CMS Physics Analysis Summary

Contact: cms-pag-conveners-exotica@cern.ch 2018/06/01

Search for pair-produced resonances decaying to quark pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV

The CMS Collaboration

Abstract

A search for the pair production of resonances decaying to two quarks is reported. The search is conducted separately for lighter resonances between 80 and 400 GeV in mass, when the resulting diquark decay products are collimated and reconstructed as a single jet producing a dijet final state, and for heavier resonances above 400 GeV in mass, when the decay products generate pairs of hadronic jets producing a four-jet final state. In addition, a b-tagged selection is applied to target resonances with a bottom quark in the final state. The analysis uses data collected with the CMS detector at the LHC, corresponding to an integrated luminosity of 35.9 fb^{-1} from proton-proton collisions at a center-of-mass energy of 13 TeV. The mass spectra are analyzed for the presence of new resonant particles, and are found to be consistent with standard model expectations. The results are interpreted in the framework of R-parity-violating supersymmentry assuming the pair production of scalar top quarks decaying via the λ''_{312} or the λ''_{323} hadronic couplings, and upper limits are placed on the pair production cross section of top squarks as a function of the top squark mass for the two hadronic coupling scenarios.

1 Introduction

New particles which decay into quarks and gluons and produce fully hadronic signatures are predicted in many models of physics beyond the standard model [1–3]. For instance, the violation of baryon number in some supersymmetric (SUSY) models leads to colored superpartners producing hadronic final states [4]. We report on a search for pair-produced resonances decaying to two light quarks (qq′) or one light quark and one bottom quark (bq′).

Minimal SUSY models introduce R parity, associated with a Z_2 symmetry group called R symmetry, to naturally forbid terms in the supersymmetric potential that lead to the violation of baryon or lepton numbers [5]. However, if R-parity is violated, some SUSY models allow supersymmetric Yukawa interactions through the terms

$$W_{\not R_p} = \frac{1}{2}\lambda_{ijk} L_i L_j E^c_k + \lambda'_{ijk} L_i Q_j D^c_k + \frac{1}{2}\lambda''_{ijk} U^c_i D^c_j D^c_k + \mu_i H_u L_i, \quad (1)$$

where λ, λ', λ'' are Yukawa couplings, and i, j, k are quark and lepton generation indices following the summation convention, while c denotes charge conjugation. The doublet superfields of the lepton and quark are denoted by L_i and Q_i, respectively, while the E_i, U_i and D_j represent the singlet superfields of the lepton, up-type and down-type quarks, respectively. The Higgs superfield is denoted by H_u and couples to up-type quarks. The first and third terms in Eq. (1) are antisymmetric in $\{i, j\}$ and $\{j, k\}$, respectively. The trilinear λ''_{ijk} couplings permit vertices of sfermions interacting with two fermions, and the bilinear μ_i coupling represents a dimensional mass parameter. In baryonic R-parity-violating (RPV) models, the only nonzero coupling in Eq. (1) is λ''_{ijk}, which lead to vertices of squarks interacting with two quarks.

We consider the pair production of top squarks ($\widetilde{\mathrm{t}}$) as a benchmark model, assuming the $\widetilde{\mathrm{t}}$ is the lightest of the colored SUSY partners and is allowed to decay via the baryonic RPV coupling to quarks. In this case, $\lambda''_{ijk} = \lambda''_{3\mathrm{DD}}$ and each index reflects the squark or quark generation of the process, two of which are down-type quarks. We study two possible choices of hadronic RPV coupling scenarios: $\widetilde{\mathrm{t}} \to \mathrm{qq}'$ through the coupling λ''_{312}, and $\widetilde{\mathrm{t}} \to \mathrm{bq}'$ through the coupling λ''_{323}. These two models are schematically depicted in Fig. 1.

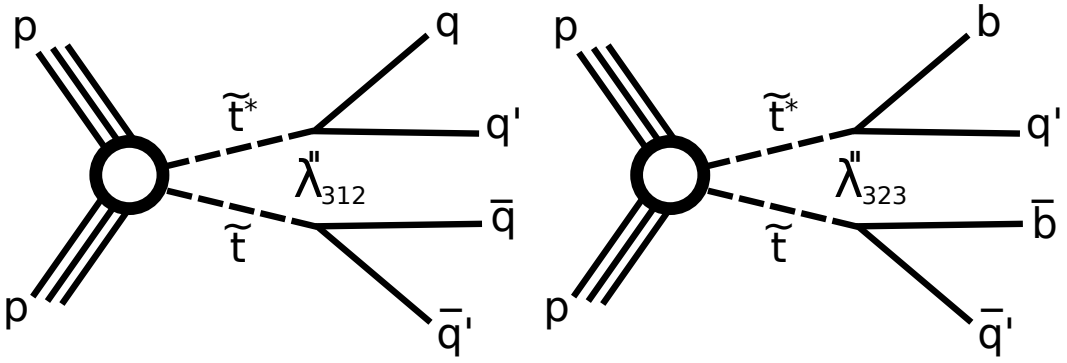

Figure 1: Diagrams for the benchmark models used in this analysis: pair production of top squarks decaying into qq′ via the RPV λ''_{312} coupling (left), and bq′ via the RPV coupling λ''_{323} (right).

Previous search for $\widetilde{\mathrm{t}} \to \mathrm{qq}'$ via RPV were performed by the ALEPH experiment at LEP [6] placing 95% confidence level (CL) limits on the production cross section as a function of mass, and later extended by the CDF experiment [7]. Similar searches have been performed at the LHC by both CMS [8, 9] at $\sqrt{s} = 7$ TeV and 8 TeV, and ATLAS [10–12] at $\sqrt{s} = 7$ TeV, 8 TeV and

13 TeV . For the $\widetilde{t} \to bq'$ scenario, mass exclusion limits at $\sqrt{s} = 8$ TeV have been reported by both CMS [9] and ATLAS [11], and at $\sqrt{s} = 13$ TeV by ATLAS [12].

The analysis reported in this result uses pp collision data at $\sqrt{s} = 13$ TeV collected with the CMS detector [13] at the LHC in 2016, corresponding to an integrated luminosity of 35.9 fb^{-1} [14]. The search is conducted in two mass ranges: for lighter resonances between 80 and 400 GeV the decay products are collimated enough to be reconstructed as a single jet (boosted search), and for heavier resonances above 400 GeV four jets are reconstructed in the final state (resolved search). Together they target resonant masses between 80 and 1500 GeV. In addition, we apply b tagging requirements to both searches, refer to the selection as b-tagged, and interpret the results in the $\widetilde{t} \to bq'$ scenario. When no b tagging is applied, we refer to the selection as inclusive, and interpret the results in the $\widetilde{t} \to qq'$ scenario.

The low-mass boosted search exploits the internal structure of the jets to differentiate between signal jets (two-pronged structure) and jets coming from quantum chromodynamics (QCD) multijet processes (predominantly with no internal structure). In this search, we use the average mass of the two jets with the highest transverse momentum (p_T), after removing soft and large-angle QCD multijet radiation, to look for evidence of a signal consistent with localized deviations from the estimated standard model (SM) backgrounds. The primary SM background component is QCD multijet events estimated from data control samples. Subdominant SM processes, such as the single and double production of W, Z, and top quarks decaying hadronically are also taken into account. The presence of these backgrounds create resonances in the mass spectrum, and we henceforth refer to them as resonant backgrounds. The subdominant resonant SM backgrounds are estimated with simulated samples.

For the resolved search, the high-mass resonances are produced with insufficient boost to be merged into single jets, and events with four individual high-p_T jets are selected. The dijet mass spectrum in this search is dominated by QCD multijets production. The mass spectrum is parameterized as a steeply falling smooth distribution that is explored for signal-like localized excesses.

2 The 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 solenoid volume are a silicon pixel and strip tracker, 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 coverage provided by the barrel and endcap detectors. Muons are detected in gas-ionization chambers embedded in the steel flux-return yoke outside the solenoid. Energy deposits from hadronic jets are measured using the ECAL and HCAL. Events of interest are selected using a two-tiered trigger system [15]. A 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. [13].

3 Simulated event samples

Top squark signal events are simulated using a hybrid of MADGRAPH5_aMC@NLO 2.2.2 [16] and PYTHIA 8.212 [17]. The calculation of the production of a pair of top squarks with up to two additional initial-state radiation jets is performed at leading order with MADGRAPH5_aMC@NLO, while PYTHIA is used for the prompt decay of each top squark to either $\widetilde{t} \to qq'$ or $\widetilde{t} \to bq'$

through the λ''_{3DD} hadronic RPV couplings. For the two coupling models considered, all other λ''_{UDD} couplings are set to zero so that the branching fraction to the desired quarks is 100%. The PYTHIA simulation is also used for the parton showering with the CUETP8M1 tune [18]. For each coupling, top squarks are generated with masses between 80 GeV and 1500 GeV, in 20 GeV increments up to 300 GeV, in 50 GeV steps up to 1 TeV, and in 100 GeV increments onwards. All other superpartners are set to higher masses in order not to produce intermediate sparticles in the top squark production. The natural width of the top squark is taken to be much smaller than the detector energy resolution.

Processes from QCD multijets are simulated and showered through PYTHIA using the CUETP8M1 tune [18]. Hadronically decaying W or Z boson production with additional jets from initial- and final-state radiation ($W\to q'\bar{q}$ +jets or $Z\to q\bar{q}$ +jets), and ZZ diboson samples are generated with MADGRAPH5_aMC@NLO, WZ processes are generated with PYTHIA, and $t\bar{t}$ +jets and WW samples are generated with POWHEG v2 [19, 20]. For $W\to q'\bar{q}/Z\to q\bar{q}$ +jets events, higher-order p_T-dependent electroweak NLO corrections are applied to improve the modeling of the kinematic distributions of those events [21–25].

Additional pp interactions in the same bunch crossings is referred to as pileup. The average number of pileup interactions was 27 for the 2016 data taking. Simulated signal and background samples are weighted in order to have the same pileup profile as the recorded data by superimposing minimum bias interactions onto all simulated events. The simulation of the CMS detector for all samples is handled by GEANT4 [26].

4 Jet reconstruction and selection

Physics objects in CMS are reconstructed using the particle-flow (PF) algorithm [27], which identifies muons, electrons, photons, and neutral and charged hadrons through a combination of information from the various subdetectors. The reconstructed vertex with the largest value of summed physics-object p_T^2 is taken to be the primary pp interaction vertex. The physics objects are the jets, clustered using the anti-k_T jet finding algorithm [28, 29], with the tracks assigned to the vertex as inputs, and the associated missing transverse momentum, taken as the negative vector sum of the p_T of those jets. The particle candidates identified as originating from pileup are removed prior to the jet clustering [30, 31]. We use jets with a distance parameter R of 0.4 (AK4 jets) and 0.8 (AK8 jets), for the resolved and the boosted searches, respectively. Corrections are applied to jet energies as a function of pseudorapidity (η) and p_T of the jet to account for the combined response function of the detector to physics objects [31, 32].

For the boosted search, jet grooming techniques are used to eliminate soft, large-angle QCD multijet radiation that improves the jet mass resolution and reduces the pileup contributions to the jet mass. Two grooming algorithms are used: trimming [33] at the trigger stage and pruning [34] at the analysis stage. The trimming technique discriminates particles within the constituents of the jet based on a dynamic p_T threshold. In pruning, the constituents of the original jet are reclustered with the same cone size but using a modified Cambridge-Aachen (CA) algorithm [28] with relative p_T and angular requirements. To further discriminate jets originating from SM background processes and those from boosted hadronic resonances, we use N-subjetiness variables (τ_N) [35], which quantify the number of N prongs of energy inside a jet. In particular, ratios of N-subjetiness variables, $\tau_{MN} = \tau_M/\tau_N$, are found to provide better discrimination between signal and background. In this analysis, we use $\tau_{21} = \tau_2/\tau_1$ to distinguish two-pronged signal-like jets from and one-prong background-like jets which are overwhelmingly from QCD multijets events, and $\tau_{32} = \tau_3/\tau_2$ to separate two-prong jets from three-pronged jets from hadronically decaying top quarks.

Jets produced by the hadronization of bottom quarks are identified with the loose selection of the combined secondary vertex v2 b tagging algorithm [36]. This algorithm uses a multivariate discriminator with inputs information related to the secondary vertex, and the track-based lifetime to differentiate between jets from bottom quarks and from light-flavor quarks and gluons. The loose working point of the b tagging algorithm used in this analysis has an $\sim$81% b tagging efficiency, a $\sim$10% misidentification rate for light-quarks and gluons, and a $\sim$40% misidentification rate for charm quark jets [36].

5 Boosted search

Events are first selected with a trigger based on the total hadronic transverse momentum in the event (H_T), defined as the scalar sum of the p_T of AK4 jets (H_T^{AK4}) with a $p_T > 30\,\text{GeV}$ and $|\eta| < 2.5$. The H_T^{AK4} trigger threshold for the early data-taking period was set to 800 GeV, and increased to 900 GeV for the last $8\,\text{fb}^{-1}$ of data. Additionally, we include a logical *OR* of triggers based on AK8 jets with a selection on the mass of the jet after the trimming algorithm applied: one of these triggers requires an AK8 jet with $p_T > 360\,\text{GeV}$ and trimmed mass above 30 GeV, while the other requires $H_T^{AK8} > 750\,\text{GeV}$ with jet $p_T > 150\,\text{GeV}$ and a jet with trimmed mass above 50 GeV. The selection efficiency of the chosen triggers is determined relative to unbiased data samples with different triggers, and are found to have an efficiency >98% with respect to the analysis level selection for events satisfying $H_T^{AK8} > 900\,\text{GeV}$. After the trigger requirements, we select events with at least two AK8 jets with $p_T > 150\,\text{GeV}$ and in the central region of the detector with $|\eta| < 2.5$, and $H_T^{AK8} > 900\,\text{GeV}$.

The boosted search assumes that the decay products of the resonance would be fully contained in a very energetic AK8 jet, and therefore we select the two most energetic AK8 jets in the event. The pruning algorithm is used to compute the mass of each of these two jets (m_{j1} and m_{j2}). We then examine the spectrum of the average pruned jet mass of these two jets, $\overline{m} = (m_{j1} + m_{j2})/2$, for the presence of new physics in the mass range 60–450 GeV.

The following optimized selection criteria are applied to reduce SM background events. The mass asymmetry variable, defined as $m_{asym} = |m_{j1} - m_{j2}|/(m_{j1} + m_{j2}) < 0.1$, reduces the number of events with large mass imbalance between the two signal jet candidates. We require both jets to satisfy $\tau_{21} < 0.45$ and $\tau_{32} > 0.57$, to reject backgrounds from QCD multijets events and those from hadronically decaying top quarks, respectively. Signal events would be predominantly produced in the central η region, compared to the more wide spread QCD multijet production, and thus we retain events based on the absolute value of the difference in η between the two jets: $\Delta\eta = |\eta_{j1} - \eta_{j2}| < 1.5$. All the selection criteria are summarized in Table 1. For the b-tagged selection, both jets are required to pass the loose b tagging criteria described in Section 4. The discriminating power of each one of these kinematic variables is illustrated in Fig. 2 where normalized distributions between data, different simulated background components, and selected simulated signal samples are presented.

Figure 3 (left) shows the mass distributions for simulated signals after the inclusive selection; the signal mass shapes for the b-tagged analysis are similar. Additionally, the signal efficiency for the boosted search is reported in Fig. 3 (right) for both the inclusive and b-tagged selections. The fraction of $\widetilde{t} \to qq'$ signal events remaining after applying the inclusive selection is 0.003% for $m_{\widetilde{t}} = 80\,\text{GeV}$, increases to 0.106% for $m_{\widetilde{t}} = 180\,\text{GeV}$, and drops again to 0.055% for $m_{\widetilde{t}} = 400\,\text{GeV}$ due to the decrease in the production of top squarks with large Lorentz boosts at higher masses. Although the fraction of boosted resonances is higher for $m_{\widetilde{t}} \lesssim 170\,\text{GeV}$, the H_T and p_T trigger requirements considerably reduce the selected events, and are the main source

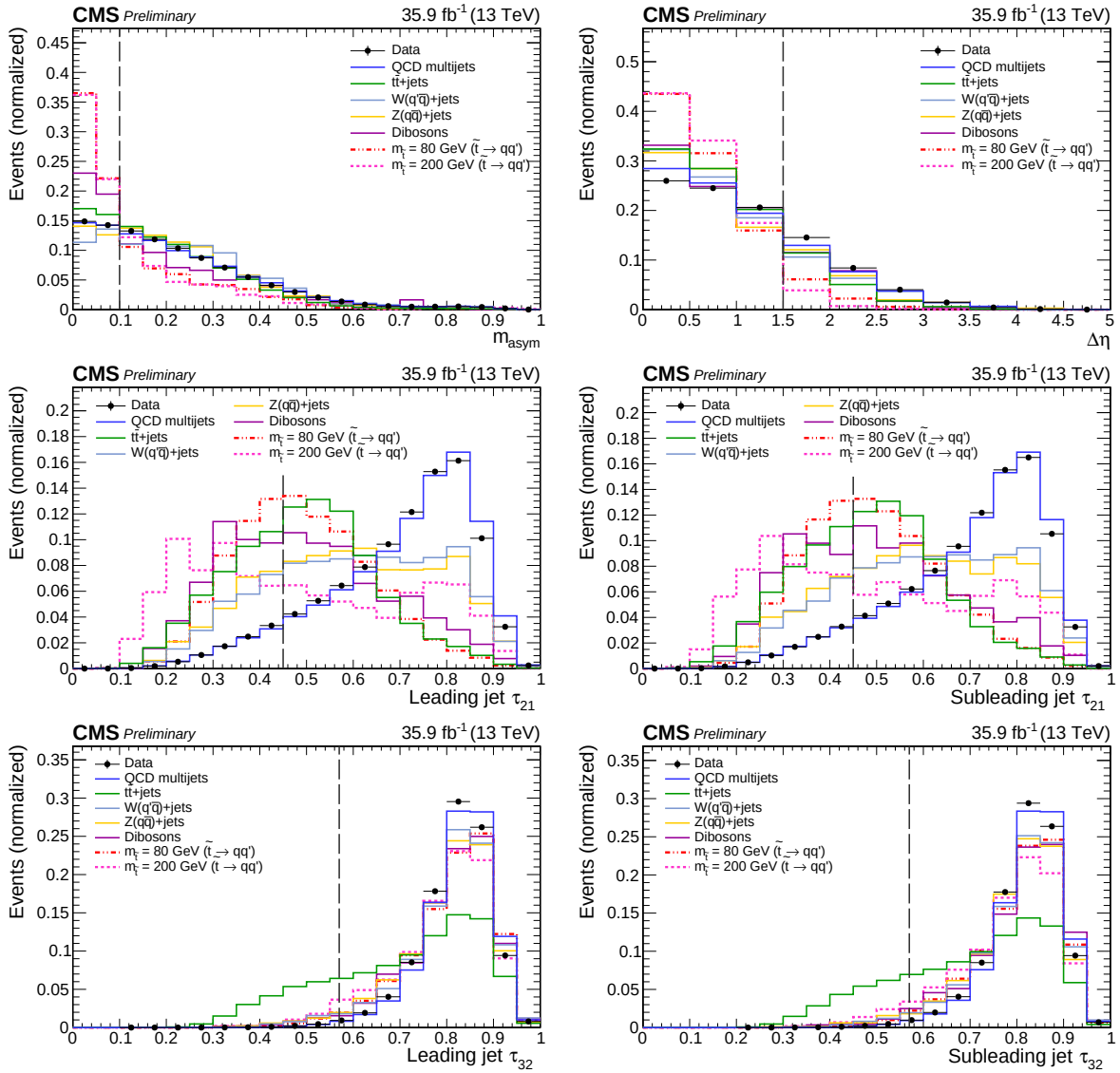

Figure 2: Boosted search: Kinematic distributions normalized to unity showing the comparison between data (black dots), backgrounds (solid colored lines), and a few selected $\tilde{t} \to qq'$ signal simulated samples (dashed colored lines). All selection criteria are applied apart from that on the variable being presented. In the case of the τ_{21} and τ_{32} variables both τ_{21} and τ_{32} requirements are removed. The black dashed lines represent where the selection is applied. Top left: m_{asym}. Top right: $\Delta\eta$. Middle left: leading jet τ_{21}. Middle right: subleading jet τ_{21}. Bottom left: leading jet τ_{32}. Bottom right: subleading jet τ_{32}.

Table 1: Signal selection criteria.

	Boosted search $60 < \overline{m} < 450\,\text{GeV}$ $(80 \leq m_{\tilde{t}} < 400\,\text{GeV})$	Resolved search $\overline{M} > 350\,\text{GeV}$ $(m_{\tilde{t}} \geq 400\,\text{GeV})$
Inclusive selection	AK8 jets jet $p_T > 150\,\text{GeV}$ jet $\lvert\eta\rvert < 2.5$ $N_{jet} \geq 2$ $H_T^{AK8} > 900\,\text{GeV}$ $m_{asym} < 0.1$ $\tau_{21} < 0.45$ $\tau_{32} > 0.57$ $\Delta\eta < 1.5$	AK4 jets jet $p_T > 80\,\text{GeV}$ jet $\lvert\eta\rvert < 2.5$ $N_{jet} \geq 4$ $H_T^{AK4} > 900\,\text{GeV}$ $M_{asym} < 0.1$ $\Delta\eta_{dijet} < 1.0$ $\Delta > 200\,\text{GeV}$
B-tagged selection	Inclusive selection plus two loose b-tagged jets	Inclusive selection plus two loose b-tagged jets

of the signal efficiency loss. The low signal selection efficiencies for boosted resonances are compensated by the large signal cross sections for low-mass top squarks [37, 38]. The b-tagged selection presents a similar pattern, where the fraction of remaining events for $\tilde{t} \to bq'$ is between 0.0009% for $m_{\tilde{t}} = 80\,\text{GeV}$, 0.0350% for $m_{\tilde{t}} = 200\,\text{GeV}$, and 0.0134% for higher resonance masses near $m_{\tilde{t}} = 400\,\text{GeV}$.

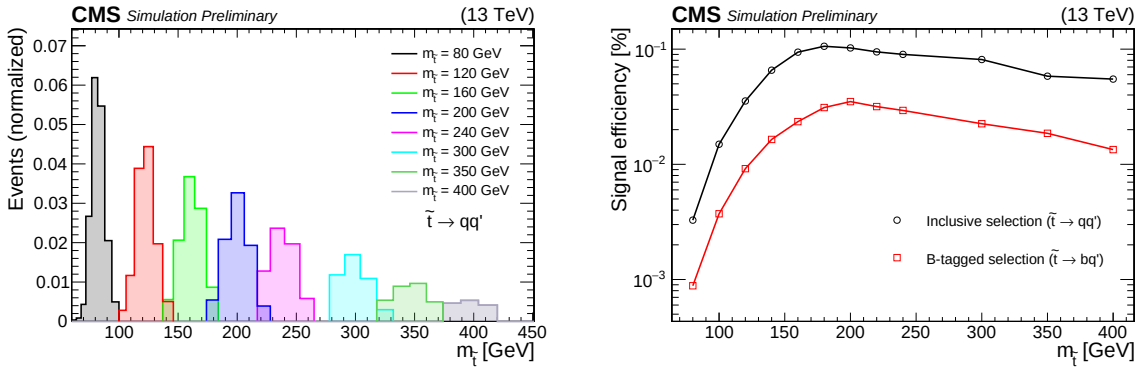

Figure 3: Boosted search: Left: Signal mass distributions for various simulated $\tilde{t} \to qq'$ masses probed in this analysis after applying the inclusive selection. Right: Signal efficiency as a function of $m_{\tilde{t}}$ for the inclusive and b-tagged selections.

After all the selection criteria are applied, the dominant remaining SM background is QCD multijet production. Subdominant resonant backgrounds are estimated from simulation and they include: $t\bar{t}$ +jets, $W \to q'\bar{q}$ +jets, $Z \to q\bar{q}$ +jets, and diboson (WW, ZZ, WZ) production. A normalization correction factor is introduced for the largest resonant background, $t\bar{t}$ +jets. To determine this correction, we first select a $t\bar{t}$ enriched sample by applying the inclusive selection except requiring $\tau_{32} < 0.57$, which increases the number of events with three-prong jets, and thus $t\bar{t}$. We then compare the $\overline{m}$ spectrum between data and simulation. We evaluate a correction factor obtained from a first-order polynomial fit of the $t\bar{t}$ component in the data after subtracting all other backgrounds. We find this correction to be flat as a function of $\overline{m}$ and consistent with unity within 10%. The uncertainty in the fit value is taken as a source of systematic uncertainty on the modeling of the simulated samples.

Events originating from QCD multijets are estimated from data control samples using two

uncorrelated variables in a so-called *ABCD* method described here. We use the m_{asym} and $\Delta\eta$ variables since their correlation is less than 0.1% in both real and simulated data. Then, we use these two variables to define four regions summarized in Table 2. Region A is a signal dominated region defined by the nominal inclusive selection criteria, while the other three regions are background dominated. Regions B and C are sideband regions where the selection of only one of the two variables is applied, and region D is defined as the sideband region when both selection criteria fail.

Table 2: Definition of the regions used in the QCD multijet background estimation for the boosted analysis. Region A is the signal dominated region while regions B, C, D are background dominated sideband regions.

	$m_{\text{asym}} < 0.1$	$m_{\text{asym}} > 0.1$
$\Delta\eta > 1.5$	Region B	Region D
$\Delta\eta < 1.5$	Region A	Region C

The yield and the shape of the $\overline{m}$ spectrum for the QCD multijet background in the signal region (A) is extrapolated using the mass spectra in sideband regions such that: $A = BC/D$. We define the transfer factor as the ratio B/D and parameterize it empirically as a function of $\overline{m}$ using a sigmoid function of the form

$$\frac{1}{p_0 + \exp(p_1 + p_2\overline{m}^2 - p_3\overline{m}^3)}, \tag{2}$$

where the coefficients p_0 to p_3 are free parameters of the function. Resonant background contributions estimated from simulations are subtracted from the data prior to the extrapolation. We find that the fit of the transfer factor gives consistent results in data and simulations. The resulting fit in the data illustrated in Fig. 4 is applied to events in region C to estimate the final $\overline{m}$ distribution for QCD multijet events in region A for the inclusive selection. The uncertainty of the fitted transfer factor and the statistical uncertainty in the $\overline{m}$ distribution in region C are assigned as systematic uncertainties.

For the b-tagged selection, an equivalent procedure is performed. However, due to insufficient statistics, we instead use the fitted transfer factor from the inclusive selection, and apply it to region C selected with b tagging requirements. By comparing the fit parameters of the transfer factors obtained with the inclusive and the b-tagged selections, an additional uncertainty is applied to cover their differences corresponding to the red band of Fig. 4.

The robustness of the *ABCD* background estimate is tested on simulated events using the resulting transfer factor fit to the QCD multijet simulation and applied to the mass spectrum of region C of the same sample. The prediction is then compared to the mass spectrum of the expected background in signal region A using the simulated sample. The level of agreement, or closure, between these two distributions is found to be within $\pm$10% over the entire $\overline{m}$ spectrum. This is used as a source of systematic uncertainty of the QCD multijet background prediction for both the inclusive and b-tagged selections.

Figure 5 shows the average pruned jet mass spectrum for data and the background predictions for the inclusive (left) and the b-tagged (right) selections. The resonant backgrounds correspond to less than 8% of the total background prediction for the inclusive category, and less than 6% for the b-tagged one, over the entire mass range.

Systematic uncertainties affecting the signal yield come from the integrated luminosity measurement (2.7%) [14], the trigger efficiency (3.0%), the modeling of the pileup interactions (1.0%), the effect from uncertainties on the parton distribution functions (PDF) (1.0%) [39], and

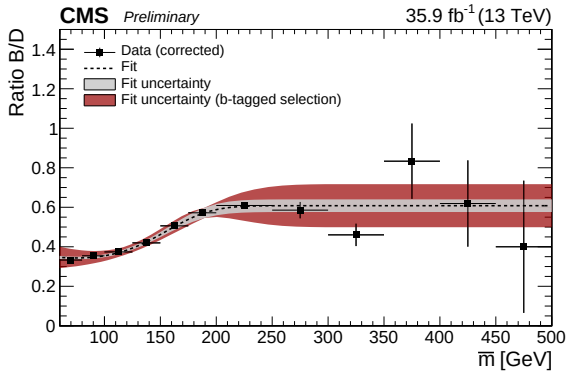

Figure 4: Boosted search: Transfer factor B/D as a function of $\overline{m}$ for data (black points) corrected for the resonant background component. Fit to the data (black dotted line) with the sigmoid function described in Eq. (2) is also presented. Gray and red bands represent the uncertainties of the fit for the inclusive and b-tagged selection, respectively, and are used as systematic uncertainties.

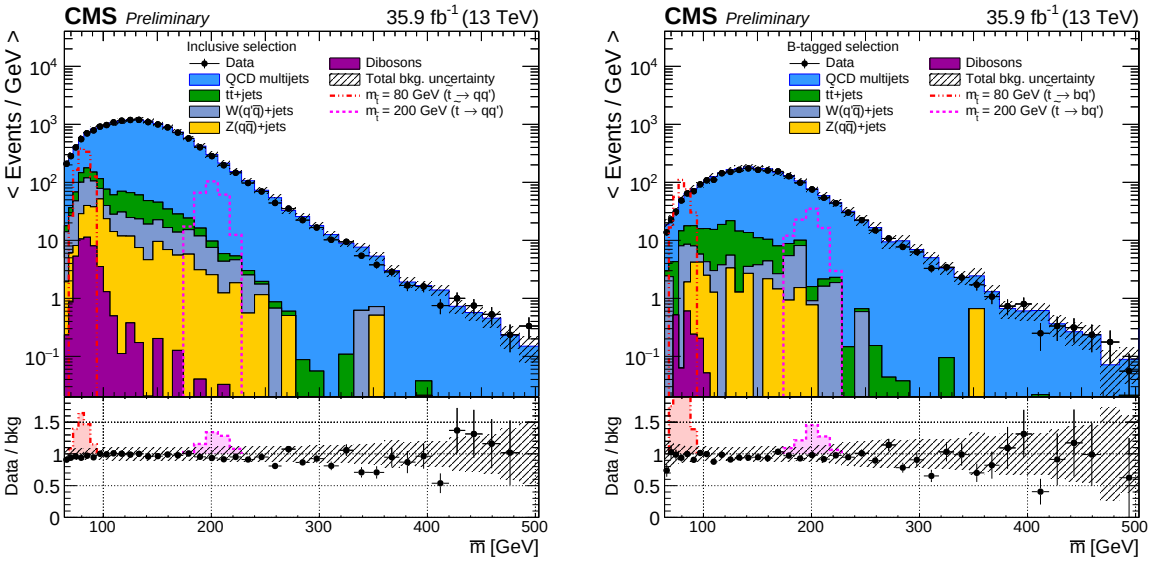

Figure 5: Boosted search: $\overline{m}$ distribution shown for data (black points) and the total background prediction. Left: inclusive selection. Right: b-tagged selection. The different background components are illustrated with different colors while the grey hashed band displays the total background uncertainty. On the bottom we show the ratio between data and the background prediction. The shaded colored regions on the bottom illustrate the effect when including a top squark signal for two different top squark masses.

the measurement of the jet energy scale (1.2%) and jet resolution (1.8%) [31, 32]. The values for these uncertainties are summarized in Table 3. For the b-tagged selection, the efficiency in the identification of bottom quarks is considered as a source of systematic uncertainty affecting the expected signal yield (1.0%) [36].

Shape uncertainties are also taken into account, coming from jet energy scale and resolution uncertainties of 2% and 14% respectively, as evaluated in Ref. [40]. These values are a result of differences between data and simulation in the reconstructed jet mass, when selecting hadronically decaying boosted W bosons.

After the τ_{21} requirement, previous studies [40, 41] have shown disagreement in the pruned

jet mass spectra between data and simulation. The method used to quantify this discrepancy is described in Ref. [41], and is based on measuring the efficiency of identifying boosted two-prong W bosons in semileptonic $\mathrm{t\bar{t}}$ events. For $\tau_{21} < 0.45$, the ratio of the efficiency in data and simulation is measured to be 1.10 ± 0.13. Since this search requires two jets to satisfy the same τ_{21} selection, we apply this scale factor squared to the signal efficiency in simulations, giving a total two-prong scale factor of 1.21 ± 0.29. A similar effect has been reported for the τ_{32} requirement [42]. In this case, we use a tag-and-probe procedure to measure the efficiency of identifying boosted three-pronged hadronic top quarks in semileptonic $\mathrm{t\bar{t}}$ events. For $\tau_{32} < 0.54$, the ratio of the efficiency between data and simulation is 1.07 ± 0.05, and the efficiency of misidentified boosted top quarks is 20%. From this ratio, we evaluate the efficiency of the $\tau_{32} > 0.54$ selection which is found to be 0.99 ± 0.01 and translates to an anti-three-prong scale factor of 0.96 ± 0.02 since we require two jets with this selection. The uncertainties on the two-prong (τ_{21}) and the anti-three-prong (τ_{32}) scale factors are taken as systematic uncertainties on the signal efficiency.

The systematic uncertainties on the background estimates are summarized in Table 4. For the QCD multijets background estimate, we consider three possible sources of uncertainty. First, we assign a systematic uncertainty based on the level of disagreement in the closure procedure described above. This is found to be 10% and it is applied to the overall yield of the background prediction. In addition, we apply an uncertainty on the transfer factor shape from the fit parameter uncertainties presented in Fig. 4. Lastly, the statistical uncertainties on the number of events in region *C* are applied bin-by-bin as a systematic uncertainty which affects the shape of the $\overline{m}$ distribution.

We also assign systematic uncertainties to the resonant background estimates. From the data-to-simulation scale factor study we performed using the $\mathrm{t\bar{t}}$ enriched selection, we assign a 10% systematic uncertainty on the modeling of the SM simulated samples. Finally, the statistical uncertainties of the simulated samples are also considered as systematic uncertainties.

6 Resolved search

Events are selected using a logical *OR* of the $H_{\mathrm{T}}^{\mathrm{AK4}}$ trigger, described in Section 5, and two additional triggers: one requiring at least four AK4 jets with $p_{\mathrm{T}} > 50\,\mathrm{GeV}$, $|\eta| < 2.5$, and $H_{\mathrm{T}}^{\mathrm{AK4}} > 800\,\mathrm{GeV}$, and another requiring at least four jets with $p_{\mathrm{T}} > 70\,\mathrm{GeV}$, $|\eta| < 2.5$, and $H_{\mathrm{T}}^{\mathrm{AK4}} > 750\,\mathrm{GeV}$. After the trigger requirements, we select events with at least four AK4 jets with $p_{\mathrm{T}} > 80\,\mathrm{GeV}$, $|\eta| < 2.5$, and $H_{\mathrm{T}}^{\mathrm{AK4}} > 900\,\mathrm{GeV}$. The selection efficiency of the chosen triggers is determined using unbiased data samples with different triggers. We measure a trigger efficiency greater than 98% with respect to this analysis level selection.

In this analysis, we construct two dijet systems compatible with signal resonances produced in pairs and each decaying into two jets. We create three unique combinations of dijet pairs per event using the leading four jets ordered in p_{T}. Out of the three combinations, the dijet configuration with the smallest $\Delta R_{\mathrm{dijet}}$ is chosen, and is defined as: $\Delta R_{\mathrm{dijet}} = \sum_{\mathrm{i}=1,2} |\Delta R^{\mathrm{i}} - 0.8|$, where ΔR^{i} represents the separation between two jets in the i^{th} dijet pair, $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$, and $\Delta\eta$ and $\Delta\phi$ are the differences in η and ϕ between the two jets under consideration. This variable exploits the expectation that the decay products of the signal resonance will be closer together compared to uncorrelated jets. An offset of 0.8 has been chosen in the definition of $\Delta R_{\mathrm{dijet}}$ to avoid overlaps between jets in the dijet systems, and to minimize the selection of dijet systems composed of jets from radiated gluons.

Once a configuration is selected the average mass of the dijet system, $\overline{M} = (m_{\mathrm{jj1}} + m_{\mathrm{jj2}})/2$, is

used to search for new resonances, where m_{jji} is the dijet mass of the i^{th} dijet pair. To further reject backgrounds from QCD multijet events and incorrect signal pairings, two additional requirements are applied to the event selection. Since the dijet systems in signal events are expected to be more centrally located in η space than in QCD multijet events, the pseudorapidity difference between the two dijet systems is required to be $\Delta\eta_{\text{dijet}} = |\eta_{jj1} - \eta_{jj2}| < 1.0$. In addition, we require a low-mass imbalance between the dijet systems by selecting events with a mass asymmetry between the dijet pairs, $M_{\text{asym}} = |m_{jj1} - m_{jj2}|/(m_{jj1} + m_{jj2}) < 0.1$. Figures 6 and 7 shows the discriminating power of these two kinematic variables between data, QCD multijets simulation, and a selected simulated signal sample.

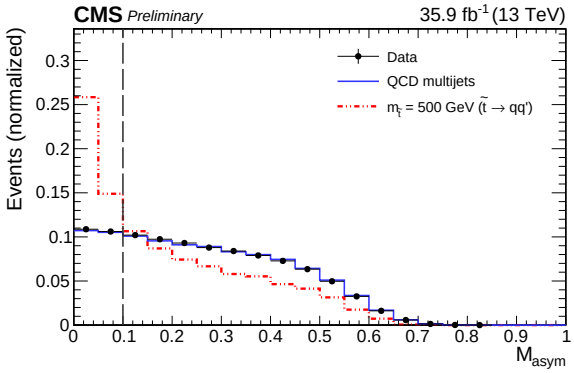

Figure 6: Resolved search: The M_{asym} distribution normalized to unity showing the comparison between data (black dots), background (solid blue line), and a selected signal $\tilde{t} \to qq'$ with a $m_{\tilde{t}} = 500\,\text{GeV}$ (dashed red line). All selection criteria are applied apart from that on the variable being presented, and the region to the left of the black dashed line indicates the optimized region of selected M_{asym} values.

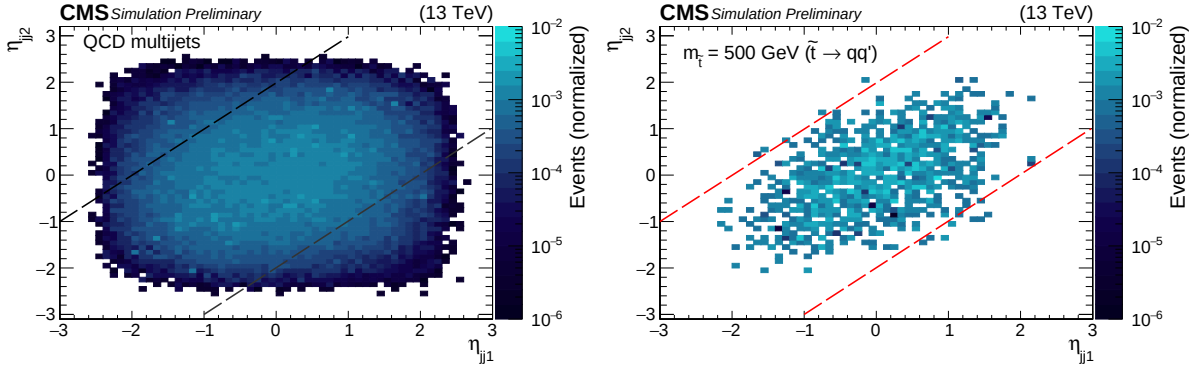

Figure 7: Resolved search: The η_{jj1} value of the higher p_{T} dijet system in the selected pair as a function of the η_{jj2} value of the lower p_{T} dijet system. The distribution is shown for simulated QCD multijet events (left) and a selected signal $\tilde{t} \to qq'$ with a $m_{\tilde{t}} = 500\,\text{GeV}$ (right). All selection criteria are applied apart from that on the variable being presented, and the region between the two red dashed lines indicates the optimized region of selected $\Delta\eta_{\text{dijet}}$ values.

An additional variable defined as $\Delta = \left(\sum_{i=1,2}|p_{\text{T}}^{i}|\right) - \overline{M}$ is calculated for each dijet system, where the p_{T} sum is over the two jets in the dijet configuration. The distributions of the Δ variable as a function of $\overline{M}$ for a selected signal sample and QCD multijets simulation are illustrated in Fig. 8. This variable has been previously used in hadronic resonance searches at both the Tevatron and the LHC [8, 9, 43–46]. In addition to rejecting background events, requiring a minimum value of Δ results in a lowering of the peak position of the $\overline{M}$ distribution

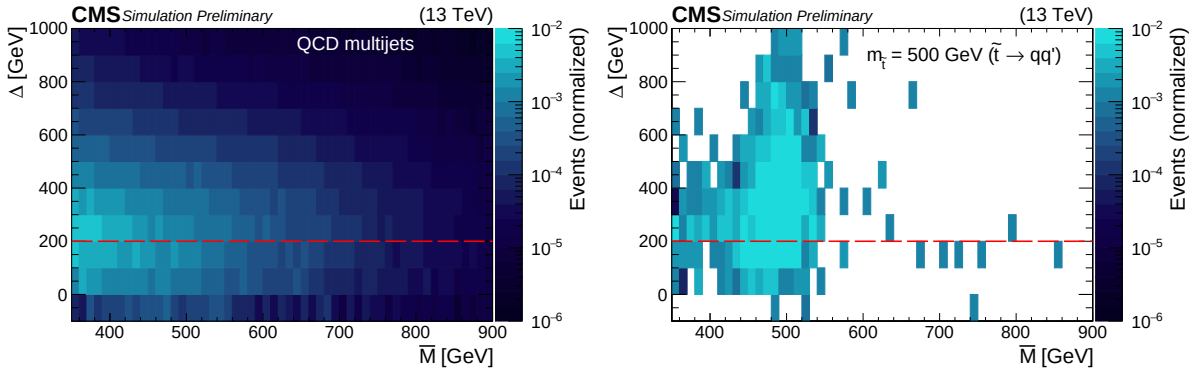

Figure 8: Resolved search: The distribution of Δ as a function of $\overline{M}$, shown for simulated QCD multijet events (left) and a selected signal $\tilde{t} \to qq'$ with a $m_{\tilde{t}} = 500\,\text{GeV}$ (right). All selection criteria are applied apart from that on the variable being presented, and the region above the red dashed line indicates the optimized region of selected Δ values.

in SM QCD multijet events, and allows the search to extend to lower mass resonances. We select events with $\Delta > 200\,\text{GeV}$.

Finally, for the b-tagged selection, we require a loose b-tagged jet in each dijet pair candidate.

A summary of the selection for this search is summarized in Table 1.

Events originating from QCD multijet processes dominate the $\overline{M}$ spectrum and are modeled with the following function

$$\frac{\mathrm{d}N}{\mathrm{d}\overline{M}} = \frac{p_0(1-x)^{p_1}}{(x)^{p_2}}, \tag{3}$$

where $x = \overline{M}/\sqrt{s}$, $\sqrt{s}$ is the center of mass energy, N is the number of events, and p_0 through p_2 are parameters of the function. The functional form in Eq. (3) successfully models the steeply falling dijet mass distribution of QCD multijet production and has been extensively used in similar previous dijet resonance searches [9, 40, 47].

Figure 9 (left) shows the fitted $\overline{M}$ distributions in data using the inclusive (top) and the b-tagged (bottom) selections for the resolved analysis. The parameterized fit is performed for $\overline{M} > 350\,\text{GeV}$ for both selections, which was determined to be the region well modeled by the background parameterization, as well as where the triggers are greater than 98% efficient as a function of $\overline{M}$. Figure 9 (right) presents the bin-by-bin difference between the data and the fit divided by the statistical uncertainty (pull distribution) and the same difference divided by the value of the fit (residual distribution). The data are observed to be in agreement with SM expectations.

The potential bias introduced by the choice of the background parameterization is investigated. Signal injection tests are performed in pseudo-experiments generated from QCD multijet simulations fit with the function of Eq. (3). Each pseudo-experiment is then fit with alternative parameterizations and the effect in the strength of the injected signal is estimated and found to be negligible.

The $\overline{M}$ distributions of the simulated signal samples are parameterized with Gaussian functions, and are presented in Fig. 10 (left) for the inclusive selection; the signal mass shapes for the b-tagged analysis are found to be similar. The signal efficiency for the resolved search is illustrated in Fig. 10 (right) for both the inclusive and b-tagged selections. The fraction of $\tilde{t} \to qq'$ signal events remaining in simulation after applying the inclusive selection is between

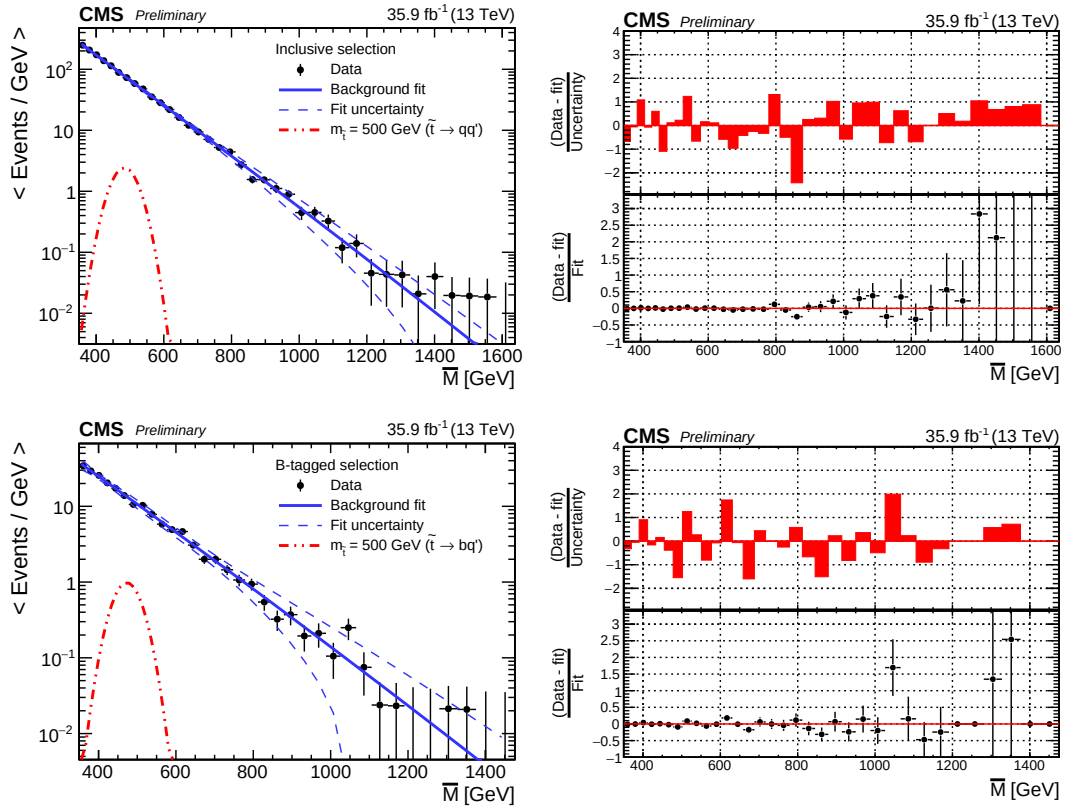

Figure 9: Resolved search: Left: The $\overline{M}$ distributions for the data (black points), along with the resulting fit of the functional form in Eq. (3) (blue line) for the inclusive selection (top) and b-tagged (bottom) selections. The expected signals from the $\widetilde{t} \to qq'$ and $\widetilde{t} \to bq'$ simulated samples at $m_{\widetilde{t}} = 500\,\text{GeV}$ are also presented (red lines) for the inclusive selection and b-tagged selections, respectively. Right: The bin-by-bin pull and residual distributions, as described in the text.

0.66% to 1.16% for $m_{\widetilde{t}}$ between 400 GeV to 1500 GeV. In the b-tagged selection, the fraction of remaining events in the $\widetilde{t} \to bq'$ simulation is between 0.12% to 0.42% for $m_{\widetilde{t}}$ between 400 GeV to 1400 GeV.

The sources of systematic uncertainties affecting the normalization of the expected signal contribution are the integrated luminosity measurement (2.7%) [14], the trigger efficiency (3.0%), the modeling of the pileup interactions (1.0%), and the effect from PDFs (1.0%) [39]. The uncertainties in the measurement of the jet energy scale (1.5%) and jet resolution (6.0%) [31, 32] introduce both a change in the yield and the shape of the $\overline{M}$ spectrum. For the b-tagged selection, the efficiency in the identification of bottom quarks is considered as a source of systematic uncertainty affecting the expected signal yield (1.0%) [36]. Finally, the statistical uncertainties of the simulated samples are also considered as a systematic uncertainty. The systematic uncertainties affecting the signal are summarized in Table 3.

The uncertainties in the fitted parameters of Eq. (3) are also taken into account as systematic uncertainties affecting both the background shape and normalization, and are summarized in Table 4.

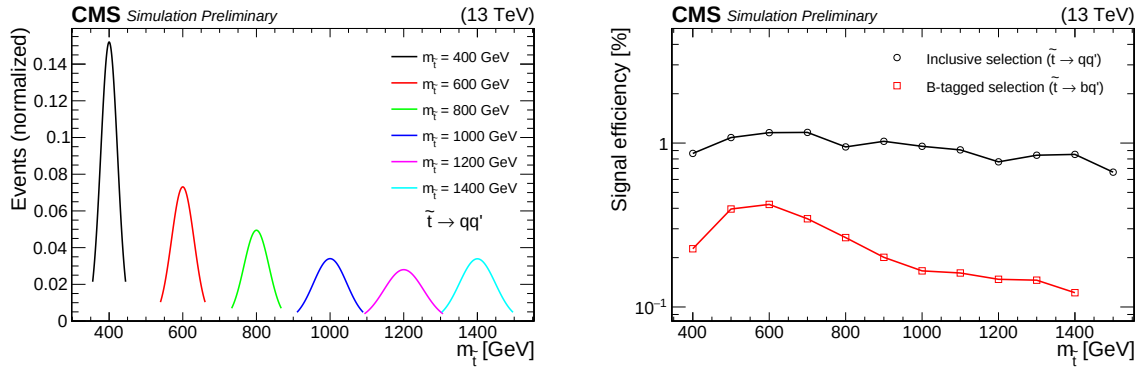

Figure 10: Resolved search: Left: Gaussian fits on the mass of the simulated signals for various $\tilde{t} \to qq'$ masses probed in this search after applying the inclusive selection. Right: Signal efficiency as a function of $m_{\tilde{t}}$ for the inclusive and b-tagged selections.

Table 3: Summary of the systematic uncertainties on the signal acceptance. For the shape uncertainties, the value represents the percentage difference in the nominal value of the systematic uncertainty.

Search	Source of Systematic	Effect	Value
Boosted and resolved	Luminosity	Yield	2.7%
	Trigger	Yield	3.0%
	Pileup	Yield	1.0%
	PDF	Yield	1.0%
	Jet energy scale	Yield	1.2–1.5%
		Shape	2.0%
	Jet energy resolution	Yield	1.8–6.0%
		Shape	10.0–14.0%
	Simulation statistical precision	-	bin-by-bin
	b tagging (only for b-tagged selection)	Yield	1.0%
Boosted	Two-prong scale factor	Yield	23.0%
	Anti-three-prong scale factor	Yield	2.0%

7 Results

Figures 5 and 9 show the mass spectra for the boosted and resolved analyses, respectively, and we find that they are in agreement with SM expectations. The mass spectra are used to set limits on the pair production of top squarks decaying via the RPV couplings λ''_{312} and λ''_{323}. The exclusion limits are computed using the modified frequentist approach for confidence levels, with a binned profile likelihood as the test statistic [48, 49] in the asymptotic approximation [50].

Results for the boosted search are obtained from combined signal and background binned likelihood fits to the $\overline{m}$ distribution in data. For each value of $m_{\tilde{t}}$ considered, only bins of $\overline{m}$ within two standard deviations of the mean of the generated top squark mass fit to a Gaussian are included in the likelihood. For each bin used in the likelihood the individual background components and signal are allowed to float within statistical uncertainties. Systematic uncertainties affecting the yield and the shape as summarized in Tables 3 and 4 are assumed to be correlated between bins. These uncertainties are treated as nuisance parameters, which are profiled, and modeled with log normal priors, except for the uncertainty on the number of events in sideband region C which is modeled with a gamma prior.

Table 4: Summary of the systematic uncertainties on the background prediction by source.

Search	Background	Source of Systematic	Effect	Value
Boosted	QCD multijets	Closure	Yield	10.0%
		Transfer factor fit	Shape	1.0–8.0%
		Statistics in region (C)	Shape	bin-by-bin
	Resonant	Simulation modeling	Yield	10.0%
		Simulation statistical precision	Shape	bin-by-bin
Resolved	QCD multijets	Fit parameters	Shape and Yield	bin-by-bin

For the resolved search, the $\overline{M}$ spectrum in data is compared to the background fit to search for localized deviations consistent with a resonance. For each $m_{\widetilde{t}}$, a likelihood fit is used to compare the data to the shapes for the signal and background within a mass window of two standard deviations around the true value of $m_{\widetilde{t}}$. Here, all systematic uncertainties are modeled with log normal priors.

Figure 11 illustrates the observed and expected 95% CL upper limits on the top squark pair production cross section times branching ratio as a function of $m_{\widetilde{t}}$ for the boosted and resolved analyses. The boosted analysis probes $80 \leq m_{\widetilde{t}} < 400\,\text{GeV}$, while the resolved analysis probes $m_{\widetilde{t}} \geq 400\,\text{GeV}$. Figure 11 (left) shows the resulting limits using the inclusive selection for the λ''_{312} coupling scenario, while Fig. 11 (right) displays the limits using the b-tagged selection assuming the λ''_{323} coupling. The dashed pink line represents the theoretical predictions for top squark pair production at $\sqrt{s} = 13\,\text{TeV}$ evaluated at next-to-leading order (NLO) with next-to-leading-logarithm (NLL) corrections [37, 38]. We note that for the range of top squark masses considered in this result, the b-tagged selection provides an added value to the $\widetilde{t} \to bq'$ results when compared to the inclusive selection. We exclude top squark masses from 80 to 520 GeV assuming the λ''_{312} coupling, and from 80 to 270 GeV, 285 to 340 GeV, and 400 to 525 GeV for the λ''_{312} coupling. The corresponding expected mass limits obtained are 80 to 520 GeV for top squarks decaying via λ''_{323}, and 80 to 270 GeV, 285 to 320 GeV, and 400 to 505 GeV for the λ''_{323} coupling.

8 Summary

A search has been performed for the pair production of diquark resonances in two jet events in a boosted jet topology and in four-jet events in a resolved jet topology. Data is analyzed from proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$ collected in 2016 with the CMS detector corresponding to an integrated luminosity of $35.9\,\text{fb}^{-1}$. In the boosted search the distribution of the average mass of the selected two jets has been investigated for localized disagreements between the data and the background estimate, consistent with a new resonance, while in the resolved analysis the average mass of the selected dijet pairs is examined for localized disagreements between data and the background expectations. The boosted search explores resonance masses between 60 and 450 GeV, while the resolved one covers masses above 350 GeV. We find agreement between observations and standard model expectations. These results are interpreted in the framework of RPV SUSY assuming the pair production of top squarks decaying to quarks via the λ''_{312} or the λ''_{323} couplings, assuming 100% branching ratios to $\widetilde{t} \to qq'$ or $\widetilde{t} \to bq'$, respectively. Upper limits are set at 95% confidence level on the pair production cross section of top squarks as a function of the top squark mass. We exclude top squark masses from 80 to 520 GeV assuming the λ''_{312} coupling, and from 80 to 270 GeV, 285 to 340 GeV, and 400 to 525 GeV assuming the λ''_{323} coupling.

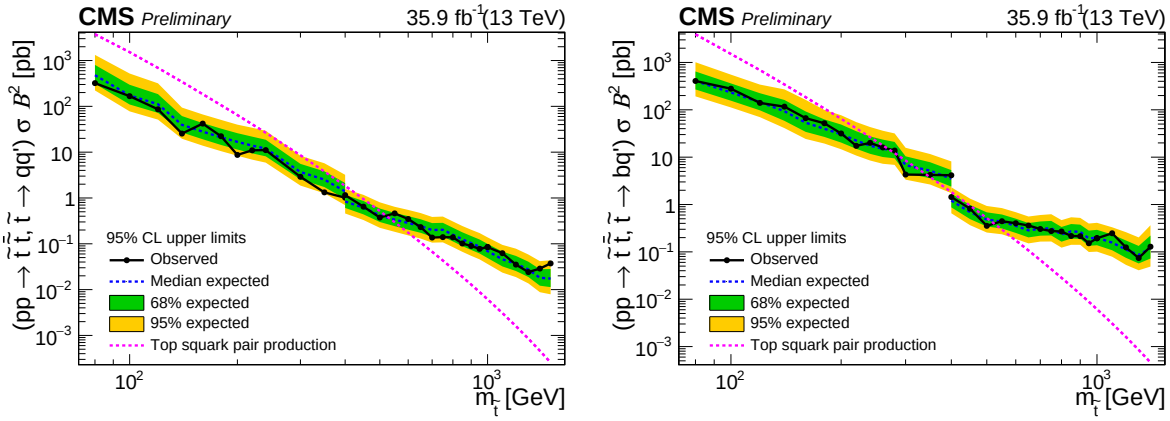

Figure 11: Observed and expected 95% CL upper limits on the product of the cross section times the branching ratio ($\mathcal{B}^2$) as a function of $m_{\tilde{t}}$. The branching ratio to quarks is assumed to be 100%. The boosted analysis probes $80 \leq m_{\tilde{t}} < 400\,\text{GeV}$, while the resolved analysis searches for $m_{\tilde{t}} \geq 400\,\text{GeV}$. Left: limits using the inclusive selection for $\tilde{t} \to qq'$ assuming the RPV coupling λ''_{312}. Right: limits using the b-tagged selection for $\tilde{t} \to bq'$ assuming the RPV coupling λ''_{323}. The dashed pink line shows the NLO + NLL theoretical predictions for top squark pair production.

Acknowledgments

We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centres and personnel of the Worldwide LHC Computing Grid for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC and the CMS detector provided by the following funding agencies: BMWFW and FWF (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MES (Bulgaria); CERN; CAS, MoST, and NSFC (China); COLCIENCIAS (Colombia); MSES and CSF (Croatia); RPF (Cyprus); SENESCYT (Ecuador); MoER, ERC IUT, and ERDF (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); GSRT (Greece); NKFIA (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); MSIP and NRF (Republic of Korea); LAS (Lithuania); MOE and UM (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI (Mexico); MBIE (New Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Dubna); MON, RosAtom, RAS and RFBR (Russia); MESTD (Serbia); SEIDI, CPAN, PCTI and FEDER (Spain); Swiss Funding Agencies (Switzerland); MST (Taipei); ThEPCenter, IPST, STAR, and NSTDA (Thailand); TUBITAK and TAEK (Turkey); NASU and SFFR (Ukraine); STFC (United Kingdom); DOE and NSF (USA).

References

[1] C. Kilic, T. Okui, and R. Sundrum, "Colored resonances at the Tevatron: phenomenology and discovery potential in multijets", *JHEP* **07** (2008) 038, doi:10.1088/1126-6708/2008/07/038, arXiv:0802.2568.

[2] C. T. Hill, "Topcolor: top quark condensation in a gauge extension of the standard model", *Phys. Lett. B* **266** (1991) 419, doi:10.1016/0370-2693(91)91061-Y.

[3] G. D. Kribs, E. Poppitz, and N. Weiner, "Flavor in supersymmetry with an extended R-symmetry", *Phys. Rev. D* **78** (2008) 055010, doi:10.1103/PhysRevD.78.055010, arXiv:0712.2039.

[4] J. A. Evans and Y. Kats, "LHC coverage of RPV MSSM with light stops", *JHEP* **04** (2013) 028, doi:10.1007/JHEP04(2013)028, arXiv:1209.0764.

[5] R. Barbier et al., "R-parity violating supersymmetry", *Phys. Rept.* **420** (2005) 1, doi:10.1016/j.physrep.2005.08.006, arXiv:hep-ph/0406039.

[6] ALEPH Collaboration, "Search for supersymmetric particles with R parity violating decays in e^+e^- collisions at $\sqrt{s}$ up to 209 GeV", *Eur. Phys. J. C* **31** (2003) 1, doi:10.1140/epjc/s2003-01311-5, arXiv:hep-ex/0210014.

[7] CDF Collaboration, "Search for pair production of strongly interacting particles decaying to pairs of jets in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV", *Phys. Rev. Lett.* **111** (2013) 031802, doi:10.1103/PhysRevLett.111.031802, arXiv:1303.2699.

[8] CMS Collaboration, "Search for pair-produced dijet resonances in four-jet final states in proton-proton collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. Lett.* **110** (2013) 141802, doi:10.1103/PhysRevLett.110.141802, arXiv:1302.0531.

[9] CMS Collaboration, "Search for pair-produced resonances decaying to jet pairs in proton-proton collisions at $\sqrt{s} = 8$ TeV", *Phys. Lett. B* **747** (2015) 98, doi:10.1016/j.physletb.2015.04.045, arXiv:1412.7706.

[10] ATLAS Collaboration, "Search for pair-produced massive coloured scalars in four-jet final states with the ATLAS detector in proton-proton collisions at $\sqrt{s} = 7$ TeV", *Eur. Phys. J. C* **73** (2013) 2263, doi:10.1140/epjc/s10052-012-2263-z, arXiv:1210.4826.

[11] ATLAS Collaboration, "A search for top squarks with R-parity-violating decays to all-hadronic final states with the ATLAS detector in $\sqrt{s} = 8$ TeV proton-proton collisions", *JHEP* **06** (2016) 067, doi:10.1007/JHEP06(2016)067, arXiv:1601.07453.

[12] ATLAS Collaboration, "A search for pair-produced resonances in four-jet final states at $\sqrt{s} = 13$ TeV with the ATLAS detector", *Eur. Phys. J. C* **78** (2018) 250, doi:10.1140/epjc/s10052-018-5693-4, arXiv:1710.07171.

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

[14] CMS Collaboration, "CMS luminosity measurements for the 2016 data taking period", CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017.

[15] CMS Collaboration, "The CMS trigger system", *JINST* **12** (2017) P01020, doi:10.1088/1748-0221/12/01/P01020, arXiv:1609.02366.

[16] J. Alwall et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **07** (2014) 079, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[17] T. Sjöstrand et al., "An Introduction to PYTHIA 8.2", *Comput. Phys. Commun.* **191** (2015) 159, doi:10.1016/j.cpc.2015.01.024, arXiv:1410.3012.

[18] CMS Collaboration, "Event generator tunes obtained from underlying event and multiparton scattering measurements", *Eur. Phys. J. C* **76** (2016) 155, doi:10.1140/epjc/s10052-016-3988-x, arXiv:1512.00815.

[19] S. Alioli, S.-O. Moch, and P. Uwer, "Hadronic top-quark pair-production with one jet and parton showering", *JHEP* **01** (2012) 137, doi:10.1007/JHEP01(2012)137, arXiv:1110.5251.

[20] T. Melia, P. Nason, R. Rontsch, and G. Zanderighi, "W^+W^+ plus dijet production in the POWHEGBOX", *Eur. Phys. J. C* **71** (2011) 1670, doi:10.1140/epjc/s10052-011-1670-x, arXiv:1102.4846.

[21] S. Kallweit et al., "NLO QCD+EW predictions for V+jets including off-shell vector-boson decays and multijet merging", *JHEP* **04** (2016) 021, doi:10.1007/JHEP04(2016)021, arXiv:1511.08692.

[22] S. Kallweit et al., "NLO QCD+EW automation and precise predictions for V+multijet production", in *Proceedings, 50th Rencontres de Moriond, QCD and high energy interactions: La Thuile, Italy, March 21-28, 2015*, p. 121. 2015. arXiv:1505.05704.

[23] J. M. Lindert et al., "Precise predictions for V+jets dark matter backgrounds", *Eur. Phys. J. C* **77** (2017) 829, doi:10.1140/epjc/s10052-017-5389-1, arXiv:1705.04664.

[24] CMS Collaboration, "Search for dark matter produced with an energetic jet or a hadronically decaying W or Z boson at $\sqrt{s} = 13\,\text{TeV}$", *JHEP* **07** (2017) 014, doi:10.1007/JHEP07(2017)014, arXiv:1703.01651.

[25] S. Kallweit et al., "NLO electroweak automation and precise predictions for W+multijet production at the LHC", *JHEP* **04** (2015) 012, doi:10.1007/JHEP04(2015)012, arXiv:1412.5157.

[26] GEANT4 Collaboration, "Geant4 — a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[27] CMS Collaboration, "Particle-flow reconstruction and global event description with the CMS detector", *JINST* **12** (2017) P10003, doi:10.1088/1748-0221/12/10/P10003, arXiv:1706.04965.

[28] 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.

[29] 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.

[30] M. Cacciari and G. P. Salam, "Pileup subtraction using jet areas", *Phys. Lett. B* **659** (2008) 119, doi:10.1016/j.physletb.2007.09.077, arXiv:0707.1378.

[31] CMS Collaboration, "Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV", *JINST* **12** (2017) P02014, doi:10.1088/1748-0221/12/02/P02014, arXiv:1607.03663.

[32] 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.

[33] D. Krohn, J. Thaler, and L.-T. Wang, "Jet trimming", *JHEP* **02** (2010) 084, doi:10.1007/JHEP02(2010)084, arXiv:0912.1342.

[34] S. D. Ellis, C. K. Vermilion, and J. R. Walsh, "Recombination algorithms and jet substructure: pruning as a tool for heavy particle searches", *Phys. Rev. D* **81** (2010) 094023, doi:10.1103/PhysRevD.81.094023, arXiv:0912.0033.

[35] 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.

[36] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", *JINST* **13** (2018) P05011, doi:10.1088/1748-0221/13/05/P05011, arXiv:1712.07158.

[37] C. Borschensky et al., "Squark and gluino production cross sections in pp collisions at $\sqrt{s} = 13$, 14, 33 and 100 TeV", *Eur. Phys. J. C* **74** (2014) 3174, doi:10.1140/epjc/s10052-014-3174-y, arXiv:1407.5066.

[38] W. Beenakker et al., "NLO+NLL squark and gluino production cross-sections with threshold-improved parton distributions", *Eur. Phys. J. C* **76** (2016) 53, doi:10.1140/epjc/s10052-016-3892-4, arXiv:1510.00375.

[39] J. Rojo, "PDF4LHC recommendations for Run II", in *Proceedings, 24th International Workshop on Deep-Inelastic Scattering and Related Subjects (DIS 2016): Hamburg, Germany, April 11-15, 2016*, volume DIS2016, p. 018. 2016. arXiv:1606.08243.

[40] CMS Collaboration, "Search for massive resonances decaying into WW, WZ, ZZ, qW, and qZ with dijet final states at $\sqrt{s} = 13$ TeV", *Phys. Rev. D* **97** (2018) 072006, doi:10.1103/PhysRevD.97.072006, arXiv:1708.05379.

[41] CMS Collaboration, "Identification techniques for highly boosted W bosons that decay into hadrons", *JHEP* **12** (2014) 017, doi:10.1007/JHEP12(2014)017, arXiv:1410.4227.

[42] CMS Collaboration, "Jet algorithms performance in 13 TeV data", CMS Physics Analysis Summary CMS-PAS-JME-16-003, 2017.

[43] CDF Collaboration, "First search for multijet resonances in $\sqrt{s} = 1.96$ TeV $\mathrm{p\overline{p}}$ collisions", *Phys. Rev. Lett.* **107** (2011) 042001, doi:10.1103/PhysRevLett.107.042001, arXiv:1105.2815.

[44] CMS Collaboration, "Search for three-jet resonances in pp collisions at $\sqrt{s} = 7$ TeV", *Phys. Rev. Lett.* **107** (2011) 101801, doi:10.1103/PhysRevLett.107.101801, arXiv:1107.3084.

[45] CMS Collaboration, "Search for three-jet resonances in pp collisions at $\sqrt{s} = 7$ TeV", *Phys. Lett. B* **718** (2012) 329, doi:10.1016/j.physletb.2012.10.048, arXiv:1208.2931.

[46] CMS Collaboration, "Searches for light- and heavy-flavour three-jet resonances in pp collisions at $\sqrt{s} = 8$ TeV", *Phys. Lett. B* **730** (2014) 193, doi:10.1016/j.physletb.2014.01.049, arXiv:1311.1799.

[47] CMS Collaboration, "Search for narrow resonances decaying to dijets in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$", *Phys. Rev. Lett.* **116** (2016) 071801, doi:10.1103/PhysRevLett.116.071801, arXiv:1512.01224.

[48] A. L. Read, "Presentation of search results: the CL(s) technique", *J. Phys. G* **28** (2002) 2693, doi:10.1088/0954-3899/28/10/313.

[49] T. Junk, "Confidence level computation for combining searches with small statistics", *Nucl. Instrum. Meth. A* **434** (1999) 435, doi:10.1016/S0168-9002(99)00498-2, arXiv:hep-ex/9902006.

[50] G. Cowan, K. Cranmer, E. Gross, and O. Vitells, "Asymptotic formulae for likelihood-based tests of new physics", *Eur. Phys. J. C* **71** (2011) 1554, doi:10.1140/epjc/s10052-011-1554-0, 10.1140/epjc/s10052-013-2501-z, arXiv:1007.1727. [Erratum: Eur. Phys. J. C **73** (2013) 2501].