

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

Doctoral Thesis

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

RTU Izdevniecība  
Rīga 2019

# 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  $3.3 \times 10^{-25}$  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.

## Anotācija

LHP, kas strādā ar 13 TeV masas centra enerģiju, ir virsotnes kvarku fabrika. Virsotnes kvarku ražošanas šķērsgriezums LHP ir 803 pb. Virsotnes kvarka mūža ilgums ir  $3,3 \times 10^{-25}$  s, un tas ir tik īss, ka atšķirībā no citiem kvarkiem virsotnes kvarks sabrūk, pirms tas hadronizējas. Virsotnes kvarks sabrūk vāmajā ceļā, izstarojot  $W$  bozonu. Ja  $W$  bozons sabrūk hadroniskajā veidā, stiprajā kodola mijiedarbībā tiek radītas daļiņu strūklas. Šo procesu apraksta kvantu hromodinamika, un varam modelēt virsotnes kvarka sabrukšanas procesu ar krāsu lādiņu un krāsu saitēm. Strūklas, kas radušās, sabrūkot  $W$  bozonam, mijiedarbojas krāsu laukā (tās ir saistītas ar krāsām). Saistība ar krāsām atstāj pamanāmus eksperimentālus nospiedumus, ko mēs varam novērot KMS detektorā, īpaši izmantojot tā trekeri, 4 T solenoīdu un kalorimetrus. KMS eksperimentā šāds pētījums tiek veikts pirmoreiz. Krāsu saistību starp strūklām, kas radušās, sabrūkot virsotnes kvarku pārim, pētām, izmantojot gala stāvokli, ko veido viens lādēts leptons, divas vieglās strūklas un divas  $b$  atzīmētās strūklas. Izmantojam vilkmes leņķi un daļiņu projicēšanu uz plaknes, ko veido divas strūklas. Tiek izmantots arī krāsu okteta  $W$  spēļu modelis, lai novērtētu dažādo metožu sniegumu.

# 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 Etienne 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

<b>1</b>	<b>Introduction</b>	<b>14</b>
<b>2</b>	<b>Physical Background</b>	<b>15</b>
<b>3</b>	<b>Experimental Setup</b>	<b>25</b>
3.1	The LHC . . . . .	25
3.2	The CMS detector . . . . .	28
<b>4</b>	<b>Methodology</b>	<b>32</b>
4.1	Pull angle . . . . .	32
4.2	LEP method . . . . .	35
<b>5</b>	<b>Data and MC Samples</b>	<b>37</b>
5.1	Corrections applied to the simulation . . . . .	40
<b>6</b>	<b>Event Selection</b>	<b>49</b>
6.1	Detector level . . . . .	49
6.2	Control plots . . . . .	54
6.3	Generator level . . . . .	58
<b>7</b>	<b>Systematic Uncertainties</b>	<b>59</b>
7.1	Experimental uncertainties . . . . .	59
7.2	Theoretical uncertainties . . . . .	66
<b>8</b>	<b>Results</b>	<b>72</b>
8.1	Pull vector . . . . .	72
8.2	Pull angle . . . . .	78
8.3	$\Delta R$ bias . . . . .	82
8.4	Sensitivity analysis . . . . .	83
8.5	Unfolding . . . . .	87
8.6	LEP method . . . . .	112
8.7	Hypothesis testing . . . . .	129
<b>9</b>	<b>Conclusions</b>	<b>134</b>
<b>A</b>	<b>The datacard for the combine tool</b>	<b>135</b>
	<b>Bibliography</b>	<b>136</b>

# List of Figures

2.1	The top quark in the Standard Model. . . . .	15
2.2	Inclusive cross section of the top pair at different centre of mass energies [16]. The result shows the cross section at $p p$ and $p \bar{p}$ collisions. The results from CMS and ATLAS are given for different channels of the decay of the $t\bar{t}$ pair. . .	16
2.3	Top quark pair production in a $p p$ collision. . . . .	17
2.4	Weak decay of the top quark $t$ . $p'$ is a quark of different flavour and $k$ and $k'$ are fermions resulting from the decay of the $W$ boson. . . . .	17
2.5	Creation of new hadrons by two energetic quarks. . . . .	19
2.6	Colour flow in the decay of a top quark pair. . . . .	19
2.7	Creation of hadrons in the colour field of two quarks. . . . .	20
2.8	Resonances at generation level corresponding to particles listed in Table 2.2 con- stituting the leading light jet and the leading $b$ jet. In both cases all jet con- stituents are included. 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. 8. . . . .	21
2.9	Total number of particles constituting the leading light jet, the second leading light jet, the leading $b$ jet and the second leading $b$ jet. In all cases all jet con- stituents are included. . . . .	22
2.10	Ratio of the number of charged particles to the total number of particles consti- tuting the leading light jet and the leading $b$ jet. . . . .	23
2.11	Ratio of the energy of charged particles to the total enery of particles constituting the leading jet and the leading $b$ jet. . . . .	23
2.12	Colour flow in the decay of a top quark pair involving a hypothetical colour octet $W$ boson. . . . .	24
3.1	The Large Hadron Collider situated underground on the French-Swiss border in Geneva area [36]. . . . .	26
3.2	The CERN accelerator complex [37]. . . . .	27
3.3	A cut-away view of the CMS detector [42]. . . . .	28
3.4	A sketch of the specific particle interactions in a transverse slice of the CMS detector, from the beam interaction region to the muon detector [43]. . . . .	29
4.1	The coordinate system used at CMS. . . . .	33
4.2	Pull angle $\theta_p$ and pull vector $\vec{v}_p$ in a $\phi$ - $y$ plane. . . . .	34

4.3	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 while the $p_T$ of the soft jet is 75 GeV. Courtesy of Cacciari, Salam and Soyez [52]. . . . .	34
4.4	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 $\eta$ intervals of 0–0.9, 0.9–1.4 and 1.4–2.5, respectively [53]. . . . .	35
4.5	Inter- $W$ and intra- $W$ planes in the process $e^+e^- \rightarrow q\bar{q}q\bar{q}$ and the relative angle $\chi_R = \chi_1/\chi_0$ . . . . .	36
4.6	Adaptation of the LEP method to $t\bar{t}$ semileptonic decay involving a colour-connected region and 3 colour-free regions. . . . .	36
5.1	The Drell–Yan process. . . . .	39
5.2	The $\rho$ - $z$ view of an event in a high pile-up run 198609 showing 78 reconstructed vertices [79]. . . . .	40
5.3	Pile-up during different periods of the operation of the LHC [81]. . . . .	41
5.4	Pile-up distributions in Monte Carlo and data after the last step of event selection. . . . .	42
5.5	Efficiency as a function of electron $p_T$ for dielectron events in data (dots) and DY simulation (triangles), for the medium working point of the sequential selection [85]. . . . .	43
5.6	Tag-and-probe efficiency for the tight PF isolation working point on top of the tight ID versus $p_T$ for muons in the acceptance of the muon spectrometer, and versus $\eta$ for muons with $p_T > 20$ GeV, for 2015 data (circles), simulation (squares), and the ratio (bottom inset). The statistical uncertainties are smaller than the symbols used to display the measurements [86]. . . . .	43
5.7	Isolated single-muon trigger efficiencies measured with 2015 data (squares), simulation (circles), and the ratio (bottom inset). Results are plotted as a function of offline reconstructed muon $p_T$ and $\eta$ [86]. . . . .	44
5.8	Isolated single-muon trigger efficiency scale factors for data taking periods BCDEFGH of 2016. . . . .	45
5.9	Jet energy response before (a) and after (b) applying JEC. The illustration is from 2012 TeV data with $\sqrt{s} = 8$ due to unavailability of public results with Run II data [95]. For a discussion of PF+CHS jets see text. . . . .	46
5.10	JER versus $p_T$ in the barrel for varying levels of pile-up $\mu$ . The results are shown separately for PF+CHS jets with jet distance parameter $R = 0.5$ [95]. . . . .	47
5.11	Jet energy resolution data/MC scale factor versus $ \eta $ for $\gamma$ + jet data collected at 8 TeV (closed circles, solid area) compared to results at 7 TeV (open circles, dashed area) [95]. We illustrate Run I results due to unavailability of public results for Run II. . . . .	48
5.12	Data-to-simulation scale factors for $b$ jets from the hadronic or leptonic side of the single-lepton $t\bar{t}$ decay as well as for their combination, as a function of the jet $p_T$ for the medium working point of the CSVv2 tagger [98]. . . . .	48

6.1	When a prompt $b$ is created in a collision at the primary vertex (PV), it hadronises and the unstable $B$ meson usually travels some discernible distance away before it decays, the charged decay products creating tracks that originate from the secondary vertex. We measure the impact parameter (IP), e.g. along the direction of the beam [98]. . . . .	51
6.2	$b$ tagging efficiency at the HLT as a function of the offline CSVv2 discriminator value [98]. . . . .	51
6.3	Comparison of the misidentification probability for light-flavour jets (left) and $c$ jets (right) versus the $b$ tagging efficiency at the HLT and offline for the CSVv2 algorithm applied on simulated $t\bar{t}$ events for which the scalar sum of the jet $p_T$ for all jets in the event exceeds 250 GeV [98]. . . . .	52
6.4	Event yields at different stages of selection: $1\ell, 1\ell+ \geq 4j, 1\ell+ \geq 4j(2b), 1\ell+ \geq 4j(2b, 2lj)$ . . . . .	52
6.5	Distribution of the lepton $p_T$ after selecting exactly one lepton. . . . .	54
6.6	Distribution of the lepton $\eta$ after selecting exactly one lepton. . . . .	54
6.7	Distribution of the number of jets after selecting exactly one lepton. . . . .	54
6.8	Distribution of the number of jets after selecting exactly one lepton and 4 jets. . . . .	55
6.9	Distribution of the number of $b$ -tagged jets after selecting exactly one lepton and 4 jets. . . . .	55
6.10	Distribution of the $p_T$ of the lepton after selecting exactly one lepton, at least 4 jets, 2 $b$ -tagged jets and 2 light jets. . . . .	55
6.11	Distribution of the $\eta$ of the lepton after selecting exactly one lepton, at least 4 jets, 2 $b$ -tagged jets and 2 light jets. . . . .	56
6.12	Distribution of the number of jets after selecting exactly one lepton, at least 4 jets, 2 $b$ -tagged jets and 2 light jets. . . . .	56
6.13	Distribution of the missing transverse momentum after selecting exactly one lepton, at least 4 jets, 2 $b$ -tagged jets and 2 light jets. . . . .	56
6.14	Distribution of the jet $p_T$ in the combined $\ell +$ jets channel after selecting exactly one lepton, at least 4 jets, 2 $b$ -tagged jets and 2 light jets. Jets are ordered according to their $p_T$ , with jet 0 having the highest $p_T$ and jet 5 having the lowest $p_T$ . . . . .	57
6.15	Distribution of the mass of the $W$ boson candidate and $t$ quark candidate (after cuts in the the $W$ mass) after selecting exactly one lepton, at least 4 jets 2 $b$ -tagged jets and 2 light jets. . . . .	57
7.1	Muon tracking efficiency scale factors from the MUO POG [87]. . . . .	61
7.2	Expected effect on the charged multiplicity by tracking efficiency [54]. . . . .	62
7.3	The nominal distribution of the pull angle from $j_1^W$ to $j_2^W$ with all jet constituents and at all $\Delta R$ and the distributions hereof with the pile-up uncertainties. . . . .	62
7.4	The nominal distribution of the pull angle from $j_1^W$ to $j_2^W$ with all jet constituents and at all $\Delta R$ and the distributions hereof with trigger and selection efficiency uncertainties. . . . .	63

7.5	The nominal distribution of the pull angle from $j_1^W$ to $j_2^W$ with all jet constituents and at all $\Delta R$ and the distributions hereof with jet energy resolution uncertainties $t\bar{t}$ <i>jer down</i> and $t\bar{t}$ <i>jer up</i> . . . . .	63
7.6	The nominal distribution of the pull angle from $j_1^W$ to $j_2^W$ with all jet constituents and at all $\Delta R$ and the distributions hereof with jet energy correction uncertainties. . . . .	64
7.7	The nominal distribution of the pull angle from $j_1^W$ to $j_2^W$ with all jet constituents and at all $\Delta R$ and the distributions hereof with $b$ tagging uncertainties. . . . .	65
7.8	The nominal distribution of the pull angle from $j_1^W$ to $j_2^W$ with all jet constituents and at all $\Delta R$ and the distributions hereof with tracking efficiency uncertainties $t\bar{t}$ <i>tracking down</i> and $t\bar{t}$ <i>tracking up</i> . . . . .	66
7.9	The nominal distribution of the pull angle from $j_1^W$ to $j_2^W$ with all jet constituents and at all $\Delta R$ and the distributions hereof with the QCD scale uncertainties. . . . .	69
7.10	The nominal distribution of the pull angle from $j_1^W$ to $j_2^W$ with all jet constituents and at all $\Delta R$ and the distributions hereof with uncertainty arising from assuming the decay model of heavy flavour particles used in EVTGEN. . . . .	69
7.11	The nominal distribution of the pull angle from $j_1^W$ to $j_2^W$ with all jet constituents and at all $\Delta R$ and the distribution hereof with uncertainty arising from assuming the hadronisation model used in HERWIG++. . . . .	69
7.12	The nominal distribution of the pull angle from $j_1^W$ to $j_2^W$ with all jet constituents and at all $\Delta R$ and the distributions hereof with uncertainties arising from $\pm 1$ GeV variations in the mass of the $t$ quark. . . . .	70
7.13	The nominal distribution of the pull angle from $j_1^W$ to $j_2^W$ with all jet constituents and at all $\Delta R$ and the distributions hereof with uncertainty from varying the hdamp parameter in the Parton Shower + Matrix Element matching scheme. . . . .	70
7.14	The nominal distribution of the pull angle from $j_1^W$ to $j_2^W$ with all jet constituents and at all $\Delta R$ and the distributions hereof with uncertainties in the Underlying Event. . . . .	70
7.15	The nominal distribution of the pull angle from $j_1^W$ to $j_2^W$ with all jet constituents and at all $\Delta R$ and the distributions hereof with parton shower scale uncertainties. . . . .	71
7.16	The nominal distribution of the pull angle from $j_1^W$ to $j_2^W$ with all jet constituents and at all $\Delta R$ and the distributions hereof with uncertainties from colour reconnection. . . . .	71
8.1	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 $p_T [\text{GeV}]/75.0$ and the radius of the circles representing constituents is equal to $p_T^{\text{constituent}}/p_T^{\text{jet}}$ . 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. . . . .	73

8.2	Results with RIVET showing SM (red) and $W$ colour octet (blue) distributions of the pull angle. The bottom inset shows the bin-per-bin ratio of the $W$ colour octet results to the SM results. . . . .	74
8.3	Distribution of the $\eta$ component of the pull vector of $j_1^W$ with all jet components.	74
8.4	Distribution of the $\eta$ component of the pull vector of $j_2^W$ with all jet components.	75
8.5	Distribution of the $\eta$ component of the pull vector of $j_1^b$ with all jet components.	75
8.6	Distribution of the $\eta$ component of the pull vector of $j_2^b$ with all jet components.	75
8.7	Distribution of the $\phi$ component of the pull vector of $j_1^W$ with all jet components.	76
8.8	Distribution of the $\phi$ component of the pull vector of $j_2^W$ with all jet components.	76
8.9	Distribution of the $\phi$ component of the pull vector of $j_1^b$ with all jet components.	76
8.10	Distribution of the $\phi$ component of the pull vector of $j_2^b$ with all jet components.	77
8.11	The magnitude of the pull vector of $j_1^W$ with all jet components. . . . .	77
8.12	The magnitude of the pull vector of $j_2^W$ with all jet components. . . . .	77
8.13	The magnitude the pull vector of $j_1^b$ with all jet components. . . . .	78
8.14	The magnitude of the pull vector of $j_2^b$ with all jet components. . . . .	78
8.15	Distribution of the pull angle from $j_1^W$ to $j_2^W$ for all $\Delta R$ and including all particles.	79
8.16	Distribution of the pull angle from $j_2^W$ to $j_1^W$ for all $\Delta R$ and including all particles.	79
8.17	Distribution of the pull angle from $j_1^b$ to $j_2^b$ for all $\Delta R$ and including all particles.	80
8.18	Distribution of the pull angle from $j_2^b$ to $j_1^b$ for all $\Delta R$ and including all particles.	80
8.19	Distribution of the pull angle from $j_1^W$ to the charged lepton for all $\Delta R$ and including all particles. . . . .	80
8.20	Distribution of the pull angle from $j_1^W$ to the hadronic $W$ for all $\Delta R$ and including all particles. . . . .	81
8.21	Distribution of the pull angle from $j_1^W$ to the positive direction of the beam including all particles. . . . .	81
8.22	Distribution of the pull angle with $\Delta R \leq 1.0$ and including all jet constituents from $j_1^W$ to $j_2^W$ . . . . .	82
8.23	Distribution of the pull angle with $\Delta R > 1.0$ and including all jet constituents from $j_1^W$ to $j_2^W$ . . . . .	82
8.24	Distribution of the pull angle with $\Delta R \leq 1.0$ and only charged jet constituents from $j_1^W$ to $j_2^W$ . . . . .	83
8.25	Distribution of the pull angle with $\Delta R > 1.0$ and only charged jet constituents from $j_1^W$ to $j_2^W$ . . . . .	83
8.26	Distribution of the transverse momentum $p_T$ of the hadronic $W$ boson. . . . .	84
8.27	The transverse momentum $p_T$ of particles constituting the leading light jet and the leading $b$ jet. In both cases all jet constituents are included. . . . .	84
8.28	Distribution of the pull angle for all $\Delta R$ and all particles from $j_1^W$ to $j_2^W$ with $p_T$ of $W > 50$ GeV. . . . .	85
8.29	Distribution of the pull angle for all $\Delta R$ and all particles from $j_1^W$ to $j_2^W$ with $p_T$ of $W \leq 50$ GeV. . . . .	85
8.30	Distribution of the pull angle for all $\Delta R$ and all particles from $j_1^W$ to $j_2^W$ with the number of jet constituents $N > 20$ . . . . .	85
8.31	Distribution of the pull angle for all $\Delta R$ and all particles from $j_1^W$ to $j_2^W$ with the number of jet constituents $N \leq 20$ . . . . .	86

8.32	Distribution of the pull angle for all $\Delta R$ and all particles from $j_1^W$ to $j_2^W$ with the $p_T$ of jet constituents $> 0.5$ GeV. . . . .	86
8.33	Distribution of the pull angle for all $\Delta R$ and all particles from $j_1^W$ to $j_2^W$ with the $p_T$ of jet constituents $\leq 0.5$ GeV. . . . .	86
8.34	Distribution of the pull angle for all $\Delta R$ and all particles from $j_1^W$ to $j_2^W$ with the magnitude of the pull vector $> 0.005$ [a.u.]. . . . .	87
8.35	Distribution of the pull angle for all $\Delta R$ and all particles from $j_1^W$ to $j_2^W$ with the magnitude of the pull vector $\leq 0.005$ [a.u.]. . . . .	87
8.36	Method of optimising the binning scheme for unfolding. . . . .	89
8.37	Unfolding plots for the $t\bar{t}$ method of the pull angle $\theta_p$ of $j_1^W$ to $j_2^W$ including all jet constituents using the original binning. . . . .	91
8.38	Unfolding plots for the $t\bar{t}$ method of the pull angle $\theta_p$ of $j_1^W$ to $j_2^W$ including all jet constituents using the optimised binning with a $\sigma$ factor of 0.15. . . . .	94
8.39	Unfolding plots for the $t\bar{t}$ method of the pull angle $\theta_p$ of $j_1^W$ to $j_2^W$ including all jet constituents using 3 regularly sized bins. . . . .	97
8.40	Unfolding plots for the $t\bar{t}$ cflip method of the pull angle $\theta_p$ of $j_1^W$ to $j_2^W$ including all jet constituents using 3 regularly sized bins. . . . .	100
8.41	Unfolding plots for the $t\bar{t}$ method of the pull angle $\theta_p$ of $j_1^b$ to $j_2^b$ including all jet constituents using 3 regularly sized bins. . . . .	103
8.42	Unfolding plots for the $t\bar{t}$ method of the magnitude of the pull vector $ \vec{P} $ of $j_1^W$ to $j_2^W$ including all jet constituents using 3 regularly sized bins. . . . .	106
8.43	Observed masses of $W$ and $t$ on the hadronic and the leptonic branch. . . . .	113
8.44	Plots of the particle flow including all jet constituents. . . . .	116
8.45	Plots of the particle flow including only charged jet constituents. . . . .	117
8.46	Plots of the energy flow including all jet constituents. . . . .	118
8.47	Plots of the energy flow including only charged jet constituents. . . . .	119
8.48	Plots of the $p_T$ flow including all jet constituents. . . . .	120
8.49	Plots of the $p_T$ flow including only charged jet constituents. . . . .	121
8.50	Graphs showing the particle flow including all jet constituents normalised to the flow at $j_1^W, j_2^W$ . . . . .	122
8.51	Graphs showing the particle flow including only charged jet constituents normalised to the flow at $j_1^W, j_2^W$ . . . . .	123
8.52	Bin-per-bin ratio of particle flow in region $(j_1^W, j_2^W)$ in the $W$ colour octet model to particle flow in region $(j_1^W, j_2^W)$ in the SM model. . . . .	123
8.53	Evaluation of hypotheses according to the Neyman–Pearson test statistic. . . . .	130
8.54	Distribution of the $q/2$ under the assumption of $t\bar{t}$ hypothesis (red), colour-flipped $t\bar{t}$ hypothesis (blue) and $q_{\text{obs}}/2$ . . . . .	130
8.55	The PLR curve as a function of $x$ . . . . .	131
8.56	Impact and pull of different nuisance parameters. . . . .	132
8.57	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 $q_{\text{obs}}/2$ . . . . .	133

# List of Tables

2.1	Branching ratios from the decay of the $W$ boson. . . . .	18
2.2	New particles created in the colour field between energetic colour-connected quarks originating from a hadronic decay of the $W$ boson. . . . .	18
5.1	Primary datasets used in this analysis. PD is an abbreviation for SingleMuon or SingleElectron [55]. . . . .	37
5.2	List of simulation samples. We quote the cross section used to normalise the sample in the analysis. Adapted after [55]. . . . .	38
5.3	Simulation samples for the colour octet $W$ boson. We quote the cross section used to normalise the sample in the analysis. . . . .	40
6.1	Trigger paths used for online selection in the analysis. . . . .	50
6.2	Event yields. . . . .	52
6.3	Event yields for the colour octet $W$ sample. . . . .	53
7.1	Tracking efficiency for tracks with $p_T > 1$ GeV based on [104, 105, 106]. Courtesy of V. Mariani. . . . .	61
7.2	Variations of the POWHEG + PYTHIA 8 setup used for the comparison with the measurements. The values changed with respect to the CUETP8M2T4 tune are given in the columns corresponding to each model. After [119]. . . . .	68
7.3	Simulation samples used for systematics [55]. . . . .	68
8.1	Bin by bin weight of uncertainty (shapes) for pull angle $\theta_p$ including all jet constituents from the $j_1^W$ to the $j_2^W$ at the generator level. The results are for unfolded output for the POWHEG + PYTHIA 8 sample. The binning method of 3 regularly sized bins is used. . . . .	109
8.2	Bin by bin weight of the additional uncertainty (shape) $t\bar{t}$ <i>cfli</i> p for pull angle $\theta_p$ including all jet constituents from the $j_1^W$ to the $j_2^W$ at the generator level. The results are for unfolded output for the POWHEG + PYTHIA 8 * sample. The binning method of 3 regularly sized bins is used. . . . .	110
8.3	$\chi^2$ and $p$ -values of pull angle $\theta_p$ including all jet constituents. The results are for the SM model. . . . .	110
8.4	$\chi^2$ and $p$ -values of pull angle $\theta_p$ including all jet constituents. The results are for the colour octet $W$ model. . . . .	111
8.5	$\chi^2$ and $p$ -values of magnitude of the pull vector $ \vec{P} $ including all jet constituents. The results are for the SM model. . . . .	111

8.6	$\chi^2$ and $p$ -values of magnitude of the pull vector $ \vec{P} $ including all jet constituents. The results are for the colour octet $W$ model. . . . .	111
8.7	$\chi^2$ and $p$ -values of pull angle $\theta_p$ including all jet constituents. The nuisances are substituted in place of the signal. The results are for the SM model. . . . .	111
8.8	Observed masses of objects for the common lepton channel at reconstruction level for MC for the SM model. . . . .	112
8.9	Observed masses of objects for the common lepton channel at reconstruction level for data for the SM model. . . . .	113
8.10	Observed masses of objects for the common lepton channel at reconstruction level for MC for the colour octet $W$ model. . . . .	114
8.11	$R$ values observed at LEP. . . . .	115
8.12	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. . . . .	124
8.13	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. . . . .	124
8.14	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. . . . .	125
8.15	Integral from 0.2 to 0.8 and the value of $R^{-1}$ for the common lepton channel at reconstruction level for energy flow in MC for the SM model. . . . .	125
8.16	Integral from 0.2 to 0.8 and the value of $R^{-1}$ for the common lepton channel at reconstruction level for energy flow in data for the SM model. . . . .	126
8.17	Integral from 0.2 to 0.8 and the value of $R^{-1}$ for the common lepton channel at reconstruction level for energy flow in MC for the colour octet $W$ model. . . . .	126
8.18	Integral from 0.2 to 0.8 and the value of $R^{-1}$ for the common lepton channel at reconstruction level for $p_T$ flow in MC for the SM model. . . . .	127
8.19	Integral from 0.2 to 0.8 and the value of $R^{-1}$ for the common lepton channel at reconstruction level for $p_T$ flow in data for the SM model. . . . .	127
8.20	Integral from 0.2 to 0.8 and the value of $R^{-1}$ for the common lepton channel at reconstruction level for $p_T$ flow in MC for the colour octet $W$ model. . . . .	128

# 1. Introduction

We investigate how colour-connected hadron jets resulting from the decay of a top quark pair can be observed in the detector. The top quark pair is produced in  $pp$  collisions at a centre of momentum energy  $\sqrt{s} = 13$  TeV. Observations are conducted at the CMS experiment of the CERN LHC. Particular focus is on the light jets resulting from the decay of the  $W$  boson. They are colour-connected and experimentally we could infer about it indirectly. We also study the decay of a hypothetical colour octet  $W$  boson. In this case the light jets are no longer colour-connected and we can use these results to compare to the colour-connected case.

We use a method where the pull angle [1] is observed. This method has been applied at the  $D\bar{0}$  experiment of the Fermilab Tevatron [2], at Run I in ATLAS [3] and also at Run II in ATLAS [4]. This method was first applied at CMS by Seidel, M. et al [5] but the 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 [6]).

Also used is an adaptation of a methodology used at the Large Electron Positron Collider LEP (hereinafter referred to as the “LEP method”) wherein jet constituents are projected onto inter-jet planes ([7], [8], [9]). This method has never been applied at the LHC.

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.

This thesis shows 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.

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. This work is the first contribution of Latvia to the experimental programme of the CERN LHC.

## 2. Physical Background

The top quark is a third generation quark with charge  $\frac{2}{3}e$ , where  $e$  is the elementary electrical charge – the magnitude of charge carried by the electron. Its place in the Standard Model is shown in Fig. 2.1. The Standard Model is the most widely accepted model to describe subatomic physics with experimental verification to spectacular agreement.

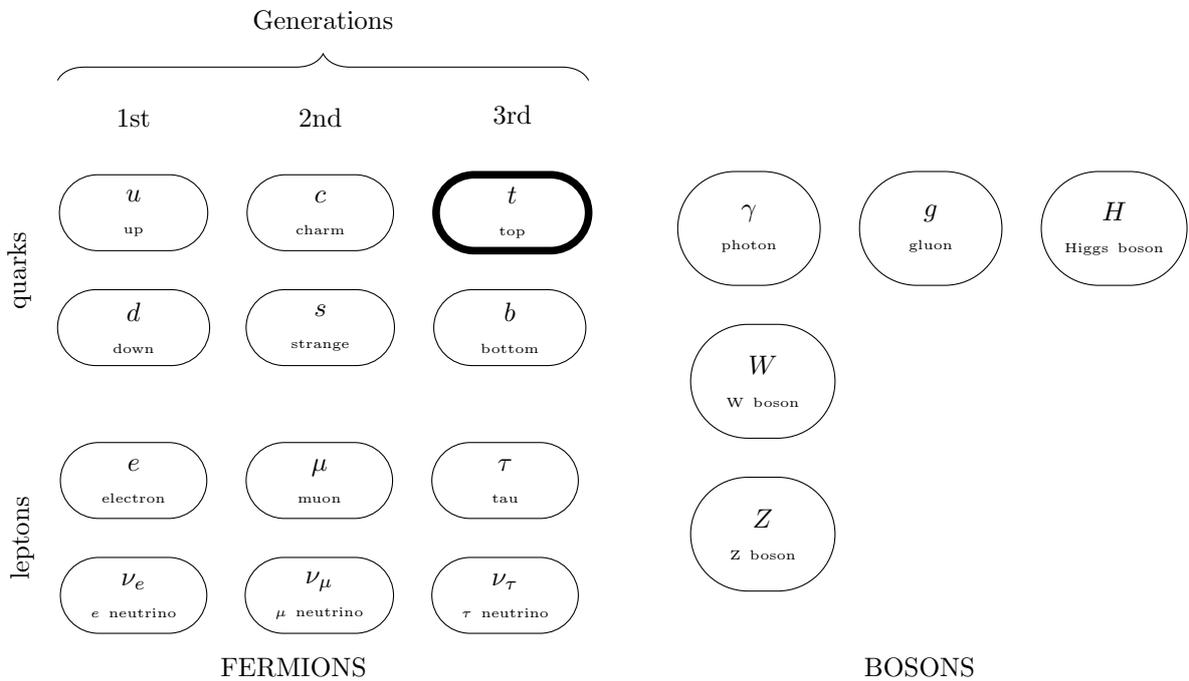


Figure 2.1. The top quark in the Standard Model.

The existence of the top quark and also its counterpart – the bottom quark was predicted by Kobayashi and Maskawa in 1973 to explain the CP violations in the decay of the kaon [10]. The discovery of the top quark was announced in 1995 by two experiments at the Fermilab Tevatron – CDF [11] and DØ [12] based on observations at  $\sqrt{s}=1.8$  TeV and integrated luminosity  $\sim 50\text{--}67$  fb $^{-1}$ .

A worldwide combination of integrated luminosity from ATLAS, CMS, CDF and DØ gives the measurement of the top quark mass of  $173.34 \pm (\text{stat}) 0.27 \pm 0.71 (\text{syst})$  GeV. The top quark is the heaviest of all known particles. It is heavier than the Higgs boson (126 GeV) and much heavier than the bottom quark whose mass is estimated around 4.2 GeV [13].

The cross section of top quark pair production in proton-proton collisions at  $\sqrt{s}=13$  TeV is measured to be  $803 \pm 2 (\text{stat}) \pm 25 (\text{syst}) \pm 20 (\text{lumi})$  pb [14]. The cross section increases once

the centre of mass energy is increased as illustrated in Fig. 2.2. At lower energies  $p\bar{p}$  colliders are better than  $pp$  colliders at producing the top quark pair.

The top quark is assumed to have occurred naturally as a constituent of the Quark-Gluon Plasma in the first picosecond after the Big Bang [15]. The stars are too cold to produce the top quark. The top quark would be produced in an environment where  $k_B T > m_t$ . Thus it is likely that the colliders on Earth are the only places in the Universe where the top quark is synthesised.

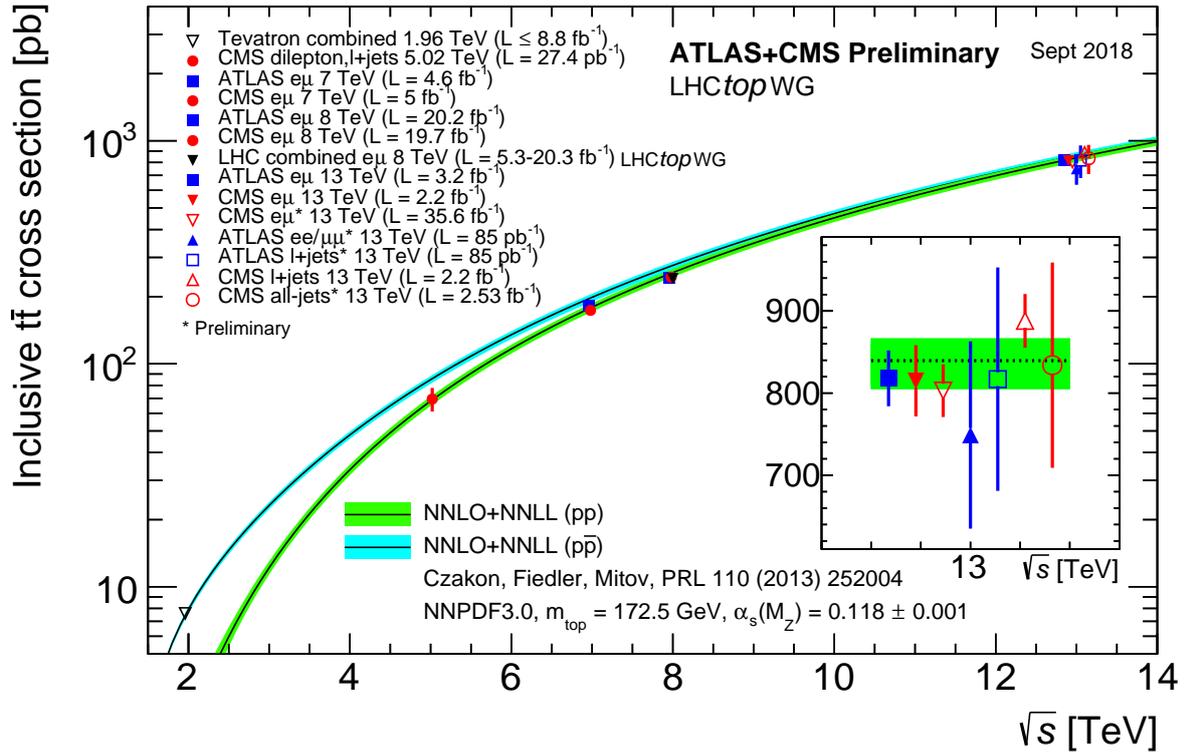


Figure 2.2. Inclusive cross section of the top pair at different centre of mass energies [16]. The result shows the cross section at  $pp$  and  $p\bar{p}$  collisions. The results from CMS and ATLAS are given for different channels of the decay of the  $t\bar{t}$  pair.

Using the relationship

$$N = \sigma \int L(t) dt \quad (2.1)$$

where  $N$  is the number of  $t\bar{t}$  pairs,  $\sigma$  is the  $t\bar{t}$  cross section,  $L$  is the instantaneous luminosity, at  $35.9 \text{ fb}^{-1}$  integrated luminosity  $26.7 \times 10^6$  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. 2.3 illustrates 2 such scenarios. The gluon exchanged is so energetic as to smash the proton into debris. Such a collision is called inelastic.

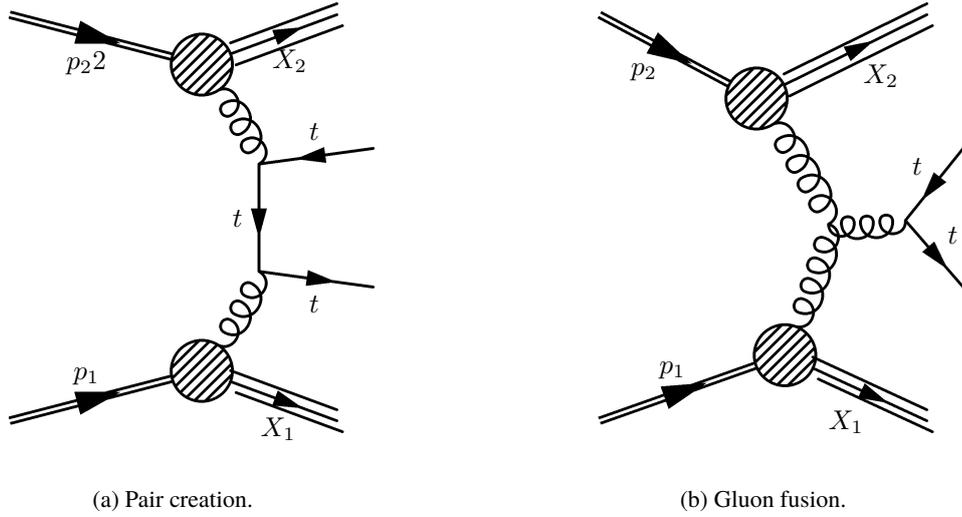


Figure 2.3. Top quark pair production in a  $pp$  collision.

The top quark decays exclusively in the weak decay process. The diagrams of the top quark decay are shown in Fig. 2.4. In the weak decay  $W$  boson and a quark of different flavour and magnitude of electrical charge  $1/3e$  is emitted.

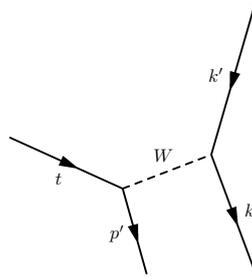


Figure 2.4. Weak decay of the top quark  $t$ .  $p'$  is a quark of different flavour and  $k$  and  $k'$  are fermions resulting from the decay of the  $W$  boson.

The average of measurements by CDF,  $D\bar{O}$  experiments of Tevatron [17], and ATLAS and CMS experiments of LHC [16] yield the result of the  $|V_{tb}|$  term of the Cabibbo–Kobayashi–Maskawa matrix

$$|V_{tb}| = 1.009 \pm 0.031. \quad (2.2)$$

This implies 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 are very small [18]:

$$|V_{td}| = 8.4 \times 10^{-3}, \quad |V_{ts}| = 40.0 \times 10^{-3}. \quad (2.3)$$

The width of the top quark as measured by the  $D\bar{O}$  collaboration [19] with  $2.3 \text{ fb}^{-1}$  of integrated luminosity is  $\Gamma = 1.99_{-0.55}^{+0.69} \text{ GeV}$ . This translates into a lifetime of  $\tau = 3.3 \times 10^{-25} \text{ s}$ .

This lifetime is smaller than the hadronisation timescale ( $1/\Lambda \sim 10^{-24}$  s), where  $\Lambda^2$  is the value of  $Q^2$  of the exchanged gluon at which the strong coupling constant  $\alpha_s$  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 [20].

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 2.1.

Table 2.1

Branching ratios from the decay of the  $W$  boson.

Mode	$\Gamma_j/\Gamma$	$\Gamma_j/\Gamma$
	simplified	observed [18]
$e\nu_e$	$1/9$	$(10.71 \pm 0.16) \%$
$\mu\nu_\mu$	$1/9$	$(10.63 \pm 0.15) \%$
$\tau\nu_\tau$	$1/9$	$(11.38 \pm 0.21) \%$
pair of quarks	$2/3$	$(67.41 \pm 0.27) \%$

colour-connected jets are emitted in the hadronic decay of the  $W$  boson (Fig. 2.6). 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. 2.5, which is based on the Lund model [21]. An alternative portrayal based on Feynman diagrams is given in Fig. 2.7.

The species of particles created in the case of a hadronic decay of the  $W$  boson are listed in Table 2.2.

Table 2.2

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
$\pi^0$	135.0	$8.5 \times 10^{-27}$	2 $\gamma$ absorbed at ECAL
$\pi^\pm$	139.6	$2.6 \times 10^{-8}$	tracker, ECAL, HCAL showers
$K_S^0$	497.6	$8.95 \times 10^{-11}$	ECAL, HCAL showers
$K_L^0$	497.6	$5.1 \times 10^{-8}$	ECAL, HCAL showers
$K^\pm$	493.7	$1.2 \times 10^{-8}$	tracker, ECAL, HCAL showers
$n$	939.6	881.5	ECAL, HCAL showers
$p$	938.3	$\infty$	tracker, ECAL, HCAL showers

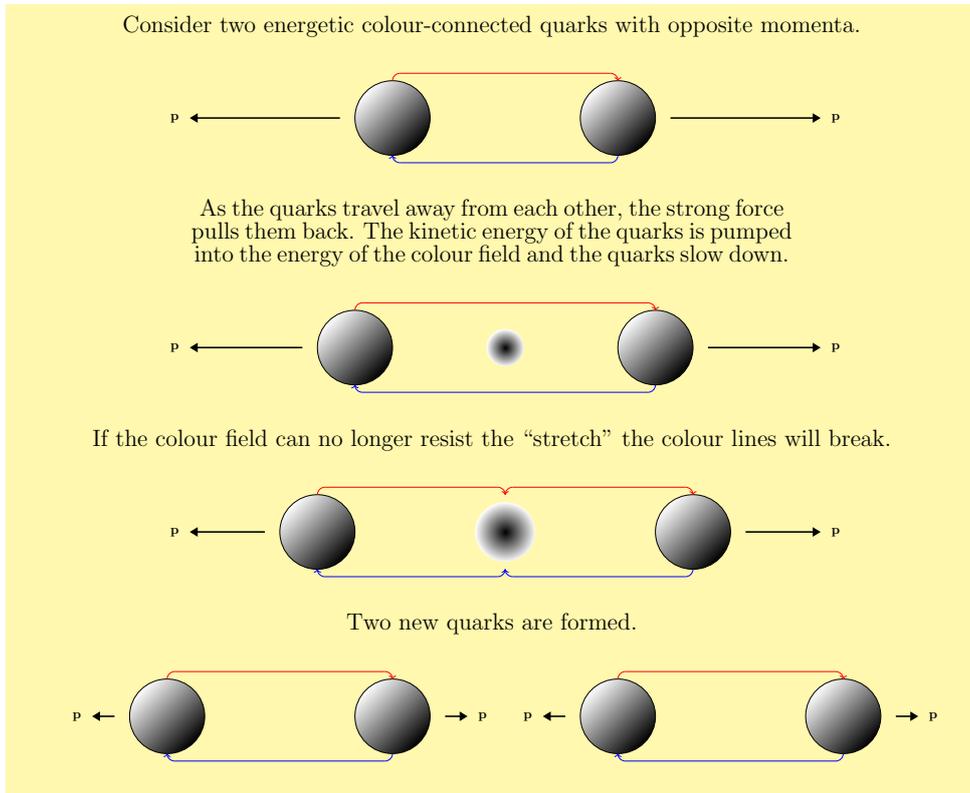


Figure 2.5. Creation of new hadrons by two energetic quarks.

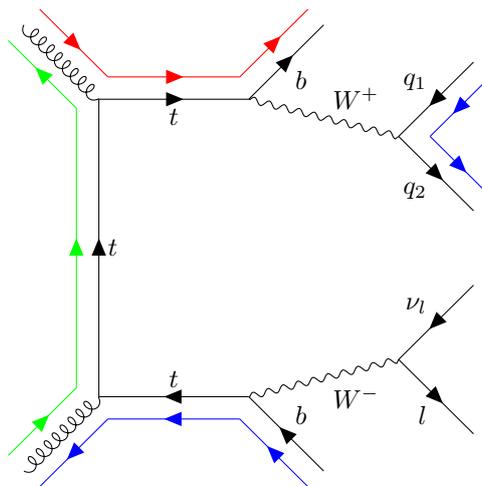


Figure 2.6. Colour flow in the decay of a top quark pair.

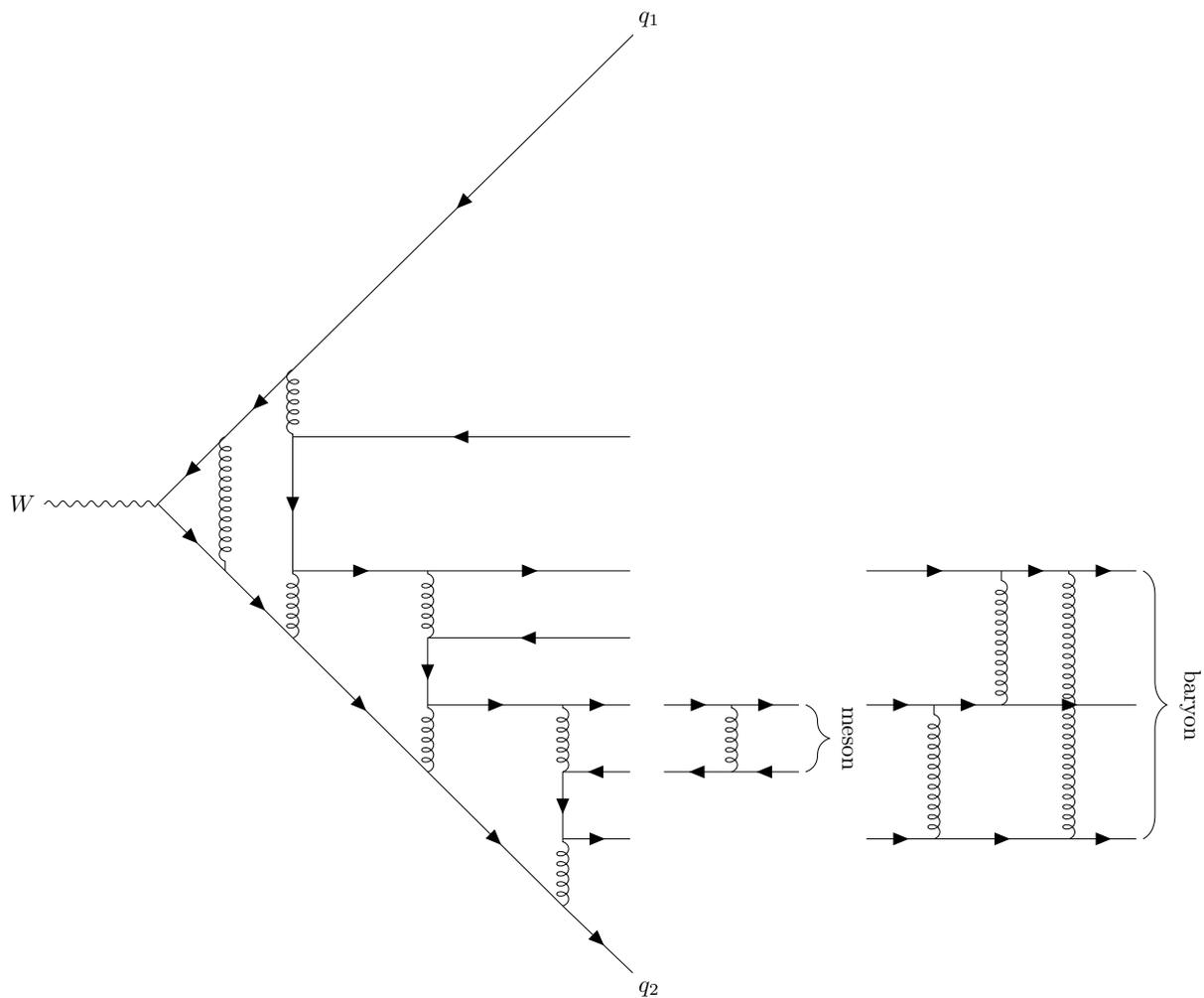


Figure 2.7. Creation of hadrons in the colour field of two quarks.

The respective resonances are clearly discernible at the generation level (Fig. 2.8). Only the neutral pion decays before being directly observed in the detector.

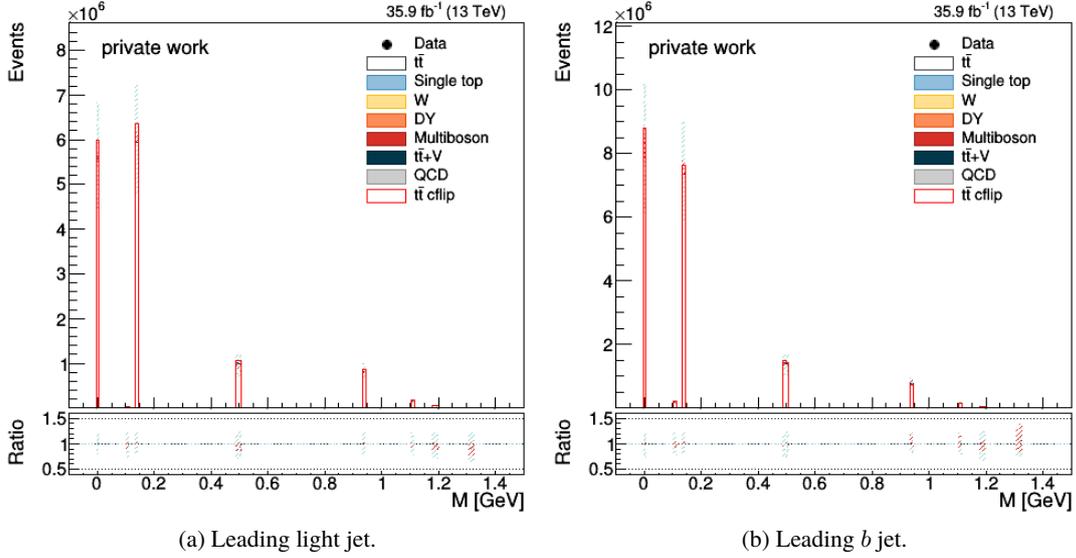


Figure 2.8. Resonances at generation level corresponding to particles listed in Table 2.2 constituting the leading light jet and the leading  $b$  jet. In both cases all jet constituents are included.

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. 8.

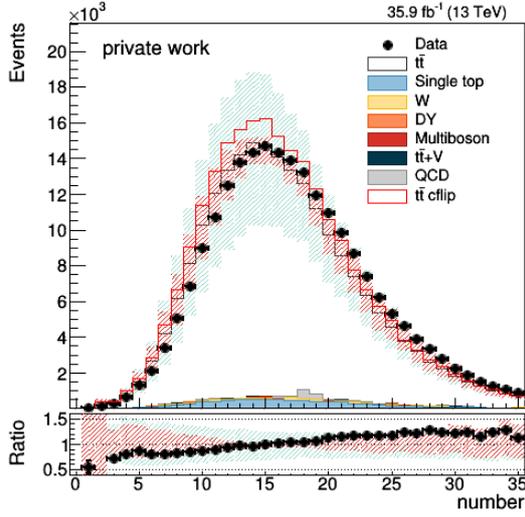
The distribution of the number of particles that constitute the leading light jet and the leading  $b$  jet is shown in Fig. 2.9, the distribution of the ratio of the number of electrically charged particles to the total number of particles is given in Fig. 2.10, and the distribution of the ratio of the energy of electrically charged particles to the total energy of particles is given in Fig. 2.11. The leading light jet is the jet from the decay of the  $W$  boson that has the highest transverse momentum  $p_T$  while the leading  $b$  jet is either of the  $b$  jets that has the highest transverse momentum.

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  $W \rightarrow qq'$  events as the leptonic decays do not have colour flow. It is hard to trigger on resolved  $W \rightarrow qq'$  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.

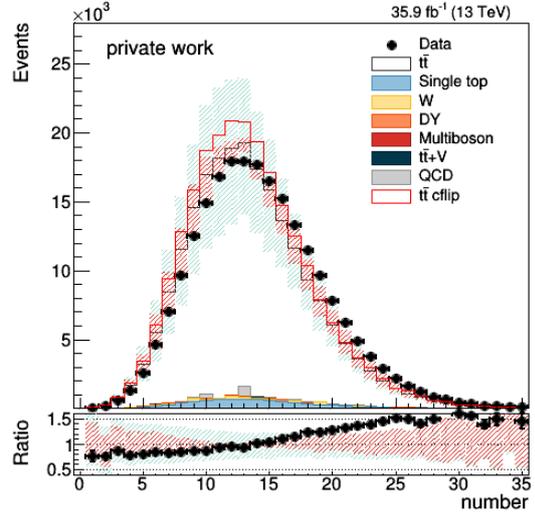
In the case of a hadronic decay the  $W$  boson decays to the colour singlet:

$$\frac{1}{\sqrt{3}} (R\bar{R} + G\bar{G} + B\bar{B}), \quad (2.4)$$

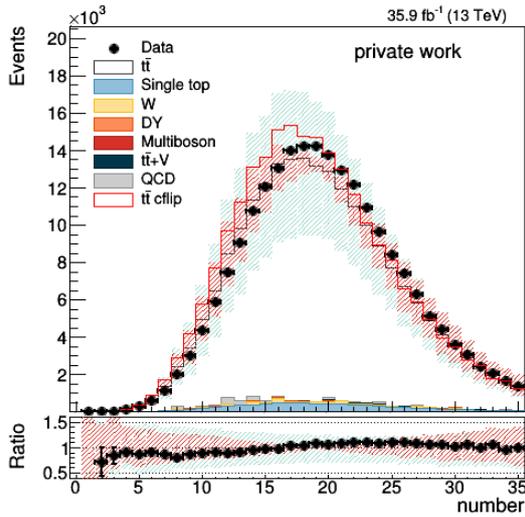
where  $R$ ,  $G$  and  $B$  are the three quantum states of the colour wave function. We mention this feature in light of our subsequent discussion of the colour octet  $W$  boson.



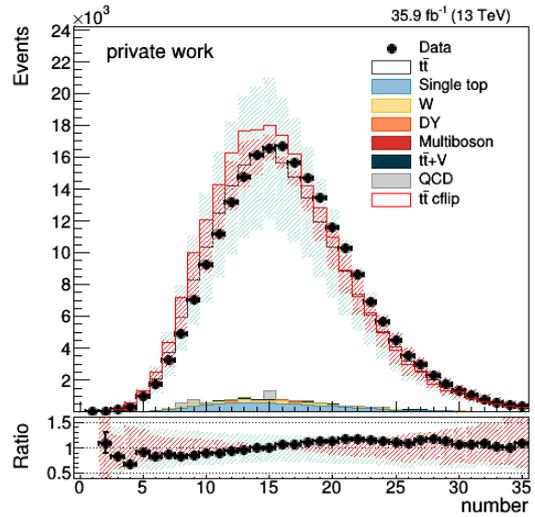
(a) Leading light jet.



(b) Second leading light jet.

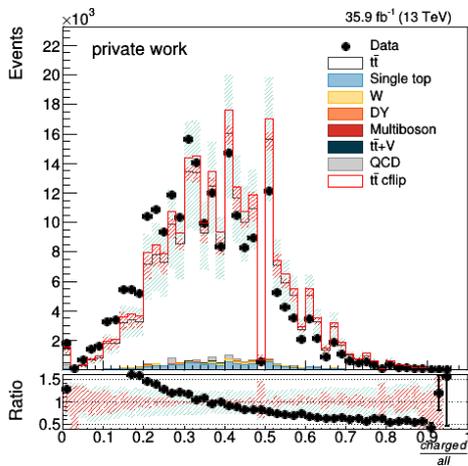


(c) Leading  $b$  jet.

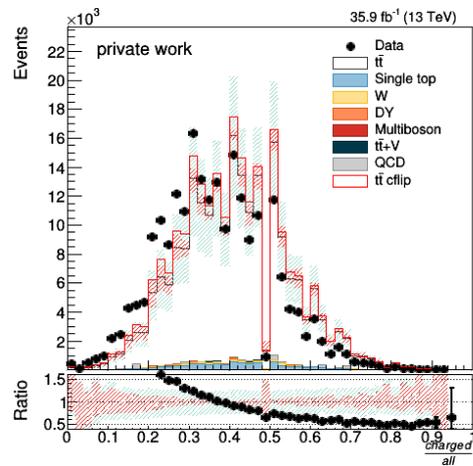


(d) Second leading  $b$  jet.

Figure 2.9. Total number of particles constituting the leading light jet, the second leading light jet, the leading  $b$  jet and the second leading  $b$  jet. In all cases all jet constituents are included.

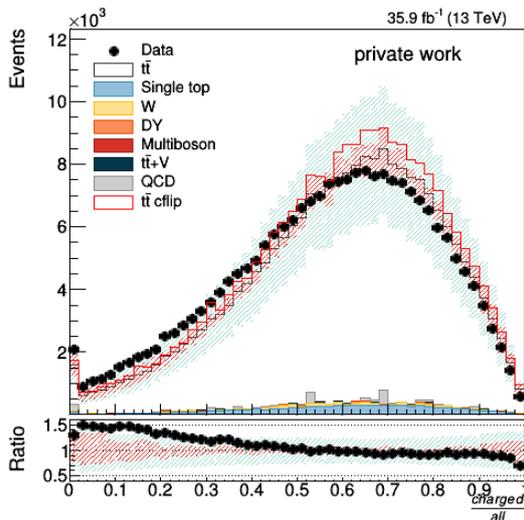


(a) Leading light jet.

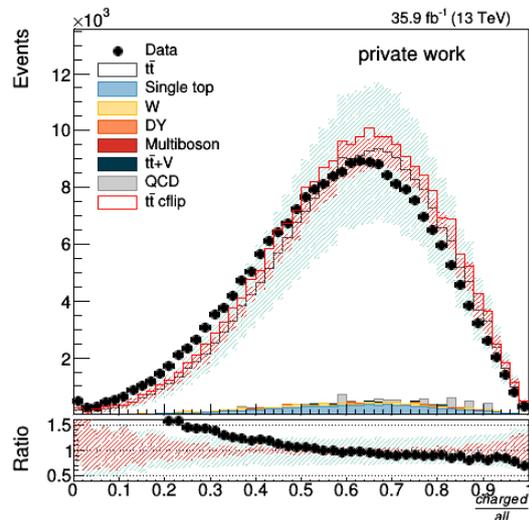


(b) Leading  $b$  jet.

Figure 2.10. Ratio of the number of charged particles to the total number of particles constituting the leading light jet and the leading  $b$  jet.



(a) Leading light jet.



(b) Leading  $b$  jet.

Figure 2.11. Ratio of the energy of charged particles to the total energy of particles constituting the leading jet and the leading  $b$  jet.

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

$$R\bar{G}, R\bar{B}, G\bar{R}, G\bar{B}, B\bar{R}, B\bar{G}, \frac{1}{\sqrt{2}}(R\bar{R} - G\bar{G}), \frac{1}{\sqrt{6}}(R\bar{R} + G\bar{G} - 2B\bar{B}). \quad (2.5)$$

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. The colour flow in the case of the colour octet  $W$  boson is shown in Fig. 2.12.

Although the existence of such a particle has not been confirmed, massive colour octet vector bosons (colourons) are predicted in a variety of models, including axigluon models, top-colour models, technicolour models with coloured technifermions, flavour-universal and chiral colouron models, and extra-dimensional models with  $KK$  gluons [22]. These states have also recently been suspected as a potential source ([23], [24]) of the top quark forward-backward asymmetry observed by the CDF collaboration ([25], [26]). Searches for resonances in the dijet mass spectrum at the LHC at  $\sqrt{s}=7-8$  TeV imply that the lower bound on such a boson is now 2–3 TeV ([27], [28], [29], [30]), while more recent LHC searches at  $\sqrt{s}=13$  TeV with integrated luminosity  $27 \text{ fb}^{-1}$  place an even higher lower bound of 6.1 TeV ([31]).

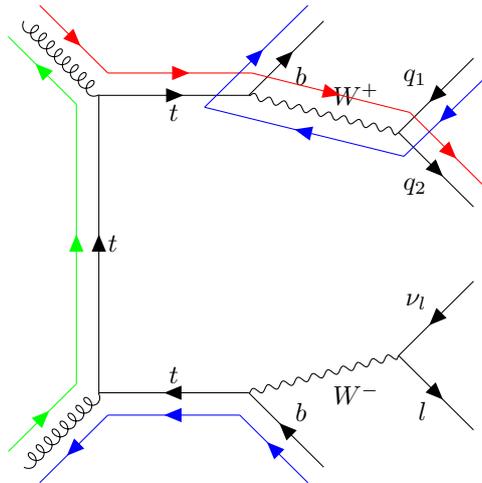


Figure 2.12. Colour flow in the decay of a top quark pair involving a hypothetical colour octet  $W$  boson.

## 3. 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 [32], [33] 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 [34], [35] 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.
- Recreate 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.

### 3.1 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. The situation of the LHC is shown in Fig. 3.1. 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.

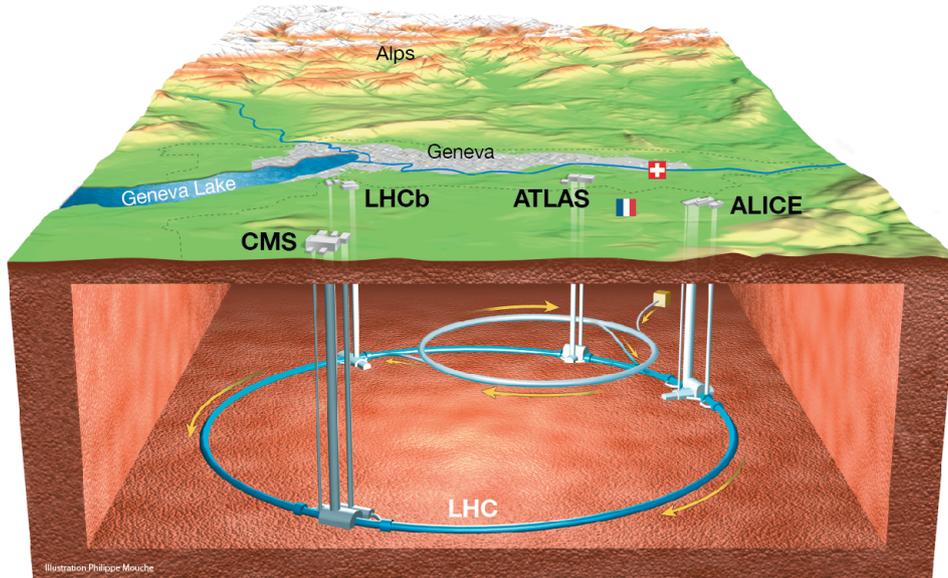


Figure 3.1. The Large Hadron Collider situated underground on the French-Swiss border in Geneva area [36].

The protons at the LHC circulate at nearly the speed of light. The per proton energy is 7 TeV, the  $\gamma$  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:

- 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).

The CERN accelerator complex is sketched in Fig. 3.2.

After the protons have been fully accelerated they are allowed to circulate in the LHC – the LHC is a storage ring. There are  $1.15 \times 10^{11}$  protons in each bunch and 2808 bunches in circulation. The revolution frequency is 11.245 kHz [38]. 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

## CERN's Accelerator Complex

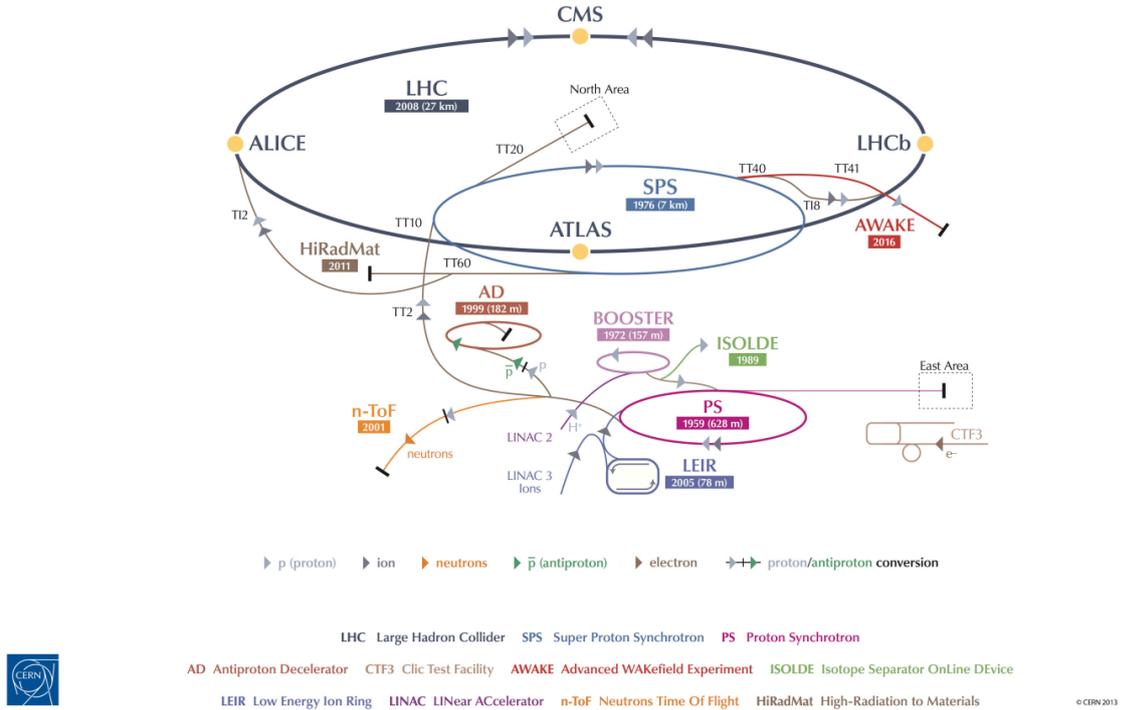


Figure 3.2. The CERN accelerator complex [37].

helium <sup>1</sup>. 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 [39], [40].

The design COM of the LHC is 14 TeV. In its first data taking period from 2010–2013 it operated at  $\sqrt{s} = 7\text{--}8$  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  $\sqrt{s} = 13\text{--}14$  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.

<sup>1</sup>The lambda point of helium is the temperature at which the normal fluid helium makes transition to a superfluid helium.

### 3.2 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 [41] is designed to operate in diverse physics programmes in the TeV range. It is an onion-type detector covering  $4\pi$  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. The shape of the CMS detector is a cylinder. It has endcaps on both ends while the central 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. 3.3.

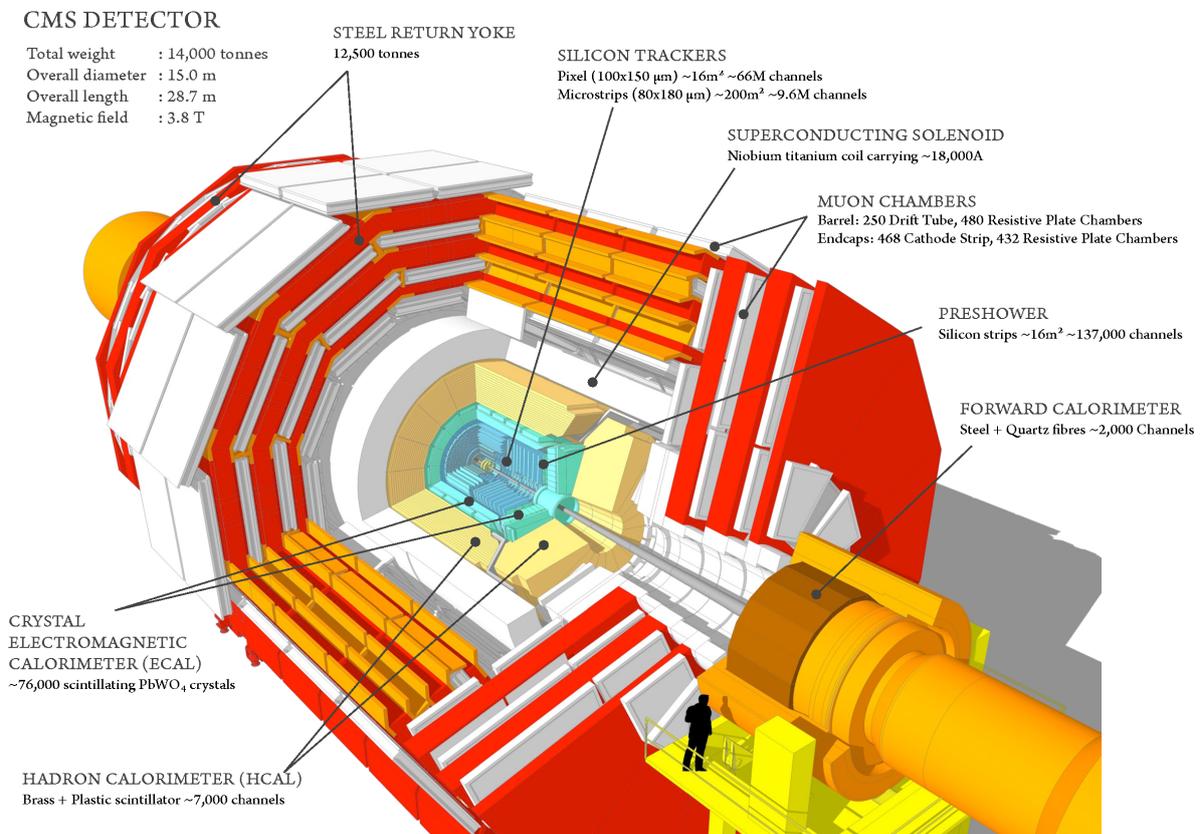


Figure 3.3. A cut-away view of the CMS detector [42].

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. 3.4, 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 [43].

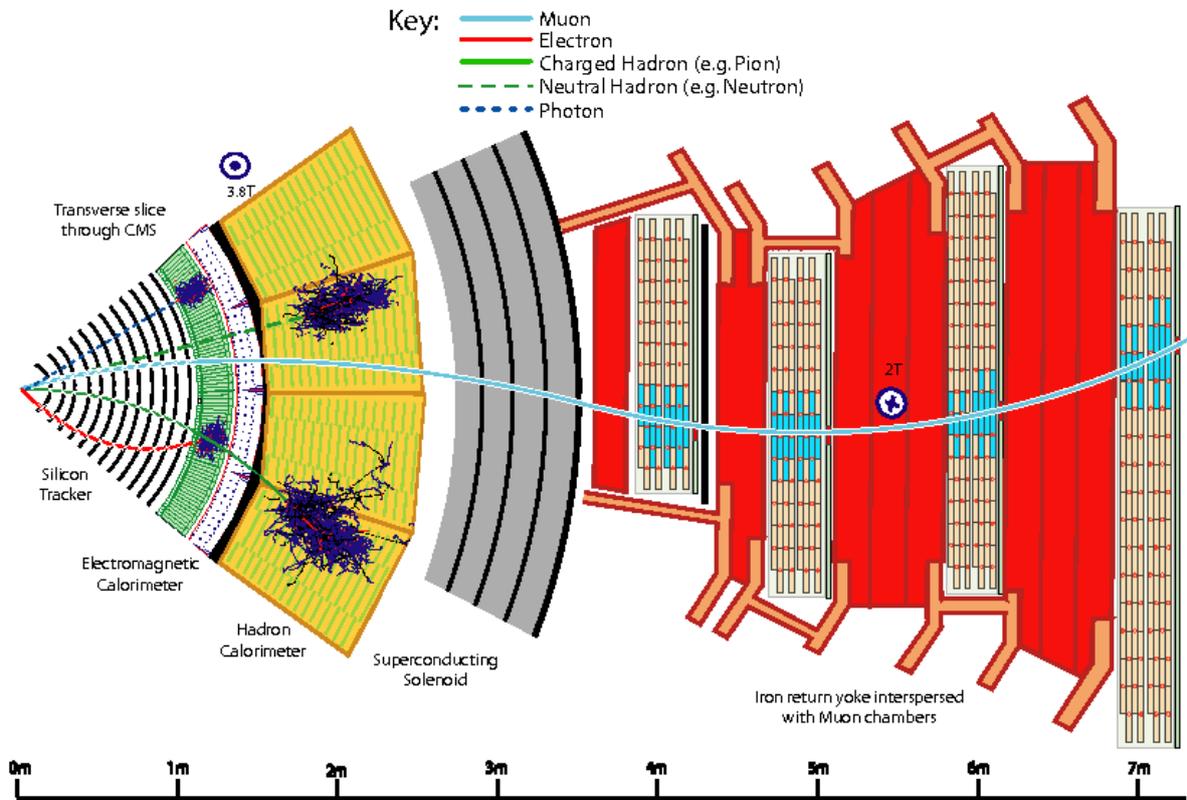


Figure 3.4. A sketch of the specific particle interactions in a transverse slice of the CMS detector, from the beam interaction region to the muon detector [43].

The fine-granularity and fast response tracker [44], [45] 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  $100 \text{ kHz/mm}^2$  is encountered. In order to achieve the desired resolution  $100 \mu\text{m} \times 150 \mu\text{m}$  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  $10 \text{ cm} \times 80 \mu\text{m}$  to  $25 \text{ cm} \times 150 \mu\text{m}$ , the size increasing with an increasing radius. There are 66 million pixels with  $1 \text{ m}^2$  active area in the pixel detector and 9.3 million strips and  $193 \text{ m}^2$  active area in the strip detector.

The electromagnetic calorimeter of CMS (ECAL) is a hermetic homogeneous calorimeter made of 61 200 lead tungstate ( $\text{PbWO}_4$ ) 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  $\text{PbWO}_4$  crystals exhibit characteristics that make them an appropriate choice for an electromagnetic calorimeter at the LHC. The high density  $8.28 \text{ g/cm}^3$ , 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  $\text{PbWO}_4$  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  $\eta$ - $\phi$ , corresponding to a front cross section  $22 \times 22 \text{ mm}^2$  and a rear cross section  $26 \text{ mm}^2 \times 26 \text{ mm}^2$ . 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  $30 \times 30 \text{ mm}^2$ , a front face cross section  $28.62 \times 28.62 \text{ mm}^2$  and a length of 220 mm corresponding to  $24.7 X_0$ . 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 [46]. 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 [47] 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 solenoid 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  $90^\circ$  is 5.82 interaction lengths ( $\lambda_I$ ). 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  $0.17 \times 0.17$  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  $11.6 \text{ kJ/kg}$ .

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 [48]. 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 [49] with radius parameter  $R = 0.4$  as implemented by the FASTJET [50] package. Distance  $d_{ij}$  between jets is determined by using  $p = -1$  in the general formula:

$$d_{ij} = \min(k_{Ti}^{2p}, k_{Tj}^{2p}) \frac{\Delta_{ij}^2}{R^2}, \quad (3.1)$$

where  $\Delta_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$  and  $k_{Ti}$ ,  $y_i$ ,  $\phi_i$  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.

## 4. Methodology

### 4.1 Pull angle

An explanation about the coordinate system used at CMS is in order. The CMS uses a coordinate system centred on the nominal collision point, the  $x$  points towards the centre of the LHC, the  $y$  axis points upwards and the  $z$  axis points along the beam in the direction of the Jura mountains. The coordinate system is illustrated in Fig. 4.1.  $\rho$  is the radial coordinate. The azimuthal angle  $\phi$  is measured from the  $x$  axis to the projection of the spatial vector  $\mathbf{p}$  in the  $x$ - $y$  plane. The polar angle  $\theta$  is measure from the positive direction of the beam to the vector  $\mathbf{p}$ . Pseudorapidity  $\eta$  is defined as

$$\eta \equiv -\ln\left(\frac{\theta}{2}\right). \quad (4.1)$$

The pseudorapidity is equal to

$$\eta = \ln\left(\frac{p + p_L}{p - p_L}\right), \quad (4.2)$$

where  $p$  is the magnitude of  $\mathbf{p}$  and  $p_L$  is the longitudinal component of  $\mathbf{p}$  along the direction of the beam. A measurement related to pseudorapidity is the rapidity  $y$  defined as:

$$y \equiv \ln\left(\frac{E + p_L}{E - p_L}\right), \quad (4.3)$$

where  $E$  is the energy of the particle. For massless particles rapidity and pseudorapidity are equal. For our present purposes, rapidity and pseudorapidity can be used interchangeably without a loss of accuracy.

We adopt the methodology proposed by [1] to use the pull angle to reveal colour connection between two quark jets. The pull angle  $\theta_p$  formed by the pull vector  $\vec{v}_p$  and difference between two jets  $\vec{J}_2 - \vec{J}_1$  is shown in Fig. 4.2. The  $\phi$ - $y$  coordinate system is used.

The pull vector is given by the formula

$$\vec{v}_p = \sum_{i \in J} \frac{p_T^i |\vec{r}_i|}{p_T^J} \vec{r}_i, \quad (4.4)$$

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.

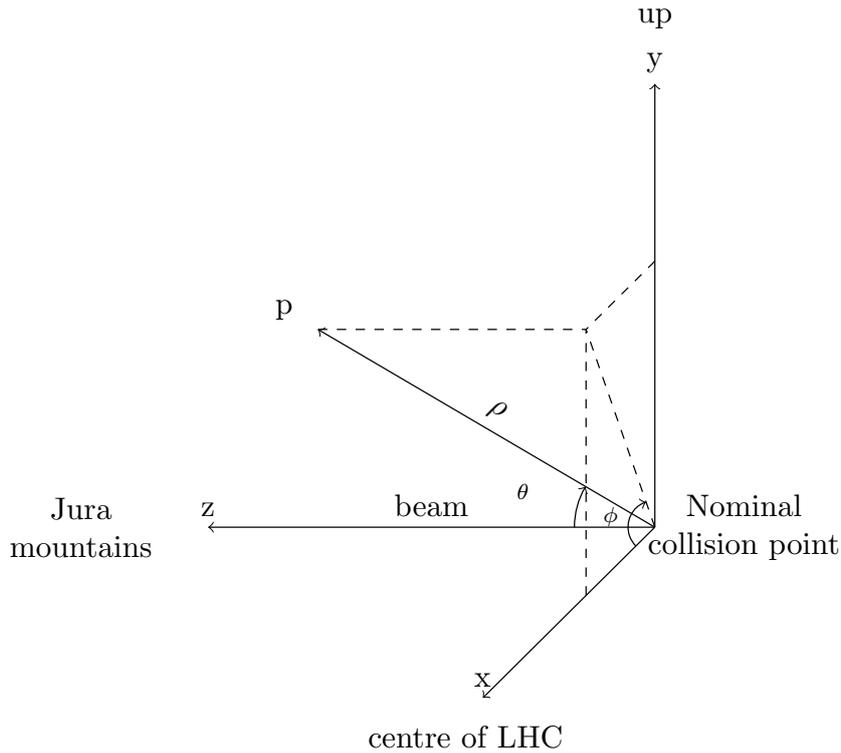


Figure 4.1. The coordinate system used at CMS.

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  $D\bar{O}$  experiment of Tevatron [2] and the ATLAS experiment at the LHC in Run I [3] and in Run II [51]. We hope to outperform all results with the methodology of the pull angle with the state-of-the-art tracker of the CMS detector immersed in the 4 T magnetic field of the superconducting solenoid.

The anti- $k_T$  clustering algorithm ensures a conical jet shape in case the jet separation  $\Delta R$  is more than double of the parameter  $R$ , which is set at 0.4 at CMS. This case is illustrated in Fig. 4.3(a). In case of separation between jets  $\Delta R$  being less than double of the parameter  $R$  the hard jet will wean constituents from the soft jet. This is illustrated in Fig. 4.3(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  $\Delta R > 2R$ .

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. 4.4 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 1.0 GeV as the threshold and exclude particles whose  $p_T$  is below it from our analysis.

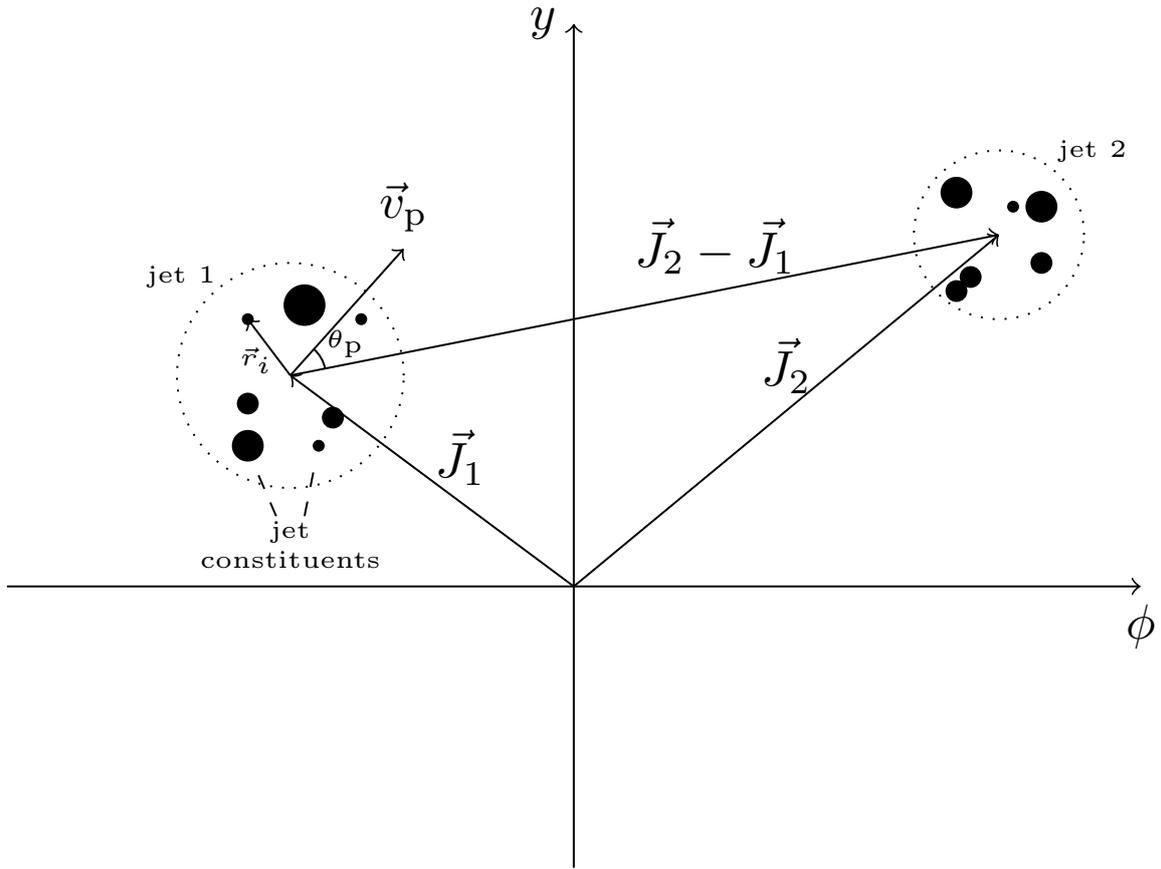


Figure 4.2. Pull angle  $\theta_p$  and pull vector  $\vec{v}_p$  in a  $\phi$ - $y$  plane.

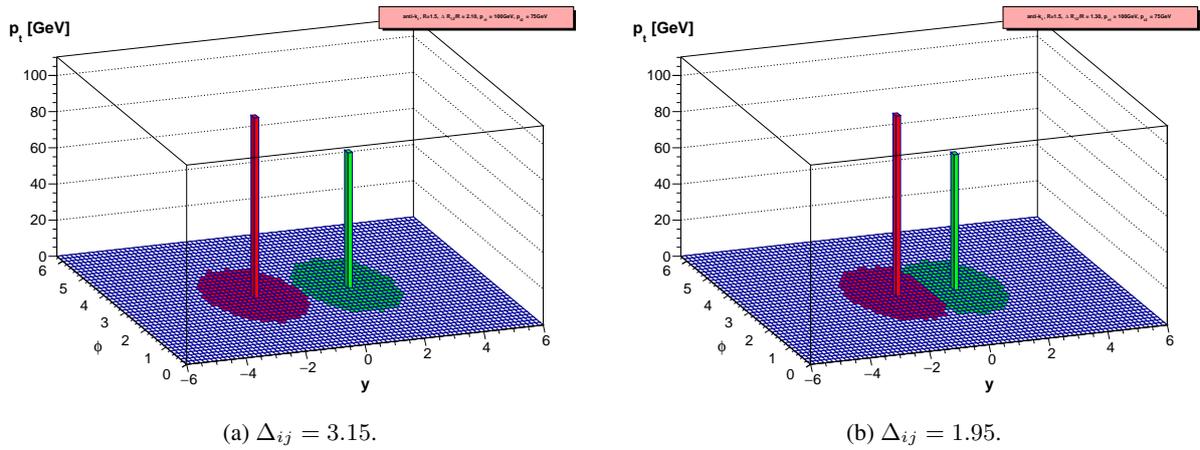


Figure 4.3. Jet shapes obtained with the anti- $k_T$  clustering.  $R = 1.5$  is used. Two cases are shown –  $\Delta_{ij} = 3.15$  and  $\Delta_{ij} = 1.95$ . The  $p_T$  of the hard jet is 100 GeV while the  $p_T$  of the soft jet is 75 GeV. Courtesy of Cacciari, Salam and Soyez [52].

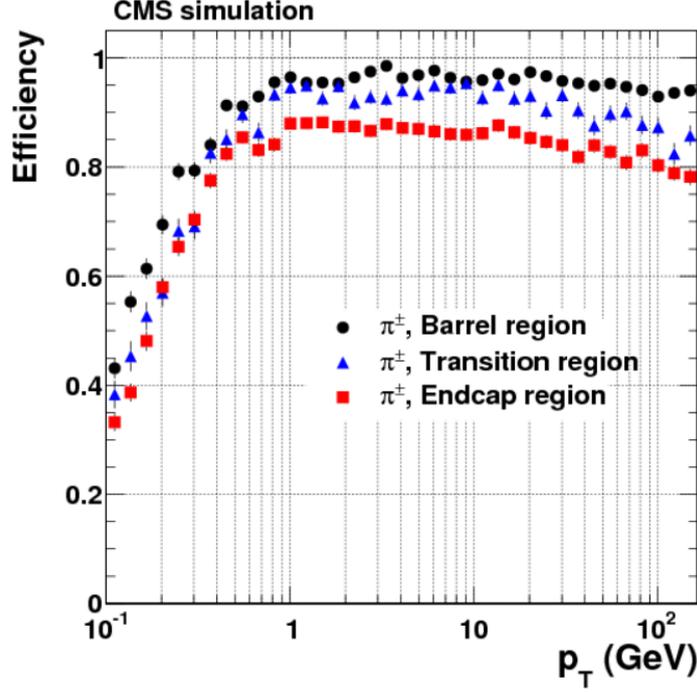


Figure 4.4. Track reconstruction efficiencies for pions passing the high-purity quality requirements. Results are shown as a function of  $p_T$ , for the barrel, transition, and endcap regions, which are defined by  $\eta$  intervals of 0–0.9, 0.9–1.4 and 1.4–2.5, respectively [53].

## 4.2 LEP method

Another methodology of studying colour-connected jets in the process  $e^+e^- \rightarrow q\bar{q}q\bar{q}$  at  $\sqrt{s} = 189\text{--}207$  GeV was used in various experiments of LEP [8], [7], [9]. 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. 4.5. Particles are projected onto these planes and the angle with the leftmost quark  $\chi_1$  is taken. If this angle is less than the angle  $\chi_0$  between the quarks forming the plane (which means the particle is projected between the respective quarks) then the normalised angle  $\chi_R = \chi_1/\chi_0$  is plotted in the region corresponding to the plane after a linear transformation

$$\chi = \chi_R + n_{\text{plane}} - 1 \quad (4.5)$$

has been performed on the normalised angle.

In the semileptonic decay of the top quark pair an arrangement as shown in Fig. 4.5 is not possible. Therefore a modification as shown in Fig. 4.6 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 hadronic  $b$  jet  $j_h^b$ , 2) the hadronic  $b$  jet  $j_h^b$  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. 4.6(b) and Fig. 4.6(c) we may hope to observe colour reconnection effects.

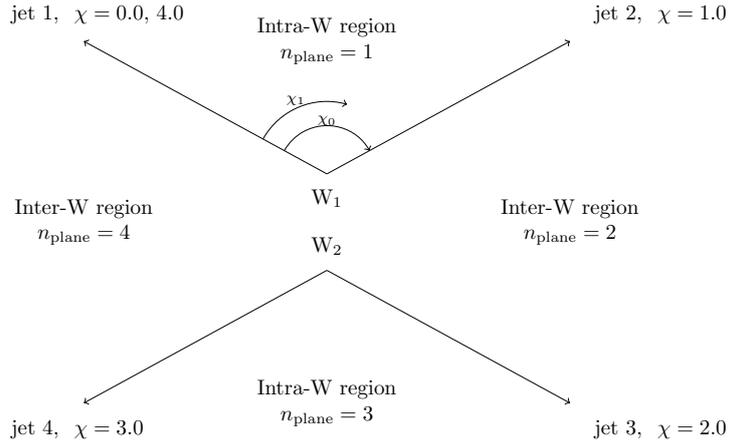


Figure 4.5. Inter- $W$  and intra- $W$  planes in the process  $e^+e^- \rightarrow q\bar{q}q\bar{q}$  and the relative angle  $\chi_R = \chi_1/\chi_0$ .

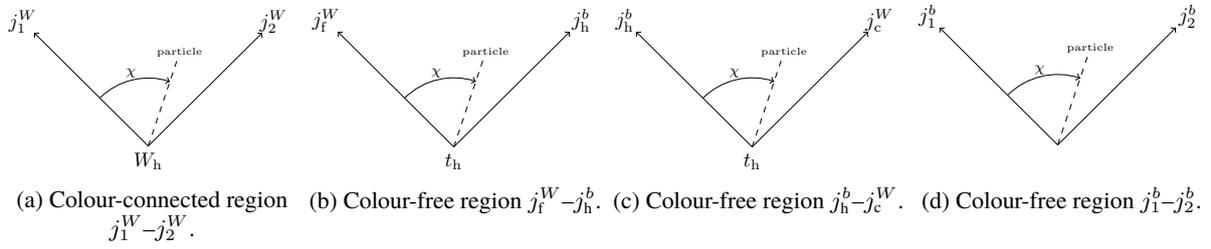


Figure 4.6. Adaptation of the LEP method to  $t\bar{t}$  semileptonic decay involving a colour-connected region and 3 colour-free regions.

The method calls for a separation of  $b$  quarks associated with hadronic or leptonic decay of the  $W$  boson (hadronic and leptonic  $b$  quarks, in our parlance). 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.

## 5. Data and MC Samples

The discussion of this section is elaborated upon [54] and [55] 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  $35.9 \text{ fb}^{-1}$  for all the channels analysed. The luminosity has been computed with the `BRILCALC` tool [56] 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 5.1. The different denominations of the data sets correspond to a different release of the reconstruction module and trigger menus in CMSSW [57] – the collection of software that is used in simulation, calibration, alignment and reconstruction so that it is possible to perform a physics analysis.

Table 5.1

Primary datasets used in this analysis. PD is an abbreviation for SingleMuon or SingleElectron [55].

Primary dataset	Integrated luminosity
/PD/Run2016B-23Sep2016-v3/MINIAOD	$35.9 \text{ fb}^{-1}$
/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	

Data is compared to samples simulated by Monte Carlo packages [58], [59] that use pseudo-random numbers to reproduce quantummechanical probabilities of a process. They rely on the numerical technique developed by Stanisław Ulam [60], [61]. The use of Monte Carlo samples in a physics analysis allows the separation of signal from background processes and to see how accurately we can model the processes in the real world.

The hardest emission of the nominal  $t\bar{t}$  sample is first generated by the `POWHEG` method [62] using full NLO accuracy and taking spin correlations of the decay products of the top quark into account [63]. The showering was implemented in `PYTHIA 8.2` [64] that is based on dipole-style  $p_T$  ordered evolution. `POWHEG` and `PYTHIA 8.2` are interfaced via the Les Houches Accord

(LHA) [65]. Les Houches Event Files (LHEF) [66], [67] are used to transfer information about the particles generated by POWHEG in the hard process of interest ( $t\bar{t}$ ) to PYTHIA 8.2. The tune CUETP8M2T4 [68] is used for PYTHIA 8.2 that specifically aims to specify the parameters of colour reconnection to produce new samples for top mass measurement. The generated events are reconstructed with the CMS detector simulation based on GEANT4 [69].

Madgraph5\_aMC@NLO [70] is a popular choice to generate hard radiation to the NLO order for background processes.

The list of simulated samples can be found in Table 5.2. They are from the RunIISummer16MiniAODv2-PUMoriond17\_80X\_mcRun2\_asymptotic\_2016\_TrancheIV\_v6 production.

The cross sections that we use are theoretical predictions. Practically, they are obtained from [71] and [72] except for  $t\bar{t}$  for which the generator cross section is quoted according to [73]. At NNLO the expected  $t\bar{t}$  cross section is  $832^{+20}_{-29}$  (scale)  $\pm 35$  (PDF +  $\alpha_s$ ) [74]. We use the NNLO reference to normalise all  $t\bar{t}$  samples.

Table 5.2

List of simulation samples. We quote the cross section used to normalise the sample in the analysis. Adapted after [55].

Process	Dataset	$\sigma$ [pb]
<b>Signal</b>		
$t\bar{t}$	TT_TuneCUETP8M2T4_13TeV-powheg-pythia8	832
<b>Background</b>		
$t\bar{t} + W$	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
$W + \text{jets}$	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
$\mu$ 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	QCD_Pt-300to470_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	797.35
	QCD_Pt-470to600_MuEnrichedPt5_TuneCUETP8M1_13TeV_pythia8	45.83

Table 5.2

Continued.

	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
	QCD_Pt-30to50_EMEnriched_TuneCUETP8M1_13TeV_pythia8	6 493 800.0
	QCD_Pt-50to80_EMEnriched_TuneCUETP8M1_13TeV_pythia8	2 025 400.0
<i>e</i> enriched	QCD_Pt-80to120_EMEnriched_TuneCUETP8M1_13TeV_pythia8	478 520.0
QCD	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 background samples contain events of processes that are different from the  $t\bar{t}$  signal event but whose reconstructed final state passes the same selection criteria as applied to the signal. For example, consider the Drell–Yan (DY) process shown in Fig. 5.1. The DY process occurs in  $pp$  collisions when a pair of oppositely charged leptons ( $e$  or  $\mu$ ) is created from the decay of an uncharged boson. The uncharged boson in turn is created from quark fusion. The cross section for the DY process is much larger than the cross section of the  $t\bar{t}$  process. The DY would form a significant background if we selected only one lepton as the other lepton could be misreconstructed as a jet. However, the importance of the DY background drops significantly once we impose the additional requirement of having at least 4 jets. A comprehensive account of the background processes is given in [75].

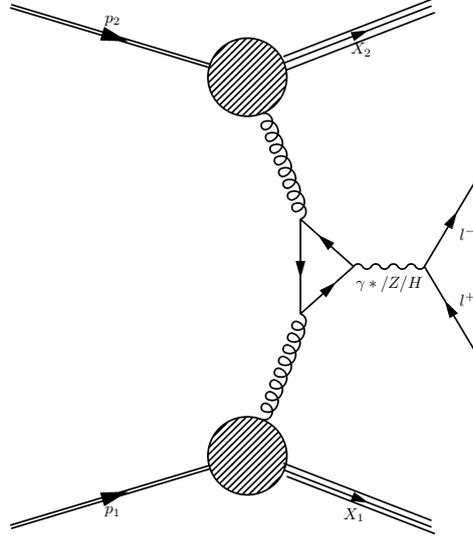


Figure 5.1. The Drell–Yan process.

Single  $t$  and single  $\bar{t}$  backgrounds have a negligible effect in the final selection stage and these samples are ignored at this stage of the analysis.

For creating the colour octet samples hard-scatter signal events were generated using POWHEG-Box v2 [76]. The colour strings in the LHE files are switched in such a way that one

quark from the decay of the  $W$  boson is colour-connected to the  $t$  quark while the other one to the  $b$  quark. The  $W$  bosons and  $t$  quarks had to be removed from the LHE files lest that PYTHIA complain of unphysical colour flow.

Initially the generator level selection of the fiducial phase space used in this analysis was implemented in a RIVET [77] routine to ease future comparison with new generators and tunes. Afterwards the colour flipped dataset with 12 million events was produced. A PYTHIA 8.2 tune adopted by the TOP group to control the number of jets in 13 TeV simulations [78] was used. This tune uses a  $h_{\text{damp}}$  factor equal to 1.5 times the top mass. The  $h_{\text{damp}}$  factor suppresses POWHEG real emissions by a factor  $\frac{h^2}{p_T^2+h^2}$ . The colour octet sample is listed in Table 5.3. We will occasionally refer to the colour octet  $W$  sample as the  $t\bar{t}$  *cflip* sample.

Table 5.3

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

Process	Dataset	$\sigma$ [pb]
	<b>Background</b>	
Colour octet $W$ boson	TT_TuneCUETP8M2T4_13TeV-powheg-colourFlip-pythia8	832

## 5.1 Corrections applied to the simulation

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

**Pile-up re-weighting.** During each bunch crossing there are multiple  $pp$  collisions. The reconstructed tracks are combined into vertices, each vertex marking the spot of a  $pp$  collision. Pile-up refers to the number of  $pp$  collisions in each bunch crossing. A pile-up recorded in a real event is shown in Fig. 5.2.

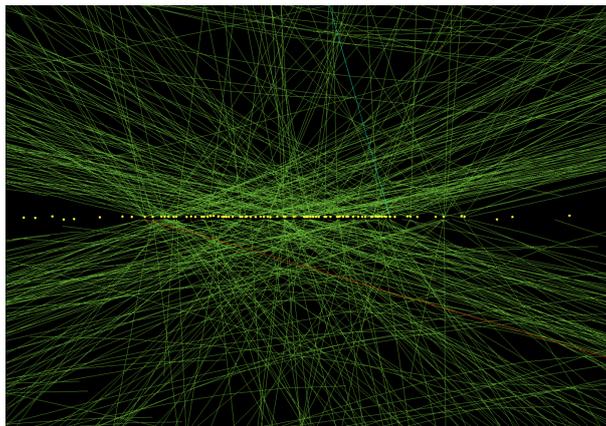


Figure 5.2. The  $\rho$ - $z$  view of an event in a high pile-up run 198609 showing 78 reconstructed vertices [79].

However, the example in Fig. 5.2 is extreme. Pile-up depends on the number of protons in the bunch and the beam emittance and has been varying at different times of the operation of

the LHC. Pile-up during different periods of the operation of the LHC is shown in Fig. 5.3. The CMS detector has been designed with an average pile-up of 25 collisions in mind but the HL-CMS will have to be prepared for a pile-up of 140–200 collisions [80].

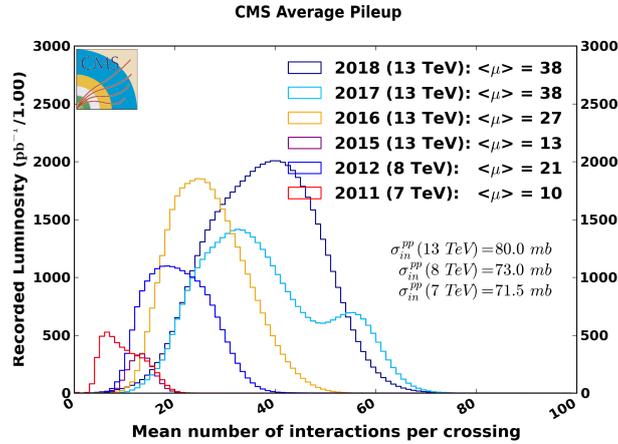


Figure 5.3. Pile-up during different periods of the operation of the LHC [81].

When preparing the MC samples also the pile-up value is determined. Additional minimum bias interactions are superimposed in order to include the effect of in-time (originating from the same bunch crossing) and out-of-time (originating from the previous bunch crossing) pile-up in the events. Minimum bias refers to events that are selected with a “loose” trigger that accepts a large fraction of the overall inelastic cross section [82]. The generated pile-up distribution is based on the configuration

SimGeneral.MixingModule.mix\_2016\_25ns\_Moriond17MC\_PoissonOOTPU\_cfi.

A minimum bias cross section of 69 mb is used to estimate the pile-up distribution following the recommendations from [83]. A 5 % uncertainty assigned to the minimum bias cross section assumed.

Fig. 5.4 shows the distribution of data compared to expectation in the number of primary vertices reconstructed. The agreement between data and Monte Carlo is not perfect, and is poorer in runs BCDEF than in GH. In order to equalise the pile-up distributions between data and MC each MC event is assigned a pile-up weight:

$$w_{\text{pu}} = \frac{N_{\text{pu}}^{\text{data}}}{N_{\text{pu}}^{\text{MC}}}. \quad (5.1)$$

**Lepton identification and isolation efficiency.** An efficiency of an algorithm to select a physics object according to a criterion is defined as

$$\epsilon \equiv \frac{\text{N of objects that pass the criterion actually passed by the algorithm}}{\text{N of objects that pass the criterion}}. \quad (5.2)$$

The efficiency is a measure of the ability of our real-world selection algorithm to select a physics object compared to an idealised selection algorithm that will select interesting physics objects flawlessly.

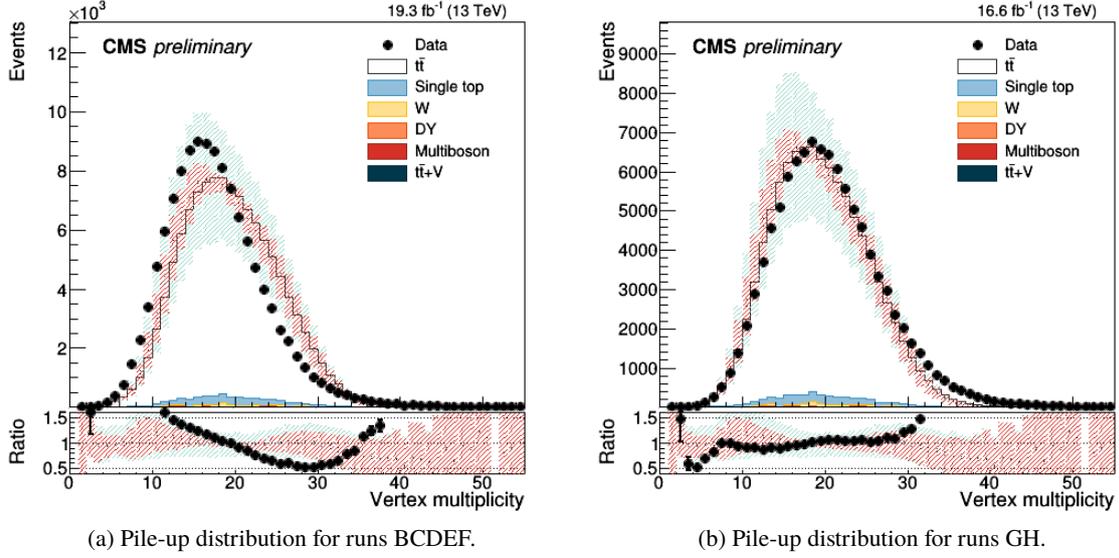


Figure 5.4. Pile-up distributions in Monte Carlo and data after the last step of event selection.

The lepton identification and isolation efficiency is measured using the tag-and-probe method [84]. This method uses known mass di-object resonances like  $Z$ ,  $J\Psi$  and  $\Upsilon$ . The “tag” is an object that passes a very tight set of criteria. The fake rate of a tag should be very small –  $\ll 1\%$ . The probe is the other object in the resonance selected according to the particular selection criteria that are much looser than the selection criteria used for the tag. The efficiency is measured as:

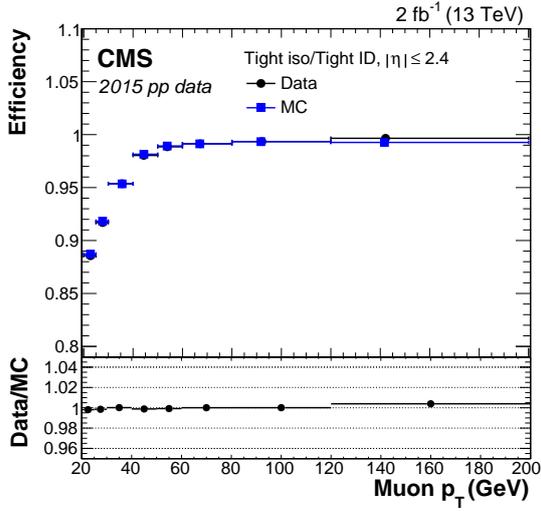
$$\epsilon = \frac{N_{\text{pass}}}{N_{\text{all}}}, \quad (5.3)$$

where  $N_{\text{pass}}$  is the number of probes passing the selection criteria, while  $N_{\text{all}}$  is the total number of probes counted.

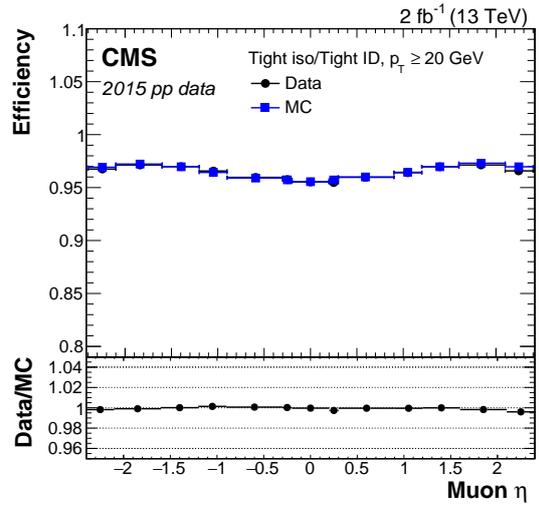
The identification efficiency refers to the ability to identify a physics object  $X$  when it actually is a physics object  $X$ . For example, a jet with large electromagnetic content could be misidentified as an electron. Or hadron shower remnants could penetrate the muon system (punch-through) and could be misidentified as a muon. The identification efficiency depends on the  $p_T$  and  $\eta$  of the physics object. Particularly it degrades for low  $p_T$ . Fig. 5.5 shows the identification efficiency of electron as a function of  $p_T$  for different ranges of  $\eta$ .

Muon isolation is used to distinguish prompt muons from weakly decaying jets. It is used for muons that have already passed the identification criterion. Muon isolation is evaluated relative to the muon  $p_T$  by summing up energy in a geometrical cone  $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$  surrounding the muon. The muon isolation efficiency degrades with lower  $p_T$ . Fig. 5.7 shows the muon isolation efficiency as a function of the muon  $p_T$  and  $\eta$ .

We correct for the difference in performance for the lepton identification in electrons and lepton identification and isolation in muons between data and simulation, by applying a  $p_T$ ,  $\eta$ -dependent scale factor. As it will be detailed later (Chap. 6) we make use of tight

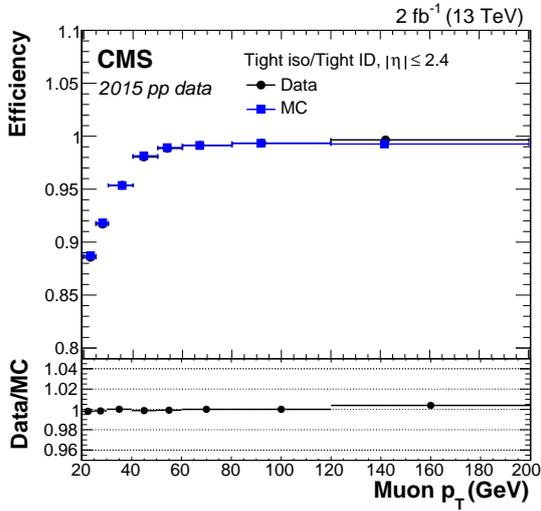


(a)  $\eta < 0.8$ .

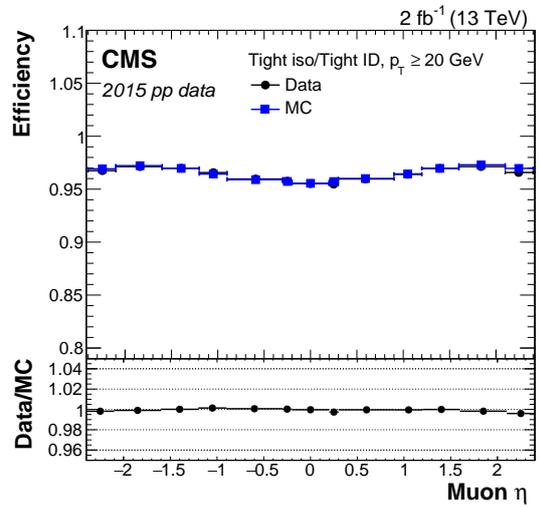


(b)  $< 1.57\eta^2$ .

Figure 5.5. Efficiency as a function of electron  $p_T$  for dielectron events in data (dots) and DY simulation (triangles), for the medium working point of the sequential selection [85].



(a) Tag-and-probe efficiency for the tight PF isolation working point on top of the tight ID versus  $p_T$  for muons in the acceptance of the muon spectrometer.



(b) Tag-and-probe efficiency for the tight PF isolation working point on top of the tight ID versus  $\eta$  for muons with  $p_T > 20$  GeV.

Figure 5.6. Tag-and-probe efficiency for the tight PF isolation working point on top of the tight ID versus  $p_T$  for muons in the acceptance of the muon spectrometer, and versus  $\eta$  for muons with  $p_T > 20$  GeV, for 2015 data (circles), simulation (squares), and the ratio (bottom inset). The statistical uncertainties are smaller than the symbols used to display the measurements [86].

muons [87] and electrons [88]. As data-to-MC scale factors we use the official values recommended by the POGs [89, 90].

**Trigger efficiency.** Trigger efficiency is measured on leptons that have already passed the ID (electrons) or ID+isolation selection (muons). We correct for the difference in performance of the High Level Trigger (HLT) [91], [92] used in data and simulation. The HLT trigger receives events from the L1 trigger at a maximum design rate 100 kHz, in practice less than a third of this value. It then further selects interesting events according to various physics programmes. It is completely implemented in software and is continuously updated. The HLT is run on a farm of about 1000 commodity PCs. For different physics programmes separate trigger paths are used. The HLT further reduces the event rate by a factor of about 1000. Its output rate is  $\mathcal{O}(1) - \mathcal{O}(100)$  Hz – a rate that is acceptable by recording devices. The HLT selection is a multistage process<sup>1</sup>. The first step uses the calorimeter information. In the second step energy deposits in the calorimeter are combined with hits in the pixel detector. In the third step full track reconstruction is used. Each step of the trigger contributes to a loss of trigger efficiency.

Trigger efficiency is a function of the  $p_T$  and  $\eta$  of the physics object. Particularly the efficiency drops at low  $p_T$ . Let us consider the isolated single-muon trigger as an example. Fig. 5.7 shows the dependence of the efficiency of this trigger as a function of  $p_T$  and  $\eta$ .

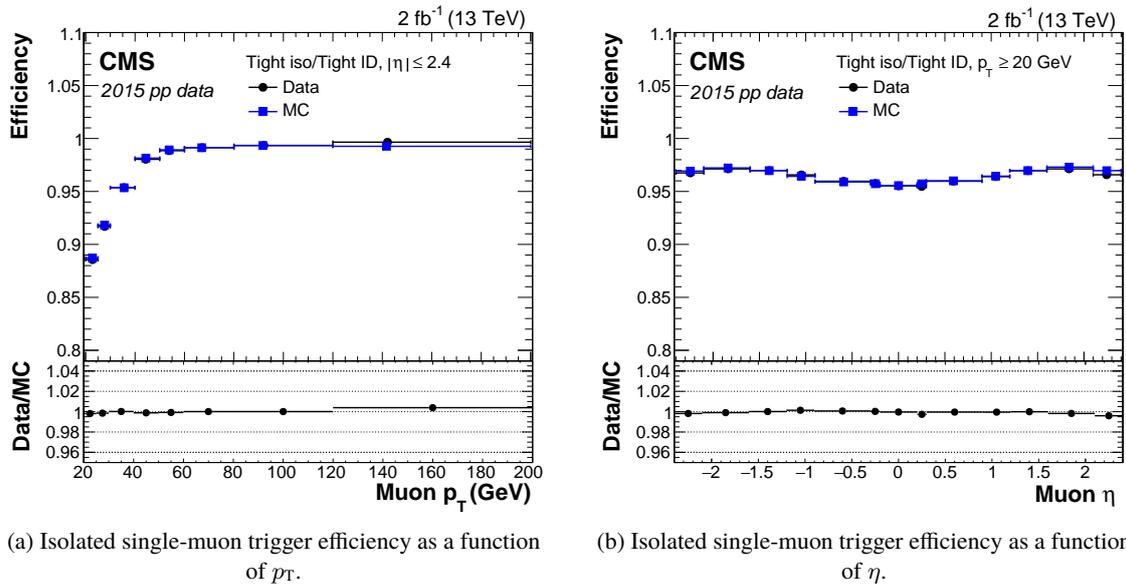
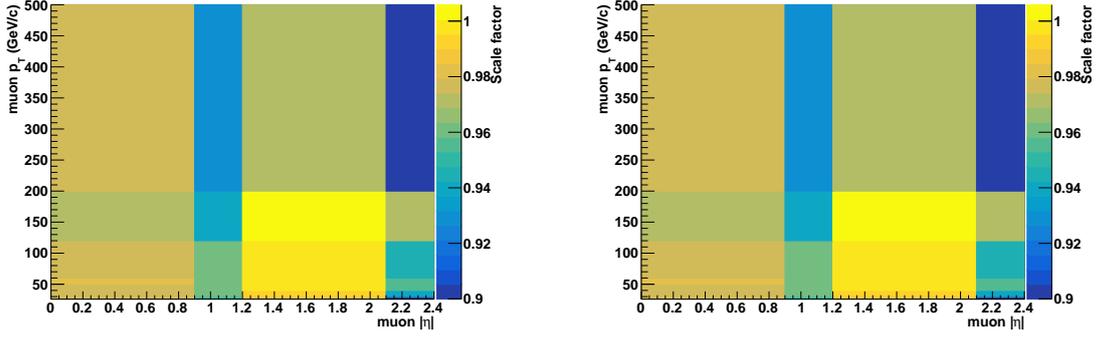


Figure 5.7. Isolated single-muon trigger efficiencies measured with 2015 data (squares), simulation (circles), and the ratio (bottom inset). Results are plotted as a function of offline reconstructed muon  $p_T$  and  $\eta$  [86].

Based on these differences  $p_T$  and  $\eta$  scale factors are used. They are different for data taking periods BCDEF and GH as each used different thresholds for tracking. The scale factors are shown in Fig. 5.8.

<sup>1</sup>It corresponds to a combination of Level 2 and Level 3 triggers used in other detector systems.



(a) Isolated single-muon trigger efficiency scale factors for data taking periods BCDEF.

(b) Isolated single-muon trigger efficiency scale factors for data taking periods GH.

Figure 5.8. Isolated single-muon trigger efficiency scale factors for data taking periods BCDEFGH of 2016.

The values used for the correction of trigger efficiencies are the ones proposed by the Top Trigger group [93].

**Generator level weights.** The generator level weights  $w_{\text{gen}}$  are assigned by the generator. They are equal to 1 for POWHEG but can be different from unity for other generators and systematic variations.

The simulated processes are scaled according to their effective integrated luminosity which can be computed from the sum of the per-event weights. The general formula applied to determine the number of events expected for a given process ( $\hat{N}$ ) is:

$$\hat{N} = \mathcal{L} \cdot \sigma \cdot \frac{\sum_{i=1}^{N_{\text{sel}}} w_i}{\sum_{i=1}^{N_{\text{gen}}} w_i}, \quad (5.4)$$

where  $\mathcal{L}$  is the integrated luminosity,  $\sigma$  is a reference theory prediction for the inclusive cross section and  $w_i$  are the per-event generator level weights.

The weight assigned to an event  $w$  is a multiplicative combination of the pile-up weight  $w_{\text{pu}}$ , trigger efficiency scale factor  $\text{sf}_{\text{trigger}}$ ,  $\text{sf}_{\text{lepton ID+isolation}}$ , generator level weight  $w_{\text{gen}}$  and the inverse of total events generated  $N_{\text{gen}}$ :

$$w = w_{\text{pu}} \cdot \text{sf}_{\text{lepton ID+isolation}} \cdot \text{sf}_{\text{trigger}} \cdot w_{\text{gen}} \frac{1}{N_{\text{gen}}} \quad (5.5)$$

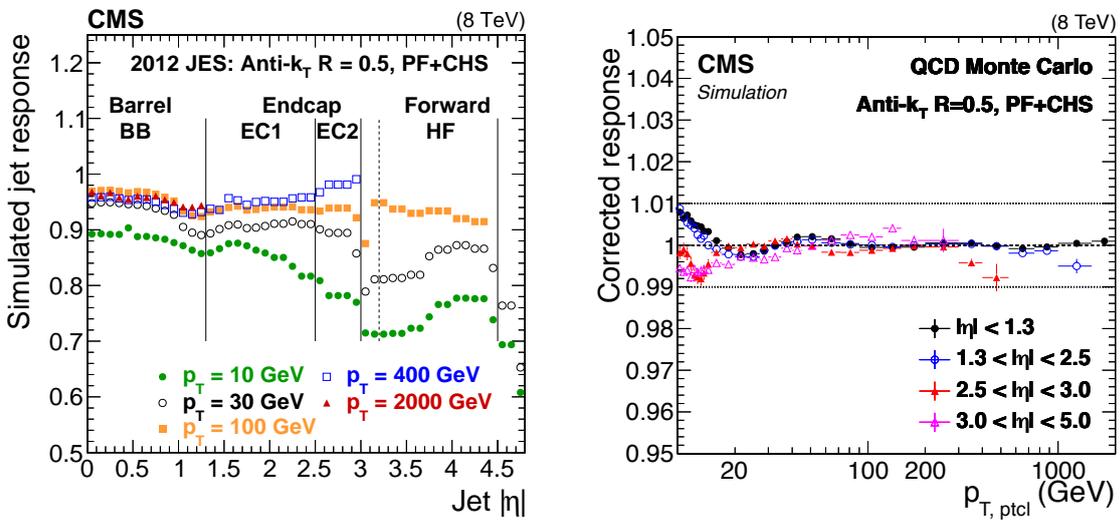
In addition to assigning event weights each distribution of an observable of sample is scaled to integrated luminosity  $\mathcal{L}$  and the theoretically predicted cross section  $\sigma$  of the process associated with the sample.

**Jet energy scale and resolution.** The jet energy scale correction (JEC) is applied to data to bring jet response to unity. Jet response  $R_{\text{ptcl}}$  is defined as the ratio of the arithmetic mean of  $p_{\text{T}}$  of reconstructed jets and the  $p_{\text{T}}$  calculated from the generated constituents of jets (particle level):

$$R_{\text{ptcl}} \equiv \frac{\langle p_T \rangle_{\text{reco}}}{\langle p_T \rangle_{\text{ptcl}}}. \quad (5.6)$$

A reconstructed jet is matched to the generated (particle level) jet if the jet separation is within half of the jet distance parameter  $R$ , where  $R \equiv \sqrt{(\eta_{\text{jet } 1} - \eta_{\text{jet } 2})^2 + (\phi_{\text{jet } 1} - \phi_{\text{jet } 2})^2}$ . If  $R = 0.5$  is used jet distance should be no more than 0.25.

The jet response is dependent on  $p_T$ ,  $\eta$  and jet size. Particularly it degrades for jet  $p_T < 30$  GeV and is lower in the endcaps as illustrated in 5.9(a). Jet energy scale corrections are applied using the so called SUMMER16\_23SEP2016V4\_DATA,MC corrections [94]. The JEC are effective to bring jet energy response to unity particularly for jets with  $p_T > 30$  GeV as shown in Fig. 5.9(b).



(a) Simulated jet response  $R$  versus  $|\eta|$  for the jet distance parameter  $R = 0.5$ .

(b) Simulated jet response  $R$ , after JEC have been applied, versus  $p_{T, \text{ptcl}}$  for jet distance parameter  $R = 0.5$  in various  $|\eta|$  regions.

Figure 5.9. Jet energy response before (a) and after (b) applying JEC. The illustration is from 2012 TeV data with  $\sqrt{s} = 8$  due to unavailability of public results with Run II data [95]. For a discussion of PF+CHS jets see text.

The PF+CHS jets referred to in Fig. 5.9 are particle flow jets to which charged hadron subtraction (CHS) has been applied. In this method tracks of charged hadrons unambiguously associated with pile-up vertices are removed before clustering jets. It is a method to mitigate in-time pile-up. CHS can remove approximately 50 % of in-time pile-up within tracker coverage.

The jet energy resolution (JER) is defined as the width of the distribution  $p_{T, \text{reco}}/p_{T, \text{ptcl}}$  determined with a gaussian fit. JEC are applied before deriving JER. Jets have the worst energy resolutions among all physics objects. JER is dependent on pile-up,  $p_T$  and  $\eta$ . Higher pile-up degrades JER due to contamination of tracks and energy deposits from other collisions. It degrades with low  $p_T$  and is worse in the endcaps than in the barrel. Fig. 5.10 shows JER versus  $p_T$  in the barrel for varying levels of pile-up.

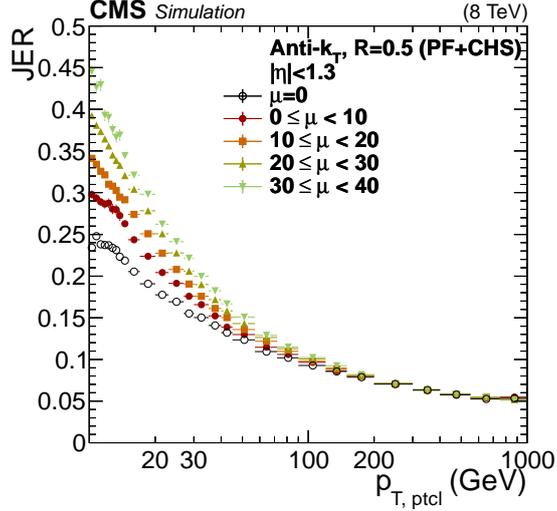


Figure 5.10. JER versus  $p_T$  in the barrel for varying levels of pile-up  $\mu$ . The results are shown separately for PF+CHS jets with jet distance parameter  $R = 0.5$  [95].

In simulation the nominal jet energy resolution is smeared using a  $p_T$ ,  $\eta$ -dependent parameterization [96]. The so called hybrid method is used. When a corresponding generated particle level jet is found the scaling method is used (Eq. 5.7).

$$c_{\text{JER}} = 1 + (s_{\text{JER}} - 1) \frac{p_T - p_{T, \text{ptcl}}}{p_T}, \quad (5.7)$$

where  $s$  is the scaling factor.

If the generated particle level jet is not found the stochastic scaling method is used (Eq. 5.8).

$$c_{\text{JER}} = 1 + \mathcal{N}(0, \sigma_{\text{JER}}) \sqrt{\max(s_{\text{JER}}^2 - 1, 0)}. \quad (5.8)$$

Fig. shows 5.11 jet energy resolution data/MC scale factor versus  $|\eta|$  for  $\gamma$  + jet data.

In both cases alternative scenarios generated by shifting the corrections according to their uncertainties are considered and shall be discussed in detail in Chapter 7.

**$b$ -tagging efficiency.** The difference in performance of the  $b$ -tagging algorithm used in the analysis is accounted for by applying a  $p_T$ -dependent scale factor. As it will be detailed later (Chap. 6), we make use of the medium working point of the Combined Secondary Vertex (v2) algorithm. The scale factors are used to correct a-posteriori the  $b$ -tagging decision in the simulation [97]. Fig. 5.12 shows the data to simulation scale factors for  $b$  jets from the hadronic and lepton side of single-lepton  $t\bar{t}$  decay.

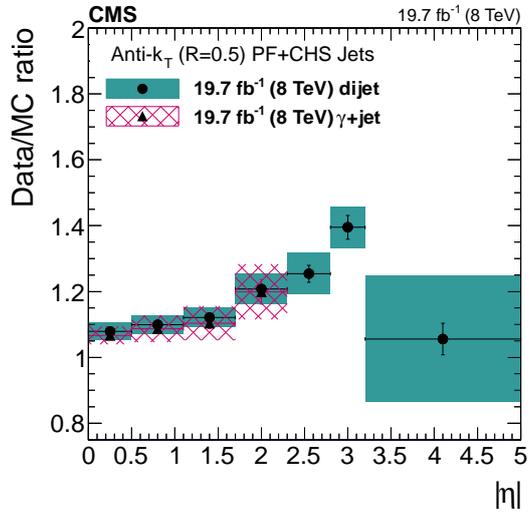


Figure 5.11. Jet energy resolution data/MC scale factor versus  $|\eta|$  for  $\gamma$  + jet data collected at 8 TeV (closed circles, solid area) compared to results at 7 TeV (open circles, dashed area) [95]. We illustrate Run I results due to unavailability of public results for Run II.

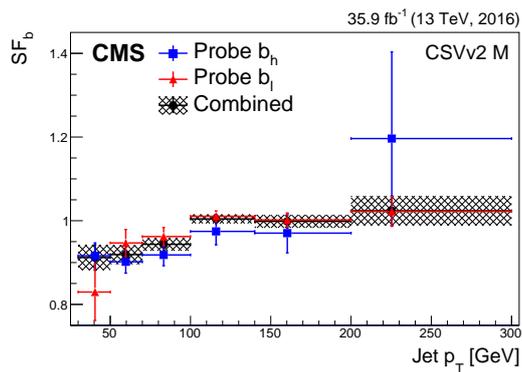


Figure 5.12. Data-to-simulation scale factors for  $b$  jets from the hadronic or leptonic side of the single-lepton  $t\bar{t}$  decay as well as for their combination, as a function of the jet  $p_T$  for the medium working point of the CSVv2 tagger [98].

## 6. 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 elaborated upon [55], which uses a similar event selection.

### 6.1 Detector level

The event selection is based on the  $t\bar{t} \rightarrow \text{lepton} + \text{jets}$  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 [43]. 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 6.1.

Offline, we require exactly one tight electron/muon with  $p_T > 34/26$  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. Additionally, the muon is required to have relative isolation of less than 0.15. The event is vetoed in the presence of a second loose lepton with  $p_T > 15$  GeV and  $|\eta| < 2.4$ . The vetoed muon is required to have relative isolation of less than 0.25.

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<sup>1</sup> (we use shorthand AK4PFchs) with  $p_T > 30$  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. Heavy quarks such as the  $b$  quark are identified by formation of secondary vertices. The  $B$  mesons have a lifetime of about 1.5 ps. It means they travel away from the primary vertex at the point of collision  $\sim 1$  cm before decaying. From the charged

---

<sup>1</sup>The method of charged hadron subtraction can be extended further by the pile-up per particle identification (PUPPI) where each particle is assigned a weight to describe the degree to which it is pile-up-like [100].

Table 6.1

Trigger paths used for online selection in the analysis.

Final state	Path	Run range	Function	L1 seed
$e + \text{jets}$	HLT_Ele32_eta2p1_WPTight_Gsf_v	all	Select $e$ with $ \eta  < 2.1$ and $p_T > 32$ GeV with the tight working point and using the GSF <sup>a</sup> to reconstruct tracks	L1_SingleEG40
				OR L1_SingleIsoEG22er OR L1_SingleIsoEG24er OR L1_SingleIsoEG24 OR L1_SingleIsoEG26
$\mu + \text{jets}$	HLT_IsoMu24_v	all	Select isolated $\mu$ with $p_T > 20$ GeV using L3 tracker algorithm <sup>b</sup>	L1_SingleMu18
	HLT_IsoTkMu24_v	all	Select isolated $\mu$ with $p_T > 20$ GeV using HLT tracker muon algorithm <sup>c</sup>	

<sup>a</sup> The bremsstrahlung energy loss distribution of electrons propagating in matter is highly non Gaussian. In such conditions the Kalman Filter which relies solely on Gaussian distributions fails. Therefore the Gaussian Sum Filter (GSF) [99] has been developed. In GSF the bremsstrahlung energy losses are modelled as a Gaussian mixture rather than single Gaussian.

<sup>b</sup> Combines muons reconstructed in HLT (Level 2) with information from the inner tracker.

<sup>c</sup> Employs an algorithm similar to the tracker muon algorithm but optimised for processing speed.

decay products of the  $B$  mesons the secondary vertex is identified. This process is illustrated in Fig. 6.1.

The CSVv2 is a retrained and optimised version of the Combined Secondary Vertex algorithm used in Run I [101], which provides discrimination also in cases when no secondary vertices are found. A higher discriminator value is associated with a higher efficiency (a higher probability that a  $b$  jet will be identified as a  $b$  jet). This is illustrated in Fig. 6.2.

However an increased efficiency comes at a cost of identifying non- $b$  objects, such  $c$ ,  $s$ ,  $u$ ,  $d$  and  $g$  jets as a  $b$  jet. The misidentification probability as a function of efficiency is identified in Fig. 6.3. A medium working point is chosen as a compromise between efficiency and misidentification probability.

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

The event yields at different selection stages are shown in Fig. 6.4 and Table 6.2. Table 6.3 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 purity of signal of our selection.

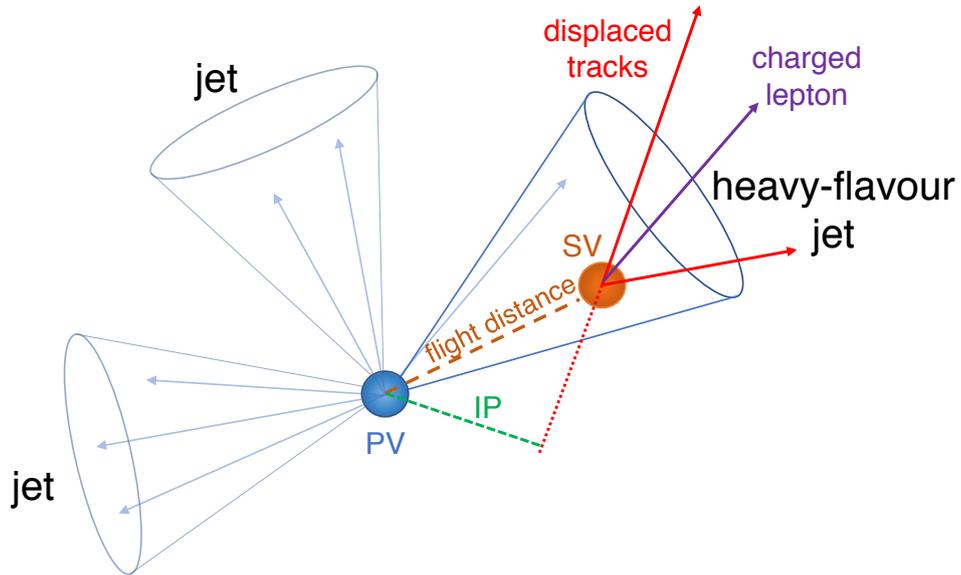


Figure 6.1. When a prompt  $b$  is created in a collision at the primary vertex (PV), it hadronises and the unstable  $B$  meson usually travels some discernible distance away before it decays, the charged decay products creating tracks that originate from the secondary vertex. We measure the impact parameter (IP), e.g. along the direction of the beam [98].

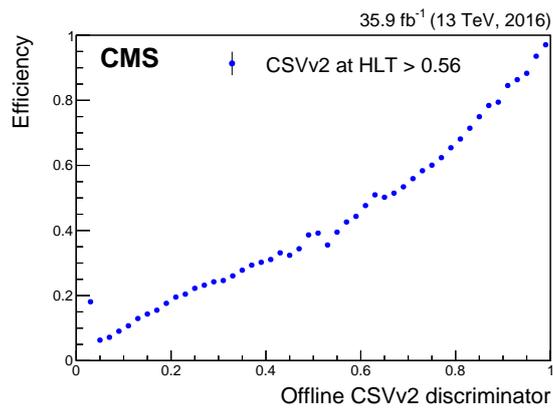
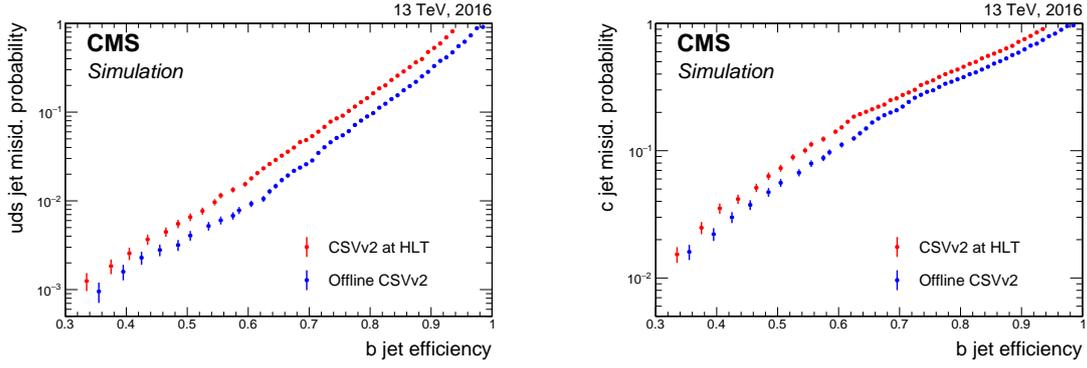


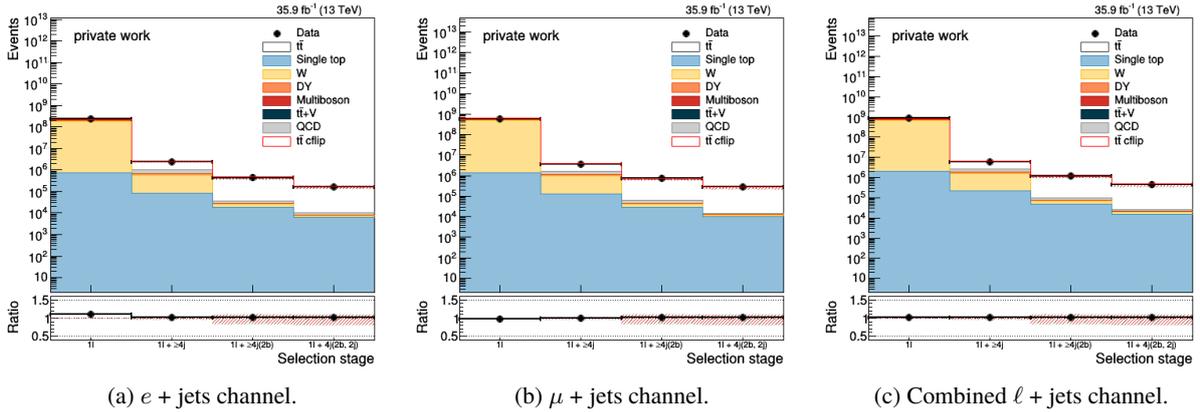
Figure 6.2.  $b$  tagging efficiency at the HLT as a function of the offline CSVv2 discriminator value [98].



(a) Comparison of the misidentification probability for light-flavour jets versus the  $b$  tagging efficiency.

(b) Comparison of the misidentification probability for  $c$  jets versus the  $b$  tagging efficiency.

Figure 6.3. Comparison of the misidentification probability for light-flavour jets (left) and  $c$  jets (right) versus the  $b$  tagging efficiency at the HLT and offline for the CSVv2 algorithm applied on simulated  $t\bar{t}$  events for which the scalar sum of the jet  $p_T$  for all jets in the event exceeds 250 GeV [98].



(a)  $e +$  jets channel.

(b)  $\mu +$  jets channel.

(c) Combined  $\ell +$  jets channel.

Figure 6.4. Event yields at different stages of selection:  $1\ell$ ,  $1\ell + \geq 4j$ ,  $1\ell + \geq 4j(2b)$ ,  $1\ell + \geq 4j(2b, 2lj)$ .

Table 6.2

Event yields.

Process	$1\ell$	$1\ell + \geq 4j$	$1\ell + \geq 4j(2b)$	$1\ell + 4j(2b, 2lj)$
<b><math>e +</math> jets channel</b>				
$t\bar{t}$	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 6.2

Continued.

Total MC	224 172 928.0	2 358 528.5	427 104.2	166 493.3
( $t\bar{t}$ uncertainty)	$\pm 459 956.0$	$\pm 89 350.9$	$\pm 64 709.8$	$\pm 27 916.7$
Data	246 644 432.0	2 411 742.0	436 962.0	169 786.0
<b><math>\mu + \text{jets channel}</math></b>				
$t\bar{t}$	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)	$\pm 489 351.3$	$\pm 145 160.4$	$\pm 107 229.4$	$\pm 46 849.0$
Data	602 190 016.0	3 728 318.0	714 552.0	277 952.0
<b>Combined <math>\ell + \text{jets channel}</math></b>				
$t\bar{t}$	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)	$\pm 673 762.4$	$\pm 233 343.6$	$\pm 171 857.7$	$\pm 74 680.0$
Data	848 834 496.0	6 140 060.0	1 151 514.0	447 738.0

Table 6.3

Event yields for the colour octet  $W$  sample.

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

## 6.2 Control plots

Figs. 6.5– 6.15 show selected control plots at various stages of event selection.

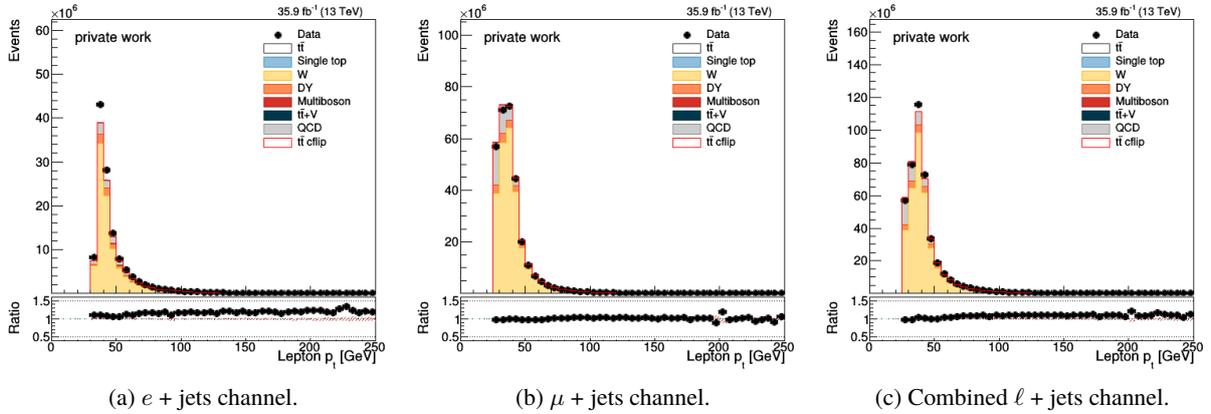


Figure 6.5. Distribution of the lepton  $p_T$  after selecting exactly one lepton.

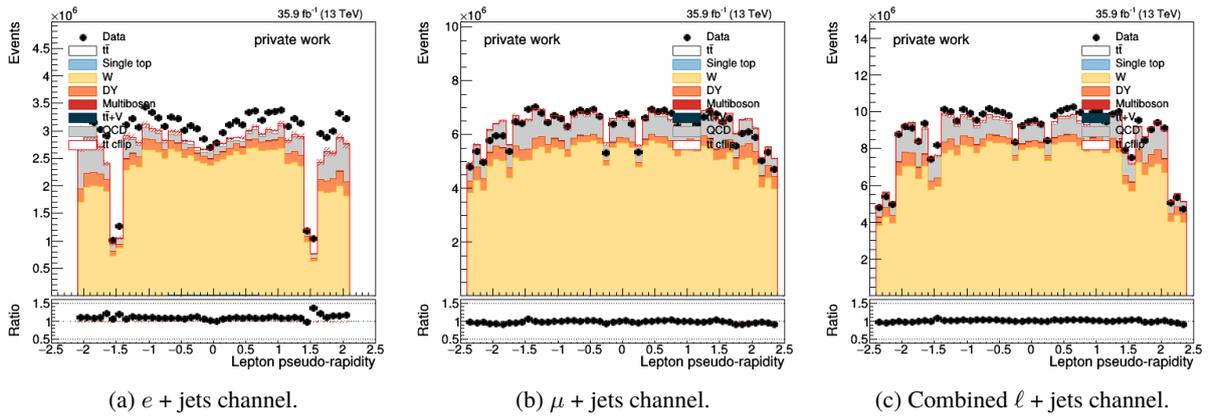


Figure 6.6. Distribution of the lepton  $\eta$  after selecting exactly one lepton.

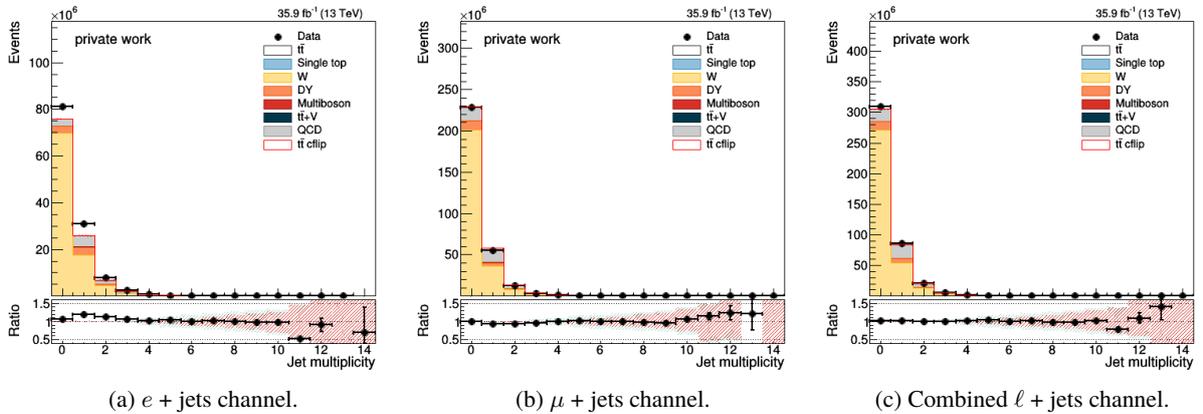


Figure 6.7. Distribution of the number of jets after selecting exactly one lepton.

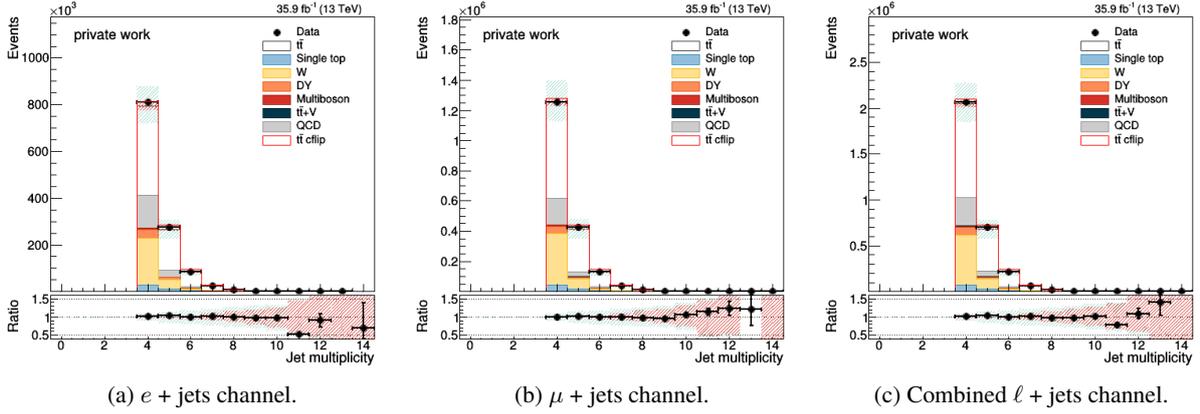


Figure 6.8. Distribution of the number of jets after selecting exactly one lepton and 4 jets.

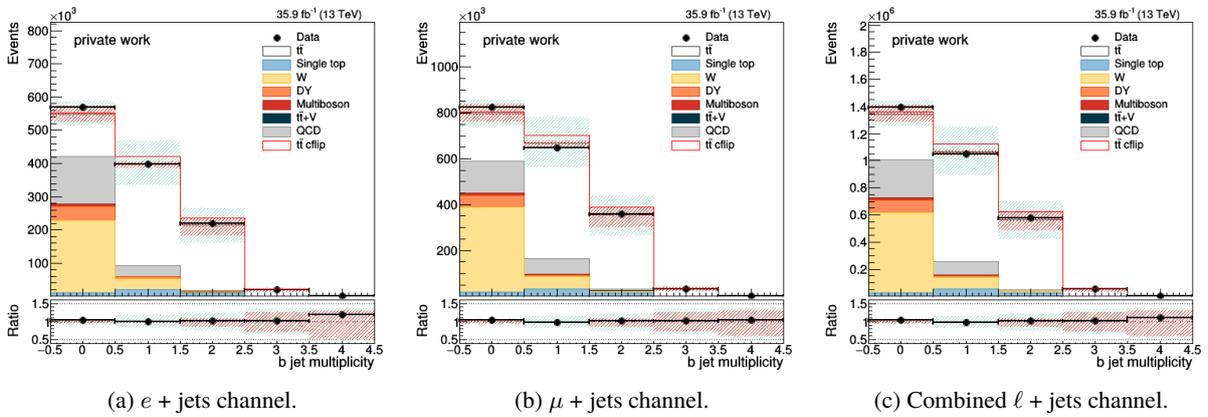


Figure 6.9. Distribution of the number of  $b$ -tagged jets after selecting exactly one lepton and 4 jets.

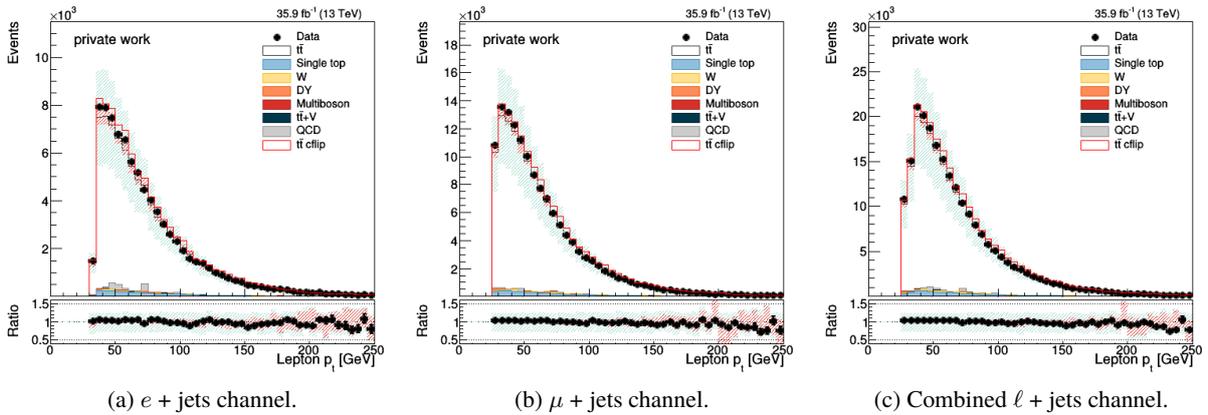


Figure 6.10. Distribution of the  $p_T$  of the lepton after selecting exactly one lepton, at least 4 jets, 2  $b$ -tagged jets and 2 light jets.

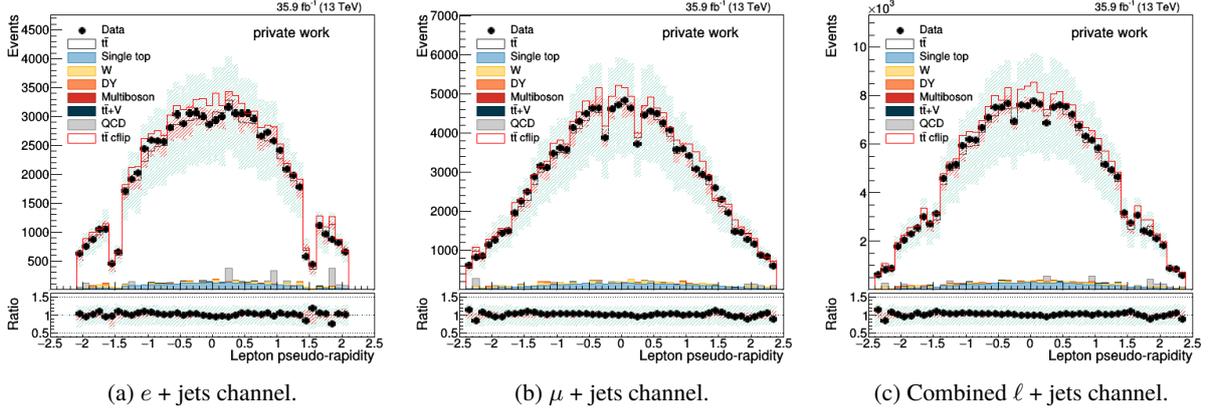


Figure 6.11. Distribution of the  $\eta$  of the lepton after selecting exactly one lepton, at least 4 jets, 2  $b$ -tagged jets and 2 light jets.

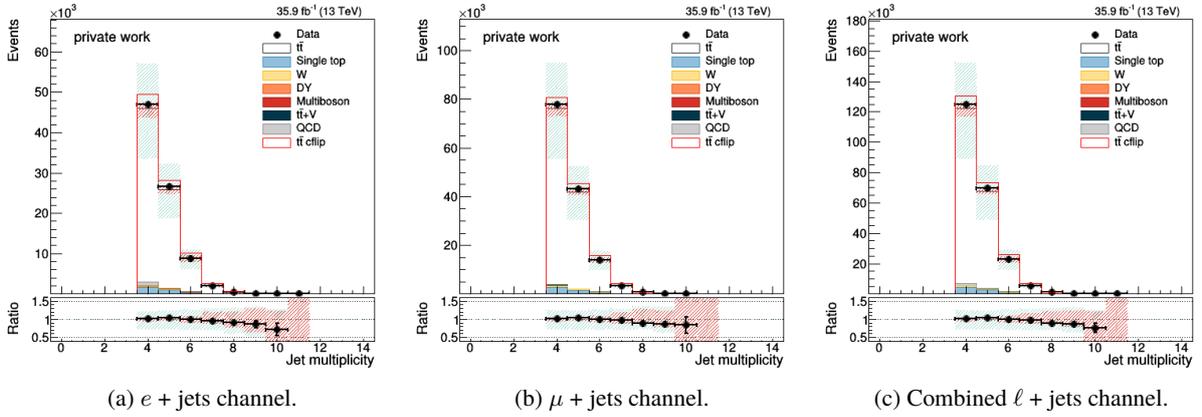


Figure 6.12. Distribution of the number of jets after selecting exactly one lepton, at least 4 jets, 2  $b$ -tagged jets and 2 light jets.

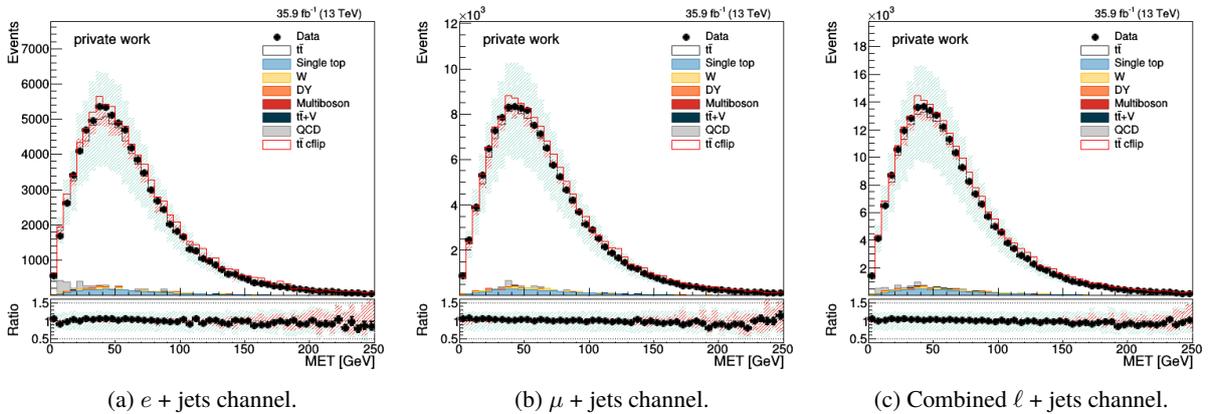


Figure 6.13. Distribution of the missing transverse momentum after selecting exactly one lepton, at least 4 jets, 2  $b$ -tagged jets and 2 light jets.

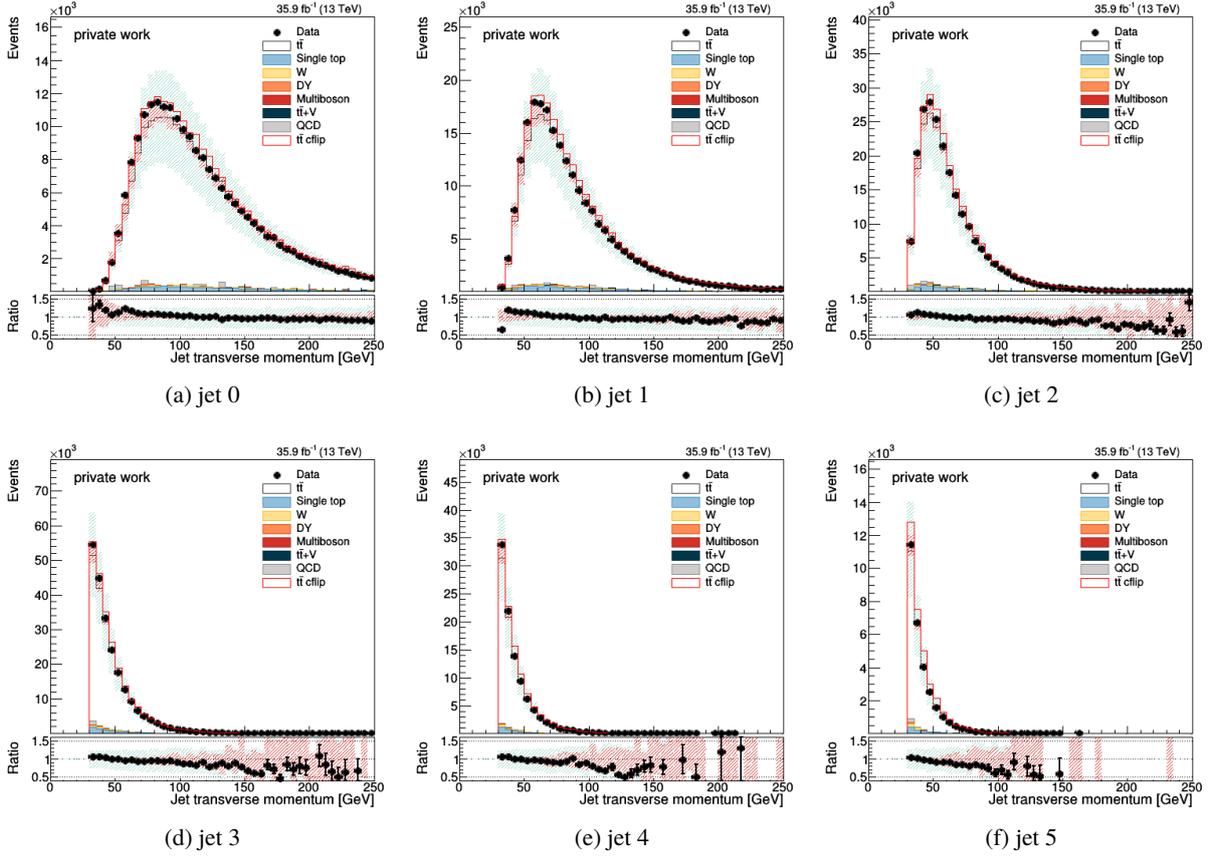


Figure 6.14. Distribution of the jet  $p_T$  in the combined  $\ell + \text{jets}$  channel after selecting exactly one lepton, at least 4 jets, 2  $b$ -tagged jets and 2 light jets. Jets are ordered according to their  $p_T$ , with jet 0 having the highest  $p_T$  and jet 5 having the lowest  $p_T$ .

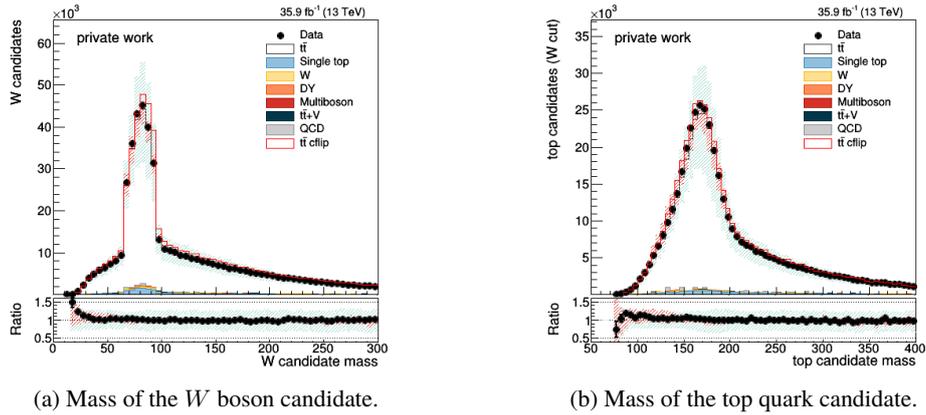


Figure 6.15. Distribution of the mass of the  $W$  boson candidate and  $t$  quark candidate (after cuts in the the  $W$  mass) after selecting exactly one lepton, at least 4 jets 2  $b$ -tagged jets and 2 light jets.

### 6.3 Generator level

We are interested in simulated events before the reconstruction by the detector as it removes detector “blur” from some details (Fig. 2.8) and are of immense importance for the unfolding procedure (Sec. 8.5). In the simulation, the offline selection is mimicked at particle level using the PSEUDOTOPPRODUCER tool [102], using a common lepton selection for both electrons and muons of  $p_T > 26$  GeV and  $|\eta| < 2.4$ , and otherwise jet  $p_T/\eta$  ( $p_T > 30$  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  $10^{-20}$  so they don’t change significantly the jet energy scale at particle level.

## 7. Systematic Uncertainties

Uncertainties are divided into statistical and systematic uncertainties. We minimise the fraction of statistical uncertainty by collecting and simulating a sufficient number of events ( $837 \times 10^6$  simulated events) so that the systematic uncertainty constitute 99.9 % of the total uncertainty. The systematic 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.

Tables summarising the importance of different systematics on the unfolded result are given in Sec. 8.5.

The discussion of this section is elaborated upon [54] and [55] as these studies use a similar set of systematics.

### 7.1 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.

The distributions of  $\theta_p$  from the leading jet  $j_1^W$  to the 2nd leading jet  $j_2^W$  under upside and downside pile-up systematics together with the nominal distribution are plotted in Fig. 7.3.

**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 determination of the scale factors has been made elsewhere, as described previously (see Chap. 5). 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.

The distributions of  $\theta_p$  from the leading jet  $j_1^W$  to the 2nd leading jet  $j_2^W$  under upside and downside trigger and selection efficiency systematics together with the nominal distribution are plotted in Fig. 7.4.

**Jet energy resolution.** We use the recommended jet energy resolution measurements [96]. Each jet is further smeared up or down depending on its  $p_T$  and  $\eta$  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.

The distributions of  $\theta_p$  from the leading jet  $j_1^W$  to the 2nd leading jet  $j_2^W$  under upside and downside jet energy resolution  $t\bar{t}jer\ up$  and  $t\bar{t}jer\ down$  systematics together with the nominal distribution are plotted in Fig. 7.5.

**Jet energy corrections.** A  $p_T, \eta$ -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 Object Group [94] 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. The following simulations are used:

**CorrelationGroup.** These are uncertainties matching the common ATLAS/CMS correlation categories grouped together [103].

**CorrelationGroupMPFInSitu.** Groups partially correlated systematic uncertainties from  $Z + jet/\gamma + jet$  absolute scale determination (e.g. radiation suppression and out-of-cone effects).

**CorrelationGroupUncorrelated.** Remaining sources which are estimated as being uncorrelated between ATLAS and CMS.

**RelativeFSR.**  $\eta$ -dependent uncertainty due to correction for initial and final state radiation, estimated from difference between MPF log-linear L2Res from PYTHIA 8 and HERWIG++, after each has been corrected for their own ISR+FSR correction [95].

**Flavour.** The flavour uncertainties are based on PYTHIA 6 Z2/HERWIG++ 2.3 differences in  $uds/c/b$  quark and gluon responses [95]. Uncertainties for the following jet flavours are used:

1. FlavorPureGluon,
2. FlavorPureQuark,
3. FlavorPureCharm,
4. FlavorPureBottom.

The distributions of  $\theta_p$  from the leading jet  $j_1^W$  to the 2nd leading jet  $j_2^W$  under the set of upside and downside jet energy correction systematics together with the nominal distribution are plotted in Fig. 7.6.

**$b$  tagging.** The nominal efficiency expected in the simulation is corrected by  $p_T$ -dependent scale factors provided by the BTV Physics Object Group [97]. 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.

The distributions of  $\theta_p$  from the leading jet  $j_1^W$  to the 2nd leading jet  $j_2^W$  under the set of upside and downside  $b$  tagging systematics  $t\bar{t}\ btag\_heavy\ down$ ,  $t\bar{t}\ btag\_heavy\ up$ ,  $t\bar{t}\ btag\_light\ down$ ,  $t\bar{t}\ btag\_light\ up$ ,  $t\bar{t}\ csv\_heavy\ down$ ,  $t\bar{t}\ csv\_heavy\ up$ ,  $t\bar{t}\ csv\_light\ down$  and  $t\bar{t}\ csv\_light\ up$  together with the nominal distribution are plotted in Fig. 7.6.

**Tracking efficiency.** As discussed in [54], the TRK and MUO Physics Object Groups have derived tracking efficiency scale factors as function of the track  $\eta$  or the reconstructed vertex multiplicity. The latter is solely available for muons and is shown in Fig. 7.1, while Table 7.1 summarises the scaling factors obtained from  $D^*$  decays. All these scale factors are run-dependent (BCDEF and GH data-taking periods are separated).

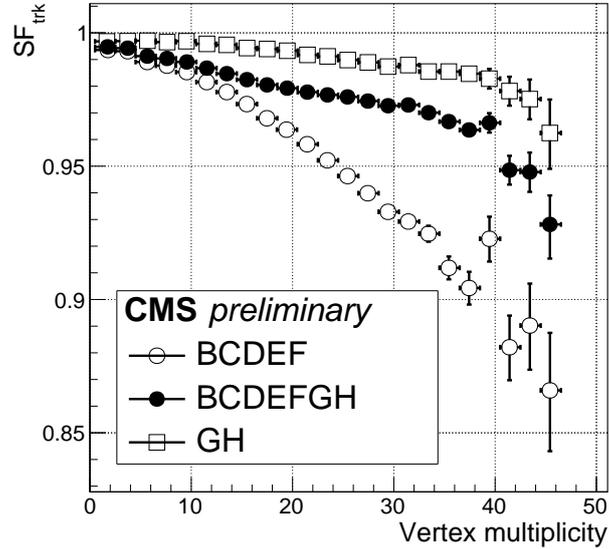


Figure 7.1. Muon tracking efficiency scale factors from the MUO POG [87].

Table 7.1

Tracking efficiency for tracks with  $p_T > 1$  GeV based on [104, 105, 106]. Courtesy of V. Mariani.

Pseudo-rapidity	$ \eta  < 0.8$	$0.8 <  \eta  < 1.5$	$ \eta  > 1.5$
BCDEF	$1.01 \pm 0.03$	$1.08 \pm 0.04$	$0.93 \pm 0.04$
GH	$1.04 \pm 0.03$	$1.07 \pm 0.06$	$1.12 \pm 0.05$

The strategy followed to assign a systematic uncertainty based on these scale factors is to either:

- remove randomly reconstructed tracks in an event if the scale factor is  $< 1$ ,
- promote an un-matched generator-level charged particle with status 1 (stable) to a track with the same spatial momentum as the generated one and assigned the mass of the charged pion if the scale factor is  $> 1$ .

To decide if a track (generator particle) should be removed (promoted) a uniform PDF in the  $[0,1]$  range is sampled randomly. If the probability exceeds the scale factor (2-scale factor), the track (generator particle) is removed (promoted). In order to reflect the possible different performance in the BCDEF and GH eras for each event a random number generator is used to assign the era (according to the relative proportion of integrated luminosity in each era) so that different scale factors are applied to evaluate the uncertainty.

In the process described above we consider the effect of applying twice the possible correction with the scale factor in order to cover the uncertainty on the scale factor itself. Thus in the evaluation above we change  $SF$  to  $SF^2$ . In each bin we then take the maximum variation induced as the estimate of the uncertainty.

Fig. 7.2 shows the expected effect on the charged multiplicity which is the distribution that is more severely affected by this systematic. The relative change induced in the charge multiplicity distribution is expected to be of the same order of that observed in data.

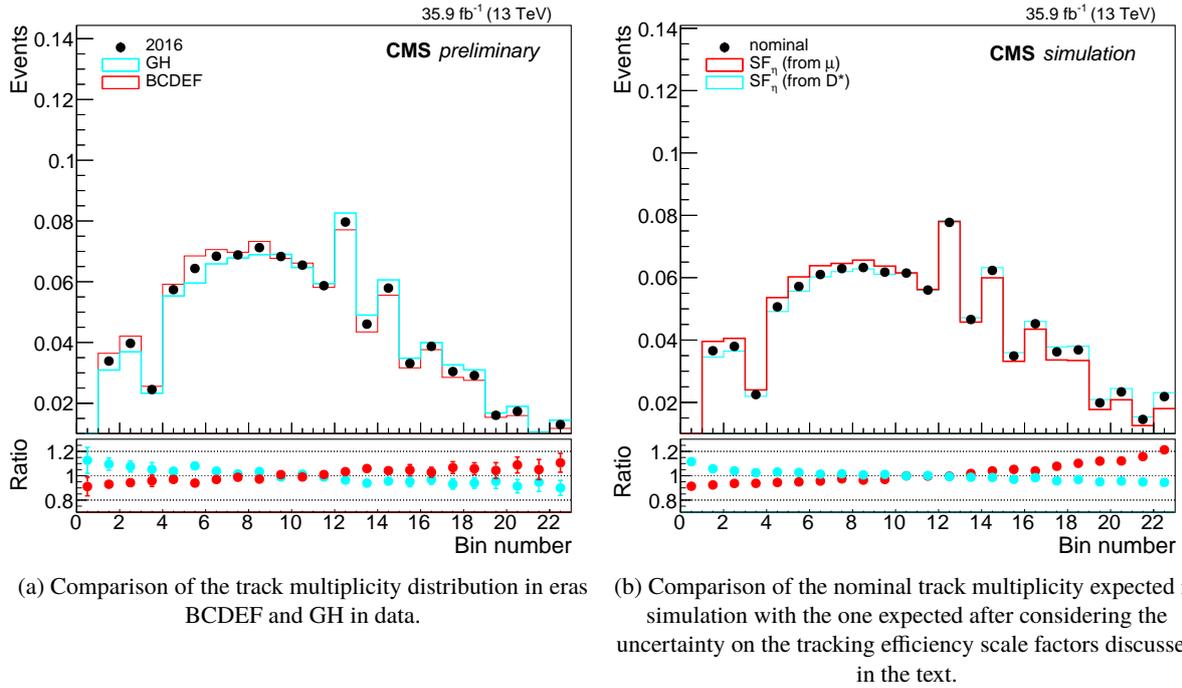


Figure 7.2. Expected effect on the charged multiplicity by tracking efficiency [54].

The distributions of  $\theta_p$  from the leading jet  $j_1^W$  to the 2nd leading jet  $j_2^W$  under the set of upside and downside tracking efficiency systematics  $t\bar{t}$  tracking down, and  $t\bar{t}$  tracking up together with the nominal distribution are plotted in Fig. 7.8.

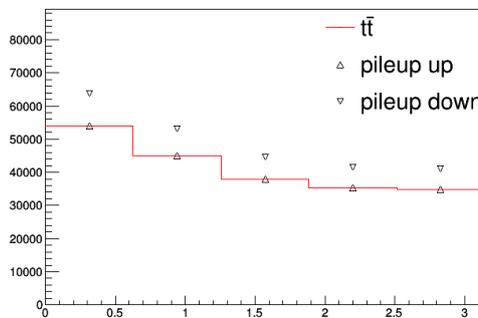
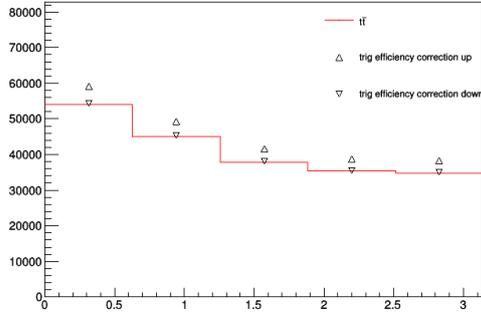
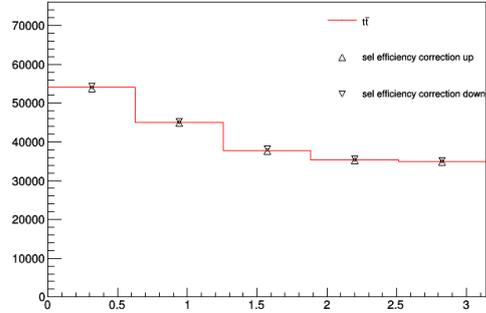


Figure 7.3. The nominal distribution of the pull angle from  $j_1^W$  to  $j_2^W$  with all jet constituents and at all  $\Delta R$  and the distributions hereof with the pile-up uncertainties.



(a) Trigger efficiency uncertainties.



(b) Selection efficiency uncertainties

Figure 7.4. The nominal distribution of the pull angle from  $j_1^W$  to  $j_2^W$  with all jet constituents and at all  $\Delta R$  and the distributions hereof with trigger and selection efficiency uncertainties.

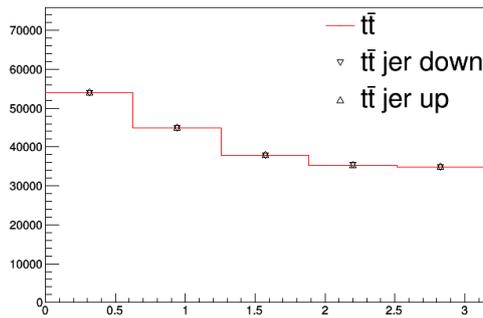
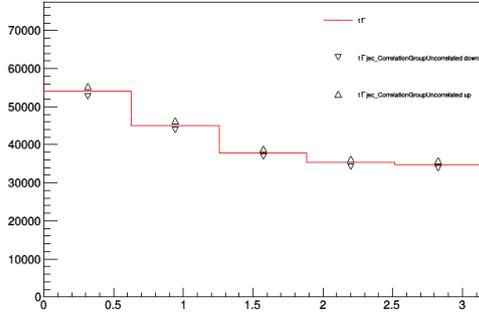
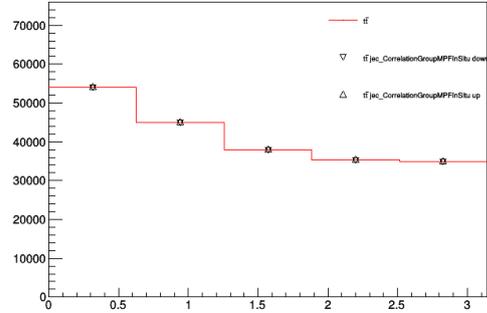


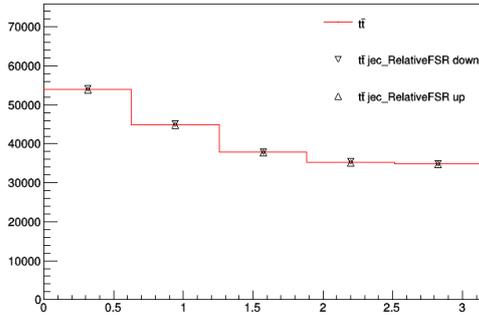
Figure 7.5. The nominal distribution of the pull angle from  $j_1^W$  to  $j_2^W$  with all jet constituents and at all  $\Delta R$  and the distributions hereof with jet energy resolution uncertainties  $t\bar{t}$  *jer down* and  $t\bar{t}$  *jer up*.



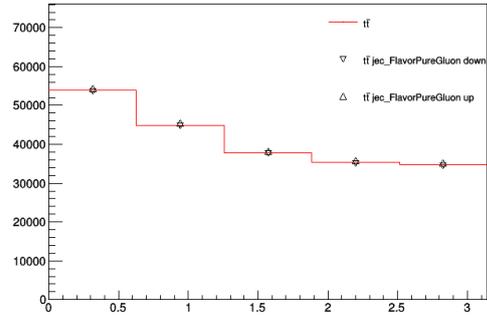
(a)  $t\bar{t}$   $jec\_CorrelationGroupUncorrelated$  down and  $t\bar{t}$   $jec\_CorrelationGroupUncorrelated$  up.



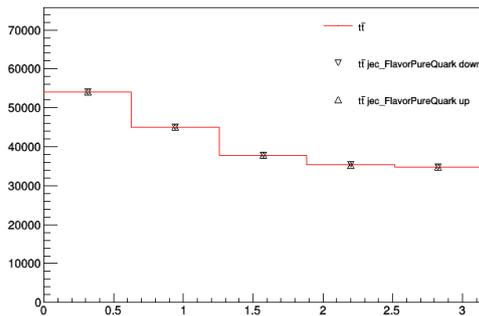
(b)  $t\bar{t}$   $jec\_CorrelationGroupMPFInSitu$  down and  $t\bar{t}$   $jec\_CorrelationGroupMPFInSitu$  up.



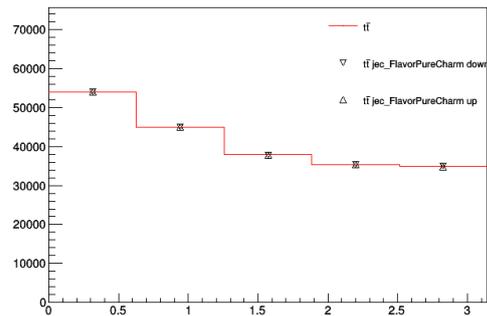
(c)  $t\bar{t}$   $jec\_RelativeFSR$  down and  $t\bar{t}$   $jec\_RelativeFSR$  up.



(d)  $t\bar{t}$   $jec\_FlavorPureGluon$  down and  $t\bar{t}$   $jec\_RelativeFSR$  up.

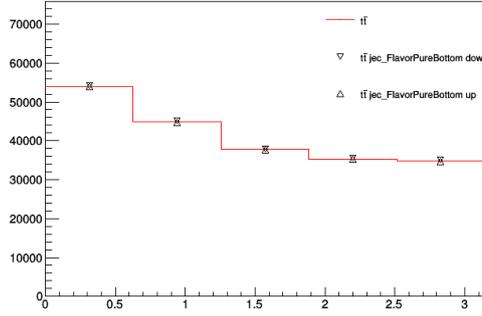


(e)  $t\bar{t}$   $jec\_FlavorPureQuark$  down and  $t\bar{t}$   $jec\_FlavorPureQuark$  up.



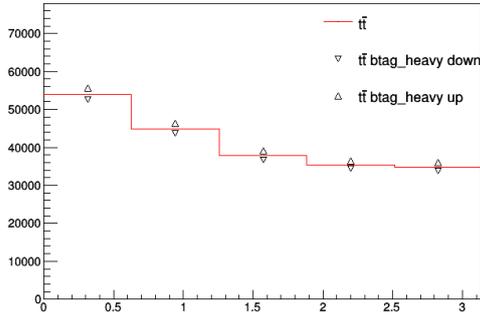
(f)  $t\bar{t}$   $jec\_FlavorPureCharm$  down and  $t\bar{t}$   $jec\_FlavorPureCharm$  up.

Figure 7.6. The nominal distribution of the pull angle from  $j_1^W$  to  $j_2^W$  with all jet constituents and at all  $\Delta R$  and the distributions hereof with jet energy correction uncertainties.

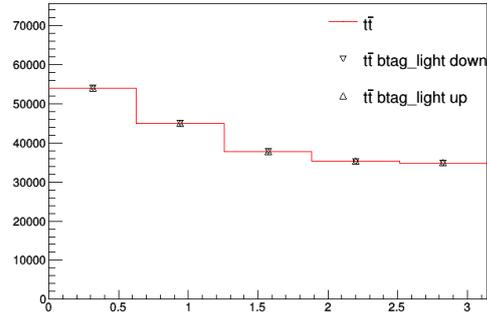


(g)  $t\bar{t}$  *jec\_FlavorPureBottom* down and  $t\bar{t}$  *jec\_FlavorPureBottom* up.

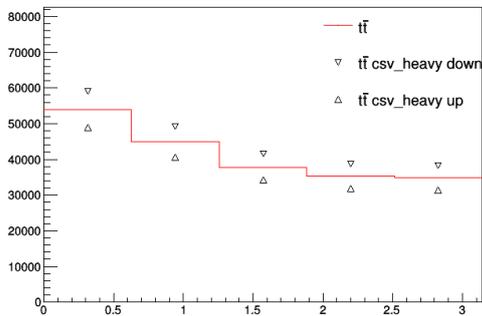
Figure 7.6. The nominal distribution of the pull angle from  $j_1^W$  to  $j_2^W$  with all jet constituents and at all  $\Delta R$  and the distributions hereof with jet energy correction uncertainties (continued).



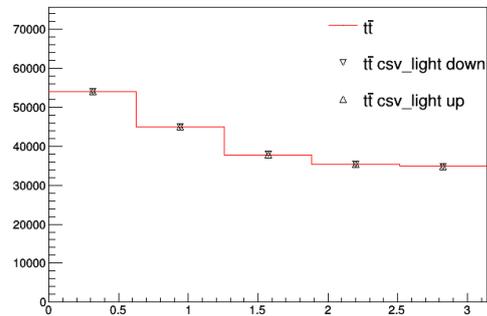
(a)  $t\bar{t}$  *btag\_heavy* down and  $t\bar{t}$  *btag\_heavy* up.



(b)  $t\bar{t}$  *btag\_light* down and  $t\bar{t}$  *btag\_light* up.



(c)  $t\bar{t}$  *csv\_heavy* down and  $t\bar{t}$  *csv\_heavy* up.



(d)  $t\bar{t}$  *csv\_light* down and  $t\bar{t}$  *csv\_light* up.

Figure 7.7. The nominal distribution of the pull angle from  $j_1^W$  to  $j_2^W$  with all jet constituents and at all  $\Delta R$  and the distributions hereof with  $b$  tagging uncertainties.

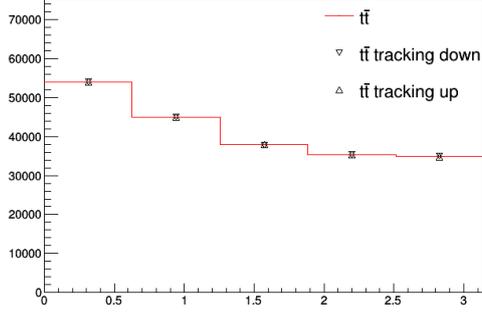


Figure 7.8. The nominal distribution of the pull angle from  $j_1^W$  to  $j_2^W$  with all jet constituents and at all  $\Delta R$  and the distributions hereof with tracking efficiency uncertainties  $t\bar{t}$  tracking down and  $t\bar{t}$  tracking up.

## 7.2 Theoretical uncertainties

**QCD scale choices.** We consider anti-correlated variations of the factorisation and renormalisation scales ( $\mu_R/\mu_F$ ) in the  $t\bar{t}$  and  $W$  + jets samples, by factors of 0.5 and 2. These variations are saved in the simulated events as an alternative set of weights which are used in the evaluation of this systematic. The envelope of 7 variations (excluding opposite variations of  $\mu_R/\mu_F$ ) is considered as a systematic.

The distributions of  $\theta_p$  from the leading jet  $j_1^W$  to the 2nd leading jet  $j_2^W$  under the set of QCD scale systematics together with the nominal distribution are plotted in Fig. 7.9.

**EVTGEN .** EVTGEN is a Monte Carlo event generator that simulates the decays of heavy flavour particles, primarily of  $B$  and  $D$  mesons. It uses amplitudes instead of probabilities. References are available in [107], [108].

The distributions of  $\theta_p$  from the leading jet  $j_1^W$  to the 2nd leading jet  $j_2^W$  under the EVTGEN systematic  $t\bar{t}$  *evtgen* together with the nominal distribution are plotted in Fig. 7.10.

**Hadroniser choice.** We check the effect of using HERWIG++ [109], tune EE5C[110], instead of PYTHIA 8 CUET2P8M4. The key difference arises from the hadronisation model – PYTHIA uses the string model, while HERWIG++ uses the cluster model [111].

The distributions of  $\theta_p$  from the leading jet  $j_1^W$  to the 2nd leading jet  $j_2^W$  under the systematic of choosing HERWIG++ as the hadroniser  $t\bar{t}$  *Herwig++* together with the nominal distribution are plotted in Fig. 7.11.

Privately produced samples based on SHERPA [112], DIRE [113], HERWIG 7 [114] could be added in a later stage.

**Top quark mass.** The most precise measurement of the top quark mass by CMS yields a total uncertainty of  $\pm 0.49$  GeV [115]. We consider however a conservative uncertainty of  $\pm 1$  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.

The distributions of  $\theta_p$  from the leading jet  $j_1^W$  to the 2nd leading jet  $j_2^W$  under the systematic of uncertainty of the top quark mass  $t\bar{t} m = 171.5$  and  $t\bar{t} m = 173.5$  together with the nominal distribution are plotted in Fig. 7.12.

**PYTHIA tunes.** The following PYTHIA tunes are used:

**Matrix Element + Parton Shower matching scheme.** The default simulation is based on POWHEG. The so-called hdamp parameter is varied according to the range determined in [116].

**Parton shower scale.** Alternative POWHEG + PYTHIA 8 samples where the parton shower scale choice is varied by a factor of 0.5 and 2 for ISR and FSR separately, are used in the analysis. This affects the fragmentation and hadronisation of the jets initiated by the matrix element calculation as well as the emission of extra jets by the hadroniser.

**Colour reconnection model.** We vary the colour reconnection model with respect to the default using alternatives including the resonant decay products in possible reconnections to the UE. The default simulation (MPI-based colour reconnection) has this effect excluded. We examine three alternative models for CR: the so-called gluon move [117], the QCD-based models[118], and ERDOn.

**Underlying Event (UE) variations.** The default parameters in the CUETP8M2T4 tune are varied according to their uncertainty and the effect on the unfolding is taken as an estimate of the systematic uncertainty.

The setups of the PYTHIA tunes described herein are summarised in Table 7.2

Table 7.3 summarises the simulation samples from the RunIISummer16MiniAODv2-PUMoriond17\_80X\_mcRun2\_asymptotic\_2016\_TracheIV\_v6 production used for the theoretical systematics.

The distributions of  $\theta_p$  from the leading jet  $j_1^W$  to the 2nd leading jet  $j_2^W$  under each upside or downside PYTHIA tune together with the nominal distribution are plotted in Fig. 7.13–7.16.

Table 7.2

Variations of the POWHEG + PYTHIA 8 setup used for the comparison with the measurements. The values changed with respect to the CUETP8M2T4 tune are given in the columns corresponding to each model. After [119].

Parameter	POWHEG + PYTHIA 8 simulation setups						
	CUETP8M2T4	MPI/CR UE up/down	Parton shower scale		CR including $t\bar{t}$		
			ISR up/down	FSR up/down	ERD on	QCD based	Gluon move
PartonLevel MPI	on						
SpaceShower renormMultFac alphaSvalue	1.0 0.1108		4/0.25			0.2521	
TimeShower renormMultFac alphaSvalue	1.0 0.1365			4/0.25		0.2521	
MultipartonInteractions pT0Ref ecmPow expPow	2.2 0.2521 1.6	2.20/2.128 1.711/1.562				2.174 0.2521 1.312	2.3 1.35
ColorReconnection reconnect range mode junctionCorrection timeDilationPar m0 flipMode m2Lambda fracGluon dLambdaCut	on 6.59 0	6.5/8.7				1 0.1222 15.86 1.204	2 0 1.89 1 0
PartonVertex setVertex							
PartonLevel earlyResDec	off				on	on	on

Table 7.3

Simulation samples used for systematics [55].

Signal variation	Dataset	$\sigma$ [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-fsdown-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

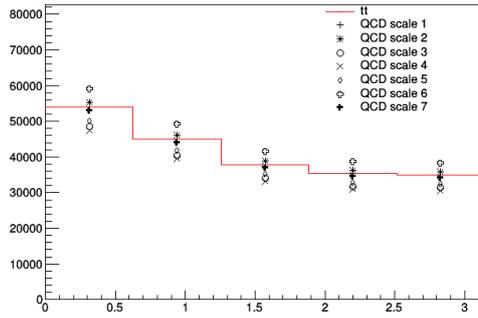


Figure 7.9. The nominal distribution of the pull angle from  $j_1^W$  to  $j_2^W$  with all jet constituents and at all  $\Delta R$  and the distributions hereof with the QCD scale uncertainties.

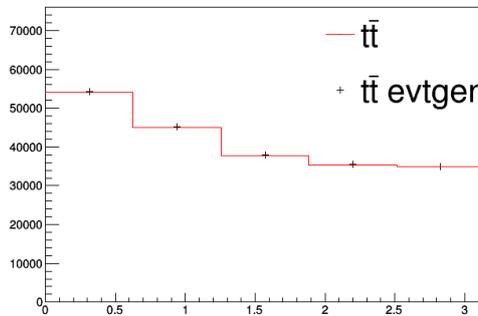


Figure 7.10. The nominal distribution of the pull angle from  $j_1^W$  to  $j_2^W$  with all jet constituents and at all  $\Delta R$  and the distributions hereof with uncertainty arising from assuming the decay model of heavy flavour particles used in EVTGEN.

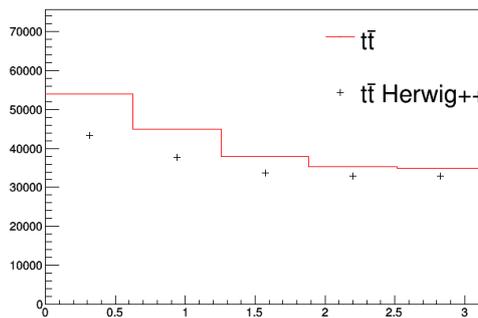


Figure 7.11. The nominal distribution of the pull angle from  $j_1^W$  to  $j_2^W$  with all jet constituents and at all  $\Delta R$  and the distribution hereof with uncertainty arising from assuming the hadronisation model used in HERWIG++.

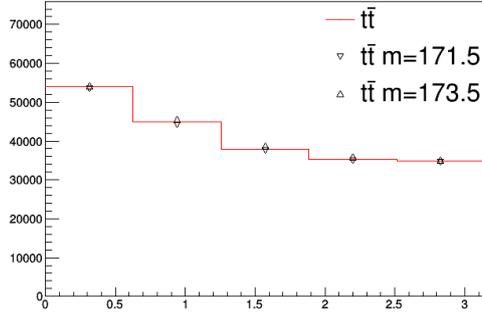


Figure 7.12. The nominal distribution of the pull angle from  $j_1^W$  to  $j_2^W$  with all jet constituents and at all  $\Delta R$  and the distributions hereof with uncertainties arising from  $\pm 1$  GeV variations in the mass of the  $t$  quark.

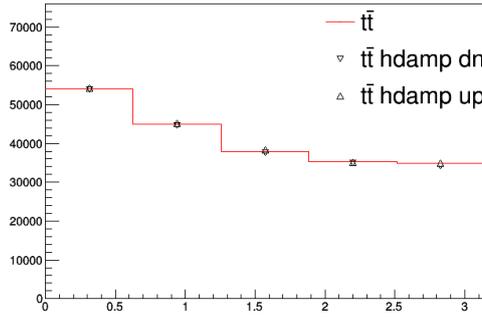


Figure 7.13. The nominal distribution of the pull angle from  $j_1^W$  to  $j_2^W$  with all jet constituents and at all  $\Delta R$  and the distributions hereof with uncertainty from varying the hdamp parameter in the Parton Shower + Matrix Element matching scheme.

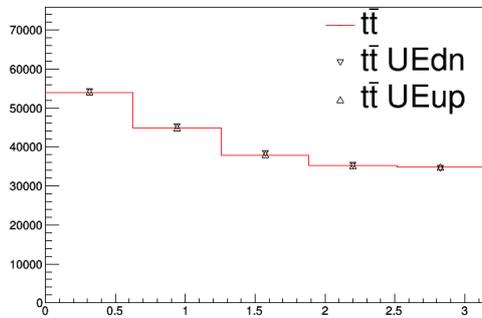
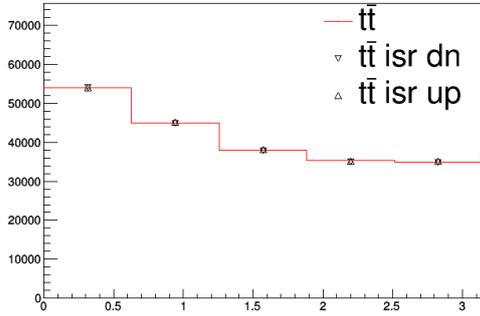
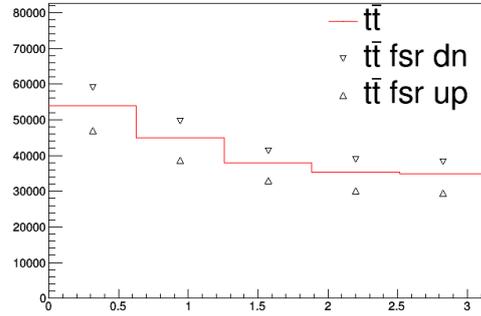


Figure 7.14. The nominal distribution of the pull angle from  $j_1^W$  to  $j_2^W$  with all jet constituents and at all  $\Delta R$  and the distributions hereof with uncertainties in the Underlying Event.

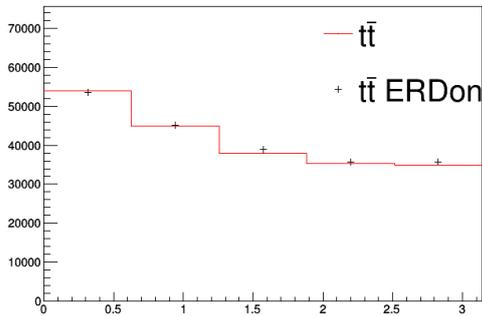


(a)  $t\bar{t}$  *ISRdown* and  $t\bar{t}$  *ISRup*.

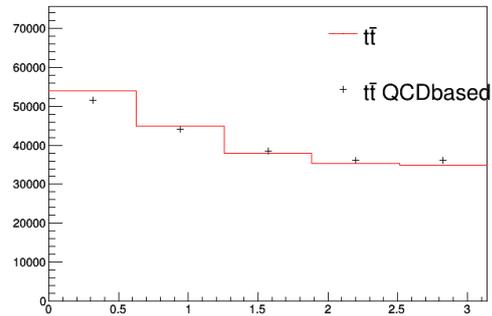


(b)  $t\bar{t}$  *FSRdn* and  $t\bar{t}$  *FSRup*.

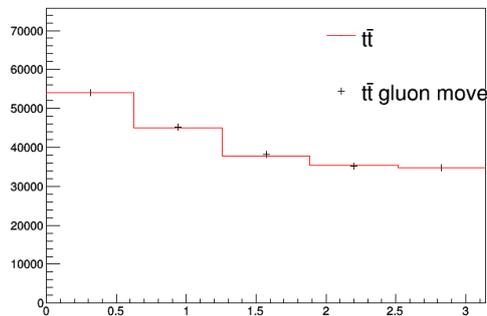
Figure 7.15. The nominal distribution of the pull angle from  $j_1^W$  to  $j_2^W$  with all jet constituents and at all  $\Delta R$  and the distributions hereof with parton shower scale uncertainties.



(a)  $t\bar{t}$  *ERDon*.



(b)  $t\bar{t}$  *QCDbased*.



(c)  $t\bar{t}$  *gluonmove*.

Figure 7.16. The nominal distribution of the pull angle from  $j_1^W$  to  $j_2^W$  with all jet constituents and at all  $\Delta R$  and the distributions hereof with uncertainties from colour reconnection.

## 8. Results

### 8.1 Pull vector

A selected event observed in the CMS detector is displayed in Fig. 8.1, showing the light jets, the  $b$ -tagged jets, the charged lepton and the pull vector in the  $\phi$ - $y$  plane in a manner analogous to Fig. 4.2.

A set of base tools CFAT [120] 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. Results with RIVET are shown in Fig. 8.2. Fig. 8.2(a) shows the distribution of  $\theta_p$  from  $j_1^W$  to  $j_2^W$ . The central peak which is the experimental signature of colour connect jets is present in the SM results but disappears in the  $W$  colour octet results. On the other hand the distribution of  $\theta_p$  suffers no alterations from  $j_1^W$  to lepton as shown in Fig. 8.2(b).

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 [121]. 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  $e + \text{jets}$ ,  $\mu + \text{jets}$  and combined  $\ell + \text{jets}$  channels.

The distribution of the  $\eta$  component of the pull vector with all jet components is given in Figs. 8.3–8.6.

An explanation of how CMS plots are represented is in order. The top plot in Fig. 8.3 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:

$$U_i = \sqrt{\sum_k (U_i^k - N_i)^2}, \quad D_i = \sqrt{\sum_k (D_i^k - N_i)^2}. \quad (8.1)$$

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).

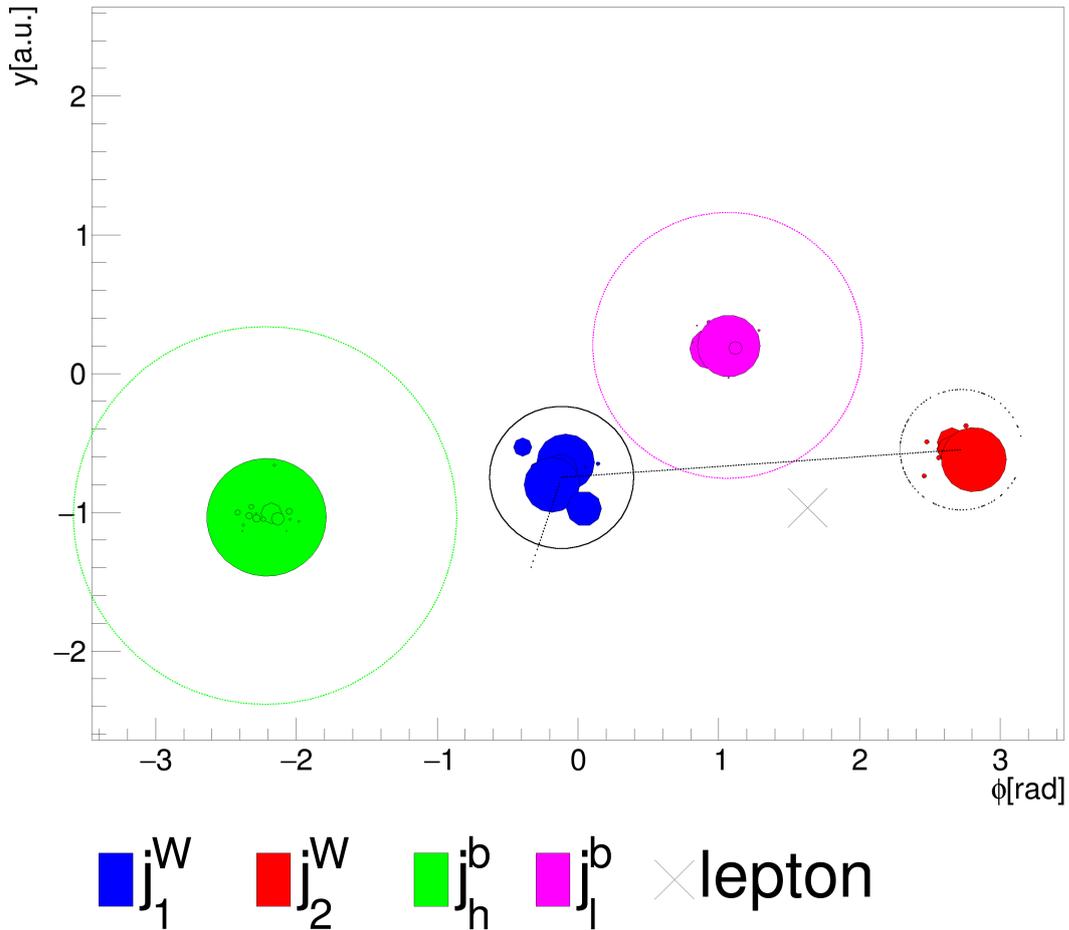


Figure 8.1. An event observed with the CMS detector. Pull vector (dash-dotted) of the leading light jet forming a pull angle of 1.96 rad with the difference between the second leading light jet and the leading light jet (dashed). Constituents of the leading light jet are marked in blue while the constituents 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  $p_T [\text{GeV}]/75.0$  and the radius of the circles representing constituents is equal to  $p_T^{\text{constituent}}/p_T^{\text{jet}}$ . 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 “x” is the charged lepton.

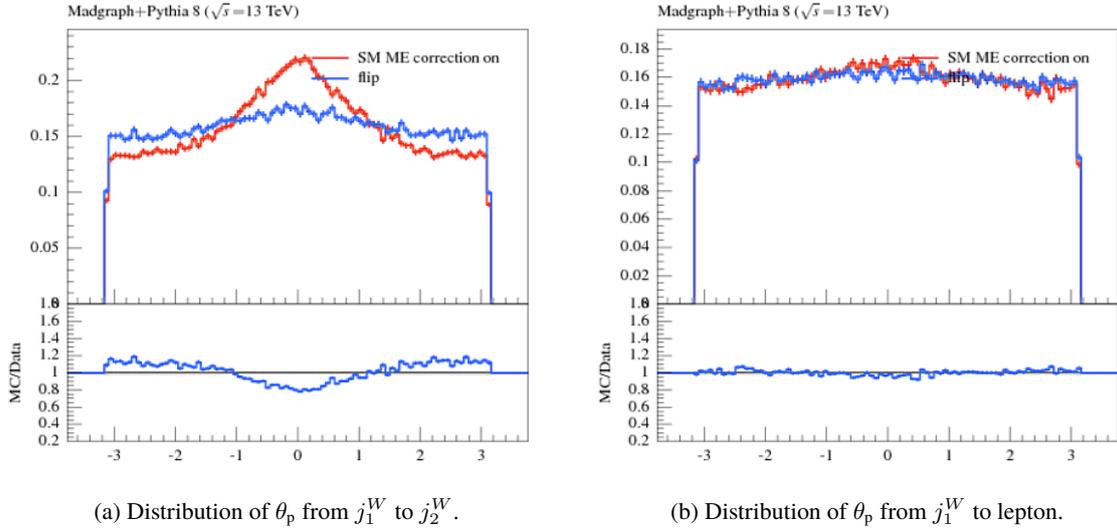


Figure 8.2. Results with RIVET showing SM (red) and  $W$  colour octet (blue) distributions of the pull angle. The bottom inset shows the bin-per-bin ratio of the  $W$  colour octet results to the SM results.

The bottom inset shows the ratio of data to Monte Carlo, as well as systematics and systematics from shapes normalised to Monte Carlo.

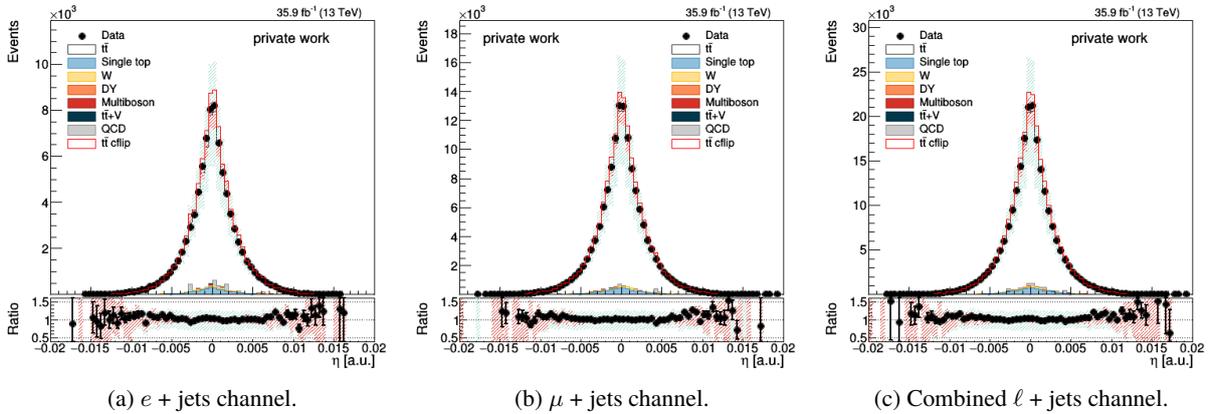


Figure 8.3. Distribution of the  $\eta$  component of the pull vector of  $j_1^W$  with all jet components.

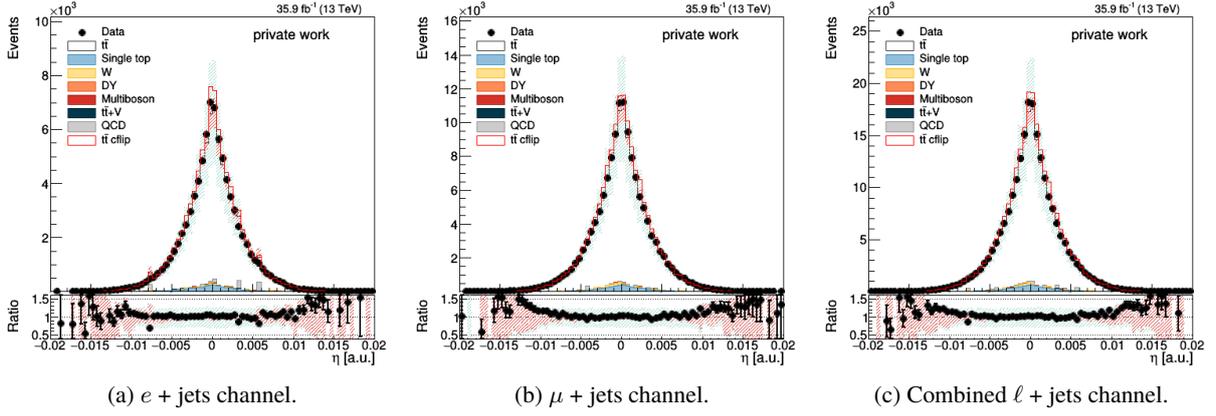


Figure 8.4. Distribution of the  $\eta$  component of the pull vector of  $j_2^W$  with all jet components.

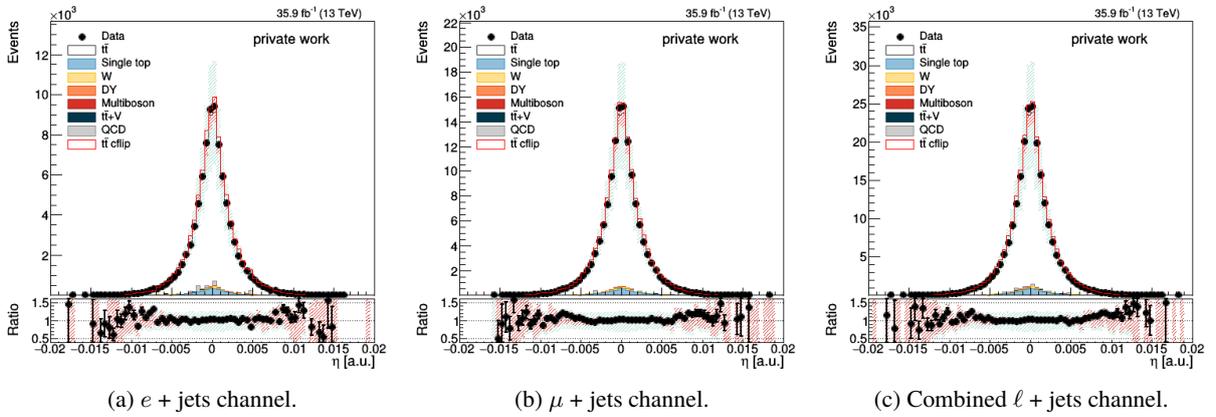


Figure 8.5. Distribution of the  $\eta$  component of the pull vector of  $j_1^b$  with all jet components.

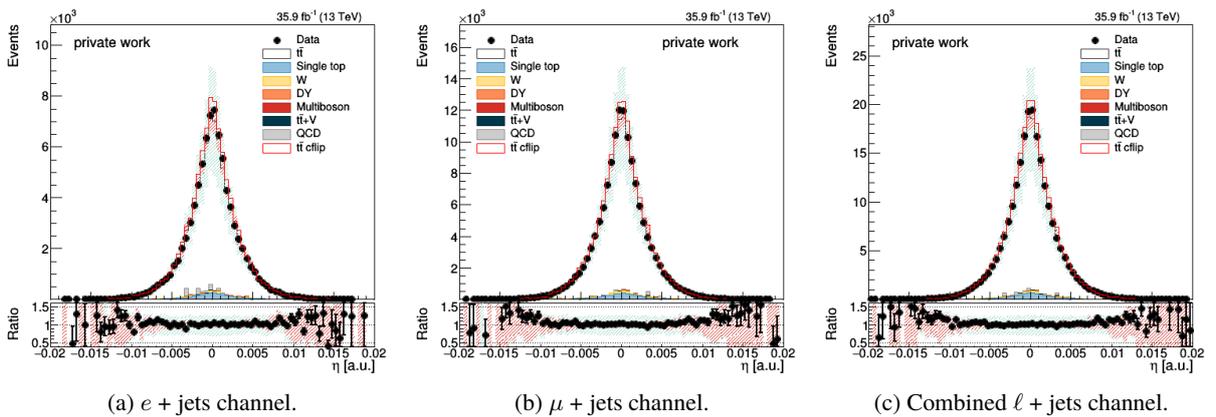


Figure 8.6. Distribution of the  $\eta$  component of the pull vector of  $j_2^b$  with all jet components.

The distribution of the  $\phi$  component of the pull vector with all jet components is given in Figs. 8.7–8.10.

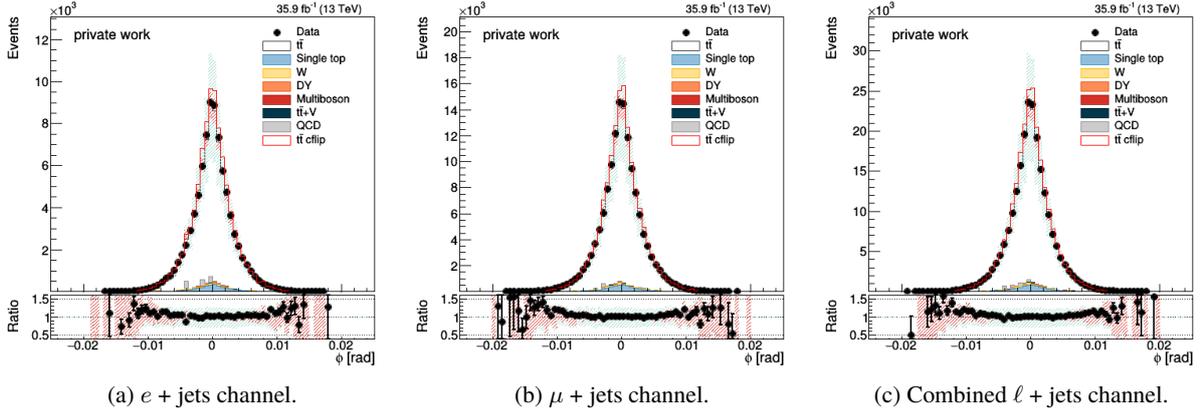


Figure 8.7. Distribution of the  $\phi$  component of the pull vector of  $j_1^W$  with all jet components.

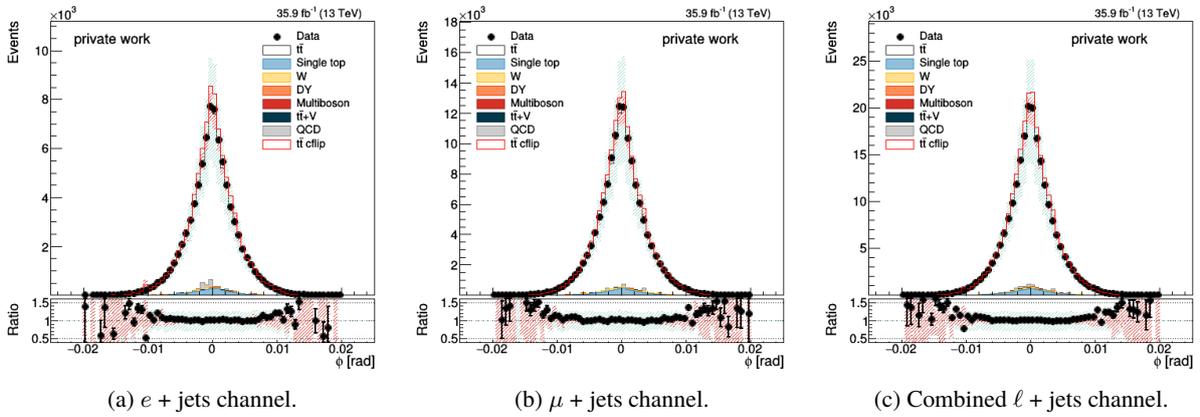


Figure 8.8. Distribution of the  $\phi$  component of the pull vector of  $j_2^W$  with all jet components.

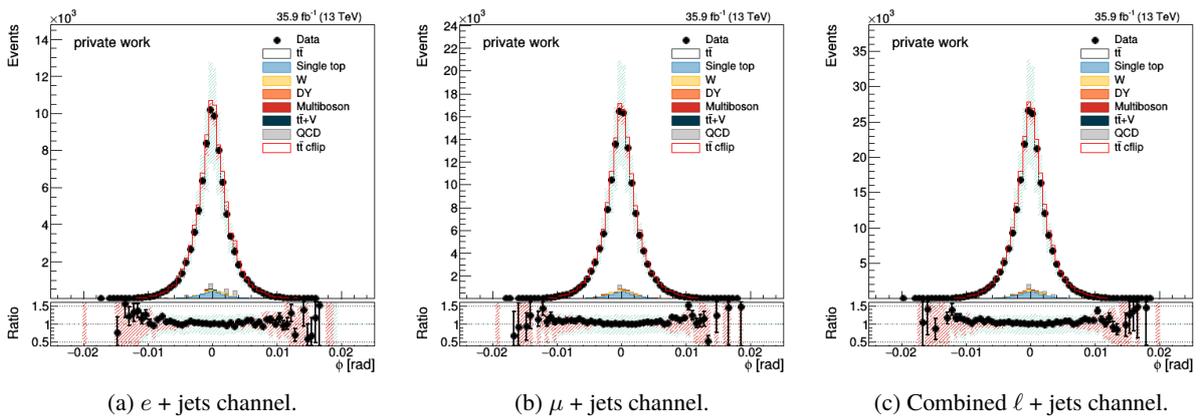


Figure 8.9. Distribution of the  $\phi$  component of the pull vector of  $j_1^b$  with all jet components.

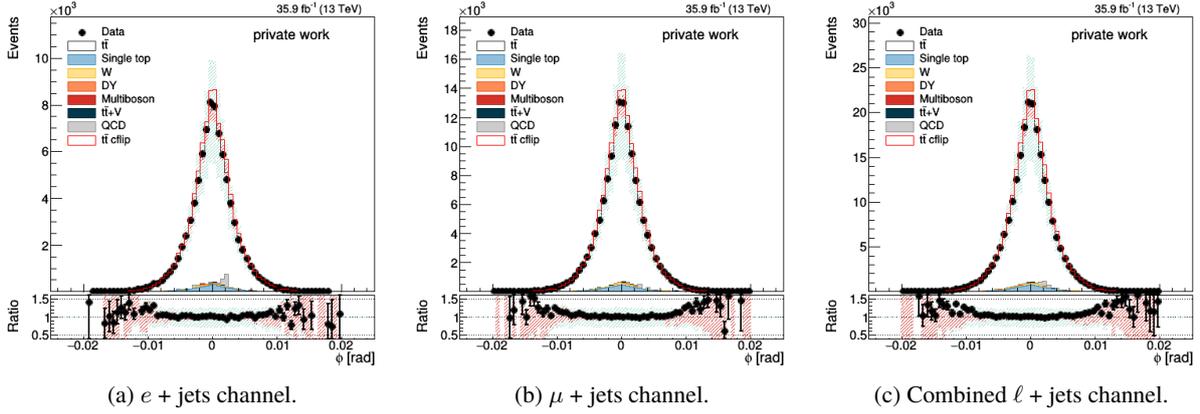


Figure 8.10. Distribution of the  $\phi$  component of the pull vector of  $j_2^b$  with all jet components.

The magnitude of the pull vector with all jet components is given in Figs. 8.11–8.14. The magnitude of the pull vector is usually contained below 0.02 [a.u.].

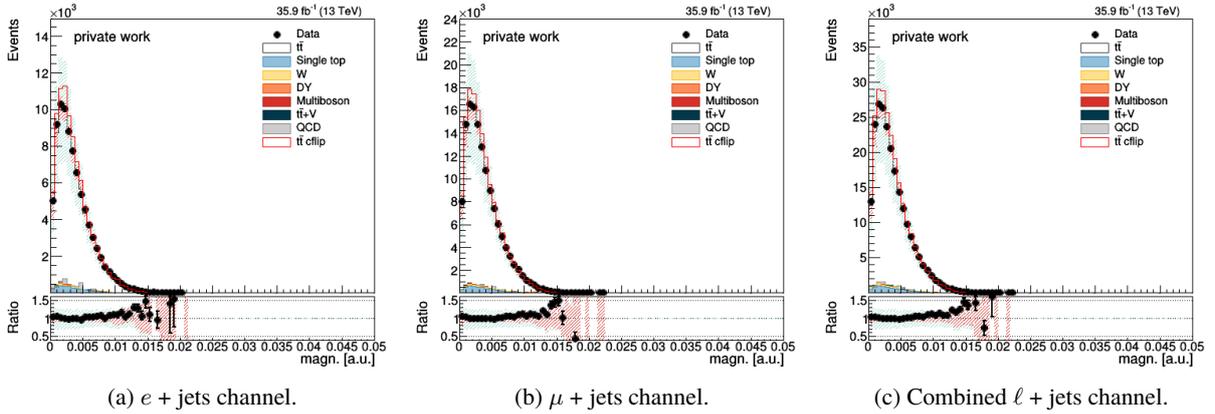


Figure 8.11. The magnitude of the pull vector of  $j_1^W$  with all jet components.

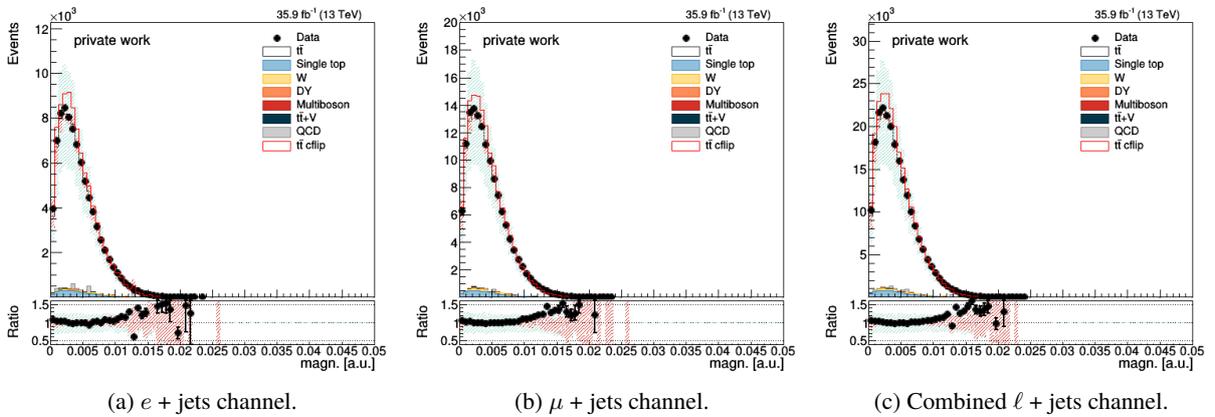


Figure 8.12. The magnitude of the pull vector of  $j_2^W$  with all jet components.

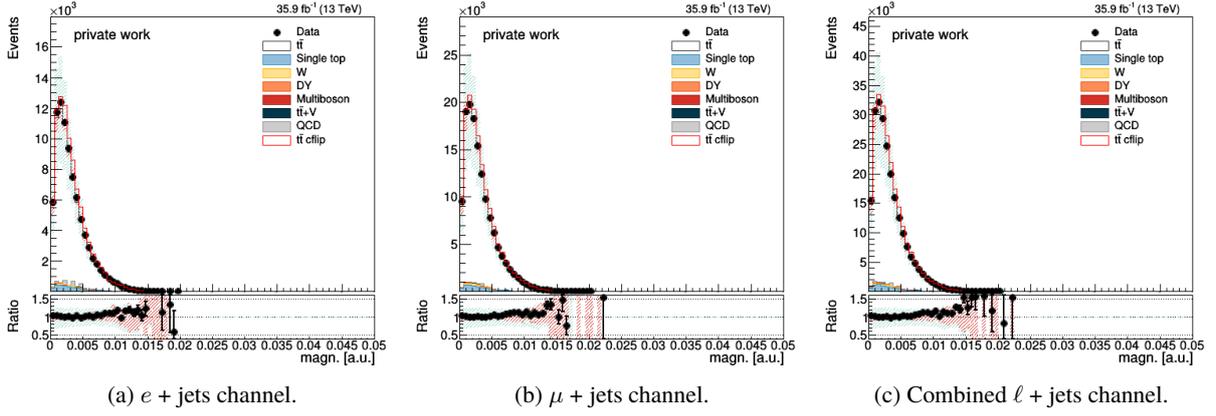


Figure 8.13. The magnitude the pull vector of  $j_1^b$  with all jet components.

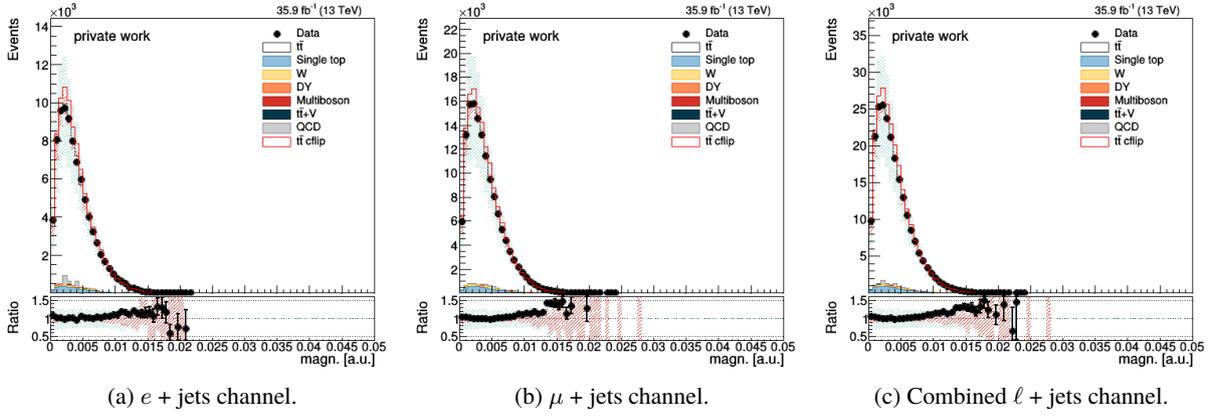


Figure 8.14. The magnitude of the pull vector of  $j_2^b$  with all jet components.

## 8.2 Pull angle

The plots of the pull angle between colour-connected jets – from  $j_1^W$  to  $j_2^W$  and  $j_2^W$  to  $j_1^W$  with all jet constituents and including all values of  $\Delta R$  are shown in Fig. 8.15 and Fig. 8.16.

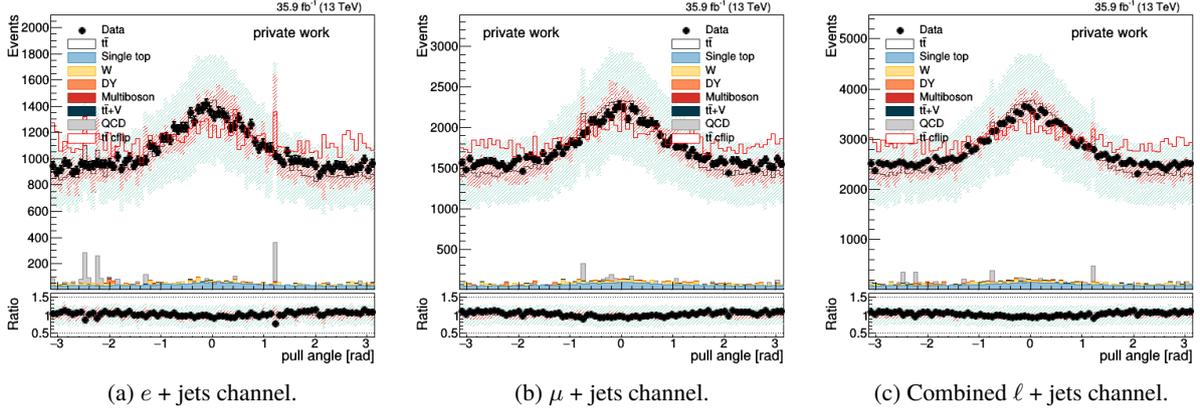


Figure 8.15. Distribution of the pull angle from  $j_1^W$  to  $j_2^W$  for all  $\Delta R$  and including all particles.

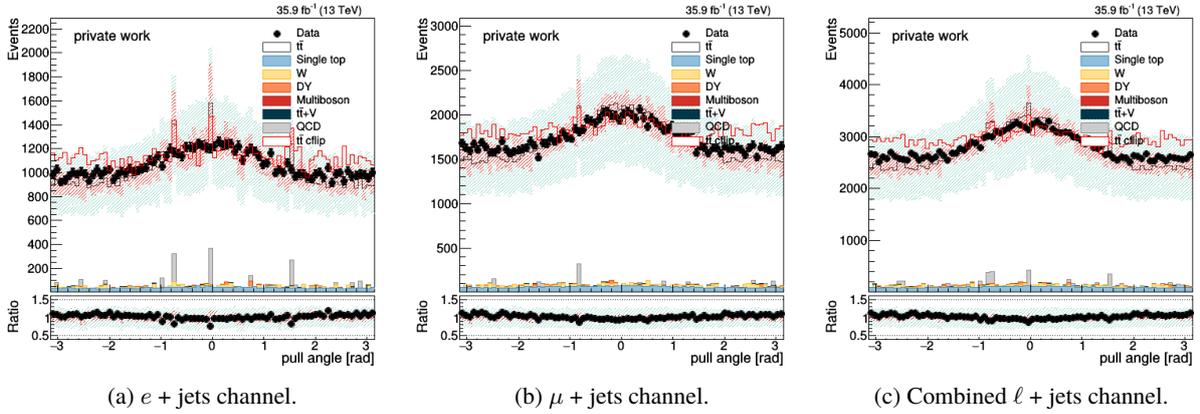


Figure 8.16. Distribution of the pull angle from  $j_2^W$  to  $j_1^W$  for all  $\Delta R$  and including all particles.

Additionally, the plots of the pull angle between jets where we expect no colour connection –  $j_1^b$  to  $j_2^b$  and  $j_2^b$  to  $j_1^b$  with all jet constituents and including all values of  $\Delta R$  are shown in Fig. 8.17 and Fig. 8.18.

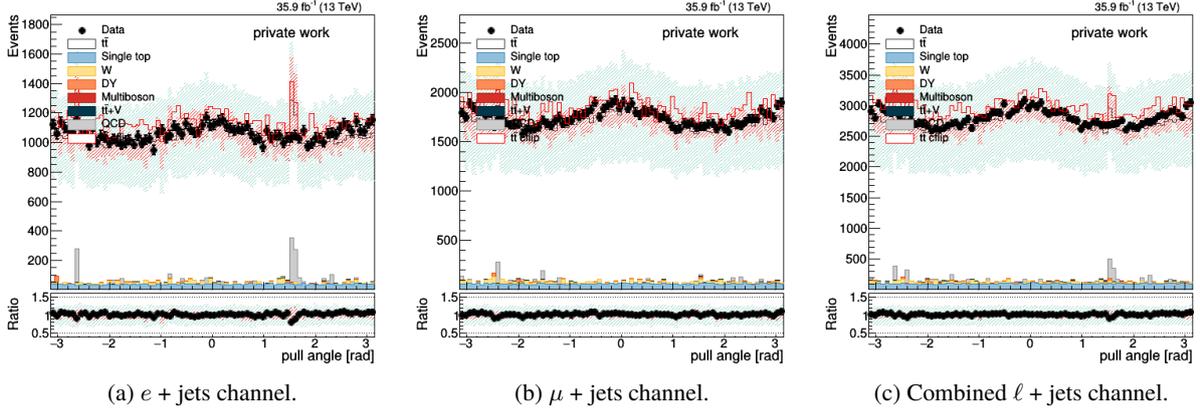


Figure 8.17. Distribution of the pull angle from  $j_1^b$  to  $j_2^b$  for all  $\Delta R$  and including all particles.

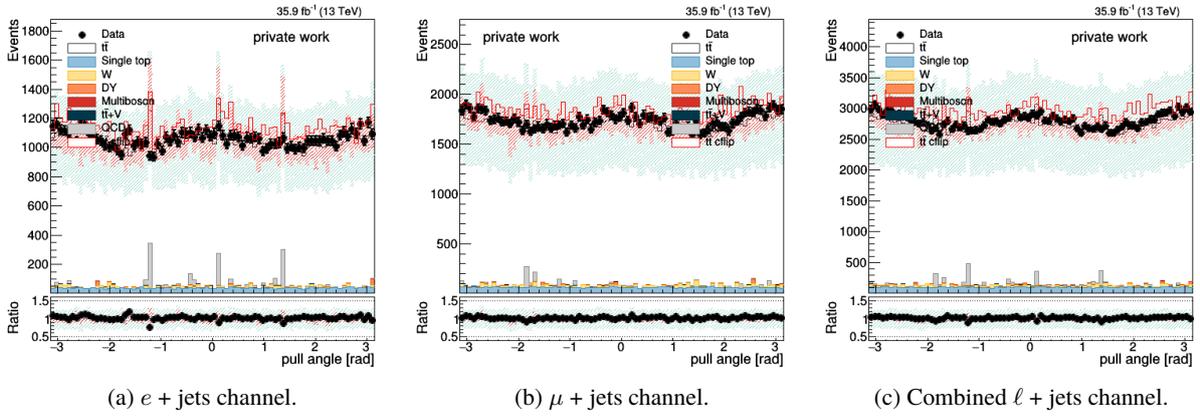


Figure 8.18. Distribution of the pull angle from  $j_2^b$  to  $j_1^b$  for all  $\Delta R$  and including all particles.

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. 8.19 shows the distribution of pull angle from  $j_1^W$  to the charged lepton.

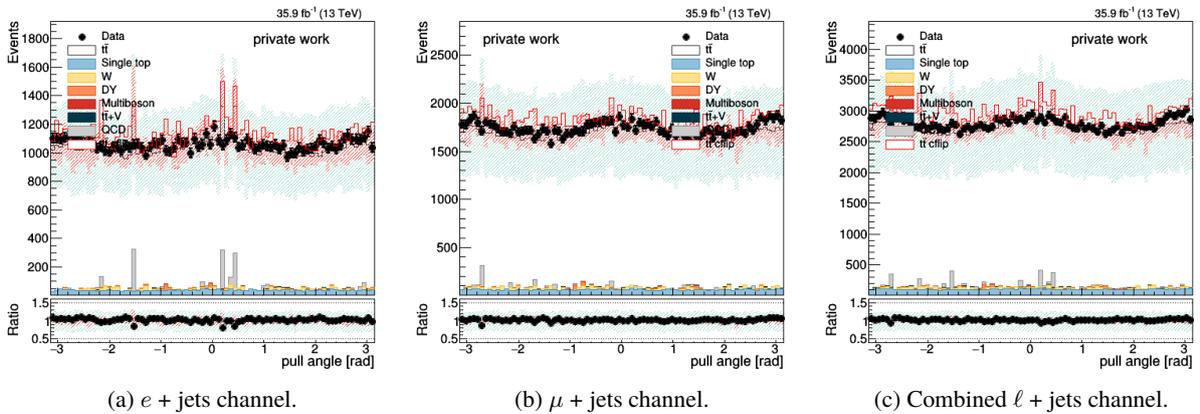


Figure 8.19. Distribution of the pull angle from  $j_1^W$  to the charged lepton for all  $\Delta R$  and including all particles.

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 from  $j_1^W$  to the hadronic  $W$  as shown in Fig. 8.20.

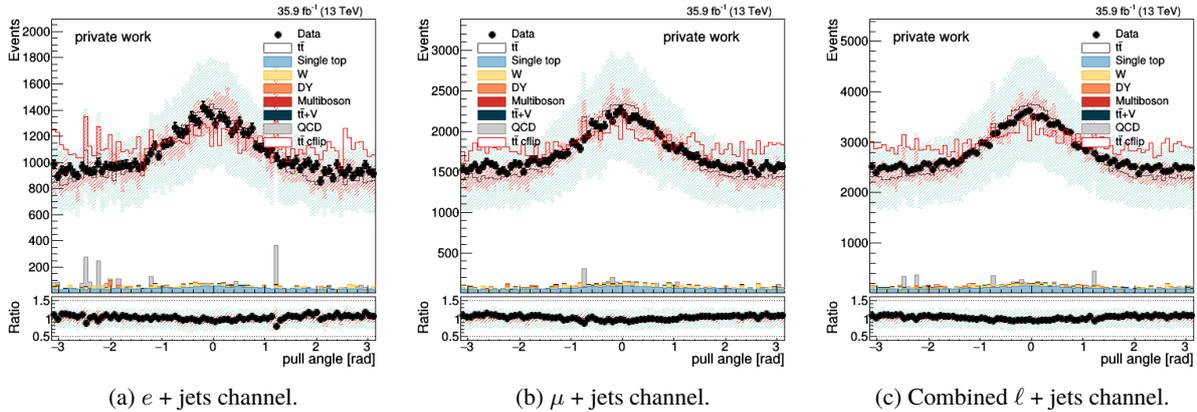


Figure 8.20. Distribution of the pull angle from  $j_1^W$  to the hadronic  $W$  for all  $\Delta R$  and including all particles.

Another interesting case is choosing the beam. In Fig. 8.21 we show the distribution of  $\theta_p$  from  $j_1^W$  to the positive direction of the beam. We see a peak at a right angle.

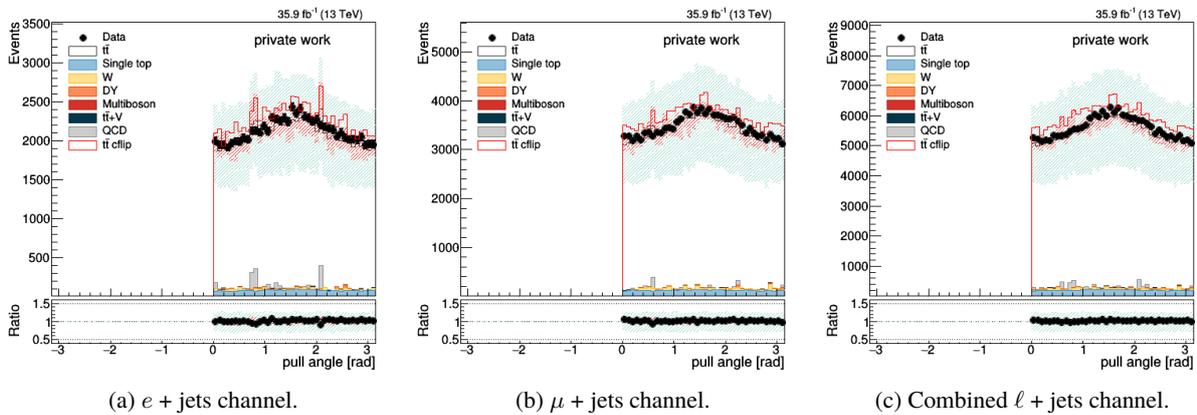


Figure 8.21. Distribution of the pull angle from  $j_1^W$  to the positive direction of the beam including all particles.

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

$$w = \mathcal{L} \cdot \sigma \frac{1}{N_{\text{gen}}}. \quad (8.2)$$

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

### 8.3 $\Delta R$ bias

When two jets are close to each other in  $\phi$ - $\eta$  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. 8.22–8.25 illustrate the distribution of pull angle for two cases – closely spaced jets with  $\Delta R \leq 1.0$  and well separated jets with  $\Delta R > 1.0$ .

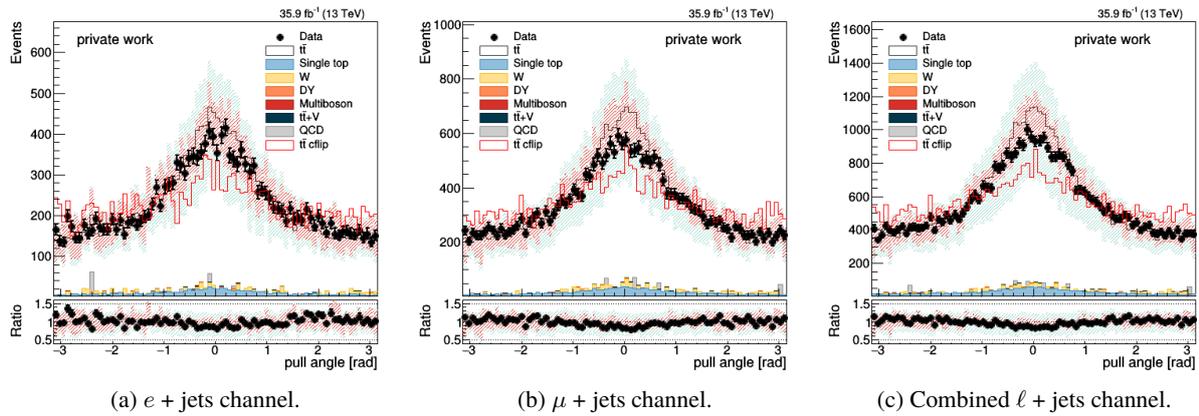


Figure 8.22. Distribution of the pull angle with  $\Delta R \leq 1.0$  and including all jet constituents from  $j_1^W$  to  $j_2^W$ .

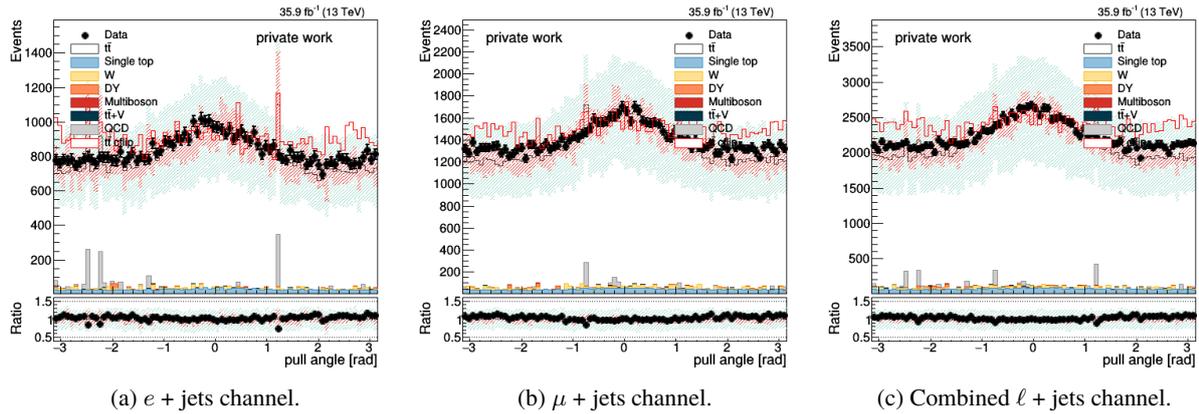


Figure 8.23. Distribution of the pull angle with  $\Delta R > 1.0$  and including all jet constituents from  $j_1^W$  to  $j_2^W$ .

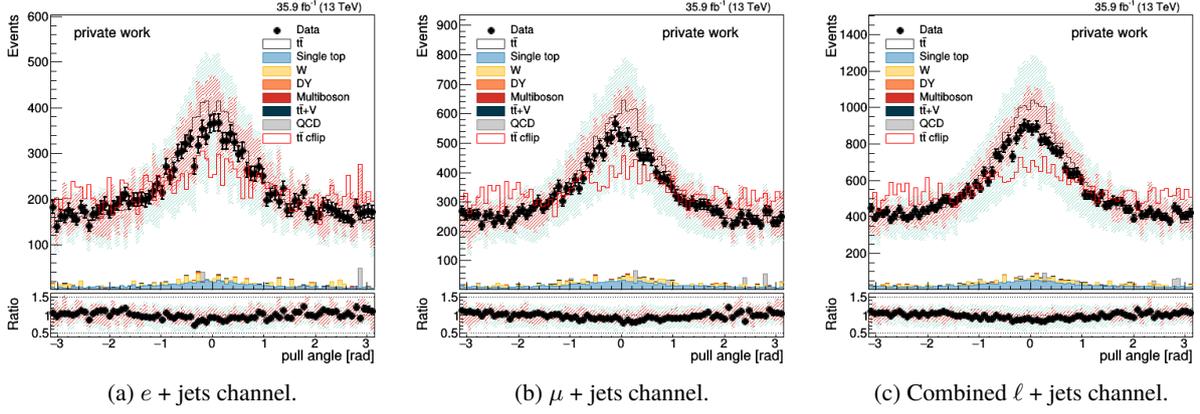


Figure 8.24. Distribution of the pull angle with  $\Delta R \leq 1.0$  and only charged jet constituents from  $j_1^W$  to  $j_2^W$ .

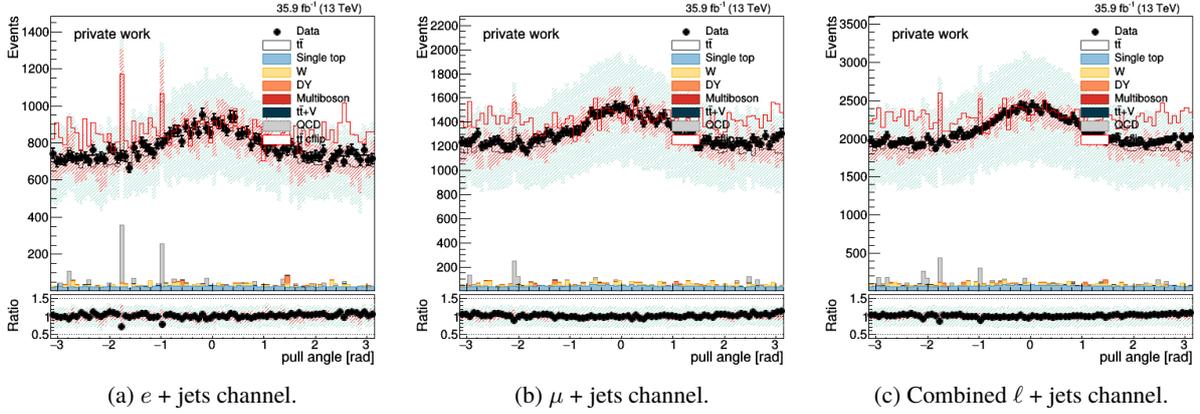


Figure 8.25. Distribution of the pull angle with  $\Delta R > 1.0$  and only charged jet constituents from  $j_1^W$  to  $j_2^W$ .

## 8.4 Sensitivity analysis

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

**The transverse momentum  $p_T$  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. This cut is near the median in the lower half of the distribution of the  $p_T$  of the hadronic  $W$  boson which is shown in Fig. 8.26. The results are shown in Figs. 8.28–8.29.

**Number of jet constituents.** A cut was chosen at the number of jet constituents  $N$  being 20 and the distribution of the pull angle was obtained at  $N$  greater than and less than or equal to this value. The distribution of the number of jet constituents is given in Fig. 2.9. The results are shown in Figs. 8.30–8.31.

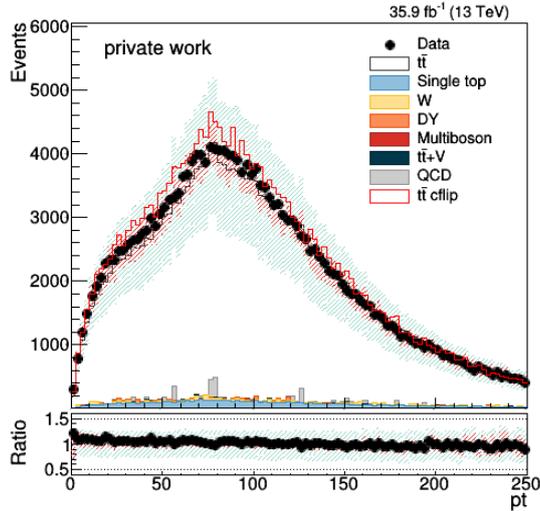


Figure 8.26. Distribution of the transverse momentum  $p_T$  of the hadronic  $W$  boson.

**The transverse momentum  $p_T$  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 obtained at  $p_T$  of the jet constituents being greater than and less than or equal to this value. The distribution of the transverse momentum of particles constituting the leading light jet and the leading  $b$  jet are shown in Fig. 8.27. The results are shown in Figs. 8.32–8.33.

**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 distribution of the magnitude of the pull vector is shown in Figs. 8.11–8.11. The results are shown in Figs. 8.34–8.35.

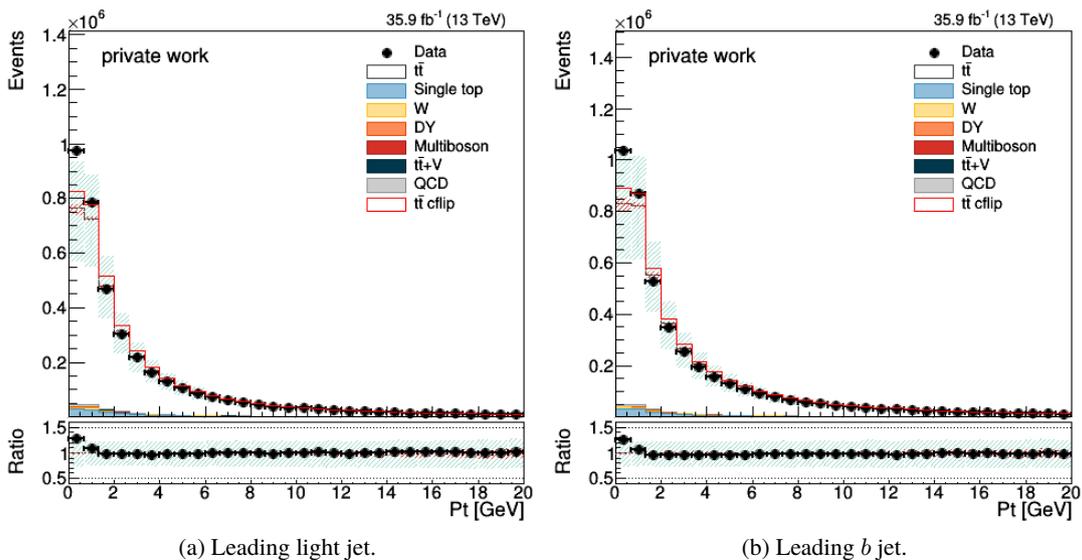


Figure 8.27. The transverse momentum  $p_T$  of particles constituting the leading light jet and the leading  $b$  jet. In both cases all jet constituents are included.

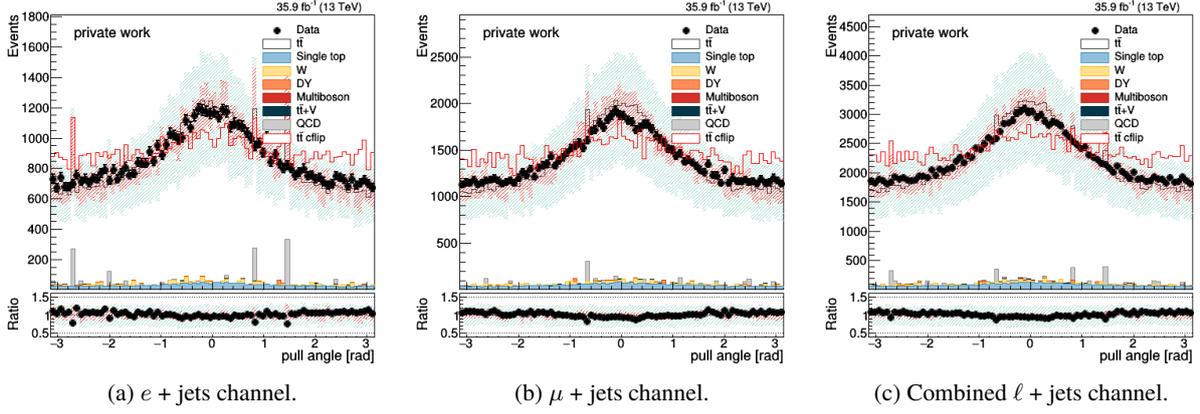


Figure 8.28. Distribution of the pull angle for all  $\Delta R$  and all particles from  $j_1^W$  to  $j_2^W$  with  $p_T$  of  $W > 50$  GeV.

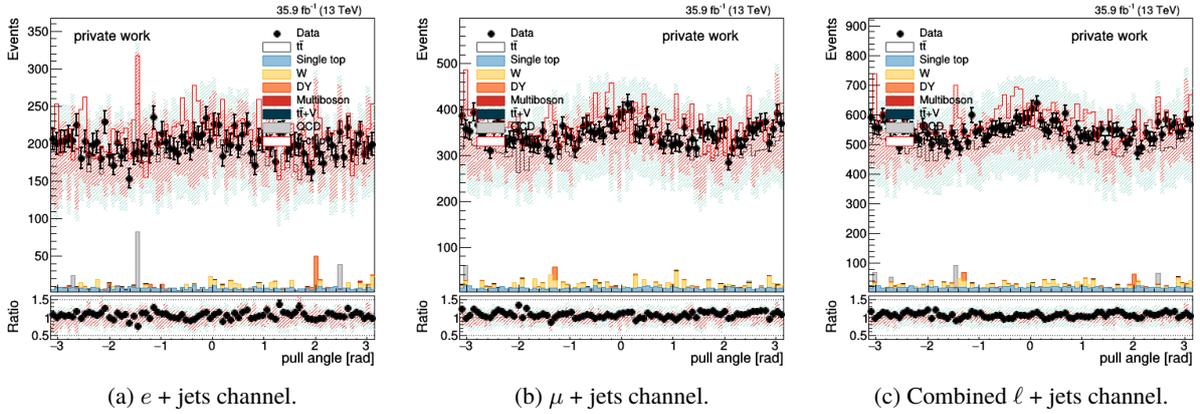


Figure 8.29. Distribution of the pull angle for all  $\Delta R$  and all particles from  $j_1^W$  to  $j_2^W$  with  $p_T$  of  $W \leq 50$  GeV.

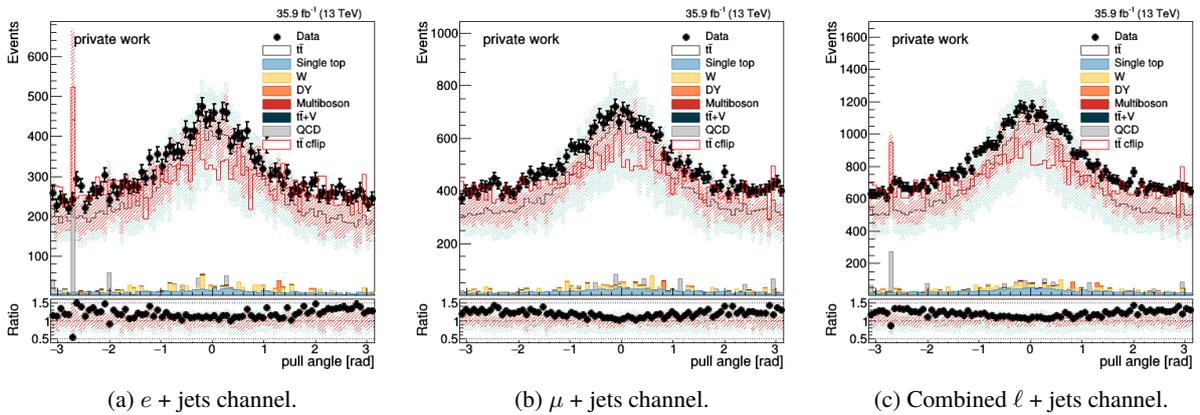


Figure 8.30. Distribution of the pull angle for all  $\Delta R$  and all particles from  $j_1^W$  to  $j_2^W$  with the number of jet constituents  $N > 20$ .

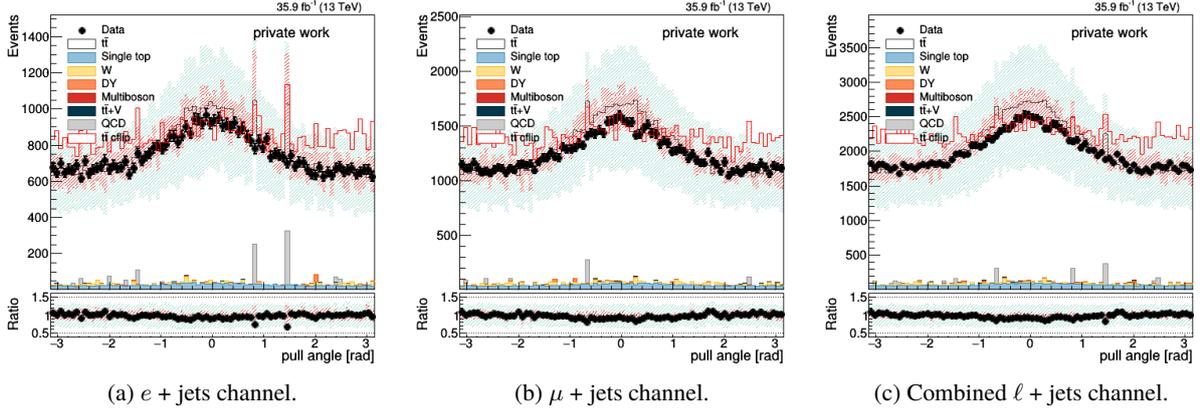


Figure 8.31. Distribution of the pull angle for all  $\Delta R$  and all particles from  $j_1^W$  to  $j_2^W$  with the number of jet constituents  $N \leq 20$ .

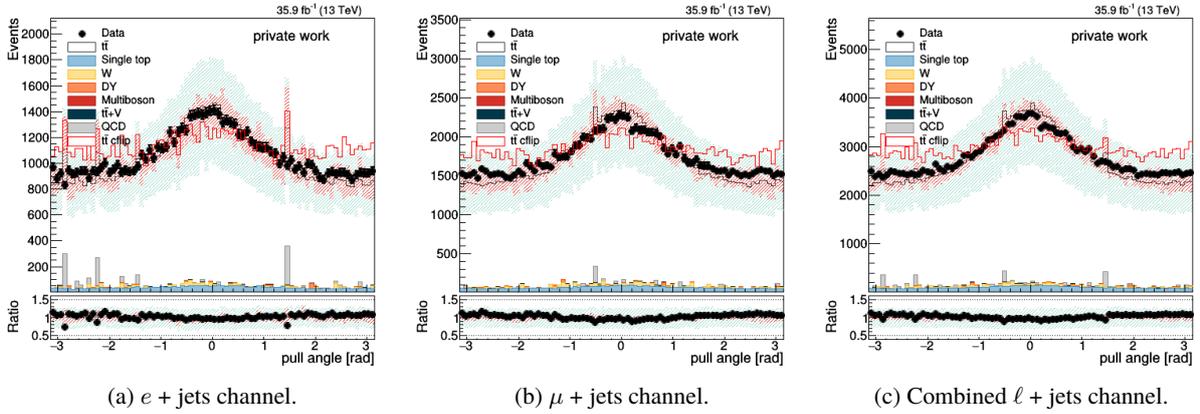


Figure 8.32. Distribution of the pull angle for all  $\Delta R$  and all particles from  $j_1^W$  to  $j_2^W$  with the  $p_T$  of jet constituents  $> 0.5$  GeV.

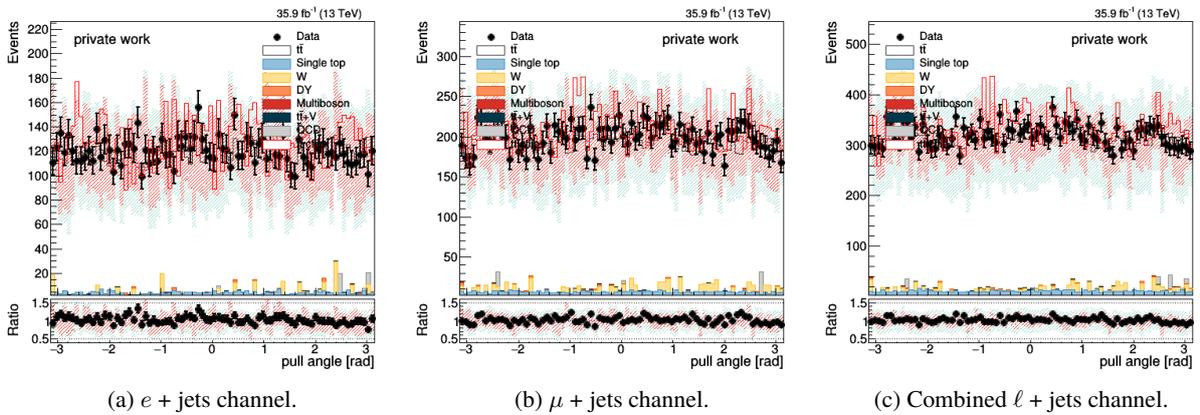


Figure 8.33. Distribution of the pull angle for all  $\Delta R$  and all particles from  $j_1^W$  to  $j_2^W$  with the  $p_T$  of jet constituents  $\leq 0.5$  GeV.

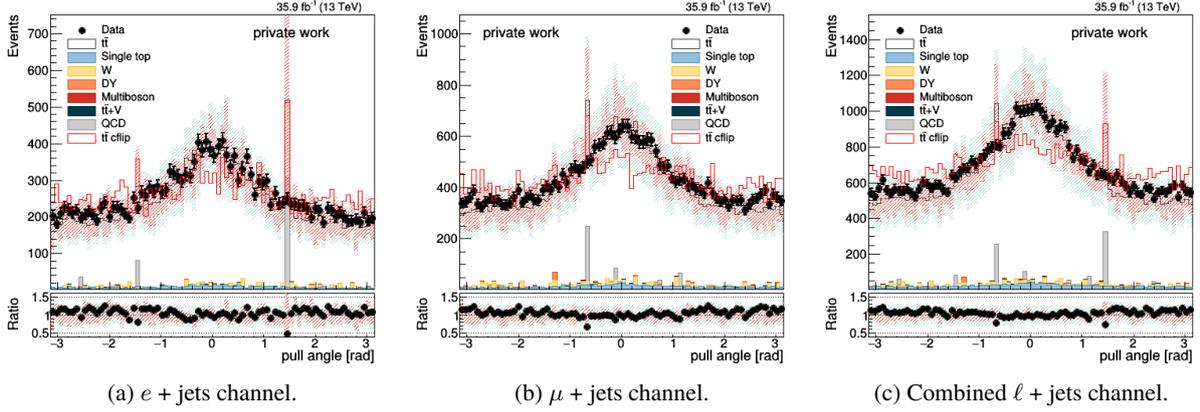


Figure 8.34. Distribution of the pull angle for all  $\Delta R$  and all particles from  $j_1^W$  to  $j_2^W$  with the magnitude of the pull vector  $> 0.005$  [a.u.].

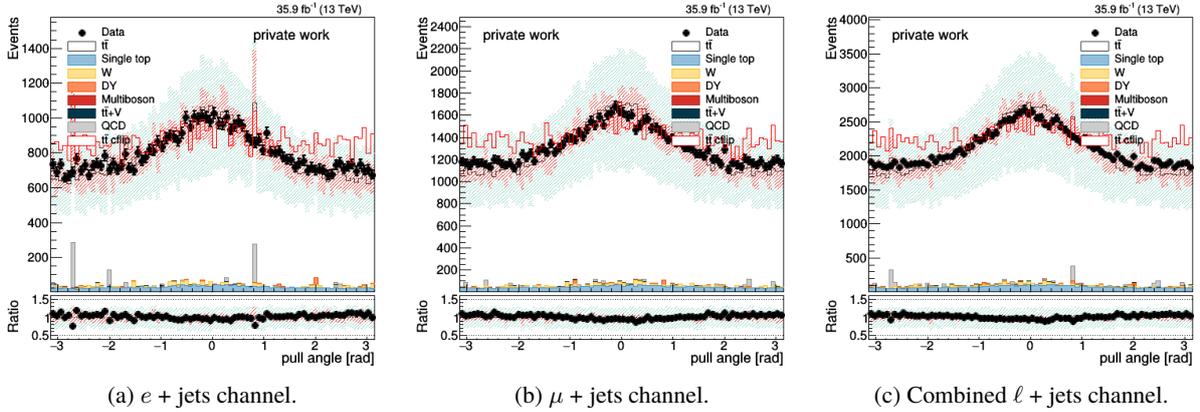


Figure 8.35. Distribution of the pull angle for all  $\Delta R$  and all particles from  $j_1^W$  to  $j_2^W$  with the magnitude of the pull vector  $\leq 0.005$  [a.u.].

From a simple qualitative observation we conclude that the pull angle methodology is sensitive to  $p_T$  of the hadronic  $W$  boson, number of jet constituents,  $p_T$  of jet constituents but not particularly sensitive to the magnitude of the pull vector.

## 8.5 Unfolding

When a detector makes an observation the end results suffer from the inefficiencies of the detector. Unfolding is a method where the observation made at the detector is corrected for detector effects. Hence we can obtain an estimate of the true distribution of the observable. However, it comes at a cost of a significant loss of granularity of the phase space of the observable.

We infer about the detector effects because in Monte Carlo samples each generated event is reconstructed. Therefore an observable in bin  $i$  at generation level migrates to bin  $k$  at reconstruction level. By accumulating a large number of events we obtain statistics of migration. In unfolding we revert the migration – given an observable at bin  $k$  we assign probabilities to the true values of the observable.

Values of  $\theta_p$  at generation level that do not have a corresponding value at reconstruction level are put in the underflow bin at reconstruction level. Values of  $\theta_p$  at reconstruction level that do not have a corresponding value at generation level are put in the underflow bin at generation level. The underflow bins at generation level are treated as background and are removed. Distributions that are not filled at generation level – data and MC backgrounds are reduced by a corresponding scale factor. The underflow bin at reconstruction level is used to constrain the underflow bin for the unfolded result.

Unfolding is performed on data from which the MC backgrounds have been subtracted. We also performed the unfolding procedure in reverse obtaining the folded back output.

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:

$$\text{stability} \equiv \frac{\theta_{\text{input}}^{\text{diag}}}{\sum_{x=1}^{N_x} \theta_{\text{input}}^x}, \quad (8.3)$$

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:

$$\text{purity} \equiv \frac{\theta_{\text{input}}^{\text{diag}}}{\sum_{y=1}^{N_y} \theta_{\text{input}}^y}, \quad (8.4)$$

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. A mathematical expression for the pull is

$$\text{pull} \equiv \frac{\theta_{\text{unf}}^{\text{gen}} - \theta_{\text{in}}^{\text{gen}}}{\sigma_{\text{unf}}^{\text{gen}}}. \quad (8.5)$$

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

The number of bins at generation level is reduced by a factor of 2 with regard to the number of bins at reconstruction level in order that unfolding be computationally feasible.

The class TUnfoldDensity [122] of Root is used to do the unfolding procedure. The binning scheme is managed with class TUnfoldBinning. No regularisation is applied. The unfolding results of  $\theta_p$  from  $j_1^W$  to  $j_2^W$  including all jet constituents are shown in Fig. 8.37. In order to create the plots shown herein a new class CompoundHistoUnfolding [123] was developed and it was added to Root complete with input and output streamers.

The unfolding results are shown in Fig. 8.37. Distributions corresponding to unfolding results with migration matrices from  $t\bar{t}$  *Herwig* ++ and  $t\bar{t}$  *cflip* as well as systematics  $t\bar{t}$  *f sr dn* and  $t\bar{t}$  *f sr up* (see Chap. 7) are laid over the unfolding plots. In the unfolded distribution there

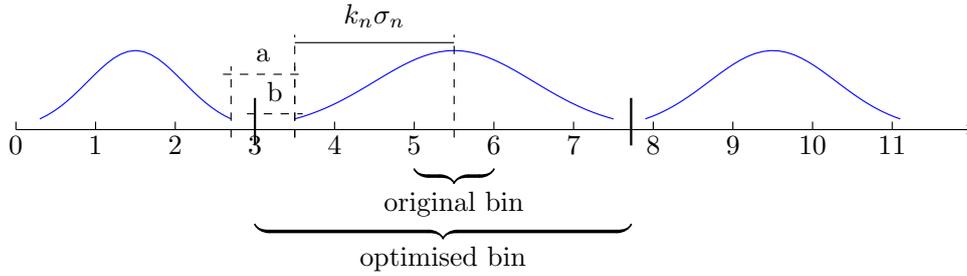


Figure 8.36. Method of optimising the binning scheme for unfolding.

are very large uncertainties and poor purity and stability in most of the bins. In order to mitigate these effects a bin optimisation algorithm was tried. The algorithm proceeds as follows:

- In each bin of the reconstructed observable, the particle-level distribution is fitted with a gaussian distribution.
- Starting from the low edge of the reconstructed distribution, a bin  $i$  is searched that fulfils  $\mu_i - f\sigma_i/2 > 0$ ,  $\mu_i$  being the mean and  $\sigma_i$  the standard distribution from the fit at generation level in each bin  $i$  at reconstruction level. The factor  $f$  is chosen so that  $f\sigma \sim (\theta_{p, \max} - \theta_{p, \min})/3$ , in order to obtain 4–3 optimised bins.  $\theta_{p, \min}$  is the lower value of  $\theta_p$  ( $=0$ ) while  $\theta_{p, \max}$  is the upper value of  $\theta_p$  ( $=\pi$ ). In practice, this factor has to be chosen very small – 0.15 for pull angle and 0.3 for the magnitude of the pull vector, indicating that  $\sigma$  is large compared to the range of the phase space of the pull angle.
- The new optimised bin is then defined from 0 to mean  $+ f\sigma_i/2$ .
- The algorithm is iterated until the edge of the histogram is reached.
- The obtained binning is used to present the result on particle (i.e. generation) level. To obtain the final migration matrix used in the unfolding each bin at particle level is split by two to obtain a suitable reconstruction level binning.

This algorithm is depicted in Fig. 8.36. The parameter  $b$  is given by  $\sigma_n/(\sigma_n + \sigma_{n+1})$ .

The unfolded result with the optimised binning is shown in Fig. 8.38. The purity and stability in the central bin is still poor. Therefore a scheme using 3 regular bins as in the ATLAS analysis [51] was tried.

The results with the regular binning scheme are shown in Fig. 8.39. The stability and purity levels with this binning scheme reach acceptable levels at each bin and it was adopted for further analysis.

The unfolding results using the migration matrix from the sample  $t\bar{t}$  *cflip* are shown in Fig. 8.40. The  $t\bar{t}$  *cflip* is included as a systematic for  $t\bar{t}$ .

The unfolding results of the  $\theta_p$  from  $j_1^b$  to  $j_2^b$  with all jet constituents are shown in Fig. 8.41.

As an additional observable the magnitude of the pull vector  $|\vec{P}|$  was unfolded. Fig. 8.42 shows the unfolding results of  $|\vec{P}|$  from  $j_1^W$  to  $j_2^W$  including all jet constituents.

The bin-per-bin significance (%) of nuisances in the total systematical error in the unfolded result is given in Table 8.1. Nuisances that directly affect the hadronisation  $t\bar{t}$  *Herwig* + +,  $t\bar{t}$  *QCDbased* and  $t\bar{t}$  *ERDon* are the most significant.

In addition to the POWHEG +PYTHIA 8 sample, we also investigate a POWHEG +PYTHIA 8 \* sample in which  $t\bar{t}$  *cflip* has been added as a systematic to  $t\bar{t}$ . Table 8.2 shows the additional bin-per-bin  $t\bar{t}$  *cflip* uncertainty for the POWHEG +PYTHIA 8 \* sample.

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  $\Sigma$  of normalised experimental uncertainties, the  $\chi^2$  is calculated as follows:

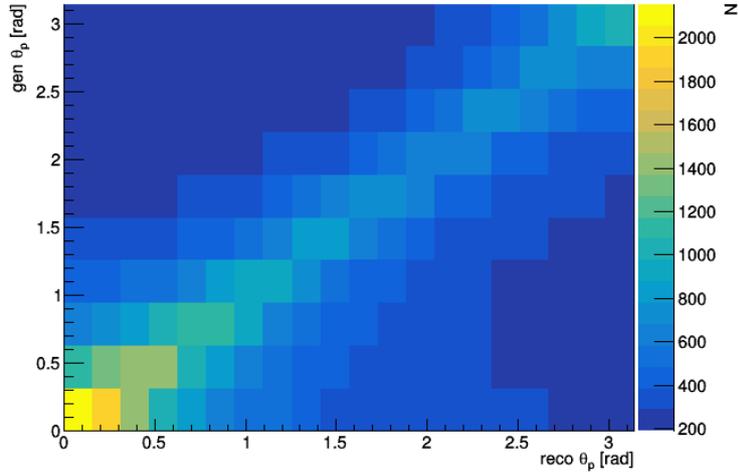
$$\chi^2 = (D^T - M^T) \cdot \Sigma^{-1} \cdot (D - M). \quad (8.6)$$

From the  $\chi^2$  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  $\Sigma$ .  $\chi^2$  value does not depend on the choice of the discarded elements.

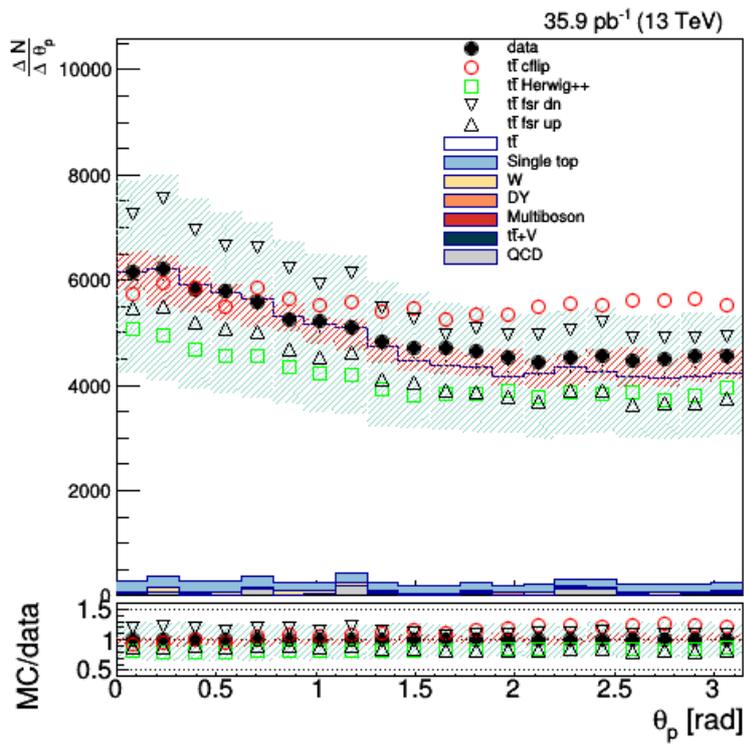
Table 8.3 shows the  $\chi^2$  values and  $p$ -values for  $\theta_p$  using all jet constituents. The results show that the pull angle distribution is poorly modelled by the MC generators. In general, the simulation predicts a more sloped distribution, i.e. a stronger colour flow effect. HERWIG++ models better the pull angle distribution than PYTHIA 8.2. Accuracy of PYTHIA 8.2 is particularly poor when predicting the distribution of  $\theta_p$  from  $j_2^W$  to  $j_1^W$ .

The  $\chi^2$  values and  $p$ -values for the  $W$  colour octet model are given in Table 8.4. In the colour flip model the distribution of  $\theta_p$  from  $j_1^W$  to  $j_2^W$  is modelled less accurately than the SM prediction.

Table 8.7 shows the values of  $\chi^2$  and if signal  $M$  in Eq. 8.6 is replaced by the respective systematic, but leaving the covariance matrix  $\Sigma$  unchanged. The agreement is better than  $t\bar{t}$  when the colour flow is modelled with the  $t\bar{t}$  *ERDOn*,  $t\bar{t}$  *Herwig ++* and  $t\bar{t}$  *QCDbased* setup.

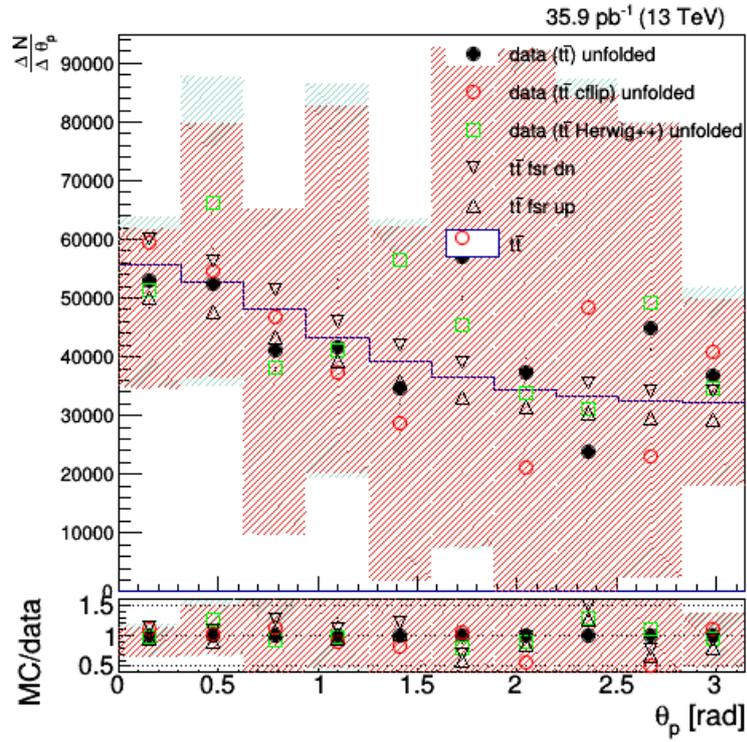


(a) Migration matrix with the reconstruction level as the  $x$  axis and the generation level as the  $y$  axis.

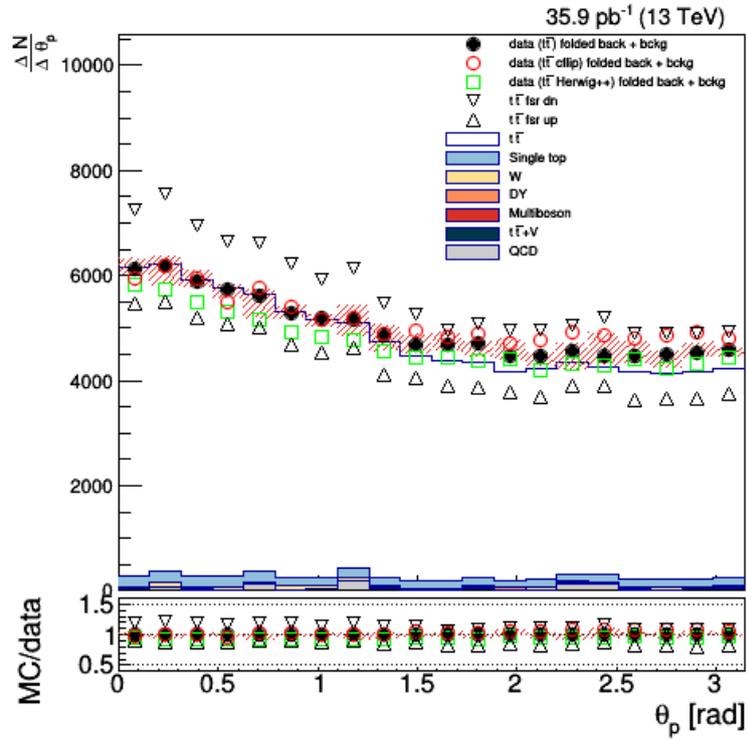


(b) The data and Monte Carlo used as input.

Figure 8.37. Unfolding plots for the  $t\bar{t}$  method of the pull angle  $\theta_p$  of  $j_1^W$  to  $j_2^W$  including all jet constituents using the original binning.

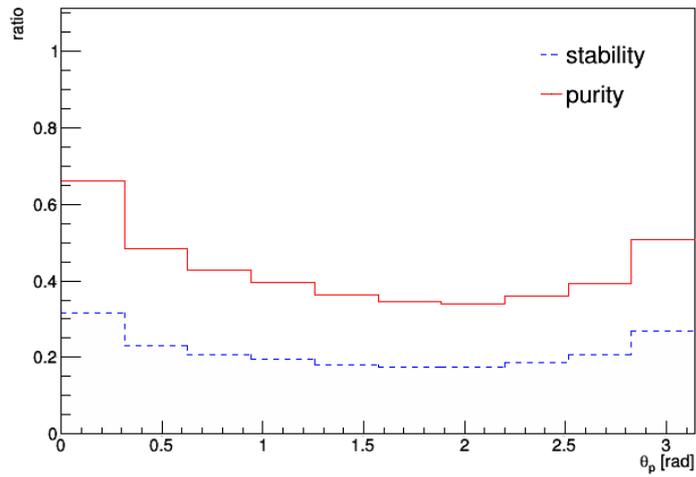


(c) The unfolded result.

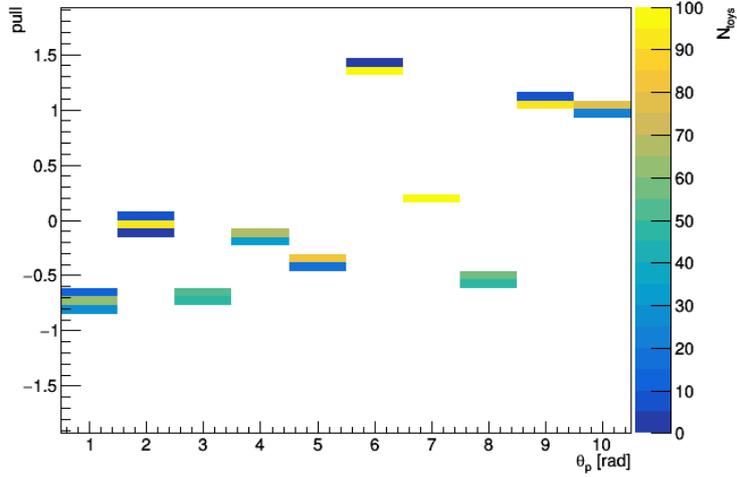


(d) The folded back result.

Figure 8.37. Unfolding plots for the  $t\bar{t}$  method of the pull angle  $\theta_p$  of  $j_1^W$  to  $j_2^W$  including all jet constituents using the original binning (continued).

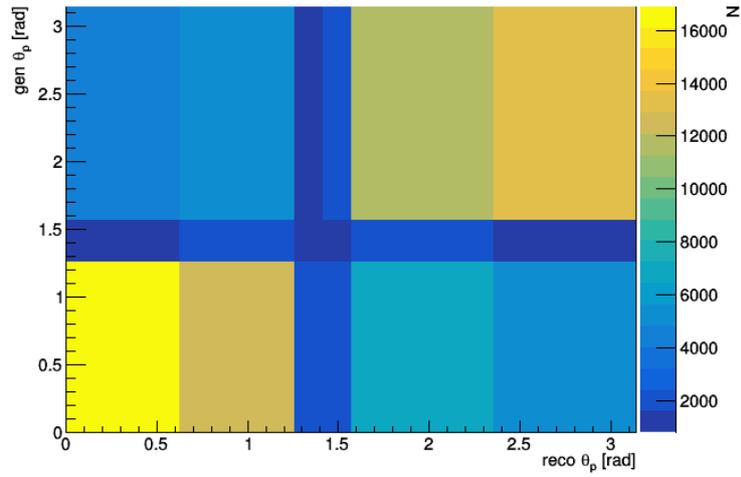


(e) Stability and purity in each bin.

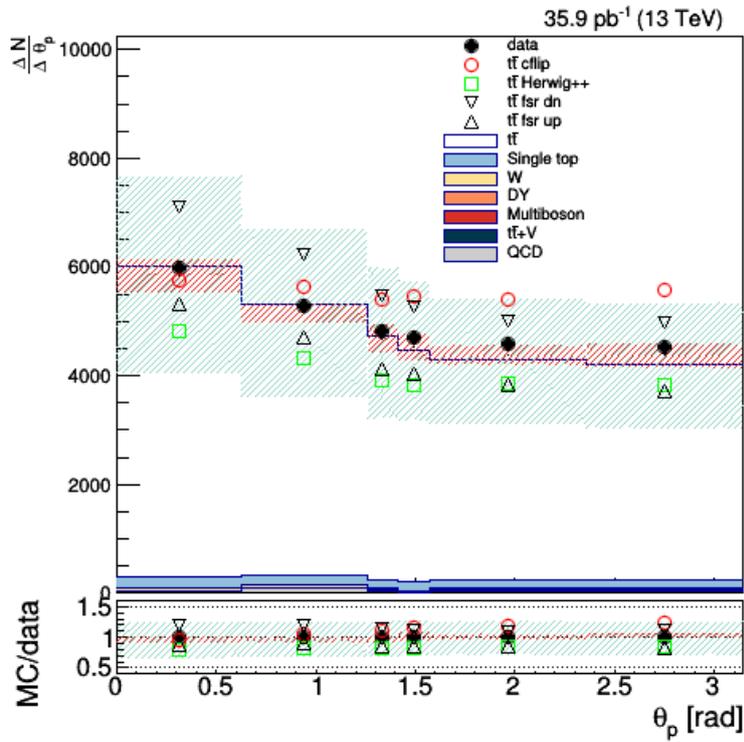


(f) The pull.

Figure 8.37. Unfolding plots for the  $t\bar{t}$  method of the pull angle  $\theta_p$  of  $j_1^W$  to  $j_2^W$  including all jet constituents using the original binning (continued).

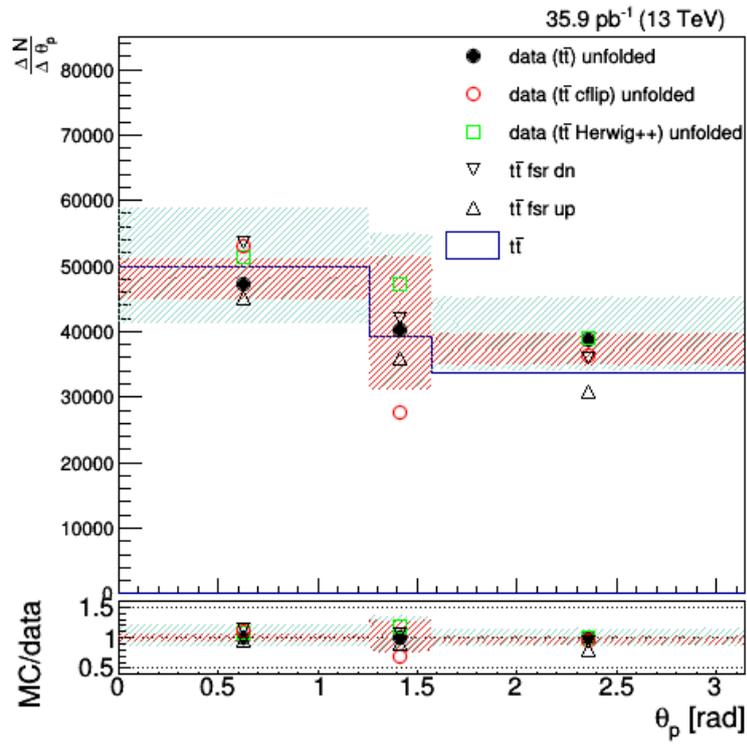


(a) Migration matrix with the reconstruction level as the  $x$  axis and the generation level as the  $y$  axis.

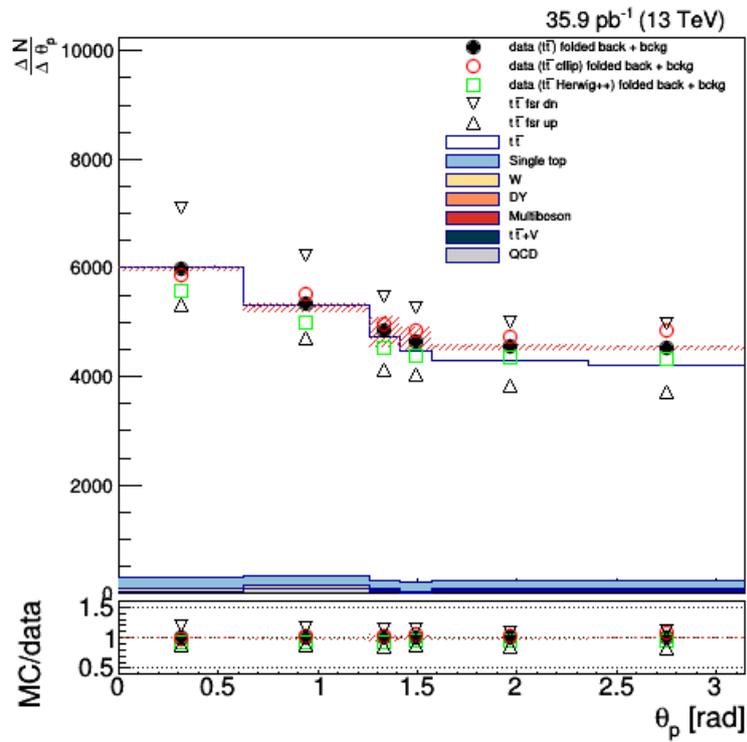


(b) The data and Monte Carlo used as input.

Figure 8.38. Unfolding plots for the  $t\bar{t}$  method of the pull angle  $\theta_p$  of  $j_1^W$  to  $j_2^W$  including all jet constituents using the optimised binning with a  $\sigma$  factor of 0.15.

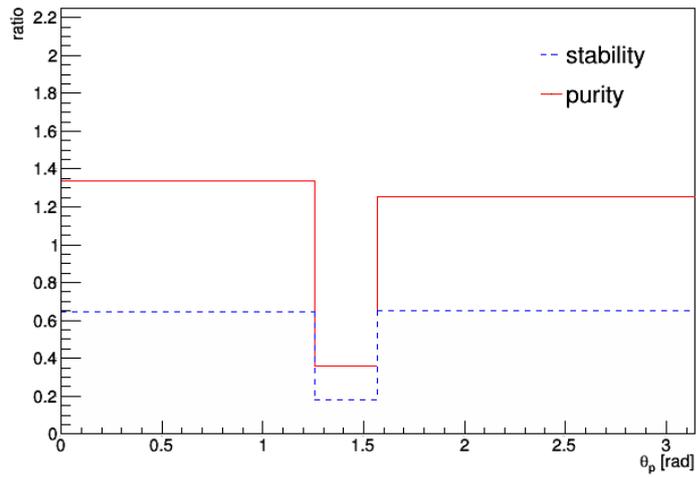


(c) The unfolded result.

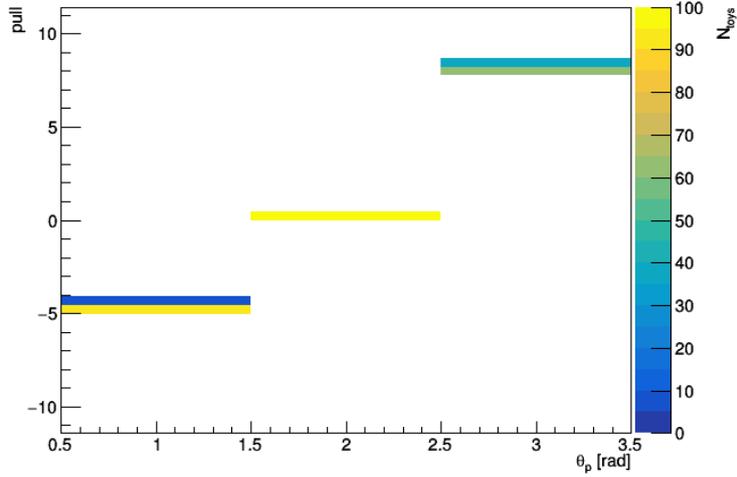


(d) The folded back result.

Figure 8.38. Unfolding plots for the  $t\bar{t}$  method of the pull angle  $\theta_p$  of  $j_1^W$  to  $j_2^W$  including all jet constituents using the optimised binning with a  $\sigma$  factor of 0.15 (continued).

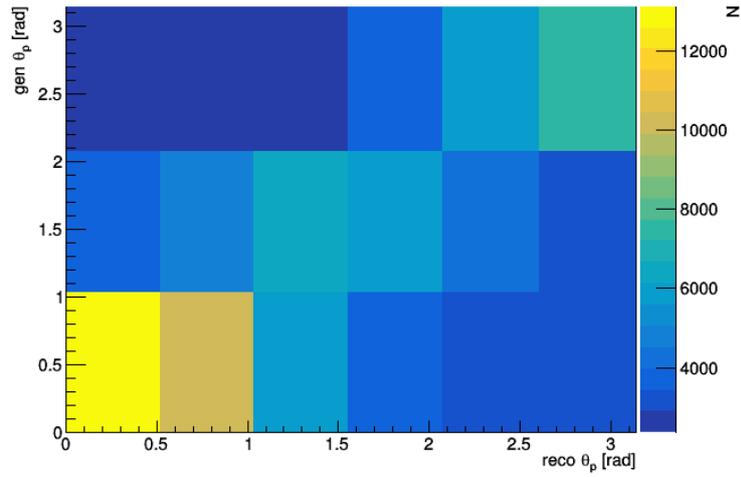


(e) Stability and purity in each bin.

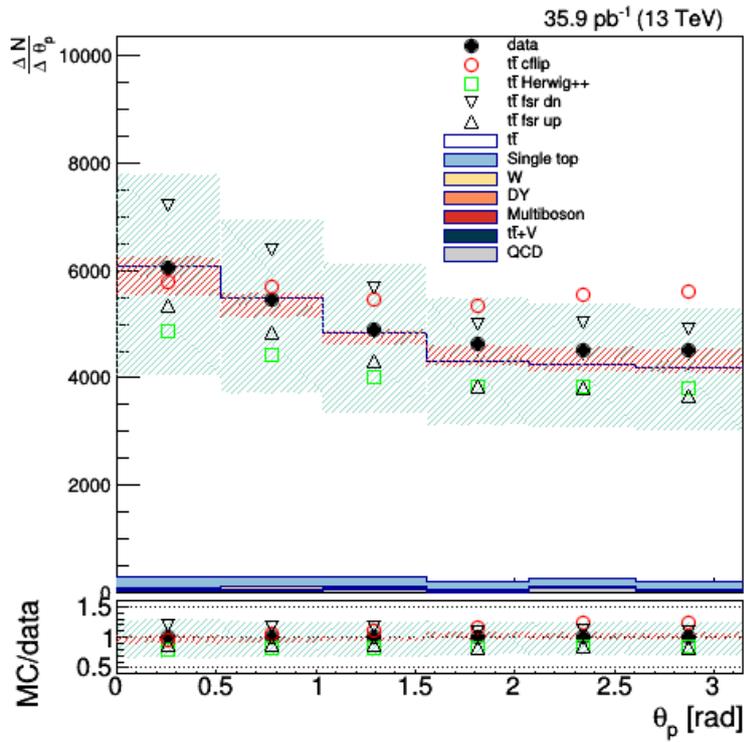


(f) The pull.

Figure 8.38. Unfolding plots for the  $t\bar{t}$  method of the pull angle  $\theta_p$  of  $j_1^W$  to  $j_2^W$  including all jet constituents using the optimised binning with a  $\sigma$  factor of 0.15 (continued).

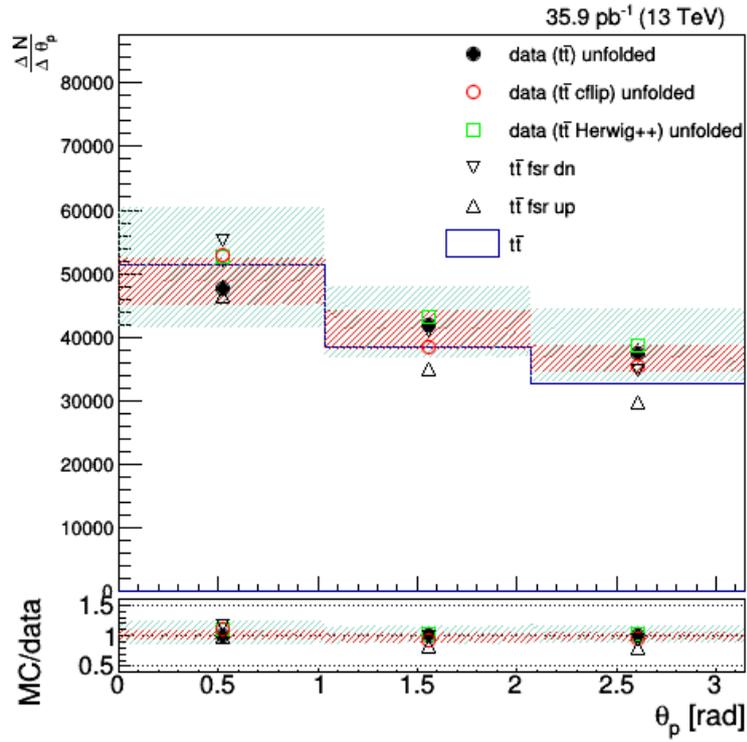


(a) Migration matrix with the reconstruction level as the  $x$  axis and the generation level as the  $y$  axis.

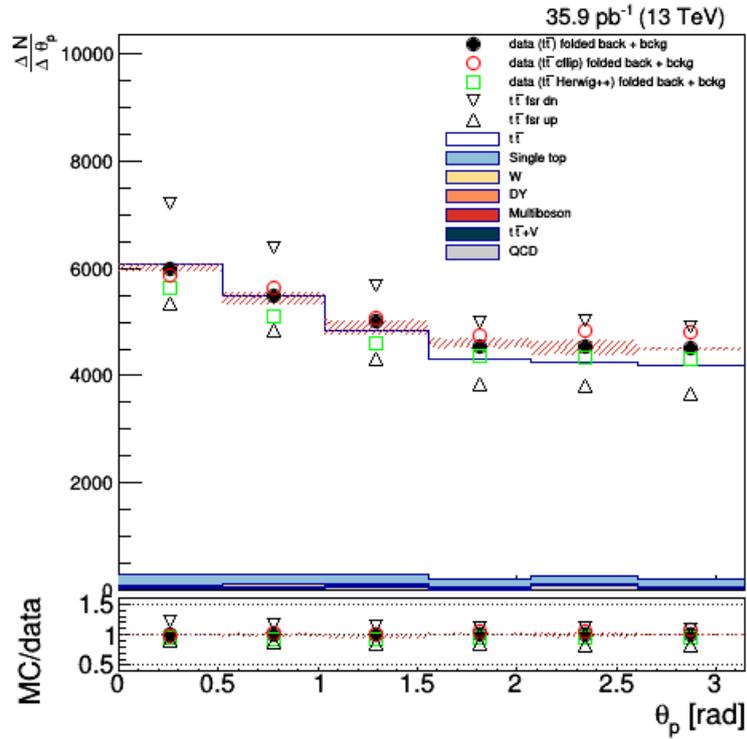


(b) The data and Monte Carlo used as input.

Figure 8.39. Unfolding plots for the  $t\bar{t}$  method of the pull angle  $\theta_p$  of  $j_1^W$  to  $j_2^W$  including all jet constituents using 3 regularly sized bins.

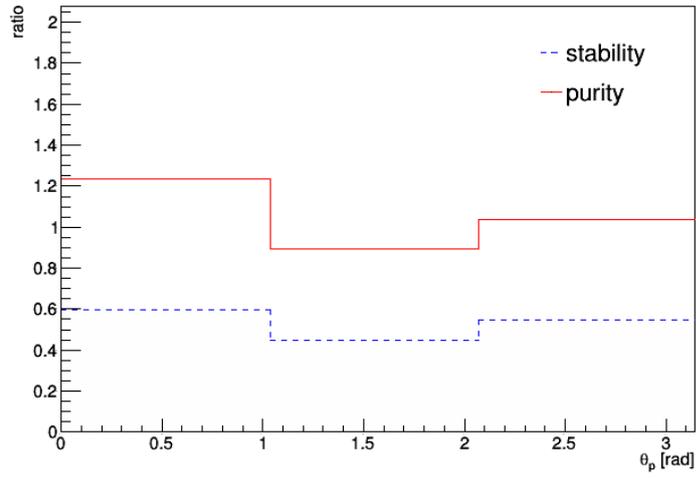


(c) The unfolded result.

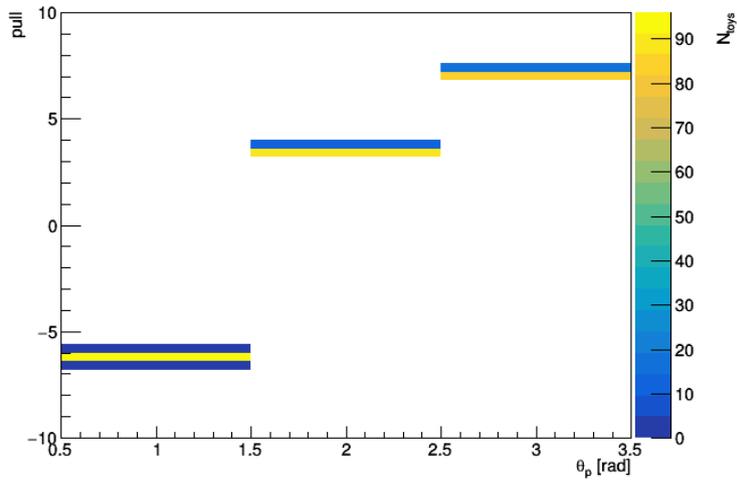


(d) The folded back result.

Figure 8.39. Unfolding plots for the  $t\bar{t}$  method of the pull angle  $\theta_p$  of  $j_1^W$  to  $j_2^W$  including all jet constituents using 3 regularly sized bins (continued).

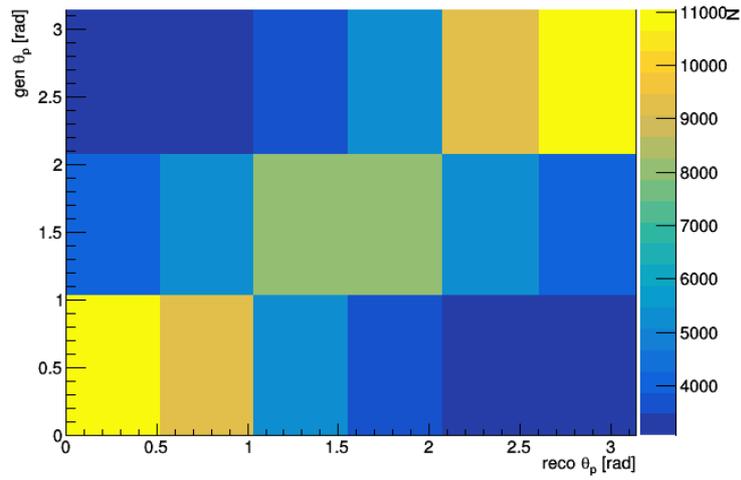


(e) Stability and purity in each bin.

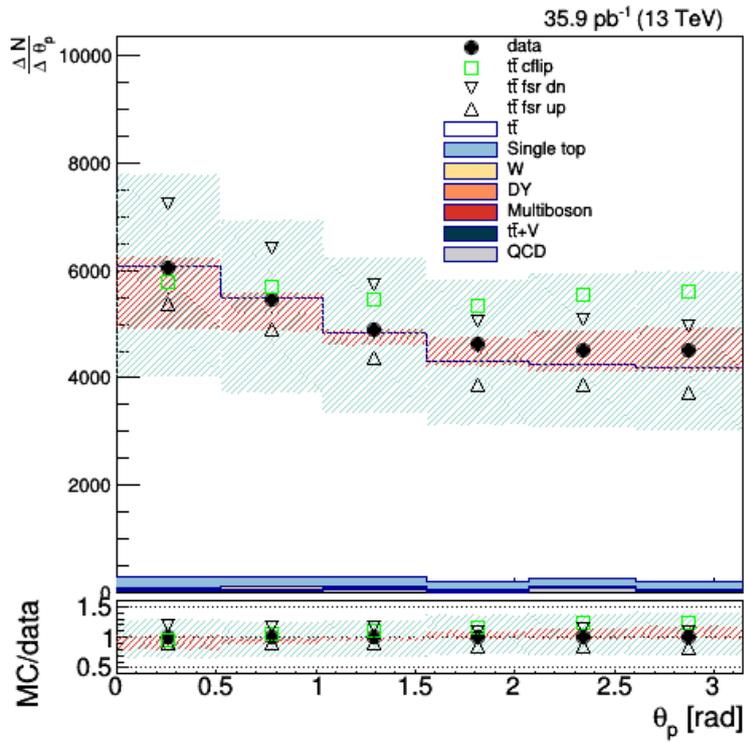


(f) The pull.

Figure 8.39. Unfolding plots for the  $t\bar{t}$  method of the pull angle  $\theta_p$  of  $j_1^W$  to  $j_2^W$  including all jet constituents using 3 regularly sized bins (continued).

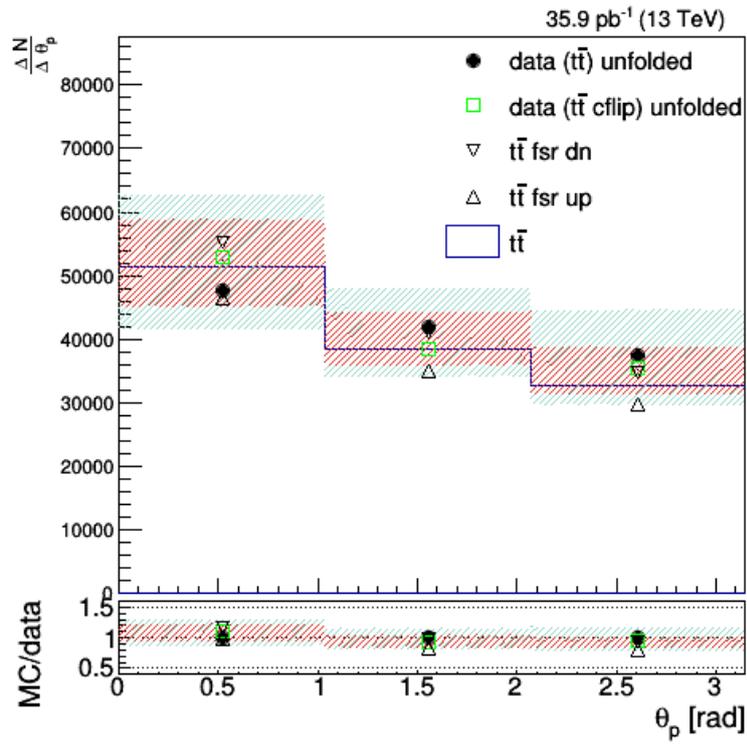


(a) Migration matrix with the reconstruction level as the  $x$  axis and the generation level as the  $y$  axis.

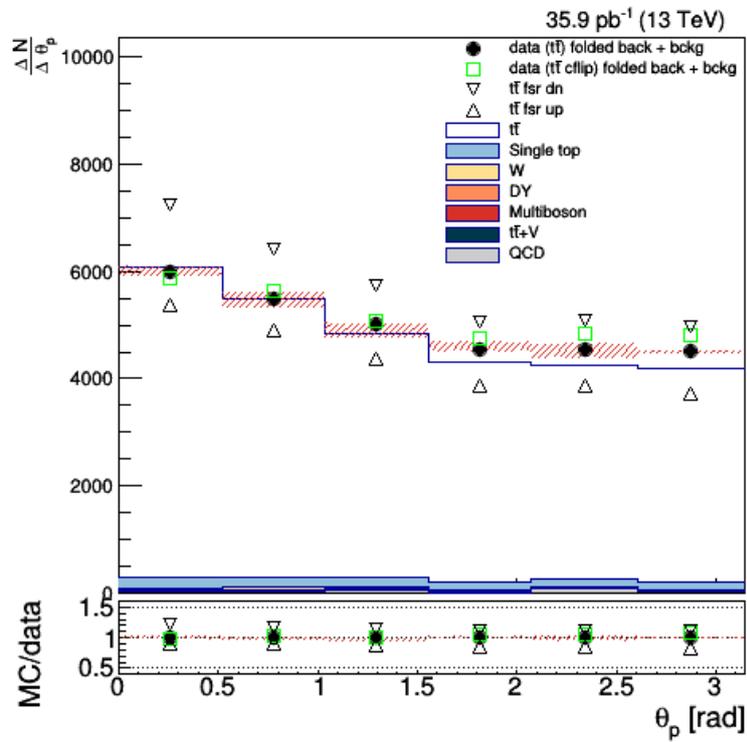


(b) The data and Monte Carlo used as input.

Figure 8.40. Unfolding plots for the  $t\bar{t}$  cflip method of the pull angle  $\theta_p$  of  $j_1^W$  to  $j_2^W$  including all jet constituents using 3 regularly sized bins.

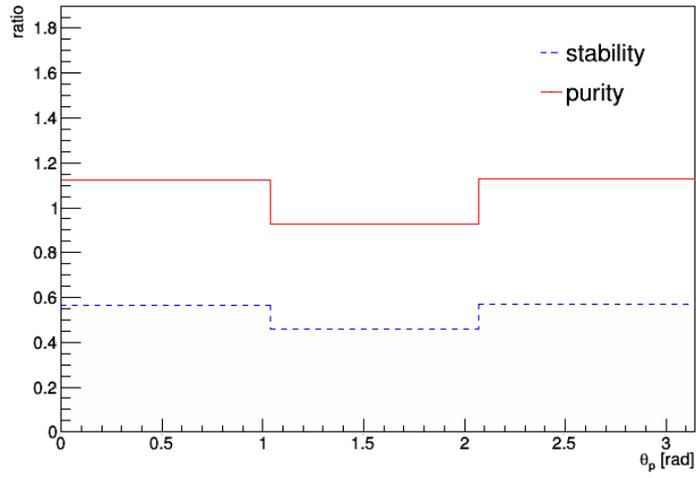


(c) The unfolded result.

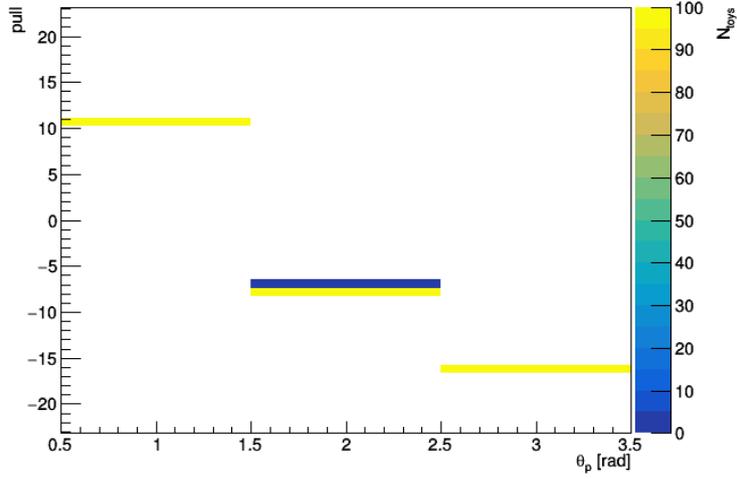


(d) The folded back result.

Figure 8.40. Unfolding plots for the  $t\bar{t}$  cflip method of the pull angle  $\theta_p$  of  $j_1^W$  to  $j_2^W$  including all jet constituents using 3 regularly sized bins (continued).

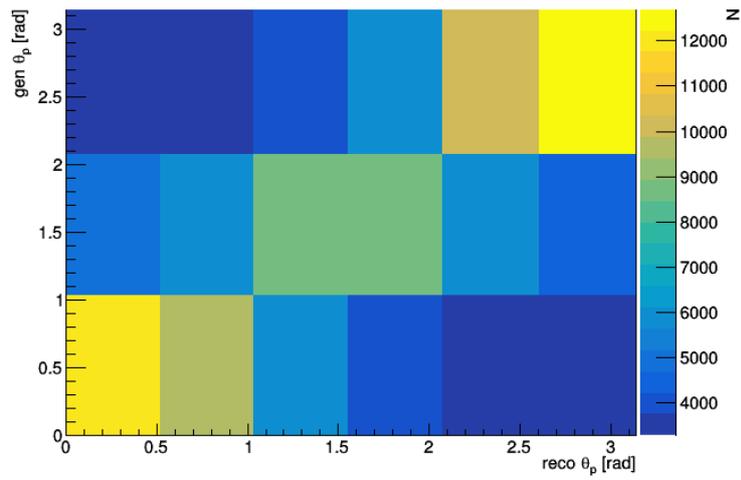


(e) Stability and purity in each bin.

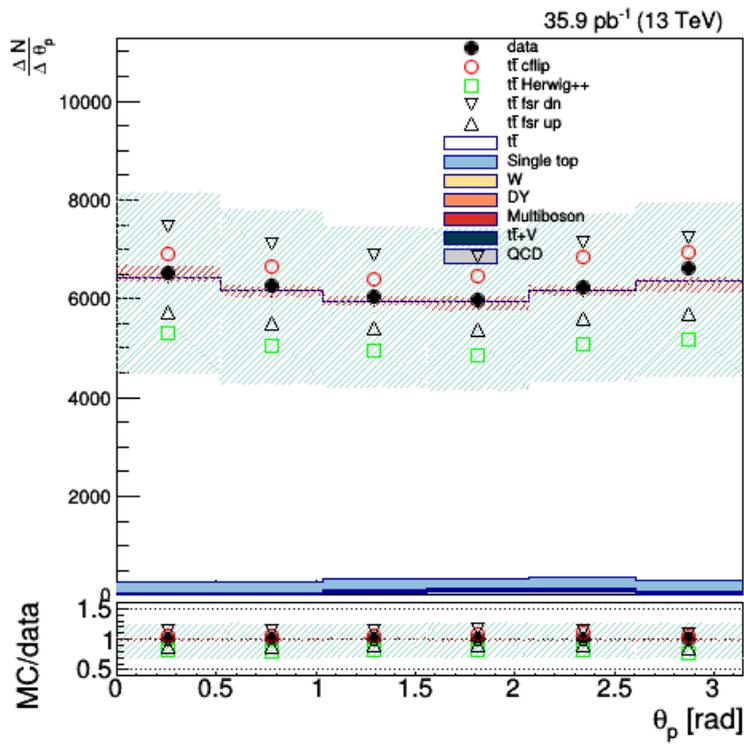


(f) The pull.

Figure 8.40. Unfolding plots for the  $t\bar{t}$  cflip method of the pull angle  $\theta_p$  of  $j_1^W$  to  $j_2^W$  including all jet constituents using 3 regularly sized bins (continued).

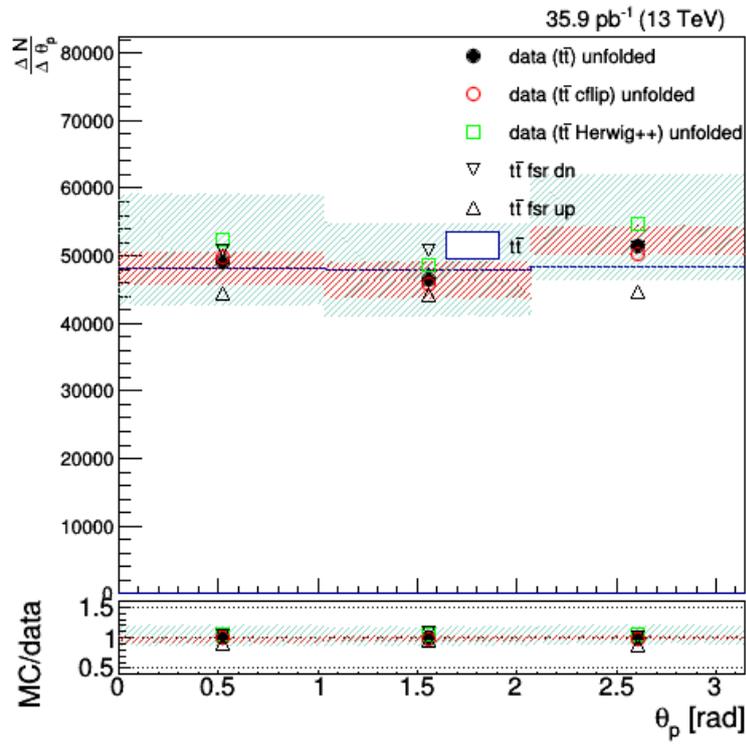


(a) Migration matrix with the reconstruction level as the  $x$  axis and the generation level as the  $y$  axis.

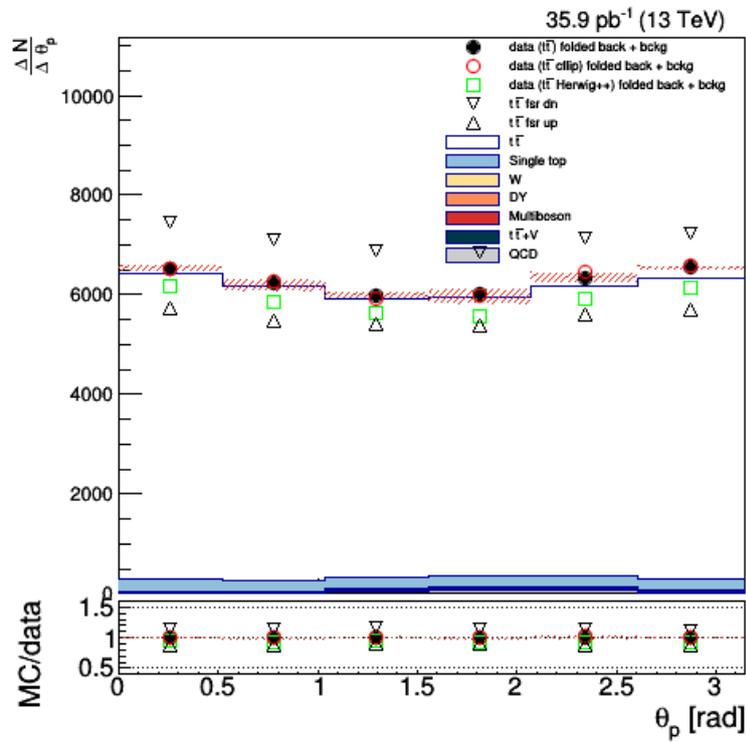


(b) The data and Monte Carlo used as input.

Figure 8.41. Unfolding plots for the  $t\bar{t}$  method of the pull angle  $\theta_p$  of  $j_1^b$  to  $j_2^b$  including all jet constituents using 3 regularly sized bins.

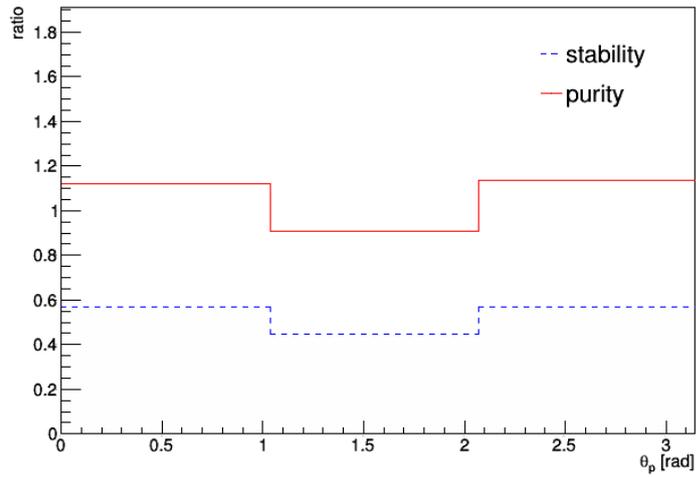


(c) The unfolded result.

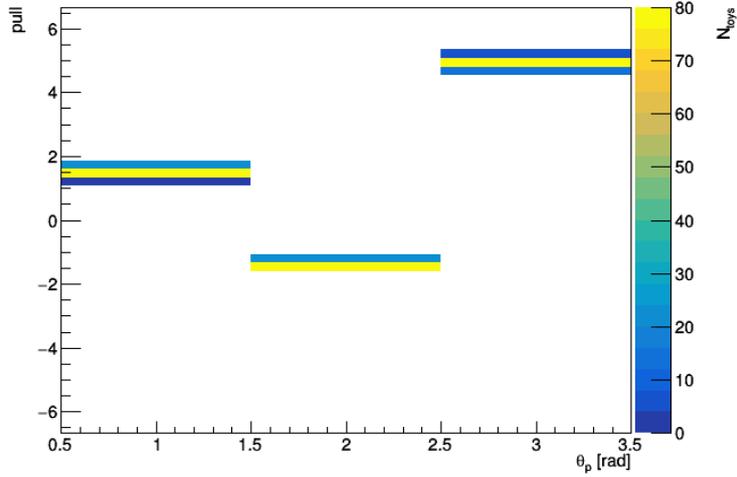


(d) The folded back result.

Figure 8.41. Unfolding plots for the  $t\bar{t}$  method of the pull angle  $\theta_p$  of  $j_1^b$  to  $j_2^b$  including all jet constituents using 3 regularly sized bins (continued).

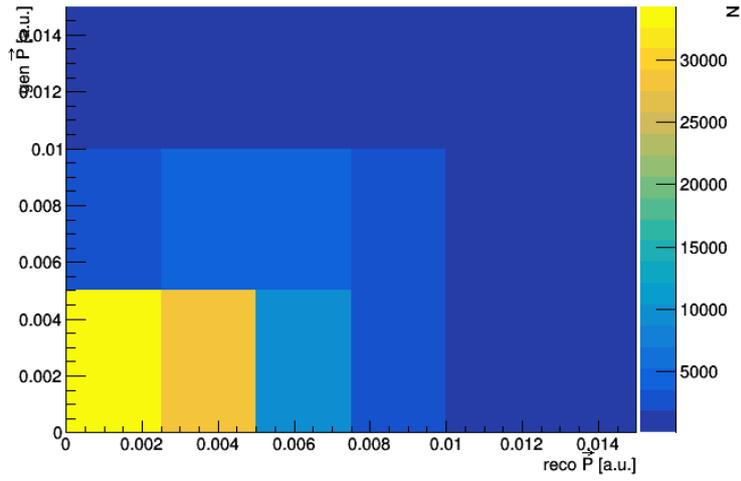


(e) Stability and purity in each bin.

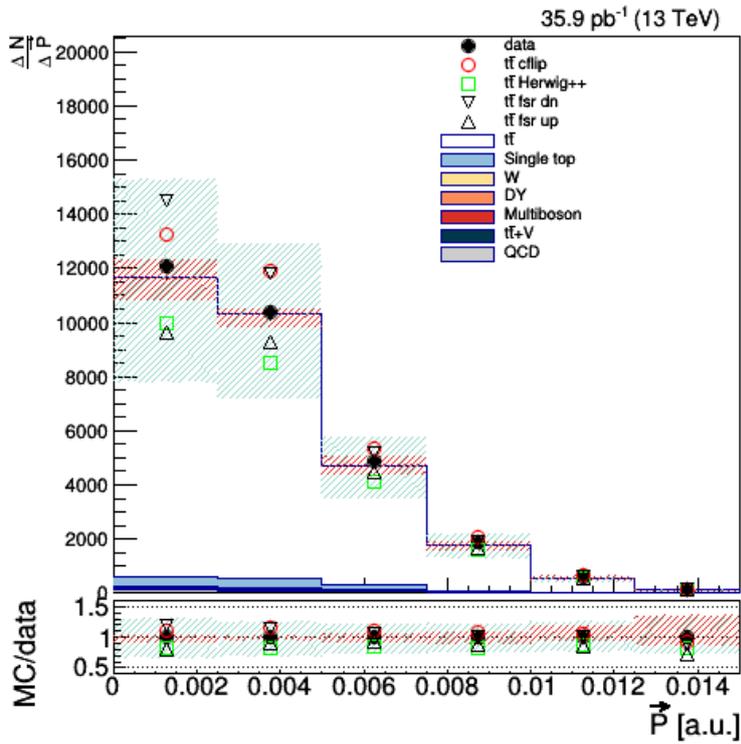


(f) The pull.

Figure 8.41. Unfolding plots for the  $t\bar{t}$  method of the pull angle  $\theta_p$  of  $j_1^b$  to  $j_2^b$  including all jet constituents using 3 regularly sized bins (continued).

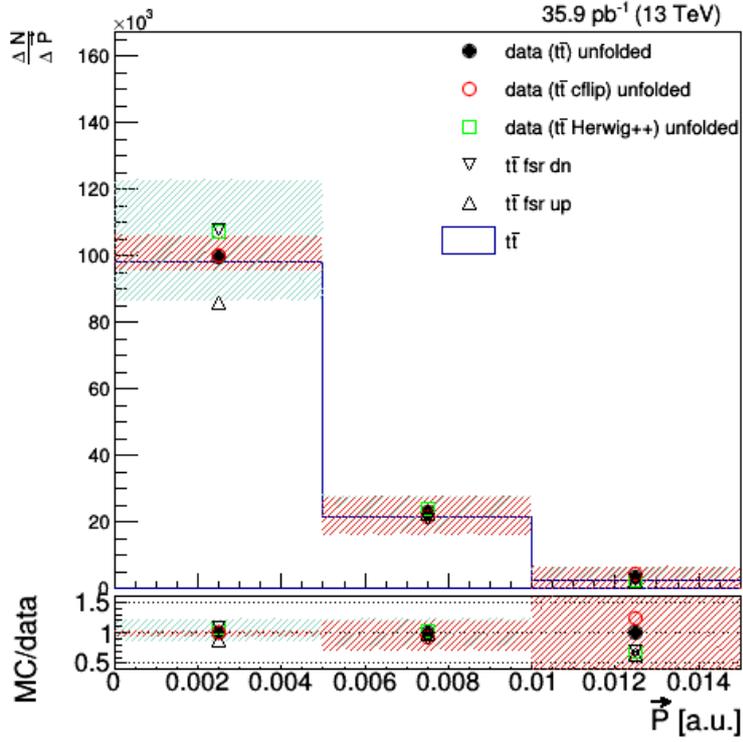


(a) Migration matrix with the reconstruction level as the  $x$  axis and the generation level as the  $y$  axis.

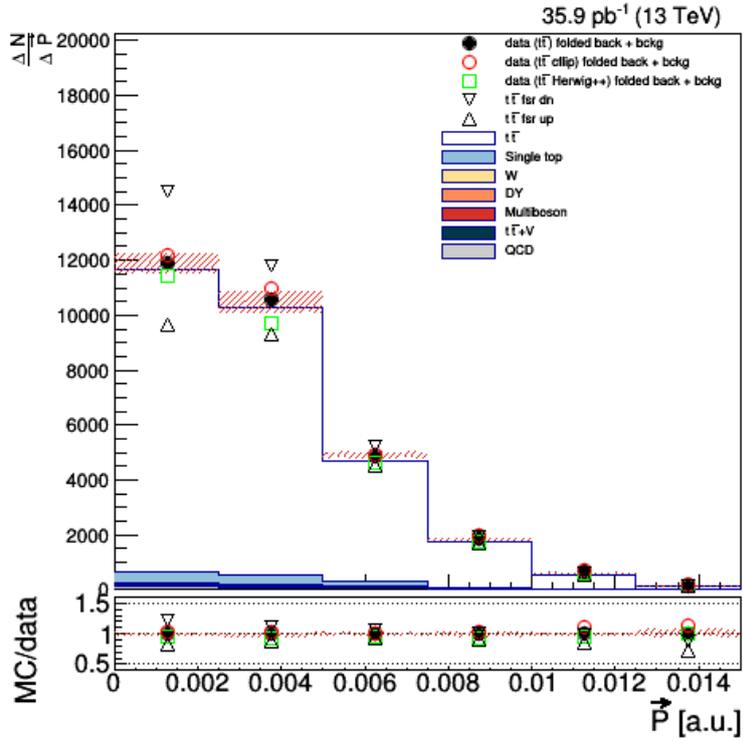


(b) The data and Monte Carlo used as input.

Figure 8.42. Unfolding plots for the  $t\bar{t}$  method of the magnitude of the pull vector  $|\vec{P}|$  of  $j_1^W$  to  $j_2^W$  including all jet constituents using 3 regularly sized bins.

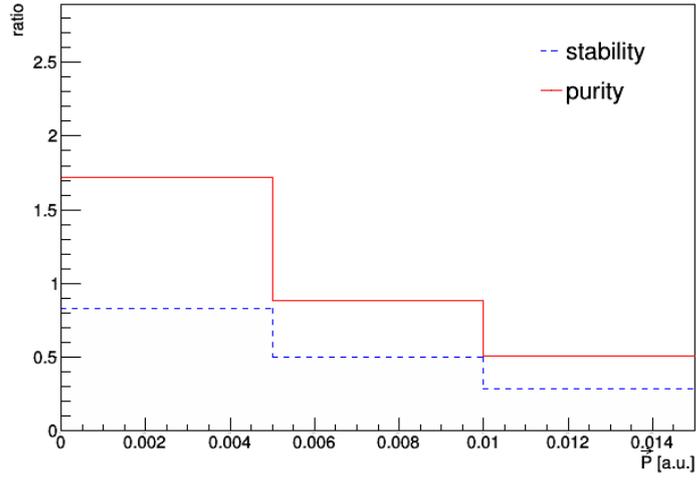


(c) The unfolded result.

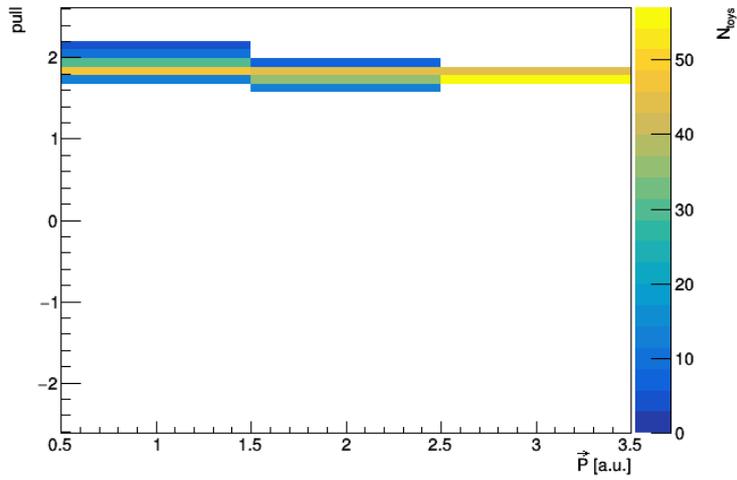


(d) The folded back result.

Figure 8.42. Unfolding plots for the  $t\bar{t}$  method of the magnitude of the pull vector  $|\vec{P}|$  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 8.42. Unfolding plots for the  $t\bar{t}$  method of the magnitude of the pull vector  $|\vec{P}|$  of  $j_1^W$  to  $j_2^W$  including all jet constituents using 3 regularly sized bins (continued).

Table 8.1

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

Nuisance	Uncertainty in bins [%]		
	bin 1 0.00–1.04	bin 2 1.04–2.07	bin 3 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 8.1

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 8.2

Bin by bin weight of the additional uncertainty (shape)  $t\bar{t}$  *cflip* for pull angle  $\theta_p$  including all jet constituents from the  $j_1^W$  to the  $j_2^W$  at the generator level. The results are for unfolded output 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 8.3

$\chi^2$  and  $p$ -values of pull angle  $\theta_p$  including all jet constituents. The results are for the SM model.

Sample	$\theta_p(j_1^W, j_2^W)$			$\theta_p(j_2^W, j_1^W)$			$\theta_p(j_1^b, j_2^b)$			$\theta_p(j_2^b, j_1^b)$		
	$\chi^2$	/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 8.4  
 $\chi^2$  and  $p$ -values of pull angle  $\theta_p$  including all jet constituents. The results are for the colour octet  $W$  model.

Sample	$\theta_p(j_1^W, j_2^W)$			$\theta_p(j_2^W, j_1^W)$			$\theta_p(j_1^b, j_2^b)$			$\theta_p(j_2^b, j_1^b)$		
	$\chi^2$	/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

Table 8.5  
 $\chi^2$  and  $p$ -values of magnitude of the pull vector  $|\vec{P}|$  including all jet constituents. The results are for the SM model.

Sample	$ \vec{P}(j_1^W) $			$ \vec{P}(j_2^W) $			$ \vec{P}(j_1^b) $			$ \vec{P}(j_2^b) $		
	$\chi^2$	/Ndf	$p$ -value									
Powheg + Pythia8	1.48	2	0.476	1.23	2	0.541	15.22	2	0.000	6.83	2	0.033
Powheg + Herwig++	0.11	2	0.948	2.10	2	0.349	0.18	2	0.913	6.02	2	0.049

Table 8.6  
 $\chi^2$  and  $p$ -values of magnitude of the pull vector  $|\vec{P}|$  including all jet constituents. The results are for the colour octet  $W$  model.

Sample	$ \vec{P}(j_1^W) $			$ \vec{P}(j_2^W) $			$ \vec{P}(j_1^b) $			$ \vec{P}(j_2^b) $		
	$\chi^2$	/Ndf	$p$ -value									
Powheg + Pythia8 *	1.37	2	0.503	1.19	2	0.552	8.44	2	0.015	2.21	2	0.331
Powheg + Pythia8 cf	1.13	2	0.567	1.80	2	0.407	0.39	2	0.823	0.94	2	0.624

Table 8.7  
 $\chi^2$  and  $p$ -values of pull angle  $\theta_p$  including all jet constituents. The nuisances are substituted in place of the signal. The results are for the SM model.

Sample	$\theta_p(j_1^W, j_2^W)$			$\theta_p(j_2^W, j_1^W)$			$\theta_p(j_1^b, j_2^b)$			$\theta_p(j_2^b, j_1^b)$		
	$\chi^2$	/Ndf	$p$ -value									
$t\bar{t}$	86.38	2	0.000	85.33	2	0.000	3.98	2	0.136	20.03	2	0.000
$t\bar{t}$ ERDon	44.89	2	0.000	50.05	2	0.000	4.06	2	0.131	20.01	2	0.000
$t\bar{t}$ Herwig++	0.81	2	0.667	4.94	2	0.085	5.15	2	0.076	24.92	2	0.000
$t\bar{t}$ Peterson Frag	83.13	2	0.000	82.73	2	0.000	3.94	2	0.139	19.93	2	0.000
$t\bar{t}$ QCDbased	12.99	2	0.002	16.40	2	0.000	4.86	2	0.088	22.98	2	0.000
$t\bar{t}$ UEdn	83.80	2	0.000	81.50	2	0.000	3.43	2	0.180	17.90	2	0.000
$t\bar{t}$ UEup	86.34	2	0.000	79.09	2	0.000	4.11	2	0.128	19.81	2	0.000
$t\bar{t}$ b fragmentation down	87.96	2	0.000	86.71	2	0.000	4.04	2	0.132	20.35	2	0.000
$t\bar{t}$ b fragmentation up	82.38	2	0.000	82.17	2	0.000	3.89	2	0.143	19.55	2	0.000
$t\bar{t}$ btag_heavy down	87.31	2	0.000	85.36	2	0.000	3.99	2	0.136	20.06	2	0.000
$t\bar{t}$ btag_heavy up	87.18	2	0.000	84.86	2	0.000	3.96	2	0.138	20.02	2	0.000
$t\bar{t}$ btag_light down	86.71	2	0.000	84.87	2	0.000	4.02	2	0.134	20.13	2	0.000
$t\bar{t}$ btag_light up	87.33	2	0.000	85.01	2	0.000	4.03	2	0.133	20.16	2	0.000
$t\bar{t}$ csv_heavy down	86.98	2	0.000	85.22	2	0.000	4.04	2	0.132	20.14	2	0.000
$t\bar{t}$ csv_heavy up	87.76	2	0.000	85.24	2	0.000	4.03	2	0.134	19.97	2	0.000
$t\bar{t}$ csv_light down	86.92	2	0.000	85.29	2	0.000	4.00	2	0.135	20.26	2	0.000
$t\bar{t}$ csv_light up	87.16	2	0.000	85.45	2	0.000	4.00	2	0.135	20.05	2	0.000
$t\bar{t}$ evtgen	88.64	2	0.000	84.17	2	0.000	3.72	2	0.156	19.89	2	0.000
$t\bar{t}$ fsr dn	94.49	2	0.000	93.76	2	0.000	4.15	2	0.125	20.13	2	0.000
$t\bar{t}$ fsr up	77.47	2	0.000	71.19	2	0.000	4.10	2	0.129	19.95	2	0.000
$t\bar{t}$ gluon move	90.92	2	0.000	87.97	2	0.000	4.17	2	0.124	21.44	2	0.000
$t\bar{t}$ hdamp dn	88.43	2	0.000	87.02	2	0.000	4.27	2	0.118	19.54	2	0.000
$t\bar{t}$ hdamp up	82.08	2	0.000	81.53	2	0.000	4.10	2	0.129	20.54	2	0.000
$t\bar{t}$ id1002muR1muF2hdampmt272.7225	85.85	2	0.000	85.11	2	0.000	3.98	2	0.137	20.05	2	0.000

$t\bar{t}$ id1003muR1muF0.5hdampmt272.7225	86.82	2	0.000	86.37	2	0.000	3.91	2	0.141	19.98	2	0.000
$t\bar{t}$ id1004muR2muF1hdampmt272.7225	86.80	2	0.000	86.29	2	0.000	3.93	2	0.140	19.99	2	0.000
$t\bar{t}$ id1005muR2muF2hdampmt272.7225	86.63	2	0.000	86.37	2	0.000	3.90	2	0.143	19.98	2	0.000
$t\bar{t}$ id1007muR0.5muF1hdampmt272.7225	85.02	2	0.000	84.06	2	0.000	4.04	2	0.133	20.27	2	0.000
$t\bar{t}$ id1009muR0.5muF0.5hdampmt272.7225	85.95	2	0.000	84.95	2	0.000	3.98	2	0.137	20.05	2	0.000
$t\bar{t}$ isr dn	91.23	2	0.000	87.69	2	0.000	4.28	2	0.118	20.23	2	0.000
$t\bar{t}$ isr up	77.17	2	0.000	78.83	2	0.000	4.23	2	0.121	20.10	2	0.000
$t\bar{t}$ jec_CorrelationGroupMPFInSitu down	87.33	2	0.000	85.15	2	0.000	4.00	2	0.135	20.21	2	0.000
$t\bar{t}$ jec_CorrelationGroupMPFInSitu up	86.74	2	0.000	85.18	2	0.000	4.04	2	0.133	20.26	2	0.000
$t\bar{t}$ jec_CorrelationGroupUncorrelated down	86.93	2	0.000	85.43	2	0.000	4.06	2	0.132	20.06	2	0.000
$t\bar{t}$ jec_CorrelationGroupUncorrelated up	87.28	2	0.000	85.19	2	0.000	4.01	2	0.135	20.21	2	0.000
$t\bar{t}$ jec_FlavorPureBottom down	87.28	2	0.000	85.75	2	0.000	4.05	2	0.132	20.27	2	0.000
$t\bar{t}$ jec_FlavorPureBottom up	87.18	2	0.000	85.19	2	0.000	4.08	2	0.130	20.20	2	0.000
$t\bar{t}$ jec_FlavorPureCharm down	87.13	2	0.000	84.73	2	0.000	4.02	2	0.134	20.23	2	0.000
$t\bar{t}$ jec_FlavorPureCharm up	86.90	2	0.000	85.00	2	0.000	4.05	2	0.132	19.99	2	0.000
$t\bar{t}$ jec_FlavorPureGluon down	86.96	2	0.000	84.86	2	0.000	4.02	2	0.134	20.24	2	0.000
$t\bar{t}$ jec_FlavorPureGluon up	87.15	2	0.000	85.25	2	0.000	4.02	2	0.134	20.16	2	0.000
$t\bar{t}$ jec_FlavorPureQuark down	87.22	2	0.000	85.35	2	0.000	4.00	2	0.135	20.13	2	0.000
$t\bar{t}$ jec_FlavorPureQuark up	87.12	2	0.000	85.45	2	0.000	4.00	2	0.135	20.11	2	0.000
$t\bar{t}$ jec_RelativeFSR down	87.28	2	0.000	85.20	2	0.000	4.02	2	0.134	20.05	2	0.000
$t\bar{t}$ jec_RelativeFSR up	86.98	2	0.000	85.26	2	0.000	4.05	2	0.132	20.03	2	0.000
$t\bar{t}$ jer down	86.97	2	0.000	85.13	2	0.000	4.01	2	0.135	20.10	2	0.000
$t\bar{t}$ jer up	86.92	2	0.000	85.25	2	0.000	4.02	2	0.134	20.15	2	0.000
$t\bar{t}$ m=171.5	82.62	2	0.000	83.40	2	0.000	3.98	2	0.136	20.51	2	0.000
$t\bar{t}$ m=173.5	84.34	2	0.000	80.90	2	0.000	4.66	2	0.097	18.29	2	0.000
$t\bar{t}$ pileup down	85.48	2	0.000	85.31	2	0.000	3.98	2	0.137	20.02	2	0.000
$t\bar{t}$ pileup up	86.20	2	0.000	85.47	2	0.000	3.96	2	0.138	19.85	2	0.000
$t\bar{t}$ sel efficiency correction down	86.32	2	0.000	85.30	2	0.000	3.98	2	0.136	20.03	2	0.000
$t\bar{t}$ sel efficiency correction up	86.40	2	0.000	85.35	2	0.000	3.98	2	0.137	20.03	2	0.000
$t\bar{t}$ semilep BR down	86.16	2	0.000	85.16	2	0.000	3.99	2	0.136	20.05	2	0.000
$t\bar{t}$ semilep BR up	86.12	2	0.000	85.15	2	0.000	3.98	2	0.136	20.05	2	0.000
$t\bar{t}$ tracking down	86.58	2	0.000	85.32	2	0.000	4.02	2	0.134	20.34	2	0.000
$t\bar{t}$ tracking up	87.45	2	0.000	85.27	2	0.000	4.01	2	0.135	20.19	2	0.000
$t\bar{t}$ trig efficiency correction down	86.33	2	0.000	85.30	2	0.000	3.98	2	0.136	20.04	2	0.000
$t\bar{t}$ trig efficiency correction up	86.43	2	0.000	85.37	2	0.000	3.98	2	0.137	20.02	2	0.000

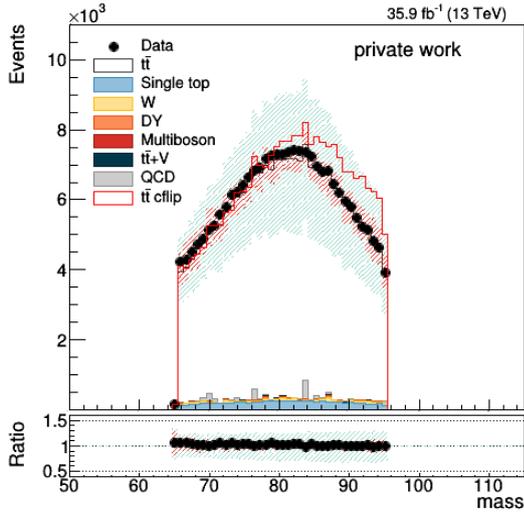
## 8.6 LEP method

In order to correctly apply the LEP methodology one needs to separate the  $b$  quarks on the hadronic and on the leptonic branch. The methodology to achieve this goal was described in Chap. 4. As a test of validity of this methodology, one can use the invariant mass of the  $t$  quark, formed by the sum of the  $b$  quark and the  $W$  boson. Fig. 8.43 shows the resonance of the  $t$  quark. Tables 8.8–8.10 provide measurements of the masses of the  $t$  quark and the  $W$  boson on the hadronic and leptonic branches using a polynomial fit. The mass of the  $W$  boson on the leptonic branch is brought at  $m_W$  in the process of calculation of the 4-vector of the neutrino.

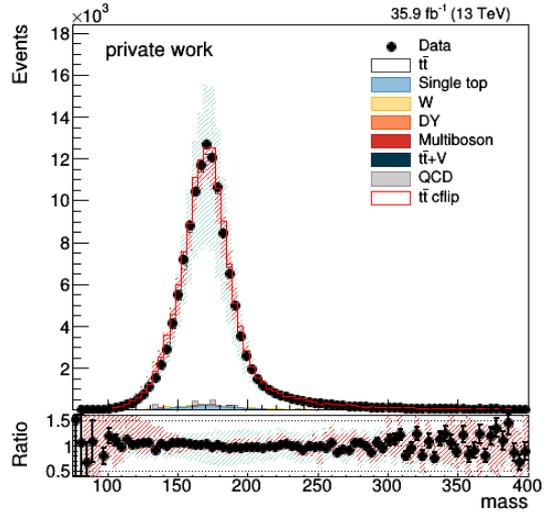
Table 8.8

Observed masses of objects for the common lepton channel at reconstruction level for MC for the SM model.

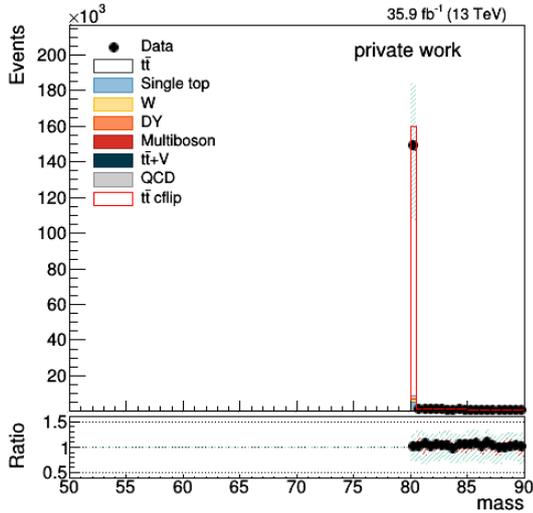
Object	Mass $\pm$ (stat) $\pm$ (syst) [GeV]	FWHM $\pm$ (stat) $\pm$ (syst) [GeV]
hadronic $W$	$8.162\text{e}+01 \pm 3.235\text{e}-06 \pm 2.067\text{e}-01$	$2.914\text{e}+01 \pm 4.089\text{e}-04 \pm 7.004\text{e}-02$
hadronic $t$	$1.726\text{e}+02 \pm 1.953\text{e}-09 \pm 3.756\text{e}-01$	$3.512\text{e}+01 \pm 6.212\text{e}-04 \pm 2.978\text{e}-01$
leptonic $W$	$8.704\text{e}+01 \pm 1.695\text{e}-04 \pm 1.542\text{e}-02$	$5.000\text{e}-01 \pm 0.000\text{e}+00 \pm 0.000\text{e}+00$
leptonic $t$	$1.730\text{e}+02 \pm 4.574\text{e}-07 \pm 9.835\text{e}-02$	$4.611\text{e}+01 \pm 1.168\text{e}-03 \pm 1.039\text{e}-01$



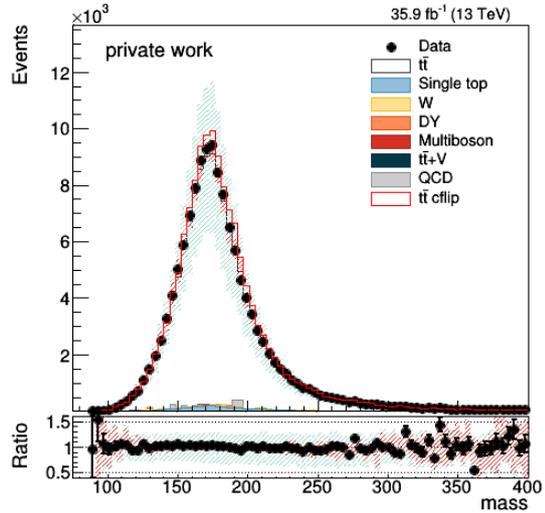
(a) Observed mass of  $W$  on the hadronic branch.



(b) Observed mass of  $t$  on the hadronic branch.



(c) Observed mass of  $W$  on the leptonic branch.



(d) Observed mass of  $t$  on the leptonic branch.

Figure 8.43. Observed masses of  $W$  and  $t$  on the hadronic and the leptonic branch.

Table 8.9  
Observed masses of objects for the common lepton channel at reconstruction level for data for the SM model.

Object	Mass $\pm$ (stat) [GeV]	FWHM $\pm$ (stat) [GeV]
hadronic $W$	$8.141\text{e}+01 \pm 2.366\text{e}-07$	$2.867\text{e}+01 \pm 4.307\text{e}-04$
hadronic $t$	$1.709\text{e}+02 \pm 2.751\text{e}-07$	$3.480\text{e}+01 \pm 5.541\text{e}-04$
leptonic $W$	$8.563\text{e}+01 \pm 1.053\text{e}-05$	$5.000\text{e}-01 \pm 0.000\text{e}+00$
leptonic $t$	$1.730\text{e}+02 \pm 3.380\text{e}-07$	$4.563\text{e}+01 \pm 1.081\text{e}-03$

Table 8.10

Observed masses of objects for the common lepton channel at reconstruction level for MC for the colour octet  $W$  model.

Object	Mass $\pm$ (stat) $\pm$ (syst) [GeV]	FWHM $\pm$ (stat) $\pm$ (syst) [GeV]
hadronic $W$	$8.347\text{e}+01 \pm 1.629\text{e}-06 \pm 1.825\text{e}-02$	$2.783\text{e}+01 \pm 1.977\text{e}-04 \pm 8.414\text{e}-02$
hadronic $t$	$1.724\text{e}+02 \pm 1.469\text{e}-04 \pm 4.000\text{e}-02$	$3.711\text{e}+01 \pm 6.257\text{e}-04 \pm 3.821\text{e}-01$
leptonic $W$	$8.581\text{e}+01 \pm 1.862\text{e}-05 \pm 8.210\text{e}-03$	$5.000\text{e}-01 \pm 0.000\text{e}+00 \pm 0.000\text{e}+00$
leptonic $t$	$1.730\text{e}+02 \pm 1.094\text{e}-06 \pm 3.321\text{e}-03$	$4.685\text{e}+01 \pm 1.247\text{e}-03 \pm 4.481\text{e}-02$

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  $p_T$  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. 8.44 with all jet constituents and in Fig. 8.45 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  $j_h^b$  and the furthest light quark  $j_f^W$  (jet distance is measured with the angle between the spatial components of the 4-vectors of the jets), the closest light quark  $j_c^W$  and the hadronic  $b$  jet  $j_h^b$ , and the leading light jet  $j_1^W$  and the second leading light jet  $j_2^W$ .

The results using energy flow are shown in Fig. 8.46 with all jet constituents and in Fig. 8.47 including only charged jet constituents.

The results using  $p_T$  flow are shown in Fig. 8.48 with all jet constituents and in Fig. 8.49 including only charged jet constituents.

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  $(j_1^b, j_2^b)$ ,  $(j_h^b, j_f^W)$ ,  $(j_c^W, j_h^b)$  to the flow in the colour-connected region  $(j_1^W, j_2^W)$  are given in Fig. 8.50 including all jet constituents and Fig. 8.51 including only charged jet constituents. Significant colour reconnection is noticeable in the region  $(j_c^W, j_h^b)$  assuming the colour octet  $W$  model.

Fig. 8.52 shows the bin-per-bin ratio of the particle flow in the region  $(j_1^W, j_2^W)$  in the colour octet  $W$  model to the particle flow in the region  $(j_1^W, j_2^W)$  in the SM 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:

$$R = \frac{\int_{0.2}^{0.8} f^{\text{inter } W \text{ region}} d\chi}{\int_{0.2}^{0.8} f^{\text{intra } W \text{ region}} d\chi}, \quad (8.7)$$

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  $625 \text{ pb}^{-1}$  of data in the range  $\sqrt{s} = 189 - 209 \text{ GeV}$  are given in Table 8.11. 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 8.11

$R$  values observed at LEP.

LEP experiment	$R$ value – data	Reference
OPAL	1.243	[7]
Delphi	0.889 ( $\sqrt{s} = 183 \text{ GeV}$ )–1.039 ( $\sqrt{s} = 207 \text{ GeV}$ )	[8]
L3	0.911	[9]

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 in Table 8.12, for data in Table 8.13 and for the colour octet  $W$  model in Table 8.14.

The integral of energy flow from 0.2 to 0.8 in different regions and the inverse of  $R$  values for the SM model is given in Table 8.15, for data in Table 8.16 and for the colour octet  $W$  model in Table 8.17.

The integral of  $p_T$  flow from 0.2 to 0.8 in different regions and the inverse of  $R$  values for the SM model is given in Table 8.18, for data in Table 8.19 and for the colour octet  $W$  model in Table 8.20.

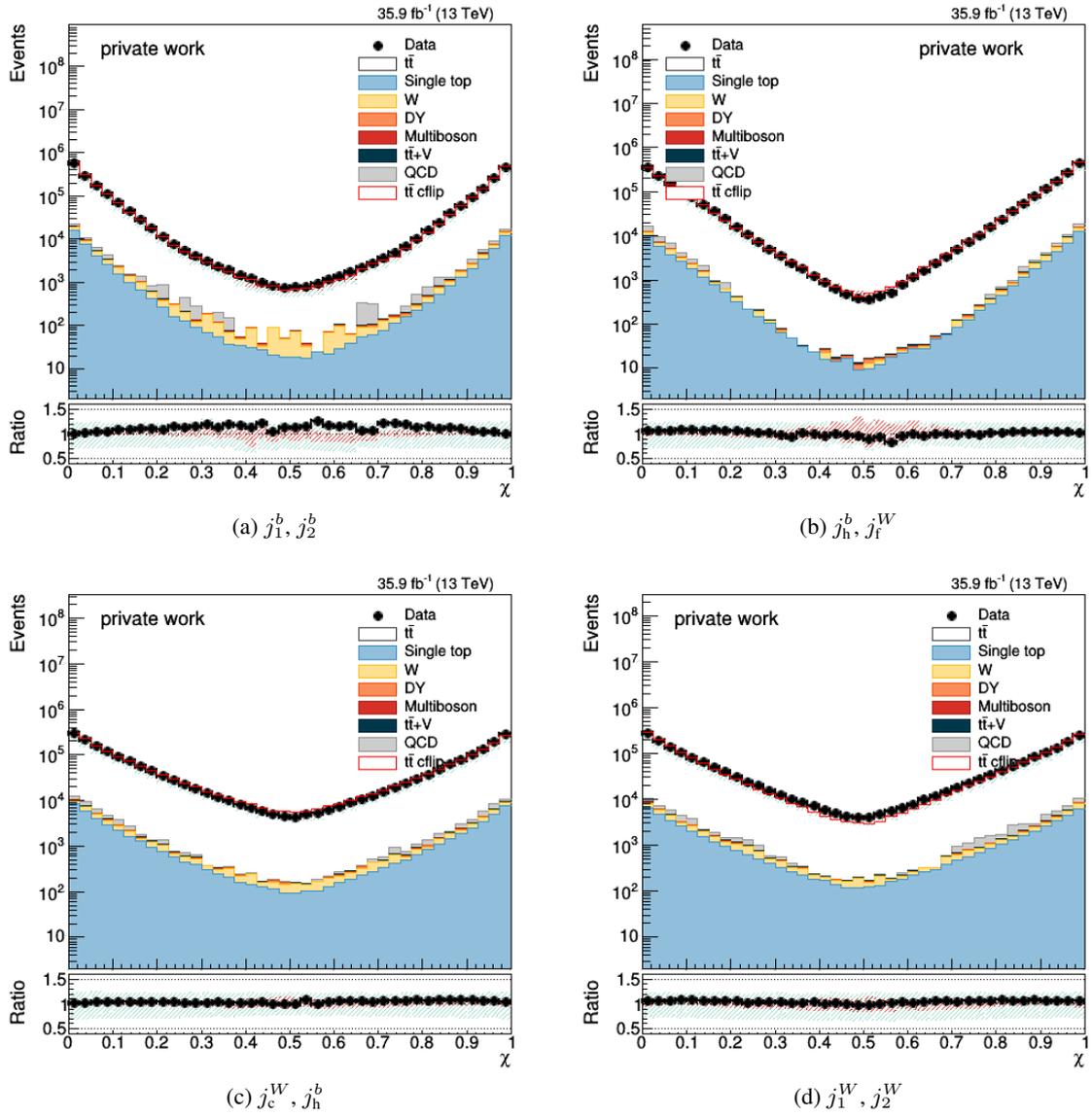


Figure 8.44. Plots of the particle flow including all jet constituents.

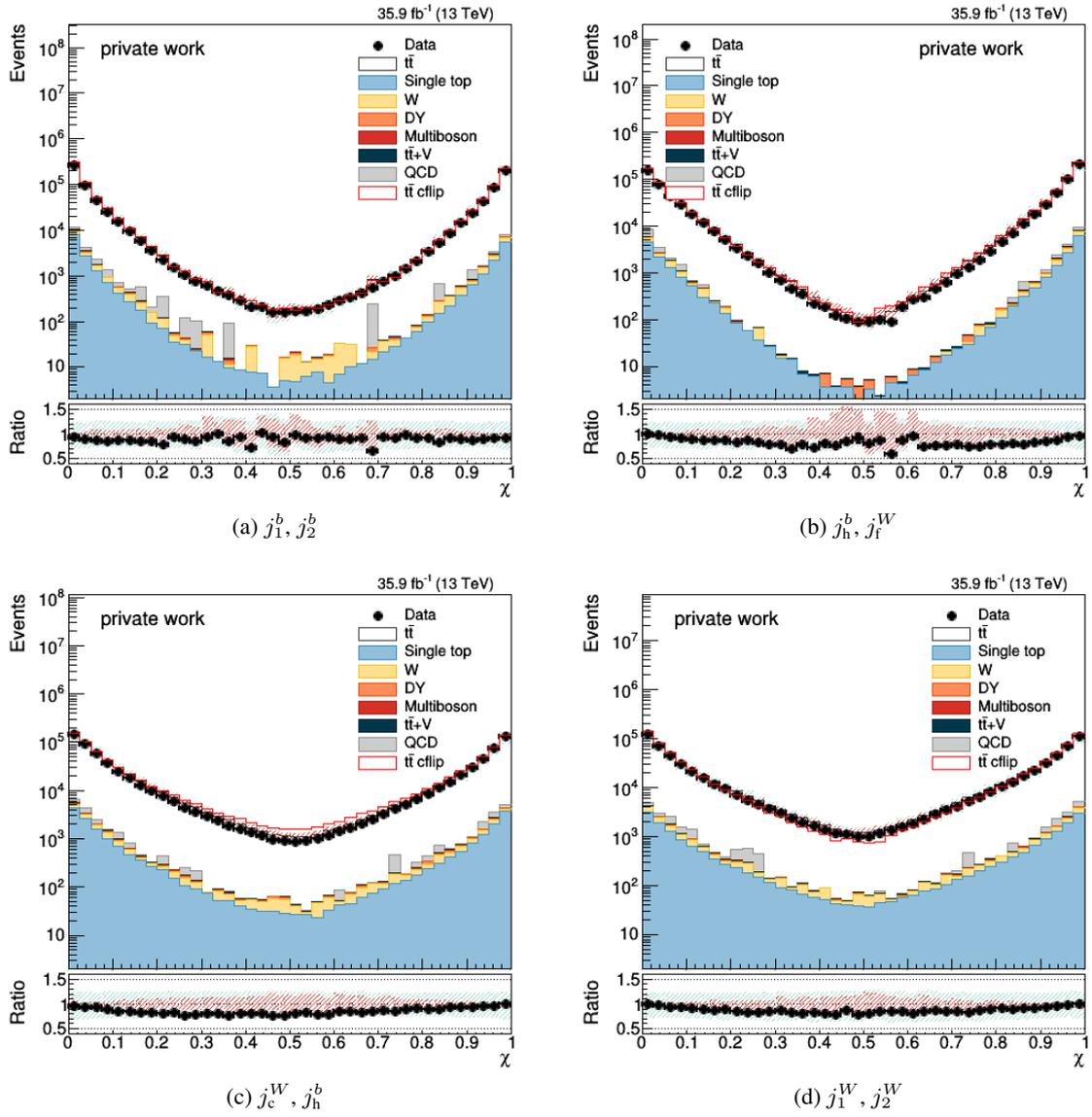


Figure 8.45. Plots of the particle flow including only charged jet constituents.

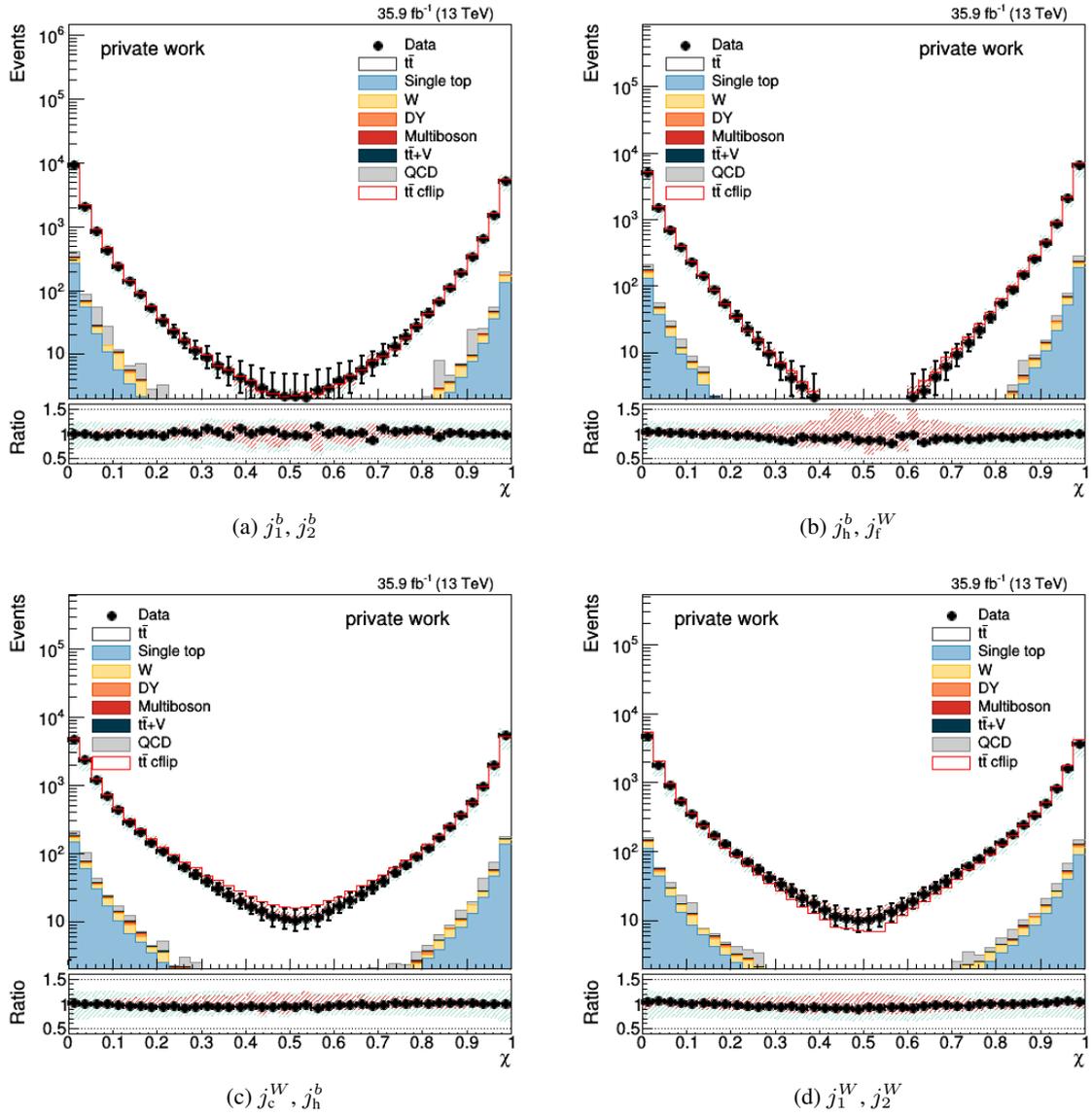


Figure 8.46. Plots of the energy flow including all jet constituents.

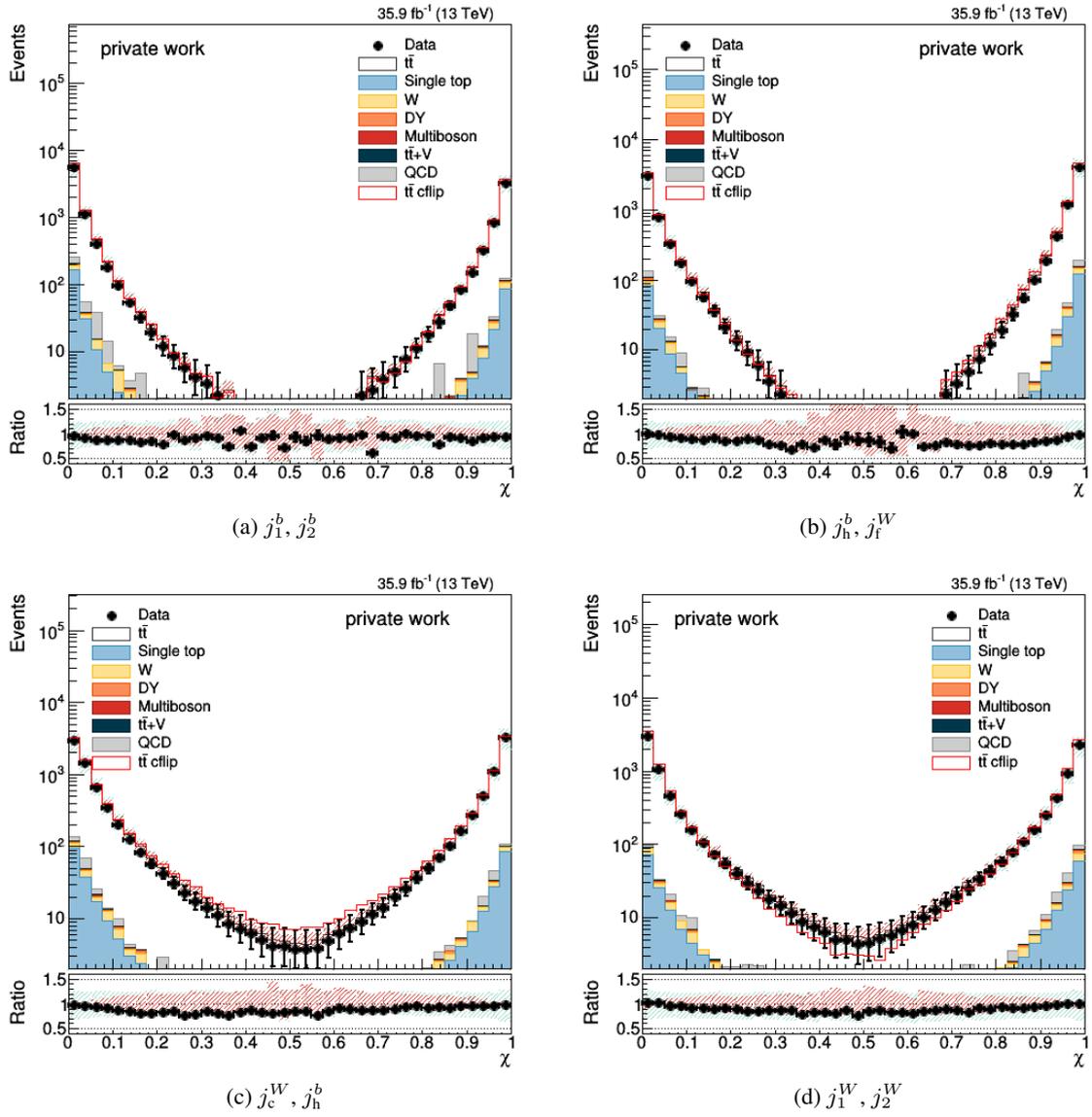
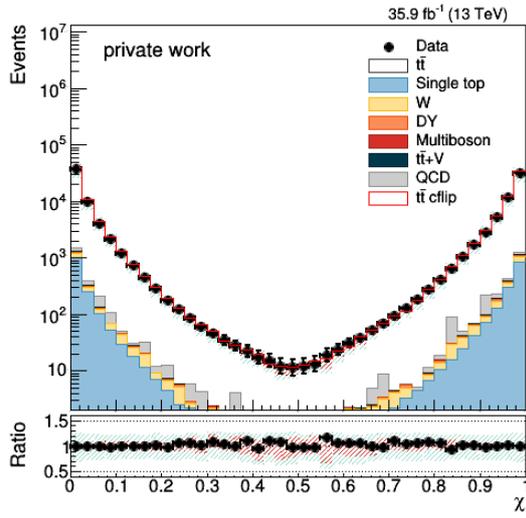
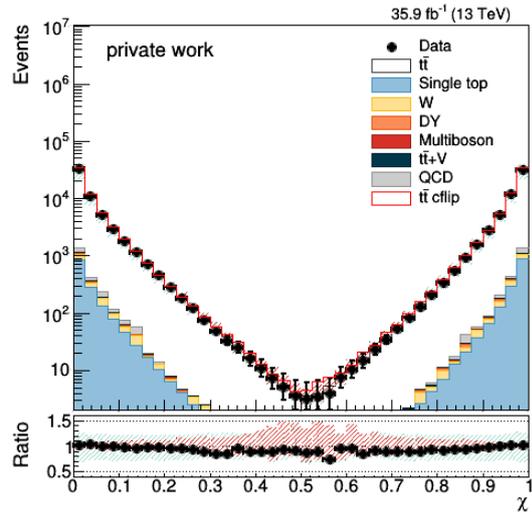


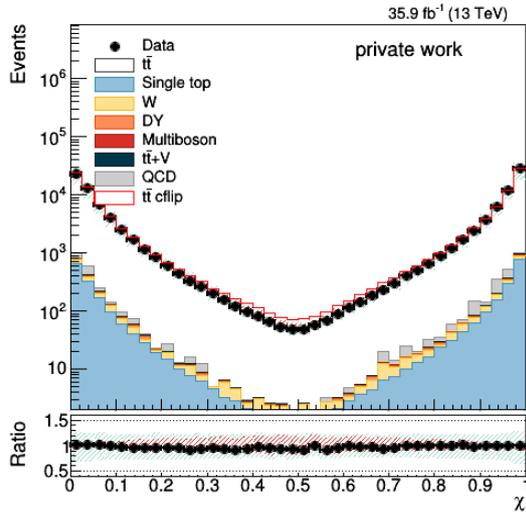
Figure 8.47. Plots of the energy flow including only charged jet constituents.



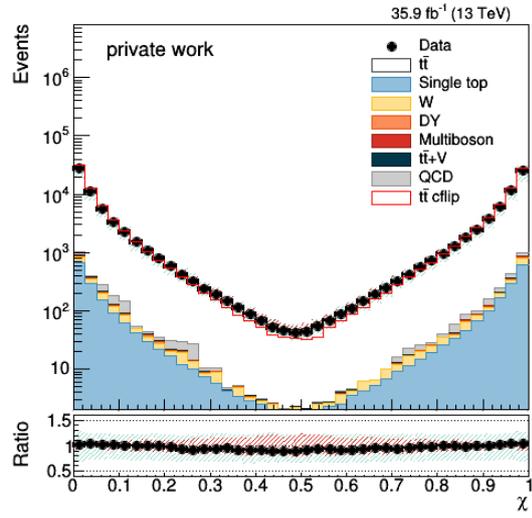
(a)  $j_1^b, j_2^b$



(b)  $j_h^b, j_f^W$



(c)  $j_c^W, j_h^b$



(d)  $j_1^W, j_2^W$

Figure 8.48. Plots of the  $p_T$  flow including all jet constituents.

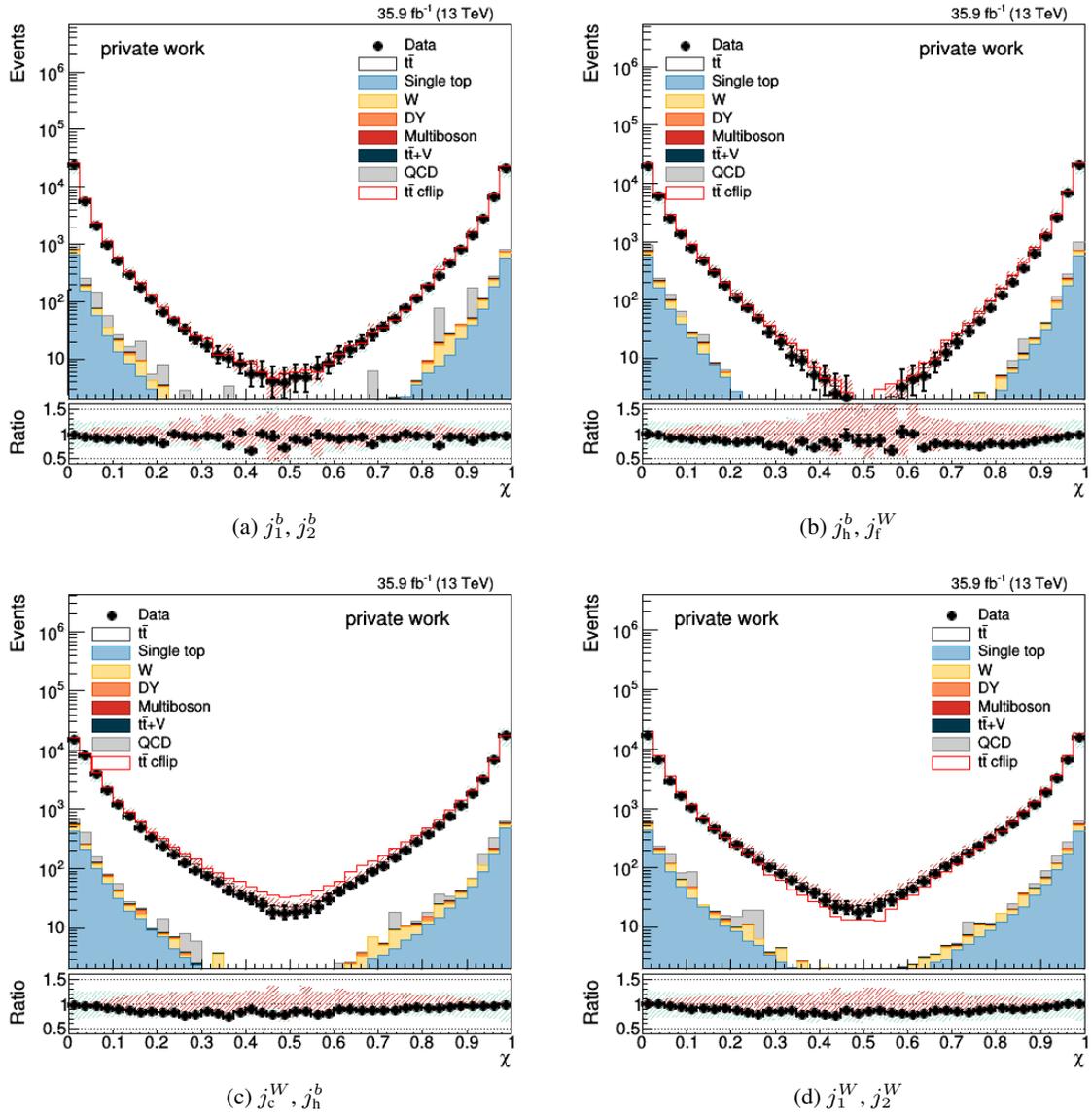


Figure 8.49. Plots of the  $p_T$  flow including only charged jet constituents.

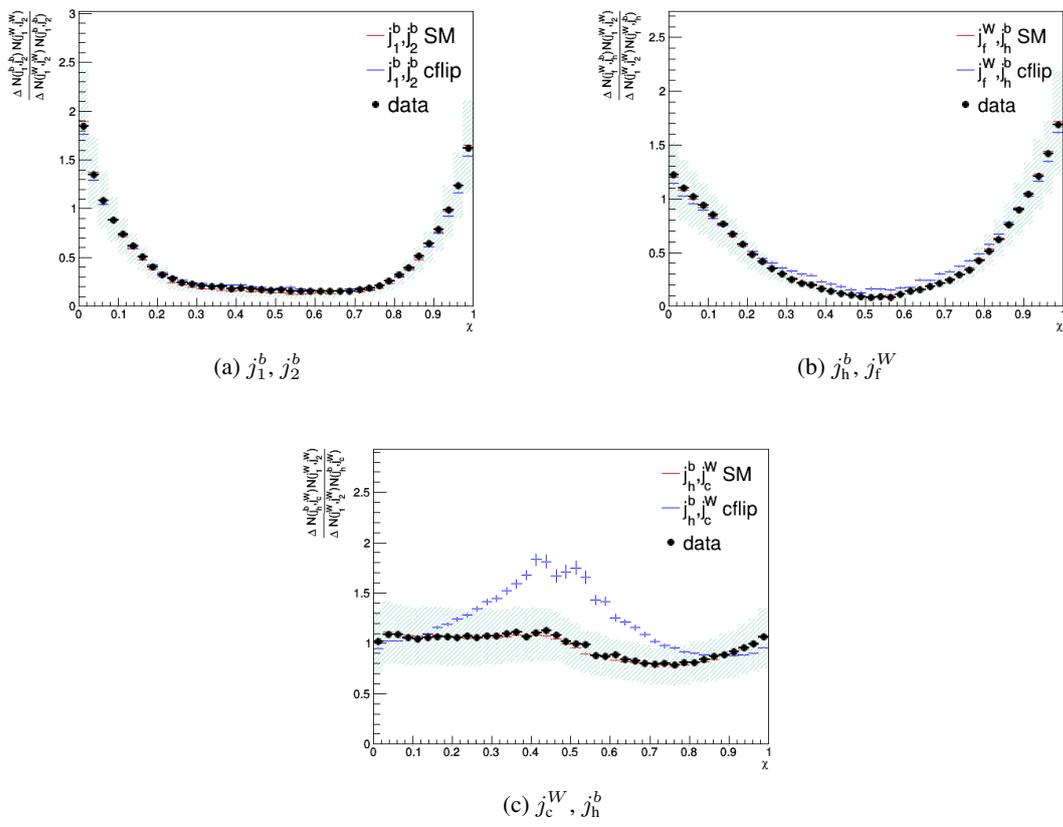


Figure 8.50. Graphs showing the particle flow including all jet constituents normalised to the flow at  $j_1^W, j_2^W$ .

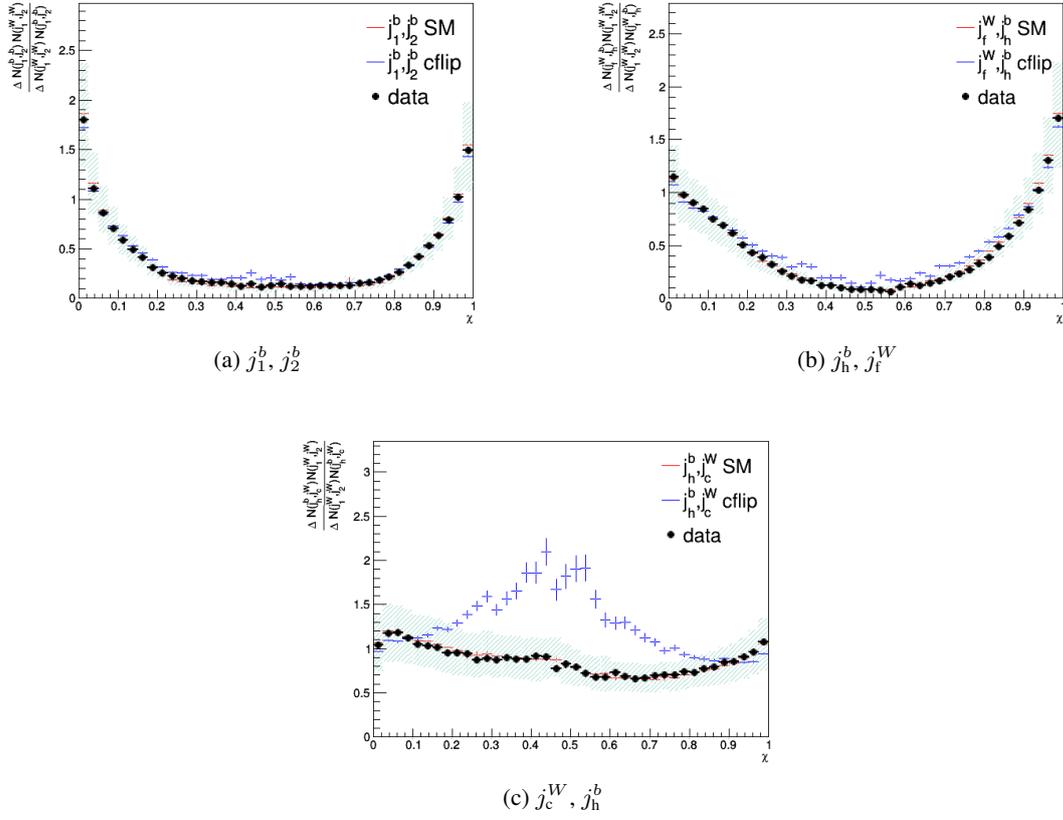


Figure 8.51. Graphs showing the particle flow including only charged jet constituents normalised to the flow at  $j_1^W, j_2^W$ .

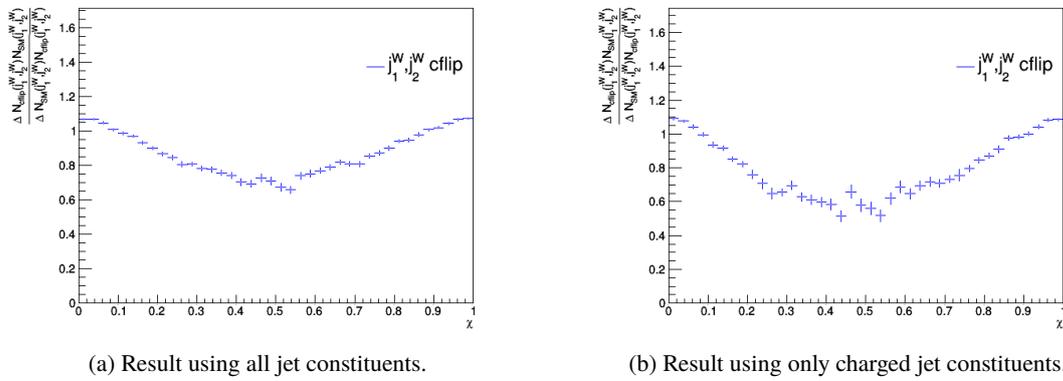


Figure 8.52. Bin-per-bin ratio of particle flow in region  $(j_1^W, j_2^W)$  in the  $W$  colour octet model to particle flow in region  $(j_1^W, j_2^W)$  in the SM model.

Table 8.12

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	$I \pm (\text{stat}) \pm (\text{syst})$ [rad]	$R^{-1} \pm (\text{stat}) \pm (\text{syst})$
$j_1^b, j_2^b$		
all	$0.010 \pm 0.000 \pm 0.000$	$0.151 \pm 0.000 \pm 0.002$
charged	$0.007 \pm 0.000 \pm 0.000$	$0.127 \pm 0.000 \pm 0.003$
$j_f^W, j_h^b$		
all	$0.012 \pm 0.000 \pm 0.000$	$0.183 \pm 0.000 \pm 0.001$
charged	$0.008 \pm 0.000 \pm 0.000$	$0.152 \pm 0.000 \pm 0.003$
$j_h^b, j_c^W$		
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 8.13

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	$I \pm (\text{stat})$ [rad]	$R^{-1} \pm (\text{stat})$
$j_1^b, j_2^b$		
all	$0.011 \pm 0.000$	$0.173 \pm 0.000$
charged	$0.006 \pm 0.000$	$0.139 \pm 0.000$
$j_f^W, j_h^b$		
all	$0.011 \pm 0.000$	$0.175 \pm 0.000$
charged	$0.007 \pm 0.000$	$0.144 \pm 0.000$
$j_h^b, j_c^W$		
all	$0.062 \pm 0.000$	$0.944 \pm 0.000$
charged	$0.035 \pm 0.000$	$0.766 \pm 0.000$
$j_1^W, j_2^W$		
all	$0.065 \pm 0.000$	$1.000 \pm 0.000$
charged	$0.046 \pm 0.000$	$1.000 \pm 0.000$

Table 8.14

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	$I \pm (\text{stat}) \pm (\text{syst})$ [rad]	$R^{-1} \pm (\text{stat}) \pm (\text{syst})$
$j_1^b, j_2^b$		
all	$0.010 \pm 0.000 \pm 0.000$	$0.189 \pm 0.000 \pm 0.002$
charged	$0.006 \pm 0.000 \pm 0.000$	$0.175 \pm 0.000 \pm 0.004$
$j_f^W, j_h^b$		
all	$0.013 \pm 0.000 \pm 0.000$	$0.250 \pm 0.000 \pm 0.001$
charged	$0.008 \pm 0.000 \pm 0.000$	$0.243 \pm 0.000 \pm 0.003$
$j_h^b, j_c^W$		
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	$0.034 \pm 0.000 \pm 0.000$	$1.000 \pm 0.000 \pm 0.008$

Table 8.15

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

Jet constituents	$I \pm (\text{stat}) \pm (\text{syst})$ [rad]	$R^{-1} \pm (\text{stat}) \pm (\text{syst})$
$j_1^b, j_2^b$		
all	$0.004 \pm 0.000 \pm 0.000$	$0.156 \pm 0.001 \pm 0.003$
charged	$0.003 \pm 0.000 \pm 0.000$	$0.139 \pm 0.002 \pm 0.004$
$j_f^W, j_h^b$		
all	$0.003 \pm 0.000 \pm 0.000$	$0.129 \pm 0.001 \pm 0.002$
charged	$0.002 \pm 0.000 \pm 0.000$	$0.110 \pm 0.002 \pm 0.003$
$j_h^b, j_c^W$		
all	$0.018 \pm 0.000 \pm 0.000$	$0.786 \pm 0.004 \pm 0.005$
charged	$0.012 \pm 0.000 \pm 0.000$	$0.664 \pm 0.006 \pm 0.008$
$j_1^W, j_2^W$		
all	$0.023 \pm 0.000 \pm 0.000$	$1.000 \pm 0.005 \pm 0.013$
charged	$0.018 \pm 0.000 \pm 0.000$	$1.000 \pm 0.010 \pm 0.018$

Table 8.16

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

Jet constituents	$I \pm (\text{stat})$ [rad]	$R^{-1} \pm (\text{stat})$
$j_1^b, j_2^b$		
all	$0.004 \pm 0.000$	$0.178 \pm 0.002$
charged	$0.002 \pm 0.000$	$0.150 \pm 0.003$
$j_f^W, j_h^b$		
all	$0.003 \pm 0.000$	$0.125 \pm 0.001$
charged	$0.002 \pm 0.000$	$0.106 \pm 0.002$
$j_h^b, j_c^W$		
all	$0.017 \pm 0.000$	$0.827 \pm 0.004$
charged	$0.011 \pm 0.000$	$0.685 \pm 0.007$
$j_1^W, j_2^W$		
all	$0.020 \pm 0.000$	$1.000 \pm 0.005$
charged	$0.015 \pm 0.000$	$1.000 \pm 0.011$

Table 8.17

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

Jet constituents	$I \pm (\text{stat}) \pm (\text{syst})$ [rad]	$R^{-1} \pm (\text{stat}) \pm (\text{syst})$
$j_1^b, j_2^b$		
all	$0.003 \pm 0.000 \pm 0.000$	$0.233 \pm 0.002 \pm 0.003$
charged	$0.002 \pm 0.000 \pm 0.000$	$0.224 \pm 0.004 \pm 0.007$
$j_f^W, j_h^b$		
all	$0.003 \pm 0.000 \pm 0.000$	$0.218 \pm 0.002 \pm 0.002$
charged	$0.002 \pm 0.000 \pm 0.000$	$0.221 \pm 0.004 \pm 0.004$
$j_h^b, j_c^W$		
all	$0.022 \pm 0.000 \pm 0.000$	$1.575 \pm 0.007 \pm 0.009$
charged	$0.016 \pm 0.000 \pm 0.000$	$1.669 \pm 0.014 \pm 0.015$
$j_1^W, j_2^W$		
all	$0.014 \pm 0.000 \pm 0.000$	$1.000 \pm 0.006 \pm 0.007$
charged	$0.010 \pm 0.000 \pm 0.000$	$1.000 \pm 0.012 \pm 0.011$

Table 8.18

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

Jet constituents	$I \pm (\text{stat}) \pm (\text{syst})$ [rad]	$R^{-1} \pm (\text{stat}) \pm (\text{syst})$
$j_1^b, j_2^b$		
all	$0.005 \pm 0.000 \pm 0.000$	$0.231 \pm 0.000 \pm 0.004$
charged	$0.003 \pm 0.000 \pm 0.000$	$0.203 \pm 0.001 \pm 0.005$
$j_f^W, j_h^b$		
all	$0.003 \pm 0.000 \pm 0.000$	$0.156 \pm 0.000 \pm 0.002$
charged	$0.002 \pm 0.000 \pm 0.000$	$0.134 \pm 0.000 \pm 0.003$
$j_h^b, j_c^W$		
all	$0.019 \pm 0.000 \pm 0.000$	$0.951 \pm 0.001 \pm 0.006$
charged	$0.013 \pm 0.000 \pm 0.000$	$0.814 \pm 0.001 \pm 0.009$
$j_1^W, j_2^W$		
all	$0.020 \pm 0.000 \pm 0.000$	$1.000 \pm 0.001 \pm 0.014$
charged	$0.016 \pm 0.000 \pm 0.000$	$1.000 \pm 0.002 \pm 0.019$

Table 8.19

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

Jet constituents	$I \pm (\text{stat})$ [rad]	$R^{-1} \pm (\text{stat})$
$j_1^b, j_2^b$		
all	$0.005 \pm 0.000$	$0.263 \pm 0.000$
charged	$0.003 \pm 0.000$	$0.221 \pm 0.001$
$j_f^W, j_h^b$		
all	$0.003 \pm 0.000$	$0.149 \pm 0.000$
charged	$0.002 \pm 0.000$	$0.127 \pm 0.000$
$j_h^b, j_c^W$		
all	$0.018 \pm 0.000$	$0.991 \pm 0.001$
charged	$0.011 \pm 0.000$	$0.835 \pm 0.002$
$j_1^W, j_2^W$		
all	$0.018 \pm 0.000$	$1.000 \pm 0.001$
charged	$0.013 \pm 0.000$	$1.000 \pm 0.002$

Table 8.20

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

<b>Jet constituents</b>	$I \pm (\text{stat}) \pm (\text{syst})$ [rad]	$R^{-1} \pm (\text{stat}) \pm (\text{syst})$
	$j_1^b, j_2^b$	
all	$0.004 \pm 0.000 \pm 0.000$	$0.332 \pm 0.000 \pm 0.005$
charged	$0.003 \pm 0.000 \pm 0.000$	$0.310 \pm 0.001 \pm 0.008$
	$j_f^W, j_h^b$	
all	$0.003 \pm 0.000 \pm 0.000$	$0.250 \pm 0.000 \pm 0.002$
charged	$0.002 \pm 0.000 \pm 0.000$	$0.251 \pm 0.001 \pm 0.004$
	$j_h^b, j_c^W$	
all	$0.023 \pm 0.000 \pm 0.000$	$1.790 \pm 0.001 \pm 0.010$
charged	$0.017 \pm 0.000 \pm 0.000$	$1.874 \pm 0.003 \pm 0.015$
	$j_1^W, j_2^W$	
all	$0.013 \pm 0.000 \pm 0.000$	$1.000 \pm 0.001 \pm 0.006$
charged	$0.009 \pm 0.000 \pm 0.000$	$1.000 \pm 0.002 \pm 0.011$

## 8.7 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 [124]. The first hypothesis is called the null hypothesis  $H_0$  while the latter one is called the alternative hypothesis  $H_{\text{alt}}$ .

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

$$n = \mu \left( (1 - x) f_{t\bar{t}} + x f_{t\bar{t}_{\text{flip}}} \right) + b, \quad (8.8)$$

where  $n$  is the expected number of events,  $\mu$  is the signal strength,  $x$  is 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  $\mu$  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:

$$q^{\text{TEV}} = -2 \ln \frac{L(H_0)}{L(H_{\text{alt}})} = -2 \ln \frac{L(\text{data}|p = 0, \hat{\theta}_0)}{L(\text{data}|p = P, \hat{\theta}_P)}, \quad (8.9)$$

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

$$L(\mu, \theta_s, \theta_b) = \prod_{i=1}^N \frac{(\mu s_i(\theta_s) + b_i(\theta_b))^{n_i}}{n_i!} e^{-(\mu s_i(\theta_s) + b_i(\theta_b))}, \quad (8.10)$$

where  $i$  is the phase space parameter (bin index) and  $n_i$  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 described in Eq. 8.8 and  $P$  is set to 1, it happens that when applying the  $q^{\text{TEV}}$  statistic  $H_0$  (with  $x = 0$ ) is defined as the  $t\bar{t} + b$  distribution while  $H_{\text{alt}}$  is defined as the  $t\bar{t}_{\text{flip}} + b$  distribution.

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

$$p \equiv \int_{q_{\text{obs}}}^{\infty} f(q) dq, \quad (8.11)$$

where  $q_{\text{obs}}$  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  $2.87 \times 10^{-7}$ . For the Neyman–Pearson test statistic the  $p$ -value for  $H_0$  is right-handed while the  $p$ -value for  $H_{\text{alt}}$  is left-handed. This is illustrated in Fig. 8.53.

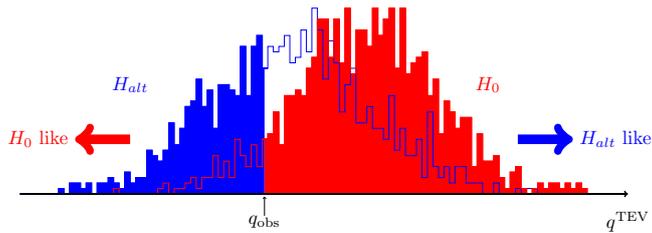


Figure 8.53. Evaluation of hypotheses according to the Neyman–Pearson test statistic.

For testing the hypothesis and doing all background work we use the CMS combine tool [125]. The datacard for creating the RooFit [126] workspace is given in Appendix A. 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. 8.8 to be equal to 1 in  $H_{\text{alt}}$ . We at the present stage use only 500 toys. The distribution of  $q/2$  where  $q$  is the test statistic under the assumption of  $H_0$ ,  $H_{\text{alt}}$  and  $q_{\text{obs}}/2$  is given in Fig. 8.54.

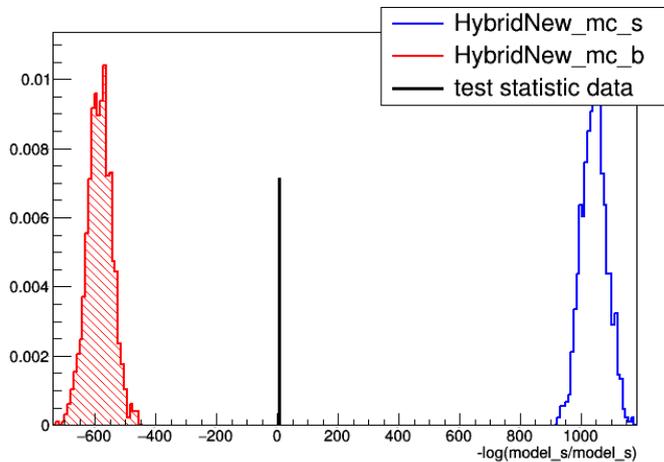


Figure 8.54. Distribution of the  $q/2$  under the assumption of  $t\bar{t}$  hypothesis (red), colour-flipped  $t\bar{t}$  hypothesis (blue) and  $q_{\text{obs}}/2$ .

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

The combine tool has a method MultiDimFit to determine the curve of the profile likelihood ratio PLR, defined in Eq. 8.12.

$$\text{PLR}(x, \theta) \equiv -2 \ln \frac{L(x=0, \theta)}{L(\hat{x}, \hat{\theta})}. \quad (8.12)$$

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. 8.55 and has a minimum at  $x = 0.335$ .

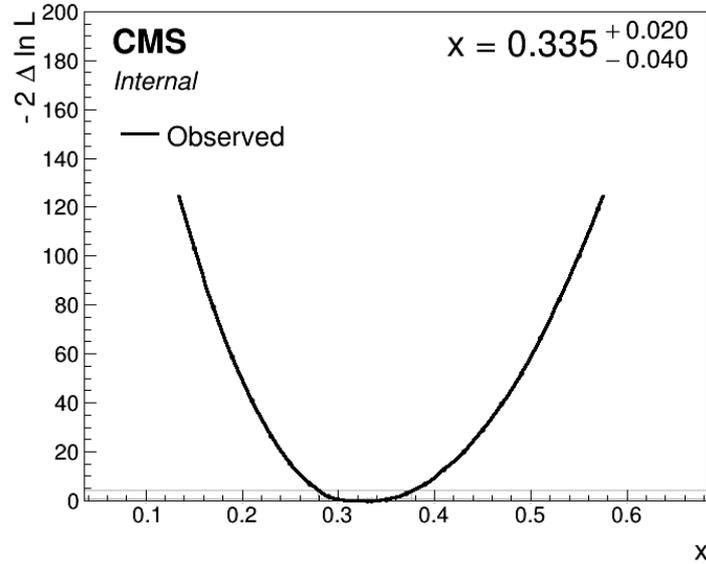


Figure 8.55. The PLR curve as a function of  $x$ .

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  $\theta$  is defined as the shift  $\Delta x$  in the parameter of interest when the nuisance is included at its  $\pm\sigma$  values:

$$\Delta x = x \Big|_{\theta \text{ at } \pm \sigma} - x_0. \quad (8.13)$$

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

$$P = \frac{\hat{\theta} - \theta_0}{\delta\theta}, \quad (8.14)$$

where  $\hat{\theta}$  is the  $\theta$  that maximises the profile likelihood,  $\theta_0$  is the pre-fit value,  $\delta\theta$  is the pre-fit uncertainty.

In order to 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. 8.56.

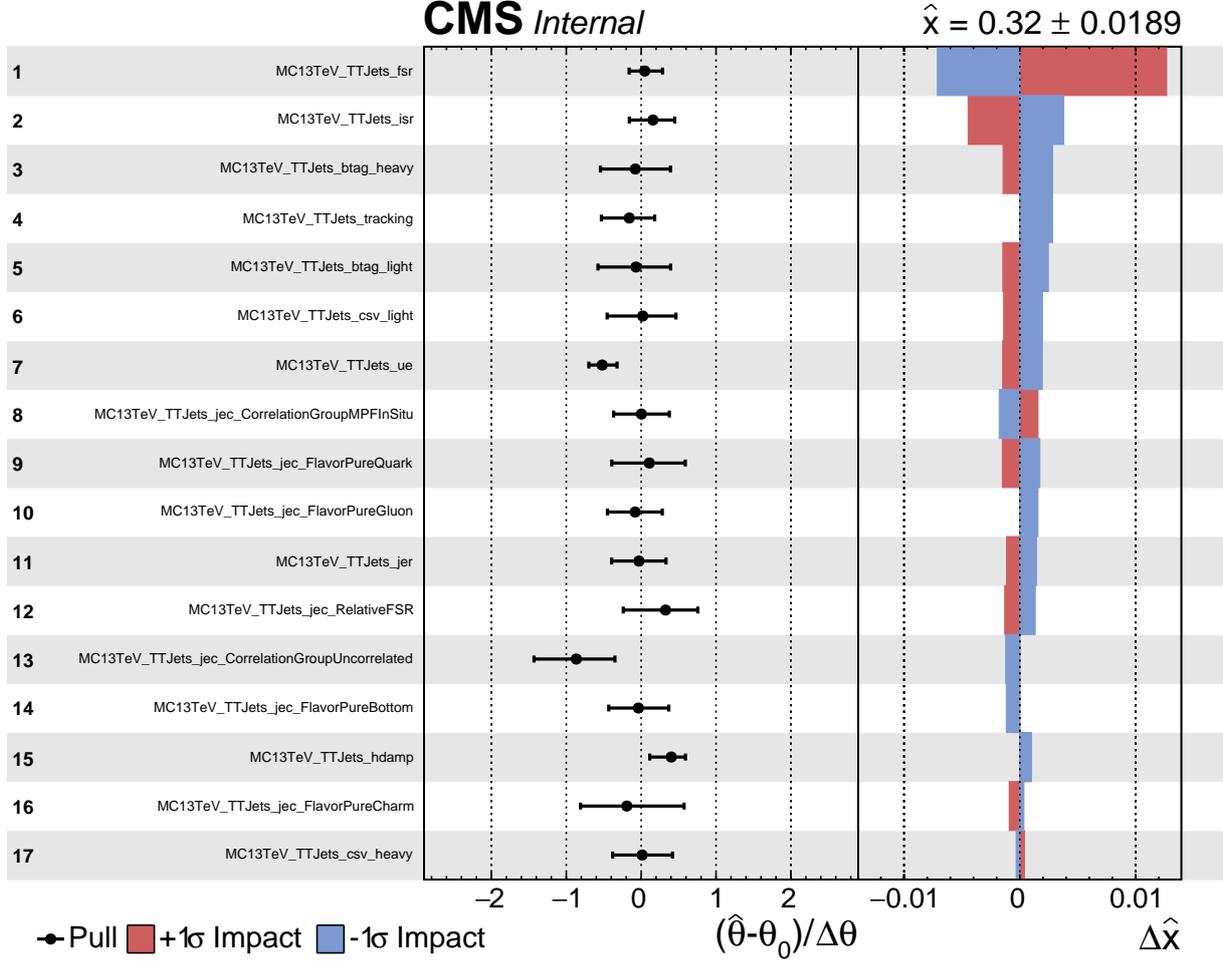


Figure 8.56. Impact and pull of different nuisance parameters.

Having obtained the value  $\hat{x} = 0.335$  (Fig. 8.55) 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 ( $H_{\text{alt}}$ ). The distribution of the test statistic for  $x = \hat{x}$  is plotted in Fig. 8.57.

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

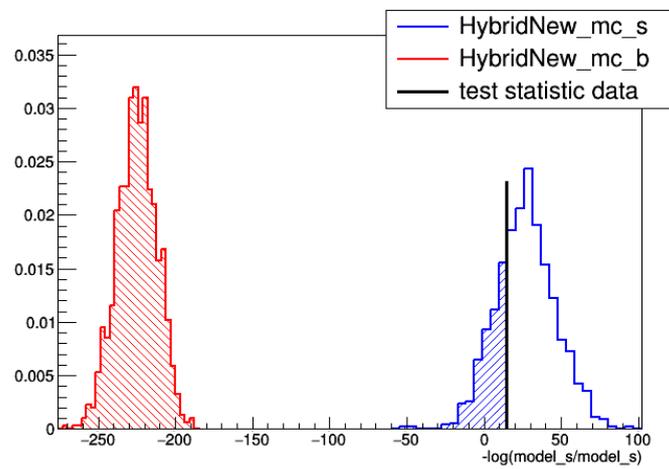


Figure 8.57. Distribution of the  $q/2$  under the assumption of  $t\bar{t}$  only hypothesis (red), a hypothesis of the signal being mixed of 66.5 %  $t\bar{t}$  and 33.5 % colour-flipped  $t\bar{t}$  process (blue) and  $q_{\text{obs}}/2$ .

## 9. 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.

# A. The datacard for the combine tool

## Datacard

```

imax 1
jmax *
kmax *

shapes * * files/TwoHypo.root
bin b1
observation -1

process
rate
MC13TeV_TTJets_jec_CorrelationGroupMPFIoSitu
MC13TeV_TTJets_jec_RelativeFSR
MC13TeV_TTJets_jec_CorrelationGroupUncorrelated
MC13TeV_TTJets_jec_FlavorPureGluon
MC13TeV_TTJets_jec_FlavorPureQuark
MC13TeV_TTJets_jec_FlavorPureCharm
MC13TeV_TTJets_jec_FlavorPureBottom
MC13TeV_TTJets_jer
MC13TeV_TTJets_btag_heavy
MC13TeV_TTJets_btag_light
MC13TeV_TTJets_csv_heavy
MC13TeV_TTJets_csv_light
MC13TeV_TTJets_tracking
MC13TeV_TTJets_isr
MC13TeV_TTJets_fsr
MC13TeV_TTJets_hdamp
MC13TeV_TTJets_ue
theory_group = MC13TeV_TTJets_isr MC13TeV_TTJets_fsr MC13TeV_TTJets_hdamp MC13TeV_TTJets_ue

$CHANNEL/$PROCESS $CHANNEL/$PROCESS_$SYSTEMATIC
b1 ttbarflip b1 ttbar b1 Single_top b1 W b1 DY b1 Multiboson b1 ttbar_V b1 QCD
-1 -1 0 0 1 1 1 2 3 4 5 6
-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1
1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0
shape shape

```

# Bibliography

- [1] 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].
- [2] 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. D* 83 (2011), p. 092002. DOI: 10.1103/PhysRevD.83.092002. arXiv: 1101.0648 [hep-ex].
- [3] 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].
- [4] 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].
- [5] 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](https://indico.cern.ch/event/276523/contributions/622482/attachments/502591/694047/colourflow_mseidel.pdf).
- [6] 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.
- [7] G. Abbiendi et al. “Colour reconnection in  $e^+e^- \rightarrow W^+W^-$  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].
- [8] 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].
- [9] P. Achard et al. “Search for color reconnection effects in  $e^+e^- \rightarrow W^+W^- \rightarrow$  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].
- [10] Makoto Kobayashi and Toshihide Maskawa. “CP Violation in the Renormalizable Theory of Weak Interaction”. In: *Prog. Theor. Phys.* 49 (1973), pp. 652–657. DOI: 10.1143/PTP.49.652.
- [11] F. Abe et al. “Observation of top quark production in  $p\bar{p}$  collisions”. In: *Phys. Rev. Lett.* 74 (1995), pp. 2626–2631. DOI: 10.1103/PhysRevLett.74.2626. arXiv: hep-ex/9503002 [hep-ex].

- [12] S. Abachi et al. “Observation of the top quark”. In: *Phys. Rev. Lett.* 74 (1995), pp. 2632–2637. DOI: 10.1103/PhysRevLett.74.2632. arXiv: hep-ex/9503003 [hep-ex].
- [13] A. H. Hoang. “1S and MS-bar bottom quark masses from Upsilon sum rules”. In: *Phys. Rev. D* 61 (2000), p. 034005. DOI: 10.1103/PhysRevD.61.034005. arXiv: hep-ph/9905550 [hep-ph].
- [14] 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].
- [15] Lars Husdal. “On Effective Degrees of Freedom in the Early Universe”. In: *Galaxies* 4.4 (2016), p. 78. DOI: 10.3390/galaxies4040078. arXiv: 1609.04979 [astro-ph.CO].
- [16] *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\\_toplhcg\\_sep18.pdf](https://twiki.cern.ch/twiki/pub/LHCPhysics/LHCTopWGSummaryPlots/tt_curve_toplhcg_sep18.pdf).
- [17] 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].
- [18] C. Patrignani et al. “Review of Particle Physics”. In: *Chin. Phys.* C40.10 (2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001.
- [19] 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].
- [20] 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].
- [21] 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.
- [22] R. Sekhar Chivukula et al. “Hadron Collider Production of Massive Color-Octet Vector Bosons at Next-to-Leading Order”. In: *Phys. Rev. D* 87.9 (2013), p. 094011. DOI: 10.1103/PhysRevD.87.094011. arXiv: 1303.1120 [hep-ph].
- [23] Paola Ferrario and German Rodrigo. “Constraining heavy colored resonances from top-antitop quark events”. In: *Phys. Rev. D* 80 (2009), p. 051701. DOI: 10.1103/PhysRevD.80.051701. arXiv: 0906.5541 [hep-ph].
- [24] Paul H. Frampton, Jing Shu, and Kai Wang. “Axigluon as Possible Explanation for p anti-p  $\rightarrow$  t anti-t Forward-Backward Asymmetry”. In: *Phys. Lett.* B683 (2010), pp. 294–297. DOI: 10.1016/j.physletb.2009.12.043. arXiv: 0911.2955 [hep-ph].
- [25] T. Aaltonen et al. “Forward-Backward Asymmetry in Top Quark Production in  $p\bar{p}$  Collisions at  $\sqrt{s} = 1.96$  TeV”. In: *Phys. Rev. Lett.* 101 (2008), p. 202001. DOI: 10.1103/PhysRevLett.101.202001. arXiv: 0806.2472 [hep-ex].

- [26] T. Aaltonen et al. “Evidence for a Mass Dependent Forward-Backward Asymmetry in Top Quark Pair Production”. In: *Phys. Rev. D* 83 (2011), p. 112003. doi: 10.1103/PhysRevD.83.112003. arXiv: 1101.0034 [hep-ex].
- [27] Tao Han, Ian Lewis, and Zhen Liu. “Colored Resonant Signals at the LHC: Largest Rate and Simplest Topology”. In: *JHEP* 12 (2010), p. 085. doi: 10.1007/JHEP12(2010)085. arXiv: 1010.4309 [hep-ph].
- [28] Ulrich Haisch and Susanne Westhoff. “Massive Color-Octet Bosons: Bounds on Effects in Top-Quark Pair Production”. In: *JHEP* 08 (2011), p. 088. doi: 10.1007/JHEP08(2011)088. arXiv: 1106.0529 [hep-ph].
- [29] Serguei Chatrchyan et al. “Search for Resonances in the Dijet Mass Spectrum from 7 TeV pp Collisions at CMS”. In: *Phys. Lett. B* 704 (2011), pp. 123–142. doi: 10.1016/j.physletb.2011.09.015. arXiv: 1107.4771 [hep-ex].
- [30] Georges Aad et al. “Search for New Physics in the Dijet Mass Distribution using 1 fb<sup>-1</sup> of pp Collision Data at  $\sqrt{s} = 7$  TeV collected by the ATLAS Detector”. In: *Phys. Lett. B* 708 (2012), pp. 37–54. doi: 10.1016/j.physletb.2012.01.035. arXiv: 1108.6311 [hep-ex].
- [31] CMS Collaboration. “Searches for dijet resonances in pp collisions at  $\sqrt{s} = 13$  TeV using data collected in 2016.” In: (2017).
- [32] 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.
- [33] 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.
- [34] 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. B* 716 (2012), pp. 30–61. doi: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235 [hep-ex].
- [35] 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. B* 716 (2012), pp. 1–29. doi: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214 [hep-ex].
- [36] *Overall view of the LHC*. CERN Document Server. 2014. URL: <http://cds.cern.ch/record/1708847>.
- [37] *CERN’s Accelerator Complex*. espace. 2014. URL: <https://espace.cern.ch/acc-sector/Pictures/CERN’s%20accelerator%20complex2015bis.png>.
- [38] Oliver S. Bruning et al. “LHC Design Report Vol.1: The LHC Main Ring”. In: (2004).
- [39] Stephen Myers. “The Large Hadron Collider 2008-2013”. In: *Int. J. Mod. Phys. A* 28 (2013), p. 1330035. doi: 10.1142/S0217751X13300354.
- [40] Lyndon Evans and Philip Bryant. “LHC Machine”. In: *JINST* 3 (2008), S08001. doi: 10.1088/1748-0221/3/08/S08001.
- [41] S. Chatrchyan et al. “The CMS Experiment at the CERN LHC”. In: *JINST* 3 (2008), S08004. doi: 10.1088/1748-0221/3/08/S08004.

- [42] 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].
- [43] Albert M Sirunyan et al. “Particle-flow reconstruction and global event description with the CMS detector”. In: (2017). arXiv: 1706.04965 [physics.ins-det].
- [44] 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>.
- [45] *The CMS tracker: addendum to the Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 2000. URL: <https://cds.cern.ch/record/490194>.
- [46] 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.
- [47] *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: <https://cds.cern.ch/record/357153>.
- [48] *The CMS muon project: Technical Design Report*. Technical Design Report CMS. Geneva: CERN, 1997. URL: <https://cds.cern.ch/record/343814>.
- [49] 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].
- [50] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “FastJet User Manual”. In: *Eur. Phys. J. C* 72 (2012), p. 1896. DOI: 10.1140/epjc/s10052-012-1896-2. arXiv: 1111.6097 [hep-ph].
- [51] 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).
- [52] G. Salam. *Anti-kt jet boundaries*. GitHub page. 2018. URL: <https://github.com/gavinsalam/antikt-jet-boundaries>.
- [53] 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].
- [54] TOP Physics Object Group. *Measurement of the underlying event in  $t$ - $t$  dilepton events at  $\sqrt{s}=13$  TeV*. draft CMS Analysis Note CMS-AN-2017-175. Geneva: CERN, 2018.
- [55] TOP Physics Object Group. *Jet shapes in  $t\bar{t}$  events at 13 TeV (2016 data)*. draft CMS Analysis Note CMS-AN-2017-159. Geneva: CERN, 2018.
- [56] *BRIL Work Suite*. site. 2015. URL: <http://cms-service-lumi.web.cern.ch/cms-service-lumi/brilwsdoc.html>.
- [57] *CMSSW Application Framework*. site. 2018. URL: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookCMSSWFramework>.
- [58] B. R. Webber. “Monte Carlo Simulation of Hard Hadronic Processes”. In: *Ann. Rev. Nucl. Part. Sci.* 36 (1986), pp. 253–286. DOI: 10.1146/annurev.ns.36.120186.001345.

- [59] Torbjorn Sjostrand. “Monte Carlo Generators”. In: *High-energy physics. Proceedings, European School, Aronsborg, Sweden, June 18-July 1, 2006*. 2006, pp. 51–74. arXiv: hep-ph/0611247 [hep-ph]. URL: <http://weblib.cern.ch/abstract?CERN-LCGAPP-2006-06>.
- [60] R. Eckart. “Stan Ulam, John von Neumann, and the Monte Carlo Method”. In: *Los Alamos Science, Special issue 15* (1987), pp. 131–137.
- [61] N. Metropolis and S. Ulam. “The Monte Carlo Method”. English. In: *Journal of the American Statistical Association* 44.247 (1949). Cited By :2264, pp. 335–341. URL: [www.scopus.com](http://www.scopus.com).
- [62] Stefano Frixione, Paolo Nason, and Carlo Oleari. “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”. In: *JHEP* 11 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv: 0709.2092 [hep-ph].
- [63] Stefano Frixione, Paolo Nason, and Giovanni Ridolfi. “A Positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction”. In: *JHEP* 09 (2007), p. 126. DOI: 10.1088/1126-6708/2007/09/126. arXiv: 0707.3088 [hep-ph].
- [64] Torbjörn Sjöstrand et al. “An Introduction to PYTHIA 8.2”. In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].
- [65] E. Boos et al. “Generic user process interface for event generators”. In: *Physics at TeV colliders. Proceedings, Euro Summer School, Les Houches, France, May 21-June 1, 2001*. 2001. arXiv: hep-ph/0109068 [hep-ph]. URL: <http://lss.fnal.gov/archive/preprint/fermilab-conf-01-496-t.shtml>.
- [66] Johan Alwall et al. “A Standard format for Les Houches event files”. In: *Comput. Phys. Commun.* 176 (2007), pp. 300–304. DOI: 10.1016/j.cpc.2006.11.010. arXiv: hep-ph/0609017 [hep-ph].
- [67] J. R. Andersen et al. “Les Houches 2013: Physics at TeV Colliders: Standard Model Working Group Report”. In: (2014). arXiv: 1405.1067 [hep-ph].
- [68] N. Kovalchuk. *New Colour Reconnection Tune for Pythia8 CUETP8M2*. slides. 2017. URL: [https://indico.cern.ch/event/632192/contributions/2555918/attachments/1445594/2226844/CRTune\\_CUETP8M2T4.pdf](https://indico.cern.ch/event/632192/contributions/2555918/attachments/1445594/2226844/CRTune_CUETP8M2T4.pdf).
- [69] S. Agostinelli et al. “GEANT4: A simulation toolkit”. In: *Nucl. Instrum. Meth.* A506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.
- [70] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv: 1405.0301 [hep-ph].
- [71] *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>.
- [72] *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>.

- [73] *McM Monte-Carlo Request Management*. site. 2015. URL: <https://cms-pdmv.cern.ch/mcm/>.
- [74] *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>.
- [75] E. Eichten et al. “Super Collider Physics”. In: *Rev. Mod. Phys.* 56 (1984). [Addendum: *Rev. Mod. Phys.* 58, 1065 (1986)], pp. 579–707. DOI: 10.1103/RevModPhys.56.579, 10.1103/RevModPhys.58.1065.
- [76] Simone Alioli et al. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv: 1002.2581 [hep-ph].
- [77] 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].
- [78] M. Seidel and Diaz C. *Powheg+Pythia8 tt̄ tuning:  $\alpha^{ISR}$  and  $h_{damp}$* . slides. 2016. URL: [https://indico.cern.ch/event/567994/contributions/2299854/attachments/1336057/2009649/tt\\_tuning\\_mseidel.pdf](https://indico.cern.ch/event/567994/contributions/2299854/attachments/1336057/2009649/tt_tuning_mseidel.pdf).
- [79] Holzner, A. *78 reconstructed vertices in event from high-pileup run 198609*. CERN CDS. 2012. URL: <http://cdsweb.cern.ch/record/1479324>.
- [80] G Apollinari et al. “High-Luminosity Large Hadron Collider (HL-LHC) : Preliminary Design Report”. In: (2015). DOI: 10.5170/CERN-2015-005.
- [81] CMS Collaboration. *CMS Luminosity - Public Results*. twiki. 2018. URL: [https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#2016\\_proton\\_proton\\_13\\_TeV\\_collis](https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults#2016_proton_proton_13_TeV_collis).
- [82] Rick Field. “Min-Bias and the Underlying Event at the LHC”. In: *Acta Phys. Polon.* B42 (2011), pp. 2631–2656. DOI: 10.5506/APhysPolB.42.2631. arXiv: 1110.5530 [hep-ph].
- [83] *Utilities for Accessing Pileup Information for Data*. twiki. 2015. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/PileupJSONFileforData>.
- [84] Adam, N. and others. *Generic Tag and Probe Tool for Measuring Efficiency at CMS with Early Data*. CMS Analysis Note CMS-AN-2009/111. 2009. URL: [http://cms.cern.ch/iCMS/jsp/openfile.jsp?tp=draft&files=AN2009\\_111\\_v1.pdf](http://cms.cern.ch/iCMS/jsp/openfile.jsp?tp=draft&files=AN2009_111_v1.pdf).
- [85] Vardan Khachatryan et al. “Performance of Electron Reconstruction and Selection with the CMS Detector in Proton-Proton Collisions at  $\sqrt{s} = 8$  TeV”. In: *JINST* 10.06 (2015), P06005. DOI: 10.1088/1748-0221/10/06/P06005. arXiv: 1502.02701 [physics.ins-det].
- [86] A. M. Sirunyan et al. “Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at  $\sqrt{s} = 13$  TeV”. In: *JINST* 13.06 (2018), P06015. DOI: 10.1088/1748-0221/13/06/P06015. arXiv: 1804.04528 [physics.ins-det].
- [87] *Reference muon id, isolation and trigger efficiencies for Run-II*. Twiki. 2015. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffsRun2>.
- [88] *Cut Based Electron ID for Run 2*. Twiki. 2015. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/CutBasedElectronIdentificationRun2>.

- [89] *Reference muon id, isolation and trigger efficiencies for Run-II*. Twiki. 2015. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MuonReferenceEffsRun2>.
- [90] *Electron Scale Factors for Run 2*. Twiki. 2015. URL: <https://twiki.cern.ch/twiki/bin/view/CMS/ElectronScaleFactorsRun2>.
- [91] W. Adam et al. “The CMS high level trigger”. In: *Eur. Phys. J. C* 46 (2006), pp. 605–667. DOI: 10.1140/epjc/s2006-02495-8. arXiv: hep-ex/0512077 [hep-ex].
- [92] P. Sphicas. “CMS: The TriDAS project. Technical design report, Vol. 2: Data acquisition and high-level trigger”. In: (2002).
- [93] *TOP Trigger (Run2)*. twiki. 2017. URL: <https://twiki.cern.ch/twiki/bin/view/CMS/TopTrigger>.
- [94] *Recommended Jet Energy Corrections and Uncertainties For Data and MC*. Twiki Page. 2017. URL: <https://twiki.cern.ch/twiki/bin/view/CMS/JECDataMC>.
- [95] Vardan Khachatryan et al. “Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV”. In: *JINST* 12.02 (2017), P02014. DOI: 10.1088/1748-0221/12/02/P02014. arXiv: 1607.03663 [hep-ex].
- [96] *Jet Energy Resolution Measurement*. Twiki Page. 2017. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetResolution>.
- [97] *BTagCalibration Documentation*. twiki. 2015. URL: <https://twiki.cern.ch/twiki/bin/view/CMS/BTagCalibration>.
- [98] A. M. Sirunyan et al. “Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV”. In: *JINST* 13.05 (2018), P05011. DOI: 10.1088/1748-0221/13/05/P05011. arXiv: 1712.07158 [physics.ins-det].
- [99] W. Adam et al. “Reconstruction of electrons with the Gaussian sum filter in the CMS tracker at LHC”. In: *eConf* C0303241 (2003). [*J. Phys.*G31,N9(2005)], TULT009. DOI: 10.1088/0954-3899/31/9/N01. arXiv: physics/0306087 [physics.data-an].
- [100] Daniele Bertolini et al. “Pileup Per Particle Identification”. In: *JHEP* 10 (2014), p. 059. DOI: 10.1007/JHEP10(2014)059. arXiv: 1407.6013 [hep-ph].
- [101] Serguei Chatrchyan et al. “Identification of b-quark jets with the CMS experiment”. In: *JINST* 8 (2013), P04013. DOI: 10.1088/1748-0221/8/04/P04013. arXiv: 1211.4462 [hep-ex].
- [102] Goh, J. and Seidel, M. *PseudoTopProducer*. github repository. 2017. URL: [https://github.com/intrepid42/cmssw/tree/pseudotoprivet\\_80x/TopQuarkAnalysis/TopEventProducers](https://github.com/intrepid42/cmssw/tree/pseudotoprivet_80x/TopQuarkAnalysis/TopEventProducers).
- [103] CMS collaboration. *Jet energy scale uncertainty sources*. Twiki Page. 2017. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JECUncertaintySources>.
- [104] Mariani, V. and others. *Measurement of the relative tracking efficiency for charged pions using the ratio of branching fractions for D0 decays to Kππ and Kπ*. CMS Analysis Note CMS-AN-2015/048. 2017. URL: [http://cms.cern.ch/iCMS/jsp/openfile.jsp?tp=draft&files=AN2015\\_248\\_v4.pdf](http://cms.cern.ch/iCMS/jsp/openfile.jsp?tp=draft&files=AN2015_248_v4.pdf).

- [105] *CMS Tracking POG Performance Plots for year 2016*. twiki. 2016. URL: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/TrackingPOGPlots2016>.
- [106] “Tracking POG plot results on 2015 data”. Detector Performance Summary. 2016. URL: <https://cds.cern.ch/record/2155558>.
- [107] *EvtGen*. HepForge Page. 2018. URL: <https://evtgen.hepforge.org/>.
- [108] D. J. Lange. “The EvtGen particle decay simulation package”. In: *Nucl. Instrum. Meth.* A462 (2001), pp. 152–155. DOI: 10.1016/S0168-9002(01)00089-4.
- [109] 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].
- [110] Michael H. Seymour and Andrzej Siodmok. “Constraining MPI models using  $\sigma_{eff}$  and recent Tevatron and LHC Underlying Event data”. In: *JHEP* 10 (2013), p. 113. DOI: 10.1007/JHEP10(2013)113. arXiv: 1307.5015 [hep-ph].
- [111] Sjostrand, T. *Particle Physics Phenomenology. Hadronization*. Lecture note. 2011. URL: <http://home.thep.lu.se/~torbjorn/ppp2011/lec7.pdf>.
- [112] T. Gleisberg et al. “Event generation with SHERPA 1.1”. In: *JHEP* 02 (2009), p. 007. DOI: 10.1088/1126-6708/2009/02/007. arXiv: 0811.4622 [hep-ph].
- [113] Stefan Höche and Stefan Prestel. “The midpoint between dipole and parton showers”. In: *Eur. Phys. J.* C75.9 (2015), p. 461. DOI: 10.1140/epjc/s10052-015-3684-2. arXiv: 1506.05057 [hep-ph].
- [114] Johannes Bellm et al. “Herwig 7.0/Herwig++ 3.0 release note”. In: *Eur. Phys. J.* C76.4 (2016), p. 196. DOI: 10.1140/epjc/s10052-016-4018-8. arXiv: 1512.01178 [hep-ph].
- [115] 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].
- [116] *Study of the underlying event, b-quark fragmentation and hadronization properties in  $t\bar{t}$  events*. CMS Physics Analysis Summary CMS-PAS-TOP-13-007. 2013. URL: <http://cds.cern.ch/record/1600599>.
- [117] Spyros Argyropoulos and Torbjörn Sjöstrand. “Effects of color reconnection on  $t\bar{t}$  final states at the LHC”. In: *JHEP* 11 (2014), p. 043. DOI: 10.1007/JHEP11(2014)043. arXiv: 1407.6653 [hep-ph].
- [118] Jesper R. Christiansen and Peter Z. Skands. “String Formation Beyond Leading Colour”. In: *JHEP* 08 (2015), p. 003. DOI: 10.1007/JHEP08(2015)003. arXiv: 1505.01681 [hep-ph].
- [119] Albert M. Sirunyan et al. “Study of the underlying event in top quark pair production in pp collisions at 13 TeV”. In: (2018). arXiv: 1807.02810 [hep-ex].
- [120] Veckalns, V. *CFAT*. Address. 2018. URL: <https://github.com/vveckaln/CFAT>.
- [121] R. Brun and F. Rademakers. “ROOT - An Object Oriented Data Analysis Framework”. In: *Nucl. Inst.* A 389 (1997), pp. 81–86.

- [122] 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].
- [123] Veckalns, V. *CompoundHistoUnfolding*. Address. 2018. URL: <https://github.com/vveckaln/COMPOUNHISTOUNFOLDING>.
- [124] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *Eur. Phys. J. C* 71 (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].
- [125] *The CMS combine tool*. Address. 2018. URL: <https://cms-hcomb.gitbooks.io/combine/>.
- [126] W. Verkerke and D. Kirkby. *RooFit Users Manual v2.91*. Address. 2008. URL: [https://root.cern.ch/download/doc/RooFit\\_Users\\_Manual\\_2.91-33.pdf](https://root.cern.ch/download/doc/RooFit_Users_Manual_2.91-33.pdf).