Searches for new physics using the $t\bar{t}$ invariant mass distribution in $pp$ collisions at $\sqrt{s} = 8$ TeV

S. Chatrchyan et al.*
(CMS Collaboration)

(Received 8 September 2013; published 22 November 2013)

Searches for anomalous top quark-antiquark production are presented, based on $pp$ collisions at $\sqrt{s} = 8$ TeV. The data, corresponding to an integrated luminosity of 19.7 fb$^{-1}$, were collected with the CMS detector at the LHC. The observed $t\bar{t}$ invariant mass spectrum is found to be compatible with the standard model prediction. Limits on the production cross section times branching fraction probe, for the first time, a region of parameter space for certain models of new physics not yet constrained by precision measurements.

DOI: 10.1103/PhysRevLett.111.211804

With the discovery of a Higgs boson with a mass around 125 GeV [1–3], the focus of particle physics has shifted towards understanding the properties of the new boson, uncovering the nature of the underlying electroweak symmetry breaking (EWSB) mechanism, and finding new physics. The standard model (SM) is believed to be an effective theory, i.e., a low-energy approximation of a more complete theory incorporating gravity and explaining the origin of many parameters that are simply postulated within the SM. Many models beyond the SM (BSM) have been proposed in order to alleviate the hierarchy problem of the SM, which stems from the fact that quantum-loop corrections to the Higgs boson mass diverge quadratically with the highest energy scale of the model, requiring an enormous degree of fine-tuning to ensure that the Higgs mass remains close to the $W$-boson mass up to the Planck scale. Since the largest quantum correction to the Higgs boson mass involves a top quark loop, it is natural to suppose that these BSM mechanisms would involve interactions with the top quark.

Potential solutions to the hierarchy problem include models with extra spatial dimensions, either flat [4] or warped [5,6]. In these models, gravity is allowed to permeate the multidimensional space, which results in its apparent scale in the SM from the Planck scale to the TeV scale, thus eliminating the hierarchy between the EWSB scale and the highest scale in the theory. Such models often contain Kaluza-Klein excitations of particles, including gravitons and gluons, both of which can have enhanced couplings to $t\bar{t}$ pairs [7]. Other new gauge bosons have been proposed, referred to generically as $Z'$, that also couple preferentially to $t\bar{t}$ pairs [8–13]. Furthermore, there may be additional spin-zero resonances that preferentially decay to $t\bar{t}$ pairs [13,14]. These various resonances may be observable as enhancements in the $t\bar{t}$ invariant mass spectrum.

Discrepancies have been observed in the forward-backward asymmetry of top quark production at the Tevatron [15]. Assuming this anomaly is due to new physics above the TeV scale, an enhancement of the $t\bar{t}$ rate at high invariant mass could be visible at the Large Hadron Collider (LHC) [16,17].

In this Letter, a search for anomalous production of $t\bar{t}$ events is presented, from data corresponding to an integrated luminosity of 19.7 fb$^{-1}$ of $pp$ collisions at $\sqrt{s} = 8$ TeV, recorded with the Compact Muon Solenoid (CMS) detector [18] at the LHC. These results represent a significant improvement over the previous searches [19–24], due primarily to the large increase in the high-$x$ parton luminosity from the higher LHC energy in 2012, but also because of the increased size of the data sample and the combination of several statistically independent channels. Specific comparisons are made to the resonant production of Randall-Sundrum Kaluza-Klein (RS KK) gluons [7], of a $Z'$ boson in the topcolor model [10], and of a scalar Higgs-like boson produced via gluon fusion through its couplings to the top quarks. In addition, enhancements of the $t\bar{t}$ invariant mass ($M_{tt}$) spectrum are constrained for $M_{tt} > 1$ TeV. These results probe, for the first time, a region of parameter space of models with warped extra dimensions not yet constrained by precision measurements [25].

Since the top quark decays primarily to a $W$ boson and a bottom ($b$) quark, top pair production signatures are classified based on whether the $W$ bosons decay to leptons or quarks. This measurement combines analyses utilizing the final states where one or both $W$ bosons from $t\bar{t}$ events decay to quarks (“semileptonic” and “all-hadronic” events, respectively). The events are classified into two categories based on the expected kinematics of the top quark decay products. In the first category, the $t\bar{t}$ pair is produced near the kinematic threshold, resulting in a topology where each parton is matched to a single jet.
electron with a $p_T$ threshold of 17 or 25 GeV, respectively, in combination with three jets with $p_T > 30$ GeV. Offline, we select events containing exactly one isolated muon (with $p_T^\mu > 26$ GeV), or electron (with $p_T^e > 30$ GeV), and at least four jets with $p_T > 70, 50, 30, 30$ GeV, respectively. The non-$W$ multijet background (NWMJ) is suppressed further by requiring the transverse missing momentum $E_T^{\text{miss}}$, the modulus of the vector sum of all measured particle $p_T$, to be larger than 20 GeV.

For the semileptonic boosted category, data were recorded with triggers requiring one muon ($p_T^\mu > 40$ GeV), or one electron ($p_T^e > 35$ GeV) in conjunction with two jets ($p_T > 100, 25$ GeV). Since the top quark decay products can be collinear in this regime, no isolation requirements on the leptons are imposed in either the trigger or offline selections. Offline, we select events containing exactly one muon with $p_T^\mu > 45$ GeV, or exactly one electron with $p_T^e > 35$ GeV and at least two jets with $p_T > 150, 50$ GeV, respectively. To reduce the contamination of the NWMJ processes, we follow the techniques of Ref. [23], placing requirements on the angle and relative momentum between the lepton and the nearest jet, and also requiring $E_T^{\text{miss}} > 50$ GeV and $E_T^{\text{miss}} + E_T^{\mu} > 150$ GeV.

The events satisfying the two semileptonic selections are separated into categories determined by the lepton flavor (electron or muon) and the number of $b$-tagged jets $N_{b\text{-tag}}$ (1 or $\geq 2$ $b$-tagged jets for the resolved analysis, and 0 or $\geq 1$ $b$-tagged jets for the boosted analysis). The purpose of this classification is to separate the sample into regions dominated by different background processes. The events in the categories with fewer $b$-tagged jets have a higher fraction of $W + \text{jets}$ events, whereas those with more $b$-tagged jets have a higher fraction of $t\bar{t}$ events. This characterization is used to constrain the various background components by imposing self-consistency among the channels. The reconstruction of semileptonic $t\bar{t}$ candidates relies on a $\chi^2$ variable built by enforcing kinematic consistency (within uncertainties) with the $t\bar{t}$ hypothesis, imposing constraints on the reconstructed $W$ and top candidates. In the semileptonic boosted regime we follow the techniques of Ref. [23], and allow candidates with more than one parton merged into a single jet.

For the boosted all-hadronic analysis, data were recorded with a trigger requiring the scalar sum of the transverse momenta of reconstructed AK5 jets to be greater than 750 GeV. In the offline analysis selection, we require two CA8 jets, each with $p_T > 400$ GeV.

The reconstruction of the boosted all-hadronic analysis relies on “top-tagging” techniques similar to those used in the previous analysis [20]. This algorithm [32], aiming to identify the top quark decay products within CA8 jets, reverses the jet-clustering sequence by iteratively separating the jet into subjects until three or four subjects with sufficient $p_T$ are found. The algorithm is validated on a sample of $t\bar{t}$ events selected by requiring one muon and
additional jets [33]. The reconstructed single jet top quark mass is found to be consistent with the expectation from simulated events, as is the disubjet $W$ mass, obtained from the minimum mass pairing of the leading three subjets. The two selected top-tagged jets are then required to be back to back, with $|\Delta \phi| > \pi/2$ and $|\Delta y| < 1.0$, to suppress nontop multijet (NTMJ) backgrounds.

SM top quark production is modeled with the next-to-leading-order (NLO) generator POWHEG (v1.0) [34], interfaced with PYTHIA 6 (v6.2.24) [35] for parton showering with tune Z2* [36]. MADGRAPH (v5.1.1) [37] interfaced to PYTHIA 6 is used for simulating $W$ and $Z$ boson production in association with jets. Diboson processes are generated with PYTHIA 6 to compute both the matrix element and background estimate is taken for all SM components together directly from the data, with the SM $t\bar{t}$ component the dominant one. The number of signal events is extracted from a binned maximum likelihood fit to the $M_{t\bar{t}}$ distribution, assuming a smoothly falling probability density function (pdf) for the SM backgrounds and a parametrization of the signal pdf based on a Breit-Wigner shape. Only events with $M_{t\bar{t}} > 550$ GeV are considered; below this value the SM backgrounds are not described by a smoothly falling pdf.

The MADGRAPH -PYTHIA 6 combination is also used to generate signal Monte Carlo (MC) simulation events for limit setting, including high-mass SM-like $Z'$ resonances with $\Gamma'_{Z}/M_{Z'} = 1\%$ and $\Gamma'_{Z}/M_{Z'} = 10\%$, where $\Gamma'_{Z}$ is the width of the resonance, and $M_{Z'}$ is the mass. This relative width can be compared to the detector resolution of about 10% for a $t\bar{t}$ resonance mass. Hence, limits set for the $Z'$ with a width of 1% would apply to a larger class of models in which the resonance width is below the experimental resolution. The MADGRAPH -PYTHIA 6 combination is also used to generate a simplified model of a spin-zero resonance produced via gluon fusion through its couplings to top quarks, with the SM interference effects neglected in the model. A Kaluza-Klein excitation of a gluon with a width of approximately 15%–20% [7] is generated with PYTHIA 6 (v1.5.3) [38].

<table>
<thead>
<tr>
<th>TABLE I. Constraints used in the likelihood maximization. The $M_{t\bar{t}}$ distributions of the boosted channels are combined into a single joint likelihood, imposing consistency of the various background and signal components.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution semileptonic</td>
</tr>
<tr>
<td>Constraints on normalization</td>
</tr>
<tr>
<td>Luminosity [41]</td>
</tr>
<tr>
<td>Pileup</td>
</tr>
<tr>
<td>$t\bar{t}$ [42]</td>
</tr>
<tr>
<td>Parton distribution functions [43]</td>
</tr>
<tr>
<td>Single top</td>
</tr>
<tr>
<td>$W +$ light-flavor jets</td>
</tr>
<tr>
<td>$W +$ heavy-flavor jets</td>
</tr>
<tr>
<td>$Z +$ jets</td>
</tr>
<tr>
<td>Lepton selection</td>
</tr>
<tr>
<td>Top-tagging efficiency</td>
</tr>
<tr>
<td>Constraints on shape</td>
</tr>
<tr>
<td>$t\bar{t}$ renormalization, factorization, and matching scales</td>
</tr>
<tr>
<td>Jet energy scale</td>
</tr>
<tr>
<td>Jet energy resolution</td>
</tr>
<tr>
<td>$b$-tagging efficiency</td>
</tr>
<tr>
<td>$b$-tagging mis-ID</td>
</tr>
<tr>
<td>Top-tagging mis-ID</td>
</tr>
<tr>
<td>Signal and background pdf</td>
</tr>
</tbody>
</table>
similar to the signal region. It is then used to weight events in the signal region. Furthermore, the efficiency for identifying true top quark jets is corrected in the signal MC simulations using measurements in a signal-depleted sideband region containing events with one isolated muon and additional jets. It is found that the efficiencies in data and MC simulations agree, having a ratio of $93 \pm 5\%$. The methods described above were validated using simulated samples and it was verified that signal contamination was minimal in the signal-depleted regions.

In the likelihood maximization, systematic uncertainties are treated as nuisance parameters. Those that are common among the channels are treated as 100% correlated, while those that are channel-specific are treated as uncorrelated.

### TABLE II. Number of expected and observed events in the boosted analyses.

<table>
<thead>
<tr>
<th>Sample</th>
<th>$N_{\text{b-tag}} = 0$</th>
<th>$N_{\text{b-tag}} \geq 1$</th>
<th>$M_{t}\geq 1 \text{ TeV}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$t\bar{t}$</td>
<td>5440 ± 520</td>
<td>9090 ± 870</td>
<td>510 ± 90</td>
</tr>
<tr>
<td>NTMJ</td>
<td>··</td>
<td>··</td>
<td>6600 ± 200</td>
</tr>
<tr>
<td>Others</td>
<td>5880 ± 820</td>
<td>1070 ± 380</td>
<td>··</td>
</tr>
<tr>
<td>Total SM</td>
<td>11320 ± 1300</td>
<td>10160 ± 1300</td>
<td>7110 ± 410</td>
</tr>
<tr>
<td>Data</td>
<td>10305</td>
<td>10159</td>
<td>6887</td>
</tr>
</tbody>
</table>

The normalizations of the backgrounds are allowed to vary within log-normal constraints in the maximization of the joint likelihood. The shapes of the backgrounds are also allowed to vary within their uncertainties. The shapes and normalizations also account for systematic variations due to efficiency and misidentification rates. The constraints used in the joint likelihood maximization are listed in Table I.

The event yields from the various background components and data are shown in Table II. The yields of the simulated samples are quoted after the likelihood maximization procedure, and the individual background uncertainties include only the uncertainty in the individual normalization. The total SM contribution includes all uncertainties, including the correlations not quoted in the individual components. Figure 1 shows the $M_{t\bar{t}}$ distributions for all channels along with the expectation from a $Z'$ signal.

In all cases, the data are well described by the SM-only background hypothesis. The absence of a signal in the $M_{t\bar{t}}$ distribution is quantified by deriving Bayesian upper limits on the signal cross section times branching fraction at 95% confidence level (C.L.), using pseudoexperiments. The resolved semileptonic analysis has some overlapping phase space with the boosted semileptonic analysis, and there is a transition point ($\sim 1 \text{ TeV}$) where the expected sensitivities of the boosted and resolved analyses are equal, above which the boosted analysis result is quoted, and below which the resolved analysis result is quoted.

Figure 2 shows the expected and observed limits for a narrow resonance, as a function of the invariant mass of the resonance. The specific example shown in Fig. 2 and given by the dashed line refers to a topcolor $Z'$ with $\Gamma_{Z'}/M_{Z'} = 1.2\%$ based on predictions from Ref. [10]. The cross section limits for this case are obtained from the MC models with $\Gamma_{Z'}/M_{Z'} = 1.0\%$, scaled by the ratio of theoretical...
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); MES (Bulgaria); CERN; CAS, MoST, and NSFC (China); COLCIENCIAS (Colombia); MSES (Croatia); RPF (Cyprus); MoER, SF0690030s09 and ERDF (Estonia); Academy of Finland, MED, 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); MBIE (New Zealand); PAEC (Pakistan); MSHE and NSC (Poland); FCT (Portugal); JINR (Dubna); MON, RosAtom, RAS, and RFBR (Russia); MESTD (Serbia); SEIDI and CPAN (Spain); Swiss Funding Agencies (Switzerland); NSC (Taipei); ThEPCenter, IPST, STAR, and NSTDA (Thailand); TUBITAK and TAEK (Turkey); NASU (Ukraine); STFC (United Kingdom); DOE and NSF (USA).


TABLE III. 95% C.L. lower limits on the masses of new particles in specific models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Observed limit</th>
<th>Expected limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z'$, $\Gamma_{Z'}/M_{Z'} = 1.2%$</td>
<td>2.1 TeV</td>
<td>2.1 TeV</td>
</tr>
<tr>
<td>$Z'$, $\Gamma_{Z'}/M_{Z'} = 10%$</td>
<td>2.7 TeV</td>
<td>2.6 TeV</td>
</tr>
<tr>
<td>RS KK gluon</td>
<td>2.5 TeV</td>
<td>2.4 TeV</td>
</tr>
</tbody>
</table>

cross sections. This scaling is done to compare to theoretical results and previous measurements. As the cross section calculation is available for this model at LO only, the predictions are multiplied by a factor of 1.3 to account for higher-order effects [44]. The vertical dash-dotted line indicates the transition between the resolved and boosted analyses. Table III shows additional model-specific limits. The combination of the semileptonic and all-hadronic boosted analyses improves the expected cross section limits at 2 TeV by ~25%. Compared to the results of previous analyses [20–23] for specific models [7,10], the lower limits on the masses of these resonances have been improved by several hundred GeV. For the semileptonic resolved analysis, assuming a spin-zero resonance with narrow width, produced via gluon fusion with no interference with the SM background, the cross section limits are 0.8 pb and 0.3 pb for a spin-zero resonance of mass 500 and 750 GeV, respectively. These are the first limits at CMS for heavy Higgs-like particles decaying into $t\bar{t}$.

In addition to investigating possible resonant structures in the $M_{t\bar{t}}$ spectrum, the presence of new physics that causes a nonresonant enhancement of the $M_{t\bar{t}}$ spectrum is also tested. The boosted all-hadronic analysis is used to set limits on such new production for events with $M_{t\bar{t}} > 1$ TeV, since the NTMJ background can be predicted entirely from data. The limit is expressed as a ratio of the total SM + BSM $t\bar{t}$ cross section to the SM-only cross section ($S_t$, as defined in Ref. [20]). The efficiency to select SM $t\bar{t}$ events with $M_{t\bar{t}} > 1$ TeV is $(3.4 \pm 1.7) \times 10^{-4}$. We find $S_t < 1.2$ at the 95% C.L., with a credible interval of 1.1–2.0 at 68% C.L., a factor of 2 improvement over the previously published limits [20].

In summary, we have performed searches for anomalous $t\bar{t}$ production using events in the semileptonic and all-hadronic topologies. In addition to new limits on nonresonant enhancements to top quark production, limits are set on the production cross section times branching fraction for several resonance hypotheses, for resonances in the mass range 0.5–3.0 TeV.

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.


PRL 111, 211804 (2013) PHYSICAL REVIEW LETTERS week ending 22 NOVEMBER 2013

126 University of Colorado at Boulder, Boulder, Colorado, USA
127 Cornell University, Ithaca, New York, USA
128 Fairfield University, Fairfield, Connecticut, USA
129 Fermi National Accelerator Laboratory, Batavia, Illinois, USA
130 University of Florida, Gainesville, Florida, USA
131 Florida International University, Miami, Florida, USA
132 Florida State University, Tallahassee, Florida, USA
133 Florida Institute of Technology, Melbourne, Florida, USA
134 University of Illinois at Chicago (UIC), Chicago, Illinois, USA
135 The University of Iowa, Iowa City, Iowa, USA
136 Johns Hopkins University, Baltimore, Maryland, USA
137 The University of Kansas, Lawrence, Kansas, USA
138 Kansas State University, Manhattan, Kansas, USA
139 Lawrence Livermore National Laboratory, Livermore, California, USA
140 University of Maryland, College Park, Maryland, USA
141 Massachusetts Institute of Technology, Cambridge, Massachusetts, USA
142 University of Minnesota, Minneapolis, Minnesota, USA
143 University of Mississippi, Oxford, Mississippi, USA
144 University of Nebraska-Lincoln, Lincoln, Nebraska, USA
145 State University of New York at Buffalo, Buffalo, New York, USA
146 Northeastern University, Boston, Massachusetts, USA
147 Northwestern University, Evanston, Illinois, USA
148 University of Notre Dame, Notre Dame, Indiana, USA
149 The Ohio State University, Columbus, Ohio, USA
150 Princeton University, Princeton, New Jersey, USA
151 University of Puerto Rico, Mayaguez, Puerto Rico
152 Purdue University, West Lafayette, Indiana, USA
153 Purdue University Calumet, Hammond, Indiana, USA
154 Rice University, Houston, Texas, USA
155 University of Rochester, Rochester, New York, USA
156 The Rockefeller University, New York, New York, USA
157 Rutgers, The State University of New Jersey, Piscataway, New Jersey, USA
158 University of Tennessee, Knoxville, Tennessee, USA
159 Texas A&M University, College Station, Texas, USA
160 Texas Tech University, Lubbock, Texas, USA
161 Vanderbilt University, Nashville, Tennessee, USA
162 University of Virginia, Charlottesville, Virginia, USA
163 Wayne State University, Detroit, Michigan, USA
164 University of Wisconsin, Madison, Wisconsin, USA

a Deceased.
b Also at Vienna University of Technology, Vienna, Austria.
c Also at CERN, European Organization for Nuclear Research, Geneva, Switzerland.
d Also at Institut Pluridisciplinaire Hubert Curien, Université de Strasbourg, Université de Haute Alsace Mulhouse, CNRS/IN2P3, Strasbourg, France.
e Also at National Institute of Chemical Physics and Biophysics, Tallinn, Estonia.
f Also at Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia.
g Also at Universidade Estadual de Campinas, Campinas, Brazil.
h Also at California Institute of Technology, Pasadena, CA, USA.
i Also at Laboratoire Leprince-Ringuet, Ecole Polytechnique, IN2P3-CNRS, Palaiseau, France.
j Also at Zewail City of Science and Technology, Zewail, Egypt.
k Also at Suez Canal University, Suez, Egypt.
l Also at Cairo University, Cairo, Egypt.
m Also at Fayoum University, El-Fayoum, Egypt.
n Also at British University in Egypt, Cairo, Egypt.
o Also at Ain Shams University, Cairo, Egypt.
p Also at National Centre for Nuclear Research, Swierk, Poland.
q Also at Université de Haute Alsace, Mulhouse, France.
r Also at Joint Institute for Nuclear Research, Dubna, Russia.