# Inclusive Search for Supersymmetry Using Razor Variables in pp Collisions at $s=7$TeV

The MIT Faculty has made this article openly available. *Please share* how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1103/PhysRevLett.111.081802">http://dx.doi.org/10.1103/PhysRevLett.111.081802</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>American Physical Society</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Tue Oct 16 06:16:18 EDT 2018</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/81377">http://hdl.handle.net/1721.1/81377</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
Inclusive Search for Supersymmetry Using Razor Variables in \( pp \) Collisions at \( \sqrt{s} = 7 \) TeV

S. Chatrchyan et al.*
(CMS Collaboration)
(Received 8 January 2013; published 23 August 2013)

An inclusive search is presented for new heavy particle pairs produced in \( \sqrt{s} = 7 \) TeV proton-proton collisions at the LHC using 4.7 ± 0.1 fb\(^{-1}\) of integrated luminosity. The selected events are analyzed in the 2D razor space of \( M_R \), an event-by-event indicator of the heavy particle mass scale, and \( R \), a dimensionless variable related to the missing transverse energy. The third-generation sector is probed using the event heavy-flavor content. The search is sensitive to generic supersymmetry models with minimal assumptions about the superpartner decay chains. No excess is observed in the number of events beyond that predicted by the standard model. Exclusion limits are derived in the CMSSM framework as well as for simplified models. Within the CMSSM parameter space considered, gluino masses up to 800 GeV and squark masses up to 1.35 TeV are excluded at 95% confidence level depending on the model parameters. The direct production of pairs of top or bottom squarks is excluded for masses as high as 400 GeV.

DOI: 10.1103/PhysRevLett.111.081802

Models with softly broken supersymmetry (SUSY) [1] predict heavy superpartners of the standard model (SM) particles. Experimental searches for \( R \)-parity [2] conserving SUSY have focused on signatures combining energetic hadronic jets and leptons or photons from the decays of pair-produced squarks and gluinos, with large missing transverse energy (\( E_T^{\text{miss}} \)) from the two weakly interacting lightest neutral superpartners (LSPs) produced in separate decay chains. Recent publications include results from both the Tevatron [3,4] and the Large Hadron Collider (LHC) [5–26].

In SUSY models, the scale of soft SUSY breaking is related to the scale of electroweak symmetry breaking. This implies either that the soft-breaking mass parameters cannot be too large, or that the smallness of the electroweak scale is explained by large cancellations arising from relations among these parameters in the high-energy theory. The latter possibility is complicated by large radiative corrections, particularly those induced by the soft-breaking parameters that are responsible for the masses of the top and bottom squarks, the superpartners of the third-generation quarks. It is thus of special importance to search for the lightest allowed top and bottom squarks, whose decays will be enriched in heavy-flavor quarks.

In this Letter we present results of an inclusive search for new heavy particles. The analysis is designed to be largely independent of the details of the decay chains and measures deviations from the characteristic distributions of the relevant SM processes in the razor variable plane [27,28]. It is generically sensitive to the production of pairs of heavy particles, provided that the decays of these particles produce significant \( E_T^{\text{miss}} \), that these particles are substantially heavier than any SM particle, and that they are strongly produced in high-energy proton-proton collisions. The selection requires only two or more energetic reconstructed calorimeter objects [29]. The selected events are sorted hierarchically into exclusive data samples, categorized according to the lepton multiplicity in the event. The analysis is repeated with the requirement of the presence of a bottom-quark jet (\( b \)-jet) to search for third-generation-enhanced SUSY signatures. The major backgrounds are top production and vector boson production in association with jets. Using Monte Carlo simulation, we verified that the contribution from other SM processes (e.g., single top production or the pair production of electroweak vector bosons) is negligible.

The razor kinematic variables are based on the generic process of the pair production of two heavy particles, each decaying to an undetected particle plus visible decay products. The razor kinematic variables are used to test, event by event, the hypothesis that the reconstructed particles in the events represent the visible portion of the decays of two heavy particles, each producing also an invisible particle. Regardless of its complexity, each event is treated as a dijetlike event by grouping all the physics objects detectable in the calorimeters (hadronic jet candidates and isolated electrons) into two megajets [28]. Muons are considered invisible objects, in order to minimize the differences between the razor variables computed after the event reconstruction and the corresponding values derived from the calorimetric jets at the trigger level. Assuming the pair of megajets accurately reconstructs the visible portion of the parent particle decays, the signal...
kinematics is equivalent, for example, to pair production of heavy squarks $q_1^*, q_2^*$, with $q_i \rightarrow j_i X^0$, where the $X^0$ are LSPs and $j_i$ denotes the visible products of the decays.

The $M_R$ razor kinematic variable is defined in terms of the momentum of the two megajets as

$$ M_R = \frac{(p_T^j + p_T^k)^2 - (p_T^l + p_T^m)^2}{2p_T^j p_T^m} $$

and is, by construction, invariant under longitudinal boosts. In the approximation of massless megajets and negligible initial-state energy- and $\eta$-dependent correction factors. Only jet candidates with $p_T > 40$ GeV within $|\eta| < 3.0$ are retained. The jet energy scale uncertainty for these corrected jets is 5% [34]. To match the trigger requirements, the $p_T$ of the two leading jets is required to be greater than 60 GeV. The transverse momentum imbalance in the event, $p_T^{miss}$, is reconstructed using the particle flow algorithm [35].

The reconstructed jets are grouped into two megajets [29]. The megajets are constructed as a sum of the four-momenta of their constituent objects. After the baseline selection and calculation of the variables $R$ and $M_R$, the events are assigned to one of six final-state boxes according to whether the event has zero, one, or two isolated leptons, divided according to lepton flavor (electrons and muons) as shown in Table I.

The requirements given in Table I define the full analysis regions of the $R^2$-$M_R$ plane, where the analysis is performed for each box. They are the loosest possible requirements that allow for the valid background description, while at the same time maintaining fully efficient triggers. To prevent ambiguities for events satisfying the selection requirements of more than one box [29], the boxes are arranged in a predefined hierarchy, as given in Table I. Each event is uniquely assigned to the first box whose criteria are satisfied by the event.

<table>
<thead>
<tr>
<th>Lepton boxes $M_R &gt; 300$ GeV, $0.11 &lt; R^2 &lt; 0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELE-MU $p_T^e &gt; 20$ GeV, $p_T^\mu &gt; 15$ GeV</td>
</tr>
<tr>
<td>MU-MU $p_T^{e1} &gt; 15$ GeV, $p_T^{\mu2} &gt; 10$ GeV</td>
</tr>
<tr>
<td>ELE-ELE $p_T^{e1} &gt; 20$ GeV, $p_T^{\mu2} &gt; 10$ GeV</td>
</tr>
<tr>
<td>MU $p_T^e &gt; 12$ GeV, $p_T^\mu &gt; 20$ GeV</td>
</tr>
<tr>
<td>ELE $p_T^e &gt; 20$ GeV</td>
</tr>
<tr>
<td>HAD box $M_R &gt; 400$ GeV, $0.18 &lt; R^2 &lt; 0.5$</td>
</tr>
</tbody>
</table>

TABLE I. Razor boxes definition. The variables and requirements are explained in the text.
rate [36,37]. These six boxes define the razor inclusive analysis of data samples with enhanced heavy-flavor content. The typical b-tagging efficiency is 65% [38].

The razor analysis is guided by studies of simulated events generated with the PYTHIA6 [39] and MADGRAPH [40] Monte Carlo programs, implemented using the CMS GEANT4-based [41] detector simulation, and then processed by the same software as that used to reconstruct data. Events with QCD multijets, top quarks, and electroweak (V) vector bosons are generated with MADGRAPH interfaced with PYTHIA for parton showering, hadronization, and the underlying event description. V + jets events are generated with up to four additional tree-level strong emissions and t¯t + jets with up to three. To generate Monte Carlo samples for SUSY, the mass spectrum is first calculated with SOFTSUSY [42] and the decays with SUSYHIT [43]. The PYTHIA program is used with the SUSY Les Houches Accord (SLHA) interface [44] to generate the events. Next-to-leading order (NLO) plus next-to-leading-logarithm (NLL) cross section calculations are used [45–50].

For each of the major SM backgrounds, a control data sample is defined from a subset of the data dominated by this particular background. It is used to obtain a description of the shapes of the background components from data. In both simulation and control data samples, the distributions of the major SM background events (QCD multijets, t¯t + jets, Z + jets, and W + jets) are found to have a simple exponential dependence on the razor variables R^2 and M_R over a large part of the R^2-M_R plane. A full 2D SM background representation is built using statistically independent data samples. This representation is used as input to the final fit performed in each fit region (FR) defined by an L-shaped area in the R^2-M_R plane such as shown in Fig. 2. The fit region is dominated by SM processes. Signal contamination has only a small impact on the determination of the background shape in the fit region, as demonstrated through studies based on simulation. The 2D background model obtained in the fit region is extrapolated to the rest of the R^2-M_R plane where the analysis is sensitive to a potential SUSY signal.

The fit function has parameters describing the shapes and normalization of the R^2 and M_R distributions of the SM backgrounds. The two-dimensional probability density function (pdf) P_j(M_R, R^2) describing the R^2 versus M_R distribution of each considered SM process j is found to be well approximated by instances of the same function F_j(M_R, R^2):

\[ F_j(M_R, R^2) = [k_j(M_R - M^0_{R,j})(R^2 - R^2_{0,j}) - 1] \times e^{-k_j(M_R - M^0_{R,j})(R^2 - R^2_{0,j})}. \]  

(1)

The scaling of the exponent as a function of the thresholds on M_R and R^2 is described by the k_j parameters of the function. When integrated over M_R (R^2), this function recovers the exponential dependence on R^2 (M_R).

Each SM process in a given final-state box is well described by a pdf P_j that is the sum of a first and a second component of the functional form Eq. (1) with separate normalizations. Studies of simulated events and fits to control data samples with either a b-jet requirement or a b-jet veto indicate that the parameters corresponding to the first components of these backgrounds (with steeper slopes at low M_R and R^2) are box dependent. The parameters describing the second component are box independent, and at the current precision of the background model, they are identical among the dominant backgrounds considered in these final states.

These sets of independent data control samples are used to derive a priori the background shape parameters. The results are incorporated in the final fits as a set of Gaussian penalty terms [51,52] for the parameters k_j, M^0_{R,j}, and R^0_{0,j} multiplying the final likelihood [Eq. (2)]. The rms values of the penalty terms for the k_j parameters are typically ~30%.

An extended and unbinned maximum likelihood (ML) fit is performed in each box using ROOFIT [52]. The fit performed in the fit region of the R^2-M_R plane provides the description of the SM background in the full plane. The likelihood function for a given box is written as [53]

\[ L_b = \frac{e^{\left(\sum_{j \in SM} N_j \right) \prod_{i=1}^{N} \left( \sum_j N_j P_j(M_{R,i}, R^2_i) \right)}}{N!}. \]  

(2)

where N is the total event yield in the box, the sum runs over all the SM processes relevant for that box, and N_j is the yield of a given fit sample in the box.

The values of the shape parameters that maximize the likelihood in these fits, along with the corresponding covariance matrix, are used to define the background model and the uncertainty associated with it. Additional background shape uncertainties due to the choice of the functional form were found to be negligible [29].

The result of the ML fit projected on M_R and R^2 is shown in Fig. 1 for the HAD box. No significant discrepancy is observed between the data and the fit model in any of the six boxes [29].

To establish the compatibility of the background model with the observed data set, we define six signal regions (SRi) in the tail of the background distribution. Using the background model returned by the ML fit, we derive the distribution of the expected yield in each SRi using pseudoexperiments accounting for correlations and uncertainties on the parameters describing the background model. For each of the SRi the distribution of the number of events derived by the pseudoexperiments is used to calculate a two-sided p value (as shown for the HAD box in Fig. 2), corresponding to the probability of observing an equal or less probable outcome for a counting experiment in each signal region. The p values test the compatibility of the observed number of events in data with the SM expectation obtained from the background parametrization. We quote
Gluino-gluino production dominates at low and high mass across the parameter space with squark-squark and changing relevant SUSY strong production processes. The shape of the observed exclusion curves reflect the cutoff first component in the median-expected limit at lower mass, respectively. The observed limit is less constraining than the median and the mode of the yield distribution for each SR, together with the observed yield.

For each box we consider the test statistic given by the logarithm of the likelihood ratio $\ln Q = \ln (L(s+b|H)/L(b|H))$, where $H$ is the hypothesis under test: $H_1$ (signal plus background) or the null hypothesis $H_0$ (background only). Given the distribution of $\ln Q$ for background-only and signal-plus-background pseudoexperiments, and the value of $\ln Q$ observed in the data, we calculate $CL_{s+b}$ and $1 - CL_b$ [54,55]. From these values the $CL_s = CL_{s+b}/CL_b$ is computed for that model point. A point in the constrained minimal supersymmetric standard model (CMSSM) plane is excluded at 95% confidence level (C.L.) if $CL_s < 0.05$. The result is shown in Fig. 3. The shape of the observed exclusion curves reflect the changing relevant SUSY strong production processes across the parameter space with squark-squark and gluino-gluino production dominating at low and high $m_0$, respectively. The observed limit is less constraining than the median-expected limit at lower $m_0$ due to an excess of observed events in the HAD box at large $R^2$, where squark-pair production dominates over gluino-pair production.

Cascading decays of gluinos yield more leptons than decays of squarks. Thus, relative to hadronic boxes, the contribution of lepton boxes increases with $m_0$.

We estimate the systematic uncertainty on the signal shape model due to parton density functions (point by point up to 30%), jet energy scale (point by point up to 1%), and lepton identification (using $Z \to \ell^+\ell^-$ data, 1% per lepton), as well as on the signal yield due to the luminosity uncertainty (2.2%) [56], the theoretical cross section (point by point up to 15%), razor trigger efficiency uncertainty (2.2%) [56], and the variation of the nuisance parameters.

FIG. 1 (color online). Projection of the 2D fit result on $M_R$ (left) and $R^2$ (right) for the HAD box. The blue histogram is the total standard model prediction. The red and green histograms represent a steep-slope component denoted as $V + jets$ first component, and a component that encapsulates both the steep-slope first component in $t\bar{t} + jets$ and the effective second component, which is indistinguishable for the different SM background processes. The fit is performed in the $R^2$-$M_R$ fit region (FR as shown in Fig. 2) and projected into the full analysis region. The full error on the total background prediction is drawn in these projections, including the one due to variation of the nuisance parameters.

FIG. 2 (color online). The $p$ values corresponding to the observed number of events in the HAD box signal regions (SR$i$). The green region indicates the fit region in the HAD box. Similar results are obtained for the other boxes.

FIG. 3 (color online). Observed (solid blue curve) and median-expected (dashed curve, shown with its ±1 standard deviation uncertainty band) 95% C.L. limits in the $(m_0, m_{1/2})$ CMSSM plane (drawn according to [61]) with $\tan \beta = 10$, $M_A = 0$ GeV, and $sgn(\mu) = +1$. Shown separately are the observed HAD-only (solid crimson) and leptonic-only (solid green) 95% C.L. limits.
FIG. 4 (color online). Summary of the 95% CL excluded largest parent mass as a function of the LSP mass in each of the simplified models studied. Results from the inclusive razor analysis (upper bars) and the $b$-jet razor analysis (lower bars) are shown.

(2%), and lepton trigger efficiency uncertainty (3%). In the $b$-tag analysis path an additional systematic is considered for the $b$-tagging efficiency (between 6% and 20% in $p_T$ bins [36]). We consider variations of the function modeling the signal uncertainty (log-normal versus Gaussian) as well as the $R^2$ and $M_\pi$ binning choice, finding negligible deviations in the result.

The results are also interpreted as cross section limits on a number of simplified models [57], where a limited set of hypothetical particles and decay chains are introduced to produce a given topological signature. Specific applications of these ideas have appeared in Refs. [58–60]. For each model studied, the excluded cross section at 95% C.L. is derived as a function of the mass of the produced particles (gluinos or squarks, depending on the model) and the LSP mass, as well as the exclusion curve corresponding to the NLL-NLO cross section. Exclusion curves for a factor of 3 cross section enhancement or reduction are corresponding to the NLL-NLO cross section. FIG. 4 shows the 95% C.L. excluded largest parent mass as a function of the LSP mass in each of the simplified model studies, for both the inclusive and $b$-jet versions of the analysis.

In summary, we performed a search for squarks and gluinos using a data sample of 4.7 fb$^{-1}$ of CMS data at $\sqrt{s} = 7$ TeV proton-proton collisions in the razor variable space using a 2D shape description of the relevant standard model processes.

No significant excess over the background expectations is observed, and the results are presented as a 95% C.L. limit in the $(m_0, m_{1/2})$ CMSSM parameter space. For $m(\tilde{q}) \sim m(\tilde{g})$ we exclude squarks and gluinos up to 1.35 TeV, and for $m(\tilde{q}) > m(\tilde{g})$ we exclude gluinos up to 800 GeV. For simplified models we exclude up to 1 TeV for the gluino mass and up to 800 GeV for the first and second generation squark masses. For direct production of pairs of top and bottom squarks we exclude top and bottom squark masses up to 400 GeV depending on the LSP mass.

We congratulate our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centers and personnel of the Worldwide LHC Computing Grid for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC and the CMS detector provided by the following funding agencies: BMWF and FWF (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MEYS (Bulgaria); CERN; CAS, MoST, and NSFC (China); COLCIENCIAS (Colombia); MSES (Croatia); RPF (Cyprus); MoER, SF0690030s09 and ERDF (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, DFG, and HGF (Germany); GSRT (Greece); OTKA and NKTH (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); NRF and WCU (Republic of Korea); LAS (Lithuania); CINVESTAV, CONACYT, SEP, and UASLP-FAI (Mexico); MูN (New Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Armenia, Belarus, Georgia, Ukraine, Uzbekistan); MON, RosAtom, RAS, and RFBR (Russia); MSTD (Serbia); SEIDI and CPAN (Spain); Swiss Funding Agencies (Switzerland); NSC (Taipei); ThEPCenter, IPST, and NSTDA (Thailand); TUBITAK and TAEK (Turkey); NASU (Ukraine); STFC (United Kingdom); DOE and NSF (USA). Individuals have received support from the Marie-Curie program and the European Research Council (European Union); the Leventis Foundation; the A.P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; Fonds pour la Formation a` la Recherche dans l’Industrie et dans l’Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); the Ministry of Education, Youth and Sports (MEYS) of Czech Republic; the Council of Science and Industrial Research, India; the Compagnia di San Paolo (Torino); the Weston Havens Foundation (US) and the HOMING PLUS program of Foundation for Polish Science, cofinanced from European Union, Regional Development Fund.

PRL 111, 081802 (2013)  PHYSICAL REVIEW LETTERS week ending 23 AUGUST 2013

[38] CMS Collaboration, JINST 8, P04013 (2013).
aDeceased.
bAlso at Vienna University of Technology, Vienna, Austria.
cAlso at National Institute of Chemical Physics and Biophysics, Tallinn, Estonia.
dAlso at California Institute of Technology, Pasadena, USA.
eAlso at CERN, European Organization for Nuclear Research, Geneva, Switzerland.
fAlso at Laboratoire Leprince-Ringuet, Ecole Polytechnique, IN2P3-CNRS, Palaiseau, France.
gAlso at Suez Canal University, Suez, Egypt.
hAlso at Zewail City of Science and Technology, Zewail, Egypt.
iAlso at Cairo University, Cairo, Egypt.
jAlso at Fayoum University, El-Fayoum, Egypt.
kAlso at British University in Egypt, Cairo, Egypt.
lNow at Ain Shams University, Cairo, Egypt.
mAlso at National Centre for Nuclear Research, Swierk, Poland.
nAlso at Université de Haute Alsace, Mulhouse, France.
oAlso at Joint Institute for Nuclear Research, Dubna, Russia.
pAlso at Moscow State University, Moscow, Russia.
qAlso at Brandenburg University of Technology, Cottbus, Germany.
rAlso at The University of Kansas, Lawrence, KS, USA.
sAlso at Institute of Nuclear Research ATOMKI, Debrecen, Hungary.
tAlso at Eötvös Loránd University, Budapest, Hungary.
uAlso at Tata Institute of Fundamental Research-HECR, Mumbai, India.
vAlso at University of Visva-Bharati, Santiniketan, India.
wAlso at Sharif University of Technology, Tehran, Iran.
xAlso at Isfahan University of Technology, Isfahan, Iran.
yAlso at Plasma Physics Research Center, Science and Research Branch, Islamic Azad University, Tehran, Iran.
zAlso at Facoltà Ingegneria, Università di Roma, Roma, Italy.
aaAlso at Università degli Studi Guglielmo Marconi, Roma, Italy.
abAlso at Università degli Studi di Siena, Siena, Italy.
cAlso at University of Bucharest, Faculty of Physics, Bucuresti-Magurele, Romania.
dAlso at Faculty of Physics of University of Belgrade, Belgrade, Serbia.
eeAlso at University of California, Los Angeles, CA, USA.
ffAlso at Scuola Normale e Sezione dell’INFN, Pisa, Italy.
fgAlso at INFN Sezione di Roma, Università di Roma, Roma, Italy.
ghAlso at University of Athens, Athens, Greece.
iAlso at Rutherford Appleton Laboratory, Didcot, United Kingdom.
jAlso at Paul Scherrer Institut, Villigen, Switzerland.
kAlso at Institute for Theoretical and Experimental Physics, Moscow, Russia.
lAlso at Albert Einstein Center for Fundamental Physics, Bern, Switzerland.
mAlso at Gaziosmanpasa University, Tokat, Turkey.
nAlso at Adiyaman University, Adiyaman, Turkey.
ooAlso at Izmir Institute of Technology, Izmir, Turkey.
ppAlso at The University of Iowa, Iowa City, IA, USA.
qqAlso at Mersin University, Mersin, Turkey.rAlso at Ozyegin University, Istanbul, Turkey.
sAlso at Kafkas University, Kars, Turkey.
tAlso at Suleyman Demirel University, Isparta, Turkey.
uuAlso at Ege University, Izmir, Turkey.
vvAlso at School of Physics and Astronomy, University of Southampton, Southampton, United Kingdom.
wwAlso at INFN Sezione di Perugia, Università di Perugia, Perugia, Italy.
xAlso at University of Sydney, Sydney, Australia.
yAlso at Utah Valley University, Orem, UT, USA.
zAlso at Institute for Nuclear Research, Moscow, Russia.
aaaAlso at University of Belgrade, Faculty of Physics and Vinca Institute of Nuclear Sciences, Belgrade, Serbia.
bbbAlso at Argonne National Laboratory, Argonne, IL, USA.
Also at Erzincan University, Erzincan, Turkey.
Also at Mimar Sinan University, Istanbul, Istanbul, Turkey.
Also at KFKI Research Institute for Particle and Nuclear Physics, Budapest, Hungary.
Also at Kyungpook National University, Daegu, Korea.