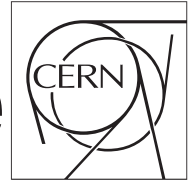


The Compact Muon Solenoid Experiment

# CMS Draft Note

Mailing address: CMS CERN, CH-1211 GENEVA 23, Switzerland



2019/10/01

Archive Hash: 95ba81d-D

Archive Date: 2019/10/01

## Extraction of CKM matrix elements in the single-top $t$ -channel events at 13 TeV

T. Chwalek<sup>1</sup>, N. Faltermann<sup>1</sup>, A. De Iorio<sup>2</sup>, A. O. M. Iorio<sup>2</sup>, W.A. Khan<sup>2</sup>, L. Lista<sup>2</sup>, Th. Müller<sup>1</sup>, P. Ott<sup>1</sup>, and T. Pambor<sup>1</sup>

<sup>1</sup> Institut für Experimentelle Kernphysik, Karlsruher Institut für Technologie, Germany

<sup>2</sup> Istituto Nazionale di Fisica Nucleare sez. Napoli, Italy

### Abstract

This analysis notes describes the analysis steps taken for the extraction of CKM matrix elements  $|V_{tq}|$  in the single-top-quark  $t$ -channel at 13 TeV. In the final state we have the events with one lepton ( $\mu/e$ ) coming from the leptonic decay of W-boson, along with two or three jets, one or two of which is a b-jet. The signal region is defined by having exactly two jets, one of which is b-tag and the events with three jets are used to estimated  $t\bar{t}$  contribution. Data driven technique is used for the extraction of QCD shape from a QCD enriched side band region. We deploy MVA to separate the  $t$ -channel events in which top-quark decays to a d- or s- quark from the standard background.

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

PDFAuthor:	single-top t-channel team
PDFTitle:	Extraction of CKM matrix elements in the single-top-quark t-channel events at 13 TeV
PDFSubject:	CMS
PDFKeywords:	CMS, physics, software, computing

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



# 1 Introduction

The three most important mechanisms for production of single-top quarks involve charged-current electroweak interactions, and thus happen mostly through a  $tWb$  vertex, with a  $b$ -quark coming from the splitting of a gluon from PDF. The main production mechanism happens through exchange of a  $W$ -boson in the  $t$ -channel. Fig. 1 shows the four-flavour scheme diagram at leading order for  $t$ -channel production.

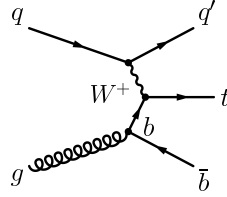


Figure 1: Leading order Feynman diagrams for single-top quark production via the  $t$ -channel in the four flavour scheme. The latter can also be seen as an NLO contribution in the five flavour scheme.

All the past analyses from Tevatron and LHC, e.g. [1–3], do exploit the cross section to measure  $|V_{tb}|$  under the assumption that other matrix elements are negligible, and only consider variation of  $V_{tb}$  in the decay.

This assumption stems from the indirect measurements of  $V_{td}$  and  $V_{ts}$  from independent lower energy measurements. However, this assumption ignores any BSM modifications to the magnitude of  $V_{td}$  and  $V_{ts}$  happening in production, and does introduce biases in measurements of  $|V_{tb}|$ . In this note we aim at measuring the elements  $V_{tb}$ ,  $V_{td}$ , and  $V_{ts}$  considering all contributions in the  $t$ -channel production and decay.

The note is organized as follows:

- Sec. 2 reports all the datasets and the MonteCarlo samples used for the analysis.
- Sec. 3 describes the event selection and object definition
- Sec. 4 describes the rationale of the measurement, the signal processes considered by including  $V_{tb}$ ,  $V_{td}$ ,  $V_{ts}$  and defines the different signal and control regions used for the measurements and to validate the modelling of the most important processes.
- Sec. 5 describes the fit procedure developed, including the impact of systematics.
- Sec. 6 describes in details all the source of systematics uncertainties for the measurement.
- Sec. 7 summarizes the result and its interpretations.

In all past single-top analyses the contribution of  $\beta|V_{td}|^2 + \gamma|V_{ts}|^2 = \lambda \cdot \epsilon \ll \alpha|V_{tb}|^2$ . This in practice means:

$$\sigma_{measured} = \alpha|V_{tb}|_{measured}^2. \quad (1)$$

Of course, for the theoretical cross section one can simply reduce the calculation to the part where the  $tWb$  vertex is present, thus we write:

$$\sigma_{theory} = \alpha|V_{tb}|_{theory}^2, \implies \quad (2)$$

$$\implies |V_{td}|_{measured} = |V_{tb}|_{theory} \sqrt{\sigma_{measured} / \sigma_{theory}} \quad (3)$$

This formula is subject to several approximations, as first of all we neglect the presence of the  $tW$ s and  $tWd$  production vertices, which give rise to new terms in the production:

$$\sigma_{t\text{-channel}} = \sigma_{t\text{-ch.,b}} + \sigma_{t\text{-ch.,s}} + \sigma_{t\text{-ch.,d}} \quad (4)$$

The other issue is that we currently measure the single-top cross section through events with the  $t \rightarrow Wb$  decay, meaning that we assume that the  $\text{BR}(t \rightarrow Wb)$  is identically 1.

In this section we will give a rough estimate of the bias that this procedure can bring and define topologies enriched with events having  $|V_{tq}| \neq 1$  in either production or decay.

## 1.1 Naming conventions

We refer here only to the  $t$ -channel single-top process. Usually a  $t$ -channel process foresees a  $tWb$  vertex in production and decay. Considering the decays  $\text{BR}(t \rightarrow Wq)$  where the  $W$ -boson decays leptonically, ( $W \rightarrow l\nu$ ), we have 3 jets in the final state, one stemming from the quark produced in the top-quark decay, one from the quark recoiling directly against the top-quark via  $W$ -boson exchange, and the last one from the initial state gluon splitting. Table 1 summarizes the acronyms that are introduced to make the note more readable.

Acronym	Description	Equivalent to
Jet selection acronyms		
NJMT	jet selection: N number of jets M number of tags (e.g. 2J1T)	
$\epsilon_j^{\text{sel}}$	efficiency for a jet $j$ to pass the kinematic requirements	
$\epsilon_j^{\text{tag}}$	efficiency for a jet $j$ to pass the b-tagging requirement	
$\epsilon_j^T$	efficiency for a jet to pass kin. AND tagging requirements,	$\epsilon^{\text{sel}} * \epsilon^{\text{tag}}$
$\epsilon_j^J$	efficiency for a jet to pass kin. AND NOT tagging requirements	$\epsilon^{\text{sel}} * (1 - \epsilon^{\text{tag}})$
Channel acronyms used to define the physics process/decay channel		
$\sigma_{t\text{-ch.,q}}$	cross section with a $tWq$ vertex, $q=b, s, \text{ or } d$ quark.	$\propto V_{tb(q)}$
$R_q$	top-quark decay fraction in $Wq$ , $q=b, s, \text{ or } d$ quark.	$\text{BR}(t \rightarrow Wq) \propto$
$ST^{\text{NJMT}}_{(q_1, q_2)}$	selection efficiency for $(\sigma_{t\text{-ch.,q}_1} \cdot R_{q_2})$ , with $q_1$ or $q_2 = b, s, \text{ or } d$	$\propto  V_{tq_1} ^2 \cdot \frac{ V_{tq_2} ^2}{ V_{tb} ^2 +  V_{td} ^2 +  V_{ts} ^2}$
Final state jets acronyms(from the quark they originate from)		
$q_{\text{top}}$	jet from a quark $q$ stemming from top-quark decay	
$q_{\text{light}}$	jet from a quark $q$ recoiling against the top	
$q_{2\text{nd}}$	jet from a quark $q$ produced the hard gluon splitting	

Table 1: Table with the naming convention used in this document.

## 1.2 Processes involved in the $t$ -channel and selection efficiencies

There are three contributions to the inclusive single top quark  $t$ -ch. cross section:

$$\begin{aligned} \sigma_{t\text{-ch.}} &= \sigma_{t\text{-ch.,b}} + \sigma_{t\text{-ch.,d}} + \sigma_{t\text{-ch.,s}} = \\ &= \alpha |V_{tb}|^2 + \beta |V_{td}|^2 + \gamma |V_{ts}|^2, \end{aligned} \quad (5)$$

where the term  $|V_{tq}|^2$ , where  $q$  stands for  $b, s, \text{ or } d$  is factored out from the cross section. The values of the quantities in Eq. 5 are reported in Table 2.

When the top quark decay modes are considered, the cross section is multiplied for the branch-

Vertex	tWb	tWd	tWs
$ V_{tq} $	$0.999119^{+0.000024}_{-0.000012}$	$0.008575^{+0.000076}_{-0.000098}$	$0.04108^{+0.00030}_{-0.00057}$
$ V_{tq} ^2$	$0.998239^{+0.000048}_{-0.000024}$	$0.000074^{+0.000013}_{-0.000017}$	$0.0016876^{+0.0000025}_{-0.0000047}$
$\sigma_{t\text{-ch},q}$	$216.99 \pm 5.8$	$1046.29 \pm 10$	$498.61 \pm 10$

Table 2: Values of matrix elements inferred from low energy regime measurements, with the respective values of top quark decay branching fractions in the first two rows. Cross sections for inclusive  $t$ -ch. production in the third row, obtained with POWHEG 2.0 when the respective  $|V_{tq}|$  is put equals to unity.

ing fraction  $\text{BR}(t \rightarrow Wq)$  that can be evaluated<sup>1</sup> as:

$$\text{BR}(t \rightarrow Wq) = \frac{|V_{tq}|^2}{|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2}. \quad (6)$$

39 All the previous approaches consider  $|V_{tb}| \gg |V_{td,s}|$  and neglect the latter terms. This assumption  
40 stems from the indirect determinations of  $V_{td}$  and  $V_{ts}$  from independent lower energy  
41 measurements. However, this assumption ignores any BSM modifications to the magnitude of  
42  $V_{td}$  and  $V_{ts}$  happening in production, and does introduce biases in measurements of  $|V_{tb}|$ .

### 43 1.2.1 The scenario with unitarity preserved

Within the SM scenario, the branching ratio can be written as in equation 6. If one considers also any other model in which other decays of the top quark can occur, then the same quantity could be written as  $\text{BR}tq = \frac{|V_{tq}|^2}{|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2 + X}$ , with  $X$  being any other final state. If the top quark decays only to quarks, with enhanced couplings by BSM scenarios, or if for any other reason the denominator is still equals to 1, one can use the same formulas as in the SM case. With such assumptions one obtains:

$$\begin{aligned} \sigma_{t\text{-ch.}} \times \text{BR}(t \rightarrow Wq) &\cong \alpha |V_{tb}|^2 |V_{tb}|^2 + \\ &+ \alpha |V_{tb}|^2 (|V_{td}|^2 + |V_{ts}|^2) + \\ &+ \beta |V_{td}|^2 |V_{tb}|^2 \\ &+ \gamma |V_{ts}|^2 |V_{tb}|^2, \end{aligned} \quad (7)$$

44 where the  $|V_{tq}|^2 |V_{tq}|^2$ , with  $q=d, s$ , contributions were neglected in this approximation.

When the measurement is performed the selection efficiencies,  $\epsilon_{q,q}^{\text{sel.}}$ , obtained from the MC simulations, have to been taken into account and, finally, one has:

$$\begin{aligned} \sigma_{t\text{-ch.}} \times \text{BR}(t \rightarrow Wq) &\cong \alpha |V_{tb}|^2 |V_{tb}|^2 \epsilon_{b,b}^{\text{sel.}} + \\ &+ \alpha |V_{tb}|^2 (|V_{td}|^2 \epsilon_{b,d}^{\text{sel.}} + |V_{ts}|^2 \epsilon_{b,s}^{\text{sel.}}) + \\ &+ \beta |V_{td}|^2 |V_{tb}|^2 \epsilon_{d,b}^{\text{sel.}} \\ &+ \gamma |V_{ts}|^2 |V_{tb}|^2 \epsilon_{s,b}^{\text{sel.}}. \end{aligned} \quad (8)$$

<sup>1</sup>This is an approximation neglecting all non-W decays of top quarks and the difference in mass between  $b$ ,  $s$ , and  $d$  quarks.

Summing non standard production and decay:

$$\begin{aligned} \sigma_{t\text{-ch.}} \times \text{BR}(t \rightarrow Wq) &\cong \alpha |V_{tb}|^2 |V_{tb}|^2 + \\ &+ \alpha |V_{tb}|^2 \left( |V_{td}|^2 + \frac{\beta}{\alpha} |V_{td}|^2 + |V_{ts}|^2 + \frac{\gamma}{\alpha} |V_{ts}|^2 \right) \end{aligned} \quad (9)$$

By fixing the ratios  $\frac{\beta}{\alpha}$  and  $\frac{\gamma}{\alpha}$  to the SM prediction, as reported in Table 3:

$$\begin{aligned} \sigma_{t\text{-ch.}} \times \text{BR}(t \rightarrow Wq) &\cong \alpha |V_{tb}|^2 |V_{tb}|^2 + \alpha |V_{tb}|^2 (7.64 \cdot |V_{td}|^2 + 3.3 \cdot |V_{ts}|^2) \cong \\ &\cong \alpha |V_{tb}|^2 |V_{tb}|^2 + 3.3 \cdot \alpha |V_{tb}|^2 (2.29 \cdot |V_{td}|^2 + |V_{ts}|^2) \end{aligned} \quad (10)$$

and this leads to:

$$\begin{aligned} \sigma_{t\text{-ch.}} \times \text{BR}(t \rightarrow Wq) &\cong \alpha |V_{tb}|^2 |V_{tb}|^2 + \alpha |V_{tb}|^2 (7.64 \cdot |V_{td}|^2 + 3.3 \cdot |V_{ts}|^2) \cong \\ &\cong \alpha |V_{tb}|^2 |V_{tb}|^2 + 3.3 \cdot \alpha |V_{tb}|^2 (|V_{td}|^2 + |V_{ts}|^2). \end{aligned} \quad (11)$$

This approximation is valid under the assumption:

$$\begin{aligned} 3.3 \cdot \alpha |V_{tb}|^2 (2.29 |V_{td}|^2 + |V_{ts}|^2) &= 3.3 \cdot \alpha |V_{tb}|^2 (|V_{td}|^2 + |V_{ts}|^2) + \\ &+ \mathcal{O}(3.3 \cdot \alpha |V_{tb}|^2 \cdot 1.29 |V_{td}|^2), \end{aligned} \quad (12)$$

since the ratio  $\frac{|V_{td}|^2}{|V_{ts}|^2}$  is well known from low energy measurements and it is of order  $4 \times 10^{-2}$ .

### 1.2.2 Non-unitarity scenarios

In case other decays of the top could occur, the unitarity relation  $|V_{tb}|^2 + |V_{td}|^2 + |V_{ts}|^2 = 1$  would be broken. In this case, by following the same steps of the previous lines, the computation leads to:

$$\begin{aligned} \sigma_{t\text{-ch.}} \times \text{BR}(t \rightarrow Wq) &\cong \alpha |V_{tb}|^2 \frac{|V_{tb}|^2}{|V_{tb}|^2 + |V_{td}|^2 + |V_{ts}|^2} + \\ &+ \alpha \frac{|V_{tb}|^2}{|V_{tb}|^2 + |V_{td}|^2 + |V_{ts}|^2} (7.64 \cdot |V_{td}|^2 + 3.3 \cdot |V_{ts}|^2) \cong \\ &\cong \alpha |V_{tb}|^2 \frac{|V_{tb}|^2}{|V_{tb}|^2 + |V_{td}|^2 + |V_{ts}|^2} + \\ &3.3 \cdot \alpha \frac{|V_{tb}|^2}{|V_{tb}|^2 + |V_{td}|^2 + |V_{ts}|^2} (|V_{td}|^2 + |V_{ts}|^2). \end{aligned} \quad (13)$$

### 1.3 Cross section interpretation: previous approaches

In this section we describe the previous approach and its shortcomings. With the naming convention described in 1.1 do write our hypothesis as

$$\text{BR}_{(t \rightarrow Wb)} = 1 \quad (14)$$

$$\text{BR}_{(t \rightarrow Wd)} = \text{BR}_{(t \rightarrow Ws)} = 0 \quad (15)$$

$$|V_{td}| = |V_{ts}| = 0 \quad (16)$$

And therefore the only possible floating parameter is  $|V_{tb}|$ . In the standard model case the  $|V_{td}|$  and  $|V_{ts}|$  are much smaller than  $|V_{tb}|$ , but this effect is balanced by the fact that the factors  $\beta$  and  $\gamma$  are much bigger with respect to  $\alpha$ , as their initial state factor in different proton PDFs.

Studies at 7 TeV allow to infer the factors  $\beta$  and  $\gamma$ . Table 3 shows the contribution to 7, 8, and 13 TeV cross sections, roughly extrapolated from the cross sections measured at 7 TeV with  $|V_{td}|^2 = 1$  and  $|V_{ts}|^2 = 1$ . Note that we take as  $\alpha$  the value of the cross section calculated by Kidonakis in [4], where in fact the computation is performed in the case of  $|V_{td}|^2 = 1$ . The value at 8 and 13 TeV has been obtained with a simple scaling of the cross section, using the same scaling as  $\alpha$ . We also define a parameter  $\lambda$ , defined as  $(\beta|V_{td}|^2 + \gamma|V_{ts}|^2) / (|V_{td}|^2 + |V_{ts}|^2)$ . We will often use in the calculations in this chapter  $\lambda$  as an approximation for  $(\beta|V_{td}|^2 + \gamma|V_{ts}|^2)$ , considering the ratio  $|V_{td}|^2 / |V_{ts}|^2$  fixed for an “effective” approach. The values for  $\beta, \gamma$ , and  $\lambda$  are also reported in 3.

	$ V_{td} $	$ V_{ts} $	$ V_{td} ^2 +  V_{ts} ^2$
Value inferred from independent measurements	$0.0084 \pm 0.0006$	$0.0429 \pm 0.0026$	-
Squared value	$7.06 \cdot 10^{-5}$	$1.84 \cdot 10^{-3}$	$1.91 \cdot 10^{-3}$
$\sigma$ at 7 TeV (pb)	0.030	0.28	0.31
$\sigma$ at 8 TeV (pb)	0.040	0.37	0.41
$\sigma$ at 13 TeV (pb)	0.10	0.92	1.02
Mult. parameter at 7 TeV (pb):	$\beta = 429$	$\gamma = 151$	$\lambda = 161$
Mult. parameter at 8 TeV (pb):	$\beta = 570$	$\gamma = 201$	$\lambda = 214$
Mult. parameter at 13 TeV (pb):	$\beta = 1441$	$\gamma = 500$	$\lambda = 534$
7 TeV $\alpha$	64.57		
8 TeV $\alpha$	87.20		
13 TeV $\alpha$	216.99		

Table 3: Extrapolated values of the cross section contribution due to  $V_{td}$  and  $V_{ts}$ .

Usually what is done in single-top analyses is to consider the contribution of  $\beta|V_{td}|^2 + \gamma|V_{ts}|^2 = \lambda \cdot \epsilon \ll \alpha|V_{tb}|^2$ . This in practice means:

$$\sigma_{measured} = \alpha|V_{tb}|_{measured}^2, \quad (17)$$

We will discuss here what happens when we take in the dependence of  $BR(t \rightarrow Wb)$  from  $|V_{tb}|$  in SM and in BSM scenarios. The particular case depends essentially on the behaviour of  $\lambda_{bsm}$ , and  $\epsilon_{bsm}$  with respect to their standard model counterparts. Moreover, we consider the effect of a non-zero acceptance of  $t \rightarrow Wb$  in the selection. The acceptance due to the final state  $t \rightarrow W + d, s$  is  $\sim 0.1$  the acceptance for  $t \rightarrow Wb$ . This is mainly due to the spectator  $b$  quark entering the selection rather than to mis-tagging of a  $d, s$  quark, thus for all calculations we henceforth consider this acceptance factor  $\delta$  to be  $\sim 0.1 \times$  (the acceptance of  $t \rightarrow Wb$ ) for all cases where we have to make a calculation. We can also conservatively add an uncertainty of  $\sim 50\%$  on this acceptance.

#### 1.4 Full approach: 2J1T efficiencies decomposition

We start with the 2J1T sample, where we do have 1 jet passing the tight jet selection AND tight b-tag selection, i.e. having efficiency  $\epsilon_{j1}^T$ , and another one with efficiency  $\epsilon_{j2}^J$ . Those jets can come either from the top-quark, the light quark recoil, or the gluon splitting, as reported also in table 1.

Follow the decomposition of the efficiency of events in the 2-jet 1-tag category,  $ST_{(b,b)}^{2J1T}$ ,  $ST_{(b,q)}^{2J1T}$ , and  $ST_{(q,b)}^{2J1T}$ , where  $q=s,d$ .

$$\begin{aligned}
ST_{(b,b)}^{2J1T} \propto & [\epsilon_{b_{top}}^T \cdot \epsilon_{q_{light}}^J \cdot (1 - \epsilon_{b_{2nd}}^{sel}) + \epsilon_{b_{top}}^T \cdot \epsilon_{b_{2nd}}^J \cdot (1 - \epsilon_{q_{light}}^{sel}) + \\
& + \epsilon_{b_{2nd}}^T \cdot \epsilon_{q_{light}}^J \cdot (1 - \epsilon_{b_{top}}^{sel}) + \epsilon_{b_{2nd}}^T \cdot \epsilon_{b_{top}}^J \cdot (1 - \epsilon_{q_{light}}^{sel}) + \\
& + \epsilon_{q_{light}}^T \cdot \epsilon_{b_{2nd}}^J \cdot (1 - \epsilon_{b_{top}}^{sel}) + \epsilon_{q_{light}}^T \cdot \epsilon_{b_{top}}^J \cdot (1 - \epsilon_{b_{2nd}}^{sel})]
\end{aligned} \tag{18}$$

$$\begin{aligned}
ST_{(b,q)}^{2J1T} \propto & [\epsilon_{b_{2nd}}^T \cdot \epsilon_{q_{light}}^J \cdot (1 - \epsilon_{q_{top}}^{sel}) + \epsilon_{b_{2nd}}^T \cdot \epsilon_{q_{top}}^J \cdot (1 - \epsilon_{q_{light}}^{sel}) + \\
& + \epsilon_{q_{top}}^T \cdot \epsilon_{q_{light}}^J \cdot (1 - \epsilon_{b_{2nd}}^{sel}) + \epsilon_{q_{top}}^T \cdot \epsilon_{b_{2nd}}^J \cdot (1 - \epsilon_{q_{light}}^{sel}) + \\
& + \epsilon_{q_{light}}^T \cdot \epsilon_{b_{2nd}}^J \cdot (1 - \epsilon_{q_{top}}^{sel}) + \epsilon_{q_{light}}^T \cdot \epsilon_{q_{top}}^J \cdot (1 - \epsilon_{b_{2nd}}^{sel})]
\end{aligned} \tag{19}$$

$$\begin{aligned}
ST_{(q,b)}^{2J1T} \propto & [\epsilon_{b_{top}}^T \cdot \epsilon_{q_{light}}^J \cdot (1 - \epsilon_{q_{2nd}}^{sel}) + \epsilon_{b_{top}}^T \cdot \epsilon_{q_{2nd}}^J \cdot (1 - \epsilon_{q_{light}}^{sel}) + \\
& + \epsilon_{q_{2nd}}^T \cdot \epsilon_{q_{light}}^J \cdot (1 - \epsilon_{b_{top}}^{sel}) + \epsilon_{q_{2nd}}^T \cdot \epsilon_{b_{top}}^J \cdot (1 - \epsilon_{q_{light}}^{sel}) + \\
& + \epsilon_{q_{light}}^T \cdot \epsilon_{q_{2nd}}^J \cdot (1 - \epsilon_{b_{top}}^{sel}) + \epsilon_{q_{light}}^T \cdot \epsilon_{b_{top}}^J \cdot (1 - \epsilon_{q_{2nd}}^{sel})]
\end{aligned} \tag{20}$$

81 Which collecting the b-tag efficiencies becomes:

$$\begin{aligned}
ST_{(b,b)}^{2J1T} \propto & [\epsilon_{b_{top}}^T \cdot (\epsilon_{q_{light}}^J \cdot (1 - \epsilon_{b_{2nd}}^{sel}) + \epsilon_{b_{2nd}}^J \cdot (1 - \epsilon_{q_{light}}^{sel})) + \\
& + \epsilon_{b_{2nd}}^T \cdot (\epsilon_{q_{light}}^J \cdot (1 - \epsilon_{b_{top}}^{sel}) + \epsilon_{b_{top}}^J \cdot (1 - \epsilon_{q_{light}}^{sel})) + \\
& + \epsilon_{q_{light}}^T \cdot (\epsilon_{b_{2nd}}^J \cdot (1 - \epsilon_{b_{top}}^{sel}) + \epsilon_{b_{top}}^J \cdot (1 - \epsilon_{b_{2nd}}^{sel}))]
\end{aligned} \tag{21}$$

$$\begin{aligned}
ST_{(b,q)}^{2J1T} \propto & [\epsilon_{b_{2nd}}^T \cdot (\epsilon_{q_{light}}^J \cdot (1 - \epsilon_{q_{top}}^{sel}) + \epsilon_{q_{top}}^J \cdot (1 - \epsilon_{q_{light}}^{sel})) + \\
& + \epsilon_{q_{top}}^T \cdot (\epsilon_{q_{light}}^J \cdot (1 - \epsilon_{b_{2nd}}^{sel}) + \epsilon_{b_{2nd}}^J \cdot (1 - \epsilon_{q_{light}}^{sel})) + \\
& + \epsilon_{q_{light}}^T \cdot (\epsilon_{b_{2nd}}^J \cdot (1 - \epsilon_{q_{top}}^{sel}) + \epsilon_{q_{top}}^J \cdot (1 - \epsilon_{b_{2nd}}^{sel}))]
\end{aligned} \tag{22}$$

$$\begin{aligned}
ST_{(q,b)}^{2J1T} \propto & [\epsilon_{b_{top}}^T \cdot (\epsilon_{q_{light}}^J \cdot (1 - \epsilon_{q_{2nd}}^{sel}) + \epsilon_{q_{2nd}}^J \cdot (1 - \epsilon_{q_{light}}^{sel})) + \\
& + \epsilon_{q_{2nd}}^T \cdot (\epsilon_{q_{light}}^J \cdot (1 - \epsilon_{b_{top}}^{sel}) + \epsilon_{b_{top}}^J \cdot (1 - \epsilon_{q_{light}}^{sel})) + \\
& + \epsilon_{q_{light}}^T \cdot (\epsilon_{q_{2nd}}^J \cdot (1 - \epsilon_{b_{top}}^{sel}) + \epsilon_{b_{top}}^J \cdot (1 - \epsilon_{q_{2nd}}^{sel}))]
\end{aligned} \tag{23}$$

82 Our goal is to focus on the most discriminating terms between  $ST_{(b,b)}^{2J1T}$  vs  $ST_{(q,b)}^{2J1T}$  and  $ST_{(b,q)}^{2J1T}$ , so  
83 we can simplify by considering only the leading terms in those. A few considerations can be  
84 made:

- 85 •  $\epsilon_q^T \ll \epsilon_b^T$ , with any q that is not a b quark. Those terms are suppressed by 2-3 orders  
86 of magnitude.
- 87 •  $\epsilon_b^J < \epsilon_q^T$ , in general by a factor that is dictated by the inefficiency of the b-tag dis-  
88 criminator if applied on b-jets. However, if we additionally require that the  $\epsilon^J$  has  
89 to include a cut on the pseudorapidity of the jet, e.g.  $|\eta| > 2.5$ , we can reject jets  
90 coming from the top-quark.



91 The resulting equations, removing the terms indicated above in the rare processes, will be:

$$ST_{(b,b)}^{2J1T} \propto [\epsilon_{b_{\text{top}}}^T \cdot (\epsilon_{q_{\text{light}}}^J \cdot (1 - \epsilon_{b_{2\text{nd}}}^{\text{sel}}) + \epsilon_{b_{2\text{nd}}}^J \cdot (1 - \epsilon_{q_{\text{light}}}^{\text{sel}})) + \\ + \epsilon_{b_{2\text{nd}}}^T \cdot (\epsilon_{q_{\text{light}}}^J \cdot (1 - \epsilon_{b_{\text{top}}}^{\text{sel}}) + \epsilon_{b_{\text{top}}}^J \cdot (1 - \epsilon_{q_{\text{light}}}^{\text{sel}})) + \\ + \epsilon_{q_{\text{light}}}^T \cdot (\epsilon_{b_{2\text{nd}}}^J \cdot (1 - \epsilon_{b_{\text{top}}}^{\text{sel}}) + \epsilon_{b_{\text{top}}}^J \cdot (1 - \epsilon_{b_{2\text{nd}}}^{\text{sel}}))] \quad (24)$$

$$ST_{(b,q)}^{2J1T} \propto [\epsilon_{b_{2\text{nd}}}^T \cdot (\epsilon_{q_{\text{light}}}^J \cdot (1 - \epsilon_{q_{\text{top}}}^{\text{sel}}) + \epsilon_{q_{\text{top}}}^J \cdot (1 - \epsilon_{q_{\text{light}}}^{\text{sel}}))] \quad (25)$$

$$ST_{(q,b)}^{2J1T} \propto [\epsilon_{b_{\text{top}}}^T \cdot (\epsilon_{q_{\text{light}}}^J \cdot (1 - \epsilon_{q_{2\text{nd}}}^{\text{sel}}) + \epsilon_{q_{2\text{nd}}}^J \cdot (1 - \epsilon_{q_{\text{light}}}^{\text{sel}}))] \quad (26)$$

92 This means that the leading term for  $ST_{(b,q)}^{2J1T}$  is the one where the secondary quark is not selected,  
93 and the  $b_{2\text{nd}}$  is tagged, while the leading term for the  $ST_{(q,b)}^{2J1T}$  will share the same final state hard  
94 jets as the  $ST_{(b,b)}^{2J1T}$ , except for the third jet which will not be b-tagged and the top pt spectrum  
95 which will stem from the gluon splitting to dd and ss pairs rather than bb pairs.

96 For  $ST_{(b,q)}^{2J1T}$  the dominating contribution is the one where either the jet from the top-quark  
97 doesn't pass the selection, or the  $q_{\text{light}}$  doesn't pass the selection. To gain discriminating power  
98 with respect to the  $ST_{(b,b)}^{2J1T}$  topologies, one can thus involve variables used for the correct top-  
99 quark assignment. To gain discriminating power with respect to the SM backgrounds, one can  
100 also use the variables of  $q_{\text{light}}$  instead.

101 For  $ST_{(q,b)}^{2J1T}$  the dominating contribution is the one where the leading b-jet is the correct one for  
102 the choice of the top-quark, as for  $ST_{(b,b)}^{2J1T}$ , but the jet from gluon splitting is not b-tagged. To  
103 gain discriminating power with respect to the  $ST_{(b,b)}^{2J1T}$  topologies, one can thus involve variables  
104 of the top-quark kinematics and tighten the veto towards secondary b-quarks.

105 Those considerations drive us to defining different sub-regions in the 2-jets–1-tag region: be-  
106 sides the selection on QCD we aim at defining a region where we enhance  $ST_{(b,q)}^{2J1T}$  and  $ST_{(q,b)}^{2J1T}$   
107 and focus on the difference in topology with respect to  $ST_{(b,b)}^{2J1T}$ .

108 Similar considerations can lead us to the definition of the efficiencies in the 3-jets–1-tag region  
109 and 3-jets–2-tags region.

110 For the 3j1t we have:

$$ST_{(b,b)}^{3J1T} \propto [\epsilon_{b_{\text{top}}}^T \cdot (\epsilon_{q_{\text{light}}}^J \cdot \epsilon_{b_{2\text{nd}}}^J + \epsilon_{b_{2\text{nd}}}^J \cdot \epsilon_{q_{\text{light}}}^J) + \\ + \epsilon_{b_{2\text{nd}}}^T \cdot (\epsilon_{q_{\text{light}}}^J \cdot \epsilon_{b_{\text{top}}}^J + \epsilon_{b_{\text{top}}}^J \cdot \epsilon_{q_{\text{light}}}^J) + \\ + \epsilon_{q_{\text{light}}}^T \cdot (\epsilon_{b_{2\text{nd}}}^J \cdot \epsilon_{b_{\text{top}}}^J + \epsilon_{b_{\text{top}}}^J \cdot \epsilon_{b_{2\text{nd}}}^J)] \quad (27)$$

$$ST_{(b,q)}^{3J1T} \propto \epsilon_{b_{2\text{nd}}}^T \cdot (\epsilon_{q_{\text{light}}}^J \cdot \epsilon_{q_{\text{top}}}^J + \epsilon_{q_{\text{top}}}^J \cdot \epsilon_{q_{\text{light}}}^J) \quad (28)$$

$$ST_{(q,b)}^{3J1T} \propto \epsilon_{b_{\text{top}}}^T \cdot (\epsilon_{q_{\text{light}}}^J \cdot \epsilon_{q_{2\text{nd}}}^J + \epsilon_{q_{2\text{nd}}}^J \cdot \epsilon_{q_{\text{light}}}^J) \quad (29)$$

Where the same considerations as 2j1t hold, the main difference being the higher energy allowed for the third jet can increase the chance a  $q_{\text{top}}$  of passing the selection.

The 3-jets–2-tags region instead is the most pure in terms of events with  $V_{\text{tb}}$  in production and decay.

## 2 Data and simulated samples

Collision data at  $\sqrt{s}=13$  TeV, corresponding to an integrated luminosity of  $35.9 \text{ fb}^{-1}$  and collected with triggers requiring either one muon or electron in the final state, are used for this measurement.

Dataset	Run range	Integrated luminosity ( $\text{fb}^{-1}$ )
/SingleMuon(SingleElectron)/Run2016B-03Feb2017_ver2-v2/MINIAOD	272007-275376	5.8
/SingleMuon(SingleElectron)/Run2016C-03Feb2017-v1/MINIAOD	275657-276283	2.6
/SingleMuon(SingleElectron)/Run2016D-03Feb2017-v1/MINIAOD	276315-276811	4.2
/SingleMuon(SingleElectron)/Run2016E-03Feb2017-v1/MINIAOD	276831-277420	4.0
/SingleMuon(SingleElectron)/Run2016F-03Feb2017-v1/MINIAOD	277772-278808	3.1
/SingleMuon(SingleElectron)/Run2016G-03Feb2017-v1/MINIAOD	278820-280385	7.5
/SingleMuon(SingleElectron)/Run2016H-03Feb2017_ver2-v1/MINIAOD	} 280919-284044	} 8.6
/SingleMuon(SingleElectron)/Run2016H-03Feb2017_ver3-v1/MINIAOD		

Table 4: Certified (golded) data samples [5].

Monte Carlo (MC) event generators are used to simulate signal and background samples. Single-top quark  $t$ -channel events are generated at NLO with POWHEG 2.0 in the four-flavour scheme (4FS), and the top quark decays are simulated with MADSPIN [6].

Single-top quark  $t$ -channel events with one  $V_{\text{td}}/V_{\text{ts}}$  vertex in production are generated using the next-to-leading order (NLO) generator POWHEG 2.0 in the five-flavor scheme (5FS) and the top quark decays are simulated with MADSPIN [6]. The  $t\bar{t}$  background process is also generated with POWHEG 2.0. The associated top quark and W boson production is simulated with POWHEG in the 5FS. The value of the top quark mass used in the simulated samples is  $m_t = 172.5 \text{ GeV}$ . To simulate the parton shower, hadronization, and the underlying event, PYTHIA 8.180 [7] with tune CUETP8M1 [8] is used, for all single-top productions and Drell-Yan samples, while PYTHIA 8.180 [7] with tune CUETP8M2T4 [9] is used for the  $t\bar{t}$  samples. Simulated event samples with W and Z bosons in association with jets are generated using MG5\_aMC@NLO and the FxFx merging scheme [10], where up to two additional partons are generated at the matrix-element level. The quantum chromodynamics (QCD) multijet events, generated with PYTHIA 8.180, are used to validate the estimation of this background with a technique based on control samples in data.

On top of these, samples with different values of theoretical scale parameters, (PSQ<sup>2</sup>, hdamp, ISR&FSR, Underlying events) have been used to considered possible deviations in the shape of kinematical distributions for the single-top  $t$ -ch. and the main background  $t\bar{t}$ . These are listed in Table 6.

The default parametrization of the PDF used in all simulations is NNPDF30\_nlo\_as.0118 [11]. All generated events undergo a full simulation of the detector response according to the implementation of the CMS detector within GEANT4 [12]. Additional proton-proton interactions within the same or nearby bunch crossing (pileup) are included in the simulation with the same distribution as observed in data.

Dataset	Events	Cross Section $\times$ BR (pb)
/ST.t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1*	67240808	136.02 (NLO) [13]
/ST.t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1*	38811017	80.95 (NLO) [13]
/ST.tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1*	6952830	35.6 (NNLL) [4]
/ST.tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1*	6933094	35.6 (NNLL) [4]
/ST.s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1*	1000000	10.32 (NLO) [13]
/TT_TuneCUETP8M2T4_13TeV-powheg-pythia8*	77081156	831.8 (NLO) [13]
/WToLNu_0J_13TeV-amcatnloFXFX-pythia8*	49142195	49670.0 (NLO)
/WToLNu_1J_13TeV-amcatnloFXFX-pythia8*	92024405	8264.0 (NLO)
/WToLNu_1J_13TeV-amcatnloFXFX-pythia8*	46924291	8264.0 (NLO)
/WToLNu_2J_13TeV-amcatnloFXFX-pythia8*	53135013	2628.0 (NLO)
/WToLNu_2J_13TeV-amcatnloFXFX-pythia8*	102093848	2628.0 (NLO)
/DYjetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8*	122055388	5765.4 (NLO)
/WWTo1L1Nu2Q_13TeV-amcatnloFXFX_madspin-pythia8*	5176114	45.85 (NLO)
/WWTo2L2Nu_13TeV-powheg*	1999000	12.178 (NLO)
/WZTo1L1Nu2Q_13TeV-amcatnloFXFX_madspin-pythia8*	24221923	10.71 (NLO)
/WZTo2L2Q_13TeV-amcatnloFXFX_madspin-pythia8*	26517272	5.595 (NLO)
/ZZTo2L2Q_13TeV-amcatnloFXFX_madspin-pythia8*	15345572	3.22 (NLO)
/ST.t-channel_top_4f_Vtd_Vts_prod_leptonicDecays_13TeV-powhegV2-madspin-pythia8*	11277196	0.231 (s) – 0.136 (d)
/ST.t-channel_antitop_4f_Vts_Vtd_prod_leptonicDecays_13TeV-powhegV2-madspin-pythia8*	7609410	0.025 (s) – 0.008 (d)
/ST.t-channel_top_4f_Vtd_Vts_decay_leptDecays_13TeV-powhegV2-madspin-pythia8*	5863856	0.2720
/ST.t-channel_antitop_4f_Vts_Vtd_decay_leptDecays_13TeV-powhegV2-madspin-pythia8*	3847666	0.1619

\*/RunIISummer16MiniAODv2-PUMoriond17.80X.mcRun2\_asymptotic.2016.TrancheIV.v6\*/MINIAODSIM

Table 5: Nominal simulation data sets with prompt leptons for the analysis of 2016 data. If not stated otherwise, the cross section obtained from the generator is used.

Dataset	Events
PSQ <sup>2</sup>	
/ST.t-channel_antitop_4f_scaleup_inclusiveDecays_13TeV-powhegV2-madspin-pythia8*	3970546
/ST.t-channel_antitop_4f_scaledown_inclusiveDecays_13TeV-powhegV2-madspin-pythia8	38361721
/ST.t-channel_top_4f_scaleup_inclusiveDecays_13TeV-powhegV2-madspin-pythia8*	69067916
/ST.t-channel_top_4f_scaledown_inclusiveDecays_13TeV-powhegV2-madspin-pythia8*	5709148
Underlying events	
/TT_TuneCUETP8M2T4down_13TeV-powheg-pythia8*	28354188
/TT_TuneCUETP8M2T4up_13TeV-powheg-pythia8*	29310620
hdamp	
/ST.t-channel_top_4f_hdampdown_inclusiveDecays_13TeV-powhegV2-madspin-pythia8*	6000000
/ST.t-channel_top_4f_hdampup_inclusiveDecays_13TeV-powhegV2-madspin-pythia8*	6000000
/ST.t-channel_antitop_4f_hdampdown_inclusiveDecays_13TeV-powhegV2-madspin-pythia8*	3999346
/ST.t-channel_antitop_4f_hdampup_inclusiveDecays_13TeV-powhegV2-madspin-pythia8*	4000000
/TT_hdampDOWN_TuneCUETP8M2T4_13TeV-powheg-pythia8*	29117820
/TT_hdampUP_TuneCUETP8M2T4_13TeV-powheg-pythia8*	29689380
Initial and final state radiation	
/TT_TuneCUETP8M2T4_13TeV-powheg-isrup-pythia8*	97436211
/TT_TuneCUETP8M2T4_13TeV-powheg-isrdowndown-pythia8*	90763410
/TT_TuneCUETP8M2T4_13TeV-powheg-fsrup-pythia8*	93386680
/TT_TuneCUETP8M2T4_13TeV-powheg-fsrdowndown-pythia8*	96684680

\*/RunIISummer16MiniAODv2-PUMoriond17.80X.mcRun2\_asymptotic.2016.TrancheIV.v6\*/MINIAODSIM

Table 6: Systematically varied  $t$ -ch. and  $t\bar{t}$  samples. The cross section obtained from the generator is used.

### 3 Physics object selection

The final-state topology in the  $t$ -channel is characterized by the presence of exactly one isolated charged lepton and a  $b$  jet from the top quark decay, as well as a light-flavoured jet produced in the forward region. Potential gluon splitting in the initial state leads to a second  $b$  quark (Fig. 1) that recoils against the top quark and thus has generally a softer  $p_T$  spectrum and a broader  $|\eta|$  distribution compared to the  $b$  quark produced in the top-quark decay. The acceptance of events with two  $b$  jets reconstructed in the final state is therefore expected to be small. In fact we can anticipate that using the selection described in this section, the number of signal events with two  $b$  jets reconstructed in the detector is smaller than the number of events with just one  $b$  jet. In the following, we describe the object definitions and selection criteria.

#### 3.1 Object definition

In this section the basic analysis objects are defined on which the event selection and the kinematic reconstruction are based on. The candidates have been reconstructed with the particle-flow (PF) algorithm [14] which combines various subdetector information to perform a global particle type identification and momentum measurement.

##### 3.1.1 Primary vertex

The reconstruction of tracks and primary vertices is detailed in Ref. [15]. If more than one primary vertex is identified, the one with largest sum of the squared transverse momenta of associated tracks is taken. Additionally, a “good” primary vertex candidate has to be fitted with  $\text{ndof} > 4$  and reconstructed in the interaction region defined by  $|d_z| < 24$  cm and  $|d_{xy}| < 2$  cm with respect to the nominal interaction point.

##### 3.1.2 Tight Muons

Events with exactly one muon within a pseudorapidity range of  $|\eta| < 2.4$  and with a transverse momentum  $p_T > 26$  GeV are selected. The baseline muon selection contains “global muons” and has to meet additional muon quality requirements referred to as “tight muon ID” to select the most genuine muons from the PF muon collection.

More specifically, tight muons must have  $\chi^2/\text{ndof} < 10$  and at least one valid hit in the muon chambers, in order to suppress hadronic punch-through and muons from decays in flight. To guarantee a good  $p_T$  measurement, muon candidates are required to have more than 5 valid hits in the silicon tracker, out of which at least one in the pixel detector so as to further suppress muons from decays in flight. At least two muon chamber segments must match the global muon object to suppress punch-through and accidental track-to-segment matches. Furthermore, the transverse and longitudinal impact parameters must be smaller than  $|d_{xy}| = 0.2$  cm and  $|d_z| = 0.5$  cm with respect to the primary vertex in order to suppress muons from cosmic rays, muons from decays of long-lived hadrons (i.e.  $B, C$  hadrons), or from pileup interactions.

We define the “particle flow (relative) isolation” ( $I_{\text{rel}}$ ) with the so-called “DeltaBeta” correction as

$$I_{\text{rel}} = \frac{I^{\text{ch.h}} + \max((I^\gamma + I^{\text{n.h}} - I^{\text{PU}}), 0)}{p_T}, \quad (30)$$

where  $I^{\text{ch.h}}$ ,  $I^\gamma$ , and  $I^{\text{n.h}}$  are the sum of the transverse energies deposited by stable particles like charged hadrons, photons and neutral hadrons respectively, in a cone of size  $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} = 0.4$  around the muon direction;  $I^{\text{PU}} \equiv \Delta\beta \times \sum p_T^{\text{PU}} \equiv 0.5 \times \sum p_T^{\text{PU}}$  is the sum of transverse momenta of tracks associated to non-leading, i.e. pileup, vertices, used to estimate

the contribution of neutral particles from pileup events by applying a multiplicative factor of 0.5 that takes into account the neutral-to-charged particles ratio expected from isospin invariance. Therefore, the  $\Delta\beta$  factor maps the expected neutral contribution in the isolation cone from the observed PU charged contribution. Tight muons are accepted by the requirement, which is expected to select signal events with  $\sim 70\%$  efficiency. The data-to-MC correction factors for muon are obtained using a “Tag and Probe” method and are applied to simulated events to account for differences in efficiencies between data and simulation. The correction factors for identification are provided by the muon object group while for isolation a similar method is used to extract the numbers corresponding to  $I_{\text{rel}} < 0.06$  working point.

### 3.1.3 Tight Electrons

Events containing exactly one tight electron candidate are selected by requiring the presence of a “GsfElectron”, i.e. an electron track fitted with the Gaussian-Sum-Filter that is matched to an active ECAL cell by the PF algorithm, with  $E_T > 35$  GeV,  $|\eta| < 2.1$  while excluding the ECAL barrel-endcap transition region of  $1.44 < |\eta| < 1.57$ . Additional quality criteria are applied based on nine variables through an optimized cut-based approach. Due to the presence of strong correlation between the  $\Delta\phi$  and  $|1/E - 1/p|$  (with E the ECAL supercluster energy and p the track momentum at the point of closest approach to the beam spot) variables, optimization has been achieved only for one of them, albeit making a reasonable cut for the other. The analyses make use of the “tight” working point yielding an efficiency of 80–90% for electrons above 45 GeV.

### 3.1.4 Veto of additional muons or electrons

Events containing additional muons or electrons besides either one tight muon or electron candidate are vetoed. This reduces the contamination by Z+jets and dileptonic  $t\bar{t}$  background events. The selection requirements for these additional leptons are loosened with respect to the tight criteria.

Loose muons are required to fall within  $|\eta| < 2.4$  while having  $p_T > 10$  GeV and being reconstructed as “global” or “tracker” muon together with a relative isolation of  $I_{\text{rel}} < 0.2$ . Loose electron candidate have to pass a cut-based quality selection similar to the tight electron selection defined above whereas the kinematic requirements are  $E_T > 15$  GeV,  $|\eta| < 2.5$ . The selection chain for electron identification used for such electrons is tuned to have a higher efficiency, thus resulting in a tighter cut on the veto. The present analysis then makes use of the cut-based “electron veto” working point yielding an efficiency of 98% at a plateau which is reached around  $E_T > 45$  GeV.

### 3.1.5 Jets

Jets are clustered from PF candidate using the anti- $k_t$  clustering algorithm [16] with a cone size of  $R = 0.4$ . The influence of pileup is mitigated using the charged hadron subtraction technique (CHS) [17] in which PF candidates not associated to the primary vertex are ignored in the clustering. A potential overlap of jets with a selected tight muon or electron is omitted by ignoring jets which are closer than  $\Delta R = 0.4$  with respect to tight muon or electron candidate.

The resulting jets have the standard multi-level jet energy corrections (JEC) applied (L1FastJet, L2, L3). Technically, we apply the Summer16v4 corrections determined from simulation on both data and simulation. For data we further apply residual corrections derived from data themselves. Furthermore the jet energy resolution (JER) is corrected in simulation to match the one observed in data by smearing the jet energy using dedicated scale factors. The analysis



considers jets within  $|\eta| < 4.7$  whose calibrated transverse energy is greater than 40 GeV and which pass a set of loose quality cuts which are detailed in the following. The cuts are different depending on the pseudorapidity of the jets as different PF information are available.

- For PF jets with  $|\eta| < 2.7$ , neutral hadronic and neutral electro-magnetic energy fractions must be smaller than 99% and more than one constituent must be present. Additionally, for PF jets with  $|\eta| < 2.4$  the number of charged hadrons and the respective energy fraction must be greater than 0, and the charged electro-magnetic energy fraction must be smaller than 99%.
- For PF jets with  $2.7 < |\eta| < 3.0$  the definition is different: the neutral electro-magnetic energy fraction is required to be greater than 0.01 and the neutral hadron fraction must be smaller than 98%, while the number of neutral particles must be greater than 2.
- Finally, for PF jets with  $|\eta| > 3.0$  neutral electro-magnetic energy fraction must be  $< 90\%$  and the number of neutral particles greater than 10.

### 3.1.6 b Tagging

Several b-tagging, i.e. identification of jets originating from b quarks, algorithms are available in CMS. Some exploit the long B-hadrons lifetime, others their semi-leptonic decay modes and others use kinematic variables related to the high B-meson mass and hard b-quark fragmentation function. Details are provided elsewhere [18]. More specifically, b-tagging algorithms are based on displaced tracks (track counting taggers, jet probability tagger), the presence of “secondary” vertices (secondary vertex taggers) or soft leptons (soft lepton taggers) or on a combination of these (combined secondary vertex taggers). The combined secondary vertex algorithm uses track-based lifetime information together with secondary vertices inside the jet to provide a MVA discriminator for b jet identification (Combined MVA v2) [19]. For this study we use the CMVA v2 algorithm at the “tight” working point corresponding to a threshold set to 0.9432 and a 0.1% gds mistag efficiency—estimated based on simulated jets with  $p_T > 60$  GeV—as recommended by the b-tagging Physics Object Group (POG) [20].

In order to correct the whole b-tagging CMVA v2 discriminant distribution in MC to match that in data a reweighting method is applied. The goal of this method is to predict correct b-tagging discriminant distributions as well as event yields in data by only changing the weight of selected MC events, i.e., MC events that did not pass the selection do not need to be added back. It is important to point out that with this method there is no migration of events from one b-tag multiplicity bin to another, as each event gets a weight based on jet scale factors. In general, the jet scale factors depend on  $p_T$ ,  $\eta$  and the b-tagging discriminant D.

The b-tag event weight  $\omega$  is obtained by

$$\omega_{\text{event}} = \prod_i^{N_{\text{jets}}} SF(D_i, p_{Ti}, \eta_i),$$

which takes into account all selected jets.

As the name already suggests, the event weights extracted with this method are supposed to change the shape of the b-tagging discriminant.

For the b-tag efficiency correction we applied the B-tagging discriminant re-shape using event weights, as reported in [21]. We report the procedure in details as described in [22], for reference. Continuous scale factors are calculated as a function of the b-tagging discriminant D in different jet  $p_T$  and  $\eta$  bins in independent regions:

- Heavy flavor (HF):  $t\bar{t}$  enriched region, dilepton, exactly two jets
- Light flavor (LF): Z+jets enriched region, dilepton, exactly two jets

In the HF (LF) region, one jet is required to be (anti-) tagged. Per lepton channel ( $ee, e\mu, \mu\mu$ ), the expected yield in MC is normalized to data in order to only correct the shape. Then, per  $(p_T, \eta)$  bin, MC is normalized again to data and histograms of the probe jet b-tagging discriminant  $D$  are created. Contaminations from udsb (b) and c jets, as determined in MC, are subtracted from data, which is then divided by the MC distribution of b (udsg) probe jets,

$$\text{HF : } SF(D, p_T, \eta) = \frac{\text{DATA} - \text{MC}_{uds+g+c}}{\text{MC}_b}$$

$$\text{LF : } SF(D, p_T, \eta) = \frac{\text{DATA} - \text{MC}_{b+c}}{\text{MC}_{uds+g}}$$

### 3.1.7 Missing transverse energy

Defined in an analogous way as PF-based jets, PF-based  $\vec{\cancel{E}}_T$  is the opposite of the vectorial sum of the transverse momenta of the identified PF particles

$$\vec{\cancel{E}}_T^{\text{raw}} = |\vec{\cancel{p}}_T^{\text{raw}}| = \left| - \sum_i^{\text{all PF}} \vec{p}_i \right|. \quad (31)$$

The  $\vec{\cancel{E}}_T$  receives corrections (type I) by propagating the vectorial difference between uncalibrated and calibrated jets (only L2 and L3) to the  $\vec{\cancel{E}}_T$ . The correction can be expressed as

$$\vec{\cancel{p}}_T^{\text{corr.}} = - \sum_i^{\text{unclustered}} \vec{p}_i - \sum_i^{\text{jets}} \vec{p}_i^{\text{corr.}} \quad (32)$$

$$= \vec{\cancel{p}}_T^{\text{raw}} + \sum_i^{\text{jets}} (\vec{p}_i^{\text{L123}} - \vec{p}_i^{\text{L1}}), \quad (33)$$

which yields an improved  $\vec{\cancel{E}}_T$  energy scale and resolution [23].

## 3.2 High level trigger selection

At high level trigger, in the muon channel events are selected if they pass the `HLT_IsoMu24` OR the `HLT_IsoTkMu24` path where the presence of an isolated online muon candidate with  $p_T > 24$  GeV is required. Efficiencies for this trigger in simulation are corrected using data-to-MC correction factors obtained from a “Tag and Probe” method. In the electron channel events need to pass the trigger path `HLT_Ele32_eta2p1_WPTight_Gsf_vX`, where the online electron is restricted to have  $p_T > 32$  GeV and  $|\eta| < 2.1$ . The respective trigger correction factors were not provided by the E/gamma Physics Object Group and have been hence calculated.

## 3.3 Signal and control regions

The signature of the  $t$ -channel single-top production includes one light quark recoiling against the virtual W boson, one b quark from the top-quark decay, and a second b quark from the initial gluon splitting.

The second b quark has a softer  $p_T$  and a harder  $\eta$  spectrum with respect to the one coming from the top-quark decay. As a result, jets stemming from the hadronization of the latter are less likely to be selected due to the  $p_T$  cut on the jet, and if selected they are less likely to be tagged, due to the intrinsic limit on the acceptance in  $\eta$  of the b-tagging algorithm, which is limited to the tracker acceptance ( $|\eta| < 2.4$ ). For that reason the region with two jets, with one of them being tagged as b jet, provides the largest fraction of signal events. In order to test the modelling of the main background processes it is useful to define further regions, which are enriched in certain background processes. For that purpose we use the notation "nJmT" or the wording "n-jets-m-tags" to refer to a sample that has exactly n reconstructed jets, exactly m of which pass the b tag threshold. Notable samples which are studied and used in this analysis are the 2J1T sample (signal region for the  $|V_{tb}|^2$  processes with the largest signal fraction among all regions), the 3J1T (signal region for the  $|V_{td,s}|^2$  processes), and the 3J2T region (enriched in  $t\bar{t}$ ).

### 3.4 Transverse W boson mass

To further suppress contributions from processes where the muon does not come from a leptonically decaying W boson, a selection based the reconstructed transverse W-boson mass  $m_T$  can be applied. It is defined as

$$m_T = \sqrt{(p_{T,\ell} + p_{T,\nu})^2 - (p_{x,\ell} + p_{x,\nu})^2 - (p_{y,\ell} + p_{y,\nu})^2}, \quad (34)$$

where the transverse momentum of the neutrino is approximated by the missing transverse energy vector,  $\vec{p}_T$ .

In order to analyze the kinematics of singly produced top quarks, the fourvector of the top quarks have to be reconstructed from the decay products. All top-quark decay products are reconstructed in the detector, except for the neutrino which escapes unobserved. While the transverse momentum of the neutrino can be inferred from the missing transverse energy, its longitudinal momentum has to be derived based on extra assumptions. Once the leptonically decaying W boson is reconstructed the selected jets have to be assigned to the final state quarks in the top quark decay chain.

#### 3.4.1 W-boson reconstruction

The first step in the reconstruction of the top quark from its decay products is the reconstruction of the W boson. We assume that the  $x$  and  $y$  components of the missing transverse energy are entirely due to the escaping neutrino, and apply the W-mass constraint in order to extract the unknown  $z$  component ( $p_{z,\nu}$ ):

$$m_W^2 = (E_\ell + \sqrt{\vec{E}_T^2 + p_{z,\nu}^2})^2 - (\vec{p}_{T,\ell} + \vec{p}_T)^2 - (p_{z,\ell} + p_{z,\nu})^2. \quad (35)$$

This equation has in general two solutions:

$$p_{z,\nu} = \frac{\Lambda \cdot p_{z,\ell}}{p_{T,\ell}^2} \pm \sqrt{\frac{\Lambda^2 \cdot p_{z,\ell}^2}{p_{T,\ell}^4} - \frac{E_\ell^2 \cdot \vec{E}_T^2 - \Lambda^2}{p_{T,\ell}^2}}, \quad (36)$$

with



$$\Lambda = \frac{m_W^2}{2} + \vec{p}_{T,\ell} \cdot \vec{p}_T. \quad (37)$$

In the case of two real solutions for  $p_{z,\nu}$  (in 65% of all cases), different choice criteria have been proposed [24, 25]. The solution with the smallest absolute value is chosen in the present analysis. By looking at truth information in simulated events it is found that, in 63.4% of the selected events with real solutions, the smallest  $|p_{z,\nu}|$  solution is closer to the true neutrino  $p_z$  than the other solution.

If the discriminant in Eq. (36) becomes negative, or equivalently  $m_T$  is larger than  $m_W$ , the solutions have an imaginary component. This happens in 35% of the cases, mostly due the finite  $\vec{E}_T$  resolution. Lepton momentum resolution and the finite W intrinsic width give negligible contributions. Several schemes have been used to deal with this situation [24, 25]. In this analysis the imaginary component is eliminated by modifying  $\vec{E}_T$  such to give  $m_T = m_W$ , still respecting the  $m_W$  constraining from Eq. (35). This is obtained by imposing that the discriminator, and thus the square-root term in Eq. (36), are null. This condition gives a quadratic relation between  $p_{x,\nu}$  and  $p_{y,\nu}$ , with two possible solutions, among which the one with minimal distance between  $p_{T,\nu}$  and  $\vec{E}_T$  is chosen.

### 3.4.2 Jet-parton assignment

In the second step the selected jets have to be assigned to the final state quarks from the top-quark decay. In the 2J1T region the procedure is straight forward: the tagged jet is assigned to the b-quark from the top-quark decay, the non-tagged jet is assigned to the light quark. The b-tagged jet matches the true b quark from top-quark decay in simulated events in about 83.9% of the selected signal events, using as matching criterion a distance of  $\Delta R < 0.3$  between the jet and the parton. The fraction of events with wrong assignment are events in which the selected b-tagged jet stems from the second b-quark from the initial gluon splitting and not from the decay of the top quark.

In the 3J2T region the non-tagged jet is again assigned to the light quark. From the two tagged jets the one with the larger value of the b-tag discriminator is assigned to the b quark from the top-quark decay, while the other tagged jet is assigned to the second b quark from the gluon splitting. This choice is correct in 48.4% of all cases, estimated on simulated signal events using the same matching criterion as described above.

## 4 Discriminant variables and Multivariate Analysis

### 4.1 Data driven QCD extraction

The production of multijets through QCD processes has a huge cross section. On the other hand, only a small fraction of these events mimics the lepton+jets final state of the applied event selection and the selection efficiency for QCD multijet events is therefore tiny. The combination of huge cross section and tiny selection efficiency would require the generation of extremely large MC samples for this process in order to produce sufficient events that survive the event selection to guarantee a proper QCD modeling with decent statistics in the signal region. An alternative way of modeling QCD is to define sideband regions in data that are enriched in QCD events and take the distributions of the relevant kinematic variables directly from these data.

The sideband region in the muon channel is defined by inverting the isolation criterion for the muon ( $I_{\text{rel}} > 0.2$ ). Using samples of simulated events from all relevant signal and background processes, the QCD-purity of the anti-isolated sideband region has been estimated to be 93.5%, for muon, and 92.3%, for electron in the 2-jets-1-tag region. The small contributions from non-QCD-processes are subtracted from the data to generate the QCD templates.

The contribution to the sample of selected events from QCD processes is obtained from a fit to the distribution of the transverse W boson mass in the muon channel and to the distribution of the missing transverse momentum in the electron channel, using the QCD templates derived from sideband regions in data as described above. An extended maximum-likelihood fit with miss two parameters is performed to the  $m_T^W$  distribution:

$$F(m_T^W) = N_{\text{QCD}} \cdot Q(m_T^W) + N_{\text{Non-QCD}} \cdot B(m_T^W), \quad (38)$$

where  $Q(m_T^W)$  stands for the QCD template taken from the sideband region in data,  $B(m_T^W)$  is the non-QCD template obtained by summing up the simulated contributions from all other processes with prompt muons in the isolated region according to their predicted cross section. Both templates,  $Q(m_T^W)$  and  $B(m_T^W)$  are taken as pdfs for the fit. The fit parameter  $N_{\text{QCD}}$  denotes the number of QCD events and  $N_{\text{Non-QCD}}$  represents the total number of non-QCD events. The  $N_{\text{QCD}}$  and  $N_{\text{Non-QCD}}$  parameters are allowed to float during the fit to the  $m_T^W$  distribution. The entire range of the  $m_T^W$  distribution is fitted. This fit is performed separately for the 2-jets-1-tag and 3-jets-1-tag regions in the muon channel and in the electron channel. The QCD contribution to the 3-jets-2-tags region is negligible and no dedicated fit is performed. For the 3-jets-1-tag region an extra scale factor is needed in order to have a better QCD yield estimation in the signal region ( $m_T^W > 50 \& |\eta_{\gamma'}| > 2.5$ ) and this is derived from the sideband as:

$$SF = \frac{N_{\text{Data}_{\text{SB}}}^{|\eta_{\gamma'}| > 2.5}}{N_{\text{Data}_{\text{SB}}}}. \quad (39)$$

The resulting QCD multijet contributions to the various regions are summarized in Tables 7.

Table 7: Number of QCD events extracted from the sideband region for both muon and electron channels in the 2-jets-1-tag and 3-jets-1-tag regions.

Channel	Process	2j1t		3j1t	
		Full $m_T^W$	$m_T^W > 50$	Full $m_T^W$	$m_T^W > 50 \&  \eta_{\gamma'}  > 2.5$
$\mu$	QCD	$69664 \pm 284$	$15286 \pm 89$	$15378 \pm 259$	$431 \pm 7$
	non-QCD	$468021 \pm 302$	$317665 \pm 1010$	$417035 \pm 270$	$59951 \pm 326$
$e$	QCD	$25838 \pm 423$	$12005 \pm 64$	$20 \pm 1$	$2 \pm 1$
	non-QCD	$259751 \pm 438$	$182087 \pm 812$	$47138 \pm 65$	$32698 \pm 250$

370

371 The fitted distributions for all relevant regions can be found in Figure 2 for the muon channel  
372 and in Figure 3 for the electron channel.

## 373 4.2 QCD rejection in 2j1t and in 3j1t

374 In the 2-jets-1-tag and in the 3-jets-1-tag to discriminate between QCD and other channels the  
375 variable used is the transverse mass of the W-boson. Firstly, by reverting the isolation (muon)  
376 and veto (electron) cut on the lepton, a region enriched enriched with QCD events is used to  
377 estimate the shape for the QCD directly from the data. Then this shape is taken as input for the

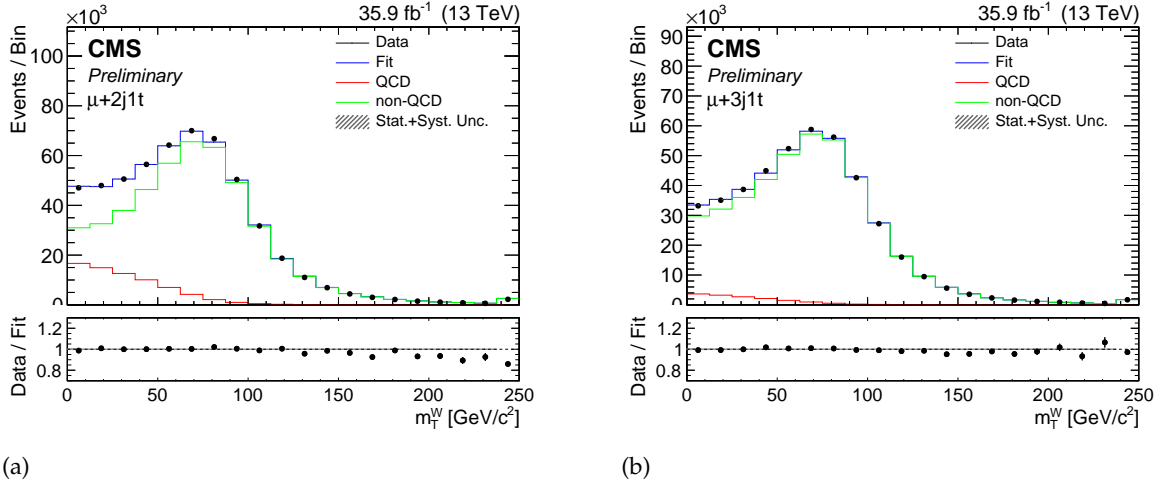


Figure 2: Results of the fit to the transverse mass of the W-boson in the 2-jets-1-tag and in the 3-jets-1-tag for the muon channel.

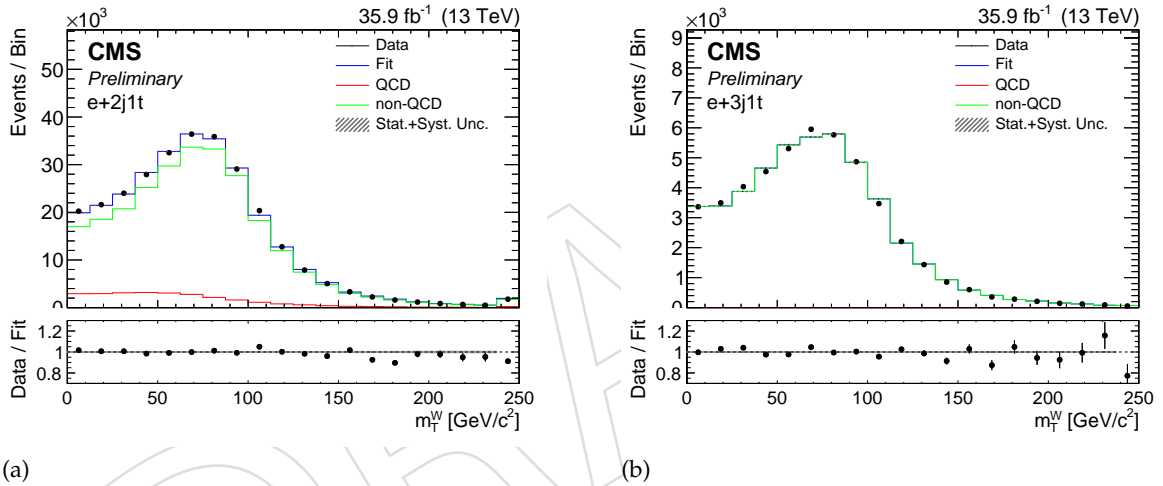


Figure 3: Results of the fit to the transverse mass of the W-boson in the 2-jets-1-tag and in the 3-jets-1-tag for the electron channel.

isolated region and the transverse mass of the W-boson is fitted with the Maximum Likelihood approach to have the right number of events in the region of interest.

Note: in the subsequent plots, all QCD is taken from data both for the muon channel and the electron channel.

In order to deplete the QCD events in the 2-jets-1-tag and 3-jets-1-tag, a  $m_T^W > 50$  GeV/ $c^2$  cut is applied. Figure 4 shows the discriminating power of the  $m_T^W$  cut against QCD events.

It is henceforth referred to the region obtained after the  $m_T^W > 50$  GeV/ $c^2$  cut as **QCD-depleted region**. In the 3-jets-1-tag an extra scale factor taking into account the extrapolation to the  $|\eta_{j'}| > 2.5$  region is added; this is extracted by the control region as the ratio between the yield of the data with and without the  $|\eta_{j'}| > 2.5$  cut applied.

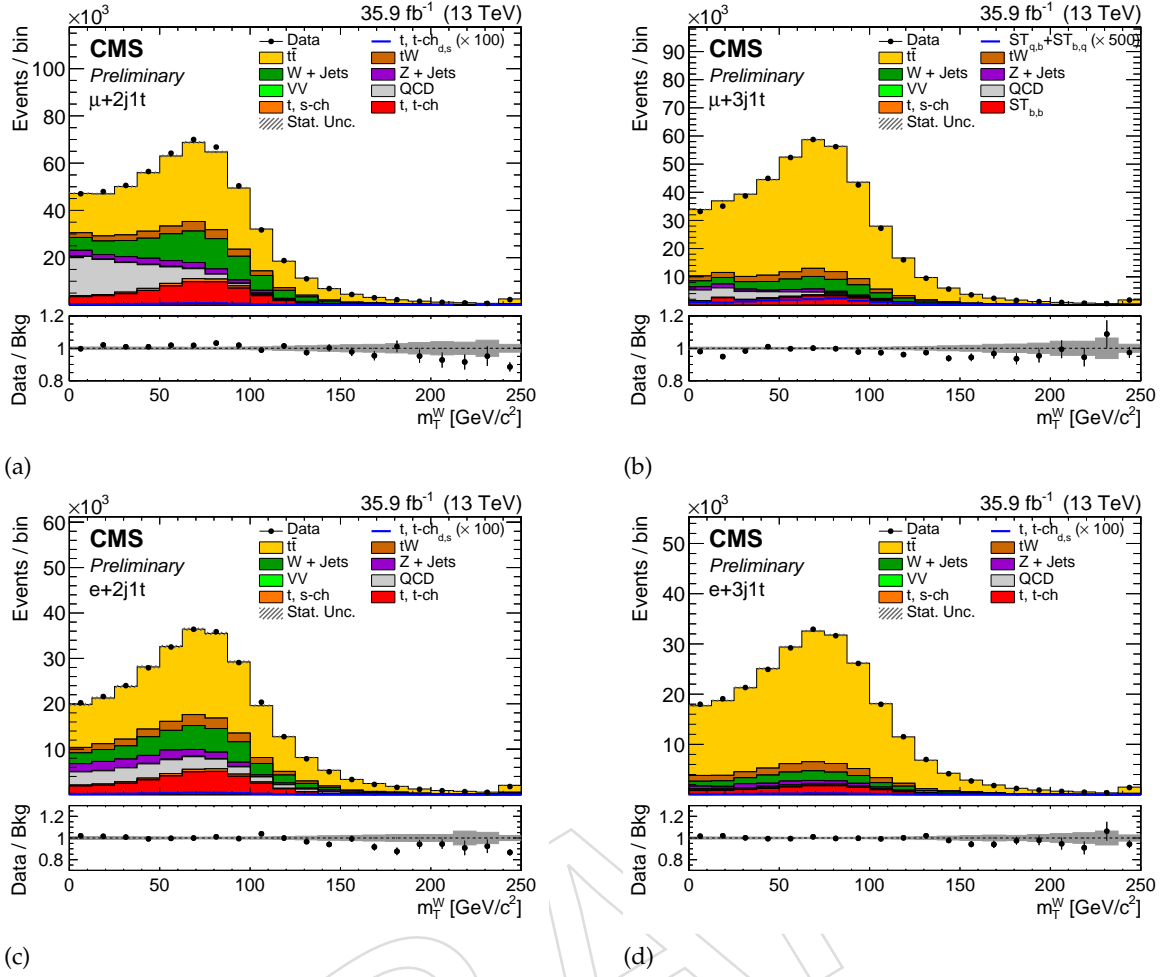


Figure 4: Transverse mass of the W-boson in the 2-jets-1-tag and in the 3-jets-1-tag after the fit to the QCD data driven sample.

### 4.3 Variables in 2j1t

The 2-jets-1-tag is enriched in single-top  $t$ -ch. events and it is used as a control region for the estimation of this class of events. A  $m_T^W > 50 \text{ GeV}/c^2$  is necessary to deplete the QCD events. The main source of background comes from  $t\bar{t}$  and  $W$ + jets events that have the same topology in the final state. In order to discriminate against these two backgrounds a multivariate analysis has been performed.

Figure 5 shows the most discriminating variables for the single-top  $t$ -ch. and two most important backgrounds in this region:  $t\bar{t}$  and  $W$ + jets. These variables are used in the discriminator with other variables in order to maximise the separation between signal and background.

The BDT is trained on exclusively dedicated samples that are not used in any other stage of the analysis. Due to limited amount of MC statistic for the  $W$ + jets sample in the electron sample the training has been done only in the muon channel and the same discriminator is used for both the samples in the analysis.

Figure 16 shows the shape of the variables used as input to the BDT trained in the 2-jets-1-tag region.

Figure 17 shows the linear correlation coefficients between the input variables.

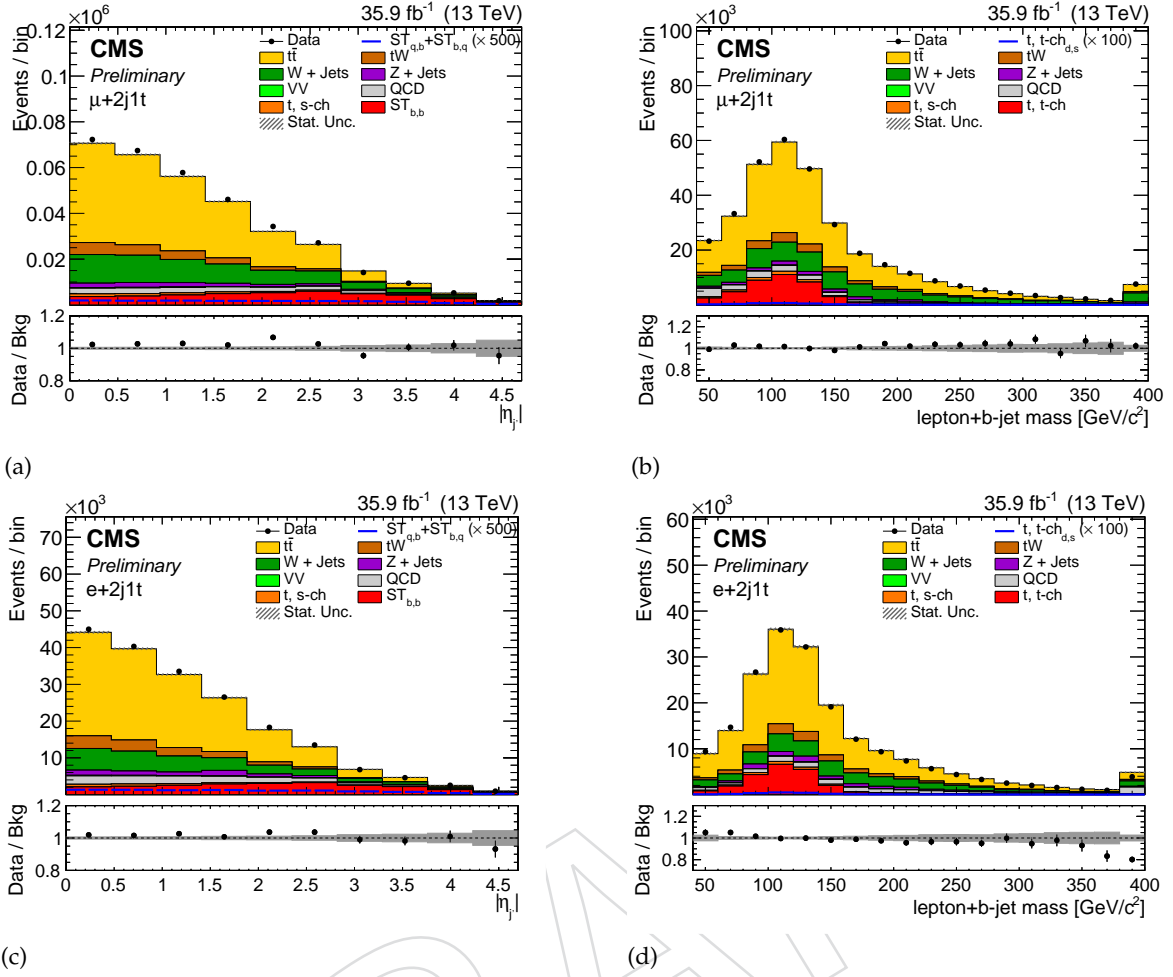


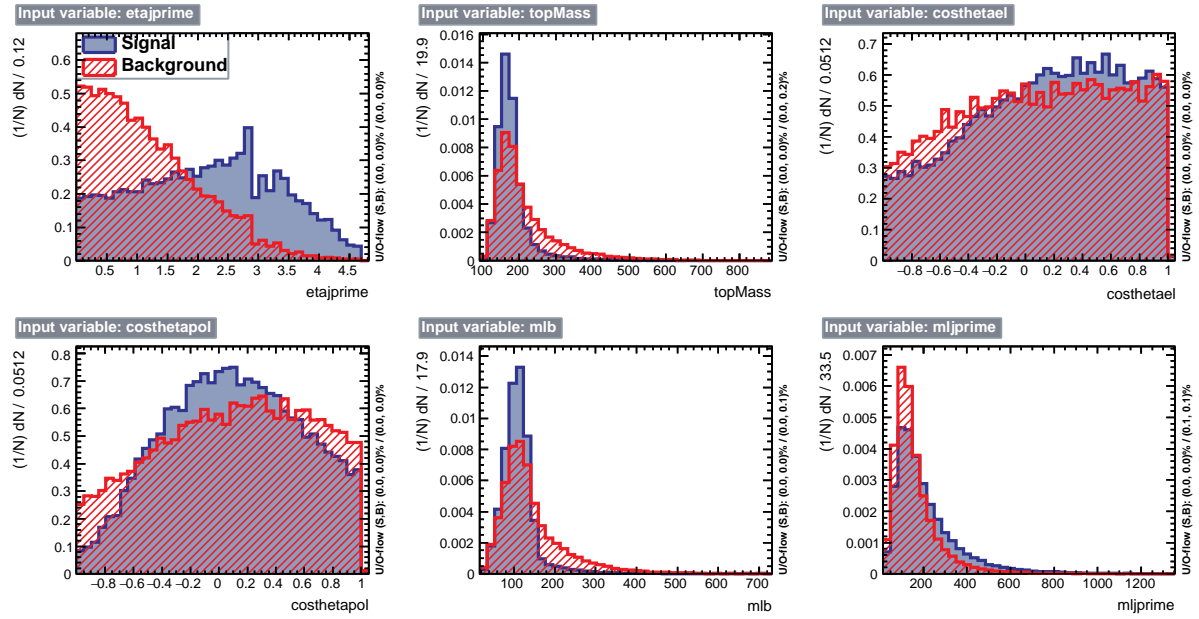
Figure 5: Distributions of the two most discriminating variables in the 2-jets-1-tag QCD-depleted region, which are used in the MVA analysis: on the top for the muon channel and on the bottom for the electron channel.

The table 8 shows the results of the BDT discriminator in the QCD-depleted region in which single-top  $t$ -ch. is considered as signal and  $t\bar{t}$  and  $W$ + jets are considered as background.

Table 8: Input variables for the BDTs trained for the single-top-quark  $t$ -ch. versus  $t\bar{t}$  and  $W$ + jets ranked according to their importance in the muon channel in the 2-jets-1-tag QCD-depleted region.

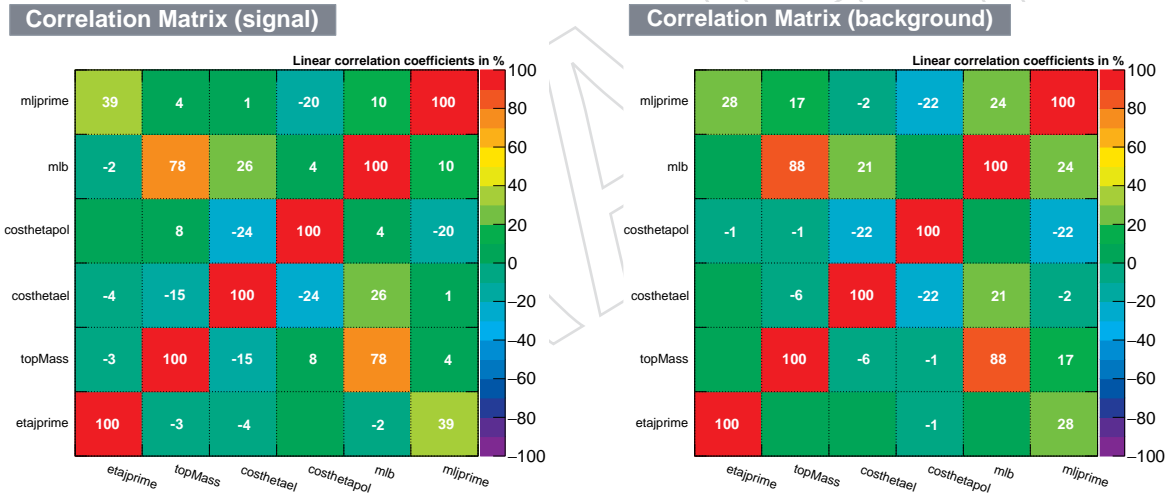
Rank	Variable	Rel. imp. [%]	Correlation to BDT	
			Signal	Background
1	$ \eta $ light jet	21.40	+0.82	+0.53
2	lepton + b-jet mass	19.01	-0.29	-0.57
3	lepton + light jet mass	18.31	+0.37	+0.07
4	$\cos \theta_{pol}^*$	15.06	+0.05	+0.07
5	top-quark mass	15.02	-0.33	-0.54
6	$\cos \theta_{hel}^*$	11.20	+0.07	+0.07

Figure 8 shows the overtraining test result and the ROC curve for the BDT obtained.



(a)

Figure 6: the plots show the shape of the variables used as input of the BDT training in the 2-jets-1-tag region.

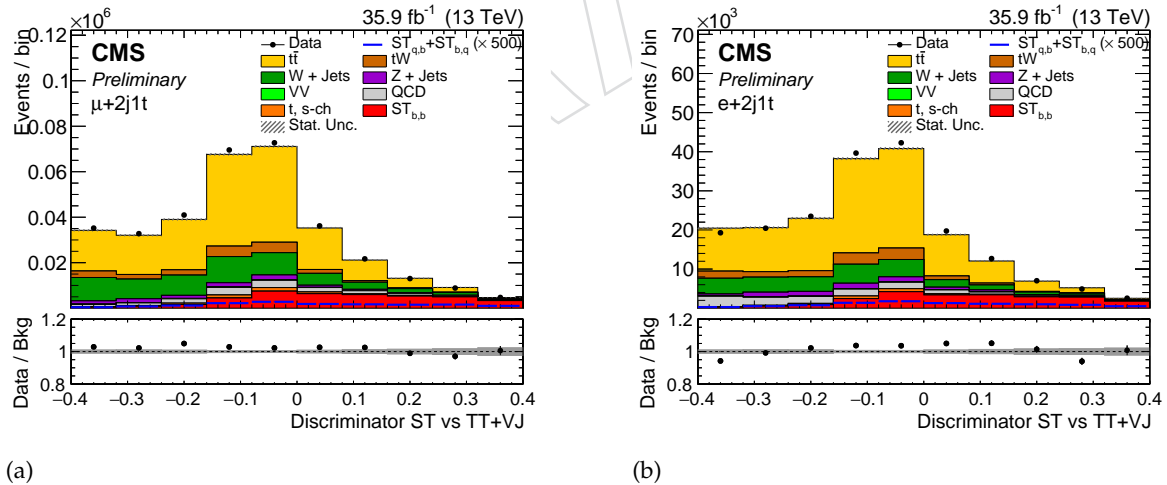
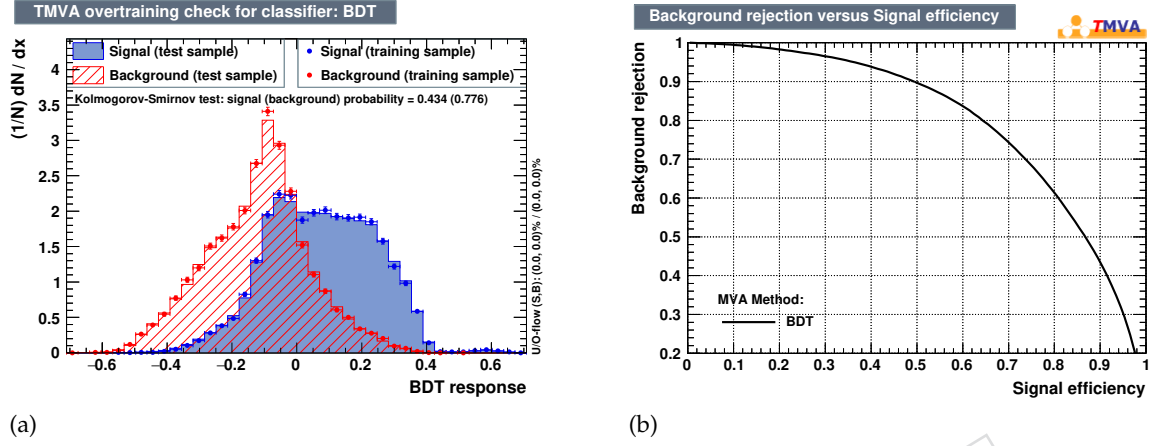


(a)

(b)

Figure 7: Linear correlation coefficients between the input variables: (a) for the signal single-top  $t$ -ch. and (b) for the background  $t\bar{t}$  and  $W$ + jets.

Figure 9 shows the pre-fit distributions for the discriminator trained in the 2-jets-1-tag QCD-depleted region.





#### 4.4 Variables in 3j1t

The 3-jets–1-tag region represents the signal region. After the  $m_T^W > 50 \text{ GeV}/c^2$  cut the most discriminating variables are, for example, the pseudo-rapidity of the light jet, the W-boson transverse mass and the response of the CMVAv2 tagger algorithm when applied to the extra jet selected.

Additional jets are selected within the range  $20 < p_T < 40 \text{ GeV}$  and having  $|\eta| < 2.4$ . The leading extra jet is used to reconstruct a top-quark candidate. In case it is an event of the  $ST_{(b,q)}^{2J1T}$  type, the b-jet of the 3-jets–1-tag will be stemming from gluon splitting, and the additional jet will have a high chance of being the one stemming from top-quark decay to s/d quarks.

Variables of interest in this sense are for example the mass between the lepton and either jet (the b-jet or the leading extra jet), the top-quark kinematic distributions built using the combination of the leading extra jet, MET and the lepton.

Then the QCD-depleted region is been divided into two subregions with a  $|\eta_{j'}|$  cut: a QCD-depleted-central region with a  $|\eta_{j'}| < 2.5$  cut and a QCD-depleted-forward region with  $|\eta_{j'}| > 2.5$  cut. The only region of interest in this case is the forward one because it's enriched in signal events and so it's the only one used in the analysis.

Figure 10 shows the most discriminating variables for the single-top  $t$ -ch. and two most important backgrounds in this region:  $t\bar{t}$  and  $W+$  jets. These variables are used in the discriminator with other variables in order to maximise the separation between signal and background.

Figure 16 shows the shape of the variables used as input to the BDT trained in the 3-jets–1-tag region.

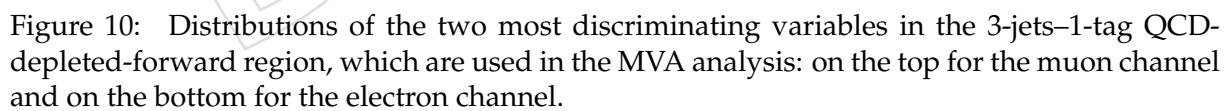
Figure 17 shows the linear correlation coefficients between the input variables.

The table 9 shows the results of the BDT discriminator in the QCD-depleted-forward region in which single-top  $t$ -ch. with  $|V_{tq}|^2$  vertex in decay is considered as signal and single-top  $t$ -ch.,  $t\bar{t}$  and  $W+$  jets are considered as background.

Figure 13 shows the overtraining test result and the ROC curve for the BDT obtained.

Figure 14 shows the pre-fit distributions for the discriminator trained in the 3-jets–1-tag QCD-depleted-forward region.





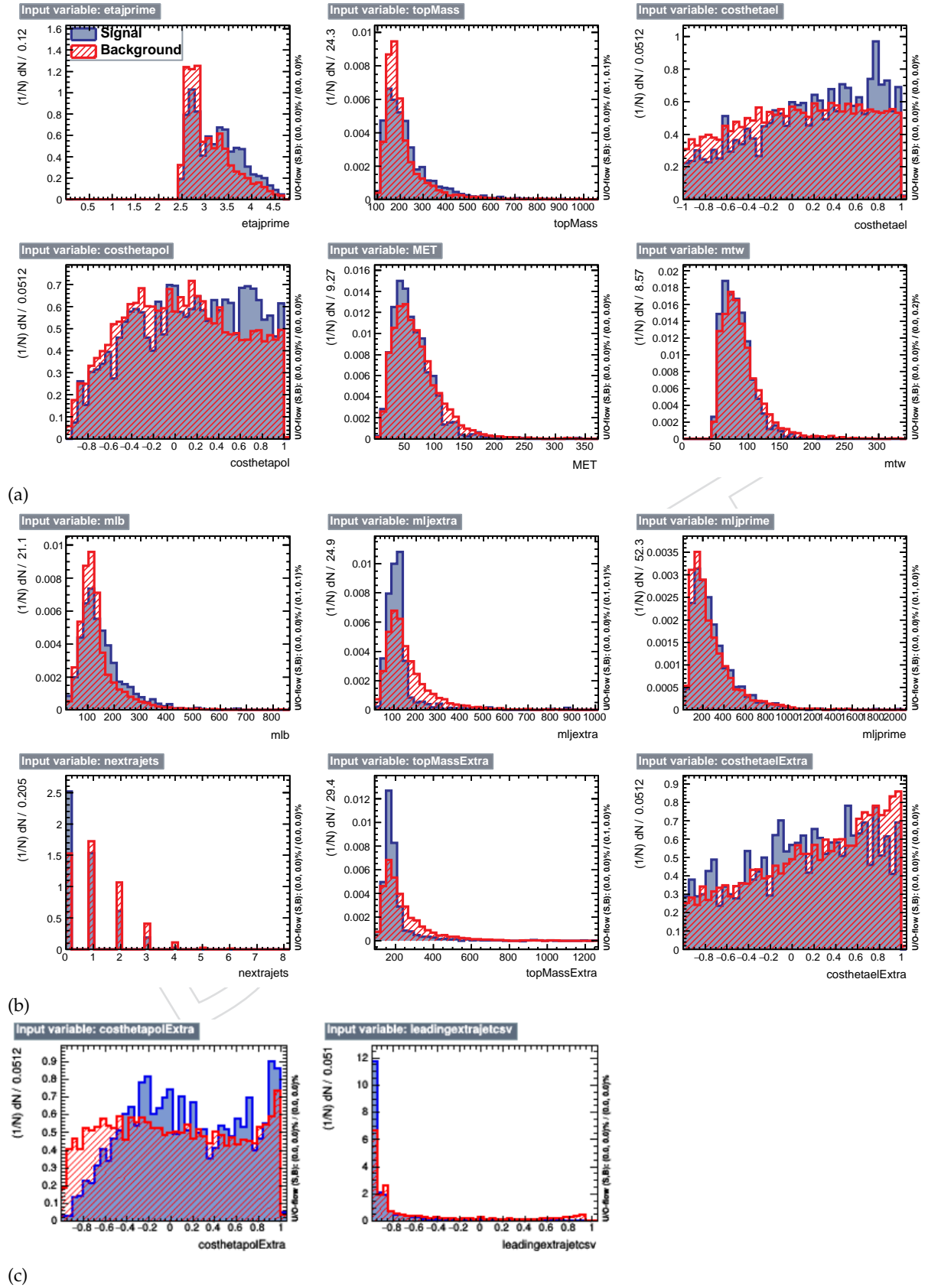


Figure 11: the plots show the shape of the variables used as input of the BDT training in the 3-jets-1-tag region.

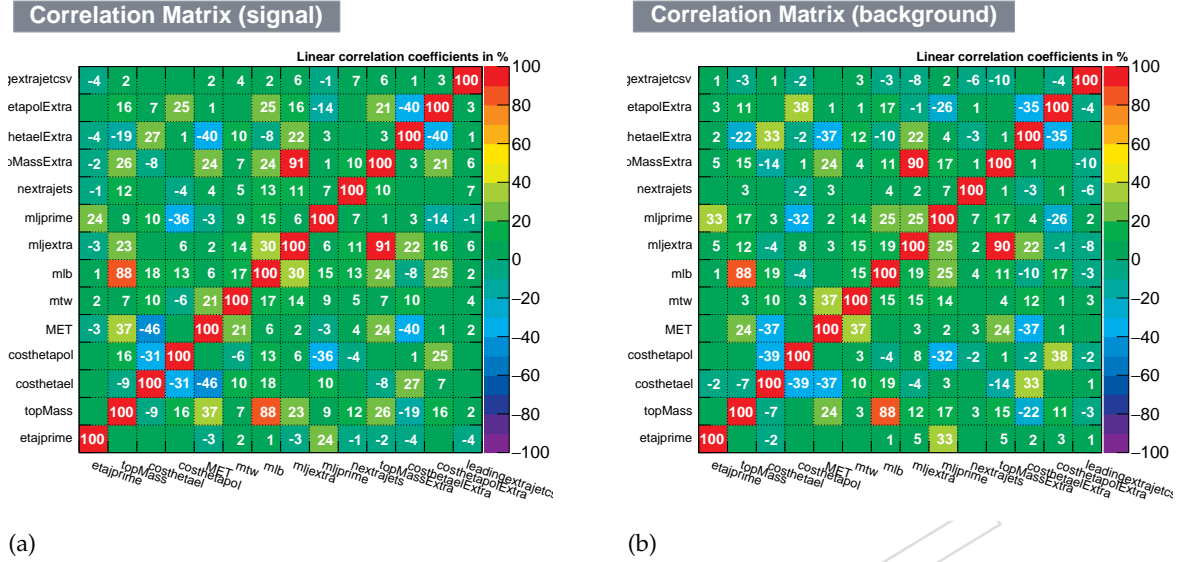
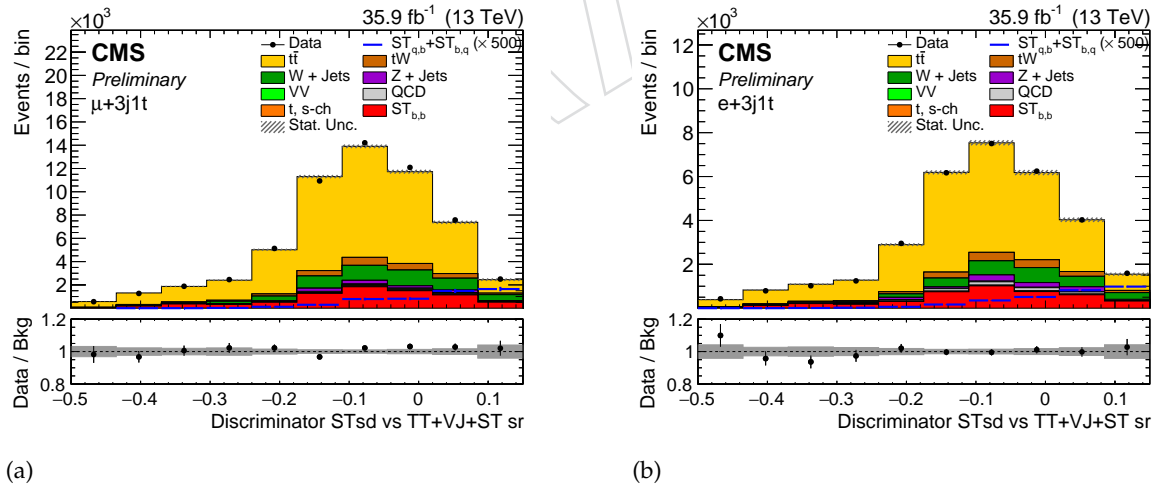
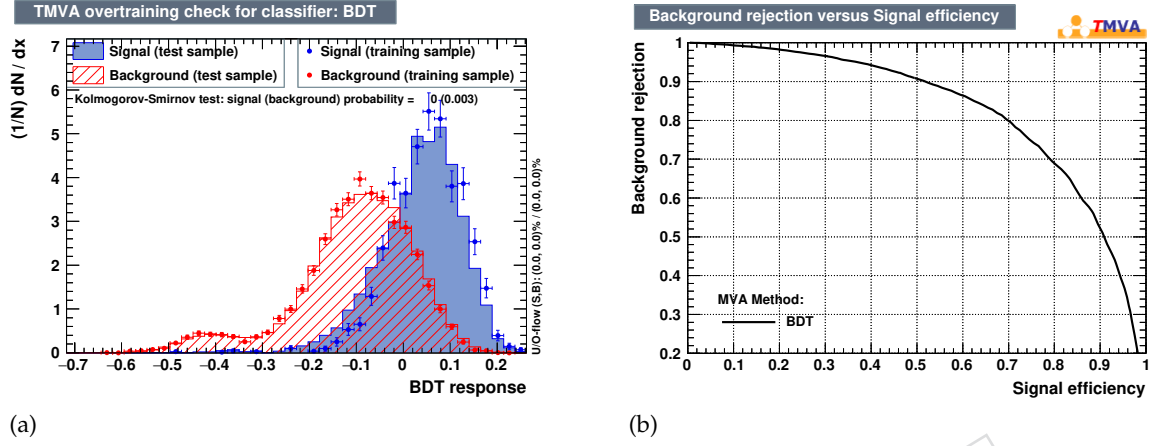


Figure 12: Linear correlation coefficients between the input variables: (a) for the signal single-top  $t$ -ch. with  $|V_{tq}|^2$  vertex in decay and (b) for the background single-top  $t$ -ch.,  $t\bar{t}$  and  $W+$  jets.

Table 9: Input variables for the BDTs trained for the single-top-quark  $t$ -ch. with  $|V_{tq}|^2$  vertex in decay versus standard single-top  $t$ -ch.,  $t\bar{t}$  and  $W+$  jets ranked according to their importance in the muon channel in the 3-jets-1-tag QCD-depleted-forward region.

Rank	Variable	Rel. imp. [%]	Correlation to BDT	
			Signal	Background
1	CMVA extra jet	9.2	-0.41	-0.61
2	W-boson transverse mass	8.6	-0.06	-0.15
3	$\vec{E}_T$	8.4	-0.15	-0.11
4	$\cos \theta_{hel}^*$	8.3	+0.17	+0.05
5	$ \eta $ light jet	7.6	+0.33	+0.08
6	top-quark mass extra	7.1	-0.29	-0.18
7	no. extra jets	7.1	-0.39	-0.21
8	$\cos \theta_{pol}^*$	6.9	+0.16	+0.14
9	lepton + extra jet mass	6.9	-0.23	-0.19
10	$\cos \theta_{pol}^*$ extra	6.6	+0.09	+0.22
11	lepton + light jet mass	6.0	+0.15	-0.07
12	lepton + b-jet mass	6.0	+0.10	+0.07
13	$\cos \theta_{hel}^*$ extra	5.9	+0.01	-0.06
14	top-quark mas	5.4	+0.04	+0.07



### 4.5 Variables in 3j2t

In the 3-jets–2-tags there are two  $b$  jets one stemming from the top-quark decay and one stemming from the gluon splitting. Both  $b$  jets are used to reconstruct a top-quark candidate and its correlated variables. In this case it is unnecessary to apply the  $m_T^W > 50 \text{ GeV}/c^2$  cut. Also the cut on  $|\eta_j|$  is irrelevant because this case is dominated by  $t\bar{t}$  sample. So only the study of single-top  $t$ -ch. versus  $t\bar{t}$  is performed. In this region variables of interest are the ones involving the kinematic variables of the light jet, like as the pseudorapidity of the light jet and the invariant mass between the lepton and the light jet. Figure 15 shows the most discriminating variables for the single-top  $t$ -ch. and the  $t\bar{t}$  background. These variables are used in the discriminator with other variables in order to maximise the separation between signal and background.

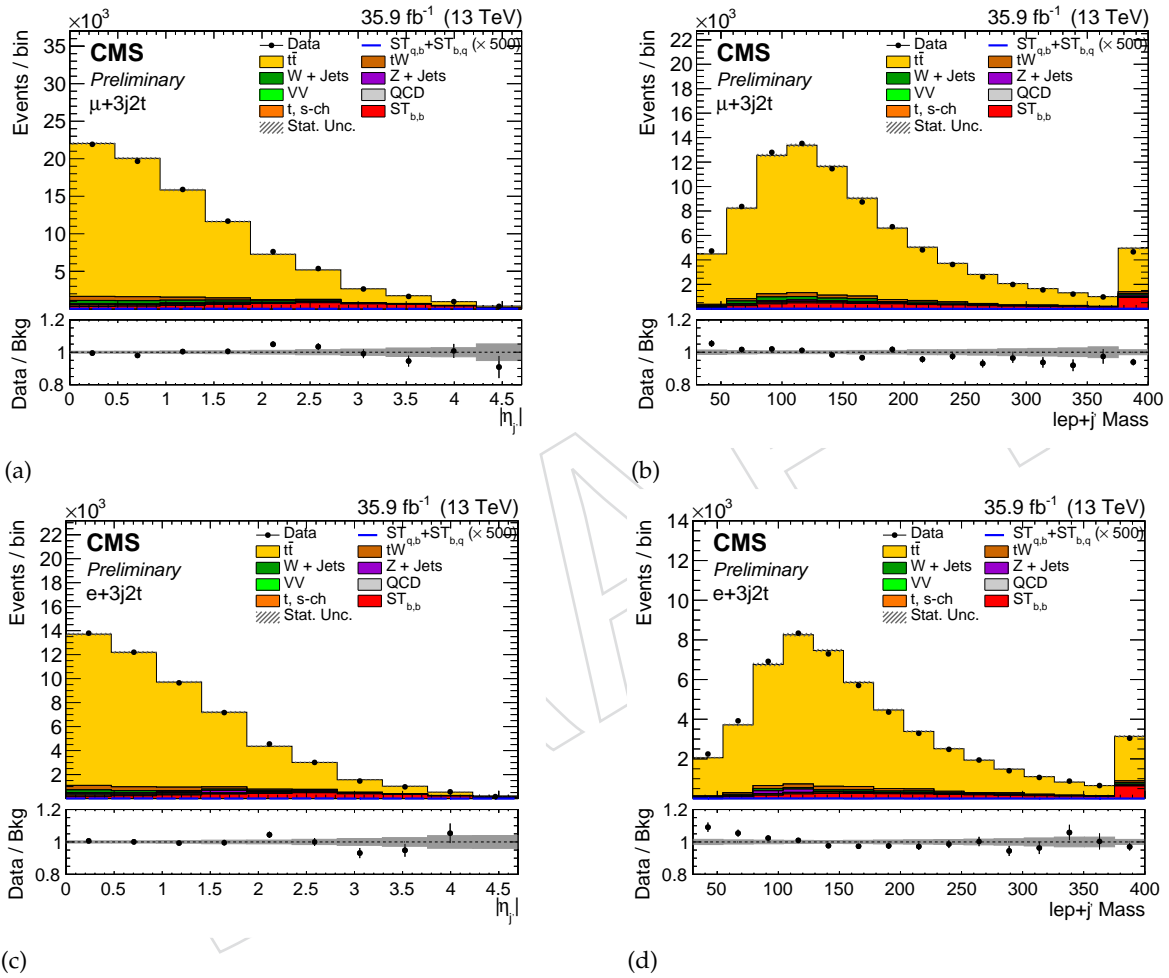


Figure 15: Distributions of the two most discriminating variables in the 3-jets–2-tags region, which are used in the MVA analysis: on the top for the muon channel and on the bottom for the electron channel.

Figure 16 shows the shape of the variables used as input to the BDT trained in the 3-jets–2-tags region.

Figure 17 shows the linear correlation coefficients between the input variables.

The table 10 shows the results of the MVA in which single-top  $t$ -ch. is considered as signal and  $t\bar{t}$  is considered as background.

Figure 18 shows the overtraining test result and the ROC curve for the BDT obtained.

Table 10: Input variables for the BDTs trained for the single-top-quark  $t$ -ch. versus  $t\bar{t}$  ranked according to their importance in the muon channel.

Rank	Variable	Rel. imp. [%]	Correlation to BDT	
			Signal	Background
1	$ \eta $ light jet	12.0	+0.77	+0.66
3	lepton + light jet mass	9.3	+0.38	+0.26
2	W-boson transverse mass	9.2	-0.06	+0.26
4	no. extra jets	9.1	-0.47	-0.43
5	$\cos \theta_{pol}^*$ second	7.0	+0.20	+0.14
6	$ \Delta\eta_{b-b} $	6.9	+0.02	-0.02
7	$\cos \theta_{hel}^*$ leading	6.7	+0.11	+0.06
8	top-quark mass leading	6.6	-0.04	+0.01
9	lepton + second b-jet mass	6.6	-0.00	-0.13
10	lepton + leading b-jet mass	6.4	+0.01	-0.00
11	top-quark mass second	5.2	-0.03	-0.06
12	$\cos \theta_{pol}^*$ leading	5.2	+0.11	+0.11
13	$\cos \theta_{hel}^*$ second	5.0	-0.07	+0.03
14	$\vec{E}_T$	4.9	-0.036	-0.10

Figure 19 shows the pre-fit distributions for the discriminator trained in the 3-jets-2-tags region.

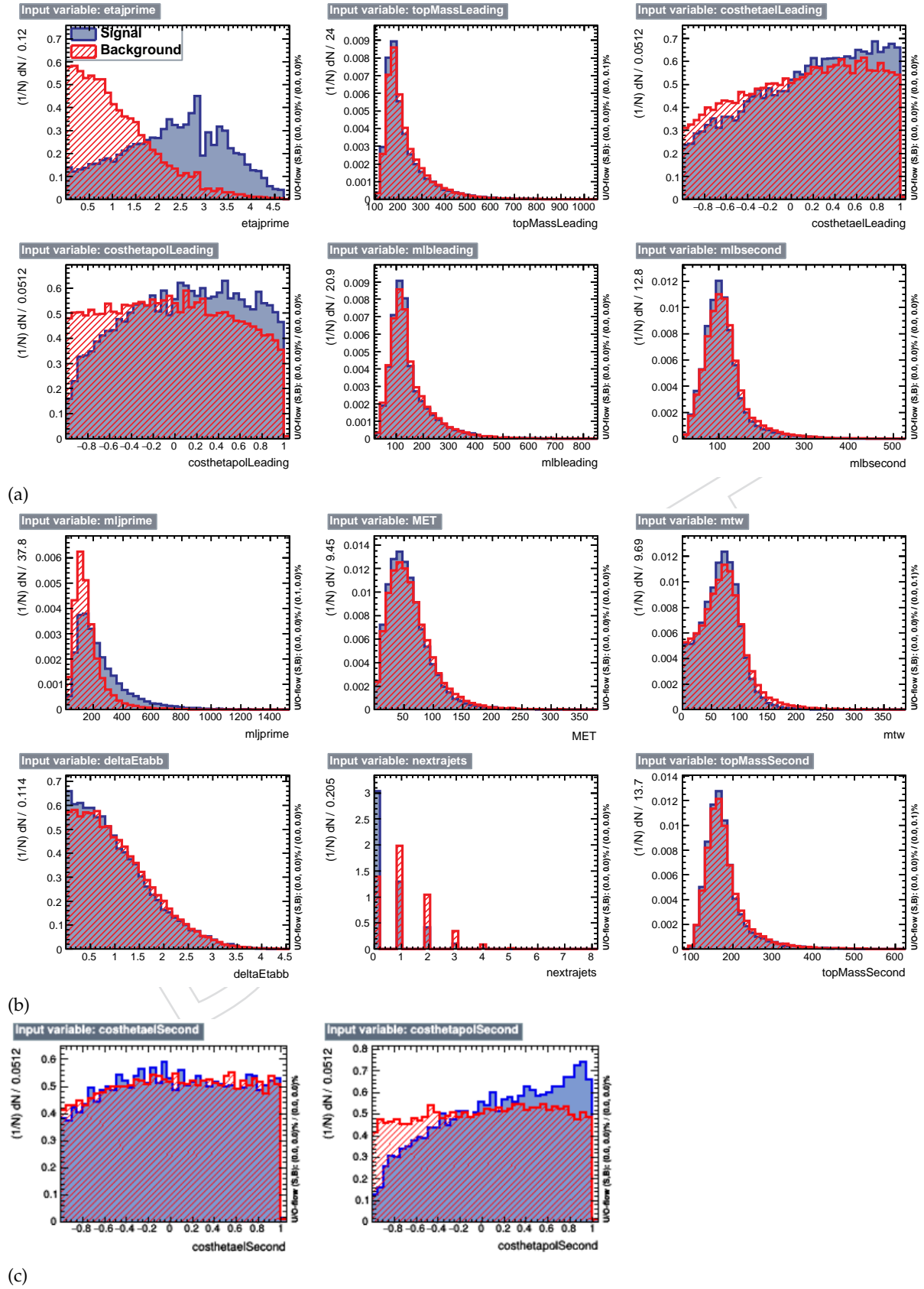
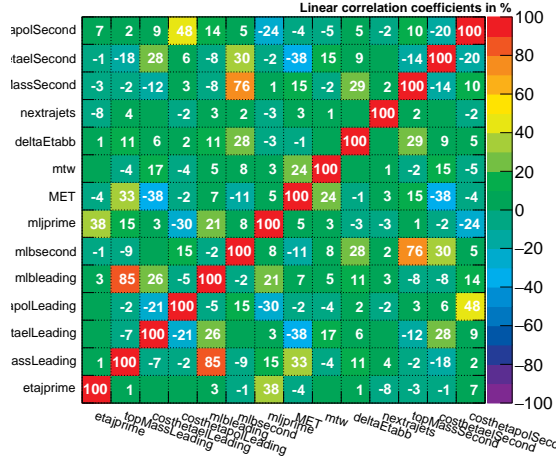


Figure 16: the plots show the shape of the variables used as input of the BDT training in the 3-jets-2-tags region.

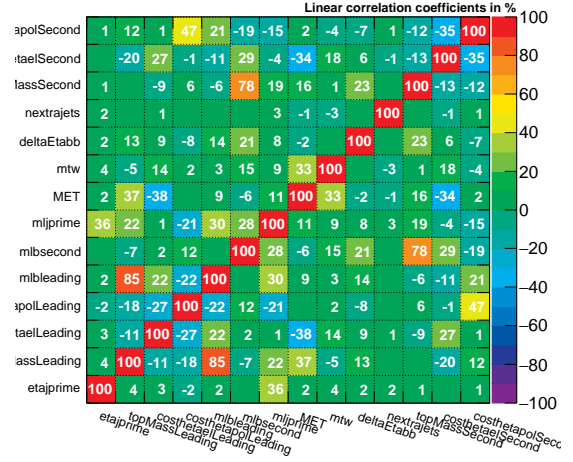


Correlation Matrix (signal)



(a)

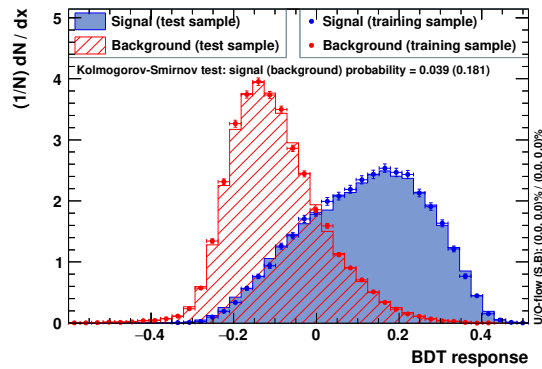
Correlation Matrix (background)



(b)

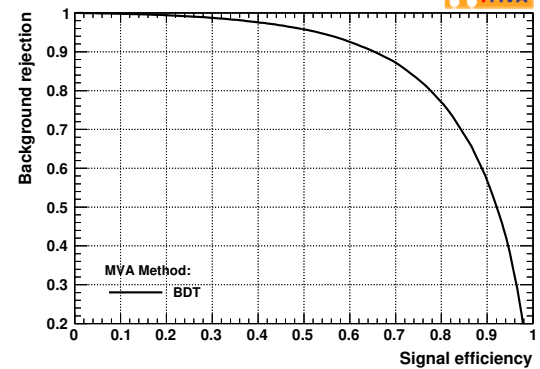
Figure 17: Linear correlation coefficients between the input variables: (a) for the signal single-top  $t$ -ch. and (b) for the background  $t\bar{t}$ .

TMVA overtraining check for classifier: BDT



(a)

Background rejection versus Signal efficiency



(b)

Figure 18: (a) Classifier output distributions for the test (histograms) and training samples (dots) superimposed; (b) Classifier background rejection vs signal efficiency.



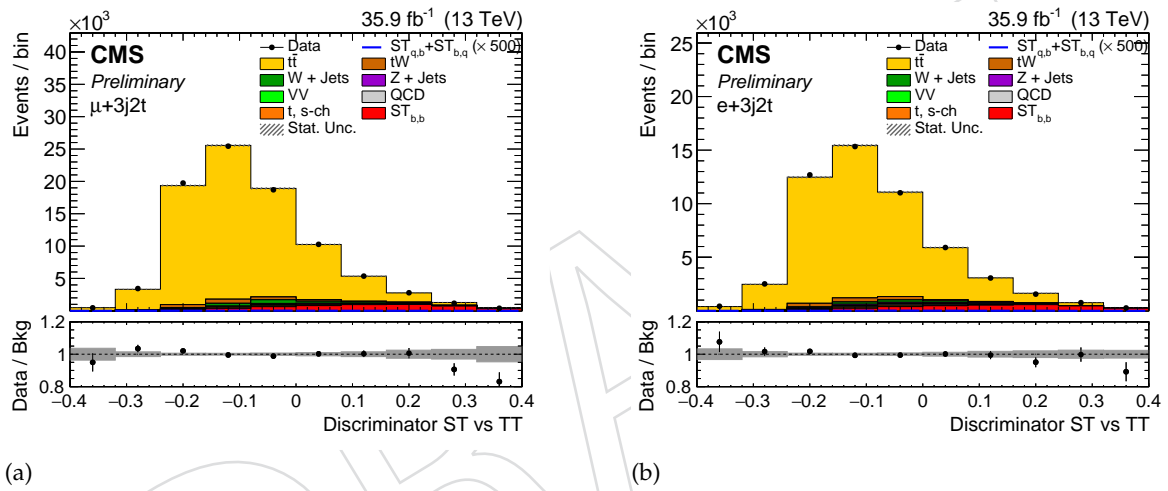


Figure 19: Distributions of the discriminators in the 3-jets-2-tags region: on the left for the muon channel and on the right for the electron channel.

## 5 Fit procedure

### 5.1 Maximum Likelihood fit

The maximum likelihood fit is performed on the distribution of the multivariate discriminators with the scale factor being the free parameter of the fit in order to normalize the  $t$ -ch. signal process. The scale factor is defined by

$$S_i = \frac{N_i}{N_i^{\text{exp}}} \quad (40)$$

where  $N_i$  is the number of events after the fit,  $N_i^{\text{exp}}$  the expected number of events and  $i$  the process category. The distribution of the multivariate discriminators in the signal region (2-jets–1-tag) are fitted simultaneously with the ones from two control regions (3-jets–1-tag and 3-jets–2-tags). As the latter are dominated by  $t\bar{t}$  events, including these control regions improves the precision on the determination of the  $t\bar{t}$  contribution.

The single-top-quark  $t$ -channel cross section is then calculated with the formula:

$$\sigma_t = \frac{N_s}{\epsilon \cdot \mathcal{B}(t \rightarrow \ell \nu b) \cdot L} \quad (41)$$

where  $N_s$  is the number of signal events,  $\epsilon$  the selection efficiency,  $\mathcal{B}(t \rightarrow \ell \nu b) = 0.326$  [26] the overall leptonic branching ratio and  $L$  the integrated luminosity.

### 5.2 Fit strategy

The CKM matrix elements are extracted from the measured cross-section of the single-top  $t$ -ch. process. For the extraction of single-top  $t$ -ch. production cross-section we perform a simultaneous maximum likelihood fit to the multivariate distributions in 2-jets–1-tag, 3-jets–1-tag and 3-jets–2-tags categories. Higgs combination tool [27] is used to perform the statistical inference. The tool makes use of ROOSTATS for the calculations. The further details on ROOSTATS can be found at [28].

The fit procedure is divided in three steps:

- In the first step, a Maximum Likelihood fit to the  $m_T^W$  distribution is performed for the QCD scale factor extraction. The scale factor is then used for the normalisation of QCD template.
- In the second step, the light quark distribution  $|\eta_{j'}|$  and  $m_T^W$  are used to define different regions:
  - 2-jets–1-tag in order to control the  $t\bar{t}$ ,  $W$ + jets together with the  $t$ -channel a  $m_T^W > 50$  GeV is required to deplete the QCD.
  - 3-jets–1-tag in order to have a region enriched in  $t$ -channel  $V_{td}$ - $V_{ts}$  events with  $|\eta_{j'}| > 2.5$  and  $m_T^W > 50$  GeV are required.
  - 3-jets–2-tags No selection is performed in this region.
- In the last step, the Maximum Likelihood fit is performed simultaneously in the regions defined in step 2, while the QCD prior uncertainty and central value are taken from step 1. The presence of standard top-quark decay products,  $|V_{tb}|$ , is now one of the major background to the  $|V_{td}|$  or  $|V_{ts}|$  extraction.

### 5.3 Fit variables

The  $m_T^W$  variable is used for a first estimation of the QCD-Multijet background and the flat prior fitted is been used as scale factor for the QCD template. In the QCD-depleted region a MVA analysis is performed, as described in Sec. 4.5, in order to obtain one MVA variables in the QCD-depleted region in the 2-jets–1-tag category and another MVA variable in the QCD-depleted-forward region in the 3-jets–1-tag category. The MVAs variables are obtained in the following way:

- the first is obtained considering the single-top  $t$ -ch. as signal and the  $t\bar{t}$  and the  $W$ + jets are considered as background in the QCD-depleted region in the 2-jets–1-tag category;
- the second is obtained considering the single-top  $t$ -ch. decaying in  $s/d$  quarks as signal and the  $t\bar{t}$ , standard single-top  $t$ -ch. and  $W$ + jets are considered as background in the QCD-depleted-forward region in the 3-jets–1-tag category;

At the end, for the 3-jets–2-tags, the multivariate analysis consists in only one BDT variable obtained considering the single-top  $t$ -channel as signal and the  $t\bar{t}$  is considered as background without any cut.

### 5.4 Maximum Likelihood fit

The maximum likelihood fit is performed simultaneously in the 2-jets–1-tag region (QCD-depleted), in the 3-jets–1-tag region (QCD-depleted-forward), and in the 3-jets–2-tags region.

The QCD and the  $t$ -ch., splitted for the muon and the electron case, are floating with a flat prior, all systematic uncertainties defined in AN-17-068 are treated as nuisance parameters. The yield of  $t\bar{t}$ ,  $W$ + jets is left floating together with the respective scale uncertainties.

### 5.5 Fit results

The Table 11 shows the values of the pulls with pseudo-data for any nuisance parameter, the ratio of error in the model before and after the fit and  $\rho$  gives the correlation coefficient between the signal strength  $r$  and for each nuisance parameter.

In the Figure 20 the same parameters reported in the Table 11 are plotted.

Figure 21 shows the stacks obtained after the fit procedure for muon channel and for electron channel.

The scale factor post/pre fit for the single-top-quark  $t$ -ch. produced via and decaying to  $b$  quarks is:

$$\mu_{ST_{b,b}} = 0.99 \pm 0.03 \pm 0.12.$$

For the single-top-quark  $t$ -ch. decaying to, or produced via  $s/d$  quarks the fit returns the scale factor:

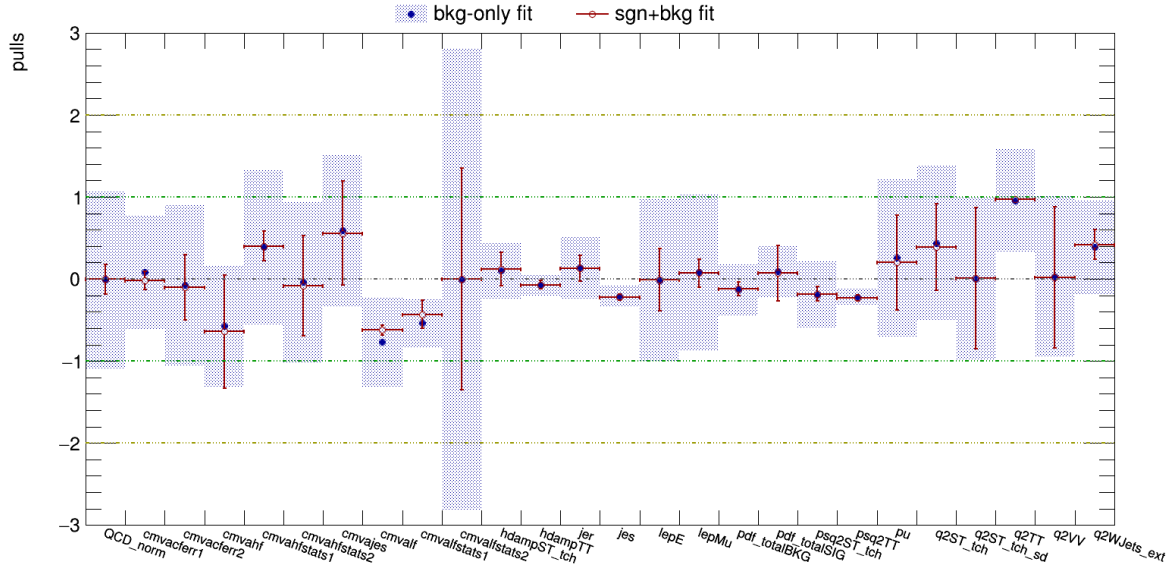
$$\mu_{ST_{b,q}} < 86 \text{ at } 95\% \text{ C.L.},$$

with an expected value of:

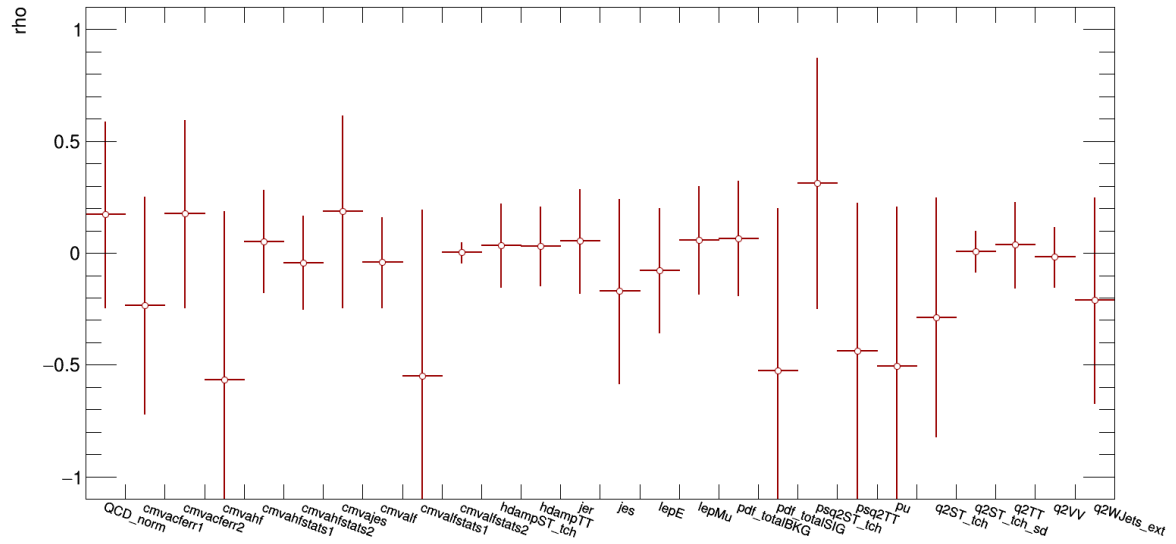
$$\mu_{ST_{b,q}}^{exp} = 72.$$

Table 11: Pulls with pseudo-data for any nuisance parameter, the ratio of error in the model before and after the fit and  $\rho$  gives the correlation coefficient between the signal strength  $r$  and for each nuisance parameter.

name	$b$ -only fit		$s + b$ fit		$\rho(\theta_\mu)$
	$\Delta x / \sigma_{\text{in}}$	$\sigma_{\text{out}} / \sigma_{\text{in}}$	$\Delta x / \sigma_{\text{in}}$	$\sigma_{\text{out}} / \sigma_{\text{in}}$	
QCD_norm	-0.01	1.08	+0.00	0.18	+0.1718
cmvacferr1	+0.08	0.69	-0.02	0.11	-0.2350
cmvacferr2	-0.07	0.98	-0.10	0.40	+0.1751
cmvahf	-0.57	0.74	-0.64	0.69	-0.5670
cmvahfstats1	+0.39	0.94	+0.40	0.18	+0.0523
cmvahfstats2	-0.04	0.98	-0.08	0.61	-0.0434
cmvajes	+0.59	0.92	+0.56	0.63	+0.1853
cmvalf	-0.77	0.55	-0.62	0.06	-0.0407
cmvalfstats1	-0.54	0.29	-0.43	0.17	-0.5487
cmvalfstats2	-0.01	2.81	-0.00	1.35	+0.0022
hdampST_tch	+0.10	0.34	+0.12	0.20	+0.0345
hdampTT	-0.07	0.13	-0.07	0.05	+0.0313
jer	+0.14	0.38	+0.13	0.16	+0.0539
jes	-0.21	0.13	-0.22	0.04	-0.1708
lepE	-0.02	0.99	-0.01	0.38	-0.0780
lepMu	+0.08	0.96	+0.07	0.17	+0.0583
pdf_totalBKG	-0.13	0.32	-0.12	0.08	+0.0656
pdf_totalSIG	+0.09	0.32	+0.07	0.34	-0.5272
psq2ST_tch	-0.19	0.40	-0.18	0.09	+0.3134
psq2TT	-0.22	0.10	-0.23	0.04	-0.4387
pu	+0.26	0.96	+0.20	0.58	-0.5053
q2ST_tch	+0.44	0.94	+0.39	0.53	-0.2868
q2ST_tch_sd	+0.00	0.99	+0.01	0.86	+0.0084
q2TT	+0.95	0.63	+0.97	0.02	+0.0367
q2VV	+0.03	0.98	+0.02	0.86	-0.0181
q2WJets_ext	+0.39	0.57	+0.42	0.18	-0.2116

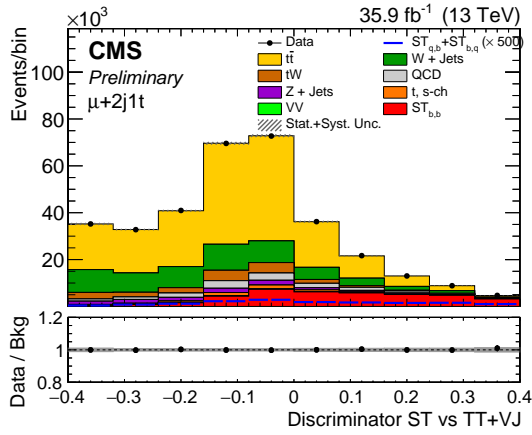


(a)

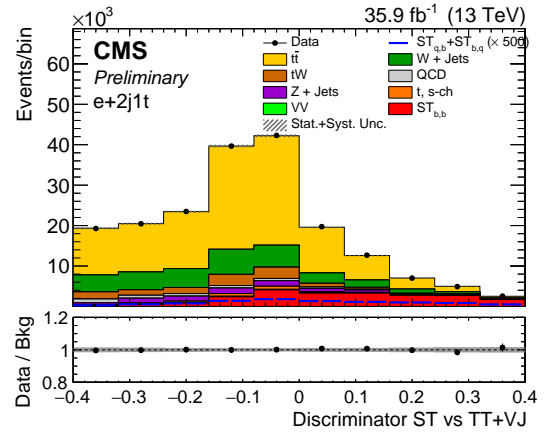


(b)

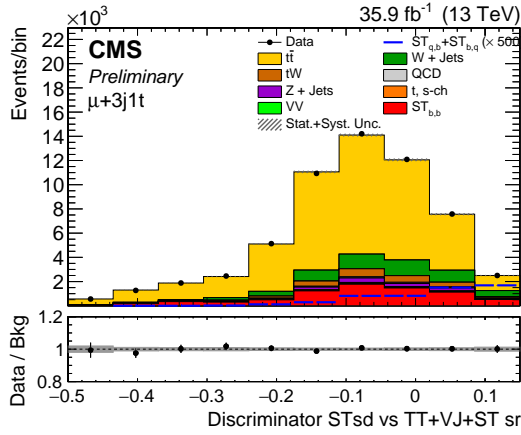
Figure 20: Pulls with pseudo-data for each nuisance parameter and the correlation coefficient between the signal strength  $r$  and for each nuisance parameter.



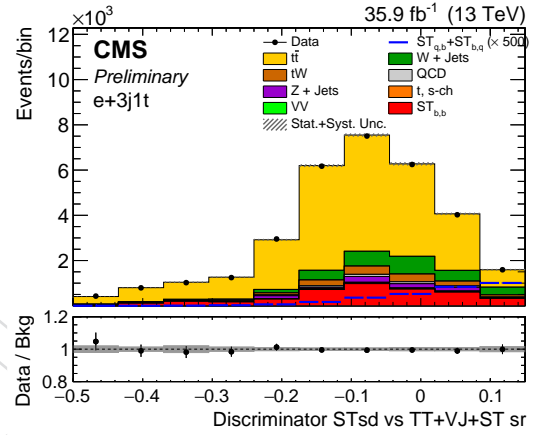
(a)



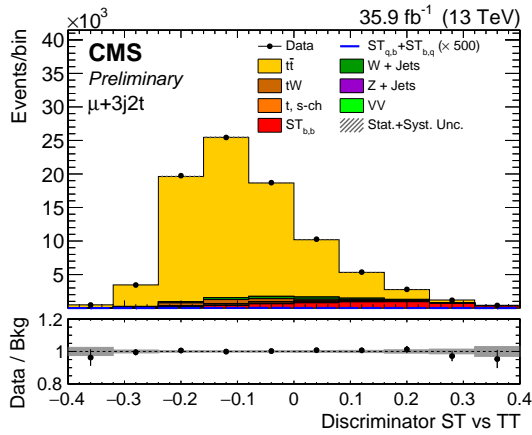
(b)



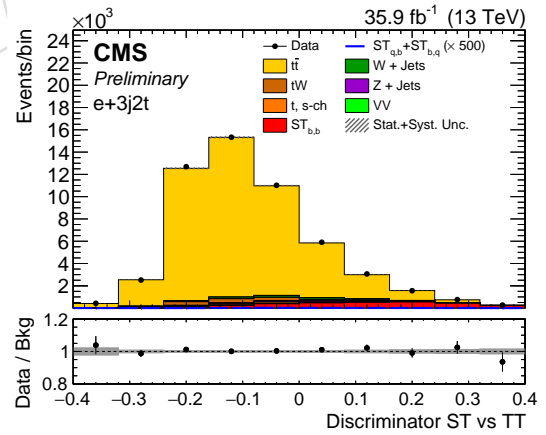
(c)



(d)



(e)



(f)

Figure 21: Stacks obtained after the fit procedure for muon channel on the left and for the electron on the right.

## 6 Systematics

Several sources of systematic uncertainties are considered in the analysis, either as nuisance parameters in the fit of the BDT discriminator distributions (profiled uncertainties) or as external uncertainties. The latter category includes the uncertainty sources related to the modeling of the signal process, which can not be constrained from the measurement as they apply to the full phase space and not only to the region in which the measurement is performed. Their impact is determined by repeating the analysis using varied templates according to the systematic uncertainty sources under study in the fit instead of the nominal templates. The uncertainty due to a certain source is then taken as half the difference between the results for up and down variations of the effect. In the following, the different uncertainty sources that are considered in the analysis are briefly described. For the sake of simplicity and better readability, they are grouped in categories of related sources.

### Profiled uncertainties

- **Limited size of samples of simulated events:** the statistical uncertainty due to limited simulated sample size is evaluated for each bin and each process with the Barlow-Beeston light method [29, 30].
- **Jet energy resolution (JER):** A smearing is applied to account for the difference in the JER between simulation and data [31], increasing or decreasing the resolutions by their uncertainties.
- **Lepton trigger and reconstruction:** Single-muon and single electron trigger efficiency and reconstruction efficiency are estimated with a “tag-and-probe” method [32] from Drell–Yan events in the Z boson mass peak.
- **Pileup:** The uncertainty in the average expected number of pileup interactions is propagated as a source of systematic uncertainty to this measurement by varying the minimum bias cross section by  $\pm 4.6\%$ . The effect on the result is found to be negligible and is therefore not considered further.
- **$t\bar{t}$  modelling:** The following uncertainty sources cover potential mismodeling of the  $t\bar{t}$  process. The effect of those uncertainties both on the acceptance and on the cross section, is considered. For the maximum likelihood fit, the uncertainties are considered as nuisance parameters in the fit.
  - **Renormalization and factorization scale uncertainty ( $\mu_R/\mu_F$ ):** The uncertainties caused by variations in the renormalization and factorization scales are considered by reweighting the BDT response distributions with different combinations of doubled/halved renormalization and factorization scales with the nominal value set at 172.5 GeV.
  - **Matching of matrix element and parton shower (ME-PS matching):** The parameter that controls the matching between the matrix-element level calculation and the parton shower and regulates the high- $p_T$  radiation in the simulation is varied within its uncertainties.
  - **Initial and final state radiation:** The impact of variations in the initial-state and final-state radiation, depending on the choice of  $\alpha_s$ , is studied by comparing the nominal sample with dedicated  $t\bar{t}$  samples.
- **Underlying event:** The effect of uncertainties on the modeling of the underlying event is studied by comparing the nominal sample with dedicated  $t\bar{t}$  samples.

- **QCD yield uncertainty:** The yield of the QCD multijets process is taken with a 50% uncertainty, which is chosen conservatively to be much larger w.r.t. the  $m_T^W$  fit uncertainty.
- **Renormalisation and factorisation scale uncertainty ( $\mu_R/\mu_F$ ):** The uncertainties due to variations in the renormalisation and factorisation scales are studied for the  $tW$ , single-top  $t$ -ch. with a  $|V_{td}|$ ,  $|V_{ts}|$  vertex, and  $W$ + jets by reweighting the distributions with different combinations of halved or doubled factorisation and renormalisation scales. The effect is estimated for each process separately.
- **PDF for background processes:** The uncertainty due to the choice of PDFs is estimated using reweighted histograms derived from all PDF sets of NNPDF 3.0 [33].
- **b-tagger algorithm uncertainties:**
  - JES<  $x$  >: Uncertainty due to JES source  $x$ .
  - LF: Contamination from  $udsg+c$  jets in HF region.
  - HF: Contamination from  $b+c$  jets in LF region.
  - Linear and quadratic statistical fluctuations:
    - HFstats1/HFstats2:  $b$  jets.
    - LFstats1/LFstats2:  $udsg$  jets.
    - cErr1/ cErr2:  $c$  jets.

#### Externalized uncertainties

- **Luminosity:** The integrated luminosity is known with a relative uncertainty of  $\pm 2.6\%$  [34].
- **Jet energy scale (JES):** All reconstructed jet four-momenta in simulated events are simultaneously varied according to the  $\eta$ - and  $p_T$ -dependent uncertainties in the JES [31]. This variation in jet four-momenta is also propagated to  $\cancel{p}_T$ .
- **PDF for signal process:** The uncertainty due to the choice of PDFs is estimated using reweighted histograms derived from all PDF sets of NNPDF 3.0 [33]. Only the effect on the acceptance, and not on the cross section, is considered.
- **Signal modelling:** The following uncertainty sources cover potential mismodeling of the single top quark  $t$ -channel signal process. The effect of those uncertainties on the acceptance, and not on the cross section, is considered. For the maximum likelihood fit, the uncertainties are not considered as nuisance parameters in the fit but evaluated by repeating the full analysis using samples of simulated signal events that feature variations in the modeling parameters covering the systematic uncertainty sources under study. In order to perform the limit extraction uncertainties are profiled to provide an upper limit on  $|V_{td}|$  and  $|V_{ts}|$ .
  - **Renormalization and factorization scale uncertainty ( $\mu_R/\mu_F$ ):** The uncertainties caused by variations in the renormalization and factorization scales are considered by reweighting the BDT response distributions with different combinations of doubled/halved renormalization and factorization scales with the nominal value set at 172.5 GeV.
  - **Matching of matrix element and parton shower (ME-PS matching):** The parameter that controls the matching between the matrix-element level calculation and the parton shower and regulates the high- $p_T$  radiation in the simulation is varied within its uncertainties.



- **Parton shower scale:** The parton shower scale in the simulation is varied within its uncertainties.

Figures 22 and 23 show the impacts of the profiled systematics on the measurement of the single top  $t$ -ch. cross section.

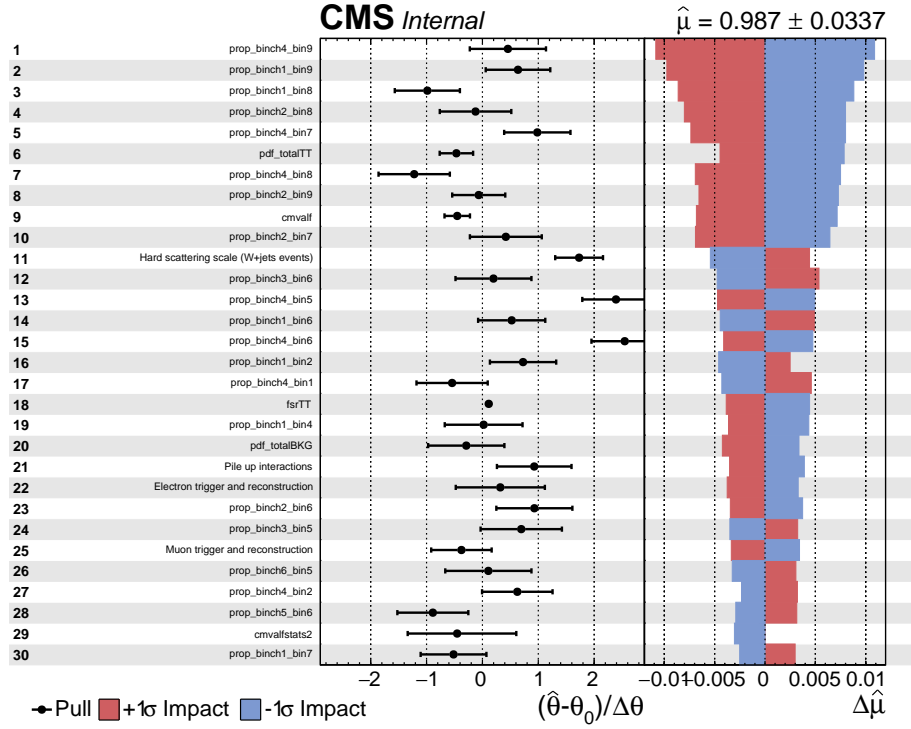
DRAFT

613 For the externalized uncertainties, the change of the result due to the respective variation is listed in % in Table 12.

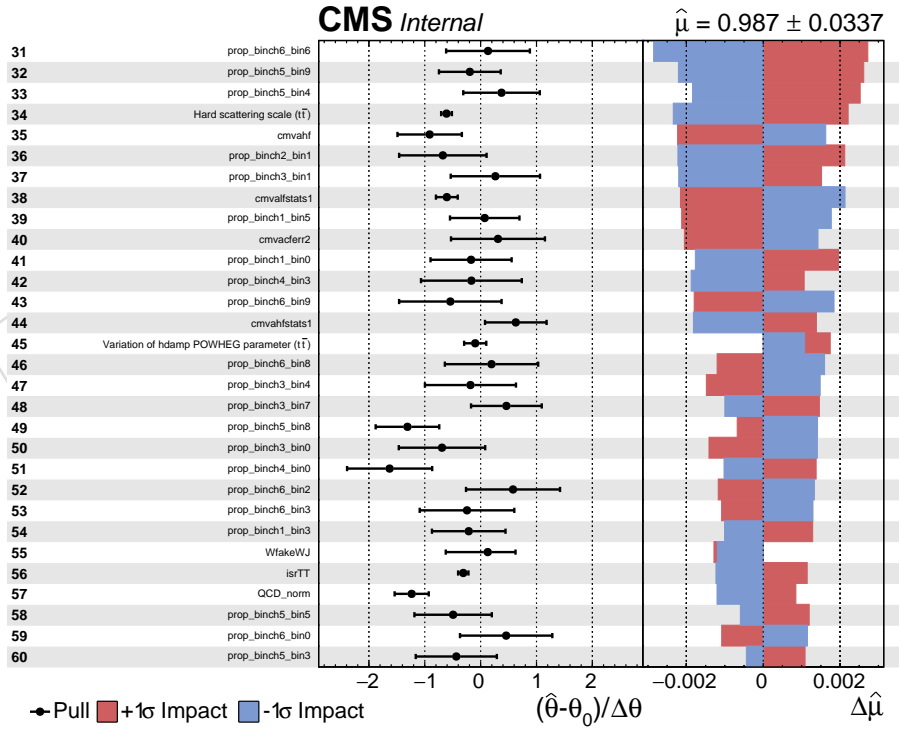
Table 12: Estimated relative contributions of the listed uncertainty sources in % to the total uncertainties of the measured cross sections for single-top quark production.

Treatment	Uncertainty	$\Delta\sigma/\sigma(\%)$
Profiled	Total (Stat+Syst)	3.8
Externalized	PDF $_{ST(b,b)}^{2j1T}$	3.8
	JES	8.0
	Parton shower scale	6.1
	ME-PS matching	3.7
	JER	2.8
	Luminosity	2.6
	Renormalization and factorization scale uncertainty ( $\mu_R/\mu_F$ )	2.4
	Total	11.5

614

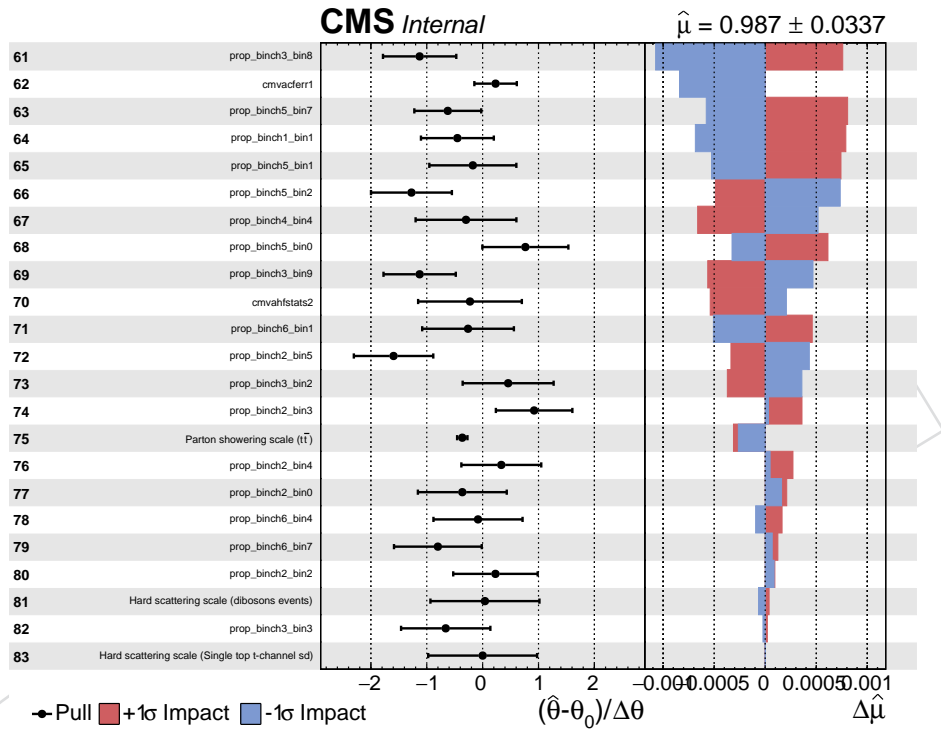


(a)



(b)

Figure 22: Impacts of the primary sources of systematic uncertainty treated as nuisance parameters in the fit of the BDT discriminator distributions (profiled uncertainties). Uncertainties are ranked by impact on the single top  $ST_{(b,b)}^{nJmT}$  component, ranks 1-30 are shown in (a), ranks 31-60 are shown in (b).



(a)

Figure 23: Impacts of the secondary sources of systematic uncertainty treated as nuisance parameters in the fit of the BDT discriminator distributions (profiled uncertainties). Uncertainties are ranked by impact on the single top  $ST_{(b,b)}^{n|_{mT}}$  component, ranks from 61 below are shown.

## 7 Results and interpretation

In order to extract information on the different matrix elements, the contributions of each matrix element to the different  $ST_{(b,b)}^{2J1T}$ ,  $ST_{(b,q)}^{2J1T}$ , and  $ST_{(q,b)}^{2J1T}$  cross sections, extracted from the fit procedure, are considered.

In the standard model, top quarks only decay to W bosons plus b, s, or d quarks, and their branching ratios are proportional to the respective matrix element, and relations in Table ?? apply.

We write the results of the fit procedure in terms of two signal strength parameters: the first one,  $\mu_b$  refers to the  $ST_{(b,b)}^{2J1T}$  process, and the second one,  $\mu_{sd}$ , refers to the sum of the of the  $ST_{(q,b)}^{2J1T}$  and  $ST_{(b,q)}^{2J1T}$  contributions. By considering the ratio of  $|V_{td}|/|V_{ts}|$  small, we can write the entire  $ST_{(q,b)}^{2J1T}$  term as proportional to  $|V_{td}|^2 + |V_{ts}|^2$  plus orders of  $|V_{td}|/|V_{ts}| \cdot \sigma_{t\text{-ch},d}/\sigma_{t\text{-ch},s}$ . The signal strength thus becomes:

$$\begin{aligned}\mu_b &= \frac{\sigma_{t\text{-ch},b}^{\text{obs}} \times \text{BR}(t \rightarrow Wb)^{\text{obs}}}{\sigma_{t\text{-ch},b} \times \text{BR}(t \rightarrow Wb)} \\ \mu_{sd} &= \frac{\sigma_{t\text{-ch},b}^{\text{obs}} \times (\text{BR}(t \rightarrow Wd,s))^{\text{obs}} + \sigma_{t\text{-ch},s,d}^{\text{obs}} \times (\text{BR}(t \rightarrow Wb))^{\text{obs}}}{\sigma_{t\text{-ch},b} \times (\text{BR}(t \rightarrow Wd,s)) + \sigma_{t\text{-ch},s,d} \times (\text{BR}(t \rightarrow Wb))}\end{aligned}\quad (42)$$

In a general case, one can write Eq. 42 in terms of the decay amplitudes, or widths. The partial width of top quark to Wq as  $\Gamma_b \cdot |V_{tq}|^2$ , where  $\Gamma_b$  is the partial width of the top quark with  $|V_{tb}| = 1$ , while we write the total top width as  $\Gamma_{\text{top}}$ :

$$\begin{aligned}\mu_b &= \frac{|V_{tb}|^4 \cdot \Gamma_q^{\text{obs}} \cdot \Gamma_{\text{top}}}{|V_{tb}|^4 \cdot \Gamma_q \cdot \Gamma_{\text{top}}^{\text{obs}}} \\ \mu_{sd} &= \frac{|V_{tb}|^2 \cdot (|V_{ts}|^2 \cdot \Gamma_q^{\text{obs}} + |V_{td}|^2 \cdot \Gamma_q^{\text{obs}}) \cdot \Gamma_{\text{top}}}{|V_{tb}|^2 \cdot (|V_{ts}|^2 + |V_{td}|^2) \cdot \Gamma_q \cdot \Gamma_{\text{top}}^{\text{obs}}}\end{aligned}\quad (43)$$

One can extract from the fit procedure the value of the signal strengths or directly the value of the CKM matrix elements  $|V_{td}|^2$ ,  $|V_{ts}|^2$ , and  $|V_{tb}|^2$ .

### 7.1 Measurement in unconstrained SM scenario

The simplest approach to derive the different matrix elements is to substitute  $\text{BR}_{(t \rightarrow Wq)}$  with  $|V_{tq}|^2 / (|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2)$ , with q indicating d, s, or b quarks, and the observed cross sections and branching fractions  $\sigma_{t\text{-ch},q}^{\text{obs}} \times \text{BR}(t \rightarrow Wq)^{\text{obs}}$ .

The results obtained from the two fits are:

$$\mu_b = 0.99 \pm 0.03(\text{stat.} + \text{syst.}) \pm 0.12(\text{externalized})$$

$$\mu_{sd} \leq 87 \text{ at } 95\% \text{ CL} \quad (44)$$

By making use of Eq. 42, and of the values reported in Table ?? and Table ??, and considering  $\text{BR}(t \rightarrow Wq) = R_q = |V_{tq}|^2$ , the fit yields are converted in terms of the elements of the CKM matrix, thus measuring:

$$\begin{aligned} |V_{tb}| &= 1.00 \pm 0.01 \pm 0.03 \\ |V_{tb}|^2 &= 0.99 \pm 0.02 \pm 0.06 \\ |V_{td}|^2 + |V_{ts}|^2 &\leq 0.17 \text{ at } 95\% \text{ C.L.} \end{aligned} \quad (45)$$

Results in Eq. 45 do not consider any constraint between  $|V_{tb}|, |V_{td}|, |V_{ts}|$  from unitarity of the CKM matrix in the standard model. Also, the above result does not consider the effect of  $R_b + R_s + R_d$  being different from one. The best estimation for the magnitude of the CKM matrix elements is the one obtained by the global fit of the SM:

$$|V_{tb}|^2 = 0.998239^{+0.000048}_{-0.000024} \quad |V_{ts}|^2 = 0.0016876^{+0.0000025}_{-0.0000047} \quad |V_{td}|^2 = 0.000074^{+0.000013}_{-0.000017}.$$

When  $|V_{td}|^2$  and  $|V_{ts}|^2$  are summed together one obtains:

$$|V_{td}|^2 + |V_{ts}|^2 = 0.0017611^{+0.0000028}_{-0.0000050}$$

## 7.2 Measurement in constrained SM scenario

In the SM case, one can simplify Eq. 42 by considering that  $|V_{tb}|^2 + |V_{td}|^2 + |V_{ts}|^2 = 1$ . By imposing strictly  $|V_{tb}|^2 + |V_{td}|^2 + |V_{ts}|^2 = 1$ , i.e. a pure SM scenario in Eq. 46, the fit is repeated by considering a single parameter,  $|V_{tb}|$ , and obtaining  $|V_{td}|^2 + |V_{ts}|^2$  as  $1 - |V_{tb}|^2$ .

$$\mu_b = \frac{|V_{tb}|^4 \text{ obs}}{|V_{tb}|^4}$$

$$\mu_{sd} = \frac{|V_{tb}|^2 \text{ obs} \cdot (1 - |V_{tb}|^2 \text{ obs})}{|V_{tb}|^2 \cdot (1 - |V_{tb}|^2)}$$
(46)

The results in this case are:

$$\begin{aligned} |V_{tb}| &= 0.980^{+0.014}_{-0.011}(\text{stat.} + \text{syst.}) \pm 0.031(\text{externalized}) \\ |V_{td}|^2 + |V_{ts}|^2 &= 0.040^{+0.023}_{-0.028}(\text{stat.} + \text{syst.}) \pm 0.059(\text{externalized}) \end{aligned} \quad (47)$$

### 7.3 Measurement in BSM scenarios

Any new physics contribution potentially enhancing the  $|V_{tb}|^2$ ,  $|V_{ts}|^2$ , or  $|V_{td}|^2$  can act either in production, in decay, or in both.

New physics could foresee this extension of the CKM matrix to new families would imply the SM part not to be unitary. If quarks from these new families are heavier than the top quark, new physics would therefore alter the CKM matrix element without appearing as top quark decay product, thus not counting for the top quark decay width  $\Gamma_{\text{top}}$  directly. The effect on the BR would be a reduced absolute value of all CKM matrix elements.

A first interpretation we consider does assume the top quark does decay through the same channels as in the SM case, and that the partial widths of each decay do only vary because of a modified CKM matrix element. This leads to considering  $R_q$  as a function of both  $|V_{tb}|^2$  and  $|V_{td}|^2 + |V_{ts}|^2$ .

This means Eq. 43 becomes:

$$\begin{aligned}\mu_b &= \frac{|V_{tb}|^{4\text{obs}}}{|V_{tb}|^4 \cdot (|V_{tb}|^{2\text{obs}} + |V_{ts}|^{2\text{obs}} + |V_{td}|^{2\text{obs}})} \\ \mu_{sq} &= \frac{|V_{tb}|^{2\text{obs}} \cdot (|V_{ts}|^{2\text{obs}} + |V_{td}|^{2\text{obs}})}{(|V_{ts}|^2 + |V_{td}|^2) \cdot (|V_{tb}|^{2\text{obs}} + |V_{ts}|^{2\text{obs}} + |V_{td}|^{2\text{obs}})}\end{aligned}\quad (48)$$

Also in this case a new fit is performed in order to extract  $|V_{tb}|$  and  $|V_{td}|^2 + |V_{ts}|^2$ , resulting in:

$$\begin{aligned}|V_{tb}| &= 0.988 \pm 0.027(\text{stat.} + \text{syst.}) \pm 0.043(\text{externalized}) \\ |V_{td}|^2 + |V_{ts}|^2 &= 0.06 \pm (0.05)(\text{stat.} + \text{syst.})^{+0.04}_{-0.03}(\text{externalized})\end{aligned}\quad (49)$$

Finally, we consider a case for which new physics would leave the partial width of the top quark decaying to  $Wb$  or  $Wq$  unchanged, but modifies the overall total width of the top quark either via additional, undetected, decays. This means that we assume the partial width of known quarks to be unchanged, but the total width of the quark top is free to change within its known experimental boundaries [35, 36]. In this approximation, we consider negligible effects of variations of  $|V_{tb}|^2$ ,  $|V_{td}|^2$ , and  $|V_{ts}|^2$  on  $\Gamma_{\text{top}}$ .

We do therefore modify the Eq. 43 by considering a generic width for the top quark, obtaining:

$$\begin{aligned}\mu_b &= \frac{|V_{tb}|^{4\text{obs}} \cdot \Gamma_{\text{top}}}{|V_{tb}|^4 \cdot \Gamma_{\text{top}}^{\text{obs}}} \\ \mu_{sd} &= \frac{|V_{tb}|^{2\text{obs}} \cdot (|V_{ts}|^{2\text{obs}} + |V_{td}|^{2\text{obs}}) \cdot |V_{tb}|^4 \cdot \Gamma_{\text{top}}}{(|V_{ts}|^2 + |V_{td}|^2) \cdot \Gamma_{\text{top}}^{\text{obs}}}\end{aligned}\quad (50)$$



666 Which results in the following measurement:

$$\begin{aligned}
 |V_{tb}| &= 0.988 \pm 0.011(\text{stat.} + \text{syst.}) \pm 0.021(\text{externalized}) \\
 |V_{td}|^2 + |V_{ts}|^2 &= 0.06 \pm (0.05)(\text{stat.} + \text{syst.}) \pm 0.04(\text{externalized}) \\
 \Gamma_{top}^{\text{obs}} &= 0.99 \pm 0.42(\text{stat.} + \text{syst.}) \pm 0.03(\text{externalized.}) \cdot \Gamma_{top}
 \end{aligned} \tag{51}$$

667 It is worthy of note that in the SM interpretation, both measurements are strongly limited by  
 668 externalized systematic uncertainties, as their uncertainties are of course fully anti-correlated  
 669 and the  $ST_{(b,b)}^{2J1T}$  modeling uncertainties dominate given the relative abundance of such signal  
 670 with respect to the  $ST_{(q,b)}^{2J1T}$  and  $ST_{(b,q)}^{2J1T}$ . In this last case, given the absence of a correlated term in  
 671 the ratio, the statistical precision of the  $|V_{tb}|$  determination is close to the one in Eq. 47.

## 672 8 Summary

The first direct measurement of the Cabibbo-Kobayashi-Maskawa matrix elements  $|V_{tb}|$ ,  $|V_{ts}|$ , and  $|V_{td}|$ , making use of single top quark  $t$ -ch. events in proton-proton collision data with a centre-of-mass energy of 13 TeV, has been presented. The subset of data analysed corresponds to an integrated luminosity of  $35.89 \text{ fb}^{-1}$ .

The dominant electroweak production mechanism for single top quarks is the  $t$ -ch. and it features a  $tWq$  vertex, where  $q$  stands for  $b$ ,  $s$  or  $d$  quarks, both in production and in decay. For this reason its cross section and branching fractions are sensitive to the strength of the electroweak coupling, making it a suitable channel for direct measurements of the magnitude of Cabibbo-Kobayashi-Maskawa matrix elements  $|V_{tb}|$ ,  $|V_{ts}|$ , and  $|V_{td}|$ . A precise determination of the magnitude of these parameters of the Standard Model allows to search for hints of potential contributions from new physics beyond the Standard Model.

In the present analysis, a selection is performed by requiring one isolated lepton, muon or electron and two or three jets, one or two of which must pass the  $b$ -tagging requirement. The events are distinguished into three different categories according to the number of jets and  $b$  tagged jets. The signal is composed of single top  $t$ -ch. produced via or decaying to  $s/d$  quarks. A data driven estimation technique is applied to obtain the QCD-multijets contribution directly from data.

For a further optimisation of the discrimination between signal and backgrounds the data set is divided into two regions, characterized by different event topologies for the categories with one  $b$ -jet. In each of these regions a multivariate analysis is performed in order to fully exploit the most important kinematic variables.

A simultaneous maximum likelihood fit to the multivariate distributions in all categories is performed in order to extract  $|V_{tb}|^2$  and  $|V_{ts}|^2 + |V_{td}|^2$ .

The results obtained with the analysis presented are:

$$|V_{tb}| = 1.00 \pm 0.01 \pm 0.03 \quad |V_{td}|^2 + |V_{ts}|^2 \leq 0.17 \text{ at } 95\% \text{ C.L.,}$$

673 and no significant deviation from the prior Standard Model expectations is observed.

## References

- [1] CMS Collaboration, “Measurement of the single-top-quark t-channel cross section in pp collisions at  $\sqrt{s} = 7$  tev”, *JHEP* **12** (2012) 035, doi:10.1007/JHEP12(2012)035.
- [2] CMS Collaboration, “Measurement of the t-channel single-top-quark production cross section and of the  $|V_{tb}|$  ckm matrix element in pp collisions at smeasurement of the single-top-quark t-channel cross section in pp collisions at  $\sqrt{s} = 8$  tev”, *JHEP* **06** (2014) 090, doi:10.1007/JHEP06(2014)090.
- [3] CMS Collaboration, “Measurement of the single-top-quark t-channel cross section in pp collisions at  $\sqrt{s} = 13$  tev”,.
- [4] N. Kidonakis, “Differential and total cross sections for top pair and single top production”, arXiv:1205.3453.
- [5] <https://cms-service-dqm.web.cern.ch/cms-service-dqm/CAF/certification/>.
- [6] J. Alwall et al., “MadGraph 5: Going Beyond”, *JHEP* **06** (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522.
- [7] T. Sjöstrand, S. Mrenna, and P. Skands, “PYTHIA 6.4 physics and manual”, *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.
- [8] P. Skands, S. Carrazza, and J. Rojo, “Tuning PYTHIA 8.1: the Monash 2013 Tune”, *Eur. Phys. J.* **C74** (2014), no. 8, 3024, doi:10.1140/epjc/s10052-014-3024-y, arXiv:1404.5630.
- [9] CMS Collaboration Collaboration, “Investigations of the impact of the parton shower tuning in Pythia 8 in the modelling of tt at  $\sqrt{s} = 8$  and 13 TeV”, Technical Report CMS-PAS-TOP-16-021, CERN, Geneva, 2016.
- [10] R. Frederix and S. Frixione, “Merging meets matching in MC@NLO”, *JHEP* **12** (2012) 061, doi:10.1007/JHEP12(2012)061, arXiv:1209.6215.
- [11] NNPDF Collaboration, “Parton distributions for the LHC Run II”, *JHEP* **04** (2015) 040, doi:10.1007/JHEP04(2015)040, arXiv:1410.8849.
- [12] J. Allison et al., “Geant4 developments and applications”, *IEEE Transactions on Nuclear Science* **53**, 1 (2006) 270–278.
- [13] TOPLHCWG, “Nlo single-top t channel cross sections”, <https://twiki.cern.ch/twiki/bin/view/LHCPhysics/SingleTopRefXsec>.
- [14] CMS Collaboration, “Particle-flow event reconstruction in CMS and performance for jets, taus, and  $\tilde{E}_T$ ”, CMS Physics Analysis Summary CMS-PAS-PFT-09-001, 2009.
- [15] CMS Collaboration, “Description and performance of track and primary-vertex reconstruction with the CMS tracker”, *JINST* **9** (2014), no. 10, P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.
- [16] M. Cacciari, G. P. Salam, and G. Soyez, “The anti- $k_t$  jet clustering algorithm”, *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:0802.1189.
- [17] CMS Collaboration, “Pileup Removal Algorithms”, CMS Physics Analysis Summary CMS-PAS-JME-14-001, 2014.

- [18] T. C. collaboration, "Identification of b-quark jets with the cms experiment", *Journal of Instrumentation* **8** (2013), no. 04, P04013.
- [19] CMS Collaboration, "Identification of b quark jets at the CMS Experiment in the LHC Run 2", CMS Physics Analysis Summary CMS-PAS-BTV-15-001, CERN, 2015.
- [20] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation80XReReco>.
- [21] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018), no. 05, P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.
- [22] <https://twiki.cern.ch/twiki/bin/view/CMS/BTagShapeCalibration>.
- [23] CMS Collaboration, "Performance of missing energy reconstruction in 13 TeV pp collision data using the CMS detector", CMS Physics Analysis Summary CMS-PAS-JME-16-004, 2016.
- [24] D0 Collaboration, "Observation of Single Top-Quark Production", *Phys. Rev. Lett.* **103** (2009) 092001, doi:10.1103/PhysRevLett.103.092001, arXiv:0903.0850.
- [25] CDF Collaboration, "First Observation of Electroweak Single Top Quark Production", *Phys. Rev. Lett.* **103** (2009) 092002, doi:10.1103/PhysRevLett.103.092002, arXiv:0903.0885.
- [26] Particle Data Group Collaboration, "Review of Particle Physics", *Chin. Phys.* **C38** (2014) 090001, doi:10.1088/1674-1137/38/9/090001.
- [27] Petrucciani, G. and others, "Documentation of the roostats-based statistics tools for higgs pag", twiki, 2012.
- [28] L. Moneta et al., "The RooStats Project", in *13<sup>th</sup> International Workshop on Advanced Computing and Analysis Techniques in Physics Research (ACAT2010)*. SISSA, 2010. arXiv:1009.1003. PoS(ACAT2010)057.
- [29] R. J. Barlow and C. Beeston, "Fitting using finite Monte Carlo samples", *Comput.Phys.Commun.* **77** (1993) 219–228, doi:10.1016/0010-4655(93)90005-W.
- [30] J. S. Conway, "Nuisance parameters in likelihoods for multisource spectra", in *Proceedings of PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding*, H. Propser and L. Lyons, eds., number CERN-2011-006, pp. 115–120. CERN, 2011.
- [31] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", *JINST* **6** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.
- [32] CMS Collaboration, "Measurements of inclusive W and Z Cross Sections in pp Collisions at  $\sqrt{s} = 7$  TeV", *JHEP* **01** (2011) 080, doi:10.1007/JHEP01(2011)080, arXiv:1012.2466.
- [33] M. Botje et al., "The PDF4LHC Working Group interim recommendations", (2011). arXiv:1101.0538.
- [34] CMS Collaboration, "Cms luminosity based on pixel cluster counting — summer 2013 update", CMS Physics Analysis Summary CMS-PAS-LUM-13-001, 2013.

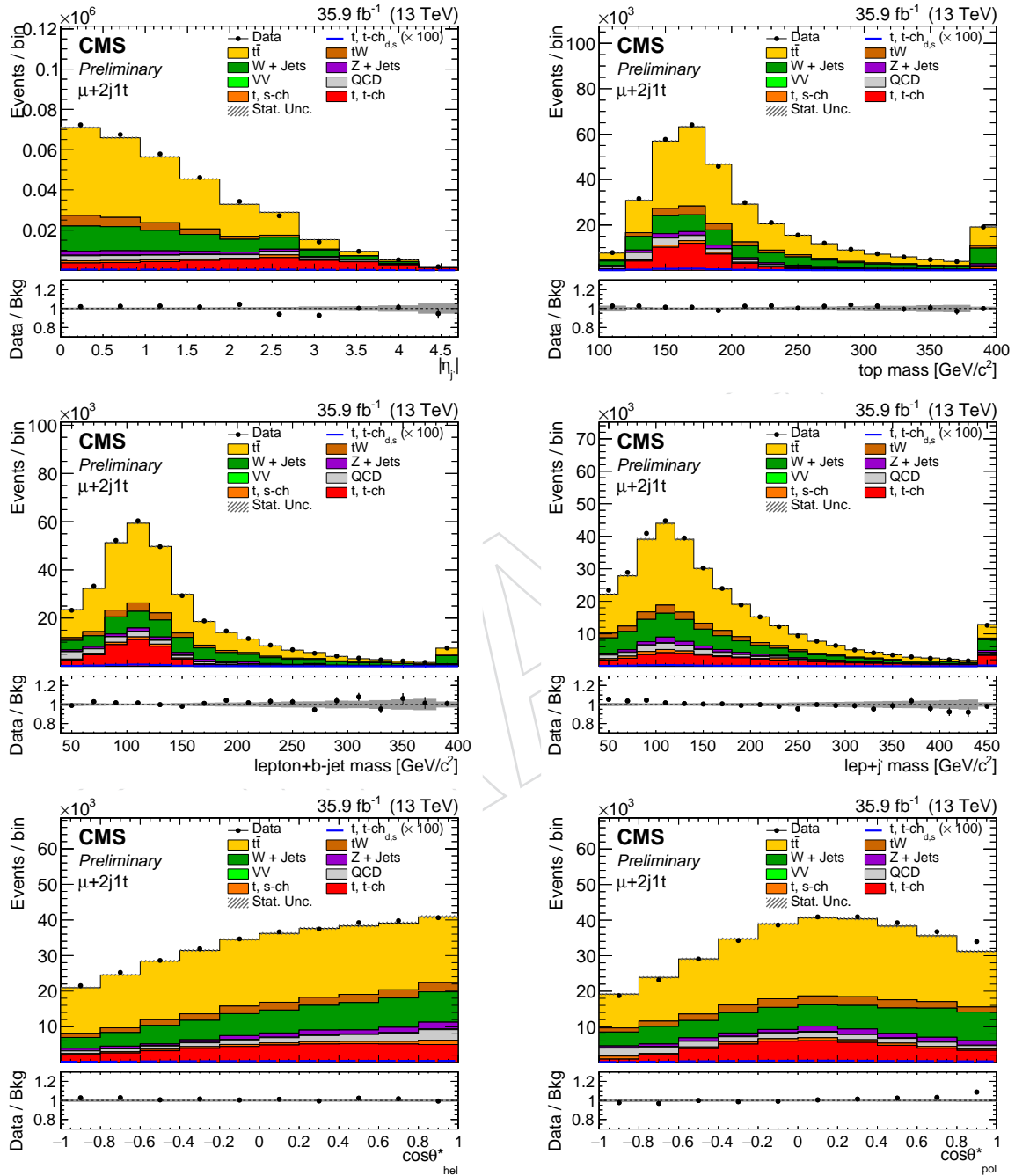
- [35] CMS Collaboration Collaboration, “Bounding the top quark width using final states with two charged leptons and two jets at  $\sqrt{s} = 13$  TeV”, Technical Report CMS-PAS-TOP-16-019, CERN, Geneva, 2016.
- [36] ATLAS Collaboration Collaboration, “Measurement of the top-quark decay width in top-quark pair events in the dilepton channel at  $\sqrt{s} = 13$  TeV with the ATLAS detector”, Technical Report ATLAS-CONF-2019-038, CERN, Geneva, Aug, 2019.

DRAFT

## A BDT input variables for the muon-channel

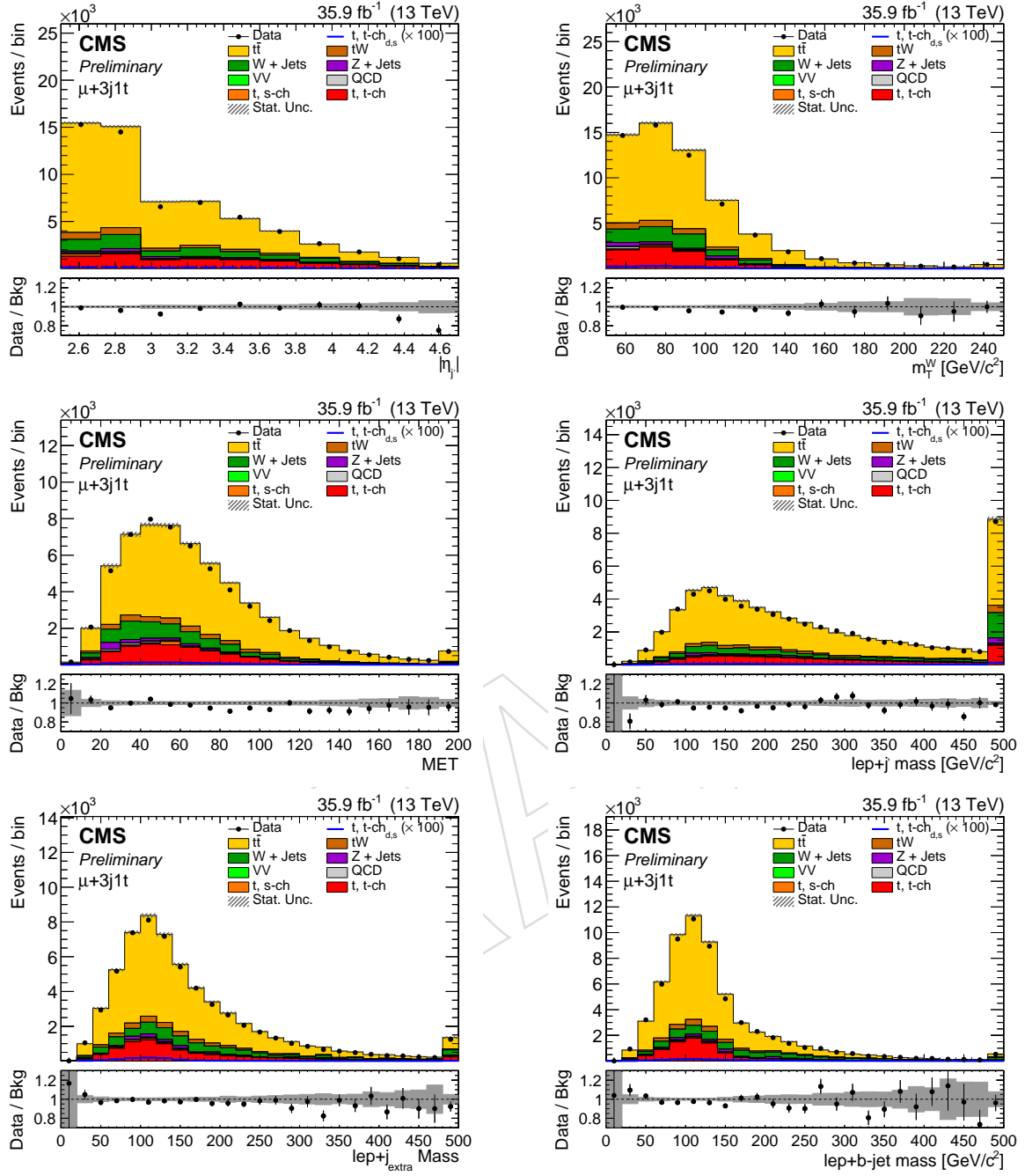
The distributions of all input variables used for the training of the BDTs in the muon-channel are shown in this section. All MC samples are normalized to the expected cross section.

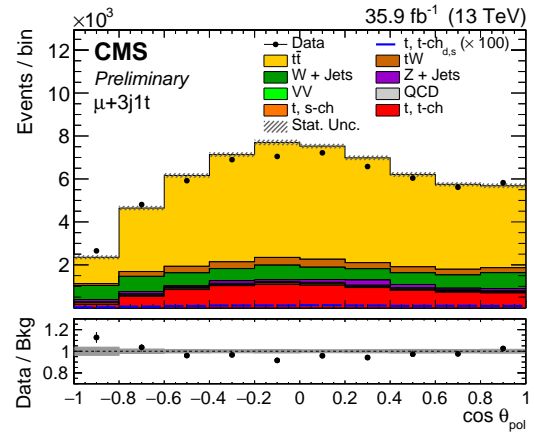
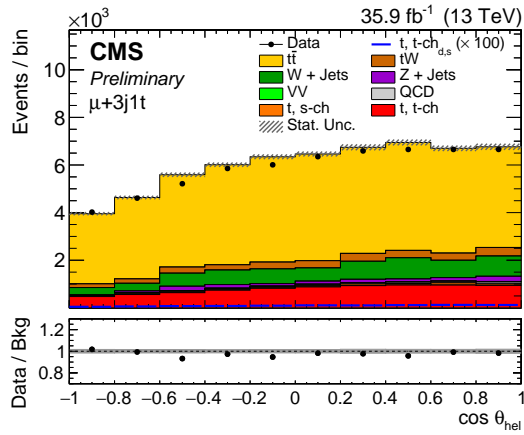
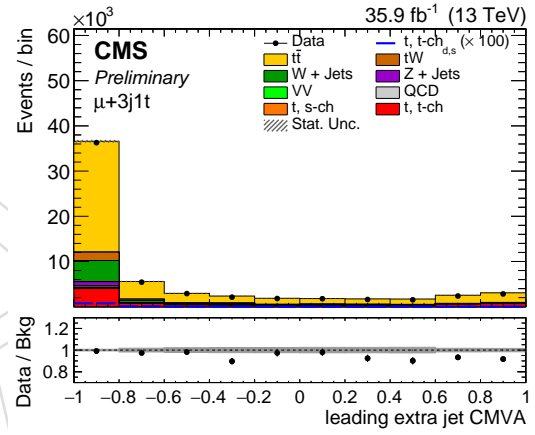
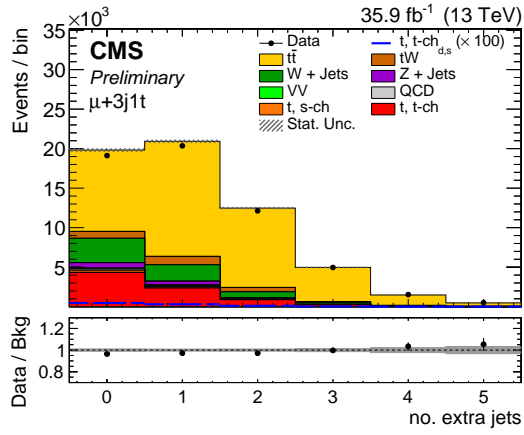
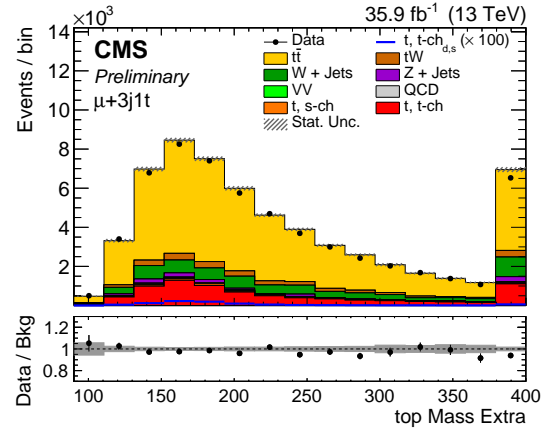
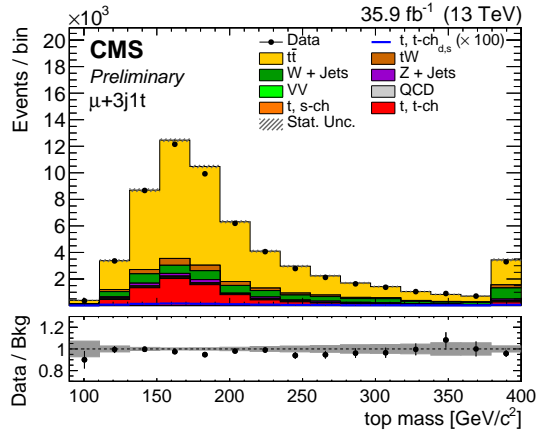
### A.1 Variables for the 2j1t



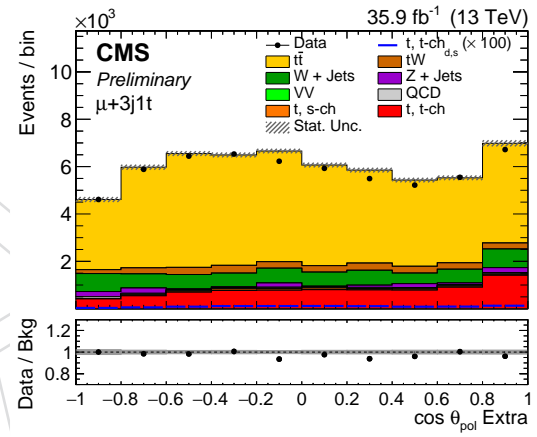
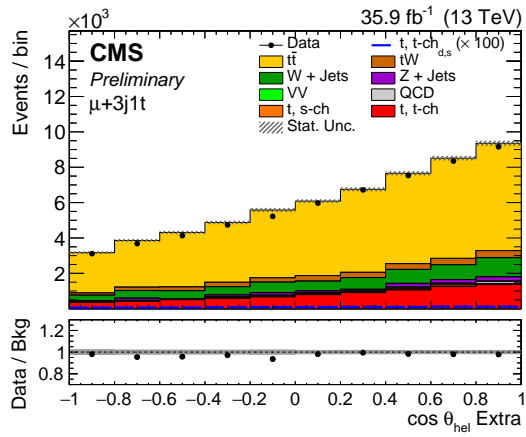
763

## A.2 Variables for the 3j1t

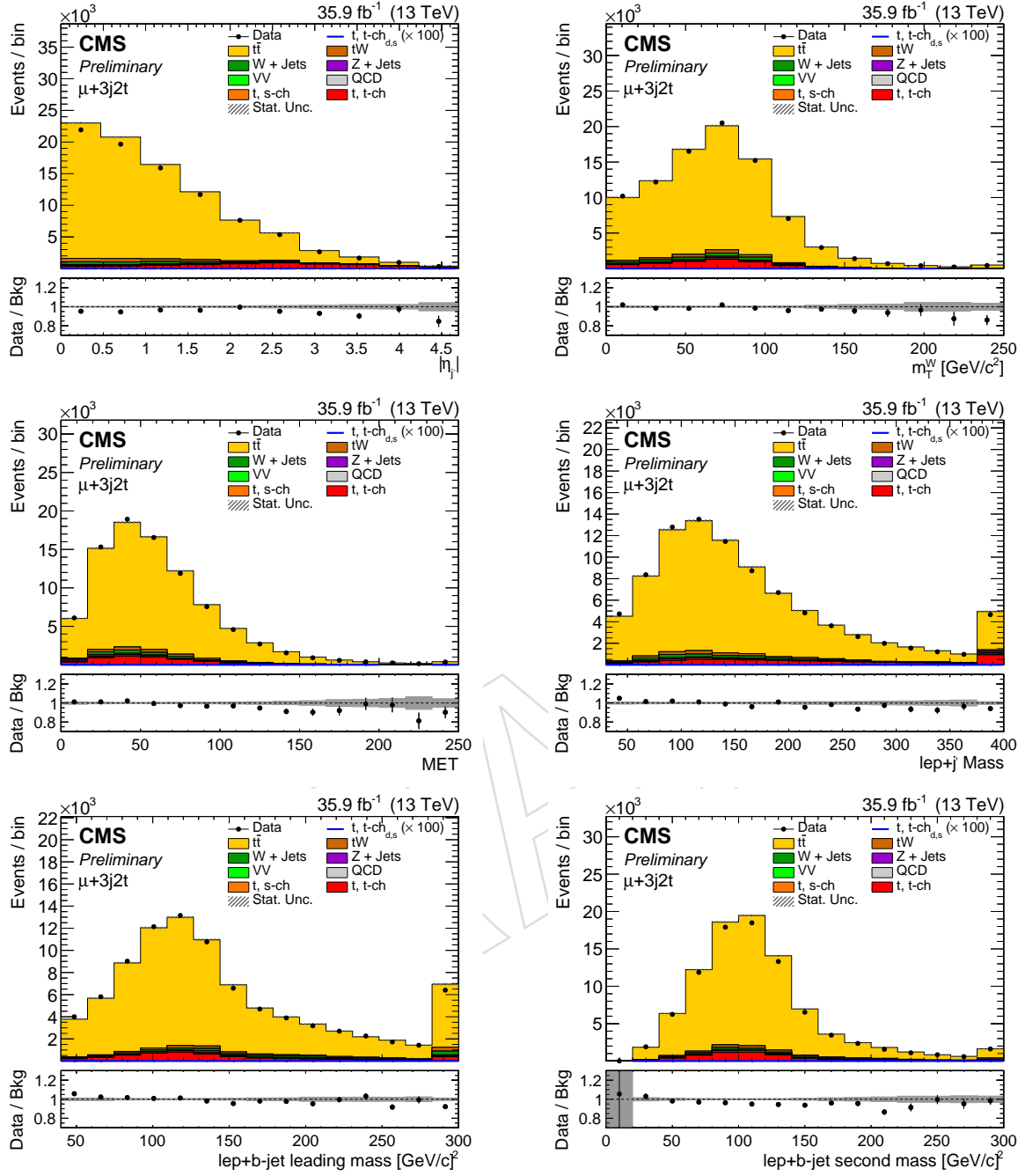


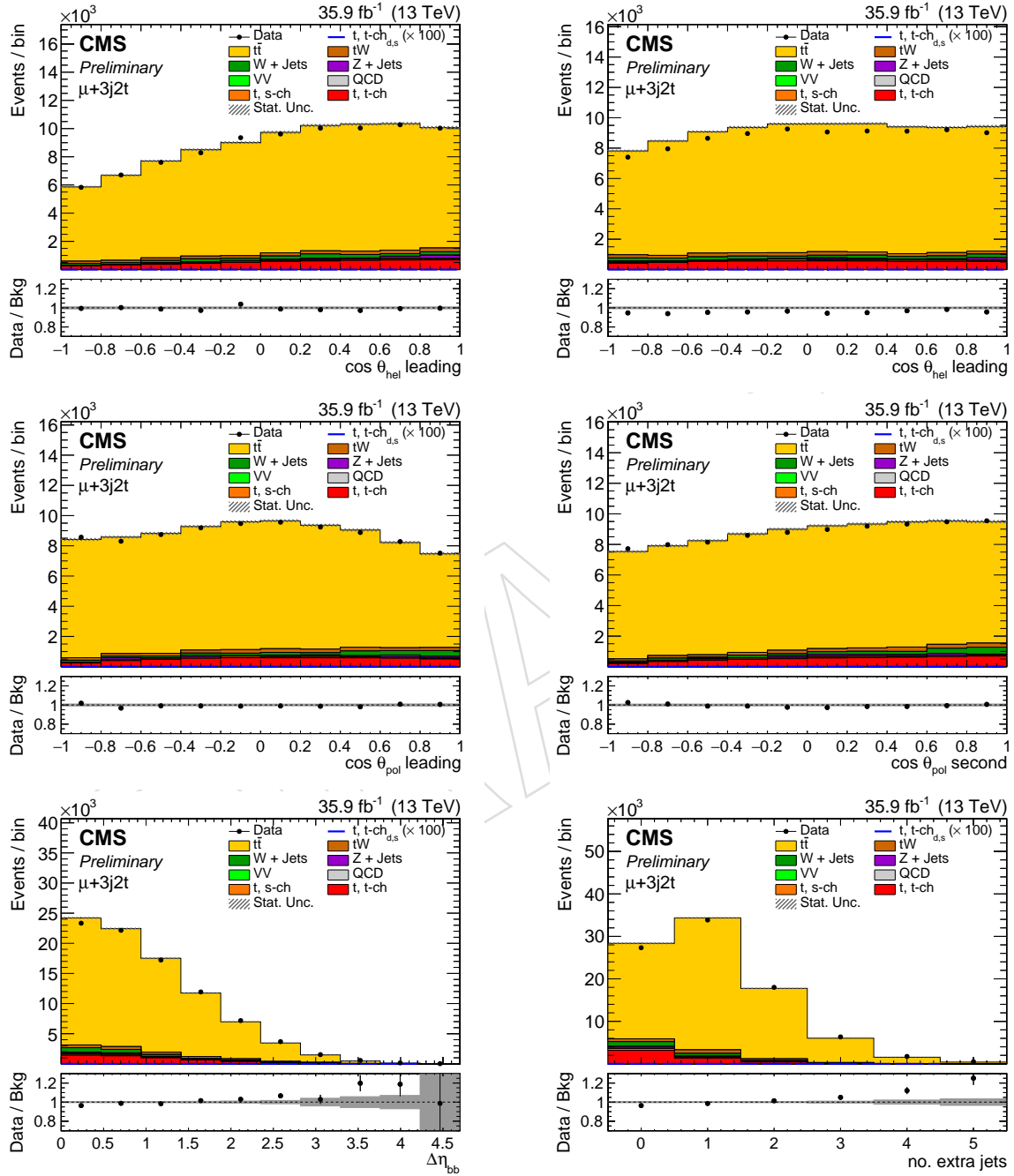


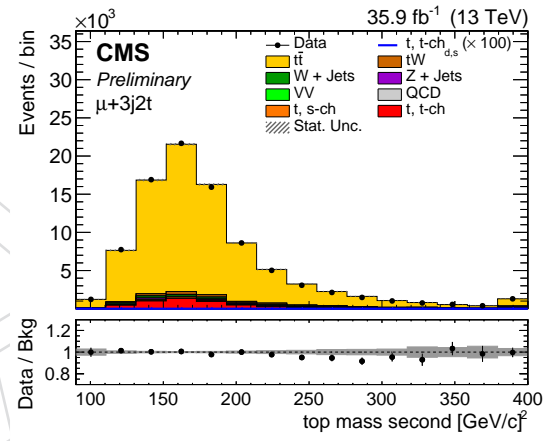
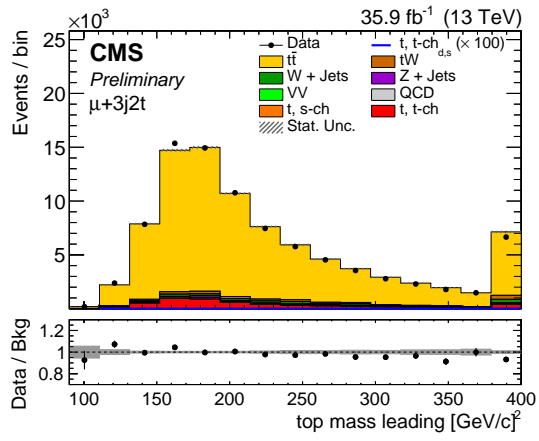




### A.3 Variables for the 3j2t



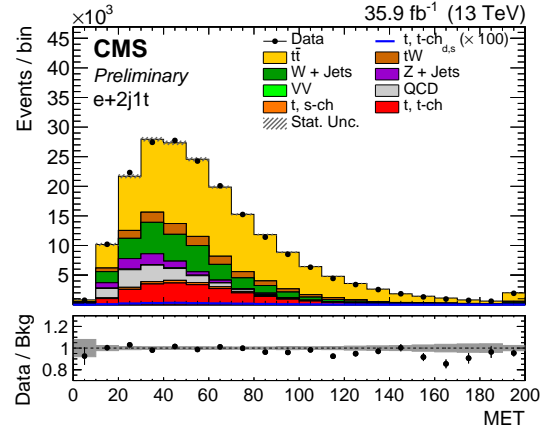
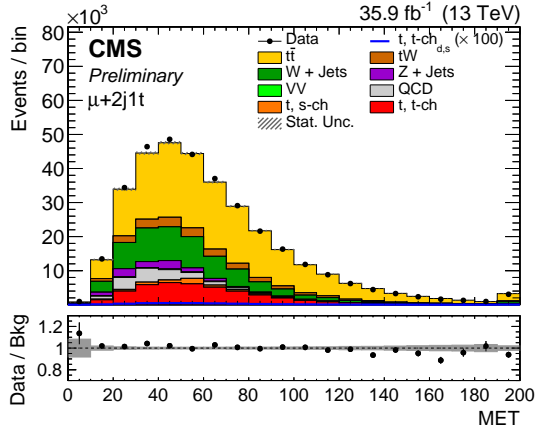
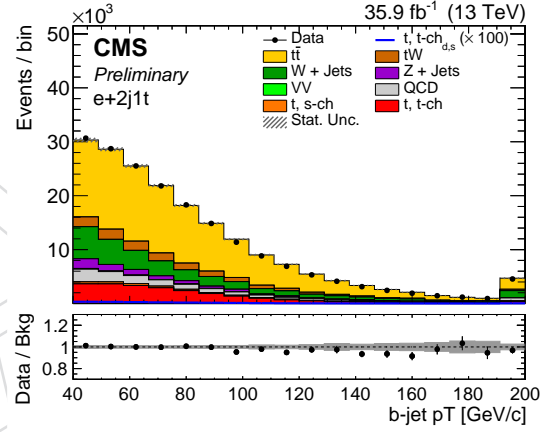
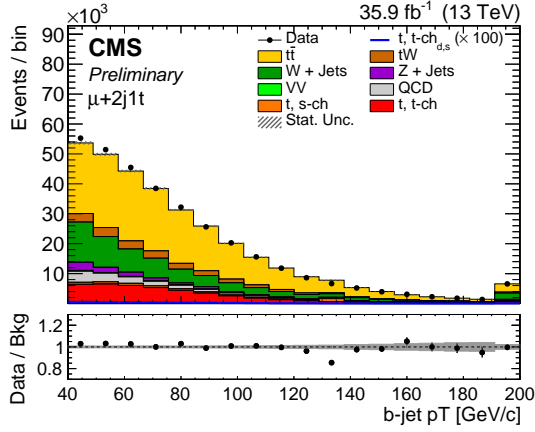
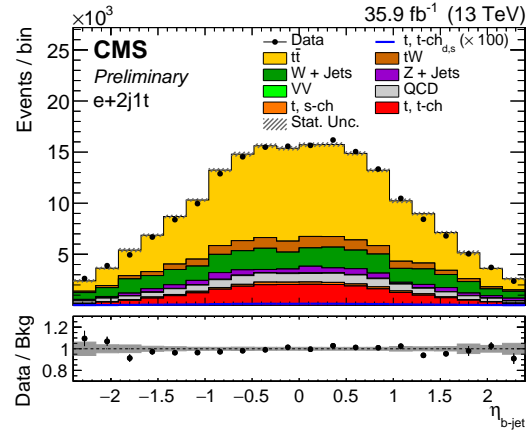
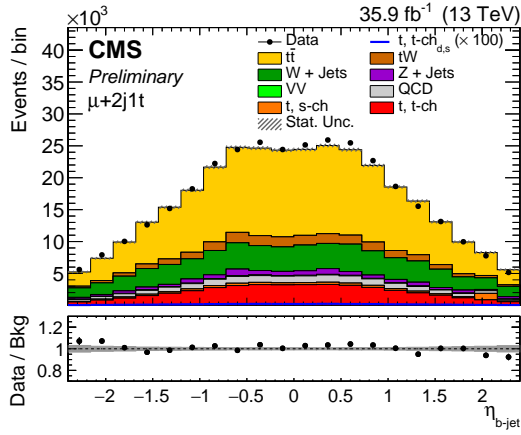


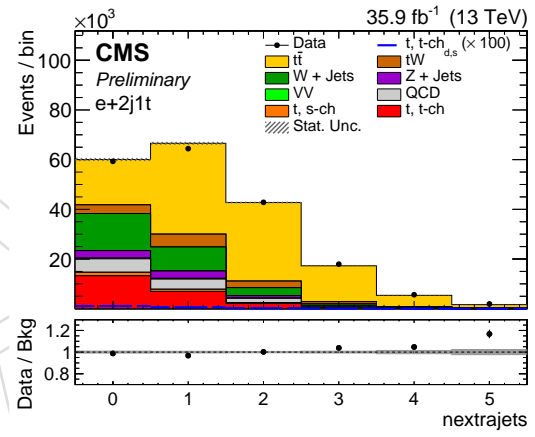
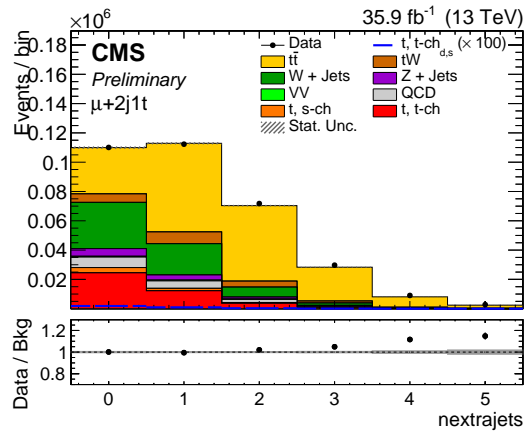


## B Fundamental variables

The distributions of some fundamental variables for both muon and electron channel are shown in this section. All MC samples are normalized to the expected cross section.

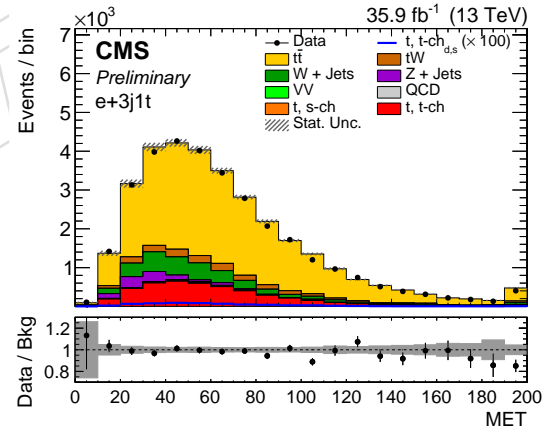
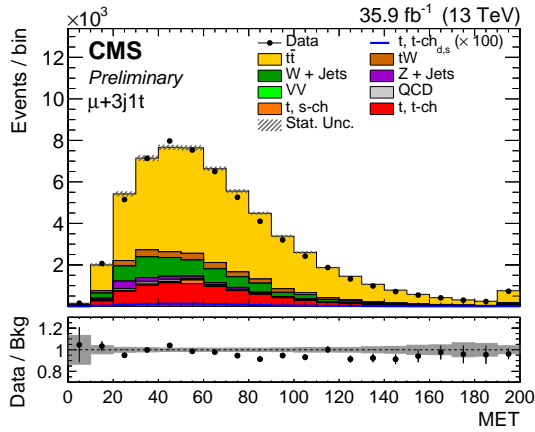
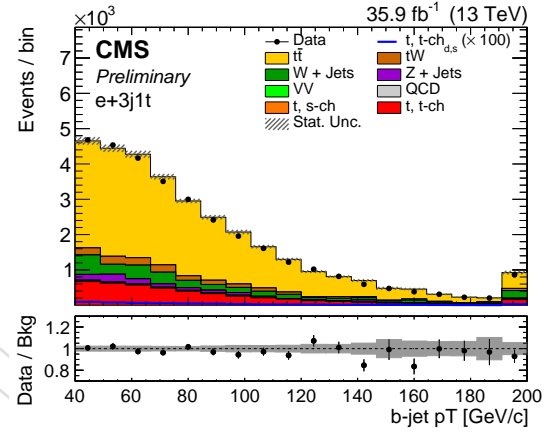
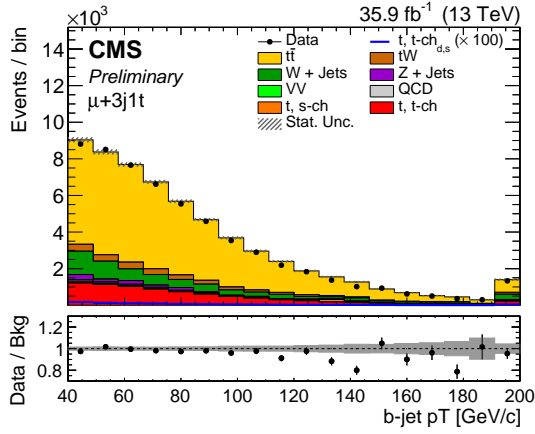
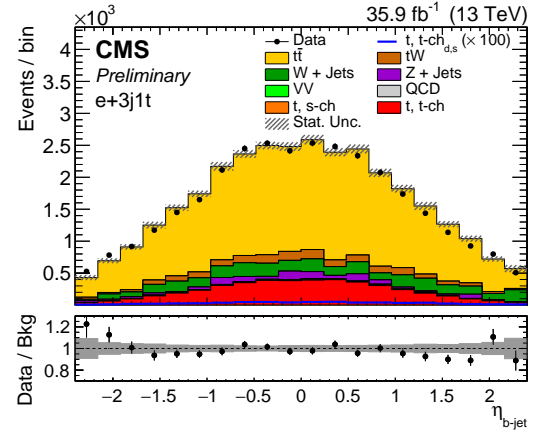
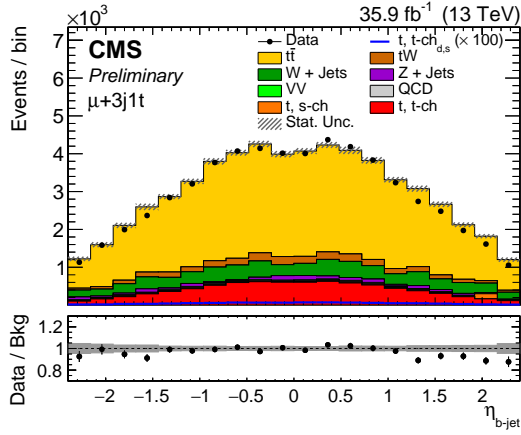
### B.1 Variables for the 2j1t



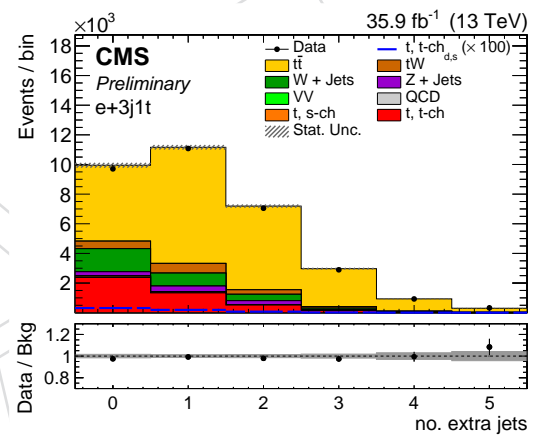
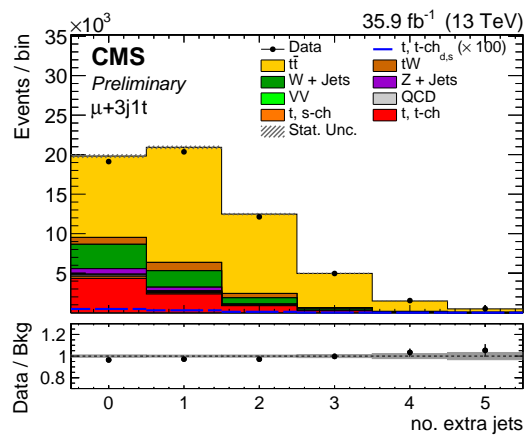


769

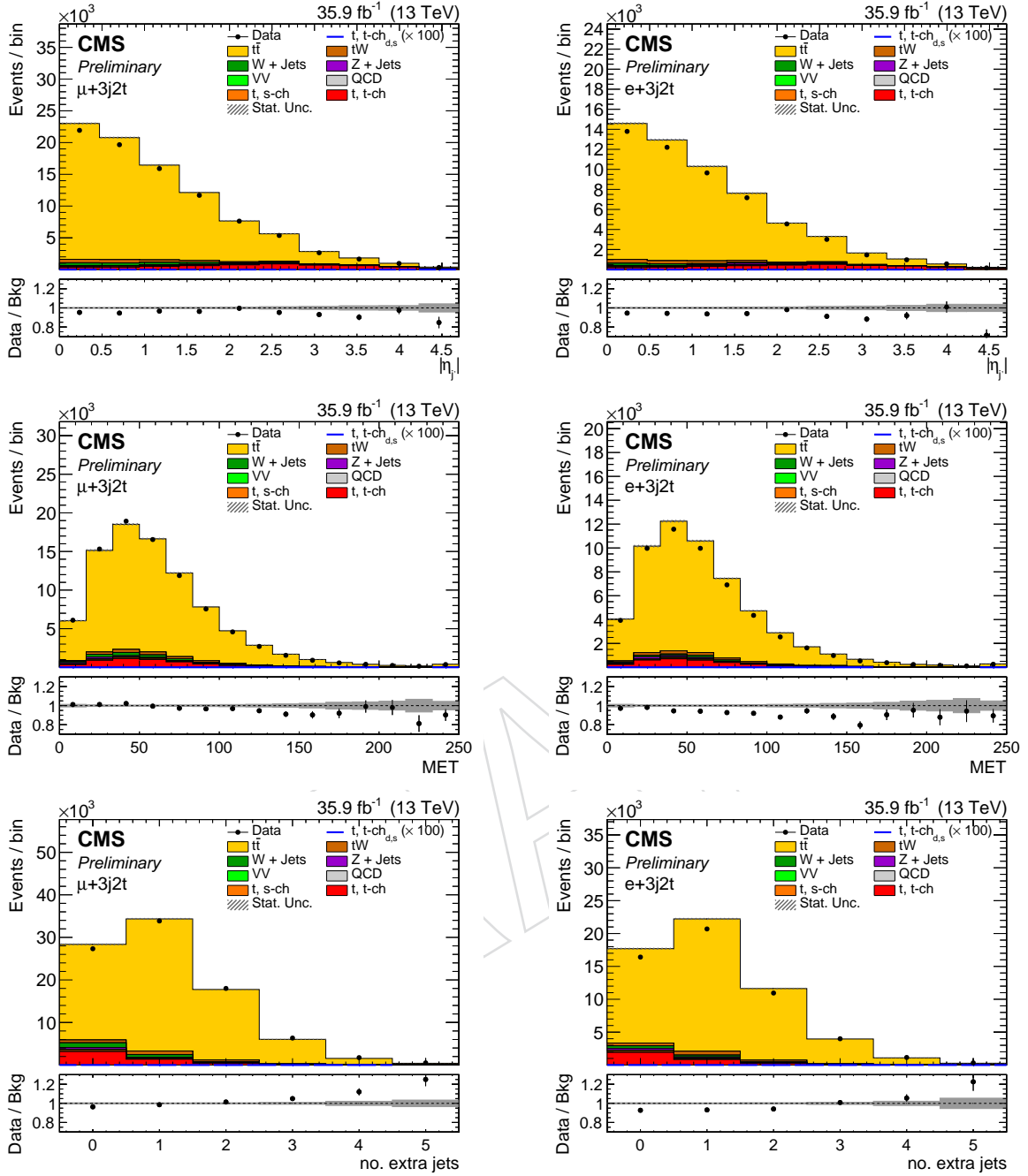
## B.2 Variables for the 3j1t







### B.3 Variables for the 3j2t



## C Fit in the single regions

The following Section presents the fits performed in each of the regions used for the analysis.

### C.1 Muon

For each region of the muon channel, the pulls and the results obtained are reported. The systematics included are just the profiled ones as reported in the Section 6.

### C.1.1 2-jets–1-tag

In this region the fit to the single-top  $t$ -ch. gives:

$$\mu_{ST_{b,b}} = 1.01 \pm 0.16 \text{ (prof.)}$$

Figure 24 shows the values of the pulls with pseudo-data for any nuisance parameter for this region.

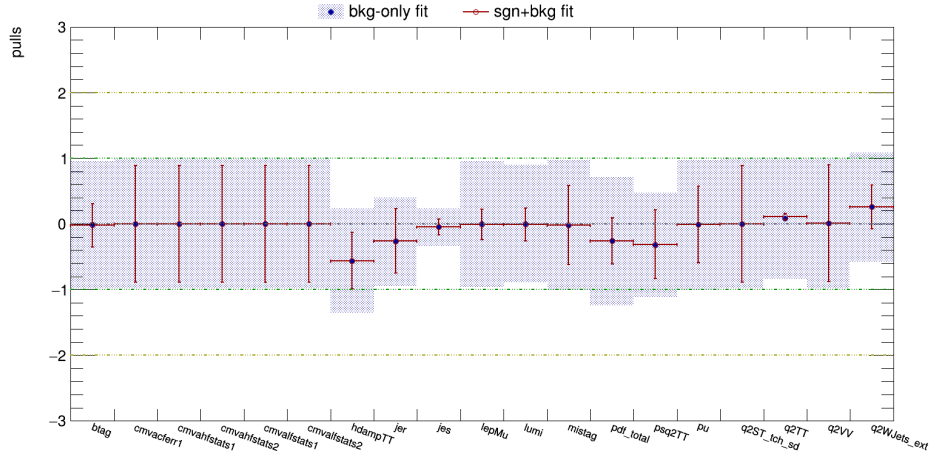


Figure 24: Pulls with pseudo-data for each nuisance parameter for the 2-jets–1-tag region in the muon channel.

On this single region no particular over-constraints are observed, except for the JECs and the hard scattering scale for the  $t\bar{t}$  sample.

### C.1.2 3-jets–1-tag

In this region the fit to the single-top  $t$ -ch. gives:

$$\mu_{ST_{b,b}} = 0.95 \pm 0.35 \text{ (prof.)}$$

Figure 25 shows the values of the pulls with pseudo-data for any nuisance parameter for this region.

On this single region no particular over-constraints are observed, except for the systematics related to the CMVAv2 tagger algorithm.

### C.1.3 3-jets–2-tags

In this region the fit to the single-top  $t$ -ch. gives:

$$\mu_{ST_{b,b}} = 0.80 \pm 0.16 \text{ (prof.)}$$

Figure 26 shows the values of the pulls with pseudo-data for any nuisance parameter for this region.

On this single region no particular over-constraints are observed, except for the JECs and the  $t\bar{t}$  modelling uncertainties.

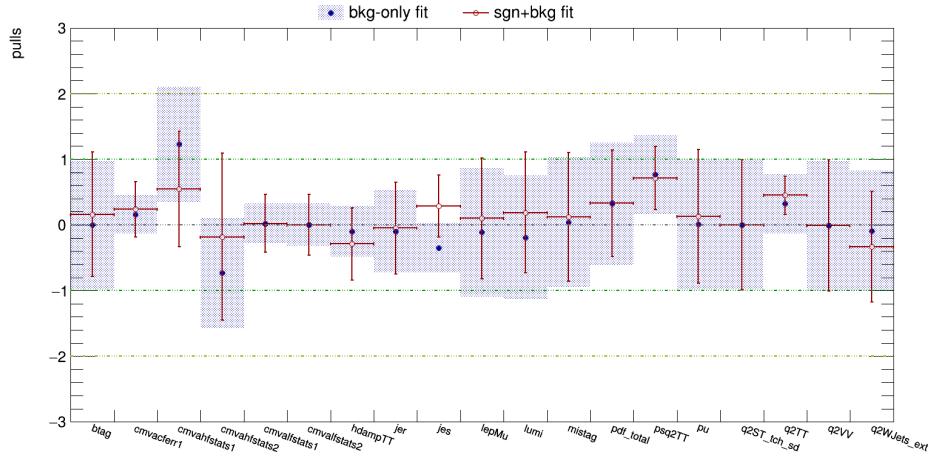


Figure 25: Pulls with pseudo-data for each nuisance parameter for the 3-jets-1-tag region in the muon channel.

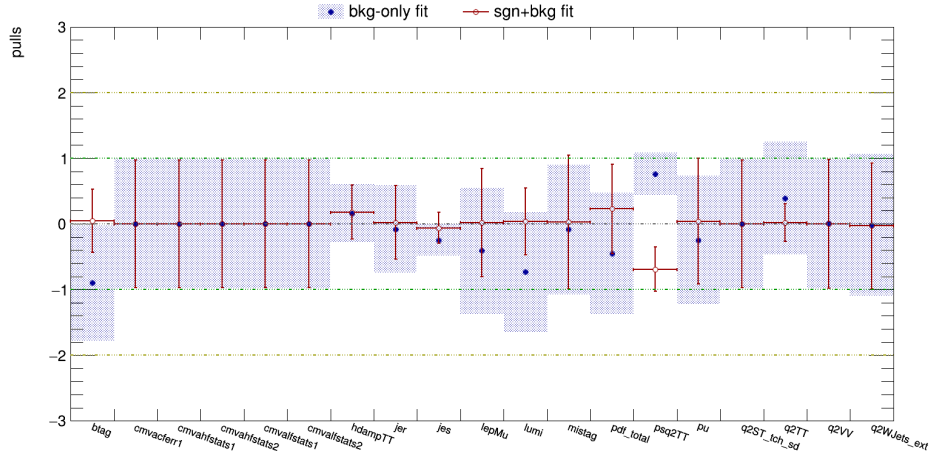


Figure 26: Pulls with pseudo-data for each nuisance parameter for the 3-jets-2-tags region in the muon channel.

#### C.1.4 Combined

By combining all the previous region reported together for the muon channel, the fit to the single-top  $t$ -ch, gives:

$$\mu_{ST_{b,b}} = 0.97 \pm 0.09 \text{ (prof.)}$$

Figure 27 shows the values of the pulls with pseudo-data for any nuisance parameter for the combination of the three regions in the muon channel.

When combining the three regions more constraints appear.

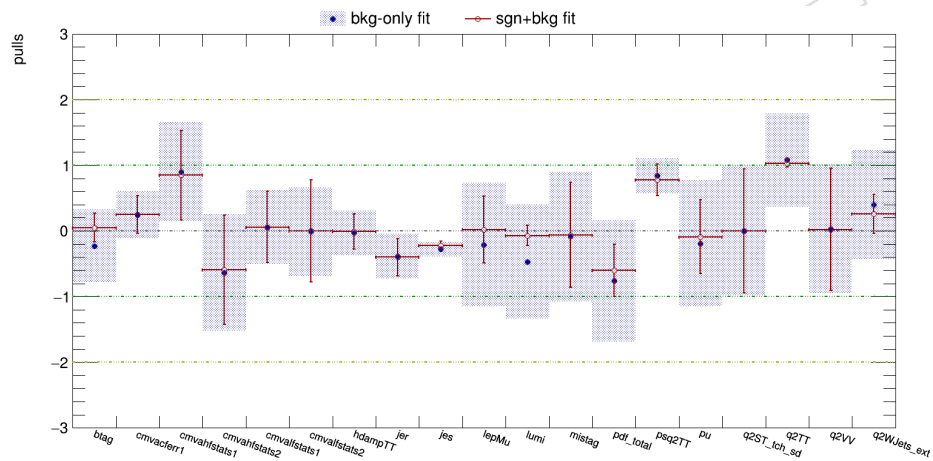


Figure 27: Pulls with pseudo-data for each nuisance parameter for the combination of the three regions in the muon channel.

## C.2 Electron

For each region of the electron channel, the pulls and the results obtained are reported. The systematics included are just the profiled ones as reported in the Section 6.

### C.2.1 2-jets–1-tag

In this region the fit to the single-top  $t$ -ch. gives:

$$\mu_{ST_{b,b}} = 0.94 \pm 0.18 \text{ (prof.)}$$

Figure 28 shows the values of the pulls with pseudo-data for any nuisance parameter for this region.

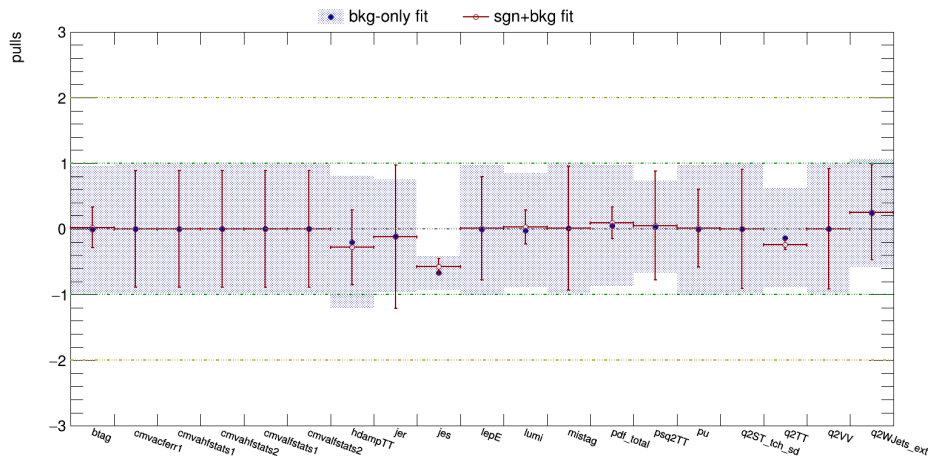


Figure 28: Pulls with pseudo-data for each nuisance parameter for the 2-jets–1-tag region in the electron channel.

On this single region no particular over-constraints are observed, except for the JECs and the hard scattering scale for the  $t\bar{t}$  sample.

### C.2.2 3-jets–1-tag

In this region the fit to the single-top  $t$ -ch. gives:

$$\mu_{ST_{b,b}} = 0.82 \pm 0.38 \text{ (prof.)}$$

Figure 29 shows the values of the pulls with pseudo-data for any nuisance parameter for this region.

On this single region no particular over-constraints are observed, except for the systematics related to the CMVA2 tagger algorithm.

### C.2.3 3-jets–2-tags

In this region the fit to the single-top  $t$ -ch. gives:

$$\mu_{ST_{b,b}} = 0.95 \pm 0.14 \text{ (prof.)}$$

Figure 30 shows the values of the pulls with pseudo-data for any nuisance parameter for this region.

On this single region no particular over-constraints are observed, except for the JECs and the  $t\bar{t}$  modelling uncertainties.

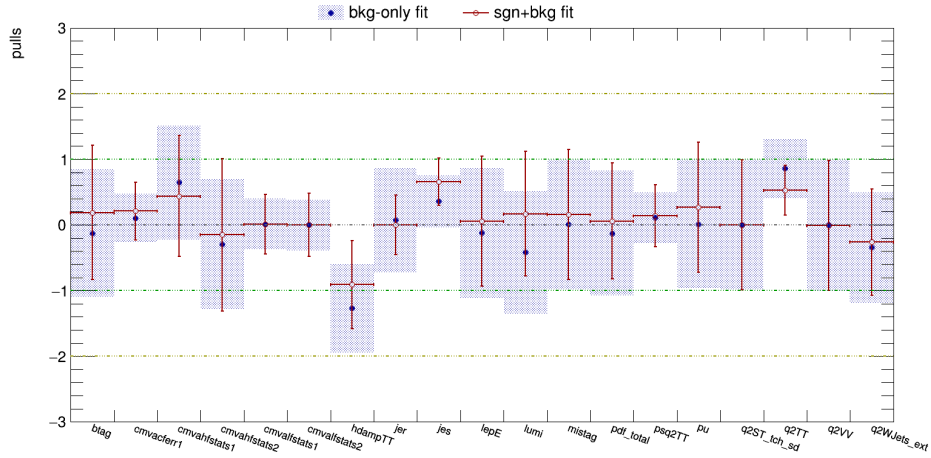


Figure 29: Pulls with pseudo-data for each nuisance parameter for the 3-jets-1-tag region in the electron channel.

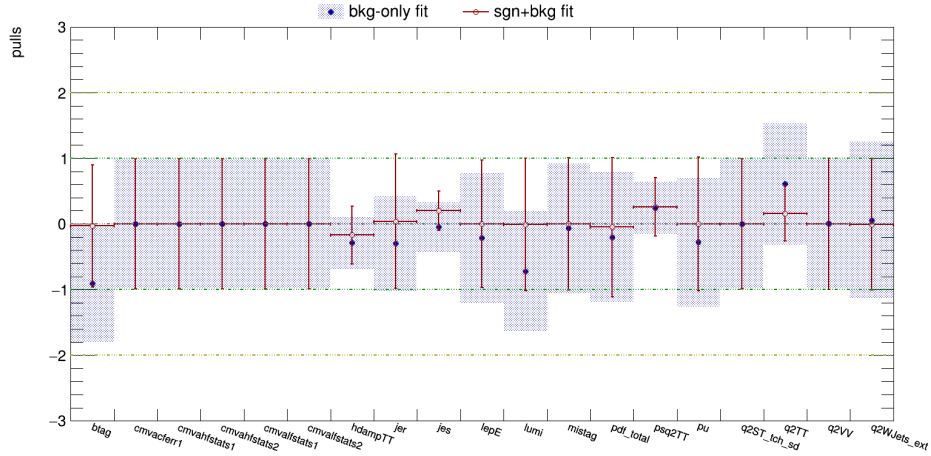


Figure 30: Pulls with pseudo-data for each nuisance parameter for the 3-jets-2-tags region in the electron channel.

## C.2.4 Combined

By combining all the previous region reported together for the electron channel, the fit to the single-top  $t$ -ch, gives:

$$\mu_{ST_{b,b}} = 1.00 \pm 0.10 \text{ (prof.)}$$

Figure 31 shows the values of the pulls with pseudo-data for any nuisance parameter for the combination of the three regions in the electron channel.

When combining the three regions more constraints appear.

All the previous results are summarized in Table 13.

## D Fit procedure checks

In this section a series of checks of the fit procedure and of the fit stability is reported.



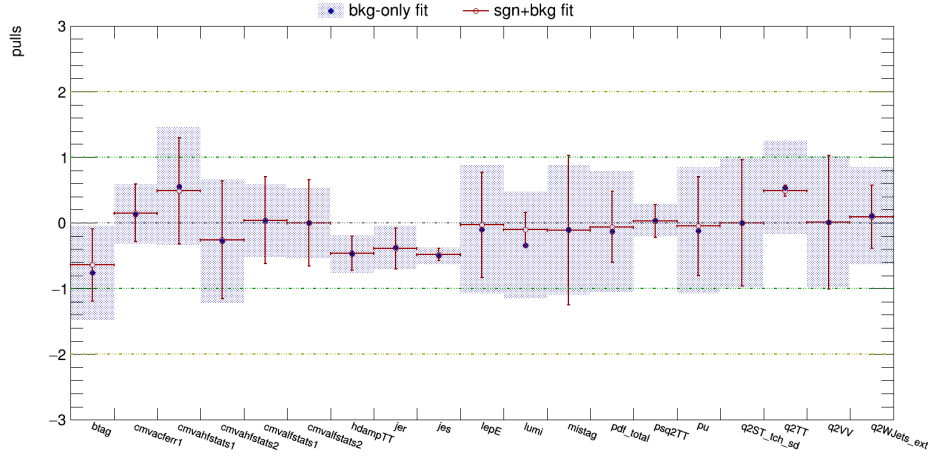


Figure 31: Pulls with pseudo-data for each nuisance parameter for the combination of the three regions in the electron channel.

Table 13: Single regions fit results in 2-jets-1-tag, 3-jets-1-tag, 3-jets-2-tags, and combination of the three region for both muon and electron channel.

Region	Fit result	
	Muon	Electron
2j1t	$1.01 \pm 0.16$	$0.94 \pm 0.18$
3j1t	$0.95 \pm 0.35$	$0.82 \pm 0.38$
3j2t	$0.80 \pm 0.16$	$0.95 \pm 0.14$
Combined	$0.97 \pm 0.09$	$1.00 \pm 0.10$

## D.1 $t\bar{t}$ systematics uncertainties externalization

In order to evaluate the relative importance of the  $t\bar{t}$  modelling systematic uncertainties in the total profiled uncertainty, the fit is repeated with and without these uncertainties in. The relative importance results to be of 1.04% and it has a marginal impact with respect to signal modelling uncertainties that are of order 20%. Figure 32 shows the pulls for the fit without the  $t\bar{t}$  modelling systematics.

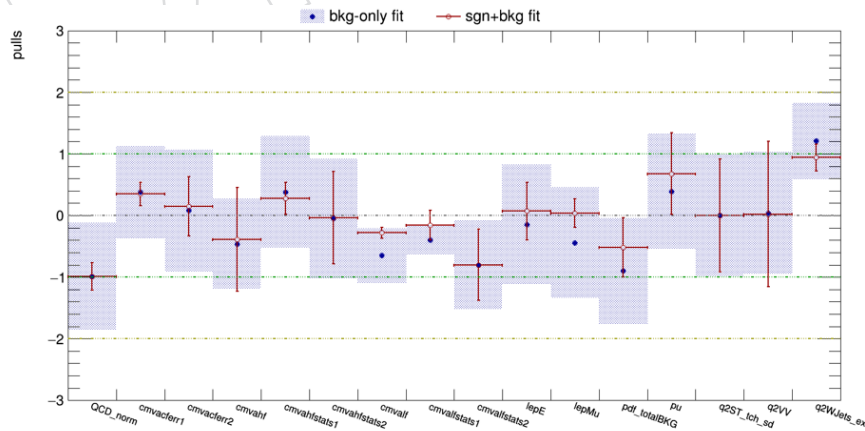


Figure 32: Impacts of the primary sources of systematic uncertainty treated as nuisance parameters in the fit of the BDT discriminator distributions (profiled uncertainties). Uncertainties are ranked by impact on the single top  $ST_{(b,b)}^{nImT}$  component, ranks 1-30 only shown in the plot.

## D.2 Impact of the JES on ttbar systematics

Looking at the plot reported in Figure 33 only for the FSR the correction is relevant ( $\approx 1\%$ ), and even by overestimating this one, the effect is negligible on the result.

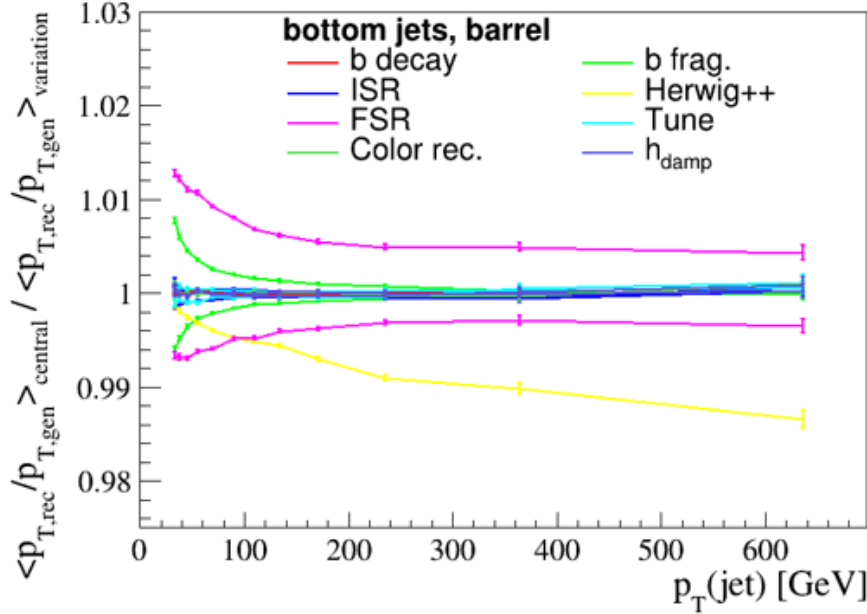


Figure 33: The figures show the differences in the jet energy response for various simulations. These MC specific correction factors are derived as the ratios of the means obtained by a Gaussian fit on the distributions of  $p_{T, reco} / p_{T, gen}$  for the standard Pythia8 MC and an alternative MC. The values are derived separately for the b and udcsg jets in the barrel end endcap regions. These MC specific correction factors are small for most of the alternative simulations. However, for the modified FSR shower scales and the Herwig++ sample they are not negligible compared to the overall effect typically observed for the variations.

We tried the fit by applying the Up/Down (JER)JES correction to the tt FSR Up/Down variation alone. This constitutes a conservative over-coverage of  $p_T$  correction effect w.r.t. the one reported in the plot in Figure 33. It has been verified with this test that this correction does not affect the result of the fit.

## D.3 Combine autoMCstats threshold

The autoMCstats threshold of the Combine tool is the maximum number of events for a process in a bin for which the statistical uncertainty of that process in that bin to be consider as a Poissonian uncertainty. Above the threshold, the statistical uncertainties of all process are summed together and treated as a single nuisance parameter in the fit, according to the Barlow-Beeston lite (BB-lite) method. In the present configuration of the fit the threshold is set to 1 so all the processes in a each bin are treated with a pure BB-lite method. In order to avoid biases due to this procedure, a scan for different values of the threshold has been performed and the results are reported in the Table 14. As can be seen no bias is observed and this justify the approach used.

## D.4 Pre-approval fit pulls

Three of the most important pulls of the pre-approval fit version, reported in Figure 34, are related to bins in which a discrepancy is observed, as one can see from the plot in Figure 35.

Threshold	Fit result
10000000	$0.94 \pm 0.04$
10000	$0.94 \pm 0.04$
1000	$0.94 \pm 0.04$
100	$0.96 \pm 0.03$
50	$0.96 \pm 0.03$
25	$0.96 \pm 0.03$
1	$0.96 \pm 0.03$

Table 14: Results of the fit to the  $V_{tb}$  signal obtain by varying the threshold of the Barlow-Beeston lite method.

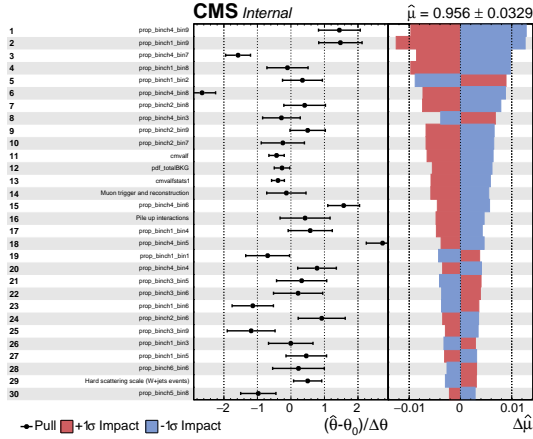


Figure 34: Most important impacts of the nuisances on the final result for the  $V_{tb}$  signal. Three of the most pulled nuisances account for the residual mismodeling observed in the 2-jets-1-tag region of the electron channel.

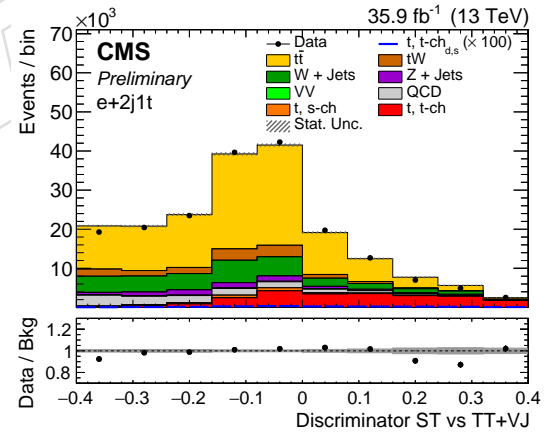


Figure 35: Data-MC comparison for the 2-jets-1-tag region for the electron channel. Mismodeling observed in few bins.

846 In these bins there is a discrete amount of QCD (data driven statistical uncertainty is equal  
 847 to propagated uncertainty from control region statistics) and W+ jets that are the two samples  
 848 with the largest MC uncertainties. Most of the W+ jets events come from W+2 jets events and  
 849 it is observed that this discrepancy is pronounced in regions enriched with W+0/1 generated  
 850 jets, which also has highest MC stat uncertainty, as shown in Figure 36.

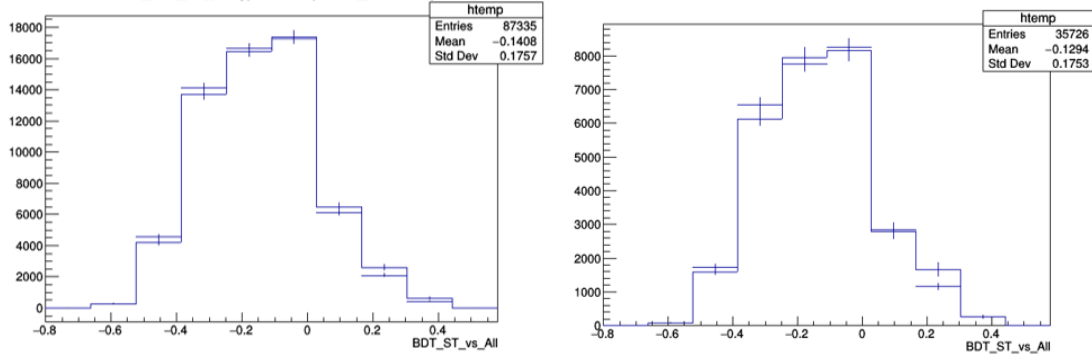


Figure 36: Shapes of the W+ jets sample plotted by considering only the W+2jets component or the W+0,1,2jets, muon to the left and electron to the right. The difference is taken as input for a one-side systematic.

851 In order to take into account this mismodelling a new systematics is introduced to cover the  
 852 difference between the shape of the total W+ jets and the W+2 jets only. With this new nuisance  
 853 the pulls, reported in Figure 37 are mitigated in the signal region: still some slight tension in  
 854 the first bins as MC limited statistic nuisances are covering residual mismodeling. This can  
 855 come from residual signal modeling uncertainties, that are not included in the fit, since they  
 856 are externalized.

857 It is checked also stability against re-binning and, as can be appreciated from Figure 38, no  
 858 significant improvement came from this in terms of pulls reduction since the fit is still molding  
 859 the shape thanks to bin-by-bin uncertainties.

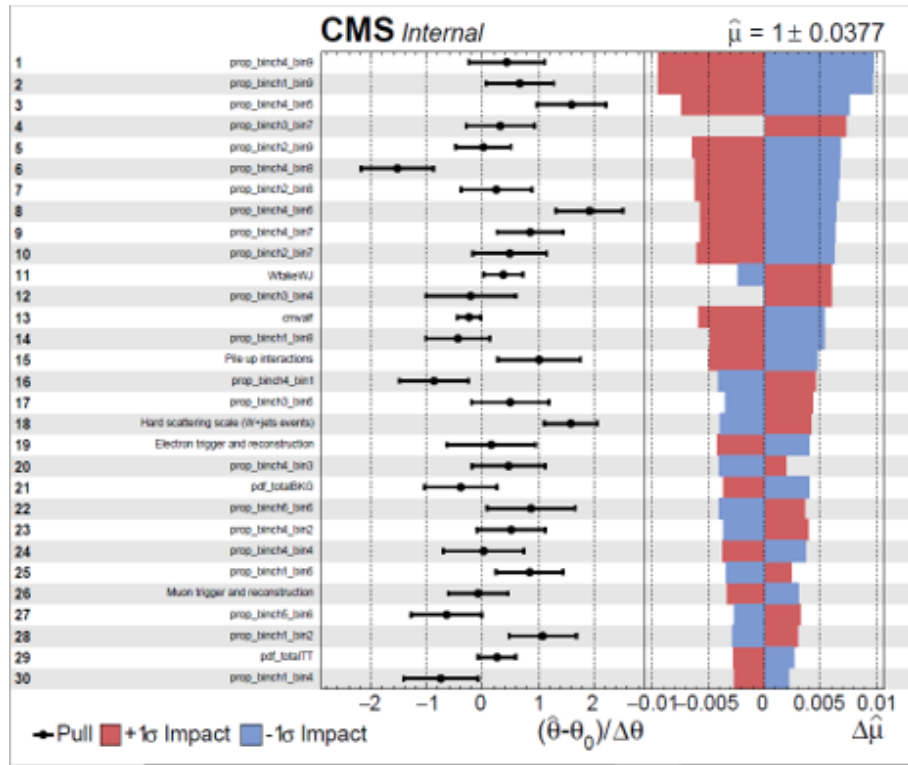


Figure 37: Most important impacts of the nuisance after the introduction of the new systematics on the  $W^+$  jets rate.

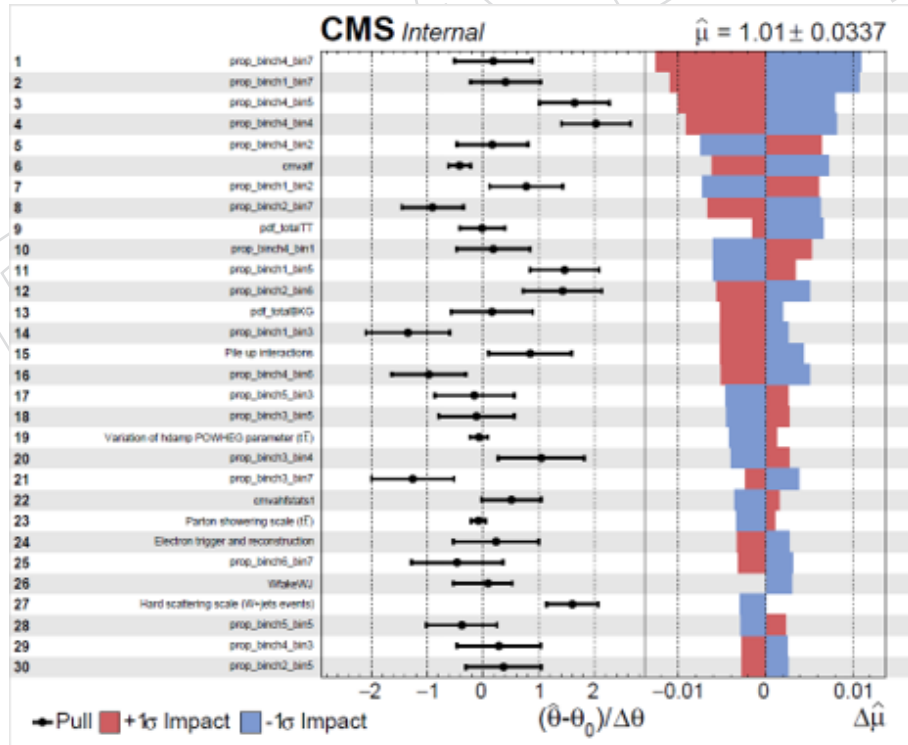


Figure 38: Most important impacts of the nuisance obtained by reducing the numbers of bins from 10 to 8 in all the six distributions used in the fit.