CMS Physics Analysis Summary

Contact: cms-pag-conveners-exotica@cern.ch 2018/03/13

A search for light pair-produced resonances decaying into at least four quarks

The CMS Collaboration

Abstract

Results are presented of a search for light pair-produced particles that each decay into at least four quarks, targeting particles with masses above 100 GeV. The representative signal model is a supersymmetric quark that decays into four quarks through an intermediate Higgsino with a hadronic R-parity-violating coupling. Using 38.2 fb^{-1} of data collected by the CMS experiment at the LHC from 2015 to 2016 with proton-proton collisions at a center-of-mass energy $\sqrt{s} = 13$ TeV, reconstructed particles are clustered into two jets with large radius and high transverse momentum, each consistent with four-parton substructure and with similar mass. A novel, data-driven technique is used to estimate the dominant QCD multijet event background. Background contributions from top-quark decays are also taken into account and additionally serve for the calibration of systematic uncertainties. No statistically significant excess is observed, and the production of supersymmetric quark pairs and supersymmetric gluon pairs with masses below 721 GeV and 1.0 TeV, respectively, are excluded at the 95% confidence level.

Despite the many successes of the LHC, a solution to the electroweak–Planck hierarchy problem remains elusive. One theoretically attractive resolution is the introduction of supersymmetry (SUSY) [1–9], but direct evidence for SUSY has yet to materialize. For SUSY to be a viable natural theory, supersymmetric particles—specifically gluinos, third generation squarks, and Higgsinos—should be accessible at the current LHC energies [10–12]. Long-standing theoretical prejudices have emphasized R-parity-conserving models of SUSY, whereas R-parity-violating (RPV) models still lack constraints in a number of important scenarios [13].

The analysis presented probes a previously unexamined region of the SUSY model space, where pair-produced squarks with masses as light as 100 GeV each decay into four light quarks through hadronic (UDD) RPV couplings [14]. Such decays procede through a two-body decay of the squark into a Higgsino and a quark, and the Higgsino subsequently undergoes an RPV-mediated three-body decay (via an off-shell squark) into quarks. This decay mode is expected to dominate over direct squark decays into two quarks if the Higgsino mass is lower than the squark mass (as expected in natural scenarios) and if the RPV coupling is not large. A related topology with five-quark resonances occurs when the gluino undergoes a three-body decay into two quarks and a Higgsino, and the Higgsino decays into three quarks as described. The experiments at LEP largely constrain squark and Higgsino masses below 100 GeV, while existing LHC searches are only sensitive to multijet scenarios above $\sim$ 600 GeV [15–18]. Light pair-produced squarks and gluinos that decay through RPV couplings into two and three quarks, respectively, have also been directly constrained by a variety of Tevatron and LHC searches [19–22].

We take advantage of the large total cross section for squark pair production by focusing on the fraction of phase space in which the squarks are boosted back-to-back and in which the decay products of each squark reside entirely within a single jet with a large clustering radius. This requirement reduces the combinatorial background and recovers the resonant structure of the squark decay. Furthermore, the masses of the two jets are required to be approximately the same and to evince substructure consistent with at least four non-top quarks. These conditions help to suppress the immense background of events consisting of non-top-quark jets produced by the strong interaction, called quantum chromodynamics (QCD) multijet events, which is estimated with a novel, data-driven technique. Although the background from top quarks is also significant, $\mathrm{t\bar{t}}$ events are harnessed as a standard candle to calibrate selection efficiency, jet-mass scale, and jet-mass resolution.

The motivation for this search comes from natural SUSY, but the analysis aims more broadly for new phenomena in general. This search makes no assumptions about the helicity or the intermediate decay topology, nor do we take advantage of the possibility of heavy flavor quarks in the final state so as to be more inclusive and tackle the experimentally more challenging fully light quark mode. In fact, we do not even require the final state particles to be quarks; they only must result in particles that are detected and reconstructed by the standard particle flow (PF) algorithms [23]. Constraints on pair-produced gluino decays to five light quarks are also presented, in part to demonstrate the power and inclusive nature of the search.

The CMS detector consists of a silicon tracker, a lead-tungstate electromagnetic calorimeter, an interleaved brass and plastic scintillator hadron calorimeter, and a gas-ionization and resistive plate muon detector. A superconducting solenoid provides a uniform magnetic field within the detector. Observations from each of these detector components are combined and reconstructed by a PF algorithm into particle objects with relativistic 4-momenta; these objects are the basic elements that we cluster to form the jets used in this analysis. A more detailed description of the CMS detector can be found in Ref. [24].

We analyze data collected by the CMS detector in 2015 and 2016 using proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$ and 25 ns bunch crossing, consisting of an integrated luminosity of $38.2\,\text{fb}^{-1}$. Analyzed events were selected by the high-level trigger that imposes requirements on the event $H_\text{T} \equiv \sum_i p_{\text{T},i}$, where the index i iterates over the jets in the event with $p_\text{T} > 150\,\text{GeV}$ and $|\eta| < 2.5$. The jets used in the trigger are "AK8" jets, that is they are clustered with the anti-k_T algorithm [25] and have a jet radius of $R = 0.8$.

Signal events, consisting of squark decays to four light quarks, are simulated using MADGRAPH 5 [26] for generation with the NNPDF2.3 parton distribution function [27] and PYTHIA 8 [28] to simulate the subsequent parton showering and decays. Events for 11 squark mass points between 100 and 800 GeV are produced. In each case, the Higgsino mass is set to 75% of the squark mass in order to evenly distribute the squark energy among the final state quarks. The production cross section assumed in this analysis corresponds to top squarks with decoupled gluinos [29]. Simulated gluino signal events are produced for 14 gluino mass points between 100 GeV and 1 TeV using the same generation procedure as described for the squark events. In this case, the squark mass is decoupled from the gluino mass, and the Higgsino mass is 60% of the gluino mass.

We use simulated standard model processes to characterize background events. QCD multijet events are simulated using MADGRAPH 5 to generate up to four partons, which are showered and hadronized with PYTHIA 8. Background from $\text{t}\bar{\text{t}}$ production is generated with POWHEG v2 [30], and the p_T distribution of the top quarks is re-weighted to measurement [31]. Other backgrounds are also considered, such as hadronic $\text{W} + \text{jets}$ and $\text{Z} + \text{jets}$, but their contribution can be neglected for this analysis. The simulated events incorporate additional proton-proton interactions within the same or a nearby bunch crossing (pileup) and are weighted to match the measured distribution of the number of interactions per bunch crossing.

In order to capture as many of the final state constituents of the squark decay products as possible in a single jet, we cluster the PF objects of each event into jets with the Cambridge-Aachen (CA) algorithm [25] and a jet radius of $R = 1.2$, using the FASTJET software package [32]. We call such objects "CA12" jets. In order to reduce the effect of pileup interactions, charged hadrons associated with the vertices other than the primary vertex are not considered in this clustering. The jet energies are corrected to compensate for the non-linear response of the CMS calorimeters with calibrations designed for AK8 jets. For simulated events, this correction includes an additional scale factor to account for the residual difference in detector response between simulation and data. Jet masses are also corrected for this residual difference. Later, we measure and apply an uncertainty that includes any miscorrection derived from the differences between AK8 jets and the CA12 jets we use. Finally, to reject misreconstructed jets or jets derived from calorimeter noise, we impose restrictions on the neutral and charged jet constituents by applying a loose PF Jet ID as described in Ref. [33].

We analyze the substructure of the CA12 jets using the N-subjettiness variables, denoted τ_N [34]. These quantities are determined by reclustering the constituents of the jet with the k_T algorithm [25] until N objects (subjets) remain and then calculating

$$\tau_N \equiv \frac{1}{d_0} \sum_i p_{\text{T},i} \min\{\Delta R_{1,i}, \ldots, \Delta R_{N,i}\}, \tag{1}$$

where i iterates over the jet constituents, $d_0 \equiv \sum_i p_{\text{T},i} R$ with R representing the radius of the jet, and $\Delta R_{j,i}$ the distance in η-ϕ space between the j^th subjet and the i^th jet constituent. The relative N-subjettiness $\tau_{kl} \equiv \tau_k / \tau_l$, for some integers k and l, is a useful characterization of the substructure of a jet. In particular, τ_{42} is used to discriminate against QCD multijet events

and τ_{43} to discriminate against $\mathrm{t\bar{t}}$ events. The term "fatjet" is used to refer to a CA12 jet with substructure.

Finally, we apply a pruning algorithm that reclusters the original constituents of a jet using a CA algorithm modified to ignore any objects with small relative p_{T} ($z_{\mathrm{cut}} = 0.1$) and large kinematically-weighted displacement ($R_{\mathrm{cut}} = 0.5$) [35]. This refinement serves to suppress perturbative radiation characteristic of QCD multijet events as well as to reduce the contribution to the masses of jets from detector noise, pileup, and the underlying event. In this analysis, the mass of a jet m always refers to the mass after pruning.

We consider events with $H_{\mathrm{T}} > 900\,\mathrm{GeV}$ in order to guarantee that we analyze a kinematic region with a fully-efficient trigger. In addition, we require that the two p_{T}-leading CA12 jets of the events must each satisfy the following three conditions: $p_{\mathrm{T}} > 400\,\mathrm{GeV}$ to ensure that the jets are boosted and therefore more likely to contain all signal products; $|\eta| < 2$ to require the jets are dominated by objects in the central detector region; and $\tau_{21} < 0.75$ to select CA12 jets with potential for substructure.

From these selected events, we define two analysis regions. Each analysis region is formed by events in which the two p_{T}-leading CA12 jets each satisfy a certain requirement on τ_{42} and τ_{43}, called the region's "fatjet tag," and also pass relational criteria of $A_{\mathrm{m}} \equiv |m_1 - m_2| / (m_1 + m_2) < 0.1$ and $\Delta\eta \equiv |\eta_1 - \eta_2| < 1.0$, defining a "fatjet pair." The signal region fatjet tag requires $\tau_{42} < 0.50$ and $\tau_{43} < 0.80$, choices that were made to maximize the statistical significance of a potential signal. The control region fatjet tag requires $\tau_{42} < 0.55$, $\tau_{43} < 0.90$, and $\tau_{42} > 0.50$ or $\tau_{43} > 0.80$. In addition to this control region, we define a b-tagged control region by imposing a loose b-tag [36] requirement in the control region fatjet tag in order to measure a $\mathrm{t\bar{t}}$-enriched sample. The final discriminating variable of this analysis is the average pruned mass of the leading two fatjets in an event: $\overline{m} \equiv (m_1 + m_2)/2$.

The fraction of squark events that are selected into the signal region ranges from $< 0.001\%$ for the 100 GeV squark mass point to 1% for the 800 GeV mass point. For very light squarks, the small selection efficiency is compensated by a large production cross section (1520 pb for 100 GeV top squarks). Similar numerical values for the selection efficiency are found for gluinos, although the $\overline{m}$ distribution for gluinos is broader than for squarks because of the extra quark in the decay.

We extract a particular signal's contribution to the observed events by performing a maximum-likelihood fit of the signal and two background $\overline{m}$ probability density functions (PDFs) to the $\overline{m}$ distribution of the data. The signal and $\mathrm{t\bar{t}}$ background $\overline{m}$ PDFs are taken from simulation; the QCD multijet background $\overline{m}$ PDF is constructed using a data-driven method detailed below. Each $\overline{m}$ PDF is assigned three fit parameters: a normalization factor, a shift of the median of the PDF from the nominal value, and a width factor (stretch) about the median of the PDF. These fit parameters allow us to account for shape differences between simulation and data.

This procedure is applied to the control regions of the data, and we use the post-fit values of the $\overline{m}$ PDF parameters as in-situ measurements of the background systematic uncertainties. Figure 1 shows the post-fit background $\overline{m}$ PDFs and associated fit parameters assuming no signal contribution. Signal contamination in the control region is generally small, and we repeat the same procedure including in turn the $\overline{m}$ PDF for each signal mass point, concluding via an F-test that the presented null hypothesis is the most reliable fit. The $\mathrm{t\bar{t}}$ $\overline{m}$ PDF fit parameters have physical interpretations: the normalization corresponds to the data-to-simulation fatjet tagging scale factor, the shift to the tagged-fatjet mass scale, and the stretch to the tagged-fatjet mass resolution. As presented in the legends of Fig. 1, the $\mathrm{t\bar{t}}$ shifts are generally consistent with zero

and the $t\bar{t}$ stretches close to unity, demonstrating that the $t\bar{t}$ $\overline{m}$ PDF is predicted well by simulation in both the inclusive and b-tagged regions. The QCD multijet $\overline{m}$ PDF needs a small shift downward in the inclusive control region. We study a variety of alternative choices of control regions, which similarly demonstrate good closure. The background estimation method reproduces the data in the control region, which has greater statistical precision than the signal region, lending confidence that it provides reliable results in the signal region as well.

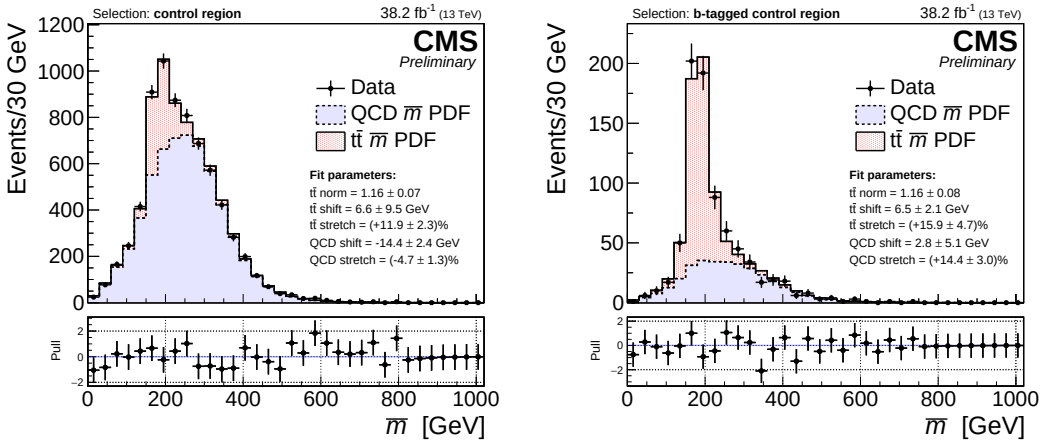

Figure 1: The background estimation results in the control region (left) and b-tagged control region (right). The relevant fit parameters are displayed on each plot.

We derive the QCD multijet background $\overline{m}$ PDF from a high-statistics collection of data events in order to avoid the significant mismodeling of simulated QCD multijet events. Events are identified in which at least the p_{T}-leading jet satisfies the fatjet tag; the number of such events is more than an order of magnitude greater than those required to have a fatjet pair. We form a collection of fatjets by taking the p_{T}-leading fatjet from each of these events. From this collection of p_{T}-leading fatjets from different events, we form all possible fatjet pairs. The $\overline{m}$ distribution of these artificial fatjet pairs is an $\overline{m}$ PDF for QCD multijet events. This construction treats the m distribution of p_{T}-leading fatjets as an m PDF $P(m)$ from which two fatjet masses are sampled to form an $\overline{m}$ PDF $P_{\mathrm{avg}}(\overline{m})$:

$$P_{\mathrm{avg}}(\overline{m}) = \int_0^{2\overline{m}} P(x) \cdot P(2\overline{m} - x) \cdot \theta\left(0.1 - \left|\frac{x - \overline{m}}{\overline{m}}\right|\right) dx, \tag{2}$$

where θ is the Heaviside step function that imposes the A_{m} fatjet pair requirement; while not represented in Eq. (2), we impose the $\Delta\eta$ fatjet pair requirement by hand when we implement this construction.

Fatjets from QCD multijet events do not contain the products of partons with well-defined masses, and so the $P(m)$ derived from them has a significant p_{T} dependence that must be taken into account. We accommodate this effect by constructing a different $P_{\mathrm{avg}}(\overline{m})$ for consecutive H_{T} windows, which we then combine with weights according to the observed H_{T} distribution in data. Figure 2 (left) shows $P(m)$ and the derived $P_{\mathrm{avg}}(\overline{m})$ with and without this H_{T} re-weighting, demonstrating the significant effect of the re-weighting on the high-$\overline{m}$ tail of the distribution. Corrections for the contamination of signal or other backgrounds to the QCD multijet $\overline{m}$ PDF are unnecessary given the dominance of QCD multijet events over all other processes in the events used to construct $P(m)$.

We demonstrate the basic accuracy of this QCD multijet $\overline{m}$ PDF construction with a closure test. We construct QCD multijet $\overline{m}$ PDFs from simulated QCD multijet events in the signal

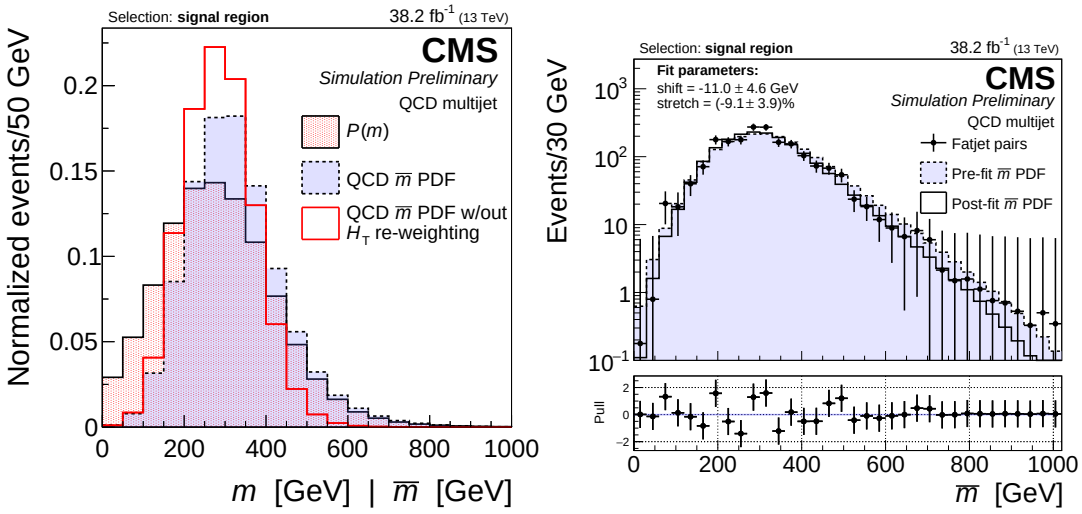

Figure 2: Left: the m distribution of the tagged p_{T}-leading fatjets $P(m)$ in the signal region and the $\overline{m}$ PDFs derived from these fatjets with and without H_{T} re-weighting. Right: a comparison between the QCD multijet $\overline{m}$ PDF and the tagged fatjet pair selection in simulated QCD multijet events that pass the signal selection.

and control regions and compare them to the $\overline{m}$ distributions of the regions' measured tagged fatjet pairs; the result for the signal region is shown in Fig. 2 (right). The signal region PDF needs to be shifted downward by approximately 11 GeV and compressed approximately 9% to best match the closure prediction, and the $\overline{m}$ PDF for the control region requires a similar adjustment that is consistent with the results in the data control regions. Similar tests with alternative choices for the N-subjettiness variables and different simulation procedures for the QCD multijet events also demonstrated good closure.

Using the data in the signal region, we apply a Bayesian construction with a flat prior on the signal strength [37] to compute 95% confidence level (CL) upper limits on squark (gluino) pair production and decay to four (five) quarks. Markov chain Monte Carlo techniques are used to integrate out the nuisance parameters. We analyze the data binned in 30 GeV steps, as shown in Fig. 3.

We assign nuisance parameters corresponding to the shift, stretch, and normalization of the QCD multijet, $\mathrm{t\bar{t}}$, and signal $\overline{m}$ PDFs. The nuisance parameters have Gaussian PDFs. The QCD multijet $\overline{m}$ PDF shape nuisance parameter uncertainties are taken from the control region, but no initial shift or stretch are applied, since these parameters in the signal region might be different from those in the control region. We do not have a robust prediction for the total rate of QCD multijet events in the signal region, so the QCD multijet $\overline{m}$ PDF normalization is left to float arbitrarily; this procedure is reasonable because the QCD multijet shape is much broader than the signals', allowing the data in the signal region to be used to directly constrain this background. The $\mathrm{t\bar{t}}$ rate and shape nuisance parameter uncertainties are determined by their measured values in the control regions; we initialize them with no shift or stretch, as that is consistent with measurement. Again interpreting the $\mathrm{t\bar{t}}$ shift and stretch as jet-mass scale and resolution uncertainties, respectively, we apply the same uncertainties to the signal shape by constraining them to match the $\mathrm{t\bar{t}}$ nuisance parameters. For the signal acceptance nuisance parameter, we assign the same uncertainty value that we measured for $\mathrm{t\bar{t}}$ events, since the signal and $\mathrm{t\bar{t}}$ production and decay topologies are similar. The uncertainty in the acceptance combines uncertainties in the integrated luminosity and the fatjet tagging scale factor. The central values and standard deviations of all of the nuisance parameters are shown in Table 1.

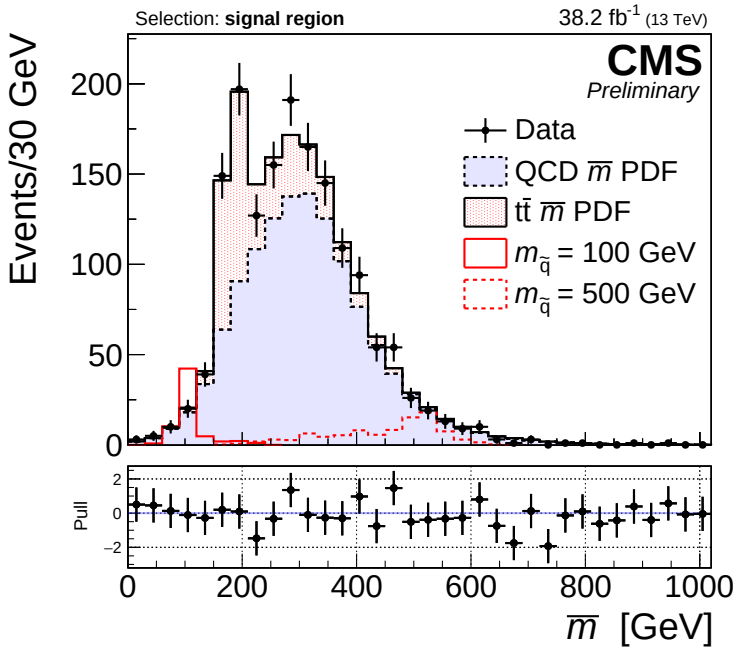

Figure 3: Distributions in $\overline{m}$ and the predicted background contributions. Shown also are the signal expected from a squark with masses of 100 and 500 GeV.

Table 1: The nuisance parameters corresponding to each rate and shape parameter of the background and signal distributions before and after the statistical fit. Except for the QCD multijet $\overline{m}$ PDF normalization, which is floating (and whose value is simply the event yield with statistical uncertainty), each nuisance parameter has a Gaussian PDF and is reported as its mean $\pm$ its standard deviation.

Parameter	**Pre-fit value**	**Post-fit value**
QCD multijet $\overline{m}$ PDF normalization	floating	1222 ± 35 events
QCD multijet $\overline{m}$ PDF shift	0 ± 17 GeV	-8 ± 4 GeV
QCD multijet $\overline{m}$ PDF stretch	(0 ± 18) %	(-1 ± 3) %
$t\bar{t}$ $\overline{m}$ PDF normalization	1.00 ± 0.24	1.08 ± 0.14
$t\bar{t}$ $\overline{m}$ PDF and signal shift	0 ± 16 GeV	-10 ± 6 GeV
$t\bar{t}$ $\overline{m}$ PDF and signal stretch	(0 ± 20) %	(15 ± 9) %
Signal $\overline{m}$ PDF normalization	1.00 ± 0.24	-

The resulting limits on the production cross section and the predicted background components in the signal region are shown in Fig. 4. Assuming the top squark production cross section, squark masses between 100 and 721 GeV are excluded. Gluinos decaying to 5 quarks with masses between 0.10 and 1.0 TeV are also excluded. The post-fit total background estimation agrees with the data. The posterior distributions of the nuisance parameters confirm that the background component predictions are not significantly different from the estimates.

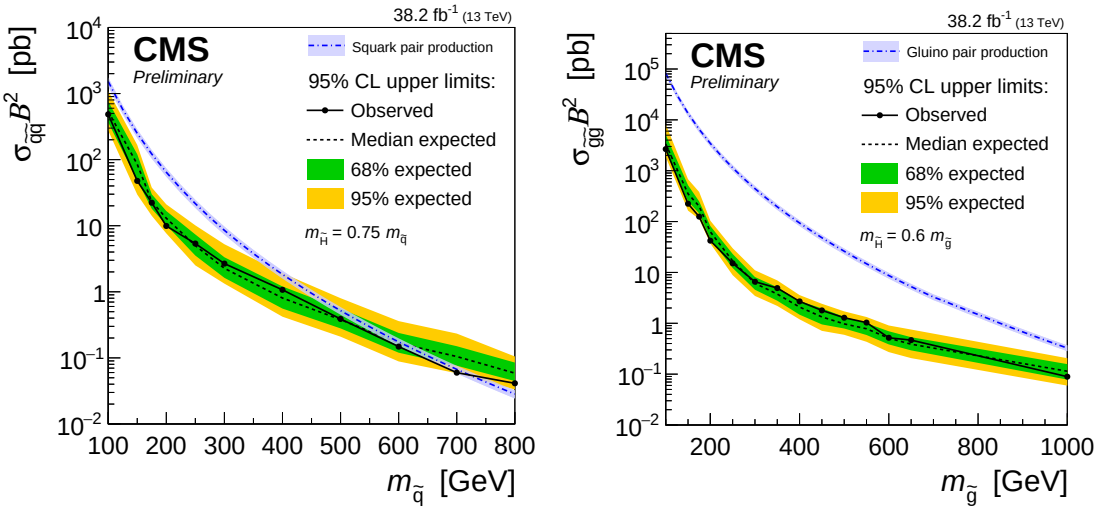

Figure 4: The expected and observed limits on the squark (left) and gluino (right) signal production.

In summary, a search was conducted for light pair-produced resonances that decay into at least four quarks. No statistically significant excess over standard model expectations was observed. The data impose limits on RPV SUSY [13] pair production, excluding squark masses between 100 and 721 GeV and gluino masses between 0.10 and 1.0 TeV. This is the first time that constraints have been placed on pair-produced particles decaying into four or five quarks in this mass range, covering a significant gap in the parameter space of RPV SUSY. This analysis is sufficiently general that other models of strongly produced particles decaying into four or more detector-visible objects are also likely to be constrained.

References

[1] R. Barbieri, S. Ferrara, and C. A. Savoy, "Gauge models with spontaneously broken local supersymmetry", *Phys. Lett. B* **119** (1982) 343, doi:10.1016/0370-2693(82)90685-2.

[2] J. Wess and B. Zumino, "Supergauge transformations in four dimensions", *Nucl. Phys. B* **70** (1974) 39, doi:10.1016/0550-3213(74)90355-1.

[3] Y. A. Gol'fand and E. P. Likhtman, "Extension of the algebra of poincaré group generators and violation of p invariance", *JETP Lett.* **13** (1971) 323, doi:10.1142/9789814542340_0001.

[4] D. V. Volkov and V. P. Akulov, "Possible universal neutrino interaction", *JETP Lett.* **16** (1972) 438.

[5] A. H. Chamseddine, R. L. Arnowitt, and P. Nath, "Locally supersymmetric grand unification", *PRL* **49** (1982) 970, doi:10.1103/PhysRevLett.49.970.

[6] G. L. Kane, C. F. Kolda, L. Roszkowski, and J. D. Wells, "Study of constrained minimal supersymmetry", *Phys. Rev. D* **49** (1994) 6173, doi:10.1103/PhysRevD.49.6173, arXiv:hep-ph/9312272.

[7] P. Fayet, "Supergauge invariant extension of the higgs mechanism and a model for the electron and its neutrino", *Nucl. Phys. B* **90** (1975) 104, doi:10.1016/0550-3213(75)90636-7.

[8] L. J. Hall, J. D. Lykken, and S. Weinberg, "Supergravity as the messenger of supersymmetry breaking", *Phys. Rev. D* **27** (1983) 2359, doi:10.1103/PhysRevD.27.2359.

[9] P. Ramond, "Dual theory for free fermions", *Phys. Rev. D* **3** (1971) 2415, doi:10.1103/PhysRevD.3.2415.

[10] R. Kitano and Y. Nomura, "Supersymmetry, naturalness, and signatures at the LHC", *Phys. Rev.* **D73** (2006) 095004, doi:10.1103/PhysRevD.73.095004, arXiv:hep-ph/0602096.

[11] C. Brust, A. Katz, S. Lawrence, and R. Sundrum, "SUSY, the Third Generation and the LHC", *JHEP* **03** (2012) 103, doi:10.1007/JHEP03(2012)103, arXiv:1110.6670.

[12] M. Papucci, J. T. Ruderman, and A. Weiler, "Natural susy endures", *JHEP* **09** (2012) 035, doi:10.1007/JHEP09(2012)035.

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

[14] R. Barbier et al., "R-parity-violating supersymmetry", *Physics Reports* **420** (2005) 1, doi:https://doi.org/10.1016/j.physrep.2005.08.006.

[15] CMS Collaboration, "Search for new phenomena in events with high jet multiplicity and low missing transverse momentum in proton-proton collisions at $\sqrt{s} = 8$ TeV", *Phys. Lett. B* **770** (2017) 257, doi:10.1016/j.physletb.2017.01.073, arXiv:1608.01224.

[16] CMS Collaboration, "Searches for *r*-parity-violating supersymmetry in pp collisions at $\sqrt{s} = 8$ tev in final states with 0–4 leptons", *Phys. Rev. D* **94** (2016) 112009, doi:10.1103/PhysRevD.94.112009, arXiv:1606.08076.

[17] ATLAS Collaboration, "Search for new phenomena in final states with large jet multiplicities and missing transverse momentum with ATLAS using $\sqrt{s} = 13$ TeV proton-proton collisions", *Phys. Lett. B* **757** (2016) 334, doi:10.1016/j.physletb.2016.04.005, arXiv:1602.06194.

[18] ATLAS Collaboration, "Search for massive supersymmetric particles decaying to many jets using the ATLAS detector in pp collisions at $\sqrt{s} = 8$ TeV", *Phys. Rev. D* **91** (2015) 112016, doi:10.1103/PhysRevD.91.112016, arXiv:1502.05686.

[19] CMS Collaboration, "Search for low mass vector resonances decaying into quark-antiquark pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JHEP* **01** (2018) 097, doi:10.1007/JHEP01(2018)097, arXiv:1710.00159.

[20] ATLAS Collaboration, "A search for pair-produced resonances in four-jet final states at $\sqrt{s}$=13 TeV with the ATLAS detector", arXiv:1710.07171.

[21] 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.

[22] 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* **6** (2016) 67, doi:10.1007/JHEP06(2016)067, arXiv:1601.07453.

[23] 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.

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

[25] G. P. Salam, "Towards jetography", *The European Physical Journal C* **67** (2010) 637, doi:10.1140/epjc/s10052-010-1314-6, arXiv:hep-ph/10906.1833.

[26] J. Alwall and et al., "The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations", *JHEP* **7** (2014) 79, doi:10.1007/JHEP07(2014)079, arXiv:1405.0301.

[27] N. Collaboration, "Parton distributions with LHC data", *Nucl. Phys. B* **867** (2013) 244, doi:10.1016/j.nuclphysb.2012.10.003, arXiv:1207.1303.

[28] T. Sjöstrand, S. Mrenna, and P. Skands, "A brief introduction to PYTHIA 8.1", *Computer Physics Communications* **178** (2008) 852, doi:10.1016/j.cpc.2008.01.036, arXiv:0710.3820.

[29] 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. D* **74** (2014) 3174, doi:10.1140/epjc/s10052-014-3174-y.

[30] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with parton shower simulations: the POWHEG method", *JHEP* **11** (2007) 70, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[31] CMS Collaboration Collaboration, "Measurement of differential cross sections for top quark pair production using the lepton + jets final state in proton-proton collisions at 13 TeV", *Phys. Rev. D* **95** (May, 2017) 092001, doi:10.1103/PhysRevD.95.092001.

[32] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *The European Physical Journal C* **72** (2012), no. 3, 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[33] CMS Collaboration, "Jet algorithms performance in 13 TeV data", CMS physics analysis summary CMS-PAS-JME-16-003, 2016.

[34] J. Thaler and K. Van Tilburg, "Maximizing boosted top identification by minimizing N-subjettiness", *JHEP* **2** (2012) 93, doi:10.1007/JHEP02(2012)093, arXiv:1108.2701.

[35] 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.

[36] CMS Collaboration, "Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV", (2017). arXiv:1712.07158. Submitted to *JINST*.

[37] G. Cowan, "Statistics", Ch. 39 in Particle Data Group, C. Patrignani et al., "Review of particle physics", *Chin. Phys. C* **40** (2016) 100001, doi:10.1088/1674-1137/40/10/100001.