The following full text is a publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/195238

Please be advised that this information was generated on 2019-07-25 and may be subject to change.
Search for Low-Mass Dijet Resonances Using Trigger-Level Jets with the ATLAS Detector in pp Collisions at $\sqrt{s} = 13$ TeV

M. Aaboud et al.
(ATLAS Collaboration)

(Received 11 April 2018; published 22 August 2018)

Searches for dijet resonances with sub-TeV masses using the ATLAS detector at the Large Hadron Collider can be statistically limited by the bandwidth available to inclusive single-jet triggers, whose data-collection rates at low transverse momentum are much lower than the rate from standard model multijet production. This Letter describes a new search for dijet resonances where this limitation is overcome by recording only the event information calculated by the jet trigger algorithms, thereby allowing much higher event rates with reduced storage needs. The search targets low-mass dijet resonances in the range $450 - 1800$ GeV. The analyzed data set has an integrated luminosity of up to $29.3 \text{ fb}^{-1}$ and was recorded at a center-of-mass energy of 13 TeV. No excesses are found; limits are set on Gaussian-shaped contributions to the dijet mass distribution from new particles and on a model of dark-matter particles with axial-vector couplings to quarks.

DOI: 10.1103/PhysRevLett.121.081801

Introduction.—If new particles beyond those of the standard model (SM) are directly produced in proton-proton ($pp$) collisions at the Large Hadron Collider (LHC), they must interact with the constituent partons of the proton, and can therefore also decay into the same partons, resulting in two-jet final states. Quantum chromodynamics (QCD) predicts that dijet events have an invariant mass distribution ($m_{jj}$) that falls smoothly, whereas a new state decaying to two partons would emerge as a localized excess in the distribution.

Traditional dijet searches at the LHC focus on the production of heavy particles with masses above 900 GeV [1–3]. LHC searches for lighter resonances with small production cross sections have been hampered by restrictions in the data-taking rate of the ATLAS and CMS detectors. Single-jet triggers with a jet $p_T$ threshold below roughly 380 GeV are prescaled, a procedure whereby only a fraction of the events passing the trigger are recorded; hence, dijet events with an invariant mass below 1 TeV are largely discarded by the trigger system, as indicated in Fig. 1. Therefore, despite the large number of $pp$ collisions produced by the LHC, traditional ATLAS and CMS searches are less sensitive to dijet resonances below 900 GeV than searches at the SPS and Tevatron colliders [4–9]. Alternative trigger strategies to search for low-mass resonances include selecting events with jets recoiling against either an energetic photon or an additional energetic jet [10–12], or selecting events with decays to heavy-flavor jets [13,14]. In these cases, additional features in the events reduce the data-taking rates, reducing the sensitivity to low-mass resonances.

This Letter describes an innovative data-taking approach to access the invariant mass region below 1 TeV; only a reduced set of information from the trigger system is recorded and subsequently analyzed. The Trigger-object Level Analysis (TLA) approach allows jet events to be recorded at a peak rate of up to twice the total rate of events using the standard approach, while using less than 1% of

![FIG. 1. Comparison between the number of dijet events in the data used by this analysis (black points), the number of events selected by any single-jet trigger (thicker, blue line), and the events selected by single-jet triggers but corrected for the trigger prescale factors (thinner, red line) as a function of the dijet invariant mass ($m_{jj}$). The definition of $y^{*}$ is $(y_1 - y_2)/2$, where $y_1$ and $y_2$ are the rapidities of the highest- and second-highest $p_T$ jets.](image)
the total trigger bandwidth [15]. This strategy is used in dijet resonance searches in $\sqrt{s} = 8$ and 13 TeV LHC $pp$ collision data by the CMS Collaboration [16,17], and it is used by the LHCb Collaboration (e.g. [18]). The analysis presented here uses 29.3 fb$^{-1}$ of $\sqrt{s} = 13$ TeV $pp$ collision data recorded in 2016 by the ATLAS detector.

**ATLAS detector and data sample.**—The ATLAS detector [19] is a multipurpose detector with a forward-backward symmetric cylindrical geometry and nearly 4$\pi$ coverage in solid angle [20], consisting of tracking detectors, calorimeters, and a muon spectrometer. In the pseudorapidity region $|\eta| < 3.2$, high-granularity lead and liquid-argon (LAr) electromagnetic sampling calorimeters are used. A steel and scintillator hadronic tile calorimeter provides coverage in the range $|\eta| < 1.7$. Hadronic calorimetry in the endcap region, $1.5 < |\eta| < 3.2$, and electromagnetic and hadronic calorimetry in the forward region, $3.1 < |\eta| < 4.9$, are provided by LAr sampling calorimeters. A two-level trigger system is used to select events for offline storage [15]. A first-level (L1) trigger based on dedicated hardware identifies jets from $J$ in the range $|\eta| < 1.7$, with an anti-$k_T$ algorithm [21,22] with a radius parameter $R = 0.4$. The inputs to this algorithm are calorimeter segments (topological clusters), in which each cell’s inclusion is based on the significance of its energy deposit over calorimeter noise [23]. Jet four-momenta are computed by summing over the four-momenta of the topological clusters that compose the jet, with each cluster pointing to the center of the ATLAS detector and being treated as massless. The HLT jet reconstruction uses the same techniques that the ATLAS offline jet reconstruction applies to similar inputs from recorded data events that include the full detector information [15].

After execution of the HLT jet algorithm, only trigger-level jets with $p_T > 20$ GeV are stored. The stored information includes the four-momentum of each jet and a set of calorimeter variables characterizing the jet [24], such as information about the jet quality and structure. The average size of these events is less than 0.5% of the size of full offline events that contain all detector information. For this analysis, all events containing at least one L1 jet with $E_T > 100$ GeV are selected and recorded, corresponding to a total luminosity of 29.3 fb$^{-1}$. Events containing a L1 jet with $E_T > 75$ GeV are also selected, in a subset of these data corresponding to 3.6 fb$^{-1}$ that were collected at the end of the data-taking period. Events with at least one L1 jet with $E_T > 100$ GeV are therefore included in both data sets.

**Calibration procedure.**—After the events are recorded, the trigger-level jet energy and direction are corrected to those of simulated particle-level jets built from stable particles with a lifetime longer than 10 ps, excluding muons and neutrinos. Before any calibration, the jet $p_T$ response, defined as the $p_T$ ratio of a trigger-level jet to the same jet [25] reconstructed offline (offline jet), is between 0.95 and 1.05 in the $p_T$ range considered by this analysis. Since the energy scale for trigger-level and offline jet are very similar, the trigger-level jet calibration employs the same procedure and constants as derived for offline jets [26], with some modifications to account for the unavailability of tracking information for trigger-level jets.

In the calibration procedure, summarized in Fig. 2, an event-by-event jet-area-based calibration [27] is used to correct for contributions from additional proton-proton interactions (pileup) in the same and neighboring crossings of proton bunches. Then, the simulation-based calibration derived for offline jets is applied to trigger-level jets to correct both jet energy and direction. Next, calorimeter-based variables are used to reduce the dependence on the trigger-level jet flavor and to minimize the impact of energy leakage. Only variables related to the trigger-level jet energy fractions in the electromagnetic and hadronic calorimeters and the minimum number of calorimeter cells containing 90% of the trigger-level jet energy are used here since track-based variables, which are normally used in the offline calibration, are not available. With this correction, the trigger-level jet energy resolution is improved by 8% at jet $p_T$ values of 85 GeV and up to 40% for jet $p_T$ values of 1 TeV relative to the previous calibration step. Next, the calibration corrections that restore the relative calibration between central and forward jets in data and simulation are derived for offline jets and applied to trigger-level jets. After these calibration steps, any residual difference between trigger-level jets and offline jets is accounted for in a dedicated correction, based on the $p_T$ response and derived from data in bins of jet $\eta$ and $p_T$. The size of this correction is on average 1%, with values reaching up to 4% in the endcap regions of the calorimeter.

Finally, an *in situ* calibration is obtained from the data-to-simulation ratio of the $p_T$ balance between offline jets and well-calibrated objects against which the jets recoil. Three different types of well-calibrated objects are used to span the full $p_T$ range of the jets: $Z$ bosons decaying to electrons or muons, photons, and multijets. A polynomial in $\log(p_T)$ is simultaneously fit to the three input measurements to combine them. The resulting curve is taken as

![Calibration stages for EM-scale trigger-level jets](image)

**FIG. 2.** Calibration stages for EM-scale trigger-level jets, each applied to the four-momentum of the jet. MC refers to the simulation.
contain at least two trigger-level jets with the same criteria as those used offline for jets with energies \( [7,28] \) have found a variety of simple functional forms to describe this shape; however, given the statistical precision of the data and the wide invariant mass range covered by this search, none of the single, simple functional forms can provide a good description of the data.

The SM background distribution is determined using a sliding-window fit [3], where a fitted functional form is evaluated at the bin at the center of a window, which then slides in one-bin steps along the \( m_{jj} \) spectrum. The evaluated background estimates evaluated in each bin are then collated to form the final background estimate. The signal selection with \( |y^*| < 0.6 \) selects higher-\( p_T \) jets at a given invariant mass and thus provides a mass distribution that is unbiased by the leading-jet selection from \( m_{jj} = 450 \) GeV.

**Background estimation.**—The invariant mass spectrum expected from SM dijet production is predicted to be smooth and falling. Prior dijet searches at various collision energies [7,28–32] have found a variety of simple functional forms to describe this shape; however, given the statistical precision of the data and the wide invariant mass range covered by this search, none of the single, simple functional forms can provide a good description of the data.

The SM background distribution is determined using a sliding-window fit [3], where a fitted functional form is evaluated at the bin at the center of a window, which then slides in one-bin steps along the \( m_{jj} \) distribution. The evaluated background estimates evaluated in each bin are then collated to form the final background estimate. The signal selection with \( |y^*| < 0.6 \) uses a window size of 19 bins in the \( m_{jj} \) spectrum from 531 to 2080 GeV, which spans 34 bins in total. The signal selection with \( |y^*| < 0.3 \) uses a window size of 27 bins over a total of 40 bins, in the range \( 400 < m_{jj} < 2080 \) GeV. The bin sizes have been chosen according to the simulated invariant mass resolution: \( \sigma_{p_T}/p_T = 10.6/p_T \pm 0.27/\sqrt{p_T} \pm 0.039 \). The sliding window, however, cannot be extended beyond the lower edge of the \( m_{jj} \) range used in each signal selection. Therefore, for the first 9 (13) bins in the \( |y^*| < 0.6 \) (\( |y^*| < 0.3 \)) signal selection, which corresponds to one half of the window size, the window is fixed to the lower edge of the spectrum and instead of evaluating the fitted functional form at the window center, it is evaluated for each bin in turn. For invariant masses higher than the \( m_{jj} \) range used for the search, the window is allowed to extend
beyond the range, to 2970 (3490) GeV for the $|y^*| < 0.6$ (0.3) signal selection, and the fit is evaluated at the center of the window.

In each sliding window, three functional forms are fit to the data: a five-parameter function of the form

$$f(x) = p_1(1 - x)p_2x^{p_3} + p_4\ln x + p_5\ln x^2,$$

where $p_i$ are free parameters and $x \equiv m_{jj}/\sqrt{s}$; a four-parameter function, which is the same as Eq. (1) but with $p_4 = 0$; and a four-parameter function used by the UA2 Collaboration [28], defined as

$$f(x) = \frac{p_1}{x^{p_2}}e^{-p_3x - p_4x^2}.$$

The function used for each signal selection is the one that yields the best $\chi^2$ over the full fitted $m_{jj}$ range. An alternative function is chosen to evaluate a systematic uncertainty. For the signal selection with $|y^*| < 0.6$, Eq. (1) is used, yielding a $\chi^2$ $p$ value of 0.13, while the alternative function is the four-parameter function with a $\chi^2$ $p$ value of 0.11. For the signal selection with $|y^*| < 0.3$, the four-parameter version of Eq. (1) yields the best $\chi^2$ $p$ value of 0.42 and the alternative function is Eq. (2), with a $\chi^2$ $p$ value of 0.35.

The size of the sliding window is optimized to yield the best $\chi^2$ value for the full $m_{jj}$ range while still being larger than the width of the expected signals and therefore insensitive to potential signal contributions. This latter requirement is checked by including signal models in pseudo-data samples and studying the dependence of the signal sensitivity on different window sizes.

Systematic uncertainties in the estimate of the background used in setting limits include the uncertainty due to the choice of functional form and uncertainties in the fit parameter values. The effect of the choice of functional form is evaluated by comparing the nominal function to the alternative. The uncertainties in the fit parameter values are evaluated using pseudoeperiments, where the pseudodata are drawn from Poisson fluctuations around the nominal background model.

**Results and limits.**—Figure 4 shows the invariant mass distributions for dijet events in each signal region including the results from the sliding-window background estimates. The $\chi^2$ $p$ value of the overall background is 0.13 for the $|y^*| < 0.6$ signal selection and 0.42 for the $|y^*| < 0.3$ signal selection, indicating that the data agrees well with the background estimate. The most discrepant interval identified by the BumpHunter algorithm [33,34] is 889–1007 GeV for events with $|y^*| < 0.6$. Accounting for statistical uncertainties only, the probability of observing a deviation at least as significant as that observed in data, anywhere in this distribution, is 0.44 and corresponds to significance of 0.16$\sigma$. Thus, there is no evidence of any localized excess.

Limits are set on both a leptophobic $Z'$ simplified dark-matter model [36] and a generic Gaussian model. The $Z'$ simplified model assumes axial-vector couplings to SM quarks and to a Dirac fermion dark-matter candidate. No interference with the SM is simulated. Signal samples were generated so that the decay rate of the $Z'$ into dark-matter particles is negligible and the dijet production rate and resonance width depend only on the coupling of the $Z'$ to quarks, $g_q$, and the mass of the resonance, $m_{Z'}$ [9]. The model’s matrix elements were calculated in MADGRAPH 5 [37] and parton showering was performed in PYTHIA 8 [38] with the A14 set of tuned parameters for underlying event [39] and NNPDF2.3 parton distribution functions [40].

The width of a $Z'$ resonance with $g_q = 0.10$, including parton shower and detector resolution effects, is approximately 7%. Limits are set on the cross section, $\sigma$, times acceptance, $A$, times branching ratio, $B$, of the model, and then displayed in the $(g_q, m_{Z'})$ plane [41]. The acceptance for a mass of 550 GeV is 20% for a $Z'$ simplified model with $g_q = 0.10$ for the $|y^*| < 0.3$ signal selection, and 41% for a signal of mass equal to 750 GeV for the $|y^*| < 0.6$ signal selection.

Limits are also set on a generic model where the signal is modeled as a Gaussian contribution to the observed $m_{jj}$ distribution. For a given mean mass, $m_G$, four different Gaussian widths are considered: a width equal to the simulated mass resolution (which ranges between 4% and 6%), and the fixed fractions 5%, 7%, and 10% of $m_G$. As the width increases, the expected signal contribution is distributed across more bins. Wider signals are
therefore less affected by statistical fluctuations from the data in a single bin. The results can be used to set limits on models of new phenomena besides that of the $Z'$ simplified model and are applicable when the resonance is sufficiently narrow and the parton distribution function and nonperturbative effects can be safely truncated or neglected, as described in Ref. [31]. These criteria are often met if the $m_{jj}$ distribution for a signal approaches a Gaussian distribution after applying the kinematic selection criteria of the resonance analysis, so that 95% of the signal lies within 20% of the Gaussian mean mass. Models of new resonances with an intrinsic width much smaller than 5% of its mass should be compared to the results with a width equal to the experimental resolution. For models with a larger width, the limit that best matches their width should be used. More-detailed instructions can be found in Appendix A of Ref. [31].

A Bayesian method is applied to the data and simulation of the signal models at a series of discrete masses to set 95% credibility-level upper limits on the cross section times acceptance [30] for the signals described above. The method uses a constant prior for the signal cross section and Gaussian priors for nuisance parameters corresponding to systematic uncertainties. For both observed and expected limits, the sliding window fit is performed for each value of the mass parameter, adding the tested signal shape with a floating normalization to the functional forms stated above. The expected limits are calculated using pseudoexperiments generated from the fit parameters of those functional forms. The uncertainties on the $Z'$ signal model include the jet energy scale and the luminosity. The impact of the jet energy resolution uncertainty is negligible. For the Gaussian model, a constant jet energy scale uncertainty of 3% is applied in accordance with the measured impact of this uncertainty on the $Z'$ samples. The uncertainty in the integrated luminosity is ±2.2%, derived following a methodology similar to that detailed in Ref. [42]. The systematic uncertainties in the background estimate include the choice of the fit function and the uncertainty in the fit parameter values, as described above.

Figure 5 shows limits on the coupling to quarks, $g_q$, as a function of the mass $m_{Z'}$ for the $Z'$ model. Figure 6 shows limits on a possible Gaussian contribution with a width equal to the detector resolution as a function of the mean mass, $m_G$. In both the $Z'$ and Gaussian models, upper limits for masses from 450 to 700 GeV are derived using the distribution with $|y^*| < 0.3$, which is sensitive to the lower masses. Limits for masses above 700 GeV are derived from the $m_{jj}$ distribution with $|y^*| < 0.6$, except for Gaussian signals with a width of 10% where only the $|y^*| < 0.3$ distribution is used.

The limit results show an upward fluctuation at masses of approximately 1 TeV in the $|y^*| < 0.6$ signal region. This is not seen in Fig. 4; when searching for excesses in the data, a background-only hypothesis is used for the sliding window fit. In the observed and expected limits, the fit includes the signal shape in addition to the background parameterization, and can adapt to local data fluctuations that mimic a signal shape. The $|y^*| < 0.6$ signal region, which uses a smaller sliding-window size, is especially sensitive to the difference in the two approaches. Therefore,
limits were not set on signals with a width of 10% for the $|y^*| < 0.6$ signal region as the signal is too wide for the sliding-window size.

Conclusions.—In conclusion, this analysis searches for resonances with masses between 450 GeV and 1800 GeV in dijet events using trigger-level jets in 29.3 fb$^{-1}$ of $\sqrt{s} = 13$ TeV proton-proton collision data recorded by the ATLAS detector at the LHC. The invariant mass distribution presents no significant local excesses compared to the estimated SM background. This analysis provides 95% credibility-level limits on $Z'$ signals and cross sections for new processes that would produce a Gaussian contribution to the dijet mass distribution. Over much of the mass range, the sensitivity to the coupling to quarks, $g_q$, is improved by a factor of 2 or more compared to pre-LHC and $\sqrt{s} = 8$ and 13 TeV ATLAS results, and is comparable to CMS searches at $\sqrt{s} = 8$ and 13 TeV using this technique. Gaussian contributions with effective cross sections times acceptance ranging from approximately 6.5 pb at 450 GeV, to 0.4 pb at 700 GeV, to 0.05 pb at 1800 GeV are excluded.

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently. We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWFW and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS, CEA-DRF/IRFU, France; SNSF, Switzerland; Alexander von Humboldt Foundation, Germany; INFN, Italy; BRF, Norway; CERCA Programme Generalitat de Catalunya, Generalitat Valenciana, Spain; the Royal Society and Leverhulme Trust, United Kingdom. The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [43].
