PROCEEDINGS OF SCIENCE

Top rediscovery at ATLAS and CMS

Julien Donini* **on behalf of ATLAS and CMS collaborations**
CNRS/IN2P3 & UJF/ENSPG, LPSC, Grenoble, France
E-mail: julien.donini@lpsc.in2p3.fr

We describe the plans and strategies of the ATLAS and CMS collaborations to measure the top-antitop cross section using early LHC data, assuming center-of-mass proton-proton collisions at 10 TeV. The methods rely on data-driven approaches to extrapolate the rates of the main background processes. These studies show that simple and robust selections can be employed to extract a clear and sizable $t\bar{t}$ signal and to measure the cross section with a precision of about 20-30% with a few tens of pb^{-1}.

XVIII International Workshop on Deep-Inelastic Scattering and Related Subjects, DIS 2010
April 19-23, 2010
Firenze, Italy

*Speaker.

PoS(DIS 2010)154

ATL-PHYS-PROC-2010-045
26 July 2010

1. Introduction

Since its discovery in 1995 at Tevatron [1] the top quark has been the object of many studies. The top-quark mass has been measured at 173.1 ± 1.3 GeV, with a remarkable precision of 0.7%, thus allowing further indirect constraints on the Higgs boson mass; the top-antitop production cross section was shown to be in good agreement with Standard Model (SM) expectations [2]; several properties of the top quark have been investigated (charge, width, production mechanisms, etc) with a precision getting to the point that Beyond Standard Model phenomena could be unveiled in the near future; finally the observation of single top quark has been reported recently [3] serving as a direct probe of the W-t-b interaction and of the quark mixing matrix element $|V_{tb}|$.

At the Large Hadron Collider (LHC) the top quark is produced with a sizably larger cross section with respect to the Tevatron and the physical backgrounds are expected to be more manageable. For the 7 TeV center-of-mass proton-proton collisions the LHC is currently providing the theoretical top-antitop cross section is 160 pb and the amount of events produced by 2011, when the LHC is expected to have collected 1 fb^{-1}, should exceed the total $t\bar{t}$ statistics produced at Tevatron.

The abundant top-quark production and the relatively clear signature will allow to use the decay products of the top quark, a b-quark and a W boson in the SM, as a tool to calibrate several crucial physical objects which are essential for claiming potential discoveries of new physics. For instance the hadronic decay of W bosons can be used for in-situ jet energy scale measurement while the properties of b-quark jets can be exploited to calibrate heavy-flavor tagging algorithms. The good understanding of top decay signature will then open the way to precise tests of the Standard Model and to an extensive search for new phenomena in the rich top-quark environment, such as the presence of new bosons (W'/Z', $H^{\pm}$ ...) or new quarks (4^{th} generation).

This proceeding addresses the first step to this long and rich physics program and describes the strategy developed by the ATLAS and CMS collaborations to “rediscover” the top quark using early LHC data. These studies are based on simulated events that represent collisions data corresponding to $10-200$ pb^{-1} at a center-of-mass energy of 10 TeV. Simple and robust methods are presented to select top-enriched samples and to determine the top-antitop cross section. The analyses focus on $t\bar{t}$ decays in the lepton+jets and dilepton channels and on data-driven techniques to estimate the rates of the main background processes.

2. Prospects for $t\bar{t}$ measurement in the dilepton channel

The dilepton channel, where both of the W bosons coming from the (anti)top-quark decay, themselves decay leptonically into electrons or muons, provide the cleanest $t\bar{t}$ state. Despite a small branching ratio to ee, $e\mu$ and $\mu\mu$ final states of 6.4% (including leptonic τ decays) this channel already shows, after the application of simple selection criteria, a sizable signal and a reduced background yield. The latter is dominated by Drell-Yan processes and QCD or W+jets events with jets misidentified as leptons which can mimic the $t\bar{t}$ signature.

2.1 Prospects for ATLAS

The prospects for $t\bar{t}$ cross-section measurement using the dilepton final states was estimated by the ATLAS collaboration for an integrated luminosity of 200 pb^{-1} [4]. The general selection of top-antitop dilepton candidate requires that events have two oppositely-charged isolated lepton (e or μ) candidates with a high transverse momentum of p_T>20 GeV and survive a missing E_T threshold of 20 GeV for the $e\mu$ channel and 35 GeV for the $ee, \mu\mu$ channels, to reduce the amount of Drell-Yan events (one of the largest background in the ee and $\mu\mu$ channels). Finally events are selected if they contain at least two jets with p_T>20 GeV and pass a veto on the Z-boson mass. The left plot in figure 1 shows the expected number of events of signal and background per jet multiplicity after the event selection (but without the cut on the number of jets) for the $e\mu$ channel only. The distribution is normalized to 200 pb^{-1} of data. Already a large signal is visible on top of a small background for events which have two or more jets. If all dilepton channels are combined a total of 1220 $t\bar{t}$ events are expected in the selected sample with about 260 background events.

Data-driven methods were studied for the determination of the rates of the main backgrounds surviving the event selection. In particular, strategies to determine Drell-Yan and fake backgrounds using control samples were developed. The DY contribution to the signal region is estimated by scaling the MC prediction to match the observed number of events in the data in sideband regions constructed in the E_T^{miss} vs dilepton invariant mass plane. The probability for a jet to fake a lepton can be determined by defining an auxiliary loose lepton selection, and measuring the efficiencies of both this loose and the default (tight) lepton selection in two independent event samples, dominated by real and fake leptons respectively. With these methods the uncertainty on the rates of DY and QCD/W+jets backgrounds was estimated to be 15% and 50-100% respectively.

A range of potential uncertainties was studied in addition to the ones related to the data-driven methods. In particular, uncertainties on jet energy scale, lepton efficiency and Monte-Carlo model turned out to be the largest contribution to the systematics after the uncertainty on luminosity, which was conservatively taken as 20% and is by far the leading constraint on the measurement. All uncertainties were combined by constructing a likelihood function for each channel. They were fit on the nominal prediction from Monte-Carlo samples and the final sensitivity was obtained from a profile likelihood ratio. The three channels were finally combined by performing a simultaneous fit incorporating the correlations between uncertainties. The combined $t\bar{t}$ cross section measurement results in an average statistical precision of 3% and a systematic uncertainty of 10%, not including the uncertainty on the luminosity measurement.

2.2 Prospects for CMS

CMS put the emphasis on an analysis strategy for the early phase of LHC operation with data corresponding to an integrated luminosity of 10 pb^{-1} [5]. As in the ATLAS study the Standard Model signal and background sources are selected based on the requirement of two isolated leptons (electron or muon) in the final state, at least two hadronic jets and missing transverse energy. A clear signal stands out as can be seen in the right distribution of Figure 1.

Data-driven methods are used to estimate the DY and fake leptons backgrounds. As previously the expected number of DY events is determined from the observed data and MC predictions in signal and sidebands regions in the dilepton mass spectrum. However this approach does not rely on jet and E_T^{miss} simulation. A 30% systematic uncertainty on the DY rate is estimated with this method. The prediction of the number of fake leptons passing the final identification and isolation selections relies on events with lepton candidates failing these selections while passing substantially looser selections. The precision of this method is estimated to 50%.

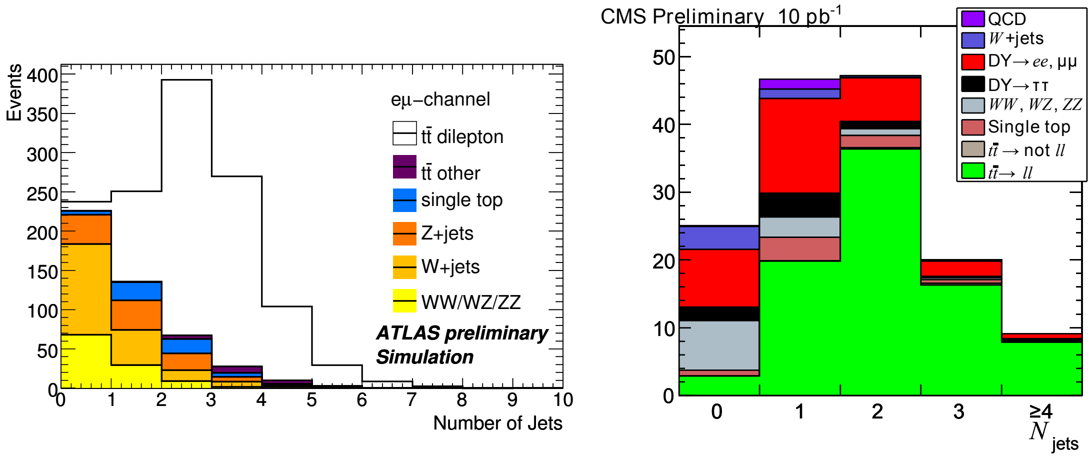

Figure 1: Expected jet multiplicity distribution after event selection for ATLAS in the $e\mu$ channel (left) and CMS in the combined dilepton channel (right). The samples are normalized to 200 pb^{-1} and 10 pb^{-1} for ATLAS and CMS respectively.

The signal cross section can be measured with a simple cut and count method with a statistical uncertainty of about 15% and a systematic uncertainty close to 10% excluding the uncertainty on the integrated luminosity, which is taken to be 10%. A complementary approach is also presented for the early phase of operation which relies on jets reconstructed in the tracker and does not make use of the missing transverse energy observable. A good signal sensitivity can be obtained with such a selection which relies minimally on the calorimeter measurements.

3. Prospects for $t\bar{t}$ measurement in the lepton+jets channel

The top-antitop lepton-plus-jets channel, where one W boson decays leptonically and the other hadronically benefits from a larger branching ratio (45% for all lepton flavors) than the dilepton channel and a reduced background with respect to the all-hadronic $t\bar{t}$ channel. This channel represent thus a good compromise between visibility and statistics for the observation of $t\bar{t}$ pairs. Moreover if the three jets originating from the hadronic decaying top-quark can be identified the top quark can be reconstructed and exploited for the cross section measurement.

3.1 Prospects for CMS

CMS explored two feasibility studies to measure the $t\bar{t}$ cross section in the μ+jets and e+jets channels separately, with 20 pb^{-1} of collected data. The goal of both analyses is to design a simple and robust method to identify top-quark pairs with the lowest possible integrated luminosity.

In the first analysis [6] events are selected which contain one highly energetic isolated muon with $p_T > 20$ GeV and at least four jets with $p_T > 30$ GeV. No E_T^{miss} criteria is required since this measured quantity was assumed poorly known for the early data. Several techniques are used to identify the three jets originating from the hadronic top decay. Figure 2 (left) shows the reconstructed hadronic invariant top mass based on a $\chi^2 - sorting$ method where jet permutations are chosen such as to minimize a χ^2 between the reconstructed hadronic and leptonic top masses, the reconstructed dijet W boson mass and their nominal values.

Two different methods are presented in which the amount of QCD background is extracted in a data-driven way. One is using two uncorrelated variables which distinguish between signal and QCD background (ABCD method), whereas the other one employs an extrapolation of the muon isolation variable. Both methods give a precision of about 50% on the amount of predicted QCD background. The $t\bar{t}$ cross section is extracted using a binned likelihood template fit to observables (such as the distribution shown in Fig. 2) which are sensitive to the composition of the sample in terms of $t\bar{t}$ signal and background processes. The statistical and main systematical uncertainties are addressed employing ensemble tests. With 20 pb^{-1} of data, the $t\bar{t}$ pair production cross section in the muon plus jets channel is expected to be measurable with 12-18% statistical and 20-25% systematic error, where the latter is dominated by the jet energy scale uncertainty. An alternative method is also proposed to subtract the background from W boson+jets production from the selected event sample using the charge asymmetry in W events. The total uncertainty expected with this method is 30%, for 100 pb^{-1}.

The precision on the $t\bar{t}$ cross section measurement in the electron-plus-jets channel was estimated for 20 pb^{-1} using a similar strategy [7]. Data-driven methods were explored to estimate the contributions from QCD and W+jets processes. The QCD background was derived from a one-dimensional extrapolation method using an electron isolation distribution. For W+jets, a template fit method using a discriminating variable (invariant mass of a three-jet combination) was studied. Results of this analysis in the electron channel indicate that the $t\bar{t}$ cross section can be measured with a statistical uncertainty of about 20% and a systematic uncertainty of 20% (not including a 10% uncertainty assumed on luminosity).

3.2 Prospects for ATLAS

The prospects for measuring the $t\bar{t}$ cross-section with the ATLAS detector using final states consisting of a single lepton (e or μ) plus jets were studied for a luminosity corresponding to 200 pb^{-1} of data [8]. Methods for obtaining the cross-section based on two simple but different event selections are explored, neither of which require the ability to tag jets originating from a b-quark and one of which does not require the use of a missing energy cut. These event selections can be used to clearly demonstrate the existence of top quarks in the early data.

A method to measure the dominant W+jets background from the data, based on expected and observed W/Z ratio in one-jet and four-jets events, is developed. This method allows to control the W+jets background to within 20% and to significantly reduce the systematic uncertainties of the $t\bar{t}$ cross-section measurements.

The signal cross section is measured in the electron and muon channels separately using either a simple cut and count method or a template fit to the reconstructed hadronic top mass (see Figure 2 right). These methods yield a statistical precision of 3% to 15% and a systematic uncertainty of less than 15% (not including the luminosity uncertainty), assuming the Standard Model scenario and 200 pb^{-1} of data.

PoS(DIS 2010)154

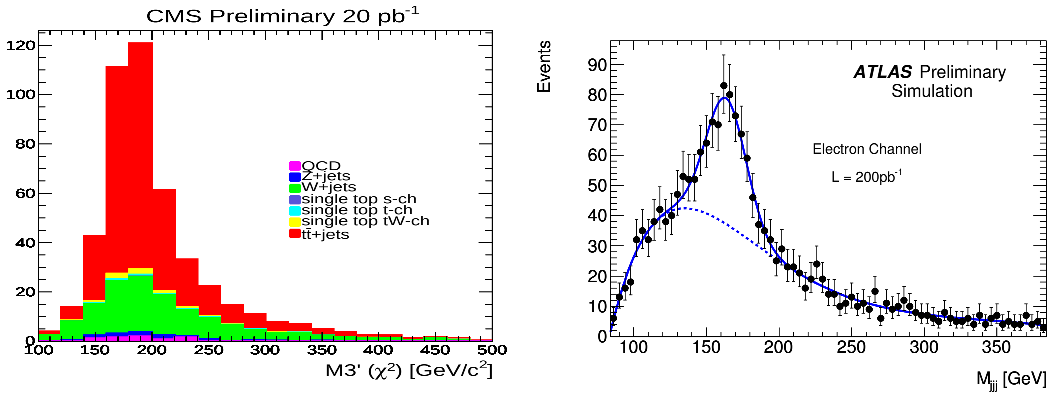

Figure 2: Left: expected reconstructed hadronic top mass in CMS using the best χ^2 combination, with 20 pb^{-1} of data. Right: likelihood fit in the three-jet invariant mass for the electron channel, for 200 pb^{-1} of data collected by ATLAS.

4. Conclusion

We presented the prospects for the ATLAS and CMS experiments to measure the top-antitop cross section in the dilepton and lepton-plus-jets channel with early data at LHC (10-200 pb^{-1}). These studies, based on 10 TeV center-of-mass proton-proton collisions, show that the observation of the top quark is feasible with a few tens of pb^{-1}. For the 7 TeV center-of-mass collisions the LHC is currently performing the situation is a little less favorable since the signal cross section is reduced by a factor 2.5 while the main backgrounds (QCD, W/Z+jets) are reduced by factors of 1.5-2.0. These different operating conditions emphasize the importance of precise measurements of the main background rates and the role of *b*-jet identification algorithm in future analysis.

References

[1] CDF Coll., *Phys. Rev.* **D 50** (1994) 2966; CDF Coll., *Phys. Rev. Lett.* **74** (1995) 2626; CDF Coll., *Phys. Rev. Lett.* **73** (1994) 225; D0 Coll., *Phys. Rev. Lett.* **74** (1995) 2632.

[2] CDF Collaboration, *Conference Note* **9913** (2009).

[3] CDF Collaboration, *Phys. Rev. Lett.* **103** (092002), 2009; D0 Collaboration, *Phys. Rev. Lett.* **103** (092001), 2009.

[4] ATLAS Collaboration, *ATLAS Note*, ATL-PHYS-PUB-2009-086.

[5] CMS Collaboration, *CMS Physics Analysis Summary*, CMS PAS Top-09-002.

[6] CMS Collaboration, *CMS Physics Analysis Summary*, CMS PAS Top-09-003.

[7] CMS Collaboration, *CMS Physics Analysis Summary*, CMS PAS Top-09-004.

[8] ATLAS Collaboration, *ATLAS Note*, ATL-PHYS-PUB-2009-087.