Expectations for first measurements of top-antitop pair production using early CMS data

A.-C. Le Bihan [a] *on behalf of the CMS collaboration*

[a]Institut Pluridisciplinaire Hubert Curien, Strasbourg, France

The top quark study will be a fundamental element of the early physics program at the Large Hadron Collider (LHC). Given the complex signature of the *most exotic* of all known Standard Model particles, the pair production of top quarks will be crucial for commissioning the LHC experiments for physics analysis. Only when the first top quark signal will be established, will the experiments be able to further probe the Standard Model and to begin the searches for new physics. We shall discuss the plans and analysis strategies of CMS to pursue this physics program, and shall show the expected performance of the experiment with an emphasis on an early cross-section measurement.

1. Introduction

The LHC *pp* collider is widely known to be a top quark factory. About one top is expected to be produced per second at the reference luminosity of $10^{33}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and center-of-mass energy of 14 TeV. At 7 TeV (10 TeV), the top-antitop production rate remains high with an expected cross-section of about 160 pb (385 pb). First top-antitop cross-section measurements are thus expected to be performed using the first data collected at 7 TeV by the LHC since april 2010.

Top-antitop pairs are mainly produced by strong interaction through gluon-gluon fusion (about 90% of the time at 14 TeV). Single-top production can occur via the electroweak interaction, mainly in the t-channel, with an expected cross-section of about one third of the top-antitop cross-section. Top quarks decay into Wb before hadronization and the top-antitop final state is denoted di-leptonic (9%), semi-leptonic (45%) and fully-hadronic (46%), according to the W bosons decay modes. Top-antitop pairs are thus characterized by the presence of high momentum isolated leptons, (b-)jets and transverse missing energy (MET). Those final states can mainly be mimicked by QCD multi-jets events, W+jets or Z+jets events, where jet(s) are misidentified as isolated leptons and fake MET can be reconstructed due to jet energy mismeasurements.

In the following we will review the CMS strategy to select top-antitop pairs in early data (di-leptonic and semi-leptonic final states) and to estimate the remaining backgrounds in a data driven way. Most results are based on 10 TeV simulation, the strategy remaining valid for 7 TeV data.

2. Early cross-section measurements

Events are recorded using unprescaled inclusive single lepton triggers. The hardware based level-1 trigger selects candidates with electromagnetic deposits in the calorimeter or muon segments in the muon chambers. The high level trigger filters out electron and muon candidates with transverse energy or momentum thresholds of $E_T > 15$ GeV and $p_T > 9$ GeV/c respectively.

The leptons are required to be isolated in order to distinguish leptons from W decays from leptons in jets originating from semi-leptonic b/c-hadrons decays. Both tracker and calorimeter isolation criteria are considered. A lepton is considered as isolated if the energy/momentum deposit in a $\Delta R = \sqrt{(\Delta\phi^2 + \Delta\eta^2)} \approx 0.3$ cone around its axis is low.

2.1. Di-lepton channel

The golden channel final state is selected requiring two oppositely charged isolated leptons of

$p_T > 20$ GeV/c and at least two jets of $p_T > 20$ GeV/c. A Z mass veto ($M_{ll} \notin [76-106]$ GeV/c^2) is applied to reject Drell-Yan events. The presence of two neutrinos in the final state allows to require moderate MET, $> 20\,(30)$ GeV depending on the final state, eμ (ee, $\mu\mu$).

The cross-section is then extracted using a simple counting experiment: $\sigma_{t\bar{t}} = \frac{N_{obs}-N_{bkg}}{A\epsilon L_{int}}$, where N_{obs} describes the number of observed events, A the signal acceptance, ϵ the signal efficiency and L_{int} the integrated luminosity. N_{bkg} refers to the expected background and is evaluated using data-driven methods.

The **lepton fake rate** is estimated using a data sample dominated by QCD multi-jets events, where the identified leptons are assumed to be fakes. The fake rate is applied in early data ($L_{int} \approx 10$ pb^{-1}) regardless of the considered process. With a higher integrated luminosity $L_{int} \approx 100$ pb^{-1} a set of equations for loose and tight lepton pair combinations can be solved to estimate the number of W+jets and QCD multi-jets events with fake leptons.

The **Drell-Yan** contribution remaining outside the Z mass region at high MET is estimated by scaling the number of lepton pairs in the Z mass region to the expectations taken from simulation.

With 10 pb^{-1} of data at 10 TeV, the $t\bar{t}$ pair production cross-section is expected to be measured with $\Delta\sigma_{(t\bar{t})}/\sigma_{(t\bar{t})} = 15\%\,(stat) \pm 10\,\%(sys) \pm 10\,\%(lumi)$ [1]. Figure 1 shows the expected jet mutiplicity by combining all three di-leptons channels, once all selection cuts applied.

2.2. Lepton+jets channel

The semi-leptonic $t\bar{t}$ events are expected to suffer from a larger QCD multi-jet background than the di-lepton $t\bar{t}$ events, especially in the electron+jets channel due to the presence of photon conversions. The high branching ratio and absence of neutrinos in this final state allow this channel to be used to reconstruct the top mass. The event selection requires the presence of exactly one central electron or muon with $p_T > 20$ GeV/c and at least four jets with $p_T > 30$-40 GeV/c, the latter being used to reconstruct the hadronic top mass. No attempt is made to apply MET or b-tagging criteria in early data.

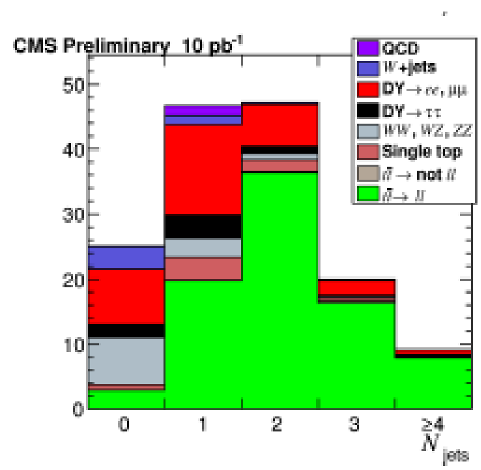

Figure 1. Number of jets in the di-lepton channel - $\sqrt{s} = 10$ TeV, $L_{int} = 10$ pb^{-1}.

The main backgrounds, QCD multi-jets and W+jets are evaluated in a data-driven way. Two methods have been developed to estimate the **QCD multi-jet background**. The first one, called ABCD method, relies on two independent variables, the muon impact parameter significance and muon isolation (muon+jets channel). Assuming that the ratio of background events is the same in the A and C regions and B and D regions, $\frac{N_A}{N_C} = \frac{N_B}{N_D}$, the background contribution in the signal region A is estimated : $N_A = N_B N_C / N_D$ (see Figure 2). The second method performs a sideband region fit to the relative lepton isolation variable from the QCD multi-jet region to the signal region. The boundaries of the sideband region are varied according to the jet multiplicity. The uncertainty is conservatively assumed as 50% for both methods.

The W+jets background is estimated using the W charge asymmetry in pp collisions.

This method is statistically limited and requires 200 pb^{-1} of integrated luminosity of data to obtain a 30% error estimate. Measuring the number of leptons N^+ to antileptons N^- in the data sample and the W asymmetry in simulation $R_{\pm}(W)$, one can estimate the number of events leading to charge asymmetry : $(N^+ + N^-) = R_{\pm}(W) \times (N^+ - N^-)$. In addition to W+jets events, V$b\bar{b}$, single-top and WZ events leading to charge asymmetry are considered.

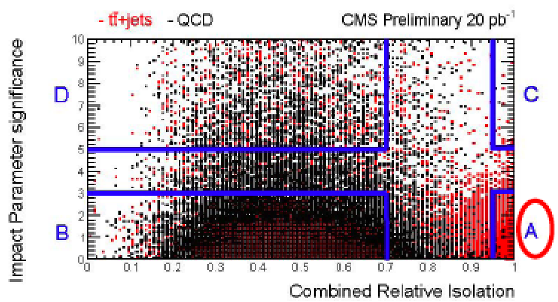

Figure 2. ABCD method used to estimate the QCD multi-jet background in the muon+jets channel.

The cross-section is extracted from a binned likelihood fit to a discriminating variable between signal and background. Three variables are considered : the pseudorapidity of the lepton $\eta(\mu)$, $M3$ and $M3'$. $M3$ denotes the invariant mass of the three jets leading to the highest invariant mass which is deemed to describe the hadronic top mass, see Figure 3. $M3'$ refers to the $M3$ variable where the jets have been χ^2 sorted according to the W and top masses (hadronic and leptonic legs). The expected amount of QCD background being small and the Z+jets and W+jets shapes almost indistinguishable, three templates are used : W/Z+jets, $t\bar{t}$ and single-top.

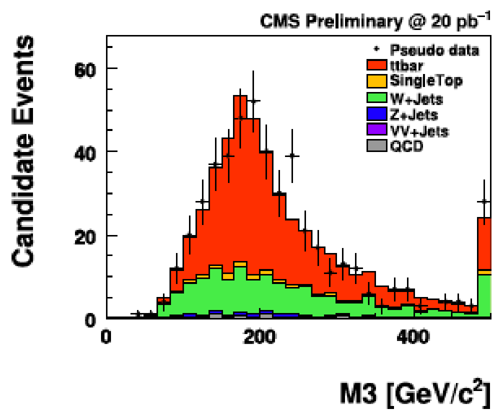

Figure 3. M3 distribution : invariant mass of the three jets leading to the highest invariant mass (muon+jets channel).

With 20 pb^{-1} of data at 10 TeV, the $t\bar{t}$ pair production cross-section is expected to be measured with $\Delta\sigma_{(t\bar{t})}/\sigma_{(t\bar{t})} = 12 - 18\% \,(stat) \pm 20 - 25\,\%(sys)$, the dominant systematic error being the jet energy scale [2], [3].

3. First analysis with L_{int} 100 pb^{-1}

3.1. Top as a calibration tool

Studies performed respectively at 10 TeV and 14 TeV are showing that a precision on the b-tagging efficiency of $\pm 2\%(stat) \pm 4\%(sys)$ and $\pm 1\%$ on the jet energy scale could be reached with respectively 250 pb^{-1} and 100 pb^{-1} of data.

The b-tagging efficiency ϵ_b, is expected to be measured in a quite pure $t\bar{t} \rightarrow e\mu b\bar{b}$ sample where the expected background is estimated to be less than 10 %. Considering the number of b-tagged jets distribution, which depend on ϵ_b, the ratio $R = BR(t \rightarrow Wb)/BR(t \rightarrow Wq_{(q=d,s,b)})$, and b-jets misassignments, ϵ_b can be measured assuming R known. Alternatively the R ratio can be measured with a $\pm 2\%(stat) \pm 9\%(sys\, from\, \epsilon_b) \pm 3\%$ accuracy if ϵ_b is known [4]. R is expected to be close to one according to the Standard Model and may allow to probe the existence of a fourth family of quarks.

Reference [5] describes how the jet energy scale of light and b jets could be measured with a precision of 1 % by constraining the W and top masses.

3.2. Single top

An early single-top evidence is expected to be seen selecting μ+jets events in 200 pb^{-1} of data (about 700 pb^{-1} of data at 7 TeV). A fit to the polarization angle $\theta^*(\bar{q}\mu)$ in the top rest frame is used to extract the cross-section. The main background are the $t\bar{t}$ events, once the QCD and W+jets events rejected by a selection on the W transverse mass $M_T(W) > 50$ GeV [6].

3.3. Boosted top

Several Beyond Standard Model extensions predict contributions to top-pair production leading to a distorted $M_{t\bar{t}}$ spectrum. Different strategies are investigated to reconstruct the collimated decay products of boosted top quarks. A dedicated jet algorithm has been developed to reconstruct monojets with a substructure compatible with the W and top masses [7]. The fully hadronic final state could be selected requiring the presence of two monojets with $p_T > 250$ GeV/c. Such a selection would lead to a discovery for signals with $\sigma \times BR > 4.0(1.6)$ pb and $M_{t\bar{t}} = 2(3)$ GeV with 100 pb^{-1} of data at 10 TeV, or alternatively to a 95% CL exclusion for $\sigma \times BR > 1.5(0.7)$ pb [8]. Boosted tops decaying into semi-leptonic final states are expected to be identified using a standard top-antitop pair selection with relaxed isolation criteria [9]. The expected sensitivity is a 95% CL exclusion for $M_{t\bar{t}}$ = 9 TeV, $\sigma \times BR > 9$ pb using 100 pb^{-1} of data at 10 TeV.

4. Conclusion

We have presented the early expectations about top-antitop pair measurements at CMS. The first collisions at 7 TeV have demonstrated the excellent ability of the CMS experiment to reconstruct the top ingredients, namely isolated leptons, (b-)jets, transverse missing energy. About 25 events of $t\bar{t}$ di-leptonic events and 110 semi-leptonic events are expected with 10 pb^{-1} of integrated luminosity at 7 TeV, which should allow the first cross-sections measurements in the di-leptonic and semi-leptonic channels. Once large amount of data collected, an exciting physics program will follow : the top quark will allow to perform precision tests of the Standard Model and may reveal its sensitivity to new physics.

REFERENCES

1. The CMS collaboration, Expectations for observation of top pair production in the dilepton final state with early data at 10 TeV, CMS PAS TOP-09-002.
2. The CMS collaboration, Plans for an early measurement of the $t\bar{t}$ cross-section in the muon + jets channel at 10 TeV, CMS PAS TOP-09-003.
3. The CMS collaboration, Plans for an early measurement of the $t\bar{t}$ cross-section in the electron + jets channel at 10 TeV, CMS PAS TOP-09-004.
4. The CMS collaboration, Probing the heavy flavour content of $t\bar{t}$ di-leptons at 10 TeV, CMS PAS TOP-09-001.
5. The CMS collaboration, Measurement of jet energy scale corrections using top quark events, CMS PAS TOP-07-004.
6. The CMS collaboration, Prospects for the measurement of the single-top t channel cross-section in the muon channel with 200 pb^{-1} at 10 TeV, CMS PAS TOP-09-005.
7. The CMS collaboration, A Cambridge-Aachen (C-A) based Jet Algorithm for boosted top-jet tagging, CMS PAS JME-09-001
8. The CMS collaboration, Search for High-Mass Resonances Decaying into Top-Antitop Pairs in the All-Hadronic Mode, CMS PAS EXO-09-002.
9. The CMS collaboration, Search for heavy narrow $t\bar{t}$ resonances in muon-jets final states with the CMS detector, CMS PAS EXO-09-008.