Search for New Phenomena in Dijet Events using 139 fb$^{-1}$ of $pp$ collisions at $\sqrt{s} = 13$ TeV collected with the ATLAS Detector

The ATLAS Collaboration

A search for new resonances decaying into two hadronic jets is reported using the entire dataset of proton-proton collisions recorded at $\sqrt{s} = 13$ TeV with the ATLAS detector at the Large Hadron Collider between 2015 and 2018, corresponding to an integrated luminosity of 139 fb$^{-1}$. The dijet invariant mass distribution is compared to a smoothly-falling background prediction obtained by fitting the data. No significant excess is observed. Excited quarks with masses below 6.7 TeV are excluded at the 95% confidence level. Model-independent limits on Gaussian-shaped signals of various widths in dijet mass distribution are also set.

© 2019 CERN for the benefit of the ATLAS Collaboration.
Reproduction of this article or parts of it is allowed as specified in the CC-BY-4.0 license.
1 Introduction

New heavy particles that couple to partons are predicted in many beyond the Standard Model (BSM) theories and can be produced directly in proton-proton (pp) collisions and decay to partons. Partons shower and hadronize, creating collimated jets of particles with four-momentum that is approximately equivalent to that of the parton. Hence, BSM phenomena may produce a resonance signal in two-jet (dijet) final states, allowing searches for BSM signals up to masses that are a significant fraction of the total hadronic collision energy. Excited quarks ($q^*$) [1, 2] are predicted in models of compositeness and are a typical benchmark for quark-gluon resonances used in many past dijet searches [3–10].

In the Standard Model (SM), production of jet pairs in hadronic collisions primarily results from $2 \rightarrow 2$ parton scattering processes via strong interactions, described by quantum chromodynamics (QCD). Outgoing particles hadronize, resulting in jets with high transverse momentum ($p_T$) with respect to the colliding partons. QCD [11] predicts a smooth and monotonically decreasing distribution for the dijet invariant mass, $m_{jj}$, on top of which the presence of a new resonant state may appear as a localized excess near the mass of the resonance. Results from previous investigations of dijet distributions carried out at the $\sqrt{s}$ [12], the Tevatron [5, 13, 14], and the Large Hadron Collider (LHC) [15, 16] were found to be in agreement with the SM QCD predictions.

In this note we present a model-independent search for deviations from SM expectations in the dijet invariant mass spectrum using a dataset of pp collisions with a center-of-mass energy of $\sqrt{s} = 13$ TeV collected at the LHC [17] at CERN from 2015 to 2018 with a total integrated luminosity of 139 fb$^{-1}$. After the model-independent search, limits are set on the masses of excited quarks, as well as on generic Gaussian-shaped signals with relative widths up to 15% of the signal mass.

2 ATLAS detector

The ATLAS experiment [18] at the LHC is a multi-purpose particle detector with a forward-backward symmetric cylindrical geometry with layers of tracking, calorimeter, and muon detectors over nearly $4\pi$ coverage in solid angle around the pp collision point. The directions and energies of high-$p_T$ hadronic jets are measured using silicon tracking detectors and a transition radiation straw-tube tracker, hadronic and electromagnetic calorimeters, and a muon spectrometer. Hadronic energy measurements are provided by a calorimeter with scintillator active layers and steel absorber material for the pseudorapidity range ($|\eta| < 1.7$), while electromagnetic (EM) energy measurement are provided by a calorimeter with liquid argon (LAr) active material and lead absorber material covering the pseudorapidity range ($|\eta| < 3.2$). The end-cap and forward regions are instrumented with LAr calorimeters for both EM and hadronic energy measurements up to $|\eta| = 4.9$. A two-level trigger system is used to select events [19], the lower-level being implemented in hardware using a subset of the detector information to reduce the accepted rate to at most 100 kHz, and followed by a software-based level that reduces the rate of recorded events to 1 kHz.

---

1 ATLAS uses a right-handed coordinate system with its origin at the nominal interaction point (IP) in the center of the detector and the z-axis along the beam pipe. The x-axis points from the IP to the center of the LHC ring, and the y-axis points upwards. Cylindrical coordinates ($r, \phi$) are used in the transverse plane, $\phi$ being the azimuthal angle around the z-axis. The pseudorapidity is defined in terms of the polar angle $\theta$ as $\eta = -\ln \tan(\theta/2)$. Angular distance is measured in units of $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$. 

3 Event Selection

Calorimeter cells with an energy deposit significant with respect to the calorimeter noise are grouped together based on their proximity to form topological clusters [20, 21]. These are then grouped into jets using the anti-$k_t$ algorithm [22] with radius parameter $R = 0.4$ as implemented in FastJet [23]. Jet four momenta are computed by summing over the topological clusters associated to the jet, treating the energy deposit of each cluster as resulting from a zero-mass four momentum. Jets with $p_T$ above 20 GeV are reconstructed in the ATLAS calorimeters with an efficiency of approximately 100% [24]. Calorimeter jets are calibrated using simulated events according to the procedure described in Refs. [25–29], followed by a residual calibration that takes into account the differences between data and simulation, and different detector response in the central ($|\eta| < 0.8$) versus forward ($|\eta| > 0.8$) region through the *in situ* calibration techniques described in Refs. [30, 31]. For jets with $p_T$ larger than 2.3 TeV the calibration is fixed to its last value due to the lack of statistics to extend the *in situ* calibrations to higher momenta.

The resulting uncertainty on the jet energy scale varies from 1% for central jets with $p_T$ of 500 GeV to 3% for jets with $p_T$ of 2 TeV, after which the residual uncertainties are due to the single-particle response measurement uncertainty [31]. The dijet invariant mass resolution is 2.9% and 2.4% for dijet masses of 2 and 5 TeV respectively, determined as described in Ref. [9].

Dijet events are selected from $pp$ collisions using a trigger that requires at least one jet reconstructed with a $p_T$ greater than 420 GeV. This is the lowest-$p_T$ single-jet trigger for which all selected events are recorded.

Events containing at least two jets are considered in this analysis, provided that the $p_T$ of the two leading ones are greater than 150 GeV. To reduce the dominant contribution from QCD $t$-channel processes by selecting central events, a kinematic requirement has been applied on the rapidity difference between the two leading jets, $|y^*| = \frac{1}{2}|y_1 - y_2| < 0.6$. The dijet invariant mass for which the trigger requirement is fully efficient within the detector acceptance is 1.1 TeV, which is the lower bound of the $m_{jj}$ range to be investigated. The dijet pair with highest invariant mass in the dataset has $m_{jj} = 8.02$ TeV and was collected in 2016.

4 Simulated Event Samples

Monte Carlo (MC) events from multijet production described by QCD are generated with Pythia 8.186 [32] using the A14 [33] set of tuned parameters for the underlying event and the leading-order (LO) NNPDF2.3 [34] parton distribution functions (PDFs). The renormalization and factorization scales are set to the average $p_T$ of the two leading jets. Pythia calculations use matrix elements that are at leading order in the QCD coupling constant, with simulation of higher-order contributions partially covered by the parton shower modeling. Pythia also models hadronization effects. The distributions of events predicted by Pythia are reweighted to next-to-leading-order QCD predictions of NLOJET++ [35–37] using mass- and angle-dependent correction factors defined as in Ref. [38]. Electroweak effects are included as additional mass- and angle-dependent correction factors [39].

---

2 The rapidity of an outgoing parton is $y = \frac{1}{2} \ln [(E + p_z)/(E - p_z)]$, where $E$ is its energy and $p_z$ is the component of its momentum along the $z$-axis.
All simulated samples include the effects of multiple $pp$ interactions in the same and neighboring bunch crossings (pileup)\(^3\) and are processed through the ATLAS detector simulation [42] based on Geant4 [43]. The same software used to reconstruct data is also used to reconstruct simulated events. The simulated QCD events are only used to test the data-based background estimate used for the $m_{jj}$ distribution, and to provide qualitative comparisons to kinematic distributions in data.

Excited quark signal samples are generated for masses in the 2 TeV to 8 TeV range, in 0.5 TeV steps, with parton-level generators assuming spin-$\frac{1}{2}$ excited quarks with the same coupling constants as SM quarks; no interference with the SM is included in the simulation. The signal model is simulated using Pythia 8.186, in an identical manner to QCD processes, using the same PDFs and parameters for non-perturbative effects. The signal samples are processed through the ATLAS detector simulation using the same framework as used for QCD processes but with a simplified parameterization of the calorimeter [44] to reduce processing time. No difference between full simulation and this faster simulation is observed in the kinematic variables relevant for this search. Only the decay of the excited quark to a gluon and an up- or down-type quark is simulated, which corresponds to a branching ratio of 85%. Before parton shower effects are taken into account, the intrinsic width of the $q^*$ signals is comparable to the detector resolution, while after showering a radiative tail is present. This effect increases in strength for higher $q^*$ masses, augmented by the impact of PDFs decreasing towards higher mass. For the range of $q^*$ masses considered, the reconstruction efficiency is close to unity, henceforth acceptance times efficiency is referred to as acceptance and is computed from all events which pass the analysis selection (see Sec. 4), including distribution tails caused by the sharp rise of PDFs at low Bjorken $x$ [45]. For $q^*$ masses greater than 2 TeV the selection acceptance is approximately constant. The selection acceptance, excluding the branching ratio to quark-gluon, for a $q^*$ with a mass of 4 TeV is 58%.

For the simulated signal samples described above, systematic uncertainties due to jet energy scale, acceptance uncertainties associated to the choice of PDF, and luminosity are included in the limit setting. The jet energy uncertainty ranges from 1.5% at the lowest masses to 3% for masses above 4.5 TeV. A flat 1% uncertainty is considered as a conservative estimate of the PDF, factorization and renormalization scale uncertainties on the acceptance [10]. The uncertainty in the combined 2015-2018 integrated luminosity is 1.7%. It is derived, following a methodology similar to that detailed in Ref. [46], and using the LUCID-2 detector for the baseline luminosity measurements [47], from calibration of the luminosity scale using x-y beam-separation scans.

5 Analysis of the dijet invariant mass spectrum

The $m_{jj}$ distribution formed from the two leading jets in events satisfying the criteria described in Section 3 is analyzed to search for contributions due to BSM resonances. Bin widths for this distribution are chosen to reflect the evolution of the $m_{jj}$ resolution and therefore widen as the mass increases, from about 30 GeV at the lowest $m_{jj}$ values (1.1 TeV) to about 140 GeV at the highest values (8 TeV). The background estimate is derived from the data using the sliding-window method described in [10] using a background parametrization of the form

$$f(x) = p_1 (1 - x)^{p_2} x^{p_3 + p_4 \text{ln} x + p_5 (\text{ln} x)^2} \quad (1)$$

\(^3\) This is reproduced by overlaying simulated soft QCD processes from Pythia 8.205 [32] using the A2 set of tuned parameters [40] and the MSTW2008LO PDF set [41] onto the hard-scattering process.
where \( x \equiv m_{jj}/\sqrt{s} \) and \( p_5 = 0 \) in the nominal fit. Validation checks, including results from pseudo-experiments based simulated data, varying the average size of the search window, linearity checks and signal injection studies as described in Refs. [10] have been performed successfully. Introduction of spurious signal by background fit has been tested by fitting hundreds of representative background data sets, and comparing the strength of the extracted signal to the background prediction uncertainty. The fit has been shown to be robust against spurious signals apart from the case of a 15% wide gaussian signal with \( m_G = 6 \text{ TeV} \), in which case a systematic uncertainty has been considered to take into account this effect. The size of the sliding window is approximately half of the total number of bins, which is wide enough to fit the excited quark signal within an individual window. The uncertainty on the values of the parameters in Eq. (1) is estimated by repeating the sliding-window fitting procedure on pseudo-data built as Poisson fluctuations from the nominal background prediction, that is, the fit result in data. The uncertainty in each \( m_{jj} \) bin is measured as the root mean square of the fit results for all pseudo-experiments in that bin. The uncertainty due to the choice of background parametrization is estimated by performing the background fitting procedure using a fit function with \( p_5 \) floating in Eq. 1. The difference between this alternative background prediction and the nominal one, averaged across a set of pseudo-data, is considered as a systematic uncertainty.

Figure 1 shows the observed \( m_{jj} \) distribution for the selected events and the background estimate from the sliding-window fit. The reduced \( \chi^2 \) of the sliding window fit is approximately 1 for 85 degrees of freedom.

The BumpHunter algorithm [48, 49] quantifies the statistical significance of any localized excess in the \( m_{jj} \) distribution. The algorithm compares the binned \( m_{jj} \) distribution of the data to the background prediction, considering contiguous mass intervals in all possible locations, up to a width of half of the distribution. From computing the significance of any excess for each interval in the scan the most discrepant interval is found to be \( 7.052 - 7.326 \text{ TeV} \), indicated by the two vertical lines in Figure 1. The global significance of this outcome is evaluated using the ensemble of possible outcomes across all intervals scanned, by applying the algorithm to pseudo-data samples drawn randomly from the background fit. Without including systematic uncertainties, the probability that fluctuations of the background model would produce an excess at least as significant as the one observed in the data anywhere in the distribution (the BumpHunter probability) is 0.8, thus there is no evidence of a localized contribution from BSM phenomena.

### 6 Statistical interpretation

The HistFitter [50] framework is used to extract 95% confidence-level (CL) upper limits on the cross-section times acceptance for a discrete set of \( q^* \) mass hypotheses. The limits were calculated using the CLs method [51] with a binned profile likelihood ratio as the test statistic and approximated by the asymptotic formulae [52] to speed up the evaluation process. For signal masses of 6.5 TeV or larger the test statistic distribution is calculated using pseudo-experiments given the poor reliability of the asymptotic approach in the tail of the spectrum, where the effective number of events present is small. The expected limits are calculated using pseudo-experiments generated from the maximum-likelihood values of the background uncertainties in the sliding-window background model. These are evaluated comparing the \( m_{jj} \) spectrum of data events and the background estimate from sliding-window fit against the different \( q^* \) simulated signals. The calculated limit is logarithmically interpolated between the different mass hypotheses. No uncertainty is taken into account on the theoretical cross-section of the signal. The effects of the systematic uncertainties on the data-driven background together with the systematic uncertainties
Figure 1: The reconstructed dijet mass distribution, $m_{jj}$, is shown for events with $p_T > 150$ GeV for the two leading jets, with $|y^*| < 0.6$, and $m_{jj}$ greater than 1.1 TeV (filled points). The solid line depicts the background prediction from the sliding-window fit. The vertical lines indicate the most discrepant interval identified by the BumpHunter algorithm [48, 49], for which the $p$-value is reported in the figure. The expected contributions for $q^*$ signal with a mass of 4 and 5 TeV are overlaid, normalized to 0.1 times their predicted cross section. The lower panel shows the bin-by-bin significance of the data-fit discrepancy, based only on statistical uncertainties.

on the signal sample acceptance, described in Section 4, are considered in the limit-setting procedure. These uncertainties are incorporated into the limits by varying all of the uncertainty sources according to Gaussian probability distributions.

Confidence intervals are then calculated from the resulting profile of the parameter-of-interest of the likelihood. The limit obtained on $q^*$ signal cross-section is shown in Figure 2(a). Thus, $q^*$ signals with mass below 6.7 TeV (6.4 TeV) are excluded (expected to be excluded) at 95% CL.

Exclusion limits are also expressed on the cross-section times acceptance times branching ratio to two jets, $\sigma \times A \times BR$, of a hypothetical signal modeled as a Gaussian peak in the particle-level $m_{jj}$ distribution (as in Ref. [9]). Gaussian signal models are tested for different mass hypotheses, $m_G$, and different possible widths of the signal, $\sigma_G$, at the detector reconstruction level. Signal widths range from the detector resolution width of approximately 3%, which would correspond to the case of a resonance with an intrinsic width negligible with respect to the detector resolution, up to a relative width of $\sigma_G/m_G = 15\%$. For resonances broader than the considered widths the presence of the signal would significantly affect the background estimation obtained using the sliding-window fit. To evaluate the effect of systematic uncertainties on the
Figure 2: The 95% CL upper limit obtained from the dijet invariant mass ($m_{jj}$) distribution on cross-section times acceptance times branching ratio to two jets, $\sigma \times A \times BR$, as a function of (a) the mass of a $q^*$ signal ($m_{q^*}$) and (b) the mass of a hypothetical signal that produces a Gaussian-shaped contribution to the $m_{jj}$ distribution. For Gaussian-shaped signals the observed limits are reported for different width hypotheses ($\sigma_G$). The expected limit and corresponding $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands are also indicated for the $q^*$ model in (a). Limits corresponding to a Gaussian-shaped signal with a relative width of 15% are set up to $m_G = 6$ TeV due to the poor background estimation when a broad signal overlaps the upper end of the $m_{jj}$ spectrum.

(a) $q^*$

(b) Gaussian peak

7 Conclusion

This note presents a search for BSM phenomena producing a localized excess in the dijet invariant mass spectrum using the dataset collected between 2015 to 2018 by the ATLAS experiment from the Large Hadron Collider proton-proton collisions at $\sqrt{s} = 13$ TeV, corresponding to a total integrated luminosity of 139 fb$^{-1}$.

No significant excess has been observed over a smoothly-falling background prediction. The data are found to be in agreement with a background only hypothesis with $p$-value of 0.8. This analysis excludes $q^*$ models with masses below 6.7 TeV at 95% CL. It also sets 95% CL upper limits on the cross-section times branching ratio for new processes that would produce a Gaussian-shaped contribution to the dijet mass distribution. These results substantially extend the excluded ranges by around 700 GeV for mass limits of.
$q^*$ models obtained using the data set collected in 2015 and 2016 dataset, corresponding to a luminosity of 37.4 fb$^{-1}$.

**References**


Appendix

Figure 3: A visualization of the highest-mass dijet event, (Event 4144227629, Run 305777) recorded in 2016: the two central high-$p_T$ jets each have transverse momenta of 3.74 TeV, they have a $y^*$ of 0.38 and their invariant mass is 8.02 TeV.
Figure 4: A visualization of the highest-mass dijet event, (Event 4144227629, Run 305777) recorded in 2016: the two central high-$p_T$ jets each have transverse momenta of 3.74 TeV, they have a $y^*$ of 0.38 and their invariant mass is 8.02 TeV.

Figure 5: A visualization of the highest-mass dijet event, (Event 4144227629, Run 305777) recorded in 2016: the two central high-$p_T$ jets each have transverse momenta of 3.74 TeV, they have a $y^*$ of 0.38 and their invariant mass is 8.02 TeV.