CMS Analysis Note

The content of this note is intended for CMS internal use and distribution only

February 20, 2009

Towards a measurement of single-top cross section in the t channel, with the first inverse femtobarn of CMS data at 14 TeV

A. Giammanco

Université Catholique de Louvain, Center for Particle Physics and Phenomenology, Louvain-la-Neuve, Belgium
Fonds National de Recherche Scientifique, Belgium

J. Weinelt, T. Müller, J. Wagner-Kuhr, P. Sturm[a)], W .Wagner[a)]

Institut für Experimentelle Kernphysik, Universität Karlsruhe, Germany

D. Konstantinov, V. Molchanov

Institute for High Energy Physics, Protvino, Russia

A. Jafari, M. Mohammadi Najafabadi

IPM and Sharif University of Technology, Tehran, Iran

E. Boos, L. Dudko, A. Markina

Skobeltsyn Institute of Nuclear Physics, Moscow State University, Moscow, Russia

Abstract

We report on a study aiming at an early observation of single-top events produced in t channel, at a centre-of-mass energy of $\sqrt{s} = 14$ TeV.

Different analysis methods are proposed and a technique for estimating the multi-jet QCD background from data is developed.

The simulation scenario and the detector systematic uncertainties in this note assume a detector knowledge as foreseen to be available at the time when the first inverse femtobarn will be on tape. With the corresponding statistics, an evidence for signal at a confidence level of more than 3σ is found to be reachable.

[a)] Now at Bergische Universität Wuppertal, Germany

1 Introduction

The theory of electroweak interactions predicts three different production mechanisms for single top quarks in hadron-hadron collisions, supplementary to the more abundant pair production due to the strong interaction. They are classified by the virtuality of the involved W boson: s-channel production ($q^2_W > 0$, Fig. 1(a)), W-associated, or tW, production ($q^2_W = M^2_W$, Fig. 1(b)) and t-channel production ($q^2_W < 0$, Fig. 1(c)). Recently the D0 and CDF experiments at the Tevatron $p\bar{p}$ collider both provided a 3σ evidence of the existence of the electroweak mode of production (see, e.g., Refs. [1, 2]), and are expected to provide a 5σ observation very soon thanks to the fast accumulation of integrated luminosity. At the Large Hadron Collider (LHC) the reobservation is expected to happen first in the t-channel mode, by far the most abundant of the three at the energies soon to be available at the new machine, and the one with the most striking final state topology. This note treats this production mode as signal, including the other two in the definition of background.

The study of single top quark production provides a unique possibility to investigate many aspects of top-quark physics that cannot be easily studied in $t\bar{t}$ production. All the three channels are directly related to the modulus squared of the CKM matrix element V_{tb}, allowing for a direct measurement of this quantity and thus for a further test of the Standard Model [3]. One can investigate the tWb vertex structure and FCNC couplings in the production processes, and the single-top topologies are also a window for searches for anomalous couplings and s-channel resonances like W' bosons. A review of many opportunities to observe new physics from deviations in the expected cross sections of the t- and s-channel modes can be found, for example, in Ref. [4].

At a proton-proton collider like the LHC, the cross sections for single t and $\bar{t}$ quark production are not equal in the t and s channels. The cross section for $\bar{t}$ quark production is suppressed due to the fact that the initial state light quark has to be an anti-quark, which is only possible in the form of a sea quark in the incoming protons. The NLO cross sections, for a top-quark mass of 171 GeV and a centre-of-mass energy of 14 TeV, are 150 pb and 92 pb for t-channel top and anti-top quark production respectively [5], and 6.55 pb and 4.07 pb in the s-channel [6]. Only in the tW production mode, where the final state top-quark content is independent of the initial-state quark content, the cross sections for top and anti-top quarks are equal, each around 33 pb at NLO [7]. The different contributions add up to a total cross section of electroweak top-quark production of 319 pb.

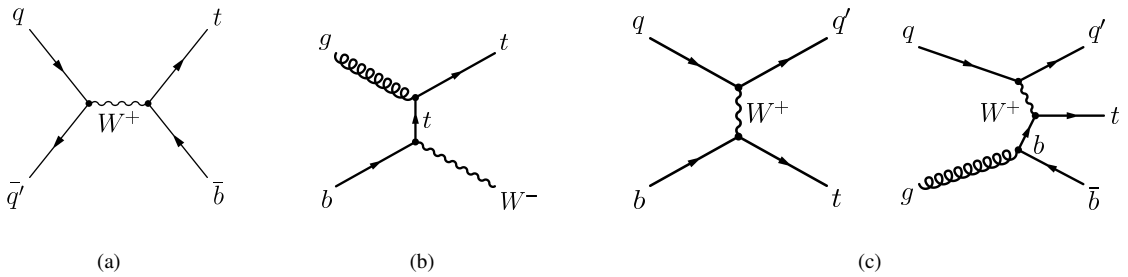


Figure 1: Leading order Feynman diagrams for single top quark production in s-channel (a), W-associated or "tW" (b), and t-channel production (c), the latter is shown in the LO description together with the dominating NLO diagram.

The goal of this note is to present many recent improvements with respect to the study of Ref. [8], from the points of view of the statistical performance, of the robustness against backgrounds and of the realism. A more extensive list of background samples was considered, the simulation was improved, and more recent reconstruction and software tools were used. This study is performed at a pp centre-of-mass energy of 14 TeV and concentrates on the muonic decay channel, in which the W boson from the top quark decays further into a muon and a neutrino.

This paper is organized as follows: Sec. 2 describes the Monte Carlo samples and the software framework used; Sec. 3 discusses the event selection; Sec. 4 presents a data-driven technique for estimating the QCD background; Sec. 5 explains in details how the systematic uncertainties have been estimated; Sec. 6 presents the cross section extraction based on a simple event counting, as well as a new technique to extract the signal in a robust way by exploiting the characteristic $t/\bar{t}$ asymmetry; Sec. 7 discusses the reconstruction of top quarks in the selected sample and shows how to confirm the presence of signal by studying the reconstructed mass and the polarization of the top quark; Sec. 8 presents an alternative selection based on the use of Neural Networks for the event classification, exploiting several kinematical features expected for the signal; Section 9 draws the conclusions and presents the plans for future analyses.

2 Data samples and software framework

The t-channel Monte Carlo events used in this study have been generated with two different generators: MadGraph [9], and the SingleTop generator [10] based on CompHEP [11]. In order to give a fair approximation of the full NLO properties of the signal, the dominant NLO contribution to the t channel (Figure 1(c), right-hand side) is combined with the LO 2→2 process (Figure 1(c), left-hand side) by a special procedure, obtaining a merged sample that describes the common phase space as completely as possible while avoiding double counting. The matching method is based on the indications of Ref. [10]. The separately generated sub-samples for the 2→2 and 2→3 processes are matched in such a way to give a smooth $p_T^{2nd\ b}$ spectrum, choosing an optimal matching threshold in this variable. This threshold is determined under the assumption that the soft transverse momentum region of the additional b quark is best described by the 2→2 process, whereas the modeling of the high-p_T tail of the spectrum by the 2→3 process. More details on the normalization of the spectrum as well as a comparison of the outcome to other generators, including MC@NLO [12], can be found in Ref. [13]. This comparison resulted in general in a good agreement between the different generators. The technical implementation of the matching in the CMS software in the MadGraph case is realised on top of the MadGraphInterface package.

Both signal samples were used for the studies presented in this note. The standard analysis (Sec. 3-Sec. 7) is based on the MadGraph signal sample, whereas the Neural Network study (Sec. 8) uses the SingleTop signal MC. As documented in the MC comparison note mentioned above, the differences between the two signal samples are very small. Therefore, the fact of having two different signal descriptions is expected to have a negligible impact on the results presented here. In both cases, the detector response was simulated by the CMS fast simulation [14], with settings equivalent to those used for the official samples described below, which are all passed through the full detector simulation based on GEANT 4 [15].

Several Standard Model background processes are taken into account in order to get a realistic observation scenario. The AlpGen [16] generator was used for the parton-level modeling of $t\bar{t}$, $W/Z + jets$, and $W/Z + Q\bar{Q}(Q = b, c)$ [1) 2)], interfaced to PYTHIA [17] for showering and hadronization. A sample for single top in the W-associated production channel (tW) was obtained from TopReX [18], a specialized generator for top-quark physics based on PYTHIA. The μ-enriched QCD events were entirely generated with PYTHIA, with a filter at generator level selecting only interactions with $\hat{p}_T > 20$ GeV that produce muons with $p_T > 15$ GeV and $|\eta| < 2.5$. It is important to point out that neither the contribution of erroneously reconstructed muons due to so-called punch-through hadrons, nor contributions of decay-in-flight muons are simulated. Di-boson samples were generated with PYTHIA, where the production of WW, WZ, and ZZ was considered separately. The contribution of single-top quark production in the s channel is expected to be negligible and therefore ignored in this study.

The aforementioned background samples were produced in the frame of the Computing, Software and Analysis Challenge 2007 (CSA07) [19]. Therefore, certain simulated processes were combined into one mixed data set by the official CMS production. In this study, the so-called "chowder soup" was used, containing $t\bar{t}$ and W/Z+jets events. Other "soups" were produced, containing multi-jet QCD events (in the following only called "QCD" for short) from several sub-processes, but not considered for this analysis since the μ-enriched QCD sample described above is equivalent to a much larger integrated luminosity; it has been checked that no event from the QCD soups survive our complete selection.

Different detector misalignment/miscalibration scenarios were simulated in the CSA07 exercise [20]. In this analysis the "$100\ pb^{-1}$" scenario is assumed, meaning the alignment and calibration precisions that are expected to be obtained from the analysis of the first $100\ pb^{-1}$ of data, as recommended for analyses scaled to $1\ fb^{-1}$. The same scenario has been used in the fast simulation. No pile-up was included in these samples. More information on the CSA07 data sets can be found elsewhere [19].

[1)] The CSA07 production has a problem with double-counting between the $W/Z + jets$ events contained in the "soup" and the $W/Z + Q\bar{Q}$ samples, since the former also contain part of the heavy-flavour contribution from gluon splitting modeled by the parton shower. A technical solution is now available for more recent versions of the CMS software but there are no plans to back-port it to the releases needed to analyze the CSA07 samples. The approach chosen here is to accept, conservatively, the small overestimation of the background coming from ignoring the problem; this is justified *a posteriori* by the fact that $W + X$ events are not a major background after the proposed selection, meaning that this choice will not affect the qualitative conclusions of the study.

[2)] Due to a bug in the generation, the official $Wc\bar{c}$, $Zb\bar{b}$ and $Zc\bar{c}$ samples had no muon in the final state. For $Wc\bar{c}$ we have used a private FastSim production based on the (debugged) generation cards, while we chose to ignore the $ZQ\bar{Q}$ processes in this analysis due to the smallness of their cross section with respect to the corresponding $WQ\bar{Q}$ and the fact that events with two reconstructed muons are vetoed in this analysis, see Sec. 3.

Table 1 summarizes the Monte-Carlo data samples for signal and backgrounds, and provides the number of events and cross section for each sample. The cross section of the $t\bar{t}$ process is normalized to 833 pb [21], which corresponds to a NLO/LO K-factor of 1.85 with respect to the AlpGen cross section. An NLO/LO K-factor of 1.12 is uniformly applied on the $W/Z+jets$ cross sections. For the signal sample generated by MadGraph, the theoretical cross section of 242 pb is multiplied by a branching ratio of 0.324 [22]. Since the signal sample from SingleTop was generated using a top-quark mass of 175 GeV/c^2, the theoretically predicted cross section was rescaled to 235 pb, and then multiplied by a branching ratio of 0.108 for the muonically decaying W boson. The other cross sections are at LO, as obtained from the corresponding generators.

The study presented in this note is performed within the CMS software release 1_6_12. The reconstructed data samples are further processed using the so called layers 0 and 1 of the Physics Analysis Toolkit (PAT) [23]. The PAT is a high-level analysis framework providing common software tools. The layer 0 performs "cleaning" tasks like isolation and duplicate removal of the standard reconstructed objects, and computes related information (e.g., lepton isolation). The output of this layer is a consistent set of AOD objects and associated information. Layer 1 then collapses the output into compact objects. Jet corrections, b tagging, Monte Carlo matching and other high-level tasks are also applied at this level. Making use of these PAT objects, we apply a pre-selection and save the surviving events in a customized ROOT "tree". The final selection is then applied in a second independent step. Both the pre-selection and the final selection criteria are described in the next section.

3 Event selection

This section provides a detailed description of the selection applied to the simulated data in order to suppress backgrounds and enrich signal events. The study presented here focuses on the muonic decay channel. The electronic channel could be considered, but we lack a sufficiently reliable electron-enriched QCD sample; the tauonic and the all hadronic channels have probably no chance to emerge from the QCD background. The final state topology in the t channel is then characterised by exactly one isolated muon and a b jet from the top quark decay, as well as a light flavour jet produced in the forward region. In the following we present the definition of the selected muon candidates, the requirements on the jets, and the tagging method used to identify b jets. The feature of the forward jet is exploited to significantly reduce the remaining background contributions. A sizable amount of QCD events seems to survive the selection chain up to this point, but we are able to control them by an additional requirement on the transverse W boson mass. The outcome of the presented selection is scaled to the expectation corresponding to an analysed data amount of $1 fb^{-1}$.

3.1 Leptons

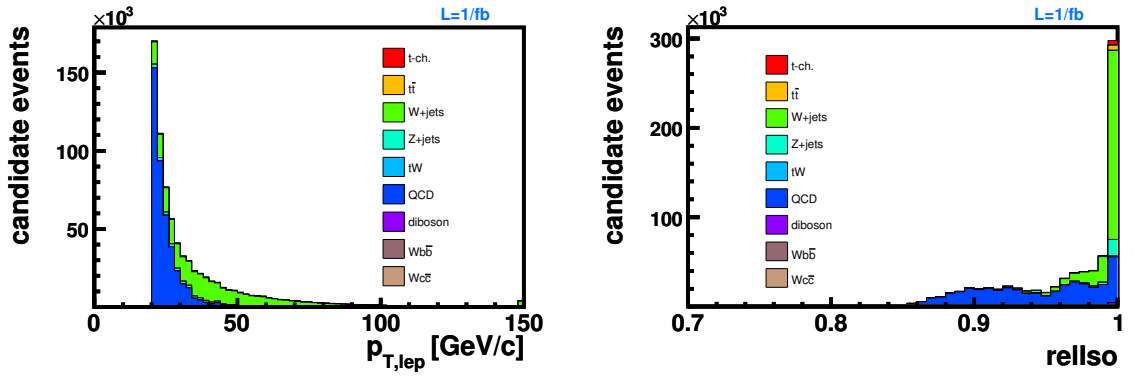
The first analysis step is the definition of a charged lepton. We base this analysis on muons with a transverse momentum $p_{T,lep} > 20$ GeV/c (see Fig. 2(a)) within the trigger acceptance range ($|\eta| < 2.1$). In order to reduce the contribution of background events where no real $W \rightarrow \mu\nu$ was produced, and in particular of QCD events, where the muon candidate can be a fake or can come from the decay of hadrons, we apply absolute isolation requirements on the muon candidates. In the pre-selection step (layer 1) the tracker isolation ($tkIso$), defined as the sum of transverse momenta of tracks with $p_T > 1$ GeV/c within a cone of radius ΔR$= 0.3$ around the lepton candidate's track, has to be smaller that 3.0 GeV/c. The isolation in the calorimeter ($caloIso$), defined as the sum of E_T of the ECAL and HCAL deposits with $E_T > 1$ GeV inside a Δ R< 0.3 cone around the lepton direction, is required to be below 1.0 GeV.

In order to reduce the contribution of dilepton events, which can come from $t\bar{t}$, we veto events with more than 1 lepton, where also electrons are considered in the count. Electron candidates are defined, in CMS, by a basic set of requirements in ECAL and in the Tracker [24]. These criteria are quite loose and serve as the starting point of an additional classification, where three different electron categories are distinguished: loose, medium and tight; a detailed description can be found in Ref. [25]. For the purpose of top-quark analyses the tight electron category has shown to give the best fake electron rejection at a reasonable cost in efficiency [26]. Therefore, in this analysis only electron candidates of the tight ID category, and passing the same isolation criteria as the muons, enter the lepton count.

After this selection a considerable amount of QCD events remains, as can be seen from Figure 2; we define the combined relative isolation as $relIso = p_{T,lep}/(p_{T,lep} + tkIso + caloIso)$, shown in Fig. 2(b) for the selected muon candidates in the 2 jet sub-sample. In contrast to the processes containing a muon in the final state, the QCD contribution doesn't accumulate at 1. Therefore, the QCD rejection can effectively be improved by requiring the relative isolation of the muon candidate to be above 0.95.

Process	$\sigma[pb]$	Tot. events	Generator	Dataset name
single top t channel ($W \to l\nu, l = e, \mu, \tau$)	77.4 (NLO)	209,909	MadGraph	Private FastSim sample
single top t channel ($W \to l\nu, l = \mu$)	25.4 (NLO)	60,000	SingleTop	Private FastSim sample
$t\bar{t}$	833 (NLO+NLL)	1,775,038	AlpGen	
$W + jets$	61,400 (NLO)	20,437,728	AlpGen	/CSA07AllEvents/CMSSW_1_6_7-CSA07-Tier0-A4-Chowder/AODSIM
$Z + jets$	5,300 (NLO)	4,466,181	AlpGen	
tW	62 (LO)	131,000	TopReX	/tW_inclusive/CMSSW_1_6_7-CSA07-1198096703/AODSIM
$Wbb0j(W \to l\nu, l = e, \mu, \tau)$	5.1 (LO)	15,000	AlpGen	/Genwbb-alpgen/CMSSW_1_6_7-CSA07-1211715927/RECO
$Wb\bar{b}1j(W \to l\nu, l = e, \mu, \tau)$	7.4 (LO)	16,162	AlpGen	/Genwbbj-alpgen/CMSSW_1_6_7-CSA07-1211715980/RECO
$Wb\bar{b}2j(W \to l\nu, l = e, \mu, \tau)$	3.5 (LO)	15,894	AlpGen	/Genwbbjj-alpgen/CMSSW_1_6_7-CSA07-1211716032/RECO
$Wb\bar{b}3j(W \to l\nu, l = e, \mu, \tau)$	3.8 (LO)	11,543	AlpGen	/Genwbbjjj-alpgen/CMSSW_1_6_7-CSA07-1211716084/RECO
$Wc\bar{c}0j(W \to l\nu, l = e, \mu, \tau)$	5.6 (LO)	21,000	AlpGen	Private FastSim sample
$Wc\bar{c}1j(W \to l\nu, l = e, \mu, \tau)$	7.5 (LO)	16,958	AlpGen	Private FastSim sample
$Wc\bar{c}2j(W \to l\nu, l = e, \mu, \tau)$	4.4 (LO)	10,057	AlpGen	Private FastSim sample
$Wc\bar{c}3j(W \to l\nu, l = e, \mu, \tau)$	4.0 (LO)	7,451	AlpGen	Private FastSim sample
WW	128 (LO)	744,261	PYTHIA	/WW_incl/CMSSW_1_6_7-CSA07-1198096665/AODSIM
WZ	52 (LO)	362,291	PYTHIA	/WZ_incl/CMSSW_1_6_7-CSA07-1198096684/AODSIM
ZZ	18.9 (LO)	143,113	PYTHIA	/ZZ_incl/CMSSW_1_6_7-CSA07-1198096655/AODSIM
QCD (μ-enriched)	226,000 (LO)	1,991,330	PYTHIA	/ppMuPt20-15/CMSSW_1_6_7-CSA07-1204357236/RECO

Table 1: Datasets used in this analysis. The samples are generated either inclusively or with a final state restricted to the leptonic mode, including electrons, muons, and taus. The cross sections are given according to the event content. The references for the cross sections obtained from generators can be found in Ref. [19].



(a) (b)

Figure 2: (a) Transverse momentum distribution of the pre-selected muons ($p_{T,lep}$, $|\eta|$) in the inclusive jet sample. (b) Relative isolation ($relIso$) of the pre-selected muons in the 2 jets sample.

3.2 Jets

Jets are reconstructed using the iterative cone algorithm with a cone size of 0.5. This choice is only historical and doesn't come from an optimization; nevertheless, previous studies in the $t\bar{t} \rightarrow 2l + X$ channel [27], whose final state has the same number of partons as our signal, suggest that the choice of the algorithm and of its resolution parameter has a minor impact on the analysis. The jet energy is scaled by a factor that describes the detector response depending on the transverse energy and the pseudo-rapidity of the jet. In this analysis the transverse momentum of the calibrated [32] jet has to be greater than 30 GeV/c, and the absolute value of the pseudo-rapidity smaller than 5.0. Furthermore, a jet is removed from the collection if it is closer than $\Delta\mathrm{R} = 0.3$ to the selected charged lepton. The resulting jet multiplicity is shown in Fig. 3. Since for most of the signal events two jets are reconstructed with these conditions, the study focuses on the 2-jets sub-sample. It can also be seen that at this stage of the selection the sample is still dominated by processes without b quarks. This problem is specifically addressed in the next subsection.

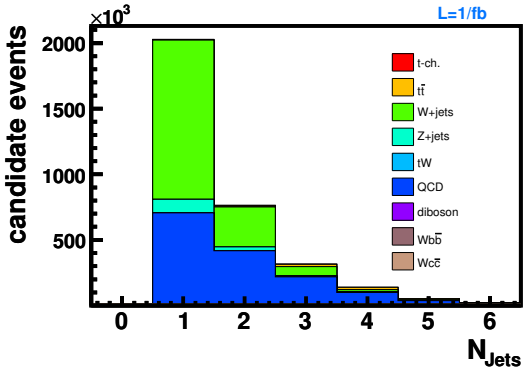


Figure 3: Jet multiplicity after the leptonic pre-selection ($p_{T,lep}$, $|\eta|$) and the jet-lepton ambiguity removal; only events with at least one jet are shown.

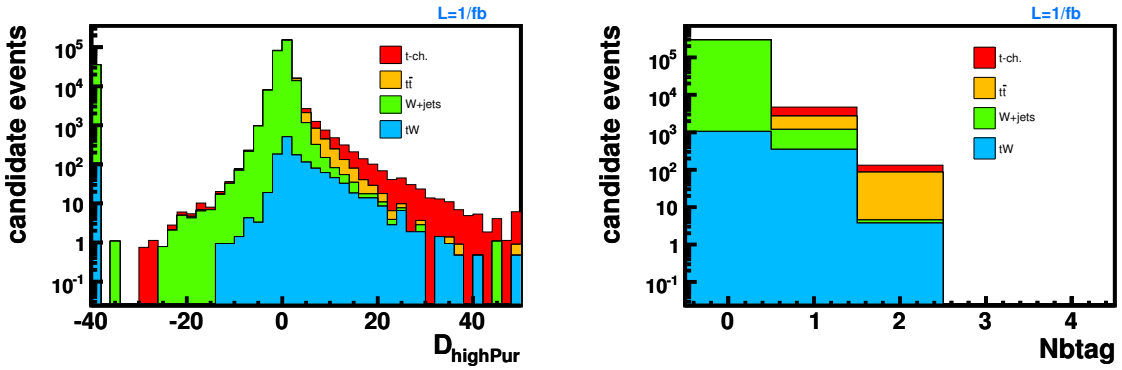
3.3 b tagging

Several b tagging algorithms are available in CMSSW. Some exploit the long B-hadrons lifetime, others their semi-leptonic decay modes and others use kinematic variables related to the high B-meson mass and hard b-quark fragmentation function. Details are provided elsewhere [33]. For this study we use the *high-purity track counting* (HPTC). It calculates the signed 3D impact parameter significance (IP/σ_{IP}) of all the tracks associated to the jet

that pass tight quality criteria, orders them by decreasing values of this observable, and outputs as jet discriminator the value of IP/σ_{IP} for the third track. The highest discriminator value in the event is shown in Fig. 4(a), where the underflow bin contains all events where no jet is found with enough high-quality tracks to allow the computation of this discriminator.

The b-tagging physics object group (POG) proposes a set of three reference thresholds (tight, medium and loose working points) for each algorithm [33], and for the track-counting family the tight working point corresponds to choose the high-purity algorithm with threshold set to 4.8, yielding a mistag rate of 0.10% for uds quarks in QCD events, and a b-tag efficiency of around 32% for the same kind of events.

The signature of the t-channel process includes 3 partons in the final state: one light quark recoiling against the virtual W boson, one b quark from the top-quark decay, and a second b quark from the initial gluon splitting. Since the second b quark is most likely produced at very high rapidities, i.e., outside the tracker acceptance of $|\eta| < 2.5$ and thus not allowing b tagging to be performed, we expect most signal events to have only one b-tagged jet. The b-tag multiplicity of the dominant processes is shown in Fig. 4(b). The contribution of processes without b quarks in the final state is strongly suppressed in the 1-tag sub-sample, showing the largest population of signal events at the same time. Therefore, selected events are required to have exactly one b-tagged jet.



(a) (b)

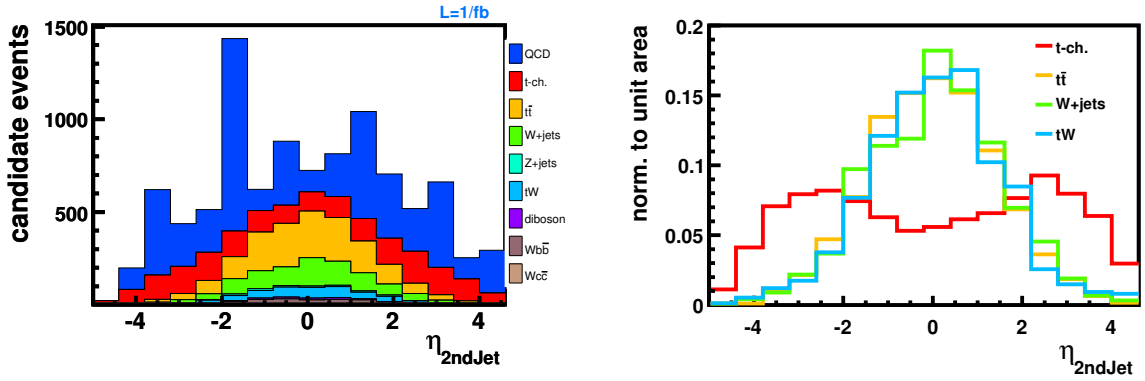
Figure 4: Highest discriminator value for the high-purity track-counting algorithm (a) and number of tags for $D > 4.8$ (b) after the preceding selection, corresponding to the forth column of table 2 ("$N_{jets} = 2$"). The underflow bin in (a) represents events where no jet has enough good tracks to calculate this discriminator (≥ 3 good tracks are required).

3.4 Forward jet

One striking feature of the single top t-channel process is the presence of an energetic light jet in the final state, stemming from the quark recoiling against the virtual W boson. This jet tends to have a rather forward direction instead of showing a maximum at $\eta = 0$ as most of the hard-interaction processes.

As mentioned above, this analysis exploits the 2-jets sub-sample, whereof one jet is identified as a b jet. Figure 5(b) compares the shape of the pseudo-rapidity of the second, non-b-tagged, jet. A characteristic difference is visible between the signal, where the second jet tends to be produced in the forward region, and the background processes. Due to the fact that the second jet is produced rather centrally for all the background processes, a veto on a second central jet reduces their contribution by a factor of 10, while almost half of the signal events are retained. The threshold is chosen as $|\eta| > 2.5$, i.e, outside of the tracker acceptance region, meaning that the b-tagging discriminator is undefined.

From Fig. 5(b) it is possible to see that this veto threshold greatly increase the purity of the selection. This figure also shows the QCD contribution as estimated from MC. Although the low MC statistics makes hard to draw quantitative conclusions, it is apparent that in QCD events the second jet tends to be distributed rather flatly in η.



(a) (b)

Figure 5: Pseudo-rapidity distribution of the second jet after the selection corresponding to the fifth column of table 2 ("$N_{btag} = 1$"), normalized to (a) $1 fb^{-1}$, and (b) to unit area.

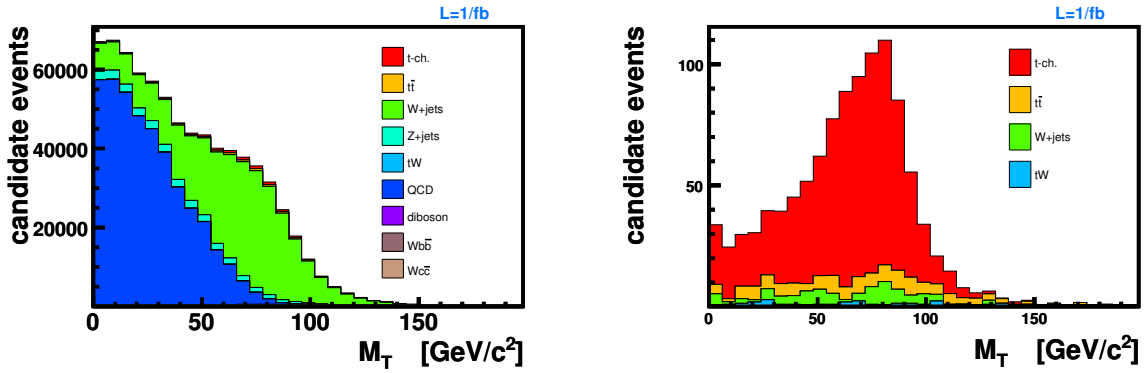
3.5 Transverse W boson mass

To further suppress contributions from processes where the lepton does not come from a leptonically decaying W boson, a selection based on the reconstructed transverse W-boson mass is applied. This quantity is defined as:

$$M_T = \sqrt{\left(p_{T,lep} + p_{T,\nu}\right)^2 - \left(p_{x,lep} + p_{x,\nu}\right)^2 - \left(p_{y,lep} + p_{y,\nu}\right)^2}\,, \tag{1}$$

where the transverse momentum components of the neutrino are approximated by the components of the missing transverse energy, $\vec{\not{E}}_{\mathrm{T}}$. This is corrected for the presence of muons, and the calorimetric clusters associated to jets are given the calibrated jet momenta. More details are given in Ref. [34].

Figure 6 shows the shape of the M_T distribution after leptonic (a) and full (b) selection. The QCD background can be nicely distinguished, since the transverse mass of the alleged W boson accumulates at low values while all processes with real W bosons tend to cluster around the W mass (this feature is know in the literature as "Jacobian peak"). Hence, the reconstructed transverse W-boson mass is required to be above 50 GeV/c^2 for the event to be kept.



(a) (b)

Figure 6: (a) Transverse mass after the leptonic selection ($p_{T,lep}, |\eta|, relIso$) and (b) after the entire selection minus the M_T cut, corresponding to the sixth column of table 2 ("$|\eta_{j2}| > 2.5$").

3.6 Results after 1 fb^{-1}

The relative impact of the presented event selection, step by step, on the signal and the considered background contributions is shown in Table 2. The pre-selected sample is still dominated by QCD, despite the isolation requirements on the track and the calorimeter properties of the muon candidates. This motivates the use of the combined isolation variable $relIso$, and an additional selection on the transverse W-boson mass allows a good QCD reduction. Nevertheless, QCD remains one of the dominant background contributions, and, due to very limited MC statistics, a huge uncertainty on its estimation has to be kept in mind. The second most dominating background after the pre-selection, W+jets, is reduced significantly by the use of b tagging. Remaining background contributions, mostly containing b quarks in the final state and therefore not distinguished from the signal by the b-tagging algorithm, are further suppressed by a veto on a second central jet. The expected final event yield, scaled to an integrated luminosity of 1 fb^{-1}, is shown in the last column of Tab. 2.

It can be seen that one of the dominant backgrounds is $t\bar{t}$, therefore it is interesting to look at the relative importance of the different final states of this process. As shown in Table 3, the subset of $t\bar{t}$ events passing our selection is almost equally composed of $2l$, $\tau + l$ and $l + jets$ ($l = e, \mu$) events.

The same selection, apart from the central second-jet veto, applied to the other jet multiplicities, yields the background contaminations shown in Fig. 7. It can be seen that, despite the b tagging, the 1-jet bin is still dominated by $W + jets$ events. Since the normalization of this process has a large uncertainty, it is envisageable to use this bin to constrain the W+jets contamination in the 2-jets bin.

Based on the estimations for 1 fb^{-1}, $S/\sqrt{B} = 27$ is obtained, while $S/\sqrt{S+B} \approx 18$. These numbers do not take into account systematic uncertainties, which will be discussed in Sec. 5, not to mention the huge uncertainty on QCD coming from the insufficient MC statistics. The significance for the simple event counting will be estimated more realistically, with and without systematics, by the use of pseudoexperiments in sec. 6.2, where also a comparison of the method with the alternative Charge Ratio technique is presented.

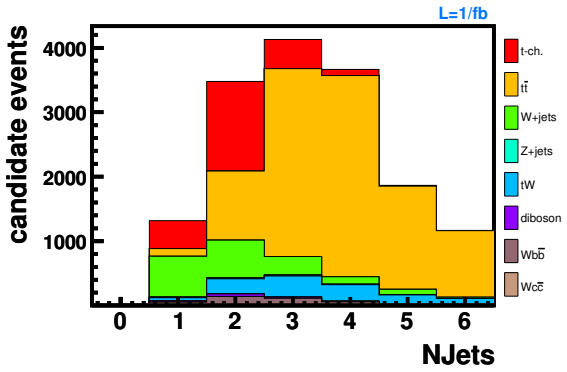


Figure 7: Jet multiplicity after complete selection apart from the central second-jet veto.

4 QCD background estimation from data

The relatively large number of expected QCD events surviving our selection, shown in Table 2, has to be taken with care, since no optimization has been attempted against this specific background due to the lack of an adequate MC statistics. It has to be noted that, for example, two of the three MC events surviving our selection in the QCD sample have a muon p_T slightly higher than our threshold of 20 GeV, so it would be easy to reduce the estimated QCD background by a factor of three in the final column of the table with a small loss of signal, but this would not teach us much since an *a posteriori* bias would have been unavoidably introduced in the analysis.

Moreover, even in case of much larger MC statistics, Monte Carlo estimations of the QCD contamination have to be considered particularly unreliable for the purposes of our analysis, because only very extreme kinematical regions have a chance to pass this kind of selection, and tail effects are the most difficult to simulate properly. These arguments lead to the conclusion that only *in situ* data-driven estimations will give the needed confidence on the amount of this background. In this section we report about one of such methods, the so called "$ABCD$" method [37].

Process	N_{tot}^{MC}	$N_{lep}=1$	$N_{jet}=2$	$N_{btag}=1$	$\|\eta(j_2)\|>2.5$	$M_T > 50$ GeV/c^2	N_{sel}^{MC}	ϵ_{sel}^{tot}	N_{evt} in 1 fb^{-1}
signal	209,909	15.4%	54.2%	30.0%	43.6%	72.4%	1651	0.79%	617
$t\bar{t}$	1,775,038	8.3%	13.5%	20.4%	7.8%	69.4%	195	0.01%	85
W+jets	20,437,728	7.1%	19.9%	0.3%	8.8%	63.9%	46	0.0002%	48
Z+jets	4,466,181	3.5%	21.2%	0.3%	8.7%	66.7%	4	0.0001%	3
tW	131,000	9.4%	24.2%	25.0%	8.3%	66.1%	41	0.03%	19
$Wb\bar{b}$	58,599	9.5%	30.1%	28.5%	3.3%	81.3%	13	0.02%	5
$Wc\bar{c}$	56,116	11.8%	29.9%	2.8%	1.8%	100%	0	0.002%	0
$WW/WZ/ZZ$	1,249,665	5.7%	38.5%	1.4%	9.8%	67.6%	25	0.002%	4
QCD	1,991,330	0.3%	27.9%	2.5%	35.7%	20.2%	3	0.001%	345

Table 2: Relative selection efficiencies calculated for each selection step with respect to the preceding one. The last column shows the number of selected events rescaled to 1 fb^{-1}.

Channel	fraction
$e/\mu + jets$	34%
$e/\mu + e/\mu$	32%
$e/\mu + \tau$	31%
$\tau + \tau$	2%
$\tau + jets$	1%
fully hadronic	-

Table 3: Final state composition of the $t\bar{t}$ events that pass the full selection.

4.1 Principle of ABCD Method

To estimate the QCD background in data we divide the $relIso/M_T$ plane into four regions, where the quadrant defined by $relIso > 0.95$ and $M_T > 50$ GeV corresponds to the "signal region" of this analysis, here labeled A (see Fig. 8). Region B is defined by $relIso > 0.95$ and $M_T < 50$ GeV, C by $relIso < 0.80$ and $M_T > 50$ GeV, and D by $relIso < 0.80$ and $M_T < 50$ GeV.
The basic idea of the $ABCD$ method is that the number of QCD background events in the signal region N_A^{QCD} can be predicted from the relative abundances of QCD events in regions where this background dominates (B, C and D):

$$N_A^{QCD} = \frac{N_B^{QCD} \cdot N_C^{QCD}}{N_D^{QCD}} \,. \tag{2}$$

Here, the assumption enters that $relIso$ and M_T are totally uncorrelated for the QCD background.
It has been proven in a more detailed study performed in Ref. [37] (where no restriction was applied on the number of jets) that the correlation between the relative isolation and M_T is small. Indeed, one of the reasons to prefer M_T over $\not{E}_{\mathrm{T}}$ in the present selection has been the much smaller correlation found between $relIso$ and M_T than between $relIso$ and $\not{E}_{\mathrm{T}}$. This can be heuristically explained by considering that in $Q\bar{Q} \to l\nu X$ events, which constitute a major part of the surviving QCD background, both $p_{T,\nu}$ and $\Delta R(l, Q)$ depend on $p_{T,Q}$, while M_T is boost-invariant.
The choice of non-contiguous ranges in $relIso$ (see Fig. 8) is suggested by the convenience of keeping negligible the contamination of signal-like events in the QCD-enriched regions.

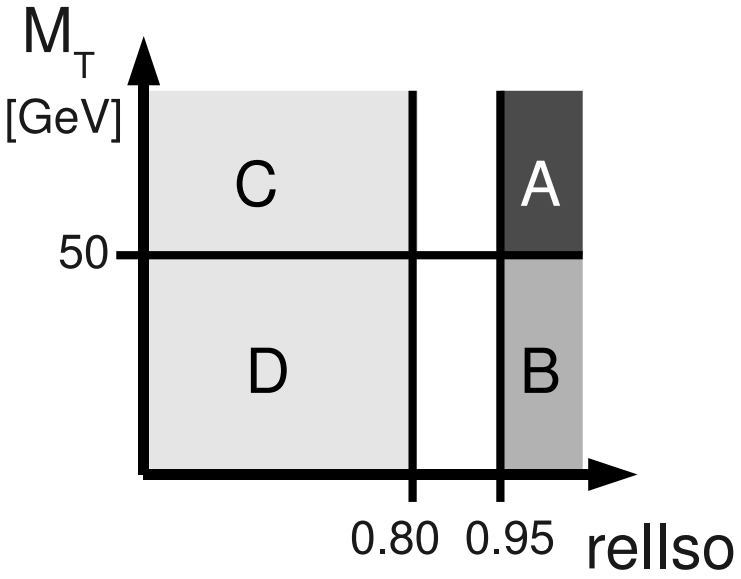
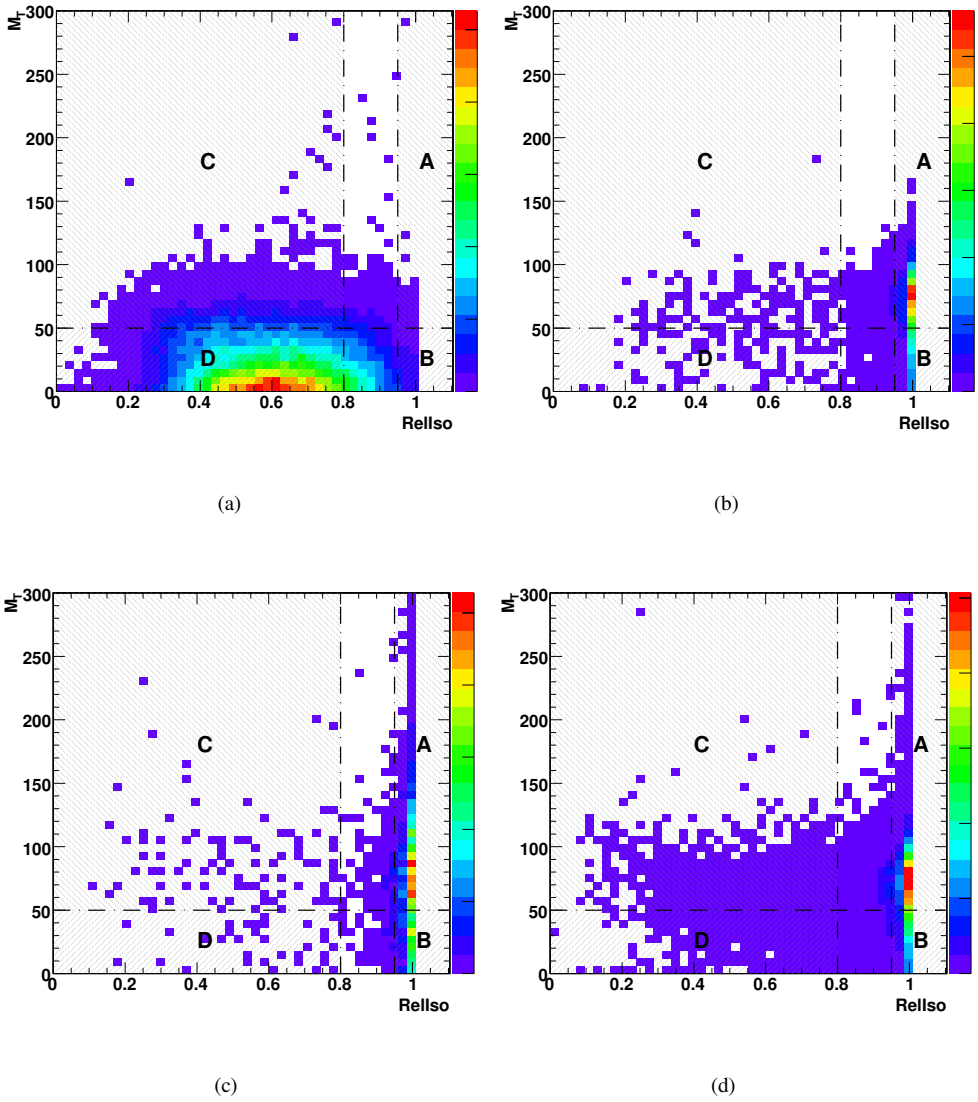


Figure 8: "$ABCD$" method: The four regions in the relative isolation $relIso$ versus M_T plane. Region A is the signal region.

For the QCD background study we apply a selection which differs from the full selection of Sec. 3 only for the absence of the b-tagging requirement, and of course the $relIso$ and M_T requirements themselves. The b-tagging selection is not applied in this section just for pragmatical reasons, dictated by the limited statistics of the available QCD Monte Carlo sample: too few events survive in the signal region, when $N_{btag} = 1$, to make any



(a) (b)

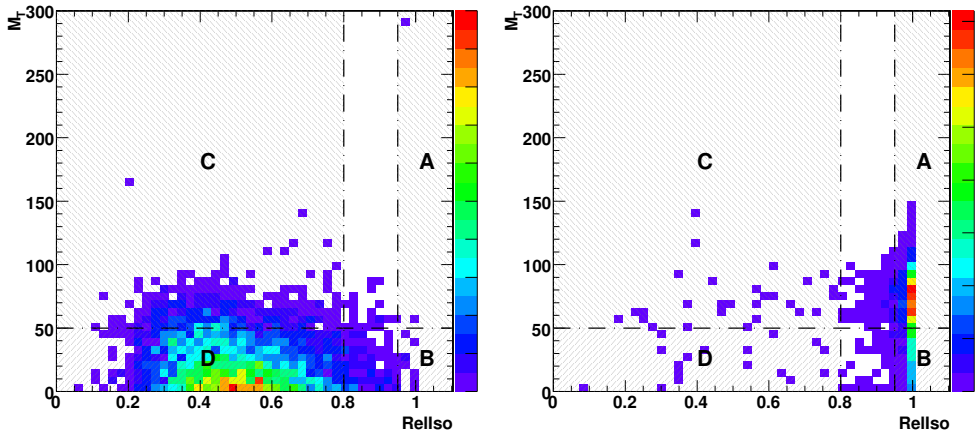
(c) (d)

Figure 9: The M_T versus relative isolation plane for the b-tagless muonic selection of the QCD estimation study: (a) QCD (μ-enriched), (b) signal, (c) $t\bar{t}$, (d) $W + jets$.

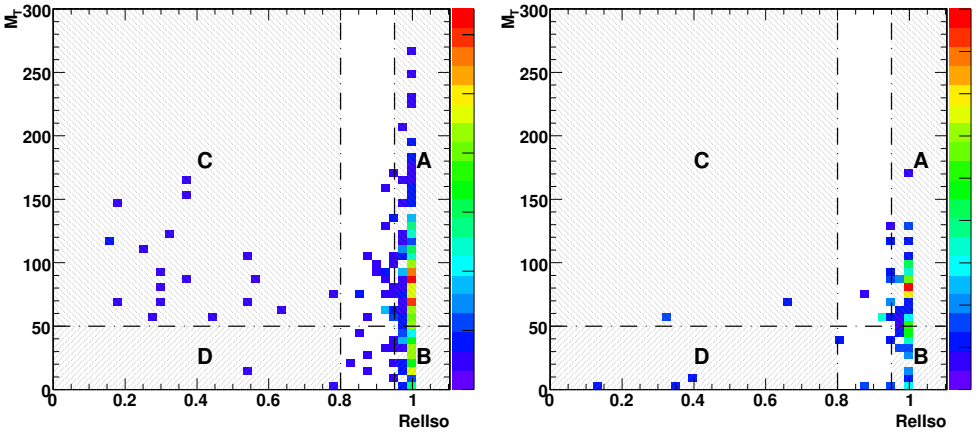
quantitative assessment of the performance of the method. The 2-dimensional $relIso/M_T$ distributions for the muonic selection without b tagging are presented in Figs. 9 for (a) QCD (muon-enriched), (b) signal, (c) $t\bar{t}$ and (d) $W + jets$ events respectively; the same after b tagging is shown in Figs. 10. Figure 11 shows the two variables individually for the four processes considered here.

Table 4 shows the expected numbers of events for the muonic selection in 1 fb^{-1}, without and with the b-tagging requirement. It can be seen that signal-like events (meaning $W + jets$, $t\bar{t}$ and the signal itself) give a negligible contribution to the regions C, D, which are both largely dominated by QCD, while region B is still dominated by QCD, but the signal-like contributions are not negligible.

We use the QCD MC for a self-consistency test by comparing the number of QCD events predicted by the $ABCD$ method (Eq. 2) with the actual number of QCD events in region A. With the numbers given in Table 4(a) an extrapolation of 10.6×10^3 QCD events in the A region is obtained, i.e., correct within $\approx 6\%$. Table 4(b) gives a prediction of 247 QCD events in A, to be compared with 345, with a huge uncertainty from MC statistics (these 345 events correspond to only 3 events in the simulated sample). Within the precision of the available statistics the "$ABCD$ method" seems to be self consistent.



(a) (b)



(c) (d)

Figure 10: The M_T versus relative isolation plane for standard muonic selection with b tagging: (a) QCD (μ-enriched), (b) signal, (c) $t\bar{t}$, (d) $W + jets$.

4.2 ABCD Method in Data

In real data we don't know the contributions of the different processes in our four regions, so it is impossible to just use equation 2. In the "*naïve $ABCD$ method*" the contribution of non-QCD processes in the QCD-enriched regions B, C, D is ignored. By assuming $N_B^{tot} = N_B^{QCD}$, $N_C^{tot} = N_C^{QCD}$, $N_D^{tot} = N_D^{QCD}$ equation 2 modifies to:

$$N_A^{QCD} = \frac{N_B^{tot} \cdot N_C^{tot}}{N_D^{tot}} \,. \tag{3}$$

The result from Table 4(a) is 13.4×10^3 events, i.e., $\approx 20\%$ too many. The systematic shift of about 20% is due to $N_B^{QCD} \approx 0.79 \cdot N_B^{tot}$, while the contamination of region C, D with signal-like processes is tiny and doesn't affect the result.

Modifying the naïve $ABCD$ method (eq. 3) by subtracting the number of signal-like events $N_B^{pred-sig}$ in region B as predicted by MC leads to:

$$N_A^{QCD} = \frac{(N_B^{tot} - N_B^{pred-sig}) \cdot N_C^{tot}}{N_D^{tot}} \,. \tag{4}$$

The relative uncertainty $\Delta N_A^{QCD} / N_A^{QCD}$ of the number of QCD events in region A depends on the uncertainty

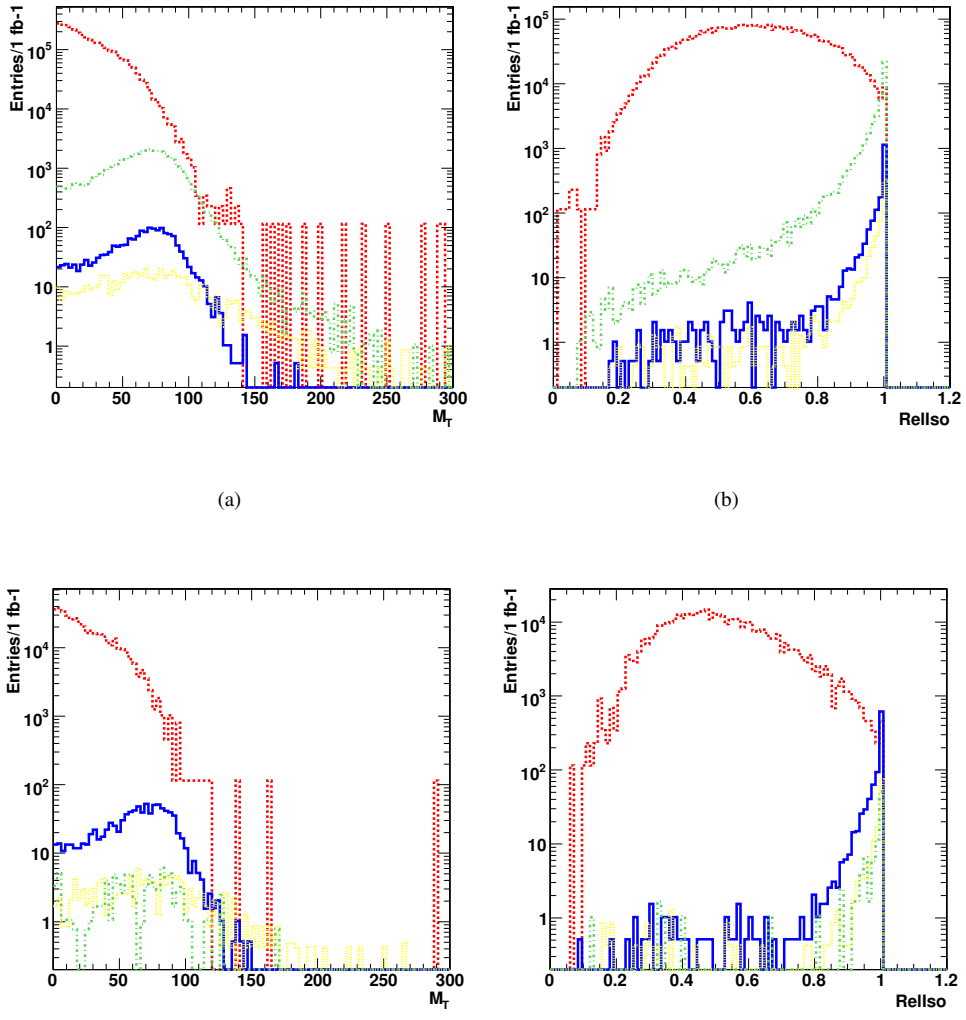


Figure 11: The M_T (left) and relative isolation (right) distributions without (top row) and with (bottom row) b tagging. Red is QCD (μ-enriched), green is $W + jets$, blue is signal, yellow is $t\bar{t}$; all samples are normalized to 1 fb^{-1}.

$\Delta N_B^{pred-sig}$:

$$\frac{\Delta N_A^{QCD}}{N_A^{QCD}} \approx \frac{\Delta N_B^{pred-sig}}{N_B^{QCD}} = \frac{\Delta N_B^{pred-sig}}{N_B^{pred-sig}} \frac{N_B^{pred-sig}}{N_B^{QCD}} \approx f_{sig/QCD} \frac{\Delta N_B^{pred-sig}}{N_B^{pred-sig}} \tag{5}$$

Here, $f_{sig/QCD}$ is the ratio of signal-like events and QCD events in the region B. Because the regions B, C, D are well populated, the relative statistical uncertainties have been neglected in the above consideration. We obtain $f_{sig/QCD} = 0.27$ in the case of the b-tagless analysis and $f_{sig/QCD} = 0.19$ with b tagging. In an early analysis, when the Monte Carlo models are not yet tuned on LHC data, the total number of signal-like events will only be known with a sizeable uncertainty leading to a sizeable uncertainty of N_A^{QCD}, if determined via equation 4.

A better way to improve the naïve $ABCD$ method is to scale the number of predicted signal-like events such that $N_A^{pred-sig} = N_A^{sig}$, where $N_A^{sig} = N_A^{tot} - N_A^{QCD}$ is the unknown number of signal-like events in data. By doing so, we only rely on the relative fractions $f_{X/A} = N_X^{pred-sig}/N_A^{pred-sig}$ ($X : B, C, D$), which are supposed to be much better known than the absolute numbers of predicted events $N_X^{pred-sig}$.

(a) No b tagging				
Process	A	B	C	D
signal	2238	861	78	59
QCD (μ-enriched)	11.2×10^3	74.1×10^3	6.52×10^5	4.57×10^6
$t\bar{t}$	740	306	51	18
$W + jets$	45.0×10^3	18.5×10^3	1630	863
N^{tot}	59.2×10^3	93.8×10^3	6.54×10^5	4.57×10^6
(b) b tagging, $D > 4.8$				
Process	A	B	C	D
signal	617	248	14	10
QCD (μ-enriched)	345	1607	52.8×10^3	344×10^3
$t\bar{t}$	85	38	9	1
$W + jets$	48	26	3	3
N^{tot}	1095	1919	52.8×10^3	344×10^3

Table 4: Number of events expected in the A, B, C, D regions defined in the text, for the muonic selection without (a) and with (b) b tagging in 1 fb^{-1}.

In this approximation we end up with two equations:

$$N_A^{QCD} = \frac{\left(N_B^{tot} - f_{B/A} \cdot N_A^{sig}\right) \cdot \left(N_C^{tot} - f_{C/A} \cdot N_A^{sig}\right)}{\left(N_D^{tot} - f_{D/A} \cdot N_A^{sig}\right)} , \tag{6}$$

$$N_A^{tot} = N_A^{QCD} + N_A^{sig} . \tag{7}$$

Inserting 7 in 6 yields a quadratic equation in N_A^{sig} (see, for example, Sec. 7.2.1 of Ref. [38] and Appendices D,E of Ref. [39]). Instead of solving this quadratic equation we follow another ansatz suggested by the negligible amount of signal-like events in C and D. We set $N_C^{tot} = N_C^{QCD}$, $N_D^{tot} = N_D^{QCD}$ and equation 6 simplifies to:

$$N_A^{QCD} = \frac{\left(N_B^{tot} - f_{B/A} \cdot N_A^{sig}\right) \cdot N_C^{tot}}{N_D^{tot}} \tag{8}$$

Inserting Eq.7 into 8 and solving for N_A^{QCD} yields:

$$N_A^{QCD} = N_C^{tot} \frac{(N_B^{tot} - f_{B/A} N_A^{tot})}{(N_D^{tot} - f_{B/A} N_C^{tot})} . \tag{9}$$

In the following we refer to formula 9 as *improved $ABCD$ method*. The improved $ABCD$ method gives, with the numbers taken from Table 4(a), the prediction 10.6×10^3, different only at per mil level from what had been obtained from the consistency test of Eq. 2. Neglecting the statistical uncertainties again, error propagation leads to a relative uncertainty of

$$\left(\frac{\Delta N_A^{QCD}}{N_A^{QCD}}\right)^2 = \left[\frac{N_B^{tot} N_C^{tot} - N_A^{tot} N_D^{tot}}{(N_D^{tot} - f_{B/A} N_C^{tot})(N_B^{tot} - f_{B/A} N_A^{tot})}\right]^2 \times (\Delta f_{B/A})^2 , \tag{10}$$

leading, with the numbers from Table 4(a), to $\Delta N_A^{QCD}/N_A^{QCD} \approx 0.7 \cdot \Delta f_{B/A}$ for the b-tagless case, and $\Delta N_A^{QCD}/N_A^{QCD} \approx 0.6 \cdot \Delta f_{B/A}$ with b tagging. Assuming a relative uncertainty $\Delta f_{B/A}/f_{B/A}$ of 10% (justified in the next subsection), the relative uncertainty on N_A^{QCD} is 2.9% without b tagging, and 2.2% with b tagging. The actual value of $\Delta f_{B/A}$ will have to be determined from control samples, but this exercise indicates that we can tolerate a relative uncertainty as large as $20-30\%$ on this ratio (since it would give an uncertainty comparable with the 6% discrepancy between the predicted and observed N_A^{QCD}), which for the b-tagless case gives an uncertainty comparable with the deviation between the prediction and the actual count of QCD events in region A.

As soon as larger simulated QCD samples will be available, the effectiveness of the method will also be better tested in the case when b tagging is applied.

4.3 $f_{B/A}$ from data

The value of $f_{B/A}$ can be affected by several detector or physics effects distorting the shape of the M_T distribution. It is then important to have reliable handles to check the MC prediction and eventually tune it to data.

One such handle, for the purposes of the standard selection with b tagging, is the b-tagless selection used as illustration through this section, since it permits a high statistics of $W + jets$ events in both regions A and B, as shown in Table 4(a). With the prediction for N_A^{QCD} coming from the method described, automatically N_B^{QCD} follows, and then we can check whether N_B^{tot}/N_A^{tot} is consistent with the assumed value for $f_{B/A}$. This strategy assumes:

- no correlation between M_T and the b-tagging discriminator;
- same M_T shape for single top and $W + jets$.

Figure 12(a,b) indicates that the first condition is well respected for both signal and $W + jets$. Quantitatively, variating the b-tagging threshold for the single-top sample ($D > -\infty, 2.0, 4.8$) gives the same $f_{B/A}$ ratio within 3%. The same exercise for $W+jets$ ($D > -\infty$ and $D > 2.0$, not enough statistics for a tighter threshold) gives the same result within 10%. Figure 12(c) shows that the M_T shapes are quite similar, although not identical, between signal and $W + jets$. The MC prediction for is $f_{B/A} = 37.9\%$ for signal and $f_{B/A} = 41.4\%$ for $W + jets$, therefore within the 10% uncertainty that we assumed in the previous subsection.

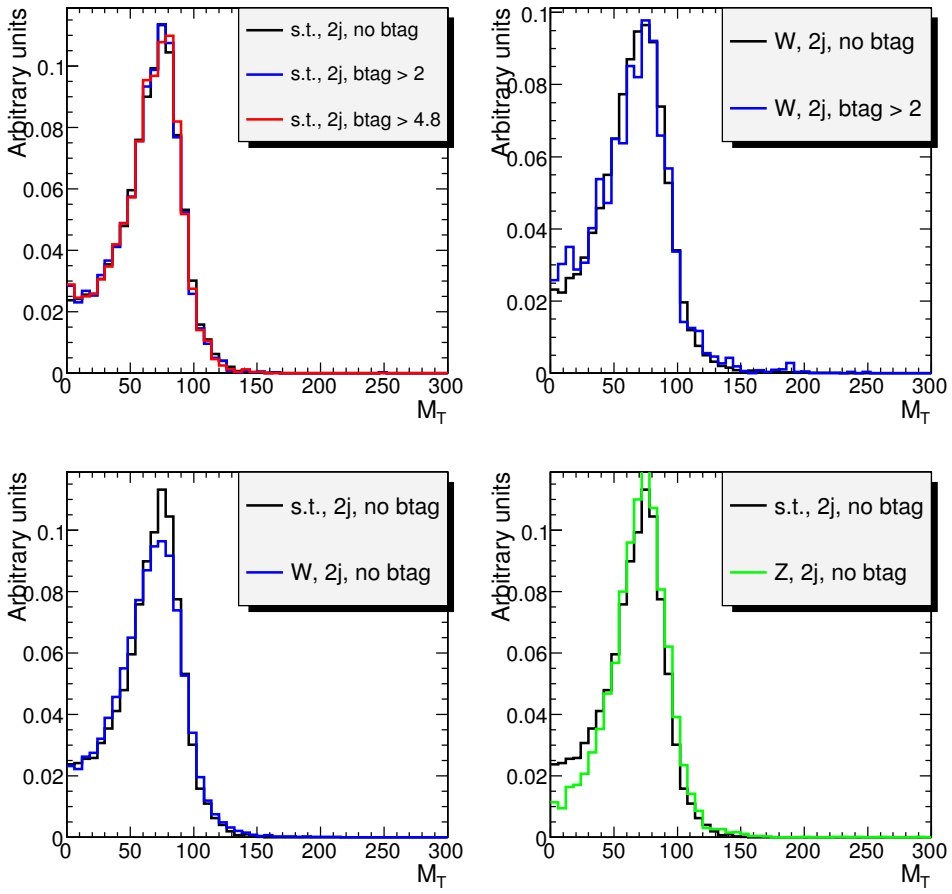


Figure 12: M_T distribution for (a) signal events with standard selection and $D > -\infty, 2.0, 4.8$; (b) $W + jets$ events with standard selection and $D > -\infty, 2.0$; (c) signal and $W + jets$ events with standard selection and $D > -\infty$; (d) signal (standard selection and $D > -\infty$) and $Z + jets$ (standard selection but $N_\mu = 2$ and $D > -\infty$), with the procedure described in the text.

An alternative way to cross-check the M_T distribution coming from MC or to tune it to data is offered by a dedicated $Z \to \mu^+\mu^-$ selection, which would have the benefit of a negligible QCD contamination. The selection

would be the standard one, apart from the request of a second muon with opposite charge and, again, the absence of the b-tagging selection in order to increase the statistics. The muon momenta would be rescaled by M_W/M_Z and one of the two muons, randomly chosen, would be treated as missing in order to mimick a neutrino (its rescaled momentum would be vectorially summed to the missing energy). Figure 12(d) shows that this procedure yields a M_T distribution very similar to the signal one in the high-M_T region, while it severely deviates for low M_T. This is heuristically attributed to the fact that the muon treated as neutrino has been necessarily selected by p_T and isolation requirements, while no such requirement can be imposed on neutrinos in real data; the effect is expected to be more visible in the soft part of the spectrum, as observed.

4.4 Results

In this section we studied the applicability of the "$ABCD$" method for single-top event candidates in the muonic channel. Here, the number of QCD events in the signal region A is estimated from the number of QCD events in QCD-enriched sections in the $relIso/M_T$ plane (regions B, C, D). One fundamental assumption of the $ABCD$ method is the uncorrelation between the variables used for defining the four sections. It has been shown in Ref. [37] that the correlation among the relative isolation and M_T is small and in our QCD study we could confirm this observation for our muonic selection and a slightly different definition of relative isolation.

For the QCD background study we apply a selection which differs from the full selection of Sec. 3 only for the absence of the b-tagging requirement, and of course the $relIso$ and M_T cuts themselves. Applying the $ABCD$ method to pure QCD MC events (Eq. 2) shows that the $ABCD$ method is self consistent for selected single-top event candidates.

We have studied several possibilities to extend Eq. 2 making it suitable for real data. In the naïve $ABCD$ method the signal-like contamination in the QCD-enriched regions is completely neglected, leading to a bias of the QCD prediction in the signal region of the size of the fraction of signal-like events in the region B, predicted to be $\approx 20\%$.
We have seen that subtracting the absolute number of signal-like events in region B as predicted by the MC from the total number of events in region B (Eq. 4) leads to a sizeable uncertainty on the number of QCD events in the signal region, although there is no bias in the method itself.
We found that a better way to improve the naïve $ABCD$ method is to use only the relative fraction $f_{B/A} = N_B^{pred-sig}/N_A^{pred-sig}$ of signal-like events in regions A and B (Eq. 9). This improved $ABCD$ method yields a bias-free QCD background estimation in the signal region and the systematic error due to the uncertainty of the signal-like contamination is small. The possibility to constrain the $f_{B/A}$ value with data has been discussed.

5 Systematic uncertainties

This section considers the sensitivity of the selection of Sec. 3 to systematic uncertainties of instrumental or physics origin.

5.1 Normalization uncertainty for backgrounds

It's particularly hard to assess the expected uncertainties on the background cross sections at the time when we will be ready to publish the first single top results from LHC collisions. The reader must be aware that the following list contains much guesswork:

- $t\bar{t}$: the single-muonic analysis with no b tagging will reach a $\pm 10\%$ uncertainty very soon [28], then the Jet Energy Scale uncertainty is expected to become the dominant error; the method of fitting the reconstructed top mass is expected to yield a cross section uncertainty of $\approx \pm 10\%$ for $\pm 10\%$ error on JES [29]; other systematic uncertainties have not been studied yet. Here we apply a $\pm 20\%$ variation on this process.

- $W/Z + jets$ and $W/Z + Q\bar{Q}$: the most important issues from the point of view of our analysis are the radiation modeling, affecting the number of jets with a chance to pass the p_T threshold, and the heavy flavour fraction; both of them will require much work to be extracted from data. Here we apply a conservative $\pm 50\%$ uncertainty on both.

- tW: for a long time the LHC will be forced to rely on the theory expectations for this process, due to the difficulty to disentangle it from the very similar and more abundant $t\bar{t}$ process [30]. This process is even out

Process	Uncertainty
$t\bar{t}$	$\pm 20\%$
$W/Z + jets$	$\pm 50\%$
$W/Z + QQ + X$	$\pm 50\%$
tW	$\pm 50\%$
$WW/WZ/ZZ$	$\pm 50\%$
QCD	$\pm 30\%$

Table 5: Uncertainties considered for the total cross sections of the background samples.

of the Tevatron's reach, making it impossible to test its modeling with real data before LHC. Here we apply a $\pm 50\%$ uncertainty, motivated by the LO/NLO difference taken from Ref. [7].

- $WW/WZ/ZZ$: the same considerations hold as for tW, and here we quote a $\pm 50\%$ uncertainty on the basis of the LO/NLO difference taken from Ref. [31].
- QCD: its a priori uncertainty is probably the worst, because it only passes the selection through the very extreme tails of the distributions, but it will be very abundantly produced in pp collisions and we demonstrated in Sec. 4 our strategy to constrain it. Our method, applied to Monte Carlo, yields a less than 6% deviation and a small systematic uncertainty coming from the signal-like contamination of the control regions; nevertheless, here we apply a conservative $\pm 30\%$ uncertainty based on the Tevatron experience.

Table 5 summarizes this list.

5.2 Jet Energy Scale (JES) and $\not\!\!E_{\rm T}$ scale uncertainty

In a scenario with $\approx\ 1\ fb^{-1}$ of data on tape, from the intrinsical uncertainty of the data-driven methods to extract jet calibrations we expect a JES uncertainty of $\approx \pm 5\%$, and an intercalibration systematic of $\pm 1\%$ for jets in the barrel ($|\eta| < 1.3$), $\pm 2-3\%$ in the endcaps ($1.3 < |\eta| < 3$) and $\pm 5-10\%$ in the Hadron Forward (HF) calorimeter ($3 < |\eta| < 5$) [32].

In this analysis we apply a simultaneous variation of $1 + \alpha(\eta)$ on all jet 4-momenta, where $\alpha(|\eta| < 3)$ can take the values $+5\%$ and -5% (barrel and endcaps), and $\alpha(|\eta| > 3) = 2 \times \alpha(|\eta| < 3)$ (HF).

Since $\not\!\!E_{\rm T}$ has "Type I" corrections, i.e., the calorimetric clusters associated to jets are given the calibrated jet momenta, its uncertainty is correlated with the JES uncertainty discussed in the previous section.

Here two independent sources of $\not\!\!E_{\rm T}$ systematics are considered:

- correlated with JES: all the jets above 20 GeV are corrected by the same factors discussed before, and $\not\!\!E_{\rm T}$ is recalculated accordingly;
- uncorrelated with JES: after subtracting the jet corrections, $\not\!\!E_{\rm T}$ is variated by 10%.

5.3 b tagging and mistagging uncertainties

For the tight b-tagging working point used in this analysis, estimates of the expected uncertainty for the efficiencies of true and fake b-jets identification can be found, respectively, in Refs. [35] and [36], in different integrated luminosity scenarios. For the $1\ fb^{-1}$ case, the b-tagging efficiency at this working point for the HPTC algorithm is expected to be known within $\pm 5.3\%$, while for mistagging the corresponding uncertainty is expected to be $\pm 15.5\%$.

These uncertainties have been translated into variations of the b-tagging discriminator, selecting from MC truth the taggable jets associated or not to real b quarks in signal events. This procedure gives a ± 0.35 range around the 4.8 threshold, corresponding to $\epsilon_b = \pm 5.3\%$, and ± 0.45 corresponding to $\epsilon_{non\ b} = \pm 15.5\%$.

5.4 Summary of systematic errors of instrumental origin

The effects on selection efficiencies of the systematic uncertainties of instrumental origin (JES, $\not\!\!E_{\rm T}$, b tagging) are summarized in Table 6 for the signal and the main backgrounds. For other samples the limited number of surviving MC events makes it difficult to draw conclusions.

Process	JES up	JES down	MET up	MET down	ϵ_{btag} down	ϵ_{btag} up	ϵ_{mis} down	ϵ_{mis} up
Signal	$+3.7\%$	-6.5%	$+1.1\%$	-2.4%	-4.5%	$+5.5\%$	0.0%	$+0.1\%$
$t\bar{t}$	-12%	$+17\%$	$+1.0\%$	-0.5%	-8.8%	$+16.2\%$	0.0%	$+0.5\%$
$Wjets$	$+22\%$	-6%	0%	$+4\%$	-6%	$+4\%$	-18%	$+33\%$
Wt	0%	0%	$+2\%$	0%	-9%	$+2\%$	-5%	$+2\%$

Table 6: Impact of the considered systematic errors of instrumental origin on the selection efficiencies for the signal and the main backgrounds.

5.5 Charge asymmetry uncertainty

In Sec. 6.2 the asymmetry between t and $\bar{t}$ quarks is exploited as a mean of evidencing the signal, and the $W+jets$ processes become an important background because the same causes yield a corresponding asymmetry between W^+ and W^-.

Here we assume an uncertainty $\Delta A_W/A_W = \Delta A_s/A_s = \pm 5\%$ (where $A_x = (N_x^+ - N_x^-)/(N_x^+ + N_x^-)$), 100% correlated for the two asymmetries, for the sake of simplicity, since they come from the same physical causes. This value is justified by the expectations for the W-asymmetry analysis discussed in Ref. [43], and the fact that once the PDF fits will have taken this information into account, the single-top asymmetry will be possible to be predicted accordingly.

This variation has of course no effect on the number of selected events, under our assumption of equal selection efficiencies for leptons of both signs, but in Sec. 6.2 it will be taken into account through:

$$N_X^+ \to N_X^+ + \Delta A_X \times \frac{N_X^+ + N_X^-}{2}\,, \tag{11}$$

$$N_X^- \to N_X^- - \Delta A_X \times \frac{N_X^+ + N_X^-}{2}\,, \tag{12}$$

for X corresponding to $W + jets$ (including $WQ\bar{Q}$) or signal.

6 Signal extraction and cross section measurement

This section describes two complementary methods to extract an evidence for signal after the complete selection:

- simple event counting (EC);
- exploitation of the charge asymmetry expected for the signal (CR).

In both cases, "pseudoexperiments" are used for the evaluation of the statistical significance, including the systematic uncertainties as estimated in Sec. 5.

6.1 Event counting (EC)

The basic idea is to assess the degree of incompatibility of the total number of selected events with the expectation for the background-only hypothesis.

For each process 30,000 random numbers are drawn from a Gaussian distribution centered on the number of selected events predicted in Table 2. Further Gaussian smearings are applied in order to take into account the systematic errors estimated in Sec. 5.

Calling $G(m,\sigma)$ a random number belonging to a Gaussian distribution with mean m and standard deviation σ, each pseudoexperiment gives, for each process i:

$$N_i = G(N_i^{sel}, \sqrt{N_i^{sel}}) \times G(1, \Delta_i^{bkg}) \times G(1, \Delta_i^{JES}) \times G(1, \Delta_i^{MET}) \times G(1, \Delta_i^{btag}) \times G(1, \Delta_i^{mistag}), \tag{13}$$

where Δ_i^{bkg}, discussed in Sec. 5.1, comes from Table 5, and the other variations from Table 6. Since the variations are in general asymmetric (the upward and the downward variations are of different size), each Δ_i^{syst} is chosen as $|N_i^{sel} - N_i^{sel,upward}|$ or $|N_i^{sel} - N_i^{sel,downward}|$ according to another draw.

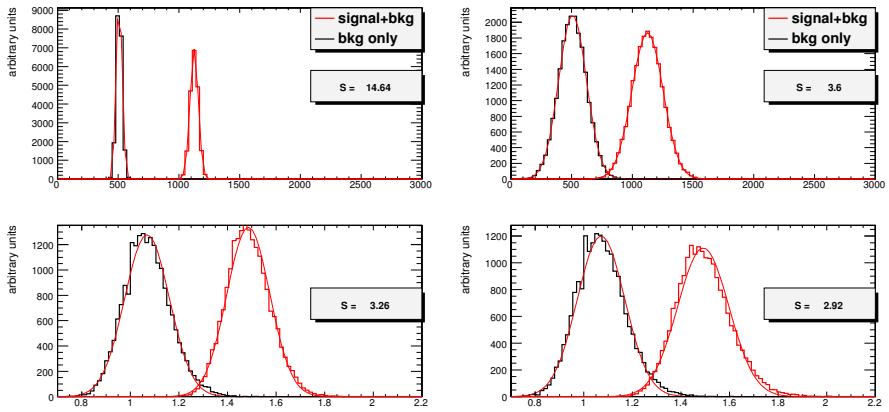


Figure 13: Result of 30,000 pseudoexperiments for EC (top row) and CR (bottom row), for the muon selection, in the background-only and signal+background hypotheses, without (left column) and with (right column) the systematic uncertainties discussed in Sec. 5 apart from asymmetry.

Source	EC sign.	CR sign.
statistical	14.6	3.26
background normalization	3.86	2.94
jet energy scale	11.9	3.23
$\not\!\!E_{\mathrm{T}}$ scale	14.2	3.26
b-tagging efficiency	9.7	3.25
mistagging	12.5	3.24
asymmetry	-	2.67
total (no asymmetry)	3.63	2.92
total (with asymmetry)	3.63	2.49

Table 7: Impact of the considered sources of uncertainty on the significance for the EC and CR methods.

Figure 13 (top row) shows the outcome of the pseudoexperiments for the muon selection (Table 2), with and without considering the systematic errors, for the simple event counting in the background-only and signal plus background hypotheses. The peak positions (M) and standard deviations (σ) of the resulting distributions are extracted by Gaussians fits, and the significance is defined as

$$\frac{M_{sig+bkg} - M_{bkg}}{\sqrt{(\sigma^2_{sig+bkg} + \sigma^2_{bkg})}} \,. \tag{14}$$

A break-down of the individual impact of each considered systematic error on the significance is shown in the first column of Table 7.

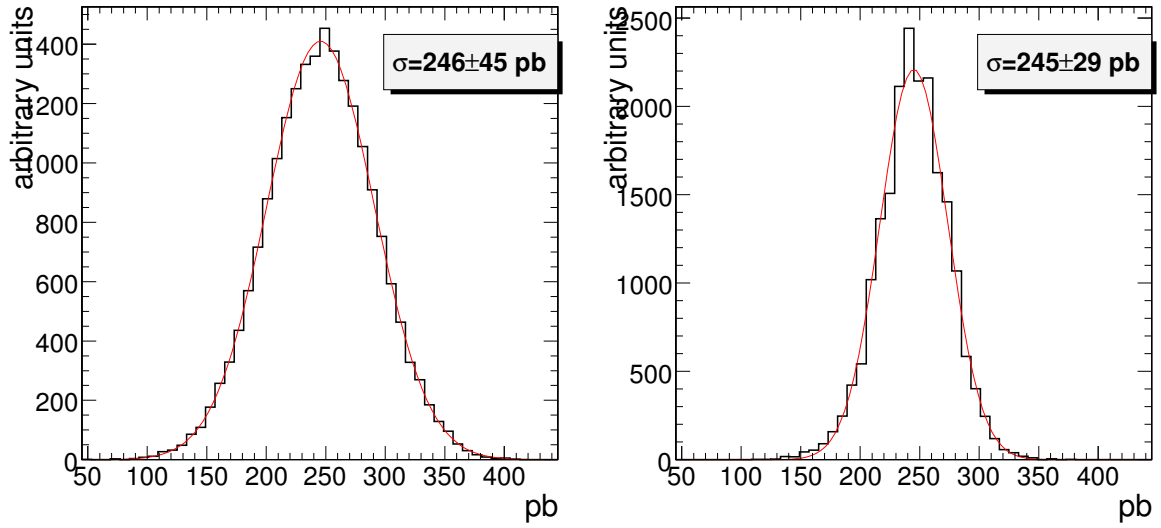
6.1.1 Cross section extraction

The outcome of the same pseudoexperiments can be straightforwardly used for calculating the signal cross section, according to the formula

$$\sigma_s = \frac{N_{sel} - B}{\epsilon_s \cdot BR \cdot L} \,, \tag{15}$$

where N_{sel} is the total number of signal and background events, B is the number of expected background events, ϵ_s is the signal efficiency and L is the integrated luminosity; B and ϵ_s come from the pseudoexperiments.

Figure 14(a) shows the distribution of σ_s obtained by the pseudoexperiments, taking into account the systematics; it can be seen that no significant bias is introduced by the systematic uncertainties. The uncertainty on the cross section is extracted by a Gaussian fit to this distribution, and Table 8 shows the individual impact of the different systematics on the cross section measurement. The background normalization is by far the dominant error for the EC method.



(a) (b)

Figure 14: Cross section extraction for EC (a) and CR (b), with all systematics included apart from asymmetry.

Source	$\Delta\sigma$ [pb], EC	$\Delta\sigma$ [pb], CR
statistical	9.2	27
background normalization	43	9.3
jet energy scale	3.2	4.5
$\not\!\!E_{\mathrm{T}}$ scale	≈ 0	3.5
b-tagging efficiency	5.5	3.1
mistagging	7.1	5.2
asymmetry	-	34
total (no asymmetry)	45	29
total (with asymmetry)	45	47

Table 8: Impact of the considered sources of uncertainty on the EC and CR cross section measurements.

6.2 Charge ratio analysis (CR)

A striking feature of single top production at pp colliders is charge asymmetry. At a centre-of-mass energy of 14 TeV, the cross sections for top and antitop in the t channel are respectively expected to be 150 pb and 92 pb for a top-quark mass of 171 GeV [5]. The top/antitop asymmetry directly translates in a corresponding lepton charge asymmetry, under the reasonable assumption that lepton selection efficiencies and fake lepton contamination do not depend on charge.

The $t\bar{t}$, tW, $\gamma/Z + jets$ and QCD backgrounds are perfectly charge-symmetric, and the only source of charge asymmetry among the major backgrounds is $W + jets$. This suggests the possibility of assessing the presence of signal in the form of an excess of positive over negative leptons in the selected sample, after subtraction of the expected $W + jets$ contribution.

The W asymmetry is expected to be known with a precision comparable to the current PDF uncertainty after $\approx 50\ pb^{-1}$ of CMS data [43]. This asymmetry will be measured from a very different sample, with a jet veto (in order to minimize the contamination from top-quark events, including single top), but Figure 15 shows no significant dependence of the observed asymmetry on jet multiplicity.

The pseudoexperiments in the CR case are performed by fluctuating independently the number of events with positive and negative lepton. The charge ratio is then calculated, for each pseudoexperiment, as

$$R = \frac{N^{+}_{sel}}{N^{-}_{sel}}\ ; \tag{16}$$

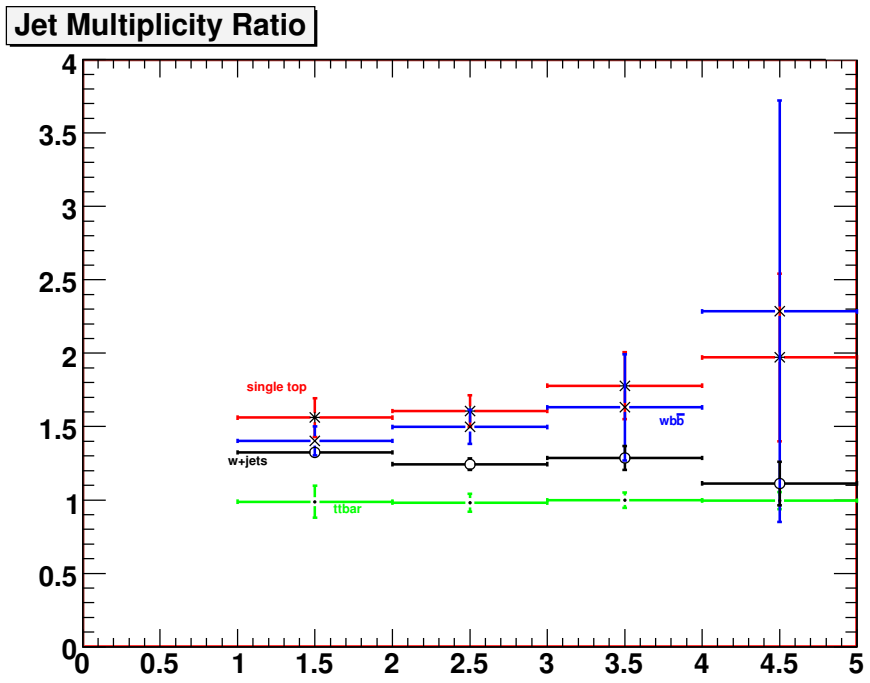


Figure 15: Ratio of positive over negative leptons, after preselection (see Sec. 3.1), as a function of the number of jets (see Sec. 3.2).

in principle it would be necessary to take into account the probability of charge misassignment, and the possible differences in efficiency between positive and negative leptons, but for muons below 100 GeV both effects are estimated to be at less than % level [44].

Figure 13 (bottom row) shows the outcome of the pseudoexperiments for the muon selection, the background-only and signal plus background hypotheses. The significance is defined as in the EC case.

Table 7 breaks down the impact of the individual sources of uncertainty on the significance obtained with the EC and CR methods. One of the greatest impacts comes in both cases from the background normalization uncertainty, where the largest contributor is the QCD uncertainty. For EC, a relatively large impact also comes from JES and b-tagging uncertainties, which in CR mostly cancel out in the ratio, while a very large impact for CR comes from the uncertainty on the asymmetries A_W and A_s.

It is apparent that while EC has a much higher discriminating power between the background-only and the signal+background hypotheses when systematic errors are ignored, as soon as realistic systematics are introduced the EC greatly decreases its significance, still yielding the largest, while the CR shows a great stability against the normalization errors and the systematic uncertainties of instrumental origin, being limited mostly by the uncertainty on the asymmetries A_W and A_s (see Sec. 5.5), here taken 100% correlated. The two methods are therefore complementary.

6.2.1 Cross section extraction

The implementation of the charge ratio method discussed in this section assumes that:

1. the W asymmetry (A_W) is extracted from data in different selections;
2. the signal asymmetry (A_s) is then calculated using as input the new global PDF fit constrained by the W measurements;
3. the W cross section is known.

Given these assumptions, the signal cross section σ_s can be extracted from the following formula:

$$\sigma_s = \frac{(N^+_{sel} - N^-_{sel}) - (B^+ - B^-)}{A_s \cdot \epsilon_s \cdot L} \,, \tag{17}$$

where $B^+ - B^-$ depends on the product $A_W \cdot \epsilon_W \cdot \sigma_W$, but its fluctuations must be calculated by taking into account also the charge-symmetric backgrounds.

The outcome of the pseudoexperiments is shown in Figure 14(b) and a complete break-down of the errors is in Table 8. It can be seen that this method, with an inverse femtobarn, is limited by the knowledge of the asymmetry and by statistics, while detector systematics have a negligible impact.

6.2.2 Outlook

It is apparent from Fig. 7 that the 1-jet bin is dominated by $W + jets$ (although single top itself is an important component of the sample). This makes possible to think about extracting the amount of $W + jets$ contamination in the 2-jets bin by an extrapolation from the 1 jet bin, through a simultaneous extraction of signal and $W + jets$ abundances by solving the following system:

$$A_s \epsilon_1^s \sigma_s + A_W \epsilon_1^W \sigma_W = \frac{N_1^+ - N_1^-}{L} \tag{18}$$

$$A_s \epsilon_2^s \sigma_s + A_W R \epsilon_1^W \sigma_W = \frac{N_2^+ - N_2^-}{L} \, , \tag{19}$$

where the unknowns to be extracted simultaneously are σ_s and $\epsilon_1^W \sigma_W$, and we assumed $\epsilon_2^W = R\epsilon_1^W$, where $R \equiv \sigma_{V(n+1)j}/\sigma_{Vnj}$ ($V = W, Z$), which is expected to be quickly extracted from data with Z selection [45], and is supposed to be also approximately equal to $\sigma_{Wb1j}/\sigma_{Wb0j}$. An end-to-end study of the uncertainties from this method is beyond the scope of this note.

7 Study of single top properties in selected events

Once a supposedly high-purity sample (as we expect to be able to achieve with our selection) has been selected from real collision data, it is important to check that its properties correspond to the features expected for single top. The most obvious is the presence of a mass peak when properly combining the final state objects, interpreted as top-quark decay products. Another useful feature, exploitable after the reconstruction of the top quark, is the angular distribution of the lepton, related with the top-quark polarization.

7.1 W mass constraint

The first step in the reconstruction of the top quark from its decay products is the reconstruction of the W boson. Since this analysis considers only muonic decays of this boson, we must assume that the x and y components of the missing energy (properly corrected) are entirely due to the escaping neutrino, and apply the W-mass constraint in order to extract the z component ($P_{z,\nu}$):

$$M_W^2 = (E_l + \sqrt{\not{E}_{\rm T}^2 + P_{z,\nu}^2})^2 - (\vec{P}_{T,l} + \vec{\not{E}}_{\rm T})^2 - (P_{z,l} + P_{z,\nu})^2 \, . \tag{20}$$

This equation has in general two solutions:

$$P_{z,\nu}^{A,B} = \frac{\mu \cdot P_{z,l}}{P_{T,l}^2} \pm \sqrt{\frac{\mu^2 \cdot P_{z,l}}{P_{T,l}^4} - \frac{E_l^2 \cdot \not{E}_{\rm T}^2 - \mu^2}{P_{T,l}^2}} \, , \tag{21}$$

with

$$\mu = \frac{M_W^2}{2} + \vec{P}_{T,l} \cdot \vec{\not{E}}_{\rm T} \, . \tag{22}$$

If the discriminant in equation 21 becomes negative, or equivalently M_T is larger than the W pole mass used in the constraint, the solutions have an imaginary part. This happens in 37.2% of the cases, mostly due the finite resolution of $\not{E}_{\rm T}$ (lepton momentum resolution and the finite W width give negligible contributions; see the s-channel single-top analysis in Ref. [8] for a more detailed study).

Several schemes exist to deal with this situation; here the imaginary component is eliminated by modifying $\not{E}_{\rm T}$ such to give $M_T = M_W$, still respecting Eq. 20. This is obtained by imposing that the discriminator, and thus the square root in Eq. 21, are null; this gives a quadratic relation between $P_{x,\nu}$ and $P_{y,\nu}$, with two solutions, among which the one with minimal distance between $P_{T,\nu}$ and $\not{E}_{\rm T}$ is chosen.

Assignment	$\Delta R < 0.3$	$\Delta R < 0.5$
b-tagged jet is b from top quark:	93.0%	95.3%
b-tagged jet is second b:	4.3%	4.0%
b-tagged jet is recoil light quark:	0%	0%
b-tagged jet is not matched:	2.7%	0.7%
light jet is b from top quark:	0.5%	0.7%
light jet is second b:	1.2%	1.3%
light jet is recoil light quark:	93.5%	96.0%
light jet is not matched:	4.8%	2.0%

Table 9: Matching of the b-tagged and light jets to MC truth, in selected signal events, for two different values of ΔR in the calculation of the matching fractions.

7.2 Ambiguity resolution and event interpretation

In the "normal" case of two real solutions for $P_{z,\nu}$, different choice criteria have been tried in the literature. Here we choose the solution with the smallest absolute value. The W boson is thus reconstructed by this procedure when the discriminant of Eq. 21 is positive, and by the procedure of the preceding section when it is negative. A similar two-fold ambiguity presents when reconstructing a top-quark hypothesis, since two jets are selected. The ambiguity is resolved by assigning the b-tagged jet to the top-quark decay.

By looking at Monte-Carlo truth it is found that, in 97.4% of the selected events with real solutions, the smallest $|P_{z,\nu}|$ solution is closer in η, ϕ to the true neutrino direction than the other solution. The b-tagged jet matches the true b quark from top-quark decay in 93% of the selected signal events, using as matching criterion a distance of $\Delta R < 0.3$ from the parton; more details are given in Table 9.

7.3 Reconstructed top-quark mass

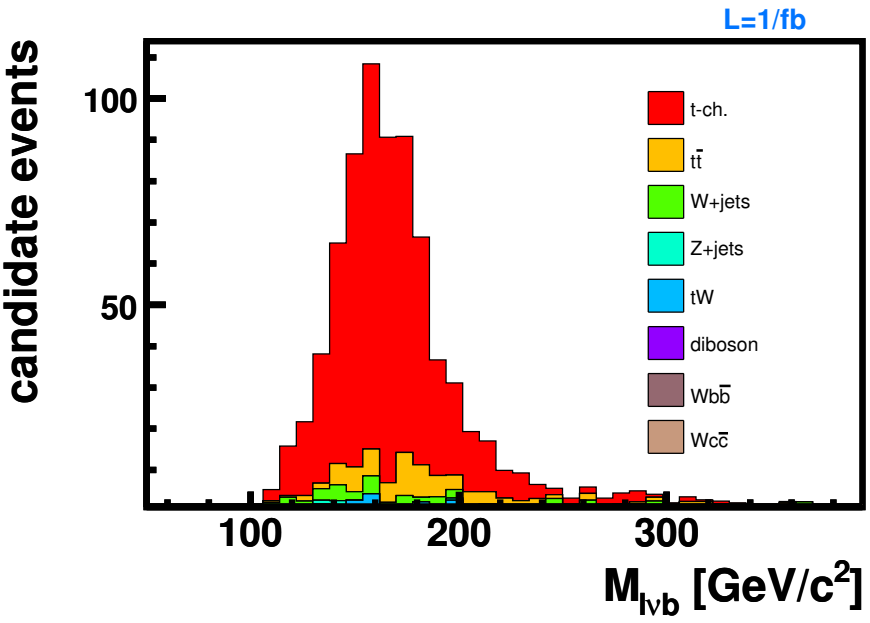


Figure 16: Reconstructed top-quark mass after full selection.

Figure 16 shows the mass of the reconstructed top quark for events passing the full selection. The observation of a peak at $\approx M_{top}^{Tevatron}$ in the selected sample from real collision data will be a smoking gun of the presence of top quarks. On the other hand, this variable is relatively sensitive to detector systematics, as shown in Fig. 17.

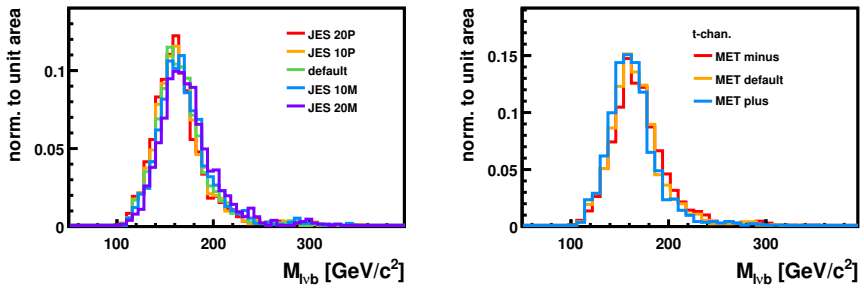


Figure 17: Reconstructed top quark mass, for signal events, with variations on jet energy scale by $\pm 20\%$ and $\pm 10\%$ (left) and $\not\!\!E_{\mathrm{T}}$ scale by $\pm 10\%$ (right).

7.4 Top quark polarization

A specific feature of signal, stemming from the $V - A$ structure of the weak interaction, is the almost 100% left-handed polarization of the top quark with respect to the spin axis [40]. The direction of the top quark spin is reflected in angular correlations in its decay products, which are distributed according to the formula

$$\frac{1}{\Gamma}\frac{d\Gamma}{d\cos\theta^*} = \frac{1}{2}(1 + A\cos\theta^*)\,, \tag{23}$$

where θ^* is the angle between the direction of the outgoing particle and the spin axis, in the top rest frame, and A is a coefficient of spin asymmetry, which depends on the identity of the particle: $+1$ for charged leptons, -0.40 for b quarks, -0.33 for neutrinos [40]. Being the particle with the best possible angular resolution in the detector, the charged muon is chosen for this analysis.

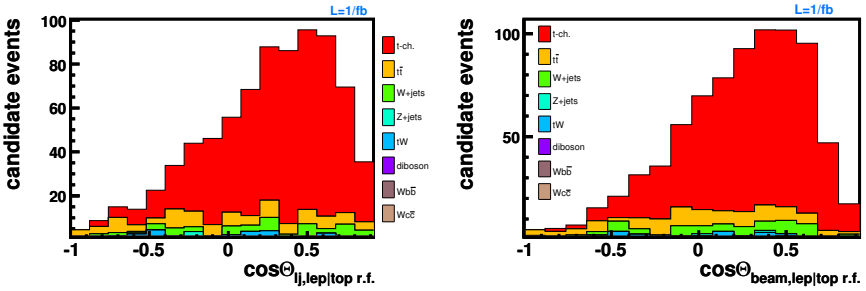


Figure 18: Angle between charged lepton and untagged forward jet (a) and between charged lepton and beam axis (b), with the sign convention explained in the text, in the reconstructed top rest frame after full selection.

Two definitions of θ^* are proposed in the literature (see, e.g., Ref. [41]):

- θ_{lj} (or "spectator basis"), see Fig. 18(a): angle between the muon direction and the light quark recoiling against the virtual W boson; in practice, the direction of the untagged forward jet is chosen;
- θ_{lb} (or "η-beamline basis"), see Fig. 18(b): angle between the muon direction and one of the two LHC beam axes, chosen as the one with the smallest angle with respect to the untagged forward jet direction;

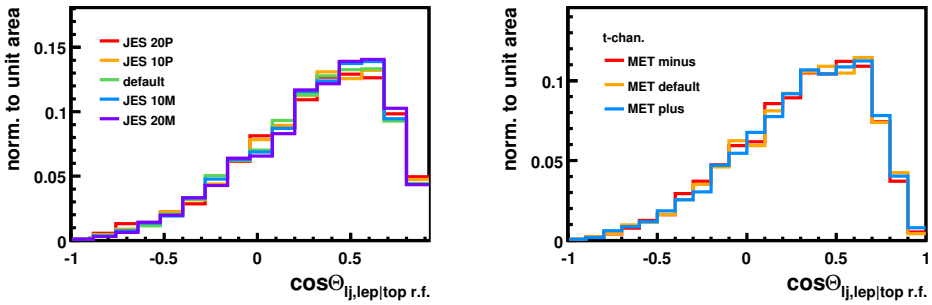
in both cases a boost of all the 4-vectors is performed in the rest frame of the reconstructed top quark.

Both variables are quite stable against JES and $\not\!\!E_{\mathrm{T}}$ systematics, as shown in Fig. 19, differently from the top-quark mass (compare with Fig. 17).

Given the relative flatness of the backgrounds, these distributions can be used for a data-driven estimation of the signal cross section, with no a priori assumption on the size of the backgrounds.

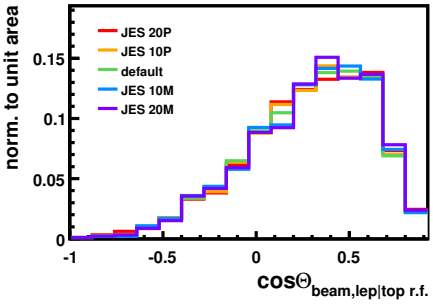
8 Neural Network analysis

In this section we discuss an implementation of the neural network technique for the t-channel event selection. This method offers several advantages, e.g., an easy combination of a very large number of sensitive variables

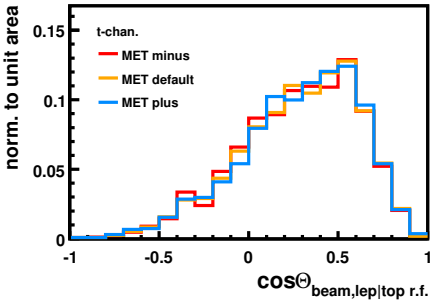


(a)

(b)



(c)



(d)

Figure 19: Angle between charged lepton and untagged forward jet direction (a,b) and between charged lepton and beam axis (c,d), for signal events, with variations on jet energy scale by $\pm 20\%$ and $\pm 10\%$ (a,c) and of $\not\!\!E_{\rm T}$ scale by $\pm 10\%$ (b,d).

in the analysis, at the same time taking into account any non-linear correlation of these variables; this universal method can be applied in different tasks in the same manner.

This analysis repeats and extends a previous experience described in [46] and references therein. For the signal Monte-Carlo model we use the SingleTop generator [47, 10], based on CompHEP [11]. For Neural Networks training we use the package MLPfit [48].

8.1 Preselection

As first analysis step we apply preselection criteria to reject kinematic regions where the experimental reconstruction of physics objects is not stable or where the Monte-Carlo model is known to be less reliable. These criteria are looser than the selection described in section 3 since we want to recover statistics by exploiting additional kinematic regions, for example the $N_j = 3$ subsample where the S/B ratio is a priori less favourable. The preselection criteria for the NN analysis are the following:

- $\not\!\!E_{\rm T} > 30$ GeV;
- only one μ with $p_T > 20$ GeV, $|\eta_\mu| < 2.1$;
- at least one b-tagged jet with $p_T > 30$ GeV$/c$, $|\eta| < 2.5$, $D > 4.8$;
- at least one untagged jet with $p_T > 30$ GeV$/c$;
- $N_j = 2 \div 3$;
- $relIso > 0.95$.

We don't apply any M_T requirement at this stage since we prefer to exploit the discriminating power of this observable in the Neural Networks; the $\not{E}_{\rm T}$ threshold has thus just the purpose of eliminating events that would have no chance to pass the further selection anyway, in a range where the resolution of this variable is very poor.

The numbers of expected events at 1 fb^{-1} that survive our preselection criteria are shown in Table 10; the sensitivity of these numbers to systematic uncertainties is shown in Table 15.

Table 10: Number of expected events that survive our preselection criteria for 1 fb^{-1}. We present the sum of the yields and separate numbers for each Jet bin ($N_{jet} = 2,\ 3$). For comparison purposes we include statistical significance as well, defined as $S/\sqrt{S+B}$.

process	$N_{Nj=2}^{expected}$	$N_{Nj=3}^{expected}$	$N_{tot}^{expected}$ (N_{tot}^{MC})
$t\bar{t}$	2339	7029	9368 (21848)
$Wb\bar{b}j$	212	196	407 (1118)
W+jets	3152	1424	4576 (4353)
tW	477	727	1205 (2546)
QCD	4653	5221	9874 (87)
Bkgd. total	10833	14597	25430
t-channel Signal	1706	1187	4576 (6839)
$S/\sqrt{S+B}$	15	9	17

In this section we consider only one signal process (t-channel single top) and the five main background processes: $t\bar{t}$, $Wb\bar{b}+jets$, $W+jets$, tW and QCD (muon-enriched).

8.2 QCD background rejection with a dedicated Neural Network

Due to the special importance of QCD background we pay special attention to its rejection in addition to the requirement on $relIso$ already included at the preselection level. Due to the difficulty to properly model this background we choose a small set of training variables, which is expected to be robust and reflects the main properties that we can expect for this background. These variables are listed in Table 11. The corresponding distributions are shown in Fig. 20. The output of the trained QCD network is shown in the Fig. 21 for the signal and QCD background, normalized to unit. We apply a threshold at 0.8 on the output of the QCD network and consider for the following analysis only the events passing this requirement. The estimated rejection of QCD background is about 86%, i.e., only 1362 QCD events survive this selection (to be compared with 9874 events from Table 10). The other backgrounds and the signal are not significantly affected by this selection. The numbers of surviving events after this step are shown in Table 12. Further rejection of the QCD background will be provided by Neural Networks trained for other backgrounds, in particular by a NN trained to reject $W+jets$ events. These NNs and the combined Super Neural Network will be described in the following subsection; here we anticipate, in Fig. 22, that the QCD events passing the QCD Neural Network threshold tend to have a flat distribution in the Super Neural Network output.

8.3 Dedicated Neural Networks for the rejection of $t\bar{t}$, $W+jets$ and $Wb\bar{b}+jets$ backgrounds.

The considered backgrounds have significantly different kinematic properties among them. In this case it is more efficient to train different networks with different sets of input variables for each background process. Therefore, we can achieve a more effective separation of the signal and backgrounds by combining a set of NNs, where each network is trained to separate the signal from only one of the background processes.

QCD sensitive variables
$\log \not{E}_{\rm T}$
$\log M_T^W$
$\log p_T^\mu$
η_μ

Table 11: Set of variables sensitive to QCD background.

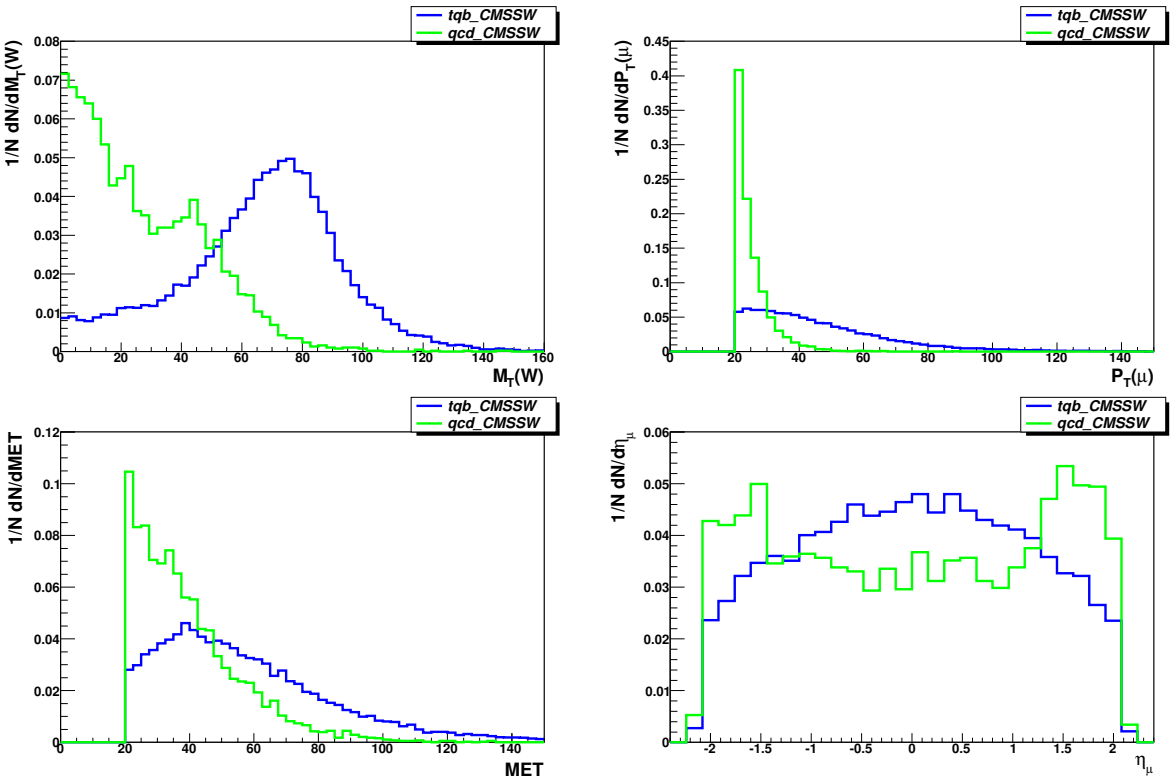


Figure 20: Distributions of the variables used for the training of the QCD network.

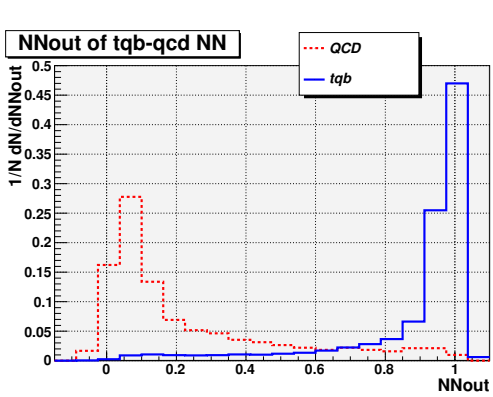


Figure 21: Distribution of QCD neural network output for signal and QCD background.

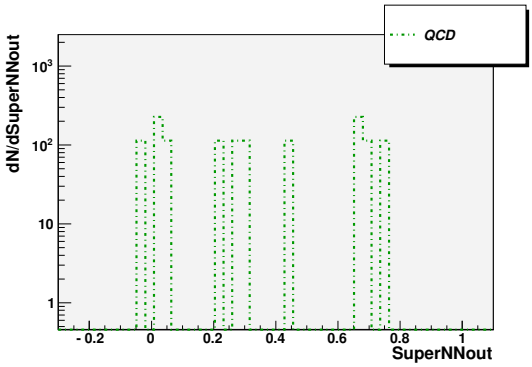


Figure 22: Distribution of Super Neural Network output of QCD events passing the QCD Neural Network threshold. The SuperNN is trained to reject other backgrounds and is described in Sec. 8.3.

One of the main questions in an experimental analysis is how to choose the most appropriate set of variables in order to extract the signal in the most optimal way. In this analysis we use the so called method of "Optimal Observables" [49], based on the analysis of Feynman diagrams contributing to signal and background processes. On the basis of this method we chose the variables listed in Table 13.

In order to construct variables marked with "not best j" we exclude the jet that gives the closest invariant mass $M_{top}(Wj)$ to the known top quark mass, 175 GeV; in the following it will be called "Best" jet. This procedure is similar in spirit to the "top window cut", but in these sets of variables we exclude the jet (supposed to come from the top quark decay) and check the properties of other jets. For example, $H_T(Jets\ without\ best\ j)$ means a scalar sum over jet p_T except "Best" jet. We believe the variables can be important for some backgrounds. For example, the other jets in the $t\bar{t}$ events are produced from the decay of the second top. So, they are more energetic than in the t-channel signal. From the other side, we know that we have got an additional energetic jet in t-channel signal, which is produced not from top quark decay and this jet is distributed in the forward regions of η. So, we want to

Table 12: Number of expected events that survive our preselection criteria and have QCD NN output greater than 0.8 (1 fb^{-1}). We present the sum of the yields and separate numbers for each jet multiplicity ($N_{jet} = 2,\ 3$). For comparison purposes we include statistical significance as well, defined as $S/\sqrt{S+B}$.

process	$N_{Nj=2}^{expected}$	$N_{Nj=3}^{expected}$	$N_{tot}^{expected}$ (N_{tot}^{MC})
$t\bar{t}$	1998	6063	8061 (18798)
$Wb\bar{b}j$	179	158	337 (924)
W+jets	2659	1227	3886 (3732)
tW	407	627	1034 (2186)
QCD	567	794	1362 (12)
Bkgd. total	5811	8869	14680
t-channel Signal	1444	1012	2456 (5806)
$S/\sqrt{S+B}$	17	10	19

Table 13: Sensitive variables for the dedicated Neural Networks.

$t\bar{t}$ sensitive	$Wbbj$ sensitive	$W + jets$ sensitive
$\Delta R(j1, j2)$	$\Delta R(j1, j2)$	$\Delta R(j1, j2)$
$\log H_T(Jets\ without\ best\ j)$	$\log p_T(\overline{j1 + j2})$	$\log p_T(\overline{j1 + j2})$
$\log p_T(j1)(not\ best\ j)$	$\log H_T(all\ j)$	$\log H_T(all\ j)$
$\log p_T(j2)(not\ best\ j)$	$\log M_{all\ j}$	$\log M_{all\ j}$
$\log p_T(\overline{Jets\ without\ best\ j})$	$\log M_{j1,j2}$	$Disc_b^{id}(b1) + Disc_b^{id}(b2)$
$\eta_j(Light\ jet)$	$\log p_T(j1)$	$\log p_T(j2)$
$\eta_j(\eta_{max})$	$\log p_T(b)$	$\log p_T(b)$
$\eta_{j1}(not\ best\ j)$	$\eta_j(Light\ jet)$	$\eta_j(Light\ jet1)$
$\cos(LightJet, \mu)\|_{top}$	$\cos(LightJet, \mu)\|_{top\ j\ best}$	$\cos(LightJet, \mu)\|_{top\ j\ best}$
$\log M_{j1,j2}$	$\cos(LightJet, \mu)\|_{top\ b}$	$\cos(LightJet, \mu)\|_{top\ b}$
$\log \sqrt{\hat{s}}$	$\log M_{top}(Wb)$	$\log M_{top}(Wb_{best})$
$\log M_{top}(Wb_{best})$	$\log \sqrt{\hat{s}}$	$Disc_b^{id}(b)$
		$\eta_b(b_{best})$

take into account this property and introduce two variables $\eta_{j1}(not\ best\ j)$ and $\eta_j(\eta_{max})$, which are the same for some part of the events, but uncorrelated for another significant part ($\eta_j(\eta_{max})$ means η of jet with highest η and $p_T > 30$ GeV).

In our notation we sort jets according to p_T and $j1$ means the highest p_T jet (the same for $b1$). M means an invariant mass of reconstructed objects (e.g. $M_{top}(Wb)$ is the invariant mass of W and the highest p_T b-tagged jet). $cos(LightJet, \mu)|_{top\ j\ best}$ is the cosine of the angle between the muon and untagged jet in the rest frame of the reconstructed top quark with "best" jet. $p_T(\overline{Jets\ without\ best\ j})$ or $p_T(\overline{j1 + j2})$ are a vector sum of p_T of jets. $Disc_b^{id}(b1)$ is a b-tagging probability for the corresponding jet. To be consistent with the recommendations on jet reconstruction we use jets with $p_T > 30$ GeV only, for the variables where we sum over the jets ($M_{all\ j}, H_T(all\ j), \sqrt{\hat{s}}$). The corresponding distributions of the variables above are shown in the plots 23, 24, and 25.

We use a feed-forward NN with supervised training in the analysis. For our task this is the simplest and the most effective type of NN. The next step in the NN analysis is to find the most effective architecture of the NN and a set of training parameters. The optimal number of hidden nodes is usually chosen within the range $[n, 2n + 1]$, where n is the number of input variables. One hidden layer is usually sufficient for most of the tasks in HEP.

To avoid the NN overfitting problem we use the standard prescription of splitting all samples into two parts, the training and testing samples, using the testing samples to check whether the ability of the network to recognize the signal and the background is not improving any more from a cycle to the next. With real pp collision data the main test will be the comparison of NN output distributions between data and simulation. After the above checks the NN is ready for the analysis.

A scheme of NNs for this analysis is shown in Fig. 26. We use three initial networks to separate t-channel signal from each of the backgrounds and then combine three outputs of three initial NNs by the Super NN and get 1D discriminator output for each event. We expect, that signal-like events come to a region close to unit in Super NN output distributions, but background-like events to a group close to zero. The outputs of trained NNs are shown

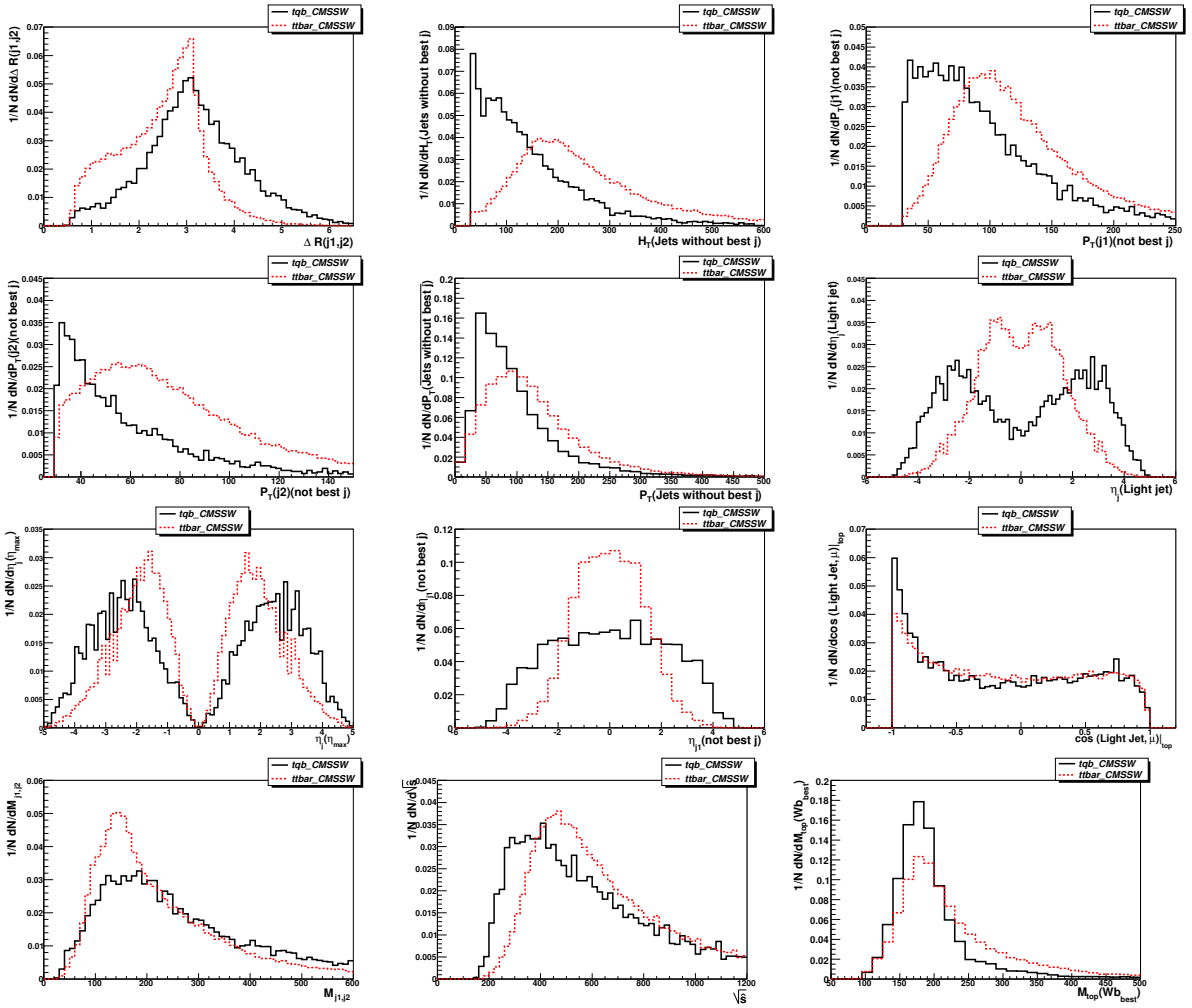


Figure 23: Distributions of variables selected for training of NN to separate t-channel signal from $t\bar{t}$ background.

in the Fig. 27. On these plots, each of $NN_{ttbar}, NN_{Wbbj}, NN_{Wjj}$ distributions are normalized to the same area. The Super NN outputs for all of the signal and background processes are shown on the same plot, all of the curves are normalized to the histogram integral. Figure 28(left) shows the $dN/dSNNout$ distribution with proper weight and contribution of the processes for 1 fb^{-1} luminosity. Figure 28(right) shows the efficiency of different cuts on Super NN output (each point on the curves corresponds to one possible cut on the NN output). The efficiency is defined as $N_{passed}/N_{initial}$ and shown as the function Signal eff. versus Background eff. for the three initial NNs and corresponding background and Super NN with total background (four left upper curves). The single right down curve shows the Signal efficiency versus Super NN cut value (x-axis).

There are different ways to apply NNs and get final results. In this note we simply apply a threshold on the NN output and count the events passing this selection. Super NN output is the 1 dimensional discriminator of signal events. Therefore we can find the most optimal threshold on this discriminator based on the necessary criteria. There are several possible criteria for the optimization. The first possible criteria is the ratio of Signal (S) to Background (B) events, the second is the statistical significance of the selection $\mathrm{Significance}(stat) = S/\sqrt{S+B}$. Our analysis is not statistics-limited but systematics-limited, therefore a sensible optimization criterion should take into account systematic uncertainties, but this is computationally complex since a proper estimation of systematic uncertainties should be performed, in principle, for each possible selection threshold. In this analysis we skip the step of optimizing the final SuperNN threshold and present only the results which can be compared directly with the standard analysis described in Sec. 3.6. We find the threshold on the Super NN output, 0.92, which yields the same number of signal events (617 events, Table 2) and calculate the event yields for the considered processes. Table 14 shows the expected number of events for signal and backgrounds. For comparison purposes we include the Statistical Significance as well, defined as $S/\sqrt{S+B}$. Because of the limited MC statistics for QCD background there are no events survived the cuts, we have mentioned the statistical uncertainty corresponding to one event in the exist MC model to pass the cut.

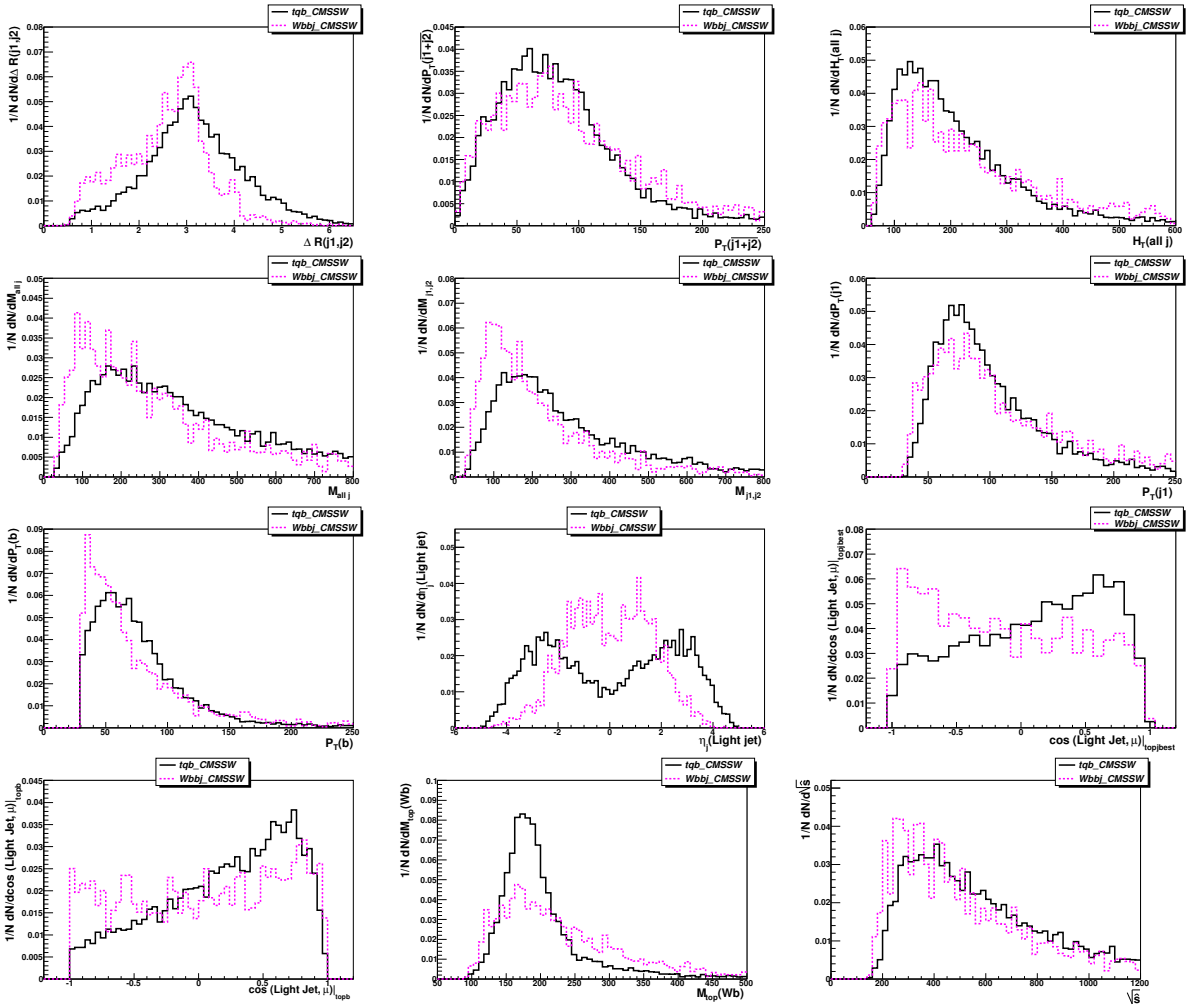


Figure 24: Distributions of variables selected for training of NN to separate t-channel signal from $Wbbj$ background.

8.4 Systematic uncertainties in the Neural Network analysis

In order to investigate the stability of our NNs with systematic uncertainties we apply the systematic shifts for the different types of sources of uncertainty. In the table 16 we show the applied shifts and the difference in the number of events with the corresponding systematic shifts described in the section 5.

In the Fig. 29 we present the changes in SuperNN output distributions after the $\pm\sigma$ shifts with possible sources of systematic uncertainties marked in the table 16.QCD is not included in this table because no events survive the selection; nevertheless, with the upward variation of $\not\!\!E_{\mathrm{T}}$ two QCD events survive in Monte Carlo, leading to a large uncertainty when rescaled to 1 fb^{-1} (2 MC events correspond to roughly 230 expected events). This is not taken into account in the following.

The statistical significance and cross section are calculated as in Sec. 6.1 and Sec. 6.1.1, leading to a significance of 6.8 and a cross section uncertainty of around 15 pb. This has to be compared with the results from Sec. 6.1; the improvement is mostly due to the much lower background contamination, and in particular the reduction in QCD background. The complete breakdown of the impact of the individual systematics is shown in Table 17 (to be compared with Tables 7, 8 for the cut-based analysis).

8.5 Conclusions

We have demonstrated how to combine a simple set of neural networks for the selection of single top t-channel events, and how this can very effectively improve the performance of the analysis. In this note we did not consider the optimization of the Super NN output threshold and only compare the efficiency with the same amount of survived signal events as for the cut-based analysis, postponing to the first real collision data the task of finding the optimal selection. This section also demonstrates the stability of these NNs with respect to the systematic

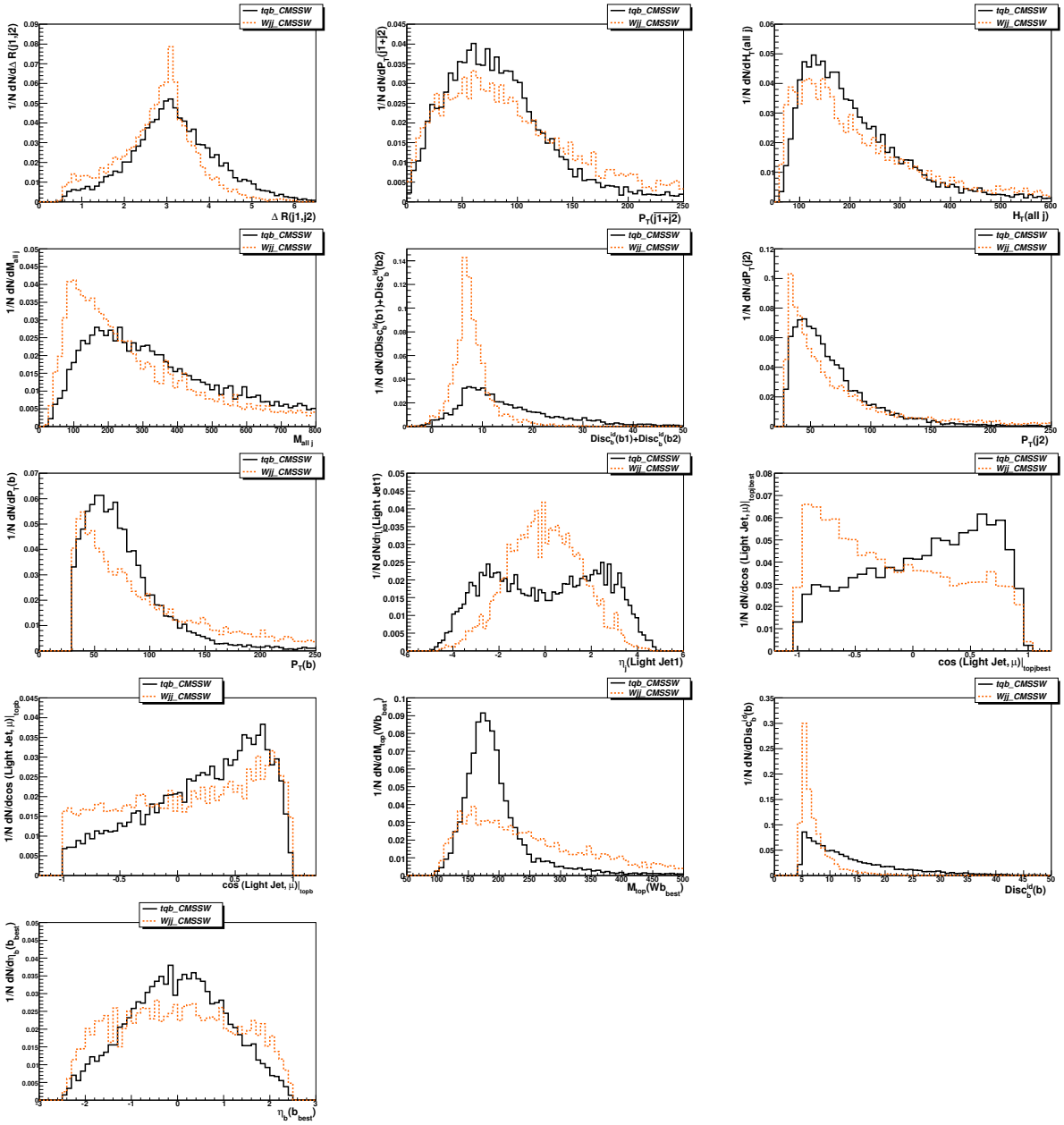


Figure 25: Distributions of variables selected for training of NN to separate t-channel signal from $W + jets$ background.

uncertainties. This method will be easily adapted to real collision data by training the NNs on background templates tuned on control samples, along the same lines as the background control techniques needed for the standard cut-based analysis. The concrete implementation of the background template extraction will be investigated in detail in a forthcoming note.

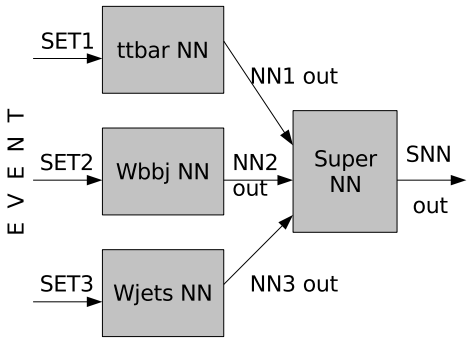


Figure 26: Scheme of NNs in the analysis. SET1 etc. are described in the Table 13.

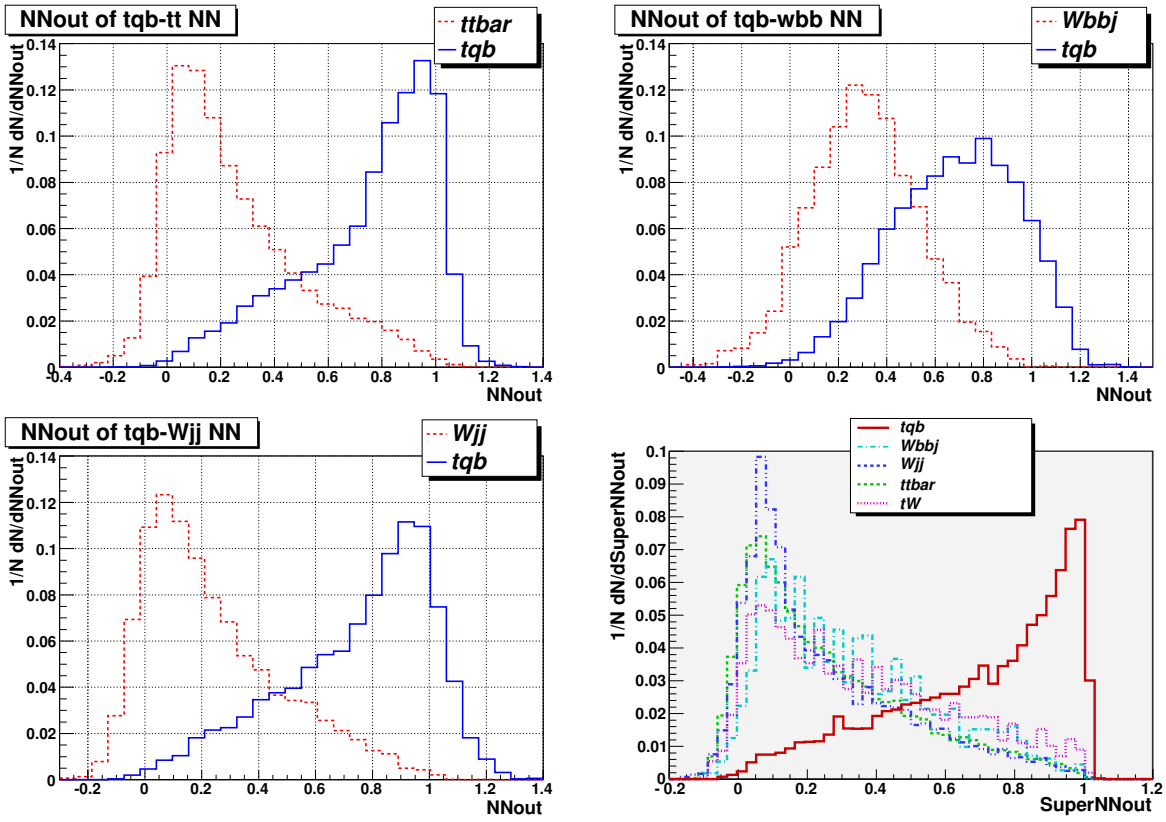


Figure 27: NN outputs of ttbar, Wbbj, Wjj and Super NN (SNN) networks. For the Super NN plot, outputs of all processes are shown. Distributions are normalized to the integral of the histogram.

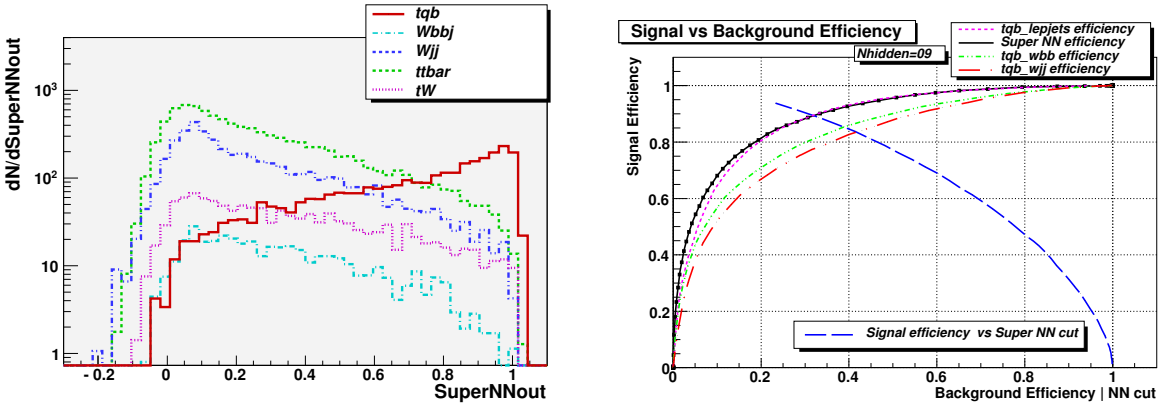


Figure 28: On the left plot we demonstrate the $dN/dSNNout$ distribution with proper weight and contribution of the processes for 1 fb^{-1}. The right plot has shown the efficiency of cuts on Super NN output (each point on the curves corresponds to one possible cut on a NN output). The efficiency is defined as $N_{passed}/N_{initial}$ and shown as the function of Signal eff. versus Background eff. for the three initial NNs and corresponding background and Super NN with total background (four left upper curves). The single right down curve shows the Signal efficiency versus Super NN cut value (x-axis).

Table 14: The number of expected events which survive cut on Super NN output greater than 0.92 for 1 fb^{-1}. We present the sum of the yields and separate numbers for each Jet bin ($N_{jet} = 2,\ 3$). For comparison purposes we include statistical significance as well, defined as $S/\sqrt{S+B}$.

process	$N^{expected}_{Nj=2}$	$N^{expected}_{Nj=3}$	$N^{expected}_{tot}$ (N^{MC}_{tot})
$t\bar{t}$	35	50	85 (195)
$Wb\bar{b}j$	1	2	3 (9)
W+jets	35	6	41 (38)
tW	14	17	30 (64)
QCD	0	0	$0 \pm 113\ (mc\ stat.)$ (0)
Bkgd. total	85	74	159
t-channel Signal	427	190	618 (1460)
$S/\sqrt{S+B}$	19	12	22

Process	JES up	JES down	MET up	MET down	ϵ_{btag} down	ϵ_{btag} up	ϵ_{mis} down	ϵ_{mis} up
Signal	$+0.5\%$	-1.7%	$+4.5\%$	-5.3%	-3.3%	$+3.1\%$	-0.2%	$+0.1\%$
$t\bar{t}$	-8.3%	$+8.6\%$	$+2.7\%$	-3.7%	-5.9%	$+6.3\%$	-0.4%	$+0.7\%$
$Wbbj$	$+1.8\%$	-0.4%	$+5.3\%$	-7.1%	-3.6%	$+3.6\%$	-0.2%	≈ 0
$W+jets$	$+4\%$	-4.3%	$+5.9\%$	-7.5%	-1.8%	$+2.2\%$	-20%	$+28.8\%$
tW	-6.4%	$+4.9\%$	$+3.3\%$	-4.2%	-3.1%	$+2.9\%$	-1.3%	$+1.2\%$
QCD	-1.2%	-3.5%	$+10.3\%$	-12.6%	-1.2%	$+3.5\%$	-3.5%	$+2.3\%$

Table 15: Impact of the considered systematic errors of instrumental origin on the preselection efficiencies for the signal and the main backgrounds in neural network analysis.

Process	JES up	JES down	MET up	MET down	ϵ_{btag} down	ϵ_{btag} up	ϵ_{mis} down	ϵ_{mis} up
Signal	$+3.6\%$	-7.1%	$+4.5\%$	-7.0%	-4.3%	$+3.9\%$	-0.1%	$+0.1\%$
$t\bar{t}$	-7.3%	$+12.3\%$	$+1.7\%$	-5.5%	-4.3%	$+6.8\%$	$+0.5\%$	≈ 0
$Wbbj$	$+40\%$	$+24\%$	$+21\%$	-4%	-8%	$+4\%$	≈ 0	≈ 0
$W+jets$	$+31\%$	-22%	$+19\%$	-8%	-3%	≈ 0	-13%	$+20\%$
tW	-6.3%	$+6.3\%$	$+4.7\%$	$+1.6\%$	≈ 0	$+3.1\%$	-1.6%	-1.6%

Table 16: Impact of the considered systematic errors of instrumental origin on the final selection efficiencies for the signal and the main backgrounds in neural network analysis. QCD is not listed for lack of statistics.

Source	significance	$\Delta\sigma$ [pb]
statistical	18.7	5.6
background normalization	11.4	11.9
jet energy scale	11.0	4.6
$\not\!\!E_{\mathrm{T}}$ scale	16.7	4.5
b-tagging efficiency	10.3	5.0
mistagging	13.5	2.2
total	6.8	14.8

Table 17: Impact of the considered sources of uncertainty on the significance and the cross section uncertainty for the EC method (see Sec. 6.1) applied to the NN selection outcome.

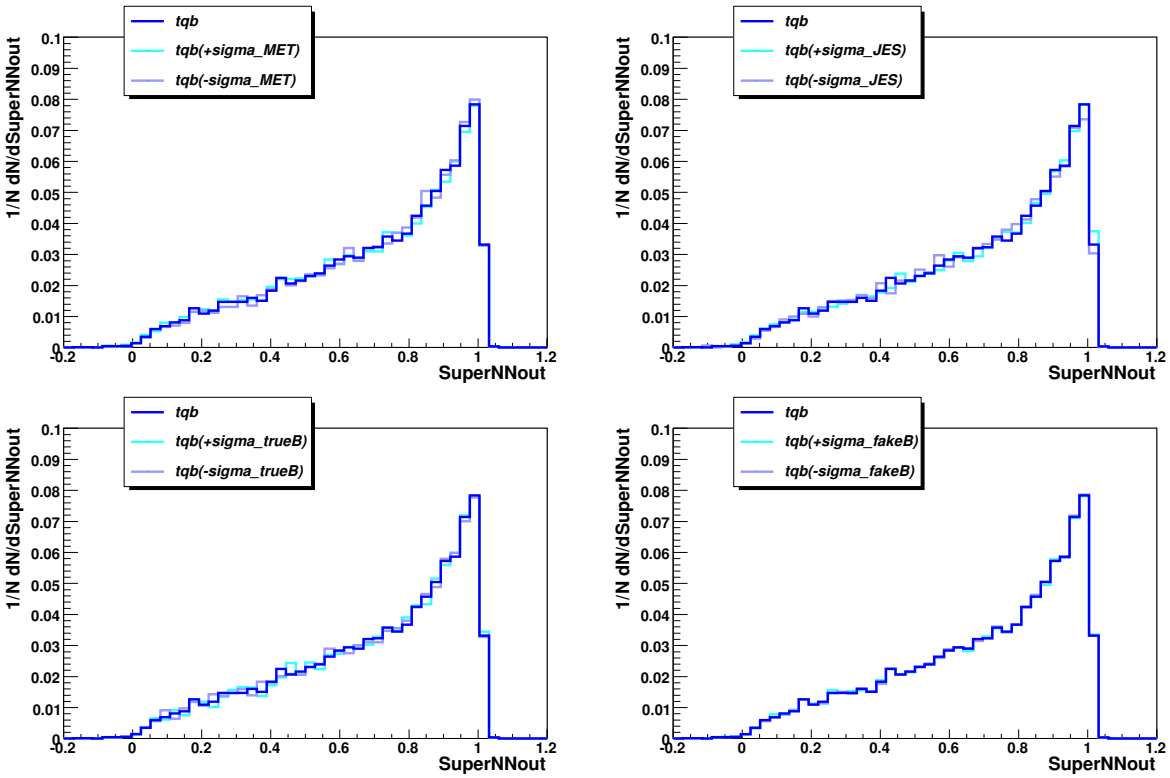


Figure 29: These plots demonstrate how the SuperNN output changes after the $\pm\sigma$ shifts with different source of systematic uncertainties (mentioned in the table 16.)

9 Conclusions and Outlook

The main outcome of the analysis presented here is that it is realistic to provide a confirmation of the Tevatron observation of single top with $\approx 1\ fb^{-1}$ of data at 14 TeV.

Different complementary methods would get us there, with a selection optimized for the dominant t channel:

- a simple event counting would give a 3.6σ evidence of the presence of signal over the expected backgrounds, in the muonic channel, the drawback being a large sensitivity to systematics and background contamination;
- under the same assumptions, by exploiting the ratio of positive over negative leptons it is possible to achieve a 2.9σ evidence, very robust against systematic uncertainties and with most backgrounds cancelling out naturally;
- an additional method exploits the most striking feature of the signal, namely the $\approx 100\%$ polarization of the top quark which is entirely propagated to the decay lepton; this note only presents a preliminary investigation, yielding very promising results regarding the stability of the method with respects to systematic effects;
- several other features of the signal can provide confirmation to the presence of signal, like the presence of a clear peak in the mass of the reconstructed top quark; a Neural Network analysis has been devised, in order to make the maximal use of the information available in the preselected events and obtain a very high purity at a relatively low cost in statistics.

Although some of the methods described are intrinsically robust against the general level of background contamination, special attention has been paid to the data-driven estimation of QCD events. The method chosen is an extrapolation to the selected region from the QCD-enriched and signal-depleted regions in the isolation versus M_T plane. It gives a correct prediction within 6%.

9.1 Plans

While the focus of this note has been explicitly set on updating and improving the previous results at 14 TeV [8], the highest priority is now to prepare the analyses to be run with the very first LHC data, which will be at a lower centre-of-mass energy for some time. New simulated datasets are starting to be available with a centre-of-mass energy of 10 TeV, and this analysis is now being redone under these new assumptions, and with a more up-to-date CMS software (with many small improvements from the point of view of reconstruction and analysis tools, and some more realism in the detector simulation which has been further tuned to cosmic data).

Nevertheless, we believe that the present selection, although already satisfactory from the point of view of the statistical significance, could be further improved if more time were to be devoted to the task, by intervening in the following area:

- better fake-muon rejection: some progress has been shown, in studies related to the $t\bar{t} \rightarrow 1\mu + jets$ analysis, on the QCD rejection with a better μ-id making use of the expected energy loss in the calorimeters [50];
- tighter second lepton veto: the $t\bar{t}$ background, as shown in Sec. 3.6, is almost equally composed of $2l$ (with $l = e, \mu$), $\tau + l$ and $l + jets$ events; so far we vetoed only the presence of a second lepton passing the same selection defining the first, and while this has the benefit of simplicity, it is probably not optimal against $t\bar{t} \rightarrow 2l$, which would be better rejected by applying a looser identification on the veto lepton;
- τ-jet veto: in the current analysis, τ-id is not used at all, and a τ jet is usually identified as a normal jet from quark or gluon;
- data-driven $t\bar{t} \rightarrow 1\mu + jets$ estimation: the remaining major component of the $t\bar{t}$ contamination is probably very hard to further reduce, but in the context of the $WW \rightarrow 2l$ analysis a method has been developed to estimate this contamination from data, by counting the number of muons inside jets [51];
- veto of electromagnetic objects in HF: the HF calorimeter acts simultaneously as an electromagnetic and hadronic calorimeter and, although this feature has not been exploited in the present analysis, this can further reduce the $t\bar{t} \rightarrow 2l$ background when the second lepton is an electron and falls in the HF acceptance, giving a fake "forward jet";

- use of the Particle Flow technique: this is expected to give more reliable $\not\!\!E_{\mathrm{T}}$ and jet-multiplicity measurements, with potential benefits to the analysis as shown by a preliminary investigation [52].

None of the items of this list are crucial for the analysis, and realistically they would only provide a marginal improvement in the results. Nonetheless, it is useful to keep in mind that there are additional handles for selection, and possibly some of them will be explored very soon in the context of the new 10 TeV analysis to which we intend to devote all our interest.

The next logical step after the re-discovery of single top will be the measurement of the cross section of the t channel. All of the methods discussed in this note (event counting and Charge Ratio, for both the cut- and NN-based analyses) can be applied to that end. The exploitation of the $\approx 100\%$ polarization of the top quark is currently being worked out in more details, and will be prominent in the forthcoming note at 10 TeV. This feature can be not only seen as a smoking gun for the confirmation that single top is actually what is being selected, but also as an effective way to estimate the backgrounds simultaneously with the signal cross section (with also a good robustness against the main systematics, as demonstrated in Sec. 7.4 of this note). In a later stage, with more accumulated statistics and an independent data-driven estimation of all the main background components, the same distribution can be used to constrain possible contributions from physics beyond the Standard Model.

Acknowledgements

We wish to thank Jorgen D'Hondt, Claudio Campagnari, Tim Christiansen, Roberto Tenchini, Frank-Peter Schilling, Fabio Maltoni and the MadGraph team, Roberto Chierici, Jérémy Andrea, Robert Harris, Philipp Schieferdecker.

References

[1] The D0 Collaboration, "Evidence for production of single top quarks", Phys.Rev. D 78, 012005 (2008)

[2] The CDF Collaboration, "Measurement of the Single Top Quark Production Cross Section at CDF", arXiv:0809.2581v2[hep-ex]

[3] J. Alwall et al., "Is $|V_{tb}| \approx 1$?", Eur Phys.J., C49, 791-801 (2007), arXiv:hep-ph/0607115v2

[4] T.M.P. Tait, C.-P. Yuan, "Single Top Production as a Window to Physics Beyond the Standard Model", Phys.Rev.D63:014018 (2001), arXiv:hep-ph/0007298

[5] N. Kidonakis, "Higher-order soft gluon corrections in single top quark production at the LHC", Phys.Rev. D75 (2007) 071501, arXiv:hep-ph/0701080

[6] B.W. Harris et al., "The fully differential single-top-quark cross section in next-to-leading order QCD", Phys.Rev. D66 (2002) 054024, arXiv:hep-ph/0207055

[7] J. Campbell, F. Tramontano, "Next-to-leading order corrections to Wt production and decay", Nucl.Phys. B726 (2005) 109-130, arXiv:hep-ph/0506289

[8] V. Abramov et al., "Selection of Single Top Events with the CMS Detector at LHC", CMS NOTE 2006/084

[9] F. Maltoni and T. Stelzer, "MadEvent: Automatic event generation with MadGraph," JHEP 0302, 027 (2003), arXiv:hep-ph/0208156

[10] E. Boos et al., "Method for simulating electroweak top-quark production events in the NLO approximation: SingleTop event generator", Phys.Atom.Nucl. 69, 8 (2006) 1317

[11] E. Boos et al. [CompHEP Collaboration], "CompHEP 4.4: Automatic computations from Lagrangians to events," Nucl. Instrum. Meth. A **534**, 250 (2004), arXiv:hep-ph/0403113

[12] S. Frixione and B. R. Webber, "Matching NLO QCD computations and parton shower simulations", JHEP **06**, 029 (2002), arXiv:hep-ph/0204244

[13] L. Dudko et al., "Comparison of the different Monte-Carlo models for the t-channel single top quark production", CMS Analysis Note 2009/024

[14] `https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideFastSimulation`

[15] J. Allison et al., "Geant4 developments and applications", IEEE Transactions on Nuclear Science 53 No. 1 (2006) 270-278

[16] M.L. Mangano et al., "ALPGEN, a generator for hard multiparton processes in hadronic collisions", JHEP 07 (2003) 001

[17] T. Sjostrand et al., "PYTHIA 6.4 physics and manual", JHEP 05 (2006) 026

[18] S. R. Slabospitsky and L. Sonnenschein, "TopReX generator (version 3.25): Short manual," Comput. Phys. Commun. 148, 87 (2002), arXiv:hep-ph/0201292

[19] https://twiki.cern.ch/twiki/bin/view/CMS/CSA07
https://twiki.cern.ch/twiki/bin/view/CMS/GeneratorProduction2007CSA07
https://twiki.cern.ch/twiki/bin/view/Main/AlpgenSummer07

[20] https://twiki.cern.ch/twiki/bin/view/CMS/DetectorPerformanceCSA07AlCaConstants

[21] R. Bonciani, S. Catani, M.L.Mangano, "NLL resummation of the heavy-quark hadroproduction cross-section", Nucl.Phys. B529 (1998) 424-450

[22] Review of Particle Physics C. Amsler, et. al., Physics Letters B 667, 1 (2008)

[23] https://twiki.cern.ch/twiki/bin/view/CMS/SWGuidePAT

[24] The CMS Collaboration, "CMS Physics TDR: Volume I (PTDR1), Detector Performace and Software", CERN-LHCC-2006-001; Chapter 10, "Electrons and Photons"

[25] https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideElectronID

[26] F. Bostock, talk given in http://indico.cern.ch/conferenceDisplay.py?confId=32732

[27] J. Caudron, A. Giammanco, D. Kcira, V. Lemaitre, "Early measurements of top quark pair events in the $e + \mu$ channel with the first data of CMS", CMS Analysis Note 2008/017

[28] J. Kiefer et al., "Observability of Top Quark Pair Production in the Semileptonic Muon Channel with the first 10 pb^{-1} of CMS Data", CMS Analysis Note 2008/014. The CMS Collaboration, CMS PAS TOP-08-005

[29] J. Kiefer, talk given in http://indico.cern.ch/contributionDisplay.py?contribId=2&confId=51557

[30] P. Yeh et al., "Search for W-Associated Production of Single Top Quarks in CMS", CMS NOTE 2006/086

[31] M. Dobbs, "Incorporating next-to-leading order matrix elements for hadronic diboson production in showering event generators", Phys.Rev. D64 (2001) 034016, arXiv:hep-ph/0103174

[32] The CMS Collaboration, "Plans for Jet Energy Corrections at CMS", CMS PAS JME-07-002

[33] https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideBTagging

[34] S. Esen et al., "Missing E_T Performance in CMS", CMS Analysis Note 2007/041

[35] M. Narain et al., "Performance Measurement of b tagging Algorithms Using Data containing Muons within Jets", CMS Analysis Note 2007/046

[36] J. Andrea et al., "Evaluation of $udsg$ Mistag Rate of b-tag Jets using Negative Tags", CMS Analysis Note 2007/048

[37] M. Cepeda et al., "Data driven methods for QCD background estimation in electroweak muon analysis: the ABCD and the Template methods", CMS Analysis Note 2008/113

[38] Y. Kemp, "Identification of Electrons in the Forward Region of the CDF Experiment for the Search for Electroweak Top Quark Production", PhD Thesis (2006),
http://www-ekp.physik.uni-karlsruhe.de/pub/web/thesis/iekp-ka2006-1.pdf

[39] N. Adam et al., "Towards a measurement of the inclusive $W \to e\nu$ and $Z \to e^+e^-$ cross section in pp collisions at $\sqrt{s} = 14$ TeV", CMS Analysis Note 2007/026.

[40] G. Mahlon and S. Parke, "Improved Spin Basis for Angular Correlation Studies in Single Top Quark Production at the Tevatron", Phys. Rev. D55, 7249 (1997), arXiv:hep-ph/9611367

[41] G. Mahlon and S. Parke, "Single Top Quark Production at the LHC: Understanding Spin", Phys.Lett. B476 (2000), arXiv:hep-ph/9912458

[42] S. Cerci, D. d'Enterria, "Low-x QCD studies with jets in the CMS Hadron Forward calorimeter in proton-proton collisions at $\sqrt{s}$ = 14 TeV", CMS Analysis Note 2008/060. The CMS Collaboration, CMS PAS FWD-08-001.

[43] J. Alcaraz et al., "Measurement of muon charge asymmetry in the $W \to \mu\nu$ channel", CMS Analysis Note 2008/074

[44] M. Chen, talk available at `https://twiki.cern.ch/twiki/pub/CMS/MuonChargeRatio/20081207_ChargeConfusion_mingshui.pdf`

[45] M. Bona et al., "The $Z \to \mu\mu$ + jets data candle", CMS Analysis Note 2008/095

[46] V.Abramov et al., "Investigation of the single top quark production in t-channel process by a Neural Network approach with the CMS detector", CMS IN-2007/045

[47] E. Boos, L. Dudko, V. Savrin, "*SingleTop* - an event generator for the single top quark production at the LHC", CMS NOTE 2000/065

[48] J. Schwindling, B. Mansoulie, "MLPfit: a Tool for Multi-Layer Perceptrons (version 1.40)", `http://schwind.home.cern.ch/schwind/MLPfit.html`

[49] E. Boos et al., "Method of *Optimum Observables* and Implementation of Neural Networks in Physics Investigations", Phys. Atom. Nucl. **71**, 2 (2008) 383-393

[50] P. Van Mulders, talk given in `http://indico.cern.ch/conferenceDisplay.py?confId=43145`

[51] D. Kovalskyi, talk given in `http://indico.cern.ch/conferenceDisplay.py?confId=40788`

[52] A. Giammanco, talk given in `http://indico.cern.ch/conferenceOtherViews.py?view=cdsagenda&confId=35611`