Measurement of the top quark mass with lepton+jets final states in pp collisions at $\sqrt{s} = 13$ TeV

NATALIIA KOVALCHUK ON BEHALF OF THE CMS COLLABORATION[1]

University of Hamburg
Luruper Chaussee 149, 22761 Hamburg, Germany

The mass of the top quark is measured using a sample of $t\bar{t}$ candidate events with one lepton, muon or electron, and at least four jets in the final state, collected by the CMS detector in pp collisions at $\sqrt{s} = 13$ TeV at the CERN LHC. The candidate events are selected from data corresponding to an integrated luminosity of 35.9 fb^{-1}. For each event the mass is reconstructed from a kinematic fit of the decay products to a $t\bar{t}$ hypothesis. The top quark mass is determined simultaneously with an overall jet energy scale factor (JSF), constrained by the mass of the W boson in $q\bar{q}$ decays. The measurement is calibrated on samples simulated at next-to-leading order matched to parton shower. The top quark mass is found to be $172.25 \pm 0.08\,(\text{stat} + \text{JSF}) \pm 0.62\,(\text{syst})$ GeV. The dependence of this result on event kinematical properties is studied and compared to predictions of different models of $t\bar{t}$ production.

PRESENTED AT

10th International Workshop on Top Quark Physics
Braga, Portugal, September 17–22, 2017

---
[1]E-mail: nataliia.kovalchuk@desy.de

1 Introduction

The first top quark mass (m_t) measurement in lepton+jets final state using the full data sample collected in 2016 by the CMS detector [1] at a centre-of-mass energy $\sqrt{s} = 13\,\text{TeV}$ is presented [2], corresponding to an integrated luminosity of $35.9\pm0.9\,\text{fb}^{-1}$ [3]. The implementation and evaluation of the ideogram method [4], which is used for the m_t extraction utilizing the Monte Carlo (MC) simulated events, is performed, analogously to the mass extraction technique employed for the most precise CMS measurement [5] using Run 1 data.

Since the publication of the CMS Run 1 measurements, new versions of Monte Carlo generators have been integrated and a larger data set has been collected at $\sqrt{s} = 13\,\text{TeV}$. The new simulations allow more refined estimations of the modeling uncertainties, while the analysis techniques and final state are the same as in [5]. Additionally, the dependence of the measured mass value on the kinematic properties of the events is evaluated.

2 Data Samples and Event Selection

The measurement is performed by analyzing events in top quark pair production ($t\bar{t}$) with lepton+jets in the final state, $t\bar{t} \rightarrow \text{WbWb} \rightarrow (l\nu_l b)(q\bar{q}b)$, which combines both μ+jets and e+jets decay channels. Hence, events are required to pass a single-muon trigger with a minimum threshold on the transverse momentum (p_T) of an isolated muon of $24\,\text{GeV}$ or a single-electron trigger with a p_T threshold for isolated electrons of $32\,\text{GeV}$.

Simulated $t\bar{t}$ signal events are generated with POWHEG v2 [6, 7] and the PYTHIA 8.219 parton-shower generator [8] using the CUETP8M2T4 tune [9] for seven different top quark mass values: 166.5, 169.5, 171.5, 172.5, 173.5, 175.5, and $178.5\,\text{GeV}$.

Events with exactly one isolated muon with $p_T > 26\,\text{GeV}$ and $|\eta| < 2.4$ or exactly one isolated electron with $p_T > 34\,\text{GeV}$ and $|\eta| < 2.1$ are selected. In addition, at least four jets with $p_T > 30\,\text{GeV}$ and $|\eta| < 2.4$ are required. Additionally, exactly two jets originating from the b quarks are required, being clustered with the CSVv2 [10] algorithm that combines reconstructed secondary vertices and track-based lifetime information [11].

For the final selection, the events are reconstructed and the parton-jet assignments for simulated $t\bar{t}$ events are classified as *correct* (cp), *wrong* (wp), and *unmatched* (un) permutations, where, in the latter, at least one quark from the $t\bar{t}$ decay is not unambiguously matched with a distance of $\Delta R < 0.4$ in the η-ϕ space to any of the four selected jets. In order to check the compatibility of an event with the $t\bar{t}$ hypothesis and to improve the resolution of the reconstructed quantities, the characteristic topology is exploited by the kinematic fit [12]. The fit is repeated for every jet permutation and constrains the mass of two reconstructed W boson to be equal $80.4\,\text{GeV}$ and the

mass of t quark to be equal to the ones from $\bar{\mathrm{t}}$. In order to increase the fraction of correct permutations, the goodness-of-fit (gof) probability for the kinematic fit with two degrees of freedom $P_{\mathrm{gof}} = \exp\left(-\frac{1}{2}\chi^2\right)$ is required to be at least 0.2. Figure 1 shows the distributions of the reconstructed mass $m_{\mathrm{W}}^{\mathrm{reco}}$ of the W boson decaying to a $\mathrm{q}\overline{\mathrm{q}}'$ pair and the top quark mass from the kinematic fit $m_{\mathrm{t}}^{\mathrm{fit}}$ for all possible permutations. These two observables are used in the mass extraction.
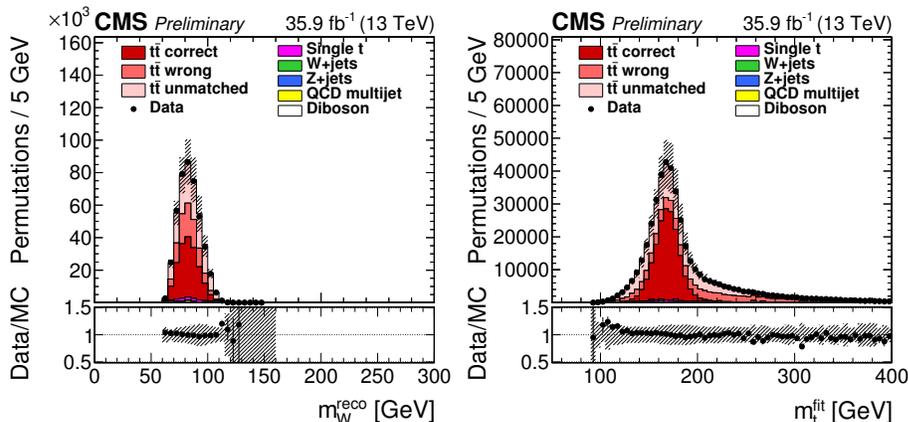

Figure 1: (left) The reconstructed W boson masses $m_{\mathrm{W}}^{\mathrm{reco}}$ and (right) the fitted top quark masses $m_{\mathrm{t}}^{\mathrm{fit}}$ after the goodness-of-fit selection and the weighting by P_{gof}. The vertical bars show the statistical uncertainty and the hatched bands show the statistical and systematic uncertainties added in quadrature. The lower portion of each panel shown the ratio of the yields between the collision data and the simulation. Figure taken from [2].

3 Ideogram Method

The joint likelihood function is constructed to determine the top quark mass simultaneously with the jet energy scale factor (JSF) in the selected events, where the JSF aims to reduce uncertainty stemming from the standard CMS jet energy corrections [13]. The observable used for measuring m_{t} is the mass $m_{\mathrm{t}}^{\mathrm{fit}}$ found in the kinematic fit, while the reconstructed W boson mass $m_{\mathrm{W}}^{\mathrm{reco}}$, before it is constrained by the kinematic fit, is used to measure *in situ* the JSF. From these distributions, probability density functions P_j are derived separately for the different permutation cases j: *cp*, *wp*, or *un*. These functions depend on m_{t} and JSF and are labeled $P_j(m_{\mathrm{t,i}}^{\mathrm{fit}}|m_{\mathrm{t}}, \mathrm{JSF})$ and $P_j(m_{\mathrm{W,i}}^{\mathrm{reco}}|m_{\mathrm{t}}, \mathrm{JSF})$ respectively for the ith permutation of an event in the final likelihood. The most likely m_{t} and JSF values are obtained by minimizing $-2\ln\left[\mathcal{L}\left(\mathrm{sample}|\mathrm{m_t}, \mathrm{JSF}\right)\right]$. The likelihood $\mathcal{L}\left(\mathrm{sample}|\mathrm{m_t}, \mathrm{JSF}\right)$ is defined as:

$$\mathcal{L}\left(\mathrm{sample}|\mathrm{m_t}, \mathrm{JSF}\right) = P(\mathrm{JSF}) \cdot \mathcal{L}\left(m_{\mathrm{t}}, \mathrm{JSF}|\mathrm{sample}\right) =$$

$$
P(\mathrm{JSF}) \cdot \prod_{\mathrm{events}} \left(\sum_{i=1}^{n} P_{\mathrm{gof}}\left(i\right) \left(\sum_{j} f_j P_j(m_{\mathrm{t,i}}^{\mathrm{fit}}|m_{\mathrm{t}}, \mathrm{JSF}) \times P_j(m_{\mathrm{W,i}}^{\mathrm{reco}}|m_{\mathrm{t}}, \mathrm{JSF}) \right) \right)^{w_{\mathrm{evt}}},
$$

where n denotes the number of permutations in each event, j labels the permutation cases, and f_j represents their relative fraction. The event weight $w_{\mathrm{evt}} = c \sum_{i=1}^{n} P_{\mathrm{gof}}\left(i\right)$ is introduced to reduce the impact of events without correct permutations, where c normalizes the average w_{evt} to 1. The choice of the prior $P(\mathrm{JSF})$ in the likelihood fit defines the three set of the solution: 1D approach that estimates only the most probable m_{t}, since the JSF is fixed to unity, 2D approach with an unconstrained JSF estimates both m_{t} and JSF. Finally, in the "'hybrid"' approach uses the prior knowledge on the JSF derived from the uncertainties on the jet energy corrections [13] and computes the prior $P(\mathrm{JSF})$ of a Gaussian form centered at 1.0 with a width σ_{prior} calculated from the relative weight $w_{\mathrm{hyb}} = 0.3$. This results in a width for $P(\mathrm{JSF})$ of $\sigma_{\mathrm{prior}} = \delta\mathrm{JSF}_{\mathrm{stat}}^{\mathrm{2D}}\sqrt{1/w_{\mathrm{hyb}} - 1} = 0.0012$, where $\delta\mathrm{JSF}_{\mathrm{stat}}^{\mathrm{2D}}$ is the statistical uncertainty on the JSF obtained from the 2D method and is expected to be 0.0008.

Due to some simplifications in the used ideogram method, like fixed permutation number and neglecting of the background, the 2D method is calibrated. This is performed by conducting 10 000 pseudo-experiments for each combination of the seven $m_{\mathrm{t,gen}}$ and the five JSF values for the muon and electron channels separately using simulated $t\bar{t}$ and background events. Corrections for the statistical uncertainty of the method are derived from the widths of the corresponding pull distributions and have a size of 5%.

4   Systematic uncertainties

The results on the systematic uncertainty estimation is shown in Table 1. Only the largest systematic uncertainties are discussed in this work, more detailed information is given in [2].

The systematic uncertainties are larger than for the Run 1 result of $m_{\mathrm{t}} = 172.35 \pm 0.16\,(\mathrm{stat} + \mathrm{JSF})\pm 0.48\,(\mathrm{syst})\,\mathrm{GeV}$, due to the more evolved treatment of the modeling uncertainties. This is mainly caused by the evaluation of a broader set of color reconnection models that were not available in Run 1. In Run 2 the effects of color reconnection on the top quark decay products can be turned on in PYTHIA 8 by enabling early resonance decays (ERD). In the default sample the early resonance decays are turned off. Additionally, the uncertainties that arise from ambiguities in modeling color reconnection effects are estimated by comparing the default model in PYTHIA 8 with two alternative models of color reconnection, a model with string formation beyond leading color ("QCD inspired") [14] and a model in that the gluons can be moved to another string ("gluon move") [15]. The influence of the matrix-element generator is estimated by using a sample from the aMC@NLO generator and the FxFx matching [16] instead of the POWHEG v2 generator used as default. The

Table 1: List of systematic uncertainties for the fit to the combined data set.

	2D approach		1D approach	Hybrid	
	(GeV)		(GeV)	(GeV)	
Experimental uncertainties					
JEC (quad. sum)	0.13	0.002	0.85	0.19	0.003
JER	0.08	0.001	0.04	0.04	0.001
Pileup	0.08	0.001	0.02	0.05	0.001
b tag. & non-t$\bar{\text{t}}$ BG	0.07	0.001	0.05	0.06	0.001
Modeling uncertainties					
JEC: Flavor (linear sum)	0.42	0.001	0.31	0.39	<0.001
b-jet modeling (quad. sum)	0.13	0.001	0.09	0.12	<0.001
ME generator	0.19	0.001	0.29	0.22	0.001
ISR & FSR PS scale	0.25	0.004	0.24	0.14	0.003
PDF, μ_R & μ_F, ME/PS	0.08	0.001	0.04	0.05	0.001
Underlying event	0.07	0.001	0.10	0.06	0.001
Early resonance decays	0.22	0.008	0.23	0.03	0.005
CR modeling	0.34	0.001	0.42	0.31	0.001
Total systematic	**0.71**	**0.010**	**1.09**	**0.62**	**0.008**
Statistical (expected)	0.09	0.001	0.05	0.07	0.001
Total (expected)	**0.72**	**0.010**	**1.09**	**0.62**	**0.008**

systematic uncertainty coming from the jet energy correction is still the dominant one.

5 Results

Out of 35.9 fb^{-1} of 2016 data, 161 496 lepton+jets events are selected, which yields for the 2D method:

$$
\begin{aligned}
m_\text{t}^\text{2D} &= 172.40 \pm 0.09 \text{ (stat+JSF)} \pm 0.68 \text{ (syst) GeV,} \\
\text{JSF}^\text{2D} &= 0.994 \pm 0.001 \text{ (stat)} \pm 0.010 \text{ (syst),}
\end{aligned}
$$

for the 1D method:

$$
m_\text{t}^\text{1D} = 171.93 \pm 0.07 \text{ (stat)} \pm 1.09 \text{ (syst) GeV,}
$$

and for the hybrid method:

$$
\begin{aligned}
m_\text{t}^\text{hyb} &= 172.25 \pm 0.08 \text{ (stat+JSF)} \pm 0.62 \text{ (syst) GeV,} \\
\text{JSF}^\text{hyb} &= 0.996 \pm 0.001 \text{ (stat)} \pm 0.008 \text{ (syst).}
\end{aligned}
$$

As expected, the hybrid method offers the lowest overall uncertainty. Due to the larger integrated luminosity and the higher t$\bar{\text{t}}$ cross section at $\sqrt{s} = 13$ TeV the statistical uncertainty is halved compared to the Run 1 result.

6 Measured top quark mass as a function of kinematic observables

The comparison of different models for soft and perturbative QCD are the main source of systematic uncertainties on the presented analysis. Differential measurements of m_t as a function of the kinematic properties of the $t\bar{t}$ system allow the validation of the different models and identification of possible biases in the presented measurement. Variables are selected that probe potential effects from color reconnection, initial- and final-state radiation, and the kinematics of the jets coming from the top quark decays. Figure 2 demonstrates the dependence of the measured m_t on the ΔR between the light quark jets ($\Delta R_{q\bar{q}}$) and on the invariant mass of the $t\bar{t}$ system ($m_{t\bar{t}}$) compared to different generator and different CR models choices. Even with the improved statistical precision from the new data, no significant deviation in the value of the measured m_t is observed. The results for POWHEG v2 + HERWIG++ [17] differ from the data and all other setups, which might be explained by the missing matrix-element corrections to the top quark decay.
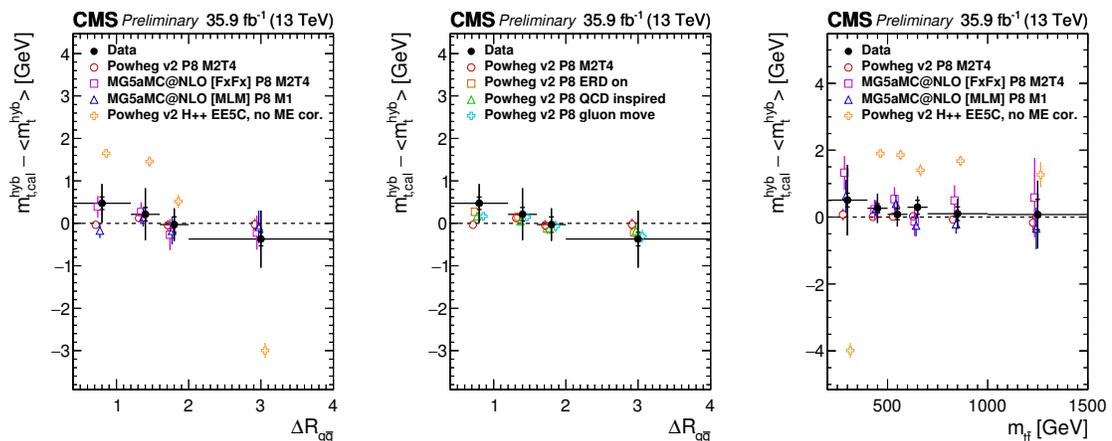

Figure 2: Measurements of m_t as a function of the ΔR between the light quark jets ($\Delta R_{q\bar{q}}$) compared to different generator (left) and different CR models (middle) choices, and m_t as a function of the invariant mass of the $t\bar{t}$ system ($m_{t\bar{t}}$) compared to different generator (right). Figure taken from [2].

7 Summary

This result of $m_t = 172.25 \pm 0.08 \,(\text{stat} + \text{JSF}) \pm 0.62 \,(\text{syst})\,\text{GeV}$ is the first result of the top quark mass measured with full 2016 data and the new NLO generator setups. The evaluation of a broader set of color reconnection models for the systematic uncertainty estimation is done. This measurement is consistent with the Run 1 result. No shift in

the measured top quark mass from the new simulation at next-to-leading order with POWHEG v2 and PYTHIA 8 and the new experimental setup is observed.

References

[1] CMS Collaboration. "The CMS Experiment at the CERN LHC". *JINST*, 3:S08004, 2008.

[2] CMS Collaboration. "Measurement of the top quark mass with lepton+jets final states in pp collisions at $\sqrt{s}$ = 13 TeV". CMS Physics Analysis Summary CMS-PAS-TOP-17-007, CERN, Geneva, 2017, `https://cds.cern.ch/record/2284594`.

[3] CMS Collaboration. "CMS Luminosity Measurements for the 2016 Data Taking Period". CMS Physics Analysis Summary CMS-PAS-LUM-17-001, 2017, `http://cds.cern.ch/record/2257069`.

[4] J. Abdallah et al. "Measurement of the mass and width of the W boson in e^+e^- collisions at $\sqrt{s} = 161 - 209$ GeV". *Eur. Phys. J. C*, 55:1, 2008.

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

[6] Paolo Nason. "A new method for combining NLO QCD with shower Monte Carlo algorithms". *JHEP*, 11:040, 2004.

[7] Campbell et al. "Top-pair production and decay at NLO matched with parton showers". *JHEP*, 04:114, 2015.

[8] Torbjorn Sjostrand, Stephen Mrenna, and Peter Z. Skands. "A Brief Introduction to PYTHIA 8.1". *Comput. Phys. Commun.*, 178:852–867, 2008.

[9] Peter Skands, Stefano Carrazza, and Juan Rojo. "Tuning PYTHIA 8.1: the Monash 2013 Tune". *Eur. Phys. J. C*, 74:3024, 2014.

[10] CMS Collaboration. "Identification of b quark jets at the CMS experiment in the LHC Run 2". CMS Physics Analysis Summary CMS-PAS-BTV-15-001, CERN, Geneva, 2016, `https://cds.cern.ch/record/2138504`.

[11] CMS Collaboration. "Identification of b-quark jets with the CMS experiment". *JINST*, 8:P04013, 2013.

[12] D0 Collaboration. "Direct measurement of the top quark mass at D0". *Phys. Rev. D*, 58:052001, 1998.

[13] CMS Collaboration. "Jet energy scale and resolution in the CMS experiment in pp collisions at $8\,\mathrm{TeV}$". *JINST*, 12:P02014, 2017.

[14] Jesper R. Christiansen and Peter Z. Skands. "String Formation Beyond Leading Colour". *JHEP*, 08:003, 2015.

[15] Spyros Argyropoulos and Torbjörn Sjöstrand. "Effects of color reconnection on $t\bar{t}$ final states at the LHC". *JHEP*, 11:043, 2014.

[16] Rikkert Frederix and Stefano Frixione. "Merging meets matching in MC@NLO". *JHEP*, 12:061, 2012.

[17] Manuel et al. Bähr. "Herwig++ physics and manual". *Eur. Phys. J. C*, 58:639, 2008.