Search for New Phenomena in Photon+Jet Events Collected in Proton-Proton Collisions at $\sqrt{s} = 8$ TeV with the ATLAS Detector

The ATLAS Collaboration

Abstract

This note describes a model-independent search for the production of new resonances in photon + jet (γ + jet) events using 20 fb$^{-1}$ of proton-proton LHC data recorded with the ATLAS detector at a center-of-mass energy of $\sqrt{s} = 8$ TeV. The γ + jet mass distribution is compared to a background model fit from data; no significant deviation from the background-only hypothesis is found. Limits are set at 95% credibility level on generic Gaussian-shaped signals and two benchmark phenomena beyond the Standard Model: non-thermal quantum black holes and excited quarks. Non-thermal quantum black holes are excluded below masses of 4.6 TeV and excited quarks are excluded below masses of 3.5 TeV.
Several exotic production mechanisms have been proposed that produce massive $\gamma + \text{jet}$ final states, such as non-thermal quantum black holes (QBH) [1–3], excited quarks [4–5], quirks [6–8]. Regge excitations of string theory [9–12], and topological pions [13]. Of the past searches [14–18], the sole LHC search was done by the ATLAS experiment with 2.11 fb$^{-1}$ in proton-proton ($pp$) collision data obtained at a center-of-mass energy of $\sqrt{s} = 7$ TeV. It found no evidence of new physics and placed limits on the visible signal cross-section in the range 1.5–100 fb and on excited-quark masses up to 2.46 TeV at the 95% credibility level (CL) [18]. This note describes an improved model-independent search over the earlier ATLAS search for $s$-channel $\gamma + \text{jet}$ production. It makes use of improved selection criteria at high mass, reduced background systematic uncertainties, a higher center-of-mass energy of $\sqrt{s} = 8$ TeV, and an order-of-magnitude larger data sample (20.3 fb$^{-1}$). In the absence of new resonances, this note presents the first limits on QBH decaying to the $\gamma + \text{jet}$ final state, as well as generic Gaussian-shaped signals and excited quarks.

The Standard Model (SM) of particle physics lacks a mechanism whereby $pp$ collisions produce resonances that subsequently decay to a $\gamma + \text{jet}$ final state. Direct $\gamma + \text{jet}$ production can occur at tree level via Compton scattering of a quark and a gluon, or through quark-antiquark annihilation. The former accounts for the majority of direct $\gamma + \text{jet}$ production at all center-of-mass energies. Events with a high transverse momentum photon and one or more jets can also arise from radiation off final-state quarks, or from dijet or multijet processes, where secondary photons are produced during fragmentation of the hard-scattered quarks or gluons [19–22]. The $\gamma + \text{jet}$ invariant mass $m_{\gamma j}$ distribution resulting from this mixture of processes is smooth and rapidly falling, and is therefore well-suited to revealing high-mass resonances decaying to $\gamma + \text{jet}$. The $m_{\gamma j}$ distribution is used to search for a peak over the SM background, estimated by fitting a smoothly falling function to the $m_{\gamma j}$ distribution in the region $m_{\gamma j} > 426$ GeV. In the absence of a signal, Bayes’ theorem is used to set limits on Gaussian-shaped signals and on two benchmark models: QBH and excited quarks.

QBHs are predicted by theories with low-scale quantum gravity, offering a solution to the mass hierarchy problem of the SM by lowering the scale of quantum gravity $M_D$ from the Planck scale ($\sim 10^{16}$ TeV) to a value of about 1 TeV. For example, in the Arkani-Hamed-Dimopoulos-Dvali (ADD) model [23,24], extra dimensions are flat and compactified on a torus. Gravity becomes strong and allows for the formation of QBHs with masses close to $M_D$. Such objects evaporate faster than they thermalize, resulting in non-thermal decays into a few particles, rather than a high-multiplicity final state [25,26]. It is expected that the mass threshold $M_{\text{th}}$ for QBH production cannot be smaller than $M_D$ [25,26]. However, the formation mass can be larger than $M_D$. Regardless of the number of extra dimensions $n$, the signal would appear as a local excess over the steeply falling $m_{\gamma j}$ spectrum near $M_{\text{th}}$ and fall slowly but exponentially at higher masses. This note assumes $M_D = M_{\text{th}}$ and $n = 6$, as used in the literature [3]. Dijet searches for QBH performed by the CMS Collaboration with high-multiplicity energetic final states yielded limits in the range of 4.3–6.2 TeV, for $n = 1–6$ and different model assumptions [27]. The cross section times branching fraction for QBH production and decay to $\gamma + \text{jet}$ final states at $M_{\text{th}} = 1, 3$ and 5 TeV is 200, 0.3 and $6.3 \times 10^{-5}$ pb [3], respectively. For decay to dijet final states at these same threshold masses the rates are larger by factors of 11, 39, and 125.

Excited quark ($q^*$) states, which the ATLAS and CMS Collaborations have also searched for in dijet final states [28,29], could be produced via the absorption of a gluon by a quark. The model is defined by one parameter, the excited quark mass $m_{q^*}$, setting the compositeness scale equal to $m_{q^*}$ and the SU(3), SU(2), and U(1) coupling multipliers to $f_f = f = f' = 1$ [3]. Only gauge interactions are considered. This results in branching fractions for $q^* \rightarrow qg$ and $q^* \rightarrow qq'$ of 0.85 (0.85) and 0.02 (0.005), respectively, for $q = u$ ($q = d$) [1]. The cross section times branching fractions for excited quarks for $m_{q^*} = 1, 3$ and 5 TeV are $4.2 \times 10^{-3}$, and $3 \times 10^{-5}$ pb, respectively.

Factorization and renormalization scale uncertainties and uncertainties on parton distribution func-

---

1 In the model considered for this analysis, only first generation quarks are assumed to be composite.
tions (PDFs) are neglected for both signal types in order to provide a convenient benchmark process for theoretical reinterpretation.

For cross-checking data-driven background estimates, the SM prompt photon processes are simulated with PYTHIA 8.165 [30] and SHERPA 1.4.0 [31]. The PYTHIA and SHERPA prompt photon samples use CTEQ6L1 [32] and CT10 [33] leading-order and next-to-leading-order PDFs, respectively. The simulated samples of QBH are obtained from the qbh 1.05 generator [34], followed by parton showering using PYTHIA 8.165. The $q^*$ simulated signal samples are generated with the excited quark model in PYTHIA 8.165. Both signal generators use the MSTW2008LO [35] leading-order PDF set with the AU2 underlying-event tune [36]. Additional inelastic $pp$ interactions, termed pileup, are included in the event simulation. The mean number of pileup interactions in the simulation is approximately 22. All the above Monte Carlo (MC) simulated samples are produced using the ATLAS full GEANT4 [37] detector simulation [38]. Supplementary studies of the background shape are also performed with the next-to-leading-order JETPHOX 1.3.0 generator [19–21] at parton level using CTEQ10 PDFs.

A detailed description of the detector is available in Ref. [39], and the event selection is similar to that described in Ref. [18]. Photons are detected by a lead-liquid-argon sampling electromagnetic calorimeter (EMC) with an accordion geometry. The EMC has a pre-sampler layer and three additional layers; only the first two are used in photon identification. Upstream of the EMC, the inner detector allows an accurate reconstruction of tracks from the primary $pp$ collision point and also from secondary vertices, permitting an efficient reconstruction of photon conversions in the inner detector. For $|\eta| < 1.37$ an iron-scintillator-tile calorimeter behind the EMC provides hadronic coverage. The endcap and forward regions, $1.5 < |\eta| < 4.9$, are instrumented with liquid-argon calorimeters for both electromagnetic and hadronic measurements. Events are collected with a trigger requiring at least one photon candidate with transverse momentum ($p_T$) above 120 GeV [40]. The integrated luminosity of the data sample is $(20.3 \pm 0.6) \text{ fb}^{-1}$.

Each event is required to contain a primary vertex with at least two tracks with $p_T > 400$ MeV. If more than one vertex is found, the primary vertex is defined as the one with the highest scalar sum $p_T^2$ of associated tracks.

Jets are reconstructed from clusters of calorimeter cells [42] using the anti-$k_T$ clustering algorithm [43] with radius parameter $R = 0.6$. The effects on jet energies due to multiple $pp$ collisions in the same or in neighboring bunch crossings are accounted for by a jet-area-based correction [44,45]. Jet energies are calibrated to the hadronic energy scale using MC and the combination of several in situ techniques applied to data [46]. Events are discarded if the leading jet is affected by noise or hardware problems in the detector, or is identified as arising from non-collision backgrounds. Only jets with $|\eta_j| < 2.8$ are considered further.

Photon candidates are reconstructed from clusters of electromagnetic calorimeter cells and tracking information provided by the inner detector. Inner detector tracking information is used to reject electrons and to recover photons converted to $e^+e^-$ pairs [47]. Candidates satisfy standard ATLAS selection criteria that are designed to reject backgrounds from hadrons [48]. The photon candidates must meet $\eta$-dependent requirements on hadronic leakage and shower shapes in the first two sampling layers of the electromagnetic calorimeter. Energy calibrations are applied to photon candidates to account for energy loss upstream of the electromagnetic calorimeter and for both lateral and longitudinal leakage. The simulation is corrected for differences between data and MC for each photon shower shape variable. Events are discarded if the leading photon consists of calorimeter cells affected by noise bursts or transient
hardware problems.

These photon identification criteria reduce instrumental backgrounds to a negligible level, but some background from fragmentation photons and hadronic jets remains. This background is further reduced with requirements on nearby calorimeter activity. Energy deposited in the calorimeter near the photon candidate, $E_T^{\text{isol}}$, must be no larger than $0.011 p_T^\gamma + 3.65 \text{ GeV}$, a criterion that provides constant efficiency for all pileup conditions and over the entire $p_T$ range explored. This transverse isolation energy is calculated by summing the energy as measured in electromagnetic and hadronic calorimeter cells inside a cone of radius $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} = 0.4$ centered on the photon cluster, but excluding the energy of the photon cluster itself, and is corrected on an event-by-event basis for the ambient energy density due to pileup and the underlying event, as well as energy leakage from the photon cluster into the cone.

Photon energy deposits in the calorimeter are also reconstructed as jets. To further suppress background from fragmentation photons, where the angular separation between the photon and the corresponding jet can be large, the leading photon candidate is required to have exactly one jet reconstructed within $\Delta R(\gamma, \text{jet}) < 0.1$, and the angular separation between the leading photon and all other jets with $p_T > 30 \text{ GeV}$ is required to be $\Delta R(\gamma, \text{jet}) > 1.0$.

Events containing at least one photon candidate and at least one jet candidate, each with $p_T > 125 \text{ GeV}$, are selected for final analysis. The photon trigger is fully efficient for these events. Sub-leading photons or jets are allowed; when more than one photon or jet is found, the highest $p_T$ candidates are selected to constitute the photon and jet pair to compute $m_{\gamma j}$.

The sensitivity of the search is improved with requirements on photon and jet pseudorapidities. Dijet production rates increase with jet pseudorapidity whereas rates for an s-channel signal would diminish. Photon acceptance is restricted to the barrel calorimeter, $|\eta_\gamma| < 1.37$, and requires $|\eta_\gamma - \eta_j| < 1.6$ between the photon and jet. The latter requirement was chosen by optimizing the expected significance using the $|\eta_\gamma - \eta_j|$ distribution found in QBH and excited quark signal simulations, as well as PYTHIA prompt photon simulation. The overall acceptance is about 55% and the overall efficiency falls from 90% to 82% for masses from 1 TeV to 6 TeV for QBH signals and from 90% to 87% for excited quark signals over the same mass range. There are 285356 events in the data sample after all event selections. The highest $m_{\gamma j}$ observed is 2.57 TeV. ATLAS event displays for this event and the event with the second-highest observed $m_{\gamma j}$ are included at the end of this note.

Figure [1] shows the resulting distribution of the $\gamma + \text{jet}$ invariant mass. The bin widths are chosen to be twice the mass resolution at the central value of each bin. The resolution is about 4% of $m_{\gamma j}$ at 1 TeV, improving to about 3% at 2 TeV. The combined SM and instrumental background to the search is determined by fitting this distribution to the four-parameter ansatz function

$$f(x \equiv m_{\gamma j}/\sqrt{s}) = p_1 (1 - x)^{p_2} x^{-p_3 + p_4 \ln x}.$$  

The motivation for this function is discussed in Ref. [49]. The functional form has been tested with PYTHIA and SHERPA prompt photon simulations and next-to-leading-order JETPHOX predictions with comparable event statistics. The functional form also describes the $m_{\gamma j}$ distribution in data control samples. Specifically, one control sample is defined by reversing two of the photon identification criteria, $\Delta E$ and $\Delta R$, which compare the lateral and longitudinal shower shapes in the first two layers of the calorimeter of photons to that of mesons, with the goal of rejecting mesons. Another data control sample is defined by reversing the photon isolation criteria.

The result of the fit to the observed mass distribution is shown in Fig. [1]. The bottom panel of the figures shows the statistical significance of the difference between data and the fit in each bin [50]. The fit quality is quantified using a negative log-likelihood test statistic. The probability of the fit quality to be at least as good as the observed fit is 74%, indicating that the data are consistent with the functional form.
Figure 1: Invariant mass of the $\gamma + \text{jet}$ pair for events passing the final selections. The bin widths are chosen to be twice the mass resolution at the central value of each bin. Overlaid is the fitted background function integrated over each bin (solid line), with three examples of QBH signals (upper) and three examples of $q^*$ signals (lower). For better visibility the QBH signals are only drawn for $m_{\gamma j}$ starting at half the threshold mass, and for the 1.5 TeV signal it ends at twice the threshold mass. The $q^*$ signals are only drawn for $m_{\gamma j}$ within $\pm 25\%$ of the nominal signal mass. The bottom panel shows the statistical significance of the difference between data and background in each bin.
The BumpHunter algorithm [51] is used to search for statistical evidence of a resonance. The algorithm operates on the binned \(m_{\gamma j}\) distribution, comparing the background estimate with the data in mass intervals of varying numbers of adjacent bins across the entire distribution. For each interval in the scan, it computes the significance of any excess found. The algorithm identifies the two-bin interval 785–916 GeV as the single most discrepant interval. The significance of the outcome is evaluated using the ensemble of possible outcomes for the significance of any region in the distribution in the background-only hypothesis, obtained by repeating the analysis on pseudodata drawn from the background function. Before including systematic uncertainties, the probability of finding a fluctuation as large as actually observed (\(p\)-value) is 61\%, including the trials factor, or “look-elsewhere” effect. Thus, the excess is not significant and the data are consistent with a smoothly falling background.

In the absence of any signal, three types of \(\gamma + \text{jet}\) signals are excluded: a generic Gaussian-shaped signal with arbitrary production cross-section, resulting from resonances with varying intrinsic widths convolved with the detector resolution; the QBH model; and the excited quark model. For each signal mass considered in this phase, the fit to the observed mass distribution is repeated with the sum of the four-parameter background function Eq. [1] and a signal template with a normalization determined during the fit. Bayesian limits at the 95\% CL are computed as described in Ref. [28] using a prior probability density that is constant for positive values of the signal production cross-section and zero for unphysical, negative values.

Systematic uncertainties affecting the limits on production of new signals are evaluated. The signal yield is subject to systematic uncertainties on the integrated luminosity (2.8\%), photon isolation efficiency (1.2\%), trigger efficiency (0.5\%), and photon identification efficiencies (1.5\%). The last of these includes extrapolation to high \(p_T\) (0.1\%) and pileup effects (0.1\%). Uncertainties on the jet and photon energy scale contribute 1.0\%–1.5\% and 0.3\%, respectively, through their effects on the shape and yield of the signal distribution. The sizes of the systematic uncertainties are similar for both q* and QBH signals. These systematic uncertainties are treated as marginalized nuisance parameters in the limit calculation. Several other fit functions from Ref. [49] have been tested and a negligible systematic uncertainty was found. To account for the statistical uncertainties on the background fit parameters, the background function is repeatedly fit to pseudodata for which each bin has been drawn from Poisson distributions. The mean of the Poisson distribution for a given bin corresponds to the number of entries actually observed in that bin in the data. The variations in the fit predictions for a given bin, 1\% of the background at 1 TeV to about 20\% of the background at 3 TeV, are taken as indicative of the systematic uncertainty. This bin-by-bin uncertainty is treated in the limit as fully correlated, using a single nuisance parameter that scales the entire background distribution.

Figure 2 shows the model-independent limits on the visible cross-section, defined as the product of the cross section (\(\sigma\)) times branching fraction (BR) times acceptance (A) times efficiency (\(\varepsilon\)), of a potential signal as a function of the mass of each signal template, and includes the systematic uncertainties discussed above. The signal line shape is modeled as a Gaussian distribution, with one of four relative widths: \(\sigma_G/m_G = 5\%, 7\%, 10\%,\) and 15\%, where \(\sigma_G(m_G)\) is the width (mass) of the Gaussian. The limit weakens as the width increases and the peak becomes less distinct. At 1 and 4 TeV the limit is 8 and 0.1 fb for \(\sigma_G/m_G\) of 5\%. The differences in shapes of the limits for different widths are driven by the increased sensitivity to local fluctuations for the narrower signals. Beyond the highest mass event recorded, 2.57 TeV, the limits for different widths converge due to the absence of observed events. At 3 TeV, the new limit improves the earlier ATLAS result [18] in this channel by an order of magnitude.

The limit on the visible cross-section in the QBH model is shown in Fig. 3 as a function of the mass threshold \(M_{th}\). Also shown are the \(\pm 1\sigma\) and \(\pm 2\sigma\) uncertainty bands indicating the underlying distribution of possible limit outcomes in the background-only hypothesis. The solid blue line indicates the prediction for the QBH signal described earlier. The observed (expected) lower limit on the QBH mass threshold is found to be 4.6 (4.6) TeV, at 95\% CL.
The limit on the visible cross-section in the excited quark model as a function of the $q^*$ mass is shown in Fig. 4. Also shown are the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands indicating the underlying distribution of possible limit outcomes in the background-only hypothesis. The solid blue line indicates the prediction from the PYTHIA excited quark implementation. The rise in the expected and observed limits at high $m_{\gamma j}$ is due to the increased fraction of off-shell production of the $q^*$, shifting the signal distribution to lower masses. The observed (expected) lower limit on the excited quark mass is found to be 3.5 (3.4) TeV, at 95% CL. With a much lower branching fraction than the dijet channel but also smaller background, this result improves on the exclusion limit of 3.19 TeV from CMS in the dijet final state with 4.0 fb$^{-1}$ of data [29] and the exclusion limit from ATLAS in the dijet final state below 2.83 TeV with 4.8 fb$^{-1}$ [28].

In conclusion, the $\gamma +$ jet mass distribution measured in 20.3 fb$^{-1}$ of $pp$ collision data collected at $\sqrt{s} = 8$ TeV is well described by the background model, and no evidence for new phenomena is found. Limits at 95% CL using Bayesian statistics are presented for signal processes yielding a Gaussian line shape, non-thermal quantum black holes, and excited quarks. The limits on Gaussian resonances
exclude 4 TeV resonances with visible cross-section near 0.1 fb. Non-thermal black hole and excited quark models in the γ + jet final state are excluded for masses up to 4.6 TeV and 3.5 TeV, respectively. The limits reported here on the production of new resonances in the γ + jet final state are the most stringent limits set to date in this channel.

References


Figure 3: The 95% CL upper limits on $\sigma \times B \times A \times \epsilon$ for QBH decaying to a photon and a jet, as a function of the threshold mass $M_{\text{th}}$, assuming $M_D = M_{\text{th}}$ and $n = 6$. The limit takes into account statistical and systematic uncertainties. Points along the solid black line indicate the mass of the signal where the limit is measured. Also shown are the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands indicating the underlying distribution of possible limit outcomes in the background-only hypothesis. The predicted visible cross section for QBH is shown as the solid blue line.
Figure 4: The 95% CL upper limits on $\sigma \times B \times A \times \epsilon$ for excited quarks, decaying to a photon and a jet, as a function of the signal mass $m_{q^*}$. The limit takes into account statistical and systematic uncertainties. Points along the solid black line indicate the mass of the signal where the limit is measured. Also shown are the $\pm 1\sigma$ and $\pm 2\sigma$ uncertainty bands indicating the underlying distribution of possible limit outcomes in the background-only hypothesis. The blue solid line shows the predicted visible cross section for excited quark production from PYTHIA.
A Event Displays

Figure 5: Event display for the event with the highest $m_{\gamma j}$ among events passing the selection (event 122752008 in run 214758). The photon shower in the calorimeter, consistent with an unconverted photon candidate, is red and the jet is yellow. The mass of the $\gamma$+jet pair is 2.57 TeV. The photon and jet $p_T$ are well balanced, 1.31 and 1.25 TeV, respectively. The $\Delta\phi$ and $\Delta\eta$ between photon and jet are 3.11 and 0.083. The missing transverse energy is 64.4 GeV and the $\Delta\phi$ between jet and missing transverse energy is 2.82.
Figure 6: Event display for the event with the second-highest $m_{\gamma j}$ among events passing the selection (event 77335209 in run 212144). The photon shower in the calorimeter, consistent with an unconverted photon candidate, is red and the jet is yellow. The mass of the $\gamma$+jet pair is 2.50 TeV. The photon and jet $p_T$ are well balanced, 1.12 and 1.36 TeV, respectively. The $\Delta\phi$ and $\Delta\eta$ between photon and jet are 2.83 and 0.75. The missing transverse energy is 59.7 GeV and the $\Delta\phi$ between jet and missing transverse energy is 2.38.