PROCEEDINGS
OF SCIENCE

Expectations for first single-top studies in CMS in proton-proton collisions

Jeannine WAGNER-KUHR[*]

on behalf of the CMS Collaboration
Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
E-mail: Jeannine.Wagner-Kuhr@kit.edu

The first long physics run of the LHC is taking place at a center-of-mass energy of 7 TeV, and is expected to go on until an integrated luminosity of $1\,\text{fb}^{-1}$ will have been collected. We present an analysis technique to measure the t-channel cross section for single top quark production at CMS, that can confirm the recent observation of single top quark production by the Tevatron experiments. Events leading to a signature of exactly one muon and two jets are selected and specific data-driven methods have been developed to reduce the sensitivity to the unknown level of background contamination.

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

---

[*]Speaker.

1. Introduction

In 1995 the top quark was discovered at the Tevatron proton-antiproton collider at Fermilab by the CDF and DØ collaborations [1]. The Large Hadron Collider (LHC) started operations in March 2010 with proton-proton collisions at a center of mass energy of 7 TeV. At the hadron colliders, Tevatron and LHC, the top quarks are dominantly pair produced via strong interactions. Beside top pair production top quarks can be singly produced in electroweak interaction. In spring 2009, single top quark production has been discovered by both Tevatron experiments [2]. Measuring the single top quark production cross section allows for a direct determination of the CKM matrix element V_{tb}.

The theory of electroweak interaction predicts three different mechanisms for the production of single top quarks: t-channel, s-channel, and associated tW production. At the LHC, the single top quark cross section via the t-channel is predicted to be about a factor five larger than that via the second largest channel (tW). In addition, the t-channel has a clear signature and hence it is considered to be the ideal channel to establish single top quark production in pp-collisions at the LHC. Throughout this article, the t-channel mode is considered signal while the other two single top quark production modes are considered background.

The top quark decays with nearly 100% branching fraction into a b-quark and a W-boson. The Monte-Carlo study described in this article uses the muonic decays of the W-boson ($W \rightarrow \mu\nu$), which have a branching ratio of about 11%. A scenario of an integrated luminosity of $200\,\text{pb}^{-1}$ recorded by the CMS detector [3] at a center of mass energy of 10 TeV is assumed. Although the predicted NLO cross section for t-channel production of $124\,\text{pb}$ [4] at $\sqrt{s} = 10$ TeV is reduced by about a factor of two at $\sqrt{s} = 7$ TeV, most of the analysis strategies presented in this paper are also valid at 7 TeV. We note the differences between the two when they are expected.

2. Modeling of events and selection of single top candidates

The single top quark production, $t\bar{t}$+jets and W/Z+jets production are modeled using the MADGRAPH event generator [5] for the generation of the matrix element. The PYTHIA [6] generator is used for hadronization and showering. In order to give a fair approximation of the full next-to-leading order (NLO) properties of t-channel production, the $2 \rightarrow 3$ diagram, corresponding to the dominant NLO contribution to the t-channel, is combined with the leading order (LO) $2 \rightarrow 2$ process based on the transverse momentum distribution of the spectator b-jet [7]. To avoid double counting between matrix element and parton shower generated jets in case of $t\bar{t}$+jets and W/Z+jets production, a procedure based on the so called "MLM prescription" [8] is used. The remaining background samples, due to di-boson production (WW, WZ, ZZ) and multi-jet QCD production enriched in events with muons coming from the decay of b and c quarks or long-lived hadrons, were simulated using PYTHIA Monte-Carlo. All generated events undergo a full simulation of the CMS detector response.

All events are required to pass the high-level single muon trigger, which includes a 15 GeV transverse momentum threshold. Further, exactly one reconstructed and well isolated muon with

2

a transverse momentum $p_{T,\mu} > 20\,\text{GeV/c}$ and pseudorapidity $|\eta_\mu| < 2.1$ is required. A cut on the muon isolation is applied to reduce QCD multi-jet background by a factor of about 50. Jets are defined according to the iterative cone algorithm with a cone size of 0.5 [9]. We consider jets within $|\eta| < 5$ and calibrated transverse momentum greater than $30\,\text{GeV/c}$. Since most of the t-channel events have two such jets, the event is accepted only if it has exactly two reconstructed jets. For the analysis at 7 TeV, we will use the anti-k_T algorithm instead of the iterative cone algorithm. The anti-k_T algorithm combines the nice features of infrared safety and cone-like shapes.

An important ingredient of the single top quark event selection is the identification of b-jets (b-tagging). We use the track counting (TC) algorithm for identifying b-jets. The TC algorithm calculates the signed 3D impact parameter significance (IP/σ_{IP}) of all the tracks associated with a jet, orders them by decreasing values of this observable, and defines the value of IP/σ_{IP} for the second (high-efficiency TC) or third (high-purity TC) track as the b-tagging discriminator for that jet. The event is accepted only if exactly one of the selected jets passes a tight threshold on the high-purity TC. This requirement reduces W/Z boson events with associated light flavor jets by 3 orders of magnitude. Since we expect most of the signal events to have only one b-quark inside the tracker acceptance ($|\eta| < 2.5$), we reject the event if the remaining jet passes a loose threshold on the high-efficiency TC. In Figure 1, IP/σ_{IP} is presented for $\sqrt{s} = 900\,\text{GeV}$ data together with Monte Carlo simulation. Due to the reasonable description of b-tagging variables in the data by Monte Carlo simulation, the use of a more sophisticated b-tagger is envisaged for the analysis of the data collected at 7 TeV.

To further suppress contributions from processes where the muon does not originate from a leptonically decaying W-boson, we select events with a transverse W-boson mass $M_T > 50\,\text{GeV/c}^2$. Applying this selection, a total of 102 single top quark t-channel events and 229 background events are expected in a sample with an integrated luminosity of $200\,\text{pb}^{-1}$ collected at 10 TeV. This yields an expected signal fraction of about 30%. The dominant background contribution arises from top-quark pair production ($\sim 40\%$ of all events). At 7 TeV, we expect that the signal fraction stays almost the same, but the top pair production fraction is expected to be slightly smaller due to the large mass of the top quark.

The expected number of QCD multi-jet events is 12 events, small compared to other backgrounds. As the model uncertainty on QCD multi-jet events is large, a data-driven method to estimate this background has been developed. The cut on M_T is relaxed and a fit to the M_T distribution is performed (see Figure 1, right panel). The fit function is the sum of parameterizations for signal-like processes (W/Z+jets, $t\bar{t}$, single top) $S(M_T)$ and for the QCD multi-jet process $B(M_T)$ where only the normalizations for both processes (a, b) are free parameters. The parameterization for the signal-like processes is obtained from a selected Z+jets sample, while the parameterization for the QCD background is obtained from a sample without applying any b-tagging requirement and with an inverted isolation cut. By comparing the QCD multi-jet contribution determined in fits with parameterizations obtained from different control samples (W+jets enriched sample, MC prediction) to the actual value, an uncertainty of this method of about 45% is estimated.
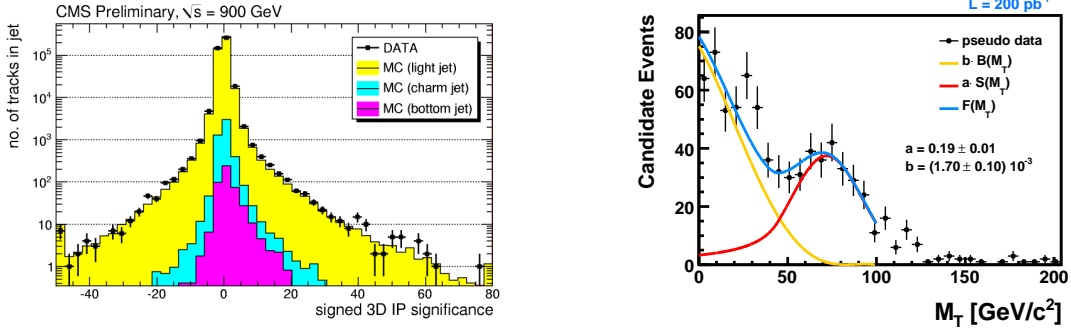
Figure 1: Left: 3D impact parameter significance for $\sqrt{s} = 900\,\text{GeV}$ data. Right: Estimation of QCD multi-jet background by performing a fit to the transverse mass of the W-boson M_T.

3. Prospects for observing single top quark production using the polarization angle

To reconstruct the four-momentum of the top quark, the muon, the missing transverse energy, and the jet tagged as *b*-jet are used. The W-boson is constructed from the four-vector of the muon and the missing transverse energy vector, which is assumed to represent the transverse momentum vector of the neutrino. Using the known W-boson mass as constraint, a quadratic equation in $p_{z,\nu}$ is obtained. In case of a complex solution, $p_{x,\nu}$ and $p_{y,\nu}$ are varied such that the imaginary part vanishes and only the real part remains. In case of two real solutions, the solution with the smallest $|p_{z,\nu}|$ is chosen. This choice is correct in 70% of all cases. To reconstruct the top quark, the *b*-jet is assigned to the W-boson, which is the correct assignment 92% of the times. The reconstructed four-momentum of the top quark is then used to construct a kinematic variable which discriminates well between signal and background.

In particular, we exploit the fact that the top quarks in electroweak single top quark production are almost 100% left-handed polarized [10]. The information of the top quark spin is passed to its decay particles and can be measured by the variable $\cos\theta^*_{lj}$, where θ^*_{lj} is the angle between the charged lepton and the light quark jet in the top quark rest frame. Figure 2 shows the distribution of the reconstructed polarization angle, $\cos\theta^*_{lj}$, for events passing the event selection.

To extract the signal content, we perform a binned likelihood fit to $\cos\theta^*_{lj}$. To be less sensitive to the kinematic effects due to the cuts on muon p_T and M_T at $\cos\theta^*_{lj} \approx 1$, the fit range is restricted to $[-1, 0.75]$. The signal template is taken from simulation, while the overall background is assumed to be flat. This assumption has been verified with background-enriched control samples. Due to the flatness of all backgrounds the method is insensitive to the detailed composition of the background. The statistical sensitivity of the signal extraction has been determined by simulating 500,000 pseudo-experiments. This procedure yields a 35% statistical uncertainty on the cross section for a data sample equivalent to $200\,\text{pb}^{-1}$ at 10 TeV center of mass energy. Assuming that the true value is the one predicted by the standard model, the expected sensitivity of the method is 2.8σ. The inclusion of systematic uncertainties coming from PDFs and from detector knowledge contribute an additional 14%, and the expected sensitivity is lowered to 2.7σ. In Figure 2 the expected sensitivity (stat. uncertainties only) as a function of the integrated luminosity is shown for
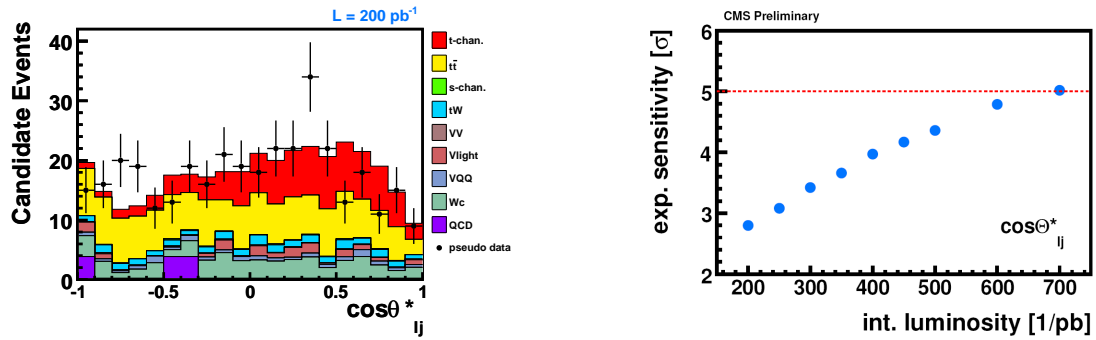
Figure 2: Left: expectation for the reconstructed polarization angle $\cos\theta_{lj}^*$ assuming an integrated luminosity of $200\,\mathrm{pb}^{-1}$ at a center of mass energy of 10 TeV. Right: prospects for single top quark production at 10 TeV.

$\sqrt{s} = 10$ TeV. Using a naive rescaling, we expect that a factor 2 larger integrated luminosity is needed at 7 TeV to obtain the expected sensitivity as demonstrated in the 10 TeV study.

4. Conclusion and outlook

In this article, an analysis strategy to measure the t-channel single top quark production cross section is presented [11]. The analysis strategy exploits the polarization ($\cos\theta_{lj}^*$) of singly produced top quarks. It turns out that $\cos\theta_{lj}^*$ is robust against systematic uncertainties. Using a naive rescaling, evidence of single top quark production in *pp* collisions assuming an integrated luminosity of $1\,\mathrm{fb}^{-1}$ at 7 TeV seems to be well within reach. With our ongoing improvements (inclusion of *e*-channel, use of a more sophisticated *b*-tagger and the combination of several variables) a rediscovery of single top quark production at $1\,\mathrm{fb}^{-1}$ at 7 TeV seems likely.

References

[1] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 2626 (1995); S. Abachi *et al.* (DØ Collaboration), Phys. Rev. Lett. **74**, 2632 (1995).

[2] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **103** (2009) 092002; V. M. Abazov *et al.* (DØ Collaboration), Phys. Rev. Lett. **103** (2009) 092001.

[3] CMS Collaboration, JINST 0803, S08004 (2008).

[4] J. M. Campbell, R. Frederix, F. Maltoni, F. Tramontano, JHEP 0910, 042 (2009).

[5] F. Maltoni and T. Stelzer, JHEP **0302**, 027 (2003).

[6] T. Sjostrand *et al.*, JHEP **05**, 026 (2006).

[7] E. Boos *et al.*, Phys. Atom. Nucl. **69, 8**, 1317 (2006).

[8] J. Alwall *et al.*, Eur. Phys. J. **C53**, 473 (2008).

[9] UA1 Collaboration, Phys. Lett. **B123**, 115 (1983).

[10] G. Mahlon and S. Parke, Phys. Rev. **D55**, 7249 (1997).

[11] CMS Collaboration, PAS TOP-09-005
(http://cms-physics.web.cern.ch/cms-physics/public/TOP-09-005-pas.pdf).