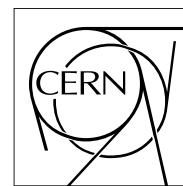


The Compact Muon Solenoid Experiment Analysis Note

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



23 May 2017 (v5, 18 March 2019)

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

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 is presented of the top-anti-top quark production cross section, in the boosted region of the hadronic decay, differentially, as a function of various kinematic variables. The data correspond to an integrated luminosity of 35.9 fb^{-1} from proton-proton collisions at 13 TeV center-of-mass energy and were collected by the CMS detector during the 2016 LHC run. The results are presented in the fiducial region and unfolded to the particle and parton levels.

1 Contents

2	1	Introduction	2
3	2	Samples	2
4	2.1	Data	2
5	2.2	Simulation	3
6	3	Trigger	4
7	4	Reconstruction and Selection	6
8	4.1	Object Reconstruction	6
9	4.2	Multivariate Discriminant	6
10	4.3	Selection	9
11	4.4	Reconstruction of Particle Jets	11
12	5	Data vs Monte Carlo	11
13	6	Background	15
14	7	Signal Extraction	18
15	8	Systematic Uncertainties	21
16	9	Fiducial Measurement	29
17	10	Unfolded Measurement	32
18	10.1	Parton Level	32
19	10.2	Particle Level	38
20	11	Summary	44
21	A	Control Distributions	46
22	B	Background Sensitivity to Pileup	51
23	C	Unfolding Tests	52

DRAFT

24 1 Introduction

25 The top quark is fundamental in the standard model (SM) due to its large mass, and the precise knowledge of its properties is critical for the overall understanding of the theory. Measurements of the top quark pair ($t\bar{t}$) production cross section confront the quantum chromodynamics (QCD) predictions, and have the potential to constrain the QCD parameters, while being sensitive to physics beyond the SM. Also, the $t\bar{t}$ process is a dominant SM background to searches for new physics phenomena, and therefore its precise knowledge is essential for new discoveries.

32 The large $t\bar{t}$ yield expected in proton-proton (pp) collisions at the LHC allows to perform more precise measurements of the $t\bar{t}$ production rate, in extended parts of the phase space, and, more importantly, differentially, as a function of the $t\bar{t}$ kinematic properties. The measurement presented here is performed with pp collision data at $\sqrt{s} = 13$ TeV, collected with the CMS detector in 2016. The differential $t\bar{t}$ cross section, measured in the all-jets decay channel, is reported at parton and particle level, and compared to various QCD predictions.

38 In the all-jets channel each W boson arising from the top quark decays into a pair of light quarks. As a result, the final state consists of at least six partons (more are possible due to initial- and final-state radiation), two of which are b quarks. Despite the large number of permutations, it is possible to reconstruct fully the $t\bar{t}$ decay products, in contrast to the leptonic channels, where the presence of one or two neutrinos makes the full event interpretation ambiguous. Due to the presence of multiple jets in the final state, the measurement of the $t\bar{t}$ cross section in the all-jets channel involves larger uncertainties (from the jet energy scale and the b-tagging) than the leptonic channels, but it is complementary and unique in its ability to reconstruct fully the $t\bar{t}$ system.

47 The analysis presented in this Note follows the boosted reconstruction approach where the p_T of the produced top quarks is high and the decay products are merged. In this case the final state consists of at least two wide, high- p_T jets. This reconstruction method allows to perform the measurement in the top p_T range from 400 GeV to 1.5 TeV.

51 This note is organized as follows: section 2 describes the data and Monte Carlo samples used in the analysis and section 3 discusses the triggers and the measurement of the corresponding efficiency. Then, section 4 presents the details of the jet reconstruction and the event selection. A comparison between data and simulation is reported in section 5. The background estimation technique is presented in section 6. The signal extraction and the systematic uncertainties are discussed in sections 7 and 8, respectively, while the results are presented in sections 9 and 10. Finally, section 11 summarizes the results.

58 2 Samples

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

60 2.1 Data

61 The collision events used for the measurement of the $t\bar{t}$ cross section have been collected with the triggers described in Section 3 that are part of the JetHT primary dataset. Table 1 lists the samples that correspond to different data-taking eras and have been reconstructed in the 03Feb2017 reprocessing. The good Run and luminosity section list used is contained in the certification file Cert_271036-284044_13TeV_23Sep2016ReReco_Collisions16_JSON.txt.
 66 The total integrated luminosity of the analyzed data is 35.922 fb^{-1} .

67 In addition to the data from the JetHT dataset, we have used the SingleMuon dataset for the
 68 measurement of the trigger efficiency. The eras and run ranges are the same as the ones in
 69 Table 1.

Table 1: Data samples.

Sample	Run range	Luminosity (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

70 2.2 Simulation

71 The Monte Carlo samples include the simulation of $t\bar{t}$ production, as well as the simulation
 72 of background processes. In the first part of Table 2 we show the signal samples, while in
 73 the second part we show the background ones that include QCD multijet production,
 74 associated production of vector bosons (W^\pm, Z) with jets, and single-top production. The list
 75 shows the total number of events analyzed, including all samples of the same kind (nomi-
 76 nal, extensions, backup). The reconstruction of the Monte Carlo samples belongs to the era
 77 RunIIISummer16MiniAODv2-PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6-v1 and
 78 we have used the MINIAODSIM data tier.

Table 2: Monte Carlo samples.

Sample	Events ($\times 10^6$)	$\sigma (\text{pb})$
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	1.712×10^6
QCD_HT300to500.TuneCUETP8M1_13TeV-madgraphMLM-pythia8, ext1	54.5	3.477×10^5
QCD_HT500to700.TuneCUETP8M1_13TeV-madgraphMLM-pythia8, ext1	62.3	3.21×10^4
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

79 3 Trigger

80 The trigger path employed for the collection of signal events uses single-jet L1 seeds that re-
 81 quire the presence of a jet with $p_T > 180 \text{ GeV}$. At HLT jets are reconstructed from (online)
 82 particle flow candidates using the anti-kt algorithm with distance parameter $R = 0.8$ and their
 83 mass, after trimming of soft particles, must be greater than 30 GeV . Interesting events are re-
 84 quired to have at least two such jets with $p_T > 280(200) \text{ GeV}$ for the leading (trailing) one.
 85 Finally, at least one of the two jets should be tagged as a b-jet, using the online CSV algorithm.
 86 The aforementioned trigger path ran unprescaled for the duration of the 2016 run, collecting an
 87 integrated luminosity of 35.5 fb^{-1} . A second, prescaled, path, using the same L1 seed, was also
 88 employed, with identical kinematic requirements but no b-tagging cut, which ran in parallel
 89 and collected an integrated luminosity of 1.67 fb^{-1} . This path is used for the selection of a
 90 control QCD sample, as described later. All the triggers described above are summarized in
 91 Table 3. Finally, it should be noted that the pileup profile of the prescaled control trigger is
 92 shifted to lower number of interactions (Fig. 1), because the path tended to collect more data
 93 towards the end of the fills when the instantaneous luminosity was lower.

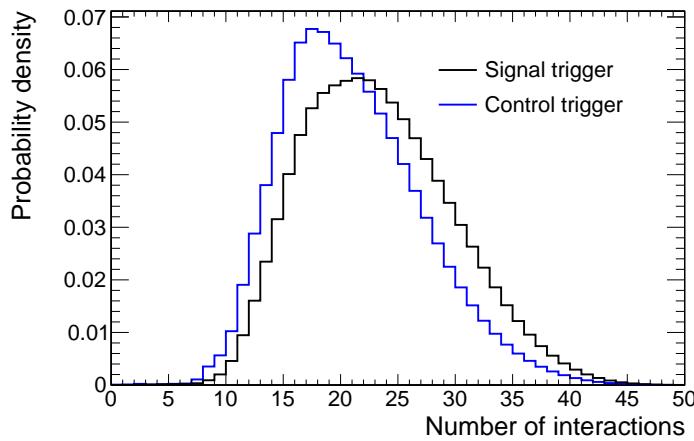


Figure 1: Pileup profile for the signal and control trigger paths.

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

94 The efficiency of the signal trigger path is measured with respect to an orthogonal path that
 95 requires the presence of an isolated muon with $p_T > 27 \text{ GeV}$ (HLT_IsoMu27). Figure 2 shows
 96 the trigger efficiency as a function of the second jet p_T in events with at least two reconstructed
 97 jets, with at least one of them containing a b-tagged subjet. For details about the offline recon-
 98 struction see Section 4. The efficiency measured in data is compared to the simulated efficiency,
 99 showing an excellent agreement.

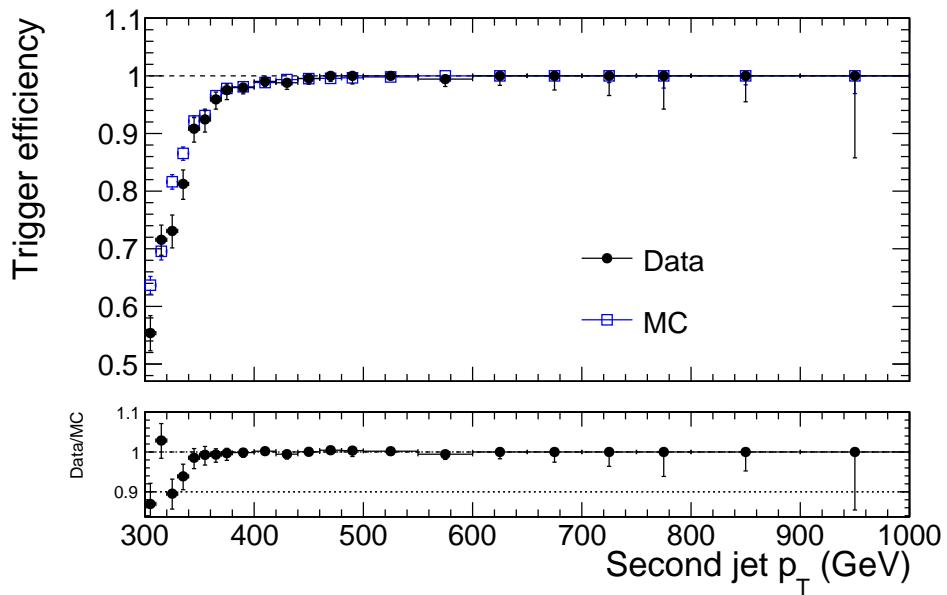


Figure 2: Trigger efficiency for the signal path of the analysis.

100 4 Reconstruction and Selection

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

103 4.1 Object Reconstruction

104 The leptons (**muons**, **electrons**) used in the analysis come from the default reconstructed col-
 105 lections in CMS samples ('**slimmedMuons**' and '**slimmedElectrons**') and must have
 106 $p_T > 20 \text{ GeV}$. **Muons** are required to pass the **medium ID working point** and **electrons** should
 107 pass the **tight working point**, while both lepton types should have a relative mini-isolation less
 108 than 0.1.

109 **Jets** are reconstructed from **particle-flow (PF)** candidates that have undergone charged-hadron
 110 **subtraction (CHS)** in order to **suppress the contribution from multiple soft interactions in the**
 111 **same bunch crossing (pileup)**. The momentum 4-vectors of the PF candidates are clustered
 112 with the **anti-kt algorithm** with **distance parameter $R = 0.8$** , hence we refer to AK8 PFCHS
 113 jets. These jets are required to pass the **tight jet ID**. Furthermore, an algorithm is run to **identify**
 114 **subjets with distance parameter $R = 0.4$ within the AK8 jets**. Finally, the **soft-drop** technique
 115 is used in order to **evaluate the mass of the AK8 jet with suppressed pileup contribution**. For
 116 all the aforementioned algorithms we have used the default definitions in the CMS software
 117 and the default collections in the MINIAOD data and Monte Carlo samples. For the energy
 118 calibration of the AK8 jets we have used the recommended corrections by the JME POG in the
 119 80X_dataRun2_2016SeptRepro_v7 and 80X_mcRun2_asymptotic_2016_TrancheIV_v8 global
 120 tags.

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

125 Since **leptons are also reconstructed as jets**, we perform cross cleaning by **removing** each iden-
 126 **tified lepton** from the jet collection with geometrical matching in the $\eta - \phi$ space: if a jet has
 127 $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.4$ from any accepted lepton candidate it is **removed** from the jet
 128 collection.

129 4.2 Multivariate Discriminant

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

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

133 where N denotes the reconstructed candidate subjets and k runs over the constituent particles
 134 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-
 135 didate subjet i and the particle k. The variable R is the characteristic jet radius ($R = 0.8$ in our
 136 case).

137 Since the signal final state ($t\bar{t}$ hypothesis) involves two jets with a top-quark decay signature,
 138 we consider the N-subjettiness variables $\tau_{1,2,3}$ for both leading jets in the event, resulting in six

discriminating variables per event. The distributions of these variables in $t\bar{t}$ and QCD events are shown in Fig. 3, indicating that there is significant separating power in them. The linear correlation between the variables is shown in Fig. 4. In order to exploit fully the differences of the N-subjettiness variables we turn to multivariate discriminants, implemented in the TMVA package. As a baseline we consider a simple Fisher discriminant, which is a linear combination of the variables. Then, a feed-forward neural network (NN) is constructed to enhance the discriminating power quantified by the integral of the receiver-operator-characteristic (ROC) curve. The network (Fig. 5) is composed of two hidden layers with 16 and 4 nodes, respectively, each of one using a sigmoid activation function. The distribution of the NN output and the convergence of its training are shown in Fig. 6. As shown in Fig. 7, the NN is clearly better than the simple Fisher discriminant and it also outperforms slightly a Boosted Decision Tree composed of 500 trees and trained with the Gradient Boost method with shrinkage parameter equal to 0.1. More complicated NN architectures have not been found to improve the performance.

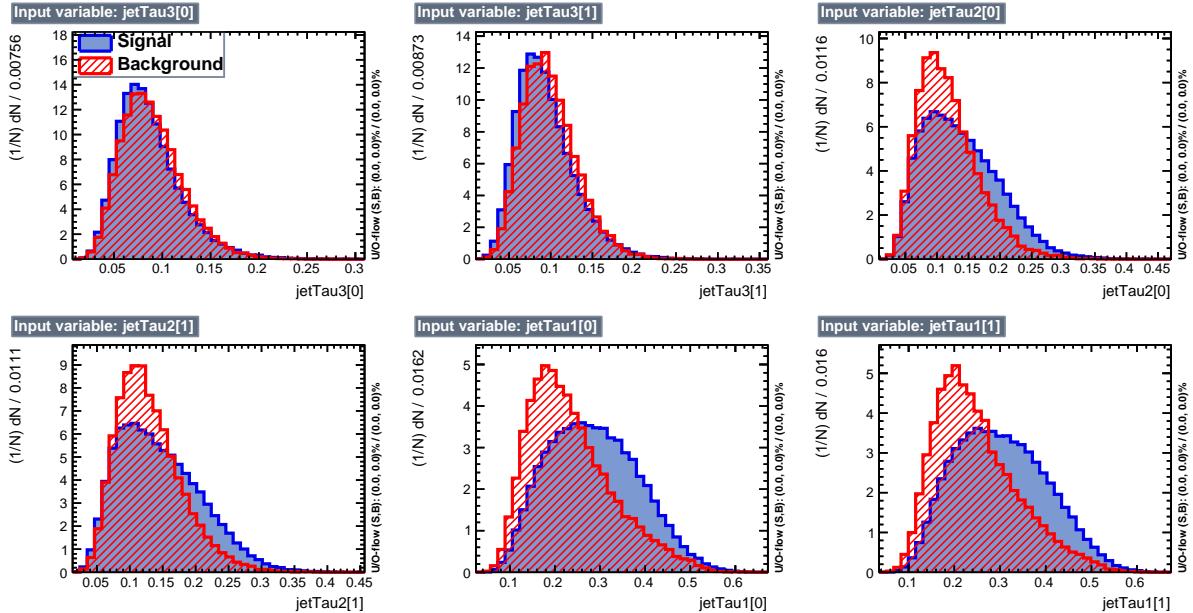


Figure 3: Discriminating variables used for the separation of $t\bar{t}$ from QCD events.

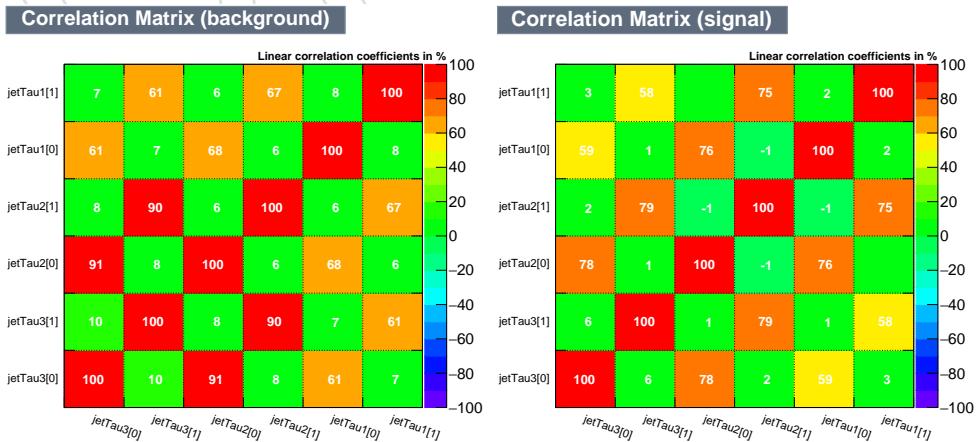


Figure 4: Linear correlation between the discriminating variables in background (left) and signal (right) events.

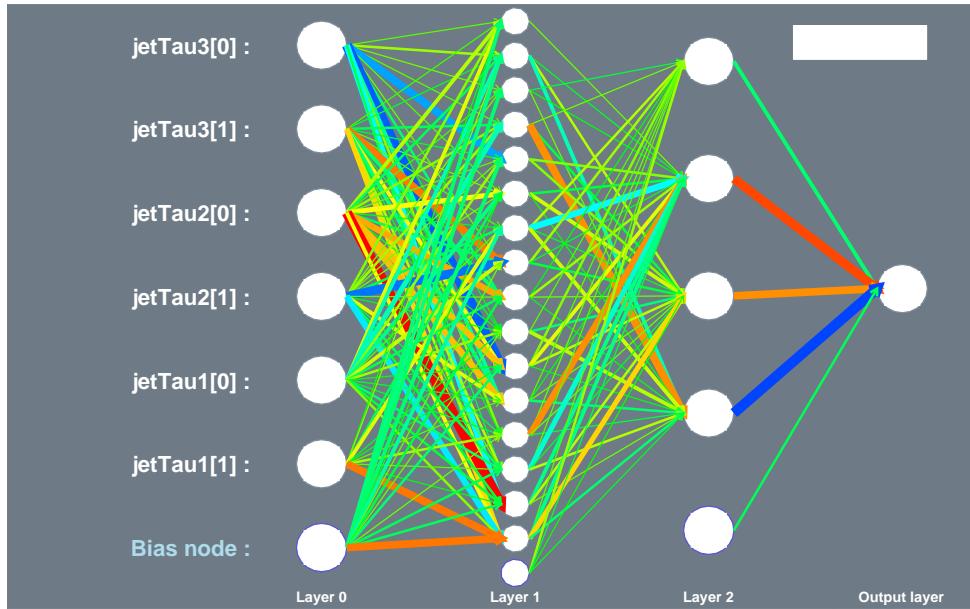


Figure 5: Neural network architecture.

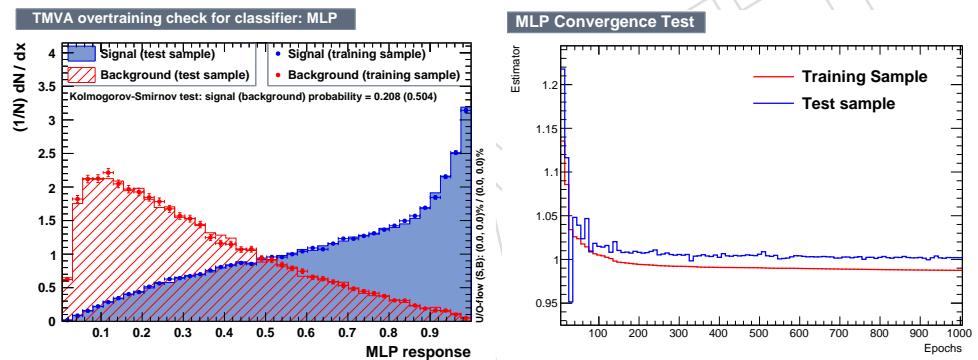


Figure 6: Output of the NN (left). Convergence of the NN training (right).

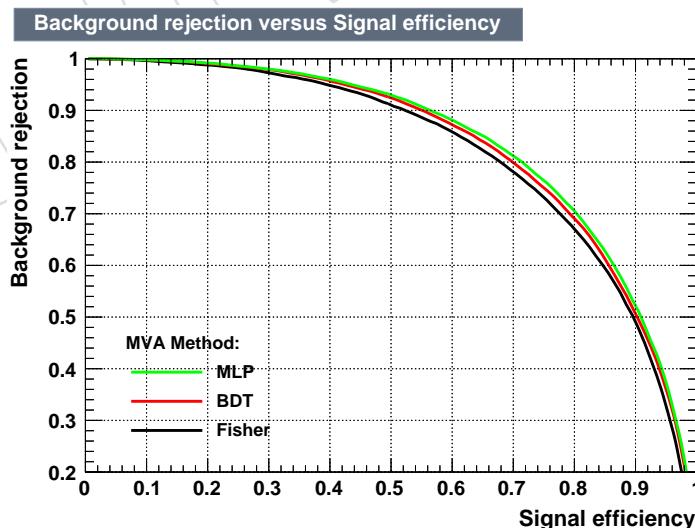


Figure 7: Performance of the multivariate discriminants in the form of a receiver-operator-characteristic curve.

4.3 Selection

The baseline selection, summarized in Table 4 is common for all regions used in the analysis and it requires at least two jets in the event with $p_T > 400 \text{ GeV}$ and softDrop masses in the range $(50, 300) \text{ GeV}$. Also, a lepton veto is applied in order to minimize the probability to select leptonic top decays. Then, on top of the baseline selection, we define four specific regions (Table 4) based on the NN output, the jets' softDrop masses, and the number of b-tagged subjets in each jet, that serve different analysis purposes. The signal region (SR) is where we perform the differential measurements and it requires both jets to have a b-tagged subjet, a tighter selection on the jets' masses, and a high mva value. Figure 8 shows the top decay mode of the selected $t\bar{t}$ events, where more than 95% come from the hadronic channel. The QCD control region (CR) is the same as the SR but with the b-tagging requirement reverted (the jets should not contain a b-tagged subjet) and it is used to get from data the shape of the QCD background for each variable of interest. Then, we employ the signal region A (SR_A), which is used to determine the normalization of the QCD background, and finally the signal region B (SR_B), which is used to constrain some of the signal modelling uncertainties. Figure 9 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_{jets}	> 1
N_{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	$\text{Base+NN} > 0.8 + \text{cat.} = 2 + m_{SD}^{\text{jet}1,2} \in (120, 220) \text{ GeV}$	signal region
SR_A	signal	$\text{Base+NN} > 0.8 + \text{cat.} = 2$	QCD fit region
SR_B	signal	$\text{Base+cat.} = 2 + m_{SD}^{\text{jet}1,2} \in (120, 220) \text{ GeV}$	signal systematics region
CR	control	$\text{Base+NN} > 0.8 + \text{cat.} = 0 + m_{SD}^{\text{jet}1,2} \in (120, 220) \text{ GeV}$	QCD control region

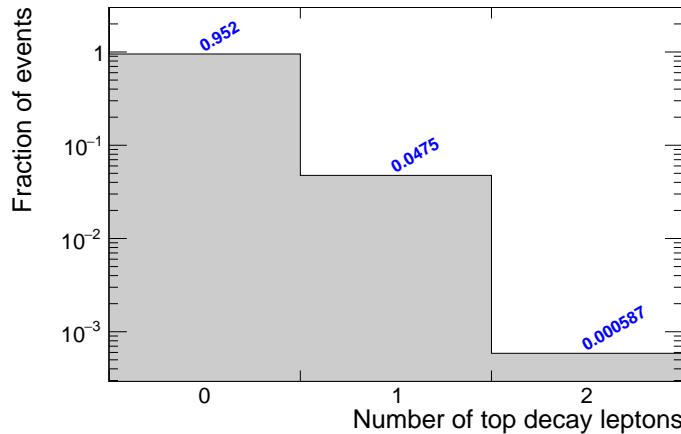


Figure 8: Distribution of $t\bar{t}$ decay modes after the signal selection of the analysis.

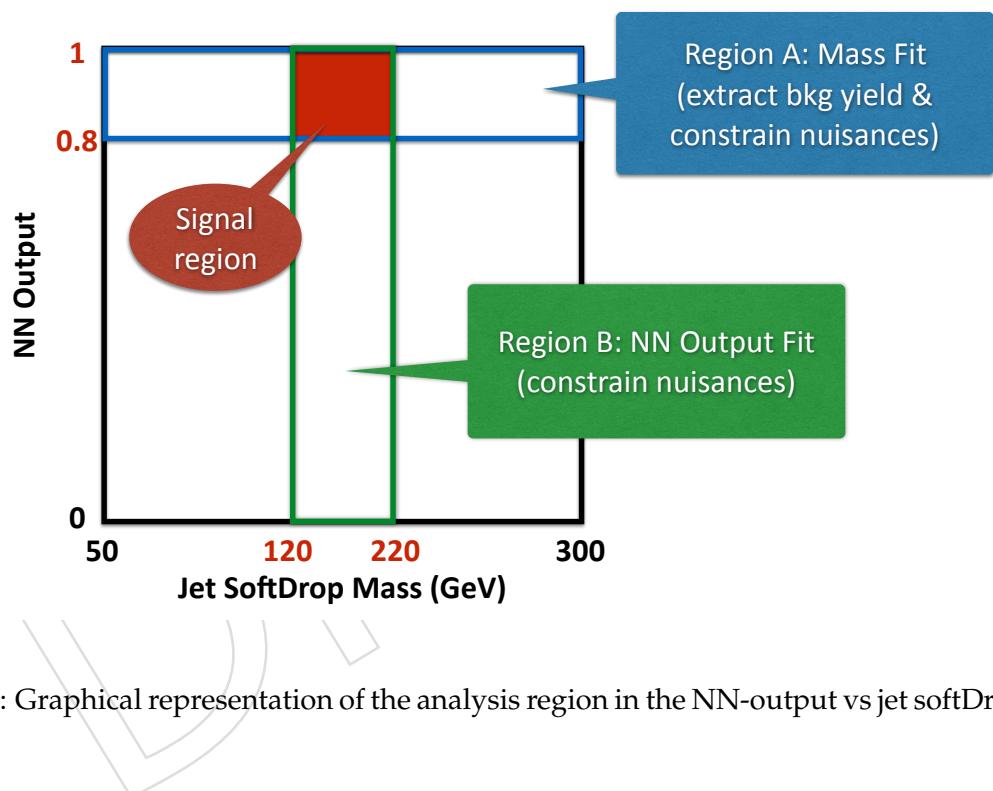


Figure 9: Graphical representation of the analysis region in the NN-output vs jet softDrop mass.

168 **4.4 Reconstruction of Particle Jets**

169 Particle jets, commonly known as “genjets” in CMS, are reconstructed with the anti- k_T clustering
 170 algorithm, with distance parameter $R = 0.8$, similar to the detector jets. The input for
 171 the particle jets are the four-momentum vectors of all the stable particles, after hadronization,
 172 originating from the primary interaction, and excluding neutrinos.

173 **5 Data vs Monte Carlo**

174 In this section we present comparisons between data and Monte Carlo for various observables.
 175 It should be noted that the $t\bar{t}$ normalization is scaled by a factor 0.6, consistent with the findings
 176 of Section 9, and the QCD background is multiplied by a factor 0.9-1.0, such that the total
 177 Monte Carlo events to be equal to the events in data. Table 6 shows the event yields for the
 178 various processes after applying the baseline selection plus the requirement that both AK8 jets
 179 contain a b-tagged subjet, with or without the NN cut (see Section 4 for details on the selection
 180 requirements). Already after the b-tagging requirement the $t\bar{t}$ signal is sufficiently enhanced
 181 that it becomes visible. Finally, the NN cut suppresses QCD by a factor ~ 17 with a signal loss
 182 of $\approx 45\%$. In the signal region the signal-over-background ratio is ≈ 1.8 .

Table 6: Expected and observed event yields in the m_{top} distribution.

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 7: 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

183 Figure 10 shows the softDrop mass of the leading jet (one entry per event), which serves as a
 184 proxy for the top mass. Figure 11 shows the mass of the W candidate. This is defined as the
 185 leading AK4 subjet in the leading AK8 jet and it is selected if the event satisfies the signal region
 186 selection and if it has $p_T > 300 \text{ GeV}$ while it is not b-tagged ($\text{CSV} < 0.8484$). The distribution
 187 of the NN output is shown in Fig. 12.

188 The kinematic properties of jets (two entries per event) in the signal region are shown in Fig. 13,
 189 while those of the $t\bar{t}$ system are shown in Fig. 14. Finally, the jet substructure properties (n-
 190 subjettiness variables) are shown in Fig. 15.

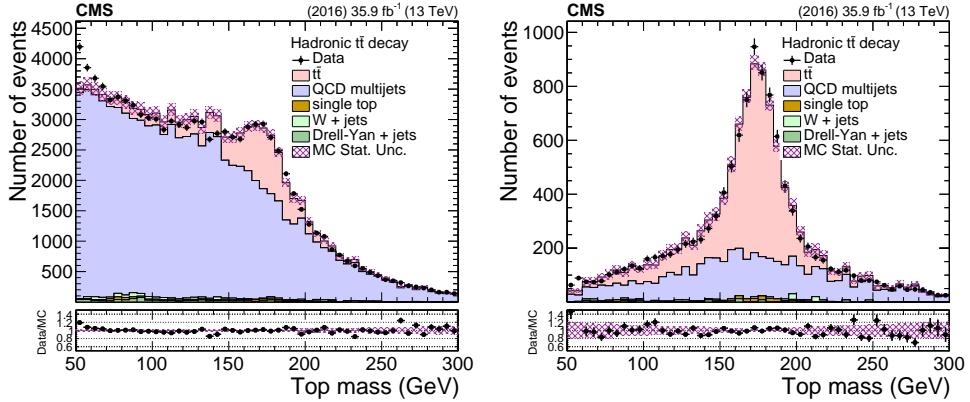


Figure 10: SoftDrop mass of the leading AK8 jet after the baseline selection with both AK8 jets containing a b-tagged subjets, with (right) and without (left) the $NN > 0.8$ cut.

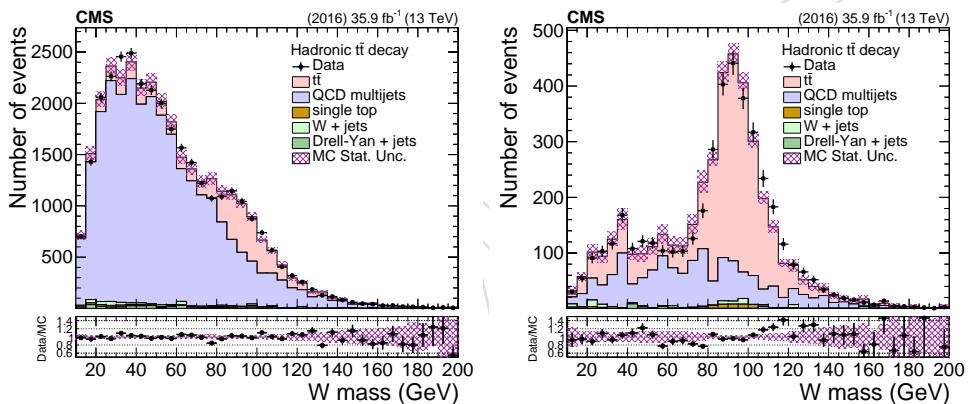


Figure 11: SoftDrop mass of the W candidate jet after the baseline selection with both AK8 jets containing a b-tagged subjets, with (right) and without (left) the $NN > 0.8$ cut.

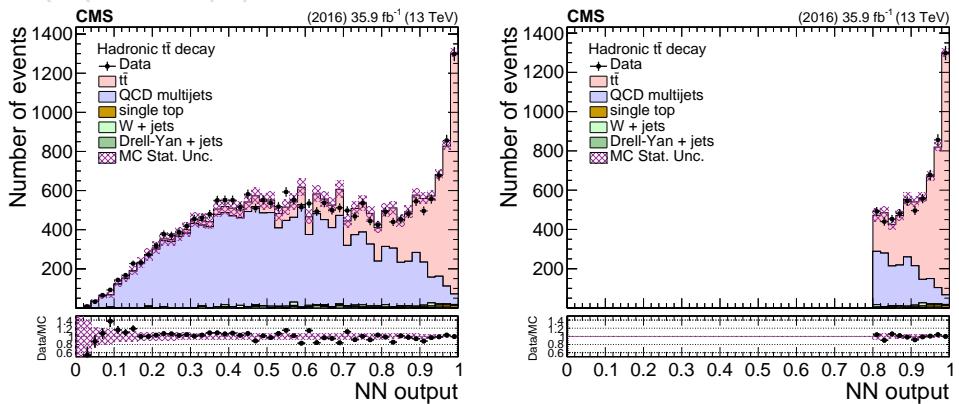


Figure 12: NN output after the baseline selection with both AK8 jets containing a b-tagged subjets, with (right) and without (left) the $NN > 0.8$ cut.

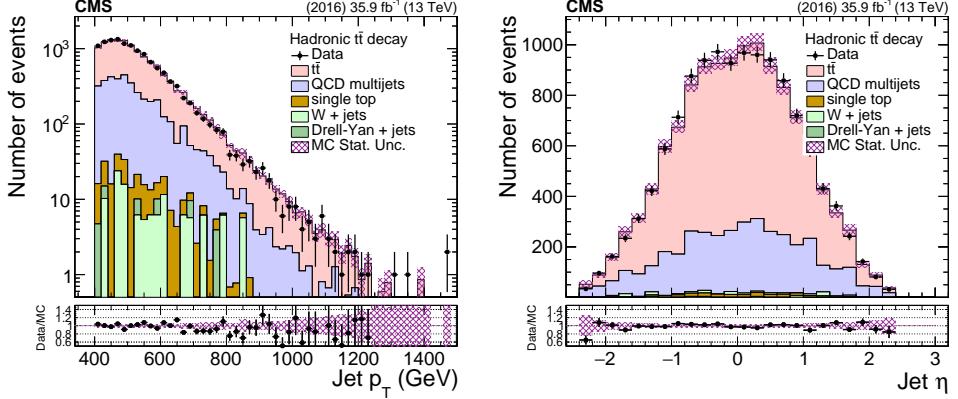


Figure 13: Jet p_T (left) and η (right) distributions in the signal region. The plots contain two entries per event (for the leading and trailing jets).

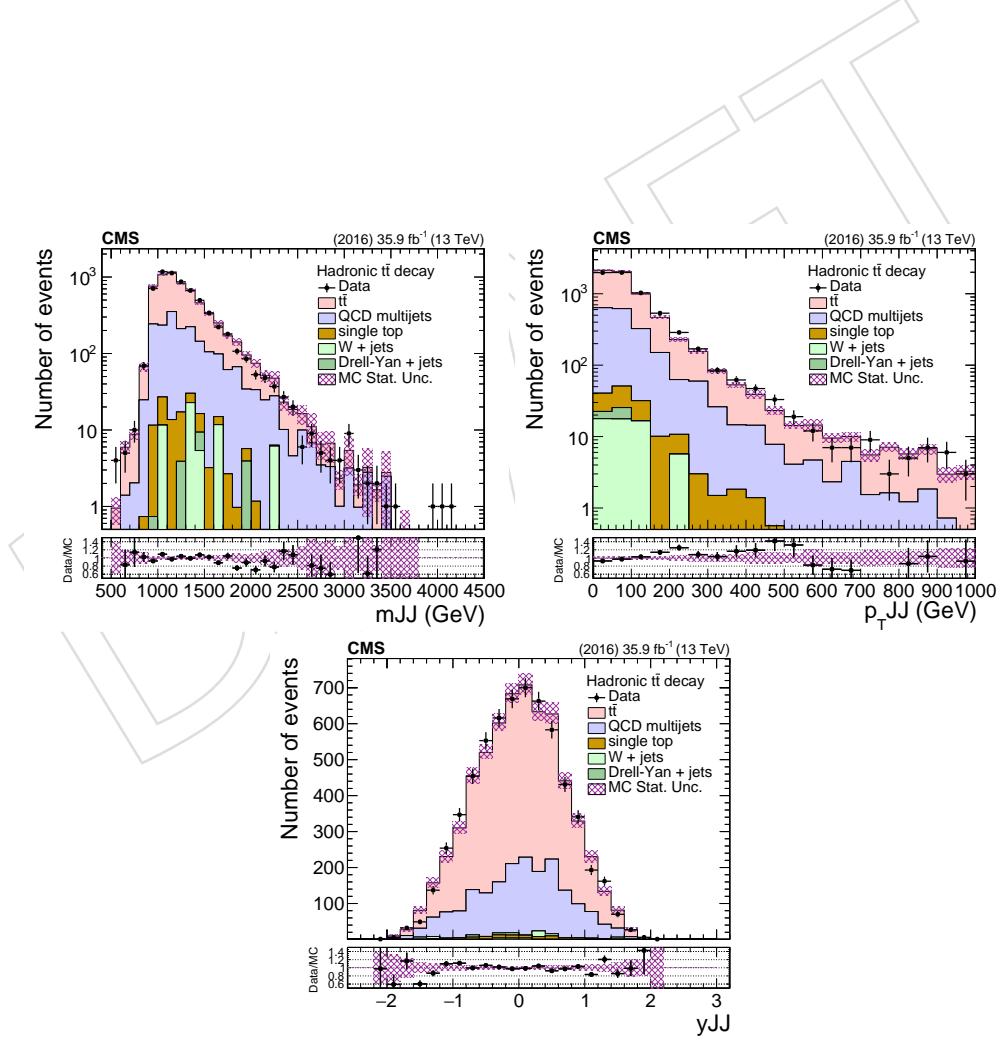


Figure 14: Kinematic distributions of the $t\bar{t}$ system in the signal region.

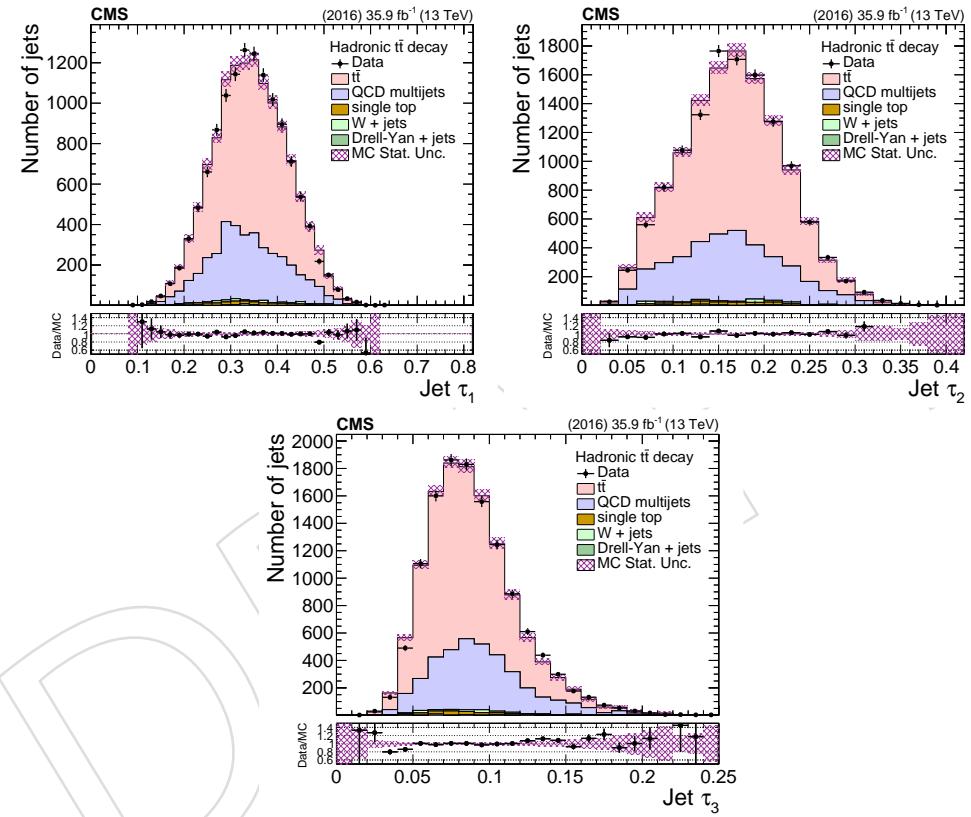


Figure 15: N-subjettiness variables for the two leading jets in the signal region.

191 6 Background

192 The, by far, dominant background in this analysis is the QCD multijet production, as there is
 193 a finite probability that ordinary jets, from single parton radiation, will mimic the topological
 194 substructure of a top-decay jet. With a combination of b-tagging requirements on the subjets
 195 and n-subjettiness variables, we are able to suppress significantly the QCD background, as
 196 shown in Section 5. In order to estimate the remaining contribution we employ a data-driven
 197 technique based on the assumptions that if the b-tagging requirement is reverted we a) get a
 198 pure QCD sample, and b) the jet kinematic properties are not affected.

199 The two elements that need to be determined from data are the shape $Q(x)$ of the QCD back-
 200 ground as a function of an observable of interest x , and the absolute normalization N_Q . The
 201 QCD template $Q(x)$ is taken from the QCD control sample by applying the signal region selec-
 202 tion with reverted b-tagging requirement (none of the leading AK8 jets is allowed to contain a
 203 b-tagged subjet). Figures 16,17 show the closure test in MC as a function of jet and dijet kin-
 204 ematic variables. We observe that, within the statistical precision of the simulation, the shapes
 205 are compatible, regardless of the b-tagging requirement. Moreover, Figs. 19,20 show that in-
 206 deed the 0-btag sample has negligible $t\bar{t}$ contamination.

207 The normalization of the QCD background is determined from a fit to the data in signal region
 208 A (Section 4) on the softDrop mass of the top candidate. However, since this is an extended
 209 signal region (in order to allow for QCD dominated sidebands), a transfer factor R_{yield} is needed
 210 in order to get the QCD normalization in the signal region. This transfer factor is different for
 211 the $t\bar{t}$ kinematic variables (one entry per event) and the top variables (two entries per event).
 212 Figure 18 shows a closure test performed in QCD Monte Carlo events, where the value of R_{yield}
 213 is independent of the b-tagging requirement (within the statistical precision of the simulation),
 214 and therefore can be taken from the QCD control sample.

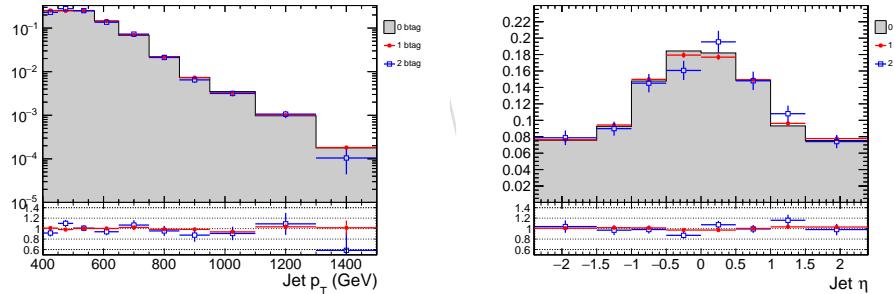


Figure 16: Closure test for the data-driven QCD prediction method for the shape of the top p_T and η variables.

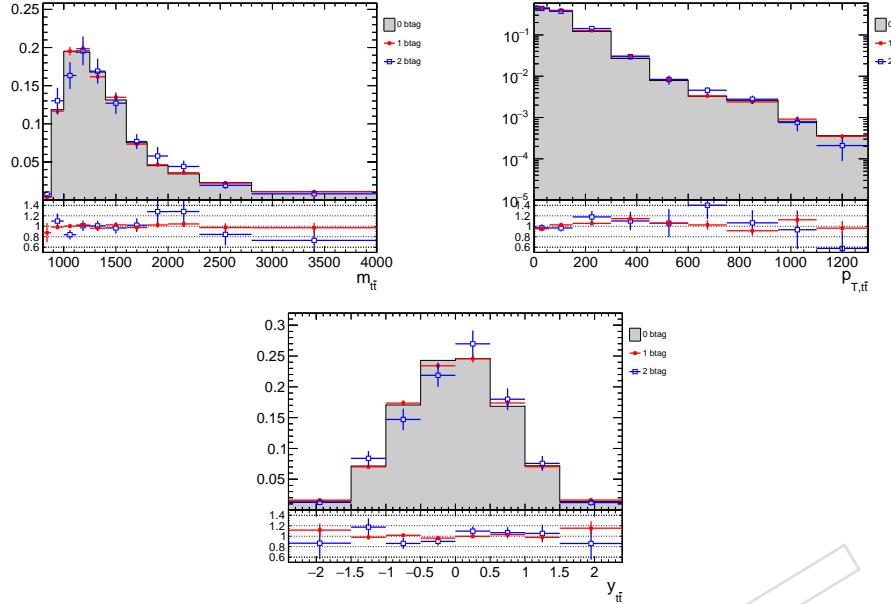


Figure 17: Closure test for the data-driven QCD prediction method for the shape of the $t\bar{t}$ variables.

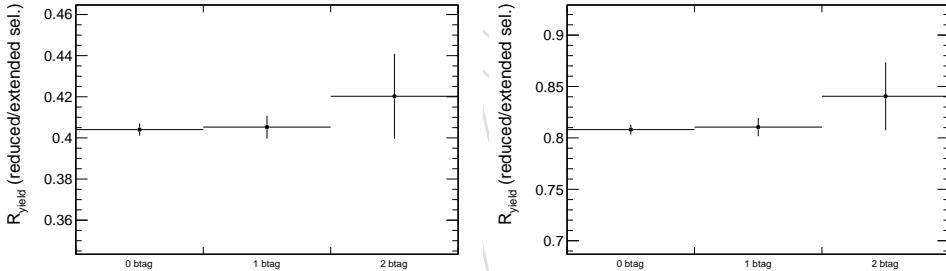


Figure 18: Closure test for the data-driven QCD prediction method for event yield ratio between the control and signal regions.

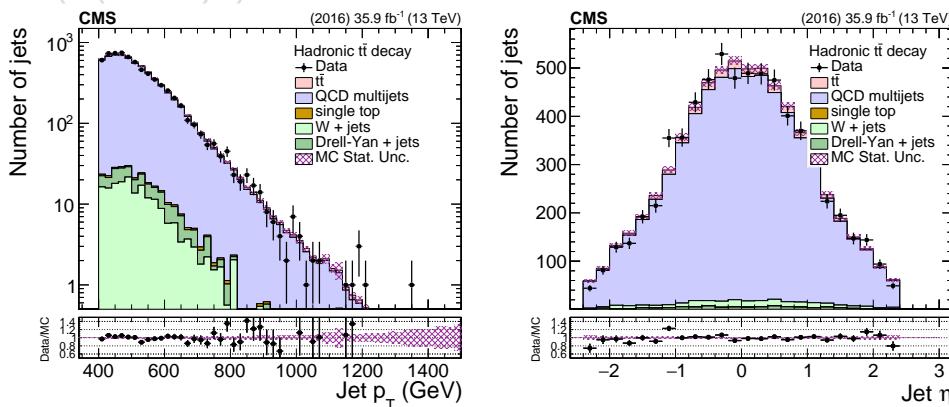


Figure 19: Jet p_T (left) and η (right) distributions in the QCD control region. The plots contain two entries per event (for the leading and trailing jets).

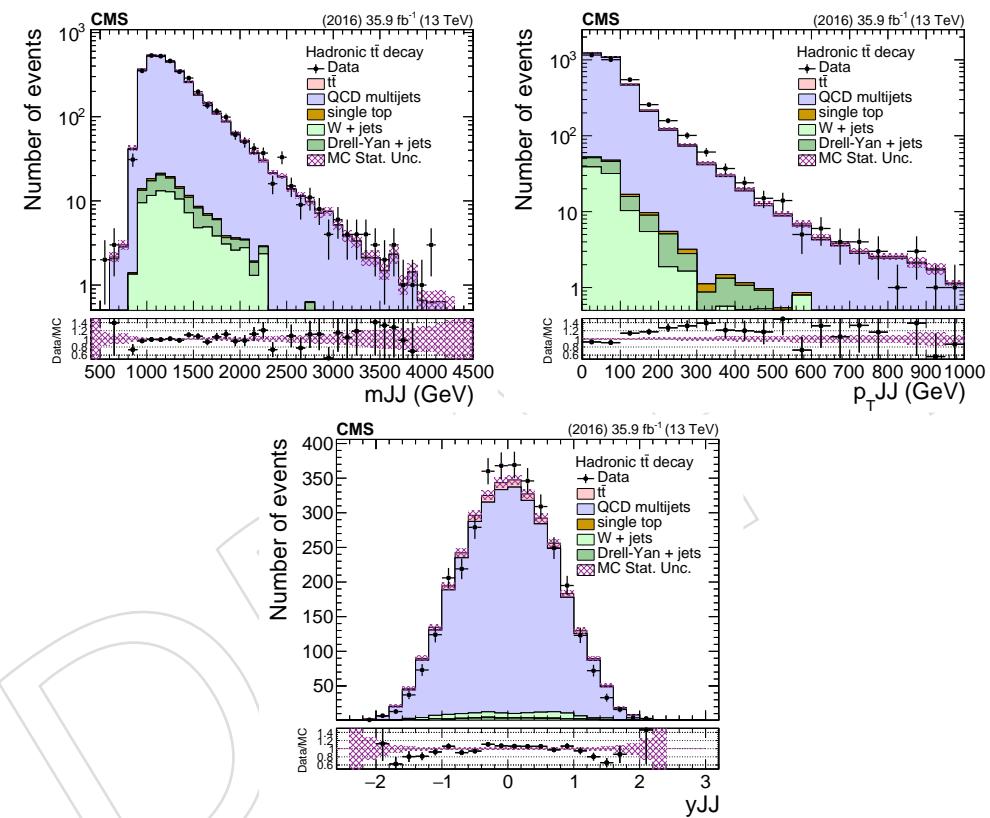


Figure 20: Kinematic distributions of the $t\bar{t}$ system in the QCD control region.

215 7 Signal Extraction

216 The $t\bar{t}$ signal is extracted from data, differentially, as a function of five variables (top p_T and η ,
 217 $t\bar{t}$ mass, p_T , and rapidity) by subtracting the contribution of the backgrounds:

$$S(x) = D(x) - R_{\text{yield}} N_{\text{qcd}} Q(x) - B(x), \quad (2)$$

218 where $x = p_T^t, \eta^t, m^{t\bar{t}}, p_T^{t\bar{t}}, y^{t\bar{t}}$, $S(x)$ is the signal, $D(x)$ is the measured distribution in data, $Q(x)$
 219 is the QCD shape, $B(x)$ is the subdominant backgrounds' contribution (both the shape and the
 220 normalization are taken from the simulation), $R_{\text{yield}} = N_{0b}^{\text{SR}} / N_{0b}^{SR_A}$ is the yield ratio between
 221 the signal region and the fit region, and N_{qcd} is the fitted number of QCD events in SR_A .

222 The quantity N_{qcd} is estimated from a fit to the data in SR_A on the mass of the top candidate,
 223 which is the softDrop mass of the leading jet. The fit is described by the equation below:

$$D(m^t) = N_{t\bar{t}} T(m^t; k_{\text{scale}}, k_{\text{res}}) + N_{\text{qcd}} (1 + k_{\text{slope}} m^t) Q(m^t) + N_{\text{bkg}} B(m^t) \quad (3)$$

224 which contains the shapes (templates) $T(m^t)$, $B(m^t)$ of the signal and the subdominant back-
 225 grounds, respectively, taken from the simulation, and the shape $Q(m^t)$ of QCD taken from the
 226 control sample in data. The templates of the various components are shown in Figs. 21,22. To
 227 account for the difference observed in the closure test of QCD for the m^t variable (Fig. 23) we
 228 introduced the linear modification factor $(1 + k_{\text{slope}} m^t)$, inspired by the simulation, but with
 229 the slope parameter k_{slope} left free in the fit. Also free in the fit are the normalization factors
 230 $N_{t\bar{t}}$, N_{qcd} , and N_{bkg} . Finally, we introduce two more nuisance parameters in the $t\bar{t}$ simulation,
 231 k_{scale} and k_{res} , which account for possible differences between data and simulation in the scale
 232 and resolution of the m^t parameter. The fit model is imported to the RooFit package and the fit
 233 result is shown in Fig. 24 while the fitted parameters are summarized in Table 8. We observe
 234 that the fitted $t\bar{t}$ yield (6238) is significantly lower than the expectation (9604), which implies
 235 that the fiducial cross section is $\approx 35\%$ lower compared to the Powheg+Pythia8 prediction.
 236 The nuisance parameters related to the m^t scale and resolution are consistent with one and the
 237 slope of the QCD modification factor $(5.7 \pm 1.4) \times 10^{-3}$ is very close to the value from the QCD
 238 simulation $((5.4 \pm 1.2) \times 10^{-3})$. Finally, despite the large uncertainty, the yield of the subdomi-
 239 nant backgrounds (400 ± 247) is very close to the Monte Carlo prediction (380) in Table 6. The
 240 overall sanity of the fit and the consistency of the results allows us to use the measured N_{qcd}
 241 for the extraction of the signal distribution.

Table 8: Results of the fit in SR_A .

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

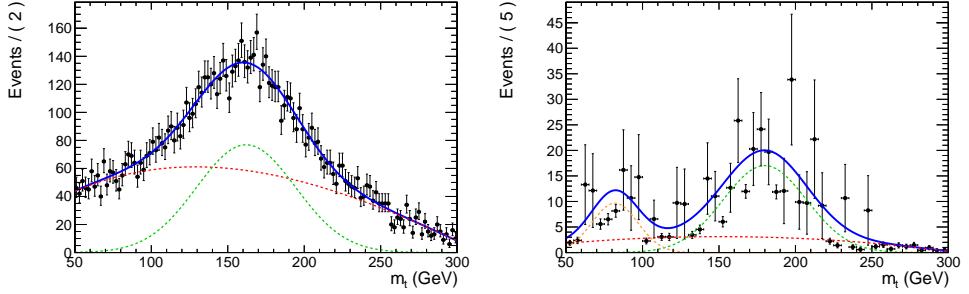


Figure 21: Templates of QCD (left), taken from data, and of the subdominant backgrounds (right), taken from the simulation. The different lines show the individual components (frozen in the fit) used to describe the shapes

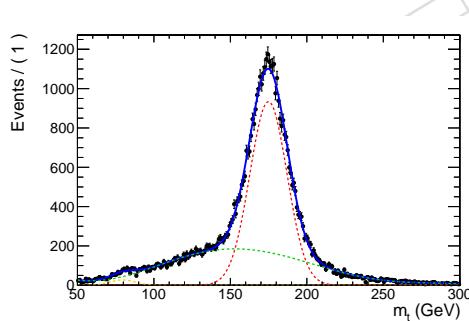


Figure 22: Template of the $t\bar{t}$ signal taken from the simulation. The different lines show the individual components (frozen in the fit) used to describe the shape.

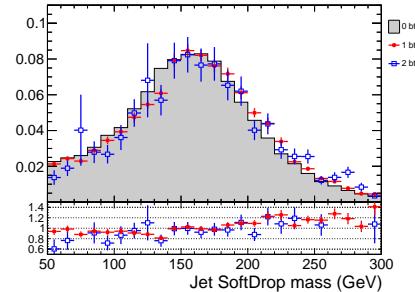


Figure 23: Closure test in the QCD simulation for the shape of the m^t variable in the three possible b-tagging requirements (none, exactly one, or both jets contain a b-tagged subjet).

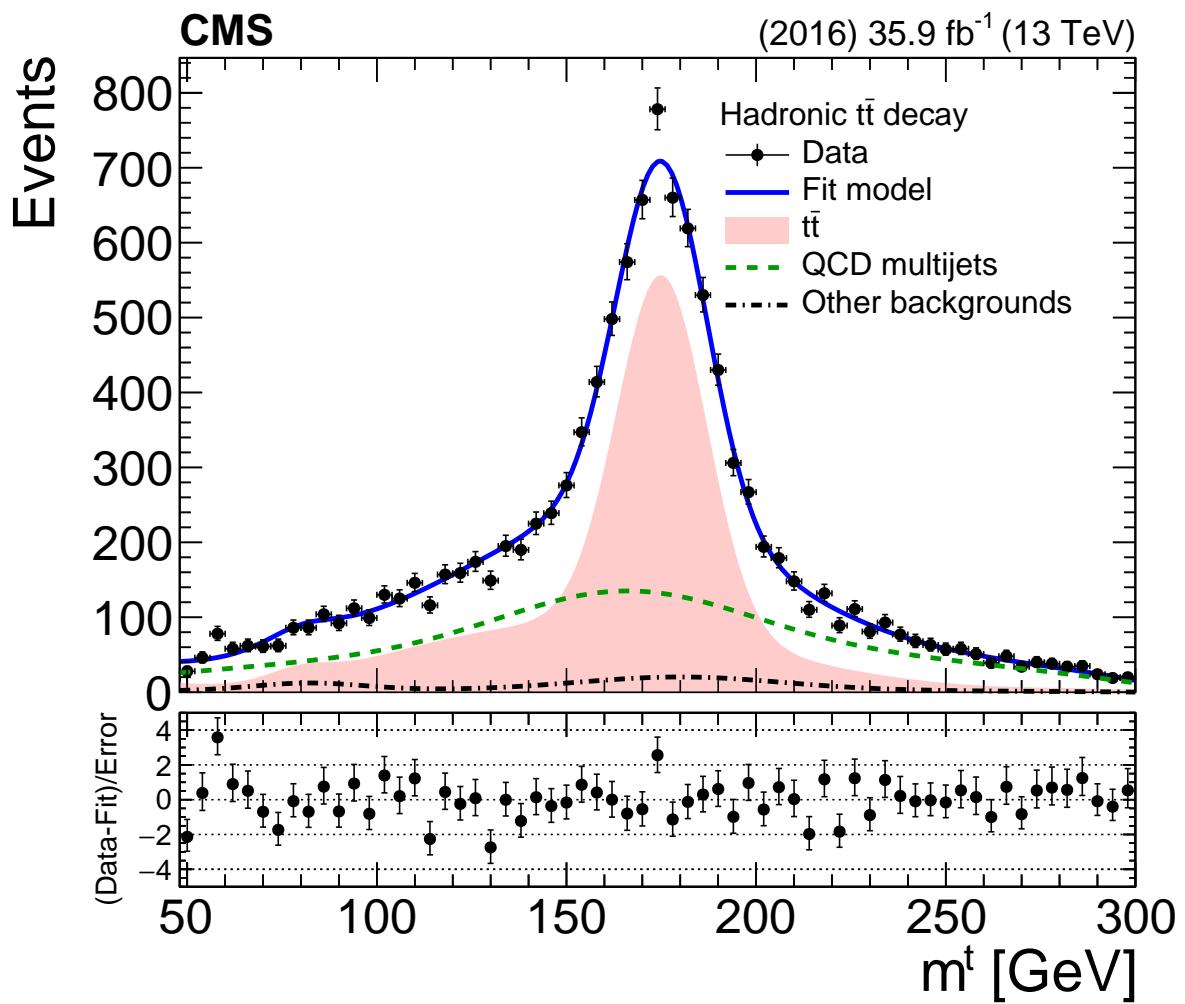


Figure 24: Result of the template fit on data. The red line shows the $t\bar{t}$ contribution, the green line shows the QCD, and the brown line shows the subdominant backgrounds.

242 8 Systematic Uncertainties

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

251 1. Experimental Uncertainties

- 252 • QCD background prediction: we use the fitted QCD yield uncertainty. The
 253 shape uncertainties due to the closure test in the simulation and the different
 254 pileup profiles in the control and signal regions are very small, at the 1% level.
- 255 • Jet energy scale (JES): this is the uncertainty on the energy scale of each recon-
 256 structed jet and it is a leading experimental uncertainty. Following the recom-
 257 mendations of the JME group, we have considered 24 independent JES sources
 258 as follows: for each variation a new jet collection is created and the event inter-
 259 pretation is repeated. This results not only in variations of the p_T scale itself,
 260 but may also lead to different top candidates. The JES uncertainty, per jet, is
 261 of the order 1 – 2%, p_T and η dependent. The effect on the measured cross
 262 section is typically of the order 10% but it can be much larger at very high jet
 263 p_T .
- 264 • Jet energy resolution (JER): Similar to what is done with the JES, jets are smeared
 265 according to the JER uncertainty and the event is re-interpreted. The effect on
 266 the cross section is relatively small, at the level of 2%.
- 267 • B-jet tagging (Btag): this is also a leading experimental uncertainty and it
 268 reflects the uncertainty on the identification of b-subjets within the boosted
 269 AK8 jets. We have followed the official BTV recommendation, where an event
 270 weight is assigned to the simulation based on the per-jet b-tag efficiency and
 271 fake-rate scale factors. The effect on the cross sections is of the order 10% rela-
 272 tively flat in all the observables. So, unlike the JES, the Btag uncertainty largely
 273 cancels in the normalized cross sections.
- 274 • Pileup: this is a subdominant uncertainty related to the profile of the true num-
 275 ber of interactions taking place in data. The simulated events (generated with
 276 a different profile of interactions compared to the one observed in data) are
 277 weighted to match the data with a varied inelastic cross section. The effect on
 278 the cross sections is negligible (below 1%).
- 279 • Trigger: this accounts for the difference between the simulated and observed
 280 trigger efficiency. Based on Fig. 2 the uncertainty is well below 1% in the phase
 281 space of this analysis.
- 282 • Luminosity: the official CMS uncertainty of 2.4% is assigned to the integrated
 283 luminosity.

284 2. Theoretical Uncertainties

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

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

- 291 • Parton distribution functions (PDF): estimated by applying event weights cor-
 292 responding to the 100 replicas of the NNPDF set. For each observable we com-
 293 pute its standard deviation from the 100 variations.
- 294 • Renormalization/ Factorization scales (Scale): estimated by applying event
 295 weights corresponding to different factorization and renormalization scale op-
 296 tions. For each observable we take the largest deviation from the nominal
 297 scales.
- 298 • Strong coupling constant (α_S): estimated by applying event weights corre-
 299 sponding to higher and lower strong-coupling constant for the matrix-element.
- 300 • Final state radiation (FSR): this uncertainty is estimated from alternative MC
 301 samples with reduced and increased value for the strong coupling constant
 302 used by Pythia8 to generate final state radiation. The original variations are
 303 largely exaggerated, as seen in Fig. 25, where the alternative samples fail mis-
 304 erably to describe the data in two sensitive observables: the mass of the top
 305 candidate jet and the NN output, both of which rely on the dynamics of the jet
 306 formation. In order to use a more realistic estimate of this uncertainty, a max-
 307 imum likelihood fit is performed (independently) on these variables, using
 308 the systematic uncertainties as independent nuisance parameters. The fits are
 309 implemented through the “Combine” toolbox using templates for the various
 310 components and are shown in Fig. 26. The fit on the m^t variable is well be-
 311 haved in the entire fit range, while the fit on the NN output is less successful
 312 in the background-enriched region. Nevertheless, since it is the signal uncer-
 313 tainties that we wish to constrain, the signal region is fitted perfectly. As a re-
 314 sult of these fits, the nuisances are constrained in-situ. Figures 27 and 28 show
 315 the comparison of the pre-fit and post-fit uncertainties, as well as the pulls of
 316 the nuisance parameters. Both fits indicate that the post-fit FSR nuisance is
 317 actually close to its pre-fit (nominal) value with a significant reduction of its
 318 uncertainty. Based on the observations from the fits, we have scaled down the
 319 effects of the FSR variations by a (conservative) factor 0.3.
- 320 • Initial state radiation (ISR): this uncertainty is estimated from alternative MC
 321 samples with reduced and increased value for the strong coupling constant
 322 used by Pythia8 to generate initial state radiation.
- 323 • Matrix element - parton shower matching: (hdamp) this uncertainty is esti-
 324 mated from alternative MC samples.
- 325 • Underlying event tune (Tune): this uncertainty is estimated from alternative
 326 MC samples with modified Pythia8 tunes.

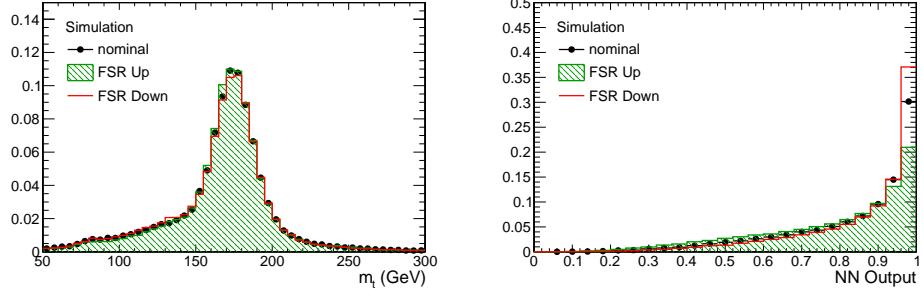


Figure 25: Variations of the m^t (left) and NN output (right) distributions due to the alternative FSR MC samples.

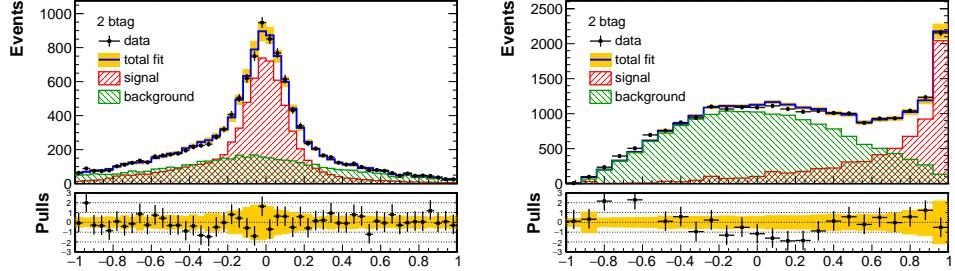


Figure 26: Fit to the m^t (left) and NN output (right) distributions aiming to constrain the MC modelling uncertainties.

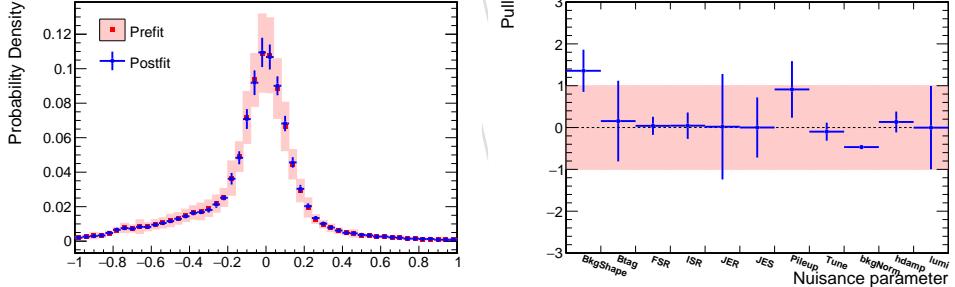


Figure 27: (left) Comparison between the prefit and postfit distributions and uncertainties of the m^t distribution. (right) Fit pulls of the nuisance parameters.

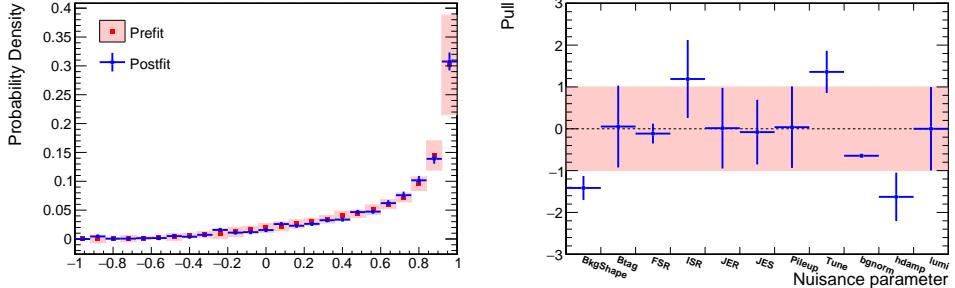


Figure 28: (left) Comparison between the prefit and postfit distributions and uncertainties of the NN output distribution. (right) Fit pulls of the nuisance parameters.

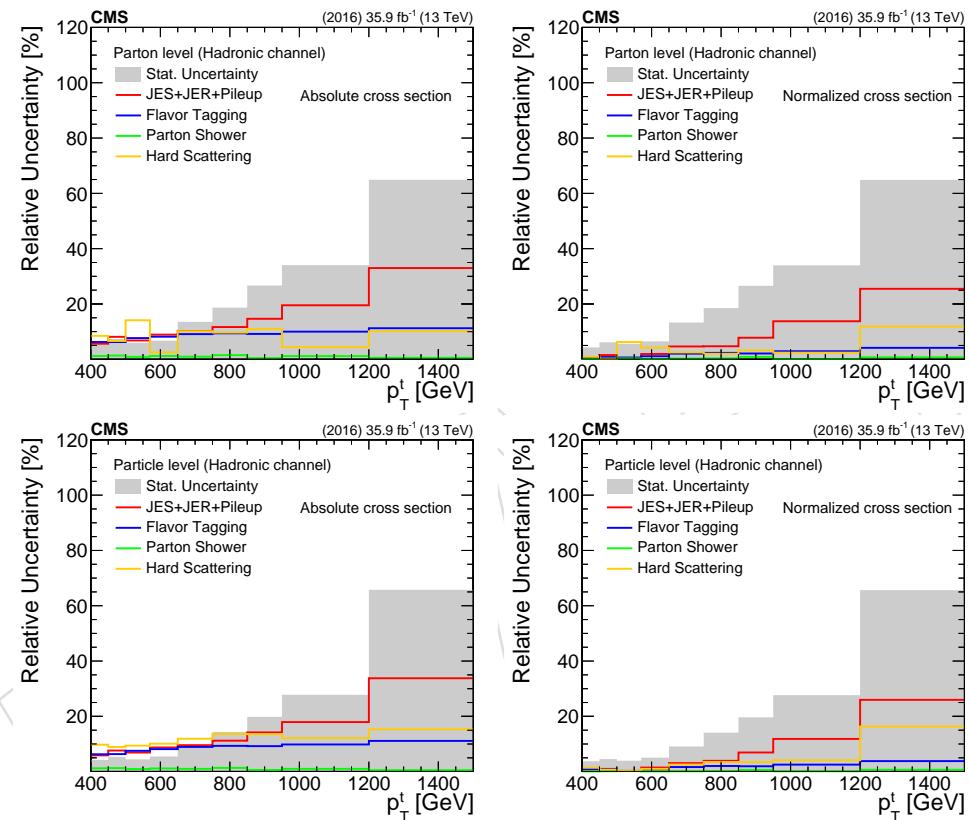


Figure 29: Decomposition of uncertainties for the parton- and particle-level measurement (left: absolute, right: normalized) as a function of top p_T .

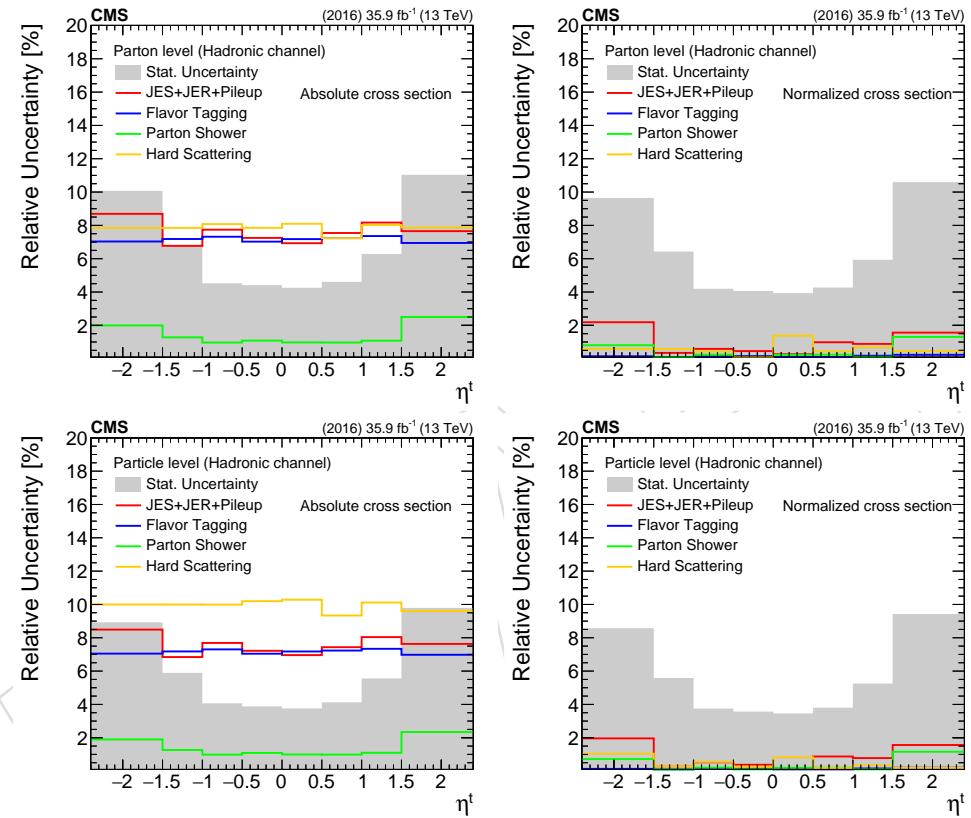


Figure 30: Decomposition of uncertainties for the parton- and particle-level measurement (left: absolute, right: normalized) as a function of top η^t .

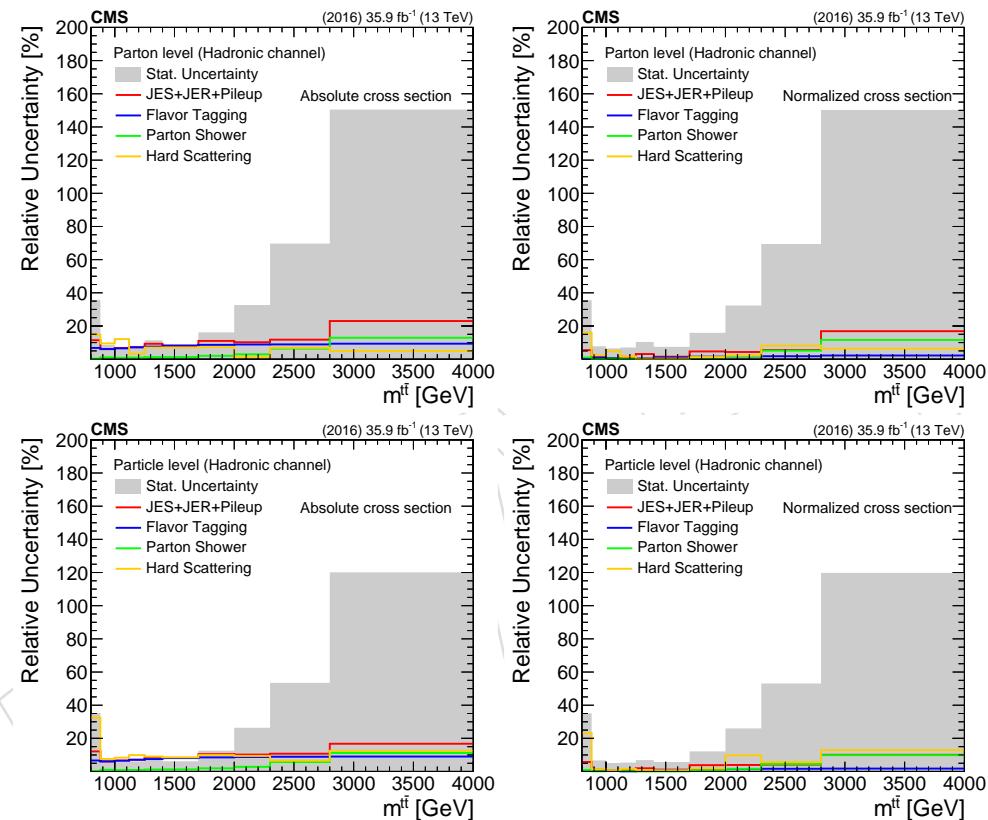


Figure 31: Decomposition of uncertainties for the parton- and particle-level measurement (left: absolute, right: normalized) as a function of $m^{t\bar{t}}$.

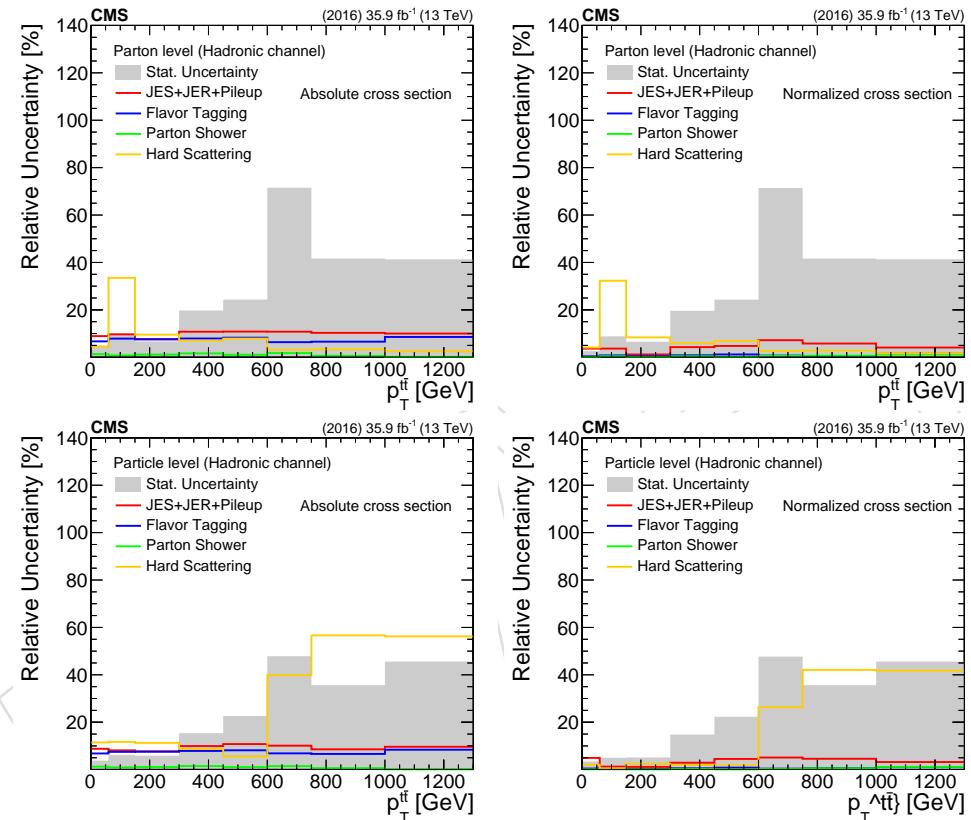


Figure 32: Decomposition of uncertainties for the parton- and particle-level measurement (left: absolute, right: normalized) as a function of p_T^t .

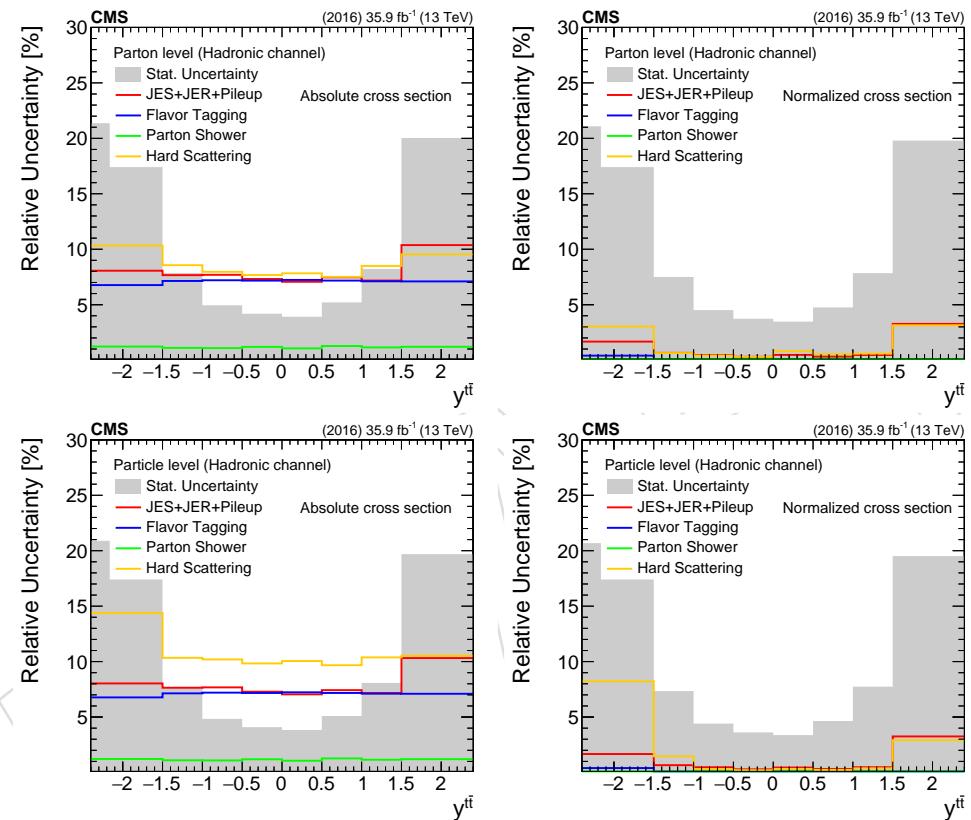


Figure 33: Decomposition of uncertainties for the parton- and particle-level measurement (left: absolute, right: normalized) as a function of $y_{t\bar{t}}$.

327 9 Fiducial Measurement

328 The fiducial differential cross section is derived in bin i of the variable x from the signal yield
 329 S_i (eq. 2) in the bin as follows:

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

330 where \mathcal{L} is the total integrated luminosity. Of particular interest is also the normalized cross
 331 section, computed as:

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

332 which is used to confront the modelling of the differential cross section regardless of the overall
 333 normalization. In order to estimate the uncertainty on the measurement, the entire procedure is
 334 repeated for every source of uncertainty described in Section 8. Both the experimental and the
 335 theoretical uncertainties affect primarily the $t\bar{t}$ singal shape, which is used to fit the data for the
 336 QCD background normalization. As a result the total systematic uncertainty of the measure-
 337 ment is small and the statistical uncertainty dominates. This is true both for the absolute and
 338 the normalized cross sections. Figures 34-38 show the fiducial cross sections as a function of
 339 the variables of interest. Firstly, we observe the expected offset of about 40% in the total cross
 340 section between the data and the Powheg+Pythia8 prediction, which is considerably smaller
 341 for the other theory predictions. In particular the Powheg+Herwigpp seems to agree best (this
 342 is most visible in Fig. 35). In terms of shape comparisons, the Powheg+Pythia8 prediction is in
 343 excellent agreement with data for all the variables, except for the invariant mass $m^{t\bar{t}}$ of the $t\bar{t}$
 344 system, where the theoretical spectrum appears to be harder above ≈ 2 TeV. A similar behavior
 345 is observed for the other theory models with hints of somewhat larger differences, although the
 346 statistical precision of the predictions does not allow quantitative comparisons.

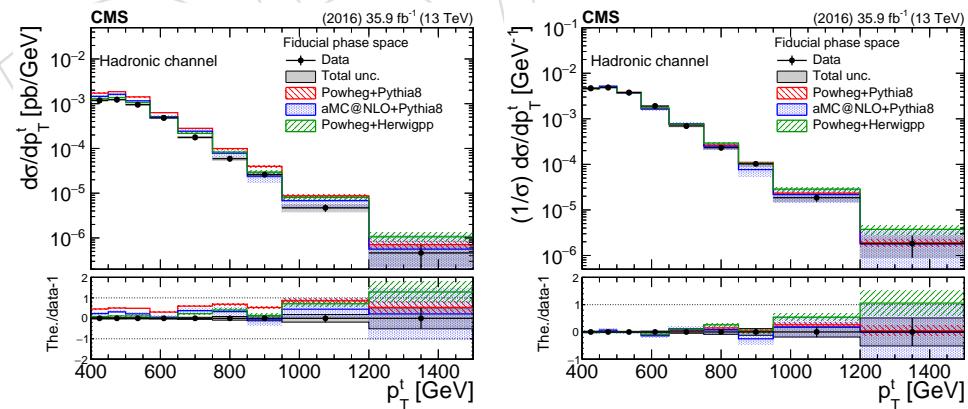


Figure 34: Fiducial differential cross section, absolute (left) and normalized (right), as a function of top p_T . The bottom panel shows the ratio (theory - data)/data.

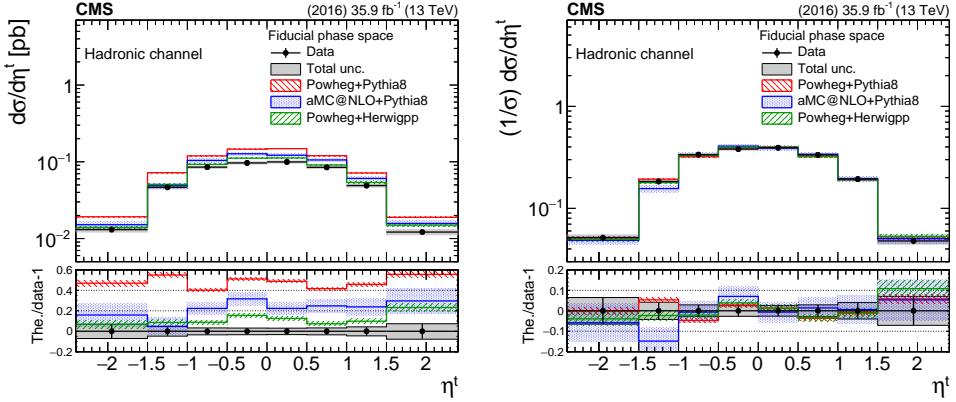


Figure 35: Fiducial differential cross section, absolute (left) and normalized (right), as a function of top η . The bottom panel shows the ratio (theory - data)/data.

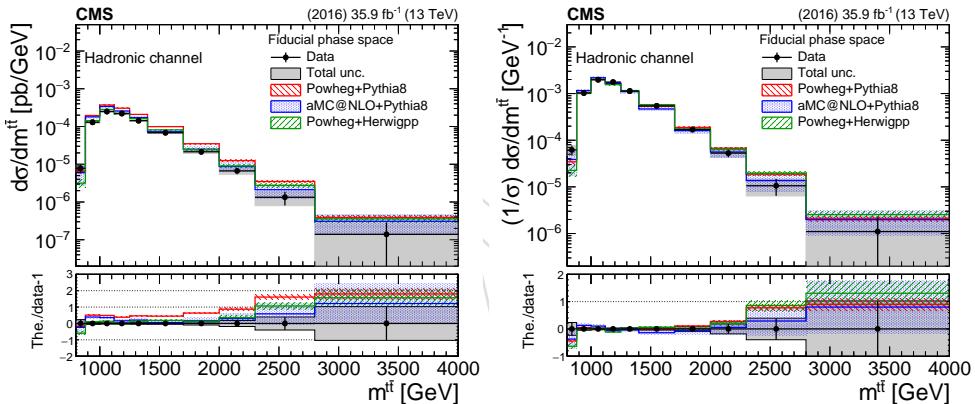


Figure 36: Fiducial differential cross section, absolute (left) and normalized (right), as a function of $m_{\bar{t}t}$. The bottom panel shows the ratio (theory - data)/data.

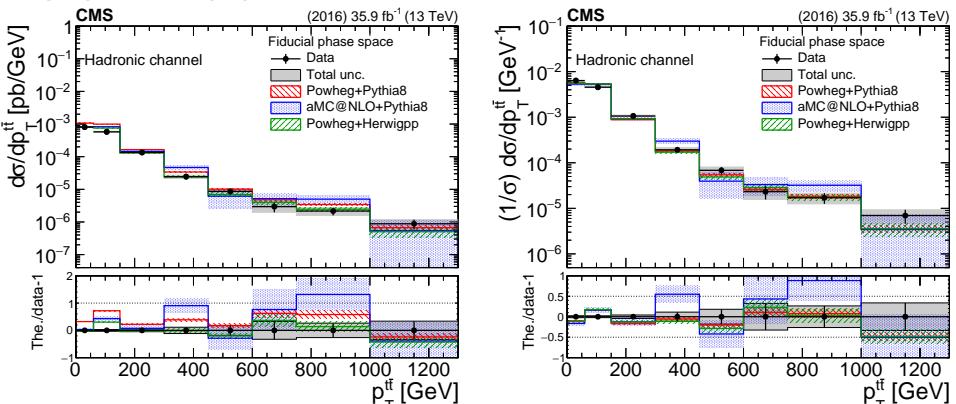


Figure 37: Fiducial differential cross section, absolute (left) and normalized (right), as a function of $p_{\bar{t}t}$. The bottom panel shows the ratio (theory - data)/data.

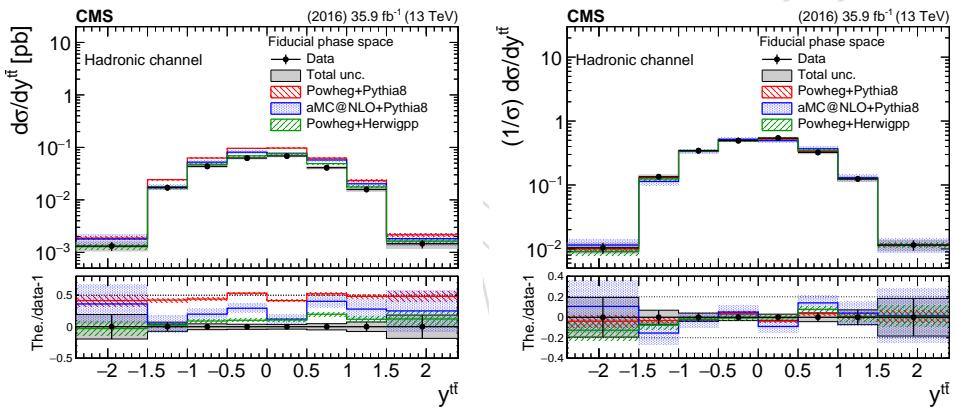


Figure 38: Fiducial differential cross section, absolute (left) and normalized (right), as a function of top $y^{t\bar{t}}$. The bottom panel shows the ratio $(\text{theory} - \text{data})/\text{data}$.

347 10 Unfolded Measurement

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

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

350 where \mathcal{L} is the total integrated luminosity and Δx_i is the width of the i-th bin of the observable
 351 x . The quantity $f_{1,j}$ is the fraction of reconstructed events in the j-th bin that have an equiva-
 352 lence event at the unfolded level (parton or particle), whereas the quantity $f_{1,i}$ is the fraction of
 353 events at the unfolded level that have an equivalent reconstructed event. Figures 39 and 48 be-
 354 low show the aforementioned fractions at parton and particle level, respectively, as a function
 355 of all the observables. The quantity R_{ij}^{-1} is the inverse of the migration matrix between the i-th
 356 and j-th bins. Due to the finite resolution of the detector, the migration matrix is non-diagonal
 357 and thus the application of an unfolding procedure is necessary. The binning of the various
 358 observables has been chosen such that the purity (fraction of reconstructed events that the true
 359 value of the observable lies in the same bin) and the stability (fraction of true events that the re-
 360 constructed observable lies in the same bin) are well above 50% (Figures 40 and 49). This choice
 361 results in highly diagonal migration matrices, shown in Figs. 41 and 50. In order to avoid the
 362 biases introduced by the various unfolding methods with some type of regularization, we have
 363 used simple migration matrix inversion, as written in Eq. 6 at a price of a moderate increase of
 364 the statistical uncertainty (see details in Appendix C).

365 10.1 Parton Level

366 The partonic phase space to which we unfold the measurement is constrained by the kinematic
 367 requirements of the detector-level fiducial region. Namely, the top and anti-top partons must
 368 have $p_T > 400 \text{ GeV}$ and $|\eta| < 2.4$, while the invariant mass of the $t\bar{t}$ system must be greater
 369 than 800 GeV in order to avoid extreme events with high top p_T and very low $m_{t\bar{t}}$. These cuts
 370 are summarized in Table 9.

Table 9: Definition of parton-level phase space.

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

371 The results of the unfolded measurement at parton level are shown in Figs. 42- 46. The compar-
 372 ison with the theory predictions follows the trends observed in the fiducial measurement. That
 373 is, there is an 20 – 40% lower inclusive cross section, while the shapes of the differential distri-
 374 butions are reasonably reproduced by all models. A hint of a deviation at very high values of
 375 $m^{t\bar{t}}$ is observed but the measurement is not very precise at this part of the phase space.

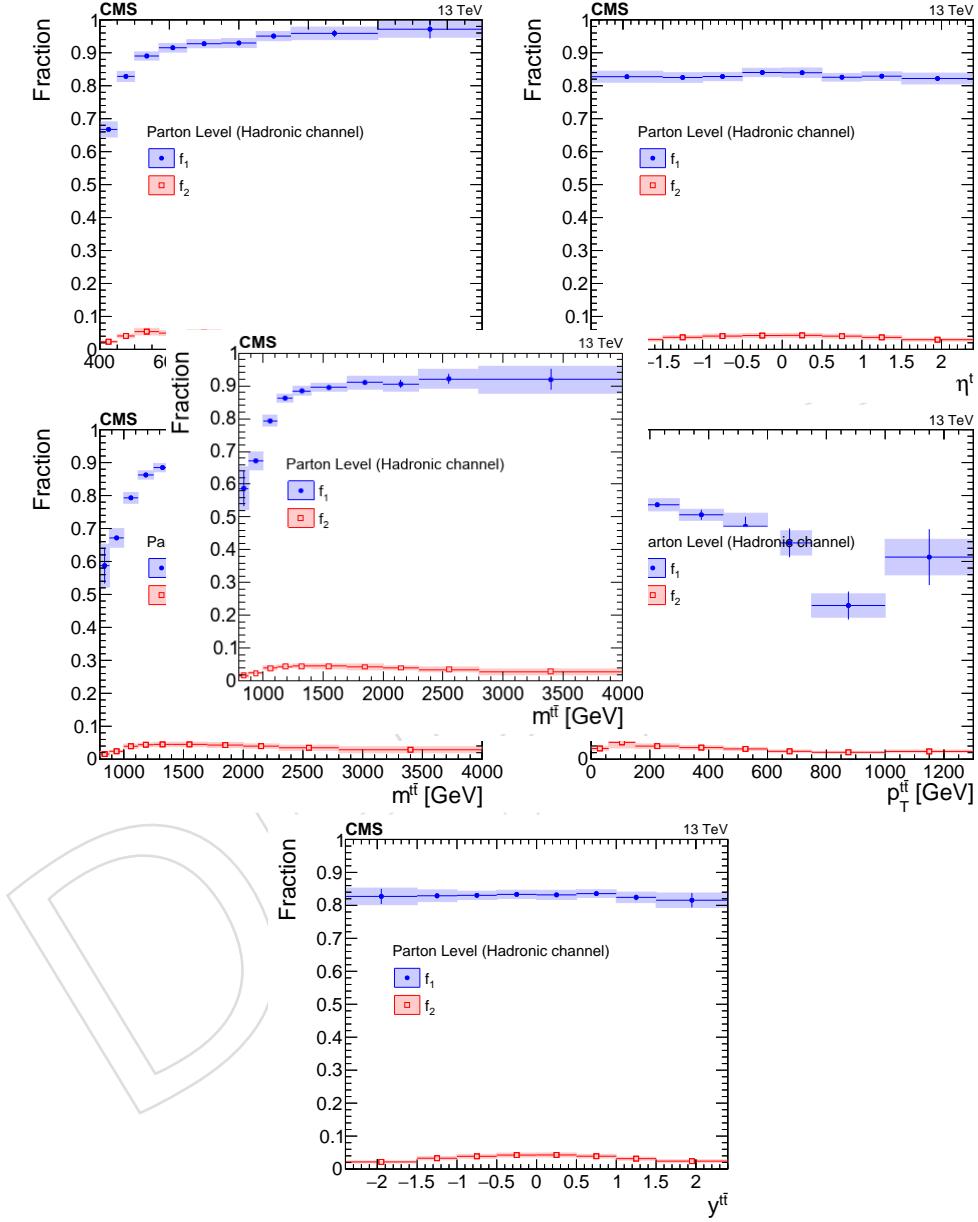


Figure 39: Simulated fractions $f_{1,2}$ for the parton-level selection as a function of the various observables.

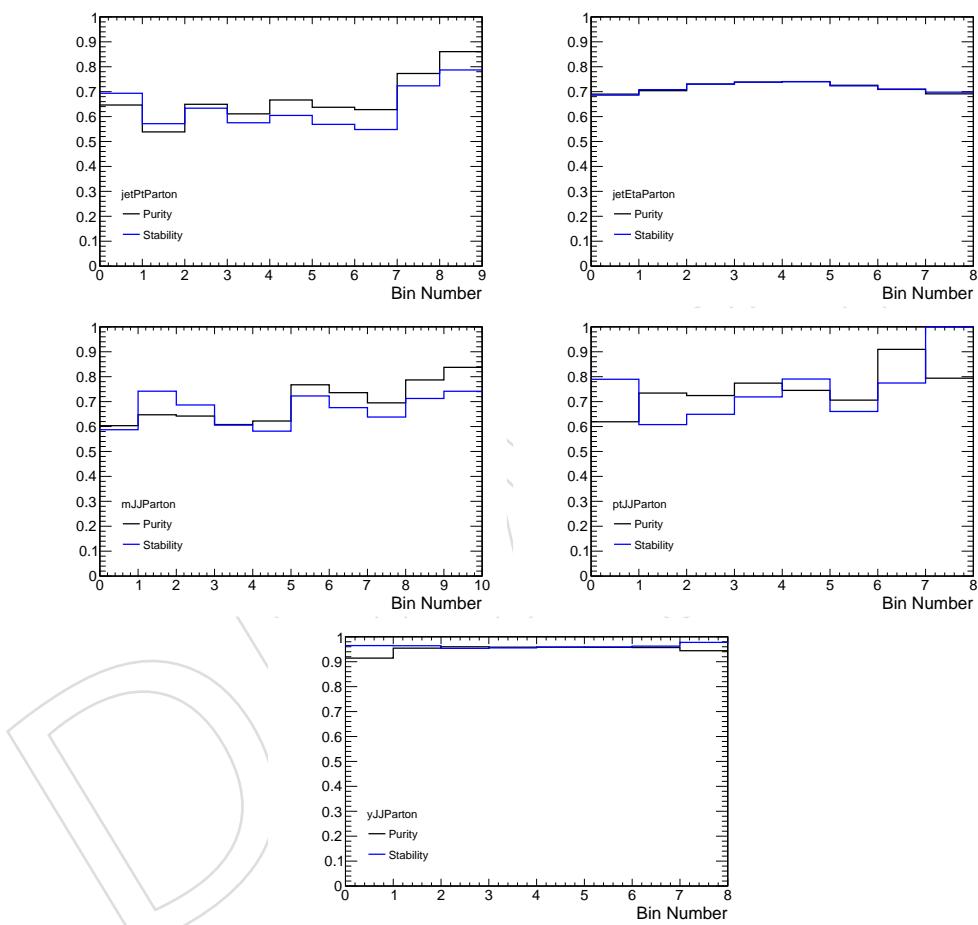


Figure 40: Simulated purity and stability of each bin at parton level.

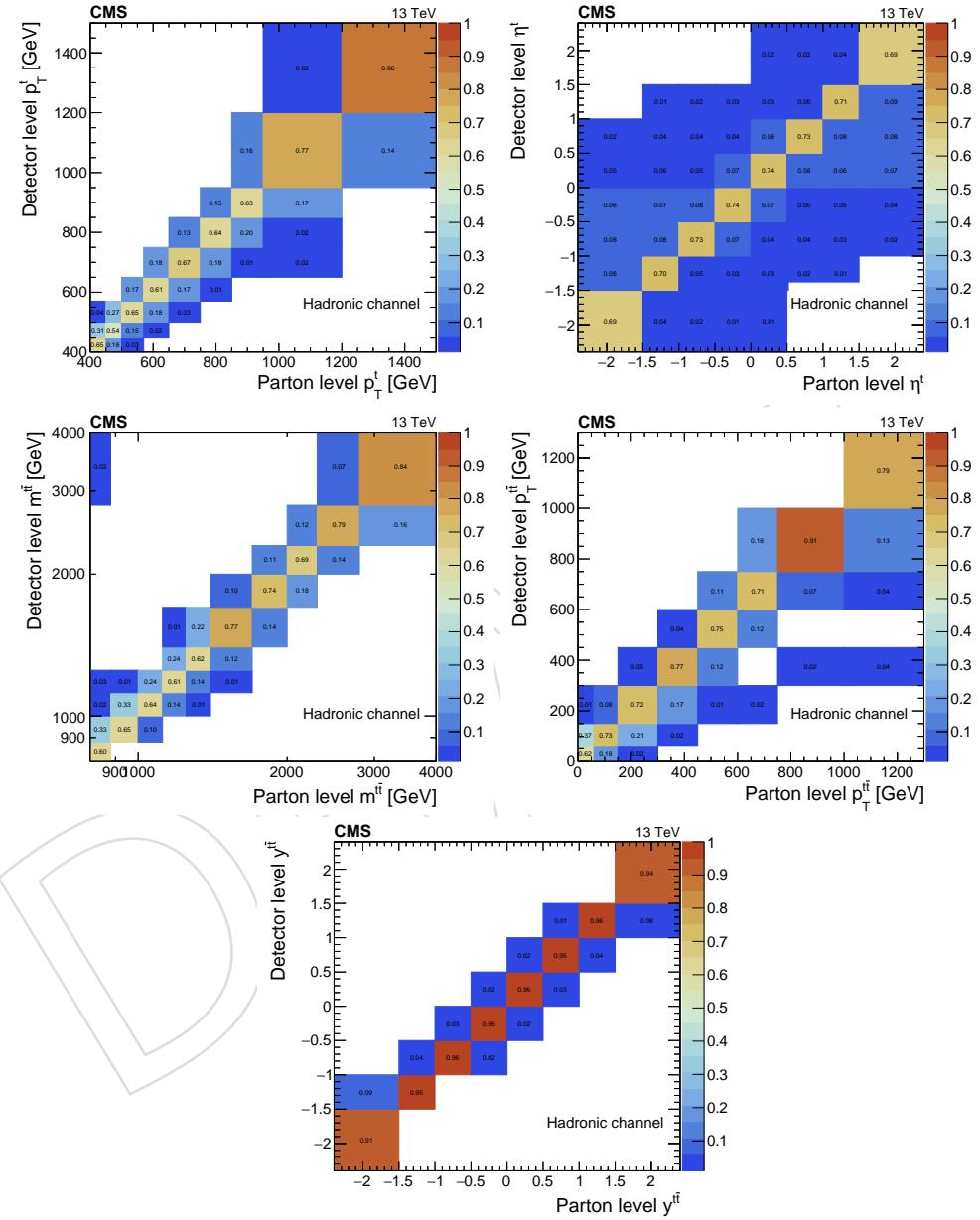


Figure 41: Simulated migration matrices at parton level. Each column is normalized to unity.

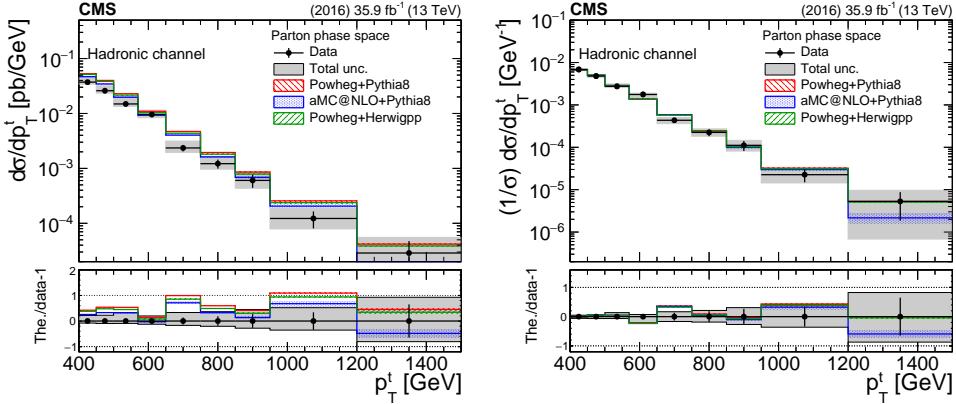


Figure 42: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of top p_T^t . The bottom panel shows the ratio (theory - data)/data.

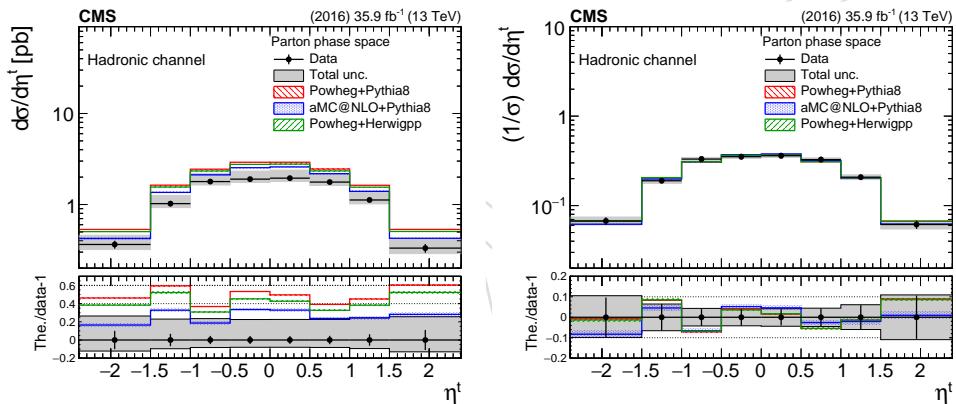


Figure 43: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of top η^t . The bottom panel shows the ratio (theory - data)/data.

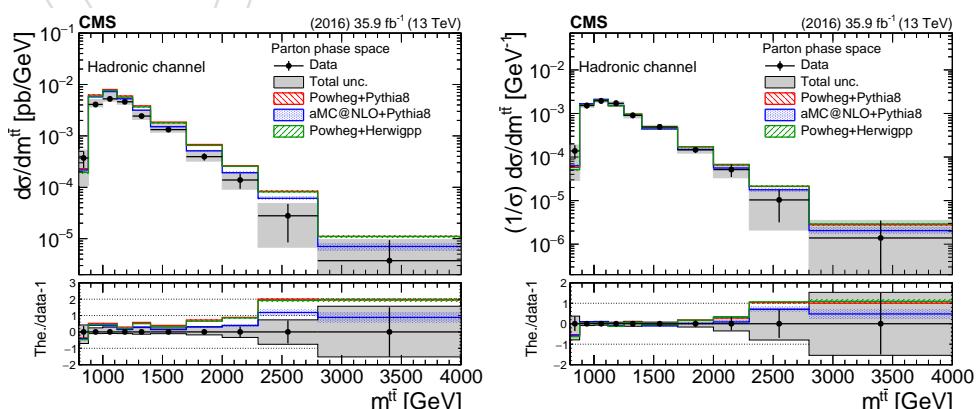


Figure 44: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of m_t^t . The bottom panel shows the ratio (theory - data)/data.

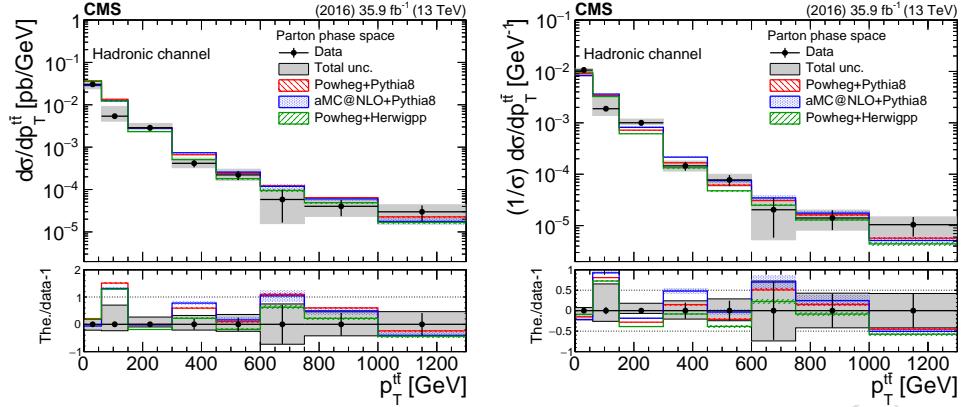


Figure 45: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of $p_T^t\bar{t}$. The bottom panel shows the ratio (theory - data)/data.

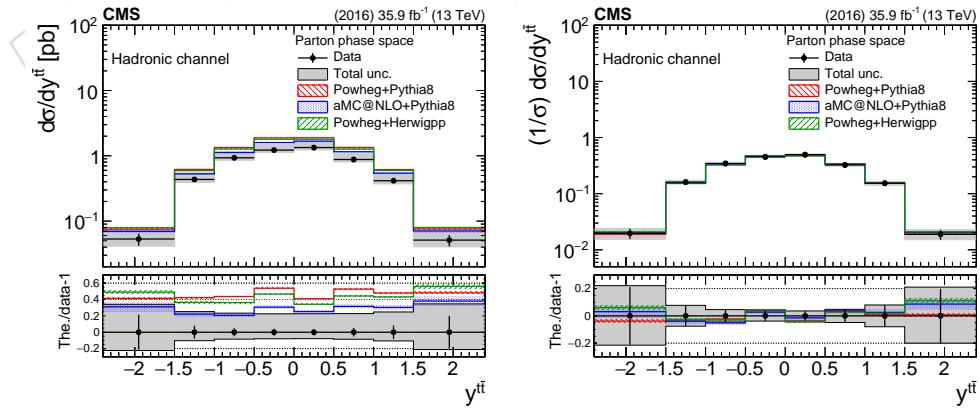


Figure 46: Differential cross section unfolded to parton level, absolute (left) and normalized (right), as a function of $y^t\bar{t}$. The bottom panel shows the ratio (theory - data)/data.

376 10.2 Particle Level

377 The so-called "particle level" represents the state that consists of stable particles originating
 378 from the proton-proton collision, after the hadronization process, and before the interaction of
 379 these particles with the detector. The observables computed from the particles' momenta are
 380 thought to be better defined compared to the ones computed from parton information and ac-
 381 companied by smaller theoretical uncertainties. Also, the associated phase space is closer to the
 382 fiducial phase space of the measurement at detector level. In the context of this analysis, parti-
 383 cle jets are reconstructed from stable particles, excluding neutrinos, with the anti-kt algorithm
 384 of distance parameter $R = 0.8$, identical to the detector-level reconstruction. It should be noted
 385 that only particles originating from the primary interaction are considered. Subsequently, jets
 386 that are geometrically matched, within $\Delta R < 0.4$ in $\eta - \phi$ from generated leptons (i.e. from the
 387 leptonic decays of the W boson) are removed from the particle-jet collection. Finally, the two
 388 particle jets with the highest p_T are considered the particle-level top-quark candidates. In order
 389 to match as closely as possible the fiducial phase space, the same kinematic cuts are applied as
 390 for detector-level events. These requirements are summarized in Table 10.

Table 10: Definition of particle-level phase space.

Observable	Requirement
N_{jets}	> 1
$p_T^{\text{jet1,2}}$	$> 400 \text{ GeV}$
$ \eta^{\text{jet1,2}} $	< 2.4
$m_{SD}^{\text{jet1,2}}$	(120, 220) GeV
m_{jj}	$> 800 \text{ GeV}$

391 In order to verify the sanity of the adopted particle-level definition, in each event that passes
 392 the requirements above, the top candidates are matched within $\Delta R < 0.4$ in $\eta - \phi$ to the original
 393 top quarks at parton level. Figure 47 shows this efficiency, which varies between 96% and 98%,
 394 as a function of particle-level top p_T and η .

395 The results of the unfolded measurement at particle level are shown in Figs. 51- 55. We observe
 396 that the uncertainties are similar to those of the parton level, while the comparison with the
 397 theory predictions follows the trends observed in the fiducial measurement. That is, there is
 398 an 20 – 40% lower inclusive cross section, while the shapes of the differential distributions
 399 are reasonably reproduced by all models. A hint of a deviation at very high values of $m^{t\bar{t}}$ is
 400 observed but the measurement is not very precise at this part of the phase space.

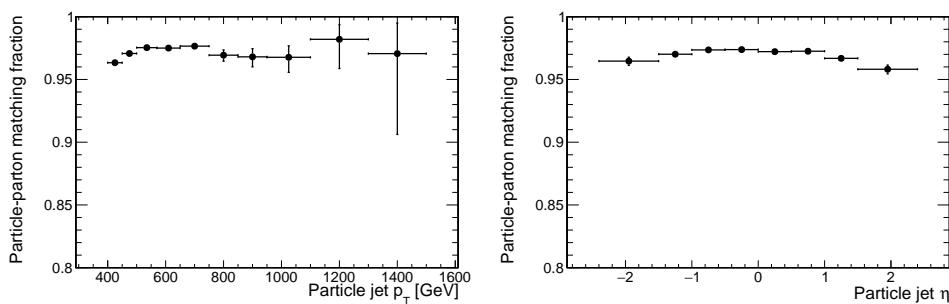


Figure 47: Simulated efficiency of the particle-level top candidates as a function of p_T and η .

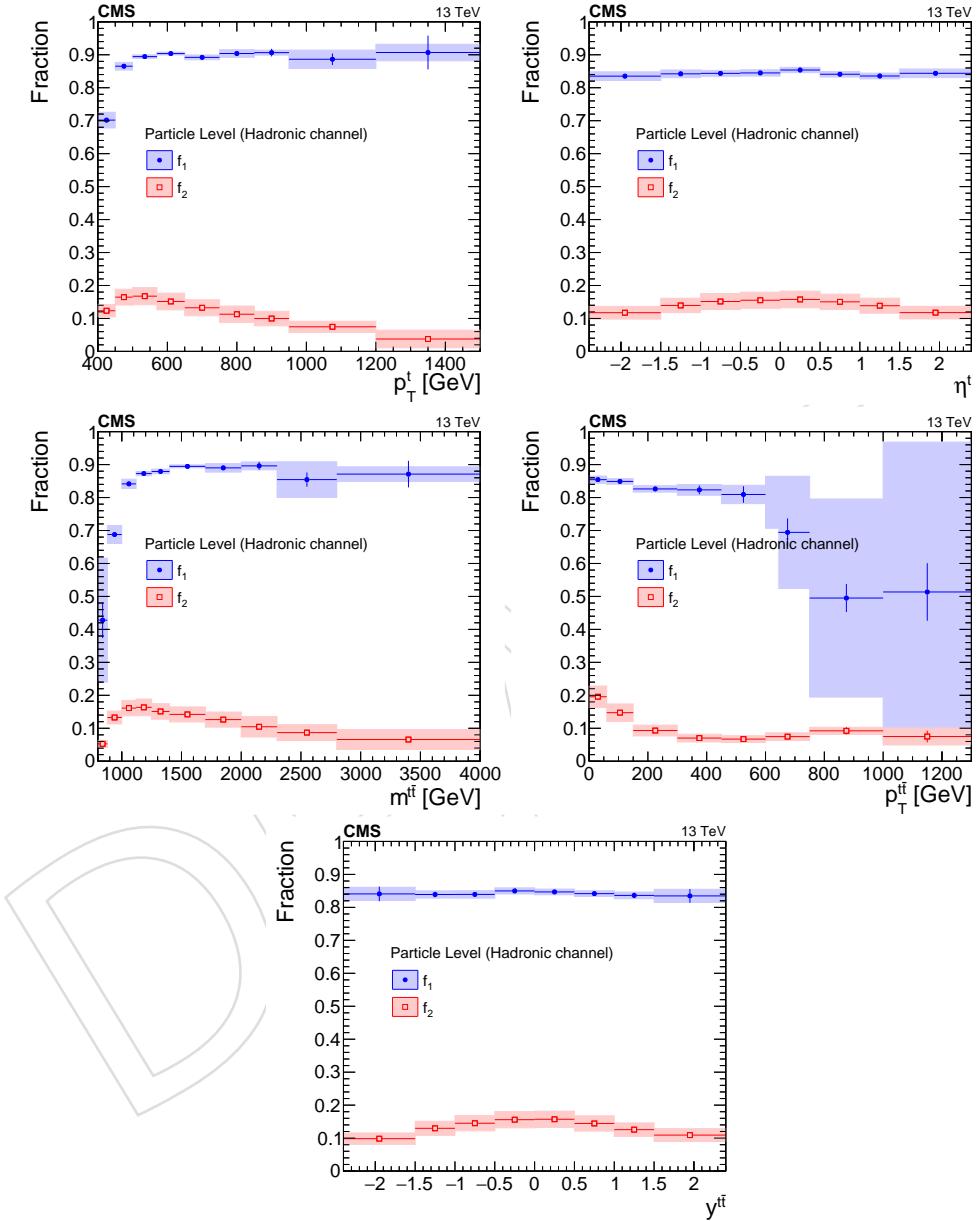


Figure 48: Simulated acceptance and efficiency for the particle-level selection as a function of the various observables.

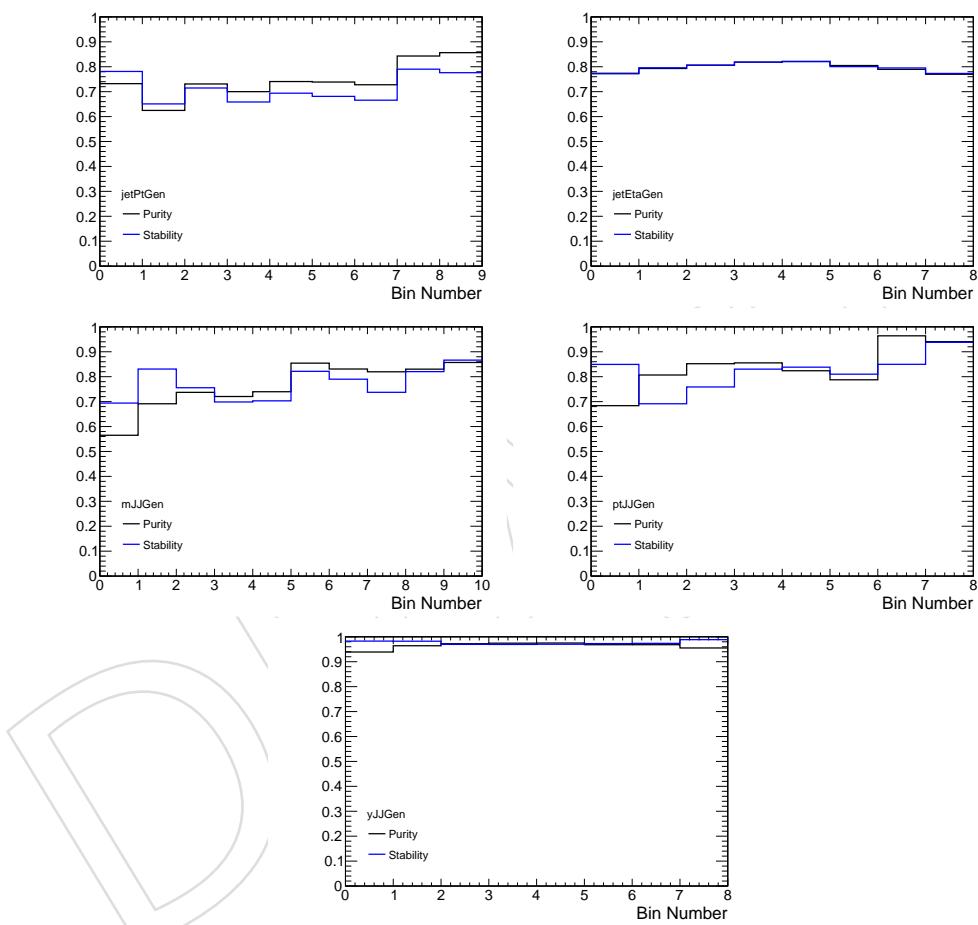


Figure 49: Simulated purity and stability of each bin at particle level.

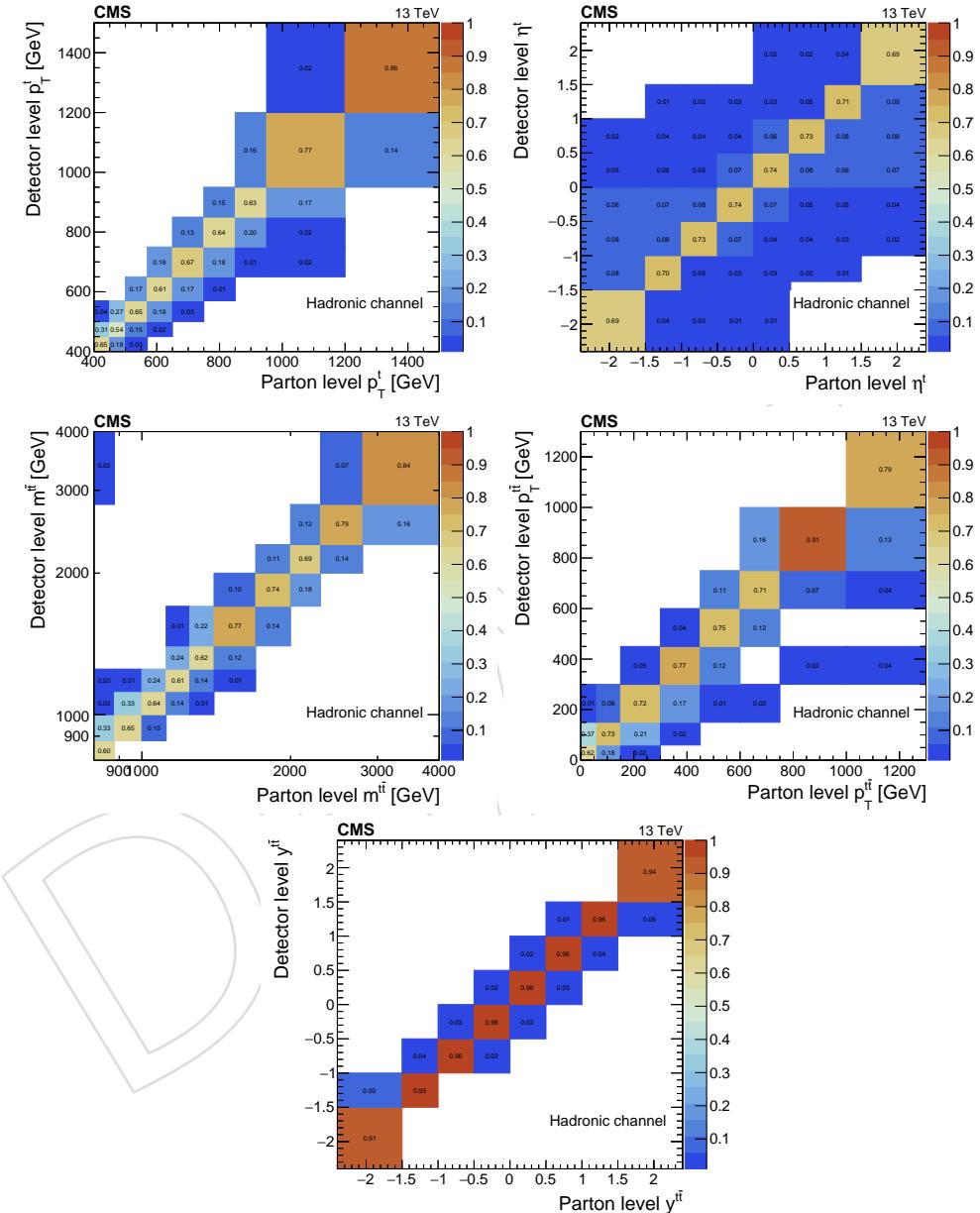


Figure 50: Simulated migration matrices at particle level. Each column is normalized to unity.

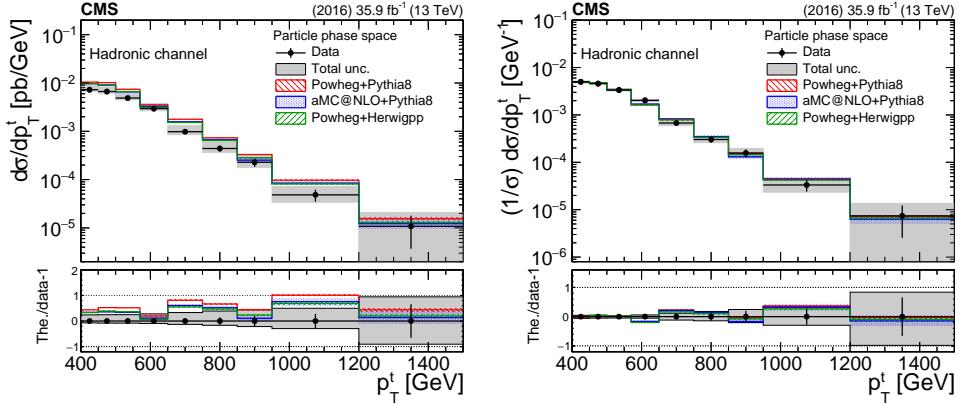


Figure 51: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of top p_T . The bottom panel shows the ratio (theory - data)/data.

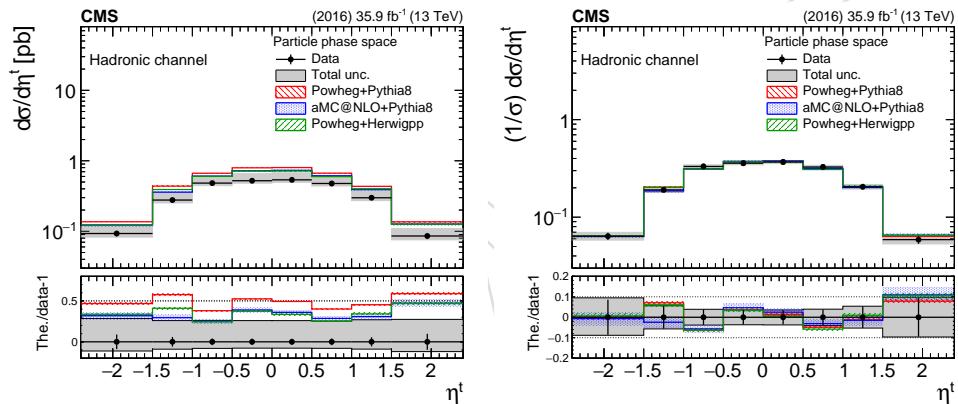


Figure 52: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of top p_T . The bottom panel shows the ratio (theory - data)/data.

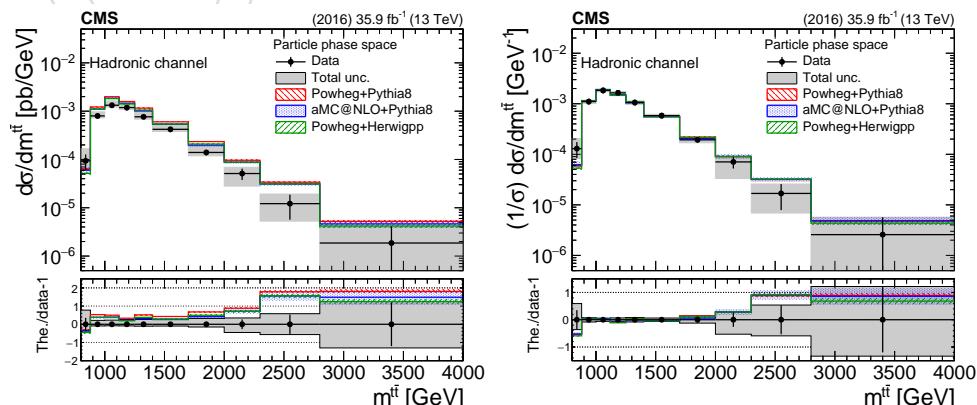


Figure 53: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of $m_t\bar{t}$. The bottom panel shows the ratio (theory - data)/data.

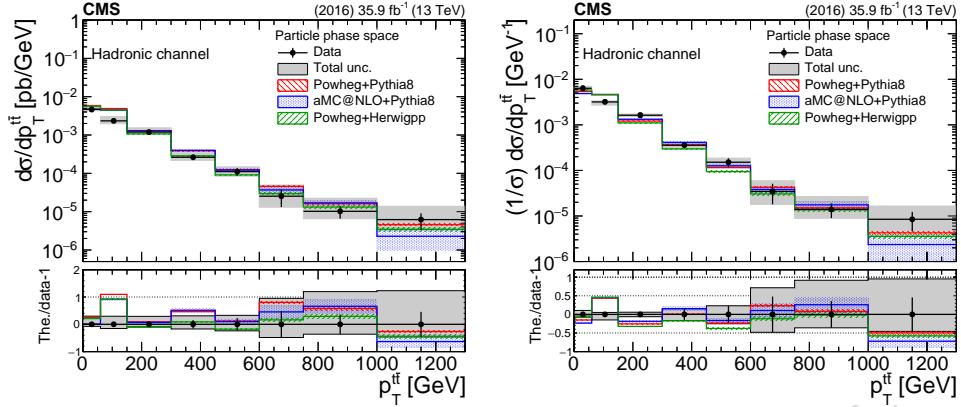


Figure 54: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of p_T^t . The bottom panel shows the ratio (theory - data)/data.

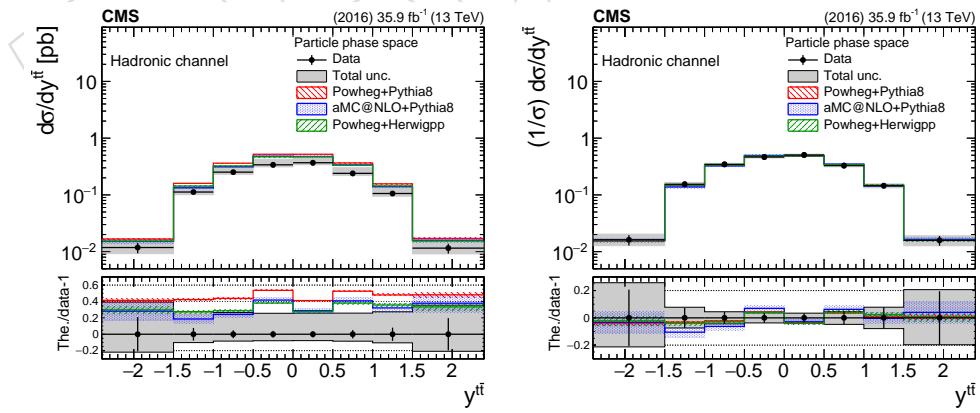


Figure 55: Differential cross section unfolded to particle level, absolute (left) and normalized (right), as a function of y^t . The bottom panel shows the ratio (theory - data)/data.

401 11 Summary

402 We have performed a measurement of the $t\bar{t}$ production cross section in the high top p_T region,
403 where the top decay products cannot be resolved and are instead clustered in a large radius jet,
404 in the all-hadronic final state. The cross section is reported differentially as a function of top
405 p_T and η (inclusive), and as a function of the invariant mass, p_T , and rapidity of the $t\bar{t}$ system,
406 unfolded to the parton and particle levels, absolute and normalized. The results are compared
407 to theoretical predictions from the POWHEG matrix-element generator, interfaced with PYTHIA
408 8 or for the underlying event and parton shower, and from the MC@NLO matrix-element gen-
409 erator, interfaced with PYTHIA 8. All the models overpredict significantly the absolute cross
410 section in the phase space of the measurement (up to 40%), while they are able to describe con-
411 sistently the differential shapes in all the variables. The most notable discrepancy can be seen
412 on the invariant mass of the $t\bar{t}$ system, where the theoretical models predict higher cross section
413 at high mass values. However, in order to estimate the significance of the discrepancy, more
414 data are needed.

A large, semi-transparent watermark consisting of the word "DRAFT" in a bold, sans-serif font. The letters are oriented diagonally, sloping upwards from the bottom-left towards the top-right. The watermark is rendered in a very light gray color, making it less prominent against the white background of the page.

415 **References**

DRAFT

416 A Control Distributions

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

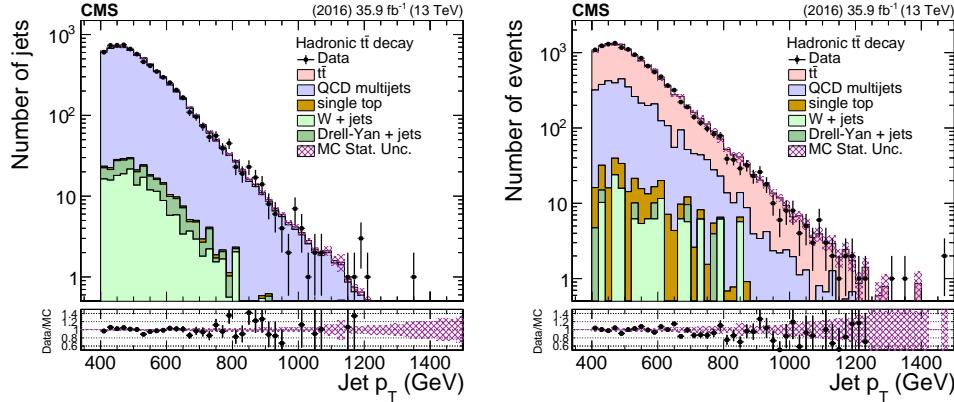


Figure 56: Data vs simulation in the control (left) and signal (right) regions for the p_T of the two leading jets.

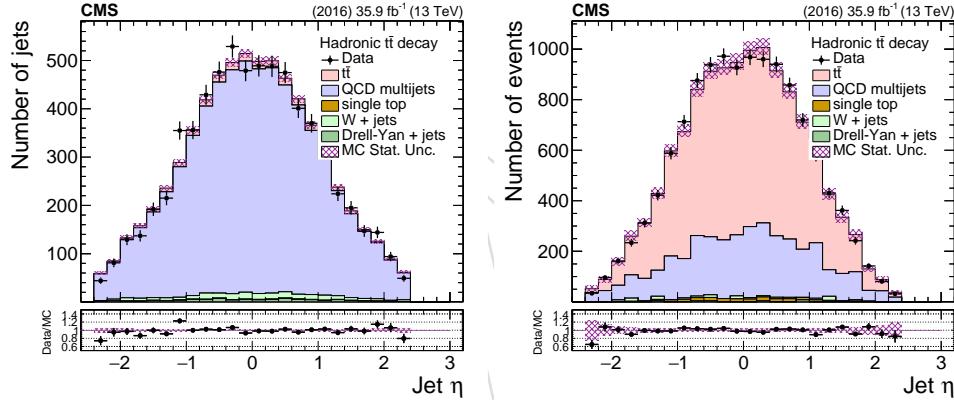


Figure 57: Data vs simulation in the control (left) and signal (right) regions for the η of the two leading jets.

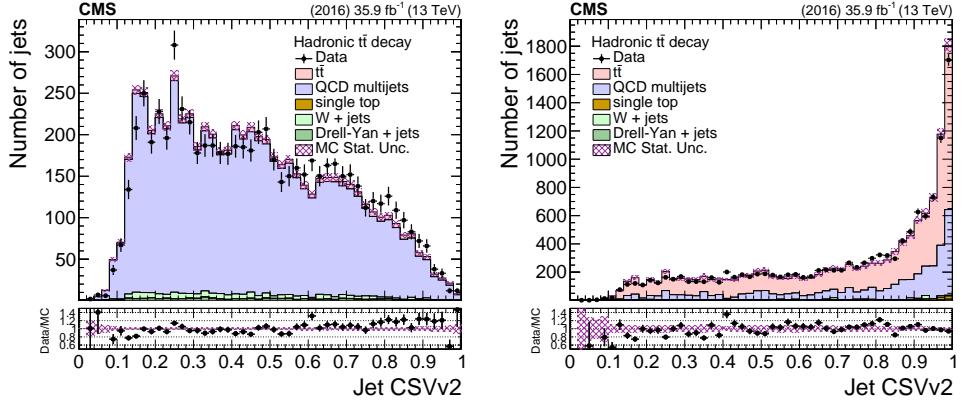


Figure 58: Data vs simulation in the control (left) and signal (right) regions for the CSVv2 of the two leading jets.

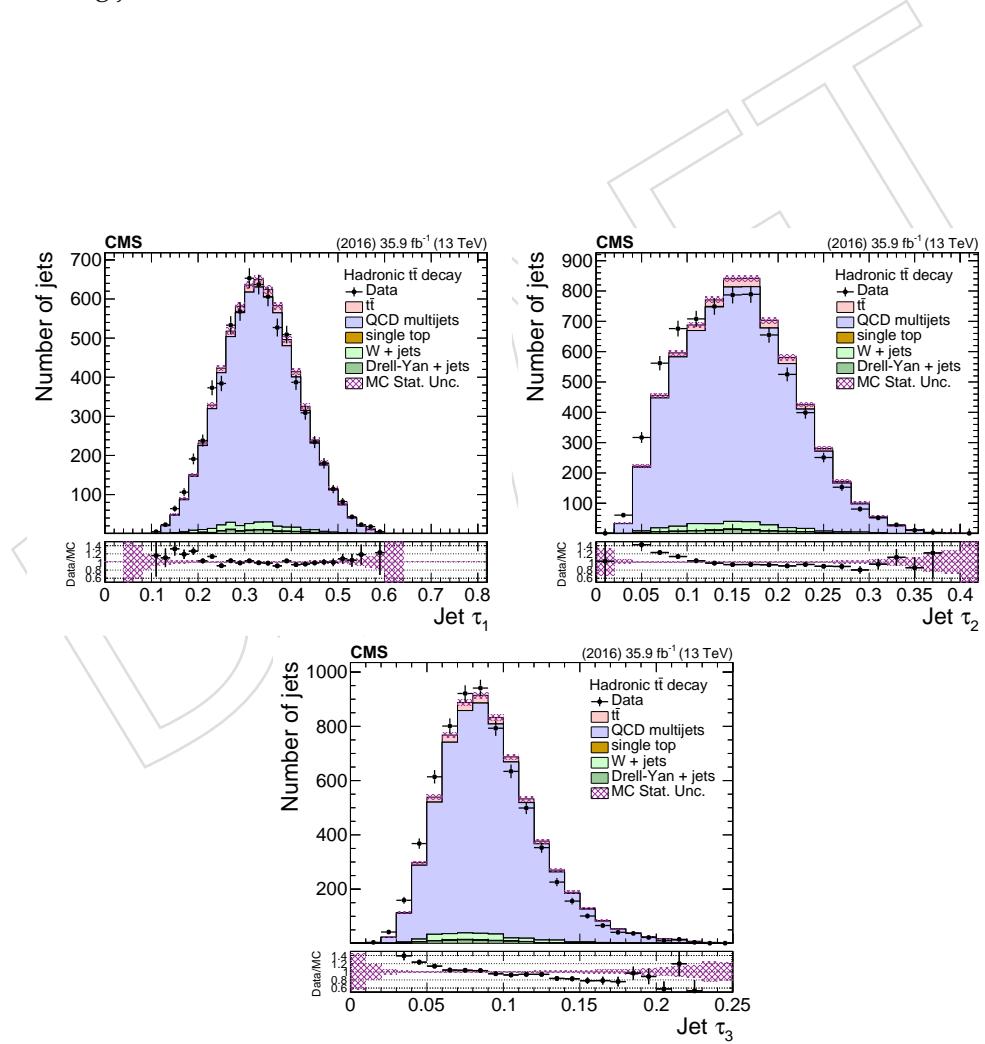


Figure 59: Data vs simulation in the control region for the n-subjettiness of the two leading jets.

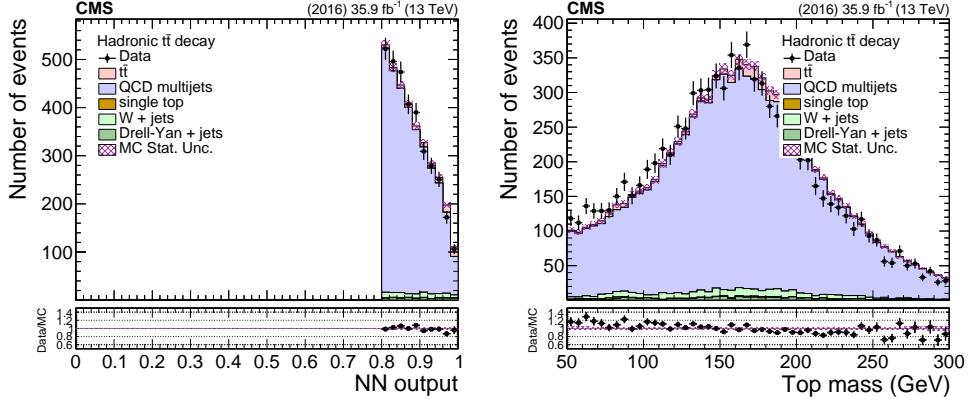


Figure 60: Data vs simulation in the control region for the NN output (left) and the top mass estimator (right).

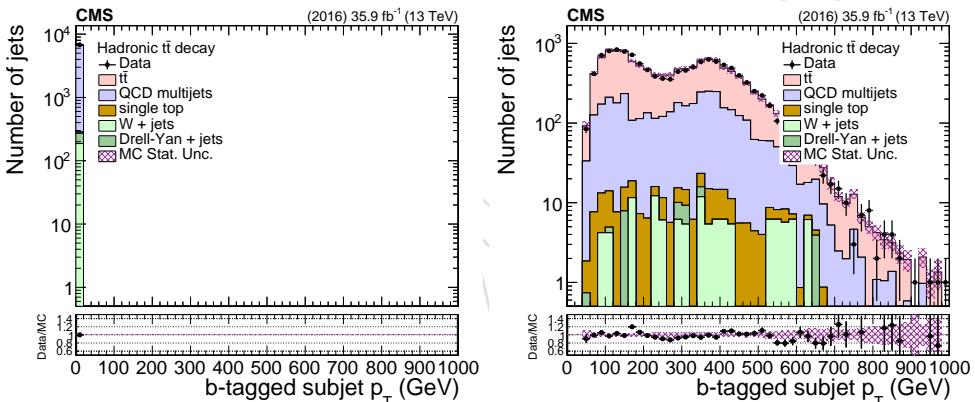


Figure 61: Data vs simulation in the control (left) and signal (right) regions for the p_T of the b-tagged subjet of the two leading jets.

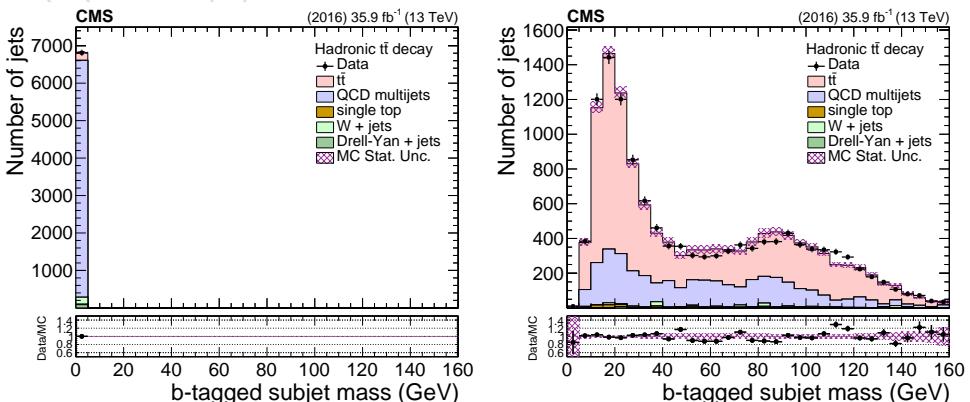


Figure 62: Data vs simulation in the control (left) and signal (right) regions for the mass of the b-tagged subjet of the two leading jets.

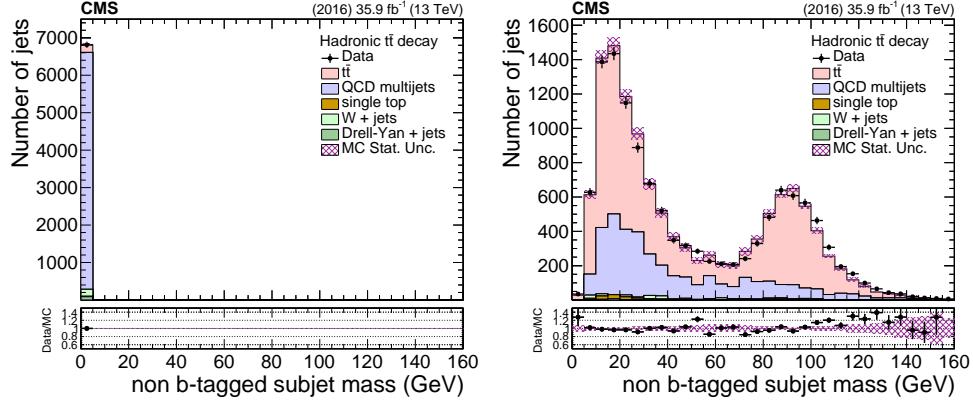


Figure 63: Data vs simulation in the control (left) and signal (right) regions for the mass of the non b-tagged subjet of the two leading jets.

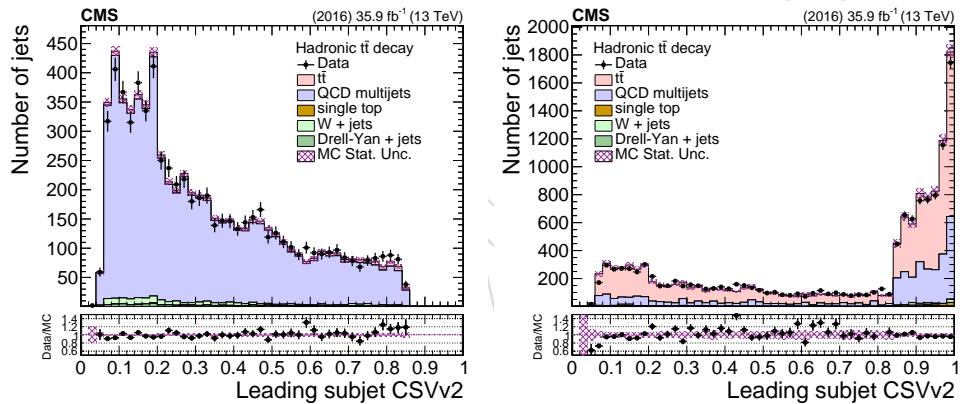


Figure 64: Data vs simulation in the control (left) and signal (right) regions for the CSVv2 of the leading subjet of the two leading jets.

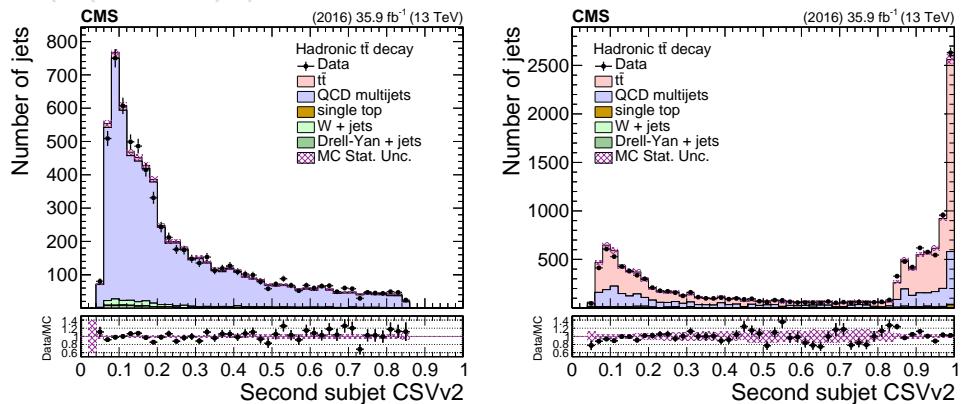


Figure 65: Data vs simulation in the control (left) and signal (right) regions for the CSVv2 of the subleading subjet of the two leading jets.

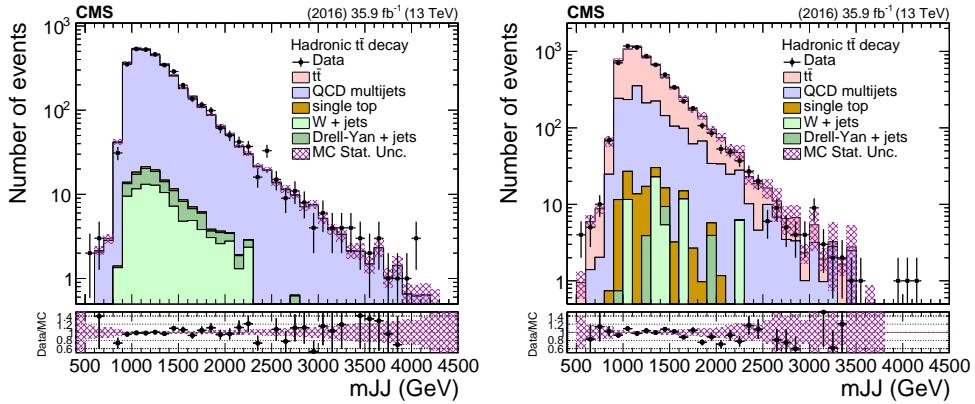


Figure 66: Data vs simulation in the control (left) and signal (right) regions dijet invariant mass.

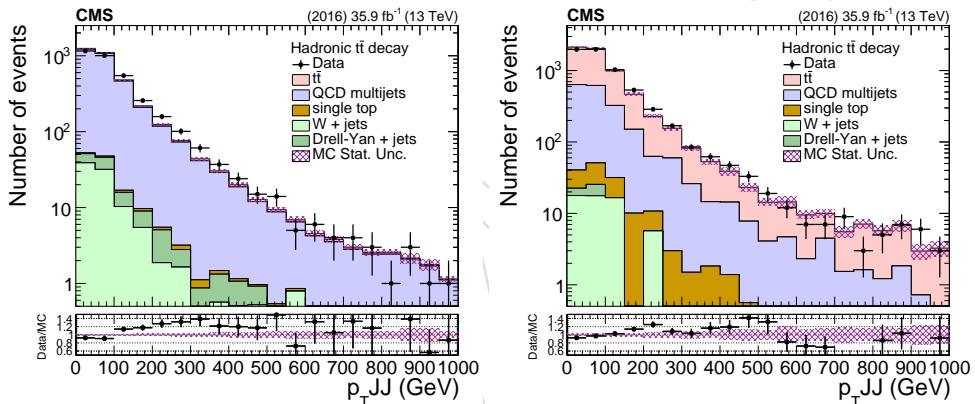


Figure 67: Data vs simulation in the control (left) and signal (right) regions dijet p_T .

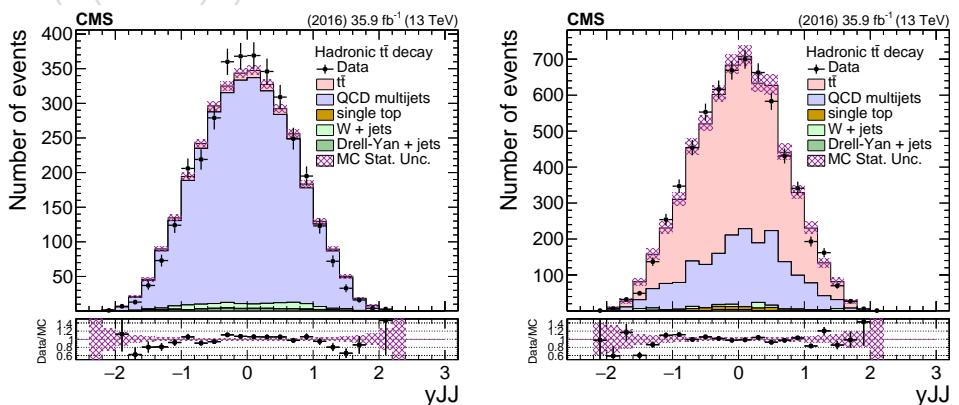


Figure 68: Data vs simulation in the control (left) and signal (right) regions dijet rapidity.

419 B Background Sensitivity to Pileup

420 The QCD background shape is taken from the corresponding control region in data, as de-
 421 scribed in Section 6. Since the data are collected with a prescaled trigger path, the pileup profile
 422 is different from the one of the signal events (Fig. 1). In order to demonstrate the robustness
 423 of the QCD prediction between the signal and control regions, we have performed the follow-
 424 ing test: the data from the control region are split in two subsets according to the number of
 425 reconstructed vertices (< 15 and ≥ 15) and we compare in Fig. 69 the distributions of the ob-
 426 servables used in this analysis. A systematic difference is observed in the softdrop mass of the
 427 jets, which is expected to be the most sensitive to pileup. This difference is naturally accounted
 428 for in the extraction of the QCD normalization (Eq. 3) with the unconstrained nuisance param-
 429 eter k_{slope} . In contrast, the shape of the jet and dijet kinematic observables are fully compatible
 430 within the statistical uncertainty. Therefore, taking also into account that the difference of the
 431 pileup profile of the two trigger paths is not as extreme as the test conducted here, we conclude
 432 that no significant bias is introduced by using the QCD shapes from the control region.

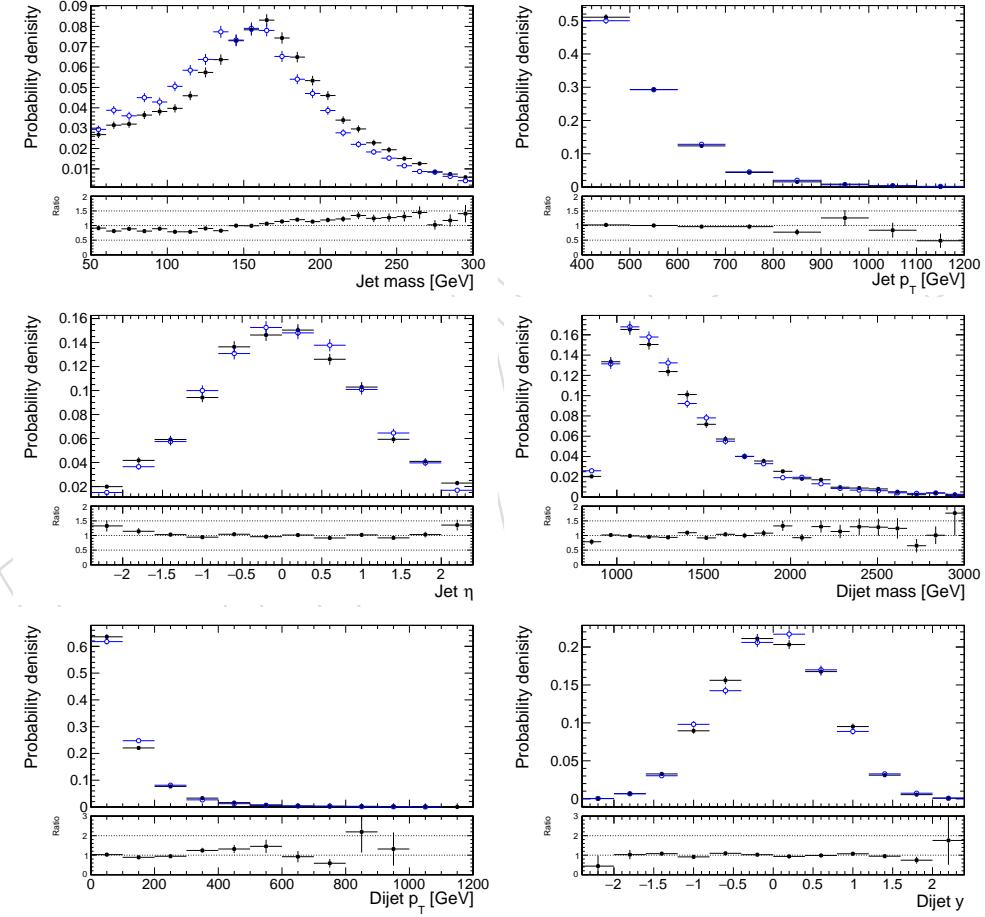


Figure 69: Distributions of jet and dijet observables from low pileup events (black, $\text{nvtx} < 15$) and high pileup events (blue, $\text{nvtx} \geq 15$). The bottom panel shows the ratio of the two.

433 **C Unfolding Tests**

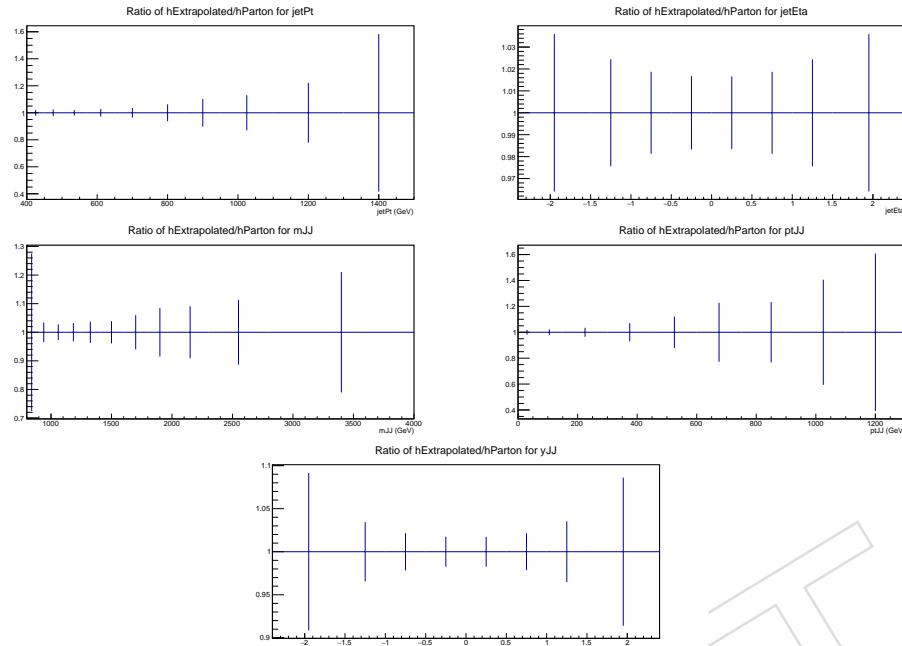


Figure 70: Parton unfolding closure test in simulation for all the reported observables. Ratio of the unfolded over the parton spectra.

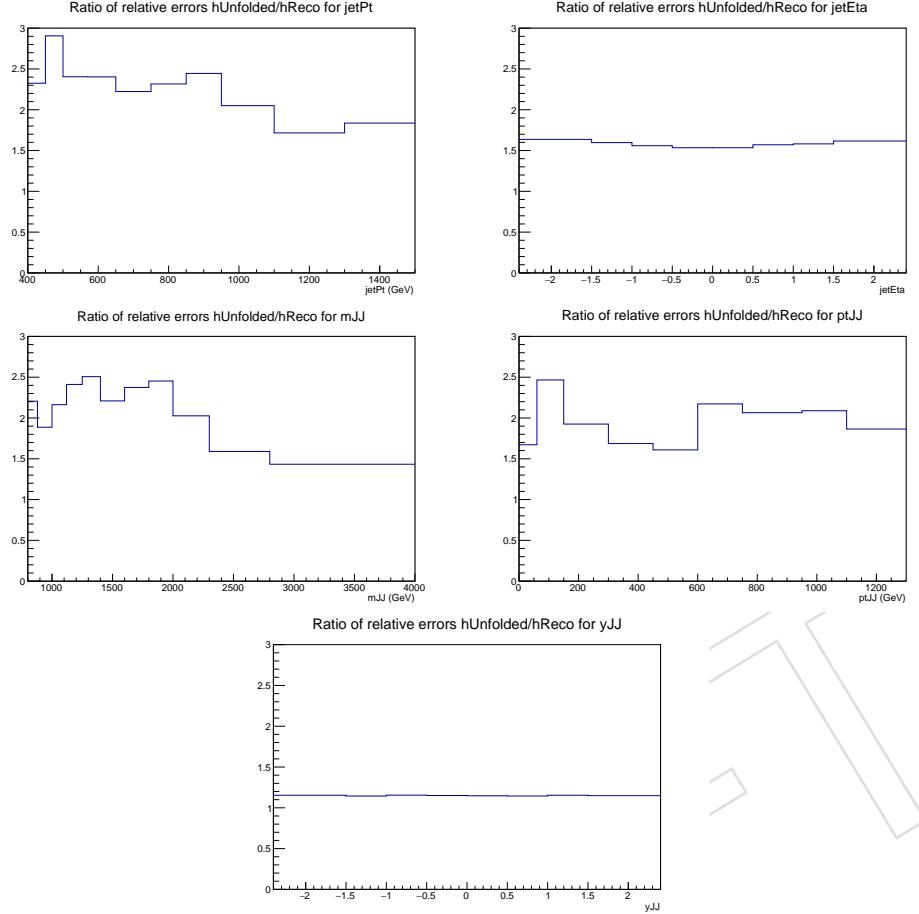


Figure 71: Ratio of the unfolded of the measured relative statistical uncertainties in the simulation (parton level).

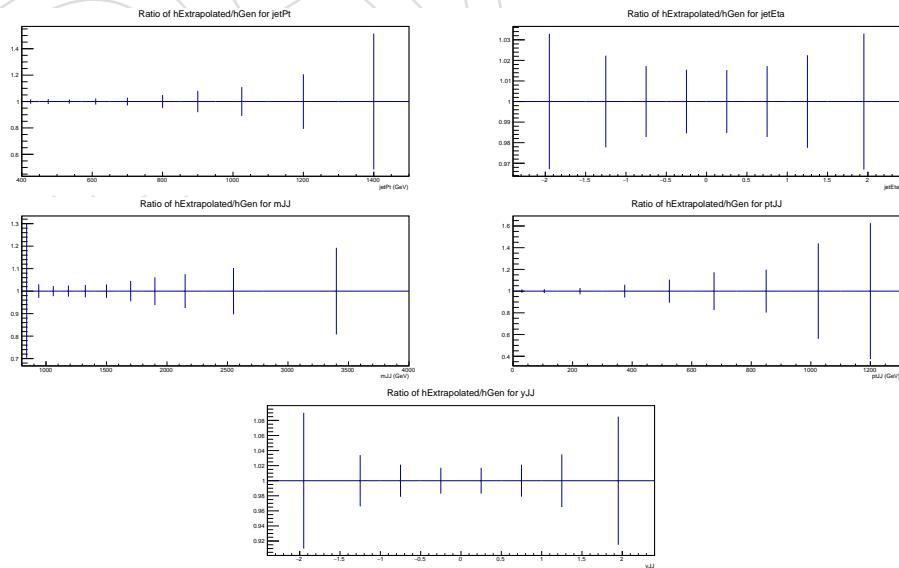


Figure 72: Particle unfolding closure test in simulation for all the reported observables. Ratio of the unfolded over the particle spectra.

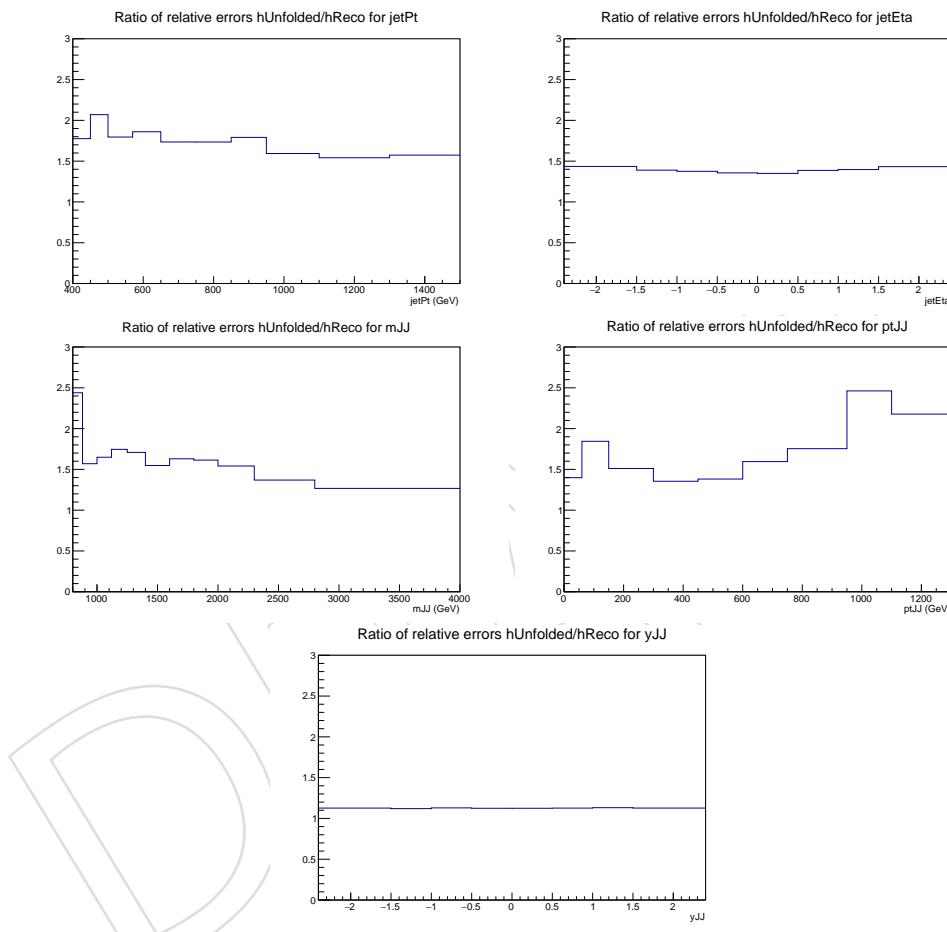


Figure 73: Ratio of the unfolded of the measured relative statistical uncertainties in the simulation (particle level).

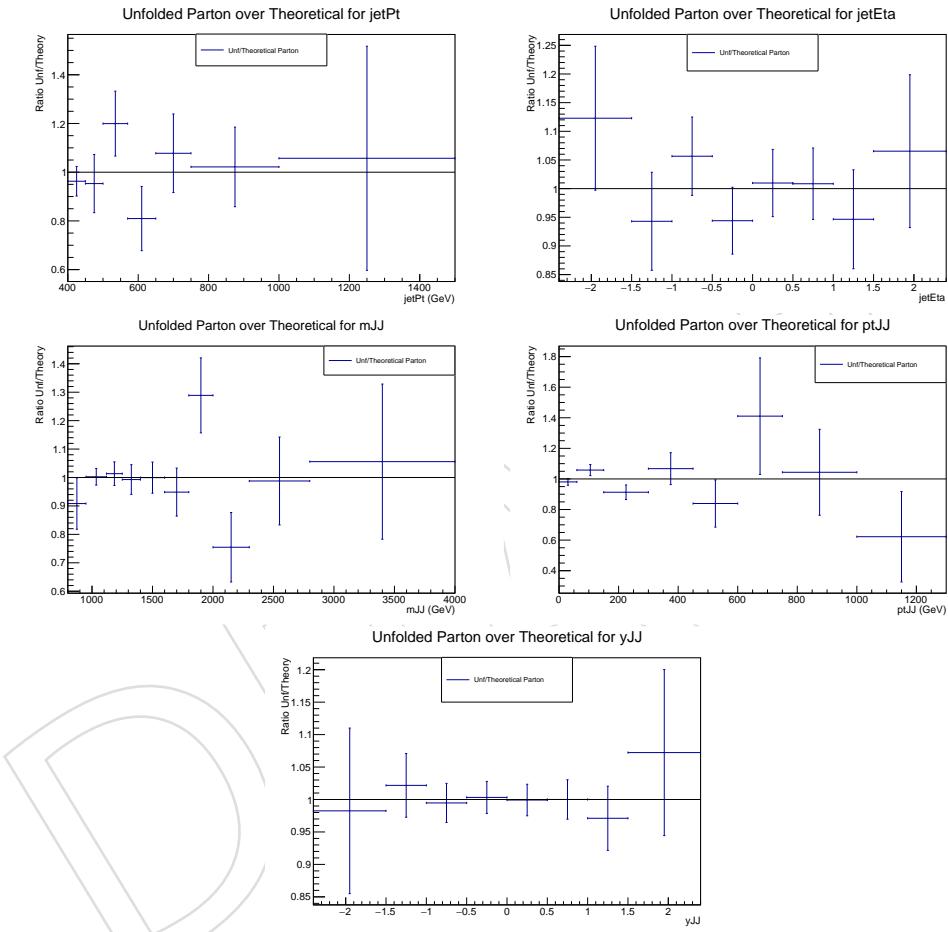


Figure 74: Stability of the unfolding at parton level in the simulation. The response matrix is derived from the first half of the sample and applied to the other half.

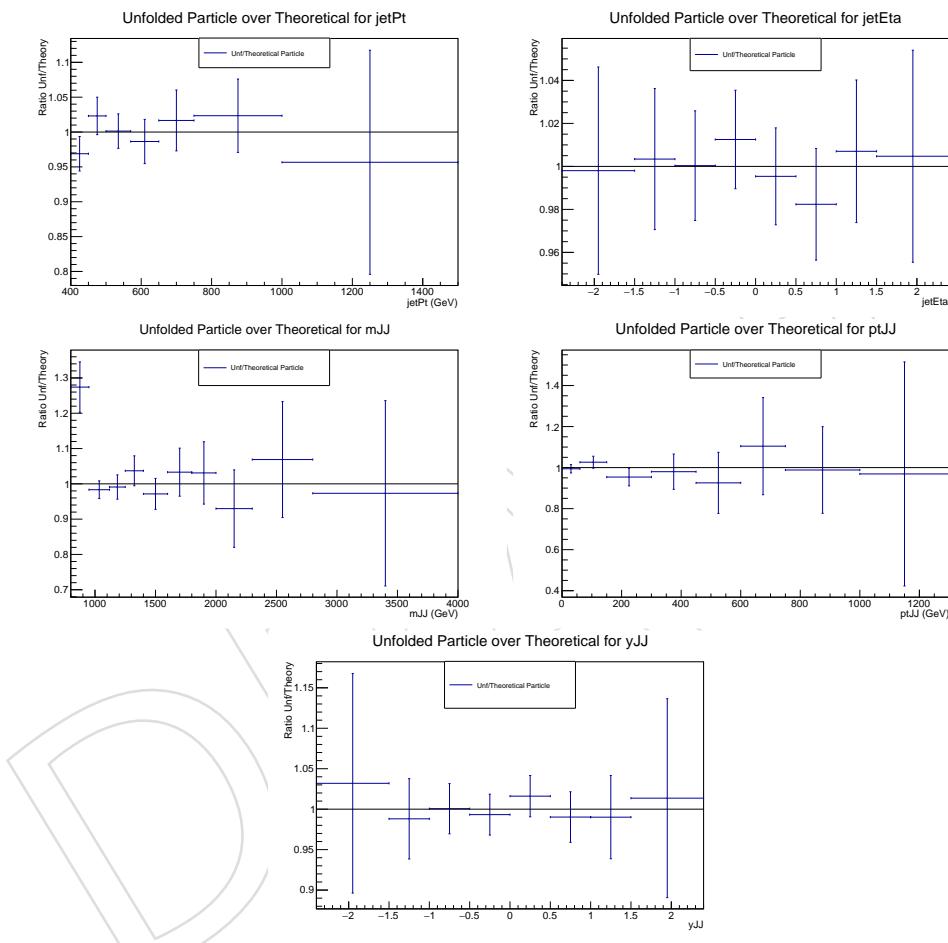


Figure 75: Stability of the unfolding at particle level in the simulation. The response matrix is derived from the first half of the sample and applied to the other half.