CMS Physics Analysis Summary

Contact: cms-pag-conveners-top@cern.ch 2016/05/17

Determination of the normalised invariant mass distribution of $t\bar{t}$+jet and extraction of the top quark mass

The CMS Collaboration

Abstract

A measurement of the top quark mass from top quark pair ($t\bar{t}$) events produced in association with additional hard jets is performed in pp collisions at $\sqrt{s} = 8$ TeV with the CMS detector using data recorded in 2012, corresponding to an integrated luminosity of 19.7 fb^{-1}. The mass is extracted from the normalised invariant mass distribution of the $t\bar{t}$+jet system at reconstruction level as well as from the related normalised differential cross section. Both measurements are performed in the dileptonic decay channels (e^+e^-, $\mu^+\mu^-$ and $e^\pm\mu^\pm$) of the $t\bar{t}$ quark pairs.

1 Introduction

The mass of the top quark (m_t) is an important parameter for precision tests of the standard model and is either measured directly, based on the kinematic reconstruction of the decay products of the quark, or based on the mass dependency of the inclusive $t\bar{t}$ production cross section. Direct determinations of the top quark mass have been performed at the Tevatron and LHC colliders and the current world average of direct measurements is $m_t = 173.34 \pm 0.27\,(\text{stat}) \pm 0.71\,(\text{syst})$ GeV [1], while the most precise combination of direct measurements by CMS yields $m_t = 172.44 \pm 0.13\,(\text{stat}) \pm 0.41\,(\text{syst})$ GeV [2].

In this note an alternative approach to measure the mass is presented, following a method proposed in [3] using the normalised differential cross section as a function of the invariant mass of the $t\bar{t}$ system and the leading additional jet in the event, which does not stem from the top quark decays. The observable is defined as

$$\rho_s = \frac{2 \cdot m_0}{\sqrt{s_{t\bar{t}+\text{jet}}}} ,$$

where m_0 is a scale of the order of the top quark mass and $\sqrt{s_{t\bar{t}+\text{jet}}}$ denotes the invariant mass of the two top quark candidates and the additional leading jet.

First the measurement of the reconstructed ρ_s distribution is described and the extraction of the top quark mass using a traditional template technique is presented. For this, the observed ρ_s distribution is compared to simulations at reconstruction level generated with different top quark masses. Then the normalised differential cross section is derived by unfolding the distribution of ρ_s in order to correct for detector effects. Finally, the extraction of the top quark mass from the measured differential cross section is performed using theoretical predictions assuming different top quark masses. The extraction of the top quark mass from this differential cross section has been previously measured by the ATLAS Collaboration in the lepton+jets decay channel [4].

The measurements are performed using $t\bar{t}$ events in pp collisions recorded at $\sqrt{s} = 8$ TeV with the CMS detector [5], corresponding to an integrated luminosity of $19.7\,\text{fb}^{-1}$. The events are selected in the dileptonic decay channel with two oppositely charged isolated leptons (electrons or muons) and at least two jets and one b-tagged jet. The analysis presented in this document follows the strategy of other CMS measurements of differential $t\bar{t}$ cross sections and the $t\bar{t}$+jet studies in the same decay channel [6, 7]. This note focuses on the sensitivity of the method proposed, studying in detail the experimental and model uncertainties associated to the measurements.

The document is organized as follows. In Section 2 the process of event simulation and the employed MC simulations are described, the event selection is outlined in Section 3, and the sources of systematic uncertainties are discussed in Section 4. The extraction of the top quark mass from the ρ_s distribution is presented in Section 5. In Section 6, the calculation of the normalised differential cross section as function of ρ_s is described, and Section 7 details the top quark mass extraction from the cross section. A summary of the measurements is given in Section 8.

2 Event Simulation

In this analysis, the reference $t\bar{t}$ sample is simulated using the MADGRAPH event generator (v. 5.1.5.11) with the MADSPIN [8] package to account for spin correlation effects. The $t\bar{t}$ signal

was generated with up to three additional partons. The value of the top quark mass in the reference sample is chosen to be $m_t = 172.5$ GeV and the proton structure is described by the CTEQ6L1 [9] set of parton density functions (PDFs). The generated events are subsequently processed with PYTHIA (v. 6.424) [10] for fragmentation and hadronization using the MLM prescription for the matching of jets with parton showers [11]. The PYTHIA Z2* tune [12] is used to describe the underlying event. The CMS detector response is simulated using GEANT4 (v. 9.4) [13]. In addition to the reference sample, MADGRAPH+PYTHIA samples generated with m_t values of 166.5, 169.5, 171.5, 173.5, 175.5, and 178.5 GeV are used to evaluate the dependence of the measurement on the top quark mass and to extract the result.

Standard model backgrounds are simulated with MADGRAPH, POWHEG or PYTHIA, depending on the process. The main background contributions stem from Z/γ^* (referred to as Drell–Yan, DY in the following), single top quark (tW-channel) and W-boson production with additional jets (W+jets in the following). Smaller background contributions arise from diboson (WW, WZ and ZZ), $t\bar{t}$ production in association with a Z, W, or γ boson (referred to as $t\bar{t}$+Z/W/γ in the following), and QCD multijet events. For comparison with the measured distributions, the events in the simulated samples are normalised to an integrated luminosity of 19.7 fb^{-1} according to their cross section predictions. They are taken from NNLO (W+jets and DY), NLO+NNLL (single top quark tW-channel [14]), NLO (diboson [15], $t\bar{t}$+W [16]) and LO (QCD multijet [10]) calculations. The predicted cross section for the $t\bar{t}$+γ sample is obtained by scaling the LO cross section as obtained with the WHIZARD event generator [17] with an NLO *k*-factor [18].

The top quark mass is extracted through a fit to the data using the predicted normalised differential $t\bar{t}$+jet cross section from a NLO calculation combined with parton showering. The samples are generated using POWHEG (POWHEGBOX ttJ) with the PYTHIA (v. 8.205) 4C tune [19] for hadronization, and the proton structure is described by the CT10 PDF set. The values of the top quark mass for the different samples are $m_t = 172.5$ GeV and ± 1 GeV, ± 3 GeV, ± 6 GeV, and an additional sample is provided at $m_t = 163.5$ GeV.

3 Event Selection

The events are reconstructed using a particle-flow technique in which signals from all subdetectors are combined [20]. Charged hadron candidates from pileup events, i.e. originating from a vertex other than the one of the hard interaction, are removed before jet clustering on an event by event basis. Subsequently, the remaining component of neutral particle candidates from pileup events is evaluated [21]. Electron candidates are reconstructed from a combination of their charged track and their energy deposition in the ECAL, while electrons from identified photon conversions are rejected. Muon candidates are reconstructed from tracks which can be linked to both the silicon tracker and the muon system. The leptons, both electrons and muons, are required to have a transverse momentum of $p_T > 20$ GeV within the pseudorapidity region $|\eta| < 2.4$, and to be isolated with $I_{rel} < 0.15$. I_{rel} is defined as the sum of the transverse momenta of all neutral and charged reconstructed particle candidates, excluding the lepton itself, inside a cone of $\Delta R \equiv \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} < 0.3$ in the $\eta - \phi$ space, divided by the transverse momentum of the lepton under consideration. Jets are reconstructed by clustering the particle-flow candidates [22] using the anti-k_T clustering algorithm with size parameter $R = 0.5$ [23]. Muons and electrons passing less stringent selection criteria compared to the ones mentioned above have been identified and are excluded from the clustering process. Jets are selected in the pseudorapidity interval $|\eta| < 2.4$ and with a requirement of $p_T > 30$ GeV. Jets originating from bottom quarks are identified using combined secondary vertex and track-based lifetime

information. The b-tagging efficiency for the working point chosen for this analysis is about 80–85% and the mis-tag rate around 10% [24]. The missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ is defined as the magnitude of the imbalance of the transverse momentum of all reconstructed particle candidates [25].

Events are selected if they contain at least two isolated leptons (electrons or muons) of opposite charge and two jets of which at least one is identified as b jet. These events are triggered using combinations of two leptons fulfiling transverse momentum thresholds and isolation criteria. Events with a lepton pair invariant mass smaller than 20 GeV are removed in order to suppress events from heavy flavour resonance decays and DY processes. In the $\mu\mu$ and ee channels, the dilepton invariant mass is required to be outside a Z-boson mass window of 91 ± 15 GeV and $E_{\mathrm{T}}^{\mathrm{miss}}$ is required to be larger than 40 GeV.

A kinematic reconstruction method [26] is used to determine the $\mathrm{t\bar{t}}$ kinematic properties and to identify the two b jets originating from the decays of the top quarks. This allows to unambigously assign the additional jets in the event. This method has previously been used in [27, 28]. In the kinematic reconstruction the following constraints are imposed: the balance of the transverse momentum of the two neutrinos; the W-boson invariant mass of 80.4 GeV; and the equality of the top and antitop quark masses. The remaining ambiguities are resolved by prioritising those event solutions with two or one b-tagged jets over solutions using jets without b tags. The top quark mass can be experimentally reconstructed in a broad range due to resolution effects. In order to account for this in the reconstruction, the assumed top quark mass for each lepton-jet combination is varied between 100 GeV and 300 GeV in steps of 1 GeV. Among the physical solutions, the solution of highest priority according to the aforementioned criteria and with the most probable neutrino energies according to a simulated neutrino energy spectrum is chosen. The efficiency of the method is about 90% and only events with a valid solution for the kinematic reconstruction are used for the subsequent analysis.

Events are finally selected if they contain at least one additional jet which has not been selected by the kinematic reconstruction and features a $p_T > 50$ GeV within $|\eta| < 2.4$, following the criteria described in [3].

Dominant backgrounds to the $\mathrm{e^+e^-}$ and $\mu^+\mu^-$ channels originate from DY+jets processes. Their contribution is estimated from data following the procedure described in [27, 29] using events rejected by the Z-boson veto. The contributions are scaled by the ratio of events failing and passing the selection criteria as estimated from simulation ($R_{\mathrm{out/in}}$) [29]. The remaining backgrounds, including tW, W+jets, diboson, and QCD multijet events, are estimated from simulation.

Figure 1 shows the invariant mass of the $\mathrm{t\bar{t}}$ system, the transverse momentum of the leading additional jet, and the ρ_s observable with $m_0 = 170$ GeV [3] for the dilepton event sample. The distributions are compared to the standard model predictions. Only $\mathrm{t\bar{t}}$ events with two leptons (electron or muon) in the final state are considered as signal. All other $\mathrm{t\bar{t}}$ events, specifically the dominant contribution originating from decays via τ leptons, are considered as background.

4 Systematic Uncertainties

Systematic uncertainties in the measurement arise from detector effects, background modelling, and theoretical assumptions. Each systematic uncertainty is investigated separately and estimated for each bin of the measurement by varying the corresponding efficiency, resolution, or scale within its uncertainty. The different sources of systematic uncertainty described in the

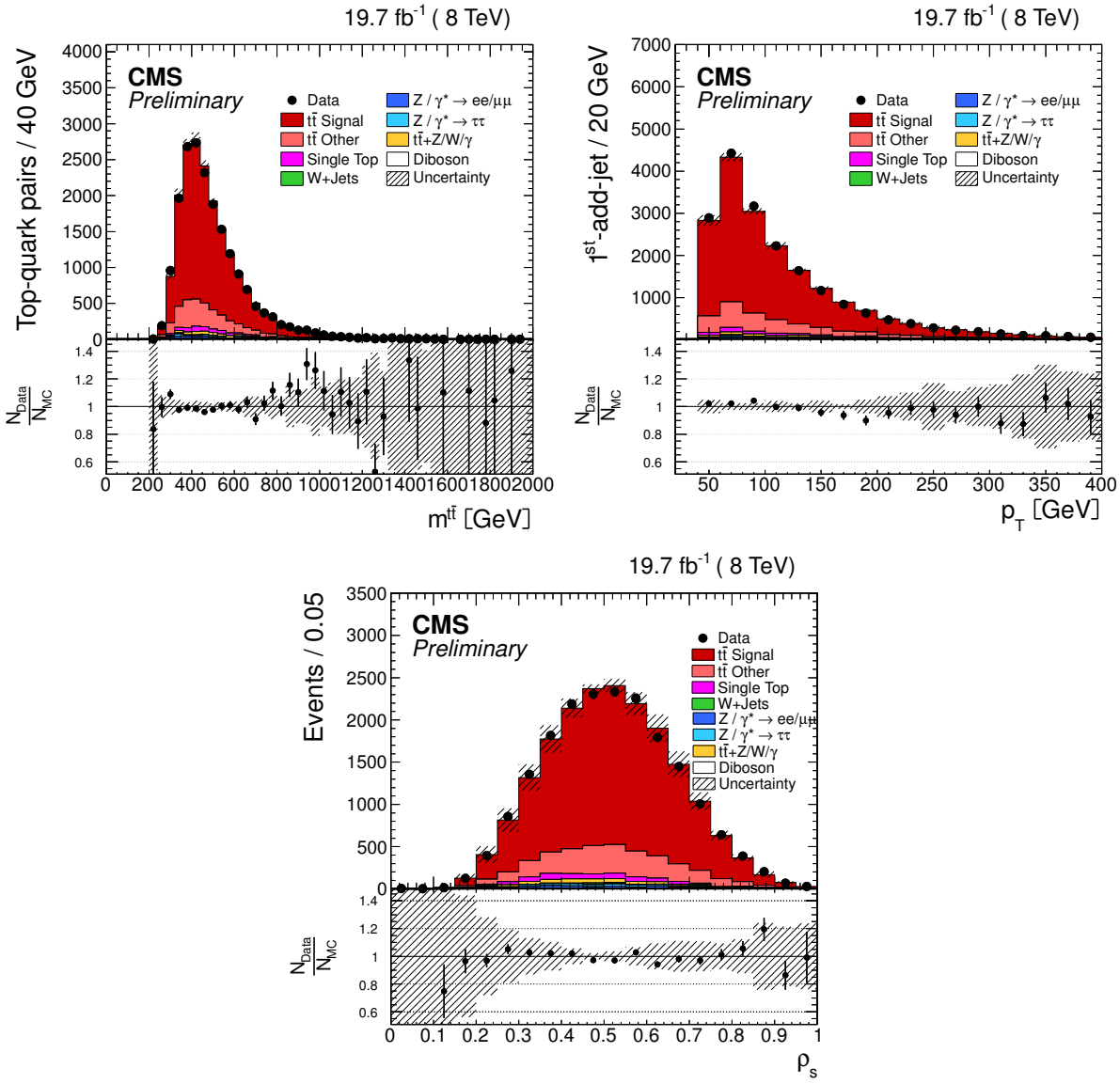

Figure 1: Invariant mass of the $t\bar{t}$ system (top left), transverse momentum of the leading additional jet (top right) and ρ_s (bottom) at reconstruction level for the combined dilepton channel. The $t\bar{t}$ sample is simulated using MADGRAPH and is assuming a top mass of $m_t = 172.5\,\text{GeV}$. The label "$t\bar{t}$ signal" refers to the events decaying dileptonically, while "$t\bar{t}$ other" refers to the other decay modes including $t\bar{t}$ decays into prompt τ-leptons. The hatched regions correspond to all shape uncertainties of the simulation (cf. Section 5).

following are assumed to be uncorrelated.

Experimental Uncertainties

The experimental sources considered are jet energy scale (JES) and jet energy resolution (JER), background normalisation, b-tag efficiency, pileup modeling and the kinematic reconstruction. The experimental uncertainty on the JES [30] is determined by varying the reconstructed energy scale as a function of the transverse momentum and the pseudorapidity of the jet, typically by a few percent. The uncertainty on the JER [31] is estimated by varying the simulated JER within its uncertainty for different pseudorapidity regions.

The impact of the background normalisation is obtained by scaling the background contributions by $\pm 30\%$ [29, 32].

The shape uncertainty on the b-tag efficiency is determined by dividing the b jet distributions for transverse momentum and pseudorapidity into two bins at their respective median value. The b-tag scale factors for the b jets in the first bin are scaled up by half of the assigned uncertainties [24], while those in the second bin are scaled down and vice versa. The difference between the scale factors in the two bins amounts to the full uncertainty.

The effect of pileup events is evaluated by weighting the simulation to the minimum bias cross section determined from data. The pileup model estimates the mean number of additional pp interactions to be about 20 events for the data analysed, based on the total inelastic proton-proton cross section which is determined to be 69.4 mb [33]. The systematic uncertainty is evaluated by varying this cross section within its uncertainty of $\pm 5\%$.

Other uncertainties taken into account stem from lepton trigger and identification efficiencies and the kinematic reconstruction efficiency, which have a negligible impact on the normalised measurements presented. The dependence of the result on the MC top quark mass used as reference for the measurement is found to be negligible.

Modelling Uncertainties

Modelling uncertainties originating from theoretical assumptions on the renormalisation and factorization scales, the jet-parton matching threshold, the hadronization model, the colour reconnection [34] and the underlying event (UE) modelling are determined by repeating the analysis, replacing the reference MADGRAPH signal simulation by dedicated simulation samples.

In particular, the impact of the first source is assessed with MADGRAPH samples with the renormalisation and factorisation scales simultaneously varied from the nominal values of μ_R and μ_F, which are defined by the Q^2 scale in the event as $\mu_R^2 = \mu_F^2 = Q^2 = m_t^2 + \sum p_T^2(\text{jet})$, where the sum runs over all the additional jets in the event not stemming from the $t\bar{t}$ decay. The samples with varied scales use $\mu_R = \mu_F = 4Q^2$ and $Q^2/4$, respectively. For the nominal MADGRAPH sample, a jet-parton matching threshold of 20 GeV is chosen, while for the varied samples, values of 40 and 10 GeV are employed, respectively.

The UE modeling is evaluated by comparing the two different Perugia 11 (P11) PYTHIA tunes `mpHi` and `TeV` to the standard P11 tune.

The uncertainties from ambiguities in modeling colour reconnection effects are estimated by comparing simulations of an underlying event tune including colour reconnection to a tune without the effect (the P11 and P11 `noCR` tunes described in [35]).

The hadronisation model uncertainty is estimated by comparing samples simulated with POWHEG

and MC@NLO, using PYTHIA and HERWIG, respectively, for hadronisation.

The uncertainty arising from the PDFs is assessed by reweighting the $t\bar{t}$ signal sample according to the CT10 error PDF sets at 68% confidence level. The effects of these variations are added in quadrature.

5 Extraction of m_t from ρ_s using a Template Technique

The top quark mass value is determined by comparison of the experimentally observed yields in individual bins of the normalized ρ_s distribution as presented in Fig. 1 with the predicted yields for different values of m_t. Each of the $t\bar{t}$ samples employed for the comparison are normalized such that the total number of expected events including backgrounds corresponds to the events selected in data.

The most probable value of top quark mass is extracted by means of a χ^2_i distribution for each bin i. For each of them, the predicted yields for the different top quark masses are fitted using a second order polynomial function, $f(m_t)$, in order to obtain a continuous distribution as a function of m_t. The χ^2_i distributions are then calculated as

$$\chi^2_i(m_t) = \frac{(N_i^{\text{data}} - f_i^{\text{sim}}(m_t))^2}{(\delta N_i^{\text{data}})^2 + (\delta f_i^{\text{sim}}(m_t))^2}\,, \tag{1}$$

where $f_i^{sim}(m_t)$ represents the polynomial fit to the simulation distributions, δf_i^{sim} corresponds to the statistical errors on the simulation obtained from the confidence interval of the fit, and N_i^{data} is the number of selected events in each bin with δN_i^{data} being the corresponding statistical uncertainty. The χ^2_i does not include the systematic uncertainties, which are treated as external variations as described below. The binning chosen corresponds to the one in Fig. 1 except for the regions $\rho_s < 0.2$ and $\rho_s > 0.85$ which are integrated into two bins in order to increase statistics.

The global χ^2 is calculated by summing all bins to $\chi^2 = \sum_i \chi^2_i(m_t)$ since the individual bins are statistically uncorrelated. The number of degrees of freedom is reduced by one owing to the normalisation of the utilized distributions. In order to accommodate for this in the global χ^2, the bin with the lowest statistical significance is removed from the calculation, i.e., the first bin of the ρ_s distribution. The resulting χ^2 distribution is presented in Fig. 2. The minimum of the distribution is taken as the measured top quark mass value, the statistical uncertainty is calculated as the $\pm 1\sigma$ deviation from the minimum by applying a $\chi^2 + 1$ variation.

The systematic uncertainties described in Section 4 are propagated to the mass measurement by repeating the extraction with each varied input distributions, either by replacing the reference simulation with the dedicated samples, or by varying the corresponding efficiency, resolution, or scale within the respective uncertainties. For each source of uncertainty, the altered simulation is used as input for the χ^2 distribution from which the top quark mass is obtained. The difference between the mass extracted for each source and the nominal measurement is taken as the systematic uncertainty. In the case of the model uncertainties, the reference MADGRAPH $t\bar{t}$ sample with nominal top quark mass is replaced by the corresponding systematic samples, for instance P11 or varied jet-parton matching sample. The same model uncertainty is then assigned to all varied mass MADGRAPH samples.

The breakdown of the systematic uncertainties for the top quark mass is presented in Table 1. All systematic uncertainties quoted are statistically significant, the statistical uncertainties are

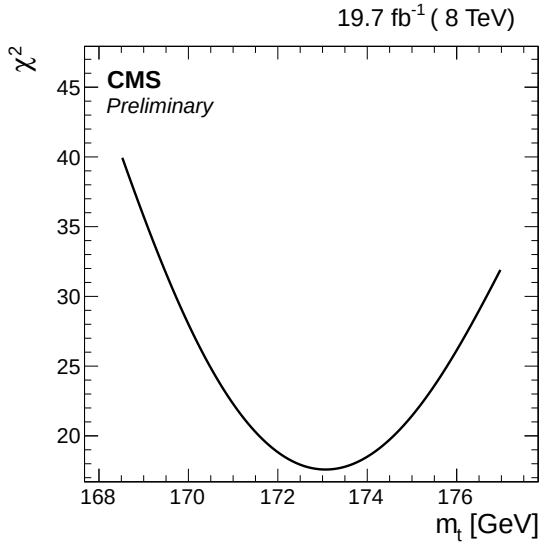

Figure 2: Global χ^2 distribution obtained as the sum of the χ^2_i distributions of the individual bins for the dilepton combined channel. The most probable top quark mass is extracted from the minimum, the statistical uncertainty from a $\chi^2 + 1$ variation around the minimum.

found to be below 0.1 GeV. The size of the uncertainties is in agreement with the expected values from the blinded analysis, performed using simulated samples as pseudo-data.

The total uncertainty amounts to 2.2 GeV and is calculated as the quadratic sum of all individual contributions. It is dominated by the renormalisation and factorization scales and jet-parton matching uncertainties, while the most relevant experimental contribution arises from uncertainties in the b-tagging. The top quark mass results to $m_t = 173.1 \pm 1.0\,(\text{stat})\,^{+2.0}_{-2.3}\,(\text{syst})$ GeV, including both the statistical uncertainty of data and the MADGRAPH samples used to extract the mass.

In order to calibrate the mass extraction technique, the measurement has been repeated with pseudo data generated from the simulated samples with different top quark masses. The measured top quark masses are compared to the true mass used to generate the input samples as shown in Fig. 3. The measured top quark mass agrees with the input mass within the statistical uncertainty, which indicates that the method does not favour certain values of the mass or introduces a bias toward higher or lower mass values. Propagating the uncertainty of a linear regression to the measured top quark mass would yield an effect of below 0.1 GeV for a mass shift of 1 GeV and can thus be neglected.

6 Differential Cross Section Measurement

The normalized differential $\bar{t}t$ production cross section as a function of ρ_s is measured from the event yields determined in Section 3. In order to avoid additional uncertainties due to the extrapolation of the measurement outside of the phase space region probed experimentally, the differential cross section is determined in a visible phase space defined at particle level by the kinematic and geometrical acceptance of the final-state leptons and jets. The charged leptons from the $\bar{t}t$ decay are selected with $|\eta| < 2.4$ and $p_T > 20$ GeV. A jet is defined at particle level by applying the anti-k_t clustering algorithm to all stable particles except leptons and neutrinos stemming from the W boson decays. A jet is defined as b jet if it has at least one b hadron

Table 1: Breakdown of the systematic uncertainties for the top quark mass extracted from the normalized event yield for the dilepton combined channels. All systematic uncertainties are found to be statistically significant. For the asymmetric uncertainties due to scale variations, the first reported value corresponds to an increase of the corresponding scale and the second one to a decrease.

Source	Δm_t [GeV]
Jet-Parton Matching	$^{-1.3}_{+0.1}$
Q^2 Scale	± 2.0
ME/Showering	$^{+0.2}_{-0.3}$
Color Reconnection	< 0.1
Underlying Event	± 0.1
PDF	± 0.1
Background	± 0.4
Jet Energy Scale	± 0.1
Jet Energy Resolution	± 0.1
Pile-Up	± 0.1
Trigger Eff.	< 0.1
Kinematic Reconstruction	< 0.1
Lepton Eff.	< 0.1
B-Tagging	± 0.2
Total syst.	$^{+2.0}_{-2.3}$
Stat.	± 1.0
Total unc.	$^{+2.2}_{-2.5}$

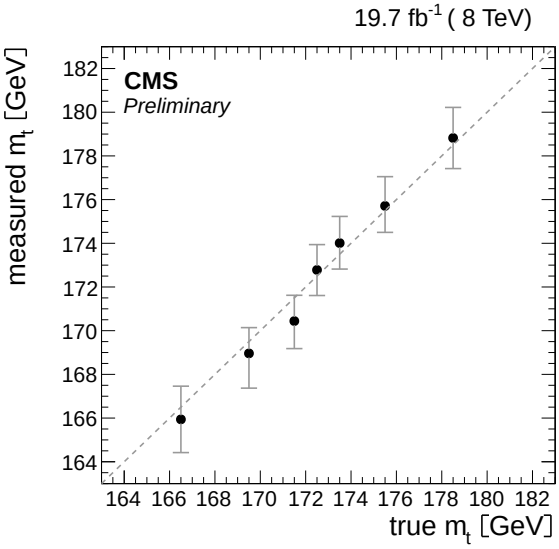

Figure 3: Top quark mass obtained from pseudo data generated from each of the MADGRAPH samples with varied mass values. Neither a favoured mass value nor a bias towards higher or lower masses can be observed.

associated to it, which can be assigned to the corresponding original b quark. The two b jets from the $t\bar{t}$ decay have to fulfill the kinematic requirements $|\eta| < 2.4$ and $p_T > 30\,\text{GeV}$, while the additional jet is required to have $|\eta| < 2.4$ and $p_T > 50\,\text{GeV}$.

The differential cross section in each bin is defined as

$$\frac{1}{\sigma}\frac{d\sigma^i}{d\rho_s} = \frac{1}{\sigma}\frac{\sum_j A_{ij}^{-1}(N^j_{\text{data}} - N^j_{\text{non}-t\bar{t}\,\text{bkg}}) \cdot f^{\text{signal}}}{\Delta^i_x \mathcal{L}}, \tag{2}$$

where j indicates the bin index of the reconstructed variable x, while i denotes the index of the corresponding generator-level bin. N^j_{data} and $N^j_{\text{non}-t\bar{t}\,\text{bkg}}$ represent the number of data events and estimated background events from processes different than $t\bar{t}$ in bin j, respectively, while $\mathcal{L}$ is the integrated luminosity and Δ^i_x denotes the bin width. The contribution from non-signal $t\bar{t}$ decays is taken into account by correcting $N_{\text{data}} - N_{\text{non}-t\bar{t}\,\text{bkg}}$ with the signal fraction, f^{signal}, defined as the ratio of selected $t\bar{t}$ signal events to the total number of selected $t\bar{t}$ events. This avoids the dependence on the inclusive $t\bar{t}$ cross section used for normalisation. The distribution is normalized by the sum of the differential cross section per bin σ^i over all bins, $\sigma = \sum \sigma^i$.

Effects from detector efficiency and resolution in each bin i of the measurement are corrected are corrected using a regularized unfolding method [27, 36, 37]. The generalized inverse of the response matrix, denoted as A^{-1}_{ij}, is used to obtain the unfolded values from the measured distribution. In order to avoid nonphysical fluctuations, a smoothing prescription (regularization) is applied. The regularization level is determined using the averaged global correlation method [38].

The binning was chosen such that the purity and stability are above 40% and the sensitivity of the distribution on the top quark mass is enhanced. The purity p^i denotes the number of events generated and correctly reconstructed in a given bin i relative to the number of events that are reconstructed in bin i but generated anywhere. The stability s^i represents the number of events which are generated and correctly reconstructed in a given bin i relative to the number of events

that are generated in bin i but reconstructed anywhere. In particular for this observable, p^i is between 40% and 60% for all bins, while s^i ranges between 60% to 80%.

All systematic sources described in Section 4 are propagated to the measured differential cross section by replacing the nominal simulation samples used for the unfolding matrix, determination of the background etc. by the samples with the systematic effect applied. For each systematic variation the differential cross section is evaluated. In each bin, the systematic variations are added in quadrature. Due to the normalisation, systematic uncertainties which are correlated across all bins of the measurement, such as the uncertainty on the integrated luminosity as well as all other normalisation uncertainties, cancel out. The final result is obtained as the weighted average of the normalized differential cross sections measured in each of the individual dileptonic channels.

The simulation used to extract the cross sections is the reference sample with a top quark mass of 172.5 GeV. Figure 4 shows the unfolded distribution together with the predicted values for different top quark masses ($\pm$1, 3 and 6 GeV from the nominal value). The total systematic uncertainty is displayed as yellow band, while the statistical uncertainty is represented as grey band. The distribution shows a discriminating power between the different top quark masses, except for the region $\rho_s = 0.5$ where the curves intersect due to the normalisation. The distribution is most sensitive to mass variations in the regions around $\rho_s = 0.4$ and $\rho_s \geq 0.6$ as expected according to the reference [3].

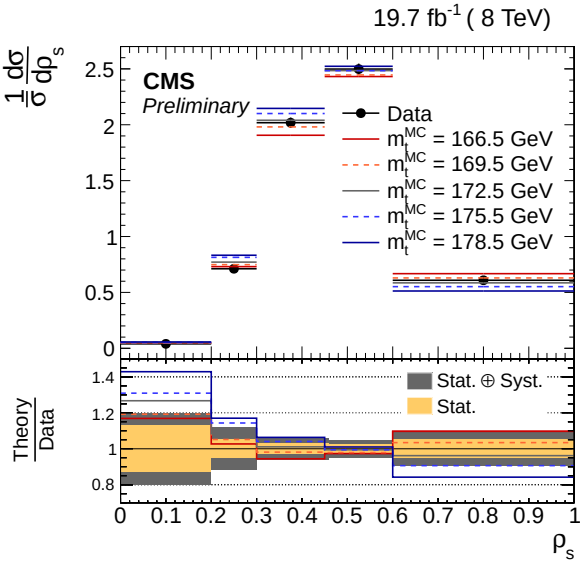

Figure 4: Normalized differential $t\bar{t}$ cross section in the visible phase space after unfolding as a function of the observable ρ_s in the dilepton channels, compared to the predictions from POWHEG $t\bar{t}$+jet simulated with a top quark mass of 172.5 GeV as well as $\pm$3 and 6 GeV variations with respect to the central value. The grey band represents the statistical uncertainty, the yellow band corresponds to the total systematic uncertainty.

7 Extraction of m_t from the Differential Cross Section

The unfolded differential cross section obtained in Section 6 is used to measure the top quark mass by comparing it to theoretical predictions of the ρ_s distribution. Following the approach described in Section 5, the measured differential cross section is compared to the predicted

cross sections for each bin of the ρ_s distribution using different top quark masses as shown in Fig. 5, and the most probable top quark mass is extracted from a global χ^2 estimator.

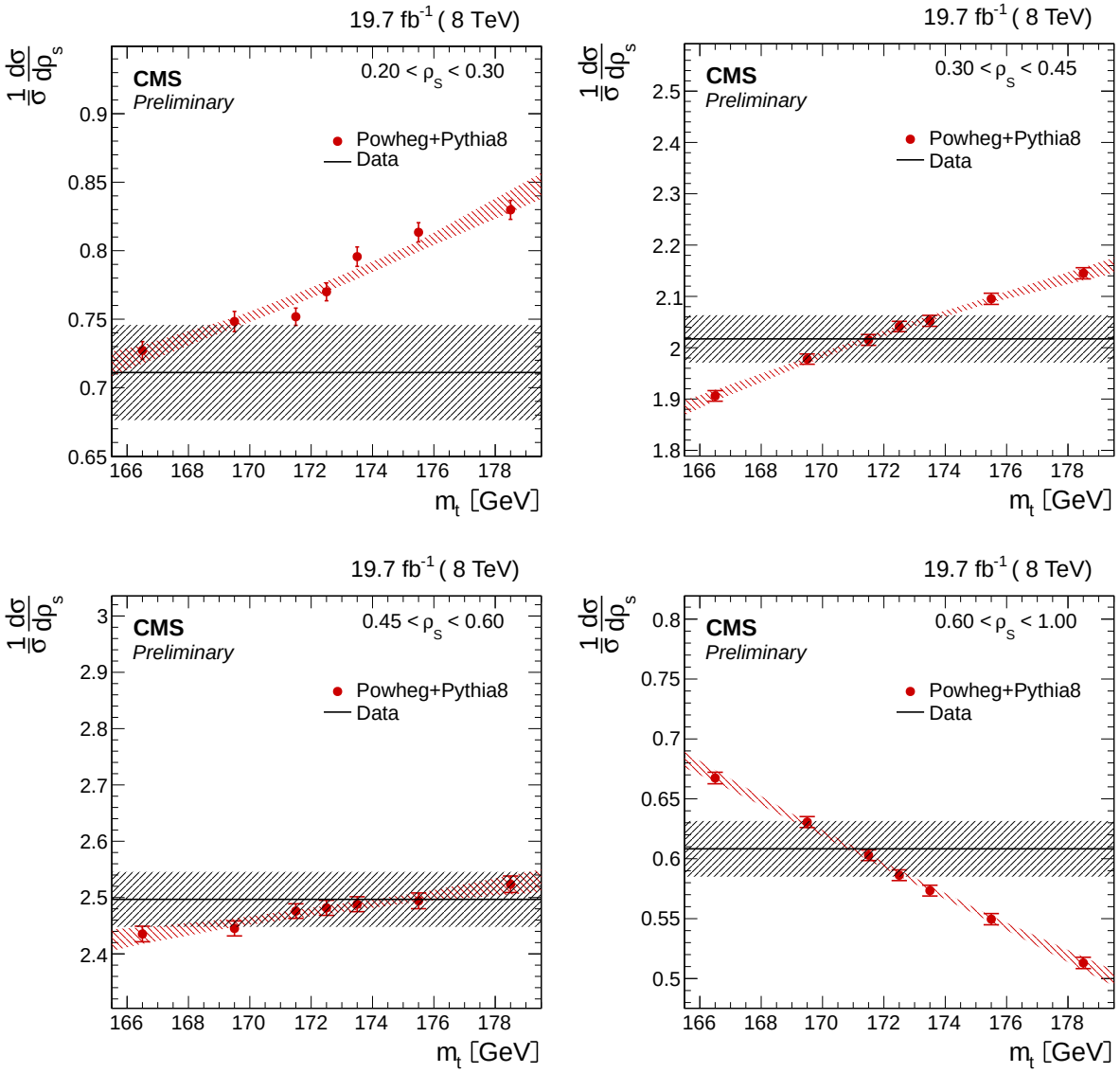

Figure 5: Distributions of the differential cross section for simulation and data for all mass samples in the different bins of the ρ_s distribution, shown for the three dilepton final states combined. The error bands correspond to the statistical error on data and the confidence interval of the second order polynomial for the simulation.

However, owing to the unfolding procedure the individual bins of the ρ_s cross section distribution are correlated and these bin-to-bin correlations need to be taken into account in the global χ^2 distribution via the covariance matrices obtained from the unfolding procedure. Since the individual bins of the simulation samples are uncorrelated, the corresponding statistical uncertainties only contribute to the diagonal elements of the covariance matrix. The global χ^2 estimator is thus obtained as

$$\chi^2(m_t) = V^T(m_t) \times \boldsymbol{COV}(m_t)^{-1} \times V(m_t), \tag{3}$$

where $\boldsymbol{COV}^{-1}$ denotes the inverted and normalized unfolding covariance matrix including the statistical uncertainties from simulation. The vector $V_i(m_t) = f_i^{\text{data}}(m_t) - f_i^{\text{sim}}(m_t)$ represents the difference between the measured cross section and the cross section function obtained by a second order polynomial fit to the simulations with different top quark masses.

The first bin ($0 < \rho_s < 0.2$) is excluded from the global χ^2 in order to satisfy the reduced number of degrees of freedom from the normalisation of the differential cross section. The resulting global χ^2 distribution is presented in Fig. 6. The minimum of the global χ^2 corresponds to the most probable top quark mass with the statistical uncertainty taken as the $\pm 1\,\sigma$ deviation from the minimum by applying a $\chi^2 + 1$ variation.

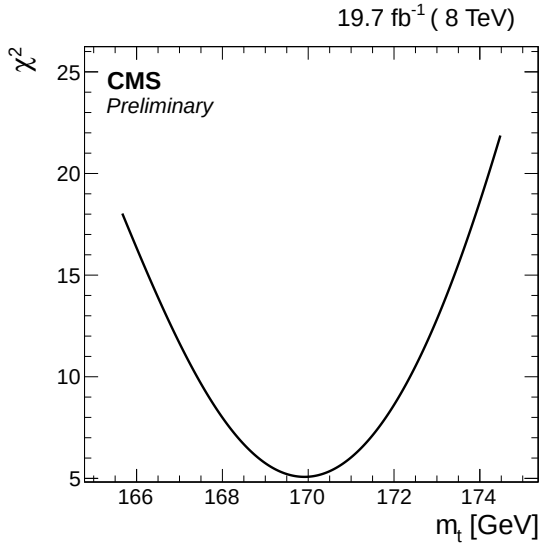

Figure 6: Global χ^2 distribution for the normalized differential cross section as a function of ρ_s in the dilepton combined channel.

Possible biases of the mass extraction have been studied by repeating the measurement using each of the available MADGRAPH samples with different top quark masses as pseudo data, and by confronting the measurement with the MC truth, following the approach presented in Section 5. The values of the top quark mass obtained from the global χ^2 distribution are compared to the mass of the sample used as input as shown Fig. 7. A good agreement within the statistical uncertainty can be observed, indicating that the method is unbiased. Considering a variation of the top quark mass of 1 GeV, the effect would correspond to a ± 0.05 GeV deviation in the final result, which is well below the statistical uncertainty of the measurement. The dependence of the result on the MC top quark mass used to evaluate the background contribution and to unfold the distribution is found to be negligible.

The systematic uncertainties are propagated to the mass measurement as follows. For each uncertainty source the normalized differential cross section is evaluated replacing the simulated samples by the corresponding variations as well as changing the correlation matrix for unfolding accordingly. For each of the obtained cross sections, the χ^2 distribution is calculated and the top quark mass is extracted. The difference between the mass obtained for each source of uncertainty and the nominal mass value is taken as systematic uncertainty. This difference is found to be statistically significant for each source of systematic uncertainty considered.

Additionally, the uncertainty in modelling of the hard-production process is assessed through a simultaneous variation of the renormalisation and factorisation scales in the POWHEG $\bar{\text{t}}$t+jet

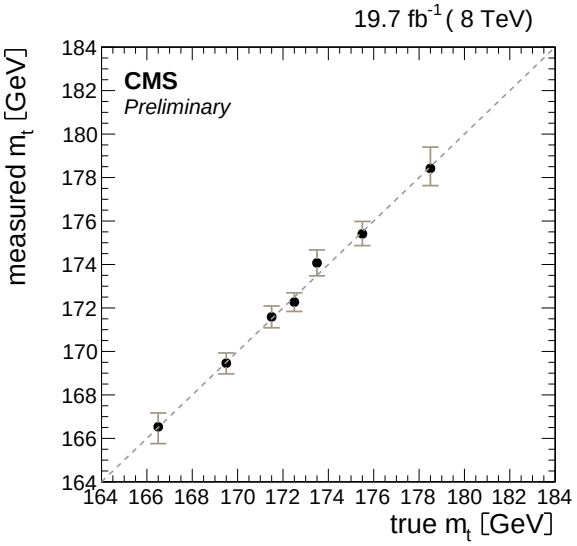

Figure 7: Top quark mass obtained by using each of the MADGRAPH samples with varied mass values as pseudo data, for the dilepton combined channel.

simulation ("POWHEG $t\bar{t}$+jet modelling") with respect to their common nominal value, which is set to m_t. For each top quark mass value considered in the mass measurement, two dedicated samples have been produced with $\mu_R = \mu_F$ varied coherently by factors of 2 and 0.5. The systematic uncertainty is then evaluated by repeating the mass measurement using the dedicated samples as theoretical prediction.

The breakdown of systematic uncertainties for the top quark mass is shown in Table 2. The first row represents the POWHEG $t\bar{t}$+jet modelling uncertainty while the other uncertainties are related to the measurement of the cross section. The total systematic uncertainty is dominated by model uncertainty sources, in particular, jet-parton matching and Q^2 and theoretical uncertainties in the POWHEG $t\bar{t}$+jet simulation. The most relevant experimental uncertainties arise from the background and JES sources.

The top quark mass obtained is $169.9 \pm 1.1\,(\text{stat})\,^{+2.5}_{-3.1}\,(\text{syst})\,^{+3.6}_{-1.6}\,(\text{theo})$ GeV, where the systematic uncertainty is calculated as the quadratic sum of all sources. The systematic uncertainty corresponds to $^{+1.5}_{-1.8}$ %, while the theoretical uncertainty on the modelling assumptions arising from the POWHEG $t\bar{t}$+jet simulations yields an additional $^{+2.1}_{-0.9}$ %.

8 Summary

The top quark mass is measured from the inverse of the invariant mass of the $t\bar{t}$+jet system, an observable proposed in [3]. The mass extraction has been performed with a global template fit using the shape of the distribution at reconstruction level as well as using the normalised differential cross section in the visible phase space. The first approach avoids statistical correlations and uncertainties arising from the unfolding procedure, however it cannot be compared to predictions given at generator level, while the second eases the comparisons with theory models.

The top quark mass obtained from the normalized differential $t\bar{t}$+jet cross section using an NLO calculation interfaced with parton shower yields $169.9 \pm 1.1\,(\text{stat})\,^{+2.5}_{-3.1}\,(\text{syst})\,^{+3.6}_{-1.6}\,(\text{theo})$ GeV. The

Table 2: Breakdown of the systematic uncertainties for the top quark mass measured from the dileptonic channel. All systematic uncertainties are found to be statistically significant. For the asymmetric uncertainties due to scale variations, the first reported value corresponds to an increase of the corresponding scale and the second one to a decrease.

Source	Δm_t [GeV]
POWHEG $t\bar{t}$+jet modelling	$^{-1.6}_{+3.6}$
Jet-Parton Matching	$^{-0.1}_{+1.6}$
Q^2 Scale	$^{+1.0}_{-2.8}$
ME/Showering	± 0.4
Color Reconnection	± 0.7
Underlying Event	± 0.3
PDF	$^{+0.9}_{-0.1}$
Background	± 1.0
Jet Energy Scale	± 0.1
Jet Energy Resolution	± 0.1
Pile-Up	± 0.3
Trigger Eff.	< 0.1
Kinematic Reconstruction	< 0.1
Lepton Eff.	± 0.1
B-Tagging	± 0.3
Syst. uncertainty	$^{+2.5}_{-3.1}$
Stat. untertainty	± 1.1

precision is mostly limited by the systematic uncertainties arising from modelling sources and the theory uncertainties in the POWHEG $t\bar{t}$+jet simulation. The result is in agreement within the uncertainties with other measurements performed following the same approach [4] as well as complementary measurements of the mass from the inclusive $t\bar{t}$ production cross section [39–41].

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] CMS Collaboration, "Measurement of the top quark mass using proton-proton data at $\sqrt{s}$ = 7 and 8 TeV", arXiv:1509.04044.

[3] S. Alioli et al., "A new observable to measure the top-quark mass at hadron colliders", *Eur.Phys.J.* **C73** (2013) 2438, doi:10.1140/epjc/s10052-013-2438-2, arXiv:1303.6415.

[4] ATLAS Collaboration, "Determination of the top-quark pole mass using $t\bar{t}$ + 1-jet events collected with the ATLAS experiment in 7 TeV pp collisions", *JHEP* **10** (2015) 121, doi:10.1007/JHEP10(2015)121, arXiv:1507.01769.

[5] CMS Collaboration, "The CMS experiment at the CERN LHC", *JINST* **03** (2008) S08004, doi:10.1088/1748-0221/3/08/S08004.

[6] CMS Collaboration, "Measurement of Top Quark Pair Differential Cross Sections at sqrt(s) = 8 TeV", *CMS Physics Analysis Summary* **TOP-12-028** (2013).

[7] CMS Collaboration, "Measurement of the Jet Multiplicity in dileptonic Top Quark Pair Events at 8 TeV", Technical Report CMS-PAS-TOP-12-041, CERN, Geneva, 2013.

[8] P. Artoisenet et al., "Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations", *JHEP* **03** (2013) 015, doi:10.1007/JHEP03(2013)015, arXiv:1212.3460.

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

[10] 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.

[11] M. L. Mangano, M. Moretti, F. Piccinini, and M. Treccani, "Matching matrix elements and shower evolution for top-quark production in hadronic collisions", *JHEP* **0701** (2007) 013, doi:10.1088/1126-6708/2007/01/013, arXiv:hep-ph/0611129.

[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] S. Agostinelli et al., "GEANT4– a simulation toolkit", *Nucl. Instrum. Meth. A* **506** (2003) 250, doi:10.1016/S0168-9002(03)01368-8.

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

[15] J. M. Campbell, R. K. Ellis, and C. Williams, "Vector boson pair production at the LHC", *JHEP* **1107** (2011) 018, doi:10.1007/JHEP07(2011)018, arXiv:1105.0020.

[16] J. Campbell and R. Ellis, "$t\bar{t}W^{+-}$ production and decay at NLO", *JHEP* **07** (2012) 052, doi:10.007/JHEP07(2012)052, arXiv:1204.5678.

[17] W. Kilian, T. Ohl, and J. Reuter, "WHIZARD: Simulating multi-particle processes at LHC and ILC", *Eur. Phys. J. C* **71** (2011) 1742, doi:10.1140/epjc/s10052-011-1742-y, arXiv:hep-ph/9905386.

[18] K. Melnikov, M. Schulze, and A. Scharf, "QCD corrections to top quark pair production in association with a photon at hadron colliders", *Phys. Rev. D* **83** (2011) 074013, doi:10.1103/epjc/PhysRevD.83.074013, arXiv:hep-ph/1102.1967.

[19] R. Corke and T. Sjostrand, "Interleaved Parton Showers and Tuning Prospects", *JHEP* **03** (2011) 032, doi:10.1007/JHEP03(2011)032, arXiv:1011.1759.

[20] CMS Collaboration, "Commissioning of the Particle-Flow Reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV", *CMS Physics Analysis Summary* **PFT-10-002** (2010).

[21] M. Cacciari, G. P. Salam, and G. Soyez, "The Catchment Area of Jets", *JHEP* **04** (2008) 005, doi:10.1088/1126-6708/2008/04/005, arXiv:0802.1188.

[22] CMS Collaboration, "Determination of the jet energy scale in CMS with pp Collisions at $\sqrt{s} = 7$ TeV", *CMS Physics Analysis Summary* **JME-10-010** (2010).

[23] M. Cacciari, G. P. Salam, and G. Soyez, "The anti-k_t jet clustering algorithm", *JHEP* **04** (2008) 063, doi:10.1088/1126-6708/2008/04/063, arXiv:hep-ph/0802.1189.

[24] CMS Collaboration, "Identification of b-quark jets with the CMS experiment", *JINST* **08** (2013) P04013, doi:10.1088/1748-0221/8/04/P04013, arXiv:hep-ex/1211.4462.

[25] CMS Collaboration, "Missing transverse energy performance of the CMS detector", *JINST* **06** (2011) P09001, doi:10.1088/1748-0221/6/09/P09001, arXiv:1106.5048.

[26] D0 Collaboration, "Measurement of the top quark mass using dilepton events", *Phys. Rev. Lett.* **80** (1998) 2063, doi:10.1103/PhysRevLett.80.2063, arXiv:hep-ex/9706014.

[27] CMS Collaboration, "Measurement of differential top-quark pair production cross sections in pp collisions at $\sqrt{s} = 7$ TeV", *Eur. Phys. J. C* **73** (2013) 2339, doi:10.1140/epjc/s10052-013-2339-4, arXiv:hep-ph/1211.2220.

[28] CMS Collaboration, "Measurement of jet multiplicity distributions in $t\bar{t}$ production in pp collisions at $\sqrt{s} = 7$ TeV", *Eur. Phys. J. C* **74** (Apr, 2014) 3014, doi:10.1140/epjc/s10052-014-3014-0, arXiv:1404.3171.

[29] CMS Collaboration, "Measurement of the $t\bar{t}$ production cross section and the top quark mass in the dilepton channel in pp collisions at $\sqrt{s} = 7$ TeV", *JHEP* **07** (2011) 049, doi:10.1007/JHEP07(2011)049, arXiv:hep-ex/1105.5661.

[30] 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.

[31] CMS Collaboration, "Jet energy resolution in CMS at $\sqrt{s} = 7$ TeV", *CMS Physics Analysis Summary* **JME-10-014** (2010).

[32] CMS 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.

[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] P. Z. Skands and D. Wicke, "Non-perturbative QCD effects and the top mass at the Tevatron", *Eur. Phys. J. C* **52** (2007) 133, doi:10.1140/epjc/s10052-007-0352-1, arXiv:hep-ph/0703081.

[35] P. Z. Skands, "Tuning Monte Carlo generators: The Perugia tunes", *Phys. Rev. D* **82** (2010) 074018, doi:10.1103/PhysRevD.82.074018, arXiv:1005.3457.

[36] A. Hoecker and V. Kartvelishvili, "SVD approach to data unfolding", *Nucl. Instrum. Meth. A* **372** (1996) 469, doi:10.1016/0168-9002(95)01478-0, arXiv:hep-ph/9509307.

[37] V. Blobel, "An unfolding method for high energy physics experiments", (2002). arXiv:hep-ex/0208022.

[38] F. James, "Statistical methods in experimental physics". World Scientific, 2^{nd} edition, 2006.

[39] CMS Collaboration Collaboration, "Measurement of the ttbar production cross section in the emu channel in pp collisions at 7 and 8 TeV", Technical Report CMS-PAS-TOP-13-004, CERN, Geneva, 2015.

[40] R. Astalos et al., "Proceedings of the sixth international workshop on multiple partonic interactions at the Large Hadron Collider", (2015). arXiv:1506.05829.

[41] CMS Collaboration, "Determination of the top-quark pole mass and strong coupling constant from the tbart production cross section in pp collisions at $\sqrt{s}$ = 7 TeV", *Phys. Lett. B* **728** (2014) 496–517, doi:10.1016/j.physletb.2014.08.040,10.1016/j.physletb.2013.12.009, arXiv:1307.1907. [Erratum: Phys. Lett.B728,526(2014)].