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 $m_t = 172.52 \pm 0.14\,(\text{stat}) \pm 0.30\,(\text{syst})$ 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 linear 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 ($t\bar{t} \to \ell^+\nu b \ell^-\bar{\nu}\bar{b}$), lepton+jets ($t\bar{t} \to \ell^\pm\nu b q\bar{q}'\bar{b}$), and all-jets ($t\bar{t} \to q\bar{q}' b q\bar{q}'\bar{b}$) channels. In addition, CMS has performed a measurement using single top quark ($t \to \ell^+\nu b$, $\bar{t} \to \ell^-\bar{\nu}\bar{b}$) 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/ψ analysis uses the invariant mass of the lepton and the two muons from the J/ψ 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

(LO) terms for $t\bar{t}$ production with up to three additional partons, also interfaced with PYTHIA6. The top quark mass is a renormalization-scheme-dependent parameter in perturbative quantum field theory. The precise identification 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 $m_t = \sum_i w^i m_t^i$ for the input measurements m_t^i. 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 lepton+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 systematic uncertainties differs between ATLAS and CMS, for each measurement they are mapped onto 25 categories that group together similar sources of uncertainties. Uncertainty categories can influence m_t in opposite directions for different measurements, as seen in the ATLAS combinations [15, 23], and this effect is included via negative correlations. The correlations between pairs of measurements 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 relative 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 ρ 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 identified, with corresponding assumed correlation strengths: $\rho = 0.85$ (strongly correlated), $\rho = 0.5$ (partially correlated), and $\rho = 0$ (uncorrelated). No category was identified to have $\rho = 1$, which reflects the many differences between the two experiments. The correlation coefficient between an ATLAS and CMS measurement for each category is the product of the respective correlation strength and the signs of the impacts of that category on each measurement. 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 determined. 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 ATLAS and CMS for each systematic uncertainty category, and the Supplemental Material [43] provides tables 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 I. Correlation strengths ρ of the systematic uncertainty categories between ATLAS and CMS, as used in the combination. The categories are defined in the text. Categories indicated with the symbol $\cdots$ in the second column correspond to uncertainties specific to a single experiment. The third column shows the range of ρ scanned for stability checks. The changes in the combination's central value m_t and uncertainty σ_{m_t} corresponding to each correlation variation are shown in the last two columns.

Uncertainty category	ρ	Scan range	$\Delta m_t/2$ [MeV]	$\Delta\sigma_{m_t}/2$ [MeV]
JES 1	0	$\cdots$	$\cdots$	$\cdots$
JES 2	0	$[-0.25, +0.25]$	8	7
JES 3	0.5	$[+0.25, +0.75]$	1	<1
b-JES	0.85	$[+0.5, +1]$	26	5
g-JES	0.85	$[+0.5, +1]$	2	<1
l-JES	0	$[-0.25, +0.25]$	1	<1
CMS JES 1	$\cdots$	$\cdots$	$\cdots$	$\cdots$
JER	0	$[-0.25, +0.25]$	5	1
Leptons	0	$[-0.25, +0.25]$	2	2
b tagging	0.5	$[+0.25, +0.75]$	1	1
$p_{\mathrm{T}}^{\mathrm{miss}}$	0	$[-0.25, +0.25]$	<1	<1
Pileup	0.85	$[+0.5, +1]$	2	<1
Trigger	0	$[-0.25, +0.25]$	<1	<1
ME generator	0.5	$[+0.25, +0.75]$	<1	4
QCD radiation	0.5	$[+0.25, +0.75]$	7	1
Hadronization	0.5	$[+0.25, +0.75]$	1	<1
CMS b hadron $\mathcal{B}$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
Color reconnection	0.5	$[+0.25, +0.75]$	3	1
Underlying event	0.5	$[+0.25, +0.75]$	1	<1
PDF	0.85	$[+0.5, +1]$	1	<1
CMS top quark p_{T}	$\cdots$	$\cdots$	$\cdots$	$\cdots$
Background (data)	0	$[-0.25, +0.25]$	8	2
Background (MC)	0.85	$[+0.5, +1]$	2	<1
Method	0	$\cdots$	$\cdots$	$\cdots$
Other	0	$\cdots$	$\cdots$	$\cdots$

The JES uncertainty is important in many m_t measurements. 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 close-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 γ/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 η 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 ($p_{\mathrm{T}}^{\mathrm{miss}}$) is estimated in the two experiments with different algorithms. Thus, the uncertainty in the $p_{\mathrm{T}}^{\mathrm{miss}}$ 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_{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-

certainties by separately varying the b quark fragmentation function and the semileptonic branching ratios (CMS b hadron $\mathcal{B}$). 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 $\mathcal{B}$ 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 $\mathcal{B}$ 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 $m_t = 172.71 \pm 0.25\,(\mathrm{stat}) \pm 0.41\,(\mathrm{syst})\,\mathrm{GeV}$ and the CMS combination giving $m_t = 172.52 \pm 0.14\,(\mathrm{stat}) \pm 0.39\,(\mathrm{syst})\,\mathrm{GeV}$. 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/ψ 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.

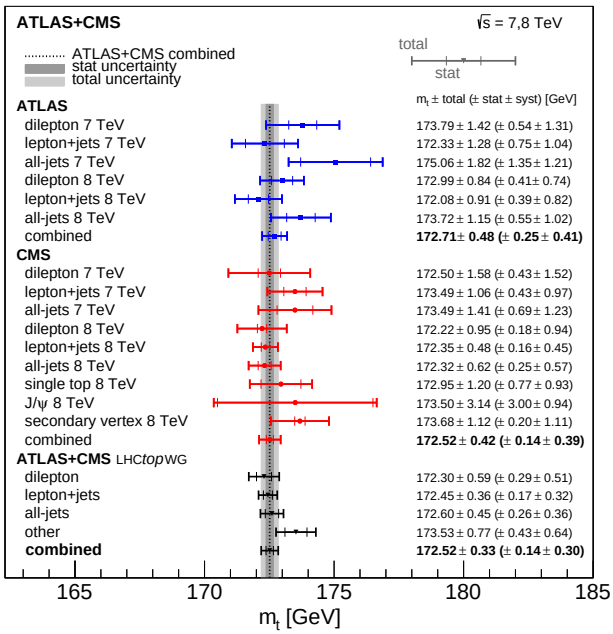

FIG. 1. Comparison of the individual m_t measurements and the result of the m_t combination. Also shown are the separate combinations of each experiment and the result of the simultaneous combination for the different decay channels, where the "other" category covers the single top, J/ψ, and secondary vertex measurements.

The combination of all 15 input measurements gives

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

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 (σ_{m_t}) is found to be 63 (19) MeV, demonstrating the stability of the combination.

TABLE II. Uncertainties on the m_t values extracted in the LHC, ATLAS, and CMS combinations arising from the same categories as listed in Table I, sorted in order of decreasing value of the combined LHC uncertainty.

Uncertainty category	Uncertainty impact [GeV]		
	LHC	ATLAS	CMS
b-JES	0.18	0.17	0.25
b tagging	0.09	0.16	0.03
ME generator	0.08	0.13	0.14
JES 1	0.08	0.18	0.06
JES 2	0.08	0.11	0.10
Method	0.07	0.06	0.09
CMS b hadron $\mathcal{B}$	0.07	···	0.12
QCD radiation	0.06	0.07	0.10
Leptons	0.05	0.08	0.07
JER	0.05	0.09	0.02
CMS top quark p_{T}	0.05	···	0.07
Background (data)	0.05	0.04	0.06
Color reconnection	0.04	0.08	0.03
Underlying event	0.04	0.03	0.05
g-JES	0.03	0.02	0.04
Background (MC)	0.03	0.07	0.01
Other	0.03	0.06	0.01
l-JES	0.03	0.01	0.05
CMS JES 1	0.03	···	0.04
Pileup	0.03	0.07	0.03
JES 3	0.02	0.07	0.01
Hadronization	0.02	0.01	0.01
$p_{\mathrm{T}}^{\mathrm{miss}}$	0.02	0.04	0.01
PDF	0.02	0.06	<0.01
Trigger	0.01	0.01	0.01
Total systematic	0.30	0.41	0.39
Statistical	0.14	0.25	0.14
Total	0.33	0.48	0.42

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); MOSTR (Sri Lanka); Swiss Funding Agencies (Switzerland); MST (Taipei); MHESI and NSTDA (Thailand); TUBITAK and TENMAK (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; BMWFW and FWF, Austria; ANAS, Azerbaijan; CNPq and FAPESP, 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 Cantons 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 Labex, 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].

[1] ALEPH, CDF, D0, DELPHI, L3, OPAL, and SLD Collaborations, LEP Electroweak Working Group, Tevatron Electroweak Working Group, and SLD electroweak and heavy flavour groups, Precision electroweak measurements and constraints on the standard model, arXiv:1012.2367 [hep-ex] (2010).

[2] G. Degrassi, S. Di Vita, J. Elias-Miró, J. R. Espinosa, G. F. Giudice, G. Isidori, and A. Strumia, Higgs mass and vacuum stability in the standard model at NNLO, J. High Energy Phys. **2012** (08), 098.

[3] F. Bezrukov, M. Y. Kalmykov, B. A. Kniehl, and M. Shaposhnikov, Higgs boson mass and new physics, J. High Energy Phys. **2012** (10), 140.

[4] F. L. Bezrukov and M. Shaposhnikov, The standard model Higgs boson as the inflaton, Phys. Lett. B **659**, 703 (2008).

[5] A. De Simone, M. P. Hertzberg, and F. Wilczek, Running inflation in the standard model, Phys. Lett. B **678**, 1 (2009).

[6] F. Abe *et al.* (CDF), Observation of top quark production in $\bar{p}p$ 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

D0 results on the mass of the top quark using up to 9.7 fb^{-1} at the Tevatron, arXiv:1608.01881 [hep-ex] (2016).

[9] ATLAS Collaboration, The ATLAS experiment at the CERN Large Hadron Collider, J. Instrum. **2008** (3), S08003.

[10] CMS Collaboration, The CMS experiment at the CERN LHC, J. Instrum. **2008** (3), S08004.

[11] CMS Collaboration, Measurement of the top-quark mass in $t\bar{t}$ events with lepton+jets final states in pp collisions at $\sqrt{s} = 7\,\text{TeV}$, J. High Energy Phys. **2012** (12), 105.

[12] CMS Collaboration, Measurement of the top-quark mass in $t\bar{t}$ events with dilepton final states in pp collisions at $\sqrt{s} = 7\,\text{TeV}$, Eur. Phys. J. C **72**, 2202 (2012).

[13] CMS Collaboration, Measurement of the top-quark mass in all-jets $t\bar{t}$ events in pp collisions at $\sqrt{s} = 7\,\text{TeV}$, Eur. Phys. J. C **74**, 2758 (2014).

[14] ATLAS Collaboration, Measurement of the top-quark mass in the fully hadronic decay channel from ATLAS data at $\sqrt{s} = 7\,\text{TeV}$, Eur. Phys. J. C **75**, 158 (2015).

[15] ATLAS Collaboration, Measurement of the top quark mass in the $t\bar{t} \to$ lepton+jets and $t\bar{t} \to$ dilepton channels using $\sqrt{s} = 7\,\text{TeV}$ ATLAS data, Eur. Phys. J. C **75**, 330 (2015).

[16] CMS Collaboration, Measurement of the top quark mass using proton-proton data at $\sqrt{s} = 7$ and $8\,\text{TeV}$, Phys. Rev. D **93**, 072004 (2016).

[17] CMS Collaboration, Measurement of the top quark mass using charged particles in pp collisions at $\sqrt{s} = 8\,\text{TeV}$, Phys. Rev. D **93**, 092006 (2016).

[18] ATLAS Collaboration, Measurement of the top quark mass in the $t\bar{t} \to$ dilepton channel from $\sqrt{s} = 8\,\text{TeV}$, Phys. Lett. B **761**, 350 (2016).

[19] CMS Collaboration, Measurement of the mass of the top quark in decays with a J/ψ meson in pp collisions at $8\,\text{TeV}$, J. High Energy Phys. **2016** (12), 123.

[20] ATLAS Collaboration, Top-quark mass measurement in the all-hadronic $t\bar{t}$ decay channel at $\sqrt{s} = 8\,\text{TeV}$ with the ATLAS detector, J. High Energy Phys. **2017** (09), 118.

[21] CMS Collaboration, Measurement of the top quark mass using single top quark events in proton-proton collisions at $\sqrt{s} = 8\,\text{TeV}$, Eur. Phys. J. C **77**, 354 (2017).

[22] CMS Collaboration, Measurement of the top quark mass in the dileptonic $t\bar{t}$ decay channel using the mass observables $M_{b\ell}$, M_{T2}, and $m_{b\ell\nu}$ in pp collisions at $\sqrt{s} = 8\,\text{TeV}$, Phys. Rev. D **96**, 032002 (2017).

[23] ATLAS Collaboration, Measurement of the top quark mass in the $t\bar{t} \to$ lepton+jets channel from $\sqrt{s} = 8\,\text{TeV}$ ATLAS data and combination with previous results, Eur. Phys. J. C **79**, 290 (2019).

[24] CMS Collaboration, Review of top quark mass measurements in CMS, arXiv:2403.01313 [hep-ex] (2024), to be submitted to *Phys. Rept.*

[25] L. Lyons, D. Gibaut, and P. Clifford, How to combine correlated estimates of a single physical quantity, Nucl. Instrum. Meth. A **270**, 110 (1988).

[26] R. Nisius, BLUE: Combining correlated estimates of physics observables within ROOT using the best linear unbiased estimate method, SoftwareX **11**, 100468 (2020).

[27] CMS Collaboration, Measurement of the top quark mass with lepton+jets final states using pp collisions at $\sqrt{s} = 13\,\text{TeV}$, Eur. Phys. J. C **78**, 891 (2018).

[28] CMS Collaboration, Measurement of the top quark mass in the all-jets final state at $\sqrt{s} = 13\,\text{TeV}$ and combination with the lepton+jets channel, Eur. Phys. J. C **79**, 313 (2019).

[29] CMS Collaboration, Measurement of the $t\bar{t}$ production cross section, the top quark mass, and the strong coupling constant using dilepton events in pp collisions at $\sqrt{s} = 13\,\text{TeV}$, Eur. Phys. J. C **79**, 368 (2019).

[30] CMS Collaboration, Measurement of the jet mass distribution and top quark mass in hadronic decays of boosted top quarks in pp collisions at $\sqrt{s} = 13\,\text{TeV}$, Phys. Rev. Lett. **124**, 202001 (2020).

[31] CMS Collaboration, Measurement of the top quark mass using events with a single reconstructed top quark in pp collisions at $\sqrt{s} = 13\,\text{TeV}$, J. High Energy Phys. **2021** (12), 161.

[32] ATLAS Collaboration, Measurement of the top-quark mass using a leptonic invariant mass in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector, J. High Energy Phys. **2023** (06), 019.

[33] CMS Collaboration, Measurement of the top quark mass using a profile likelihood approach with the lepton+jets final states in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$, Eur. Phys. J. C **83**, 963 (2023).

[34] P. Nason, A new method for combining NLO QCD with shower Monte Carlo algorithms, J. High Energy Phys. **2004** (11), 040.

[35] S. Frixione, P. Nason, and C. Oleari, Matching NLO QCD computations with parton shower simulations: the POWHEG method, J. High Energy Phys. **2007** (11), 070.

[36] S. Alioli, P. Nason, C. Oleari, and E. Re, A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX, J. High Energy Phys. **2010** (06), 043.

[37] T. Sjöstrand, S. Mrenna, and P. Z. Skands, PYTHIA 6.4 physics and manual, J. High Energy Phys. **2006** (05), 026.

[38] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer, and T. Stelzer, MADGRAPH5: going beyond, J. High Energy Phys. **2011** (06), 128.

[39] A. H. Hoang, S. Plätzer, and D. Samitz, On the cutoff dependence of the quark mass parameter in angular ordered parton showers, J. High Energy Phys. **2018** (10), 200.

[40] P. Azzi *et al.*, *Report from working group 1: Standard model physics at the HL-LHC and HE-LHC*, CERN Report CERN-LPCC-2018-03 (2019).

[41] A. H. Hoang, What is the top quark mass?, Ann. Rev. Nucl. Part. Sci. **70**, 225 (2020).

[42] B. Dehnadi, A. H. Hoang, O. L. Jin, and V. Mateu, Top quark mass calibration for Monte Carlo event generators—an update, J. High Energy Phys. **2023** (12), 065.

[43] Supplemental material, link to be supplied by publisher.

[44] E. Maguire, L. Heinrich, and G. Watt, HEPData: a repository for high energy physics data, in *Proc. 22nd International Conference on Computing in High Energy and Nuclear Physics (CHEP2016): San Francisco CA, USA, October 14–16, 2016* (2017) [J. Phys. Conf. Ser. 898 (2017) 102006].

[45] HEPData record for this analysis (2024).

[46] ATLAS and CMS Collaborations, *Jet energy scale uncertainty correlations between ATLAS and CMS*, ATLAS PUB Note ATL-PHYS-PUB-2014-020, CMS Physics Analysis Summary CMS-PAS-JME-14-003 (2014).

[47] ATLAS and CMS Collaborations, *Jet energy scale un-*

certainty correlations between ATLAS and CMS at 8 TeV, ATLAS PUB Note ATL-PHYS-PUB-2015-049, CMS Physics Analysis Summary CMS-PAS-JME-15-001 (2015).

[48] ATLAS Collaboration, Determination of jet calibration and energy resolution in proton-proton collisions at $\sqrt{s} = 8\,\text{TeV}$ using the ATLAS detector, Eur. Phys. J. C **80**, 1104 (2020).

[49] CMS Collaboration, Jet energy scale and resolution in the CMS experiment in pp collisions at 8 TeV, J. Instrum. **2017** (12), P02014.

[50] ATLAS Collaboration, Performance of b-jet identification in the ATLAS experiment, J. Instrum. **2016** (11), P04008.

[51] CMS Collaboration, Identification of b-quark jets with the CMS experiment, J. Instrum. **2013** (8), P04013.

[52] S. Frixione and B. R. Webber, Matching NLO QCD computations and parton shower simulations, J. High Energy Phys. **2002** (06), 029.

[53] S. Frixione, P. Nason, and B. R. Webber, Matching NLO QCD and parton showers in heavy flavour production, J. High Energy Phys. **2003** (08), 007.

[54] G. Corcella, I. G. Knowles, G. Marchesini, S. Moretti, K. Odagiri, P. Richardson, M. H. Seymour, and B. R. Webber, HERWIG 6: an event generator for hadron emission reactions with interfering gluons (including supersymmetric processes), J. High Energy Phys. **2001** (01), 010.

[55] ATLAS Collaboration, Measurements of top-quark pair differential cross-sections in the lepton+jets channel in pp collisions at $\sqrt{s} = 8\,\text{TeV}$ using the ATLAS detector, Eur. Phys. J. C **76**, 538 (2016).

[56] J. Kieseler, A method and tool for combining differential or inclusive measurements obtained with simultaneously constrained uncertainties, Eur. Phys. J. C **77**, 792 (2017).

[57] T. Ježo, J. M. Lindert, P. Nason, C. Oleari, and S. Pozzorini, An NLO+PS generator for $t\bar{t}$ and Wt production and decay including non-resonant and interference effects, Eur. Phys. J. C **76**, 691 (2016).

[58] ATLAS Collaboration, *Improvements in $t\bar{t}$ modelling using NLO+PS Monte Carlo generators for Run 2*, ATLAS PUB Note ATL-PHYS-PUB-2018-009 (2018).

[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.

[61] J. R. Christiansen and P. Z. Skands, String formation beyond leading colour, J. High Energy Phys. **2015** (08), 003.

[62] J. Mazzitelli, P. F. Monni, P. Nason, E. Re, M. Wiesemann, and G. Zanderighi, Top-pair production at the LHC with $\text{MINNLO}_{\text{PS}}$, J. High Energy Phys. **2022** (04), 079.

[63] H. Brooks and P. Skands, Coherent showers in decays of colored resonances, Phys. Rev. D **100**, 076006 (2019).

[64] ATLAS Collaboration, *ATLAS computing acknowledgements*, Tech. Rep. (CERN, Geneva, 2023).

[65] ATLAS Collaboration, The ATLAS simulation infrastructure, Eur. Phys. J. C **70**, 823 (2010).

[66] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, and P. Z. Skands, An introduction to PYTHIA 8.2, Comput. Phys. Commun. **191**, 159 (2015).

[67] R. Nisius, On the combination of correlated estimates of a physics observable, Eur. Phys. J. C **74**, 3004 (2014).