CMS Physics Analysis Summary

Contact: cms-pag-conveners-exotica@cern.ch 2023/08/01

Search for pair-produced multijet resonances using data scouting

The CMS Collaboration

Abstract

Searches for pair-produced multijet signatures are presented. The analyses use a data sample corresponding to an integrated luminosity of 128 fb^{-1} of proton-proton collisions at $\sqrt{s} = 13$ TeV to measure the average mass distributions of pairs of tri-jets, both merged and resolved, and pairs of merged dijets. The data were collected using data scouting, which saves events with trigger-level reconstruction, enabling the recording of events corresponding to relatively low multi-jet masses. The results are interpreted within an R-parity violating supersymmetric framework in which the pair productions of higgsinos, gluinos, and top squarks are used as benchmark models. Decays to three resolved jets, as well as to highly Lorentz-boosted pairs of two or three quarks yielding single large radius jets, are considered. The search excludes RPV gluinos with mass less than 1.7 TeV, and extends prior exclusions of RPV squarks and gluinos to the low mass region between 70 and 200 GeV. The electroweak production of R-parity violating supersymmetry is probed for the first time in fully hadronic final states, yielding the first exclusions on prompt hadronically decaying mass-degenerate higgsinos.

Despite the discovery of the Higgs boson [1–3], the last particle predicted by the standard model (SM), many phenomena in particle physics remain unexplained [4–7]. Several theories of physics beyond the standard model (BSM) that may resolve these questions posit the existence of new resonances, which would be produced in high-energy proton-proton (pp) collisions and subsequently decay to multi-quark final states [8–10]. This note presents searches for the pair-production of several R-parity violating (RPV) supersymmetric (SUSY) particles which decay to pairs or triplets of quarks (q): stop quarks ($\tilde{t} \to qq$), higgsinos ($\tilde{h} \to qqq$), and gluinos ($\tilde{g} \to qqq$) [7]. Feynman diagrams for each signature are shown in Fig.1. The searches are based on data from pp collisions at the CERN LHC, collected with the CMS detector in 2016–2018, corresponding to an integrated luminosity of 128 fb^{-1}.

The high center-of-mass energy ($\sqrt{s} = 13$ TeV) of these LHC collisions means that low mass resonances are produced with significant Lorentz boosts, such that the hadronization products of the individual final state quarks may overlap in the detector. The resulting signatures are large radius jets with identifiable substructure. The searches presented in this note consider three signatures: pairs of well-resolved triplets of jets (henceforth called *resolved trijets*), pairs of large jets with substructure consistent with three underlying quarks (henceforth called *boosted trijets*), and pairs of large jets with two-pronged substructure (henceforth called *boosted dijets*). The first two signatures are sensitive to both $\tilde{h}$ and $\tilde{g}$ production, while the third is sensitive to $\tilde{t}$ pair production. A separate, dedicated search for resolved paired dijet signatures has previously been presented by the CMS Collaboration [11].

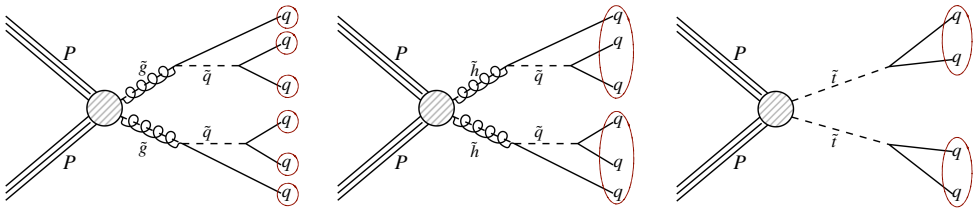

Figure 1: Benchmark models from RPV SUSY for resolved trijets (left), merged trijets (middle) and merged dijets (right). The red circles group the final state quarks according to the expected jet clustering of their hadronization products.

While the three signatures are analyzed independently, they use similar techniques and share a single method for estimating the dominant background from quantum chromodynamical multijet processes (QCD). In particular, the analyses use the CMS *scouting dataset* [12], in which only the event data reconstructed by the high-level trigger (described below) is saved. The resulting reduced event size allows for higher event rates at lower event energy thresholds, thus increasing the sensitivity of the searches for lower resonance masses. Jet substructure variables developed explicitly for data scouting are used to reduce the QCD background for the boosted signatures. Additional resonant backgrounds are estimated from simulations: the fully hadronic decay of the SM production of top-antitop pairs ($t\bar{t}$) is an irreducible background to the paired trijets search. Similarly, events containing hadronically decaying Z or W bosons are a background to the paired dijet search.

Searches for pairs of BSM particles each decaying to three resolved jets have been performed by the CDF [13], ATLAS [14, 15], and CMS [16–18] collaborations. Searches for pairs of particles which decay to two quarks have also been reported by the ATLAS [19–21], and CMS [22, 23] collaborations. However, no limits have previously been set on the production of mass degenerate RPV higgsinos at colliders, or on the production of RPV gluinos with masses below 200 GeV at $\sqrt{s} = 13$ TeV. The search presented in this analysis is the first to probe those regimes.

The CMS apparatus [24] is a multipurpose, nearly hermetic detector, designed to trigger on [25, 26] and identify electrons, muons, photons, and hadrons [27–29]. A global "particle-flow" (PF) algorithm [30] aims to reconstruct all individual particles in an event, combining information provided by the all-silicon inner tracker and by the crystal electromagnetic and brass-scintillator hadron calorimeters, operating inside a 3.8 T superconducting solenoid, with data from the gas-ionization muon detectors embedded in the flux-return yoke outside the solenoid. The reconstructed particles are used to build τ leptons, jets, and missing transverse momentum [31–33].

While the LHC provides collisions every 25 ns, the CMS detector and data acquisition systems do not have the bandwidth to record every event. Instead, events of interest are selected using a two-tiered trigger system. The first level (L1), composed of custom hardware processors, uses information from the calorimeters and muon detectors to select events at a rate of around 100 kHz within a fixed latency of about 4 μs [25]. The second level, known as the high-level trigger (HLT), consists of a farm of processors running a version of the PF algorithm optimized for fast processing, and reduces the event rate to around 1 kHz before data storage [26]. If the partial reconstruction by the HLT indicates that the event has passed some specific requirements, such as the total hadronic transverse momentum in the event (H_T) being above some threshold, the event is fully reconstructed and stored for offline analysis.Due to significant QCD production rates, the threshold for trigger selection based purely on H_T was set to 800 GeV in 2016 and 1.05 TeV in 2017-2018. For the mass ranges considered in these analyses, the multijet signatures do not produce jets with sufficient transverse momenta (p_T) for the events to meet these thresholds. To address this issue CMS implemented data scouting, where some of the partially reconstructed events analyzed by the HLT are saved for further analysis. In particular, these analyses use a dataset of scouting events with $H_T > 410$ GeV [12, 34].

The pair production of RPV SUSY particles is used as the benchmark models for these analyses, where gluinos and higgsinos decay to the three jet final state and top squarks to the dijet final state. Signal events are generated using MADGRAPH5_aMC@NLO 2.6.5 [35] and PYTHIA 8.212 [36, 37] with CP5 as the underlying event tune. The masses of virtual squarks in gluino decays are set to sufficiently high values that the gluinos decay without an internal resonance. The natural width of the RPV gluinos and RPV top squarks is assumed to be much smaller than the experimental resolution. QCD, $\mathrm{t\bar{t}}$ and W/Z boson events are simulated with MADGRAPH5_aMC@NLO 2.6.1. QCD events are simulated to leading order with up to four partons in the matrix element calculation. The QCD simulation is used only to optimize the selection variables. The QCD background for each search is estimated using Gaussian processes (GP) regressions as described below.

All the PF candidates in an event are clustered into jets using FASTJET [38, 39] with the anti-k_T algorithm [40] and distance parameters of 0.4 and 0.8 (AK4 and AK8 jets) for the resolved and boosted categories respectively. The jet energies are corrected to compensate for the combined response functions of the CMS calorimeters and the presence of neutral hadrons from additional pp interactions in the same or adjacent bunch crossings (pileup) [32, 41]. Jets in data are further corrected to account for the residual difference between simulation and data. The corrections are derived from data that has undergone full offline reconstruction, rather than from the scouting dataset. The "tight jet identification criteria" [42] are applied to remove poorly reconstructed jets and jets consisting mainly of calorimeter noise.

The trimming algorithm described in Ref. [43] is applied to the AK8 jets. The algorithm reclusters the constituents of the jet with a radius of 0.2 using the k_T algorithm. Clusters with p_T smaller than 3% of the original jet p_T are removed. This trimming procedure reduces contribu-

tions from perturbative QCD radiation, pileup, and particles from the underlying event to the jet mass m. The jet mass after trimming is further corrected so that the top quark peak observed in data matches that from $\mathrm{t\bar{t}}$ simulations.

In the boosted analyses, jet substructure [44] is used to separate signal jets (trimmed AK8 jets containing multiple hadronized quarks from a boosted resonance decay) from the QCD background. In particular, the variables τ_{32} [44] and N_2^1 [45] are used to identify boosted resonances decaying to 3 and 2 quarks, respectively. Jets from three overlapping quarks tend to have low values of τ_{32}. Similarly, jets clustered from two quarks tend to have low values of N_2^1. Both variables tend to have higher values when evaluated on jets from QCD. These substructure variables are correlated with the kinematic distributions of the jets, such that imposing a selection based on them can preferentially enhance certain regions in the mass distributions of backgrounds, even in the absence of signal. To avoid such correlations, the designed decorrelated tagger (DDT) [46] procedure is used to remove the dependence of the variables on the p_{T} and $\rho = \ln(m^2/p_T^2)$ of the jets. The background efficiency of the DDT selection is fixed at 5% for each (p_{T}, ρ) bin as determined by the QCD simulation. The resulting variables are called $N^1_{2,\mathrm{DDT}}$ and $\tau_{32,\mathrm{DDT}}$. Both the leading and sub-leading jets are required to have $\tau_{32,\mathrm{DDT}} < 0$ or $N^1_{2,\mathrm{DDT}} < 0$, for selecting pair-produced resonances decaying to three or two quarks, respectively. For signals with resonance masses between 70 and 200 GeV, the acceptance of these selections varies between 2.5% and 0.71% for RPV gluinos and mass degenerate higgsinos, and between 2.9% and 1.1% for RPV top squarks.

For the resolved analysis, a quark-gluon discriminator (QGD) based on a neural network (NN) is used to distinguish between the narrow quark-initiated AK4 jets of the trijet resonances and gluons produced by QCD, which yield wider AK4 jets. The NN architecture is based on the "Particle-Flow network" as described in Ref. [47]. The NN inputs are the normalized four-momenta information (p_T, η, ϕ,m) along with the particle type of each jet constituent. Each constituent's p_{T} is divided by the p_{T} of the entire jet. The NN is trained on an equal number of quark and gluon jets, sampled from the QCD simulation for a total of 900,000 jets. The NN outputs a QGD score between 0 and 1, with 1 corresponding to more quark-like jets. The analysis uses loose, medium, and tight QGD selections corresponding to quark acceptance and gluon rejection rates of 98% and 31%, 83% and 70%, and 61% and 87% respectively

The analyses suffer from immense background from QCD multijet processes. In the case of the merged signatures (boosted trijets and boosted paired dijets) substructure in the jets from resonance decay is used to suppress backgrounds. To improve the sensitivity in the resolved trijet search, the jet-ensemble technique [18] along with the QGD and internal kinematics (Dalitz variables) are used. These techniques allow us to suppress the huge background contribution from the QCD multijets. The remaining QCD background in the analyses is estimated using GP regression [48], a novel technique for fitting data without assuming a prior function hypothesis. Backgrounds from $\mathrm{t\bar{t}}$ are estimated using Monte Carlo simulation.

For the merged resonance searches, the leading two AK8 jets in the events are considered. For a fully efficient trigger, these jets are required to have $p_{\mathrm{T}} > 300\,\mathrm{GeV}$ and $|\eta| < 2.4$. The mass asymmetry between the leading jet (j_1) and sub-leading jet (j_2), $A_m \equiv |j_1^m - j_2^m|/(j_1^m + j_2^m)$ is required to be less than 0.15.

In the resolved trijets search, AK4 jets that pass the thresholds of $p_{\mathrm{T}} > 30\,\mathrm{GeV}$ and $|\eta| < 2.4$ are considered in each event. The events are required to have $H_{\mathrm{T}} > 600$ GeV to ensure a fully

efficient trigger. In the passing events, jets that fail the loose quark threshold of the QGD are rejected, and every event is required to have at least six surviving jets. In events with more than six jets, we find that restricting the set of considered triplets to the ones involving only the six quark jets of the highest p_T maximizes the sensitivity to the signal, while keeping the combinatorial background manageable. These six jets are grouped into combinations of three, making twenty triplets, and triplets are grouped into ten triplet pairs. The selections are placed at event, pair, and triplet levels to fully utilize the various kinematic features, and select decays with desired internal dynamics using the Dalitz variables described in [18].

After the pre-selection in the resolved trijet search, event-level selection on the six jets is imposed. This selection includes a sixth-jet p_T requirement and an event-level Dalitz selection using the variable $D^2_{[(6,3)+(3,2)]}$, which measures the geometric spread in the six jet topology. Next, at the triplet pair level, pairs based on the mass-asymmetry between the constituent triplet objects are selected. Delta selection, defined as $\Delta = p_{T_{jjj}} - m_{jjj}$, is applied on the triplets from the surviving pairs. Further, a triplet-level requirement on the Dalitz variable $D^2_{[3,2]}$ is imposed to enforce symmetry of decay in the triplets. All jets within a triplet are further required to pass the medium QGD selection and at least one jet to pass the tight QGD selection. This search is performed in three regions of gluino mass, and the selection variables are optimized individually for each, as shown in Table 1.

Table 1: Selection requirements for the resolved trijet resonance search are listed for the three regions of gluino mass.

Region	Gluino mass range (GeV)	Jet p_T (GeV)	H_T (GeV)	Sixth jet p_T (GeV)	$D^2_{[(6,3)+(3,2)]}$	A_m	Δ (GeV)	$D^2_{[3,2]}$
1	200–500	>30	>600	>40	<1.25	<0.25	>250	<0.05
2	500–900	>30	>600	>50	<1.00	<0.175	>180	<0.175
3	900–2000	>50	>900	>125	<0.9	<0.15	>175	<0.2

GP regression is used to estimate the dominant background from QCD [49]. The advantages of GP regression over more conventional methods, which employ a linear expansion over a fixed set of basis functions such as polynomials or Gaussians, are its non-parametric flexibility and a principled Bayesian framework. GP regression is defined using a kernel that directly encodes an understanding of the underlying physics, manifesting as covariance among the bin counts [50]. In these analyses the Radial Bias Function (RBF) kernel is used to model the QCD multijet invariant mass distributions as described in Ref. [49]. GP regression was observed to accurately model multijet invariant mass distributions from QCD. The signal shapes are modeled by simple Gaussian distributions. A bias test was performed at various masses by injecting signal shapes to validate the fitting procedure, and yielded no indication of significant bias in the signal extraction process. The background contributions from processes with SM multijet resonances ($t\bar{t}$, W/Z + jets) are taken from simulation, with their cross sections constrained by the GP regression in data.

For the boosted trijet and boosted dijet searches, this background estimation procedure is applied to the average jet mass ($\bar{m} = (m_1 + m_2)/2$) distribution for events passing the selection criteria for pair-produced boosted resonances decaying to three or two quarks, respectively, as illustrated in Fig. 2. Similarly, in the resolved trijet search, this background estimation procedure is applied to the trijet invariant mass (m_{jjj}) distributions for jet triplets passing the selection

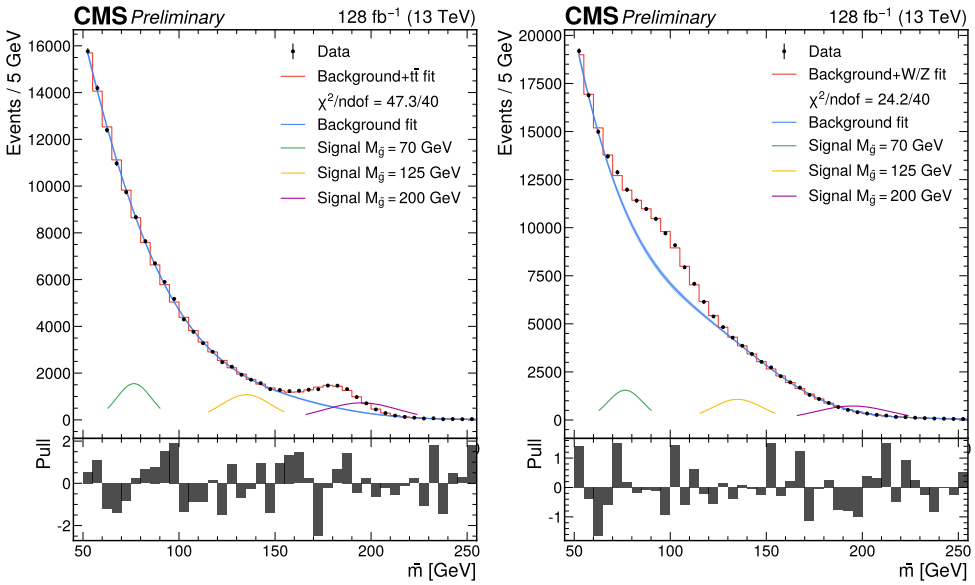

Figure 2: The distribution of average jet mass in the data (points), for the search for pair-produced boosted resonances decaying to trijets (left) and dijets (right), compared to a background-only prediction from GP regression (blue), and the full background fit including simulations (red) of SM resonances from $t\bar{t}$ (left) and W/Z +jets (right). Also shown are the expected shapes of signals from R-parity violating gluinos with resonance masses 70 (green), 125 (yellow) and 200 GeV (purple).

criteria, as illustrated in Fig. 3. The GP regression and added $t\bar{t}$ and W/Z simulations describes the data well, with p-values from χ^2 metric of 0.75, 0.12, and 0.97 for the three regions of the resolved trijet search, and 0.21 and 0.96 for the boosted trijet and boosted dijet searches. The rates of $t\bar{t}$ and W/Z production observed in the trijet and dijet searches respectively are within SM expectations.

The widths of the Gaussian distributions used to model the signal are extracted from the mass spectra of the simulated gluino and top squarks, and found to be 9.1% and 7.5% of the resonance mass in the boosted and resolved topologies, respectively. The search is performed by fitting these signal templates and the backgrounds simultaneously to the data using a Bayesian approach described in detail below. No significant deviation from the SM is found; the largest excess occurs at a reconstructed trijet mass of approximately 721 GeV as shown in Fig. 3 (middle) with a local significance below 3 standard deviations. This corresponds to a resonance mass of 768 GeV.

This analysis uses a Bayesian procedure with a uniform prior signal strength to evaluate significance and determine the 95% confidence level upper limit. Markov chain Monte Carlo (MCMC) method is used to marginalize the nuisance parameters [49, 51]. The MCMC is built with 12,500 samples in each of 20 independent MC trajectories. MCMC sampling procedure is applied to 1000 synthetic datasets. The posterior distributions from these are used to obtain the estimated limits.

Nuisance parameters are assigned for factors that define the signal shape: mean, width, and a nuisance parameter for the uncertainty in signal rate. For each of the parameters, various factors contribute to its uncertainty. The shape and rate uncertainty that affect the signal also affects the $t\bar{t}$ background in a correlated manner. The uncertainty for efficiency of the selection based on the substructure variables is measured using the semi-leptonic decay of $t\bar{t}$. This

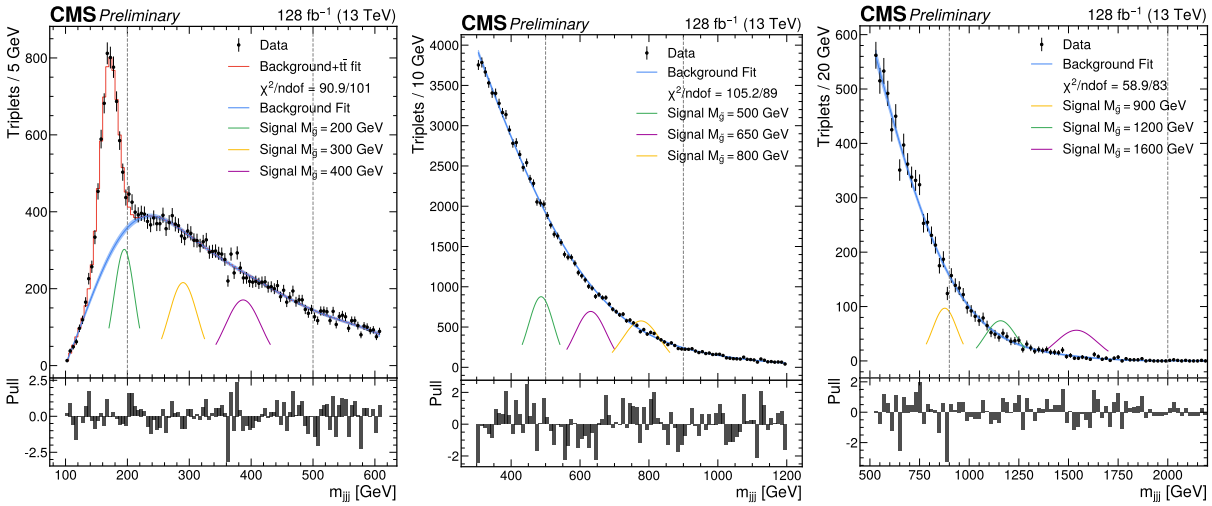

Figure 3: The distribution of trijet mass in the data (points), for the three search regions (left, middle and right) for pair-produced resolved resonances decaying to trijets, compared to a background-only fit from GP regression (blue), and (left) the full background fit including simulation of a SM resonance from t tbar production (red). Also shown are the expected shapes of signals from R-parity violating gluinos at various masses (green, yellow and purple).

is the dominant systematic uncertainty. The uncertainty in the luminosity measurement also contributes to the signal rate uncertainty. However, the contributions from the choice of parton distribution function for the generation have a negligible effect. All the signal rate uncertainties are represented as log-normal probability density functions (PDFs). The uncertainties arising from the jet energy corrections and resolution affect the shape of the signal. They are manifested as uncertainties in the mean and width of the signal, respectively. The nuisance parameters corresponding to the signal and background shape uncertainty are assigned to Gaussian PDFs. The standard deviations for all nuisance parameters are shown in Table 2.

Table 2: Summary of the systematic uncertainties in the signal yield and shape for the searches for resolved and merged, three and two quark resonances.

Source of systematic	Effect	Resolved	3 merged quarks	2 merged quarks
Trigger efficiency	Yield	2.5%	2.5%	2.5%
Acceptance	Yield	5 %	5%	5%
Jet energy correction	Signal mean	3.5%	3.5%	3.5%
Jet energy correction	Signal Width	12%	12%	12%
QGD efficiency	Yield	5%	-	-
$\tau_{32,\mathrm{DDT}}$ SF	Yield	-	4%	-
$N^1_{2,\mathrm{DDT}}$ SF	Yield	-	-	20%

Upper limits on the product of cross section, branching fraction, and acceptance for the narrow resonances searches are shown in Fig. 4. The results were compared to asymptotic $\mathrm{CL_s}$ [52] limits for a subset of signal masses and are found to be in close agreement. The total acceptance varies between 10^{-5} and 6×10^{-5}, 4×10^{-5} and 1.7×10^{-4}, and 2×10^{-3} and 6×10^{-2} over the mass intervals shown in Fig. 4 for the merged three quark, merged two quark, and resolved three quark resonance searches, respectively. The most significant deviation from SM expectations occurs at a resonance mass of 768 GeV for the resolved search, corresponding to the

aforementioned effect at a trijet mass of 721 GeV. The upper limits are compared to predictions for RPV gluinos, RPV top squarks and mass degenerate RPV higgsinos [53]. The trijet searches exclude RPV gluinos with mass between 70 GeV and 1.7 TeV, extending to both lower and higher mass compared to the existing exclusions between 0.2 and 1.5 TeV at 95% CL [18]. The boosted dijet resonance search excludes RPV top squarks with masses between 70 and 200 GeV at 95% CL, extending to lower mass and yielding higher sensitivity [23]. Finally, the boosted trijet resonance search sets the first limits on mass degenerate RPV higgsinos, excluding masses between 70 and 75 GeV and between 95 and 105 GeV at 95% CL.

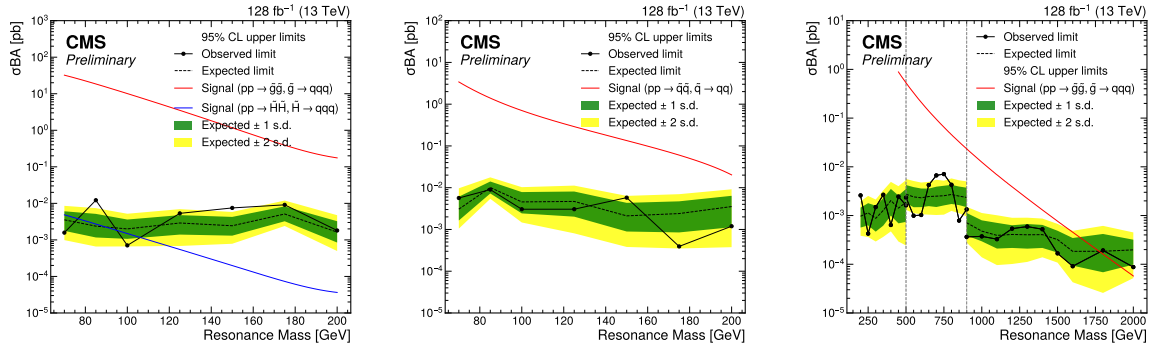

Figure 4: Observed limits (points) and expected limits (dashes) on the product of cross section (σ), branching fraction (B), and acceptance (A) for pair produced merged three quark resonances (left), pair produced merged two quark resonances (middle), and pair produced resolved three quark resonances (right). The vertical lines on the resolved three quark resonance limits (right) indicate the different search mass regions. The variations at the one and two standard deviation levels in the expected limits (shaded bands) are also shown. Limits are compared to predictions for three signals of pair production of R-parity violating resonances: gluinos (red in left and right), mass degenerate higgsinos (blue in left), and top squarks (red in middle).

A search has been performed for pair-produced multijet resonances, where the jets in the final state can be either merged or individually resolved. This search observes the hadronic decays of standard model resonances, such as the top quark and W, Z bosons. New, additional resonances were not observed, with the largest effect seen for a resonance mass of 768 GeV, but with a local significance below 3 standard deviations. This search extends the previous limits on R-parity violating models of top squarks and gluinos, and sets the first limits on R-parity violating mass-degenerate, prompt, and hadronically decaying higgsinos.

References

[1] CMS Collaboration, "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC", *Phys. Lett. B* **716** (2012) 30, doi:10.1016/j.physletb.2012.08.021, arXiv:1207.7235.

[2] CMS Collaboration, "Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV", *JHEP* **06** (2013) 081, doi:10.1007/JHEP06(2013)081, arXiv:1303.4571.

[3] ATLAS Collaboration, "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC", *Phys. Lett. B* **716** (2012) 1, doi:10.1016/j.physletb.2012.08.020, arXiv:1207.7214.

[4] Y. Fukuda et al., "Evidence for oscillation of atmospheric neutrinos", *Phys. Rev. Lett.* **81** (1998) 1562, doi:10.1103/physrevlett.81.1562.

[5] Muon $g-2$ Collaboration, "Measurement of the positive muon anomalous magnetic moment to 0.46 ppm", *Phys. Rev. Lett.* **126** (2021) 141801, doi:10.1103/PhysRevLett.126.141801.

[6] LHCb Collaboration, "Test of lepton universality in beauty-quark decays", *Nature Phys.* **18** (2022) 277, doi:10.1038/s41567-021-01478-8, arXiv:2103.11769.

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

[8] P. Asadi, "New solutions to the charged current B-anomalies". PhD thesis, Rutgers U., Piscataway, 2019. doi:10.7282/t3-yhaa-dm09.

[9] H. Easa, T. Gregoire, and D. Stolarski, "New limits on coloured three jet resonances", *JHEP* **09** (2020) 131, doi:10.1007/JHEP09(2020)131, arXiv:2003.00014.

[10] B. A. Dobrescu, R. M. Harris, and J. Isaacson, "Ultraheavy resonances at the LHC: beyond the QCD background", Technical Report FERMILAB-PUB-18-561-T, 10, 2018. arXiv:1810.09429.

[11] CMS Collaboration, "Search for resonant and nonresonant production of pairs of dijet resonances in proton-proton collisions at $\sqrt{s}$ = 13 TeV", *JHEP* **07** (2023) 161, doi:10.1007/JHEP07(2023)161, arXiv:2206.09997.

[12] CMS Collaboration, S. Mukherjee, "Data scouting : A new trigger paradigm", in *5th Large Hadron Collider Physics Conference*. 8, 2017. arXiv:1708.06925.

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

[14] 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. [Erratum: doi:10.1103/PhysRevD.93.039901].

[15] ATLAS Collaboration, "Search for R-parity-violating supersymmetric particles in multi-jet final states produced in p-p collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC", *Phys. Lett. B* **785** (2018) 136, doi:10.1016/j.physletb.2018.08.021, arXiv:1804.03568.

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

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

[18] CMS Collaboration, "Search for pair-produced three-jet resonances in proton-proton collisions at $\sqrt{s}$ =13 TeV", *Phys. Rev. D* **99** (2019) 012010, doi:10.1103/PhysRevD.99.012010, arXiv:1810.10092.

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

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

[21] ATLAS Collaboration, "Search for pair production of massive particles decaying into three quarks with the ATLAS detector in $\sqrt{s} = 7$ TeV pp collisions at the LHC", *JHEP* **12** (2012) 086, doi:10.1007/JHEP12(2012)086, arXiv:1210.4813.

[22] 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.

[23] CMS Collaboration, "Search for pair-produced resonances decaying to quark pairs in proton-proton collisions at $\sqrt{s} = 13$ TeV", *Phys. Rev. D* **98** (2018) 112014, doi:10.1103/PhysRevD.98.112014, arXiv:1808.03124.

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

[25] CMS Collaboration, "Performance of the CMS level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **15** (2020) P10017, doi:10.1088/1748-0221/15/10/P10017, arXiv:2006.10165.

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

[27] CMS Collaboration, "Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC", *JINST* **16** (2021) P05014, doi:10.1088/1748-0221/16/05/P05014, arXiv:2012.06888.

[28] CMS Collaboration, "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s} = 13$ TeV", *JINST* **13** (2018) P06015, doi:10.1088/1748-0221/13/06/P06015, arXiv:1804.04528.

[29] CMS Collaboration, "Description and performance of track and primary-vertex reconstruction with the CMS tracker", *JINST* **9** (2014) P10009, doi:10.1088/1748-0221/9/10/P10009, arXiv:1405.6569.

[30] 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.

[31] CMS Collaboration, "Performance of reconstruction and identification of τ leptons decaying to hadrons and ν_τ in pp collisions at $\sqrt{s} = 13$ TeV", *JINST* **13** (2018) P10005, doi:10.1088/1748-0221/13/10/P10005, arXiv:1809.02816.

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

[33] CMS Collaboration, "Performance of missing transverse momentum reconstruction in proton-proton collisions at $\sqrt{s} = 13$ TeV using the CMS detector", *JINST* **14** (2019) P07004, doi:10.1088/1748-0221/14/07/P07004, arXiv:1903.06078.

[34] 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.

[35] J. Alwall et al., "MadGraph/MadEvent v4: The new web generation", *JHEP* **09** (2007) 028, doi:10.1088/1126-6708/2007/09/028, arXiv:0706.2334.

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

[37] S. Mrenna and P. Richardson, "Matching matrix elements and parton showers with HERWIG and PYTHIA", *JHEP* **05** (2004) 040, doi:10.1088/1126-6708/2004/05/040, arXiv:hep-ph/0312274.

[38] M. Cacciari, G. P. Salam, and G. Soyez, "FastJet user manual", *Eur. Phys. J. C* **72** (2012) 1896, doi:10.1140/epjc/s10052-012-1896-2, arXiv:1111.6097.

[39] M. Cacciari and G. P. Salam, "Dispelling the N^3 myth for the k_t jet-finder", *Phys. Lett. B* **641** (2006) 57, doi:10.1016/j.physletb.2006.08.037, arXiv:hep-ph/0512210.

[40] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-kt jet clustering algorithm", *JHEP* **2008** (2008) 063, doi:10.1088/1126-6708/2008/04/063.

[41] CMS Collaboration, "Pileup mitigation at CMS in 13 TeV data", *JINST* **15** (2020) P09018, doi:10.1088/1748-0221/15/09/P09018, arXiv:2003.00503.

[42] CMS Collaboration, "Jet performance in pp collisions at 7 TeV", CMS Physics Analysis Summary CMS-PAS-JME-10-003, 2010.

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

[44] J. Thaler and K. Van Tilburg, "Identifying boosted objects with N-subjettiness", *JHEP* **03** (2011) 015, doi:10.1007/JHEP03(2011)015, arXiv:1011.2268.

[45] I. Moult, L. Necib, and J. Thaler, "New angles on energy correlation functions", *JHEP* **12** (2016) 153, doi:10.1007/JHEP12(2016)153, arXiv:1609.07483.

[46] J. Dolen et al., "Thinking outside the ROCs: Designing decorrelated taggers (DDT) for jet substructure", *JHEP* **05** (2016) 156, doi:10.1007/JHEP05(2016)156, arXiv:1603.00027.

[47] P. T. Komiske, E. M. Metodiev, and J. Thaler, "Energy flow networks: Deep sets for particle jets", *JHEP* **01** (2019) 121, doi:10.1007/JHEP01(2019)121, arXiv:1810.05165.

[48] C. E. Rasmussen and C. K. I. Williams, "Gaussian processes for machine learning". The MIT Press, 2005. doi:10.7551/mitpress/3206.001.0001, ISBN 9780262256834.

[49] A. Gandrakota, A. Lath, A. V. Morozov, and S. Murthy, "Model selection and signal extraction using Gaussian process regression", *JHEP* **02** (2023) 230, doi:10.1007/JHEP02(2023)230, arXiv:2202.05856.

[50] M. Frate et al., "Modeling smooth backgrounds and generic localized signals with Gaussian processes", arXiv:1709.05681.

[51] D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman, "emcee: The MCMC hammer", *Publ. Astron. Soc. Pac.* **125** (2013) 306, doi:10.1086/670067.

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

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