

Measurement of the top pair invariant mass distribution and search for New Physics using the CMS experiment

Freya Blekman*

On behalf of the CMS Collaboration

Interuniversity Institute for High Energies (ULB-VUB)

Vrije Universiteit Brussel, Belgium

E-mail: freya.blekman@vub.ac.be

Searches are presented for a massive particle decaying to a $t\bar{t}$ pair, using the full CMS 2011 dataset of proton-proton collisions at a center of mass energy of 7 TeV, corresponding to an integrated luminosity of up to 5 fb^{-1} . The searches focus on resonant particles decaying to a top quark pair, generally referred to as a Z' . Top quark decays to electrons, muons and hadrons are considered, leading to searches in final states using fully hadronic final states and final states with one and two leptons. The searches for Z' resonances look for resonant particles that are sufficiently massive to produce highly Lorentz-boosted top quarks, which yield collimated decay products that may be partially or fully merged. 95% confidence level upper limits up to $2.0 \text{ TeV}/c^2$ are set on the product of the production cross section and branching fraction for topcolor Z' modelled for several widths, as well as for Randall–Sundrum Kaluza–Klein gluons, which are excluded to values up to $1.5 \text{ TeV}/c^2$.

*36th International Conference on High Energy Physics,
July 4-11, 2012
Melbourne, Australia*

*Speaker.

1. Introduction

The top quark is the heaviest known fermion, and plays a central role in several theories beyond the standard model (SM). Some theories beyond the SM predict the existence of heavy resonances that manifest themselves as a resonant component to SM $t\bar{t}$ production. Examples of such resonances include models with massive color-singlet Z-like bosons in extended gauge theories, colorons, axigluons, models with a pseudo-scalar Higgs boson that couples strongly to top quarks and models predicting Kaluza-Klein excitations of gluons or gravitons in extensions of extra-dimensions Randall–Sundrum models[1, 2].

In this article, three searches for the production of heavy resonances decaying into $t\bar{t}$ are presented, using data collected by the CMS experiment in pp collisions at $\sqrt{s}=7$ TeV at the Large Hadron Collider. Using samples corresponding to an integrated luminosity of 4.4-5.0 fb⁻¹, we focus on three common $t\bar{t}$ decay modes, which are traditionally classified by the decay products of the W bosons from the top quark decay as in the SM the $\text{BR}(t \rightarrow bW) \approx 1$ [3, 4, 5]. After an introduction of the shared experimental aspects of the analyses, the search for a resonance in the invariant mass spectrum of $t\bar{t}$ pairs decaying fully hadronically to jets is presented in Sec. 3. We then proceed by considering the final states where one lepton is present in Sec. 4, and conclude with a search in the dilepton final state in Sec. 5.

2. CMS experiment and event samples

The examination of the $t\bar{t}$ invariant mass spectrum is performed using the Compact Muon Solenoid (CMS) detector at the Large Hadron Collider. The CMS detector uses a polar coordinate system with the polar angle θ defined in the direction of the counter-clockwise proton beam. The pseudorapidity η is defined as $\eta = -\ln \tan(\theta/2)$. The CMS detector is described in detail elsewhere[6].

The data analysed for the analyses summarized in this document were collected with the two-tier trigger system, selecting on the presence of two lepton in the case of dilepton analyses, on the presence of a single lepton in combination with a number of jets for the lepton+jets analyses and a hadronic trigger selecting on the level of jet activity in the case of the all-hadronic analyses. The integrated luminosity is equivalent to 5.0 fb⁻¹ unless otherwise noted[7].

Offline, CMS uses a particle-flow based event reconstruction, which combines information from each subdetector, including charged particle tracks and energy deposits from the electromagnetic and hadronic calorimeter and information from the muon system to reconstruct all particles in the event[8]. The particles selected by the particle flow algorithm are then passed through the standard CMS identification procedure for electron and muon candidates. Hadronic candidates are clustered to jets using the anti-kT jet algorithm[13] with a distance parameter of $R=0.5$, except for the fully hadronic analysis where the Cambridge-Aachen (CA) jet reconstruction algorithm is used[14], with a distance parameter of $R=0.8$. R that is defined in $\eta - \phi$ space as $R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$. All jets, electrons and muons are required to satisfy quality criteria and are fully corrected for remaining small differences between the behaviour of the reconstruction algorithms in simulation and data. Due to the track-based nature of the particle flow algorithm, only particle flow objects are considered within the tracking fiducial region $|\eta| < 2.5$. The missing transverse momentum E_T^{miss} is

reconstructed taking the negative vector sum of all particle flow objects in the event, taking into account improved knowledge of clustered objects such as jets. Only events with at least one primary vertex are considered and only charged particles identified as originating from the leading primary vertex are considered in the reconstruction, drastically reducing the effect of multiple interactions per beam crossing.

There are several simulated samples used. The continuum SM $t\bar{t}$ background, and Standard Model processes such as Z and W boson production are simulated with the MADGRAPH[9] generator, using the PYTHIA generator[10] for the hadronization modelling. Diboson backgrounds, important in the dilepton analysis, are modeled using PYTHIA for the production and decay. For the signal model, resonance widths of 1%, 3% and 10% were considered. The MADGRAPH generator is used to model a generic Z' that has SM-like fermion couplings. However, only the decay of this Z' to $t\bar{t}$ is simulated. Randall–Sundrum KK gluons are generated using PYTHIA. The CTEQ6L parton density functions are used [11]. The detector response is simulated using the CMS detector simulation based on GEANT4[12]. The simulation is corrected such that the primary vertex multiplicity is the same as observed in data. More details on the full selection procedure and simulation procedures can be obtained from Ref. [3, 4, 5].

3. Search for $Z' \rightarrow t\bar{t}$ production in the fully hadronic channel

The analysis designed for cases where both top quarks decay to hadrons is described in Ref. [3]. The largest background in this search is the background from multijet production. This background is modelled using data, where parametrizations are derived using a sample selected with an inverted top identification cut.

For this final state, two different topologies are considered: both top quarks are boosted and have each merged in a single jet, resulting in two highly energetic jets in the final state, and one top quark was fully merged into one jet while the other top quark candidate consists of a jet containing the hadronic W decay and a separate b jet. The former is referred to as the “1+1” topology, while the later is named the “1+2” event topology.

The 1+1 events are required to have at least two top-quark candidates, each reconstructed with transverse momentum $p_T > 350 \text{ GeV}/c$. In the case of more than two jets, the two leading jets are considered. The top tagging algorithm then considers the decomposition of a top quark candidate into subjects, by reversing the final steps of the CA jet clustering. At least three subjects are required in each jet, and mass constraints on the minimum pairwise invariant mass are required to be consistent with a W decay while the three-subjet system is required to be consistent with the mass of a top quark $m_t \approx 175 \text{ GeV}/c^2$ ($140 < m_{\text{jet}} < 250 \text{ GeV}/c^2$).

The 1+2 events are required to have exactly one hemisphere containing a top-tagged candidate with the selection also applied for 1+1 candidates. In the opposite hemisphere there must be at least two jets, one identified as a W-jet candidate with $p_T > 200 \text{ GeV}/c$ and the other is assumed to come from a b quark (although no identification algorithm is used) with $p_T > 30 \text{ GeV}/c$.

The $M_{t\bar{t}}$ distribution is then fit assuming a resonance hypothesis, using a likelihood fit that takes into account the shapes and uncertainties of the signal and background. As no excess is observed over the SM production we proceed by setting limits on resonant production of new particles in this distribution, using generic Z' hypotheses with two widths and KK gluon production as the

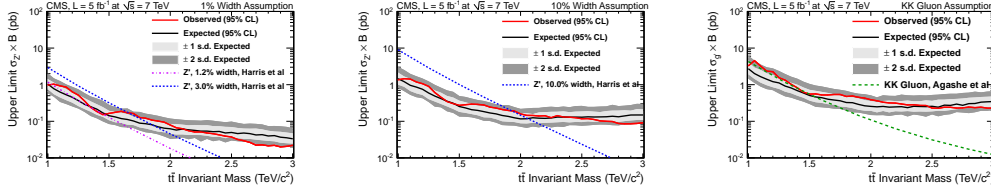


Figure 1: All hadronic limits: The 95% CL upper limits on the product of the production cross section $\sigma_{Z'}$ and the branching fraction B of hypothesized resonances that decay into $t\bar{t}$ as a function of the invariant mass of the resonance. The Z' production with $\Gamma_{Z'}/m_{Z'} = 1.2\%$ and 10% , and the branching fraction B of Kaluza–Klein excitation of gluon production are compared to theory prediction. The vertical dashed line indicates the transition between the threshold and the boosted analyses, chosen based on the sensitivity of the expected limit. The ± 1 and ± 2 s.d. excursions from the expected limits are also shown.

benchmark models. The observed cross section limits are compared to expected cross section limits from Refs. [1, 2]. The upper limits at 95% confidence level (CL) are shown in Fig. 1. Model-dependent mass limits to exclude heavy resonances are listed in Table 1.

4. Search for $Z' \rightarrow t\bar{t}$ production in the lepton+jets channel

To study a very wide kinematic range of top pair production, two complementary strategies are pursued in the case where only one of the top quarks from the $t\bar{t}$ pair decays to a lepton. In the case where the invariant mass is in the range 0.5 - 1 TeV/c^2 , the analysis is optimized to identify top quarks produced with small boost in the detector frame and hence with well-separated decay products. In this threshold region we expect a final state to contain one isolated lepton, four jets (including two b jets) and E_T^{miss} . When the $t\bar{t}$ invariant mass distribution is examined above 1 TeV/c^2 , the Lorentz-boosted top quarks will yield collimated decay products that are partially or fully overlapping. As a consequence, the products of the hadronically decaying W boson from one of the top quarks might be reconstructed as fewer than three jets, and the leptons might not be isolated. In the boosted regime, the search thus selects events containing exactly one, not necessarily isolated, electron or muon and at least two jets.

In the threshold analysis, events are required to contain at least three anti-kT jets with $p_T > 50$ GeV/c , a leading jet with $p_T > 70$ GeV/c , and additional jets have to pass a $p_T > 30$ GeV/c requirement. E_T^{miss} is required to be above 20 GeV/c . The choices of neutrino and jet association are then solved by forming a χ^2 using information from the leptonic and hadronic top quark mass, the hadronic W mass, the p_T of the $t\bar{t}$ system and the ratio of the p_T of the selected jets over the p_T of all jets in the event. The χ^2 is calculated for each possible combination, and b-tagged jets may only be associated with b quarks. The combination with the smallest value of χ^2 is chosen and used for the examination of the $t\bar{t}$ invariant mass distribution.

In the boosted analysis, events are selected when they contain one muon(electron) with $p_T > 42(50)$ GeV/c and $|\eta| < 2.1(2.5)$. At least two jets are required with $p_T > 50$ GeV/c , of which one jet is required to be above $250(15)$ GeV/c in the muon(electron) channel. No isolation requirements is applied to the leptons, and multijet production is instead rejected by requiring that the $\Delta R(\text{lepton, closest jet}) > 0.5$ or $p_T^{\text{rel}}(\text{lepton, closest jet}) > 25$ GeV/c . Here, p_T^{rel} is the magni-

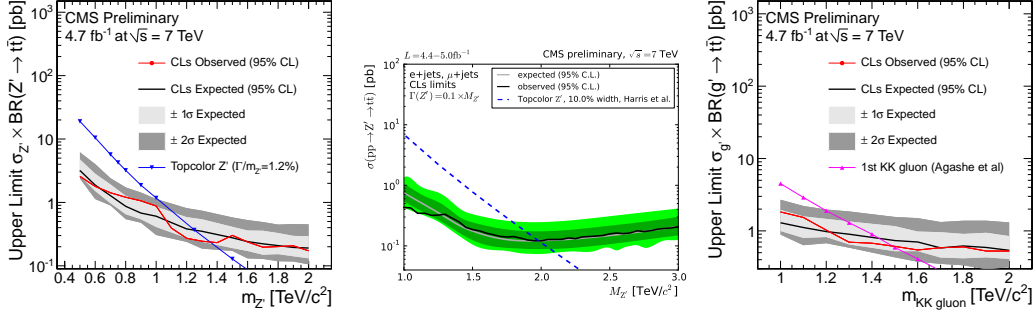


Figure 2: Lepton+jets limits: The 95% CL upper limits on the product of the production cross section $\sigma_{Z'}$ and the branching fraction B of hypothesized resonances that decay into $t\bar{t}$ as a function of the invariant mass of the resonance. The Z' production with $\Gamma_{Z'}/m_{Z'} = 1.2\%$ and 10% are shown for the threshold (left) and boosted (center) analyses, and are compared to theory prediction. Limits on KK gluon production from the threshold analysis are also presented (right). The ± 1 and ± 2 s.d. excursions from the expected limits are also shown.

tude of the lepton momentum orthogonal to the jet axis. In addition there is a requirement on $E_T^{miss} + p_T^{lepton} > 150 \text{ GeV}/c$. In the electron channel further selection is applied to further reduce the multijet background. Again, a χ^2 is constructed to determine the mass hypothesis of the event, but in the boosted hypothesis it might be possible that two jets have merged, so only the leptonic W boson information is combined with one jet, while the remaining jets are used as an estimate of $m_{t\bar{t}}$.

Both analyses separate the event samples in multiple classes determined by lepton flavor, jet and b-tag multiplicities, to constrain the various backgrounds in a simultaneous likelihood fit of all bins. Data driven background estimates are used for the not well modelled backgrounds such as multijet production and good agreement is observed in all cases, so limits are set on the presence of a resonance in the data. All systematic uncertainties applicable to $t\bar{t}$ cross section measurements are taken into account as nuisance parameters on either the yield of signal, background or on the shape of the signal or background, and the cross section limits on Z' production are shown in Fig. 2. Model-dependent mass limits to exclude heavy resonances are listed in Table 1.

5. Search for $Z' \rightarrow t\bar{t}$ production in the dilepton channel

In the search for Z' production in fully leptonic decays, events are required to contain at least two jets and two opposite sign leptons, where only muons and electrons are considered. At least one of the jets is required to be tagged as a b jet using a b-tagging algorithm that reconstructs the secondary vertex of B meson decays. In order to reduce the background from low-mass dilepton resonances and Z-boson production, events are rejected if the dilepton invariant mass $M_{\ell\ell} < 12 \text{ GeV}/c^2$, and a veto is applied $76 < M_{\ell\ell} < 106 \text{ GeV}/c^2$ in the ee and $\mu\mu$ channels. In order to further reduce the background from DY and multijet production, a requirement of $E_T^{miss} > 30 \text{ GeV}/c$ is applied in these channels. Remaining SM backgrounds are $t\bar{t}$ production, Z and W boson production, single top production and diboson production. The DY and multijet backgrounds are estimated from data using the Z resonance sideband and isolation inversion techniques.

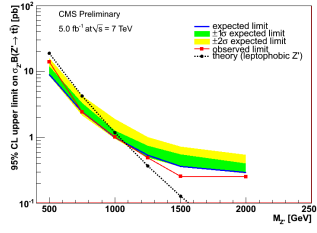


Figure 3: Dilepton limit: The 95% CL upper limit on the product of the production cross section $\sigma_{Z'}$ and the branching fraction $B(Z' \rightarrow t\bar{t})$ as a function of the invariant mass of the resonance. The Z' production with $\Gamma_{Z'}/m_{Z'} = 1.2\%$ is compared to theory prediction. The ± 1 and ± 2 s.d. excursions from the expected limits are also shown.

Table 1: Summary of all mass exclusion limits, in TeV/c^2 , on $M_{Z'}$ for widths $\Gamma_{Z'}/M_{Z'}=0.01$ (or 0.012) and 0.1, and for $M_{G_{KK}}$ obtained from CMS analyses searching for new physics in the $t\bar{t}$ invariant mass spectrum.

analysis	$M_{Z'}$ (1-1.2% width)	$M_{Z'}$ (10% width)	KK gluon
dilepton	> 1.1	\times	\times
lepton+jets (thresh.)	> 1.3	> 1.7	1.4
lepton+jets (boost)	> 1.55	> 1.98	\times
all hadronic	> 1.5	> 2.0	excl: 1.4 – 1.5

The $t\bar{t}$ invariant mass is constructed using the four-vectors of the two leading leptons, the two leading jets and the E_T^{miss} . The longitudinal momenta of the two neutrinos were observed to be small and are set to zero as an experimental reconstruction of these is difficult. A multivariate analysis based on Bayesian neural networks (BNN) [15] using as input 17 variables such as the P_T , η of the jets and leptons, the number of b-tagged jets, the E_T^{miss} and the angles between the objects are used to discriminate between the SM background and a possible Z' signal. The BNN is trained for a $M_{Z'}$ of 750(1000) GeV/c^2 for the 1.2(3)% width hypothesis, and a separate discriminant is constructed for the ee , $e\mu$ and $\mu\mu$ final states. The shape of the BNN discriminant is used for the final limit calculation, and all correlations between the different samples and channels are included.

The shape and normalization changes of signal and background are affected by systematic uncertainties, which are propagated into the limit calculation. With no excess observed, upper limits on $\sigma_{Z'} \cdot B(Z' \rightarrow t\bar{t})$ are set at 95% DL. Using the BNN distribution improves the expected limit by 29% compared to using the invariant mass distribution. The resulting expected and observed limits are shown in Fig. 3. Model-dependent limits on the production of a leptophobic Z' for two widths are listed in Table 1.

6. Summary and Conclusion

Using the full dataset collected in 2011, the CMS experiment has examined the invariant mass spectrum of $t\bar{t}$ production and has not observed any deviation from the Standard Model prediction. By converting the observed cross sections to an 95% CL exclusion limit, the CMS experiment excludes generic Z' resonances up to $M_{Z'}$ values of 1.5-2.0 TeV/c^2 , depending on the width of the resonance. Production of Kaluza–Klein gluon production is also excluded to 95% CL for $M_{G_{KK}}$

values up to $1.82 \text{ TeV}/c^2$. The exclusion limits for various models are summarized in Table 1. The dilepton channel still remains competitive using multivariate techniques, and sets the strongest limit when Z' models with a width of the order of 1% are considered. In general, the lepton+jets and hadronic results produce the most stringent mass limits for narrow resonances. The mass limits set for 1% width Z' production are of order $1.1\text{-}1.5 \text{ GeV}/c^2$ at 95% CL, depending on the final state, while the lepton+jets and all hadronic analyses are competitive in setting a limit of around $2 \text{ TeV}/c^2$ for 10% width Z' production. This is a remarkable achievement in an environment where substructure techniques or non-orthodox lepton reconstruction are necessary to isolate a pure $t\bar{t}$ signal.

References

- [1] C. T. Hill, Phys. Lett. B 345 (1995) 483–489, doi:10.1016/0370-2693(94)01660-5, arXiv:hep-ph/9411426.
- [2] K. Agashe et.al, Phys. Rev. D 77 (2008), 015003, doi:10.1103/PhysRevD.77.015003, arXiv:hep-ph/0612015.
- [3] CMS Collaboration, JHEP **1209** (2012) 029 [arXiv:1204.2488 [hep-ex]].
- [4] CMS Collaboration, CMS Physics Analysis Summary, CMS-PAS-TOP-11-009 (2012).
CMS Collaboration, CMS Physics Analysis Summary, CMS-PAS-EXO-11-093 (2012).
- [5] CMS Collaboration, CMS Physics Analysis Summary, CMS-PAS-TOP-11-010 (2012)
- [6] CMS Collaboration, JINST 03 (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.
- [7] CMS Collaboration, CMS Physics Analysis Summary, CMS-PAS-SMP-12-008 (2012).
- [8] CMS Collaboration, CMS Physics Analysis Summary, CMS-PAS-PFT-10-002 (2010).
- [9] J. Alwall et al., JHEP 06 (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522.
- [10] T.Sjöstrand, S.Mrenna, and P.Skands, JHEP05 (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175
- [11] J. Pumplin et al., JHEP 07 (2002) 012, doi:10.1088/1126-6708/2002/07/012, arXiv:hep-ph/0201195.
- [12] S. Agostinelli et al., Nucl. Instrum. Methods A 506 (2003) 250, doi:10.1016/S0168-9002(03)01368-8.
- [13] M. Cacciari, G. P. Salam, and G. Soyez, JHEP 04 (2008) 063, arXiv:0802.1189.
- [14] M. Wobisch and T. Wengler, DESY-PROC 02 (1998) arXiv:hep-ph/9907280.
Y. L. Dokshitzer, G. D. Leder, S. Moretti et al., JHEP 08 (1997) 001, doi:10.1088/1126-6708/1997/08/001, arXiv:hep-ph/9707323.
M. Cacciari, G. P. Salam, and G. Soyez, (2011). arXiv:1111.6097.
- [15] R. M. Neal, Springer-Verlag, New York, 1996.