet separation $R = 0.4$ and charged hadron subtraction (we use shorthand AK4PFchs) with $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 2.4$. The motivation for selecting high p_{T} physics objects is that the detector efficiency drops at low p_{T}.

At least two jets are required to be b-tagged by the Combined Secondary Vertex algorithm (CSVv2) medium working point.

At least two untagged (light) jets are required to yield a W boson candidate with an invariant mass $|m_{jj} - 80.4| < 15$ GeV.

The event yields at different selection stages are shown in Fig. 5.1 and Table 5.2. Table 5.3 shows the event yields for the colour octet W sample. The estimated fraction of the signal increases from 0.1 % in the initial selection stage to 94.2 % at the final selection stage – this is a measure of signal purity of our selection.



(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

Figure 5.1. Event yields at different stages of selection: 1ℓ, $1\ell+ \geq 4j$, $1\ell+ \geq 4j(2b)$, $1\ell+ \geq 4j(2b, 2lj)$.

Table 5.2

Event yields.

Process	$\mathbf{1\ell}$	$\mathbf{1\ell+ \geq 4j}$	$\mathbf{1\ell+ \geq 4j(2b)}$	$\mathbf{1\ell + 4j(2b, 2lj)}$
		e + jets channel		
$t\bar{t}$	3 227 135.8	1 291 471.9	390 503.9	155 959.6
Single top	752 333.9	82 315.0	18 364.4	6154.0
W	185 518 128.0	508 317.5	7024.2	1124.9
DY	14 212 676.0	96 568.3	2219.6	546.8
Multiboson	371 750.3	15 821.5	213.3	50.4
$t\bar{t}$ +V	3070.7	2270.4	631.1	211.7
QCD	20 087 832.0	361 764.0	8147.7	2445.9

Table 5.2

Continued.

Total MC	224 172 928.0	2 358 528.5	427 104.2	166 493.3
($t\bar{t}$ uncertainty)	±459 956.0	±89 350.9	±64 709.8	±27 916.7
Data	246 644 432.0	2 411 742.0	436 962.0	169 786.0
μ + jets channel				
$t\bar{t}$	5 461 582.0	2 130 612.0	645 764.0	258 003.4
Single top	1 353 846.9	131 111.2	29 449.2	9769.0
W	495 234 272.0	860 706.3	13 434.8	2545.6
DY	32 335 596.0	123 329.8	2576.6	572.5
Multiboson	730 682.8	25 606.6	298.8	77.0
$t\bar{t}$ +V	5072.9	3752.8	1052.4	339.5
QCD	77 839 968.0	425 565.2	14 498.0	1428.3
Total MC	612 961 024.0	3 700 684.0	707 073.7	272 735.3
($t\bar{t}$ uncertainty)	±489 351.3	±145 160.4	±107 229.4	±46 849.0
Data	602 190 016.0	3 728 318.0	714 552.0	277 952.0
Combined ℓ + jets channel				
$t\bar{t}$	8 688 431.0	3 422 067.2	1 036 266.9	413 962.9
Single top	2 106 152.8	213 426.9	47 813.6	15 923.0
W	680 901 824.0	1 369 021.4	20 459.0	3670.5
DY	46 548 016.0	219 898.0	4796.1	1119.3
Multiboson	1 102 146.8	41 428.2	512.1	127.3
$t\bar{t}$ +V	8143.4	6023.3	1683.5	551.2
QCD	97 927 816.0	787 329.3	22 645.7	3874.2
Total MC	837 282 560.0	6 059 194.5	1 134 177.0	439 228.5
($t\bar{t}$ uncertainty)	±673 762.4	±233 343.6	±171 857.7	±74 680.0
Data	848 834 496.0	6 140 060.0	1 151 514.0	447 738.0

Table 5.3

Event yields for the colour octet W sample.

Process	$\mathbf{1\ell}$	$\mathbf{1\ell + \geq 4j}$	$\mathbf{1\ell + \geq 4j(2b)}$	$\mathbf{1\ell + 4j(2b, 2lj)}$
e + jets channel				
$t\bar{t}$ cflip	3 286 355.8	1 402 535.2	435 571.4	171 860.5
($t\bar{t}$ cflip uncertainty)	±457 826.7	±45 171.2	±47 044.9	±17 862.5
μ + jets channel				
$t\bar{t}$ cflip	5 498 127.0	2 291 341.8	715 210.2	279 979.4
($t\bar{t}$ cflip uncertainty)	±488 185.1	±73 663.0	±76 578.6	±28 351.2
Combined ℓ + jets channel				
$t\bar{t}$ cflip	8 784 295.0	3 693 872.2	1 150 780.1	451 839.8
($t\bar{t}$ cflip uncertainty)	±669 284.6	±117 188.9	±123 525.4	±46 115.0

5.2 Generator level

In the simulation, the offline selection is mimicked at particle level using the PseudoTopProducer tool [47], using a common lepton selection for both electrons and muons of $p_{\mathrm{T}} > 26\,\mathrm{GeV}$ and $|\eta| < 2.4$, and otherwise jet p_{T}/η ($p_{\mathrm{T}} > 30$ GeV, $|\eta| < 2.4$) and W mass requirements ($|m_{jj} - 80.4| < 15$ GeV) identical to the offline selection.

Charged leptons stemming from the hard process are dressed with nearby photons in a $R = 0.1$ cone, and jets are clustered with the anti-k_{T} algorithm with $R = 0.4$ cone after removing the dressed leptons as well as all neutrinos. In order to identify the flavour of the jet at particle level, "ghost" B hadrons are included in the clustering after scaling their momentum by 10^{-20} in order that they do not change significantly the jet energy scale at particle level.

6. Systematic Uncertainties

Uncertainties are divided into experimental and theoretical uncertainties. When including an uncertainty from the first group we vary some parameter in the event selection, such as a data-to-MC scale factor. Theoretical uncertainties reflect our lack of knowledge about the real world, e.g. the true top quark mass or details of the hadronisation process.

The discussion of this section is adapted from [40] and [41] as these studies use a similar set of systematics.

6.1 Experimental uncertainties

Pile-up. Although pile-up is included in the simulation, there is an intrinsic uncertainty in modelling it appropriately. To estimate the effect of mismodelling the pile-up we vary the average pile-up scenario, through the choice of the minimum bias cross section parameter, by 5 % with respect to its initial estimate.

Trigger and selection efficiency. The uncertainty on the trigger efficiency and on the lepton identification and isolation efficiency scale factors are propagated by re-weighting the simulation after shifting the nominal values up or down. The uncertainty on the muon tracker efficiency is included in this category and added in quadrature, although its effect is expected to be negligible. The impact on the rate is fully absorbed by normalising the distributions in the end, and only the impact on the shape (by weighting more/less some events) is relevant in this analysis.

Jet energy resolution. We use the recommended jet energy resolution measurements [48]. Each jet is further smeared up or down depending on its p_{T} and η with respect to the central value measured in data. The main effect of this systematic is related to the exclusion/inclusion of events with jets near the offline thresholds.

Jet energy corrections. A p_{T}, η-dependent parameterization of the jet energy scale is used to vary the calibration of the jets in the simulation. The parameterization is provided by the JetMET Physics Objet Group [49] for the Spring16 V3 corrections. The main effect of this systematic is related to the exclusion/inclusion of events with jets near the offline thresholds.

b tagging. The nominal efficiency expected in the simulation is corrected by p_{T}-dependent scale factors provided by the BTV Physics Object Group [50]. Depending on the flavour of each jet, the b tagging decision is updated according to the scale factor measured. The scale factor is also varied according to its uncertainty. The main effect of this systematic is the demotion/promotion of candidate b jets and thus a migration of events used for analysis.

Tracking efficiency. The TRK and MUO Physics Object Groups have derived tracking efficiency scale factors as function of the track η or the reconstructed vertex multiplicity. All these scale factors are run-dependent (BCDEF and GH data-taking periods are separated).

6.2 Theoretical uncertainties

QCD scale choices. : We consider anti-correlated variations of the factorisation and renormalisation scales (μ_R/μ_F) in the $t\bar{t}$ sample, by factors of 0.5 and 2. These variations are saved in the simulated events as alternative sets of weights which are used in the evaluation of this systematic. The envelope of 7 variations (excluding opposite variations of μ_R/μ_F) is considered as a systematic.

EVTGEN. The systematic comes about by using EVTGEN Monte Carlo to simulate the decays of heavy flavour particles, primarily B and D mesons.

Hadroniser choice. The systematic comes about by using HERWIG++ [51] instead of PYTHIA 8.

Top quark mass. The most precise measurement of the top quark mass by CMS yields a total uncertainty of ± 0.49 GeV [52]. We consider however a conservative ± 1 GeV. In the possibility that some of these results are used in the future we would like to avoid that they bias too much to a specific top mass.

PYTHIA tunes. The following PYTHIA tunes are used:

1) matrix Element + Parton Shower matching scheme,
2) parton shower scale,
3) colour reconnection model,
4) Underlying Event (UE) variations.

Table 6.1 summarises the simulation samples from the RunIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6 production used for the theoretical systematics.

Table 6.1

Simulation samples used for systematics [41].

Signal variation	Dataset	σ [pb]
Parton shower scale	TT_TuneCUETP8M2T4_13TeV-powheg-isrup-pythia8	832
	TT_TuneCUETP8M2T4_13TeV-powheg-isrdown-pythia8	832
	TT_TuneCUETP8M2T4_13TeV-powheg-fsrup-pythia8	832
	TT_TuneCUETP8M2T4_13TeV-powheg-fsrup-pythia8	832
Underlying event	TT_TuneCUETP8M2T4up_13TeV-powheg-pythia8	832
	TT_TuneCUETP8M2T4down_13TeV-powheg-pythia8	832
ME-PS matching scale (hdamp)	TT_hdampUP_TuneCUETP8M2T4_13TeV-powheg-pythia8	832
	TT_hdampDOWN_TuneCUETP8M2T4_13TeV-powheg-pythia8	832
Colour reconnection	TT_TuneCUETP8M2T4_erdON_13TeV-powheg-pythia8	832
	TT_TuneCUETP8M2T4_QCDbasedCRTune_erdON_13TeV-powheg-pythia8	832
	TT_TuneCUETP8M2T4_GluonMoveCRTune_13TeV-powheg-pythia8	832
Top mass	TT_TuneCUETP8M2T4_mtop1715_13TeV-powheg-pythia8	832
	TT_TuneCUETP8M2T4_mtop1735_13TeV-powheg-pythia8	832
HERWIG++	TT_TuneEE5C_13TeV-powheg-herwigpp	832

7. Results

7.1 Pull vector

A selected event observed in the CMS detector is displayed in Fig. 7.1, showing the light jets, the b-tagged jets, the charged lepton and the pull vector in the ϕ-y plane in a manner analogous to Fig. 3.1.

A set of base tools CFAT [53] was developed having in mind that the analysis can be implemented both in RIVET and CMSSW. Initial tests were done with RIVET because before the colour octet W samples were developed this procedure provided the only means to generate colour-flipped events.

A more comprehensive analysis with data and simulated events at generator and reconstruction level was implemented in CMSSW version CMSSW_8_0_26_patch1. The plots are rendered with ROOT [54]. The pull vectors were obtained for all observable jets – the leading light jet j_1^W (highest p_T), the second leading light jet j_2^W, the leading hadronic b jet j_1^b and the second leading hadronic b jet j_2^b. In each case it was differentiated whether all jet particles or only charged ones should be included in determining the pull vector. The results are separated into e + jets, μ + jets and combined ℓ + jets channels.

The η component of the pull vector of the leading light jet with all jet components is plotted in Fig. 7.2.

An explanation of how CMS plots are represented is in order. The top plot in Fig. 7.2 shows data and Monte Carlo simulations. Unless otherwise specified the Monte Carlo is in reconstruction level. The blue band shows systematics. Given a systematic with index k we identify it as an upside systematic U_i^k if in bin i the systematic S_i^k exceeds the nominal value N_i. In the opposite case we classify the systematic as a downside systematic D_i^k. The total upside and downside systematic is given as a sum of squares:

$$U_i = \sqrt{\sum_k \left(U_i^k - N_i\right)^2}, \qquad D_i = \sqrt{\sum_k \left(D_i^k - N_i\right)^2}. \tag{7.1}$$

The width of the blue band corresponds to the systematical error calculated as $(U_i+D_i)/2$. It is centred on $N_i + (U_i-D_i)/2$. The same applies to the pink band except that the systematics are normalised to the integral of the signal (such normalised histograms are referred to as shapes). The bottom inset shows the ratio of data to Monte Carlo, as well as systematics and systematics from shapes normalised to Monte Carlo.

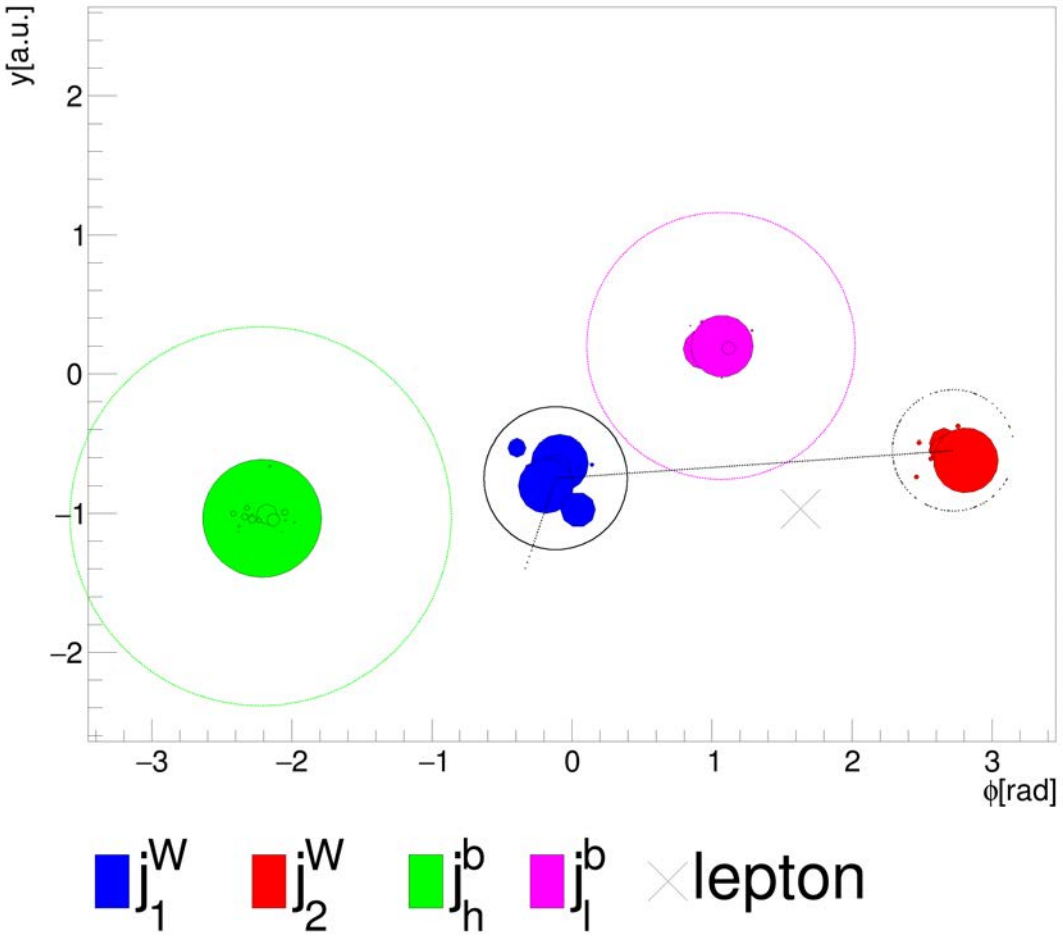
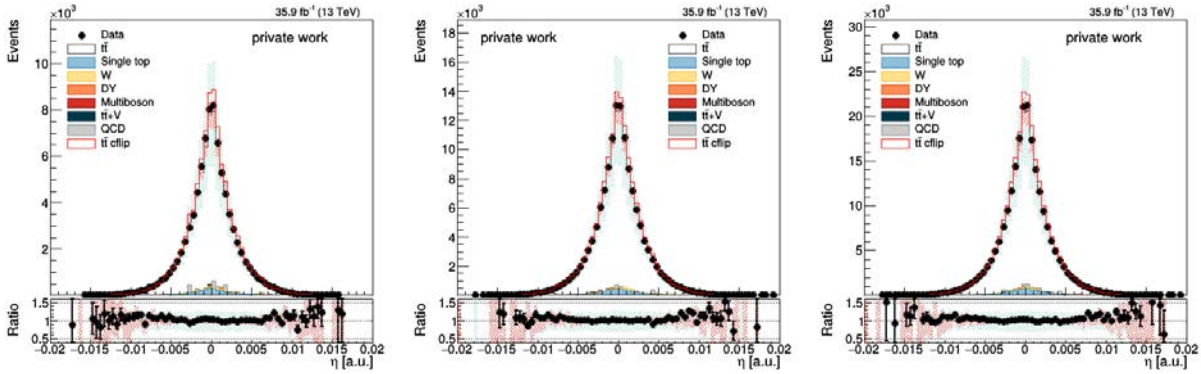


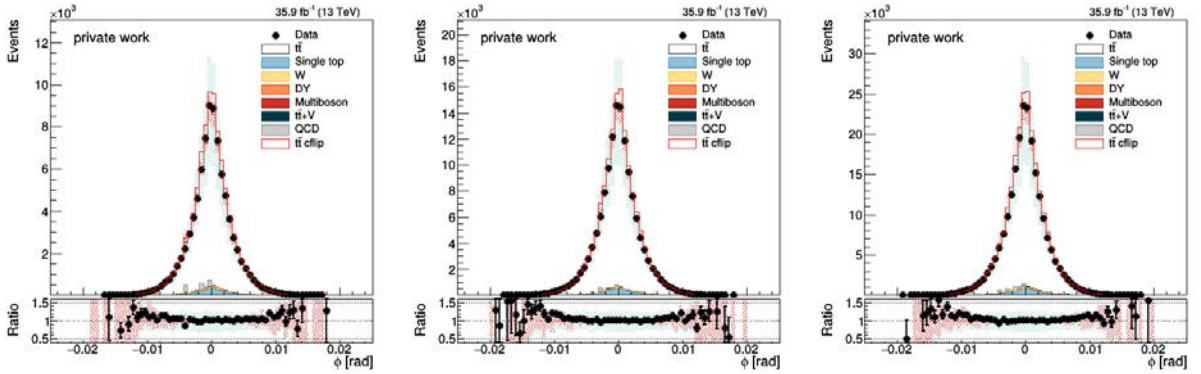
Figure 7.1. An event observed with the CMS detector. Pull vector (dash-dotted) of the leading light jet forming a pull angle of 1.96 rad with the difference between the second leading light jet and the leading light jet (dashed). Constituents of the leading light jet are marked in blue while the constituents of the second leading light jet are marked in red. The leading light jet is marked with a solid line while the second leading light jet is marked with a dotted line. The pull vector is enhanced 200 times, while the radius of the circles representing jets is equal to $p_{\mathrm{T}}\ [\mathrm{GeV}]/75.0$ and the radius of the circles representing constituents is equal to $p_{\mathrm{T}}^{\mathrm{constituent}}/p_{\mathrm{T}}^{\mathrm{jet}}$. The hadronic b jet and its constituents are marked in green, while the leptonic b jet and its constituents are marked in magenta. Also shown with "$\times$" is the charged lepton.



(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

Figure 7.2. Distribution of the η component of the pull vector of j_1^W with all jet components.

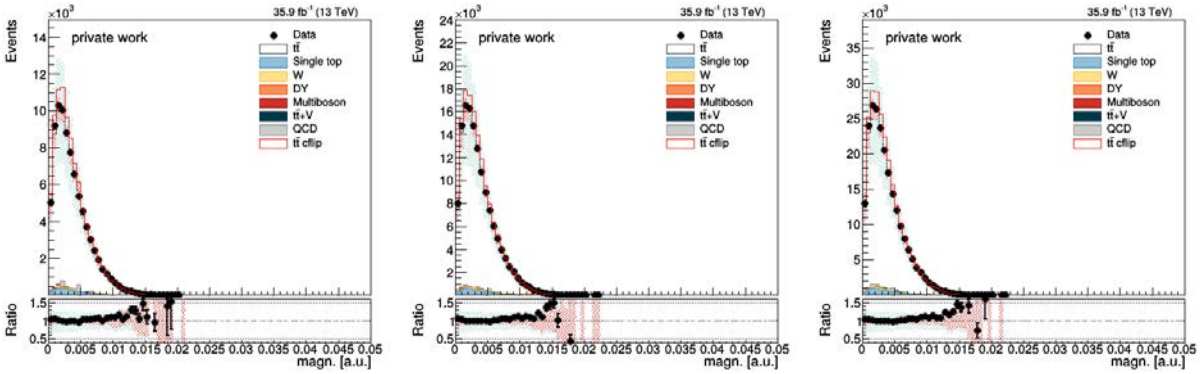
The ϕ component of the pull vector of the leading light jet with all jet components is plotted in Fig. 7.3.



(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

Figure 7.3. Distribution of the ϕ component of the pull vector of j_1^W with all jet components.

The magnitude of the pull vector of the leading light jet with all jet components is given in Fig. 7.4. The magnitude of the pull vector is usually contained below 0.02 [a.u.].

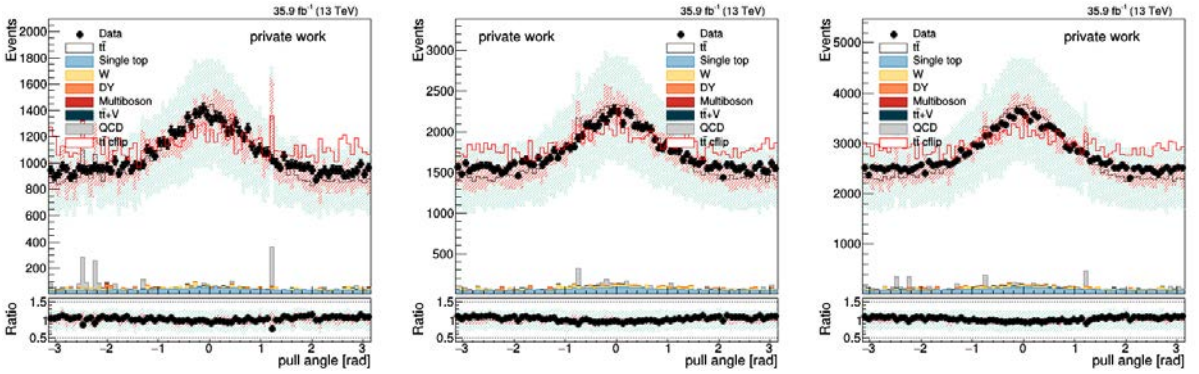


(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

Figure 7.4. Distribution of the magnitude of the pull vector of j_1^W with all jet components.

7.2 Pull angle

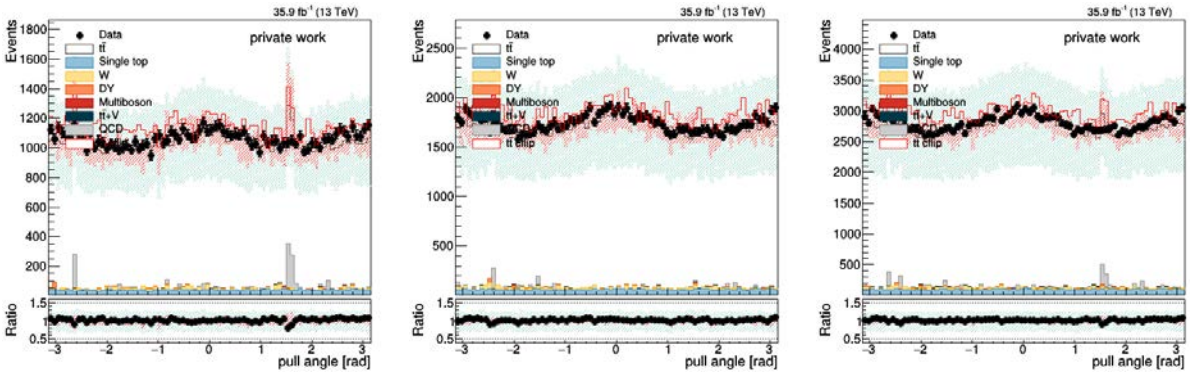
The plots of the pull angle between colour-connected jets – from j_1^W to j_2^W with all jet constituents and including all values of ΔR are shown in Fig. 7.5.



(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

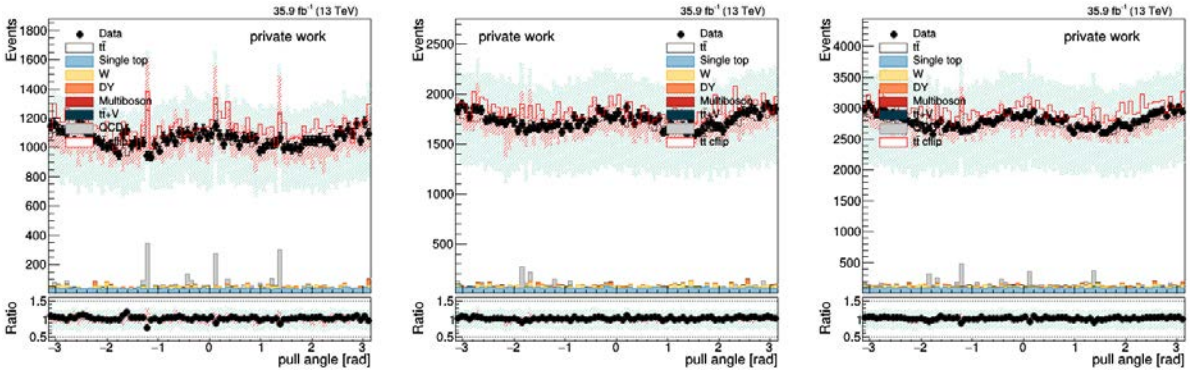
Figure 7.5. Distribution of the pull angle from j_1^W to j_2^W for all ΔR and including all particles.

Additionally, the plots of the pull angle between jets where we expect no colour connection – from j_1^b to j_2^b and from j_2^b to j_1^b with all jet constituents and including all values of ΔR are shown in Fig. 7.6 and Fig. 7.7.



(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

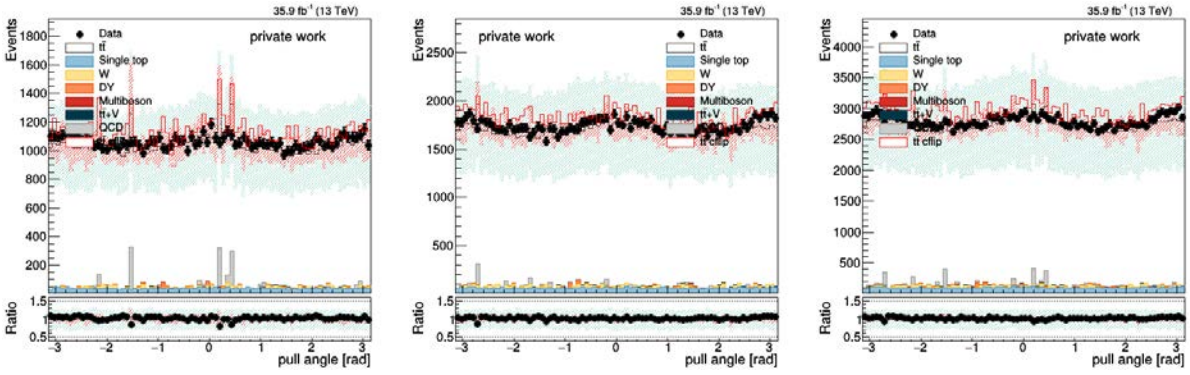
Figure 7.6. Distribution of the pull angle from j_1^b to j_2^b for all ΔR and including all particles.



(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

Figure 7.7. Distribution of the pull angle from j_2^b to j_1^b for all ΔR and including all particles.

Another chance to look at the distribution of pull angle between objects that are not colour-connected is to choose a jet and a lepton. Fig. 7.8 shows the distribution of pull angle from j_1^W to the charged lepton.

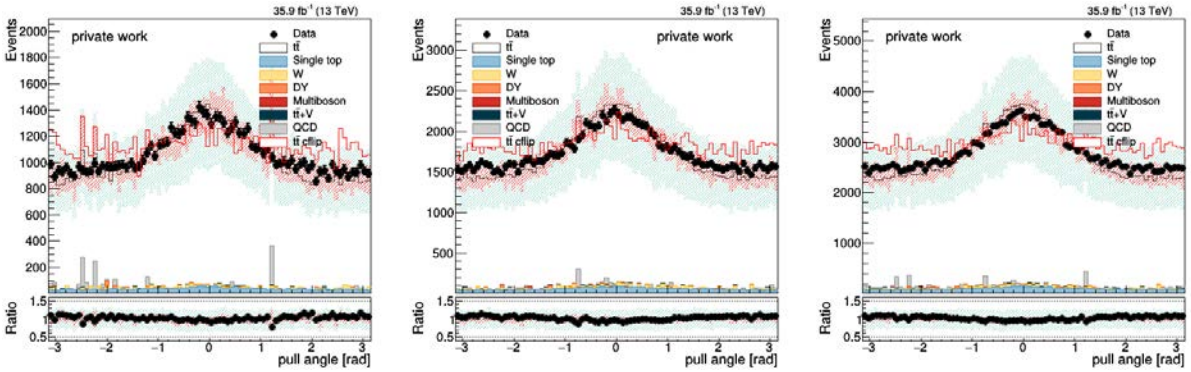


(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

Figure 7.8. Distribution of the pull angle from j_1^W to the charged lepton for all ΔR and including all particles.

As can be readily observed, the central peak in the distribution of the pull angle is prominent in case of colour-connected jets and flattens out in the case of objects that are not colour-connected.

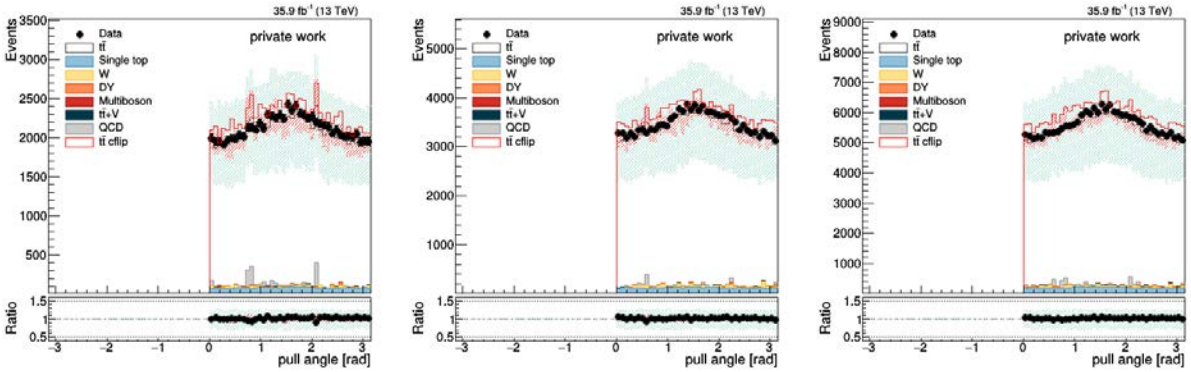
The central peak can reappear in the case of collinearities of the vectors of physics objects even though they are not colour-connected. Such a case is seen in the distribution of the pull angle of j_1^W to hadronic W (Fig. 7.9). .



(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

Figure 7.9. Distribution of the pull angle from j_1^W to the hadronic W for all ΔR and including all particles.

Another interesting case is choosing the beam. In Fig. 7.10 we show the distribution of θ_{p} of j_1^W to the positive direction of the beam. We see a peak at a right angle.



(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

Figure 7.10. Distribution of the pull angle from j_1^W to the positive direction of the beam including all particles.

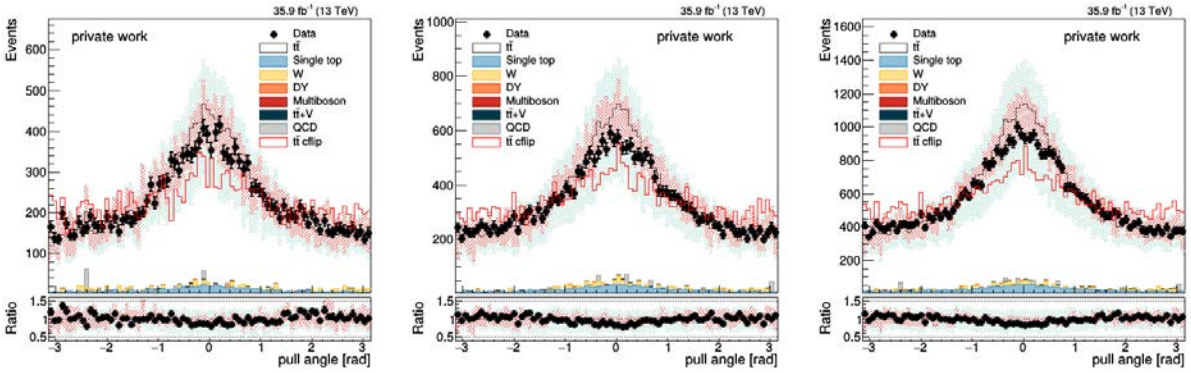
The QCD samples contribute peaks to the plots because only a few QCD events pass the selection criteria, but they are assigned a large weight. Each event gets effectively assigned a weight

$$w = \mathcal{L} \cdot \sigma \frac{1}{N_{\mathrm{gen}}}. \tag{7.2}$$

The cross section σ for QCD events is very large but the number of generated MC events N_{gen} is very low. Therefore a few QCD events represent an entire distribution.

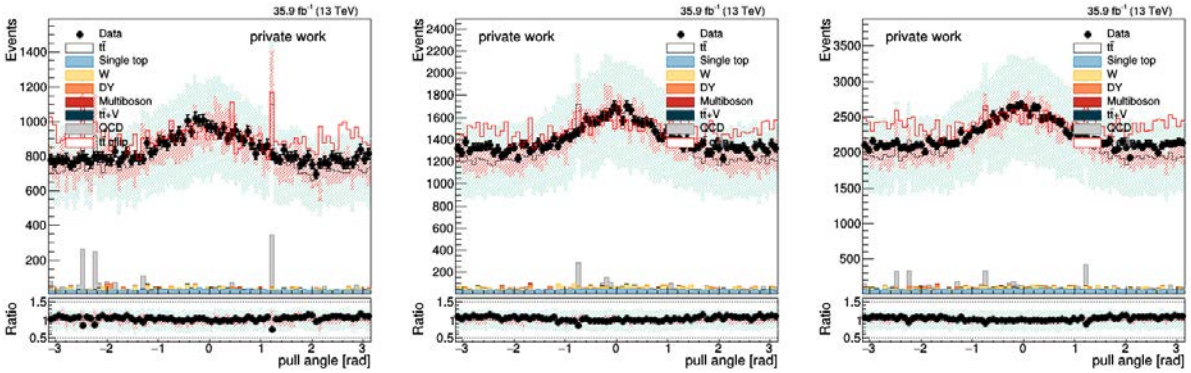
7.3 ΔR bias

When two jets are close to each other in ϕ-η space, the jet clustering algorithm is inclined to associate particles of one jet (lowest p_{T} jet) to another (highest p_{T} jet). This effect creates a bias in the pull angle analysis as the pull vector is more likely to point to the jet from which the particles were weaned. Figs. 7.11–7.12 illustrates the distribution of pull angle for two cases – closely spaced jets with $\Delta R \leq 1.0$ and well separated jets with $\Delta R > 1.0$.



(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

Figure 7.11. Distribution of the pull angle with $\Delta R \leq 1.0$ and including all jet constituents from j_1^W to j_2^W.



(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

Figure 7.12. Distribution of the pull angle with $\Delta R > 1.0$ and including all jet constituents from j_1^W to j_2^W.

7.4 Sensitivity analysis

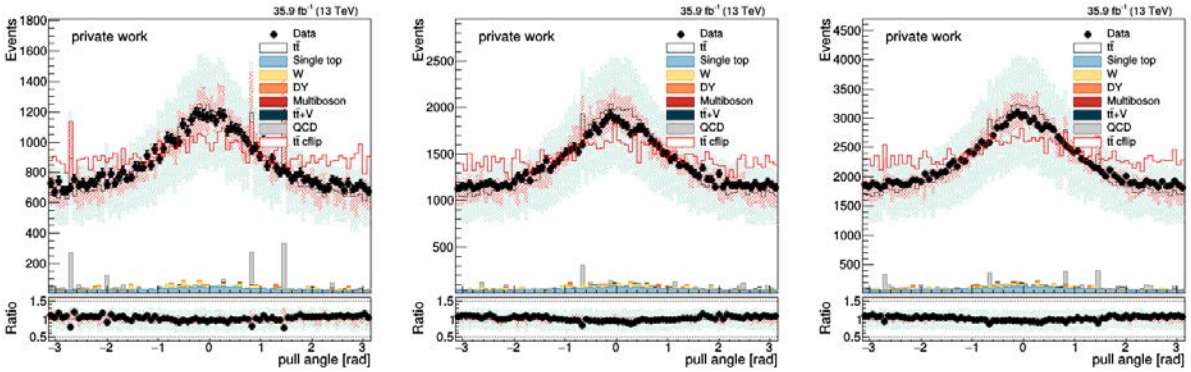
Sensitivity of the pull angle methodology was studied by applying cuts to the following parameters:

p_{T} of the hadronic W boson. A cut was chosen at 50 GeV and the distribution of the pull angle was obtained at p_{T} of the hadronic W boson greater than and less than or equal to this value. The results are shown in Figs. 7.13–7.14.

Number of jet constituents. A cut was chosen at the number of jet constitutents N being 20 and the distribution of the pull angle was obtained at N greater than and less than or equal to this value. The results are shown in Figs. 7.15–7.16.

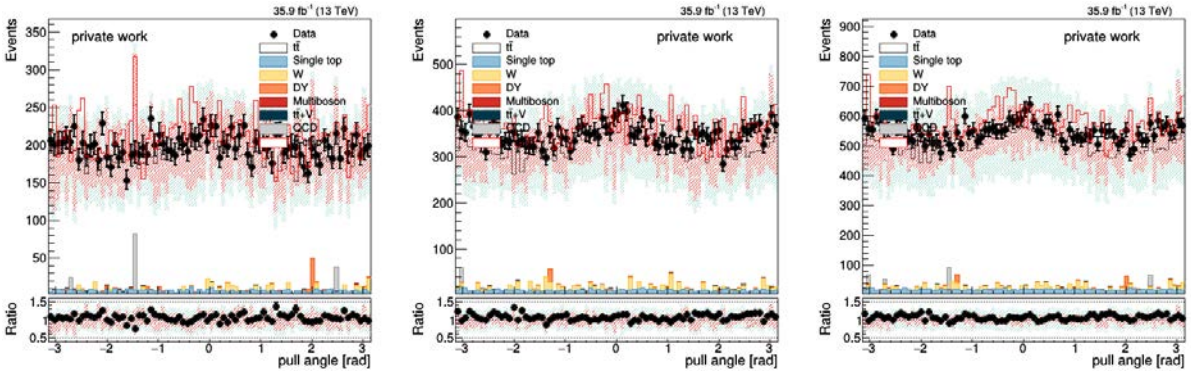
p_{T} of jet constituents. A cut was chosen at p_{T} of the jet constituents being 0.5 GeV and the distribution of the pull angle was at obtained at p_{T} of the jet constituents being greater than and less than or equal to this value. The results are shown in Figs. 7.17–7.18.

Magnitude of the pull vector. A cut was chosen at the magnitude of the pull vector of 0.005 [a.u.] and the distribution of the pull angle was obtained at the magnitude of the pull vector being greater than and less than or equal to this value. The results are shown in Figs. 7.19–7.20.



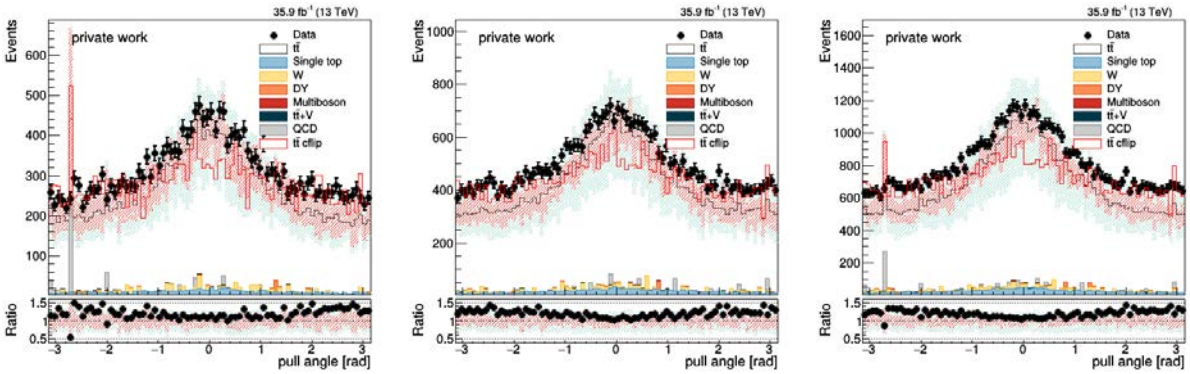
(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

Figure 7.13. Distribution of the pull angle for all ΔR and all particles from j_1^W to j_2^W with p_{T} of $W > 50$ GeV.



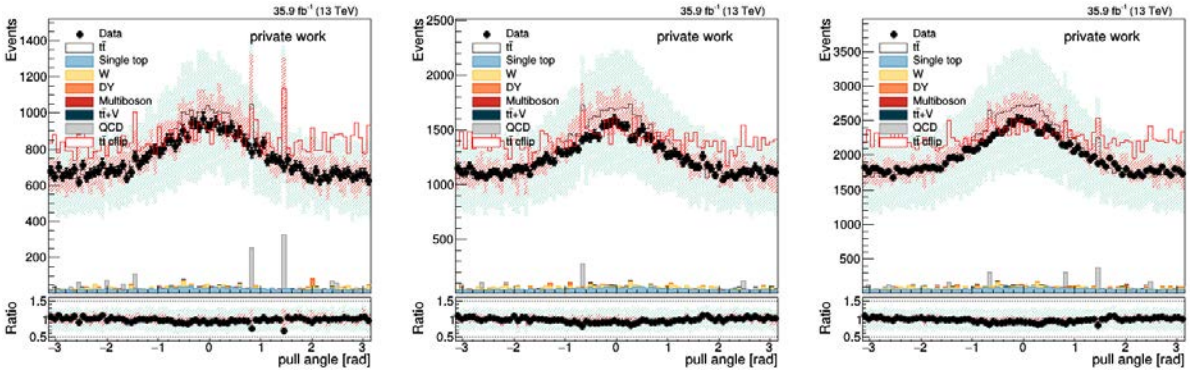
(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

Figure 7.14. Distribution of the pull angle for all ΔR and all particles from j_1^W to j_2^W with p_{T} of $W \leq 50$ GeV.



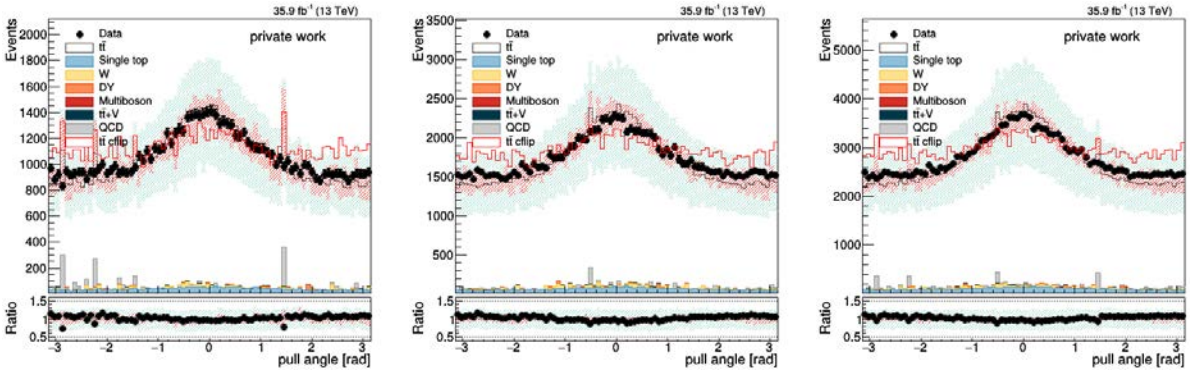
(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

Figure 7.15. Distribution of the pull angle for all ΔR and all particles from j_1^W to j_2^W with the number of jet constituents $N > 20$.



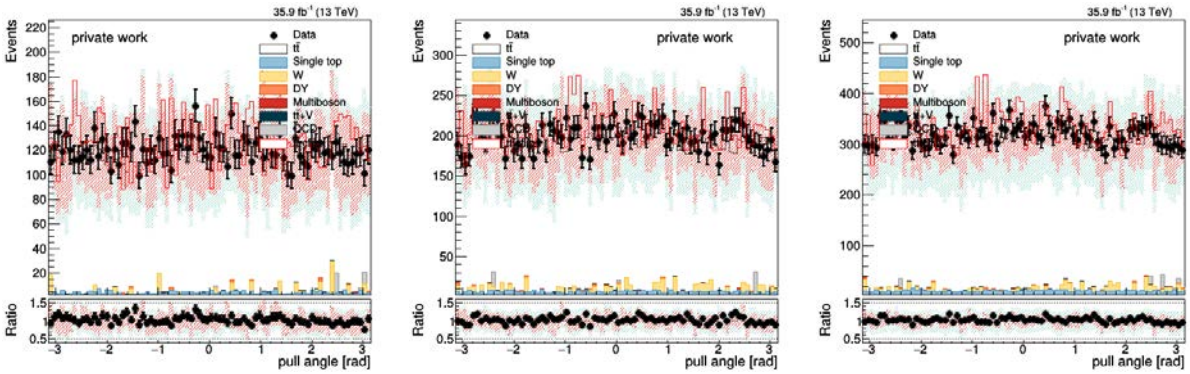
(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

Figure 7.16. Distribution of the pull angle for all ΔR and all particles from j_1^W to j_2^W with the number of jet constituents $N \leq 20$.



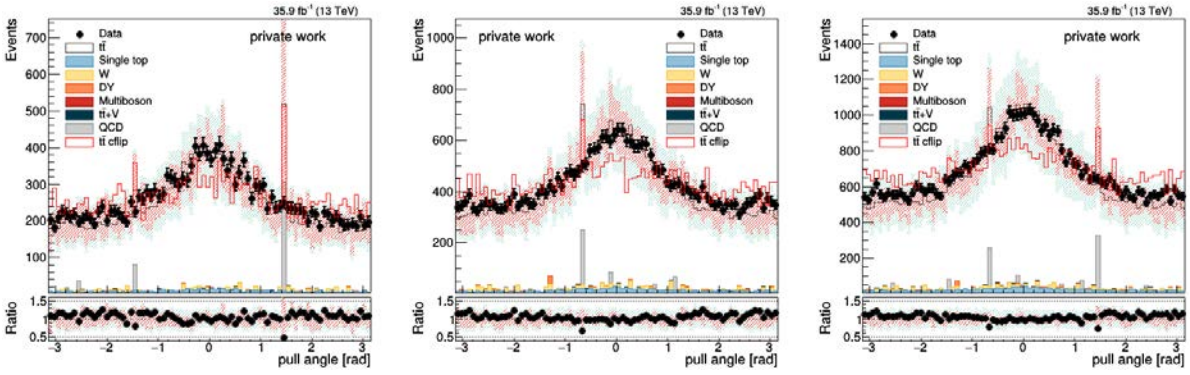
(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

Figure 7.17. Distribution of the pull angle for all ΔR and all particles from j_1^W to j_2^W with the p_{T} of jet constituents > 0.5 GeV.



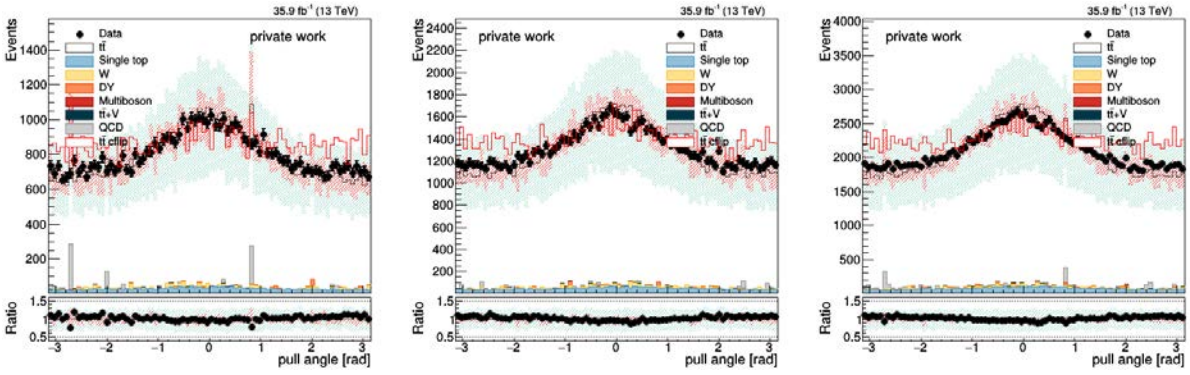
(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

Figure 7.18. Distribution of the pull angle for all ΔR and all particles from j_1^W to j_2^W with the p_{T} of jet constituents ≤ 0.5 GeV.



(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

Figure 7.19. Distribution of the pull angle for all ΔR and all particles from j_1^W to j_2^W with the magnitude of the pull vector > 0.005 [a.u.].



(a) e + jets channel. (b) μ + jets channel. (c) Combined ℓ + jets channel.

Figure 7.20. Distribution of the pull angle for all ΔR and all particles from j_1^W to j_2^W with the magnitude of the pull vector ≤ 0.005 [a.u.].

It was observed that the pull angle methodology is sensitive to p_{T} of hadronic W boson, number of jet constituents, p_{T} of jet constituents but not particularly sensitive to the magnitude of the pull vector.

7.5 Unfolding

The class TUnfoldDensity[55] of ROOT is used to do the unfolding procedure. The binning scheme is managed with class TUnfoldBinning. No regularisation is applied. We are interested to have the migration matrix as diagonal as possible to reduce statistical uncertainties on the unfolding result. Two measures are used to characterise the share of statistics on the diagonal of the migration matrix – stability and purity. Stability is the ratio of the contents of the diagonal element to the total number of events at reconstruction level in the bin:

$$\text{stability} \equiv \frac{\theta^{\text{diag}}_{\text{input}}}{\sum_{x=1}^{x=N_x} \theta^{x}_{\text{input}}}, \tag{7.3}$$

where x is the bin index at reconstruction level, starting the numbering from 1 and N_x is the number of bins at reconstruction level. Purity is the ratio of the contents of the diagonal element to the total number of events at generation level in the bin:

$$\text{purity} \equiv \frac{\theta^{\text{diag}}_{\text{input}}}{\sum_{y=1}^{y=N_y} \theta^{y}_{\text{input}}}, \tag{7.4}$$

where we have used y as the bin index at generation level. The values of purity and stability are recommended to exceed 50 % at each bin.

An interesting measure is the amount by which the unfolded result is different from the generated result at MC (an ideal result would be 0), normalised to statistical uncertainty of the unfolded result. This measure is called the pull:

$$\text{pull} \equiv \frac{\theta^{\text{gen}}_{\text{unf}} - \theta^{\text{gen}}_{\text{in}}}{\sigma^{\text{gen}}_{\text{unf}}}. \tag{7.5}$$

We generate random toy distributions of the observable at generation level, thus obtaining a distribution of the pull.

The unfolding results with 3 regular-sized bins are shown in Fig. 7.21. Distributions corresponding to unfolding results with migration matrices from $t\bar{t}\ Herwig++$ and $t\bar{t}\ cflip$ as well as systematics $t\bar{t}\ fsr\ dn$ and $t\bar{t}\ fsr\ up$ (see Chap. 6) are laid over the unfolding plots. The stability and purity levels with this binning scheme reach acceptable levels at each bin and it was adopted for further analysis. In order to create the plots shown herein a new class CompoundHistoUnfolding [56] was developed and it was added to ROOT complete with input and output streamers.

The bin-per-bin significance (%) of nuisances in the total systematical error in the unfolded result is given in Table 7.1. Nuisances that directly affect the hadronisation $t\bar{t}\ Herwig++$, $t\bar{t}\ QCDbased$ and $t\bar{t}\ ERDon$ are the most significant. Table 7.2 shows the bin-per-bin uncertainties including the $t\bar{t}\ cflip$ sample as a systematic to $t\bar{t}$.

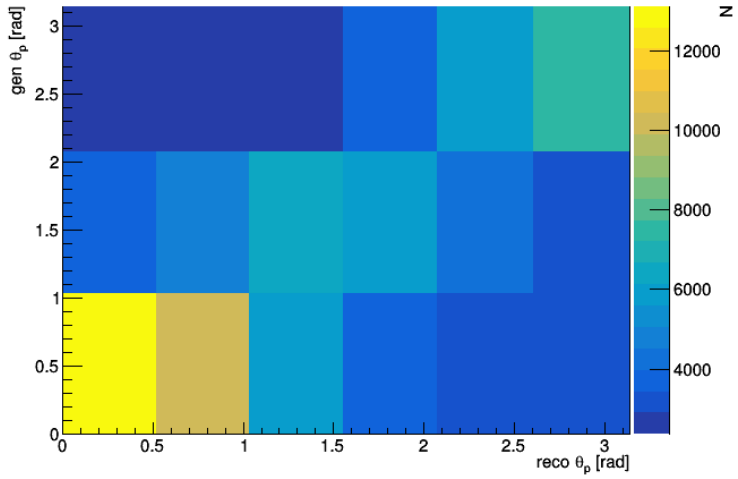
The agreement between the unfolded result and MC prediction at generation level is quantified using a goodness of fit method. Given the normalised unfolded detector observation D, the normalised MC prediction M, the full covariance matrix Σ of normalised experimental uncertainties, the χ^2 is calculated as follows:

$$\chi^2 = (D^T - M^T) \cdot \Sigma^{-1} \cdot (D - M). \tag{7.6}$$

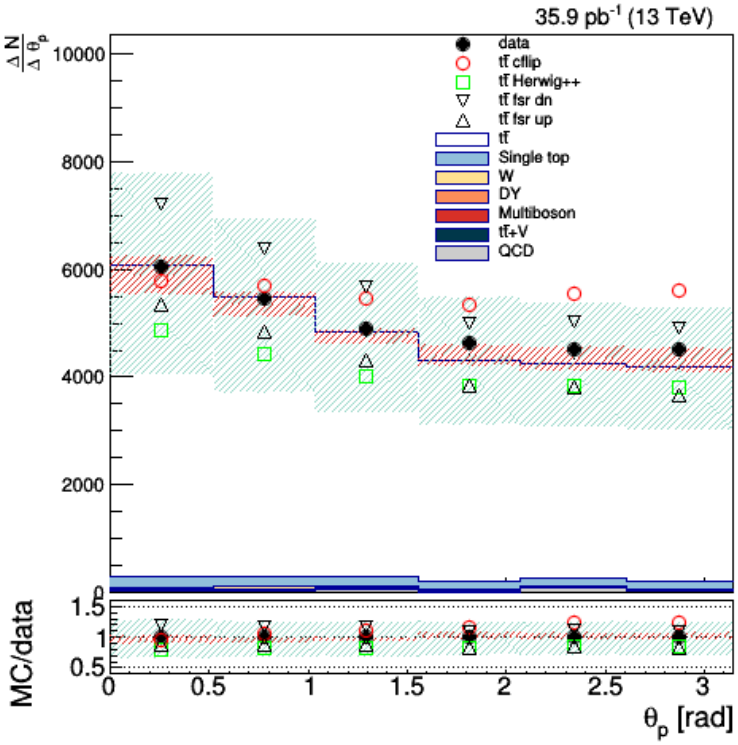
From the χ^2 value the p-values can be computed using the number of degrees of freedom equal to the number of bins in the unfolded distribution subtracted by 1 to account for a loss of freedom when normalising the distributions. One row and one column is discarded from the covariance matrix Σ. χ^2 value does not depend on the choice of the discarded elements.

Table 7.3 shows the χ^2 values and p-values for θ_{p} using all jet constituents. The results show that the pull angle distribution is poorly modelled by the MC generators. In general, the simulation predicts a more sloped distribution, i.e. a stronger colour flow effect. HERWIG++ models better the pull angle distribution than PYTHIA 8.2. Accuracy of PYTHIA 8.2 is particularly poor when predicting the distribution of θ_{p} from j_2^W to j_1^W.

The χ^2 values and p-values for the W colour octet model are given in Table 7.4. In the POWHEG +PYTHIA 8 * entry $t\bar{t}\ cflip$ has been added as a systematic to $t\bar{t}$. In the colour flip model the distribution of θ_{p} from j_1^W to j_2^W is modelled less acurately than the SM prediction.

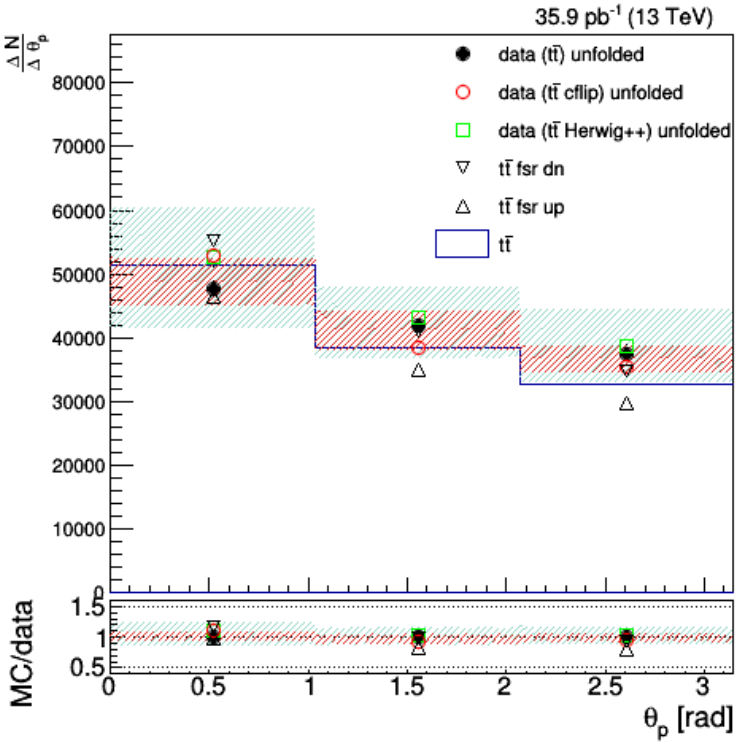


(a) Migration matrix with the reconstruction level as the x axis and the generation level as the y axis.

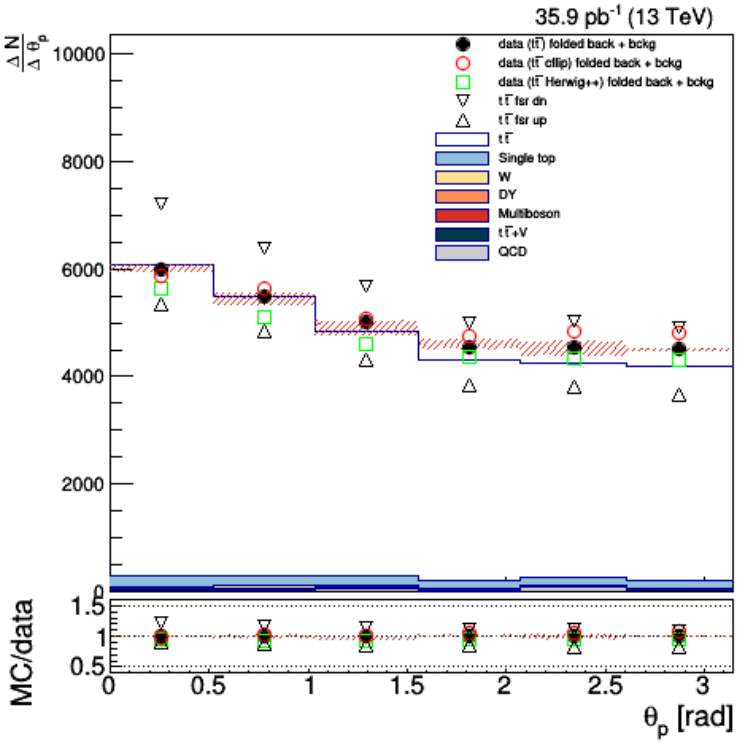


(b) The data and Monte Carlo used as input.

Figure 7.21. Unfolding plots for the $t\bar{t}$ method of the pull angle θ_{p} of j_1^W to j_2^W including all jet constituents using 3 regularly sized bins.

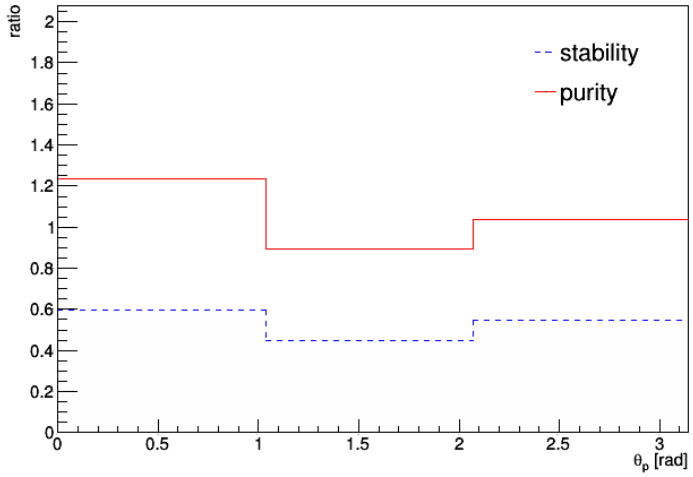


(c) The unfolded result.

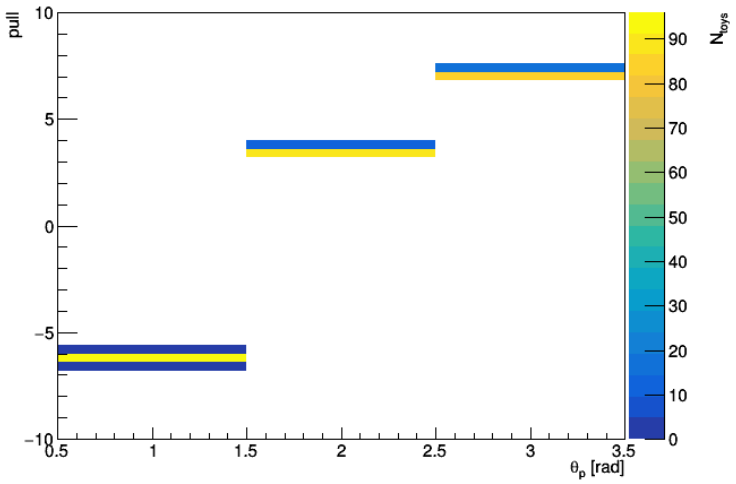


(d) The folded back result.

Figure 7.21. Unfolding plots for the $t\bar{t}$ method of the pull angle θ_{p} of j_1^W to j_2^W including all jet constituents using 3 regularly sized bins (continued).



(e) Stability and purity in each bin.



(f) The pull.

Figure 7.21. Unfolding plots for the $t\bar{t}$ method of the pull angle θ_{p} of j_1^W to j_2^W including all jet constituents using 3 regularly sized bins (continued).

Table 7.1
Bin by bin weight of uncertainty (shapes) for pull angle θ_{p} including all jet constituents from the j_1^W to the j_2^W at the generator level. The results are for unfolded ouput for the POWHEG +PYTHIA 8 sample. The binning method of 3 regularly sized bins is used.

Nuisance	**Uncertainty in bins [%]**		
	bin 1	**bin 2**	**bin 3**
	0.00–1.04	**1.04–2.07**	**2.07–3.14**
$t\bar{t}$ ERDon	2.416	1.669	1.171
$t\bar{t}$ Herwig++	7.151	2.877	5.713
$t\bar{t}$ Peterson Frag	0.323	0.297	0.076
$t\bar{t}$ QCDbased	6.792	5.576	2.329
$t\bar{t}$ UEdn	1.895	1.509	0.701
$t\bar{t}$ UEup	0.112	0.298	0.187
$t\bar{t}$ b fragmentation down	0.031	0.056	0.100
$t\bar{t}$ b fragmentation up	0.056	0.085	0.162
$t\bar{t}$ btag_heavy down	0.215	0.256	0.012
$t\bar{t}$ btag_heavy up	0.197	0.191	0.451
$t\bar{t}$ btag_light down	0.219	0.112	0.392
$t\bar{t}$ btag_light up	0.177	0.251	0.491
$t\bar{t}$ csv_heavy down	0.728	1.260	0.471
$t\bar{t}$ csv_heavy up	0.328	1.230	0.933
$t\bar{t}$ csv_light down	0.113	0.207	0.086
$t\bar{t}$ csv_light up	0.022	0.134	0.120
$t\bar{t}$ evtgen	3.121	3.496	0.054
$t\bar{t}$ fsr dn	2.512	3.451	0.650
$t\bar{t}$ fsr up	1.225	4.823	3.735
$t\bar{t}$ gluon move	0.973	0.345	0.828
$t\bar{t}$ hdamp dn	0.636	0.036	0.747
$t\bar{t}$ hdamp up	0.831	0.673	0.295
$t\bar{t}$ id1002muR1muF2hdampmt272.7225	0.063	0.194	0.133
$t\bar{t}$ id1003muR1muF0.5hdampmt272.7225	0.147	0.145	0.024
$t\bar{t}$ id1004muR2muF1hdampmt272.7225	0.195	0.290	0.074
$t\bar{t}$ id1005muR2muF2hdampmt272.7225	0.089	0.037	0.150
$t\bar{t}$ id1007muR0.5muF1hdampmt272.7225	0.064	0.068	0.005
$t\bar{t}$ id1009muR0.5muF0.5hdampmt272.7225	0.054	0.023	0.092
$t\bar{t}$ isr dn	2.203	2.097	0.443
$t\bar{t}$ isr up	0.373	0.653	0.250
$t\bar{t}$ jec_CorrelationGroupMPFInSitu down	0.245	0.104	0.190
$t\bar{t}$ jec_CorrelationGroupMPFInSitu up	0.029	0.251	0.237
$t\bar{t}$ jec_CorrelationGroupUncorrelated down	0.377	0.247	0.198
$t\bar{t}$ jec_CorrelationGroupUncorrelated up	0.348	0.783	0.422
$t\bar{t}$ jec_FlavorPureBottom down	0.043	0.278	0.249
$t\bar{t}$ jec_FlavorPureBottom up	0.056	0.033	0.105

Table 7.1

Continued.

$t\bar{t}$ jec_FlavorPureCharm down	0.165	0.174	0.394
$t\bar{t}$ jec_FlavorPureCharm up	0.073	0.093	0.191
$t\bar{t}$ jec_FlavorPureGluon down	0.080	0.032	0.133
$t\bar{t}$ jec_FlavorPureGluon up	0.784	1.401	0.555
$t\bar{t}$ jec_FlavorPureQuark down	0.040	0.224	0.194
$t\bar{t}$ jec_FlavorPureQuark up	0.064	0.137	0.229
$t\bar{t}$ jec_RelativeFSR down	0.113	0.218	0.377
$t\bar{t}$ jec_RelativeFSR up	0.050	0.149	0.101
$t\bar{t}$ jer down	0.668	0.711	0.052
$t\bar{t}$ jer up	0.279	0.057	0.283
$t\bar{t}$ m=171.5	0.149	0.917	0.814
$t\bar{t}$ m=173.5	2.187	0.565	2.090
$t\bar{t}$ pileup down	0.009	0.106	0.104
$t\bar{t}$ pileup up	0.104	0.323	0.481
$t\bar{t}$ sel efficiency correction down	0.003	0.003	0.001
$t\bar{t}$ sel efficiency correction up	0.006	0.005	0.002
$t\bar{t}$ semilep BR down	0.013	0.007	0.008
$t\bar{t}$ semilep BR up	0.022	0.034	0.010
$t\bar{t}$ tracking down	0.489	0.187	0.402
$t\bar{t}$ tracking up	0.151	0.175	0.004
$t\bar{t}$ trig efficiency correction down	0.002	0.009	0.008
$t\bar{t}$ trig efficiency correction up	0.002	0.010	0.008

Table 7.2

Bin by bin weight of the additional uncertainty (shape) $t\bar{t}\ cflip$ for pull angle θ_{p} including all jet constituents from the j_1^W to the j_2^W at the generator level. The results are for unfolded ouput for the POWHEG +PYTHIA 8 * sample. The binning method of 3 regularly sized bins is used.

Nuisance	**Uncertainty in bins [%]**		
	bin 1	**bin 2**	**bin 3**
	0.00–1.04	**1.04–2.07**	**2.07–3.14**
$t\bar{t}$ cflip	20.924	11.031	13.870

Table 7.3

χ^2 and p-values of pull angle θ_{p} including all jet constituents. The results are for the SM model.

Sample	$\theta_{\mathrm{p}}(j_1^W, j_2^W)$			$\theta_{\mathrm{p}}(j_2^W, j_1^W)$			$\theta_{\mathrm{p}}(j_1^b, j_2^b)$			$\theta_{\mathrm{p}}(j_2^b, j_2^b)$		
	χ^2	/Ndf	p-value	χ^2	/Ndf	p-value	χ^2	/Ndf	p-value	χ^2	/Ndf	p-value
Powheg + Pythia8	86.38	2	0.000	85.33	2	0.000	3.98	2	0.136	20.03	2	0.000
Powheg + Herwig++	0.30	2	0.861	1.14	2	0.565	3.11	2	0.211	8.05	2	0.018

Table 7.4
χ^2 and p-values of pull angle $\theta_{\rm p}$ including all jet constituents. The results are for the colour octet W model.

Sample	$\theta_{\rm p}(j_1^W, j_2^W)$			$\theta_{\rm p}(j_2^W, j_1^W)$			$\theta_{\rm p}(j_1^b, j_2^b)$			$\theta_{\rm p}(j_2^b, j_2^b)$		
	χ^2	/Ndf	p-value	χ^2	/Ndf	p-value	χ^2	/Ndf	p-value	χ^2	/Ndf	p-value
Powheg + Pythia8 *	1.05	2	0.591	8.85	2	0.012	2.89	2	0.235	19.75	2	0.000
Powheg + Pythia8 cf	3.24	2	0.198	1.12	2	0.570	0.99	2	0.611	1.36	2	0.506

7.6 LEP method

Three types of flows are analysed:

- in particle flow all particles are assigned a weight equal to 1.0,
- in energy flow particles are assigned a weight proportional to their energy normalised to the sum of the energy of the top quarks,
- in $p_{\rm T}$ flow particles are assigned a weight proportional to their transverse momentum normalised to the transverse momentum of the respective jet.

The results of the LEP methodology using particle flow are shown in Fig. 7.22 with all jet constituents and in Fig. 7.23 including only charged jet constituents. The flow is plotted between the leading b jet j_1^b and the 2nd leading b jet j_2^b, the hadronic b jet $j_{\rm h}^b$ and the furthest light quark $j_{\rm f}^W$ (jet distance is measured with the angle between the spatial components of the 4-vectors of the jets), the closest light quark $j_{\rm c}^W$ and the hadronic b jet $j_{\rm h}^b$, and the leading light jet j_1^W and the second leading light jet j_2^W.

In all cases the density drops in the middle area between jets compared to the jet centre with the central density varying between colour-connected jets and jets not connected in colour.

The bin-per-bin ratios of the flow in colour-free regions (j_1^b, j_2^b), ($j_{\rm h}^b$, $j_{\rm f}^W$), ($j_{\rm c}^W$, $j_{\rm h}^b$) to the flow in the colour-connected region (j_1^W, j_2^W) are given in Fig. 7.24 including all jet constituents and Fig. 7.25 including only charged jet constituents. Significant colour reconnection is noticeable in the region ($j_{\rm c}^W$, $j_{\rm h}^b$) assuming the colour octet W model.

Fig. 7.26 shows the bin-per-bin ratio of the particle flow in the region (j_1^W, j_2^W) in the SM model to the particle flow in the region (j_1^W, j_2^W) in the colour octet W model. A loss of colour connection in this region in the colour octet W model is evident.

As a quantitative result from the LEP methodology one can use the parameter R which is defined as the ratio between the integral from 0.2 to 0.8 in the colour-connected region to the integral from 0.2 to 0.8 in the region not connected in colour:

$$R = \frac{\int_{0.2}^{0.8} f^{\text{inter } W \text{ region}} d\chi}{\int_{0.2}^{0.8} f^{\text{intra } W \text{ region}} d\chi}, \tag{7.7}$$

where $f(\chi)$ is the density of the flow distribution.

This parameter was used at LEP to quantify colour connection effects and their values from different experiments corresponding to 625 pb^{-1} of data in the range $\sqrt{s}$ =189–209 GeV are given in Table 7.5. We note inconsistency in the R values reported by different experiments.

Furthermore, R should exceed 1 on theoretical basis. The range 0.2–0.8 is identified as sensitive to colour connection effects.

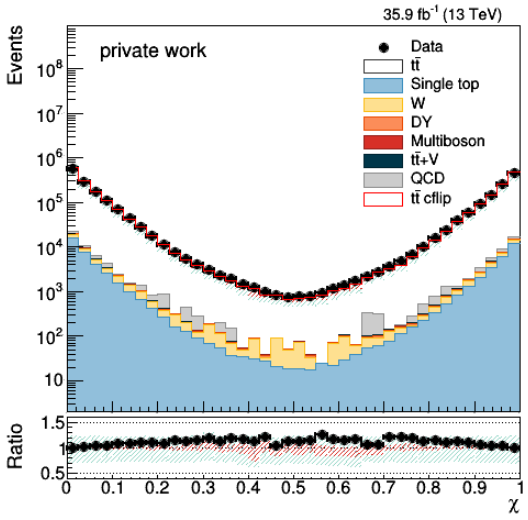
Table 7.5

R values observed at LEP.

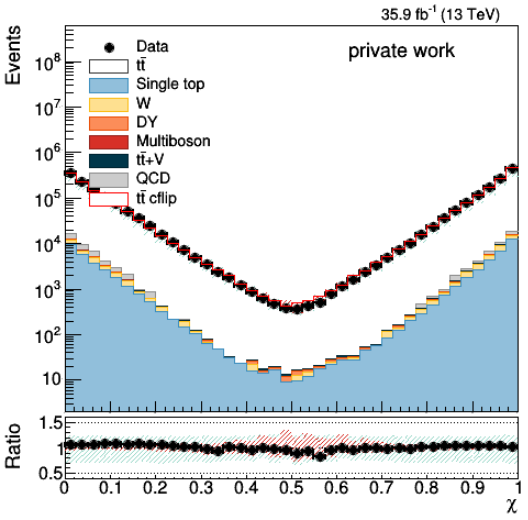
LEP experiment	$\boldsymbol{R}$ **value – data**	**Reference**
OPAL	1.243	[4]
Delphi	0.889 ($\sqrt{s} = 183$ GeV)–1.039 ($\sqrt{s} = 207$ GeV)	[5]
L3	0.911	[6]

In our case we use 3 R values for any of the regions not connected in colour with normalisation to the colour-connected region (j_1^W, j_2^W).

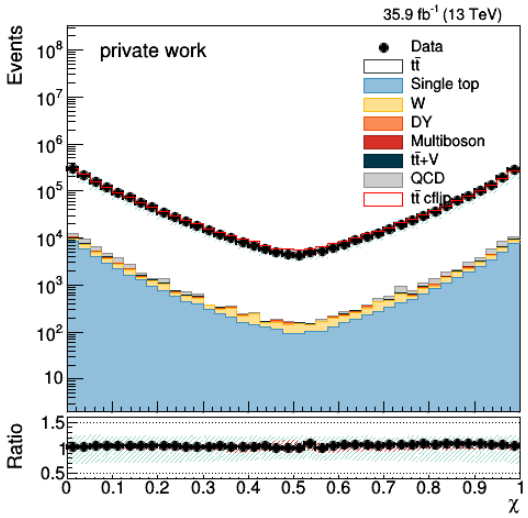
The integral of particle flow from 0.2 to 0.8 in different regions and the inverse of R values for the SM model is given Table 7.6, for data in Table 7.7 and for the W colour octet model in Table 7.8. It is observed that the plane between colour-connected jets is more densely filled than the planes between jets not connected in colour.



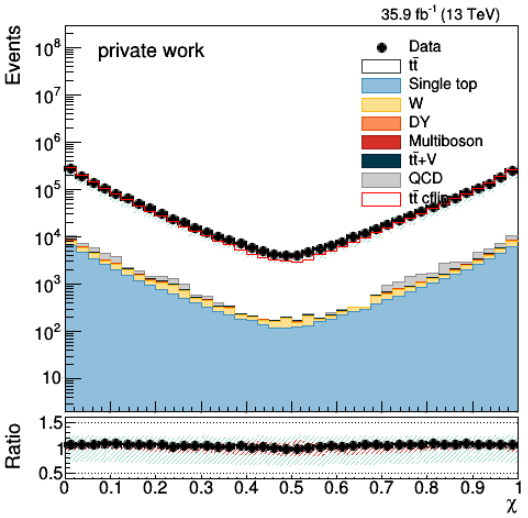
(a) j_1^b, j_2^b



(b) j_h^b, j_f^W

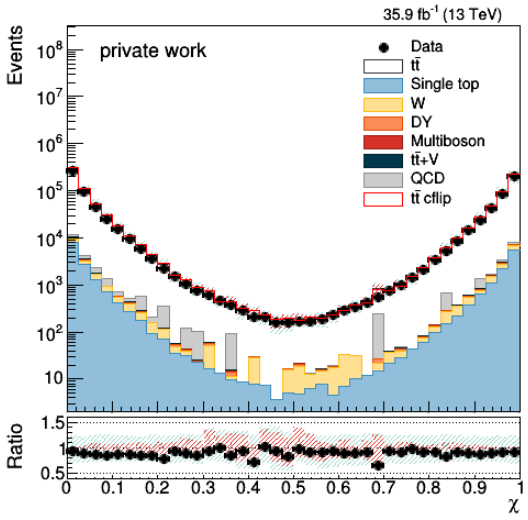


(c) j_c^W, j_h^b

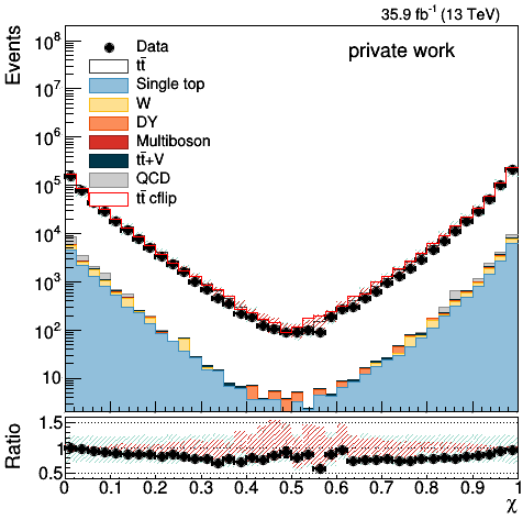


(d) j_1^W, j_2^W

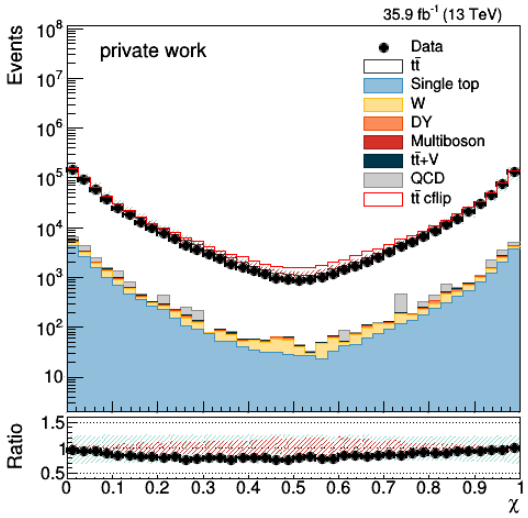
Figure 7.22. Plots of the particle flow including all jet constituents.



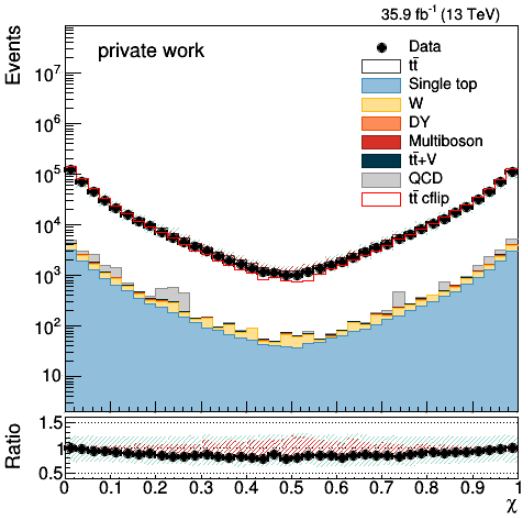
(a) j_1^b, j_2^b



(b) j_h^b, j_f^W

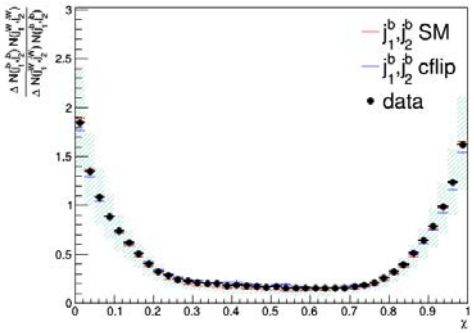


(c) j_c^W, j_h^b

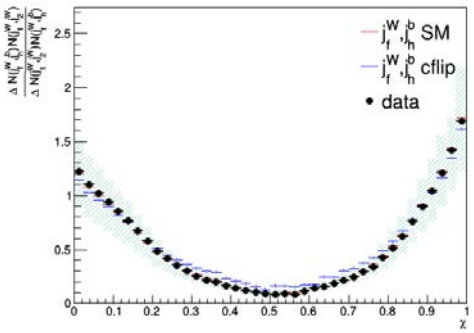


(d) j_1^W, j_2^W

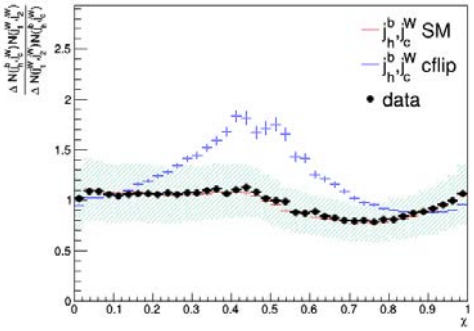
Figure 7.23. Plots of the particle flow including only charged jet constituents.



(a) j_1^b, j_2^b

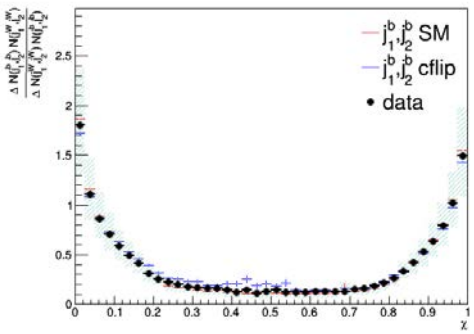


(b) j_{h}^b, j_{f}^W

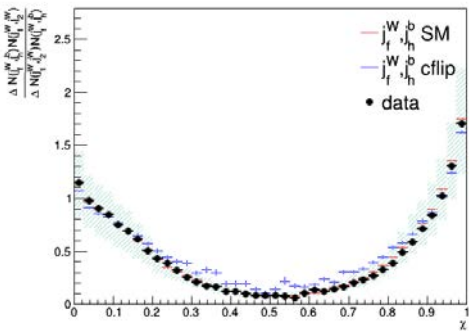


(c) j_{c}^W, j_{h}^b

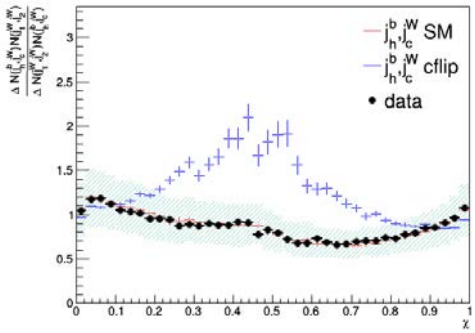
Figure 7.24. Graphs showing the particle flow including all jet constituents normalised to the flow at j_1^W, j_2^W.



(a) j_1^b, j_2^b

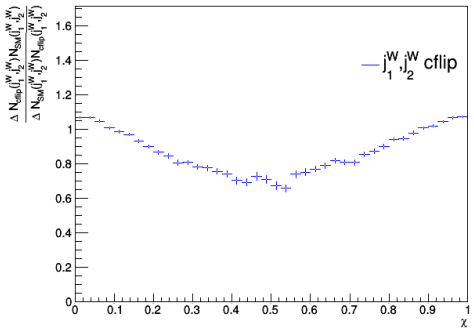


(b) j_h^b, j_f^W

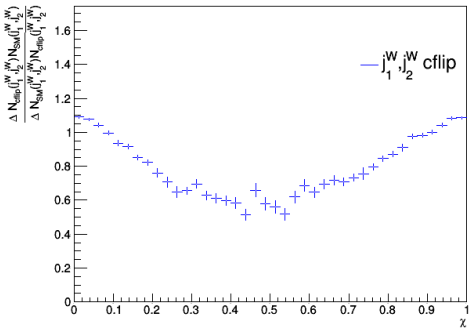


(c) j_c^W, j_h^b

Figure 7.25. Graphs showing the particle flow including only charged jet constituents normalised to the flow at j_1^W, j_2^W.



(a) Result using all jet constituents.



(b) Result using only charged jet constituents.

Figure 7.26. Bin-per-bin ratio of particle flow in region (j_1^W, j_2^W) in the SM model to particle flow in region (j_1^W, j_2^W) in the W colour octet model.

Table 7.6
Integral from 0.2 to 0.8 and the value of R^{-1} for the common lepton channel at reconstruction level for particle flow in MC for the SM model.

Jet constituents	$I \pm$(stat)$\pm$(syst) [rad]	$R^{-1} \pm$(stat)$\pm$(syst)
	j_1^b, j_2^b	
all	$0.010 \pm 0.000 \pm 0.000$	$0.151 \pm 0.000 \pm 0.002$
charged	$0.007 \pm 0.000 \pm 0.000$	$0.127 \pm 0.000 \pm 0.003$
	$j_{\mathrm{f}}^W, j_{\mathrm{h}}^b$	
all	$0.012 \pm 0.000 \pm 0.000$	$0.183 \pm 0.000 \pm 0.001$
charged	$0.008 \pm 0.000 \pm 0.000$	$0.152 \pm 0.000 \pm 0.003$
	$j_{\mathrm{h}}^b, j_{\mathrm{c}}^W$	
all	$0.062 \pm 0.000 \pm 0.000$	$0.923 \pm 0.000 \pm 0.002$
charged	$0.040 \pm 0.000 \pm 0.000$	$0.768 \pm 0.000 \pm 0.005$
	j_1^W, j_2^W	
all	$0.067 \pm 0.000 \pm 0.001$	$1.000 \pm 0.000 \pm 0.010$
charged	$0.052 \pm 0.000 \pm 0.001$	$1.000 \pm 0.000 \pm 0.016$

Table 7.7
Integral from 0.2 to 0.8 and the value of R^{-1} for the common lepton channel at reconstruction level for particle flow in data for the SM model.

Jet constituents	$I \pm$(stat) [rad]	$R^{-1} \pm$(stat)
	j_1^b, j_2^b	
all	0.011 ± 0.000	0.173 ± 0.000
charged	0.006 ± 0.000	0.139 ± 0.000
	$j_{\mathrm{f}}^W, j_{\mathrm{h}}^b$	
all	0.011 ± 0.000	0.175 ± 0.000
charged	0.007 ± 0.000	0.144 ± 0.000
	$j_{\mathrm{h}}^b, j_{\mathrm{c}}^W$	
all	0.062 ± 0.000	0.944 ± 0.000
charged	0.035 ± 0.000	0.766 ± 0.000
	j_1^W, j_2^W	
all	0.065 ± 0.000	1.000 ± 0.000
charged	0.046 ± 0.000	1.000 ± 0.000

Table 7.8
Integral from 0.2 to 0.8 and the value of R^{-1} for the common lepton channel at reconstruction level for particle flow in MC for the colour octet W model.

Jet constituents	$I \pm$(stat)$\pm$(syst) [rad]	$R^{-1} \pm$(stat)$\pm$(syst)
	j_1^b, j_2^b	
all	$0.010 \pm 0.000 \pm 0.000$	$0.189 \pm 0.000 \pm 0.002$
charged	$0.006 \pm 0.000 \pm 0.000$	$0.175 \pm 0.000 \pm 0.004$
	$j_{\mathrm{f}}^W, j_{\mathrm{h}}^b$	
all	$0.013 \pm 0.000 \pm 0.000$	$0.250 \pm 0.000 \pm 0.001$
charged	$0.008 \pm 0.000 \pm 0.000$	$0.243 \pm 0.000 \pm 0.003$
	$j_{\mathrm{h}}^b, j_{\mathrm{c}}^W$	
all	$0.071 \pm 0.000 \pm 0.000$	$1.375 \pm 0.000 \pm 0.003$
charged	$0.049 \pm 0.000 \pm 0.000$	$1.433 \pm 0.000 \pm 0.008$
	j_1^W, j_2^W	
all	$0.052 \pm 0.000 \pm 0.000$	$1.000 \pm 0.000 \pm 0.003$
charged	$0.034 \pm 0.000 \pm 0.000$	$1.000 \pm 0.000 \pm 0.008$

7.7 Hypothesis testing

Our present work with the colour-flipped MC samples provides some means to resolve if we can see the colour octet W signal in the data. Such results are to be treated cautiously because the agreement between data and SM MC samples is not particularly good. Here we will revert to the tool used by the particle physicist to announce a discovery: testing the background only hypothesis against a signal + background hypothesis with a significance Z of at least 5 [57]. The first hypothesis is called the null hypothesis H_0 while the latter one is called the alternative hypothesis H_{alt}.

We construct a two-hypothesis model to combine background, $t\bar{t}$ and colour-flipped $t\bar{t}$ signals:

$$n = \mu \left((1-x)\, f_{t\bar{t}} + x f_{t\bar{t}_{\text{cflip}}} \right) + b, \tag{7.8}$$

where n is the expected number of events, μ is the signal strength, x a parameter to assign weight to the $t\bar{t}$ and colour-flipped $t\bar{t}$ signal so that their total weight sums up to 1. b is the MC backgrounds. In the subsequent computer analysis μ is set to 1 and x is defined as the parameter of interest.

As the test statistic we choose the Tevatron test statistic. It is also known as the Neyman–Pearson test statistic. The Tevatron test statistic is defined as:

$$q^{\text{TEV}} = -2 \ln \frac{L(H_0)}{L(H_{\text{alt}})} = -2 \ln \frac{L\left(\text{data}|p=0, \hat{\theta}_0\right)}{L\left(\text{data}|p=P, \hat{\theta}_P\right)}, \tag{7.9}$$

where p is the parameter of interest, θ is the nuisance factor and $\hat{\theta}$ is the nuisance factor that maximises the profile likelihood. The likelihood L is defined the probility of the hypothesis given the data. Assuming a hypothesis with signal strength μ the likelihood is evaluated as:

$$L(\mu, \theta_s, \theta_b) = \prod_{i=1}^{N} \frac{(\mu s_i(\theta_s) + b_i(\theta_b))^{n_i}}{n_i!} e^{-(\mu s_i(\theta_s) + b_i(\theta_b))}, \tag{7.10}$$

where i is the phase space parameter (bin index), n_i is the observation (data) in the relevant phase (bin).

The Tevatron test statistic is of interest to us because if x is defined as the parameter of interest in the two hypothesis model Eq. 7.8 and P is set to 1, it happens that when applying the q^{TEV} statistic H_0 (with $x = 0$) is defined as the $t\bar{t} + b$ distribution while H_{alt} is defined as the $t\bar{t}_{\text{cflip}} + b$ distributions.

In order to test the H_0 and H_{alt} hypotheses one needs to calculate their p-values . A right-handed p-value is defined as

$$p \equiv \int_{q_{\text{obs}}}^{\infty} f(q) dq, \tag{7.11}$$

where q_{obs} is the value of the test statistic observed from the data, and f is the probability distribution function (pdf) under the assumption of the hypothesis. A low p-value is an indicator against the assumed hypothesis. A significance of $Z = 5$ corresponds to a p-value of 2.87×10^{-7}. For the Neyman–Pearson test statistic the p-value for H_0 is right-handed while the p-value for H_{alt} is left-handed. This is illustrated in Fig. 7.27.

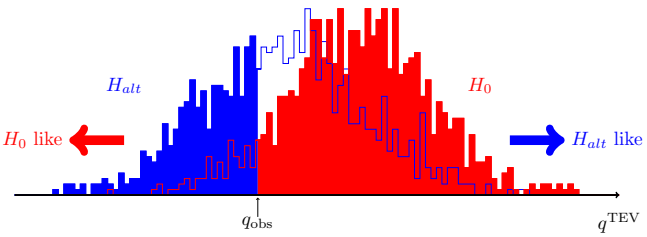


Figure 7.27. Evaluation of hypotheses according to the Neyman–Pearson test statistic.

For testing the hypothesis and doing all background work we use the CMS combine tool [58]. For the generation of the test statistic the HybridNew method of the combine tool is used. To calculate the theoretical test statistic distributions data is estimated from the MC samples in the frequentist approach. Invocation of the HybridNew method is given in the following listing:

```
combine -M HybridNew -T 500 -i 2 --fork 6 --clsAcc 0 --
    fullBToys -m 125.7 TwoHypo.root --seed 8192 --testStat=TEV
      --saveHybridResult --singlePoint 1
```

where TwoHypo.root is the ROOT file containing the workspace. --singlePoint 1 means that we require x – the parameter of interest in Eq. 7.8 to be equal to 1 in H_{alt}. We at the present stage use only 500 toys. The distribution of $q/2$ where q is the test statistic under the assumption of H_0, H_{alt} and $q_{\text{obs}}/2$ is given in Fig. 7.28.

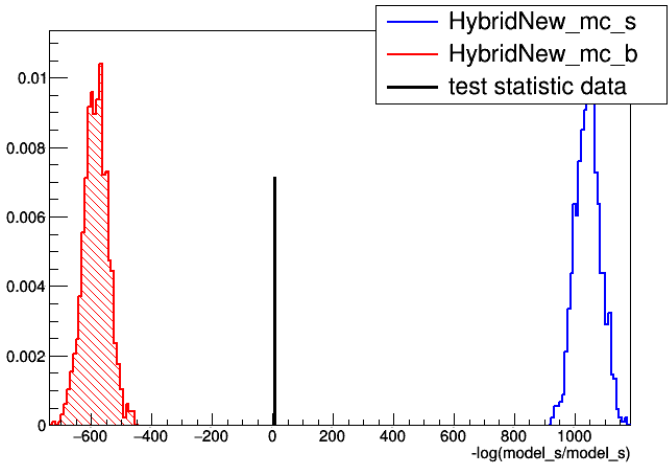


Figure 7.28. Distribution of the $q/2$ under the assumption of $t\bar{t}$ hypothesis (red), colour-flipped $t\bar{t}$ hypothesis (blue) and $q_{\text{obs}}/2$.

The p-values of H_{alt} and H_0 are infinitessimal. Thus we cannot make a conclusion – we fail to reject H_0 in favour of H_{alt} and fail to reject H_{alt} in favour of H_0.

The combine tool has a method MultiDimFit to determine the curve of the profile likelihood ratio PLR:

$$\text{PLR}(x,\theta) = -2\ln\frac{L(x=0,\theta)}{L(\hat{x},\hat{\theta})}. \tag{7.12}$$

At $\hat{x}$ and $\hat{\theta}$ the PLR has a minimum. At this point the MC best fits the data. The PLR curve can be obtained by invoking

```
combine -M MultiDimFit --algo grid --points 50 TwoHypo.root
```

The PLR curve is plotted in Fig. 7.29 and has a minimum at $x = 0.335$.

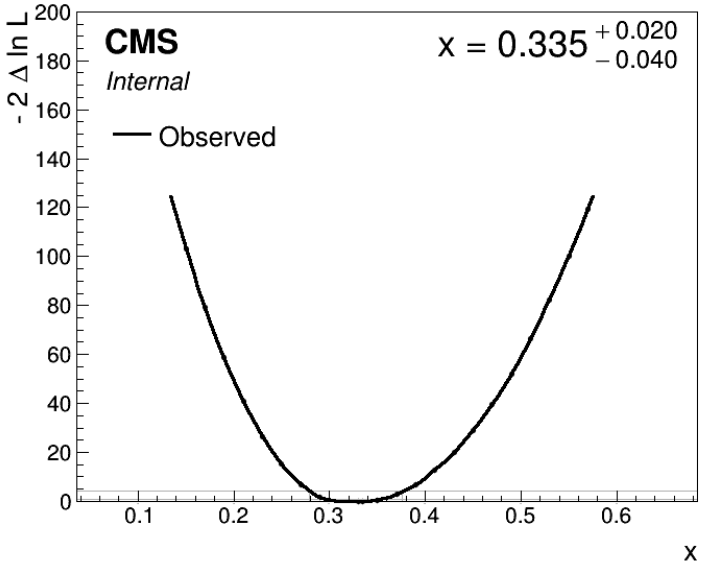


Figure 7.29. The PLR curve as a function of x.

When calculating the likelihood the combine tool combines the nominal signal with the nuisances and looks for the combination that maximises the profile likelihood. Different nuisances have a different impact. The impact of a nuisance parameter θ is defined as the shift Δx in the parameter of interest when the nuisance is included at its $\pm\sigma$ values:

$$\Delta x = x\Big|_{\theta \text{ at } \pm \sigma} - x_0. \tag{7.13}$$

In order to achieve the maximum profile likelihood different nuisances have to be stretched to a different amount. The pull of a nuisance parameter θ that quantifies this stretch is defined as:

$$P = \frac{\hat{\theta} - \theta_0}{\delta\theta}, \tag{7.14}$$

where $\hat{\theta}$ is the θ that maximises the profile likelihood, θ_0 is the pre-fit value, $\delta\theta$ is the pre-fit uncertainty.

In order the measure the impact and pull of the nuisance parameters we use the Impact method of the combine tool with the following recipe:

```
combineTool.py -M Impacts -d TwoHypo.root -m 125.7 --
    doInitialFit --robustFit 1
```

```
combineTool.py -M Impacts -d TwoHypo.root -m 125.7 --
    robustFit 1 --doFits
combineTool.py -M Impacts -d TwoHypo.root -m 125.7 -o impacts
    .json
plotImpacts.py -i impacts.json -o impacts
```

The impacts and pulls of the different nuisance parameters are plotted in Fig. 7.30.

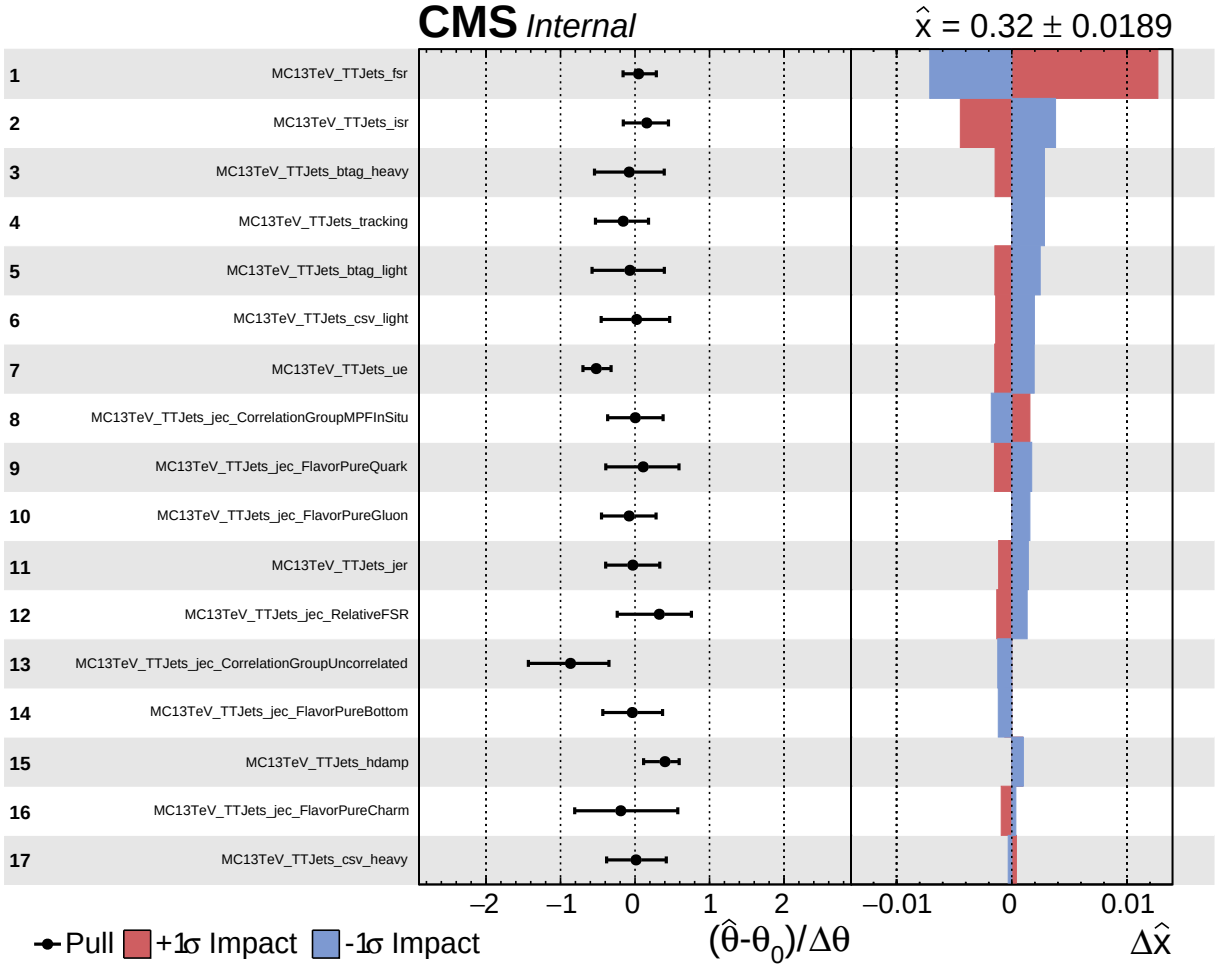


Figure 7.30. Impact and pull of different nuisance parameters.

Having obtained the value $\hat{x} = 0.335$ (Fig. 7.29) we can return to the hypothesis testing this time setting $x = \hat{x}$. In this case we will test the $t\bar{t}$ only hypothesis (H_0) against the hypothesis where the signal is composed of 66.5 % $t\bar{t}$ process and 33.5 % colour-flipped $t\bar{t}$ process (H_{alt}). The distribution of the test statistic for $x = \hat{x}$ is plotted in Fig. 7.31.

Under $x = \hat{x}$ the p-value for H_0 is 0 while the p-value for H_{alt} is 0.25. Thus we are able to reject H_0 in favour of H_{alt}.

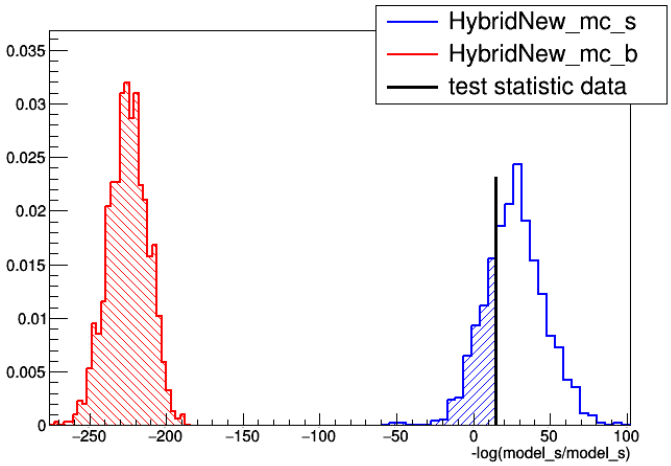


Figure 7.31. Distribution of the $q/2$ under the assumption of $t\bar{t}$ only hypothesis (red), a hypothesis of the signal being mixed of 66.5 % $t\bar{t}$ and 33.5 % colour-flipped $t\bar{t}$ process (blue) and $q_{\text{obs}}/2$.

8. Conclusions

We have been convinced that the method of pull angle based on good track reconstruction is sensitive to identify colour-connected jets. In the distribution of the pull angle there is a discernible peak centred on 0 rad for colour-connected jets while the distribution is flat for jets not connected in colour.

Convincing results have also been obtained applying the "LEP method". The density of particles is higher between colour-connected jets than in colour-free regions.

We were able to test the results with W colour octet samples in which the colour connection between the hadronic decay products of the W boson was removed. Hence, these jets appeared as jets not connected in colour in the pull angle method and the "LEP method".

We did the exercise of unfolding the pull angle as it is a valid model to identify the true value of the observable before the reconstruction at detector. Unfolding did not bring any change into our conclusions.

We noticed that the POWHEG + PYTHIA MC simulation overemphasises colour connection compared to detector observations of real world events. This is represented in a more prominent central peak in the distribution of the pull angle in MC simulations. HERWIG++ and several PYTHIA tunes turn out to be better modellers of colour connection in hadronisation.

Overall, the fit between data and MC results is not particularly good. A combination of $\sim {}^{2}/_{3}$ $t\bar{t}$ results and $\sim {}^{1}/_{3}$ $t\bar{t}$ $cflip$ results best fit the detector observations. This result was obtained in the hypothesis testing exercise.

Bibliography

[1] Albert M Sirunyan et al. "Measurement of the $t\bar{t}$ production cross section, the top quark mass, and the strong coupling constant using dilepton events in pp collisions at $\sqrt{s} = 13$ TeV". In: (2018). arXiv: 1812.10505 [hep-ex].

[2] Albert M Sirunyan et al. "Particle-flow reconstruction and global event description with the CMS detector". In: (2017). arXiv: 1706.04965 [physics.ins-det].

[3] Jason Gallicchio and Matthew D. Schwartz. "Seeing in Color: Jet Superstructure". In: *Phys. Rev. Lett.* 105 (2010), p. 022001. DOI: 10.1103/PhysRevLett.105.022001. arXiv: 1001.5027 [hep-ph].

[4] G. Abbiendi et al. "Colour reconnection in $e^+e^- \to W^+W^-$ at $\sqrt{s}$ = 189-GeV - 209-GeV". In: *Eur. Phys. J.* C45 (2006), pp. 291–305. DOI: 10.1140/epjc/s2005-02439-x. arXiv: hep-ex/0508062 [hep-ex].

[5] J. Abdallah et al. "Investigation of colour reconnection in WW events with the DELPHI detector at LEP-2". In: *Eur. Phys. J.* C51 (2007), pp. 249–269. DOI: 10.1140/epjc/s10052-007-0304-9. arXiv: 0704.0597 [hep-ex].

[6] P. Achard et al. "Search for color reconnection effects in $e^+e^- \to W^+W^- \to$ hadrons through particle flow studies at LEP". In: *Phys. Lett.* B561 (2003), pp. 202–212. DOI: 10.1016/S0370-2693(03)00490-8. arXiv: hep-ex/0303042 [hep-ex].

[7] Andy Buckley et al. "Rivet user manual". In: *Comput. Phys. Commun.* 184 (2013), pp. 2803–2819. DOI: 10.1016/j.cpc.2013.05.021. arXiv: 1003.0694 [hep-ph].

[8] Victor Mukhamedovich Abazov et al. "Measurement of color flow in $t\bar{t}$ events from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV". In: *Phys. Rev.* D83 (2011), p. 092002. DOI: 10.1103/PhysRevD.83.092002. arXiv: 1101.0648 [hep-ex].

[9] Georges Aad et al. "Measurement of colour flow with the jet pull angle in $t\bar{t}$ events using the ATLAS detector at $\sqrt{s} = 8$ TeV". In: *Phys. Lett.* B750 (2015), pp. 475–493. DOI: 10.1016/j.physletb.2015.09.051. arXiv: 1506.05629 [hep-ex].

[10] Morad Aaboud et al. "Measurement of colour flow using jet-pull observables in $t\bar{t}$ events with the ATLAS experiment at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J.* C78.10 (2018), p. 847. DOI: 10.1140/epjc/s10052-018-6290-2. arXiv: 1805.02935 [hep-ex].

[11] M. Seidel, P. Schleper, and H. Stadie. *Study of colour flow in l+jets*. site. 2013. URL: https://indico.cern.ch/event/276523/contributions/622482/attachments/502591/694047/colourflow_mseidel.pdf.

[12] G. Aad et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), S08003. DOI: 10.1088/1748-0221/3/08/S08003.

[13] *Summary of LHC and Tevatron measurements of the top-pair production cross-section as a function of the centre-of-mass energy compared to the NNLO QCD calculation complemented with NNLL resummation (top++2.0)*. Twiki Page. 2018. URL: https://twiki.cern.ch/twiki/pub/LHCPhysics/LHCTopWGSummaryPlots/tt_curve_toplhcwg_sep18.pdf.

[14] Timo Antero Aaltonen et al. "Tevatron Combination of Single-Top-Quark Cross Sections and Determination of the Magnitude of the Cabibbo-Kobayashi-Maskawa Matrix Element V_{tb}". In: *Phys. Rev. Lett.* 115.15 (2015), p. 152003. DOI: 10.1103/PhysRevLett.115.152003. arXiv: 1503.05027 [hep-ex].

[15] C. Patrignani et al. "Review of Particle Physics". In: *Chin. Phys.* C40.10 (2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001.

[16] Victor Mukhamedovich Abazov et al. "Determination of the width of the top quark". In: *Phys. Rev. Lett.* 106 (2011), p. 022001. DOI: 10.1103/PhysRevLett.106.022001. arXiv: 1009.5686 [hep-ex].

[17] Markus Cristinziani and Martijn Mulders. "Top-quark physics at the Large Hadron Collider". In: *J. Phys.* G44.6 (2017), p. 063001. DOI: 10.1088/1361-6471/44/6/063001. arXiv: 1606.00327 [hep-ex].

[18] Bo Andersson et al. "Parton Fragmentation and String Dynamics". In: *Phys. Rept.* 97 (1983), pp. 31–145. DOI: 10.1016/0370-1573(83)90080-7.

[19] Peter W. Higgs. "Broken symmetries, massless particles and gauge fields". In: *Phys. Lett.* 12 (1964), pp. 132–133. DOI: 10.1016/0031-9163(64)91136-9.

[20] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (1964). [,157(1964)], pp. 321–323. DOI: 10.1103/PhysRevLett.13.321.

[21] Serguei Chatrchyan et al. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Phys. Lett.* B716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].

[22] Georges Aad et al. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Phys. Lett.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].

[23] *Overall view of the LHC*. CERN Document Server. 2014. URL: http://cds.cern.ch/record/1708847.

[24] *CERN's Accelerator Complex*. espace. 2014. URL: https://espace.cern.ch/acc-tec-sector/Pictures/CERN's\%20accelerator\%20complex2015bis.png.

[25] Oliver S. Bruning et al. "LHC Design Report Vol.1: The LHC Main Ring". In: (2004).

[26] Stephen Myers. "The Large Hadron Collider 2008-2013". In: *Int. J. Mod. Phys.* A28 (2013), p. 1330035. DOI: 10.1142/S0217751X13300354.

[27] Lyndon Evans and Philip Bryant. "LHC Machine". In: *JINST* 3 (2008), S08001. DOI: 10.1088/1748-0221/3/08/S08001.

[28] S. Chatrchyan et al. "The CMS Experiment at the CERN LHC". In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.

[29] Tai Sakuma and Thomas McCauley. "Detector and Event Visualization with SketchUp at the CMS Experiment". In: *J. Phys. Conf. Ser.* 513 (2014), p. 022032. DOI: 10.1088/1742-6596/513/2/022032. arXiv: 1311.4942 [physics.ins-det].

[30] V Karimäki et al. *The CMS tracker system project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/368412.

[31] *The CMS tracker: addendum to the Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 2000. URL: https://cds.cern.ch/record/490194.

[32] P. Adzic et al. "Energy resolution of the barrel of the CMS electromagnetic calorimeter". In: *JINST* 2 (2007), P04004. DOI: 10.1088/1748-0221/2/04/P04004.

[33] *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/357153.

[34] *The CMS muon project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/343814.

[35] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The Anti-k(t) jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[36] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "FastJet User Manual". In: *Eur. Phys. J.* C72 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph].

[37] The ATLAS collaboration. "Measurement of colour flow using jet-pull observables in $t\bar{t}$ events with the ATLAS experiment at $\sqrt{s} = 13$ TeV". In: (2017).

[38] G. Salam. *Anti-kt jet boundaries*. GitHub page. 2018. URL: https://github.com/gavinsalam/antikt-jet-boundaries.

[39] Serguei Chatrchyan et al. "Description and performance of track and primary-vertex reconstruction with the CMS tracker". In: *JINST* 9.10 (2014), P10009. DOI: 10.1088/1748-0221/9/10/P10009. arXiv: 1405.6569 [physics.ins-det].

[40] TOP Physics Object Group. *Measurement of the underlying event in t-t dilepton events at √s=13 TeV*. draft CMS Analysis Note CMS-AN-2017-175. Geneva: CERN, 2018.

[41] TOP Physics Object Group. *Jet shapes in ttbar events at 13 TeV (2016 data)*. draft CMS Analysis Note CMS-AN-2017-159. Geneva: CERN, 2018.

[42] *BRIL Work Suite*. site. 2015. URL: http://cms-service-lumi.web.cern.ch/cms-service-lumi/brilwsdoc.html.

[43] *NLO single-top channel cross sections : ATLAS-CMS recommended predictions for single-top cross sections using the Hathor v2.1 program*. twiki. 2015. URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec.

[44] *Summary table of samples produced for the 1 Billion campaign, with 25ns bunch-crossing*. twiki. 2015. URL: https://twiki.cern.ch/twiki/bin/view/CMS/SummaryTable1G25ns.

[45] *McM Monte-Carlo Request Management*. site. 2015. URL: https://cms-pdmv.cern.ch/mcm/.

[46] *NNLO+NNLL top-quark-pair cross sections : ATLAS-CMS recommended predictions for top-quark-pair cross sections using the Top++v2.0 program (M. Czakon, A. Mitov, 2013)*. twiki. 2015. URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/TtbarNNLO.

[47] Goh, J. and Seidel, M. *PseudoTopProducer*. github repository. 2017. URL: https://github.com/intrepid42/cmssw/tree/pseudotoprivet_80x/TopQuarkAnalysis/TopEventProducers.

[48] *Jet Energy Resolution Measurement*. Twiki Page. 2017. URL: https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution.

[49] *Recommended Jet Energy Corrections and Uncertainties For Data and MC*. Twiki Page. 2017. URL: https://twiki.cern.ch/twiki/bin/view/CMS/JECDataMC.

[50] *BTagCalibration Documentation*. twiki. 2015. URL: https://twiki.cern.ch/twiki/bin/view/CMS/BTagCalibration.

[51] M. Bahr et al. "Herwig++ Physics and Manual". In: *Eur. Phys. J.* C58 (2008), pp. 639–707. DOI: 10.1140/epjc/s10052-008-0798-9. arXiv: 0803.0883 [hep-ph].

[52] Vardan Khachatryan et al. "Measurement of the top quark mass using proton-proton data at $\sqrt{(s)} = 7$ and 8 TeV". In: *Phys. Rev.* D93.7 (2016), p. 072004. DOI: 10.1103/PhysRevD.93.072004. arXiv: 1509.04044 [hep-ex].

[53] Veckalns, V. *CFAT*. Address. 2018. URL: https://github.com/vveckaln/CFAT.

[54] R. Brun and F. Rademakers. "ROOT - An Object Oriented Data Analysis Framework". In: *Nucl. Inst.* A 389 (1997), pp. 81–86.

[55] Stefan Schmitt. "TUnfold: an algorithm for correcting migration effects in high energy physics". In: *JINST* 7 (2012), T10003. DOI: 10.1088/1748-0221/7/10/T10003. arXiv: 1205.6201 [physics.data-an].

[56] Veckalns, V. *CompoundHistoUnfolding*. Address. 2018. URL: https://github.com/vveckaln/COMPOUNHISTOUNFOLDING.

[57] Glen Cowan et al. "Asymptotic formulae for likelihood-based tests of new physics". In: *Eur. Phys. J.* C71 (2011). [Erratum: Eur. Phys. J.C73,2501(2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0, 10.1140/epjc/s10052-013-2501-z. arXiv: 1007.1727 [physics.data-an].

[58] *The CMS combine tool*. Address. 2018. URL: https://cms-hcomb.gitbooks.io/combine/.