Measurement of the top quark mass using charged particles in pp collisions at root s=8 TeV

Khatchatryan, V.

2016-05-18


http://hdl.handle.net/10138/165006
https://doi.org/10.1103/PhysRevD.93.092006

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.
Measurement of the top quark mass using charged particles in $pp$ collisions at $\sqrt{s} = 8$ TeV

V. Khachatryan et al.*
(CMS Collaboration)
(Received 21 March 2016; published 18 May 2016)

A novel technique for measuring the mass of the top quark that uses only the kinematic properties of its charged decay products is presented. Top quark pair events with final states with one or two charged leptons and hadronic jets are selected from the data set of 8 TeV proton-proton collisions, corresponding to an integrated luminosity of 19.7 fb$^{-1}$. By reconstructing secondary vertices inside the selected jets and computing the invariant mass of the system formed by the secondary vertex and an isolated lepton, an observable is constructed that is sensitive to the top quark mass that is expected to be robust against the energy scale of hadronic jets. The main theoretical systematic uncertainties, concerning the modeling of the fragmentation and hadronization of $b$ quarks and the reconstruction of secondary vertices from the decays of $b$ hadrons, are studied. A top quark mass of $173.68 \pm 0.20$ (stat)$^{+1.58}_{-0.97}$ (syst) GeV is measured. The overall systematic uncertainty is dominated by the uncertainty in the $b$ quark fragmentation and the modeling of kinematic properties of the top quark.

DOI: 10.1103/PhysRevD.93.092006

I. INTRODUCTION

The top quark is the heaviest known elementary particle and as such has a privileged interaction with the Higgs boson. Its mass, $m_t$, is hence an important input to global fits of electroweak parameters together with measurements of the $W$ boson and Higgs boson masses, and serves as an important cross-check of the consistency of the standard model (SM). Moreover, by comparing precision electroweak measurements and theoretical predictions, a precisely measured $m_t$ can place strong constraints on contributions from physics beyond the SM. The top quark is the only colored particle that decays before forming a color-neutral state through hadronization and thus presents a unique opportunity to directly probe the properties of color charges.

Direct determinations of the mass of the top quark have been carried out with ever-increasing precision since it was discovered at the Tevatron by the CDF and D0 experiments \[1,2\]. More recently, the most precise measurements reconstruct top quarks in hadronic decays and calibrate the energy of hadronic jets in situ, using constraints from the reconstructed $W$ boson mass \[3–5\]. Other analyses exploit the purity of leptonic top quark decays and constrain the neutrino momenta analytically \[5,6\]. All four experiments where the top quark mass is being studied (ATLAS, CDF, CMS, and D0) have combined their results in a world average \[7\]. A recent combination of measurements at 7 and 8 TeV by the CMS experiment yields the best determination of the top quark mass to date, with a result of $172.44 \pm 0.48$ GeV, i.e. reaching a precision of 0.28% \[8\].

The most precise top quark mass measurements are systematically limited by experimental uncertainties related to the calibration of reconstructed jet energies and their resolution, with other important uncertainties concerning the modeling of the fragmentation and hadronization of bottom quarks. To improve further the precision of the value of the top quark mass and our understanding of the modeling of top quark decays, the development and application of alternative and complementary methods is essential. Complementarity to “standard” methods can be gained by using observables with reduced sensitivity to certain sources of systematic uncertainties, such as the $b$ hadron decay length \[9–11\] or kinematic properties of leptons \[12\], or by extracting the mass from end points of kinematic distributions \[13\] or from the production cross section \[14\].

This paper describes a measurement performed with the CMS experiment at the CERN LHC that minimizes the sensitivity to experimental systematic uncertainties such as jet energy scale. This is achieved by constructing a mass-dependent observable that uses only the individually measured momenta of charged decay products (tracks) of the top quark. The mass of the top quark is estimated by measuring the invariant mass of a charged lepton from the $W$ boson decay and the tracks used in the reconstruction of a secondary vertex (SV) resulting from the long lifetime of $b$ hadrons. The dependence of the observable on the top quark mass is calibrated using simulated Monte Carlo (MC)
events. This approach is similar to a proposed measurement using the invariant mass of leptons and reconstructed $J/\psi$ mesons [15], but requires a lower integrated luminosity to become sensitive.

The paper is organized as follows: Sec. II describes the experiment, the collected and simulated data, and the event reconstruction and selection; Sec. III describes control region studies of $b$ quark fragmentation and secondary vertex reconstruction; Sec. IV describes the measurement of the top quark mass and the assigned systematic uncertainties; and Sec. V concludes and gives an outlook of prospects in the ongoing LHC run.

II. EXPERIMENTAL SETUP

A. The CMS detector

The central feature of the CMS apparatus is a superconducting solenoid of 6 m internal diameter, providing a magnetic field of 3.8 T. Within the solenoid volume are a silicon pixel and strip tracker, a lead tungstate crystal electromagnetic calorimeter (ECAL), and a brass and scintillator hadron calorimeter (HCAL), each composed of a barrel and two end cap sections. The tracker has a track-finding efficiency of more than 99% for muons with transverse momentum $p_T > 1$ GeV and pseudorapidity $|\eta| < 2.5$. The ECAL is a fine-grained hermetic calorimeter with quasiprojective geometry and is segmented in the barrel region of $|\eta| < 1.48$ and in two end caps that extend up to $|\eta| < 3.0$. The HCAL barrel and end caps similarly cover the region $|\eta| < 3.0$. In addition to the barrel and end cap detectors, CMS has extensive forward calorimetry. Muons are measured in gas-ionization detectors which are embedded in the flux-return yoke outside of the solenoid. A more detailed description of the CMS detector, together with a definition of the coordinate system used and the relevant kinematic variables, can be found in Ref. [16].

B. Data and simulation

This analysis makes use of a large sample of top quark pair, $t\bar{t}$, event candidates with either one or two isolated charged leptons (electrons or muons) in the final state. In the semileptonic (only one lepton) case, at least four reconstructed hadronic jets are required, whereas in the dilepton case at least two jets are required. Events are selected from the data sample acquired in proton-proton ($pp$) collisions at a center-of-mass energy of $\sqrt{s} = 8$ TeV by the CMS experiment throughout 2012, corresponding to an integrated luminosity of 19.7 fb$^{-1}$.

At that energy the predicted $t\bar{t}$ cross section in $pp$ collisions, computed at the next-to-next-to-leading-order (NNLO) quantum chromodynamics (QCD) and including corrections and next-to-next-to-leading-logarithmic resummation accuracy [17], is $245.8^{+4.7}_{-10.6}$ pb for a top quark mass of 173 GeV, where the uncertainty covers missing higher orders in the calculation as well as variations of the parton distribution functions (PDFs). Signal $t\bar{t}$ events are simulated with the leading-order (LO) MadGraph (v5.1.3.30) generator [18] matched to LO Pythia (v6.426) [19] for parton showering and fragmentation. The $t$ lepton decays are simulated with the Tauola package (v27.121.5) [20]. The LO CTEQ6L1 PDF set [21] and the $Z^* + \text{t~t}$ underlying event tune [22] are used in the generation. The $Z^*$ tune is derived from the Z1 tune [23], which uses the CTEQ5L PDF set, whereas $Z^* + \text{t~t}$ adopts CTEQ6L. Matrix elements describing up to three partons in addition to the $t\bar{t}$ pair are included in the generator used to produce the simulated signal samples, and the MLM prescription [24] is used for matching of matrix-element jets to parton showers. Following the generator chain, the response of the CMS detector is simulated using Geant4 (v9.4p03) for both signal and background samples [25].

The most relevant background for the semileptonic channel is the production of a $W$ boson in association with hadronic jets. This background is modeled with MadGraph and normalized to a total cross section of 36.3 nb, computed with Fewz (v3.1) [26] at NNLO. Multijet QCD processes are also relevant and studied in simulations using Pythia. Single top quark processes are modeled with Powheg (v1.0, r1380) [27–31] with the CTEQ6M PDF set and normalized to the cross sections of 22.2, 86.1, and 5.6 pb for the $tW$, $t$, and $s$ production channels, respectively [32]. Charged-lepton production from Drell–Yan (DY) processes is modeled with MadGraph for dilepton invariant masses above 10 GeV and is normalized to a cross section of 4.4 nb, computed with Fewz [33,34]. The production of $WW$, $WZ$, and $ZZ$ pairs is modeled with Pythia and normalized to cross sections of 54.8, 33.2, and 17.7 pb, respectively, computed at next-to-leading order (NLO) accuracy using MCFM (v6.6) [35].

All simulated samples include the effects of pileup, i.e. multiple $pp$ collisions in the same and neighboring beam crossings (within 50 ns) as the generated hard interaction. The distribution of the number of pileup events in simulation matches that in the data and has an average of about 21 interactions per bunch crossing.

C. Event reconstruction and selection

The event selection is designed to identify the $t\bar{t}$ final state in the semileptonic and dileptonic channels. Single-lepton triggers are used to collect the data samples for the semileptonic channels with a minimum $p_T$ of 27 for electrons and 24 GeV for muons. In the dilepton channel double-lepton triggers are required with a minimum $p_T$ of 17 and 8 GeV for the leading and subleading leptons, respectively. In both cases isolation and identification criteria are required at the trigger level. More information can be found in Refs. [14,36].

The events are reconstructed using a particle-flow (PF) algorithm that optimally combines the information from all
subdetectors to reconstruct and identify all individual particles in the event [37,38]. In addition, improved electron and muon reconstruction, identification and calibration algorithms have been employed as described in [39,40]. Electron candidates are required to have $p_T > 30 \text{ GeV}$ and to be in the fiducial region of the detector, i.e. $|\eta| \leq 2.4$. Muon candidates are selected with $p_T > 26 \text{ GeV}$ and $|\eta| \leq 2.1$.

In the dilepton channel these requirements are relaxed to $p_T > 20 \text{ GeV}$ and $|\eta| \leq 2.4$ for all lepton candidates. The track associated with each lepton candidate is required to have an impact parameter compatible with prompt production. A particle-based relative isolation is computed for each lepton and is corrected on an event-by-event basis for contributions from pileup events [14]. The scalar sum of the transverse momenta of all reconstructed particle candidates—except for the leptons themselves—within a cone of size $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} < 0.3$ (less than 0.4 for muons) built around the lepton direction must be less than 10% of the electron $p_T$ and less than 12% of the muon $p_T$. In the dilepton channels, the electron isolation threshold is relaxed to less than 15%. Events in the semileptonic channel are required to have exactly one selected lepton. In the dilepton channel, at least two selected leptons are required.

Jets are reconstructed using the anti-$k_T$ algorithm with a distance parameter of 0.5 and taking PF candidates as input to the clustering [41]. The jet momentum is defined as the vectorial sum of all particle momenta associated to the jet and is determined from the simulation to be within 5%—10% of the generated jet momentum at particle level over the whole $p_T$ range and detector acceptance. An offset correction is applied to take into account the extra energy clustered into the jets due to pileup, following the procedure described in Refs. [42,43]. Jet energy scale corrections are derived from the simulation and are cross-checked with in situ measurements of the energy balance in dijet and photon + jet events. The selected jets are required to have a corrected $p_T$ greater than 30 GeV and $|\eta| \leq 2.5$. Jets within $\Delta R = 0.4$ of any selected lepton are rejected, but the event is retained if it passes the other selection criteria. The magnitude of the vectorial sum of the transverse momenta of all PF candidates reconstructed in the event is used as an estimator of the energy imbalance in the transverse plane, $E_T^{\text{miss}}$.

For each jet, the charged PF candidates used in the clustering are given as input to an adaptive vertex fitter algorithm to reconstruct secondary vertices [44]. Secondary vertex candidates that share 65% or more of their tracks with the primary vertex (defined as the vertex with highest $\sum p_T^2$ of its associated tracks) or that have a flight direction outside a $\Delta R = 0.5$ cone around the jet momentum are rejected. Furthermore, if the radial distance from the primary vertex is greater than 2.5 cm, candidates with an invariant mass consistent with that of a $K^0$, or higher than 6.5 GeV, are rejected (assuming each decay particle to have the rest mass of a charged $K$). In case an event does not have any jet with a valid secondary vertex candidate it is discarded from the analysis.

Secondary vertices are used together with track-based lifetime information in a likelihood ratio algorithm to provide a discriminant for jets originating from the hadronization of a b quark (“b jets”) [45]. The chosen threshold on the discriminant output value has an efficiency for selecting a genuine b jet of about 60%, selects charm-initiated jets with an efficiency of about 15%, while the probability to misidentify a light-flavor jet as a b jet is about 1.5%. Jets passing this selection are referred to as b-tagged.

Events in the three dilepton channels ($e\mu$, $ee$, and $\mu\mu$) are selected with at least two jets, of which at least one is required to have a reconstructed secondary vertex. The dilepton invariant mass is required to be greater than 20 GeV to remove low-mass QCD resonances. To suppress contributions from DY production in the $ee$ and $\mu\mu$ channels, the dilepton mass is further required to differ by at least 15 GeV from the Z boson mass (91 GeV), and $E_T^{\text{miss}} > 40 \text{ GeV}$ is required. In the two semileptonic channels, events are selected with at least four jets, of which at least one has a reconstructed secondary vertex and one more has either another secondary vertex or is b-tagged.

Table I shows the number of selected data events in the five channels and the purity of events containing top quarks as expected from simulation. Figure 1 shows the distribution of the transverse decay length, $L_{xy}$, between the secondary vertex reconstructed from charged-particle tracks inside the jets selected for this analysis and the primary vertex of each event. Good agreement is observed between data and expectations based on $m_t = 172.5$ GeV. The background expectations are obtained from the simulation, except for the multijet background which is determined from a control region in the data, as described in Sec. IV B.

### III. ANALYSIS OF b QUARK FRAGMENTATION IN DATA

The crucial objects used in this measurement are the charged leptons from a W boson decay and the charged decay products of a b hadron, forming a reconstructed secondary vertex. While the reconstruction of leptons is well controlled in the experiment, the modeling of

| Table I. Number of observed events and expected purity of top quark production ($t\bar{t}$ and single top quarks) for the five channels investigated in this analysis. |
|---------------------------------|----------|----------|----------|----------|----------|
|                                 | $e\mu$   | $ee$     | $\mu\mu$ | $e$      | $\mu$    |
| Observed events                 | 31 639   | 9 558    | 10 674   | 103 586  | 117 198  |
| Expected purity                 | 98.6%    | 95.8%    | 95.4%    | 93.7%    | 92.8%    |
hadronization of the colored decay products of the top quark is subject to theoretical uncertainties. These uncertainties affect the kinematic properties of the produced tracks, as well as their flavor composition and multiplicity.

The parton-to-hadron momentum transfer in the hadronization of b quarks—referred to in the following as b quark fragmentation—has been measured before in $e^+e^-$ collisions by the ALEPH, DELPHI, OPAL, and SLD Collaborations [46–50], and in $p\bar{p}$ collisions by the CDF Collaboration [51]. However, no measurement at the LHC has been published so far.

In this section, two complementary studies are presented that attempt to constrain the uncertainties from the modeling of $b$ quark fragmentation, which are expected to be the main contributors to the final uncertainty in this top quark mass measurement. These studies constitute a first step towards measuring the $b$ quark fragmentation using $t\bar{t}$ events, but, as will become clear, the 2012 LHC data do not provide the necessary statistical precision, and significant constraints on the $b$ quark fragmentation will be possible only with future data.

In this study we compare the PYTHIA $Z^+Z^-$ tune, used by the CMS experiment at 8 TeV [22] with an updated version which includes the $e^+e^-$ data to improve the description of the fragmentation. Without the inclusion of this data, the default $Z^+Z^-$ $b$ quark fragmentation function is found to be too soft. The $r_b$ parameter in PYTHIA [FARJ (47)] can be optimized to fit the $e^+e^-$ data using the PROFESSOR tool [52], resulting in a value of $0.591^{+0.216}_{-0.224}$. In contrast, the default central value used in $Z^+Z^-$ is 1.0 [53]. In this analysis, the improved tune using the $r_b$ central value of 0.591 (and variations within the uncertainty band) is denoted as $Z^+Z^-$ LEP $r_b$ ($Z^+Z^-$ LEP $r_b^{\pm}$) and is used to calibrate the measurement and evaluate the systematic uncertainty associated with the calibration. For completeness, we also include other alternatives of the $Z^+Z^-$ tuning using the Peterson and Lund parameterizations [19]. All the considered PYTHIA tunes use the so-called Lund string fragmentation model [54]. The impact on the measurement of $m_t$ when using the alternative cluster model [55,56] is discussed in Sec. IV C 1.

A. Secondary vertex properties in $Z + jets$ and $t\bar{t}$ events

Events with a leptonically decaying $Z$ boson recoiling against hadronic jets provide an independent and low-background sample to study the properties of secondary vertices. Candidate $Z$ events are selected by requiring two opposite-sign leptons with an invariant mass compatible with the $Z$ boson mass within 15 GeV. To minimize effects from mismodeling of kinematic properties of the $Z$ boson, events are reweighted such that the predicted $p_T(Z)$ distribution reflects the one observed in the data. Furthermore, events are required to have a leading jet with $p_T > 30$ GeV that is spatially separated from the $Z$ boson candidate by $\Delta R > 2.1$.

The flavor of jets with reconstructed secondary vertices in such events changes with increasing number of tracks associated with the vertex. From simulation, we expect vertices with two tracks to predominantly correspond to jets from light and $c$ quarks, with the fraction of jets from $b$ quarks increasing to above 90% for vertices with five or more tracks.

Several observables of secondary vertex kinematic properties are investigated for their sensitivity to modeling of $b$ quark fragmentation. Of those, the highest sensitivity is achieved when studying the ratio of SV momentum—i.e. the transverse component of the vectorial sum of all charged particle momenta used in the reconstruction of the vertex—to the total transverse momentum of the jet carried by charged particles.
\[ F_{ch} = \frac{p_T(SV)}{\sum_{ch} p_T}. \]

Effects arising from mismodeling of the overall kinematic properties of the event are canceled, to first approximation, by studying the ratio of the two momenta, in which the secondary vertex serves as a proxy for the b hadron and the charged particles represent the full momentum of the initial b quark. Note that this observable is not sensitive to variations in the jet energy scale, as it makes use only of the charged constituents of the selected jets. The observed and predicted distributions for \( F_{ch} \) in Z + jets events are shown in Fig. 2 (top), separately for vertices with three, four, and five tracks. For each plot the average of the distribution in the data is compared to the MC prediction using different b fragmentation tunes. The data appear to favor softer fragmentation shapes such as the Z2* and Peterson tunes. However, in this selection a significant fraction of the selected jets stems from the hadronization of light and charm quarks which are not changed by the event reweighting procedure used to compare the different tunes. Likewise, the Z2* LEP \( r_b \) tune only affects the simulated fragmentation of b quarks and was obtained using data from LEP enriched in jets from b quark hadronizations, and hence is not expected to correctly describe charm and light quark fragmentation.

In the sample of \( t\bar{t} \) events, selected as described in Sec. II C, and used later for the top quark mass extraction, the selected jets are expected to contain a significantly larger fraction of b quarks. From simulation, we expect a negligible dependence of \( F_{ch} \) on the kinematic properties and mass of the top quarks, making this distribution appropriate to compare different fragmentation models. The equivalent distributions of secondary vertex properties in \( t\bar{t} \) events are shown in Fig. 2 (bottom).

The observed distributions in this signal selection are generally well described by the central (Z2* LEP \( r_b \)) tune, but the comparison of the mean values of \( F_{ch} \)—as shown in the top panels of the plots—reveals differences between the various fragmentation shapes. Unlike in the Z + jets data, the Z2* tune shows the largest deviation with respect to the \( t\bar{t} \) data among the studied variations, whereas the Z2* LEP \( r_b \) fragmentation shape is in better agreement. Furthermore, the hard and soft variations of Z2* LEP \( r_b \) corresponding to

![Images of distributions with various tracks and jets](attachment:image.png)

**FIG. 2.** Distributions of the ratio of the transverse momentum of secondary vertices to the charged component of the jet with three, four, and five tracks (from left to right) in Z + jets dilepton (top) and \( t\bar{t} \) dilepton events (bottom), compared to the expected shape using the Z2* LEP \( r_b \) fragmentation tune. In each plot, the top panels compare the average of the distribution measured in data and its statistical uncertainty (shaded area) with that expected from different choices of the b quark fragmentation function in PYTHIA. For Z2* LEP \( r_b \), the error bar represents the \( \pm \) variations of Z2* LEP \( r_b \).
one standard deviation variations of the \( r_b \) parameter, provide a bracketing that encloses or approaches the data. The Z* LEP \( r_b \) tune is therefore used as the nominal \( b \) quark fragmentation shape in the following analysis, with the shape variations used to estimate systematic uncertainties in the top quark mass measurement.

**B. Inclusive charm mesons in \( \ell\ell \) events**

Kinematic properties of inclusively reconstructed charmed mesons inside \( b \) jets from top quark decays are expected to be sensitive to the modeling of \( b \) quark fragmentation. We limit the study to meson decays with large branching fractions and high expected signal-to-background ratios: \( J/\psi \rightarrow \mu^+\mu^- \), \( D^0 \rightarrow K^-\pi^+ \) in semileptonic \( B \) decays, and inclusive \( D^*(2010)^+ \rightarrow D^0\pi^+ \), with \( D^0 \rightarrow K^-\pi^+ \).

Top quark pair signal events are selected as described above; with the requirement of at least one \( b \)-tagged jet replacing that of the presence of a reconstructed secondary vertex. In the dilepton channels the \( b \) tagging algorithm output threshold is relaxed, as the expected background is lower. All five leptonic decay channels of the \( \ell\ell \) state are considered, as discussed above. To gather as much data as possible, both \( b \) jets in each event are considered, selected by their tagging discriminant value and their transverse momentum. All charged PF candidates used in the jet clustering are used to reconstruct mesons, with particle identification restricted to distinguishing electrons and muons from charged hadrons.

Candidates for \( J/\psi \) mesons are reconstructed by requiring two opposite-sign muon candidates among the charged jet constituents, and fitting their invariant mass in the range of 2.5–3.4 GeV, as shown in Fig. 3. The distribution is modeled with the sum of two Gaussian functions for the \( J/\psi \) signal and a falling exponential for the combinatorial backgrounds.

Neutral charm mesons, \( D^0 \), are produced in the majority of \( b \) hadron decays and are reconstructed via their decay to a \( K^- \) and \( \pi^+ \). To reduce combinatorial backgrounds they are selected together with a soft lepton from a semileptonic \( b \) hadron decay, whose charge determines the respective flavor of the two hadron tracks. All opposite-sign permutations of the three leading charged constituents of the jet are considered for \( K \) and \( \pi \) candidates, and no additional vertex reconstruction is attempted. The \( K\pi \) invariant mass is then fitted between 1.7 and 2.0 GeV, using a crystal ball [57] shape for the signal and an exponential for the combinatorial backgrounds, as shown in Fig. 3.

A large fraction of \( D^0 \) mesons is produced in the decays of intermediate excited charmed hadron states, such as the \( D^*(2010)^+ \), which can be reconstructed by considering the difference in invariant mass between the three-track (\( K\pi\pi \)) and the two-track (\( K\pi \)) systems, where a soft pion is emitted in the \( D^*(2010)^+ \rightarrow D^0\pi^+ \) decay. The \( D^0 \) mesons are reconstructed among the three leading tracks as described in the previous paragraph, and selected in a mass window of 50 MeV around the nominal \( D^0 \) mass. A third track of the same charge as the \( \pi \) candidate from the \( D^0 \) decay is then added, and the mass difference is fitted in a range of 140–170 MeV, as shown in Fig. 3. The shape of the mass difference showing the \( D^*(2010)^+ \) resonance is modeled using a sum of two Gaussian functions for the signal and a threshold function for the combinatorial backgrounds.

The position of the fitted invariant mass peaks—reconstructed purely in the silicon tracker—agree with the expected meson rest masses within about 0.05% for the \( D^0 \) and \( D^*(2010)^+ \), indicating that the pion and kaon momentum scales are very well described. The observed \( J/\psi \) meson mass, reconstructed using muons, agrees with the expectation [58] within about 0.3%, well within the muon momentum scale uncertainty.

The fitted signal and background distributions are then used to extract the kinematic properties of the reconstructed mesons using the \( \mathcal{P} \) lob technique [59], where a discriminating observable (in this case the invariant mass of the candidates) is used to separate the signal and background.

![FIG. 3. Fits to the invariant mass peaks of the three considered charmed mesons in \( \ell\ell \) events in the data, as described in the text: \( J/\psi \) (left), \( D^0 \) (middle), and \( D^*(2010)^+ \) (right).](image-url)
contributions to the distribution of an observable of interest. The same method is applied to simulated events with different generator tunes and a range of different $b$ quark fragmentation functions, and the results are compared with data. Among several investigated kinematic properties of the charm meson candidates, the fraction of transverse momentum relative to the charged component of the jet momentum shows the highest sensitivity to variations in the $b$ quark fragmentation shape. The results are displayed in Fig. 4.

The reconstructed mesons are observed to carry about 50%--60% of the overall charged jet momentum. These results are in good agreement with the predictions obtained from simulated $t\bar{t}$ events for the central fragmentation function choice and corresponding variations. The conclusions from the study of secondary vertex properties in the previous section are confirmed by the charm meson candidates, the fraction of transverse momentum showing better agreement with the data than the nominal $Z^*$ shape, albeit with a large statistical uncertainty.

The numbers of meson candidates observed in the data are reproduced within about 10% when PYTHIA with the $Z^*$ tune is used in the parton shower and hadronization, whereas HERWIG 6 [60] with the AUET2 tune [61] underestimates both the $D^*(2010)^+$ and $J/\psi$ yields by more than 50%, and overestimates $D^0$ production by about 30%.

IV. TOP QUARK MASS MEASUREMENT

Observables that are dependent on the top quark mass are constructed using the kinematic properties of the decay products of the top quark. The choice of observable is a compromise between sensitivity to the mass on the one hand and susceptibility to systematic uncertainties on the other hand. The most precise measurements to date have approached this trade-off by fully reconstructing the top quark from three jets in hadronic decays, heavily relying on precise calibrations of the reconstructed jet energies. In the analysis presented here, a different approach is used that sacrifices some sensitivity to minimize the reliance on detector calibrations. This exposes the result to uncertainties in the modeling of top quark decays and $b$ hadronization, but has reduced experimental uncertainties. The analysis will therefore immediately benefit from a future improvement of our understanding of these effects.

A. Observable and measurement strategy

The observable exploited in this analysis is built from the measured properties of the charged lepton from the $W$ boson decay and the charged constituents of a hadronic jet compatible with originating from a common secondary vertex. The invariant mass of the secondary vertex-lepton system, $m_{\nu_\ell}$, then serves as a proxy for the top quark mass. In building the invariant mass, the vertex constituents are assumed to be charged pions. The $m_{\nu_\ell}$ variable shows a strong dependence on the mass of the top quark despite not accounting for the neutrino from the $W$ boson decay or from semileptonic $b$ hadron decays, nor for neutral products of the $b$ quark hadronization. Using only charged particles and well-modeled leptons reduces the main experimental uncertainties to acceptance effects.

For each selected event, all possible combinations of leptons and secondary vertices—up to two in semileptonic events and up to four in dileptonic events—are taken into account in the measurement. Hence, by construction, the same number of correct and wrong combinations (i.e. pairing the lepton with the vertex associated with the other top quark decay) enter the analysis. In simulation, in about 11% of cases the selected vertex could not be attributed to the decay products of either $b$ quarks and is most likely spurious, either from a light quark from a hadronic $W$ boson decay, or from a gluon or light quark from initial-state radiation.
The observed $m_{svl}$ observable depends considerably on the number of tracks associated with the secondary vertex, shifting to higher values as more tracks are included. The analysis is therefore carried out in three exclusive track multiplicity categories of exactly three, four, or five tracks. Vertices with only two tracks show an increased level of backgrounds and reduced sensitivity to $m_t$ and are therefore excluded from the analysis. Furthermore, when evaluating systematic uncertainties, the results from the individual categories are assigned weights corresponding to the observed event yields in each, to absorb any mismodeling of the vertex multiplicity distribution in simulated events. Hence the analysis is carried out in fifteen mutually exclusive categories—three track multiplicities and five lepton flavor channels—and combined to yield the final result.

B. Signal and background modeling

The observed $m_{svl}$ distributions in each category are fitted with a combination of six individual components:

(a) “correct” pairings for the $t\bar{t}$ signal where leptons and vertices are matched to the same top quark decay;
(b) “wrong” pairings for the $t\bar{t}$ signal where leptons and vertices are matched to the opposite top quark decay products;
(c) “unmatched” pairings for the $t\bar{t}$ signal where leptons are paired with vertices that cannot be matched to a $b$ quark hadronization, i.e. either from a hadronic $W$ boson decay or from initial- or final-state radiation;
(d) “correct” pairings for the single top quark signal;
(e) “unmatched” pairings for the single top quark signal, where there can be no “wrong” pairs in the sense of the above;
(f) leptons and vertices from background processes.

Among those, the “correct” pairings both for $t\bar{t}$ and single top quarks, and the “wrong” pairings in the $t\bar{t}$ signal carry information about the top quark mass and are parametrized as a function of $m_t$. The relative fractions of correct, wrong, and unmatched pairings for both $t\bar{t}$ and single top quarks and their dependence on $m_t$ are determined from simulated events. Furthermore, the relative contributions of $t\bar{t}$ and single top quark events are calculated using the top quark mass-dependent theoretical predictions of the production cross sections at NNLO for $t\bar{t}$, and single top quark $t$ channel as well as $tW$ channel. The overall combined signal strength of $t\bar{t}$ and single top quark signal is left floating in the final fit, together with $m_t$.

The background contribution is a combination of different processes, depending on the channel, with dominant contributions from DY + jets in the dilepton channels, and $W + jets$ and QCD multijet processes in the semileptonic channels. The overall background yields are fixed to the predictions from simulation, with the exception of QCD multijets, the normalization of which is determined from a fit to the $E_T^{miss}$ distribution in the data, and DY + jets, which is normalized in a data control sample selecting dilepton pairs compatible with a $Z$ boson decay. The total (statistical plus systematic) uncertainty in the normalization of the QCD multijets and DY + jets backgrounds is about 30%.

For each channel and track multiplicity category, the full signal model is given by:

$$N[m_{svl}|m_t, \mu, \theta_{bkg}] = \mu N_{exp}^{top}[\alpha_{cor} f_{\bar{t}}^{cor}(m_{svl}|m_t) + \alpha_{wro} f_{\bar{t}}^{wro}(m_{svl}|m_t)] + (1 - \alpha_{cor} - \alpha_{wro}) f_{\bar{t}}^{unm}(m_{svl}) + \kappa_t [\alpha_{cor} f_{\bar{t}}^{cor}(m_{svl}|m_t) + (1 - \alpha_{cor}) f_{\bar{t}}^{noncor}(m_{svl})] + N_{exp}^{bkg}(1 + \theta_{bkg}) f_{bkg}(m_{svl}),$$

where $N_{exp}^{top}$ and $N_{exp}^{bkg}$ are the number of top quark events ($t\bar{t}$ and single top quarks) and background events expected from simulation; the $f_k$ are the six $m_{svl}$ templates of which three are parametrized in $m_t$; $\alpha_{cor}$, $\alpha_{wro}$, and $\alpha_{unm}$ are the fractions of correct and wrong lepton-vertex pairings for $t\bar{t}$ and single top quark production, determined from simulated events as a function of $m_t$; $\kappa_t$ is the relative fraction of single top quark events, fixed as a function of $m_t$ from the theoretical prediction; $\theta_{bkg}$ is a Gaussian penalty for a correction of the background yield; and finally $\mu$ is the overall signal strength of top quark events, determined in the fit.
The parameters of each of the $f^i_j$ templates and their possible $m_t$ dependence is determined in a fit to $m_{svl}$ distributions of simulated events in the corresponding category and pairing classification. The combined background template is built from fits to dedicated samples of simulated events of the corresponding processes, weighted by the expected event yields. The shape for QCD multijet processes is determined from a control sample of nonisolated leptons in the data and normalized using a fit to the $E_T^{miss}$ distribution. For correct and wrong pairings in $\bar{t}t$ and for correct pairings in single top quark events, the fit is done for a range of generated top quark mass points in the range $163.5$–$181.5$ GeV, from which a linear dependence of the parameters on $m_t$ is extracted. The $m_{svl}$ distributions for unmatched pairings and background events do not depend on $m_t$. Each distribution is fitted with the sum of an asymmetric Gaussian ($\Gamma_{asym}$) and a Gamma distribution ($\Gamma$), of which four of the six total parameters are found to provide sensitivity to the top quark mass:
FIG. 7. Template fits to the observed $m_{svl}$ distributions for the semileptonic channels ($\mu$ and $\mu$ from top to bottom row), and for exactly three, four, and five tracks assigned to the secondary vertex (from left to right column). The top panels show the bin-by-bin difference between the observed data and the fit result, divided by the statistical uncertainty (pull). The inset shows the scan of the negative log-likelihood as a function of the calibrated top quark mass, accounting only for the statistical uncertainty, when performed exclusively in each event category.

$$ f_i^m(m_{svl}|m_i) = \lambda G_{\text{asy}}(m_{svl}|\mu(m_i), \sigma_L(m_i), \sigma_R(m_i)) + (1 - \lambda) \Gamma(m_{svl}|\gamma, \beta, \nu(m_i)). $$

The shape parameters are the mean of the Gaussian peak ($\mu$), the left and right width parameters of the Gaussian ($\sigma_L$ and $\sigma_R$), the shape parameter of the Gamma distribution ($\gamma$), its scale ($\beta$), and its shift ($\nu$). Of these, all but $\gamma$ and $\beta$ show some usable sensitivity to the top quark mass.

The results of the fits to the observed $m_{svl}$ distributions in all fifteen categories are shown in Figs. 6 and 7 for the dilepton and semileptonic channels, respectively.

The final results for the top quark mass are then extracted by performing a binned maximum-likelihood estimation where the observed data are compared to the expectations using Poisson statistics. The combined likelihood is then written as:

$$ \mathcal{L}(m_r, \mu, \bar{\theta}_bkg) = \prod_{c=1}^{5} \prod_{n=3}^{5} \prod_{i=1}^{N_{\text{bins}}} \mathcal{P}[N_{\text{obs}}(m_{svl}^i), N_{\text{exp}}(m_r, \mu, \bar{\theta}_bkg, m_{svl}^i)] \mathcal{G}(0, \bar{\theta}_bkg, 0.3), $$

where the products of the Poisson-distributed yields ($\mathcal{P}$) over every channel ($c$), track multiplicity category ($n$), and $m_{svl}$ bin ($i$) are multiplied by a penalty Gaussian function for the correction of the expected background yields ($\mathcal{G}$), with a fixed width of 30%, corresponding to the uncertainty in the background normalization. Finally, the combined likelihood is maximized to obtain the final $m_r$ result. The analysis has been developed using simulated events, without performing the final fit on the data until the full measurement procedure had been validated.

The method is calibrated separately in each channel and track multiplicity bin before combining them by running pseudoeperiments for each generated top quark mass point and calculating a linear calibration function from the respective extracted mass points. Pseudodata are generated from the combined expected shape of the top quark signals and the mixture of backgrounds with the number of generated events taken from a Poisson distribution around the expected number of events in each category. The width of the pull distributions, i.e. the observed bias of each fit...
divided by its uncertainty, indicate a proper coverage of the statistical uncertainty. The postcalibration mass difference is below 100 MeV for the entire range of generated \(m_t\) values, well within the statistical uncertainty of the overall measurement of 200 MeV.

**C. Systematic uncertainties**

The size of the systematic uncertainties is evaluated from their impact on the \(m_{svl}\) shape and its propagation to the extracted \(m_t\) value in the combined fit. Modified pseudo-data are generated for each variation of the signal shape at their impact on the \(m_{svl}\) shape and its propagation to the extracted \(m_t\) value in the combined fit. Modified pseudo-data are generated for each variation of the signal shape at

### Table II. Summary of the systematic uncertainties in the final measurement.

<table>
<thead>
<tr>
<th>Source</th>
<th>(\Delta m_t, \text{[GeV]})</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Theoretical uncertainties</strong></td>
<td></td>
</tr>
<tr>
<td>(\mu_R/\mu_F) scales (t\bar{t})</td>
<td>+0.22</td>
</tr>
<tr>
<td>(\mu_R/\mu_F) scales (t) ((t) channel)</td>
<td>−0.04</td>
</tr>
<tr>
<td>(\mu_R/\mu_F) scales (tW)</td>
<td>+0.21</td>
</tr>
<tr>
<td>Parton shower matching scale</td>
<td>−0.04</td>
</tr>
<tr>
<td>Single top quark fraction</td>
<td>−0.07</td>
</tr>
<tr>
<td>Single top quark diagram interference (*)</td>
<td>+0.24</td>
</tr>
<tr>
<td>Parton distribution functions</td>
<td>+0.06</td>
</tr>
<tr>
<td>Top quark (p_T)</td>
<td>+0.82</td>
</tr>
<tr>
<td>Top quark decay width (*)</td>
<td>−0.05</td>
</tr>
<tr>
<td>b quark fragmentation</td>
<td>+1.00</td>
</tr>
<tr>
<td>Semileptonic B decays</td>
<td>−0.16</td>
</tr>
<tr>
<td>b hadron composition (*)</td>
<td>−0.09</td>
</tr>
<tr>
<td>Underlying event</td>
<td>+0.07</td>
</tr>
<tr>
<td>Color reconnection (*)</td>
<td>+0.08</td>
</tr>
<tr>
<td>Matrix element to parton shower matching scale</td>
<td>−0.42</td>
</tr>
<tr>
<td>(\sigma(t\bar{t} + \text{heavy flavor}))</td>
<td>+0.46</td>
</tr>
<tr>
<td><strong>Total theoretical uncertainty</strong></td>
<td>+1.52</td>
</tr>
<tr>
<td><strong>Experimental uncertainties</strong></td>
<td></td>
</tr>
<tr>
<td>Jet energy scale</td>
<td>+0.19</td>
</tr>
<tr>
<td>Jet energy resolution</td>
<td>−0.05</td>
</tr>
<tr>
<td>Unclustered energy</td>
<td>+0.07</td>
</tr>
<tr>
<td>Lepton energy scale</td>
<td>−0.26</td>
</tr>
<tr>
<td>Lepton selection efficiency</td>
<td>+0.01</td>
</tr>
<tr>
<td>b tagging</td>
<td>−0.02</td>
</tr>
<tr>
<td>Pileup</td>
<td>−0.05</td>
</tr>
<tr>
<td>Secondary-vertex track multiplicity (*)</td>
<td>−0.06</td>
</tr>
<tr>
<td>Secondary-vertex mass modeling (*)</td>
<td>−0.29</td>
</tr>
<tr>
<td>Background normalization</td>
<td>&lt; 0.03</td>
</tr>
<tr>
<td><strong>Total experimental uncertainty</strong></td>
<td>+0.43</td>
</tr>
<tr>
<td><strong>Total systematic uncertainty</strong></td>
<td>+1.58</td>
</tr>
<tr>
<td><strong>Statistical uncertainty</strong></td>
<td>±0.20</td>
</tr>
</tbody>
</table>

#### 1. Modeling and theoretical uncertainties

(i) **Choice of renormalization and factorization scales:**

The factorization and renormalization scales used in the signal simulation are set to a common value, \(Q\), defined by \(Q^2 = m_t^2 + \sum (p_{\text{parton}}^2)\), where the sum runs over all extra partons in the event. Two alternative data sets with a variation \(\mu_R = \mu_F = 2Q\) or \(Q/2\) are used to estimate the systematic effect from the choice of scales. These variations are observed to provide a conservative envelope of the additional jet multiplicity observed in data [62]. The scale choice for single top quark \(t\) and \(tW\) channels has a smaller effect on the measurement because the production happens through an electroweak interaction and because single top quark events only make up about 5% of the total yield. Dedicated single top quark data samples with \(\mu_F\) and \(\mu_R\) varied by a factor 2 or 1/2 are generated and used to estimate the effect.

(ii) **Matrix element to parton shower matching scale:**

The choice of the threshold in the event generation at which additional radiation is produced by the \(\text{PYTHIA}\) showering instead of matrix element calculations in \(\text{MADGRAPH}\) is expected to have a small impact on the shape of \(m_{svl}\), affecting mostly the “unmatched” lepton-SV pairings, which constitute only about 5% of the total. Variations of this threshold are furthermore observed to have small impact on the kinematic properties of extra jets [62]. The effect is estimated using dedicated samples with the nominal threshold (20 GeV) varied up and down by a factor of 2.

(iii) **Single top quark fraction:**

The overall signal shapes in each category are constructed from \(t\bar{t}\) events and events from single top quark production, with their relative fractions fixed to the expectation from theory. Because of a relative difference in their respective shapes, a deviation in this fraction can have an impact on the final mass measurement. The effect is estimated by repeating the fits with the relative fraction of single top quark events in the signal shape varied by \(\pm 20\%\). The size of the variation reflects the experimental uncertainty in the overall cross section of single top quark production [63,64].

(iv) **Single top quark interference:**

Interference between \(t\bar{t}\) pair production and single top quark production in the \(tW\) channel at next-to-leading order in QCD is resolved in the \(tW\) signal generation by removing all doubly
resonant diagrams in the calculation [65–67]. A different scheme for the resolution of the diagram interference can be defined where a gauge-invariant subtraction term modifies the $tW$ cross section to cancel the contributions from $t\bar{t}$. Samples using the second scheme are generated and compared and the difference is quoted as a systematic uncertainty [65,68].

(v) **Parton distribution functions:** Uncertainties from the modeling of parton momentum distributions inside the incoming protons are evaluated using the diagonalized uncertainty sources of the CT10 PDF set [21]. Each source is used to derive event-by-event weights, which are then applied to obtain a variation of the signal $m_{svl}$ shape. The maximal difference with respect to the nominal signal sample is quoted as the systematic uncertainty.

(vi) **Top quark $p_T$ modeling:** Measurements of the differential $t\bar{t}$ production cross section reveal an observed top quark $p_T$ spectrum that is softer than what is predicted from simulation [69]. The difference between the unfolded data and the simulation based on MadGraph is parametrized and can be used to calculate event-by-event weights correcting the spectrum. This reweighting is not applied when calibrating the measurement, as it introduces a dependence on the true top quark mass. The impact of the full difference between the predicted spectrum used in the calibration (at $m_t = 172.5$ GeV) and the data-corrected spectrum is estimated by comparing the result from reweighted pseudodata to the nominal value. The difference is then added as a one-sided systematic uncertainty in the extracted mass value. The effect of the reweighting on the simulated $m_{svl}$ shape for correct and wrong lepton-vertex pairings is shown in Fig. 8.

(vii) **Top quark decay width:** The decay width of the top quark has been experimentally determined with a precision of about 10% [70]. A dedicated sample with an increased width is used to estimate the impact on the mass measurement, and the difference is quoted as an uncertainty.

(viii) **$b$ quark fragmentation:** A variation in the momentum transfer from $b$ quark to $b$ hadron has a direct impact in the $m_{svl}$ distribution, and correspondingly, the uncertainty from the used $b$ quark fragmentation function on the extracted top quark mass is expected to be significant. As shown in Sec. III, the average momentum transfer in the nominal PYTHIA $Z^+$ tune is found to be significantly softer than that seen in $t\bar{t}$ events in the data, whereas the $Z^+$ LEP $r_b$ variation that follows a fragmentation function measured at LEP is in better agreement. Its soft and hard variations provide one standard deviation variations of the shape parameters, and are used to estimate the systematic uncertainty. Variations of the $m_{svl}$ shape for the central $Z^+$ LEP $r_b$ fragmentation function, its soft and hard variations, as well as the nominal $Z^+$ fragmentation are shown in Fig. 8. The impact of the choice of $b$ quark fragmentation function on the extracted top quark mass is shown in Fig. 9. To first order the measured $m_t$ value depends only on the average momentum transfer, as indicated by the linear dependence on $\langle p_T(B)/p_T(b) \rangle$. The extracted mass changes by about 0.61 GeV for each percent of change in the average momentum transfer.

FIG. 8. Variation of the simulated $m_{svl}$ shape with systematic effects: reweighting of the simulated top quark $p_T$ shape to the observed distribution, separately for correct and wrong lepton-vertex pairings (top); and different $b$ quark fragmentation function shapes (bottom). The bottom panels in the two plots show the ratios between the top quark $p_T$ reweighted and nominal shapes for the correct and wrong pairings (top), and between the various fragmentation models and the central $Z^+$ LEP $r_b$ tune (bottom).
FIG. 9. Impact of the average \( b \) quark fragmentation, \( \langle p_T(B)/p_T(b) \rangle \), on the extracted \( m_t \) value, for various different fragmentation shapes. The horizontal band represents the contribution of the \( b \) quark fragmentation model to the systematic uncertainty in the measurement of the top quark mass, which is estimated from the Z2* LEP \( r_b \) variations. A linear fit to the effects on the different variations (the line in the figure) suggests a relative change in the measured top quark mass of 0.61 GeV for each percent change in average momentum transfer.

(ix) Semileptonic \( B \) meson branching fractions: The effect of the uncertainties in semileptonic \( b \) hadron branching fractions is estimated by varying the fraction of \( b \) jets containing neutrinos down by 0.45% and up by 0.77%, covering the uncertainties in the experimentally measured semileptonic branching fractions of \( B^0 \) and \( B^\pm \) mesons [58].

(x) \( b \) hadron composition: The PYTHIA Z2* tune produces an average composition of about 40% \( B^0 \), 40% \( B^0 \), 12% \( B^\pm \), and 8% heavier \( b \) hadron states in the hadronization of \( b \) quarks. An improved version of this tune that takes into account hadron multiplicity measurements [58] is used to estimate the uncertainty due to the composition of \( b \) hadrons in the \( b \) jets.

(xi) Hadronization model cross-check: To test for additional uncertainties arising from the usage of the Lund string hadronization model in PYTHIA [54] in the default simulation, additional cross-checks are performed with alternative hadronization models as used in HERWIG. However, an inclusive comparison of the two parton shower and hadronization frameworks entangles various different effects in an incoherent and nontransparent manner and includes uncertainties that are already evaluated in dedicated studies in more sound ways. The inclusive PYTHIA-HERWIG difference is therefore not included as a systematic uncertainty. Evaluating whether there are indeed additional sources of uncertainty arising when comparing different hadronization models requires a comparison without changing the parton shower model, the hard-scattering simulation, or the \( b \) quark fragmentation functions. Such a check is possible in the SHERPA 2.1.0 framework [71], which permits a \( p_T \)-ordered parton shower model to be used, interfaced with a cluster hadronization model as used in HERWIG or with the Lund string model of PYTHIA. The change in hadronization model entails a difference in hadron flavor multiplicities, with the cluster model tending to yield more heavy \( B_c \) mesons and \( \Lambda_b \) baryons. Restricting the study to the dominant production of \( B^0 \) and \( B^\pm \) mesons reveals a different \( b \) quark fragmentation function shape between the two models. As the uncertainty from this effect is already covered by a more extreme variation in the dedicated \( b \) quark fragmentation uncertainty, the distributions are reweighted to a common \( b \) parton to \( b \) hadron momentum transfer distribution to remove any difference in fragmentation shapes. The resulting lepton + \( b \) jet invariant mass distributions for cluster and Lund string fragmentation are found to be in very good agreement and do not warrant any additional uncertainty in the top quark mass measurement.

(xii) Underlying event and color reconnection: Effects from the modeling of the proton collision remnants and multiparton interactions (the underlying event) and from the color connection of the \( b \) quark fragmentation products to the rest of the event (color reconnection) are estimated using dedicated samples with variations of the Perugia 11 (P11) underlying event tunes [72]. Two variations, one with altered multiparton interactions and one based on the Tevatron data are used to evaluate the effect of the underlying event modeling. A separate sample, in which color reconnection effects are not simulated, is used to gauge the impact from the modeling of this effect. In both cases, the full difference of the results obtained on the modified samples and the case of using pseudodata from the central P11 tune are quoted as the systematic uncertainty.

(xiii) Matrix element generator: The default Born-level matrix element generator, MADGRAPH, is substituted by a POWHEG simulation based on the heavy-quark pair production (hvq) model [73] at NLO accuracy for \( t\bar{t} \) production and at leading order for the top quark decays. In both cases, the matrix element generators are interfaced with PYTHIA for parton showering. The difference, propagated to the mass measurement, is reported as a systematic uncertainty.

Furthermore, the effect of including NLO corrections in the modeling of the top quark decay is studied.
using the parton-level MCFM program [35,74]. Since no fragmentation or parton shower evolution is included in the simulation and therefore the actual impact on the mass measurement is uncertain, the result is only reported here but not included as a systematic uncertainty. By reweighting the mass of the lepton-\( b \)-jet system generated by MADGRAPH to the differential cross sections predicted by MCFM, with and without applying NLO corrections to the top quark decay, a + 1.29 GeV shift in the calibrated mass in the \( ee \) channel is observed.

(xiv) Modeling of the associated production of \( t\bar{t} \) with heavy flavors: While the simulation is observed to describe the shape of the different distributions for \( t\bar{t} \) + heavy flavors well (most notably \( t\bar{t} \) + \( b\bar{b} \)), these predictions tend to underestimate the total cross section [62,75]. To evaluate the impact on the measurement, the nominal simulation is compared to the one obtained after reweighting the contribution from extra \( b \) jets in the simulation by the data-to-theory scale factor measured in [62]. A symmetric variation of the expected extra heavy-flavor content is used to estimate this uncertainty.

2. Experimental uncertainties

(i) Jet energy scale and jet energy resolution: By design, the reconstructed jet energy does not affect the \( m_{svl} \) observable. However jet momenta are used in the event selection and therefore variations of the jet energy have an effect on the event yields that enter the bins of the \( m_{svl} \) distributions. The effects are estimated by rescaling the reconstructed jet energies depending on \( p_T \) and \( \eta \) before performing the event selection. The effect of jet energy resolution on the measured distributions is estimated by inflating or deflating the resolution within the measured uncertainties and propagating the effects to the final distributions. The varied \( m_{svl} \) distributions are used to generate pseudodata, and the full differences to the nominal sample are quoted as the systematic uncertainties.

(ii) Unclustered energy: The missing transverse energy is used only in the event selection for the \( ee \) and \( \mu\mu \) channels to suppress events containing neutrinoless \( Z \) boson decays. Since the DY yield is normalized from a dedicated data control region, the effect from the \( E_T^{\text{miss}} \) resolution is expected to be small. It is estimated by varying the amount of energy that is not clustered into jets in the \( E_T^{\text{miss}} \) calculation by ±10% and studying its impact on the observed \( m_{svl} \) distributions.

(iii) Lepton momentum scale: The reconstructed lepton momenta directly affect the \( m_{svl} \) spectrum. The uncertainty in the measured energy scale for electrons depends on \( p_T \) and \( \eta \) and varies between about 0.6% in the central barrel region and about 1.5% in the forward region [39]. The muon momentum scale is known within an uncertainty of about 0.2% [40]. Varying the scales up and down within their measured uncertainties—as a function of \( p_T \) and \( \eta \) for electrons—produces a shift in the \( m_{svl} \) distribution that is propagated to the final mass measurement and quoted as a systematic uncertainty.

(iv) Lepton selection efficiency: Similar to the jet energy scales, the requirements applied when selecting lepton candidates for the analysis affect the event yields in the \( m_{svl} \) distributions and can cause a slight change in the extracted top quark mass. The measured electron and muon selection efficiencies are varied within their uncertainties and the difference is quoted as a systematic uncertainty.

(v) \( b \) tagging efficiency and mistag rate: The \( t\bar{t} \) event selection relies on the use of a \( b \) tagging algorithm to select jets originating from the hadronization of a \( b \) quark. The impact on \( m_{svl} \) from the uncertainties in the signal and background efficiencies of this algorithm are estimated by varying the efficiencies within their measured uncertainties and propagating the effect to the final result.

(vi) Pileup: The effect of additional concurrent \( pp \) interactions on the measured precision is estimated by varying the cross section for inelastic \( pp \) collisions used in the pileup generation by ±5%, and propagating the difference to the extracted \( m_t \) result.

(vii) Secondary-vertex track multiplicity: The distribution of the number of tracks assigned to secondary vertices is not well described by simulation, as has been observed in several processes involving \( b \) quarks. Generally, the data shows about 5%–10% fewer tracks than the simulation. As the analysis is carried out in exclusive bins of track multiplicity to minimize the impact of this issue, it only enters as a second-order effect when combining the results from different bins, as the individual bins would be assigned slightly different weights in simulation. This is corrected for by reweighting each bin content by the yield observed in the data, and the impact of this reweighting on the final result is quoted as a remaining systematic uncertainty.

(viii) Secondary-vertex mass modeling: A discrepancy between the observed secondary vertex mass (i.e. the invariant mass of the tracks used to reconstruct the vertex) and the one predicted in the simulation is observed. The effect is propagated in the \( m_{svl} \) shape by weighting the simulated events to reflect the observed distributions in each bin of track multiplicity, and the resulting shift in the extracted top quark mass is quoted as a systematic uncertainty.

(ix) Background normalization: Processes not involving top quarks constitute about 5% of the overall selected events and their combined yield is allowed to float within about 30% in the fit. The
normalization of the main background processes is furthermore determined in dedicated control samples in the data. To estimate the uncertainty in the result stemming from the uncertainty in the background normalization, the expected yields of backgrounds are varied within their uncertainties, and the resulting change in the $m_{\text{svl}}$ shape is propagated to the final result. These variations are observed to have a negligible impact on the measurement as they are absorbed by upward/downward variations of the background yields in the fit.

D. Results

The top quark mass is measured from the invariant mass distribution of leptons and reconstructed secondary vertices from $b$ hadron decays using only charged particles. After calibrating the measurement with simulated events, a value of

$$m_t = 173.68 \pm 0.20\,(\text{stat})^{+1.59}_{-0.99}\,(\text{syst}) \text{ GeV}$$

is obtained from the data, with a combined uncertainty of $+1.59$ GeV. The overall systematic uncertainty is dominated by the $b$ quark fragmentation modeling uncertainty of $+1.00/−0.54$ GeV and the uncertainty in the modeling of the top quark $p_T$ of $+0.82$ GeV. Experimental uncertainties related to the understanding of jet energy scales only affect the event acceptance and are therefore virtually irrelevant to the final result. Studies of the $b$ quark fragmentation with reconstructed secondary vertices and inclusively reconstructed charm quark mesons are used to select the central $b$ quark fragmentation shape and to validate the systematic uncertainty.

With the significantly larger data sets becoming available for analysis from the current 13 TeV run of the LHC, this method could be extended to constrain the $b$ quark fragmentation, using the properties of the secondary vertices or charmed mesons, while measuring the top quark mass. This is expected to lead to a significant reduction of the overall uncertainty. Furthermore, theoretical uncertainties related to kinematic properties of top quarks and scale choices in QCD calculations are expected to decrease with the next generation of Monte Carlo event generators.

Finally, this result is complementary to standard measurements relying on kinematic properties of jets. The precision of such analyses is typically limited by the uncertainty from the modeling of hadronization effects, influencing the understanding of the jet energy scale, while not much affected by the choice of $b$ quark fragmentation model and the modeling of top quark kinematic properties. Therefore, a combination of this result with standard measurements could optimally benefit from independent sources of systematic uncertainties.

V. SUMMARY AND PROSPECTS

A novel measurement of the top quark mass has been presented, using an observable that relies entirely on the reconstruction of charged particles. It shows minimal sensitivity to experimental sources of uncertainty. The final result yields a value of $m_t = 173.68 \pm 1.59$ GeV, equivalent to a precision of well below one percent. The overall uncertainty is dominated by the $b$ quark fragmentation modeling uncertainty of $+1.00/−0.54$ GeV and the uncertainty in the modeling of the top quark $p_T$ of $+0.82$ GeV. Experimental uncertainties related to the understanding of jet energy scales only affect the event acceptance and are therefore virtually irrelevant to the final result. Studies of the $b$ quark fragmentation with reconstructed secondary vertices and inclusively reconstructed charm quark mesons are used to select the central $b$ quark fragmentation shape and to validate the systematic uncertainty.

We congratulate 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 centers and personnel of the Worldwide LHC Computing Grid 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 and the CMS detector provided by the following funding agencies: the Austrian Federal Ministry of Science, Research and Economy and the Austrian Science Fund; the Belgian Fonds de la Recherche Scientifique, and Fonds voor Wetenschappelijk Onderzoek; the Brazilian Funding Agencies (CNPq, CAPES, FAPERJ, and FAPESP); the Bulgarian Ministry of Education and Science; CERN; the Chinese Academy of Sciences, Ministry of Science and Technology, and National Natural Science Foundation of China; the Colombian Funding Agency (COLCIENCIAS); the Croatian Ministry of Science, Education and Sport, and the Croatian Science Foundation; the Research Promotion Foundation, Cyprus; the Ministry of Education and Research, Estonian Research Council via IUT23-4 and IUT23-6 and European Regional Development Fund, Estonia; the Academy of Finland, Finnish Ministry of Education and Culture, and Helsinki Institute of Physics; the Institut National de Physique Nucléaire et de Physique des Particules / CNRS, and Commissariat à l’Énergie Atomique et aux Énergies Alternatives / CEA, France; the Bundesministerium für Bildung und Forschung, Deutsche Forschungsgemeinschaft, and Helmholtz-Gemeinschaft Deutscher Forschungszentren, Germany; the General Secretariat for Research and Technology, Greece; the National Scientific Research Foundation, and National Innovation Office, Hungary; the Department of Atomic Energy and the Department of Science and Technology, India; the Institute for Studies in Theoretical Physics and Mathematics, Iran; the Science Foundation, Ireland; the Istituto Nazionale di Fisica Nucleare, Italy; the Ministry of Science, ICT and Future Planning, and National Research Foundation (NRF), Republic of Korea; the Lithuanian Academy of Sciences; the Ministry of Education, and University of Malaya (Malaysia); the Mexican Funding Agencies (CINVESTAV, CONACYT, SEP, and UASLP-FAI); the Ministry of Business, Innovation and Employment, New Zealand; the Pakistan Atomic Energy Commission; the Ministry of Science and Higher Education and the National Science Center, Poland; the Fundação para a Ciência e a Tecnologia, Portugal; JINR, Dubna; the Ministry of Education and Science of the Russian Federation, the Federal Agency of Atomic Energy of the Russian Federation, Russian Academy of Sciences, and the Russian Foundation for Basic Research; the Ministry of Education, Science and Technological Development of Serbia; the Secretaría de Estado de Investigación, Desarrollo e Innovación and Programa Consolider-Ingenio 2010, Spain; the Swiss Funding Agencies (ETH Board, ETH Zurich, PSI, SNF, UniZH, Canton Zurich, and SER); the Ministry of Science and Technology, Taipei; the Thailand Center of Excellence in Physics, the Institute for the Promotion of Teaching Science and Technology of Thailand, Special Task Force for Activating Research and the National Science and Technology Development Agency of Thailand; the Scientific and Technical Research Council of Turkey, and Turkish Atomic Energy Authority; the National Academy of Sciences of Ukraine, and State Fund for Fundamental Researches, Ukraine; the Science and Technology Facilities Council, UK; the U.S. Department of Energy, and the U.S. National Science Foundation. Individuals have received support from the Marie-Curie program and the European Research Council and EPLANET (European Union); the Leventis Foundation; the A.P. Sloan Foundation; the Alexander von Humboldt Foundation; the Belgian Federal Science Policy Office; the Fonds pour la Formation à la Recherche dans l’Industrie et dans l’Agriculture (FRIA-Belgium); the Agentschap voor Innovatie door Wetenschap en Technologie (IWT-Belgium); the Ministry of Education, Youth and Sports (MEYS) of the Czech Republic; the Council of Science and Industrial Research, India; the HOMING PLUS program of the Foundation for Polish Science, cofinanced from European Union, Regional Development Fund; the OPUS program of the National Science Center (Poland); the Compagnia di San Paolo (Torino); MIUR Project No. 20108T4XTM (Italy); the Thalis and Aristeia programs cofinanced by EU-ESF and the Greek NSRF; the National Priorities Research Program by Qatar National Research Fund; the Rachadapisek Somphot Fund for Postdoctoral Fellowship, Chulalongkorn University (Thailand); and the Welch Foundation, Contract No. C-1845.

Measurement of the Top Quark Mass Using ... 


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


[10] A. Abulencia et al. (CDF Collaboration), Measurement of the top quark mass in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV using the decay length technique, Phys. Rev. D 75, 071102 (2007).


[14] CMS Collaboration, Measurement of the $t\bar{t}$ production cross section in the $e\mu$ channel in proton-proton collisions at $\sqrt{s} = 7$ and 8 TeV, arXiv:1603.02303.


[18] CMS Collaboration, Measurements of the $t\bar{t}$ production cross section in lepton + jets final states in $pp$ collisions at 8 TeV and ratio of 8 to 7 TeV cross sections, arXiv:1602.09024.


[22] CMS Collaboration, Measurement of the underlying event activity at the LHC with $\sqrt{s} = 7$ TeV and comparison with $\sqrt{s} = 0.9$ TeV, J. High Energy Phys. 09 (2011) 109.

[51] T. Affolder et al. (CDF), Measurement of b Quark Fragmentation Fractions in pp Collisions at \(\sqrt{s} = 1.8\) TeV, Phys. Rev. Lett. 84, 1663 (2000).


[62] CMS Collaboration, Measurement of \(\bar{t}t\) production with additional jet activity, including b quark jets, in the dilepton channel using \(pp\) collisions at \(\sqrt{s} = 8\) TeV, arXiv:1510.03072.
[64] CMS Collaboration, Measurement of the t-channel single-top-quark production cross section and of the \(|V_{tb}|\) CKM matrix element in \(pp\) collisions at \(\sqrt{s} = 8\) TeV, J. High Energy Phys. 06 (2014), 090.
[70] CMS Collaboration, Measurement of the ratio \(B(t \to Wb)/B(t \to Wq)\) in \(pp\) collisions at \(\sqrt{s} = 8\) TeV, Phys. Lett. B 736, 33 (2014).
[75] ATLAS Collaboration, Measurements of fiducial cross-sections for \(\bar{t}t\) production with one or two additional b-jets in \(pp\) collisions at \(\sqrt{s} = 8\) TeV using the ATLAS detector, Eur. Phys. J. C 76, 11 (2016).
MEASUREMENT OF THE TOP QUARK MASS USING ...

PHYSICAL REVIEW D 93, 092006 (2016)

(CMS Collaboration)

1Yerevan Physics Institute, Yerevan, Armenia
2Institut für Hochenergiephysik der OeAW, Wien, Austria
3National Centre for Particle and High Energy Physics, Minsk, Belarus
4Universiteit Antwerpen, Antwerpen, Belgium
5Vrije Universiteit Brussel, Brussel, Belgium
6Université Libre de Bruxelles, Bruxelles, Belgium
7Ghent University, Ghent, Belgium
8Université Catholique de Louvain, Louvain-la-Neuve, Belgium
9Université de Mons, Mons, Belgium
10Centro Brasileiro de Pesquisas Fisicas, Rio de Janeiro, Brazil
11Universidade do Estado do Rio de Janeiro, Rio de Janeiro, Brazil
12Universidade Estadual Paulista, São Paulo, Brazil
13Institute for Nuclear Research and Nuclear Energy, Sofia, Bulgaria
14University of Sofia, Sofia, Bulgaria
15Beihang University, Beijing, China
16Institute of High Energy Physics, Beijing, China
17State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing, China
18Universidad de Los Andes, Bogota, Colombia
19University of Split, Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture, Split, Croatia
20University of Split, Faculty of Science, Split, Croatia
21Institute Rudjer Boskovic, Zagreb, Croatia
22University of Cyprus, Nicosia, Cyprus
23Charles University, Prague, Czech Republic
24Universidad San Francisco de Quito, Quito, Ecuador
25Academy of Scientific Research and Technology of the Arab Republic of Egypt, Egyptian Network of High Energy Physics, Cairo, Egypt
26National Institute of Chemical Physics and Biophysics, Tallinn, Estonia
27Department of Physics, University of Helsinki, Helsinki, Finland
28Helsinki Institute of Physics, Helsinki, Finland
29Lappeenranta University of Technology, Lappeenranta, Finland
30DSM/IRFU, CEA/Saclay, Gif-sur-Yvette, France
31Laboratoire Leprince-Ringuet, Ecole Polytechnique, IN2P3-CNRS, Palaiseau, France
32Institut Pluridisciplinaire Hubert Curien, Université de Strasbourg, Université de Haute Alsace Mulhouse, CNRS/IN2P3, Strasbourg, France
33Centre de Calcul de l’Institut National de Physique Nucleaire et de Physique des Particules, CNRS/IN2P3, Villeurbanne, France
34Université de Lyon, Université Claude Bernard Lyon 1, CNRS-IN2P3, Institut de Physique Nucléaire de Lyon, Villeurbanne, France
35Georgian Technical University, Tbilisi, Georgia
36Tbilisi State University, Tbilisi, Georgia
37RWTH Aachen University, I. Physikalisches Institut, Aachen, Germany
38RWTH Aachen University, III. Physikalisches Institut A, Aachen, Germany
39RWTH Aachen University, III. Physikalisches Institut B, Aachen, Germany
40Deutsches Elektronen-Synchrotron, Hamburg, Germany
41University of Hamburg, Hamburg, Germany
42Institut für Experimentelle Kernphysik, Karlsruhe, Germany
43Institute of Nuclear and Particle Physics (INPP), NCSR Demokritos, Aghia Paraskevi, Greece
44National and Kapodistrian University of Athens, Athens, Greece
45University of Ioánnina, Ioánnina, Greece
46MTA-ELTE Lendület CMS Particle and Nuclear Physics Group
47Wigner Research Centre for Physics, Budapest, Hungary
48Institute of Nuclear Research ATOMKI, Debrecen, Hungary
49University of Debrecen, Debrecen, Hungary
50National Institute of Science Education and Research, Bhubaneswar, India
51Panjab University, Chandigarh, India
52University of Delhi, Delhi, India

092006-25
Saha Institute of Nuclear Physics, Kolkata, India
Bhabha Atomic Research Centre, Mumbai, India
Tata Institute of Fundamental Research, Mumbai, India
Indian Institute of Science Education and Research (IISER), Pune, India
Institute for Research in Fundamental Sciences (IPM), Tehran, Iran
University College Dublin, Dublin, Ireland
INFN Sezione di Bari, Bari, Italy
Università di Bari, Bari, Italy
Politecnico di Bari, Bari, Italy
INFN Sezione di Bologna, Bologna, Italy
Università di Bologna, Bologna, Italy
INFN Sezione di Catania, Catania, Italy
Università di Catania, Catania, Italy
INFN Sezione di Firenze, Firenze, Italy
Università di Firenze, Firenze, Italy
INFN Laboratori Nazionali di Frascati, Frascati, Italy
INFN Sezione di Genova, Genova, Italy
Università di Genova, Genova, Italy
INFN Sezione di Milano-Bicocca, Milan, Italy
Università di Milano-Bicocca, Milan, Italy
INFN Sezione di Napoli, Napoli, Italy
Università di Napoli ‘Federico II’, Napoli, Italy
Università della Basilicata, Potenza, Italy, Università G. Marconi, Roma, Italy
Università G. Marconi, Potenza, Italy, Università G. Marconi, Roma, Italy
INFN Sezione di Padova, Padova, Italy
Università di Padova, Padova, Italy
Università di Trento, Trento, Italy
INFN Sezione di Pavia, Pavia, Italy
Università di Pavia, Pavia, Italy
INFN Sezione di Perugia, Perugia, Italy
Università di Perugia, Perugia, Italy
INFN Sezione di Pisa, Pisa, Italy
Università di Pisa, Pisa, Italy
Scuola Normale Superiore di Pisa, Pisa, Italy
INFN Sezione di Roma
Università di Roma
INFN Sezione di Torino, Torino, Italy
Università di Torino, Torino, Italy
Università del Piemonte Orientale, Novara, Italy
INFN Sezione di Trieste, Trieste, Italy
Università di Trieste, Trieste, Italy
Kangwon National University, Chunchon, Korea
Kyungpook National University, Daegu, Korea
Chonbuk National University, Jeonju, Korea
Chonnam National University, Institute for Universe and Elementary Particles, Kwangju, Korea
Korea University, Seoul, Korea
Seoul National University, Seoul, Korea
University of Seoul, Seoul, Korea
Sungkyunkwan University, Suwon, Korea
Vilnius University, Vilnius, Lithuania
National Centre for Particle Physics, Universiti Malaya, Kuala Lumpur, Malaysia
Centro de Investigacion y de Estudios Avanzados del IPN, Mexico City, Mexico
Universidad Iberoamericana, Mexico City, Mexico
Benemerita Universidad Autonoma de Puebla, Puebla, Mexico
Universidad Autónoma de San Luis Potosí, San Luis Potosí, Mexico
University of Auckland, Auckland, New Zealand
University of Canterbury, Christchurch, New Zealand
National Centre for Physics, Quaid-I-Azam University, Islamabad, Pakistan
National Centre for Nuclear Research, Swierk, Poland
Institute of Experimental Physics, Faculty of Physics, University of Warsaw, Warsaw, Poland
University of Nebraska-Lincoln, Lincoln, USA
State University of New York at Buffalo, Buffalo, USA
Northeastern University, Boston, USA
Northwestern University, Evanston, USA
University of Notre Dame, Notre Dame, USA
The Ohio State University, Columbus, USA
Princeton University, Princeton, USA
University of Puerto Rico, Mayaguez, USA
Purdue University, West Lafayette, USA
Purdue University Calumet, Hammond, USA
Rice University, Houston, USA
University of Rochester, Rochester, USA
Rutgers, The State University of New Jersey, Piscataway, USA
University of Tennessee, Knoxville, USA
Texas A&M University, College Station, USA
Texas Tech University, Lubbock, USA
Vanderbilt University, Nashville, USA
University of Virginia, Charlottesville, USA
Wayne State University, Detroit, USA
University of Wisconsin—Madison, Madison, WI, USA

Deceased.

1Also at Vienna University of Technology, Vienna, Austria.
2Also at State Key Laboratory of Nuclear Physics and Technology, Peking University, Beijing, China.
3Also at Institut Pluridisciplinaire Hubert Curien, Université de Strasbourg, Université de Haute Alsace Mulhouse, CNRS/IN2P3, Strasbourg, France.
4Also at Universidade Estadual de Campinas, Campinas, Brazil.
5Also at Centre National de la Recherche Scientifique (CNRS) - IN2P3, Paris, France.
6Also at Université Libre de Bruxelles, Bruxelles, Belgium.
7Also at Laboratoire Leprince-Ringuet, Ecole Polytechnique, IN2P3-CNRS, Palaiseau, France.
8Also at Joint Institute for Nuclear Research, Dubna, Russia.
9Also at British University in Egypt, Cairo, Egypt.
10Also at Zewail City of Science and Technology, Zewail, Egypt.
11Also at Ain Shams University, Cairo, Egypt.
12Also at Université de Haute Alsace, Mulhouse, France.
13Also at CERN, European Organization for Nuclear Research, Geneva, Switzerland.
14Also at Skobeltsyn Institute of Nuclear Physics, Lomonosov Moscow State University, Moscow, Russia.
15Also at Tbilisi State University, Tbilisi, Georgia.
16Also at RWTH Aachen University, III. Physikalisches Institut A, Aachen, Germany.
17Also at University of Hamburg, Hamburg, Germany.
18Also at Brandenburg University of Technology, Cottbus, Germany.
19Also at Institute of Nuclear Research ATOMKI, Debrecen, Hungary.
20Also at MTA-ELTE Lendület CMS Particle and Nuclear Physics Group.
21Also at University of Debrecen, Debrecen, Hungary.
22Also at Indian Institute of Science Education and Research, Bhopal, India.
23Also at University of Visva-Bharati, Santiniketan, India.
24Also at King Abdulaziz University, Jeddah, Saudi Arabia.
25Also at University of Ruhuna, Matara, Sri Lanka.
26Also at Isfahan University of Technology, Isfahan, Iran.
27Also at University of Tehran, Department of Engineering Science, Tehran, Iran.
28Also at Plasma Physics Research Center, Science and Research Branch, Islamic Azad University, Tehran, Iran.
29Also at Università degli Studi di Siena, Siena, Italy.
30Also at Purdue University, West Lafayette, USA.
31Also at Hanyang University, Seoul, Korea.
32Also at International Islamic University of Malaysia, Kuala Lumpur, Malaysia.
33Also at Malaysian Nuclear Agency, MOSTI, Kajang, Malaysia.
34Also at Consejo Nacional de Ciencia y Tecnología, Mexico city, Mexico.
35Also at Warsaw University of Technology, Institute of Electronic Systems, Warsaw, Poland.
36Also at Institute for Nuclear Research, Moscow, Russia.
37Also at National Research Nuclear University "Moscow Engineering Physics Institute" (MEPhI), Moscow, Russia.