

Contents

1	The CMS experiment at LHC	1
1.1	The Large Hadron Collider (LHC)	1
1.1.1	The Proton Proton Collision	2
1.1.2	The Pile Up	2
1.1.3	The Luminosity	3
1.2	The Compact Muon Solenoid (CMS)	4
1.2.1	Coordinate conventions	5
1.2.2	Inner tracking system	6
1.2.3	Electromagnetic calorimeter	7
1.2.4	Hadronic calorimeter	10
1.2.5	Magnet	13
1.2.6	Muon system	14
1.2.7	Trigger	15
1.3	Summary	16
2	Event reconstruction	17
2.1	Electron and photon reconstruction	17
2.1.1	Clustering	17
2.1.2	Super-cluster (SC)	19
2.1.3	Position measurement	20
2.1.4	Electron reconstruction	21
2.1.5	Photon reconstruction	23
2.2	Muon reconstruction	24
2.2.1	Standalone muon reconstruction	24
2.2.2	Global muon reconstruction	24
2.2.3	Tracker muon reconstruction	25
2.3	Particle-flow algorithm	25
2.4	Tau reconstruction	25
2.5	Jet and bjet reconstruction	26
2.5.1	b-jet reconstruction	27
2.6	MET reconstruction	28
2.7	Summary	28
3	Searching for high mass resonances in dielectron final state	29
3.1	Data and MC samples	29
3.2	Trigger	32
3.2.1	Method for Measuring Trigger Efficiencies in Data	32
3.2.2	Primary Signal Trigger: L1 Efficiency	33
3.2.3	Primary Signal Trigger:HLT Efficiency	34
3.2.4	Other Trigger Efficiencies	35

3.3	Object and Event Selection	36
3.4	Mass resolution and scale	37
3.5	HEEP ID efficiency and scale factor	44
3.5.1	Tag and probe method	44
3.5.2	HEEP ID efficiencies and scale factors	56
3.6	Standard Model backgrounds	64
3.6.1	SM Drell-Yan background	64
3.6.2	t <bar>t} and t<bar>t}-like background</bar></bar>	68
3.6.3	Jet background	70
3.7	Invariant mass spectra	75
3.7.1	Complementary plot	75
3.8	Statistical interpretation	84
3.8.1	Upper limits	88
3.9	Summary	91
4	Search for new physics via top quark production in dilepton final state	93
4.1	Introduction	93
4.2	Data-sets and Monte Carlo Samples	96
4.2.1	Data samples	96
4.2.2	MC samples	96
4.3	Triggers	97
4.4	Event reconstruction and object identification	98
4.4.1	Lepton selection	98
4.4.2	Jet selection	99
4.4.3	Missing Transverse Energy	100
4.4.4	Scale factors	100
4.4.5	Top P_T reweighting	100
4.5	Event selection and cut optimization	101
4.5.1	Event selection (step 1)	101
4.5.2	Event selection (step 2)	101
4.6	Background predictions	103
4.6.1	Prompt Background	103
4.6.2	Fake Background	104
4.7	Data/MC comparison	105
4.8	Signal extraction using neural networks tools	113
4.8.1	Data/MC comparison for MVA input variables	114
4.9	Systematic uncertainties	125
4.10	Results	133
4.10.1	Limit setting procedure	133
4.10.2	Exclusion limits on C_G effective coupling	134
4.10.3	Exclusion limits on C_{tG} , $C_{\phi q}$ and C_{tW} effective couplings	135
4.10.4	Exclusion limits on C_{uG} and C_{cG} effective couplings	138
4.11	Summary	141
Appendices		141
.1	Mass resolution fit results	143
.2	For 2016 HEEP ID scale factor	143
.2.1	N-1 (or N-2, N-3) efficiency for HEEP variables	143
.2.2	HEEP efficiency versus η for different E_T bins	158
.2.3	Cross check with DYJetsToLL amcatnlo sample	160

.2.4	HEEP efficiecy for mc matched electron for different DY samples .	162
.3	For 2017 HEEP ID scale factor	162
.3.1	N-1 (or N-2, N-3) efficiency for HEEP variables	162
.3.2	Cross check with fit method	175
.4	MET disagreement investigation	180
.5	SM tW cross section measurement	184

Chapter 1

The CMS experiment at LHC

This chapter introduces the Compact Muon Solenoid (CMS) experiment at Large Hadron Collider (LHC). The design of CMS detector is briefly exposed starting from the innermost region to the outermost one. Before that a primary introduction about the LHC is presented including the design of the LHC as well as the phenomenology of the proton-proton interactions.

1.1 The Large Hadron Collider (LHC)

The Large Hadron Collider (LHC) is the world's largest and most powerful particle accelerator. It was built by the European Organization for Nuclear Research (CERN) between 1998 and 2008 in collaboration with over 10,000 scientists and engineers from over 100 countries, as well as hundreds of universities and laboratories. It has 26.7 kilometres circumference and is as deep as 175 metres beneath the France-Switzerland border near Geneva shown in left plot of Figure 1.1 and it first started up on 10 September 2008. There are four main experiments at LHC which are shown in right plot of 1.1, the general description about the four experiments are in the following:

1. **ATLAS** : One of two general-purpose detectors. ATLAS investigates many different types of physics that might become detectable in the energetic collisions of the LHC. Some of these are confirmations or improved measurements of the SM (like study the Brout-Englert-Higgs boson and top quark), while many others are possible clues for new physical theories (like supersymmetry, extra dimensions and microscopic black holes theories).
2. **CMS** : The other general-purpose detector, like ATLAS, studies the SM and look for clues of new physics. Will be described in Section 1.2.
3. **ALICE** : ALICE is optimized to study heavy-ion (Pb-Pb nuclei) collisions and is focusing on the physics of strongly interacting matter at extreme energy densities. ALICE is studying a “fluid” form of matter called quark-gluon plasma which are believed to have existed a fraction of the second after the Big Bang before quarks and gluons bound together to form hadrons and heavier particles and its properties are key issues in QCD physics.
4. **LHCb** : The experiment has wide physics program covering many important aspects of heavy flavor (both beauty and charm), electroweak and QCD physics. Like measuring the parameters of CP violation in the interactions of b-hadrons (hadrons containing a bottom quark) and such studies can help to explain the matter-antimatter asymmetry of the Universe. The detector is also able to perform measurements of production cross sections and electroweak physics in the forward region.

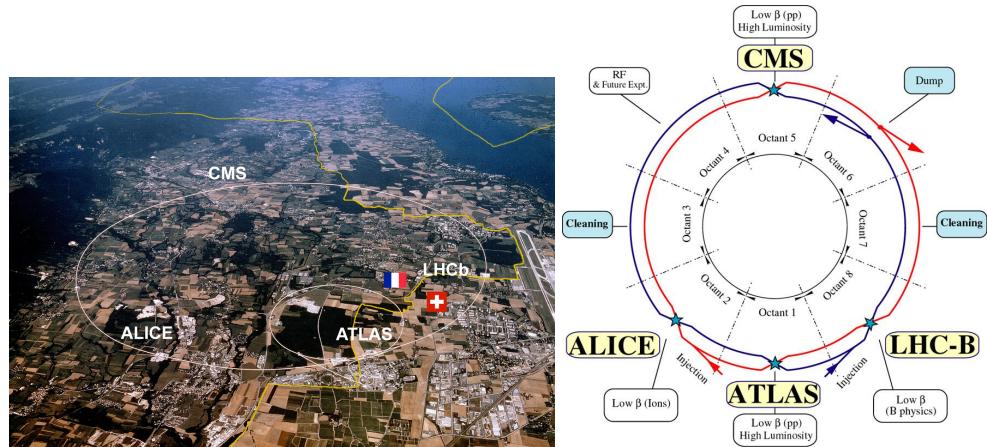


Figure 1.1: The overview of LHC (left) and the four main experiments in LHC (right).

1.1.1 The Proton Proton Collision

The reason for LHC choosing proton as accelerating particle is that proton is stable, easy to get and very low synchrotron radiation comparing with electron, so it can be accelerated to very high energy. The reason for choosing colliding beams is that it gives much higher effective collision energy (or the centre of mass energy E_{cm}) which is shown in equation 1.1 assuming \mathbf{p}_1 is four-vector $\mathbf{p} = (E, \vec{p})$ for proton 1 and \mathbf{p}_2 is four-vector for proton 2. For example when two 7 TeV protons colliding the E_{cm} is 14 TeV while for when one of protons is at rest the E_{cm} is 114.6 GeV.

$$\begin{aligned} E_{cm}^2 &= (\mathbf{p}_1 + \mathbf{p}_2)^2 = (E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2)^2 \\ E_{cm}^2 &= (E_1 + E_2)^2 \text{ when proton 1 and proton 2 collision} \\ E_{cm}^2 &= (m_1^2 + m_2^2 + 2m_2 E_{1,lab}) \text{ when proton 2 is at rest} \end{aligned} \quad (1.1)$$

The acceleration chain for the protons is shown in Figure 1.2. At first the linear particle accelerator LINAC 2 generating 50 MeV protons, which feeds the Proton Synchrotron Booster (PSB). In the PSB the protons are accelerated to 1.4 GeV and injected into the Proton Synchrotron (PS), where they are accelerated to 26 GeV and the proton bunches are formed with the correct 25 ns spacing. Then the proton beam is subsequently accelerated to 450 GeV in the Super Proton Synchrotron (SPS) and transferred to the LHC main ring. In the main ring the proton beam is accelerated in two adjacent parallel beam pipes with opposite travel direction and until they reach the target energy. Finally the proton proton or Pb Pb collisions occur at the points of the four main experiments (shown in right plot of 1.1).

1.1.2 The Pile Up

In the LHC the collisions is between proton proton bunch and there are $\mathcal{O}(10^{10})$ protons per bunch, therefore there are many proton proton collisions in the same bunch crossing and this phenomenon is so-called pile up which is shown in left plot of Figure 1.3. The number of pile up distribution of proton proton collision in 2016 and 2017 are shown in middle plot and right plot of Figure 1.3, the average number of pile up is 27 for 2016 and it is 33 for 2017.

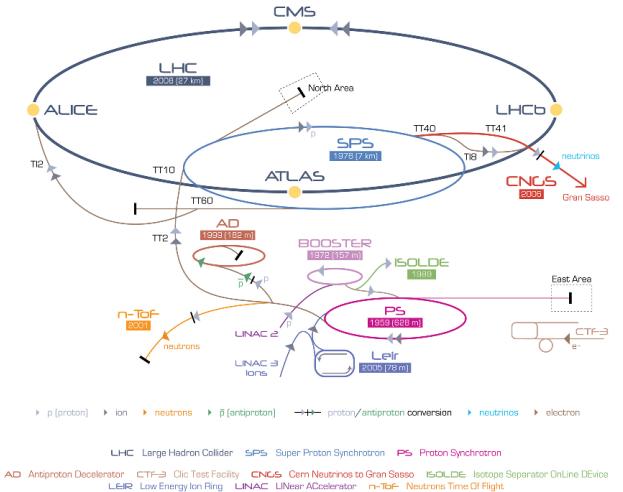


Figure 1.2: The overview of LHC accelerator chain.

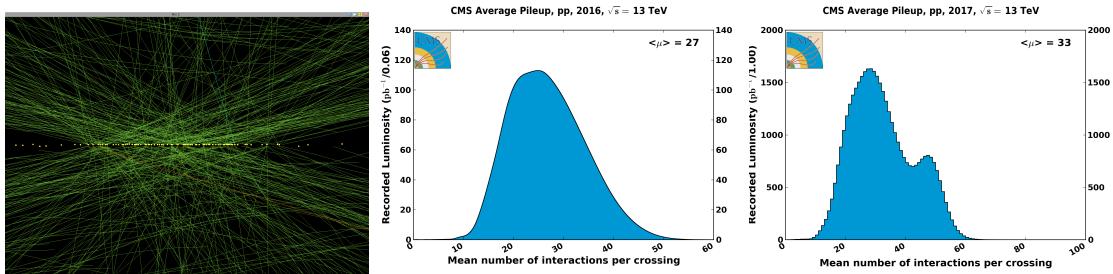


Figure 1.3: The phenomenon of pile up (left), the number of pile up distribution for 2016 (middle) and 2017 (right) [1].

1.1.3 The Luminosity

The quantity that measures the ability of a particle accelerator to produce the required number of interactions is called the luminosity and is the proportionality factor between the number of events per second dR/dt and the cross section σ_p :

$$\frac{dR}{dt} = \mathcal{L} \cdot \sigma_p \quad (1.2)$$

The unit of the luminosity is $cm^2 s^{-1}$ and it is so-called instantaneous luminosity. The instantaneous luminosity of two Gaussian beams colliding head-on can be calculated by formula 1.3 [2], where the N_1 and N_2 are the number of particles of two colliding bunches, N_b is the number of bunches in one beam, f is the beam revolution frequency and σ_x and σ_y is the width of the bunch in x axis and y axis direction respectively (here the beam moving direction is z axis).

$$\mathcal{L} = \frac{N_1 N_2 N_b f}{4\pi \sigma_x \sigma_y} \quad (1.3)$$

After integrating the instantaneous luminosity over time it gives so-called integrated luminosity:

$$\mathcal{L}_{int} = \int_0^T \mathcal{L}(t) dt \quad (1.4)$$

The cumulative online integrated luminosity of LHC delivered and CMS recorded for 2016 and 2017 are shown in Figure 1.4.

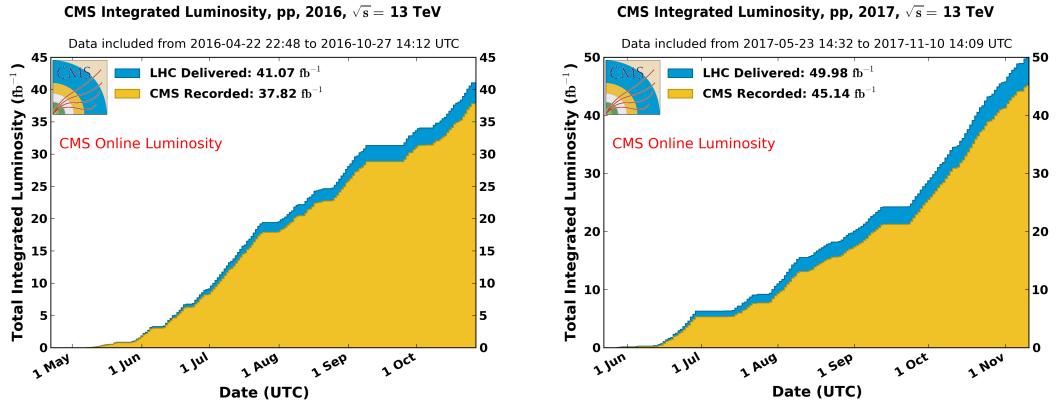


Figure 1.4: The cumulative online integrated luminosity of LHC delivered and CMS recorded for 2016 (left) and 2017 (right) [1].

The plan of LHC and High Luminosity LHC (HL-LHC) are shown in Figure 1.5. For Run 2 (from year 2015 to 2018) LHC plans to deliver 150 fb^{-1} data which is achievable because from 2015 to 2017 LHC already delivered $\sim 100 \text{ fb}^{-1}$ data. For HL-LHC the goal is to deliver 3000 fb^{-1} which is a factor of 10 increased compared with 300 fb^{-1} data which is the total number of luminosity delivered by LHC.

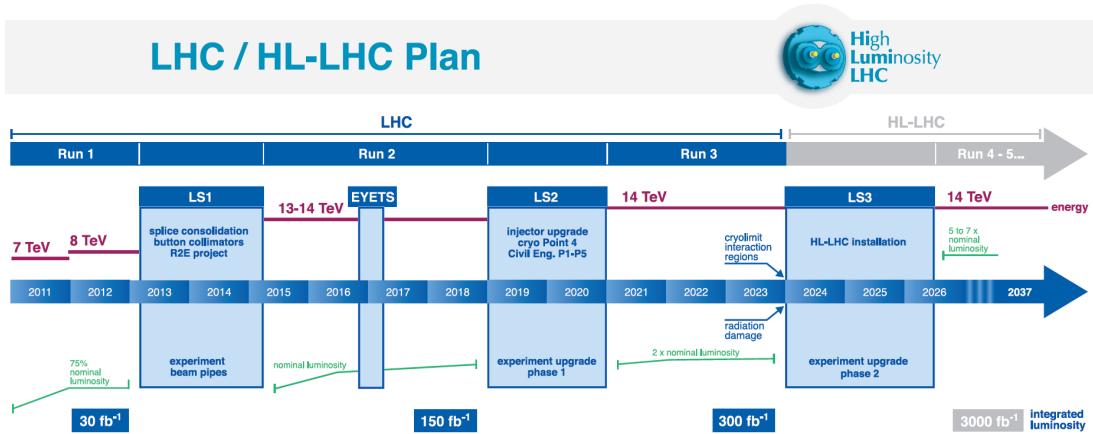


Figure 1.5: The plan of LHC and HL-LHC [3].

1.2 The Compact Muon Solenoid (CMS)

The Compact Muon Solenoid (CMS) is a general-purpose detector at the LHC. It has a broad physics programme ranging from studying the SM to searching for extra dimensions and particles that could make up dark matter. The CMS detector is built around a huge solenoid magnet. This takes the form of a cylindrical coil of superconducting cable that generates a field of 4 tesla and the field is confined by a steel “yoke” which forms the bulk of the detector’s 14,000-tonne weight. The complete detector is 21 metres long, 15 metres wide and 15 metres high. A one-quarter cross-sectional view of the CMS detector is shown in Figure 1.6 together with some labels used to name the detector elements. The transversal view of the CMS detector is shown in Figure 1.7 and the key elements from the innermost to the outermost are:

1. **Inner tracking system** which measures the trajectory of charged particles and reconstructs secondary vertices;
2. **Electromagnetic calorimeter** which measures and absorbs the energy of electrons and photons;
3. **Hadronic calorimeter** which measures and absorbs the energy of hadrons;
4. **Superconducting magnet** which provides a 3.8 T magnetic field parallel to the beam axis to bend the tracks of charged particles;
5. **Muon system** which measures the energy of muons and reconstructs their tracks.

In addition, because of the high collision rate at the LHC, a trigger system has been designed to only record data interesting for physics analyses.

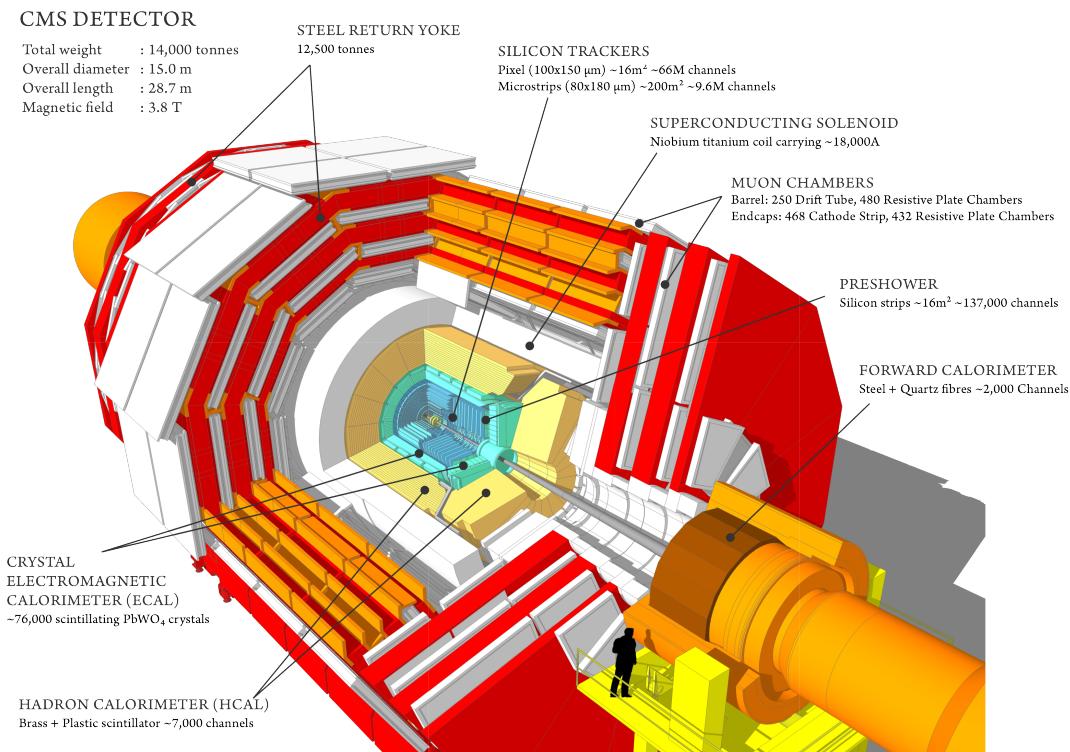


Figure 1.6: The exploded view of the CMS detector

Before moving to the description of the CMS subdetectors, the coordinate conventions is described in the following section.

1.2.1 Coordinate conventions

The coordinate system adopted by CMS has the origin centered at the nominal collision point inside the experiment, the y-axis pointing vertically upward, and the x-axis pointing radially inward toward the center of the LHC. Thus, the z-axis points along the beam direction toward the Jura mountains. The azimuthal angle ϕ is measured from the x-axis in the x-y plane. The polar angle θ is measured from the z-axis. The coordinate system is shown in Figure 1.8. Pseudorapidity is defined as $\eta = -\ln \tan(\theta/2)$. The momentum and energy measured transverse to the beam direction, denoted by P_T and E_T , respectively, are computed by $P_T(E_T) = P(E) \cdot \sin\theta$. The imbalance of energy measured in the transverse plane is denoted by E_T^{miss} (or \cancel{E}_T) which is computed by $\overrightarrow{E_T^{miss}} = -\sum \overrightarrow{E_T^{exists}}$ (more details in Section 2.6).

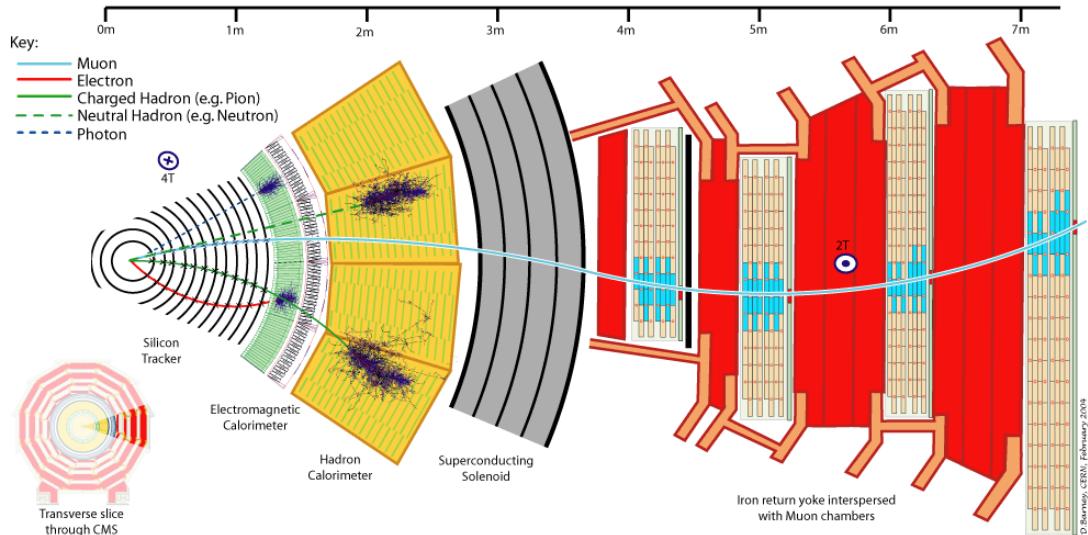


Figure 1.7: The transversal view of the CMS detector

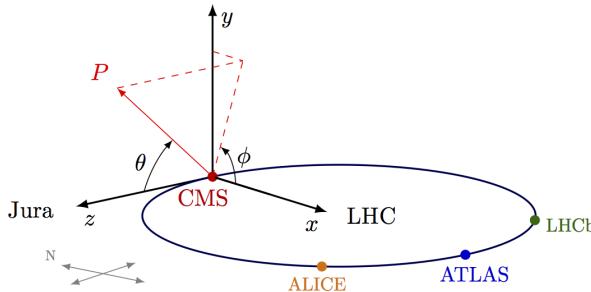


Figure 1.8: The CMS coordinate system

1.2.2 Inner tracking system

The CMS tracking system is used for reconstructing the tracking of charged particle. The tracker system is composed of two sub-detectors immersed in a 3.8 T magnetic field produced by an external solenoidal magnet (detailed in Section 1.2.5). Referring to Figure 1.9, the innermost detector, closest to the beam pipe, is the silicon pixel detector, while the outer detector is the silicon strip detector.

The silicon pixel detector is the innermost part of the CMS detector. It provides space point measurements of charged particle trajectories within a pseudorapidity up to $\eta = 2.5$. The pixel detector has been designed to withstand an instantaneous luminosity of $1 \times 10^{34} cm^{-2}s^{-1}$ with a bunch spacing of 25 ns. The pixel detector comprises three layers in the barrel (TPB) region and two endcap (TPE) disks at each side of the barrel which is shown in Figure 1.10. The three barrel layers are located at mean radii of 4.4 cm, 7.3 cm and 10.2 cm, and have a length of 53 cm. The two end disks, extending from 6 to 15 cm in radius, are placed on each side at $|z| = 34.5$ cm and 46.5 cm. In order to achieve the optimal vertex position resolution, the designed size of pixels is $100 \times 150 \mu m^2$. It has a sensitive area of $1.1 m^2$ and is segmented into 66 (48 in TPB and 18 in TPE) million pixels. The endcap disks are assembled in a turbine-like geometry with blades rotated by 20° to benefit from the large Lorentz drift angle in the magnetic field. The measured hit resolution in the TPB is $9.4 \mu m$ in the $r\phi$ plane and $20\text{--}40 \mu m$ in the longitudinal

direction which depends on the angle of the track relative to the sensor.

It is worth mention that during Run 2 the instantaneous luminosity of LHC is significantly increased and it is around to $2 \times 10^{34} \text{ cm}^{-2} \text{s}^{-1}$. The number of pileup events is then increase to 50 or more which, together with ongoing radiation damage, will potentially lead to a loss in tracking efficiency. To maintain the high tracking efficiency, the pixel detector was updated during an extended winter shutdown in 2016/2017. The new detector (hereafter referred to as the Phase 1 pixel detector) is consist of four layers in the barrel (BPIX), which represents an additional layer compared to the current detector. The radius of the innermost layer will be reduced from 44 mm to 30 mm. In the endcap region (FPIX) a third disk is added per side. The new detector is therefore allow a four-point coverage in the whole tracking region and the number of channels is almost double from 66 million to 124 million with keeping the same pixel size. A comparison of the current and upgrade designs is shown in Figure 1.11.

The silicon strip tracker is placed outside of the pixel tracker. It has 10 layers in the barrel region, four in the inner barrel (TIB) and 6 in the outer barrel (TOB). It also has two endcaps, each one made up of 3 inner disks (TID) and 9 outer disks (TEC). It is composed of 9.6 million silicon strips with a pitch varying from 80 to 205 μm . The total area of the Si detectors is around 200 m^2 , providing a coverage up to $\eta = 2.5$. In order to provide 3-dimensional information, several layers in the barrel and in the endcap have stereo modules with two silicon strip modules mounted back-to-back and rotated by 100 mrad with respect to each other. This leads to a single point resolution between 23-34 μm in the $r\phi$ direction and 23 μm in z direction for TIB, for TOB it is from 35-52 μm in the $r\phi$ direction and 52 μm in z direction. The single point resolution that can be achieved depends strongly on the size of the cluster and on the pitch of the sensor and varies not only as a function of the cluster width, but also as a function of pseudorapidity, as the energy deposited by a charged particle in the silicon depends on the angle at which it crosses the sensor plane.

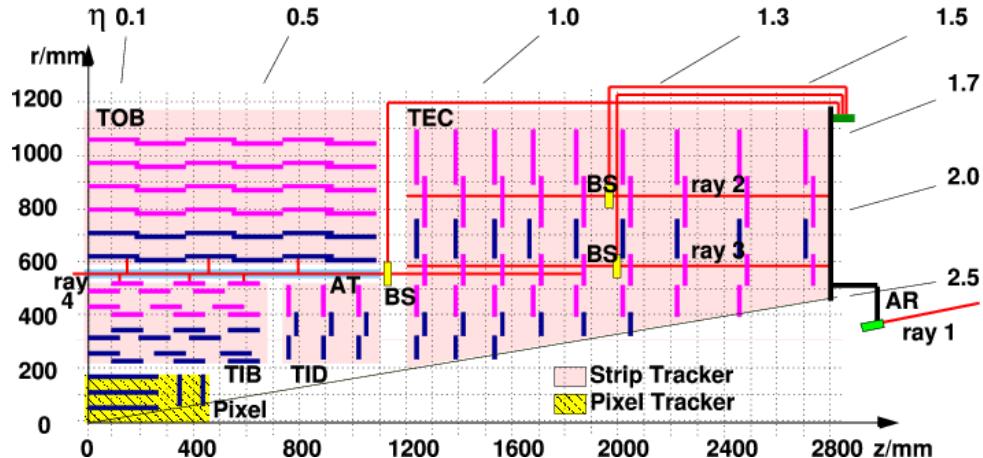


Figure 1.9: The tracker layout (1/4 of the z view).

1.2.3 Electromagnetic calorimeter

In order to measure the energy of electron and photon the CMS detector use electromagnetic calorimeter (ECAL) which is placed outside of the tracker system. The ECAL is a hermetic, homogeneous calorimeter comprising 61200 lead tungstate ($PbWO_4$) crystals mounted in the central barrel part, closed by 7324 crystals in each of the 2 endcaps with

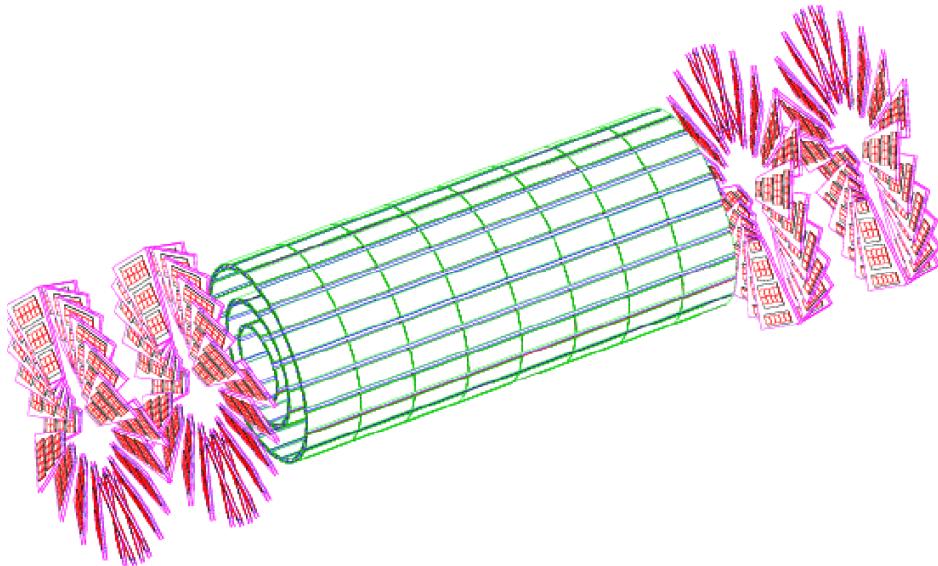


Figure 1.10: Layout of pixel detectors in the CMS tracker.

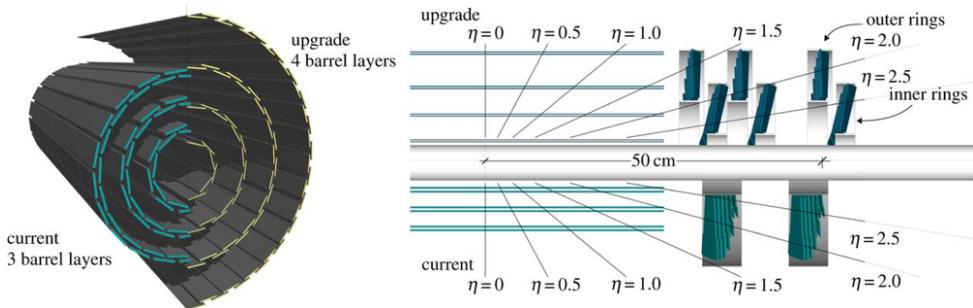


Figure 1.11: Comparison of the Phase 1 pixel detector (top) and the current detector layout (bottom).

coverage in pseudorapidity up to $|\eta| < 3.0$. A preshower system is installed in front of the edges of ECAL endcap for π^0 rejection. A 3D view of the barrel and endcap electromagnetic calorimeter is shown in Figure 1.12. The reason to choose lead tungstate scintillating crystals for ECAL is because it has short radiation ($X_0 = 0.89$ cm) and Moliere (2.2 cm) lengths, also its radiation is fast (80% of the light is emitted within 25 ns) and hard (up to 10 mrad). However, the relatively low light yield (30 γ/MeV) requires use of photodetectors with intrinsic gain that can operate in a magnetic field. Silicon avalanche photodiodes (APDs) are used as photodetectors in the barrel and vacuum phototriodes (VPTs) in the endcaps. In addition, the sensitivity of both the crystals and the APD response to temperature changes requires a temperature stability (the goal is 0.1°C). The use of ($PbWO_4$) crystals has thus allowed the design of a compact calorimeter inside the solenoid that is fast, has fine granularity, and is radiation resistant.

The barrel part of the ECAL covers the pseudorapidity range $|\eta| < 1.479$ (see Figure 1.14). The front face of the crystals is at a radius of 1.29 m and each crystal has a square cross-section of $\approx 22 \times 22 \text{ mm}^2$ and a length of 230 mm corresponding to $25.8 X_0$. The truncated pyramid-shaped crystals are mounted in a geometry which is off-pointing with respect to the mean position of the primary interaction vertex, with a 3° tilt in both ϕ and in η , in order to avoid the scenario in which a particle could go right along the separation between two center-pointing crystals. The crystal cross-section corresponds to

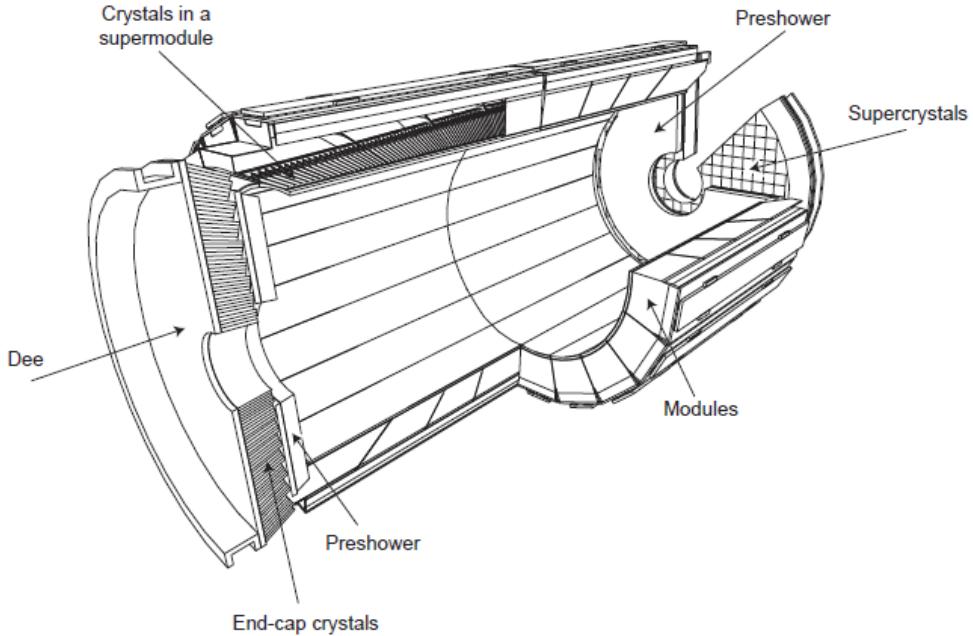


Figure 1.12: A 3D view of the barrel and endcap electromagnetic calorimeter.

$\Delta\eta \times \Delta\phi = 0.0175 \times 0.0175(1^\circ)$. The barrel granularity is 360-fold in ϕ and 2×85 -fold in η , resulting in a total number of 61200 crystals. The crystal volume in the barrel amounts to 8.14 m^3 (67.4 t). Crystals for each half-barrel will be grouped in 18 supermodules each subtending 20° in ϕ . Each supermodule will comprise four modules with 500 crystals in the first module and 400 crystals in each of the remaining three modules (see Figure 1.13). For simplicity of construction and assembly, crystals have been grouped in arrays of 2×5 crystals which are contained in a very thin wall ($200 \mu\text{m}$) alveolar structure and form a submodule.

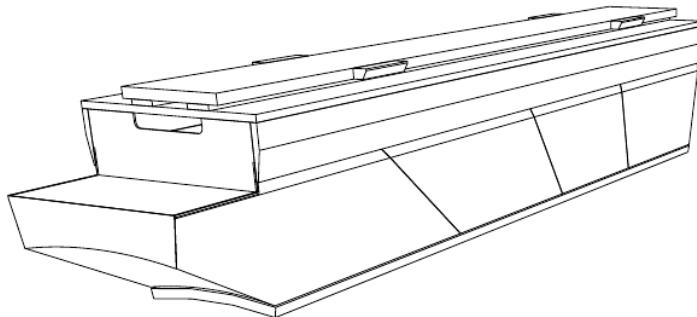


Figure 1.13: A view of the ECAL barrel supermodule which comprises four modules with 500 crystals in the first module and 400 crystals in each of the remaining three modules.

The endcap part of the crystal calorimeter covers a pseudorapidity range from 1.48 to 3.0 (see Figure 1.14). The design of the endcaps provides precision energy measurement to $|\eta| = 2.6$. However, the crystals will be installed up to $|\eta| = 3$ in order to augment the energy-flow measurement in the forward direction. The mechanical design of the endcap calorimeter is based on an off-pointing pseudoprojective geometry using tapered crystals of the same shape and dimensions ($24.7 \times 24.7 \times 220 \text{ mm}^3$) grouped together into units of 36, referred to as supercrystals. A total of 268 identical supercrystals will be used to

cover each endcap with a further 64 sectioned supercrystals used to complete the inner and outer perimeter. Each endcap contains 10764 crystals, corresponding to a volume of 1.52 m^3 (12.6 t). Both endcaps are identical and each endcap detector is constructed using Dee-shaped sections as seen Figure 1.15.

The endcap preshower covers a pseudorapidity range from $|\eta| = 1.65$ to 2.61 (see Figure 1.14). Its main function is to provide $\pi^0 - \gamma$ separation. The preshower detector, placed in front of the crystals, contains two lead converters of a total thickness of $2X_0$ and $1X_0$ respectively, followed by detector planes of silicon strips with a pitch of $< 2 \text{ mm}$. The impact position of the electromagnetic shower is determined by the barycenter of the deposited energy and accuracy is typically $300 \mu\text{m}$ at 50 GeV. In order to correct for the energy deposited in the lead converter, the energy measured in the silicon is used to apply corrections to the energy measurement in the crystal. The fraction of energy deposited in the preshower (typically 5% at 20 GeV) decreases with increasing incident energy. Figure 1.16 shows the layout of the preshower.

Figure 1.17 shows the total thickness (in radiation lengths) of the ECAL as a function of pseudorapidity where the endcap part also includes the preshower detector.

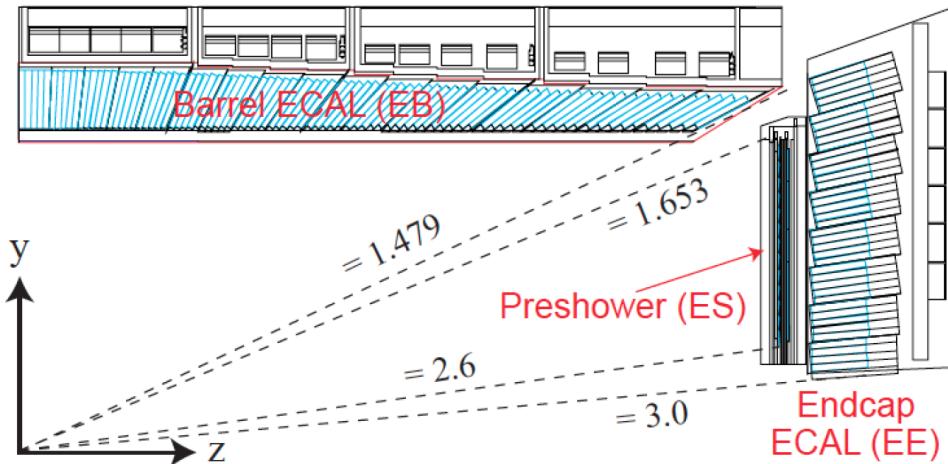


Figure 1.14: Longitudinal section of the electromagnetic calorimeter (one quadrant).

The energy resolution of the ECAL can be parameterized by the following expression:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{S}{\sqrt{E}}\right)^2 + \left(\frac{N}{E}\right)^2 + C^2 \quad (1.5)$$

where S is the stochastic term, N is the noise term and C is the constant term. Figure 1.18 summarizes the different contributions expected for the energy resolution. The stochastic term includes fluctuations in the shower containment as well as a contribution from photostatistics. The noise term contains the contributions from electronic noise and pile-up energy; the former is quite important at low energy, the latter is negligible at low luminosity. The curve labeled *intrinsic* in Figure 1.18 includes the shower containment and a constant term of 0.55%. The parameters measured in an electron test beam, for incident electrons of different energies from 20 to 250 GeV, with a 3×3 crystal configuration, considering E in GeV, correspond to $S=0.028 \sqrt{(\text{GeV})}$, $N=0.12 \text{ GeV}$ and $C=0.003$.

1.2.4 Hadronic calorimeter

In order to measure the energy of hadron (charged or neutral), CMS detector use hadronic calorimeter (HCAL) which surrounds the ECAL system. The design of the

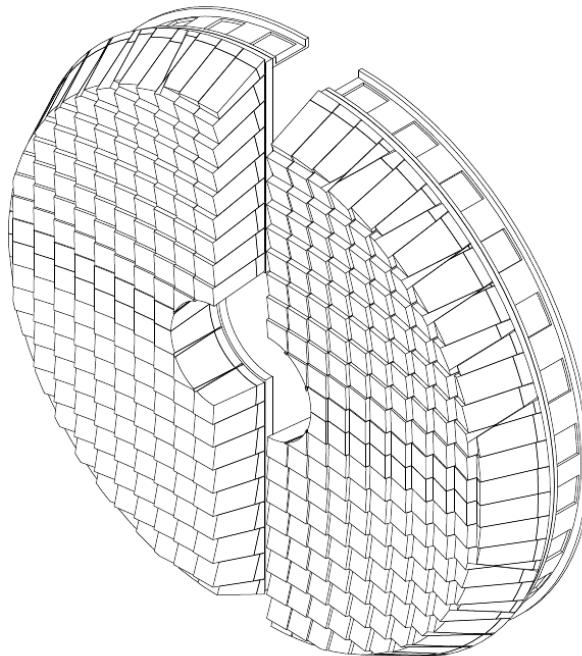


Figure 1.15: A single endcap with Dees apart.

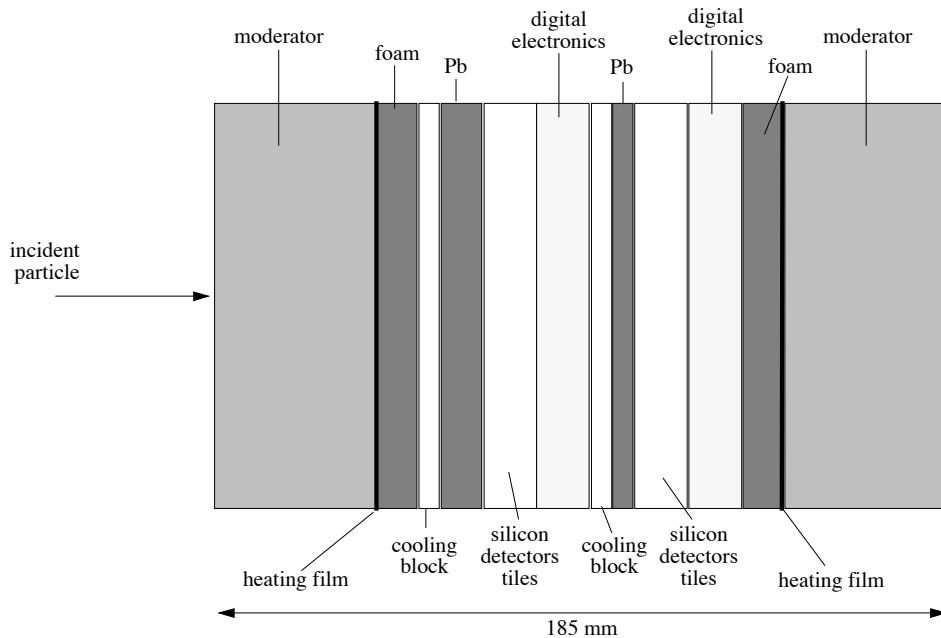


Figure 1.16: Schematic section through the endcap preshower.

HCAL is strongly influenced by the choice of the magnet parameters since most of the CMS calorimetry is located inside the magnet coil (see Figure 1.7). An important requirement of HCAL is to minimize the non-Gaussian tails in the energy resolution and to provide good containment and hermeticity for the E_T^{miss} measurement. Hence, the HCAL design maximizes material inside the magnet coil in terms of interaction lengths. Brass has been chosen as absorber material as it has a reasonably short interaction length (5.15 interaction length in 79 cm) and is relatively easy to mold and it is non-magnetic.

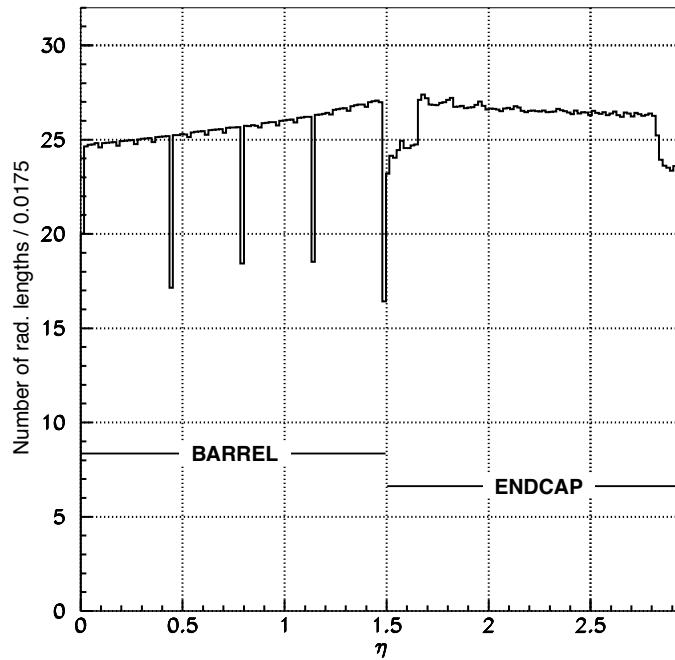


Figure 1.17: Total thickness in X_0 of the ECAL as a function of pseudorapidity, averaged over ϕ .

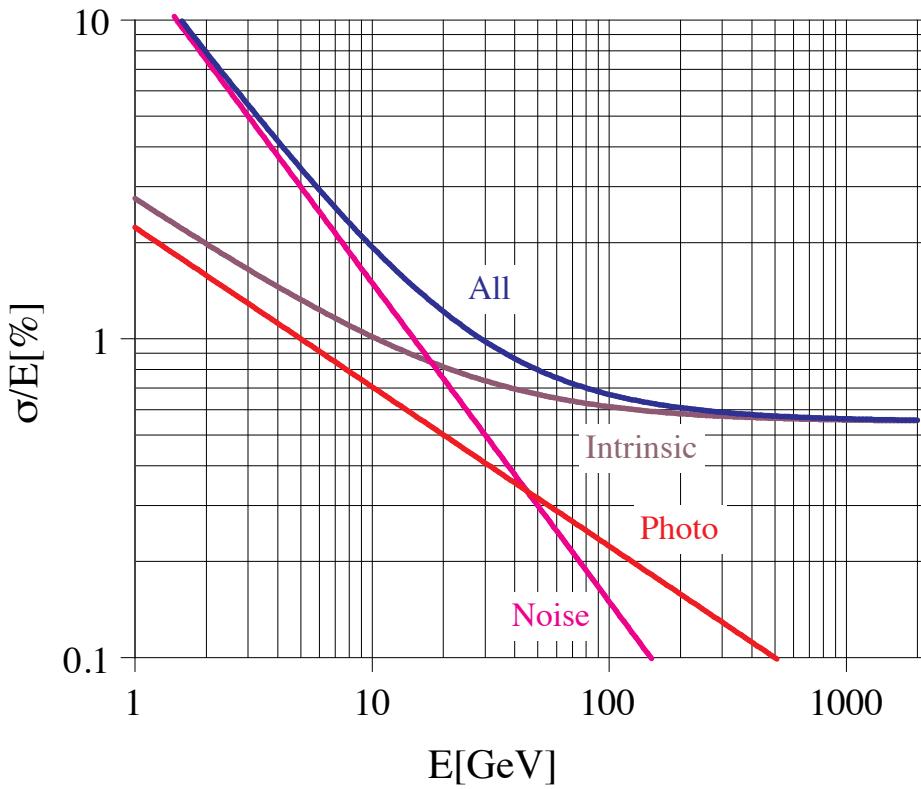


Figure 1.18: Different contributions to the energy resolution of the $PbWO_4$ calorimeter. The noise term contains the contributions from electronic noise and pile-up energy. The curve labelled “photo” describes the contribution from photostatistics, while the curve labelled “intrinsic” includes the shower containment and a constant term of 0.55%.

The architecture of the HCAL is illustrated in Figure 1.19. The hadron barrel calorimeter (HB), located inside the magnet coil, covers pseudorapidity to 1.3 and is divided in $\eta \times \phi$ towers of dimension 0.087×0.087 . The HB is complemented by an additional layer of scintillators, referred to as the hadron outer (HO) detector which lining the outside of the magnet coil. The total thickness of the combination of the HB and the HO is around twelve interaction lengths. The hadron endcap calorimeter (HE) covers a pseudorapidity range $1.3 < \eta < 3.0$ and its thickness corresponds to approximately ten interaction lengths. Forward hadron calorimeters (HF) cover the high pseudorapidity regions ($3.0 < \eta < 5.2$), as the particle flux in this very forward region is extremely high, a radiation hard technology, using Cherenkov light in quartz fibers was chosen and using steel as an absorber. The HF detector is also used as a real-time monitor for the luminosity on a bunch-by-bunch basis. The overall assembly enables the HCAL to be built with essentially no uninstrumented cracks or dead areas in ϕ . The gap between the HB and the HE, through which the services of the ECAL and the inner tracker pass, is inclined at 53° and points away from the center of the detector. The HCAL baseline single-particle energy resolution is:

$$\frac{\sigma}{E} = \frac{X}{\sqrt{E}} \oplus 5\%, \quad X=65\% \text{ (in barrel)}, \quad 83\% \text{ (in endcap)}, \quad 100\% \text{ (in forward)} \quad (1.6)$$

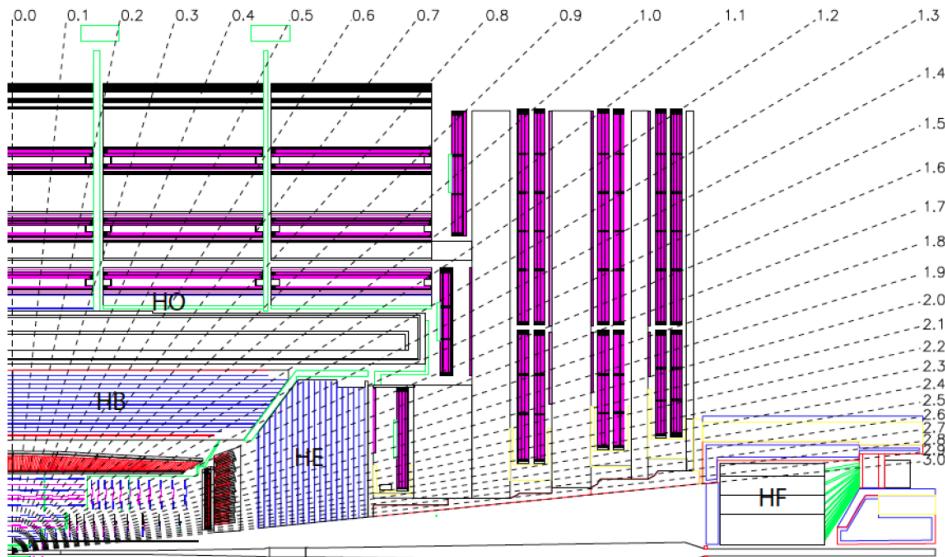


Figure 1.19: Longitudinal view of the CMS detector. The locations of the hadron barrel (HB), endcap (HE), outer (HO) and forward (HF) calorimeters are indicated.

1.2.5 Magnet

The design of the CMS magnet, in particular its bending power, was driven by the required performance of the muon system. To achieve the goal of the unambiguous determination of the curvature (hence the electric charge) for muons with momentum of 1 TeV, the requirement on the momentum resolution is: $\Delta p/p \sim 10\%$ at $p = 1$ TeV. CMS preferred indeed a higher field within a relatively compact space. Therefore a superconducting solenoid magnet, with a length of 12.9 m and an inner diameter of 5.9 m, providing a 3.8 T magnetic field, with a large bending power is used to curve the tracks

of charged particles. The magnet is made of 2168 turns carrying a 19.5 kA current and is kept at a temperature of 4.5 K with liquid helium.

1.2.6 Muon system

The muon system is used to identify muons, to measure their momenta, and to contribute to the event triggering. It relies on three types of gaseous detectors, located outside the magnet solenoid: drift tubes (DT), cathode strip chambers (CSC) and resistive plate chambers (RPC). The DT and the CSC provide an excellent spatial resolution for the measurement of charged particle momentum; the RPC are used for trigger issues because of the very good timing. The active parts of the muon system are hosted into stations which are interleaved by the iron layers of the return yoke of the magnet. The longitudinal view of a quarter of the muon system is given in Figure 1.20. The barrel of muon system (MB) is extends up to $\eta < 1.2$, the endcap of muon system (ME) is up to $\eta < 2.4$.

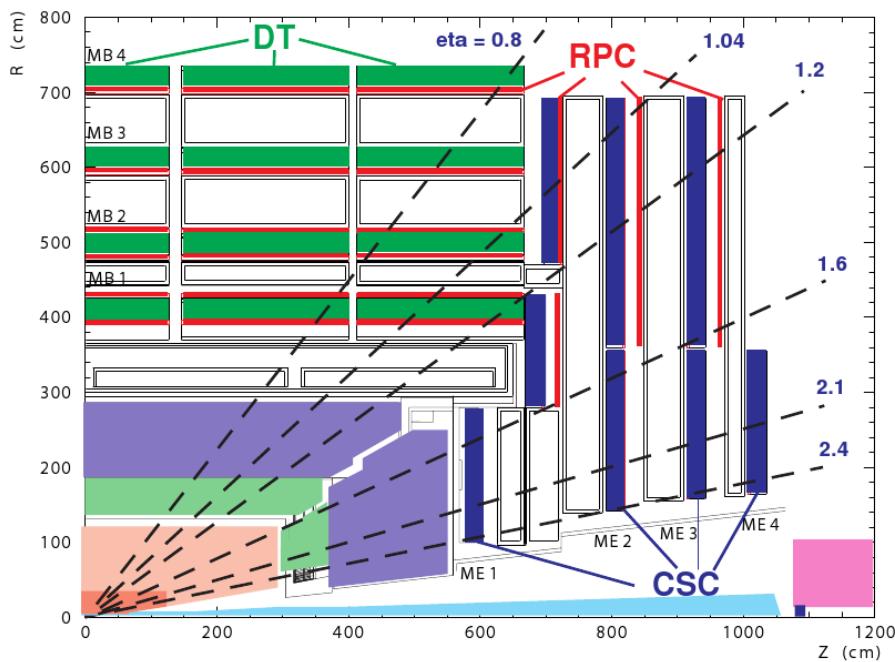


Figure 1.20: Longitudinal view of one quarter of the CMS muon system.

In the MB there are four concentric muon stations (labeled MB1, MB2, MB3 and MB4 with the last being the outermost) consisting of 250 chambers inside the magnet return yoke. The MB is further divided into five wheels around the beam axis, which are themselves divided in twelve sectors, with each covering a 30° azimuthal angle. The exact composition of the muon stations in terms of the number of DTs and their orientation, depends on the position of the station, and is chosen in such a way as to provide a good efficiency for reconstructing muon tracks from muon hits in different stations. The resolution of a single station is close to $100 \mu\text{m}$ in position and 1 mrad in direction.

The ME comprises 468 CSCs in the 2 endcaps and is divided in four stations per endcap (labeled ME1, ME2, ME3 and ME4 with the last being the outermost). The CSCs, which consist in multiwire proportional chambers, have a trapezoidal shape and count six gas gaps, each gap having a plane of radial cathode strips and a plane of anode wires running almost perpendicularly to the strips. Unlike DT, they can support the high rate of neutron-induced background and cope with a large and non-uniform magnetic

field. The spatial resolution provided by CSC is typically about $200\ \mu\text{m}$ and the angular resolution in ϕ is in the order of 10 mrad.

For low-momenta muons, the momentum resolution is by far dominated by the tracker measurements, while for particles with high momenta (around 1 TeV), the tracker and the muon system both provide a momentum resolution of about 5%. Combining the inner tracker and the muon system, the transverse momentum resolution for particles up to 1 TeV lies between 1 and 5%. Although DTs and CSCs can be used to trigger events based on the P_T of the muons with a good efficiency, their time response is comparable to the design bunch crossing space. Therefore, RPCs, which are double-gap chambers operated in avalanche mode, composed of parallel anode and cathode plates with a gas gap in between, have been introduced in the barrel and endcaps as a dedicated trigger system with a fast response and good time resolution. The position resolution of RPCs is however coarser than that of DTs and CSCs. Six layers of RPC are embedded in the barrel, whereas three layers of RPCs are in part of each endcap muon system.

Without complementary information from the tracker, the muon system provides a resolution of about 10% for muons with $\eta < 2.4$ and $P_T < 200\ \text{GeV}$.

1.2.7 Trigger

The bunch crossing frequency at the CMS interaction point is 40 MHz (bunch spacing of 25 ns) while technical difficulties in handling, storing and processing extremely large amounts of data impose a limit of about 1 kHz on the rate of events that can be written to permanent storage, as the average event size is about 1 MB. The trigger system reduces the rate by selecting events that have a physical interest, based on the characteristics of these events such as the transverse momentum of the particles. The rate reduction by at least a factor of 10^6 is done in two steps: Level-1 (L1) Trigger and High-Level Trigger (HLT).

The L1 triggers relies coarse information from calorimeters and the muon systems as well as some correlated information between these systems. The L1 trigger has a design output rate of 100 kHz and a response time of $3.2\ \mu\text{s}$. The L1 decision is based on the presence of “trigger primitive” objects such as e/γ , muons and jets with E_T or P_T above a predefined threshold. It also employs global sums of E_T and E_T^{miss} . The simplest triggers are in general based on the presence of one object with an E_T or P_T above a predefined threshold (single-object triggers) or based on the presence of two objects of the same type (di-object triggers) with either symmetric or asymmetric thresholds. Other requirements are those for multiple objects of the same or different types (“mixed” and multiple-object triggers). The high resolution data from the inner tracker are not used to generate the L1 decision, which means that there is no information about the vertices and no distinction between electrons and photons available at this level. Field Programmable Gate Array (FPGA) technology is mainly exploited for the L1 Trigger hardware, while application-specific integrated circuits (ASICs) and programmable memory lookup tables (LUT) are also used in special cases where speed, density and radiation resistance are of high importance.

Events passing the L1 Trigger are then processed by the HLT, which performs more complex calculations, based on a combination of information from the different subdetectors. The goal of the HLT is to reduce the event rate from the maximum L1 output ($\sim 100\ \text{kHz}$) to 600 Hz which is the maximum rate for mass storage. Once the L1 trigger has accepted an event, the data of this event are transferred from the buffer memory to the surface, where they are reconstructed in the HLT. The HLT is a special part of the CMS software and runs on a farm of several thousand processors. Each processor works

on the reconstruction of one event at a time, to get to a trigger decision within 100 ms on average. Since the time budget for one event is much larger than at the L1 trigger, more complicated algorithms, including tracking, can be executed at the HLT level. Once an event is accepted, it is stored on disk and fully reconstructed offline at a later time. The HLT path starts from the seed of the L1 trigger which looks for different objects and signatures in the event. One trigger path is built from reconstruction modules and filter modules. After some parts of the data are reconstructed, a filter module decides either the reconstructed objects pass the thresholds and the next step in reconstruction is started, or the event is not accepted by the path. In the latter case, the execution of the path is stopped and the following reconstruction steps and filter steps are not performed to save computation time.

If the acceptance rate from trigger is too high (for example in case of a trigger path with very low thresholds), the trigger path can be prescaled to lower the rate. For example, a prescale value 10 of HLT trigger means that the path is executed only for 1 over 10 events (randomly chosen to avoid biases) that were accepted by the L1 trigger and, consequently, the trigger rate for that HLT path is 10 times smaller. The prescale value for one trigger path has several predefined levels, depending on the instantaneous luminosity of the LHC machine. During an LHC fill, the instantaneous luminosity decreases, and the prescale values can be changed during a CMS run to keep the global trigger rate at an optimal level.

1.3 Summary

In this chapter, a basic introduction about LHC is delivered including the main experiments at LHC, the LHC accelerator, the phenomenon of proton-proton collision and pile up as well as the luminosity of LHC. After that a detailed description of CMS detector is given which is composed of several subdetectors: tracker system, electromagnetic calorimeter, hadronic calorimeter and muon chambers.

Chapter 2

Event reconstruction

This chapter describes how the physical objects are reconstructed using the combined information coming from the CMS subdetectors. The electron and photon reconstruction is described in Section 2.1, the muon reconstruction in Section 2.2, the tau reconstruction in Section 2.4, the jet and b-tagged jet reconstruction in Section 2.5 and finally the missing transverse energy reconstruction in Section 2.6.

2.1 Electron and photon reconstruction

The beginning of reconstruction of electron or photon is to cluster its energy deposition in the ECAL and then to estimate its real energy and position from this information. In fact that an electrons can radiate bremsstrahlung photons when it traverses the material between the interaction point and the ECAL, that photons can convert into electron pairs when it traverses the material, which in turn can radiate bremsstrahlung photons. The bending of the electron in the CMS magnetic field results in a spread of energy for both electron and photon in the large ϕ direction in the ECAL. The energy of electron or photon can be collected by making a cluster of ECAL clusters along a ϕ road which is called a super-cluster (SC). Finally, the presence a track which match to the SC in ECAL allows one to distinguish a electron from a photon.

2.1.1 Clustering

There are two algorithms to cluster the electromagnetic shower in the ECAL. One is Island algorithm which is designed to search for small deposits of energy in individual clusters, for example when making a calorimetric isolation cut, the basic clusters of the Island algorithm are more appropriate objects to work with. Another is Hybrid algorithm which is designed to reconstruct relatively high energy electrons in the barrel (for electrons with $E_T > 10$ GeV). The details of these two algorithms are described below.

2.1.1.1 The Island algorithm

The island algorithm starts by searching for crystals which have transverse energy above a certain threshold. These crystals are called “seeds” and are listed by decreasing energy. The algorithm then loops over seeds and removes those seeds that are adjacent to higher energy ones, after this process only the seeds with local maximum transverse energy remained. Then starting from the most energetic seed, the algorithm collects crystals belonging to a certain cluster (which contains the seed). The sequence is sketched in Figure 2.1: starting from the seed position, the algorithm moves in both directions in ϕ and collects all crystals until it sees a rise in the energy or a hole (a crystal has very

low energy which comparable to a noise). Then it moves one step in η and makes another ϕ search. The η -steps are stopped when a rise in energy or a hole is encountered. When one direction in η is completed, the algorithm goes back to the seed position and works in the other η direction. All the collected crystals are marked as belonging to that one cluster and cannot be used anymore. This procedure guarantees that there is no double counting of crystal energy.

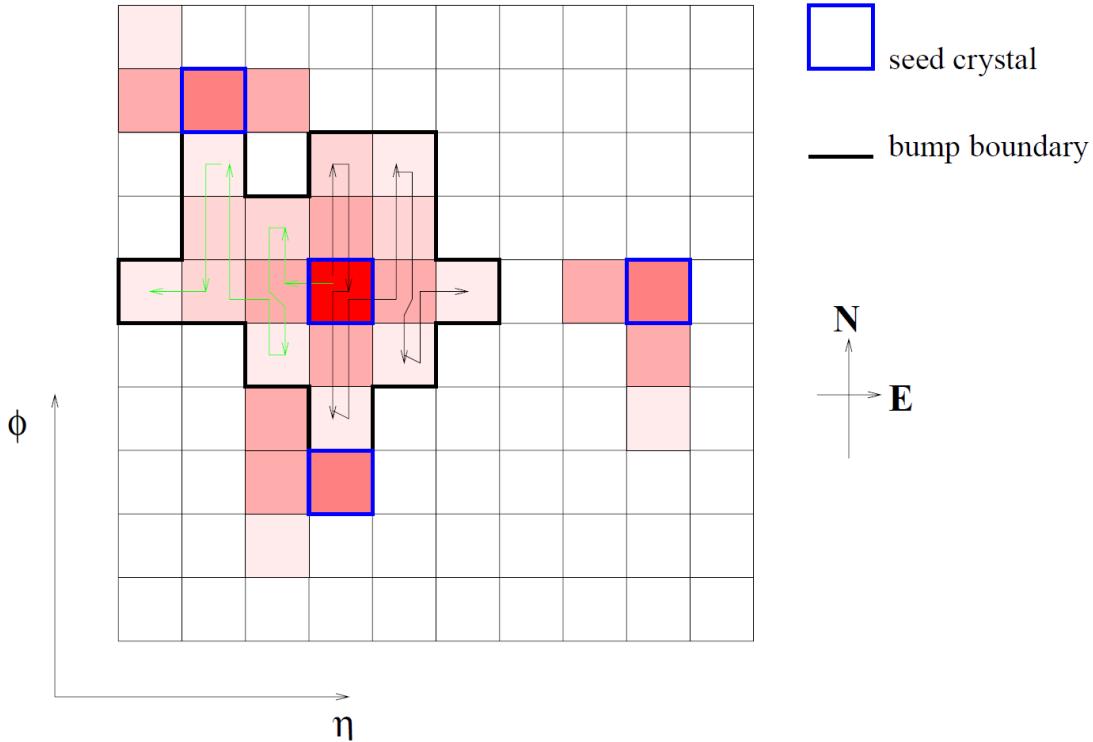


Figure 2.1: Illustration of the Island clustering algorithm in the barrel ECAL [4].

Because much of the endcap is covered by a preshower device with two planes of silicon strip readout. The energy deposited in the preshower detector (which is about $3 X_0$ thick) needs to be added to the crystal clusters. A preshower cluster is constructed in each plane, in front of each crystal cluster. The search area in the preshower is centred on the point determined by extrapolating the crystal cluster position to the preshower plane in the direction of the nominal vertex position.

There are only one parameter for the island algorithm which is the E_T threshold of the seed. This value has to be a trade-off between an optimal energy resolution and cutting off noisy hits, low pile up energy as well as keeping the execution time low.

2.1.1.2 The Hybrid algorithm

In the case of a unconverted photon shower or electrons in test beam conditions, the summed energy from fixed arrays of crystals seems to consistently give better results in terms of energy resolution, than energy sums of crystals collected dynamically according to a cluster finding algorithm. This seems to be because containment variation as a function of impact position is amplified by dynamic cluster finding (e.g. at the shower borders, where energy depositions are comparable to noise, energy belonging to the shower may be noise-suppressed, or a large noise fluctuation may fake the presence of a secondary seed). The Hybrid algorithm attempts to use the η - ϕ geometry of the barrel crystals to exploit

the knowledge of the lateral shower shape in the η direction (taking a fixed domino of three or five crystals in η), while searching dynamically for separated (bremsstrahlung) energy in the ϕ direction.

The algorithm starts from a seed crystal (the maximum energy crystal in the region being searched which must also satisfy the condition $E_T > E_T^{hybseed}$), 1×3 crystal dominoes are made, each with their central crystal aligned in η with the seed crystal. If the energy of the central crystal of a domino is greater than E^{wing} then a 1×5 domino is used. This making of dominoes proceeds N_{step} crystals in each η direction from the original seed crystal. Dominoes with energy less than E^{thresh} are eliminated. The domino construction step of the algorithm is illustrated in Figure 2.2.

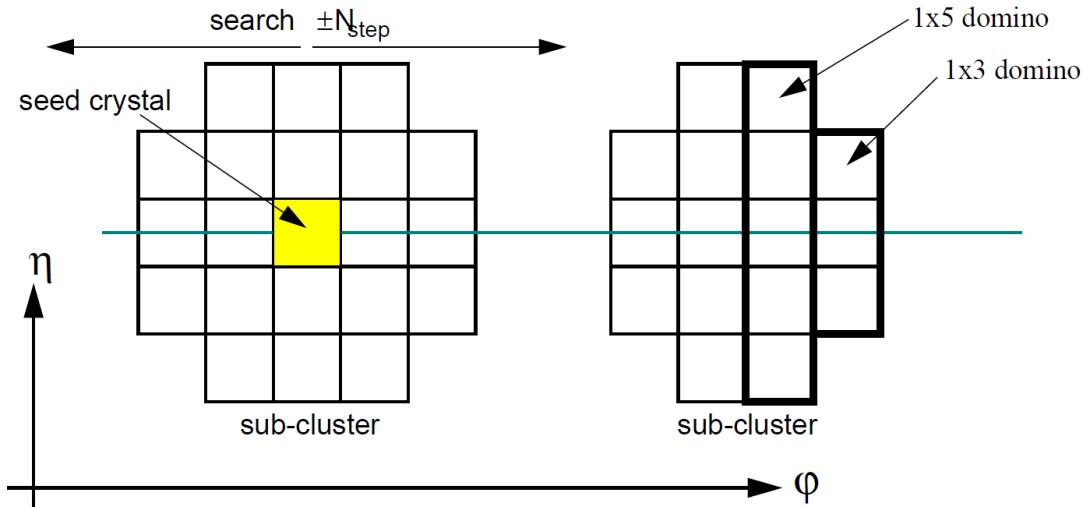


Figure 2.2: Domino construction step of Hybrid algorithm [4].

The dominoes are then clustered in ϕ . Each distinct cluster of dominoes is required to have a seed domino with energy greater than E^{seed} . The default values of the control parameters for Hybird algorithm are shown in Table 2.1.

Table 2.1: Default values of control parameters for Hybrid algorithm .

Parameter description	label used in text	default value
Minimum E_T for Hybrid super-cluster seed crystal	$E_T^{hybseed}$	1 GeV
Number steps (crystals) for search in ϕ (in each direction)	N_{step}	10
Threshold for using 1×5 crystals (rather than 1×3)	E^{wing}	1 GeV
Threshold for using domino	E^{thresh}	0.1 GeV
Minimum domino to make a disconnected subcluster	E^{seed}	0.35 GeV

2.1.2 Super-cluster (SC)

A possible approach to recollecting energy radiated by an electron or photon that falls outside the main shower cluster is to build a cluster of clusters. In much the same way as cluster energy is clustered at the level of calorimeter cells, non-overlapping clusters can in turn be clustered into SC. The procedure is started by searching for the most energetic cluster (seed cluster) and then by recollecting the others based on some geometric criterion, e.g. a fixed search area around the seed cluster. In a purely axial magnetic field

the clusters belonging to radiation from a single electron will be nicely aligned in narrow η region, but spread in ϕ . In this case, one can hope that collecting all the clusters in a narrow η window, whose size is dictated by the η position resolution of the detector, it is possible to recover most of the radiated energy (at least all that is clustered), as illustrated in the Figure 2.3.

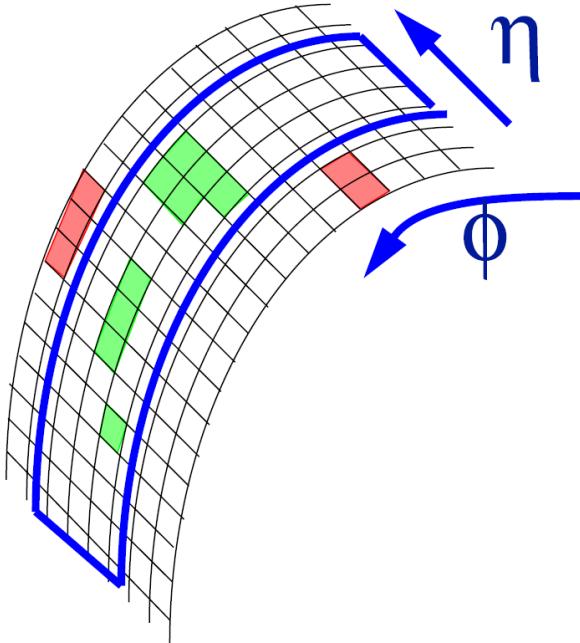


Figure 2.3: Illustration of a super-cluster algorithm collects all clusters which satisfies a given geometric condition (e.g. lying in a certain region around the seed cluster) [4].

2.1.3 Position measurement

A simple position measurement of the shower can be obtained by calculating the energy weighted mean position of the crystals in the cluster. However, there are two issues need to be considered in more detail. Firstly, the meaning of *crystal position* needs to be defined. The crystals in the CMS ECAL are quasi-projective, and do not exactly point to the nominal interaction vertex. So the lateral position (η, ϕ) of the crystal axis depends on depth as illustrated in Figure 2.4. A depth t_{max} thus needs to be defined and it is also dependent on particle type, e.g. electron showers have a short radiation length comparing with photon showers.

The second issue is related to the lateral shower shape. Since the energy density does not fall away linearly with distance from the shower axis, but rather exponentially, a simple energy weighted mean of crystal energies is distorted and the measured position is biased towards the centre of the crystal containing the largest energy deposit. Therefore, a new algorithm is used which delivers almost as good precision by calculating the weighted mean using the logarithm of the crystal energy:

$$x = \frac{\sum x_i \cdot W_i}{\sum W_i} \quad (2.1)$$

where x_i the position of crystal i , and W_i is the log weight of the crystal which is the log of the fraction of the cluster energy contained in the crystal, calculated with the formula:

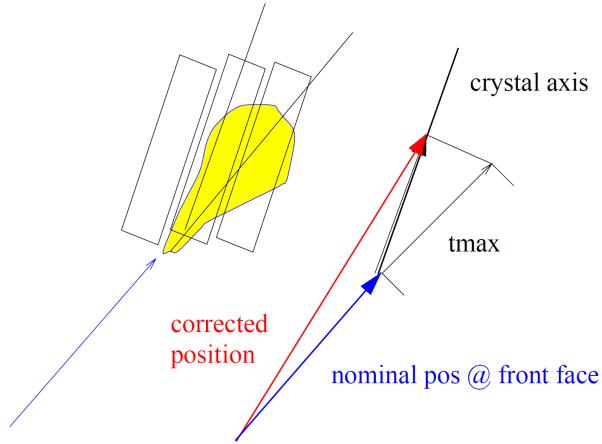


Figure 2.4: Illustration of the crystal offpointing [4].

$$W_i = W_0 + \ln\left(\frac{E_i}{\sum E_j}\right) \quad (2.2)$$

where the weight W_i is constrained to be positive otherwise it is set to zero. W_0 controls the smallest fractional energy that a crystal can have and still contribute to the position measurement. Its default value which is obtained after optimization studies is 4.2, so that crystals in the cluster containing more than 1.5% of the cluster energy will contribute to the position measurement.

2.1.4 Electron reconstruction

In order to reconstruct electron we need to find a track in the tracker which match to the SC and of course this is not required for photon. To search for the track, four steps are involved: the track seed selection, the track building and the track fitting (the last two are usually referred to as “tracking”) and the track SC matching. They are described below.

2.1.4.1 Track seeding

Track seeds are the starting point for the track reconstruction which are built from doublets or triplets of hits in the pixel detector. There are two different approaches for track seeding: the ECAL driven seeding and the tracker driven seeding.

For the ECAL driven seeding, the procedure starts from a SC in ECAL, with at least 4 GeV of transverse energy and a veto of 0.15 on the ratio of hadronic energy to SC energy. Hits in the pixel layers are predicted by propagation of the energy weighted mean position of the SC (see Section 2.1.3) backward through the magnetic field under both charge hypotheses towards the pixel detector. The reason for this step is that the SC and pixel matching takes advantage of the fact that the energy weighted average impact point of the electron and associated bremsstrahlung photons, as calculated using information from the SC in the ECAL, coincides (assuming a successful collection of photons) with the impact point that would have been measured for a non-radiating electron of the same initial momentum. It is this space-point that the position measurement of the SC attempts to determine. This point can be propagated back through the field to obtain an estimate of the direction of the electron at the vertex, and the hit positions expected in the pixel

detector. Since most of the tracker material lies after the pixel detector, most electrons do not radiate significantly before it, and most photon conversions take place after it.

Next a first compatible hit is looked for in the innermost pixel layer within a loose $\Delta\phi$ window and loose Δz interval, when a first compatible hit is found a new estimate for the z coordinate of the primary track vertex is calculated combining the found pixel hit and calorimetry information in the Rz plane. The predicted trajectory is then propagated to look for a second pixel hit in the next pixel layer(s), within some narrower $\Delta\phi$ and Δz windows. If the first two hits are matched with the prediction from the SC, then the seed is selected.

For the tracker driven seeds, they are selected from tracks that were reconstructed with the Kalman filter (KF) algorithm [5]. This algorithm is not suited for electrons that emit bremsstrahlung photons since the curvature of the track changes in that case. All seeds of KF tracks that match a SC in the ECAL and pass a matching criterion are selected.

The choice of two approaches is analysis dependent, a seeding strategy could be preferred with respect to the other. For example in the search for high mass resonances decaying in dielectron final state (see Chapter 3), the ECAL driven seeding is required at selection level.

2.1.4.2 Tracking

Once the track seed has been obtained, the tracking procedure can take place. The tracking procedure consists of the “track building” which outwards from the seed and the combinatorial track finder method (CTF) [6] is used (which is an extension of the standard KF method), followed by the “track fitting” which uses a Gaussian sum filter (GSF) method [7] in a backward fit. For the track building the CTF method makes use of a specific Bethe-Heitler (BH) [8] model (modeling the electron energy losses during track building) when collecting matching hits in successive silicon layers, with a tolerance of 1 layer without hits. Since the distribution of the energy loss after the BH model is non-Gaussian, fitting the track with the KF algorithm that uses Gaussian distributions will not give good results. For this reason, the GSF algorithm models the BH energy loss distribution as a sum of six Gaussian distributions with different means, widths and amplitudes. After passing through a layer, six new trajectory components are generated with the weight according to the weight of the initial trajectory multiplied by the weight of the Gaussian component in the BH energy loss distribution estimation. To limit the maximal number of followed trajectories to 12, the ones with low weight are dropped or merged if they are similar. Finally, the track parameters obtained have their uncertainty distributed according to the sum of Gaussian distributions from the trajectory components.

One of the great benefit of the GSF tracks comes from the combined facts that hits are collected efficiently along the full trajectory through the tracker volume, and that meaningful track parameter errors are available at both track ends. Thus, a good estimation of the electron track parameters at ECAL entrance is made available. Moreover, the fractional amount of momentum carried away by bremsstrahlung photons can be evaluated from the outermost and innermost track parameters. This will be very useful information in distinguishing various electron patterns, to improve electron energy measurements and electron identification.

2.1.4.3 Track-supercluster matching

In order to build GSF electron candidates, a track has to be associated to a SC.

For ECAL driven tracks, the difference between energy weighted position of SC and the position of the track at the SC which is the extrapolated from the innermost track

should be smaller than 0.02 in the η direction and 0.15 rad in the ϕ direction.

For tracker driven tracks a multivariate technique, using a boosted decision tree (BDT), is used which combines track observables and SC observables to get a global identification variable. For a successful matching, the track-SC combination should have a value which is higher than a threshold of this variable.

A schematic view of the electron reconstruction procedure, considering also multiple bremsstrahlung emissions, is shown in Figure 2.5.

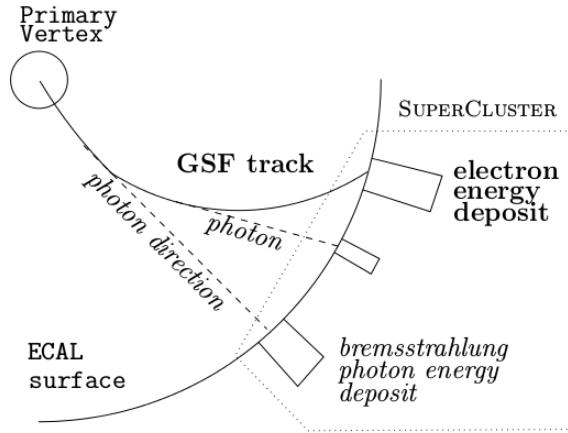


Figure 2.5: Schematic view of an electron reconstructed in CMS. The track is reconstructed by the GSF algorithm taking into account the trajectory kinks due to energetic bremsstrahlung photon emission. The energy deposits belonging to the emitted photons are collected together to the electron cluster by the clustering algorithm.

2.1.5 Photon reconstruction

The photon is the simplest electromagnetic object. Any reconstructed SC with $P_T > 10$ GeV is considered as a photon candidate [9]. In order to reduce the fake photon from other objects like electron, hadron and jet, the most important tool to be used is isolation requirement. Fake photon due to jets can usually be rejected by looking for additional energetic particles in a cone around the reconstructed ECAL cluster. Charged particles like electron, pions and kaons can be detected in the tracker or in the calorimeter. Neutral pions and other particles decaying to photons can be detected in the ECAL. The hadron calorimeter is important for detecting hadrons which do not efficiently reconstruct in the tracker, particularly at high pseudorapidity or particles like neutrons. Therefore, the basic isolation variables considered are based on charged tracks reconstructed in the tracker, electromagnetic energy deposits observed in the ECAL, and hadronic energy deposits in the HCAL.

Photon candidates which are converted (produce an electron pair) in the material upstream ECAL are tagged using an algorithm which searches for conversion tracks matching the ECAL SC, as will be described in the following subsection.

2.1.5.1 Photon conversion

A large number of photons originating from the primary interaction vertex convert in the tracker material. Identification of converted photons allows a better choice of energy clustering algorithm. The approach to reconstruct conversion track pairs are similar with

standard tracks reconstruction as described before. The energy deposits in the ECAL are used as a starting point for inward track seed finding. For each cluster both positive and negative charge hypotheses are tried. The final opposite-charge track pairs are combined together and asked to satisfy the photon conversion topology.

Before finishing this section it is worth mention that at analysis level many other variables (like shower shape, isolation and so on) will be used to distinguish electron and photon from jet efficiently. Several variables are helpful in this task and can be combined in a set of selection criteria used by the CMS analyses.

2.2 Muon reconstruction

The muon reconstruction use the information from muon system and silicon tracker. For the muon which is reconstructed only using muon system information is called “standalone muon”. The muon which is reconstructed by combining the information coming from the muon system inwards to the inner tracker are referred to as “global muon candidates”. Finally, muon candidates reconstructed by combining the information coming from the inner tracker outward to the muon station are referred to as “tracker muon candidates”. The global muon reconstruction is especially efficient for muons leaving hits in several muon stations, the tracker muon reconstruction is more efficient for low P_T muon candidates. The efficiency for reconstructing a muon as global or tracker muon is as high as 99%.

2.2.1 Standalone muon reconstruction

The standalone muon reconstruction uses only information from the muon detectors (the silicon tracker is not used). Both tracking detectors (DT and CSC) and RPCs (see Section 1.2.6) participate in the reconstruction. Despite the coarser spatial resolution, the RPCs complement the tracking chambers, especially where the geometrical coverage is problematic (mostly in the barrel-endcap overlap region). The reconstruction of muon tracks in the muon system starts from seeds, generated by the DT and the CSC, which are fitted to produce track segments. The state vectors (track position, momentum, and direction) associated with the segments found in the innermost chambers are used to seed the muon trajectories, working from inside to outside using the KF technique. The predicted state vector at the next measurement surface is compared with existing measurements and will be updated accordingly. In case no matching hits are found (e.g. due to detector inefficiencies, geometrical cracks) the search is continued in the next station. The track parameters and the corresponding errors are updated at each step. The procedure is iterated until the outermost measurement surface of the muon system is reached. Then a backward KF is applied, working from outside to inside and the track parameters are defined at the innermost muon station. Finally, the track is extrapolated to the nominal interaction point and a vertex-constrained fit to the track parameters is performed.

2.2.2 Global muon reconstruction

The reconstruction of global muon consists extending the standalone muon to include hits in the silicon tracker. It starts with extrapolating muon trajectory from the innermost muon station to the outer tracker surface and taking into account the muon energy loss in the material as well as the effect of multiple scattering. Silicon layers compatible with the muon trajectory are then determined and a region of interest (ROI) within them is

defined in which the track reconstruction will be performed. The determination of the ROI is based on the standalone muon parameters and their corresponding uncertainties. The a well defined ROI can improve the reconstruction efficiency, reduce fake rate and save CPU reconstruction time.

Inside the ROI, initial candidates for the muon track (regional seeds) are built from pairs of reconstructed hits. Starting from the regional seeds, a track-reconstruction algorithm, based on the KF technique is used to reconstruct tracks from outside to inside. After the tracker track is obtained, it will be asked to match with standalone muon. If there is a suitable match between the tracker track and the standalone muon track, a final fit is performed all over the hits (otherwise no global muon is produced). However it is also possible to combine only a subset of the hits for the final fit. In particular, choosing a subset of the muon hits provides a better momentum resolution for high energy muons, when the measurements in the muon system are frequently contaminated by electromagnetic showers. Finally, the global muon reconstruction ends with the matching of the global muon track and the energy deposits in the calorimeters.

2.2.3 Tracker muon reconstruction

For low P_T muons (below 6-7 GeV), large fraction of them does not leave enough hits in the muon detector to be reconstructed as standalone muons. In addition, some muons can escape in the gap between the wheels. Therefore a complementary approach which starts from the tracker tracks to reconstruct muon has been designed [10] and hence improve the muon reconstruction efficiency. The algorithm starts extrapolating each reconstructed tracker track outward to other subdetectors like ECAL, HCAL and muon system. After collecting the associated signals from each subdetector, the algorithm gives a decision if the tracker track is produced by a muon according to how well the observed signals fit with the hypothesis. By the way, if the extrapolated track matches at least one muon segment in the muon detector, the track is qualified as a “tracker muon”.

2.3 Particle-flow algorithm

Before moving forward to describe remaining objects reconstruction, it is needed to give a brief introduction about the particle-flow (PF) algorithm [11, 12] in CMS. This algorithm allows to reconstruct stable particles which are called “PF particles” (including electron, photon, muon, charged and neutral hadrons) using the combined information from all subdetectors under the form of calorimeter clusters and tracks (individually denoted as PF elements).

Since a single particle can create multiple PF elements, such as a track and several calorimeter clusters. Therefore, a link algorithm is used to link different PF elements to fully reconstruct a PF particle and limit double counting, via computing the distance between the tracks and the calorimeter clusters to determine whether they correspond to the same physical objects. Finally, the individual PF particles are then combined to form more complex objects such as hadronically decaying taus, jets or transverse missing energy.

2.4 Tau reconstruction

Good tau lepton identification is very important for analyses at the LHC that involve searches for Higgs bosons or evidence of SUSY. While tau leptons have a very short lifetime ($2.9 \times 10^{-13} s$) and decay right away when they are produced. Therefore, tau leptons

are reconstructed from their visible decay products. In the case of a leptonic decay (about one third of the cases), the electron or the muon is the only reconstructed particle and the standard tools for electron and muon reconstruction (described in the previous sections) are used. For the remaining two thirds of cases, taus decay hadronically with the production of a jet with very specific features which contain one or three charged hadrons (mostly pions) and several neutral pions which decay into two photons immediately. Moreover, jets coming from tau decays are on average more collimated and isolated than quark and gluon jets. The taus with hadronically decay are reconstructed and identified using the hadrons-plus-strips (HPS) algorithm [13, 14]. The algorithm is designed to reconstruct individual decay modes of the tau by taking advantage of the excellent performance of the PF algorithm in reconstructing individual charged and neutral particles. The HPS algorithm is performed in two steps:

1. **Reconstruction** : combining charged and neutral particles reconstructed by the PF algorithm that are compatible with specific tau decays and the four-momentum which is expressed in terms of (P_T , η , ϕ , and mass) of tau candidates is computed.
2. **Identification** : discriminators that separate tau decays from quark and gluon jets, and from electrons and muons, are computed. This provides a reduction of misidentified taus from the jet, electron and muon.

The efficiency in identifying hadronically decaying taus is typically between 45% and 70% for misidentification rates of a jet as a tau of the order of 1%. There are some reasons for the hard reconstruction of tau: in LHC the cross section of QCD process is several order of magnitude larger than tau's and the jet is the most important background. Since at least one neutrino is produced during the tau decay, only a fraction of its total momentum is measured. Although tau has lower reconstruction efficiency compared with electron and muon, it is still very important in some analysis, e.g. in Higgs boson related analysis, because the branch ratio of Higgs decays to tau pair is much larger than Higgs decays to electron or muon pair, it makes tau being the most sensitive lepton in Higgs related analysis.

2.5 Jet and bjet reconstruction

The jet producing cross section is the largest one at the LHC. A well designed jet reconstruction algorithm is important because jets will not only provide a benchmark for understanding the detector, but will also serve as an important tool in the search for physics beyond the standard model (e.g. searching Z' in di-jets final state). Jets coming from quarks and gluons are reconstructed by PF objects. There are several jet reconstruction methods and the most commonly used is the anti- k_T algorithm [15]. In order to cluster PF objects into the jets, this algorithm defines two variables, one is the distance (d_{ij}) between the PF object (i) and the pseudojets (j), another is the distance (d_{iB}) between the PF object (i) and the beam (B):

$$d_{ij} = \min(k_{ti}^{-2}, k_{tj}^{-2}) \frac{\Delta_{ij}^2}{R^2} \quad (2.3)$$

$$d_{iB} = k_{ti}^{-2} \quad (2.4)$$

where k_{ti} and k_{tj} are the transverse momenta of the i and j entities respectively, $\Delta_{ij}^2 = (\phi_i - \phi_j)^2 + (\eta_i - \eta_j)^2$, and R is the cone size parameter which can be chosen in the range from 0.4-0.7 depending on the analyses choices. If the smallest distance is of d_{ij} -type, the entities i and j are combined into a new single entity, while if it is of d_{iB} -type,

the i entity is considered as a jet and removed from the list of entities. The procedure continues until the entity list is empty. Unlike other jet clustering algorithms, the anti- k_T algorithm produces jets with a conical shape, clustered around the hardest particles and with boundaries resilient with respect to soft radiation.

The raw jet energies are corrected to ensure a uniform response in η and an absolute calibration in P_T . The target of the calibration is that the reconstructed jet energy matches the energy of the generated jet. The correction to the raw P_T of the jet can be decomposed in four multiplicative terms [15]:

- An offset correction, to remove the energy due to particles not involved in the hard-scattering process (pile-up, detector noise);
- A MC calibration factor, which corrects the reconstructed energy to match the generated MC particle jet energy, based on simulations;
- A residual calibration for the relative energy scale, to correct the energy response as a function of the pseudorapidity;
- A residual calibration for the absolute energy scale, to make the energy response uniform as a function of the transverse momentum.

The final energy resolution for a jet of 100 GeV of P_T is around 10%.

2.5.1 b-jet reconstruction

All the quarks will generate a jet except the top quark which decays before hadronizing. However, the jets from b quark can be distinguished from other jets coming from gluons, light-flavor quarks (u, d, s) and c quark fragmentation using track, vertex and identified lepton information. There are different algorithms to tag b-jets but only the Combined Secondary Vertex (CSV) algorithm is described here since it is the most commonly used in the CMS analyses. Because b hadrons typically have a lifetime of $c\tau \simeq 450 \mu\text{m}$, a powerful handle to discriminate between b jets and other jets is the existence of a secondary vertex. A secondary vertex is defined as a vertex sharing less than 65% of its tracks with the primary vertex and separated radially from the primary vertex with a significance at least 3σ . In addition, if the radial distance exceeds 2.5 cm and if the mass is compatible with a K^0 or greater than 6.5 GeV, the secondary vertex is rejected. The last condition for secondary vertices is that the flight direction of each candidate is in a cone with $\Delta R = 0.5$ around the jet direction. In case no secondary vertex is found (about 35% of cases for real b-jets), the CSV algorithm can use so-called “pseudo-vertices”, from tracks whose impact parameter is more than 2σ away. If no pseudo-vertex is found, the CSV algorithm proceeds from simple track variables. The list of variables used by the CSV algorithm to identify b-jets are summarized as follows [16]:

- The presence of a secondary vertex, a pseudo-vertex or none of them;
- The flight distance significance between the primary and the secondary (or pseudo-) vertex in the transverse plane;
- The number of tracks at the secondary or pseudo-vertex;
- The ratio of the energy carried by tracks at the vertex with respect to all tracks in the jet;
- The pseudorapidities of the tracks at the vertex with respect to the jet axis;
- The number of tracks in the jet.

A likelihood ratio to reject c-jets and to reject light-parton jets are combined to form the final CSV discriminator. The efficiency of the CSV algorithm in data and simulations for the medium working point is close to 70% with a mistagging rate of about 1.5%.

2.6 MET reconstruction

Neutrinos and other hypothetical neutral weakly interacting particles cannot be detected by CMS. However, some information about their presence can be obtained from the detection of a momentum imbalance in the transverse plane to the beam axis. The missing transverse energy is noted \vec{E}_T , while its magnitude is referred to as E_T . The measurement of E_T is strongly relies on the reconstruction of all other physics objects and is sensitive to a wide range of effects: mismeasurement or misidentification of physics objects, detector noise or malfunctions, pileup interactions and so on.

There are several types of \vec{E}_T in CMS and the most widely used is the particle-flow (PF) \vec{E}_T , which is the negative vectorial sum over the transverse momenta of all PF particles. A bias in the E_T measurement can be caused by several reasons, such as the non-linearity of the response of the calorimeter for hadronic particles, or the minimum energy thresholds in the calorimeters. In addition, another bias comes from the pile-up interactions and it can be corrected by subtracting the \vec{E}_T by a certain fraction ($f(\vec{v})$) of \vec{v} which is the vectorial P_T sum of charged particles associated to each pile-up vertex:

$$\vec{E}_T^{corr} = \vec{E}_T - \sum_{PU} f(\vec{v})\vec{v} \quad (2.5)$$

Finally, an asymmetry in the ϕ variable is observed in data and simulated events, and is found to be related to the number of reconstructed vertices N_{vtx} . The \vec{E}_T projection in the x - and y -directions are independently corrected by some functions of N_{vtx} .

As mentioned before, the \vec{E}_T measurement [17] strongly relies on the reconstruction of all other physics objects, and is sensitive to a wide range of effects. Artificially large E_T can be measured because of spurious detector signals. Sources of fake E_T include:

- Dead cells in the ECAL;
- Beam-halo particles;
- Particles striking sensors in the ECAL barrel detector;
- Noise from HCAL hybrid photodiode;
- Direct particle interactions with light guides and photomultipliers tubes in the forward calorimeter;
- High-amplitude anomalous pulses in the ECAL endcaps;
- A misfire of the HCAL laser calibration system;
- A defective track reconstruction, from coherent noise in the silicon strip tracker.

Dedicated algorithms are used to identify and remove these events with fake E_T .

2.7 Summary

This chapter describes how the objects used for analyses are reconstructed in CMS. The electron and photon reconstruction has been described in details in Section 2.1, while the muon, tau, jet and missing transverse energy reconstructions are described in Sections 2.2-2.6.

Chapter 3

Searching for high mass resonances in dielectron final state

In order to address the shortcoming of the standard model (SM), there are several theories beyond the SM predicting the existence of heavy resonances at the TeV scale that can couple to quarks or gluons and can decay to dilepton pairs. Examples include a heavy Z boson-like particle, denoted Z' , such as the Z'_{SSM} boson of the sequential standard model with SM-like couplings, or the Z'_ψ boson which inspired by superstring models and the spin 2 Kaluza-Klein graviton which comes from the Randall-Sundrum (RS) model with one extra dimension. This analysis uses proton-proton collision data at the centre-of-mass energy of 13 TeV collected by the CMS experiment at the LHC in 2016 and 2017, corresponding to an integrated luminosity of 35.9fb^{-1} and 41.4fb^{-1} respectively. The strategy of the analysis is looking for a “bump” in the dielectron invariant mass distribution, in particular in the high mass tail.

3.1 Data and MC samples

The name and integrated luminosity of the all datasets used in this analysis are summarized in Table 3.1. The DoubleEG dataset requires at least two trigger level electrons or photons for each event and it is used for main analysis. The SingleElectron dataset requires at least one trigger level electron for each event and it is used for trigger efficiency and electron selection efficiency measurement. The SingleMuon dataset requires at least one trigger level muon for each event and it is used for $t\bar{t}$ background cross check using ‘e μ ’ method. The SinglePhoton dataset requires at least one trigger level photon for each event and it is used for fake electron study. From eras 2016B through 2016G and 2017B through 2017F the re-reconstruction datasets are used, the prompt reconstruction datasets are used for era 2016H. The total integrated luminosity of the data sample is 35.9fb^{-1} and 41.4fb^{-1} collected by the CMS experiment in 2016 and 2017 respectively. Only certified data which recommended for physics analysis is used.

The Monte Calor (MC) simulation samples used in the main analysis for 2016 and 2017 are summarized in Table 3.2 with the corresponding cross section and the precision of the cross section. It is organized as follows for 2016 MC samples: the top part of the samples are for Drell-Yan (DY) process simulation which is the main background in this analysis, then it is followed by $t\bar{t}$ process simulation samples and then is for di-boson (WW, WZ, ZZ) process simulation samples, finally is for gravitation signal simulation sample. Similar organization for 2017 MC samples. The 2016 MC samples are produced from RunIISummer16MiniAODv2*PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6 campaign and for 2017 it is from RunIIFall17MiniAOD-94X_mc2017_realistic_v10_v1(v2)

Table 3.1: Datasets (X) used in this analysis. X = DoubleEG is for the main analysis, X = SingleElectron is for trigger and electron selection efficiency, X = SinglePhoton is for the fake electron study and X = SingleMuon is for e μ study.

Year	Datasets	Integrated luminosity (fb^{-1})
2016	/X/Run2016B-03Feb2017_ver2-v2/MINIAOD	5.788
	/X/Run2016C-03Feb2017-v1/MINIAOD	2.573
	/X/Run2016D-03Feb2017-v1/MINIAOD	4.248
	/X/Run2016E-03Feb2017-v1/MINIAOD	4.009
	/X/Run2016F-03Feb2017-v1/MINIAOD	3.102
	/X/Run2016G-03Feb2017-v1/MINIAOD	7.540
	/X/Run2016H-03Feb2017_ver2-v1/MINIAOD	8.391
	/X/Run2016H-03Feb2017_ver3-v1/MINIAOD	0.215
	Sum 2016	35.867
2017	/X/Run2017B-17Nov2017-v1/MINIAOD	4.802
	/X/Run2017C-17Nov2017-v1/MINIAOD	9.629
	/X/Run2017D-17Nov2017-v1/MINIAOD	4.235
	/X/Run2017E-17Nov2017-v1/MINIAOD	9.268
	/X/Run2017F-17Nov2017-v1/MINIAOD	13.433
	Sum 2017	41.368

campaign. The most samples are generated by POWHEG v2 [18, 19, 20, 21, 22, 23] at next-to-leading order (NLO), few are generated by MadGraph5_aMC@NLO [24] at NLO. The Pythia8 [25] is used to simulate the parton showering and hadronization. For detector response it is simulated by Geant4 [26].

The pile up distributions for MC and data which is calculated by using 69.2 mb as the minimum bias cross section are shown in Figure 3.1 for 2016 and 2017. The average number of pile up in 2016 (2017) is 27 (33). MC events are re-weighted to account for the pile up difference between data and MC.

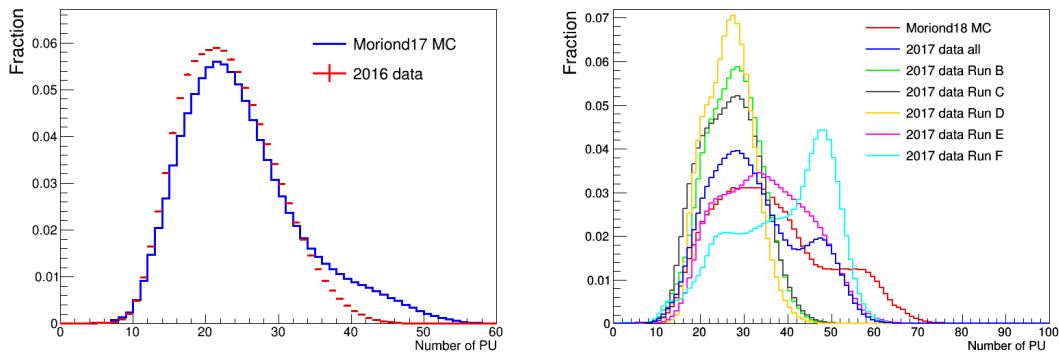


Figure 3.1: Pileup distribution for data and MC samples in 2016 (left) and in 2017 (right).

In order to improve data-mc agreement, in 2016 the data energy scale has been corrected by 1.0012 in the barrel and 1.0089 in the endcap using the mean values of the official EGamma scale corrections (except in Section 3.4 study which we measured the mean data energy correction and found it agree with offical EGamma value). In 2017 the official EGamma energy scale in data and energy smearing in MC is applied in all studies.

Year	Sample	xsection(pb)	xs precision
2016	ZToEE_NNPDF30_13TeV-powheg_M_50_120	1975	NLO
	ZToEE_NNPDF30_13TeV-powheg_M_120_200	19.32	NLO
	ZToEE_NNPDF30_13TeV-powheg_M_200_400	2.73	NLO
	ZToEE_NNPDF30_13TeV-powheg_M_400_800	0.241	NLO
	ZToEE_NNPDF30_13TeV-powheg_M_800_1400	1.68E-2	NLO
	ZToEE_NNPDF30_13TeV-powheg_M_14000_2300	1.39E-3	NLO
	ZToEE_NNPDF30_13TeV-powheg_M_2300_3500	8.948E-5	NLO
	ZToEE_NNPDF30_13TeV-powheg_M_3500_4500	4.135E-6	NLO
	ZToEE_NNPDF30_13TeV-powheg_M_4500_6000	4.56E-7	NLO
	ZToEE_NNPDF30_13TeV-powheg_M_6000_Inf	2.06E-8	NLO
	DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8 (for Z → ττ)	5765.4	NNLO
	TTTo2L2Nu_TuneCUETP8M2_ttHtranche3_13TeV-powheg	87.31	NNLO
	TTTToLL_MLL_500To800_TuneCUETP8M1_13TeV-powheg-pythia8	0.326	NLO
	TTTToLL_MLL_800To1200_TuneCUETP8M1_13TeV-powheg-pythia8	3.26E-2	NLO
	TTTToLL_MLL_1200To1800_TuneCUETP8M1_13TeV-powheg-pythia8	3.05E-3	NLO
	TTTToLL_MLL_1800ToInf_TuneCUETP8M1_13TeV-powheg-pythia8	1.74E-4	NLO
	ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1/	19.47	app.NNLO
	ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1/	19.47	app.NNLO
	WWTo2L2Nu_13TeV-powheg	12.178	NNLO
	WWTo2L2Nu_Mll_200To600_13TeV-powheg	1.39	NNLO
	WWTo2L2Nu_Mll_600To1200_13TeV-powheg	5.7E-2	NNLO
	WWTo2L2Nu_Mll_1200To2500_13TeV-powheg	3.6E-3	NNLO
	WWTo2L2Nu_Mll_2500ToInf_13TeV-powheg	5.4E-5	NNLO
	WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8	4.42965	NLO
	WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.595	NLO
	ZZTo2L2Nu_13TeV_powheg_pythia8	0.564	NLO
	ZZTo4L_13TeV_powheg_pythia8	1.212	NLO
	ZZTo2L2Q_13TeV_powheg_pythia8	1.999	NLO
	RSGravToEEMuMu_kMpl-001_M-*_TuneCUETP8M1_13TeV-pythia8	-	-
2017	ZToEE_NNPDF31_13TeV-powheg_M_50_120	1975	NLO
	ZToEE_NNPDF31_13TeV-powheg_M_120_200	19.32	NLO
	ZToEE_NNPDF31_13TeV-powheg_M_200_400	2.73	NLO
	ZToEE_NNPDF31_13TeV-powheg_M_400_800	0.241	NLO
	ZToEE_NNPDF31_13TeV-powheg_M_800_1400	1.68E-2	NLO
	ZToEE_NNPDF31_13TeV-powheg_M_14000_2300	1.39E-3	NLO
	ZToEE_NNPDF31_13TeV-powheg_M_2300_3500	8.948E-5	NLO
	ZToEE_NNPDF31_13TeV-powheg_M_3500_4500	4.135E-6	NLO
	ZToEE_NNPDF31_13TeV-powheg_M_4500_6000	4.56E-7	NLO
	ZToEE_NNPDF31_13TeV-powheg_M_6000_Inf	2.06E-8	NLO
	DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8 (for Z → ττ)	5765.4	NNLO
	TTTo2L2Nu_TuneCP5_PSweights_13TeV-powheg-pythia8	87.31	NNLO
	ST_tW_top_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	19.47	app.NNLO
	ST_tW_antitop_5f_NoFullyHadronicDecays_TuneCP5_13TeV-powheg-pythia8	19.47	app.NNLO
2018	WW_TuneCP5_13TeV-pythia8	118.7	NLO
	WZ_TuneCP5_13TeV-pythia8	47.13	NLO
	ZZ_TuneCP5_13TeV-pythia8	16.523	NLO

Table 3.2: MC samples used in main analysis

3.2 Trigger

The primary high level trigger (HLT) used in main analysis for 2016 and 2017 is HLT_DoubleEle33_CaloIdL_MW which requires two electron candidates with E_T of the supercluster higher than 33 GeV and passing loose calorimeter identification (CaloIdL) requirements and Medium Window (MW) matching between the gaussian sum filter (gsf) [27] track and the hits in pixel detector. In the run period of 276453 to 278822 of 2016 this trigger was prescaled and HLT_DoubleEle33_CaloIdL_GsfTrkIdVL was used as the primary signal trigger. The HLT_DoubleEle33_CaloIdL_GsfTrkIdVL is the same as HLT_DoubleEle33_CaloIdL_MW except replacing MW pixel matching by very loose matching between gsf track and the supercluster in ECAL (GsfTrkIdVL).

The level 1 trigger (L1) seeding of the primary high level trigger is always seeded by the OR of a DoubleEG (two deposit in ECAL) seed, a SingleEG (one deposit in ECAL) seed and a SingleJet (one L1 object compatible with a jet) seed, after run 275319 in 2016 and in full 2017 it is also seeded by a SingleTau (one L1 object compatible with a τ) seed. The presence of the SingleJet and SingleTau seeds are mean to mitigate the loss of efficiency for high E_T electron. The exact unprescaled threshold of each of those seeds changing in time. The lowest threshold for the SingleEG which was always unprescaled was 40 GeV with the corresponding thresholds for the DoubleEG seed being 24 GeV, 17 GeV.

The trigger efficiency will be split in L1 trigger efficiency, HLT supercluster E_T filter efficiency (the HLT turn on curve) and online electron identification (CaloIdL+MW or GsfTrkIdVL) efficiency components. For final result only E_T dependent efficiency will be used to weight MC events, others will be cancel in the normalisation of MC events to data in the Z peak region (M_{ee} in 60-120 GeV). The method to measure the efficiency will be described in Section 3.2.1. The L1 trigger efficiency of primary signal trigger will be shown in Section 3.2.2. The HLT E_T turn on curve and HLT identification efficiency of primary signal trigger will be shown in Section 3.2.3, Other trigger efficiencies will be shown in Section 3.2.4.

3.2.1 Method for Measuring Trigger Efficiencies in Data

The tag and probe method [28] is used to measure the efficiency in data. The event is selected by HLT_Ele27_eta2p1_WPTight (require one HLT electron candidate with supercluster E_T higher than 27 GeV and $|\eta|$ less than 2.1 and passing tight online electron cut) for 2016 and HLT_Ele35_WPTight (require one HLT electron candidate with supercluster E_T higher than 35 GeV and passing tight online electron cut) for 2017. The tag is the electron which passing HLT_Ele27_eta2p1_WPTight (HLT_Ele35_WPTight) in 2016 (2017) and passing HEEP ID (as defined in Section 3.3) and in barrel of ECAL. The probe is the electron which passing the HEEP ID as well as any other requirements necessary to measure the given efficiency such as being matched to the E_T filter to measure the trigger identification efficiency.

To simplify the computation, tags can not be probes. In the case of the probe being in the barrel, the tag is required to have a smaller supercluster ϕ than the probe for even number events and a larger supercluster ϕ for odd number events. As the sample is already very pure given there are two electrons passing HEEP ID, no background subtraction is applied, nor any mass window cut imposed. When measuring efficiencies involving the unseeded leg of the HLT_DoubleEle33_CaloIdL_MW (or HLT_DoubleEle33_CaloIdL_GsfTrkIdVL) trigger path, the tag should additionally pass the L1 seeded leg of that trigger and be matched to a L1 EG object, using a ΔR cone of 0.1 to be completely sure that the un-

seeded leg is unbiased by L1 seeded trigger. The efficiency is equal to the number of passing probes divided by all probes shown in 3.1.

$$\epsilon = \frac{N_{\text{passing probes}}}{N_{\text{all probes}}} \quad (3.1)$$

The L1 trigger efficiency and HLT turn on curves are fitted with either a single or double turn on function (defined in terms of 'error function' erf). The double turn on function is shown in equation 3.2, with the single turn on function being identical except that the B terms are removed. The A0 and B0 parameters can be interpreted as the efficiency at the plateau, the A1 and B1 as the value where the efficiency reaches half maximum and A2 and B2 are the turn on of the curve.

$$f(E_T) = 0.5 \cdot A0 \cdot (1 + \text{erf}(\frac{E_T - A1}{\sqrt{2} \cdot A2})) + 0.5 \cdot B0 \cdot (1 + \text{erf}(\frac{E_T - B1}{\sqrt{2} \cdot B2})) \quad (3.2)$$

3.2.2 Primary Signal Trigger: L1 Efficiency

In 2016 the efficiency for a HEEP electron to pass the lowest unprescaled L1 SingleEG seed is shown in Figure 3.2. From applying to MC events, this translates to an efficiency of 99.5% to select barrel-barrel and 98.8% to select barrel-endcap events in a mass range of 60 to 120 GeV and a $\sim 100\%$ efficiency above 120 GeV. This is a lower bound on the efficiency, because there is the DoubleEG L1 seed which will further increase the efficiency. So it can be assumed the L1 seed trigger efficiency is 100% with a 0.5% uncertainty in the barrel and 1.2% in the endcap.

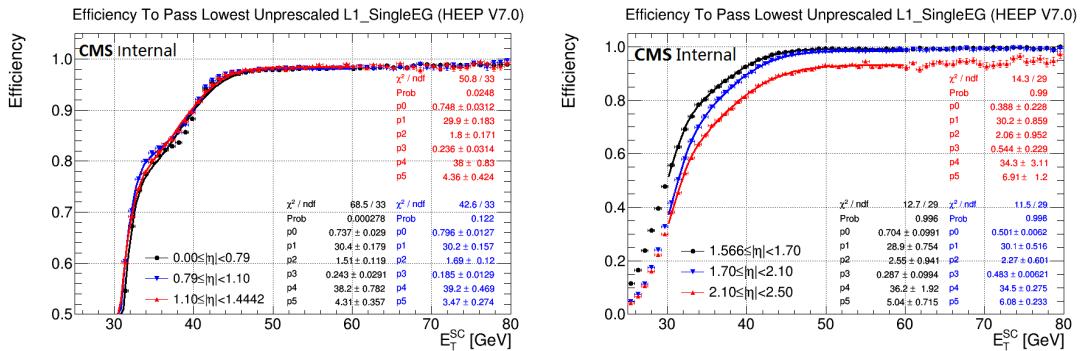


Figure 3.2: The efficiency for an electron passing HEEP to pass the lowest unprescaled L1 SingleEG seed versus supercluster E_T and η in barrel (left) and in endcap (right) for 2016 [29].

In 2017 the efficiency for a HEEP electron to pass the lowest unprescaled L1 SingleEG seed is shown in Figure 3.3. In both barrel and endcap, a slow threshold related turn on and a slow general increase in efficiency in the plateau due to increasing efficiency of the L1 ID requirements are observed. From applying to MC events, this translates to an efficiency of 71% to select barrel-barrel and 67% to select barrel-endcap events with the worst case where the supercluster E_T of both electrons is 35 GeV. The efficiency is higher than 99.5% for two electrons with supercluster E_T more than 42 GeV in barrel and more than 47 GeV in endcap. This is a lower bound on the efficiency, because there is the DoubleEG L1 seed which will further increase the efficiency.

Due to the fact that the L1 efficiency is not 100% for low E_T electrons in 2017, the L1 seeded trigger turn on is considered in the 2017 analysis as described below. In data, we require for at least one of the selected electrons to be matched with the object of L1 seed trigger filter of the HLT_DoubleEle33_CaloIdL_MW and with that object having a L1

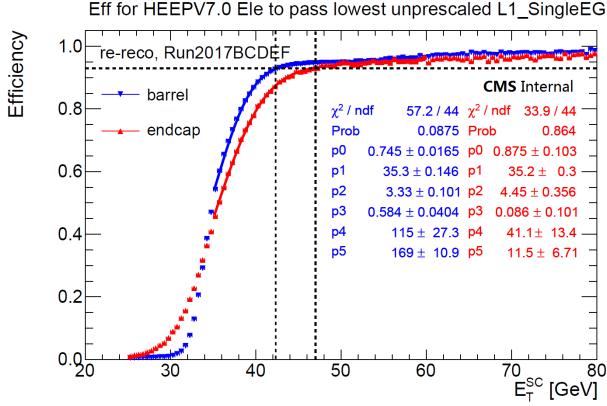


Figure 3.3: The efficiency for an electron passing HEEP to pass the lowest unprescaled L1 SingleEG seed versus supercluster E_T and η in barrel and endcap for 2017 [30].

E_T greater than the lowest unprescaled L1 SingleEG seed E_T threshold. Therefore, only L1 SingleEG seed turn on shown in Figure 3.3 is used to weight MC events. Since only one electron is seeded by L1 in HLT_DoubleEle33_CaloIdL_MW, the L1 weight value of selected MC events are shown in equation 3.3 where P_1 and P_2 are the L1 SingleEG efficiencies (shown in Figure 3.3) for leading and sub-leading selected HEEP electrons.

$$\text{weight}(L1) = 1 - (1 - P_1) \cdot (1 - P_2) = P_1 + P_2 - P_1 P_2 \quad (3.3)$$

3.2.3 Primary Signal Trigger:HLT Efficiency

The HLT efficiency is divided into two components, the efficiency of the supercluster $E_T > 33$ GeV cut (the turn on curve) and the efficiency of the CaloIdL plus MW matching (or GsfTrkIdVL) identification requirements. The turn on curves of the E_T cut for 2016 and 2017 are shown in Figure 3.4 and Figure 3.5 respectively. These turn on curves are used to weight the MC events. The efficiency of the CaloIdL plus MW matching (or GsfTrkIdVL) identification requirements for 2016 and 2017 are shown in Figure 3.6 and Figure 3.7 respectively. As the efficiencies are flat versus E_T , there is no need to weight MC events with this factor as it will automatically be included in the Z peak normalisation.

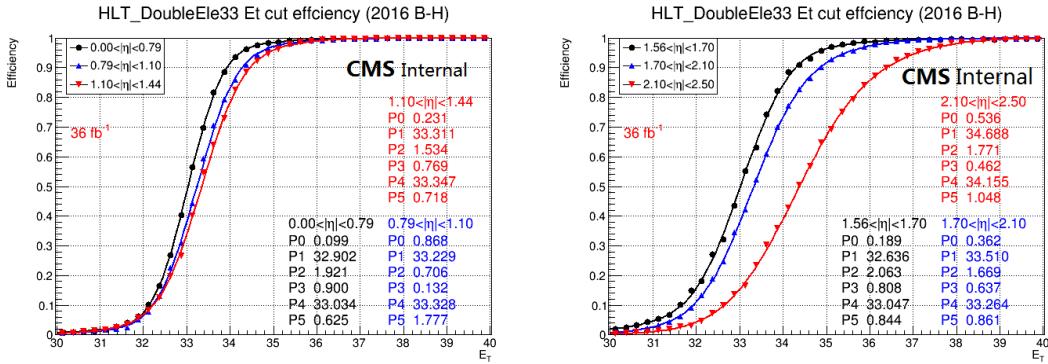


Figure 3.4: The efficiency for electron in the barrel (left) and endcap (right) passing HEEP to pass an online supercluster $E_T > 33$ GeV cut for 2016 [29].

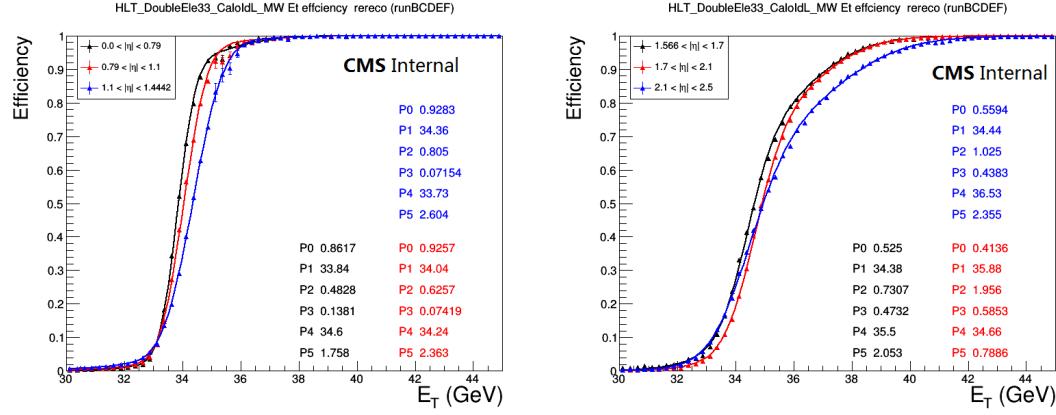


Figure 3.5: The efficiency for electron in the barrel (left) and endcap (right) passing HEEP to pass an online supercluster $E_T > 33$ GeV cut for 2017 [30].

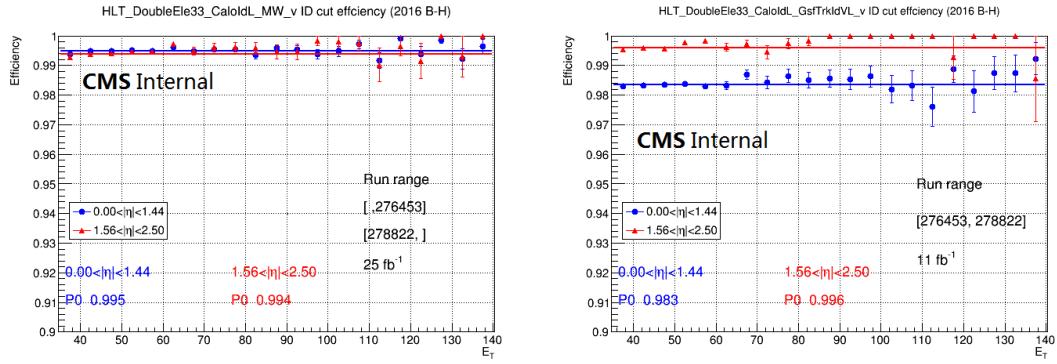


Figure 3.6: The efficiency for electron in the barrel and endcaps passing HEEP to pass the CaloIdL+MW ID requirement(left) and CaloIdL+GsfTrkIdVL ID requirement (right) for 2016 [29].

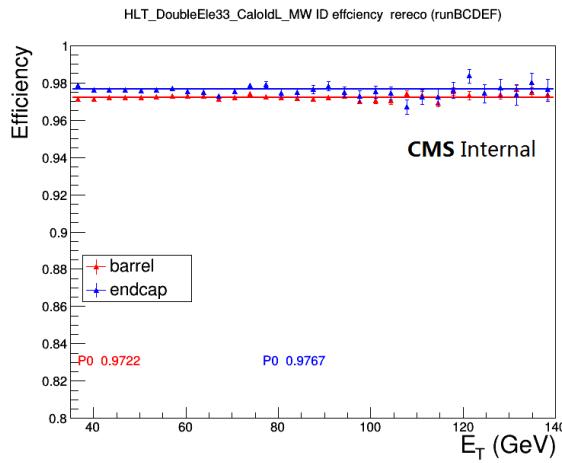


Figure 3.7: The efficiency for electron in the barrel and endcaps passing HEEP to pass the CaloIdL+MW ID requirement for 2017 [30].

3.2.4 Other Trigger Efficiencies

In 2016 the data-MC HEEP ID efficiency scale factor study uses events selected by the HLT_Ele27_eta2p1_WPTight trigger path. The efficiency of this path in data is shown

in Figure 3.8. Similar for 2017 the HLT_Ele35_WPTight trigger path is used and the efficiency of this trigger is shown in Figure 3.9. These curves are used to weight MC events to simulate the effect of the trigger requirement in data.

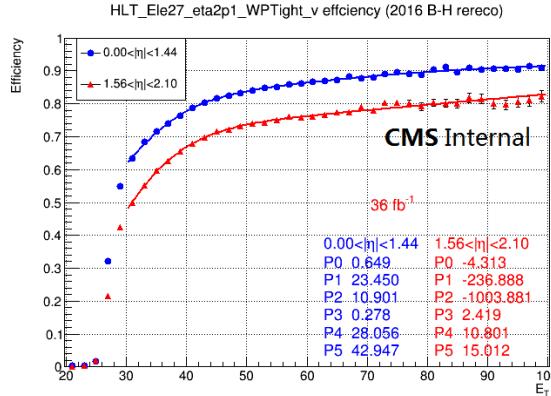


Figure 3.8: The efficiency for an electron passing HEEP to pass the HLT_Ele27_eta2p1_WPTight trigger for 2016 [29].

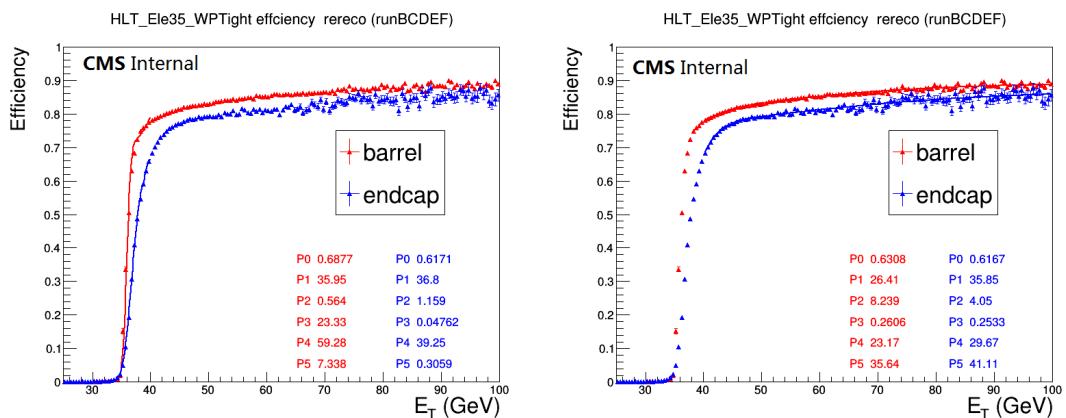


Figure 3.9: The efficiency for an electron passing HEEP to pass the HLT_Ele35_WPTight for E_T less than 40 GeV (left) and E_T more than 40 GeV (right) for 2017 [30].

3.3 Object and Event Selection

Electrons are required to pass the HEEP selection which is listed in Table 3.3 and then are combined to form dielectron candidates. If more than one dielectron candidate is found in the event, only the pair with the two largest electron E_T is retained. There is no charge requirement for dielectron candidates and this is made to avoid efficiency losses at high mass for the main analysis. Besides, at least one of the electron candidates has to be in barrel (events with both electron candidates in endcaps regions are rejected). Events in data are required to satisfy the trigger selection described in Section 3.2. MC events are weighted using turn on curves shown in figures 3.3, 3.4 and 3.5 for considering the L1 or HLT effects.

Variable	Barrel	Endcap
Acceptance selections		
E_T	$E_T > 35 \text{ GeV}$	$E_T > 35 \text{ GeV}$
η	$ \eta < 1.4442$	$1.566 < \eta < 2.5$
Identification selections		
$\Delta\eta_{in}^{seed}$	$ \Delta\eta_{in}^{seed} < 0.004$	$ \Delta\eta_{in}^{seed} < 0.006$
$\Delta\phi_{in}$	$ \Delta\phi_{in} < 0.06$	$ \Delta\phi_{in} < 0.06$
H/E	$H/E < 1/E + 0.05$	$H/E < 5/E + 0.05$
$\sigma_{i\eta i\eta}$	-	$\sigma_{i\eta i\eta} < 0.03$
$\frac{E_{1\times 5}}{E_{5\times 5}}$ and $\frac{E_{2\times 5}}{E_{5\times 5}}$	$\frac{E_{1\times 5}}{E_{5\times 5}} > 0.83$ or $\frac{E_{2\times 5}}{E_{5\times 5}} > 0.94$	-
Inner lost layer hits	lost hits ≤ 1	lost hits ≤ 1
Impact parameter d_{xy}	$ d_{xy} < 0.02 \text{ cm}$	$ d_{xy} < 0.05 \text{ cm}$
Isolation selections		
Calorimeter isolation Iso	$Iso < 2 + 0.03E_T[\text{GeV}] + 0.28\rho$	$Iso < 2.5 + 0.28\rho$ ($E_T < 50 \text{ GeV}$) else $Iso < 2.5 + 0.03(E_T[\text{GeV}] - 50) + 0.28\rho$
p_T isolation $Isopt$	$Isopt < 5 \text{ GeV}$	$Isopt < 5 \text{ GeV}$

Table 3.3: Definitions of HEEP selection cuts.

3.4 Mass resolution and scale

The mass resolution function is a crucial point of the analysis, since its outcome enters in the signal model definition. Its estimation follows two steps: a data-MC comparison, and a MC-only study. The first step consists in the comparison between the broadness of the invariant mass distribution of the electron pairs m_{ee} selected requiring the HEEP ID at the Z peak ($80 \text{ GeV} < M_{ee} < 100 \text{ GeV}$) between data and DY MC. Both distributions are fitted using a Breit-Wigner (B-W) function, whose parameters are fixed to the PDG value of the Z boson, convoluted with a double-sided crystal ball function (dCB), which is defined as a Gaussian core connected with two power-law functions on both sides. The σ parameter of the dCB function is then compared between data and MC in different η categories. For the BB category both electrons are required to be in the ECAL barrel, while for the BE category one electron is required to be in the ECAL barrel, the other one in the ECAL endcap. Fit results for data and MC in 2016 (2017) are shown in Figure 3.10 (3.11).

The σ_{MC} of the dCB which fits the MC distribution is subtracted in quadrature from the σ_{data} coming from the fit to the data, thus defining the σ_{extra} parameter through the relation: $\sigma_{extra} = \sqrt{\sigma_{data}^2 - \sigma_{MC}^2}$. In Table 3.4 the results for the σ_{extra} parameter are shown for the different categories for 2016 and 2017. Note that the numbers in Table 3.4 are expressed in percentage [%] of the Z peak mass value (M_Z DPG value of 91.1876 GeV). The quoted mean values are the ones of the dCB function used to fit the mass spectra after convolution with the B-W function. For 2017 the official EGamma energy scale and smearing is applied for data and MC, therefore the mean values difference between data and MC in 2017 are much smaller than it in 2016 and for the σ_{extra} which is 0 in 2017 because of the larger σ_{MC} than σ_{data} .

Year	Category	$\frac{\Delta M}{M} [\%]$	$\sigma_{data} [\%]$	$\sigma_{MC} [\%]$	$\sigma_{extra} [\%]$
2016	BB	-0.19 ± 0.02	1.45 ± 0.00	1.20 ± 0.03	0.81 ± 0.04
	BE	-0.40 ± 0.02	2.49 ± 0.01	2.15 ± 0.03	1.26 ± 0.05
2017	BB	0.04 ± 0.01	1.63 ± 0.01	1.74 ± 0.01	0 ± 0
	BE	-0.03 ± 0.00	2.91 ± 0.00	2.93 ± 0.00	0 ± 0

Table 3.4: Results per category for the data-MC scale shift $\frac{\Delta M}{M}$ and the σ_{extra} parameters.

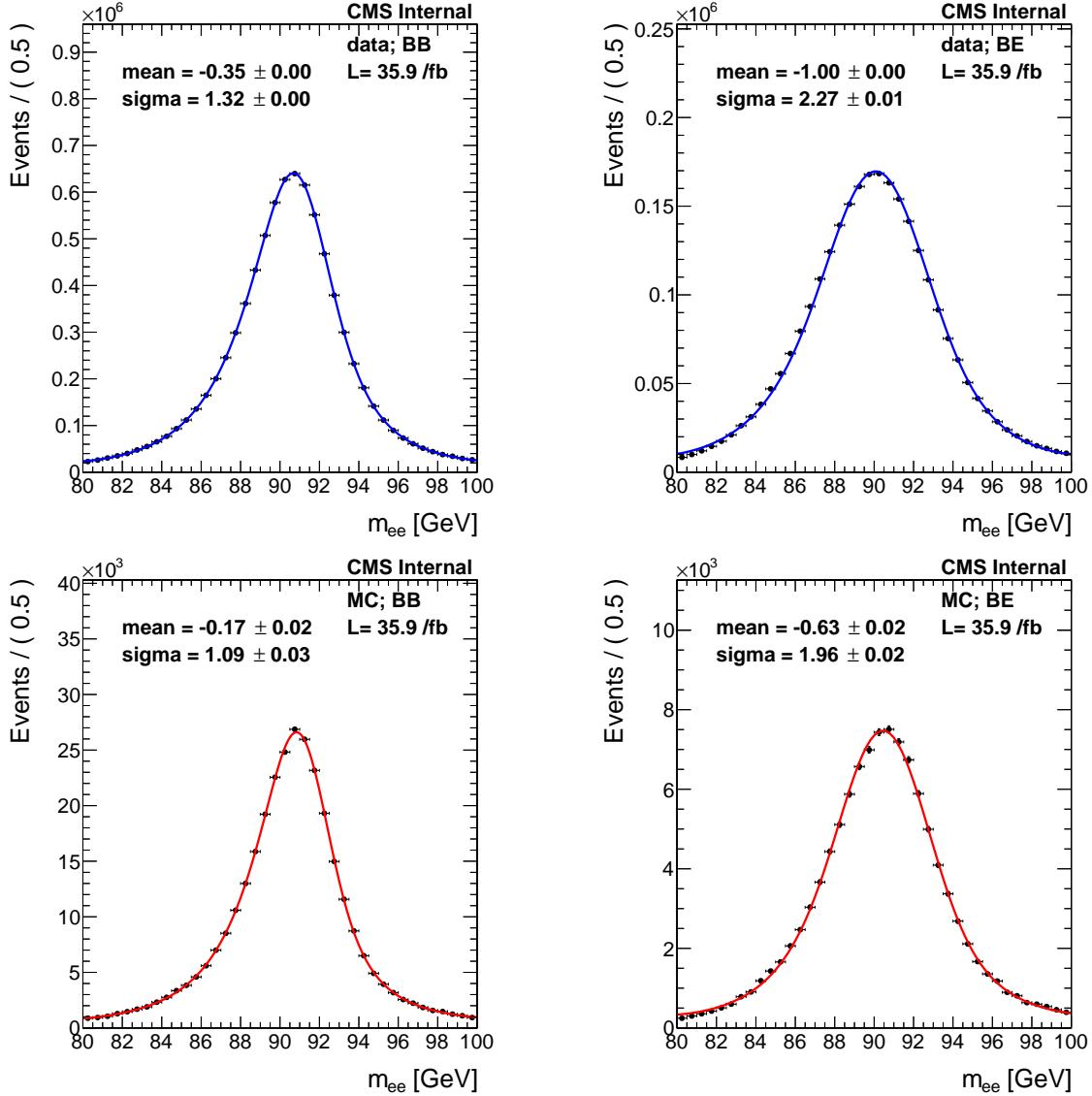


Figure 3.10: Invariant mass distribution at the Z peak in data (top) and MC (bottom) for the BB region (left) and BE (right) channel for 2016.

In 2017 we also checked mass scale and resolution for data and MC as a function of E_T , energy and η of electron using the method explained above. Results are shown from Figure 3.12 to 3.14. In all E_T and η ranges, the data and MC show good agreement. The energy scale at high E_T is validated to within 2% for barrel electrons and 1% for endcap electrons.

The second step of the study is MC-only based. In particular, the mass resolution has been studied as a function of the generated invariant mass of the electron pair m_{gen} . In order to maximise the statistics, different Drell-Yan (DY) samples are used, all of the them generated for a centre-of-mass energy $\sqrt{s} = 13$ TeV and 25 ns of bunch spacing.

For each bin of the generated invariant mass m_{gen} , the distribution of the difference between the reconstructed and the generated invariant mass, divided for the generated invariant mass is analysed. Defining the variable $resolution = \frac{m_{reco} - m_{gen}}{m_{gen}}$, its distribution is obtained as a function of m_{gen} and a binned maximum-likelihood fit is performed using a “cruijff function” (Gaussian core, connected with an exponential tail on each side) for 2016, while for 2017 the dCB function is used because it is found that the cruijff is not

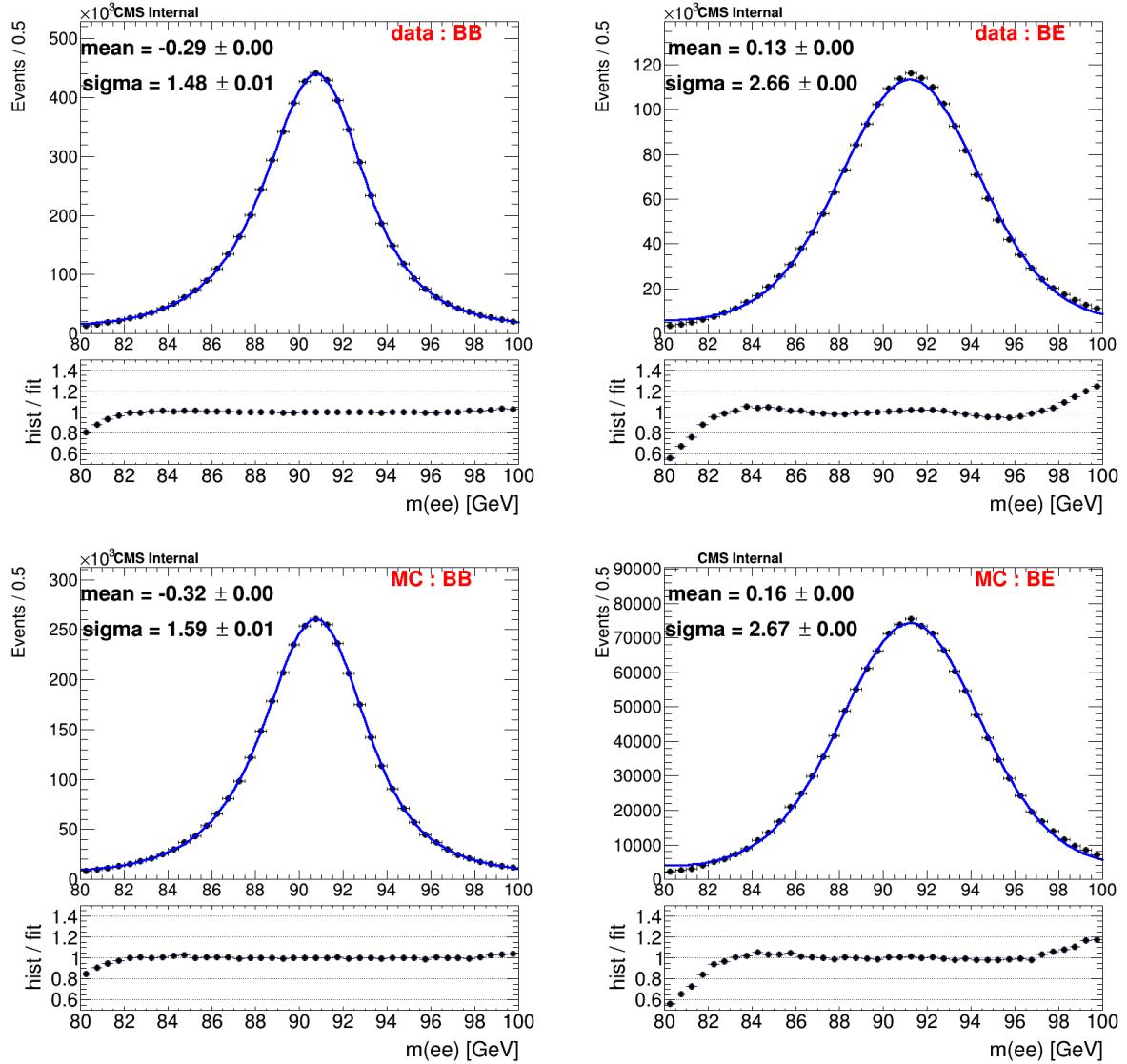


Figure 3.11: Invariant mass distribution at the Z peak in data (top) and MC (bottom) for the BB region (left) and BE (right) channel for 2017.

able to correctly describe the invariant mass shape at very high mass due to saturation effects. The fitted parameters are studied in their behaviour versus the corresponding generated mass and an analytic parametrisation is provided and used as an input for the limit setting procedure (more details are given in appendix .1). In particular, for the mass resolution, the sigma parameter of the fit function σ_{fit} is added in quadrature with the σ_{extra} parameter. Results for the BB region and BE region are shown in Figure 3.15.

In the BB region there is a small linear rise in the mass resolution starting around ≈ 1.5 TeV. The effect has been already studied in [31] and it is due to leakage of the electromagnetic ECAL shower in the HCAL subdetector which worsen in this way the energy reconstruction driven by the ECAL detector. In fact, the effect of the leakage in the HCAL subdetector is visible as an increase in the H/E variable, which is the ratio between the energy in the HCAL over the energy contained in the ECAL detector around the electron direction. For the increase of mass resolution from 4.5 TeV to 5 TeV in BB region is because the saturation effect becomes significant.

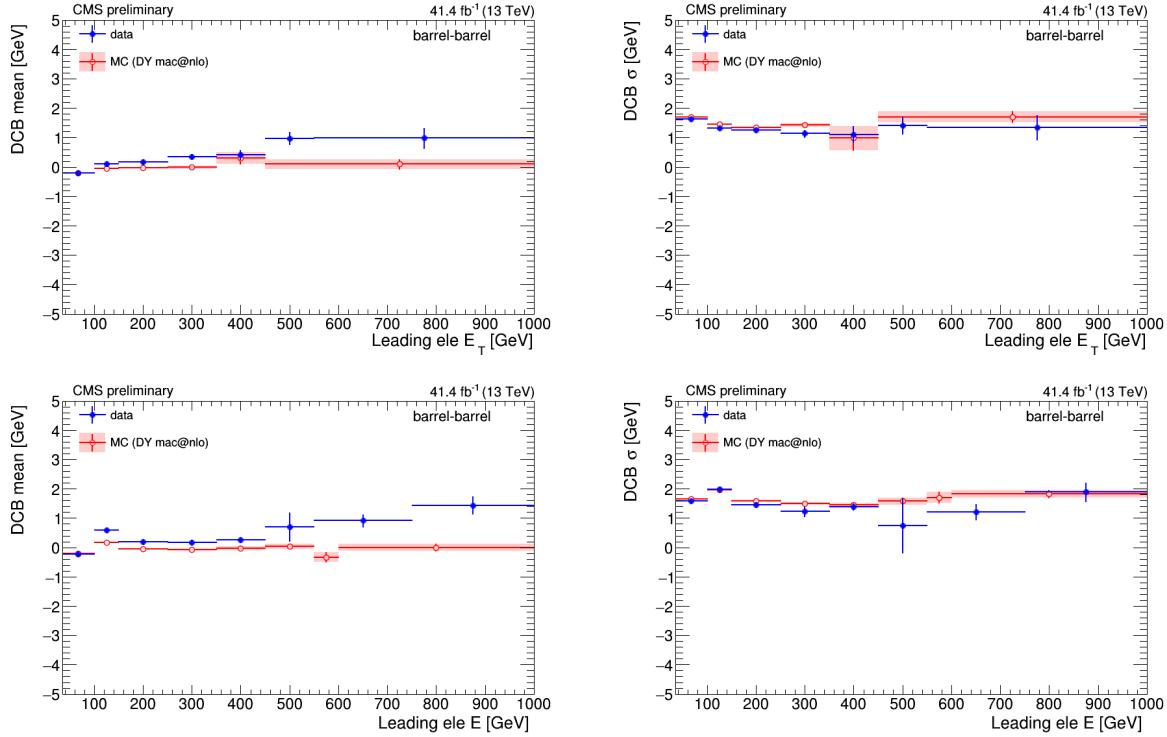


Figure 3.12: The mean (left) and sigma (right) value of dCB as the function of E_T (top) and energy (bottom) of leading electron for barrel-barrel channel.

Also the mass scale of the ECAL detector has been studied as a function of the generated invariant mass of the electron pair m_{gen} , using the same generated samples taken into account for the mass resolution determination.

For each bin of the generated invariant mass m_{gen} , the distribution of the reconstructed invariant mass, divided by the generated invariant mass is analysed. The mass scale variable is defined as $scale = \frac{m_{reco}}{m_{gen}}$. From the *resolution* variable defined above it can be seen that $scale = \frac{m_{reco}}{m_{gen}} = 1 + resolution$. The mean parameter of the crystal ball function used to fit the *resolution* distribution is taken as the mass scale simply adding the unity to it. The error on the parameter is taken accordingly.

Results for the BB region and BE region are shown in Figure 3.16.

In the BB region there is a drop in the mass scale parameter starting around ≈ 1 TeV ($\approx 0.5\%$ effect), which is the counterpart of the rise observed in the mass resolution, and is again due to leakage of a small part of the high-energetic e.m. shower in the HCAL subdetector.

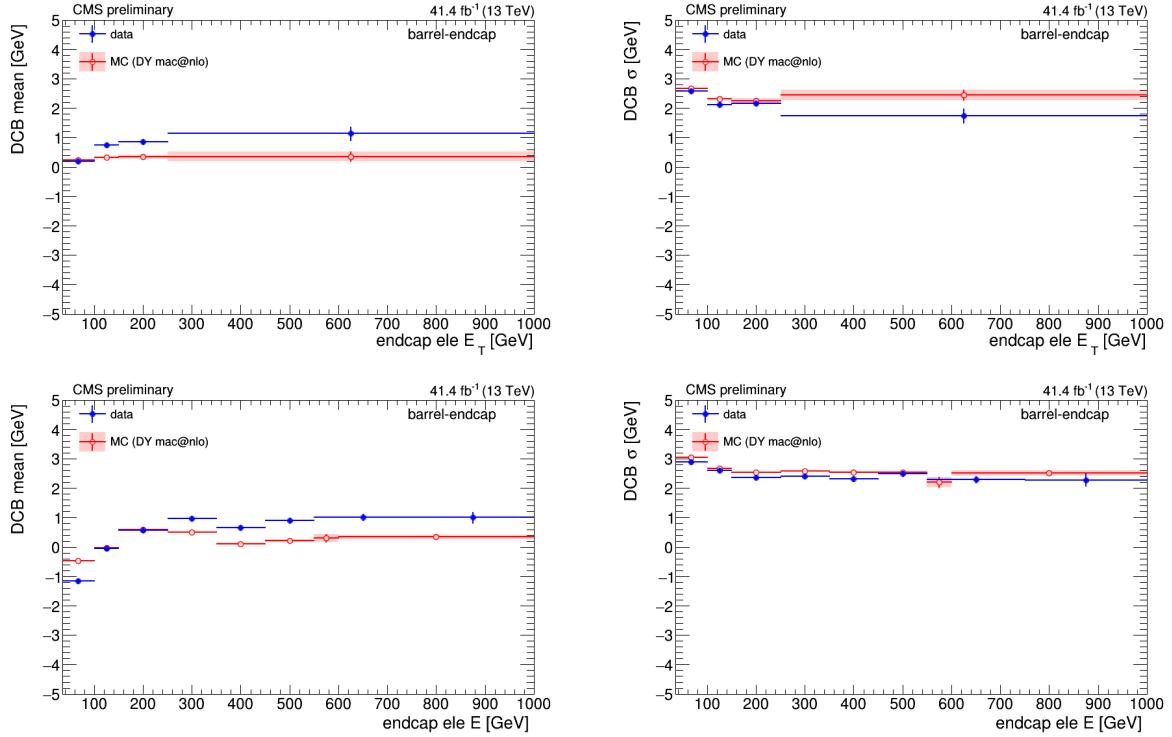


Figure 3.13: The mean (left) and sigma (right) value of dCB as the function of E_T (top) and energy (bottom) of endcap electron for barrel-endcap channel.

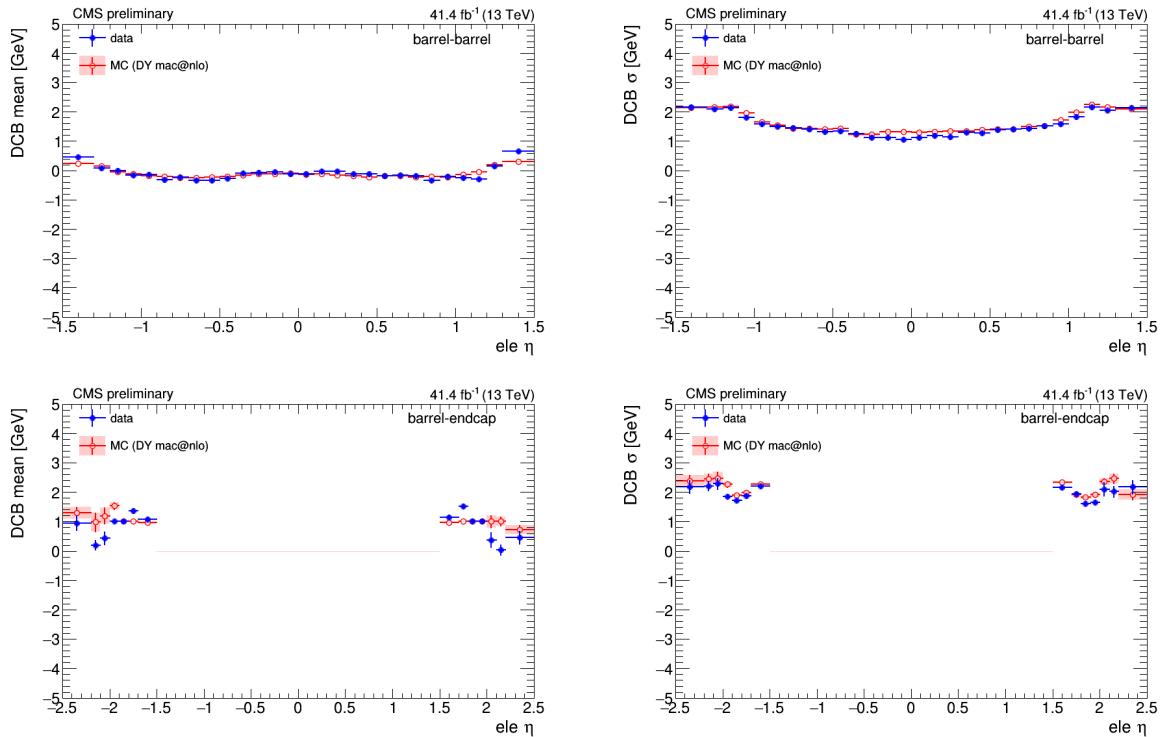


Figure 3.14: The mean (left) and sigma (right) value of dCB as the function of η of second electron (the first electron is asked to have $|\eta| < 0.5$) for barrel-barrel (top) and barrel-endcap (bottom) channel.

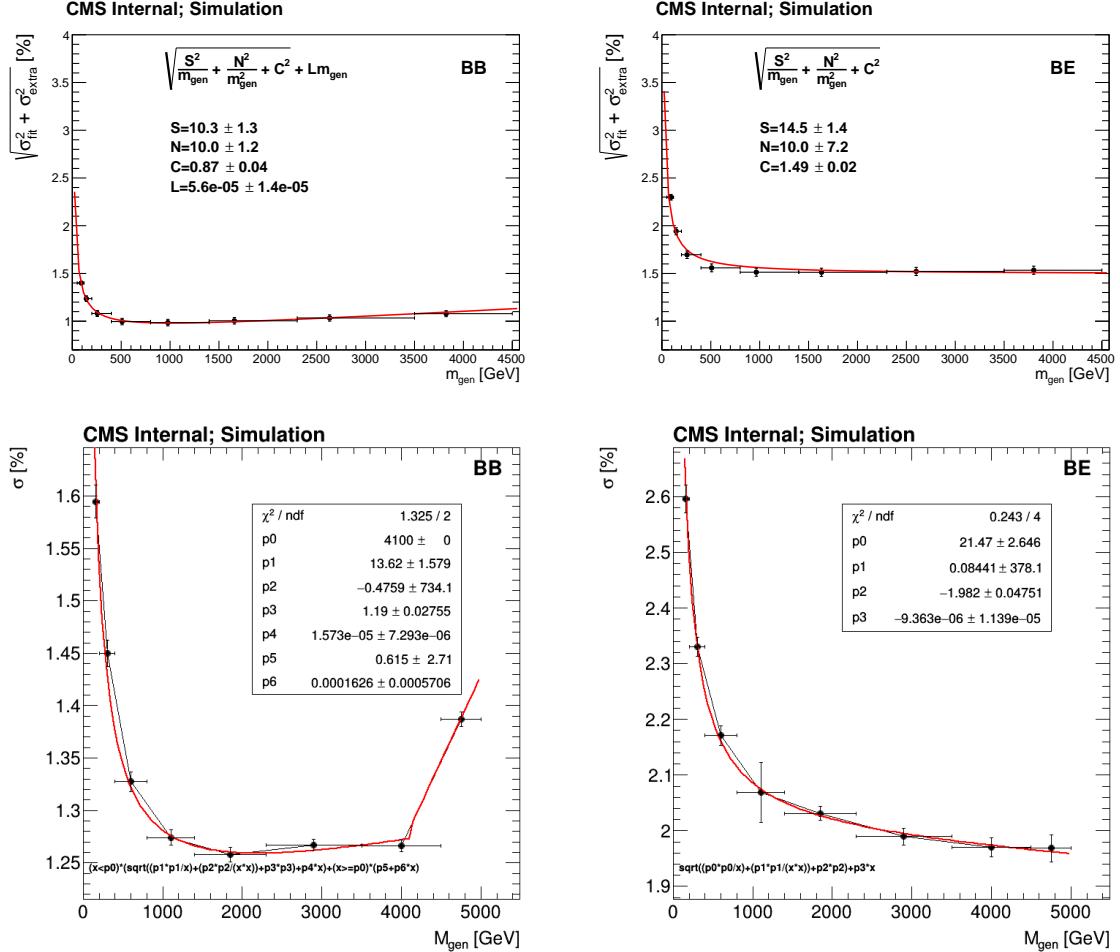


Figure 3.15: Mass resolution (σ_{fit}) as a function of the generated invariant mass for the BB region (left) and BE (right) channel for 2016 (top) and 2017 (bottom).

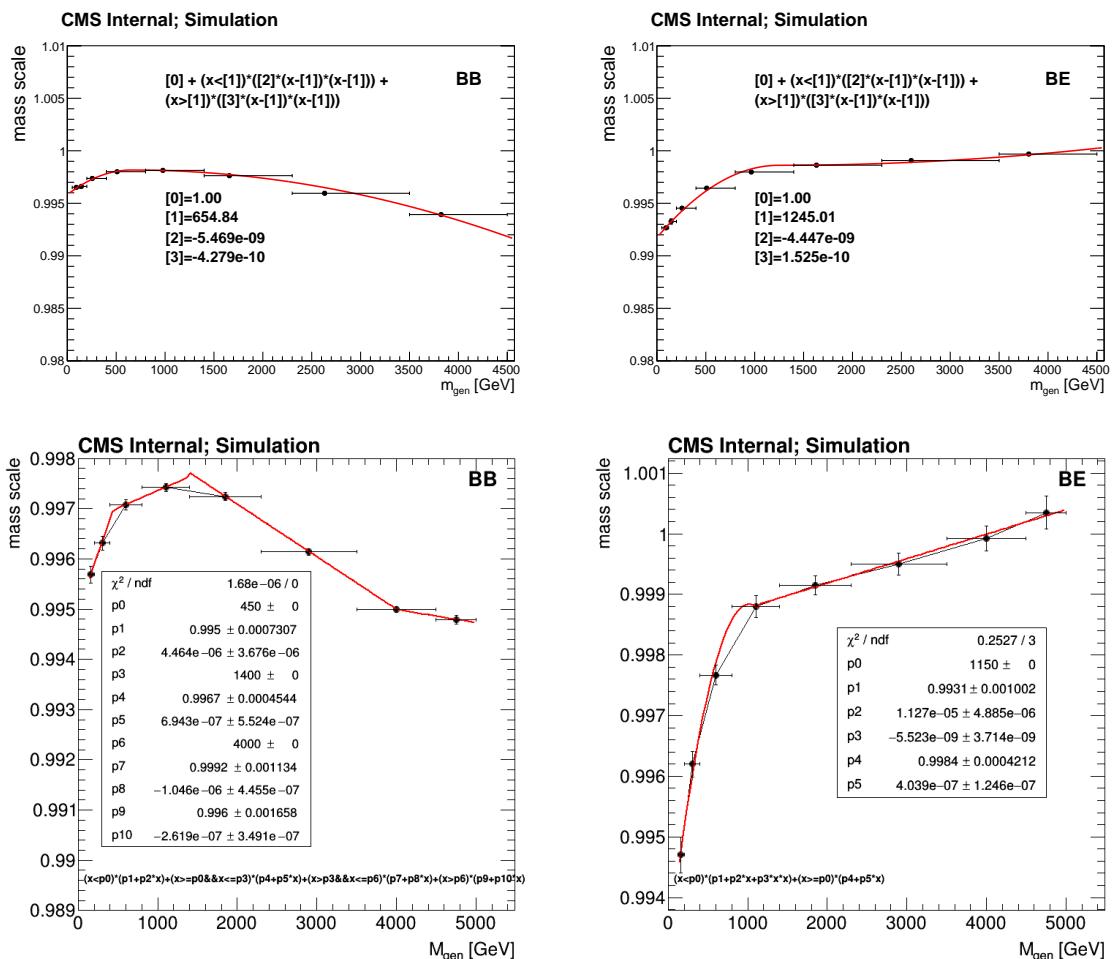


Figure 3.16: Mass scale as a function of the generated invariant mass for the BB region (left) and BE region (right) channel for 2016 (top) and 2017 (bottom).

3.5 HEEP ID efficiency and scale factor

The MC samples used for High Energy Electron Pair (HEEP) selection efficiency and scale factor measurement are listed in Table 3.5

sample	xsection(pb)	comments
DYToEE_NNPDF30_13TeV-powheg-pythia8	1921.8	for 2016 $Z \rightarrow ee$
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	5765.4	for 2016 $Z \rightarrow \tau\tau$
DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8	5765.4	for 2017 $Z \rightarrow ee, Z \rightarrow \tau\tau$
WJetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	61526.7	for 2016, 2017
TT_TuneCUETP8M2T4_13TeV-powheg	831.76	for 2016
TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	87.31	for 2017
TTToSemiLeptonic_TuneCP5_PSweights_13TeV-powheg-pythia8	366.6	for 2017
ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8	35.6	for 2016, 2017
ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8	35.6	for 2016, 2017
GJets_HT-40To100_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	20790	for 2016, 2017
GJets_HT-100To200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	9238	for 2016, 2017
GJets_HT-200To400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	2305	for 2016, 2017
GJets_HT-400To600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	274.4	for 2016, 2017
GJets_HT-600ToInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	93.46	for 2016, 2017
WW_TuneCUETP8M1_13TeV-pythia8	118.7	for 2016, 2017
WZ_TuneCUETP8M1_13TeV-pythia8	47.13	for 2016, 2017
ZZ_TuneCUETP8M1_13TeV-pythia8	16.523	for 2016, 2017

Table 3.5: MC samples used in HEEP efficiency and scale factor study

3.5.1 Tag and probe method

To calculate the efficiency of a certain selection we use a method called 'tag and probe' using $Z \rightarrow ee$ events. This method starts by searching for a good electron, 'tag', which satisfies certain types of criteria. Then the efficiency of interest is measured testing the cuts on a second electron candidate called 'probe' electron. For this study, we measure HEEP ID efficiency and scale factor.

Tag is required to fulfill the following conditions and is paired with every other gsf electron candidate in the event that passes the E_T and η acceptance cuts of the HEEP ID (probe).

- tag must pass the HEEP ID
- tag must be a barrel electron
- tag must be matched to the `HLT_Ele27_eta2p1_WPTight` (`HLT_Ele35_WPTight`) trigger in data in 2016 (2017)

The invariant mass of the tag and probe must satisfy $70 < m(ee) < 110$ GeV/c². If there are multiple tag and probe candidates in an event then all pairs are selected. When there are two tags in a pair, both are considered to also be probes. What we get is a very clear peak around the mass of the Z boson. Therefore, we are confident that the electrons we have chosen are in fact products of a Z boson decay and are real electrons although there are low contamination from other standard model backgrounds.

The efficiency is defined as

$$\epsilon_{\text{HEEP}} = \frac{N_{\text{passing probes}}}{N_{\text{all probes}}} \quad (3.4)$$

where $N_{\text{all probes}}$ is the total number of all selected probes and $N_{\text{passing probes}}$ is the total number of selected probes which pass HEEP ID selection criteria. The efficiency can be measured in data and MC as a function of different variables like electron E_T , η , ϕ , etc.

For finding HEEP ID efficiency in data, events are selected from SingleElectron dataset using `HLT_Ele27_eta2p1_WPTight` (`HLT_Ele35_WPTight`) trigger in 2016 (2017).

For finding HEEP ID efficiency in MC, various MC samples are used which can be found in tables 3.5. MC events are weighted using the trigger turn on curve of `HLT_Ele27_eta2p1_WPTight` (`HLT_Ele35_WPTight`) path (see figure 3.8 and 3.9 of section 3.2.4) to emulate the trigger efficiency instead of matching tag with the `HLT_Ele27_eta2p1_WPTight` (`HLT_Ele35_WPTight`) trigger object. MC events are weighted to correct for difference between data and MC pileup distribution according to the standard procedure.

The contribution of QCD background, where tag and probe are misidentified jets, is extracted from data using the 'same-sign' method. In same-sign method, we use the fact that the probability of assigning positive or negative charge to the misidentified jet should be equal (LHC is proton proton collider and positive charged jets should be produced a bit more but we are not sensitive to that amount in this study). Therefore, opposite sign and same sign tag and probe pairs are similar for QCD in total number and distribution shape for many variables. On the other hand, all other standard model processes have opposite sign electron pairs and do not contribute to same sign control region.

In Figure 3.17, the invariant mass distributions of selected tag and probe pairs are shown for all tag and probe pairs (top), tag and probe pairs in which probe passes HEEP ID selection cuts 'pass-pass' (middle) and tag and probe pairs in which probe fails passing HEEP ID selection cuts 'pass-fail' (bottom). Due to the different ID efficiencies in barrel and endcaps, the invariant mass distribution of selected tag and probe pairs are separated for probe electrons in barrel and endcaps. From Figure 3.18 to 3.21 (3.22 to 3.26) are invariant mass distributions of selected tag and probe pairs, E_T , η , ϕ of probes and number of primary vertex for the selected tag and probe pairs event in 2016 (2017).

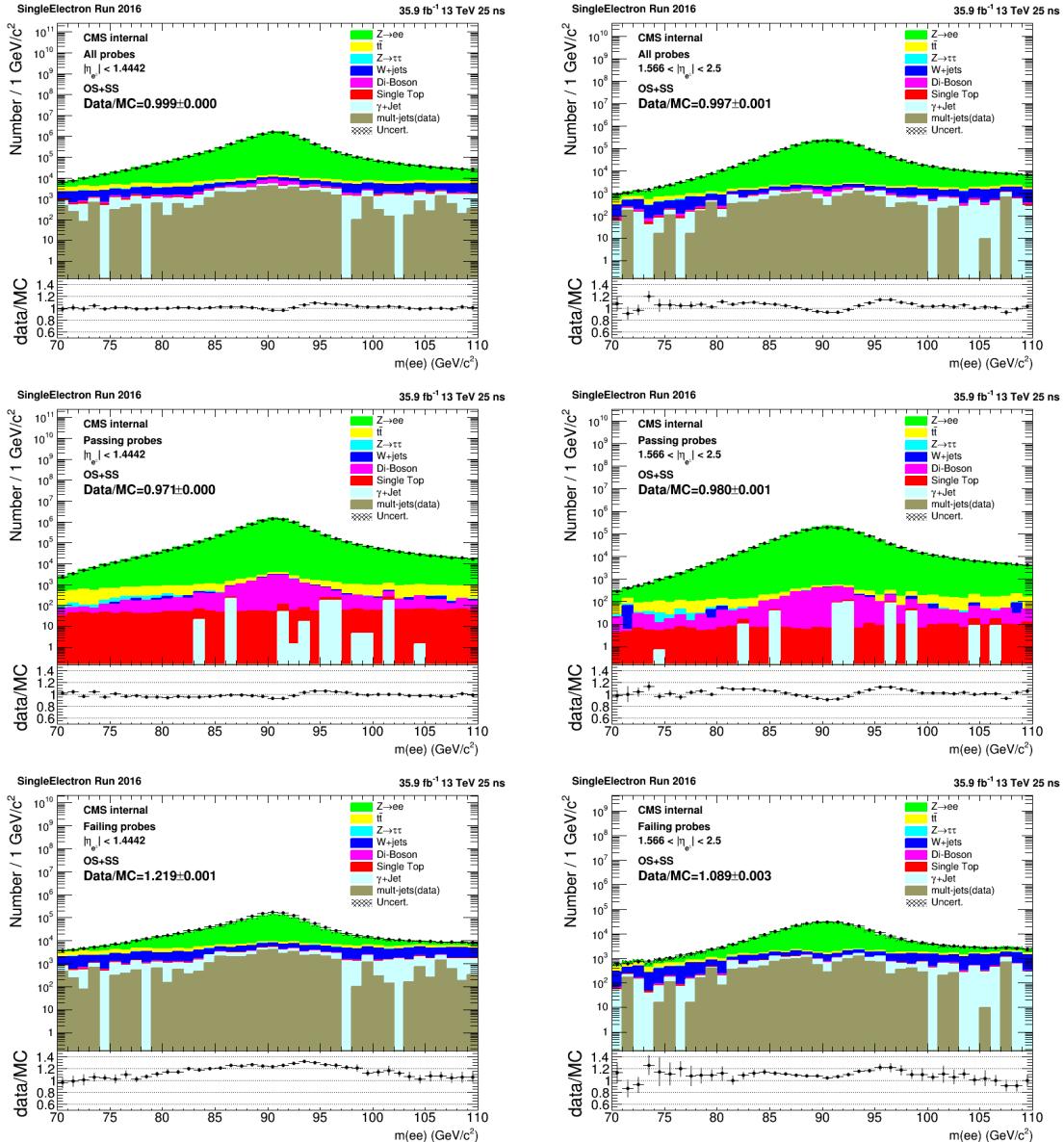


Figure 3.17: Invariant mass of tag and probe for probe in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle) and only failed probes are included (bottom) for 2016.

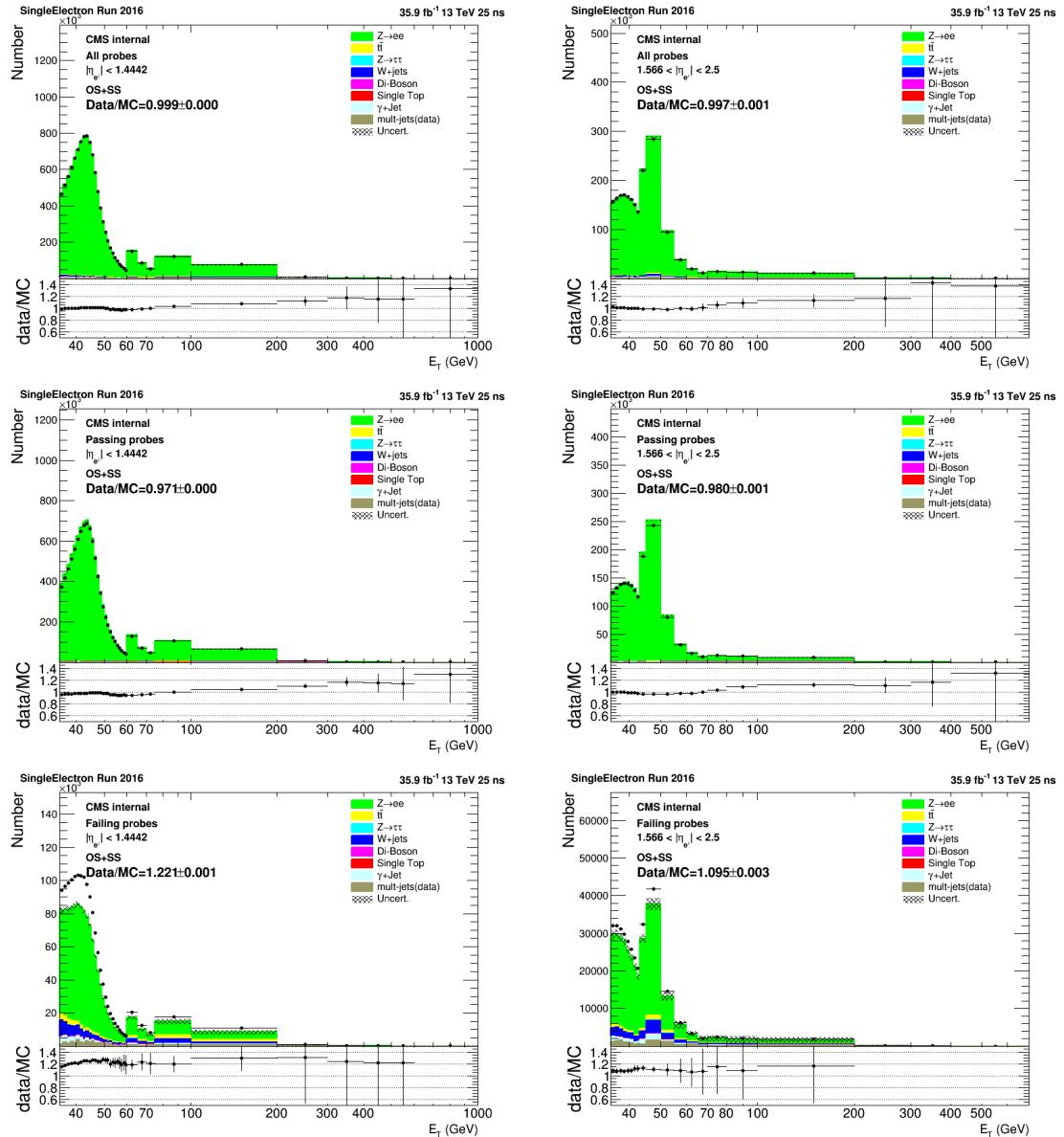


Figure 3.18: E_T of probe in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle) and only failed probes are included (bottom) for 2016.

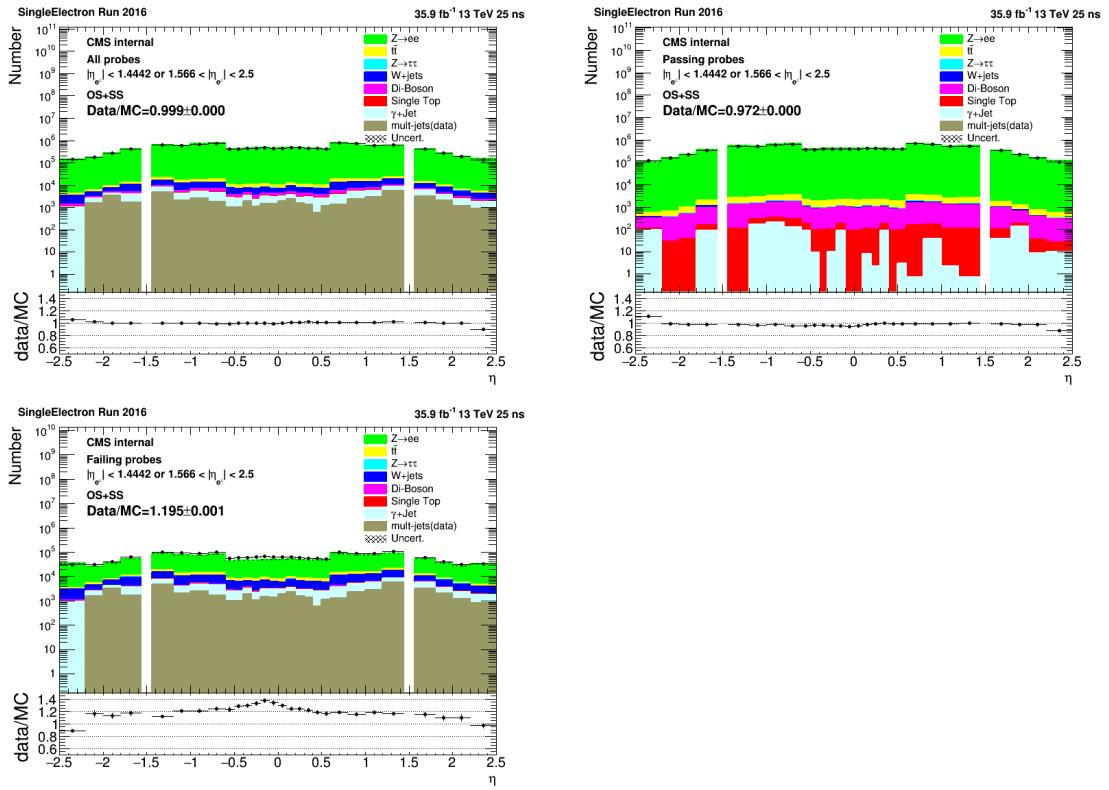


Figure 3.19: η of probe in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle) and only failed probes are included (bottom) for 2016.

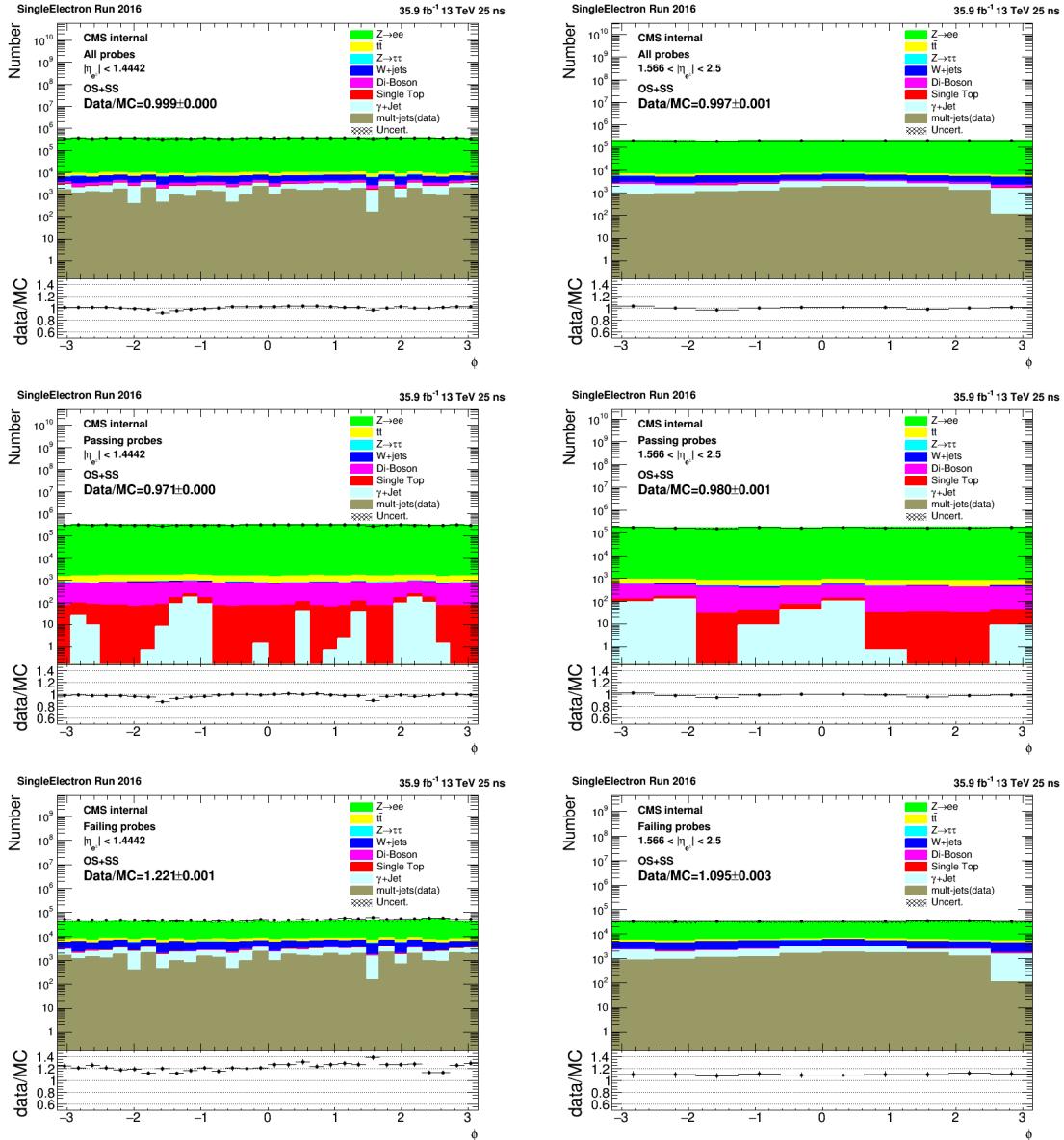


Figure 3.20: ϕ of probe in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle) and only failed probes are included (bottom) for 2016.

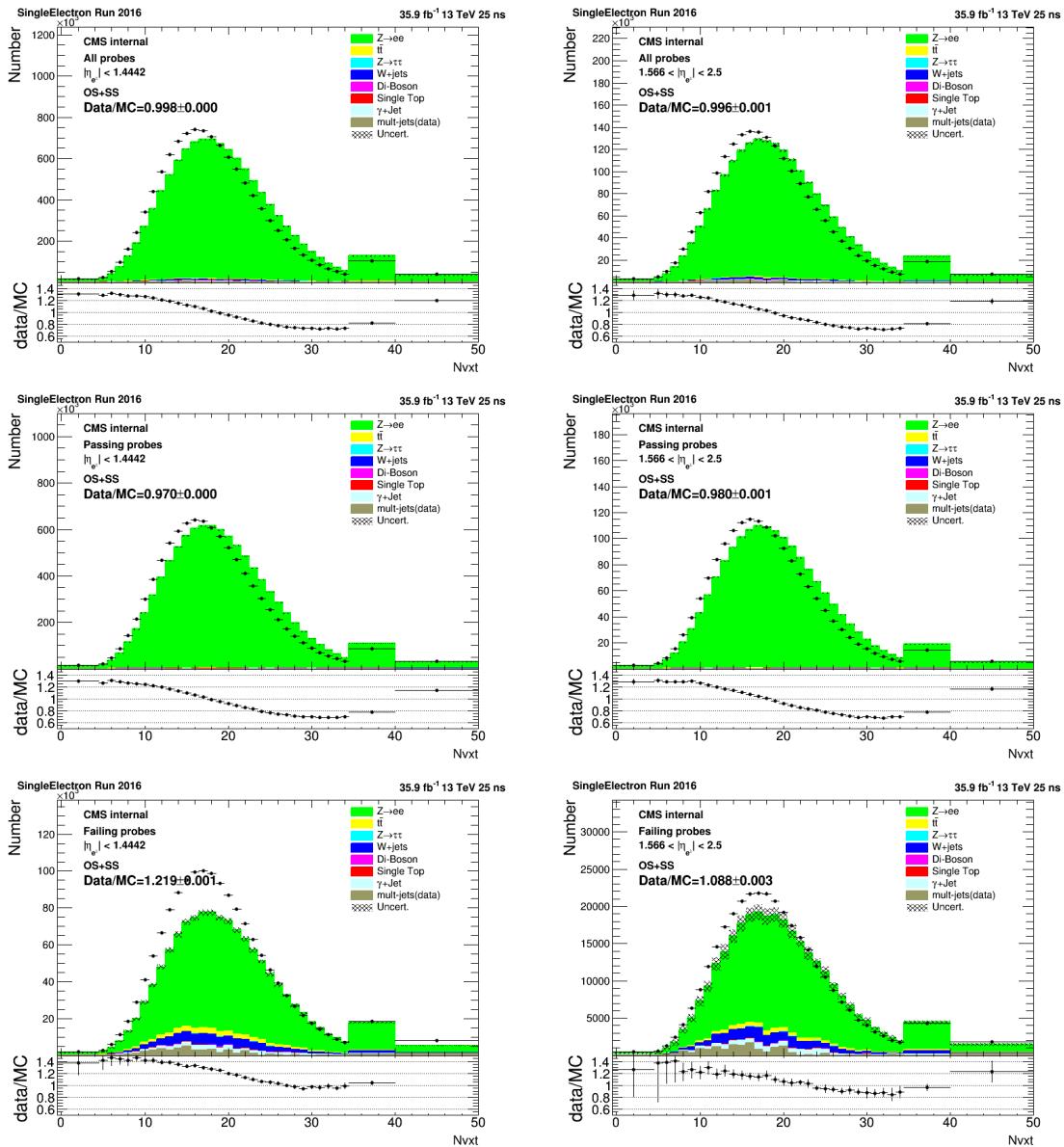


Figure 3.21: Number of primary vertex of tag and probe pair event in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle) and only failed probes are included (bottom) for 2016.

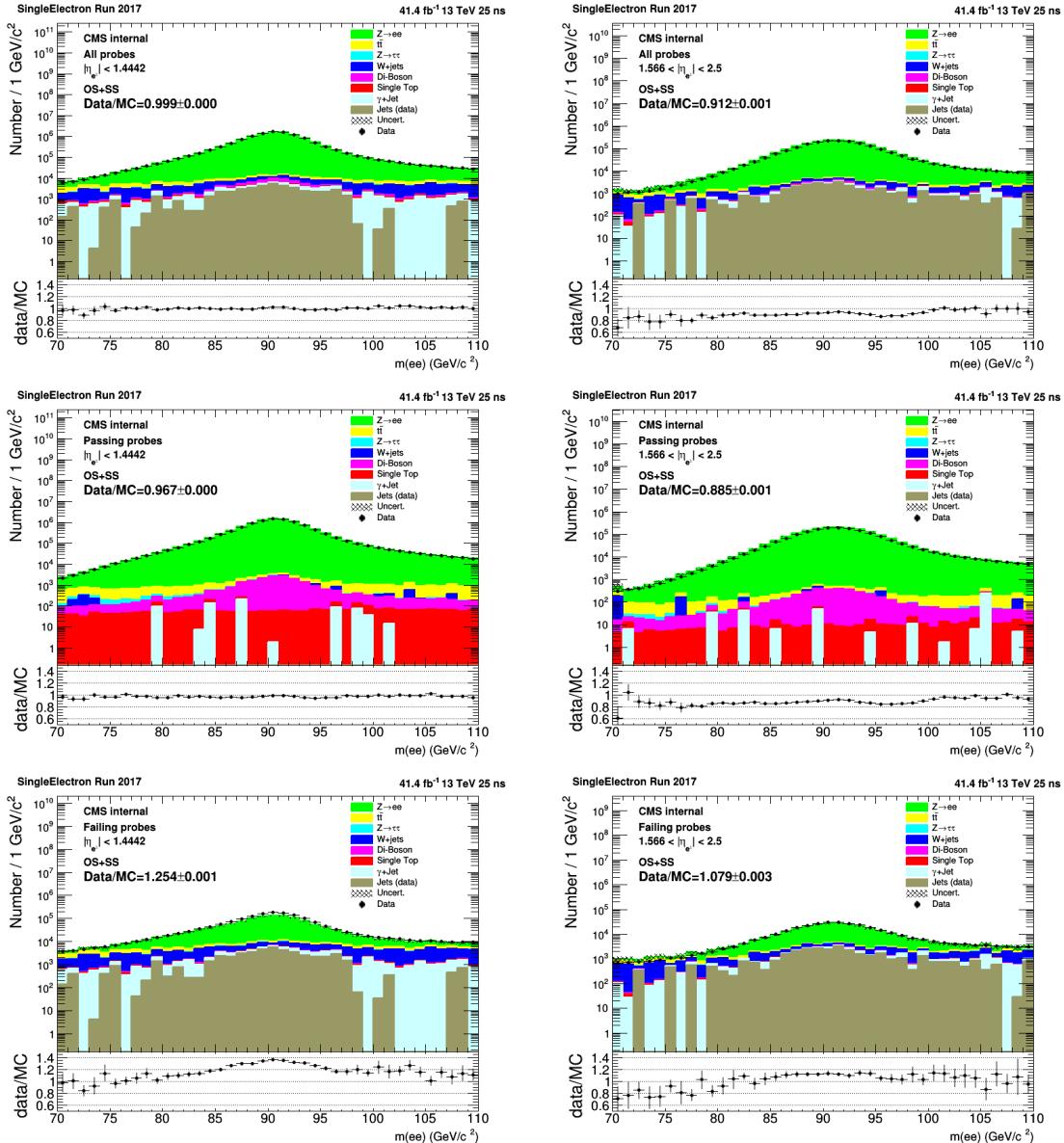


Figure 3.22: Invariant mass of tag and probe for probe in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle) and only failed probes are included (bottom) for 2017.

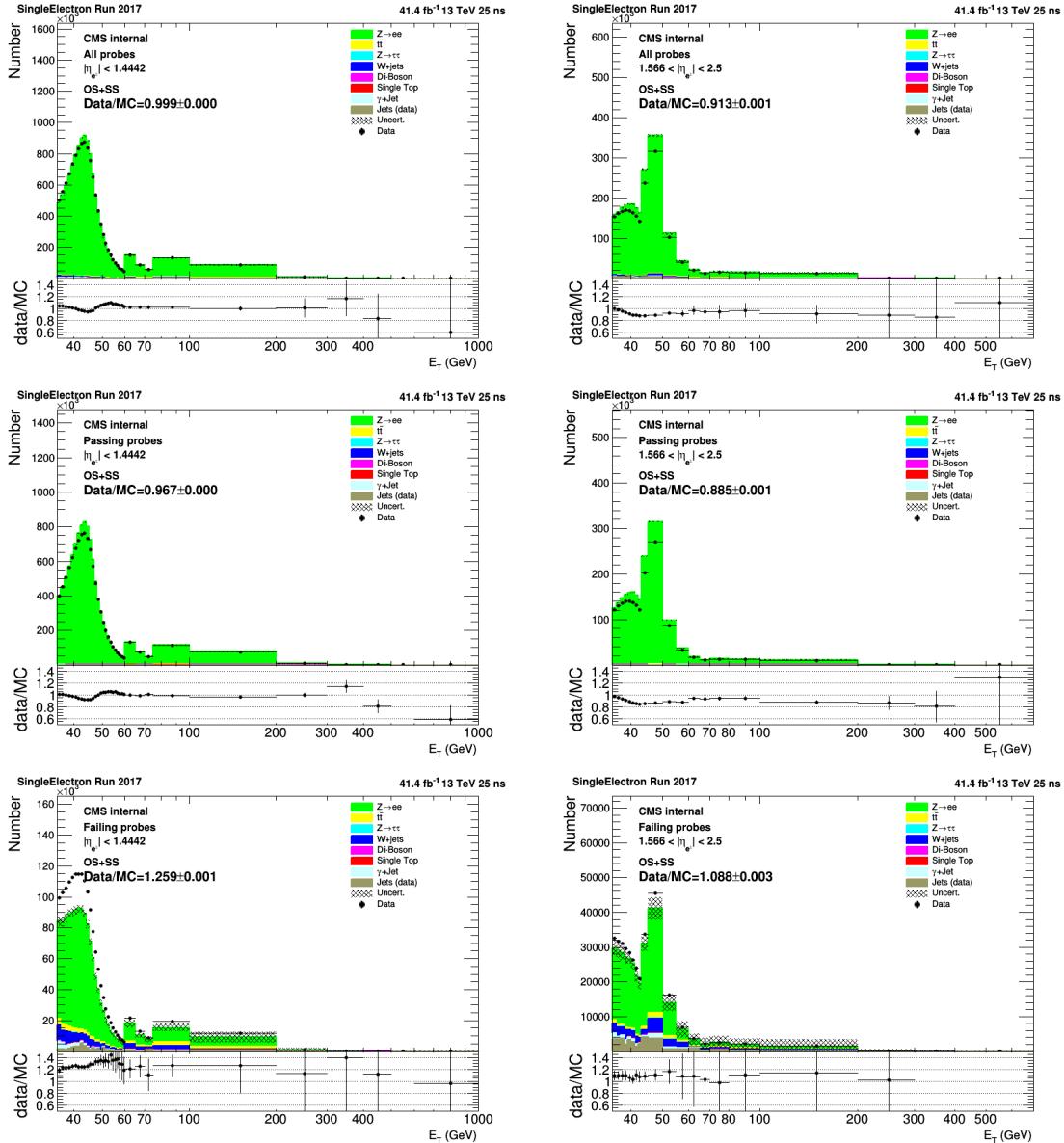


Figure 3.23: E_T of probe in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle) and only failed probes are included (bottom) for 2017.

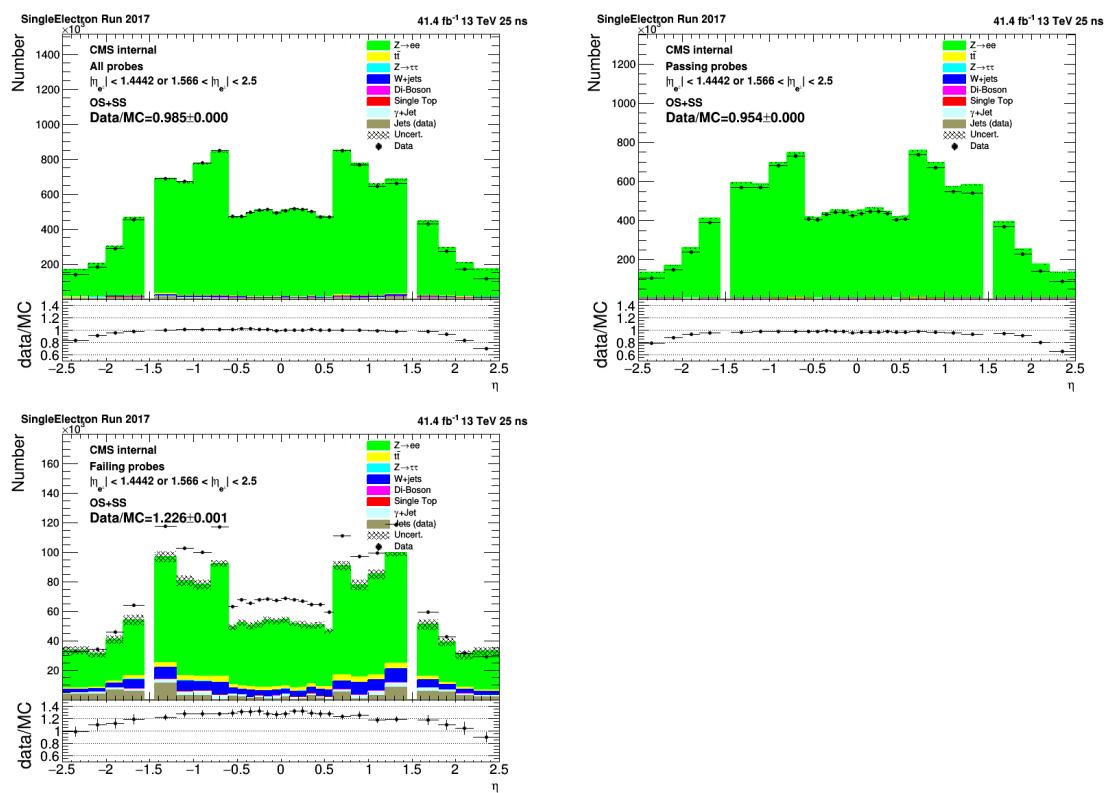


Figure 3.24: η of probe in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle) and only failed probes are included (bottom) for 2017.

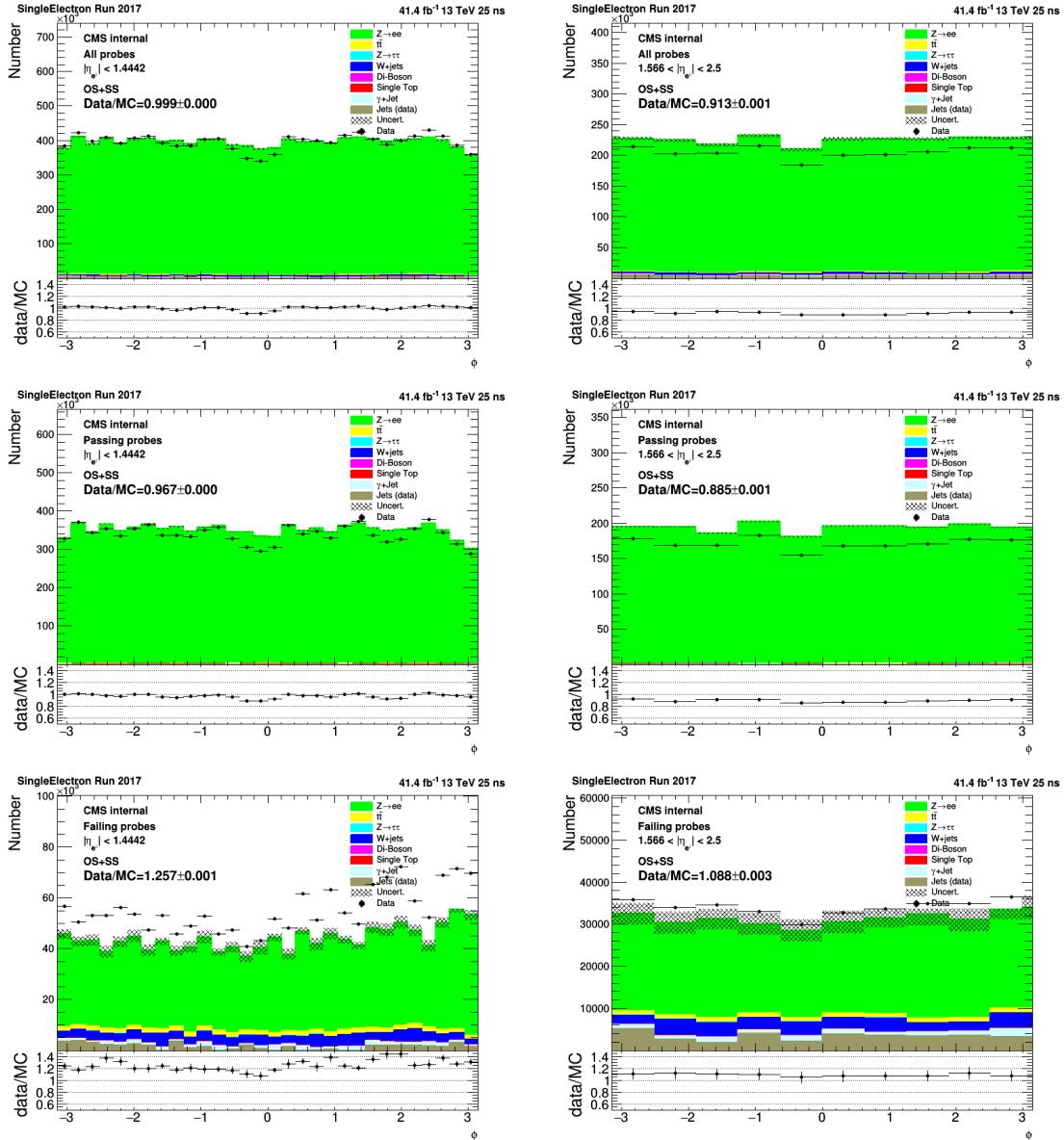


Figure 3.25: ϕ of probe in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle) and only failed probes are included (bottom) for 2017.

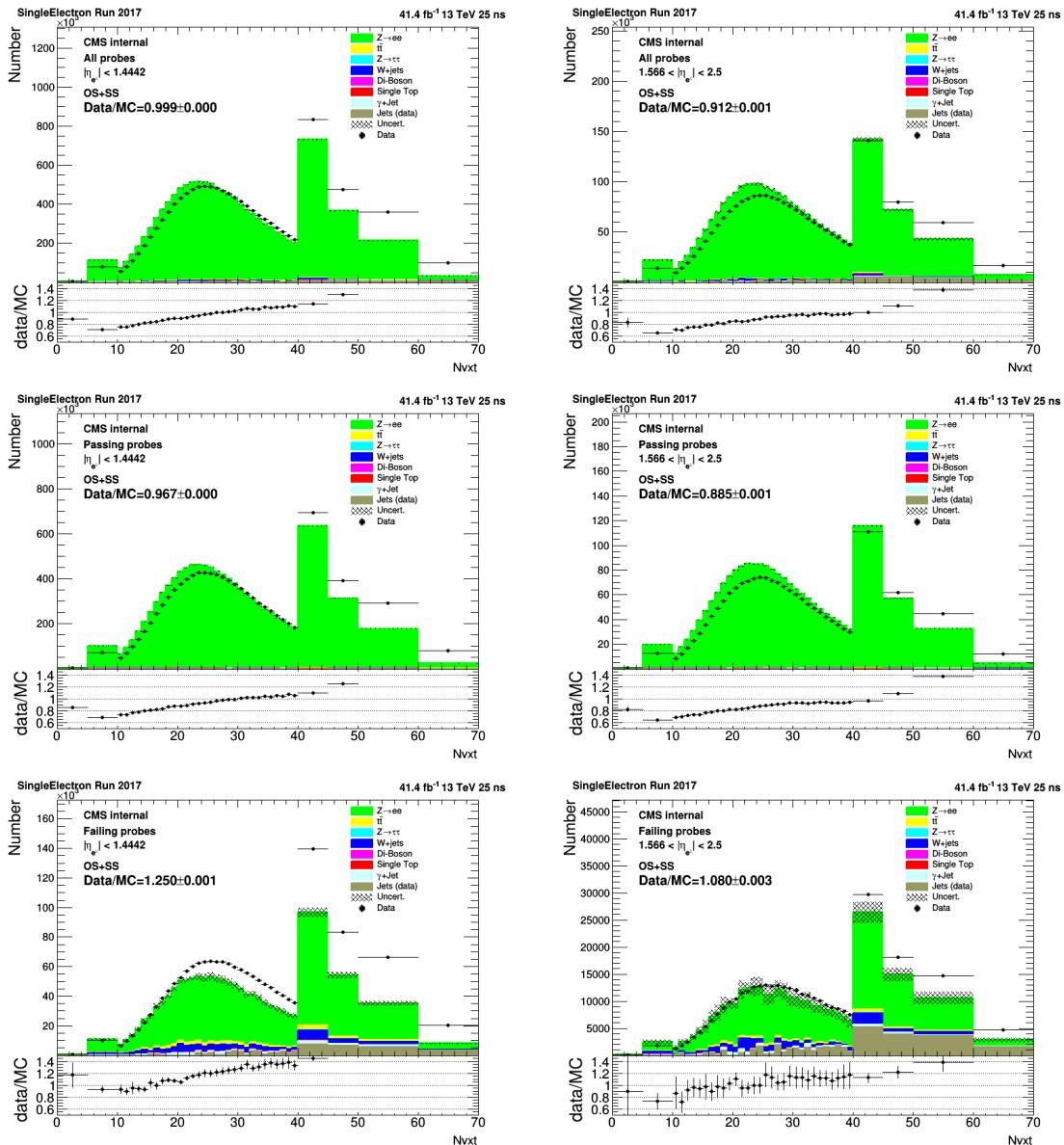


Figure 3.26: Number of primary vertex of tag and probe pair event in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle) and only failed probes are included (bottom) for 2017.

3.5.2 HEEP ID efficiencies and scale factors

In 2016 the HEEP ID efficiencies for data and MC are shown as functions of E_T , η , and ϕ of the probe, and of N_{vtx} in figures 3.27-3.30 (left), here data includes both DY and non-DY events. In order to check if the contribution of non-DY backgrounds are estimated correctly, the DY efficiency is compared to data while non-DY contributions are subtracted from data in figures 3.27-3.30 (right). The scale factors obtained with these two approaches are in a good agreement in all variables. In addition, data to MC scale factor per bin and the mean value are shown in the bottom pad of the same plot. Same plots in 2017 are shown in figures 3.31-3.34.

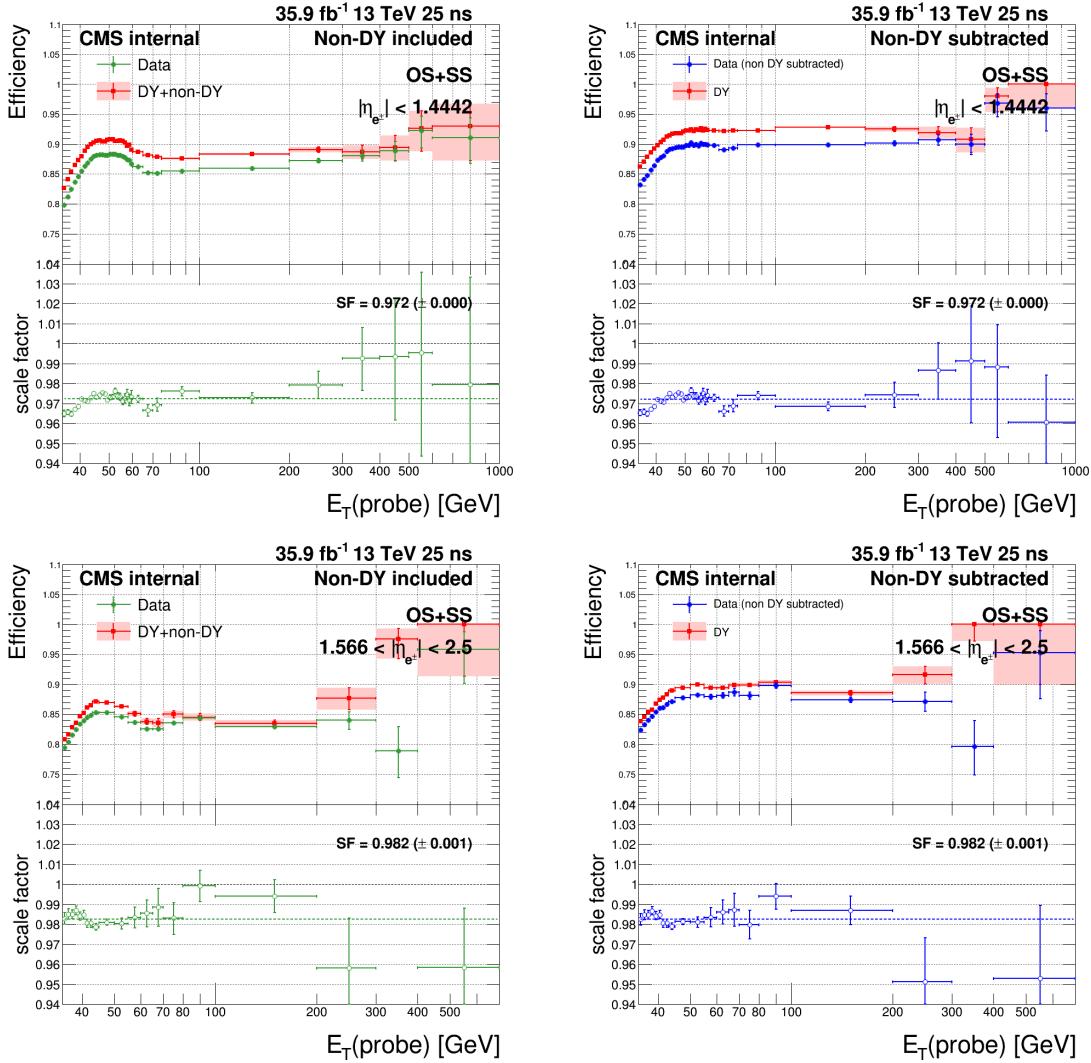


Figure 3.27: Efficiencies and scale factors in MC and data in the barrel (top) and endcap (bottom) where the non-DY processes are included (left) and subtracted (right) as functions of probe E_T in 2016.

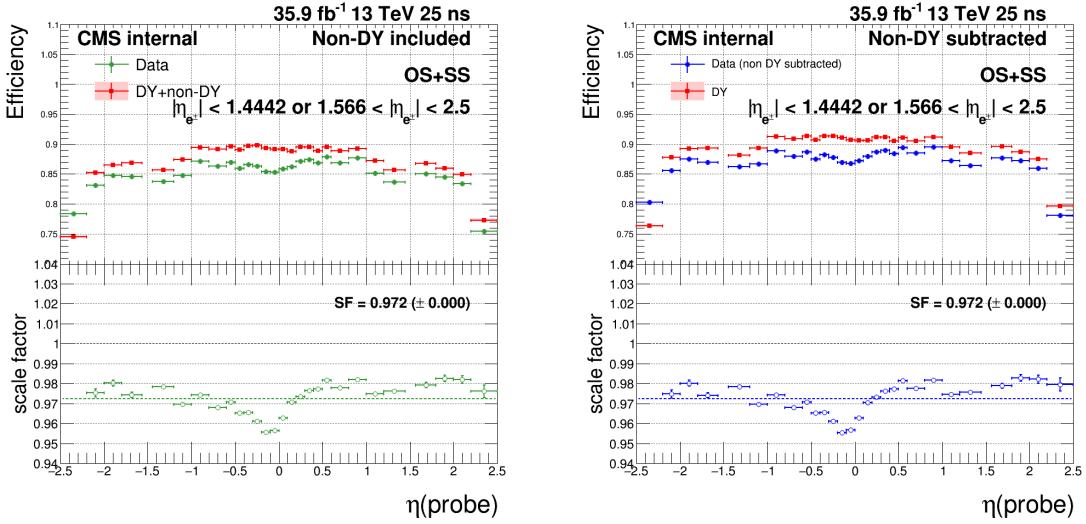


Figure 3.28: Efficiencies and scale factors in MC and data where the non-DY processes are included (left) and subtracted (right) as functions of probe η in 2016.

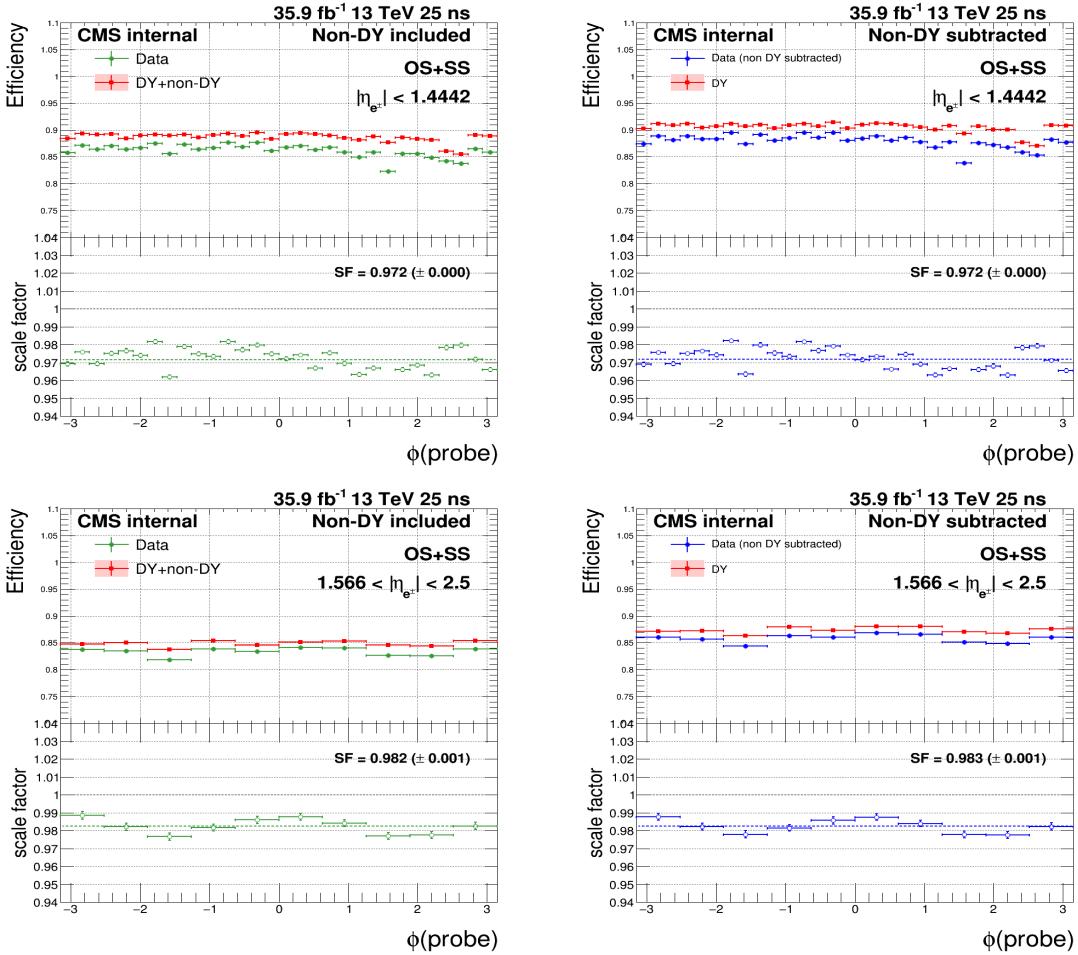


Figure 3.29: Efficiencies and scale factors in MC and data in the barrel (top) and endcap (bottom) where the non-DY processes are included (left) and subtracted (right) as functions of probe ϕ in 2016.

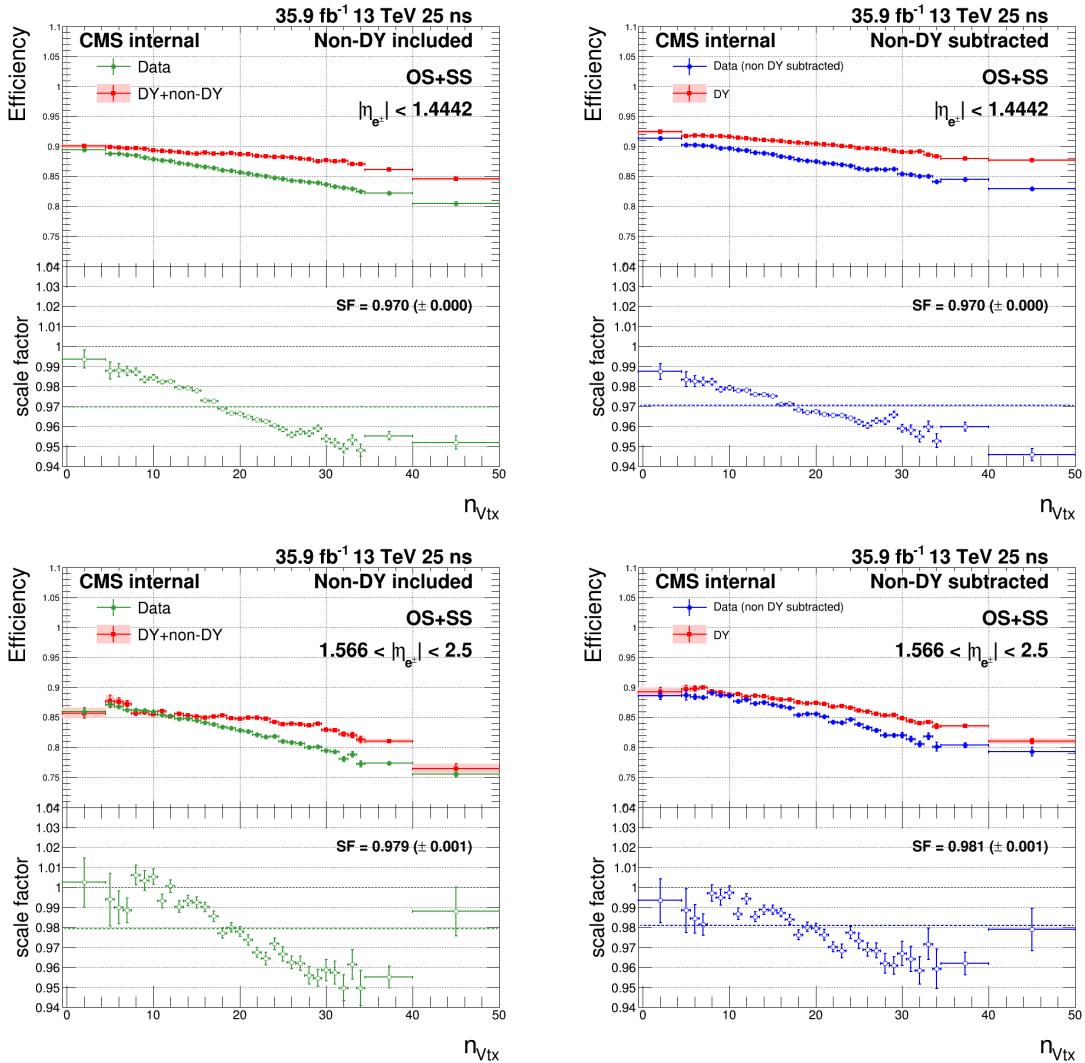


Figure 3.30: Efficiencies and scale factors in MC and data in the barrel (top) and endcap (bottom) where the non-DY processes are included (left) and subtracted (right) as functions of the number of primary vertices, n_{Vtx} in 2016.

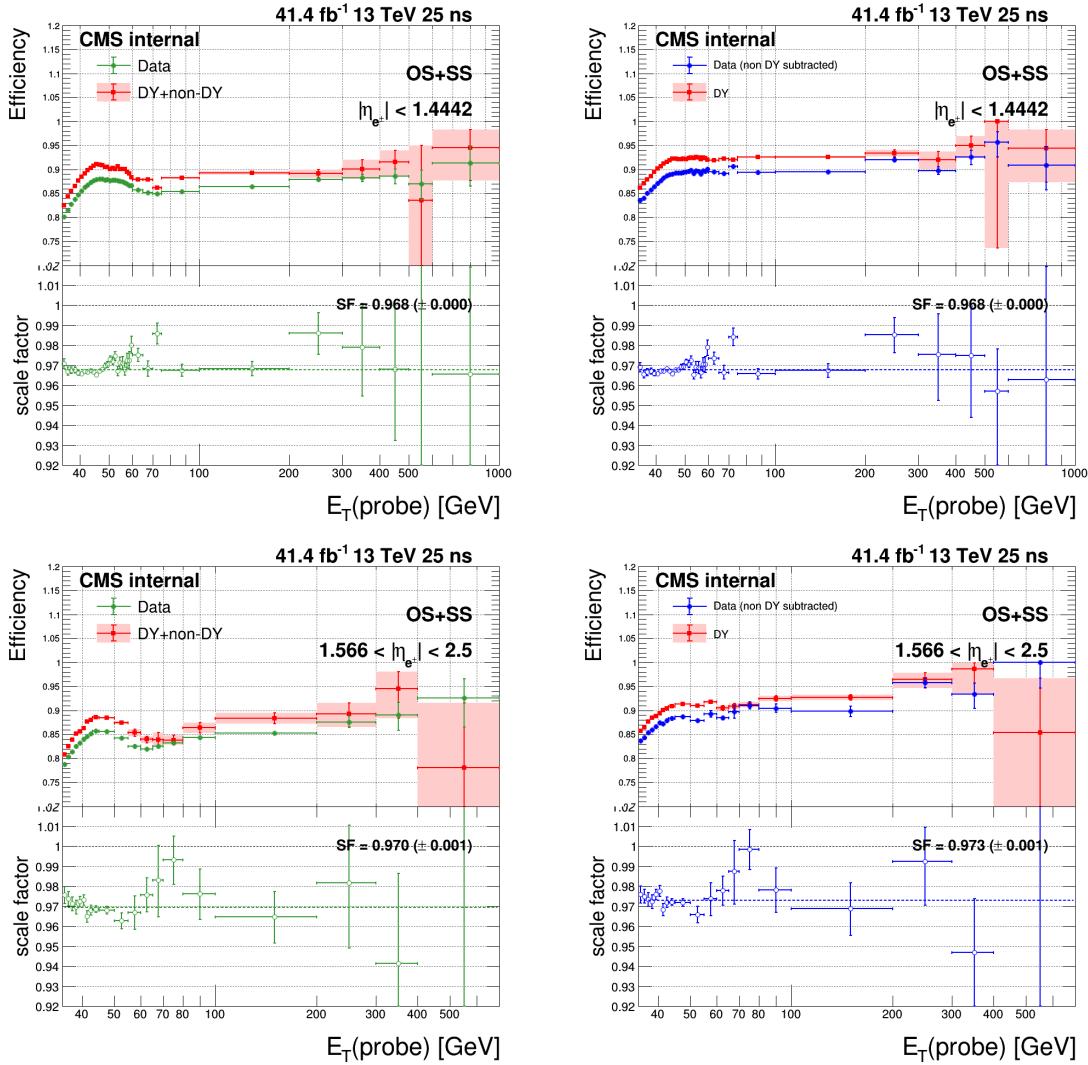


Figure 3.31: Efficiencies and scale factors in MC and data in the barrel (top) and endcap (bottom) where the non-DY processes are included (left) and subtracted (right) as functions of probe E_T in 2017.

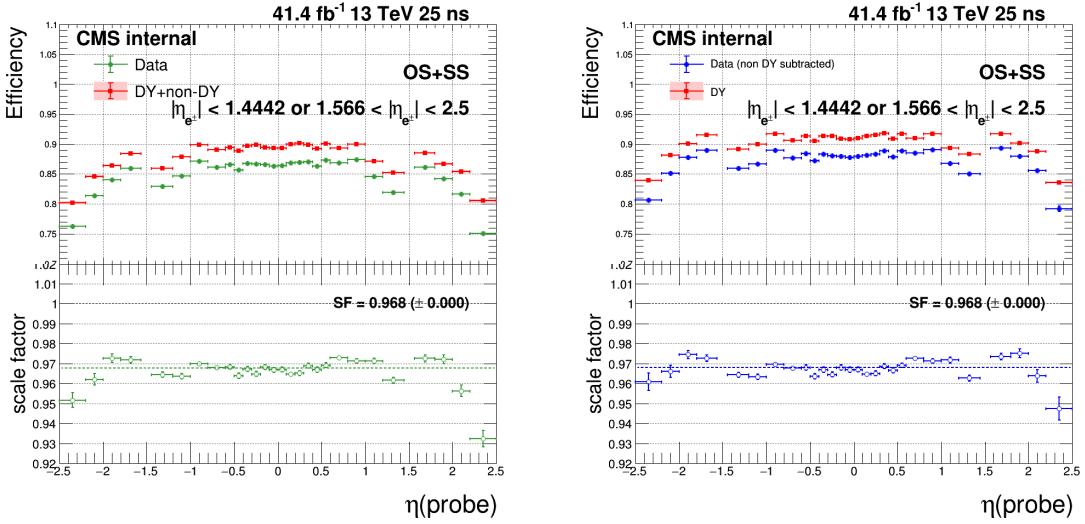


Figure 3.32: Efficiencies and scale factors in MC and data where the non-DY processes are included (left) and subtracted (right) as functions of probe η in 2017.

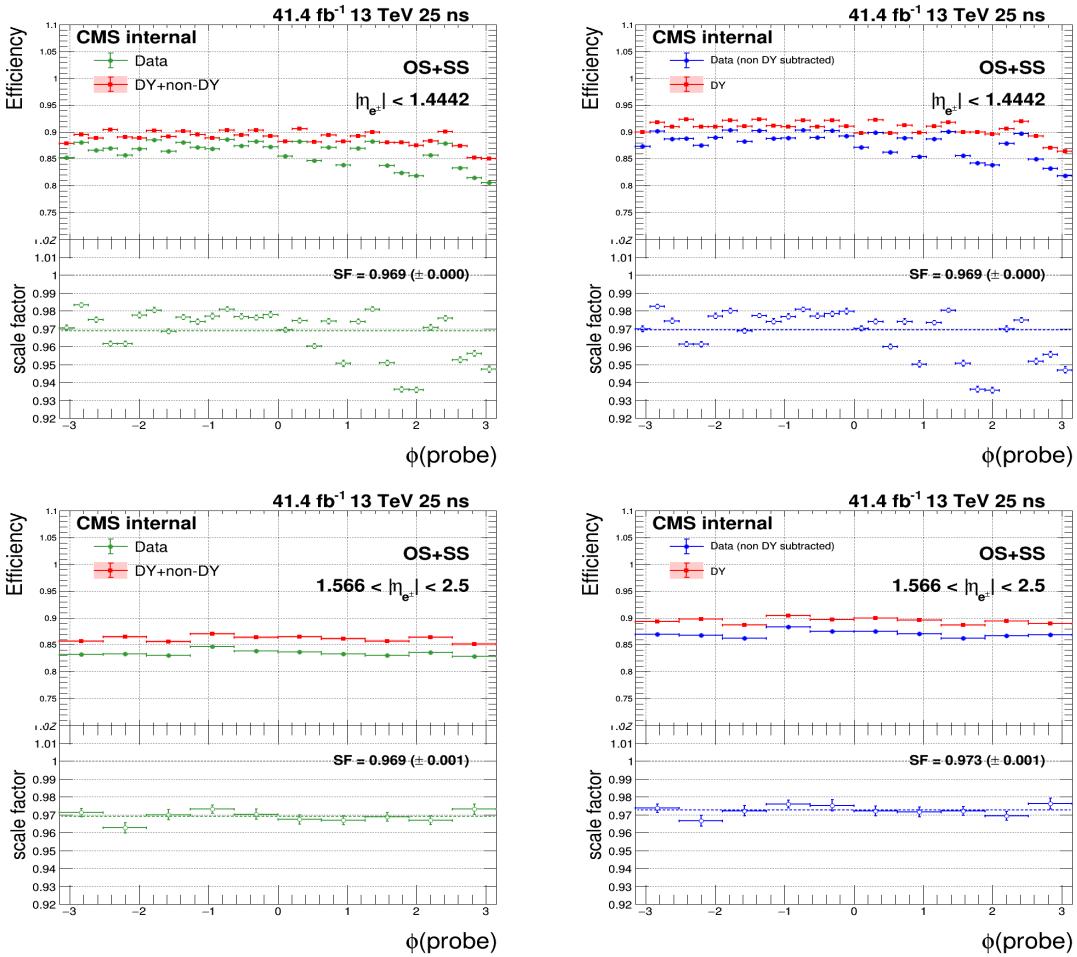


Figure 3.33: Efficiencies and scale factors in MC and data in the barrel (top) and endcap (bottom) where the non-DY processes are included (left) and subtracted (right) as functions of probe ϕ in 2017.

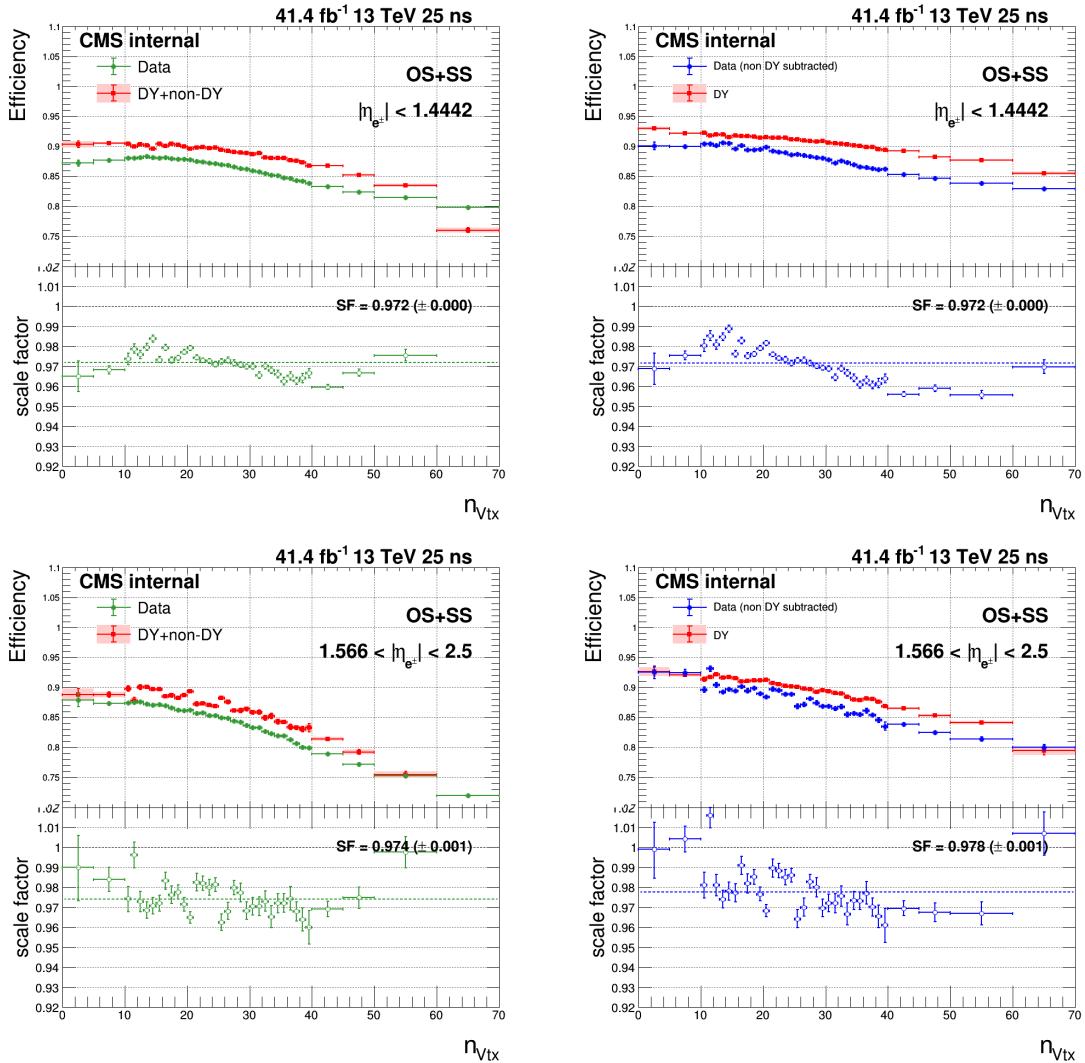


Figure 3.34: Efficiencies and scale factors in MC and data in the barrel (top) and endcap (bottom) where the non-DY processes are included (left) and subtracted (right) as functions of the number of primary vertices, n_{Vtx} in 2017.

The main sources of systematic uncertainties on the scale factor originate from non-DY processes. The nominal value of the cross sections are varied by 10% and 50% for $t\bar{t}$ and $W+jets$ processes, respectively. For QCD estimation, we also consider 50% uncertainty. The summary of the efficiencies and scale factors are shown in Table 3.6.

Year		Barrel	Endcap
2016	Data	$86.13\% \pm 0.01\%(stat.)$	$83.38\% \pm 0.03\%(stat.)$
	DY + non-DY	$88.65\% \pm 0.03\%(stat.)$	$84.85\% \pm 0.09\%(stat.)$
	Scale factor	$0.972 \pm 0.000(stat.) \pm 0.006(syst.)$	$0.983 \pm 0.001(stat.) \pm 0.007(syst.)$
	Data - non-DY	$87.92\% \pm 0.03\%(stat.)$	$85.83\% \pm 0.09\%(stat.)$
	DY	$90.50\% \pm 0.01\%(stat.)$	$87.35\% \pm 0.03\%(stat.)$
	Scale factor	$0.971 \pm 0.000(stat.) \pm 0.006(syst.)$	$0.983 \pm 0.001(stat.) \pm 0.007(syst.)$
2017	Data	$86.01\% \pm 0.01\%(stat.)$	$83.46\% \pm 0.03\%(stat.)$
	DY + non-DY	$88.89\% \pm 0.05\%(stat.)$	$86.12\% \pm 0.14\%(stat.)$
	Scale factor	$0.968 \pm 0.001(stat.) \pm 0.005(syst.)$	$0.969 \pm 0.002(stat.) \pm 0.01(syst.)$
	Data - non-DY	$87.81\% \pm 0.05\%(stat.)$	$87.02\% \pm 0.16\%(stat.)$
	DY	$90.77\% \pm 0.02\%(stat.)$	$89.48\% \pm 0.04\%(stat.)$
	Scale factor	$0.967 \pm 0.001(stat.) \pm 0.005(syst.)$	$0.973 \pm 0.002(stat.) \pm 0.01(syst.)$

Table 3.6: Efficiencies and scale factors in MC and data in the barrel and endcap for non-DY processes included and subtracted.

It is worth mentioning that many complementary studies are done to understand HEEP scale factor better. Important points are summarized in the following and related plots can be found in Appendix.

- In 2016
- the DYToEE Monte Carlo sample used had a special global tag which was discovered late in the process. This tag has a different ECAL noise profile and different transparency corrections which could impact our scale factor. The efficiency vs gen E_T for this inclusive sample was compared to the mass binned samples in figure48 and they were found to be similar, with a deviation of only 0.3% at low E_T in the barrel and up to 1% in the endcap. It should be noted that the deviation in the endcap will act to flatten the scale factor. Thus it is concluded that impact of the special global tag used in the DYToEE sample does not impact the scale factor measurement significantly.
- The HEEP scale factor in the last two bins of E_T in the endcap seems unusual and DY efficiency is 100%. The reason could be either a statistical fluctuations or something in Moriond17 MC DY sample. In Appendix .2.3, we cross checked the HEEP efficiency and scale factor using DYJetsToLL amcatnlo sample which shows the DY efficiency in endcap in last two bins are normal. From figure 48, it can be seen that the efficiency in the MC for 500 GeV electrons is normal if you don't apply the Z constraint and so it is thought to be a statistical fluctuation.
- HEEP scale factor drops 2% around $|\eta| = 0$. This is mostly related to 'HIP' (Heavy Interacting Particle) problem which means the heavy interacting particle (most are hadron) produced a huge current in silicon strips and make them off for few bunch crossing and we lose hits from electrons. This problem is present in runs B-F and is removed in runs G-H. This issue is discussed in .2.2.
- there is a turn on effect of scale factor at low E_T .This comes mainly from shower shape variable (which can be seen in Appendix .2.1 Figure 41). This effect is small ($<1\%$) so it is ignored.
- In 2017
- In run F the pixel detector has a lower efficiency in some region which can be seen

in Figure 62 in Appendix .3. This problem causes the lower HEEP ID efficiency in run F.

- The HEEP ID efficiency in data after non-DY contributions are subtracted for different runs for barrel and endcap in 2017 is shown in Figure 3.35. The efficiency is lower in run B this is because at the beginning of the data taking the detector does not work in very good condition. The efficiency decreases in run E this is because the pile up is significant increased after run E which is shown in Figure 3.1. Comparing 2016 the average HEEP ID efficiency in barrel is close in 2017, for endcap it is improved in 2017.
- A 'fit method' is preformed to cross check the HEEP ID efficiency which shown in Appendix .3.2. Comparing the nominal 'cut count method' and 'fit method' they are consistent within 1% for barrel and 2% for endcap.

More details can be found in ref. [32] ([33]) for 2016 (2017).

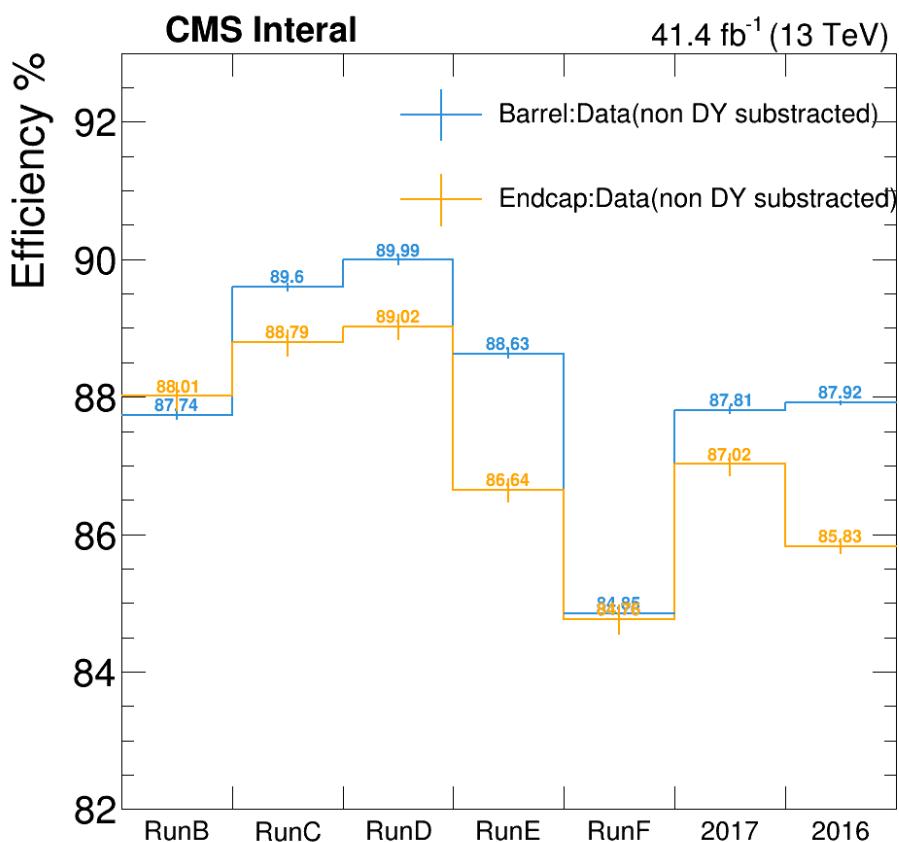


Figure 3.35: HEEP ID efficiency in data after non-DY contributions are subtracted for different runs for barrel and endcap in 2017.

3.6 Standard Model backgrounds

There are three main types of SM background to new physics search in the di-electron channel. The most significant is the irreducible SM Drell-Yan process. New physics can interfere with this process, if not mitigated, the effects can be significant.

The second most important background type comes from electrons from non-singularly produced W and Z bosons. The dominant source of these electrons are from $t\bar{t}$ events although WW events become increasingly important at high mass as the boost of the top quark means that the b-jet enters the electrons isolation cone and so the electron fails isolation requirements. Other sources include tW , WZ , ZZ and $Z/\gamma^* \rightarrow \tau^+\tau^-$ events although they are small compared to $t\bar{t}$ and WW , mainly entering at the Z peak. This background is referred to $t\bar{t}$ and $t\bar{t}$ -like background as it is dominated by $t\bar{t}$.

The third background is the jet background, where one or more jets is misidentified as an electron, mainly arising from $W+jets$ and multijets where one more jets is reconstructed as an electron.

3.6.1 SM Drell-Yan background

The SM Drell-Yan background is estimated using simulated events generated by the POWHEG event generator interfaced to PYTHIA8. Like all previous analyses, the Monte Carlo samples are normalised to the data in Z peak region of $60\text{-}120 \text{ GeV}/c^2$ unless otherwise stated. They are also corrected with the trigger turn on curve. Figure 3.36 shows the data and MC at the Z peak for the barrel-barrel and barrel-endcap. In this case, in order to see the normalisation agreement between data and MC, MC events are weighted to the luminosity, trigger turn on curve is applied and the scale factor between data and MC for gsf electron reconstruction efficiency and HEEP ID efficiency is applied. Further to this, official electron energy scale and smearing is applied for data and MC in order to get better data and MC agreement.

A cross-section measurement including the trigger efficiencies and data/MC efficiency scale factors is shown in table 3.7. The SM Z cross section in mass range 60 to 120 GeV at NNLO is $1928\pm72 \text{ pb}$. The result has a good agreement with theory value. it is $\sim 2\%$ for barrel-barrel and $\sim 0.3\%$ for barrel-endcap in 2016, for 2017 it is $\sim 2\%$ for barrel-barrel and $\sim 4\%$ for barrel-endcap.

Year	2016		2017	
Channel	barrel-barrel	barrel-endcap	barrel-barrel	barrel-endcap
N data events	5760345 ± 2400	2051759 ± 1432	6189746 ± 2488	2095959 ± 1448
N expect bkg	32805	11336	32092	10540
MC acc \times eff	$0.0880\pm0.001(\text{stat.})$	$0.0315\pm0.001(\text{stat.})$	$0.0807\pm0.001(\text{stat.})$	$0.0289\pm0.001(\text{stat.})$
Data/MC gsf RECO SF	$0.979\pm X$	$0.985\pm X$	-	-
Data/MC ID Eff SF	$0.943\pm0.001(\text{stat.})$	$0.953\pm0.002(\text{stat.})$	$0.935\pm0.002(\text{stat.})$	$0.947\pm0.004(\text{stat.})$
Luminosity (pb^{-1})	35867	35867	41368	41368
DY cross-section (pb)	$1967\pm3(\text{stat})\pm51(\text{lumi})$	$1922\pm3(\text{stat})\pm50(\text{lumi})$	$1974\pm3(\text{stat})$	$1854\pm3(\text{stat})$
Ratio to theory (1928 pb)	1.02	0.997	1.024	0.962

Table 3.7: The DY cross-section measurement in the range of $60 < M_{ee} < 120 \text{ GeV}/c^2$. The HEEP efficiency data/MC scale factor is taken from table 3.6. Note that in table 3.6 the scale factors are for individual electrons while here they are for electron pairs. For 2017 the data/MC gsf electron reconstruction efficiency scale factor is already included in MC acc \times eff.

mass range (GeV)	relative uncertainty
200-300	1.21%
400-500	1.54%
900-1000	2.16%
1400-1500	2.73%
1900-2000	3.24%
2400-2500	3.72%
2900-3000	4.27%
3400-3500	5.00%
3900-4000	5.94%
4400-4500	7.47%
4900-5000	10.2%
5400-5500	14.3%
5900-6000	19.9%

Table 3.8: The PDF uncertainties relative to the Z peak region as a function of mass.

3.6.1.1 DY Background Correction and Uncertainty

The main uncertainties on the Drell-Yan background originate from PDF and higher-order effects. The cross-section has been evaluated using FEWZ 3.1.b2 with NNLO accuracy in QCD and NLO in EWK. Photon induced effects were taken into account by using a special PDF set, namely the LUXqed_plus_PDF4LHC15_nnlo_100. Cross section ratios relative to the Z peak were estimated together with their uncertainties by taking into account possible correlations of the PDF uncertainties between the various mass bins. Full details of this calculation are in [34]. The cross-sections in various mass bins were evaluated in the analysis acceptance ($E_T > 35$ GeV, $|\eta| < 2.5$, excluding the $1.4442 - 1.566$ region). The ratio of these cross-sections to that predicted by our POWHEG samples generated with NNPDF3.0 is shown in figure 3.37.

It is immediately noted that the POWHEG NNPDF3.0 prediction is increasingly higher than the FEWZ prediction as the mass increases. It is unclear what the analysis should do about this and current approach is to apply a FEWZ to POWHEG k-factor. This would slightly improve data/MC agreement at high mass. The functional form for this k-factor (accounting for the fact we normalise in the 60-120 GeV region) is shown in figure 3.37. This is now applied in the mass spectrum plots and the background estimations for the limits.

The PDF uncertainties for FEWZ 3.1.b2 with the LUXqed_plus_PDF4LHC15_nnlo_100 PDF set are shown in table 3.8. The uncertainties quoted are on the ratio to the Z peak region of 60 to 120 GeV to the invariant mass region in question. The uncertainties are fitted by polynomial which is shown in Figure 3.38.

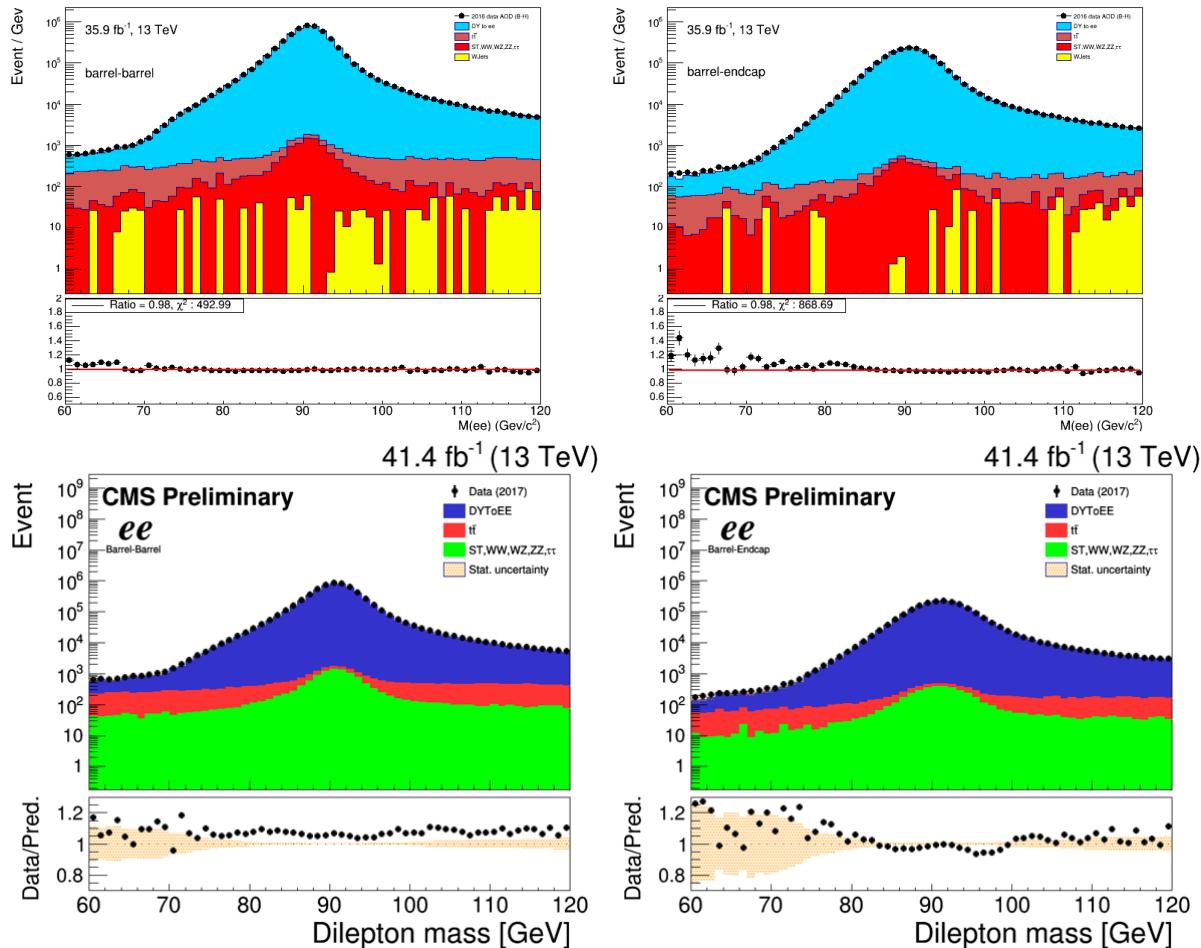


Figure 3.36: Data - MC agreement in the Z peak region with the MC normalised to the luminosity of data for the barrel-barrel (left) and barrel-endcap (right) regions. The trigger turn on curve, gsf reconstruction and HEEP ID scale factor are applied for MC in 2016 (top) and 2017 (bottom).

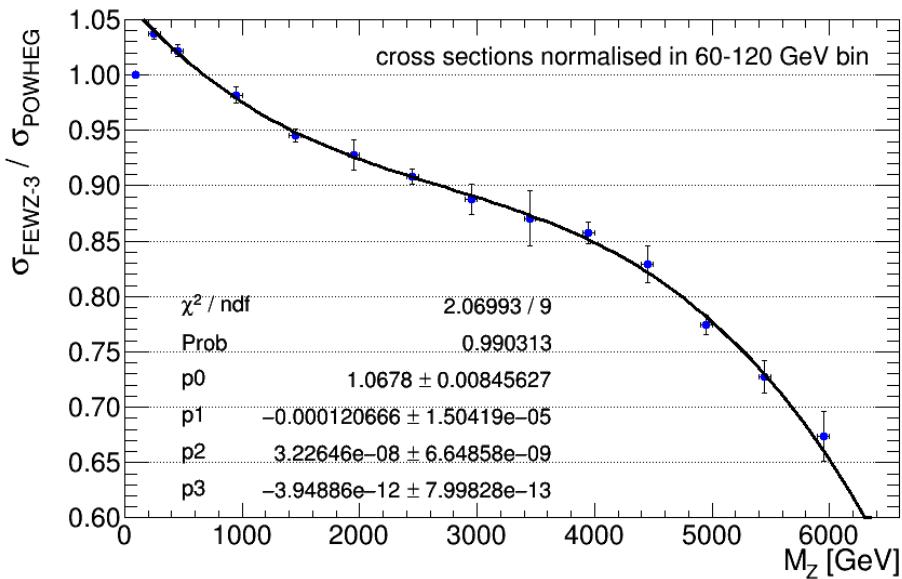


Figure 3.37: The ratio of the $Z/\gamma^* \rightarrow e^+e^-$ cross-section as predicted by FEWZ 3.1.b2 using the LUXqed_plus_PDF4LHC15_nnlo_100 PDF set and that predicted by POWHEG at NLO using the NNPDF3.0 PDF set in various mass bins. The cross-sections are normalised to each other in the 60 to 120 GeV bin.

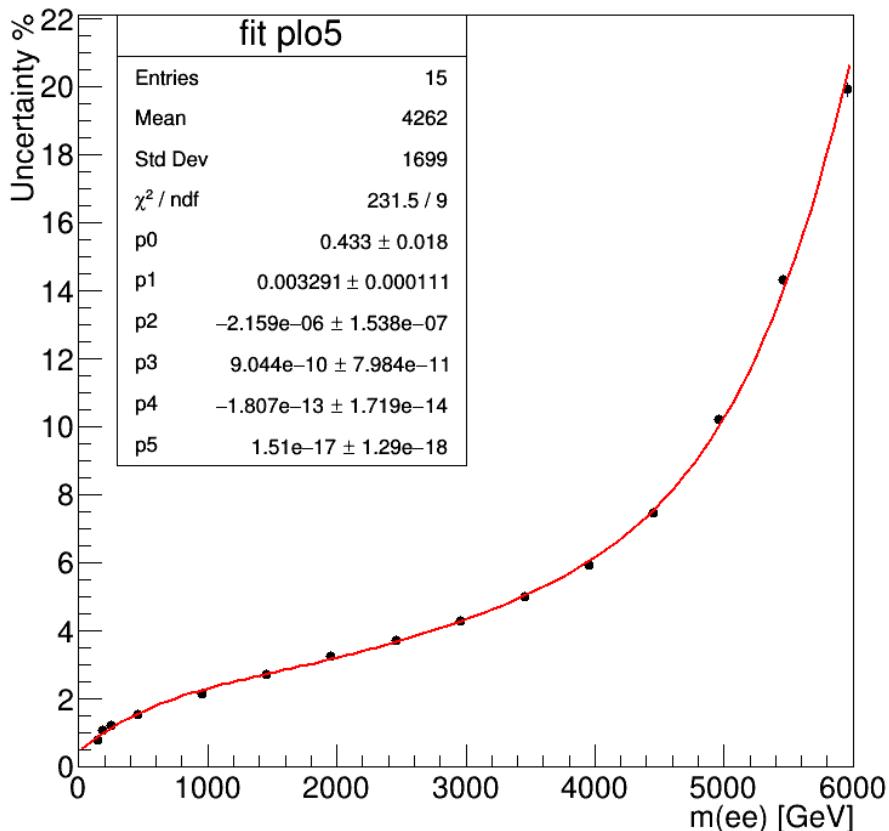


Figure 3.38: The PDF uncertainties relative to the Z peak region as a function of mass.

3.6.2 $t\bar{t}$ and $t\bar{t}$ -like background

The background from electrons arising from non-singularly produced W and Z bosons is estimated directly from Monte Carlo simulation as in the 13 TeV analysis. The leptonic decay of $t\bar{t}$ and tW processes are generated using powheg, while the WW , WZ and ZZ are generated using pythia8. The $Z/\gamma^* \rightarrow \tau\tau$ processes are generated using amc@NLO.

All these processes are flavour-symmetric and have a branching ratio to a pair of leptons of different flavour, $e\mu$, twice as large as the branching ratio to e^+e^- . This means the $e - \mu$ mass spectrum provides an excellent control region to validate the Monte Carlo predictions for these backgrounds. This validation is described below. In order to study the $e - \mu$ mass spectrum, we use the SingleMuon datasets and 'HLT_Mu50' trigger is applied on data events. The electron-muon events are selected such that the first object is a global muon passing the high p_T muon identification criteria[35], and the second object is an electron passing the HEEP ID selection. Electron is required to have $p_T > 35$ GeV while muon is required $p_T > 60$ GeV. Since high energetic muons can produce bremsstrahlung and an associated super-cluster in the ECAL in the direction of the muon's inner track, the selected muons can lead to fake electron candidates. Therefore, an electron veto is applied such that if there is a global muon with $p_T > 5$ GeV within $\Delta R < 0.1$ of the electron, the electron is not selected. Finally the electron and muon should have opposite sign. The Monte Carlo samples and the cross-sections used are documented in table 3.2, all scale factors have been applied to mc event.

The estimation of the multi-jet background from simulated samples is not feasible because of the small misreconstruction rate for the jets. Instead, the multi-jet background is obtained from the same sign $e\mu$ spectrum, where the electron and muon have the same charge. The contributions of the other SM processes (estimated from simulations) are subtracted from the data spectrum in same sign $e\mu$ spectrum and the remaining spectrum is taken to come from multi-jet events. For the multi-jet background, the spectrum for the same sign or opposite sign $e\mu$ pair should be the same.

The invariant mass spectrum of $e\mu$ events in data and MC spectrum are shown for both datasets in Figure 3.39.

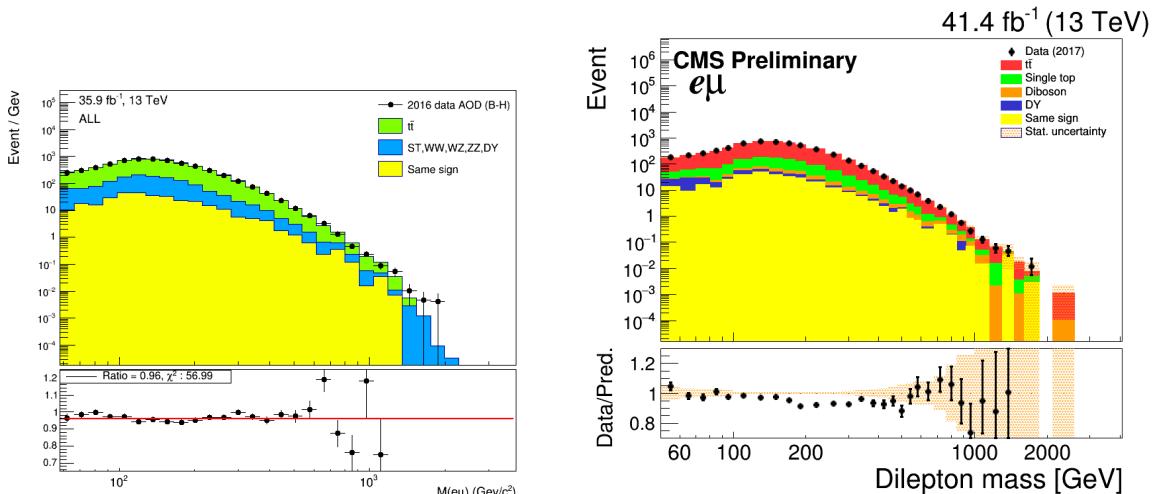


Figure 3.39: Invariant mass spectra of the $e\mu$ events in data and MC for 2016 (left) and 2017 (right).

After having checked that the simulations describe the sample of $e\mu$ events well, the contribution of these backgrounds to the ee spectrum is estimated from Monte Carlo. Tendency of data/MC ratio is known effect from top p_T , which is not corrected in above

plot.

3.6.3 Jet background

The jet background is the smallest of the significant SM backgrounds considered in this search. The primary components of the jet background are: di-jet events, where both jets have passed electron ID criteria, $W+jets$ events, where the W decays to an $e\nu$ pair and a jet passes the electron ID criteria and $\gamma+jets$ events, where the photon and a jet passed electron ID criteria. The jet background is estimated using the ‘fake rate method’ which is explained below.

3.6.3.1 Jet to electron misidentification probability measurement

The probability for a jet reconstructed as an electron to pass the HEEP ID selection, here after referred to as a ‘fake rate’, is measured using data events triggered by single photon triggers.

The fake rate is measured in bins of E_T^{HLT} and η_{SC} , where E_T^{HLT} is the transverse energy of the electron as obtained by the HLT and η_{SC} is the η of the RECO supercluster of the electron w.r.t 0,0,0. The fake rate is relatively flat across the barrel but increases with increasing η in the endcap. Therefore, while the barrel is one η bin, the endcap is split into two η bins. Given the rate of change of the fake rate with E_T and that the overall precision is not high, using RECO E_T on a HLT E_T parameterized function is not expected to have any significant effect.

The fake rate is measured with respect to an electron candidate passing both the fake rate pre-selection in table 3.9 and the first leg of either the HLT_DoubleEle33_CaloIdL_GsfTrkIdVL or HLT_DoubleEle33_CaloIdL_MW trigger, with the exact trigger requirement being run dependent. There can be only one ECAL driven reconstructed electron with $E_T > 10$ GeV and $H/E < 0.15$ in the event to reduce contamination from $Z/\gamma^* \rightarrow e^+e^-$ events. The fake rate is therefore simply the number of misidentified jets in this sample which then go on to pass the HEEP selection.

The number of misidentified jets in this sample is estimated using track isolation template for jets normalised to the observed N-1 track isolation distribution for each bin the fake rate is calculated in. The N-1 track isolation distribution is simply the track isolation distribution for electron candidates in the fake rate sample which pass the HEEP selection with the track isolation cut removed.

The distribution of the track isolation for jets is obtained by requiring the electron candidate to pass the H/E and calorimeter isolation cuts but fail at least one other cut. In practice the cuts which it is possible for the electron candidate to fail are the $\Delta\eta_{in}^{seed}$, $\Delta\phi_{in}$, and shower shape cuts. The H/E and calorimeter isolation variables are strongly correlated with the track isolation variable and misidentified jets which pass these cuts will have a different track isolation distribution to misidentified jets which fail this cuts.

The jet track isolation template is then normalised to the observed N-1 track isolation distribution in the range of $10 < \text{Isol } p_T \text{ Tracks} < 20$ GeV. Any signal contamination in this region is small and is predicted to be a maximum of a few percent from Monte Carlo simulated events. To be clear, it is this feature that allows us to forget a signal template.

Then the prediction of the normalised jet template for the number electron candidates with $\text{Isol } p_T \text{ Tracks} < 5$ GeV is taken as the number of misidentified jets passing HEEP selection.

To summarise the method, the number of jets passing HEEP ID is the number of events in the jet tracker isolation template below 5 GeV once that template has been normalised to observed N-1 track isolation distribution in a region where there are no signal events.

The measured fake rates in 2016 (2017) are shown in figure 3.40 (3.41), together with simple fitted functions to allow the fake rate to be easier to apply in the analysis. The fit

variable	barrel	endcap
$\sigma_{inj\eta}$	<0.013	<0.034
H/E	<0.15	<0.10
nr. missing hits	≤ 1	≤ 1
$ dxy $	< 0.02	< 0.05

Table 3.9: The selection requirements for the starting point of the fake rate calculation.

Year	region	E_T range (GeV)	functional form
2016	barrel	$35.0 \leq E_T < 131.6$ GeV	$0.106 - 0.0025 \times E_T + 2.26 \times 10^{-5} \times E_T^2 - 7.11 \times 10^{-8} \times E_T^3$
		$131.6 \leq E_T < 359.3$ GeV	$0.0139 - 0.000104 \times E_T + 3.6 \times 10^{-7} \times E_T^2 - 4.13 \times 10^{-10} \times E_T^3$
		$E_T \geq 359.3$ GeV	$0.00264 + 3.38 \times 10^{-6} \times E_T$
	endcap $ \eta < 2.0$	$35.0 \leq E_T < 122.0$ GeV	$0.117 - 0.0013 \times E_T + 4.67 \times 10^{-6} \times E_T^2$
		$122.0 \leq E_T < 226.3$ GeV	$0.0345 - 4.76 \times 10^{-5} \times E_T$
		$E_T \geq 226.3$ GeV	$0.0258 - 9.09 \times 10^{-6} \times E_T$
	endcap $ \eta > 2.0$	$35.0 \leq E_T < 112.5$ GeV	$0.0809 - 0.000343 \times E_T$
		$E_T \geq 112.5$ GeV	0.0423
2017	barrel	$35.0 \leq E_T < 131.6$ GeV	$0.140 - 0.0029 \times E_T + 2.56 \times 10^{-5} \times E_T^2 - 8.48 \times 10^{-8} \times E_T^3$
		$131.6 \leq E_T < 359.3$ GeV	$0.020 - 0.00013 \times E_T + 3.50 \times 10^{-7} \times E_T^2 - 2.90 \times 10^{-10} \times E_T^3$
		$E_T \geq 359.3$ GeV	$0.00514 + 4.73 \times 10^{-7} \times E_T$
	endcap $ \eta < 2.0$	$35.0 \leq E_T < 125.0$ GeV	$0.1012 - 0.00094 \times E_T + 3.37 \times 10^{-6} \times E_T^2$
		$125.0 \leq E_T < 226.3$ GeV	$0.0488 - 11.37 \times 10^{-5} \times E_T$
		$E_T \geq 226.3$ GeV	$0.0241 - 1.24 \times 10^{-6} \times E_T$
	endcap $ \eta > 2.0$	$35.0 \leq E_T < 152.$ GeV	$0.0622 - 0.00012 \times E_T$
		$E_T \geq 152.$ GeV	0.0387

Table 3.10: The functional approximation of the measured fake rate for HEEP electrons in the barrel and endcap vs E_T .

parameters are summarised in table 3.10. A 50% uncertainty on the method is assumed and this seems cover deviations from the arbitrarily chosen fits. More details can be found in [29].

3.6.3.2 Jet mass spectrum estimates

The jet background is estimated by selecting electron pairs passing the primary analysis trigger with one electron passing the HEEP selection and one electron passing the fake rate (FR) pre-selection in table 3.9 but failing the HEEP selection. This is referred to as the 1FR estimate. The events are then weighted by $FR/(1-FR)$ where FR is the E_T and η appropriate fake rate for the electron failing the HEEP selection. In the case of more than one electron pair in the same event satisfying these conditions, all valid pairs are allowed to enter the estimation. There is a residual contamination of the $Z/\gamma^* \rightarrow e^+e^-$ events which is corrected for by directly subtracting off the MC estimate.

The 1FR estimate includes the background from $W+jets$, $\gamma+jets$ and di-jets but due to combinatorial effects, the 1FR estimate overestimates the di-jet contribution by a factor 2. The di-jet component can be estimated by selecting electron pairs where both electrons pass the FR pre-selection but fail the HEEP selection, again selected using the primary analysis trigger. This is referred to as the 2FR estimate. These events are weighted by $FR_1/(1 - FR_1) \times FR_2/(1 - FR_2)$ where FR_1 (FR_2) is the E_T and η appropriate fake rate for the first (second) electron. This estimate is then subtracted off the 1FR estimate to estimate the total jet background without any double counting.

Due to fake rate measurement uncertainties and statistical effects, the 2FR estimate can sometimes be greater than half the 1FR estimate. This implies that the entirety of the 1FR estimate is from di-jets and therefore the true estimate of the di-jet background is simply 50% of its value. Therefore the 1FR estimate can only be reduced to a minimum

of 50% of its uncorrected value.

The fake rate is tested using the invariant mass of two electrons passing HEEP ID and both in endcap region. Because there are more jets in endcap comparing with barrel region. The results are shown in Figure 3.42 (the uncertainty band is explained in 3.7) and one can see the data and MC agree within uncertainty.

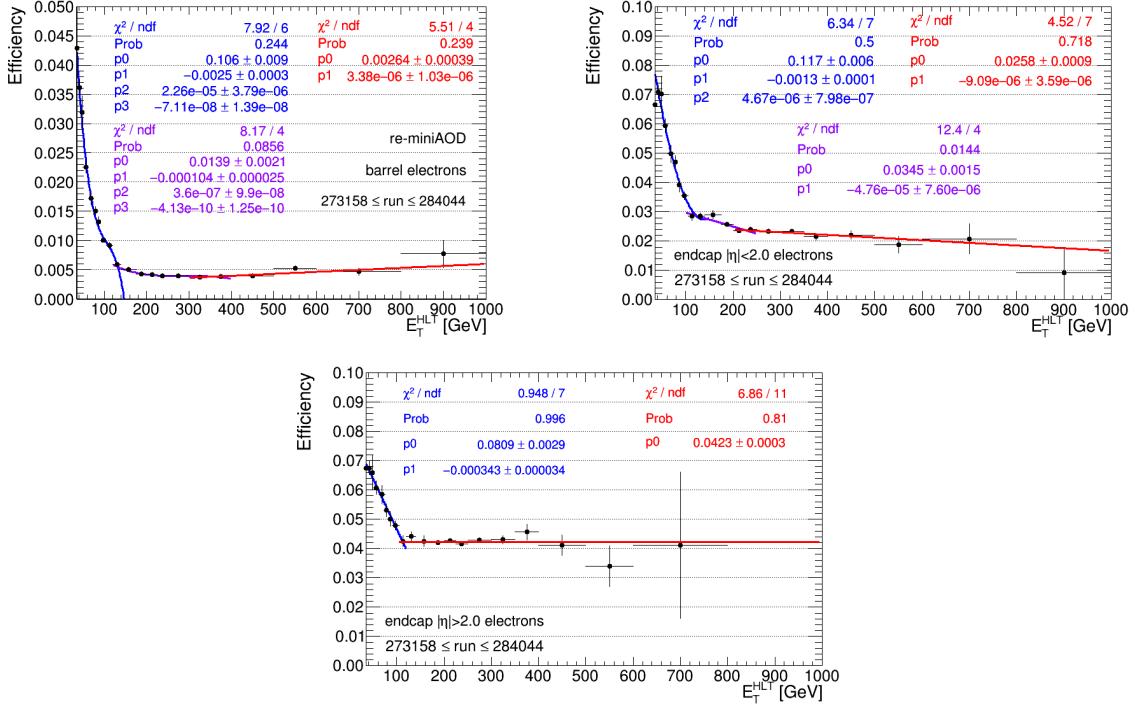


Figure 3.40: The measured HEEP ID fake rate vs E_T for the barrel region (top left), the endcap low η region (top right) and the endcap high η region (bottom) in 2016.

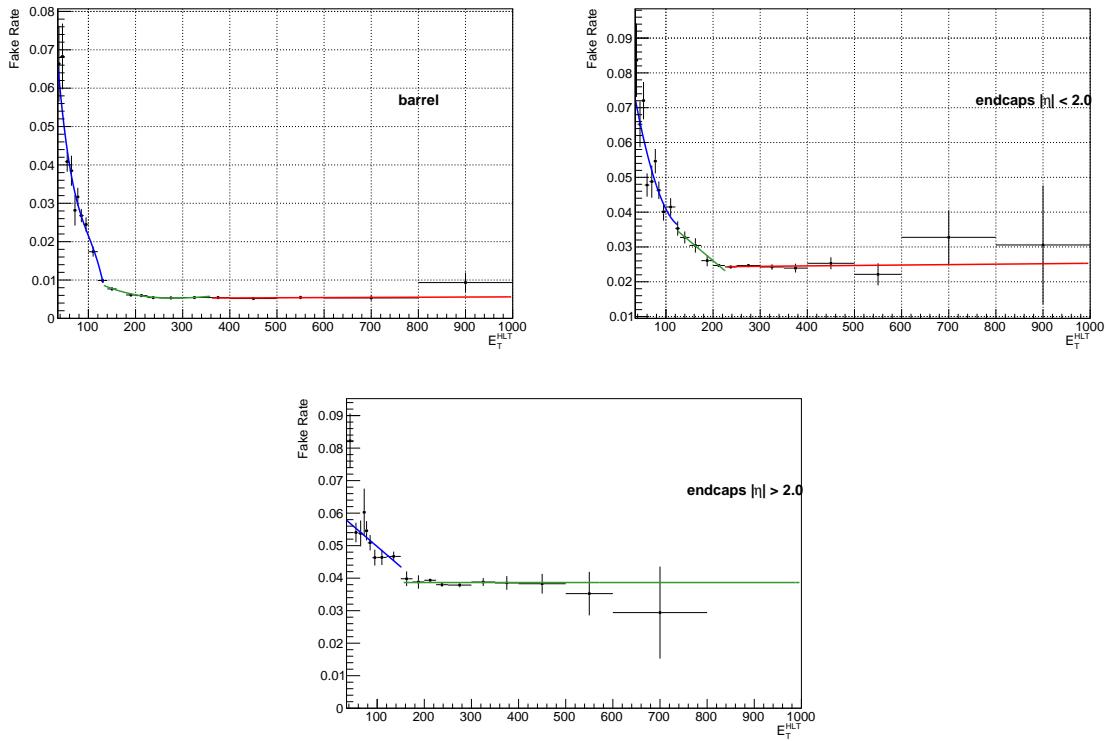


Figure 3.41: The measured HEEP ID fake rate vs E_T for the barrel region (top left), the endcap low η region (top right) and the endcap high η region (bottom) in 2017.

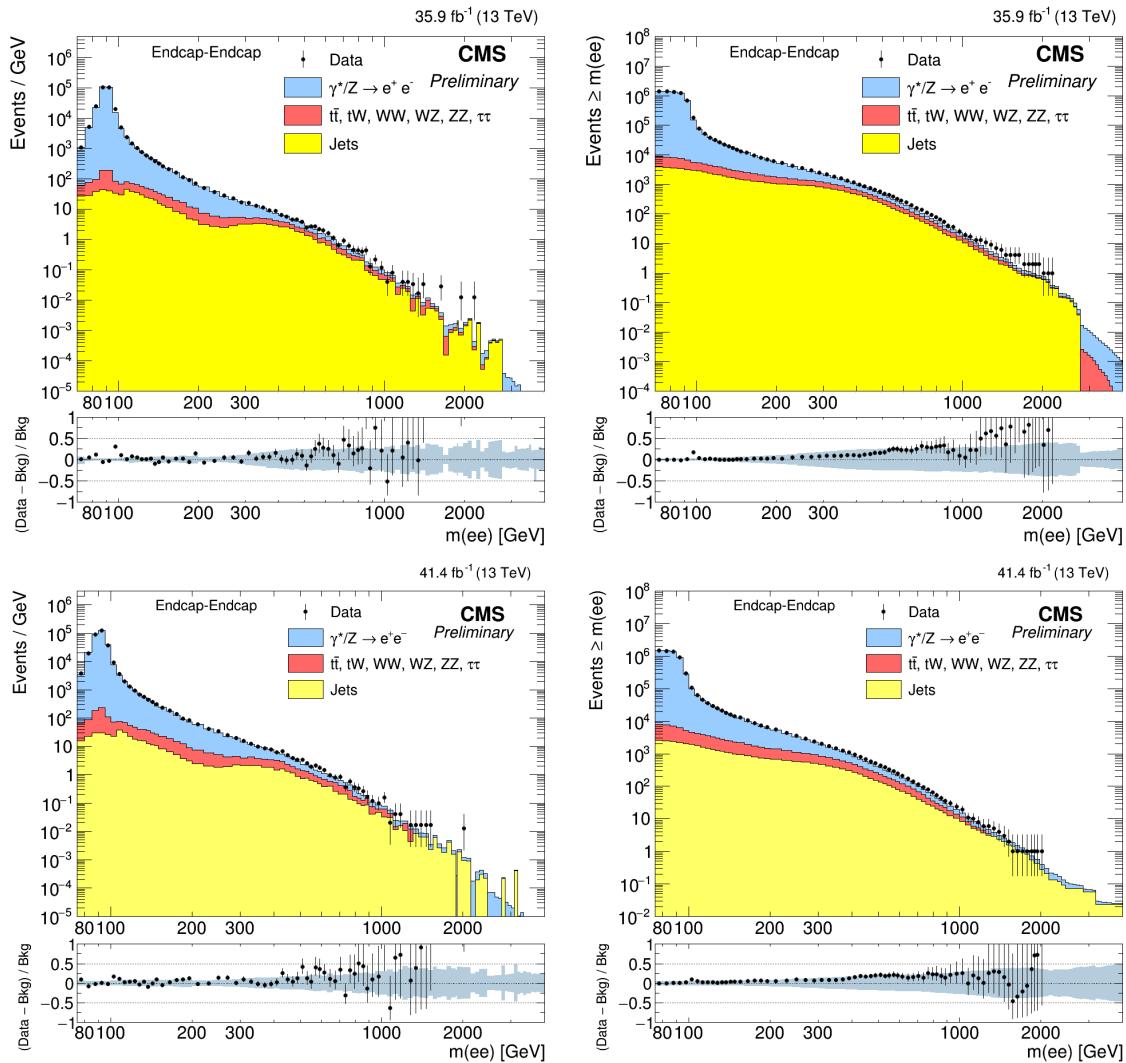


Figure 3.42: The dielectron mass spectrum (left) and the cumulated distribution (right) for both electrons in endcap in 2016 (top) and 2017 (bottom).

3.7 Invariant mass spectra

The observed dielectron mass spectrum together with the predicted standard model backgrounds and the integral of the mass spectrum are shown in figure 3.43 (3.45) for the barrel-barrel and barrel-endcap channels and combined channels in 2016 (2017). In Figure 3.44 (3.46), the ratio of observed dielectron mass spectrum in signal region for the barrel-barrel, barrel-endcap and sum of the barrel-barrel and the barrel-endcap together is shown with different binning in 2016 (2017). The following systematic uncertainties sources are considered.

- normalization to Z peak: 1% uncertainty is considered on the normalization factor estimated from Z peak in 2016. In 2017 2% (4%) uncertainty is considered for barrel-barrel (barrel-endcap).
- pileup reweighting: scaling the minimum bias cross section up/down by 4.6%
- DY PDF: relative mass dependent uncertainty from DY PDF (see Figure 3.38)
- cross section of background processes: 7% uncertainty is considered on the cross section of the Non-DY background estimated from MC for different processes.
- jets: 50% uncertainty is considered on fake jet estimation
- electron energy scale: 2% in Barrel-Barrel and 1% in Barrel-Endcap (for mass > 120)
- HEEP ID scale factor: For barrel it is 1% below 90 GeV and 1-3% linearly increase for 90GeV-1TeV range and 3% for higher than 1 TeV. For endcap it is 1% below 90 GeV and 1-4% linearly increase for 90GeV-300 GeV and 4% for higher than 300 GeV in 2016. In 2017 it is the same for barrel, for endcap it is 2% below 90 GeV and 1-5% linearly increase for 90GeV-300 GeV and 5% for higher than 300 GeV.

In Table 3.11, predicted SM background and observed data yields are shown as a function of dielectron invariant mass.

In Table 3.12, the relative effect of each systematic uncertainty is shown as a function of dielectron invariant mass.

3.7.1 Complementary plot

In addition to the invariant mass plot, the distributions of the following variables are shown in figures 3.47 (3.48) for 2016 (2017).

- invariant mass of the selected electron pair in Z peak region for different categories
- E_T , η and ϕ of leading electron and sub-leading electron for $M_{ee} > 200$ GeV
- ΔR and $\Delta\phi$ between selected electrons for $M_{ee} > 200$ GeV
- P_T of the reconstructed Z for $M_{ee} > 200$ GeV

Table 3.11: Predicted SM background and observed data yields as a function of dielectron invariant mass for Barrel-Barrel, Barrel-Endcap and Barrel-Barrel plus Barrel-Endcap. The uncertainty contains statistic uncertainty and systematic uncertainty

Year	Mass range	data	total bkg	$\gamma^*/Z \rightarrow ee$	$t\bar{t}$ and tt-like bkg	Jets	
Barrel-Barrel							
2016	60 – 120	5760346	5762889.7 \pm 133911.3	5730973.8 \pm 133096.7	29369.6 \pm 1277.6	2546.3 \pm 1273.2	
	120 – 400	146598	152496.1 \pm 11452.7	120819.5 \pm 10364.5	29824.2 \pm 1572.3	1852.5 \pm 926.2	
	400 – 600	2163	2295.3 \pm 183.6	1636.5 \pm 124.6	643.5 \pm 63.2	15.4 \pm 7.7	
	600 – 900	523	520.1 \pm 51.1	425.9 \pm 39.2	91.8 \pm 13.3	2.4 \pm 1.2	
	900 – 1300	100	107.9 \pm 10.6	96.5 \pm 9.4	10.9 \pm 1.8	0.6 \pm 0.3	
	1300 – 1800	24	21.5 \pm 2.8	20.5 \pm 2.6	0.9 \pm 0.2	0.1 \pm 0.0	
	1800 – 6000	3	5.4 \pm 0.9	5.2 \pm 0.9	0.2 \pm 0.0	0.0 \pm 0.0	
	Barrel-Endcap						
	60 – 120	2051759	2054401.1 \pm 40405.8	2042472.6 \pm 40296.8	9270.3 \pm 345.6	2658.2 \pm 1329.1	
	120 – 400	98503	99151.9 \pm 4158.3	77350.1 \pm 3306.5	17813.2 \pm 728.6	3988.6 \pm 1994.3	
	400 – 600	2134	2117.0 \pm 112.6	1243.8 \pm 58.0	751.3 \pm 44.6	121.9 \pm 61.0	
	600 – 900	420	463.3 \pm 27.9	311.7 \pm 18.5	128.3 \pm 9.9	23.2 \pm 11.6	
	900 – 1300	82	78.4 \pm 5.9	59.2 \pm 4.6	15.9 \pm 1.6	3.3 \pm 1.7	
	1300 – 1800	9	12.8 \pm 1.2	10.3 \pm 1.0	1.9 \pm 0.4	0.6 \pm 0.3	
	1800 – 6000	6	2.1 \pm 0.3	1.8 \pm 0.2	0.1 \pm 0.0	0.1 \pm 0.1	
	Barrel-Barrel + Barrel-Endcap						
2017	60 – 120	7812105	7817300.8 \pm 166656.9	7773476.7 \pm 165766.3	38627.6 \pm 1524.1	5196.5 \pm 2598.2	
	120 – 400	245101	2520714.4 \pm 12933.9	198526.1 \pm 11326.7	47709.5 \pm 2148.8	5835.8 \pm 2917.9	
	400 – 600	4297	4424.5 \pm 232.4	2887.1 \pm 147.9	1400.2 \pm 87.6	137.2 \pm 68.6	
	600 – 900	943	985.9 \pm 63.9	739.3 \pm 48.7	221.0 \pm 17.2	25.6 \pm 12.8	
	900 – 1300	182	186.6 \pm 13.5	155.9 \pm 11.9	26.8 \pm 2.3	3.9 \pm 1.9	
	1300 – 1800	33	34.3 \pm 3.4	30.9 \pm 3.2	2.8 \pm 0.5	0.6 \pm 0.3	
	1800 – 6000	9	7.5 \pm 1.1	7.0 \pm 1.1	0.3 \pm 0.0	0.1 \pm 0.1	
	Barrel-Barrel						
	60 – 120	6190697	6194808.2 \pm 178177.2	6156571.2 \pm 177324.0	34116.3 \pm 1921.1	4120.7	
	120 – 400	162005	167925.2 \pm 13932.9	132981.8 \pm 12618.8	33128.5 \pm 2365.8	1815.0	
	400 – 600	2503	2404.7 \pm 215.9	1782.7 \pm 144.5	605.2 \pm 79.6	16.8	
	600 – 900	588	560.2 \pm 51.5	478.6 \pm 44.9	78.7 \pm 15.6	3.0	
	900 – 1300	118	113.4 \pm 13.1	105.1 \pm 11.4	7.8 \pm 2.8	0.5	
	1300 – 1800	28	23.1 \pm 3.0	23.0 \pm 3.0	0.0 \pm 0.0	0.1	
	1800 – 6000	7	5.7 \pm 1.0	5.7 \pm 1.0	0.0 \pm 0.0	0.0	
	Barrel-Endcap						
	60 – 120	2096490	2098260.5 \pm 96902.6	2086010.3 \pm 96566.8	10473.3 \pm 675.3	1777.0	
	120 – 400	109771	110357.8 \pm 6227.6	87277.3 \pm 5107.0	19860.0 \pm 1450.6	3220.6	
	400 – 600	2365	2364.5 \pm 164.6	1442.8 \pm 99.9	810.5 \pm 68.3	111.3	
	600 – 900	518	488.3 \pm 37.9	341.8 \pm 26.5	124.1 \pm 13.9	22.4	
	900 – 1300	75	86.5 \pm 7.7	69.5 \pm 6.3	14.0 \pm 2.7	3.0	
	1300 – 1800	16	14.2 \pm 1.7	11.7 \pm 1.3	2.0 \pm 1.0	0.6	
	1800 – 6000	3	2.2 \pm 0.3	2.1 \pm 0.3	0.0 \pm 0.0	0.1	
	Barrel-Barrel + Barrel-Endcap						
2017	60 – 120	8287187	8290233.1 \pm 369104.3	8242650.8 \pm 364680.2	44535.5 \pm 2834.7	3046.7	
	120 – 400	271776	280802.3 \pm 18298.7	222376.6 \pm 15710.7	53411.9 \pm 3976.3	5013.9	
	400 – 600	4868	4841.2 \pm 329.4	3267.6 \pm 216.8	1446.0 \pm 128.0	127.6	
	600 – 900	1106	1062.0 \pm 77.4	829.4 \pm 63.0	207.4 \pm 22.6	25.3	
	900 – 1300	193	201.3 \pm 18.0	176.2 \pm 16.0	21.6 \pm 3.9	3.5	
	1300 – 1800	44	37.6 \pm 4.2	34.8 \pm 4.0	2.1 \pm 1.1	0.7	
	1800 – 6000	10	7.9 \pm 1.2	7.8 \pm 1.2	0.0 \pm 0.0	0.1	

Table 3.12: The relative effect of each systematic uncertainty as a function of dielectron invariant mass.

Year	Uncertainty	60-120	120-400	400-600	600-900	900-1300	1300-1800	1800-6000
2016	normalization_scale_up	0.998%	0.972%	0.966%	0.971%	0.977%	0.980%	0.982%
	normalization_scale_down	0.998%	0.972%	0.966%	0.971%	0.977%	0.980%	0.982%
	pdf_scale_up	0.147%	0.165%	1.122%	2.215%	3.753%	6.279%	10.351%
	pdf_scale_down	0.096%	0.025%	1.000%	1.774%	3.769%	6.397%	9.928%
	energy_scale_up	0.201%	4.473%	4.147%	4.925%	4.830%	6.523%	8.699%
	energy_scale_down	0.095%	4.325%	3.930%	3.942%	5.073%	5.614%	6.873%
	PU_scale_up	0.472%	0.446%	0.405%	0.827%	0.323%	0.050%	0.729%
	PU_scale_down	0.424%	0.262%	0.353%	0.264%	0.394%	0.367%	0.349%
	bkg_scale_up	0.036%	1.334%	1.870%	1.460%	1.092%	0.958%	0.888%
	bkg_scale_down	0.036%	1.334%	1.870%	1.460%	1.092%	0.958%	0.888%
	SF_scale_up	1.797%	1.778%	2.078%	2.893%	3.065%	3.590%	4.949%
	SF_scale_down	1.711%	1.510%	1.991%	2.144%	3.072%	3.836%	4.355%
	total_scale_up	2.124%	5.111%	5.231%	6.426%	7.004%	9.836%	14.477%
	total_scale_down	2.031%	4.877%	4.996%	5.141%	7.188%	9.443%	12.909%
2017	normalization_scale_up	4.033%	3.900%	3.882%	3.895%	3.922%	3.922%	3.937%
	normalization_scale_down	3.992%	3.900%	3.882%	3.895%	3.922%	3.922%	3.937%
	pdf_scale_up	0.114%	0.125%	1.021%	2.102%	3.941%	6.456%	10.549%
	pdf_scale_down	0.114%	0.125%	1.021%	2.102%	3.941%	6.456%	10.549%
	energy_scale_up	0.217%	4.312%	4.009%	4.562%	4.382%	5.875%	8.415%
	energy_scale_down	0.143%	4.660%	4.312%	4.144%	5.083%	5.335%	7.396%
	PU_scale_up	0.446%	0.315%	0.215%	0.510%	0.344%	0.000%	0.414%
	PU_scale_down	0.480%	0.338%	0.247%	0.427%	0.283%	0.096%	0.422%
	bkg_scale_up	0.029%	1.391%	2.036%	1.491%	0.959%	0.952%	0.773%
	bkg_scale_down	0.029%	1.391%	2.036%	1.491%	0.959%	0.952%	0.773%
	SF_scale_up	1.805%	1.515%	1.670%	2.327%	3.477%	4.648%	5.198%
	SF_scale_down	1.806%	1.798%	2.627%	3.422%	4.391%	4.872%	5.162%
	total_scale_up	4.448%	6.176%	6.258%	6.950%	7.953%	10.681%	15.012%
	total_scale_down	4.412%	6.498%	6.768%	7.133%	8.776%	10.496%	14.453%

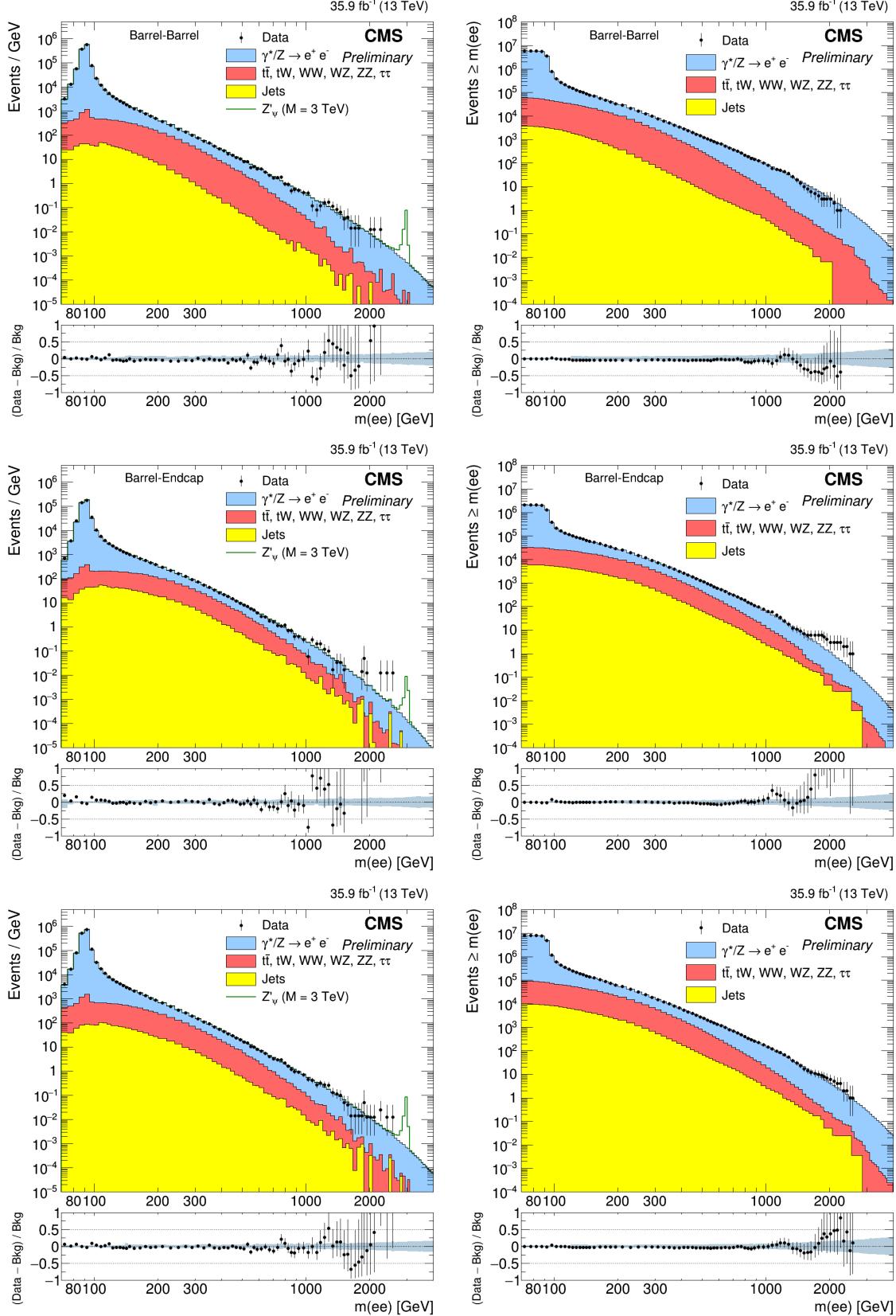


Figure 3.43: The observed dielectron mass spectrum (left) and the integral of the measured dielectron mass spectrum (right) for barrel-barrel (top), barrel-endcap (middle) and sum of the barrel-barrel and the barrel-endcap (bottom) together with the predicted standard model backgrounds in 2016.

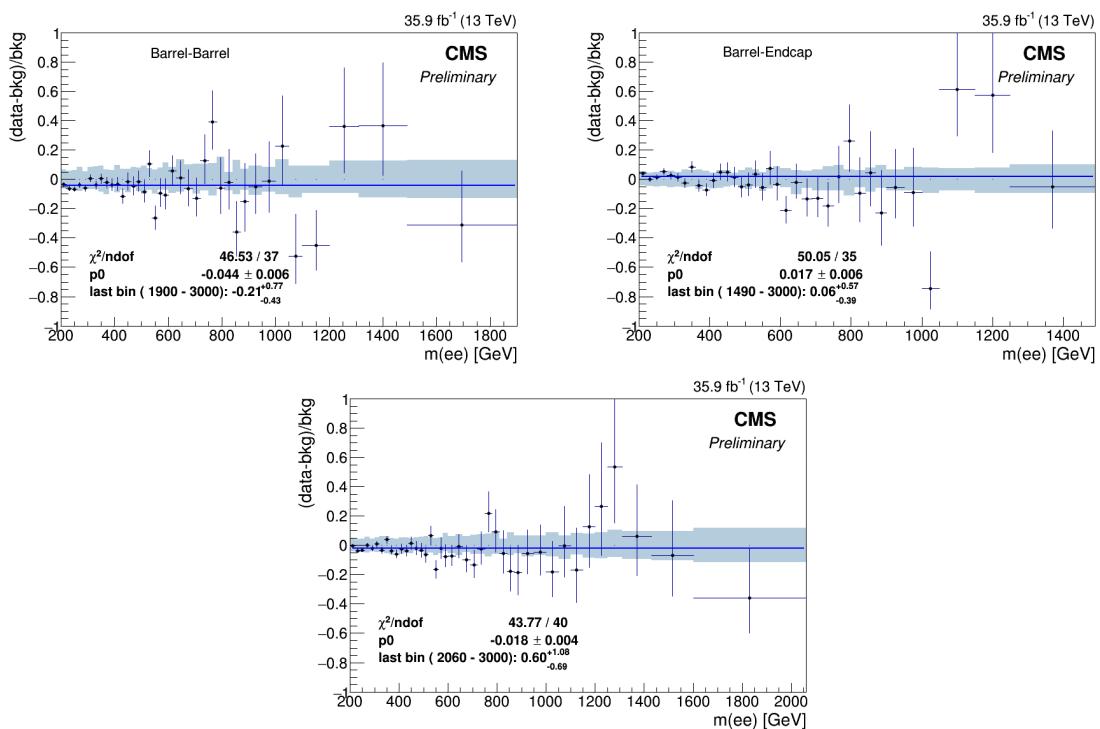


Figure 3.44: The ratio of observed dielectron mass spectrum in signal region for barrel-barrel (top left), barrel-endcap (top right) and sum of the barrel-barrel and the barrel-endcap together (bottom) in 2016.

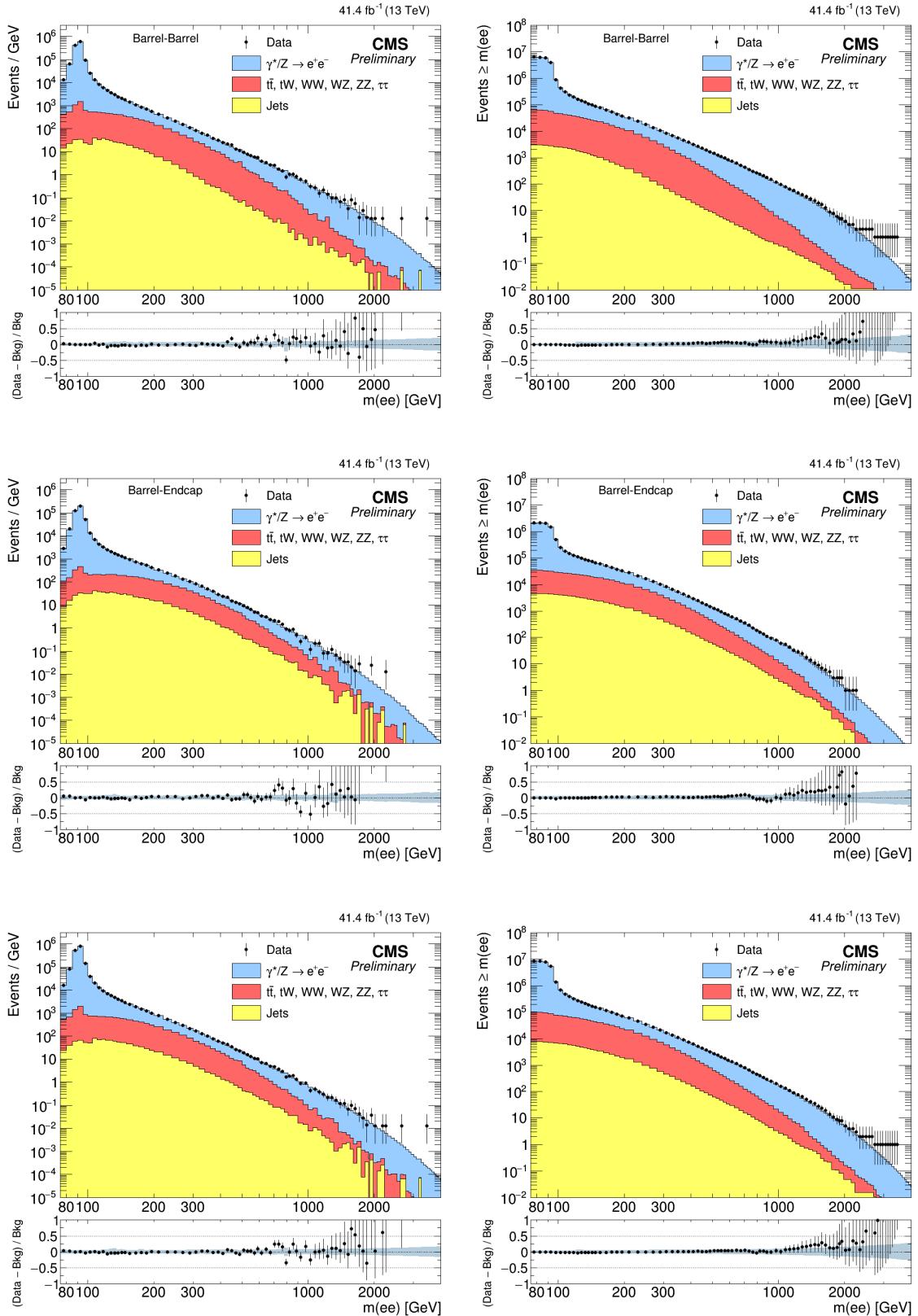


Figure 3.45: The observed dielectron mass spectrum (left) and the integral of the measured dielectron mass spectrum (right) for barrel-barrel (top), barrel-endcap (middle) and sum of the barrel-barrel and the barrel-endcap (bottom) together with the predicted standard model backgrounds in 2017.

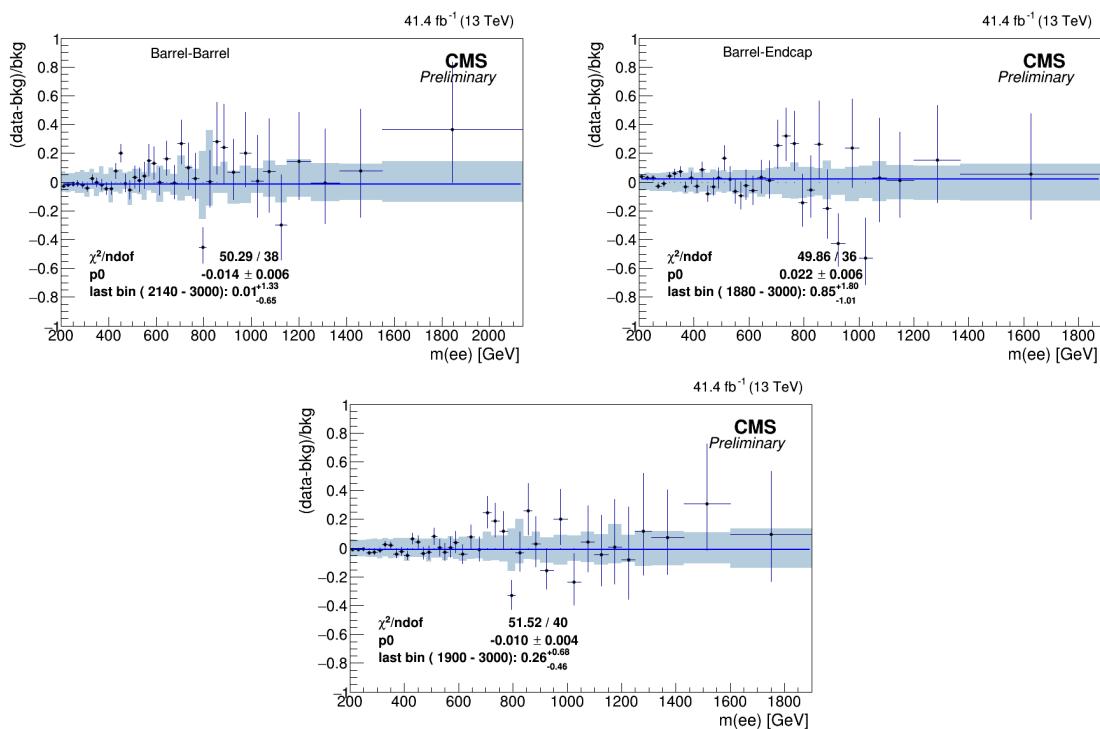


Figure 3.46: The ratio of observed dielectron mass spectrum in signal region for barrel-barrel (top left), barrel-endcap (top right) and sum of the barrel-barrel and the barrel-endcap together (bottom) in 2017.

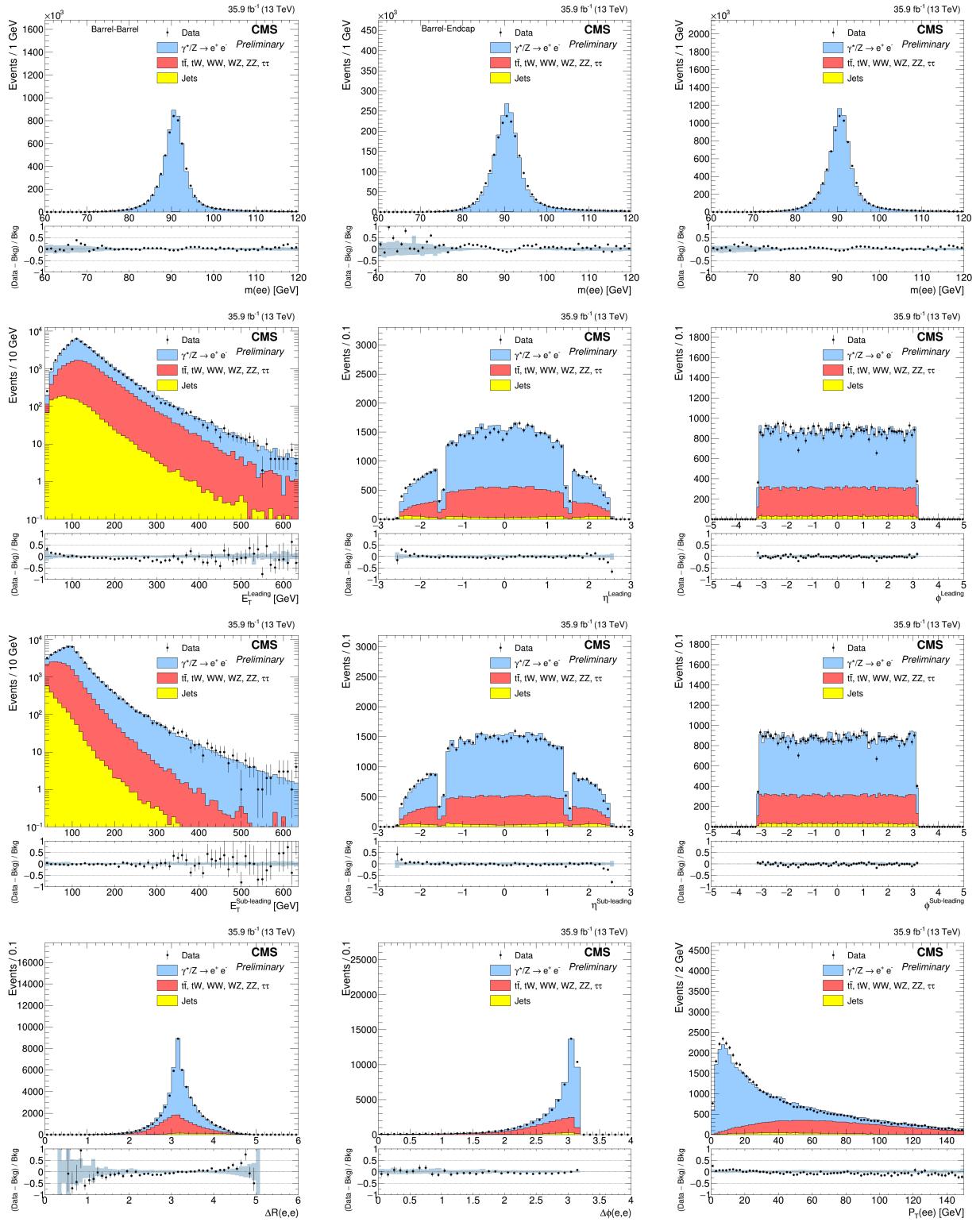


Figure 3.47: The distributions of invariant mass of two electrons in barrel-barrel, barrel-endcap and barrel-barrel + barrel-endcap (first row), E_T , η and ϕ of leading electron (second row) and E_T , η and ϕ of sub-leading electron (third row), ΔR , $\Delta\phi$ between two electrons and P_T of Z (fourth row) in 2016.

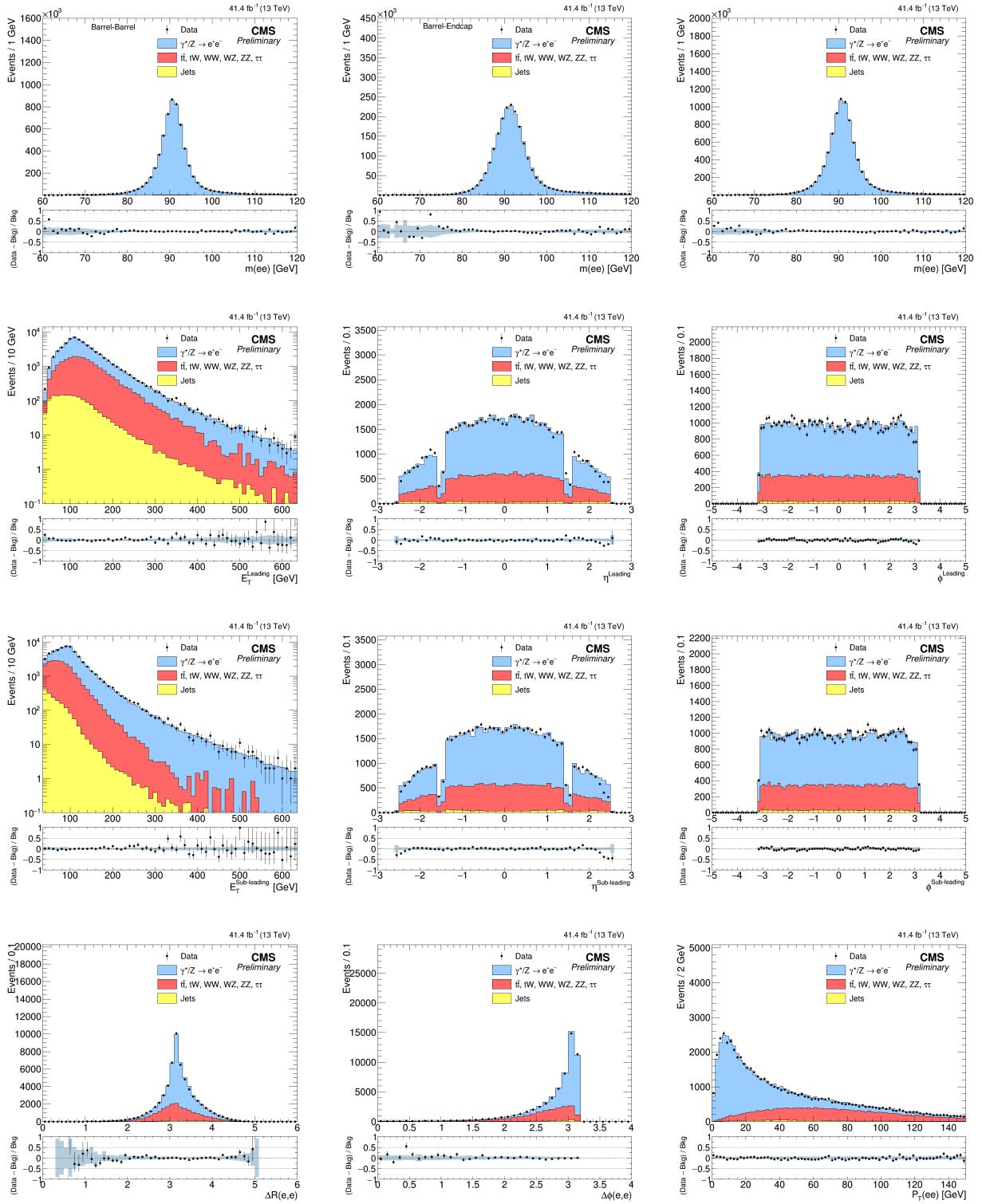


Figure 3.48: The distributions of invariant mass of two electrons in barrel-barrel, barrel-endcap and barrel-barrel + barrel-endcap (first row), E_T , η and ϕ of leading electron (second row) and E_T , η and ϕ of sub-leading electron (third row), ΔR , $\Delta\phi$ between two electrons and P_T of Z (fourth row) in 2017.

3.8 Statistical interpretation

Given that there are not evidence for a significant deviation from the SM expectations is observed, upper limits on the ratio of cross sections of a new resonance to the Z resonance are computed.

Methodology

The statistical treatment of the results follows a Bayesian method with an unbinned extended likelihood function [36, 37, 38]. The mass range considered in the limit is a window of $\pm 6\sigma$ with the window expanded symmetrically so it includes a minimum of 100 events. The barrel-barrel and barrel-endcap channels are treated as two separate channels and then combined together. The probability density function (pdf) is modeled as the sum of a resonant signal pdf and a steeply falling background pdf as follows:

$$f(m|\boldsymbol{\theta}, \boldsymbol{\nu}) = q_1 \cdot f_S(m|\boldsymbol{\theta}, \boldsymbol{\nu}) + (1 - q_1) \cdot f_B(m|\boldsymbol{\theta}, \boldsymbol{\nu}) \quad (3.5)$$

where m is the dilepton invariant mass, $\boldsymbol{\theta}$ is the vector of parameters of interest and $\boldsymbol{\nu}$ the vector of nuisance parameters. The probability of a signal event is given by q_1 . The signal pdf f_S is modeled as a Breit-Wigner convoluted with a resolution function $\text{Res}(m|\sigma, \boldsymbol{\theta})$:

$$f_S(m|\Gamma, \sigma, \boldsymbol{\theta}, \boldsymbol{\nu}) = \text{BW}(m|\Gamma) \otimes \text{Res}(m|\sigma, \boldsymbol{\theta}, \boldsymbol{\nu}) \quad (3.6)$$

where Γ is the intrinsic width of the signal and σ is the mass resolution. As described in Section 3.4 the resolution function $\text{Res}(m|\sigma, \boldsymbol{\theta}, \boldsymbol{\nu})$ is described by a Cruijff (double-sided crystal ball) function in 2016 (2017).

The background pdf f_B has instead an ad-hoc shape derivation computed using simulated background events. An analytic function is used to describe the background shape in the search region above 140 GeV of dilepton invariant mass and it can be expressed by:

$$\begin{aligned} m^\kappa \exp\left(\sum_{i=0}^3 \alpha_i m^i\right), & \text{ if } m \leq 600 \text{ GeV} \\ m^\lambda \exp\left(\sum_{i=0}^3 \beta_i m^i\right), & \text{ if } m > 600 \text{ GeV}, \end{aligned} \quad (3.7)$$

The parameters is determined by a fit to the simulated dilepton mass spectrum. The background spectra together with the fitted function are shown in Figure 3.49 for the barrel-barrel and barrel-endcap categories.

The unbinned likelihood is defined as:

$$\mathcal{L}(\mathbf{m}|R_\sigma, \boldsymbol{\nu}) = \prod_{i=1}^N f(m_i|R_\sigma, \boldsymbol{\nu}) \quad (3.8)$$

where the product is over the events in the dataset and \mathbf{m} is the vector of corresponding dilepton masses and the R_σ is the parameter of interest in this analysis, defined as the ratio between the cross section times branching ratio (BR) to electron pairs of a generic new resonance and the same quantity for the Z resonance in the mass region 60-120 GeV:

$$R_\sigma = \frac{\sigma_{Z'} \cdot \text{BR}(Z' \rightarrow ee)}{\sigma_Z \cdot \text{BR}(Z \rightarrow ee)} \quad (3.9)$$

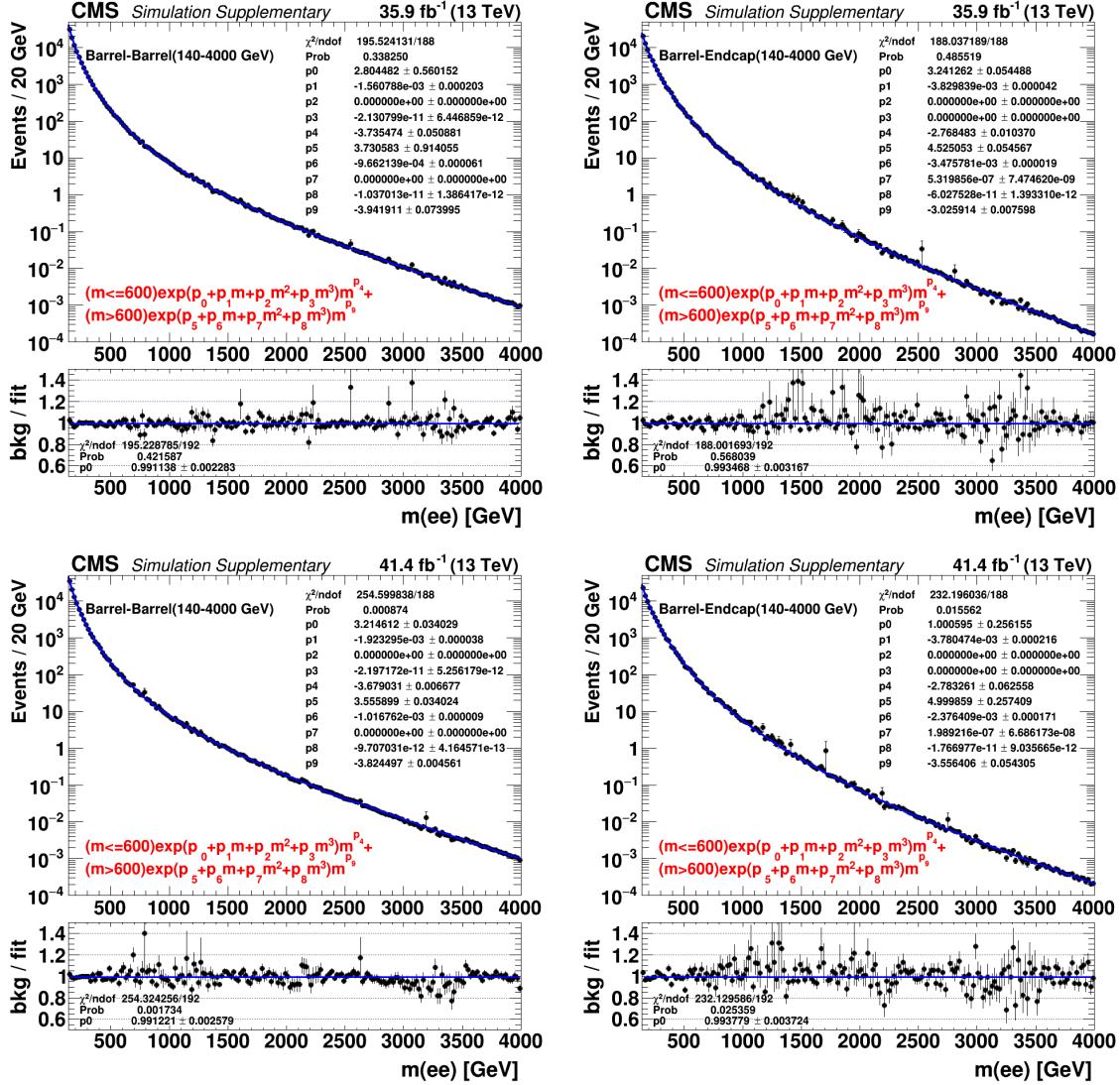


Figure 3.49: The total SM background together with the fitted functional form used to enter it into the limit setting tools for the barrel-barrel (left) and barrel-endcap (right) channels in 2016 (top) and 2017 (bottom).

This choice has the important advantage that certain uncertainties, e.g. the uncertainty on the integrated luminosity or any other E_T -independent effect will be canceled out or are at least greatly reduced. Proceeding in extending the likelihood with a poissonian normalization component in front of equation (3.8) and inserting the equation for the signal and background pdfs detailed above, one obtains:

$$\mathcal{L}(\mathbf{m}|R_\sigma, \boldsymbol{\nu}) = \frac{\mu^N e^{-\mu}}{N!} \cdot \prod_{i=1}^N \left(\frac{\mu_S(R_\sigma, \boldsymbol{\nu})}{\mu} f_S(m_i|R_\sigma, \boldsymbol{\nu}) + \frac{\mu_B(R_\sigma, \boldsymbol{\nu})}{\mu} f_B(m_i|R_\sigma, \boldsymbol{\nu}) \right) \quad (3.10)$$

with μ_S and μ_B being the signal and background yields and μ the sum of the two yields. The R_σ can be connected to the signal event yield μ_S via the following relation:

$$\mu_S = R_\sigma \frac{(\text{Acc} \times \epsilon)_{Z'}}{(\text{Acc} \times \epsilon)_Z} N_Z \quad (3.11)$$

where $(\text{Acc} \times \epsilon)_{Z'}$ and $(\text{Acc} \times \epsilon)_Z$ are the acceptance times efficiency of the Z' (shown in Figure 3.50) and the Z boson respectively and N_Z is the number of selected Z events, defined in the mass region 60-120 GeV.

The uncertainties on the nuisance parameters in the vector $\boldsymbol{\nu}$ are taken into account by modeling the nuisance parameter as

$$\nu = \hat{\nu} \cdot (1 + \delta\nu)^\beta \quad (3.12)$$

where $\hat{\nu}$ is the estimate of ν , $\delta\nu$ is the corresponding systematic uncertainty and β is a random number drawn from a gaussian distribution with mean value at zero and second order moment equal to 1 (denoted as $Gauss(\beta|0, 1)$). The likelihood is then weighted by $Gauss(\beta|0, 1)$ for each nuisance parameter giving

$$\mathcal{L}(\mathbf{m}|R_\sigma, \boldsymbol{\nu}) = \mathcal{L}(\mathbf{m}|R_\sigma, \boldsymbol{\nu}) \cdot \prod_j Gauss(\beta_j|0, 1) \quad (3.13)$$

where the product is done over the nuisance parameters. The two categories of the analysis are independent categories, hence the total likelihood can be obtained by multiplying the two separated likelihoods. With this definition of the likelihood function, 95% confidence level (CL) upper limits can be computed using the Bayes theorem, which states:

$$f(R_\sigma, \boldsymbol{\nu} | \mathbf{m}) \cdot p(\mathbf{m}) = \mathcal{L}(\mathbf{m}|R_\sigma, \boldsymbol{\nu}) \cdot p(R_\sigma, \boldsymbol{\nu}) \quad (3.14)$$

where $p(R_\sigma, \boldsymbol{\nu})$ is the prior pdf for the parameter of interest of the model. In this analysis, the prior is taken as a log-normal distribution for the uncertainties, and a uniform (positive) prior for the parameter of interest. After integrating over the nuisance parameters $\boldsymbol{\nu}$, one obtains

$$p(R_\sigma | \mathbf{m}) \cdot p(\mathbf{m}) = \mathcal{L}(\mathbf{m}|R_\sigma) \cdot p(R_\sigma) \quad (3.15)$$

The expression for the posterior pdf immediately follows as

$$p(R_\sigma | \mathbf{m}) = \frac{\mathcal{L}(\mathbf{m}|R_\sigma) \cdot p(R_\sigma)}{p(\mathbf{m})} = \frac{\mathcal{L}(\mathbf{m}|R_\sigma) \cdot p(R_\sigma)}{\int \mathcal{L}(\mathbf{m}|R_\sigma) \cdot p(R_\sigma) dR_\sigma} \quad (3.16)$$

Given the posterior pdf for the parameter of interest R_σ , the 95% C.L. upper limit R_σ^{95} is defined by the following constraint:

$$\int_0^{R_\sigma^{95}} p(R_\sigma | \mathbf{m}) dR_\sigma = 0.95 \quad (3.17)$$

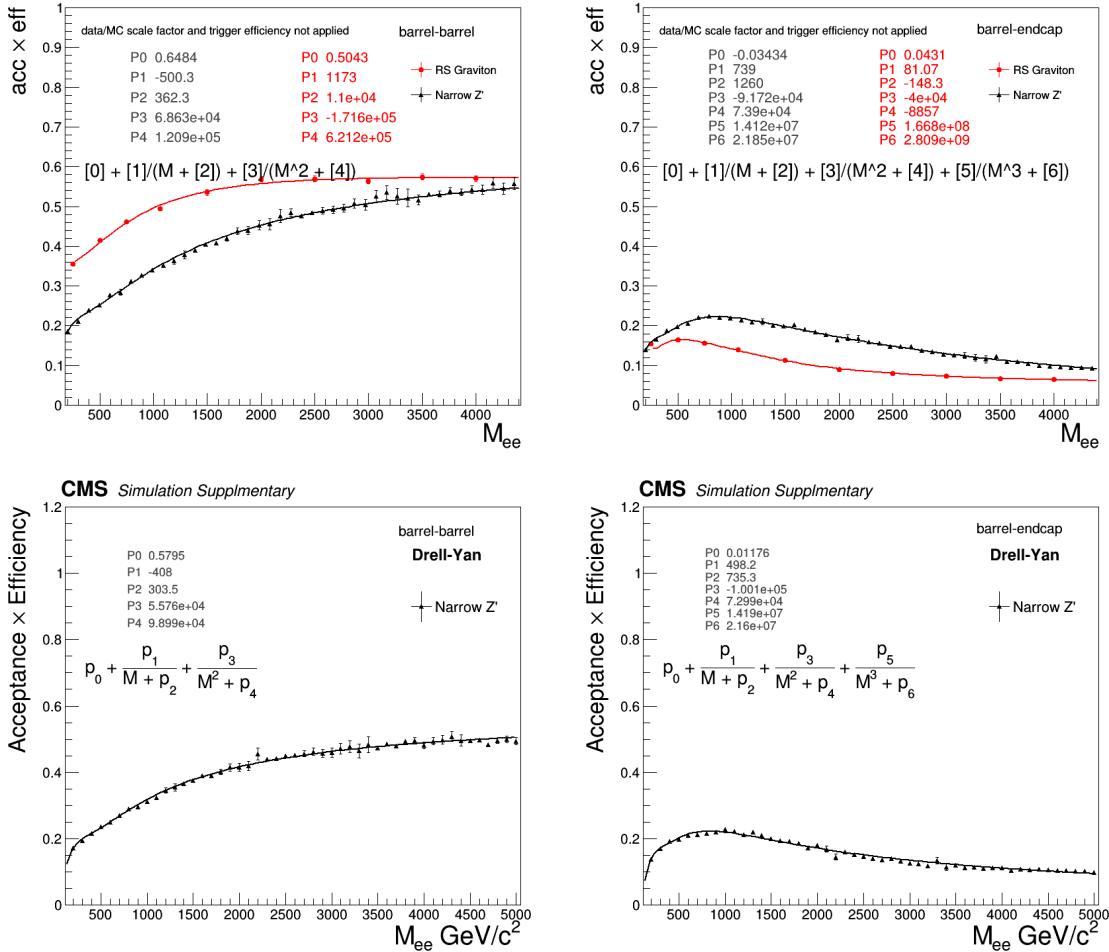


Figure 3.50: The acceptance times efficiency of for a spin-1 or spin-2 particle to be selected by the analysis in the barrel-barrel region (left) and barrel-endcap region(right) together with the fitted functional form used to enter it into the limit setting tools in 2016 (top) and 2017 (bottom). E_T independent effects which will cancel in the ratio to the acceptance times efficiency at the Z peak are not included. These are primarily the data/MC efficiency scale factor and the trigger ID efficiency.

where the integration is done using the Metropolis-Hastings algorithm [39, 40].

Finally, it is worth to mention that (see Equation (3.11)) the parametrization of the acceptance times efficiency ($\text{Acc} \times \epsilon$) $_{Z'}$ is shown in figure 3.50. Other inputs required for the limit setting tool are listed in table 3.13, as well as the number of data events and acceptance at the Z peak region. Since the limits are normalised to the Z peak, any E_T independent effects on the efficiency cancel and are not included in the acceptance times efficiency parametrisation nor the Z peak acceptance x efficiency. The effects not included are the data/MC efficiency scale factor and the trigger ID efficiency. The uncertainty for $(\text{Acc} \times \epsilon)_{Z'}/(\text{Acc} \times \epsilon)_Z$ is mainly due to data/MC HEEP selection scale factor at high E_T as well as NLO and PDF effects on the Drell-Yan background. A 2% (1%) energy scale uncertainty is assigned for barrel-barrel (barrel-endcap) channel.

Expected limits

Expected upper limits on R_σ under the background-only hypothesis are obtained by computing the median of a set of limits derived using an ensemble of randomly drawn

Year	Variable	EB-EB	EB-EE
2016	N_Z (60-120 GeV)	5730976	2042478
	$(\text{Acc} \times \epsilon)_Z$ (60-120 GeV)	0.0895	0.0318
	$(\text{Acc} \times \epsilon)_Z' / (\text{Acc} \times \epsilon)_Z$ err	6%	8%
	energy scale uncertainty	2%	1%
2017	N_Z (60-120 GeV)	6156571	2086010
	$(\text{Acc} \times \epsilon)_Z$ (60-120 GeV)	0.0811	0.0294
	$(\text{Acc} \times \epsilon)_Z' / (\text{Acc} \times \epsilon)_Z$ err	6%	8%
	energy scale uncertainty	2%	1%

Table 3.13: The input parameters to the limit setting code. The MC efficiencies do not have the data/MC scale factor applied or any E_T independent efficiency like HLT identification efficiency although E_T dependent effects like L1 and HLT turn on are included.

pseudo-data. The limits for the pseudo-data are estimated using the same procedure as described for the observed limits. The pseudo-data are generated by drawing the event yield as a random number from a Poisson distribution whose mean is

$$\mu_B = \hat{\mu}_B \cdot (1 + \delta\mu_B)^{\beta_B} \quad (3.18)$$

where β_B is again a random number extracted from a normal distribution as the case of Equation (3.12). The value of $\hat{\mu}_B$ is estimated by integrating the background shape over the observable range, where the shape is normalized over a sideband in the data below 200 GeV of dielectron invariant mass. Repeating this procedure many times, the distribution of the expected limits under the background-only hypothesis is built, therefore the median and the $\pm 1\sigma$ and $\pm 2\sigma$ bands of the limit can be computed.

3.8.1 Upper limits

Using the method described above, the observed and expected 95% upper limits on R_σ for a resonance width of 0.6% is shown in Figure 3.51. The signal R_σ curves are shown on the plot in order to obtain a mass limit on two specific Z' signal model.

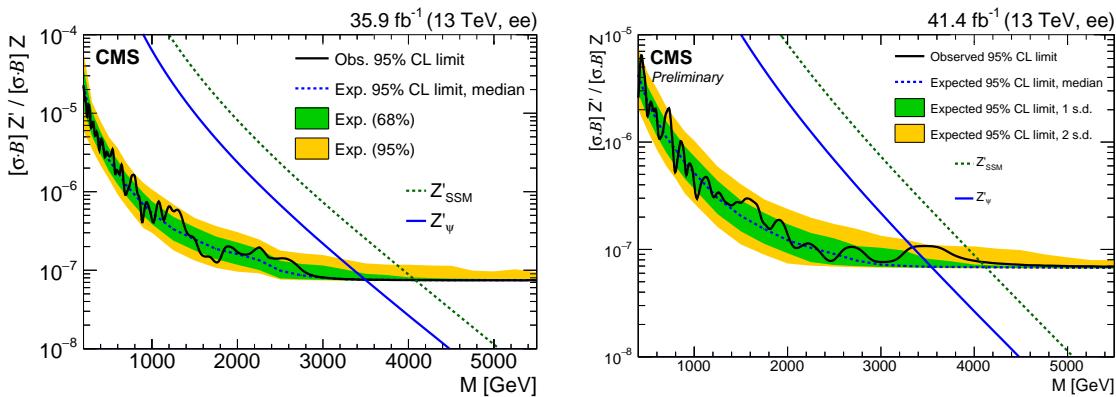


Figure 3.51: The 95% CL upper limits on R_σ for a spin 1 resonance with a width equal to 0.6% of the resonance mass for 2016 (left) and 2017 (right). The shaded bands correspond to the 68% and 95% quantities for the expected limits. Theoretical predictions for the spin 1 Z'_SSM and Z'_ψ resonances are also shown.

In parallel to the search for new resonances in the dielectron final state, a similar search was performed using the dimuon final state [35]. Given that the sensitivity of the two searches is comparable and under the assumption that the BR to dielectron and

dimuon final state is the same, the upper limits coming from the two separate searches can be combined. The upper limits for the combination of the dielectron (using 2016 and 2017 data) and dimuon (using 2016 data) analysis is shown in Figure 3.52.

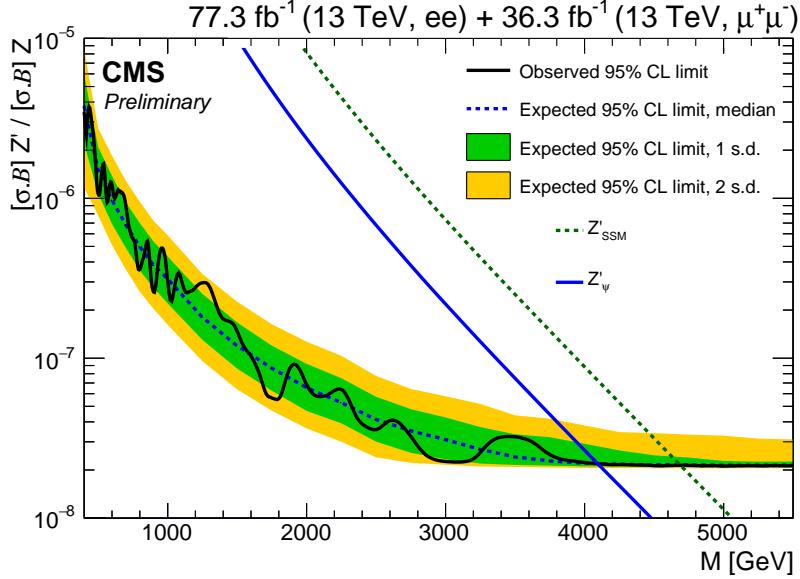


Figure 3.52: The 95% CL upper limits on R_σ for a spin 1 resonance with a width equal to 0.6% of the resonance mass for the dielectron (using 2016 and 2017 data) and dimuon (using 2016 data) final states combined. The shaded bands correspond to the 68% and 95% quantities for the expected limits. Theoretical predictions for the spin 1 Z'_{SSM} and Z'_{ψ} resonances are also shown.

In Addition, In 2016 the results for widths equal to 0.6%, 3%, 5% and 10% of the resonance mass are shown in figure 3.53 and one can see at high masses the limits do not exhibit any dependence on the assumed resonance width.

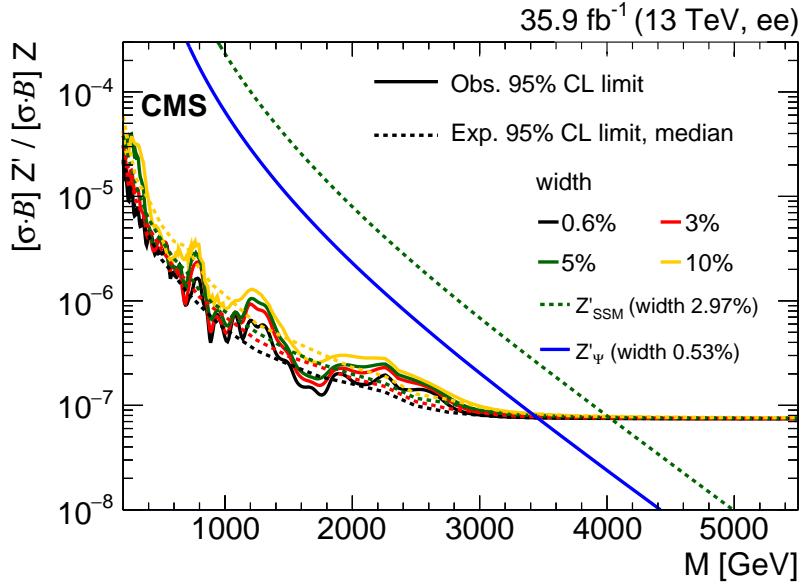


Figure 3.53: The 95% CL upper limits on R_σ for a spin 1 resonance with a width equal to 0.6%, 3%, 5% and 10% of the resonance mass in 2016.

Table 3.14 lists the lower limit on the resonance mass for the Z'_{SSM} and Z'_{ψ} models, for the dielectron final states in 2016 and 2017 and for its combination with dimuon final

states from 2016.

Channel	Z'_{SSM}		Z'_{ψ}	
	Observed (TeV)	Expected (TeV)	Observed (TeV)	Expected (TeV)
ee (2016)	4.10	4.10	3.45	3.45
ee (2017)	4.10	4.15	3.35	3.55
ee (2016 and 2017) + $\mu\mu$ (2016)	4.70	4.70	4.10	4.10

Table 3.14: The observed and expected 95% CL lower limits on the masses of spin 1 Z'_{SSM} and Z'_{ψ} bosons, assuming a signal width of 0.6% (3%) of the resonance mass for Z'_{ψ} (Z'_{SSM}).

Finally, in 2016 the expected and observed limits for a spin-2 resonance with intrinsic widths of 0.01, 0.36, and 1.42 GeV corresponding to coupling parameters k/\bar{M}_{Pl} of 0.01, 0.05, and 0.10 are shown in Figure 3.54 for the dielectron channel and dielectron dimuon combined. Table 3.15 presents the values of the observed and expected 95% CL lower limits of the aforementioned models.

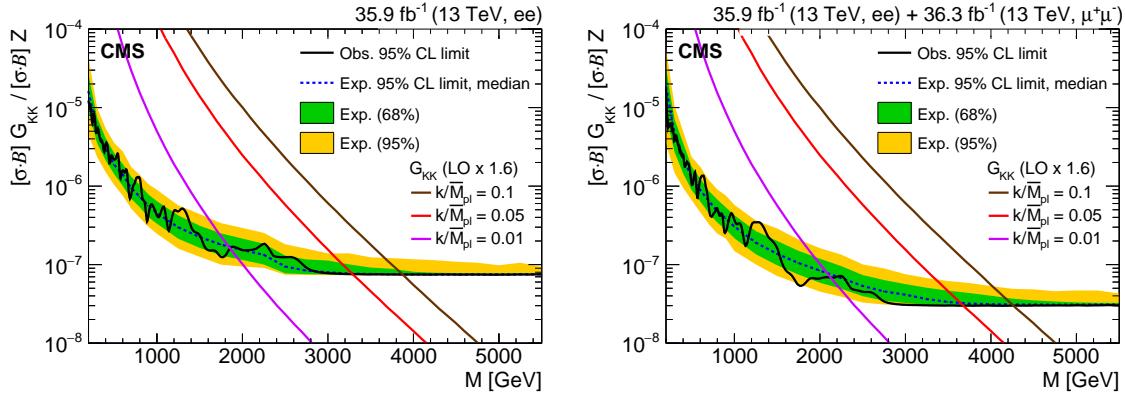


Figure 3.54: The 95% CL upper limits on R_σ for a spin 2 resonance for the dielectron (left) and its combination with dimuon (right) final states in 2016. The shaded bands correspond to the 68% and 95% quantities for the expected limits. Theoretical predictions for the spin 2 resonances with width equal to 0.01, 0.36 and 1.42 GeV corresponding to coupling parameters k/\bar{M}_{Pl} of 0.01, 0.005 and 0.10 are shown for comparison.

Channel	$k/\bar{M}_{Pl}=0.01$		$k/\bar{M}_{Pl}=0.05$		$k/\bar{M}_{Pl}=0.1$	
	Obs. (TeV)	Exp. (TeV)	Obs. (TeV)	Exp. (TeV)	Obs. (TeV)	Exp. (TeV)
ee (2016)	1.85	1.85	3.30	3.30	3.90	3.90
ee (2016) + $\mu\mu$ (2016)	2.10	2.05	3.65	3.60	4.25	4.25

Table 3.15: The observed and expected 95% CL lower limits on the masses of spin 2 resonance with width equal to 0.01, 0.36 and 1.42 GeV corresponding to coupling parameters k/\bar{M}_{Pl} of 0.01, 0.005 and 0.10.

3.9 Summary

A search for narrow resonances in dielectron invariant mass spectra has been performed using data recorded by CMS in 2016 and 2017 from proton-proton collisions at $\sqrt{s} = 13$ TeV. The results of the analysis have been also combined with those of the analogous search in the dimuon final state. The integrated luminosity for the dielectron sample is 35.9 fb^{-1} in 2016 and 41.4 fb^{-1} in 2017, for the dimuon sample it is 36.3 fb^{-1} . Observations are in agreement with standard model expectations. Upper limits at 95% confidence level on the parameter of interest R_σ have been derived.

Limits are set on the masses of various hypothetical particles. For the Z' SSM particle, which arises in the sequential standard model, and for the superstring-inspired Z'_ψ particle, 95% confidence level lower mass limits for the combined channels are found to be 4.70 and 4.10 TeV, respectively. These limits extend the previous ones from CMS by 1.3 TeV in both models. The corresponding limits for Kaluza-Klein gravitons arising in the Randall-Sundrum model of extra dimensions with coupling parameters \bar{M}_{Pl} of 0.01, 0.05, and 0.10 are 2.10, 3.65, and 4.25 TeV, respectively. The limits extend previous published CMS results by 0.6 (1.1) TeV for a \bar{M}_{Pl} value of 0.01 (0.10).

My personal contributions to the analysis include performing regular checks of the detector response by providing invariant mass plots and HEEP selection efficiency using both data and simulations, and on the extraction of the data to simulation scale factor during the data taking periods. Because the HEEP ID scale factor is the vital part of this analysis, many studies and checks have been performed to make sure we have understand it. The HEEP ID is also used in other CMS analyses, so providing HEEP ID scale factor becoming one of my EPR work. Besides, I was responsible for the studies on the mass scale and resolution of the ECAL detector, which are key inputs for the computation of the limits, given that they are used to define the signal model. Moreover, I was responsible for a study of the electromagnetic calorimeter saturation effects although we don't have saturated events in our final mass distribution in data. In addition, I was worked on the fit to the background contribution and estimated the various systematics uncertainties of the analysis. Last but not least I was responsible of providing the final mass spectra.

This results in several publications by the CMS collaboration: two PAS documents, one corresponding to the analysis of 35.9 fb^{-1} of integrated luminosity collected in full 2016 [41] and the other corresponding to the analysis of 41.4 fb^{-1} of integrated luminosity collected in full 2017 [42]. There are one paper [43] where the results obtained analyzing the full 2016 dataset with 35.9 fb^{-1} of integrated luminosity.

The ATLAS collaboration has also provided several publications on the same topic and the latest available results from ATLAS [44] put a 95% confidence level lower mass limit of 4.5 TeV for the Z' SSM model and 3.8 TeV for the Z'_ψ one after combining both dielectron and dimuon final states using 36.1 fb^{-1} data. These results are well in agreement with the ones obtained by CMS and showed in this chapter.

Chapter 4

Search for new physics via top quark production in dilepton final state

This chapter introduced searching for new physics via top quark production in ee and $\mu\mu$ final state. The data-set corresponds to an integrated luminosity of 35.9 fb^{-1} of proton-proton collisions at 13 TeV, and was collected in 2016 by the CMS detector. The search is sensitive to new physics in top quark pair production and in single top quark production in association with a W boson. This is the first search for new physics that uses the tW process. No significant deviation from the standard model expectation is observed. Results are interpreted in the framework of an effective field theory and constraints on the relevant effective couplings are set using a dedicated multivariate analysis.

4.1 Introduction

Due to its large mass, close to the electroweak symmetry breaking scale, the top quark is expected to play an important role in several new physics scenarios. If the new physics scale is in the available energy range of the LHC, the existence of new physics could be directly observed via the production of new particles. Otherwise, new physics could affect standard model (SM) interactions indirectly, through modifications of SM couplings or enhancements of rare SM processes. In the latter case, it is useful to introduce a model-independent approach to parameterize and to constrain possible deviations from SM predictions, independently of the fundamental theory of new physics.

An effective field theory (EFT) approach is followed to search for new physics in the top quark sector in the ee and $\mu\mu$ final states. In Refs. [45, 46] all dimension-six operators that contribute to the top quark pair production ($t\bar{t}$) and the single top quark production in association with a W boson (tW) are investigated. The operators and the related

effective Lagrangian, which are relevant for dilepton final states, can be written as [47]:

$$O_{\phi q}^{(3)} = (\phi^+ \tau^I D_\mu \phi)(\bar{q} \gamma^\mu \tau^I q), \quad L_{\text{eff}} = \frac{C_{\phi q}^{(3)}}{\sqrt{2}\Lambda^2} g v^2 \bar{b} \gamma^\mu P_L t W_\mu^- + h.c., \quad (4.1)$$

$$O_{tW} = (\bar{q} \sigma^{\mu\nu} \tau^I t) \tilde{\phi} W_{\mu\nu}^I, \quad L_{\text{eff}} = -2 \frac{C_{tW}}{\Lambda^2} v \bar{b} \sigma^{\mu\nu} P_R t \partial_\nu W_\mu^- + h.c., \quad (4.2)$$

$$O_{tG} = (\bar{q} \sigma^{\mu\nu} \lambda^A t) \tilde{\phi} G_{\mu\nu}^A, \quad L_{\text{eff}} = \frac{C_{tG}}{\sqrt{2}\Lambda^2} v (\bar{t} \sigma^{\mu\nu} \lambda^A t) G_{\mu\nu}^A + h.c., \quad (4.3)$$

$$O_G = f_{ABC} G_\mu^{A\nu} G_\nu^{B\rho} G_\rho^{C\mu}, \quad L_{\text{eff}} = \frac{C_G}{\Lambda^2} f_{ABC} G_\mu^{A\nu} G_\nu^{B\rho} G_\rho^{C\mu} + h.c., \quad (4.4)$$

$$O_{u(c)G} = (\bar{q} \sigma^{\mu\nu} \lambda^A t) \tilde{\phi} G_{\mu\nu}^A, \quad L_{\text{eff}} = \frac{C_{u(c)G}}{\sqrt{2}\Lambda^2} v (\bar{u} (\bar{c}) \sigma^{\mu\nu} \lambda^A t) G_{\mu\nu}^A + h.c., \quad (4.5)$$

where $C_{\phi q}^{(3)}$, C_{tW} , C_{tG} , C_G and $C_{u(c)G}$ stand for the dimensionless Wilson coefficients, also called effective couplings. The variable Λ represents the energy scale beyond which new physics becomes relevant. The detailed description of the operators is given in Refs. [45, 46]. The operators $O_{\phi q}^{(3)}$ and O_{tW} modify the SM interaction between W boson, top quark, and b quark (Wtb). The operator O_{tG} is called chromomagnetic dipole moment operator of the top quark. The triple gluon field strength operator O_G represents the only genuinely gluonic CP conserving term which can appear at dimension 6 within an effective strong interaction Lagrangian [48]. The operators O_{uG} and O_{cG} lead to flavor changing neutral current (FCNC) interactions of top quark and contribute to the tW production. The effect of introducing the new couplings $C_{\phi q}^{(3)}$, C_{tW} , C_{tG} and $C_{u(c)G}$ can be investigated in the tW production. The chromomagnetic dipole moment operator of the top quark affects also the $t\bar{t}$ production. In the case of the C_G coupling, only the $t\bar{t}$ production is modified. It should be noted that the O_{tW} and O_{tG} operators with imaginary coefficient lead to CP-violating effects. Representative Feynman diagrams for new physics contributions in the tW and $t\bar{t}$ production are shown in Fig. 4.1. In this analysis we only probe CP-even dimension six operators via top quark production.

Several searches for new physics in the top quark sector including new non-SM couplings have been performed at the Tevatron and LHC colliders. Results can be interpreted in two ways. Most of the previous analyses followed the anomalous coupling approach in which SM interactions are extended for possible new interactions. In this study, the EFT framework with effective couplings is used for the interpretation of the results. Constraints obtained on anomalous couplings can be translated to effective coupling bounds [45, 49]. A variety of limits have been set on the Wtb anomalous coupling through single top quark t -channel production and measurements of the W boson polarisation from top quark decay by the D0 [49], ATLAS [50, 51] and CMS [52, 53] Collaborations. Direct limits on the top chromomagnetic dipole moment have been obtained by the CMS Collaboration at 7 TeV using top quark pair events [54]. Searches for top quark FCNC interactions have been performed at Tevatron [55, 56] and at LHC [52, 57] via single top quark production and limits are set on related anomalous couplings.

In this chapter, a search for new physics in top quark production using an EFT framework is reported. Final states with two same flavor opposite-sign isolated leptons (electrons or muons) in association with jets identified as originated from the fragmentation of a b quark (“b jets”) are analysed. The search is sensitive to new physics contributions to the tW and $t\bar{t}$ production, and the six effective couplings, C_G , $C_{\phi q}^{(3)}$, C_{tW} , C_{tG} , C_{uG} , and C_{cG} , are constrained independently. Kinematic distributions of final state particles and the production rate of the tW and $t\bar{t}$ processes are both affected by the effective couplings. For the $C_{\phi q}^{(3)}$, C_{tW} , C_{tG} , and C_G effective couplings, the deviation from the SM prediction

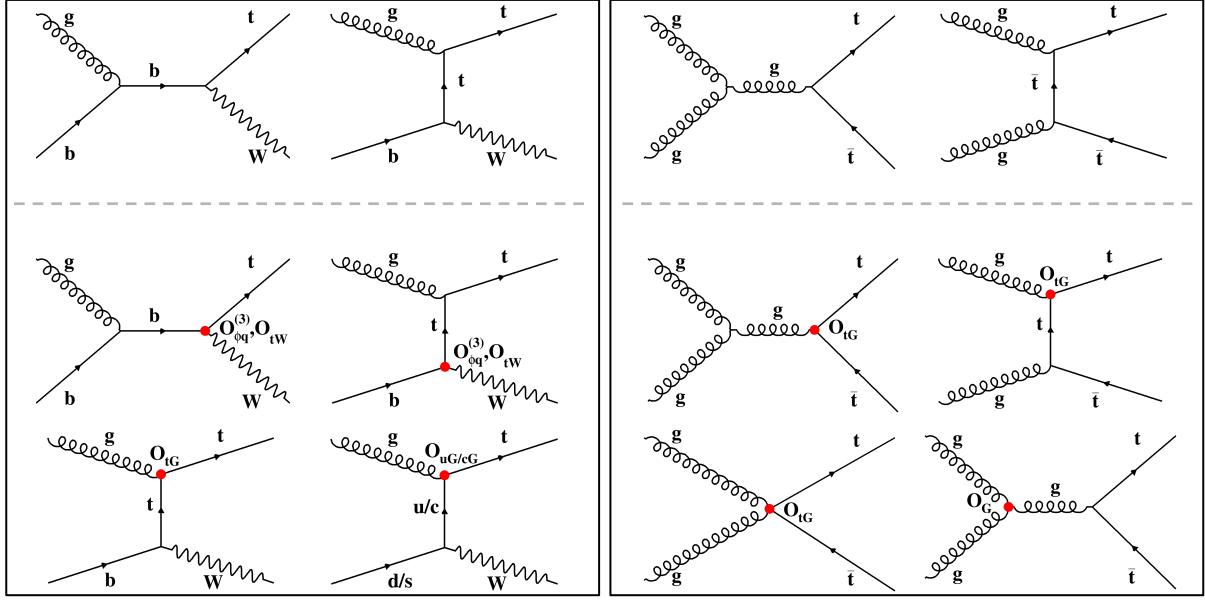


Figure 4.1: Representative Feynman diagrams for the tW (left panel) and $t\bar{t}$ (right panel) production at leading order. The upper row gives the SM diagrams, the middle and lower rows present diagrams corresponding to the $O_{\phi q}^{(3)}$, O_{tW} , O_{tG} , O_G and $O_{u/cG}$ contributions.

is dominated by the interference term between SM and new physics diagrams, which is linear with respect to the effective coupling. Therefore, kinematic distributions of final state particles vary as a function of Wilson coefficients and a small value of the effective couplings leads to distributions similar to the SM predictions. On the other hand, the new physics terms for the C_{uG} and C_{cG} effective couplings do not interfere with the SM tW process, and the kinematic distributions of final state particles are determined by the new physics term independently of the SM prediction. In this analysis, we use the rates of tW and $t\bar{t}$ production for probing the $C_{\phi q}^{(3)}$, C_{tW} , C_{tG} , and C_G effective couplings, while both variation in rate and kinematic distributions of final state particles are employed for probing the C_{uG} and C_{cG} effective couplings.

The chapter is structured as follows. In Section 4.2, a description of the data and simulated samples used in the analysis is detailed. The trigger used in this analysis is detailed in Section 4.3. The object selection and event selection are presented in Section 4.4 and 4.5, respectively. The SM background estimation and data mc distribution are presented in sections 4.6 and Section 4.7, respectively. Section 4.8 presents a description of the signal extraction procedure. An overview of the systematic uncertainty treatment is given in Section 4.9. A cross check for measuring SM tW process cross section is provided in appendix .5. Finally, the constraints on the effective couplings are presented in Section 4.10, and a summary is given in Section 4.11.

4.2 Data-sets and Monte Carlo Samples

4.2.1 Data samples

The primary data sets used in this analysis are summarised in Table 4.1. The data from the Moriond17 rereco campaign (Run2016 03Feb2017 Re-miniAOD) is used for eras 2016B through 2016G and the prompt reconstruction is used for era 2016H (Run2016H-03Feb2017_ver2 and Run2016H-03Feb2017_ver3). The integrated luminosity of the data sample used in this analysis is 35.9 fb^{-1} collected by the CMS experiment in 2016. Only certified data recommended for analysis by the PdmV group is used. The corresponding JSON file is the following:

Cert_271036-284044_13TeV_23Sep2016ReReco_Collisions16_JSON.txt

Datasets	run range	integrated luminosity (pb^{-1})
/X/Run2016B-03Feb2017_ver2-v2/MINIAOD	273158 – 275376	5788.348
/X/Run2016C-03Feb2017-v1/MINIAOD	275657 – 276283	2573.399
/X/Run2016D-03Feb2017-v1/MINIAOD	276315 – 276811	4248.384
/X/Run2016E-03Feb2017-v1/MINIAOD	276831 – 277420	4009.132
/X/Run2016F-03Feb2017-v1/MINIAOD	277981 – 278808	3101.618
/X/Run2016G-03Feb2017-v1/MINIAOD	278820 – 280385	7540.488
/X/Run2016H-03Feb2017_ver2-v1/MINIAOD	281613 – 284035	8390.540
/X/Run2016H-03Feb2017_ver3-v1/MINIAOD	284036 – 284044	215.149
Sum	273158 – 284044	35.867

Table 4.1: Data sets (X) used in this analysis. X = 'DoubleEG' and 'SingleElectron' for the ee channel. X = 'DoubleMuon' and 'SingleMuon' for $\mu\mu$ channel analysis.

4.2.2 MC samples

The analysis uses centrally produced Monte Carlo (MC) samples from 'RunIISummer16 MiniAODv2' campaign as are recommended for Moriond17. For the MC samples used, GEN-SIM was produced using CMSSW_7_1_X and the DIGI-RECO was produced using CMSSW_8_0_X. All the MC are listed in Table 4.2.

sample	xsection(pb)	xs precision
DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	18610	NLO
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	5765.4	NNLO
WWTo2L2Nu_13TeV-powheg	12.178	NNLO
WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8	4.42965	NLO
WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.595	NLO
ZZTo2L2Nu_13TeV_powheg_pythia8	0.564	NLO
ZZTo4L_13TeV_powheg_pythia8	1.212	NLO
WGToLNuG_TuneCUETP8M1_13TeV-amcatnloFXFX-pythia8	489	NLO
ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	19.47	app.NNLO
ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	19.47	app.NNLO
TTTo2L2Nu_TuneCUETP8M2_ttHtranche3_13TeV-powheg	87.31	NNLO
TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.4062	NLO
TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.2043	NLO
TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.2529	NLO
TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.5297	NLO
TTGJets_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	3.697	NLO
WJetsToLNu_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	61526.7	NNLO
GluGluHToWWTo2L2Nu_M125_13TeV_powheg_pythia8	2.5	NLO
VBFHToWWTo2L2Nu_M125_13TeV_powheg_pythia8	0.175	NLO

Table 4.2: MC samples (`dataset=/*/RunIISummer16MiniAODv2*PUMoriond17_80X_mcRun2_asymptotic_2016_TrancheIV_v6*/MINIAODSIM`)

The pile up distributions for MC and data which is calculated by using 69.2 mb as the minimum bias cross section are shown in left of Figure 3.1. MC events are re-weighted to account for the pile up difference between data and MC.

4.3 Triggers

In this analysis we use various sets of triggers in order to achieve an optimal selection efficiency. For each dilepton channel (ee , $\mu\mu$) we take the logical OR of the dilepton triggers listed in Table 4.3. The increase of the instantaneous luminosity delivered by the LHC in Run H needed the (re)introduction of the DZ cut for pure dimuon trigger.

We complement the partially inefficient dilepton triggers by single lepton triggers and remove the overlap between two primary data sets by vetoing events that fired a dilepton trigger in the single lepton primary data sets. Table 4.3 gives an overview of all triggers used in this analysis. Using $t\bar{t}$ and tW MC samples, it is found that by adding single lepton triggers, trigger efficiency is increased by around 5%.

channel	path	dataset
ee	HLT_Ele23_Ele12_caloIdL_TrackIdL_IsoVL_DZ	data & MC
	HLT_Ele27_WPTight_Gsf	data & MC
$\mu\mu$	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL	data runs B-G & MC
	HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL	data runs B-G & MC
	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ	data only run H
	HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ	data only run H
	HLT_IsoMu24	data & MC
	HLT_IsoTkMu24	data & MC

Table 4.3: Summary of the signal triggers

Since the trigger information is available in simulated samples, we require FullSim MC samples to fire the trigger and finally apply corrections for any data/MC disagreement using scale factors. TOP group recommended trigger scale factor (described in [58]) are used. We also have found trigger scale factors as a cross checked which gives similar result (although they are not used in this analysis).

4.4 Event reconstruction and object identification

The tW process in dilepton final states is characterised by the presence of two high P_T leptons associated with missing transverse energy (E_T^{miss}) and one b-jet. The reconstruction of the different objects is based on the Particle-Flow (PF) algorithm. The selection includes the following requirements:

4.4.1 Lepton selection

4.4.1.1 Muon

The muons used in this analysis are selected inside the fiducial region of the muon spectrometer, $|\eta| < 2.4$, with a minimum P_T of 20 GeV, and using standard identification criteria, suggested by the Muon POG. Furthermore, they are required to be particle-flow muons. Cuts are applied on the quality of the track fit, number of hits in the pixel, tracker and number of matched muon segments for the muons to be considered for the dilepton candidate. These requirements are summarised in the following and correspond to the so-called Tight muon identification [59].

- $P_T > 20$ GeV and $|\eta| < 2.4$,
- is a GlobalMuon and PFMuon,
- number of matched Stations > 1 ,
- number of pixel hits > 0 ,
- number of hits in the inner tracker > 5 ,
- number of muon hits > 0 ,
- $\frac{\chi^2}{NDF}$ of the global-muon track fit < 10 ,
- Impact parameter constraints between the muon track and the selected primary vertex $dZ < 0.5$ and $d0 < 0.2$ cm.

For the muon isolation, a cone of $\Delta R = 0.4$ is built to compute the flux of particle flow candidates, the delta-beta correction is applied to correct for pileup contamination. This correction is achieved by subtracting half the sum of the P_T of the charged particles in the cone of interest but with particles not originating from the primary vertex. The muon isolation is therefore defined as:

$$I_{\text{rel}}^\mu = \frac{1}{P_T^\mu} \left(\Sigma P_T(\text{ch-had from PV}) + \max(0, \Sigma E_T(\text{neut-had}) + \Sigma E_T(\text{photon}) - 0.5 \Sigma P_T(\text{ch-had from PU})) \right) \quad (4.6)$$

The factor 0.5 corresponds to a approximate average of neutral to charged particles and has been measured in jets in [60]. In our analysis the muon candidates must have $I_{\text{rel}}^\mu < 0.15$ to be considered as isolated [59]. Scale factors are used to correct for differences in the reconstruction, ID and Isolation efficiencies in data and Monte Carlo. They are evaluated using the tag and probe technique, and both the scale factors and their uncertainty prescriptions are provided by the Muon POG [61]. In addition, muon energy scale and smearing is applied based on Rochester group recommendations [62].

4.4.1.2 Electron

Electron candidates are selected from the reconstructed GSF electrons with $P_T > 20$ GeV and $|\eta| < 2.4$ while gap is removed ($1.4442 < |\eta_{\text{SuperCluster}}| < 1.566$). Electrons need to pass the tight cut based POG recommended working point [63] which includes the requirement to pass the conversion veto. Additionally, we apply cuts on the longitudinal and transverse impact parameter with respect to the primary vertex which are removed

from the official working point variable list. It is shown by TOP PAG group that electron ID scale factors are not changed by adding these two extra variables with respect to official tight working point [63]. The following Table contains the cuts used for the tight electron identification, derived using the Spring16 samples for 80X-based data and simulation. The details can be found in [63].

	Barrel($ \eta_{SuperCluster} $) ≤ 1.479	Endcap($ \eta_{SuperCluster} $) > 1.479
full5 \times 5 $\sigma_{I_n I_n} <$	0.00998	0.0292
$ \Delta\eta_{Inseed} <$	0.00308	0.00605
$ \Delta\Phi_{In} <$	0.0816	0.0394
H/E <	0.0414	0.0641
relative electron isolation <	0.0588	0.0571
$ (1/E - 1/P) <$	0.0129	0.0129
expected missing inner hits \leq	1	1
pass conversion veto	yes	yes
$d_0 <$	0.05	0.1
$dZ <$	0.1	0.2

Table 4.4: The cuts for the electron identification in the barrel and endcap.

An MVA regression technique is used to find the corrections for the super cluster energy to account for the effects like energy leakage into the gaps between crystals, energy leakage into the HCAL downstream the ECAL, etc [64]. The regression which is used in RunIISummer16 MiniAODv2 MC samples is trained using CMSSW74X. EGamma POG has provided 80 regression which should be applied offline to find the correct electron energy. 80 electron regression is applied as is explained in [65]. In addition, electron energy scale and smearing is applied based on Egamma POG recommendations [66].

4.4.2 Jet selection

Particle candidates found by the PF algorithm are clustered into jets using the Anti-kT algorithm with distance parameter $R = 0.4$ (AK4). The influence of pileup is mitigated by the Charged Hadron Subtraction (CHS) technique which removes tracks identified as originating from pileup vertices. Jets are calibrated in simulation and in data separately, accounting for deposits from pile-up and the imperfect detector response. L1Fastjet corrections, Level 2 and Level 3 jet energy corrections and L2L3Residual corrections (only for data) are applied using the latest set of JECs (Summer16_23Sep2016V4). Jets in MC are smeared using the latest set of JERs (Spring16_25nsV10). Corrected jets with $P_T > 30$ GeV and $|\eta| < 2.4$ are selected if they pass the loose jet identification criteria [67], i.e. the neutral electromagnetic and hadron fractions are $< 99\%$ and the jet consists of at least two PF candidates. Furthermore, both the charged hadron fraction and multiplicity are required to be > 0 and the charged electromagnetic fraction has to be $< 99\%$.

Selected jet may still overlap with the selected leptons. This is possible because the lepton can be clustered into a jet as well. To prevent such cases, jets that are found within a cone of $R = 0.4$ around any of the selected signal leptons are removed from the set of selected jets.

Jets originating from the hadronization of b-quarks are identified using the Combined Secondary Vertex algorithm (CSVv2). The CSVv2 algorithm combines information from track impact parameters and secondary vertices identified within a given jet. In this analysis, a jet is b-tagged when its CSVv2 value passes the medium working point (i.e. $CSVv2 > 0.8484$) [68].

4.4.3 Missing Transverse Energy

Missing transverse energy is calculated as the negative of the vectorial sum of the transverse momentum vectors of all PF candidates in an event. To make MET a better estimate of true invisible particles, so called Type-1 plus smeared corrections are applied, which propagate the jet energy corrections to the raw MET [69].

In order to reduce the instrumental noise in the detector, MET filters are applied as is recommended by Jet-MET POG [70]. These filters are summarised below:

- HBHENoiseFilter (data and MC),
- HBHENoiseIsoFilter (data and MC),
- globalTightHalo2016Filter (data and MC),
- goodVertices (data and MC),
- EcalDeadCellTriggerPrimitiveFilter (data and MC),
- BadChargedCandidateFilter (data and MC),
- BadPFMuonFilter (data and MC),
- eeBadScFilter (only data)

4.4.4 Scale factors

In the following, object related scale factors used in this analysis are listed:

Muon: https://twiki.cern.ch/twiki/bin/view/CMS/MuonWorkInProgressAndPagResults#Results_on_the_full_2016_data

- Tracking efficiency https://test-calderona.web.cern.ch/test-calderona/MuonPOG/2016dataReRecoEfficiencies/tracking/Tracking_EfficienciesAndSF_BCDFFGH.root
- Identification efficiency https://gaperrin.web.cern.ch/gaperrin/tnp/TnP2016/2016Data_Moriond2017_6_12_16/JSON/RunBCDEF/EfficienciesAndSF_BCDEF.root, https://gaperrin.web.cern.ch/gaperrin/tnp/TnP2016/2016DataMoriond2017_6_12_16/JSON/RunGH/EfficienciesAndSF_GH.root
- Isolation efficiency https://test-calderona.web.cern.ch/test-calderona/MuonPOG/2016dataReRecoEfficiencies/isolation/EfficienciesAndSF_BCDEF.root, https://test-calderona.web.cern.ch/test-calderona/MuonPOG/2016dataReRecoEfficiencies/isolation/EfficienciesAndSF_GH.root

Electron: https://twiki.cern.ch/twiki/bin/view/CMS/EgammaIDRecipesRun2#Efficiencies_and_scale_factors

- Reconstruction efficiency http://fcouderc.web.cern.ch/fcouderc/EGamma/scaleFactors/Moriond17/approval/RECO/passingRECO/egammaEffi.txt_EGM2D.root
- Identification + isolation efficiency http://fcouderc.web.cern.ch/fcouderc/EGamma/scaleFactors/Moriond17/approval/EleID/passingTight80X/egammaEffi.txt_EGM2D.root

B-tagging: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation80XReReco>

- b-tagging efficiency

4.4.5 Top P_T reweighting

In order to better describe the transverse momenta (P_T) distribution of the top quark in $t\bar{t}$ events, the top quark transverse momentum spectrum simulated with POWHEG is reweighted to match the differential top quark P_T distribution at NNLO QCD accuracy and including EW corrections calculated in Ref. [71].

4.5 Event selection and cut optimization

The selection of the events to be used for our search is done in two steps which are explained in the following sections.

4.5.1 Event selection (step 1)

At the first step, we focus on selecting events by requiring trigger and lepton selection. Events should fire one of the triggers summarized in Table 4.3. At least 1 pair of opposite charge leptons with invariant mass > 20 GeV is required while leading lepton should have $P_T > 25$ GeV. The first two selected leptons which are sorted due to the P_T should have same flavor opposite sign. If the two highest P_T leptons are different flavor (e.g. $e\mu$ event) or same sign, event is rejected. The events are divided in the ee and $\mu\mu$ channels according to the flavours of the two leptons with the highest P_T and are further categorised in different bins depending on the number of jets (“n-jets”) in the final state and the number of them which are b-tagged (“m-tags”). The largest number of tW events is expected in the category with exactly one b-tagged jet (1-jet,1-tag) followed by the category with two jets, one of which being a b jet (2-jets,1-tag). Events in the categories with more than two jets and exactly two b-tagged jets are dominated by $t\bar{t}$ process (≥ 2 -jets,2-tags).

4.5.2 Event selection (step 2)

In order to suppress the contribution of Drell-Yan events, we reject events with dilepton invariant mass around Z peak [76, 106]. In addition, a MET cut > 60 GeV is applied. In Figure 4.2, one can see that the Data/MC agreement for MET distribution in ee and $\mu\mu$ channels are not good. We have investigated this problem (see Appendix .4). Because of that the normalization of the DY background simulation is estimated from data which is described in Section 4.6.1. The final event selection criteria is shown in Table 4.5.

		ee and $\mu\mu$
Step 1	$p_T^{\text{leading lepton}} > 25$ GeV and $p_T^{\text{sub-leading lepton}} > 20$ GeV	✓
	Mass(ll) > 20 GeV	✓
	The two highest p_T leptons should have same flavor and opposite charge	✓
Step 2	MET > 60 GeV	✓
	Mass(ll) < 76 or > 106 GeV	✓

Table 4.5: Event Selections

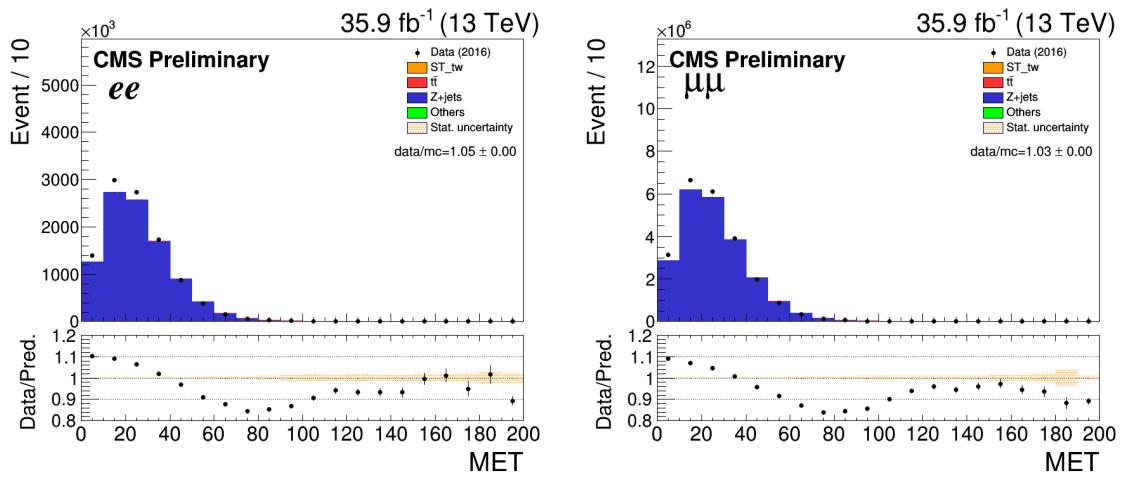


Figure 4.2: The distributions of MET for ee (left) and $\mu\mu$ (right) channels after step 1 selection. All backgrounds are estimated from MC.

4.6 Background predictions

4.6.1 Prompt Background

The background from processes giving two prompt leptons is taken from Monte Carlo samples and normalized to the luminosity. It consists mostly of events from $t\bar{t}$ production, Drell-Yan and WW production. Other background processes considered are the other diboson processes like WZ and ZZ .

In the $\mu\mu$ and ee final states, the normalization of the DY background simulation is estimated from data using the method described in [72, 73, 74, 75], extracting the events outside the Z-veto region from the events inside. As described above, events in this region have a dilepton invariant mass between 76 GeV and 106 GeV and are rejected for the analysis. Since contamination from non-DY background contributions can still be present in the Z-veto region, this contribution is subtracted from the $e\mu$ channel and then scaled according to the event yields in the ee and $\mu\mu$ channels.

The expected number of events outside the Z-veto can be measured from data as:

$$N_{out}^{l^+l^-, Z+jets \text{ data}} = R_{out/in}^{l^+l^-}(N_{in}^{l^+l^-, \text{data}} - 0.5N_{in}^{e\mu, \text{data}}k_{ll})$$

where $ll = \mu\mu$ or ee and $R_{out/in}$ is the ratio of the number of events outside/inside the Z-veto region taken from the DY simulated sample:

$$R_{out/in} = \frac{N_{out}^{l^+l^-, Z+jets \text{ MC}}}{N_{in}^{l^+l^-, Z+jets \text{ MC}}}.$$

Here, k_{ll} is a correction factor that takes into account the differences between electron and muon reconstruction. This correction can be determined from the number of ee and $\mu\mu$ events in the Z peak region after applying the MET requirement (labeled as *loose*). Since $N^{e^+e^-_{in,loose}}$ and $N^{\mu^+\mu^-_{in,loose}}$ are proportional to the square of the corresponding single-lepton candidate selection efficiencies, the correction factor can be expressed as:

$$k_{ee} = \sqrt{\frac{N^{e^+e^-_{in,loose}}}{N^{\mu^+\mu^-_{in,loose}}}}$$

$$k_{\mu\mu} = \sqrt{\frac{N^{\mu^+\mu^-_{in,loose}}}{N^{e^+e^-_{in,loose}}}}$$

The value of k_{ll} for different njet-nbjet regions are shown in Table 4.6. We use explicitly k_{ll} for different njet-nbjet regions.

Table 4.6: The value of k_{ll} for different njet-nbjet regions. Errors are statistical uncertainties only.

Channel	all	1jet,1tag	2jet,1tag	$>=2\text{jet},2\text{tag}$
$N^{e^+e^-_{in,loose}}$ (data)	220435	3805	3735	2493
$N^{\mu^+\mu^-_{in,loose}}$ (data)	501781	8291	7550	5180
k_{ee}	0.66 ± 0.001	0.68 ± 0.007	0.70 ± 0.007	0.69 ± 0.008
$k_{\mu\mu}$	1.51 ± 0.002	1.48 ± 0.014	1.42 ± 0.014	1.44 ± 0.018

The global scaling factors

$$C_{Z+jets} = \frac{N_{out}^{l^+l^-, Z+jets \text{ data}}}{N_{out}^{l^+l^-, Z+jets \text{ MC}}}$$

are determined. The results and scaling factors are summarized in Table 4.7.

	all		1jet,1tag		2jet,1tag		>= 2jet, 2tag	
	ee	$\mu\mu$	ee	$\mu\mu$	ee	$\mu\mu$	ee	$\mu\mu$
$N_{in}^{l^+l^-, Z+jets \text{ MC}}$	243878.3	562506.9	2712.3	6490.7	1550.6	3734.2	306.8	712.7
$N_{out}^{l^+l^-, Z+jets \text{ MC}}$	22376	56494.6	301.4	878.7	280.8	590.9	53.7	94.1
$R_{out/in}$	0.092	0.100	0.111	0.135	0.181	0.158	0.175	0.132
$N_{in}^{l^+l^-, data}$	220435	501781	3805	8291	3735	7550	2493	5180
$N_{in}^{e\mu, data}$	34322	34322	4453	4453	6230	6230	6259	6259
$N_{out}^{l^+l^-, Z+jets \text{ data}}$	19185.9	47793.1	254.6	676.3	281.6	494.8	58.4	89.0
C_{Z+jets}	0.857 ± 0.004	0.846 ± 0.003	0.845 ± 0.049	0.770 ± 0.030	1.002 ± 0.085	0.837 ± 0.048	1.087 ± 0.28	0.945 ± 0.181

Table 4.7: Data-driven $Z+jets$ background estimation in the $\mu\mu$ and ee channels after the “Step1+Step2” selection requirement. Errors are statistical uncertainties only.

4.6.2 Fake Background

Another source of events with misidentified leptons is the $W\gamma$ process, where the W decays to a muon or electron and the photon is either misidentified as electron, or the photon converts and gives an electron. This background contribution is taken from MC simulation. For the backgrounds which involve a jet that are misidentified as an electron or muon, a data-driven technique is used.

The jet background (or called Nonprompt background) consists of events where a jet is reconstructed as an electron or muon that passes the selection mostly from $W+jets$ process and QCD. The method which is used to estimate jets is called same sign method. In the same sign method, we use the fact that the probability of assigning positive or negative charge to the misidentified jet should be equal. Therefore, opposite and same-sign pairs are similar for fake jets in total number and distribution shape for many variables. On the other hand, all other standard model processes have opposite sign electron pairs and do not contribute to same sign control region. The contributions of the prompt backgrounds are subtracted from data in same sign region using MC samples to find the jets contribution in opposite sign region (signal region).

4.7 Data/MC comparison

After all selections and background estimation, the expected numbers of events from tW , $t\bar{t}$, DY and remaining background contributions mentioned above, as well as the total number of background events are reported in table 4.8 for the ee and $\mu\mu$ channels and for the various (n-jets,m-tags) categories. The data and MC comparison are shown in figures 4.3-4.8. In Figure 4.9, the data in the six (three ee plus three $\mu\mu$) search regions are shown together with the predictions for the SM backgrounds. The sources of systematic uncertainties which are considered in the Table 4.8 and the plots of figures 4.3-4.9 are explained in Section 4.9.

Table 4.8: Numbers of expected events from tW , $t\bar{t}$ and DY production, from the remaining backgrounds (other and nonprompt backgrounds), total background contribution and observed events in data after all selections for the ee and $\mu\mu$ channels and for various (n-jets,m-tags) categories. The uncertainties correspond to the statistical contribution only for the individual background predictions and to the quadratic sum of the statistical and systematic contributions for the total background predictions.

Channel	(n-jets,m-tags)	Prediction					Data
		tW	$t\bar{t}$	DY	Other + nonprompt	Total predicted yield	
ee	(1,1)	884 \pm 8	4741 \pm 15	258 \pm 50	53 \pm 5	5936 \pm 470	5902 \pm 76
	(2,1)	518 \pm 6	7479 \pm 19	241 \pm 53	94 \pm 5	8331 \pm 597	8266 \pm 90
	(\geq 2,2)	267 \pm 4	7561 \pm 18	46 \pm 24	99 \pm 4	7973 \pm 819	7945 \pm 89
$\mu\mu$	(1,1)	1738 \pm 12	9700 \pm 21	744 \pm 90	183 \pm 5	12366 \pm 879	12178 \pm 110
	(2,1)	989 \pm 9	14987 \pm 27	501 \pm 75	275 \pm 5	16751 \pm 1276	16395 \pm 128
	(\geq 2,2)	508 \pm 6	15136 \pm 26	82 \pm 24	163 \pm 5	15889 \pm 1714	15838 \pm 125

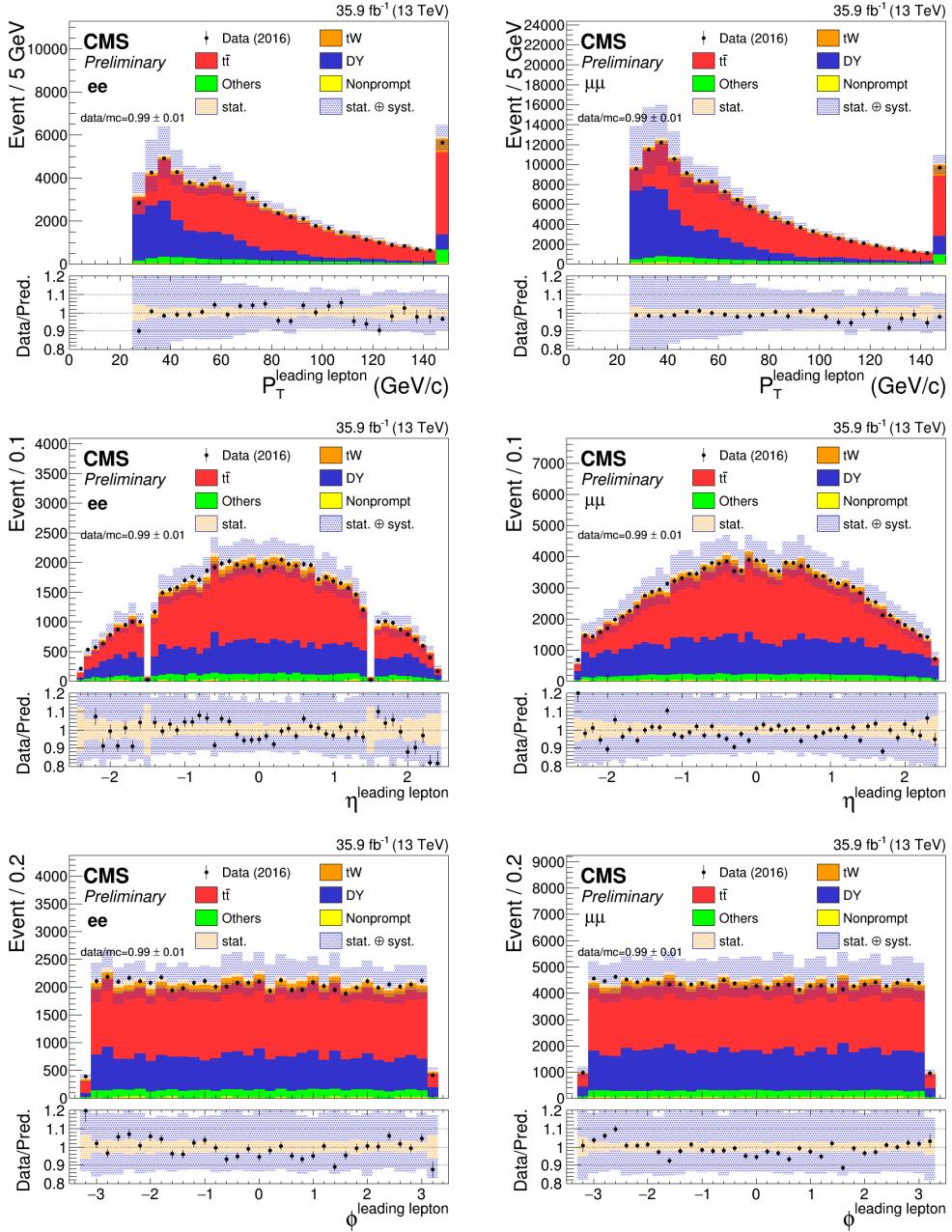


Figure 4.3: The distributions of P_T (top), η (middle) and ϕ (bottom) of leading lepton for ee (left) and $\mu\mu$ (right) channels after step 2 (full selection).

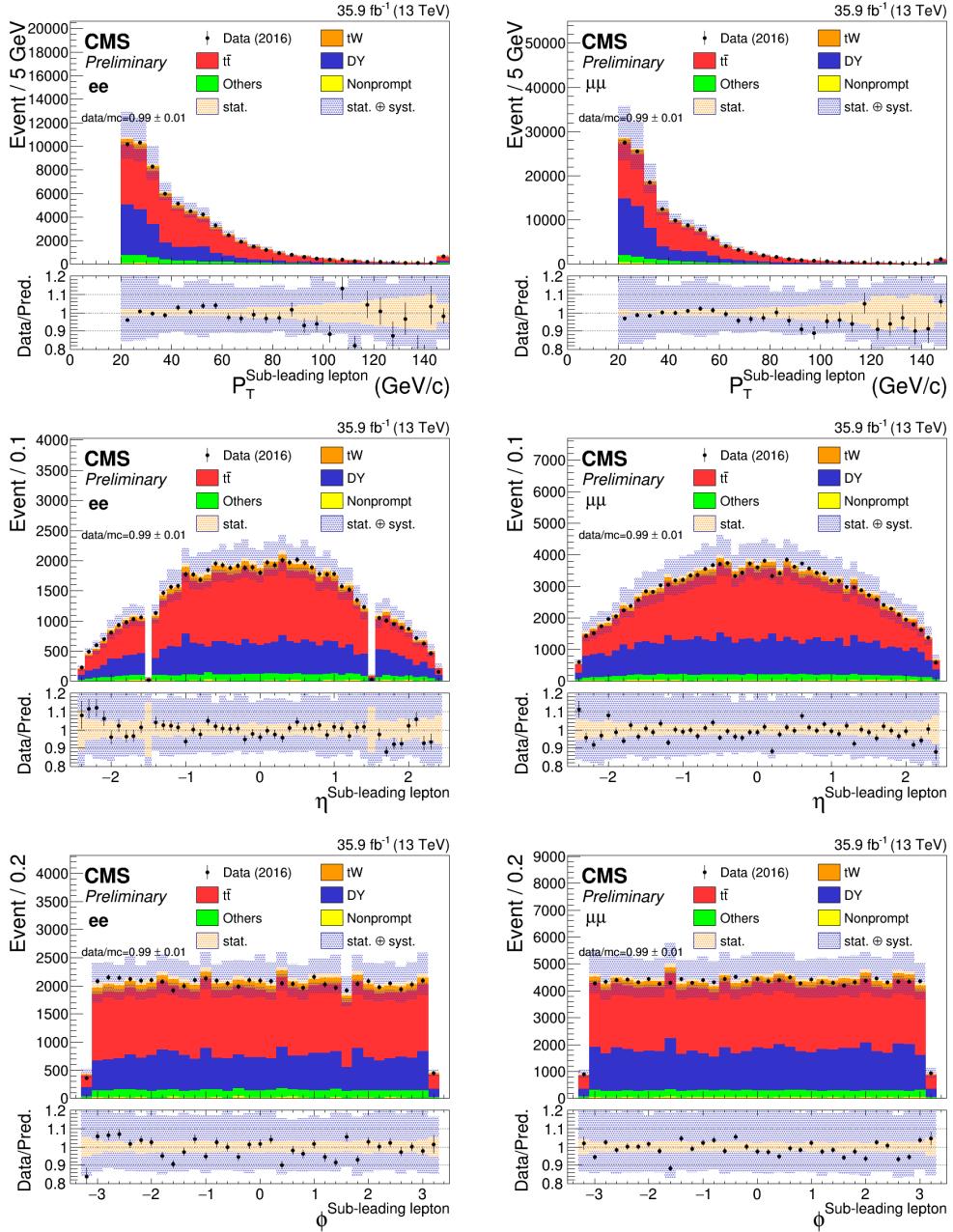


Figure 4.4: The distributions of P_T (top), η (middle) and ϕ (bottom) of sub-leading lepton for ee (left) and $\mu\mu$ (right) channels after step 2 (full selection).

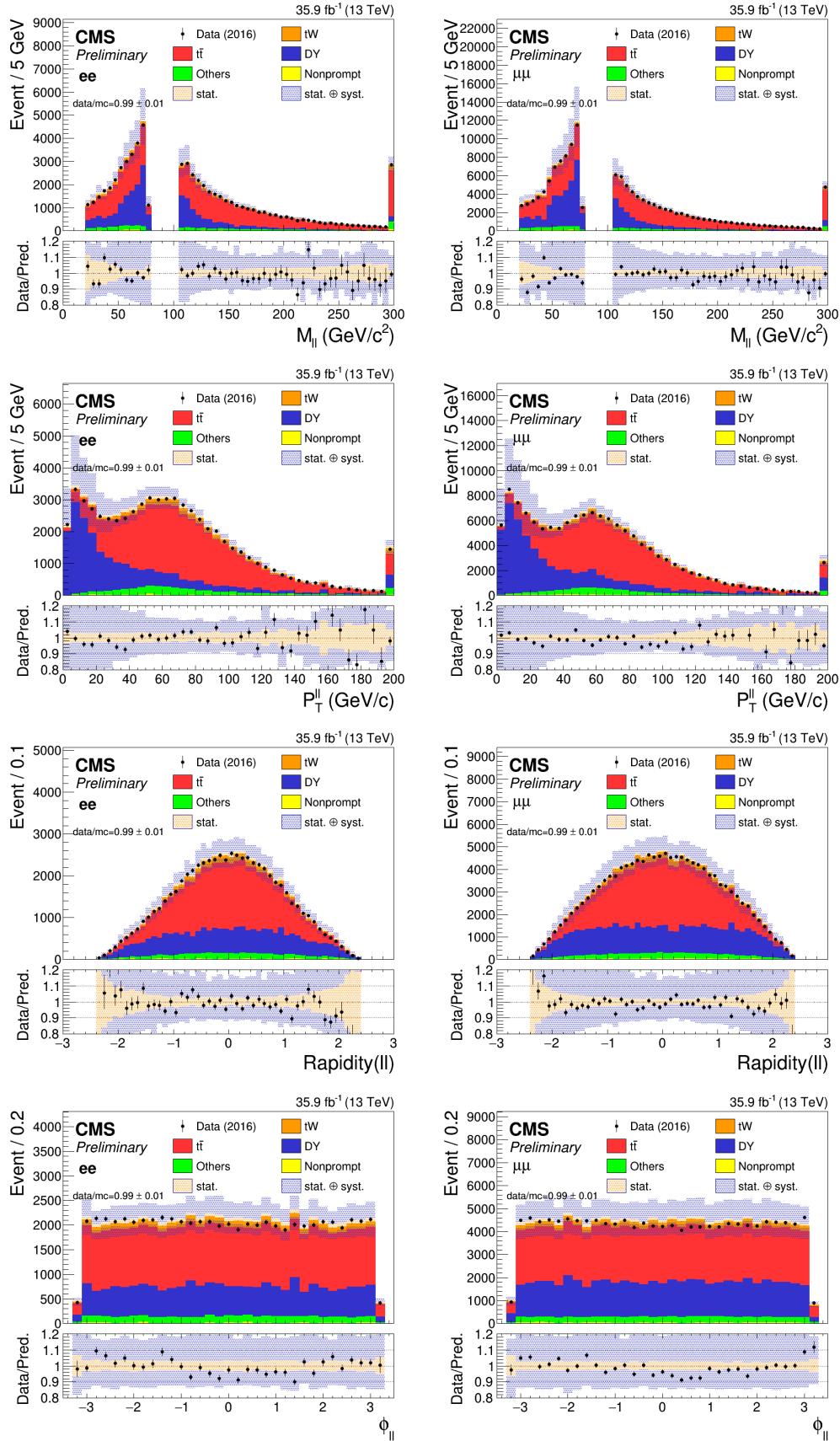


Figure 4.5: The distributions of invariant mass of two leptons (first row), P_T of two leptons (second row), Rapidity of two leptons (third row) and ϕ of two leptons (last row) for ee (left) and $\mu\mu$ (right) channels after step 2 (full selection).

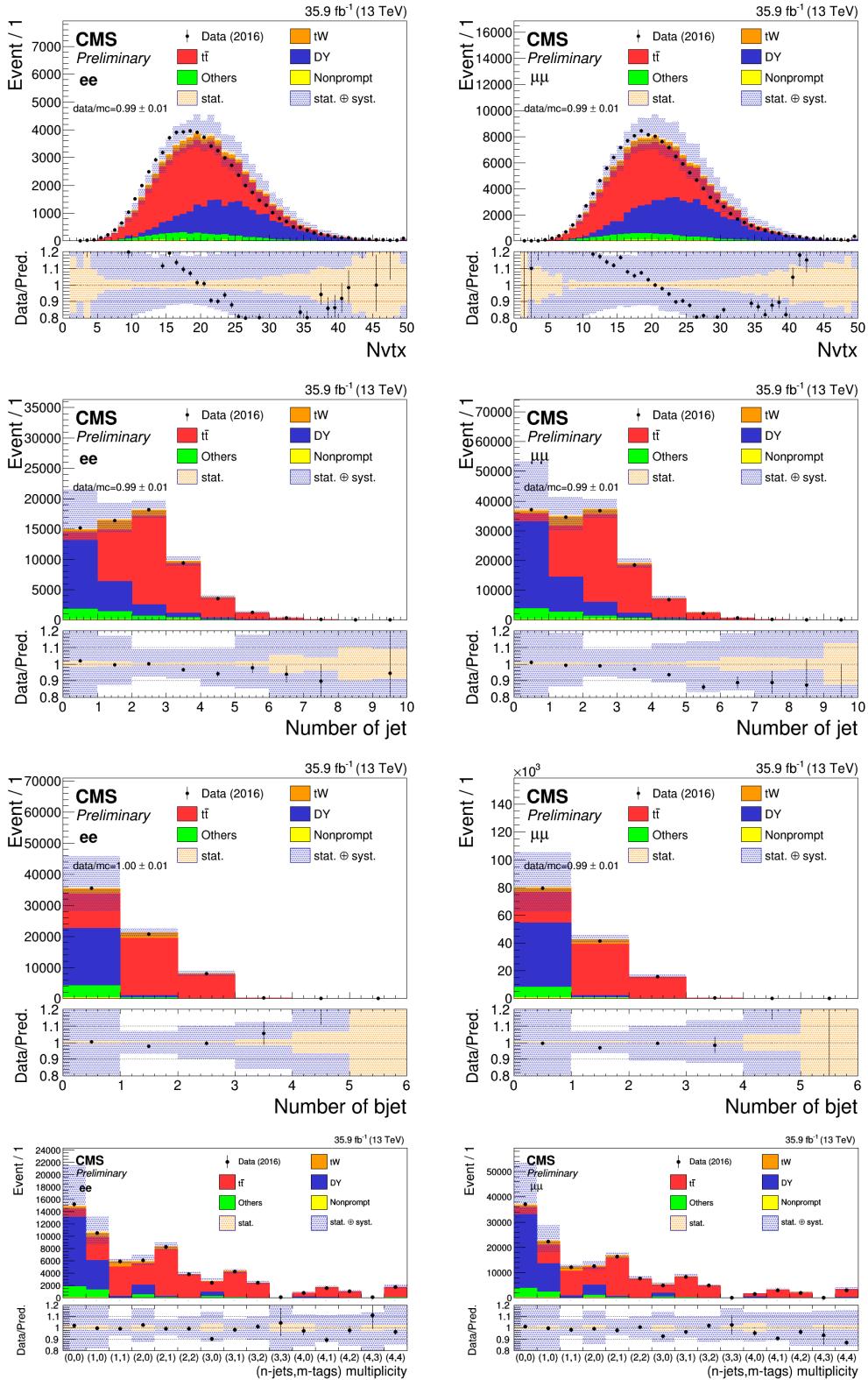


Figure 4.6: The distributions of number of vertices (first row), number of jet (second row), number of b jet (third row) and number of jet-bjet (last row) for ee (left) and $\mu\mu$ (right) channels after step 2 (full selection).

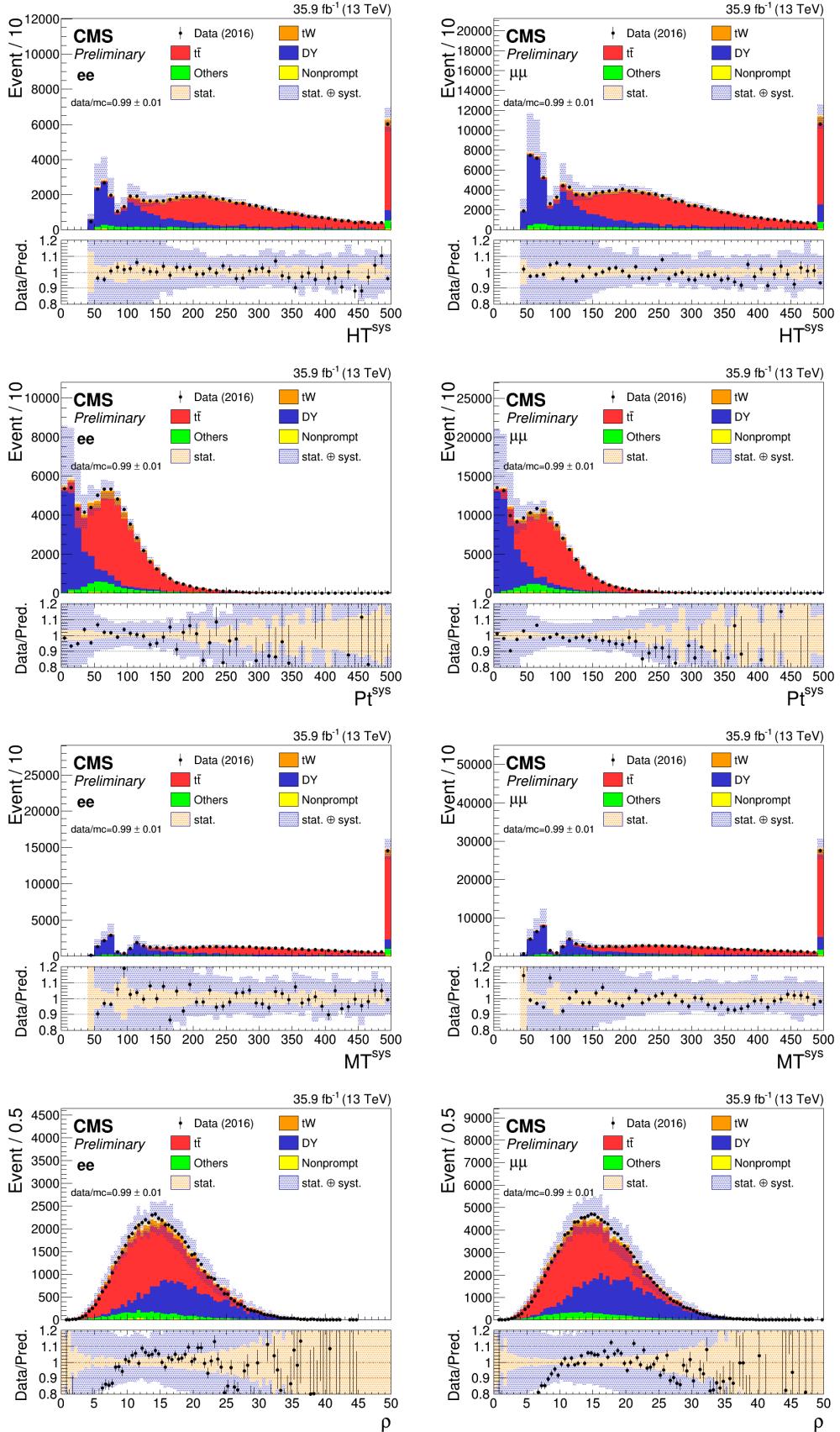


Figure 4.7: The distributions of HT^{system} (first row), Pt^{system} (second row), Mt^{system} (third row) and ρ for ee (left) and $\mu\mu$ (right) channels after step 2 (full selection).

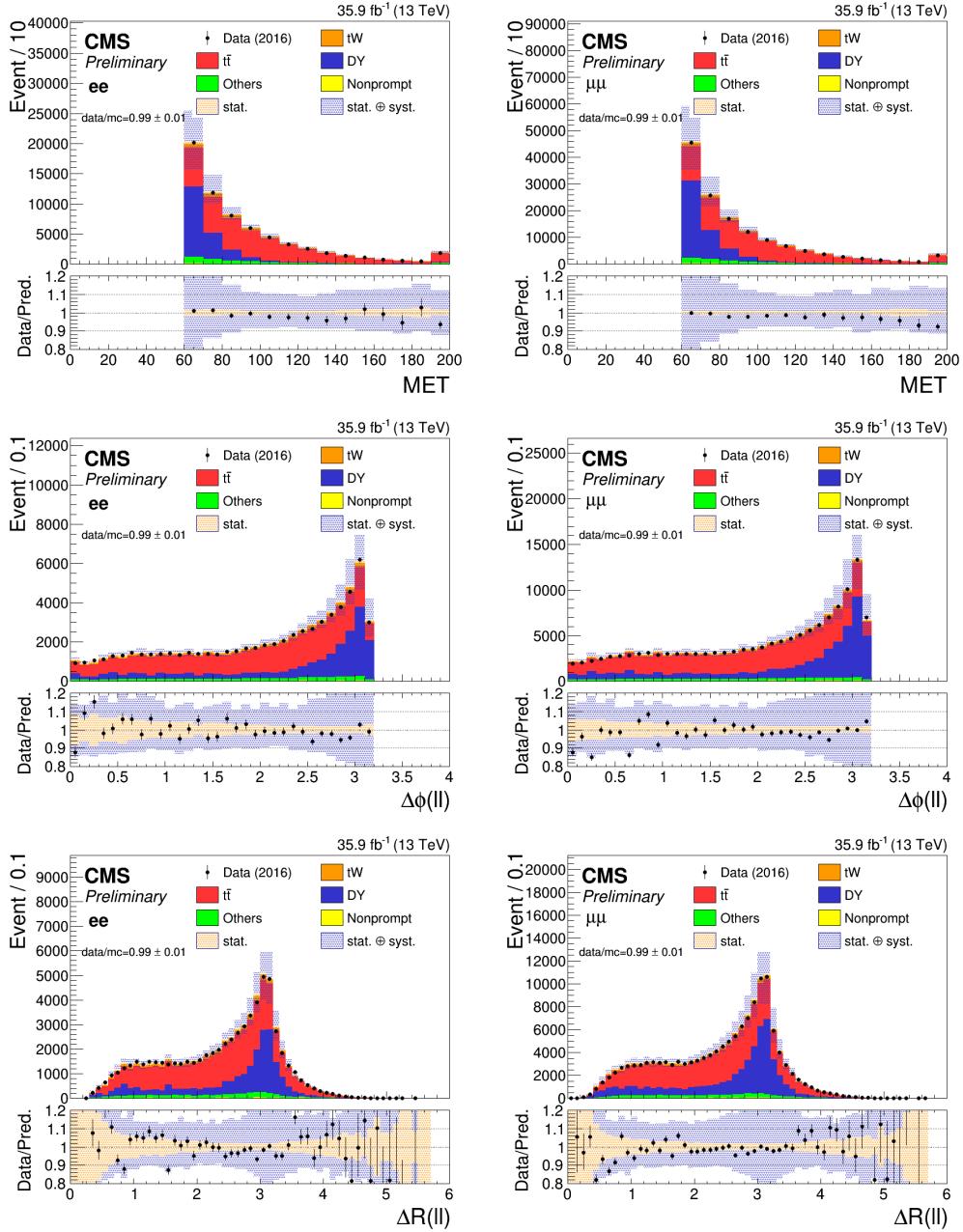


Figure 4.8: The distributions of MET (top row), $\Delta\phi$ between two leptons (middle row) and ΔR between two leptons (bottom row) for ee (left) and $\mu\mu$ (right) channels after step 2 (full selection).

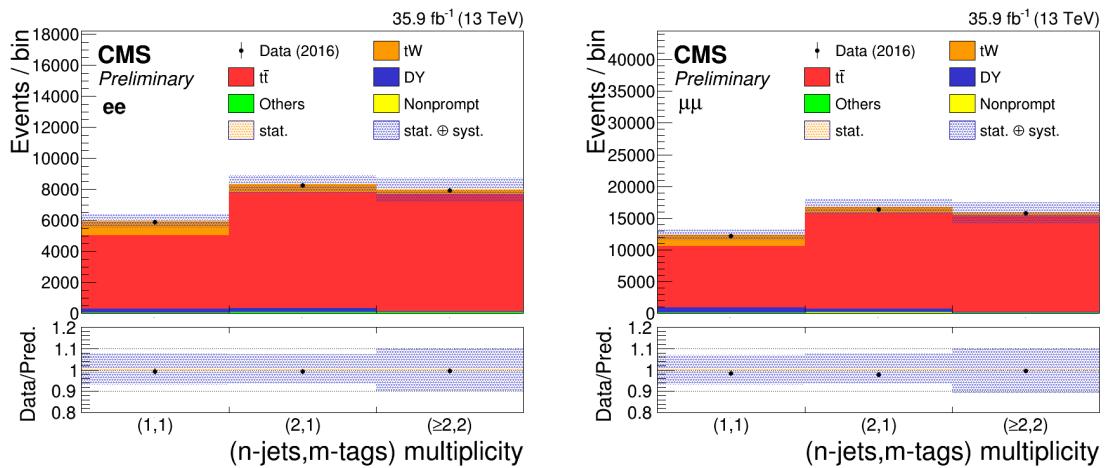


Figure 4.9: The observed numbers of events and SM background predictions in the search regions of the analysis for the ee (left), $\mu\mu$ (right) channels. The hatched band corresponds to the quadratic sum of statistical and systematic uncertainties in the event yield for the SM background predictions. The ratios of data to the sum of the predicted yields are shown at the bottom of each plot. The narrow hatched band represents the contribution from the statistical uncertainty in the MC simulation.

4.8 Signal extraction using neural networks tools

The purpose of the analysis is to search for deviations from the SM predictions in the tW and $t\bar{t}$ production due to new physics, parameterized with the presence of new effective couplings. In order to investigate the effect of the introduction of the new effective couplings, it is important to find suitable variables with high discrimination power between the signal and the background. Depending on the couplings, the total rate (yield) or the distribution of the output of a neural network algorithm (NN) is employed. The NN algorithm used in this analysis is a multilayer perceptron [76].

All the effective couplings introduced in Section 4.1 can contribute to tW production except the O_G operator. The introduction of the O_G operator affects only the $t\bar{t}$ production. It was shown in Ref. [48] and checked in this analysis that the top quark transverse momentum distribution is sensitive to the triple gluon field strength operator. The kinematic distributions of final-state particles show less discrimination power than the top quark transverse momentum distribution. In addition, they vary as a function of C_G and tend to the SM prediction for decreasing values of the C_G coupling. Therefore, we use the total number of events (yield) in various (n -jets, m -tags) categories to constrain the C_G effective coupling.

The deviation from the SM tW production from the interference terms between the SM and the O_{tG} , $O_{\phi q}^{(3)}$ and O_{tW} operators is of the order of $1/\Lambda^2$. It is assumed that the new physics scale Λ is larger than the scale we probe. Therefore, $1/\Lambda^4$ contributions from the new physics terms are small compared to the contribution from the interference term. The operator $O_{\phi q}^{(3)}$ is similar to the SM Wtb operator and leads to a rescaling of the SM Wtb vertex. The O_{tW} and O_{tG} operators provide new interactions compared to the SM Wtb vertex and the top-top-gluon (ttG) vertex. However, their effects have been investigated through the various kinematic distributions of the final-state particles and are found to be not distinguishable from the SM tW and $t\bar{t}$ processes for unconstrained values of the effective couplings. After the selection described in Section 4.5, the dominant background comes from the $t\bar{t}$ production, with a contribution of about 90%. In order to observe deviations from SM tW production in the presence of the $O_{\phi q}^{(3)}$, O_{tW} and O_{tG} effective operators, we need to separate tW events from the large number of $t\bar{t}$ events. Two independent NN are trained to separate $t\bar{t}$ events (the background) and tW events (considered as the signal) in the (1-jet,1-tag) (NN_{11}) and (2-jets,1-tag) (NN_{21}) categories which have significant signal contributions [77].

The presence of the O_{uG} and O_{cG} operators changes the initial-state particle (see Figure 4.1), and leads to different kinematic distributions for the final-state particles, compared to the SM tW process. For these FCNC operators, new physics effects on final-state particle distributions are expected to be distinguishable from SM processes. In order to search for new physics due to the O_{uG} and O_{cG} effective operators, a NN (NN_{FCNC}) is used to separate the SM backgrounds ($t\bar{t}$ and tW events together) and the new physics signals for events with exactly one b-tag jet with no requirement on the number of light jets (n -jets,1-tag).

The observables used in the analysis for probing new physics are summarised in Table 4.9. The various input variables for the NN introduced above are shown in Table 4.10.

The MLP input variables distributions of signal and background events are shown from Figure 4.10 to 4.12. The MLP output for test and train samples are shown in Figure 4.13.

Table 4.9: Summary of the observables used for probing effective couplings in various (n-jets,m-tags) categories in the ee and $\mu\mu$ channels.

Eff. coupling	Channel	Categories			
		1-jet,1-tag	2-jets,1-tag	n-jets,1-tag	≥ 2 -jets,2-tags
C_G	ee	Yield	Yield	-	Yield
	$\mu\mu$	Yield	Yield	-	Yield
$C_{\phi q}^{(3)}, C_{tW}, C_{tG}$	ee	NN ₁₁	NN ₂₁	-	Yield
	$\mu\mu$	NN ₁₁	NN ₂₁	-	Yield
C_{uG}, C_{cG}	ee	-	-	NN _{FCNC}	-
	$\mu\mu$	-	-	NN _{FCNC}	-

Table 4.10: Input variables for the NN used in the analysis in various bins of n-jets and m-tags. The symbols “√” indicate the variable is used.

Variable	Description	NN ₁₁	NN ₂₁	NN _{FCNC}
M_{ll}	Invariant mass of dilepton system			√
$p_T^{\ell\ell}$	p_T of dilepton system		√	√
$\Delta p_T(\ell, \ell)$	$p_T^{\text{leading lepton}} - p_T^{\text{sub-leading lepton}}$			√
$p_T^{\text{leading lepton}}$	p_T of leading lepton		√	√
Centrality($\ell^{\text{leading}}, \text{jet}^{\text{leading}}$)	Scalar sum of p_T of the leading lepton and leading jet, over total energy of selected objects			√
Centrality(ℓ, ℓ)	Scalar sum of p_T of the leading and sub-leading leptons, over total energy of selected objects			√
$\Delta\phi(\ell\ell, \text{jet}^{\text{leading}})$	$\Delta\phi$ between dilepton system and leading jet	√	√	
$p_T(\ell\ell, \text{jet}^{\text{leading}})$	p_T of dilepton and leading jet system	√		√
$p_T(\ell^{\text{leading}}, \text{jet}^{\text{leading}})$	p_T of leading lepton and leading jet system	√		
Centrality($\ell\ell, \text{jet}^{\text{leading}}$)	Scalar sum of p_T of the dilepton system and leading jet, over total energy of selected objects	√		
$\Delta R(\ell, \ell)$	ΔR between leading and sub-leading leptons	√		
$\Delta R(\ell^{\text{leading}}, \text{jet}^{\text{leading}})$	ΔR between leading lepton and leading jet	√		
$M(\ell^{\text{leading}}, \text{jet}^{\text{leading}})$	Invariant mass of leading lepton and leading jet		√	
$M(\text{jet}^{\text{leading}}, \text{jet}^{\text{sub-leading}})$	Invariant mass of leading jet and sub-leading jet		√	
$\Delta R(\ell^{\text{leading}}, \text{jet}^{\text{sub-leading}})$	ΔR between leading lepton and sub-leading jet		√	
$\Delta R(\ell\ell, \text{jet}^{\text{leading}})$	ΔR between dilepton system and leading jet		√	√
$\Delta p_T(\ell^{\text{sub-leading}}, \text{jet}^{\text{sub-leading}})$	$p_T^{\text{sub-leading lepton}} - p_T^{\text{sub-leading jet}}$		√	
$M(\ell^{\text{sub-leading}}, \text{jet}^{\text{leading}})$	Invariant mass of sub-leading lepton and leading jet			√

4.8.1 Data/MC comparison for MVA input variables

The data and MC comparison for MVA input variables are shown from Figure 4.14 to 4.15 for NN₁₁, and from Figure 4.16 to 4.17 for NN₂₁ and from Figure 4.18 to 4.20 for NN_{FCNC}. Sources of systematic uncertainties which are considered in the plots are explained in Section 4.9. Considering the systematic uncertainties, data/MC are in good agreement and no obvious large mis-modelling is observed.

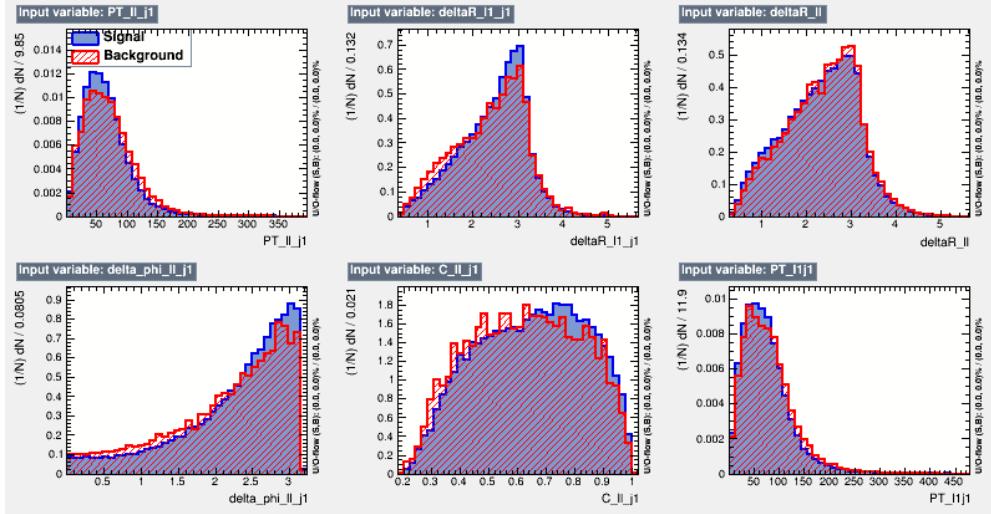


Figure 4.10: MLP input variable distributions for NN₁₁. Signal distributions are drawn in blue and the background distributions in red.

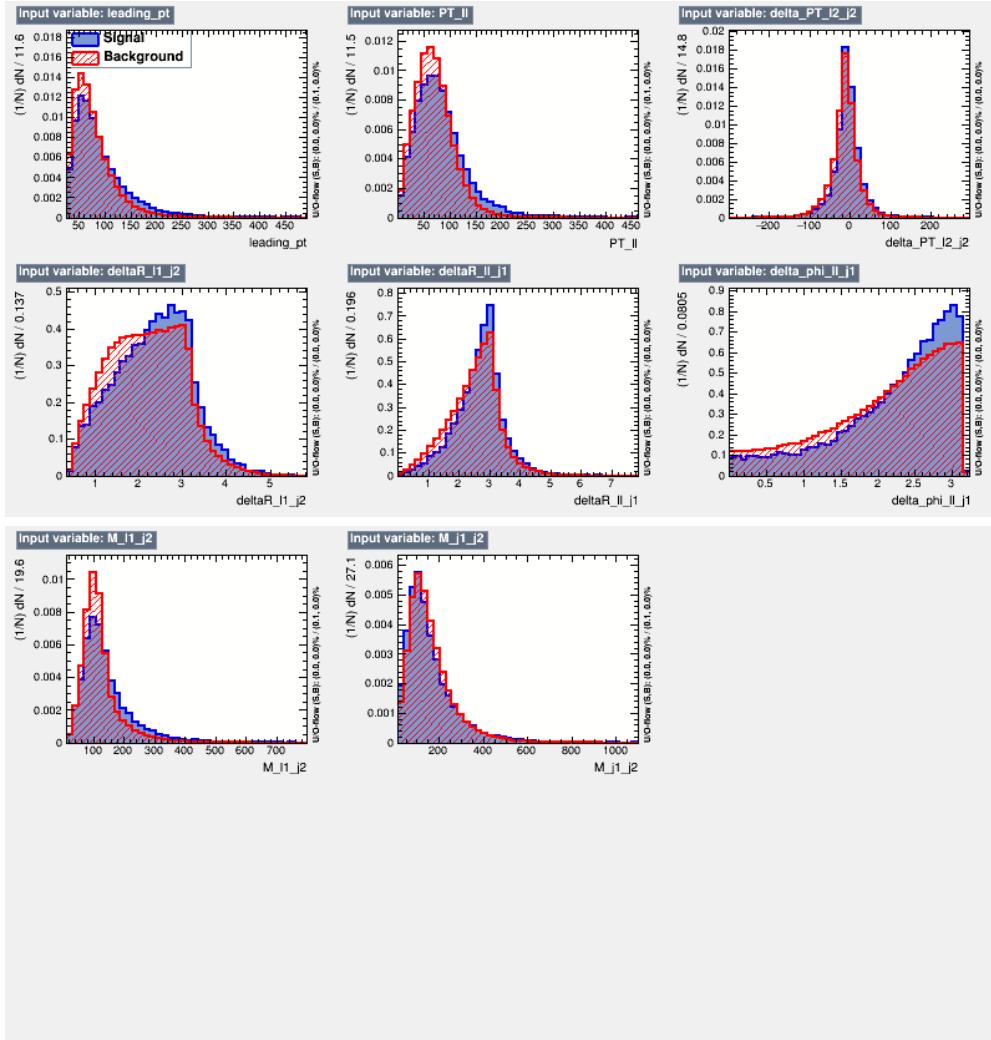


Figure 4.11: MLP input variable distributions for NN₂₁. Signal distributions are drawn in blue and the background distributions in red.

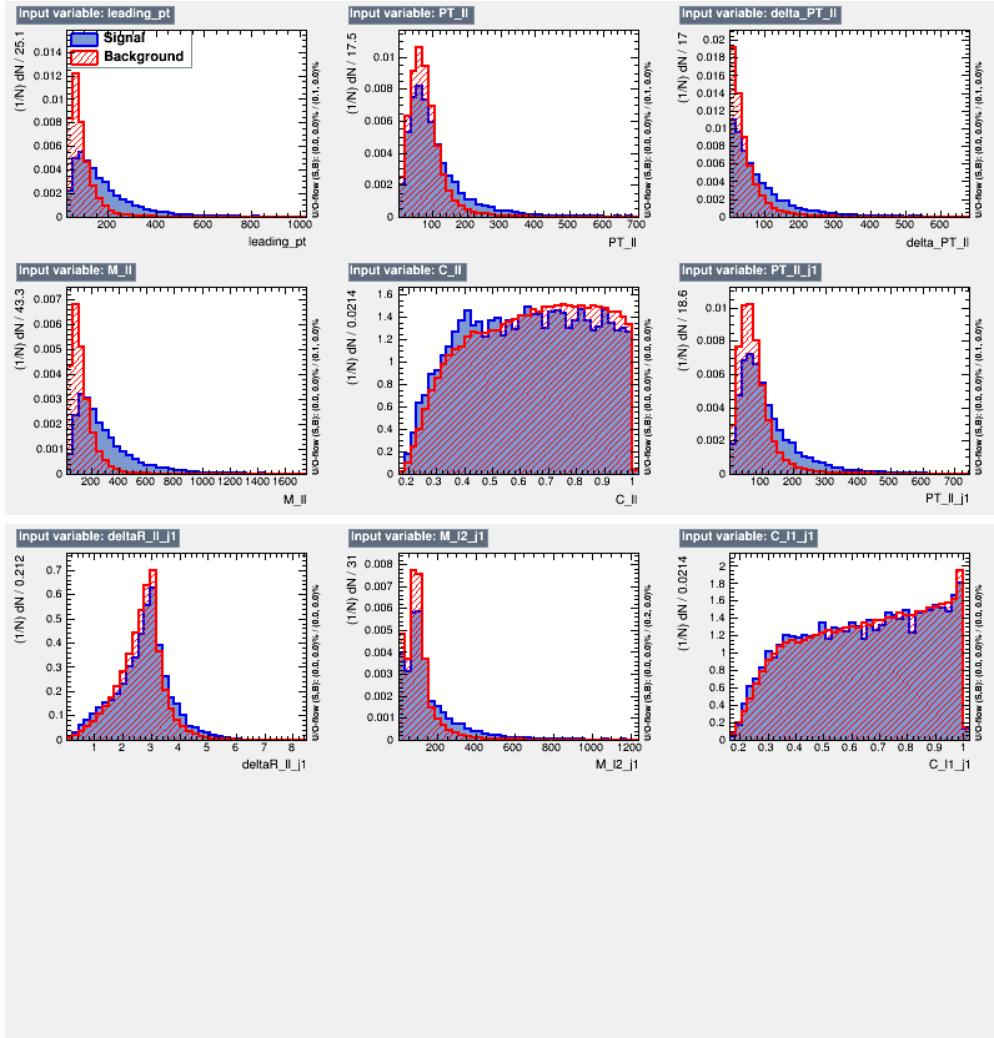


Figure 4.12: MLP input variable distributions for NN_{FCNC} . Signal distributions are drawn in blue and the background distributions in red.

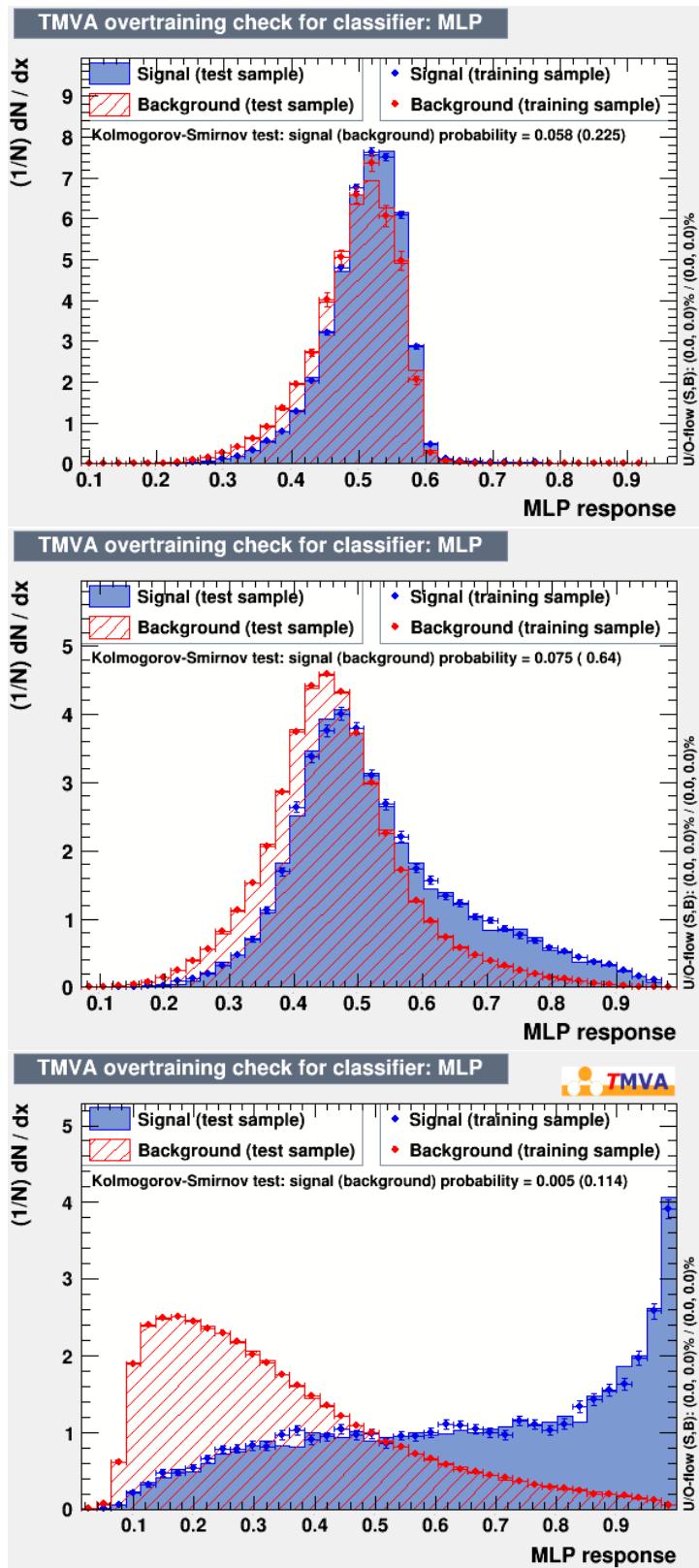


Figure 4.13: MLP output for test and train samples for NN₁₁ (top), NN₂₁ (middle) and NN_{FCNC} (bottom)

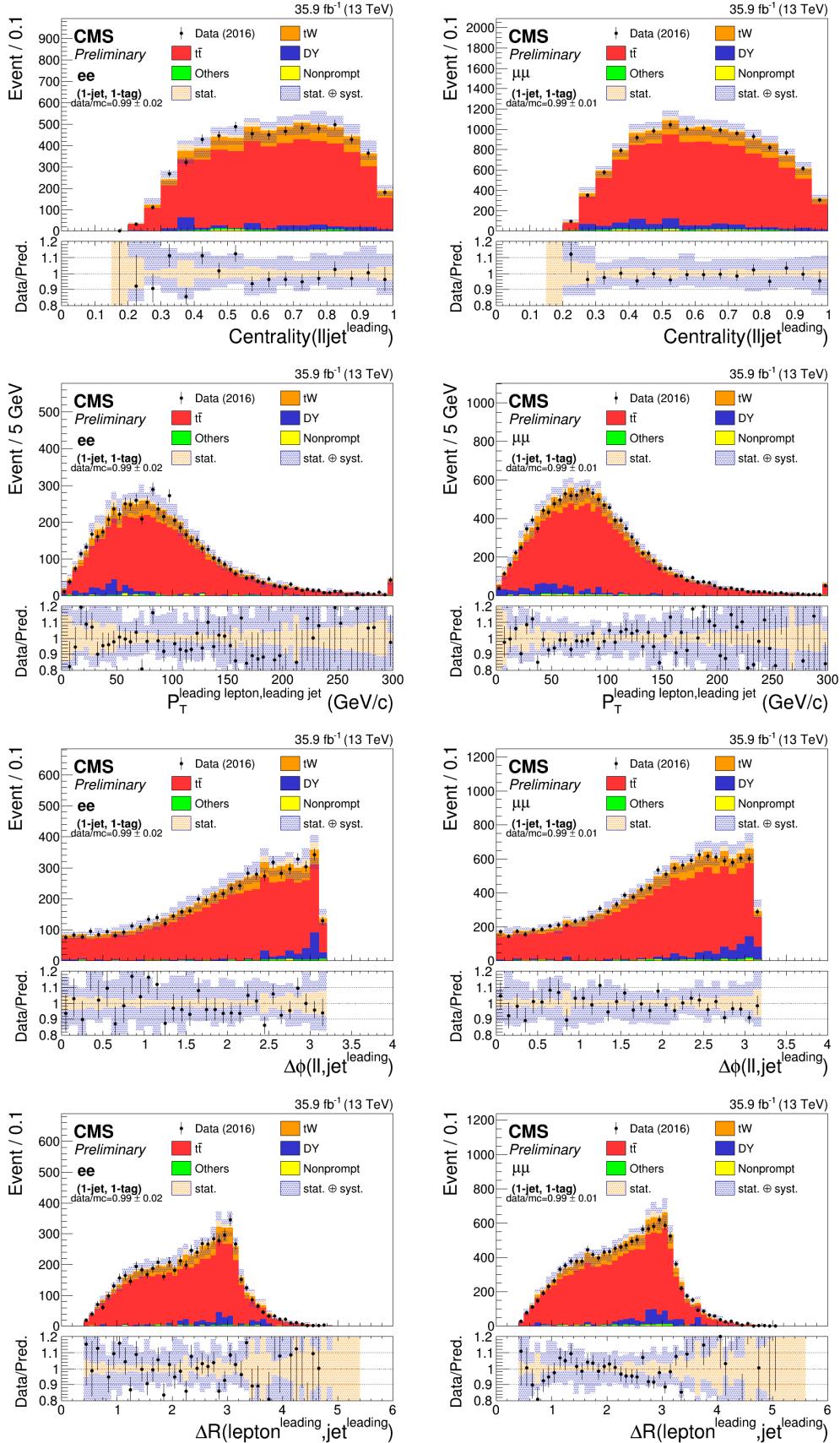


Figure 4.14: The distributions of variables used for MVA input for ee (left) and $\mu\mu$ (right) channels for NN_{11} .

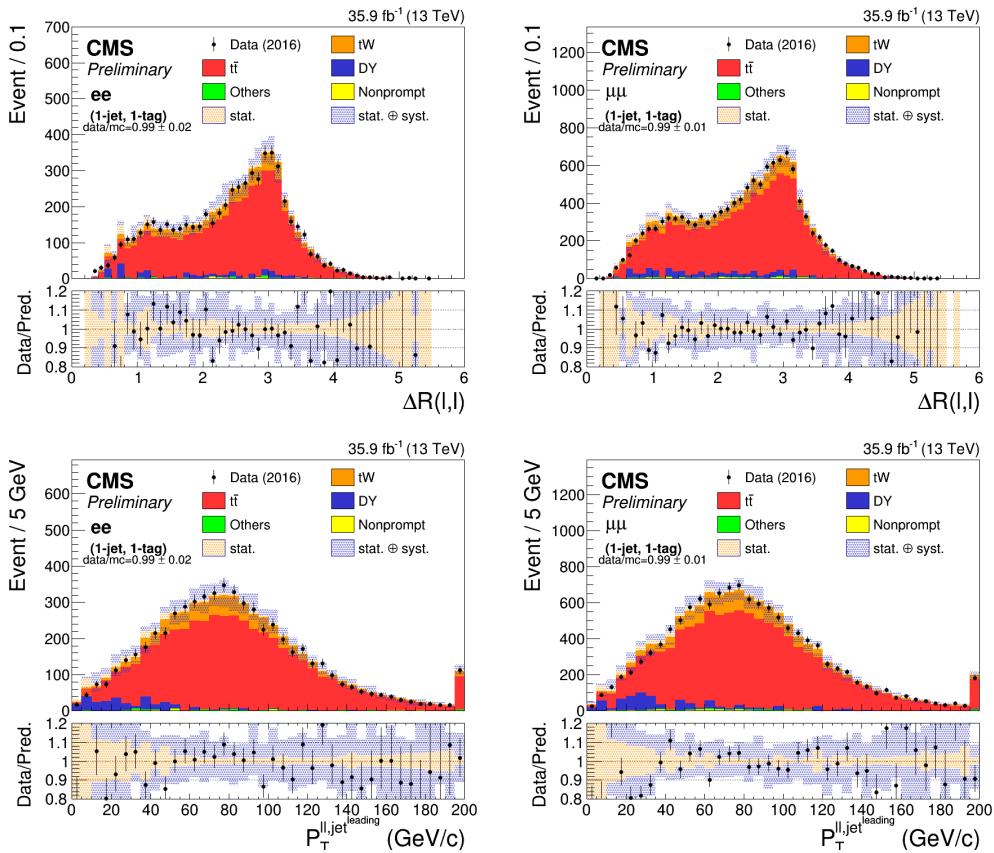


Figure 4.15: The distributions of variables used for MVA input for ee (left) and $\mu\mu$ (right) channels for NN₁₁.

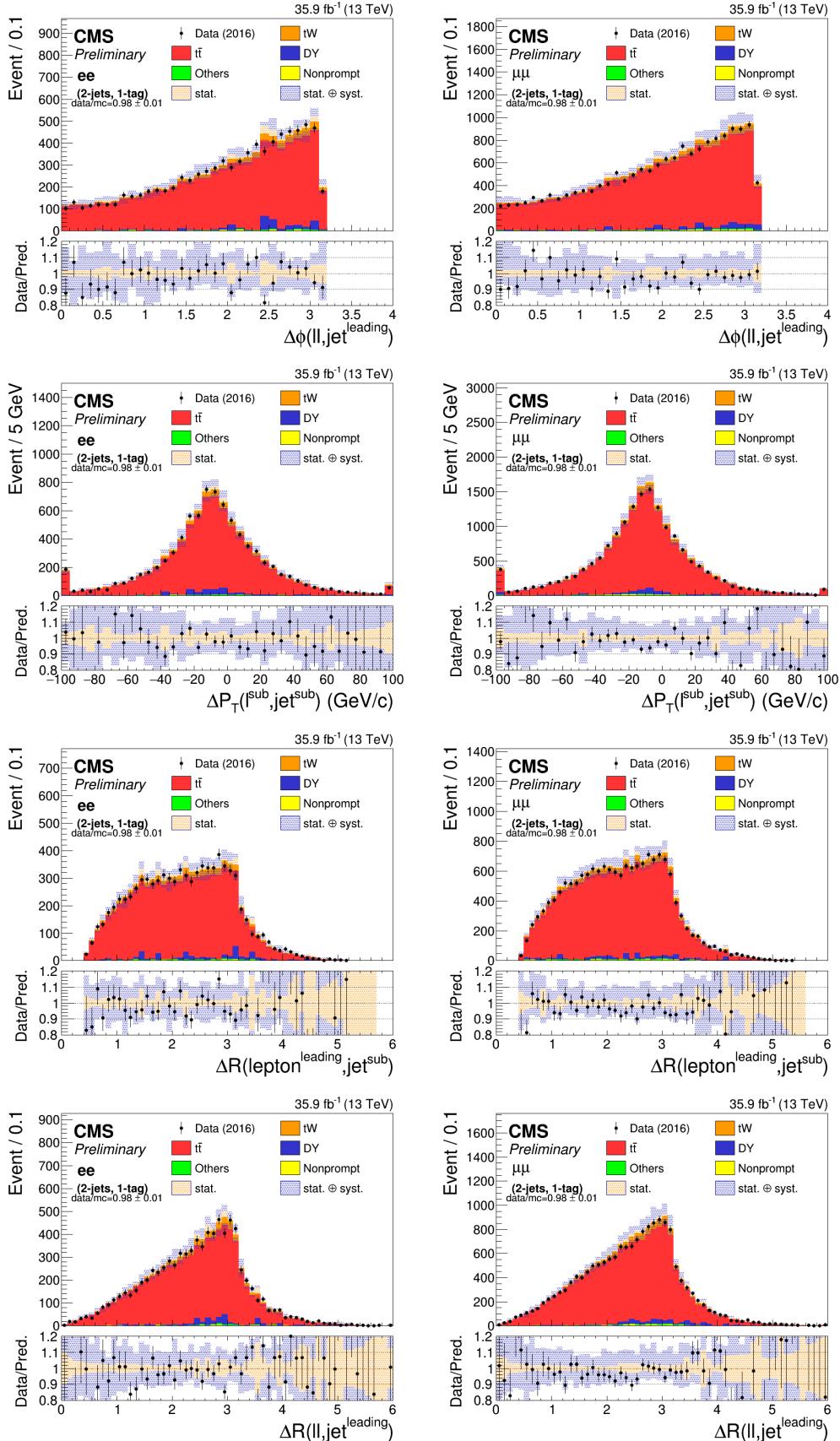


Figure 4.16: The distributions of variables used for MVA input for ee (left) and $\mu\mu$ (right) channels for NN_{21} .

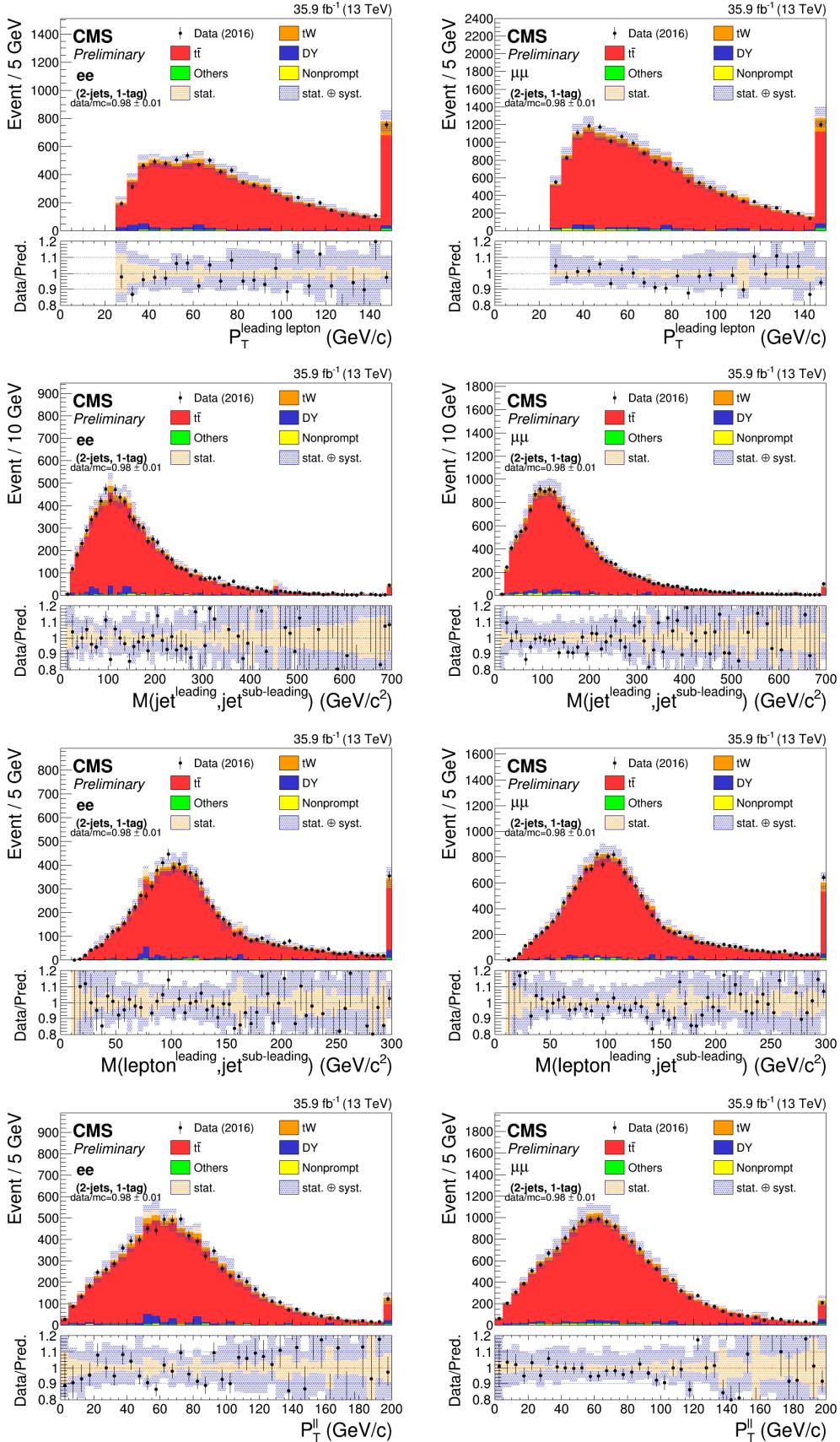


Figure 4.17: The distributions of variables used for MVA input for ee (left) and $\mu\mu$ (right) channels for NN_{21} .

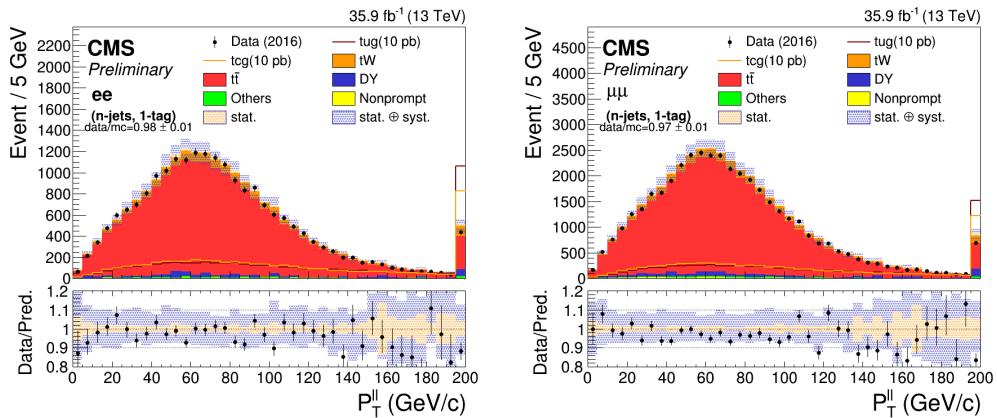


Figure 4.18: The distributions of variables used for MVA input for FCNC study for ee (left) and $\mu\mu$ (right) channels for NN_{FCNC} .

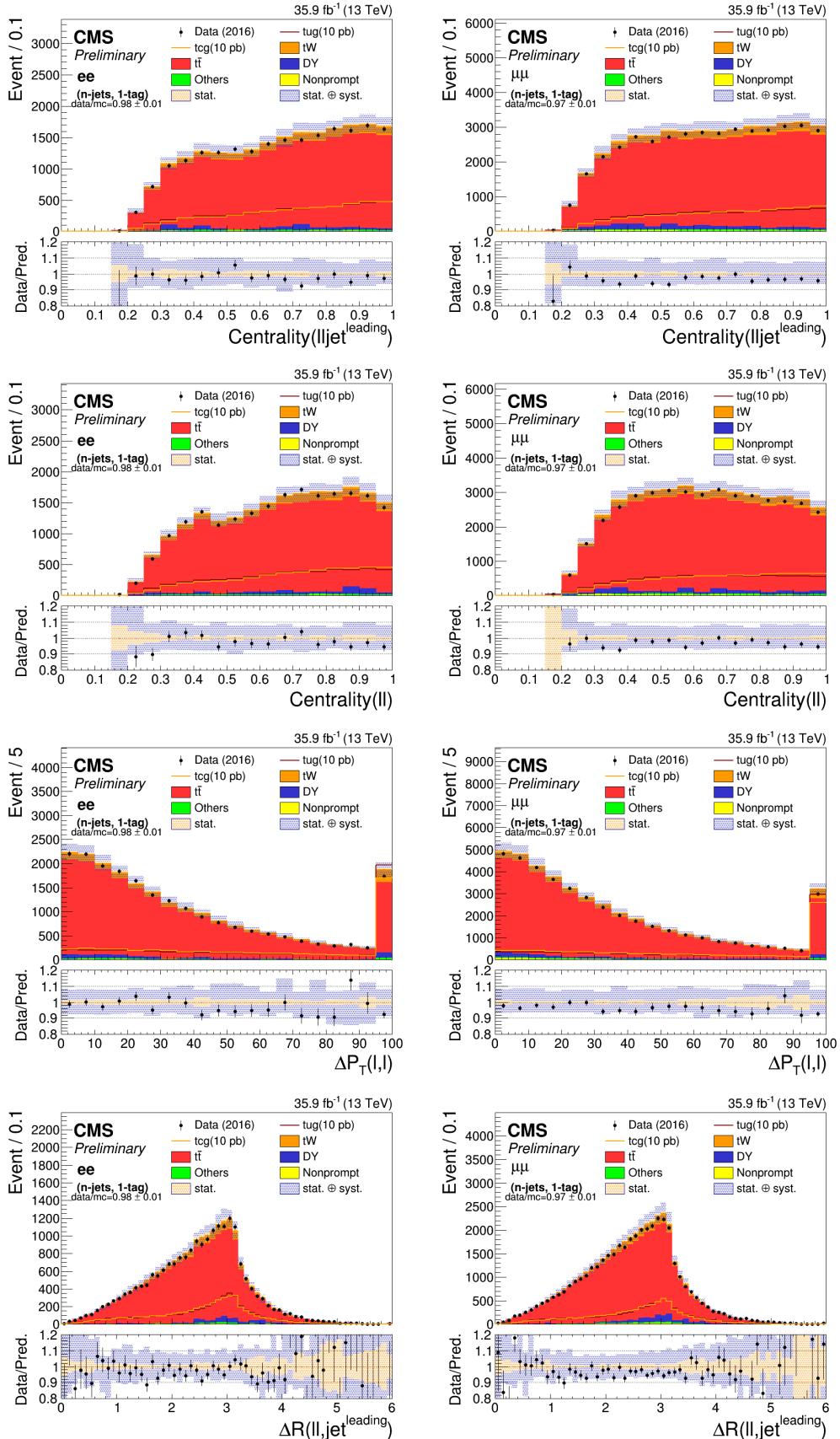


Figure 4.19: The distributions of variables used for MVA input for FCNC study for ee (left) and $\mu\mu$ (right) channels for NN_{FCNC} .

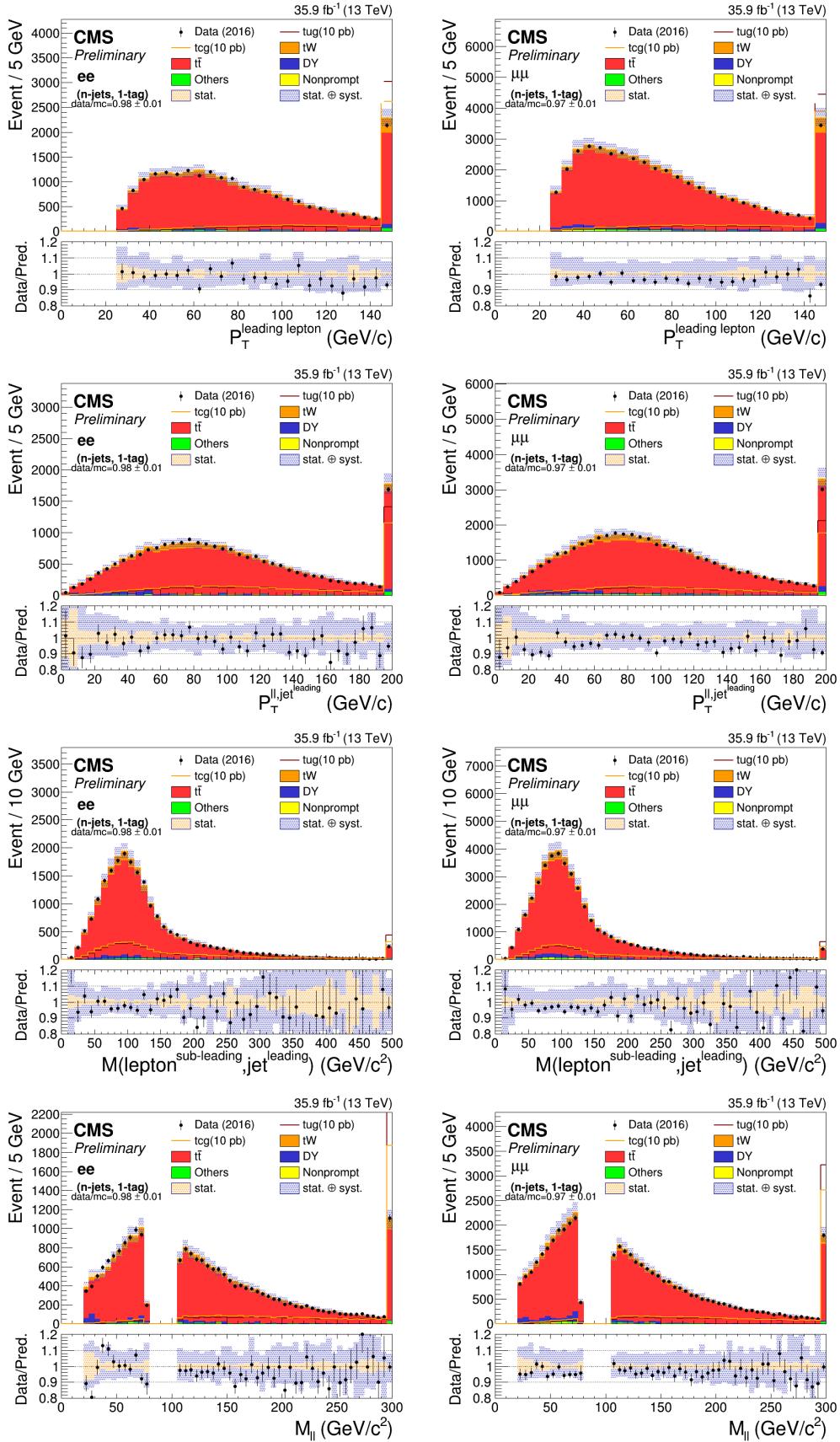


Figure 4.20: The distributions of variables used for MVA input for FCNC study for ee (left), $e\mu$ (middle) and $\mu\mu$ (right) channels for NNFCNC.

4.9 Systematic uncertainties

This analysis depends on both the normalization and shape of the background and signal expectations. We consider the following sources of systematic uncertainties which affect the shape and normalization of the templates used in the statistical evaluation:

- **$t\bar{t}$ and tW modeling uncertainty:** there are various input for generating and simulating $t\bar{t}$ and tW processes. In Table 4.11, related samples used for studying the effect of the modeling uncertainties are sorted. The $t\bar{t}$ and tW modeling uncertainty sources are
 - ***QCD scale uncertainty:*** This uncertainty is estimated by varying the renormalization and the factorization scales, used during the MC generation of the sample, independently by a factor 0.5, 1 or 2. Unphysical cases, where one scale fluctuate up while the other fluctuate down, are not considered. An envelope is built from all the 6 possible variations by taking in each bin of the distribution the maximum (minimum) variation and is used as an estimate of the QCD scale uncertainties for $t\bar{t}$ sample. For tW , independent samples are used (see Table 4.11).
 - ***Parton Distribution Functions uncertainty:*** The magnitude of the uncertainties related to the parton distribution functions and the variation of the strong coupling constant for each simulated signal processes is obtained using the replicas of the NNPDF 3.0 set (**NOTE no weight for calculating the PDF uncertainty is presented in tW samples. So this uncertainty is not included for tW).
 - ***Top mass:*** The most recent measurement of the top quark mass by CMS yields a total uncertainty of ± 0.49 GeV. We consider a sample with varied top mass at $m_t = 172.5 \pm 3.0$ GeV and reduce the obtained uncertainty by a factor of 6.
 - ***$tW/t\bar{t}$ interference:*** At NLO QCD, tW production is expected to interfere with $t\bar{t}$ production [78]. Two schemes for defining the tW signal in a way which distinguishes it from $t\bar{t}$ production have been compared in our analysis: 'diagram removal' (DR), in which all doubly resonant NLO tW diagrams are removed, and 'diagram subtraction' (DS), where a gauge-invariant subtractive term modifies the NLO tW cross section to locally cancel the contribution from $t\bar{t}$. The difference between the samples simulated using the two approaches is taken as a systematic uncertainty.
 - ***ME/PS matching:*** The model parameter h_{damp} in POWHEG that controls the matching of the matrix elements to the PYTHIA parton showers is varied from its default value of 172.5 GeV by factors of 0.5 and 2. This source of uncertainty is only considered for $t\bar{t}$.
 - ***scale variations of initial state radiation (ISR) and final state radiation FSR:*** From a practical point of view, we vary the renormalization scale for QCD emissions in FSR and ISR by a factor of 0.5 and 2. The uncertainty for the FSR shower scale variation is reduced by a factor of $\sqrt{2}$ following the TOP PAG recommendations.
 - ***Underlying event:*** the default parameters in the CUETP8M2T4 are varied according to their uncertainty and the effect on the unfolding is taken as an estimate of the systematic uncertainty. This source of uncertainty is only considered for $t\bar{t}$.
 - ***Color reconnection:*** we vary the color reconnection model with respect to the default using alternatives including the resonant decay products in possible reconstructions to the UE. The default simulation (MPI-based color reconnection) has this effect excluded. We examine three alternative models for CR: the so-called gluon move, early resonance decay and the QCD-inspired models. The envelope of the differences is considered as a systematic uncertainty. This source of uncertainty is only considered for $t\bar{t}$.

Sample	Events
Nominal:TTTo2L2Nu_TuneCUETP8M2_ttHtranche3_13TeV-powheg	64910035
Nominal:ST_tW_top_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	8609398
Nominal:ST_tW_antitop_5f_NoFullyHadronicDecays_13TeV-powheg_TuneCUETP8M1	8681265
ST_tW_top_5f_DS_NoFullyHadronicDecays_13TeV-powheg-pythia8	3192538
ST_tW_antitop_5f_DS_NoFullyHadronicDecays_13TeV-powheg-pythia8	3098002
ST_tW_top_5f_MEscaleup_NoFullyHadronicDecays_13TeV-powheg	3188665
ST_tW_antitop_5f_MEscaleup_NoFullyHadronicDecays_13TeV-powheg	1606961
ST_tW_top_MEscaledown_5f_NoFullyHadronicDecays_13TeV-powheg	3051991
ST_tW_antitop_MEscaledown_5f_NoFullyHadronicDecays_13TeV-powheg	1575142
ST_tW_top_5f_mtop1695_NoFullyHadronicDecays_13TeV-powheg-pythia8	3178900
ST_tW_antitop_5f_mtop1695_NoFullyHadronicDecays_13TeV-powheg-pythia8	2968744
ST_tW_top_5f_mtop1755_NoFullyHadronicDecays_13TeV-powheg-pythia8	2938402
ST_tW_antitop_5f_mtop1755_NoFullyHadronicDecays_13TeV-powheg-pythia8	3194626
ST_tW_top_5f_isrup_NoFullyHadronicDecays_13TeV-powheg	3110339
ST_tW_top_isrdown_5f_NoFullyHadronicDecays_13TeV-powheg	3181500
ST_tW_antitop_5f_isrup_NoFullyHadronicDecays_13TeV-powheg	3076275
ST_tW_antitop_isrdown_5f_NoFullyHadronicDecays_13TeV-powheg	3101321
ST_tW_top_5f_fsrup_NoFullyHadronicDecays_13TeV-powheg	3192325
ST_tW_top_fsrdown_5f_NoFullyHadronicDecays_13TeV-powheg	2935595
ST_tW_antitop_5f_fsrup_NoFullyHadronicDecays_13TeV-powheg	3001527
ST_tW_antitop_fsrdown_5f_NoFullyHadronicDecays_13TeV-powheg	3234964
TT_TuneCUETP8M2T4_QCDbasedCRTune_erdON_13TeV-powheg-pythia8	57788977
TT_TuneCUETP8M2T4_erdON_13TeV-powheg-pythia8	58448827
TT_TuneCUETP8M2T4_GluonMoveCRTune_13TeV-powheg-pythia8	56456001
TT_TuneCUETP8M2T4_mtop1695_13TeV-powheg-pythia8	58281931
TT_TuneCUETP8M2T4_mtop1755_13TeV-powheg-pythia8	38909457
TT_TuneCUETP8M2T4_13TeV-powheg-fsrdown-pythia8	57563666
TT_TuneCUETP8M2T4_13TeV-powheg-fsrup-pythia8	58475264
TT_TuneCUETP8M2T4_13TeV-powheg-isrdown-pythia8	58421030
TT_TuneCUETP8M2T4_13TeV-powheg-isrup-pythia8	57577179
TT_hdampDOWN_TuneCUETP8M2T4_13TeV-powheg-pythia8	55809842
TT_hdampUP_TuneCUETP8M2T4_13TeV-powheg-pythia8	58320199
TT_TuneCUETP8M2T4down_13TeV-powheg-pythia8	57721717
TT_TuneCUETP8M2T4up_13TeV-powheg-pythia8	58144172

Table 4.11: MC samples used for systematic study

In figures 4.21-4.22, the relative effect of each source of $t\bar{t}$ and tW modeling uncertainties for the MLP output distributions are shown. Relative uncertainties stands for the ratio of MLP for tW ($t\bar{t}$) systematic sample to nominal tW ($t\bar{t}$) sample. All modeling uncertainties are shape dependent uncertainties.

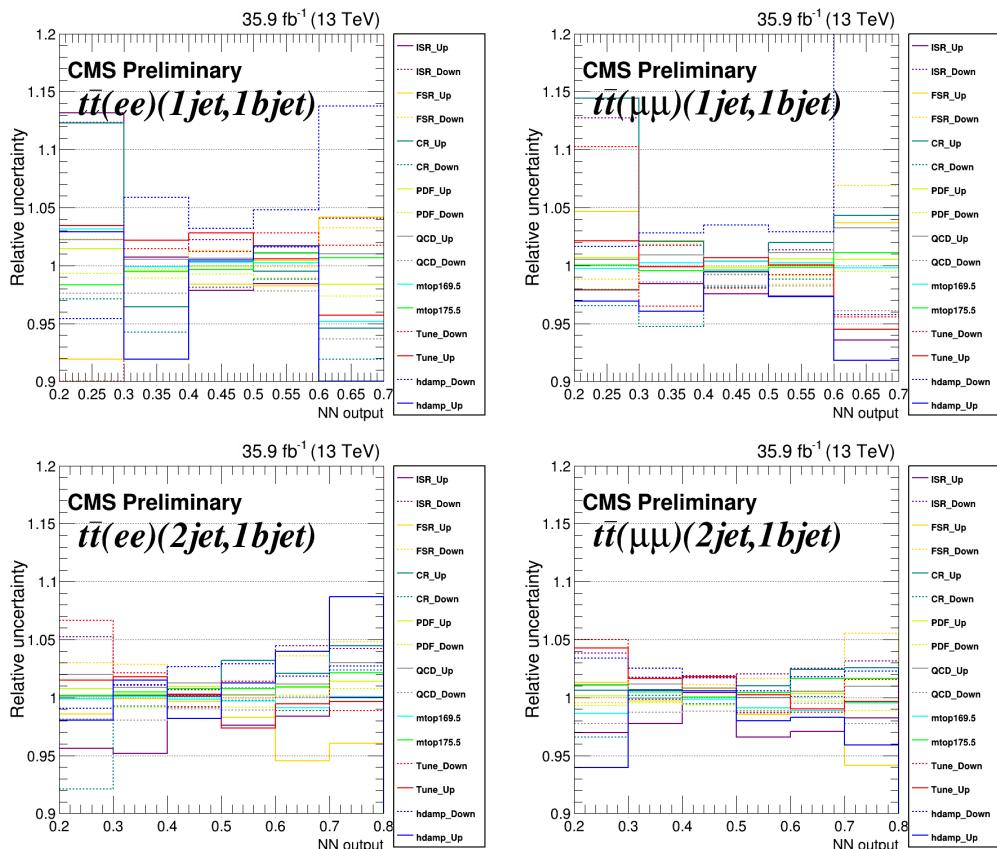


Figure 4.21: The relative effect of the $t\bar{t}$ modeling uncertainties in MLP distribution for different (1jet,1b-jet) region (top row), (2jet,1b-jet) region (bottom row) for ee channel (left column) and $\mu\mu$ channel (right column).

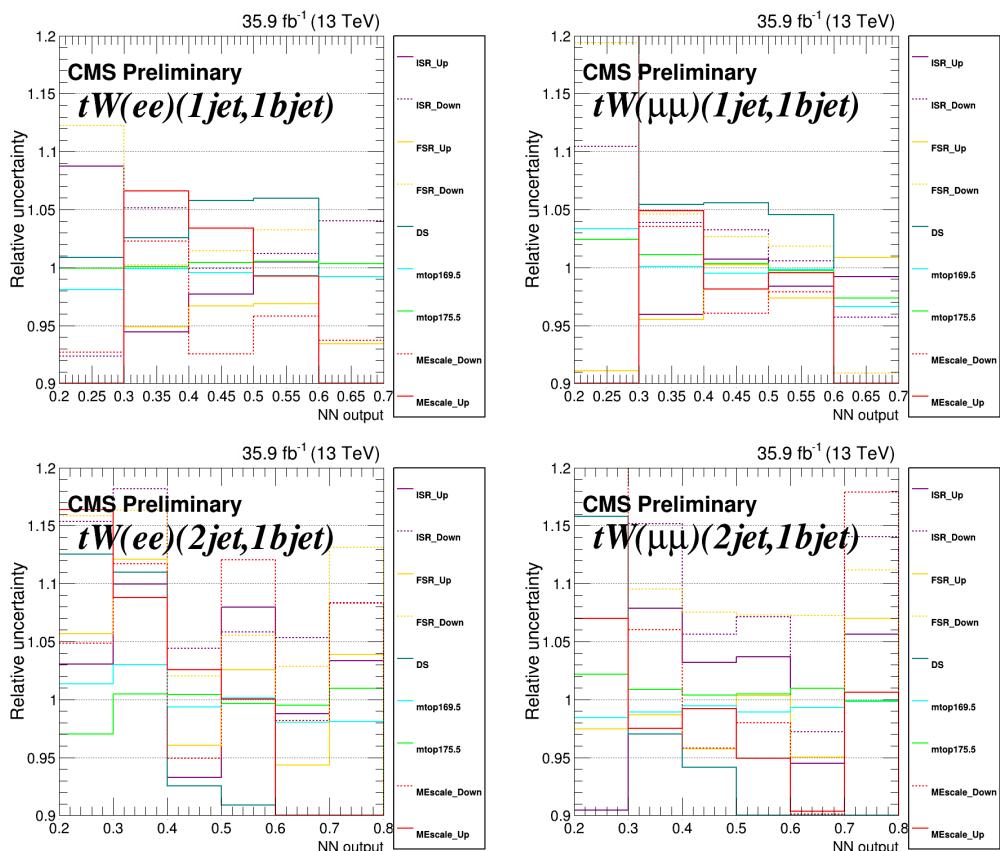


Figure 4.22: The relative effect of the tW modeling uncertainties in MLP distribution for different (1jet,1b-jet) region (top row), (2jet,1b-jet) region (bottom row) for ee channel (left column) and $\mu\mu$ channel (right column).

- **Lepton reconstruction, identification and isolation scale factors:** Electrons and muons reconstruction, isolation and identification scale factors and uncertainties are provided centrally by related POGs, extracted with a tag-and-probe analysis on Z to ll events.
- **Jet energy scale and resolution:** uncertainties in the jet energy scale and resolution is provided officially by JET/MET POG in the recommended global tag [79]. In order to find the latest jet energy scale uncertainty, '80X_mcRun2_asymptotic_2016_TrancheIV_v8' global tag is used. Variation of jet energy scale and resolution are propagated to MET and MET is corrected due to the changes.
- **Unclustered energy uncertainty:** uncertainty on unclustered energy is considered which propagated to MET and MET is corrected due to the changes.
- **Trigger scale factor:** uncertainty on the trigger scale factor are provided in TOP recommended scale factor root files.
- **b-tagging:** The efficiency for b-tagging is determined for the baseline selection and then scaled up and down according to their uncertainties given by the BTV group. The b-quark and c-quark jet efficiencies are varied simultaneously, while the efficiencies for the light quarks are varied independently [80].
- **Pile-up reweighting:** The measured minimum-bias cross section (69.2 mb) is varied by 4.6% to produce different expected pileup distributions for data (up and down).
- **Luminosity:** A systematic uncertainty of 2.5% is assigned to the integrated luminosity and is used for background rates [81].
- **$t\bar{t}$ normalization:** uncertainty on $t\bar{t}$ normalization is considered to be 5% [82] for $O_{\phi q}^{(3)}$, O_{tW} , O_{uG} and O_{cG} study, 3% for O_G study due to the observed difference between the $t\bar{t}$ kinematic distribution with and without O_G .
- **tW normalization:** uncertainty on tW normalization is considered to be 10% for O_{uG} , O_{cG} and O_G study.
- **Non-top background normalization:** uncertainty on DY normalization in ee and $\mu\mu$ channels (data-driven normalization) is considered to be 30% in all njet-mtag regions. DY normalization uncertainty is considered to be uncorrelated between various njet-mbtag regions. The uncertainty on other and jet backgrounds are considered to be 50%.

For FCNC signal (O_{uG} and O_{cG}) study, additional uncertainties are considered in following.

- **Parton Distribution Functions uncertainty:** The magnitude of the uncertainties related to the parton distribution functions and the variation of the strong coupling constant for FCNC tW simulated signal processes is obtained using the replicas of the NNPDF 3.0 set. Each event is weighted with respect to the LHE weights provided for each replicas of the NNPDF 3.0 set and final NN distribution is found. One sigma UP/DOWN uncertainty from the distribution of the NN output due to the various PDF set with respect to the nominal set is assigned as PDF error.

• **QCD scale uncertainty:** This uncertainty is estimated by varying the renormalization and the factorization scales for FCNC tW simulated signal, used during the MC generation of the sample by a factor 0.5 and or 2. Each event is weighted with respect to the LHE weights provided for renormalization and factorization scale variation. The largest deviation from the nominal value is taken as QCD scale error.

- **Parton shower Q scale uncertainty:** The scales of the initial (ISR) and final (FSR) state shower are varied up and down by a factor of two FCNC tW simulated signal samples. MC samples used to estimate the parton shower Q scale uncertainties are summarised in Table 4.12.

In Figure 4.23, the effect of shape dependent uncertainties except $t\bar{t}$ and tW modeling are shown.

sample
ST_tW_tcgFCNC_scaledown_leptonDecays_Madgraph
ST_tW_tcgFCNC_scaleup_leptonDecays_Madgraph
ST_tW_tugFCNC_scaledown_leptonDecays_Madgraph
ST_tW_tugFCNC_scaleup_leptonDecays_Madgraph

Table 4.12: Systematic samples for FCNC signal study

The relative effect of the all uncertainties on the total normalization are summarized in tables 4.13, and 4.14 for ee and $\mu\mu$ channel respectively.

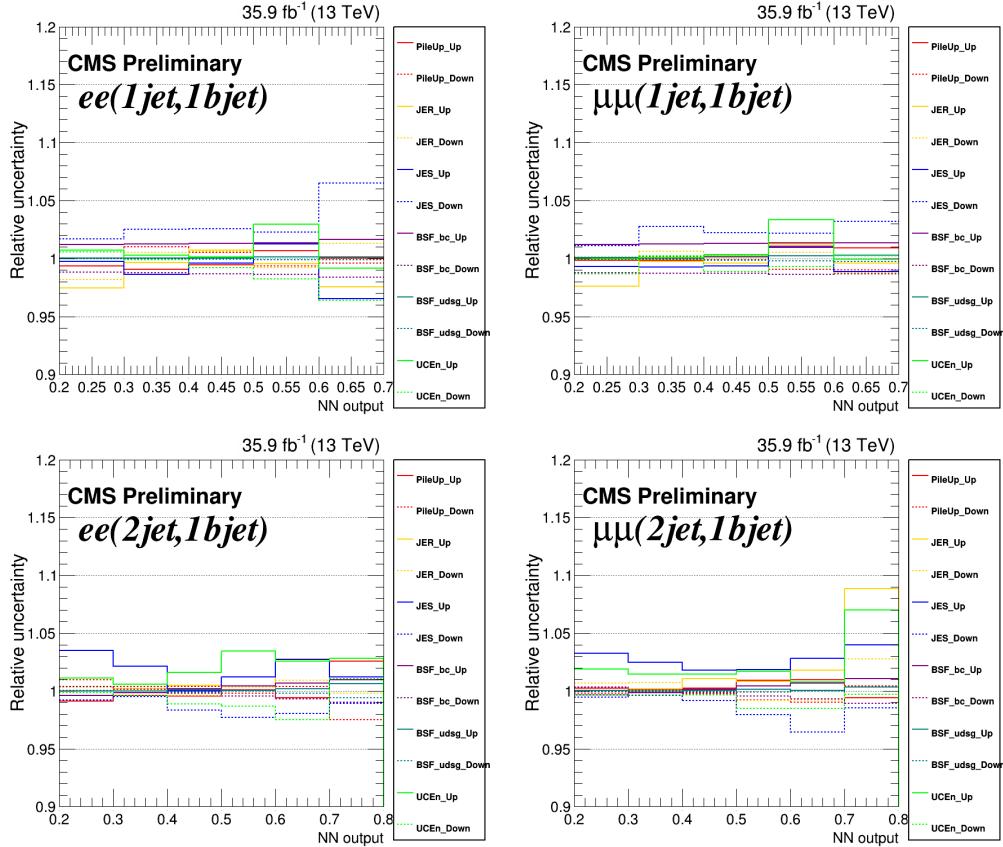


Figure 4.23: The relative effect of the shape dependent uncertainties in MLP distribution for different (1jet,1b-jet) region (top row), (2jet,1b-jet) region (bottom row) for ee channel (left column) and $\mu\mu$ channel (right column).

Table 4.13: all systematic uncertainty effect for ee channel

all	all	1jet, 1bjet	2jet, 1bjet	>= 2jet, 2bjet
nominal	64928.258	5936.414	8330.690	7973.390
luminosity_up	2.438%	2.480%	2.459%	2.459%
luminosity_down	-2.436%	-2.453%	-2.459%	-2.459%
TT_up	2.877%	3.993%	4.489%	4.742%
TT_down	-2.877%	-3.993%	-4.489%	-4.742%
DY_up	8.886%	1.305%	0.867%	0.173%
DY_down	-8.886%	-1.305%	-0.867%	-0.173%
Jets_up	0.393%	0.010%	0.105%	0.136%
Jets_down	-0.393%	-0.010%	-0.105%	-0.136%
other_up	3.730%	0.287%	0.290%	0.313%
other_down	-3.730%	-0.287%	-0.290%	-0.313%
TriggerSF_up	0.834%	0.829%	0.809%	0.792%
TriggerSF_down	-0.834%	-0.829%	-0.809%	-0.792%
PileUp_up	2.883%	-0.035%	-0.220%	-0.630%
PileUp_down	-2.757%	0.026%	0.162%	0.616%
JER_up	6.354%	-0.022%	0.520%	0.736%
JER_down	1.454%	-0.055%	-0.229%	-0.024%
JES_up	5.257%	0.215%	1.167%	3.161%
JES_down	-4.670%	2.485%	-1.465%	-2.931%
BTagSF_bc_up	0.073%	1.333%	0.207%	2.520%
BTagSF_bc_down	-0.073%	-1.333%	-0.236%	-2.491%
BTagSF_udsg_up	-0.022%	0.082%	0.081%	0.127%
BTagSF_udsg_down	0.022%	-0.082%	-0.081%	-0.128%
EleRecoSF_up	1.002%	1.009%	1.006%	1.004%
EleRecoSF_down	-0.997%	-1.004%	-1.001%	-0.999%
EleIDIsoSF_up	2.601%	2.618%	2.623%	2.620%
EleIDIsoSF_down	-2.568%	-2.584%	-2.589%	-2.586%
UnclusteredEn_up	10.246%	1.408%	1.779%	0.896%
UnclusteredEn_down	-3.109%	-1.304%	-0.891%	-0.534%
ISR_down	0.596%	1.777%	1.501%	-0.052%
ISR_up	-0.217%	-1.427%	-1.480%	-0.689%
FSR_down	0.862%	0.779%	1.462%	6.013%
FSR_up	-0.832%	-1.802%	-1.003%	-7.126%
TT_CR_Up	1.435%	-0.552%	0.502%	1.733%
TT_CR_Down	-0.475%	-1.515%	-0.235%	0.418%
TT_PDF_Up	0.203%	0.007%	0.556%	0.602%
TT_PDF_Down	-0.306%	-0.668%	-0.401%	-0.553%
TT_QCD_Up	0.076%	0.631%	0.740%	0.315%
TT_QCD_Down	-0.164%	-1.691%	-0.833%	-0.271%
TT_TopMass_down	-0.106%	0.161%	-0.172%	-0.752%
TT_TopMass_up	0.313%	0.168%	0.669%	0.746%
TT_Tune_down	0.544%	1.322%	0.711%	0.320%
TT_Tune_up	0.252%	1.386%	-0.059%	0.446%
TT_hdamp_down	0.105%	3.320%	1.957%	0.024%
TT_hdamp_up	0.789%	-0.098%	0.269%	1.191%
TW_DS	-0.287%	0.699%	-0.656%	-0.776%
TW_TopMass_down	-0.025%	-0.080%	-0.026%	-0.035%
TW_TopMass_up	0.024%	0.064%	0.008%	0.056%
TW_MEscale_down	-0.021%	-0.764%	0.212%	0.275%
TW_MEscale_up	-0.118%	0.169%	-0.129%	-0.307%
total_up	17.429%	7.865%	7.131%	9.870%
total_down	-12.504%	-6.885%	-6.564%	-10.266%

Table 4.14: all systematic uncertainty effect for $\mu\mu$ channel

all	all	1jet, 1bjet	2jet, 1bjet	≥ 2 jet, 2bjet
nominal	138879.681	12365.806	16751.491	15888.515
luminosity_up	2.457%	2.469%	2.469%	2.490%
luminosity_down	-2.457%	-2.469%	-2.469%	-2.490%
TT_up	2.688%	3.922%	4.473%	4.763%
TT_down	-2.688%	-3.922%	-4.473%	-4.763%
DY_up	10.341%	1.806%	0.896%	0.154%
DY_down	-10.341%	-1.806%	-0.896%	-0.154%
Jets_up	0.483%	0.351%	0.399%	0.101%
Jets_down	-0.483%	-0.351%	-0.399%	-0.101%
other_up	3.655%	0.330%	0.291%	0.261%
other_down	-3.655%	-0.330%	-0.291%	-0.261%
TriggerSF_up	0.639%	0.663%	0.650%	0.643%
TriggerSF_down	-0.639%	-0.663%	-0.650%	-0.643%
PileUp_up	3.973%	0.568%	0.372%	0.038%
PileUp_down	-3.749%	-0.484%	-0.364%	-0.045%
JER_up	6.743%	0.354%	1.174%	0.769%
JER_down	2.153%	0.367%	0.051%	-0.012%
JES_up	6.038%	-0.217%	2.174%	3.160%
JES_down	-4.781%	2.199%	-1.209%	-2.834%
BTagSF_bc_up	0.072%	1.316%	0.201%	2.535%
BTagSF_bc_down	-0.072%	-1.316%	-0.230%	-2.506%
BTagSF_udsg_up	-0.019%	0.191%	0.064%	0.141%
BTagSF_udsg_down	0.019%	-0.191%	-0.064%	-0.140%
MuIDSF_up	2.378%	2.316%	2.316%	2.331%
MuIDSF_down	-2.350%	-2.289%	-2.290%	-2.304%
MuIsoSF_up	1.015%	1.015%	1.014%	1.020%
MuIsoSF_down	-1.010%	-1.010%	-1.009%	-1.015%
MuTrackSF_up	0.035%	0.034%	0.033%	0.031%
MuTrackSF_down	-0.035%	-0.034%	-0.033%	-0.031%
UnclusteredEn_up	11.621%	1.484%	1.730%	1.095%
UnclusteredEn_down	-3.669%	-0.959%	-0.694%	-0.315%
ISR_down	0.663%	0.720%	2.338%	-0.018%
ISR_up	-0.121%	-1.984%	-1.059%	-0.906%
FSR_down	1.006%	-0.185%	1.897%	6.035%
FSR_up	-0.433%	0.220%	-0.286%	-7.915%
TT_CR_Up	1.409%	0.843%	0.866%	1.612%
TT_CR_Down	-0.229%	-1.486%	-0.298%	0.702%
TT_PDF_Up	0.169%	0.073%	0.320%	0.548%
TT_PDF_Down	-0.275%	-0.526%	-0.591%	-0.541%
TT_QCD_Up	0.006%	0.554%	0.487%	0.166%
TT_QCD_Down	-0.098%	-1.349%	-0.961%	-0.139%
TT_TopMass_down	-0.058%	0.205%	-0.378%	-0.454%
TT_TopMass_up	0.248%	-0.226%	0.382%	0.989%
TT_Tune_down	0.517%	-1.182%	0.669%	1.669%
TT_Tune_up	0.337%	0.247%	1.175%	-1.057%
TT_hdamp_down	0.307%	2.577%	1.558%	-0.604%
TT_hdamp_up	0.279%	-1.488%	-0.518%	-0.526%
TW_DS	-0.193%	0.629%	-0.636%	-0.596%
TW_TopMass_down	-0.009%	-0.047%	-0.043%	-0.004%
TW_TopMass_up	0.012%	0.009%	0.028%	-0.003%
TW_MEscale_down	0.042%	-0.349%	0.026%	0.160%
TW_MEscale_up	-0.063%	-0.134%	-0.161%	-0.068%
total_up	19.567%	7.066%	7.604%	9.886%
total_down	-13.846%	-6.835%	-6.275%	-10.783%

4.10 Results

Note: before going to measure the limit on EFT couplings, it is worth to mention that a cross check on measuring SM tW cross section is performed (see appendix .5) which makes sure we have understand the SM tW part.

In order to calculate the total cross sections for the $t\bar{t}$ and tW processes and generate events in the presence of new effective interactions, we implement the operators of equations 4.1 in the universal FeynRules output format [83] through the FeynRules package [84] and use the MADGRAPH5_aMC@NLO event generator [85, 86] at the leading order (LO). If we allow for the presence of one operator at a time, the total cross section up to $\mathcal{O}(\Lambda^{-4})$ can be parameterised as

$$\sigma = \sigma_{SM} + C_i \sigma_i^{(1)} + C_i^2 \sigma_i^{(2)}, \quad (4.7)$$

where the C_i 's are effective couplings introduced in equations 4.1. Here, $\sigma_i^{(1)}$ is the cross section of the interference term between the SM diagrams. The cross section $\sigma_i^{(2)}$ is the pure new physics contribution. We use the most precise available SM predictions for σ_{SM} , which are $\sigma_{SM}^{t\bar{t}} = 832_{-29}^{+20}(scales) \pm 35(PDF + \alpha_s)pb$ and $\sigma_{SM}^{tW} = 71.7_{-1.8}^{+1.8}(scales) \pm 3.4(PDF + \alpha_s)pb$ for $t\bar{t}$ and tW production, respectively [87, 88]. The scales reflects uncertainties in the factorization and renormalization scales. In the framework of EFT, the $\sigma_i^{(1)}$ and $\sigma_i^{(2)}$ terms are calculated at NLO accuracy for all of the operators except O_G . The values of $\sigma_i^{(1)}$ and $\sigma_i^{(2)}$ for various effective couplings at LO and available K-factors are given in Table 4.15.

Table 4.15: Cross sections for $t\bar{t}$ and tW production [in pb] for the various effective couplings for $\Lambda = 1$ TeV. The respective available K-factors are also shown.

Channel	Variable	C_G	$C_{\phi q}^{(3)}$	C_{tW}	C_{tG}	C_{uG}	C_{cG}
$t\bar{t}$	$\sigma_i^{(1)-LO}$	31.9	-	-	137	-	-
	$\sigma_i^{(1)-NLO}/\sigma_i^{(1)-LO}$	-	-	-	1.48 [89]	-	-
	$\sigma_i^{(2)-LO}$	102.3	-	-	16.4	-	-
	$\sigma_i^{(2)-NLO}/\sigma_i^{(2)-LO}$	-	-	-	1.44 [89]	-	-
tW	$\sigma_i^{(1)-LO}$	-	6.7	-4.5	3.3	0	0
	$\sigma_i^{(1)-NLO}/\sigma_i^{(1)-LO}$	-	1.32 [90]	1.27 [90]	1.27 [90]	0	0
	$\sigma_i^{(2)-LO}$	-	0.2	1	1.2	16.2	4.6
	$\sigma_i^{(2)-NLO}/\sigma_i^{(2)-LO}$	-	1.31 [90]	1.18 [90]	1.06 [90]	1.27 [46]	1.27 [46]

4.10.1 Limit setting procedure

For those operators which interfere with the SM, $C_G-C_{tG}-C_{\phi q}-C_{tW}$, normalization of the $t\bar{t}$ or tW process is directly extracted from a fit to data. Normalization of the signal ($tW/t\bar{t}$) is parameterized using equation 4.7 in which σ_{SM} , $\sigma_i^{(1)}$ and $\sigma_i^{(2)}$ are fixed parameters (see Table 4.15) and C is the parameter of interest in the fit. In order to evaluate the effect of the uncertainties on σ_{SM} , $\sigma_i^{(1)}$ and $\sigma_i^{(2)}$, fit is performed when these parameters are varied $\pm\sigma$ because of the Q scale uncertainties. All three terms are considered fully correlated for Q-scale variation based on the recommendation from theorists. In addition, uncertainties due to PDF is considered. Results of the mentioned variations are only shown for observed limits for comparison to the nominal results. In Table 4.16, nominal values for $\sigma_i^{(1)}$ and $\sigma_i^{(2)}$ are shown together with errors.

Table 4.16: Cross sections for $t\bar{t}$ and tW production [in pb] for the various effective couplings for $\Lambda = 1$ TeV together with Q-scale errors.

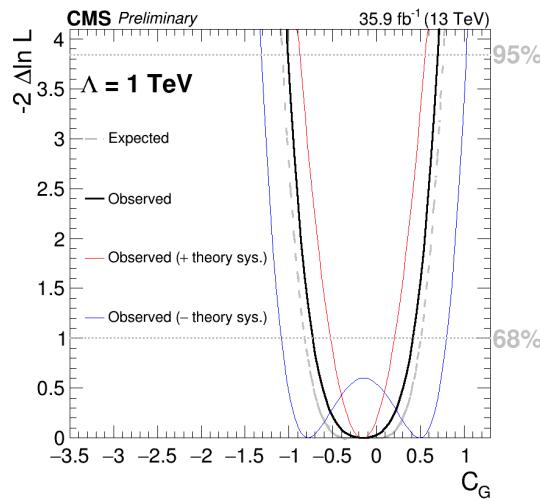
Channel	σ_{SM} (scale unc.) (PDF+alphaS unc.)
tt	831.76 (+19.77, -29.20), (+35.06, -35.06)
tW	71.7 (+1.80, -1.80), (-3.40 -3.40)
Channel	$\sigma_i^{(1)}$ (scale unc.)
tt	202.83 C_{tG} (+24.54,-26.98) , 31.9 C_G (+8.1,-6.9)
tW	8.844 $C_{\phi q}^2$ (+0,-0) ,-5.65 C_{tW} (+0.08317,-0.061846) , 4.223 C_{tG} (+0.0294,-0.0398)
Channel	$\sigma_i^{(2)}$ (scale unc.)
tt	23.545 C_{tG}^2 (+0,-0) , 102.3 C_G^2 (+22.7,-15.3)
tW	0.275 $C_{\phi q}^2$ (+0,-0) , 1.18 C_{tW}^2 (+0.0283,-0.0257) , 1.322 C_{tG}^2 (+0.0558,-0.0335) , 21.209 C_{uG}^2 (+1.485,-1.273) , 5.804 C_{cG}^2 (+0.255,-0.250)

Note: Before going the limit result it is worth mention that all these studies are also performed using $e\mu$ final states. Therefore results from $e\mu$ channel are shown for comparison and the results for ee, $e\mu$ and $\mu\mu$ channels combined are also shown.

4.10.2 Exclusion limits on C_G effective coupling

It was discussed in Section 4.1 that the operator O_G operator only contributes to $t\bar{t}$ production process. It was found that the shapes of some variables are affected in the presence of the O_G operator. On the other hand, the effect is not big enough to be observed experimentally in $t\bar{t}$ kinematic distributions as a shape effect. Therefore, the fit is performed simultaneously on the observed event yield in (1-jet,1-tag), (2-jet,1-tag) and (≥ 2 -jet,2-tag) categories for ee and $\mu\mu$ channels. The results for ee, $\mu\mu$ and $e\mu$ individual channels and combined channels are listed in Table 4.17. The results of the combined likelihood scans of the C_G coupling is shown in Figure 4.24.

	Regions	best fit exp./obs.	68% exp./obs.limit	95% exp./obs.limit
C_G	ee yield (1j1t), (2j1t), (≥ 2 j,2t)	0.00 / -0.14	[-0.90 to 0.59]/[-0.82 to 0.51]	[-1.20 to 0.88]/[-1.14 to 0.83]
	$\mu\mu$ yield (1j1t), (2j1t), (≥ 2 j,2t)	0.00 / -0.14	[-0.88 to 0.57]/[-0.75 to 0.44]	[-1.16 to 0.85]/[-1.06 to 0.75]
	$e\mu$ yield (1j0t), (1j1t), (2j1t), (≥ 2 j,2t)	0.00 / -0.18	[-0.82 to 0.51]/[-0.73 to 0.42]	[-1.08 to 0.77]/[-1.01 to 0.70]
	Combined	0.00 / -0.18	[-0.82 to 0.51]/[-0.73 to 0.42]	[-1.07 to 0.76]/[-1.01 to 0.70]

 Table 4.17: Summary of allowed 68% CL and 95% CL intervals on C_G effective coupling obtained in ee, $e\mu$, $\mu\mu$ and combined channels ($\Lambda = 1$ TeV).

 Figure 4.24: Likelihood scan of C_G effective coupling for combined channels.

4.10.3 Exclusion limits on C_{tG} , $C_{\phi q}$ and C_{tW} effective couplings

The deviation from the SM tW production from the interference terms between the SM and the O_{tG} , $O_{\phi q}$ and O_{tW} operators is of the order of $\frac{1}{\Lambda^2}$. It is assumed that the new physics scale, Λ , is larger than the scale we probe. Therefore, $\frac{1}{\Lambda^4}$ contributions, pure new physics term, would be small compared to the contribution from the interference term. Following the strategy described in Section 4.10.1, in the likelihood fit, the signal probability density function (pdf) originates from the sum of the SM term, the interference term and the pure new physics term is assumed to be the same as the SM tW (or $t\bar{t}$ for O_{tG}) pdf.

In order to set limits on the effective couplings C_{tG} , $C_{\phi q}$ and C_{tW} , we utilize the MLP output distributions for both data and MC expectation in the (1jet,1b-jet) and (2jet,1b-jet) regions and event yield in the (≥ 2 jet,2b-jet) region for ee and $\mu\mu$ channels. The MLP is trained to separate tW from $t\bar{t}$ events as was discussed in Section 4.8. The inclusion of the (≥ 2 jet,2b-jet) and (2jet,1b-jet) regions helps to constrain the normalization and systematic uncertainties of the $t\bar{t}$ background. Comparison between observed data and the SM background prediction for the MLP output shape in various jet-bjet regions are shown in Figure 4.25.

Three Wilson coefficients sensitive to new physics contribution in top quark interactions, as defined in equations 4.1 are tested in observed data. The results for individual channels and combined channels are listed in Table 4.18. The results of the combined likelihood scans of the Wilson coefficients on the full 13 TeV dataset are shown in Figure 4.26 for combined channels.

		best fit exp./obs.	68% exp./obs.limit	95% exp./obs.limit
$C_{\phi q}$	ee NN output for (1j1t+2j1t) + yields(≥ 2 j,2t)	0.00 / 1.12	[-2.53 to 1.74] / [-1.18 to 2.89]	[-6.40 to 3.27] / [-4.03 to 4.37]
	$\mu\mu$ NN output for (1j1t+2j1t) + yields(≥ 2 j,2t)	0.00 / 1.13	[-2.20 to 1.92] / [-0.87 to 2.86]	[-4.68 to 3.66] / [-3.58 to 4.46]
	$e\mu$ NN output for (1j0t+1j1t+2j1t) + yields(≥ 2 j,2t)	0.00 / -0.70	[-1.34 to 1.12] / [-2.16 to 0.59]	[-2.57 to 2.15] / [-3.74 to 1.61]
	Combined	0.00 / -1.52	[-1.05 to 0.88] / [-2.71 to -0.33]	[-2.04 to 1.63] / [-3.82 to 0.63]
C_{tW}	ee NN output for (1j1t+2j1t)+yields(≥ 2 j,2t)	0.00 / 6.18	[-2.02 to 6.81] / [-3.02 to 7.81]	[-3.33 to 8.12] / [-4.16 to 8.95]
	$\mu\mu$ NN output for (1j1t+2j1t)+yields(≥ 2 j,2t)	0.00 / -1.40	[-2.18 to 6.97] / [-3.00 to 7.79]	[-3.63 to 8.42] / [-4.23 to 9.01]
	$e\mu$ NN output for (1j0t+1j1t+2j1t)+yields(≥ 2 j,2t)	0.00 / 1.64	[-1.40 to 6.19] / [-0.80 to 5.59]	[-2.39 to 7.18] / [-1.89 to 6.68]
	Combined	0.00 / 2.38	[-1.14 to 5.93] / [0.22 to 4.57]	[-1.91 to 6.70] / [-0.96 to 5.74]
C_{tG}	ee NN output for (1j1t+2j1t)+yields(≥ 2 j,2t)	0.00 / -0.19	[-0.22 to 0.21] / [-0.40 to 0.02]	[-0.44 to 0.41] / [-0.65 to 0.22]
	$\mu\mu$ NN output for (1j1t+2j1t)+yields(≥ 2 j,2t)	0.00 / -0.15	[-0.19 to 0.18] / [-0.34 to 0.02]	[-0.40 to 0.35] / [-0.53 to 0.19]
	$e\mu$ NN output for (1j0t+1j1t+2j1t)+yields(≥ 2 j,2t)	0.00 / -0.03	[-0.17 to 0.15] / [-0.19 to 0.11]	[-0.34 to 0.29] / [-0.34 to 0.27]
	Combined	0.00 / -0.13	[-0.15 to 0.14] / [-0.27 to 0.02]	[-0.30 to 0.28] / [-0.41 to 0.17]

Table 4.18: Summary of allowed 68% CL and 95% CL intervals on C_{tG} , $C_{\phi q}$ and C_{tW} effective couplings obtained in ee , $e\mu$, $\mu\mu$ and combined channels ($\Lambda = 1 TeV$).

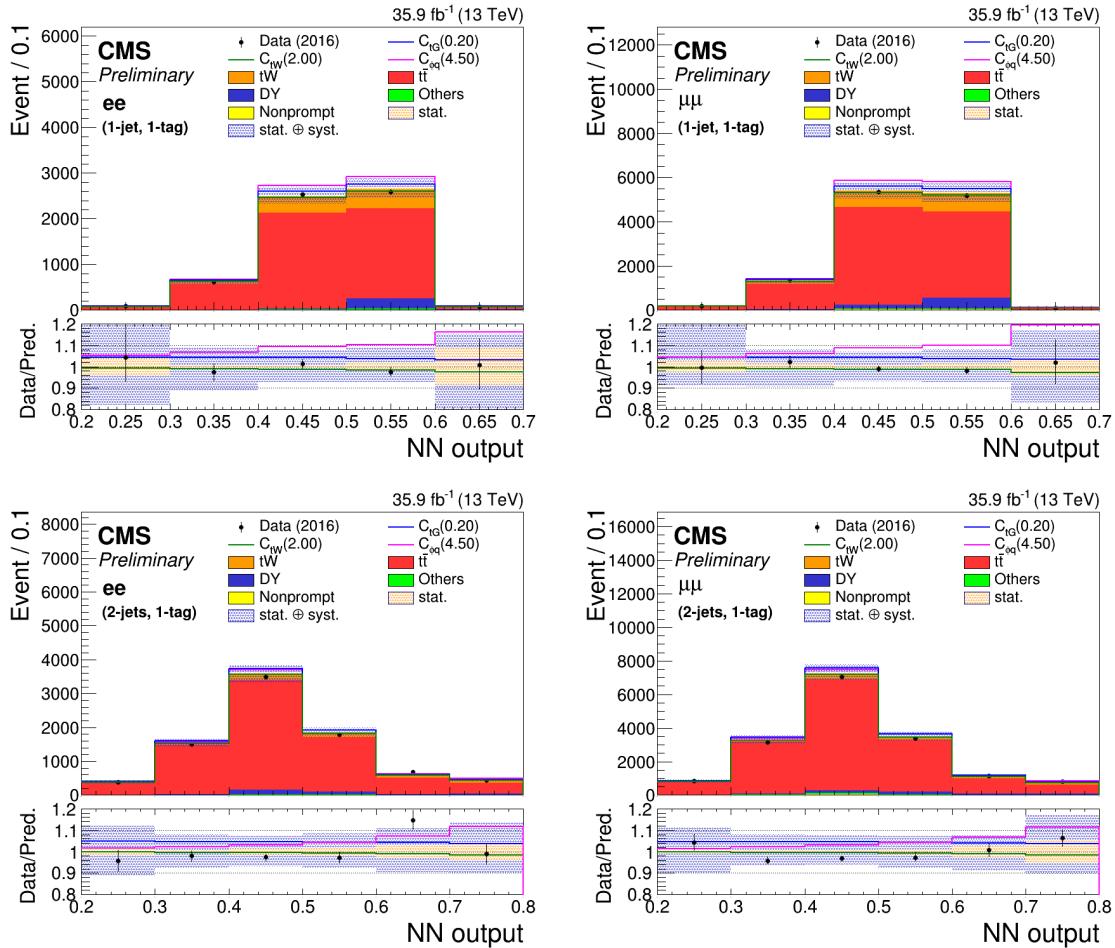


Figure 4.25: The MLP distribution of data and MC in different regions: (1jet,1b-jet) (top row), (2jet,1b-jet) (bottom row) used in limit setting for ee channel (left column) and $\mu\mu$ channel (right column). The blue hatched bands correspond to the sum of statistical and systematic uncertainties in the event yield for the sum of signal and background predictions. The ratios of data to the sum of the predicted yields are shown at the bottom of each plot. Here, an additional solid yellow band represents the contribution from the statistical uncertainty in the MC simulation.

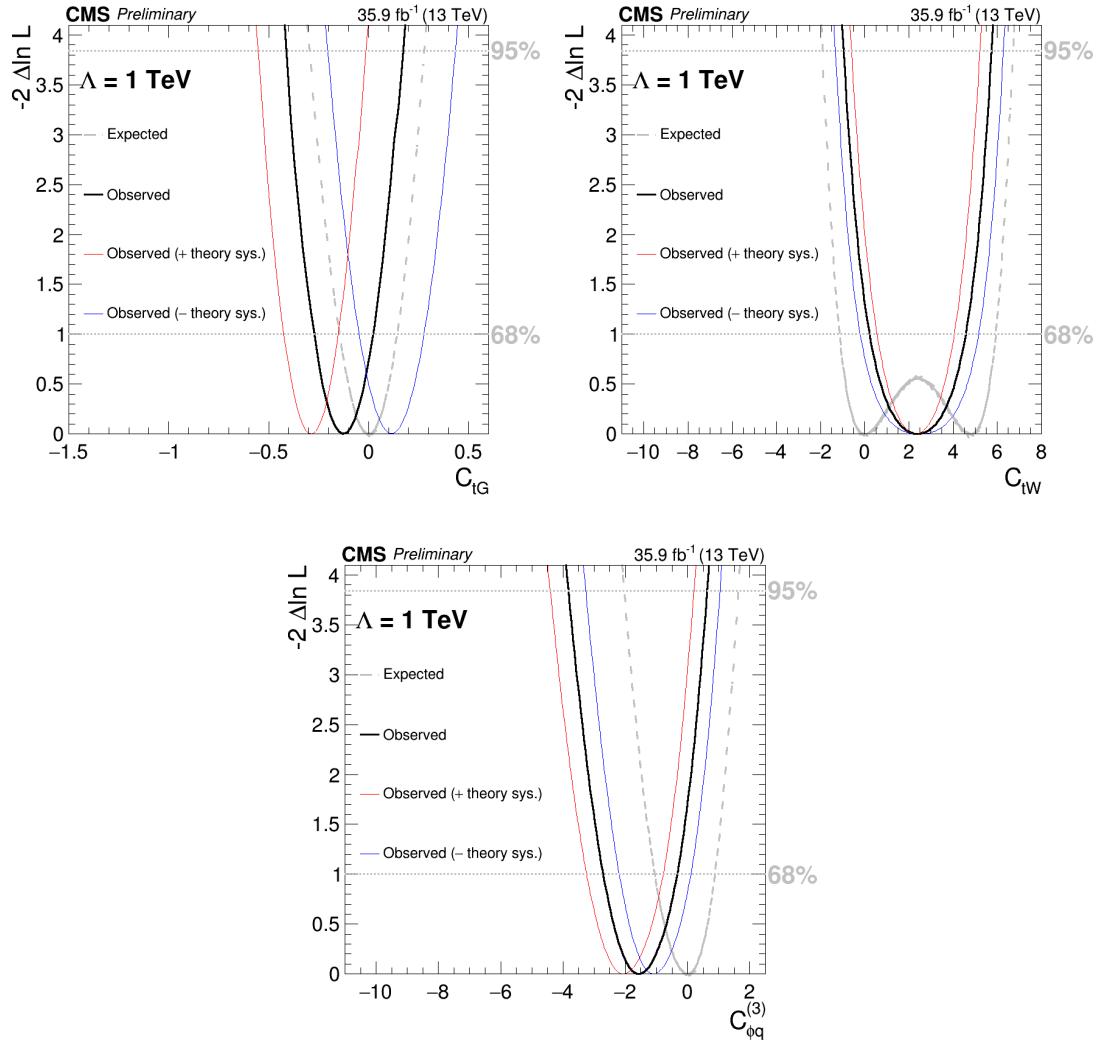


Figure 4.26: Likelihood scans of C_{tG} , $C_{\phi q}$ and C_{tW} effective couplings for combined channels.

4.10.4 Exclusion limits on C_{uG} and C_{cG} effective couplings

Since the tW production via FCNC interactions doesn't interfere with the SM (with the assumption of $|V_{td}| = |V_{ts}| = 0$), independent pdf for signal is considered to set upper bound on related Wilson coefficients. The comparison of the MLP output for the data, SM background and signal (tW events via FCNC interactions) in 1b-jet region are shown in Figure 4.27. Here the MLP is trained to separate FCNC tW events from SM tW and $t\bar{t}$ events as discussed in Section 4.8.

The limit results for individual channels and combined channels are listed in Table 4.19. The results of the likelihood scans of the Wilson coefficients on the full 13 TeV dataset are shown in Figure 4.28 for combined channels. The observed and median expected 95% CL upper limits on the $\sigma(pp \rightarrow tW) \times B(W \rightarrow \ell\nu)^2$ for FCNC signals are given for combined channel in Table 4.20.

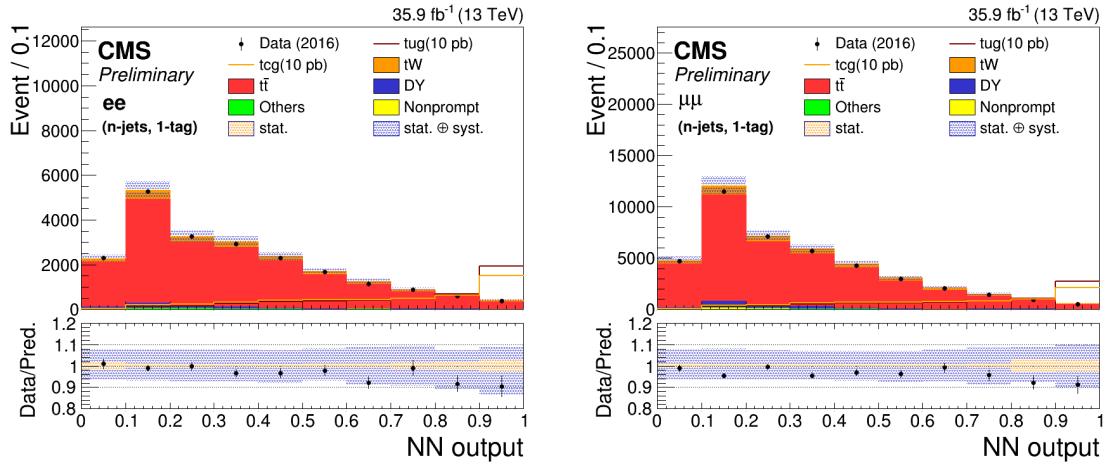


Figure 4.27: The MLP distribution of data, MC and FCNC signals in 1b-jet region used in limit setting for ee channel (left column) and $\mu\mu$ channel (right column). The blue hatched bands correspond to the sum of statistical and systematic uncertainties in the event yield for the sum of signal and background predictions. The ratios of data to the sum of the predicted yields are shown at the bottom of each plot. Here, an additional solid yellow band represents the contribution from the statistical uncertainty in the MC simulation.

		best fit exp./obs.	68% exp./obs.limit	95% exp./obs.limit
C_{uG}	ee NN output for (1t)	0.00 / -0.017	[-0.29 to 0.29] / [-0.224 to 0.224]	[-0.42 to 0.42] / [-0.368 to 0.368]
	$\mu\mu$ NN output for (1t)	0.00 / -0.017	[-0.27 to 0.27] / [-0.167 to 0.167]	[-0.38 to 0.38] / [-0.289 to 0.289]
	$e\mu$ NN output for (1t)	0.00 / -0.017	[-0.26 to 0.26] / [-0.167 to 0.167]	[-0.38 to 0.38] / [-0.290 to 0.290]
	Combined	0.00 / -0.017	[-0.21 to 0.21] / [-0.125 to 0.125]	[-0.30 to 0.30] / [-0.221 to 0.221]
C_{cG}	ee NN output for (1t)	0.00 / -0.032	[-0.63 to 0.63] / [-0.471 to 0.471]	[-0.92 to 0.92] / [-0.778 to 0.778]
	$\mu\mu$ NN output for (1t)	0.00 / -0.032	[-0.58 to 0.58] / [-0.363 to 0.363]	[-0.84 to 0.84] / [-0.628 to 0.628]
	$e\mu$ NN output for (1t)	0.00 / -0.032	[-0.56 to 0.56] / [-0.341 to 0.341]	[-0.81 to 0.81] / [-0.599 to 0.599]
	Combined	0.00 / -0.032	[-0.46 to 0.46] / [-0.259 to 0.259]	[-0.65 to 0.65] / [-0.464 to 0.464]

Table 4.19: Summary of allowed 68% CL and 95% CL intervals on C_{uG} and C_{cG} effective coupling obtained in ee , $e\mu$, $\mu\mu$ and combined channels ($\Lambda = 1TeV$).

The expected and observed upper limits on the Wilson coefficients obtained from the combination of all channels and signal regions are visualized in Figure 4.29.

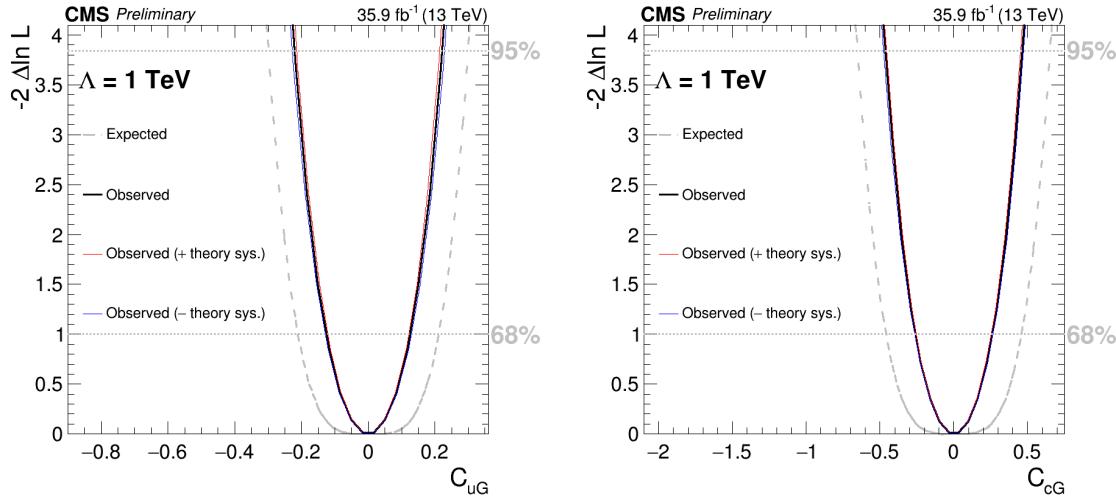


Figure 4.28: Likelihood scans of C_{uG} and C_{cG} effective couplings for combined channels.

	68% exp./obs.limit	95% exp./obs.limit
$\sigma(pp \rightarrow tW) \times B(W \rightarrow \ell\nu)^2$	[0,0.10] / [0,0.03] pb	[0,0.20] / [0,0.11] pb
$C_{uG}(\Lambda = 1 \text{ TeV})$	[-0.21,0.21] / [-0.13,0.13]	[-0.30,0.30] / [-0.22,0.22]
$B(t \rightarrow ug)$	[0,0.10918%] / [0,0.03897%]	[0,0.22068%] / [0,0.12136%]
$\sigma(pp \rightarrow tW) \times B(W \rightarrow \ell\nu)^2$	[0,0.13] / [0,0.04] pb	[0,0.26] / [0,0.13] pb
$C_{cG} (\Lambda = 1 \text{ TeV})$	[-0.46,0.46] / [-0.26,0.26]	[-0.65,0.65] / [-0.46,0.46]
$B(t \rightarrow cg)$	[0,0.51612%] / [0,0.16617%]	[0,1.05509%] / [0,0.53367%]

Table 4.20: The expected and observed 95% CL upper limits on the cross section of tW production via C_{uG} and C_{cG} effective couplings times square of the branching fraction $B(W \rightarrow \ell\nu)$, the effective couplings C_{uG} and C_{cG} , and the corresponding branching fractions $B(t \rightarrow ug)$ and $B(t \rightarrow cg)$.

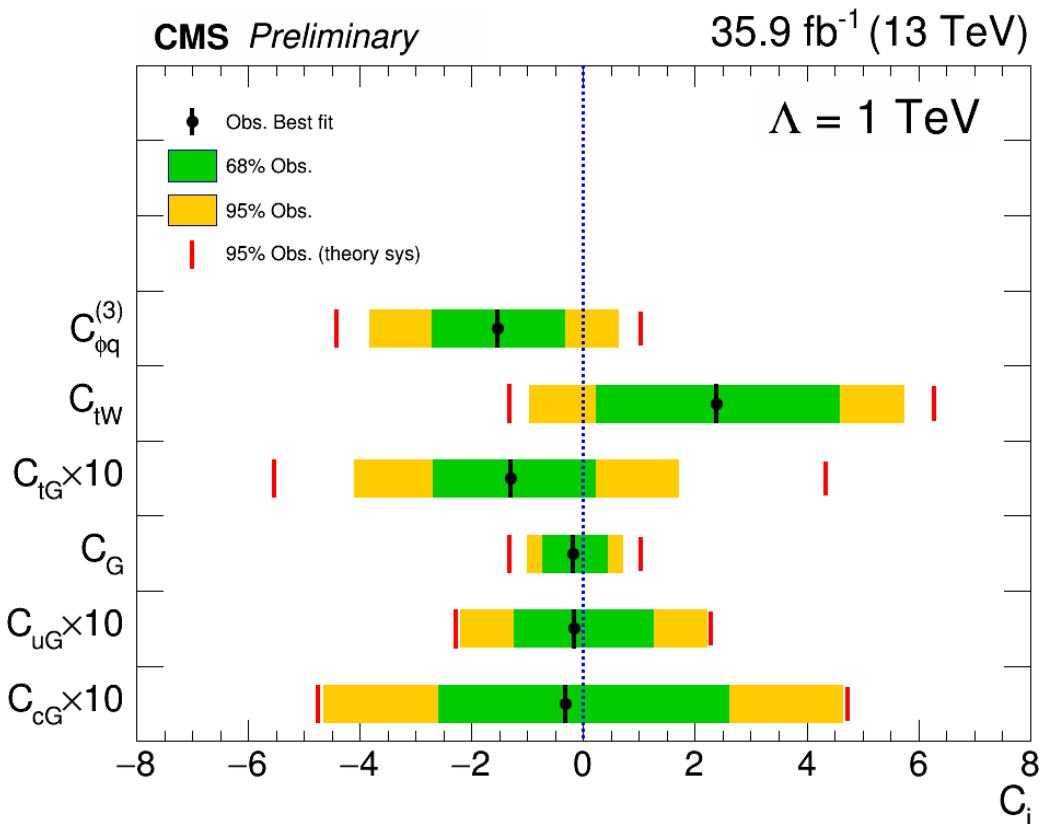


Figure 4.29: Observed and expected 95% CL upper limits on the top quark effective couplings for combined channel ($\Lambda = 1 \text{ TeV}$).

4.11 Summary

A search for new physics via top quark production in ee and $\mu\mu$ final state has been performed using data recorded by CMS in 2016 from proton-proton collisions at $\sqrt{s} = 13$ TeV with 35.9 fb^{-1} integrated luminosity. The single top quark production in association with a W boson is probed for the first time together with the top quark pair production to find the new physics signatures. The results of the analysis have been also compared and combined with those of the analogous search in the $e\mu$ final state. Observations are in agreement with standard model expectations. Limit interval at 68% and 95% confidence level on the effective couplings have been derived. The results are interpreted to constrain the relevant effective couplings using a dedicated multivariate analysis. The observed 95% CL limit band on effective couplings are found to be $[-1.01, 0.70]$ for C_G , $[-0.41, 0.17]$ for C_{tG} , $[-0.96, 5.74]$ for C_{tW} and $[-3.82, 0.63]$ for $C_{\phi q}^{(3)}$. The corresponding expected limits are $[-1.07, 0.76]$ for C_G , $[-0.30, 0.28]$ for C_{tG} , $[-1.91, 6.70]$ for C_{tW} and $[-2.04, 1.63]$ for $C_{\phi q}^{(3)}$. For the FCNC effective couplings, the observed limits are $[-0.22, 0.22]$ and $[-0.46, 0.46]$ for C_{uG} and C_{cG} , respectively; the expected limit being $[-0.30, 0.30]$ and $[-0.65, 0.65]$. The extracted values give the first experimental bound on the C_G coupling and improve upon limits previously obtained at 8 TeV for C_{tG} . The limits obtained on the C_{tW} , $C_{\phi q}^{(3)}$, C_{uG} and C_{cG} couplings from the tW process are complementary to the limits from the single top t-channel process.

My personal contributions to the analysis include performing all the steps of the analysis and cross check of tW cross section measurement. In addition I am responsible to calculate all the systematical uncertainties. Moreover, calculating the limit is my key work in this analysis.

This results is in publications by the CMS collaboration with one PAS documents [91].

Appendices

.1 Mass resolution fit results

As described in 3.4, for 2017 the quantity $\frac{(M_{reco}-M_{gen})}{M_{gen}}$ is fitted in bins of M_{gen} using a double-sided crystal ball (dCB) function. The fits results are shown for the different mass points in fig. 30 for the Barrel-Barrel category (BB) and in fig. 31 for the Barrel-Endcap (BE) category.

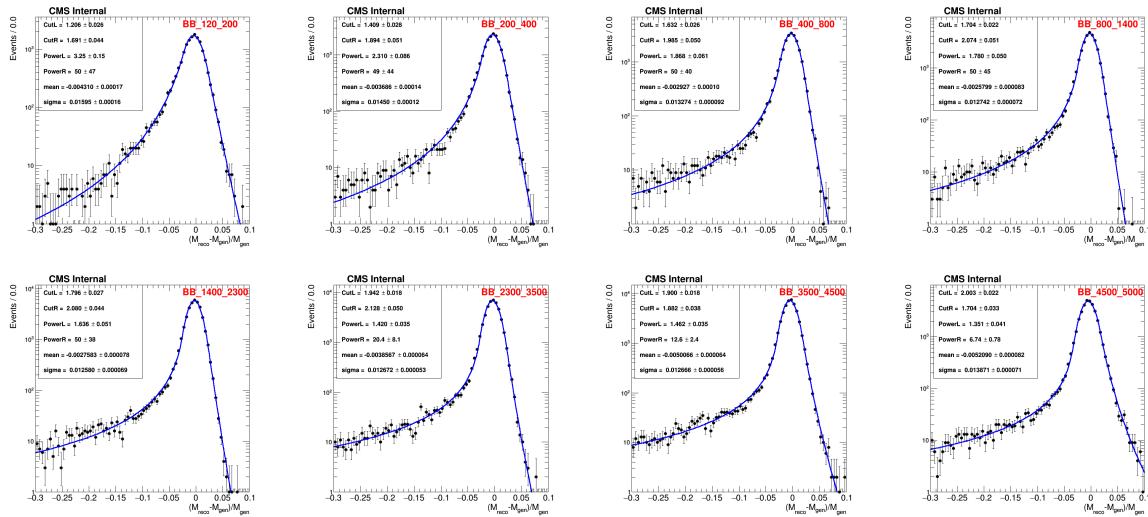


Figure 30: Fit results of the $\frac{(M_{reco}-M_{gen})}{M_{gen}}$ histograms in the BB category per M_{gen} bins.

The parameters of the fitted dCB functions are then drawn as functions of M_{gen} and a fit is superimposed in order to get an analytic description of their behaviour (see fig. 32 and 33).

At this point, in the limit setting procedure, the mass resolution is treated as a dCB function whose parameters are described by the analytic functions derived from 32 and 33.

.2 For 2016 HEEP ID scale factor

.2.1 N-1 (or N-2, N-3) efficiency for HEEP variables

In order to find variables which cause the HEEP efficiency drop, one can look at N-1 or N-2 efficiencies and scale factors for various HEEP variables. In 2016 the N-1, N-2 and N-3 efficiencies and scale factors are shown as functions of E_T , η of the probe and of N_{vtx} in figures 34-44.

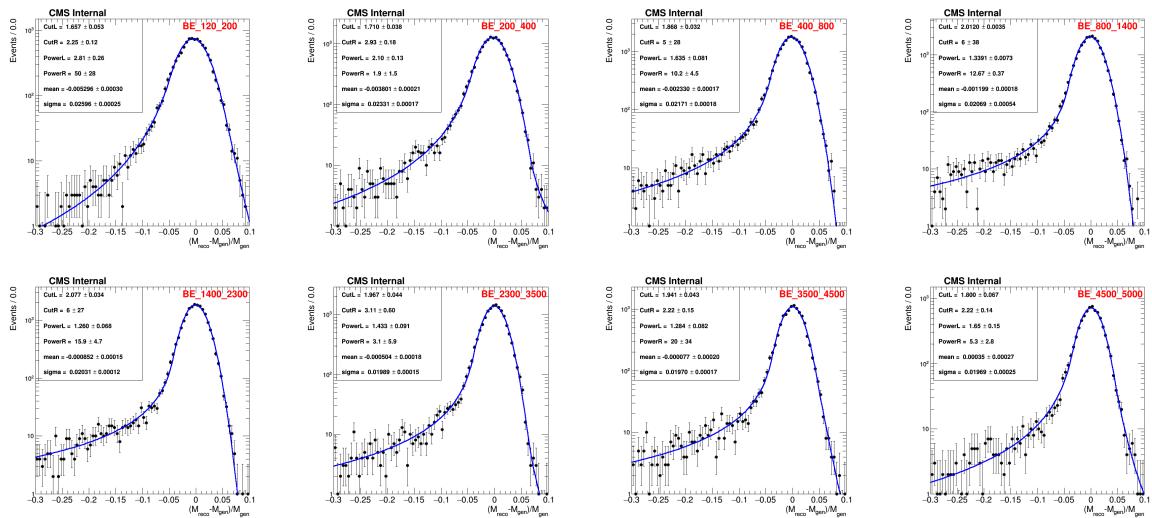


Figure 31: Fit results of the $\frac{(M_{reco} - M_{gen})}{M_{gen}}$ histograms in the BE category per M_{gen} bins.

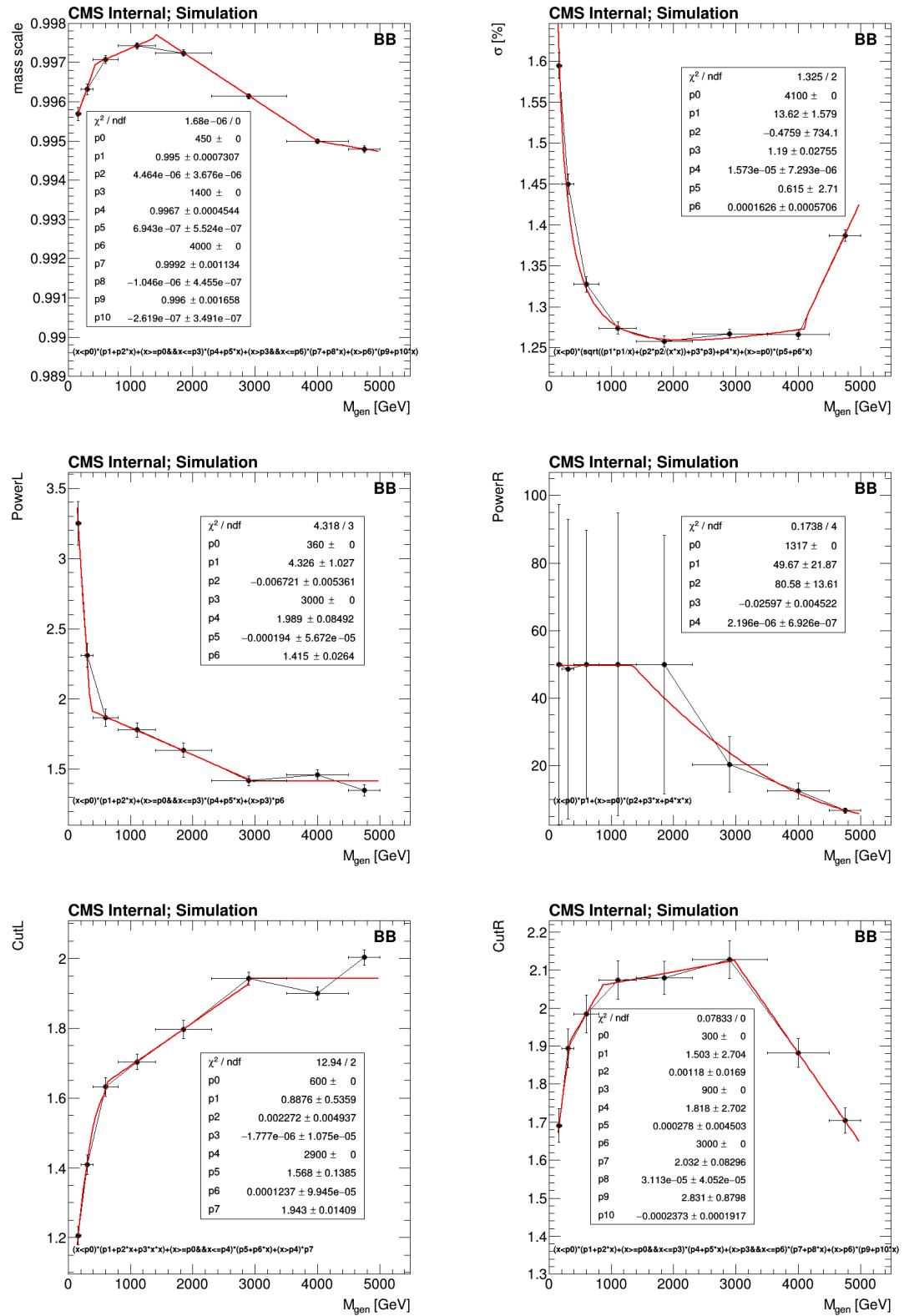


Figure 32: Parameters of the dCB as a function of M_{gen} for BB category.

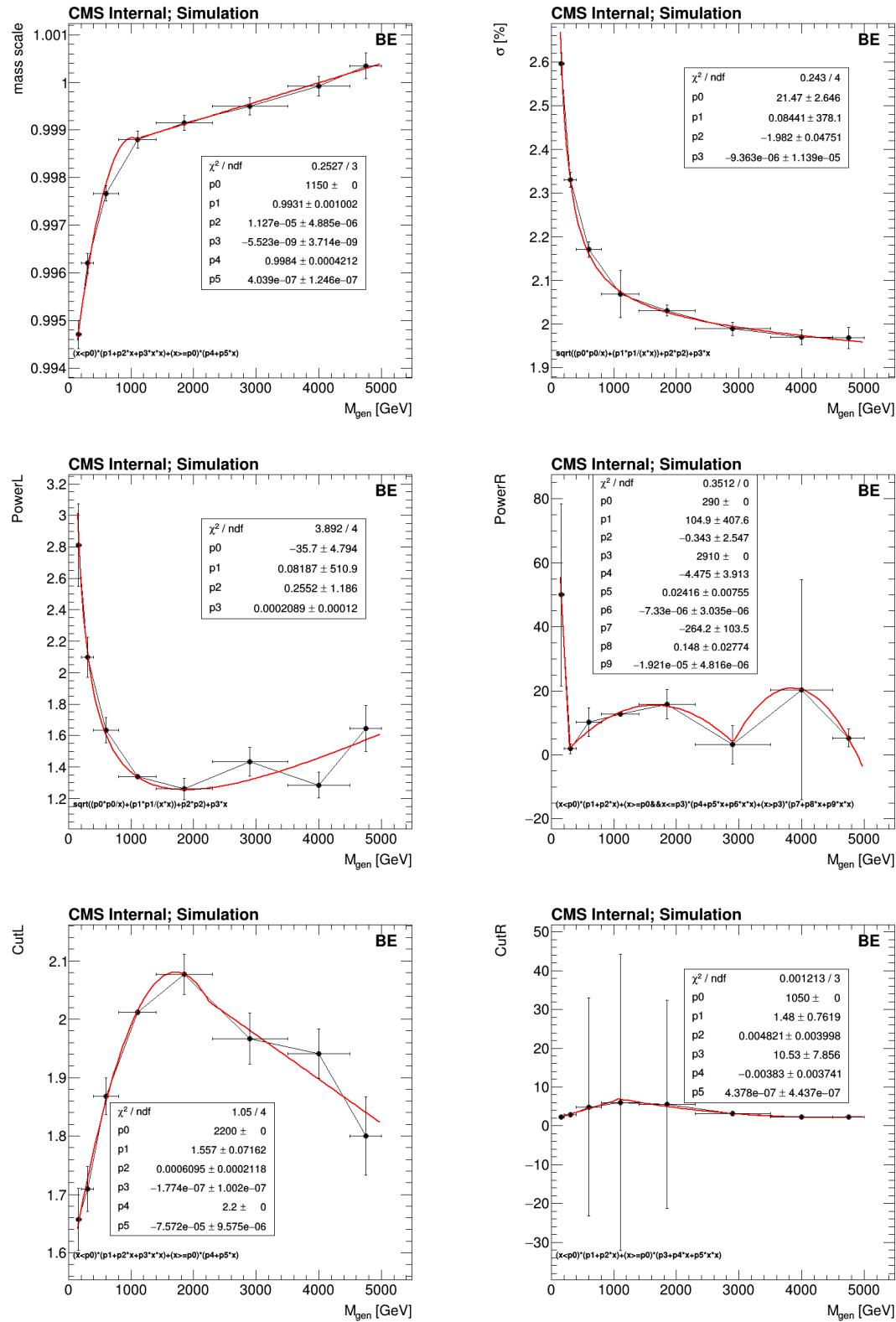


Figure 33: Parameters of the dCB as a function of M_{gen} for BE category.

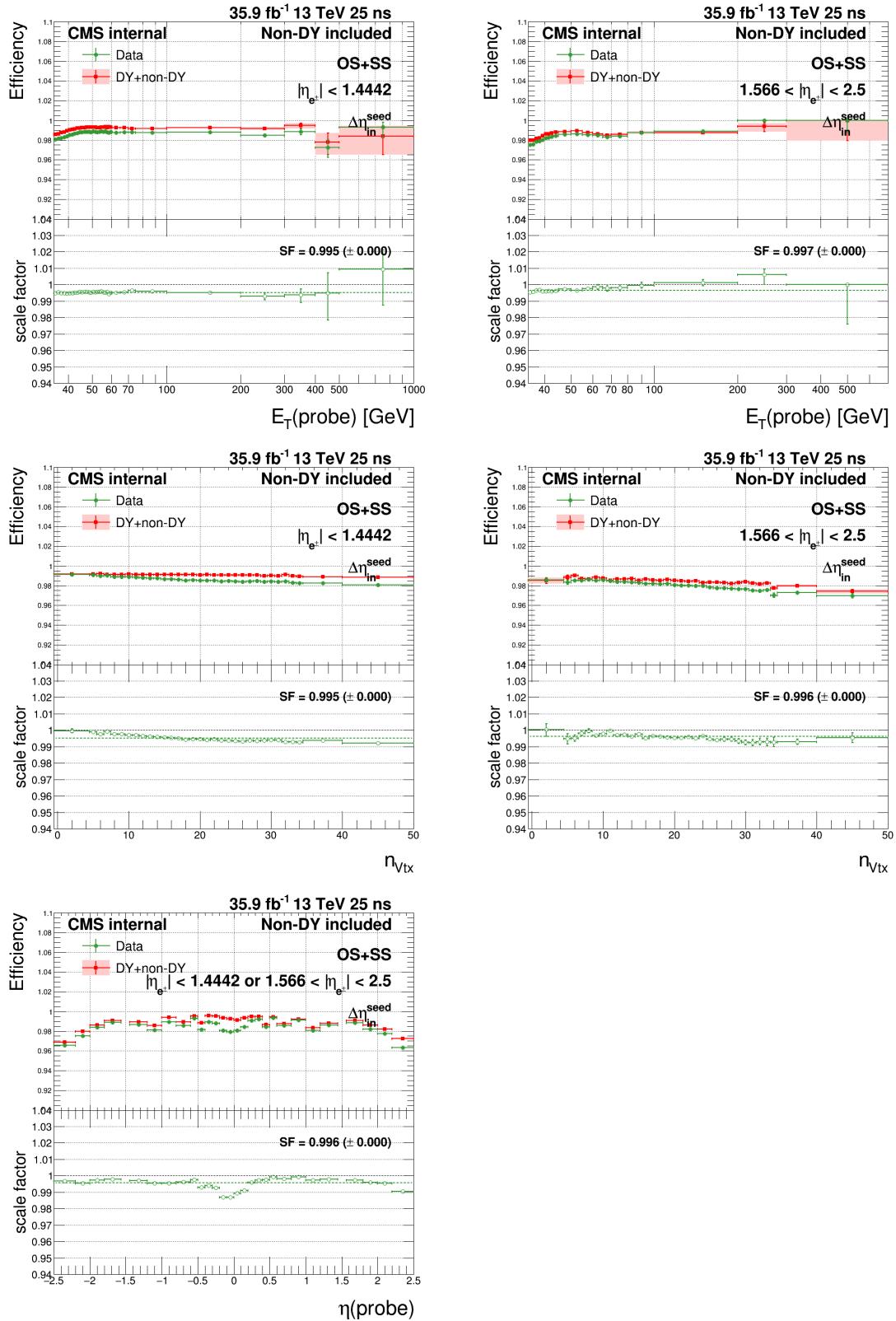


Figure 34: $\Delta\eta_{in}^{seed}$ N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

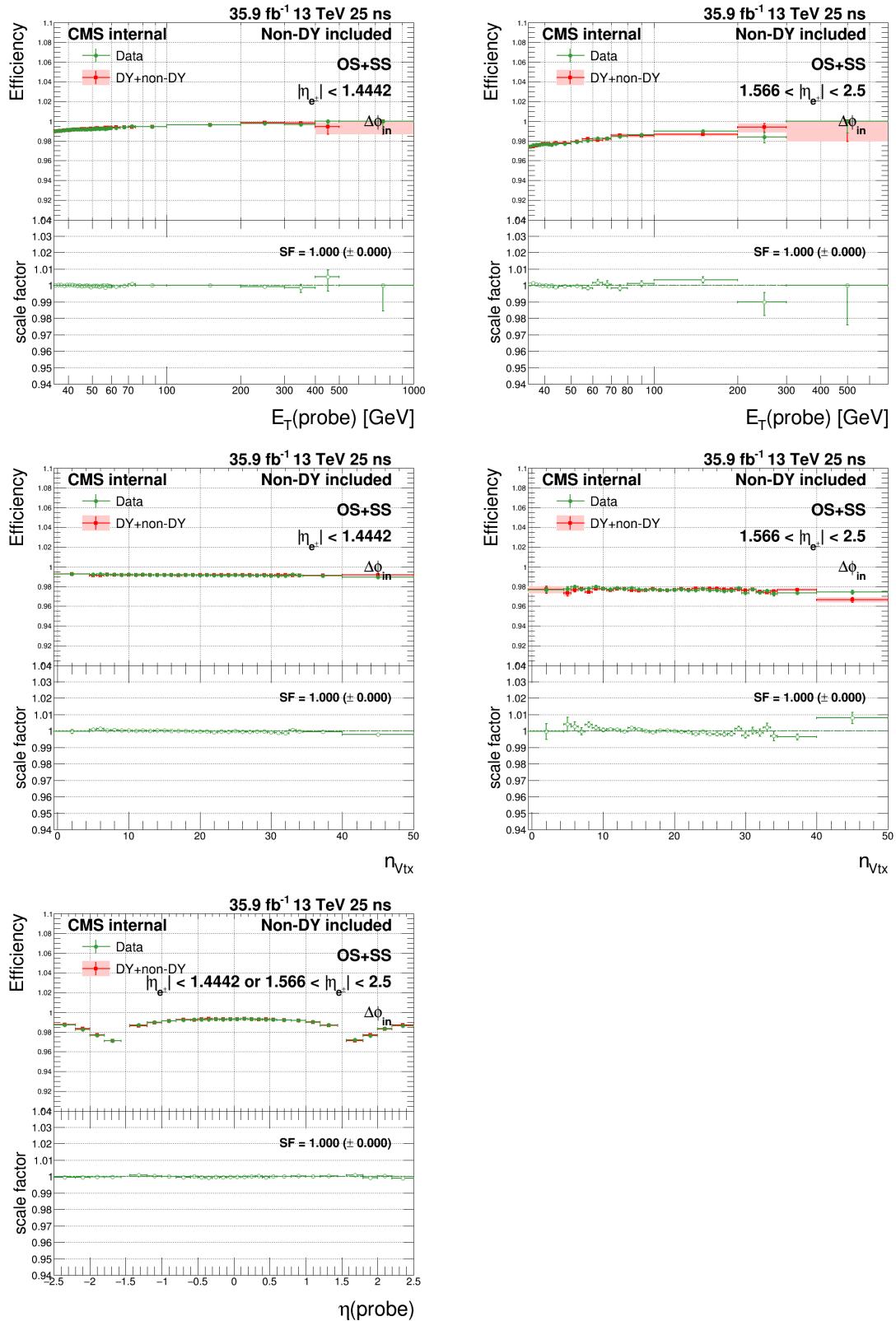


Figure 35: $\Delta\phi_{in}$ N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

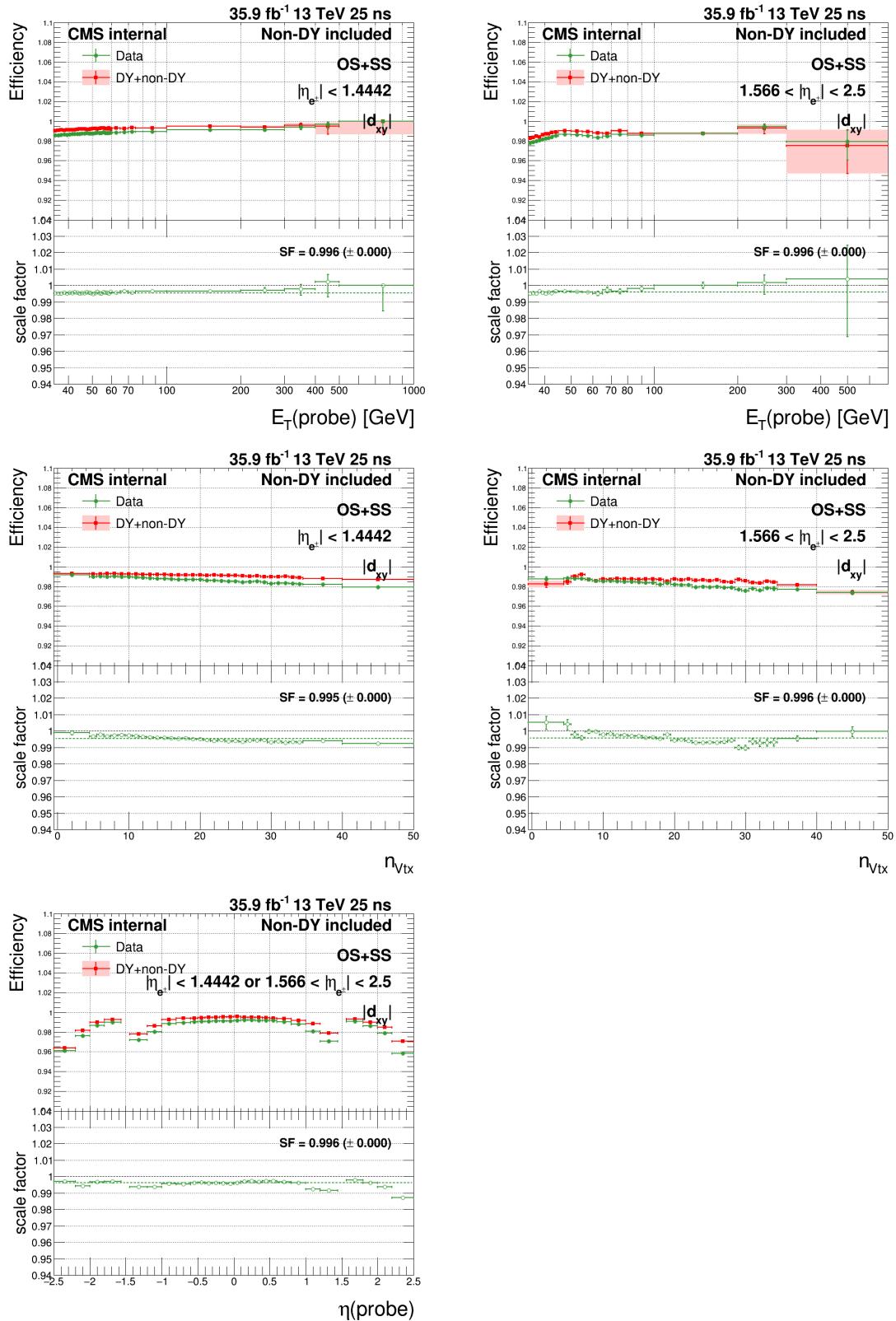


Figure 36: $|d_{xy}|$ N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

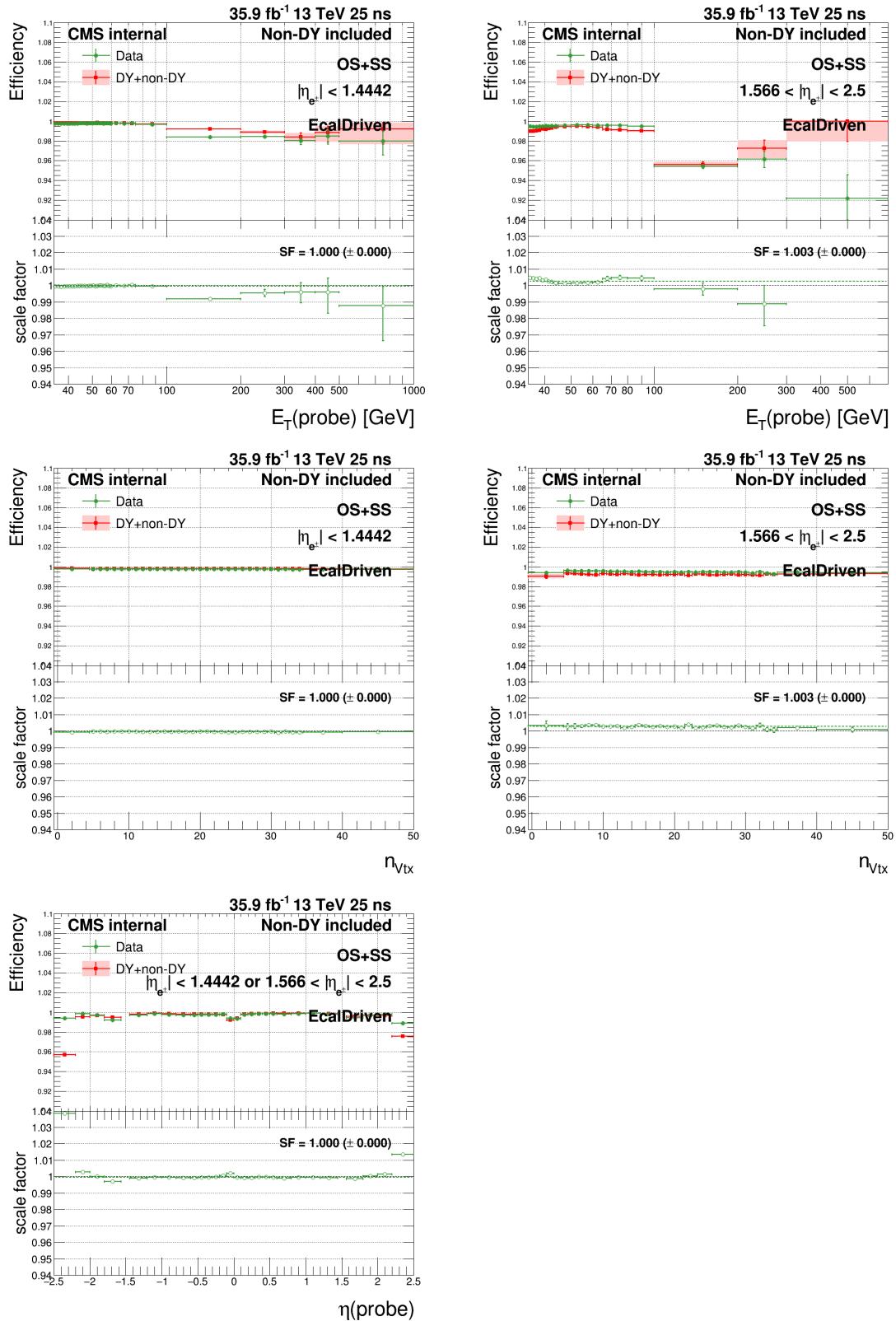


Figure 37: *EcalDriven* N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

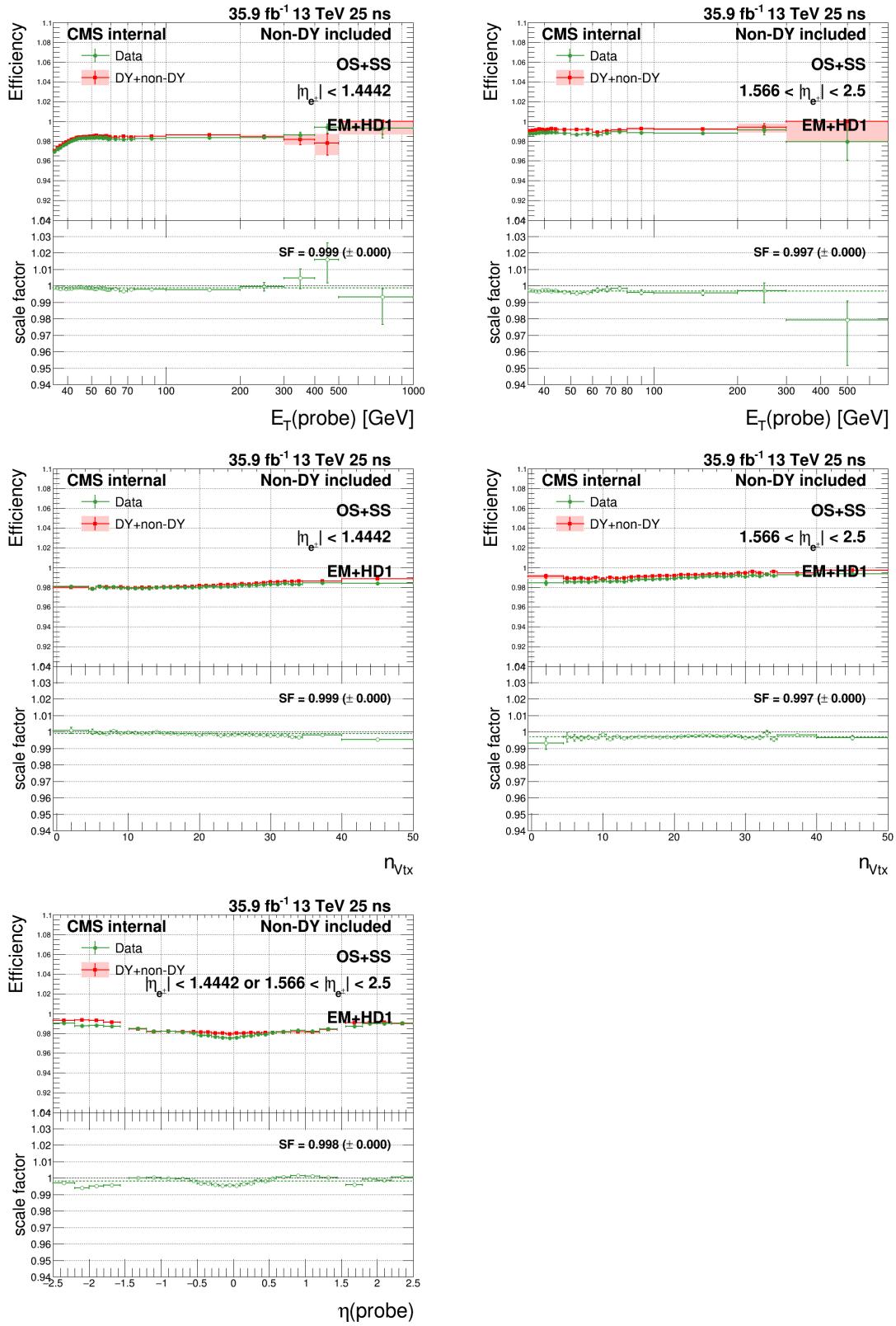


Figure 38: $EM + HD1$ N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

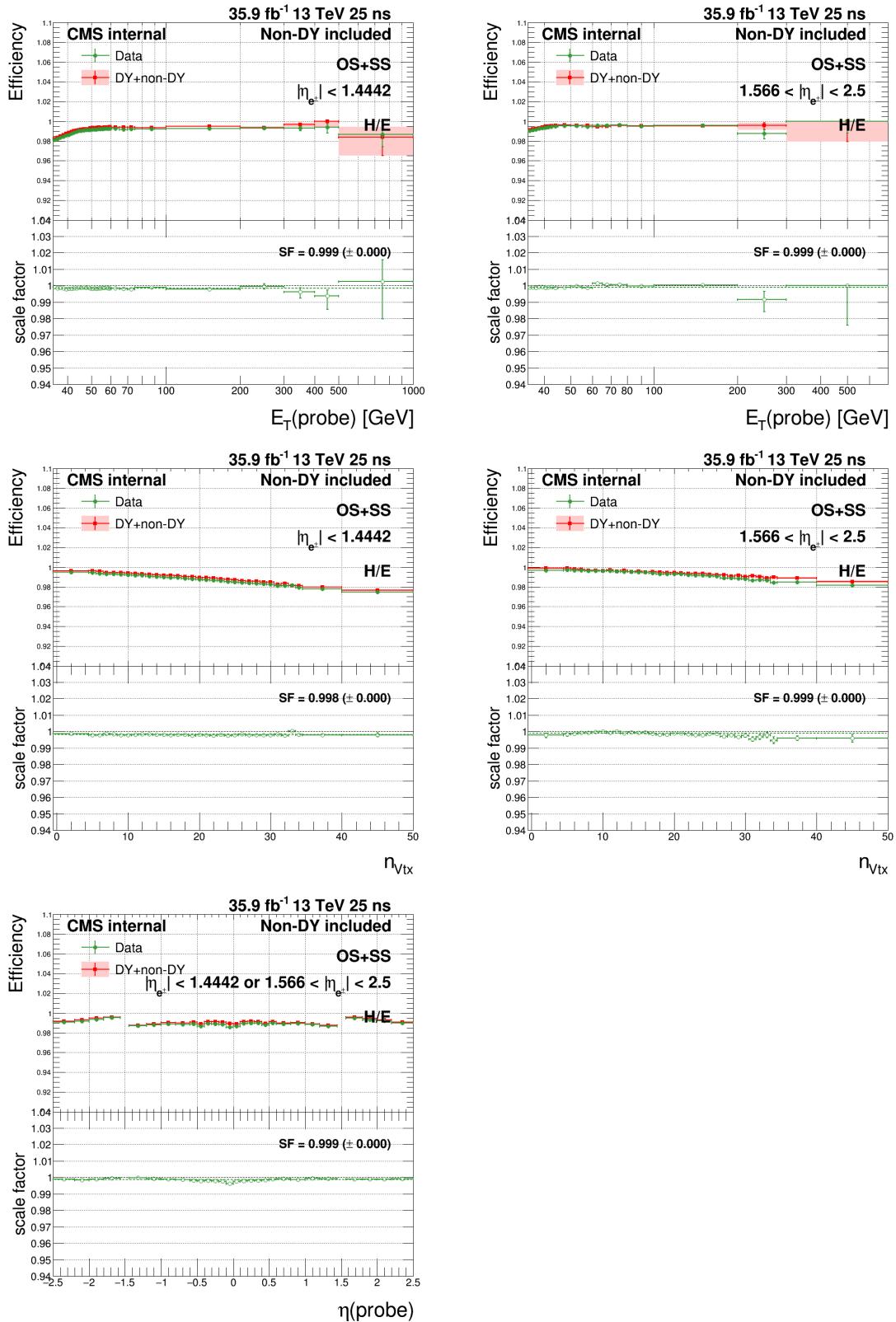


Figure 39: H/E N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

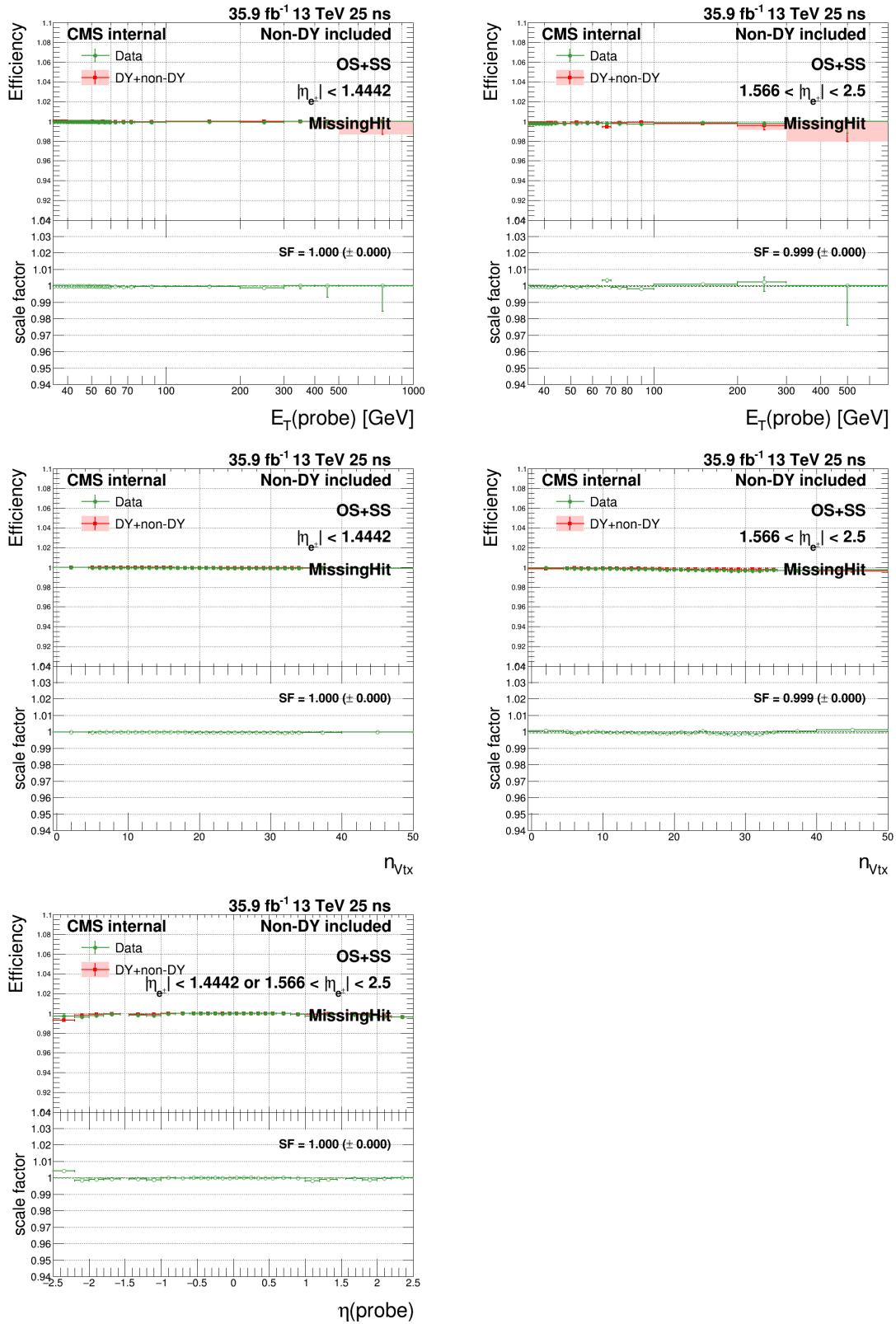


Figure 40: *MissingHit* N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

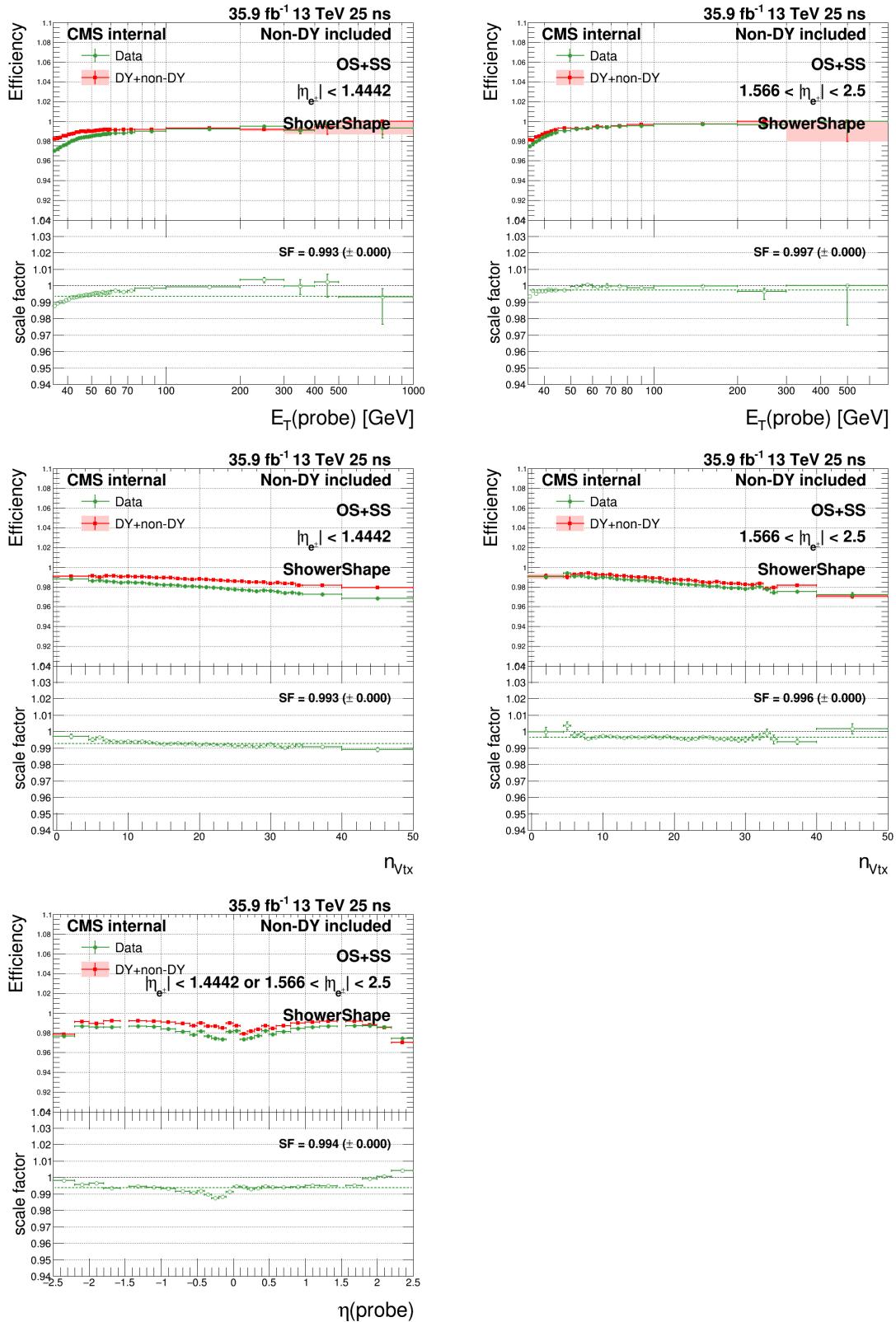


Figure 41: *ShowerShape* N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

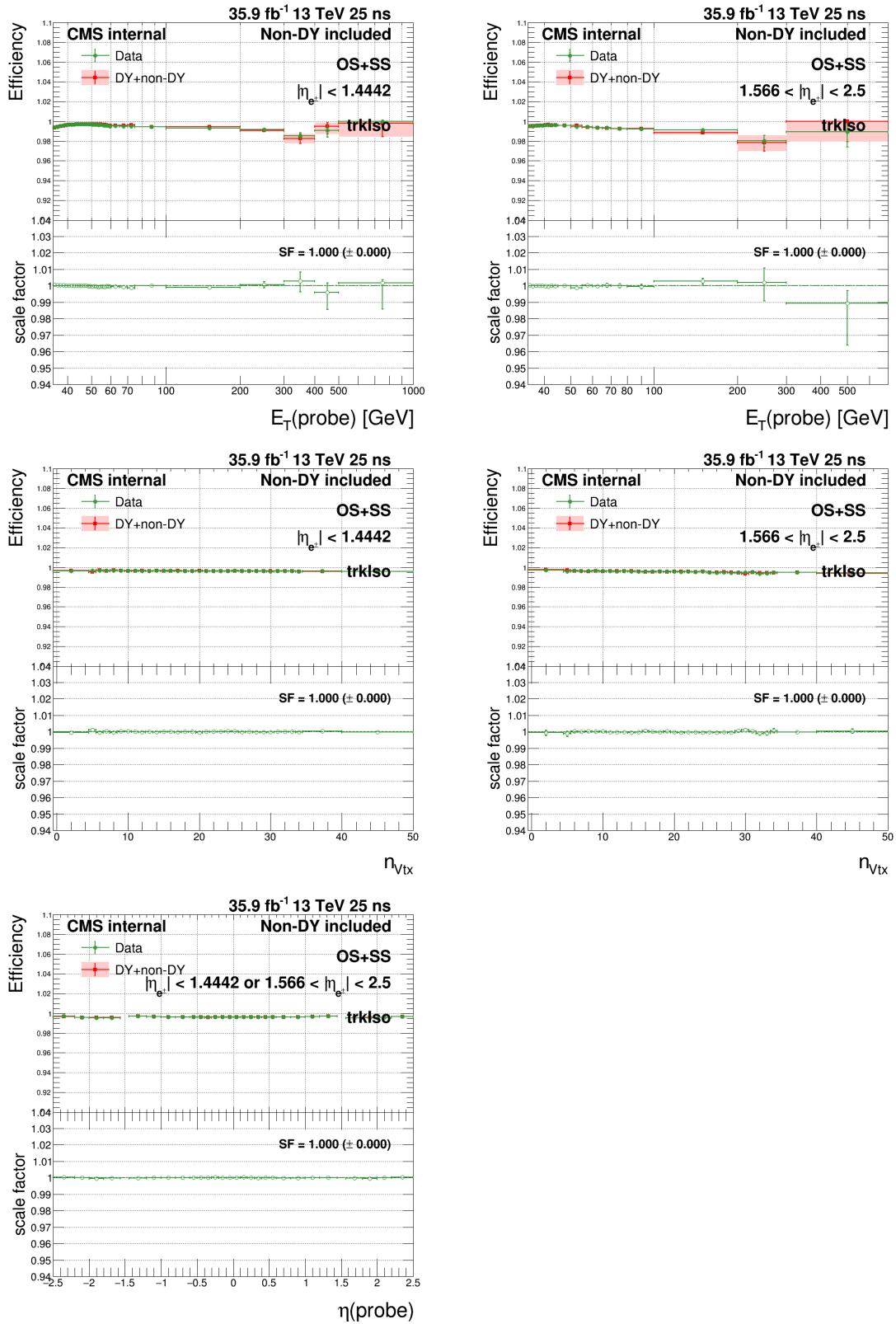


Figure 42: *trackisolation* N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

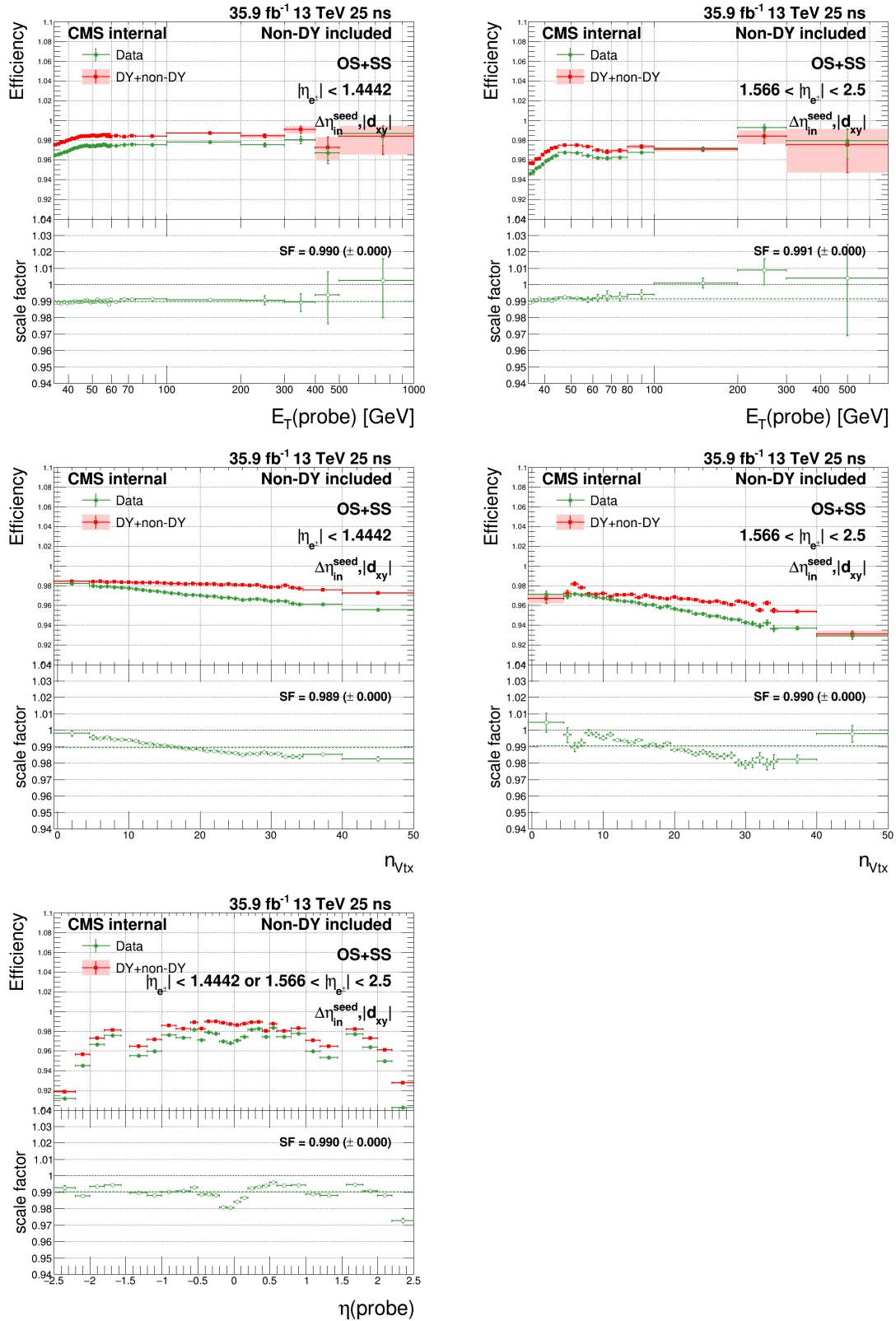


Figure 43: $\Delta\eta_{in}^{\text{seed}}$ and $|d_{xy}|$ N-2 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

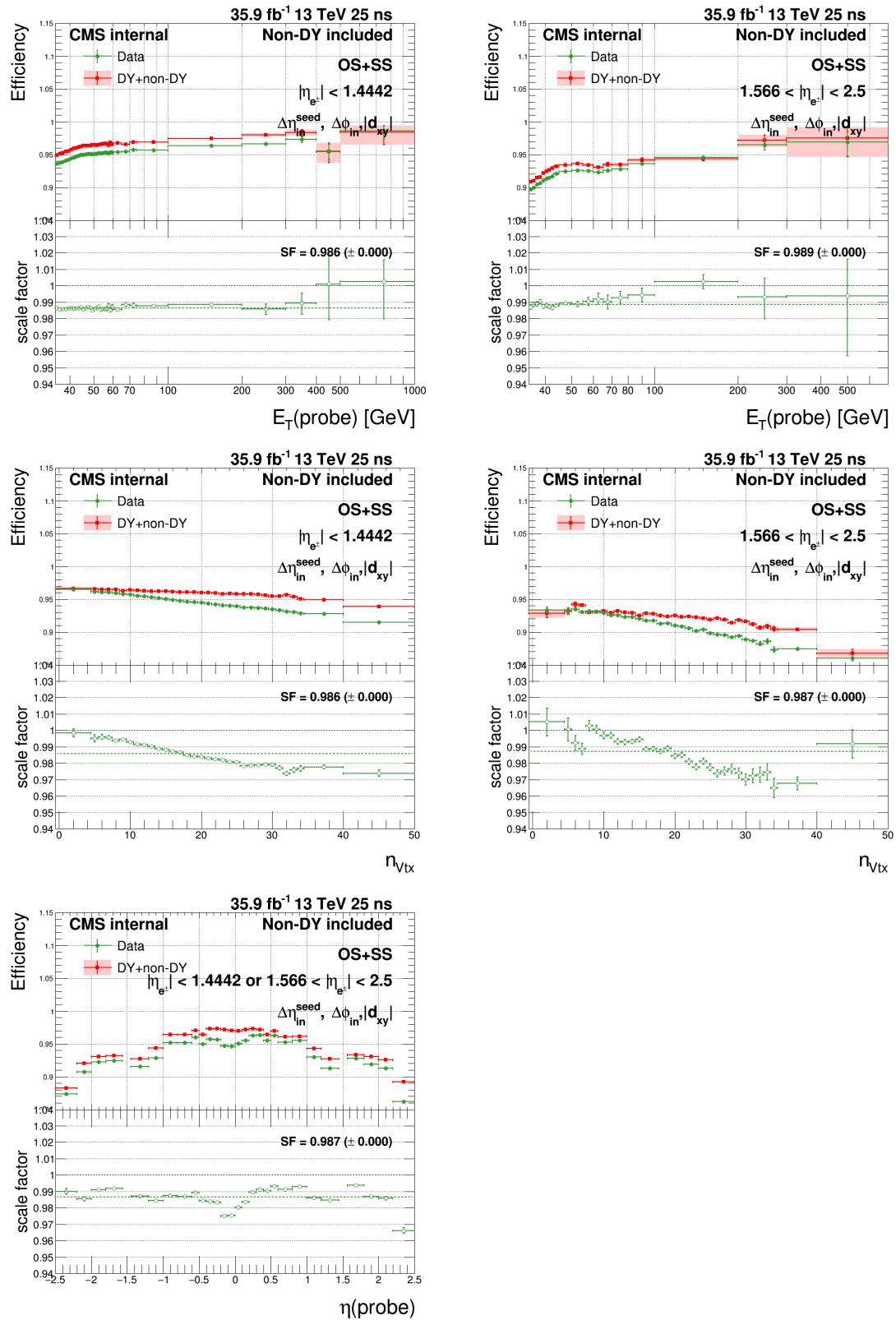


Figure 44: $\Delta\eta_{in}^{seed}$, $|d_{xy}|$ and $\Delta\phi_{in}$ N-3 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

2.2 HEEP efficiency versus η for different E_T bins

As can be seen in Figure 3.28, scale factor drops by 2-3% around $|\eta| = 0$. The efficiencies and scale factors are shown as functions of η for E_T of probe from 35–50(GeV), 50 – 100(GeV) and > 100(GeV) in Figure 45. One can see the behaviour of scale factor for η close to 0 for different E_T bins are similar, so it is E_T independent.

In addition, HEEP scale factor is measured in three η bins and summarized in Table 21

Table 21: HEEP scale factors for different η bins.

η	−2.5 to −1.566	−1.4442 to −0.5	−0.5 to 0
Scale factor	$0.985 \pm 0.001(\text{stat.})$	$0.971 \pm 0.001(\text{stat.})$	$0.961 \pm 0.001(\text{stat.})$
η	0 to 0.5	0.5 to 1.4442	1.566 to 2.5
Scale factor	$0.974 \pm 0.001(\text{stat.})$	$0.978 \pm 0.001(\text{stat.})$	$0.982 \pm 0.002(\text{stat.})$

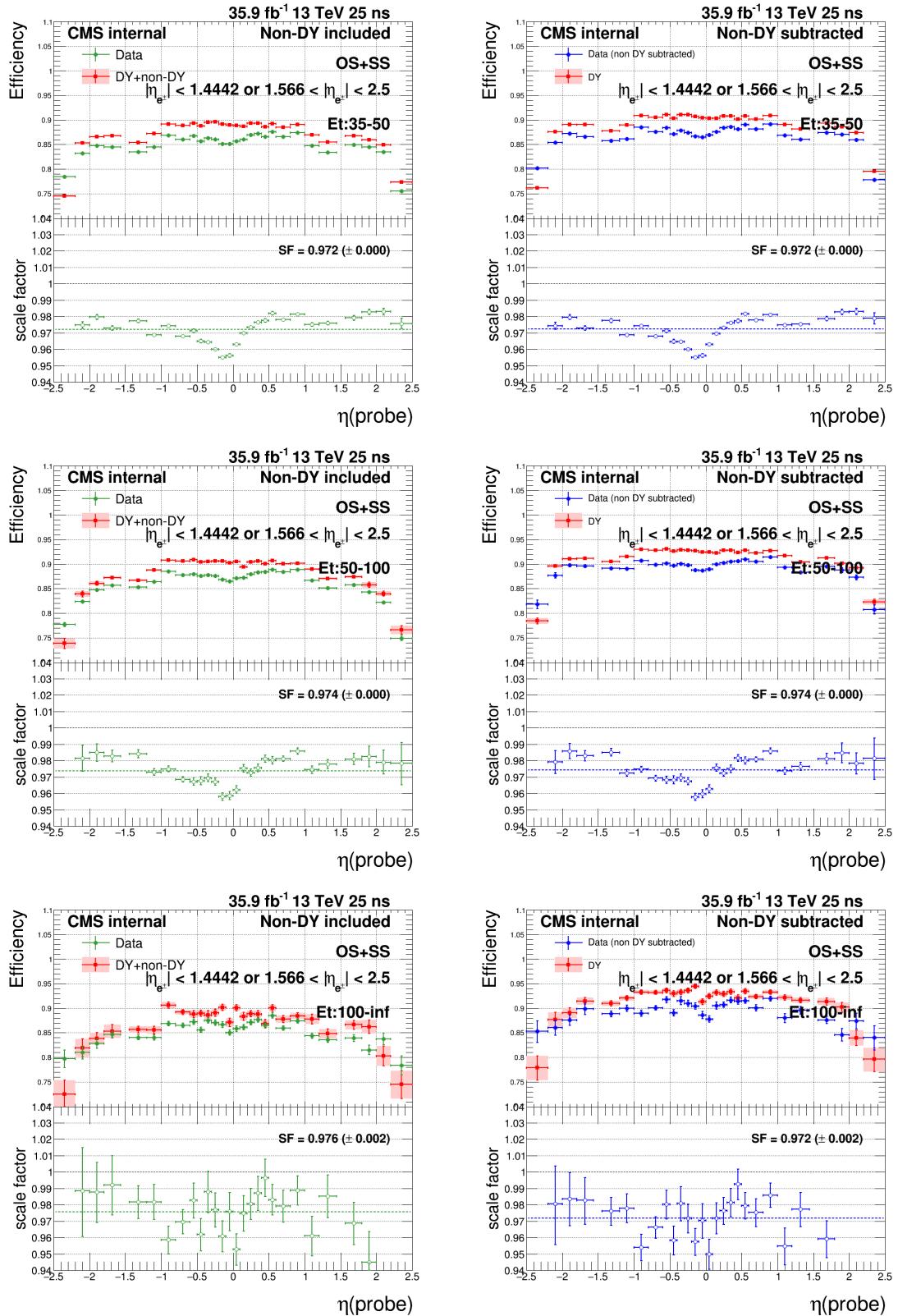


Figure 45: Efficiencies and scale factors in MC and data for E_T of probe bin 35 – 50(GeV) (top), 50 – 100(GeV) (middle) and > 100 (GeV) (bottom) where the non-DY processes are included (left) and subtracted (right) as functions of probe η .

2.3 Cross check with DYJetsToLL amcatnlo sample

Here we cross checked the HEEP efficiency and scale factor using DYJetsToLL amcatnlo sample samples. The E_T of probes are shown in Figure 46. The HEEP efficiency and scale factor for different E_T are shown in figures 47.

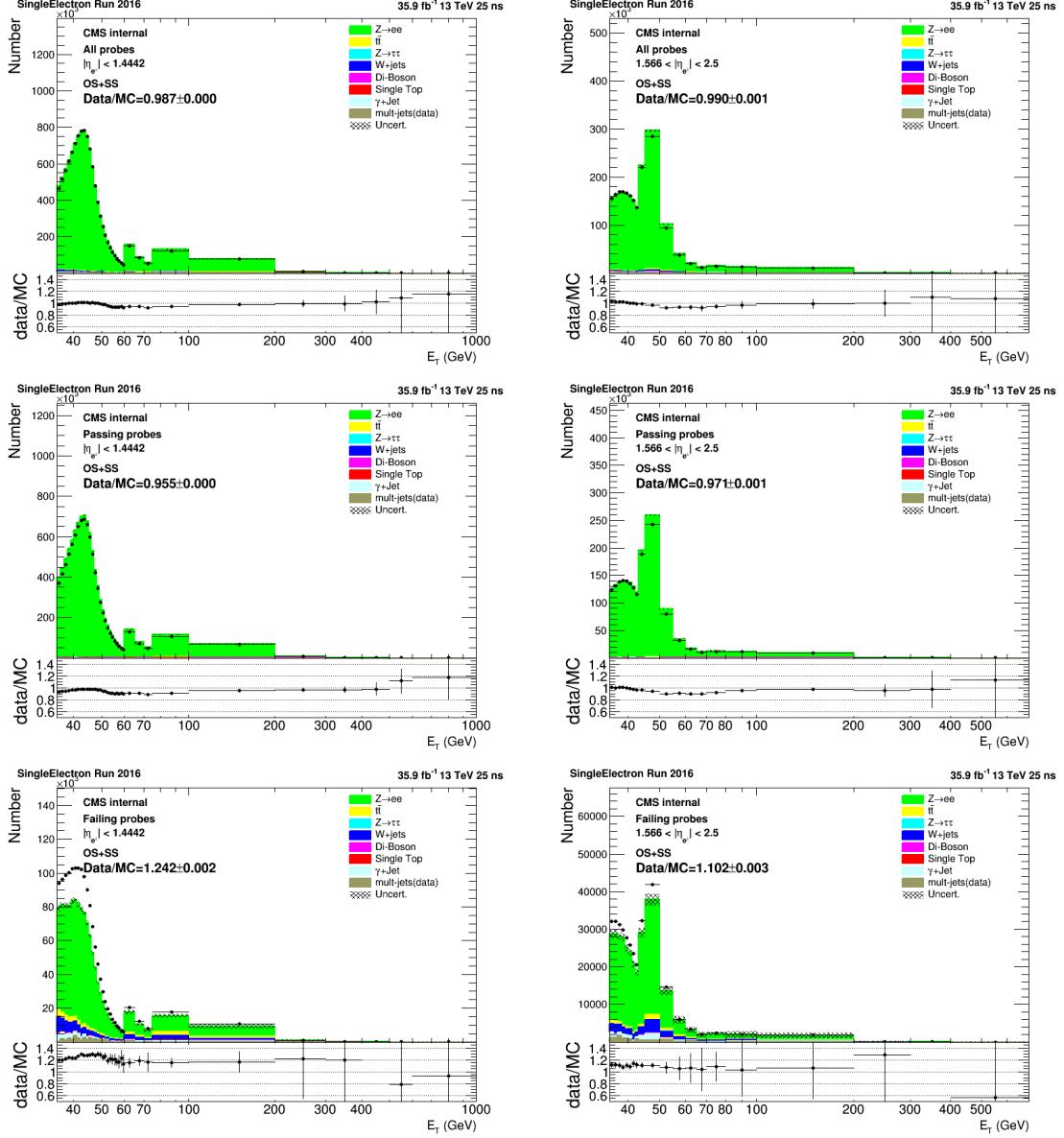


Figure 46: E_T of probe in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle) and only failed probes are included (bottom).

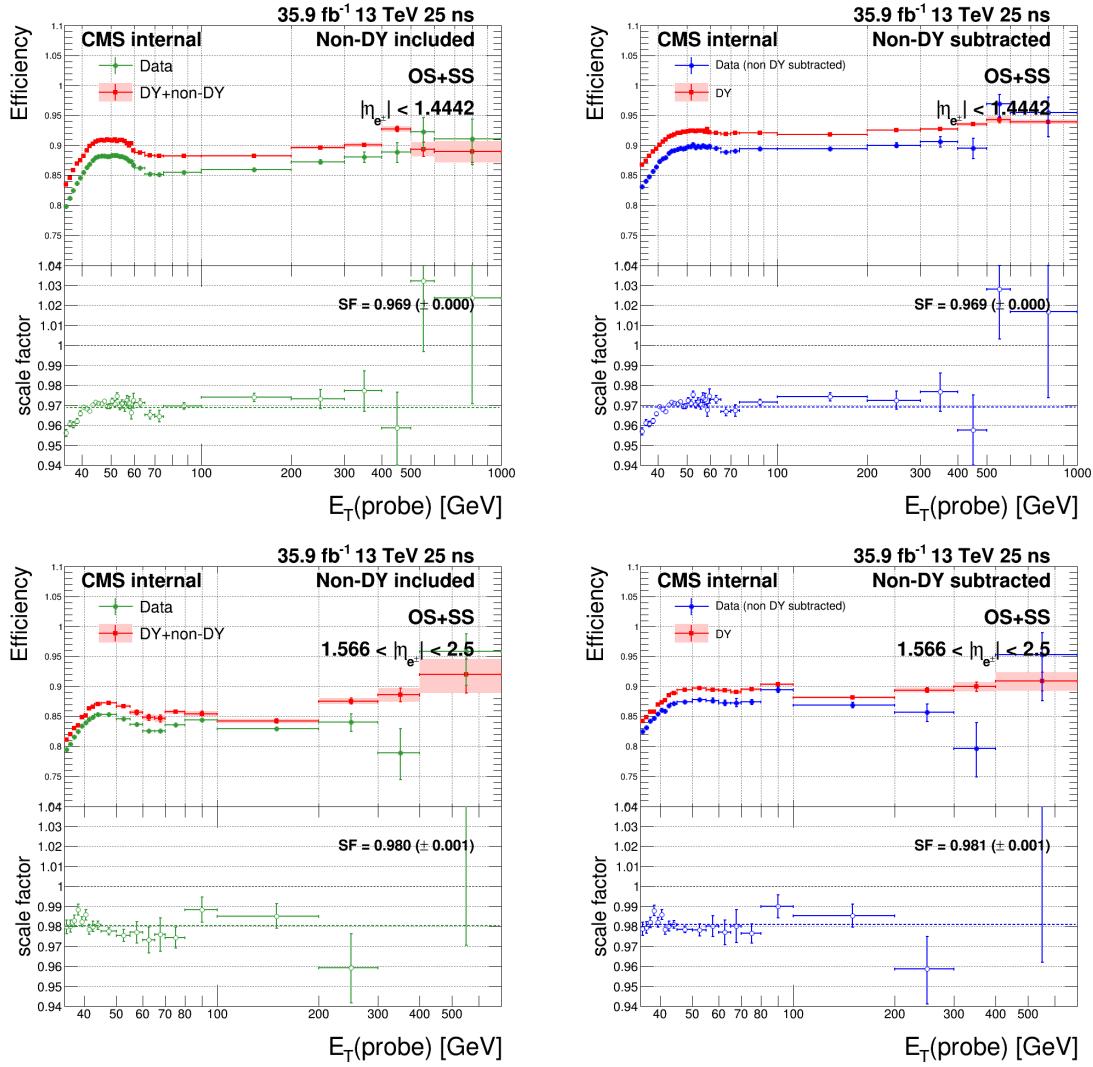


Figure 47: Efficiencies and scale factors in MC and data in the barrel (top) and endcap (bottom) where the non-DY processes are included (left) and subtracted (right) as functions of probe E_T .

2.4 HEEP efficiency for mc matched electron for different DY samples

Here we checked the HEEP efficiency using mc matched electron ($\Delta R < 0.1$ between gsf electron and mc electron) for DYToEE powheg, ZToEE mass bin powheg, DYJetsToLL amcatnlo and DYJetsToLL madgraph samples versus E_T of mc electron. Besides we ask the minimum ΔR spacing for two generated electron to be 0.5. The results are shown in figures 48-50. From Figure 48 one can see for high E_T the DYToEE and ZToEE agree well but for low E_T there are small difference because of the different global tag for these two samples. From Figure 49 one can see for low E_T DYToLL madgraph and ZToEE powheg agree well but for high E_T there are small difference. From Figure 50 one can see ZToEE mass bin powheg and DYJetsToLL amcatnlo agree well.

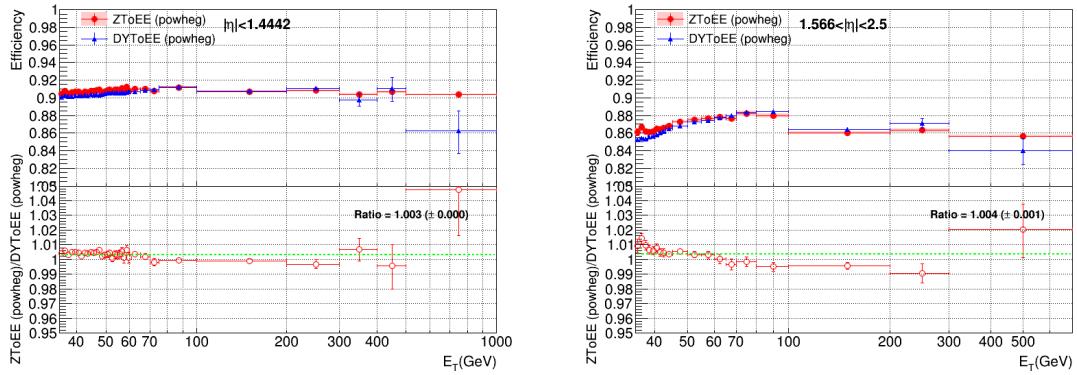


Figure 48: The HEEP efficiencies for mc matched electron for DYToEE powheg sample and ZToEE powheg sample for barrel (left) and endcap (right) for different generated E_T of electron.

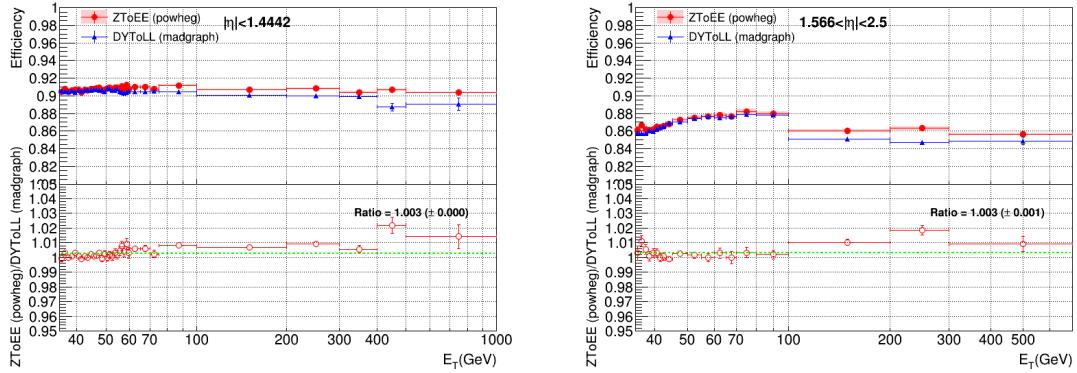


Figure 49: The HEEP efficiencies for mc matched electron for DYToLL madgraph sample and ZToEE powheg sample for barrel (left) and endcap (right) for different generated E_T of electron.

3 For 2017 HEEP ID scale factor

3.1 N-1 (or N-2, N-3) efficiency for HEEP variables

In order to find variables which cause the HEEP efficiency drop, one can look at N-1 or N-2 efficiencies and scale factors for various HEEP variables. In 2017 N-1, N-2 and N-3 efficiencies and scale factors are shown as functions of E_T , η of the probe and of N_{vtx} in figures 51-61.

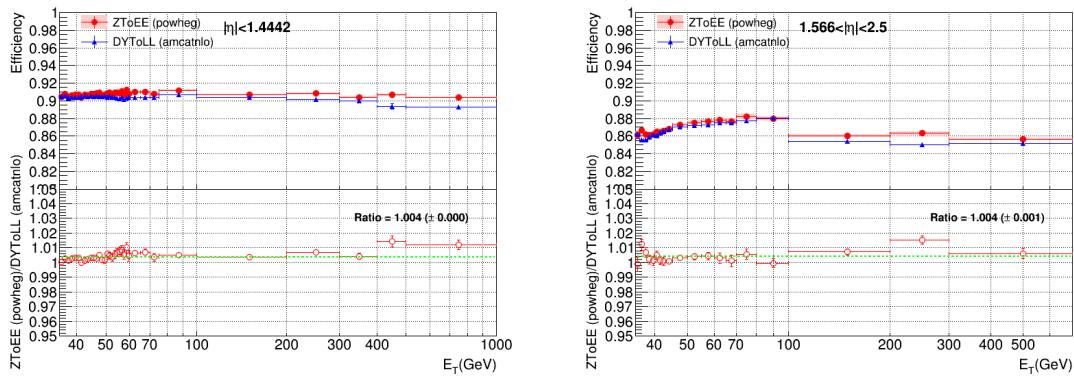


Figure 50: The HEEP efficiencies for mc matched electron for DYToLL amcatnlo sample and ZToEE powheg sample for barrel (left) and endcap (right) for different generated E_T of electron.

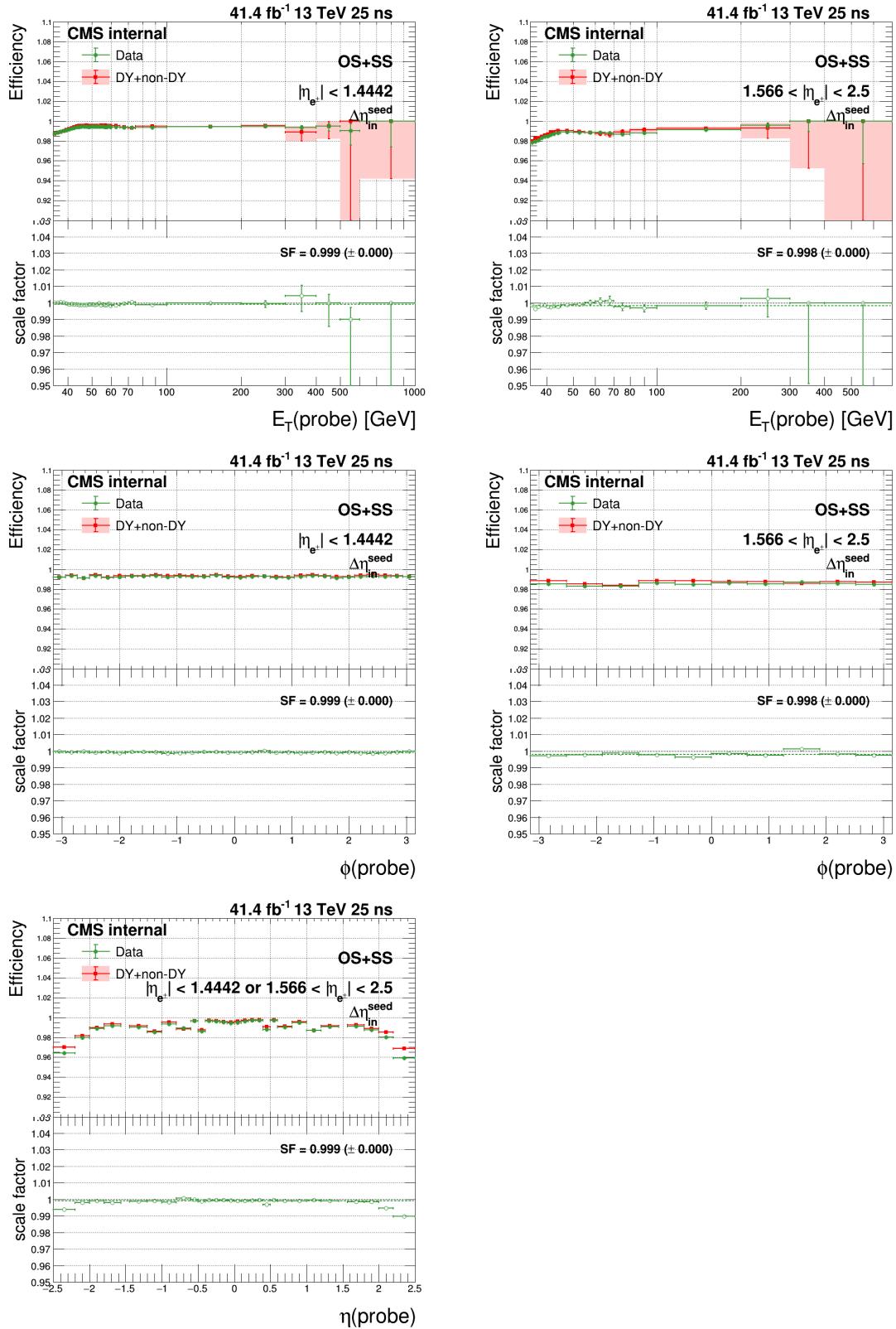


Figure 51: $\Delta\eta_{in}^{seed}$ N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

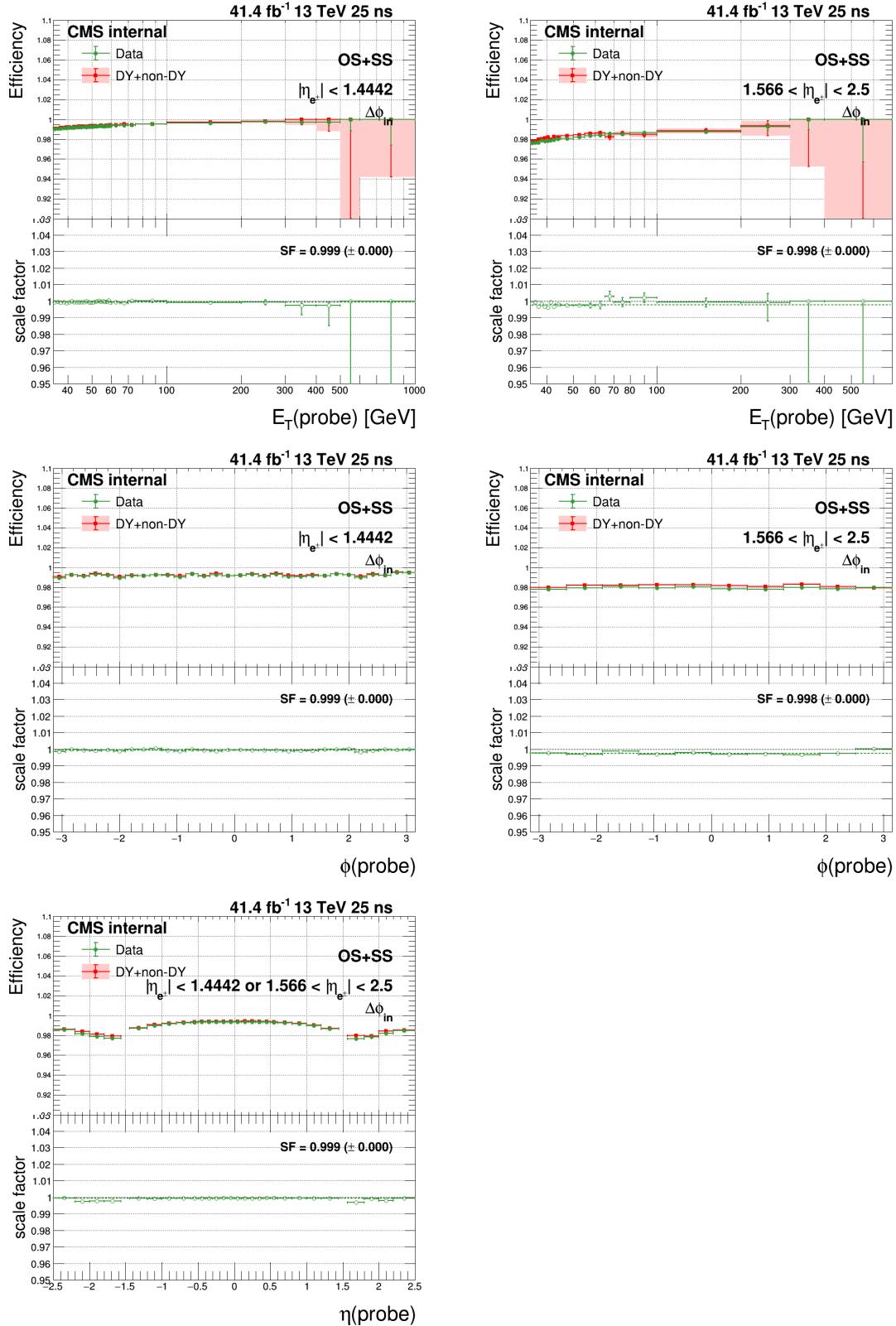


Figure 52: $\Delta\phi_{in}$ N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

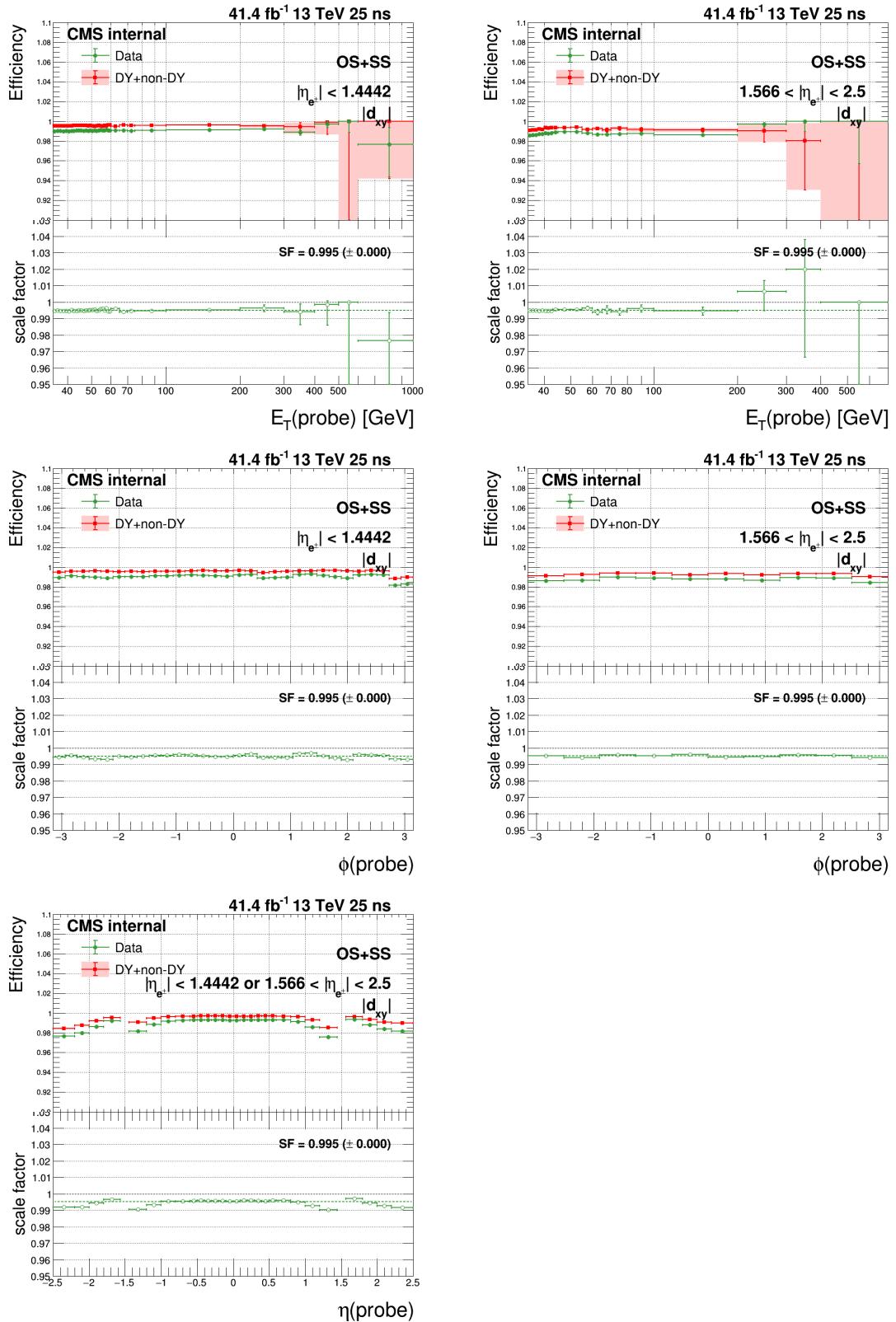


Figure 53: $|d_{xy}|$ N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

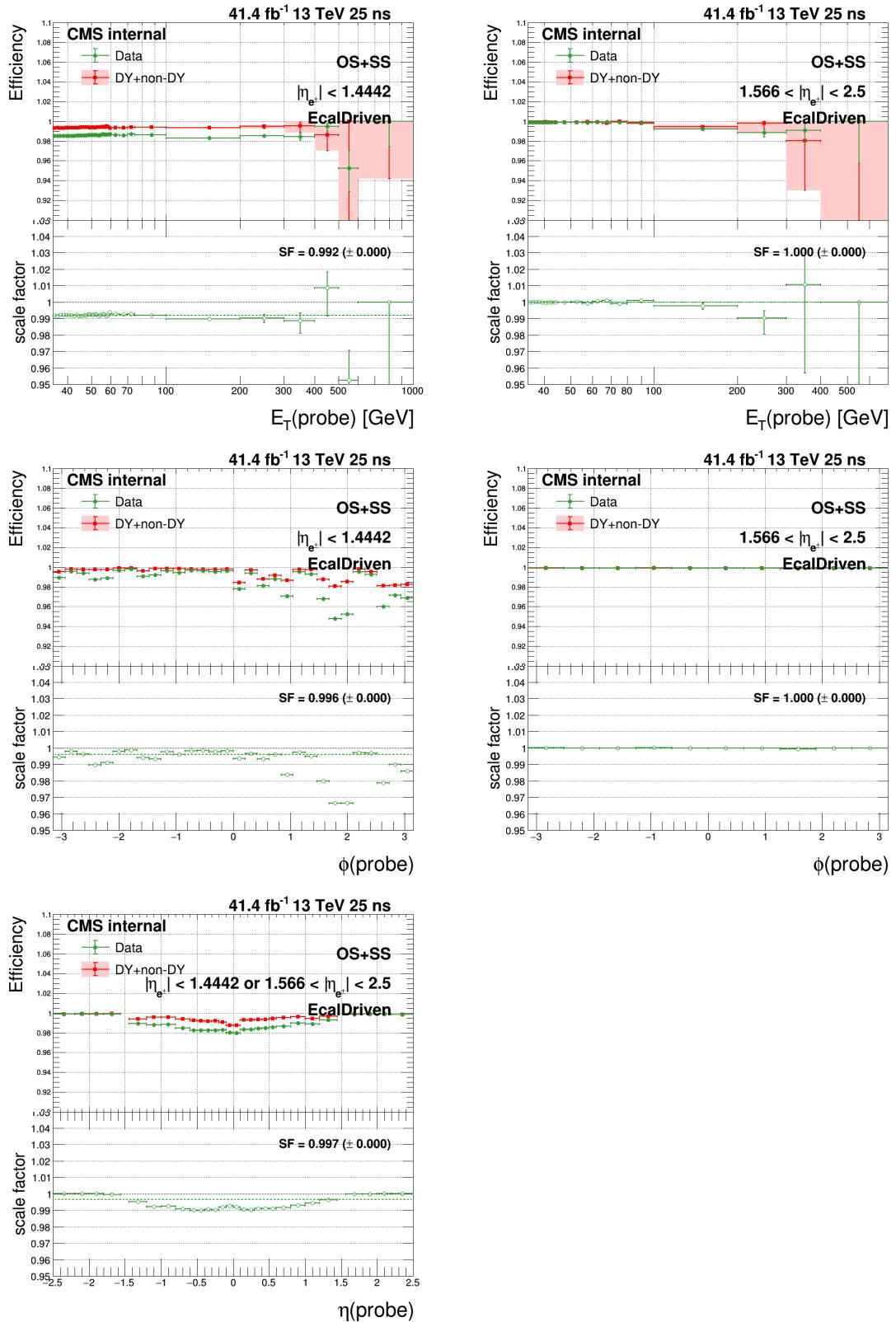


Figure 54: *EcalDriven* N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

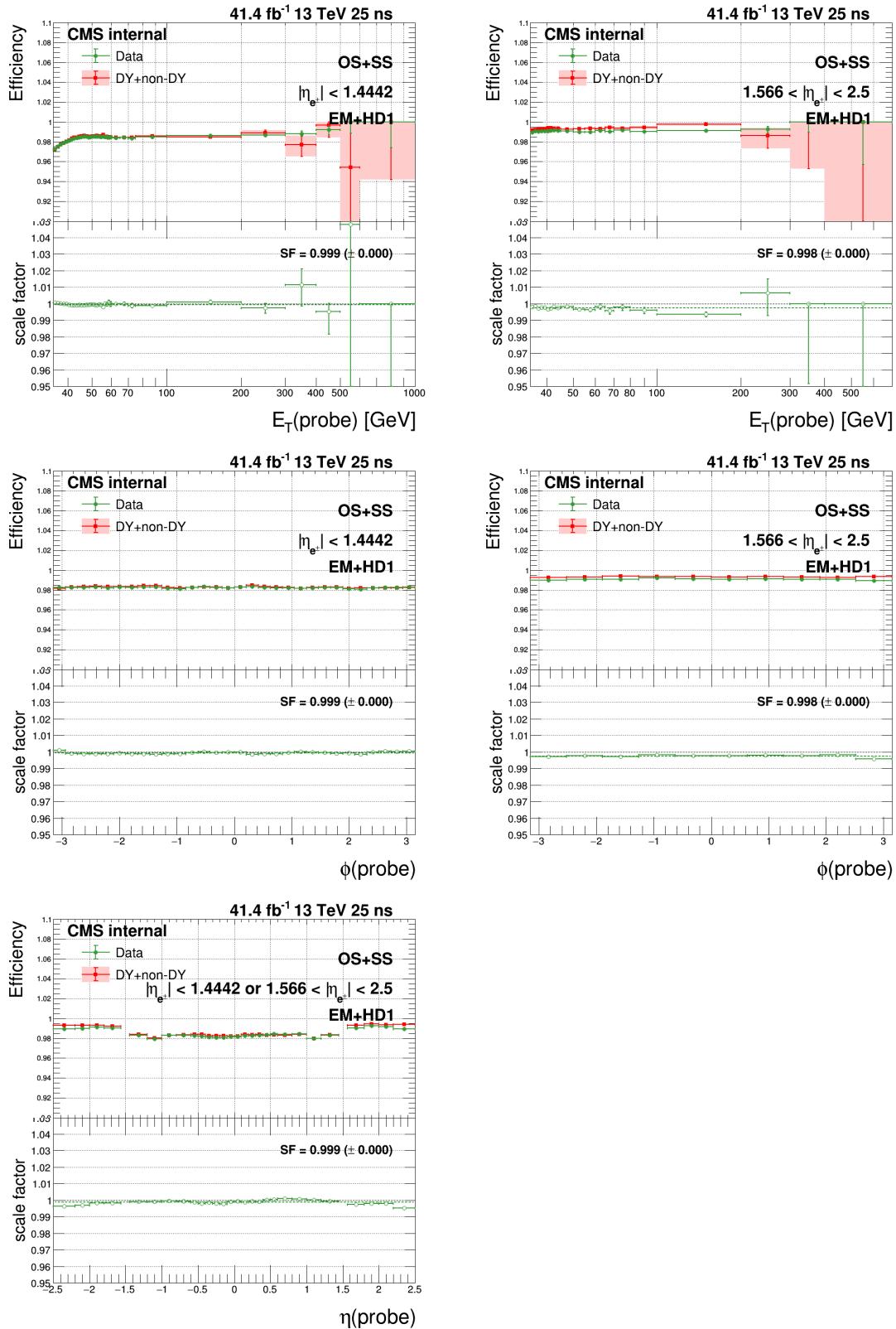


Figure 55: $EM + HD1$ N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

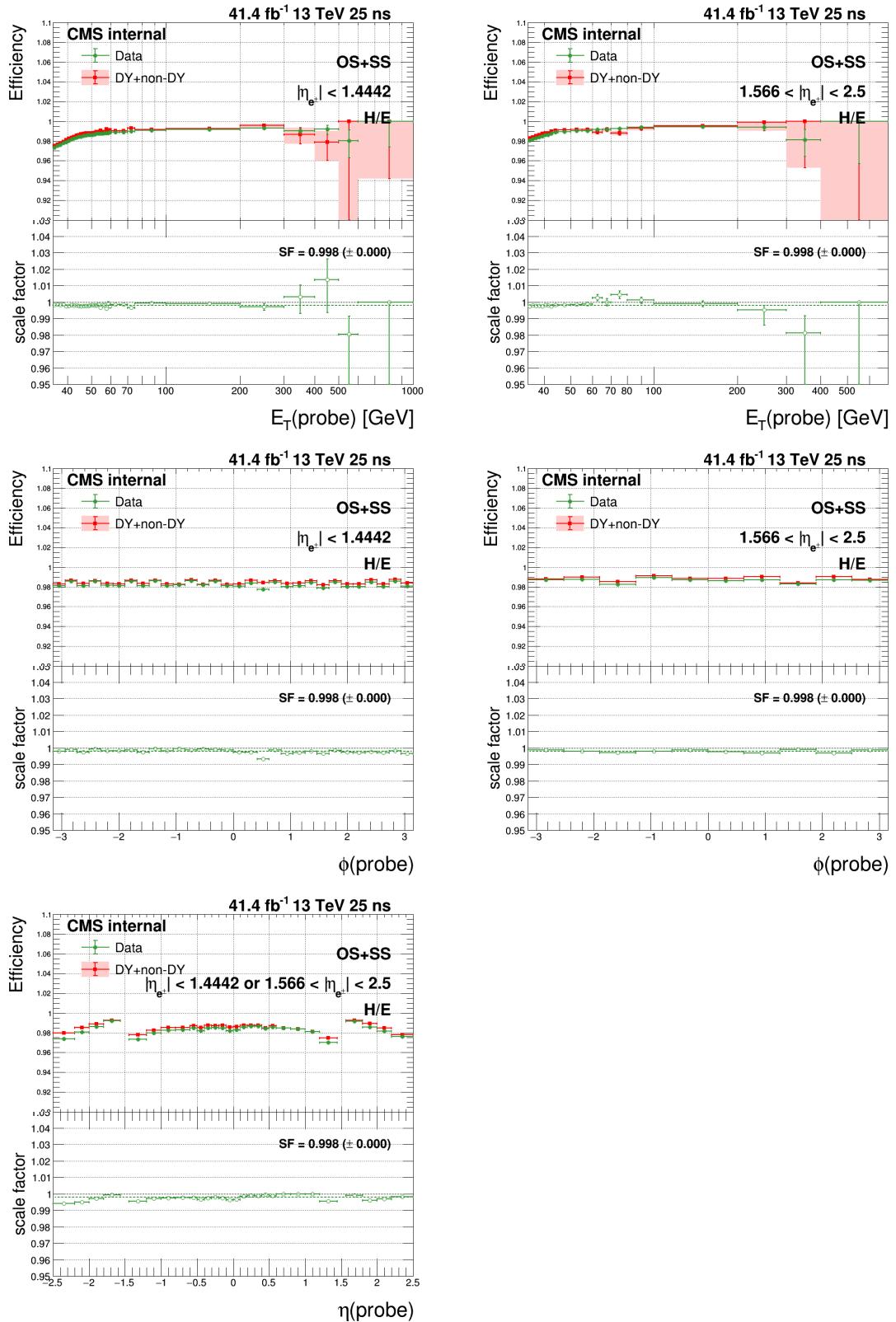


Figure 56: H/E N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

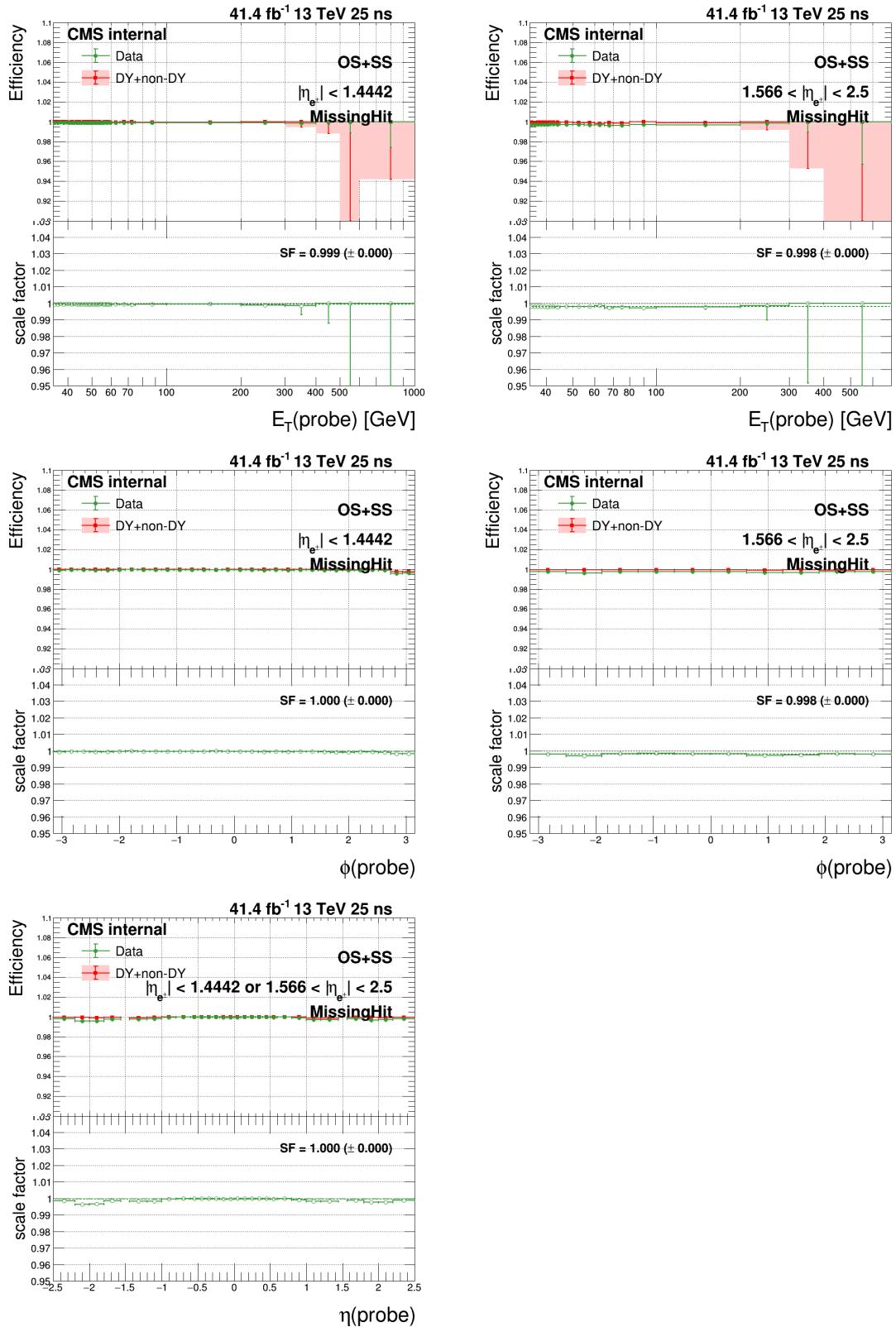


Figure 57: *MissingHit* N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

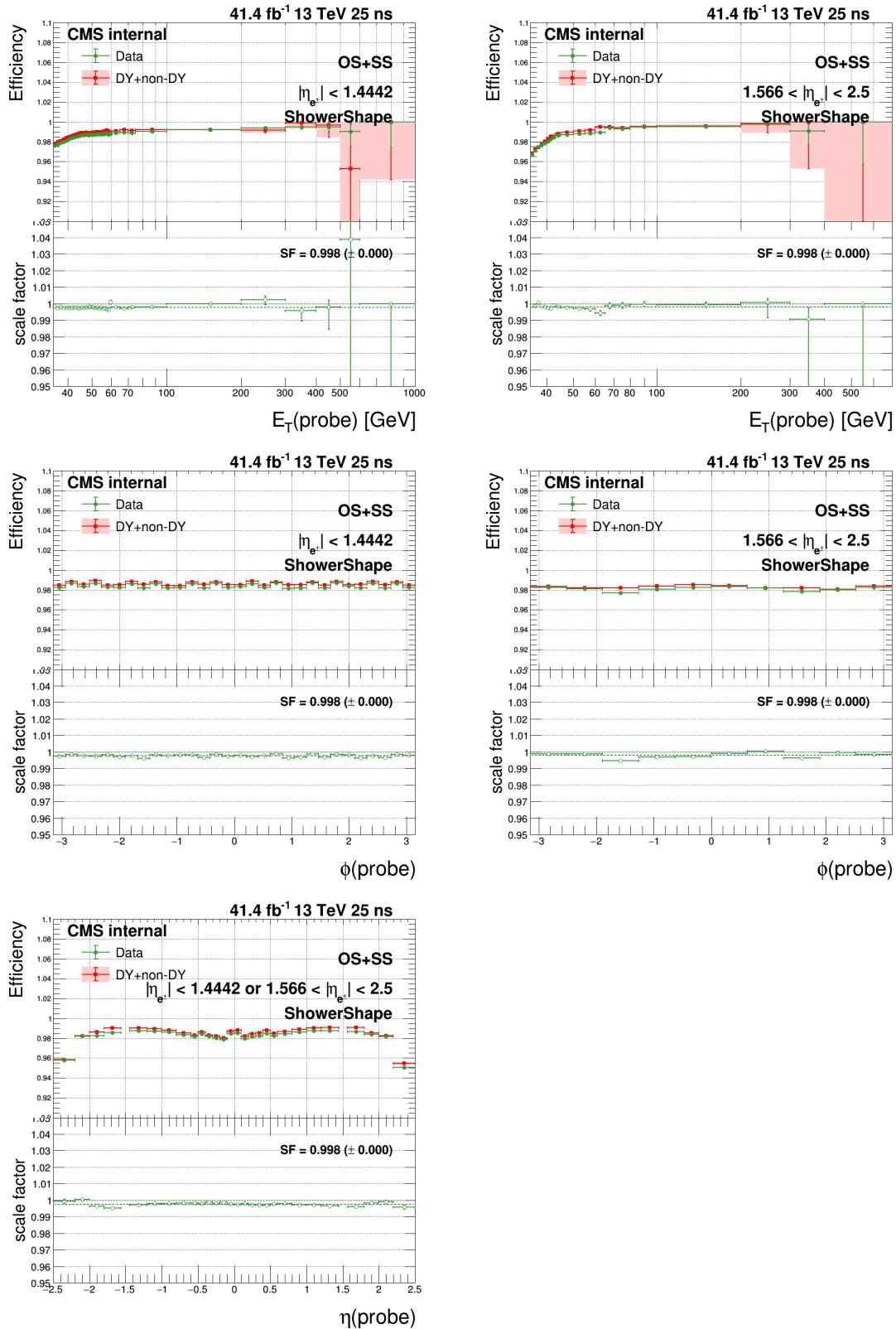


Figure 58: *ShowerShape* N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

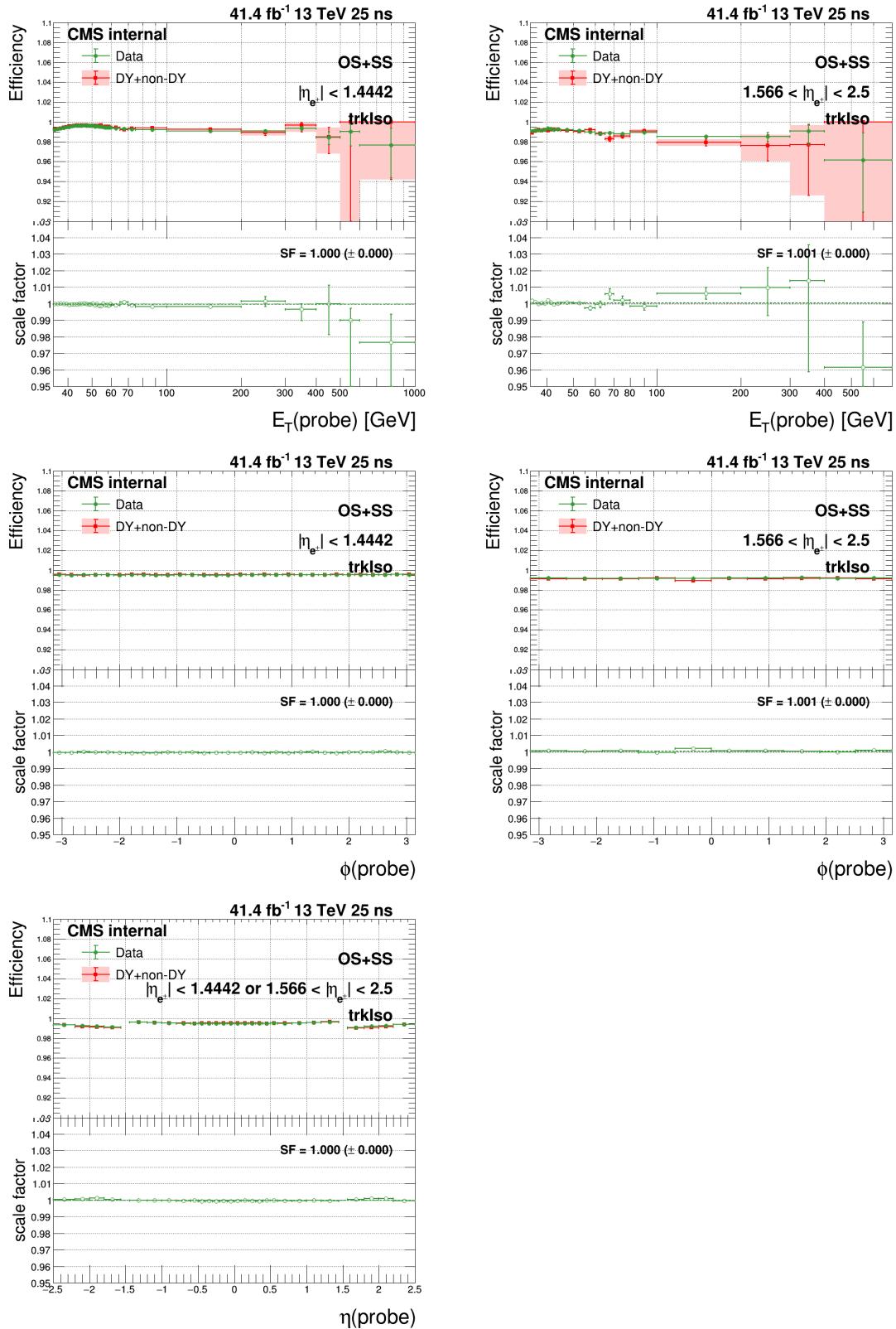


Figure 59: *trackisolation* N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

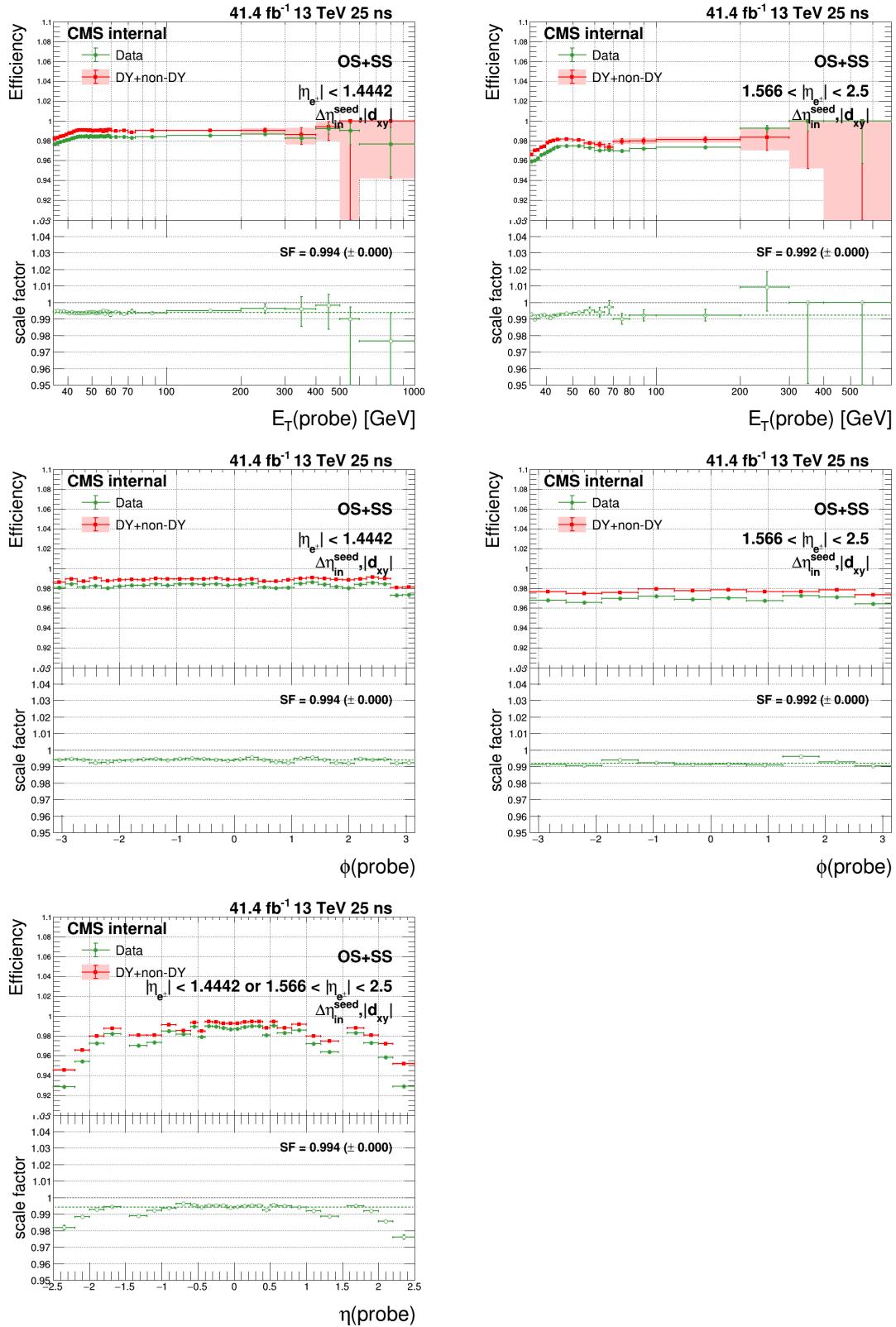


Figure 60: $\Delta\eta_{in}^{seed}$ and $|d_{xy}|$ N-2 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

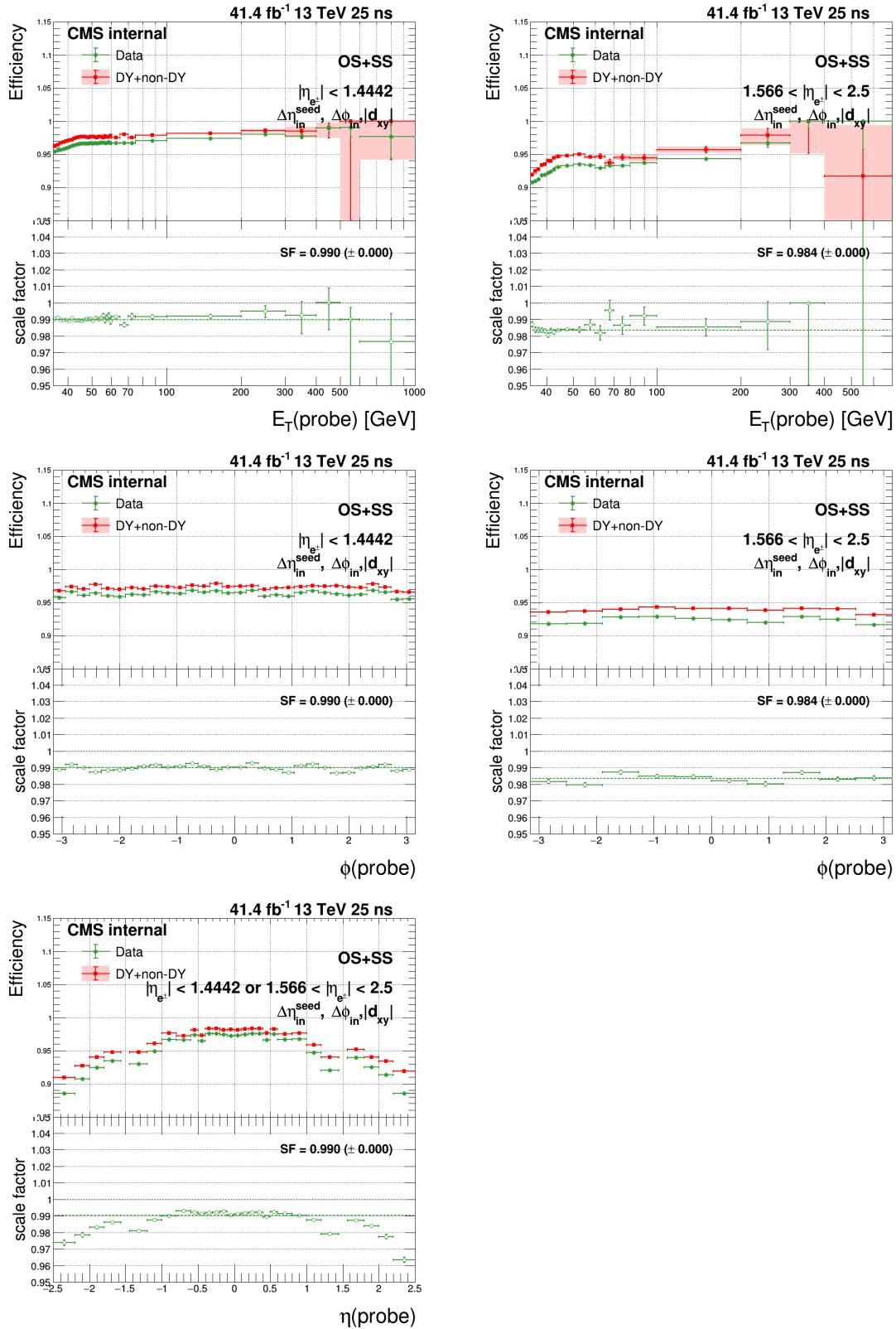


Figure 61: $\Delta\eta_{in}^{seed}$, $|d_{xy}|$ and $\Delta\phi_{in}$ N-3 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

.3.2 Cross check with fit method

Here we check the HEEP ID efficiency and scale factor using fit method. For data the invariant mass distribution of tag and probe pair in 70-110 GeV region is fitted by signal (DY mc template convoluted by Gaussian) + background (CMS shape) and the efficiency is equal to the number of passing signal events divided by the sum of passing signal events and failing signal events. For DY mc the efficiency is calculated by cut and count method, the tag and probe are required to match generated electron.

The fit package is from EGM, some fit plots are shown in Figure 63 and 64. The HEEP ID efficiency and scale factor is shown in Figure 65. The comparation of the HEEP ID efficiency and scale factor from cut method and fit method are shown in Figure 66.

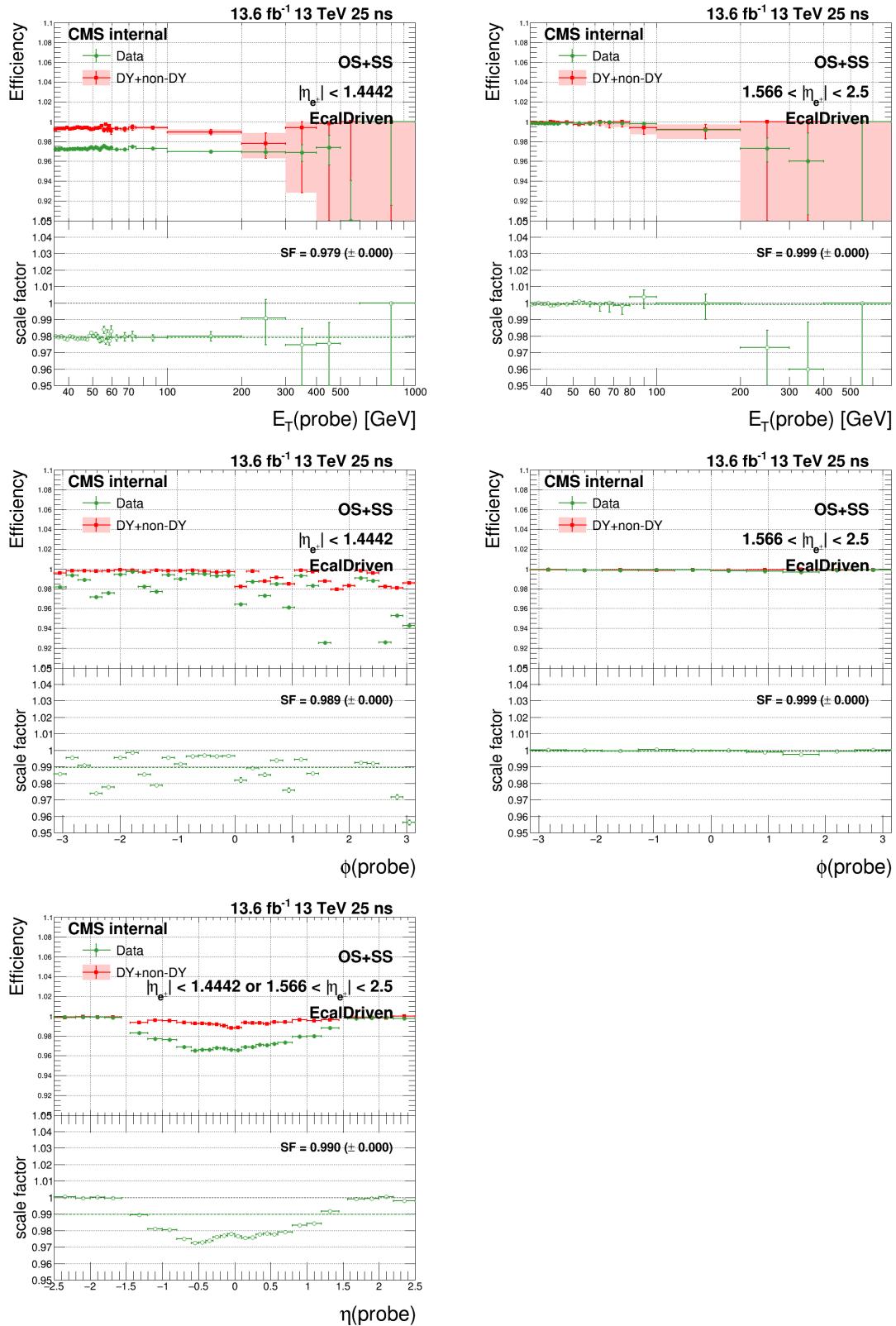


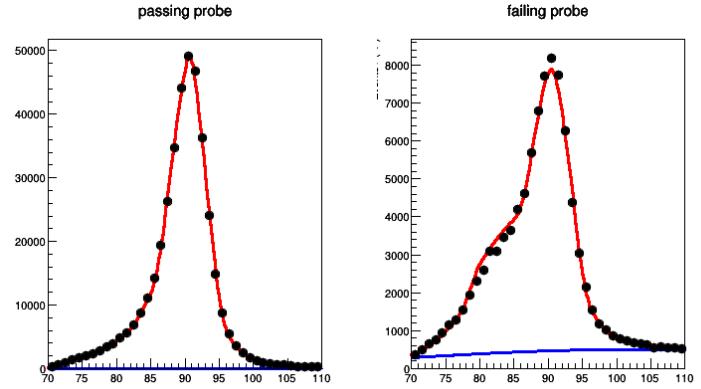
Figure 62: *EcalDriven* N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom) for run F.

* fit status pass: 0, fail : 0

* eff = 0.8285 ± 0.0008

--- parameters

- acmsP = 141.175 ± 55.034
- betaP = 0.035 ± 0.020
- gammaP = 0.097 ± 0.059
- meanP = 0.704 ± 0.006
- nBkgP = 1112.553 ± 379.268
- nSigP = 393384.838 ± 723.234
- sigmaP = 0.866 ± 0.013
- acmsF = 169.333 ± 10.771
- betaF = 0.025 ± 0.001
- gammaF = 0.097 ± 0.012
- meanF = 0.698 ± 0.019
- nBkgF = 17554.511 ± 321.887
- nSigF = 81421.981 ± 408.589
- sigmaF = 0.929 ± 0.040

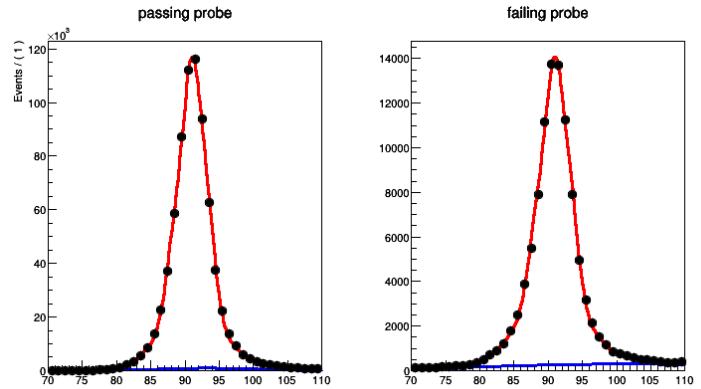


* fit status pass: 4, fail : 0

* eff = 0.8843 ± 0.0006

--- parameters

- acmsP = 154.590 ± 0.315
- betaP = 0.100 ± 0.000
- gammaP = 1.246 ± 0.010
- meanP = 0.768 ± 0.002
- nBkgP = 16228.204 ± 5424.191
- nSigP = 719041.608 ± 2363.712
- sigmaP = 0.773 ± 0.012
- acmsF = 139.006 ± 28.426
- betaF = 0.037 ± 0.003
- gammaF = 0.125 ± 0.065
- meanF = 0.723 ± 0.010
- nBkgF = 9091.370 ± 283.559
- nSigF = 94107.446 ± 406.815
- sigmaF = 0.773 ± 0.026



* fit status pass: 4, fail : 0

* eff = 0.8948 ± 0.0058

--- parameters

- acmsP = 152.955 ± 1.712
- betaP = 0.100 ± 0.008
- gammaP = 1.246 ± 0.062
- meanP = 1.135 ± 0.019
- nBkgP = 380.954 ± 768.104
- nSigP = 9510.594 ± 390.440
- sigmaP = 0.929 ± 0.104
- acmsF = 192.542 ± 156.181
- betaF = 0.016 ± 0.017
- gammaF = 0.020 ± 0.107
- meanF = 0.681 ± 0.148
- nBkgF = 239.889 ± 41.228
- nSigF = 1118.076 ± 50.655
- sigmaF = 2.237 ± 0.166

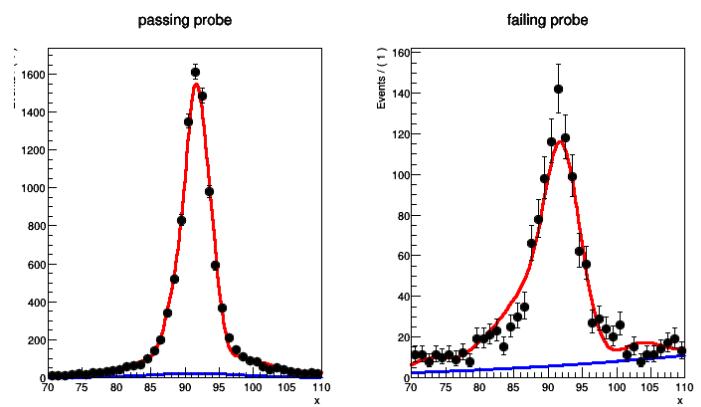


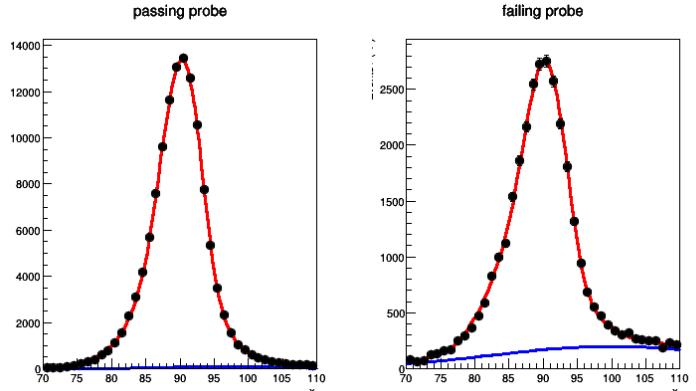
Figure 63: The fit plot of passing probes (left) and failing probes (right) for E_T of probe between 35-36 GeV (top), 44-45 GeV (middle) and 200-1000 GeV (bottom) in the barrel.

* fit status pass: 4, fail : 4

* eff = 0.8196 ± 0.0018

--- parameters

- acmsP = 44.268 ± 0.622
- betaP = -0.100 ± 0.003
- gammaP = -1.051 ± 0.016
- meanP = 0.141 ± 0.019
- nBkgP = 1718.972 ± 1758.863
- nSigP = 122472.103 ± 888.126
- sigmaP = 1.370 ± 0.043
- acmsF = 147.342 ± 7.036
- betaF = 0.042 ± 0.002
- gammaF = 0.185 ± 0.022
- meanF = 0.256 ± 0.032
- nBkgF = 5870.512 ± 210.544
- nSigF = 26953.397 ± 255.250
- sigmaF = 1.283 ± 0.085

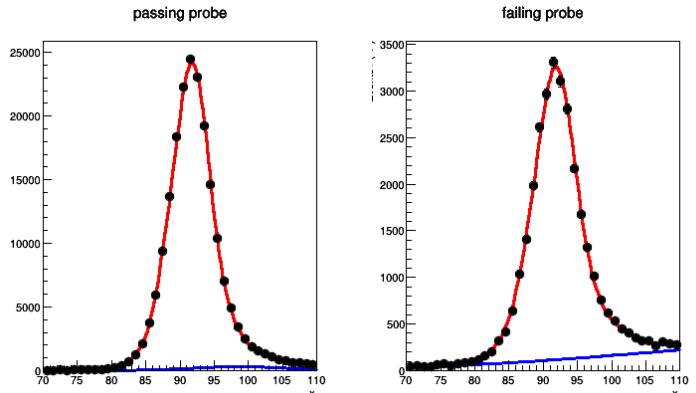


* fit status pass: 4, fail : 0

* eff = 0.8724 ± 0.0010

--- parameters

- acmsP = 153.012 ± 1.474
- betaP = 0.100 ± 0.000
- gammaP = 1.106 ± 0.042
- meanP = 0.579 ± 0.011
- nBkgP = 5150.073 ± 2235.321
- nSigP = 192796.231 ± 999.003
- sigmaP = 1.292 ± 0.017
- acmsF = 181.220 ± 24.825
- betaF = 0.023 ± 0.003
- gammaF = 0.059 ± 0.026
- meanF = 0.605 ± 0.026
- nBkgF = 4619.477 ± 143.402
- nSigF = 28202.728 ± 209.856
- sigmaF = 1.483 ± 0.052



* fit status pass: 4, fail : 0

* eff = 0.8761 ± 0.0061

--- parameters

- acmsP = 151.873 ± 3.662
- betaP = 0.100 ± 0.007
- gammaP = 1.131 ± 0.103
- meanP = 0.911 ± 0.058
- nBkgP = 525.915 ± 691.049
- nSigP = 9885.514 ± 374.159
- sigmaP = 1.426 ± 0.107
- acmsF = 152.710 ± 103.199
- betaF = 0.001 ± 0.040
- gammaF = -0.029 ± 0.051
- meanF = -0.054 ± 0.163
- nBkgF = 376.240 ± 49.424
- nSigF = 1397.431 ± 58.840
- sigmaF = 2.253 ± 0.238

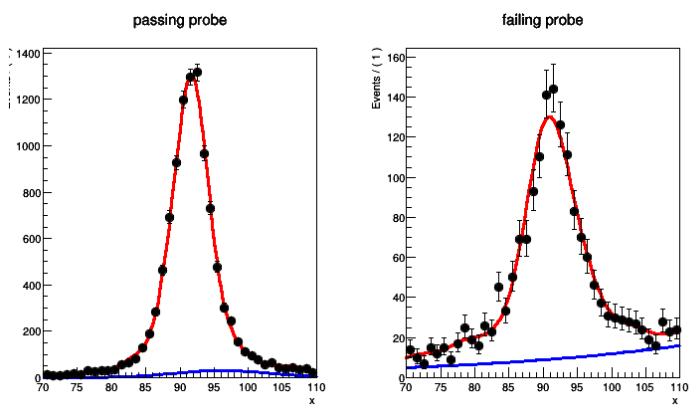


Figure 64: The fit plot of passing probes (left) and failing probes (right) for E_T of probe between 35-36 GeV (top), 43-45 GeV (middle) and 100-1000 GeV (bottom) in the endcap.

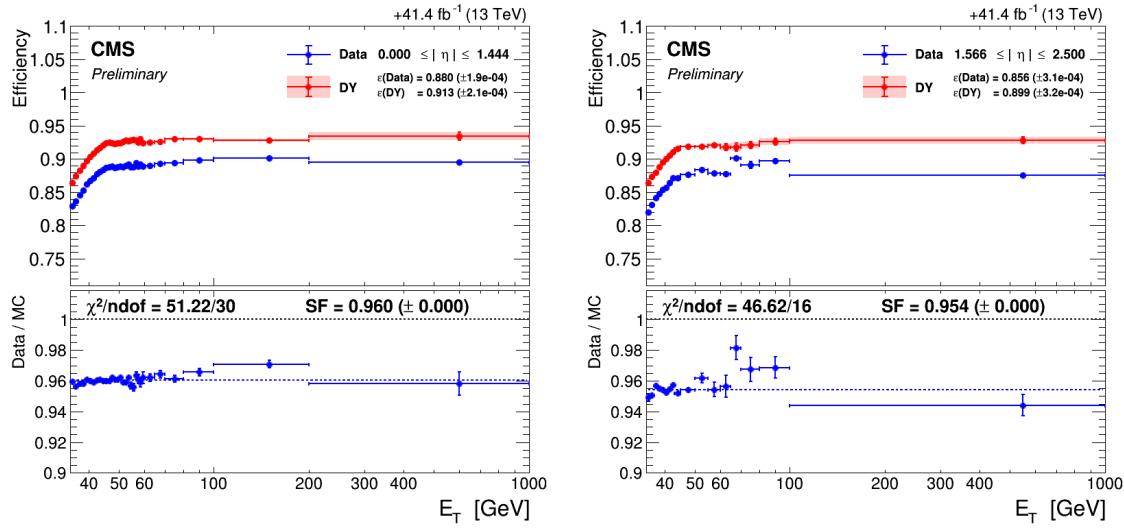


Figure 65: Efficiencies and scale factors in MC and data as functions of probe E_T .

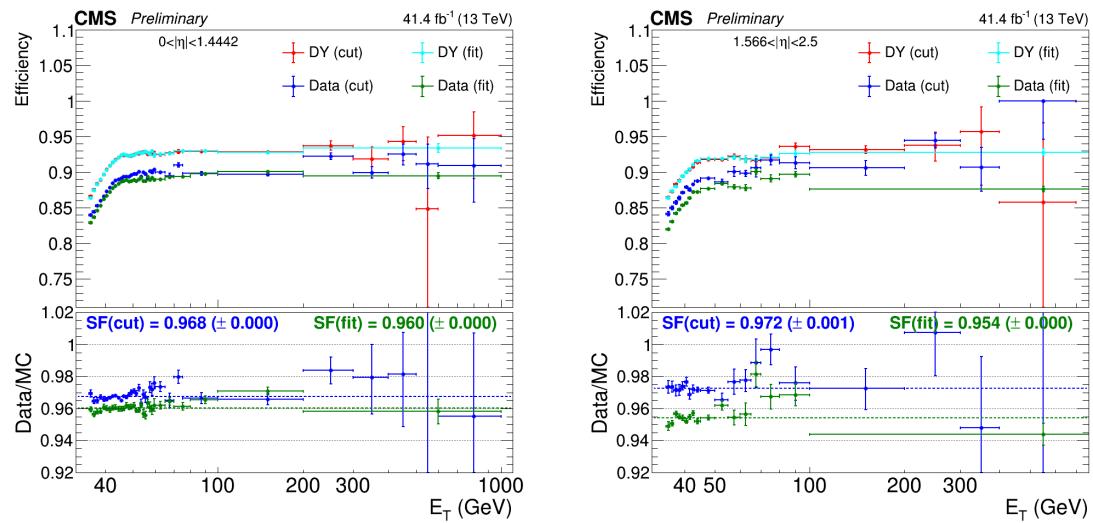


Figure 66: The comparation of efficiencies and scale factors in MC and data between cut method and fit method as functions of probe E_T .

.4 MET disagreement investigation

***NOTE: The following mitigation method for Data/MC disagreement is not applied to the analysis anymore.

The MET agreement between data and MC is not good which can be seen in Figure 4.2. The problem is present in same flavor channel mostly from DY process which has not real MET.

.4.0.1 Source of MET disagreement

Due to the fact that there is no real MET (comes from neutrino in W decay) in DY process and MET in this process originate from jets, we conclude that the problem can come from the following sources.

- modeling of the vector-boson recoil and detector effects, which can be difficult to simulate accurately and deficiencies in the modeling of the calorimeter response and resolution [92]

- poor simulation of number of pile up

In order to investigate the simulation of recoiled jet, we looked at the $|\Delta\phi|$ between MET and dilepton. We expect to see MET disagreement, because of the poor simulation of the recoiled jets, mostly close to 0 or π . In Figure 67, the distribution of $|\Delta\phi|(\text{MET}, \text{ll})$ is shown. We see disagreement close to the π . So if the source of MET disagreement is related to the events with a high p_T jet recoiling against jet, we should see good agreement for events without such a jet. We put a cut on $|\Delta\phi|(\text{MET}, \text{ll}) < 2$ and looked at MET distribution which can be seen in Figure 67. It can be seen that MET has similar shape and the disagreement source is not the poor simulation of recoiled jet.

As it can be seen in Figure 68, the distribution of number of reconstructed vertices is different for data and MC and pileup reweighting does not work properly. Due to the fact that there is a correlation between MET and number of pileup, disagreement in number of vertex distribution can be propagated into the MET distribution. In Figure 69, MET distributions are shown for events with number of vertices in various bins after step1 selection. One can see that the ratio of data and MC for the MET (MET shape) is much more flat compare with the nominal plot (see figure 4.2) although the ratio of data MC normalization is not 1 because of the fact that pile-up reweighting doesn't work well. After finding the relation between MET and number of vertices, we do a reweighting using number of vertices in same flavor channels. The results are shown in Figure 70. The data MC agreement has improved clearly.

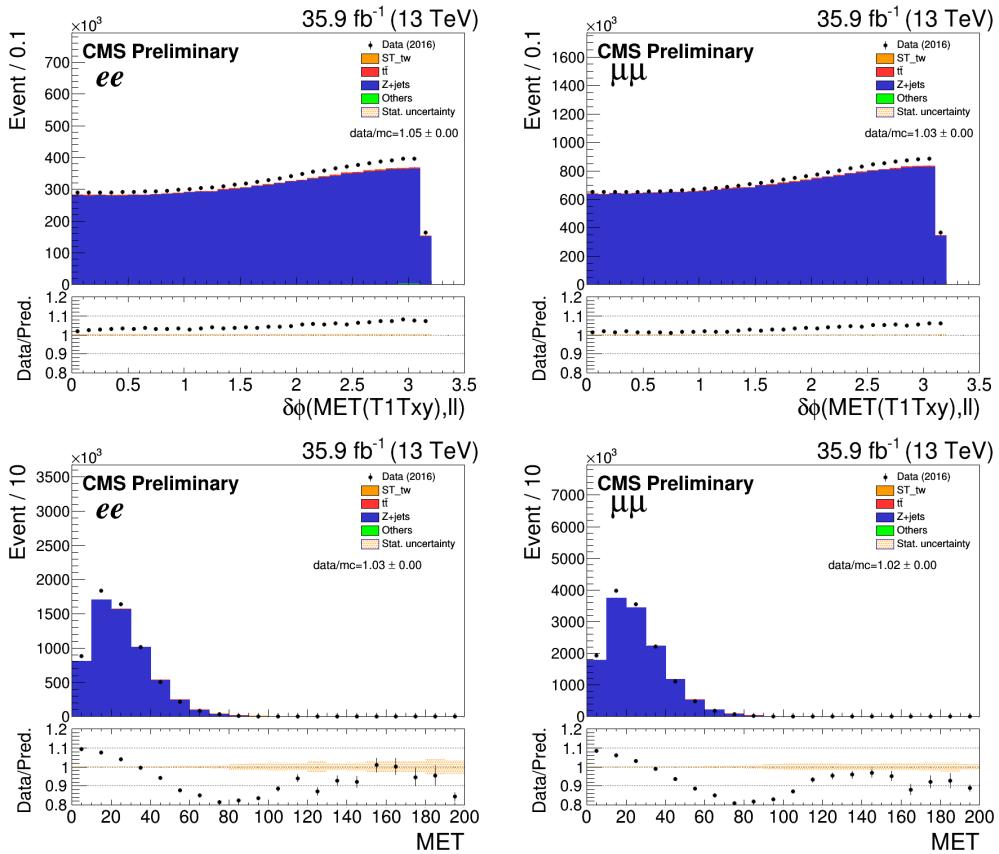


Figure 67: The distributions of the $\Delta\phi$ (MET, ll) (first row), MET (second row) for events with $\Delta\phi$ (MET, ll) < 2 , for ee (left) and $\mu\mu$ (right) channels after step 1.

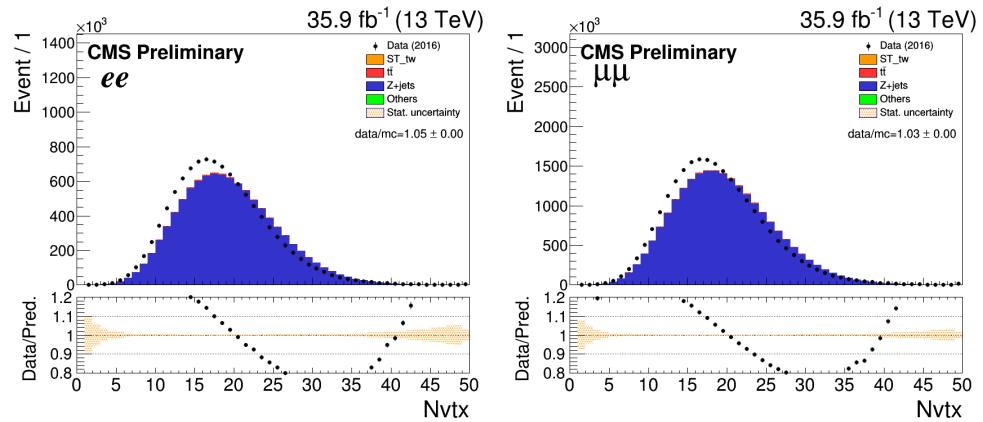


Figure 68: The distributions of number of reconstructed vertices for ee (left) and $\mu\mu$ (right) channels after step 1.

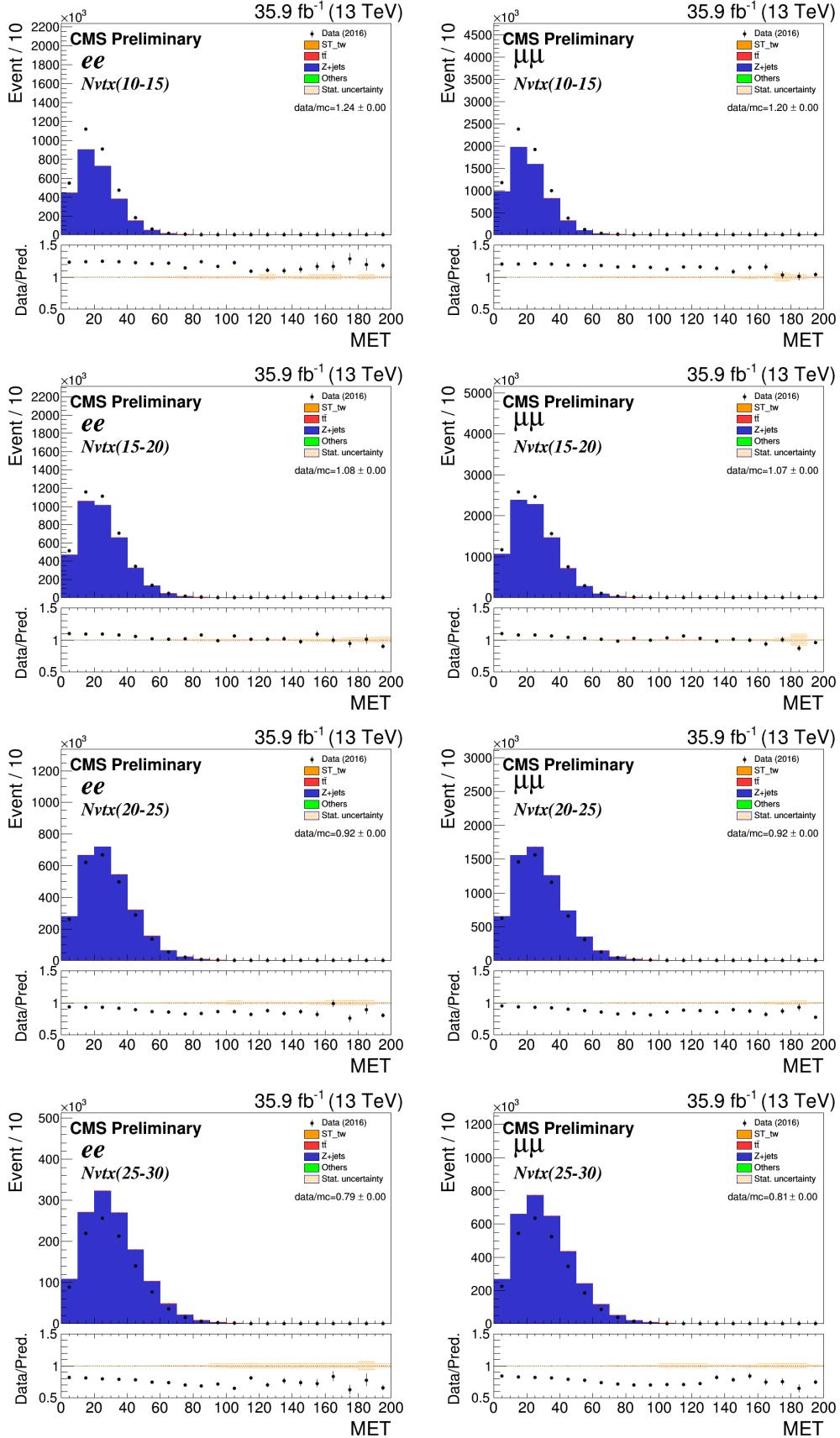


Figure 69: The distributions of MET for events with number of vertices between 10-15 (first row), 15-20 (second row), 20-25 (third row) and 25-30 (last row) for ee (left) and $\mu\mu$ (right) channels.

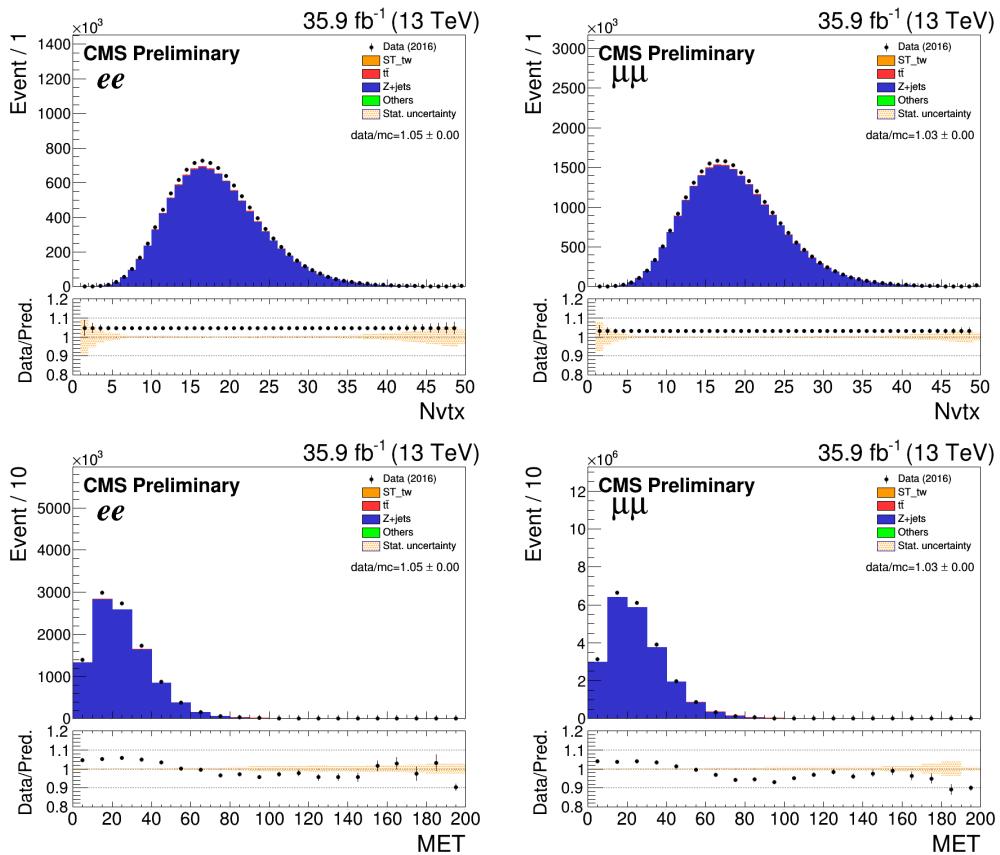


Figure 70: The distributions of number of vertices (top) and MET (bottom) after number of vertices re-weighting for ee (left) and $\mu\mu$ (right) channels.

5 SM tW cross section measurement

In order to do the fit for measuring the tW cross section, we utilize the MLP (which is discussed in Section 4.8) output distributions for both data and MC expectation in the (1jet,1b-jet) and (2jet,1b-jet) regions and event yield in the (≥ 2 jet,2b-jet) region for ee and $\mu\mu$ channels. The inclusion of the (≥ 2 jet,2b-jet) and (2jet,1b-jet) regions helps to constrain the normalization and systematic uncertainties of the $t\bar{t}$ background.

Comparison between observed data and the SM background prediction for the MLP output shape in various jet-bjet regions are shown in Figure 71. All sources of systematic uncertainties discussed in Section 4.9 are included in our results.

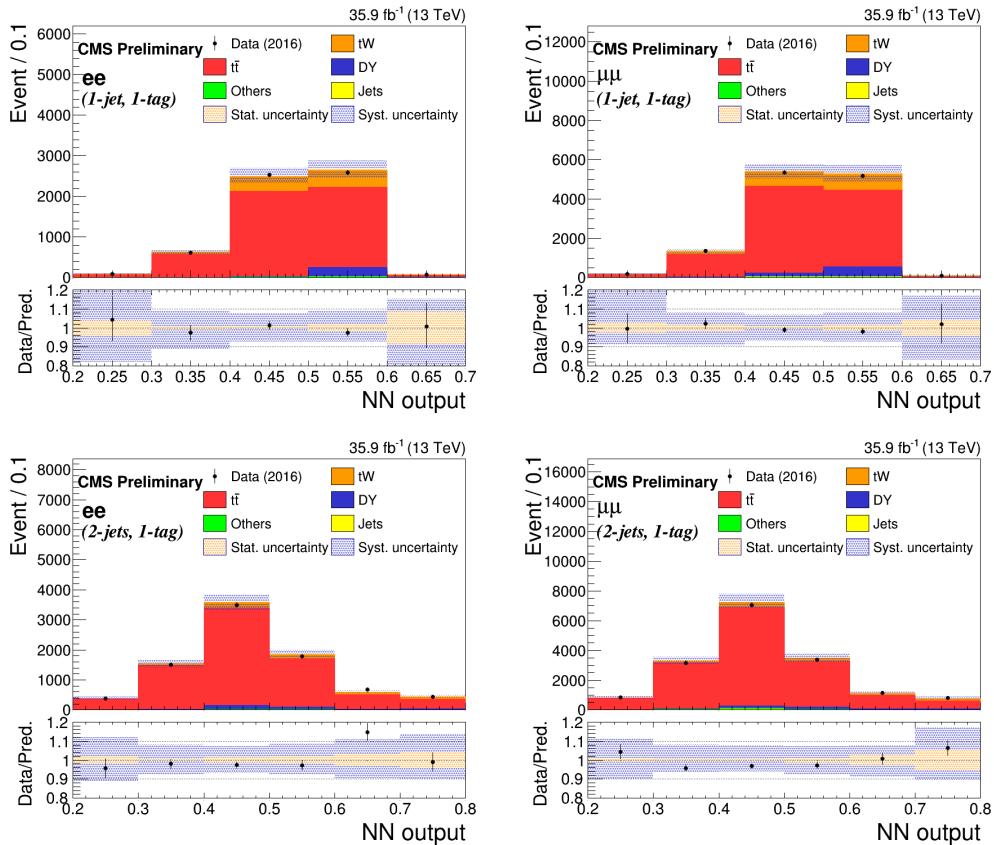


Figure 71: The MLP distributions for different (1jet,1b-jet) region (top row), (2jet,1b-jet) region (bottom row) for ee channel (left column) and $\mu\mu$ channel (right column).

In Figure 72, likelihood scan for the signal strength are shown for different region for ee and $\mu\mu$ channels. In Table 22, tW measured cross section is reported for ee and $\mu\mu$ channels compared with the result from $e\mu$ channel and combined channels. In figures 73 and 74, impact of individual systematic sources for ee, $\mu\mu$, $e\mu$ and combined channels are shown for expected and observed. In Table 23, the effect of each systematic uncertainty source to the combined fit is shown.

To summarize, the tW cross-section is measured to be $58.08 \pm 9.32(syst.) \pm 1.43(stat.)\text{ pb}$ with a 6.9 (7.3) σ significance for observed (expected) and in agreement with the standard model prediction of $\sigma_{tW}^{ref} = 71.7 \pm 1.8(scale) \pm 3.4(PDF)\text{ pb}$.

Table 22: The expected significance and best fit of tW cross section measurement for ee , $e\mu$, $\mu\mu$ channels and combined

region	Exp./Obs. significance σ	Exp./Obs. best fit
ee MLP output for $(1j1t + 2j1t) +$ yields ($>=2j, 2t$)	3.2 / 3.5	$1.00^{+0.23}_{-0.29} / 1.14^{+0.25}_{-0.28}$
$\mu\mu$ MLP output for $(1j1t + 2j1t) +$ yields ($>=2j, 2t$)	2.7 / 4.2	$1.00^{+0.25}_{-0.25} / 1.14^{+0.24}_{-0.27}$
$e\mu$ MLP output for $(1j0t + 1j1t + 2j1t) +$ yields ($>=2j, 2t$)	6.4 / 5.7	$1.00^{+0.14}_{-0.16} / 0.92^{+0.16}_{-0.16}$
combined	7.3 / 6.9	$1.00^{+0.11}_{-0.12} / 0.82^{+0.14}_{-0.12}$

Table 23: The effect of systematical uncertainties for combined channel

Source	Uncertainty
TT_PDF	2.451%
ISR	2.734%
TW_DS	8.205%
FSR	3.824%
Trigger	2.801%
ElectronIDIso	3.355%
PileUp	3.848%
TW_mtop	1.674%
DY_normalisation	6.578%
MuonIso	2.385%
MuonID	2.765%
MuonTrack	2.015%
TT_CR	3.879%
Missingtag	2.608%
DY_PDF	2.569%
UnclusteredEn	5.394%
JER	3.395%
JES	12.475%
TW_ME	2.451%
Btag	6.093%
TT_QCD	3.034%
ElectronReco	2.698%
TT_normalisation	2.378%
Other_normalisation	3.188%
TT_mtop	2.910%
TT_Tune	3.305%
DY_QCD	1.866%
Jets_normalisation	1.998%
TT_hdamp	3.311%
Luminosity	4.431%
MC_stat	6.820%
Data_stat	2.435%
Total	16.176%

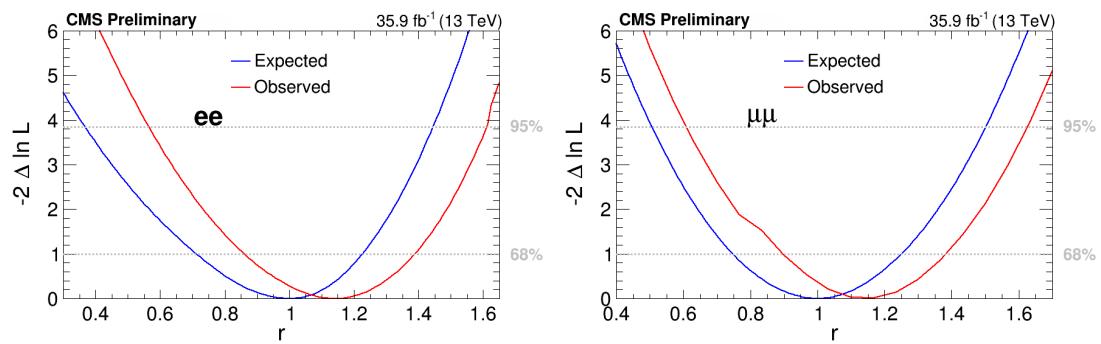


Figure 72: The continue likelihood scan for various regions for ee channel (left) and $\mu\mu$ (right) channels.

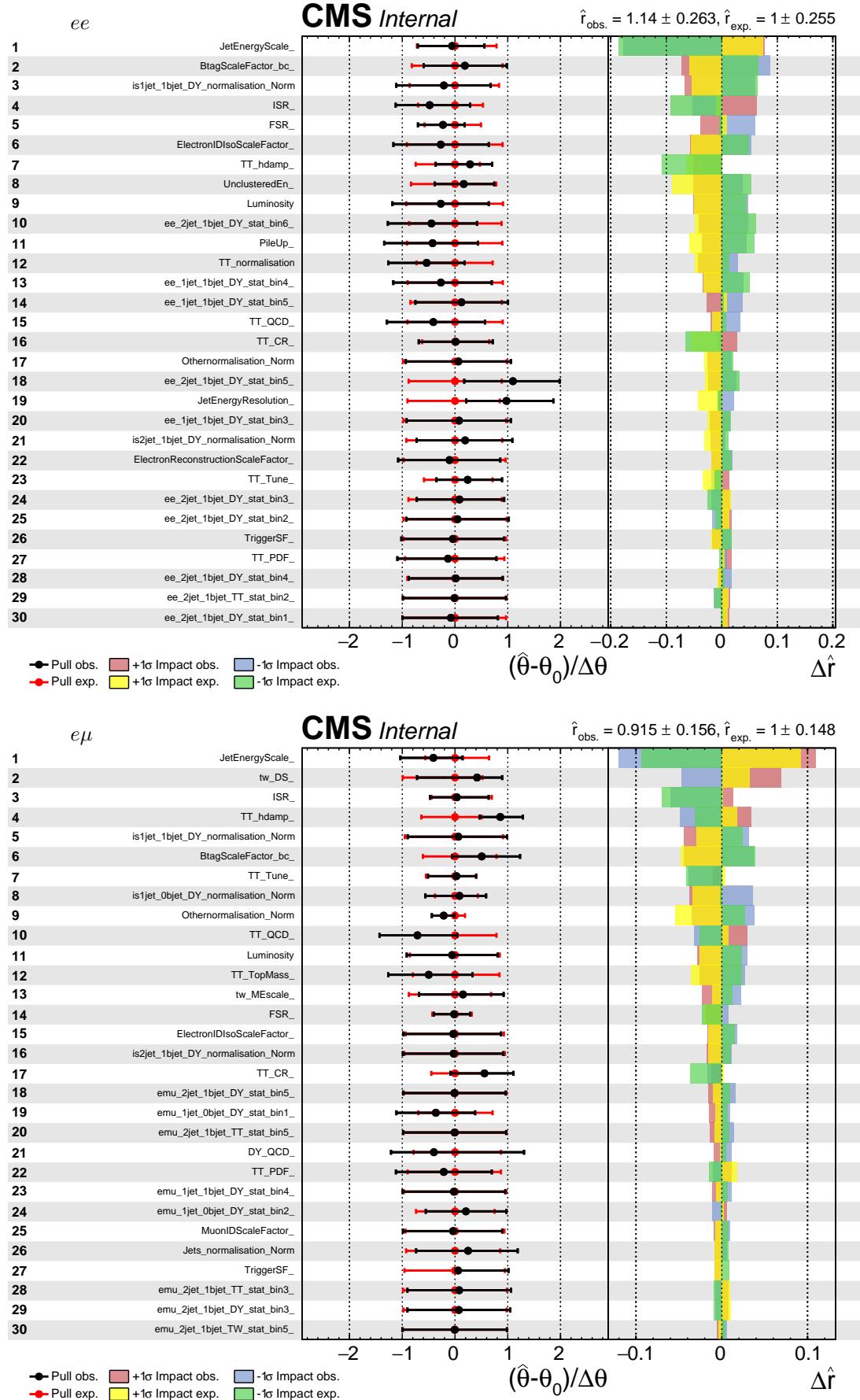


Figure 73: The expected and observed impacts of the most important uncertainty sources on the measurement of tW cross section in ee, eμ channels.

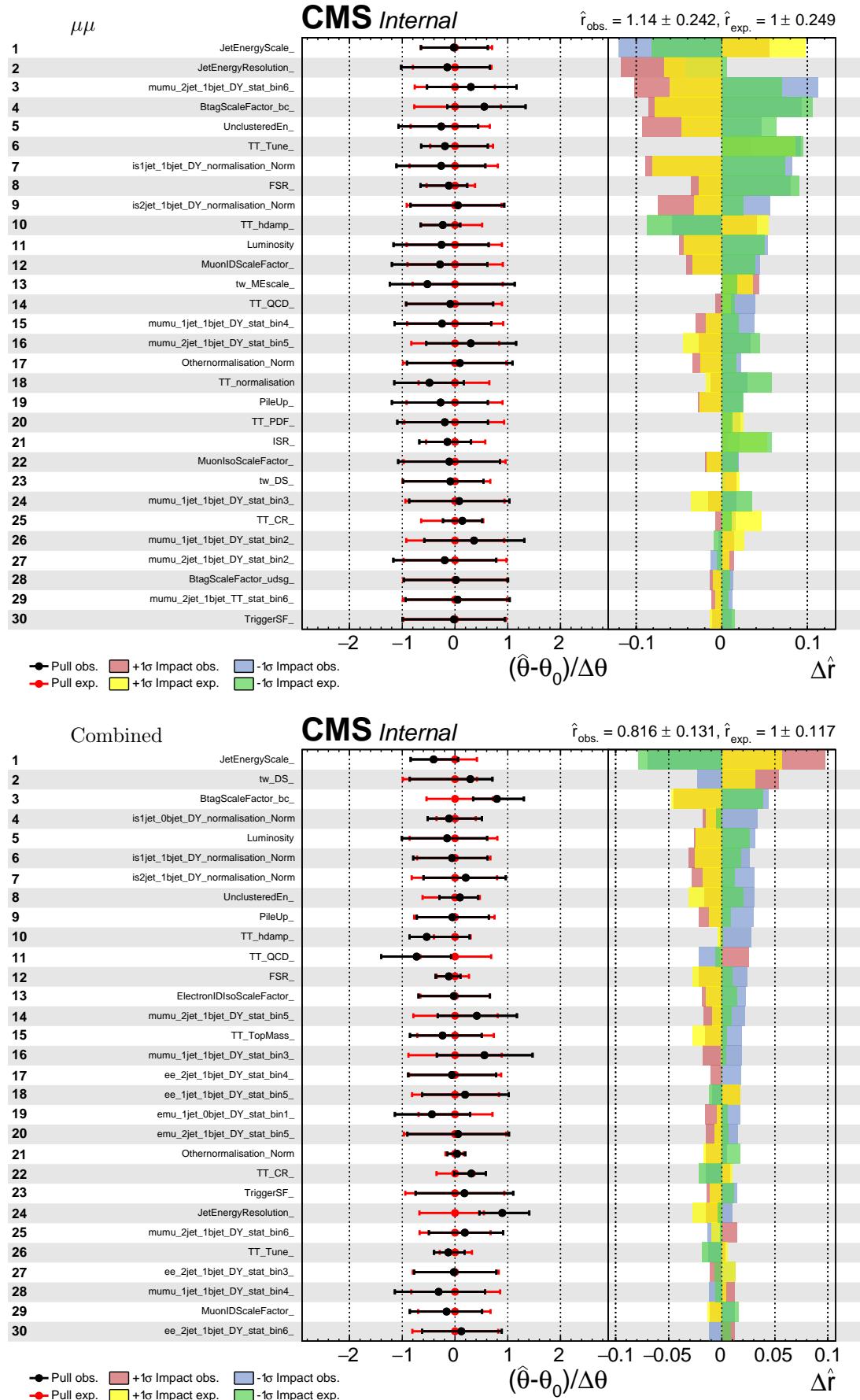


Figure 74: The expected and observed impacts of the most important uncertainty sources on the measurement of tW cross section in $\mu\mu$ channel and combined.

Bibliography

- [1] CMS. “Public CMS Luminosity Information”. In: *CMS Twiki* <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults> ().
- [2] W. Herr, B. J. Holzer, and B. Muratori. “6.4 Concept of Luminosity”. In: *Accelerators and Colliders*. Ed. by S. Myers and H. Schopper. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 140–146. ISBN: 978-3-642-23053-0. DOI: 10.1007/978-3-642-23053-0_9. URL: https://doi.org/10.1007/978-3-642-23053-0_9.
- [3] Antonella Del Rosso. “HL-LHC updates in Japan. Projet HL-LHC : une réunion fait le point au Japon”. In: BUL-NA-2014-272. 51/2014 (2014), p. 4. URL: <https://cds.cern.ch/record/1975962>.
- [4] E. Meschi et al. “Electron Reconstruction in the CMS Electromagnetic Calorimeter”. In: *CMS Notes* CMS Note-2001-034 (2001).
- [5] Franz Hamilton, Tyrus Berry, and Timothy Sauer. “Kalman-Takens filtering in the presence of dynamical noise”. In: Eur. Phys. J. C (2016). arXiv:1611.05414 [physics.data-an].
- [6] The CMS Collaboration. “Description and performance of track and primary-vertex reconstruction with the CMS tracker”. In: *J. Instrum.* 9 (2014), P10009. DOI: 10.1088/1748-0221/9/10/P10009.
- [7] Afiq A. Anuar. “Electrons and photons at High Level Trigger in CMS for Run II”. In: *J. Phys. Conf. Ser.* 664.8 (2015), p. 082001. DOI: 10.1088/1742-6596/664/8/082001.
- [8] K. Krajewska, C. Müller, and J. Z. Kamiński. “Bethe-Heitler pair production in ultrastrong short laser pulses”. In: *Phys. Rev. A* 87 (6 2013), p. 062107. DOI: 10.1103/PhysRevA.87.062107. URL: <https://link.aps.org/doi/10.1103/PhysRevA.87.062107>.
- [9] Livia Soffi. “Search for new exotic particles decaying to photons with the CMS experiment at the LHC”. In: *The European Physical Journal Plus* 131.8 (2016), p. 283. ISSN: 2190-5444. DOI: 10.1140/epjp/i2016-16283-3. URL: <http://dx.doi.org/10.1140/epjp/i2016-16283-3>.
- [10] The CMS Collaboration. *The CMS Physics Technical Design Report*. CERN/LHCC 2006-001, CMS TDR 8.1. 2006.
- [11] The CMS Collaboration. “Particle-Flow event reconstruction in CMS and performance for jets, taus and MET”. In: *CMS Physics Analysis Summaries* CMS-PAS-PFT-09-001 (2009).
- [12] The CMS Collaboration. “Commissioning of the particle-flow event reconstruction with leptons from J/Ψ and W decays at 7 TeV ”. In: *CMS Physics Analysis Summaries* CMS-PAS-PFT-10-003 (2010).
- [13] The CMS Collaboration. “Performance of τ -lepton reconstruction and identification in CMS”. In: *J. Instrum.* 7.arXiv:1109.6034. CMS-TAU-11-001. CERN-PH-EP-2011-137 (2011), P01001. 33 p. URL: <https://cds.cern.ch/record/1385560>.
- [14] E K Friis. “Tau reconstruction and identification at the compact muon solenoid”. In: *Nucl. Phys. B, Proc. Suppl.* 218 (2011), pp. 256–261. URL: <https://cds.cern.ch/record/2002999>.
- [15] Sanmay Ganguly. *Jet Measurements In CMS*. Tech. rep. arXiv:1303.6038. 2013. URL: <https://cds.cern.ch/record/1532824>.
- [16] Ivan Marchesini. *Identification of b -quark Jets in the CMS experiment*. Tech. rep. CMS-CR-2013-412. Geneva: CERN, 2013. URL: <https://cds.cern.ch/record/1629516>.
- [17] Saranya Ghosh. “Performance of MET reconstruction and pileup mitigation techniques in CMS”. In: *Nucl. Part. Phys. Proc.* 273-275.arXiv:1502.05207 (2015). Proceedings of the 37th International Conference on High Energy Physics, ICHEP 2014, 2512–2514. 3 p. URL: <https://cds.cern.ch/record/1992977>.

- [18] Paolo Nason. “A new method for combining NLO QCD with shower Monte Carlo algorithms”. In: *JHEP* 11 (2004), p. 040. DOI: 10.1088/1126-6708/2004/11/040. arXiv:0409146.
- [19] Stefano Frixione, Paolo Nason, and Carlo Oleari. “Matching NLO QCD computations with Parton Shower simulations: the POWHEG method”. In: *JHEP* 0711 (2007), p. 070. DOI: 10.1088/1126-6708/2007/11/070. arXiv:0709.2092 [hep-ph].
- [20] Simone Alioli et al. “A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX”. In: *JHEP* 06 (2010), p. 043. DOI: 10.1007/JHEP06(2010)043. arXiv:1002.2581 [hep-ph].
- [21] Simone Alioli et al. “NLO vector-boson production matched with shower in POWHEG”. In: *JHEP* 07 (2008), p. 060. DOI: 10.1088/1126-6708/2008/07/060. arXiv:0805.4802 [hep-ph].
- [22] Stefano Frixione, Paolo Nason, and Giovanni Ridolfi. “A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction”. In: *JHEP* 09 (2007), p. 126. DOI: 10.1088/1126-6708/2007/09/126. arXiv:0707.3088 [hep-ph].
- [23] Emanuele Re. “Single-top Wt-channel production matched with parton showers using the POWHEG method”. In: *Eur. Phys. J. C* 71 (2011), p. 1547. DOI: 10.1140/epjc/s10052-011-1547-z. arXiv:1009.2450 [hep-ph].
- [24] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections and their matching to parton shower simulations”. In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv:1405.0301 [hep-ph].
- [25] Torbjörn Sjöstrand et al. “An Introduction to PYTHIA 8.2”. In: *Comput. Phys. Commun.* 191 (2015), p. 159. DOI: 10.1016/j.cpc.2015.01.024. arXiv:1410.3012 [hep-ph].
- [26] S. Agostinelli et al. “GEANT4: a simulation toolkit”. In: *Nucl. Instrum. Meth. A* 506 (2003), p. 250. DOI: 10.1016/S0168-9002(03)01368-8.
- [27] W Adam et al. “Reconstruction of electrons with the Gaussian-sum filter in the CMS tracker at the LHC”. In: *Journal of Physics G: Nuclear and Particle Physics* 31.9 (2005), N9. URL: <http://stacks.iop.org/0954-3899/31/i=9/a=N01>.
- [28] The CMS Collaboration. “Generic Tag and Probe Tool for Measuring Efficiency at CMS with Early Data”. In: *CMS Notes* CMS AN-2009-111 (2009).
- [29] B. Clerbaux et al. “Search for high mass dielectron resonances with the full 2016 data”. In: *CMS Notes* CMS AN-2016-404 (2017).
- [30] B. Clerbaux et al. “Search for high-mass resonances in the di-electron final state with 2017 data”. In: *CMS Notes* CMS AN-2018-021 (2018).
- [31] B. Clerbaux et al. “Dielectron resonance search in Run 2 at $\sqrt{s} = 13$ TeV pp collisions”. In: *CMS Notes* CMS AN-2015-222 (2015).
- [32] B. Clerbaux, W. Fang, and R. Goldouzian. *HEEP selection efficiency scale factor study for CMS with the full 2016 data sample*. CMS Note 2017/077. CERN, 2017.
- [33] B. Clerbaux, W. Fang, and R. Goldouzian. *HEEP selection efficiency scale factor study for CMS with the full 2017 data sample*. CMS Note 2018/143. CERN, 2018.
- [34] D. Bourilkov and G. Daskalakis. *PDF Uncertainties for Z' searches at 13 TeV with Electron Pair or Muon Pair Final States*. CMS Note 2016/053. CERN, 2016.
- [35] Abbiendi et al. “Search for High-Mass Resonances Decaying to Muon Pairs in pp Collisions at $\sqrt{s} = 13$ TeV with the full 2016 data set of 37 fb^{-1} and combination with 2015 result”. In: *CMS Notes* CMS-AN-2016-391 (2016).
- [36] S. Folgueras, N. Neumeister, and J.-F. Schulte. *Statistical Analysis for a Search for a Narrow Resonance*. CMS Note 2016/307. CERN, 2016.
- [37] S. Schmitz et al. “Statistical Inference in a Search for a Narrow Resonance”. In: *CMS Notes* CMS-AN-2012-185 (2012).
- [38] V. Giakoumopoulou et al. “Z' stats meeting”. In: *CMS Indico Pages* (2015). URL: <https://indico.cern.ch/event/387187/>.
- [39] Nicholas Metropolis et al. “Equation of State Calculations by Fast Computing Machines”. In: *The Journal of Chemical Physics* 21.6 (1953), pp. 1087–1092. DOI: <http://dx.doi.org/10.1063/1.1699114>. URL: <http://scitation.aip.org/content/aip/journal/jcp/21/6/10.1063/1.1699114>.

- [40] W. K. Hastings. “Monte Carlo sampling methods using Markov chains and their applications”. In: *Biometrika* 57.1 (1970), pp. 97–109. DOI: 10.1093/biomet/57.1.97. eprint: <http://biomet.oxfordjournals.org/content/57/1/97.full.pdf+html>. URL: <http://biomet.oxfordjournals.org/content/57/1/97.abstract>.
- [41] The CMS Collaboration. “Search for high-mass resonances in dilepton final states in proton-proton collisions at fb^{-1} TeV”. In: *CMS Physics Analysis Summaries* CMS-PAS-EXO-16-047 (2016).
- [42] The CMS Collaboration. “Search for high mass resonances in dielectron final state”. In: *CMS Physics Analysis Summaries* CMS-PAS-EXO-18-006 (2018).
- [43] The CMS Collaboration. “Search for high-mass resonances in dilepton final states in proton-proton collisions at fb^{-1} TeV”. In: *Journal of High Energy Physics* 2018.6 (2018), p. 120. ISSN: 1029-8479. DOI: 10.1007/JHEP06(2018)120. URL: [https://doi.org/10.1007/JHEP06\(2018\)120](https://doi.org/10.1007/JHEP06(2018)120).
- [44] The ATLAS Collaboration. “Search for new high-mass phenomena in the dilepton final state using proton-proton collisions $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *ATLAS Conference Notes* ATLAS-CONF-2017-027 (2017).
- [45] Cen Zhang and Scott Willenbrock. “Effective-Field-Theory Approach to Top-Quark Production and Decay”. In: *Phys. Rev.* D 83 (2011), p. 034006. DOI: 10.1103/PhysRevD.83.034006. arXiv:1008.3869 [hep-ph].
- [46] Gauthier Durieux, Fabio Maltoni, and Cen Zhang. “Global approach to top-quark flavor-changing interactions”. In: *Phys. Rev. D* 91 (2015), p. 074017. DOI: 10.1103/PhysRevD.91.074017. arXiv:1412.7166 [hep-ph].
- [47] B. Grzadkowski et al. “Dimension-Six Terms in the Standard Model Lagrangian”. In: *JHEP* 10 (2010), p. 085. DOI: 10.1007/JHEP10(2010)085. arXiv:1008.4884 [hep-ph].
- [48] Peter L. Cho and Elizabeth H. Simmons. “Searching for G3 in $t\bar{t}$ production”. In: *Phys. Rev. D* 51 (1995), p. 2360. DOI: 10.1103/PhysRevD.51.2360. arXiv:hep-ph/9408206 [hep-ph].
- [49] Victor Mukhamedovich Abazov et al. “Combination of searches for anomalous top quark couplings with 5.4 fb^{-1} of $p\bar{p}$ collisions”. In: *Phys. Lett.* B713 (2012), pp. 165–171. DOI: 10.1016/j.physletb.2012.05.048. arXiv:1204.2332 [hep-ex].
- [50] Morad Aaboud et al. “Probing the Wtb vertex structure in t -channel single-top-quark production and decay in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector”. In: *JHEP* 04 (2017), p. 124. DOI: 10.1007/JHEP04(2017)124. arXiv:1702.08309 [hep-ex].
- [51] Morad Aaboud et al. “Measurement of the W boson polarisation in $t\bar{t}$ events from pp collisions at $\sqrt{s} = 8$ TeV in the lepton+jets channel with ATLAS”. In: *Eur. Phys. J.* C77.4 (2017), p. 264. DOI: 10.1140/epjc/s10052-017-4819-4. arXiv:1612.02577 [hep-ex].
- [52] Vardan Khachatryan et al. “Search for anomalous Wtb couplings and flavour-changing neutral currents in t -channel single top quark production in pp collisions at $\sqrt{s} = 7$ and 8 TeV”. In: *JHEP* 02 (2017), p. 028. DOI: 10.1007/JHEP02(2017)028. arXiv:1610.03545 [hep-ex].
- [53] Vardan Khachatryan et al. “Measurement of the W boson helicity in events with a single reconstructed top quark in pp collisions at $\sqrt{s} = 8$ TeV”. In: *JHEP* 01 (2015), p. 053. DOI: 10.1007/JHEP01(2015)053. arXiv:1410.1154 [hep-ex].
- [54] CMS Collaboration. “Search for Anomalous Top Chromomagnetic Dipole Moments from angular distributions in $t\bar{t}$ Dileptonic events at $\sqrt{s} = 7$ TeV with the CMS detector”. In: *CMS-PAS-TOP-14-005* (2014).
- [55] Victor Mukhamedovich Abazov et al. “Search for flavor changing neutral currents via quark-gluon couplings in single top quark production using 2.3 fb^{-1} of $p\bar{p}$ collisions”. In: *Phys. Lett.* B693 (2010), pp. 81–87. DOI: 10.1016/j.physletb.2010.08.011. arXiv:1006.3575 [hep-ex].
- [56] T. Aaltonen et al. “Search for top-quark production via flavor-changing neutral currents in $W+1$ jet events at CDF”. In: *Phys. Rev. Lett.* 102 (2009), p. 151801. DOI: 10.1103/PhysRevLett.102.151801. arXiv:0812.3400 [hep-ex].
- [57] Georges Aad et al. “Search for single top-quark production via flavour-changing neutral currents at 8 TeV with the ATLAS detector”. In: *Eur. Phys. J. C* 76 (2016), p. 55. DOI: 10.1140/epjc/s10052-016-3876-4. arXiv:1509.00294 [hep-ex].
- [58] T Arndt, C.D Pardos, and A Meyer. “Measurement of the 2016 Trigger Efficiencies for a dilepton selection for a $t\bar{t}$ analysis”. In: *AN-2016-392* (2016).

- [59] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideMuonIdRun2>.
- [60] “Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV”. In: (2010).
- [61] <https://twiki.cern.ch/twiki/bin/view/CMS/MuonWorkInProgressAndPagResults>.
- [62] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/RochcorMuon>.
- [63] <https://twiki.cern.ch/twiki/bin/view/CMS/EgammaIDRecipesRun2>.
- [64] Vardan Khachatryan et al. “Performance of Electron Reconstruction and Selection with the CMS Detector in Proton Proton Collisions at 8 TeV”. In: *JINST* 10 (2015), P06005. DOI: 10.1088/1748-0221/10/06/P06005.
- [65] <https://twiki.cern.ch/twiki/bin/view/CMS/EGMRegression>.
- [66] <https://twiki.cern.ch/twiki/bin/view/CMS/EGMSmearer>.
- [67] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID>.
- [68] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation80XReReco>.
- [69] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETRun2Corrections>.
- [70] <https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFiltersRun2>.
- [71] Michal Czakon et al. “Top-pair production at the LHC through NNLO QCD and NLO EW”. In: *JHEP* 10 (2017), p. 186. DOI: 10.1007/JHEP10(2017)186. arXiv:1705.04105 [hep-ph].
- [72] “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”. In: *JHEP* 1107, 049 (2011). arXiv:1105.5661 [hep-ex].
- [73] CMS Collaboration. “Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 7$ TeV”. In: *JHEP* 11 (2012), p. 067. DOI: 10.1007/JHEP11(2012)067. arXiv:1208.2671 [hep-ex].
- [74] Serguei Chatrchyan et al. “Measurement of the $t\bar{t}$ production cross section in the dilepton channel in pp collisions at $\sqrt{s} = 8$ TeV”. In: *JHEP* 02 (2014), p. 024. DOI: 10.1007/JHEP02(2014)024. arXiv:1312.7582 [hep-ex].
- [75] CMS Collaboration. *Measurement of the top quark pair production cross section in proton-proton collisions at $\sqrt{s} = 13$ TeV with the CMS detector*. CMS Physics Analysis Summary CMS-PAS-TOP-15-003. 2015.
- [76] Bruce H. Denby. “The use of neural networks in high-energy physics”. In: *Neural Comput.* 5 (1993), p. 505. DOI: 10.1162/neco.1993.5.4.505.
- [77] Albert M Sirunyan et al. “Measurement of the production cross section for single top quarks in association with W bosons in proton-proton collisions at $\sqrt{s} = 13$ TeV”. In: *Submitted to: JHEP* (2018). arXiv:1805.07399 [hep-ex].
- [78] Federico Demartin et al. “tWH associated production at the LHC”. In: *Eur. Phys. J.* C77.1 (2017), p. 34. DOI: 10.1140/epjc/s10052-017-4601-7.
- [79] https://cms-condbd.cern.ch/cmsDbBrowser/diff/Prod/gts/80X_mcRun2_asymptotic_2016_TrancheIV_v6/80X_mcRun2_asymptotic_end2016_forEGM_v0.
- [80] <https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation80XReReco>.
- [81] *CMS Luminosity Measurements for the 2016 Data Taking Period*. Tech. rep. CMS-PAS-LUM-17-001. Geneva: CERN, 2017.
- [82] Vardan Khachatryan et al. “Measurement of the $t\bar{t}$ production cross section using events in the $e\mu$ final state in pp collisions at $\sqrt{s} = 13$ TeV”. In: *Eur. Phys. J.* C77 (2017), p. 172. DOI: 10.1140/epjc/s10052-017-4718-8. arXiv:1611.04040 [hep-ex].
- [83] Celine Degrande et al. “UFO - The Universal FeynRules Output”. In: *Comput. Phys. Commun.* 183 (2012), p. 1201. DOI: 10.1016/j.cpc.2012.01.022. arXiv:1108.2040 [hep-ph].
- [84] Adam Alloul et al. “FeynRules 2.0 - A complete toolbox for tree-level phenomenology”. In: *Comput. Phys. Commun.* 185 (2014), p. 2250. DOI: 10.1016/j.cpc.2014.04.012. arXiv:1310.1921 [hep-ph].
- [85] Johan Alwall et al. “MadGraph/MadEvent v4: The New Web Generation”. In: *JHEP* 09 (2007), p. 028. DOI: 10.1088/1126-6708/2007/09/028. arXiv:0706.2334 [hep-ph].

- [86] J. Alwall et al. “The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations”. In: *JHEP* 07 (2014), p. 079. DOI: 10.1007/JHEP07(2014)079. arXiv:1405.0301 [hep-ph].
- [87] Michal Czakon and Alexander Mitov. “Top++: A Program for the Calculation of the Top-Pair Cross-Section at Hadron Colliders”. In: *Comput. Phys. Commun.* 185 (2014), p. 2930. DOI: 10.1016/j.cpc.2014.06.021. arXiv:1112.5675 [hep-ph].
- [88] Nikolaos Kidonakis. “Theoretical results for electroweak-boson and single-top production”. In: *PoS DIS2015* (2015), p. 170. arXiv:1506.04072 [hep-ph].
- [89] Diogo Buarque Franzosi and Cen Zhang. “Probing the top-quark chromomagnetic dipole moment at next-to-leading order in QCD”. In: *Phys. Rev.* D91.11 (2015), p. 114010. DOI: 10.1103/PhysRevD.91.114010.
- [90] Cen Zhang. “Single Top Production at Next-to-Leading Order in the Standard Model Effective Field Theory”. In: *Phys. Rev. Lett.* 116 (2016), p. 162002. DOI: 10.1103/PhysRevLett.116.162002. arXiv:1601.06163 [hep-ph].
- [91] The CMS Collaboration. “Search for new physics via top quark production in dilepton final state at 13 TeV”. In: *CMS Physics Analysis Summaries* CMS-PAS-TOP-17-020 (2018).
- [92] <https://twiki.cern.ch/twiki/bin/view/CMS/MissingETRun2RecoilCorrection>.