CMS Physics Analysis Summary

Contact: cms-pag-conveners-b2g@cern.ch 2016/08/24

Search for $W' \to tb$ in pp collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration

Abstract

A search is performed for the production of a massive W′ boson decaying to a top and a bottom quark. The data analyzed correspond to an integrated luminosity of 2.55 fb^{-1} collected with the CMS detector at the LHC in proton-proton collisions at $\sqrt{s} = 13$ TeV. The hadronic decay products of the top quark produced in the W′ boson decay with high Lorentz boost are detected as a single top-flavored jet. The use of jet substructure algorithms allows the top quark jet to be distinguished from standard model QCD multi-jet background. The production of a right-handed W′ boson with a mass below 2.0 TeV decaying to a hadronic final state is excluded at 95% confidence level.

1 Introduction

Many extensions of the standard model (SM) predict new massive charged gauge bosons [1–3]. The W′ boson is a heavy partner of the SM W boson that could manifest itself in proton-proton collisions at the CERN Large Hadron Collider (LHC). Searches for a high-mass W′ boson resonance have been performed at the Tevatron [4, 5] and the LHC [6–15] in the lepton-neutrino, diboson and diquark final states.

We present a search using the $W'^{+} \to t\bar{b}$ and $W'^{-} \to \bar{t}b$ decays which we generally denote as $W' \to tb$. This decay channel is of particular interest because the SM backgrounds can be greatly reduced when compared to W′ decays to light quarks, and some models predict a stronger W′ coupling to third generation quarks [16]. Here, the hadronic decay channel ($W' \to tb \to qqbb$) is presented.

The most general, lowest dimension effective Lagrangian that describes the interaction of the W′ boson with quarks [17] can be written as:

$$\mathcal{L} = \frac{V_{q_i q_j}}{2\sqrt{2}} g_w \overline{q}_i \gamma_\mu \left(a^R_{q_i q_j}(1+\gamma^5) + a^L_{q_i q_j}(1-\gamma^5)\right) W' q_j + \text{h.c.}\,, \tag{1}$$

where the parameters $a^L_{qi,qj}$ and $a^R_{qi,qj}$ represent the left-handed and right-handed couplings of the W′ boson to quarks, g_w is the SM weak coupling constant, and $V_{q_i q_j}$ is the Cabibbo-Kobayashi-Maskawa matrix. For this analysis, we investigate the case of a right-handed W′ boson with using the parameters $a^L_{qi,qj} = 0$ and $a^R_{qi,qj} = 1$.

In the kinematic region of interest ($M_{W'} \gtrsim 1.2$ TeV), the top quark is highly energetic ($p_T \gtrsim 350\,\text{GeV}$). Due to the Lorentz boost, the angular separation between the top quark decay products (W boson and b quark) is small. The final-state particles resulting from the hadronization of the b quark and the decay of the W boson into light quarks usually overlap, resulting in a single jet with top flavour, the "top quark jet", or t jet. Dedicated methods are applied to resolve the substructure of this t jet that strongly suppress background processes. We apply b jet identification algorithms (b tagging) to the b jet from the W′ decay in order to further reduce the SM background.

We reconstruct the W′ boson mass using the invariant mass of the top and bottom quarks (M_{tb}), and use the measured M_{tb} distribution to derive limits on the production cross section of the W′ boson.

2 CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the superconducting solenoid volume are silicon pixel and strip tracker detectors, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two endcap sections. Forward calorimeters extend the pseudorapidity [18] coverage provided by the barrel and endcap detectors. Muons are measured in gas-ionization detectors embedded in the steel flux-return yoke outside the solenoid.

In the region $|\eta| < 1.74$, the HCAL cells have widths of 0.087 in pseudorapidity and 0.087 in azimuth (ϕ). In the η-ϕ plane, and for $|\eta| < 1.48$, the HCAL cells map on to 5×5 ECAL crystals arrays to form calorimeter towers projecting radially outwards from close to the nominal interaction point. At larger values of $|\eta|$, the size of the towers increases and the matching ECAL

arrays contain fewer crystals. Within each tower, the energy deposits in ECAL and HCAL cells are summed to define the calorimeter tower energies, subsequently used to provide the energies and directions of hadronic jets.

The particle-flow (PF) event algorithm [19, 20] reconstructs and identifies each individual particle with an optimized combination of information from the various elements of the CMS detector. The energy of photons is directly obtained from the ECAL measurement, corrected for zero-suppression effects. The energy of electrons is determined from a combination of the electron momentum at the primary interaction vertex (vertex with the largest sum of the squared p_T of the clustered objects), the energy of the corresponding ECAL cluster, and the energy sum of all bremsstrahlung photons spatially compatible with originating from the electron track. The momentum of muons is obtained from the curvature of the corresponding track. The energy of charged hadrons is determined from a combination of their momentum measured in the tracker and the matching ECAL and HCAL energy deposits, corrected for zero-suppression effects and for the response function of the calorimeters to hadronic showers. Finally, the energy of neutral hadrons is obtained from the corresponding corrected ECAL and HCAL energy. When combining information from the entire detector, the jet energy resolution amounts typically to 15% at 10 GeV, 8% at 100 GeV, and 4% at 1 TeV, to be compared to about 40%, 12%, and 5% obtained when the ECAL and HCAL calorimeters alone are used.

Jet momentum is determined as the vectorial sum of all particle momenta in the jet, and is found from simulation to be within 5 to 10% of the true momentum over the whole p_T spectrum and detector acceptance. An offset correction is applied to jet energies to take into account the contribution from additional proton-proton interactions within the same or nearby bunch crossings. Jet energy corrections are derived from simulation, and are confirmed with in situ measurements of the energy balance in dijet and photon + jet events. Additional selection criteria are applied to each event to remove spurious jet-like features originating from isolated noise patterns in certain HCAL regions.

A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [18].

3 Event samples

The data used for this analysis correspond to an integrated luminosity of 2.55 fb^{-1} of pp collisions provided by the LHC at a centre-of-mass energy of 13 TeV. We select events online using a trigger algorithm that requires the scalar p_T sum of reconstructed jets in the detector to be greater than 800 GeV. Additionally, an event can be selected if it passes a dijet trigger that requires at least one jet to pass trimmed jet mass and b tagging selections that are looser than the offline selection designed to identify top and b jets respectively. This trigger additionally requires one jet with $p_T > 280$ GeV, and one jet with $p_T > 200$ GeV.

For SM top-quark pair production ($t\bar{t}$) and single top production, we use the Powhegv2 [21] generator, and for QCD multi-jet estimation we use Madgraph5 v1.5.11 [22]. These samples are interfaced with PYTHIA 8.2 for showering and hadronization, use the CUET8M1 underlying event tune, and use the NNPDF 3.0 parton distribution function (PDF) sets. The SM $t\bar{t}$ and single top Monte Carlo samples are used for background estimation when setting limits, but the QCD samples are only used for cross checks of the self-consistency of the data driven QCD background estimation procedure. The signal samples are generated with the CompHEP 4.5.2rc10 [23] package, which is used for the leading order cross section calculation, which is then scaled to next-to-leading order using a factor of 1.2 [17]. Signal samples are generated

using W′ masses between 1.2 and 3.0 TeV with cross sections from 2.17 to 0.02 pb, and use the CTEQ6l1 PDFs. Simulated events are generated with additional proton-proton interactions in the same bunch crossing (pileup). Detector simulation for all generated samples is performed using GEANT4 [24].

The signal samples use the right-handed coupling, and are generated using various W′ mass hypotheses. The signal resonance width is set to 3% of the W′ mass. We re-weight all MC samples to account for differences due to proton-proton interactions from sub-leading primary vertices (pileup) by comparing MC truth information and the expected number of interactions from data.

4 Event reconstruction

Jets are reconstructed using the anti-k_T (AK) jet clustering algorithm with a distance parameter of 0.8 (AK8 jets) as implemented by FastJet [25, 26] to cluster PF candidates into jets. This algorithm clusters constituents (the reconstructed PF candidates in each event) to form jets based on the angular distance between them and their p_T. We select jets with $p_T > 350$ GeV and $|\eta| < 2.4$ in order to achieve a high trigger efficiency in the analysis.

We use the charged hadron subtraction method [27] to remove charged hadrons that originate from a non-leading vertex prior to the application of the jet clustering algorithm.

4.1 t tagging algorithm

The W′ decay produces a final state with two high p_T back-to-back jets. One of these jets forms from the merged top quark decay products and the other results from the hadronisation of a bottom quark.

When the W boson decays to quarks which hadronize, the top quark can be detected as three jets. The high boost of the top quark from a W′ boson decay causes the three jets to merge into one large jet with a distinct substructure. The CMS t tagging algorithm [28] discriminates signal from background by using this characteristic substructure. We use the 0.3% mis-tagging rate working point of the t tagging algorithm, which is based on the following Variables:

- **Jet mass** - The softdrop algorithm [29] declusters the AK8 jet until the following requirement is met:

$$\frac{min(p_{T1}, p_{T2})}{p_{T1} + p_{T2}} > z(\Delta R_{12}/R_0)^{\beta}. \tag{2}$$

  For this analysis we use z = 0.1 and β = 0, and require the mass of the softdrop declustered jet to be between 110 and 210 GeV in order to put the mass in the range of a top quark.
- **N-subjettiness** - The N-subjettiness algorithm [30] defines τ_N variables, which describe the consistency between the jet energy and the number of assumed subjets, N:

$$\tau_N = \frac{1}{d}\sum_i p_{T_i} \min\{\Delta R_{1,i}, \Delta R_{2,i}, ..., \Delta R_{N,i}, \} \tag{3}$$

  where $\Delta R_{J,i}$ is the angular distance ($\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$) measured between the subjet candidate (J) axis and a specific constituent particle (i), and d is the normalization

p_T region	Scale Factor
$p_T < 550\,GeV$	0.88 ± 0.11
$p_T > 550\,GeV$	1.00 ± 0.23

Table 1: The top tagging scale factor as parametrized in p_T.

factor

$$d = \sum_i p_{T_i} R, \tag{4}$$

where R is the characteristic distance parameter used by the jet clustering algorithm. A jet with energy consistent with N subjets will typically have a low τ_N variable. A t jet is more consistent with three subjets than two (when compared to jets originating from a gluon or a light quark), therefore the ratio of τ_3 and τ_2 allows top jets to be distinguished from QCD background. We select events with $\tau_3/\tau_2 < 0.61$.

- **Subjet b tagging** - We apply the combined secondary vertex [31] (CSVIVFv2) b tagging algorithm to all of the subjets found by the t tagging algorithm. We require the maximum discriminator value ($SJ_{CSVIVFv2MAX}$) to register a value of at least 0.76.

We use a MC-to-data scale factor to account for differences in top tagging between data and simulation (Table 1). This scale factor is extracted using a semileptonic $t\bar{t}$ selection, and takes into account known data/MC discrepancies due to all t tagging variables described above.

4.2 b jet identification

To identify the b-quark daughter of the W′ boson, we start with the b-candidate jet, which is the highest p_T jet that is hemispherically separated from the top-tagged jet. We apply the loose operating point of the CSVIVFv2 algorithm to this b-candidate jet. A data-derived scale factor is applied to correct for differences in b tagging efficiency between data and simulation. We apply a scale factor for b jets extracted for the loose operating point to re-weight our MC samples to better agree with data.

For a W′ signal event, the mass of this b-candidate jet reflects the mass of the b quark, whereas the b-candidate jet identified in SM $t\bar{t}$ background generally includes non-b components from the top decay, and is commonly reconstructed close to the W boson or top quark mass. Therefore, SM $t\bar{t}$ background is greatly reduced by requiring the softdrop mass of the b-candidate jet to be below 70 GeV. This selection reduces $t\bar{t}$ by 83% while removing around 35% of the signal.

The difference in rapidity ($|\Delta y|$) between jets emerging from the decay of a heavy resonance is lower than for QCD in the high invariant mass region, and we require $|\Delta y|$ less than 1.3. This selection increases the sensitivity of the analysis in the high M_{tb} region and has been optimized based on the expected limit.

4.3 Reconstruction of W' mass

We select candidate $W' \to tb$ events by using the following criteria, which is applied to the two leading jets:

- one jet with $p_T > 350\,GeV$ identified with the CMS t tagging algorithm,
- one jet with $p_T > 350\,GeV$ with a b tag at the loose operating point and softdrop mass $< 70\,GeV$,
- the two jets are in opposite hemispheres ($|\Delta\phi| > \pi/2$),

- the difference in rapidity between the two jets is less than 1.3.

The number of events remaining after each successive selection in data, data-derived QCD, simulated $t\bar{t}$ events, simulated tW-channel single top events, and right-handed W′ boson signal MC is shown in Table 2.

Table 2: Numbers of observed and expected events at successive stages of the event selection. The expected numbers are scaled to an integrated luminosity of 2.55 fb^{-1}. Statistical uncertainties in the event yields are quoted. The QCD background contribution is only reported for the final selection due to the fact that the data driven QCD estimate requires the b tagging and subjet b tagging selections be applied. The signal events, shown for several values of the W′ boson mass, are normalized to the theoretical cross section. Scale factors are applied to rows after the corresponding selection is applied.

Selection	Data	QCD	$t\bar{t}$	Single top	$M_{W'_R}$ = 1400 GeV	$M_{W'_R}$ = 2000 GeV	$M_{W'_R}$ = 2600 GeV
$p_T > 350\,GeV$	5896714±2428.3	–	21325±74.7	1808±20.2	2257±11.1	522±2.1	133±0.5
$\|\Delta y\| < 1.3$	3562920±1887.6	–	14950±62.5	1396±17.8	1613±9.4	315±1.6	75±0.4
$110\,GeV < M_t < 210\,GeV$	965506±982.6	–	11174±54.1	712±12.9	1005±7.4	186±1.3	42±0.3
$SJ_{CSVMAX} > 0.76$	253963±503.9	–	8826±48.0	512±11.0	787±6.6	138±1.1	29±0.2
τ_3/τ_2	41599±204.0	–	4165±31.8	199±6.4	452±4.9	75±0.9	15±0.2
$M_b < 70\,GeV$	27499±165.8	–	1204±17.2	87±4.2	342±4.3	54±0.7	11±0.1
CSV	6811±82.5	6377±48	411±9.9	36±2.8	242±3.6	29±0.5	5±0.1

5 Background modeling

The primary sources of background are SM QCD multi-jet and $t\bar{t}$ production. This is due to the abundance of QCD multi-jet background after selecting an all-jet final state and the large contribution from $t\bar{t}$ production that remains after t-jet discrimination criteria are applied.

The QCD multi-jet production background is estimated using a data-driven technique to extract both the shape and normalization. The average b tagging pass-to-fail ratio (average b tagging rate), measured from events with an enhanced QCD multi-jet component and small signal contamination component, is used to estimate the QCD multi-jet contribution in the signal region.

5.1 QCD background estimate

In order to measure the average b tagging rate for QCD jets, a control sample is obtained by re-defining the selection criteria used to identify t jets:

$$50 < M_{jet} < 170\ GeV; \tag{5}$$

$$\tau_3/\tau_2 > 0.75; \tag{6}$$

$$SJ_{CSVIVFv2MAX} \geq 0.76. \tag{7}$$

The shifted top mass window along with the inverted N-subjettiness requirement ensure a low signal contamination ($< 1\%$), while the subjet b tagging discrimination is kept in this control sample in order to ensure similar parton flavour distributions in the signal and control regions. The b tagging criteria is applied to the b-candidate jet in this control region to define the numerator of the average b tagging rate for QCD jets. Because of the similar parton flavour distributions, and the fact that this region has a low $t\bar{t}$ and signal contamination component,

this control sample is an ideal selection to extract the average b tagging rate. The average b tagging rate is parameterized as a function of the p_T of the b-candidate jets (which explicitly fail the b tagging requirement) in three $|\eta|$ regions:

- Low $(0.0 < |\eta| \leq 0.5)$
- Transition $(0.5 < |\eta| \leq 1.15)$
- High $(1.15 < |\eta| \leq 2.4)$

Events that pass all of the signal region selection except fail the b tagging requirement are then weighted by this average b tagging rate to estimate the QCD background contribution in the signal region. Events in the signal region that fail b tagging are largely from QCD background, but the small $t\bar{t}$ background component is subtracted when deriving the QCD background contribution to avoid double counting.

We use a bifurcated polynomial to fit the average b tagging rate in each of these $|\eta|$ ranges. This fitting function, which provides a satisfactory description of the data, is defined as follows:

$$f(x) = \begin{cases} p_0 + p_1 x + p_2 (x-a)^2, & \text{if } x < a \\ p_0 + p_1 x + p_3 (x-a)^2, & \text{if } x \geq a \end{cases} \tag{8}$$

Here, the parameters p_0 to p_3 are the polynomial coefficients, and x is the p_T of the b-candidate jet. The parameter a is the bifurcation point, and is optimized for each region in $|\eta|$. The average b tagging rate and fit can be seen in Figure 1.

The uncertainty in the average b tagging rate is extracted using the full covariance matrix obtained from the output of the fitting algorithm. Additionally, we assign a systematic uncertainty to cover the choice of the fit function (see Section 6.1) based on several alternative functional forms (such as second degree polynomial or exponential functions).

There is a correlation between b tagging and the softdrop mass of the b-candidate jet, which creates a shape discrepancy. The b-candidate mass shape is corrected by taking the ratio of the b tagging pass and fail templates in QCD MC sideband region and using this ratio to weight the QCD background template. The change to the M_{tb} spectrum is small (around a 3% efffect) and the uncertainty in this method is taken as half of the difference of the unweighted and weighted background templates.

5.2 Background estimate validation

To investigate the applicability of the QCD background estimation in data, we apply the procedure first to QCD MC. For this, we extract the average b tagging rate from QCD MC, and apply it to the pre b tagged signal region selection in QCD MC. The agreement from this closure test is investigated by using QCD MC in place of data when forming the M_{tb} distribution as can be seen in Figure 2.

We also investigate the agreement in data by applying the average b tagging rate to a control region defined by inverting the subjet b tagging selection used for top tagging. This region is orthogonal to both the signal region and the control region that is used for the determination of the average b tagging rate, and has a very low yield of $t\bar{t}$ production, which allows for a precise measurement of the QCD background contribution. The average b tagging rate used for this closure test is extracted from the same control region that is used for extracting the average b tagging rate in the signal region, except the subjet b tagging selection is inverted.

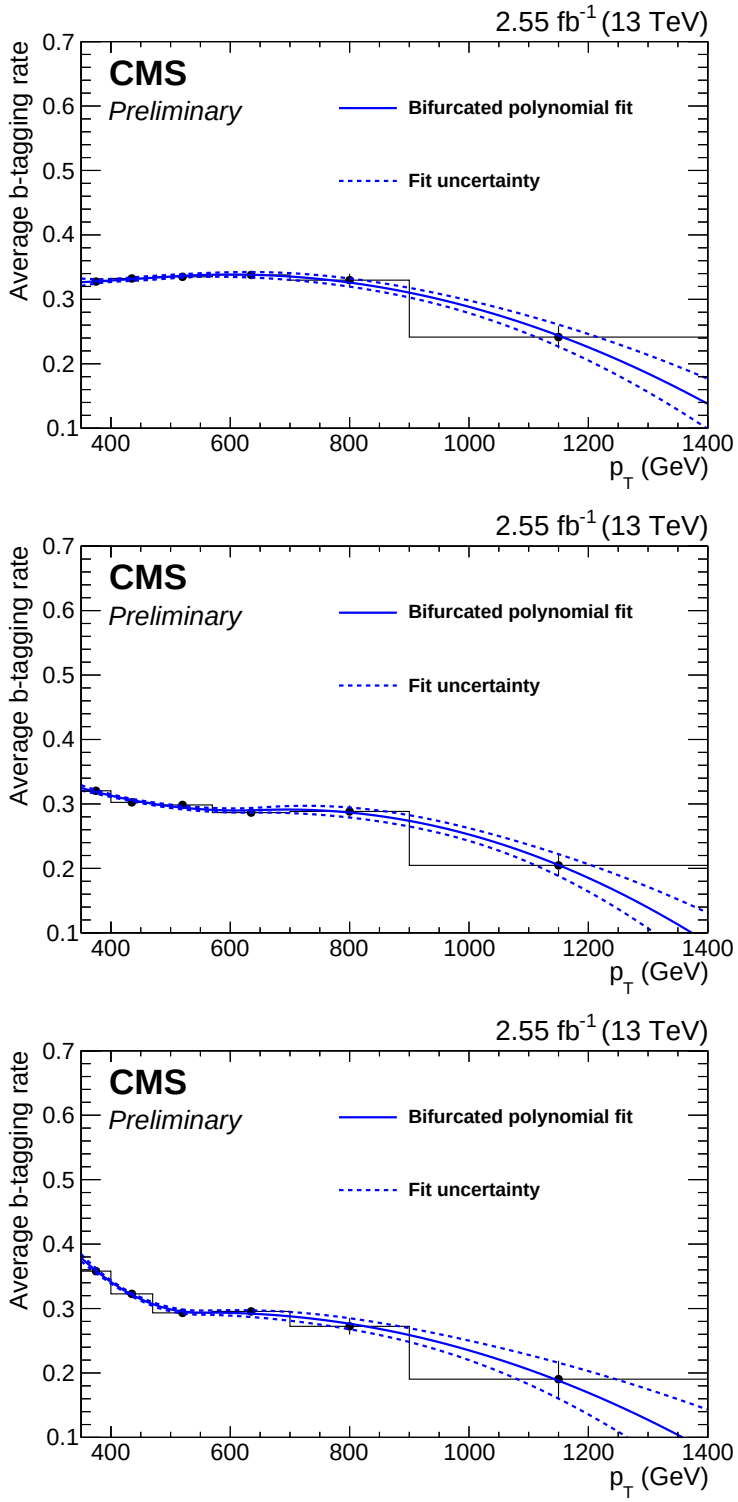

Figure 1: The average b tagging rate for QCD jets parameterized as a function of p_{T} from the low (top), transition (middle), and high (bottom) $|\eta|$ regions. The measured average b tagging rate is represented by the data points, the polynomial fit is shown as a solid line, and the propagated uncertainties from the fit are shown as the dashed lines.

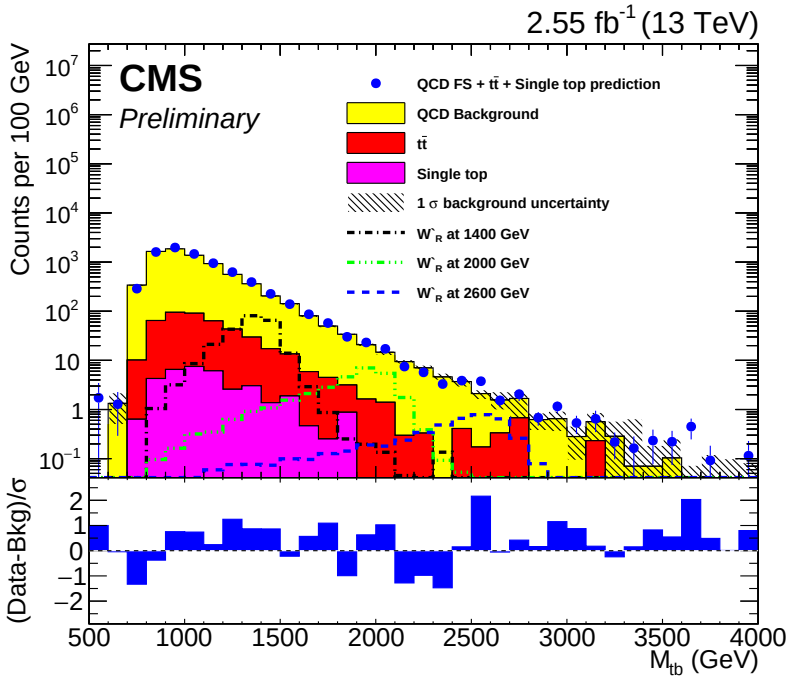

Figure 2: MC-only comparison of M_{tb} in the signal region for demonstrating that the QCD background shape can be determined correctly using the b tagging rate method described in the text. The data points are the summed QCD, single top, and $\mathrm{t\bar{t}}$ signal regions, representing an approximation of the signal region in data given the null signal hypothesis. The bottom plot shows the pull ((data-background)/σ) between the data and the background estimate distributions. All distributions are normalized to the expected number of events.

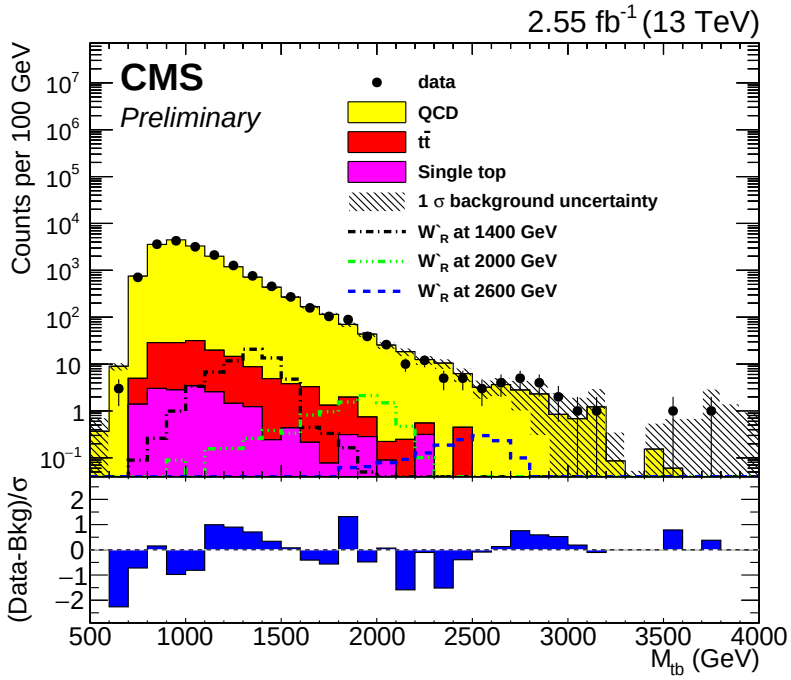

Figure 3: Distribution of M_{tb} in the inverted subjet b tag control region defined in the text. The bottom plot shows the pull ((data-background)/σ) between the data and the background estimate distributions. All distributions are normalized to the expected number of events.

The agreement from this closure test is investigated by forming the M_{tb} distribution in this control region as can be seen in Figure 3.

After investigating agreement and extracting a correction with the QCD closure tests, we proceed to characterize the agreement of the background contribution from simulated $\mathrm{t\bar{t}}$ events using a control sample in data. A sideband region is defined by inverting the b-candidate mass requirement (see Section 4.2) in the signal region. This region has an enhanced fraction of $\mathrm{t\bar{t}}$ events and is statistically independent from all other sidebands in the analysis.

We compare the sum of the QCD background estimate from data and the SM $\mathrm{t\bar{t}}$ contribution obtained from MC to the observed yield in data. Good agreement is observed as can be seen in Figure 4.

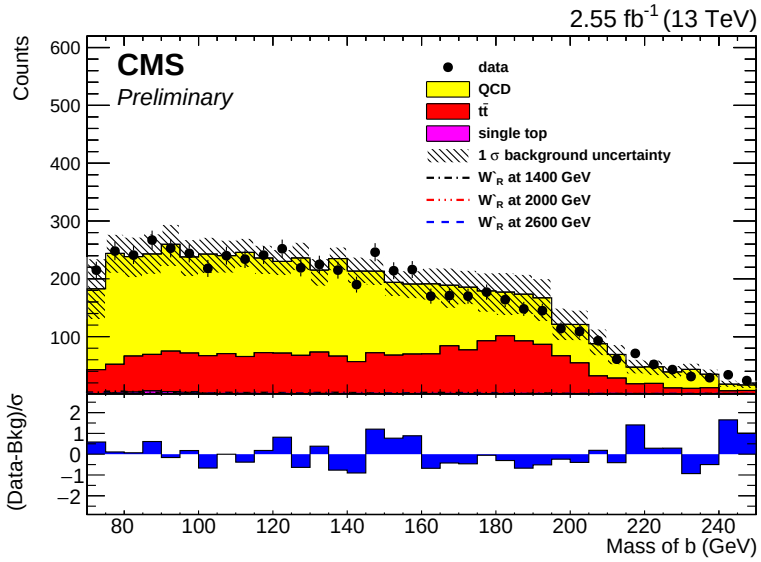

Figure 4: Distribution of b candidate softdrop mass in the inverted b candidate softdrop mass control region defined in the text. This region is used to investigate the agreement of data and background, specifically with respect to the MC $t\bar{t}$ estimate, which has an increased fraction in this region. The bottom plot shows the pull ((data-background)/σ) between the data and the background estimate distributions. All distributions are normalized to the expected number of events.

6 Results

6.1 Systematic uncertainties

The primary source of systematic uncertainty on the background estimate results from the data driven QCD background estimation method. We use a fit to the average b tagging rate to evaluate the QCD background prediction, and the uncertainty in this fit is used as a QCD background shape uncertainty. The functional form of this fit is a subjective choice, and the uncertainty on this choice is extracted from studying alternative functional forms and is evaluated as the mean squared error of the resulting M_{tb} spectra. The uncertainty in the b softdrop mass correction mentioned in Section 5.1 is taken as half of the difference of the unweighted and weighted background templates. We extract a correction based on residual differences in the closure test of the background estimate when investigated in QCD MC. The uncertainty on this correction is taken as the 100% of the M_{tb} change, and corresponds to approximately a 4% rate effect.

The uncertainty arising from the variation of the Q^2 factorization and renormalization scale in $\mathrm{t\bar{t}}$ production is evaluated by investigating simulated $\mathrm{t\bar{t}}$ samples generated with twice and half of the nominal Q^2 scale to derive the $\pm 1\sigma$ template shapes. This amounts to an approximately 5% rate effect.

The b tagging scale factor $\pm 1\sigma$ variations are used to estimate the b tagging scale factor uncertainty for $\mathrm{t\bar{t}}$ production and signal MC which amounts to an approximately 9% rate effect. We include an additional 3% uncertainty due to the fact that AK8 jets are used in the analysis as opposed to AK4 jets, which the scale factor and uncertainty are derived for.

To estimate the uncertainty in the jet energy scale, we use the $\pm 1\sigma$ variations on jet energy scale, which are η and p_{T} dependent, and result in a scaling of the jet four momentum prior to the analysis selection. The effect of this uncertainty is a change in shape that causes a rate change at low M_{tb} due to jets falling below the kinematic threshold.

In the analysis we apply the approved jet energy resolution correction, which is η dependent and is used to correct for differences in p_{T} resolution between data and Monte Carlo. The recommended $\pm 1\sigma$ variations are also taken into into account.

To correct for differences in data an Monte Carlo pileup distributions, a correction and uncertainty is evaluated by comparing the expected number of interactions in data to the true number of interactions in Monte Carlo and varying the minimum bias cross section by $\pm 5\%$.

We estimate the uncertainty arising from PDFs in a shape dependent manner by evaluating the RMS of the distribution of the 100 NNPDF MC replicas.

The uncertainties in the top tagging scale factor used in this analysis is listed in Table 1. The uncertainty is p_{T} dependent, but we take the larger of the two regions as a conservative rate uncertainty.

We also include the recommended 2.7% uncertainty on the measured luminosity of 2.55 fb.

Finally, the uncertainty in the trigger turn-on efficiency that is applied to the MC samples is taken as one half of the trigger inefficiency.

A summary of the systematic uncertainties used in the analysis can be found in table 3.

Table 3: Sources of uncertainty that affect the M_{tb} distribution that are taken into account when setting limits. Because of the low yield, shape uncertainties are neglected for single top.

Source	Variation	Samples
$\mathrm{t\bar{t}}$ cross section	+4.8%,-5.5%	$\mathrm{t\bar{t}}$
single top cross section (tW)	$\pm$5.4%	single top
Luminosity	$\pm$2.7%	$\mathrm{t\bar{t}}$, single top, signal
t tagging	$\pm$23%	$\mathrm{t\bar{t}}$, single top, signal
AK4 to AK8 b tagging	$\pm$3%	$\mathrm{t\bar{t}}$, single top, signal
Pileup	$\pm 1\sigma$ (σ_{mb})	$\mathrm{t\bar{t}}$, signal
PDF	$\pm 1\sigma$ (x, Q^2)	$\mathrm{t\bar{t}}$, signal
Jet Energy Scale	$\pm 1\sigma(p_{\mathrm{T}})$	$\mathrm{t\bar{t}}$, signal
Jet Energy Resolution	$\pm 1\sigma(p_{\mathrm{T}}, \eta)$	$\mathrm{t\bar{t}}$, signal
Q^2 Scale	$\pm 1\sigma(Q^2)$	$\mathrm{t\bar{t}}$
b tagging	$\pm 1\sigma(p_{\mathrm{T}})$	$\mathrm{t\bar{t}}$, signal
Trigger	$\pm 1\sigma(H_T)$	$\mathrm{t\bar{t}}$, signal
Average b tagging rate fit	$\pm 1\sigma(p_{\mathrm{T}}, \eta)$	QCD (from data)
QCD MC closure	$\pm 1\sigma(M_{\mathrm{tb}})$	QCD (from data)
b candidate mass	$\pm 1\sigma(M_b)$	QCD (from data)
Alternate functional forms	$\pm 1\sigma(p_{\mathrm{T}}, \eta)$	QCD (from data)

6.2 Cross section limits

After investigating the agreement of the simulated $\mathrm{t\bar{t}}$ events estimate with data using a control region, and investigating the agreement of the QCD background estimate by using both a control region in data and the signal region in QCD MC, the background estimate is used to predict the W′ boson signal region in data. Since the data yields are consistent with SM expectations, we proceed to compute limits on the W′ boson cross section.

To set limits on the production cross section of the $\mathrm{W'_R}$ boson model described in Eq. 1, we compare, for each bin in the M_{tb} distribution, the numbers of expected events. The following expression is used to compute the expected contribution from $\mathrm{W'_R}$ boson production:

$$N_{\mathrm{expected}} = \sigma_{\mathrm{W'_R}} \times \mathcal{B}_{\mathrm{W'_R}\to\mathrm{tb;W}\to\mathrm{hadrons}} \times \varepsilon \times \int Ldt, \tag{9}$$

where $\sigma_{\mathrm{W'_R}}$ is the $\mathrm{W'_R}$ cross section, $\mathcal{B}_{\mathrm{W'_R}\to\mathrm{tb;W}\to\mathrm{hadrons}}$ is the branching fraction $\mathrm{W'_R} \to \mathrm{tb}$ with the W boson decay constrained to the hadronic branching fraction, ε is the signal efficiency and acceptance, and $\int Ldt$ is the integrated luminosity of the dataset. We perform a binned maximum likelihood fit to compare the M_{tb} distribution from data with the $\mathrm{W'_R}$ boson signal hypothesis, summed together with the SM distribution obtained from the background estimation procedure described in Section 5.

A Poisson model is used for each bin of the M_{tb} distribution. The mean of the Poisson distribution for each bin is taken to be:

$$\mu_{\mathrm{i}} = \sum_{\mathrm{k}} \beta_{\mathrm{k}} \times T_{\mathrm{k,i}}, \tag{10}$$

where k includes both the signal and background models, β_{k} is the Poisson mean for process k, and $T_{\mathrm{k,i}}$ represents the fraction of events expected for each process k in bin i.

Using Bayesian statistics with a flat prior for the signal cross section, we obtain 95% CL upper limits on the production cross section of W'_R. Pseudo-experiments are used to derive the $\pm 1\sigma$ deviations in the expected limit. The systematic uncertainties described above are accounted for as nuisance parameters and the posterior probability is refitted for each pseudo-experiment. The Theta package [32] is used to set cross section upper limits, which are shown in Figure 6.

7 Summary

A search for a new massive gauge boson W' decaying to a bottom quark and a hadronically decaying top quark is performed using pp collisions recorded by the CMS detector at $\sqrt{s} = 13$ TeV.

The analysis uses jet substructure algorithms to allow the top quark jet to be distinguished from standard model hadronic jet backgrounds. The main background from QCD multi-jet production is estimated from data using the average b tagging rate measured in a QCD-enhanced control region.

Limits are placed on the production cross section of a right-handed W' boson using 2.55 fb^{-1} of luminosity. This is the first search in this channel at 13 TeV, and a right-handed W' with mass below 2.0 TeV is excluded at 95% CL.

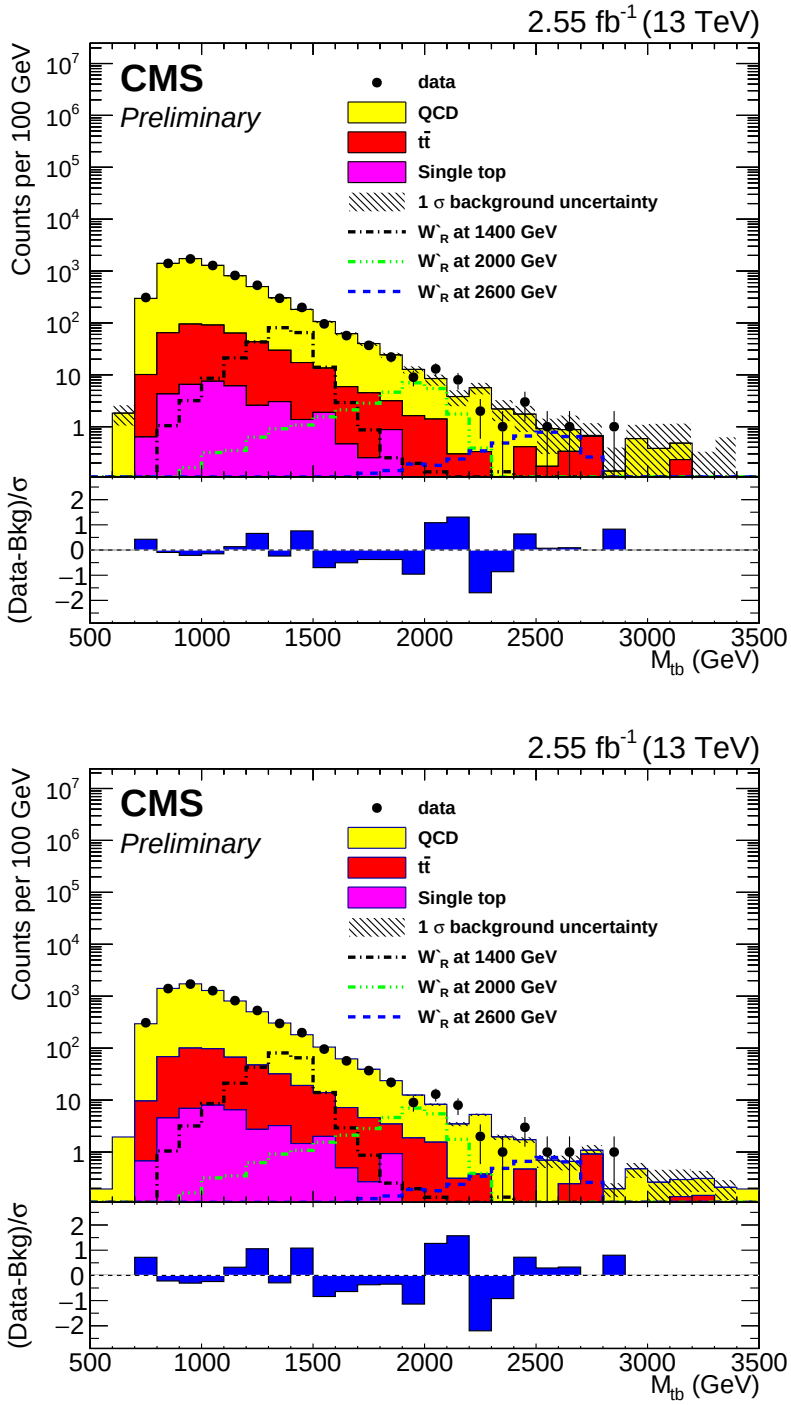

Figure 5: The distribution of M_{tb} in the signal region shown for data, data-derived QCD background, several MC background contributions, and several example signal W′ boson MC samples. The W′ signal and background MC samples are normalized to the cross section and the luminosity of the full dataset. The distributions are shown after the application of all selections, and are plotted using the background and uncertainty estimates from both before (top) and after (bottom) the limit setting fit.

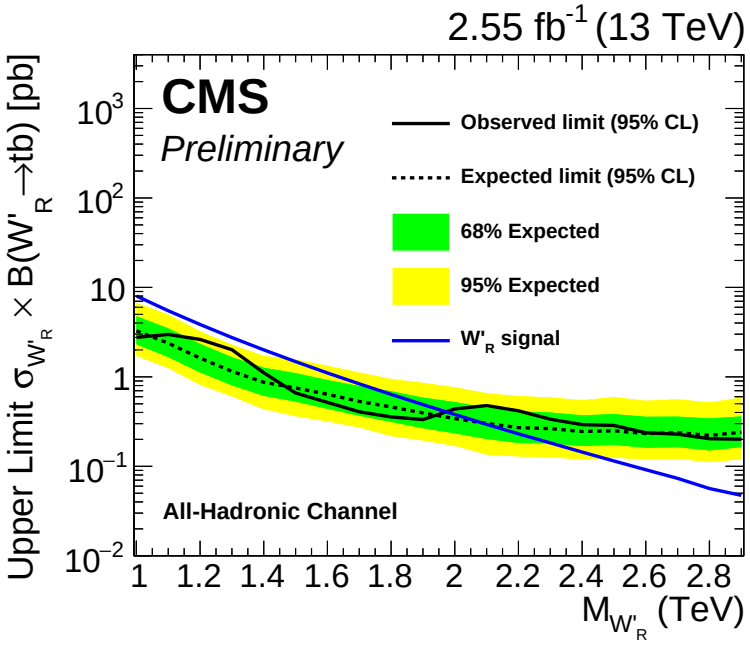

Figure 6: The W'_R boson 95% CL production cross section limits. The observed and expected limits, as well as W'_R boson theoretical cross section are shown. The uncertainty in the expected limit band represents the 68% and 95% confidence intervals.

References

[1] M. Schmaltz and D. Tucker-Smith, "Little Higgs Theories", *Annual Review of Nuclear and Particle Science* **55** (2005) 229, doi:10.1146/annurev.nucl.55.090704.151502, arXiv:hep-ph/0502182.

[2] T. Appelquist, H.-C. Cheng, and B. A. Dobrescu, "Bounds on universal extra dimensions", *Phys. Rev. D* **64** (2001) 035002, doi:10.1103/PhysRevD.64.035002, arXiv:hep-ph/0012100.

[3] R. N. Mohapatra and J. C. Pati, "Left-right gauge symmetry and an 'isoconjugate' model of CP violation", *Phys. Rev. D* **11** (1975) 566, doi:10.1103/PhysRevD.11.566.

[4] D0 Collaboration, "Search for W' Boson Resonances Decaying to a Top Quark and a Bottom Quark", *Phys. Rev. Lett.* **100** (2008) 211803, doi:10.1103/PhysRevLett.100.211803, arXiv:0803.3256.

[5] D0 Collaboration, "Search for $W' \to tb$ resonances with left- and right-handed couplings to fermions", *Phys. Lett. B* **699** (2011) 145, doi:10.1016/j.physletb.2011.03.066, arXiv:1101.0806.

[6] CMS Collaboration, "Search for $W' \to tb$ decays in the lepton + jets final state in pp collisions at $\sqrt{s}$ = 8 TeV", *JHEP* **05** (2014) 108, doi:10.1007/JHEP05(2014)108, arXiv:1402.2176.

[7] CMS Collaboration, "Search for physics beyond the standard model in final states with a lepton and missing transverse energy in proton-proton collisions at $\sqrt{s}$ = 8 TeV", (2014). arXiv:1408.2745.

[8] CMS Collaboration, "Search for new resonances decaying via WZ to leptons in proton-proton collisions at $\sqrt{s} = 8$ TeV", *Phys. Lett. B* **740** (2015) 83, doi:10.1016/j.physletb.2014.11.026, arXiv:1407.3476.

[9] ATLAS Collaboration, "ATLAS search for new phenomena in dijet mass and angular distributions using pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **01** (2013) 029, doi:10.1007/JHEP01(2013)029, arXiv:1210.1718.

[10] ATLAS Collaboration, "Search for resonant diboson production in the WW/WZjj decay channels with the ATLAS detector at $\sqrt{s} = 7$ TeV", *Phys. Rev. D* **87** (2013) 112006, doi:10.1103/PhysRevD.87.112006, arXiv:1305.0125.

[11] ATLAS Collaboration, "Search for $W' \to t\bar{b}$ in the lepton plus jets final state in proton-proton collisions at a centre-of-mass energy of $\sqrt{s}$ = 8 TeV with the ATLAS detector", *Phys. Lett. B* **743** (2015) 235–255, doi:10.1016/j.physletb.2015.02.051, arXiv:1410.4103.

[12] ATLAS Collaboration, "Search for $W' \to tb \to qqbb$ decays in pp collisions at $\sqrt{s}$ = 8 TeV with the ATLAS detector", *Eur. Phys. J. C* **75** (2015), no. 4, 165, doi:10.1140/epjc/s10052-015-3372-2, arXiv:1408.0886.

[13] ATLAS Collaboration, "ATLAS search for a heavy gauge boson decaying to a charged lepton and a neutrino in pp collisions at $\sqrt{s} = 7$ TeV", *Eur. Phys. J. C* **72** (2012) 2241, doi:10.1140/epjc/s10052-012-2241-5, arXiv:1209.4446.

[14] CMS Collaboration, "Search for massive resonances in dijet systems containing jets tagged as W or Z boson decays in pp collisions at $\sqrt{s}$ = 8 TeV", *JHEP* **08** (2014) 173, doi:10.1007/JHEP08(2014)173, arXiv:1405.1994.

[15] CMS Collaboration, "Search for new physics in final states with a lepton and missing transverse energy in pp collisions at the LHC", *Phys.Rev.D* **87** (2013), no. 7, 072005, doi:10.1103/PhysRevD.87.072005, arXiv:1302.2812.

[16] D. J. Muller and S. Nandi, "Top flavor: A Separate SU(2) for the third family", *Phys. Lett. B* **383** (1996) 345, doi:10.1016/0370-2693(96)00745-9, arXiv:hep-ph/9602390.

[17] Z. Sullivan, "Fully differential W' production and decay at next-to-leading order in QCD", *Phys. Rev. D* **66** (2002) 075011, doi:10.1103/PhysRevD.66.075011, arXiv:hep-ph/0207290.

[18] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **3** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[19] CMS Collaboration, "Particle–Flow Event Reconstruction in CMS and Performance for Jets, Taus, and $E_{\mathrm{T}}^{\mathrm{miss}}$", CMS Physics Analysis Summary CMS-PAS-PFT-09-001, 2009.

[20] CMS Collaboration, "Commissioning of the Particle-flow Event Reconstruction with the first LHC collisions recorded in the CMS detector", CMS Physics Analysis Summary CMS-PAS-PFT-10-001, 2010.

[21] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[22] J. Alwall et al., "MadGraph 5 : Going Beyond", *JHEP* **06** (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522.

[23] CompHEP Collaboration, "CompHEP 4.4: Automatic computations from Lagrangians to events", *Nucl.Instrum.Meth. A* **534** (2004) 250, doi:10.1016/j.nima.2004.07.096, arXiv:hep-ph/0403113.

[24] G. Collaboration, "Geant4a simulation toolkit", *Nucl. Instrum. Meth.* **506** (2003), no. 3, 250 – 303, doi:http://dx.doi.org/10.1016/S0168-9002(03)01368-8.

[25] M. Cacciari and G. P. Salam, "Dispelling the N**3 myth for the k(t) jet-finder", *Phys. Lett. B* **641** (2006) 57, doi:10.1016/j.physletb.2006.08.037, arXiv:hep-ph/0512210.

[26] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet User Manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[27] CMS Collaboration, "Pileup Removal Algorithms", CMS Physics Analysis Summary CMS-PAS-JME-14-001, 2014.

[28] CMS Collaboration, "Top Tagging with New Approaches", CMS Physics Analysis Summary CMS-PAS-JME-15-002, 2016.

[29] A. J. Larkoski, S. Marzani, G. Soyez, and J. Thaler, "Soft drop", *JHEP* **2014** (2014), no. 5, 1–46, doi:10.1007/JHEP05(2014)146.

[30] J. Thaler and K. Van Tilburg, "Maximizing Boosted Top Identification by Minimizing N-subjettiness", *JHEP* **02** (2012) 093, doi:10.1007/JHEP02(2012)093, arXiv:1108.2701.

[31] 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, 2016.

[32] T. Muller, J. Ott, and J. Wagner-Kuhr, "theta - a framework for template-based modeling and inference", CMS Internal Note 2010/017, 2010.