Measurement of the mass difference between top and anti-top quarks in pp collisions at $\sqrt{s} = 7$ TeV using the ATLAS detector


Published in:
Physics Letters B

DOI:
10.1016/j.physletb.2013.12.010

Citation for published version (APA):
Measurement of the mass difference between top and anti-top quarks in \( pp \) collisions at \( \sqrt{s} = 7 \) TeV using the ATLAS detector

ATLAS Collaboration*

1. Introduction

The CPT symmetry\(^1\) required by a locally gauge-invariant quantum field theory dictates that the masses of all particles and their anti-particles be exactly equal. Any deviation from this would have major implications for particle physics, implying a non-local field theory [1]. Searches for CPT violation both in the \( B \) meson sector [2–5] and with \( K \) mesons [6–8] have not yielded any deviations from the Standard Model (SM). The top quark has the unique property of decaying before hadronization, making it the only quark for which a direct measurement of its mass is possible. The CDF Collaboration measured the mass difference between top and anti-top quarks to be \( \Delta m \approx m_t - m_{\bar{t}} = 3.3 \pm 1.4 \pm 1.0 \) GeV [9], approximately 2 standard deviations away from zero. The D0 Collaboration measured \( \Delta m = 0.8 \pm 1.8 \pm 0.5 \) GeV [10], in agreement with the SM value. The CMS Collaboration recently measured \( \Delta m = -0.44 \pm 0.46 \pm 0.27 \) GeV [11], also in agreement with the SM value. The CDF and D0 analyses used both the top and anti-top quarks within each event to measure \( \Delta m \). In the CMS measurement, the masses of the top and anti-top quarks with hadronic \( W \) boson decays are extracted from two separate samples, split using the lepton charge, and subtracted from one another. In this Letter, the ATLAS Collaboration presents a measurement of this mass difference. The top and anti-top quarks are each taken from the same event, in which a \( t\bar{t} \) pair is produced and decays in the lepton+jets channel.

2. ATLAS detector

ATLAS [12] is a general-purpose particle physics detector with cylindrical geometry covering nearly the entire solid angle around the collision point. Cylindrical coordinates \((r, \phi)\) are used in the transverse plane, where \( \phi \) is the azimuthal angle around the beam pipe. The pseudorapidity is defined as \( \eta \equiv -\ln \tan(\theta/2) \), where \( \theta \) is the polar angle. The transverse mass \((m_\text{T})\) of any two objects is defined as \( m_\text{T} \equiv \sqrt{\not{E}_T^2 + E_T^2 (1 - \cos \Delta \phi)} \), where \( \not{E}_T \) is the object's transverse energy, defined in the plane transverse to the beam axis.

The inner detector (ID) systems, located closest to the interaction region, are immersed in a 2 T axial magnetic field and provide charged particle tracking in the range \(|\eta| < 2.47\). The ID systems consist of a high-granularity silicon pixel detector and a silicon microstrip tracker, as well as a transition radiation tracking (TRT) system. The TRT system provides a high-precision tracking chamber, the muon spectrometer measures the deflection of muons in a magnetic field with a field integral from 2–8 Tm, generated by one barrel and two endcap superconducting air-core toroids. A three-level trigger system is used to select and record interesting events. The level-1 hardware trigger uses a subset of detector information to reduce the event rate resulting from the peak LHC bunch crossing rate of 20 MHz in 2011 to a value of at most 65 kHz. The level-1 trigger is followed by two software-based trigger levels, level-2 and the event filter, which together reduce the event rate to a few hundred Hz for permanent storage and offline analysis.

---

* E-mail address: atlas.publications@cern.ch.

\(^1\) CPT is the combination of three symmetries; Charge conjugation (C), Parity (P) and Time reversal (T).
3. Data sample and event selection

This analysis uses 4.7 ± 0.2 fb⁻¹ [13] of proton–proton collision data recorded by the ATLAS experiment at √s = 7 TeV in 2011. The selected events used in this analysis must contain the signature of a ττ event decaying in the lepton + jets channel. Exactly one charged lepton is required – either a single electron with E_T > 25 GeV, or a single muon with p_T > 20 GeV, where p_T is the object’s transverse momentum, defined in the plane transverse to the beam axis. Energy deposits are selected as electron candidates based on their shower shapes in the electromagnetic calorimeters and on the presence of a good-quality track pointing to them. Electron candidates are required to pass the “tight” quality cuts described in Ref. [14], to fall inside a well-instrumented region of the detector, and to be well isolated from other objects in the event. Muons are required to pass “tight” muon quality cuts [15–17], to be well measured in both the ID and the muon spectrometer, and to be isolated from other objects in the event. Events with an electron (muon) are required to have been triggered by an electron (muon) trigger with an E_T (p_T) threshold of 20 (18) GeV. The selection requirements ensure that triggered events are on the trigger efficiency plateau [18,19].

Jets are reconstructed in the calorimeter using the anti-kT algorithm [20,21] with a radius parameter of 0.4, starting from energy deposits grouped into noise-suppressed topological clusters [22,23]. Jets are required to satisfy p_T > 25 GeV and |η| < 2.5. Events with jets arising from problematic regions in the calorimeters, beam backgrounds and cosmic rays are rejected[24]. Additional corrections are applied after the default ATLAS jet energy calibration [24] to restore on average the partonic energies in ττ events. Jets from the decay of long-lived heavy-flavor hadrons are selected using a multivariate tagging algorithm (b-tagging) [25,26]. The transverse momentum of neutrinos is inferred from the magnitude of the missing transverse momentum (E_T^{miss}) [27].

In addition to the requirement of exactly one charged lepton, the signal selection for this analysis requires four or more jets, at least two of which must be b-tagged. The selected lepton is required to match a trigger object that caused the event to be recorded. To suppress backgrounds from multi-jet events, E_T^{miss} must be larger than 30 (20) GeV in the electron (muon) channel. Further reduction of the multi-jet background in the electron channel is achieved by requiring the transverse mass (m_T) of the lepton and E_T^{miss} to be > 30 GeV. In the muon channel, E_T^{miss} + m_T > 60 GeV is required.

4. Simulated samples and background estimation

The ATLAS detector simulation [28], based on GEANT4 [29], is used to process simulated signal and background events. Simulated minimum bias collisions are overlaid on top of the hard scatter process, and events are reweighted so that the distribution of the average number of interactions (typically 5–20, see Ref. [30]) per bunch crossing matches the distribution observed in data.

Simulated samples of ττ events are produced using PYTHIA v6.425 [31] with Δm ranging from −15 GeV to +15 GeV. In total, 15 such samples are used, with decreasing granularity at large |Δm|. Near Δm = 0, the granularity is 0.3 GeV. In these samples, the average top quark mass (m_{top}) is set to 172.5 GeV. The underlying-event tune used is AUET2B [32], and the parton distribution function (PDF) set is MRST [33]. Despite being a leading order generator, PYTHIA is used because it allows generation of events where the masses of the top and anti-top quarks are not equal. Non-zero widths as predicted by the SM for the corresponding top and anti-top quark masses are included in the event generation.

Pseudo-experiments and additional checks for systematic uncertainties are performed with a SM ττ sample with Δm = 0 generated using MC@NLO [34,35] v4.01 interfaced to HERWIG v6.520 [36] and JIMMY v4.31 [37]. Except for multi-jet processes, Monte Carlo simulations are used to study and estimate the backgrounds. The background from production of single W bosons in association with jets is studied using ALPGEN v2.13 [38] interfaced to HERWIG and JIMMY. The MLM matching scheme [39] is used to form inclusive W+jets samples, taking appropriate care to remove overlapping events in heavy-flavor phase space stemming from both the hard scatter and the showering. Diboson events are generated using HERWIG. Single-top events are generated using MC@NLO in the s- and Wt-channels, and AcerMC v3.8 [40] in the t-channel. The distribution of the multi-jet background is taken from a control region in data where leptons are required to be semi-isolated and have large impact parameter (d_0) and impact parameter significance (d_0 divided by its uncertainty) with respect to the collision vertex. The semi-isolated selection requires the scalar p_T sum for tracks in a cone of 0.3 around the electron (muon) divided by its E_T to be between 0.1 and 0.3. The normalization of this background is obtained from a likelihood fit to the E_T^{miss} distribution in data [41].

5. Kinematic fits

In order to measure a quantity sensitive to the mass difference Δm between the top and anti-top quarks, the kinematic χ² fitter described below is used to reconstruct the ττ system from the observed lepton, E_T^{miss} and jets. The assignment of the selected jets to the partons from the ττ decay uses knowledge of the over-constrained ττ system with the reconstructed top/anti-top quark mass difference (Δm_{fit}) as a free parameter in each event. In the kinematic fitter, the p_T of the lepton and jets is allowed to fluctuate within uncertainties determined from simulated ττ events. The average top quark mass is fixed, but the individual t and ¯t masses are allowed to fluctuate while being constrained by the predicted top quark width. The masses of the two reconstructed W bosons are also allowed to vary within the W boson width. The fit is applied by examining all jet-parton assignments (from among the five leading jets) consistent with the b-jet assignment and minimizing the following χ²:

\[ \chi^2 = \sum_{i=1}^{4} \frac{(p_T^{i,fit} - p_T^{i,meas})^2}{\sigma_T^i} + \sum_{j=x,y} \frac{(p_T^{j,fit} - p_T^{j,meas})^2}{\sigma_T^j} + \sum_{k=1}^{9} \frac{(m_{b,j}^{meas} - m_W)^2}{\sigma_W^2} + \frac{(m_{b,j}^{meas} - m_t)^2}{\sigma_t^2} + \frac{(m_{b,j}^{meas} - m_{t\bar{t}})^2}{\sigma_{t\bar{t}}^2} \]

(1)

\[ \Delta m_{fit} = q_t \times (m_{b,j}^{fit} - m_{b,j}^{meas}) \]

(2)

In total, 15 signal samples were generated with a Δm of ±15, ±10, ±5, ±3, ±1, ±0.6, ±0.3 and 0 GeV.
where \( p_{T}^{i,\text{fit}} \) and \( p_{T}^{i,\text{meas}} \) are the fitted and measured \( p_{T} \) of the jets and the charged lepton, and \( \sigma_{i} \) is the uncertainty on those values. The unclustered energy in the calorimeter \( (E_{U}) \) is defined as a quantity that includes all energy not associated with the primary lepton or the jets and is used to correct \( E_{T}^{\text{miss}} \). The width of the W boson (\( \sigma_{W} \)) is set to the PDG value \([42]\), and the top quark width \( (\sigma_{t}) \) is set to the value predicted from theory. The top quark mass \( (m_{t}) \) is fixed to 172.5 GeV, and the W boson mass \( (m_{W}) \) is set to \( m_{W} = 80.42 \) GeV. The value of \( m_{t}^{\text{fit}} \) is the fitted dijet (lepton-neutrino) mass from the hadronic (leptonic) W boson decay, and \( m_{tj}\) and \( m_{tjj} \) are the fitted top quark masses with leptonically and hadronically decaying W bosons. The value of the mass difference between the hadronic- and leptonic-side top quarks is a free parameter in the fit. In each event, the single jet–parton assignment with the lowest \( \chi^{2} \) is used, and the fitted value of \( \Delta m_{t}^{i,\text{fit}} \) is taken as an observable to measure the true \( \Delta m_{t} \). As seen in Eq. (2), \( \Delta m_{t}^{i,\text{fit}} \) is calculated from the product of the lepton charge \( (q_{l}) \) and the difference between \( m_{tj}^{i,\text{fit}} \) and \( m_{tjj}^{i,\text{fit}} \). Events with \( \chi^{2} > 10 \) for the best jet–parton assignment are considered to be poorly measured or background, and are rejected. The value of this cut is chosen based on studies of simulated signal events, and the efficiency of the \( \chi^{2} \) selection is estimated in simulation to be 55% for \( t\bar{t} \) signal events and 31% for background events. Table 1 shows the expected and observed number of events after all selection requirements, including the \( \chi^{2} \) cut, are applied.

Distributions of \( \Delta m_{t}^{i,\text{fit}} \) are produced for all background samples as well as for a number of simulated \( t\bar{t} \) samples generated with different \( \Delta m \).

The \( \Delta m_{t}^{i,\text{fit}} \) distributions in the signal samples are parameterized in templates by fitting the sum of two Gaussians, where the narrow one corresponds to the correct jet–parton pairing, and the wide one corresponds to an incorrect pairing. The widths of the two Gaussians are quadratic functions of \( \Delta m \) (symmetric about \( \Delta m = 0 \)). The means of the two Gaussians are fit to linear functions of \( \Delta m \). The relative weight of the two Gaussians is fit to \( \Delta m \) using a quadratic function symmetric about \( \Delta m = 0 \). Fig. 1 shows the parameterization for five different values of \( \Delta m \).

The choice of background parameterization has only a small impact on the fits due to the small background in the double b-tag channel. The signal and background templates are used to model the probability density distributions in \( \Delta m \).

6. Likelihood fit

An unbinned extended maximum likelihood fit to the distribution of \( \Delta m_{t}^{i,\text{fit}} \) is performed to extract \( \Delta m \), as well as the expected number of signal \( (n_{s}) \) and background \( (n_{b}) \) events in the data.

Given the data \( D \), which contain \( N \) values of \( \Delta m_{t}^{i,\text{fit}} \), the probability distribution function for signal \( (p_{s}) \) and background \( (p_{b}) \) are used to write down a likelihood \( (L) \):

\[
L(D|n_{s}, n_{b}, \Delta m) = q(N, n_{s} + n_{b}) \times \prod_{i=1}^{N} \frac{n_{p_{s}}(\Delta m_{t}^{i,\text{fit}}|\Delta m_{t}^{i,\text{fit}} + n_{b} p_{b}(\Delta m_{t}^{i,\text{fit}}))}{n_{s} + n_{b}}
\]

where \( q(N, n_{s} + n_{b}) \) is the Poisson probability to observe \( N \) events given \( n_{s} + n_{b} \) expected events and the product over \( i \) is over the \( N \) reconstructed events. The likelihood is maximized over all three parameters \( (n_{s}, n_{b}, \Delta m) \). Ensembles of pseudo-experiments are run to ensure that the fits are unbiased and return correct statistical uncertainties. The widths of pull distributions are consistent with unity. Due to the use of \textsc{Pythia} to generate templates and \textsc{mc@nlo} to run ensemble tests, a 175 MeV offset is applied to all pseudo-experiments (and to the nominal fit result) to return an unbiased measurement, with the statistical uncertainty of 50 MeV on this calibration taken as a systematic uncertainty. The 175 MeV offset is the average difference between the \textsc{mc@nlo} samples with the top and anti-top quark masses reweighted to the distributions in \textsc{Pythia} for a given mass difference. When running pseudo-experiments, events are drawn directly from the simulated samples and not from the parameterizations in order to check for any potential bias.

The extended maximum likelihood fit is applied to the full 2011 dataset, yielding the result shown in Fig. 3. The value of 175 MeV

---

Table 1

<table>
<thead>
<tr>
<th>Channel</th>
<th>Muon</th>
<th>Electron</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data</td>
<td>8854</td>
<td>4941</td>
</tr>
<tr>
<td>SM t\bar{t} \rightarrow W^{+} b W^{-} b</td>
<td>7700±1600</td>
<td>4500±1000</td>
</tr>
<tr>
<td>W/Z + jets</td>
<td>320 ± 90</td>
<td>160 ± 40</td>
</tr>
<tr>
<td>Single top</td>
<td>300 ± 50</td>
<td>170 ± 30</td>
</tr>
<tr>
<td>Diboson</td>
<td>5 ± 1</td>
<td>3 ± 1</td>
</tr>
<tr>
<td>Multijet</td>
<td>220 ± 110</td>
<td>110 ± 60</td>
</tr>
<tr>
<td>Total expected (SM)</td>
<td>8550±1600</td>
<td>4900±1000</td>
</tr>
</tbody>
</table>

---

![Figure 1](ATLAS Collaboration / Physics Letters B 728 (2014) 363–379: Parameterization of \( \Delta m_{t}^{i,\text{fit}} \) for simulated \( t\bar{t} \) samples with different values of \( \Delta m \).)

![Figure 2](ATLAS Collaboration / Physics Letters B 728 (2014) 363–379: Parameterized background template.)
Systematic uncertainties on $\Delta m$ due to differences in the detector response to jets from $b$- and $b$-quarks are difficult to evaluate with in-situ means in the $t\bar{t}$ environment due to correlations with $\Delta m$. Based on the evaluation of the jet energy scale uncertainty from single-hadron response measurements [43], most differences between the calorimeter response to the two types of jets are expected to be small; exceptions are discussed below. One such difference could come from the different responses to positively and negatively charged kaons, which occur at different rates in jets from $b$- and $b$-quarks. The interaction cross sections for $K^+$ and $K^−$ in the calorimeters are different. Such effects are studied by comparing convolutions of the kaon spectra in $b$- and $b$-jets from $t\bar{t}$ events with the expected calorimeter response to kaons simulated with various hadron shower simulation models, as specified in Ref. [43]. The resulting uncertainty is 80 MeV. Uncertainties due to fragmentation and the decay of $b$-hadrons can also lead to uncertainties in the particle content and hadron momentum spectra, and thus in the calorimeter response. This uncertainty is evaluated by comparing powheg samples that use EvtGen [44] and Pythia to decay $b$-hadrons, and is estimated to be 340 MeV. The EvtGen particle decay simulation implements different hadron decay models and up-to-date $b$-hadron decay tables. An additional 80 MeV is assigned to account for any residual difference in response between jets from $b$ and $b$ quarks due to effects not considered above. Parton shower and additional fragmentation uncertainties are estimated by comparing powheg samples interfaced with Herwig to those interfaced with Pythia.

Other uncertainties are small compared to those from differences between jets from $b$- and $b$-quarks. The uncertainty on $\Delta m$ from the uncertainty on the $b$-tagging efficiency is measured by varying the $b$-tag scale factors, which correct simulated efficiencies to those measured in data, within 1σ of their uncertainties. The systematic effects from uncertain light- and $b$-jet energy scales and resolutions are small, as they affect the top and anti-top quark masses in the same way [45,46]. Generator uncertainties are estimated by comparing pseudo-experiments using MC@NLO and Powheg. A systematic uncertainty on the amount of QCD radiation is derived from AcerMC $t\bar{t}$ samples that have varying amounts of initial- and final-state radiation [47]. Uncertainties from the template parameterization are estimated by varying the parameters within their uncertainties, and are found to be small. The systematic uncertainties due to background shape and rate are estimated by replacing the $W$ + jets background used in pseudo-experiments with the shape from the multi-jet background and by varying the normalization within uncertainties. A small systematic uncertainty due to the parton distribution functions of the proton is evaluated by taking the envelope of the MSTW2008NLO [48], NNPDF2.3 [49] and CTEQ6.6 [50] PDF set uncertainties, following the PDF4LHC recommendations [51]. Asymmetries due to lepton energy scales are negligible. A systematic uncertainty on the top quark mass of 40 MeV is estimated by comparing pseudo-experiments where the input average top quark mass is shifted up and down by 1.5 GeV. Other systematic uncertainties considered are those caused by the uncertainty on the lepton identification and reconstruction.

8. Conclusions

The analysis described in this Letter is the first measurement by ATLAS of the mass difference between the top and anti-top quarks using event-by-event quantities in $t\bar{t}$ events. It is based on 4.7 fb$^{-1}$ of 7 TeV proton–proton collisions at the LHC. The mass difference, $\Delta m$, is calculated using a kinematic $\chi^2$ fitter. The measured mass difference is $\Delta m \equiv m_t - m_{\bar{t}} = 0.67 \pm 0.61$ (stat) $\pm 0.41$ (syst) GeV, consistent with the SM expectation of no mass difference.
Acknowledgements

We thank CERN for the very successful operation of the LHC, as well as the support staff from our institutions without whom ATLAS could not be operated efficiently.

We acknowledge the support of ANPCyT, Argentina; YerPhI, Armenia; ARC, Australia; BMWF and FWF, Austria; ANAS, Azerbaijan; SSTC, Belarus; CNPq and FAPESP, Brazil; NSERC, NRC and CFI, Canada; CERN; CONICYT, Chile; CAS, MOST and NSFC, China; COLCIENCIAS, Colombia; MSMT CR, MPO CR and VSC CR, Czech Republic; DNRF, DNSRC and Lundbeck Foundation, Denmark; EPLANET, ERC and NSRF, European Union; IN2P3-CNRS, CEA-DSM/IRFU, France; GNSF, Georgia; BMBF, DFG, MPG and AvH Foundation, Germany; GSRT and NSRF, Greece; ISF, MINERVA, GIF, DIP and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; FOM and NWO, Netherlands; BRF and RCN, Norway; MNISW, Poland; GRICES and FCT, Portugal; MNE/IFA, Romania; MES of Russia and ROSATOM, Russian Federation; JINR; MSTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DST/NRF, South Africa; MINECO, Spain; SRC and Wallenberg Foundation, Sweden; SER, SNSF and Cantons of Bern and Geneva, Switzerland; NSC, Taiwan; TAEK, Turkey; STFC, the Royal Society and Leverhulme Trust, United Kingdom; DOE and NSF, United States of America.

The crucial computing support from all WLCG partners is acknowledged gratefully, in particular from CERN and the ATLAS Tier-1 facilities at TRIUMF (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), ASGC (Taiwan), RAL (UK) and BNL (USA) and in the Tier-2 facilities worldwide.

References


