



Measurement of the relative fraction of $t\bar{t}$ events produced via gluon-gluon fusion

The CDF Collaboration
URL <http://www-cdf.fnal.gov>
(Dated: March 6, 2007)

We present a measurement of the relative fraction of $t\bar{t}$ events produced via gluon-fusion to the total number of $t\bar{t}$ events, which we call C_f^{true} , using a template method based on Neural Networks. Using a total integrated luminosity of 695 pb^{-1} we find $C_f^{true} < 0.23$ at 68% confidence level, and $C_f^{true} < 0.51$ at 95% confidence level.

Preliminary Results for Winter 2007 Conferences

| Sample | Acceptance(%) 1-tag | Acceptance(%) ≥ 2-tag |
|-----------------|------------------------|--------------------------|
| $t\bar{t}^{gg}$ | 2.36 ± 0.01 | 1.05 ± 0.01 |
| $t\bar{t}^{qq}$ | 1.83 ± 0.01 | 0.84 ± 0.01 |

TABLE I: Acceptance for gg and $q\bar{q}$ produced $t\bar{t}$ events for the 1 and 2-tag samples.

I. INTRODUCTION

This note describes a measurement of the fraction of $t\bar{t}$ events produced through gluon-fusion interactions (C_f^{true}) in $\bar{p}p$ collisions at $\sqrt{s} = 1.96$ TeV with the CDF detector at the Fermilab Tevatron. The Standard Model predicts that gluon-fusion produced $t\bar{t}$ events ($t\bar{t}^{gg}$) and quark annihilation produced $t\bar{t}$ events ($t\bar{t}^{qq}$) to occur with relative fractions of 15% and 85% respectively. However, this does have large uncertainties [1, 2]. For example, the contribution from gluon-fusion at the Tevatron can vary up to a factor of 2 (from 10-20%). This measurement will help us better understand the composition of the proton and may suggest new top production mechanisms.

To discriminate between $t\bar{t}^{gg}$ and $t\bar{t}^{qq}$, in part we use the fact that gg produced events tend to produce $t\bar{t}$ with unlike spin, while qq produced events tend to produce $t\bar{t}$ with like spin. The top quark, with a mass of about 175 GeV, has an expected lifetime of $\approx 0.5 \times 10^{-24}$ s; which is an order of magnitude faster than the typical hadronization time. As a consequence, the spin information carried by the top quark is preserved in the decay products, allowing the different production processes to retain their kinematic characteristics in the final state. This, along with information from kinematics of the production, is used to discriminate $t\bar{t}^{gg}$ from $t\bar{t}^{qq}$. See Section II B.

II. DATA SAMPLE & EVENT SELECTION

This analysis is based on an integrated luminosity of 695 pb^{-1} collected with the CDFII detector between March 2002 and September 2005. The data are collected with an inclusive lepton trigger that requires an electron or muon with $E_T > 18$ GeV ($P_T > 18$ GeV/c for the muon). From this inclusive lepton dataset we select events offline with a reconstructed isolated electron E_T (muon P_T) greater than 20 GeV, missing $E_T > 20$ GeV and at least 4 jets with $E_T > 15$ GeV and $|\eta| < 2$. To improve the sensitivity the sample is broken into two subsamples, one with 1 secondary vertex tag (b-tag) and one with 2 or more b-tags. After the tagging requirement we have a total of 167 events. The CDF detector is described in detail in [3].

A. Total $t\bar{t}$ Acceptance

The total acceptance is measured in Monte Carlo using the HERWIG Monte Carlo program [4]. The acceptance for $t\bar{t}^{gg}$ and $t\bar{t}^{qq}$ is shown in Table I, the uncertainties are due to Monte Carlo statistics.

B. Kinematic Variables

The top quarks are reconstructed using a kinematic fitter [5]. This provides us with the fully reconstructed kinematics of the $t\bar{t}$ system, Equation 1.

$$t\bar{t} \rightarrow W(\rightarrow ud) + b + W(\rightarrow l\nu) + b \quad (1)$$

The top mass is constrained to 175 GeV. All permutations of jet to partons, which are consistent to b-tag information, are tried. No requirement is made of the χ^2 of the fit.

The permutation with the lowest χ^2 is used to extract the kinematic variables of the event. A total of 8 variables, two of which describe the production and six of which describe the decay of a given $t\bar{t}$ event, are calculated.

The two kinematic variables we use that describe the production of the event are evaluated in the $t\bar{t}$ rest frame [6]:

- $\cos\theta^*$: the angle between the top quark and the right incoming parton.
- β : the top quark velocity relative to c .

| Sample | Signal Fraction (C_s) | b-tag Fraction | N_{data} | N_{bkg} |
|----------------|---------------------------|----------------|------------|----------------|
| 1-tag | 0.845 ± 0.025 | 0.725 | 121 | 19.9 ± 2.5 |
| ≥ 2 -tags | 0.959 ± 0.015 | 0.275 | 46 | 1.9 ± 0.5 |

TABLE II: Estimated signal fraction (C_s), b-tag fraction and number of events observed in data (N_{data}) and estimated background (N_{bkg}) for events with 1 and two or more b-tags.

The remaining six variables describe the decay and contain information about the spin correlations. These variables are angles with respect to the “off-diagonal” spin basis in the top (or anti-top) rest frame. This “off-diagonal” basis is defined using the zero momentum frame of the $t\bar{t}$ system, and it is designed such that the like spin components, $\uparrow\uparrow$ or $\downarrow\downarrow$, vanishes on average over large number of events. For this choice of spin basis the top pairs are, on average, in a state of unlike spins independent of their production angle and rest frame speed. For a further description of this basis, see [7]. The decay variables we use are as follows:

- $\cos\theta_{Lep}$: angle between lepton and the “off-diagonal” basis.
- $\cos\theta_\nu$: angle between neutrino and the “off-diagonal” basis.
- $\cos\theta_{Wlep}$: angle between leptonically decaying W and the “off-diagonal” basis.
- $\cos\theta_{Whad}$: angle between hadronically decaying W and the “off-diagonal” basis.
- $\cos\theta_{Down}$: angle between down quark and the “off-diagonal” basis.
- $\cos\theta_{Up}$: angle between up quark and the “off-diagonal” basis.

III. BACKGROUNDS

The backgrounds we consider for this analysis come from QCD production of W boson plus multijet events ($W + 4jets$). We use the Monte Carlo program ALPGEN [8] to generate $W + bb + qq$, $W + cc + qq$, $W + c + qqq$, and, $W + qqqq$. These are mixed in approximately equal proportion, and taken to represent the total background. The estimated number of background events can be seen in Table II.

IV. ANALYSIS METHOD

The measurement is performed over the set of events selected according to the selection criteria defined in Section II. This sample is composed mainly of three processes: $t\bar{t}^{gg}$, $t\bar{t}^{qq}$, and background from $W + 4jets$ (WJ).

The strategy of the measurement is based upon the different kinematic properties of these three processes. The kinematic properties we use are characterized by 8 event variables defined in Section II B. These variables are fed into a Neural Network (NN) which is trained on distinguishing gluon-fusion produced $t\bar{t}$ events from quark annihilation produced $t\bar{t}$ events. This NN is then used to make three templates (T^{gg} , T^{qq} , T^{WJ}) which are obtained by running over $t\bar{t}^{gg}$ events, $t\bar{t}^{qq}$ events, and WJ samples, Figure 1. Monte Carlo generated by HERWIG is used to train the NN and to construct T^{gg} and T^{qq} . Monte Carlo generated by ALPGEN is used to construct T^{WJ} , see Section III.

A likelihood, as a function of the gluon content C_f , is constructed based on the templates, Section IV A. This likelihood function is used to determine the most probable value of C_f (C_f^{fit}). To deal with possible negative values of the C_f^{fit} the Feldman-Cousins (FC) [9] statistical approach is used. The application of FC statistics to this analysis is discussed in Section IV B.

A. Likelihood

A likelihood as a function of C_f can be obtained for any given sample of events. Recall from Section II our sample has been split into two exclusive subsamples with 1 and 2 or more b-tagged events. The likelihood of the full sample of events is then simply calculated as the product of the likelihoods of the two exclusive samples:

$$\mathcal{L}(C_f) = \mathcal{L}^{1T}(C_f, C_s^{1T}) * \mathcal{L}^{2T}(C_f, C_s^{2T}) \quad (2)$$

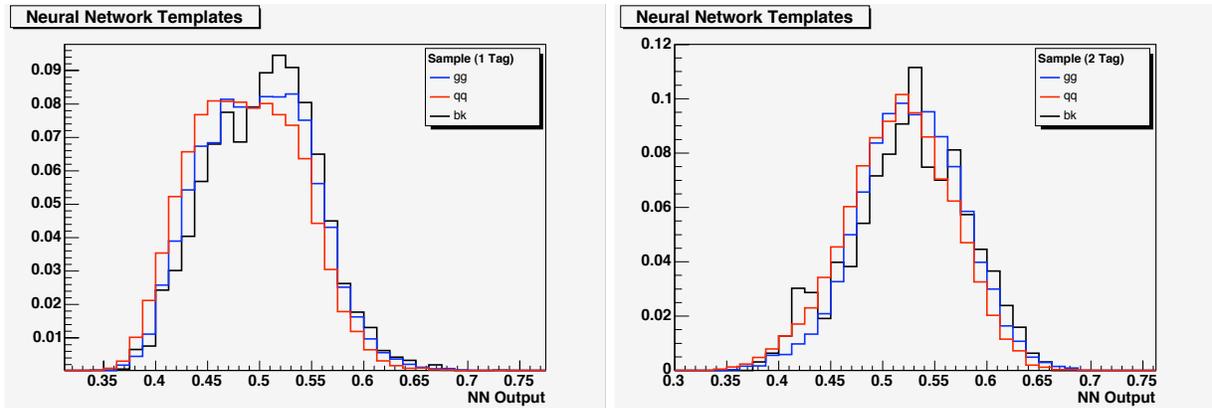


FIG. 1: Distributions of templates for 1-tag (left) and 2-tag (right) events for $q\bar{q}$ (red), gg (blue) and background (black).

where $\mathcal{L}^{NT}(C_f)$ represent the likelihood for the subsample of events with \mathcal{N} b-tags and C_s^{NT} , or signal fraction, is the relative fraction of $t\bar{t}$ events (both $t\bar{t}^{gg}$ and $t\bar{t}^{qq}$) to the total number of events for the given b-tag subsample.

The b-tag subsample likelihoods are defined as:

$$\mathcal{L}^{NT}(C_f, C_s^{NT}) = e^{-\frac{(C_s^{NT} - C_s^{NT})^2}{2.0 * \sigma_{C_s^{NT}}^2}} \prod \{C_s^{NT} [C_f T_{NT}^{gg} + (1 - C_f) T_{NT}^{qq}] + (1 - C_s^{NT}) T_{NT}^{WJ}\}$$

where the product is over the events with \mathcal{N} b-tags, and T_{NT}^{gg} , T_{NT}^{qq} and T_{NT}^{WJ} represent the template probability for the given event assuming its is $t\bar{t}^{gg}$, $t\bar{t}^{qq}$ or $W + 4jets$ background respectively. The values of C_s^{NT} can be found in Table II and $\sigma_{C_s^{NT}}$ are set to 0.1.

To evaluate the likelihood we scan over values of C_f between -1 and 2. For each value of C_f the likelihood is maximized by fitting the three fractions C_s^{NT} . In a given sample the C_f value for which the likelihood is maximum is called C_f^{fit} .

B. Feldman Cousins Method

The likelihood defined above is not constrained to a physically allowable range. However, the final result must have a value in closed interval from 0.0 to 1.0. We use the Feldman-Cousins Method (FC) to ensure a physical result.

FC has several properties that make it especially useful. A Frequentist approach, it ensures a result in the physically allowable region, and gives proper coverage for all values. It also provides a consistent way to quote either a measured value or an upper or lower limit.

In brief, FC produced a confidence band (FC band) based on a large number of simulated experiments, what we call pseudo-experiments(PE). We generate sets of PE varying C_f^{true} from 0.0 to 1.0. The PE are used to create a mapping of the fitted fraction, C_f^{fit} , to the actual value of the $t\bar{t}^{gg}$ fraction, C_f^{true} . For each set of PE the distribution of C_f^{fit} is gaussian, its width is use in FC to calculate the uncertainty. A full description of the Feldman-Cousins Method can be found here [9].

V. SYSTEMATIC UNCERTAINTIES

Systematic uncertainties in this analysis come several sources. We consider uncertainties on the jet energy scale, uncertainties in the parton distribution function (PDF), uncertainties in initial/final state radiation (ISR/FSR), uncertainties on the background shape and composition, and next to leading order correction (NLO). All systematics are derived from Monte Carlo.

To estimate the uncertainty of PDFs and jet energy scale, we compare different choices for PDFs and varying the jet energy scale in the Monte Carlo. Estimating the effect of ISR/FSR was done by increasing and decreasing the amount of ISR/FSR in $t\bar{t}$ events using PYTHIA [10]. The uncertainty due to the background composition, was estimated by

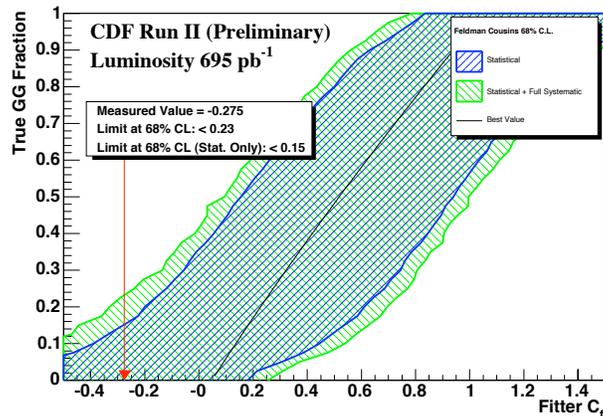


FIG. 2: FC with all systematics compared to FC with statistical only for a confidence level of 68%.

assuming 100% of the background is from one process (i.e. all $W + bb + qq$). We do this for all the processes we use to compose the background. To address the uncertainty in the background shape we vary the p_t of the W in our $W + jets$ samples. Finally, we consider the case where the background contains events without W bosons, multijet QCD. After all these calculations, we take the variation with the largest effect as our background systematic uncertainty.

To estimate the uncertainty due to next to leading order corrections, we use $t\bar{t}$ Monte Carlo generated with MC@NLO [11]. This is then used in place of the HERWIG sample. The variation is taken as a systematic uncertainty.

A. Systematics in Feldman-Cousins

Recall from Section IV B, FC uses the width of the pseudo-experiments (PE) to determine the uncertainty. To include systematics, PE are run with the systematically varied Monte Carlo. The difference in the mean of these PE from our default samples are added in quadrature and used to broaden C_f used to generate our final FC band.

Because the systematics uncertainties vary slightly with the value of C_f^{true} , it is difficult to quantify them in a table. However, you can see the total effect of all the sources of systematic uncertainties in Figure 2.

VI. RESULTS

Fitting the 167 events we have in data to the NN templates we derived from Monte Carlo we find:

$$C_f^{fit} = -0.275.$$

At 68% confidence level this corresponds, from Figure 2, to a purely statistical result of $C_f^{true} < 0.15$, and to a result of $C_f^{true} < 0.23$ when systematic uncertainties are taken into account.

Figure 3 shows the limits, including systematic uncertainties, for bands of 68% and 95% confidence level. At the 95% confidence level we find $C_f^{true} < 0.51$.

As a cross check you can see the fit in the NN output as well as the other individual variables that go into the NN with their components' fractions set from the NN fit in Appendix A.

VII. CONCLUSIONS

We present a measurement of the relative fraction of $t\bar{t}$ events produced via gluon-fusion to the total number of $t\bar{t}$ events using a template method based on Neural Networks probabilities. Using a total integrated luminosity of 695 pb^{-1} we find $C_f^{true} < 0.23$ at 68% confidence level.

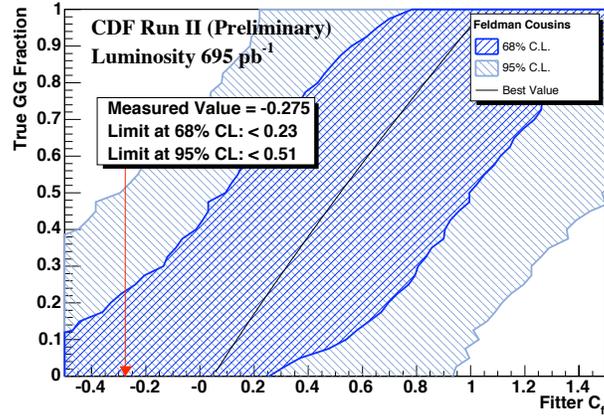


FIG. 3: Final FC with all systematics.

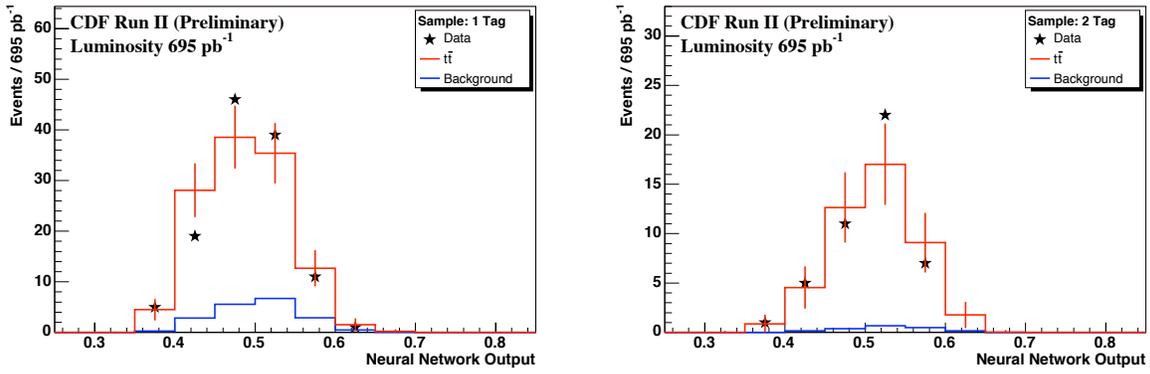


FIG. 4: Neural network templates fit to the Data (left 1 tag, right 2 tag).

APPENDIX A: APPENDIX: NEURAL NETWORK VARIABLES IN DATA

All the variable that go into the neural network can be seen below with the appropriate fitted fractions.

- Neural Network: Figure 4
- β : the top quark velocity relative to c . Figure 5
- $\cos\theta^*$: the angle between the top quark and the right incoming parton. Figure 6
- $\cos\theta_{lep}$: angle between lepton and the “off-diagonal” basis. Figure 7
- $\cos\theta_\nu$: angle between neutrino and the “off-diagonal” basis. Figure 8
- $\cos\theta_{up}$: angle between up quark and the “off-diagonal” basis. Figure 9
- $\cos\theta_{down}$: angle between down quark and the “off-diagonal” basis. Figure 10
- $\cos\theta_{Wlep}$: angle between leptonically decaying W and the “off-diagonal” basis. Figure 11
- $\cos\theta_{Whad}$: angle between hadronically decaying W and the “off-diagonal” basis. Figure 12

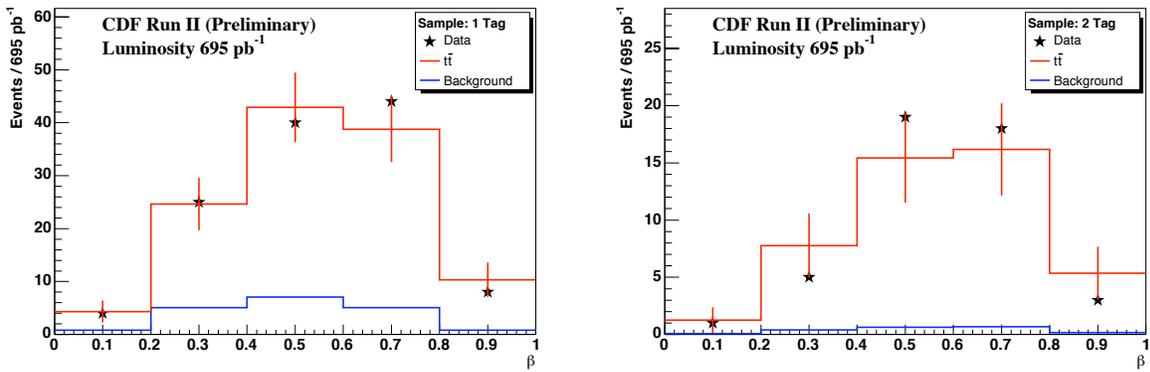


FIG. 5: β fractions set to the neural network fit to the Data (left 1 tag, right 2 tag).

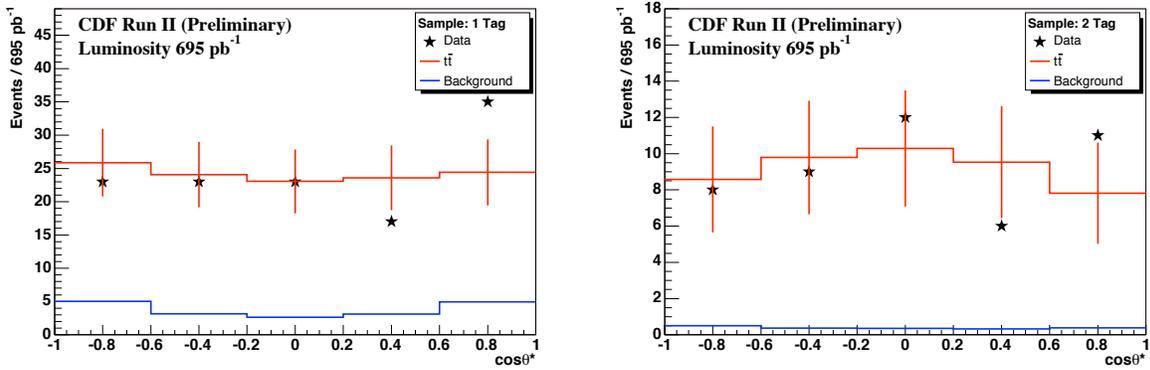


FIG. 6: $\cos\theta^*$ fractions set to the neural network fit to the Data (left 1 tag, right 2 tag).

ACKNOWLEDGMENTS

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium fuer Bildung und Forschung, Germany; the Korean Science and Engineering Foundation and the Korean Research Foundation; the Particle Physics and Astronomy Research Council and the Royal Society, UK; the Russian Foundation for Basic Research; the Comision Interministerial de Ciencia y Tecnologia, Spain; and in part by the European Community's Human Potential Programme under contract HPRN-CT-20002, Probe for New Physics.

-
- [1] M. Cacciari, *et al.*, “The t anti- t cross-section at 1.8-TeV and 1.96-TeV: A Study of the systematics due to parton densities and scale dependence”, hep-ph/0303085
[2] N. Kidonakis, *et al.*, “Next-to-next-to-leading order soft gluon corrections in top quark hadroproduction”, Phys. Rev. D **68** 114014 (2003).
[3] F. Abe, *et al.*, Nucl. Instrum. Methods Phys. Res. A **271**, 387 (1988); D. Amidei, *et al.*, Nucl. Instrum. Methods Phys. Res. A **350**, 73 (1994); F. Abe, *et al.*, Phys. Rev. D **52**, 4784 (1995); P. Azzi, *et al.*, Nucl. Instrum. Methods Phys. Res. A **360**, 137 (1995); The CDFII Detector Technical Design Report, Fermilab-Pub-96/390-E
[4] G. Corcella *et al.*, HERWIG 6: An Event Generator for Hadron Emission Reactions with Interfering Gluons (including

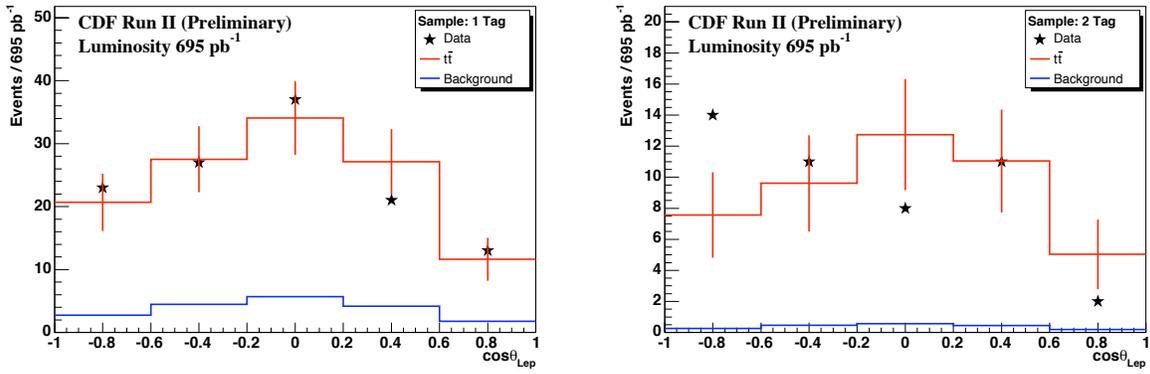


FIG. 7: $\cos\theta_{Lep}$ fractions set to the neural network fit to the Data (left 1 tag, right 2 tag).

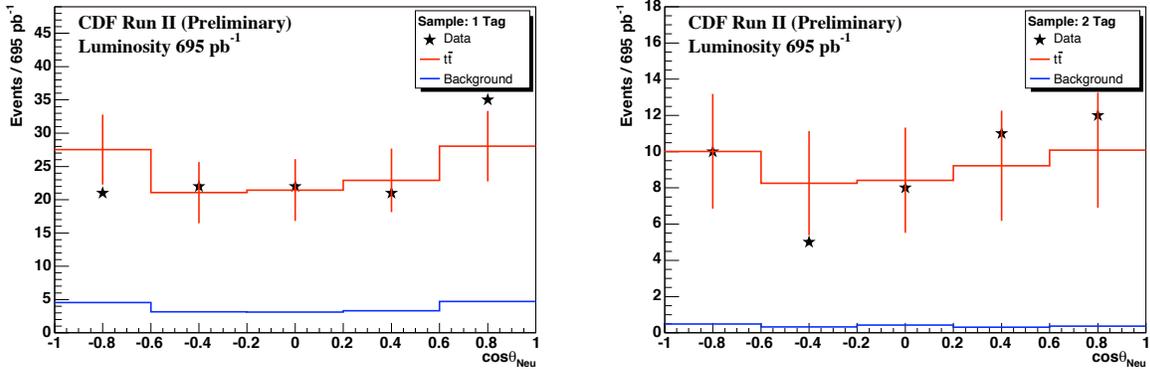


FIG. 8: $\cos\theta_{\nu}$ fractions set to the neural network fit to the Data (left 1 tag, right 2 tag).

supersymmetric processes), *JHEP* **01**, 10 (2001).

- [5] A. Abulencia, *et al* (CDF collaboration), *Phys. Rev.D* **73**,032003 (2006).
- [6] G. Mahlon, S. Parke, *Phys. Rev. D* **53**, 4886 (1996)
- [7] S. Parke and Y. Shadmi, *Phys Lett.* **B387**, 199 (1996); G. Mahlon and S. Parke, *Phys Lett.* **B411**, 173 (1997); G. Mahlon, hep-ph/9811281
- [8] M.L. Mangano, M. Moretti, F. Piccinini, R. Pittau, A. Polosa, ALPGEN, a generator for hard multiparton processes in hadronic collisions, *JHEP* **0307**, 001 (2003).
- [9] G. Feldman and R. Cousins, *Phys. Rev. D* **57**, 3873 (1998)
- [10] T. Sjostrand *et al.*, High-Energy-Physics Event Generation with PYTHIA 6.1, *Comput. Phys. Commun.* **135**, 238 (2001).
- [11] S. Frixione, P. Nason and B.R. Webber, Matching NLO QCD and parton showers in heavy flavour production, *JHEP* **0308** 007 (2003).

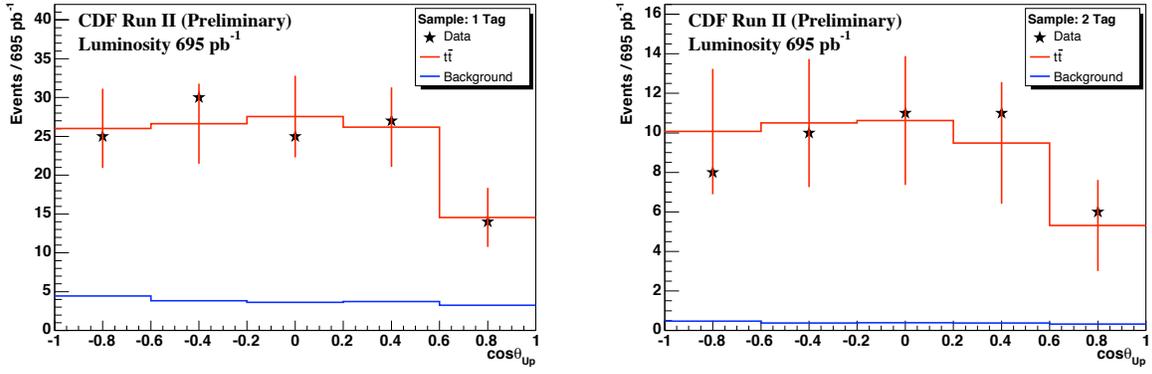


FIG. 9: $\cos\theta_{Up}$ fractions set to the neural network fit to the Data (left 1 tag, right 2 tag).

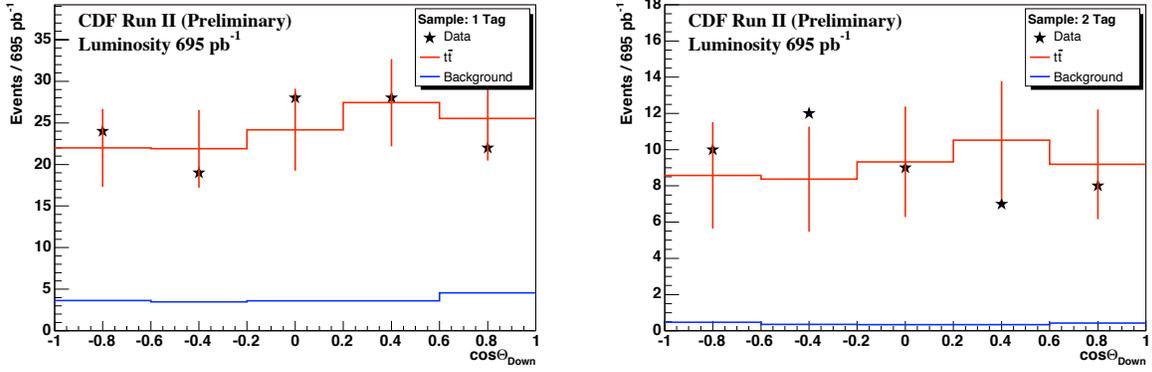


FIG. 10: $\cos\theta_{Down}$ fractions set to the neural network fit to the Data (left 1 tag, right 2 tag).

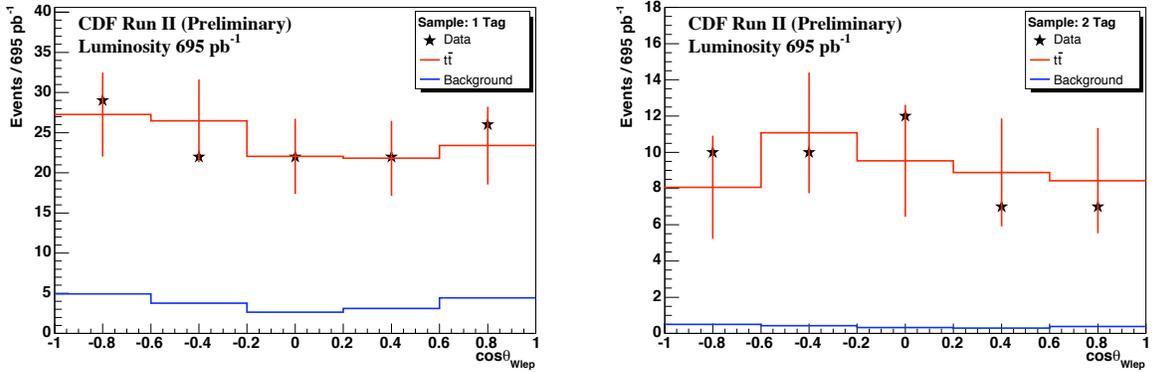


FIG. 11: $\cos\theta_{Wlep}$ fractions set to the neural network fit to the Data (left 1 tag, right 2 tag).

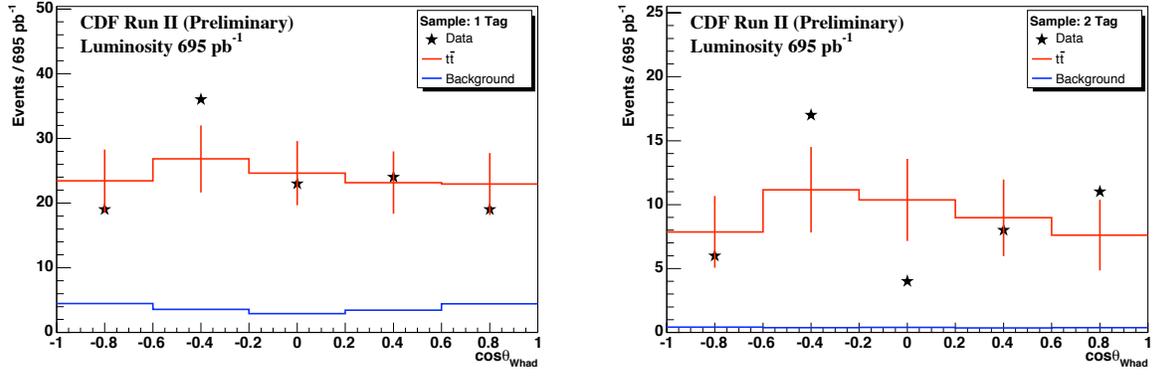


FIG. 12: $\cos\theta_{Whad}$ fractions set to the neural network fit to the Data (left 1 tag, right 2 tag).