Extraction of CKM matrix elements in single top quark $t$-channel events in proton-proton collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration

Abstract

This note presents a model-independent extraction of the modulus of the Cabibbo Kobayashi Maskawa matrix elements $V_{tb}$, $V_{td}$, and $V_{ts}$ using an event sample enriched in single top quark $t$-channel events. The analysis uses proton-proton collisions data from the LHC collected during 2016 by the CMS experiment at a centre-of-mass energy of 13 TeV, corresponding to an integrated luminosity of 35.9 fb$^{-1}$. Processes directly sensitive to the matrix elements $V_{tb}$, $V_{td}$, and $V_{ts}$ are considered in both the production and decay vertices of the top quark in single top quark $t$-channel production, as well as in the background processes containing top quarks. Final states are investigated where a muon or electron stems from the leptonic decay chain of the top quark, and two or three jets are selected, one or two of which are identified as coming from the hadronisation of a $b$ quark. The event sample is divided into categories according to the number of jets. Multivariate classifier variables are built in each category in order to discriminate between signal and other standard model processes, and a simultaneous maximum-likelihood fit to data is performed on all categories. The measured value of $|V_{tb}|$ is $1.00 \pm 0.03$, where the uncertainty includes both statistical and systematic uncertainties, and the upper limit derived on $|V_{ts}|^2 + |V_{td}|^2$ is 0.17 at 95% confidence level.
1 Introduction

A defining feature of the electroweak sector of top quark physics is the relative magnitude of the Cabibbo Kobayashi Maskawa (CKM) [1] matrix element V_{tb} with respect to V_{td} and V_{ts}, which leads to a strong suppression of processes involving the mixing between the third quark family and the other two. This feature can be probed at the LHC by studying the couplings of top quarks with d, s, and b quarks in electroweak charged-current interactions, where such couplings can manifest itself in either the production or decay vertices of the top quark. In general, top quarks are produced in proton-proton collisions through strong interactions, predominantly via gluon fusion, creating a top quark-antiquark pair, and through electroweak interactions. The dominant mechanism in the latter case is the production of a single top quark through exchange of a W boson in the $t$-channel, which has been precisely measured at the LHC [2–11]. The dominant decay process for top quarks is to a W boson and a b quark, via an electroweak charged-current interaction. All single top quark processes therefore allow the direct probing of the tWq vertex, with q being a b, d, or s quark, both in production and decay of the top quark. Top quarks produced in the $t$-channel are accompanied by a spectator light quark. Figure 1 shows typical Feynman diagrams at leading order for the different production and decay modes considered in the analysis.

The elements of the CKM matrix V_{tb}, V_{td}, and V_{ts} can be indirectly constrained from experiments in the B and K meson sectors [12], but those determinations rely crucially on model assumptions such as the existence of only three generations of quarks and the absence of particles beyond the standard model (SM) in the loops [13]. This model dependence motivates alternative inferences based on different sets of hypotheses. In particular, given that these three CKM elements connect the top quark with down-type quarks, it is natural to use events enriched in top quarks to set constraints on them. Two complementary approaches have been pursued, so far, by Tevatron and LHC experiments to extract $|V_{tb}|$: measuring the branching fraction $B(t \to Wb) = |V_{tb}|^2/(|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2)$ in $tt$ events [14–17], or exploiting the cross section of single top quark production. The $B(t \to Wb)$ measurement is sensitive to the CKM elements of interest through the decay vertex of the top quark, and can be turned into a measurement of $|V_{tb}|$ only under the hypothesis of the unitarity of the $3 \times 3$ CKM matrix. Instead, the method based on single top quark production cross section, instead, is sensitive in principle through both the production and decay of the top quark. To disentangle the effects at the two vertices, in all past measurements from the Tevatron [18–27] and LHC [3–10, 28], $|V_{tb}|$ was extracted in the $t$-channel under the assumption that the values of $|V_{ts}|$ and $|V_{td}|$ are negligible. Some proposals have appeared in the theory literature to extract the three CKM matrix elements simultaneously from a combination of measurements of $B(t \to Wb)$ and either inclusive [13, 29] or differential [30, 31] cross sections of single top quark production in $t$ channel. Some theoretical studies specifically address the determination of $|V_{td}|$ [30, 32]. These studies rely on rather strong assumptions and approximations, and cannot exploit the discriminating power of multivariate analyses (see Ref. [33] for a discussion).

This note aims at a first direct and model independent simultaneous measurement of $|V_{tb}|$, $|V_{td}|$, and $|V_{ts}|$, by considering their respective contributions to $t$-channel production and top quark decay. It is organised as follows: Section 2 provides the details on the data set collected and simulated samples used in the analysis. Section 3 gives insights on the object selection and reconstruction of top quark candidates, while Section 4 illustrates the peculiarities of the sought after signals and the strategy to exploit them in the measurement. Sections 5 and 6 describe the signal extraction procedure, and the treatment of systematic uncertainties considered in the measurement, respectively. Finally, Section 7 summarizes the theoretical framework used in the signal extraction, followed by a short summary on the results.
Figure 1: Leading-order Feynman diagrams for single top quark production via the $t$-channel featuring: (a) a $tWb$ vertex in production and decay, (b) a $tWb$ vertex in production and a $tWq$ in decay, with $q$ being a $s$ or $d$ quark, (c) a $tWq$ vertex in production and a $tWb$ in decay, and (d) a $d$ quark-initiated process, enhanced thanks to contributions from valence $d$ quarks.

2 Data and simulated samples

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

Monte Carlo (MC) event generators are used to simulate signal and background samples. Single top quark $t$-channel events are generated at next-to-leading order (NLO) with POWHEG 2.0 [34] in the four-flavour scheme (4FS), and the top quark decays are simulated with MADSPIN [35]. Single top quark $t$-channel events with one $V_{td}$ or $V_{ts}$ vertex in production are generated at NLO with POWHEG 2.0 in the five-flavour scheme (5FS) and the top quark decays are simulated with MADSPIN.

The $t\bar{t}$ background process is also generated with POWHEG 2.0, as well as the double vector boson production [36, 37]. Associated top quark and $W$ boson production are simulated with POWHEG in the 5FS [38]. The value of the top quark mass used in the simulated samples is $m_t = 172.5$ GeV. For all samples PYTHIA 8.180 [39] with tune CUETP8M1 [40] is used to simulate the parton shower, the hadronisation, and the underlying event, except for $t\bar{t}$, where
3. Event selection and reconstruction

the tune CUETPM2T4 is used [41]. Simulated event samples with W and Z bosons in association with jets are generated using MG5_aMC@NLO [42] and the FxFx merging scheme [43], 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.

The default parametrisation of the parton distribution functions (PDFs) used in all simulations is NNPDF30_nlo_as_0118 [44]. All generated events undergo a full simulation of the detector response according to the implementation of the CMS detector within GEANT4 [45]. 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.

Except for the QCD multijet process, which is determined from a fit to data and given a corresponding systematic uncertainty, all simulated samples are normalized to the expected cross sections with uncertainties corresponding to the size of the samples.

3 Event selection and reconstruction

The signal event selection is based on the single top quark final states where the top quark decays to a b, s, or d quark, and a W boson, which then decays to a lepton-neutrino pair. Events with exactly one muon or electron and at least two jets are considered in this analysis, as was done in the CMS single top quark cross section measurement [10]. The neutrino accompanying the lepton cannot be directly detected, and manifests itself as a momentum imbalance in the detector. Depending on the CKM matrix element involved in the decay, the top quark decay cascade will include a jet from the hadronisation of either a b, s, or d quark. A spectator jet recoiling against the top quark is present, and it is produced usually at low angle with respect to the beam axis. A third jet can stem from the second quark produced in the gluon splitting (as shown in Fig. 1). The quark from gluon splitting generates a jet that usually has a softer $p_T$ spectrum than that of the jet from the top quark decay products. Depending on the number of $tWb$ vertices in an event, one can have one jet coming from a b quark if the $tWb$ vertex occurs in production or in decay but not both, or two jets coming from a b quark if the $tWb$ vertex occurs both in production and in decay. The selection accommodates all cases where either two or three high-$p_T$ jets are present, and either one or two are identified as coming from a b quark.

Events are retained for offline analysis if they were selected online by an HLT path that requires the presence of either an isolated muon with $p_T > 26$ GeV or an electron with $p_T > 32$ GeV. Further identification criteria for the HLT electron are applied, with an efficiency of 85% for prompt electrons. From the sample of triggered events, only those with at least one primary vertex reconstructed from at least four tracks, with the longitudinal (radial) distance of less than 24 (2) cm from the centre of the detector, are considered for the analysis. Among all primary vertices in the event, the one with the largest scalar sum of $p_T^2$ of associated particles is selected. The particle-flow (PF) algorithm [46] is used to reconstruct and identify individual particles in the event using combined information from the subdetectors of the CMS experiment, allowing identification of muons, electrons, photons, and charged and neutral hadrons. After triggering, muons are considered for further analysis if they have $p_T > 26$ GeV, and $|\eta| < 2.1$, while electrons are required to have a transverse energy $p_T > 35$ GeV and $|\eta| < 2.4$. Additional isolation requirements are used to discriminate between prompt leptons and leptons coming from hadronic decays within jets by requiring the lepton transverse momentum to be larger than about 6% of the scalar sum of hadronic $p_T$ in a cone of $\Delta R = \sqrt{\Delta \phi^2 + (\Delta \eta)^2} = 0.4$ (0.3) around the muon (electron), where the contribution from hadrons from pileup vertices is subtracted from the scalar sum.
Jets are reconstructed by using the anti-$k_T$ clustering algorithm [47, 48] with a distance parameter of 0.4 on the collection of PF particle candidates. To be included, charged particle candidates must be closer along the z axis to the primary vertex than to any other vertex.

A correction to account for pileup interactions is estimated on an event-by-event basis using the jet area method described in Ref. [49], and is applied to the reconstructed jet $p_T$. Further jet energy corrections, derived from the study of dijet events and photon plus jet events in data, are applied. Jets are required to have $|\eta| < 4.7$ and $p_T > 40\text{ GeV}$.

Once the jets have been selected according to the above criteria, they can be further categorized using a b tagging discriminator variable in order to distinguish between jets stemming from the hadronisation of b quarks and those from the hadronisation of light partons. The combined multivariate discriminator algorithm (MVA) uses track-based lifetime information together with secondary vertices inside the jet to provide a MVA discriminator for b jet identification [50, 51]. For values of the discriminator above the chosen threshold, the efficiency of the tagging algorithm to correctly find b jets is about 45% with a rate of 0.1% for mistagging light-parton jets [50, 51].

Events are divided into “regions” according to the number of selected jets and b-tagged jets. In the following, regions are labelled as "$n_j m_t$", referring to events with $n$ jets, $m$ of which are tagged as b jets. To reject events from QCD multijet background processes, a requirement on the transverse mass of the W boson of $m_W^T > 50\text{ GeV}$ is imposed, where

$$m_W^T = \sqrt{(p_{T,\ell} + p_{T,\nu}^{\text{miss}})^2 - (p_{x,\ell} + p_{x,\nu}^{\text{miss}})^2 - (p_{y,\ell} + p_{y,\nu}^{\text{miss}})^2}. \quad (1)$$

Here, $p_T^{\text{miss}}$ is defined as the magnitude of $\vec{p}_T^{\text{miss}}$, which is the negative of the vectorial $p_T$ sum of all the PF particles. The $p_x^{\text{miss}}$ and $p_y^{\text{miss}}$ quantities are the $\vec{p}_T^{\text{miss}}$ components along the x and y axes, respectively.

To analyse the kinematics of single top quark production, the momentum four-vectors of the top quarks are reconstructed from the decay products: muons, neutrinos, and b jet candidates. The $p_T$ of the neutrino can be inferred from the missing transverse momentum. The longitudinal momentum of the neutrino, $p_{z,\nu}$, is inferred assuming energy-momentum conservation at the $W\ell\nu$ vertex and constraining the W boson mass to $m_W = 80.4\text{ GeV}$ [12]:

$$p_{z,\nu}^\pm = \frac{\Lambda p_{z,\ell}}{p_{T,\ell}^2} \pm \frac{1}{p_{T,\ell}^2} \sqrt{\Lambda^2 p_{z,\ell}^2 - p_{T,\ell}^2 (E_{\ell}^2 p_{T,\nu}^2 - \Lambda^2)}, \quad (2)$$

where

$$\Lambda = \frac{m_W^2}{2} + \vec{p}_{T,\ell} p_{T,\nu}^{\text{miss}}, \quad (3)$$

and $E_{\ell}^2 = p_{T,\ell}^2 + p_{z,\ell}^2$ denotes the square of the lepton energy. In most of the cases this leads to two real solutions for $p_{z,\nu}$ and the solution with the smallest absolute value is chosen [21, 23]. For some events the discriminant in Eq. (2) becomes negative, leading to complex solutions for $p_{z,\nu}$. In this case the imaginary component is eliminated by modification of $p_{x,\nu}$ and $p_{y,\nu}$ so that $m_W^T = m_W$, while still respecting the $m_W$ constraint. This is achieved by requiring the determinant, and thus the square-root term in Eq. (2), equal zero. This condition gives a quadratic relation between $p_{x,\nu}$ and $p_{y,\nu}$ with two possible solutions and one remaining degree of freedom. The solution is chosen by finding the neutrino transverse momentum $\vec{p}_{T,\nu}$ that has the minimum vectorial distance from the $p_{T,\nu}^{\text{miss}}$ in the $p_x^{\text{miss}} - p_y^{\text{miss}}$ plane.
Top quark candidates are reconstructed by selecting one of the jets to accompany the W boson decay. Multiple top quark candidates can be reconstructed in the different regions, depending on the hypothesis for the origin of the jet in the event.

4 Signal description and event categorisation

The predicted branching fractions of top quarks to d, s, and b quarks can be written as a function of the overall magnitude of 

\[ B(t \rightarrow Wq) = \frac{|V_{tq}|^2}{|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2} \]

The value of |V_{tq}| and all possible combinations of B(t \rightarrow Wq), taken from [12], are shown in Table 1.

<table>
<thead>
<tr>
<th>Quark</th>
<th>b</th>
<th>s</th>
<th>d</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.999119^{+0.000024}_{-0.000012}</td>
<td>0.04108^{+0.000048}_{-0.000024}</td>
<td>0.008575^{+0.000076}_{-0.000098}</td>
</tr>
<tr>
<td>B(t \rightarrow Wq)</td>
<td>0.998239^{+0.000048}_{-0.000024}</td>
<td>0.0016876^{+0.0000025}_{-0.0000047}</td>
<td>0.000074^{+0.000013}_{-0.000017}</td>
</tr>
</tbody>
</table>

Table 1: Values of the third-row elements of the CKM matrix inferred from low-energy measurements, taken from [12], with the respective values of the top quark decay branching fractions. The q in |V_{tq}| and B(t \rightarrow Wq) of the first column refers to b, s, and d quarks accordingly to the quark present in the first row.

The channels considered as signal and the corresponding cross sections are reported in Table 2. The cross sections are evaluated at NLO in the five-flavour scheme using POWHEG 2.0 [34, 52–54] for \( \sigma_{t-ch,b} \), \( \sigma_{t-ch,s} \), and with HATHOR [55] for \( \sigma_{t-ch,b} \).

<table>
<thead>
<tr>
<th>Process</th>
<th>Cross section \times\text{branching fraction (pb)}</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma_{t-ch,b}B(t \rightarrow Wb) )</td>
<td>216.99 \pm 8.37(scale \oplus pdf)</td>
</tr>
<tr>
<td>( \sigma_{t-ch,b}(B(t \rightarrow Ws) + B(t \rightarrow WD)) )</td>
<td>0.41 \pm 0.05(scale \oplus pdf \oplus exp)</td>
</tr>
<tr>
<td>( \sigma_{t-ch,q}B(t \rightarrow Wb) )</td>
<td>0.102 \pm 0.015(scale \oplus pdf \oplus exp)</td>
</tr>
<tr>
<td>( \sigma_{t-ch,q}B(t \rightarrow Wb) )</td>
<td>0.92 \pm 0.11(scale \oplus pdf \oplus exp)</td>
</tr>
</tbody>
</table>

Table 2: Production cross section of signal processes taken into consideration, where the “scale” component of the uncertainty refers to factorisation and renormalisation scale uncertainties, “pdf” refers to uncertainties due to the PDFs, and “exp” denotes the experimental component of the uncertainty from low-energy measurements.

The signatures for t-channel processes involving \( V_{tb}, V_{td}, \) and \( V_{ts} \) either in production or decay do differ in three aspects: the number of b-tagged jets produced actually stemming from a b quark, the features of the jet involved in the reconstruction of the correct top quark candidate, and the kinematic features of the events as a result of different PDFs involved for b, s, or d incoming quarks.

Henceforth, the t-channel process involving \( V_{tb} \) in both production and decay will be referred to as \( ST_{b,b} \), while t-channel processes involving \( V_{tb} \) in only production or decay will be referred to as \( ST_{b,q} \) and \( ST_{q,b} \) respectively.

Strong interaction production of t\( \bar{t} \), where one top quark decays through tWd and tWs vertices, was studied. The process was found to be indistinguishable from the more frequent case where both top quarks decay through the tWb coupling. For this reason, \( t\bar{t}_{b,q} \) processes are absorbed in the background contribution.

Multiple regions are defined in order to extract the contribution of the different t-channel processes, while at the same time discriminating against the background processes, mainly t\( \bar{t} \) pair
production and W boson plus jets associated production (W+jets). The majority of t-channel events populate regions with 2 or 3 jets, as defined above. The main backgrounds arise from tt (all regions), W+jets (in the 2j1t and 3j1t regions), and QCD multijet (in the 2j1t region). The signal processes taken into consideration give different contributions to the three regions, and it is possible to identify the most sensitive regions with respect to each process based on the respective signatures, as summarized in Table 3. The physics motivations leading to this strategy are described in this section.

<table>
<thead>
<tr>
<th>Region</th>
<th>Enriched in Corresponding cross section × BR</th>
<th>Feynman diagram</th>
</tr>
</thead>
<tbody>
<tr>
<td>2j1t</td>
<td>ST_{b,b} $\sigma_{t-ch,b} B(t \to Wb)$</td>
<td>1a</td>
</tr>
<tr>
<td>3j1t</td>
<td>ST_{b,q}, ST_{q,b} $\sigma_{t-ch,b} B(t \to Wq), \sigma_{t-ch,q} B(t \to Wb)$</td>
<td>1b,1c,1d</td>
</tr>
<tr>
<td>3j2t</td>
<td>ST_{b,b} $\sigma_{t-ch,b} B(t \to Wb)$</td>
<td>1a</td>
</tr>
</tbody>
</table>

Table 3: For each region, the corresponding signal process, the cross section times branching fraction determined, and the specific Feynman diagram from Fig. 1 involved.

The discrimination between the three signals ST_{b,q}, ST_{q,b}, and ST_{b,b} is based on three characteristics. First, for ST_{b,q} events, only a single b quark is present in the final state stemming from gluon splitting, thus resulting in a low-energy b-tagged jet, while the jet coming from the top quark decay will not be b-tagged. For ST_{q,b} events, a single b-tagged jet will be produced in the top quark decay, and additional jets from gluon splitting will not be b-tagged. Both ST_{q,b} and ST_{b,q} processes therefore differ from ST_{b,b} by means of having a single b quark in the final state, as opposed to two for the latter process. However, this feature can only be exploited when the jet from gluon splitting is energetic enough to be reconstructed.

Second, further discrimination is achieved by exploiting the features of the reconstructed top quark candidates. The kinematic and angular properties of the decay products will exhibit significant differences depending on whether the correct jet is chosen, or if the jet that originated from the gluon splitting-quark is used instead. The top quark reconstructed with the correct jet assignment will usually not be the one with the b-tagged jet in the event, while in the case of ST_{b,b} and ST_{q,b} the top quark candidate will be reconstructed by using the b-tagged jet in the majority of cases. It is therefore possible to differentiate between the ST_{b,b} and ST_{b,q} processes by comparing the features of top quark candidates reconstructed with b-tagged jets and non-b-tagged jets.

Finally, different PDFs are involved in ST_{b,b} and ST_{q,b} processes, the latter drawing contributions from valence d quarks as well.

The second characteristic, related to the correctness of the top quark reconstruction hypothesis, proves to be the strongest in terms of discrimination amongst the three mentioned. However, since the signature of the ST_{q,b} features the same decay as the standard ST_{b,b} production, the two channels cannot be differentiated using this characteristic.

The region 2j1t is populated by events with V_{tb} in production and decay, where the single reconstructed b jet comes in the majority of cases (85%) from top quark decays, and for the remaining cases from the second b jet from gluon splitting. This means that the jet from the second b quark fails either the jet $p_T$ requirement or the b tag requirement, or both. Events for which V_{ld} or V_{ls} are present, either in production or in decay, populate this region as well, with either the b-tagged jet coming from top quark decay or the secondary b quark from gluon splitting.

For t-channel signal events from all four processes, the most distinctive features that allow the discrimination against backgrounds in this region rely on the fact that the second jet stems
4. Signal description and event categorisation

from the light recoil quark, and for this reason the non-b-tagged jet is not used for top quark reconstruction. This region is the one where the highest discrimination power for ST_{b,b} against backgrounds is achieved by making use of the features of the top quark decay products, like the reconstructed top quark mass or the W boson transverse mass, and of the light recoil jet. However, the discrimination power with respect to other t-channel mechanisms is poor since jets from gluon splitting are typically not energetic enough to pass the p_T threshold, making it impossible to reconstruct two different top quark candidates.

The region 3j1t is also populated by all of our interesting t-channel processes, and it differs from 2j1t in the fact that it accommodates events in the higher end of the p_T spectrum of the jet from gluon splitting. For both the 2j1t and 3j1t regions, for processes where the top quark decays proceed through tW_d,s vertices, the jet coming from the top quark should not pass the b-tagging requirement since it stems from the hadronisation of a light quark, while it is expected to pass the b-tagging requirement in all other cases according to the efficiency of the tagging algorithm.

The 3j1t region is enriched in t-channel events by using the variable \( \eta_{j'} \), the pseudorapidity of the most forward jet, and by selecting events with \( |\eta_{j'}| > 2.5 \). The two jets other than the most forward one are used to reconstruct the two top quark candidates. In case it is an event of the ST_{b,q} process, the b-tagged jet in 3j1t will stem from gluon splitting, and the additional jet will have a higher chance of being the one coming from the top quark decay to an s or d quark. Variables of interest in this case include the invariant mass of the lepton and either jet (the b jet or the leading extra jet), and several top quark kinematic variables constructed using a combination of the extra jet, the missing transverse momentum, and the lepton.

In both the 2j1t and the 3j1t regions the transverse mass of the W boson is also used to discriminate between QCD multijet background and other samples. A QCD multijet-depleted region is defined by adding the requirement \( m_W^T > 50 \text{ GeV} \). Figure 2 shows the \( m_W^T \) distribution from data in the 2j1t and 3j1t for the muon (upper plots) and electron (lower plots) channels, with the result of the fit.

In the 3j2t region there are two b jets, one produced from the top quark decay and 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 \( m_W^T > 50 \text{ GeV} \) requirement since the QCD multijet contamination is negligible and the dominant background process is t\bar{t}. No categorisation on \( \eta_{j'} \) is performed as there is no need to reduce the combinatorial top quark background, as this region is dominated by ST_{b,b}.

Multivariate analyses are then performed by using boosted decision trees (BDT) in order to obtain appropriate discriminating variables, henceforth referred to as BDT discriminants, in the three regions, separately for muons and electrons. The choices of the processes to use as signal or background in the training are the following:

- In the 2j1t region the single top quark ST_{b,b} process is considered as signal and t\bar{t} and W+jets processes as background.
- In the 3j1t region the single top quark ST_{q,b} process is considered as signal and the ST_{b,b}, t\bar{t} and W+jets processes as background.
- In the 3j2t region the single top quark ST_{b,b} process is considered as signal and t\bar{t} as background.

Figures 3, 4, and 5 show the most discriminating variables in the 2j1t, 3j1t, and 3j2t regions, respectively.
The CKM matrix elements are extracted by measuring the production cross sections and branching fractions of single top quark t-channel processes drawing contributions from $V_{tb}$, $V_{td}$, and $V_{ts}$ in production and decay. The vast majority of single top quark t-channel events comes from the $ST_{bb}$ process, while $ST_{b,q}$ and $ST_{q,b}$ constitute subdominant production mechanisms. The $t\bar{t}_{b,q}$ contribution is taken into account in the background estimation.

The fit procedure is divided into two steps. In the first step a maximum likelihood (ML) fit to the $m_{T}^{W}$ distribution is performed separately for the 2j1t and 3j1t regions in order to extract the QCD multijet contribution. The QCD multijet normalisation and the relative uncertainty are extrapolated to the QCD multijet-depleted regions and used as an input to the second fit procedure. In the second step in order to discriminate between $ST_{bb}$, $ST_{b,q}$, $ST_{q,b}$, the multivariate discriminants described in Section 4 are used in a simultaneous ML fit across the respective regions, while the QCD multijet prior uncertainty and central value are taken from the first step. The ML fit is performed simultaneously in the 2j1t, 3j1t, and 3j2t regions. The t-channel single top quark signals are parametrized with a flat prior representing the coupling strength, and all systematic uncertainties defined in Section 6 are treated as nuisance parameters. The smaller background yields are allowed to float in the fit, along with the respective scale uncertainties. The QCD multijet background is fitted with a flat prior nuisance, while $t\bar{t}$ and W+jets are left floating according to the respective systematic uncertainties. The modelling uncertainties im-
Figure 3: Distributions of the two most discriminant variables in the 2j1t region: the absolute value of the pseudorapidity of the non b-tagged jet $\eta_{j'}$ (left) and the invariant mass of the vectorial sum of the lepton and b jet momenta (right), shown for the muon (upper) and electron (lower) channel, respectively.

Impacting the rates of those processes are large enough that no additional uncertainty is needed to fit the yield. The $t$-channel $ST_{b,q}$ and $t\bar{t}b,q$ cannot distinguish amongst topologies with $V_{td}$ or $V_{ts}$ in the decay, while $ST_{q,b}$ is sensitive to the different PDFs contributing to the processes. Figure 6 shows the distributions after the fit procedure has been applied for (left) muons and (right) electrons.
Figure 4: Distributions of the two most discriminant variables in the 3j1t region: the missing momentum in the transverse plane (left) and the response of the CMVAv2 b-tagger discriminator when applied to the extra jet (right) are shown for the muon (upper) and electron (lower) channel, respectively.
Figure 5: Distributions of the two most discriminant variables in the 3j2t region: the absolute value of the pseudorapidity of the non b-tagged jet $\eta_j'$ (left) and the invariant mass of the vectorial sum of the lepton and non b-tagged jet (right) are shown for the muon (upper) and electron (lower) channel, respectively.
Figure 6: Distribution of the multivariate discriminants, comparing data to simulation normalized after the fit procedure for muon channel on the left and for the electron on the right, for 2j1t (top), 3j1t (middle), and 3j2t (bottom). The systematic bands correspond to the profiled uncertainties constrained by the fit procedure.
6 Systematic uncertainties

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. These uncertainties include the uncertainty sources related to the modelling of the signal process, which cannot be constrained from the measurement since they apply to the full phase space and not only to the region in which the measurement is performed, and jet energy scale and resolution uncertainties, which play a major role for events featuring hadronic activity in the high-pseudorapidity region of the detector, and are also intertwined with uncertainties in the modelling of the hadronisation. For this reason, a more conservative approach is preferred for these uncertainties. In this analysis the jet related uncertainties are not profiled because this measurement is sensitive to the effect of uncertainties in the total signal acceptance, if compared with other measurements [10, 11].

The impact of nonprofiled uncertainties 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 simulated event samples**: The statistical uncertainty due to the limited size of the simulated event samples is evaluated for each bin and each process with the Barlow-Beeston “light” method [56, 57].

- **Lepton trigger and reconstruction**: Single-muon and single-electron trigger and reconstruction efficiencies are estimated with a “tag-and-probe” method [58] from Drell–Yan events with the dilepton invariant 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 in this measurement by varying the total proton-proton inelastic 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 mismodelling of the $t\bar{t}$ process. Their effect is considered on both the acceptance and the cross section. For the ML fit, the uncertainties are considered as nuisance parameters.
  - **$t\bar{t}$ renormalisation and factorisation scale uncertainty ($\mu_R/\mu_F$)**: The uncertainties caused by variations in the renormalisation and factorisation scales are considered by reweighting the BDT response distributions with different combinations of doubled/halved renormalisation and factorisation scales, with respect to 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.

- **QCD multijet background normalization**: The yield of the QCD multijets process is assigned a 50% uncertainty, which is chosen conservatively to be much larger with respect to the $m_T^{WW}$ fit uncertainty.
- **W+jets composition**: A separate uncertainty is dedicated to the fraction of W+jets events where the forward jet is not matched to a quark from the hard-scattering process.

- **Other backgrounds $\mu_R/\mu_F$**: In addition to $t\bar{t}$, the uncertainties due to variations in the renormalisation and factorisation scales are studied for the tW and W+jets processes 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 [59].

- **b tagging**: The uncertainties in the b tagging and mistagging efficiency measurements are split into different components.

### Nonprofiled uncertainties

- **Luminosity**: The integrated luminosity is known with a relative uncertainty of $\pm 2.6\%$ [60].

- **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 [61]. This variation in jet four-momenta is also propagated to $p_T^{\text{miss}}$.

- **Jet energy resolution (JER)**: A smearing is applied to account for the difference in the JER between simulation and data [61], and its uncertainty is estimated by increasing or decreasing the resolutions by their uncertainties.

- **Signal modelling**: The following uncertainty sources cover potential mismodelling of the single top quark $t$-channel signal processes. The effect of those uncertainties on the acceptance, and not on the cross section, is considered. For the ML 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 modelling parameters covering the systematic uncertainty sources under study.

  - **Signal $\mu_R/\mu_F$**: The uncertainties caused by variations in the renormalisation and factorisation scales are considered by reweighting the BDT response distributions with different combinations of doubled/halved renormalisation and factorisation 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.

  - **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 [59]. Only the effect on the acceptance, and not on the cross section, is considered.

In order to perform the limit extraction, uncertainties are profiled to provide an upper limit on $|V_{td}|$ and $|V_{ts}|$. 

6. Systematic uncertainties

The partial and total contributions of profiled and nonprofiled uncertainties are given in Table 4.

Table 4: Estimated relative contributions of uncertainty sources in % to the total uncertainties of the measured cross sections for single top quark production.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Uncertainty</th>
<th>$\Delta\sigma/\sigma$(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profiled</td>
<td>Lepton trigger and reconstruction</td>
<td>0.50</td>
</tr>
<tr>
<td>Profiled</td>
<td>Limited size of samples of simulated events</td>
<td>3.13</td>
</tr>
<tr>
<td>Profiled</td>
<td>$t\bar{t}$ modelling</td>
<td>0.66</td>
</tr>
<tr>
<td>Profiled</td>
<td>Pileup</td>
<td>0.35</td>
</tr>
<tr>
<td>Profiled</td>
<td>QCD background normalization</td>
<td>0.08</td>
</tr>
<tr>
<td>Profiled</td>
<td>$W+$jets composition</td>
<td>0.13</td>
</tr>
<tr>
<td>Profiled</td>
<td>Other backgrounds $\mu_R/\mu_F$</td>
<td>0.44</td>
</tr>
<tr>
<td>Profiled</td>
<td>PDF for background processes</td>
<td>0.42</td>
</tr>
<tr>
<td>Profiled</td>
<td>b-tagging</td>
<td>0.73</td>
</tr>
<tr>
<td>Profiled</td>
<td>Total profiled</td>
<td>3.4</td>
</tr>
<tr>
<td>Nonprofiled</td>
<td>Luminosity</td>
<td>2.6</td>
</tr>
<tr>
<td>Nonprofiled</td>
<td>JER</td>
<td>2.8</td>
</tr>
<tr>
<td>Nonprofiled</td>
<td>JES</td>
<td>8.0</td>
</tr>
<tr>
<td>Nonprofiled</td>
<td>PDF for signal process</td>
<td>3.8</td>
</tr>
<tr>
<td>Nonprofiled</td>
<td>Signal $\mu_R/\mu_F$</td>
<td>2.4</td>
</tr>
<tr>
<td>Nonprofiled</td>
<td>ME-PS matching</td>
<td>3.7</td>
</tr>
<tr>
<td>Nonprofiled</td>
<td>Parton shower scale</td>
<td>6.1</td>
</tr>
<tr>
<td>Nonprofiled</td>
<td>Total nonprofiled</td>
<td>11.5</td>
</tr>
<tr>
<td>Total</td>
<td>Total uncertainty</td>
<td>12.0</td>
</tr>
</tbody>
</table>
7 Results and interpretation

The contributions of each CKM matrix element to the different \( ST_{b,b} \), \( ST_{b,q} \), and \( ST_{q,b} \) cross sections, extracted from the fit procedure, are considered.

In the SM, top quarks only decay to W bosons plus b, s, or d quarks, and their branching fractions are proportional to the respective matrix element, and the relations in Table 2 apply.

We write the results of the fit procedure in terms of two signal strength parameters: the first one, \( \mu_b \), refers to the \( ST_{b,b} \) process, and the second one, \( \mu_{sd} \), refers to the sum of the \( ST_{q,b} \) and \( ST_{b,q} \) contributions.

By neglecting terms proportional to \( |V_{td}|^4 \) and \( |V_{ts}|^4 \) and considering the ratio \( |V_{td}| / |V_{ts}| \) to be small, the \( ST_{q,b} \) term can be written as proportional to \( |V_{td}|^2 + |V_{ts}|^2 \), plus terms of order \( |V_{td}| / |V_{ts}| \). These assumptions can be justified by taking into account the hierarchy observed in the first two rows of the CKM matrix. The signal strength thus becomes:

\[
\begin{align*}
\mu_b &= \frac{\sigma_{t-ch,b}^{\text{obs}} B(t \rightarrow Wb)^{\text{obs}}}{\sigma_{t-ch,b} B(t \rightarrow Wb)} \\
\mu_{sd} &= \frac{\sigma_{t-ch,b}^{\text{obs}} B(t \rightarrow Ws,d)^{\text{obs}} + \sigma_{t-ch,s,d}^{\text{obs}} B(t \rightarrow Wb)^{\text{obs}}}{\sigma_{t-ch,b} B(t \rightarrow Ws,d) + \sigma_{t-ch,s,d} B(t \rightarrow Wb)},
\end{align*}
\]

where, from here on, the “obs” superscript refers to the value of the quantity measured in data, and the absence of such a superscript means the theoretical value.

More generally, we can write Eq. 4 in terms of the top quark decay amplitudes or partial widths. The partial width of top quark to Wq is \( \Gamma_b |V_{tq}|^2 \), where \( \Gamma_b \) is the partial width of the top quark for \( |V_{tb}| = 1 \). Using this and the total width \( \Gamma_{\text{top}} \) of the top quark, we can write Eq. 4 as:

\[
\begin{align*}
\mu_b &= \frac{|V_{tb}|^4 \Gamma_{\text{top}}^q}{|V_{tb}|^4 \Gamma_{\text{top}}^q} \\
\mu_{sd} &= \frac{(|V_{ts}|^2 + |V_{td}|^2) \Gamma_{\text{top}}^q}{(|V_{ts}|^2 + |V_{td}|^2) \Gamma_{\text{top}}^q}.
\end{align*}
\]

One can extract from the fit either the signal strengths or the values of the squares of the CKM matrix elements \( |V_{td}|^2 \), \( |V_{ts}|^2 \), and \( |V_{tb}|^2 \).

7.1 Measurement for the unconstrained scenario

The simplest approach to derive the different CKM matrix elements is to replace \( B(t \rightarrow Wq) \) with \( |V_{tq}|^2 / (|V_{td}|^2 + |V_{ts}|^2 + |V_{tb}|^2) \), with q indicating d, s, or b quarks, and use the observed cross sections and branching fractions \( \sigma_{t-ch,q}^{\text{obs}} B(t \rightarrow Wq)^{\text{obs}} \), given in Table 1.

The results from the two fits are:

\[
\begin{align*}
\mu_b &= 0.99 \pm 0.03 \text{ (stat + syst) } \pm 0.12 \text{ (nonprofiled)} \\
\mu_{sd} &\leq 87 \text{ at 95% confidence level (CL).}
\end{align*}
\]
The above results do not consider the effect of $R_d + R_s + R_b$, where $R_q = \Gamma_q / \Gamma_{\text{top}}$, being different from one. By making use of Eq. 4 and of the values reported in Tables 1-2, and considering $B(t \to Wq) = R_q = |V_{tq}|^2$, the fit can be presented in terms of the CKM matrix elements, thus measuring:

\[ |V_{tb}| = 1.00 \pm 0.01 \text{ (stat + syst)} \pm 0.03 \text{ (nonprofiled)} \]
\[ |V_{tb}|^2 = 0.99 \pm 0.02 \text{ (stat + syst)} \pm 0.06 \text{ (nonprofiled)} \]
\[ |V_{td}|^2 + |V_{ts}|^2 \leq 0.17 \text{ at 95\% CL.} \]

The results shown in Eq. 7 do not include the unitary constraint on $|V_{tb}|, |V_{td}|, |V_{ts}|$. Moreover, the assumption $R_d + R_s + R_b = 1$ is not generally respected in many beyond-the-SM (BSM) scenarios. Therefore, in the following subsections different hypotheses are assumed.

### 7.2 Measurement for the constrained SM scenario

In the SM case, one can simplify Eq. 4 by assuming the unitarity constraint $|V_{tb}|^2 + |V_{td}|^2 + |V_{ts}|^2 = 1$. Assuming this, the fit is repeated taking $|V_{tb}|$ as the single free parameter and replacing $|V_{td}|^2 + |V_{ts}|^2$ with $1 - |V_{tb}|^2$. In this case, Eq. 4 becomes:

\[ \mu_b = \frac{|V_{tb}|^4_{\text{obs}}}{|V_{tb}|^4} \]
\[ \mu_{sd} = \frac{|V_{tb}|^2_{\text{obs}} (1 - |V_{tb}|^2_{\text{obs}})}{|V_{tb}|^2 (1 - |V_{tb}|^2)} \]

resulting in:

\[ |V_{tb}| = 0.980^{+0.014}_{-0.011} \text{ (stat + syst)} \pm 0.031 \text{ (nonprofiled)} \]
\[ |V_{td}|^2 + |V_{ts}|^2 = 0.040^{+0.023}_{-0.028} \text{ (stat + syst)} \pm 0.059 \text{ (nonprofiled).} \]

### 7.3 Measurement for BSM scenarios

Any BSM contribution potentially enhancing $|V_{tb}|^2, |V_{ts}|^2$, or $|V_{td}|^2$ can act in top quark production, decay, or both. Some BSM scenarios predict the presence of additional quark families. In this case, the CKM matrix is extended to due to the mixing between the SM quarks and the new hypothesized ones. This would imply that the CKM matrix elements $|V_{tb}|, |V_{ts}|$, and $|V_{td}|$ would not necessarily satisfy the unitarity constraint of $|V_{tb}|^2 + |V_{ts}|^2 + |V_{td}|^2 = 1$. If these BSM quarks are heavier than the top quark, these new particles would alter the CKM matrix elements without appearing as top quark decay products, and thus not contributing directly to the top quark decay width $\Gamma_{\text{top}}$. The effect on the top quark branching fractions would cause a reduction in the absolute values of the corresponding SM CKM matrix elements.

In a first interpretation including BSM physics, we assume the top quark decays through the same channels as in the SM case, and that the partial width of each decay only varies because of a modified CKM matrix element. In this case, by writing $R_q$ as a function of $|V_{tb}|^2$ and $|V_{td}|^2 + |V_{ts}|^2$, Eq. 5 becomes:
\begin{align*}
\mu_b &= \frac{|V_{tb}^4|_{obs}}{|V_{tb}^4|_{obs} (|V_{tb}^2|_{obs} + |V_{td}^2|_{obs} + |V_{ts}^2|_{obs})} \\
\mu_{sq} &= \frac{|V_{tb}^2|_{obs} (|V_{ts}^2|_{obs} + |V_{td}^2|_{obs})}{(|V_{ts}|^2 + |V_{td}|^2) (|V_{tb}|^2_{obs} + |V_{ts}|^2_{obs} + |V_{td}|^2_{obs})}.
\end{align*}

(10)

In this scenario, the measurement is performed leaving $|V_{tb}|$ and $|V_{td}|^2 + |V_{ts}|^2$ as free parameters in the fit, resulting in:

\begin{align*}
|V_{tb}| &= 0.988 \pm 0.027 \text{ (stat + syst)} \pm 0.043 \text{ (nonprofiled)} \\
|V_{td}|^2 + |V_{ts}|^2 &= 0.06 \pm 0.05 \text{ (stat + syst)} \pm 0.04 \text{ (nonprofiled)}.
\end{align*}

(11)

Finally, we consider a case for which BSM physics leaves the partial width of the top quark decaying to $W_b$ or $W_q$ unchanged, but modifies the overall total width of the top quark via additional, undetected, decays. This means that we assume the partial width of known quarks to be unchanged, but the total width of the top quark is free to change within its known experimental boundaries [62, 63]. In this scenario, we neglect the effects of variations of $|V_{tb}|^2$, $|V_{td}|^2$, and $|V_{ts}|^2$ on $\Gamma_{\text{top}}$.

We modify Eq. 5 by considering a generic width for the top quark, obtaining:

\begin{align*}
\mu_b &= \frac{|V_{tb}^4|_{obs} \cdot \Gamma_{\text{top}}}{|V_{tb}^4| \cdot \Gamma_{\text{top}}} \\
\mu_{sd} &= \frac{|V_{tb}^2|_{obs} \cdot (|V_{ts}|^2_{obs} + |V_{td}|^2_{obs}) \cdot |V_{tb}^4| \cdot \Gamma_{\text{top}}}{(|V_{ts}|^2 + |V_{td}|^2) \cdot \Gamma_{\text{top}}^{obs}}.
\end{align*}

(12)

By performing the fit leaving as free parameters $|V_{tb}|^2$, $|V_{td}|^2 + |V_{ts}|^2$, and $\Gamma_{\text{top}}^{obs} / \Gamma_{\text{top}}$, we obtain:

\begin{align*}
|V_{tb}| &= 0.988 \pm 0.011\text{ (stat. + syst.)} \pm 0.021\text{ (nonprofiled)} \\
|V_{td}|^2 + |V_{ts}|^2 &= 0.06 \pm 0.05\text{ (stat. + syst.)} \pm 0.04\text{ (nonprofiled)} \\
\frac{\Gamma_{\text{top}}^{obs}}{\Gamma_{\text{top}}} &= 0.99 \pm 0.42\text{ (stat. + syst.)} \pm 0.03\text{ (nonprofiled)}.
\end{align*}

(13)

It is worthy of note that in the SM interpretation, both measurements in Eq. 9 are strongly limited by nonprofiled systematic uncertainties, as their uncertainties are fully anticorrelated. The $S_{T_{bb}}$ modelling uncertainties dominate given the relative abundance of the signal with respect to the $S_{T_{q,b}}$ and $S_{T_{b,q}}$ processes. In this scenario, given the absence of a correlated term in the ratio, the statistical precision of the $|V_{tb}|$ determination is close to the one in Eq. 9.

As mentioned in Section 1, constraints on $|V_{td}|$ and $|V_{ts}|$ from precision physics measurements do not necessarily hold when BSM particles are present in the loops. Therefore, values of up to $|V_{ts}| \simeq 0.2$ are attainable [13]. The current measurement establish an upper limit in those...
8. Summary

A measurement of the CKM matrix elements $|V_{tb}|$, $|V_{td}|$, and $|V_{ts}|$ has been performed in an event sample enriched in $t$-channel single top quark events, featuring one muon or electron and jets in the final state. The data sample is from proton-proton collisions at $\sqrt{s} = 13$ TeV, acquired at the LHC by the CMS experiment and corresponding to an integrated luminosity of $35.9 \text{ fb}^{-1}$. The contributions from single top quark processes featuring all three matrix elements in the production vertex have been considered as separate signal processes, as well as contributions from decays of single top quarks involving all three quark families. The yields of the signal processes have been extracted through a simultaneous fit to data in different selected regions, and the values of the CKM matrix elements have been inferred from the signal strengths. All results are in agreement with the SM predictions.
References


