The $Z \to \tau\tau \to e + \mu$ standard candle with early Run2 data at CMS

I.D. Bobovnikov[1,*] and Supervisor A. Raspereza[2,†]

[1]*Belorussian State University, Minsk, Belarus*

[2]*DESY-Hamburg, Germany*

(Dated: September 10, 2015)

The Drell–Yan $\tau\tau$ production is one of the main backgrounds for the studies of the Higgs boson properties in the $\tau\tau$ decay channel. In this report we present results of an initial study of the inclusive Z boson production followed by the $Z \to \tau\tau$ decay in the $e+\mu$ final states with early Run2 data delivered by the Large Hadron Collider in July 2015. The analyzed data set corresponds to an integrated luminosity of 41 pb^{-1} collected with the CMS detector at a center-of-mass energy of 13 TeV.

*Electronic address: `boboilya@yandex.by`

†Electronic address: `rasp@mail.desy.de`

I. INTRODUCTION

A discovery of a new boson at the Large Hadron Collider (LHC) was announced by the ATLAS and CMS collaborations in July 2012 [1, 2]. Subsequent measurements performed by the two collaborations have shown that the discovered particle has properties consistent with those expected for the Standard Model Higgs boson [3, 4]. Initially signal has been observed in decays to gauge bosons, $H \to WW$, ZZ, $\gamma\gamma$. After thorough and greatly improved analysis of the entire Run1 LHC data set, collected by two experiments at center-of-mass energies of 7 and 8 TeV, an evidence of more than 3 standard deviations above background expectation has been established also for the $H \to \tau\tau$ decay [5, 6].
The study of the Drell–Yan (DY) $\tau\tau$ production is essential for analyses involving τ leptons. The $Z \to \tau\tau$ process can be used as a standard candle for commissioning of the $H \to \tau\tau$ analysis. Since τ lepton has the largest mass among leptons, the $H \to \tau\tau$ decay mode is the most sensitive channel to probe Yukawa coupling of the Higgs boson to leptons.
The $\tau\tau$ decays result in various final states.

- Hadronic decays of both τ leptons happen in 42% of the cases.
- Decays to an electron (or a muon) and hadrons constitute branching fraction of $2 \times 23\% = 46\%$.
- The $e + \mu$ final states make up 6% of all $\tau\tau$ decays.
- Finally, decays to a pair of same-flavor leptons, ee or $\mu\mu$, comprise branching fraction of $2 \times 3\% = 6\%$.

The small branching fraction of the $e + \mu$ channel is in part counter-balanced by the higher efficiency of electron and muon reconstruction and lower fake rate in comparison to the reconstruction of hadronic τ decays. This channel also benefits from the absence of very large $Z \to ee$ or $Z \to \mu\mu$ backgrounds which dominate the $\tau\tau \to ee$ and $\tau\tau \to \mu\mu$ channels.

II. DATA AND MONTE CARLO SAMPLES

In this study we analyze proton-proton collision data collected with the CMS experiment at center-of-mass energy of 13 TeV. The total integrated luminosity of the analyzed data set amounts to 41 pb^{-1}.

The event samples used in this analysis have been collected with the single lepton and $e + \mu$ cross-triggers and are selected for further analysis using two alternative options of the trigger logic:

- Mu23_Ele12 OR Mu8_Ele23;
- SingleEle23 OR SingleMu24.

The first logic exploits $e + \mu$ cross-triggers with asymmetric p_{T} thresholds (23 GeV for muon and 13 GeV for electron OR 8 GeV for muon and 23 GeV for electron). The second logic exploits single lepton triggers with p_{T} threshold of 23 GeV for electron and 24 GeV for muon.

The Monte Carlo (MC) samples used for the simulation studies are detailed in Table I.

Physics process	Theoretical cross section	Event generator
$Z \to \ell\ell$, $m_{\ell\ell} < 50$ GeV	6 nb	aMC@NLO
$t\bar{t} + jets$	0.83 nb	aMC@NLO
$t + W^-$, $\bar{t} + W^+$	71 pb	POWHEG
$W + jets$, $W \to \ell\nu$	62 nb	aMC@NLO
WW	63 pb	PYTHIA8
WZ	23 pb	PYTHIA8
ZZ	10 pb	PYTHIA8

Table I: The simulated MC samples used in the analysis.

III. EVENT SELECTION

The $Z \to \tau\tau \to e{+}\mu$ events are selected with dedicated triggers discussed in previous section. Selected events must have at least one reliably reconstructed primary vertex fulfilling the following criteria:

- $|\mathrm{z_{vtx}}| < 25$ cm,
- $|\mathrm{d_{vtx}}| < 2$ cm,

- ndof ≥ 4,

where z_{vtx} and d_{vtx} are z and xy coordinates of the vertex position with respect to the geometrical center of the CMS detector, and ndof is the number of degrees of freedom in the vertex fit.

Events are required to have oppositely charged electron and muon, where the leading (trailing) lepton is required to have $p_T > 24$ (13) GeV for electrons and $p_T > 24$ (9) GeV for muons. Electrons with $|\eta| < 2.5$ and muons with $|\eta| < 2.4$ are selected. Electrons and muons are required to have an impact parameter in the transverse plane $|d_{xy}| < 0.45$ mm and along the beam axis $|d_z| < 2$ mm. We also request muon and electron to pass identification requirements as propose by the Muon [7] and E/Gamma [8] Physics Object Groups of the CMS Collaboration.

Electrons and muons from τ decays are expected to be isolated in the detector. A measure of isolation, used to discriminate the signal from the QCD multijet background, is based on information about the charged hadrons, photons, and neutral hadrons, whose momenta lie within an isolation cone $\Delta R_{iso} = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ around the lepton momentum direction. In the analysis we use isolation variable defined as

$$\mathrm{Iso}_\ell = \frac{\sum p_{T,h^\pm} + \max\left(0, \sum p_{T,\gamma} + \sum p_{T,h^0} - 1.5 \cdot \sum p_{T,pu}\right)}{p_{T,\ell}}, \quad (1)$$

where sum runs over all charged hadrons ($h^\pm$), photons (γ), neutral hadrons (h^0). The sum is corrected for pileup effects by subtracting contribution from charged particles, originating from pileup vertexes (pu). This contribution is multiplied by a factor 1.5, which accounts for the fact that in the QCD processes the total energy of neutral hadrons and the total energy of charged hadrons are produced on average with ratio 1/2. The notation $p_{T,\ell}$ refers to p_T of the lepton ($\ell = e, \mu$). In our analysis we consider two options of the isolation cone $\Delta R_{iso} = 0.3$ and 0.4. Both muon and electron isolation is required to be $\mathrm{Iso}_\ell < 0.15$.

To discriminate di-tau like events against the $t\bar{t}$ and $W + Jets$ background processes we apply additional selection criteria. The missing transverse energy in an event is required to be $E_T^{mis} < 80$ GeV. Another cut is placed on the D_ζ variable, built from the momenta of leptons ($\vec{p}_{T,e}$ and $\vec{p}_{T,\mu}$) and the missing transverse momentum ($\vec{E}_T^{mis}$). First the bisector $\hat{\zeta}$ is constructed between the flight directions of electron and muon

$$\hat{\zeta} = \frac{\hat{p}_{\mathrm{T,e}} + \hat{p}_{\mathrm{T},\mu}}{|\hat{p}_{\mathrm{T,e}} + \hat{p}_{\mathrm{T},\mu}|}, \qquad \hat{p}_{\mathrm{T,e}} = \frac{\vec{p}_{\mathrm{T,e}}}{|\vec{p}_{\mathrm{T,e}}|}, \qquad \hat{p}_{\mathrm{T},\mu} = \frac{\vec{p}_{\mathrm{T},\mu}}{|\vec{p}_{\mathrm{T},\mu}|}, \tag{2}$$

The variable D_ζ is then defined as follows

$$D_\zeta = \hat{\zeta} \cdot \vec{E}_{\mathrm{T}}^{\mathrm{mis}} - \alpha\hat{\zeta} \cdot \left(\vec{p}_{\mathrm{T,e}} + \vec{p}_{\mathrm{T},\mu}\right) \tag{3}$$

where $\alpha = 0.85$ is the optimized value, yielding the most efficient discrimination of the signal against the $t\bar{t}$ and $W + jets$ background processes. As one can see from Fig. 1, we can significantly suppress these backgrounds by applying cuts

- $E_{\mathrm{T}}^{\mathrm{mis}} < 80$ GeV,
- $D_\zeta > -60$ GeV.

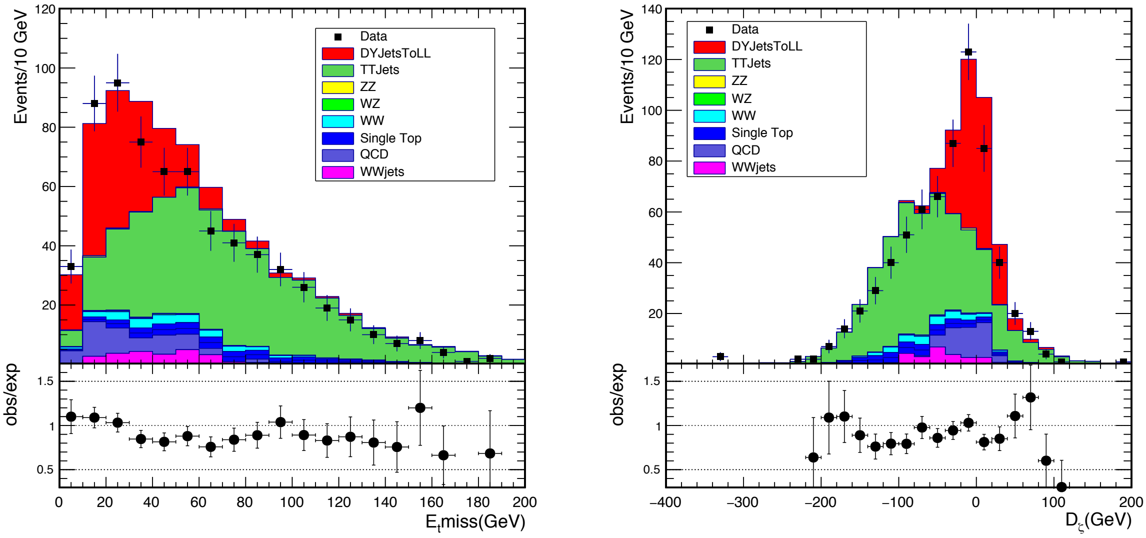

Figure 1: Missing transverse energy and D_ζ distributions after dilepton selection.

IV. PILEUP REWEIGHTING

In early Run2 data, the average number of proton-proton interactions per LHC bunch crossing was about 15. The simulated MC events are reweighted to represent the distribution of the number of pileup interactions per bunch crossing in data. The effect of reweighting is shown in Fig. 2.

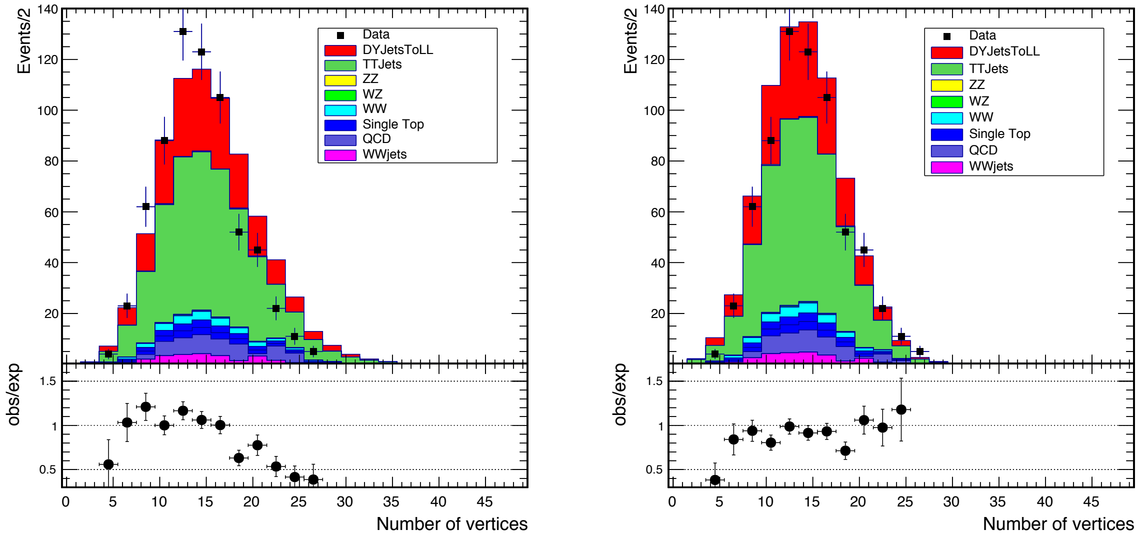

Figure 2: Vertex multiplicity distributions after dilepton selection before (left) and after (right) applying pileup reweighting.

V. QCD AND W+JETS BACKGROUND ESTIMATION

The event selection presented above, efficiently suppresses the $W + jets$ and QCD multijet backgrounds. Nonetheless, a small contribution remains and needs to be properly modeled. This is done in a data-driven way, using sample disjoint from the one used to perform statistical analysis [9]. Events are selected in this sample using the same selection as in the signal region with only one exception. The leptons are required to have same sign (SS) in contrast to the signal region, where opposite sign (OS) leptons are selected.

It is not guaranteed that the SS sample has the same yield as the OS one. Therefore two alternative corrections for possible variations in the normalization of the spectrum are implemented. Both corrections utilize the so-called ABCD method. This method uses a pair of weakly correlated variables to estimate the yield of events in one of four regions as illustrated in Fig. 3. The regions are:

- A : Opposite sign dilepton selection (signal region).
- B : Same sign dilepton selection.
- C : Opposite sign dilepton selection with inverted isolation on leptons ($\mathrm{Iso}_\ell > 0.4$).

- D : Same sign dilepton selection with inverted isolation on leptons ($\mathrm{Iso}_\ell > 0.4$).

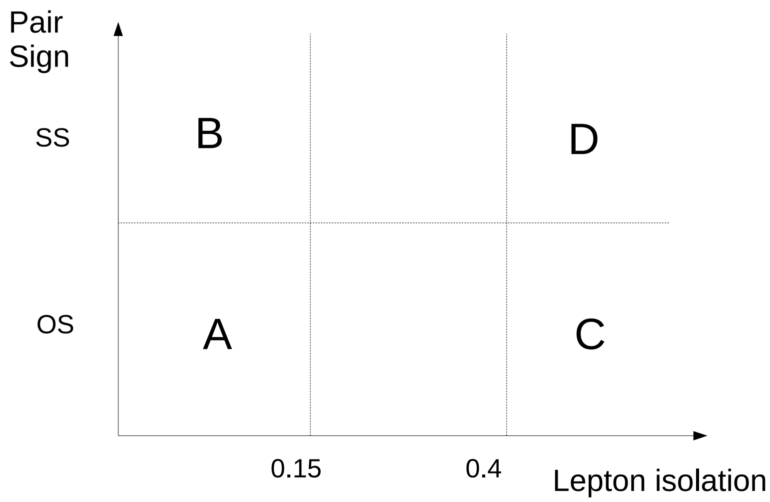

Figure 3: The ABCD method, using lepton isolation and the net charge of a $e\mu$ pair.

The correction exploits data yields in each of the QCD/$W + jets$-enriched control regions B, C and D (N_B, N_C and N_D) to determine the QCD multijet and $W + jets$ background yield in the signal region (N_A):

$$N_\mathrm{A} = N_\mathrm{B}\frac{N_\mathrm{C}}{N_\mathrm{D}}. \tag{4}$$

For the ratio $N_\mathrm{C}/N_\mathrm{D}$, we take an average of three options of the inverted isolation on leptons:

- only electron is anti-isolated;
- only muon is anti-isolated;
- both electron and muon are anti-isolated.

$$\frac{N_\mathrm{C}}{N_\mathrm{D}} = \left\langle \left(\frac{N_\mathrm{C}}{N_\mathrm{D}}\right)_{\mathrm{inv,e}}, \left(\frac{N_\mathrm{C}}{N_\mathrm{D}}\right)_{\mathrm{inv},\mu}, \left(\frac{N_\mathrm{C}}{N_\mathrm{D}}\right)_{\mathrm{inv,e}\mu} \right\rangle \tag{5}$$

For the ratio $N_\mathrm{C}/N_\mathrm{D}$ we obtain 1.4±0.3. The shape of the QCD multijet and $W + jets$ backgrounds is modeled using region B. The results presented hereafter are obtained with the QCD multijet and $W + jets$ background estimation method described in this section.

VI. ESTIMATION OF TOP-PAIR BACKGROUND NORMALIZATION

The normalization of the top-pair background is performed in a data-driven way. For this purpose, inverse cuts are applied:

- $D_\zeta < -60$ GeV,
- $E_{\mathrm{T}}^{\mathrm{mis}} > 80$ GeV.

Distributions of the electron-muon invariant mass after applying cut on D_ζ, on $E_{\mathrm{T}}^{\mathrm{mis}}$, and on both variables are shown in Fig. 4. The high-purity $t\bar{t}$-enriched samples are obtained by applying these cuts and can be used to control the normalization of the top-pair background. The $t\bar{t}$ normalization factor $\alpha_{t\bar{t}}$ is estimated from the control regions as follows

$$\alpha_{t\bar{t}} = \frac{N_{\mathrm{data}} - N_{\mathrm{other}}}{N_{t\bar{t}}}, \tag{6}$$

where N_{data} denotes the number of observed data events, $N_{t\bar{t}}$ is the number of predicted MC events for the $t\bar{t}$ production, and N_{other} is the number of predicted MC events for all other processes. Results for different combination of cuts are summarized in Table II. The average scale factor is taken to be $\alpha_{t\bar{t}} = 0.8$. Figure 5 shows distributions of $E_{\mathrm{T}}^{\mathrm{mis}}$ and D_ζ in the signal region after applying the normalization scale factor $\alpha_{t\bar{t}} = 0.8$ to the $t\bar{t}$ MC sample. The agreement between data and simulation is improved.

Cuts	Purity	$\alpha_{t\bar{t}}$	$\mathrm{d}\alpha_{\mathrm{t\bar{t}}}$	$(\mathrm{d}\alpha/\alpha)_{\mathrm{t\bar{t}}}$
$D_\zeta < -60$ GeV	0.82	0.83	0.07	0.09
$E_{\mathrm{T}}^{\mathrm{mis}} > 80$ GeV	0.84	0.79	0.08	0.10
Both cuts	0.88	0.79	0.10	0.13

Table II: Purity of the $t\bar{t}$-enriched sample, scale factor for the $t\bar{t}$ background normalization (α) and its uncertainty for different combinations of cuts on $E_{\mathrm{T}}^{\mathrm{mis}}$ and D_ζ.

VII. RESULTS

We first estimate sensitivity of our analysis for the two trigger options combined with the two variants of the lepton isolation cone $\Delta\mathrm{R}_{\mathrm{iso}} = 0.3$ and 0.4. The estimation is done after

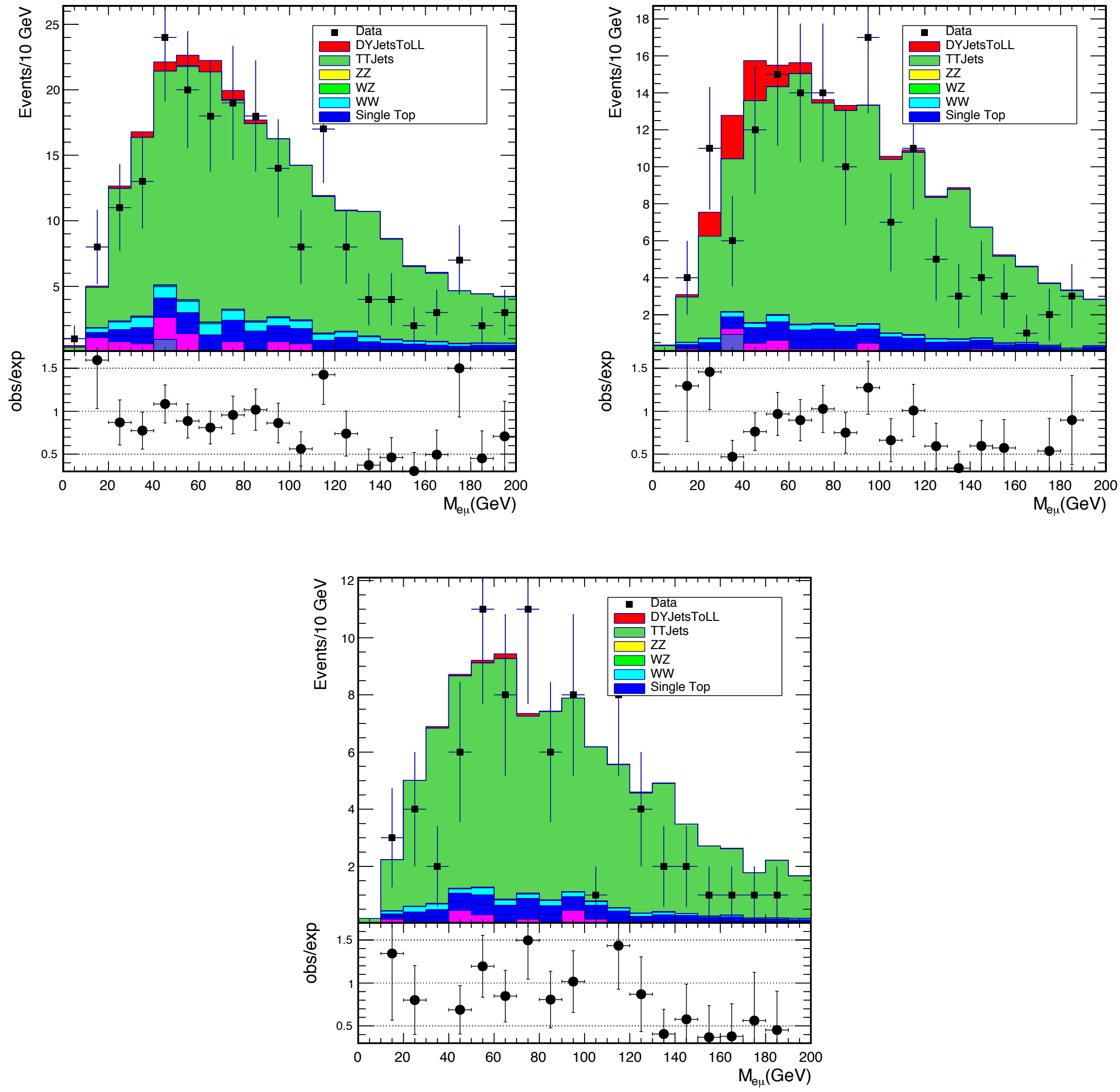

Figure 4: Dilepton mass distributions in the $t\bar{t}$-enriched control regions. Upper-left plot: cut on $E_{\mathrm{T}}^{\mathrm{mis}}$ is applied. Upper-right plot: cut on D_ζ is applied. Lower plot: cut applied on both $E_{\mathrm{T}}^{\mathrm{mis}}$ and D_ζ.

dilepton selection without applying cuts on the $E_{\mathrm{T}}^{\mathrm{mis}}$ and D_ζ variables. The distributions of the electron-muon invariant mass $m_{e\mu}$ after dilepton selection are presented in Fig. 6 for four different analysis options detailed in Table III. The expected signal significance is estimated as

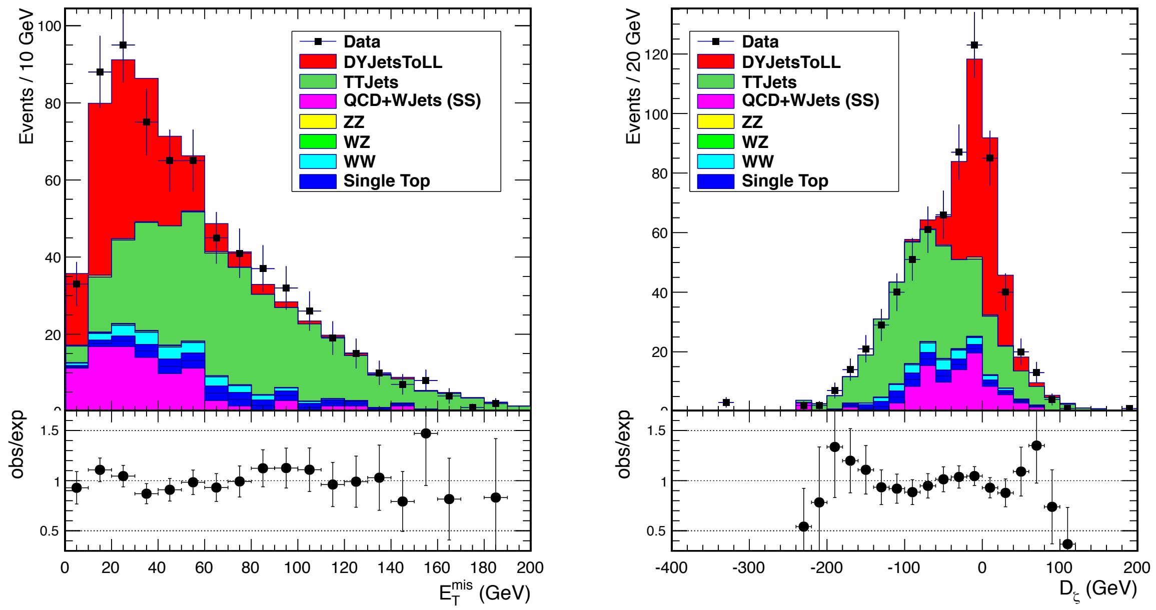

Figure 5: The $E_{\mathrm{T}}^{\mathrm{mis}}$ and D_ζ distributions after dilepton selection. The normalization scale factor is applied to the $t\bar{t}$ MC sample.

$$\mathrm{Sign}(Z \to \tau\tau) = \frac{N_{Z\to\tau\tau}}{\sqrt{N_{\mathrm{bkg}}}}, \tag{7}$$

where $N_{Z\to\tau\tau}$ is the expected number of $Z \to \tau\tau$ events, and N_{bkg} is the expected number of background events in the $m_{e\mu}$ range from 30 to 90 GeV, which corresponds to the $Z \to \tau\tau$ peak in the electron-muon invariant mass spectrum. The signal significance for various analysis options is presented in Table III.

Trigger	Isolation cone $\Delta\mathrm{R}_{\mathrm{iso}}$	$\mathrm{Sign}(Z \to \tau\tau)$
Mu23_Ele13 OR Mu8_Ele23	0.3	12.7
Mu23_Ele13 OR Mu8_Ele23	0.4	12.3
SingleEle23 OR SingleMu24	0.3	12.1
SingleEle23 OR SingleMu24	0.4	11.8

Table III: The significance of the $Z \to \tau\tau \to e + \mu$ signal for four different analysis options.

The results presented hereafter, correspond to the analysis option with the e + μ cross-

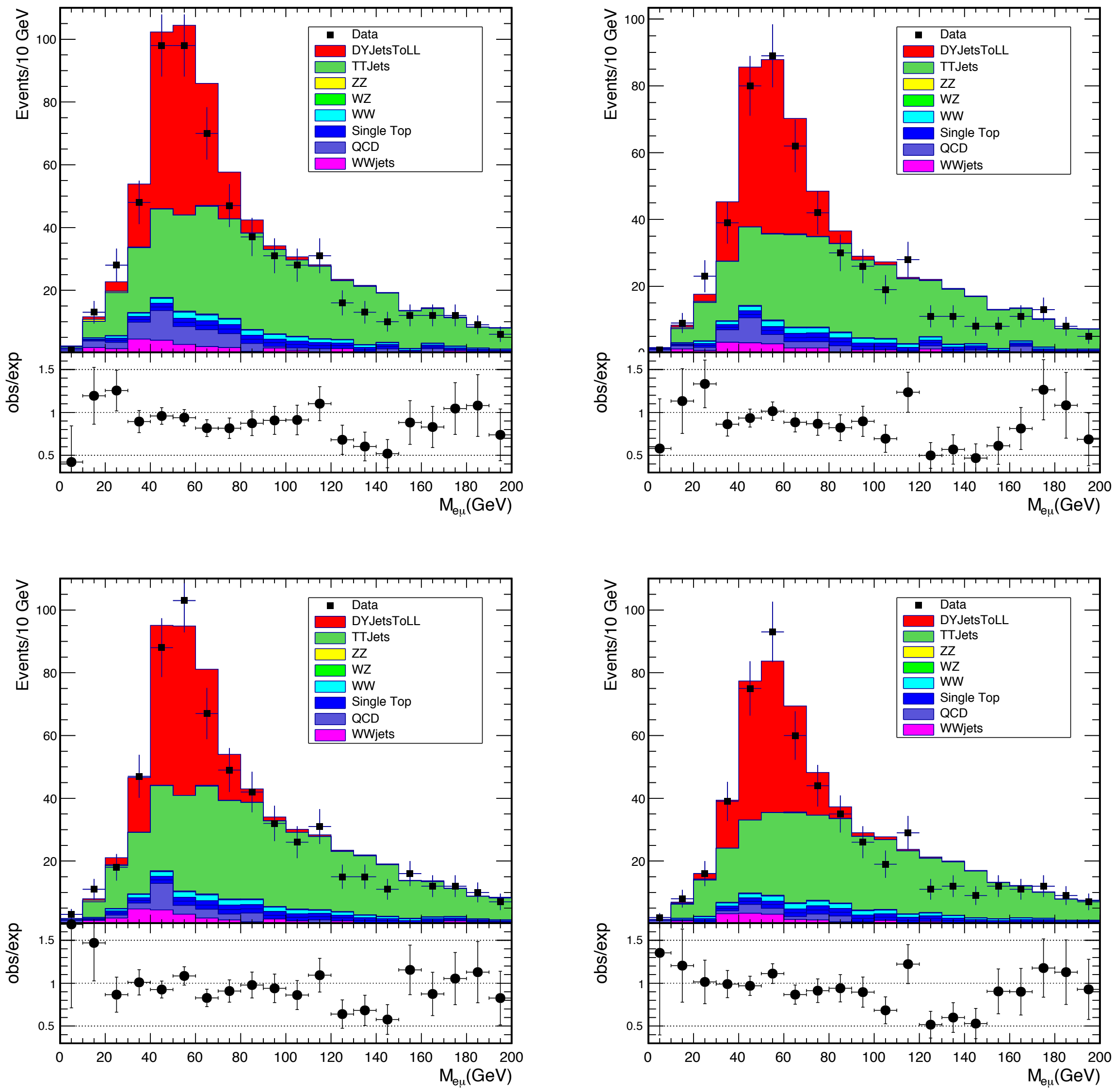

Figure 6: Dilepton mass distribution after dilepton selection for different analysis options. Top-left: $e + \mu$ triggers and $\Delta R_{iso} = 0.3$. Top-right: $e + \mu$ triggers and $\Delta R_{iso} = 0.4$. Bottom-left: single lepton triggers and $\Delta R_{iso} = 0.3$. Bottom-right: single lepton triggers and $\Delta R_{iso} = 0.4$.

triggers and lepton isolation cone of 0.3. This analysis option yields the best sensitivity to the $\tau\tau$ DY signal.

Effect of applying cuts on E_T^{mis} and D_ζ is demonstrated in Figs. 7 and 8, showing the key kinematic distributions before and after applying these additional cuts. As one can see, cuts on E_T^{mis} and D_ζ substantially suppress $t\bar{t}$ background without significant loss of the signal efficiency.

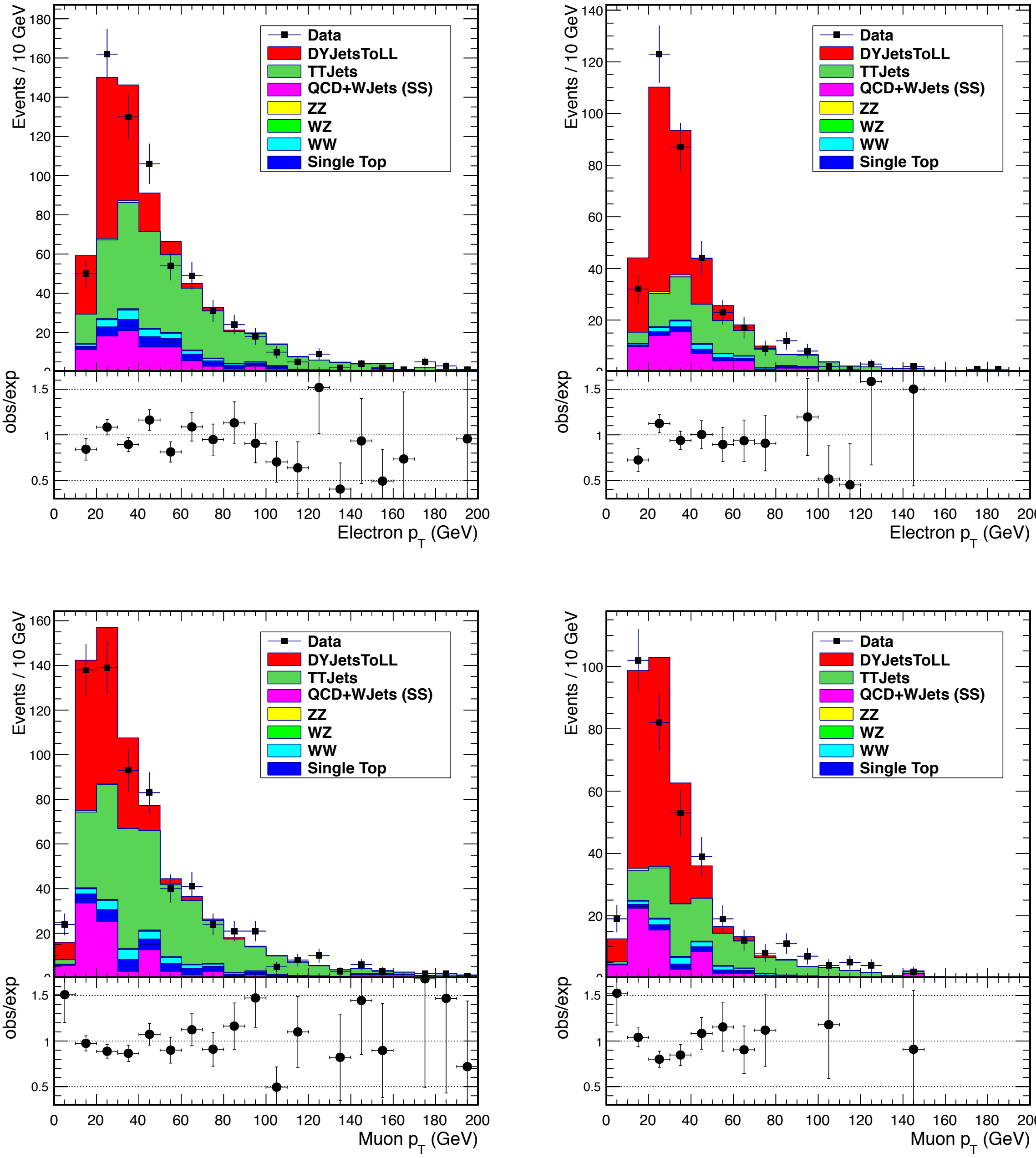

Figure 7: The distributions of the electron (top) and muon (bottom) transverse momentum before (left) and after (right) applying cuts on the $E_{\mathrm{T}}^{\mathrm{mis}}$ and D_{ζ} variables.

Finally, estimation of the cross section for the DY process at $\sqrt{s} = 13$ TeV has been made:

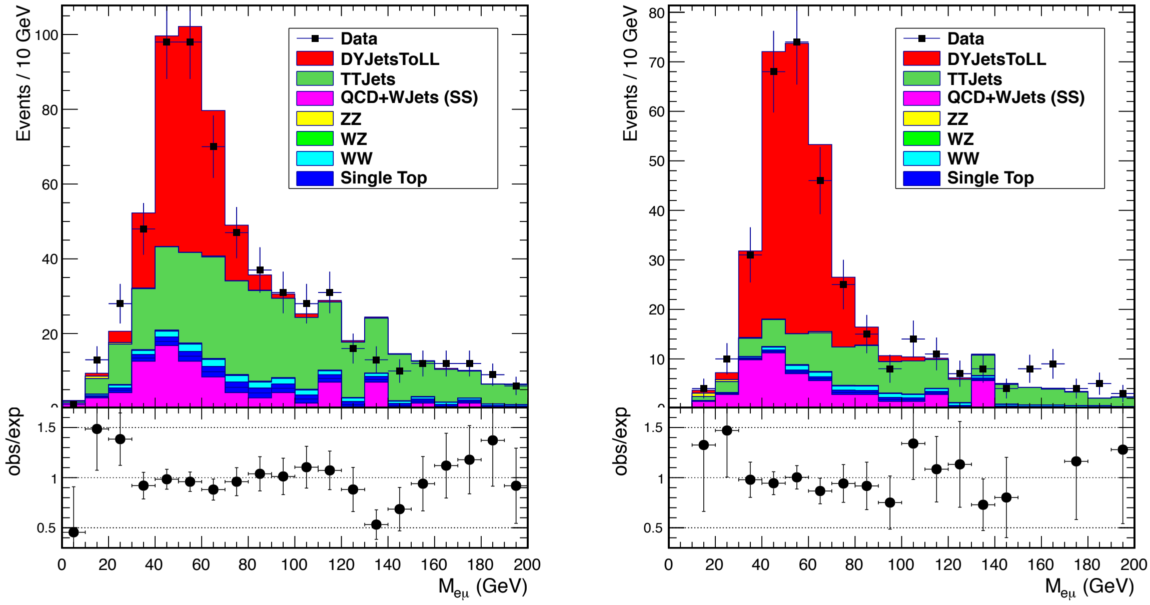

Figure 8: The electron-muon invariant mass distribution before (left) and after (right) applying cuts on $E_{\mathrm{T}}^{\mathrm{mis}}$ and D_ζ.

$$\big(\sigma(pp \to Z + X)\mathcal{B}(Z \to \ell\ell)\big)^{\text{meas}} = \frac{N_{\text{data}} - N_{\text{bkg}}}{N_{Z\to\tau\tau}}\big(\sigma(pp \to Z + X)\mathcal{B}(Z \to \ell\ell)\big)^{\text{theo}}, \quad (8)$$

where theoretical prediction for the DY cross section at 13 TeV is

$$\big(\sigma(pp \to Z + X)\mathcal{B}(Z \to \ell\ell)\big)^{\text{theo}} \approx 6\,\text{nb}.$$

The statistical uncertainty in the cross section measurement is computed as

$$\Delta\sigma_{\text{stat.}} = \sigma\frac{\sqrt{N_{\text{data}}}}{N_{\text{data}} - N_{\text{bkg}}}. \quad (9)$$

The cross section of the DY process at 13 TeV is measured to be

$$\big(\sigma(pp \to Z + X)\mathcal{B}(Z \to \ell\ell)\big)^{\text{meas}} = 5.6 \pm 0.6(\text{stat.}) \pm 0.6(\text{lumi})\,\text{nb}.$$

VIII. SUMMARY

An initial study of the inclusive Z boson production followed by the $Z \to \tau\tau$ decay in the final states characterized by the presence of isolated electron and muon, is presented. This process provides a standard candle for commissioning of the analysis, studying the Higgs boson decay into a pair of τ leptons. The analyzed data set corresponds to an integrated luminosity of 41 pb^{-1} of proton-proton collisions collected with the CMS detector at a center-of-mass energy of 13 TeV. Several analysis options are studied and various background estimation techniques are validated. The cross section for the DY process at 13 TeV, measured with the statistical and luminosity uncertainties, constitutes

$$\big(\sigma(pp \to Z + X)\mathcal{B}(Z \to \ell\ell)\big)^{\mathrm{meas}} = 5.6 \pm 0.6(\mathrm{stat.}) \pm 0.6(\mathrm{lumi})\,\mathrm{nb}.$$

The measured cross section is consistent with the theoretical prediction of about 6 nb. To complete the cross section measurement, various systematic uncertainties have to be assessed and taken into account in the measurement.

[1] G. Aad *et al.* [ATLAS Collaboration], Phys. Lett. B **716**, 1 (2012). [arXiv:1207.7214 [hep-ex]].

[2] S. Chatrchyan *et al.* [CMS Collaboration], Phys. Lett. B **716**, 30 (2012). [arXiv:1202.4083 [hep-ex]].

[3] S. Chatrchyan *et al.* [CMS Collaboration], Eur. Phys. J. C **75**, 212 (2015). [arXiv:1412.8662 [hep-ex]].

[4] G. Aad *et al.* [ATLAS Collaboration], submitted to Eur. Phys. J. C, [arXiv:1412.8662 [hep-ex]].

[5] S. Chatrchyan *et al.* [CMS Collaboration], JHEP **05** (2014) 104. [arXiv:1401.5041 [hep-ex]].

[6] G. Aad *et al.* [ATLAS Collaboration], JHEP **04** (2015) 117 [arXiv:1501.04943 [hep-ex]].

[7] `https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideMuonIdRun2`

[8] `https://twiki.cern.ch/twiki/bin/viewauth/CMS/EgammaIDRecipesRun2`

[9] L. Calligaris. "Search for a Heavy Neutral Higgs Boson decaying into Tau Pairs with the CMS Experiment at the LHC", PhD Thesis, DESY-Hamburg 2015.