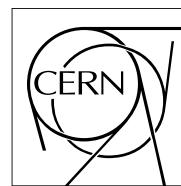


The Compact Muon Solenoid Experiment  
**Analysis Note**

The content of this note is intended for CMS internal use and distribution only



23 May 2017 (v9, 06 September 2019)

# Measurement of the top-anti-top differential production cross section of high transverse momentum top quarks in the all-hadronic final state using the 2016 proton-proton collision data at $\sqrt{s} = 13$ TeV.

Yorgos Tsipolitis  
National Technical University of Athens

Konstantinos Kousouris  
National Technical University of Athens

Georgios Bakas  
National Technical University of Athens

Garyfallia Paspalaki  
National Center of Scientific Research "Demokritos"

Andrea Castro  
Universita e INFN, Bologna

Federico Celli  
Universita e INFN, Bologna

Pralay Kumar Mal  
National Institute of Science Education and Research

## Abstract

A measurement of the production cross section of high transverse momentum ( $p_T$ ) top quark pairs is reported. The dataset was collected during 2016 with the CMS detector at the CERN LHC from proton-proton collisions at a center-of-mass energy of 13 TeV, and corresponds to an integrated luminosity of  $35.9 \text{ fb}^{-1}$ . The measurement uses events where either both top quark candidates decay hadronically and are reconstructed as large- $R$  jets with  $p_T > 400 \text{ GeV}$ . The cross section is extracted differentially as a function of kinematic variables of the top quark or top quark pair system. The results are unfolded to the particle and parton levels, and are compared to various theoretical models. The measured cross section is significantly lower, by up to 40%, in the phase space of interest, compared

to the theory predictions, while the normalized differential cross sections are consistent between data and theory.

This box is only visible in draft mode. Please make sure the values below make sense.

PDFAuthor: Konstantinos Kousouris, Georgios Tsipolitis, Georgios Bakas, Prolay Kumar Mal, Andrea Castro, Federico Celli

PDFTitle: Measurement of the top-anti-top quark differential production cross section in the all-hadronic final state using the 2016 proton-proton collision data at  $\text{sqrt}(s) = 13 \text{ TeV}$ .

PDFSubject: CMS

PDFKeywords: CMS, physics, top, qcd, jets

Please also verify that the abstract does not use any user defined symbols

DRAFT

## 1 Contents

2	1	Introduction . . . . .	2
3	2	Samples . . . . .	2
4	2.1	Data . . . . .	2
5	2.2	Simulation . . . . .	3
6	3	Trigger . . . . .	5
7	4	Reconstruction and Selection . . . . .	7
8	4.1	Object Reconstruction . . . . .	7
9	4.2	Selection . . . . .	7
10	4.3	Multivariate Discriminant . . . . .	10
11	4.4	Parton level . . . . .	13
12	4.5	Particle level . . . . .	13
13	5	Signal Extraction . . . . .	13
14	5.1	Inclusive cross section . . . . .	14
15	5.2	Differential cross sections . . . . .	15
16	6	Data vs Monte Carlo . . . . .	21
17	7	Systematic Uncertainties . . . . .	25
18	8	Fiducial Measurement . . . . .	35
19	9	Unfolded Measurement . . . . .	38
20	9.1	Parton Level . . . . .	38
21	9.2	Particle Level . . . . .	45
22	10	Summary . . . . .	52
23	A	Data vs simulation with default MC normalization . . . . .	57
24	B	NN Input Variables . . . . .	60
25	C	Control Distributions . . . . .	64
26	D	Background Sensitivity to Pileup . . . . .	68
27	E	Unfolding Tests . . . . .	69
28	F	Simulated b tagging efficiencies . . . . .	75
29	G	Correlation Matrices . . . . .	75

## 30 1 Introduction

31 The top quark completes the third generation of quarks in the standard model (SM), and the  
 32 precise knowledge of its properties is critical for the overall understanding of the theory. Mea-  
 33 surements of the top-anti-top quark pair ( $t\bar{t}$ ) production cross section confront the predictions  
 34 from quantum chromodynamics (QCD) and have the potential to constrain the QCD param-  
 35 eters, while being sensitive to physics beyond the SM. Also, the  $t\bar{t}$  production process is a  
 36 dominant SM background to searches for new physics phenomena and therefore its precise  
 37 knowledge is essential for new discoveries.

38 The large  $t\bar{t}$  yield expected in proton-proton (pp) collisions at the CERN LHC allows to perform  
 39 measurements of the  $t\bar{t}$  production rate in a large phase space, and, more importantly, differ-  
 40 entially, as a function of the  $t\bar{t}$  kinematic properties. Such measurements have been performed  
 41 by the ATLAS [1–6] and CMS [7–15] Collaborations at 7, 8, and 13 TeV center-of-mass energies,  
 42 under the hypothesis of the resolved final state, where the decay products of the  $t\bar{t}$  pair can  
 43 be reconstructed individually. This hypothesis is valid for top quark transverse momenta,  $p_T$ ,  
 44 up to approximately 500 GeV. However, at higher  $p_T$  ( $p_T/m \approx 1$ ), the top quark decay prod-  
 45 ucts are highly collimated ("boosted") and they can no longer be reconstructed separately. In  
 46 order to explore the highly boosted phase space, hadronic top quark decays are reconstructed  
 47 as large-radius jets. Previous efforts in this domain by the ATLAS [16, 17] and CMS [18–20]  
 48 Collaborations confirm that it is feasible to perform precise differential measurements of the  $t\bar{t}$   
 49 production and have also shown interesting deviations from the theory predictions.

50 In this note, a measurement of the differential boosted  $t\bar{t}$  production cross section in the hadronic  
 51 final state is presented, using pp collisions at  $\sqrt{s} = 13$  TeV recorded with the CMS detector  
 52 during the 2016 LHC run and amounting to a total integrated luminosity of  $35.9 \text{ fb}^{-1}$ . In the  
 53 hadronic decay channel, each W boson arising from the top quark decays into a pair of light  
 54 quarks. As a result, the final state consists of at least six partons (more are possible due to  
 55 initial- and final-state radiation), two of which are b quarks. Due to the high boost consid-  
 56 ered in this measurement ( $p_T > 400$  GeV), the top quarks are reconstructed unambiguously as  
 57 large-radius jets and the final state consists of at least two such jets.

58 This note is organized as follows: section 2 describes the data and Monte Carlo samples used  
 59 in the analysis and section 3 discusses the triggers and the measurement of the corresponding  
 60 efficiency. Then, section 4 presents the details of the jet reconstruction and the event selection.  
 61 The signal extraction is discussed in section 5 and a comparison between data and simulation is  
 62 reported in section 6. The systematic uncertainties are discussed in section 7, while the results  
 63 are presented in sections 8 and 9. Finally, section 10 summarizes the results.

## 64 2 Samples

65 In this section we describe the data and Monte Carlo samples that are used in the analysis.

### 66 2.1 Data

67 The collision events used for the measurement of the  $t\bar{t}$  cross section have been collected with  
 68 the triggers described in Section 3 that are part of the JetHT primary dataset. Table 1 lists  
 69 the samples that correspond to different data-taking eras and have been reconstructed in the  
 70 03Feb2017 reprosessing. The good Run and luminosity section list used is contained in the  
 71 certification file Cert\_271036-284044\_13TeV\_23Sep2016ReReco\_Collisions16\_JSON.txt.  
 72 The total integrated luminosity of the analyzed data is  $35.9 \text{ fb}^{-1}$ .

<sup>73</sup> In addition to the data from the JetHT dataset, we have used the SingleMuon dataset for the  
<sup>74</sup> measurement of the trigger efficiency. The eras and run ranges are the same as the ones in  
<sup>75</sup> Table 1.

Table 1: Data samples.

Sample	Run range	Luminosity ( $\text{pb}^{-1}$ )
/JetHT/Run2016B-03Feb2017_ver2-v2/MINIAOD	273150-275376	5750
/JetHT/Run2016C-03Feb2017-v1/MINIAOD	275656-276283	2573
/JetHT/Run2016D-03Feb2017-v1/MINIAOD	276315-276811	4242
/JetHT/Run2016E-03Feb2017-v1/MINIAOD	276947-277420	4025
/JetHT/Run2016F-03Feb2017-v1/MINIAOD	277932-278808	3105
/JetHT/Run2016G-03Feb2017-v1/MINIAOD	278820-280385	7576
/JetHT/Run2016H-03Feb2017_ver2-v1/MINIAOD	281613-284035	8435
/JetHT/Run2016H-03Feb2017_ver3-v1/MINIAOD	284036-284044	216

## <sup>76</sup> 2.2 Simulation

<sup>77</sup> Monte Carlo simulation is used to generate samples for the  $t\bar{t}$  signal and to model the kinematic distributions of some of the background processes. Samples of simulated  $t\bar{t}$  events have  
<sup>78</sup> been generated at next-to-leading order (NLO) in QCD using POWHEG v2 [21–25], assuming  
<sup>79</sup> a top quark mass of  $m_t = 172.5 \text{ GeV}$ . Single top quark production in the  $t$  channel or in association with a  $W$  boson are simulated at NLO with POWHEG [26]. The production of  $W$  or  
<sup>80</sup>  $Z$  bosons in association with jets (+jets), as well as QCD multijet events, are simulated with  
<sup>81</sup> MG5\_AMC@NLO [27] at leading order (LO), with the MLM matching algorithm [28].

<sup>82</sup> All simulated events are processed with PYTHIA 8.212 [29, 30] for modeling of the parton showering, hadronization, and underlying event (UE). The NNPDF 3.0 [31] Parton Distribution  
<sup>83</sup> Functions (PDF) are used throughout, and the CUETP8M1 UE tune [32] is used for all processes  
<sup>84</sup> except for the  $t\bar{t}$ ,  $t\bar{t}H$  and single top quark processes, for which the tune CUETP8M2T4 [33] is  
<sup>85</sup> used. The Simulation of the CMS detector response is based on GEANT4 [34]. Additional pp  
<sup>86</sup> interactions in the same or neighbouring bunch crossings (pileup) are simulated with PYTHIA  
<sup>87</sup> and overlaid with generated events according to the pileup distribution measured in data.

<sup>88</sup> The various simulated processes are normalized to the best known theoretical cross sections,  
<sup>89</sup> namely the  $t\bar{t}$ , +jets, and single top quark samples are normalized to NNLO precision in QCD [35–  
<sup>90</sup> 37].

<sup>91</sup> The measured cross sections for the  $t\bar{t}$  process are compared to theoretical predictions provided  
<sup>92</sup> by the following Monte Carlo models: POWHEG combined with PYTHIA for the parton showering,  
<sup>93</sup> as described above, or combined with HERWIG ++ [38] and the corresponding EE5C UE  
<sup>94</sup> tune [39]. In addition,  $t\bar{t}$  events were also generated with MC@NLO [27] combined with PYTHIA  
<sup>95</sup> for the parton showering.

<sup>96</sup> In the first part of Table 2 we show the signal samples, while in the second part we show the  
<sup>97</sup> background ones that include QCD multijet production, associated production of vector bosons  
<sup>98</sup> ( $W^\pm, Z$ ) with jets, and single-top production. The list shows the total number of events analyzed,  
<sup>99</sup> including all samples of the same kind (nominal, extensions, backup). The reconstruction of the Monte Carlo samples belongs to the era RunIIISummer16MiniAODv2-PUMoriond17\_80X\_mcRun2\_asymptotic\_v10, and we have used the MINIAODSIM data tier. The reported cross sections have been used in  
<sup>100</sup> the normalization of the various processes in the data vs MC comparison plots.

Table 2: Monte Carlo samples.

Sample	Events ( $\times 10^6$ )	$\sigma$ (pb)
TT_TuneCUETP8M2T4_13TeV-powheg-pythia8, backup	153.4	832
TTJets_TuneCUETP8M2T4_13TeV-amcatnloFXFX-pythia8, backup	87.9	832
TT_TuneEE5C_13TeV-powheg-herwigpp, ext2, ext3	57.7	832
TT_hdampDOWN_TuneCUETP8M2T4_13TeV-powheg-pythia8, ext1	57.9	832
TT_hdampUP_TuneCUETP8M2T4_13TeV-powheg-pythia8, ext1	58.2	832
TT_TuneCUETP8M2T4down_13TeV-powheg-pythia8, ext1	58.3	832
TT_TuneCUETP8M2T4up_13TeV-powheg-pythia8, ext1	58.9	832
TT_TuneCUETP8M2T4_13TeV-powheg-fsrdown-pythia8, ext1, ext2	155.7	832
TT_TuneCUETP8M2T4_13TeV-powheg-fsrup-pythia8, ext1, ext2	152.6	832
TT_TuneCUETP8M2T4_13TeV-powheg-isrdown-pythia8, ext1, ext2	148.5	832
TT_TuneCUETP8M2T4_13TeV-powheg-isrup-pythia8_ext1, ext2	156.5	832
TT_TuneCUETP8M2T4_mtop1665_13TeV-powheg-pythia8, ext1, ext2, backup	19.4	832
TT_TuneCUETP8M2T4_mtop1715_13TeV-powheg-pythia8	19.6	832
TT_TuneCUETP8M2T4_mtop1735_13TeV-powheg-pythia8	19.4	832
TT_TuneCUETP8M2T4_mtop1755_13TeV-powheg-pythia8, ext1, ext2	59.4	832
TT_TuneCUETP8M2T4_mtop1785_13TeV-powheg-pythia8	16.4	832
QCD_HT200to300_TuneCUETP8M1_13TeV-madgraphMLM-pythia8, ext1	57.6	1.712e+6
QCD_HT300to500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8, ext1	54.5	3.477e+5
QCD_HT500to700_TuneCUETP8M1_13TeV-madgraphMLM-pythia8, ext1	62.3	3.21e+4
QCD_HT700to1000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8, ext1	45.4	6831
QCD_HT1000to1500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8, ext1	15.1	1207
QCD_HT1500to2000_TuneCUETP8M1_13TeV-madgraphMLM-pythia8, ext1	11.8	119.9
QCD_HT2000toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8, ext1	6.0	25.24
DYJetsToQQ_HT180_13TeV-madgraphMLM-pythia8	12.1	1187
WJetsToQQ_HT180_13TeV-madgraphMLM-pythia8	22.4	2788
ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	67.2	136.02
ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	38.8	80.95
ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1_ext1	6.9	35.6
ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1_ext1	6.9	35.6

### 106 3 Trigger

107 The trigger path employed for the collection of signal events uses single-jet L1 seeds that re-  
 108 quire the presence of a jet with  $p_T > 180 \text{ GeV}$ . At HLT jets are reconstructed from (online)  
 109 particle flow candidates using the anti- $\text{kt}$  algorithm with distance parameter  $R = 0.8$  and their  
 110 mass, after trimming of soft particles, must be greater than  $30 \text{ GeV}$ . Interesting events are re-  
 111 quired to have at least two such jets with  $p_T > 280(200) \text{ GeV}$  for the leading (trailing) one.  
 112 Finally, at least one of the two jets should be tagged as a b-jet, using the online CSV algorithm.  
 113 The aforementioned trigger path ran unprescaled for the duration of the 2016 run, collecting an  
 114 integrated luminosity of  $35.9 \text{ fb}^{-1}$ . A second, prescaled, path, using the same L1 seed, was also  
 115 employed, with identical kinematic requirements but no b-tagging cut, which ran in parallel  
 116 and collected an integrated luminosity of  $1.67 \text{ fb}^{-1}$ . This path is used for the selection of a  
 117 control QCD sample, as described later. All the triggers described above are summarized in  
 118 Table 3. Finally, it should be noted that the pileup profile of the prescaled control trigger is  
 119 shifted to lower number of interactions (Fig. 1), because the path tended to collect more data  
 120 towards the end of the fills when the instantaneous luminosity was lower.

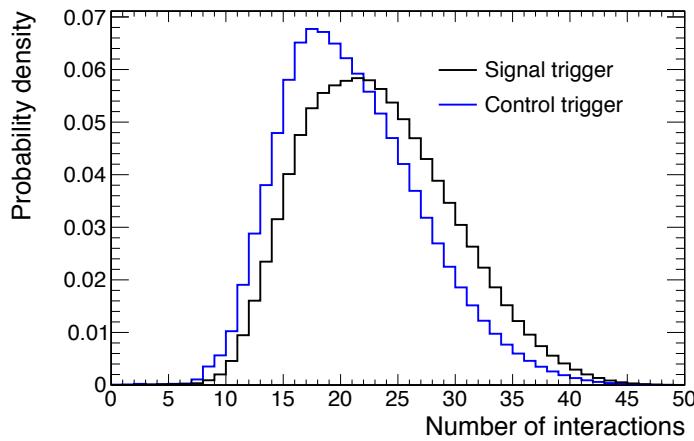


Figure 1: Pileup profile for the signal and control trigger paths.

Table 3: Summary of triggers used in the analysis.

Trigger	Purpose
L1_SingleJet180 OR L1_SingleJet200	L1 seed
HLT_AK8DiPFJet280_200_TrimMass30_BTagCSV_p20	signal HLT path
HLT_AK8DiPFJet280_200_TrimMass30	control HLT path

121 The efficiency of the signal trigger path is measured with respect to an orthogonal path that  
 122 requires the presence of an isolated muon with  $p_T > 27 \text{ GeV}$  (HLT\_IsoMu27). Figure 2 shows  
 123 the trigger efficiency as a function of the second jet  $p_T$  in events with at least two reconstructed  
 124 jets, with at least one of them containing a b-tagged subjet. For details about the offline recon-  
 125 struction see Section 4. The efficiency measured in data is compared to the simulated efficiency,  
 126 showing an excellent agreement. The offline selection requires that the second jet  $p_T$  is greater  
 127 than  $400 \text{ GeV}$ , which is at the beginning of the efficiency plateau.

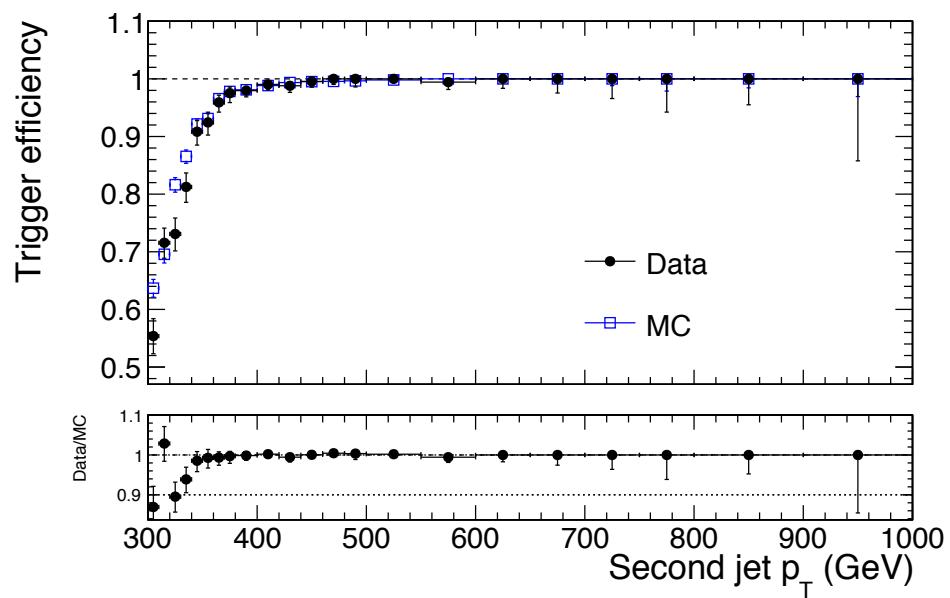


Figure 2: Trigger efficiency for the signal path of the analysis.

## 128 4 Reconstruction and Selection

129 In this section we present the objects used in the analysis, the multivariate method that dis-  
 130 criminated  $t\bar{t}$  events from the QCD multijet background, and we describe the event selection.

### 131 4.1 Object Reconstruction

132 The leptons (muons, electrons) used in the analysis come from the default reconstructed col-  
 133 lections in CMS samples ('slimmedMuons' and 'slimmedElectrons') and must have  
 134  $p_T > 20 \text{ GeV}$ . Muons are required to pass the medium ID working point and electrons should  
 135 pass the tight working point, while both lepton types should have a relative mini-isolation less  
 136 than 0.1.

137 Jets are reconstructed from particle-flow (PF) candidates that have undergone charged-hadron  
 138 subtraction (CHS) in order to suppress the contribution from multiple soft interactions in the  
 139 same bunch crossing (pileup). The momentum 4-vectors of the PF candidates are clustered  
 140 with the anti- $\text{kt}$  algorithm with distance parameter  $R = 0.8$ , hence we refer to AK8 PFCHS  
 141 jets. These jets are required to pass the tight jet ID. Furthermore, an algorithm is run to identify  
 142 subjets with distance parameter  $R = 0.4$  within the AK8 jets. Finally, the soft-drop technique  
 143 is used in order to evaluate the mass of the AK8 jet with suppressed pileup contribution. For  
 144 all the aforementioned algorithms we have used the default definitions in the CMS software  
 145 and the default collections in the MINIAOD data and Monte Carlo samples. For the energy  
 146 calibration of the AK8 jets we have used the recommended corrections by the JME POG in the  
 147 80X\_dataRun2\_2016SeptRepro\_v7 and 80X\_mcRun2\_asymptotic\_2016\_TrancheIV\_v8 global  
 148 tags.

149 The selection of AK8 jets that originate from the top decay relies on the identification of a b-jet  
 150 within them. For this purpose, we use the CSVv2 b-tagging algorithm applied on the AK4  
 151 subjets of each AK8 jet. In particular, we use the medium working point, which requires the  
 152 value of the tagger to be greater than 0.8484.

153 Since leptons are also reconstructed as jets, we perform cross cleaning by removing each iden-  
 154 tified lepton from the jet collection with geometrical matching in the  $\eta - \phi$  space: if a jet has  
 155  $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$  from any accepted lepton candidate it is removed from the jet  
 156 collection.

### 157 4.2 Selection

158 The baseline selection, summarized in Table 4 is common for all regions used in the analysis  
 159 and it requires at least two jets in the event with  $p_T > 400 \text{ GeV}$  and softDrop masses in the  
 160 range  $(50, 300) \text{ GeV}$ . Also, a lepton veto is applied in order to minimize the probability to select  
 161 leptonic top decays. Then, on top of the baseline selection, we define four specific regions  
 162 (Table 4) based on the NN output (Section 4.3), the jets' softDrop masses, and the number of  
 163 b-tagged subjets in each jet, that serve different analysis purposes. The signal region (SR) is  
 164 where we perform the differential measurements and it requires both jets to have a b-tagged  
 165 subjet, a tighter selection on the jets' masses, and a high mva value. Figure 3 shows the top  
 166 decay mode of the selected  $t\bar{t}$  events, where more than 95% come from the hadronic channel.  
 167 The QCD control region (CR) is the same as the SR but with the b-tagging requirement reverted  
 168 (the jets should not contain a b-tagged subjet) and it is used to get from data the shape of the  
 169 QCD background for each variable of interest. Then, we employ the signal region A ( $SR_A$ ),  
 170 which is used to determine the normalization of the QCD background, and finally the signal  
 171 region B ( $SR_B$ ), which is used to constrain some of the signal modelling uncertainties. Figure 4

<sup>172</sup> shows a cartoon of the signal regions on the plane defined by the NN output and the jet masses.

Table 4: Baseline selection requirements.

Observable	Requirement
$N_{\text{jets}}$	$> 1$
$N_{\text{leptons}}$	$= 0$
$p_T^{\text{jet}1,2}$	$> 400 \text{ GeV}$
$m_{SD}^{\text{jet}1,2}$	$(50, 300) \text{ GeV}$

Table 5: Selection requirements per analysis region.

Region	Trigger	Offline Requirements	Purpose
$SR$	signal	$\text{Base+NN} > 0.8 + 2\text{btags} + m_{SD}^{\text{jet}1,2} \in (120, 220) \text{ GeV}$	signal region
$SR_A$	signal	$\text{Base+NN} > 0.8 + 2\text{btags}$	QCD fit region
$SR_B$	signal	$\text{Base} + 2\text{btags} + m_{SD}^{\text{jet}1,2} \in (120, 220) \text{ GeV}$	signal systematics region
$CR$	control	$\text{Base+NN} > 0.8 + 0\text{btags} + m_{SD}^{\text{jet}1,2} \in (120, 220) \text{ GeV}$	QCD control region

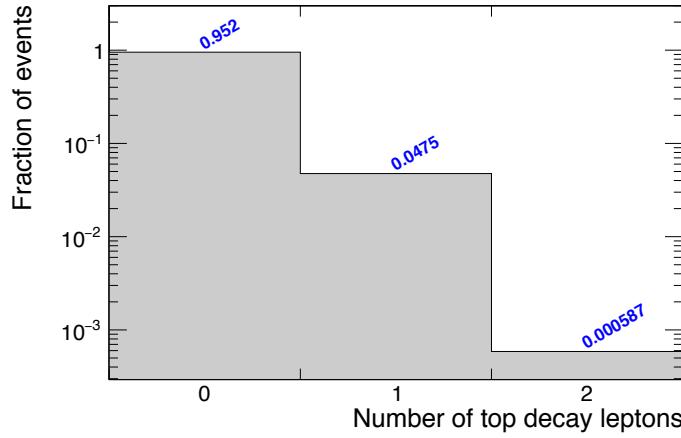


Figure 3: Distribution of  $t\bar{t}$  decay modes after the signal selection of the analysis.

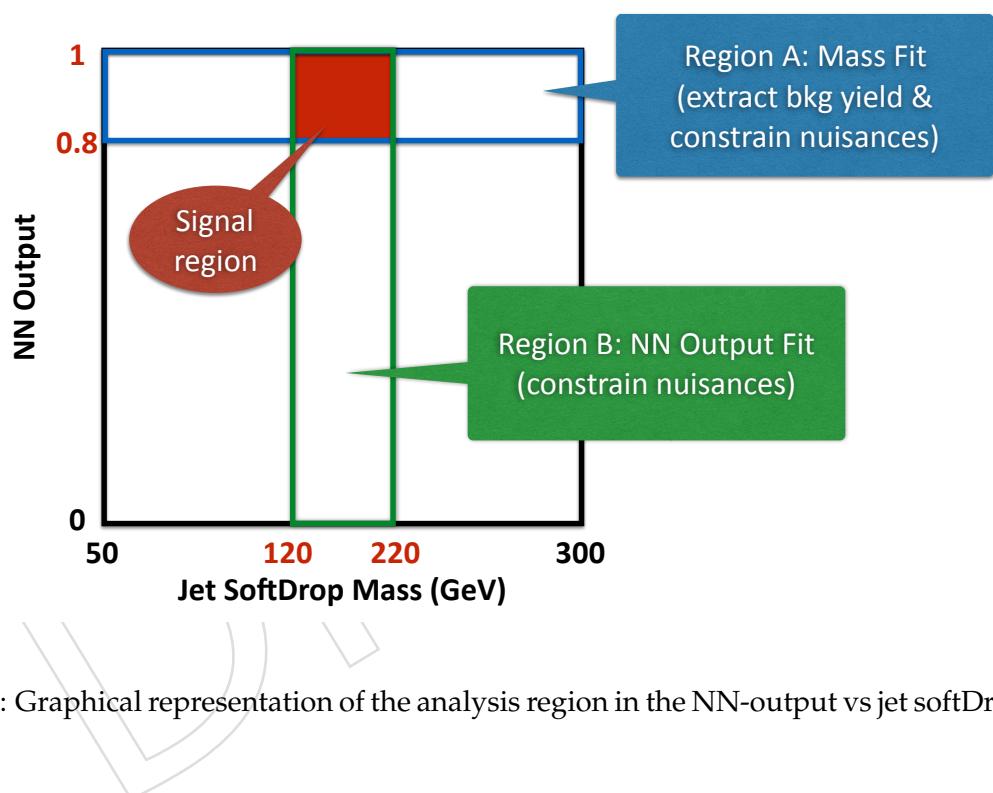


Figure 4: Graphical representation of the analysis region in the NN-output vs jet softDrop mass.

173 **4.3 Multivariate Discriminant**

174 In order to discriminate between events that come from  $t\bar{t}$  decays and QCD multijet production  
 175 we rely on variables that reveal the jet substructure. In particular, we use the "N-subjettiness"  
 176  $\tau_N$ , defined as:

$$\tau_N = \frac{1}{\sum_k p_{T,k} R} \sum_k p_{T,k} \min\{\Delta R_{1,k}, \Delta R_{2,k}, \dots, \Delta R_{N,k}\}, \quad (1)$$

177 where  $N$  denotes the reconstructed candidate subjets and  $k$  runs over the constituent particles  
 178 in the jet. The variable  $\Delta R_{i,k} = \sqrt{(\Delta y_{i,k})^2 + (\Delta \phi_{i,k})^2}$  is the angular distance between the can-  
 179 didate subjet  $i$  and the particle  $k$ . The variable  $R$  is the characteristic jet radius ( $R = 0.8$  in our  
 180 case).

181 Since the signal final state ( $t\bar{t}$  hypothesis) involves two jets with a top-quark decay signature,  
 182 we consider the N-subjettiness variables  $\tau_{1,2,3}$  for both leading jets in the event, resulting in  
 183 six discriminating variables per event. The distributions of these variables in  $t\bar{t}$  and QCD  
 184 simulated events are shown in Fig. 5, indicating that there is significant separating power in  
 185 them. The linear correlation between the variables is shown in Fig. 6. In order to exploit fully  
 186 the differences of the N-subjettiness variables we turn to multivariate discriminants, imple-  
 187 mented in the TMVA package [40]. As a baseline we consider a simple Fisher discriminant,  
 188 which is a linear combination of the variables. Then, a feed-forward neural network (NN)  
 189 is constructed to enhance the discriminating power quantified by the integral of the receiver-  
 190 operator-characteristic (ROC) curve. The network (Fig. 7) is composed of two hidden layers  
 191 with 16 and 4 nodes, respectively, each of one using a sigmoid activation function. The dis-  
 192 tribution of the NN output and the convergence of its training are shown in Fig. 8. As shown  
 193 in Fig. 9, the NN is clearly better than the simple Fisher discriminant and it also outperforms  
 194 slightly a Boosted Decision Tree composed of 500 trees and trained with the Gradient Boost  
 195 method with shrinkage parameter equal to 0.1. More complicated NN architectures have not  
 196 been found to improve the performance. The training is performed with events that pass the  
 197 baseline selection and the MC samples are divided in two halves, one used for the training and  
 198 one for the testing.

why is the training set 50% same with test set?

Usually we have a 80, 20 distribution

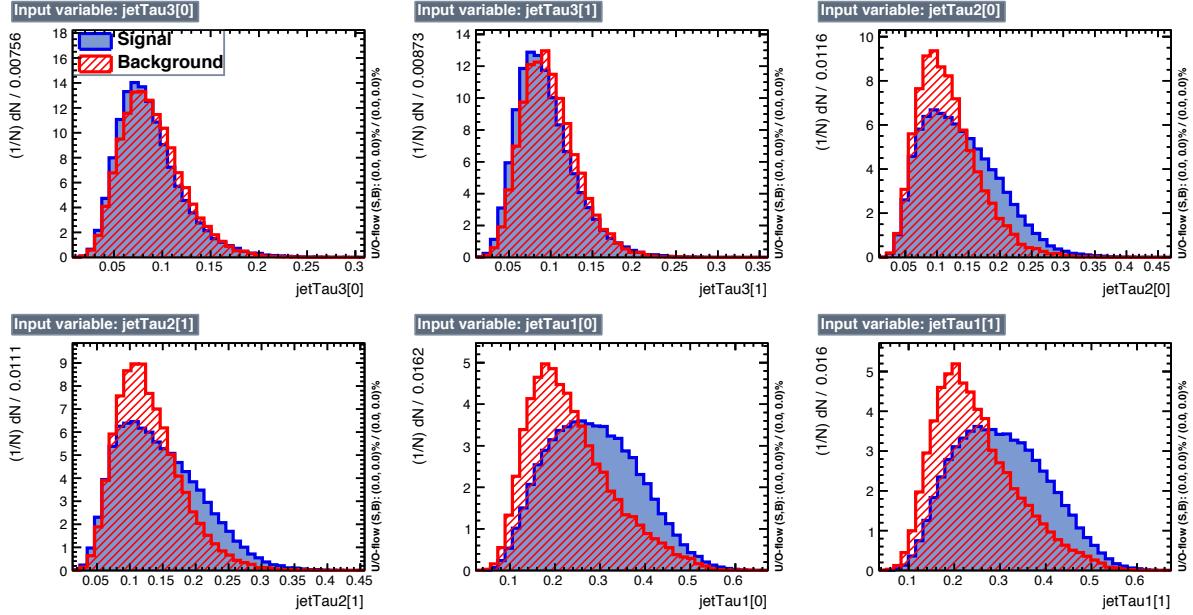


Figure 5: Discriminating variables used for the separation of  $t\bar{t}$  from QCD events.

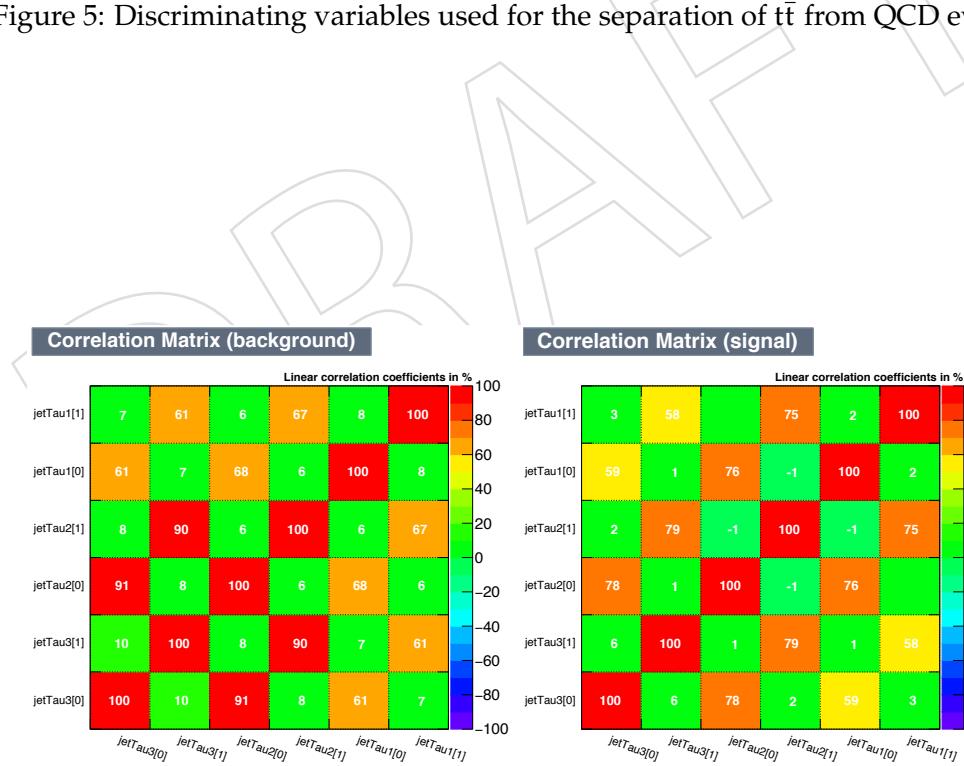


Figure 6: Linear correlation between the discriminating variables in background (left) and signal (right) events.

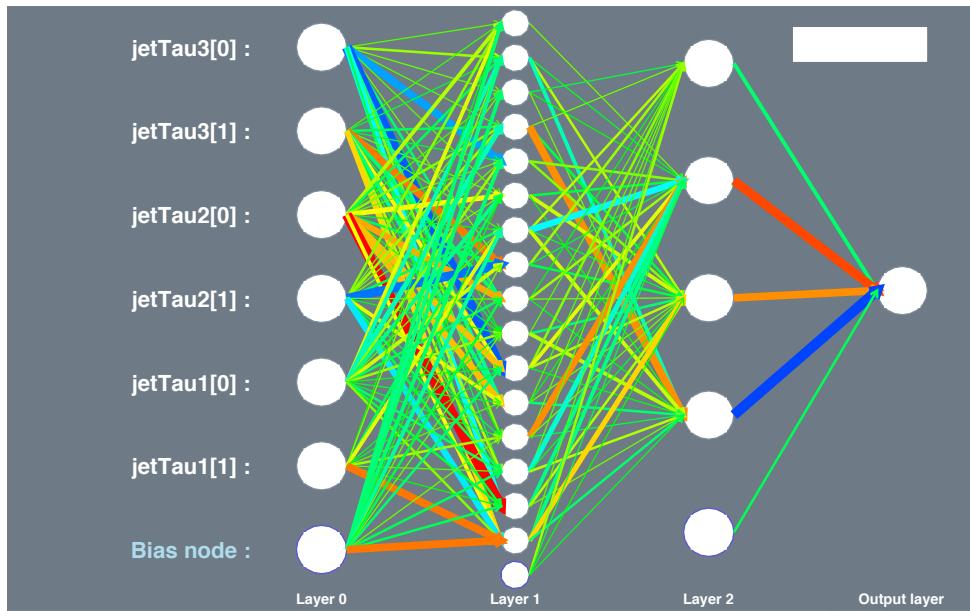


Figure 7: Neural network architecture.

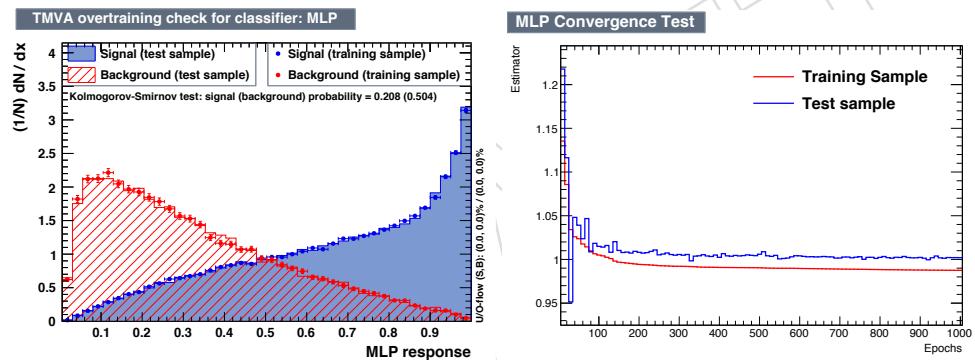


Figure 8: Output of the NN (left). Convergence of the NN training (right).

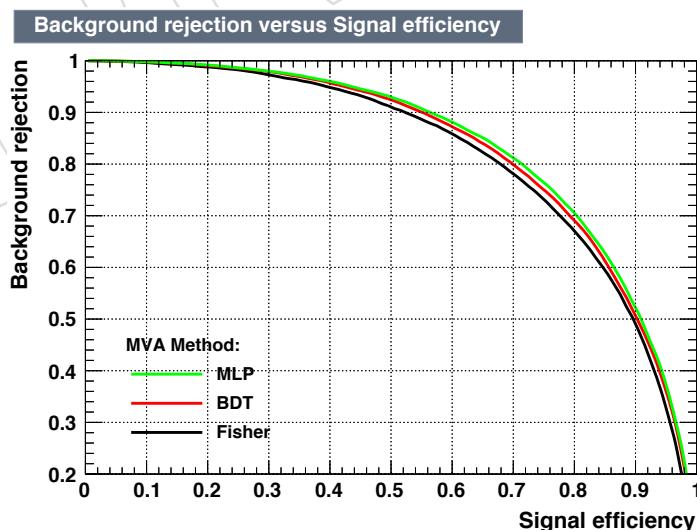


Figure 9: Performance of the multivariate discriminants in the form of a receiver-operator-characteristic curve.

199 **4.4 Parton level**

200 The partonic phase space to which we unfold the measurement is constrained by the kinematic  
 201 requirements of the detector-level fiducial region. Namely, the top and anti-top partons must  
 202 have  $p_T > 400 \text{ GeV}$  and  $|\eta| < 2.4$ , while the invariant mass of the  $t\bar{t}$  system must be greater  
 203 than  $800 \text{ GeV}$  in order to avoid extreme events with high top  $p_T$  and very low  $m_{t\bar{t}}$ . These cuts  
 204 are summarized in Table 6.

Table 6: Definition of parton-level phase space.

Observable	Requirement
$p_T^{t,\bar{t}}$	$> 400 \text{ GeV}$
$ \eta^{t,\bar{t}} $	$< 2.4$
$m_{t\bar{t}}$	$> 800 \text{ GeV}$

205 **4.5 Particle level**

206 The so-called "particle level" represents the state that consists of stable particles originating  
 207 from the proton-proton collision, after the hadronization process, and before the interaction of  
 208 these particles with the detector. The observables computed from the particles' momenta are  
 209 thought to be better defined compared to the ones computed from parton information and ac-  
 210 companied by smaller theoretical uncertainties. Also, the associated phase space is closer to the  
 211 fiducial phase space of the measurement at detector level. In the context of this analysis, parti-  
 212 cle jets are reconstructed from stable particles, excluding neutrinos, with the anti-kt algorithm  
 213 of distance parameter  $R = 0.8$ , identical to the detector-level reconstruction. It should be noted  
 214 that only particles originating from the primary interaction are considered. Subsequently, jets  
 215 that are geometrically matched, within  $\Delta R < 0.4$  in  $\eta - \phi$  from generated leptons (i.e. from the  
 216 leptonic decays of the W boson) are removed from the particle-jet collection. Finally, the two  
 217 particle jets with the highest  $p_T$  are considered the particle-level top-quark candidates. In order  
 218 to match as closely as possible the fiducial phase space, the same kinematic cuts are applied as  
 219 for detector-level events. These requirements are summarized in Table 7.

Table 7: Definition of particle-level phase space.

Observable	Requirement
$N_{\text{jets}}$	$> 1$
$p_T^{\text{jet}1,2}$	$> 400 \text{ GeV}$
$ \eta^{\text{jet}1,2} $	$< 2.4$
$m_{SD}^{\text{jet}1,2}$	(120, 220) GeV
$m_{jj}$	$> 800 \text{ GeV}$

220 In order to verify the sanity of the adopted particle-level definition, in each event that passes  
 221 the requirements above, the top candidates are matched within  $\Delta R < 0.4$  in  $\eta - \phi$  to the original  
 222 top quarks at parton level. Figure 56 shows this efficiency, which varies between 96% and 98%,  
 223 as a function of particle-level top  $p_T$  and  $\eta$ .

224 **5 Signal Extraction**

225 The, by far, dominant background in this analysis is the QCD multijet production, as there is  
 226 a finite probability that ordinary jets, from single parton radiation, will mimic the topological  
 227 substructure of a top-decay jet. With a combination of b-tagging requirements on the subjets

and n-subjettiness variables, we are able to suppress significantly the QCD background. In order to estimate the remaining contribution we employ a data-driven technique based on the assumptions that if the b-tagging requirement is reverted we a) get a pure QCD sample, and b) the jet kinematic properties are not affected.

## 5.1 Inclusive cross section

Before we attempt to extract differential cross sections, we first measure the inclusive, fiducial cross section. This is done in region  $SR_A$  that contains a QCD dominated sideband in the top candidate mass (softDrop mass of the leading jet) and allows a simultaneous fit for the  $t\bar{t}$  signal and the QCD background yields. The fit is described by the equation below:

$$D_{SR_A}(m^t) = N_{t\bar{t}} T(m^t; k_{\text{scale}}, k_{\text{res}}) + N_{\text{qcd}}(1 + k_{\text{slope}} m^t) Q_{CR_A}(m^t) + N_{\text{bkg}} B(m^t), \quad (2)$$

which contains the shapes (templates)  $T(m^t)$ ,  $B(m^t)$  of the signal and the subdominant backgrounds, respectively, taken from the simulation, and the shape  $Q(m^t)$  of QCD taken from the control sample in data ( $CR_A$ ). The templates of the various components are shown in Figs. 10,11. To account for the difference observed in the closure test of QCD for the  $m^t$  variable (Fig. 12) we introduce the linear modification factor  $(1 + k_{\text{slope}} m^t)$ , inspired by the simulation, but with the slope parameter  $k_{\text{slope}}$  left free in the fit. Also free in the fit are the normalization factors  $N_{t\bar{t}}$ ,  $N_{\text{qcd}}$ , and  $N_{\text{bkg}}$ . Finally, we introduce two more nuisance parameters in the  $t\bar{t}$  simulation,  $k_{\text{scale}}$  and  $k_{\text{res}}$  that account for possible differences between data and simulation in the scale and resolution of the  $m^t$  parameter. The fit model is imported to the RooFit package and the fit result is shown in Fig. 13 while the fitted parameters are summarized in Table 8. We observe that the fitted  $t\bar{t}$  yield (6238) is significantly lower than the expectation (9604), which implies that the fiducial cross section is  $\approx 35\%$  lower compared to the Powheg+Pythia8 prediction. The nuisance parameters related to the  $m^t$  scale and resolution are consistent with one and the slope of the QCD modification factor  $(5.7 \pm 1.4) \times 10^{-3}$  is very close to the value from the QCD simulation  $((5.4 \pm 1.2) \times 10^{-3})$ . Finally, despite the large uncertainty, the yield of the subdominant backgrounds  $(400 \pm 247)$  is very close to the Monte Carlo prediction (380) in Table 9.

The fitted  $t\bar{t}$  yield can be converted to a fiducial cross section  $\sigma_{\text{fid}} = \frac{N_{t\bar{t}}}{\mathcal{L}}$ . For an estimate of the systematic uncertainties, we repeat the measurement, independently, for each of the sources of uncertainty described in section 7. For each variation, a new  $t\bar{t}$  template is introduced and the fit is repeated. The independent results are shown in Fig. 14 and the resulting cross section is  $\sigma = 174.7 \pm 5.1(\text{fit})^{+3.3}_{-5.1}(\text{syst.}) \pm 4.2(\text{lumi.})$ . The measured cross section is found to be a factor  $r = 0.635$  lower than the theoretical prediction (Powheg+Pythia8). This factor is used to scale down the  $t\bar{t}$  contribution in the data vs simulation comparisons.

Table 8: Results of the fit in  $SR_A$ .

Parameter	Value	Error
$k_{\text{res}}$	0.960	0.026
$k_{\text{scale}}$	1.002	0.002
$k_{\text{slope}}$	5.7e-03	1.4e-03
$N_{\text{bkg}}$	400	255
$N_{\text{qcd}}$	4539	247
$N_{t\bar{t}}$	6238	181

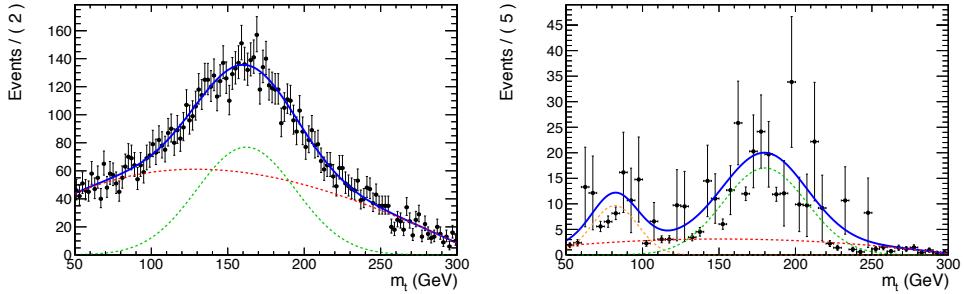


Figure 10: Templates of QCD (left), taken from data, and of the subdominant backgrounds (right), taken from the simulation. The different lines show the individual components (frozen in the fit) used to describe the shapes. The QCD shape is composed of a smooth polynomial and a Gaussian, while the shape of the subdominant backgrounds contains a smooth polynomial and two Gaussians (one describes the W resonance from the single top and WJets processes and the other describes the broader peak from the kinematic selections).

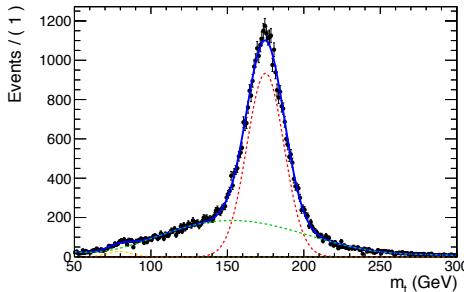


Figure 11: Template of the  $t\bar{t}$  signal taken from the simulation. The different lines show the individual components (frozen in the fit) used to describe the shape. The shape consists of a smooth polynomial and two Gaussians (one describes the W resonance from unmerged top decays and the other describes the fully merged top resonance).

## 261 5.2 Differential cross sections

262 The  $t\bar{t}$  cross section is also reported differentially, in the signal region  $SR$  as a function of several  
 263 variables: the leading and second top  $p_T$  and  $|y|$ , and the  $t\bar{t}$  mass,  $p_T$ , and rapidity. For the  
 264 derivation of the differential cross sections, the background contributions are subtracted from  
 265 the data:

$$S(x) = D(x) - R_{yield} N_{qcd} Q(x) - B(x), \quad (3)$$

266 where  $x = p_T^{t1,2}, |y^{t1,2}|, m^{t\bar{t}}, p_T^{t\bar{t}}, y^{t\bar{t}}$ ,  $S(x)$  is the signal,  $D(x)$  is the measured distribution in  
 267 data,  $Q(x)$  is the QCD shape,  $B(x)$  is the subdominant backgrounds' contribution (both the  
 268 shape and the normalization are taken from the simulation),  $R_{yield}$  is the yield ratio between  
 269 the signal region ( $SR$ ) and the fit region ( $SR_A$ ), and  $N_{qcd}$  is the fitted number of QCD events in  
 270  $SR_A$ .

271 The two elements that need to be determined from data are the shape  $Q(x)$  of the QCD back-  
 272 ground as a function of an observable of interest  $x$ , and the absolute normalization  $N_Q$ . The  
 273 QCD template  $Q(x)$  is taken from the QCD control sample by applying the signal region selec-  
 274 tion with reverted b-tagging requirement (none of the leading AK8 jets is allowed to contain  
 275 a b-tagged subjet). Figures 15, 16, 17 show the closure test in MC as a function of jet and di-

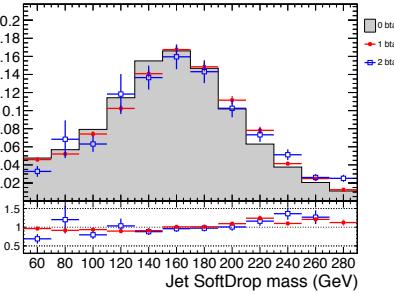


Figure 12: Closure test in the QCD simulation for the shape of the  $m^t$  variable in the three possible b-tagging requirements (none, exactly one, or both jets contain a b-tagged subjet).

276 jet kinematic variables. We observe that, within the statistical precision of the simulation, the  
 277 shapes are compatible, regardless of the b-tagging requirement. Moreover, Figs. 19,20 show  
 278 that indeed the 0-btag sample has negligible  $t\bar{t}$  contamination.

279 The normalization of the QCD background is determined from the fit to the data in  $SR_A$ . How-  
 280 ever, since this is an extended signal region, a transfer factor  $R_{yield} = N^{SR}/N^{SR_A}$  is needed in  
 281 order to get the QCD normalization in the signal region. Figure 18 shows a closure test per-  
 282 formed in QCD Monte Carlo events, where the value of  $R_{yield}$  is independent of the b-tagging  
 283 requirement (within the statistical precision of the simulation), and therefore can be taken from  
 284 the QCD control sample ( $R_{yield} = N^{CR}/N^{CR_A}$ ).

DRAFT

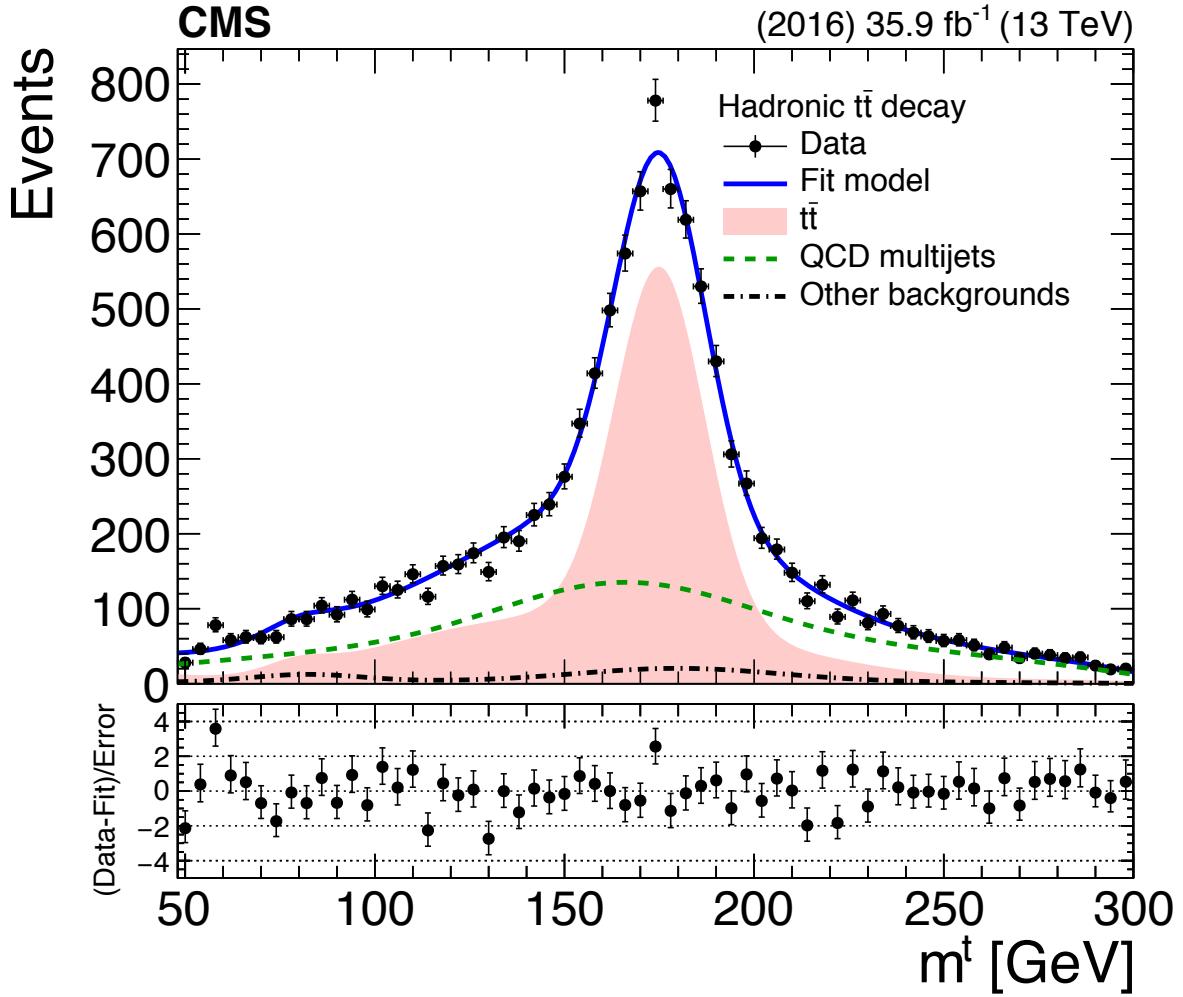


Figure 13: Result of the template fit on data. The red line shows the  $t\bar{t}$  contribution, the green line shows the QCD, and the brown line shows the subdominant backgrounds.

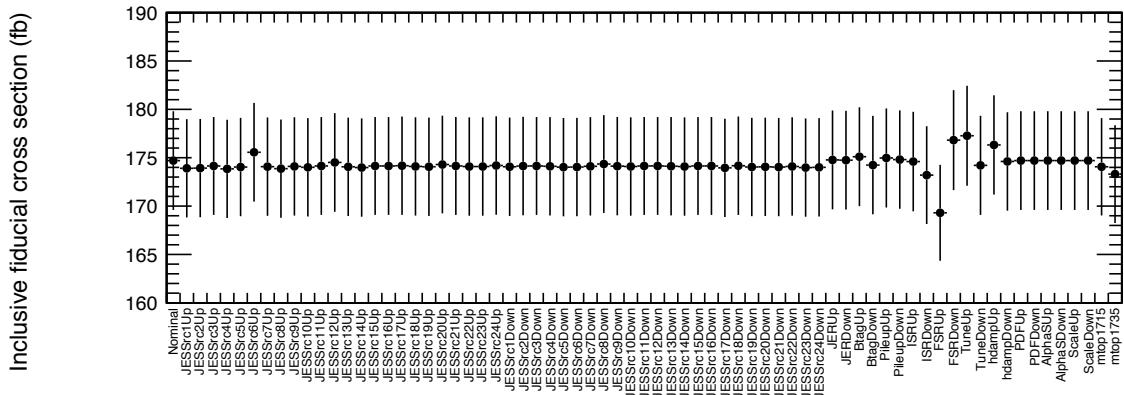


Figure 14: Measurements of the inclusive, fiducial  $t\bar{t}$  cross section for the nominal signal simulation and all the systematic variations.

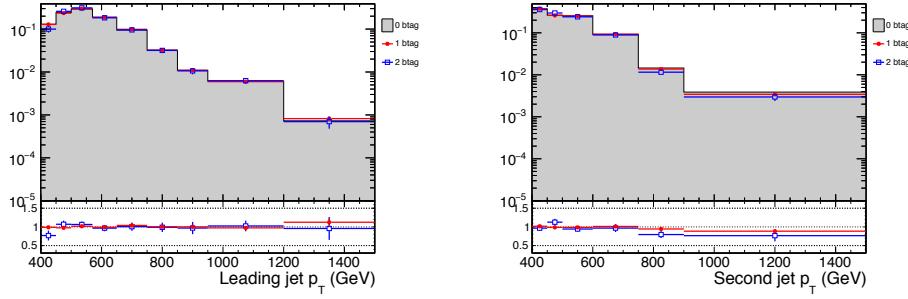


Figure 15: Closure test for the data-driven QCD prediction method for the shape of the top  $p_T$  variables.

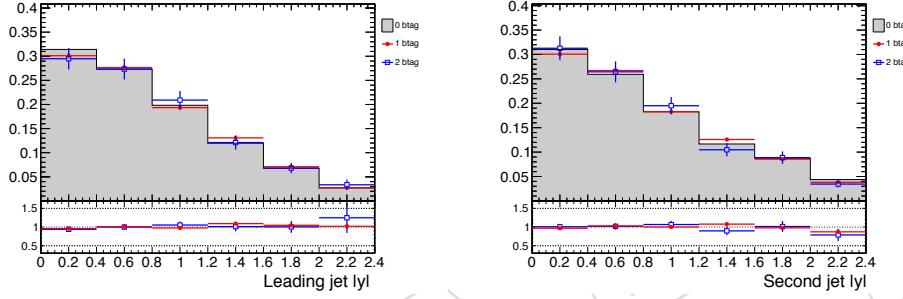


Figure 16: Closure test for the data-driven QCD prediction method for the shape of the top  $|y|$  variables.

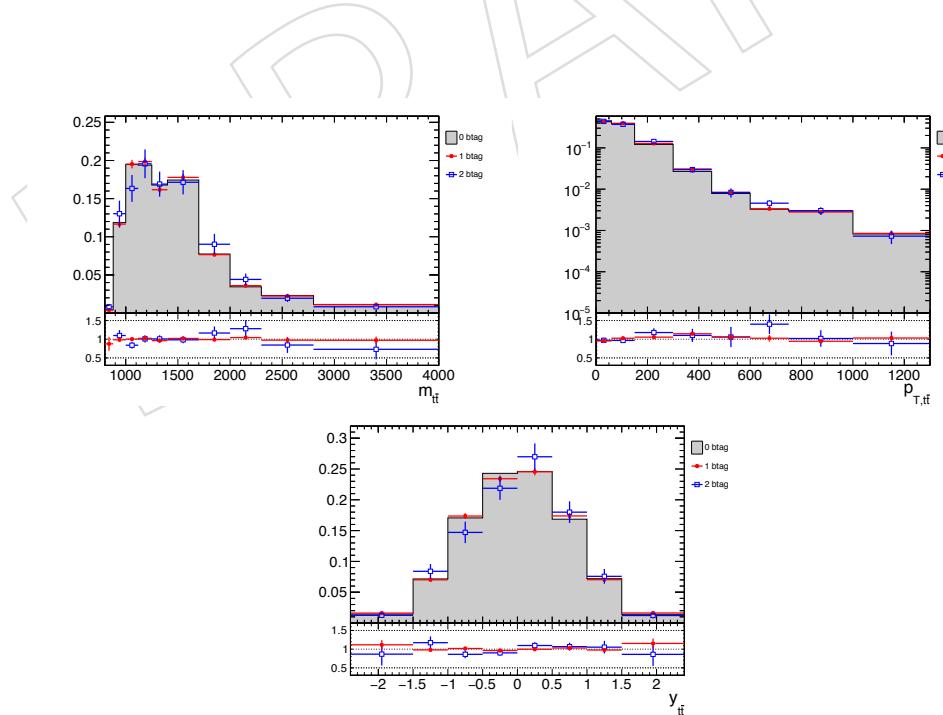


Figure 17: Closure test for the data-driven QCD prediction method for the shape of the  $t\bar{t}$  variables.

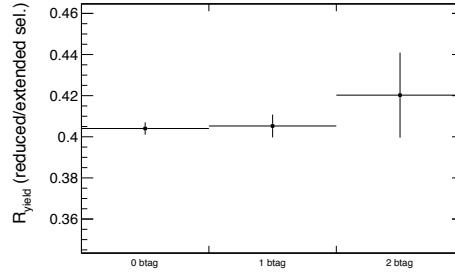


Figure 18: Closure test for the data-driven QCD prediction method for event yield ratio between the control and signal regions.

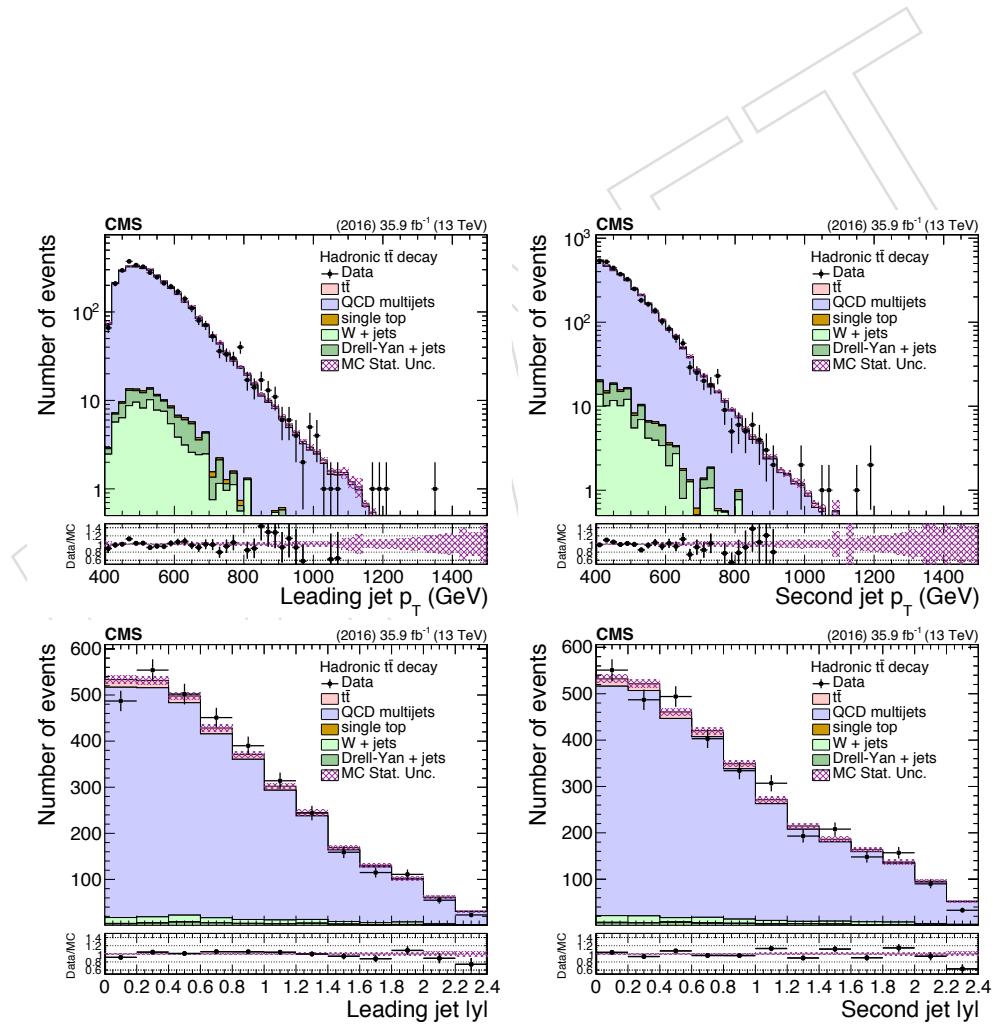


Figure 19: Jet  $p_T$  and  $|y|$  distributions in the QCD control region.

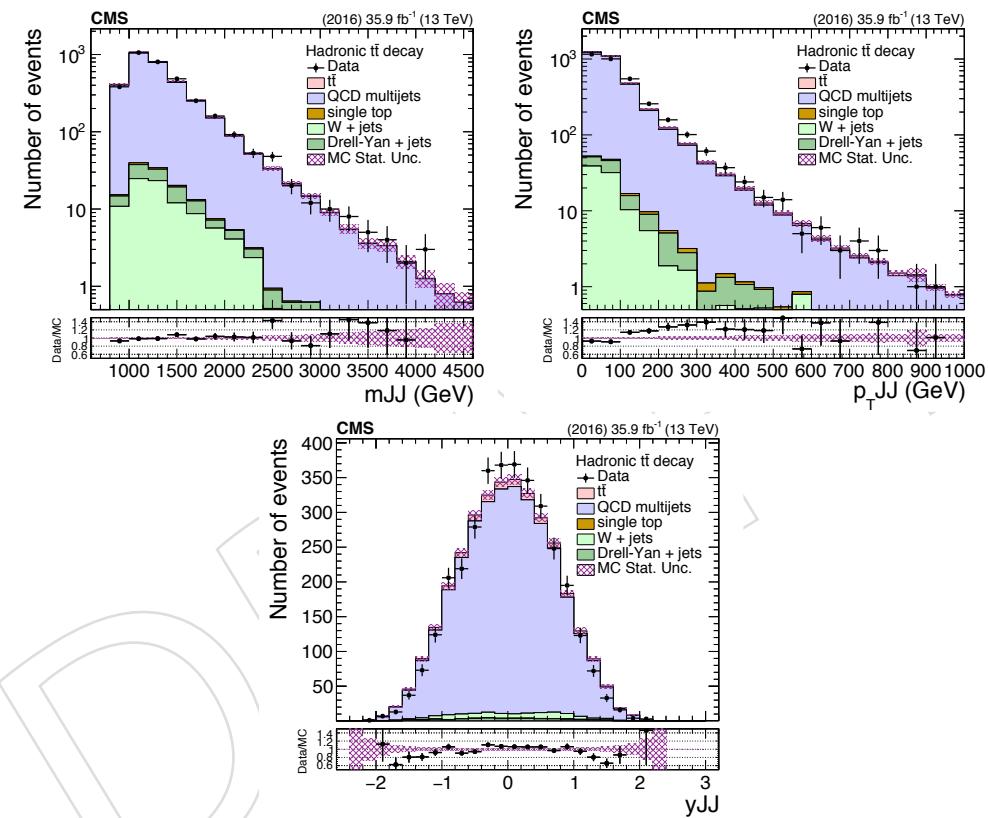


Figure 20: Kinematic distributions of the  $t\bar{t}$  system in the QCD control region.

## 285 6 Data vs Monte Carlo

286 In this section we present comparisons between data and Monte Carlo for various observables.  
 287 It should be noted that the  $t\bar{t}$  normalization is scaled by a factor 0.635, consistent with the  
 288 findings of Section 5, and the QCD background yield is adjusted such that the total Monte Carlo  
 289 events are equal to the events in data. Table 9 shows the event yields for the various processes  
 290 after applying the baseline selection plus the requirement that both AK8 jets contain a b-tagged  
 291 subjet, with or without the NN cut (see Section 4 for details on the selection requirements).  
 292 Already after the b-tagging requirement the  $t\bar{t}$  signal is sufficiently enhanced that it becomes  
 293 visible. Finally, the NN cut suppresses QCD by a factor  $\sim 17$  with a signal loss of  $\approx 45\%$ . In  
 294 the signal region the signal-over-background ratio is  $\approx 1.8$ .

Table 9: Expected and observed event yields in  $SR_A$ .

Process	No NN cut	With NN cut
$t\bar{t}$	10881	5867
QCD	85155	4930
W+jets	857	110
Z+jets	784	59
Single Top	753	211
Data	98430	11177

Table 10: Expected and observed event yields in the signal region.

Process	Yield
$t\bar{t}$	3978
QCD	2171
W+jets	51
Z+jets	12
Single Top	83
Data	6295

295 Figure 21 shows the softDrop mass of the leading jet, which serves as a proxy for the top  
 296 mass. This is defined as the leading AK4 subjet in the leading AK8 jet and it is selected if  
 297 the event satisfies the signal region selection and if it has  $p_T > 300 \text{ GeV}$  while it is not b-tagged  
 298 ( $\text{CSV} < 0.8484$ ). The distribution of the NN output is shown in Fig. 22.  
 299 The kinematic properties of jets in the signal region are shown in Figs. 23, 24, while those of the  
 300  $t\bar{t}$  system are shown in Fig. 25. Finally, the jet substructure properties (n-subjettiness variables)  
 301 are shown in Figs. 26, 27.

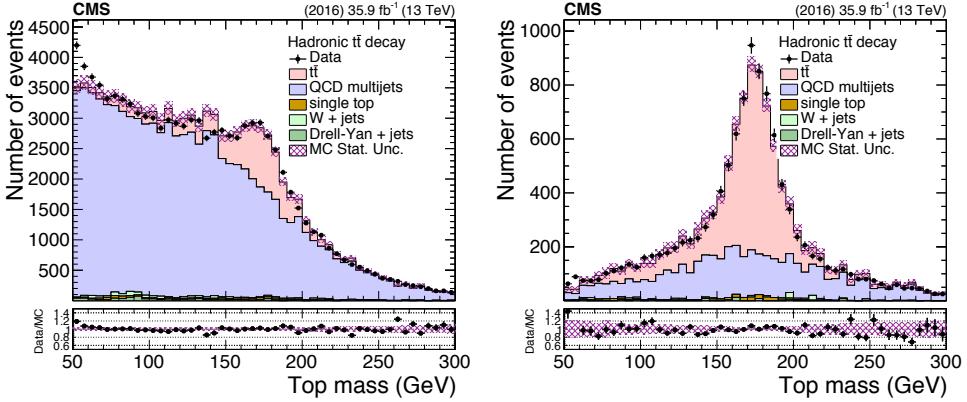


Figure 21: SoftDrop mass of the leading AK8 jet after the baseline selection with both AK8 jets containing a b-tagged subjets, with (right) and without (left) the  $NN > 0.8$  cut.

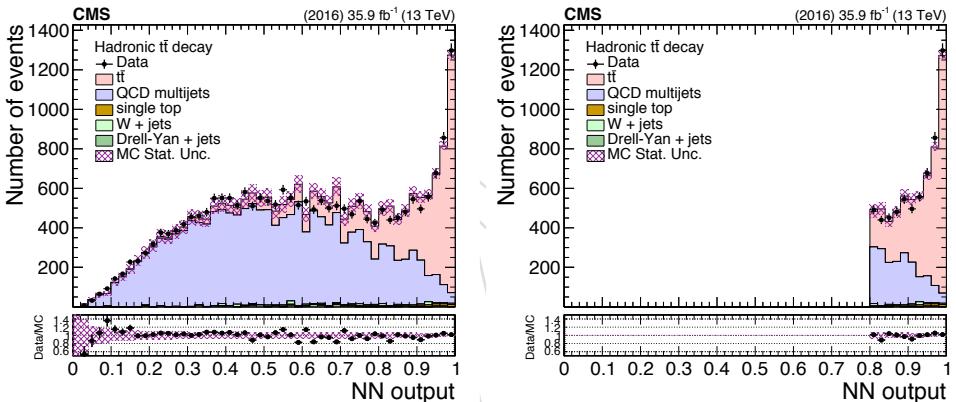


Figure 22: NN output after the baseline selection with both AK8 jets containing a b-tagged subjets, with (right) and without (left) the  $NN > 0.8$  cut.

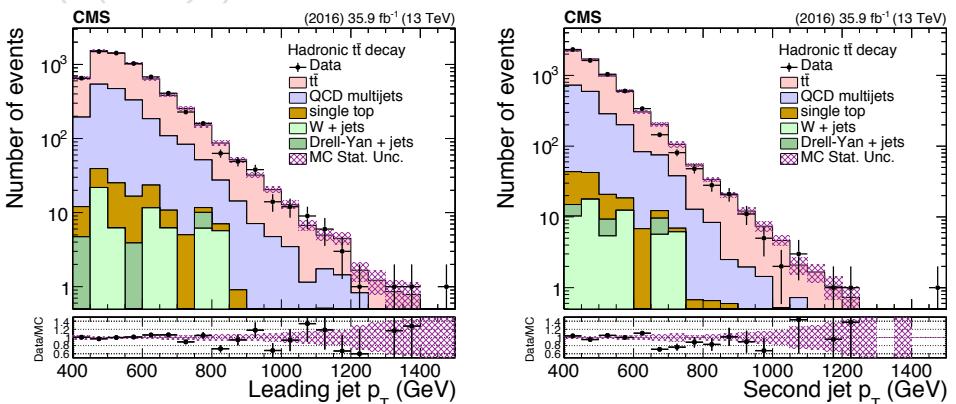


Figure 23: Distributions of the leading and second jet  $p_T$  in the signal region.

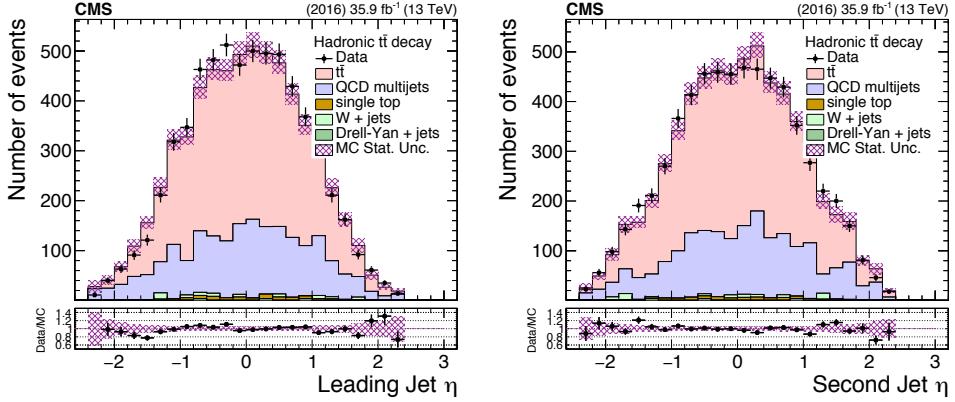


Figure 24: Distributions of the leading and second jet  $\eta$  in the signal region.

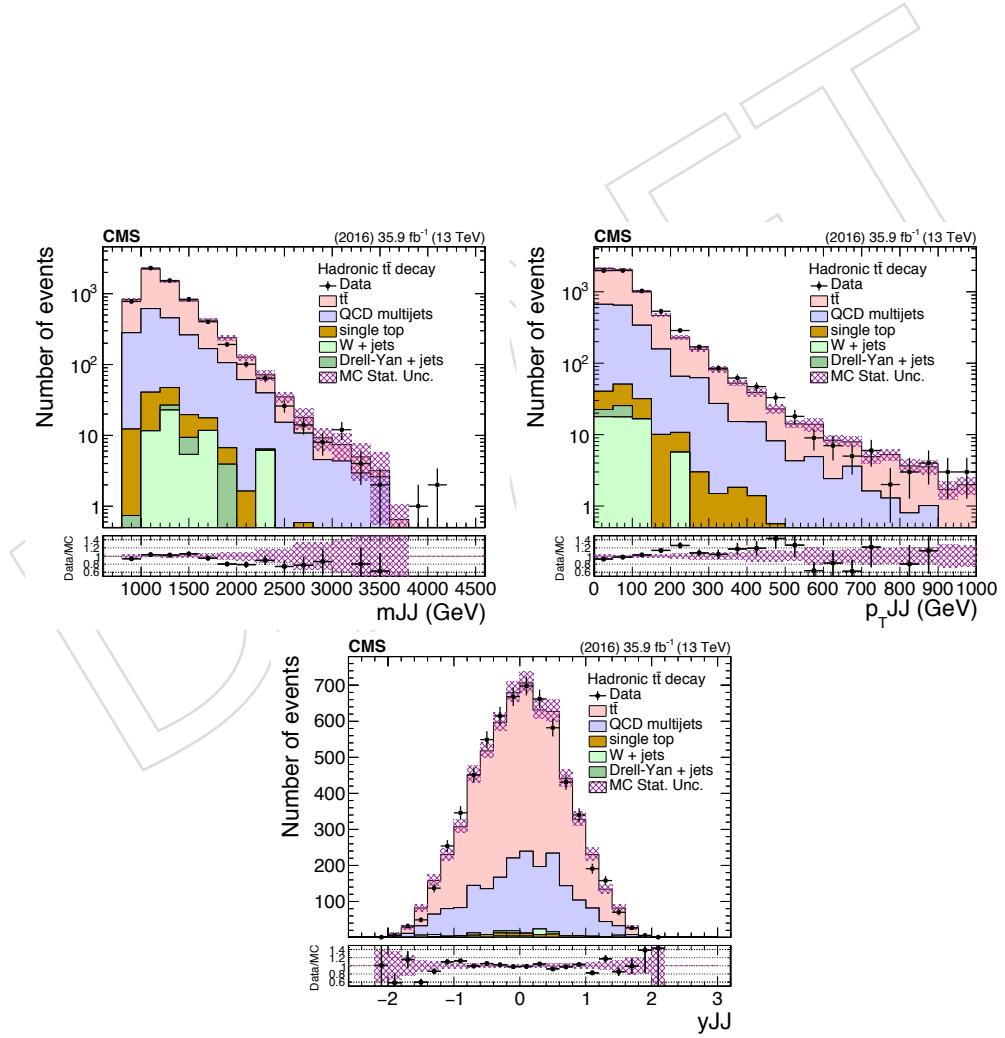


Figure 25: Kinematic distributions of the  $t\bar{t}$  system in the signal region.

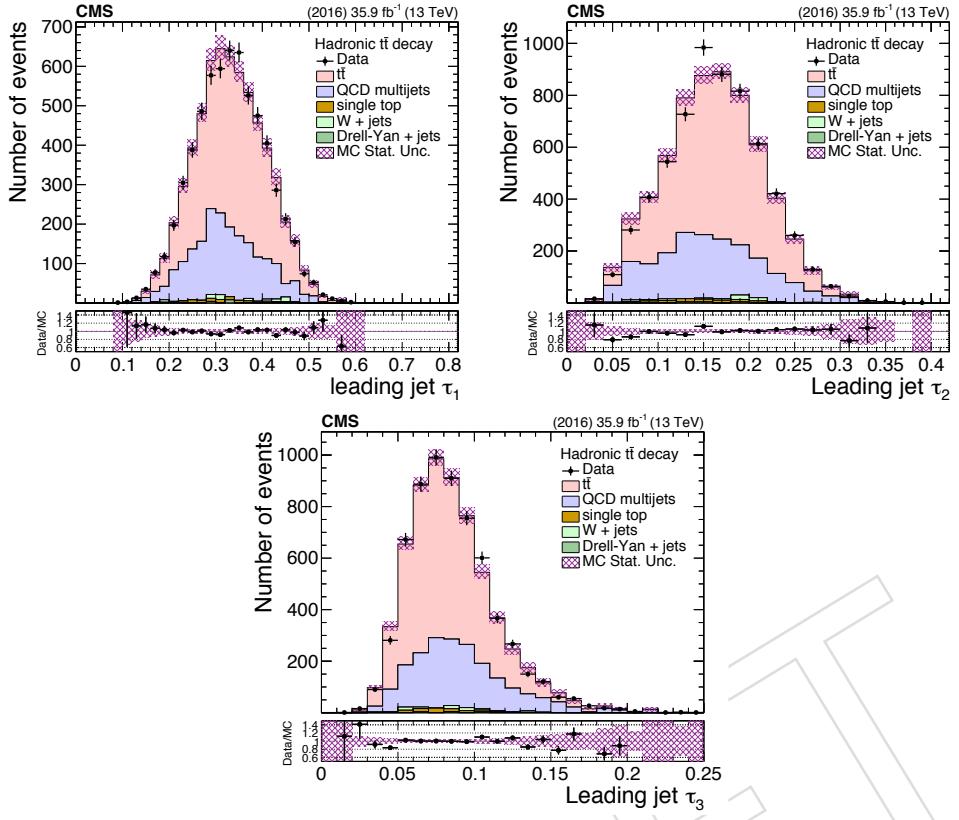


Figure 26: N-subjettiness variables for the leading jet in the signal region.

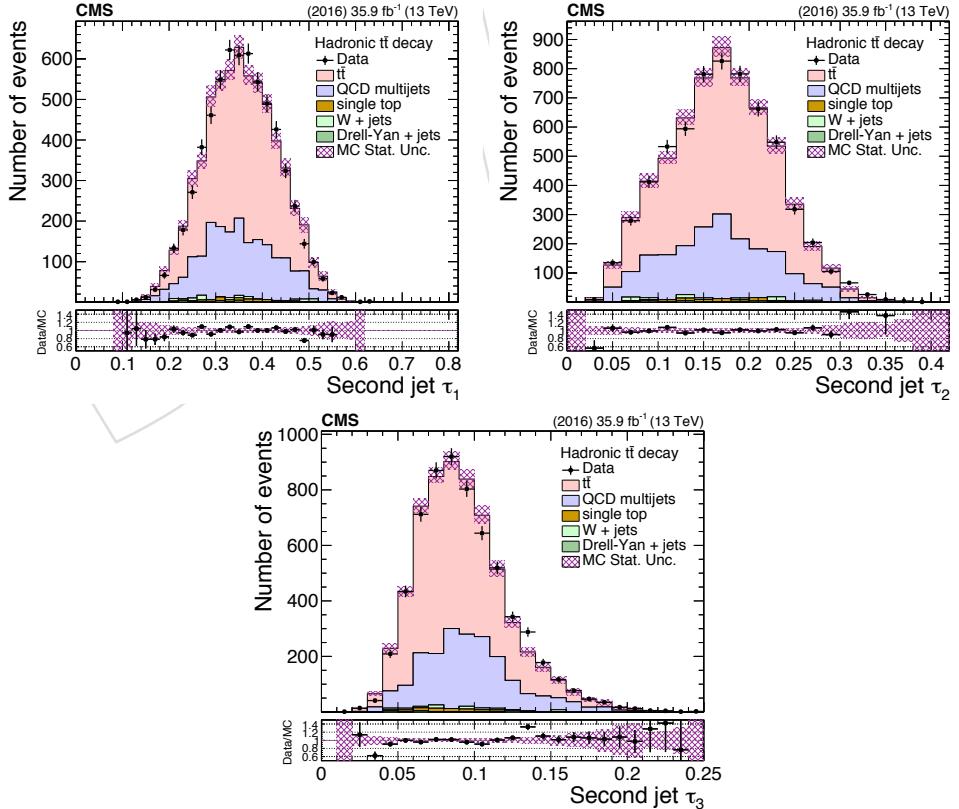


Figure 27: N-subjettiness variables for the second jet in the signal region.

## 302 7 Systematic Uncertainties

303 The systematic uncertainties considered in this analysis are devided in two categories: experimental and theoretical. The former includes all the uncertainties related to the differences in  
 304 the object performance between data and simulation. The latter are related to the simulation it-  
 305 self and affect primarily the unfolded results through the acceptance, efficiency, and migration  
 306 matrix. The list below describes briefly these uncertainties and the way they have been han-  
 307 dled here. It should be noted that for each systematic variation the differential cross sections  
 308 (fiducial and unfolded) are re-measured and the difference with respect to the nominal result  
 309 is taken as the effect of this variation to the measurement.

### 311 1. Experimental Uncertainties

- 312 • QCD background prediction: we use the fitted QCD yield uncertainty. The  
 313 shape uncertainties due to the closure test in the simulation and the different  
 314 pileup profiles in the control and signal regions are very small, at the 1% level.
- 315 • Jet energy scale (JES): this is the uncertainty on the energy scale of each recon-  
 316 structed jet and it is a leading experimental uncertainty. Following the rec-  
 317 commendations of the JME group, we have considered 24 independent JES  
 318 sources [41] as follows: for each variation a new jet collection is created and  
 319 the event interpretation is repeated. This results not only in variations of the  
 320  $p_T$  scale itself, but may also lead to different top candidates. The JES uncer-  
 321 tainty, per jet, is of the order 1 – 2%,  $p_T$  and  $\eta$  dependent. The effect on the  
 322 measured cross section is typically of the order 10% but it can be much larger  
 323 at very high jet  $p_T$ .
- 324 • Jet energy resolution: The impact on the measurement due to the jet energy res-  
 325 olution (JER) is determined by smearing the jets according to the JER uncer-  
 326 tainty. The effect on the cross section is relatively small, at the level of 2%.
- 327 • Subjet b tagging efficiency (hadronic): The uncertainty in the identification of b-  
 328 subjets within the large- $R$  jets (estimated in [42]) is a leading experimental  
 329 uncertainty in the hadronic channel. The effect on the cross sections is of the  
 330 order of 10% relatively flat in all the observables. Unlike the uncertainty asso-  
 331 ciated with the JES, the b-subjet tagging uncertainty therefore largely cancels  
 332 in the normalized cross sections.
- 333 • Pileup: The uncertainty related to the modeling of additional pileup interac-  
 334 tions is a subdominant uncertainty. The impact on the measurement is esti-  
 335 mated by varying the total inelastic cross section used to weight the simulated  
 336 events (69.2 mb) by  $\pm 4.6\%$  [43]. The effect on the cross sections is negligible  
 337 (below 1%).
- 338 • Trigger: this accounts for the difference between the simulated and observed  
 339 trigger efficiency. Based on Fig. 2 the uncertainty is well below 1% in the phase  
 340 space of this analysis.
- 341 • Luminosity: The uncertainty in the measurement of the integrated luminosity  
 342 is 2.5% [44].

### 343 2. Theoretical Uncertainties

344 The theoretical uncertainties are divided into two sub-categories: the ones related to the  
 345 matrix element of the hard process and the ones related to the modelling of the parton  
 346 shower and the underlying event. Practically, the first category (consisting of the first

347 three sources below) is evaluated by variations of LHE event weights stored in the nominal  
 348 MC simulation, while the second category is evaluated with dedicated, alternative  
 349 MC samples.

- 350 • **Parton distribution functions:** The uncertainty due to parton distribution func-  
 351 tions (PDFs) is estimated by applying event weights corresponding to the 100  
 352 replicas of the NNPDF set [31]. For each observable we compute its standard  
 353 deviation from the 100 variations.
- 354 • **Renormalization and factorization scales:** The uncertainty in the choice of renor-  
 355 malization and factorization scale is assessed by separately varying the renor-  
 356 malization and factorization scales by a factor of [2.0,0.5]. The unphysical anti-  
 357 correlated variations are discarded, yielding a total of 7 combinations of the  
 358 renormalization and factorization scales. The event weight is determined for  
 359 each variation and the envelope is taken as the scale uncertainty, according to  
 360 the prescription in [<https://indico.cern.ch/event/459797/contributions/1961581/attachments/1181555/1800214/mcaod-Feb15-2016.pdf>]. These variations are applied using the event weights provided in the  
 361 LHEEventInfo [<https://twiki.cern.ch/twiki/bin/viewauth/CMS/LHReaderCMSSW>].
- 363 • **Strong coupling constant ( $\alpha_S$ ):** The uncertainty associated with the  $\alpha_S$  is esti-  
 364 mated by applying event weights corresponding to higher and lower values  
 365 of  $\alpha_S$  for the matrix element using the variations of the NNPDF set [31].
- 366 • **Final state radiation (FSR):** this uncertainty is estimated from alternative MC  
 367 samples with reduced and increased value for the strong coupling constant  
 368 used by Pythia8 to generate final state radiation ( $\alpha_S^{FSR}$  by factors  $\sqrt{2}$  and  $1/\sqrt{2}$ ).  
 369 However, the default variations are largely exaggerated, as seen in Fig. 28,  
 370 where the alternative samples fail miserably to describe the data in two sensi-  
 371 tive observables: the mass of the top candidate jet and the NN output, both of  
 372 which rely on the dynamics of the jet formation. In order to use a more realistic  
 373 estimate of this uncertainty, a maximum likelihood fit is performed (indepen-  
 374 dently) on these variables, using the systematic uncertainties as independent  
 375 nuisance parameters. The fits are implemented through the “Combine” tool-  
 376 box using templates for the various components and are shown in Fig. 29. The  
 377 fit on the  $m_t$  variable is well behaved in the entire fit range, while the fit on the  
 378 NN output is less successful in the background-enriched region. Nevertheless,  
 379 since it is the signal uncertainties that we wish to constrain, the signal region is  
 380 fitted perfectly. As a result of these fits, the nuisances are constrained in-situ.  
 381 Figures 30 and 31 show the comparison of the pre-fit and post-fit uncertainties,  
 382 as well as the pulls of the nuisance parameters. Both fits indicate that the post-  
 383 fit FSR nuisance is actually close to its pre-fit (nominal) value with a significant  
 384 reduction of its uncertainty. Based on the observations from the fits, we have  
 385 scaled down the effects of the FSR variations by a (conservative) factor 0.3.
- 386 • **Initial state radiation (ISR):** this uncertainty is estimated from alternative MC  
 387 samples with reduced and increased value for the strong coupling constant  
 388 used by Pythia8 to generate initial state radiation ( $\alpha_S^{ISR}$  scaled by factors of 2  
 389 and 0.5).
- 390 • **Matrix element – parton shower matching:** In the POWHEG matrix element to par-  
 391 ton shower (ME-PS) matching scheme, the resummation damping factor  $h_{damp}$   
 392 is used to regulate high- $p_T$  radiation. Uncertainties in  $h_{damp}$  are parameterized  
 393 by considering alternative simulated samples with  $h_{damp}$  varied by  $h_{damp} = m_t$   
 394 and  $h_{damp} = 2.24m_t$ .

- 395     • *Underlying event tune:* This uncertainty is estimated from alternative Monte  
 396     Carlo samples with the tune CUETP8M2T4 parameters varied by  $\pm 1\sigma$ .

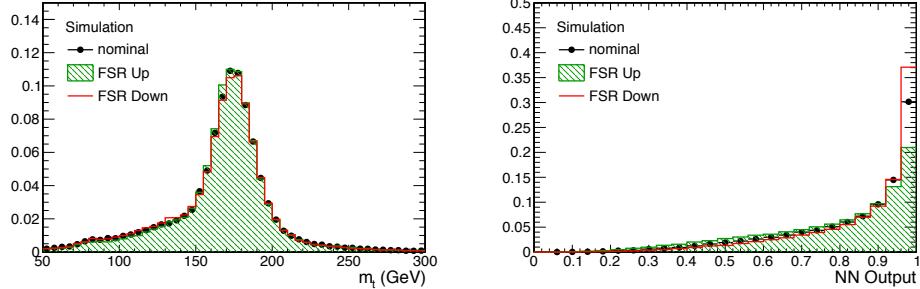


Figure 28: Variations of the  $m^t$  (left) and NN output (right) distributions due to the alternative FSR MC samples.

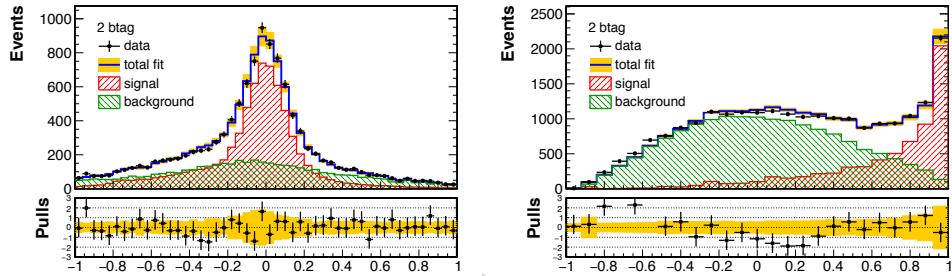


Figure 29: Fit to the  $m^t$  (left) and NN output (right) distributions aiming to constrain the MC modelling uncertainties.

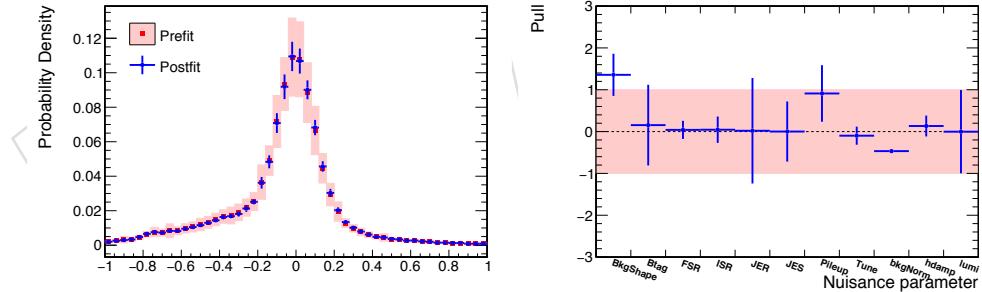


Figure 30: (left) Comparison between the prefit and postfit distributions and uncertainties of the  $m^t$  distribution. (right) Fit pulls of the nuisance parameters.

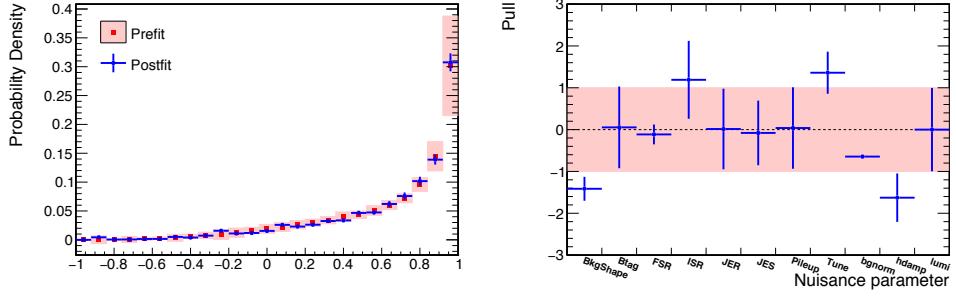


Figure 31: (left) Comparison between the prefit and postfit distributions and uncertainties of the NN output distribution. (right) Fit pulls of the nuisance parameters.

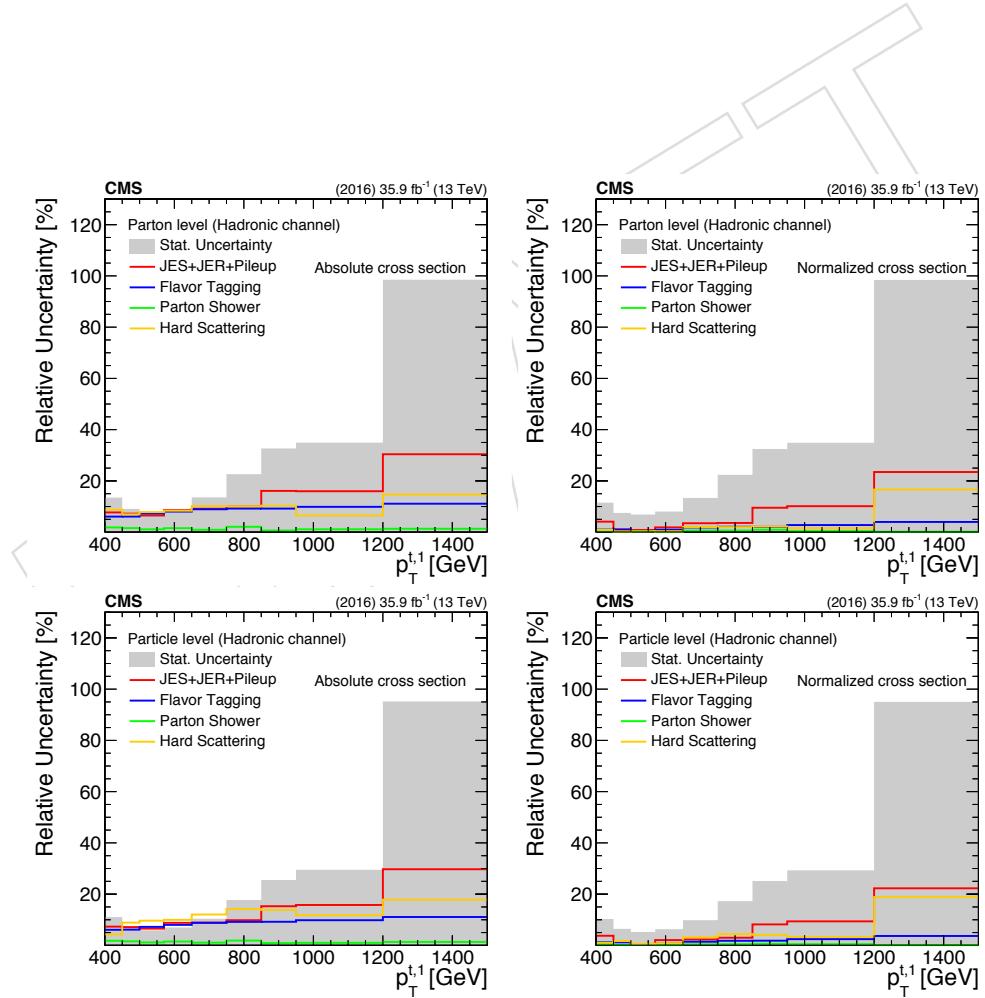


Figure 32: Decomposition of uncertainties for the parton- and particle-level measurement (left: absolute, right: normalized) as a function of the leading top  $p_T$ .

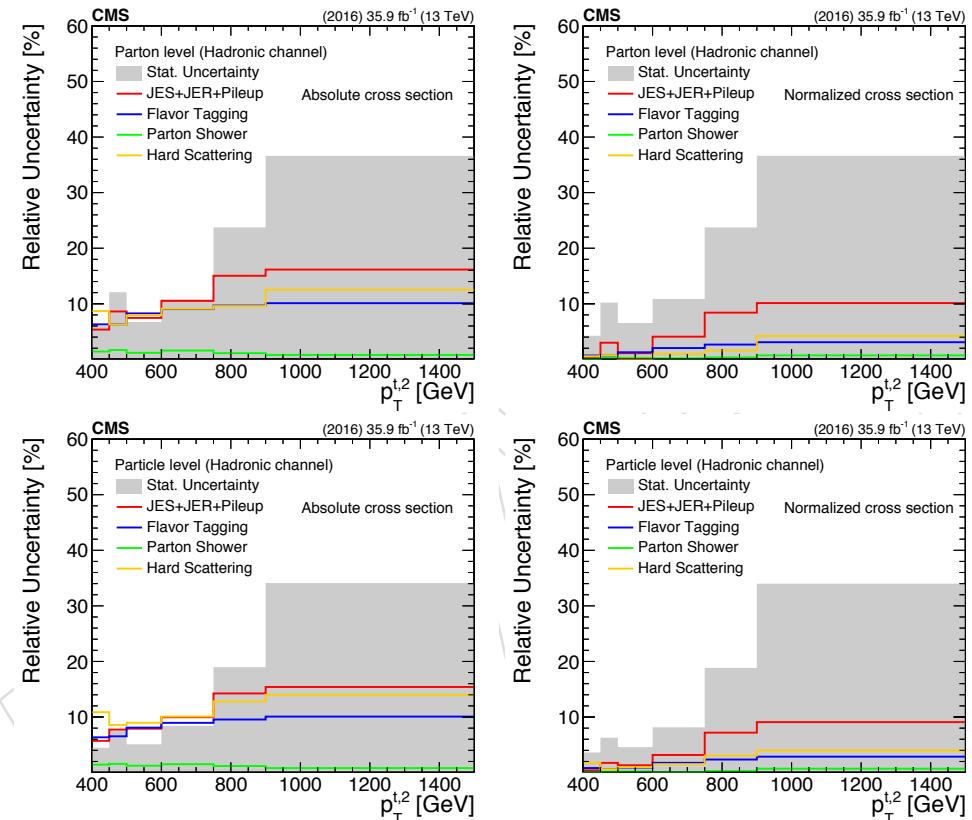


Figure 33: Decomposition of uncertainties for the parton- and particle-level measurement (left: absolute, right: normalized) as a function of the second top  $p_T$ .

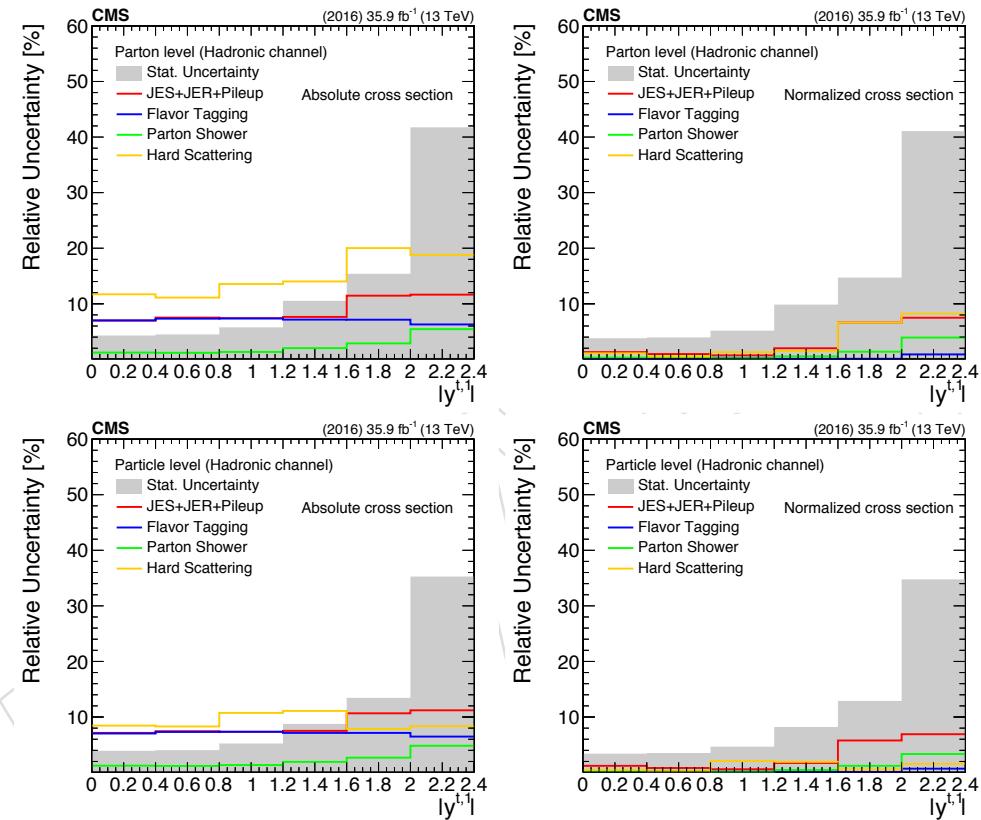


Figure 34: Decomposition of uncertainties for the parton- and particle-level measurement (left: absolute, right: normalized) as a function of the leading top  $|y|$ .

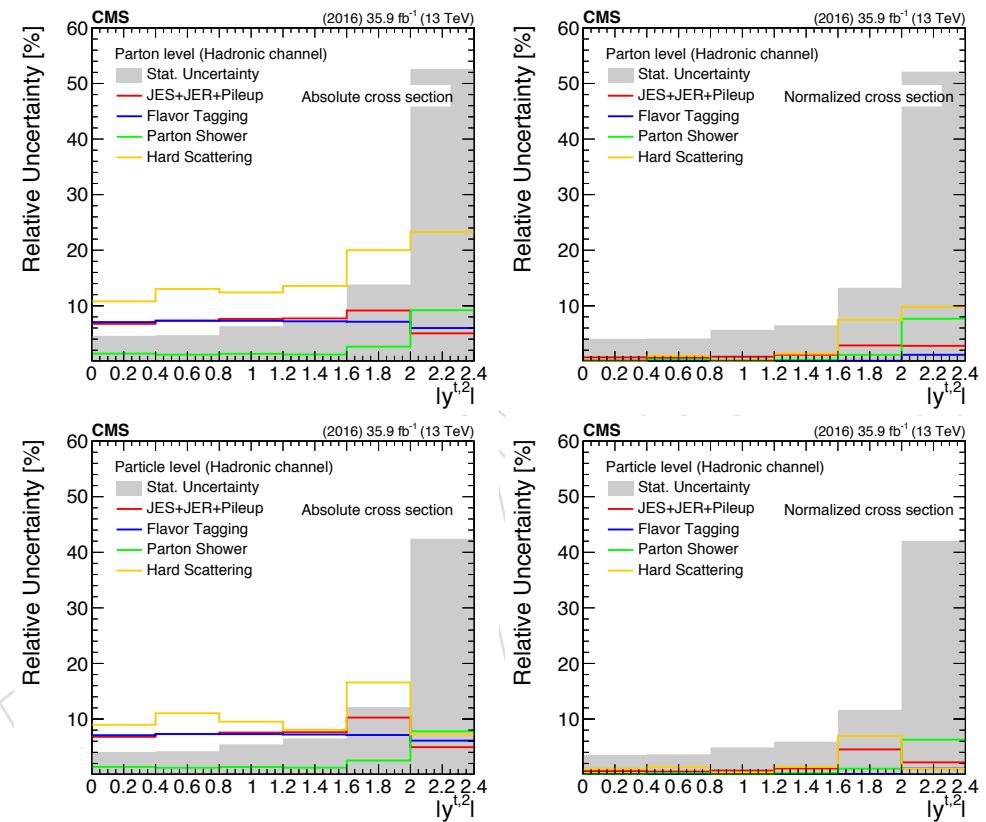


Figure 35: Decomposition of uncertainties for the parton- and particle-level measurement (left: absolute, right: normalized) as a function of the second top  $|y|$ .

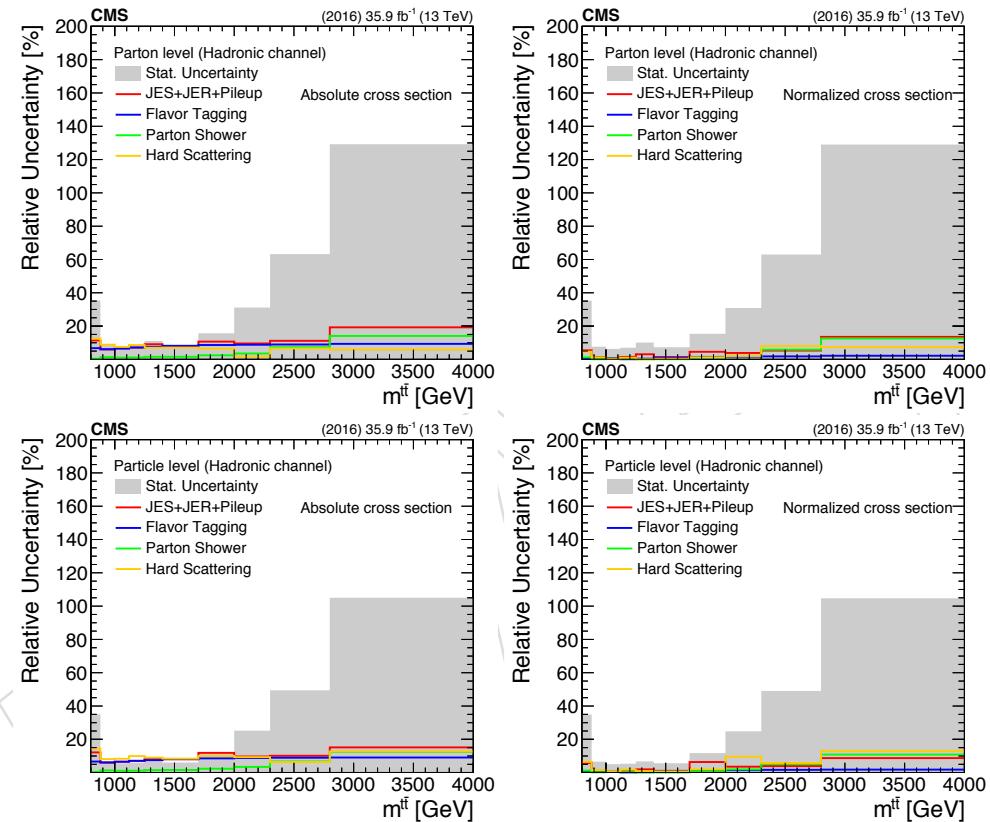


Figure 36: Decomposition of uncertainties for the parton- and particle-level measurement (left: absolute, right: normalized) as a function of  $m_{t\bar{t}}$ .

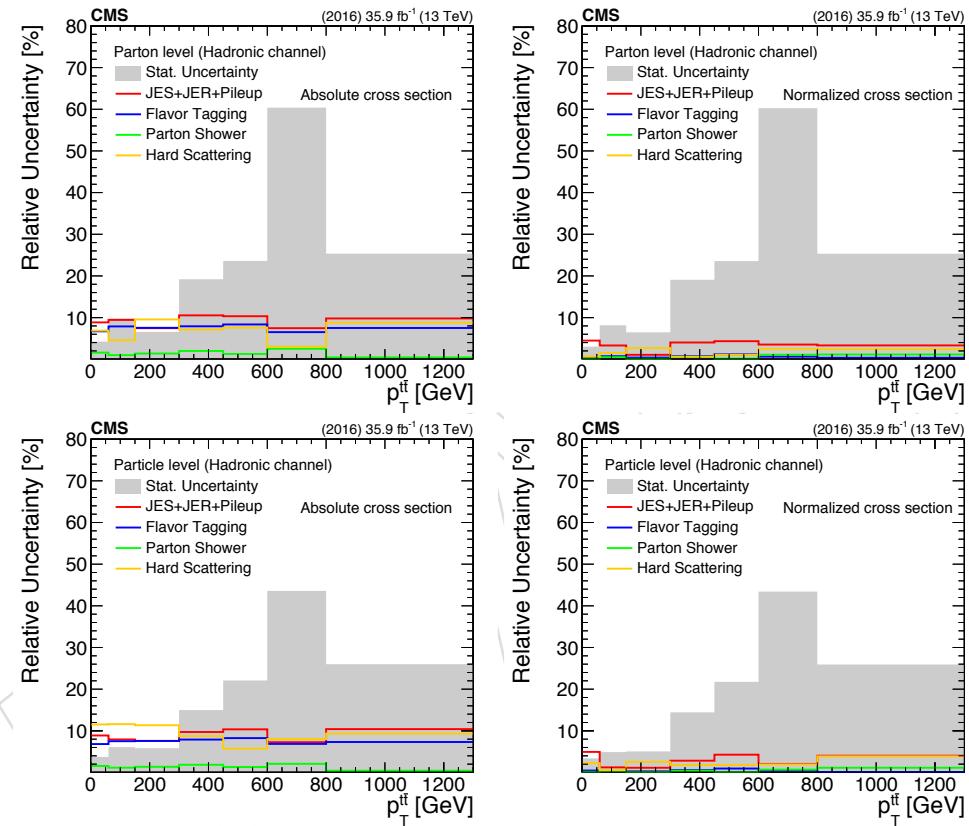


Figure 37: Decomposition of uncertainties for the parton- and particle-level measurement (left: absolute, right: normalized) as a function of  $p_T^t$ .

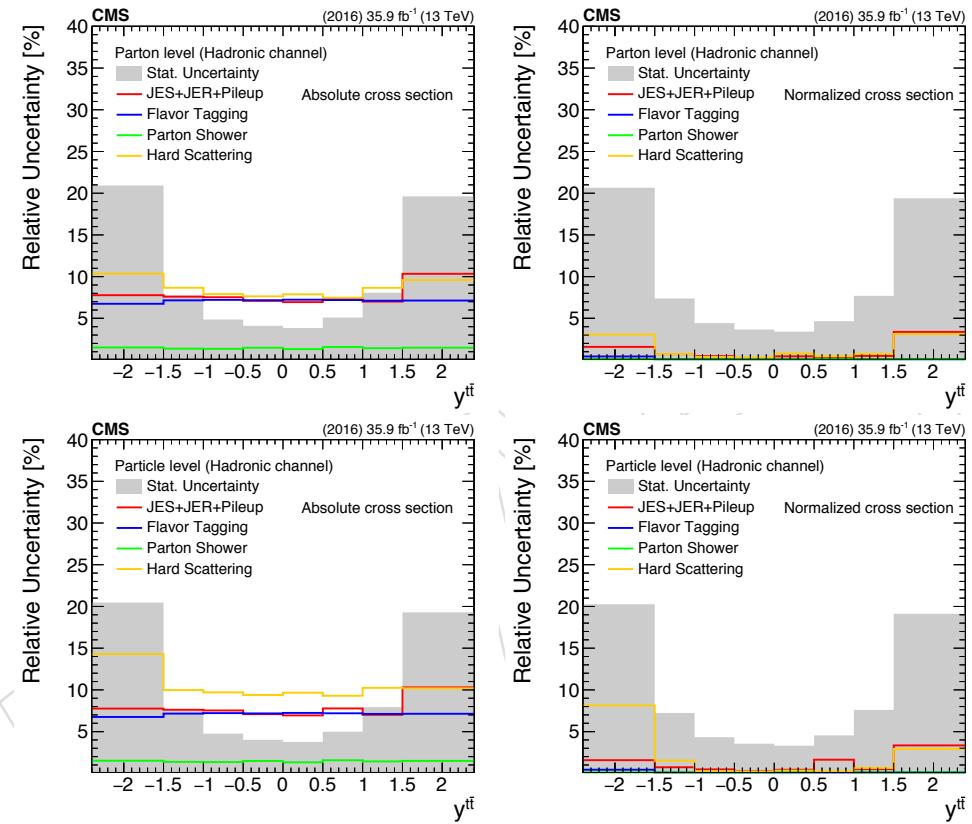


Figure 38: Decomposition of uncertainties for the parton- and particle-level measurement (left: absolute, right: normalized) as a function of  $y_{t\bar{t}}$ .

## 397 8 Fiducial Measurement

398 The fiducial differential cross section is derived in bin  $i$  of the variable  $x$  from the signal yield  
 399  $S_i$  (eq. 3) in the bin as follows:

$$\frac{d\sigma_i^{\text{fid}}}{dx} = \frac{S_i}{\mathcal{L} \cdot \Delta x_i}, \quad (4)$$

400 where  $\mathcal{L}$  is the total integrated luminosity and  $\Delta x_i$  is the width of the  $i$ -th bin of the observable  
 401  $x$ . Of particular interest is also the **normalized cross section**, computed as:

$$\frac{1}{\sigma^{\text{fid}}} \frac{d\sigma_i^{\text{fid}}}{dx} = \frac{1}{\sum_k S_k} \cdot \frac{S_i}{\mathcal{L} \cdot \Delta x_i}, \quad (5)$$

402 which is used to confront the modelling of the differential cross section regardless of the overall  
 403 normalization. In order to estimate the **uncertainty** on the **measurement**, the entire **procedure** is  
 404 **repeated** for **every source** of uncertainty described in Section 7. Both the experimental and the  
 405 theoretical uncertainties affect primarily the  **$t\bar{t}$  singal shape**, which is used to fit the data for the  
 406 QCD background normalization. As a result the total **systematic** uncertainty of the measure-  
 407 ment is **small** and the **statistical uncertainty dominates**. This is true both for the absolute and  
 408 the normalized cross sections. Figures 39-45 show the fiducial cross sections as a function of  
 409 the variables of interest. Firstly, we observe the **expected offset of about 35%** in the total **cross**  
 410 **section** between the data and the **Powheg+Pythia8** prediction, which is considerably smaller  
 411 for the other theory predictions. In particular the Powheg+Herwigpp seems to agree best (this  
 412 is most visible in Fig. 41). In terms of **shape** comparisons, the Powheg+Pythia8 prediction is in  
 413 excellent agreement with data for all the variables, **except** for the invariant mass  $m_{t\bar{t}}$  of the  $t\bar{t}$   
 414 system, where the theoretical spectrum appears to be harder above  $\approx 2$  TeV. A similar behavior  
 415 is observed for the other theory models with hints of somewhat larger differences, although the  
 416 statistical precision of the predictions does not allow quantitative comparisons.

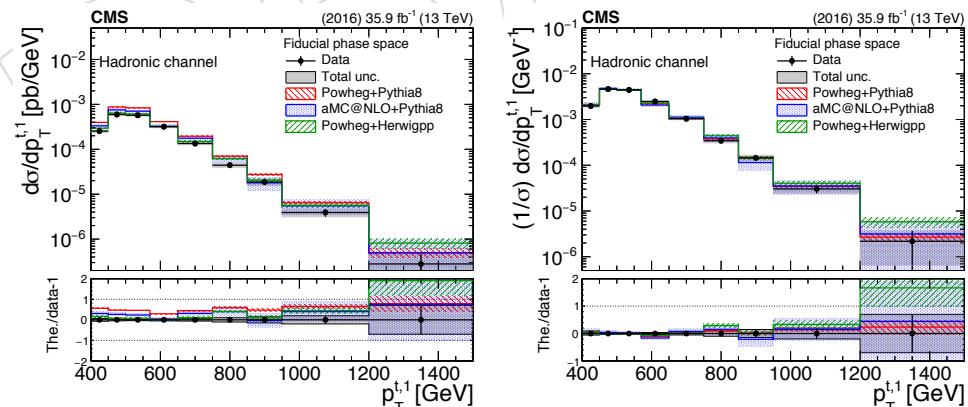


Figure 39: Fiducial differential cross section, absolute (left) and normalized (right), as a function of the leading top  $p_T$ . The bottom panel shows the ratio (theory - data)/data.

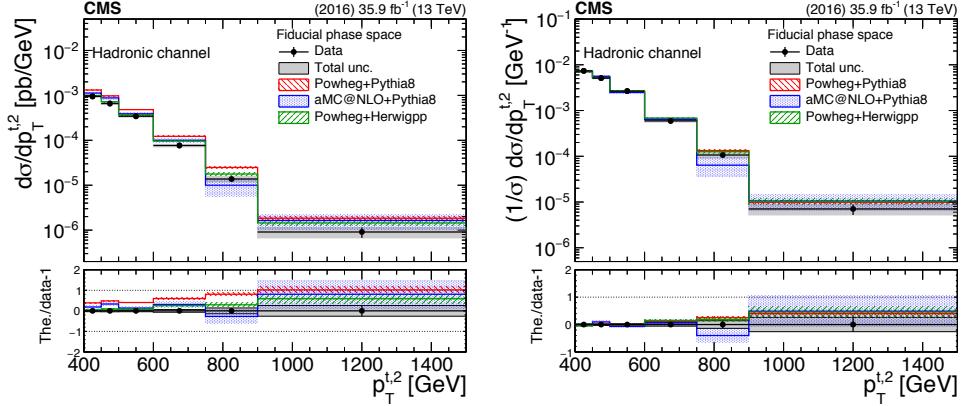


Figure 40: Fiducial differential cross section, absolute (left) and normalized (right), as a function of the second top  $p_T$ . The bottom panel shows the ratio (theory - data)/data.

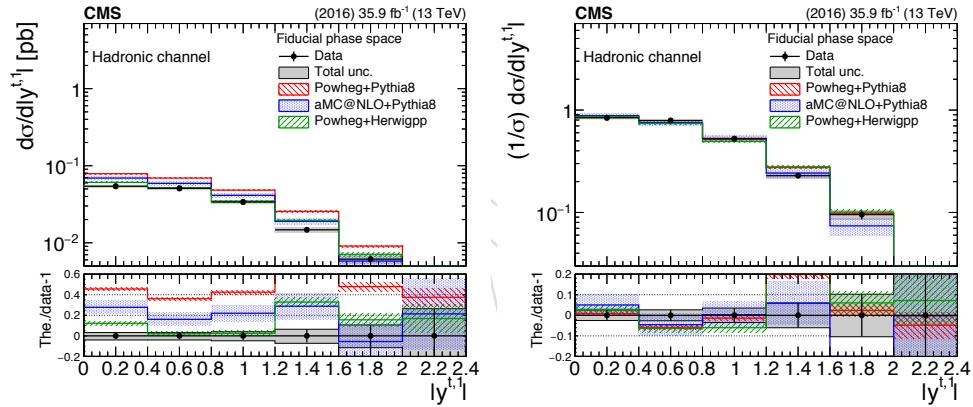


Figure 41: Fiducial differential cross section, absolute (left) and normalized (right), as a function of the leading top  $|y|$ . The bottom panel shows the ratio (theory - data)/data.

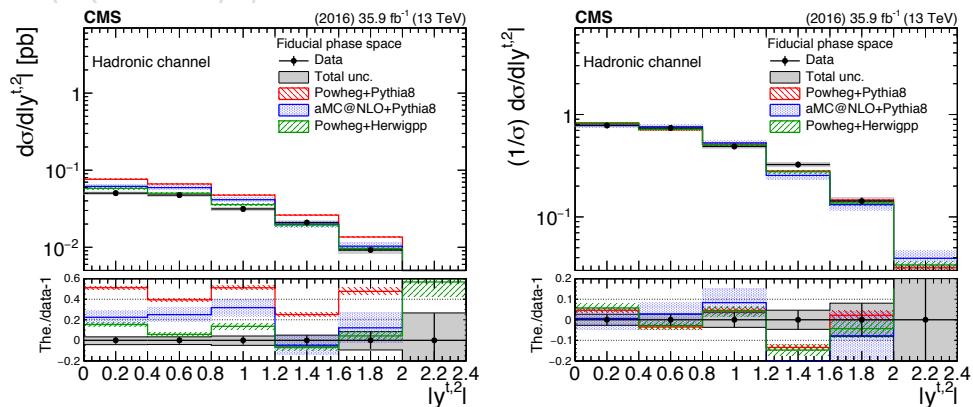


Figure 42: Fiducial differential cross section, absolute (left) and normalized (right), as a function of second top  $|y|$ . The bottom panel shows the ratio (theory - data)/data.

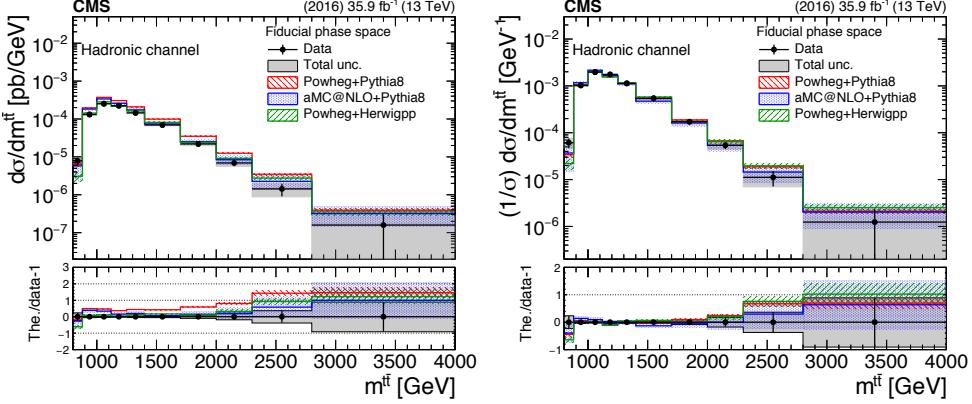


Figure 43: Fiducial differential cross section, absolute (left) and normalized (right), as a function of  $m_{t\bar{t}}^f$ . The bottom panel shows the ratio (theory - data)/data.

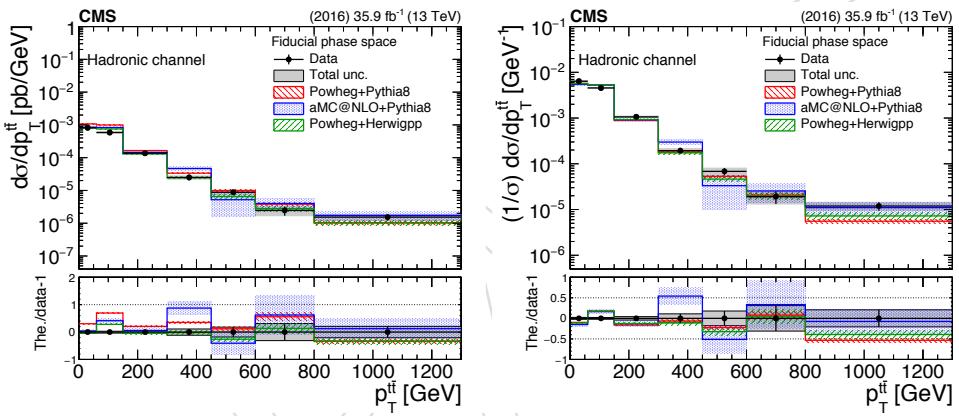


Figure 44: Fiducial differential cross section, absolute (left) and normalized (right), as a function of  $p_T^f$ . The bottom panel shows the ratio (theory - data)/data.

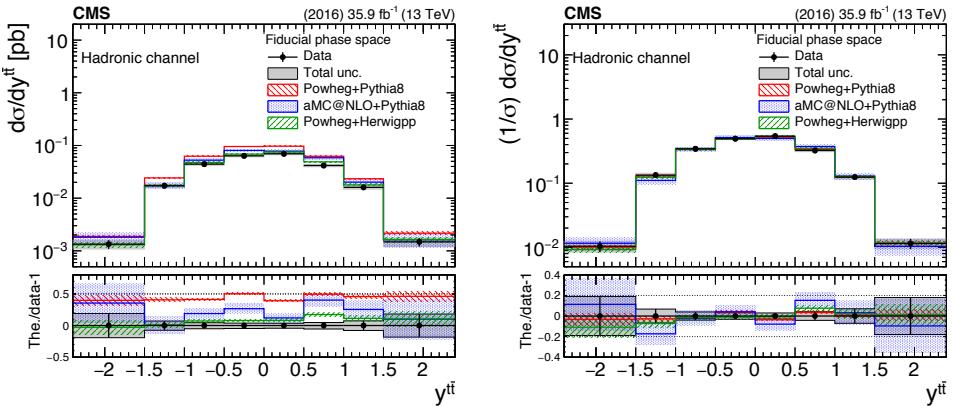


Figure 45: Fiducial differential cross section, absolute (left) and normalized (right), as a function of top  $y_{t\bar{t}}^f$ . The bottom panel shows the ratio (theory - data)/data.

## 417 9 Unfolded Measurement

418 The fiducial measurement at detector level is also reported unfolded to the parton and particle  
 419 levels. Equation 6 defines the unfolded differential cross section:

$$\frac{d\sigma_i^{\text{unf}}}{dx} = \frac{1}{\mathcal{L} \cdot \Delta x_i} \cdot \frac{1}{f_{2,i}} \cdot \sum_j \left( R_{ij}^{-1} \cdot f_{1,j} \cdot S_j \right), \quad (6)$$

420 where  $\mathcal{L}$  is the total integrated luminosity and  $\Delta x_i$  is the width of the i-th bin of the observable  
 421  $x$ . The quantity  $f_{1,j}$  is the fraction of reconstructed events in the j-th bin that have an equiva-  
 422 lence event at the unfolded level (parton or particle), whereas the quantity  $f_{2,i}$  is the fraction of  
 423 events at the unfolded level that have an equivalent reconstructed event. Figures 46 and 57 be-  
 424 low show the aforementioned fractions at parton and particle level, respectively, as a function  
 425 of all the observables. The quantity  $R_{ij}^{-1}$  is the inverse of the migration matrix between the i-th  
 426 and j-th bins. Due to the finite resolution of the detector, the migration matrix is non-diagonal  
 427 and thus the application of an unfolding procedure is necessary. The binning of the various  
 428 observables has been chosen such that the purity (fraction of reconstructed events that the true  
 429 value of the observable lies in the same bin) and the stability (fraction of true events that the re-  
 430 constructed observable lies in the same bin) are well above 50% (Figures 47 and 58). This choice  
 431 results in highly diagonal migration matrices, shown in Figs. 48 and 59. In order to avoid the  
 432 biases introduced by the various unfolding methods with some type of regularization, we have  
 433 used simple migration matrix inversion, as written in Eq. 6 at a price of a moderate increase of  
 434 the statistical uncertainty (see details in Appendix E).

### 435 9.1 Parton Level

436 The results of the unfolded measurement at parton level are shown in Figs. 49–55. The compari-  
 437 son with the theory predictions follows the trends observed in the fiducial measurement. That  
 438 is, there is an 20–40% lower inclusive cross section, while the shapes of the differential distri-  
 439 butions are reasonably reproduced by all models. A hint of a deviation at very high values of  
 440  $m^{t\bar{t}}$  is observed but the measurement is not very precise at this part of the phase space.

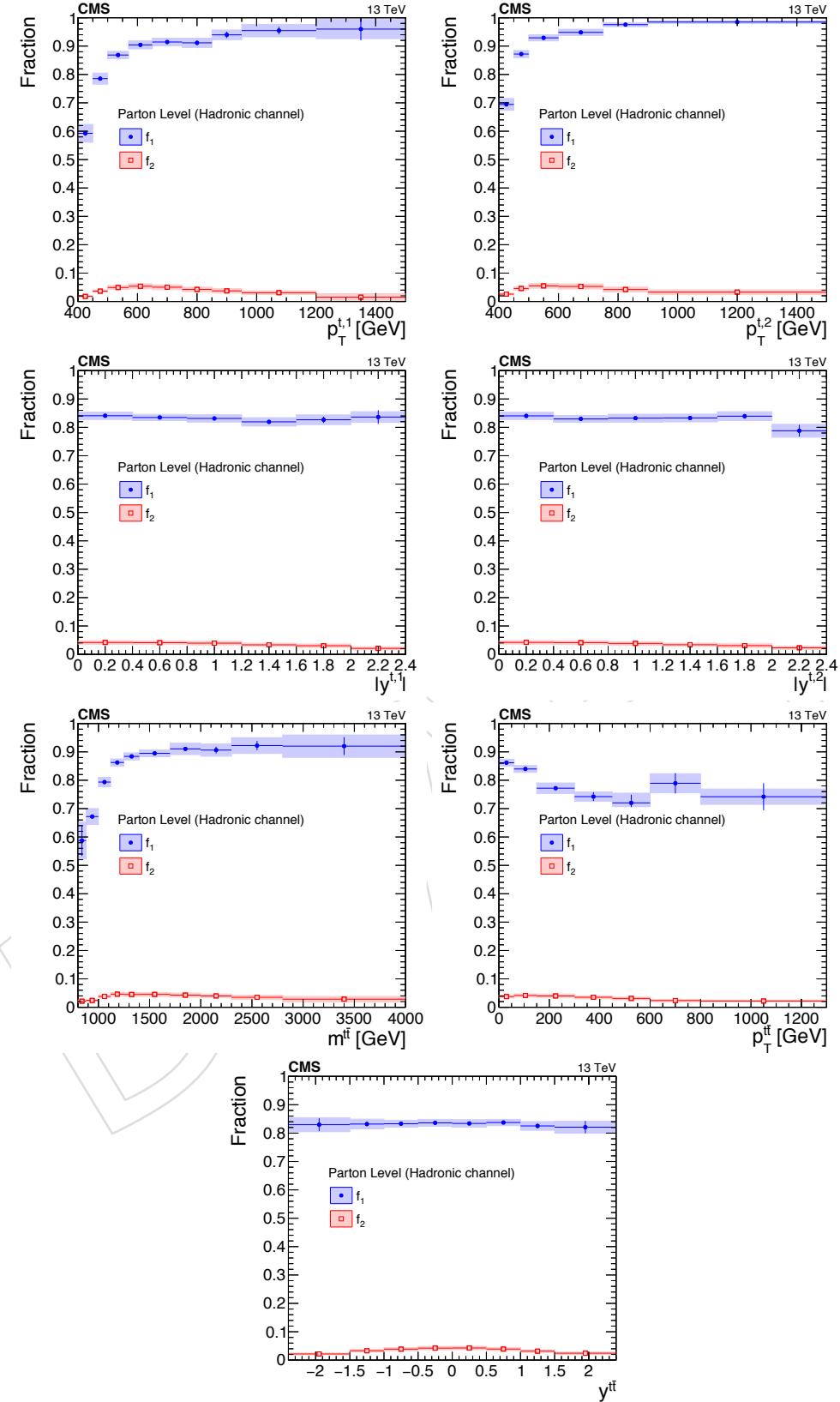


Figure 46: Simulated fractions  $f_{1,2}$  for the parton-level selection as a function of the various observables.

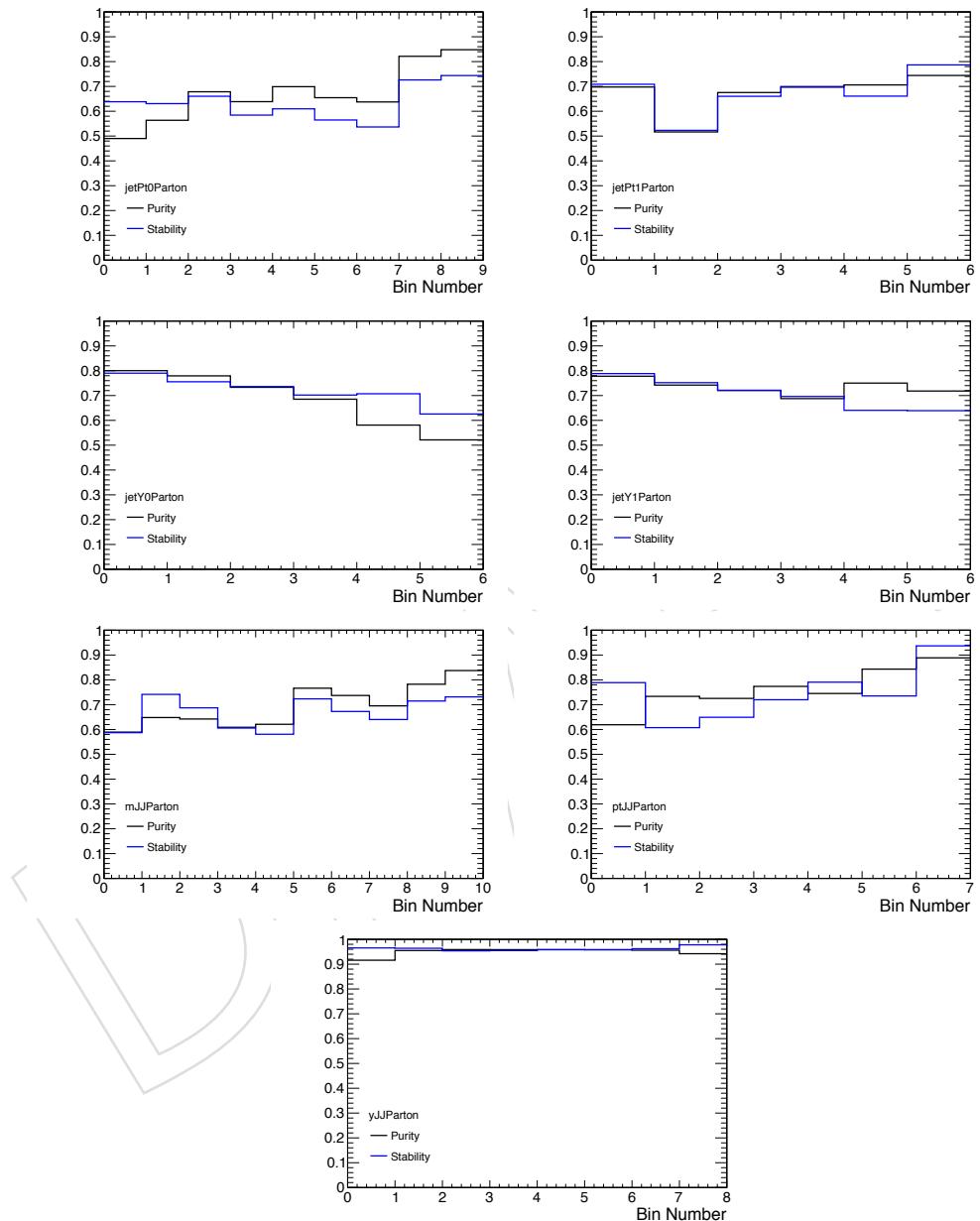


Figure 47: Simulated purity and stability of each bin at parton level.

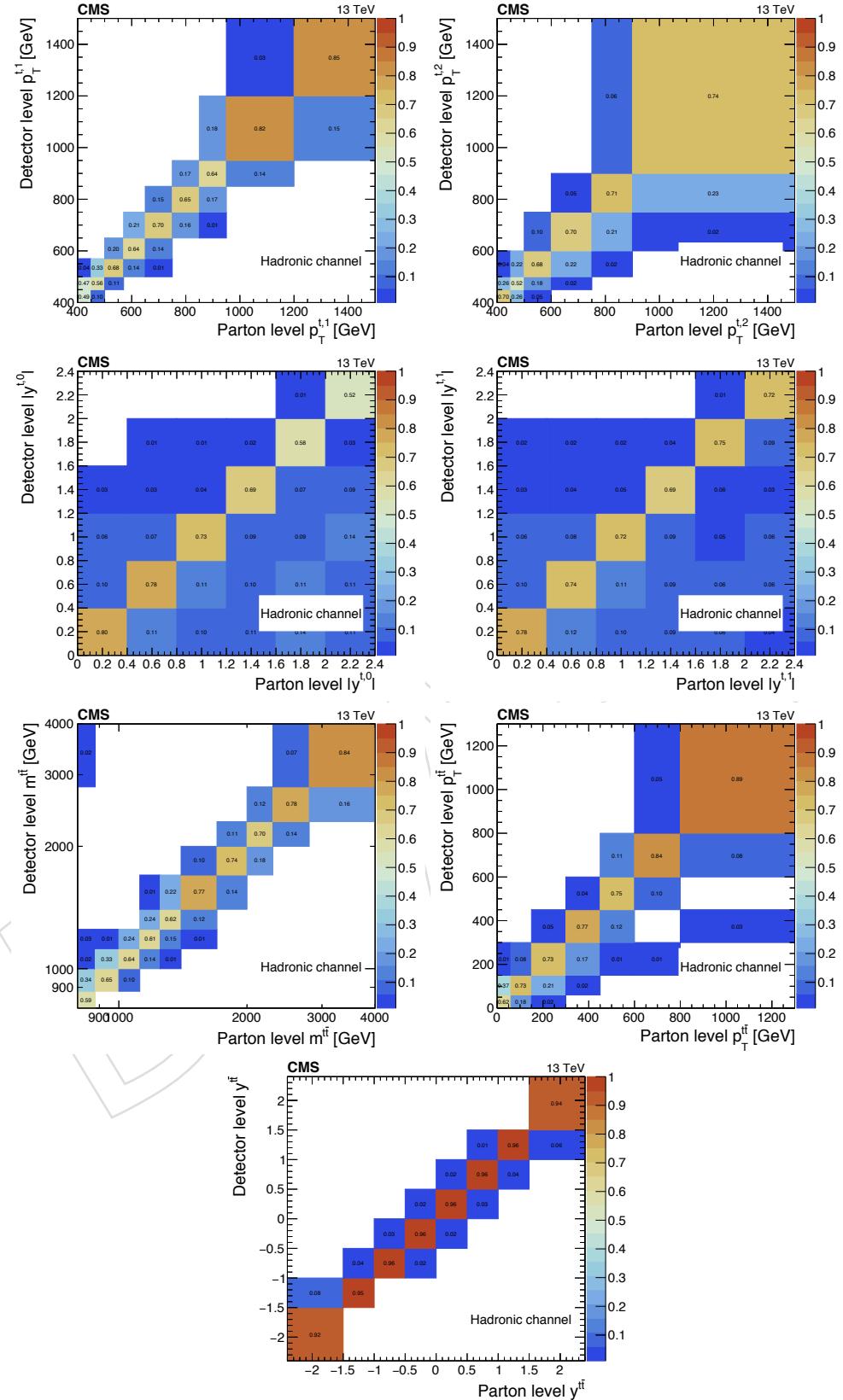


Figure 48: Simulated migration matrices at parton level. Each column is normalized to unity.

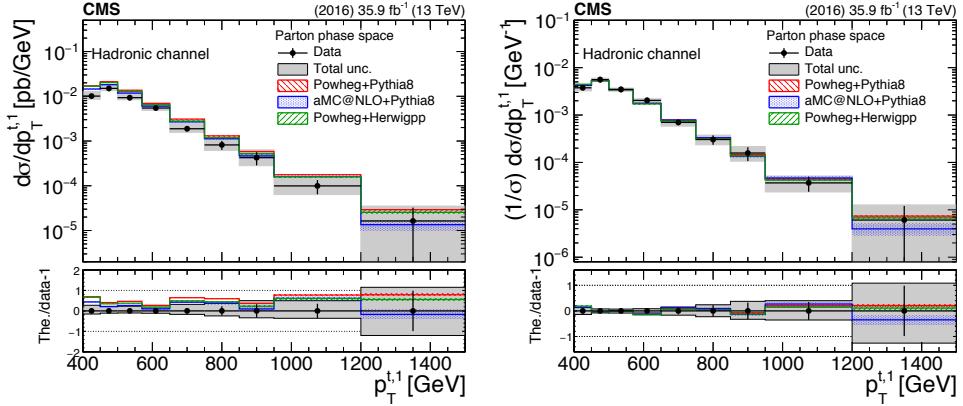


Figure 49: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of the leading top  $p_T$ . The bottom panel shows the ratio (theory - data)/data.

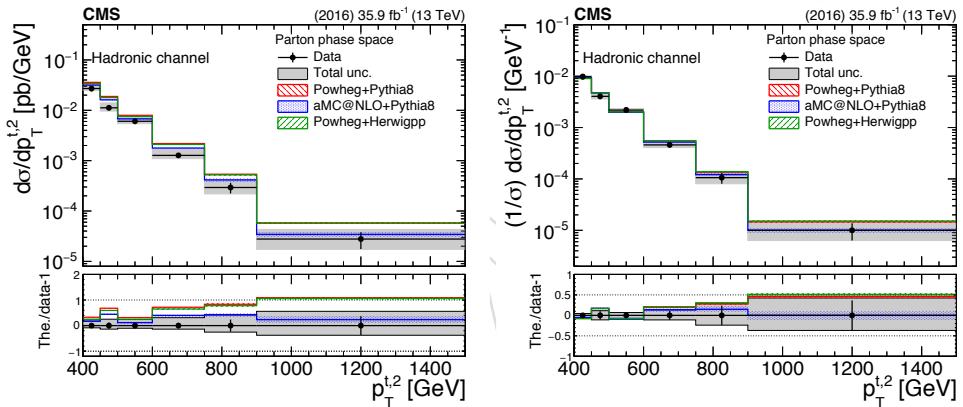


Figure 50: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of the second top  $p_T$ . The bottom panel shows the ratio (theory - data)/data.

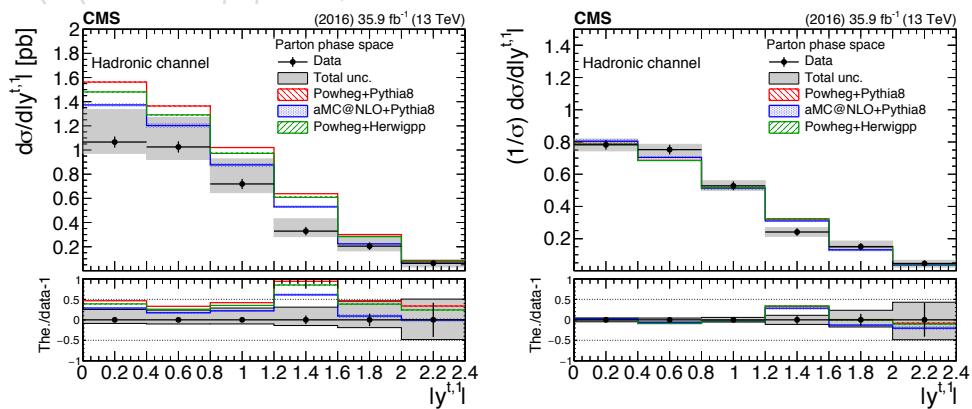


Figure 51: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of the leading top  $|y|$ . The bottom panel shows the ratio (theory - data)/data.

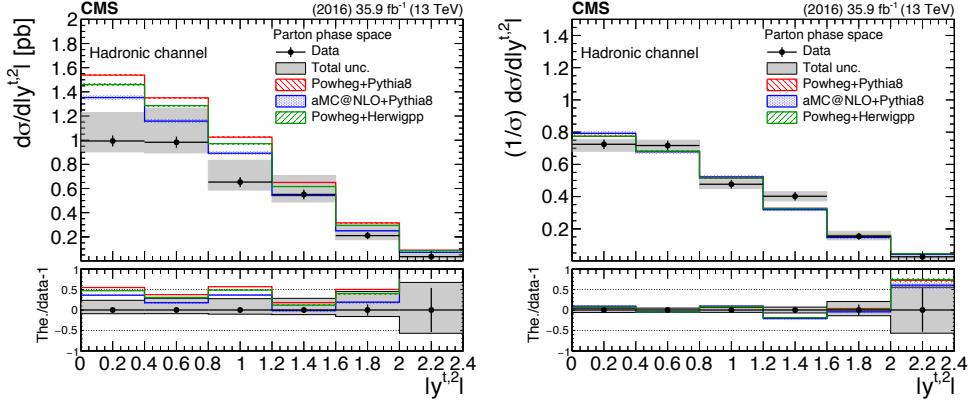


Figure 52: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of the second top  $|y|$ . The bottom panel shows the ratio (theory - data).

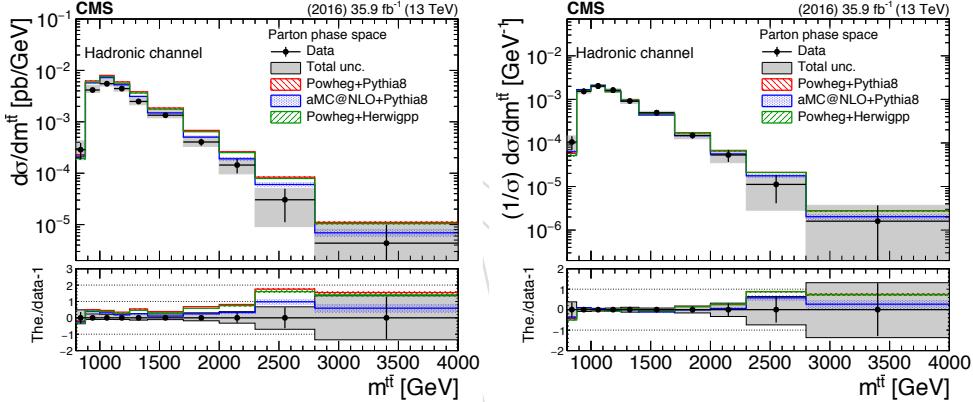


Figure 53: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of  $m_{t\bar{t}}$ . The bottom panel shows the ratio (theory - data).

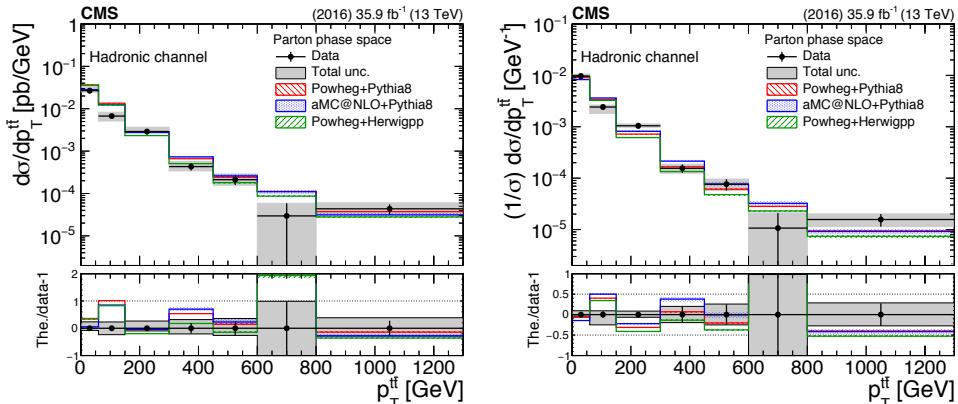


Figure 54: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of  $p_T^{t\bar{t}}$ . The bottom panel shows the ratio (theory - data).

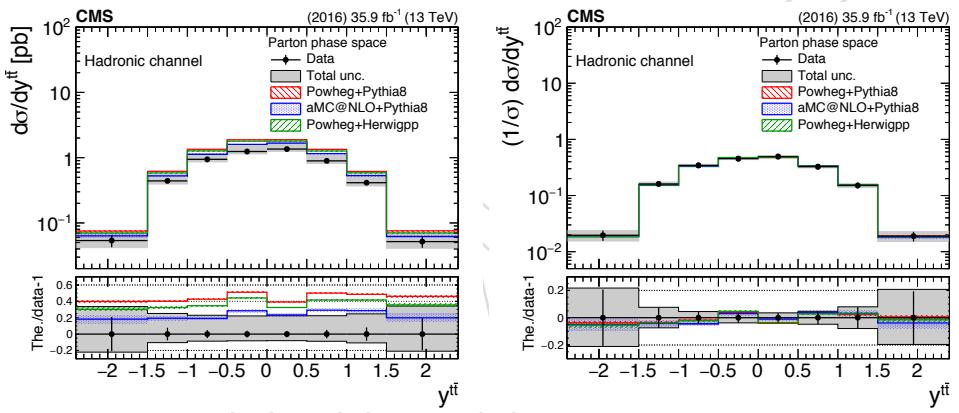


Figure 55: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of  $y^{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data.

## 441 9.2 Particle Level

442 The results of the unfolded measurement at particle level are shown in Figs. 60- 66. We observe  
 443 that the uncertainties are similar to those of the parton level, while the comparison with the  
 444 theory predictions follows the trends observed in the fiducial measurement. That is, there is  
 445 an 20 – 40% lower inclusive cross section, while the shapes of the differential distributions  
 446 are reasonably reproduced by all models. A hint of a deviation at very high values of  $m^{t\bar{t}}$  is  
 447 observed but the measurement is not very precise at this part of the phase space.

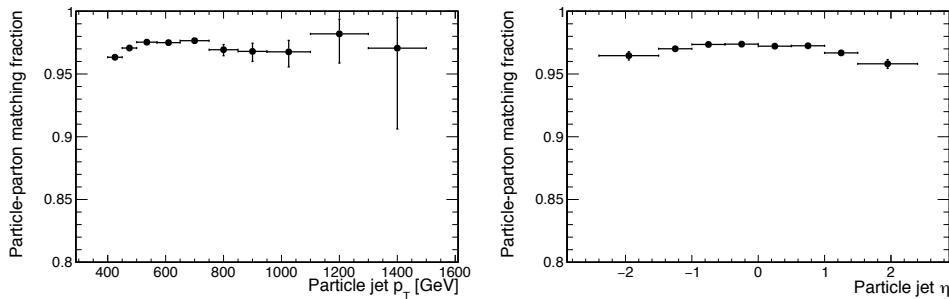


Figure 56: Simulated efficiency of the particle-level top candidates as a function of  $p_T$  and  $\eta$ .

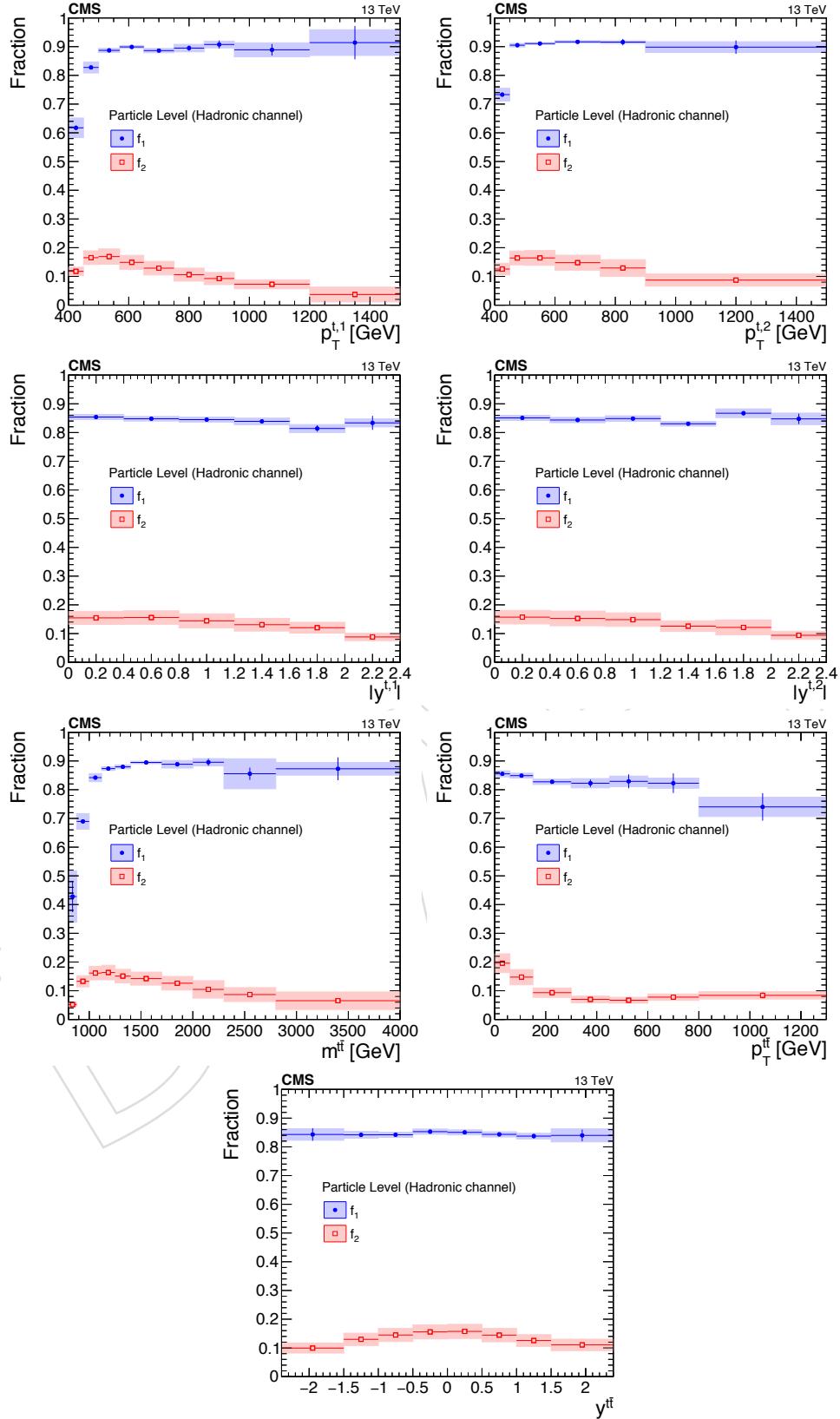


Figure 57: Simulated acceptance and efficiency for the particle-level selection as a function of the various observables.

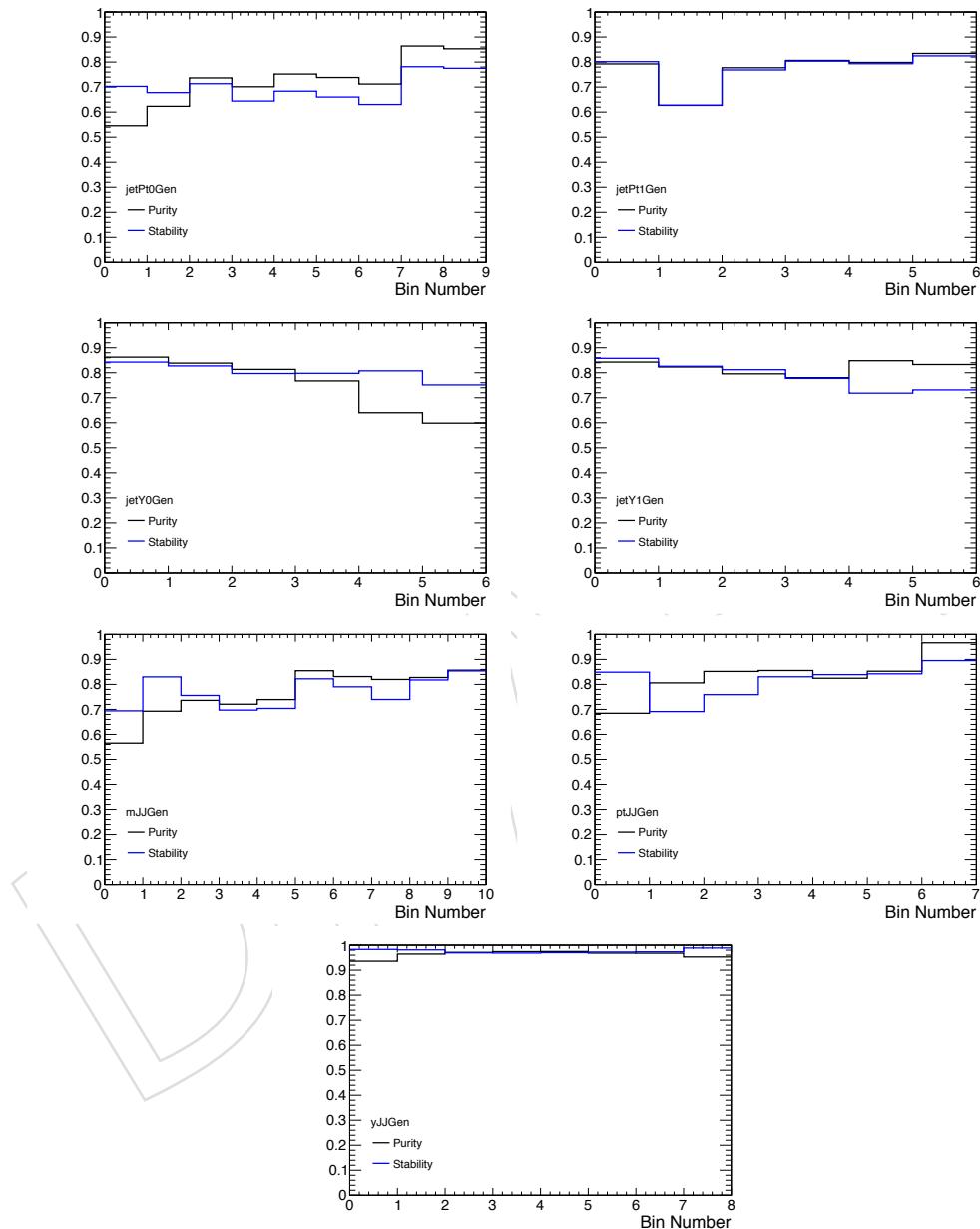


Figure 58: Simulated purity and stability of each bin at particle level.

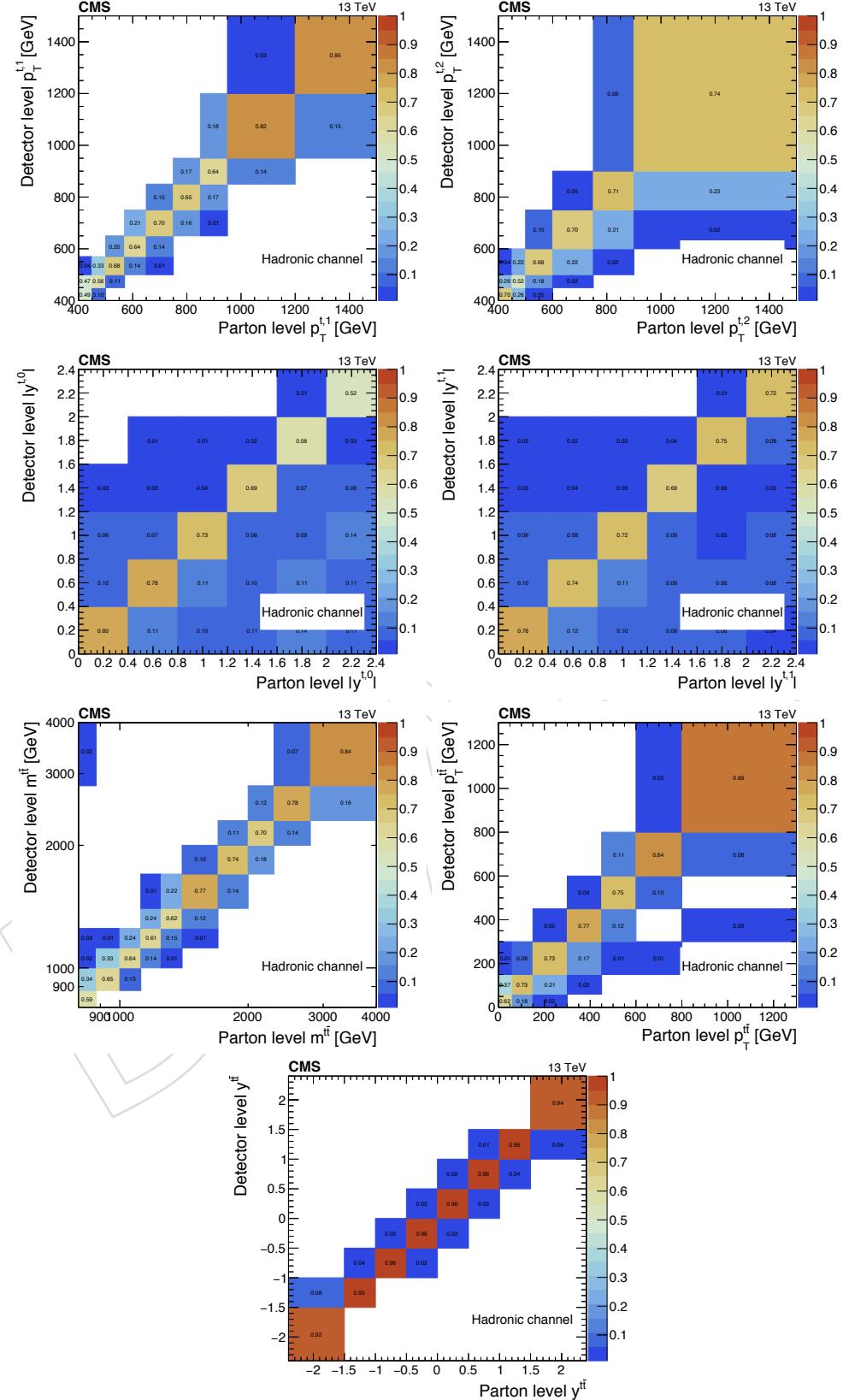


Figure 59: Simulated migration matrices at particle level. Each column is normalized to unity.

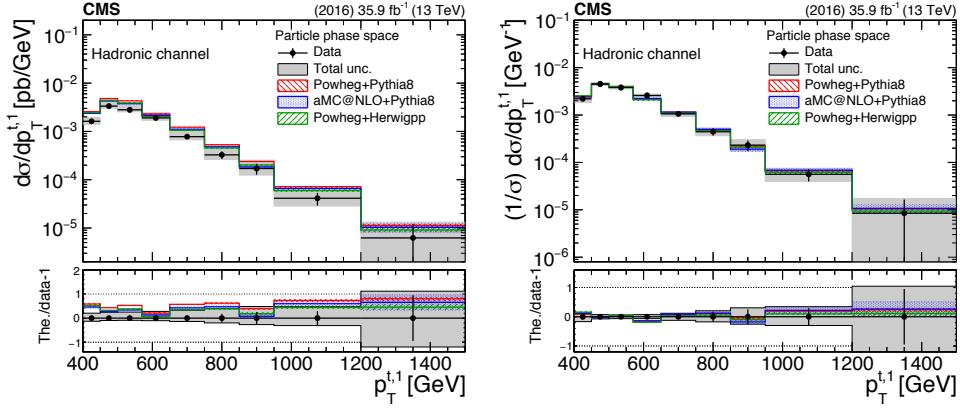


Figure 60: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of the leading top  $p_T$ . The bottom panel shows the ratio (theory - data)/data.

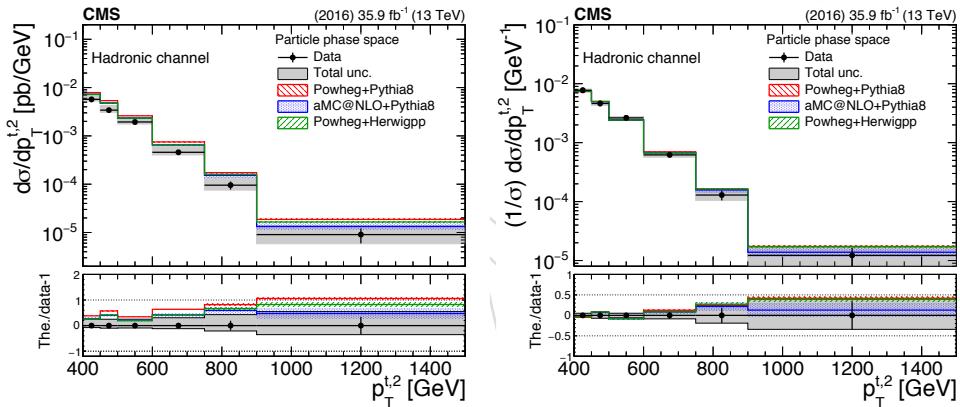


Figure 61: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of the second top  $p_T$ . The bottom panel shows the ratio (theory - data)/data.

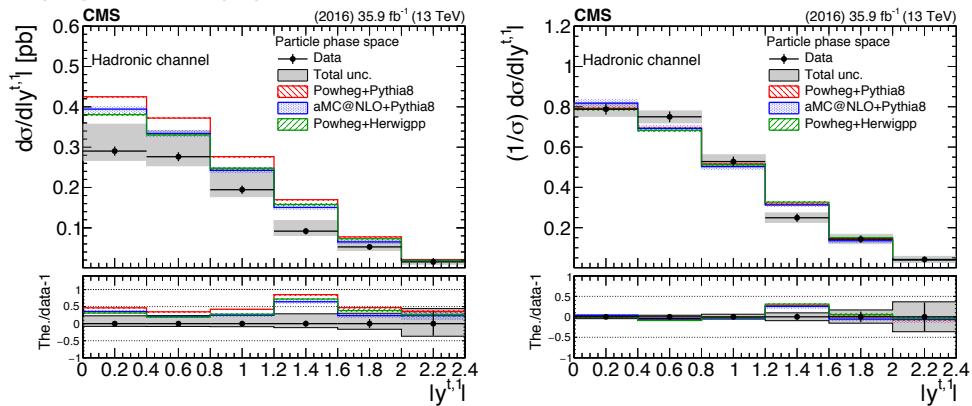


Figure 62: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of the leading top  $|y|$ . The bottom panel shows the ratio (theory - data)/data.

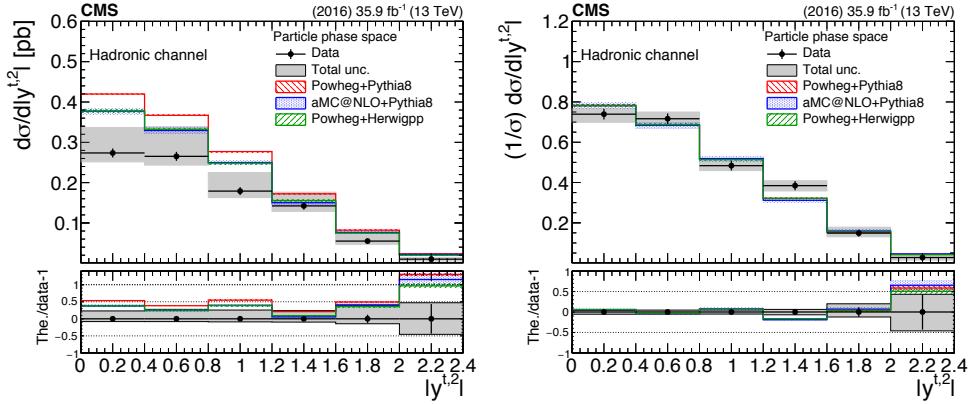


Figure 63: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of the leading top  $|y|$ . The bottom panel shows the ratio (theory - data)/data.

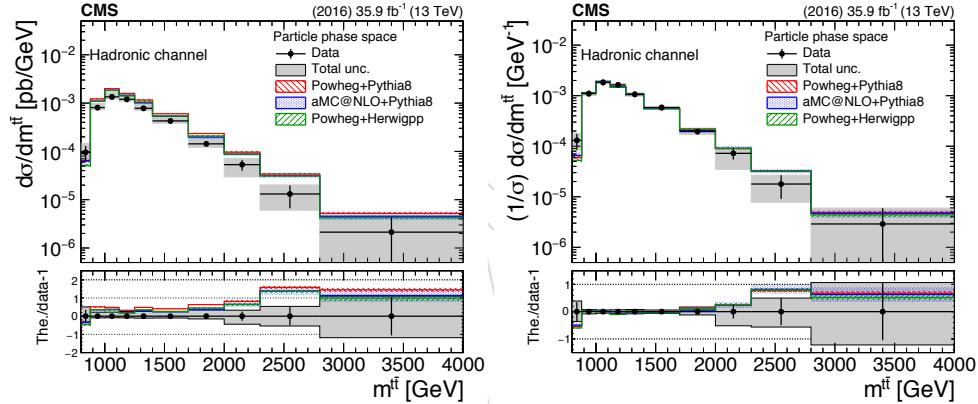


Figure 64: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of  $m_{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data.

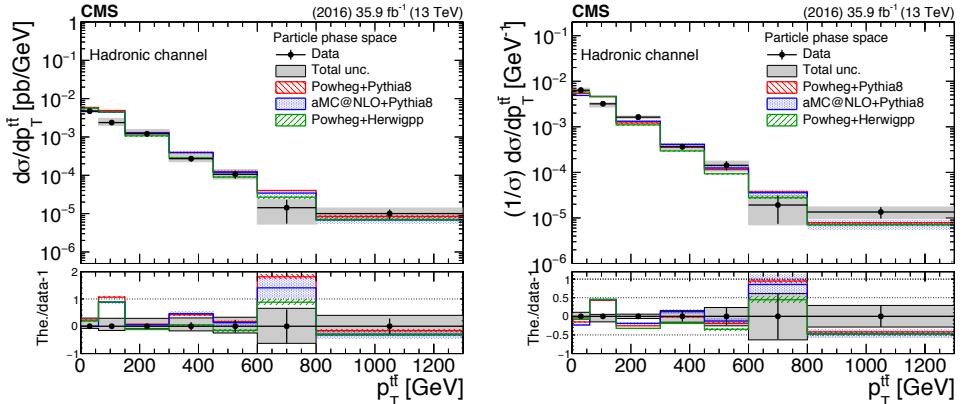


Figure 65: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of  $p_T^{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data.

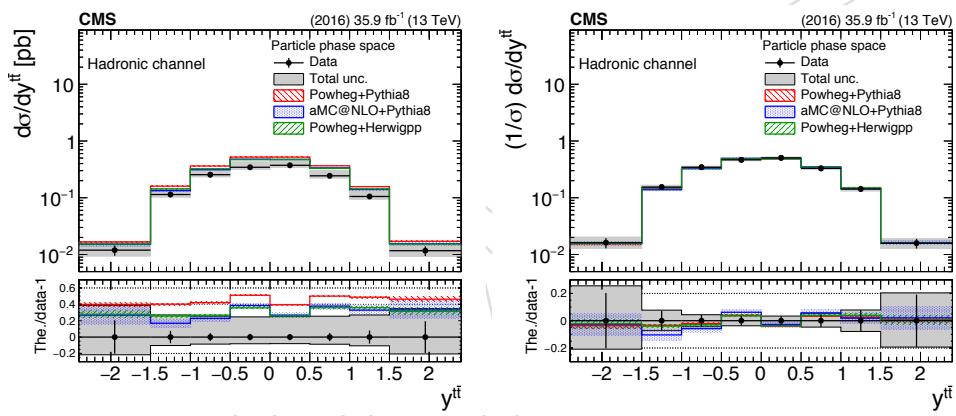
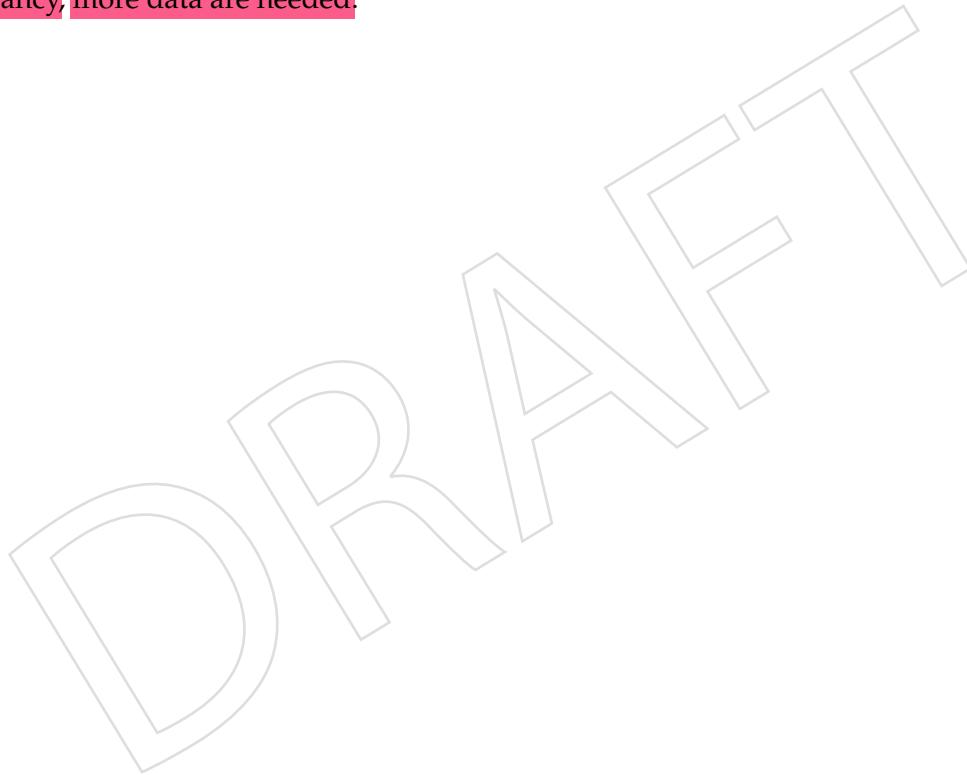


Figure 66: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of  $y_{t\bar{t}}$ . The bottom panel shows the ratio (theory - data)/data.

---

**448 10 Summary**

449 A measurement of the  $t\bar{t}$  production cross section for high- $p_T$  top quarks at 13 TeV proton-  
450 proton collisions has been presented. The measurement is performed with events where both  
451 top quarks decay hadronically, and where the hadronic decay products cannot be resolved but  
452 are instead clustered in a single large- $R$  jet with  $p_T > 400 \text{ GeV}$ . The hadronic final state thus  
453 contains two large- $R$  jets. The cross section is reported differentially as a function of (inclusive)  
454 top  $p_T$  and  $\eta$ , and as a function of the invariant mass,  $p_T$ , and rapidity of the  $t\bar{t}$  system, un-  
455 folded to the parton and particle levels, absolute and normalized. The results are compared  
456 to theoretical predictions from the POWHEG matrix-element generator, interfaced with PYTHIA  
457 8 or HERWIG ++ for the underlying event and parton shower, and from the MC@NLO matrix-  
458 element generator, interfaced with PYTHIA 8. All the models overpredict significantly the ab-  
459 solute cross section in the phase space of the measurement (up to 40%), while they are able  
460 to describe consistently the differential shapes in all the variables. The most notable discrep-  
461 ency can be seen in the invariant mass of the  $t\bar{t}$  system, where the theoretical models predict  
462 higher cross section at high mass values. However, in order to estimate the significance of the  
463 discrepancy, more data are needed.

A large, semi-transparent watermark is printed diagonally across the page. It consists of the word "DRAFT" in a bold, sans-serif font, with each letter slightly overlapping the next. The watermark is oriented from the bottom-left towards the top-right, covering approximately one-third of the page area.

## 464 References

- [1] ATLAS Collaboration, "Measurements of top-quark pair differential cross-sections in the lepton+jets channel in  $pp$  collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector", *JHEP* **11** (2017) 191, doi:10.1007/JHEP11(2017)191, arXiv:1708.00727.
- [2] ATLAS Collaboration, "Measurements of top-quark pair differential cross-sections in the  $e\mu$  channel in  $pp$  collisions at  $\sqrt{s} = 13$  TeV using the ATLAS detector", *Eur. Phys. J.* **C77** (2017), no. 5, 292, doi:10.1140/epjc/s10052-017-4821-x, arXiv:1612.05220.
- [3] ATLAS Collaboration, "Measurement of lepton differential distributions and the top quark mass in  $t\bar{t}$  production in  $pp$  collisions at  $\sqrt{s} = 8$  TeV with the ATLAS detector", *Eur. Phys. J.* **C77** (2017), no. 11, 804, doi:10.1140/epjc/s10052-017-5349-9, arXiv:1709.09407.
- [4] ATLAS Collaboration, "Measurement of top quark pair differential cross-sections in the dilepton channel in  $pp$  collisions at  $\sqrt{s} = 7$  and 8 TeV with ATLAS", *Phys. Rev.* **D94** (2016), no. 9, 092003, doi:10.1103/PhysRevD.94.092003, arXiv:1607.07281.
- [5] ATLAS Collaboration, "Differential top-antitop cross-section measurements as a function of observables constructed from final-state particles using  $pp$  collisions at  $\sqrt{s} = 7$  TeV in the ATLAS detector", *JHEP* **06** (2015) 100, doi:10.1007/JHEP06(2015)100, arXiv:1502.05923.
- [6] ATLAS Collaboration, "Measurements of top quark pair relative differential cross-sections with ATLAS in  $pp$  collisions at  $\sqrt{s} = 7$  TeV", *Eur. Phys. J.* **C73** (2013), no. 1, 2261, doi:10.1140/epjc/s10052-012-2261-1, arXiv:1207.5644.
- [7] CMS Collaboration, "Measurement of normalized differential  $t\bar{t}$  cross sections in the dilepton channel from  $pp$  collisions at  $\sqrt{s} = 13$  TeV", *JHEP* **04** (2018) 060, doi:10.1007/JHEP04(2018)060, arXiv:1708.07638.
- [8] CMS Collaboration, "Measurements of normalised multi-differential cross sections for top quark pair production in  $pp$  collisions at  $\sqrt{s} = 13$  TeV and simultaneous determination of the strong coupling strength, top quark pole mass and parton distribution functions", Technical Report CMS-PAS-TOP-18-004, CERN, Geneva, 2018.
- [9] CMS Collaboration, "Measurements of  $t\bar{t}$  differential cross sections in proton-proton collisions at  $\sqrt{s} = 13$  TeV using events containing two leptons", *Submitted to: JHEP* (2018) arXiv:1811.06625.
- [10] CMS Collaboration, "Measurement of differential cross sections for top quark pair production using the lepton+jets final state in proton-proton collisions at 13 TeV", *Phys. Rev.* **D95** (2017), no. 9, 092001, doi:10.1103/PhysRevD.95.092001, arXiv:1610.04191.
- [11] CMS Collaboration, "Measurements of differential cross sections of top quark pair production as a function of kinematic event variables in proton-proton collisions at  $\sqrt{s} = 13$  TeV", *JHEP* **06** (2018) 002, doi:10.1007/JHEP06(2018)002, arXiv:1803.03991.
- [12] CMS Collaboration, "Measurement of double-differential cross sections for top quark pair production in  $pp$  collisions at  $\sqrt{s} = 8$  TeV and impact on parton distribution

- 506       functions”, *Eur. Phys. J.* **C77** (2017), no. 7, 459,  
507       doi:[10.1140/epjc/s10052-017-4984-5](https://doi.org/10.1140/epjc/s10052-017-4984-5), arXiv:[1703.01630](https://arxiv.org/abs/1703.01630).
- 508 [13] CMS Collaboration, “Measurement of the differential cross section for top quark pair  
509 production in pp collisions at  $\sqrt{s} = 8 \text{ TeV}$ ”, *Eur. Phys. J.* **C75** (2015), no. 11, 542,  
510 doi:[10.1140/epjc/s10052-015-3709-x](https://doi.org/10.1140/epjc/s10052-015-3709-x), arXiv:[1505.04480](https://arxiv.org/abs/1505.04480).
- 511 [14] CMS Collaboration, “Measurement of the differential cross sections for top quark pair  
512 production as a function of kinematic event variables in pp collisions at  $\sqrt{s}=7$  and 8  
513 TeV”, *Phys. Rev.* **D94** (2016), no. 5, 052006, doi:[10.1103/PhysRevD.94.052006](https://doi.org/10.1103/PhysRevD.94.052006),  
514 arXiv:[1607.00837](https://arxiv.org/abs/1607.00837).
- 515 [15] CMS Collaboration, “Measurement of differential top-quark pair production cross  
516 sections in pp colisions at  $\sqrt{s} = 7 \text{ TeV}$ ”, *Eur. Phys. J.* **C73** (2013), no. 3, 2339,  
517 doi:[10.1140/epjc/s10052-013-2339-4](https://doi.org/10.1140/epjc/s10052-013-2339-4), arXiv:[1211.2220](https://arxiv.org/abs/1211.2220).
- 518 [16] ATLAS Collaboration, “Measurements of  $t\bar{t}$  differential cross-sections of highly boosted  
519 top quarks decaying to all-hadronic final states in pp collisions at  $\sqrt{s} = 13 \text{ TeV}$  using the  
520 ATLAS detector”, *Phys. Rev.* **D98** (2018), no. 1, 012003,  
521 doi:[10.1103/PhysRevD.98.012003](https://doi.org/10.1103/PhysRevD.98.012003), arXiv:[1801.02052](https://arxiv.org/abs/1801.02052).
- 522 [17] ATLAS Collaboration, “Measurement of the differential cross-section of highly boosted  
523 top quarks as a function of their transverse momentum in  $\sqrt{s} = 8 \text{ TeV}$  proton-proton  
524 collisions using the ATLAS detector”, *Phys. Rev.* **D93** (2016), no. 3, 032009,  
525 doi:[10.1103/PhysRevD.93.032009](https://doi.org/10.1103/PhysRevD.93.032009), arXiv:[1510.03818](https://arxiv.org/abs/1510.03818).
- 526 [18] CMS Collaboration Collaboration, “Measurement of the  $t\bar{t}$  production cross section at 13  
527 TeV in the all-jets final state”, Technical Report CMS-PAS-TOP-16-013, CERN, Geneva,  
528 2016.
- 529 [19] CMS Collaboration, “Measurement of the integrated and differential  $t\bar{t}$  production cross  
530 sections for high- $p_t$  top quarks in pp collisions at  $\sqrt{s} = 8 \text{ TeV}$ ”, *Phys. Rev.* **D94** (2016),  
531 no. 7, 072002, doi:[10.1103/PhysRevD.94.072002](https://doi.org/10.1103/PhysRevD.94.072002), arXiv:[1605.00116](https://arxiv.org/abs/1605.00116).
- 532 [20] CMS Collaboration Collaboration, “Measurement of the differential  $t\bar{t}$  cross section with  
533 high- $p_T$  top-quark jets in the all-hadronic channel at  $\sqrt{s} = 8 \text{ TeV}$ ”, Technical Report  
534 CMS-PAS-TOP-16-018, CERN, Geneva, 2017.
- 535 [21] P. Nason, “A new method for combining NLO QCD with shower Monte Carlo  
536 algorithms”, *JHEP* **11** (2004) 040, doi:[10.1088/1126-6708/2004/11/040](https://doi.org/10.1088/1126-6708/2004/11/040),  
537 arXiv:[hep-ph/0409146](https://arxiv.org/abs/hep-ph/0409146).
- 538 [22] S. Frixione, P. Nason, and G. Ridolfi, “A Positive-weight next-to-leading-order Monte  
539 Carlo for heavy flavour hadroproduction”, *JHEP* **09** (2007) 126,  
540 doi:[10.1088/1126-6708/2007/09/126](https://doi.org/10.1088/1126-6708/2007/09/126), arXiv:[0707.3088](https://arxiv.org/abs/0707.3088).
- 541 [23] S. Frixione, P. Nason, and C. Oleari, “Matching NLO QCD computations with parton  
542 shower simulations: the POWHEG method”, *JHEP* **11** (2007) 070,  
543 doi:[10.1088/1126-6708/2007/11/070](https://doi.org/10.1088/1126-6708/2007/11/070), arXiv:[0709.2092](https://arxiv.org/abs/0709.2092).
- 544 [24] S. Alioli, P. Nason, C. Oleari, and E. Re, “A general framework for implementing NLO  
545 calculations in shower Monte Carlo programs: the POWHEG BOX”, *JHEP* **06** (2010) 043,  
546 doi:[10.1007/JHEP06\(2010\)043](https://doi.org/10.1007/JHEP06(2010)043), arXiv:[1002.2581](https://arxiv.org/abs/1002.2581).

- [25] S. Alioli, S. O. Moch, and P. Uwer, “Hadronic top-quark pair-production with one jet and parton showering”, *JHEP* **01** (2012) 137, doi:10.1007/JHEP01(2012)137, arXiv:1110.5251.
- [26] S. Alioli, P. Nason, C. Oleari, and E. Re, “NLO single-top production matched with shower in POWHEG:  $s$ - and  $t$ -channel contributions”, *JHEP* **09** (2009) 111, doi:10.1088/1126-6708/2009/09/111, arXiv:0907.4076. [Erratum: doi:10.1007/JHEP02(2010)011].
- [27] 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”, *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.
- [28] J. Alwall et al., “Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions”, *Eur. Phys. J.* **C53** (2008) 473–500, doi:10.1140/epjc/s10052-007-0490-5, arXiv:0706.2569.
- [29] T. Sjöstrand, S. Mrenna, and P. Z. Skands, “PYTHIA 6.4 physics and manual”, *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.
- [30] T. Sjöstrand, S. Mrenna, and P. Z. Skands, “A brief introduction to PYTHIA 8.1”, *Comput. Phys. Commun.* **178** (2008) 852, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.
- [31] NNPDF Collaboration, “Parton distributions for the LHC Run II”, *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.
- [32] CMS Collaboration, “Event generator tunes obtained from underlying event and multiparton scattering measurements”, *Eur. Phys. J. C* **76** (2015) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.
- [33] CMS Collaboration, “Investigations of the impact of the parton shower tuning in Pythia 8 in the modelling of  $t\bar{t}$  at  $\sqrt{s} = 8$  and 13 TeV”, Technical Report CMS-PAS-TOP-16-021, CERN, Geneva, 2016.
- [34] GEANT4 Collaboration, “GEANT4—a simulation toolkit”, *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.
- [35] M. Czakon and A. Mitov, “Top++: A program for the calculation of the top-pair cross-section at hadron colliders”, *Comput. Phys. Commun.* **185** (2014) 2930, doi:10.1016/j.cpc.2014.06.021, arXiv:1112.5675.
- [36] Y. Li and F. Petriello, “Combining QCD and electroweak corrections to dilepton production in the framework of the FEWZ simulation code”, *Phys. Rev. D* **86** (2012) 094034, doi:10.1103/PhysRevD.86.094034, arXiv:1208.5967.
- [37] N. Kidonakis, “Top Quark Production”, (2014). arXiv:1311.0283.
- [38] M. Bahr et al., “Herwig++ Physics and Manual”, *Eur. Phys. J.* **C58** (2008) 639–707, doi:10.1140/epjc/s10052-008-0798-9, arXiv:0803.0883.
- [39] S. Gieseke, C. Rohr, and A. Siodmok, “Colour reconnections in Herwig++”, *Eur. Phys. J.* **C72** (2012) 2225, doi:10.1140/epjc/s10052-012-2225-5, arXiv:1206.0041.

- 586 [40] J. Therhaag, "TMVA Toolkit for multivariate data analysis in ROOT", *PoS ICHEP2010*  
587 (2010) 510, doi:10.22323/1.120.0510.
- 588 [41] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp  
589 collisions at 8 TeV", *JINST* **12** (2017) P02014,  
590 doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.
- 591 [42] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp  
592 collisions at 13 TeV", *JINST* **13** (2018) P05011,  
593 doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.
- 594 [43] ATLAS Collaboration, "Measurement of the Inelastic Proton-Proton Cross Section at  
595  $\sqrt{s} = 13$  TeV with the ATLAS Detector at the LHC", *Phys. Rev. Lett.* **117** (2016), no. 18,  
596 182002, doi:10.1103/PhysRevLett.117.182002, arXiv:1606.02625.
- 597 [44] CMS Collaboration Collaboration, "CMS Luminosity Measurements for the 2016 Data  
598 Taking Period", Technical Report CMS-PAS-LUM-17-001, CERN, Geneva, 2017.

DRAFT

## 599 A Data vs simulation with default MC normalization

600 Below we show the data vs simulation comparisons for various observables where the MC pro-  
 601 cesses are normalized according to the known cross sections, without using the postfit values  
 602 for the  $t\bar{t}$  and QCD yields.

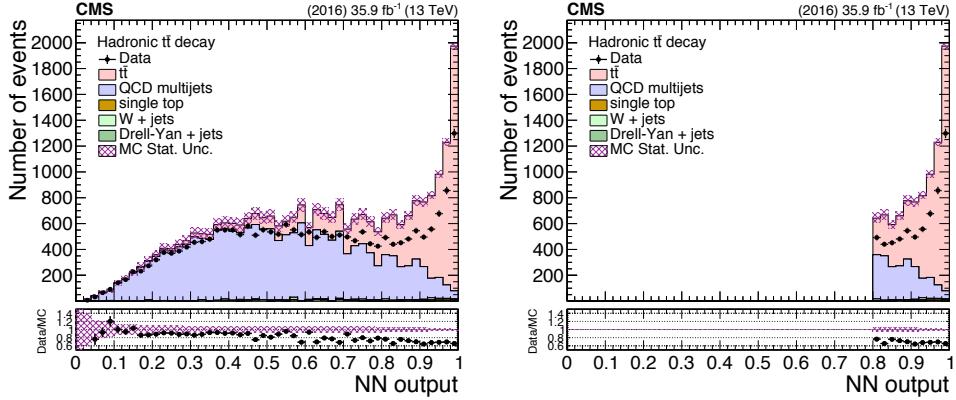


Figure 67: Data vs simulation before (left) and after (right) applying the  $NN > 0.8$  cut for the NN output.

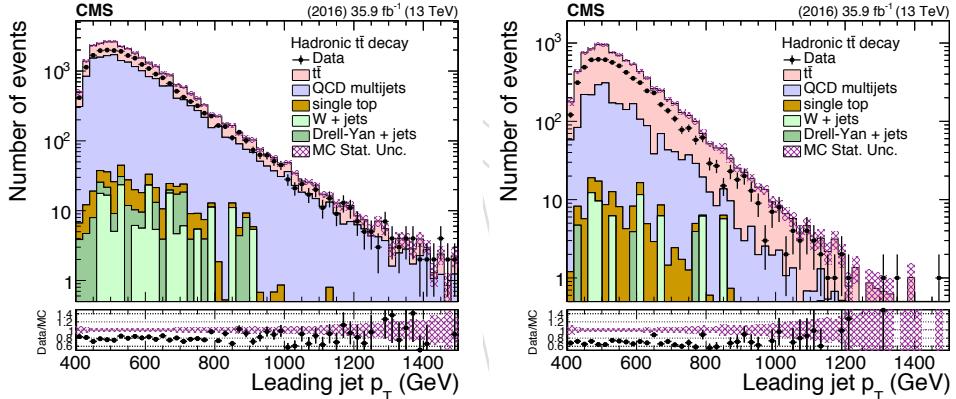


Figure 68: Data vs simulation before (left) and after (right) applying the  $NN > 0.8$  cut for the  $p_T$  of the leading jet.

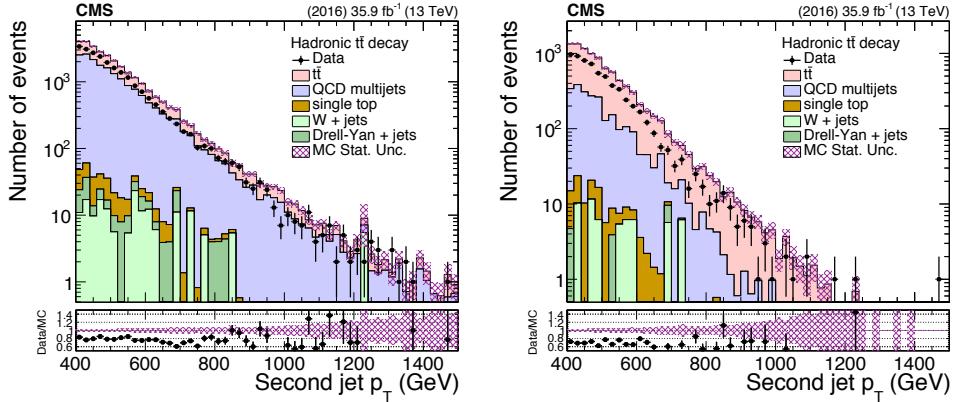


Figure 69: Data vs simulation before (left) and after (right) applying the  $NN > 0.8$  cut for the  $p_T$  of the second jet.

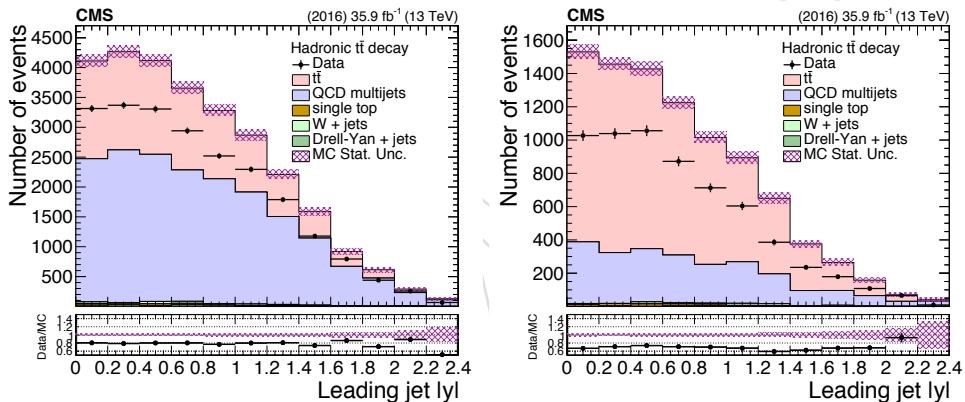


Figure 70: Data vs simulation before (left) and after (right) applying the  $NN > 0.8$  cut for the  $|y|$  of the leading jet.

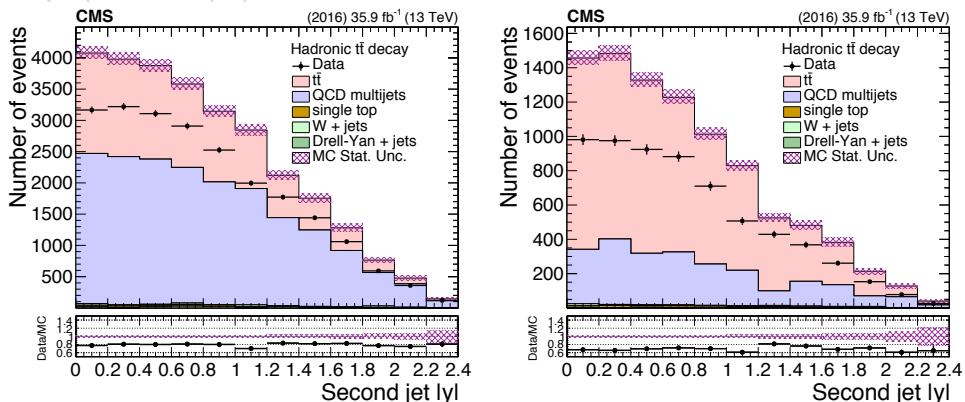


Figure 71: Data vs simulation before (left) and after (right) applying the  $NN > 0.8$  cut for the  $|y|$  of the second jet.

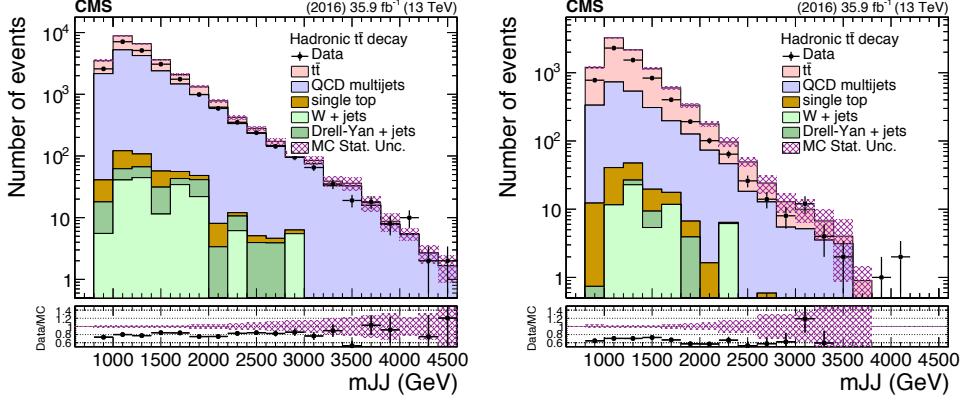


Figure 72: Data vs simulation before (left) and after (right) applying the  $NN > 0.8$  cut for the mass of the  $t\bar{t}$  system.

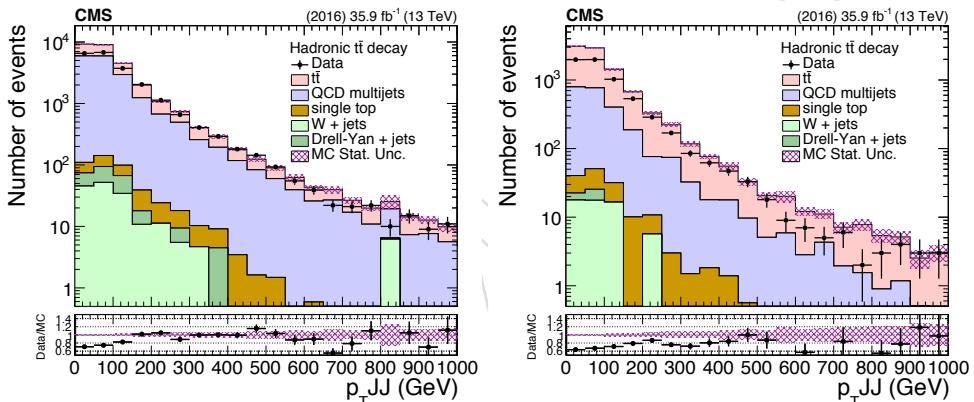


Figure 73: Data vs simulation before (left) and after (right) applying the  $NN > 0.8$  cut for the  $p_T$  of the  $t\bar{t}$  system.

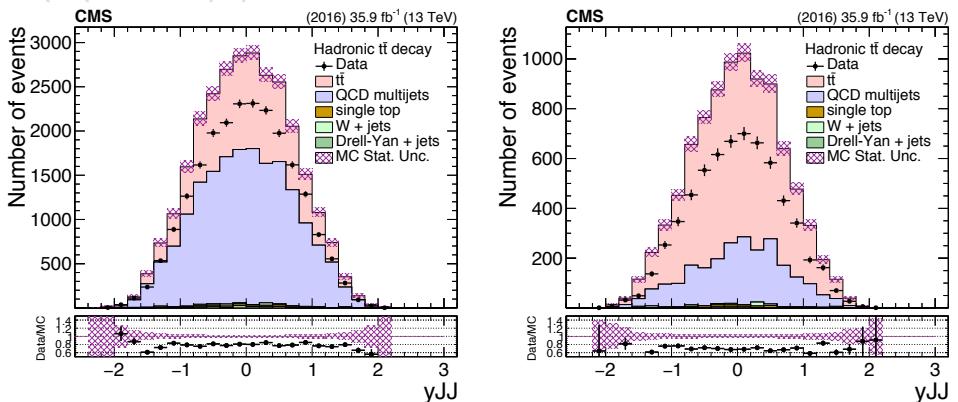


Figure 74: Data vs simulation before (left) and after (right) applying the  $NN > 0.8$  cut for the rapidity of the  $t\bar{t}$  system.

## 603 B NN Input Variables

604 In this section we provide additional comparisons between **data** and **simulation** for the six  
 605 variables used as inputs for the NN training, in three different categories: 0, 1, or 2 top jet  
 606 candidates with a b tagged subjet, prior any cut on the NN output.

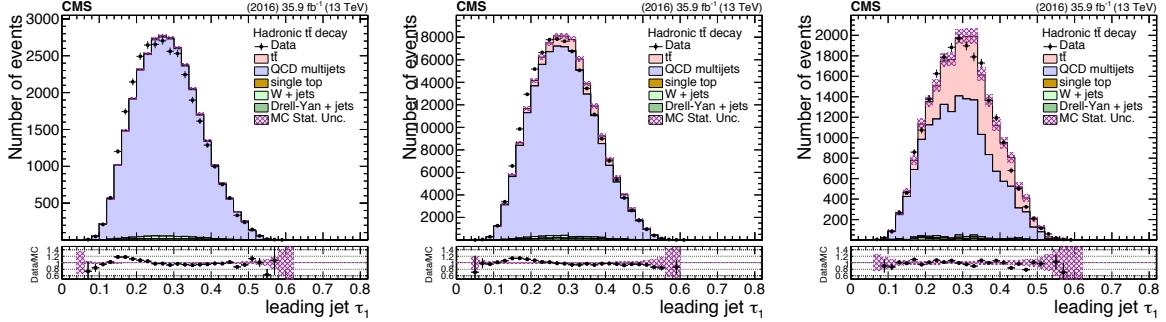


Figure 75: Data vs simulation for the  $\tau_1$  variable of the leading jet after the selection of 0 btagged subjet (left), 1 (middle) and 2 (right).

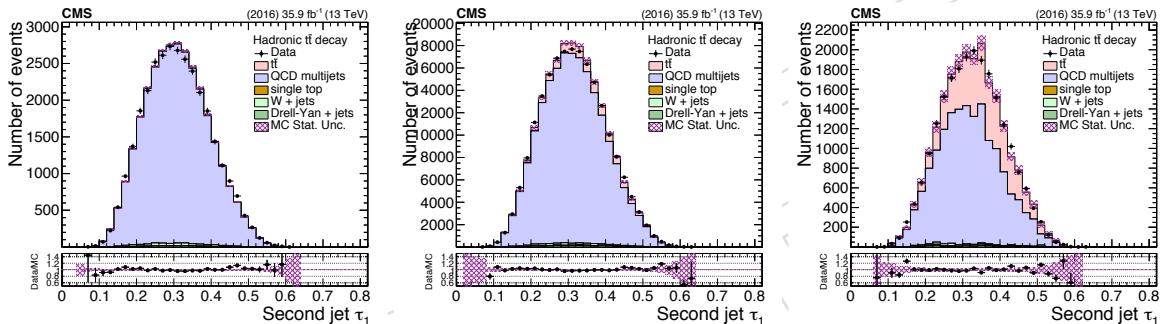


Figure 76: Data vs simulation for the  $\tau_1$  variable of the second jet after the selection of 0 btagged subjet (left), 1 (middle) and 2 (right).

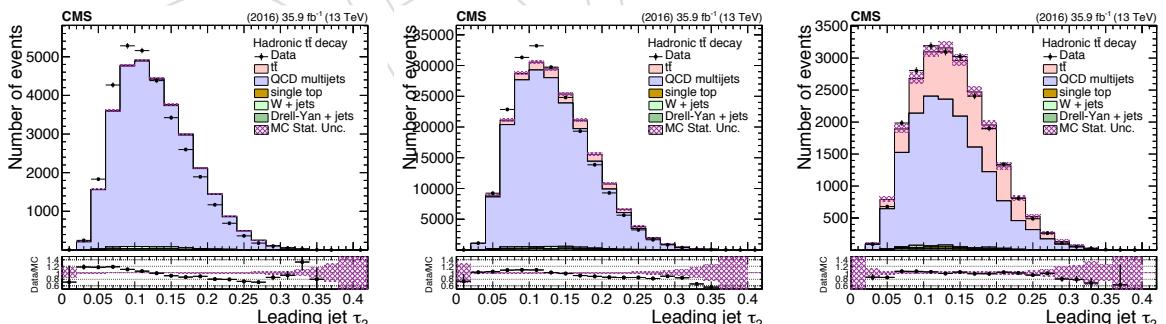


Figure 77: Data vs simulation for the  $\tau_2$  variable of the leading jet after the selection of 0 btagged subjet (left), 1 (middle) and 2 (right).

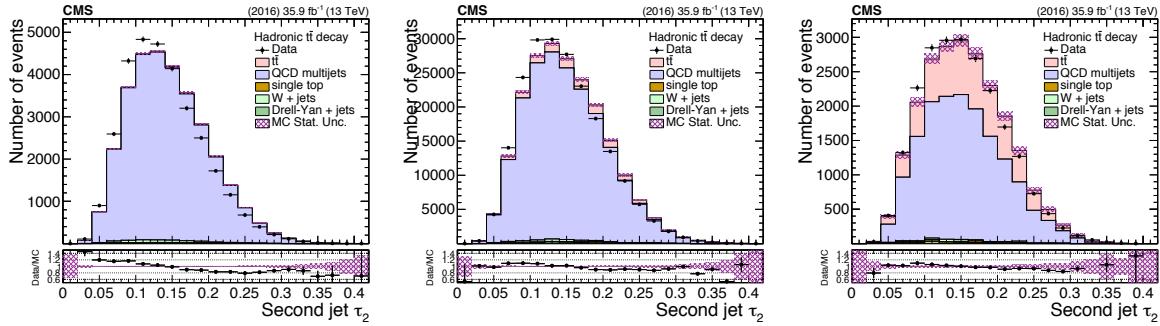


Figure 78: Data vs simulation for the  $\tau_2$  variable of the second jet after the selection of 0 btagged subjet (left), 1 (middle) and 2 (right).

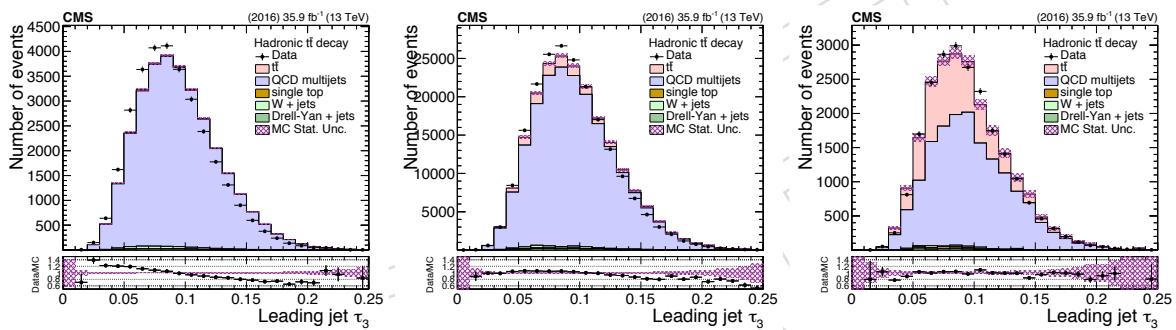


Figure 79: Data vs simulation for the  $\tau_3$  variable of the leading jet after the selection of 0 btagged subjet (left), 1 (middle) and 2 (right).

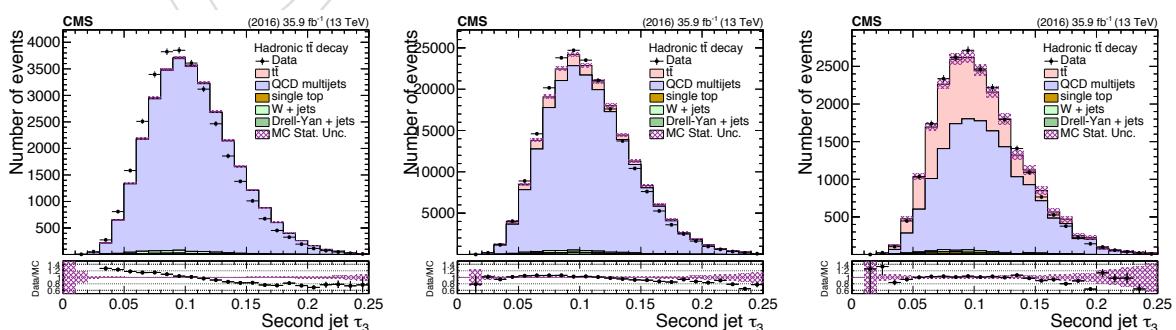


Figure 80: Data vs simulation for the  $\tau_3$  variable of the second jet after the selection of 0 btagged subjet (left), 1 (middle) and 2 (right).

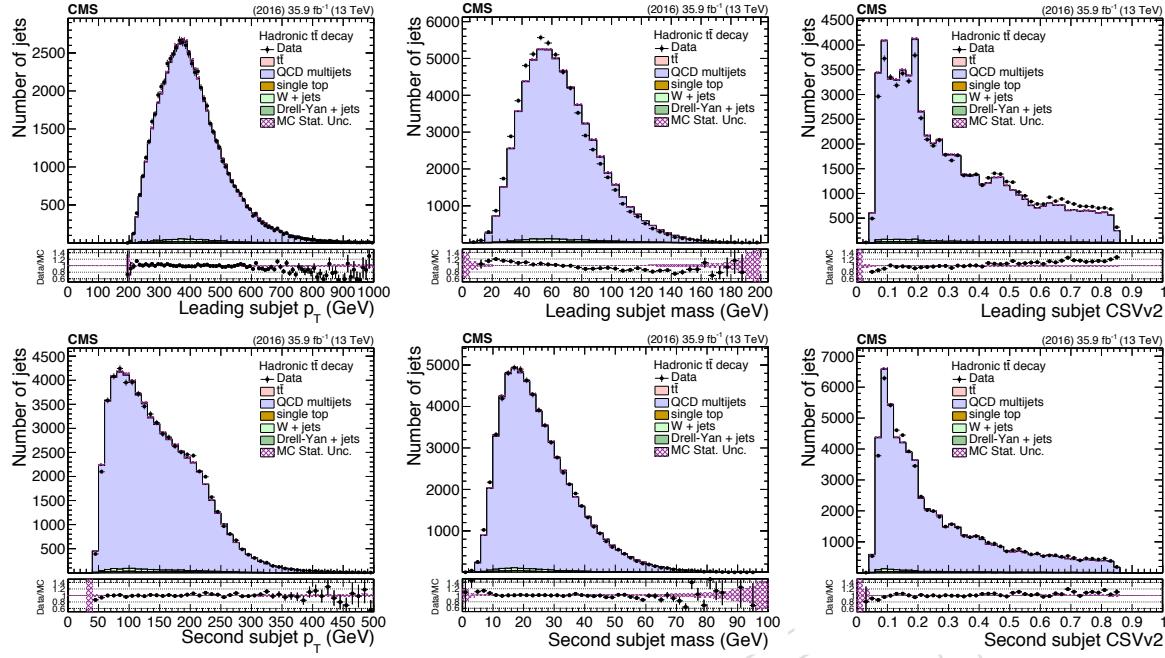


Figure 81: Data vs simulation for the  $p_T$ , mass, and CSVv2 variables for the two subjets of the top candidates after the baseline selection and the requirement of 0 b tagged subjets.

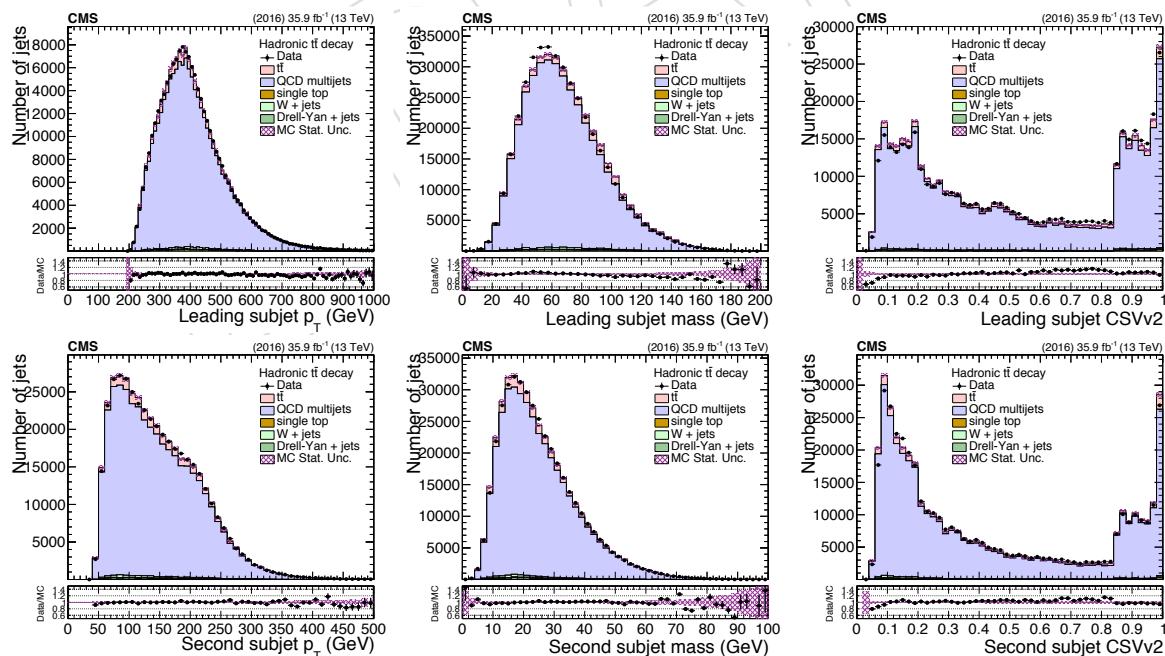


Figure 82: Data vs simulation for the  $p_T$ , mass, and CSVv2 variables for the two subjets of the top candidates after the baseline selection and the requirement of 1 b tagged subjets.

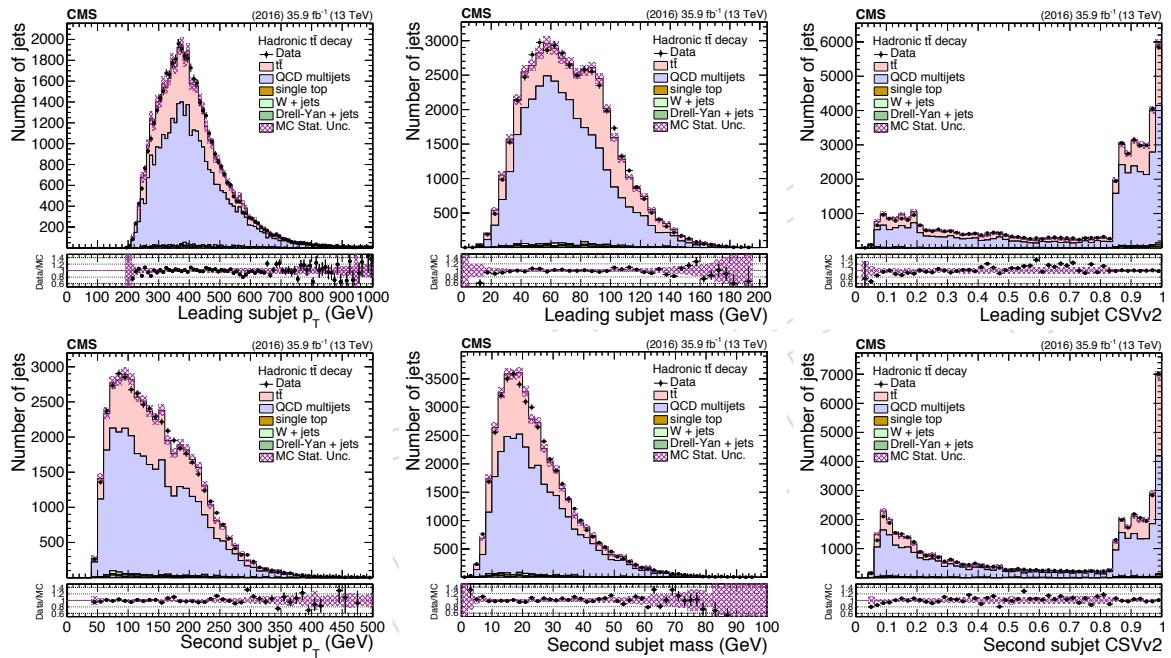


Figure 83: Data vs simulation for the  $p_T$ , mass, and CSVv2 variables for the two subjets of the top candidates after the baseline selection and the requirement of 2 b tagged subjets.

## 607 C Control Distributions

608 In this section we provide additional comparisons between **data** and **simulation**. In what fol-  
 609 lows, we show the various observables in the **control** and **signal** regions.

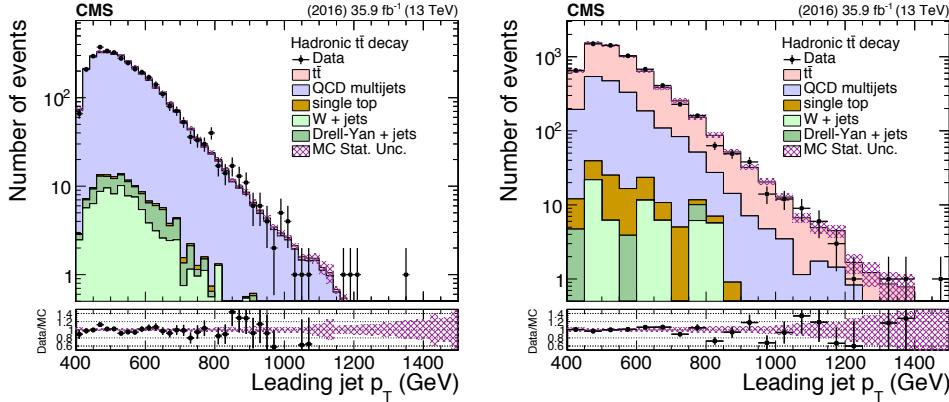


Figure 84: Data vs simulation in the control (left) and signal (right) regions for the  $p_T$  of the two leading jets.

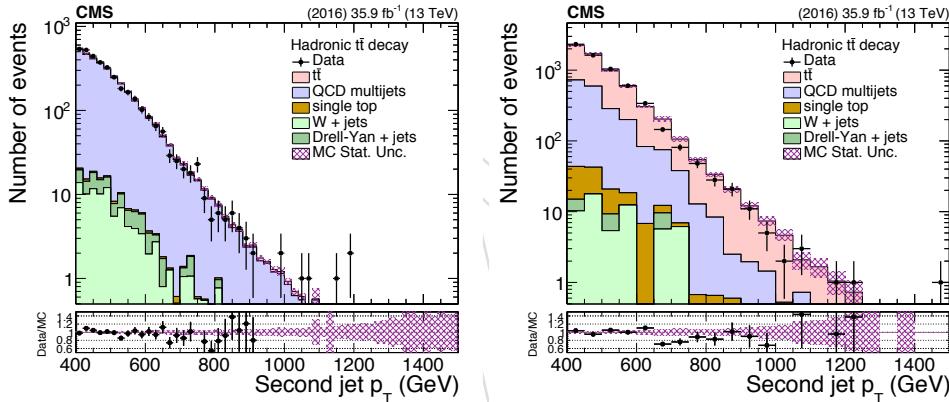


Figure 85: Data vs simulation in the control (left) and signal (right) regions for the  $p_T$  of the two leading jets.

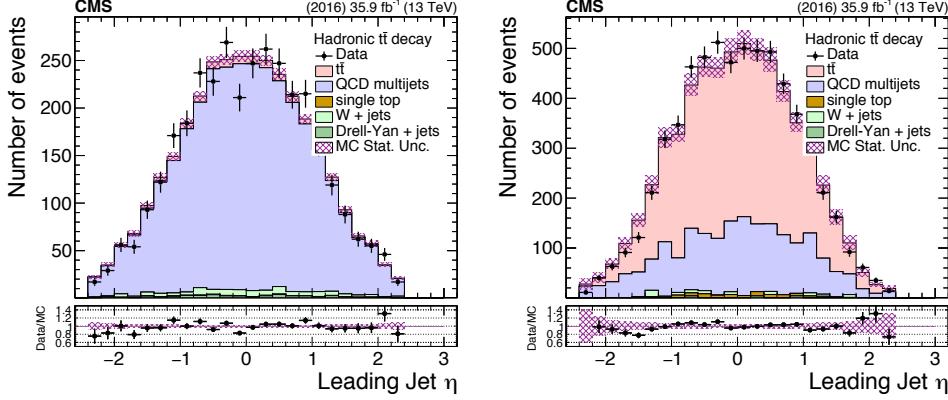


Figure 86: Data vs simulation in the control (left) and signal (right) regions for the  $\eta$  of the two leading jets.

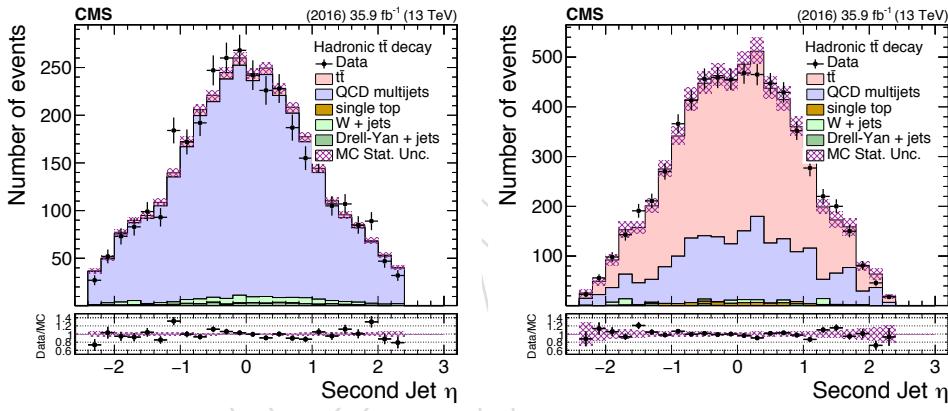


Figure 87: Data vs simulation in the control (left) and signal (right) regions for the  $\eta$  of the two leading jets.

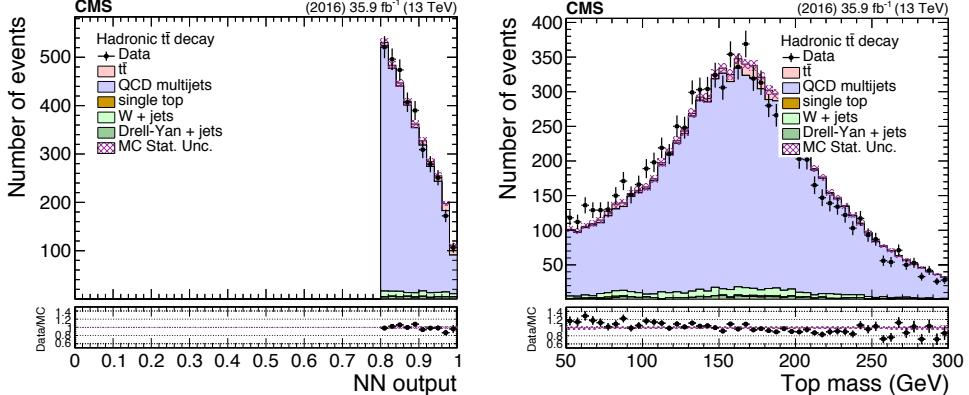


Figure 88: Data vs simulation in the control region for the NN output (left) and the top mass estimator (right).

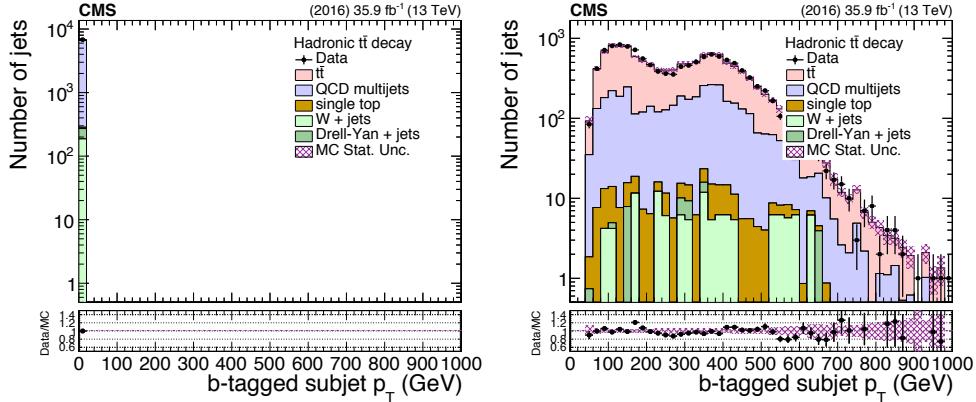


Figure 89: Data vs simulation in the control (left) and signal (right) regions for the  $p_T$  of the b-tagged subjet of the two leading jets.

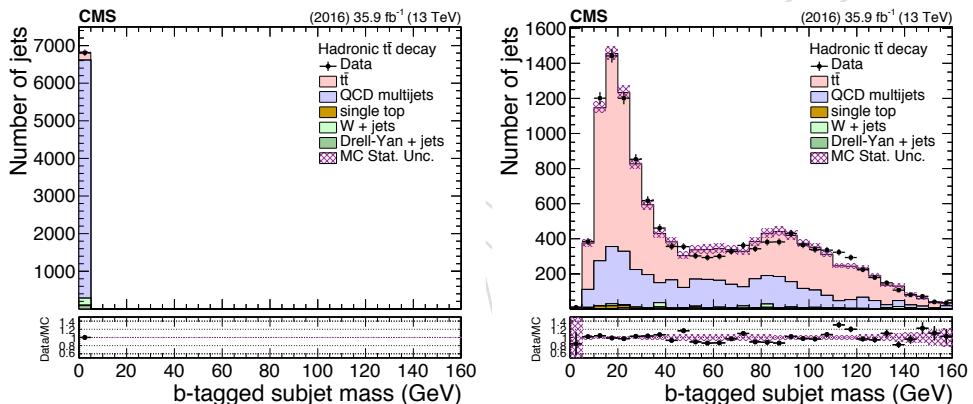


Figure 90: Data vs simulation in the control (left) and signal (right) regions for the mass of the b-tagged subjet of the two leading jets.

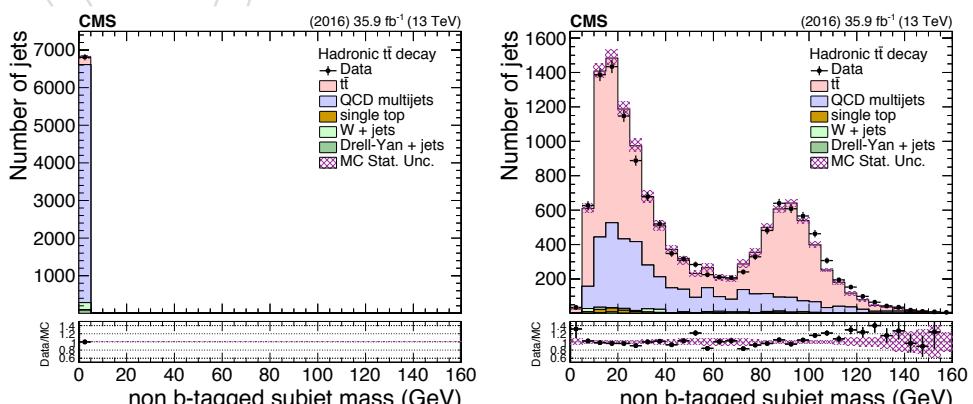


Figure 91: Data vs simulation in the control (left) and signal (right) regions for the mass of the non b-tagged subjet of the two leading jets.

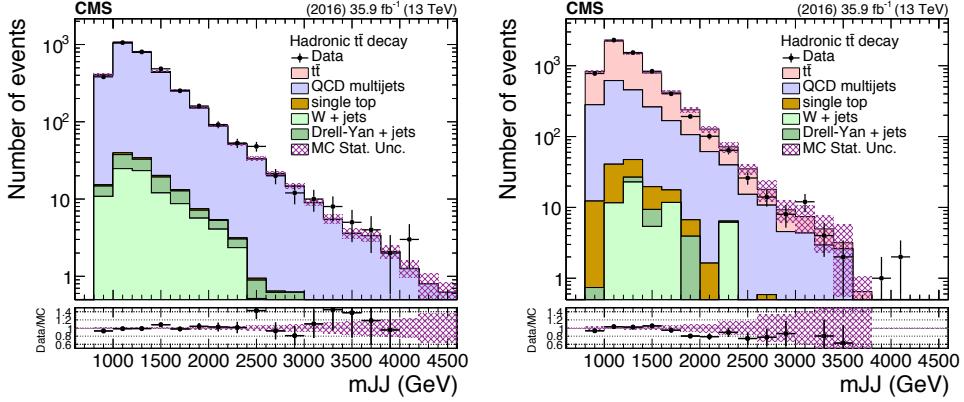


Figure 92: Data vs simulation in the control (left) and signal (right) regions dijet invariant mass.

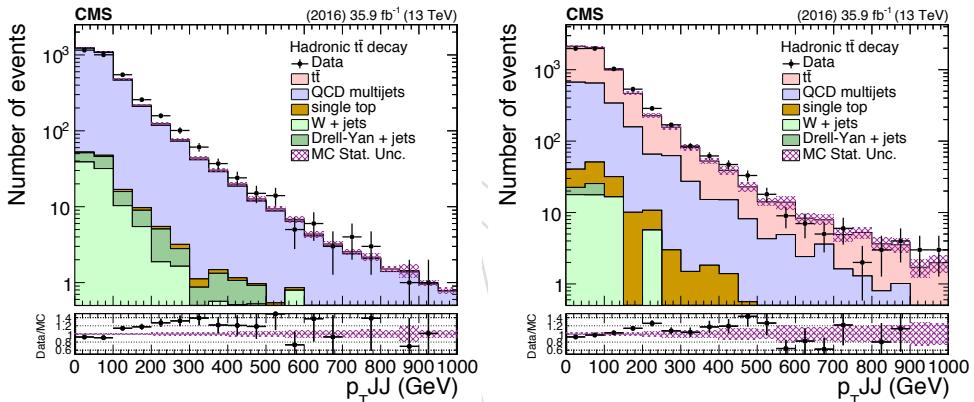


Figure 93: Data vs simulation in the control (left) and signal (right) regions dijet  $p_T$ .

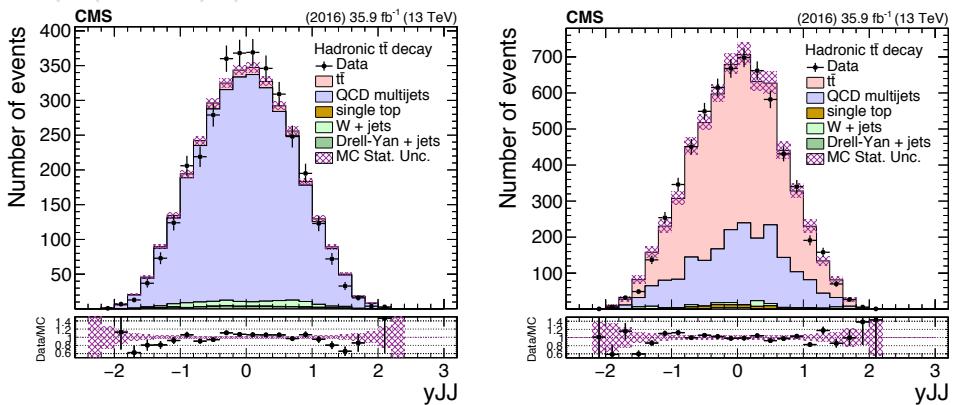


Figure 94: Data vs simulation in the control (left) and signal (right) regions dijet rapidity.

## 610 D Background Sensitivity to Pileup

611 The QCD background shape is taken from the corresponding control region in data, as de-  
 612 scribed in Section 5. Since the data are collected with a prescaled trigger path, the pileup profile  
 613 is different from the one of the signal events (Fig. 1). In order to demonstrate the robustness  
 614 of the QCD prediction between the signal and control regions, we have performed the follow-  
 615 ing test: the data from the control region are split in two subsets according to the number of  
 616 reconstructed vertices ( $< 15$  and  $\geq 15$ ) and we compare in Fig. 95 the distributions of the ob-  
 617 servables used in this analysis. A systematic difference is observed in the softdrop mass of the  
 618 jets, which is expected to be the most sensitive to pileup. This difference is naturally accounted  
 619 for in the extraction of the QCD normalization (Eq. 2) with the unconstrained nuisance param-  
 620 eter  $k_{\text{slope}}$ . In contrast, the shape of the jet and dijet kinematic observables are fully compatible  
 621 within the statistical uncertainty. Therefore, taking also into account that the difference of the  
 622 pileup profile of the two trigger paths is not as extreme as the test conducted here, we conclude  
 623 that no significant bias is introduced by using the QCD shapes from the control region.

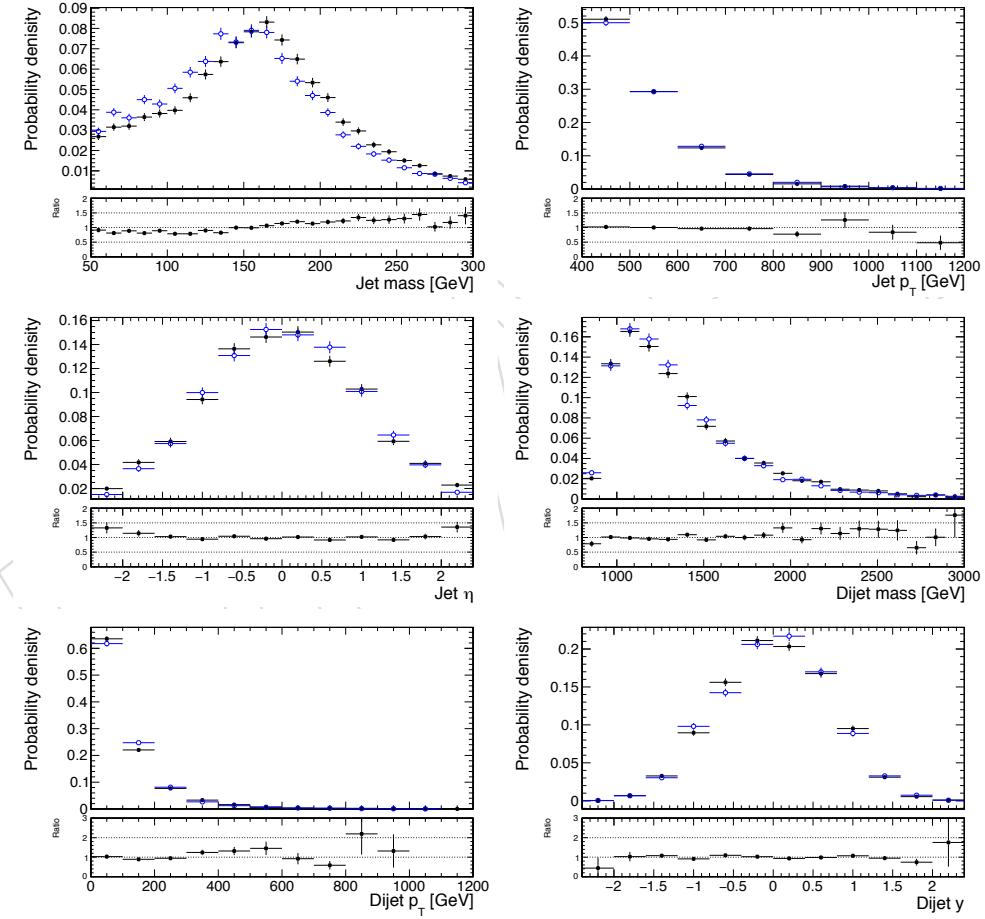


Figure 95: Distributions of jet and dijet observables from low pileup events (black,  $n_{\text{vtx}} < 15$ ) and high pileup events (blue,  $n_{\text{vtx}} \geq 15$ ). The bottom panel shows the ratio of the two.

## 624 E Unfolding Tests

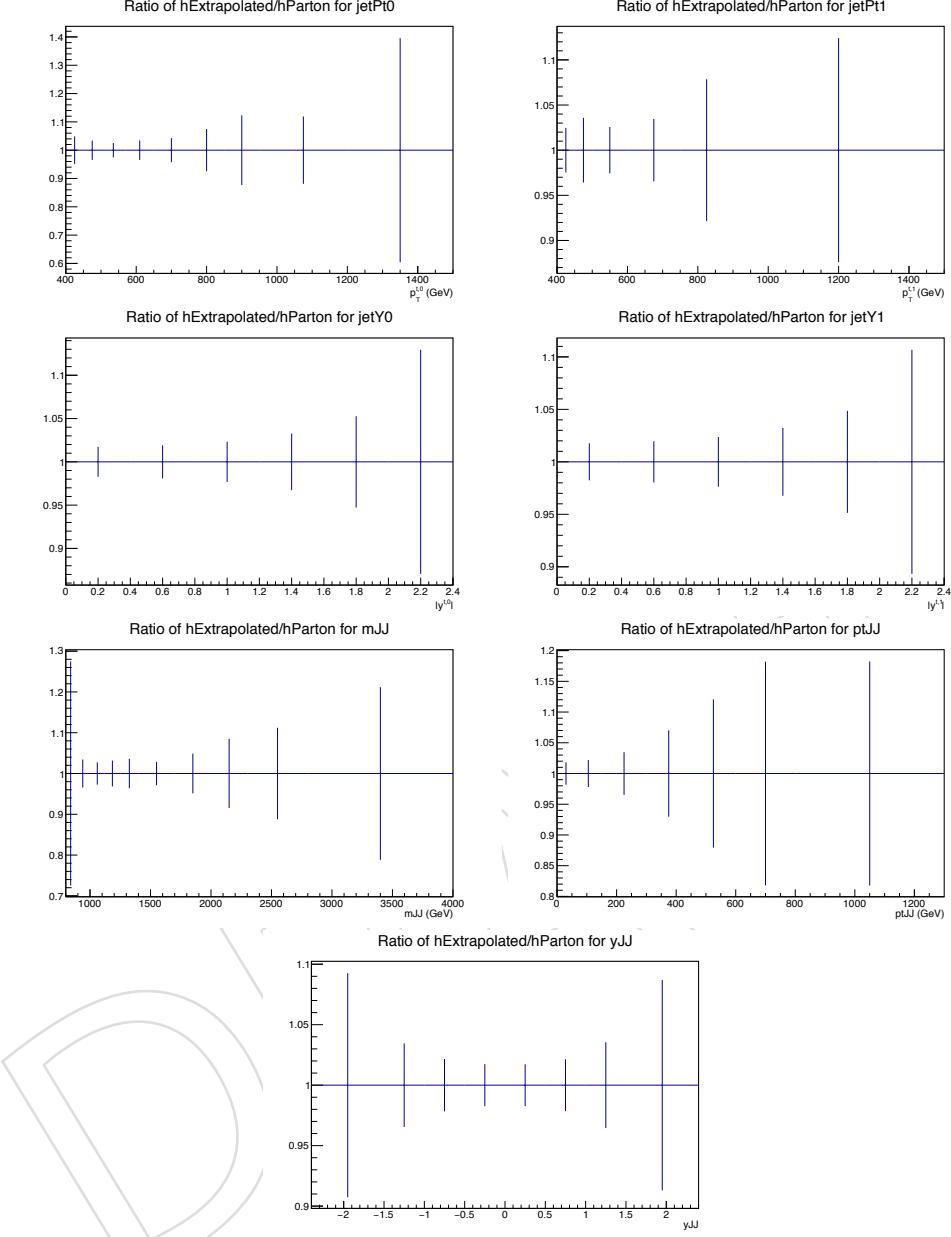


Figure 96: Parton unfolding closure test in simulation for all the reported observables. Ratio of the unfolded over the parton spectra.

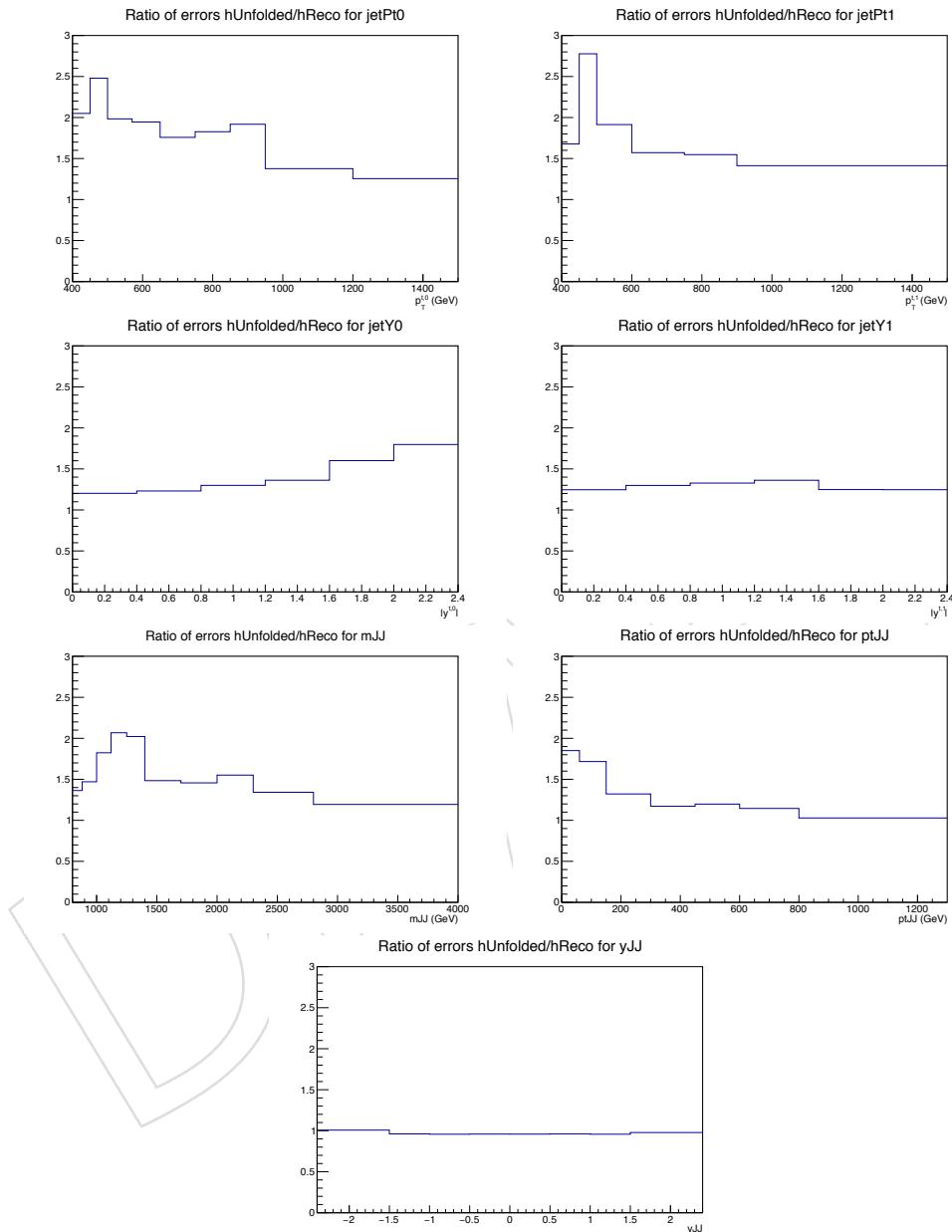


Figure 97: Ratio of the unfolded of the measured relative statistical uncertainties in the simulation (parton level).

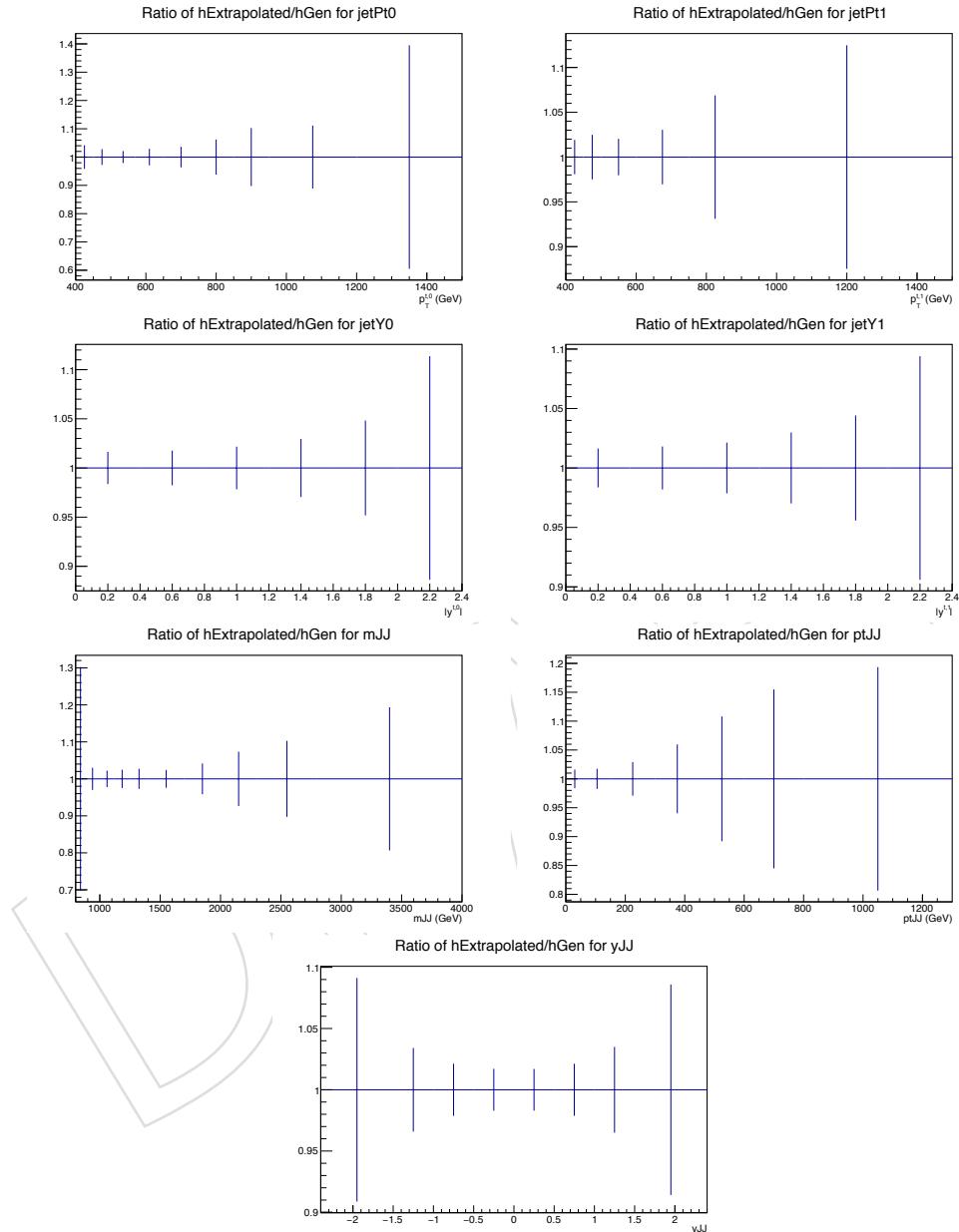


Figure 98: Particle unfolding closure test in simulation for all the reported observables. Ratio of the unfolded over the particle spectra.

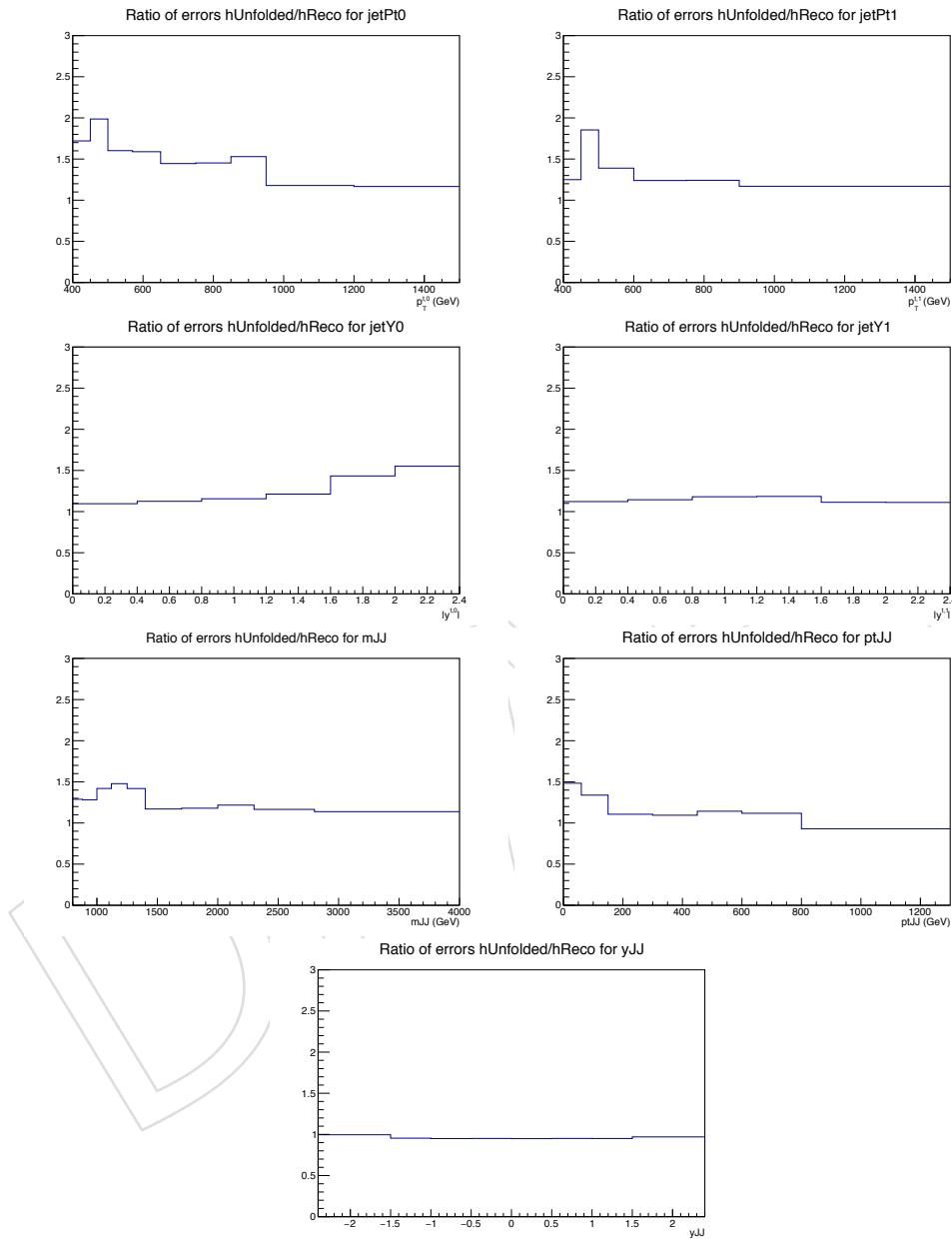


Figure 99: Ratio of the unfolded of the measured relative statistical uncertainties in the simulation (particle level).

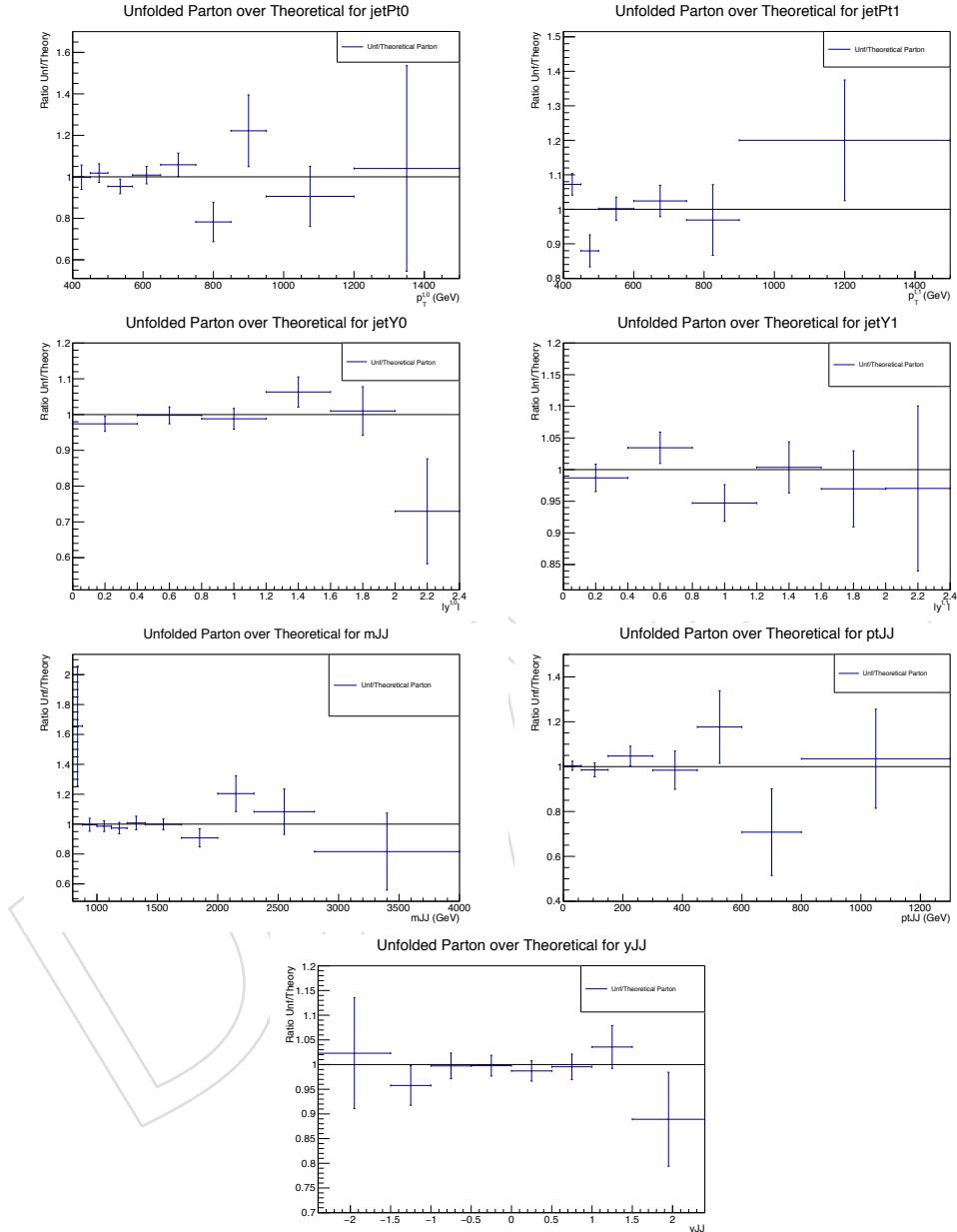


Figure 100: Stability of the unfolding at parton level in the simulation. The response matrix is derived from the first half of the sample and applied to the other half.

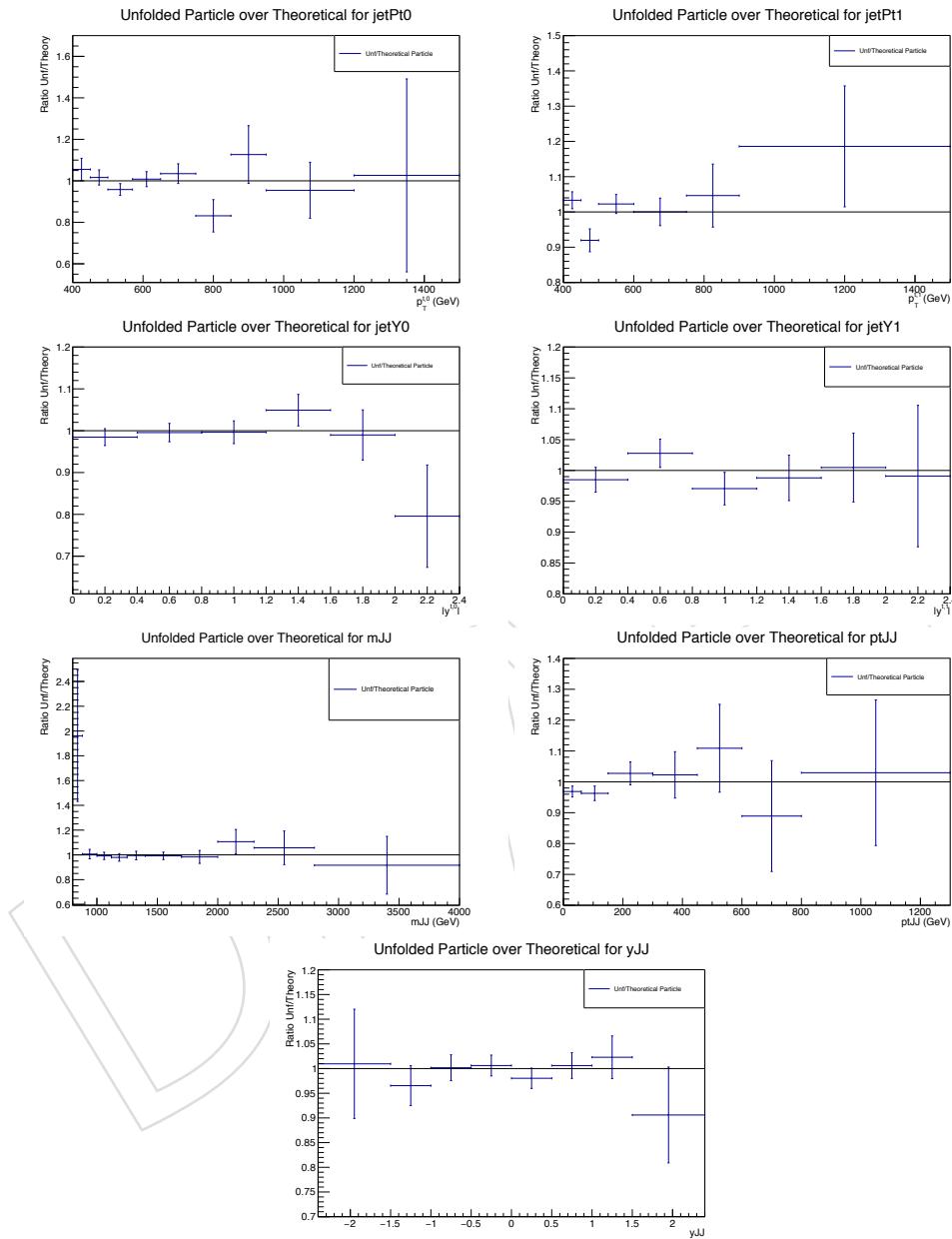


Figure 101: Stability of the unfolding at particle level in the simulation. The response matrix is derived from the first half of the sample and applied to the other half.

## 625 F Simulated b tagging efficiencies

626 In this section we show the **b** tagging efficiency of subjets for each parton in Fig. 102 and the  
 627 purity of the **b** tagged reconstructed subjets in Fig. 103.

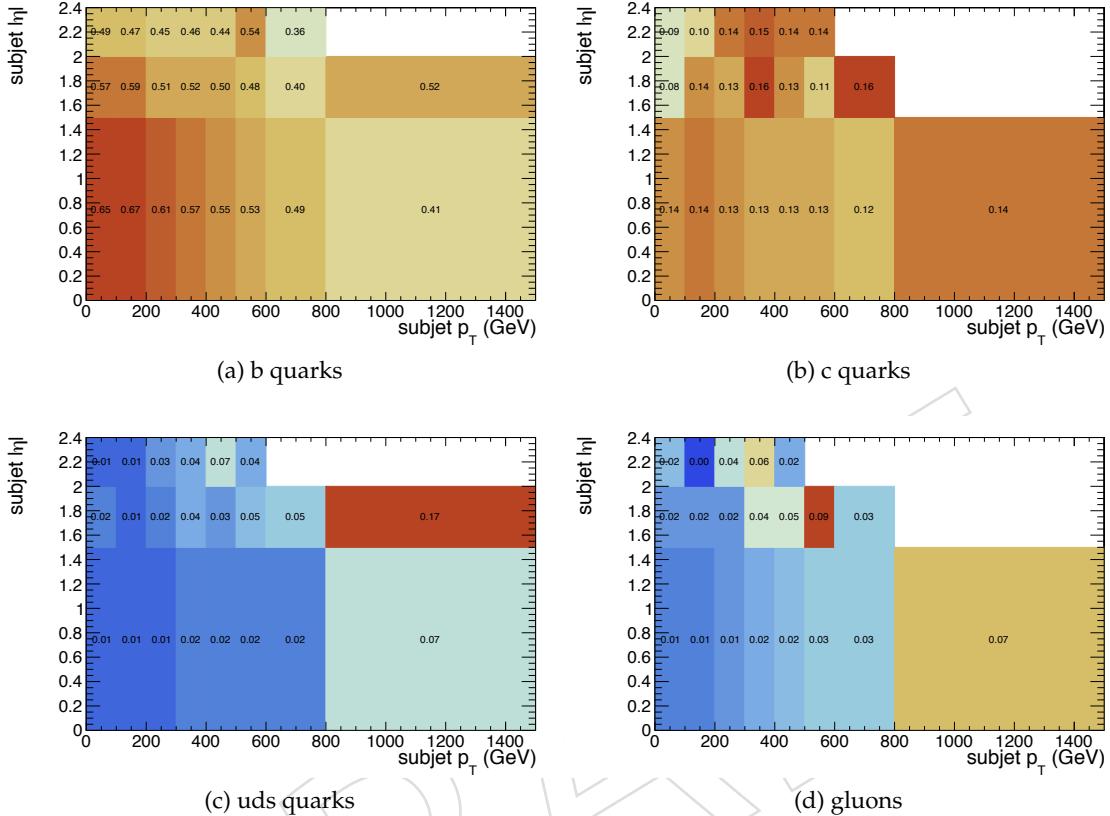


Figure 102: Simulated b tagging efficiency of subjets for different parton types.

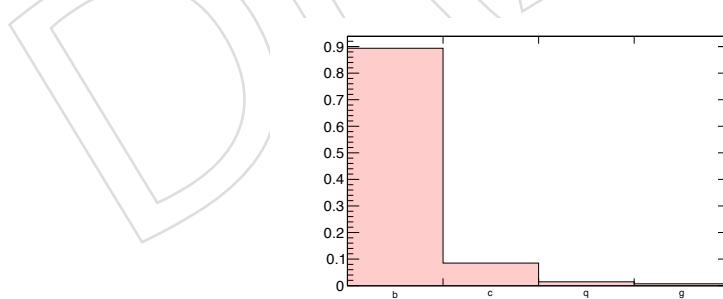


Figure 103: Distribution of the parton flavor for each b tagged subjet.

## 628 G Correlation Matrices

629 In this section we present the correlation matrices between the bins of the unfolded distribu-  
 630 tions.

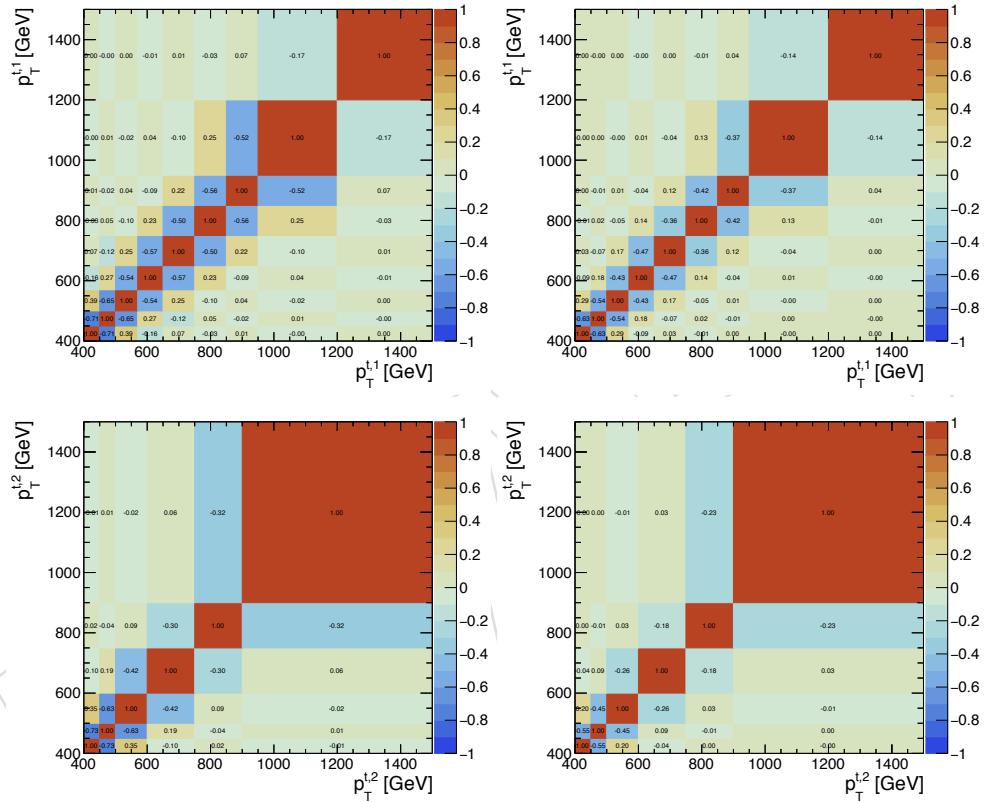


Figure 104: Correlation matrices for the unfolded top  $p_T$  distributions at parton (left) and particle (right) level.

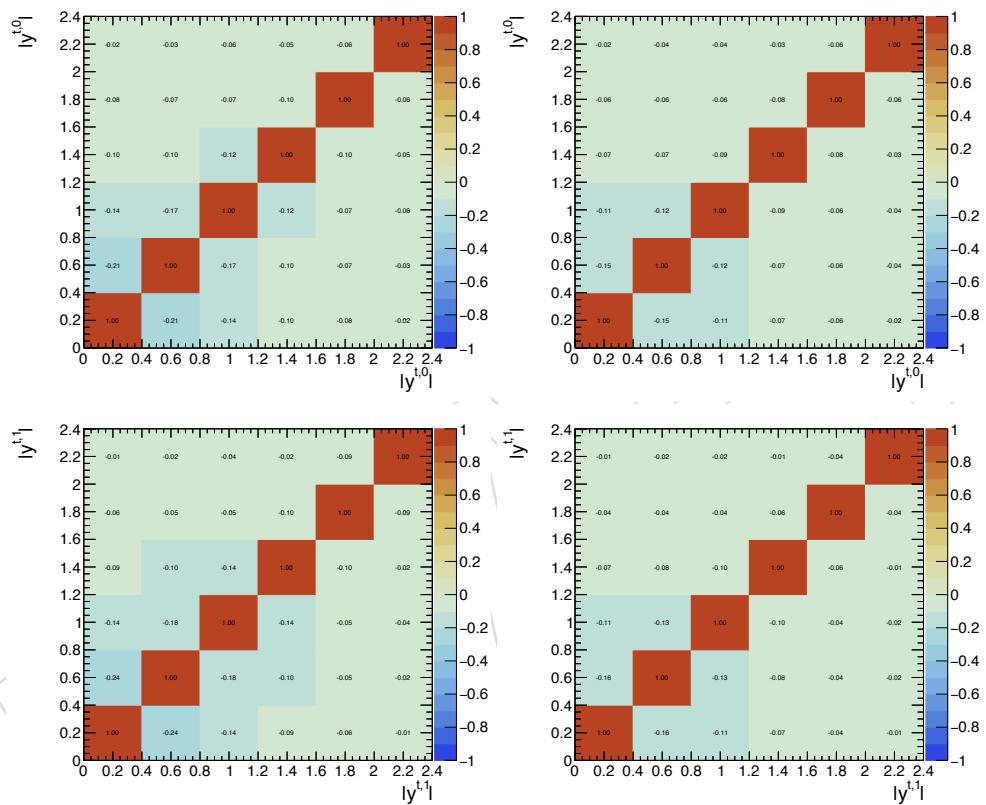


Figure 105: Correlation matrices for the unfolded top  $|y|$  distributions at parton (left) and particle (right) level.

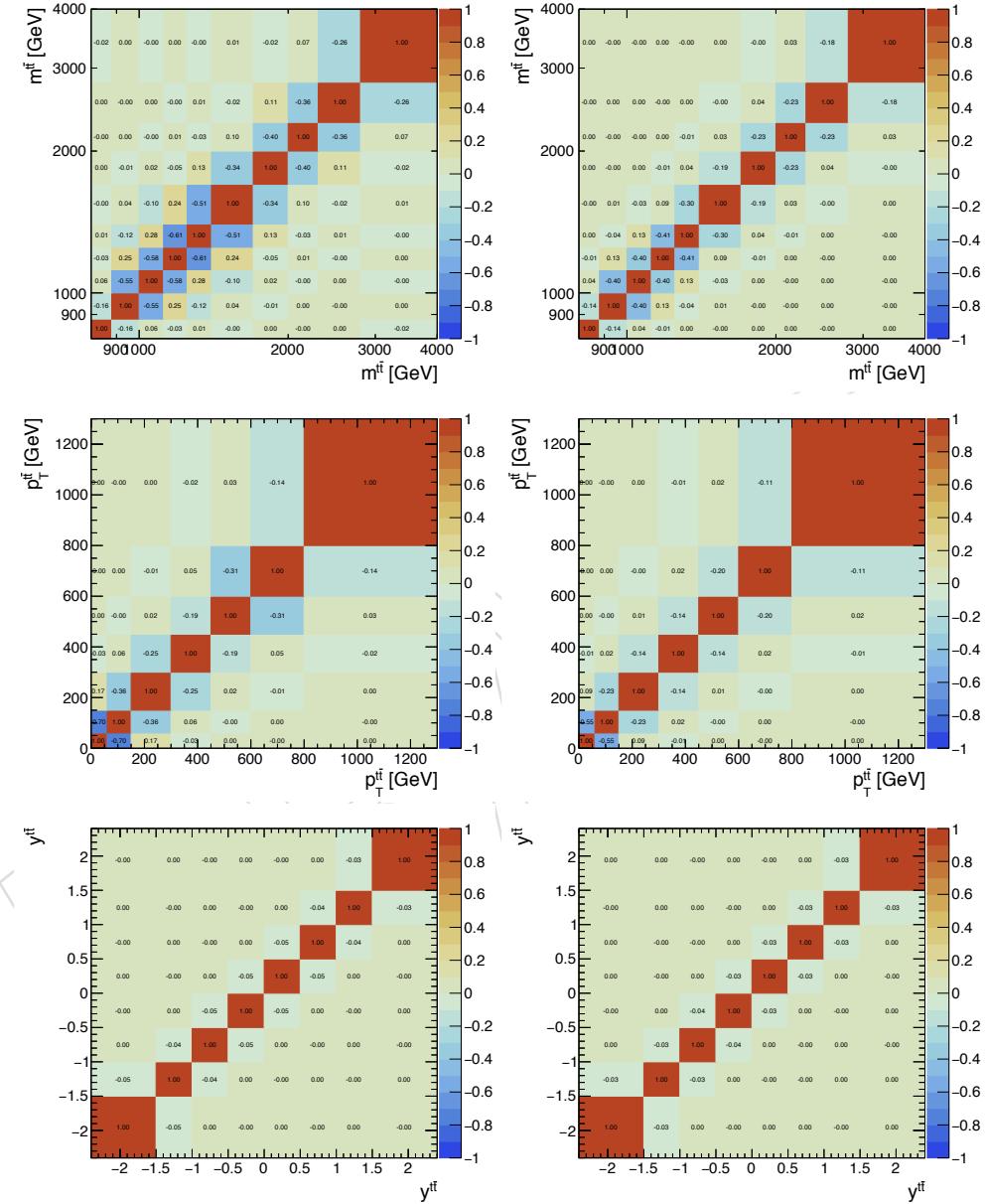


Figure 106: Correlation matrices for the unfolded  $t\bar{t}$  distributions at parton (left) and particle (right) level.