CMS Physics Analysis Summary

Contact: cms-pag-conveners-top@cern.ch 2014/11/02

Determination of the top-quark mass from the lepton-b-jet invariant mass distribution in dileptonic $t\bar{t}$ events using proton-proton collision data at $\sqrt{s}$ = 8 TeV

The CMS Collaboration

Abstract

A measurement of the top-quark mass is presented using the distribution of the invariant mass m_{lb} of the lepton and the b-quark originating from top-quark decays. The analysed dataset of pp-collisions at a centre-of-mass energy $\sqrt{s} = 8$ TeV was recorded by the CMS experiment at the LHC, and corresponds to an integrated luminosity of 19.7 fb^{-1}. Top-quark pair candidate events are selected with two opposite-charge isolated leptons, one electron and one muon, and at least two jets in the final state. Using the m_{lb} shape prediction from MadGraph a top-quark mass of 172.3 ± 1.3 GeV is obtained, consistent with previous measurements and with the world average. In addition, a study is presented in which the shape and/or normalisation of the measured m_{lb} distribution are used to extract the top-quark mass using different theory predictions, the MadGraph simulation as well as a fixed-order QCD calculation.

1 Introduction

The top quark is the heaviest of the known elementary particles. The top-quark mass, m_t, is a fundamental parameter of quantum chromodynamics (QCD) and its value affects theoretical predictions significantly. Most direct measurements of m_t are based on the full or partial kinematic reconstruction of top-quark decays in $t\bar{t}$ events and make use of the relation of reconstructed observables with the top-quark mass parameter m_t as implemented in the simulation. The current world average of direct measurements is m_t = 173.34 $\pm$0.27 (stat) $\pm$0.71 (syst) GeV [1]. Despite the remarkable precision that has been achieved already, the phenomenological applications give strong motivations to further decrease the uncertainty and to gain a better understanding of the relation between the experimental results for m_t and the mass parameters that are employed in the theory calculations [2, 3].

It is thus necessary to use alternative methods of top-quark mass determination and to study the involved theoretical uncertainties. One alternative approach is to determine m_t from the inclusive cross section measurement. This yields m_t as defined in the theory scheme of the employed calculation of top-quark pair production. However, the obtained precision is inferior compared to direct measurements, with similar contributions from experimental and theory uncertainties. In particular, sensitivity of the measurement to the theoretical description of the top-quark pair production, like missing higher order contributions, description of the proton structure and uncertainty on the value of the strong coupling constant, limit the precision of this method of top-quark mass determination [4]. Extensive theoretical studies were performed to explore different kinematic distributions involving the top-quark decays, as reviewed in [3], which could be promising in determination of an unambiguously defined top-quark mass with only small dependence on the description of the production mechanism.

Top quarks decay almost exclusively into a bottom quark (b) and a W boson. Events with top-quark pairs ($t\bar{t}$) can thus be classified according to the decay modes of the two W bosons. In the present analysis, we consider cases where top quarks and antiquarks decay semileptonically via $t \to W^+b \to l^+\nu b$. The invariant mass of the b-quark and the lepton, m_{lb}, is found [5] to be very sensitive to the top-quark mass value and not significantly affected by the details of the production mechanism or the choice of parton distribution functions (PDFs) of the proton. Already before the start of the LHC, it was suggested to employ the m_{lb} distribution in dileptonic t decays for measurements of m_t [6]. Next-to-leading-order (NLO) corrections to the top-quark decay are available [7, 8]. It has been concluded [5], that this distribution is well under theoretical control and can be used for the determination of the top-quark pole mass.

In the present analysis, the m_{lb} distribution of $t\bar{t}$ candidate events is reconstructed and confronted to theory predictions using Monte Carlo (MC) simulation and fixed-order QCD predictions at leading order (LO) and at next-to-leading order (NLO), which implies an unambiguously defined pole mass of the top quark. Pair production of top quarks yielding an electron and a muon in the final state is investigated. This channel is chosen because of its high experimental precision. The top-quark mass is determined by studying the m_t-dependence of m_{lb} and investigating the impact of both the relative shape and the absolute event rate of the m_{lb} distribution.

2 Theory and simulation

The production of top-quark pairs is simulated using MC techniques based on the matrix-element (ME) generator MADGRAPH 5.1.5.11 [9], interfaced with MADSPIN [10] for the decay of heavy resonances, PYTHIA 6.426 [11] for parton showering (PS) and hadronization using the

Z2* tune [12], and TAUOLA for the decay of τ leptons. In the simulation, the proton structure is described by the PDF set CTEQ6L1 [13]. Dedicated MC samples are employed to study the dependence of the result on the choice of the hard scattering scale (Q^2), of the ME-PS matching scale, and of the value of the top-quark mass. The $t\bar{t}$ simulation is normalized to the integrated luminosity of the data using the prediction at NNLO+NNLL QCD for the inclusive $t\bar{t}$ cross section [14]. In the following, the MC simulation will be referred to as MADGRAPH+PYTHIA.

The production of single top-quarks in association with a W boson is simulated using POWHEG 1.0 r1380 [15–18] interfaced with PYTHIA and TAUOLA. These tW samples are normalized using the cross-section prediction at approximate NNLO [19] taking the m_t dependence of the single-top production rate into account according to [20]. The production of top-quark pairs in association with vector bosons ($t\bar{t}+\gamma$, $t\bar{t}+W$, $t\bar{t}+Z$) as well as Drell-Yan and W+jet production are simulated with MADGRAPH interfaced with PYTHIA. Diboson production (WW, WZ, ZZ) as well as QCD-production of $e\mu$+jets events are simulated with PYTHIA. The MC generated events have been passed through a full simulation of the CMS detector based on GEANT 4 [21].

Alternatively, the fixed-order calculation MCFM [8, 22, 23] at LO and NLO is used for the determination of m_t. The simulated top-pair production includes semi-leptonic decays for both the top quark and antiquark. At NLO, the calculation includes radiative corrections to the decay and to the production, which allows for study their impact on the determined m_t. All calculations used in this analysis assume Wtb couplings according to the standard model.

3 Dataset and event selection

The present analysis makes use of $t\bar{t}$ candidate events with an opposite-charged lepton pair $e^{\pm}\mu^{\mp}$ in the final state. The data were recorded by the CMS experiment at a centre-of-mass energy $\sqrt{s} = 8$ TeV, corresponding to an integrated luminosity of 19.7 ± 0.9 fb^{-1}.

The events are triggered if they contain at least one isolated electron and at least one muon, with one of these lepton candidates having a transverse momentum, p_T, larger than 17 GeV and the other above 8 GeV. For the final selection, both electron and muon are required to have a transverse momentum of $p_T > 20$ GeV and a pseudo-rapidity of $|\eta| < 2.4$. Furthermore, they are required to fulfill the isolation condition [24] $I_{\text{rel}} < 0.1$ (electrons) and $I_{\text{rel}} < 0.12$ (muons), where I_{rel} is defined as the sum of the transverse energy deposits in a cone of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ around the lepton, relative to the transverse momentum of the lepton itself. Here, $\Delta\eta$ and $\Delta\phi$ are the differences in pseudorapidity and azimuthal angle between any element of energy and the axis of the lepton. The cone is chosen as $\Delta R = 0.4$ for muons and $\Delta R = 0.3$ for electrons and the isolation values are corrected for pileup effects. The invariant mass of the two leptons is required to be $m_{ll} > 20$ GeV. The pair with the largest scalar sum of lepton transverse momenta in the event is considered as a signal candidate.

In addition, the event must have at least two jets, with at least one of them being b-tagged. Jets are reconstructed using the anti-k_T clustering algorithm [25] with a distance parameter of $R = 0.5$ and are required to have $p_T > 30$ GeV and $|\eta| < 2.4$. The b-quark jets are identified by using secondary vertex information with a mis-identification rate of $\mathcal{O}(10\%)$. Electrons, muons, and jets are reconstructed using the particle-flow algorithm. The contamination from pileup is minimized by performing a charged-hadron subtraction and removing remaining neutral components. Top-quark pair events involving decays $W \to \tau\nu$ are considered as signal if the τ decays directly to e or μ and the final state of the event satisfies the $e^{\pm}\mu^{\mp}$ selection criteria described above.

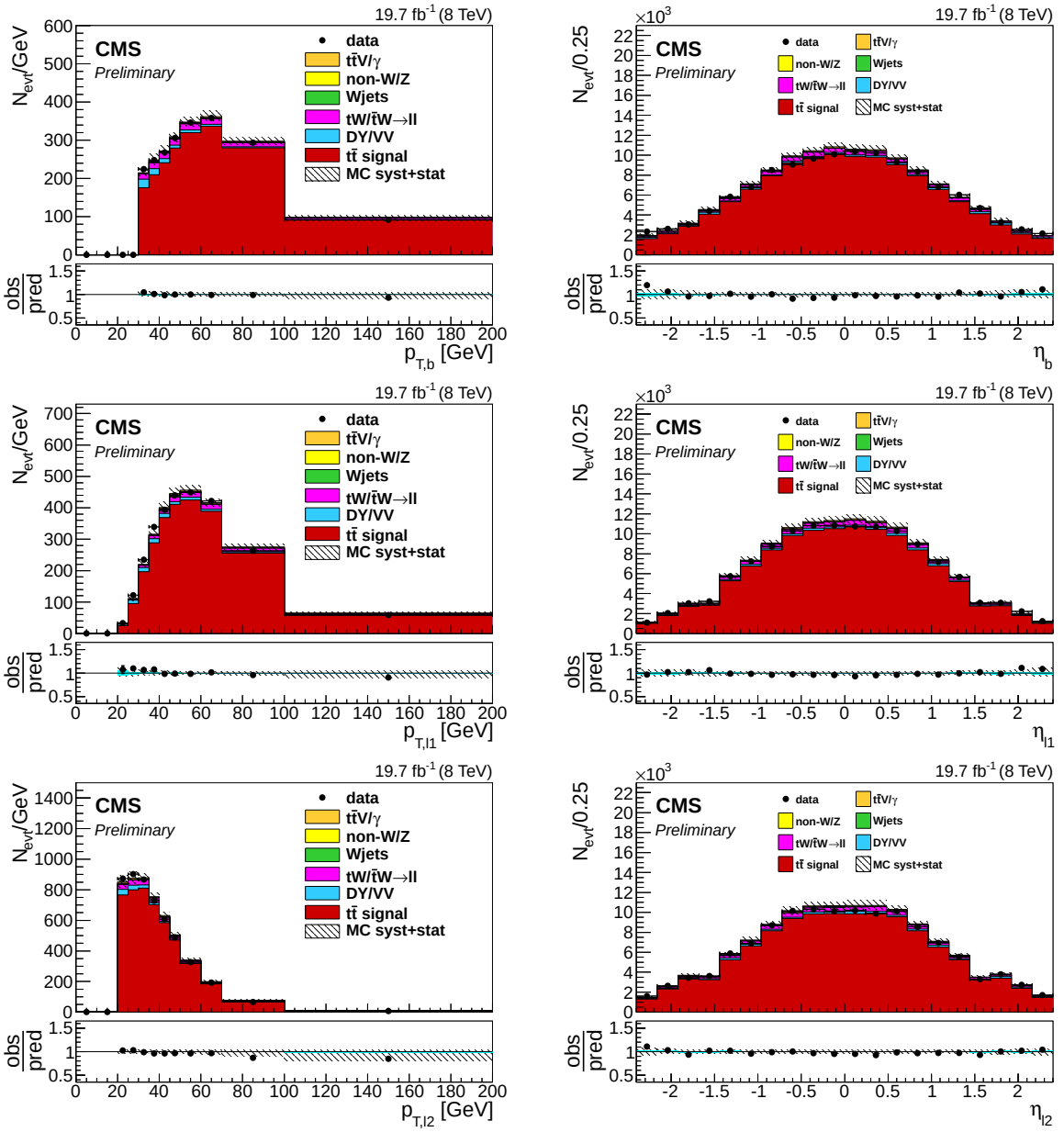

Figure 1: Event rate, shown as a function of the transverse momentum (left) and pseudorapidity (right) of the leading b-jet (top), the leading lepton (middle), and the second leading lepton (bottom). The bullets represent the experimental data with their statistical uncertainties, shown by the vertical error bars. The simulated rates for signal and different background contributions are represented by the histograms of different styles. The corresponding ratios of the observed and the predicted event rates are also shown. Here, the filled areas represent statistical uncertainties on the prediction. These are combined with systematic uncertainties due to systematic variations into a total uncertainty, indicated by the hatched areas in both, top and bottom panels. These systematic uncertainties do not account for luminosity and background normalization.

In Figure 1 comparisons of the data to the simulation are shown as functions of several variables. The simulation is corrected using data-driven estimates for trigger, lepton identification and isolation, and b-tagging efficiencies. Trigger efficiencies are derived as a function of lepton pseudorapidities ($|\eta_e|, |\eta_\mu|$) by using a set of independent triggers for both data and simulated events. Lepton identification and isolation efficiencies are extracted using a "tag and probe" technique using Z^0 decays into leptons. From the total number of the selected events, the expected fraction of $t\bar{t}$ signal is 92%, with 80% being $e\mu$ decays and 12% having at least one τ decay. The most significant background contributions arise from Drell-Yan (1.5%) and single-top quark production (4%).

4 Definition of m_{lb}

The top-quark decay chain considered in this analysis is $t \to Wb$ followed by $W \to l\nu$. With neglected lepton and b-quark masses, at leading order, m_{lb} is directly related to m_t and the mass of the W boson, m_W, as:

$$m_{lb}^2 = \frac{m_t^2 - m_W^2}{2} (1 - \cos\theta_{lb}) . \tag{1}$$

Here, θ_{lb} is the opening angle between the lepton and the b-quark in the W-boson rest frame. This relation illustrates that the m_{lb} distribution has an endpoint at

$$\max(m_{lb}) \approx \sqrt{m_t^2 - m_W^2}, \tag{2}$$

i.e. around 153 GeV for a top-quark mass of 173 GeV. The actual measured m_{lb} distribution is modified by several effects, among them higher-order corrections which include additional radiation within the top-quark decay, and experimental effects such as limited detector acceptance and finite resolution in the reconstruction of the lepton and jet four-momenta. Furthermore, the endpoint is smeared out by events in which the assignment of lepton and b-jet is not correct.

In this analysis, m_{lb} is reconstructed by choosing the permutation that minimizes the value of m_{lb} in each event and only the b-jet candidate with the highest p_T is considered together with both leptons ($e^\pm$ and $\mu^\mp$). By construction, only one top quark in each event is used. In this particular definition, referred to as m_{lb}^{min} in the following, correct ℓb combinations are found in about 85% of the cases [6]. The visible phase space of the analysis is defined by the transverse momenta of $p_T^l > 20$ GeV for the leptons and $p_T > 30$ GeV for b-jets in a pseudo-rapidity range of $|\eta| < 2.4$. Throughout this paper, the analysis is performed in this visible phase space. For the analysis using MCFM, the visible phase space is defined at parton level in the same way, i.e. $p_T^l > 20$ GeV for the leptons and $p_T^b > 30$ GeV for the b-quark in a pseudo-rapidity range of $|\eta| < 2.4$.

In Fig. 2, the reconstructed m_{lb}^{min} is compared to the simulation using m_t = 172.5 GeV. The m_t-dependence of the signal component is illustrated in Fig. 3. The value of m_t alters both, the absolute normalization of the predicted $t\bar{t}$ production rate and the shape of the m_{lb}^{min} distribution. Therefore, the top-quark mass can be determined from the absolute rate or the shape or the combination of both rate and shape information.

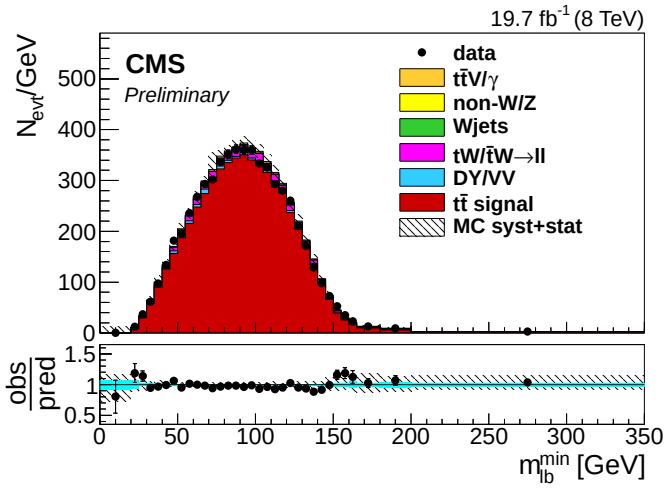

Figure 2: Observed m_{lb}^{min} distribution as compared to the simulation using m_t = 172.5 GeV. The bullets represent the experimental data with their statistical uncertainties, shown by the vertical error bars. The simulated rates for signal and different background contributions are represented by the histograms of different styles. The corresponding ratios of the observed and the predicted event rates are also shown. Here, the filled areas represent statistical uncertainties on the prediction. These are combined with systematic uncertainties due to systematic variations into a total uncertainty, indicated by the hatched areas in both, top and bottom panels. These systematic uncertainties do not account for luminosity and background normalization.

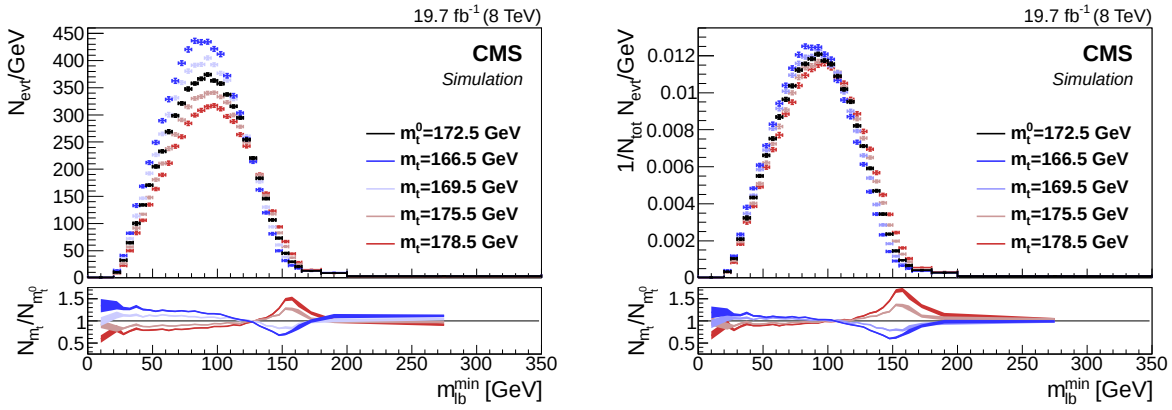

Figure 3: The absolute (left) and relative (right) event rates as a function of m_{lb}^{min} for different top-quark masses as predicted by the MADGRAPH+PYTHIA simulation.

5 Determination of m_t from the m_{lb}^{min} shape

The top-quark mass value is determined by comparison of the experimentally observed yields N_{obs} in individual bins of the (rate-normalized) m_{lb}^{min} distribution with the prediction N_{pred} for different values of m_t. For a given top-quark mass and bin i, an estimator $\chi^2(m_t, i)$ is defined as

$$\chi^2(m_t, i) = \frac{(N_{pred,i}(m_t) - N_{obs,i})^2}{\Delta^2_{pred,i} + \Delta^2_{obs,i}} \quad (3)$$

with $\Delta_{pred,i}$ and $\Delta_{obs,i}$ being the statistical uncertainties in a Gaussian approximation of the prediction or measurement, respectively.

To allow for a continuous scan of $\chi^2(m_t, i)$, the m_t-dependences of the predicted event yields $N_{pred,i}(m_t)$ are derived including both, signal and background. For this purpose, the yields are evaluated for different values of m_t = 166.5, 169.5, 171.5, 172.5, 173.5, 175.5, and 178.5 GeV. The resulting dependence is well described by a second-order polynomial, which determines the predicted number of events for any given value of m_t. The m_t-dependence of the m_{lb}^{min} shape for the background contributions is not considered. The background contribution from single-top quark events was studied and found negligible (see Section 6). The individual $\chi^2(m_t, i)$ are summed over the bins to a global $\chi^2(m_t)$ and the top-quark mass is determined from its minimum χ^2_{min}. The statistical uncertainty is obtained applying the criterium $\chi^2_{min} \pm 1$. The analysis was performed in a blind-analysis approach, using pseudo-experiments to determine the systematic uncertainties as well as possible biases of the extraction method. No bias was found.

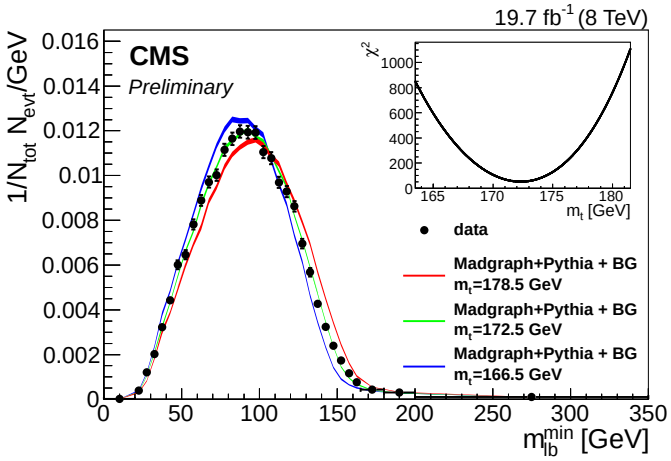

Figure 4: The normalized event yields for $t\bar{t}$ production at the LHC at $\sqrt{s}$=8 TeV, presented as a function of m_{lb}^{min}. The bullets represent the experimental data points and the error bars indicate their statistical uncertainties. The predictions from MADGRAPH+PYTHIA are presented for m_t=178.5 GeV (red band), m_t=172.5 GeV (green band) and m_t=166.5 GeV (blue band). The widths of the bands represent the statistical uncertainties of the predictions. The inset shows the χ^2 distribution as a function of the MC simulation parameter m_t as determined from the fit of the simulation to the shape of the data.

The resulting value of the top-quark mass with its total uncertainty is m_t=172.3 $\pm$1.3 GeV. The distributions of the normalized event yields as a function of the m_{lb}^{min}, together with predictions by MADGRAPH+PYTHIA is shown in Figure 4. A detailed description of the systematic uncertainties is given in the following section.

6 Systematic Uncertainties

Systematic uncertainties on the measurement arise from detector effects as well as from uncertainties on assumptions in the MC simulation. Each systematic uncertainty is investigated separately, and determined individually in each bin of the measurement, by variation of the corresponding efficiency, resolution, or scale within its uncertainty. For each variation, the event yield as a function of m_{lb}^{min} is recalculated. The respective value of m_t is derived and the difference to the nominal result is taken as systematic uncertainty, if statistically significant. Otherwise, the statistical uncertainty is used instead. The overall uncertainty on the measurement is then derived by adding the individual contributions in quadrature. The considered sources of experimental uncertainty are listed in Table 3 and described in the following.

- Lepton trigger efficiencies are measured using triggers that are only weakly correlated to the dilepton triggers used in the analysis. A dependence on the pseudorapidity of a few percent is observed and scale factors are derived, with typical variations of about 1.3%. The lepton identification and isolation uncertainties are determined using the 'tag-and-probe" method [26] with Z-boson event samples with a typical value of $\mathcal{O}(1\%)$.
- The uncertainty due to the jet energy scale is determined by variation of the jet energy in bins of transverse momentum and rapidity [27]. These variations are divided into main uncertainty sources, as listed in Table 1, with the effect of each source being estimated individually. The uncertainty on the jet energy resolution (JER) is determined by variation of the simulated JER by about $\pm 2.5\%$, $\pm 4\%$, and $\pm 5\%$, for the pseudorapidity regions $|\eta| < 1.7$, $1.7 < |\eta| < 2.3$, and $|\eta| > 2.3$, respectively [27].

Table 1: Subgroups of the JEC-sources of the correlated uncertainty

Short name	description
MPF *in-situ*	systematic uncertainties from Z+jet/γ+jet scale determinations
Intercalibration	relative uncertainties from FSR modeling
Pileup	pileup dependent terms
Uncorrelated	remaining uncertainties

- The uncertainty due to the b-tagging efficiency is determined by varying the b-tagging scale factors for the b jets by the uncertainties quoted in Ref. [28].
- The impact of the b-fragmentation modeling is evaluated by varying the Bowler-Lund b-fragmentation function in Tune Z2* to agree with x_B measurements by ALEPH [29] and DELPHI [30]. The differences are symmetrized.
- The flavor-dependent hadronization uncertainty is part of the JES uncertainty and comes from differences in the energy response for different jet flavors and flavor mixtures. It originates in differences between the Lund fragmentation model (PYTHIA) and cluster fragmentation (HERWIG++ [31]) and is evaluated for each jet flavor independently. The resulting uncertainties on m_t are summed linearly.
- The semi-leptonic branching fraction of B hadrons is varied within the uncertainties given in [32].
- The uncertainty due to the choice of the matrix element generator is estimated by comparing the central result obtained using MADGRAPH+PYTHIA with the result obtained with POWHEG+PYTHIA. The differences are symmetrized.
- The pileup model estimates the mean number of pp interactions to be about 20.5 events per bunch crossing. This value is derived from the total inelastic proton-

proton cross section [33]. The systematic uncertainty is determined by varying this cross section within $\pm 5\%$ and propagating the effect to the simulation.

- The uncertainties on the background arise from the modeling of the background processes. A global variation is applied to each background contribution to account for both, normalization and differences between different background models. In the case of single-top quark background, a variation of 30% accounts for the theory uncertainty on the absolute rate and includes additional uncertainties due to PDF. In addition, the influence of the m_t-dependence of the contribution of single-top quark production on the resulting m_{lb} shape is studied for the central and extreme values of m_t and found to be small as compared to statistical uncertainties. The same global variation is applied to the second dominant background contribution from Drell-Yan events. The predicted cross section has an uncertainty of $\approx 5\%$ and the variation by 30% used here is a conservative estimate of the uncertainty, suggested by data-driven estimates. Although negligible, the contributions of the remaining backgrounds are varied by the same amount.
- In the MADGRAPH signal event sample, the renormalization and factorization scales are set to $\mu_r = \mu_f = Q$, with $Q = \sqrt{m_t^2 + \sum p_T^2}$, where the transverse momenta sum runs over all additional final-state partons. The uncertainty due to missing higher-order corrections in the used model is determined by the variation of the renormalization and factorization scales up and down simultaneously by factors of 2 around the central values. The influence of the scale choice for the matching of the matrix element (ME) and the parton shower (PS) part in MADGRAPH is studied by scaling its reference value of 20 GeV by a factor of two up and down. The corresponding relative variations of m_{lb}^{min} for $m_t = 172.5$ GeV with respect to the central result are used also for other choices of m_t.
- The uncertainty on the luminosity measurement is 2.6% [34].
- The prediction is reweighted to describe the transverse momentum of the top quark, p_T^t, as measured in the data [35]. The difference between the resulting m_t to its central value is assigned as an additional symmetric uncertainty.
- The modelling of underlying event (UE) and color reconnection (CR) effects and their impact on the top-quark mass determination are investigated using different Perugia11 tunes of PYTHIA including the tune without CR effects, respectively. The effects of both, UE and CR are propagated into the uncertainty on the top-quark mass and are found to be small.
- The uncertainties due to PDF are estimated as follows. The PDF uncertainties, PDF_{MC}, originate mostly from the unmeasured phase space and are estimated using eigenvectors of the CT10NLO PDF at 90% CL according to prescription of the CT10 collaboration.
- The uncertainty due to choice of a second-order polynomial as parameterization function for the m_t dependence of the event yields was studied by using other functional forms. The maximum variation of the resulting m_t is assigned as an additional uncertainty.

7 Determination of m_t using MCFM

In the following, an alternative technique is presented which allows to use fixed-order calculations, such as MCFM, to determine their top-quark mass parameter. The advantage of this

technique is the unambiguous definition of m_t, e.g. in MCFM as a pole mass of the top quark. However, MCFM, like many other theory predictions, does not implement the simulation of the complete final state of the events at detector level. Here, a folding approach is presented, in which the theory prediction is folded with the detector response as described by a response matrix, such that the folded theory prediction, for a given observable, can be confronted with the data at reconstruction level. The analysis, performed here on MCFM as one example, can be extended to any higher-order differential calculation of $t\bar{t}$ production and decay, using a cleanly defined top-quark mass, once available.

The response matrix M^{resp} relates the differential cross section prediction $\vec{\sigma}_{\text{pred}}$ to the number of events $\vec{N}_{\text{pred,det}}$ expected at detector level:

$$\vec{N}_{\text{pred,det}} = \mathcal{L} \cdot M^{\text{resp}} \cdot \vec{\sigma}_{\text{pred}}, \tag{4}$$

where $\mathcal{L}$ stands for the integrated luminosity.

M^{resp} contains all corrections, including acceptance corrections, bin-to-bin migrations and detector efficiencies, and is determined from Monte Carlo simulation as

$$M_{ij}^{\text{resp}} = \frac{N_{ij}}{\sum_{j=0}^{n} N_{ij}}. \tag{5}$$

where N_{ij} represent the number of events that are generated in a certain bin i and are reconstructed in bin j, and n stands for the total number of bins. The entries N_{i0} account for events that belong to the generator bin i but, due to detector efficiencies, do not enter the reconstructed distribution.

In Fig. 5 (left), the response matrix M_{ij}^{resp} is shown for the case of m_t=172.5 GeV. A strong correlation between the generated and reconstructed distributions is observed. The contribution from the efficiency, i.e. the fraction of events that were generated in the m_{lb} bin i, that also fulfill the event selection criteria and are reconstructed in any of the visible $\text{m}_{\text{lb}}^{\text{min}}$ bins,

$$\epsilon_i = \frac{\sum_{j=1}^{n} N_{ij}}{\sum_{j=0}^{n} N_{ij}}. \tag{6}$$

is shown in Fig. 5 (right).

In this analysis, the fixed-order calculation MCFM is used which is available at leading order LO and next-to-leading order (NLO) QCD. The corresponding MCFM routines are modified to apply the visible phase-space requirements as described in section 4. The configuration of MCFM that is referred to as MCFM (NLO) in the following, includes NLO contributions only to the production of the top-quark pair, the decay of the top-quarks modelled in leading order. This configuration is expected to correspond closest to the decay definition in MADGRAPH+PYTHIA.

The fit of the $\text{m}_{\text{lb}}^{\text{min}}$ distribution as predicted by MCFM (NLO) for different values of m_t to the shape of the data results in a value of m_t =171.4 $^{+1.0}_{-1.1}$ GeV.

The experimental uncertainties, including those on the Monte Carlo mass parameter, are determined by use of individual response matrices for each source of uncertainty. The uncertainties of the response matrix and of the specific theory prediction are assumed uncorrelated and are therefore added in quadrature. The resulting distribution can be directly compared to the data, after addition of the backgrounds, obtained using the MADGRAPH + PYTHIA simulation.

Effects from radiative corrections are also studied. The result, when using the fit to the shape with MCFM at LO, is m_t =171.5 $^{+1.1}_{-1.1}$ GeV, very similar to the one above with MCFM (NLO),

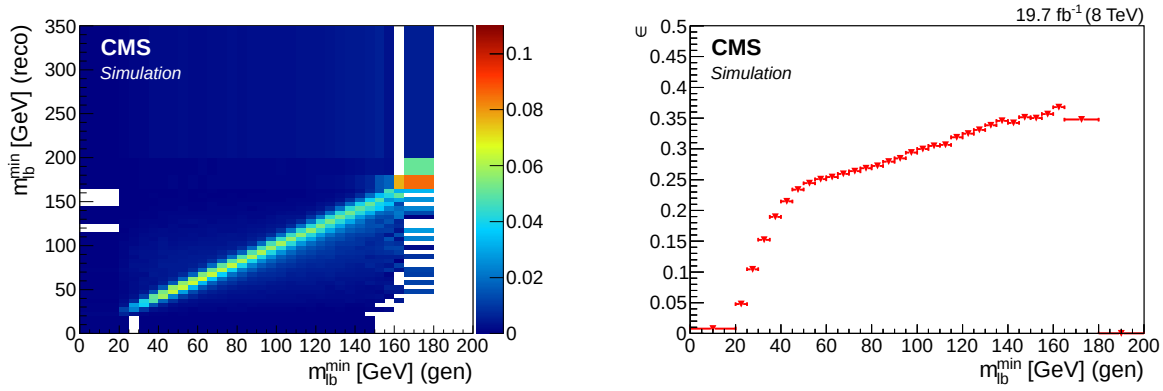

Figure 5: Left panel: the response matrix, quantifying all detector detector effects, as defined in Eq. 5, presented as a function of generated and reconstructed values of m_{lb}^{min}. Right panel: the efficiency, as defined in Eq. 6, shown as a function of the reconstructed value of m_{lb}^{min}.

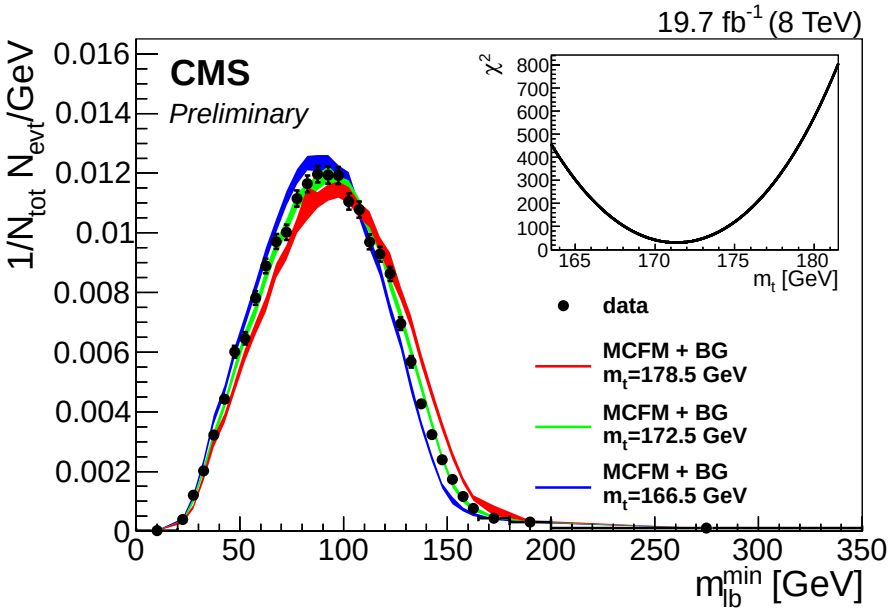

Figure 6: The normalized event yields for $t\bar{t}$ production at the LHC at $\sqrt{s}$=8 TeV, presented as a function of m_{lb}^{min}. The bullets represent the experimental data points and the error bars indicate their statistical uncertainties. The predictions from the MCFM calculation at NLO are presented for m_t=178.5 GeV (red band), m_t=172.5 GeV (green band) and m_t=166.5 GeV (blue band). The widths of the bands represent the statistical uncertainties of the predictions. The inset shows the χ^2 distribution as a function of the mass parameter m_t in MCFM as determined from the fit of the simulation to the shape of the data.

indicating that the sensitivity of this measurement to the details of the production process are indeed small [5]. Alternatively, using MCFM including real and virtual corrections to next-to-leading order also in the decay, a shift of the resulting value of the top-quark mass of $+0.9$ GeV is observed.

Additional Systematic Uncertainties Related to MCFM

The total uncertainty of the top-quark mass extracted from MCFM is composed of experimental uncertainties that affect the response matrix and theoretical uncertainties affecting the MCFM calculation. The sources of uncertainties on the response matrix are the same as those described in section 6, propagated to m_{lb}^{min} by applying a different folding matrix to the calculation for each source. The additional uncertainties due to the MCFM calculation are the following:

- The MCFM prediction is obtained using MSTW08 PDF at LO and NLO, respectively. The corresponding PDF uncertainty, PDF_{MCFM}, is calculated according to the prescription of the MSTW08 group.
- The value of $\alpha_S(M_Z)$ is varied consistently in calculation and PDF and the differences to the central result are considered as an additional uncertainty.
- The renormalization and factorization scales are set to m_t and are varied independently by a factor of 2 up and down. The maximum deviation in each m_{lb}^{min} bin is taken as the systematic uncertainty, $\text{scale}_{\text{MCFM}}$.
- The b-quark mass m_b is set to 4.75 GeV and varied by ± 0.25 GeV.

A breakdown of the systematic uncertainties is provided in Table 3. The impact of top quark p_T and PDF_{MC} uncertainties on the MCFM-based result is slightly reduced in comparison to the mass determination using MADGRAPH, since here these affect only the detector response uncertainty, and not the shape of the prediction. For lack of statistics of the corresponding samples, the uncertainties on CR and UE are taken from the result obtained using MADGRAPH+PYTHIA.

8 Determination of m_t using the top-quark pair production rate

The inclusive top-quark pair cross section, i.e. the event rate, also provides information about the value of the top-quark mass [4]. However, fixed-order LO and NLO predictions as implemented in MCFM have been seen to underestimate the absolute value of the top-quark pair cross section significantly and are thus not suitable for the extraction of the top-quark mass. Here, a study is performed in which the top-quark mass m_t is extracted from the measured reconstructed number of top-quark pair events using the predictions from the MADGRAPH+PYTHIA simulation in combination with the NNLO calculation [14].

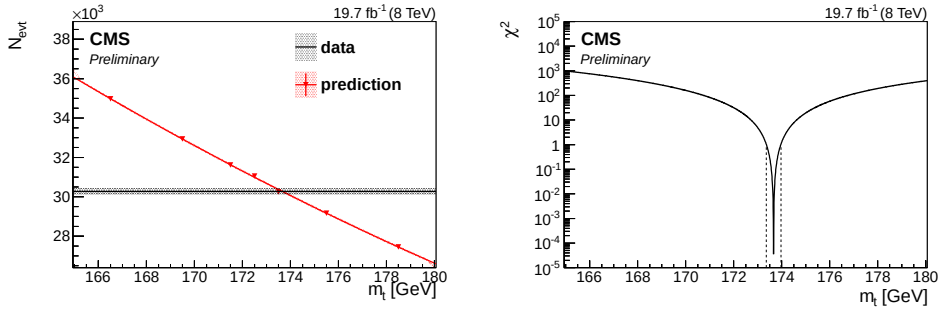

Figure 7: Left panel: dependence of the event yields on the top-quark mass hypothesis. The dark (light) bands represent N_{obs} (N_{pred}) with the width of the band showing the statistical uncertainty. Right panel: the estimator χ^2 as a function of m_t.

The extraction procedure is illustrated in Fig. 7. For each value of m_t, the normalization of MADGRAPH+PYTHIA is obtained from the NNLO prediction [14]. In both calculations the same numerical value for m_t is used. From this procedure a value of m_t =173.7 $^{+3.5}_{-3.4}$ GeV is obtained with systematic uncertainties that are significantly larger than those obtained when using the shape alone. The experimental systematic uncertainties are estimated as described in 6 and all systematic uncertainties are listed in table 3. The uncertainties due to the identification of m_t in MADGRAPH+PYTHIA and NNLO calculation are estimated by independent variation of renormalization and factorization scales (scale$_{\text{NNLO normalization}}$) and PDF uncertainties (PDF$_{\text{NNLO normalization}}$) separately for MADGRAPH+PYTHIA and NNLO calculation. Both are taken from [14].

Using MADGRAPH+PYTHIA, the rate and m_{lb}^{min}-shape information can also be combined. The prediction for the dependence of the absolute m_{lb}^{min} distribution in rate and shape is shown in Figure 8 in comparison with the data. The fit of the prediction to the data yields m_t =173.1 $^{+1.9}_{-1.8}$. In the case of the combined fit to shape and rate, partial cancellations of uncertainties are observed. In this case, experimental normalization uncertainties, such as luminosity, background normalization and electron identification and isolation efficiencies are dominant.

9 Summary

The top-quark mass is determined using the invariant mass distribution m_{lb}^{min} of the final state charged lepton (electron or muon) and the b-jet from leptonic top-quark decays. The analysis makes use of data of the CMS experiment in proton-proton collisions at a center-of-mass energy of 8 TeV at the LHC. The shape of the observed m_{lb}^{min} distribution in the data is used to extract the top-quark mass m_t in MADGRAPH+PYTHIA yielding a value $m_t = 172.3^{+1.23}_{-1.3}$ GeV consistent with other determinations at the LHC and the world average.

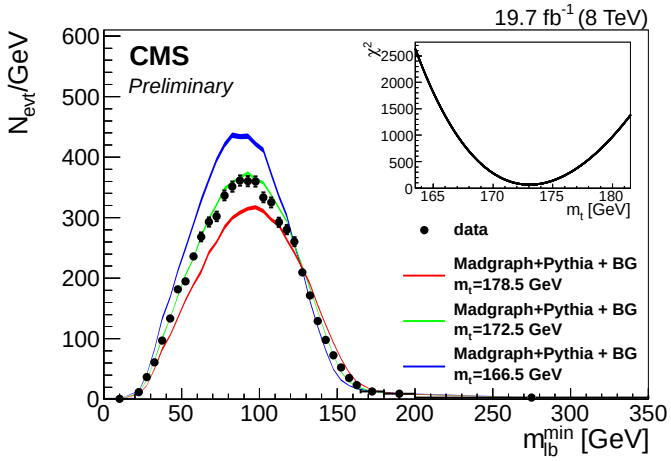

Figure 8: The absolute event yields for $t\bar{t}$ production at the LHC at $\sqrt{s}$=8 TeV, presented as a function of m_{lb}^{min}. The bullets represent the experimental data points and the error bars indicate their statistical uncertainties. The predictions from MADGRAPH+PYTHIA are presented for m_t=178.5 GeV (red band), m_t=172.5 GeV (green band) and m_t=166.5 GeV (blue band). The widths of the bands represent the statistical uncertainties of the predictions. The inset shows the χ^2 distribution as a function of the MC simulation parameter m_t as determined from the fit of the simulation to the shape and rate of the data.

The same data are also confronted with the fixed-order calculation MCFM to both leading order and next-to-leading order QCD. The results on m_t with the full uncertainties, as obtained from the shape, rate and rate+shape fits are presented in Table 2. The statistical, systematic, and theory uncertainties on m_t, as obtained from the fits using the MADGRAPH+PYTHIA simulation, for rate and/or shape as well as MCFM LO and NLO calculations, for shape, are listed in Table 3.

Table 2: The top-quark mass determined from confronting the observed event yield for $t\bar{t}$ production at CMS to theory predictions by using MADGRAPH+PYTHIA simulation and the fixed-order QCD calculation MCFM. The calculation referred to as MCFM (NLO) includes next-to-leading order contributions to the production of the top-quark pair only.

		Fitted m_t [GeV]
Prediction	Fit method	from m_{lb}^{min}
MADGRAPH+PYTHIA	shape+rate	$173.1\ ^{+1.9}_{-1.8}$
MADGRAPH+PYTHIA	rate	$173.7\ ^{+3.5}_{-3.4}$
MADGRAPH+PYTHIA	shape	$172.3\ ^{+1.3}_{-1.3}$
MCFM (LO)	shape	$171.5\ ^{+1.1}_{-1.1}$
MCFM (NLO)	shape	$171.4\ ^{+1.0}_{-1.1}$

Table 3: Breakdown of systematic uncertainties on the top-quark mass value, obtained by confronting the inclusive event rate and the shape of the m_{lb}^{min} distribution the theory predictions by MADGRAPH+PYTHIA and MCFM. The calculation referred to as MCFM (NLO) includes next-to-leading order contributions to the production of the top-quark pair only.

Source	Δm_t [GeV]				
	MADGRAPH +PYTHIA			MCFM shape	
	shape+rate	rate	shape	NLO	LO
Statistical uncertainty	$^{+0.22}_{-0.22}$	$^{+0.31}_{-0.30}$	$^{+0.32}_{-0.31}$	$^{+0.39}_{-0.40}$	$^{+0.40}_{-0.40}$
b-tagging	$^{+0.19}_{-0.10}$	$^{+0.45}_{-0.32}$	$^{-0.08}_{+0.14}$	$^{-0.08}_{+0.14}$	$^{-0.09}_{+0.14}$
b-mistag	$^{+0.14}_{+0.01}$	$^{+0.46}_{+0.03}$	$^{-0.18}_{+0.01}$	$^{-0.19}_{+0.01}$	$^{-0.19}_{-0.04}$
Electron energy scale	$^{-0.04}_{+0.03}$	$^{+0.02}_{-0.04}$	$^{-0.11}_{+0.10}$	$^{-0.12}_{+0.09}$	$^{-0.12}_{+0.09}$
Electron ID	$^{+0.45}_{-0.45}$	$^{+0.78}_{-0.78}$	$^{+0.11}_{-0.11}$	$^{+0.10}_{-0.11}$	$^{+0.10}_{-0.11}$
Muon energy scale	$^{-0.01}_{+0.01}$	$^{+0.08}_{-0.08}$	$^{-0.09}_{+0.12}$	$^{-0.09}_{+0.12}$	$^{-0.09}_{+0.14}$
Muon ID	$^{+0.24}_{-0.24}$	$^{+0.48}_{-0.48}$	$^{-0.01}_{+0.01}$	$^{-0.01}_{+0.01}$	$^{-0.01}_{+0.01}$
Jet energy resolution	$^{+0.03}_{-0.04}$	< 0.01	$^{+0.06}_{-0.07}$	$^{+0.05}_{-0.08}$	$^{+0.06}_{-0.09}$
MPF *in-situ* JES component	$^{-0.06}_{+0.06}$	$^{+0.09}_{-0.09}$	$^{-0.23}_{+0.23}$	$^{-0.24}_{+0.22}$	$^{-0.24}_{+0.23}$
Uncorrelated JES component	$^{+0.06}_{-0.07}$	$^{+0.27}_{-0.28}$	$^{+0.32}_{-0.33}$	$^{+0.30}_{-0.35}$	$^{+0.34}_{-0.35}$
Flavor-dependent hadronization	$^{-0.07}_{+0.09}$	$^{+0.09}_{-0.09}$	$^{-0.27}_{+0.30}$	$^{-0.30}_{+0.27}$	$^{-0.28}_{+0.29}$
Pileup JES component	$^{+0.17}_{-0.19}$	$^{+0.42}_{-0.42}$	$^{+0.09}_{-0.11}$	$^{+0.07}_{-0.12}$	$^{+0.07}_{-0.14}$
Intercalibration JES component	$^{-0.}_{+0.01}$	$^{+0.01}_{-0.01}$	$^{-0.01}_{+0.02}$	$^{-0.02}_{+0.01}$	$^{-0.02}_{+0.01}$
MHI (Pileup)	$^{-0.14}_{+0.14}$	$^{-0.19}_{+0.19}$	$^{-0.09}_{+0.08}$	$^{-0.09}_{+0.06}$	$^{-0.10}_{+0.07}$
Trigger	$^{+0.30}_{-0.30}$	$^{+0.58}_{-0.58}$	< 0.01	< 0.01	< 0.01
Backgrounds	$^{+0.55}_{-0.53}$	$^{+1.17}_{-1.19}$	$^{-0.12}_{+0.12}$	$^{-0.17}_{+0.17}$	$^{-0.22}_{+0.21}$
Luminosity	$^{+0.66}_{-0.67}$	$^{+1.27}_{-1.30}$	< 0.01	< 0.01	< 0.01
Top p_T	$^{+0.01}_{-0.01}$	$^{-0.60}_{+0.60}$	$^{+0.66}_{-0.66}$	$^{+0.28}_{-0.28}$	$^{+0.27}_{-0.27}$
Fit parameterization m_t	$^{-0.04}_{+0.04}$	$^{-0.06}_{+0.06}$	$^{+0.07}_{-0.07}$	$^{+0.07}_{-0.07}$	$^{+0.07}_{-0.07}$
ME-PS matching threshold	$^{-0.10}_{+0.10}$	$^{-0.13}_{+0.43}$	$^{+0.13}_{-0.21}$	$^{+0.13}_{-0.24}$	$^{+0.15}_{-0.20}$
Renormalization and factorization (Q^2) scale	$^{-0.07}_{-0.11}$	$^{-0.55}_{+0.47}$	$^{+0.50}_{-0.59}$	$^{+0.50}_{-0.50}$	$^{+0.45}_{-0.51}$
b-fragmentation	$^{+0.57}_{-0.57}$	$^{+0.54}_{-0.54}$	$^{+0.62}_{-0.62}$	$^{+0.43}_{-0.43}$	$^{+0.40}_{-0.40}$
B branching fractions	$^{+0.07}_{-0.02}$	$^{-0.05}_{+0.05}$	$^{+0.20}_{-0.10}$	$^{+0.18}_{-0.10}$	$^{+0.17}_{-0.10}$
ME generator	$^{-0.73}_{+0.73}$	$^{-1.16}_{+1.16}$	$^{-0.19}_{+0.19}$	$^{-0.10}_{+0.10}$	$^{-0.08}_{+0.08}$
Color reconnection	$^{+0.22}_{-0.22}$	$^{+0.19}_{-0.19}$	$^{+0.12}_{-0.12}$	$^{+0.12}_{-0.12}$	$^{+0.12}_{-0.12}$
Underlying event	$^{-0.10}_{+0.21}$	$^{-0.14}_{+0.14}$	$^{-0.24}_{+0.14}$	$^{-0.24}_{+0.14}$	$^{-0.24}_{+0.14}$
PDF_{MC}	$^{+0.20}_{-0.08}$	$^{+0.25}_{-0.15}$	$^{+0.20}_{-0.06}$	$^{+0.09}_{-0.02}$	$^{+0.04}_{-0.03}$
$\text{scale}_{\text{NNLO normalization}}$	$^{+0.85}_{-0.88}$	$^{+1.65}_{-1.71}$	< 0.01		
$\text{PDF}_{\text{NNLO normalization}}$	$^{+0.66}_{-0.67}$	$^{+1.27}_{-1.30}$	< 0.01		
$\text{scale}_{\text{MCFM}}$				$^{+0.05}_{-0.03}$	$^{+0.04}_{-0.09}$
PDF_{MCFM}				$^{-0.02}_{+0.03}$	$^{-0.01}_{+0.01}$
$\alpha_S(M_Z)$ variation				< 0.01	< 0.01
m_b variation				< 0.01	< 0.01
Total uncertainty	$^{+1.85}_{-1.83}$	$^{+3.45}_{-3.43}$	$^{+1.29}_{-1.33}$	$^{+1.04}_{-1.08}$	$^{+1.02}_{-1.11}$

References

[1] ATLAS, CDF, CMS, and D0 Collaborations, "First combination of Tevatron and LHC measurements of the top-quark mass", (2014). arXiv:1403.4427.

[2] A. Buckley et al., "General-purpose event generators for LHC physics", *Phys.Rept.* **504** (2011) 145–233, doi:10.1016/j.physrep.2011.03.005, arXiv:1101.2599.

[3] S. Moch et al., "High precision fundamental constants at the TeV scale", (2014). arXiv:1405.4781.

[4] CMS Collaboration, "Determination of the top-quark pole mass and strong coupling constant from the $t\bar{t}$ production cross section in pp collisions at $\sqrt{s}$ = 7 TeV", *Phys.Lett. B* **728** (2014) 496, doi:10.1016/j.physletb.2013.12.009, arXiv:1307.1907. Corrigendum to appear.

[5] S. Biswas, K. Melnikov, and M. Schulze, "Next-to-leading order QCD effects and the top quark mass measurements at the LHC", *JHEP* **08** (2010) 048, doi:10.1007/JHEP08(2010)048, arXiv:1006.0910.

[6] M. Beneke et al., "Top quark physics", (2000). arXiv:hep-ph/0003033.

[7] K. Melnikov and M. Schulze, "NLO QCD corrections to top quark pair production and decay at hadron colliders", *JHEP* **08** (2009) 049, doi:10.1088/1126-6708/2009/08/049, arXiv:0907.3090.

[8] J. M. Campbell and R. K. Ellis, "Top-quark processes at NLO in production and decay", (2012). arXiv:1204.1513.

[9] J. Alwall et al., "MadGraph 5: Going Beyond", *JHEP* **06** (2011) 128, doi:10.1007/JHEP06(2011)128, arXiv:1106.0522.

[10] P. Artoisenet, R. Frederix, O. Mattelaer, and R. Rietkerk, "Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations", *JHEP* **03** (2013) 015, doi:10.1007/JHEP03(2013)015, arXiv:1212.3460.

[11] T. Sjöstrand, S. Mrenna, and P. Skands, "PYTHIA 6.4 physics and manual", *JHEP* **05** (2006) 026, doi:10.1088/1126-6708/2006/05/026, arXiv:hep-ph/0603175.

[12] CMS Collaboration, "Measurement of the Underlying Event Activity at the LHC with $\sqrt{s}$ = 7 TeV and Comparison with $\sqrt{s}$ = 0.9 TeV", *JHEP* **09** (2011) 109, doi:10.1007/JHEP09(2011)109, arXiv:1107.0330.

[13] J. Pumplin et al., "New generation of parton distributions with uncertainties from global QCD analysis", *JHEP* **0207** (2002) 012, doi:10.1088/1126-6708/2002/07/012, arXiv:hep-ph/0201195.

[14] M. Czakon, P. Fiedler, and A. Mitov, "Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $O(\alpha_S^4)$", *Phys.Rev.Lett.* **110** (2013), no. 25, 252004, doi:10.1103/PhysRevLett.110.252004, arXiv:1303.6254.

[15] P. Nason, "A New method for combining NLO QCD with shower Monte Carlo algorithms", *JHEP* **11** (2004) 040, doi:10.1088/1126-6708/2004/11/040, arXiv:hep-ph/0409146.

[16] S. Frixione, P. Nason, and C. Oleari, "Matching NLO QCD computations with Parton Shower simulations: the POWHEG method", *JHEP* **11** (2007) 070, doi:10.1088/1126-6708/2007/11/070, arXiv:0709.2092.

[17] S. Alioli, P. Nason, C. Oleari, and E. Re, "A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX", *JHEP* **06** (2010) 043, doi:10.1007/JHEP06(2010)043, arXiv:1002.2581.

[18] E. Re, "Single-top Wt-channel production matched with parton showers using the POWHEG method", *Eur.Phys.J. C* **71** (2011) 1547, doi:10.1140/epjc/s10052-011-1547-z, arXiv:1009.2450.

[19] N. Kidonakis, "Differential and total cross sections for top pair and single top production", (2012). arXiv:1205.3453.

[20] N. Kidonakis, "Two-loop soft anomalous dimensions for single top quark associated production with a W- or H-", *Phys.Rev.* **D82** (2010) 054018, doi:10.1103/PhysRevD.82.054018, arXiv:1005.4451.

[21] GEANT4 Collaboration, "GEANT4: A Simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

[22] J. M. Campbell and R. Ellis, "MCFM for the Tevatron and the LHC", *Nucl. Phys. Proc. Suppl.* **205-206** (2010) 10, doi:10.1016/j.nuclphysbps.2010.08.011, arXiv:1007.3492.

[23] S. Badger, R. Sattler, and V. Yundin, "One-Loop Helicity Amplitudes for $t\bar{t}$ Production at Hadron Colliders", *Phys.Rev.* **D83** (2011) 074020, doi:10.1103/PhysRevD.83.074020, arXiv:1101.5947.

[24] CMS Collaboration Collaboration, "Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **1211** (2012) 067, doi:10.1007/JHEP11(2012)067, arXiv:1208.2671.

[25] M. Cacciari, G. P. Salam, and G. Soyez, "The anti- k t jet clustering algorithm", *Journal of High Energy Physics* **2008** (2008), no. 04, 063.

[26] CMS Collaboration, "Measurement of the Drell-Yan cross sections in pp collisions at $\sqrt{s}$ = 7 TeV with the CMS experiment", *JHEP* **04** (2008) 007, doi:10.1007/JHEP10(2011)007, arXiv:1108.0566.

[27] CMS Collaboration, "Determination of jet energy calibration and transverse momentum resolution in CMS", *JINST* **06** (2011) P11002, doi:10.1088/1748-0221/6/11/P11002, arXiv:1107.4277.

[28] CMS Collaboration, "b-Jet Identification in the CMS Experiment", CMS Physics Analysis Summary CMS-PAS-BTV-11-004, 2011.

[29] ALEPH Collaboration, "Study of the fragmentation of b quarks into B mesons at the Z peak", *Phys.Lett.* **B512** (2001) 30–48, doi:10.1016/S0370-2693(01)00690-6, arXiv:hep-ex/0106051.

[30] DELPHI Collaboration, "A study of the b-quark fragmentation function with the DELPHI detector at LEP I and an averaged distribution obtained at the Z Pole", *Eur.Phys.J.* **C71** (2011) 1557, doi:10.1140/epjc/s10052-011-1557-x, arXiv:1102.4748.

[31] M. Bahr et al., "Herwig++ Physics and Manual", *Eur. Phys. J* **C58** (2008) 639–707.

[32] K. Nakamura and P. D. Group, "Review of Particle Physics", *Journal of Physics G: Nuclear and Particle Physics* **37** (2010), no. 7A, 075021.

[33] TOTEM Collaboration, "First measurement of the total proton-proton cross section at the LHC energy of $\sqrt{s}$ =7 TeV", *Europhys. Lett.* **96** (2011) 21002, doi:10.1209/0295-5075/96/21002, arXiv:hep-ex/1110.1395.

[34] CMS Collaboration, "CMS Luminosity Based on Pixel Cluster Counting -Summer 2013 Update", CMS Physics Analysis Summary CMS-PAS-LUM-13-001, 2010.

[35] CMS Collaboration, "Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s}$ = 8 TeV", (2014). TOP-12-028 paper, to be submitted to EPJC.