Combination of measurements of the top quark mass from data collected by the ATLAS and CMS experiments at $\sqrt{s} = 7$ and 8 TeV

The ATLAS and CMS Collaborations

CERN

(Dated: March 19, 2024)

A combination of fifteen top quark mass measurements performed by the ATLAS and CMS experiments at the LHC is presented. The data sets used correspond to an integrated luminosity of up to 5 and 20 fb$^{-1}$ of proton-proton collisions at center-of-mass energies of 7 and 8 TeV, respectively. The combination includes measurements in top quark pair events that exploit both the semileptonic and hadronic decays of the top quark, and a measurement using events enriched in single top quark production via the electroweak $t$-channel. The combination accounts for the correlations between measurements and achieves an improvement in the total uncertainty of 31% relative to the most precise input measurement. The result is $m_t = 172.52 \pm 0.14$ (stat) $\pm 0.30$ (syst) GeV, with a total uncertainty of 0.33 GeV.

The mass of the top quark ($m_t$) is a fundamental parameter of the standard model (SM). Its precise measurement provides a crucial input to fits that probe the consistency of the SM [1–5]. The Tevatron experiments CDF and D0 were the first to measure $m_t$ [6, 7], and produced a combined result in 2016 [8]. During the 2011–2012 data-taking period of the CERN LHC, proton-proton collisions at $\sqrt{s} = 7$ and 8 TeV produced large numbers of top quarks in pairs via strong interactions or singly via electroweak processes. The two general-purpose experiments at the LHC, ATLAS [9] and CMS [10], performed multiple measurements of $m_t$ using these data [11–24]. In this Letter, a combined $m_t$ measurement from the ATLAS and CMS experiments is published for the first time. The 15 input measurements utilize up to 5 (20) fb$^{-1}$ of integrated luminosity per experiment at 7 (8) TeV. A detailed estimate of the correlations between the ATLAS and CMS measurements is performed and the measurements are combined using the best unbiased estimate (BLUE) method [25, 26]. Not included in this combination are recent measurements of $m_t$ performed with a partial 13 TeV data set [24, 27–33], for which the correlations have not yet been studied in detail. These measurements include new analysis techniques and the most precise measurement to date, made by CMS, with an uncertainty of 0.37 GeV [33].

The final state of events containing top quarks is determined by the decay of the $W$ bosons produced in the top quark decays. In the top quark pair ($t\bar{t}$) production mode, ATLAS and CMS have made measurements in the dilepton ($t\bar{t} \rightarrow l^+\nu_l b\ell^−\bar{\nu}_{\ell}$), lepton+jets ($t\bar{t} \rightarrow l^+\nu_l b q\bar{q}$), and all-jets ($t\bar{t} \rightarrow q\bar{q}^* b q\bar{q}$) channels. In addition, CMS has performed a measurement using single top quark ($t \rightarrow l^+\nu_b b \rightarrow ℓ^−\nu\bar{b}$) events.

In the dilepton channel, ATLAS uses the average invariant mass of the two lepton and $b$-tagged jet pairs as the observable sensitive to $m_t$ [15, 18], where a $b$-tagged jet is any reconstructed jet identified as being likely to originate from a $b$ quark. At $\sqrt{s} = 7$ TeV, CMS uses a kinematic reconstruction with the analytical matrix weighting technique [12] and at $\sqrt{s} = 8$ TeV CMS performs a fit to two dedicated observables [22] to simultaneously extract $m_t$ and the global jet energy scale (JES). In the lepton+jets channel [11, 15, 16, 23], both experiments perform a kinematic fit on an event-by-event basis to reconstruct the top quark mass and the invariant mass of the hadronically decaying $W$ boson. The latter is used to constrain the global JES. In addition, ATLAS fits a scale factor for the relative JES between jets originating from $b$ quarks ($b$ quark jets) and light quark / gluon jets [15, 23]. For the all-jets channel, ATLAS uses the ratio of the reconstructed top quark mass to the reconstructed $W$ boson mass [14, 20] to extract $m_t$, while CMS fits the reconstructed top quark mass [13] directly, and at 8 TeV exploits the larger data sample to constrain the global JES using the reconstructed $W$ boson mass [16].

The CMS single top quark analysis fits the invariant mass of the lepton, neutrino, and $b$-tagged jet [21] to extract $m_t$. Two additional CMS measurements [17, 19] use observables built only from leptons and charged-particle tracks, resulting in $m_t$ measurements with low sensitivity to the JES uncertainties. The $J/\psi$ analysis uses the invariant mass of the lepton and the two muons from the $J/\psi$ meson decay [19]. The secondary vertex analysis uses the invariant mass of the lepton and the charged particles from a displaced secondary vertex [17]. Both measurements use $t\bar{t}$ events from the dilepton and lepton+jets decay modes.

All $m_t$ measurements are calibrated using Monte Carlo (MC) simulation. Matrix element (ME) calculations are performed at fixed order in quantum chromodynamics (QCD) and interfaced to a parton shower (PS) algorithm that provides resummation of soft and collinear QCD radiation and a hadronization model that simulates the nonperturbative formation of hadrons. The POWHEG [34–36] generator at next-to-leading-order (NLO) in the strong coupling constant is interfaced with PYTHIA6 [37] to simulate $t\bar{t}$ production in the ATLAS measurements. The CMS measurements use the MADGRAPH5 [38] generator, which includes leading-order
partons, also interfaced with PYTHIA. The top quark
mass is a renormalization-scheme-dependent parameter
in perturbative quantum field theory. The precise iden-
tification of the \( m_t \) parameter in MC simulations within
a field-theoretic mass scheme is the subject of theoretical
studies [39–42].

The BLUE method defines the estimator \( m_t = \sum w_i m^i_t \) for the input measurements \( m^i_t \). The weights
\( w_i \) are determined by minimizing the uncertainty in \( m_t \),
where the covariance between each pair of measurements
is the crucial input. The individual analyses \( i \) are
defined to be orthogonal, such that each measurement
is statistically uncorrelated with every other measurement.
The exception is the CMS secondary vertex analysis [17],
which overlaps statistically with the dilepton and lep-
ton+jets measurements [16, 22]. Given the different
nature of the observable in the secondary vertex analysis,
the analyses are assumed to be uncorrelated. Taking
the maximal statistical correlation allowed by the overlap
produces no significant impact on the combination.

The measurements are affected by similar systematic
uncertainties, and the assessment of their correlation is
central to the combination. As the treatment of system-
atic uncertainties differs between ATLAS and CMS, for
each measurement they are mapped onto 25 categories
that group together similar sources of uncertainties. Un-
certainty categories can influence \( m_t \) in opposite direc-
tions for different measurements, as seen in the ATLAS
combinations [15, 23], and this effect is included via nega-
tive correlations. The correlations between pairs of mea-
surements from a single experiment for each category are
evaluated by summing the covariance matrices of all the
input uncertainty sources, mainly using the correlation
assumptions discussed in Refs. [16, 23]. Differences rel-
ative to Refs. [16, 23] are discussed in the Supplemental
Material [43]. Each input uncertainty source is included
irrespective of whether it is statistically significant.

The correlation strength \( \rho \) between ATLAS and CMS
for each uncertainty category is assessed based on the
similarities of the underlying models and methods, and
of the estimates used. Three different cases are identi-
fied, with corresponding assumed correlation strengths:
\( \rho = 0.85 \) (strongly correlated), \( \rho = 0.5 \) (partially
 correlated), and \( \rho = 0 \) (uncorrelated). No category was iden-
tified to have \( \rho = 1 \), which reflects the many differences
between the two experiments. The correlation coefficient
between an ATLAS and CMS measurement for each cat-
egory is the product of the respective correlation strength
and the signs of the impacts of that category on each mea-
surement. In this way, for a given pair of measurements,
categories that impact \( m_t \) in the same (opposite) direction
have a positive (negative) correlation. For categories
composed of multiple uncertainty sources (e.g., \( b \) tagging
in ATLAS), the sign of the combined impact is not de-
termined. In this case, the sign of the combined impact
is assumed to be positive and it was checked that taking
the alternative assumption of a negative sign does not
significantly impact the result, with the largest change
in the central value (uncertainty) being 41 (7) MeV. In
calculating the final covariance matrix, it is assumed that
each category is uncorrelated to the others.

Table I displays the correlation strengths between AT-
LAS and CMS for each systematic uncertainty cate-
gory, and the Supplemental Material [43] provides ta-
bles with the uncertainties for all 15 measurements.
The corresponding correlation coefficients are available
in HEPData [44, 45]. The subsequent paragraphs outline
the categorization of systematic uncertainties and their
corresponding correlation assessments.

<table>
<thead>
<tr>
<th>Uncertainty category</th>
<th>( \rho )</th>
<th>Scan range</th>
<th>( \Delta m_t/2 ) [MeV]</th>
<th>( \Delta \sigma_{m_t}/2 ) [MeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td>JES 1</td>
<td>0</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>JES 2</td>
<td>0</td>
<td>([-0.25, +0.25])</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>JES 3</td>
<td>0.5</td>
<td>([+0.25, +0.75])</td>
<td>1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>b-JES</td>
<td>0.85</td>
<td>([+0.5, +1])</td>
<td>26</td>
<td>5</td>
</tr>
<tr>
<td>g-JES</td>
<td>0.85</td>
<td>([+0.5, +1])</td>
<td>2</td>
<td>&lt;1</td>
</tr>
<tr>
<td>t-JES</td>
<td>0</td>
<td>([-0.25, +0.25])</td>
<td>1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>CMS JES 1</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>JER</td>
<td>0</td>
<td>([-0.25, +0.25])</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Leptons</td>
<td>0</td>
<td>([-0.25, +0.25])</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>b tagging</td>
<td>0.5</td>
<td>([+0.25, +0.75])</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>( p_T^{miss} )</td>
<td>0</td>
<td>([-0.25, +0.25])</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Pileup</td>
<td>0.85</td>
<td>([+0.5, +1])</td>
<td>2</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Trigger</td>
<td>0</td>
<td>([-0.25, +0.25])</td>
<td>&lt;1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>ME generator</td>
<td>0.5</td>
<td>([+0.25, +0.75])</td>
<td>&lt;1</td>
<td>4</td>
</tr>
<tr>
<td>QCD radiation</td>
<td>0.5</td>
<td>([+0.25, +0.75])</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Hadronization</td>
<td>0.5</td>
<td>([+0.25, +0.75])</td>
<td>1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>CMS b hadron ( B )</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Color reconnection</td>
<td>0.5</td>
<td>([+0.25, +0.75])</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Underlying event</td>
<td>0.5</td>
<td>([+0.25, +0.75])</td>
<td>1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>PDF</td>
<td>0.85</td>
<td>([+0.5, +1])</td>
<td>1</td>
<td>&lt;1</td>
</tr>
<tr>
<td>CMS top quark ( p_T )</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Background (data)</td>
<td>0</td>
<td>([-0.25, +0.25])</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Background (MC)</td>
<td>0.85</td>
<td>([+0.5, +1])</td>
<td>2</td>
<td>&lt;1</td>
</tr>
<tr>
<td>Method</td>
<td>0</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Other</td>
<td>0</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

The JES uncertainty is important in many \( m_t \) mea-
surements. Six categories are used to describe the uncer-
tainties associated with the calibration of the JES that are in common between the experiments [46–49]. The category JES 1 includes contributions from the limited size of the data samples used to derive the JES corrections and contributions due to pileup and its time-dependent variation. For ATLAS (7 TeV only), it also includes an uncertainty term from the effects of closely-by-jet activity. This category is uncorrelated between ATLAS and CMS measurements. The category JES 2 corresponds to the uncertainties from the absolute JES determined using $\gamma/Z+$jets events that are not included in JES 1. There are significant differences between the ATLAS and CMS approaches [46, 47], including differences in the jet radius, treatment of muons in jets, and methods to correct for additional radiation. Hence, this category is treated as uncorrelated. The category JES 3 corresponds to the modeling uncertainty in the relative $\eta$ intercalibration [48, 49]. Both experiments use dijet events for this calibration, and the modeling uncertainty originates from the use of different generators to predict the radiation patterns in these events. As similar but not identical generators and techniques are used in both experiments, JES 3 is treated as partially correlated.

The remaining JES categories correspond to the flavor-dependent calibration uncertainties. The category $b$-JES corresponds to the jet energy response uncertainty for $b$ quark jets. The category $g$-JES corresponds to the uncertainty in the jet response of gluon jets for CMS and the uncertainty in the difference of the jet response of gluons to light-quark ($u$, $d$, $s$, $c$) jets for ATLAS. In both cases, MC comparisons determine the flavor-dependent effects, hence a strong correlation is used for the $b$-JES and $g$-JES components. The category $l$-JES includes the combined CMS uncertainty in the jet response of light-quark jets and the ATLAS uncertainty for the flavor composition of jets in $t\bar{t}$ events. As these uncertainty sources are different, the $l$-JES component is treated as uncorrelated. One additional flavor uncertainty category CMS JES 1 is included for the CMS 7 TeV measurements, corresponding to the full envelope of the response dependencies for gluons and all quark flavors.

The jet energy resolution (JER) uncertainty affects all measurements, and one category is used for the corresponding uncertainties. ATLAS and CMS both measure the JER using data [48, 49], hence this category is treated as uncorrelated.

The energy scale, efficiency, and resolution of leptons affect the $m_t$ measurements, and one category is used for the corresponding uncertainties. ATLAS and CMS both calibrate the lepton energy scales, resolutions, and efficiencies using resonances that decay into dilepton pairs. Since the calibration samples are independent between the two experiments, and detector technologies and reconstruction algorithms are different, this category is treated as uncorrelated.

The selection criteria for many top quark measurements make use of $b$ tagging. The uncertainty in the efficiency and rejection rate of these algorithms can impact the $m_t$ measurements, and one category is used for the corresponding uncertainties. Both collaborations use dijet events to calibrate the $b$-tagging efficiency, employing equivalent methods [50, 51] that depend on similar simulation setups. As the ATLAS $b$ jet calibration (unlike the CMS one) also uses $t\bar{t}$ events [50], this category is assessed as partially, rather than strongly, correlated.

The missing transverse momentum ($p_{T}^{\text{miss}}$) is estimated in the two experiments with different algorithms. Thus, the uncertainty in the $p_{T}^{\text{miss}}$ scale originating from energy deposits not included in the reconstruction of jets or leptons is treated as uncorrelated.

The high instantaneous luminosity of the LHC results in multiple interactions in each bunch crossing (pileup). As the modeling of pileup relies on simulation, the correlation between ATLAS and CMS is assessed to be strong. While for other categories, the correlation strength is independent of the data set, the pileup category has zero correlation between analyses performed at 7 and 8 TeV due to the different pileup conditions in the two data sets.

The uncertainty in the efficiency of the triggers used to select events typically have a small impact on the measurements. As the triggers are calibrated in independent data sets, the uncertainty is treated as uncorrelated between ATLAS and CMS.

The $m_t$ measurements rely on MC simulation of $t\bar{t}$ events to relate the reconstructed observables to $m_t$. The corresponding modeling uncertainties are encompassed in seven uncertainty categories. The category ME generator includes uncertainties originating from the choice of the ME generator. ATLAS assesses this uncertainty by comparing the results obtained using an MC@NLO [52, 53] sample with the POWHEG sample. CMS assesses this uncertainty by comparing the results obtained using a POWHEG sample with the MadGraph sample. As the experiments employ different nominal MC models, the category is treated as partially correlated. The category QCD radiation includes uncertainty sources for the modeling of QCD radiation in $t\bar{t}$ events. For the ATLAS measurements, samples with parameter variations of the initial- and final-state radiation in PYTHIA, and the $h_{\text{damp}}$ parameter in POWHEG (which controls ME/PS matching and effectively regulates high-$p_{T}$ QCD radiation) are used to evaluate these uncertainties. For the CMS measurements, samples with variations of the factorization, renormalization, and matching scales are used. Similarly to the ME category, the QCD radiation category is treated as partially correlated between the two experiments.

In the ATLAS analyses, the uncertainty originating from the hadronization model is evaluated by using an alternative PS and hadronization generator (POWHEG+HERWIG6 [54]). CMS addresses similar un-
uncertainties by separately varying the $b$ quark fragmentation function and the semileptonic branching ratios (CMS $b$ hadron $B$). As the ATLAS approach changes many aspects of the simulation that are not changed in the two CMS uncertainty sources and the PYTHIA settings in the two experiments are not the same, there is no clear mapping and correlation for these sources. Nevertheless, some degree of correlation is expected, hence the ATLAS hadronization uncertainty is grouped with the CMS uncertainty from the fragmentation model in the category hadronization and this category is assumed to be partially correlated between the experiments. The CMS $b$ hadron $B$ uncertainty source is treated as uncorrelated with the ATLAS uncertainties. It was verified that the alternative treatment of correlating the ATLAS hadronization uncertainty with the CMS $b$ hadron $B$ uncertainty had no significant impact on the result.

The uncertainties associated with color reconnection and the underlying event tunes are included in separate categories. The experiments use different PYTHIA settings for the nominal simulation, and these uncertainty categories are taken to be partially correlated. The uncertainty in the parton distribution functions (PDFs) is driven by the input data used in the PDF extractions, and hence this category is taken as strongly correlated between ATLAS and CMS. The CMS analyses account for an uncertainty in the modeling of the top quark $p_T$ distribution, represented by a separate category, while for the ATLAS analyses, the alternative MC sample used to evaluate the hadronization uncertainty covers the disagreement between data and simulation [55], and no additional uncertainty is evaluated.

The analyses typically have small contributions from background processes, and background uncertainties have only a small impact on the measurements. Uncertainties in backgrounds estimated from data control samples are included in the category Background (data), treated as uncorrelated between the experiments. Both ATLAS and CMS rely on MC simulation for several backgrounds. The uncertainties in these are included in the category Background (MC), assumed to be strongly correlated.

Every analysis ensures that the $m_t$ fit is unbiased. This is done using simulated samples generated with different $m_t$ values. The limited sample size introduces a systematic uncertainty (Method) that is statistical and hence uncorrelated between measurements.

A few systematic uncertainties affect only a limited number of analyses (see the Supplemental Material [43]). These uncertainty sources are in the category Other, which is uncorrelated between ATLAS and CMS.

The measurements from each experiment are separately combined, with the ATLAS combination giving $m_t = 172.71 \pm 0.25 \text{ (stat)} \pm 0.41 \text{ (syst)} \text{ GeV}$ and the CMS combination giving $m_t = 172.52 \pm 0.14 \text{ (stat)} \pm 0.39 \text{ (syst)} \text{ GeV}$. The ATLAS combination is very similar to, and supersedes, the result in Ref. [23], with the slight difference originating from changes in the correlation assumptions that are discussed in the Supplemental Material [43]. The CMS measurement is improved compared to the previous combination [16] and supersedes that result. The improvement originates from including a more precise dilepton measurement at 8 TeV together with the single top, secondary vertex, and $J/\psi$ meson measurements, and from including the effect of anticorrelations of the systematic uncertainties between the input measurements. It was verified that performing the combinations with a likelihood-based approach [56] gives identical results.

The combination of all 15 input measurements gives

$$m_t = 172.52 \pm 0.14 \text{ (stat)} \pm 0.30 \text{ (syst)} \text{ GeV},$$

which is compared with the input measurements in Fig. 1. The LHC combination has the same statistical uncertainty as the CMS combination. This is because the figure of merit in BLUE is the total uncertainty, and the statistical component is a consequence of the optimized weights in the combination. The difference in statistical uncertainty between ATLAS and CMS reflects different analysis choices, as explained in the Supplemental Material [43].

The combination achieves an improvement in the total $m_t$ uncertainty of 31% relative to the most precise input measurement. The measurements with the largest weight
in the combination are the CMS 8 TeV lepton+jets (0.34), dilepton (0.12), and all-jets (0.12) results, and the ATLAS 8 TeV lepton+jets (0.17) and dilepton (0.16) measurements. The hierarchy of the weights originates from the uncertainty of each measurement, as well as the correlation between measurements. For example, the ATLAS 8 TeV lepton+jets measurement has a higher weight than the corresponding dilepton measurement, despite having a larger uncertainty. This is because of the smaller correlation with the precise CMS 8 TeV lepton+jets measurement. The combination shows good compatibility between the measurements, with $\chi^2 = 7.5$ and a corresponding $p$-value of 91%. The LHC combination is much closer to the CMS combination than the ATLAS one because the relative weights of the measurements with slightly lower measured $m_t$ are higher in the LHC combination than in the per-experiment combinations. All weights and the individual pulls can be found in the Supplemental Material [43], along with a combination where all 15 measurements are used to extract separate $m_t$ values for ATLAS and CMS.

Table II shows the breakdown of the systematic uncertainty in the combined measurement and the individual ATLAS and CMS combinations. The largest systematic uncertainties are seen to originate from JES, $b$ tagging, and $t\bar{t}$ modeling. The stability of the measurement against the correlation assumptions is checked by varying the correlation strengths for each uncertainty category as shown in Table I. The ranges reflect the extent of the understanding of the correlations. No variation is performed for categories where there is no ambiguity in the correlation assumption. Table I shows the variation in the total uncertainty and central value of the combination under those changes. Both the central value and uncertainty are observed to vary linearly under the variations and the changes are small (<30 MeV) compared to the uncertainty in $m_t$. The largest change in central value is seen for $b$-JES, which is the leading correlated uncertainty source in the combination.

The consistency of the result and the measurements from the different decay channels have been checked by performing the combination with a separate $m_t$ parameter for each $t\bar{t}$ decay channel. The results are also shown in Fig. 1, and the $m_t$ values are found to be consistent.

The impact of the limited statistical precision of the estimates of the systematic uncertainties is evaluated by performing pseudo-experiments where the systematic uncertainties of the measurements are varied according to their uncertainties and the combination procedure is repeated. In this procedure, changes in the sign of the impacts of systematic uncertainties are propagated to the signs of the corresponding correlations. The root-mean-square of the measured $m_t$ ($\sigma_{m_t}$) is found to be 63 (19) MeV, demonstrating the stability of the combination.

<table>
<thead>
<tr>
<th>Uncertainty category</th>
<th>LHC</th>
<th>ATLAS</th>
<th>CMS</th>
</tr>
</thead>
<tbody>
<tr>
<td>$b$-JES</td>
<td>0.18</td>
<td>0.17</td>
<td>0.25</td>
</tr>
<tr>
<td>$b$ tagging</td>
<td>0.09</td>
<td>0.16</td>
<td>0.03</td>
</tr>
<tr>
<td>ME generator</td>
<td>0.08</td>
<td>0.13</td>
<td>0.14</td>
</tr>
<tr>
<td>JES 1</td>
<td>0.08</td>
<td>0.18</td>
<td>0.06</td>
</tr>
<tr>
<td>JES 2</td>
<td>0.08</td>
<td>0.11</td>
<td>0.10</td>
</tr>
<tr>
<td>Method</td>
<td>0.07</td>
<td>0.06</td>
<td>0.09</td>
</tr>
<tr>
<td>CMS $b$ hadron $B$</td>
<td>0.07</td>
<td>...</td>
<td>0.12</td>
</tr>
<tr>
<td>QCD radiation</td>
<td>0.06</td>
<td>0.07</td>
<td>0.10</td>
</tr>
<tr>
<td>Leptons</td>
<td>0.05</td>
<td>0.08</td>
<td>0.07</td>
</tr>
<tr>
<td>JER</td>
<td>0.05</td>
<td>0.09</td>
<td>0.02</td>
</tr>
<tr>
<td>CMS top quark $p_T$</td>
<td>0.05</td>
<td>...</td>
<td>0.07</td>
</tr>
<tr>
<td>Background (data)</td>
<td>0.05</td>
<td>0.04</td>
<td>0.06</td>
</tr>
<tr>
<td>Color reconnection</td>
<td>0.04</td>
<td>0.08</td>
<td>0.03</td>
</tr>
<tr>
<td>Underlying event</td>
<td>0.04</td>
<td>0.03</td>
<td>0.05</td>
</tr>
<tr>
<td>$g$-JES</td>
<td>0.03</td>
<td>0.02</td>
<td>0.04</td>
</tr>
<tr>
<td>Background (MC)</td>
<td>0.03</td>
<td>0.07</td>
<td>0.01</td>
</tr>
<tr>
<td>Other</td>
<td>0.03</td>
<td>0.06</td>
<td>0.01</td>
</tr>
<tr>
<td>$t$-JES</td>
<td>0.03</td>
<td>0.01</td>
<td>0.05</td>
</tr>
<tr>
<td>CMS JES 1</td>
<td>0.03</td>
<td>...</td>
<td>0.04</td>
</tr>
<tr>
<td>Pileup</td>
<td>0.03</td>
<td>0.07</td>
<td>0.03</td>
</tr>
<tr>
<td>JES 3</td>
<td>0.02</td>
<td>0.07</td>
<td>0.01</td>
</tr>
<tr>
<td>Hadronization</td>
<td>0.02</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>$p_T^{miss}$</td>
<td>0.02</td>
<td>0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>PDF</td>
<td>0.02</td>
<td>0.06</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Trigger</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Total systematic</td>
<td>0.30</td>
<td>0.41</td>
<td>0.39</td>
</tr>
<tr>
<td>Statistical</td>
<td>0.14</td>
<td>0.25</td>
<td>0.14</td>
</tr>
<tr>
<td>Total</td>
<td>0.33</td>
<td>0.48</td>
<td>0.42</td>
</tr>
</tbody>
</table>

The understanding of top quark production and decay has continued to evolve since the publication of the measurements used in this combination. Developments in the simulations include improved modeling of off-shell effects [57], reduced uncertainties in additional QCD radiation [58, 59], new models of color reconnection [60, 61], MC simulations at next-to-NLO precision in QCD [62], and investigations into the radiation patterns in the top quark decay [63]. Advancements in the modeling, which may either increase or decrease the mass uncertainty, and improvements in analysis techniques [29, 33] are being incorporated into analyses performed at $\sqrt{s} = 13$ TeV, but this is not possible for the analyses used in this combination. A cross-check, detailed in the Supplemental Material [43], was performed to verify that potential modeling uncertainties in the recoil in the top quark decay [63] do not significantly affect the combination.

In summary, a combination of top quark mass measurements by the ATLAS and CMS experiments at the
CERN LHC in proton-proton collisions at $\sqrt{s} = 7$ and 8 TeV has been performed. The combination yields $m_t = 172.52 \pm 0.33$ GeV, which is the most precise result to date.

CMS congratulates our colleagues in the CERN accelerator departments for the excellent performance of the LHC and thank the technical and administrative staffs at CERN and at other CMS institutes for their contributions to the success of the CMS effort. In addition, we gratefully acknowledge the computing centres and personnel of the Worldwide LHC Computing Grid and other centres for delivering so effectively the computing infrastructure essential to our analyses. Finally, we acknowledge the enduring support for the construction and operation of the LHC, the CMS detector, and the supporting computing infrastructure provided by the following funding agencies: SC (Armenia); BMBWF and FWF (Austria); FNRS and FWO (Belgium); CNPq, CAPES, FAPERJ, FAPERGS, and FAPESP (Brazil); MES and BNSF (Bulgaria); CERN; CAS, MoST, and NSFC (China); MINCIENCIAS (Colombia); MSES and CSF (Croatia); RIF (Cyprus); SENESCYT (Ecuador); ERC PRG, RVTT3 and TK202 (Estonia); Academy of Finland, MEC, and HIP (Finland); CEA and CNRS/IN2P3 (France); SRNSF (Georgia); BMBF, DFG, and HGF (Germany); GSRI (Greece); NKFIH (Hungary); DAE and DST (India); IPM (Iran); SFI (Ireland); INFN (Italy); MSIP and NRF (Republic of Korea); MES (Latvia); LAS (Lithuania); MOE and UM (Malaysia); BUAP, CINVESTAV, CONACYT, LNS, SEP, and UASLP-FAI (Mexico); MOS (Montenegro); MBIE (New Zealand); PAEC (Pakistan); MES and NSC (Poland); FCT (Portugal); MESTD (Serbia); MCIN/AEI and PCTI (Spain); MEYS CR, Czech Republic; MST (Taipei); MHESI and NSTDA (Thailand); TUBITAK and TEMNAK (Turkey); NASU (Ukraine); STFC (United Kingdom); DOE and NSF (USA).

ATLAS thanks 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; CNPq and FAPERJ, Brazil; NSERC, NRC and CFI, Canada; CERN; ANID, Chile; CAS, MOST and NSFC, China; Minciencias, Colombia; MEYS CR, Czech Republic; DNRF and DNSRC, Denmark; IN2P3-CNRS and CEA-DRF/IRFU, France; SRNSFG, Georgia; BMBF, HGF and MPG, Germany; GSRI, Greece; RGC and Hong Kong SAR, China; ISF and Benoziyo Center, Israel; INFN, Italy; MEXT and JSPS, Japan; CNRST, Morocco; NWO, Netherlands; RCN, Norway; MEIN, Poland; FCT, Portugal; MNE/IFA, Romania; MESTD, Serbia; MSSR, Slovakia; ARRS and MIZŠ, Slovenia; DSI/NRF, South Africa; MICINN, Spain; SRC and Wallenberg Foundation, Sweden; SERI, SNSF and Canton of Bern and Geneva, Switzerland; MOST, Taipei; TENMAK, Türkiye; STFC, United Kingdom; DOE and NSF, United States of America. In addition, individual groups and members have received support from BCKDF, CANARIE, CRC and DRAC, Canada; PRIMUS 21/SCI/017 and UNCE SCI/013, Czech Republic; COST, ERC, ERDF, Horizon 2020, ICSC-NextGenerationEU and Marie Skłodowska-Curie Actions, European Union; Investissements d’Avenir Label, Investissements d’Avenir Idex and ANR, France; DFG and AvH Foundation, Germany; Herakleitos, Thales and Aristeia programmes co-financed by EU-ESF and the Greek NSRF, Greece; BSF-NSF and MINERVA, Israel; Norwegian Financial Mechanism 2014-2021, Norway; NCN and NAWA, Poland; La Caixa Banking Foundation, CERCA Programme Generalitat de Catalunya and PROMETEO and GenT Programmes Generalitat Valenciana, Spain; Göran Gustafssons Stiftelse, Sweden; The Royal Society and Leverhulme Trust, United Kingdom.

The crucial computing support from all WLCG partners is acknowledged gratefully by ATLAS, in particular from CERN, the ATLAS Tier-1 facilities at TRIUMF/SFU (Canada), NDGF (Denmark, Norway, Sweden), CC-IN2P3 (France), KIT/GridKA (Germany), INFN-CNAF (Italy), NL-T1 (Netherlands), PIC (Spain), RAL (UK) and BNL (USA), the Tier-2 facilities worldwide and large non-WLCG resource providers. Major contributors of computing resources are listed in Ref. [64].

6. F. Abe et al. (CDF), Observation of top quark production in $p\bar{p}$ collisions with the Collider Detector at Fermilab, Phys. Rev. Lett. 74, 2626 (1995).
7. S. Abachi et al. (D0), Observation of the top quark, Phys. Rev. Lett. 74, 2632 (1995).
8. CDF and D0 Collaborations, Combination of CDF and...


[59] CMS Collaboration, Investigations of the impact of the parton shower tuning in PYTHIA 8 in the modelling of $t\bar{t}$ at $\sqrt{s} = 8$ and 13 TeV, CMS Physics Analysis Summary CMS-PAS-TOP-16-021 (2016).

[60] S. Argyropoulos and T. Sjöstrand, Effects of color reconnection on $t\bar{t}$ final states at the LHC, J. High Energy Phys. 2014 (11), 043.


