

Measurements of the associated production of top quark pairs with bosons or other top quarks

Javier Cuevas on behalf of the CMS collaboration*

Universidad de Oviedo

E-mail: Javier.Cuevas@cern.ch

A comprehensive set of measurements of top quark pair production in association with electroweak bosons (W , Z or γ) is presented at different centre-of-mass energies. The results are compared to theoretical predictions and re-interpreted as searches for new physics that induce deviations from the standard model expectations using an effective field theory approach. The status of the search for double- $t\bar{t}$ pair production combining different final states is also reported.

*The European Physical Society Conference on High Energy Physics
5-12 July
Venice, Italy*

*Speaker.

1. Introduction

Measurements of the $t\bar{t}Z$ and $t\bar{t}\gamma$ processes provide a means of directly determining top quark couplings to bosons. $t\bar{t}Z$ and $t\bar{t}W$ represent one of the most massive final states that are currently observable. These processes form the dominant backgrounds to several searches for New Physics as well as to the measurements of standard model (SM) processes such as top quark pair production in association with the Higgs boson $t\bar{t}H$. The $t\bar{t}W$ process is among other things, a SM source of same-sign dilepton events, which are a signature of many models of physics beyond the SM. It is also interesting per-se as it tags the process $qq \rightarrow t\bar{t}X$ at the LHC.

$t\bar{t}H$ where the Higgs boson decays to $b\bar{b}$, leads to a $t\bar{t}b\bar{b}$ final state. This final state, which has not been observed yet, has an irreducible nonresonant background from the production of a top quark pair in association with a b quark pair produced via gluon splitting ($g \rightarrow b\bar{b}$).

Observing $t\bar{t}t\bar{t}$ production consistent with SM predictions would provide a valuable test of higher-order perturbative QCD calculations. In addition, many models of physics beyond the SM predict an increase in the $t\bar{t}t\bar{t}$ cross section owing either to the presence of hypothetical particles that decay into top quarks or to modified couplings. In this document, I present a review of the most recent results of $t\bar{t}$ production in association with W ($t\bar{t}W$), Z ($t\bar{t}Z$), γ ($t\bar{t}\gamma$), b quarks ($t\bar{t}b\bar{b}$), and search for four top quarks ($t\bar{t}t\bar{t}$) production in different final states. The measurements are performed by CMS [1] using data from the LHC at a centre-of-mass energies of 8 and 13 TeV.

2. Measurement of the $t\bar{t}W$ and $t\bar{t}Z$ cross-sections at $\sqrt{s} = 13\text{TeV}$

$t\bar{t}W$ and $t\bar{t}Z$ cross-section measurement is performed using events in which at least one of the W bosons originating from a top quark decays to a lepton and a neutrino, and the associated W boson decays to a lepton and a neutrino, or the Z boson decays to two charged leptons, where electron or muon is only considered as a lepton [2].

For $t\bar{t}W$ the two leptons are required to have the same electric charge, then only one third of the signal produced in the dilepton final state is retained. However, this selection significantly improves the signal- to-background ratio, as same-sign lepton pairs are produced in SM processes with relatively small cross sections. The main backgrounds to this analysis originate from misreconstruction effects: misidentification of leptons from heavy-quark decays and mismeasurement of the charge of one of the leptons in events with an opposite-charged lepton pair. In order to distinguish these backgrounds from the signal, a multivariate analysis (MVA) is developed. The MVA is trained using events from $t\bar{t}W$ signal and background samples which are selected to contain at least two jets of which one or more are b jets. A boosted decision tree (BDT) classifier is chosen as the MVA technique trained on several kinematic observables, among others, the highest- p_T (leading) and second- highest- p_T (subleading) jet p_T , and the ΔR between the trailing lepton and the nearest selected jet. A cut on BDT value is applied to suppress the background due to nonprompt leptons. The number of jets and b-tagged jets are used to form exclusive event categories to maximize the signal significance. Given the charge asymmetric production of the signal ($t\bar{t}W^+$ vs $t\bar{t}W^-$) these categories are further split into two separate sets according to the charge of the leptons: $++$ or $--$.

The $t\bar{t}Z$ process is measured in channels with three or four leptons, with a pair of same-flavour leptons of opposite-charge, with an invariant mass close to the Z boson mass. In the three lepton

channel the leading, sub-leading and trailing transverse momenta have to pass 40, 20 and 10 GeV requirement respectively. Nine categories are formed using events with $N_{jets} = 2, 3$ and ≥ 4 , where each jet multiplicity bin is further split according to the b jet multiplicity, $N_{bjets} = 0, 1$ and ≥ 2 . In the four lepton category, events are characterized by the presence of two b jets, missing transverse momentum, and four leptons, two of which form a same flavor opposite charge pair consistent with the Z boson mass. For this channel, in addition to nonprompt leptons, WZ , $t(t)X$ and rare SM processes are also important, and are evaluated by Monte Carlo (MC) simulation and validated in control regions in data. Uncertainties on the lepton reconstruction, b-tagging and trigger efficiency have the largest effect on the $t\bar{t}W$ and $t\bar{t}Z$ cross-section measurement. The uncertainty on non-prompt background gives a significant contribution to the systematic uncertainty of the $t\bar{t}W$ cross section measurement.

The observed yields, background estimates, and systematic uncertainties are used to construct a binned likelihood function $L(r, \theta)$ as a product of Poisson probabilities from all bins. The parameter r is the signal strength modifier and θ represents the full suite of nuisance parameters. The signal strength parameter $r = 1$ corresponds to a signal cross section equal to the SM prediction. An expected signal significance of 4.5 standard deviations in the same-sign two-lepton channel with the background-only hypothesis, and 8.3 and 4.7 standard deviations in the three and four-lepton channels, respectively are expected. The observed significances are 5.3, 8.6, and 4.5 respectively.

The measured signal strength parameter is found to be $1.23^{+0.19}_{-0.18}(stat)^{+0.20}_{-0.18}(sys)^{+0.13}_{-0.12}(theo)$ for $t\bar{t}W$, and $1.17^{+0.11}_{-0.10}(stat)^{+0.14}_{-0.12}(sys)^{+0.11}_{-0.12}(theo)$ for $t\bar{t}Z$. The $t\bar{t}W$ and $t\bar{t}Z$ production cross sections are measured to be $\sigma_{t\bar{t}W} = 0.77^{+0.12}_{-0.11}(stat)^{+0.13}_{-0.12}(sys)$ pb and $\sigma_{t\bar{t}Z} = 0.99^{+0.09}_{-0.08}(stat)^{+0.12}_{-0.10}(sys)$ pb, respectively. All measurements are consistent with SM calculations. Figure 1 shows the distribution of event yields in the analysis of the SS final state (left) and the final discriminating variable in the analysis of the $t\bar{t}Z$ three lepton final state (right) performed by CMS [2]. The results of the individual fits are summarized in Figure 2 as well as a simultaneous fit of the cross sections of the two processes using all dilepton, three-lepton, and four-lepton channels.

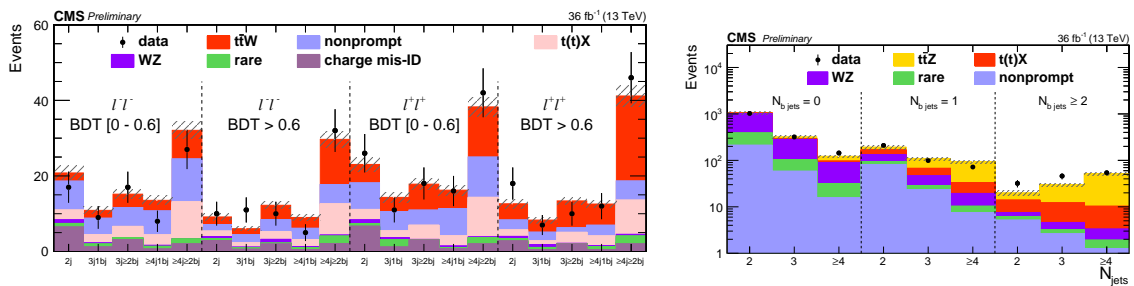


Figure 1: (left) Post-fit predicted and observed yields in each analysis bin in the same-sign dilepton analysis. The hatched band shows the total uncertainty associated to signal and background predictions. (right) Post-fit predicted and observed yields in $N_{jets} = 2, 3$ and ≥ 4 categories in the three-lepton analyses. The hatched band shows the total uncertainty associated to signal and background predictions [2].

These results are also used, in the context of effective field theory interpretations, to set constraints on the Wilson coefficients of 8 operators which would modify $t\bar{t}Z$ and $t\bar{t}W$ production.

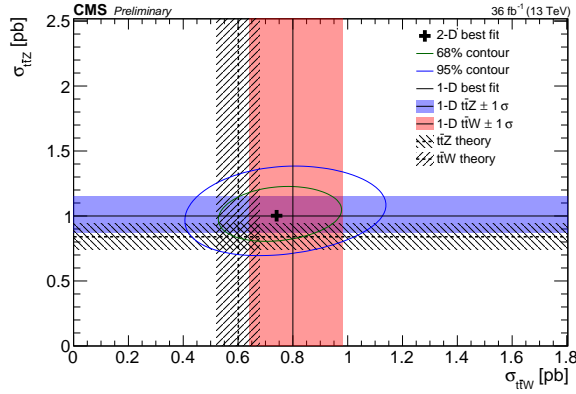


Figure 2: The result of the two-dimensional best fit for $t\bar{t}W$ and $t\bar{t}Z$ cross sections (cross symbol) is shown along with its 68 and 95% confidence level contours. The result of this fit is superimposed with the separate $t\bar{t}W$ and $t\bar{t}Z$ cross section measurements [2].

3. Measurement of the $t\bar{t}\gamma$ cross-section at $\sqrt{s} = 8$ TeV

The final state of the signal process in the semileptonic decay channel consists of a high transverse momentum of the lepton, missing transverse energy due to the presence of a neutrino, jets originating from both the b quarks and from the decay of the W boson, and an energetic γ . The cross section of the process is measured relative to the $t\bar{t}$ production cross section. In order to measure the ratio of these two cross sections, it is essential to measure how many $t\bar{t}$ and $t\bar{t}\gamma$ events are observed in data [3]. After the photon selection is applied, over half of the events in simulation come from background processes, and not from $t\bar{t}\gamma$. The two largest backgrounds are from $t\bar{t}$ events which have a non-prompt photon coming from jets in the event and events. There is not a single variable which can sufficiently discriminate both of these backgrounds for the signal. The $V + \gamma$ background can be differentiated from $t\bar{t}\gamma$ events by trying to reconstruct a top quark in the event, however $t\bar{t}$ events are very similar to the signal in this respect. Alternatively, the non-prompt photon from the $t\bar{t}$ background will tend to be less isolated than the photons from the signal, but the photon isolation variable will not be able to distinguish the $V + \gamma$ background from $t\bar{t}\gamma$ events. In order to be able to distinguish both $t\bar{t}$ and $V + \gamma$ background events, both of these methods are used and the results are combined to measure the number of the $t\bar{t}\gamma$ events observed in data. The number of events passing the photon selection containing top quark pairs can be measured by reconstructing the hadronically decaying top quark in the event. The M3 variable, defined as the invariant mass of the three jet combination that has the highest summed transverse momentum, is used for this purpose. The photon purity variable, defined as the fraction of reconstructed photons in the selection region which come from isolated photons as opposed to fake photons originating from jets, is used to discriminate between the types of real photons expected from signal and the hadronically produced photons from the $t\bar{t}$ background. The fits for extracting the top quark purity and the photon purity are performed sequentially, and then the values are used in a likelihood function, from which a fit is performed to extract the number of selected events that originate from the $t\bar{t}\gamma$ signal process. The statistical uncertainty in the number of signal events dominates the determination of the cross section for $t\bar{t}\gamma$. It includes the uncertainties on the measurement of the photon purity, top

purity after photon selection and the statistical uncertainty from the number of events in data. The ratio of the fiducial cross section of $t\bar{t}\gamma$ to $t\bar{t}$ production, R , is obtained to be: $R = (5.2 \pm 1.1) \times 10^{-4}$ for the combination of electron+jets and muon+jets final state as well as fiducial cross section $\sigma_{t\bar{t}\gamma}^{fid} = 127 \pm 27$ fb and the cross section times branching ratio $\sigma_{t\bar{t}\gamma} \times \mathcal{B} = 515 \pm 108$ fb. The measured cross section is in agreement with the theoretical prediction. Figure 3 shows the distribution of the transverse momentum and pseudorapidity of the photon in data and simulation, scaled to the result of the likelihood fit.

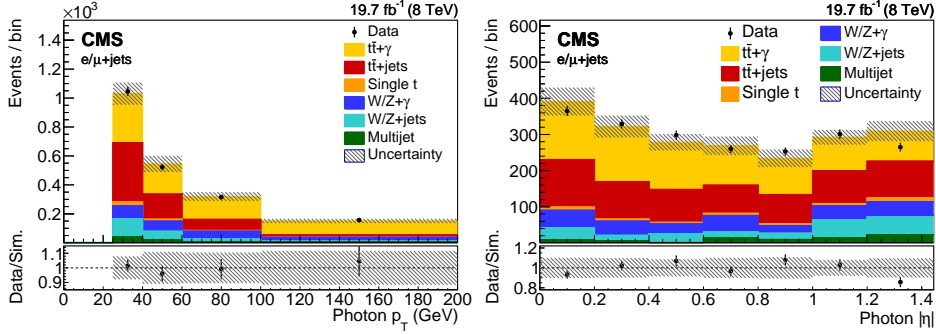


Figure 3: Distribution of the transverse momentum and pseudorapidity of the photon in data and simulation, scaled to the result of the likelihood fit in a combination of the e +jets and μ +jets channels for events passing the photon selection. The lower panel shows the ratio of the data to the prediction from simulation. The uncertainty band is a combination of statistical and systematic uncertainties in the simulation [3].

4. Measurement of the $t\bar{t}b\bar{b}$ cross-section at $\sqrt{s} = 13$ TeV

Measurements of production cross sections of the processes $pp \rightarrow t\bar{t}jj$ ($\sigma_{t\bar{t}jj}$) and $pp \rightarrow t\bar{t}b\bar{b}$ ($\sigma_{t\bar{t}b\bar{b}}$) provide an important test of NLO QCD theory calculations and important input for describing the main background in the search for the $t\bar{t}H$ process. Measurements of the cross sections $\sigma_{t\bar{t}b\bar{b}}$ and $\sigma_{t\bar{t}jj}$ and their ratio are performed using a data sample corresponding to an integrated luminosity of 2.3 fb^{-1} [4]. Events are selected with the final state consisting of two leptons (e or μ) and at least four reconstructed jets, of which at least two are identified as b quark jets. The cross section ratio is measured with a smaller systematic uncertainty exploiting the partial cancellation of uncertainties. The third and the fourth jets from $t\bar{t}jj$ events are mostly light-flavour jets, while these are heavy-flavour jets for $t\bar{t}b\bar{b}$ events. The normalized 2D distributions of the discriminators from simulation for the third and the fourth are used to separate $t\bar{t}b\bar{b}$ events from other processes. To extract the ratio of the number of $t\bar{t}b\bar{b}$ events to $t\bar{t}jj$ events, a binned maximum likelihood fit is performed on the 2D distribution of the b tagging discriminators of the third and the fourth jets, where the three event categories $e^{\pm}e^{\mp}$, $\mu^{\pm}\mu^{\mp}$ and $e^{\pm}\mu^{\mp}$ are merged. Figure 4 shows the comparisons of the b tagging discriminator distributions of the third and the fourth jets in the events from data and simulation, where the simulated histograms are scaled to the fit result. Main systematic uncertainties affecting the measurement are the jet energy scale and resolution, b tagging, and the choice of MC generator and scale in parton shower. The measured cross section ratios in the visible and the full phase space are $\sigma_{t\bar{t}b\bar{b}}/\sigma_{t\bar{t}jj} = 0.024 \pm 0.003(stat) \pm 0.007(syst)$

and $\sigma_{t\bar{t}b\bar{b}}/\sigma_{t\bar{t}jj} = 0.022 \pm 0.003(stat) \pm 0.006(syst)$ respectively, where a minimum transverse momentum for the particle-level jets of 20 GeV is required. The individual cross sections $\sigma_{t\bar{t}b\bar{b}} = 4.0 \pm 0.6(stat) \pm 1.3(syst)$ pb and $\sigma_{t\bar{t}jj} = 184 \pm 6(stat) \pm 33(syst)$ pb are also measured. These results, in particular the ratio of the cross sections, provide important information for the $t\bar{t}H$ search, permitting the reduction of a dominant systematic uncertainty that derives from the uncertainty in the $t\bar{t}b\bar{b}$ background.

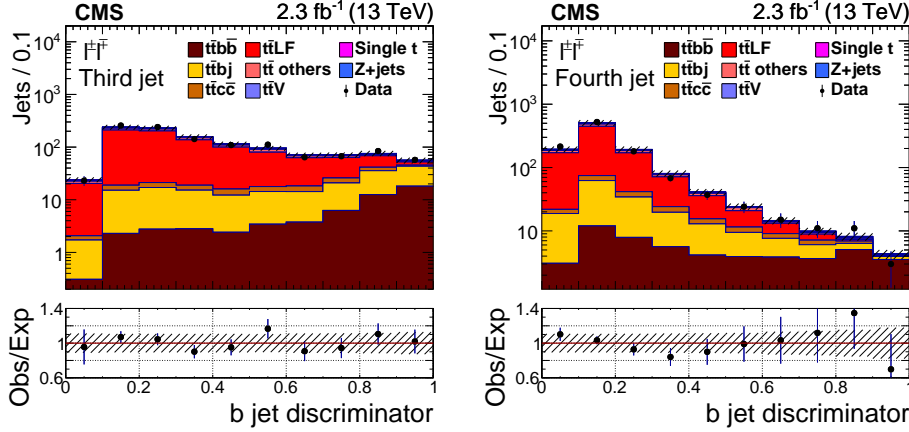


Figure 4: Distributions of b jet discriminator for the third (left) and the fourth (right) jets in decreasing order of b tagging discriminator value, after the full event selection. The points show the data and the stacked histograms are from simulated events, normalized by the results of the fit. The ratio of the number of data events to the expected number, as given by the stacked histograms, is shown in the lower panels. The hatched region indicates the modelling uncertainty in the MC simulation [4].

5. Search for standard model production of four top quarks in proton-proton collisions at $\sqrt{s} = 13$ TeV

The SM cross section for $t\bar{t}t\bar{t}$ production is approximately five orders of magnitude smaller than that for $t\bar{t}$, and $t\bar{t}t\bar{t}$ production has not yet been observed. Observing $t\bar{t}t\bar{t}$ production consistent with SM predictions would provide a valuable test of higher-order perturbative QCD calculations. Two BDTs are used to improve the discrimination between signal and background. The principal differences between $t\bar{t}t\bar{t}$ and production are in the jet multiplicity and the number of b jets. To further improve the sensitivity of the analyses, the data are split into exclusive jet multiplicity categories. The single-lepton analysis uses categories of $N_j = 6, 7, 8$, and ≥ 9 . The dilepton channel, with fewer events, uses only the 4-5, 6-7, and ≥ 8 jet categories. The largest systematic uncertainty comes from the variation of QCD scale choice in the matrix element [5]. No deviation from the background-only simulation is observed. This leads to an upper limit on the SM production cross section for $t\bar{t}t\bar{t}$ of 94 fb ($10.2 \sigma_{t\bar{t}t\bar{t}}^{SM}$), with an expected limit of 118_{-41}^{+76} fb at the 95% confidence level. This result is combined with a previous search [6] with similar sensitivity in the same-sign dilepton channel to obtain an improved limit of 69 fb, with an expected limit of 71_{-24}^{+38} fb.

References

- [1] CMS Collaboration, JINST 3, **S08004** (2008)
- [2] CMS Collaboration, CMS-PAS-TOP-17-005, [<https://cds.cern.ch/record/2264544>] 2017.
- [3] CMS Collaboration, arXiv:1706.08128
- [4] CMS Collaboration, arXiv:1705.10141
- [5] CMS Collaboration, Phys. Lett. **B 772** (2017) 336
- [6] CMS collaboration, Eur. Phys. J. **C 76** (2016) 439