The following full text is a preprint version which may differ from the publisher's version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/132588

Please be advised that this information was generated on 2017-08-30 and may be subject to change.
Measurement of the forward–backward asymmetry in top quark–antiquark production in $p\bar{p}$ collisions using the lepton+jets channel

We present a measurement of the forward–backward asymmetry in top quark–antiquark production using the full Tevatron Run II dataset collected by the D0 experiment at Fermilab. The measurement is performed in lepton+jets final states using a new kinematic fitting algorithm for events with four or more jets and a new partial reconstruction algorithm for events with only three jets. Corrected for detector acceptance and resolution effects, the asymmetry is evaluated to be \( A_{FB} = (10.6 \pm 3.0)\% \). Results are consistent with the standard model predictions which range from 5.0\% to 8.8\%. We also present the dependence of the asymmetry on the invariant mass of the top quark–antiquark system and the difference in rapidities of top quark and antiquark.

PACS numbers: 14.65.Ha,12.38.Qk,11.30.Er,13.85.-t

I. INTRODUCTION

A. Motivation and definitions

Over the last five years both experiments at the Fermilab Tevatron Collider measured positive forward–backward asymmetries in the production of top quark–antiquark pairs in proton–antiproton collisions \((p\bar{p} \rightarrow t\bar{t})\) \([1,2]\). The reported values were consistently above predictions of the standard model of particle physics (SM) \([3,4]\). In particular, the CDF Collaboration observed a strong rise of the asymmetry with the invariant mass of the \(t\bar{t}\) system, \(m_{t\bar{t}}\) \([5]\). The dependence of the asymmetry on \(m_{t\bar{t}}\) in D0 data, as measured in Ref. \([3]\), was statistically compatible with both the SM predictions and with the CDF result. Several beyond-the-SM scenarios were suggested to explain the measured \(A_{FB}\) values \([6]\), in particular using the framework of parity-violating strong interactions suggested in Ref. \([7]\). In this paper we report new results from the D0 experiment based on the full dataset collected during Run II of the Fermilab Tevatron Collider, which supersede the result of Ref. \([4]\).

In proton–antiproton collisions, top quark–antiquark pairs are predominantly produced via valence quark–antiquark annihilation. Thus, the direction of the proton (antiproton) almost always coincides with the direction of the incoming quark (antiquark). We define the difference in rapidity\(^1\) between the top quark \(y_t\) and antiquark \(y_{\bar{t}}\):

\[
\Delta y = y_t - y_{\bar{t}}. \tag{1}
\]

We refer to the events that have \(\Delta y > 0\) as “forward”, and to those with \(\Delta y < 0\) as “backward”. The forward–backward asymmetry in \(t\bar{t}\) production is defined as

\[
A_{FB} = \frac{N_f - N_b}{N_f + N_b}, \tag{2}
\]

where \(N_f\) (\(N_b\)) is the number of forward (backward) events. All \(t\bar{t}\) asymmetries reported in this paper are given after subtracting the contributions from background processes.

The rapidities of the \(t\) and \(\bar{t}\) quarks and the corresponding asymmetries can be defined at the production level (sometimes denoted as generator level or parton level), when the kinematic parameters of the generated top quarks are used. Unless stated otherwise the production-level asymmetries are defined for all signal events without imposing the selection criteria of this analysis. The

---

\(^1\) The rapidity \(y\) is defined as \(y = \frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right)\), where \(E\) is the particle’s energy and \(p_z\) is its momentum along the \(z\)-axis, which corresponds to the direction of the incoming proton.
rapidity and asymmetry can also be defined at the reconstruction level, using the reconstructed kinematics of the selected events. Similarly, the invariant mass of the $t\bar{t}$ system can be defined at the production and reconstruction levels.

B. Strategy

In the SM a top quark almost always decays to a $b$ quark and a $W$ boson, which decays either leptonically or hadronically. In this paper we identify $t\bar{t}$ events using the $t\bar{t} \rightarrow W^+bW^-\bar{b}$; $W^+ \rightarrow l^+\nu$; $W^- \rightarrow q\bar{q}'$ (and charge conjugates) decay chain. This channel is commonly referred to as the “lepton+jets” (l+jets) channel. We select events that contain one isolated lepton (electron or muon) of high transverse momentum ($p_T$) and at least three jets. The electric charge of the lepton identifies the electric charge of the leptonically decaying $W$ boson and its parent top quark. The other top quark is assumed to have the opposite charge. The event selection, sample composition determination, and modeling of the signal and background processes are identical to those used in the measurement of the leptonic asymmetry in $t\bar{t}$ production in the l+jets channel \cite{10}. The four-vectors of the top quarks and antitop quarks in the events containing at least four jets are reconstructed with a kinematic fitting algorithm, while for the events that contain only three jets a partial reconstruction algorithm is used. If a jet exhibits properties consistent with a jet originating from a $b$ quark, such as the presence of a reconstructed secondary vertex, we identify it as a $b$-tagged jet \cite{11}. The l+jets events are separated into channels defined by jet and $b$-tag multiplicities. The amount of signal and the forward–backward asymmetry at the reconstruction level are determined using a simultaneous fit to a kinematic discriminant in these channels.

The measured background-subtracted one-dimensional (1D) distribution in $\Delta y$ is corrected to the production level (“unfolded”). From this distribution we calculate the fully-corrected $A_{FB}$ as well as its dependence on $\Delta y$. To study the dependence of the asymmetry on the invariant mass of the $t\bar{t}$ system, unfolding is done on the background-subtracted data distributions in two dimensions (2D: $\Delta y$ vs $m_{t\bar{t}}$). The signal channels are unfolded simultaneously to yield the desired 1D or 2D production-level distributions, from which the production-level $A_{FB}$ values are computed using Eq. 2. The procedure is calibrated using simulated samples with varied asymmetries and input distributions in $\Delta y$ and $m_{t\bar{t}}$. The statistical and systematic uncertainties of the results are evaluated using ensembles of simulated pseudo-datasets (PDs).

II. D0 DETECTOR

We use the data collected by the D0 detector during Run II of the Tevatron in the years 2001–2011. After imposing event quality requirements ensuring that all detector systems were fully operational, this dataset corresponds to an integrated luminosity of 9.7 fb$^{-1}$. The D0 detector is described in detail elsewhere \cite{12}. The central tracking system, consisting of a silicon microstrip tracker and a scintillating fiber tracker, is enclosed within a 1.9 T superconducting solenoid magnet. Tracks of charged particles are reconstructed within a detector pseudorapidity region$^2$ of $|\eta| < 2.5$. Electrons, photons, and jets of hadrons are identified \cite{13} using a liquid-argon and uranium-plate calorimeter, which consists of a central barrel covering up to $|\eta| \approx 1.1$, and two endcap sections that extend coverage to $|\eta| \approx 4.2$ \cite{14}. Central and forward preshower detectors are positioned in front of the corresponding sections of the calorimeter. A muon system consisting of layers of tracking detectors and scintillation counters placed in front of and behind 1.8 T iron toroids \cite{15} identifies muons \cite{10} within $|\eta| < 2$. Luminosity is measured using arrays of plastic scintillators located in front of the endcap calorimeter cryostats. A three-level trigger system selects interesting events at the rate of $\approx 200$ Hz for offline analysis \cite{17}.

III. EVENT SELECTION AND MODELING

Object reconstruction and identification, as well as event selection, are the same as in Ref. \cite{10} and are briefly outlined in this section. We select events with exactly one isolated electron within the detector pseudorapidity range of $|\eta| < 1.1$ or one isolated muon within $|\eta| < 2.0$, and at least three jets within $|\eta| < 2.5$. To limit the possible contribution of poorly modeled background due to multijet production, leptons of either flavor are required to have $p_T > 1.5$. The presence of a neutrino is inferred from a transverse momentum imbalance, which is measured primarily using calorimetry and is referred to as the “missing transverse energy” $E_T$. All selected objects are required to have transverse momentum $p_T > 20$ GeV, and the jet with the largest $p_T$ (the leading jet) is also required to have $p_T > 40$ GeV.

To identify jets that are likely to be associated with $b$ quarks, we perform a multivariate analysis (MVA) that combines variables characterizing the properties of secondary vertices and of tracks with large impact parameters relative to the primary $p\bar{p}$ interaction vertex (PV) \cite{11}. The output of the MVA is a continuous variable, MVA$_b$. The requirement on MVA$_b$ ($b$ tagging) used in this analysis has an efficiency of about 64% for identifying $b$ jets originating from top quark decay, and a misidentification probability of about 7% for jets that

$^2$ The detector pseudorapidity $\eta$ is defined as $-\ln [\tan (\frac{\theta}{2})]$, where $\theta$ is the polar angle measured with respect to the center of the detector. The angle $\theta = 0$ corresponds to the direction of the incoming proton.
do not contain heavy flavor quarks and are produced in association with leptonically decaying $W$ bosons.

We simulate $t\bar{t}$ production using MC@NLO program (version 3.4) \cite{18} with the parton showering performed by HERWIG \cite{19}. This simulation is fully integrated with the D0 software, allowing for detailed studies of the kinematic dependences of $A_{FB}$ and their interplay with selection and reconstruction effects. The main source of background to the $t\bar{t}$ signal is the production of a leptonically decaying $W$ boson in association with jets ($W$+jets). The kinematic properties of this process are simulated using ALPGEN \cite{20} with hadronic showering performed by PYTHIA \cite{21}. For signal and background modeling we use the CTEQ6.1 set of parton distribution functions (PDFs) \cite{22}. The normalization of the $W$+jets contribution is a free parameter in the fitting procedure described below. Events with multiple jets can also mimic $t\bar{t}$ signal when a particle from one of the jets is misidentified as an isolated lepton. The normalization of this multijet background is extrapolated from a control sample enriched in this process using the probability for a jet to satisfy the lepton-quantity requirements \cite{23}. For the other backgrounds, $Z$+jets events are simulated with ALPGEN, diboson events are simulated with PYTHIA, and events from single-top-quark production are simulated with COMPHEP \cite{24}. The normalizations for the last three background processes are taken from NLO calculations \cite{25}. In all cases, event generation is followed by the GEANT-based D0 detector simulation \cite{26}. To model energy depositions from noise and additional $p\bar{p}$ collisions within the same bunch crossing, simulated events are overlaid with data from random $p\bar{p}$ crossings. All simulated events are reconstructed using the same code as for the reconstruction of the collider data.

IV. RECONSTRUCTION OF THE EVENT KINEMATICS

To measure the $t\bar{t}$ forward–backward asymmetry and its dependence on the invariant mass of the $t\bar{t}$ system we need to determine the four-vectors of top quark and antitop quark, which is done by summing the four-vectors of their decay products in the $t\bar{t} \rightarrow W^+bW^-\bar{b}; W^+ \rightarrow l^+\nu_l; W^- \rightarrow q\bar{q}'$ (and charge conjugates) decay chain. There are four final state quarks in this decay chain, while we select events that contain one isolated lepton and at least three jets. When an event contains at least four jets we assume that the four jets with the largest $p_T$ originate from the quarks from $t\bar{t}$ decay. If an event contains only three jets one of the jets from $t\bar{t}$ decay is missing. In either case all possible assignments of three or four jets to the final state quarks are used, with the likelihood of each assignment evaluated by the reconstruction algorithms described below.

For $l+\geq 4$ jet events, the $t\bar{t}$ system is fully reconstructed using a kinematic fitting algorithm. Previous D0 top quark analyses used the algorithm of Ref. \cite{27}. In this paper a new algorithm is employed, which utilizes an analytic solution for the neutrino momentum using the constraints on the $W$-boson ($M_W$) and top-quark masses ($m_t$) \cite{28}. The likelihood term for each jet-to-quark assignment accounts for the differences between the observed jet energy and the energy scaled to satisfy the constraints on $M_W$ and $m_t$. The jet energy resolution and the probability for a jet to be reconstructed (see “Type III” transfer function in Ref. \cite{29}) are taken into account. The $b$-tagging observables MVA$_b$ are also used to evaluate the likelihood of each assignment.

For $l+3$ jet events, a partial reconstruction algorithm of the $t\bar{t}$ decay chain is employed \cite{30}. With one jet entirely lost, no significant improvement is expected from scaling the four-vectors of the remaining objects as is done by the kinematic fitting algorithm in $l+\geq 4$ jet events, so the partial reconstruction algorithm does not attempt to modify the kinematics of the observed objects. As only the transverse components of the neutrino momentum are measured in $E_T$, the longitudinal component is calculated using a quadratic equation which results from imposing the $M_W$ constraint on the $W \rightarrow l\nu$ decay products. The two-fold ambiguity is resolved by choosing the solution that minimizes the difference between the known $m_t$ and the invariant mass of the objects assigned to the leptonic top quark decay, $m_l$. This algorithm thus assumes that the jet associated with the $b$ quark from the leptonically decaying top quark is detected. This assumption holds for 80% of the $t\bar{t}$ events. The lost jet is assumed to be associated with either a light quark or a $b$ quark from the hadronic top-quark decay chain. In the majority of cases (74%) this jet is lost due to its low energy, so this loss has little effect on the kinematics of the hadronically decaying top quark. The lost jet is neglected in the partial reconstruction algorithm. The sum of the four-vectors of the two jets assigned to the products of the hadronically decaying top quark serves as a proxy for the four-vector of the hadronically decaying top quark with the invariant mass $m_p$. Even though $m_p$ is not expected to be equal to $m_t$, the distribution in this variable is different for combinations correctly associated with the hadronically decaying top quark and combinations that include a $b$ jet from the leptonically decaying top quark. In each event we consider the following nine observables: the MVA$_b$ for each of the three jets, the three possible $m_l$, corresponding to the three possible lepton-neutrino-jet combinations, and the three possible $m_p$. The likelihood of each of the three possible jet-to-quark assignment is calculated by evaluating the consistency of the nine observables with the distributions corresponding to the hypothesized assignment. In particular, the jet hypothesized to be associated with a $b$ quark should have a value of MVA$_b$ consistent with the one expected for $b$ jets, while for a jet hypothesized to originate from a $W$ boson decay MVA$_b$ should be consistent with the distribution expected for such jets. The values of $m_l$ and $m_p$ for the jet combinations that correspond to the hypothesized assignment should be consistent with the
distributions expected for correctly assigned jets, while
the values of \( m_\ell \) and \( m_p \) for the other jet combinations
should agree with the distributions expected for wrong
assignments. When calculating \( m_{\ell t} \), we compensate for
the effect of the lost jet by applying an \( m_p \)-dependent scaling to the four-vector of the hadronically decaying
top quark.

Unlike the \( A_{FB} \) measurement in Ref. [4], where only
the jet-to-quark assignment with the lowest \( \chi^2 \) was used,
we reconstruct \( \Delta y \) by averaging its values over all possi-
ble assignments, weighted by their likelihoods evaluated
as described above for \( \ell+\geq 4 \) jet and \( \ell+3 \) jet events. The
same approach is used to reconstruct \( m_{\ell t} \) in the \( \ell+3 \) jet
channel. For \( \ell+4 \) jet events, \( m_{\ell t} \) is reconstructed us-
ing the outputs of three reconstruction algorithms: the
new kinematic fit algorithm, the kinematic fit algorithm
of Ref. [27], and a simple reconstruction algorithm [31]
that evaluates the kinematics of the leptonically decay-
ing \( W \) boson from the lepton and the neutrino by im-
posing the \( M_W \) constraint and calculates \( m_{\ell t} \) by adding
the four most energetic jets without imposing the \( m_t \)
constraint. The likelihood values calculated by the algo-
rithms give indications on how well the kinematics of a
particular event match the assumptions made by a given
algorithm. In particular, for high \( m_{\ell t} \) there is a higher
probability that two final state quarks are associated with
the same jet. Such a jet is likely to be the most energetic
jet in the event and have a large mass. The simple re-
construction algorithm, which does not assume a specific
jet-to-quark assignment, performs best for such events.
We use a multivariate regression [32] to combine the par-
tially correlated \( m_{\ell t} \) values and the likelihoods produced
by the three algorithms with supplementary observables
such as the mass of the leading jet to estimate \( m_{\ell t} \). This
combined \( m_{\ell t} \) reconstruction outperforms the individual
algorithms in all \( m_{\ell t} \) ranges.

For the asymmetry measurement the performance of a
\( t\bar{t} \) reconstruction algorithm can be characterized by the
probability \( P_c \) to correctly reconstruct the sign of \( \Delta y \).
For the algorithm employed in this analysis for \( \ell+\geq 4 \) jet
events \( P_c = 0.775 \), compared to \( P_c = 0.756 \) for the algo-
rithm of Ref. [27]. The partial reconstruction algorithm
achieves \( P_c = 0.754 \) for \( l+3 \) jet events. The dependence
of \( P_c \) on the production-level \( |\Delta y| \) is shown in Fig. 1 for
these three algorithms. The high values of \( P_c \) achieved by
the partial reconstruction algorithm, which are almost as
high as \( P_c \) for \( l+\geq 4 \) jet events, can be understood from
the following consideration. All four leading jets are asso-
ciated with the quarks from the \( t\bar{t} \) decay in only 55% of
the \( l+\geq 4 \) jet events. For the other 45% of the events one
of the jets originates from initial or final state radiation,
which can lead to badly misreconstructed \( t\bar{t} \) four-vectors.
Only 4% of the \( l+3 \) jet events contain a jet that does not
originate from the four quarks of the \( t\bar{t} \) decay. Thus,
even though some information is lost with the unrecon-
structed jet, no wrong information is added, leading to a
low probability to misreconstruct the sign of \( \Delta y \).

V. SM PREDICTIONS FOR \( A_{FB} \)

The differential cross section for \( t\bar{t} \) production is avail-
able only at order \( \alpha_s^3 \), where \( \alpha_s \) is the strong coupling
constant. Since in the SM the \( t\bar{t} \) asymmetry only appears
at this order, no full higher order prediction for
\( A_{FB} \) exists yet. The relative uncertainty on the \( \alpha_s^3 \) calcu-
lation of the asymmetry due to higher order corrections
is evaluated to be as large as \( \approx 25\% \) [33].

Recently the order \( \alpha_s^4 \) calculation for the total cross
section of \( t\bar{t} \) production [34] was made available, but the
asymmetry was not computed at this order. Several pa-
ers report calculations of the leading corrections to the
asymmetry with the predicted \( A_{FB} \) values ranging from
5.0% in \( MC@NLO \) to 8.8% once the electroweak correc-
tions and resummations of particular regions of phase
space [35] are taken into account. The dominant uncer-
tainty on these predictions is from the renormalization
and factorization scales, and is evaluated to be as high
as 2.0% (absolute) [33, 37]. The authors of Ref. [36] ob-
tain a value of \( A_{FB} = 12.7\% \) by choosing a normalization
scale that arguably stabilizes the perturbative expansion
yet differs significantly from the scales commonly used
in top quark physics calculations. Some authors suggest
that the corrections from interactions between the top
quark decay products and the proton remnants should
also be taken into account when calculating \( A_{FB} \) [39].
Given this variety of predictions, we choose to compare
our data to the well defined \( MC@NLO \) simulation.

At order \( \alpha_s^3 \), the QCD contributions to the asymme-
try in \( t\bar{t} \) production can be divided into two classes up
to divergences that cancel between these two classes [6].
The first class, which contributes to negative asymmetry,
is a result of interference between the terms that contain
gluon radiation in the initial or final states, which may
result in an extra jet in the event and typically leads to

\[
D\bar{O}, 9.7 fb^{-1}
\]

\[
0.5 \quad 0.6 \quad 0.7 \quad 0.8 \quad 0.9 \quad 1
\]

\[
0 \quad 0.5 \quad 1 \quad 1.5 \quad 2 \quad 2.5 \quad 3
\]

\[ P_c \]

-4 jet, PRD 84, 112005
-4 jet, new kinematic fit
3 jet, partial reconstruction

FIG. 1: (Color online). The probability to correctly recon-
struct the sign of \( \Delta y \) as a function of the production-level
\( |\Delta y| \) for the algorithm of Ref. [27] used to measure the \( A_{FB} \)
in Ref. [4] and the algorithms used to reconstruct \( l+\geq 4 \) jet
events and to partially reconstruct \( l+3 \) jet events in this pa-
per.
a higher transverse momentum of the $\bar{t}t$ system. The second class, which contributes to positive asymmetry, is from interference between the Born term ($\alpha_s^2$) and the term described by a box diagram ($\alpha_s^4$). The overall asymmetry is positive and depends on the jet multiplicity. Selection criteria that give preference to events with higher jet multiplicity favor the first class of events and further lower the overall expected asymmetry, while a higher asymmetry is expected for events with lower jet multiplicity. Consequently, forward events tend to have fewer jets than events in the backward category. Similarly, since a $b$-tagged jet is less likely to originate from initial or final state radiation, samples with a larger number of $b$ tags tend to have higher values of $A_{FB}$.

$\text{MC@NLO}$ predicts an overall asymmetry in $\bar{t}t$ production before selection of $(5.01 \pm 0.03)\%$. Here, and in the following sections, the quoted uncertainties on the predictions are from the finite size of the simulated samples unless otherwise stated. Table I lists the $\text{MC@NLO}$ predictions for $\bar{t}t$ events after the selection criteria are applied.

All previous measurements of $A_{FB}$ in the $l+\text{jets}$ channel selected $\bar{t}t$ events that had at least four jets in the final state. As is apparent from Table I restricting the selection to only $l+\geq 4$ jet events lowers the production-level asymmetry. Including events with three jets reduces this selection bias.

### TABLE I: Asymmetries predicted by $\text{MC@NLO}$ for $\bar{t}t$ events that pass the analysis selection criteria. Statistical uncertainties only.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Production level $A_{FB}$, %</th>
<th>Reconstruction level $A_{FB}$, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\geq3$ jets, $\geq1$ $b$ tags</td>
<td>$4.7 \pm 0.1$</td>
<td>$3.9 \pm 0.1$</td>
</tr>
<tr>
<td>$3$ jets, $1$ $b$ tag</td>
<td>$6.6 \pm 0.2$</td>
<td>$4.7 \pm 0.3$</td>
</tr>
<tr>
<td>$3$ jets, $\geq2$ $b$ tags</td>
<td>$7.3 \pm 0.2$</td>
<td>$5.6 \pm 0.2$</td>
</tr>
<tr>
<td>$\geq4$ jets, $1$ $b$ tag</td>
<td>$1.4 \pm 0.2$</td>
<td>$1.9 \pm 0.2$</td>
</tr>
<tr>
<td>$\geq4$ jets, $\geq2$ $b$ tags</td>
<td>$3.2 \pm 0.1$</td>
<td>$3.3 \pm 0.2$</td>
</tr>
</tbody>
</table>

Asymmetries after reconstruction are presented in the last column of Table I. Finite resolution in $\Delta y$ results in roughly 20% of the forward events being misreconstructed as backward, and vice versa. Since there are more forward events, $\Delta y$ smearing leads to an overall lowering of the reconstructed asymmetries. At the same time, forward $\bar{t}t$ events, which tend to have fewer jets, have a lower probability to be misreconstructed, resulting in fewer migrations into the backward category, and an upward shift in the reconstructed asymmetry. This bias is most apparent in the $l+\geq 4$ jet, one-$b$-tag channel, where the lowest asymmetry is predicted.

## VI. SAMPLE COMPOSITION AND RECONSTRUCTION-LEVEL $A_{FB}$

Reconstructed events are divided into six channels by the number of jets and $b$ tags: $l+3$ jet and $l+\geq 4$ jet with 0, 1, and $\geq 2$ $b$ tags each. The $l+3$ jet zero-$b$-tag channel is used only for the background asymmetry calibration, and not for the $\bar{t}t$ asymmetry measurement. The $l+\geq 4$ jet zero-$b$-tag channel is used only for determining the sample composition and the reconstruction-level $A_{FB}$, and is not used for measuring the production-level asymmetry.

Several well-modeled variables that have different distributions for signal and background processes, and that have minimal correlations between each other and with $\Delta y$ and $m_{t\bar{t}}$, are combined into kinematic discriminants bounded between 0 and 1 \cite{ref1}. For $l+\geq 4$ jet events a discriminant $D_4$ is built from the following input variables:

- $\chi^2$ – the test statistic of the likeliest assignment from the kinematic fit.
- $p_T^{lB}$ – the transverse momentum of the leading $b$-tagged jet, or when no jets are $b$ tagged, the $p_T$ of the leading jet.
- $k_{\text{min}} = \min (p_{T,a}, p_{T,b}) \cdot \Delta \mathcal{R}_{ab}$, where $\Delta \mathcal{R}_{ab} = \sqrt{(\eta_a - \eta_b)^2 + (\phi_a - \phi_b)^2}$ is the angular distance\(^3\) between the two closest jets, $a$ and $b$, and $p_{T,a}$ and $p_{T,b}$ are their transverse momenta.
- $M_{jj}$, the invariant mass of the jets corresponding to the $W \rightarrow q\bar{q}'$ decay in the likeliest assignment from the kinematic fit, calculated using kinematic quantities before the fit.

The variables $\chi^2$ and $M_{jj}$ are based on the full $\bar{t}t$ reconstruction using the kinematic fitting technique of Ref. \cite{ref27}.

For the $l+3$ jet events we construct a discriminant $D_3$ using a different set of input variables:

- $S$ – the sphericity \cite{ref40}, defined as $S = \frac{4}{3}(\lambda_2 + \lambda_3)$, where $\lambda_2$ and $\lambda_3$ are the two largest of the three eigenvalues of the normalized quadratic momentum tensor $M$. The tensor $M$ is defined as

$$M_{ij} = \frac{\sum_o \mathbf{p}_i^o \mathbf{p}_j^o}{\sum_o |\mathbf{p}_o|^2}, \quad (3)$$

where $\mathbf{p}_o$ is the momentum vector of a reconstructed object $o$, and $i$ and $j$ run over the three indices for the Cartesian coordinates. The sum over objects includes the three selected jets and the selected charged lepton.

\(^3\) Here the pseudorapidity $\eta$ and the azimuthal angle $\phi$ are defined relative to the PV.
The discriminants for all channels are combined into a single discriminant $D_c$, so that for the $l+3$ jet events $D_c = N_{\text{tag}} + D_3$, while for $l+4$ jet events $D_c = 3 + N_{\text{tag}} + D_4$. The variable $N_{\text{tag}}$ above is usually taken to be equal to the number of $b$-tagged jets in the event, but for events with more than two $b$-tagged jets $N_{\text{tag}} = 2$ instead. We fit the sum of the signal and background templates to the data distribution in the discriminant $D_c$ as shown in Fig. 2. This fit is identical to the fit for the sample composition in Ref. 10. The sample composition and its breakdown into individual channels are summarized in Table II. Background contributions other than $W+\text{jets}$ and multijet production are labeled “Other $\text{Bg}$” in Table II.

In the simulated $W+\text{jets}$ background, the angular distribution of leptons from $W$-boson decay has a forward-backward asymmetry, which is in part tuned to Tevatron data 11. Due to this asymmetry, when these events are reconstructed according to the $t\bar{t}$ hypothesis, there remains a residual asymmetry of $\approx 5\%$ in the $\Delta y$ distribution. To improve the modeling of this asymmetry, we apply a weight to each simulated $W+\text{jets}$ event which depends on the product of the generated lepton charge and its rapidity. These weights are chosen so that the simulation best matches control data with three jets and zero $b$ tags as in Ref. 10. The difference in the $\Delta y$ distributions predicted by the simulation with and without the applied weights is treated as a source of systematic uncertainty due to background modeling. This uncertainty exceeds the uncertainty due to PDFs by about a factor of two. We rely on the simulation to predict the variation of the asymmetry in $W+\text{jets}$ events with jet and $b$-tag multiplicities.

The distributions of the reconstructed $\Delta y$ are shown in Fig. 3. The $t\bar{t}$ asymmetry at the reconstruction level is extracted using a fit to the distributions in the discriminant $D_c$ and sign of $\Delta y$, excluding the $l+3$ jet events with zero $b$ tags. This fitting procedure is identical to the procedure used in Ref. 10. The inclusive asymmetry measured at the reconstruction level is $(7.9 \pm 2.3)\%$. The results for individual channels are listed in Table III.

### Table II: Estimated number of events from the fit of the data distribution in the discriminant $D_c$ to the sum of signal and background processes (see Fig. 2). The sum of the estimated number of signal and background events is constrained to be equal to that in data. The “Selected events” column includes the $l+3$ jet events with at least one $b$ tag and all $l+4$ jet events. The statistical uncertainties from the fit are quoted. We also present the event breakdown for the channels with at least one $b$ tag, which are used to determine the production-level $A_{\text{FB}}$. Table is from Ref. 11.

<table>
<thead>
<tr>
<th>Source events</th>
<th>3 jets</th>
<th>$\geq 4$ jets</th>
</tr>
</thead>
<tbody>
<tr>
<td>W+jets</td>
<td>4447 ± 74</td>
<td>2461 ± 2466</td>
</tr>
<tr>
<td>Multijet</td>
<td>969 ± 54</td>
<td>449 ± 95</td>
</tr>
<tr>
<td>Other $\text{Bg}$</td>
<td>786 ± 404</td>
<td>112 ± 75</td>
</tr>
<tr>
<td>Signal</td>
<td>4745 ± 570</td>
<td>2752 ± 1001</td>
</tr>
<tr>
<td>Sum</td>
<td>10947 ± 570</td>
<td>4526 ± 1560</td>
</tr>
<tr>
<td>Data</td>
<td>10947</td>
<td>4526 ± 1560</td>
</tr>
</tbody>
</table>

- $p_T^{3\text{rd}}$ — the transverse momentum of the third leading jet.
- $M_{jj}^{\text{min}}$ — the lowest of the invariant masses of two jets, out of the three possible jet pairings.
- $p_T^{B\text{B}}$, defined as for the $l+4$ jet channel, above.
- $\Delta \phi(\text{jet}_1, \text{E}_T)$, the difference in azimuthal angle between the leading jet and $\text{E}_T$.

The distributions of the reconstructed invariant mass of the $t\bar{t}$ system are shown in Fig. 3. Since the $l+3$ jet and $l+4$ jet channels have different response (both mean and shape) for $m_{t\bar{t}}$, the dependence of $A_{\text{FB}}$ on $m_{t\bar{t}}$ at the reconstruction level is difficult to interpret and is not presented here. The measurement of production-level $A_{\text{FB}}$ and its dependence on $m_{t\bar{t}}$ is described in Section VII.

We use the results of the sample composition study summarized in Table II to normalize the distributions for the background processes in the sensitive variables ($\Delta y$, and also $m_{t\bar{t}}$ for the 2D measurement), which are subtracted from the distributions observed in data. To increase the signal purity of the data used in the fully corrected measurements, we unfold only events containing at least one $b$ tag. The background-subtracted distributions of the sensitive variables in the corresponding four channels are used as inputs to the unfolding procedure.

### Table III: Reconstruction-level background-subtracted asymmetries for selected events for different channels. The last row includes the channels listed in the rows above and the $l+4$ jet, zero-$b$-tag channel. The first uncertainty is statistical, and the second one is systematic. Systematic uncertainties are discussed in Section VIII. The prediction is based on the MC@NLO simulation.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Predicted</th>
<th>Measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 jets, 1 $b$ tag</td>
<td>4.7</td>
<td>5.4 $\pm 6.0^{+3.3}_{-4.0}$</td>
</tr>
<tr>
<td>$\geq 2$ b tags</td>
<td>5.6</td>
<td>10.7 $\pm 4.2 \pm 0.8$</td>
</tr>
<tr>
<td>$\geq 3$ jets, 1 $b$ tag</td>
<td>1.9</td>
<td>11.0 $\pm 4.4 \pm 0.8$</td>
</tr>
<tr>
<td>$\geq 4$ jets, $\geq 2$ b tags</td>
<td>3.3</td>
<td>5.9 $\pm 3.3 \pm 0.1$</td>
</tr>
<tr>
<td>Combined</td>
<td>3.6</td>
<td>7.9 $\pm 2.1 \pm 0.9$</td>
</tr>
</tbody>
</table>

VII. UNFOLDING THE ASYMMETRY

The true or generated distribution of a certain variable ($\Delta y$ for the inclusive measurement) is shaped by acceptance and detector resolution, resulting in the observed distribution, which is also subject to statistical fluctuations. The goal of the unfolding procedure is to find the best estimator for the true distribution given the background-subtracted data and knowing detector acceptance and resolution from simulation. After finding the
FIG. 2: (Color online). The combined discriminant $D_c$. The region $D_c < 1$ is not used to determine the signal $A_{FB}$. The ratio between the data counts and the model expectation is shown in the lower panel, with the hashed area representing the systematic uncertainties. Figure is from Ref. [10].

FIG. 3: (Color online). Reconstructed difference between the rapidities of the top and antitop quarks, $\Delta y$. The left column shows $l+3$ jet events, and the right column shows $l+\geq 4$ jet events. Rows from top to bottom display events with 0, 1, and $\geq 2$ $b$ tags. Overflows are included in the edge bins. The ratio between the data counts and the model expectation is shown in the lower panels, with the hashed area representing the systematic uncertainties.

best estimator for the true distribution of $\Delta y$, we summarize it into the production-level $A_{FB}$ using Eq. 2. This is the same general approach used in the previous measurement [4]. For this unfolding we use TUnfold [42], which we extend as discussed below.

Each distribution is presented as event counts in a
binned histogram\textsuperscript{4}, i.e., as a vector with a dimension equal to the number of bins. Given the vector of production-level $t\bar{t}$ signal counts, $p$, and the vector of expected background counts, $b$, the expected data counts in the \textit{i}-th bin $\hat{d}_i$ is given by

$$\hat{d}_i = T_{ij}p_j + b_i,$$

where $A$ is a diagonal acceptance matrix, whose $jj$-th element is the probability for an event produced in the \textit{j}-th bin to pass the selection criteria and $M$ is the normalized migration matrix, whose $ij$-th element is the probability for a selected event produced in the \textit{j}-th bin to be observed in the \textit{i}-th bin.

Given the vector of observed counts, $d$, we can construct the vector of background-subtracted reconstruction-level counts, $r = d - b$, with its covariance error matrix $V$. The matrix $V$ is constructed to account for the expected statistical uncertainties on data and background, in particular those due to the size of the multijet-enriched control sample. We then seek to find the vector $u$, which best estimates the vector of production-level counts $p$, by minimizing

$$\chi^2 = (r - Tu)^TV^{-1}(r - Tu) + \tau^2(Ru)^TRu$$

for a given vector $r$, where $\tau$ is the regularization strength and $R$ is the regularization matrix. The first term of Eq.~(6) quantifies the consistency of $u$ with data, while the second (regularization) term quantifies the smoothness of $u$.

Without regularization, the unfolding procedure amounts to a minimization of the first term in Eq.~(6) If the numbers of reconstruction and production-level bins are equal, the problem of minimization is solved by simply inverting the matrix: $u_{\text{unregularized}} = T^{-1}r$.

Unregularized matrix inversion typically results in unphysical, rapidly varying distributions\textsuperscript{43}. Such distributions are disfavored in regularized unfolding by adding a second “regularization” term to the $\chi^2$. The regularization term in Eq.~(6) depends on the discrete second derivative of the binned distribution $u$. For constant bin widths, the regularization term is calculated using a regularized matrix with the following structure\textsuperscript{43}:

$$R = \begin{pmatrix}
0 & 0 & 0 & 0 & \cdots & 0 \\
1 & -2 & 1 & 0 & \cdots & 0 \\
0 & 1 & -2 & 1 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & 1 & -2 & 1 \\
0 & \cdots & 0 & 0 & 0 & 0
\end{pmatrix}.$$  

For this analysis we modify the structure of $R$ to regularize based on the second derivative of the event density rather than the event counts, which allows for the use of variable bin sizes. The regularization strength $\tau$ is chosen using both ensemble testing (described below) and the L-curve technique\textsuperscript{12} to balance the minimization of statistical fluctuations and bias. The difference between the two techniques is included in the evaluation of the systematic uncertainty due to the choice of the regularization strength.

As in Ref.\textsuperscript{4}, the production-level $\Delta y$ distribution is divided into 26 bins and the reconstruction-level $\Delta y$ distribution is divided into 50 bins. Both have narrower bins near $\Delta y = 0$, where the probability to misclassify forward events as backward or vice versa changes rapidly, and wider bins at high $|\Delta y|$, where statistics are low.

For the 2D measurement, we use six $m_{t\bar{t}}$ bins at the production level, with edges at 0, 400, 450, 500, 550, 650 GeV and $+\infty$. The joint distribution of $\Delta y$ and $m_{t\bar{t}}$ has a kinematic boundary at $|\Delta y| = \log \left( \frac{1 + \beta}{1 - \beta} \right)$, where $\beta = \sqrt{1 - (2m_t/m_{t\bar{t}})^2}$. A bin edge close to this boundary would result in a large difference in the event density between adjacent bins, a feature that would be smoothed by a regularization procedure, thus biasing $u$. To avoid such a bias, the $\Delta y$ edges of the bins of the 2D measurement are chosen to depend on $m_{t\bar{t}}$ as shown in Fig.~5.

The reconstruction-level histograms have similar but finer bins along both the $\Delta y$ and $m_{t\bar{t}}$ directions. In the $l+3\text{jet}$ channels 13 $m_{t\bar{t}}$ bins are used to accurately describe migrations among the six production-level bins. The $m_{t\bar{t}}$ resolution in the $l+\geq 4\text{jet}$ channels allows for 14 $m_{t\bar{t}}$ bins.

![FIG. 5: \textcopyright{}Color online. Production-level bins for the 2D measurement in the $(m_{t\bar{t}}, \Delta y)$ plane, overlaid on the distribution in these variables predicted from MC@NLO. The shading reflects the predicted event density in arbitrary units. The solid and dashed lines denote the production-level bins. The solid lines show bins that are used for the final result.](image)

We simultaneously unfold to the production level the four channels that contain at least one $b$ tag. The difference in purity among channels is accounted for in the definition of the covariance matrix $V$. The unfolding technique is calibrated, and the statistical and systematical uncertainties are determined using the results of ensemble tests. Each ensemble comprises simulated PDFs that we build according to MC@NLO, ALPGEN\textsuperscript{20}.

\textsuperscript{4} Overflows are included in the edge bins.
or MadGraph SM predictions, or according to toy models with different asymmetries. The PDs are created from the expected bin counts $d_i$ calculated using Eq. [4] by adding Poisson (statistical) and Gaussian (systematic) fluctuations, with the Gaussian width taken as one standard deviation for the corresponding systematic uncertainty.

In the toy models the input distribution $P(\Delta y)$ has the form:

$$P(\Delta y) = G(\Delta y; \mu, w\sigma_0)(1 + a \text{erf}(\Delta y/\delta)), \quad (8)$$

where $a$ and $\delta$ are shaping parameters, $G$ is a Gaussian distribution with mean $\mu$ and width $w\sigma_0$, $\sigma_0$ the width predicted by MC@NLO, and $w$ a scaling parameter. The shape of the $\Delta y$ distribution and the input asymmetry are varied using the parameters $\mu$, $a$, $\delta$, and $w$. In addition, we produce ensembles with the signal taken from simulated samples of $t\bar{t}$ production mediated by axigluons, hypothetical massive particles that arise in extensions of the SM that suggest different strong couplings for left and right-handed quarks [9]. The input asymmetry in the models used for calibration ranges from $-30\%$ to $+30\%$, while the axigluon masses are varied from 0.2 to 2 TeV.

The bias, which is the average difference between the unfolded and input $A_{FB}$ values, is shown in Fig. 6 as a function of the input $A_{FB}$. Axigluon scenarios are indicated in the legend by the mass of the axigluon, $m_a$. The toy models are labeled by parameters $w$, $a$, and $\delta$ of Eq. [8]. Unless stated otherwise $w = 1$, $a = 0$, and $\delta = 1$. For each set of $a$, $\delta$, and $w$, the value of $\mu$ is varied to produce different input asymmetries. The dashed line indicates the calibration applied to the inclusive measurement and the dotted lines indicate the assigned calibration uncertainties. The point significantly outside of the dotted lines corresponds to an axigluon mass of 0.4 TeV and is discussed in the text.

The unfolded $\Delta y$ distribution is presented in eight bins, with each bin calibrated for the expected bias observed in the MC@NLO-simulated PDs. The $A_{FB}$ value in each $|\Delta y|$ range, $A_{FB}(|\Delta y|)$, is calibrated using the same procedure as for the inclusive $A_{FB}$. Since no systematic correlation is found between the $A_{FB}$ biases in different $|\Delta y|$, as well as $m_{t\bar{t}}$ bins, they are calibrated individually.

### VIII. Systematic Uncertainties

The systematic uncertainties on the reconstruction and production-level $A_{FB}$ are summarized in Table [IV] in several categories, which are detailed below. To evaluate the systematic uncertainty on the reconstruction-level $A_{FB}$, we vary the modeling according to the estimated uncertainty in the relevant parameter of the model and propagate the effect to the result. The systematic uncertainties on the production-level $A_{FB}$ are evaluated by including the effects of systematic variations on the simulated background-subtracted PD into the ensemble tests. To find the expected uncertainty due to each category we use dedicated ensembles generated without statistical fluctuations and with only the relevant systematic effects. The total uncertainties on the production-level $A_{FB}$ are taken from ensembles built including both statistical fluctuations and systematic effects (see Sec. [VII]).

![Table IV: Systematic uncertainties on $A_{FB}$, in absolute %.
For the 2D measurement, the range of changes in $A_{FB}$ over the six $m_{t\bar{t}}$ bins is given.

<table>
<thead>
<tr>
<th>Source</th>
<th>Reco. level</th>
<th>Production level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>inclusive</td>
<td>inclusive 2D</td>
</tr>
<tr>
<td>Background model</td>
<td>$+0.7/-0.8$</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.1–2.8</td>
</tr>
<tr>
<td>Signal model</td>
<td>$&lt;0.1$</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.8–5.2</td>
</tr>
<tr>
<td>Unfolding</td>
<td>N/A</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.9–1.9</td>
</tr>
<tr>
<td>PDFs and pileup</td>
<td>0.3</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.5–2.9</td>
</tr>
<tr>
<td>Detector model</td>
<td>$+0.1/-0.3$</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.4–3.3</td>
</tr>
<tr>
<td>Sample composition</td>
<td>$&lt;0.1$</td>
<td>$&lt;0.1$</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$&lt;0.1$</td>
</tr>
<tr>
<td>Total</td>
<td>$+0.8/-0.9$</td>
<td>1.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.1–7.5</td>
</tr>
</tbody>
</table>

The background model category includes the following sources, which affect the properties predicted for background events. The leptonic asymmetry of the $W$+jets background is varied within its uncertainty of 3% [10]. The rate of heavy-flavor production within $W$+jets production is varied by $\pm 20\%$ [22, 42]. The efficiencies for lepton identification, and the probabilities
The signal model category includes the sources of uncertainty that affect the predictions for signal events other than the ones accounted for in the PDFs and pileup category. The top quark mass is varied according to the combined Tevatron measurement of Ref. [46]. The effect of higher order corrections to $t\bar{t}$ production is estimated by replacing the migration matrix $M$ from Eq. 1 simulated by MC@NLO with the one simulated by ALPGEN, which uses tree-level matrix elements. The $b$ quark fragmentation function is varied within its uncertainties [46], which also affects background modeling.

The signal model category also includes the uncertainties associated with gluon radiation. The total amount of initial state radiation is varied in a range consistent with the results of Ref. [47]. We also consider the difference in the predicted amount of initial state radiation between forward and backward events, both because of contributions at order $a_s^3$ and due to higher order effects which are modeled by the simulated parton showers [48]. We account for this uncertainty by reducing the difference in the distributions of the $p_T$ of the $t\bar{t}$ system for forward and backward events by 25%, a value derived from Ref. [18]. We also account for the possibility that the mismodeling of this variable in the $l+3$ jet final state affects $A_{FB}$ by reweighting this distribution to match the D0 data, similarly to the procedure used in Ref. [10].

The uncertainties due to unfolding are dominated by the calibration uncertainties. The uncertainties associated with the choice of the regularization strength and statistical fluctuations in the MC samples used to find the migration matrix are also included.

The PDFs and pileup category includes uncertainties on the modeling of the $p\bar{p}$ collisions. The main uncertainties are from the PDFs, which primarily affect the $\Delta y$ distribution of the $W+3$ jet background. These uncertainties are evaluated by varying the contributions of the various eigenvectors from the CTQS6.1 PDF [22] and by considering an alternative set of PDFs (MRST2001 [19]). The number of additional collisions within the same bunch crossing (pileup) affects the quality of the reconstruction. The uncertainties on the modeling of additional collisions are also included in this category.

The detector model category includes the following sources of systematic uncertainty. The efficiencies of the $b$-tagging algorithm for jets of different flavors, which are measured from collider data, are varied according to their uncertainties [11]. These variations affect the measured $A_{FB}$ mostly through the estimated sample composition, which depends strongly on the classification of data into several channels according to the number of $b$ tags. The modeling of jet energy reconstruction, including the overall energy scale and the energy resolution, as well as jet-reconstruction efficiencies and single-particle responses, are all calibrated to collider data and are varied according to their uncertainties [50]. The uncertainties due to jet reconstruction and energy measurement are significantly reduced compared to the previous measurement due to the inclusion of the $l+3$ jet events.

Lastly, the sample composition is varied according to its fitted uncertainties. This variation is performed in addition to the changes in the sample composition implicitly induced by other systematic variations.

IX. RESULTS

A. Inclusive $A_{FB}$ and $A_{FB}$ dependence on $|\Delta y|$

The calibrated production-level $\Delta y$ distribution is shown in Fig. 7. The corresponding inclusive forward–backward asymmetry in $t\bar{t}$ production is $(10.6 \pm 3.0)\%$.

The dependence of $A_{FB}$ on $|\Delta y|$ is shown in Fig. 8 and Table V with the correlation factors between bins listed in Table VII. These correlations are taken into account in the fit of the measured $A_{FB}$ ($|\Delta y|$) to a line. Since for any physical $\Delta y$ distribution the asymmetry at $\Delta y = 0$ is 0, we constrain the line to the origin and fit for its slope. For data, we find a slope of $0.154 \pm 0.043$. This slope is compatible within two standard deviations with the MC@NLO-simulated slope of 0.080, which has negligible statistical uncertainty. The difference between the slope reported by the CDF Collaboration [5] and the slope reported in

![FIG. 7: (Color online). The production-level $\Delta y$ distribution. The D0 data points are shown with their statistical uncertainty indicated by the black rectangles and their total uncertainty, based on the diagonal elements of the covariance matrix, indicated by the hashed areas. The histogram shows the MC@NLO prediction [18]. The $x$ coordinate of each data point is the observed average of the $\Delta y$ distribution in the corresponding bin.](image-url)
FIG. 8: (Color online). The dependence of the forward–backward asymmetry on $|\Delta y|$. The D0 data points are shown with the total error bars indicating the total uncertainty, based on the diagonal elements of the covariance matrix, while the statistical uncertainties are indicated by the inner error bars. The dashed line shows the fit to the data with the dotted lines indicating the fit uncertainty. The horizontal lines show the $\text{MC@NLO}$ prediction for the asymmetry in each $m_{t\bar{t}}$ bin [18]. The last bin has no upper boundary. The $x$ coordinate of each data point is the observed average of $|\Delta y|$ in the corresponding bin.

TABLE V: Variation of the production-level $A_{FB}$ on $|\Delta y|$. The measured values are calibrated and listed with their total uncertainties. The theoretical predictions are based on $\text{MC@NLO}$ simulation.

| $|\Delta y|$ | Predicted | Measured |
|---------|-----------|----------|
| < 0.25  | 1.1       | 1.8 ± 1.3|
| 0.25–0.5| 2.5       | 5.4 ± 3.3|
| 0.5–1   | 5.2       | 10.8 ± 4.8|
| > 1     | 11.4      | 21.8 ± 7.1|

The values of the asymmetry measured in six $m_{t\bar{t}}$ ranges constitute a six-dimensional vector $\mathbf{\alpha}$ with a $6 \times 6$ covariance matrix $\Sigma$. Table [X] lists the eigenvectors $\mathbf{e}_i$ ($i = 1, \ldots, 6$) of $\Sigma$ together with the corresponding components of the vector $\mathbf{\alpha}$ in the basis formed by the eigenvectors: $\alpha_i = \mathbf{\alpha} \cdot \mathbf{e}_i$. Their uncertainties $\sigma_i = \sigma \mathbf{\Sigma}^{-1/2} \mathbf{\alpha}_i$, where $\Sigma'$ is the covariance matrix transformed to the basis $\mathbf{e}_i$. The elements of Table [X] fully specify the measured six-dimensional likelihood in the Gaussian approximation, and can be used for quantitative comparison with theoretical predictions and other experimental results [51].

Using the full covariance matrix we perform a fit of the measured $A_{FB}$ to the functional form

$$A_{FB}(m_{t\bar{t}}) = \alpha \left( \frac{m_{t\bar{t}}}{\text{GeV}} - C \right) + A_0. \quad (9)$$

We choose $C = 445$ so that the correlation factor between the fit parameters $\alpha$ and $A_0$ is less than 0.01 in the fit to the data. The parameters of the fit are listed in Table [X] for the data and the $\text{MC@NLO}$ simulation. We observe a slope $\alpha$ consistent with zero and with the $\text{MC@NLO}$ prediction. The difference between slope reported by the CDF Collaboration [8] and the slope reported in this paper corresponds to 1.8 standard deviations.

$5$ When comparing to CDF results, we neglect the correlations of the systematic uncertainties between the two experiments.
TABLE IX: The eigenvectors of the covariance matrix $\Sigma$ and the result of the 2D measurement $\vec{r}$, in the basis of eigenvectors.

<table>
<thead>
<tr>
<th>$i$</th>
<th>Eigenvector $\vec{c}_i$</th>
<th>$v_i \pm \sigma_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$(-0.592, +0.770)$</td>
<td>$(-0.237, -0.007, +0.004, -0.000)$</td>
</tr>
<tr>
<td>2</td>
<td>$(+0.434, +0.099)$</td>
<td>$(-0.775, +0.448, -0.030, +0.002)$</td>
</tr>
<tr>
<td>3</td>
<td>$(+0.673, +0.591)$</td>
<td>$(+0.251, -0.339, +0.138, -0.004)$</td>
</tr>
<tr>
<td>4</td>
<td>$(+0.034, +0.192)$</td>
<td>$(+0.516, -0.826, +0.104, +0.049)$</td>
</tr>
<tr>
<td>5</td>
<td>$(-0.076, -0.099)$</td>
<td>$(-0.113, -0.040, +0.917, +0.360)$</td>
</tr>
<tr>
<td>6</td>
<td>$(-0.029, -0.030)$</td>
<td>$(-0.019, +0.031, +0.359, -0.932)$</td>
</tr>
</tbody>
</table>

TABLE X: Parameters of the fit to Eq. 9. The theoretical predictions are based on the MC@NLO simulation and have negligible statistical uncertainties.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Predicted</th>
<th>Measured</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope, $\alpha$</td>
<td>$3.8 \cdot 10^{-4}$</td>
<td>$(3.9 \pm 4.4) \cdot 10^{-4}$</td>
</tr>
<tr>
<td>Offset, $A_0$</td>
<td>$5.3 \cdot 10^{-2}$</td>
<td>$(11.9 \pm 3.6) \cdot 10^{-2}$</td>
</tr>
</tbody>
</table>

FIG. 9: (Color online). The dependence of the forward–backward asymmetry on the invariant mass of the $t\bar{t}$ system. The D0 data points are shown with the total error bars indicating the total uncertainty, based on the diagonal elements of the covariance matrix, while the statistical uncertainties are indicated by the inner error bars. The dashed line shows a linear fit to the data with the dotted curves indicating the fit uncertainty. The horizontal lines correspond to the MC@NLO prediction for the asymmetry in each $m_{t\bar{t}}$ bin. Shaded boxes correspond to the prediction of Refs. [8, 18]. The last bin has no upper boundary. The $x$ coordinate of each data point is the simulated average of the $m_{t\bar{t}}$ distribution in the corresponding bin.

X. DISCUSSION

The measured inclusive forward–backward asymmetry in $t\bar{t}$ production, $A_{FB} = (10.6 \pm 3.0)\%$ is in agreement with the SM predictions reviewed in Section VI which range from an inclusive asymmetry of 5.0% predicted by the MC@NLO simulation to (8.8 \pm 0.9)\% [35] once electroweak effects are taken into account. The measured dependences of the asymmetry on $|\Delta y|$ and $m_{t\bar{t}}$ are also in agreement with the SM predictions. Nevertheless, the observed $A_{FB}$ and the dependences of $A_{FB}$ on $m_{t\bar{t}}$ and $|\Delta y|$ do not disfavor the larger asymmetries that were previously measured in $pp$ collisions [2].

To compare the presented result with the previous D0 publication [4], Table XI presents $A_{FB}$ at the reconstruction level measured in different samples. The method discussed in this paper applied to $l+\geq 4 \text{ jet}$ events from the first $5.4 \text{ fb}^{-1}$ of integrated luminosity yields a result consistent with that in Ref. [4], but with a reduced uncertainty mainly due to the separation of data into channels based on the number of $b$ tags and the increased efficiency of the new $b$-tagging algorithm. Once the analysis is extended to include the $l+3 \text{ jet}$ events collected at that time, the uncertainty is reduced by a factor of 1.26. The result obtained in the second $4.3 \text{ fb}^{-1}$ of the Tevatron dataset is within one standard deviation from that obtained in the first $5.4 \text{ fb}^{-1}$. The statistical uncertainty obtained in the combined $9.7 \text{ fb}^{-1}$ of integrated luminosity is reduced by a factor of 1.29 with respect to the result obtained using the same method in the first $5.4 \text{ fb}^{-1}$, while the reduction expected from scaling with the integrated luminosity is 1.34. This loss of sensitivity is mainly due to higher instantaneous luminosity during the collection of the later data, which required a tighter trigger selection.

TABLE XI: Reconstruction-level asymmetries measured in different samples with different methods, with their statistical uncertainties.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Method</th>
<th>$A_{FB}$, %</th>
</tr>
</thead>
<tbody>
<tr>
<td>$l+\geq4 \text{ jet}$, first $5.4 \text{ fb}^{-1}$</td>
<td>From Ref. [4]</td>
<td>$9.2 \pm 3.7$</td>
</tr>
<tr>
<td>$l+\geq4 \text{ jet}$, first $5.4 \text{ fb}^{-1}$</td>
<td>This analysis</td>
<td>$9.9 \pm 3.4$</td>
</tr>
<tr>
<td>$l+\geq3 \text{ jet}$, first $5.4 \text{ fb}^{-1}$</td>
<td>This analysis</td>
<td>$10.1 \pm 2.7$</td>
</tr>
<tr>
<td>$l+\geq3 \text{ jet}$, additional $4.3 \text{ fb}^{-1}$</td>
<td>This analysis</td>
<td>$6.0 \pm 3.1$</td>
</tr>
<tr>
<td>$l+\geq3 \text{ jet}$, full $9.7 \text{ fb}^{-1}$</td>
<td>This analysis</td>
<td>$7.9 \pm 2.1$</td>
</tr>
</tbody>
</table>

The improved reconstruction of $\Delta y$ and the reduced acceptance bias due to the inclusion of the $l+3 \text{ jet}$ events result in further reduction of the statistical uncertainty on the unfolded result compared to Ref. [4]. The separation of the data into channels allows us to add the $l+3 \text{ jet}$ channels without losing the statistical power of the purer $l+\geq4 \text{ jet}$ channels.
XI. SUMMARY

In summary, we report the measurement of the forward–backward asymmetry in $t\bar{t}$ production using the dataset recorded by the D0 detector in Run II of the Fermilab Tevatron Collider. The results presented here supersede the ones that were based on about half of the data. The analysis is extended to include events with three jets, allowing for the loss of one jet from the $t\bar{t}$ decay and reducing the acceptance corrections. The unfolding procedure now accounts for the differences in sample compositions between channels, thus maximizing the statistical strength of the individual channels. New reconstruction techniques are used in the $l+\geq4$ jet channel, improving the experimental resolution in all variables of interest.

The asymmetry measured at the reconstruction level is $A_{FB} = (7.9 \pm 2.3)\%$. After correcting for detector resolution and acceptance, we obtain a production-level asymmetry $A_{FB} = (10.6 \pm 3.0)\%$. The observed asymmetry and the dependences of $A_{FB}$ on $m_{t\bar{t}}$ and $|\Delta y|$ are consistent with the standard model predictions.

Acknowledgments

We thank M. Mangano, P. Skands and G. Perez for enlightening discussions. We thank the staffs at Fermilab and collaborating institutions, and acknowledge support from the DOE and NSF (USA); CEA and CNRS/IN2P3 (France); MON, NRC KI and RFBR (Russia); CNPq, FAPERJ, FAPESP and FUNDUNESP (Brazil); DAE and DST (India); Colciencias (Colombia); CONACyT (Mexico); NRF (Korea); FOM (The Netherlands); STFC and the Royal Society (United Kingdom); MSMT and GACR (Czech Republic); BMBF and DFG (Germany); SFI (Ireland); The Swedish Research Council (Sweden); and CAS and CNSF (China).

[33] Private communications with W. Bernreuther, updating the error estimates of Ref. [35].