Version 0.3

Measurement of the $\mathrm{t\bar{t}}$ Production Cross-section at $\sqrt{s} = 1.96$ TeV in the Lepton+jets Final States using a Topological Method

Top Production group

(Dated: February 24, 2004)

A measurement of the $t\bar{t}$ production cross-section at $\sqrt{s} = 1.96$ TeV, based on the application of a topological method to data preselected in the e+jets (141.2 pb^{-1}) and μ+jets (143.9 pb^{-1}) channels, is presented. The preliminary cross-sections obtained for each of the lepton+jets channels are:

$$\mathrm{e + jets} \ : \quad \sigma_{p\overline{p}\rightarrow t\bar{t}+X} = 9.4^{+3.9}_{-3.5}\,(\mathrm{stat})\,^{+1.6}_{-2.6}\,(\mathrm{syst})\ \pm 0.6\,(\mathrm{lumi})\,\mathrm{pb},$$
$$\mu + \mathrm{jets} \ : \quad \sigma_{p\overline{p}\rightarrow t\bar{t}+X} = 7.5^{+3.4}_{-3.0}\,(\mathrm{stat})\,^{+1.9}_{-1.6}\,(\mathrm{syst})\ \pm 0.5\,(\mathrm{lumi})\,\mathrm{pb},$$

leading to the following preliminary combined cross-section:

$$\mathrm{l + jets} \ : \ \sigma_{p\overline{p}\rightarrow t\bar{t}+X} = 8.4^{+2.5}_{-2.3}\,(\mathrm{stat})\,^{+1.5}_{-1.8}\,(\mathrm{syst})\ \pm 0.5\,(\mathrm{lumi})\,\mathrm{pb}.$$

Preliminary Results for Winter 2004 Conferences

Contents

Trigger List	$\int \mathcal{L}$ (pb^{-1}) $\mu + jets$	$e + jets$
v8	20.8	20.2
v9	18.6	18.6
v10	14.9	14.8
v11	54.8	52.7
v12	34.7	34.9
total	143.9	141.2

TABLE I: Breakdown of integrated luminosities by trigger list version.

I. INTRODUCTION

In $p\bar{p}$ collisions at Tevatron energies the Standard Model (SM) top quark is expected to be dominantly produced via the strong interaction ($q\bar{q}$ annihilation (85%) and gluon fusion (15%)). The top quark decay proceeds quickly virtually 100% of the time to a W boson and a b quark. The final states which are visible in the detector are fully determined by the W boson decay modes. The production and decay of top quark pairs can therefore result in final states possessing one isolated, high p_T lepton, missing transverse energy ($\not\!\!E_T$) from a neutrino, and associated jets. The dominant physics background to such a signature arises from W boson production accompanied by three or four jets. Instrumental backgrounds for these channels originates primarily from QCD multijet processes with π^0's and γ's misidentified as electrons (e+jets channel) or from semileptonic heavy flavor decays faking isolated muons (μ+jets channel).

The analysis described in this note presents results from the topological analyses in the electron+jets and muon+jets channels. For information concerning how these analyses were performed with p13 data, please consult Refs. [1, 2].

II. DATA SET

The data sample used consists of data taken during the period between June 2002 and September 2003. The detailed description of the data set, run and luminosity block quality selection and total integrated luminosity can be found in [3]. Table I gives a breakdown of the integrated luminosities by trigger version for the two lepton+jets analyses and the total integrated luminosity of the data set for each analysis.

A. Triggering

The triggers used in the lepton+jets analyses are described in Section III of [3].

B. Additional Control Data Samples

1. QCD in μ+jets channel (μQCD_{presel})

In order to derive the muon isolation efficiency, ϵ_{qcd}, in QCD events using the matrix method, a sample is selected with the same requirements as for the preselection, as described in Section VII A, without applying the $\not\!\!E_T$ cut. Instead the following two cuts are applied:

- $\not\!\!E_T < 10$ GeV
- $E_T^W = \not\!\!E_T$ without muon correction < 25 GeV

2. QCD in μ+jets channel (μQCD)

To describe the QCD background which survives the preselection, as described in Section VII A, a sample is selected by applying the loose preselection and inverted isolation cut:

- Rat11 < 2.5 GeV

- RatTrk < 2.5 GeV

The definitions of muon isolation variables can be found in [3].

Since the statistics is very low in this sample also a looser version of this QCD definition, (μQCD$_{loose}$), is introduced, where the requirement is:

- Rat11 < 2.5 GeV or RatTrk < 2.5 GeV

3. QCD in e+jets channel

In order to derive the electron likelihood efficiency, ϵ_{qcd}, in QCD events using the matrix method, a sample is selected with the same requirements as for the loose preselection described in Section VIII A but without applying the $\not\!\!E_T$ cut. To describe the QCD background which survives the preselection a sample is selected by applying the loose preselection and inverted likelihood cut:

- $L_{CC} < 0.75$ or $L_{EC} < 0.8$

III. OBJECT IDENTIFICATION

In order to isolate $t\bar{t}$ events from background, we define four fundamental objects: electrons, muons, jets and missing transverse energy $\not\!\!E_T$. These objects are described briefly below, and more detailed treatment can be found in [3].

A. Vertex Selection

In order to ensure high quality of the reconstructed vertex ([4]) it is required to be within the SMT fiducial region ($|PV_z| \leq 60$ cm) and have a track multiplicity ≥ 3. The efficiency of these cuts is evaluated using two methods. The first one makes use of the preselected signal samples as described in VII A and VIII A for each channel. Primary vertex efficiencies as a function of jet multiplicity are presented in Table II. In the second method, vertex cut efficiencies are evaluated using Z+jets data sample and are then corrected for the topological difference between Z+jets and $t\bar{t}$ events using the ratio of efficiencies measured on corresponding Monte Carlo samples. Tables III and IV present the efficiencies measured in Z events in data and MC along with the correction factors. Unfortunately, due to the lack of Z events in data with three or more jets we are not able to measure the correction factor for the jet multiplicity relevant for $t\bar{t}$ events. Nevertheless, from Tables III and IV we can conclude that the data-to-MC scale factor for the primary vertex selection efficiency tends to approach unity at high jet multiplicity. Table V summarizes PV efficiencies evaluated in data for the fourth jet multiplicity bin and in $t\bar{t}$ and W+4jets MC. Since all methods give consistent results we take the primary vertex selection efficiency measurement from data (i.e. left column in Table V). The statistical error on this measurement is taken as a systematic uncertainty on the primary vertex selection efficiency.

	e+jets	μ+jets
$\mathrm{N_{jet}} = 1$	96.2 ± 0.1	91.5 ± 0.4
$\mathrm{N_{jet}} = 2$	96.7 ± 0.3	90.2 ± 0.6
$\mathrm{N_{jet}} = 3$	96.6 ± 0.7	92.6 ± 1.1
$\mathrm{N_{jet}} \geq 4$	98.9 ± 0.8	95.2 ± 1.9

TABLE II: Primary vertex cuts efficiencies (in %) as measured in the preselected signal samples as a function of jet multiplicity.

B. Electrons

We consider three steps in the electron identification: electromagnetic cluster reconstruction, track matching, and passage of a likelihood test. Detailed description of all elements of electron identification can be found in section V of [3]. The likelihood is described extensively in [5].

	$\epsilon_{data}(Z \to ee)$	$\epsilon_{MC}(Z \to ee)$	scale factor
$\mathrm{N_{jet}} = 0$	93.3 $\pm$ 0.3	97.7 $\pm$ 0.2	0.955 $\pm$ 0.003
$\mathrm{N_{jet}} = 1$	95.3 $\pm$ 0.5	98.1 $\pm$ 0.1	0.972 $\pm$ 0.006
$\mathrm{N_{jet}} = 2$	98.9 $\pm$ 0.6	98.3 $\pm$ 0.1	1.006 $\pm$ 0.007

TABLE III: Primary vertex cuts efficiencies (in %) in $Z \to ee$ data, MC and a correction factor as a function of jet multiplicity.

	$\epsilon_{data}(Z \to \mu\mu)$	$\epsilon_{MC}(Z \to \mu\mu)$	scale factor
$\mathrm{N_{jet}} = 0$	95.3 $\pm$ 0.3	98.6 $\pm$ 0.2	0.966 $\pm$ 0.004
$\mathrm{N_{jet}} = 1$	97.0 $\pm$ 0.5	98.8 $\pm$ 0.1	0.982 $\pm$ 0.005
$\mathrm{N_{jet}} = 2$	98.6 $\pm$ 0.8	98.7 $\pm$ 0.1	0.999 $\pm$ 0.008

TABLE IV: Primary vertex cuts efficiencies (in %) in $Z \to \mu\mu$ data, MC and a correction factor as a function of jet multiplicity.

The initial identification proceeds from EM cluster reconstruction and adds the following requirements on its EM fraction, isolation and shower shape:

- $f_{EM} > 0.9$
- $f_{iso} < 0.15$
- hmx8 $\chi^2 < 75$

Electron candidates are also required to have an associated track. If an electron satisfies the criteria mentioned above and have p_T >15 GeV we call it loose electron. If in addition it passes the likelihood cuts, $L_{CC} > 0.75$ or $L_{EC} > 0.8$, we call it tight electron.

The efficiencies for finding an EM cluster, cluster identification cuts, track match and likelihood as well as the methods used to evaluate them are summarized in section V of [3]. All efficiencies are measured in data using a $Z \to ee$ sample. To account for the difference in topology between $t\bar{t}$ and Z events corresponding efficiencies are also evaluated with $t\bar{t}$ and Z MC (see Table VI). In the analysis, data to MC correction factors or "topological" scale factors ($\epsilon^{MC}_{t\bar{t}}/\epsilon^{MC}_{Z}$) are applied where appropriate.

C. Muons

The details of the muon reconstruction are described in section VI of [3]. Muons accepted in the analyses are required to be well reconstructed. In particular, we select medium $n_{seg} = 3$ muons which have a track reconstructed in the central tracking system matched to it, and satisfying a cosmic veto based on scintillator timing cuts.

Isolation criteria are imposed to distinguish muons originating from W decays from the ones coming from heavy quark decays. In the μ+jets analysis we demand that a muon is separated from a jet $\Delta R(\mu, jet) > 0.5$ and satisfy tight isolation criteria based on variables involving the ratio of the visible energy surrounding the muon and its p_T (see section VI of [3]). The cuts on these variables are optimized to achieve strong QCD background suppression in the $\mu + jets$ channel.

We demand that the isolated muon is consistent with originating from the primary vertex. This involves cutting on both r-ϕ distance-of-closest approach significance ($|dca|/\sigma_{dca} < 3$) and the distance along the beam direction between the muon and the primary vertex ($|\Delta Z(\mu, PV)| < 1$ cm).

	$\epsilon^{PV,data}$	$\epsilon^{PV,MC_{t\bar{t}}}$	$\epsilon^{PV,MC_{Wjjjj}}$
e+jets	98.9 ± 0.8	97.8 ± 0.2	98.0 ± 0.1
μ+jets	95.2 ± 1.9	97.9 ± 0.2	98.2 ± 0.2

TABLE V: Efficiencies (in %) of primary vertex selection measured in the preselected signal samples and in $t\bar{t}$ and $W + jets$ Monte Carlo.

	CC			EC		
	Z^{data}	Z^{MC}	$t\bar{t}^{MC}$	Z^{data}	Z^{MC}	$t\bar{t}^{MC}$
cluster reconstruction efficiency (%)	96.0±0.4±1.5	97.6±0.1±0.2		93.5±1.2±1.5	96.6±0.2±0.2	
cluster selection efficiency (%)	99.7±0.1	98.35±0.03	96.7±0.2	99.9±0.1	99.49±0.02	99.0±0.2
track match efficiency (%)	91.3±0.4	96.63±0.04	95.6±0.2	68.0±0.8	86.5±0.1	84.3±0.9
likelihood efficiency (%)	89.1±0.4	95.68±0.05	94.4±0.2	86.3±0.7	93.79±0.08	94.2±0.6

TABLE VI: EM efficiencies in data and MC.

D. Jets

Jets are reconstructed via the Run II cone algorithm with $\Delta R = 0.5$. The new T42 algorithm [6] designed to provide finer and more uniform treatment of the calorimeter noise is applied before reconstructing the calorimeter objects. We employ further selections which remove isolated electromagnetic particle contribution and suppress jets originating from instrumental problems (hot cells, noisy regions). These include

- $0.05 < f_{EM} < 0.95$,
- $f_{CH} < 0.4$,
- hot cell ratio < 10.0,
- $n90 > 1$.

To further reject jets clustered on noise in the calorimeter a jet is required to have sufficient energy deposited at Level 1 calorimeter trigger system. This requirement was found to be very powerful discriminant against noise. Jet reconstruction and the choice of cuts is discussed in detail in section VII of [3].

For the lepton+jets analyses we accept jets with p_T >15 GeV and within $|\eta| < 2.5$.

E. Missing E_T

The presence of the neutrino in the final state is detected from the imbalance of the transverse energy in an event. We utilize the standard procedure for $\not\!\!E_T$ calculation described in section VII of [3]. If a central track-matched muon is found in an event, the transverse missing energy is corrected for its presence since it deposits only a small amount of energy in the calorimeter.

IV. SIGNAL-TO-BACKGROUND DISCRIMINATION

The two main backgrounds to $t\bar{t}$ pair production in the lepton+jets channels arise from the W multijet production (physics background) and QCD multijet events constituting the main instrumental background. In this section we describe the methods used to separate $t\bar{t}$ signal from he background contributions in the lepton+jets channels. In the discussion below we use the $\mu + jets$ channel as an example.

A. Statistical Methods to extract the Signal and Background Conributions

Only events which passed the preselection, as discussed in sections VII A and VIII A, are considered here. The preselected data set enriched in W+jets can be subdivided into three components:

- The $t\bar{t}$ signal
- $W + jets$ background, the electroweak W production plus initial state radiation (ISR) of four or more jets
- QCD background, strong production of four or more jets

Only the MC simulation of the $t\bar{t}$ signal and the $W + jets$ background are used. The MC simulation of QCD is abandoned due to the large uncertainties in the modelling and a QCD sample is obtained from data by reversing the lepton tight requirement (electron likelihood or muon isolation, respectively for e+jets and μ+jets; see Section II B). The goal is to disentangle all three components. The following methods, common to both lepton+jets channels but here exemplified in the μ+jets channel, are used:

1. *Matrix Method - Evaluation of the QCD Background*

In a first step the instrumental background from fake isolated muons is evaluated. This background is essentially due to heavy flavor QCD events where the muon originates from a semi-leptonic heavy quark decay. Real isolated muons are produced in W decays in $W + jets$ background events or $t\bar{t}$ signal events.

Two samples of events, a loose and a tight set are used, the latter being a subset of the first. The loose set (N_ℓ) corresponds to the preselection sample, with the loose muon isolation cut ($\Delta R(\mu, jet) > 0.5$). The tight sample (N_t) is defined with a requirement of the tight muon isolation. If the total number of $W + jets$ background, $t\bar{t}$ signal and QCD background events are denoted by $N^{\mathrm{W+t\bar{t}}}$ and N^{QCD} respectively, N_ℓ and N_t can be written:

$$\begin{aligned} N_\ell &= N^{\mathrm{W+t\bar{t}}} + N^{\mathrm{QCD}} \\ N_t &= \varepsilon_{sig} N^{\mathrm{W+t\bar{t}}} + \varepsilon_{QCD} N^{\mathrm{QCD}} \end{aligned} \tag{1}$$

Solving this linear system for N^{QCD} and $N^{\mathrm{W+t\bar{t}}}$ yields:

$$N^{\mathrm{W+t\bar{t}}} = \frac{N_t - \varepsilon_{QCD} N_\ell}{\varepsilon_{sig} - \varepsilon_{QCD}} \quad \text{and} \quad N^{\mathrm{QCD}} = \frac{\varepsilon_{sig} N_\ell - N_t}{\varepsilon_{sig} - \varepsilon_{QCD}} \tag{2}$$

The efficiency (ε_{sig}) for $t\bar{t}$ signal and $W + jets$ background has been studied in $Z \to \mu\mu$ events and the dependency on the number of reconstructed jets is taken from the simulation. The rate for which a muon in QCD events appears isolated was measured in the $\mu\mathrm{QCD}_{presel}$ data sample as described in Section II B 3.

2. *The 2^{nd} Matrix Method - Evaluation of the $W + jets$ Background*

One approach to separate the $t\bar{t}$ sample from $W + jets$ events uses a method, referred to as 2^{nd} matrix method, which can be reduced to the matrix method (see Section IV A 1). Instead of by the muon isolation, the $W + jets$ and $t\bar{t}$ signal sample can be distinguished by a topological selection. The number of events in the tight sample (N_t) and the selected number of events (N_{sel}) can be written as:

$$\begin{aligned} N_t &= N_t^{\mathrm{t\bar{t}}} + N_t^{\mathrm{W}} + N_t^{\mathrm{QCD}} \\ N_{sel} &= \varepsilon_{sel}^{\mathrm{t\bar{t}}} N_t^{\mathrm{t\bar{t}}} + \varepsilon_{sel}^{\mathrm{W}} N_t^{\mathrm{W}} + \varepsilon_{sel}^{QCD} N_t^{\mathrm{QCD}} \ . \end{aligned} \tag{3}$$

where the ε_{sel}^{i} are the efficiencies for a topological selection. This can be rewritten by bringing the QCD terms to the left side which yields:

$$\begin{aligned} N_t - N_t^{\mathrm{QCD}} &= N_t^{\mathrm{t\bar{t}}} + N_t^{\mathrm{W}} \\ N_{sel} - \varepsilon_{sel}^{QCD} N_t^{\mathrm{QCD}} &= \varepsilon_{sel}^{\mathrm{t\bar{t}}} N_t^{\mathrm{t\bar{t}}} + \varepsilon_{sel}^{\mathrm{W}} N_t^{\mathrm{W}} \ . \end{aligned} \tag{4}$$

N_t^{QCD} was already determined by the matrix method as described in Section IV A 1 so that the system of linear equations reduces to the matrix method (with a inhomogeneous part), which can be solved directly for $N_{\mathrm{sel}}^{\mathrm{t\bar{t}}}$:

$$N_{sel}^{\mathrm{t\bar{t}}} = \varepsilon_{sel}^{\mathrm{t\bar{t}}} \frac{N_{sel} - \varepsilon_{sel}^{\mathrm{W}} N_t - (\varepsilon_{sel}^{QCD} - \varepsilon_{sel}^{\mathrm{W}}) N_t^{\mathrm{QCD}}}{\varepsilon_{sel}^{\mathrm{t\bar{t}}} - \varepsilon_{sel}^{\mathrm{W}}} \tag{5}$$

$N_{sel}^{\mathrm{t\bar{t}}}$ is hardly sensitive to the precision with which W and QCD are separated since ε_{sel}^{QCD} and $\varepsilon_{sel}^{\mathrm{W}}$ are very close. In the limit when $\varepsilon_{sel}^{QCD} = \varepsilon_{sel}^{\mathrm{W}}$ the 2^{nd} matrix method boils down to the regular matrix method and there is no need to evaluate the relative fractions of W and QCD.

B. Topological Analysis

The preselection served to extract the $W \to l\nu$ signal with maximum efficiency while rejecting the QCD background as much as possible. In a second step, event shape variables showing a good separation power between $t\bar{t}$ and $W + jets$ are considered. Ten topological variables were tested for their discrimination power, their correlation and their sensitivity against systematic uncertainties, where only the dominant systematic uncertainty was taken into account originating from the jet energy scale (JES).

1. *The Choice of Topological Variables*

The criteria for the choice of the topological variables are the following:

- use variables with a large separation power between signal and background.
- use variables with a small sensitivity against systematic uncertainties (the JES-associated systematic uncertainty is the only one considered in this discussion).
- use variables which are not correlated or which have a small correlation. Correlated variables add only new information when the correlation is different for signal and background. However, every added variable adds noise and can enlarge both the statistical and the systematical uncertainty.

The topological variables can be roughly subdivided into two categories:

- variables which are proportional to the energy present in the event, e.g. the scalar sum over all jet transverse energies $H_T = \Sigma E_T$. These variables have a very good separation power since one peculiarity of the $t\bar{t}$ events is the large energy which is needed to produce a $t\bar{t}$ quark pair. However these variables are also very sensitive to uncertainties in the jet energy scale, which translates into a large systematic uncertainty on the cross section. With increasing integrated luminosity the statistical error will be reduced steadily. Soon the uncertainty on the cross section will be purely limited by systematical uncertainties, which excludes the usage of these variables.
- the second class of variables is characterized by the property of having reduced sensitivity to JES uncertainties. The variables are either defined purely by angular qualities, e.g. sphericity, or by ratios of energy dependent variables where the JES systematic uncertainty largely cancels out.

As described in Section IV B 2, a likelihood dicriminant can be built from the various topological variables.

The optimization procedure for the choice of topological variables can be summarized as follows:

- Only MC information from the $t\bar{t}$ and the $W + jets$ simulation is used to do the optimization. The QCD background is neglected since it represents a contribution to the overall background $\leq 20\%$ and it has been found to have an event topology very similar to W+jets. This simplification can only result in a loss of statistical sensitivity but no bias is introduced.
- It is shown that the distributions in data are reproduced by the simulation by comparing the shape of the used variables in the second and third jet multiplicity bin. In data the $W+jets$ contribution in these bins is extracted after subtracting the QCD background as determined from the matrix method.
- Due to limited MC statistics for the W+jets background, the same simulated events are used both to retrieve the shape of the discriminant variable and to determine the efficiency for the cut, respectively the performance of the fit.
- Whenever ensemble testing is performed, the events for one ensemble are drawn randomly form the lists of signal and background events, then the events are put back. Some events will be picked more than once.

Figure 1 points out the difference in separation power for likelihood dicriminants built from the two different classes of topological variables. For the given reasons only variables which are not sensitive to the JES uncertainty were considered as the input for the likelihood. Figure 2 shows that a set of the following four variables appears to have the best performance:

Sphericity: the normalized momentum tensor $\mathcal{M}$ is defined as:

$$\mathcal{M}_{ij} = \frac{\Sigma_o p_i^o p_j^o}{\Sigma_o |\vec{p^o}|^2}, \tag{6}$$

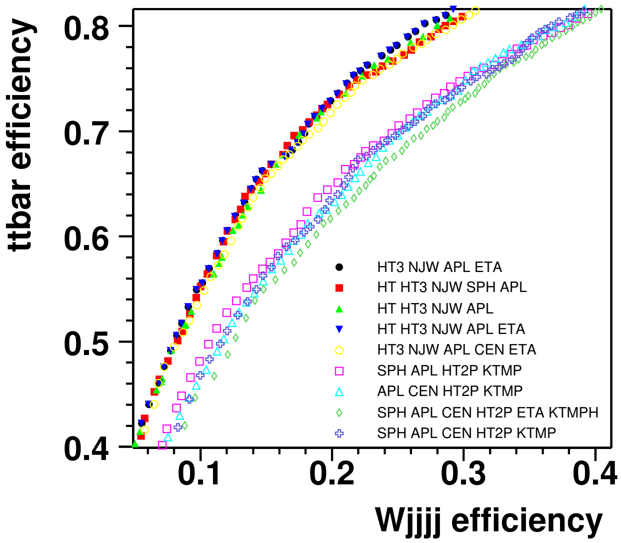

FIG. 1: Likelihood discriminant performance curves: The survival efficiency for the background ($W + jets$) is plotted vs. the survival efficiency for the signal ($t\bar{t}$) for different cut values for the different likelihood discriminant. The discriminants are built from variables given in the legend. In particular HT, HT3 and NJW belong to the category of variables which are proportional to the energy in the event, APL, SPH, CEN, HT2P, KTMP belong to the category of variables which are immune against JES uncertanties.

where $\vec{p^o}$ is the momentum-vector of a reconstructed object o, i and j are Cartesian coordinates. By standard diagonalization of $\mathcal{M}_{ij}$ one may find three eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$, with $\lambda_1 + \lambda_2 + \lambda_3 = 1$. The sphericity of the event is then defined as

$$S = \frac{3}{2}(\lambda_2 + \lambda_3), \tag{7}$$

so that $0 \geq S \geq 1$. Sphericity is essentially a measure of the summed $p_\perp^2$ with respect to the event axis; a 2-jet event corresponds to $S \approx 0$ and an isotropic event to $S \approx 1$. $t\bar{t}$ events are quite isotropic as is typical for the decay of a heavy object. $W + jets$ and QCD events are less isotropic, primarily due to the fact that the jets in these events arise from gluon radiation.

Aplanarity: the aplanarity $\mathcal{A}$, a measure of the flatness of the event, is defined as $\mathcal{A} = \frac{3}{2}\lambda_3$, where λ_3 is the smallest eigenvalue of the normalized momentum tensor $\mathcal{M}$ (see Eq. 6). Fig. 3 shows the performance of $\mathcal{A}$ for different objects in the sum. The objects included in the sum are the jets and the μ from the W decay, which have the best discrimination power between $t\bar{t}$ and $W + jets$. Large values of $\mathcal{A}$ are indicative of spherical events, whereas small values correspond to more planar events. $t\bar{t}$ events are quite spherical as is typical for the decay of a heavy object. $W + jets$ and QCD events are more planar, primarily due to the fact that the jets in these events arise from gluon radiation.

Ht2p: $H'_{T2} = H_{T2}/H_z$: provides a measure of event centrality. H_{T2} is given by the scalar sum of the p_T of the jets (H_T), excluding the leading jet. H_z is the scalar sum of the $|p_z|$ of the jets, lepton and neutrino;

Ktminp: $K'_{Tmin} = \Delta R_{jj}^{min} E_T^{min} / E_T^W$: provides a measure of the minimum jet p_T relative to another. In order to reduce the correlation of this variable with the jet energy scale, it is divided by another JES-sensitive variable: $E_T^W = E_T^{lepton} + \not{E}_T$. Only the four leading jets in an event are considered in the definition of this variable. ΔR_{jj}^{min} corresponds to the minimum separation in $\eta - \phi$ space between a pair of jets and E_T^{min} is the E_T of the lesser jet of that pair.

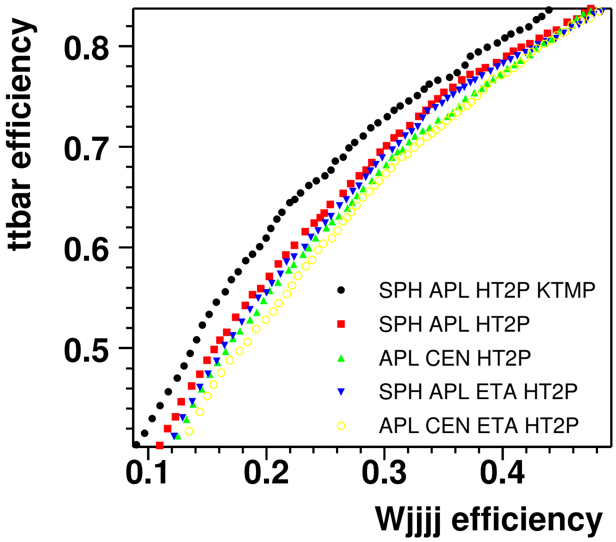

FIG. 2: Likelihood discriminant performance curves: The survival efficiency for the background ($W + jets$) is plotted vs. the survival efficiency for the signal ($t\bar{t}$) for different cut values for the different likelihood discriminant. Only variables which are immune against JES uncertanties are used here.

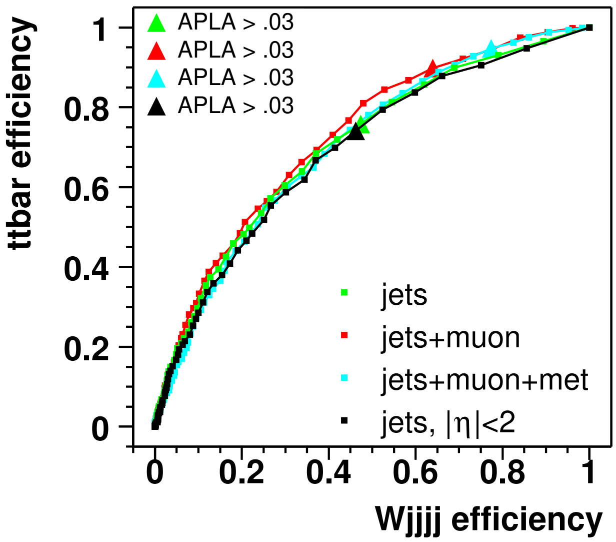

FIG. 3: Aplanarity performance curves: The survival efficiency for the background ($W + jets$) is plotted vs. the survival efficiency for the signal ($t\bar{t}$) for different cut values for the aplanarity. For convenience the efficiencies for $\mathcal{A} > 0.03$ are highlighted. Including the jets and the μ from the W decay in the sum leads to the best separation power.

2. The Topological Likelihood Discriminant

In order to combine the information of the topological variables to separate $t\bar{t}$ signal and the background a likelihood discriminant is built. Alternatively, the combination in a neural network, which takes into account correlations of the variables, was studied. The improvements of the neural network as compared to the likelihood were not considerable and therefore it will not be discussed here.

We have chosen to use a 2-class likelihood with $t\bar{t}$ as signal and $W+jets$ as background instead of a 3-class likelihood where in addition to the $W+jets$ background class a QCD background class would be considered. These are the considerations that went into such decision:

- the topological properties for the W background and the QCD background are found to be very similar and
- the QCD background fraction is relatively small.

The likelihood discriminant is built in following way:

- The MC statistics for the central mass point which was chosen to be 175 GeV is limited. It was found that the likelihood discriminant shape is rather independent of the top mass (see Table XIV and Fig. 27). The two MC samples with a top mass of 170 GeV and 180 GeV are combined and used for the signal input distributions. However, the likelihood output is evaluated with the central mass point in order to avoid biases from correlated statistical fluctuations.
- The topological variables are transformed using functions in order to be less sensitive to statistical fluctuations in rapidly varying regions. The used functions are:
  - $\ln(S)$
  - $\exp(-11 \cdot \mathcal{A})$
  - $\ln(H'_{T2})$
  - $\sqrt{K'_{Tmin}}$
- It is made sure that bins are adequately populated. Bins are combined so that the minimum number of entries per bin is larger than 25.
- The distributions are normalized to unity and the ratio of signal ($t\bar{t}$) over background ($W+jets$) is built for each of the four distributions.
- The logarithm of the ratios is built and fitted with a polinomial. Fitting the logarithm simplifies the fit function and symmetrises the errors on the points.
- Data in the tail of a distribution is consolidated including over- and underflow bins.
- A reasonable fit range is chosen. If the abscissa lies beyond this range the fit is evaluated at the range limit.
- The likelihood output $\mathcal{L}$ is given by

$$\mathcal{L} = \frac{\exp\left(\sum_i (\ln \frac{S}{B})^i_{\text{fitted}}\right)}{\exp\left(\sum_i (\ln \frac{S}{B})^i_{\text{fitted}}\right) + 1} , \tag{8}$$

  where $(\ln \frac{S}{B})^i_{\text{fitted}}$ is the fit to $\ln(\frac{\mathcal{P}^i_{\text{t}\bar{\text{t}}}}{\mathcal{P}^i_{\text{W}}})$ for each topological variables i that is used to build the discriminant, and $\mathcal{P}^i_{\text{t}\bar{\text{t}}}$ and $\mathcal{P}^i_{\text{W}}$ are the probabilities for the topological variable i to be $t\bar{t}$ signal and $W+jets$ background, respectively.

V. CROSS-SECTION EXTRACTION PROCEDURE

3. Likelihood Fit Method

The number of $t\bar{t}$ events in the preselected data sample can be directly extracted by performing a likelihood fit to the likelihood discriminant (see Section IV B 2) distribution in data using templates for the $t\bar{t}$ signal and the W and QCD backgrounds. However, the relative fractions of $W+jets$ and QCD background are already measured as a function of the $t\bar{t}$ cross section by applying the matrix method that was discussed in Section IV A 1. This circumstance allows to perform a fit where the number of $t\bar{t}$ ($N_t^{\mathrm{t\bar{t}}}$), $W + jets$ (N_t^{W}) and QCD (N_t^{QCD}) are allowed to float, but the relative fractions of $W + jets$ and QCD are constrained by the matrix method measurement. This is realized by defining the likelihood in the following way:

$$L(N_t^{\mathrm{t\bar{t}}}, N_t^{\mathrm{W}}, N_t^{\mathrm{QCD}}) = \prod_i P(n_i^{\mathrm{obs}}, \mu_i) \cdot P(N_{\ell-t}^{\mathrm{obs}}, N_{\ell-t}) \tag{9}$$

where $P(n,\mu)$ generically denotes the Poisson probability density function for n observed events given a expectation of μ. In the first term of Eq. 9, i runs over all bins of the likelihood discriminat histogram, n_i^{obs} is the content of bin i as obtained in the preselected sample, and μ_i is the expectation, which is a function of $N_t^{\mathrm{t\bar{t}}}$, N_t^{W} and N_t^{QCD}:

$$\mu_i(N_t^{\mathrm{t\bar{t}}}, N_t^{\mathrm{W}}, N_t^{\mathrm{QCD}}) = f_i^{\mathrm{t\bar{t}}} N_t^{\mathrm{t\bar{t}}} + f_i^{\mathrm{W}} N_t^{\mathrm{W}} + f_i^{\mathrm{QCD}} N_t^{\mathrm{QCD}} \tag{10}$$

where $f_i^{\mathrm{t\bar{t}}}$, f_i^{W}, f_i^{QCD} are the fractions in bin i of the $\mathrm{t\bar{t}}$, W and QCD likelihood discriminat templates.

The second term of Eq. 9 is a poisson constraint on the observed number of events in the "loose-tight" preselected sample and effectively completes the incorporation of the Matrix Method in the likelihood. By introducing $N_{\ell-t} = N_\ell - N_t$ one makes sure that $N_{\ell-t}$ and N_t (implicitly present in the likelihood discriminant histogram) are indeed uncorrelated by construction.

The next step is to express $N_{\ell-t}$ in terms of $N_t^{\mathrm{t\bar{t}}}$, N_t^{W} and N_t^{QCD}. Equation (2) can be rewritten in the following way:

$$\begin{aligned} N_\ell &= N^{\mathrm{t\bar{t}}} + N^{\mathrm{W}} + N^{\mathrm{QCD}} \\ N_t &= \varepsilon_{sig} N^{\mathrm{t\bar{t}}} + \varepsilon_{sig} N^{\mathrm{W}+} + \varepsilon_{QCD} N^{\mathrm{QCD}} \end{aligned} \tag{11}$$

Introducing also the number of preselected events as $N_t^{\mathrm{t\bar{t}}} = \varepsilon_{sig} N^{\mathrm{t\bar{t}}}$, $N_t^{\mathrm{W}} = \varepsilon_{sig} N^{\mathrm{W}}$, and $N_t^{\mathrm{QCD}} = \varepsilon_{QCD} N^{\mathrm{QCD}}$, one can finally obtain:

$$N_{\ell-t} = \frac{1-\varepsilon_{sig}}{\varepsilon_{sig}} N_t^{\mathrm{t\bar{t}}} + \frac{1-\varepsilon_{sig}}{\varepsilon_{sig}} N_t^{\mathrm{W}} + \frac{1-\varepsilon_{QCD}}{\varepsilon_{QCD}} N_t^{\mathrm{QCD}} \tag{12}$$

Thus, the task is to minimize:

$$-\log L(N_t^{\mathrm{t\bar{t}}}, N_t^{\mathrm{W}}, N_t^{\mathrm{QCD}}) = -\sum_i n_i^{\mathrm{obs}} \log \mu_i + \mu_i - N_{\ell-t}^{\mathrm{obs}} \log N_{\ell-t} + N_{\ell-t} \tag{13}$$

where, as already mentioned, μ_i and $N_{\ell-t}$ are functions of $N_t^{\mathrm{t\bar{t}}}$, N_t^{W} and N_t^{QCD} as given, respectively, by Eq. 10 and 12.

4. Likelihood Cut Method

The number of $t\bar{t}$ events in the preselected sample can also be determined by using the 2^{nd} Matrix Method (see Sect. IV A 2) after applying a cut to the likelihood discriminant distribution. The efficiencies of such topological cut are derived from MC in the case of $t\bar{t}$ and W+jets whereas, in the case of QCD, it is derived from preselected data where the tight lepton ID criteria (electron likelihood or muon isolation, respectively for e+jets or μ+jets) has been reversed.

In analogy to the way the number of $t\bar{t}$ events was extracted for the Likelihood Fit Method described above, one can apply a similar likelihood fit to the selected events in the Likelihood Cut Method. This allows to properly propagate statistical uncertainties taking into account correlations. The likelihood function is then defined as:

$$L(N_{sel}^{\mathrm{t\bar{t}}}, N_{sel}^{\mathrm{W}}, N_{sel}^{\mathrm{QCD}}) = P(N_{\ell-t-sel}^{\mathrm{obs}}, N_{\ell-t-sel}) \cdot P(N_{t-sel}^{\mathrm{obs}}, N_{t-sel}) \cdot P(N_{sel}^{\mathrm{obs}}, N_{sel}) \tag{14}$$

with $N_{\ell-t-sel}$, N_{t-sel} and N_{sel} given by:

$$\begin{aligned} N_{\ell-t-sel} &= \frac{1-\varepsilon_{sig}-\varepsilon_{sig}\varepsilon_{sel}^{\mathrm{t\bar{t}}}}{\varepsilon_{sig}\varepsilon_{sel}^{\mathrm{t\bar{t}}}} N_{sel}^{\mathrm{t\bar{t}}} + \frac{1-\varepsilon_{sig}-\varepsilon_{sig}\varepsilon_{sel}^{\mathrm{W}}}{\varepsilon_{sig}\varepsilon_{sel}^{\mathrm{W}}} N_{sel}^{\mathrm{W}} + \frac{1-\varepsilon_{QCD}-\varepsilon_{QCD}\varepsilon_{sel}^{QCD}}{\varepsilon_{QCD}\varepsilon_{sel}^{QCD}} N_{sel}^{\mathrm{QCD}} \\ N_{t-sel} &= \frac{1-\varepsilon_{sel}^{\mathrm{t\bar{t}}}}{\varepsilon_{sel}^{\mathrm{t\bar{t}}}} N_{sel}^{\mathrm{t\bar{t}}} + \frac{1-\varepsilon_{sel}^{\mathrm{W}}}{\varepsilon_{sel}^{\mathrm{W}}} N_{sel}^{\mathrm{W}} + \frac{1-\varepsilon_{sel}^{QCD}}{\varepsilon_{sel}^{QCD}} N_{sel}^{\mathrm{QCD}} \\ N_{sel} &= N_{sel}^{\mathrm{t\bar{t}}} + N_{sel}^{\mathrm{W}} + N_{sel}^{\mathrm{QCD}} \end{aligned} \tag{15}$$

as obtained from Eq. 2 and 3.

VI. SYSTEMATIC UNCERTAINTIES

The systematic uncertainties can be subdivided into uncertainties on the preselection efficiency and uncertainties on the estimated number of selected $t\bar{t}$ signal events. In the section below the uncertainties relevant for the lepton+jets channels are listed.

A. Uncertainties on the Preselection Efficiency

Uncertainty associated with the Primary Vertex Selection Efficiency The systematic uncertainty on the primary vertex selection efficiency is taken from the statistical uncertainty in the preselected sample without applying the σ_{dca} and the Δz cuts on the muon.

Uncertainty associated with the EM Reconstruction Efficiency This systematic uncertainty comes from the dependence of the measured reconstruction efficiency in $Z \to ee$ data events on the Z mass window variation.

Uncertainty associated with the EM Identification Efficiency This systematic uncertainty entirely comes from the systematic uncertainty on the measurement from $Z \to ee$ data.

Uncertainty associated with the EM Tracking Efficiency This systematic uncertainty entirely comes from the systematic uncertainty on the measurement from $Z \to ee$ data.

Uncertainty associated with the EM Likelihood Efficiency This systematic uncertainty entirely comes from the systematic uncertainty on the measurement from $Z \to ee$ data.

Uncertainty associated with the Muon Tracking Efficiency The systematic uncertainty on the tracking efficiency is obtained by varying the correction factor by $\pm 1\sigma$.

Uncertainty associated with the μ Identification Efficiency The systematic uncertainty on the μ identification efficiency is obtained by varying the correction factor by $\pm 1\sigma$.

Uncertainty associated with the μ Isolation Efficiency The systematic uncertainty on the μ isolation efficiency is obtained by varying the linear fit to the efficiency from the $Z \to \mu\mu$ data events as a function of jet multiplicity by $\pm 1\sigma$.

Uncertainty associated with the μ σ_{dca} Efficiency The systematic uncertainty on the μ σ_{dca} efficiency is obtained from the statistical uncertainty on the the correction factor.

Uncertainty associated with the Trigger Efficiency The systematic uncertainty on the trigger efficiency is obtained by varying the efficiency parametrisation by $\pm 1\sigma$.

Uncertainty associated with the Δz Cut Efficiency Conservatively, the systematic uncertainty is defined as the inefficiency for the Δz cut efficiency.

Uncertainty associated with the χ^2 Cut Efficiency Conservatively, the systematic uncertainty is defined as the inefficiency for the χ^2 cut efficiency.

Uncertainty associated with the Uncertainty on the Jet Energy Resolution The jet energy resolution (JER) is varied by $\pm 1\sigma$ and the preselection efficiency is redetermined.

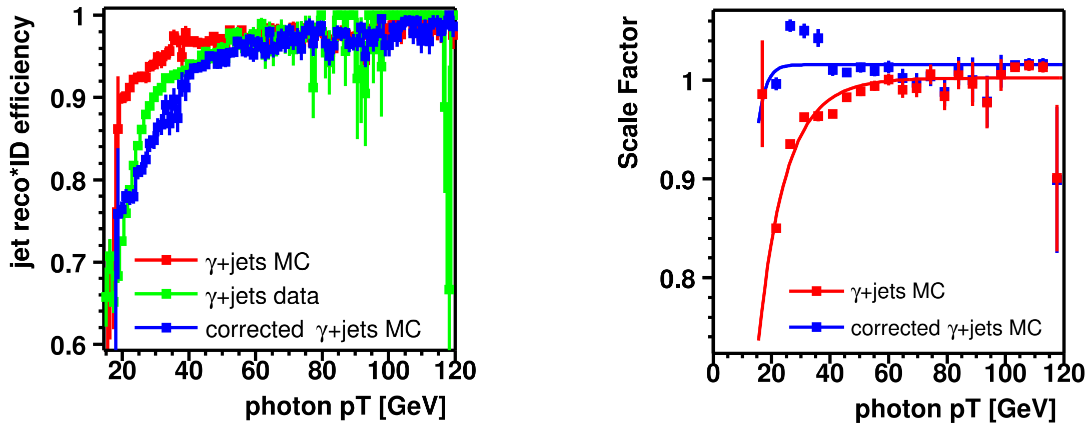

FIG. 4: Left: Jet reconstruction $\times$ jet identification efficiency as a function of the photon pT for $\gamma + jets$ events measured in data, in MC and in the corrected MC, where the difference in efficiency is simulated in MC. In the $\gamma + jets$ data we don't have the same fraction of events with the limited L1 trigger coverage, so the corrected MC does not exactly reproduce the data. Right: Correction factor between MC and data before and after correcting the MC.

Uncertainty associated with the JES Uncertainty The effect of the jet energy scale uncertainty on the preselection efficiency uncertainty is obtained by varying the JES by $\pm 1\sigma$, where, conservatively,

$$\sigma = \sqrt{\sigma^2_{stat,data} + \sigma^2_{syst,data} + \sigma^2_{stat,MC} + \sigma^2_{stat,MC}} \tag{16}$$

Uncertainty associated with the data/MC difference on the Jet Reconstruction and Identification Efficiency A correction factor is defined by dividing the combined jet reconstruction $\times$ jet identification efficiency measured in data by the efficiency measured in MC. The efficiencies and the scale factor is measured using γ+jet events and shown in Fig. 4. 37% of the data set used for these analyses were recorded with a limited L1 trigger coverage ($|\eta_{det}| < 2.4$). Triggering is not simulated in the MC. The effect of this limited L1 trigger coverage on the JID is simulated by rejecting all jets with $|\eta_{det}| > 2.4$ for 37% of the events. For the remaining fraction of events (63%), jets are removed randomly according to the estimated data-to-MC scale factor in order to reproduce the jet ID times reconstruction efficiency in data. Effects on jet counting, $\not\!\!E_T$, etc, are properly taken into account. As it can be observed in Fig. 4, the simulated inefficiency through this simplified procedure is slightly worse than that observed in data. Therefore, the quoted systematic uncertainty is believed to be conservative. Eventually, once proper tuning of data and MC are achieved, this will become a default correction and not a systematic uncertainty anymore.

The net effect on the $t\bar{t}$ preselection efficiency is used as a systematic uncertainty.

Uncertainty associated with the Soft Lepton Veto The systematic uncertainty associated with the soft lepton veto is described in section VI E of [3].

Uncertainty associated with the Top Mass The effect of the uncertainty on the top mass on the preselection efficiency uncertainty is obtained from the difference in preselection efficiencies for a top mass $m_t = 170$ GeV, respectively $m_t = 180$ GeV, instead of $m_t = 175$ GeV. The uncertainty is quantified by determining the preselection efficiency for MC samples with a top mass of $m_t = 150$ GeV, $m_t = 160$ GeV, $m_t = 170$ GeV, $m_t = 175$ GeV, $m_t = 180$ GeV, $m_t = 190$ GeV, and $m_t = 200$ GeV. A fit is performed to the efficiency versus top mass and the fit values at $m_t = 170$ GeV and $m_t = 180$ GeV are used as a systematic uncertainty.

Uncertainty due to lack of Multiple Parton Interactions An uncertainty due to lack of multiple parton interactions ("tune A uncertainty") was estimated by calculating the preselection efficiency on a $t\bar{t}$ Monte Carlo sample generated with simulation of multiple parton interactions. The results were then compared with those obtained using the default MC samples and the difference quoted as a systematic uncertainty.

B. Uncertainties on the Extracted Number of $t\bar{t}$ Events

Uncertainty associated with the Uncertainty on the Jet Energy Resolution The JER is varied by $\pm 1\sigma$ and the estimated signal is redetermined.

Uncertainty associated with the JES Uncertainty The effect of the jet energy scale uncertainty on the signal uncertainty is obtained by varying the JES by $\pm 1\sigma$, where σ is defined in Equation (16).

Uncertainty associated with the data/MC difference on the Jet Reconstruction and Jet Identification Efficiency The scale factor is applied to the $t\bar{t}$ and $W + jets$ MC and the templates are redetermined. The effect on the fitted number of $t\bar{t}$ events is used as a systematic uncertainty.

Uncertainty associated with the Top Mass The effect of the uncertainty on the top mass on the extracted uncertainty is obtained from the difference in preselection efficiencies for a top mass $m_t = 170$ GeV, respectively $m_t = 180$ GeV, instead of $m_t = 175$ GeV.

Uncertainty due to lack of Multiple Parton Interactions An uncertainty due to multiple parton interactions ("tune A uncertainty") was estimated by calculating the selection efficiency on a Monte Carlo sample generated with simulation of multiple parton interactions. The results were then compared with those obtained using the default MC samples.

Uncertainty associated with the Statistics used to build the Likelihood Templates The likelihood fit to the data is repeated 1000 times where in each bin of the likelihood discriminant one allows for Poisson fluktuations. The RMS of the distribution of the fitted number of $t\bar{t}$ events is an estimate for the systematic uncertainty.

Uncertainty associated with ε_{sig} The uncertainty on ε_{sig} is propagated to the uncertainty on the signal.

Uncertainty associated with ε_{QCD} The uncertainty on ε_{QCD} is propagated to the uncertainty on the signal.

Uncertainty associated with the Trigger Efficiency The uncertainty on the number of selected $t\bar{t}$ signal events from the systematic uncertainty on the trigger efficiency is obtained by varying the efficiency parametrisation $\pm 1\Delta$.

VII. MUON+JETS CHANNEL

A. Event Preselection

The analyzed data sample is triggered with the signal triggers (*MU_JT20_L2L0*) and (*MU_JT25_L2L0*) as discussed in Section II A. The preselection is a selection of a $W \to \mu\nu$ and jets, where the muon is well reconstructed as described in Section III C. To reject QCD events where the muon originates from a heavy flavor decay the muon is required to have a high transverse momentum, to come from the primary vertex and to be isolated from hadronic activity. The neutrino is indirectly identified by requiring a high transverse missing energy, and in order to reject background from QCD the missing energy has to be separated from the leading jet and the muon in ϕ. Only events with four or more jets are considered. Events with a soft muon tag or a second muon with high transverse momentum are analysed in [9] and in [8] , respectively and are explicitely vetoed here.

All cuts used to define the preselected sample and the corresponding efficiencies are listed in Table VIII od section VII E. The measurement of the reconstruction efficiencies and the survival efficiencies for the preselection cuts are discussed in the following subsections. The preselection efficiency $\varepsilon_{preselection}$ is factorized in the following way:

1. *Muon Acceptance, Reconstruction Efficiency and Cosmic Veto*

To measure the efficiency of the muon reconstruction in data, $Z^0 \to \mu\mu$ events are used which can be identified by one fully reconstructed muon (control muon) and a second muon only reconstructed in the tracking system (test muon). To ensure that the reconstructed track belongs to a muon from a Z^0 decay, the invariant mass of the control and the test muon has to be compatible with the Z^0 mass m_{Z^0}. The existing trigger bias is removed by selecting events recorded with a single muon trigger and matching the control muon with the muon trigger object. The rate at which the test muon is matched to a reconstructed muon is a measure of the local muon reconstruction and track-muon matching efficiency.

The same measurement is done in MC with $Z^0 \to \mu\mu$ events and the differential efficiencies as a function of detector η, ϕ and p_T are compared. The ratio of the measured efficiencies in data and MC is found to depend only on the absolute of detector η and is parameterized using a polinomial of third order.

The efficiency to reconstruct a medium muon with $N_{seg} = 3$ within the timing window in $t\bar{t} \to \mu + jets$ events is derived from the simulation where the detector η dependent correction factor is taken into account.

2. *Muon Tracking Efficiency*

The efficiency to reconstruct a track to the corresponding muon in $t\bar{t} \to \mu{+}jets$ events is measured in the simulation. The determination of a scale factor that accounts for the difference in tracking efficiency between data and MC is described in [3].

3. *Efficiency of the Second Muon Veto (SMV)*

The efficiency of the cuts used to reject events with a second muon with high transverse momentum that would have been selected by the dimuon analysis (see [8]) in $t\bar{t} \to \mu + jets$ events is measured in the simulation.

4. *Efficiency of the Soft Muon Veto (SLV)*

The efficiency of the cuts used to reject events with at least one muon-tagged jet in $t\bar{t} \to \mu{+}jets$ events is measured in the simulation. A correction factor for the reconstruction efficiency for muons with a low transverse momentum is determined in the same way as described in Section VII A 1, but using the process $J/\Psi \to \mu\mu$.

5. *Efficiency of the Muon Separation from Jets*

The efficiency to select events with $\Delta R(\mu, jet) > 0.5$ in $t\bar{t} \to \mu + jets$ events is measured in the simulation.

6. *Efficiency of the Muon Transverse Momentum Cut*

The efficiency to select events with $p_{T\mu} > 20$ GeV in $t\bar{t} \to \mu + jets$ events is measured in the simulation.

7. *Efficiency of the Muon Distance to Closest Approach Cut*

The cut on the DCA significance of the track matched to a muon $|dca|/\sigma_{dca} < 3$ is used to reject muons from semileptonic decays, while most of the muons from W decays are kept. The efficiency for the $|dca|/\sigma_{dca}$ cut is measured both in $Z \to \mu\mu$ data and simulation and found to be in very good agreement. The ratio of the measured efficiencies in data and MC is used as a correction factor, shown in Fig. 5. The efficiency to select events with $|dca|/\sigma_{dca} < 3$ in $t\bar{t} \to \mu + jets$ events is measured in the simulation taking into account the correction factor.

8. *Efficiency of the $|\Delta z(\mu,primary\ vertex)|$ Cut*

The cut $|\Delta z(\mu, PV)| < 1$ cm is used to reject both muons from semileptonic decays and cosmic muons, while most of the muons from W decays are kept. The cut efficiency is measured in $(Z \to \mu\mu) + 1jet$ data events.

9. *Efficiency of the Track χ^2 Cut*

A quality cut on the track associated to the muon is applied in order to suppress random matching of muons and tracks. The cut efficiency for $\chi^2 < 4$ is measured in $(Z \to \mu\mu) + 1jet$ data events.

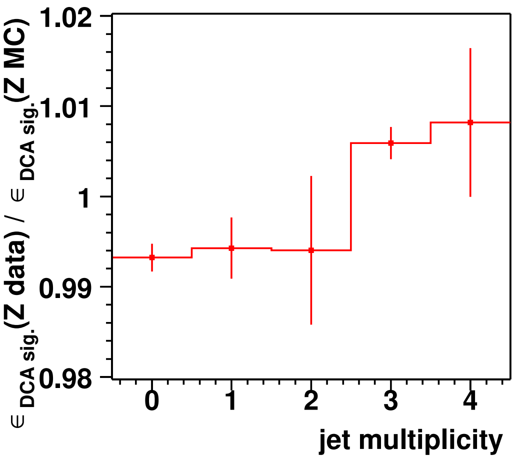

FIG. 5: Ratio of the measured efficiencies for $|dca|/\sigma_{dca} < 3$ in $Z \to \mu\mu$ data and MC as a function of exclusive jet multiplicity. The last bin is inclusive.

10. Tight Muon Isolation Efficiency

The tight muon isolation as defined in Section III C is dependent on the topology of the event and therefore has to be measured in the $t\bar{t} \to \mu + jets$ simulation. Differences between the data and the simulation are quantified using $Z \to \mu\mu$ events. The isolation efficiency is measured as a function of jet multiplicity both in simulated and in data events. The shape of this distribution in the simulation can be described by a polinomial of first order. The statistics for $(Z \to \mu\mu) + 4jets$ events in data is very low and the efficiency cannot be determined with the desired statistical precision. A polinomial of first order is fitted to the distribution and extrapolated to the $(Z \to \mu\mu) + 4jets$ bin. The ratio of this efficiency measurement in data and the measurement in the simulation is used as a correction factor and shown in Fig. 6.

11. Muon Trigger Efficiencies

The determination of the trigger efficiency per medium muon with $N_{seg} = 3$ and the parametrisation as a function of the relevant variables is described in section III of [3]. The muon trigger efficiency for $t\bar{t} \to \mu + jets$ events is measured in the simulation by folding in the trigger parametrisation per muon and deriving the event trigger efficiency as described in section III of [3].

12. Jet Reconstruction and Identification Efficiency

A correction factor that accounts for possible differences in the reconstruction efficiency for jets has been determined in a sample where events are selected with an electromagnetic object back-to-back to a jet, recorded with a single EM trigger, compared to MC. Instead of applying this “scale factor”, a systematic uncertainty will be assessed.

13. Jet Trigger Efficiencies

The determination of the trigger efficiency per jet and the parametrisation as a function of the relevant variables is described in section III of [3]. The jet trigger efficiency for $t\bar{t} \to \mu + jets$ events is measured in the simulation by folding in the trigger parametrisation per jet and deriving the event trigger efficiency as described in section III of [3].

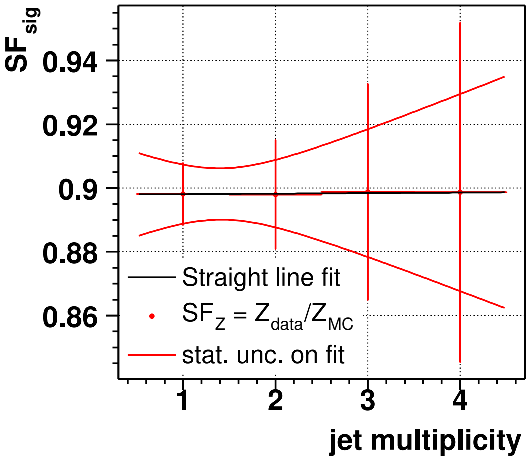

FIG. 6: Ratio of the muon isolation efficiencies in $Z \to \mu\mu$ data and MC as a function of exclusive jet multiplicity. The last bin is inclusive.

14. Efficiencies for the Missing Transverse Energy Cut

A cut on the missing transverse energy is applied in order to reject QCD background. However there are QCD heavy flavor events which survive this cut. The different angular distribution of the missing transverse energy and the muon (or the leading jet) allows to further reject QCD without suffering from a large efficiency loss.

- QCD is found to have the missing transverse energy mainly back-to-back with the leading jet where the missing transverse energy in these events is generated by the mismeasurement of the leading jet energy.
- QCD is also found to have the missing transverse energy either in the direction of the muon or back-to-back to it. This can be explained by $b\bar{b}$ production where one or both of the b-hadrons decay semileptonically and one of the jets is not reconstructed in the calorimeter.

In a grid search the cuts on the missing transverse energy and the $\Delta\phi$ between the missing transverse energy and the muon, respectively the leading jet were varied. A set of cuts was chosen which had the same efficiency for the signal ($t\bar{t} \to \mu + jets$ simulation) as in [2], but a maximum rejection for the background (QCD selection). The optimum set of cuts was found to be:

- $\not{E}_T > 17$ GeV
- $\Delta\phi(\mu,\not{E}_T) > 1.2 - \not{E}_T$[GeV]$*1.2/38$
- $\Delta\phi(\mu,\not{E}_T) < 1.3 + \not{E}_T$[GeV]$*(\pi - 1.3)/24$
- $\Delta\phi$(leading jet,$\not{E}_T$) $< 2.2 + \not{E}_T$[GeV]$*(\pi - 2.2)/26$,

which is illustrated in Fig. 7. The efficiency for this set of cuts is determined with simulated $t\bar{t}$ events.

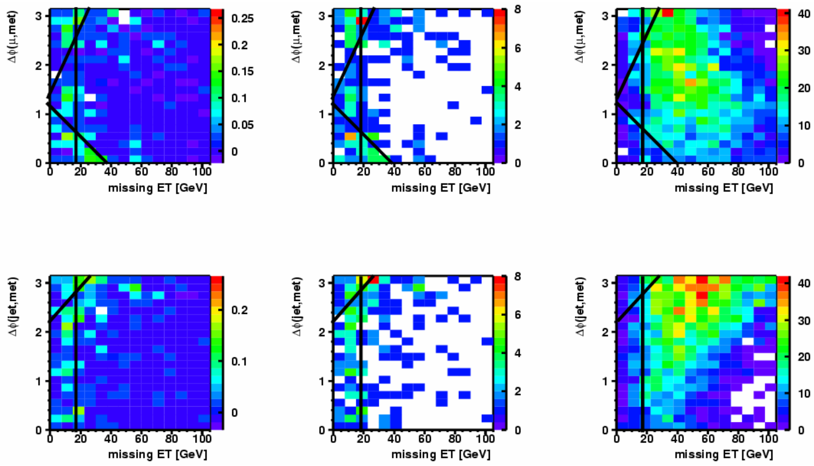

FIG. 7: Top: $\Delta\phi(\mu,\not{E}_T)$ is plotted versus $\not{E}_T$. Bottom: $\Delta\phi$(leading jet,$\not{E}_T$) is plotted versus $\not{E}_T$. Left: QCD in the preselected data sample without applying the $\not{E}_T$ cut. Middle: μQCD without applying the $\not{E}_T$ cut. Right: $t\bar{t}$ simulation without applying the $\not{E}_T$ cut. The black lines represent the cut borders.

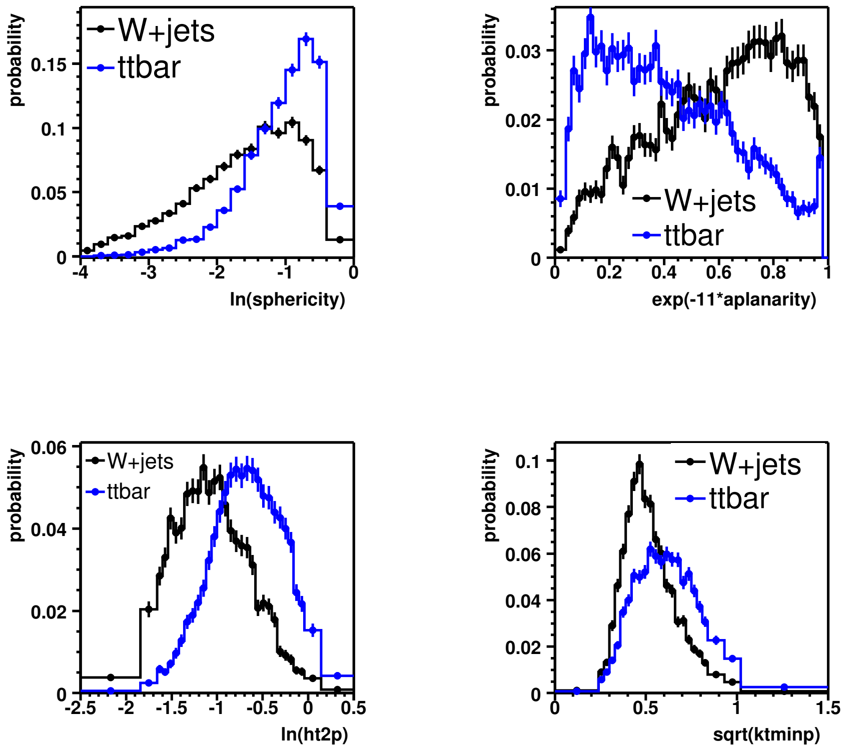

FIG. 8: Distributions of the four topological variables for $t\bar{t}$, $W + jets$ and QCD.

TABLE VII: Correlation factors for the four topological variables for $t\bar{t}$ and $W + jets$.

variables	correlation factor for $t\bar{t}$	correlation factor for $W + jets$
$\mathcal{A}, S$	0.57	0.56
$\mathcal{A}, H'_{T2}$	0.16	0.10
$\mathcal{A}, K'_{Tmin}$	0.24	0.38
S, H'_{T2}	0.09	0.03
S, K'_{Tmin}	0.28	0.40
H'_{T2}, K'_{Tmin}	0.09	0.10

B. The Topological Variables

The normalized distributions of the functions of the four topological variables that were introduced in Section IV B 1 are shown in Fig. 8 for the $t\bar{t}$ signal and the $W + jets$ background. The fit to the logarithm of the ratio of signal over background for each of the four likelihood input variables is shown in Fig.9. The likelihood function is built according to Equation 8 and evaluated for $t\bar{t}$, $W + jets$ and QCD. The resulting normalized likelihood output distributions (templates) are shown in Fig. 10.

The correlations among the four variables are shown in Fig. 11 and Fig. 12 for $t\bar{t}$, respectively for $W + jets$ and the correlation factors are summarized in Table VII. Figures 13- 14 show that the $W + jets$ MC is able to reproduce the shape of the four topological variables as measured in data. The $W + jets$ distribution in data is extracted by applying the matrix method as described in Section IV A 1. Figures 15- 16 show the same comparison between the QCD data sample, referred to as μQCD, and the QCD in the preselected sample. Here the agreement is present by construction since the the two samples are highly correlated. By applying the matrix method bin by bin, in order to get the QCD distribution in the preselected sample, one makes use of the loose preselected sample which is correlated with the μQCD sample.

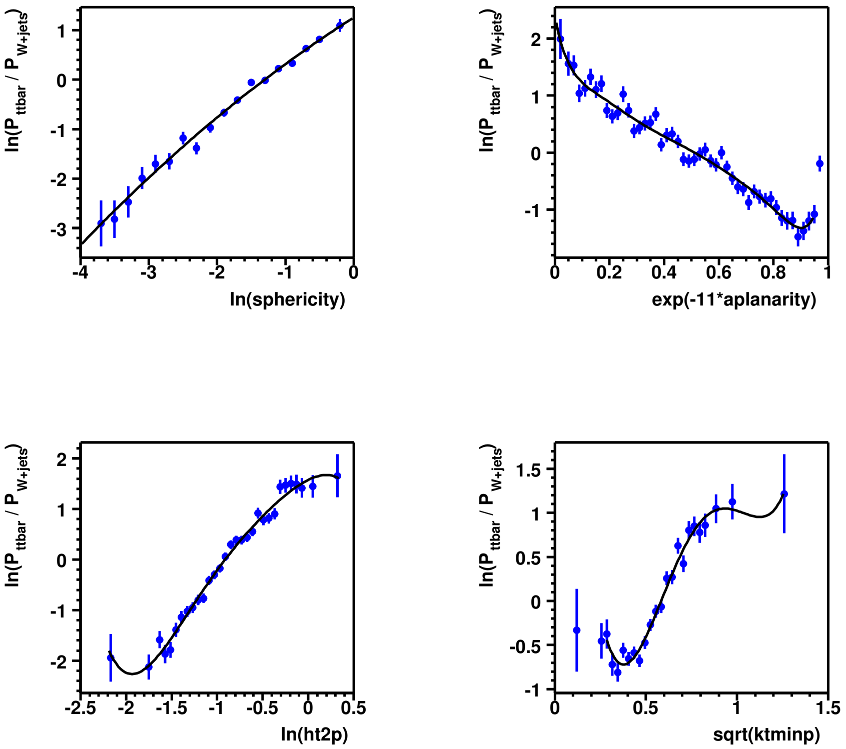

FIG. 9: Distributions of the four topological variables for $t\bar{t}$, $W + jets$ and QCD.

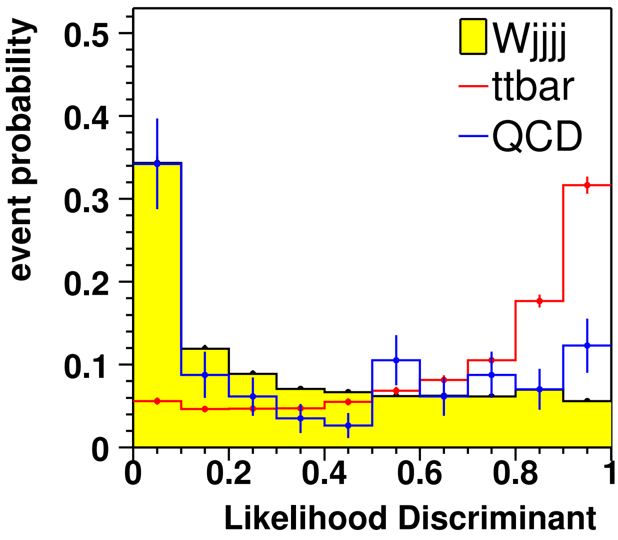

FIG. 10: Likelihood discriminant distributions for $t\bar{t}$, $W + jets$ and QCD.

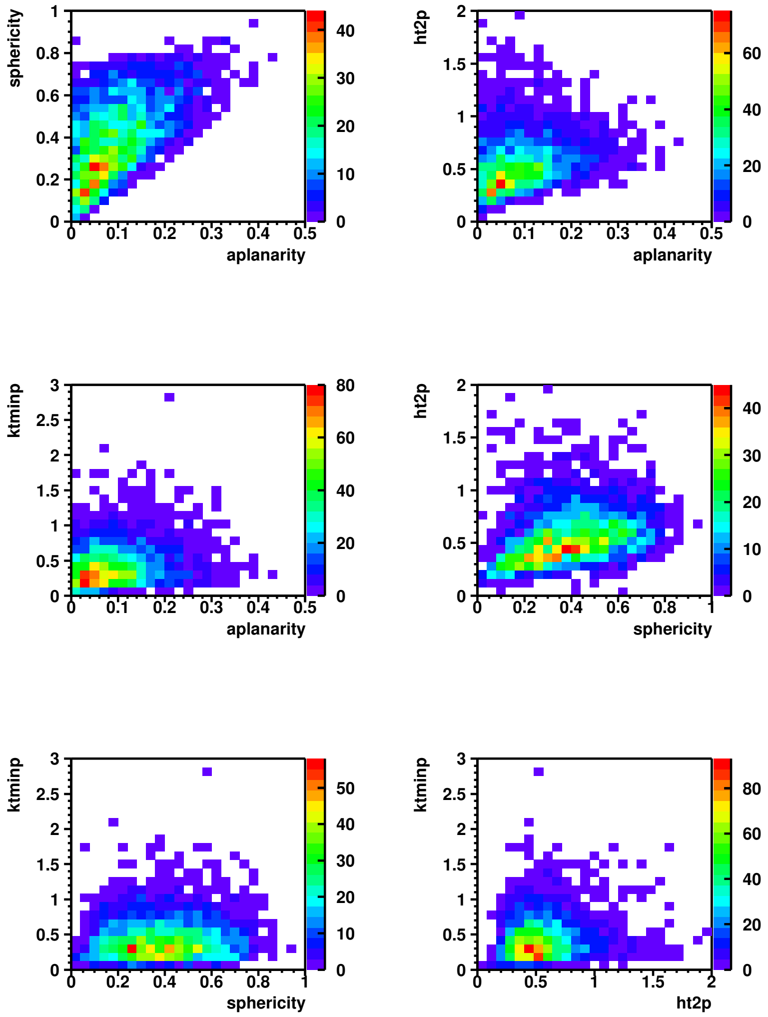

FIG. 11: Correlations among the four topological variables for $t\bar{t}$.

C. Likelihood cut Method

The expected statistical uncertainties for this method are determined using the matrix method (see Section IV A 1). One needs four ingredients to solve the system of two linear equations:

1. the efficiency for signal events to pass the likelihood cut, $\epsilon_{t\bar{t}}$, which is determined from the $t\bar{t}$ simulation.
2. the efficiency for background events to pass the likelihood cut, ϵ_{W+jets}, which is determined from the $W+jets$ simulation.
3. the number of events before applying the likelihood cut, N_{loose}, which is taken from the measurement in data.
4. the number of events which survive the likelihood cut, N_{tight}, which is dependent on the S/B at the preselection level. N_{tight} can be retrieved in two different ways, either by applying the signal and background efficiencies $\epsilon_{t\bar{t}}$ and ϵ_{W+jets}:

$$N_{tight} = N_{loose}\left(\frac{S/B}{S/B+1}\epsilon_{t\bar{t}} + \frac{1}{S/B+1}\epsilon_{W+jets}\right) \tag{17}$$

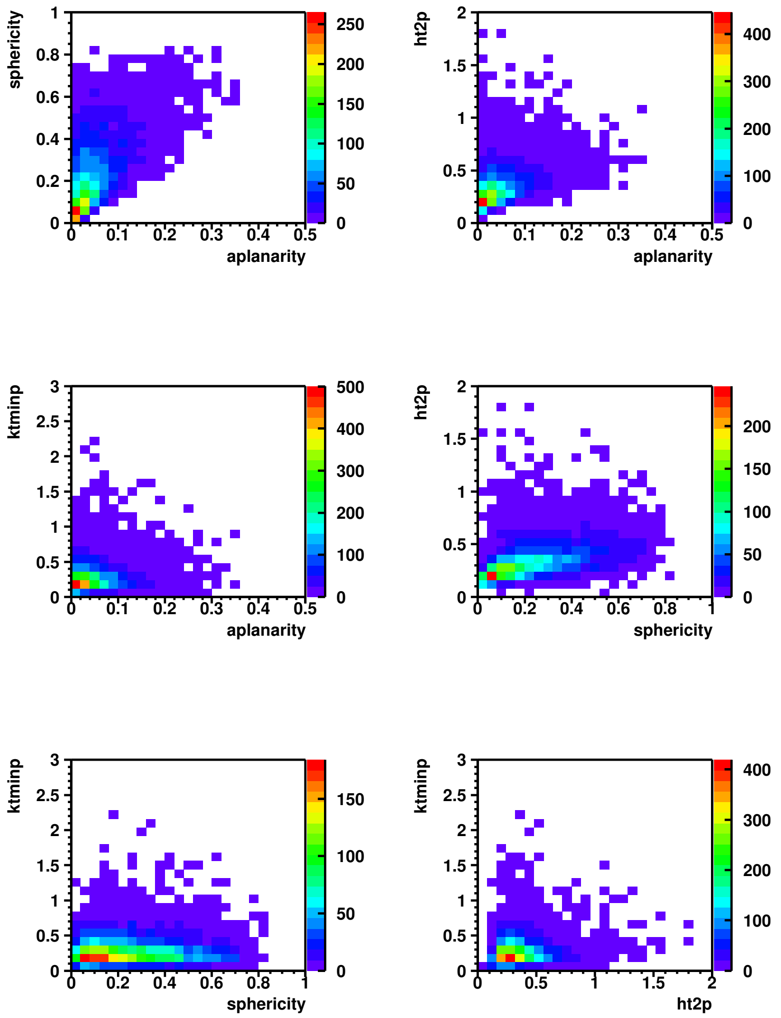

FIG. 12: Correlations among the four topological variables for $W + jets$.

The other possibility is to do ensemble testing with ensembles with the size of N_{loose}. For a given S/B, N_{tight} is the number of events in an ensemble which survive the likelihood cut. For each ensemble one gets an error using the matrix method. The mean of the distribution of these errors represents the expected statistical uncertainty. The expected uncertainties from the two methods are in good argreement as shown in Figure 17.

The expected systematical uncertainties are determined for the uncertainty on the JES and shown in Fig. 18. The expected total uncertainties by adding the expected statistical and systematical uncertainties in quadrature are shown in Fig. 19.

D. Likelihood fit Method

A standard likelihood fit using Poisson statistics is used to fit the number of $t\bar{t}$ events $f_{t\bar{t}}$ in the preselected data sample. To study the uncertainties ensemble testing is done with 1000 ensembles for different ratios of S/B, where one ensemble has the size of the number of events that survive the preselection. The fitted $f_{t\bar{t}}^{fitted}$ values are depicted in Fig. 20 and the corresponding errors are presented in Fig. 21. On average, there are only ~ 30 statistically independent ensembles possible due to the limited MC statistics. The mean of the fitted $f_{t\bar{t}}^{fitted}$ is compatible with the input $f_{t\bar{t}}^{true}$,

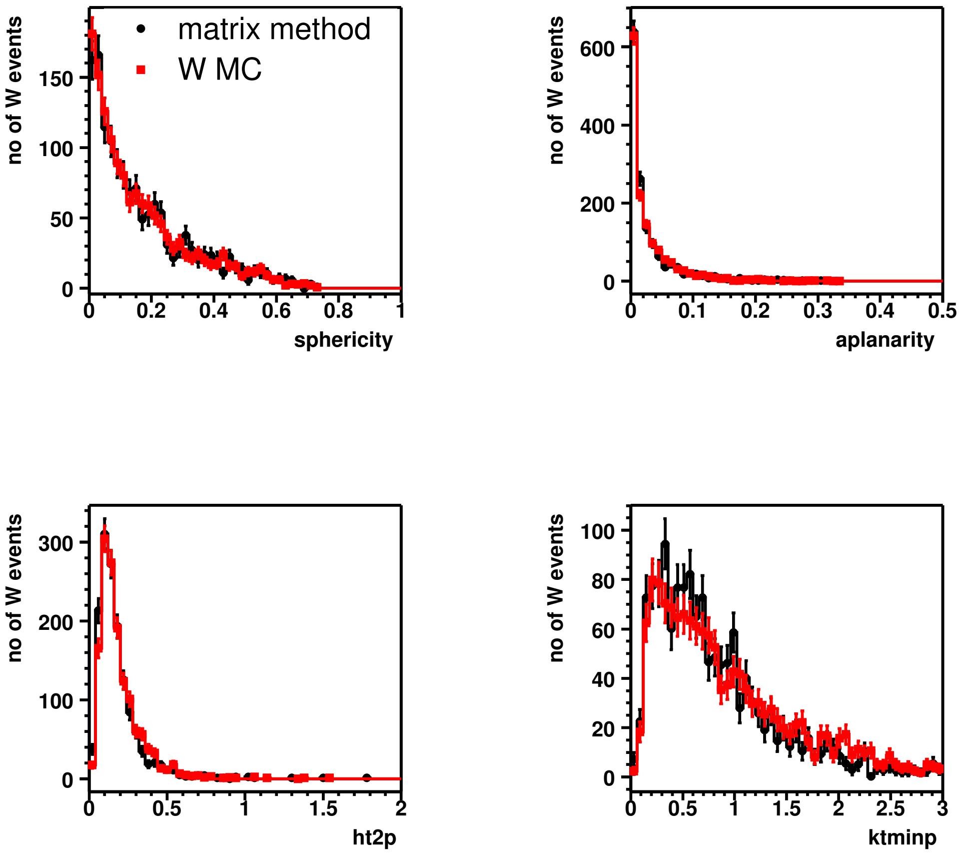

FIG. 13: Distributions of the four topological variables for $W + 2jets$ in MC and in the preselected sample.

given that the error on the mean is defined as:

$$\frac{RMS(f_{t\bar{t}}^{fitted})}{\sqrt{\text{number of statistically independent ensembles}}} \tag{18}$$

The expected statistical uncertainties are represented by the mean of the errors from the fit which is shown in the top left plot of Fig. 22. In the same figure on the top right the calibration curve is shown and on the bottom left the expected systematical uncertainties are shown.

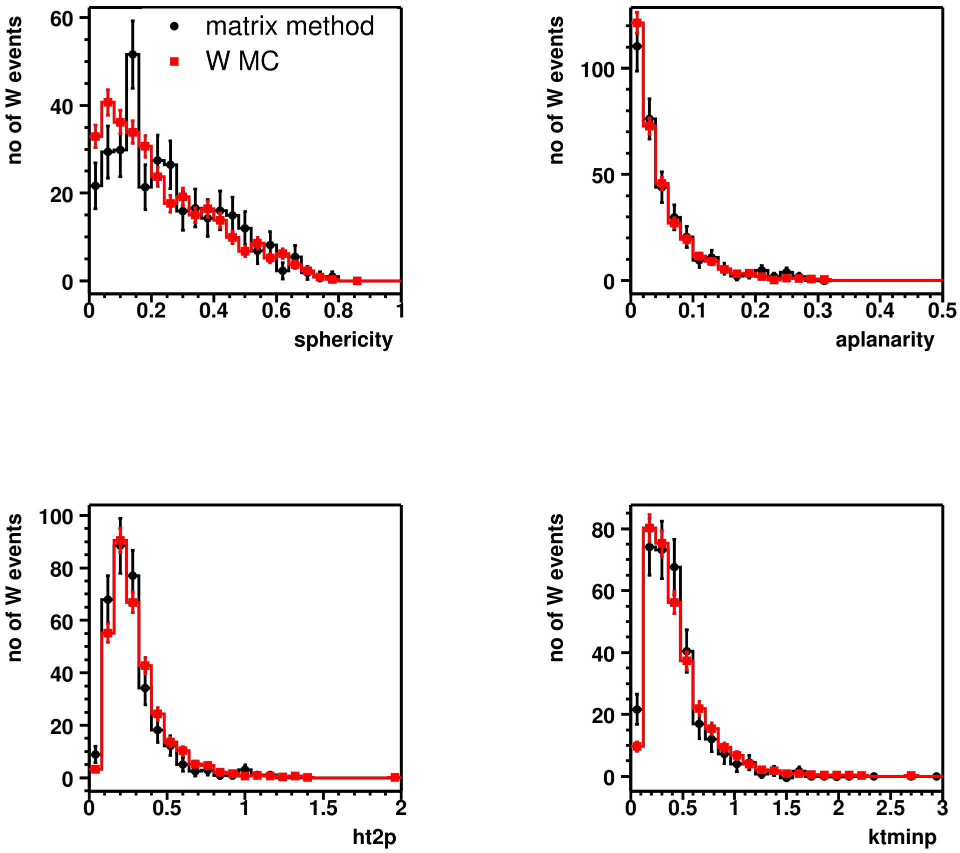

FIG. 14: Distributions of the four topological variables for $W + 3jets$ in MC and in the preselected sample.

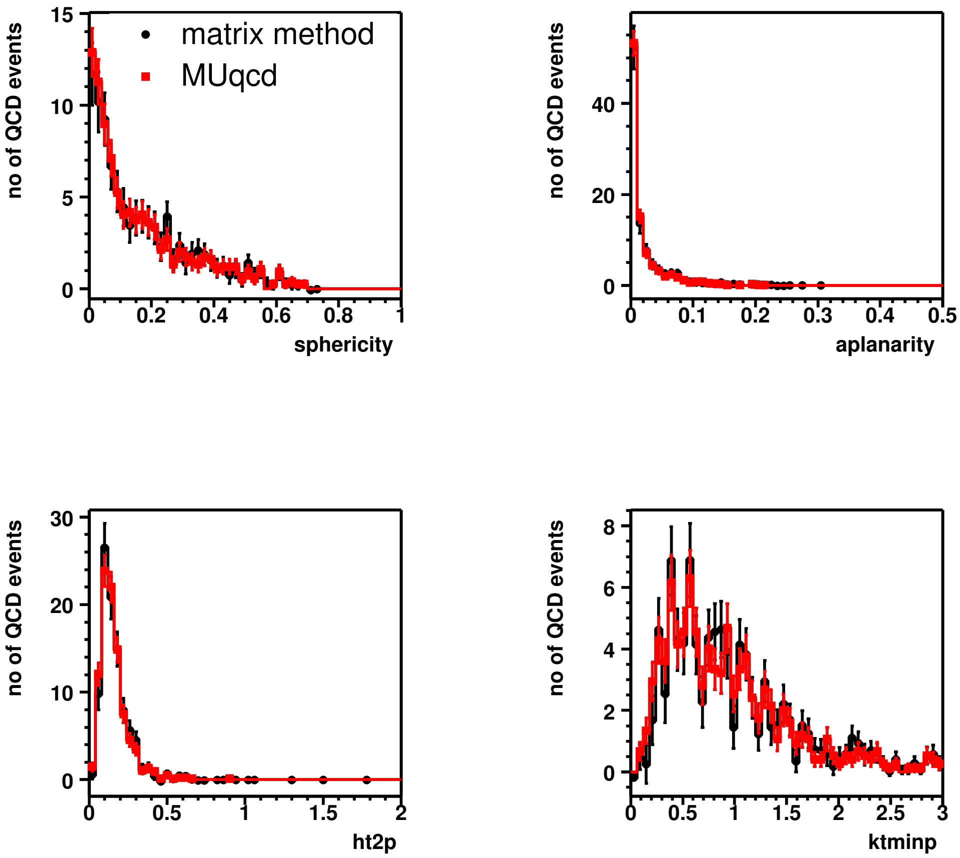

FIG. 15: Distributions of the four topological variables for $QCD + 2jets$ in the μQCD data sample and in the preselected sample.

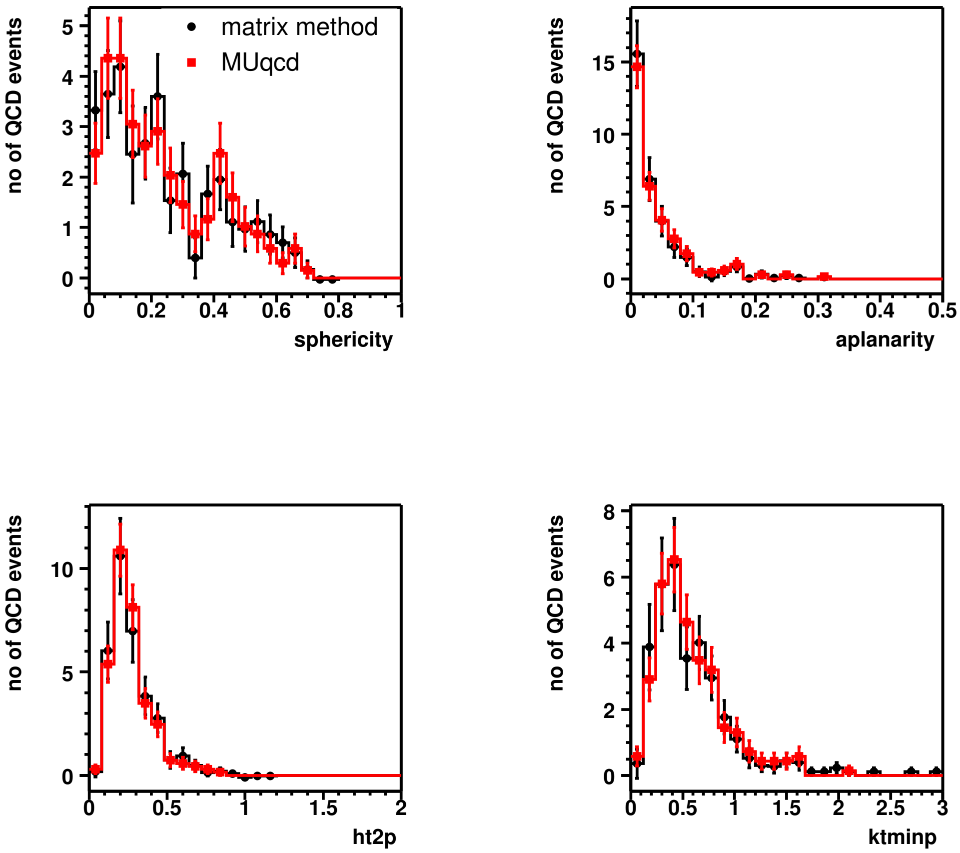

FIG. 16: Distributions of the four topological variables for $QCD+3jets$ in the μQCD data sample and in the preselected sample.

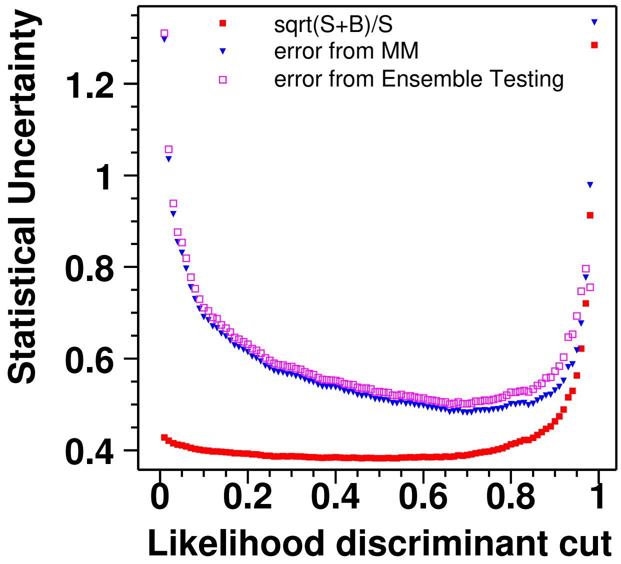

FIG. 17: Expected relative statistical uncertainty for the Likelihood cut Method as a function of the likelihood discriminant cut for $S/B = 1/3$. The uncertainties derived from the ensemble testing are compared with the uncertainties derived with Equation 17. For comparison also the uncertainties are shown when correlations between signal and background are not accounted for, $(\sqrt{(}S+B)/S)$.

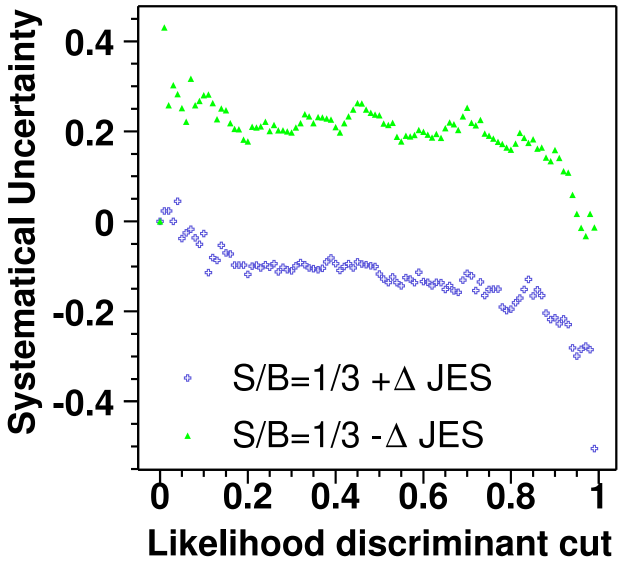

FIG. 18: Expected relative systematical uncertainty for the Likelihood cut Method as a function of the likelihood discriminant cut for $S/B = 1/3$.

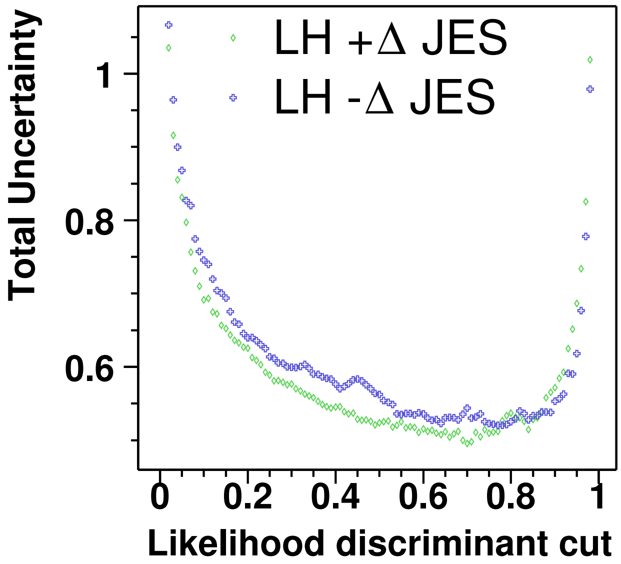

FIG. 19: Expected relative total uncertainty for the Likelihood cut Method as a function of the likelihood discriminant cut for $S/B = 1/3$.

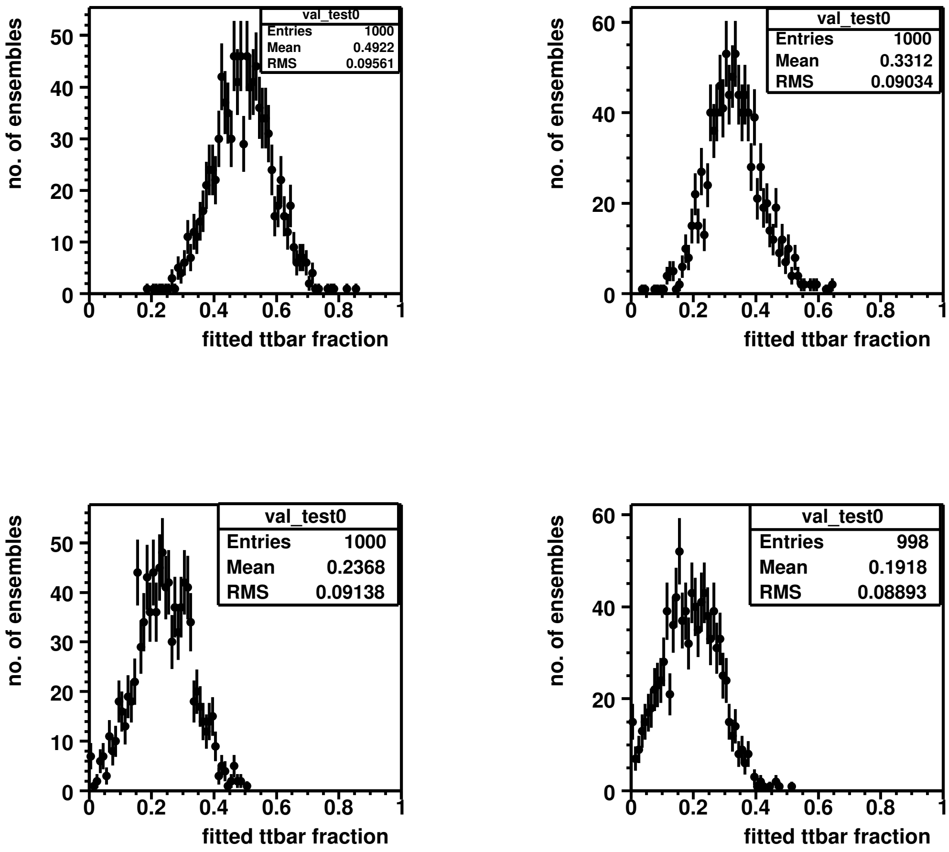

FIG. 20: Ensemble testing: Fitted $t\bar{t}$ fraction $f_{t\bar{t}}$ for 1000 ensembles for four different input $f_{t\bar{t}}$ (which can be also expressed in terms of S/B): top left: $S/B = 1/1,\ f_{t\bar{t}} = 0.5$, top right: $S/B = 1/2,\ f_{t\bar{t}} = 0.33$, bottom left: $S/B = 1/3,\ f_{t\bar{t}} = 0.25$, bottom right: $S/B = 1/4,\ f_{t\bar{t}} = 0.2$.

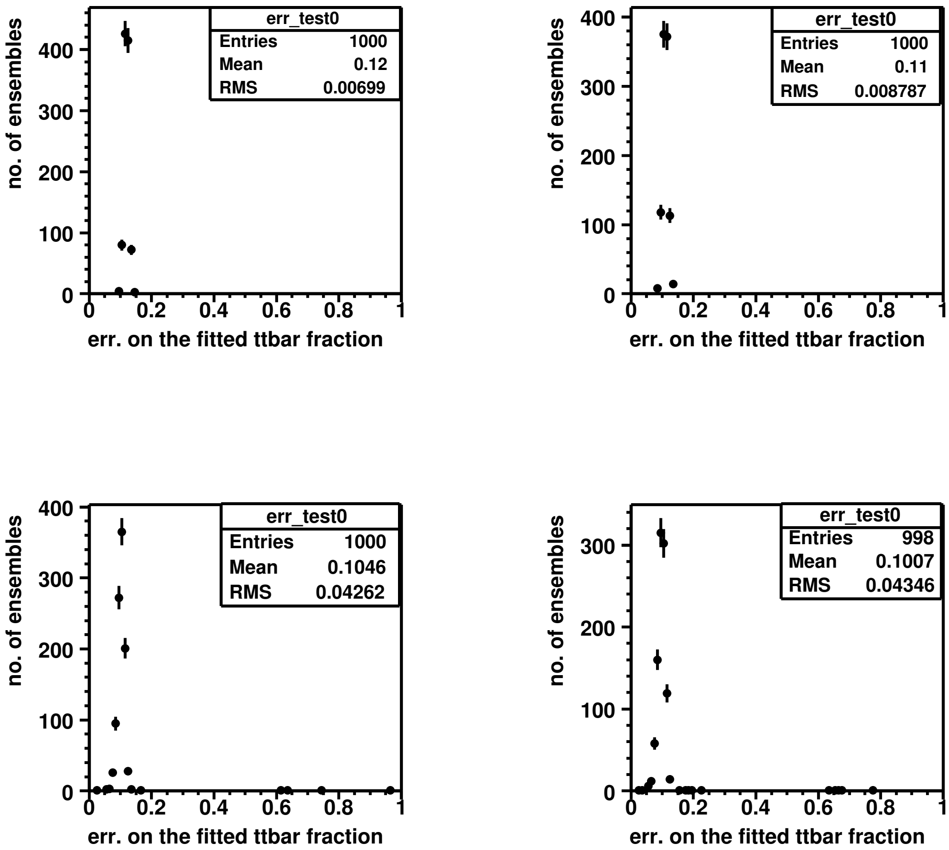

FIG. 21: Ensemble testing: Fit error on the fitted $t\bar{t}$ fraction $f_{t\bar{t}}$ for 1000 ensembles for four different input $f_{t\bar{t}}$ (which can be also expressed in terms of S/B): top left: $S/B = 1/1$, $f_{t\bar{t}} = 0.5$, top right: $S/B = 1/2$, $f_{t\bar{t}} = 0.33$, bottom left: $S/B = 1/3$, $f_{t\bar{t}} = 0.25$, bottom right: $S/B = 1/4$, $f_{t\bar{t}} = 0.2$.

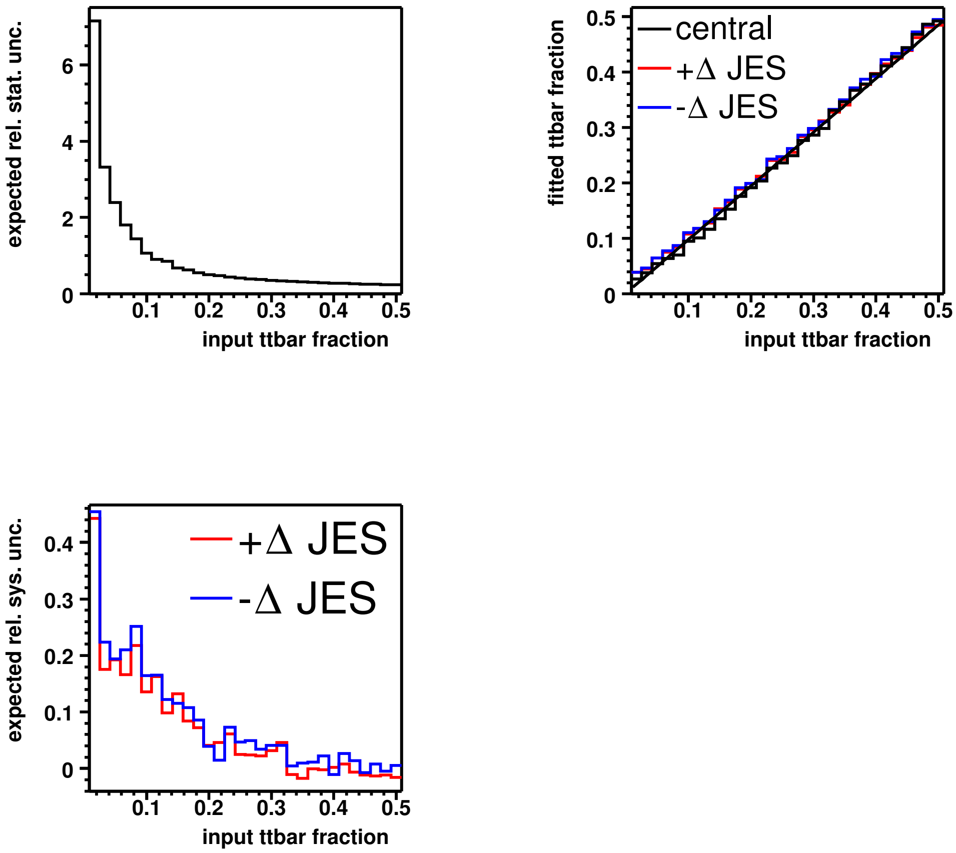

FIG. 22: Top left: Expected relative statistcal uncertainty as a function of the $t\bar{t}$ fraction, top right: calibration curve, $f_{t\bar{t}}^{fitted}$ versus $f_{t\bar{t}}^{true}$, bottom left: expected relative systematical uncertainty as a function of the $t\bar{t}$ fraction.

TABLE VIII: Summary of the $t\bar{t} \to \mu + jets$ event preselection efficiencies and the corresponding correction factors from MC to data ($SF_{MC-to-data}$), if aplicable.

cut or $SF_{MC-to-data}$	exclusive cut efficiency	cumulative cut efficiency
ε_{PV}	94.0±2.%	94.0±2.%
$\varepsilon_{muon}^{id\&acc\&cos}$	36.8±0.2%	34.6±0.7%
$SF_{muon}^{id\&acc\&cos}$	1.01±0.016	34.9±0.7%
$\varepsilon_{muon}^{track}$	96.7±0.2%	33.8±0.7%
SF_{muon}^{track}	0.96±0.02	32.4±0.7%
$\varepsilon_{muon}^{\sigma_{dca}}$	73.9±0.3%	24.0±0.6%
$SF_{muon}^{\sigma_{dca}}$	1.01±0.01	24.2±0.6%
$\varepsilon_{muon}^{p_T\ cut}$	73.2±0.4%	17.7±0.5%
$\varepsilon_{muon}^{\Delta R}$	81.1±0.4%	14.4±0.4%
ε_{muon}^{iso}	90.7±0.4%	13.0±0.4%
SF_{muon}^{iso}	0.899±0.031	11.7±0.5%
$\varepsilon_{\not{E}_T}$	90.9±0.4%	10.7±0.5%
ε_{χ^2}	98.3±0.3%	10.4±0.5%
$\varepsilon_{\Delta z}$	99.7±0.1%	10.4±0.5%
$\varepsilon_{jetmult}$	59.1±0.6%	6.2±0.3%
ε_{muon}^{SMV}	100±0%	6.2±0.3%
ε_{muon}^{SLV}	81.7±0.7%	5.0±0.2%
SF_{muon}^{SLV}	0.976±0.06	4.9±0.4%
$\varepsilon^{trigger}$	94.3±0.4%	4.6±0.4%
ε^{total}	-	4.6±0.4%

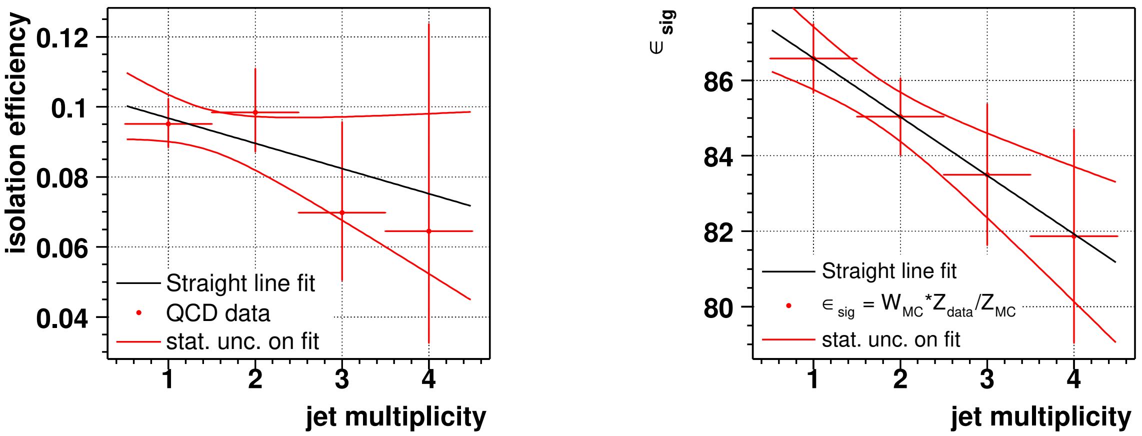

FIG. 23: ε_{QCD} and ε_{sig} as a function of jet multiplicity.

E. $t\bar{t}$ Event Efficiencies

The preselection efficiencies are summarized in Table VIII. The efficiencies are with respect to $t\bar{t} \to \ell + jets$ in order to account for possible contributions to the signal from $t\bar{t} \to \tau + jets$.

The numbers of observed events in the loose (N_ℓ) and the tight (N_t) sample are listed in Table IX, together with the values of ε_{QCD} and ε_{sig} and the number of estimated $W + jets$ background plus $t\bar{t}$ signal events and QCD background events for each inclusive jet multiplicity bin. ε_{QCD} and ε_{sig} as a function of jet multiplicity are shown in Fig. 23.

TABLE IX: Number of observed events in the loose and tight data sample, the efficiency for the tight muon isolation cut and the estimated number of $W + jets$ and QCD background events for each inclusive jet multiplicity.

	N_l	N_t	ε_{sig}	ε_{QCD}	$\varepsilon_{sig} \cdot N^{\mathrm{W+t\bar{t}}}$	$\varepsilon_{QCD} \cdot N^{\mathrm{QCD}}$
≥ 1 jet	11567	7279	86.6 ± 0.8%	9.7 ± 0.7%	6933.6 ± 88.6	345.4 ± 17.8
≥ 2 jet	4310	2466	85.0 ± 0.7%	9.0 ± 0.8%	2324.2 ± 51.1	141.8 ± 7.8
≥ 3 jet	1095	542	83.5 ± 1.1%	8.2 ± 1.5%	501.5 ± 24.0	40.5 ± 3.7
≥ 4 jet	207	100	81.9 ± 1.8%	7.5 ± 2.3%	93.0 ± 10.3	7.0 ± 1.2

TABLE X: Summary of selection efficiencies for $t\bar{t}$ signal, $W + jets$ and QCD background for the Likelihood Discriminant Cut Method.

	$t\bar{t}$	$W + jets$ bkg.	QCD bkg.
$\varepsilon_{preselection}$	4.6±0.4%	-	-
$\varepsilon_{LH>0.7\ cut}$	59.8±0.9%	18.7±0.6%	28.1±4.2%
$\varepsilon_{selection}$	2.8±0.2%	-	-

F. Results

The result is obtained by using the Likelihood Fit Method. The $\mathrm{t\bar{t}}$ production cross section in $\mu + jets$ events is given by:

$$\sigma = \frac{N_t^{t\bar{t}}}{Br \cdot \mathcal{L} \cdot \varepsilon} , \tag{19}$$

where $N_t^{t\bar{t}}$ is the number of fitted $t\bar{t}$ events, Br the branching fraction for the $\ell + jets$ final state, $\mathcal{L}$ the integrated luminosity and ε the preselection efficiency. The results from the Likelihood fit Method are shown in Figure 24, and the summary of the results is given in Table XI. As a sanity check the likelihood input distributions (see Fig. 25) and various typical distributions (see Fig. 26) are overlaid for the preselected data and for the result from the fit of $t\bar{t}$, $W + jets$ and QCD.
The result of the $\mathrm{t\bar{t}}$ production cross section in the $\mu + jets$ channel for the Likelihood fit Method at $\sqrt{s}$ =1.96 TeV yields:

$$\sigma_{p\bar{p}\to t\bar{t}+X} = 7.5\ ^{+3.4}_{-3.0}\ (\text{stat.})\ ^{+1.9}_{-1.6}\ (\text{syst.}) \pm 0.5\ (\text{lumi.})\ \text{pb}. \tag{20}$$

1. *Results from the Likelihood Discriminant Cut Method*

Alternatively, the $t\bar{t}$ cross section is extracted with the Likelihood Discriminant Cut Method. Here the $\mathrm{t\bar{t}}$ production cross section in $\mu + jets$ events is given by:

$$\sigma = \frac{N_{sel}^{t\bar{t}}}{Br \cdot \mathcal{L} \cdot \varepsilon} , \tag{21}$$

where $N_{sel}^{t\bar{t}}$ is the number of $t\bar{t}$ events after applying the likelihood cut at 0.7. Br is the branching fraction for the $\ell + jets$ final state, $\mathcal{L}$ the integrated luminosity and ε the overall efficiency. The summary of the results for the Likelihood Discriminat cut Method is given in Table XII.

TABLE XI: Summary of the results for the Likelihood Discriminat fit Method.

N^{obs}	$\mathcal{B}r$	$N_t^{t\bar{t}}$	N_t^{W}	N_t^{QCD}	$\mathcal{L}$	ε
100	44.4%	$22.1^{+9.9}_{-8.9}$	$70.6^{+12.4}_{-11.6}$	$7.0^{+0.9}_{-0.8}$	143.9 pb^{-1}	4.6%

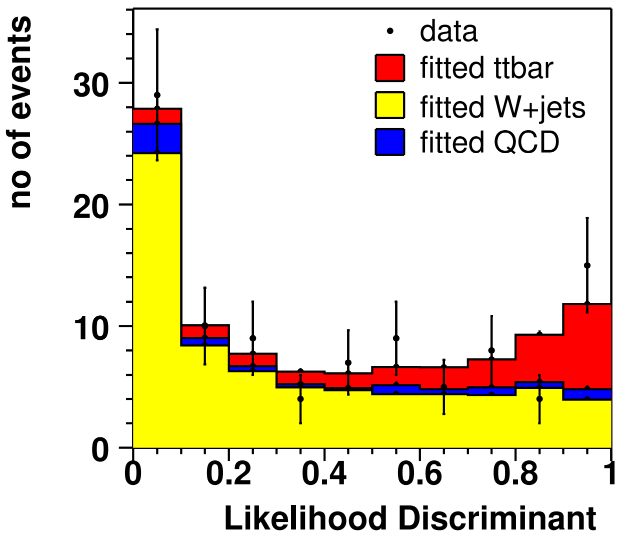

FIG. 24: Results from the likelihood fit.

TABLE XII: Summary of the results for the Likelihood Discriminat cut Method, using the second matrix method to determine $N_{sel}^{t\bar{t}}$, N_{sel}^{W} and N_{sel}^{QCD}.

N^{obs}	$\mathcal{B}r$	LH cut	$N_{sel}^{t\bar{t}}$	N_{sel}^{W}	N_{sel}^{QCD}	$\mathcal{L}$	ε
27	44.4%	0.7	$11.1^{+6.9}_{-6.2}$	$13.9^{+2.6}_{-2.5}$	$2.0^{+0.2}_{-0.2}$	143.9 pb^{-1}	2.8%

The result of the $\mathrm{t\bar{t}}$ production cross section in the $\mu + jets$ channel for the Likelihood Discriminat cut Method at $\sqrt{s}$ =1.96 TeV yields:

$$\sigma_{p\overline{p}\to t\bar{t}+X} = 6.2\ ^{+3.9}_{-3.5}\ (\mathrm{stat.})\ ^{+1.3}_{-1.3}\ (\mathrm{syst.}) \pm 0.4\ (\mathrm{lumi.})\ \mathrm{pb}. \tag{22}$$

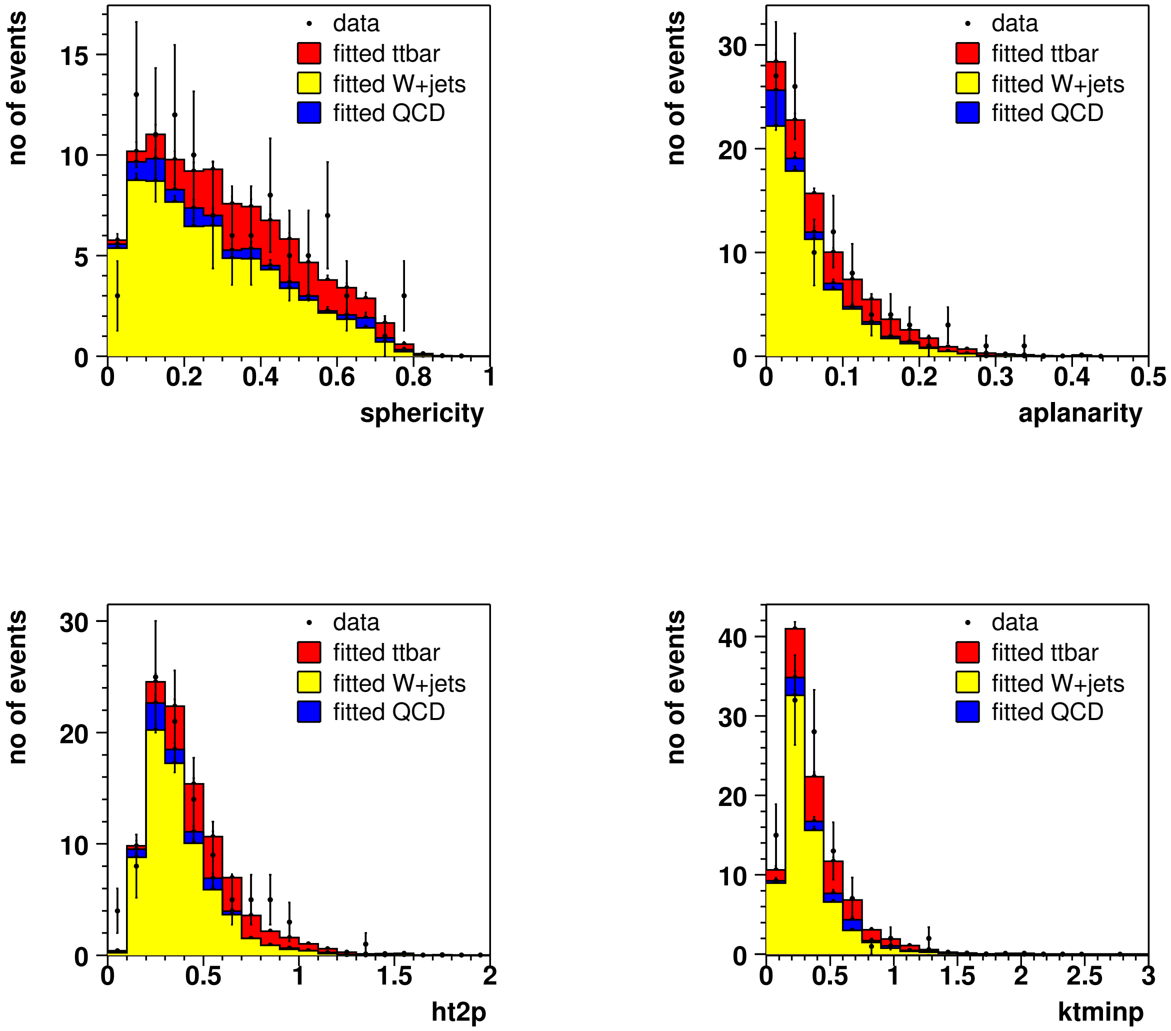

FIG. 25: Sanity Checks: The distributions of the topological variables that are used as an input for the likelihood discriminant are overlaid for the preselected data and for the result from the fit of $t\bar{t}$, $W + jets$ and QCD.

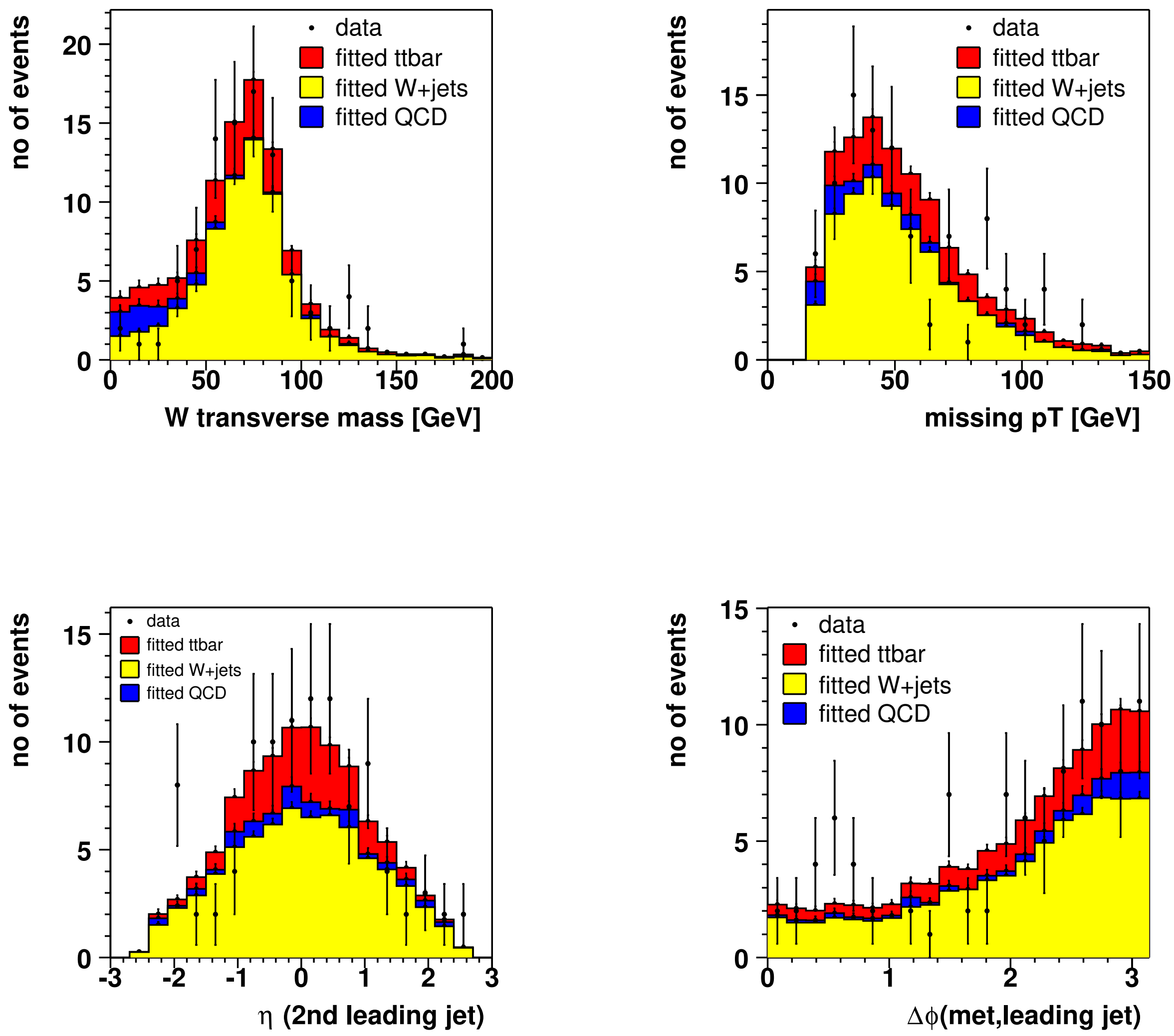

FIG. 26: Sanity Checks: The distributions of typical variables are overlaid for the preselected data and for the result from the fit of $t\bar{t}$, $W + jets$ and QCD.

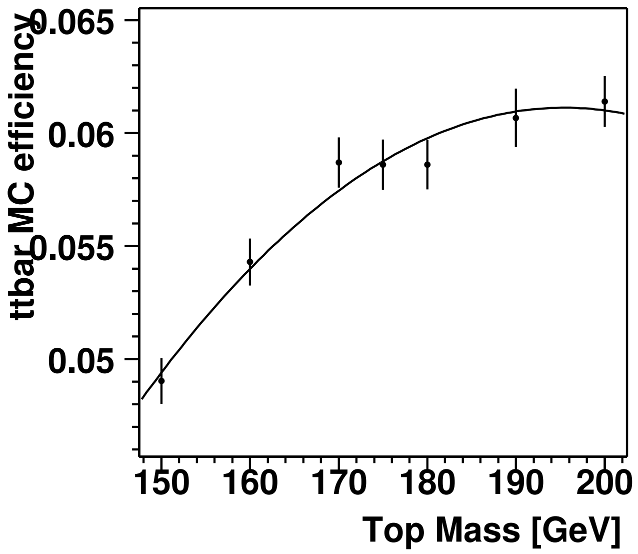

FIG. 27: mass dependence

TABLE XIII: Summary of the relative systematic uncertainties on the preselection.

source	relative systematic uncertainty
primary vertex	±2%
tracking	±2%
μ ID	±1.6%
L1 μ trigger	+2.0%-2.5%
L2 μ trigger	+3.5%-4.9%
μ isolation	±3.4%
μ σ_{dca}	±1%
Δz	±1%
χ^2	±1%
JES	+6.1%-11.7%
JER	+1.3%-1.6%
Jet ID	−9.8%
L1 jet trigger	0.0%
L2 jet trigger	0.0%
L3 jet trigger	+0.3%-2.4%
soft-μ veto	±6%
Tune A	-3.4%
top mass	+1.7%-2.2%

G. Systematic Uncertainties

The considered uncertainties are summarized in Tables XIII- XV. Figure 27 shows the mass dependence of the $t\bar{t}$ MC preselection efficiency. Figures 28- 30 show the change in the likelihood discriminant distribution for a systematic variation of the JES, the JER and the top mass. Table XIV summarizes the relative changes on the fitted number of $t\bar{t}$ events from the different top masses.

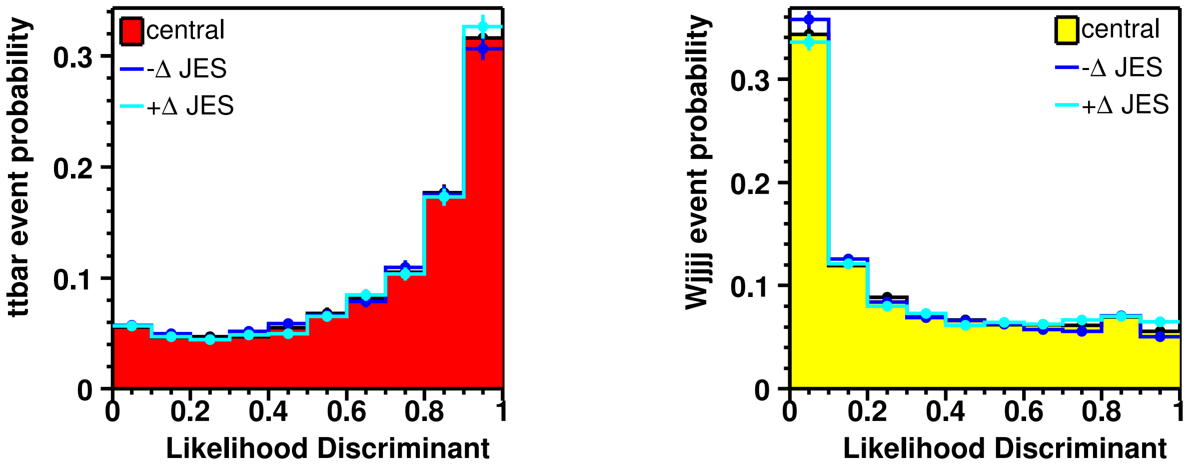

FIG. 28: Likelihood discriminant distributions using the central value for the JES and using JES+ΔJES and JES-ΔJES to correct the jet energy. Left: for $t\bar{t}$, right: for $W + jets$

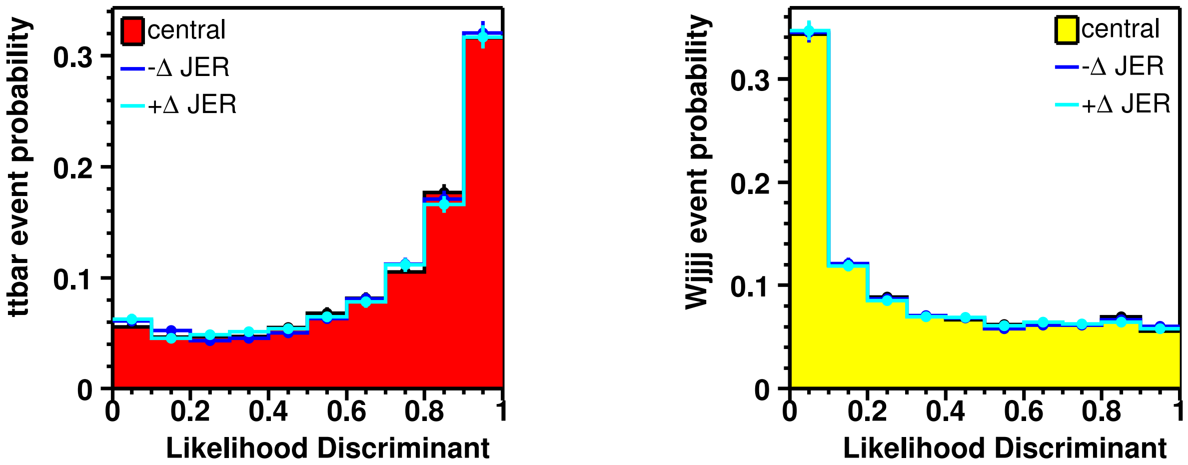

FIG. 29: Likelihood discriminant distributions using the central value for the JER and using JER+ΔJER and JER-ΔJER to correct the jet energy resolution in the simulation. Left: for $t\bar{t}$, right: for $W + jets$

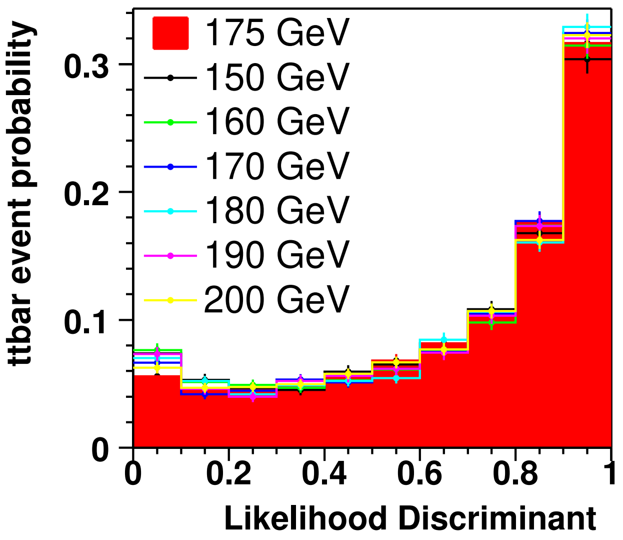

FIG. 30: Likelihood discriminant distributions for a $t\bar{t}$ mass of 175 GeV, and for a $t\bar{t}$ mass of 170 GeV and 180 GeV.

TABLE XIV: Summary of the relative systematic uncertainties originating from different masses of the top quark on the fitted number of $t\bar{t}$ events for the Likelihood fit Method.

top mass [GeV]	relative systematic uncertainty	
	Fit Method	Cut Method
150	+7.2%	+1.5%
160	+3.4%	+1.0%
170	−1.9%	−0.6%
180	+0.6%	+0.2%
190	+0.6%	+0.1%
200	+3.7%	+0.5%

TABLE XV: Summary of the relative systematic uncertainties on the number of fitted $t\bar{t}$ events for the Likelihood fit Method.

source	relative systematic uncertainty
JES	−12.8%+11.5%
Jet ID	−11.6%
Jet resolution	+0.8%−2.0%
L1 μ trigger	+0.0%−0.0%
L2 μ trigger	+0.2%−0.2%
L1 jet trigger	+0.2%−0.2%
L2 jet trigger	+0.2%−0.2%
L3 jet trigger	+0.2%−1.7%
top mass	±2.0%
Tune A	−2.1%
Template statistics	±5.%
ε_{sig}	−0.2%+0.2%
ε_{QCD}	−2.6%+2.4%

TABLE XVI: Summary of the relative systematic uncertainties on the estimated number of $t\bar{t}$ events for the Likelihood cut Method using the second matrix method.

source	relative systematic uncertainty
JES	$-15.3\%+9.9\%$
Jet ID	-12.8%
Jet resolution	$+1.7\%-1.8\%$
L1 μ trigger	0.0%
L2 μ trigger	0.0%
L1 jet trigger	0.0%
L2 jet trigger	0.0%
L3 jet trigger	$+3.7\%-2.5\%$
top mass	$\pm1.0\%$
Tune A	-0.4%
ε_{sig}	$-0.2\%+0.3\%$
ε_{QCD}	$-3.0\%+2.8\%$
$\varepsilon_{sel}^{\mathrm{t\bar{t}}}$	$-0.7\%+0.7\%$
$\varepsilon_{sel}^{\mathrm{W}}$	$-5.4\%+5.3\%$
$\varepsilon_{sel}^{\mathrm{QCD}}$	$-3.9\%+3.9\%$

TABLE XVII: Loose and tight selection criteria in e+jets channel.

Loose selection criteria
EM particle, $E_T^e > 20$ GeV
$\lvert\eta_e\rvert < 1.1$ (CC) or $1.5 < \lvert\eta_e\rvert < 2.5$ (EC)
electron has a matched track
μ tag veto
second electron veto
$\not{E}_T > 20$ GeV
cut in $\Delta\phi(\not{E}_T, e)$ vs. $\not{E}_T$ plane
Primary vertex with ≥ 3 tracks and $\lvert z\rvert < 60$ cm
Δz(primary vertex, e) < 1 cm
Tight selection criteria
electron passes likelihood

VIII. ELECTRON+JETS CHANNEL

A. Event Preselection

The characteristic topology of the electron+jets channel is one isolated high transverse momentum electron, a large $\not{E}_T$ and four or more high transverse momentum jets. The preselection cuts are designed to define a data sample enriched in W+jets and top events.

Triggered event is selected for the analysis if it satisfies primary vertex cuts (section III A). The jets are required to pass standard identification cuts, to be within $|\eta| < 2.5$ and have $p_T > 15$ GeV. We demand that a selected event has large $\not{E}_T$.

Two vetoes are imposed: the presence of a soft muon within a jet is vetoed, and the presence of a second electron with $p_T > 15$ GeV is vetoed. These requirements ensure that the e+jets topological analysis is orthogonal to the e+jets analysis with soft muon tagging, and to the dielectron analysis.

The loose preselection is used to perform an evaluation of the absolute normalization of the QCD background. The tight preselection implies that EM object passes electron likelihood cuts as discussed in section III B in addition to the loose selection. Since electrons are found to have different efficiencies in CC and EC, these regions are treated separately throughout the whole e+jets analysis. A summary of the selection cuts for both the loose and tight samples is given in Table XVII for the e+jets channel.

As was mentioned above, the main background source to W+jets production are the QCD multijet events with a jet misidentified as an electron or an electron produced inside a jet. Thus QCD events contain both real and fake electrons. At the same time energy imbalance in QCD events occurs predominately due to jet energy mismeasurements, in which case the direction of the missing transverse energy points along (or opposite) to one of the objects (electron or jet) in the event. In order to suppress the QCD background more efficiently than just by the applied requirement on missing transverse energy ($\not{E}_T$), we optimized our selection using the two event characteristics, missing transverse energy and the difference in azimutal angle between electron and missing energy ($\Delta\phi(\not{E}_T, e)$). The two-dimensional distribution in these variables is presented in Fig 31 (a) for a ttbar MC sample and in Fig 31 (b) for the QCD control sample. It is apparent that in QCD events missing energy tends to be back to back with the electron and is in general smaller than in the signal sample.

The cut was optimized using a grid search technique retaining the original signal efficiency of the previously used plain $\Delta\Phi$ cut ($\Delta\Phi > 0.5$) and improving the QCD rejection to minimize systematics due to this background which will be enhanced again in a b-tagging analysis.

The grid search results are shown in Fig 32. The selected cut:

$$\Delta\phi(\not{E}_T, e) > 1.7 - \frac{1.7\times \not{E}_T}{80GeV} \tag{23}$$

is 87% efficient to signal and 43% to QCD background. For comparison, the previously used square cut had QCD efficiency of 62%.

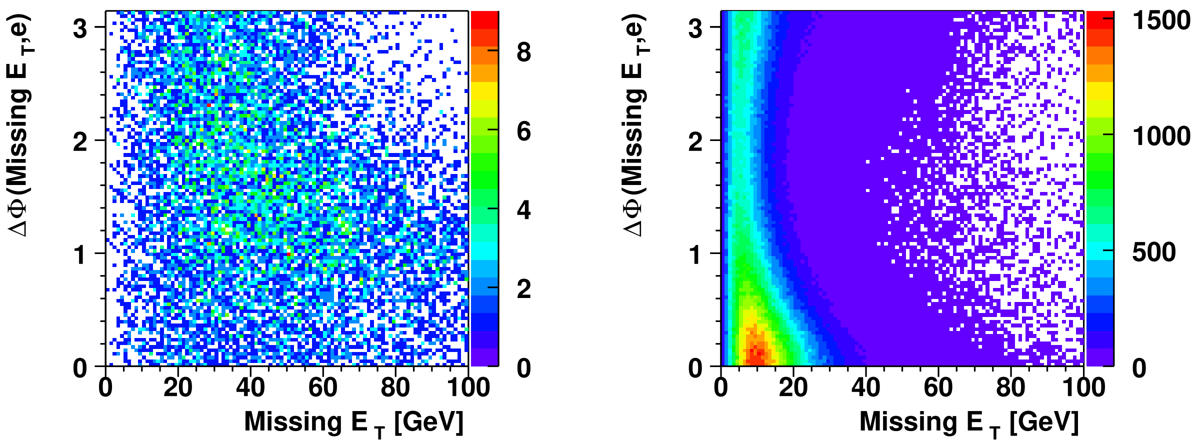

FIG. 31: $\Delta\phi(\not{E}_T, e)$ vs. $\not{E}_T$ for $t\bar{t}$ MC (a) and a QCD enriched data sample (b).

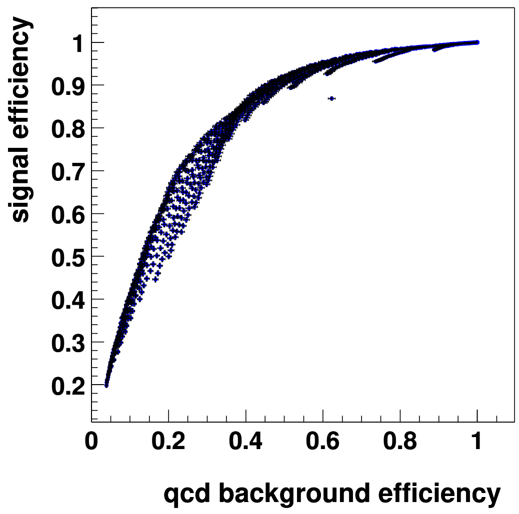

FIG. 32: Efficiencies for $t\bar{t}$ and QCD for different triangular cuts in the $\Delta\phi(\not{E}_T, e)$ vs. $\not{E}_T$ plane. The isolated point shows the efficiencies of the old cut $\Delta\Phi > 0.5$.

B. $\mathrm{t\bar{t}}$ Event Efficiencies

The efficiency of either preselection stage is determined from either $t\bar{t}$ Monte Carlo with correction factors applied where appropriate to account for the difference between data and Monte Carlo efficiencies derived from comparison of $Z \rightarrow ee$ data and Monte Carlo samples. The efficiencies are with respect to $t\bar{t} \rightarrow l + jets$ in order to account for the possible contributions to the signal from $t\bar{t} \rightarrow \tau + jets$.

The efficiencies of all preselection cuts is summarised in Table XVIII.

TABLE XVIII: Selection cuts efficiencies for the $e + jets$ channel.

Selection	ε (CC)	$\varepsilon_{cumulative}$ (CC)	ε (EC)	$\varepsilon_{cumulative}$ (EC)
Recoed EM in acceptance	0.343 ± 0.002		0.039 ± 0.001	
Reco Data-MC correction factor	0.984 ± 0.016	0.338 ± 0.006	0.968 ± 0.020	0.0378 ± 0.0012
PassEMID	0.792 ± 0.003		0.906 ± 0.007	
PassEMID Data-MC correction factor	1.014 ± 0.001	0.271 ± 0.005	1.004 ± 0.001	0.0343 ± 0.0011
Trmatch	0.811 ± 0.004		0.722 ± 0.015	
Trmatch Data-MC correction factor	0.945 ± 0.004	0.208 ± 0.004	0.786 ± 0.009	0.0195 ± 0.0008
lhood	0.914 ± 0.004		0.873 ± 0.014	
lhood Data-MC correction factor	0.931 ± 0.004	0.177 ± 0.004	0.920 ± 0.008	0.0156 ± 0.0007
2nd EM veto	0.944 ± 0.002		0.911 ± 0.007	
Corrected soft mu veto	0.809 ± 0.012	0.135 ± 0.003	0.825 ± 0.014	0.0118 ± 0.0006
EM-pT cut	0.921 ± 0.003	0.124 ± 0.003	0.771 ± 0.012	0.0091 ± 0.0005
delta Phi and MET cut	0.799 ± 0.005	0.099 ± 0.003	0.788 ± 0.016	0.0071 ± 0.0004
PV cuts (#tracks, z-pos)	0.989 ± 0.008	0.098 ± 0.003	0.989 ± 0.008	0.0070 ± 0.0004
PV cut (delta z)	0.964 ± 0.001	0.095 ± 0.003	0.964 ± 0.001	0.0068 ± 0.0004
#jets ≥ 4	0.573 ± 0.007	0.054 ± 0.002	0.615 ± 0.023	0.0042 ± 0.0003
Trigger Efficiency	0.930 ± 0.027	0.050 ± 0.002	0.800 ± 0.064	0.0033 ± 0.0003
Total signal efficiency	0.050 ± 0.002	0.050 ± 0.002	0.0033 ± 0.0003	0.0033 ± 0.0003

TABLE XIX: ε_{QCD} and ε_{sig} for different inclusive jet multiplicities in the e+jets analysis, with their respective total errors.

	CC		EC	
N_{Jets}	ε_{sig}	ε_{QCD}	ε_{sig}	ε_{QCD}
≥ 0 jets	0.897 ± 0.006	-	0.873 ± 0.008	-
≥ 1 jet	0.891 ± 0.009	0.145 ± 0.002	0.861 ± 0.014	0.116 ± 0.002
≥ 2 jets	0.881 ± 0.019	0.154 ± 0.004	0.832 ± 0.044	0.122 ± 0.004
≥ 3 jets	-	0.143 ± 0.009	-	0.120 ± 0.008
≥ 4 jets	-	0.147 ± 0.022	-	0.080 ± 0.019

C. The Matrix Method

As discussed in section IV QCD background in the preselected sample is evaluated using Matrix Method. Tight and loose samples used to apply this method are defined in the previous section. The rates at which W+$t\bar{t}$ on the one hand and QCD on the other are reduced, from one set to the other (respectively ε_{sig} and ε_{QCD}), are well separated and can be measured in data. Figures 33 and 34 show ε_{QCD} as a function of $\not{E}_T$ for the CC and EC region respectively. The tight and loose selection differs only by the requirement that the electron passes the likelihood cut in the tight sample. The distribution of ε_{QCD} is flat in the low $\not{E}_T$ region and is determined from a fit to a constant within $\not{E}_T <$ 10 GeV. Measured ε_{QCD} in the different jet multiplicity bins are summarized in Table XIX.

The corresponding rate at which signal events are reduced from the loose to the tight sample, ε_{sig}, is by definition the electron likelihood efficiency which is determined on a $Z \to ee$ data sample. ε_{sig} is found to be, within errors, constant as a function of jet multiplicity and the value used in the analysis is the one obtained in the first jet multiplicity bin. To account for the kinematic differences between the Z events and $t\bar{t}$, a correction factor derived from $Z \to ee$ and $t\bar{t}$ Monte Carlo is applied to ε_{sig} before using the matrix method. Those correction factors are found to be: 0.983 ± 0.004(CC) and 0.986 ± 0.013(EC).

The numbers of observed events in the loose (N_l) and in the tight (N_t) samples along with the estimated number of W+jets background plus $t\bar{t}$ signal events and QCD background events from the Matrix Method for each inclusive jet multiplicity bin are listed in Table XX. The QCD background in the transverse mass distributions (as predicted by the Matrix Method) is shown in Figures 35,36 for CC and EC, respectively.

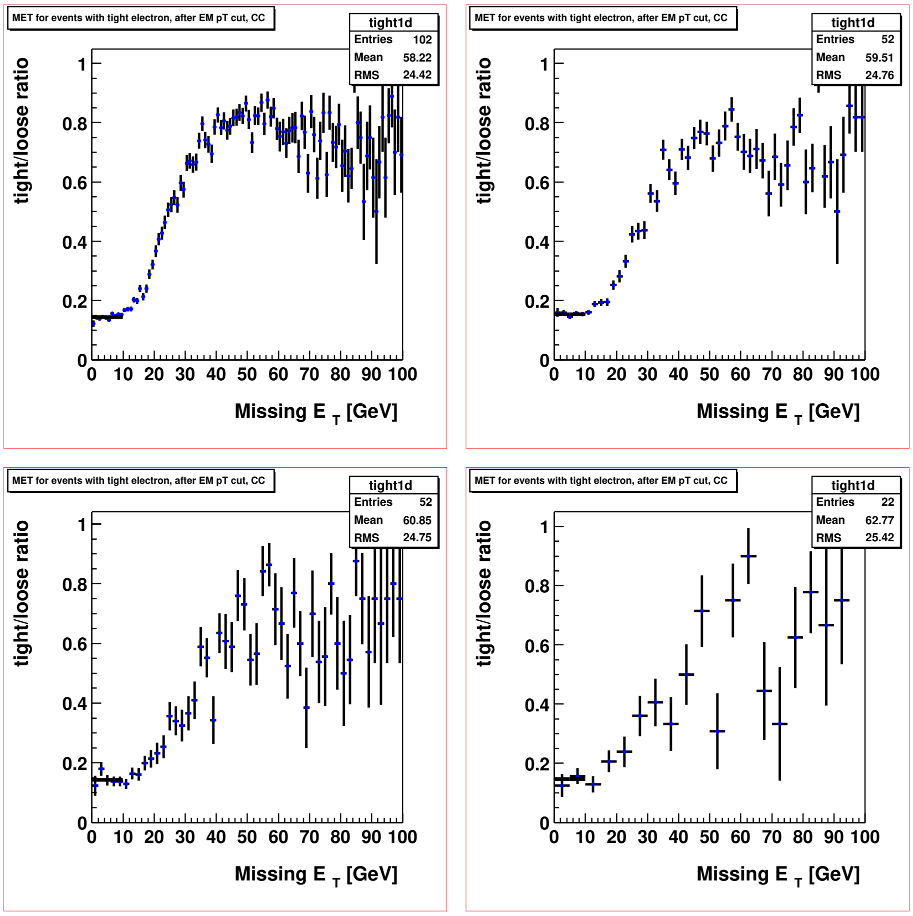

FIG. 33: ε_{QCD} for the different N_{Jets} multiplicities as a function of $\not\!\!E_T$in the CC.

TABLE XX: N_l, N_t, $N_t^{\mathrm{W+t\bar{t}}}$ and N_t^{QCD} for the $e + jets$ analysis.

		CC		
	N_l	N_t	$N_t^{\mathrm{W+t\bar{t}}}$	N_t^{QCD}
≥ 1 jet	12010	7908	7391 $\pm$ 135	517 $\pm$ 20
≥ 2 jets	4636	2580	2264 $\pm$ 62	316 $\pm$ 11
≥ 3 jets	1348	628	520 $\pm$ 28	108 $\pm$ 6
≥ 4 jets	311	136	109 $\pm$ 13	27 $\pm$ 3
		EC		
	N_l	N_t	$N_t^{\mathrm{W+t\bar{t}}}$	N_t^{QCD}
≥ 1 jet	5039	2204	1875 $\pm$ 66	329 $\pm$ 10
≥ 2 jets	2285	799	608 $\pm$ 33	191 $\pm$ 6
≥ 3 jets	712	208	143 $\pm$ 16	65 $\pm$ 3
≥ 4 jets	178	46	35 $\pm$ 7	11 $\pm$ 1

D. Topological variables

The distributions of the four topological variables that were introduced in section IV B 1 are shown in figures 37 and 38 for $t\bar{t}$ signal, Wjjjj Monte Carlo and the QCD enriched data sample selected using all preselection cuts except track match for EM object and applying inverted likelihood cut. Since similar to μ+jets channel Wjjjj and QCD backgrounds in $e + jets$ channel exhibit very similar behaviour for the chosen variables, a two-class likelihood was built following the prescription described in section IV B 2. The resulting likelihood distributions for $t\bar{t}$ and Wjjjj

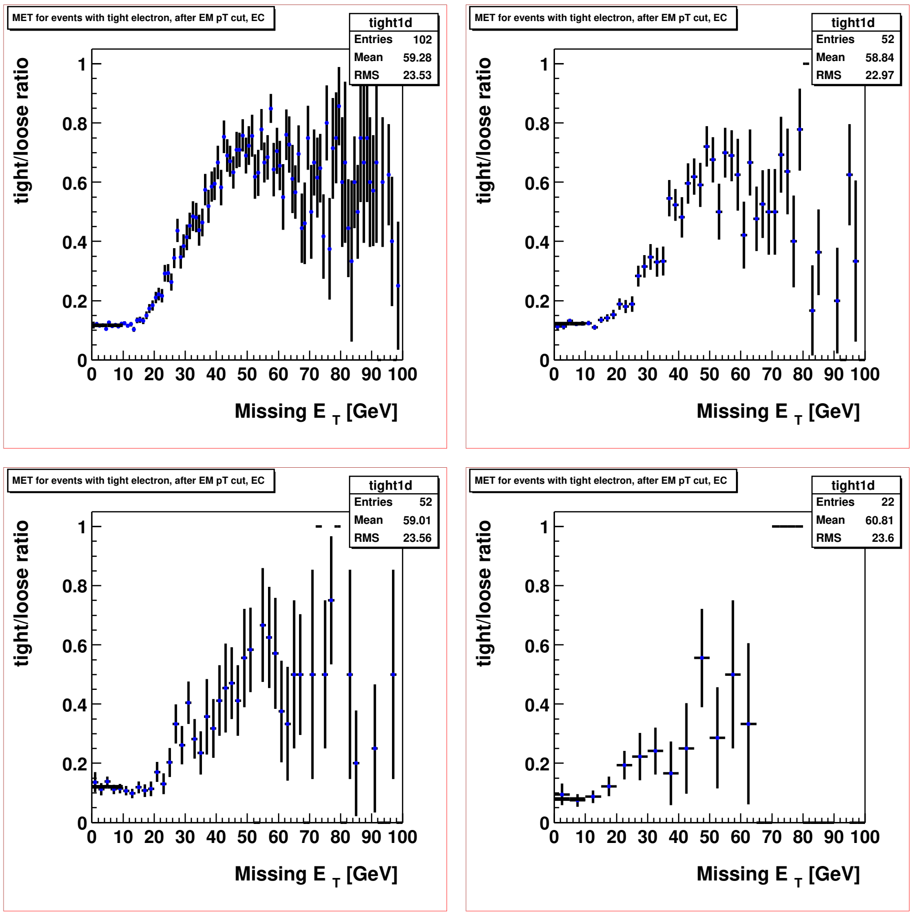

FIG. 34: ε_{QCD} for the different jet multiplicities as a function of $\not\!\!E_T$ in the EC.

TABLE XXI: Summary of the fitted values from the likelihood fit in e+jets channel.

CC

N^{obs}	$\mathcal{B}r$	$N_{t\bar{t}}$	N_W^{bkg}	N_{QCD}^{bkg}	$\mathcal{L}$	ε
136	44.4%	$30.16^{+12.4}_{-11.2}$	$78.0^{+15.1}_{-14.5}$	$27.6^{+2.4}_{-2.3}$	141.2 pb^{-1}	5.0%

EC

N^{obs}	$\mathcal{B}r$	$N_{t\bar{t}}$	N_W^{bkg}	N_{QCD}^{bkg}	$\mathcal{L}$	ε
46	44.4%	$0.8^{+4.5}_{-0.8}$	$34.1^{+7.8}_{-7.2}$	10.9 ± 1.0	141.2 pb^{-1}	0.33%

Monte Carlo events and QCD events from data are shown in figures 39, 40 for CC and EC respectively.

E. Results

The fraction of $t\bar{t}$ events in the tight preselected signal sample is obtained by fitting the data likelihood distribution with fractions of the $t\bar{t}$ and Wjjjj likelihoods and the QCD contribution set to the value obtained from the matrix method but allowing for Poissonian fluctuations (Likelihood Fit Method). The result of this fit is shown in figures 41,42 for CC and EC respectively.

The obtained fractions are summarized in Table XXI.

The result of the $\mathrm{t\bar{t}}$ production cross section in the e+jets channel for the Likelihood fit Method at $\sqrt{s}$ =1.96 TeV

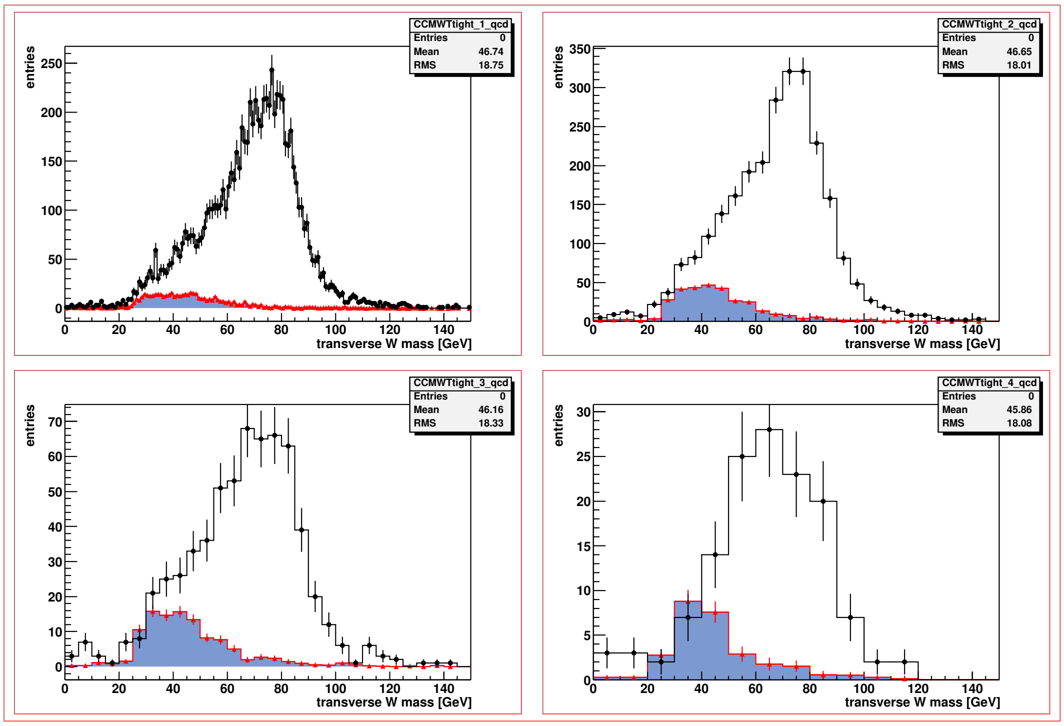

FIG. 35: Transverse mass distributions for the CC in the ≥ 1, ≥ 2, ≥ 3, and ≥ 4 jet multiplicity bins. Dots are the observed numbers of events after tight selection cuts, the colored histograms represent the QCD background contribution.

yields:

$$\begin{aligned}
CC &: \sigma_{p\bar{p}\to t\bar{t}+X} = 9.6\ ^{+4.0}_{-3.6}\ \text{(stat.)}\ ^{+1.7}_{-2.4}\ \text{(syst.)} \pm 0.6\ \text{(lumi.) pb;} \\
EC &: \sigma_{p\bar{p}\to t\bar{t}+X} = 4.0\ ^{+21.9}_{-4.0}\ \text{(stat.)}\ ^{+4.5}_{-8.7}\ \text{(syst.)} \pm 0.3\ \text{(lumi.) pb;} \\
all &: \sigma_{p\bar{p}\to t\bar{t}+X} = 9.4\ ^{+3.9}_{-3.5}\ \text{(stat.)}\ ^{+1.6}_{-2.6}\ \text{(syst.)} \pm 0.6\ \text{(lumi.) pb.}
\end{aligned}$$

F. Systematic Uncertainties

Systematic uncertainties on the preselection efficiencies in $e + jets$ channel are summarized in Tables XXII and XXIV for CC and EC, respectively. Uncertainies on the topological likelihood can be found in Tables XXIII and Tables XXV.

ε_{sig}

IX. SUMMARY

A preliminary measurement of the $t\bar{t}$ production cross-section in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV has been presented. The analysis is based on the application of a topological method to the lepton+jets preselected samples, corresponding to integrated luminosities of 141.2 pb^{-1} (e+jets) and 143.9 pb^{-1} (μ+jets).

The preliminary cross-sections obtained for each of the lepton+jets channels are:

$$\begin{aligned}
\mathrm{e + jets} &: \quad \sigma_{p\bar{p}\to t\bar{t}+X} = 9.4^{+3.9}_{-3.5}\ \text{(stat)}\ ^{+1.6}_{-2.6}\ \text{(syst)}\ \pm 0.6\ \text{(lumi) pb,} \\
\mu + \mathrm{jets} &: \quad \sigma_{p\bar{p}\to t\bar{t}+X} = 7.5^{+3.4}_{-3.0}\ \text{(stat)}\ ^{+1.9}_{-1.6}\ \text{(syst)}\ \pm 0.5\ \text{(lumi) pb,}
\end{aligned}$$

leading to the following preliminary combined cross-section:

$$\mathrm{l + jets}\ :\ \sigma_{p\bar{p}\to t\bar{t}+X} = 8.4^{+2.5}_{-2.3}\ \text{(stat)}\ ^{+1.5}_{-1.8}\ \text{(syst)}\ \pm 0.5\ \text{(lumi) pb.}$$

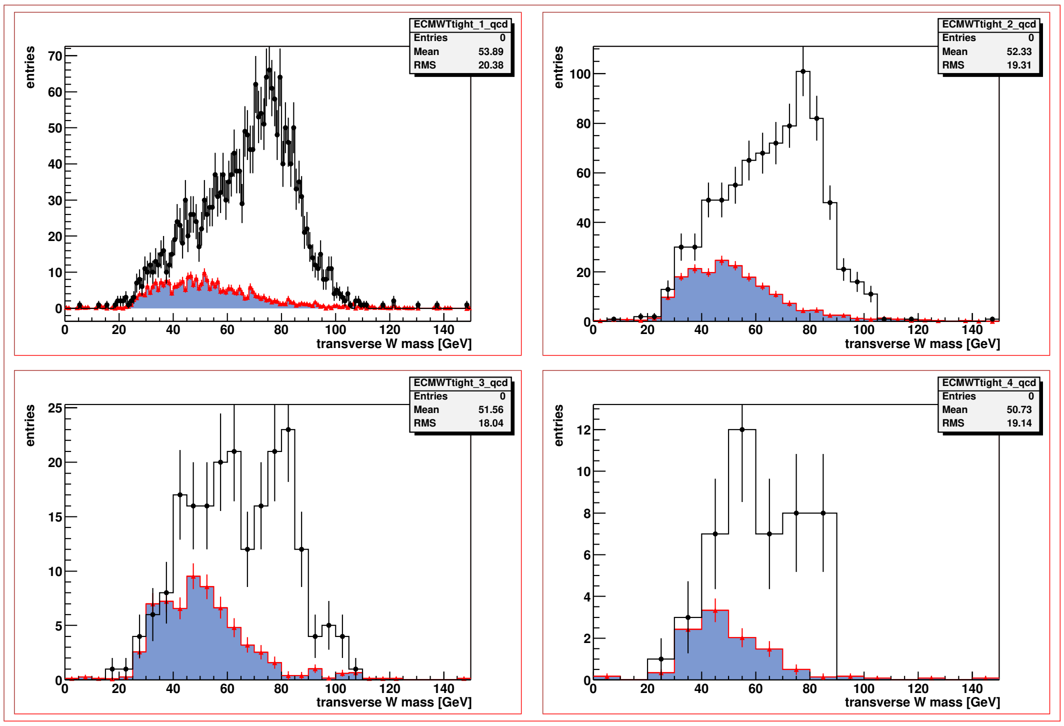

FIG. 36: Transverse mass distributions for the EC in the ≥ 1, ≥ 2, ≥ 3, and ≥ 4 jet multiplicity bins. Dots are the observed numbers of events after tight selection cuts, the colored histograms represent the QCD background contribution.

TABLE XXII: Summary of the relative systematic uncertainties on the signal preselection efficiencies(CC).

source	relative systematic uncertainty
primary vertex	1%
EM reconstruction	1.5%
EM ID	0.1%
EM tracking	0.4%
EM likelihood	0.5%
L1 EM trigger	+2.1% -2.7%
L2 EM trigger	
L3 EM trigger	1.0%
$\|\Delta z(e, PV)\|$	1%
JES	+6.7% -9.7%
Jet ID	-11%
Jet resolution	2.2%
L1 jet trigger	+0.01% -0.05%
L2 jet trigger	
L3 jet trigger	+0.01% -0.5%
soft-μ veto	6%
tune A	2.0%
top mass	+1.7% -2.2%

X. ACKNOWLEDGEMENTS

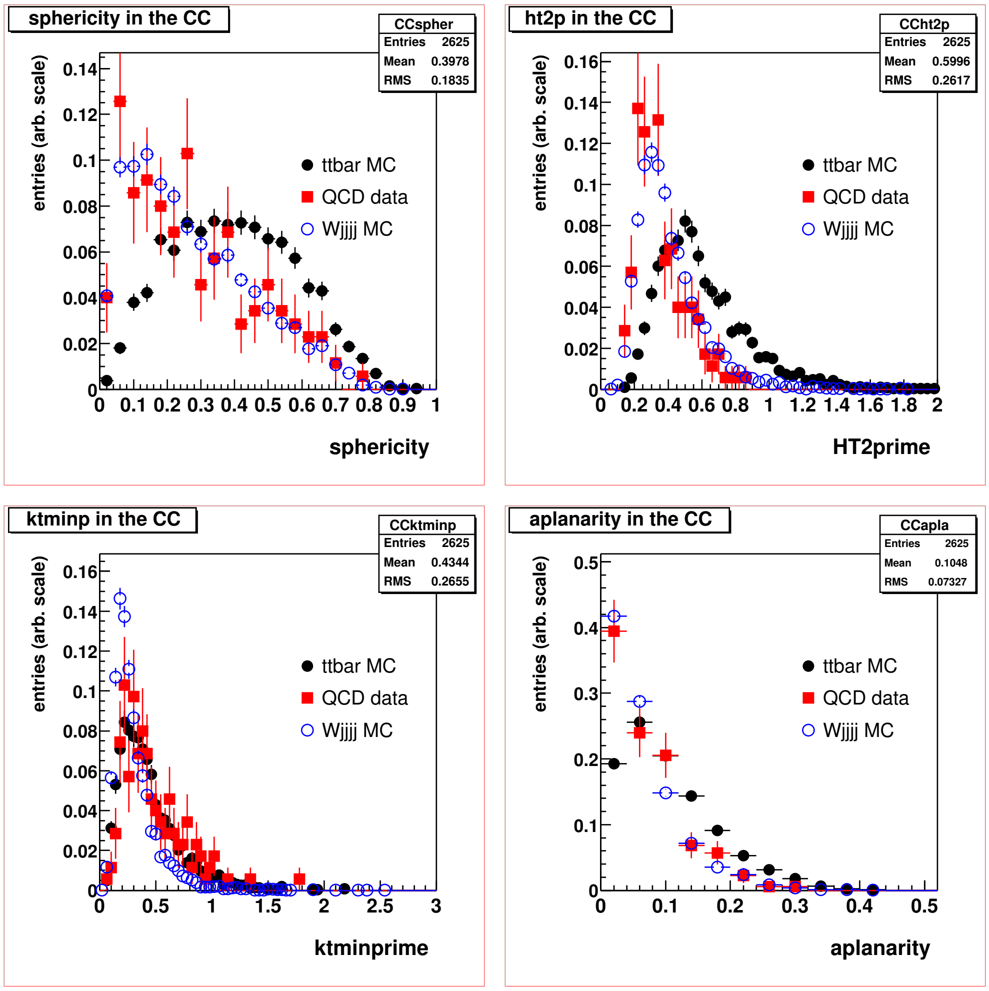

FIG. 37: Distributions of the four topological variables for $t\bar{t}$, W+jets and QCD in the CC.

TABLE XXIII: Summary of the relative systematic uncertainties on the number of fitted $t\bar{t}$ events (CC).

source	relative systematic uncertainty
L1 EM trigger	+0.03% −0.4%
L2 EM trigger	
L3 EM trigger	0.3%
JES	+7.2% −5.7%
Jet ID	-18%
Jet resolution	7%
L1 jet trigger	+0.2% −0.3%
L2 jet trigger	
L3 jet trigger	+0.04% −0.4%
tune A	1%
top mass	2%
template statistics	7%
ε_{sig}	0.07%
ε_{QCD}	−1.5% +1.4%

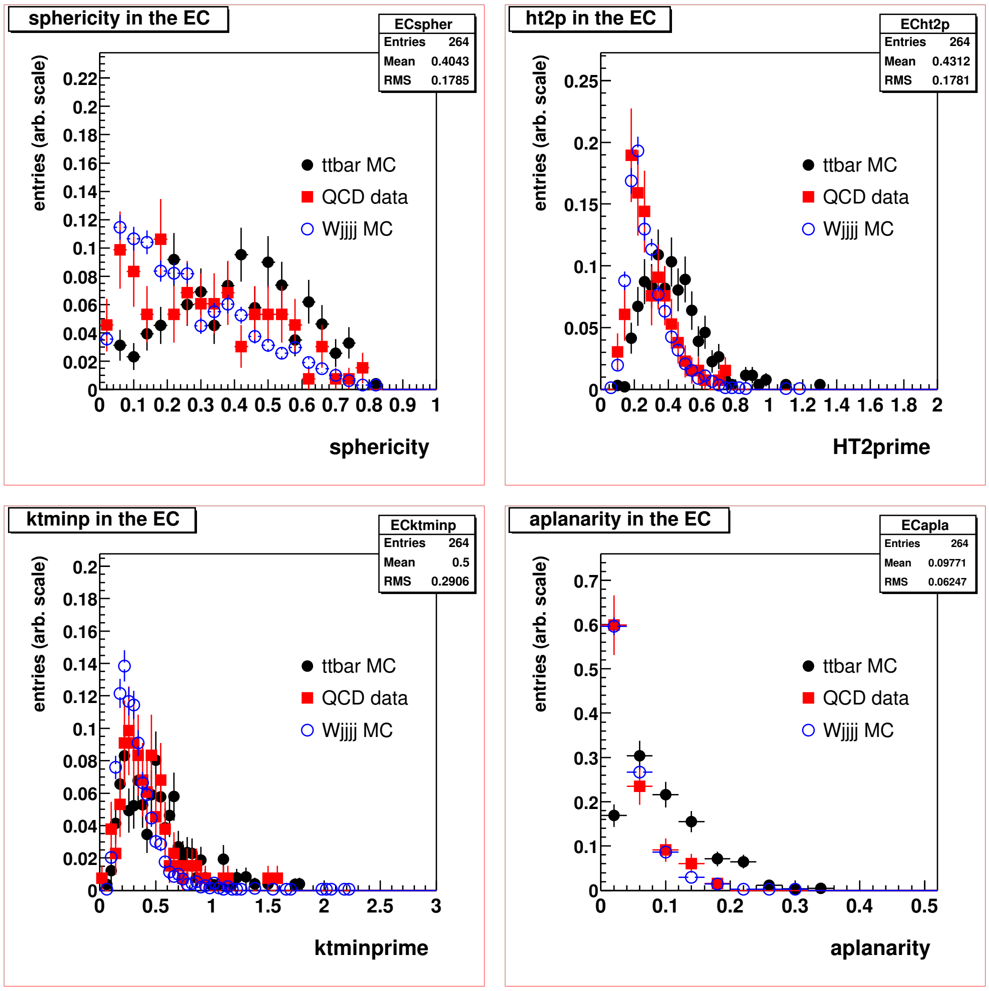

FIG. 38: Distributions of the four topological variables for $t\bar{t}$, W+jets and QCD in the EC.

TABLE XXIV: Summary of the relative systematic uncertainties on the signal preselection efficiencies (EC).

source	relative systematic uncertainty
primary vertex	1%
EM reconstruction	1.5%
EM ID	0.1%
EM tracking	0.5%
EM likelihood	0.8%
L1 EM trigger	+2.3% -4.9%
L2 EM trigger	
L3 EM trigger	6.3%
$\lvert\Delta z(e, PV)\rvert$	1%
JES	+5.7%-10.5%
Jet ID	-17%
Jet resolution	3.5%
L1 jet trigger	+0.2% -0.04%
L2 jet trigger	
L3 jet trigger	+0.01% -0.5%
soft-μ veto	6%
tune A	20%
top mass	+1.1% $-$0.9%

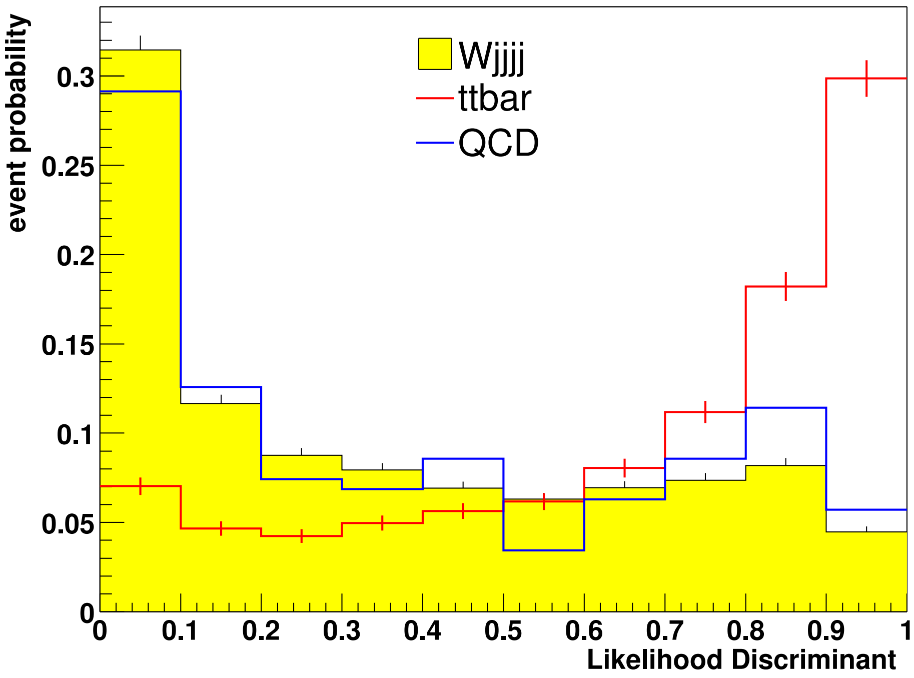

FIG. 39: Likelihood discriminant distributions for $t\bar{t}$, W+jets and QCD in the CC.

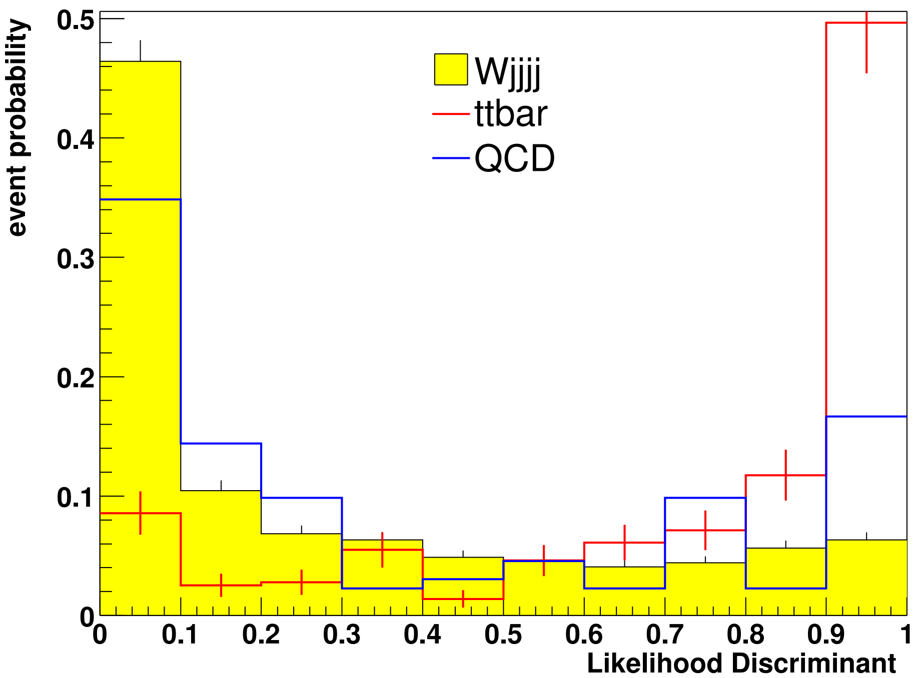

FIG. 40: Likelihood discriminant distributions for $t\bar{t}$, W+jets and QCD in the EC.

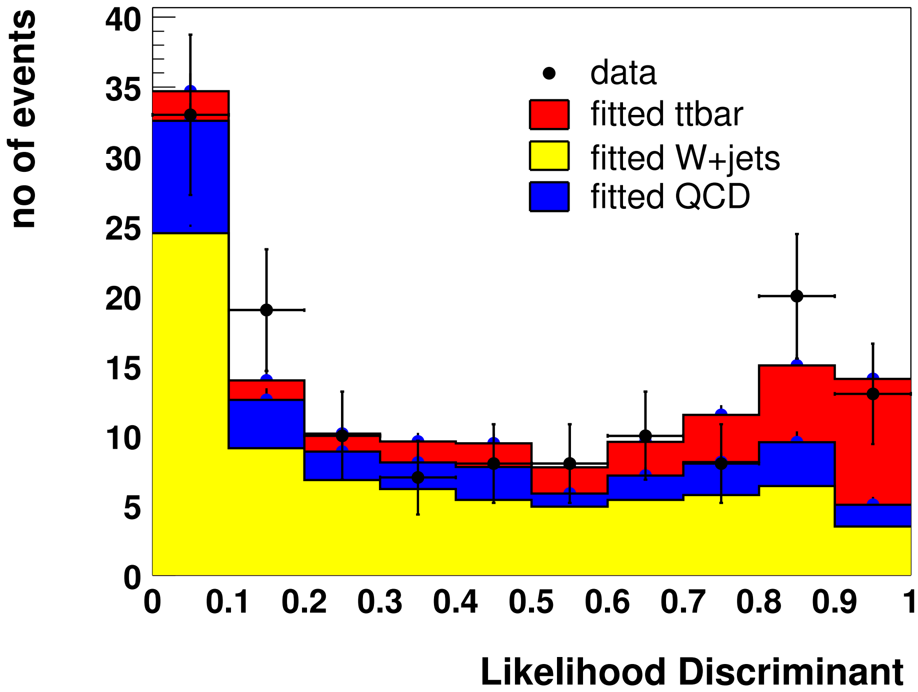

FIG. 41: The result of the topological likelihood fit showing data and fitted $t\bar{t}$, W+jets and QCD contributions in the CC.

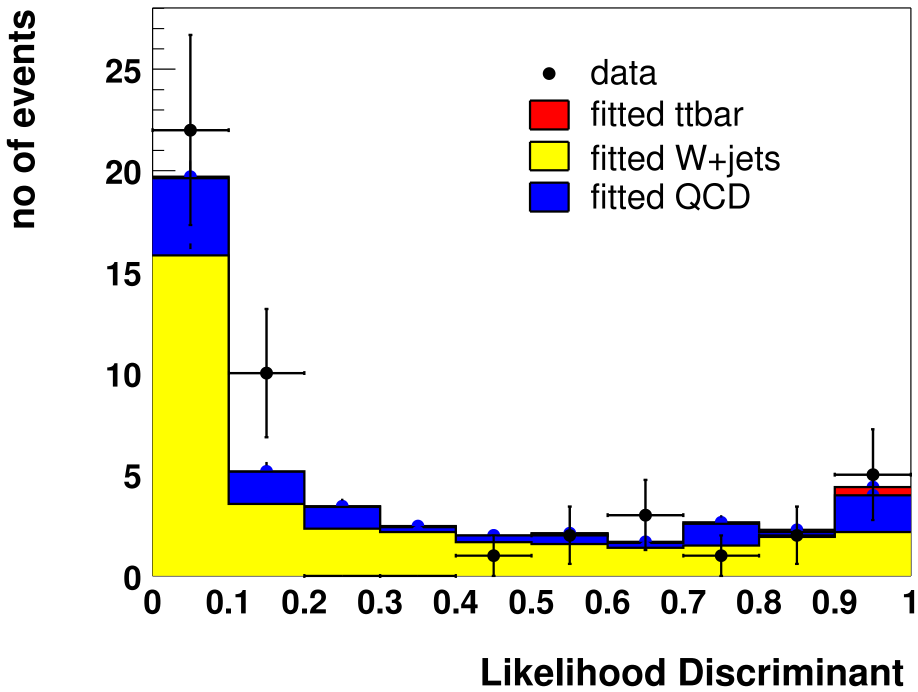

FIG. 42: The result of the topological likelihood fit showing data and fitted $t\bar{t}$, W+jets and QCD contributions in the EC.

TABLE XXV: Summary of the relative systematic uncertainties on the number of fitted $t\bar{t}$ events (EC).

source	relative systematic uncertainty
L1 EM trigger	+1.3%-0.5%
L2 EM trigger	
L3 EM trigger	+18%-16%
JES	22%
Jet ID	-100%
Jet resolution	+50.5% −46.0%
L1 jet trigger	+2.6% −4.3%
L2 jet trigger	
L3 jet trigger	+3.9% −23%
tune A	40%
top mass	2%
template statistics	100%
ε_{sig}	1.8%
ε_{QCD}	−69% +67%

APPENDIX A: ELECTRON+JETS CANDIDATE EVENTS

APPENDIX B: MUON+JETS CANDIDATE EVENTS

[1] L. Phaf and I. Iashvili, DØ Note 4194.
[2] M. Klute, DØNote 4185.
[3] The Top Physics Group of the DØ Collaboration, *DØ Top Analysis and Data Sample for the Winter Conferences 2004*, DØ Note xxxx, February 2004.
`http://www-d0.fnal.gov/Run2Physics/top/private/winter04/winter04_top_note_v0_2.ps`
[4] A. Garcia-Bellido, et al., DØ Note 4320, January 2004.
[5] J. Kozminski et al. *The electron likelihood in p14*, November 2003.
`http://www-d0.fnal.gov/Run2Physics/top/private/winter04/lh_note_v1_1.pdf`
[6] U. Bassler and G. Bernardi, DØ note 4124, March 2004.
[7] T. Golling et al., note in preparation).
[8] The Top Production Subgroup, *Measurement of the $t\bar{t}$ production cross section at $\sqrt{s} = 1.96$ TeV in dilepton final states*, DØ Note xxxx, February 2004.
`http://www-d0.fnal.gov/Run2Physics/top/private/winter04/winter04_dilepton_note_v0_0.ps`
[9] The Top Production Subgroup, note in preparation.