PROCEEDINGS OF SCIENCE

Measurements of $t\bar{t}+\mathrm{jets}$, $t\bar{t}+b\bar{b}$, and $t\bar{t}t\bar{t}$ production at CMS

Andrea Castro[*†]
University of Bologna and INFN - Italy
E-mail: Andrea.Castro@cern.ch

The most recent CMS measurements of the production of $t\bar{t}$ pairs with two additional jets, or two additional b quarks, are presented. These two processes are important for several reasons: they can allow a better modeling of Monte Carlo simulations, because of the different energy scale associated to b-jets and generic jets; they are important backgrounds to other standard model processes; and they might be related to new physics. In addition, the search for the production of two $t\bar{t}$ pairs is discussed, along with its connection to the Yukawa coupling of top quarks to the Higgs boson.

The 40th International Conference on High Energy Physics
28 July - 6 August 2020
Prague, Czech Republic

[*]Speaker.
[†]The measurements described here are presented on behalf of the CMS Collaboration.

PoS(ICHEP2020)320

1. Introduction

Because of its large mass and large Yukawa coupling y_{t} to the Higgs boson H, the top quark t lies at the boundaries of the standard model (SM) territory, and it is important to map as fully as possible this area. The production at LHC of top quarks, either singly and in pairs, is well measured in terms of total and differential cross sections. This is to be complemented by the study of the production of $\mathrm{t\bar{t}}$ pairs plus something else, which is interesting for theoretical and experimental reasons.

Specifically, $\mathrm{t\bar{t}+jj}$ (where j stands for generic jets) [1–3] and $\mathrm{t\bar{t}+b\bar{b}}$ production [4–6] represent important backgrounds for $\mathrm{t\bar{t}H}$ and for beyond-the-SM events. In addition, generic jets and b-originated jets are associated to different energy scales and described with different approaches assuming massless or massive b quarks. So, a deeper characterization of these processes might allow a better modeling of Monte Carlo (MC) simulations.

On the other hand, $\mathrm{t\bar{t}t\bar{t}}$ production is very rare, with a cross section of about 12 fb [7]. For this reason it has not been observed yet, although a strong evidence has been presented at this conference by the ATLAS Collaboration [8]. The importance of this channel is that it can constrain the t–H Yukawa coupling while being sensitive to new physics.

In this report, $\mathrm{t\bar{t}}+\mathrm{jets}$, $\mathrm{t\bar{t}+bb}$, and $\mathrm{t\bar{t}t\bar{t}}$ production cross sections are measured, in different decay channels, using recent data from pp collisions at the LHC collected by the CMS detector [9] at $\sqrt{s} = 13$ TeV.

2. Cross section for $\mathrm{t\bar{t}+jj}$ and $\mathrm{t\bar{t}+b\bar{b}}$, in the single-lepton and dilepton channels (36 fb^{-1})

For this measurement [10], standard single-lepton (SL) and dilepton (DL) selections are applied, the important point being the identification of the additional jets (i.e. those not directly associated to $\mathrm{t\bar{t}}$) and their possible association to b quarks (b-tagging), using the scores provided by the combined secondary vertex (CSV) algorithm [11]. For the DL events the additional b jets are chosen as those with the 3rd and 4th largest b-tagging scores, while in the SL case a kinematic fit to the $\mathrm{t\bar{t}}$ hypothesis helps the choice.

Two different phase space definitions are considered: the visible phase space (VPS), which corresponds to a number of jets and leptons matching the selection, and a full phase space (FPS), where all decays having at least two additional jets are considered. The b-tagging scores are used in a maximum-likelihood (ML) fit to extract the $\mathrm{t\bar{t}jj}$ cross section and the ratio $R_{\mathrm{t\bar{t}b\bar{b}/t\bar{t}jj}}$ of the $\mathrm{t\bar{t}b\bar{b}}$ vs $\mathrm{t\bar{t}jj}$ cross sections. The resulting measurements are then extrapolated to the FPS using MC acceptances. Because of different phase space definitions and different background compositions, the two channels are fitted separately.

The measured $\mathrm{t\bar{t}jj}$ cross section in the VPS for the DL channel is shown in Fig. 1-left along with the cross section ratio which is about 30% larger than what expected from MC generators, with a 3σ significance (which is slightly smaller for the SL case). Cross sections and ratio at FPS are shown instead in Fig. 1-right, with comparison to different MCs. Similar plots are obtained for the SL channel. The treatment of the systematic uncertainties as nuisance parameters in the fit provided an improvement in precision with respect to previous measurements.

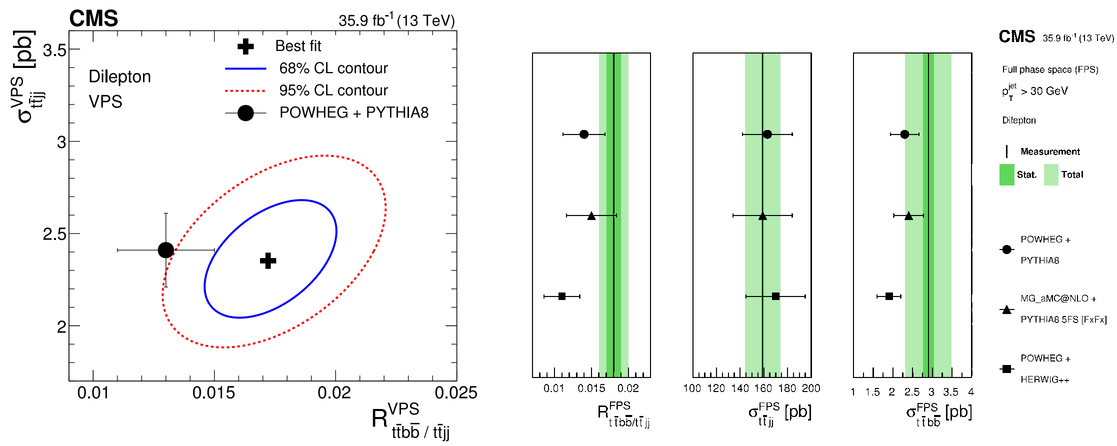

Figure 1: Dilepton channel [10]. Left: contour plot for $\sigma_{t\bar{t}jj}$ vs $R_{t\bar{t}b\bar{b}/t\bar{t}jj}$ in the VPS. Right: measured values of the $t\bar{t}b\bar{b}$ and $t\bar{t}jj$ cross sections and their ratio in the FPS, compared to different MC generators.

3. Cross section for $t\bar{t}+b\bar{b}$, in the all-hadronic channel (36 fb^{-1})

For this measurement [12], a baseline selection requiring at least 8 jets, 4 of which b-tagged, is strengthened with the help of multivariate discriminators which serve also for the identification of the additional jets. One discriminator (quark-gluon likelihood ratio, QGLR) is based on jet substructure and favors events produced from 4 quarks rather than from 4 gluons; a boosted decision tree (BDT) discriminator helps the selection of the correct jet permutation; and an additional weakly supervised BDT discriminator (CWoLa BDT) is trained for the reduction of the major multijet background.

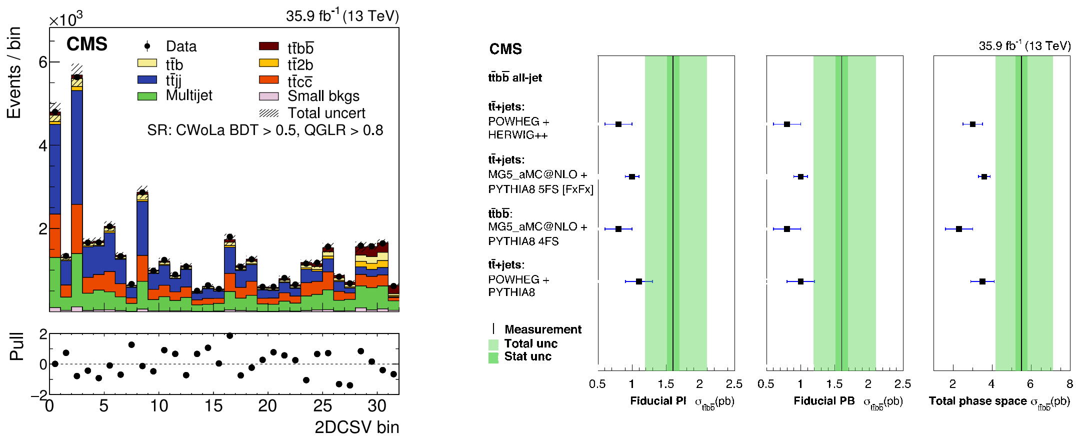

Figure 2: Left: distribution of the 2DCSV in the SR region [12]. Right: comparison of the measured $t\bar{t}b\bar{b}$ production cross sections with predictions from several MC generators, for the three phase space definitions.

Two different phase space definitions are considered: one parton-based (PB) which refers to partons after radiation, and one which is parton-independent (PI) because refers to stable particles after hadronization. A 2-dimensional discriminator (2DCSV) is built from the b-tagging scores of the additional jets; its unrolling into one dimension is shown in Fig. 2-left for a signal (SR) region.

The data are fitted to the 2DCSV discriminators, in the SR and in three control regions (CRs) defined in terms of QGLR and CWoLa BDT outputs, yielding cross sections values of $1.6 \pm 0.1(stat)^{+0.5}_{-0.4}(syst)$ pb both for the PB and the PI phase space. The two values are numerically equal at the quoted precision because of the large overlap between the two phase space definitions. The total phase space cross section is then obtained applying acceptances to the PB phase space and amounts to $5.5 \pm 0.3(stat)^{+1.6}_{-1.3}(syst)$ pb. Comparisons of these cross sections to different MC predictions are shown in Fig. 2-right. The measurements are found to be larger than MC predictions by a factor of 1.5–2.4, corresponding to 1–2 standard deviations. This is consistent with what observed in the SL/DL measurements described above.

4. Search for $\mathrm{t\bar{t}t\bar{t}}$, in the single-lepton and opposite-sign dilepton channels (36 fb^{-1})

The production of four top quarks is searched for in SL events and in DL events where the leptons have opposite-sign charges [13]. The event selection for this search is based on the number of leptons and of jets, N_j, including those b-tagged, N_b. Figure 3-left shows as an example the N_j distributions from data and MC simulations for the SL channel.

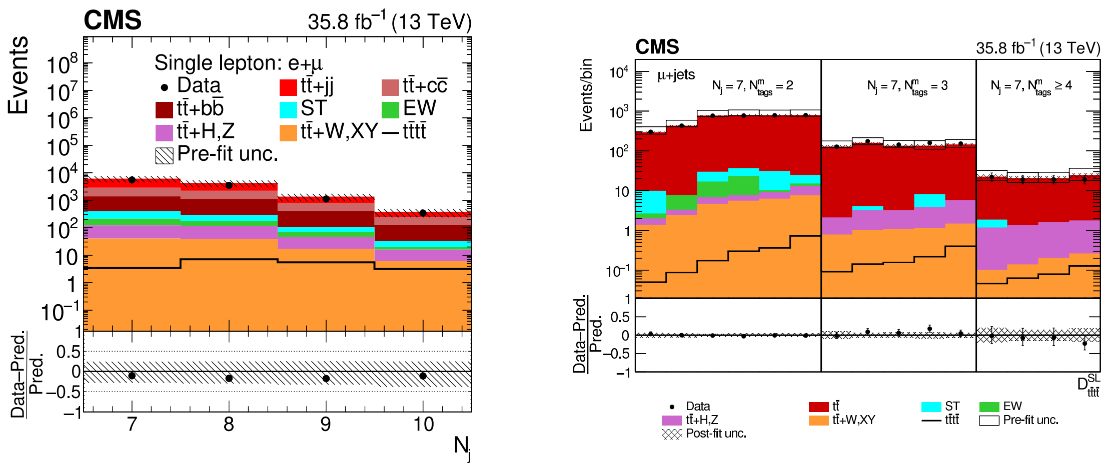

Figure 3: Left: distribution of N_j in the combined SL channels [13]. Right: post-fit $D^{\mathrm{SL}}_{\mathrm{t\bar{t}t\bar{t}}}$ distribution in the single-muon channel for events with $N_\mathrm{j} = 7$, and 2, 3, ≥ 4 tagged jets.

In order to improve the sample purity and to identify the top quarks, two BDTs are trained for the two channels, $D^{\mathrm{DL}}_{\mathrm{t\bar{t}t\bar{t}}}$ and $D^{\mathrm{SL}}_{\mathrm{t\bar{t}t\bar{t}}}$, but to increase the sentitivity the events are also categorized in terms of N_j and N_b. The choice of the input variables has been optimized separately for the two channels.

A simultaneous binned ML fit is carried to extract the signal strength μ defined as the ratio between observed and theoretical cross sections. An example of the results of the fit is shown in Fig. 3-right. The measured signal strengths amount to $1.6^{+4.6}_{-1.6}$ for the SL channel and to $0^{+2.7}$ for the DL case where no candidates are observed. The combination of the two channels gives $\mu = 0^{+2.2}$, corresponding to a cross section $\sigma(\mathrm{t\bar{t}t\bar{t}}) = 0^{+20}$ fb. The significance of this measurement is still poor, and thus a 95% confidence level (CL) upper limit of 48 fb is quoted. A combination with a previous measurement in the multijet channel [14] yields a cross section value of 13^{+11}_{-9} fb, with a significance of 1.4 standard deviations.

5. Search for $\mathrm{t\bar{t}t\bar{t}}$, in the same-sign dilepton and trilepton channels (137 fb^{-1})

The production of four top quarks is searched also [15] in events with two leptons having the same-sign charge, and in events with three leptons, using the standard lepton requirements. The event selection requires also two generic jets and two b-tagged jets.

There are two analyses: one called 'cut-based' which considers 14 SRs and 2 CRs, and a second one based on a BDT, which considers instead 17 SRs and 1 CR. They both provide consistent results, but the BDT analysis is expected to be more sensitive and precise so it is the one used for the interpretation of the results. Observed and predicted yields, as shown in Fig. 4-left, enter a profile ML fit in which the parameter of interest is the $\mathrm{t\bar{t}t\bar{t}}$ cross section, and all nuisance parameters are profiled.

The obtained cross section amounts to $\sigma_{\mathrm{BDT}} = 12.6^{+5.8}_{-5.2}$ fb. Given the small (2.6σ) significance of this measurement, a 95% CL observed upper limit of 22.5 fb is assigned. The expected upper limit of $8.5^{+3.9}_{-2.6}$ fb represents however a significant improvement relative to a previous determination by CMS in the same channel [14].

The dependence of the cross section on y_{t}, as shown in Fig. 4-right, enables us to contrain it, yielding an upper limit of 1.7 times the SM value, at 95% CL.

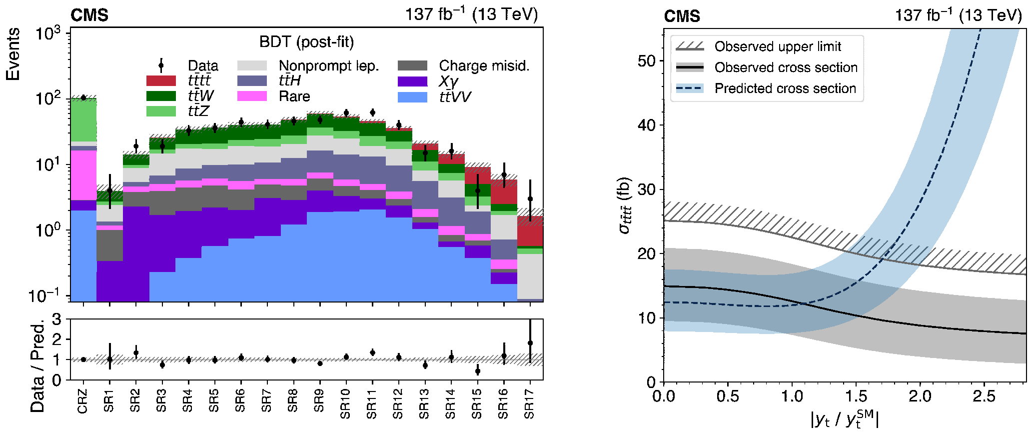

Figure 4: Left: observed yields in the control and signal regions for the BDT analysis, compared to the post-fit predictions [15]. Right: observed $\sigma(\mathrm{pp} \to \mathrm{t\bar{t}t\bar{t}})$ and 95% CL upper limit shown as a function of $|y_{\mathrm{t}}/y_{\mathrm{t}}^{\mathrm{SM}}|$.

6. Conclusions

The production of $\mathrm{t\bar{t}}$ pairs plus two generic jets, or plus two b-originated jets has been measured by the CMS Collaboration in different decay channels. The cross sections for $\mathrm{t\bar{t}+jj}$ events are consistent with SM predictions, while those for $\mathrm{t\bar{t}+b\bar{b}}$ production are approximately 30% larger, as shown in Fig. 5, begging for further experimental and theoretical studies.

As for $\mathrm{t\bar{t}t\bar{t}}$ production, this rare process has not been observed yet and upper limits of about 2–4 times the SM cross section have been set, along with a limit on the value of the Yukawa coupling of the top quark to the Higgs boson relative to its SM value.

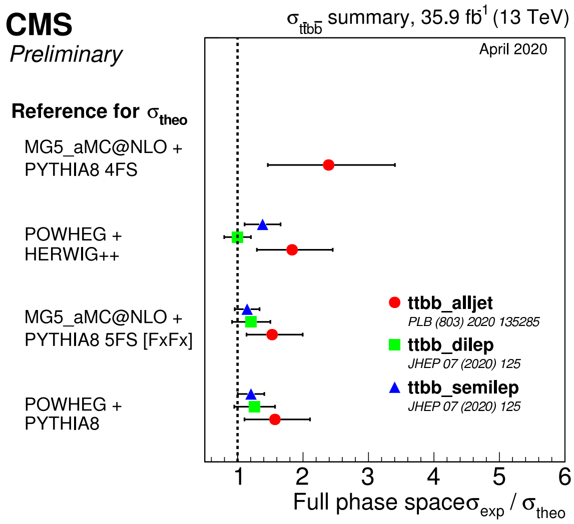

Figure 5: Summary of CMS $t\bar{t}b\bar{b}$ cross section measurements compared to SM predictions. The measurements shown here refer to the FPS.

References

[1] G. Bevilacqua, M. Czakon, C.G. Papadopoulos and M. Worek, Phys. Rev. D 84 (2011) 114017.

[2] S. Höche, F. Krauss, P. Maierhoefer, S. Pozzorini and M. Schönherr, Phys. Lett. B 748 (2015) 74.

[3] C. Gütschow, J.M. Lindert and M. Schönherr, Eur. Phys. J. C 78 (2018) 4.

[4] G. Bevilacqua and M. Worek, JHEP 07 (2014) 135.

[5] G. Bevilacqua, M.V. Garzelli and A. Karlos arxiv:1709.06915.

[6] T. Ježo, J.M. Lindert, N. Moretti and S. Pozzorini, Eur. Phys. J. C 78 (2018) 502.

[7] R. Frederix, D. Pagani and M. Zaro, JHEP 02 (2018) 031.

[8] ATLAS Collaboration, arxiv:2007.14858.

[9] CMS Collaboration, JINST 3 (2008) S08004.

[10] CMS Collaboration, JHEP 07 (2020) 125.

[11] CMS Collaboration, JINST 13 (2018) P05011.

[12] CMS Collaboration, PLB 803 (2020) 135285.

[13] CMS Collaboration, JHEP 11 (2019) 082.

[14] CMS Collaboration, EPJC 78 (2018) 140.

[15] CMS Collaboration, EPJC 80 (2020) 75.