



Measurement of the Forward-Backward Asymmetry in $t\bar{t}$ Production in 3.2 fb^{-1} of $p\bar{p}$ Collisions at $\sqrt{s} = 1.96 \text{ TeV}$

The CDF Collaboration
<http://www-cdf.fnal.gov>
(Dated: March 17, 2009)

Abstract

We measure the forward-backward asymmetry of pair produced top quarks using 776 semi-leptonic b-tagged $t\bar{t}$ events reconstructed with a χ^2 based kinematic fitter and correct the value for experimental effects. This is an update to the measurement published in PRL 101, 202001 [1]. We have increased the dataset from 1.9 fb^{-1} up to 3.2 fb^{-1} . We study the rapidity y_{had} of the hadronically-decaying top (or anti-top) system, tagging the charge with the lepton sign Q_l from the leptonically decaying system. We find the forward-backward asymmetry in the $p\bar{p}$ lab frame to be

$$A_{fb} = 0.193 \pm 0.065^{stat} \pm 0.024^{syst}$$

consistent with previous results. An independent cross-check uses a likelihood fit to templates derived from linear asymmetry in the top production angle in the $t\bar{t}$ rest frame ($1 + A \cos(\alpha)$) and finds asymmetry consistent with our unfold method and our $1 + A * \cos(\alpha)$ hypothesis. These results should be compared with the small $p\bar{p}$ frame charge asymmetry expected in QCD at NLO, $A_{fb} = 0.05 \pm 0.015$.

Contents

1	Introduction	3
1.1	Asymmetry Overview	3
1.2	Previous Measurements	4
1.3	$p\bar{p}$ Frame	4
1.4	$t\bar{t}$ Frame	5
1.4.1	Correction for Experimental Effects	5
1.4.2	Previous Published Results	6
1.5	The Current Measurement	6
2	Event Selection Summary	6
3	Backgrounds, MC Samples	7
4	Event Reconstruction	7
4.1	Matching Jets To Quarks And Reconstructing The Neutrino	8
5	The Rapidity of the Top Quark	9
6	Background Subtraction	10
7	Unfolding	12
7.1	Correction Using Inverse Matrices	13
7.2	Error Propagation	14
7.3	Preliminary Application of Correction Method	16
8	Checks of the Unfold Procedure	16
8.1	Generating a Sample Shape with Known Asymmetry	16
8.2	Pseudoexperiment Tests of Correction Procedure	18
9	Measurement	19
10	Systematic Uncertainties	19
11	Test of Linear Hypothesis	21
12	Check of Reweighted Signal Reconstructed Variables	23
13	Results and Conclusion	24

1 Introduction

In LO QCD, the top quark production angle is symmetric with respect to beam direction. At NLO, QCD predicts a small charge asymmetry, $A_{fb} = 0.050 \pm 0.015$ [2, 3], due to interference of initial-state radiation diagrams with final-state diagrams (Figure 1b and 1a) and the “box diagram” with Born processes (fig. 1c and 1d).

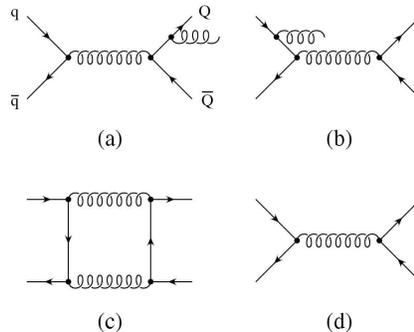


Figure 1: NLO and LO diagrams

In the CP invariant Tevatron system, the charge asymmetry is equivalent to a forward-backward asymmetry. CDF and D0 have both published preliminary measurements with non-zero asymmetries and large uncertainties (see Sec. 1.2.3). In this note we present an update on our measurement of $t\bar{t}$ production asymmetries. Our method follows that of our previous analyses, substituting the more natural top quark rapidity for the laboratory production angle. We have increased the dataset from 1.9 fb^{-1} up to 3.2 fb^{-1} . We study the rapidity y_{had} of the hadronically-decaying top (or anti-top) system, tagging the charge with the lepton sign Q_l from the leptonically decaying system. We assume CP and measure the asymmetry in $-Q \cdot y_{had}$. We subtract backgrounds and perform a model-independent correction for acceptance and reconstruction dilutions in order to find the asymmetry so it can be compared with theoretical predictions. We also show an independent cross-check using a likelihood fit to templates derived from linear asymmetry in the top production angle in the $t\bar{t}$ rest frame ($1 + A \cos(\alpha)$).

1.1 Asymmetry Overview

An integral *charge asymmetry* compares the number of top and anti-top quarks produced with momentum in a given direction. Choosing the incoming proton momentum as this direction gives

$$A_C = \frac{N_t(p) - N_{\bar{t}}(p)}{N_t(p) + N_{\bar{t}}(p)} \quad (1)$$

where $N_i(j)$ is the number of particle i observed in the direction of particle j . A non-zero value A_C implies a net top current in the proton direction.

An integral *forward-backward asymmetry* compares the number of top quarks moving for or against a given direction. A convenient choice for opposite directions are the proton and anti-proton directions, thus

$$A_{fb} = \frac{N_t(p) - N_t(\bar{p})}{N_t(p) + N_t(\bar{p})} \quad (2)$$

where $N_i(j)$ is as described above. If CP invariance is good, $N_{\bar{t}}(p) = N_t(\bar{p})$ and $A_C = A_{fb}$. Our analysis is built for A_{fb} .

1.2 Previous Measurements

Using the the angle θ between the reconstructed top quark momentum vector and the proton beam direction, Equation 2 may be written as

$$A_{fb} = \frac{N_t(\cos(\theta) > 0) - N_t(\cos(\theta) < 0)}{N_t(\cos(\theta) > 0) + N_t(\cos(\theta) < 0)} \quad (3)$$

but the definition is not complete until the frame of reference is specified. Collinear initial-state radiation (ISR) makes the fundamental $q\bar{q}$ frame inaccessible in both experiment and simulation, leaving either the $t\bar{t}$ rest frame or the $p\bar{p}$ lab frame. With the χ^2 reconstruction, measurement of angles in the lab frame is straightforward. Since the boost to the lab frame changes the top quark direction, we expect that any asymmetry would be larger (less diluted) in the $t\bar{t}$ frame, if we could reliably boost back from the lab frame reconstruction. CDF's most recent published measurement, PRL 101, 202001 [1], used 1.9 fb^{-1} of data to measure A_{fb} in both the lab and $t\bar{t}$ rest frames. We briefly explain their method and results here.

1.3 $p\bar{p}$ Frame

In the lab frame we looked at the angle between the hadronically decaying top quark and the proton direction. The hadronically decaying system was chosen because its direction was better reconstructed than the leptonic one, which is complicated by the missing energy of the neutrino. We can use events where the anti-top is the hadronic decay by invoking CP invariance. In the $t\bar{t}$ rest frame, $\vec{p}_t = -\vec{p}_{\bar{t}}$, so we can use $-\cos(\theta)$, and assume the boost to the lab frame will have the same effect on the distribution. In the lab frame then, we used the charge of the lepton in the event to tell us which of the quarks in the $t\bar{t}$ pair decayed leptonically, so multiplying the momentum of the hadronically-decaying system by $-1 \cdot Q_l$ gives us an *equivalent top momentum*. We use $-Q$ so that a net top current in the proton direction will produce a positive asymmetry. The publication used the distribution $-Q \cdot \cos(\theta_{had})$ to measure $A_{fb}^{p\bar{p}}$.

$$A_{fb}^{\text{pp}} = \frac{N(-Q_l \cdot \cos(\theta) > 0) - N(-Q_l \cdot \cos(\theta) < 0)}{N(-Q_l \cdot \cos(\theta) > 0) + N(-Q_l \cdot \cos(\theta) < 0)} \quad (4)$$

1.4 $t\bar{t}$ Frame

To make a measurement in the $t\bar{t}$ frame, we cannot just boost the reconstructed system, because the incorrect solutions create too large a dilution.

Instead, note that the difference in rapidity between the top and anti-top quark (a Lorentz-invariant quantity) at LO is directly related to the top quark production angle as seen by the following equation

$$y_t - y_{\bar{t}} = 2 \cdot \tanh^{-1} \left(\frac{\cos(\theta_{t\bar{t}})}{\sqrt{1 + \frac{4m_t^2}{\hat{s} - 4m_t^2}}} \right) \quad (5)$$

where \hat{s} is center of mass energy and m_t is the top quark mass. Note that while there is not an exact 1:1 correspondence of Δy with $\cos(\theta)$ due to \hat{s} , the forward-backward nature remains the same. That is, if $\theta < 0$, then $\tanh^{-1}(\cos(\theta)) < 0$, so $\Delta y < 0$. Therefore asymmetries in Δy will be identical to those in $\cos(\theta)$, allowing an effective measurement in the $t\bar{t}$ rest frame. In practice the difference of top and anti-top rapidities is related to the difference of the hadronic and leptonic top rapidities, modulo the lepton sign, thus

$$A_{fb}^{t\bar{t}} = \frac{N(Q_l \Delta y_{lep-had} > 0) - N(Q_l \Delta y_{lep-had} < 0)}{N(Q_l \Delta y_{lep-had} > 0) + N(Q_l \Delta y_{lep-had} < 0)} \quad (6)$$

In principal, this equation recovers the larger undiluted asymmetry than A_{fb}^{pp} , but with a larger uncertainty, as the neutrino uncertainty still enters through the involvement of y_{lepton} in the calculation of $\Delta y_{lep-had}$.

1.4.1 Correction for Experimental Effects

The raw A_{fb} values calculated directly from data using Equations 3 and 6 are the corrected for experimental effects by accounting for the presence of backgrounds and shape distortions. First we note that non-signal events that pass our selection criteria cuts (see Sections 2 and 3), particularly those from electroweak processes, may have asymmetries that affect the final A_{fb} calculation. These backgrounds are estimated and subtracted, using the method we repeat and explain in Section 6. Event selection cuts (see Section 2) remove signal events and modify the distribution shape of our production angle data. Finally, the t and \bar{t} four vectors must be reconstructed from limited information (4 jets, 1-2 btags, MET, as explained in Sections 2 and 4), and it is known that uncertainty in reconstruction causes smearing of the $\cos(\theta)$ and Δy distributions. The acceptance and smearing effects are corrected using a matrix unfold method that we also repeat here, see Section 7 for a complete description.

1.4.2 Previous Published Results

Employing these techniques, Reference [1], using 1.9 fb^{-1} of data, observed

$$A_{fb}(\cos(\theta)) = A_{fb}^{\text{p}\bar{\text{p}}} = 0.17 \pm 0.08 \quad (7)$$

$$A_{fb}(\Delta y) = A_{fb}^{\text{t}\bar{\text{t}}} = 0.24 \pm 0.14 \quad (8)$$

The $\text{p}\bar{\text{p}}$ frame value is large, but consistent with the NLO prediction of $A_{fb} = 0.05 \pm 0.015\%$ within large uncertainty. The results in the two frames are consistent with the theoretically expected dilution of 30% in passing from $\text{t}\bar{\text{t}}$ to $\text{p}\bar{\text{p}}$ frames [2].

1.5 The Current Measurement

Our measurement here uses the laboratory direction of the hadronic system and is thus a $\text{p}\bar{\text{p}}$ frame measurement. We substitute the more natural top quark rapidity for the laboratory production angle. We have increased the dataset from 1.9 fb^{-1} up to 3.2 fb^{-1} . We study the rapidity, y_{had} of the hadronically-decaying top (or anti-top) system, tagging the charge with the lepton sign Q_l from the leptonically decaying system. We assume CP and measure the asymmetry in $-Q \cdot y_{had}$

$$A_{fb} = \frac{N(-Q \cdot y_{had} > 0) - N(-Q \cdot y_{had} < 0)}{N(-Q \cdot y_{had} > 0) + N(-Q \cdot y_{had} < 0)} \quad (9)$$

We subtract backgrounds and perform a model-independent correction for acceptance and reconstruction dilutions in order to find the asymmetry so it can be compared with theoretical predictions. We also show an independent cross-check using a likelihood fit to templates derived from linear asymmetry in the top production angle in the $\text{t}\bar{\text{t}}$ rest frame ($1 + A \cos(\alpha)$).

2 Event Selection Summary

This analysis selects $\text{t}\bar{\text{t}}$ events in the lepton plus jets channel where one top decays semi-leptonically ($t \rightarrow l\nu b$) and the other hadronically ($t \rightarrow q\bar{q}b$). Selection begins by requiring a single high transverse momentum electron or muon in the central portion of the detector ($|p_t| > 20 \text{ GeV}/c$ and $|\eta| < 1.1$). In addition, we require a large amount of missing transverse energy as evidence of the presence of a neutrino ($\cancel{E}_T \geq 20 \text{ GeV}$). Each event must have four or more tight jets ($|E_t| > 20 \text{ GeV}/c$ and $|\eta| < 2.0$) and at least one jet must have two tracks that form a secondary vertex (a “tagged” jet). A tagged jet is evidence that the jet originates from a “b” quark and therefore this requirement reduces W plus light flavor background processes which dominate the event sample. The above selection produces roughly a 3.6 to 1 signal to background

ratio as calculated in the section below. For 3.2 fb^{-1} of data collected at CDF the number of events that pass through event selection is 776. A description of a similar selection process can be found in the documentation of the lepton plus jets cross-section measurement [4].

3 Backgrounds, MC Samples

Our background models and their normalizations are based on the same procedure used in Reference [4], applied to data through 3.2 fb^{-1} . The calculated contribution of each background component in our sample is given in Table 1. The total number of background events in the sample is 167.4 ± 33.5 . The sample contains 776 ± 55.6 top-pair events, and the ratio of number of signal events to background is roughly 3.6:1.

Process	==4 Jets	≥ 5 Jets
W + HF Jets	70.08 ± 21.99	16.48 ± 5.41
Mistags (W+LF)	22.52 ± 5.72	4.91 ± 1.98
Non-W (QCD)	25.04 ± 20.53	8.40 ± 7.53
Single Top	6.62 ± 0.42	1.20 ± 0.08
WW/WZ/ZZ	6.00 ± 0.57	1.57 ± 0.17
Z+Jets	3.89 ± 0.48	0.89 ± 0.11
Top	425.02 ± 58.86	144.06 ± 19.95
Total Prediction	559.15 ± 66.99	177.49 ± 22.23

Table 1: Summary of background numbers used
See Table 2 for the raw Afb values of each background sample.

4 Event Reconstruction

The measurement of A_{fb} will use the production angle of the top quark. The top quark is not directly observed in the detector, and therefore, we must reconstruct its momentum 4-vector from the final state particles: jets, charged leptons and neutrino. Unfortunately, we measure only the transverse component of the neutrino (in the \cancel{E}_T) and it is impossible to identify the parent quark of a jet based upon detector information. Because the type of parton cannot be identified by its jet, we cannot tell which jets came from which partons in a $t\bar{t}$ event. If we are to reconstruct the event we must find a method to choose the correct jet-parton assignments, as illustrated in Figures 2 and 3. We use an algorithm to match jets to the correct partons and reconstruct the full neutrino momentum by employing several constraints available in the “ $t\bar{t}$ lepton plus jets hypothesis”. This method allows us to reconstruct the complete kinematics of the $t\bar{t}$ final state.

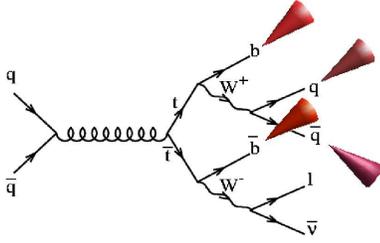
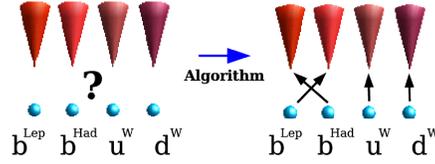
Figure 2: $t\bar{t}$ Lepton Plus Jets Event

Figure 3: Matching Jets To Quarks

4.1 Matching Jets To Quarks And Reconstructing The Neutrino

The problem of reconstructing the $t\bar{t}$ event is a combinatoric one: we must choose between a number of possible arrangements. The highest four energy jets in the event are assumed to come from the four quarks in the $t\bar{t}$ process. Matching four jets to four quarks leads to 24 possible combinations. This can be reduced by a factor of two since interchanging the two quarks from W -boson decay does not change the kinematics of the event.

Because we cannot measure the momentum of the event along the beam direction, we cannot infer the P_z of the neutrino from “missing E_z ”. However, we can calculate the neutrino P_z by requiring that the lepton and neutrino be consistent with the known mass of the W -boson. This calculation involves a quadratic equation and produces two solutions for the neutrino P_z . Both solutions are considered. Together with the jet assignments, the event has 24 possible combinations.

Our strategy is to test each combination for consistency with the “ $t\bar{t}$ hypothesis”. That hypothesis has four main components:

- The lepton and neutrino are decay products of a W -boson ($W \rightarrow l\nu$)
- Two jets are decay products of a W -boson ($W \rightarrow jj$)
- The lepton, neutrino, and a third jet are final states from a top quark decay ($t \rightarrow l\nu j$)
- The two jets from $W \rightarrow jj$ and a fourth jet are final states from the other top quark decay ($t \rightarrow jjj$)

The consistency of each combination with the $t\bar{t}$ hypothesis is assessed with a χ^2 test. The χ^2 equation is:

$$\begin{aligned}
\chi^2 = & \sum_{i=l,jets} \frac{(p_t^{i,meas} - p_t^{i,fit})^2}{\sigma_i^2} + \sum_{j=x,y} \frac{(p_j^{UE,meas} - p_j^{UE,fit})^2}{\sigma_j^2} \\
& + \frac{(M_{jj} - M_W)^2}{\Gamma_W^2} + \frac{(M_{lv} - M_W)^2}{\Gamma_W^2} + \frac{(M_{bjj} - M_{fit})^2}{\Gamma_t^2} + \frac{(M_{blv} - M_{fit})^2}{\Gamma_t^2}
\end{aligned} \tag{10}$$

While we are assessing the “goodness-of-fit” we can also take the opportunity to make modest corrections to the jet energies. The last four terms are the constraints. M_{jj} is the invariant mass of the two jets that must be consistent with the known W boson mass. M_{bjj} and M_{lv} are the invariant masses of the hadronically decaying and leptonically decaying top quark side. These should be consistent with being equal, and their common value, M_{fit} is the best estimate of the top quark mass. M_{lv} is the mass of the lepton and the neutrino which must be consistent with the mass of a W boson. All four of the constraints are particle masses, and their weights are the theoretical decay width of the particle.

The first two terms are sums over lepton and jet transverse energies and “unclustered” energy, which is the energy in the event outside the $t\bar{t}$ interaction. These values are varied within their measured error. This improves resolution on jet energies, as well as the probability of finding the correct combination. The known top quark mass may also be used as a further constraint in the fit by setting $M_{fit} = M_{known}$.

The standard package MINUIT is used to vary the independent parameters and minimize the χ^2 for each possible combination of jet-parton assignments and neutrino solutions [5]. The combination with the lowest χ^2 is chosen as the best representative of the $t\bar{t}$ hypothesis for the event. Tests with Monte Carlo simulations show that the correct assignment is chosen 45% of the time, and this improves to 60 % in the constrained fit. However, incorrect combinations still provide useful information about the event kinematics.

5 The Rapidity of the Top Quark

Before examining the raw asymmetry we first checked that the detector does not have an inherent asymmetry by examining observable detector variables, such as jet energy and rapidities, lepton Pt and rapidity, and various angles between the jets, lepton, and MET. We found good agreement in all plots. We then examined the fitted variables, that is, variables set by the kinematic fitter, such as Top, W, and b-jet Et and rapidity. These plots showed good agreement between data and MC predictions.

We now investigate our $-Q_l \cdot y_{had}$ distribution and measure a raw A_{fb} – Figure 4 is the main distribution used in the remainder of our analysis. We observe that the data shape has a noticeable shift from the background and background+signal

predictions. We want to relate the asymmetry measured in this raw distribution to the true underlying asymmetry, so we need to correct for background and account for smearing and acceptance effects. We turn now to techniques for these corrections and return to our measurement in Section 9.

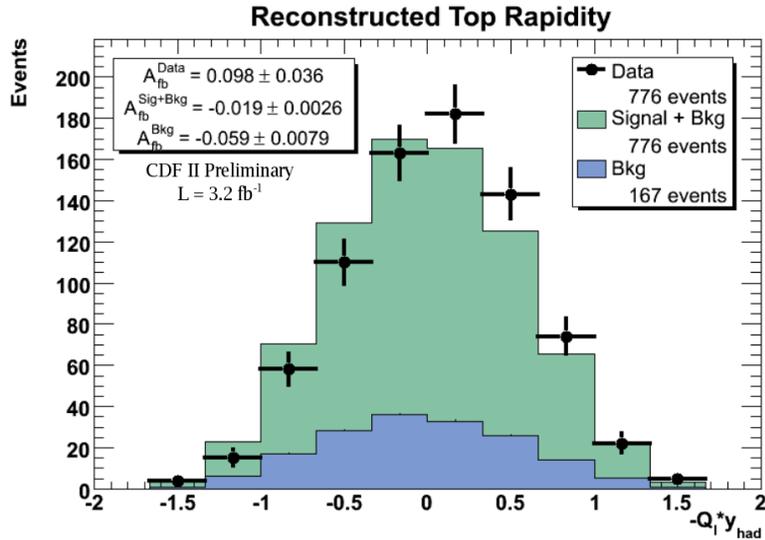


Figure 4: $-Q_l \cdot y_{had}$

6 Background Subtraction

Non- $t\bar{t}$ events that pass our event selection requirements have a small overall asymmetry (see Figure 5 and Table 2 for numbers), and this affects our measured $-Q_l \cdot y_{had}$ distribution shape. However, we can subtract off the total background shape and correct our data back to the signal shape.

Figure 5 shows the expected background shapes in our reconstructed $-Q_l \cdot y_{had}$ variable. Table 2 lists the A_{fb} values of each component. We see that some background asymmetries can be large, especially in W+HF.

A check of the background shapes is available in the antitagged sample, which is deficient in $t\bar{t}$ events, and therefore dominated by the background. In the legend of Figure 6, “Bkg” is what we expect for the backgrounds and “Sig+Bkg” includes the small amount of our Pythia $t\bar{t}$ signal model. The A_{fb} values for the data, background, and background+signal are all consistent with each other and the KS comparison is very good. Now, it is true that the asymmetry is smaller in the anti-tagged sample than the tagged sample accounted for in Table 2. This is because the anti-tags select against the W+HF which has the highest asymmetry. The anti-tags are the control sample we have, and we take the good agreement in the background dominated anti-tags as evidence that the model reproduces the data. Possible model dependence in the

shapes and normalizations will be included as systematic uncertainties (see Section 10).

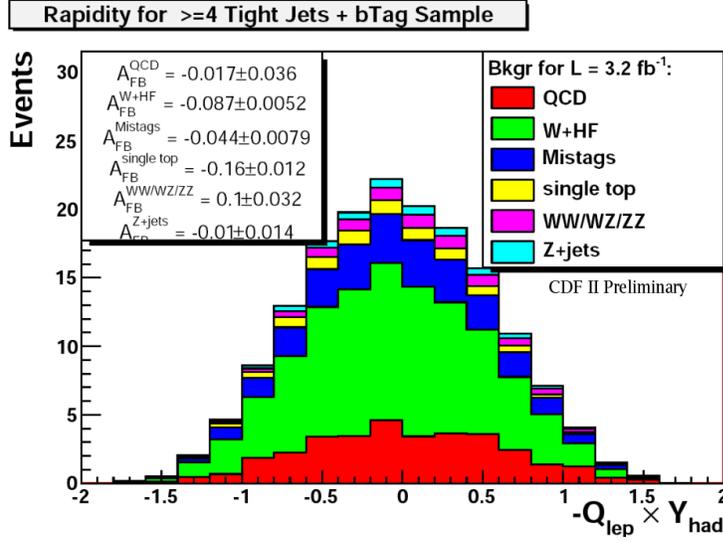
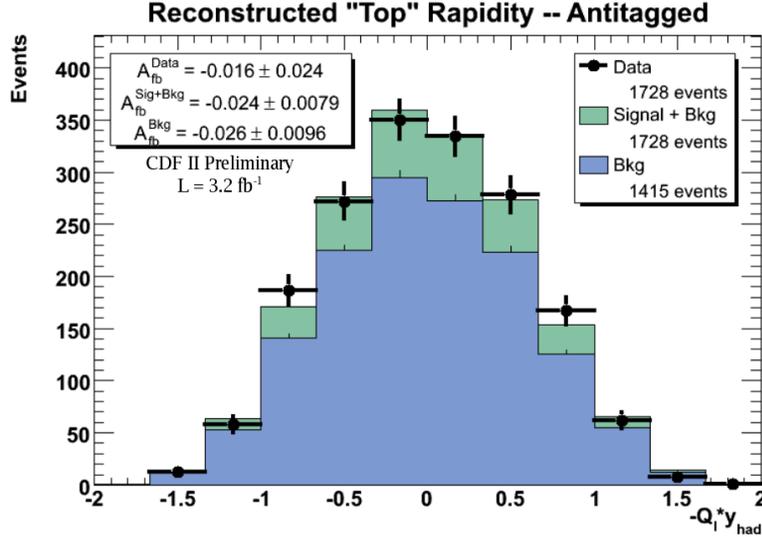


Figure 5: Background Components

Process	≥ 4 Jets	≥ 5 Jets	≥ 4 Jets
Tagged W+HF Jets	-0.095 ± 0.0078	-0.078 ± 0.007	-0.087 ± 0.0052
Tagged W+LF	-0.038 ± 0.012	-0.05 ± 0.01	-0.044 ± 0.0079
Tagged Non-W (QCD)	-0.044 ± 0.06	-0.002 ± 0.044	-0.017 ± 0.036
Tagged Single Top	-0.18 ± 0.016	-0.12 ± 0.018	-0.16 ± 0.012
Tagged WW/WZ/ZZ	0.078 ± 0.046	0.13 ± 0.045	0.1 ± 0.032
Tagged Z+Jets	-0.016 ± 0.021	-0.0041 ± 0.018	-0.01 ± 0.014
Total Prediction	-0.07 ± 0.012	-0.045 ± 0.012	-0.059 ± 0.0079

Table 2: A_{fb} Summary of A_{fb} values for different background samples

Figure 6: $-Q_l \cdot y_{had}$ for the antitagged sample

7 Unfolding

In addition to background contributions, we know of two other effects that modify the true $-Q_l \cdot y_{had}$ distribution. First, the kinematic fitter is known to smear out the true rapidities of the reconstructed top quark. We can examine this effect by using the ttop75 Pythia MC to generate a 2d histogram of $-Q_l \cdot y_{had}^{TRUE}$ vs $-Q_l \cdot y_{had}^{RECON}$. This smear matrix describes the movement of events from bin-to-bin in going from truth to reconstructed data when the kinematic fitter is applied. We noted most of the large values in the smear matrix lie close to the diagonal, meaning there is little extreme smearing of far-apart bins. Also, values are roughly symmetrical about the diagonal in the second plot, which shows that smearing will not cause an A_{fb} if none exists in the true distribution, but rather will dilute any such existing A_{fb} .

After rebinning and normalizing, our smear matrix has the values listed in Equation 15.

In addition to reconstruction smearing effects, our true distribution is also modified by event selection itself, which cuts out some $t\bar{t}$ events. If this acceptance is biased with respect to y_t , such bias would cause a change in measured A_{fb} . We look at truth info for events in the MC sample and calculate the ratio of events passed to generated number in each rapidity bin. Using these bin ratios we can then construct a matrix whose diagonal entries contain the ratio information, which describes the acceptance bias. Applying the inverse of this matrix to our data will correct our post-selection-cuts shape to the true pre-cuts shape. We describe this method in detail below, but first we will explain our re-binning procedure.

7.1 Correction Using Inverse Matrices

In order to correct for smearing we construct a 4x4 matrix S from the truth-vs-recon matrix and binning choice explained above. The entries of S are

$$S_{ij} = N_{recon}^{ij}/N_{truth}^i \quad (11)$$

where N^{ij} is the number of events in the smear matrix with i is the truth bin index and j the reconstructed bin index. N^i is the value of bin i of the truth histogram, the projection of N^{ij} , used to properly normalize the matrix.

In addition to smearing, we also know that selection cuts remove a number of $t\bar{t}$ events from our analysis. These removed events may have a different asymmetry than the remaining events, so to correct our raw asymmetry we apply the inverse of an acceptance matrix

$$A_{ii} = N_{selection}^i/N_{generated}^i \quad (12)$$

Where A only has diagonal terms with values equaling the ratio of number of events selected over the total number, using the same 4x4 binning scheme as the data and smear matrices.

So we see that if we start with a true distribution of $-Q \cdot y_{had}$, it is first modified by selection, then smearing, then background (Equation 13). We find our final value by correcting for these effects in reverse order.

$$N_{raw} = [S \cdot (A \cdot N_{true})] + N_{bkg} \quad (13)$$

$$N_{corrected} = A^{-1} \cdot S^{-1} \cdot N_{raw-bkg} \quad (14)$$

where N is a vector whose values are the number of events in each bin of our $-Q \cdot y_{had}$ distribution.

The smear matrix we use is calculated from $t\bar{t}$ ttop75 Pythia MC and is

$$S = \begin{bmatrix} 0.76 \pm 0.0096 & 0.13 \pm 0.0041 & 0.035 \pm 0.0022 & 0.011 \pm 0.0012 \\ 0.17 \pm 0.0046 & 0.66 \pm 0.0093 & 0.16 \pm 0.0046 & 0.042 \pm 0.0023 \\ 0.053 \pm 0.0025 & 0.17 \pm 0.0047 & 0.67 \pm 0.0095 & 0.16 \pm 0.0045 \\ 0.017 \pm 0.0014 & 0.042 \pm 0.0023 & 0.13 \pm 0.0042 & 0.79 \pm 0.01 \end{bmatrix} \quad (15)$$

Using this same MC sample, we find from comparing the event counts before and after cuts that

$$A = \begin{bmatrix} 0.981 \pm 0.00483 & 0 \pm 0 & 0 \pm 0 & 0 \pm 0 \\ 0 \pm 0 & 1.09 \pm 0.00555 & 0 \pm 0 & 0 \pm 0 \\ 0 \pm 0 & 0 \pm 0 & 1.05 \pm 0.00547 & 0 \pm 0 \\ 0 \pm 0 & 0 \pm 0 & 0 \pm 0 & 0.901 \pm 0.00464 \end{bmatrix} \quad (16)$$

Note that neither the S nor A are expected to be unitary, therefore we are unconcerned with array values greater than 1. In fact, because the acceptance ratio (selected/generated) is actually small, we normalize the above matrices to values near 1 in order to easily analyze number trends.

Using these matrices, we find a final correction unfold matrix

$$A^{-1} \cdot S^{-1} = \begin{bmatrix} 1.4 & -0.28 & -0.0061 & -0.0043 \\ -0.33 & 1.6 & -0.35 & -0.0084 \\ -0.016 & -0.38 & 1.6 & -0.29 \\ -0.0095 & -0.018 & -0.28 & 1.5 \end{bmatrix} \quad (17)$$

We drop the statistical error in the unsmearing matrix at this point, as it is negligible compared to the overall matrix entry values. In the next section we explain how errors are propagated for the corrected distribution shape. We see that small statistical errors in the unsmearing matrix would enter only as second-order effects and therefore are negligible compared to the data-background statistical errors and first-order error effects from the correction procedure.

Note that the unsmearing matrix has larger values along the diagonal entries and much smaller values for far off-diagonal entries. This shows us that smearing occurs mostly for events close to the bin edges in our rebinned distribution. The near-symmetry of these numbers is indicative that the correction procedure should not bias the A_{fb} value. We discuss possible bias effect in detail in Section 8.

Again, we note that the $A^{-1} \cdot S^{-1}$ matrix is not expected to be unitary.

7.2 Error Propagation

With the understanding of acceptance and reconstruction bias in hand, we can develop an overall formalism for correcting the measured A_{fb} back to the true A_{fb} of $t\bar{t}$ production. Matrices A and S are multiplied together to create a relationship between the background corrected number of forward and backward events and the true number of forward and backward events generated in Monte Carlo. We will call the corrected values that are comparable to the number of events generated $N_{corrected}$.

$$N_{bkg-sub} = S \cdot A \cdot N_{truth} \quad (18)$$

The combined matrix formed by multiplication of A and S is then inverted so that we can solve for the corrected values.

$$N_{corrected} = A^{-1} \cdot S^{-1} \cdot N_{bkg-sub} \quad (19)$$

This technique is used to calculate the final corrected asymmetry that may be compared to theoretical prediction. The forward backward asymmetry is calculated as follows. Let,

$$\alpha = [1, 1, \dots, 1, 1] \quad (20)$$

$$\zeta = [1, 1, \dots, 1, -1, \dots, -1, -1] \quad (21)$$

Then

$$A_{fb} = \frac{\zeta \cdot N_{corrected}}{\alpha \cdot N_{corrected}} \quad (22)$$

The uncertainty on this equation is slightly more complicated. To simplify some algebra let:

$$N = N_{corr} \quad (23)$$

$$n = N_{bkg-sub} \quad (24)$$

$$M = A^{-1} \cdot S^{-1} \quad (25)$$

So,

$$N = M \cdot n \quad (26)$$

is equivalent to Equation 19. A_{fb} can then be represented as a sum:

$$A_{fb} = \frac{\sum_i \zeta_i \sum_j^{nbins} M_{i,j} \cdot n_j}{\sum_i \alpha_i \sum_j^{nbins} M_{i,j} \cdot n_j} \quad (27)$$

Now we just perform simple error propagation:

$$\sigma_{A_{fb}}^2 = \sum_i \sigma_{n_i}^2 \cdot \left(\frac{\delta A_{fb}}{\delta n_i} \right)^2 \quad (28)$$

where, σ_{n_i} = the statistical uncertainty in bin ‘‘i’’ for background corrected data and,

$$\frac{\delta A_{fb}}{\delta n_x} = \frac{(\sum_i \zeta_i \cdot M_{i,x}) \cdot (\alpha \cdot N) - (\sum_i \alpha_i \cdot M_{i,x}) \cdot (\zeta \cdot N)}{(\alpha \cdot N)^2} \quad (29)$$

7.3 Preliminary Application of Correction Method

Before accepting the above correction method, we first need to validate its effect on samples with known asymmetry. This will be explained in the following section. However, we wish to first demonstrate our correction procedure graphically by presenting the $-Q \cdot y_{had}$ distribution in various stages of correction.

Figure 7 shows graphically the changes in the $-Q \cdot y_{had}$ distribution as we apply corrections. The black shape is the raw data and the solid blue histogram the background shape. Subtracting off this background yields the green-colored histogram, and after applying the inverse matrices $A^{-1} \cdot S^{-1}$, we arrive at the final corrected red-colored histogram. This process will be repeated in our final measurement section along with A_{fb} values for each shape.

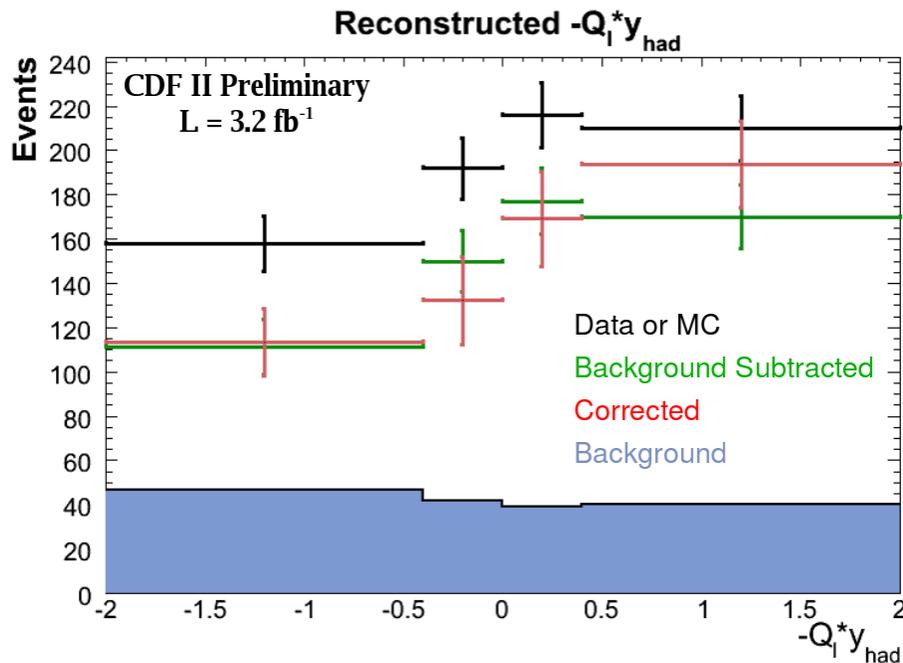


Figure 7: raw data (black), background (solid blue) data-background (green), and corrected data (red)

8 Checks of the Unfold Procedure

8.1 Generating a Sample Shape with Known Asymmetry

Before we apply this correction method to our data we would like to validate it using a sample with known asymmetry. We begin by using a different ttbar Pythia signal sample then we use for the unfold matrix to ensure our uncertainties calculation is not correlated with the unfold. We note that this sample has a raw asymmetry consistent

with 0 in the $t\bar{t}$ rest frame. We reweight the production angle distribution in the $t\bar{t}$ rest frame by the simplest possible function, which is linear in the variable $\cos(\theta)$:

$$N'_{reweighted} = N_i + n_i = N_i \cdot (1 + A \cdot \cos(\theta)) \quad (30)$$

where N_i is the initial bin $i = \cos(\theta)$ value, n_i is the change in value, and A is a free parameter. More simply,

$$n_i = A \cdot N_i \cdot \cos(\theta) \quad (31)$$

We choose the rest frame of the $t\bar{t}$ event as boosting to the lab frame introduces a dilution effect. To account for this, we use the bin ratios of reweighted to original data to make a reweight vector

$$w_i = \frac{N_i + n_i}{N_i} \quad (32)$$

We apply w_i to a $\cos(\theta_{t\bar{t}})$ vs $\cos(\theta_{lab})$ smear matrix, reweighting slices of constant $\cos(\theta_{t\bar{t}})$. By taking a projection along the lab axis, we now have a reweighted shape distribution of $\cos(\theta_{lab})$. We use the ratio of reweighted to original data in this shape to find a reweight vector w'_{lab} which we apply to the smear matrix $\cos(\theta_{lab})$ vs y_{lab} . The projection of this reweighted matrix gives a reweighted distribution in rapidity. Finally, we use this ratio w_y of reweighted to original of this histogram to reweight the shape of the reconstructed $-Q \cdot y_{had}$ signal MC after cuts. This matrix represents what we expect to see in our raw data (minus background). That is, given an original true $\cos(\theta_{t\bar{t}})$ distribution before selection cuts, this procedure gives an expected reconstructed y_{lab} distribution after selection cuts.

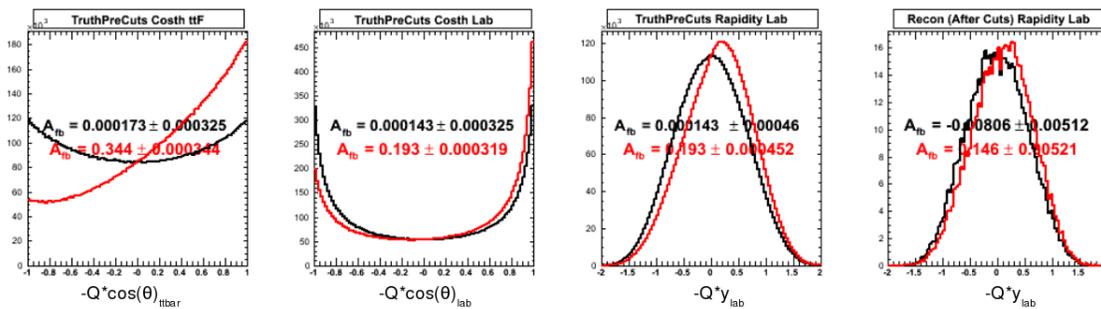


Figure 8: Evolution of angular distribution shapes during the reweight procedure

Figure 8 shows these steps graphically. In the leftmost plot, the black line shape is the original true precut top MC distribution of $-Q \cdot \cos(\theta_{t\bar{t}})$, and the red shape is modified bin-wise by ratio w_i . In the next plot we have moved to the lab frame, where the black data are the original shape and the red have been modified by w'_{lab} . The next plot changes variables to rapidity y_{lab} . Note the A_{fb} value for the reweighted data (red), which is the value we “target” when choosing our parameter A. We will use this

value to compare with our corrected A_{fb} from data as well as generated samples used in our systematic uncertainties calculation. The final plot shows the differences between the reconstructed MC shape (post-cuts) with the final reweighted shape.

8.2 Pseudoexperiment Tests of Correction Procedure

We use our reweighting procedure to test the reliability of our correction method over a range of values for the asymmetry parameter A . For each A , we generate a reweighted samples as above, noting the true precut A_{fb} value. We then correct the final reweighted reconstructed post-cut distribution with our method in Section 7 and compare the final corrected A_{fb} with the true A_{fb} expected.

In practice we use pseudoexperiments to vary our generated samples by Poisson-fluctuations consistent with bin values. For a given input A_{fb}^{t} we calculate PE values for the raw A_{fb} in the lab, and the corrected lab A_{fb} and its error. We compare the measured A_{fb} values from the corrected distribution with the expected A_{fb} in the lab, and calculate a pull-distribution.

We examine the performance over the range of A_{fb} . Plotting measured vs true A_{fb} using values from our pseudoexperiments shows us how accurate our correct method is. If our correction unfold matrix has a bias on A_{fb} , this will be evident in a slope for this plot. See Figure 9 and note the 1:1 correspondence.

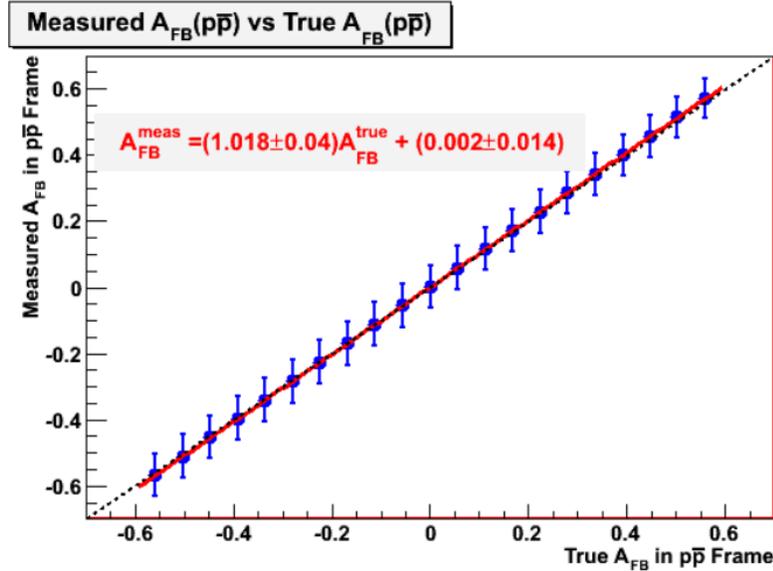


Figure 9: Measured A_{fb} vs True A_{fb}^{precut}

9 Measurement

We choose to bin our histogram in 4 bins of $-Q_l \cdot y_{had}$, with the central bins' outer edges placed at 0.4 in order to minimize any correction bias. Having validated our data and backgrounds and checked the validity of our correction procedure, we are now ready to calculate a corrected A_{fb} from our data. See Figure 4 for our $-Q_l \cdot y_{had}$ raw distribution, which has an initial asymmetry of

$$A_{fb}^{raw} = 0.098 \pm 0.036 \quad (33)$$

We subtract off the background shape in that Figure which has an A_{fb} of -0.059 ± 0.018 , calculating an intermediate corrected data asymmetry of

$$A_{fb}^{raw-bkg} = 0.141 \pm 0.046 \quad (34)$$

Applying our unsmearing correction procedure as explained in Section 7, using the $A^{-1} \cdot S^{-1}$ correction matrix in Equation 17, and propagating error as explained in section 7.2, we arrive at a final corrected asymmetry measurement of

$$A_{fb}^{corrected} = 0.193 \pm 0.065^{stat} \quad (35)$$

See Figure 7 in Section 7.3 for a graphical representation of the angular distribution's evolution with the application of our correction procedure.

10 Systematic Uncertainties

A number of systematic effects contribute to our measurement uncertainty in a way that is not yet reflected in our calculation. Each systematic is estimated in a unique way, but the general procedure is to compare the measured result of a $t\bar{t}$ Monte Carlo model with $A_{fb} = 0.193$ before and after the systematic has been varied. We generate a simulated shape for signal+background with a modified signal or background (depending on the systematic of study) as described in Sections 7 and 8, apply our correction procedure (subtract off the total background and unfold), and compare our measurement of corrected A_{fb} in the modified sample with the nominal sample.

The normalization of our background and the shape of the Monte Carlo model are varied within error and the difference in the measurement of our example models is taken as a systematic. The Jet Energy Scale is estimated by fluctuating the model by the known uncertainties in JES by $\pm 1\sigma$. Similarly, samples of Monte Carlo were generated with more and less initial and final state radiation to estimate the impact of each. We also used a different MC sample for $t\bar{t}$ signal and compared our unfold method results with the nominal sample used for our systematic uncertainties study.

Additionally, this measurement has been tested for a number of different underlying production angle distributions. The variance between different distributions has been taken as a systematic (see Figure 10 for samples of the reweight functions used).

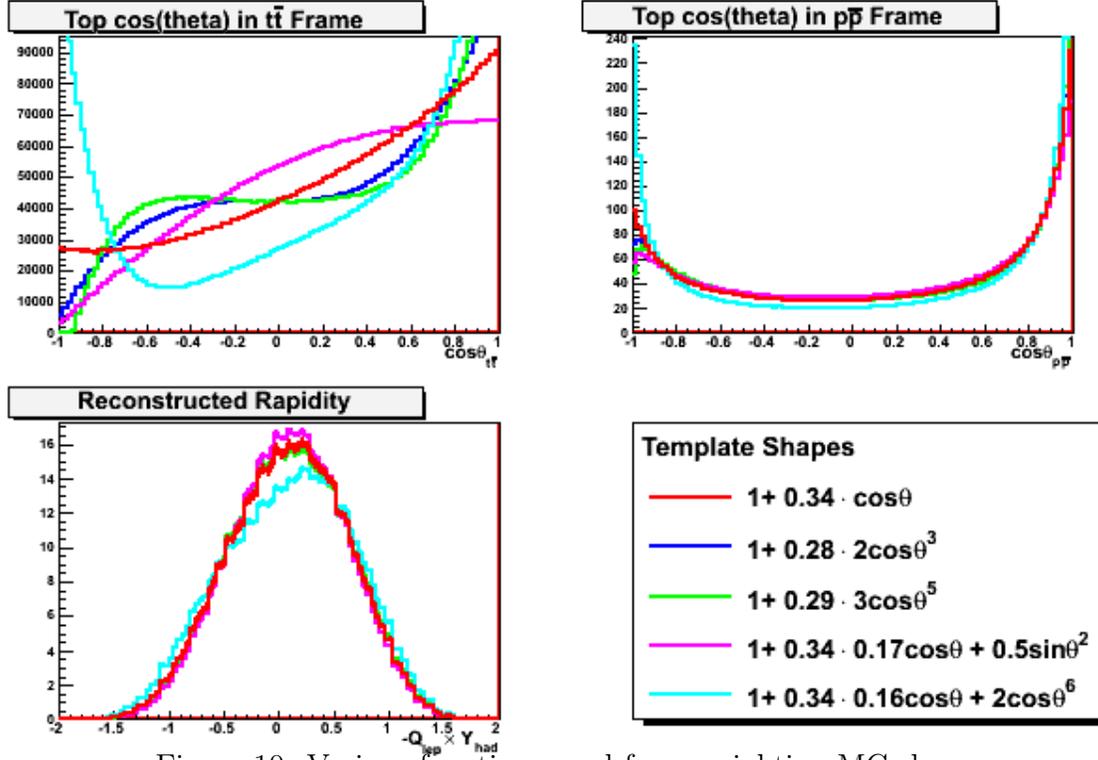


Figure 10: Various functions used for reweighting MC shapes

Finally, we use 46 different sets of PDF and compare to the default set used. Table 3 summarizes the uncertainty taken for each systematic effect. The dominant uncertainty is due to background shape and normalizations. The combined systematic uncertainty on the measurement of A_{fb} is calculated by adding each individual uncertainty in quadrature. The result is:

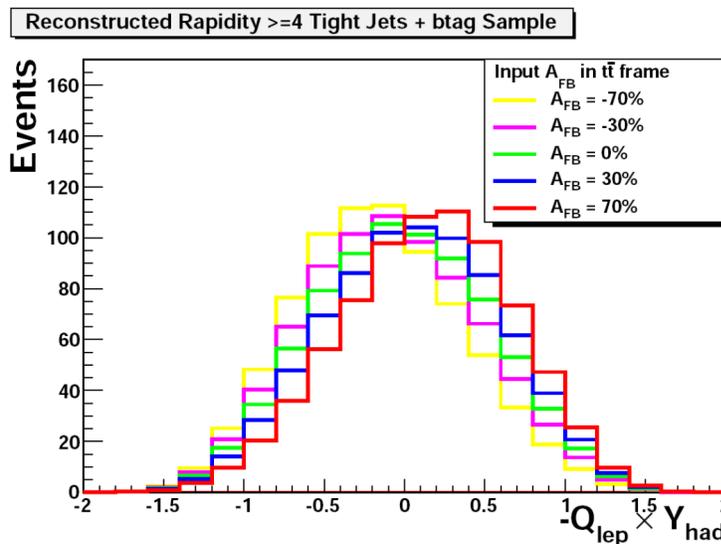
$$\sigma_{syst} = \pm 0.024 \quad (36)$$

By adding the various systematic uncertainties listed above (background size/shape, signal parameters, shape uncertainty) in quadrature, we arrive at this final value:

$$A_{fb}^{corrected} = 0.193 \pm 0.065 \pm 0.024 \quad (37)$$

Systematic	Uncertainty
background size	0.018
background shape	0.011
ISR/FSR	0.008
JES	0.002
PDF	0.001
MC Generator	0.003
Shape/Unfolding	0.007
Total Uncertainty	0.024

Table 3: Summary of Systematic Uncertainties

Figure 11: Template shapes for various values of $A_{fb}^{\bar{t}t}$

11 Test of Linear Hypothesis

The unfold correction method presented above is model independent – the reweight function only enters into our goodness test (see Sections 8 and 9) and shape systematic uncertainty, not the unfold matrices themselves. However, now that we’ve demonstrated that there *is* an asymmetry, we do wish to understand the shape.

We are conveniently prepared to test the simplest hypothesis, an asymmetry linear in $\cos(\theta)$, namely $1 + A \cdot \cos(\theta)$, by using the machinery created in Section 8. We generate reweighted signal shapes with this machinery, add on the background shapes, and compare these templates directly with data. See Figure 11 above for a plot of our template shapes, each reweighted by a different A_{fb} value in the $t\bar{t}$ rest frame.

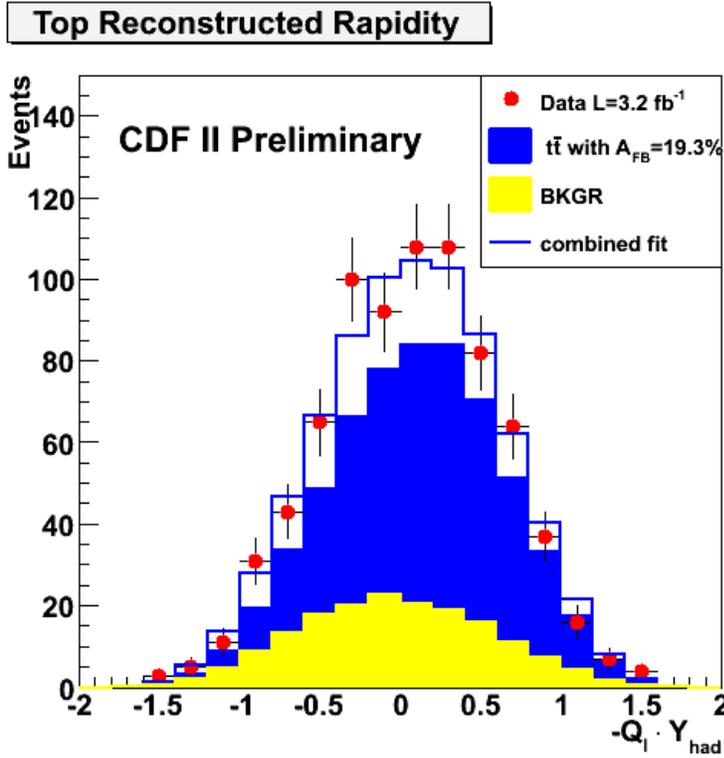


Figure 12: Comparison of data, backgrounds, signal, and the $A_{fb} = 19.3$ template

We plot $-\log(\text{likelihood})$ of the templates with respect to the data and find the minimum, which corresponds to a best fit A_{fb} . Despite using the $t\bar{t}$ rest frame for our parameter A in the template construction, our reweight method allows us to measure the true A_{fb} in the lab frame as well. Using these values we can calculate a best fit A_{fb} value for the lab frame and compare with our value obtained through the unfold method. This results in a template measurement

$$A_{fb}^{lab} = 0.173 \pm 0.052 \quad (38)$$

We note that this value is consistent with our unfold method value $A_{fb} = 19.3\%$. A plot of data, background, signal, and a template shape generated with this value is shown above in Figure 12. We see that this template agrees well with the raw data.

Finally, we can examine the pull distribution using this best fit A_{fb} . We found that the pull distributions for a template generated with parameter $A = 0.30$ had a Gaussian shape, centered near 0, having a width near 1. This is characteristic of a good fit and validates our linear hypothesis.

12 Check of Reweighted Signal Reconstructed Variables

As a final check of our analysis we use the reweight method explained above on other variables besides the $-Q_l * y_{had}$ distribution. We were interested to see if our reconstructed data would agree better with a reweighted signal (using $A_{fb}^{lab} = 19\%$) than the nominal signal. Below are plots showing $-Q_l * y_{hadTop}$, $-Q_l * y_{hadW}$, and $-Q_l * y_{hadB}$, comparing the initial signal shape (left plots) with reweighted versions (right plots). The solid yellow shape are background events, solid purple are signal events, and the points are data. We see that the reweighted shapes are indeed better.

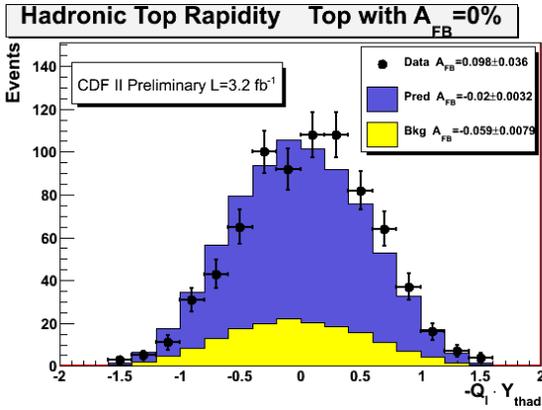


Figure 13: Hadronically-decaying Top

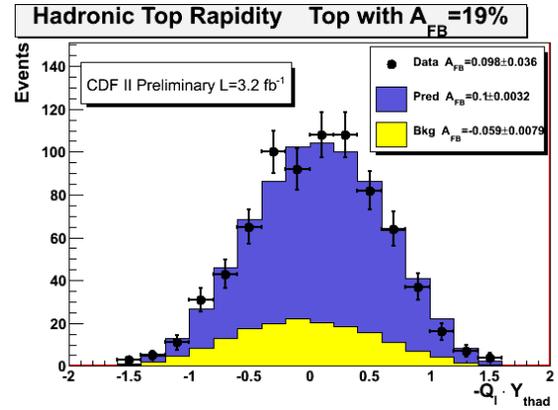


Figure 14: Reweighted Signal

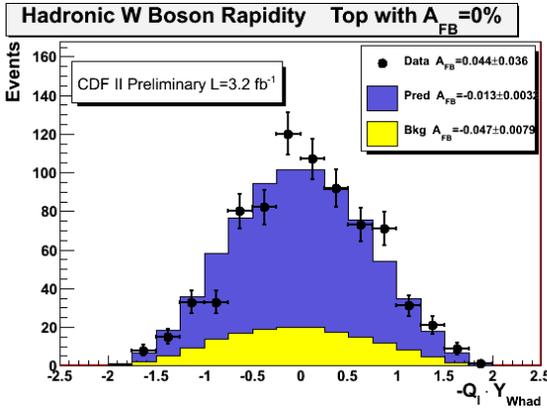


Figure 15: Hadronically-decaying W Boson

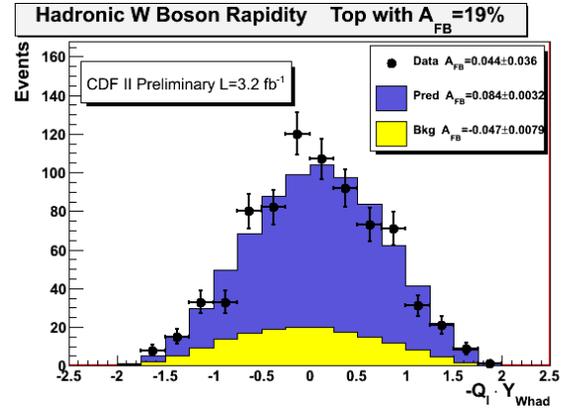


Figure 16: Reweighted Signal

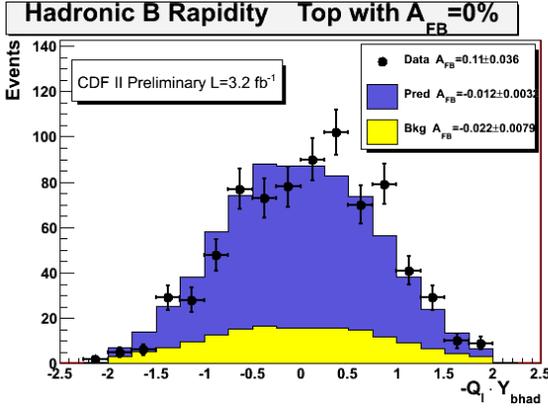


Figure 17: Hadronically-decaying B Quark

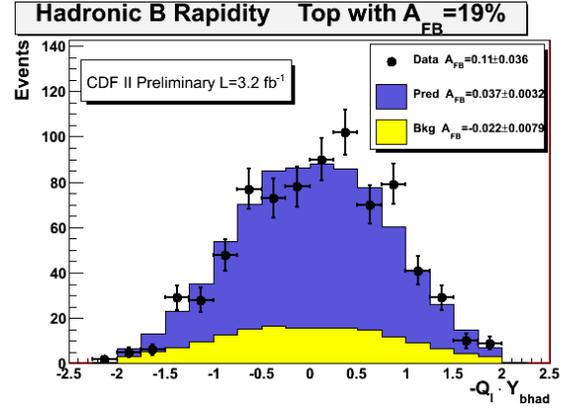


Figure 18: Reweighted Signal

13 Results and Conclusion

We have updated our previous measurement to include data through 3.2 fb^{-1} of CDF II preliminary data. Additionally, we have changed variables to measure y_{top} instead of $\cos(\theta_{top})$. Finally, we have added a template technique to directly compare reweighted shapes with our raw data. We found this template method to agree with our model-independent unfold method. Our front-back asymmetry is measured as

$$A_{fb}^{unfold} = 0.193 \pm 0.065 \pm 0.024 \quad (39)$$

and the template method result supports our hypothesis that the forward-backward asymmetry can be modeled by a $1 + A \cos(\theta)$ reweight in the $t\bar{t}$ rest frame.

References

- [1] T. Aaltonen *et al.*, CDF Collaboration, “Forward-Backward Asymmetry in Top-Quark Production in $p\bar{p}$ Collisions at $\sqrt{s}=1.96$ TeV”, Phys. Rev. Lett. 101 (2008), 202001
- [2] O. Antunano, J. H. Kuhn, and G. Rodrigo, Phys. Rev. D 77, 014003 (2008);
- [3] M. T. Bowen, S. Ellis, and D. Rainwater, Phys. Rev. D 73, 014008 (2006);
S. Dittmaier, P. Uwer, and S. Weinzierl, Phys. Rev. Lett. 98, 262002 (2007);
L. G. Almeida, G. Stermann, and W. Vogelsang, Phys. Rev. D 78, 014008 (2008).
- [4] CDF Collaboration, “Measurement of the Top Pair Cross Section in the Lepton + Jets Channel with 2.7 fb^{-1} ”, CDF-Note 9462
http://www-cdf.fnal.gov/physics/new/top/confNotes/cdf9462_xsec_secvtx_3invfb.pdf
- [5] F. James, “MINUIT: Function Minimization and Error Analysis Reference Manual”, <http://wwwasdoc.web.cern.ch/wwwasdoc/minuit/minmain.html>, Computing And Networks Division, CERN (1998).