Measurement of the charge asymmetry in top-quark pair production in proton–proton collisions at $\sqrt{s} = 7$ TeV

CMS Collaboration

CERN, Switzerland

ARTICLE INFO

Article history:
Received 21 December 2011
Received in revised form 31 January 2012
Accepted 31 January 2012
Available online 3 February 2012
Editor: M. Doser

Keywords:
CMS
Physics
Top quark

ABSTRACT

The difference in angular distributions between top quarks and antiquarks, commonly referred to as the charge asymmetry, is measured in pp collisions at the LHC with the CMS experiment. The data sample corresponds to an integrated luminosity of 1.09 fb$^{-1}$ at a centre-of-mass energy of 7 TeV. Top-quark pairs are selected in the final state with an electron or muon and four or more jets. At least one jet is identified as originating from b-quark hadronization. The charge asymmetry is measured in two variables, one based on the pseudorapidities ($\eta$) of the top quarks and the other on their rapidities ($y$). The results $A^\eta_C = -0.017 \pm 0.032$ (stat.)$^{+0.029}_{-0.036}$ (syst.) and $A^y_C = -0.013 \pm 0.028$ (stat.)$^{+0.029}_{-0.031}$ (syst.) are consistent within uncertainties with the standard-model predictions.

© 2012 CERN. Published by Elsevier B.V. Open access under CC BY-NC-ND license.

1. Introduction

The top quark is the only fundamental fermion with a mass on the order of the scale of electroweak symmetry breaking, and may therefore play a special role in physics beyond the standard model (BSM). In some BSM theories, top-quark pairs can be produced through the exchange of yet unknown heavy particles, in addition to the production through quark–antiquark annihilation and gluon–gluon fusion. Possible candidates include axigluons [1, 2], $Z'$ bosons [3], and Kaluza–Klein excitations of gluons [4, 5]. Such new particles can appear as resonances in the $t\bar{t}$ invariant mass spectrum in s-channel production of top-quark pairs. If these hypothetical particles are exchanged in the $t$ or $u$ channels, alternative approaches are needed to search for new top-quark production modes [6]. One property of $t\bar{t}$ production that can be sensitive to the presence of such additional contributions is the difference in angular distributions of top quarks and antiquarks, commonly referred to as the charge asymmetry.

In the standard model (SM), a small charge asymmetry in $t\bar{t}$ production through quark–antiquark annihilation appears in QCD calculations at next-to-leading order (NLO) [7,8]. The interference between the Born diagram and the box diagram, as well as between initial- and final-state radiation, correlates the flight directions of the top quarks and antiquarks to the directions of motion of the initial quarks and antiquarks, respectively. The asymmetric initial state of proton–antiproton collisions leads to an observable forward–backward asymmetry at the Tevatron, where the top quarks are emitted preferentially along the direction of motion of the incoming protons and the top antiquarks along the direction of the antiprotons. This asymmetry is observable in the difference in rapidity ($y$) of top quarks and antiquarks, $y_t - y_{\bar{t}}$. Recent measurements [9,10] by the CDF and D0 Collaborations report asymmetries that are about two standard deviations larger than the value of about 0.08 [7,8,11–13] predicted in the SM. At high $t\bar{t}$ invariant mass ($M_{t\bar{t}} > 450$ GeV/c$^2$), the CDF Collaboration finds an even larger asymmetry relative to the SM prediction [9], while the D0 Collaboration does not observe a significant mass dependence of the asymmetry. These results have led to speculations that the large asymmetry might be generated by additional axial couplings of the gluon [14] or by heavy particles with unequal vector and axial-vector couplings to top quarks and antiquarks [15–28].

Owing to the symmetric initial state of proton–proton collisions at the Large Hadron Collider (LHC), the charge asymmetry does not manifest itself as a forward–backward asymmetry; the rapidity distributions of top quarks and antiquarks are symmetrical around $y = 0$. However, since the quarks in the initial state are mainly valence quarks, while the antiquarks are always sea quarks, the larger average momentum fraction of quarks leads to an excess of top quarks produced in the forward directions. The rapidity distribution of top quarks in the SM is therefore broader than that of the more centrally produced top antiquarks. The same effect is visible in the purely geometrically defined pseudorapidity $\eta = -\ln(\tan(\theta/2))$, where $\theta$ is the polar angle relative to the counterclockwise beam axis. The charge asymmetry can be observed through the difference in the absolute values of the pseudorapidities of top quarks and antiquarks, $\Delta|\eta| = |\eta_t| - |\eta_{\bar{t}}|$ [29].
Another approach is motivated in Ref. [30], where the observable used by the Tevatron experiments \((y_1 - y_2)\) is multiplied by a factor that accounts for the boost of the \(t\bar{t}\) system, yielding \(\Delta y^2 = (y_1 - y_2) \cdot (y_1 + y_2) = (y_1^2 - y_2^2)\). Using either of the two variables, the charge asymmetry can be defined as

\[
A_C = \frac{N^+ - N^-}{N^+ + N^-},
\]

where \(N^+\) and \(N^-\) represent the number of events with positive and negative values in the sensitive variable, respectively. Since the SM charge asymmetry is a higher-order effect in quark–antiquark annihilation, and at the LHC the top–quark pairs are produced mainly through gluon–gluon fusion, the asymmetry expected in the SM at the LHC is smaller than at the Tevatron. For a centre-of-mass energy of 7 TeV, the NLO prediction for an asymmetry in the \(\Delta y\eta\) variable is \(A_{pT}^{\Delta y}(\text{theory}) = 0.0136 \pm 0.0008\) [12], while \(A_{pT}^{\Delta y}(\text{theory}) = 0.0115 \pm 0.0006\) [12] for the rapidity variable. The existence of new sources of physics with different vector and axial-vector couplings to top quarks and antiquarks could enhance these asymmetries up to a maximum of 0.08 [6]. The uncertainties on the above predictions reflect the variations from the choice of parton distribution functions and different choices of factorization and renormalization scales, as well as the dependence on the top-quark mass within its experimental uncertainty.

For the sake of simplicity, we focus on \(\Delta y\eta\) when describing the method but quote uncertainties and results for both variables. We discuss below the experimental setup (Section 2), the data sample (Section 3), and the selection of \(t\bar{t}\) candidate events (Section 4). This is followed by a description of the estimation of the background contamination (Section 5). The reconstruction of the kinematics of the top–quark candidates is described in Section 6. The details of the applied unfolding and measurement procedures and of the different sources of systematic uncertainties are given in Section 7 and Section 8, respectively, and the results of the analysis are presented in Section 9.

2. The CMS detector

The central feature of the Compact Muon Solenoid (CMS) apparatus is a superconducting solenoid of 6 m internal diameter, providing a field of 3.8 T. Within the field volume are the silicon pixel and strip tracker, the crystal electromagnetic calorimeter (ECAL), and the brass/scintillator hadron calorimeter (HCAL). The inner tracker measures trajectories of charged particles within the pseudorapidity range \(|\eta| < 2.5\). It consists of 1440 silicon pixel and 15148 silicon strip detector modules and provides an impact parameter resolution of \(\sim 15 \mu m\) and a transverse momentum \(p_T\) resolution of about 1.5% for 100 GeV/c particles.

The ECAL consists of nearly 76 000 lead tungstate crystals that provide coverage in pseudorapidity of \(|\eta| < 1.48\) for the ECAL barrel region and \(1.48 < |\eta| < 3.0\) for the two endcaps. A preshower detector consisting of two planes of silicon sensors interleaved with a total of three radiation lengths of lead is located in front of the endcaps. The ECAL energy resolution is 3% or better for the range of electron energies relevant for this analysis. The HCAL is composed of layers of plastic scintillator within a brass/stainless steel absorber, covering the region \(|\eta| < 3.0\). In the region \(|\eta| < 1.74\), the HCAL cells have widths of 0.087 in pseudorapidity and 0.087 rad in azimuth (\(\phi\)). In the \((\eta, \phi)\) plane, for \(|\eta| < 1.48\), the HCAL cells match the corresponding \(5 \times 5\) ECAL crystal arrays to form calorimeter towers projecting radially outwards from the centre of the detector. At larger values of \(|\eta|\), the coverage in \(\eta\) of each tower increases, although the matching ECAL arrays contain fewer crystals.

Muons are measured in the pseudorapidity range \(|\eta| < 2.4\), with detection planes made using three technologies, drift tubes, cathode strip chambers, and resistive plate chambers, all embedded in the steel return yoke. Matching the muons to the tracks measured in the silicon tracker provides a transverse momentum resolution between 1 and 5%, for \(p_T\) values up to 1 TeV/c. In addition to barrel and endcap detectors, CMS has extensive forward calorimeter. A detailed description of CMS can be found in Ref. [31].

3. Data and simulation

This analysis of \(t\bar{t}\) events produced in proton–proton collisions at a centre-of-mass energy of 7 TeV is based on data taken with the CMS detector, corresponding to an integrated luminosity of \(1.09 \pm 0.04\) fb\(^{-1}\). To translate the distributions measured with reconstructed objects to distributions for the underlying quarks, we use simulated data samples. Top-quark pair events are generated with the tree-level matrix-element generator MADGRAPH version 5 [32], interfaced to PYTHIA version 6.4 [33] for the parton showering, where the MLM algorithm [34] is used for the matching. Spin correlation in decays of top quarks is taken into account and higher-order gluon and quark production is described through matrix elements for up to three extra jets accompanying the \(t\bar{t}\) system. Although the higher-order processes leading to the \(t\bar{t}\) charge asymmetry are not taken fully into account in this leading-order (LO) simulation, its usage is still justified by the fact that these processes affect only the production of top quarks and not their decay, and that the simulated events are used only to reconstruct top-quark momenta from their decay products and to correct for resolution effects. We simulate the main SM backgrounds to top-quark pair production using the same combination of MADGRAPH and PYTHIA programs. The radiation of up to four jets in weak vector-boson production is simulated through matrix-element-based calculations (these processes are denoted as \(W + jets\) and \(Z + jets\) in the following). The background from electroweak production of single top quarks is also simulated using the MADGRAPH generator. Multijet background events are generated using PYTHIA version 6.4. On average, six additional proton–proton interactions (pile-up) per event are observed in the analysed data, and this pile-up contribution is overlaid on the simulated events, all of which are processed through the CMS detector simulation and reconstructed using standard CMS software.

4. Event selection

In the SM, a top quark decays almost exclusively into a b quark and a W boson. In this measurement we focus on \(t\bar{t}\) events, where one of the W bosons from the decay of a top–quark pair subsequently decays into a muon or electron and the corresponding neutrino, and the other W boson decays into a pair of jets. We therefore select events containing one electron or muon and four or more jets, at least one of which is identified as originating from the hadronization of a b quark. For the reconstruction of electrons, muons, jets, and any imbalance in transverse momentum due to the neutrino, we use a particle-flow (PF) algorithm [35]. This algorithm aims to reconstruct the entire event by combining information from all subdetectors, including tracks of charged particles in the tracker and the muon system, and energy depositions in the electromagnetic and hadronic calorimeters.

We select events with triggers that require an electron or muon with transverse momentum greater than 25 GeV/c and 17 GeV/c, respectively, together with at least three jets, each with \(p_T > 30\) GeV/c. In addition, we require the primary vertex reconstructed from the tracks with the largest summed transverse momentum...
to be located in a cylindrical region defined by the longitudinal distance $|z| < 24 \text{ cm}$ and radial distance $r < 2 \text{ cm}$ relative to the centre of the CMS detector.

In the electron + jets selection, the electron candidates are required to have a transverse momentum of $p_T > 30 \text{ GeV}/c$ and to be within the region $|\eta| < 2.5$, excluding the transition region between the ECAL barrel and endcaps of $1.4442 < |\eta_{\text{rel}}| < 1.5660$, where $\eta_{\text{rel}}$ is the pseudorapidity of the electron candidate's supercluster, which corresponds to the cluster of ECAL energy deposits from the electron and any accompanying bremsstrahlung photons [36]. The transverse impact parameter of the electron track relative to the beam axis is required to be smaller than $0.02 \text{ cm}$. The energy in the HCAL cell that is mapped onto the supercluster must be less than 2.5% of the total ECAL energy associated with the supercluster. Additional requirements are made on the spatial distribution of the shower and the angular separation between the ECAL supercluster and the matching track. The longitudinal position of the electron track at its closest approach to the beam line is required to lie within 1 cm of the longitudinal position of the primary vertex, to ensure that the electron is emitted from the primary interaction. Also, electron candidates must be isolated. The lepton isolation variable $I^\ell_{\text{rel}}$ is based on the reconstructed energies of particle-flow objects relative to the lepton transverse momentum ($p^\ell_T$):

$$I^\ell_{\text{rel}} = \frac{E^\ell_{\text{CH}} + E^\ell_{\text{NH}} + E^\ell_{\text{F}}}{p^\ell_T} \cdot c,$$

where $E^\ell_{\text{CH}}$ is the energy deposited by charged hadrons in a cone with radius $\Delta R = 0.4$ in $(\eta, \phi)$ around the lepton track, and $E^\ell_{\text{NH}}$ and $E^\ell_{\text{F}}$ are the respective energies of neutral hadrons and photons. We require electron candidates to have $I^\ell_{\text{rel}} < 0.125$. Events with exactly one electron candidate satisfying these quality criteria are selected for further consideration. Electron candidates that lack signals in the inner layers of the tracking system or that can be paired with a second track of opposite curvature are assumed to be the product of photon conversions and are discarded.

Muons are reconstructed using the combined information from the silicon tracker and muon system. In the selection of muon + jets events, the muons are required to have $p^\mu_T > 20 \text{ GeV}/c$ and lie within the muon trigger acceptance ($|\eta| < 2.1$). The same requirements as for electron candidates are imposed on the transverse impact parameter and the longitudinal origin of the muon track. The muon candidate is required to have a prescribed minimum number of hits in both the silicon tracking system and the muon chambers, and must be isolated from other energy deposits in the event, again defined by $I^\mu_{\text{rel}} < 0.125$. When more than one muon passes all these criteria, the event is rejected.

Dilepton events from $t\bar{t}$ and Z-boson decays are suppressed by applying a veto on additional, less stringently defined charged leptons in the event. We reject all events containing any additional electron candidates with $p_T > 15 \text{ GeV}/c$, $|\eta| < 2.5$, and $I^\ell_{\text{rel}} < 0.25$, or additional muon candidates with $p_T > 10 \text{ GeV}/c$, $|\eta| < 2.5$, and $I^\mu_{\text{rel}} < 0.25$.

We cluster all particles reconstructed through the particle-flow algorithm, excluding isolated electron and muon candidates, into jets using the anti-$k_T$ jet algorithm [37] with the distance parameter $R = 0.5$, as constructed with FastJet version 2.4 [38,39]. The jet energy is corrected for additional contributions from multiple interactions, as well as for $\eta$ and $p_T$-dependent detector response. To account for observed differences of about 10% in jet-energy resolution between data and simulation, a correction is applied to jets in the simulated samples so that their resolutions match those measured in data. Selected jets are required to be within $|\eta| < 2.4$ and have corrected $p_T > 30 \text{ GeV}/c$. At least four jets must be present in an event, and at least one of the jets must be tagged as coming from the hadronization of a b quark by an algorithm that orders the tracks in impact parameter significance and discriminates using the track with the second highest significance [40,41]. This algorithm has a tagging efficiency of about 60%, evaluated using b jets containing muons from semileptonic decays of b hadrons [41], and a misidentification rate of about 1% [41].

5. Estimation of background

Applying the selection criteria described above, we find a total of 12757 events, 5665 in the electron + jets channel and 7092 in the muon + jets channel. From an evaluation of the simulated background processes, we expect a background contribution of about 20% to the selected data sample. We estimate the contributions from the various background processes separately for the two lepton flavours. We make use of the discriminating power of the imbalance in transverse momentum in an event, $E^\text{miss}_T$, and of M3, the invariant mass of the combination of three jets that corresponds to the largest vectorially summed transverse momentum [42]. For the W + jets, Z + jets, and single-top-quark background processes, the respective simulated samples are used to model the shapes of the $E^\text{miss}_T$ and M3 distributions, while an approach based on data is pursued for the multijet background.

Background contributions are estimated in both channels separately by means of binned maximum-likelihood fits to the two distributions. The $E^\text{miss}_T$ distribution shows the largest discrimination power for small values, where it discriminates between events with and without neutrinos in the final state. We therefore separate the data sample into events with $E^\text{miss}_T < 40 \text{ GeV}$ and $E^\text{miss}_T > 40 \text{ GeV},$ and simultaneously fit the $E^\text{miss}_T$ distribution for the low-$E^\text{miss}_T$ sample and the M3 distribution from the high-$E^\text{miss}_T$ sample, to obtain estimates of the numbers of events from each process in the entire data sample ($E^\text{miss}_T > 0$).

The $t\bar{t}$ signal and all the above-listed background processes enter the likelihood function with a single fit parameter for the normalization of their respective $E^\text{miss}_T$ and M3 distributions. W + jets production is inherently asymmetric at the LHC, with more W + bosons being produced than W − bosons. As the distributions of kinematic variables for the two processes are slightly different, this could introduce an artificial contribution to the measured $t\bar{t}$ charge asymmetry. Therefore, this background process requires special care, and we measure the W + jets contributions from W + and W − bosons separately, using different fit parameters for the two sources of W + jets.

As mentioned above, for all background processes, with the exception of multijet events, we rely on simulations to model the $E^\text{miss}_T$ and M3 distributions. Since the overall cross section for multijet production is several orders of magnitude larger than that of any other process, this specific background can be modelled directly from data by defining an appropriate region enriched in multijet events. In both lepton channels, the largest suppression of multijet events in the default event selection is achieved by requiring isolation of the charged leptons. Consequently, to enrich background from multijet events, we require $0.3 < I^\ell_{\text{rel}} < 0.5$, instead of $I^\ell_{\text{rel}} < 0.125$. To avoid double counting of energy contributions, the momenta of these electron and muon candidates are removed from that of the jet to which they were assigned. The event samples obtained with these altered selections are estimated using simulated events to have multijet efficiencies of 92% in the muon + jets channel and 87% in the electron + jets channel.

The Z + jets contribution to the selected data is expected to be small, and is difficult to discriminate from the multijet background.
processes, especially in the $E_T^{\text{miss}}$ distributions. It is also very difficult to discriminate single-top-quark production from the $t\bar{t}$ signal. Both single-top-quark and $Z + \text{jets}$ production are well understood theoretically and their expected contributions are modest. We therefore constrain the numbers of $Z + \text{jets}$ and single-top-quark events in the fit to the predictions from simulation, assigning an uncertainty of 30%, as was done in Ref. [42], through Gaussian functions in the likelihood. The numbers of events for all other processes are left free in the fit.

Table 1 summarizes the results of the fits separately for the electron + jets and muon + jets channels, along with their statistical uncertainties and the sum. The largest correlation is found between the rates for the $W^+ + \text{jets}$ and $W^- + \text{jets}$ backgrounds (+17%). The rates of $W^+ + \text{jets}$ and $Z + \text{jets}$ backgrounds are anticorrelated with that from multijet background (~10%). All other correlations among the fit parameters are found to be small. Fig. 1 shows the measured $E_T^{\text{miss}}$ and $M_3$ distributions summed for the two channels, with the individual simulated contributions normalized to the results from the fit.

6. Reconstruction of $t\bar{t}$ pairs

The measurement of the $t\bar{t}$ charge asymmetry is based on the full reconstruction of the four-momenta of the top quark and antiquark in each event. This is done in two steps: first, by reconstructing the leptonically decaying $W$ boson, and then by associating the measured jets in the event with quarks in the $t\bar{t}$ decay chain.

The transverse momentum of the neutrino is taken to be the reconstructed $E_T^{\text{miss}}$ vector. To calculate the longitudinal component of the neutrino momentum, a quadratic constraint, relying on the known mass and decay kinematics of the $W$ boson, is used. This procedure leads to two solutions for the longitudinal momentum of the neutrino. If these solutions are complex, the transverse components of the neutrino momentum are adjusted such that the $p_T$ of the neutrino is as close as possible to the measured $E_T^{\text{miss}}$ and the imaginary part of the $p_z$ solution vanishes. Adding the resulting four-momentum of the neutrino to that of the charged lepton defines the four-momentum of the parent $W$ boson. Combining the four-vector of one of the jets in the event with that of the $W$ boson results in the four-vector of the top quark decaying to the charged lepton in the final state, while the other top quark is reconstructed by combining three of the remaining jets. The charge of the lepton then defines which of the two reconstructed four-momenta corresponds to the top quark, and which to the top antiquark.

The probability that an assignment of a jet to $t$ or $\bar{t}$ is correct depends on the $t\bar{t}$ production. The probability that a jet to be the best possible assignment is estimated in simulated $t\bar{t}$ events at the value of $\psi$ for each hypothesis to be the best possible one. The calculation of $\psi$ uses the masses of the two reconstructed top quarks and of the hadronically decaying $W$ boson, as well as the $b$-tag information for the four jets assigned to the four final-state quarks. The three masses are correlated, especially those of the hadronically decaying $W$ boson and $t$ quark. Assuming a linear correlation, $m_1$, $m_2$, and $m_3$, through a rotation matrix derived from simulated $t\bar{t}$ events. The mass $m_1$ is almost identical to the mass of the top quark decaying to the charged lepton in the final state, while $m_2$ and $m_3$ are mixtures of the masses of the other top quark and the hadronically decaying $W$ boson. For each of the three uncorrelated masses $m_i$ we calculate a likelihood ratio function $L_i(m_i)$, that provides a measure of the probability for a given hypothesis with a certain value of $m_i$ to be the best possible one.

In addition, we consider the $b$-tag values for the jets assigned to the two b quarks and the two light quarks. The probability that a jet with $b$-tag value $x$ is assigned to one of the b quarks in the best possible hypothesis is estimated in simulated $t\bar{t}$ events and denoted as $P_b(x)$. The probability that an assignment of a jet to one of the light quarks is the best possible assignment is then given by $(1 - P_b(x))$.

Finally, we choose in each event the hypothesis with the largest value of $\psi$.
ical predictions, since several effects bias the measurement at this
definition in Eq. (1), we find the numbers of events with positive and negative values. Using the
rion. In about 72% of all events, the values of \( x_{q1} \) and \( x_{q2} \) are shown in Fig. 2. These distributions are used to calcu-
ected with the correct sign. The observed agreement between the measured and the simulated
distributions substantiates that the simulated samples used in the
assumption Gaussian uncertainties on the background rates as well as on statistical fluctuations in the background templates. The effect of the correlations among the individual background
distributions, discussed in Section 5, is found to be negligible.

Distortions of the remaining \( \bar{t}t \) distributions relative to the true distributions can be factorized into effects from the event selec-
tion and event reconstruction. The values of \( \Delta|\eta| \) or \( \Delta y^2 \) affect the probability for any event to survive all selection criteria (see Fig. 3 (left)), and thereby the distributions even before reconstruc-
tion. Further distortions can occur because of ambiguities in the
assignment of jets to top-quark candidates, the determination of the neutrino momentum from the W-boson mass constraint, the energy resolution of the calorimeters and jet reconstruction, and the overall detector acceptance. The migration matrix, obtained from simulated \( \bar{t}t \) events and shown graphically in Fig. 3 (right), describes the migration of selected events from true values of \( \Delta|\eta| \) to different reconstructed values.

To correct for the above effects, we apply a regularized un-
folding procedure to the data [43] through a generalized matrix-
inversion method. The measured spectrum, denoted as vector \( \bar{x} \), is divided into 12 bins (see y axis of Fig. 3 (right)), where the bin
widths are chosen to contain approximately an equal number of events. Six bins are used for the unfolded spectrum \( \tilde{x} \) (see x axis stage. First, despite the application of relatively stringent \( \bar{t}t \) event selections, about 20% of all events arise from background pro-
cesses. The simulated distributions for these background processes exhibit no significant asymmetries. We normalize these distributions to the observed background rates, and subtract them from the data, assuming Gaussian uncertainties on the background rates as well as on statistical fluctuations in the background templates. The effect of the correlations among the individual background
distributions, discussed in Section 5, is found to be negligible.

Distortions of the remaining \( \bar{t}t \) distributions relative to the true distributions can be factorized into effects from the event selec-
tion and event reconstruction. The values of \( \Delta|\eta| \) or \( \Delta y^2 \) affect the probability for any event to survive all selection criteria (see Fig. 3 (left)), and thereby the distributions even before reconstruc-
tion. Further distortions can occur because of ambiguities in the
assignment of jets to top-quark candidates, the determination of the neutrino momentum from the W-boson mass constraint, the energy resolution of the calorimeters and jet reconstruction, and the overall detector acceptance. The migration matrix, obtained from simulated \( \bar{t}t \) events and shown graphically in Fig. 3 (right), describes the migration of selected events from true values of \( \Delta|\eta| \) to different reconstructed values.

To correct for the above effects, we apply a regularized un-
folding procedure to the data [43] through a generalized matrix-
inversion method. The measured spectrum, denoted as vector \( \bar{x} \), is divided into 12 bins (see y axis of Fig. 3 (right)), where the bin
widths are chosen to contain approximately an equal number of events. Six bins are used for the unfolded spectrum \( \tilde{x} \) (see x axis
of Fig. 3 (right)). Using twice as many bins for the uncorrected as for the corrected spectrum is recommended for this type of unfolding technique [43].

The smearing matrix $S$, which accounts for migration and efficiency, is derived from simulated $t\bar{t}$ events. Mathematically, this matrix is the product of the migration matrix, depicted in Fig. 3 (right), and a diagonal matrix with the efficiencies for each of the bins in Fig. 3 (left) on the diagonal, and all other elements set to zero. It defines the translation of the true spectrum $\vec{x}$ into the measured spectrum $\vec{w} = S \vec{x}$. We solve this equation for the true spectrum $\vec{x}$ using a least-squares (LS) technique and searching for the $\vec{x}$ that minimizes the LS through the use of the generalized inverse of the smearing matrix $S$.

In general, the resulting solutions are unstable, with unacceptable fluctuations for small changes in $\vec{w}$. To regularize the problem and avoid unphysical fluctuations, two additional terms, a regularization term and a normalization term, are introduced in the procedure [44,45]. For both the $\Delta y^2$ and the $\Delta \eta$ variables, we use independent unfolding procedures based on the respective observable.

The performance of the unfolding algorithm is tested in sets of pseudoexperiments, each of which provides a randomly-generated sample distribution. The number of events from each contributing process is determined through a random number from a Gaussian distribution centered around the measured event rate given in Table 1, with a width corresponding to the respective uncertainty. To take statistical variations into account, the number of expected events is defined by a Poisson distribution around the chosen Gaussian means. This final number of events for each process is drawn randomly from the appropriately simulated events to generate distributions for each pseudoexperiment. Each generated distribution is then subjected to the unfolding procedure described above. For all pseudoexperiments, we subtract the same number of background events as found in data.

We perform 50,000 pseudoexperiments and compare the unfolded spectrum with the generated distribution in each experiment. The average asymmetry from these pseudoexperiments agrees well with the true asymmetry in the sample used to model the signal component and the pull distributions agree with expectations, indicating that the treatment of uncertainties is consistent with Gaussian behaviour. To test the unfolding procedure for different asymmetries, we reweight the events of the default $t\bar{t}$ sample according to their $\Delta \eta$ or $\Delta y^2$ value, to artificially introduce asymmetries between $-0.2$ and $+0.2$, and then perform 50,000 pseudoexperiments for each of the reweighted distributions. We find a linear dependence of the ensemble mean on the input value. While for $\Delta \eta$ the agreement is excellent, for $\Delta y^2$ we observe a slope for the linear dependence of 0.94 instead of 1.0, necessitating a correction of 1/0.94 to the measured asymmetry. The statistical uncertainties of the measurements are found to be independent of the generated asymmetries.

8. Estimation of systematic uncertainties

The measured charge asymmetry $A_C$ can be affected by several sources of systematic uncertainty. Influences on the direction of the reconstructed top-quark momenta can change the value of the reconstructed charge asymmetry. Systematic uncertainties with an impact on the differential selection efficiency can also bias the result, while the overall selection efficiency and acceptance may not. Variations in the background rates can also change the asymmetry attributed to the signal. Since the uncorrected asymmetries observed in the selected data events are close to zero, such changes have only a small influence. To evaluate each source of systematic uncertainty, we perform studies on pseudoexperiments using samples with systematically shifted parameters, and unfolding the distributions of interest as done with data.

The corrections on jet-energy scale (JES) and jet-energy resolution (JER) are changed by $\pm 1$ standard deviations of their $\eta$ and $p_T$-dependent uncertainties for all simulated signal and background events to estimate their effects on the measurement. Similarly, to estimate differences between simulation and data, other dedicated $t\bar{t}$ samples are generated using different renormalization and factorization scales ($Q^2$), jet-matching thresholds [34], and initial and final-state radiation (ISR and FSR). The systematic uncertainties on the measured asymmetry from the choice of parton distribution functions (PDF) for the colliding protons are estimated using the CTEQ6.6 [46] PDF set and the LHAPDF [47] package. In addition, the impact of the uncertainty on b-tagging efficiency, lepton selections, and lepton-trigger efficiencies are examined, taking their $\eta$ dependence into account. The uncertainty on the multijet background obtained from data, and on the frequency of occurrence of pile-up events, are also estimated. A mismodelling of $E_T^{miss}$ and $M3$ in the simulation could change the estimation of the background rates. The measured charge asymmetry is found to be stable under such variations, as verified by shifting the amount of background in the pseudoexperiments. The estimations of all the systematic uncertainties are summarized in Table 2. The largest contribution arises from the jet-matching threshold, which is changed by factors of 2 and 0.5. Other important effects are from uncertainties in the $Q^2$ scale, from ISR and FSR in the simulated $t\bar{t}$ sample, and from the uncertainty in the multijet-background model.

9. Results

Table 3 gives the values of the uncorrected asymmetries, the asymmetries after background subtraction (BG-subtracted), and the

### Table 2

Listed are the positive and negative shifts on $A_C$ induced by systematic uncertainties in the pseudoexperiments from the different sources and the total.

<table>
<thead>
<tr>
<th>Source</th>
<th>$A_C$</th>
<th>$A_C$</th>
<th>$A_C$</th>
<th>$A_C$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$-\text{Shift}$</td>
<td>$+\text{Shift}$</td>
<td>$-\text{Shift}$</td>
<td>$+\text{Shift}$</td>
</tr>
<tr>
<td>JES</td>
<td>$-0.003$</td>
<td>$0.000$</td>
<td>$-0.007$</td>
<td>$0.000$</td>
</tr>
<tr>
<td>JER</td>
<td>$-0.002$</td>
<td>$0.000$</td>
<td>$-0.001$</td>
<td>$+0.001$</td>
</tr>
<tr>
<td>$Q^2$ scale</td>
<td>$-0.005$</td>
<td>$+0.008$</td>
<td>$-0.013$</td>
<td>$0.000$</td>
</tr>
<tr>
<td>ISR/FSR</td>
<td>$-0.006$</td>
<td>$+0.003$</td>
<td>$0.000$</td>
<td>$+0.024$</td>
</tr>
<tr>
<td>Jet-matching threshold</td>
<td>$-0.034$</td>
<td>$+0.021$</td>
<td>$-0.026$</td>
<td>$+0.014$</td>
</tr>
<tr>
<td>PDF</td>
<td>$-0.001$</td>
<td>$+0.001$</td>
<td>$-0.001$</td>
<td>$+0.001$</td>
</tr>
<tr>
<td>b-Tagging efficiency</td>
<td>$-0.001$</td>
<td>$+0.003$</td>
<td>$0.000$</td>
<td>$+0.001$</td>
</tr>
<tr>
<td>Lepton ID/SEL. efficiency</td>
<td>$-0.002$</td>
<td>$+0.004$</td>
<td>$-0.002$</td>
<td>$+0.003$</td>
</tr>
<tr>
<td>Multijet-background model</td>
<td>$-0.008$</td>
<td>$+0.008$</td>
<td>$-0.006$</td>
<td>$+0.006$</td>
</tr>
<tr>
<td>Pile-up</td>
<td>$-0.002$</td>
<td>$+0.002$</td>
<td>$0.000$</td>
<td>$0.000$</td>
</tr>
<tr>
<td>Total</td>
<td>$-0.036$</td>
<td>$+0.025$</td>
<td>$-0.031$</td>
<td>$+0.029$</td>
</tr>
</tbody>
</table>
Table 3
The measured asymmetries for both observables at the different stages of the analysis and the corresponding theory predictions. The final result for $A_C^\eta$ is corrected for a small bias observed in the dependence of the reconstructed value on the true value.

<table>
<thead>
<tr>
<th>Asymmetry</th>
<th>$A_C^\eta$</th>
<th>$A_C^y$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uncorrected</td>
<td>$-0.004 \pm 0.009$ (stat.)</td>
<td>$-0.004 \pm 0.009$ (stat.)</td>
</tr>
<tr>
<td>BG-subtracted</td>
<td>$-0.009 \pm 0.010$ (stat.)</td>
<td>$-0.007 \pm 0.010$ (stat.)</td>
</tr>
<tr>
<td>Final corrected</td>
<td>$-0.017 \pm 0.032$ (stat.) $^{+0.025}_{-0.026}$ (syst.)</td>
<td>$-0.013 \pm 0.028$ (stat.) $^{+0.029}_{-0.031}$ (syst.)</td>
</tr>
<tr>
<td>Theory predictions</td>
<td>$0.0136 \pm 0.0008$</td>
<td>$0.0115 \pm 0.0006$</td>
</tr>
</tbody>
</table>

Fig. 4. Unfolded $\Delta|\eta|$ (left) and $\Delta y^2$ (right) normalized spectra. The NLO prediction is based on the calculations of Ref. [12]. The last bins include the sum of all contributions for $|\Delta|\eta| > 4.0$ and $|\Delta y^2| > 4.0$, respectively. The uncertainties shown on the data are statistical, while the uncertainties on the prediction account also for the dependence on the top-quark mass, PDF, and factorization and renormalization scales.

Fig. 5. Background-subtracted asymmetries for $\Delta|\eta|$ (left) and $\Delta y^2$ (right) as functions of the reconstructed $t\bar{t}$ invariant mass.

10. Summary

A measurement of the charge asymmetry in $t\bar{t}$ production using data corresponding to an integrated luminosity of 1.09 fb$^{-1}$ has been reported. Events with top-quark pairs decaying in the lepton+jets channel were selected and a full $t\bar{t}$ event reconstruction was performed to determine the four-momenta of the top quarks and antiquarks. The measured distributions of the sensitive observables were then subjected to a regularized unfolding procedure to extract the asymmetry values, corrected for acceptance and reconstruction. The measured asymmetries are

$$A_C^\eta = -0.017 \pm 0.032 \text{ (stat.)}^{+0.025}_{-0.036} \text{ (syst.),}$$

and

$$A_C^y = -0.013 \pm 0.028 \text{ (stat.)}^{+0.029}_{-0.031} \text{ (syst.),}$$

consistent with SM predictions. The background-subtracted asymmetry shows no statistically significant dependence on the reconstructed $t\bar{t}$ invariant mass.

Acknowledgements

We thank Johann H. Kühn and German Rodrigo for very fruitful discussions, and we wish to congratulate our colleagues in the
CERN accelerator departments for the excellent performance of the LHC machine. We thank the technical and administrative staff at CERN and other CMS institutes, and acknowledge support from:

FMSR (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, and FAPESP (Brazil); MES (Bulgaria); CERN; CAS, MoST, and NSFC (China); COlCIOlENCIAS (Colombia); MSES (Croacia); RPF (Cyprus); Academy of Sciences and NICPB (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); BMBF, MPG, and HGF (Germany); GSRT (Greece); OTKA and NKTH (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); NRF and WCU (Korea); LAS (Lithuania); CINVESTAV, CONACYT, SEP, and UASLP-FAI (Mexico); MSI (New Zealand); PAEC (Pakistan); SCSR (Poland); FCT (Portugal); JINR (Armenia, Belarus, Georgia, Ukraine, Uzbekistan); MST, MAE and RFBR (Russia); MSTD (Serbia); MICINN and CPAN (Spain); Swiss Funding Agencies (Switzerland); NSE (Taipei); TUBITAK and TAEK (Turkey); STFC (United Kingdom); DOE and NSF (USA).

Individuals have received support from the Marie-Curie programme and the European Research Council (European Union); the Leventis Foundation; the A.P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; the Flemish Research Foundation (IWT); the Leventis Foundation; the A.P. Sloan Foundation; the Alexander von Humboldt Foundation; the Flemish Research Foundation (IWT); the Institute for the Promotion of Innovation through Science and Technology (IWT-Belgium); and the Council for Science and Industrial Research, India.

Open access

This article is published Open Access at sciencedirect.com. It is distributed under the terms of the Creative Commons Attribution License 3.0, which permits unrestricted use, distribution, and reproduction in any medium, provided the original authors and source are credited.

References

A. Dimitrov, R. Hadjiiska, A. Karadzhinova, V. Kozhuharov, L. Litov, B. Pavlov, P. Petkov
University of Sofia, Sofia, Bulgaria

Institute of High Energy Physics, Beijing, China

State Key Lab. of Nucl. Phys. and Tech., Peking University, Beijing, China

A. Cabrera, B. Gomez Moreno, A.F. Osorio Oliveros, J.C. Sanabria
Universidad de Los Andes, Bogota, Colombia

N. Godinovic, D. Lelas, R. Plestina, D. Polic, I. Puljak
Technical University of Split, Split, Croatia

Z. Antunovic, M. Dzelalija, M. Kovac
University of Split, Split, Croatia

V. Brigljevic, S. Duric, K. Kadija, J. Luetic, S. Morovic
Institute Rudjer Boskovic, Zagreb, Croatia

A. Attikis, M. Galanti, J. Mousa, C. Nicolaou, F. Ptochos, P.A. Razis
University of Cyprus, Nicosia, Cyprus

M. Finger, M. Finger Jr.
Charles University, Prague, Czech Republic

Y. Assran, A. Ellithi Kamel, S. Khalil, M.A. Mahmoud, A. Radi
Academy of Scientific Research and Technology of the Arab Republic of Egypt, Egyptian Network of High Energy Physics, Cairo, Egypt

A. Hektor, M. Kadastr, M. Müntel, M. Raidal, L. Rebane, A. Tiko
National Institute of Chemical Physics and Biophysics, Tallinn, Estonia

V. Azzolini, P. Eerola, G. Fedi, M. Voutilainen
Department of Physics, University of Helsinki, Helsinki, Finland

Helsinki Institute of Physics, Helsinki, Finland

K. Banzuzi, A. Korpela, T. Tuuva
Lappeenrant University of Technology, Lappeenranta, Finland

D. Sillou
Laboratoire d’Annecy-le-Vieux de Physique des Particules, IN2P3-CNRS, Annecy-le-Vieux, France

DSM/IRFU, CEA/Saclay, Gif-sur-Yvette, France

Laboratoire Leprince-Ringuet, Ecole Polytechnique, IN2P3-CNRS, Palaiseau, France


Institut Pluridisciplinaire Hubert Curien, Université de Strasbourg, Université de Haute Alsace Mulhouse, CNRS/IN2P3, Strasbourg, France

F. Fassi, D. Mercier

Centre de Calcul de l'Institut National de Physique Nucleaire et de Physique des Particules (IN2P3), Villeurbanne, France

D. Lomidze

Institute of High Energy Physics and Informatization, Tbilisi State University, Tbilisi, Georgia


RWTH Aachen University, I. Physikalisches Institut, Aachen, Germany


RWTH Aachen University, III. Physikalisches Institut A, Aachen, Germany


RWTH Aachen University, III. Physikalisches Institut B, Aachen, Germany


Deutsches Elektronen-Synchrotron, Hamburg, Germany


University of Hamburg, Hamburg, Germany

Institute for Research in Fundamental Sciences (IPM), Tehran, Iran


\[ a \] INFN Sezione di Bari, Bari, Italy
\[ b \] Università di Bari, Bari, Italy
\[ c \] Politecnico di Bari, Bari, Italy


\[ a \] INFN Sezione di Bologna, Bologna, Italy
\[ b \] Università di Bologna, Bologna, Italy

S. Albergo, G. Cappello, M. Chiorboli, S. Costa, R. Potenza, A. Tricomi, C. Tuve

\[ a \] INFN Sezione di Catania, Catania, Italy
\[ b \] Università di Catania, Catania, Italy

G. Barbaglia, V. Ciulli, C. Civinini, R. D’Alessandro, E. Focardi, S. Frosali, E. Gallo, S. Gonzi, M. Meschini, S. Paoletti, G. Sguazzoni, A. Tropiano

\[ a \] INFN Sezione di Firenze, Firenze, Italy
\[ b \] Università di Firenze, Firenze, Italy

L. Benussi, S. Bianco, S. Colafranceschi, F. Fabbri, D. Piccolo

INFN Laboratori Nazionali di Frascati, Frascati, Italy

P. Fabbricatore, R. Musenich

INFN Sezione di Genova, Genova, Italy


\[ a \] INFN Sezione di Milano-Bicocca, Milano, Italy
\[ b \] Università di Milano-Bicocca, Milano, Italy


\[ a \] INFN Sezione di Napoli, Napoli, Italy
\[ b \] Università di Napoli “Federico II”, Napoli, Italy


\[ a \] INFN Sezione di Padova, Padova, Italy
\[ b \] Università di Padova, Padova, Italy
\[ c \] Università di Trento (Trento), Padova, Italy
P. Baesso\textsuperscript{a,b}, U. Berzana\textsuperscript{a}, S.P. Ratti\textsuperscript{a,b}, C. Riccardi\textsuperscript{a,b}, P. Torre\textsuperscript{a,b}, P. Vitul\textsuperscript{a,b}, C. Vivian\textsuperscript{a,b}

\textsuperscript{a} INFN Sezione di Pavia, Pavia, Italy  
\textsuperscript{b} Università di Pavia, Pavia, Italy

M. Biasini\textsuperscript{a,b}, G.M. Bilei\textsuperscript{a}, B. Caponera\textsuperscript{a,b}, L. Fanò\textsuperscript{a,b}, P. Lariccia\textsuperscript{a,b}, A. Lucaroni\textsuperscript{a,b,1}, G. Mantovani\textsuperscript{a,b}, M. Menichelli\textsuperscript{a}, A. Nappi\textsuperscript{a,b}, F. Romeo\textsuperscript{a,b}, A. Santocchia\textsuperscript{a,b}, S. Taroni\textsuperscript{a,b,1}, M. Valdata\textsuperscript{a,b}

\textsuperscript{a} INFN Sezione di Perugia, Perugia, Italy  
\textsuperscript{b} Università di Perugia, Perugia, Italy

P. Azzurra\textsuperscript{a,c}, G. Bagliesi\textsuperscript{a}, T. Boccalia\textsuperscript{a}, G. Broccolo\textsuperscript{a,c}, R. Castaldi\textsuperscript{a}, R.T. D’Agnolo\textsuperscript{a,c}, R. Dell’Orso\textsuperscript{a}, F. Fiors\textsuperscript{a,b}, L. Foà\textsuperscript{a,c}, A. Giassi\textsuperscript{a}, A. Kraan\textsuperscript{a}, F. Ligabue\textsuperscript{a,c}, T. Lomtadze\textsuperscript{a}, L. Martini\textsuperscript{a,27}, A. Messineo\textsuperscript{a,b}, F. Palla\textsuperscript{a}, F. Palmonari\textsuperscript{a}, A. Rizzi\textsuperscript{a}, G. Segneri\textsuperscript{a}, A.T. Serban\textsuperscript{a}, P. Spagnolo\textsuperscript{a}, R. Tenchini\textsuperscript{a}, G. Tonelli\textsuperscript{a,b,1}, A. Venturi\textsuperscript{a,1}, P.G. Verdini\textsuperscript{a}

\textsuperscript{a} INFN Sezione di Pisa, Pisa, Italy  
\textsuperscript{b} Università di Pisa, Pisa, Italy  
\textsuperscript{c} Scuola Normale Superiore di Pisa, Pisa, Italy

L. Barone\textsuperscript{a,b}, F. Cavallari\textsuperscript{a}, D. Del Re\textsuperscript{a,b,1}, M. Diemoz\textsuperscript{a}, C. Fanelli\textsuperscript{a}, D. Franci\textsuperscript{a,b}, M. Grassi\textsuperscript{a,1}, E. Longo\textsuperscript{a,b}, P. Meridiani\textsuperscript{a}, F. Micheli\textsuperscript{a}, S. Nourbakhsh\textsuperscript{a}, G. Organtini\textsuperscript{a,b}, F. Pandolfi\textsuperscript{a,b}, R. Paramatti\textsuperscript{a}, S. Rahatlou\textsuperscript{a,b}, M. Sigamani\textsuperscript{a}, L. Soffi\textsuperscript{a}

\textsuperscript{a} INFN Sezione di Roma, Roma, Italy  
\textsuperscript{b} Università di Roma “La Sapienza”, Roma, Italy

N. Amapane\textsuperscript{a,b}, R. Arcidiacono\textsuperscript{a,c}, S. Argiro\textsuperscript{a,b}, M. Arneodo\textsuperscript{a,c}, C. Biino\textsuperscript{a}, C. Bott\textsuperscript{a,b}, N. Cartiglia\textsuperscript{a}, R. Castello\textsuperscript{a,b}, M. Costa\textsuperscript{a,b}, N. Demaria\textsuperscript{a}, A. Graziano\textsuperscript{a,b}, C. Mariotti\textsuperscript{a,1}, S. Maselli\textsuperscript{a}, E. Migliore\textsuperscript{a,b}, V. Monaco\textsuperscript{a,b}, M. Musich\textsuperscript{a}, M.M. Obertino\textsuperscript{a,c}, N. Pastrone\textsuperscript{a}, M. Pelliccioni\textsuperscript{a}, A. Potenza\textsuperscript{a,b}, A. Romero\textsuperscript{a,b}, M. Ruspa\textsuperscript{a,c}, R. Sacchi\textsuperscript{a,b}, V. Sol\textsuperscript{a,b}, A. Solano\textsuperscript{a,b}, A. Staiano\textsuperscript{a}, A. Vilela Pereira\textsuperscript{a}

\textsuperscript{a} INFN Sezione di Torino, Torino, Italy  
\textsuperscript{b} Università di Torino, Torino, Italy  
\textsuperscript{c} Università del Piemonte Orientale (Novara), Torino, Italy

S. Belforte\textsuperscript{a}, F. Cossutti\textsuperscript{a}, G. Della Ricca\textsuperscript{a,b}, B. Gobbo\textsuperscript{a}, M. Marone\textsuperscript{a,b}, D. Montanino\textsuperscript{a,b,1}, A. Penzo\textsuperscript{a}

\textsuperscript{a} INFN Sezione di Trieste, Trieste, Italy  
\textsuperscript{b} Università di Trieste, Trieste, Italy

S.G. Heo, S.K. Nam

Kangwon National University, Chuncheon, Republic of Korea


Kyungpook National University, Daegu, Republic of Korea

J.Y. Kim, Zero J. Kim, S. Song

Chonnam National University, Institute for Universe and Elementary Particles, Kwangju, Republic of Korea

H.Y. Jo

Konkuk University, Seoul, Republic of Korea


Korea University, Seoul, Republic of Korea

M. Choi, S. Kang, H. Kim, J.H. Kim, C. Park, I.C. Park, S. Park, G. Ryu

University of Seoul, Seoul, Republic of Korea


Sungkyunkwan University, Suwon, Republic of Korea

University of Colorado at Boulder, Boulder, USA


Cornell University, Ithaca, USA

A. Biselli, G. Cirino, D. Winn

Fairfield University, Fairfield, USA


Fermi National Accelerator Laboratory, Batavia, USA


Florida International University, Miami, USA

V. Gaultney, L.M. Lebolo, S. Linn, P. Markowitz, G. Martinez, J.L. Rodriguez

Florida State University, Tallahassee, USA

M.M. Baarmand, B. Dorney, M. Hohlmann, H. Kalakhety, I. Vodopiyano

Florida Institute of Technology, Melbourne, USA


Florida State University, Tallahassee, USA


University of Illinois at Chicago (UIC), Chicago, USA


The University of Iowa, Iowa City, USA


Johns Hopkins University, Baltimore, USA