CDF Note 9716

Search for Electroweak Single Top-Quark Production using Neural Networks with 3.2 fb^{-1} of CDF II data

The CDF Collaboration
URL http://www-cdf.fnal.gov
(Dated: March 5, 2009)

We report on a search for electroweak single top–quark production with CDF II data corresponding to 3.2 fb^{-1} of integrated luminosity. We apply neural networks to construct discriminants that distinguish between single–top–quark and background events. Two analyses are performed, assuming a top–quark mass of 175 GeV/c^2. In the first one, we combine t– and s–channel events to one single–top signal under the assumption that the ratio of the two processes is given by the standard model. A binned likelihood fit to the data in 8 distinct channels yields a cross section of $1.8^{+0.6}_{-0.6}$ pb for single top–quark production. Using ensemble tests we determine that we expect with a probability of 50% to see a single-top signal that is larger than a 5.2 σ fluctuation of the background (p–value of 0.00000011). The observed p–value is 0.00024014 which corresponds to a significance of 3.5 σ.

In the second analysis, we separate the two single top–quark production modes. A binned likelihood fit done simultanously to 2 two–dimensional and 6 one–dimensional distributions of neural network outputs yields most probable values for the cross sections of $0.7^{+0.5}_{-0.5}$ pb for the t–channel and $2.0^{+0.7}_{-0.6}$ pb for the s–channel.

Preliminary Results for Winter 2009 Conferences

I. INTRODUCTION

According to the standard model, in $p\bar{p}$ collisions at the Tevatron top quarks can be created in pairs via the strong force, of singly via the electroweak interaction. The latter production mode is referred to as "single top–quark" production and takes place mainly through the s- or t- channel exchange of a W boson. The CDF and DØ collaborations have published single top–quark results at $\sqrt{s} = 1.8$ TeV and $\sqrt{s} = 1.96$ TeV [1, 2]. The most recent results from the DØ collaboration [3, 4] has seen evidence for single top quark production and measured $\sigma_{\text{single top}} = 4.7 \pm 1.3$ pb. In the meantime the CDF collaboration has also seen evidence by measuring $\sigma_{\text{single top}} = 2.2 \pm 0.7$ pb [5].
The theoretical single top–quark production cross section is $\sigma_{\text{single top}} = 2.9 \pm 0.4$ pb for a top–quark mass of 175 GeV/c^2 [6]. Despite this small rate, the main obstacle in finding single top–quarks is in fact the large associated background. After all section requirements are imposed, the signal to background ratio is approximately 1/20. This challenging, background-dominated dataset is the main motivation for using multivariate techniques.

II. COMMON EVENT SELECTION

The CDF event selection exploits the kinematic features of the signal final state, which contains a top quark, a bottom quark, and possibly additional light quark jets. To reduce multijet backgrounds, the W originating from the top quark is required to have decayed leptonically. One therefore demands a single high-energy electron or muon ($E_T(e) > 20$ GeV, or $P_T(\mu) > 20$ GeV/c) and large missing transverse energy $\not\!\!E_T > 25$ GeV from the undetected neutrino.
The backgrounds belong to the following categories: $Wb\bar{b}$, $Wc\bar{c}$, Wc, $Wq\bar{q}$ (light quarks misidentified as heavy flavor jets), top pair production $t\bar{t}$ events (one lepton or two jets are lost due to detector acceptance), QCD (multijet events where a jet is erroneously identified as a lepton), $Z \to ll$ and diboson WW, WZ, and ZZ. We remove a large fraction of the backgrounds by demanding two or three jets with $E_T > 20$ GeV and $|\eta| < 2.8$ be present in the event. At least one of these jets has to be tagged as a b quark jet by using displaced vertex information from the silicon vertex detector (SVX) of CDF [7]. The QCD multijet content of the selected dataset is further reduced by several requirements to transverse mass of the W-boson candidate, the missing transverse energy significance, the angle between the $\not\!\!E_T$ vector and the transverse momentum vector and the angle between the charged lepton and the momentum vector of the jets.

III. NEURAL NETWORK INPUT VARIABLES

Using neural networks, many kinematic or event shape variables are combined to a powerful discriminant. In the search for single top–quark production, four different neural networks are trained in different jet and tag bins:

- 2 jets, 1 tag
- 2 jets, 2 tags
- 3 jets, 1 tag
- 3 jets, 2 tags

We devide each of the four categories into two separate channels, one containing triggered electrons and muons called Triggered Lepton Coverage (TLC), and the other containing muons from an Extended Muon Coverage (EMC) accepted through the MET + 2 jets trigger. For the separate search we include an additional network in the 2 jet 1 tag category to build a 2D discriminant. This improves the apriori sensitivity for s-channel of about 15%. The TLC 2 jet, 1 tag sample is the largest sub sample and dominates so the search for single top–quarks. The four most important variables of the the NN in the TLC 2 jet, 1tag sample are the reconstructed top quark mass, the KIT flavor separator, the invariant mass of the two jets and the product of the lepton-charge and the pseudorapidity of the light quark, see figure 1. For each variable the signal and background shapes, a data Monte-Carlo comparison and a check of the background shape in the zero-tag sample are shown in figure 1. The KIT flavor separator gives an additional handle to reduce the large background components where no real b quarks are contained, namely $Wq\bar{q}$ and charm-backgrounds. Both of them amount to about 50% in the W+2 jets data sample even after imposing the requirement that one jet is identified by the secondary vertex tagger of CDF [7].

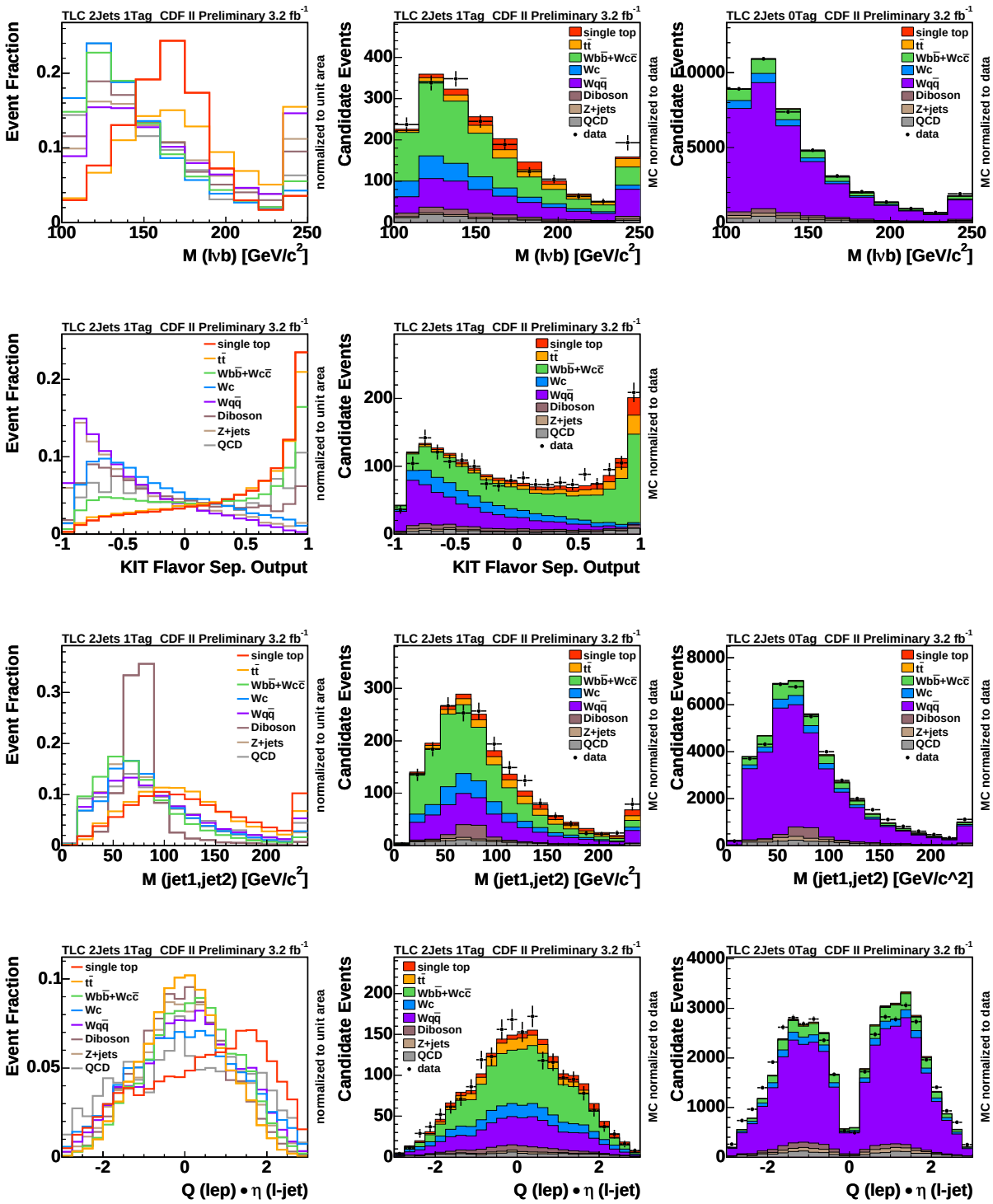


FIG. 1: Some of the most powerful variables in the TLC 2jet, 1tag channel. Left: signal and background shapes, Middle: Data-Monte Carlo comparison, Right: Check of background shapes in the W+2jet zero tag sample.

IV. TEMPLATES FOR BINNED LIKELIHOOD FIT

A. Templates for Combined Search

The training of a neural network results in one output variable continuously distributed between −1 and 1. The output of the neural networks in the different 8 channels is used to create signal and background templates which are to be fitted to the output distribution of observed events. We perform a combined single top–quark search meaning that the output distributions of both t– and s–channel events are combined into one single distribution, where the ratio between the two processes is as predicted by the standard model. In the fit, all 8 considered channels are fitted simultaneously to determine the combined single top–quark cross section. Figure 2 and 3 show the templates of the combined single top–quark search.

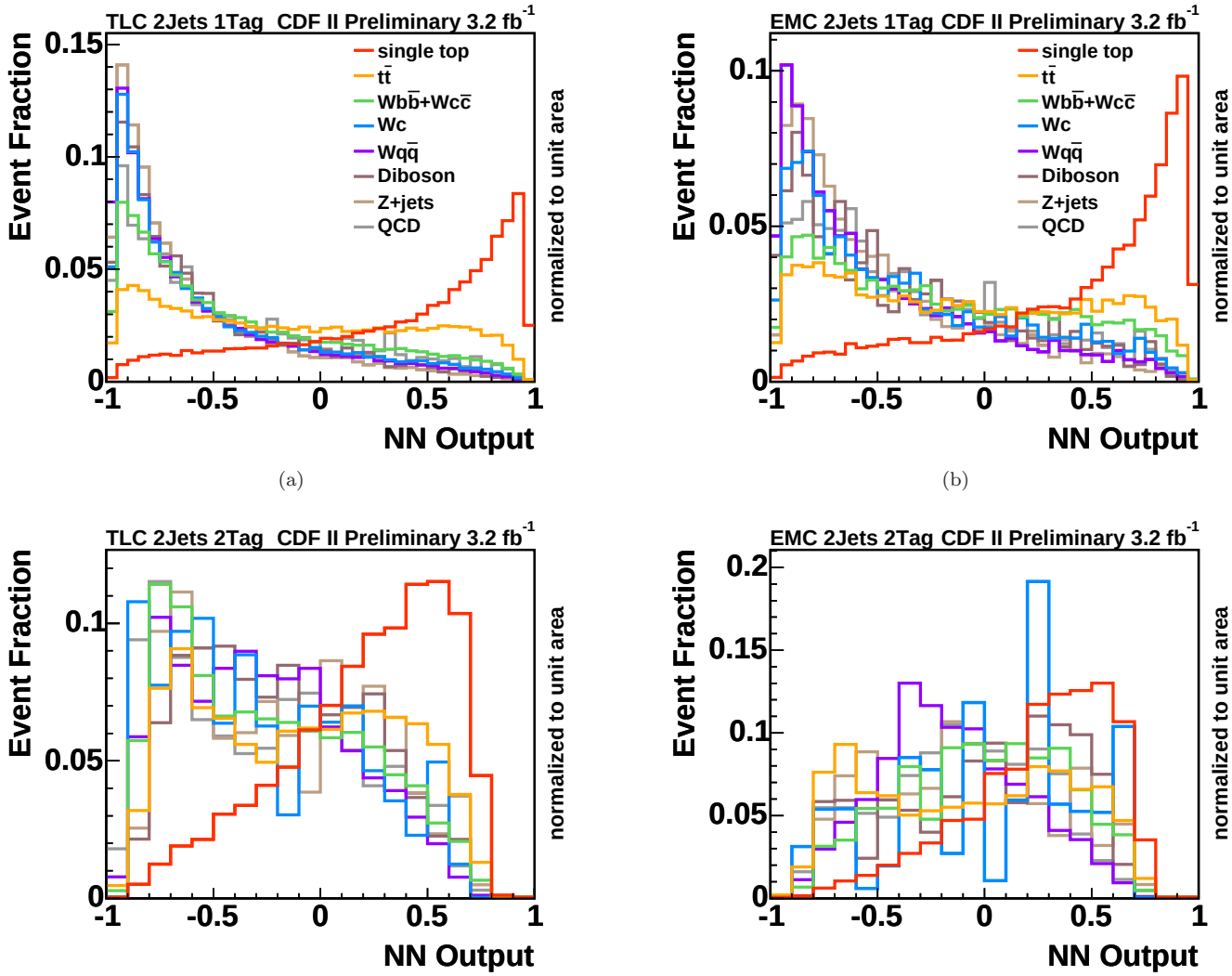
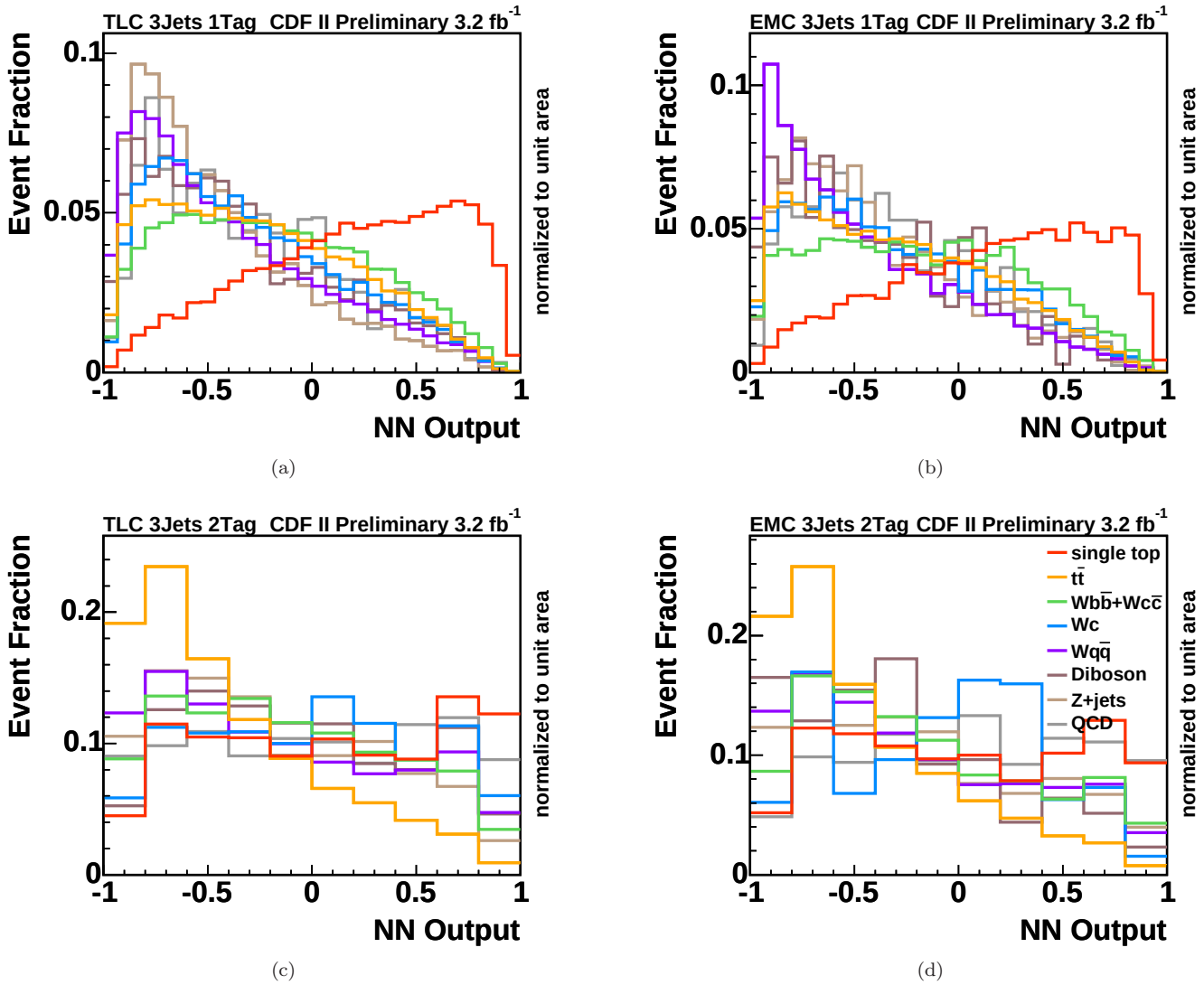


FIG. 2: The templates for the t–channel neural network in the 2–jet channel with 1 b–tag (top) and the s–channel neural network in the 2–jet channel with 2 b–tags (bottom) for TLC (left) and EMC (right) leptons.



(a) (b) (c) (d)

FIG. 3: The templates for the t–channel neural network in the 3–jet channel with 1 b–tag (top) and with 2 b–tags (bottom) for TLC (left) and EMC (right) leptons.

B. Templates for Separate Search

The templates of the separate search in the 2 jet bin with 1 b–tag are illustrated in figures 4–7 showing the output of the s–channel neural network versus the output of the t–channel neural network. For the final fit to data the 2D templates get unwinded bin by bin to have 1D distributions. The final templates of the separate search in the 2–jet bin with 1 and 2 b–tags and 3–jet bin with 1 and 2 b–tags, respectively are illustrated in figure 8 and 9. In the fit, all considered bins are fitted simultaneously to determine the t– and s–channel cross sections.

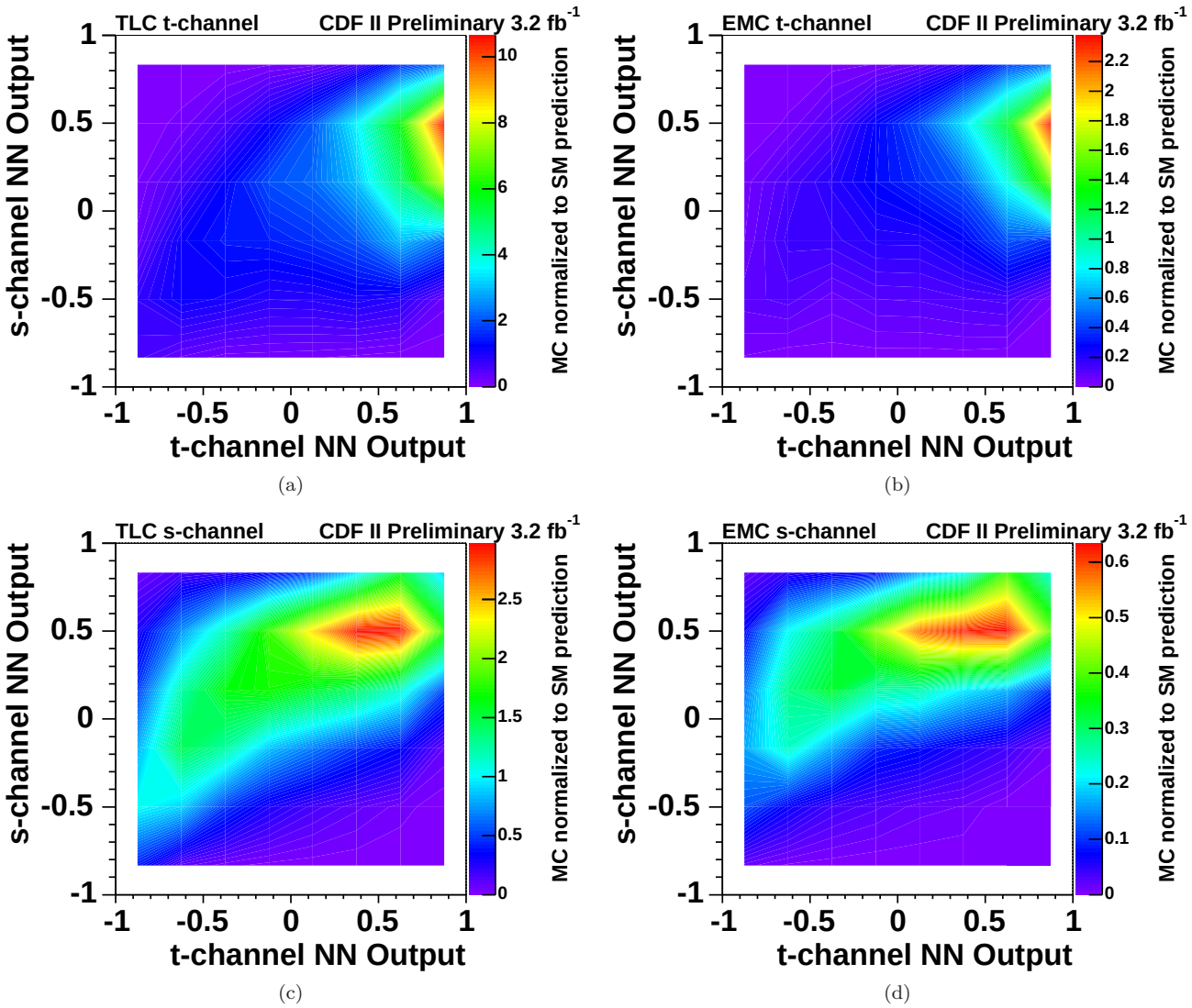


FIG. 4: The templates utilized in the separate search are illustrated. The output of the s–channel network in dependence of the output of the t–channel network is shown for t–channel and s–channel events for TLC (left) and EMC (right) leptons.

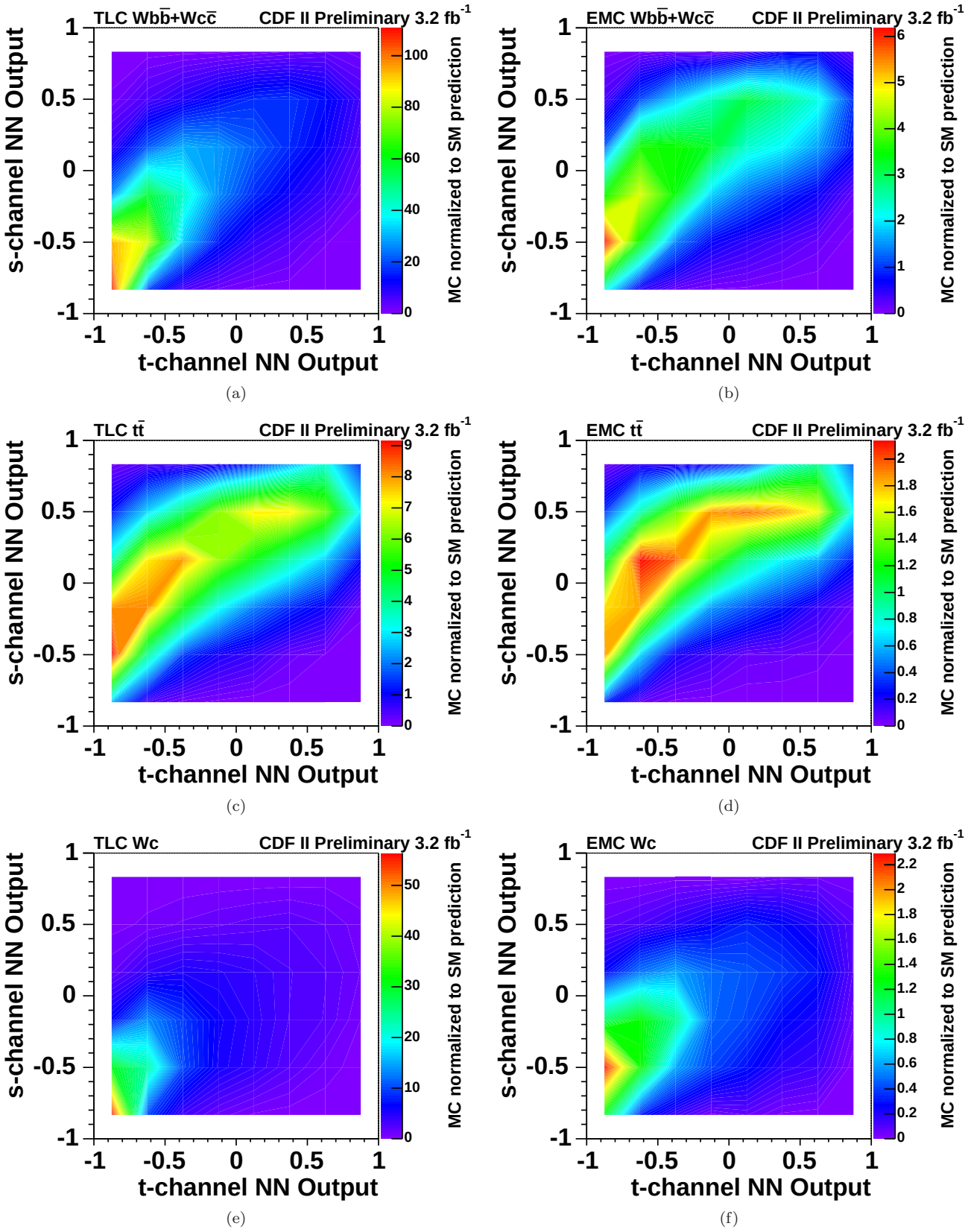


FIG. 5: The templates utilized in the separate search are illustrated. The output of the s–channel network in dependence of the output of the t–channel network is shown for $Wb\bar{b} + Wc\bar{c}$ events, $t\bar{t}$ events, and Wc events for TLC (left) and EMC (right) leptons.

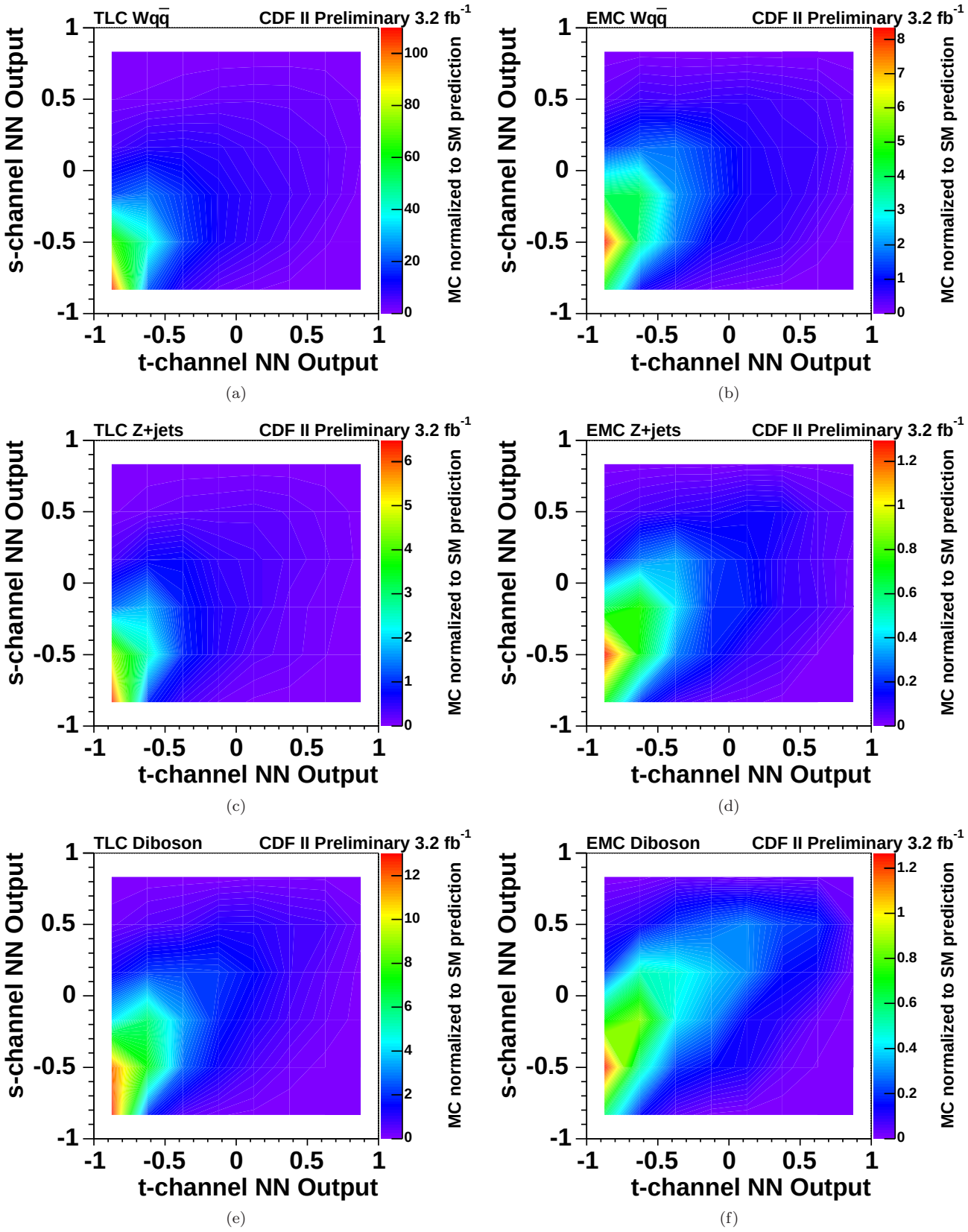
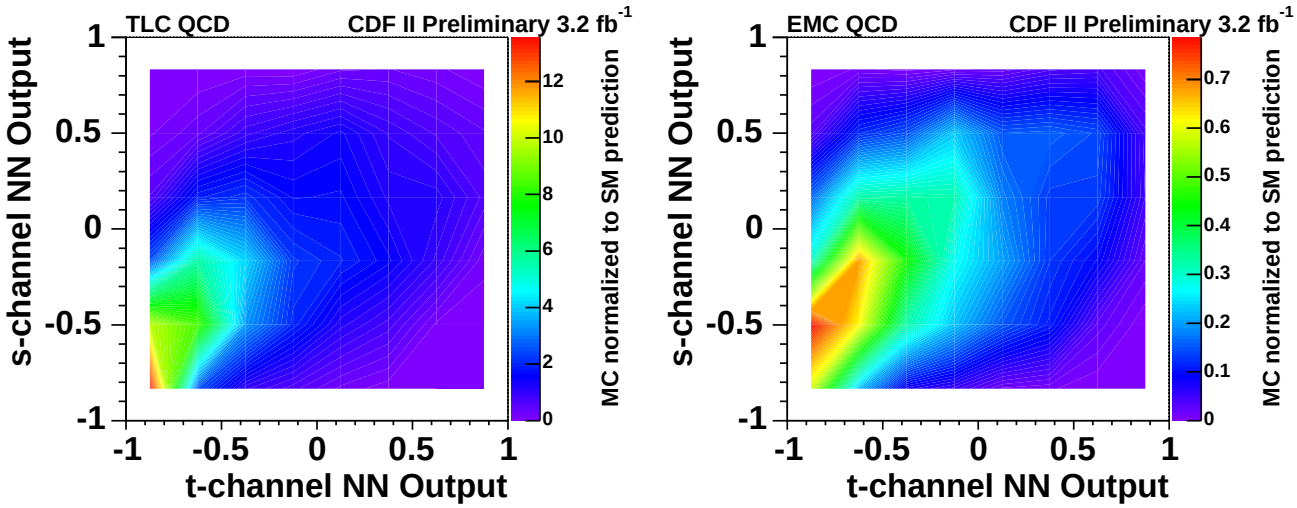
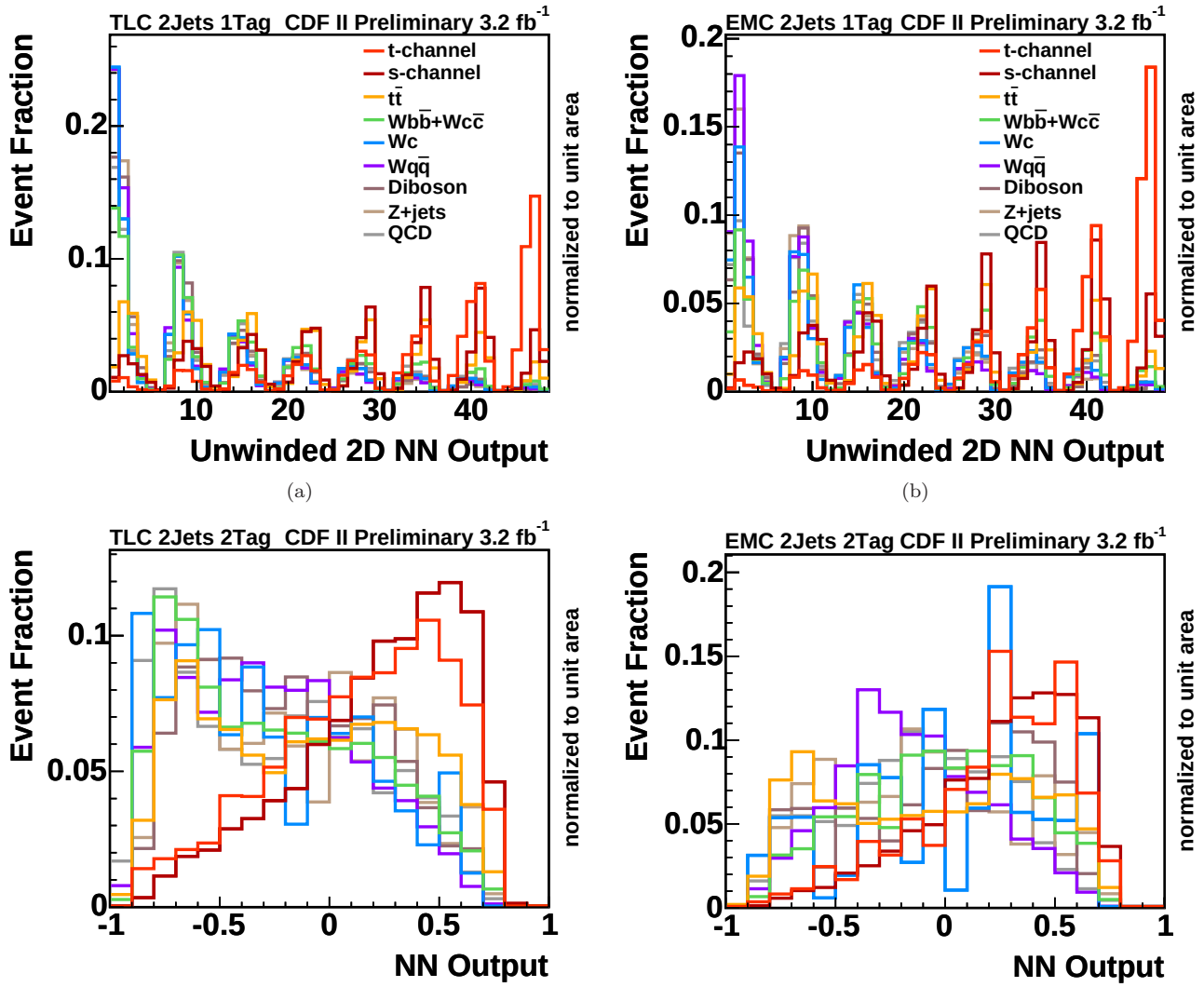


FIG. 6: The templates utilized in the separate search are illustrated. The output of the s–channel network in dependence of the output of the t–channel network is shown for $Wq\bar{q}$ events, Z+jets events, and Diboson events for TLC (left) and EMC (right) leptons.



(a) (b)

FIG. 7: The templates utilized in the separate search are illustrated. The output of the s–channel network in dependence of the output of the t–channel network is shown for QCD multijet events for TLC (left) and EMC (right) leptons.



(a) (b) (c) (d)

FIG. 8: The templates for the separate search: the unwinded 2D NN output of the s– and t–channel neural networks in the 2–jet bin with 1 b–tag (top) and the NN output of the s–channel neural network in the 2–jet bin with 2 b–tags (bottom) for TLC (left) and EMC (right) leptons.

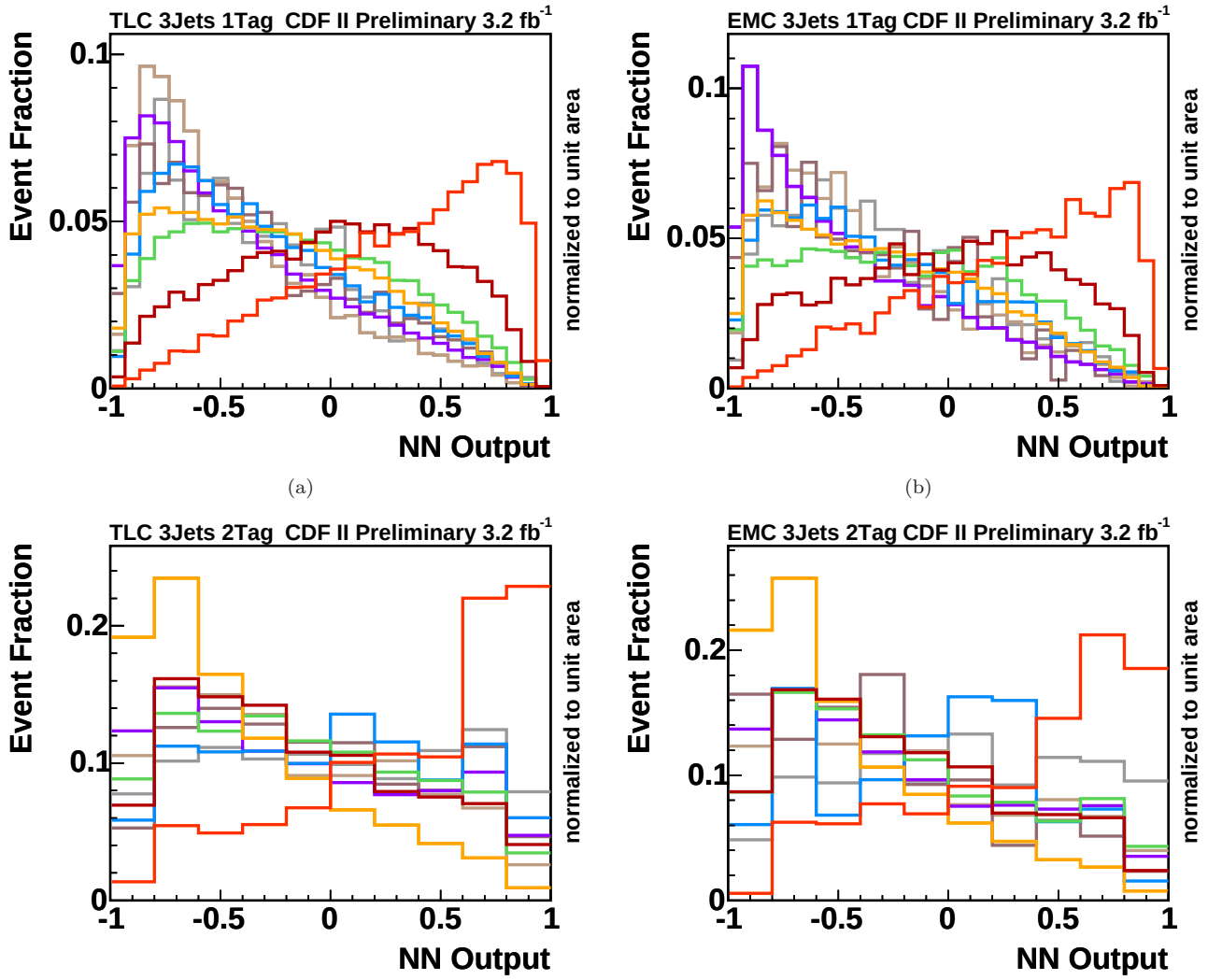


FIG. 9: The templates for the separate search: the NN output of the t–channel neural networks in the 3–jet bin with 1 b–tag (top) and the 3–jet bin with 2 b–tags (bottom) for TLC (left) and EMC (right) leptons.

V. SYSTEMATIC UNCERTAINTIES

Uncertainties in the modeling of physics processes and detector effects cause systematic uncertainties on the measurement results, affecting the rate of predicted signal and background events as well as the shape of the template histograms used in the fit to the observed data distribution. It must be noted that some effects induce only rate uncertainties, while some affect only the shape of the templates, even though most of them impact both.

The following sources of systematic uncertainties are considered: the uncertainty on the jet energy corrections, the uncertainty in modeling initial and final state gluon radiation (IFSR), the choice of the parameterization of the parton distribution functions (PDF) used for the event simulation, the choice of the Monte Carlo event generator, the uncertainty in the event detection efficiency, the uncertainty in modeling the output of the KIT flavor separator, the uncertainty in the factorization and renormalization scale (Q^2) for the simulation of W+heavy flavor processes, the modeling of instrumental backgrounds, that is mistagged $Wq\bar{q}$ events and QCD multijet events, the uncertainty in the luminosity determination, and the uncertainty on the modeling of the distributions of $\Delta R_{j_1,j_2}$, η_{j2} in the 2 jet bin and η_{j3} in the 3 jet bin in the pretag sample. The impact of these sources of uncertainties is evaluated by altering the modeling of the corresponding processes or effects within their uncertainties or by assigning a plausible alternative model. As a result, relative changes of the event rates and shifted template distributions are obtained. Most rate uncertainties are only determined for single–top–quark and $t\bar{t}$ events because the rates of the main backgrounds, W+jets and QCD events, are estimated based on the observed rate of events before b tagging or by a fit to the observed missing transverse energy distribution, respectively. The analyses are done under the assumption of a top–quark mass of $M_t = 175\,\mathrm{GeV}/c^2$. That is why the uncertainty in the top–quark mass is not taken into account as a systematic uncertainty when fitting for the single top–quark production cross section. Hence, the analyses provide rather a measurement at the specified value of the top–quark mass. However, for the extraction of the expected and observed significance, the rate and shape uncertainty on the top–quark mass is fully taken into account. In Figure 10 you find three examples of systematic shape uncertainties in the TLC 2 jet 1 tag channel: jet energy scale (JES) for the single top-quark template, factorization and renormalization scale (Q^2) for Wbb events, and modeling uncertainty on the KIT flavor separator output (KIT opt.). Tables I-V summarize the relative rate uncertainties on the event prediction of the various Monte Carlo samples for the TLC lepton channels.

Source	t-channel	s-channel	single-top	$t\bar{t}$
IFSR less/more	7.0/-1.5 %	6.2/7.1 %	6.7/1.4 %	-7.7/-9.7 %
PDF	3.1/-3.5 %	1.7/-1.4 %	2.6/-2.8 %	1.9/-2.3 %
MC	2.0/-2.0 %	1.0/-1.0 %	1.7/-1.7 %	-2.7/2.7 %
ϵ_{evt}	4.2/-4.2 %	2.3/-2.3 %	3.6/-3.6 %	2.9/-2.9 %
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %
Cross section	12.6/-12.6 %	12.4/-12.4 %	12.6/-12.6 %	12.4/-12.4 %
M_{top} 170/180	6.1/-5.3 %	9.5/-8.0 %	7.3/-6.2 %	7.8/-8.1 %
	Diboson	Z+jets		
ϵ_{evt}	7.6/-7.6 %	8.3/-8.3 %		
Luminosity	6.0/-6.0 %	6.0/-6.0 %		
Cross section	1.9/-1.9 %	10.8/-10.8 %		

TABLE I: Systematic rate uncertainties for TLC lepton events with 2 jets and 1 b tag

Source	t-channel	s-channel	single-top	$t\bar{t}$
IFSR less/more	-0.9/-13.4 %	9.5/11.4 %	7.9/7.5 %	-7.5/-11.2 %
PDF	3.0/-3.3 %	1.7/-1.5 %	1.9/-1.7 %	1.9/-2.3 %
MC	2.0/-2.0 %	1.0/-1.0 %	1.2/-1.2 %	4.6/-4.6 %
$\epsilon_{\rm evt}$	10.0/-10.0 %	8.7/-8.7 %	8.9/-8.9 %	9.0/-9.0 %
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %
Cross section	12.6/-12.6 %	12.4/-12.4 %	12.5/-12.5 %	12.4/-12.4 %
$M_{\rm top}$ 170/180	2.5/-7.3 %	9.4/-6.9 %	7.7/-6.7 %	9.9/-7.1 %
	Diboson	Z+jets	Mistags	
$\epsilon_{\rm evt}$	9.8/-9.8 %	10.6/-10.6 %		
Luminosity	6.0/-6.0 %	6.0/-6.0 %		
Double tag			22.0/-22.0%	
Cross section	1.9/-1.9 %	10.8/-10.8 %		

TABLE II: Systematic rate uncertainties for TLC lepton events with 2 jets and 2 b tags

Source	t-channel	s-channel	single-top	$t\bar{t}$
IFSR less/more	-8.4/-3.1 %	-3.6/-17.4 %	-6.5/-8.6 %	-4.0/-6.7 %
PDF	3.2/-3.7 %	1.8/-1.5 %	2.7/-2.8 %	1.9/-2.3 %
MC	1.9/-1.9 %	1.5/-1.5 %	1.7/-1.7 %	-1.7/1.7 %
$\epsilon_{\rm evt}$	3.5/-3.5 %	2.3/-2.3 %	3.0/-3.0 %	2.3/-2.3 %
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %
Cross section	12.6/-12.6 %	12.4/-12.4 %	12.6/-12.6 %	12.4/-12.4 %
$M_{\rm top}$ 170/180	6.2/-6.5 %	11.7/-8.6 %	6.4/-5.6 %	9.3/-8.4 %
	Diboson	Z+jets		
$\epsilon_{\rm evt}$	7.8/-7.8 %	7.8/-7.8 %		
Luminosity	6.0/-6.0 %	6.0/-6.0 %		
Cross section	1.9/-1.9 %	10.8/-10.8 %		

TABLE III: Systematic rate uncertainties for TLC lepton events with 3 jets and 1 b tag

Source	t-channel	s-channel	single-top	$t\bar{t}$
IFSR less/more	22.7/4.4 %	-3.1/-16.2 %	8.2/-7.2 %	-3.9/-9.3 %
PDF	3.7/-4.1 %	1.8/-1.5 %	2.6/-2.6 %	1.9/-2.3 %
MC	1.9/-1.9 %	1.5/-1.5 %	1.7/-1.7 %	2.0/-2.0 %
$\epsilon_{\rm evt}$	9.1/-9.1 %	8.8/-8.8 %	8.9/-8.9 %	9.1/-9.1 %
Luminosity	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %	6.0/-6.0 %
Cross section	12.6/-12.6 %	12.4/-12.4 %	12.5/-12.5 %	12.4/-12.4 %
$M_{\rm top}$ 170/180	7.8/-3.0 %	9.1/-11.1 %	6.9/-6.0 %	9.3/-9.5 %
	Diboson	Z+jets	Mistags	
$\epsilon_{\rm evt}$	10.8/-10.8%	11.1/-11.1 %		
Luminosity	6.0/-6.0 %	6.0/-6.0 %		
Double tag			22.0/-22.0%	
Cross section	1.9/-1.9 %	10.8/-10.8 %		

TABLE IV: Systematic rate uncertainties for TLC lepton events with 3 jets and 2 b tags

process	2jets 1tag	2jets 2tags	3jets 1tag	3jets 2tags
t-ch	-1.1/0.6 %	4.8/-3.5 %	-10.4/10.6 %	-5.7/4.3 %
s-ch	-0.1/-0.6 %	1.2/-1.9 %	-8.3/9.4%	-7.2/7.4 %
single-top	-0.8/0.2 %	1.8/-2.2 %	-9.1/9.9 %	-6.6/6.1 %
$t\bar{t}$	9.8/-9.4 %	8.1/-7.5 %	4.6/-5.1 %	5.4/-5.2 %
$Wc\bar{c}$+$Wb\bar{b}$	7.0/-6.9 %	10.8/-10.6 %	8.4/-7.7 %	11.0/-12.1 %
Wc	7.0/-6.3 %	11.3/-10.3 %	8.2/-6.9 %	13.9/-15.8 %
Z+jets	-5.3/5.4 %	5.0/-5.0 %	-10.8/14.0 %	-5.9/7.2 %
Diboson	-2.7/1.7 %	-3.0/1.5 %	-12.4/11.9 %	-12.0/12.0 %

TABLE V: Systematic JES down/up rate uncertainties for TLC lepton events.

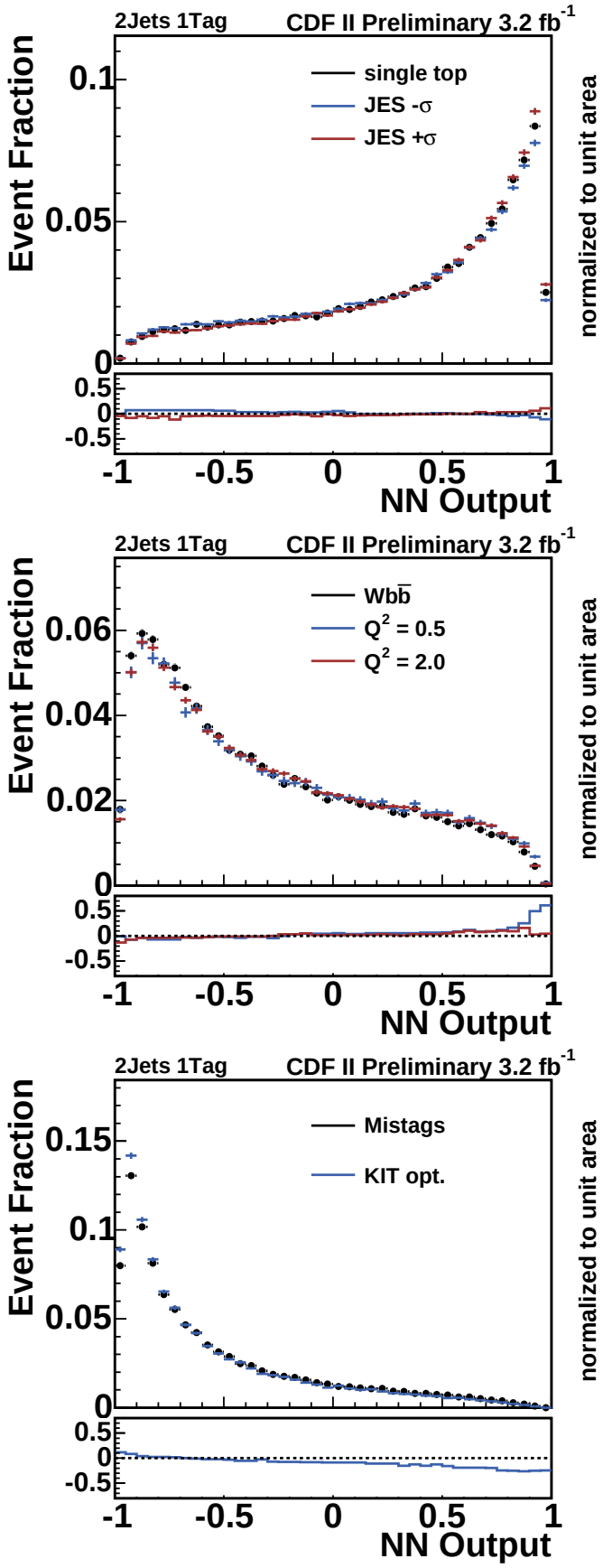


FIG. 10: Examples of shape uncertainties. Shown is the comparison between the default distribution and the shifted distribution. Top: Shape systematics due the uncertainty on the jet energy correction (JES) for single top–quark production. Center: Shape systematics due the uncertainty on the renormalization scale (Q^2) for $Wb\bar{b}$. Bottom: Shape systematics due to the uncertainty on the output of the KIT flavor separator for $Wq\bar{q}$ events.

VI. EXPECTED SIGNIFICANCE FOR COMBINED SEARCH

We use ensemble tests to compute the sensitivity of our analysis. An ensemble test consists of a set of pseudo experiments. For each pseudo experiment we determine first the number of events N_j of each process by drawing a random number from a Poisson distribution with a mean μ_j. In a second step we draw random numbers from the template distributions of the neural network output.

To compute the significance of a potentially observed signal, we perform a hypothesis test. Two hypotheses are considered. The first one, H_0, assumes that the single top–quark cross section is zero ($\beta_1 = 0$) and is called the *null hypothesis*. The second hypothesis, H_1, assumes that the single top–quark production cross section is the one predicted by the standard model ($\beta_1 = 1$). The objective of our analysis is to observe single top–quarks, that means to reject the null hypothesis H_0.

The hypothesis test is based on the Q-value,

$$Q = -2\left(\ln L_{\rm red}(\beta_1 = 1) - \ln L_{\rm red}(\beta_1 = 0)\right) , \tag{1}$$

where $L_{\rm red}(\beta_1 = 1)$ is the value of the reduced likelihood function at the standard model prediction and $L_{\rm red}(\beta_1 = 0)$ is the value of the reduced likelihood function for a single top–quark cross section of zero. Using the two ensemble tests the distribution of Q-values is determined for the case with single top–quarks included at the standard model rate, q_1, and for the case of zero single top–quark cross section, q_0. The two Q-value distributions are shown in figure 11.

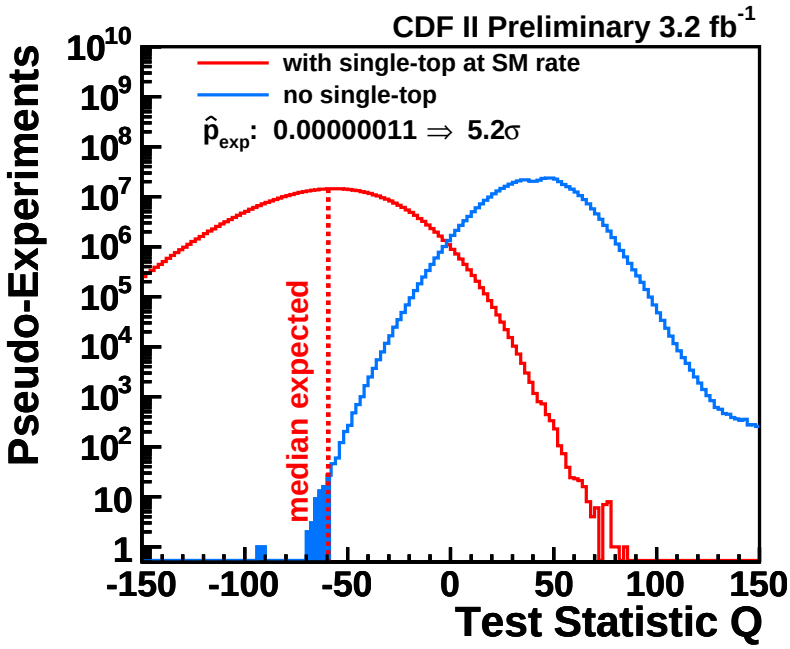
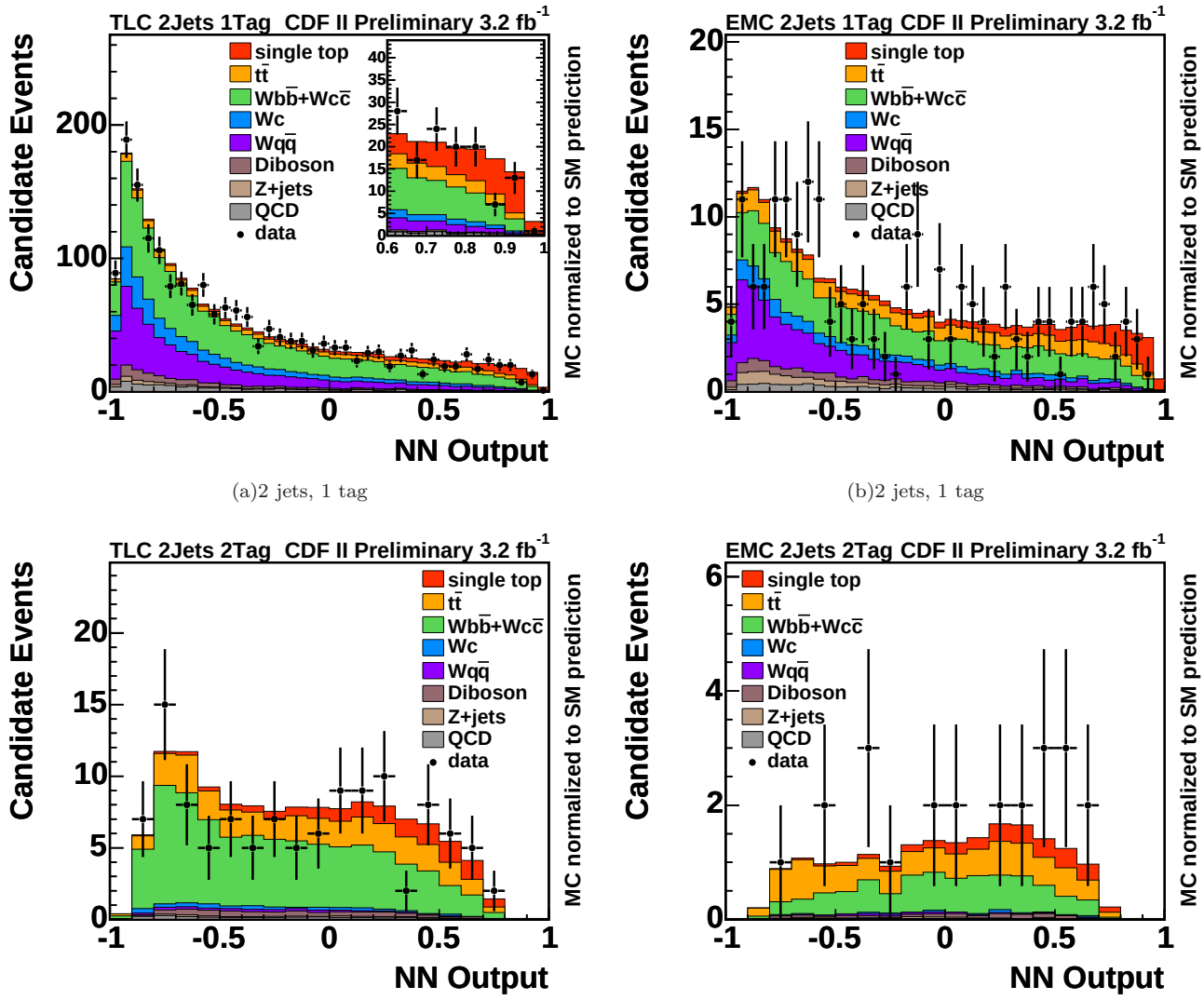


FIG. 11: Distributions of Q-values for two ensemble tests, one with single top–quark events present at the expected standard model rate, one without any single top–quark events.

To quantify the sensitivity of our analysis we define the expected p-value $\hat{p}_{exp} = p(Q_1^{\rm med})$ where $Q_1^{\rm med}$ is the median of the Q-value distribution q_1 for the hypothesis H_1. The meaning of $\hat{p}_{exp}$ is the following: Under the assumption that H_1 is correct one expects to observe $p < \hat{p}_{exp}$ with a probability of 50%. We find $\hat{p}_{exp} = 0.00000011$, including all systematic uncertainties. In other words, assuming the predicted single top–quark production cross–section, we expect with a probability of 50% , to see at least that many single top–quark events that the observed excess over the background corresponds to a background fluctuation of 5.2 σin case of the combined search.



(a)2 jets, 1 tag

(b)2 jets, 1 tag

(c)2 jets; 2 tags

(d)2 jets; 2 tags

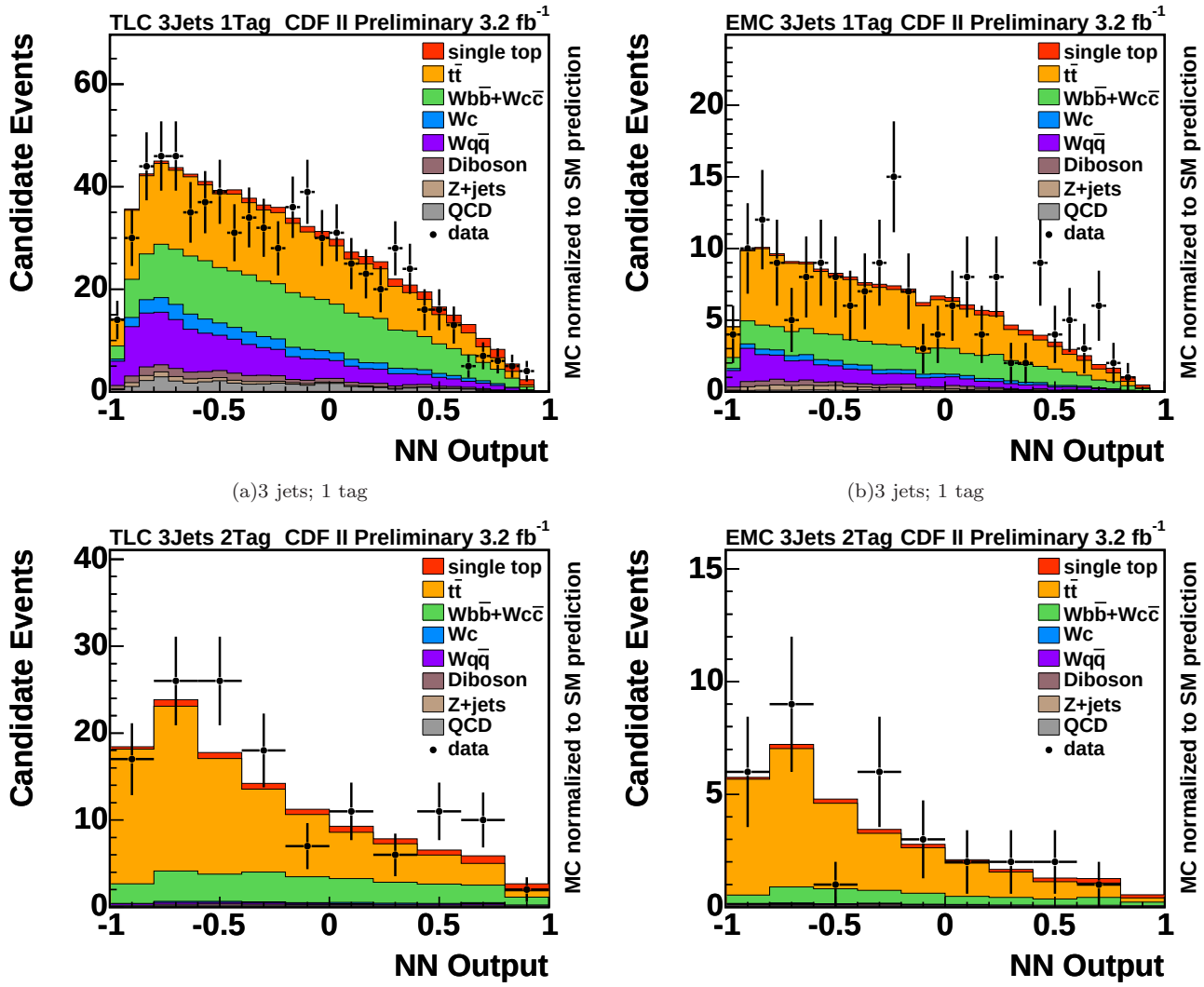
FIG. 12: The predicted and measured distributions of the combined search for the t–channel neural network in the 2–jet bin with 1 b–tag (top) and the s–channel neural network in the 2–jet bin with 2 b–tags (bottom) for TLC (left) and EMC (right) leptons.

VII. BINNED LIKELIHOOD FIT TO DATA FOR COMBINED SEARCH

The predicted and measured output distribution of all 8 channels used in the combined search are depicted in figure 12–13. In figure 14 the distributions of all 8 channels are added together. Finally, the templates are fitted to the 8 observed distributions to determine the single top–quark cross section. The fit yields a single top–quark cross section of $1.8^{+0.6}_{-0.6}$ pb.

Figure 15, compares the observed Q–value ($Q = -23.9$) with the expectation. The observed p–value is 0.00024014 corresponding to an observed significance of 3.5 σ.

Figure 16 summarizes the single top–quark cross sections if we fit the 8 channels separately and also if we fit all categories simultaneously.



(a)3 jets; 1 tag

(b)3 jets; 1 tag

(c)3 jets; 2 tag

(d)3 jets; 2 tag

FIG. 13: The predicted and measured distributions of the combined search for the t–channel neural network in the 3–jet bin with 1 b–tag (top) and in the 3–jet bin with 2 b–tags (bottom) for TLC (left) and EMC (right) leptons.

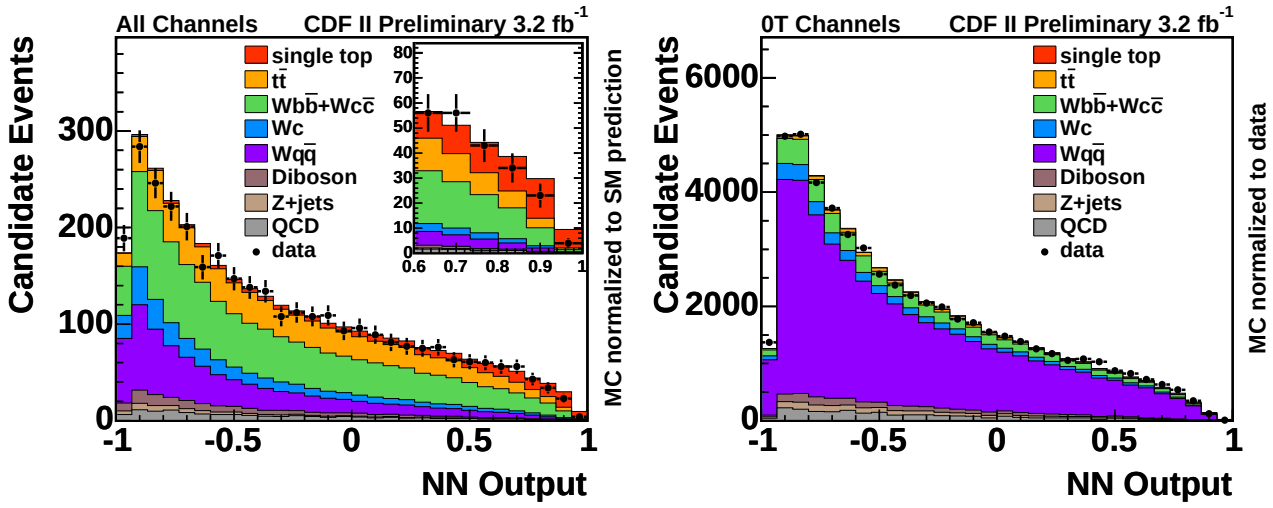


FIG. 14: The predicted and measured distributions of all 8 channels used in the combined search (left) and the MC modeling validation in all untagged sideband channels (right).

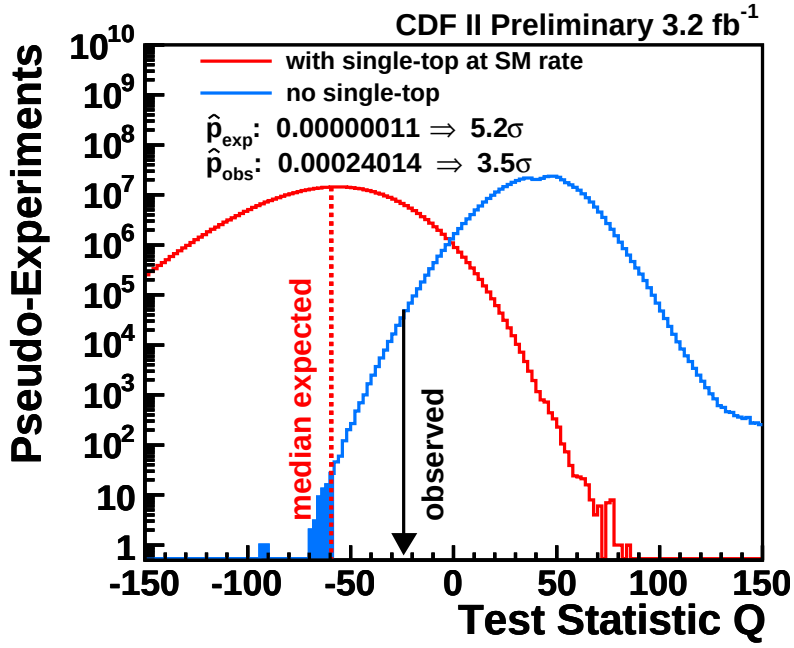


FIG. 15: Comparison of observed Q–value to the expectation in the combined search.

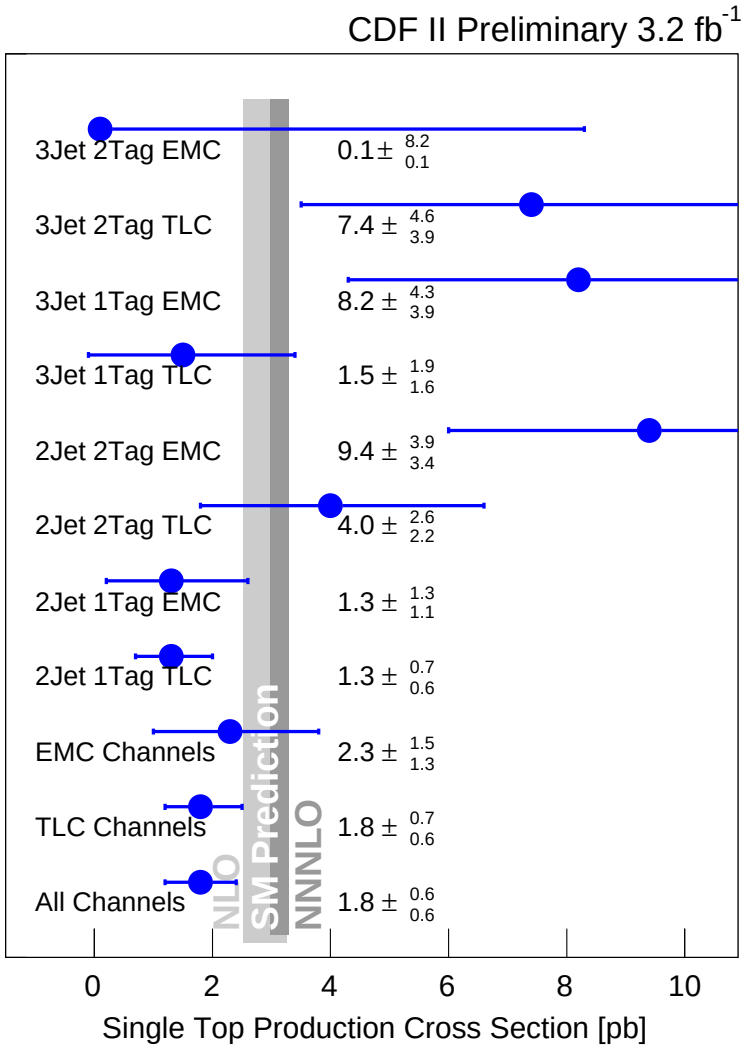


FIG. 16: Summary of the results for the eight different channels and the final result of the simultaneous fit in all channels.

VIII. BINNED LIKELIHOOD FIT TO DATA FOR SEPARATE SEARCH

For the separate search, the fit yields the cross sections $\sigma_t = 0.7^{+0.5}_{-0.5}(\text{stat.} + \text{syst.})$ pb for t–channel and $\sigma_s = 2.0^{+0.7}_{-0.6}(\text{stat.} + \text{syst.})$ pb for s–channel, respectively. The fit result is depicted in figure 17 showing the difference between the logarithm of the reduced likelihood function and its minimum in the plane of single–top–quark s–channel versus t–channel cross sections. Negative cross section values are physically meaningless and therefore not allowed. The minimum represents the best fit values and is indicated by the black dot. The error bars quote the 1σ, 2σ, and 3σ uncertainties ($\Delta ln(L)$ of 0.50, 2.0, and 4.5) on the fitted t– and s–channel cross sections. The true values of both cross sections have a probability of 68.3%, 95.5%, and 99.7% to be found in the region comprised by the corresponding contours ($\Delta ln(L)$ of 1.15, 3.09, and 5.92). The value predicted by the standard model, within its uncertainties, is illustrated by the blue rectangle.

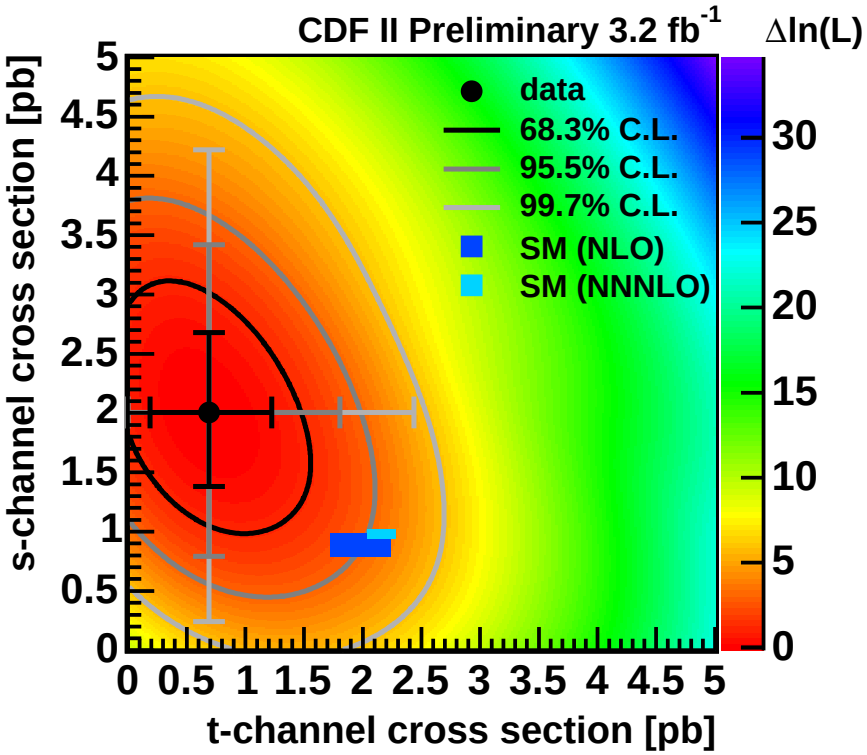


FIG. 17: The likelihood fit estimate for the simultaneous s- and t-channel production cross section measurement. The contours of the 1σ ($\Delta ln(L) = 1.15$), 2σ ($\Delta ln(L) = 3.09$), and 3σ ($\Delta ln(L) = 5.92$) uncertainties are valid for both production channels simultaneously. The error bars represent the 1σ ($\Delta ln(L) = 0.5$), 2σ ($\Delta ln(L) = 2.0$), and 3σ ($\Delta ln(L) = 4.5$) uncertainties of the given production channel without any assumptions on the other production channel.

IX. CONCLUSIONS

We present a search of single top–quark production in a CDF II data set corresponding to 3.2 fb^{-1}, assuming a top quark mass of 175 GeV/c^2. We employ neural networks to construct discriminants between single top–quark events and background events. In a combined search, where t- and s-channel single-top events are regarded as signal, we find an expected p-value of 0.00000011 which corresponds to a expected significance of 5.2 σ. In data we see evidence for single top and compute a p-value of 0.00024014, corresponding to an observed significance of 3.5 σ. We measure

$$\sigma_{\text{single top}} = 1.8^{+0.6}_{-0.6}\ \text{pb}$$

For the separate search the observed t- and s-channel cross sections are

$$\sigma_{t-ch} = 0.7^{+0.5}_{-0.5}\,(\text{stat.} + \text{syst.})\ \text{pb}$$

$$\sigma_{s-ch} = 2.0^{+0.7}_{-0.6}\,(\text{stat.} + \text{syst.})\ \text{pb}$$

[1] D. Acosta *et al.* (CDF Collaboration), *Phys. Rev.* **D65**, 091102 (2002); D. Acosta *et al.* (CDF Collaboration), *Phys. Rev.* **D69**, 052003 (2004); D. Acosta *et al.* (CDF Collaboration), *Phys. Rev.* **D71**, 012005 (2005).

[2] B. Abbott *et al.* (DØ Collaboration), *Phys. Rev.* **D63**, 031101 (2001); V. M. Abazov *et al.* (DØ Collaboration), *Phys. Lett.* **B517**, 282 (2001); V. M. Abazov *et al.* (DØ Collaboration), *Phys. Lett.* **B622**, 265 (2005).

[3] V. M. Abazov *et al.* (DØ Collaboration), *Phys. Rev. Lett.* **98**, 181802 (2007).

[4] V. M. Abazov *et al.* (DØ Collaboration), *Phys. Rev.* **D78**, 012005 (2008).

[5] T. Aaltonen *et al.* (CDF Collaboration), *Phys. Rev. Lett.* **101**, 252001 (2008).

[6] B. W. Harris *et al.*, *Phys. Rev.* **D66**, 054024 (2002); Z. Sullivan, *Phys. Rev.* **D70**, 114012 (2004); J. Campbell, R.K. Ellis, F. Tramontano, *Phys. Rev.* **D70**, 094012 (2004); N. Kidonakis, *Phys. Rev.* **D74**, 114012 (2006).

[7] D. Acosta *et al.* (CDF Collaboration), *Phys. Rev.* **D71**, 052003.