FACULTÉ
DES SCIENCES

UNIVERSITÉ LIBRE DE BRUXELLES

Search for new physics in dilepton final states at the CMS experiment

Thesis submitted by Wenxing FANG
in fulfilment of the requirements of the PhD Degree in Physics (ULB - "Docteur en Sciences") and in Particle Physics and Nuclear Physics (BUAA)
Academic year 2018-2019

Supervisors: Professor Barbara CLERBAUX (Université libre de Bruxelles)

Interuniversity Institute for High Energies (ULB-VUB)

and Professor Chengping SHEN (Beihang university)

School of Physics

Thesis jury:

Laura LOPEZ HONOREZ (Université libre de Bruxelles, Chair)
Li YUAN (Beihang University, Secretary)
Yajun MAO (Peking University)
Andrea GIAMMANCO (Université catholique de Louvain)
Reza GOLDOUZIAN (University of Notre Dame)

Beihang University
Université Libre de Bruxelles

Search for new physics in dilepton final states at the CMS experiment

Wenxing Fang

Thesis submitted for the degree of

Doctor of Philosophy

Bruxelles, 13 June 2019

Defense date:

– Private: 23th May at Beihang University, Beijing (China)
– Public: 19th June at Université Libre de Bruxelles, Brussels (Belgium)

Doctoral jury:

– Prof. Barbara CLERBAUX (Université Libre de Bruxelles, Belgium), co-promoter
– Prof. Chengping SHEN (Fudan University, China), co-promoter
– Prof. Laura LOPEZ HONOREZ (Université Libre de Bruxelles, Belgium), Chair
– Prof. Li YUAN (Beihang University, China), Secretary
– Prof. Yajun MAO (Peking University, China)
– Dr. Andrea GIAMMANCO (UCL, Belgium)
– Dr. Reza GOLDOUZIAN (University of Notre Dame, USA)

Internal Chinese readers:

– Prof. Yujie ZHANG (Beihang University, China)
– Prof. Huaxing CHEN (Beihang University, China)
– Prof. Yong BAN (Peking University, China)
– Prof. Mingshui CHEN (IHEP, China)
– Prof. Chunxu YU (Nankai University, China)

摘要

该论文介绍了利用双电子末态寻找新的重共振态和在顶夸克产生过程寻找新物理的研究。论文的第一章介绍了粒子物理中的标准模型。紧接着在第二章介绍了一些与研究相关的超标准模型。在第三章介绍了欧洲大型强子对撞机（the Large Hadron Collider，简称 LHC）和紧凑缪子线圈探测器（the Compact Muon Solenoid，简称 CMS）。在随后的一章介绍了 CMS 中粒子的重建技术和过程。最后给出这两个研究的具体介绍。

第一个研究是在双电子末态寻找重的共振态。这个新的共振态是许多超标准模型所预言的，例如大统一理论（the Grand Unified Theories，简称 GUT）和额外维理论。如果存在这种共振态，那么我们将会在双电子不变质量谱中观察到一个新的质量峰。该研究利用了 CMS 在 2016 年采集到的 35.9 fb^{-1} 和在 2017 年采集到的 41.4 fb^{-1} 的数据。采用了优化的事例选择条件以增加其对信号事例选择效率。该分析的主要本底来自 Drell-Yan 过程，该过程利用蒙特卡洛样本（MC）来模拟。对于次要的顶夸克对和类顶夸克对过程，该分析也采用 MC 来模拟，同时利用数据来对 MC 进行检查。对于喷注（jet）误判为电子的本底，该分析利用 data-driven 的方法来估计该本底的贡献。在观察研究了最终的双电子不变质量谱后发现数据的分布与标准模型的预期相符合，并没有看到新物理存在的迹象。因此，在研究的最后给出了相关的新共振态产生截面乘以衰变分支比的上限和对应的新共振态的质量下限。

第二个研究是利用双电子和双缪子末态在顶夸克产生过程中寻找新物理。由于顶夸克是基本粒子中最重的粒子，其与 Higgs 粒子和 W 玻色子有很强的耦合。因此，顶夸克在许多新物理模型中占有重要地位。该研究利用了 CMS 在 2016 年采集到的 35.9 fb^{-1} 的数据。所研究的过程包括顶夸克对（$t\bar{t}$ ）产生过程和单个顶夸克伴随一个 W 玻色子产生过程（tW）。同时，由于 $t\bar{t}$ 和 tW 过程很接近，该研究利用了多变量分析的方法去区分 $t\bar{t}$ 和 tW 过程。由于最终的数据分布和标准模型预期的分布一致，因此并没有发现新物理。最终该研究利用有效场理论给出了对可能存在的新耦合的强度的限制。

关键词: 新物理，双电子，双缪子，重共振态，顶夸克，CMS实验，有效场理论，多变量分析

Résumé

Cette thèse décrit la recherche de nouvelles résonances massives qui se désintègrent en une paires d′électrons et la recherche de nouvelle physique dans le secteur des quarks top. Le modèle standard des particules élémentaires est présentédans le premier chapitre. Ensuite, nous décrivons une sélection de théories au-delàdu modèle standard prédisant l′existence de nouvelles résonances massives, ainsi qu′une introduction à la théorie effective des champs utilisée pour la recherche de nouvelles physiques dans le secteur des quark top. Après cela, le collisionneur LHC (Large Hadron Collider) et le détecteur CMS (Compact Muon Solenoid) sont introduits, et les techniques utilisées afin de reconstruire les particules produit dans les collisions sont discutées ensuite. Finalement, deux analyses séparées sont présentées.

La première est la recherche de nouvelles résonances massives dans l′état final diélectron. Certaines théories au-delàdu modèle standard prédisent l′existence de nouvelles résonances massives pouvant se désintégrer en paires d′électrons, telles que les théories de grande unification et les théories qui introduisent des dimensions spatiales supplémentaires. L′observation d′un excès local d′événements dans le spectre de masse invariante diélectron serait la preuve de l′existence d′une nouvelle résonance massive. Les données utilisées proviennent de l′expérience CMS, et correspondent à 35.9 fb^{-1} collectés en 2016 et 41.4 fb^{-1} en 2017. La sélection d′événements est optimisée pour les électrons de haute énergie et pour éviter de perdre des événements de signal potentiels. Le processus principal est le processus Drell-Yan et il est estiméà partir de simulations. Les processus $\text{t}\bar{\text{t}}$ et $\text{t}\bar{\text{t}}$-like sont aussi important, et sont également estimés à partir de simulations. La simulation de ce bruit de fond est validée par une méthode d′analyse de données. Le dernier bruit de fond, à savoir les processus de chromodynamique quantique, est déterminé à partir des données. Après inspection du spectre de masse invariante diélectron, aucun excès significatif par rapport au bruit de fond du modèle standard n′est observé, et une limite supérieure à 95% de niveau de confiance est posée sur le rapport entre d′une part le produit de la section efficace de production d′une nouvelle résonance par son rapport de branchement en diélectron, et d′autre part ce même produit mesuréau pic du boson Z.

La deuxième analyse est la recherche de nouvelle physique dans le secteur des quarks top avec les états finaux diélectron et dimuon en utilisant les données collectées par l′expérience CMS en 2016 avec 35.9 fb^{-1}. En raison de sa masse élevée et de sa masse proche de l′énergie de brisure de la symétrie électrofaible, le quark top devrait jouer un rôle important dans plusieurs scénarios de nouvelle physique. Nous recherchons cette nouvelle physique dans la production de paires de quarks top et dans la production d′un seul quark top associéà un boson, et une analyse multivariée est utilisée pour séparer ces deux processus. Aucun écart significatif par rapport aux prédictions du modèle standard n′est observé. Les résultats sont interprétés dans le cadre d′une théorie effective des champs et les contraintes sur les couplages effectifs correspondants sont définies à un niveau de confiance de 95%.

Mots clés: nouvelle physique, diélectron, dimuon, résonances massives, quark top, Expérience CMS, théorie effective des champs, analyse multivariée.

Abstract

This thesis describes searches for new heavy resonances that decay into dielectron final state and searches for new physics in the top quark sector. The standard model of elementary particle is introduced in the first chapter. After that, a selection of theories beyond the standard model that predict the existence of new massive resonances are described together with an introduction to the effective field theory that is used to search for new physics in top quark sector. Then, the Large Hadron Collider (LHC) and the Compact Muon Solenoid (CMS) detector are introduced, and the techniques used in order to reconstruct the particles produced in the collisions are discussed afterwards. Finally, two separate analyses are presented.

The first analysis is searching for new heavy resonances using dielectron final state. As some beyond Standard Model theories predict the existence of new heavy resonances that can decay into dielectron pair, such as the grand unified theories and theories that introduce extra space-like dimensions. An observation of a local "bump" in the dielectron invariant mass spectrum will be an evidence for the existence of a new heavy resonance. The data used is from CMS experiment collected in 2016 with 35.9 fb^{-1} and in 2017 with 41.4 fb^{-1}. The event selection is optimized in order to be highly efficiency for high energy electron and avoid loosing potential signal events. The leading background is the Drell-Yan process and it is estimated from simulation. The sub-leading background is from $t\bar{t}$ and $t\bar{t}$-like processes and it is estimated from simulation also. A data-driven method is used to validate the simulation of sub-leading background. The last background from quantum chromodynamics processes is determined by data-driven approach. After having inspected the final dielectron invariant mass spectrum, no significant excess over the standard model background is observed, and upper limit at 95% confidence level is set on the ratio of production cross-section times branching ratio of a new resonance to the one at the Z boson peak.

The second analysis is the search for new physics in the top quark sector with dielectron and dimuon final states using data collected by the CMS experiment in 2016 with 35.9 fb^{-1}. Because of its high mass and close to electroweak symmetry breaking scale, the top quark is expected to play an important role in several new physics scenarios. The new physics in top quark pair production and in single top quark production in association with a W boson are investigated and a dedicated multivariate analysis is used to separate these two processes. 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 at 95% confidence level.

Key works: new physics, dielectron, dimuon, heavy resonances, top quark, CMS experiment, effective field theory, multivariate analysis.

Acknowledgements

This thesis would have not been possible without the contributions of many people who helped me during the five years of my doctoral studies.

Firstly, I would like to greatly thank Chengping Shen who is my PhD supervisor at Beihang University and supervised the work presented here. We met each other in 2014 for the first time, he was very friendly and enthusiastic. He inspired my interest in particle physics and opened the door to do a PhD thesis. He gave me great guidance, great support, and countless advices during my PhD study in Beihang University. It is him who encouraged and supported me to be a joint PhD between Beihang University and Université Libre de Bruxelles (ULB). Besides, he provided me great help and suggestions when I was searching for a job after my PhD.

Secondly, I would give great thanks to Barbara Clerbaux who is my PhD supervisor at ULB and supervised the work presented here. She is very kind, thoughtful, and supportive. She is expert in CMS and she gave me countless guidance during my PhD study in CMS. When I had some questions, she always can provide me very nice explanations and answers. It is she who leaded me to the world of searching for new physics in CMS. She also cared about my living at Brussels and provided me help without hesitation when I needed.

Then, I want to thank the people with whom I worked in searching for new physics. I want to thank Reza Goldouzian who helped me a lot both in theoretical and experimental parts of my research. Besides, I want to thank Sam Harper for the time he devoted to me in discussions and explanations. I learnt many things in these occasions. A special acknowledgement goes to Xuyang Gao who is one of my best partner of my research, we worked together efficiently and pleasantly. In addition, a great thank goes to Aidan Randle-Conde who helped me a lot at the beginning of my $Z^{'}$ search study. I would like thank to Laurent Thomas from whom I took over the high P_T electron selection efficiency study in $Z^{'}$ searching. I want to thank Giuseppe Fasanella who did a very nice work in the $Z^{'}$ search team, he is friendly and provided me a very nice Latex template for my PhD thesis.

Moreover, I want to thank the people in IIHE. Firstly, I would like to thank Laurent Favart who is director of IIHE from ULB, we met each other at first time in Beihang University in 2014. He is gentle and friendly as well as taking care of my living in ULB. Secondly, I would thank Pascal Vanlaer who is enthusiastic, easy going and willing to help when I had some questions, he gave me useful comments on my V_{tx} phenomenological study. Then, I want to thank Audrey Terrier who is the secretary at IIHE. She is very kind and helped a lot in my accommodation and living at ULB. Finally, I would like thank to IIHE IT team who works very hard in maintaining and upgrading the IIHE computer cluster which is easy to use and has very high computing efficiency.

A special thank goes to my office mates, Amandeep and Diego, we get alone very well and I wish them all the best for their future.

Last but not least, I would like to thank my family for the great support and understanding during my PhD study.

Although it is impossible to name everyone here, I would like to thank all of you who helped me.

Contents

Summary of the thesis in Chinese
中文概述

该部分简要地概述了本论文中的主要内容。其中包括对标准模型和超标准模型的介绍，同时还介绍了 LHC 和 LHC 上面的四个主要实验。接着着重地介绍了 CMS 探测器的组成和性能。最后分别介绍了利用双电子末态去寻找重的玻色子的研究课题和在顶夸克（top）产生过程寻找新物理的研究课题。

标准模型和超标准模型的介绍

众所周知，标准模型 [1, 2, 3] 是描述基本粒子性质以及粒子间相互作用机制的理论。基本粒子可以分成构成物质的粒子以及传递相互作用的粒子。其中构成物质的基本粒子包括 6 种轻子：电子（e）、电子中微子（ν_e）、缪子（μ）、缪子中微子（ν_μ）、τ 和 τ 中微子（ν_τ），以及 6 种夸克：上夸克（u）、下夸克（d）、粲夸克（c）、奇异夸克（s）、顶夸克（t）和底夸克（b）。所有这些轻子都有相应的反粒子，此外每种夸克还带有有3种颜色量子数（R、G、B）。传递相互作用的粒子包括传递电磁相互作用的光子（γ）、传递强相互作用的胶子（gluon）和传递弱相互作用的 W 玻色子和 Z 玻色子。除了以上的基本粒子外，标准模型中预言的使基本粒子获得质量的希格斯粒子（Higgs）在 2012 年最终被大型强子对撞机（LHC）上的 ATLAS 实验组和 CMS 实验组同时发现 [4, 5]。至此，标准模型中的基本粒子都已被找到，见图1。此外，截止目前粒子物理实验的测量结果都与标准模型的预言相符。因此，标准模型取得了巨大的成功。关于标准模型的更加具体的介绍可以参见第 1 章，该章节还详细的介绍了 Drell-Yan 过程 [6] 和有效场理论（Effective Field Theory，简称 EFT），这是因为它们在本论文的研究中有着重要作用。

虽然标准模型取得了巨大的成功，但是它也存在一些缺陷。例如，在天文学和宇宙学界科学家们通过实验观测的结果普遍认为存在暗物质和暗能量 [7, 8]。其中宇宙中的暗物质占约 25%，暗能量占约 70% 而可见的物质只占约 5%。但是标准模型并未涉及有关暗物质和暗能量的预言，当然也没能提供暗物质和暗能量的候选者。此外，实验上观测到了中微子在传播过程中它的味可以发生变化（例如从 ν_e 到 ν_μ、ν_μ 到 ν_τ 等）即所谓的中微子振荡现象 [9, 10]。中微子振荡现象的存在表明中微子的质量是非零的，这与标准模型中中微子质量为零的假设不相符。众所周知，我们生活在以正物质组成的世界中，那么反物质去哪儿了呢？一般认为在宇宙大爆炸时，正反物质是成对产生的。因此宇宙中应该存在相同的正物质和反物质。标准模型在 Cabibbo-Kobayashi-Maskawa （CMK）矩阵中引入了电荷宇称（CP）破坏的参数，但是该参数远不能解释目前所观测到的正反物质不对称现象。标准模型除了不能很好地解释以上这些现象外，它也存在一些瑕疵。例如，由于标准模型没有考虑引力的相互作用，因此可以假设标准模型直到普朗克能标（10^{19} GeV）都是有效的。但为了得到实验上发现的质量为 125 GeV 的 Higgs 粒子，在标准模型中则需要通过将两个 10^{38} 量级的大数相减得到一个 10^4 量级的数，这在理论上是可以成立的，但看上去非常不自然，这被称做 Higgs 质量 fine-tuning 问题。如果标准模型只到 TeV 能量有效，那么该问题就不存在了。此外，标准模型里存在 19（若加上中微子的 7 个自由参数，则总共为 26 个自

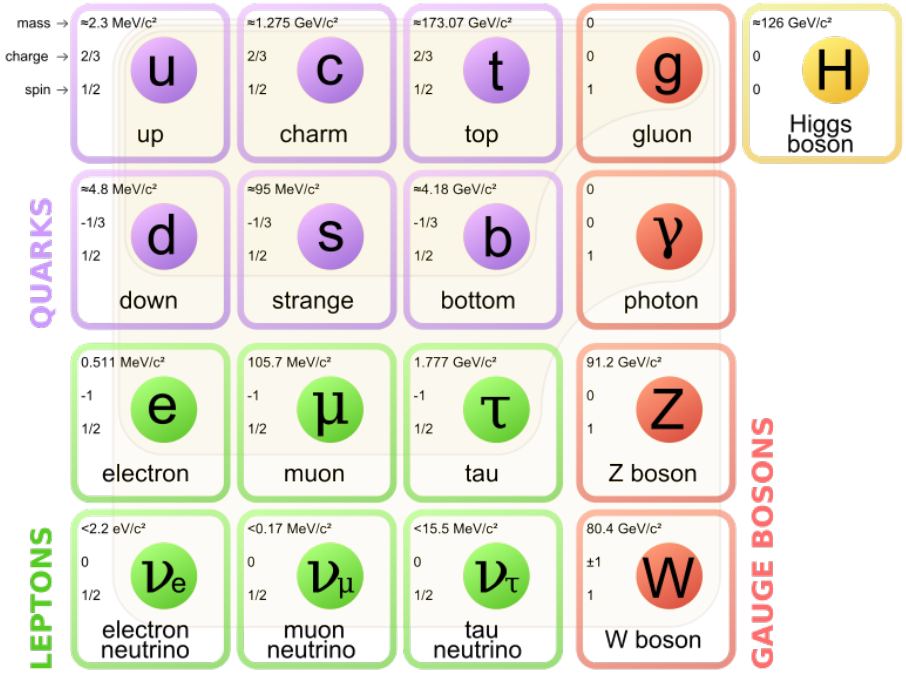

Figure 1: 标准模型中的基本粒子。

由参数） 个需要通过实验测量确定的自由参数，这使得它看起来不够简洁。关于标准模型缺陷更加详细的描述可参见 2.1。

正因如此，许多超越标准模型物理学的理论被提了出来。例如超对称模型（SUSY） [11]，在该模型中每一种基本粒子都有一种被称为超对称伙伴（Superpartner）的粒子与之匹配，超对称伙伴的自旋与原粒子相差 1/2 （也就是说玻色子的超对称伙伴是费米子，费米子的超对称伙伴是玻色子），两者质量相同，各种耦合常数间也有着十分明确的关联。这个模型可以很好的消除 Higgs 质量 fine-tuning 的问题以及等级问题（Hierarchy Problem）即为什么在电弱统一能标与大统一或 Planck 能标之间存在高达十几个数量级的差别。同时该模型还为暗物质提供了候选者。此外还有大统一理论（GUT） [12, 13, 14]，该理论想尝试将电磁相互作用、弱相互作用和强相互作用统一起来。对于额外维理论 [15, 16]，该理论想通过引入额外的维度来解释为什么引力的相互作用强度比另外三种相互作用弱 10^{30} 个量级。以上这些超标准模型都预言存在一个新的重玻色子，其可以通过双电子过程衰变。这也是本论文利用双电子末态寻找新的高质量玻色子的动机所在。关于这部分更加详细的介绍可参见 2.2。

如果新物理的能标能在实验中达到，那么新的粒子就能够被产生，可能被直接发现。反之，则需要通过间接的方式去寻找新物理。由于顶夸克是已知基本粒子中最重的粒子以及它与 Higgs 和 W 玻色子都有很强的耦合。因此顶夸克在许多超标准模型中占有很重要的地位。本论文的第二个课题对顶夸克产生过程中可能存在的新物理进行了探索研究。其中包括顶夸克对的产生过程以及单个顶夸克伴随 W 玻色子的产生过程。为了提高该分析对新物理的敏感度，该研究采用了多变量分析方法来区分本底事例和信号事例，同时使用有效场的方法对各种可能存在的新耦合做出与理论模型无关的限制。更多的相关介绍可参见 2.3。

LHC 和 CMS 的介绍

欧洲大型强子对撞机（LHC）是位于法国和瑞士边界，周长为 27 km，位于地下 50 至 150 米之间的质子-质子对撞机，见图 2 的左部分。其质心对撞能量达到了世界最高，在 2015 年到 2018 年该值为 13 TeV。LHC 上主要有四个实验组，见图 2 的右部分。其中包括大型离子对撞器（ALICE）其主要通过铅离子与铅离子对撞或铅离子与质子对撞研究夸克胶子等离子体的性质。还有 LHC 底夸克探测器（LHCb）其主要研究CP破坏、底夸克的性质等。超环面仪器（ATLAS）其是一个综合的粒子探测器，主要研究内容包括精确测量标准模型中的自由参数、寻找 Higgs （在 2012 年已经被发现）、寻找超越标准模型的新物理。紧凑缪子线圈（CMS）其作用与目的和 ATLAS 一致。ATLAS 和 CMS 两个探测器的存在使得各自的实验结果能被互相检查或确认。

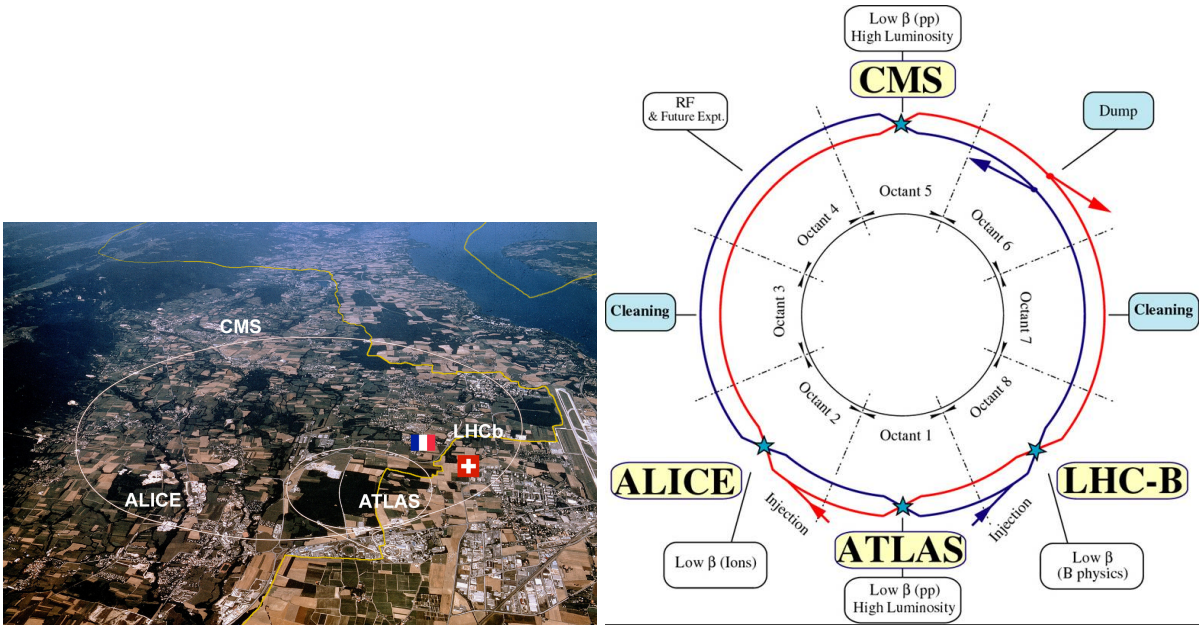

Figure 2: LHC的全貌（左）和LHC 上的四个主要实验（右） [17]。

由于一个质子束团中约有 10^{11} 个质子，因此当质子束团与质子束团发生对撞时，可能产生多个质子-质子对撞顶点，这种现象被称作“pile-up”，“pile-up”现象可见图 3 的左部分，2016 年“pile-up”的分布情况可见图 3 的中间部分以及 2017 年“pile-up”的分布情况可见图 3 的右部分。

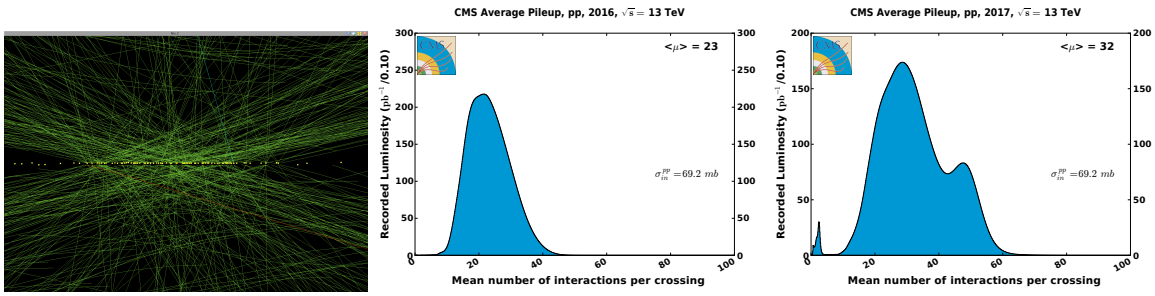

Figure 3: “pile-up”现象（左图），2016 年“pile up”的分布情况（中间）和 2017 年“pile up”的分布情况（右图 [18]。

从 2015 年到 2018 年 LHC 产生了亮度为 156 fb^{-1} 的质子-质子对撞数据，实现了预期的 150 fb^{-1} 目标。在 2024 年以后 LHC 将升级为 HL-LHC 即高亮度 LHC 。关

于 LHC 和 HL-LHC 的取数计划可见图 4。更多的关于 LHC 的介绍可以参见 3.1。

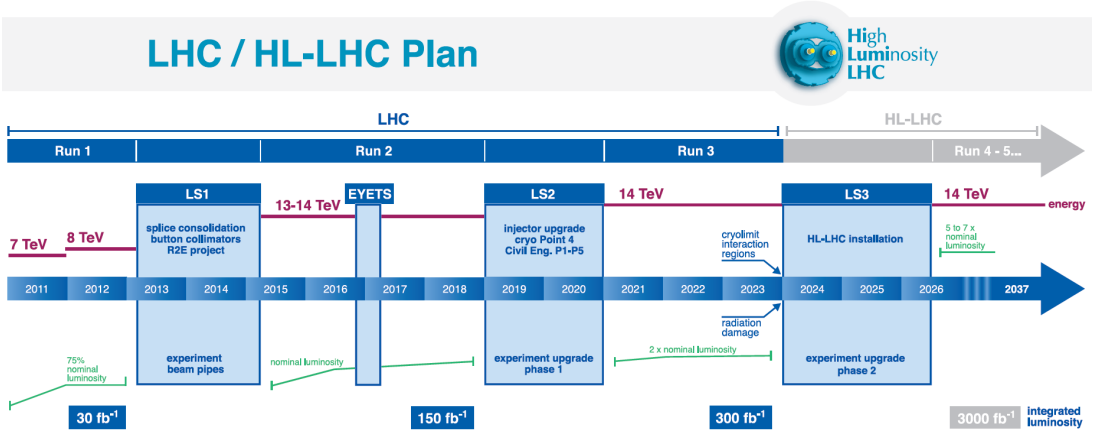

Figure 4: LHC 和 HL-LHC 的取数计划 [19]。

CMS 探测器 [20] 是一个长为 21 米，宽 15 米，高 15 米，重 14000 吨的探测器。从内到外的子探测器依次为硅像素径迹探测器、硅微条径迹探测器、电磁量能器、强子量能器、厄铁线圈以及缪子探测器。CMS 探测器的剖视图和横向截面图可见图 5 和图 6。

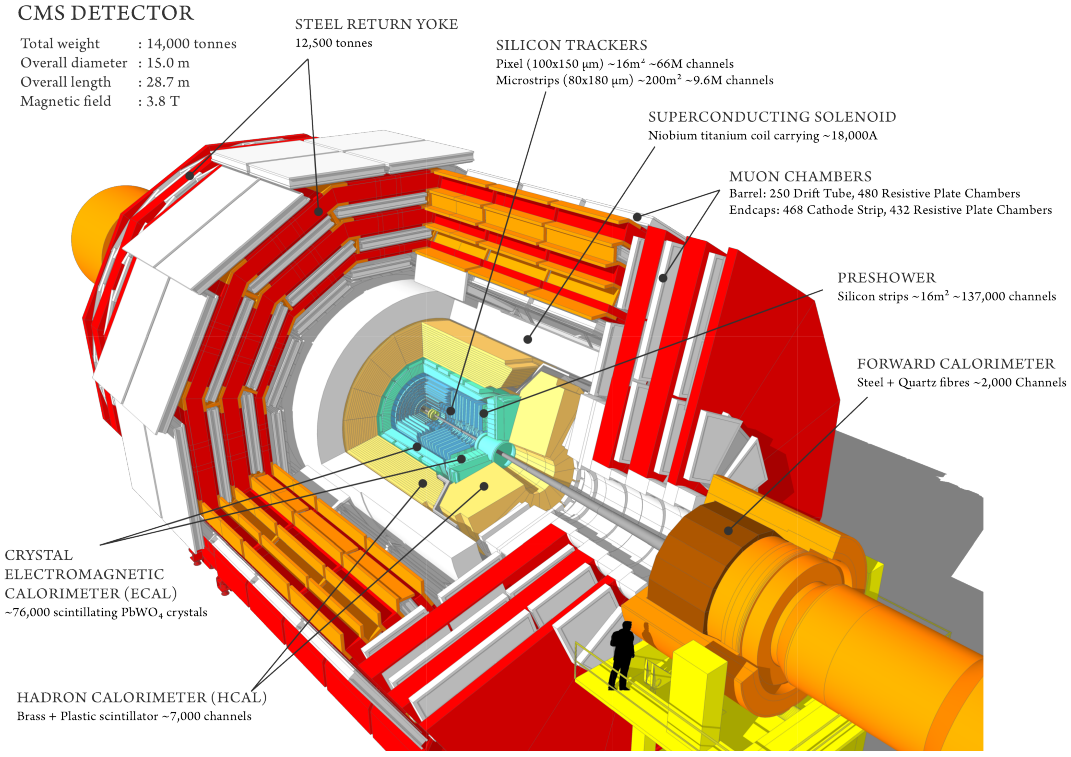

Figure 5: CMS 探测器的剖视图 [21]。

在介绍下面的内容之前有必要先介绍一下 CMS 探测器的坐标系统。CMS 的直角坐标原点为设计时质子-质子对撞的地方，Y 轴的方向向上，X 轴的方向朝向 LHC 的中心，Z 轴的方向朝向 Jura 山脉。X-Y 平面的方位角用 ϕ 表示，与 Z 轴的夹角（即极角）用 θ 表示。CMS 坐标系统定义可见图 7。赝快度 η 的定义为 $\eta = -\ln \tan\theta/2$。横

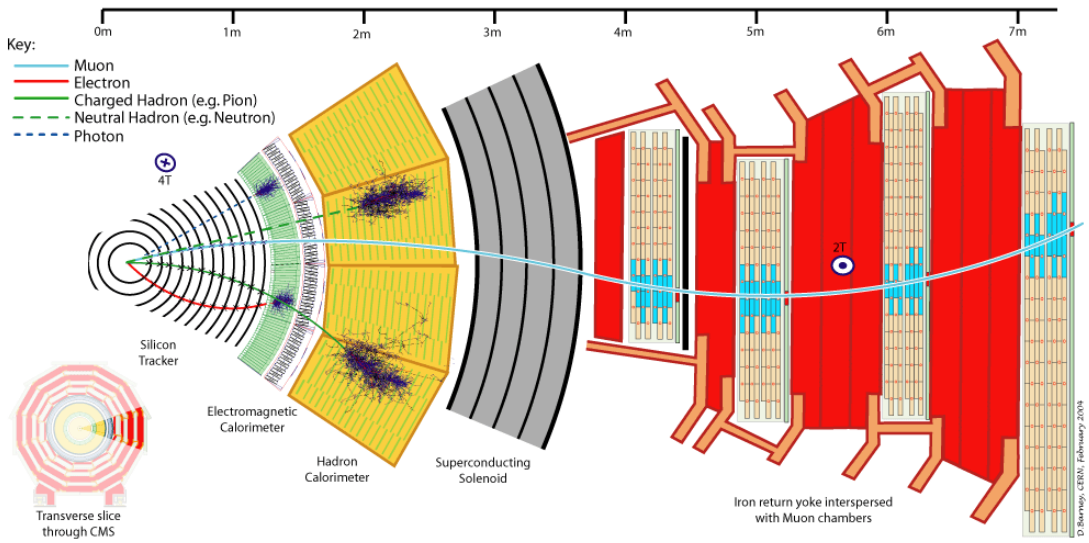

Figure 6: CMS 探测器的横向截面图 [22]。

向动量和横向能量分别用 p_T 和 E_T 表示。其中 $p_T(E_T) = p(E)\sin\theta$。在横向面上能量的不平衡值用 E_T^{miss} 或$\not{E}_T$ 表示。其值的计算公式为 $\overrightarrow{E_T}^{miss} = -\sum \overrightarrow{E_T}^{exists}$。

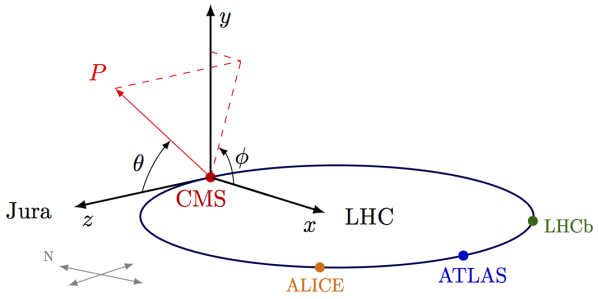

Figure 7: CMS 坐标系统定义。

下面继续介绍 CMS 的各个子探测器。CMS 的硅像素径迹探测器是最靠近输流管的子探测器，其主要目的为探测带电粒子的运动轨迹和质子-质子碰撞的顶点。每个像素的大小为$100 \times 150\ \mu m^2$，可覆盖 $|\eta|$ 到 2.5 的范围。该探测器设计为能承受瞬时亮度为$1 \times 10^{34}\ cm^{-2}s^{-1}$，对撞间隔为 25 ns 的质子-质子对撞。硅微条径迹探测器也是为了探测带点粒子的运动轨迹，根据位置的不同其分辨率在 23-52 μm 之间，同样覆盖 $|\eta|$ 到 2.5 的范围。CMS 的硅像素探测器和硅微条探测器组成了 CMS 的径迹探测器。图8 为 CMS 径迹探测器纵向截面图的四分之一，其中黄色部分为 CMS 的硅像素探测器，粉色部分为 CMS 硅微条探测器。

CMS 的电磁量能器的作用是为了测量电子和光子的能量，其由一个桶部（Barrel）和两个端盖（Endcap）组成。桶部由 61200 块（长为 230 mm，宽为 22 mm，高为 22 mm）钨酸铅（$PbWO_4$ ）晶体组成，可覆盖赝快度从 0 到 1.479。两边端盖部分分别由 7324 块（长为 220 mm，宽为 24.7 mm，高为 24.7 mm）钨酸铅晶体组成，可覆盖 $|\eta|$ 从 1.48 到 3.0 的范围。之所以选择钨酸铅是因为其辐射长度短（$X_0 = 0.89$ cm）、辐射半径小（2.2 cm）、辐射速度快（80% 的能量在 25 ns 内释放）。为了更好地区分 π^0 和光子，CMS 在每个电磁量能器的端盖前放置了一个 preshower 探测器，覆盖 $|\eta|$ 从 1.65 到 2.6 的范围。图 9 为 CMS 电磁量能器四分之一的纵向截面图。

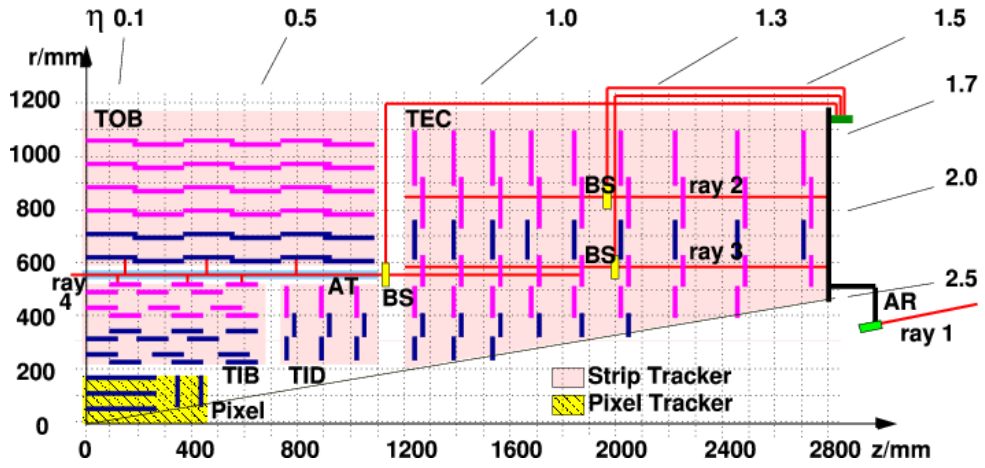

Figure 8: CMS 径迹探测器纵向截面图的四分之一 [20]。

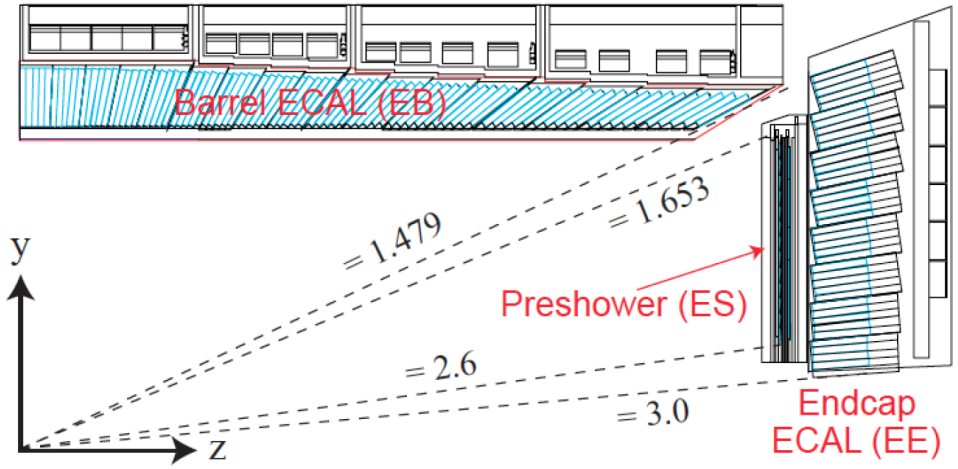

Figure 9: CMS 电磁量能器四分之一的纵向截面图 [23]。

CMS 电磁量能器的能量分辨率可以表示为 [24]：

$$\frac{\sigma}{E}^2 = \frac{S}{\sqrt{E}}^2 + \frac{N}{E}^2 + C^2 \tag{1}$$

其中，S 表示随机项，N 代表噪声项，C 代表常数项。例如对于能量在 20 到 250 GeV 的电子，利用 3×3 的晶体去测量其能量，那么对应的 S 为 0.028 $\sqrt{\text{GeV}}$，N 为 0.12 GeV，C 为 0.003。

在电磁量能器的外面是强子量能器，其目的是测量强子的能量，包括带电的和中性的强子。强子量能器有桶部（HB）、端盖（HE）、前端（HF）和外部（HO）组成。其中 HB 探测范围为 |η| 从 0 到 1.3，HE 探测范围为 |η| 从 1.3 到 3.0，HF 探测范围为 |η| 从 3.0 到 5.2，目的是为了监测质子对撞的瞬时亮度。而 HO 是放在厄铁线圈外的部分，主要是用于对 HB 的辅助测量。关于强子量能器给组成部分在 CMS 探测器中的位置可见图 10。

CMS 强子量能器对单个强子的基准能量分辨率可以表示为 [25]：

$$\frac{\sigma}{E} = \frac{X}{\sqrt{E}} \oplus 5\%, \quad \texttt{X=65\%} \ \ (\texttt{HB}) \ , \ \texttt{83\%} \ \ (\texttt{HE}) \ , \ \texttt{100\%} \ \ (\texttt{HF}) \tag{2}$$

在 CMS 强子量能器外部是 CMS 的厄铁线圈。厄铁线圈利用低温超导技术能提

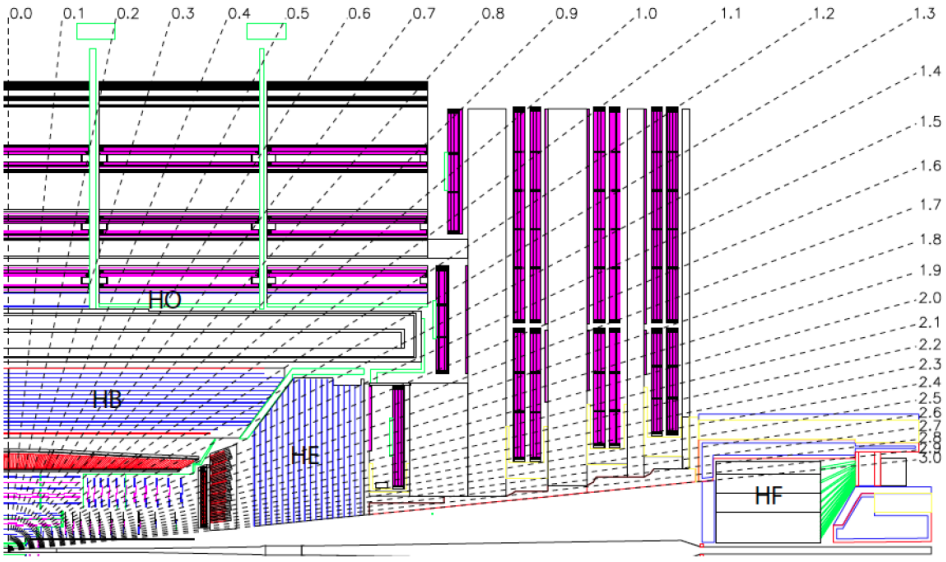

Figure 10: CMS 探测器的纵向截面图。其中分别标明了强子量能器桶部（HB）、端盖（HE）、前端（HF）以及外部（HO）的位置 [20]。

供 3.8 T 的强磁场，使得带电粒子在磁场中运动的轨迹变的弯曲，从而使得带电粒子的动量能被测量。另外厄铁线圈隔绝了几乎所有除了缪子以外的粒子进入其外部的缪子探测器。缪子探测器是 CMS 最外部的子探测器，其主要目的是探测缪子及其动量。缪子探测器由漂移管（DT）、阴极条室（CSC）和阻性板（RPC）组成。其中 DT 和 CSC 提供了很高的缪子空间分辨率，而 RPC 提供了很快的缪子时间分辨率。缪子探测器的桶部（MB）为 $|\eta|$ 从 0 到 1.2 的范围，端盖（ME）的部分为 $|\eta|$ 从 1.2 到 2.4 的范围。图 11 为 CMS 缪子探测器的纵向四分之一截面图。如果只利用缪子探测器的信息，那么对于动量约为 1 TeV的谬子，其动量分辨率在 5% 左右，如果结合 CMS 内部的径迹探测器，那么其动量分辨率可以达到 1% 到 5%。

由于 LHC 质子-质子对撞的频率为 40 MHz（周期为25 ns），每个事例的大小约为 1 MB，而 CMS 数据存储速率上限约为 1 kHz。因此 CMS 利用了两级触发（trigger）来实现对感兴趣事例进行筛选，使得不感兴趣的事例被忽略，而感兴趣的事例能被存储下来。第一级触发被称为 L1 触发，其利用电磁量能器、强子量能器以及谬子探测器来快速判断该事例中是否有能量超过阈值的电子或光子或谬子或中微子或者喷注（jet），为了缩短判断时间 L1 触发并不会使用径迹探测器的信息对光子和电子进行区分。L1 触发的反应时间为 3.2 μs，通过 L1 触发后事例的速率降为约 100 kHz。事例在通过 L1 触发后将会进入第二级 HLT 触发，HLT 将会结合各个子探测器的信息对事例进行重建包括利用径迹探测器来重建带电粒子的径迹。由于需要进行复杂的重建过程，对于每个事例 HLT 将花费约 100 ms 对其是否通过触发进行判断。最终通过 HLT 的事例速率将降为 600 Hz 左右。这些事例将被永久存储到磁盘以用于后续的物理分析。有关 CMS 探测器的介绍和触发可以参见 3.2。关于不同粒子的重建过程可以参见 4。

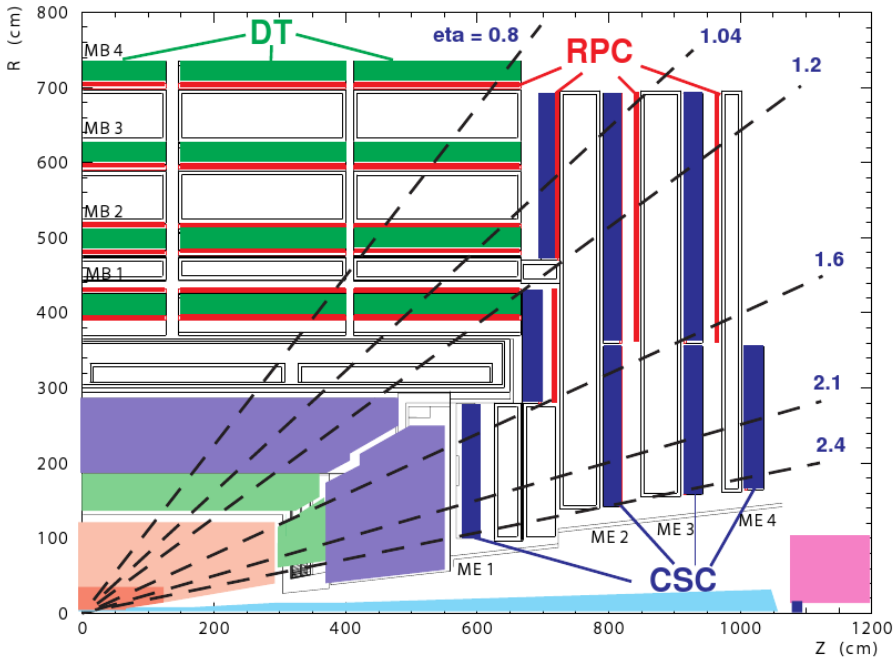

Figure 11: CMS 谬子探测器的纵向四分之一截面图 [20]。

利用双电子末态寻找重的玻色子

在介绍该物理分析之前有必要先介绍一下该物理的总体进展。 LHC 和 Tevatron 在以前就利用过双轻子末态寻找重的玻色子。例如 CMS 利用了 7 TeV [26, 27] 和 8 TeV [28, 29] 以及 2015 年的 13 TeV 加之前 8 TeV [30] 的质子质子对撞数据寻找了 Z' 。最近的 CMS 结果是来自于 2016 年 13 TeV [31] 和 2017 年 13 TeV [32] 的数据，即本论文所要介绍的研究工作。和 CMS 类似， LHC 上的 ATLAS 实验组也做了对 Z' 的寻找。其利用的数据包括： 7 TeV [33, 34]、 8 TeV [35] 和最新的 13 TeV 全部的数据 [36, 37]。在 Tevatron 上， CDF 实验组和 D0 实验组利用 1.96 TeV 的质子和反质子对撞数据对 Z' 进行了寻找 [38, 39, 40, 41, 42, 43]。

本论文的第一个研究课题是利用双电子末态寻找重的玻色子。利用的数据是来自 CMS 的 2016 年 35.9 fb^{-1} 和 2017 年 41.4 fb^{-1} 的数据。对于被选中的电子其必须在电磁量能器的桶部或端盖部分，同时其横向能量要大于 35 GeV，然后通过高能电子选择条件（即 HEEP 选择条件）。我们利用 tag 和 probe 的方法测量了HEEP 的选择效率。其中 tag 为通过 HEEP 条件的电子，并且通过触发表。probe 为普通的电子，此外 tag 和 probe 的不变质量在 Z 玻色子区间。这样做的目的是尽可能减少 probe 中的假电子成分，提高真电子的比例。最后测量到的 HEEP 对真电子的选择效率可见图 12。

对于两个电子都在端盖的事例将不会被选入，因为在高质量区间 QCD 本底对这种事例有较大贡献。最后，被选中的事例需要通过双电子触发表。优化了的事例选择条件能保证对电子具有高的选择效率以及减少非电子被鉴别为电子的误判。该分析的主要本底来源于 Drell-Yan 过程，对于该本底的估计则主要来源于蒙特卡洛（MC）模拟。次主要本底来源于顶夸克对和双玻色子过程，这些本底通过 MC 进行模拟并利用数据对其进行检查。最后一部分的本底来源于 jet 误判为电子的情况，这部分的本底通过 data-driven 的方法进行估计。图 13 为两电子都在端盖时的双电子不变质量谱（左）和质量谱积分（右）在 2016 年（上）和 2017 年（下）。从中可见 data-driven 方法还是能比较好地估计出 jet 本底。

图 14 和 15 分别为来自 2016 年和 2017 年的最终双电子不变质量谱（左）和不变质量谱积分（右）对于两电子都在桶部（上）、一个电子在桶部一个电子在端盖（中）和有一个电子在桶部（下）的情况。可惜在对得到的最终双电子不变质量谱进行研究后并没有发现明显超出标准模型预言的事例。

因此该分析利用统计学的方法给出了不同新粒子产生截面乘以衰变分支比的 95% 的置信上限。同时将该截面乘以衰变分支比上限转换为不同新粒子质量的下限。其结果可见图16。

在结合了双电子道（2016 年 35.9 fb^{-1} 加上 2017 年 41.4 fb^{-1} 的数据）和双缪子道（2016 年 36.3 fb^{-1} 的数据）后，可以给出更严格的不同模型下的 Z' 质量下限，如 Z'_{SSM}（其与标准模型中费米子的耦合和 Z 玻色子一样）的质量下限为 4.7 TeV，Z'_{ψ}（来自于 GUT 模型）的质量下限为 4.1 TeV，见图17。具体的研究过程和结果可参见 5。

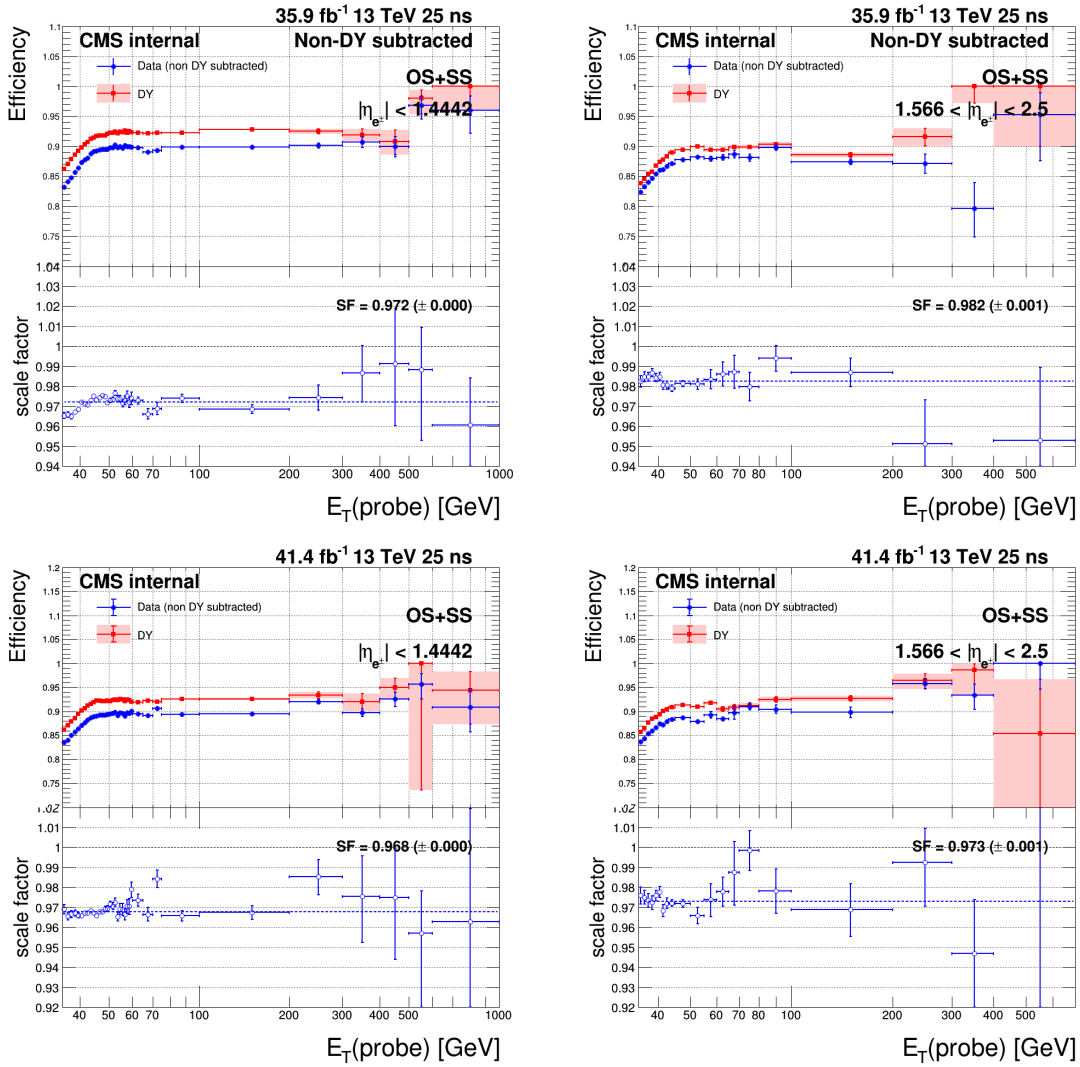

Figure 12: HEEP 条件对桶部（左侧）和端盖（右侧）电子在数据和蒙卡样本（MC）中的选择效率和 scale factor 在 2016（上面）和 2017 年（下面）。

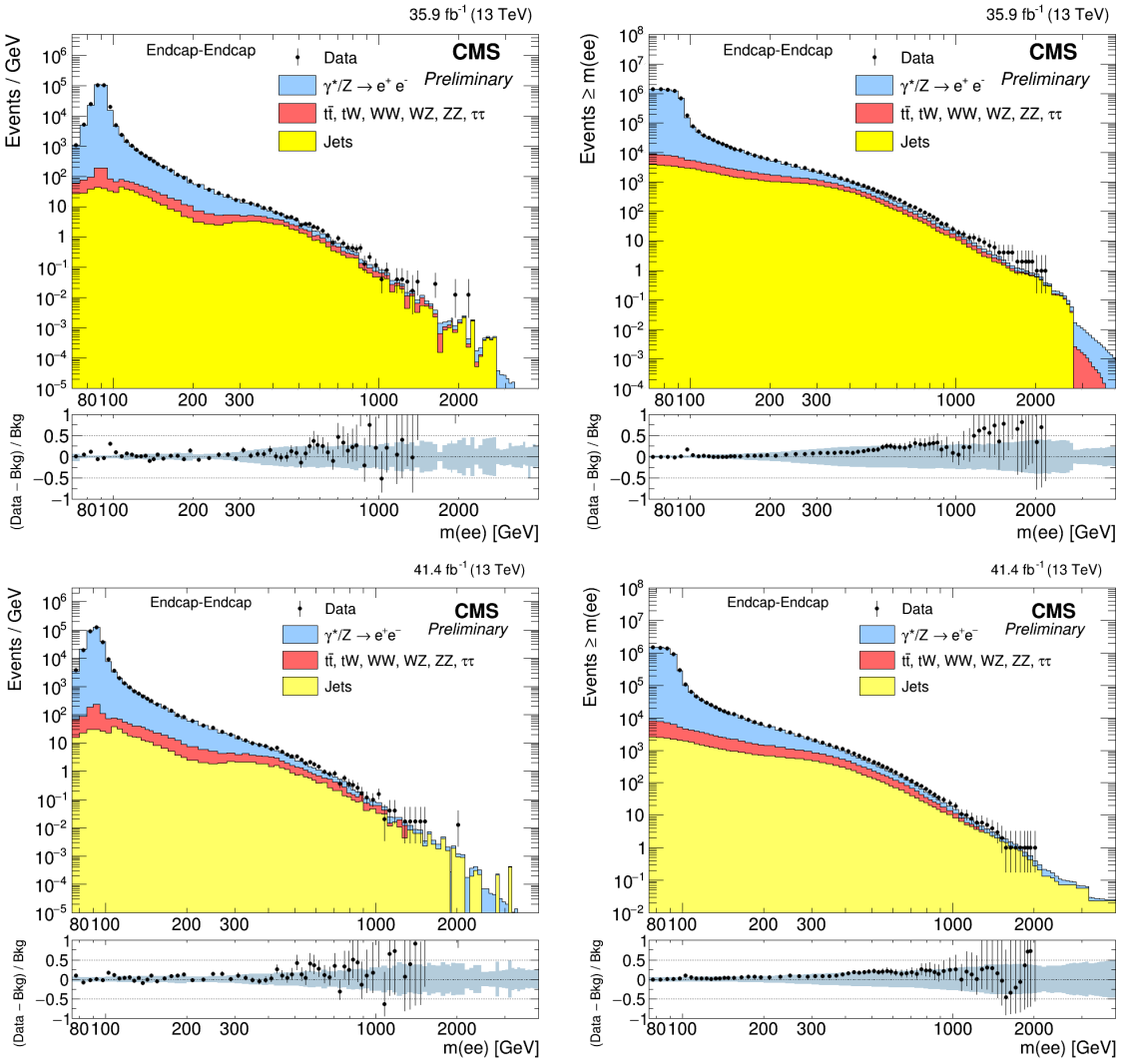

Figure 13: 两电子都在端盖时的双电子不变质量谱（左）和质量谱积分（右）在 2016 年（上）和 2017 年（下）。

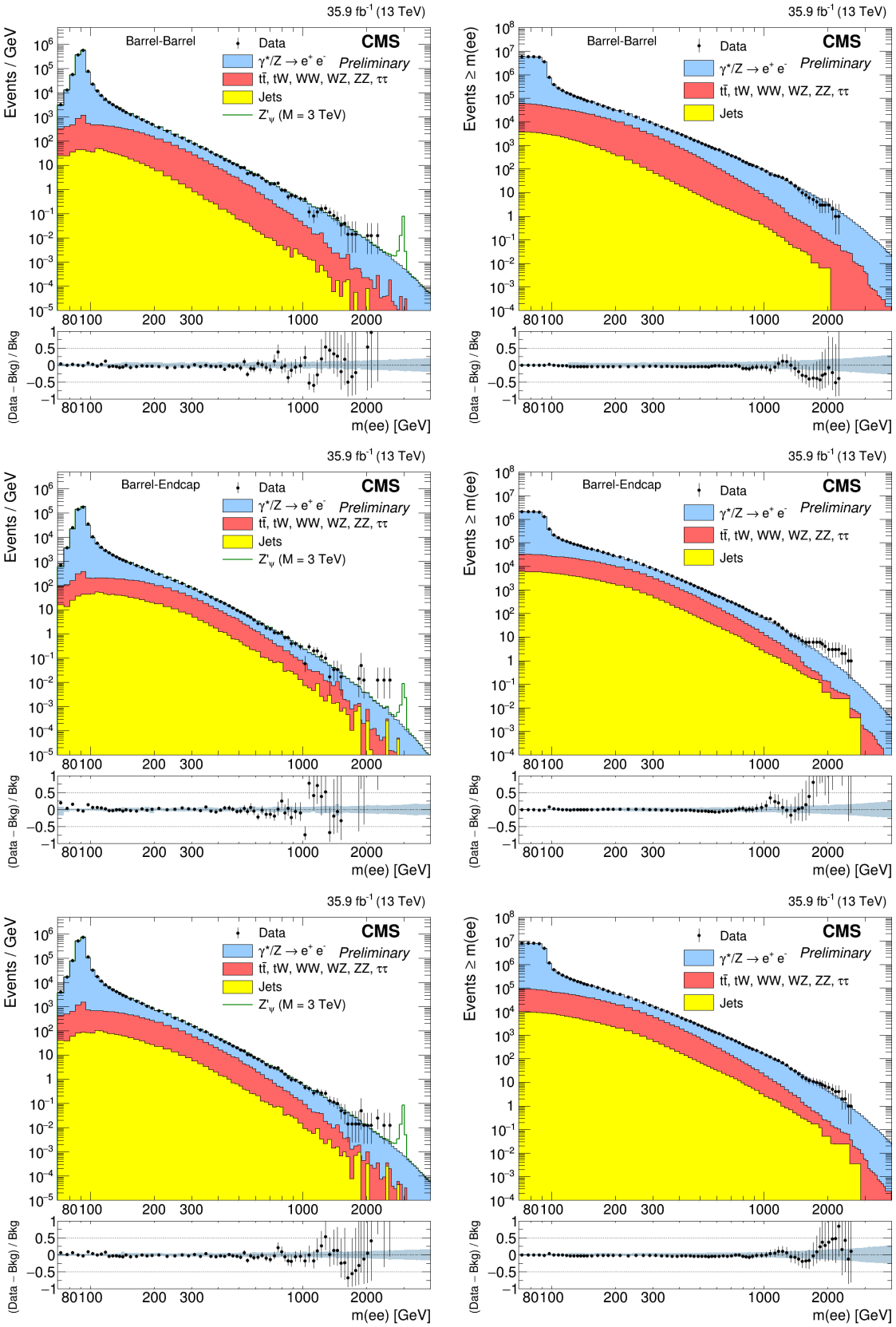

Figure 14: 2016 年的最终双电子不变质量谱（左）和不变质量谱积分（右）对于两电子都在桶部（上）、一个电子在桶部一个电子在端盖（中）和一个电子在桶部（下）的情况。

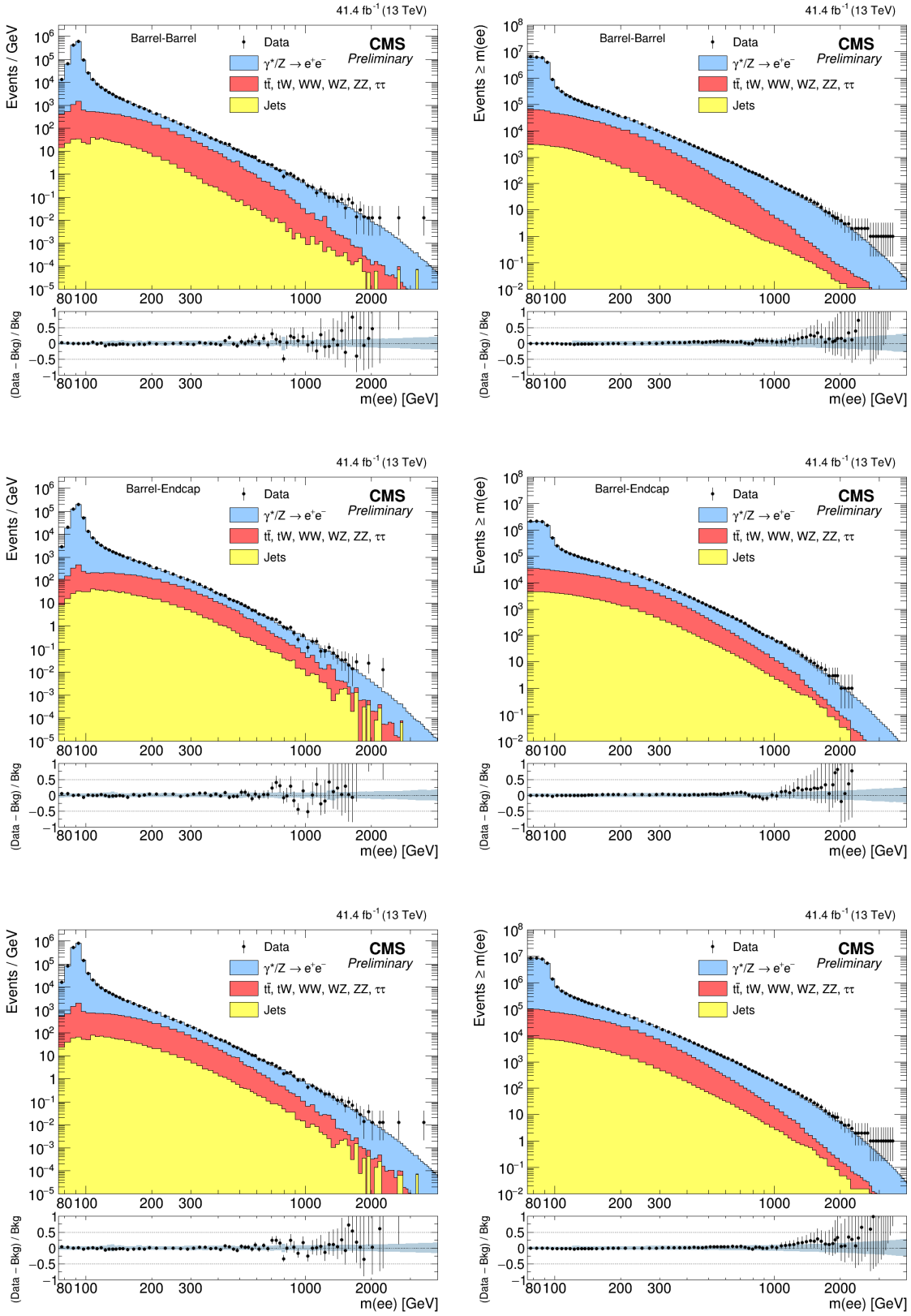

Figure 15: 2017 年的最终双电子不变质量谱（左）和不变质量谱积分（右）对于两电子都在桶部（上）、一个电子在桶部一个电子在端盖（中）和一个电子在桶部（下）的情况。

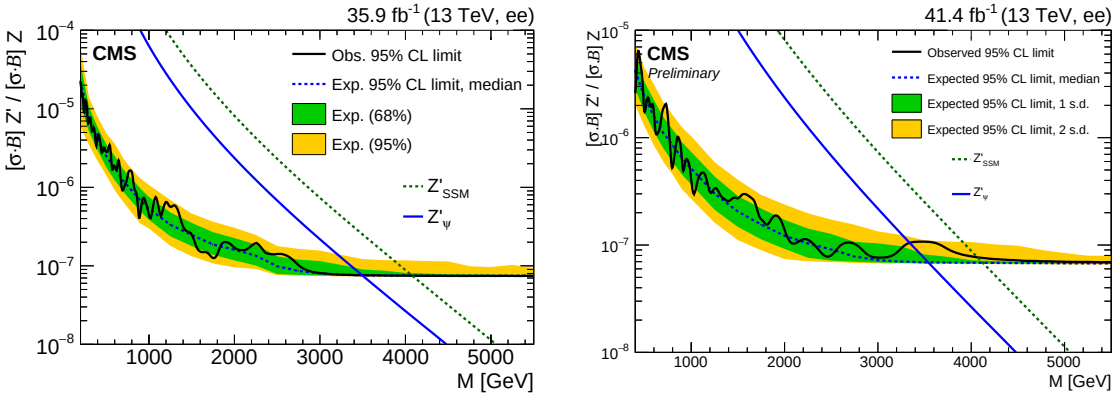

Figure 16: 观测到的和期望的 $Z^{'}$ 的产生截面乘以衰变分支比的 95% 上限在 2016 年（左）[31] 和 2017 年（右）[32]，同时给出了 $Z^{'}_{SSM}$ 和 $Z^{'}_{\psi}$ 对应的理论值。

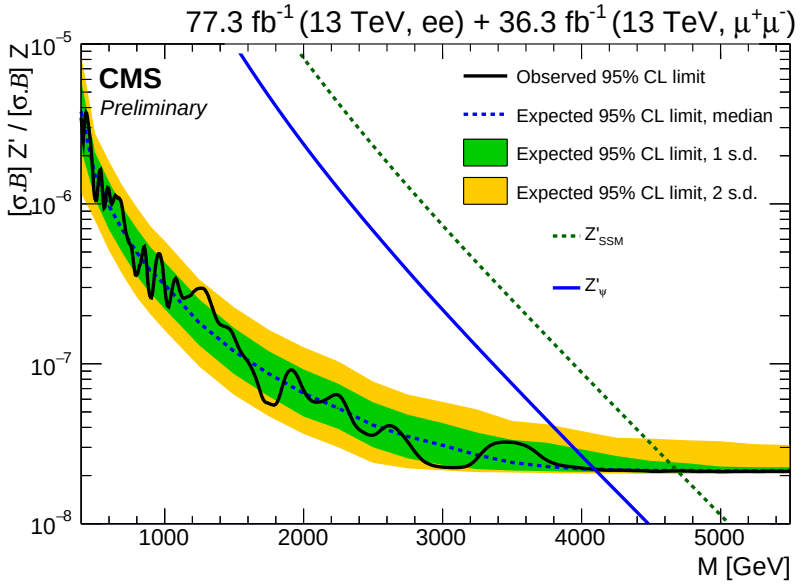

Figure 17: 在结合 2016 年双电子和双谬子道的数据以及 2017 年双电子道的数据后观测到的和期望的 $Z^{'}$ 的产生截面乘以衰变分支比的 95% 上限，同时给出了 $Z^{'}_{SSM}$ 和 $Z^{'}_{\psi}$ 对应的理论值[32]。

在 top 夸克产生过程寻找新物理

在介绍该物理分析之前有必要先介绍一下该物理的总体进展。 LHC 上的 CMS [44, 45] 和 ATLAS [46, 47] 实验组以及 Tevatron 上的 D0 [48] 实验组曾利用单顶夸克 *t*-channel 产生过程和顶夸克衰变过程寻找 Wtb 的奇异耦合。此外 CMS 曾利用 7 TeV 的数据通过顶夸克对产生过程寻找过新的夸克胶子耦合。最后， LHC [44, 49] 和 Tevatron [50, 51] 也对顶夸克参与的味改变中性流过程的进行了广泛的寻找。

本论文的第二个研究课题是通过双电子或双缪子并且伴随几个 jet（即喷注） 或 bjet（也叫 btag，为含底夸克的喷注） 的末态对顶夸克的产生过程是否存在新物理进行寻找，包括了顶夸克对的产生过程（$t\bar{t}$ ）和一个顶夸克加上一个 W 玻色子的过程（tW）。利用的数据是来自 CMS 的 2016 年 35.9 fb^{-1} 的数据。被选中的电子或缪子需要通过好电子和好缪子的鉴别，同时电子或缪子的横动量要大于 20 GeV，其中 leading 的电子或缪子的横动量要大于 25 GeV。此外，双电子或双缪子的电荷需要相反，整个事例需要通过相关的触发表。为了去除 Drell-Yan 本底过程，要求双电子或双缪子的不变质量需要在 Z 玻色子质量区间外，同时该事例的 MET 要大于 60 GeV。但在实验中发现数据的 MET 分布和蒙卡样本中 MET 的分布符合的并不好，因此采用了 data-driven 方法去估计数据中的 Drell-Yan 本底事例数。我们利用蒙特卡洛样本来模拟能产生两个真电子或真缪子的过程，例如 $t\bar{t}$ 、tW 和双玻色子过程，对于存在的 jet 误判为电子或缪子的过程，我们利用 “same-sign” 的 data-driven 方法去估计。经过事例选择后的双电子道和双缪子道中 leading 轻子的 Pt、η 和 ϕ 的分布可见图 18。双电子道和双缪子道中 sub-leading 轻子的 p_T 、η 和 ϕ 的分布可见图19。从中可见数据的分布与蒙特卡洛样本的预期一致。

最终的双电子道和双缪子道中数据和蒙卡样本在不同 jet 和 bjet 区域的事例数可见图 20。从中我们知道 tW 过程的事例主要出现在 1-jet,1-tag 区域，而 2-jet,1-tag 和 ≥2-jet,2-tag 的区域主要来自 $t\bar{t}$ 过程。

该分析主要寻找了几种可能存在的新耦合，其中包括只参与 tW 过程的 C_{tW} （此表示右手的 W 玻色子和 top 夸克及 b 夸克的相互作用），$C_{\phi q}^{(3)}$ （此与标准模型的 W 玻色子和 top 夸克及 b 夸克相互作用一致），C_{uG} （此表示胶子与 u 夸克和 top 夸克的相互作用）和 C_{cG} （此表示胶子与 c 夸克和 top 夸克的相互作用）以及只参与 $t\bar{t}$ 过程的 C_G （此表示三个胶子的相互作用）和两个过程都参与的 C_{tG} （此表示胶子与两 top 夸克的相互作用）。这些过程的费曼图可见图 21。

由于 $t\bar{t}$ 和 tW 过程末态很相似，因此该分析采用了多变量分析（神经网络）的方法来区分 $t\bar{t}$ 和 tW 事例，从而提高了该分析对新物理的敏感度。最终的神经网络输出可见图22。

可惜从图 22 中看到实验数据的分布和标准模型预测一致，并没有发现新物理的迹象。在最后该分析利用统计学的方法给出了不同耦合的 95% 的置信区间。在结合了一个类似的分析道即一个电子加一个缪子道后，最终的不同耦合的 95% 的置信区间可见图 23。该分析首次对 C_G 给出了限制，也提高了之前 C_{tG} 的结果，关于 C_{tW} ，$C_{\phi q}^{(3)}$ ，C_{uG} 和 C_{cG} 则是首次利用 tW 过程对其进行研究，也是对之前来自单顶夸克过程的研究结果的补充。关于该分析的更加具体的研究过程和结果可参见 6。

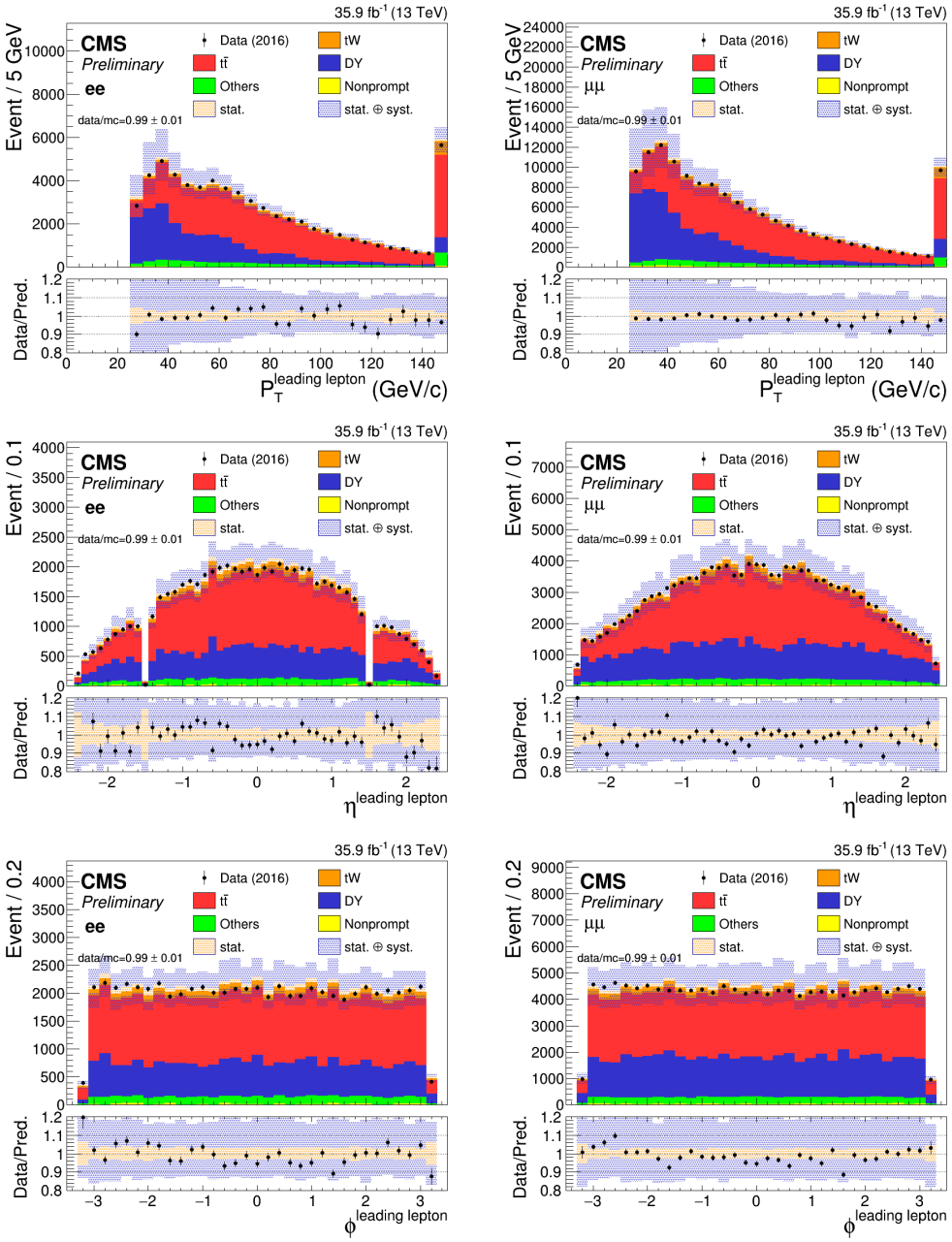

Figure 18: 双电子道（左）和双缪子道（右）中 leading 轻子的 Pt （上）、η （中）和 ϕ （下）的分布。

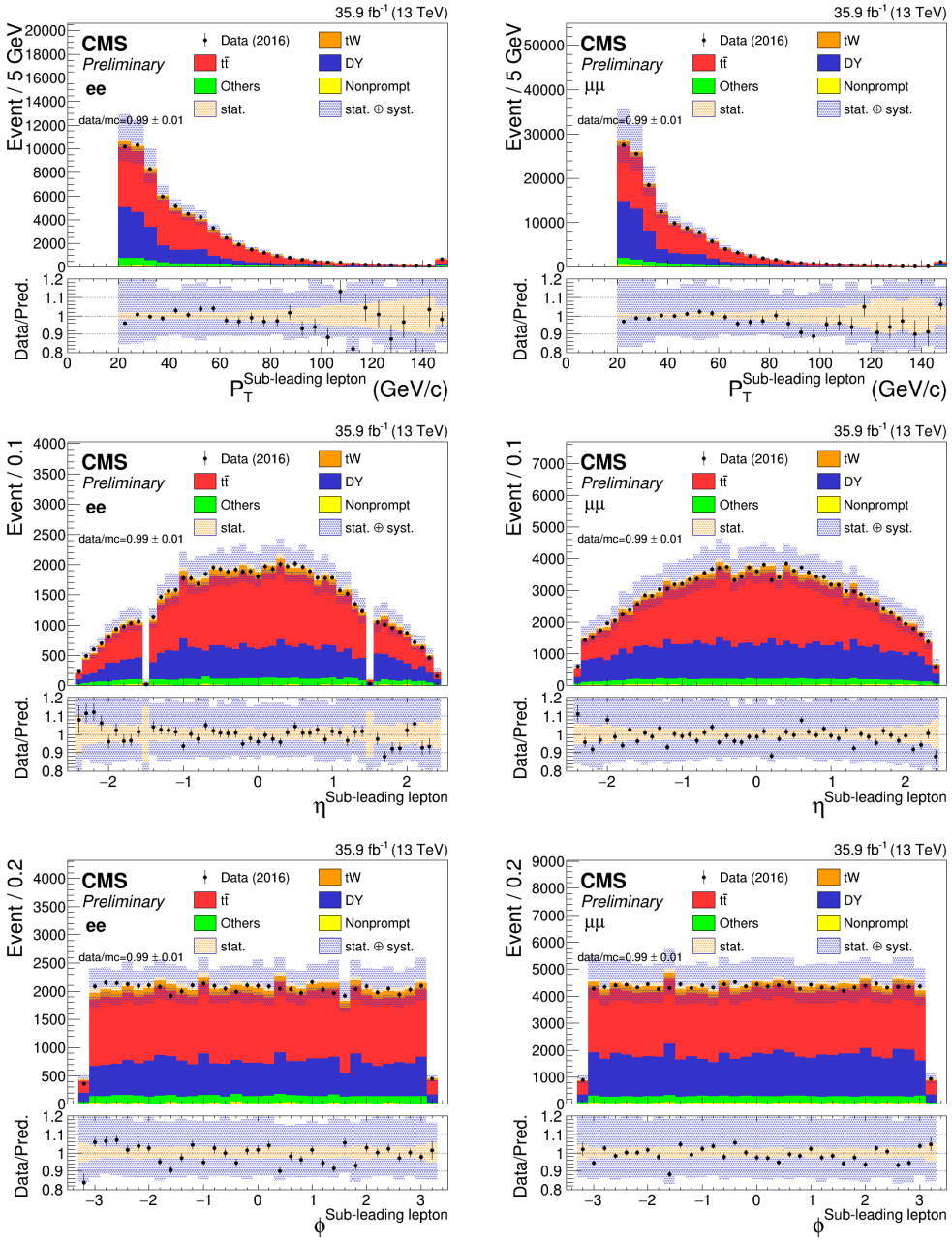

Figure 19: 双电子道（左）和双缪子道（右）中 sub-leading 轻子的 Pt （上）、η （中）和 ϕ （下）的分布。

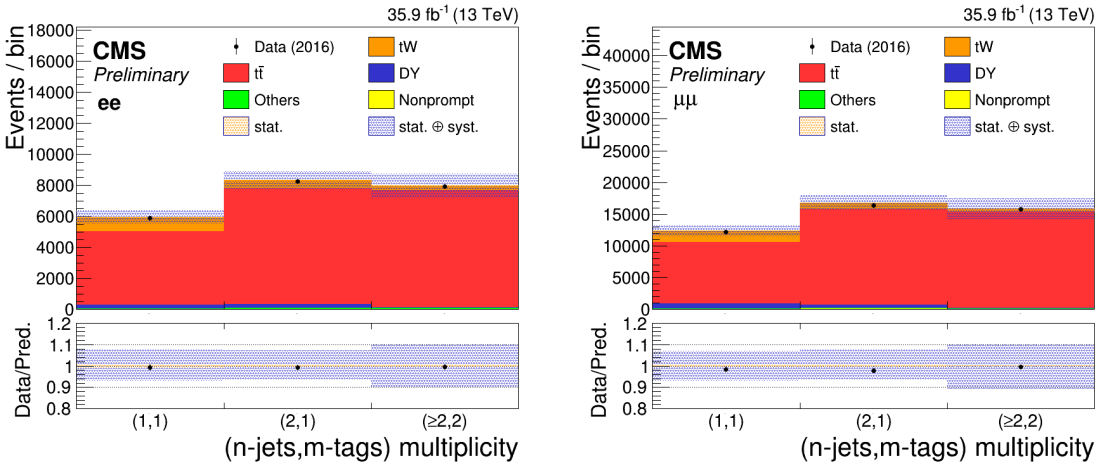

Figure 20: 双电子道（左）和双缪子道（右）中数据和蒙卡样本在不同 jet 和 bjet区间的事例数。

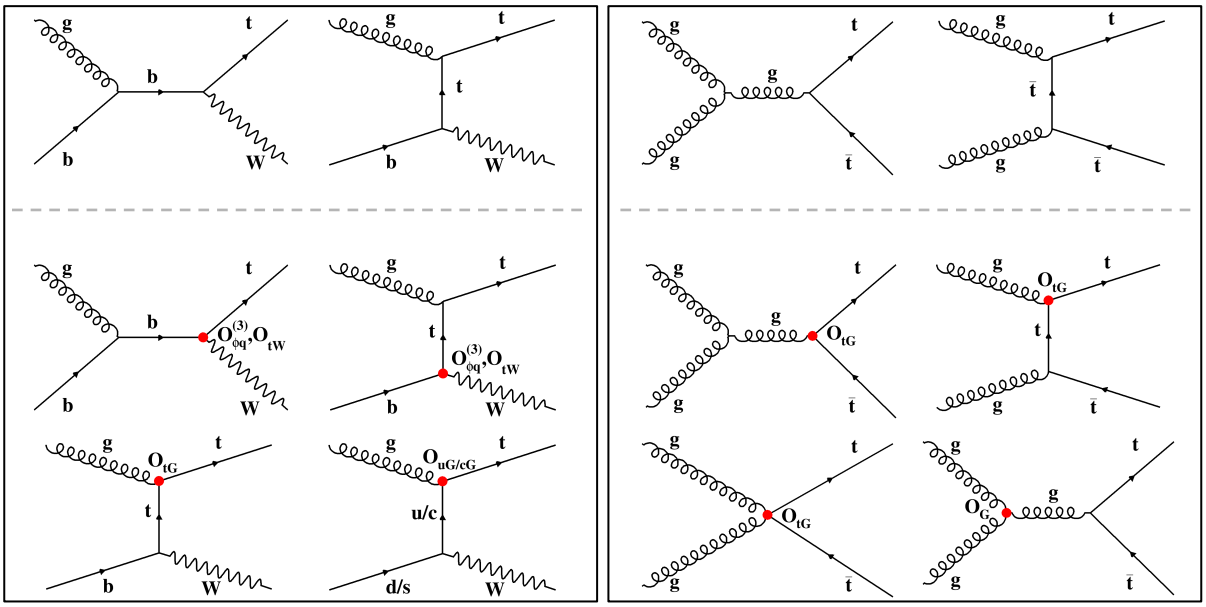

Figure 21: tW（左）和 $t\bar{t}$ （右）的领头阶费曼图。第一行为标准模型过程，第二行和第三行分别代表含有 $O^{(3)}_{\phi q}$ 、O_{tW} 、O_{tG} 、O_G 和 $O_{u(c)G}$ 的过程。

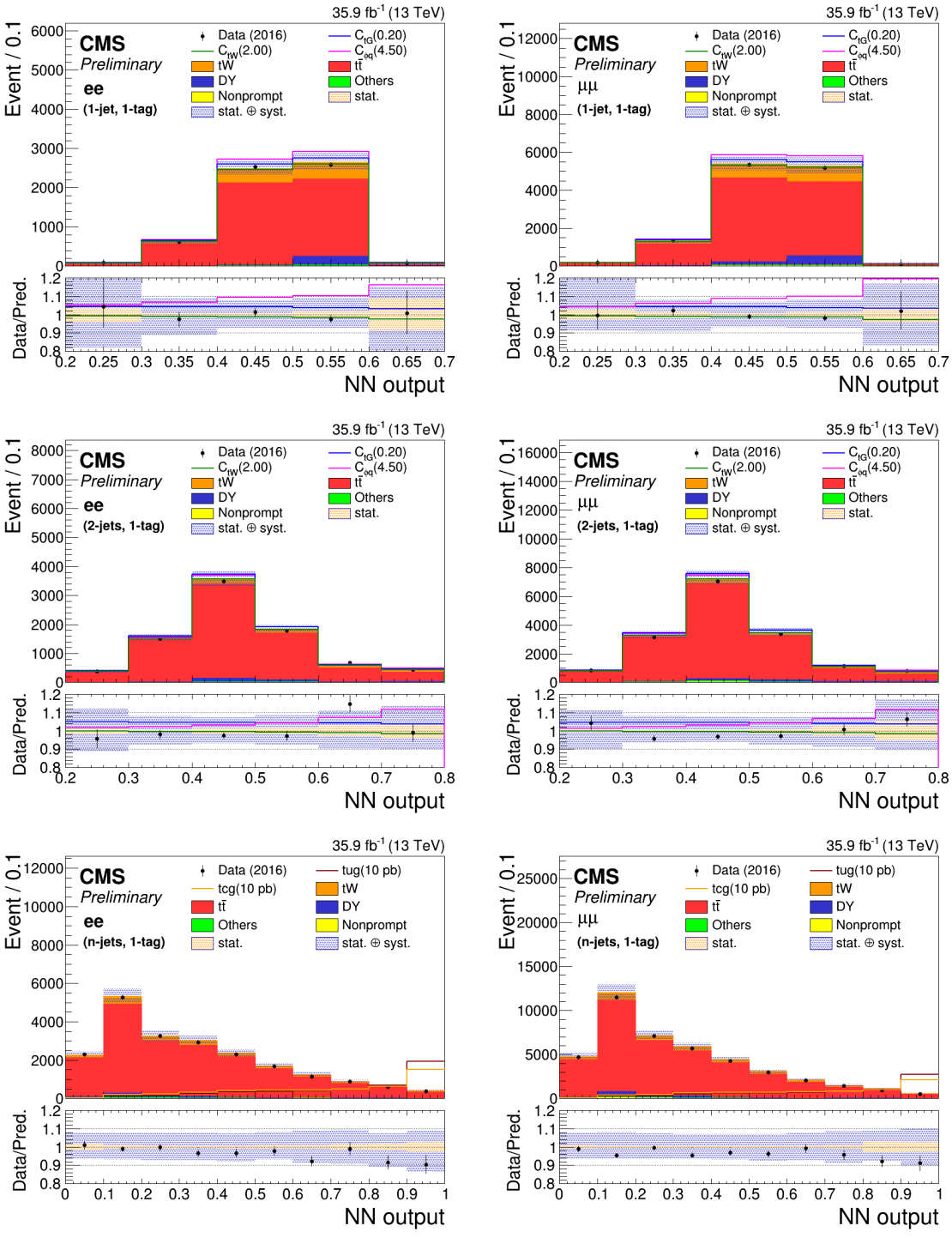

Figure 22: 数据和蒙特卡洛样本在双电子道（左）和双缪子道（右）在 1-jet,1-tag （上）、2-jet,1-tag （中）和 n-jets,1-tag （下）区域的神经网络（Neural Network）输出。

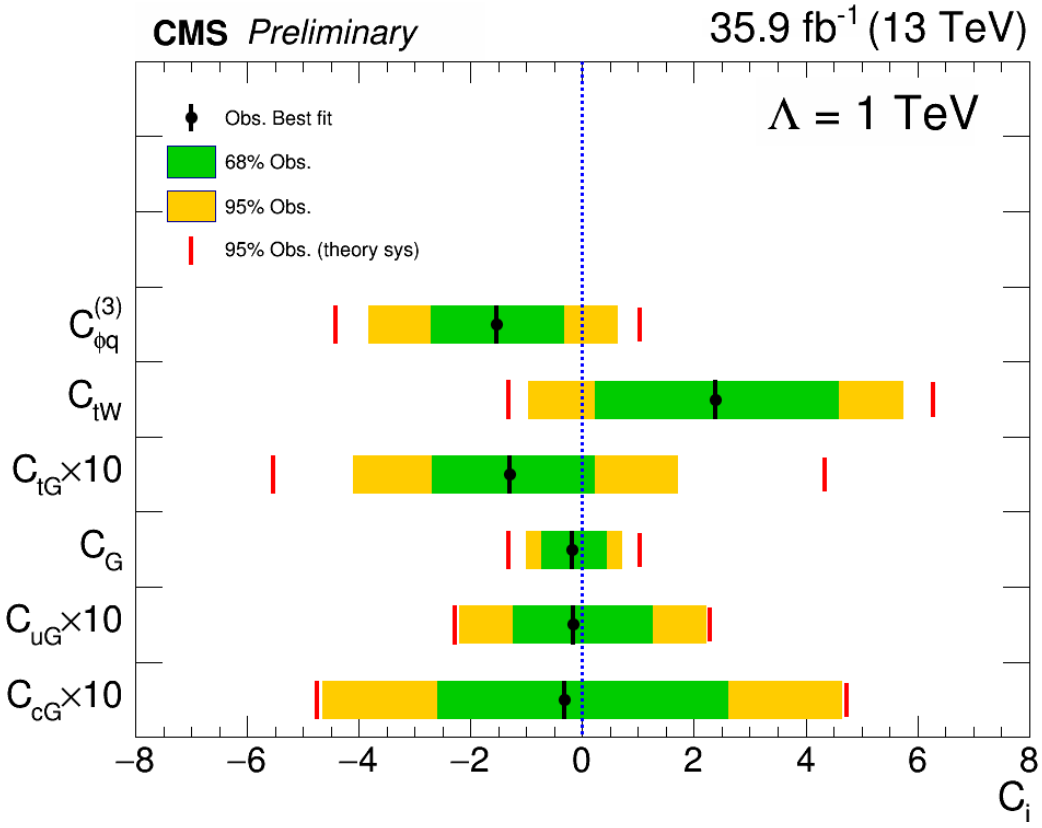

Figure 23: 在结合双电子道、双缪子道和一个电子加一个缪子道后观测到的和预测的不同耦合的 95% 的置信区间 [52]。

Introduction

What is the origin of matter? What are the most fundamental elements in our universe? What are the main forces between matter? These are interesting, basic, and important questions. Although they are difficult to be answered, we are on the way to find the answers.

Particle physics is a subject that tries to find the basic particles in the Universe and to understand the interaction mechanisms between these fundamental particles. The most well known and successful theory in particle physics is the Standard Model (SM) which managed to explain until now all the experimental observations with outstanding precision. The final missing piece of the SM, the "Higgs" particle which was introduced in 1964 by Brout, Englert, and Higgs in order to explain the origin of masses of elementary particles, has been discovered in 2012 (48 years after its prediction) at a mass of around 125 GeV by the ATLAS and CMS experiments at the Large Hadron Collider (LHC).

The LHC is the largest hadron collider in the world providing proton-proton collisions with the highest center-of-mass energy ever achieved (13 TeV from 2015 to 2018). There are four main experiments at the LHC, two of them CMS and ATLAS are general purpose detectors. The discovery of the Higgs boson by joint efforts of the ATLAS and CMS collaborations is one of the most important achievements of modern particle physics research and accomplished one of the main goals of the LHC program.

Nevertheless, besides the tremendous successes of the SM, it is not able to describe the full picture of Nature. Indeed, it does not show candidates of dark matter and dark energy, it does not predict the oscillation of neutrinos, it does not have a good explanation for the asymmetry between matter and antimatter. It has in addition some issues of internal consistency, such as the hierarchy problem, a large number of free parameters and so on. Therefore, the SM is generally considered as an effective theory of a more fundamental theory at high energy. In order to address some shortcomings of the SM, several models that go beyond the SM have been proposed, such as supersymmetric models (SUSY), which provide a candidate of dark matter and provide an explanation to the Higgs mass fine-tuning problem, or The Grand Unified Theory (GUT) which tries to unify electromagnetic, weak, and strong interactions into one interaction through extensions of the SM gauge group, or the large extra dimensions theory, which involves additional spatial dimensions to explain the weakness of the gravitational force compared to the other forces. These beyond SM models typically introduce new neutral bosons heavier than the standard model Z boson, which are generically called $Z^{'}$ bosons.

If the mass scale of such new particles are reachable in collider experiment, these particles would manifest themselves as a localised excess of events in the observed invariant mass spectra. In this thesis, direct search for new heavy resonances decaying into the dielectron final state has been performed using the CMS detector. This channel has the advantage that electrons can be reconstructed and identified with high efficiency which leads to a low background contamination coming from misreconstructed electron candidates. Besides, the main component of SM background in this channel is Drell-Yan

process, which is well understood and its rate is small in the high mass region. These facts give a strong motivation for searching for new heavy resonances in the dielectron final state. In this thesis, the analysis of data collected by the CMS experiment during years 2016 and 2017 are reported.

However, if the new physics scale is not reachable at the LHC, new physics could affect SM interactions indirectly, through modifications of SM couplings or enhancements of rare SM processes. 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. An effective field theory (EFT) approach which is a model-independent approach is used in this thesis to search for new physics in the top quark sector in the dilepton (ee, $\mu\mu$) final states using the data collected by CMS in 2016.

The thesis is organised as follows. The SM of particle physics is introduced in Chapter 1, including a description of fundamental particles, the forces between these particles, the main properties of the SM Drell-Yan process and an introduction to the EFT theory. Chapter 2 lists the shortcomings of the SM and addresses how various theories beyond the SM propose to solve them. In particular, the models that predict additional massive resonances are introduced. The motivations for searching for new physics in the top quark sector is also given. Chapter 3 presents an introduction to the LHC machine, including its design and operational parameters, as well as the phenomenological aspects of the proton-proton interactions. The CMS detector is also introduced in detail in this chapter. The reconstruction of the different particles produced in the proton proton collisions in CMS is explained in Chapter 4. Chapter 5 describes in detail the results of the search for new resonances decaying into the dielectron final state and all the aspects of the analysis are covered. The results of the search for new physics in the top quark sector are shown in Chapter 6, covering as well all the aspects of the analysis. Finally, Chapter 7 exposes the conclusions coming from both searches.

Chapter 1

The Standard Model of particle physics

This chapter introduces the Standard Model of particle physics which describes the family of elementary particles and the three of the four fundamental forces of Nature with corresponding mediators. After that, the gauge symmetry is briefly expressed. Then, the Drell-Yan process is introduced in details due to its importance in this thesis. In Addition, the photon induced process is also shortly discussed. Finally, the effective field theory which will be used in this thesis is presented.

1.1 The elementary particles

It has been long time for people to understand what are the basic objects which constitute our world. From Demokritos (470-380 BC) who thought matter was built of discrete building blocks to John Dalton (1766-1844) who came up with the matter made of atoms. In the early 1900's J.J. Thomson proposed a so called "plum pudding model" which assumes the atom was a uniform sphere of positively charged matter in which electrons were embedded. However in 1910 Ernest Rutherford and his colleagues performed α ray scattering experiments and found that the whole mass and all positive charges of the atom were concentrated in a minute space at the centre which is called "nucleus". After the discovery of the neutron in 1932 by James Chadwick, models for a nucleus composed of protons and neutrons were quickly developed by Dmitri Ivanenko and Werner Heisenberg. Furthermore in 1968 the deep inelastic scattering experiments at the Stanford Linear Accelerator Center provided the first convincing evidence of the reality of quarks in the proton or neutron. In the Standard Model (SM) [1, 2, 3] the quarks are the elementary particles and there are six different kinds of flavors for the quarks called up (u), down (d), charm (c), strange (s), top (t), and bottom (b). The mass, charge, and spin of the quarks are shown in Table 1.1, here the e means one electron's charge which equals 1.6×10^{-19} C. The quarks are categorized in there generations according to their masses and charges. Besides it is verified that the quarks have another property which is called the "color charge", a quark can be "Red" or "Blue" or "Green". Moreover all the quarks have its anti-quark partner which has opposite quantum numbers with regard to the quark including flavor, charge, and color charge. Therefore there are $6\ (\text{flavor}) \times 3\ (\text{color}) \times 2\ (\text{anti}-\text{quark}) = 36$ kinds of quarks in the SM and all the hadrons are composed by quarks or anti-quarks. For meson it is composed by $q\bar{q}$ and for baryon it is composed by qqq or $\bar{q}\bar{q}\bar{q}$.

Similar to the quark family, there is a lepton family and the most common lepton is the electron which exists as a cloud out of nucleus to form an atom. All the leptons are shown in Table 1.2 with their charges, spins, and masses. They are electron (e), electron neutrino (ν_e), muon (μ), muon neutrino (ν_μ), tau (τ), and tau neutrino (ν_τ). The leptons

Generation	Quark	Charge	Spin	Mass
First	up quark (u)	2/3 e	1/2	$2.3^{+0.7}_{-0.5}$ MeV
	down quark (d)	-1/3 e	1/2	$4.8^{+0.5}_{-0.3}$ MeV
Second	charm quark (c)	2/3 e	1/2	1.275 ± 0.025 GeV
	strange quark (s)	-1/3 e	1/2	95 ± 5 MeV
Third	top quark (t)	2/3 e	1/2	$173.21 \pm 0.51 \pm 0.71$ GeV
	bottom quark (b)	-1/3 e	1/2	4.66 ± 0.03 GeV

Table 1.1: Quarks and their properties [53].

Generation	Lepton	Charge	Spin	Mass
First	electron (e)	-e	1/2	511 MeV
	electron neutrino (ν_e)	0	1/2	< 2 eV
Second	muon (μ)	-e	1/2	105.67 MeV
	muon neutrino (ν_μ)	0	1/2	< 2 eV
Third	tau (τ)	-e	1/2	1776.99 MeV
	tau neutrino (ν_τ)	0	1/2	< 2 eV

Table 1.2: Properties of the leptons in the three generations. Neutrinos are assumed to have zero mass in SM but by the observation of neutrino's oscillations the upper limits on their masses are set [53].

are also categorized into three generations and each generation has its own lepton flavor. The first generation has electron flavor, the second generation has muon flavor, and the third generation has tau flavor. For neutrinos they are always left handed and they are very hard to be detected because of the very weak interaction between the matter and neutrinos. In SM the neutrinos are assumed to be massless while in experiment we have observed the neutrino oscillations which means the neutrinos have masses, so the study of neutrinos may bring us the new theory beyond the SM. Similar with the quarks, all the leptons also have theirs anti-particle partners which own the opposite quantum numbers. While unlike the quarks, the leptons do not have color charge property. Therefore there are $6 \times 2 = 12$ kinds of leptons in SM.

Now that all the spin 1/2 (fermion) elementary particles in SM have been introduced, the interactions between these particles and the mediators will be discussed in the next section.

1.2 The fundamental interactions

It is well known that there are four characteristic interactions among fundamental particles.

1. Electromagnetic interaction : It is mediated by massless photon ($m_\gamma = 0$) with spin = 1 among charged particles. Because of the massless of photon the interacting range of electromagnetic interaction is infinite. The theory to describe the electromagnetic interaction is quantum electrodynamics (QED) which has been well understood. QED is renormalizable, for example the divergence from vacuum polarization and higher order loop contributions can be absorbed into the physical charge of particle. The coupling constant is $\alpha = \frac{e^2}{4\pi\varepsilon_0\hbar c} \simeq \frac{1}{137}$ which characterizes the strength of the coupling of charged particle with the electromagnetic field. Because of the smallness of α, the perturbation development works well for QED.
2. Strong interaction : It is mediated by massless gluons ($m_g = 0$) with spin = 1. This interaction can only happen between quarks and gluons. The theory that describes

the strong interaction is called quantum chromodynamics (QCD). For strong interaction the coupling strength is also running but in an opposite way as the electromagnetic interaction, which means that the coupling will decrease when the interaction energy increases. Because of that there is one special phenomenon for strong interaction called "asymptotic freedom" which means quarks and gluons behave like free particles when the interaction energy is very high. At high energy the perturbation theory works well because of "asymptotic freedom" but it is not the case at low energy (below the GeV). Therefore there are still a lot of works needed to be done for understanding QCD process at low energy. Another special phenomenon of the strong interaction is called the "color confinement", which means there are no free quarks in the world. The quarks need to be grouped to form a colorless hadron. For instance the baryon is formed by red, green, and blue quarks and the meson is formed by red and anti-red quarks or blue and anti-blue quarks or green and anti-green quarks.

3. Weak interaction : It is mediated by massive weak bosons ($m_{W^{\pm}} \cong 80.4$ GeV, $m_Z \cong 91.2$ GeV) with spin $= 1$. Because of the heavy mediator, its interacting range is very short, which is $\sim 10^{-18}$ m. The coupling strength of weak interaction is the weakest among electromagnetic and strong interactions. Although the coupling strength is small, some processes can only happen via weak interaction like flavor changing or neutrino involved processes.
4. Gravitational interaction : It is mediated by massless gravitons ($m_G = 0$) with spin $= 2$ among all massive particles. Because of its very small coupling strength we normally do not consider gravitational interaction in high energy physics. In macroscopic world the gravitation is important, such as it makes an apple falling.

The summary of interaction range, relative strength, and mediator for the four fundamental interactions is shown in Table 1.3. Last but not least, in the SM the origin of mass of the elementary particles is coming from interactions between particles and the Brout-Englert-Higgs scalar field (the so called "Higgs" boson). In 2012 the ATLAS [4] and CMS [5] experiments observed such a particle and the mass is ~ 125 GeV.

The summary of all elementary particles, force carries, and Higgs boson is shown in Figure 1.1.

Interaction	Range	Relative strength	Mediators
Strong	10^{-15} m	1	8 gluons (g)
Electromagnetic	∞	10^{-3}	photon (γ)
Weak	10^{-18} m	10^{-14}	W^+, W^-, Z
Gravitational	∞	10^{-43}	graviton (G) ?

Table 1.3: Range, relative strength with respect to the strong force, and mediators of the four fundamental interactions. The gravitational force is not included in the SM, and gravitons are hypothetical particles.

An introduction about the Feynman calculus is presented in Appendix A.1. Besides, more detailed introductions about the QED and the QCD are presented in Appendix A.2 and Appendix A.3, respectively. The group theory is briefly introduced in Appendix A.4.

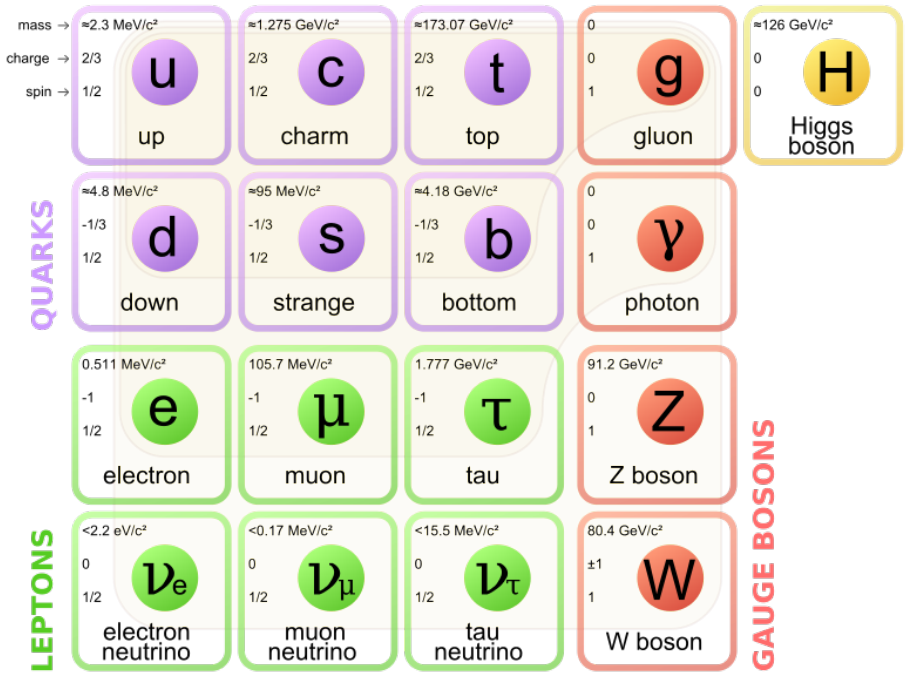

Figure 1.1: Overview of the Standard Model constituents: the quarks and leptons, the gauge bosons and the Higgs boson.

1.3 Gauge symmetries: a brief introduction

The formulas in this section are from book [54].

The Lagrangian

As we know, in classic physics, the motion equation of a particle can be obtained from the Lagrange's equation

$$\frac{d}{dt}(\frac{\partial L}{\partial \dot{q}_i}) - \frac{\partial L}{\partial q_i} = 0 \tag{1.1}$$

where the q_i are the generalized coordinates of the particle, t is the time variable and $\dot{q}_i = dq_i/dt$. The $L \equiv T - V$, where T is kinetic energy of the particle and V is the potential energy of the particle. The Lagrange's equation 1.1 can be extended from describing the motion of one particle to describe the motion of a field $\phi(t, \mathbf{x})$ (which has a value at every point in space that changes in time) by replace q_i and $\dot{q}_i$ with ϕ and $\partial\phi/\partial x_\mu$ respectively, here the $x_\mu \equiv (t, \mathbf{x})$. Therefore, we obtained the Lagrange's equation for field ϕ as

$$\frac{\partial}{\partial x_\mu}(\frac{\partial \mathcal{L}}{\partial(\partial\phi/\partial x_\mu)}) - \frac{\partial \mathcal{L}}{\partial \phi} = \frac{\partial}{\partial t}(\frac{\partial \mathcal{L}}{\partial(\partial\phi/\partial t)}) + \sum_{i=1}^{3} \frac{\partial}{\partial x_i}(\frac{\partial \mathcal{L}}{\partial(\partial\phi/\partial x_i)}) - \frac{\partial \mathcal{L}}{\partial \phi} = 0, \tag{1.2}$$

which is called Euler-Lagrange equation and the $\mathcal{L}$ is Lagrangian density with

$$L = \int \mathcal{L} d^3x.$$

Usually we call $\mathcal{L}$ itself the Lagrangian.

For example, the Dirac Lagrangian (which describes the spin $\frac{1}{2}$ particle in quantum mechanics) is

$$\mathcal{L} = i\overline{\psi}\gamma_\mu\partial^\mu\psi - m\overline{\psi}\psi \tag{1.3}$$

where ψ is the particle field and $\overline{\psi} \equiv \psi^\dagger\gamma^0$, m is the mass of the particle, the $\partial^\mu = (\frac{\partial}{\partial t}, -\nabla) = (\frac{\partial}{\partial t}, -\frac{\partial}{\partial x_1}, -\frac{\partial}{\partial x_2}, -\frac{\partial}{\partial x_3})$, $\gamma_\mu = (\gamma_0, \gamma_1, \gamma_2, \gamma_3)$ are the gamma matrices:

$$\begin{aligned} &\gamma_0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \ \gamma_i = \begin{pmatrix} 0 & -\sigma_i \\ \sigma_i & 0 \end{pmatrix} \quad \text{with } i = 1,\ 2,\ 3 \\ &\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned} \tag{1.4}$$

Symmetry and conservation

The first person who linked symmetry with conservation law is Emmy Noether [55]. She pointed out that each symmetry (which means the physics or the Lagrangian is invariant under this operation) corresponds to one conserved quantity. For example, the symmetries of transitions, time displacements, and rotations lead to the conservation of momentum, energy and angular momentum. An example about "internal" symmetry is given below.

Suppose we change the phase of electron field by

$$\psi \to e^{i\alpha}\psi, \tag{1.5}$$

where α is space and time independent. It is easily to see that this operation is a symmetry operation (usually called global $U(1)$ symmetry, because all these kind of operations with

different α value form a unitary group with one group generator, the global means α is space and time independent) due to the invariance of Lagrangian (see Equation 1.3). From Noether's theorem we know there is a conserved quantity corresponding to this symmetry.

According to the property of $U(1)$ group, when α is infinitesimal the Equation 1.5 can be written as

$$\psi \to (1 + i\alpha)\psi. \tag{1.6}$$

and by asking the invariance of Lagrangian we get

$$0 = \delta\mathcal{L} = \frac{\partial}{\partial\psi}\delta\psi + \frac{\partial}{\partial(\partial_\mu\psi)}\delta(\partial_\mu\psi) + \frac{\partial}{\partial\overline{\psi}}\delta\overline{\psi} + \frac{\partial}{\partial(\partial_\mu\overline{\psi})}\delta(\partial_\mu\overline{\psi}) \tag{1.7}$$

and finally we can get can a conserved current from

$$\partial_\mu j^\mu = 0, \tag{1.8}$$

where

$$j^\mu = \frac{ie}{2}(\frac{\partial}{\partial(\partial_\mu\psi)}\psi - \overline{\psi}\frac{\partial}{\partial(\partial_\mu\overline{\psi})}) = -e\overline{\psi}\gamma^\mu\psi. \tag{1.9}$$

It can be proved that the conserved current j^μ leads to the charge conservation of the particle.

$U(1)$ local gauge symmetry

As we know, in previous section, the Lagrangian (Equation 1.3) is invariant under $U(\alpha)$ operation, while it will not be the case for local operator $U(\alpha(t, \mathbf{x}))$ which is time and space dependent. However, the real Lagrangian should be invariant with $U(\alpha(t, \mathbf{x}))$, as we know the observation $|\langle\psi|\psi\rangle|^2 = |\langle\psi|U^\dagger U|\psi\rangle|^2$ do not dependent with the phase.

In order to maintain the Lagrangian is invariant under $U(\alpha(t, \mathbf{x}))$, it is needed to replace derivative ∂_μ by D_μ with

$$D_\mu \equiv \partial_\mu - ieA_\mu \tag{1.10}$$

where A_μ transforms as

$$A_\mu \to A_\mu + \frac{1}{e}\partial_\mu\alpha. \tag{1.11}$$

Therefore, the updated Lagrangian will be

$$\mathcal{L} = i\overline{\psi}\gamma^\mu D_\mu\psi - m\overline{\psi}\psi = \overline{\psi}(i\gamma^\mu\partial_\mu - m)\psi + e\overline{\psi}\gamma^\mu\psi A_\mu, \tag{1.12}$$

which means there is an interaction between the field ψ and field A_μ. Actually, it can be proved that the A_μ can be regarded as the photon and after include the kinetic term of the photon (not for the mass term which will break the symmetry) the final Lagrangian for QED is

$$\mathcal{L}_{QED} = \overline{\psi}(i\gamma^\mu\partial_\mu - m)\psi + e\overline{\psi}\gamma^\mu\psi A_\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{1.13}$$

with

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{1.14}$$

We have seen that after asking the local $U(1)$ symmetry, a massless gauge boson, the photon, is created.

$SU(3)$ local gauge symmetry

As we know, the quark has three colors (R, G, B) and the Lagrangian

$$\mathcal{L} = \overline{q_i}(i\gamma^\mu\partial_\mu - m)q_i, \quad \text{with } i \;=\; 1,\; 2,\; 3, \tag{1.15}$$

where the q_1, q_2, q_3 denote the three color fields, should be invariant under the color phase transformation (R $\rightleftarrows$ G $\rightleftarrows$ B). This is because we really can't distinguish the exact color of one quark. This color phase transformation can be represented by 3×3 traceless unitary matrices U, and all these matrices form a $SU(3)$ ("S" means special because of the zero trace of the matrices) group with 8 generators. The transformation of the quark field under the color phase change can be written as

$$q(t, \mathbf{x}) \to Uq(t, \mathbf{x}) \equiv e^{i\alpha_a(t,\mathbf{x})T_a}q(t, \mathbf{x}), \tag{1.16}$$

where a summation over suffix a from 1 to 8 is implied, the T_a are a set of linearly independent traceless 3×3 matrices, and the α_a are the group parameters. Because not all generators T_a commute with each other (e.g. $T_aT_b \neq T_bT_a$), this group is non-Abelian. The conventional choice of T_a matrices are the $\lambda_a/2$ (know as Gell-Mann λ matrices, see Equation A.66). It can be proved that the commutator of any two T_a follows

$$[T_a, T_b] = if_{abc}T_c, \tag{1.17}$$

where f_{abc} are real constant, called the structure constants of the group.

In order to impose the $SU(3)$ color local invariance of the Lagrangian 1.15, we can use the same method described in Section 1.3 by making

$$D_\mu = \partial_\mu + igT_aG^a_\mu, \tag{1.18}$$

and

$$G^a_\mu \to G^a_\mu - \frac{1}{g}\partial_\mu\alpha_a - f_{abc}\alpha_bG^c_\mu. \tag{1.19}$$

Similar with $U(1)$ gauge symmetry, after requiring $SU(3)$ color local symmetry, we created a new gauge field G^a_μ (a=1,...,8) which can be regarded as 8 gluons. After adding the kinetic energy terms of the gluons, the final gauge invariant Lagrangian for QCD process is

$$\mathcal{L}_{QCD} = \overline{q}(i\gamma^\mu\partial_\mu - m)q - g(\overline{q}\gamma^\mu T_aq)G^a_\mu - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a, \tag{1.20}$$

with

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - gf_{abc}G^b_\mu G^c_\nu.$$

It should be noticed that due to the non-Abelian of $SU(3)$, the kinetic energy terms $G^a_{\mu\nu}G^{\mu\nu}_a$ induce self-interactions between gauge bosons which is not the case for $U(1)$ gauge symmetry which is Abelian.

We have seen that the exact color $SU(3)$ (or simply $SU(3)_c$) local symmetry gives 8 massless gauge bosons which are gluons, and these gluons can have interactions with the quarks or have self-interactions by QCD process.

$SU(2)_L$ local gauge symmetry and Higgs mechanism

In the weak interaction, the left-handed fermions is coupled to form weak-isospin doublets (e.g. $(\nu_{eL}, e_L), (u_L, d_L)$) and the right-handed fermions form weak-isospin singlet (e.g. e_R, u_R, d_R, and there is no right-handed neutrinos). The "rotation" from one fermion to another fermion within the same doublets can be represented by $SU(2)_L$ ("L" means left-handed) group with 3 generators.

As we know, the mediators $W^{\pm}$ boson for weak interaction are heavy particles. If we impose $SU(2)_L$ local symmetry in the same way as we did for $SU(3)$ then we will obtain massless gauge bosons which conflicts with the experimental results. Therefore, we need additional mechanisms to make the gauge bosons have masses. Luckily, in SM we have a mechanism (called "Higgs mechanism") proposed by Brout, Englert and Higgs [56, 57, 58] which suppose there exist a scalar boson (called "Higgs" boson) and its the potential V is not at minimum when its field ϕ at 0, while at value v (called "vacuum expectation value" or simply VEV) the potential reaches minimum, see Figure 1.2 for example.

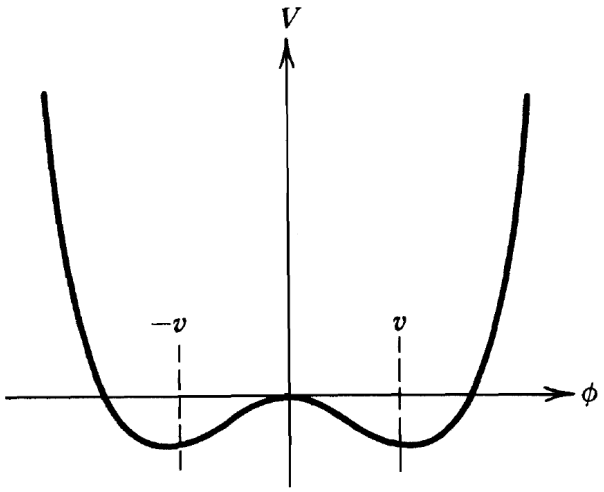

Figure 1.2: An example of one dimension "Higgs" potential [54].

In perturbative calculation we should involve expansions around the minimum energy and this can be done by expanding the ϕ around the v

$$\phi(t, \mathbf{x}) = v + h(t, \mathbf{x}).$$

It means ϕ can be expressed by h, while potential V will not be symmetry under the change of h to $-h$. This makes the "spontaneous symmetry breaking" of ϕ. After expanding the ϕ around the VEV, in new Lagrangian we have a term related to the mass of h (actually is the mass of "Higgs" boson) and a term related to the masses of gauge bosons which means the gauge bosons obtained the masses. By the way, the masses of fermions in SM are also "generated" by Higgs mechanism.

Last but not least, the only $SU(2)_L$ local symmetry does not create the physical Z boson. It is created together with $W^{\pm}$ and photon in $SU(2)_L \times U(1)_Y$ ("Y" means hypercharge, $Y = 2Q - T^3$ with Q is the charge of particle, T^3 is third component of weak-isospin) symmetry which is proposed by Weinberg, Salam, and Glashow [59, 60, 61]. This $SU(2)_L \times U(1)_Y$ gauge invariance theory unifies electromagnetic and weak interactions and is called electroweak theory. Finally, the complete SM theory is based on $SU(3)_c \times SU(2)_L \times U(1)_Y$ gauge symmetry.

1.4 The Drell-Yan process

Due to its importance in this thesis for the search for heavy resonances in the dielectron final state (see Chapter 5), the Drell-Yan (DY) [6] process is introduced with more details. The DY process is defined as the annihilation of a quark-antiquark pair into a pair of oppositely-charged leptons. The Feynman diagrams of the DY process at leading order are shown in Figure 1.3.

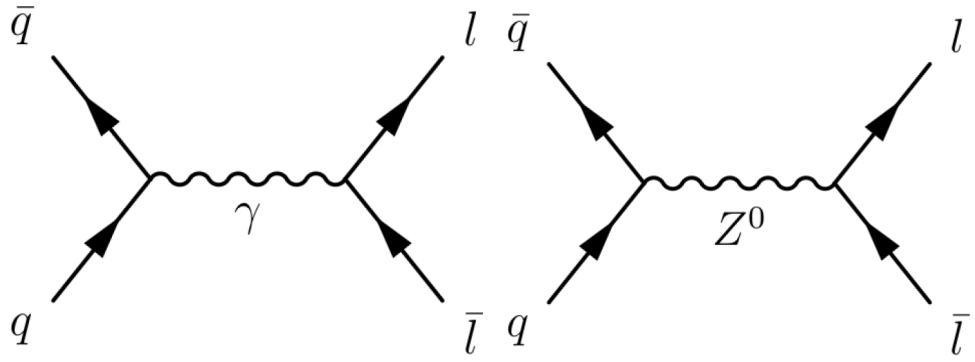

Figure 1.3: Feynman diagrams contributing to the Drell-Yan process at leading order. The left (right) diagram corresponds to the annihilation of a $q\bar{q}$ pair into a photon (a Z boson).

The diagram in Figure 1.4, with a Higgs boson exchange is neglected. Indeed, the coupling between the fermion and the Higgs boson is proportional to m_f/v, where m_f is the mass of the fermion and v is the vacuum expected value of the scalar field (≈ 246 GeV). So the amplitude of this diagram is proportional to $\frac{m_q(\text{GeV})}{246} \cdot \frac{m_\ell(\text{GeV})}{246}$ with m_ℓ be the mass of lepton. As we know, the valence quarks in the proton are the up and down quarks which are light ($m < 10$ MeV). Therefore the contribution of the third diagram is several orders of magnitude smaller than the first two diagrams in proton proton collision and it is usually neglected in the calculation.

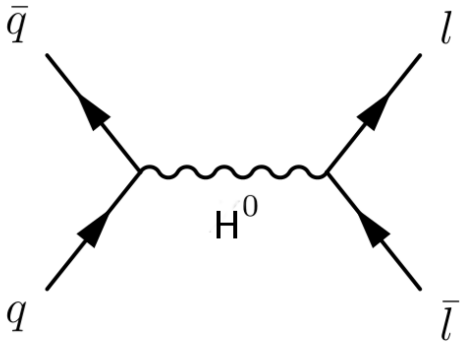

Figure 1.4: Feynman diagram of the process $q + \bar{q} \to H^0 \to \ell + \bar{\ell}$ at leading order.

Cross-section

The diagrams in Figure 1.3 give rise to two matrice terms $\mathcal{M}_\gamma$ and $\mathcal{M}_Z$, hence the cross section of DY process can be expressed as:

$$\sigma_{\gamma/Z} = \sigma_\gamma + \sigma_Z + \sigma_{int} \tag{1.21}$$

where σ_γ is the cross-section corresponding to the exchange of a photon only, σ_Z is for the exchange of a Z boson only, and σ_{int} is the cross-section from the interference of the first two processes.

The formulas in the rest of the section are taken from [62].

To be specific, the Equation 1.21 can be written as:

$$\sigma(q(p_1)\bar{q}(p_2) \to l^+l^-) = \frac{4\pi\alpha^2}{3s}\frac{1}{N}(Q_q^2 - 2Q_qV_lV_q\chi_1(s) + (A_l^2 + V_l^2)(A_q^2 + V_q^2)\chi_2(s)) \tag{1.22}$$

with

$$\chi_1(s) = \kappa\frac{s(s - M_Z^2)}{(s - M_Z^2)^2 + \Gamma_Z^2M_Z^2}$$

$$\chi_2(s) = \kappa^2\frac{s^2}{(s - M_Z^2)^2 + \Gamma_Z^2M_Z^2}$$

$$\kappa = \frac{\sqrt{2}G_FM_Z^2}{16\pi\alpha}$$

Here p_1 is the four-momenta of the quark, p_2 is the four-momenta of the anti-quark, and $s = (p_1 + p_2)^2$. The α is the electromagnetic coupling, the G_F is the Fermi constant with the value 1.1663787×10^{-5} GeV^{-2} [63], which can be precisely determined by muon lifetime experiment. The M_Z and Γ_Z are the mass and total decay width of the Z boson respectively. The $\frac{1}{N} = \frac{1}{3}$ and it is due to the color matching between quark and anti-quark. The Q_q is the charge of the quark, the V and A are the vectorial and axial couplings associated to the lepton/quark. The value of V and A are

$$V_f = T_f^3 - 2Q_f\sin^2\theta_W, \qquad A_f = T_f^3$$

with ${T_f}^3 = +\frac{1}{2}$ for $f = \nu, u, ...$ and ${T_f}^3 = -\frac{1}{2}$ for $f = e, d, ...$ The θ_W in $\sin^2\theta_W$ is so called "weak mixing angle" and it has been mathematically defined as $\cos\theta_W = \frac{M_W}{M_Z}$, here the M_W is the mass of W boson.

The first term $(\frac{4\pi\alpha^2}{3s}\frac{1}{N}Q_q^2)$ in the right side of Equation 1.22 is corresponding to σ_γ and it can be calculated using QED. The second term $(\frac{4\pi\alpha^2}{3s}\frac{1}{N}(-2Q_qV_lV_q\chi_1))$ is for σ_{int} and the last term $(\frac{4\pi\alpha^2}{3s}\frac{1}{N}(A_l^2 + V_l^2)(A_q^2 + V_q^2)\chi_2)$ is for σ_Z.

From Equation 1.22 one can see that at low centra-mass of energy $(\sqrt{s})$ the Drell-Yan cross section is dominated by photon exchange process (e.g. at $\sqrt{s} = M_Z/2$, the σ_γ is around 100 times greater than σ_Z and over 10 times larger than σ_{int}), while at Z pole $(\sqrt{s} \sim M_Z)$ it is dominated by Z boson exchange process (e.g. at $\sqrt{s} = M_Z$, the σ_Z is well over 100 times larger than σ_γ, while σ_{int} is zero). To get a feeling about the Drell-Yan cross section as the function of lepton pair mass M_{ll} one can see Figure 1.5.

1.5 The photon induced process

As we known, the production of high invariant mass opposite sign lepton pairs in proton proton collision at the LHC is dominated by the Drell-Yan process (See previous section). Addition to this, photon-photon collisions, where the photons are radiated by the quarks in the proton, can also produce lepton pairs. The Feynman diagrams for the photon induced (PI) production of lepton pairs at leading order can be seen in Figure 1.6. The left diagram corresponds to a t-channel process and the right diagram corresponds to a u-channel process. There is no s-channel for the PI process at leading order and this will give different kinematic properties for the lepton pair comparing with Drell-Yan process which is s-channel.

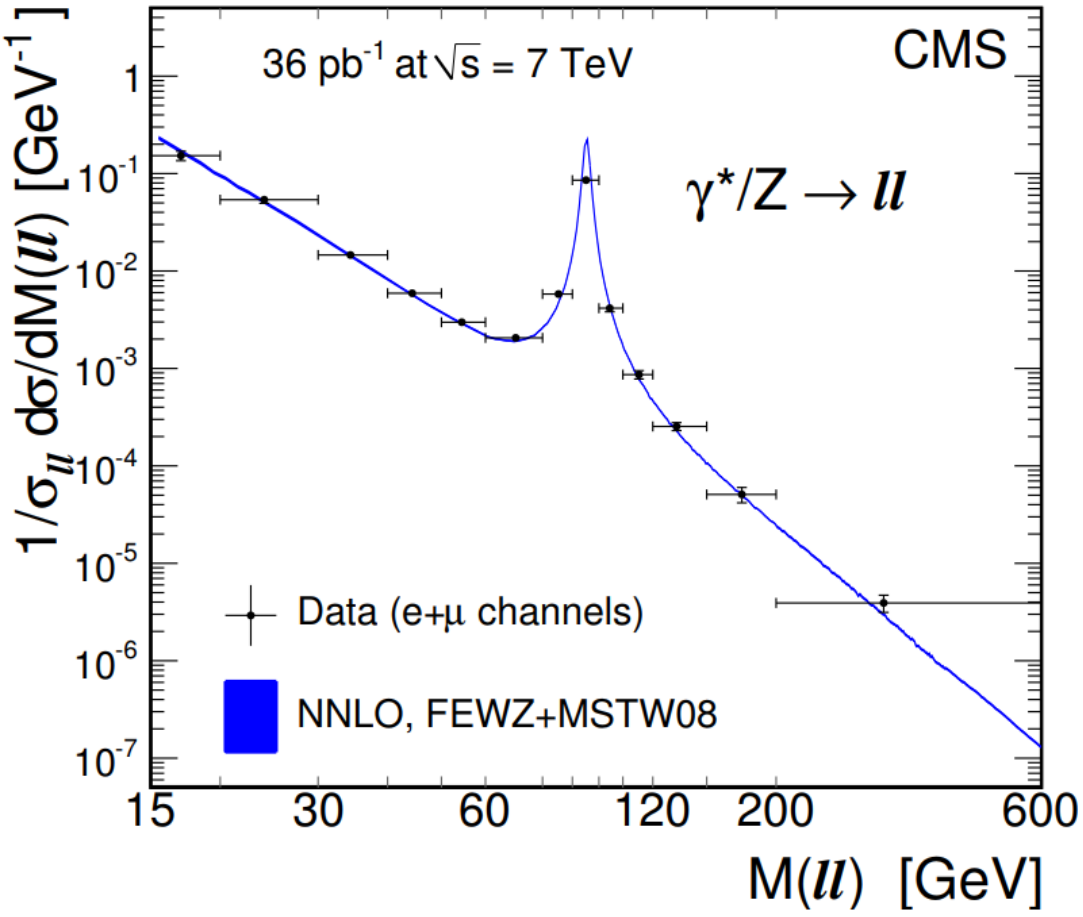

Figure 1.5: The normalized (in Z resonance region) DY cross section as the function of lepton pair mass (M_{ll}) in pp collision at $\sqrt{s} = 7$ TeV with CMS data [64].

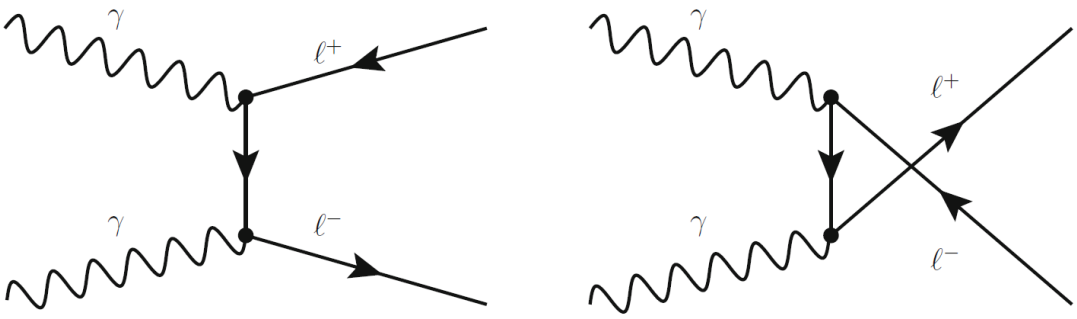

Figure 1.6: Feynman diagrams for the photon induced production of lepton pairs at leading order. The left (right) diagram corresponds to a t-channel (u-channel) process.

The contribution from PI process becomes a significant part of the dilepton production at high invariant masses. Therefore, the knowledge of this process is important input for high mass resonant (like Z′) or non-resonant searches. From [65] we know the PI effects are generally small, not above the 5% level, for masses of up to $\sim$ 2 TeV, but can reach $\sim$ 15 – 20% above 5 TeV. This effects will be taken into account in searching for Z′ study in Chapter 5.

1.6 The effective field theory

If new physics scale is reachable at the experiment then the new physics could be directly observed via the production of new particles. Otherwise, it can take part in as a virtual particle which can not be detected directly but can still have an impact on the SM interactions by modifications of SM couplings or enhancements of rare SM processes. In the latter case, the effective field theory (EFT) approach is useful to parameterize and constrain new physics. In EFT, we extend the SM by adding new terms to the Lagrangian. An example of the extended Lagrangian is shown in Equation 1.23 where Λ represents the energy scale beyond which new physics becomes relevant, $\mathcal{C}_i$ stands for the dimensionless Wilson coefficients (also called as the effective couplings), $\mathcal{O}_i^6$ are dimension-six operators. Therefore, the underlying new physics particle gets integrated out (by measuring only $\mathcal{C}_i/\Lambda^2$) and leaving only the effective vertex, such as the Fermi theory for neutron decay (see the Feynman diagram in Figure 1.7). Besides, the EFT must maintain all the necessary symmetries of the SM. There are total O(100) new EFT vertices and it can be reduced by focusing on specific physical processes.

$$\mathcal{L} = \mathcal{L}_{SM} + \sum_i \frac{\mathcal{C}_i}{\Lambda^2}\mathcal{O}_i^6 \tag{1.23}$$

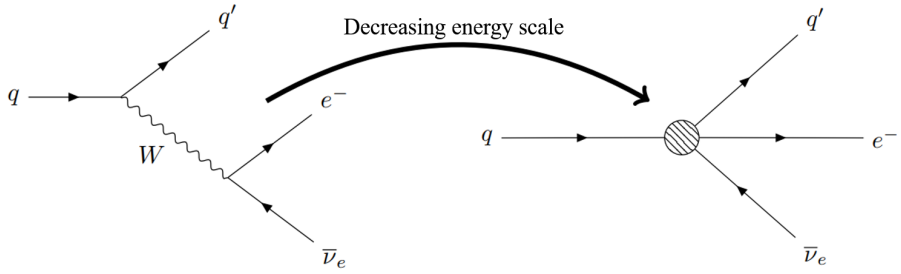

Figure 1.7: The Feynman diagram for neutron decay for high energy scale (left) and low energy scale (right).

The EFT provides an important and powerful technique for searching for areas of new physics and it has the following properties:

1. It is a model independent approach, which means it does not depend on whatever the underlying new physics is. Therefore, it allows a systematically search for new physics in SM processes;
2. It is able to impose the same strict symmetry requirements as in the SM;
3. It can recover the full SM theory in a very natural way;
4. It can be used to quantify the accuracy with which new physics can be excluded.

In this thesis the EFT approach is used for searching for new physics in top production which is reported in Chapter 6.

1.7 Summary

In this chapter, an introduction of the SM is delivered including the descriptions of the elementary particles and the three fundamental forces (electromagnetic, weak and strong). After that, a brief description about gauge symmetry is given. Due to its importance in this thesis, the Drell-Yan process is also introduced. Moreover, the photon

induced dilepton production process is shortly discussed. Finally, a basic introduction about the effective field theory is given.

Chapter 2

The beyond Standard Model of particle physics

This Chapter describes the motivations for new theories beyond the Standard Model (BSM) in Section 2.1 and introduces some BSM theories that predict the existence of heavy resonances which can decay to a dilepton final state[1] in Section 2.2. Finally, an introduction to the search for new physics in top quark production is given in Section 2.3.

2.1 Motivation for new physics

Despite the tremendous success of SM, there are still some shortcomings about it. Such as the existence of neutrino mass [9, 10], the existence of dark matter and dark energy [7, 8] and the matter-antimatter asymmetry. All of these observations can not be explained by SM. Besides, there are some limitations for the SM, such as lack of gravity description, convergence of the coupling constants [66]. Finally, the hierarchy problem and the existence of large number of free parameters in SM [53] make it looks unnatural. Therefore, it is commonly admitted that the SM is an effective model of a more fundamental theory at high energy. Each of these issues is shortly described in below.

- **Neutrino mass:** In the SM, the neutrino is assumed to be massless. However, it is observed that neutrinos can change from one flavour to another flavour which implies that they must have non-zero mass differences [9, 10] and their mass eigenstates are different from their flavour eigenstates. Experimentally, only upper limits on the neutrino masses have been set ($m < 2$ eV [53]). In addition, the differences between the neutrino squared masses have been measured: $\Delta m^2_{12} = (7.53 \pm 0.18) \times 10^{-5}$ eV^2 and $\Delta m^2_{32} = (2.44 \pm 0.06) \times 10^{-3}$ eV^2 [67].
- **Dark matter and dark energy:** Some astronomical observations show that the visible content of matter is only be $\sim 5\%$ of the total matter and energy content of the universe. Firstly, it is measured that the orbital velocities of stars around their galaxy center are too fast [7, 8] which is incompatible with the observed matter density in space. In order to solve the conflict between the experimental result and the theory prediction, the existence of "dark" matter which does not interact via electromagnetic or strong interaction has been proposed. Secondly, it is discovered that the universe is in accelerated expansion which means the galaxies are recede from each other and their escape rate increases with the distance [68, 69]. Giving these two cosmological

1. In this thesis, the expression "dilepton final state" denotes the decay in electron-positron pair (e^+e^-) or muon-antimuon pair ($\mu^+\mu^-$).

results, one can conclude that the matter (or energy) content of the universe is made of 5% ordinary matter, 25% dark matter and 70% dark energy which gives repellent force and is thought to be responsible for the observed accelerated expansion of the universe. However, the SM does not provides the candidates for dark matter as well as it can not explain the dark energy problem.

- **Asymmetry between matter and antimatter:** It is believed that matter and antimatter were produced with the same quantities at the time of Big Bang. However, we are living in a world composed with matter. So why does this happen and is it possible that some corners of the universe are dominated by antimatter? In 1967, Sakharov identified the three mechanisms necessary to obtain a global matter or antimatter asymmetry [70]:
  – Baryon and lepton number violation;
  – Interactions in the universe out of thermal equilibrium at a given moment of the universe history;
  – Charge (C) and charge-parity (CP) violation (the rate of a process $i \to f$ can be different from the CP conjugate process $\bar{i} \to \bar{f}$).

  The SM includes sources of CP violation: one is from a complex phase factor in the Cabibbo-Kobayashi-Maskawa (CKM) [71, 72] unitary matrix (which contains information on the strength of the flavour-changing weak interaction), and the other in the form of the QCD vacuum angle, θ_{QCD} [73]. However, they are not sufficient to explain the magnitude of the observed matter-antimatter asymmetry.
- **Free parameters of the SM:** There are 19 free parameters which have to be measured in the SM. The parameters include the masses of charged lepton, the masses of quark, the coupling constants of the three forces, the mass and vacuum expectation value of the Higgs boson, the mixing angles and the CP violating phase of the CKM matrix[2]. Due to this large number of free parameters, it is widely believed that there could be a more general and elegant theory than the SM. The list of parameters is summarized in Table 2.1.
- **Gravitational interaction and hierarchy problem:** The gravity, the fourth fundamental interaction, is not included in the SM because of its very small interaction strength compared with other three forces. The electromagnetic, weak and strong forces have similar strengths at the electroweak scale (energies of $\approx$ 100 GeV), but gravity is more than 10^{30} times weaker. The energy at which gravitational interactions becomes relevant is at the order of the Planck scale of $E_{Pl} = 10^{19}$ GeV, which is defined by the Planck mass, $M_{Pl} = \sqrt{\hbar c/G}$, where G is the gravitational constant. The huge difference between the electroweak scale and the Planck scale is also known as the hierarchy problem and it is deeply connected to the problem of the Higgs boson mass fine-tuning (which is expressed in the following).
- **Fine-tuning of the Higgs boson mass:** All the ingredients of the SM have been experimentally established after the discovery of the Higgs boson and the measurement of its mass ($\approx$ 125 GeV). All particles in the SM have a *bare mass* which is the mass obtained from the quantum propagator at the lowest order in perturbation theory. This is can be different from *physical mass* which contains higher order loop radiative corrections and can be measured in experiment.

  It is known [75] that the squared Higgs physical mass (m_{H}) can be obtained by the squared bare mass (m_0) of Higgs corrected with an extra term which includes higher order corrections (δm_{H}^2):

$$m_{\mathrm{H}}^2 = m_0^2 - \delta m_{\mathrm{H}}^2 \tag{2.1}$$

2. In addition, there are 7 parameters come from neutrino section: 3 neutrino masses, 3 mixing angles between different neutrinos, and 1 CP violating phase.

Quantity	Symbol	Value
Electron mass	m_{e}	511 keV
Muon mass	m_{μ}	105.7 MeV
Tau mass	m_{τ}	1.78 GeV
Up quak mass	m_{u}	2.3 MeV ($\mu_{\overline{\mathrm{MS}}}$=2 GeV)
Down quak mass	m_{d}	4.8 MeV ($\mu_{\overline{\mathrm{MS}}}$=2 GeV)
Strange quak mass	m_{s}	95 MeV ($\mu_{\overline{\mathrm{MS}}}$=2 GeV)
Charm quak mass	m_{c}	1.28 GeV ($\mu_{\overline{\mathrm{MS}}} = m_{\mathrm{s}}$)
Bottom quak mass	m_{b}	4.18 GeV ($\mu_{\overline{\mathrm{MS}}} = m_{\mathrm{b}}$)
Top quak mass	m_{t}	173.5 GeV
W boson mass	m_{W}	80.4 GeV
Z boson mass	m_{Z}	91.2 GeV
Higgs boson mass	m_{H}	125.09 GeV [74]
Higgs boson vacuum expectation value	v	246 GeV
Strong coupling constant	α_s	0.119 ($\mu_{\overline{\mathrm{MS}}} = m_{\mathrm{Z}}$)
QCD vacuum angle	θ_{QCD}	~ 0
CKM 12-mixing angle	θ_{12}	12.9°
CKM 23-mixing angle	θ_{23}	2.4°
CKM 13-mixing angle	θ_{13}	0.2°
CKM CP violating phase	δ_{13}	69°

Table 2.1: The SM parameters. The quark masses are presented in the renormalization scheme known as $\overline{\mathrm{MS}}$ [53].

where δm_{H}^2 includes all contributions from radiative corrections to the Higgs propagator. The main contributions include the top quarks, the Higgs boson itself, and the vector bosons. The related Feynman diagrams are shown in Figure 2.1, where the Higgs boson is denoted as h [53].

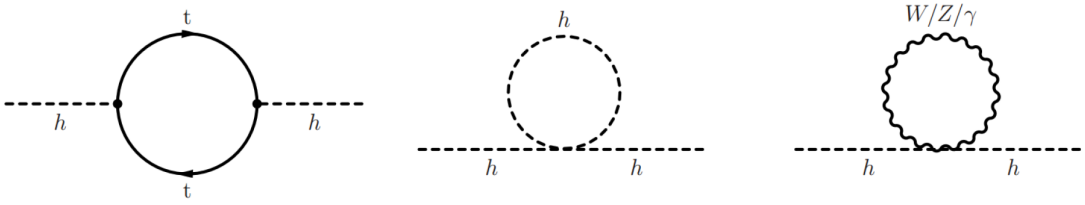

Figure 2.1: The Feynman diagrams of the processes which give the main divergent contributions to the Higgs boson mass.

However, the integrals corresponding to the amplitude of these processes are divergent, so a cut-off parameter Λ is introduced. The Λ represents the energy scale, up to which the SM is still valid. In principle, one can assume that the SM is valid up to the Planck scale at which gravitational effects cannot be neglected. With this assumption Λ would be of the order of $\approx 10^{19}$ GeV. The full calculation gives that δm_{H}^2 is proportional to Λ^2:

$$\delta m_{\mathrm{H}}^2 \propto \Lambda^2 \approx 10^{38}\ \mathrm{GeV}^2 \tag{2.2}$$

Because m_{H} is ≈ 125 ($\approx 10^2$) GeV, Equation (2.1) can be rewritten as:

$$10^4\ \mathrm{GeV}^2 \approx m_0^2 - \Lambda^2 \approx m_0^2 - 10^{38}\ \mathrm{GeV}^2$$

which means that m_0^2 is of the same order of Λ^2 (10^{38}) and these two terms cancel with a very high precision to obtain the value of the Higgs physical mass. This mathematical problem is called "Higgs mass fine-tuning" problem, although it does not

invalidate the theory. However, it seems an unnatural and implausible coincidence that m_0^2 cancels all the loop contributions up to this astonishing precision.

Actually, the problem comes from the choice of the Λ. If we choose the Λ to be ≈ 1 TeV, then the problem is solved because the cancellation is of the order of one over ten which seems is a natural and acceptable value. For this reason, if one accepts the Higgs mass fine-tuning argument, then the new physics will appear at the TeV scale, because at energy higher than $\Lambda = 1$ TeV, the SM is not valid anymore.

- **Convergence of the coupling constants**: In SM, the coupling strengths for electromagnetic interaction, weak interaction, and strong interaction have a close value at the energy scale $\mathcal{O}(10^{16})$ GeV. However, these three coupling strengths can not converge at a single point which is shown in Figure 2.2. In order to unify these couplings, an extension of the SM is needed and the new physics will be involved.

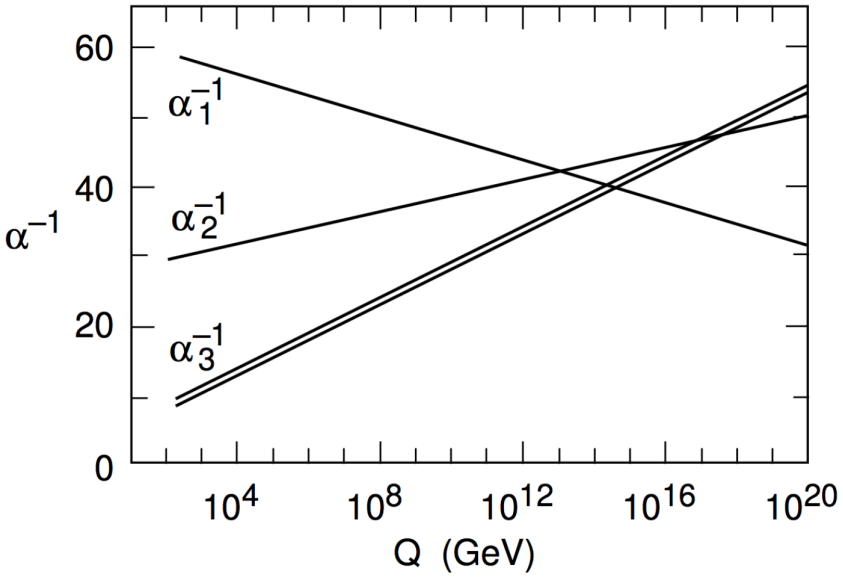

Figure 2.2: Evolution of the SM couplings $\alpha_i = \frac{g_i^2}{4\pi}$ as a function of the energy scale [66].

All these issues indicate that there must be new physics at a scale beyond the electroweak scale. Driven by the hierarchy problem, it is believed that there should be new physics at the TeV scale. Therefore, a discovery with direct searches at the Large Hadron Collider (LHC) could be possible.

2.2 New heavy particles decaying into a lepton pair

The most directly way to search for a new heavy particle decaying into a lepton pair would be searching for the "bump" (or localized excess) in dilepton mass spectrum at high mass. This study is motivated both theoretically and experimentally. From a theoretical point of view, many beyond SM (BSM) models predict the existence of new massive resonances which can decay into dilepton. Such as:

- The supersymmetric model [11] which is extremely attractive, because it can provide an explanation to the Higgs mass fine-tuning problem. In order to avoid issue of naturalness for the Higgs boson mass, the supersymmetric particles cannot be much heavier than 1 TeV which makes searching for these new particles is doable at LHC. Besides, the supersymmetric models provide a natural candidate of dark

matter. There are many supersymmetric models, the simplest one is the Minimal Supersymmetric Standard Model (MSSM) where each SM particle has a superymmetric partner. The fermions have spin 0 "sfermion" partners, gauge bosons have spin $\frac{1}{2}$ "gaugino" partners and the super-partner for higgs is "higgsinos". In SM, the baryon (B) number and lepton (L) number are conserved, where $\mathrm{B} = \frac{1}{3}(\mathrm{n_q} - \mathrm{n_{\bar{q}}})$ and $\mathrm{L} = (\mathrm{n_l} - \mathrm{n_{\bar{l}}})$. However, in the MSSM the B and L can be violated. In order to maintain the experimentally verified conservation laws, one defines a new conserved quantum number, the R-parity $\mathrm{P_R}$ as:

$$\mathrm{P_R} = (-1)^{3\mathrm{B}-\mathrm{L}+2\mathrm{s}},$$

where s is the spin of the particle. For the SM particles the $\mathrm{P_R} = +1$ and for supersymmetric particles the $\mathrm{P_R} = -1$. Due to the conserved $\mathrm{P_R}$, the supersymmetric particles can not decay into dilepton final states. However, in more complicated supersymmetric model [76] where the $\mathrm{P_R}$ can be violated, the new particles (e.g. "sneutrinos") can decay into dilepton final states.

- The Grand Unified Theory (GUT) [12, 13, 14] which tries to unify electromagnetic, weak, and strong interactions into one interaction through the extensions of the SM gauge group is also attractive. There are many GUT models, the starting point is the $SU(5)$ [77] model which was initially proposed by Georgi and Glashow in 1974. $SU(5)$ is the smallest gauge group that can contain the SM (can be expressed by $SU(5) \supset SU(3)_c \times SU(2)_L \times U(1)_Y$). It supposes the coupling strengths of electromagnetic (g_1), weak (g_2), and strong (g_3) interactions will merge into a single coupling (g_G) at the energy scale $\mathcal{O}(10^{16})$ GeV (which is called the unification scale). Besides, it predicts the $\sin^2\theta_\mathrm{W} = 0.375$ at the unification scale, and this was compatible with the measurements at that time. However, the $\sin^2\theta_\mathrm{W} = 0.375$ is now ruled out by most precise measurements and consequently g_1, g_2 and g_3 do not converge at single point. Moreover, in $SU(5)$ the decay of proton is allowed (e.g. $p \to e^+ + \pi^0$), while the predicted half time of the proton decay is several orders of magnitude smaller than the experimental lower limits. Therefore, it seems the simplest $SU(5)$ is not a correct GUT model.

  Another famous GUT model is $SO(10)$ [78] which was proposed by H. Fritzsch and P. Minkowski in 1975. In $SO(10)$ all matter particles belonging to the same generation are grouped into a single multiplet and has the nice feature to predict an half time for the proton decay which is not in contradiction with the experimental results. In $SO(10)$, the symmetries can be broken at different scale. For instance, the breaking scheme for $SO(10)$ could be:

$$SO(10) \to SU(5) \times U(1)_\chi \to G_{SM} \times U(1)_\chi$$

  where G_{SM} is $SU(3)_c \times SU(2)_L \times U(1)_Y$ and χ is the charge associated to the extra $U(1)_\chi$ group. There is no constraint on the breaking scale for this $U(1)_\chi$ and it might happen at the TeV scale.

  Moreover, the E_6 model [79] is also popular which is able to embed $SU(5)$ lie on the exceptional E_6 group. The E_6 group can break down to SM by following scheme:

$$E_6 \to SO(10) \times U(1)_\psi \to SU(5) \times U(1)_\chi \times U(1)_\psi \to G_{SM} \times U(1)_\chi \times U(1)_\psi.$$

  The new particle Z' from the linear combination of $U(1)_\chi$ and $U(1)_\psi$ is given by:

$$\mathrm{Z}' = \mathrm{Z}'_\psi \cos\theta_{\mathrm{E}_6} + \mathrm{Z}'_\chi \sin\theta_{\mathrm{E}_6}$$

  where $0 \leq \theta_{\mathrm{E}_6} \leq \pi$ is a mixing angle. Therefore, different θ_{E_6} will give different Z'.

Some specific Z′s are: 1, the Z'_{ψ} ($\theta_{E_6} = 0$) that only interacts through axial-vector couplings with the fermions and it is predicted by superstring theories [80]. 2, the Z'_{χ} ($\theta_{E_6} = -\frac{\pi}{2}$) that corresponds to a pure $U(1)_{\chi}$ group. 3, the Z'_{η} ($\theta_{E_6} = \arccos\sqrt{\frac{5}{8}}$), also suggested by superstring theories [80]. 4, the Z'_{I} (θ_{E_6}=($\arccos\sqrt{\frac{5}{8}}$)-$\frac{\pi}{2}$) that does not couple to up quarks and only couples to left-handed down quarks and right-handed leptons. The couplings between these specific Z′s and the up quarks, the down quarks, and the charged leptons can be seen in Table 2.2.

Last but not least, the Sequential Standard Model (SSM) [12] Z′ has couplings which are exactly the same as those of the SM Z (see Table 2.2), but is just heavier. This is not a real model but is very commonly used as a "standard candle" in experimental Z′ (or W′) searches.

Model	θ_{E_6}	c_V^u	c_A^u	c_V^d	c_A^d	c_V^l	c_A^l
Z'_{ψ}	0	0	0.300547	0	0.300547	0	0.300547
Z'_{η}	$\arccos\sqrt{\frac{5}{8}}$	0	0.380165	-0.285124	0.095041	0.285124	0.095041
Z'_{χ}	$-\frac{\pi}{2}$	0	0.073458	-0.416249	-0.342792	0.416249	-0.342792
Z'_{I}	$(\arccos\sqrt{\frac{5}{8}})-\frac{\pi}{2}$	0	0	0.620752	-0.620752	-0.620752	-0.620752
Z'_{SSM}	–	-0.227388	0.592979	0.410183	-0.592979	0.044592	-0.592979

Table 2.2: The specific Z′ bosons with corresponding θ_{E_6} form E_6 model together with the vector (c_V) and axial (c_A) couplings between the Z′ and the up quarks (u), the down quarks (d), and the charged leptons (l). The Z'_{SSM} from the Sequential Standard Model which has the same SM couplings is also shown.

- In order to explain the large difference between the electroweak scale ($\mathcal{O}(100)$ GeV) and the Planck scale ($\mathcal{O}(10^{19})$ GeV), the theories involve extra dimensions are proposed. It is assumed that there exist a spin 2 graviton (the carrier of the gravitational interaction and can decay into dielectron or dimuon pair) that can propagate in extra dimensions which have small radius R (R should be less than 100 μm [81] and assuming all the extra dimensions share the same radius), while SM forces are confined in usual 4-dimension spacetime. Due to the overlap of the wave functions of the SM particles with the graviton is small in 4-dimension spacetime, the gravity is much weaker than other three forces in our 4-dimension world. There are many extra dimensions models, such as ADD model (proposed by Arkani-Hamed, Dimopulous and Dvali) [15] which is one of the first solutions of the hierarchy problem by involving extra spatial dimensions. In ADD model, the R decreases with the increases of the number of extra dimension n. For $n = 1$, the R is around 10^{11} m at which distances the Newton's law is well established. However, the prediction from ADD for $n = 1$ gives deviations to the one from Newton's law. Therefore, $n = 1$ is excluded in ADD and n should start from 2 which corresponds to $R = 100$ μm which is just at the experimental limits [81]. Another type of extra dimension model using 5-dimensional warped geometry theory has been developed by Randall and Sundrum [16] called "Randall-Sundrum model". The Randall-Sundrum model uses "brane"[3] to describe SM particles ("Weakbrane") and graviton ("Planckbrane"), and gravity is much weaker on the Weakbrane than on the Planckbrane.

Generically, the spin 1 particle that can give rise to a resonance in the dilepton mass spectrum is called Z′, while for spin 2 particle it is called "graviton".

3. A brane is a physical object that generalizes the notion of a point particle to higher dimensions

From an experimental point of view, these BSM models of new physics give rise to high energy lepton pair in the final states and the SM background for such final states is relatively low at a hadron collider. Searches for heavy resonance decaying into dilepton have been performed at LHC and Tevatron. The CMS Collaboration at the LHC performed the search with proton-proton collision data collected at $\sqrt{s} = 7$ TeV [26, 27], with data collected at 8 TeV [28, 29], and using the combination of 2015 data collected at 13 TeV with data collected at 8 TeV [30]. Recently, CMS performed this search using the data collected at 13 TeV from 2016 [31] and 2017 [32], this search will be presented in detail in Chapter 5. Similar to CMS, the ATLAS Collaboration also performed the search with data collected at 7 TeV [33, 34], with data collected at 8 TeV [35], and with data collected at 13 TeV [36, 37]. At the Tevatron, the CDF and D0 Collaborations have published results based on a $p - \bar{p}$ collision sample at $\sqrt{s} = 1.96$ TeV, corresponding to an integrated luminosity of approximately 5 fb^{-1} [38, 39, 40, 41, 42, 43].

2.3 New physics in top quark production

As described in Section 2.2 if the new physics scale is available at hadron collider, the existence of new physics could be directly observed via the production of new particles. Otherwise, new physics could affect 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.

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. An effective field theory (EFT) (See Section 1.6) approach is followed in this thesis to search for new physics in the top quark sector in the ee and $\mu\mu$ final states (the experimental setup and results are shown in Chapter 6). In Refs. [82, 83] 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 [84]:

$$O_{\phi q}^{(3)} = (\phi^{+}\tau^{I}D_{\mu}\phi)(\bar{q}\gamma^{\mu}\tau^{I}q), \quad L_{eff} = \frac{C_{\phi q}^{(3)}}{\sqrt{2}\Lambda^{2}} g v^{2} \bar{b}\gamma^{\mu}P_{L}tW_{\mu}^{-} + \text{h.c.}, \tag{2.3}$$

$$O_{tW} = (\bar{q}\sigma^{\mu\nu}\tau^{I}t)\tilde{\phi}W_{\mu\nu}^{I}, \quad L_{eff} = -2\frac{C_{tW}}{\Lambda^{2}} v \bar{b}\sigma^{\mu\nu}P_{R}t\partial_{\nu}W_{\mu}^{-} + \text{h.c.}, \tag{2.4}$$

$$O_{tG} = (\bar{q}\sigma^{\mu\nu}\lambda^{A}t)\tilde{\phi}G_{\mu\nu}^{A}, \quad L_{eff} = \frac{C_{tG}}{\sqrt{2}\Lambda^{2}} v \left(\bar{t}\sigma^{\mu\nu}\lambda^{A}t\right) G_{\mu\nu}^{A} + \text{h.c.}, \tag{2.5}$$

$$O_{G} = f_{ABC}G_{\mu}^{A\nu}G_{\nu}^{B\rho}G_{\rho}^{C\mu}, \quad L_{eff} = \frac{C_{G}}{\Lambda^{2}} f_{ABC}G_{\mu}^{A\nu}G_{\nu}^{B\rho}G_{\rho}^{C\mu} + \text{h.c.}, \tag{2.6}$$

$$O_{u(c)G} = (\bar{q}\sigma^{\mu\nu}\lambda^{A}t)\tilde{\phi}G_{\mu\nu}^{A}, \quad L_{eff} = \frac{C_{u(c)G}}{\sqrt{2}\Lambda^{2}} v \left(\bar{u}\,(\bar{c})\,\sigma^{\mu\nu}\lambda^{A}t\right) G_{\mu\nu}^{A} + \text{h.c.}, \tag{2.7}$$

where $C_{\phi q}^{(3)}$ (ϕ is Higgs field, q is quark), 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. [82, 83]. The operators $O_{\phi q}^{(3)}$ and O_{tW} modify the SM interaction between W boson, top quark, and b quark (Wtb). 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 [85]. The operators O_{uG} and O_{cG}

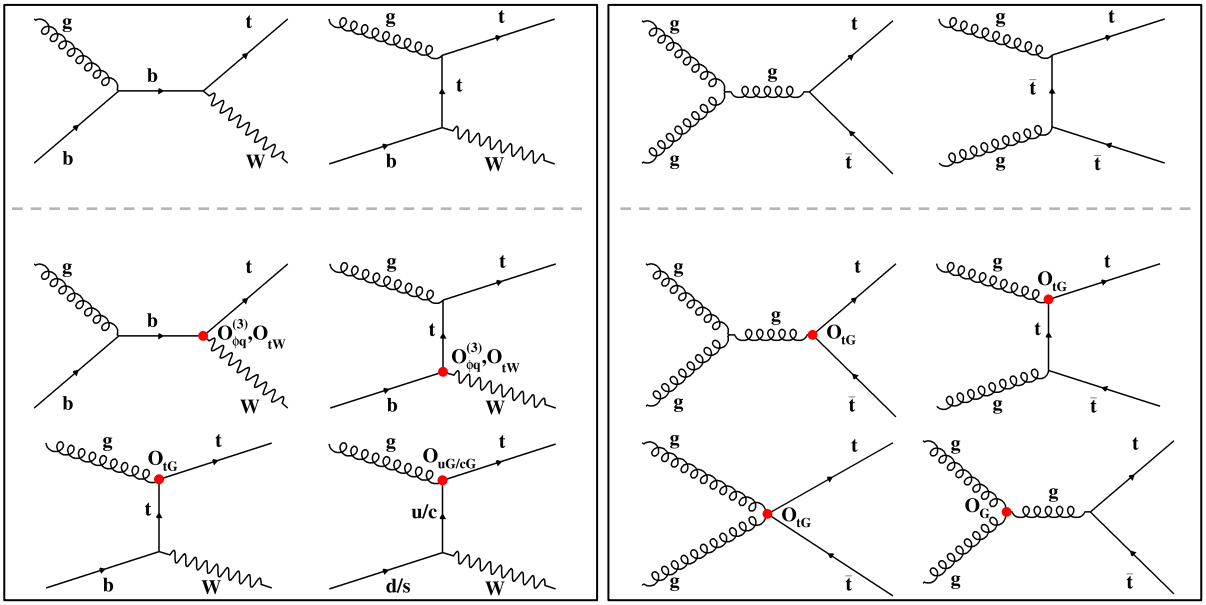

Figure 2.3: Representative Feynman diagrams for the tW (left panel) and $t\bar{t}$ (right panel) production at leading order. The upper row gives the SM diagrams, the middle and lower rows present diagrams corresponding to the $O^{(3)}_{\phi q}$, O_{tW}, O_{tG}, O_G, and $O_{u(c)G}$ contributions.

lead to flavor changing neutral current (FCNC) interactions of top quark and contribute to the tW production. As we know the FCNC processes do not exist at tree level in the SM and are induced only at loop level. Therefore the rates of FCNC processes are highly suppressed. The observation of such processes will be very important for searching new physics. The effect of introducing the new couplings $C^{(3)}_{\phi q}$, C_{tW}, C_{tG} and $C_{u(c)G}$ can be investigated in the tW production. The C_{tG} 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 Figure 2.3. 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 [82, 48]. 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 [48], ATLAS [46, 47] and CMS [44, 45] Collaborations. Direct limits on the top chromomagnetic dipole moment have been obtained by the CMS Collaboration at 7 TeV using top quark pair events [86]. Searches for top quark FCNC interactions have been performed at Tevatron [50, 51] and at LHC [44, 49] via single top quark production and limits are set on related anomalous couplings.

2.4 Summary

In this chapter the shortcomings of SM and motivations of BSM are delivered. After that some BSM theories which predict the existence of heavy resonances decaying into dilepton are introduced. Finally, an indirect search for new physics through top production is described.

Chapter 3

The CMS experiment at LHC

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

3.1 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 meters beneath the France-Switzerland border near Geneva shown in left plot of Figure 3.1 and it first started up on 10 September 2008. There are four main experiments at LHC which are shown in right plot of Figure 3.1, the general description about the four experiments are in the following:

1. **ALTAS** : 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 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 looks for clues of new physics. The CMS detector is described in Section 3.2.
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.

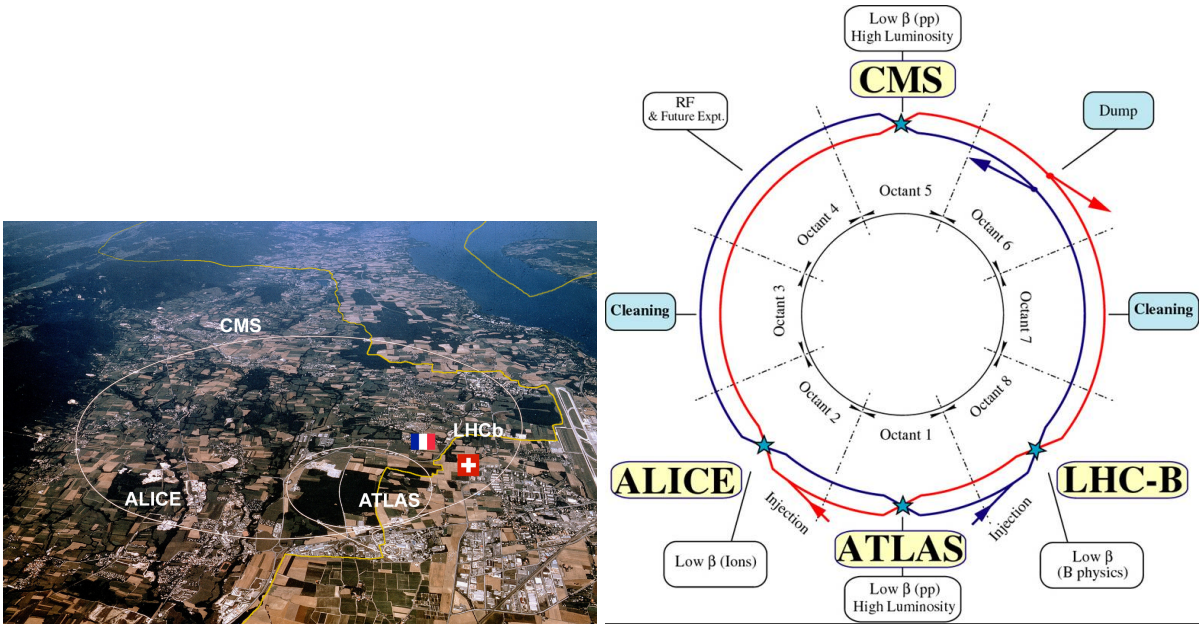

Figure 3.1: The overview of LHC (left) and the four main experiments in LHC (right) [17].

3.1.1 Proton proton collision

The reason to choose proton as accelerating particle in the LHC is that proton is stable, easy to get and has very low synchrotron radiation comparing to the 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 3.1 assuming $\mathbf{p_1}$ is the four-vector $\mathbf{p} = (E, \overrightarrow{p})$ for proton 1 and $\mathbf{p_2}$ is the four-vector for proton 2. For example when two 7 TeV protons collide, the E_{cm} is 14 TeV, but if one of the two protons is at rest, the E_{cm} is 114.6 GeV.

$$\begin{aligned} E_{cm}^2 &= (\mathbf{p_1} + \mathbf{p_2})^2 = (E_1 + E_2)^2 - (\overrightarrow{p_1} + \overrightarrow{p_2})^2, \\ E_{cm}^2 &= (E_1 + E_2)^2 \text{ when proton 1 and proton 2 collide,} \\ E_{cm}^2 &= (m_1^2 + m_2^2 + 2m_2 E_{1,lab}) \text{ when proton 2 is at rest.} \end{aligned} \tag{3.1}$$

The LHC acceleration chain for the protons is shown in Figure 3.2. At first the linear particle accelerator LINAC 2 generates 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 collision occur at the points of the main experiments (shown in right plot of Figure 3.1).

3.1.2 Pile up

In the LHC the collisions are between proton bunches and there are $\mathcal{O}(10^{11})$ protons per bunch. Depending on the instantaneous luminosity of the beam (see next section), there may be several proton proton collisions in the same bunch crossing. This phenomenon is so-called pile up, and is shown in the left plot of Figure 3.3. The pile up distributions for

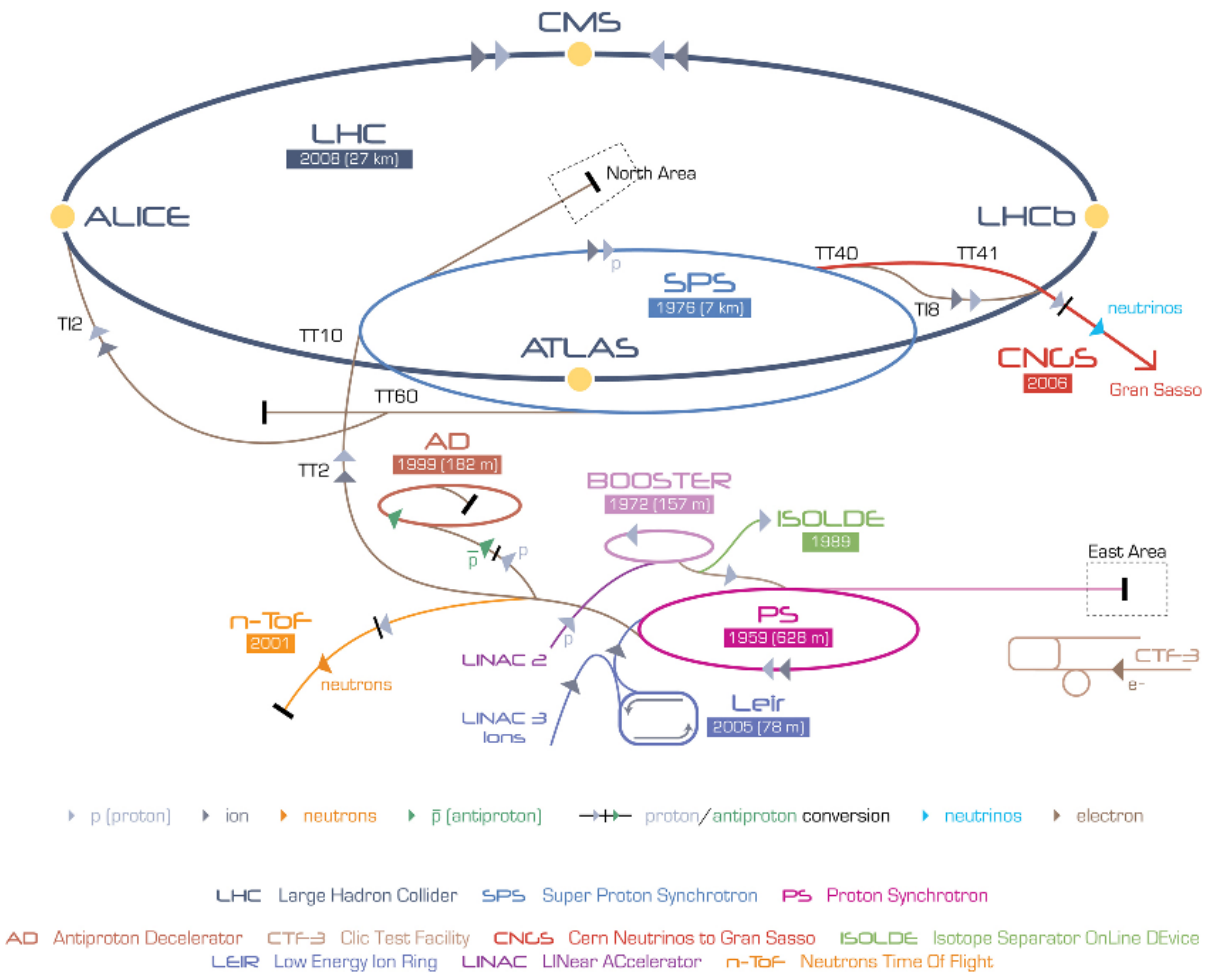

Figure 3.2: The overview of LHC accelerator chain [87].

proton proton collisions in 2016 and 2017 are shown in the middle plot and right plot of Figure 3.3.

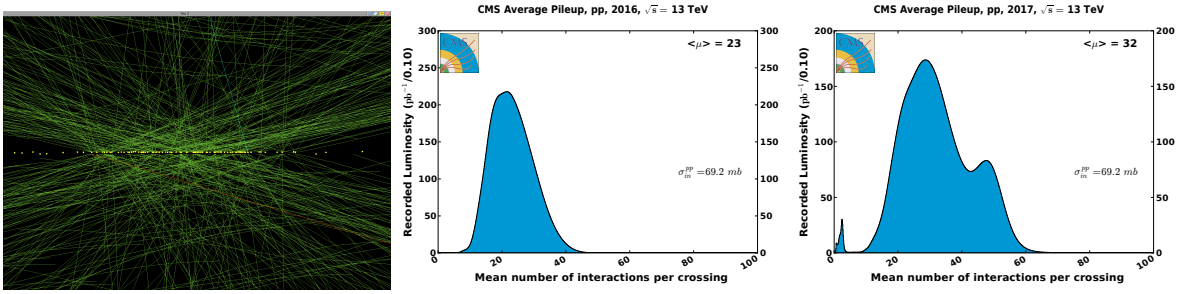

Figure 3.3: The phenomenon of pile up (left), the number of interaction per bunch crossing for 2016 (middle) and 2017 (right) [18].

3.1.3 Luminosity

The instantaneous luminosity ($\mathcal{L}$) is the proportionality factor between the number of events detected per second (dN/dt) and the interaction cross section (σ_p):

$$\frac{dN}{dt} = \mathcal{L} \cdot \sigma_p \tag{3.2}$$

The unit of the instantaneous luminosity is $cm^{-2}s^{-1}$. The instantaneous luminosity of two beams colliding head-on can be calculated by formula (3.3) [88], where N_1 and N_2 are the numbers of particles per bunch in the beam 1 and in the beam 2, N_b is the number of bunches in one beam, f is the beam revolution frequency, and σ_x and σ_y are the 1 σ gaussian widths of the bunch in the x axis and in the y axis directions, respectively (here the beam moving direction is along the z axis).

$$\mathcal{L} = \frac{N_1 N_2 N_b f}{4\pi\sigma_x\sigma_y} \tag{3.3}$$

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

$$\mathcal{L}_{int} = \int_0^T \mathcal{L}(t)dt \tag{3.4}$$

The cumulative online integrated luminosity that the LHC delivered and the one that CMS recorded for 2016 and 2017 are shown in Figure 3.4.

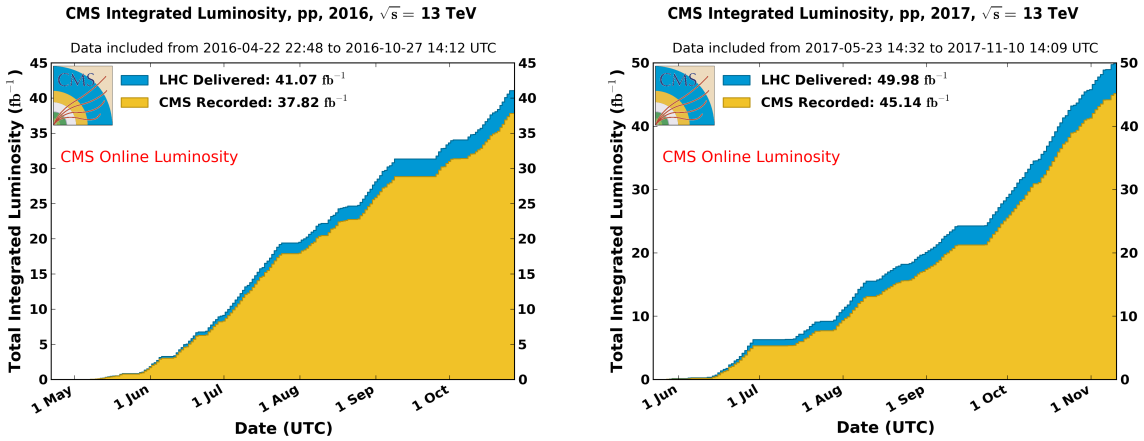

Figure 3.4: The cumulative online integrated luminosities that the LHC delivered and the one that CMS recorded for 2016 (left) and 2017 (right) [18].

The plan of data taking of the LHC and for the High Luminosity LHC (HL-LHC) phase are shown in Figure 3.5. For Run 2 (from years 2015 to 2018), LHC has delivered 156 fb^{-1} data which is more than the planned data which is 150 fb^{-1} data. The LHC data taking conditions from its start-up in 2010 to 2018 are shown in Table 3.1. 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 luminosity expected to be delivered by the LHC at the end of Run 3 in 2023.

3.2 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 new physics.

	year	$\sqrt{s}$	$\mathcal{L}_{int}$	bunch spacing	pile up
Run 1	2010	7 TeV	44.96 pb^{-1}	50 ns	-
	2011	7 TeV	6.10 fb^{-1}	50 ns	10
	2012	8 TeV	23.30 fb^{-1}	50 ns	21
Run 2	2015	13 TeV	4.21 fb^{-1}	25 ns	12
	2016	13 TeV	40.99 fb^{-1}	25 ns	23
	2017	13 TeV	49.79 fb^{-1}	25 ns	33
	2018	13 TeV	67.86 fb^{-1}	25 ns	32

Table 3.1: The LHC data taking conditions from its start-up in 2010 to 2018.

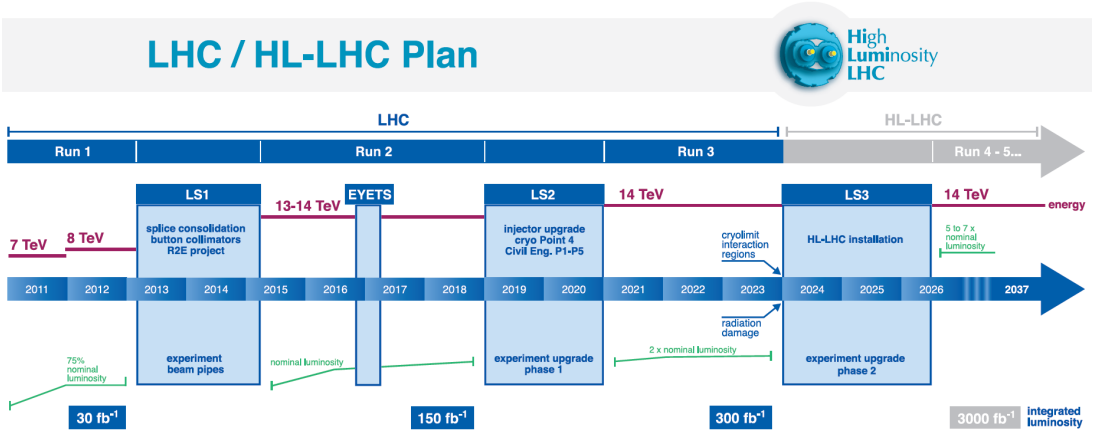

Figure 3.5: The data delivery plan of LHC and HL-LHC [19].

The CMS detector is built around a huge solenoid magnet. It is a cylindrical coil of superconducting cables that generates a field of 3.8 Tesla (T) and the field is confined by a steel "yoke" which has a 14,000-tonne weight. The complete detector is 21 meters long, 15 meters wide, and 15 meters high. A one-quarter cross-sectional view of the CMS detector is shown in Figure 3.6 together with some labels used to name the detector elements. The transversal view of the CMS detector is shown in Figure 3.7, together with the effect of the various sub-detectors response to the different incoming particles. The different elements of the CMS detector 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 the energy of electrons and photons;
3. **Hadronic calorimeter** which measures 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 identifies and measures the trajectories of muons.

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

Before moving to a detailed description of the CMS subdectors, the coordinate convention is described in the following section.

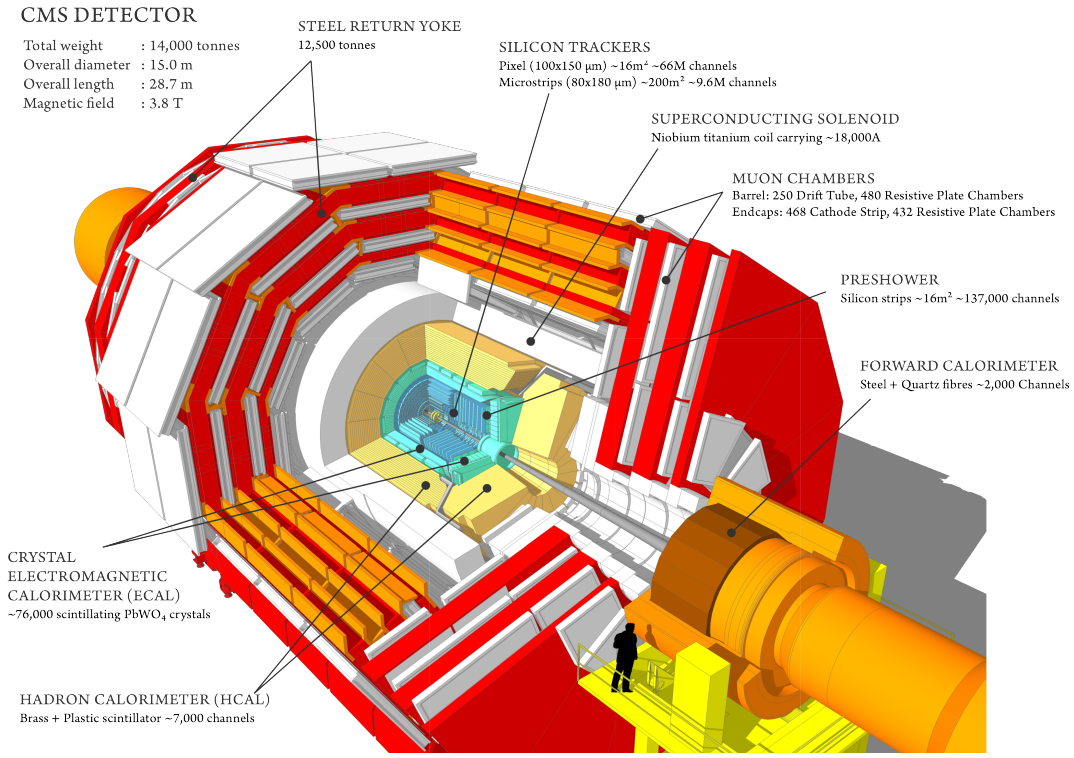

Figure 3.6: The exploded view of the CMS detector [21].

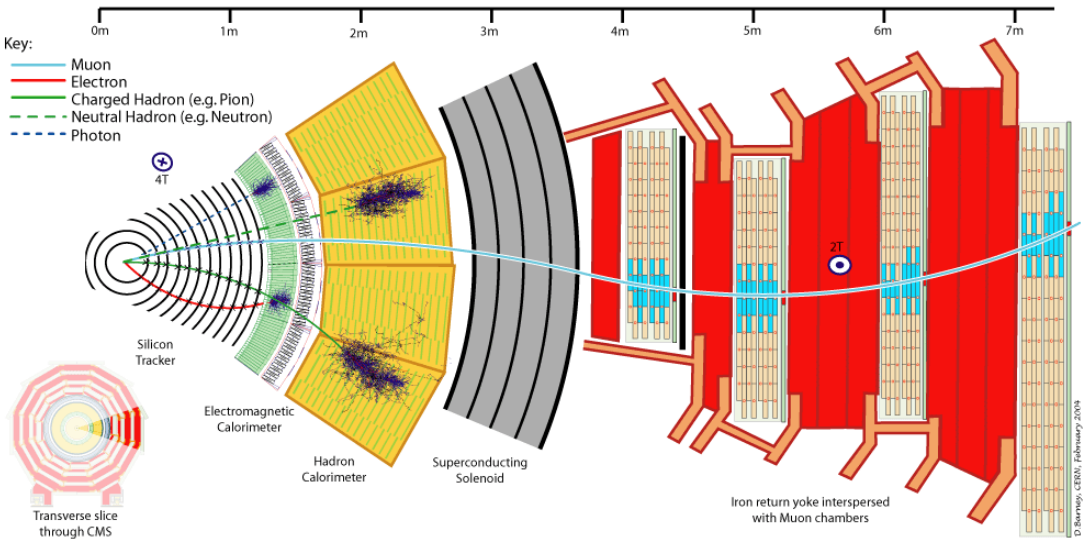

Figure 3.7: The transversal view of the CMS detector [22].

3.2.1 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 3.8. Pseudorapidity is defined as $\eta = -\ln \tan(\theta/2)$. The momentum of a particle measured transverse to the beam direction and in the beam direction, denoted by p_T and p_Z, respectively, are: $p_T = p\sin\theta$ and $p_Z = p\cos\theta$, where p is the magnitude of the 3-momentum of the particle $p = |\overrightarrow{p}|$.

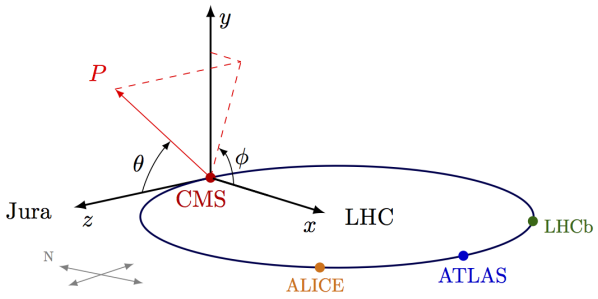

Figure 3.8: The CMS coordinate system.

3.2.2 Tracking system

The CMS tracking system is used for reconstructing the trajectories 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 3.2.5). Referring to Figure 3.9, 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 1×10^{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 3.10. 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 endcap 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 100×150 μm^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 μm in the r-ϕ plane and 20-40 μm 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 has significantly increased and it was around to 2×10^{34} $cm^{-2}s^{-1}$. The number of pileup events has then increased 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) consists of four layers in the barrel (BPIX), which represents an additional layer compared to the old detector. The radius of the innermost layer was reduced from 44 mm to 30 mm. In the endcap region (FPIX) a third disk was added per side. The new detector therefore allows a four-point coverage in the whole tracking region and the number of channels is almost double from 66 millions to 124 millions with keeping the same pixel size. A comparison of the old and new designs is shown in Figure 3.11.

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 millions 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-ϕ direction and 23 μm in z direction for TIB, for TOB it is from 35-52 μm in the r-ϕ 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.

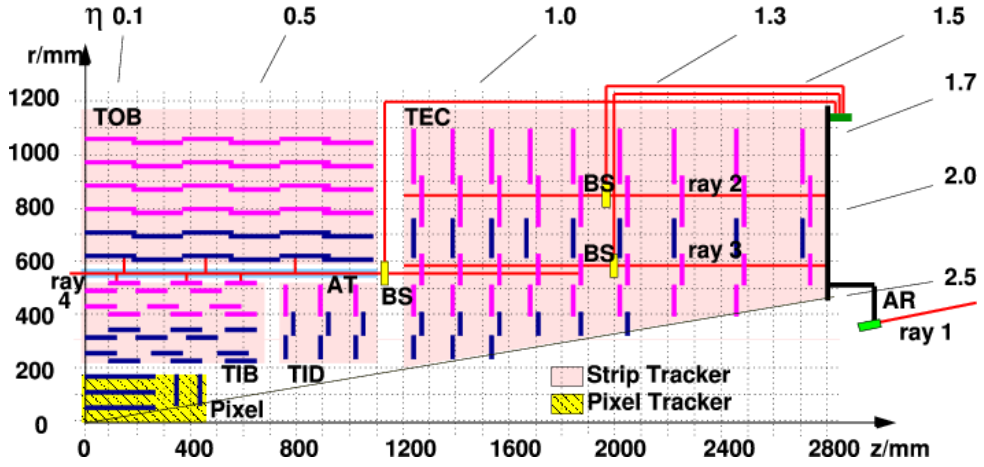

Figure 3.9: The tracker layout (1/4 of the z view) [20].

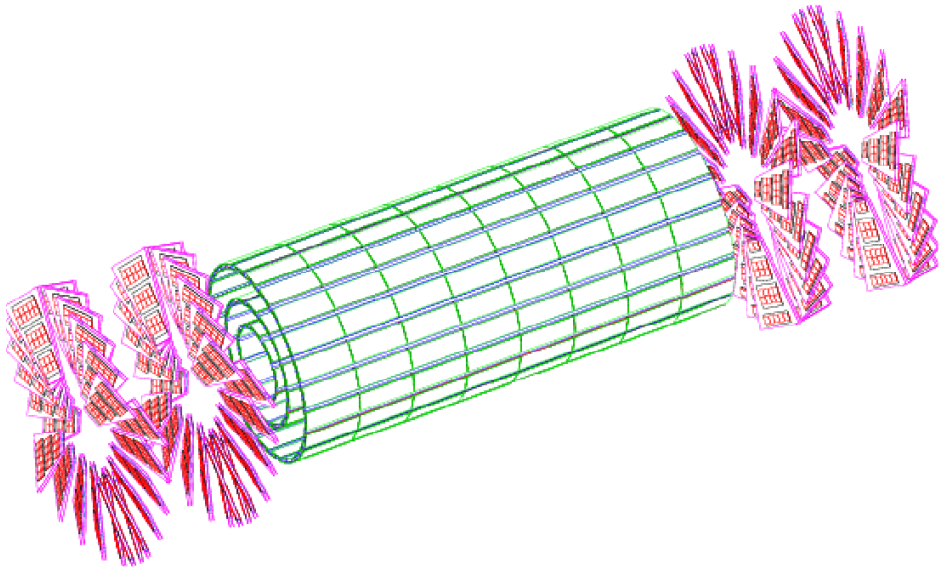

Figure 3.10: Layout of pixel detectors in the CMS tracker [20].

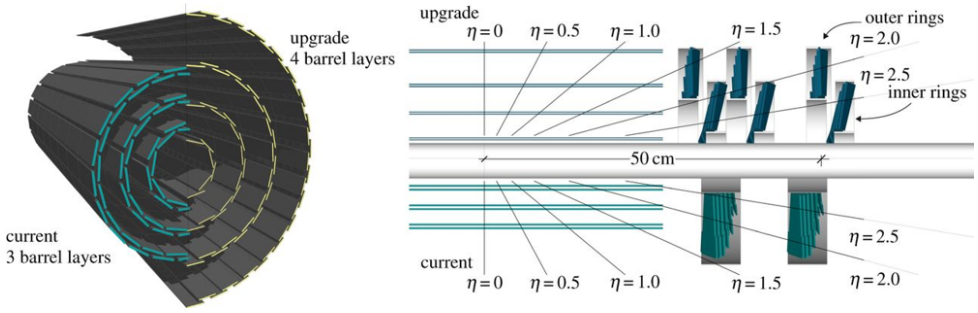

Figure 3.11: Comparison of the Phase 1 pixel detector (above the beam pipe) and the 2016 detector layout (under the beam pipe) [20].

3.2.3 Electromagnetic calorimeter

In order to measure the energy of electrons and photons the CMS detector uses 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 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 3.12. The reason to choose lead tungstate scintillating crystals for ECAL is because it has short radiation length ($X_0 = 0.89$ cm, for electron, the radiation lengths is the mean distance over which the electron loses all but 1/e of its energy by bremsstrahlung.) and Moliere radius (2.2 cm, Moliere radius is the radius of a cylinder containing on average 90% of the shower's energy deposition.). Besides, 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 3.13). The front face of the crystals is at a radius of 1.29 m and each crystal has a square cross-section of $\approx 22{\times}22$ 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 $\Delta\eta \times \Delta\phi = 0.0175 \times 0.0175$ (1°). 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 are grouped in 18 supermodules each subtending 20° in ϕ. Each supermodule comprises four modules with 500 crystals in the first module and 400 crystals in each of the remaining three modules (see Figure 3.14). For simplicity of construction and assembly, crystals are grouped in arrays of 2×5 crystals which are contained in a very thin wall (200 μm) alveolar structure and form a submodule.

The endcap part of the crystal calorimeter covers a pseudorapidity range from 1.48 to 3.0 (see Figure 3.13). The design of the endcaps provides precision energy measurement

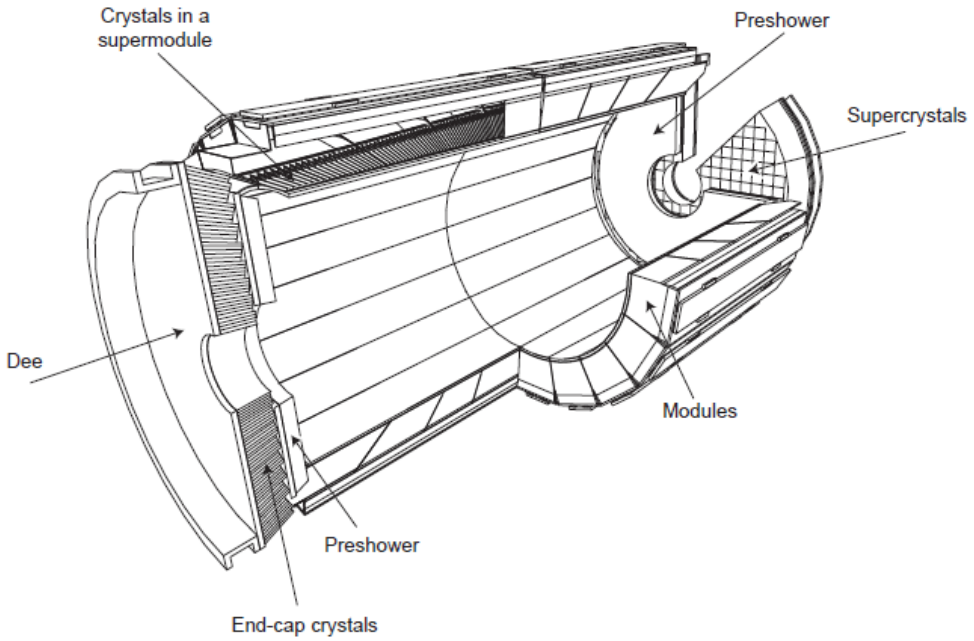

Figure 3.12: A 3D view of the barrel and endcap electromagnetic calorimeter [20].

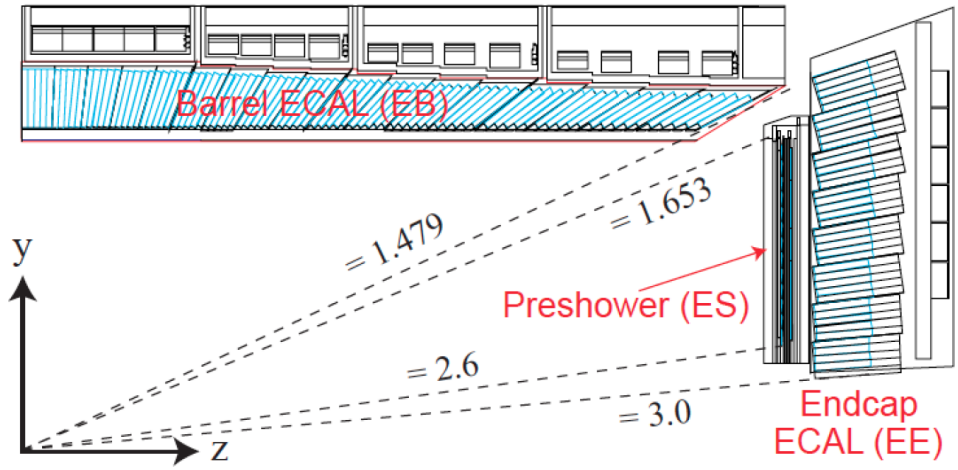

Figure 3.13: Longitudinal view of the electromagnetic calorimeter (one quadrant) [23].

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 ($24.7 \times 24.7 \times 220$ 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 3.15.

The endcap preshower covers a pseudorapidity range from $|\eta| = 1.65$ to 2.61 (see Figure 3.13). Its main function is to provide π^0-γ 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 mm. The

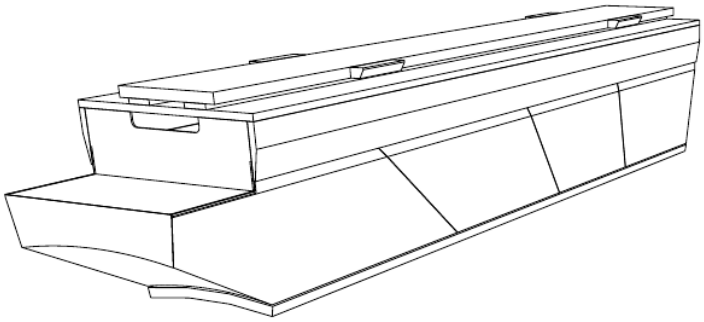

Figure 3.14: A view of the ECAL barrel supermodule which comprises four modules with 500 crystals in the first module and 400 crystals in each of the remaining three modules [24].

impact position of the electromagnetic shower is determined by the barycenter of the deposited energy and accuracy is typically 300 μm 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 3.16 shows the layout of the preshower.

Figure 3.17 shows the total thickness (in radiation lengths) of the ECAL as a function of pseudorapidity where the endcap part also includes the preshower detector.

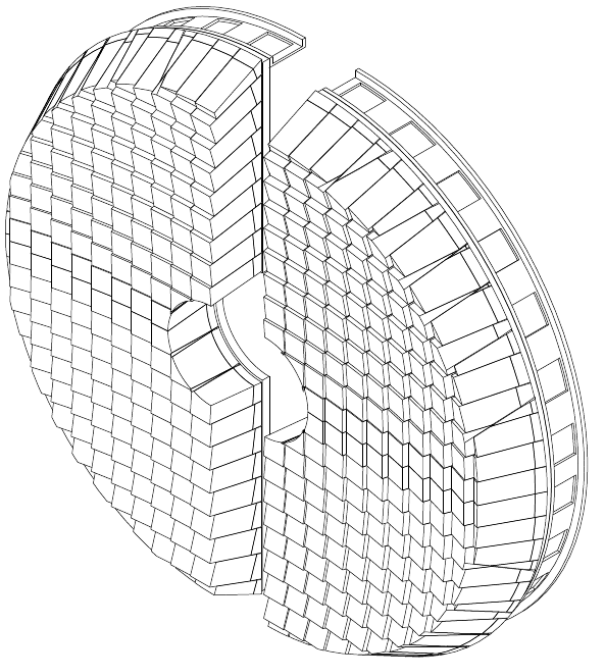

Figure 3.15: A single endcap with Dee-shaped [24].

The energy resolution of the ECAL can be parameterized by the following expression [24]:

$$(\frac{\sigma}{E})^2 = (\frac{S}{\sqrt{E}})^2 + (\frac{N}{E})^2 + C^2 \tag{3.5}$$

where σ is the energy resolution, S is the stochastic term, N is the noise term, and C is

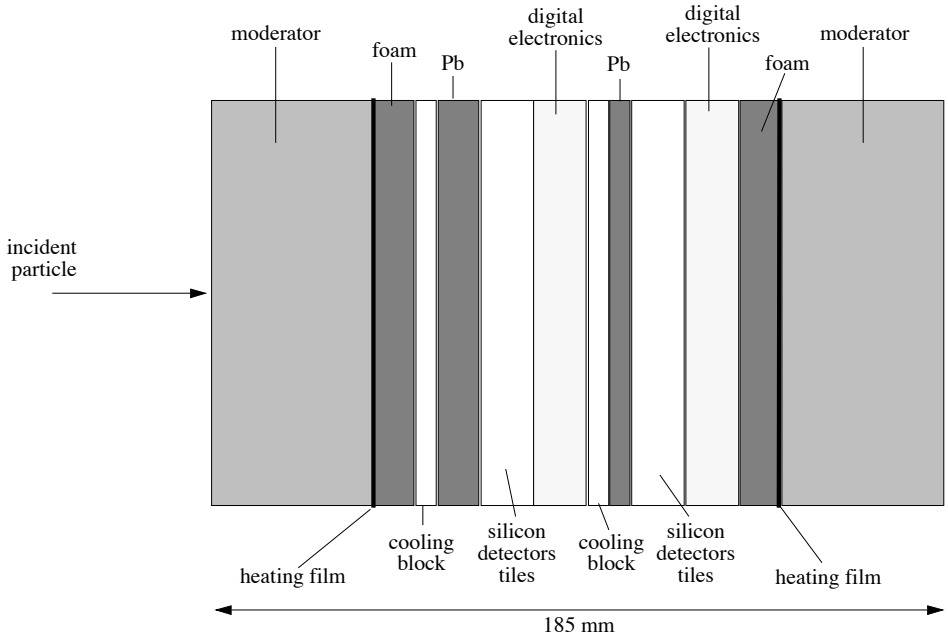

Figure 3.16: The components of the endcap preshower [24].

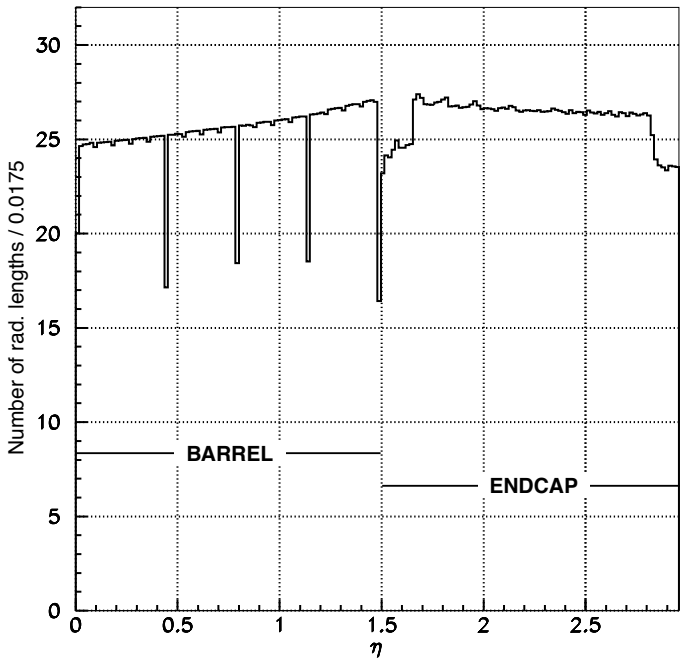

Figure 3.17: The thickness in X_0 of the ECAL as a function of pseudorapidity (averaged over ϕ) [24].

the constant term. 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; For instance, for 20 to 250 GeV of the electron test beam, with a 3×3 crystal configuration, the measured value of the parameters is S = 0.028 $\sqrt{(\text{GeV})}$, N = 0.12 GeV, and C = 0.003.

3.2.4 Hadronic calorimeter

In order to measure the energy of hadron (charged or neutral), CMS detector uses hadronic calorimeter (HCAL) which surrounds the ECAL system. The design of the 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 3.7). 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 (see Section 4.4). 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 and is relatively easy to mold and it is non-magnetic.

The architecture of the HCAL is illustrated in Figure 3.18. 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 [25]:

$$\frac{\sigma}{E} = \frac{X}{\sqrt{E}} \oplus 5\%, \quad \texttt{X=65\% (in barrel), 83\% (in endcap), 100\% (in forward)} \tag{3.6}$$

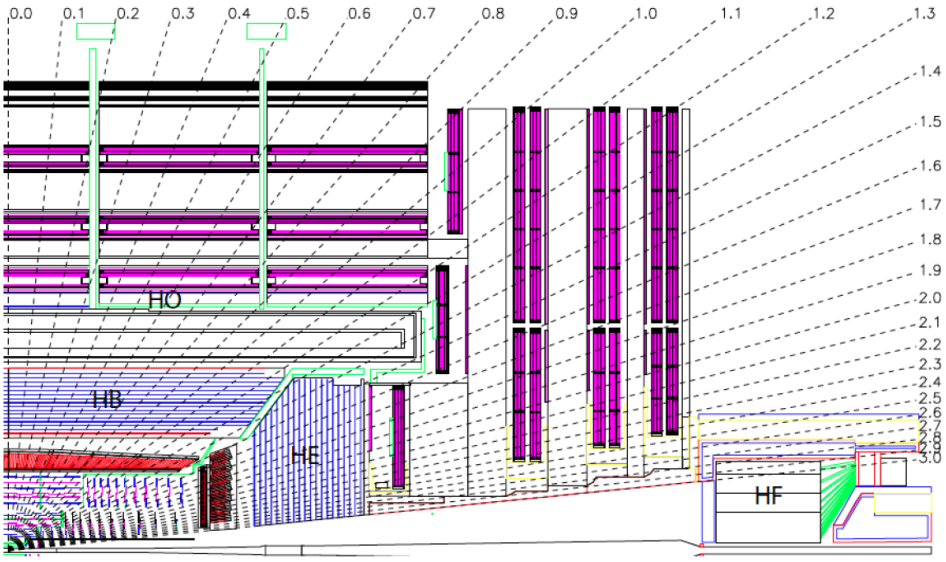

Figure 3.18: Longitudinal view of the CMS detector. The locations of the hadron barrel (HB), endcap (HE), outer (HO), and forward (HF) calorimeters are indicated [20].

3.2.5 Magnet

The CMS magnet is a superconducting solenoid magnet. Its length is 12.9 m and its inner diameter is 5.9 m. It can provide a 3.8 T strong magnetic field. The magnet in CMS is used to bend the track of charged particle and the radius of the track in transverse plane can be used to measure the transverse momentum of the charged particle by

$$p_T = 0.3R|Q|B$$

where p_T (GeV) is the transverse momentum, R (m) is the radius of the track in transverse plane, |Q| (e) is the absolute charge of the particle (e.g. for electron it is 1) and B (T) is the magnetic field strength. The design of the CMS magnet was driven by the required performance of the muon system (see next section). For example, for muon with momentum of 1 TeV, the $\Delta p/p$ should $\sim 10\%$.

3.2.6 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 its very fast time response. 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 3.19. The barrel of muon system (MB) is extended up to $|\eta|<1.2$, the endcap of muon system (ME) is for $1.2 < |\eta| < 2.4$.

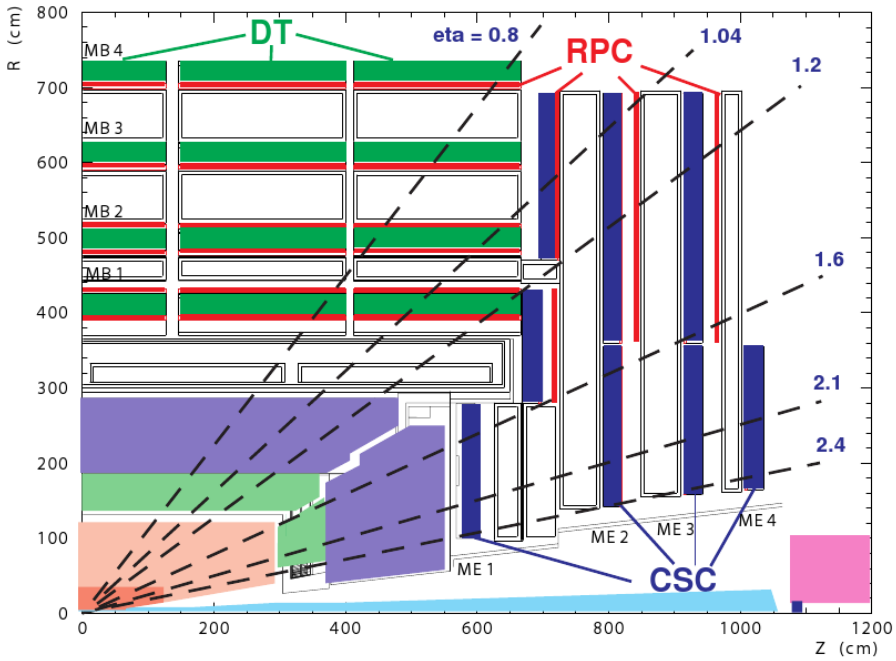

Figure 3.19: Longitudinal view (one quarter) of the muon system [20].

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 μm 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 μm 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 form the tracker, the muon system provides a resolution of about 10% for muons with $|\eta| < 2.4$ and $p_T < 200$ GeV.

3.2.7 Trigger

Because the proton bunch crossing frequency is 40 MHz at the CMS interaction point and the size of each collision event is around 1 MB, it is impossible to save all the events. Besides, most of the events from the collision are QCD events, which are less interesting for physics analysis. Therefore, the CMS developed a trigger system in order to save the events that are interesting for physics analysis. The trigger system is separated by to steps: Level-1 (L1) trigger and High-Level trigger (HLT).

The L1 triggers relies coarse information from calorimeters and the muon systems. Its decision is based on the p_T of specific objects like e/γ, muons and jets. Due to the large time consuming of tracker reconstruction algorithm, the information from track system is not used in the L1. After applying the L1 trigger, the output event rate is around 100 kHz.

The event passing the L1 will then asked to pass the HLT. At the LHT stage, the information from track system will be used and the reconstructed objects at the HLT will be close to the one used in physics analysis. After applying the HLT, the finally total event rate is around few hundred hertz. If the event rate after the HLT is still very high for some specific trigger path, then a prescale value will be applied to that trigger path to reduce the event rate. Finally, all the information from the event passing the HLT will be transferred to mass storage and saved for physics analysis.

3.3 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 collisions and pile up as well as the luminosity of LHC. After that a detailed description of the CMS detector is given which is composed of several subdetectors: the tracker system, the electromagnetic calorimeter, the hadronic calorimeter, the magnet, and the muon chambers.

Chapter 4

Object reconstruction

This chapter describes how the physical objects are reconstructed using the combined information coming from the CMS subdetectors. The electron and photon reconstructions are described in Section 4.1, the muon reconstruction in Section 4.2, the jet and b-tagged jet reconstructions in Section 4.3, and the missing transverse energy reconstruction in Section 4.4. Finally, a short introduction to the particle-flow particle is presented in Section 4.5.

4.1 Electrons and Photons

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, an electron can radiate bremsstrahlung photons when it traverses the material between the interaction point and the ECAL, those photons can convert into electron pairs when they traverse 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 electrons and photons in the 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 of a track which matches to the SC in ECAL allows one to distinguish an electron from a photon.

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.

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 4.1: 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 is 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.

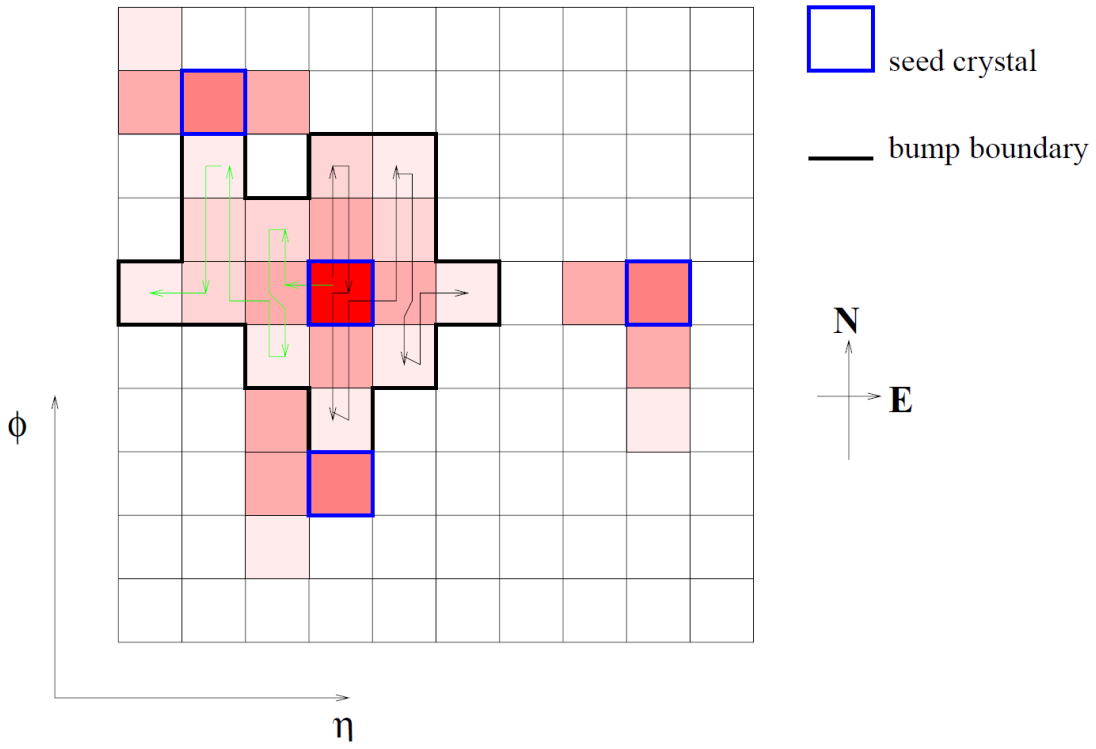

Figure 4.1: Illustration of the Island clustering algorithm in the barrel ECAL [89].

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 is 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.

The Hybrid algorithm

In the case of an 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 4.2.

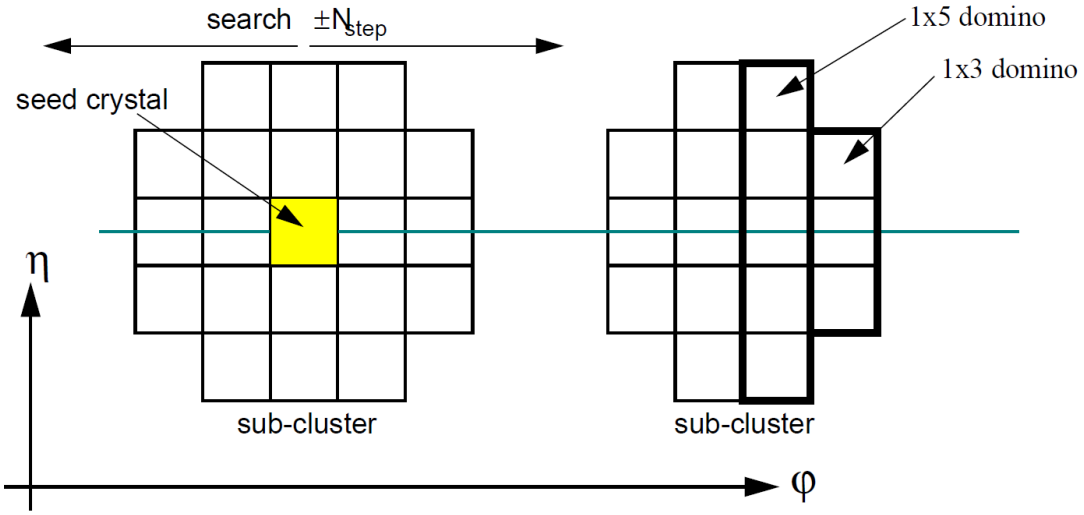

Figure 4.2: Domino construction step of Hybrid algorithm [89].

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 4.1.

Table 4.1: 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

Super-cluster (SC)

A possible approach to recollect 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 4.3.

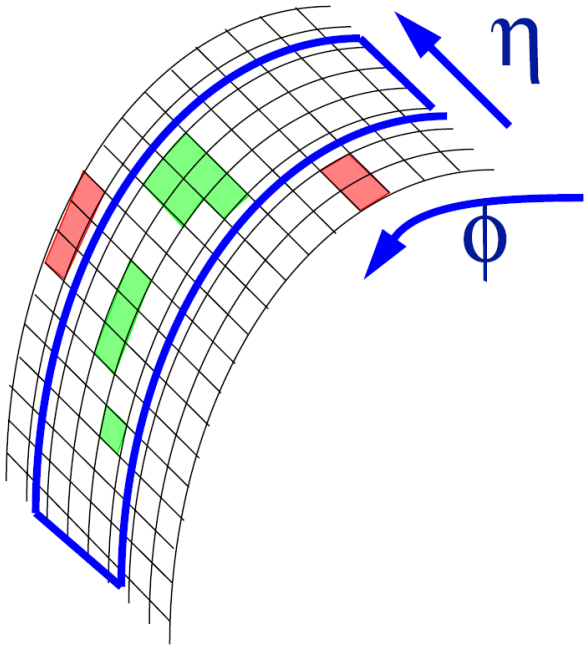

Figure 4.3: Illustration of a super-cluster algorithm collects all clusters which satisfies a given geometric condition (e.g. lying in a certain region around the seed cluster) [89].

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 4.4. A depth *tmax* 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:

$$x = \frac{\Sigma x_i \cdot W_i}{\Sigma W_i} \tag{4.1}$$

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

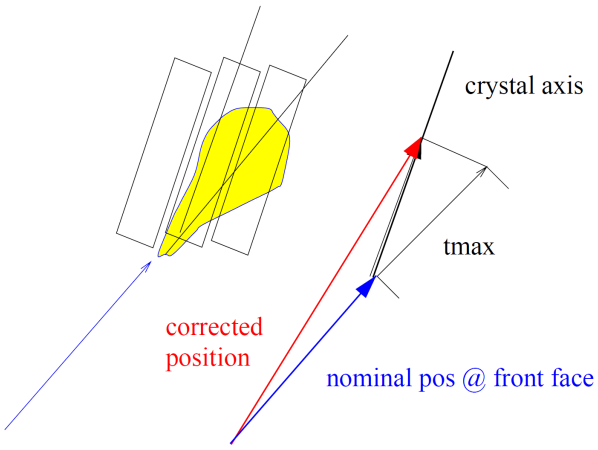

Figure 4.4: Illustration of the crystal offpointing [89].

formula:

$$W_i = W_0 + ln(\frac{E_i}{\Sigma E_j}) \tag{4.2}$$

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.

4.1.1 Electrons

In order to reconstruct electron we need to find a track in the tracker which matches to the SC. To search for the track, there are four steps: 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.

Track seeding

Track seeds are the starting points 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 4.1) 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 (assuming a successful collection of photons) coincides 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 [90]. 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. For instance, in searching for heavy resonances decaying into dielectron final state study (see Chapter 5), the ECAL driven seeding is choosed.

Tracking

After the track seeding, the tracking procedure consists of "track build" and "track fitting" starts. The "track build" uses combinatorial track finder method (CTF) [91] to build the track which outwards from the seed. The "track fitting" use a Gaussian sum filter (GSF) method [92] to fit the track in backward.

One of the great benefit of the tracks come from GSF method is that track 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.

Track-supercluster matching

The finally step to build GSF electron candidates is to associate the track to a SC.

For ECAL driven tracks, the difference between the position of the track at the SC which is the extrapolated from the innermost track and the energy weighted position of SC 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 is used which combines track observables and SC observables to get a global identification variable which gives whether the matching is successful or not.

4.1.2 Photons

The photon is the simplest particle. Any reconstructed SC with $p_T > 10$ GeV is considered as a photon candidate [93]. In order to reduce the fake photon from other objects, such as electron, hadron and jet, the most important tool can be used is isolation requirement. Fake photon from jets can usually be rejected by looking for additional energetic particles in a cone around the reconstructed ECAL cluster. Charged particles like electrons, 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 reconstructed 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 as well as hadronic energy deposits in the HCAL. It is also worth mention that at analysis level many other variables (like shower shape) will be used to distinguish electron and photon from jet efficiently.

As we known, the photon can be converted (produce an electron pair) in the material before reaching ECAL. Therefore a dedicated algorithm for tagging the converted photon by searching for conversion tracks matching the ECAL SC is used in CMS.

4.2 Muons

The muon reconstruction uses the information from the muon system and the silicon tracker. For the muon which is reconstructed using only muon system information is called "standalone muon". Muons which are 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 E_T muon candidates. The efficiency for reconstructing a muon as global or tracker muon is as high as 99%.

Standalone muons

The standalone muon reconstruction only uses information from the muon detectors (the silicon tracker is not used). Both tracking detectors (DT and CSC) and RPCs (see Section 3.2.6) 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.

Global muons

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. Then 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.

Tracker muons

For low E_T muons (below 6-7 GeV), a 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 [94] and hence improves 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".

4.3 Jets and Bjets

The jet producing (the QCD process) 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). There are several jet reconstruction methods and the most commonly used is the anti-k_T algorithm [95]. This algorithm defines two variables, one is the distance (d_{ij}) between the PF object (i) (the object reconstructed by particle-flow algorithm will be explained in Section 4.5) and the pseudojets (j), another is the distance (d_{iB}) between the PF object (i) and the beam (B):

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

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 depending on the analysis. If the d_{ij} smaller than d_{iB} then the i and j are combined to form a new single entity. Otherwise, the i entity is considered as a jet and removed from the list of entities. The procedure continues until the entity list is empty.

4.3.1 b-jets

All the quarks will generate a jet except the top quark which decays before hadronizing. The jets from b quark can be distinguished from other jets coming from gluons, light-flavor quarks (u, d, s) and c quark fragmentation by using vertex, track, and identified lepton information. There are several algorithms to identify b-jets and the most commonly used in CMS analyses is the Combined Secondary Vertex (CSV) algorithm [96]. Because b hadrons typically have a lifetime of $c\tau \simeq 450$ μm, the existence of a secondary vertex is a powerful variable to discriminate between b jets and other jets. The definition of a secondary vertex is that 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 its radial distance exceeds 2.5 cm and its mass is compatible with a K^0 or greater than 6.5 GeV, then the secondary vertex is rejected. The last requirement for secondary vertices is that the flight direction of each candidate should be in a cone with $\Delta R = 0.5$ around the jet direction. However, if no secondary vertex is found, the CSV algorithm will use so-called "pseudo-vertices" (the impact parameter of tracks is more than 2σ away). Moreover, if no pseudo-vertex is found, the CSV algorithm proceeds from simple track variables. 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%.

4.4 Missing transverse energy

As we known, at the interaction point the transverse energy of two proton beam is close to zero. Therefore, after the collision, the total transverse energy should also close to zero when all final state particles are taken into account. However, if the final state contains neutrinos and other hypothetical neutral weakly interacting particles that cannot be detected by CMS, then the total transverse energy will be imbalanced (or not be zero) which means we missed some transverse energy and this is where the name of "missing transverse energy" (MET or $\not{E}_T$) comes from. The mathematically definition of MET is

$$\overrightarrow{\not{E}_{\mathrm{T}}} = -\sum \overrightarrow{\mathrm{E_T}^{exists}}$$

with $\overrightarrow{\mathrm{E_T}^{exists}}$ be the vectorial $\mathrm{E_T}$ of all visible particles reconstructed by CMS detector.

4.5 Particle-flow algorithm

Before ending this chapter, it is worth to mention the particle-flow (PF) algorithm [97, 98] 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). Besides, it avoids double counting scenario, such as a track can be assigned to only one PF particles. Finally, the individual PF particles can be combined to form more complex objects such as hadronically decaying taus, jets or transverse missing energy.

In this thesis, the jets and MET are reconstructed based on PF particles. The muons are required to be both PF muon and global muon. The electrons are from the nominal reconstruction algorithm.

4.6 Summary

This chapter describes how the objects used for analyses are reconstructed in CMS. The electron reconstructions have been described in details in Section 4.1. The reconstruction of photon is introduced in Section 4.1.2, while the muon, jet and missing transverse energy reconstructions are described in Sections 4.2–4.4. Finally, a brief introduction about particle-flow algorithm is presented in Section 4.5.

Chapter 5

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. We present in this chapter a search for high mass resonances in the dielectron final state. This analysis uses proton-proton collision data at the centre-of-mass energy of 13 TeV collected by the CMS experiment in 2016 and 2017, corresponding to an integrated luminosity of 35.9 fb^{-1} and 41.4 fb^{-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. The data and MC samples are described in Section 5.1. Introduction to the triggers used is presented in Section 5.2. The object and event selection are expressed in Section 5.3. The mass scale and resolution studies are introduced in Section 5.4. The measurement of High Energy Electron Pair selection efficiency is presented in Section 5.5. The SM backgrounds estimation is expressed in Section 5.6. In Section the final dielectron invariant mass spectra are given. Finally, the statistical interpretation is presented in Section 5.8.

5.1 Data and MC samples

The name and integrated luminosity of all the datasets used in this analysis are summarized in Table 5.1. The DoubleEG dataset requires at least two trigger level electrons or photons for each event and it is used for the main analysis. The SingleElectron dataset requires at least one trigger level electron for each event and it is used for trigger efficiency measurement and electron selection efficiency measurement. The SingleMuon dataset requires at least one trigger level muon for each event and it is used for performing the $t\bar{t}$ background cross check using the "eμ" 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 35.9 fb^{-1} and 41.4 fb^{-1} collected by the CMS experiment in 2016 and 2017, respectively. Only certified data which are recommended for physics analysis are used.

The Monte Carlo (MC) simulated samples used in the main analysis for 2016 and 2017 are summarized in Table 5.2 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 is 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 by di-boson

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

Table 5.1: The various datasets using in the analysis and their integrated luminosities. X = DoubleEG is for the main analysis, X = SingleElectron is for the trigger and electron selection efficiency, X = SinglePhoton is for the fake electron study and X = SingleMuon is for the eμ study.

(WW, WZ, ZZ) samples, finally by the gravitation signal sample. A similar organization is given for the 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) campaign. Most of the samples are generated by POWHEG v2 [99, 100, 101, 102, 103, 104] at next-to-leading order (NLO), few are generated by MadGraph5_aMC@NLO [105] at NLO. Pythia8 [106] is used to simulate the parton showering and hadronization. For detector response it is simulated by Geant4 [107].

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 5.1 for 2016 and 2017. MC events are re-weighted to account for the pile up difference between data and MC.

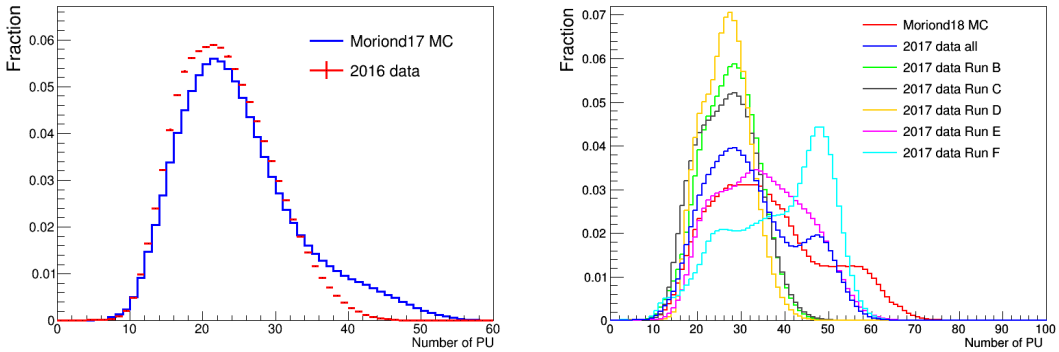

Figure 5.1: Pile up distributions for data and MC samples in 2016 (left) and in 2017 (right) .

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 electron-photon (called EGamma) group scale corrections (except for the study presented in Section 5.4, for which we measured the mean data energy correction and found it agrees with the offical EGamma value). In 2017 the official EGamma energy scale in data and energy smearing in MC are 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 $\rightarrow \tau\tau$)	5765.4	NNLO
	TTTo2L2Nu_TuneCUETP8M2_ttHtranche3_13TeV-powheg	87.31	NNLO
	TTToLL_MLL_500To800_TuneCUETP8M1_13TeV-powheg-pythia8	0.326	NLO
	TTToLL_MLL_800To1200_TuneCUETP8M1_13TeV-powheg-pythia8	3.26E-2	NLO
	TTToLL_MLL_1200To1800_TuneCUETP8M1_13TeV-powheg-pythia8	3.05E-3	NLO
	TTToLL_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-*_TuneCUEP8M1_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 $\rightarrow \tau\tau$)	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
	WW_TuneCP5_13TeV-pythia8	118.7	NLO
	WZ_TuneCP5_13TeV-pythia8	47.13	NLO
	ZZ_TuneCP5_13TeV-pythia8	16.523	NLO

Table 5.2: MC samples used in the main analysis.

5.2 Trigger

The primary high level trigger (HLT) used in the 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) [108] 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 the 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 is mean to mitigate the loss of efficiency for high E_T electron. The exact unprescaled threshold of each of those seeds changes 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 is computed as a product of 3 terms: the L1 trigger efficiency, the HLT supercluster E_T filter efficiency (the HLT turn on curve) and the online electron identification (CaloIdL+MW or GsfTrkIdVL) efficiency. For the final results only E_T dependent efficiency is used to weight MC events, other efficiencies are cancelled in the normalisation of MC events to data in the Z peak region (M_{ee} in 60-120 GeV). The method to measure the efficiencies is described in Section 5.2.1. The L1 trigger efficiency of primary signal trigger is shown in Section 5.2.2. The HLT E_T turn on curve and HLT identification efficiency of primary signal trigger are presented in Section 5.2.3. Other trigger efficiencies are given in Section 5.2.4.

5.2.1 Method for Measuring Trigger Efficiencies in Data

The tag and probe method [109] is used to measure the efficiency of a given selection in data. There are reasons to use tag and probe method. First one is the tag can be asked to match with trigger, by doing this, the probe will not be biased by trigger requirement. Second one is by requiring tight selection on tag, the probe will most likely be real electron not the jet fake electron. The events are selected by HLT_Ele27_eta2p1_WPTight (which requires 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 (which requires 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 passes HLT_Ele27_eta2p1_WPTight (HLT_Ele35_WPTight) in 2016 (2017) and passes HEEP ID (as defined in Section 5.3) and is in barrel of ECAL. The probe is the electron which passes 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 back-

ground subtraction is applied, nor any mass window cut is 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 unseeded leg is unbiased by L1 seeded trigger. The efficiency is equal to the number of passing probes divided by all probes shown in Equation 5.1.

$$\epsilon = \frac{N_{passing\ probes}}{N_{all\ probes}} \tag{5.1}$$

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 5.2, 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 values where the efficiency reaches half maximum and A2 and B2 are the turn on of the curve.

$$f(\mathrm{E_T}) = 0.5 \cdot A0 \cdot (1 + erf(\frac{\mathrm{E_T} - A1}{\sqrt{2} \cdot A2})) + 0.5 \cdot B0 \cdot (1 + erf(\frac{\mathrm{E_T} - B1}{\sqrt{2} \cdot B2})) \tag{5.2}$$

5.2.2 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 5.2. From applying to MC events, this translates to an efficiency of 99.5% to select barrel-barrel events and 98.8% to select barrel-endcap events in a mass range of 60 to 120 GeV and a ∼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.

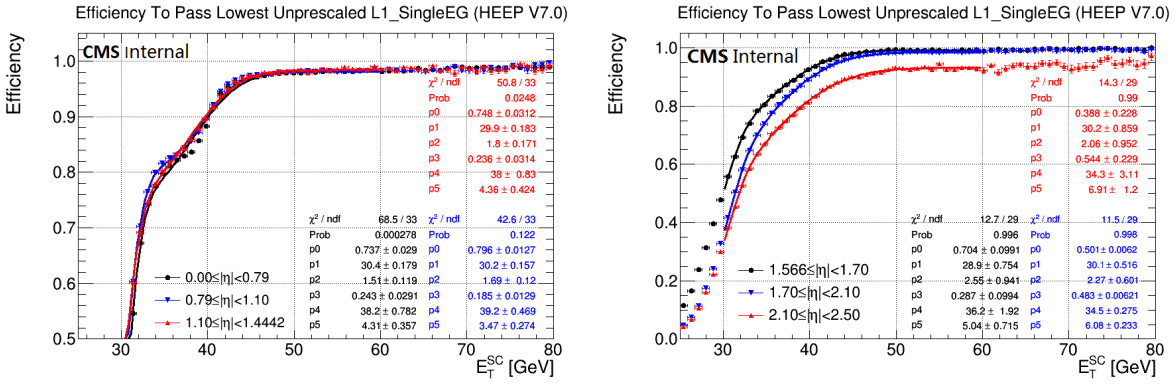

Figure 5.2: The efficiency for an electron passing HEEP to pass the lowest unprescaled L1 SingleEG seed versus supercluster $\mathrm{E_T}$ and η in barrel (left) and in endcap (right) for 2016 [110].

In 2017 the efficiency for a HEEP electron to pass the lowest unprescaled L1 SingleEG seed is shown in Figure 5.3. 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 events and 67% to select barrel-endcap events with the worst case where the supercluster $\mathrm{E_T}$ of both electrons is 35 GeV. The efficiency is higher than 99.5% for two electrons with supercluster $\mathrm{E_T}$ more than 42 GeV in barrel

and more than 47 GeV in endcap. This is a lower bound on the efficiency, becasue there is the DoubleEG L1 seed which will further increase the efficiency.

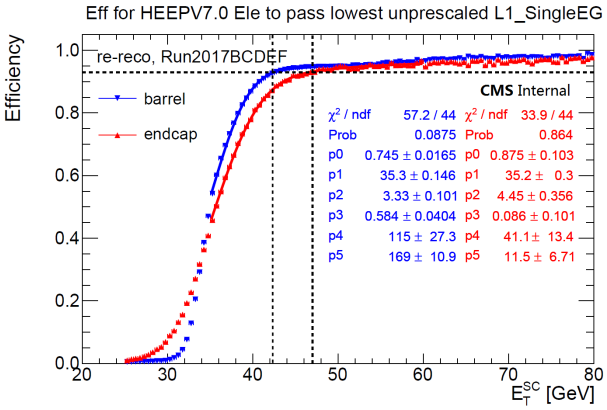

Figure 5.3: The efficiency for an electron passing HEEP to pass the lowest unprescaled L1 SingleEG seed versus supercluster E_T and η in barrel and endcap for 2017 [111].

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 E_T greater than the lowest unprescaled L1 SingleEG seed E_T threshold. Therefore, only L1 SingleEG seed turn on shown in Figure 5.3 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 is shown in Equation 5.3 where P_1 and P_2 are the L1 SingleEG efficiencies (shown in Figure 5.3) for leading and sub-leading selected HEEP electrons.

$$weight(L1) = 1 - (1 - P_1) \cdot (1 - P_2) = P_1 + P_2 - P_1 P_2 \tag{5.3}$$

5.2.3 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 5.4 and Figure 5.5, 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 5.6 and Figure 5.7, 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.

5.2.4 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 5.8. Similar for 2017 the HLT_Ele35_WPTight trigger path is used and the efficiency of this trigger is shown in Figure 5.9. These curves are used to weight MC events to simulate the effect of the trigger requirement in data.

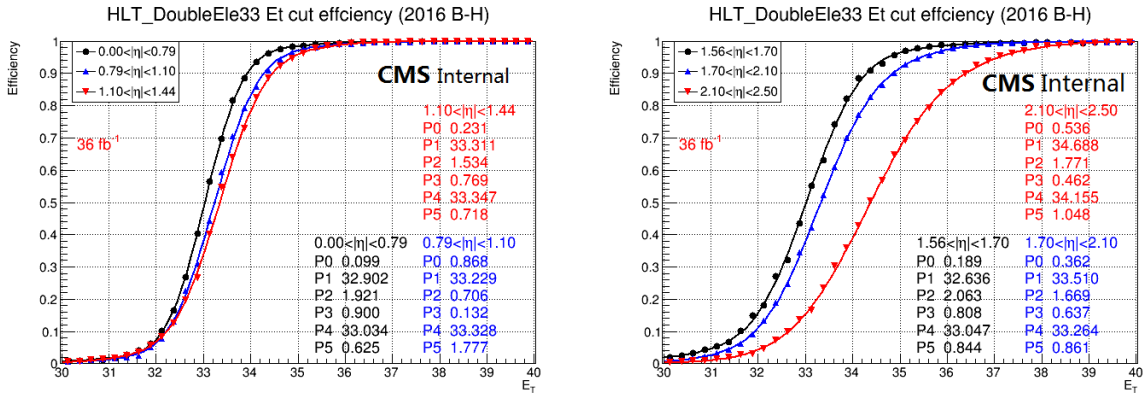

Figure 5.4: The efficiency for electron in the barrel (left) and endcap (right) passing HEEP to pass an online supercluster $E_T > 33$ GeV cut for 2016 [110].

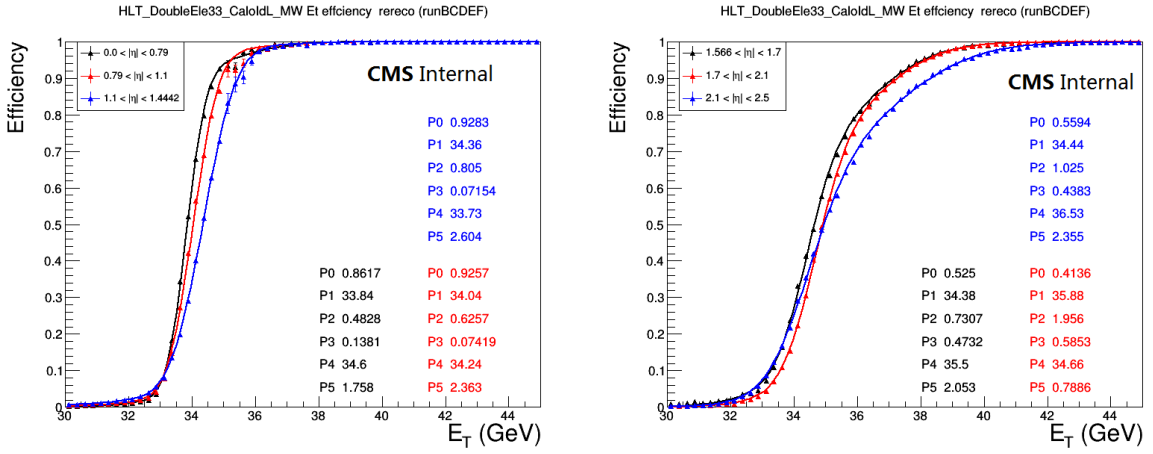

Figure 5.5: The efficiency for electron in the barrel (left) and endcap (right) passing HEEP to pass an online supercluster $E_T > 33$ GeV cut for 2017 [111].

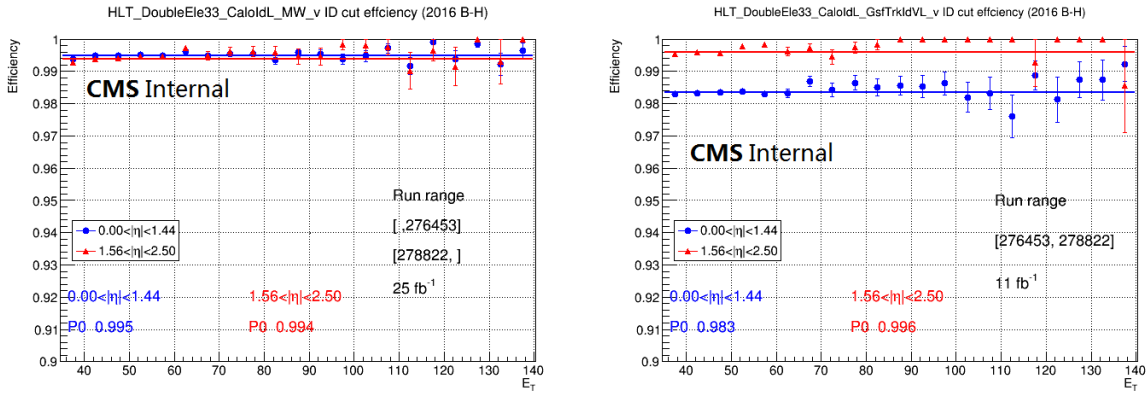

Figure 5.6: The efficiency for electron in the barrel and endcaps passing HEEP to pass the CaloIdL+MW ID requirement(left) and CaloIdL+GsfTrkIdVL ID requirement (right) for 2016 [110].

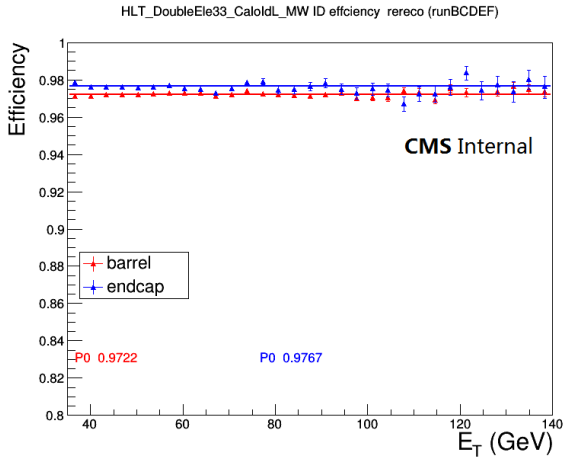

Figure 5.7: The efficiency for electron in the barrel and endcaps passing HEEP to pass the CaloIdL+MW ID requirement for 2017 [111].

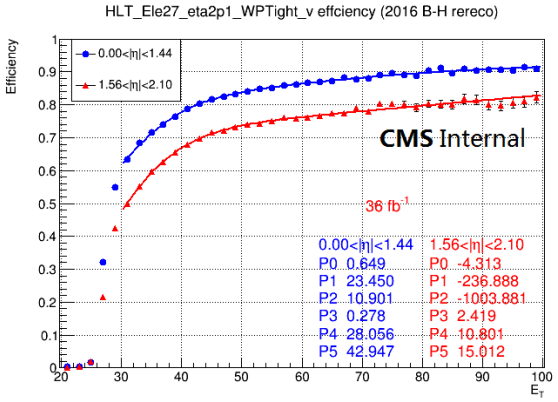

Figure 5.8: The efficiency for an electron passing HEEP to pass the HLT_Ele27_eta2p1_WPTight trigger for 2016 [110].

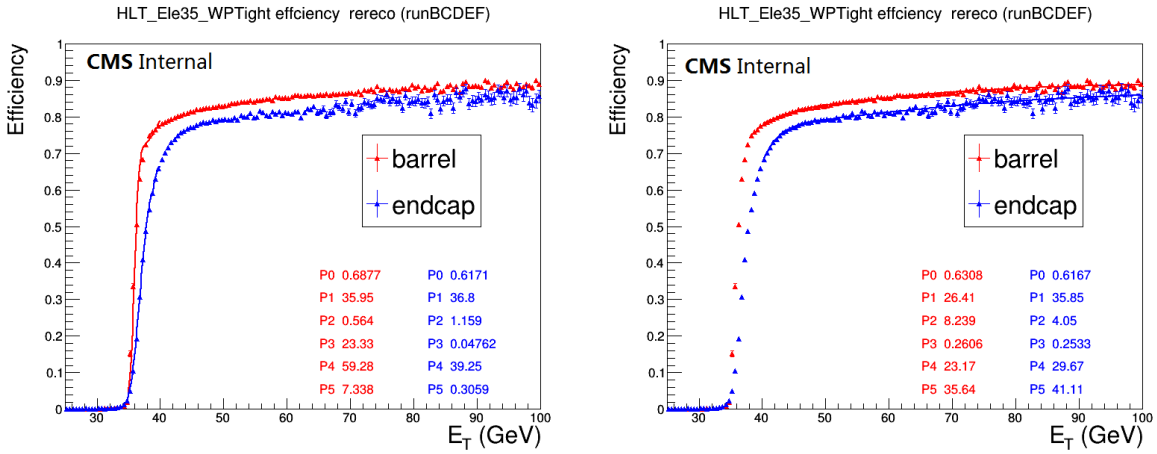

Figure 5.9: The efficiency for an electron passing HEEP to pass the HLT_Ele35_WPTight for E_T less than 40 GeV (left) and E_T more than 40 GeV (right) for 2017 [111].

5.3 Object and Event Selection

Electron candidates are required to pass the so called "HEEP" (High Energy Electron Pairs) selection which is listed in Table 5.3 followed by variables definition. In order to obtain well reconstructed electron candidates in tracker and ECAL sensitive regions, the candidates in the ECAL transition region ($1.4442 < |\eta| < 1.566$) and beyond the η coverage ($|\eta| > 2.5$) of the tracker are therefore discarded. After passing the HEEP selection the electrons 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. In addition, no charge requirement for dielectron candidates is asked 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 5.2. MC events are weighted using turn on curves shown in Figures 5.3, 5.4 and 5.5 for considering the L1 and HLT effects.

Variable	Barrel	Endcap
	Acceptance selections	
E_T	$E_T > 35$ GeV	$E_T > 35$ 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	$H/E < 1/E + 0.05$	$H/E < 5/E + 0.05$
$\sigma_{i\eta i\eta}$	-	$\sigma_{i\eta i\eta} < 0.03$
$\frac{E_{1\times5}}{E_{5\times5}}$ and $\frac{E_{2\times5}}{E_{5\times5}}$	$\frac{E_{1\times5}}{E_{5\times5}} > 0.83$ or $\frac{E_{2\times5}}{E_{5\times5}} > 0.94$	-
Inner layer lost hits	lost hits ≤ 1	lost hits ≤ 1
Impact parameter d_{xy}	$\|d_{xy}\| < 0.02$ cm	$\|d_{xy}\| < 0.05$ cm
isEcalDriven	true	true
	Isolation selections	
Calorimeter isolation Iso	$Iso < 2 + 0.03E_T[\text{GeV}] + 0.28\rho$	$Iso < 2.5 + 0.28\rho$ (E_T <50 GeV) else $Iso < 2.5 + 0.03(E_T[\text{GeV}] - 50) + 0.28\rho$
Track p_T isolation $Isopt$	$Isopt < 5$ GeV	$Isopt < 5$ GeV

Table 5.3: The definitions of HEEP selection cuts [112].

The variables used in the HEEP definition are defined as follows [112]:

- $\Delta\eta_{in}^{seed}$: it is the difference in η between the track position as measured in the inner layers, extrapolated to the interaction vertex and then extrapolated to the calorimeter and the η of the supercluster's seed. The cut value in the barrel (0.004) is tighter than in the endcaps (0.006), because the tracker material budget is thicker in the endcaps and it reduces the precision on the position measurement.
- $\Delta\phi_{in}$: it is the difference in ϕ between the track position as measured in the inner layers, extrapolated to the interaction vertex and then extrapolated to the calorimeter and the ϕ of the supercluster. Since there is a wider spread of the energy in ϕ with respect to η due to photon emissions from bremsstrahlung processes of a electron, the distribution of $\Delta\phi_{in}$ is much broader than $\Delta\eta_{in}^{seed}$. Hence, The cut value of $\Delta\phi_{in}$ (0.06 in both barrel and endcap regions) is around ten times looser than for $\Delta\eta_{in}^{seed}$.
- H/E: it is a ratio of hadronic energy of the HCAL tower in a cone of radius 0.15 centred on the electron's position in the calorimeter to the electromagnetic energy of the electron's supercluster.
- $\sigma_{i\eta i\eta}$: it is a measure of the spread in η in units of crystals of the electrons energy in

the 5×5 block centred on the seed crystal. It is computed as:

$$\sigma_{i\eta i\eta} = \sqrt{\frac{\sum_{i\in 5\times 5}(\eta_i - \bar{\eta})^2 w_i}{\sum_{i\in 5\times 5} w_i}}, \qquad w_i = \max\left(0, 4.7 + \log(E_i/E_{5\times 5})\right).$$

- $\frac{E_{1\times 5}}{E_{5\times 5}}$: it is the ratio of the energy contained in the 1×5 domino in $\eta \times \phi$ in the barrel ($x \times y$ in the endcaps) centered in ϕ on the seed crystal of the supercluster over the energy of the 5×5 matrix centered on the same seed crystal.
- $\frac{E_{2\times 5}}{E_{5\times 5}}$: it is the ratio of the energy contained in the most energetic 2×5 domino in $\eta \times \phi$ in the barrel ($x \times y$ in the endcaps) centered in ϕ on the seed crystal of the supercluster over the energy of the 5×5 matrix centered on the same seed crystal.
- Inner layer lost hits : it is defined as the number of missing hits in the innermost layers of the tracker (including the pixel) before the gsf track first hit. It is mainly designed to reject photons that convert into a pair of electrons in the tracker.
- Impact parameter d_{xy}: it is the closest distance (in the transverse plane) between the primary vertex and the track of the gsf electron candidate. The distribution is wider in the endcaps due to the poorer resolution on the track momentum in that region. Similarly to the missing hit cut, the d_{xy} cut is mainly useful to reject converted photons.
- isEcalDriven : electrons can be ecal driven (found using egamma techniques) or tracker driven (found using particle flow algorithm). Tracker driven is useful for low energy electrons, while it is not useful or validated for high energy electrons. Hence we require that the electron be ecal driven.
- ECAL isolation: it is defined as the scalar sum of the transverse energy of all the ECAL crystals with $E_T > 80$ MeV in the barrel ($E_T > 100$ MeV in the endcaps) in a cone of $\Delta R = 0.3$ centered on the gsf electron candidate position in the calorimeter, excluding those in an inner cone of radius 3 crystals and those in a $|\eta|$ strip of total width of 3 crystals (see left side of Figure 5.10). The inner cone veto removes the electron energy from the sum whereas the $|\eta|$ strip removes the energy from the bremsstrahlung photons.
- HCAL isolation: it is defined as the sum of transverse energy collected by all the towers of the first layer of the HCAL in a cone of $\Delta R = 0.3$ centered on the gsf electron candidate position in the calorimeter, excluding the towers in a cone of $\Delta R = 0.15$ (see right side of Figure 5.10).

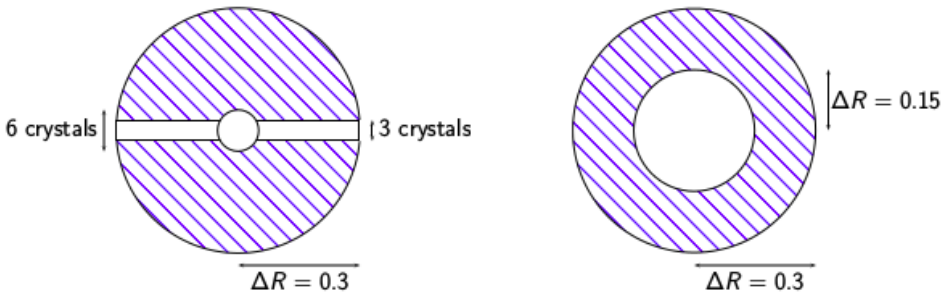

Figure 5.10: Definition of the regions used to compute the ECAL (left) and HCAL (right) isolation.

- Calorimeter isolation Iso: it is sum of the ECAL isolation and the HCAL isolation (defined in the previous points). The variable is strongly dependent on the electron energy and tends to increase for high E_T electrons due to the extension of the shower. A cut value of the form $Iso < a + b \cdot E_T$ is therefore applied.

- Track p_T isolation $Isopt$: it is defined as the sum of the p_T of all the reconstructed tracks between an inner and outer cone around the gsf electron. The ΔR for the inner cone is 0.04 and it is 0.3 for the outer cone. Besides, only tracks having $p_T >$ 700 MeV and $|\Delta z|$ with the gsf electron track < 0.2 cm are considered, where z is the minimum distance of the track to the nominal interaction point (0,0,0).

5.4 Mass Resolution and Scale

The mass resolution of heavy resonances 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 data and DY MC for the resolution and mean value of the invariant mass distribution of the electron pairs passing the HEEP ID at the Z peak (80 GeV $< M_{ee} <$ 100 GeV). The distributions are fitted using a Breit-Wigner (B-W) function (whose parameters are fixed to the PDG values 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 from the dCB function is then compared between data an MC in different η-categories. For BB category both electrons are required to be in the ECAL barrel. For BE category one electron is required to be in the ECAL barrel, and the other one in one of the ECAL endcaps. Fit results for data and MC in 2016 (2017) are shown in Figure 5.11 (5.12).

The σ_{MC} of the dCB from the fit to the MC distribution is subtracted in quadrature by the σ_{data} coming from the fit to the data, and defining the σ_{extra} through the relation: $\sigma_{extra} = \sqrt{\sigma^2_{data} - \sigma^2_{MC}}$. In Table 5.4 the results for the σ_{extra} are shown for the different categories for 2016 and 2017. Note that the numbers in Table 5.4 are presented in percentage [%] of the Z peak mass value (it is 91.1876 GeV for M_Z DPG). 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 value difference between data and MC in 2017 is much smaller than that in 2016 and for the σ_{extra} which is 0 in 2017 because of the larger σ_{MC} than σ_{data}.

Year	Category	$\frac{\Delta M}{M}$ [%]	σ_{data} [%]	σ_{MC} [%]	σ_{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 5.4: The results of scale shift $\frac{\Delta M}{M}$ between data and MC, as well as the σ_{extra} parameters per category.

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 5.13 to 5.15. 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 elections.

The second step of the study is based on MC only. 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 mass bin (from 200-5000 GeV) DY samples are used.

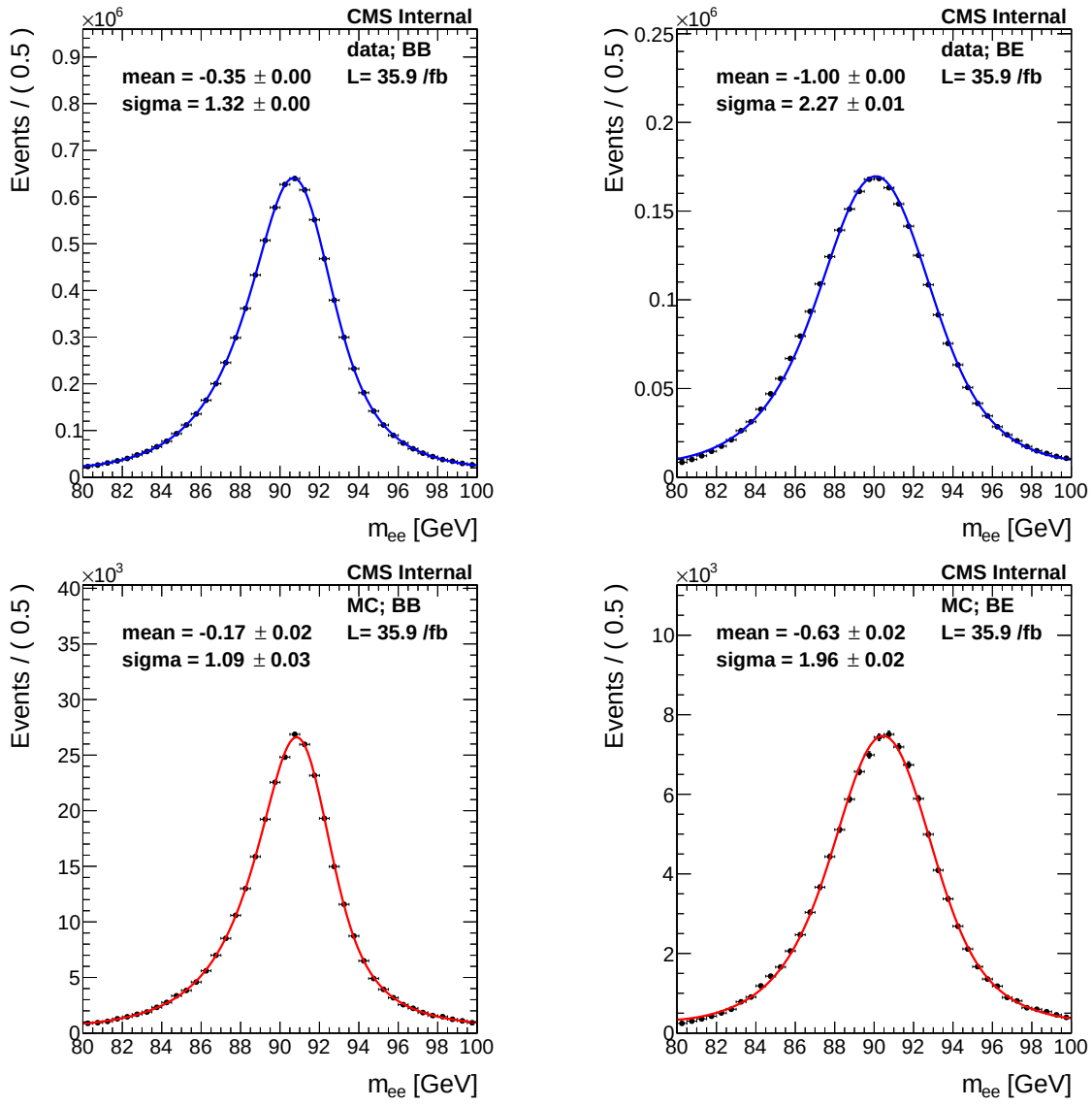

Figure 5.11: The distributions of dielectron invariant masses at the Z peak in data (top) and MC (bottom) for the BB (left) and BE (right) regions in 2016 [110].

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 analyzed. Defining a variable $resolution = \frac{m_{reco}-m_{gen}}{m_{gen}}$, its distribution is then acquired as a function of m_{gen}, and a maximum-likelihood fit to the distribution in each mass bin is performed using a "Cruijff function" (which is defined as a Gaussian core connected with an exponential tail on each side) in 2016. While in 2017, the dCB function is used for the fit because it is found that the Cruijff is not able to correctly describe the invariant mass shape at very high mass due to electron saturation effects[1]. Finally, the sigma parameter (σ_{fit}) of the fit function is added in quadrature with the σ_{extra} coming from the first step of the study.

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 B.1). Results of mass resolution for the

1. When the deposite energy in single ECAL crystal is very high, the electric readout will be saturated because of limited dynamic range. See more details in Appendix B.4.

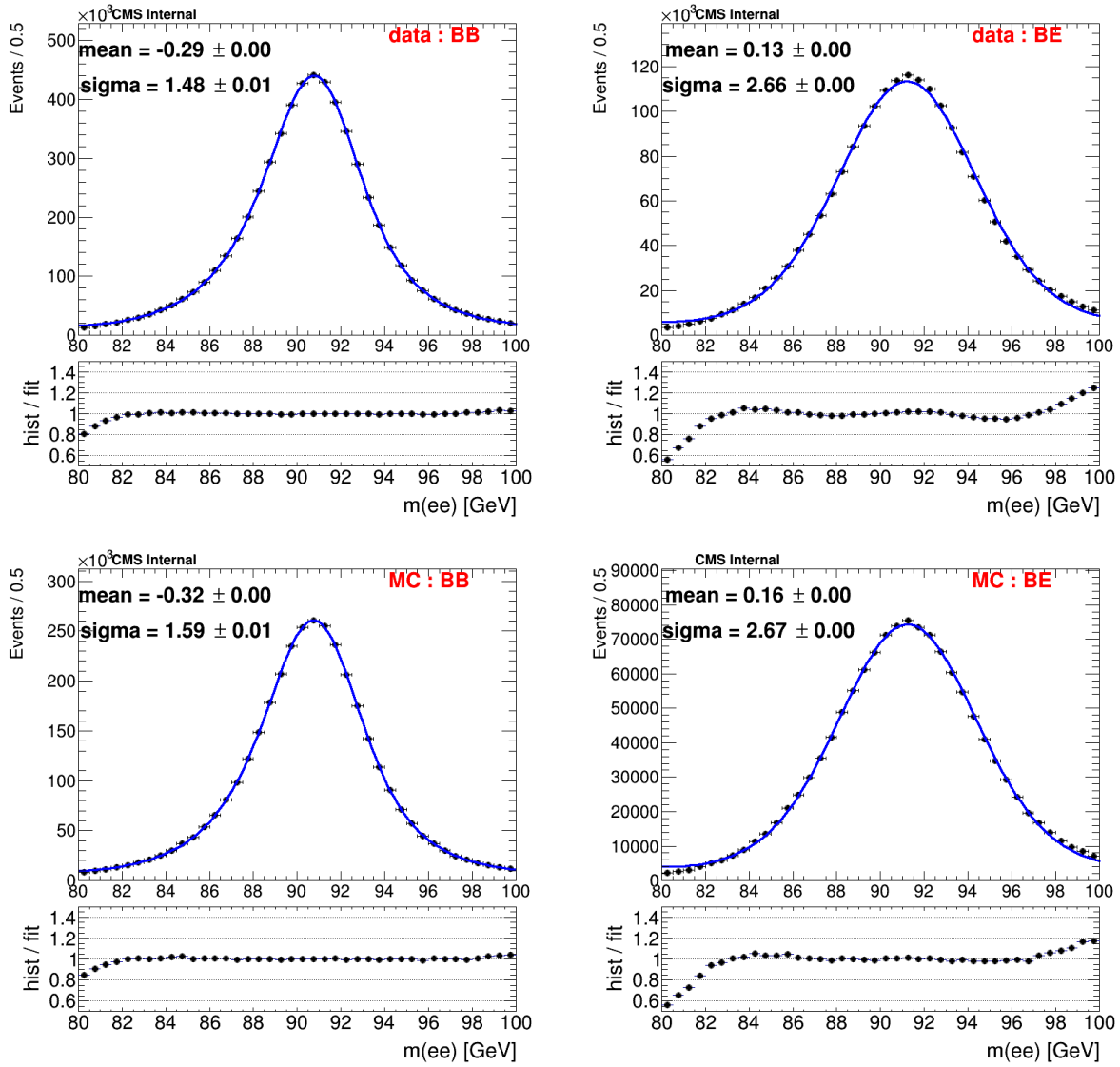

Figure 5.12: The distributions of dielectron invariant masses at the Z peak in data (top) and MC (bottom) for the BB (left) and BE (right) regions in 2017.

BB and BE regions are shown in Figure 5.16.

In the BB region there is a small linear rise in the mass resolution starting around $\approx$ 1.5 TeV. The effect has been already studied in [113] 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.

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.

Defining the mass scale variable as $scale = \frac{m_{reco}}{m_{gen}}$, where m_{gen} is the generated invariant mass, m_{reco} is the reconstructed invariant mass. From the *resolution* variable defined

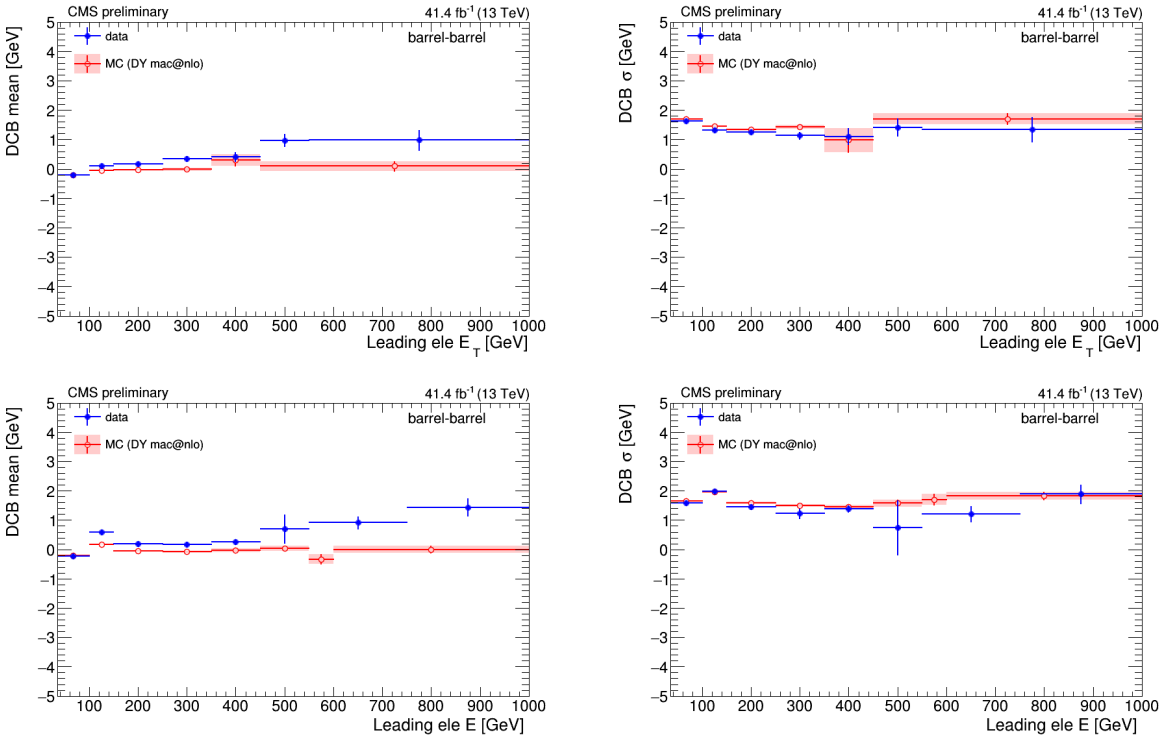

Figure 5.13: The mean (left) and sigma (right) values of dCB as the function of E_T (top) and energy (bottom) of leading electron for barrel-barrel channel.

above it is easy to see that $scale = \frac{m_{reco}}{m_{gen}} = 1 + resolution$. Therefore, the mass scale is taken from the mean parameter of the function used to fit the *resolution* distribution and simply adding the unity to it. Besides, the error on the mean parameter is taken accordingly.

Results of mass scale for the BB and BE regions are shown in Figure 5.17.

In the BB region there is a drop in the mass scale parameter starting around ≈ 1 TeV (≈ 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 electromagnetic shower in the HCAL subdetector.

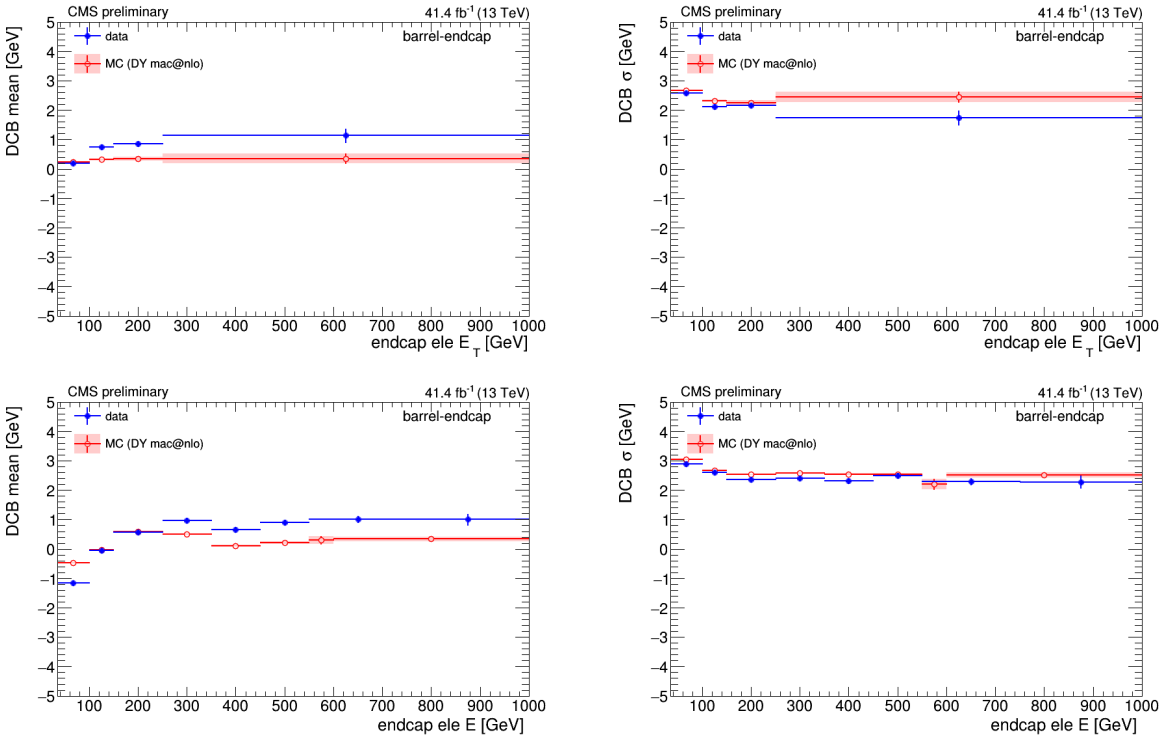

Figure 5.14: The mean (left) and sigma (right) values of dCB as the function of E_T (top) and energy (bottom) of endcap electron for barrel-endcap channel.

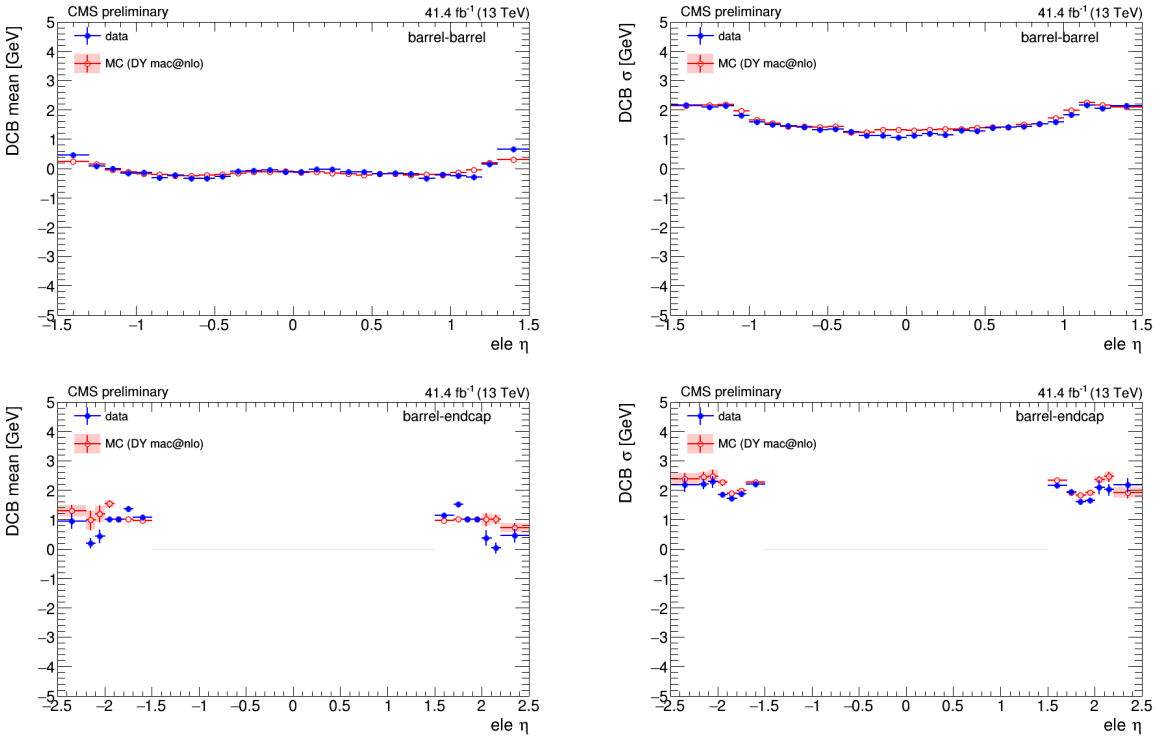

Figure 5.15: The mean (left) and sigma (right) values of dCB as the function of η of second electron (the first electron is asked to have $|\eta| < 0.5$) for barrel-barrel (top) and barrel-endcap (bottom) channels.

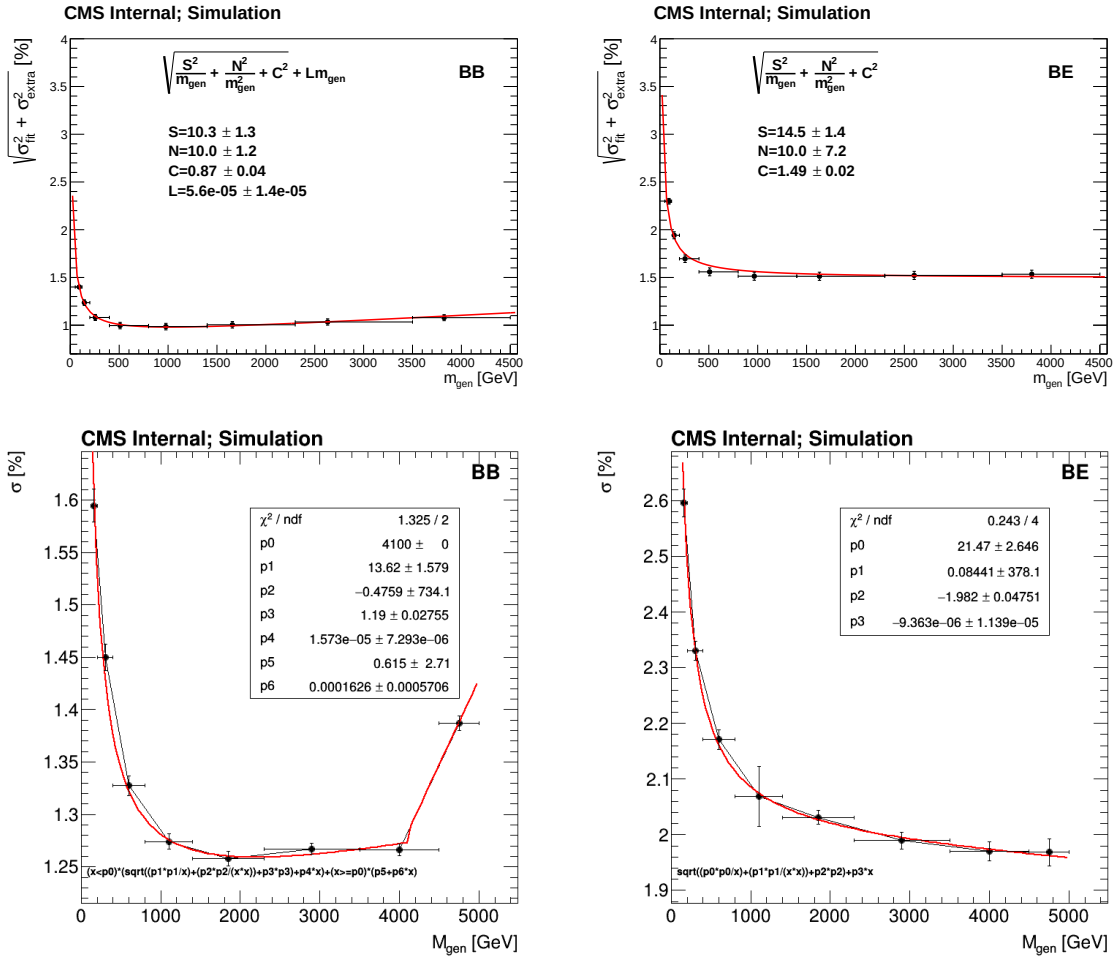

Figure 5.16: The mass resolutions (σ_{fit}) as a function of generated Z boson mass for BB (left) and BE (right) regions in 2016 (top) [110] and 2017 (bottom).

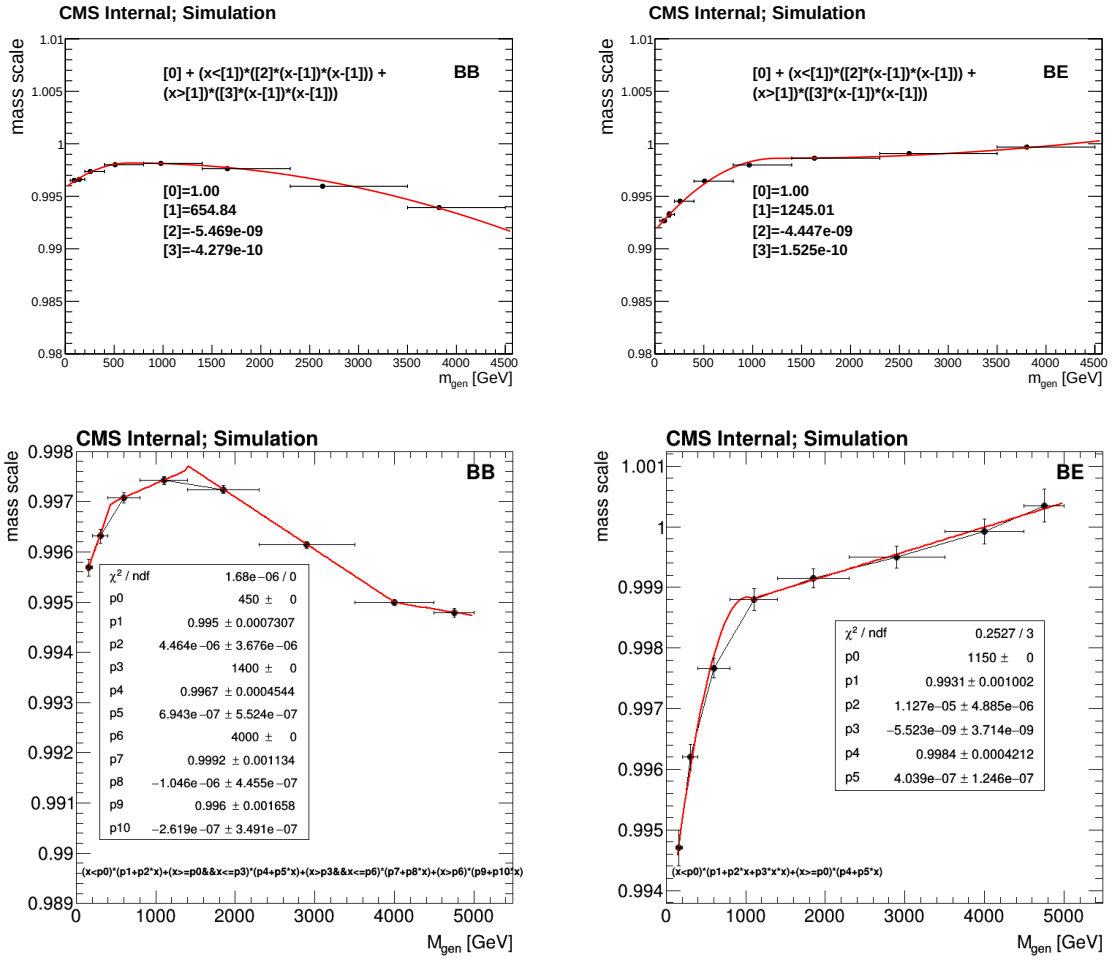

Figure 5.17: The mass scales as a function of generated Z boson mass for BB (left) and BE (right) regions in 2016 (top) [110] and 2017 (bottom).

5.5 HEEP ID Efficiency and Scale Factor

The MC samples used for the High Energy Electron Pair (HEEP) selection efficiency measurement and the scale factor measurement are listed in Table 5.5.

Sample	Xsection(pb)	Comments
DYToEE_NNPDF30_ 13TeV-powheg-pythia8	1921.8	for 2016 $Z \to ee$
DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	5765.4	for 2016 $Z \to \tau\tau$
DYJetsToLL_M-50_TuneCP5_13TeV-amcatnloFXFX-pythia8	5765.4	for 2017 $Z \to ee$, $Z \to \tau\tau$
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 5.5: MC samples used in the HEEP efficiency and scale factor studies

5.5.1 Tag and probe method

A so called "tag and probe" method is often used for calculating a certain selection efficiency. For the HEEP ID efficiency study, this method starts by searching for a good electron (the "tag") which satisfies certain types of criteria. Then the efficiency of the interested selection cut is measured on the second electron candidate called the "probe".

The 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. 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 be probes (e.g. once as "$e_1 = tag, e_2 = probe$" and once as "$e_1 = probe, e_2 = tag$") and by doing this we can get more statistics. What we get is a very clear peak around the mass of the Z boson. Therefore, we are confident that the electrons we have selected are almost real electrons coming from Z boson decay. Although there is a low contamination from other standard model backgrounds.

The efficiency is defined as

$$\epsilon_{\text{HEEP}} = \frac{N_{\text{passing probes}}}{N_{\text{all probes}}} \tag{5.4}$$

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

For measuring the 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 Table 5.5. MC events are weighted using the trigger turn on curve of `HLT_Ele27_eta2p1_WPTight` (`HLT_Ele35_WPTight`) path (see Figures 5.8 and 5.9 of section 5.2.4) 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 distributions 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, the number of probe and variable distributions of probe are similar for opposite sign and same sign tag and probe pairs from QCD process. On the other hand, all other standard model processes have opposite sign electron pairs and do not contribute to same sign control region. The other SM backgrounds are considered using MC simulations.

Because the HEEP ID efficiency is different in barrel and endcaps regions, the probes are separated in barrel and endcaps. In Figure 5.18, 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). Figures 5.19 to 5.22 present the E_T, η, ϕ distributions of the probes and the number of primary vertex for the selected tag and probe pairs event in 2016. The corresponding distributions for the 2017 data are presented in Figures 5.23 to 5.27.

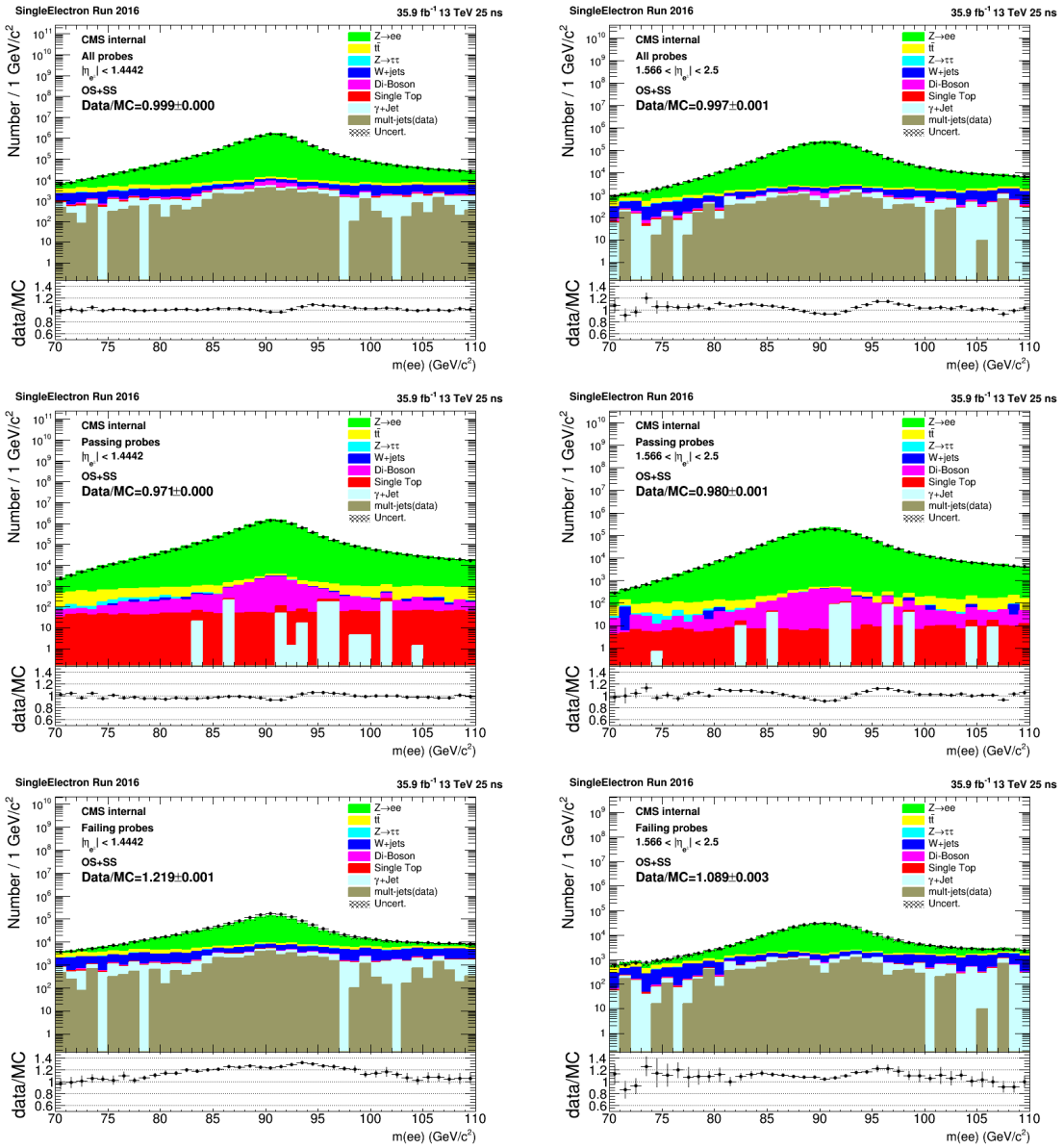

Figure 5.18: Invariant mass distributions of tag and probe for probe in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle), and only failed probes are included (bottom) for 2016.

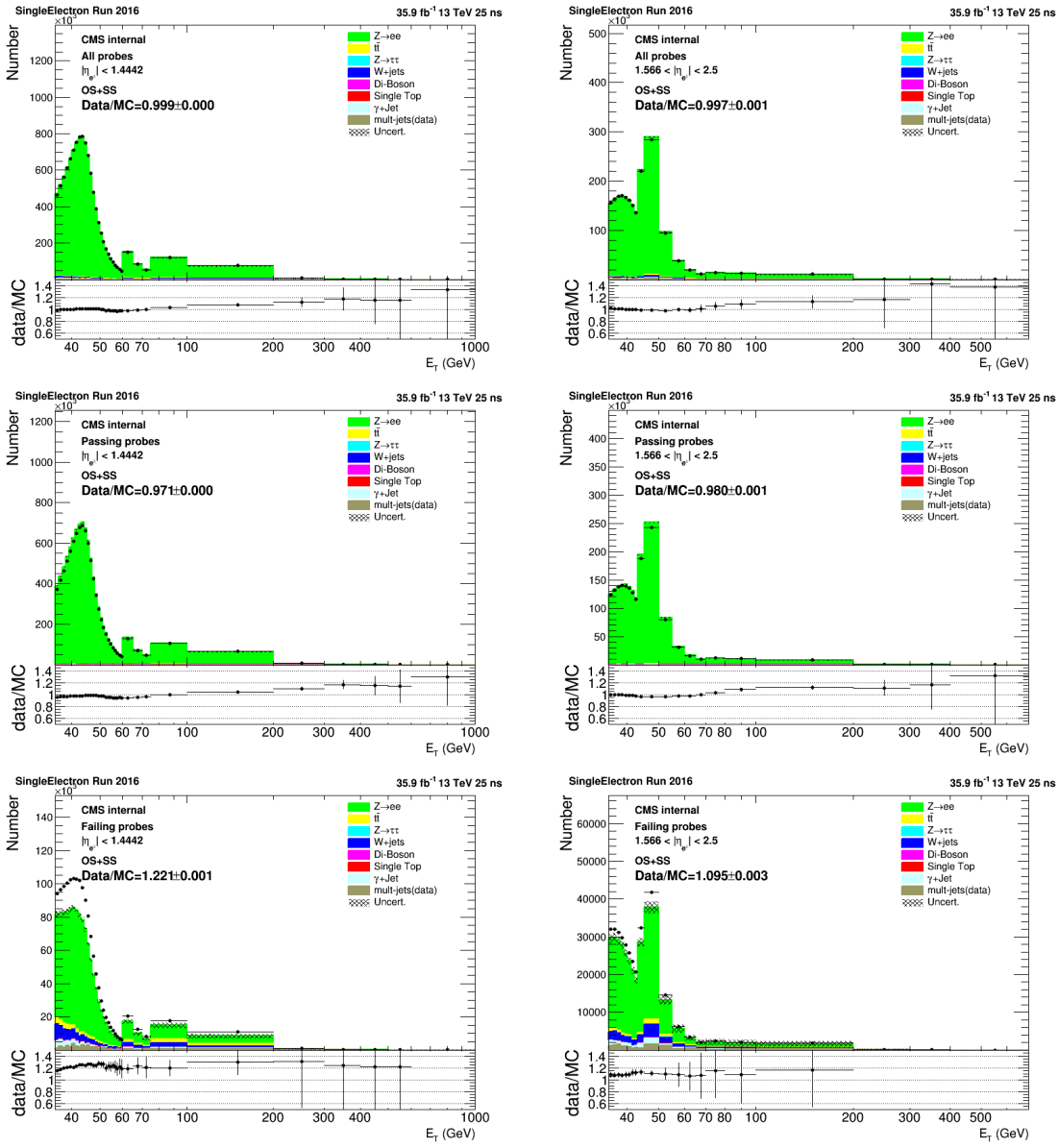

Figure 5.19: E_T of probe in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle), and only failed probes are included (bottom) for 2016.

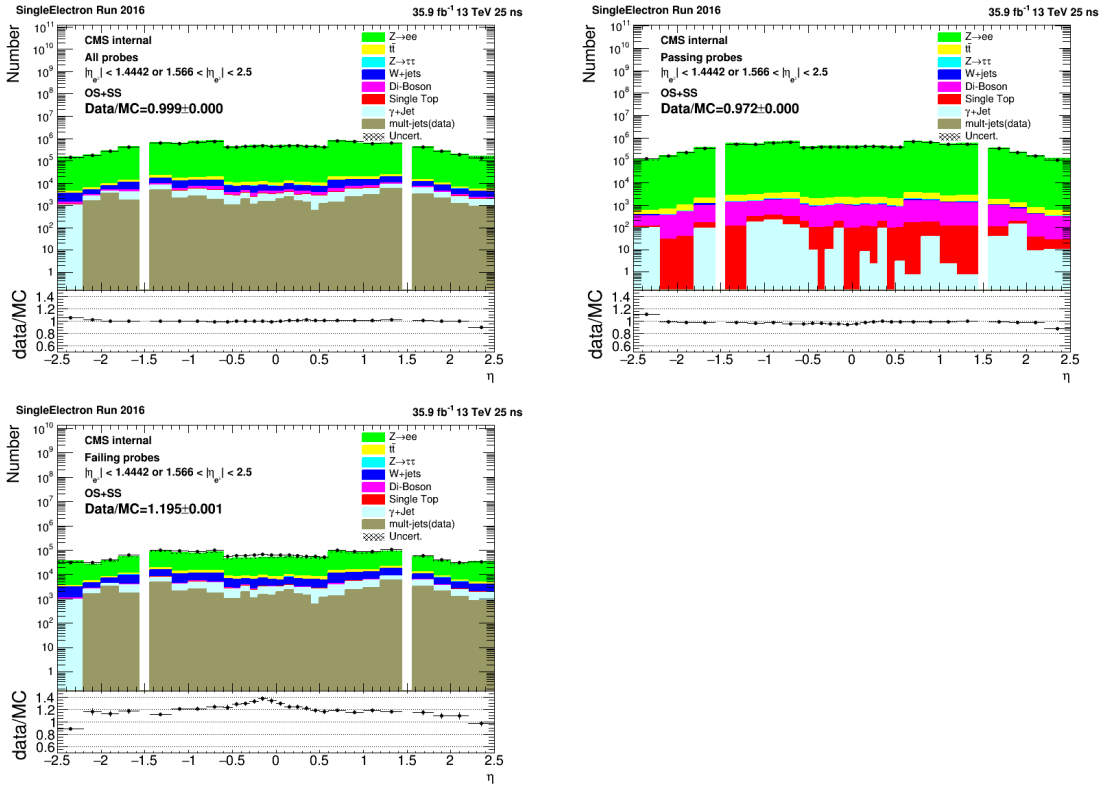

Figure 5.20: η of probe where all the probes are included (top left), only passing probes are included (top right), and only failed probes are included (bottom) for 2016.

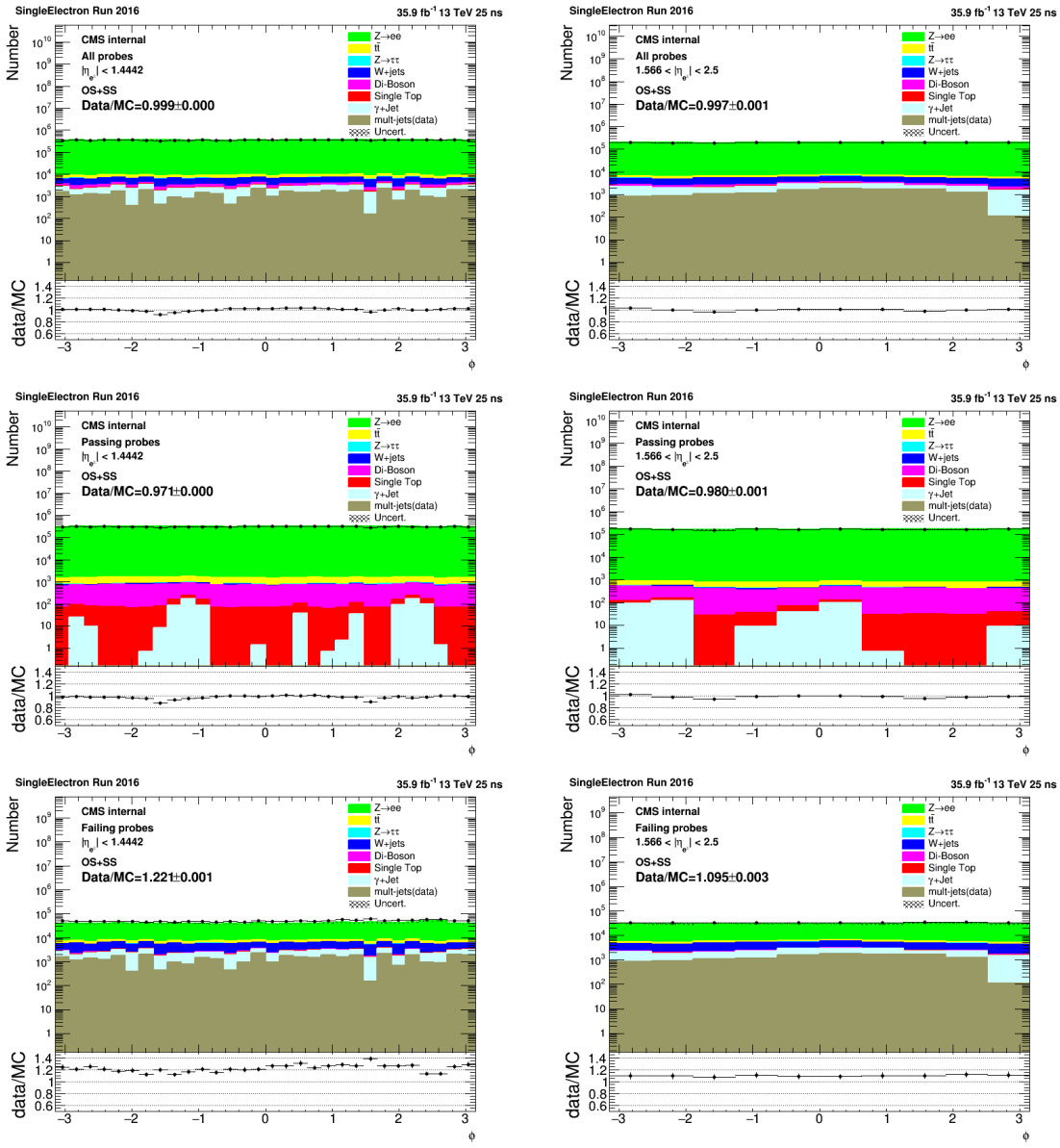

Figure 5.21: ϕ of probe in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle), and only failed probes are included (bottom) for 2016.

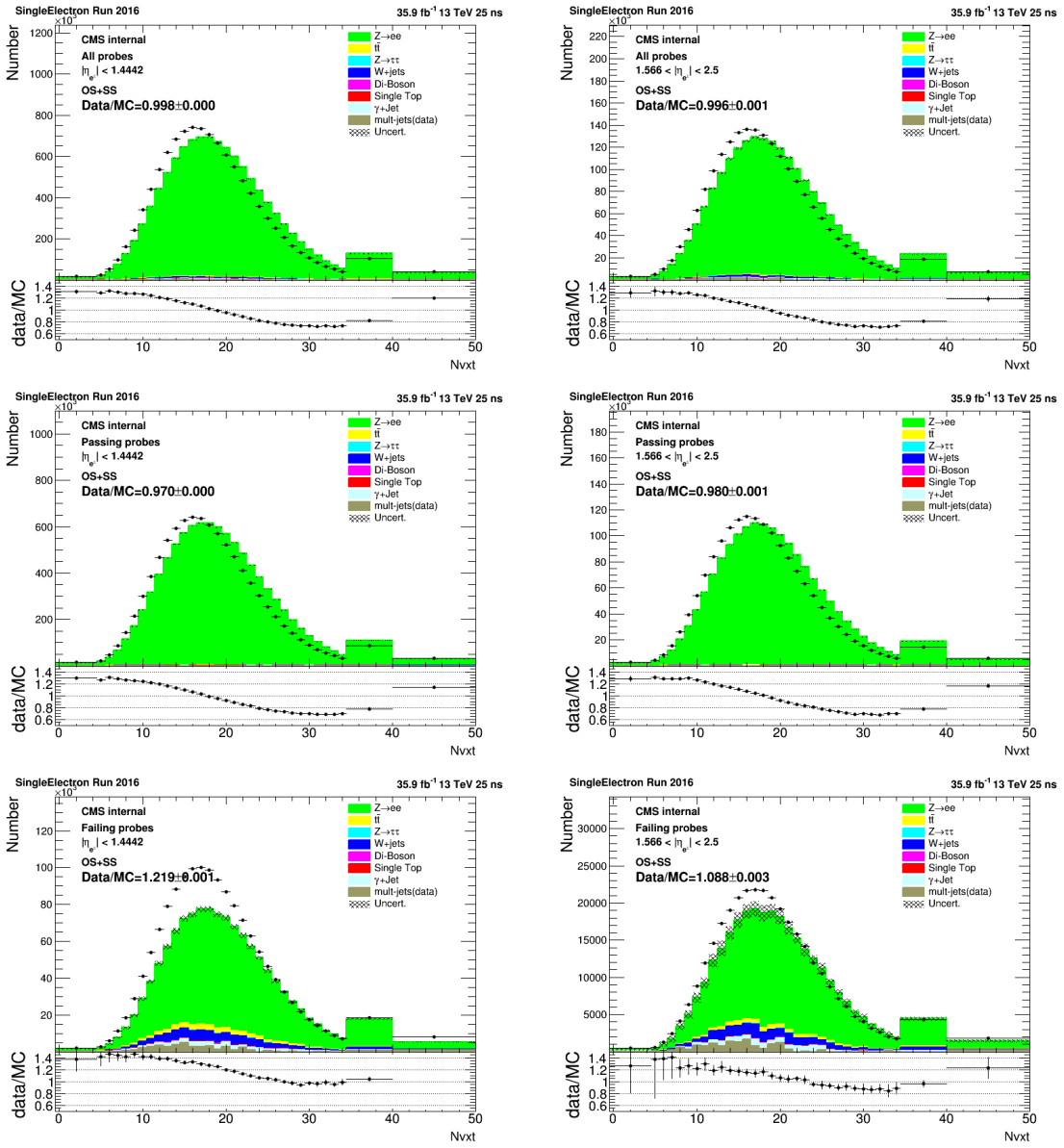

Figure 5.22: Number of primary vertex of tag and probe pair event in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle), and only failed probes are included (bottom) for 2016.

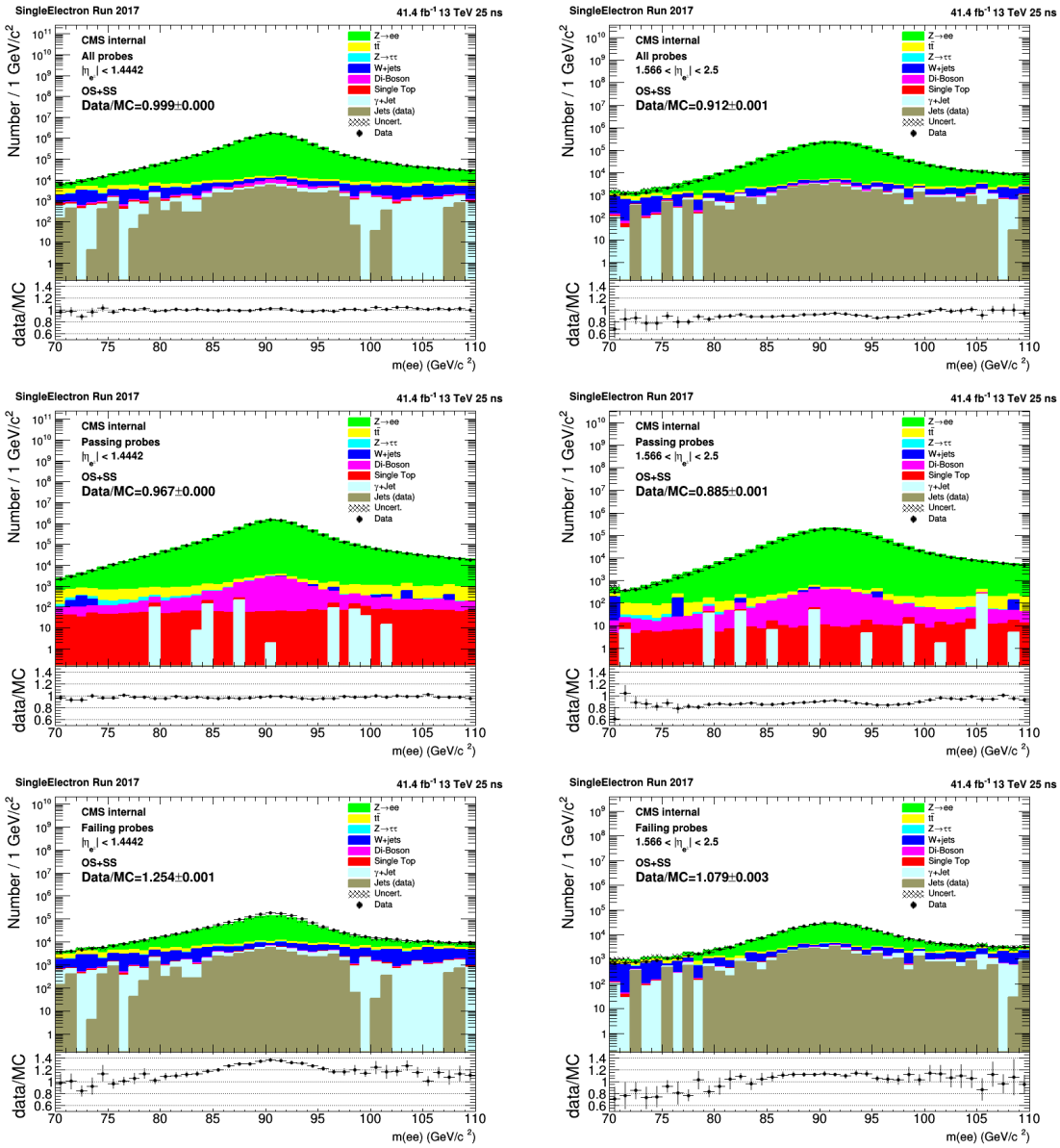

Figure 5.23: Invariant mass of tag and probe distributions for probe in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle), and only failed probes are included (bottom) for 2017.

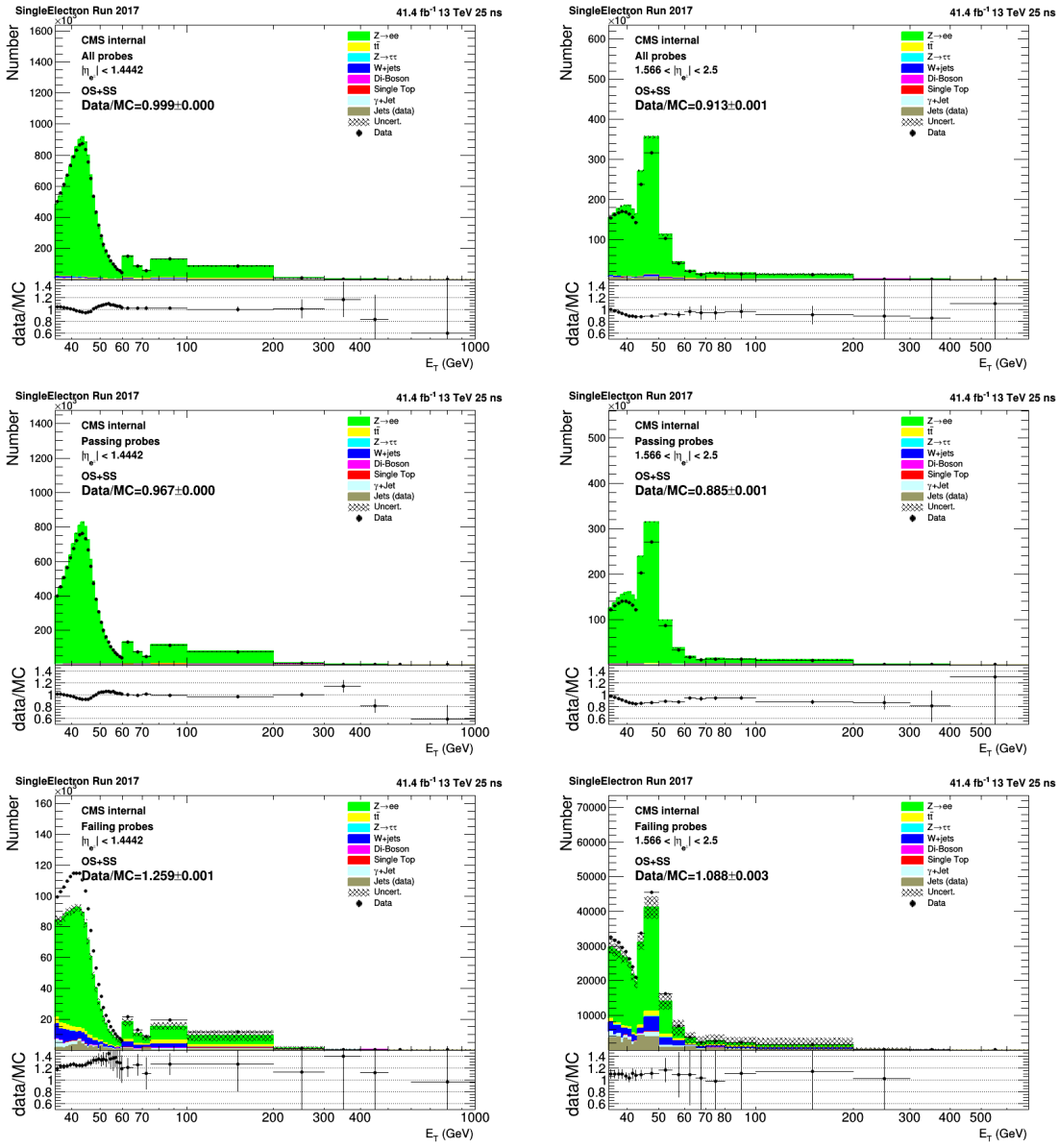

Figure 5.24: E_T of probe in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle), and only failed probes are included (bottom) for 2017.

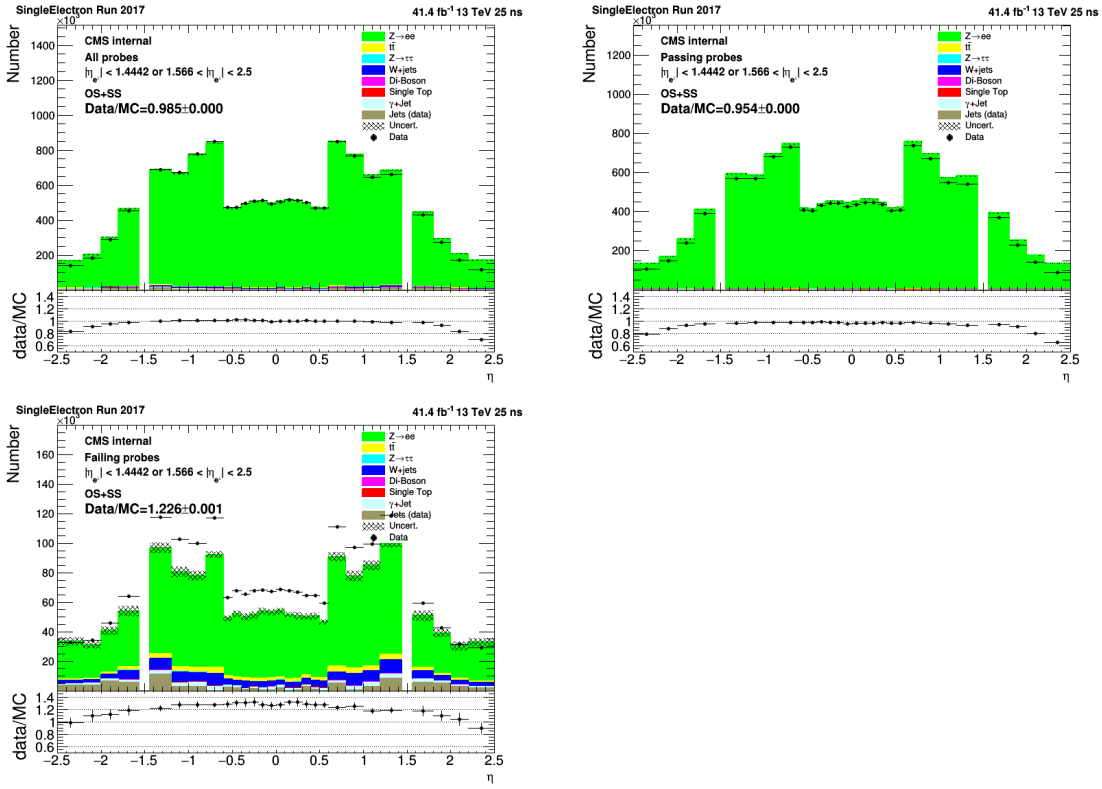

Figure 5.25: η of probe where all the probes are included (top left), only passing probes are included (top right), and only failed probes are included (bottom) for 2017.

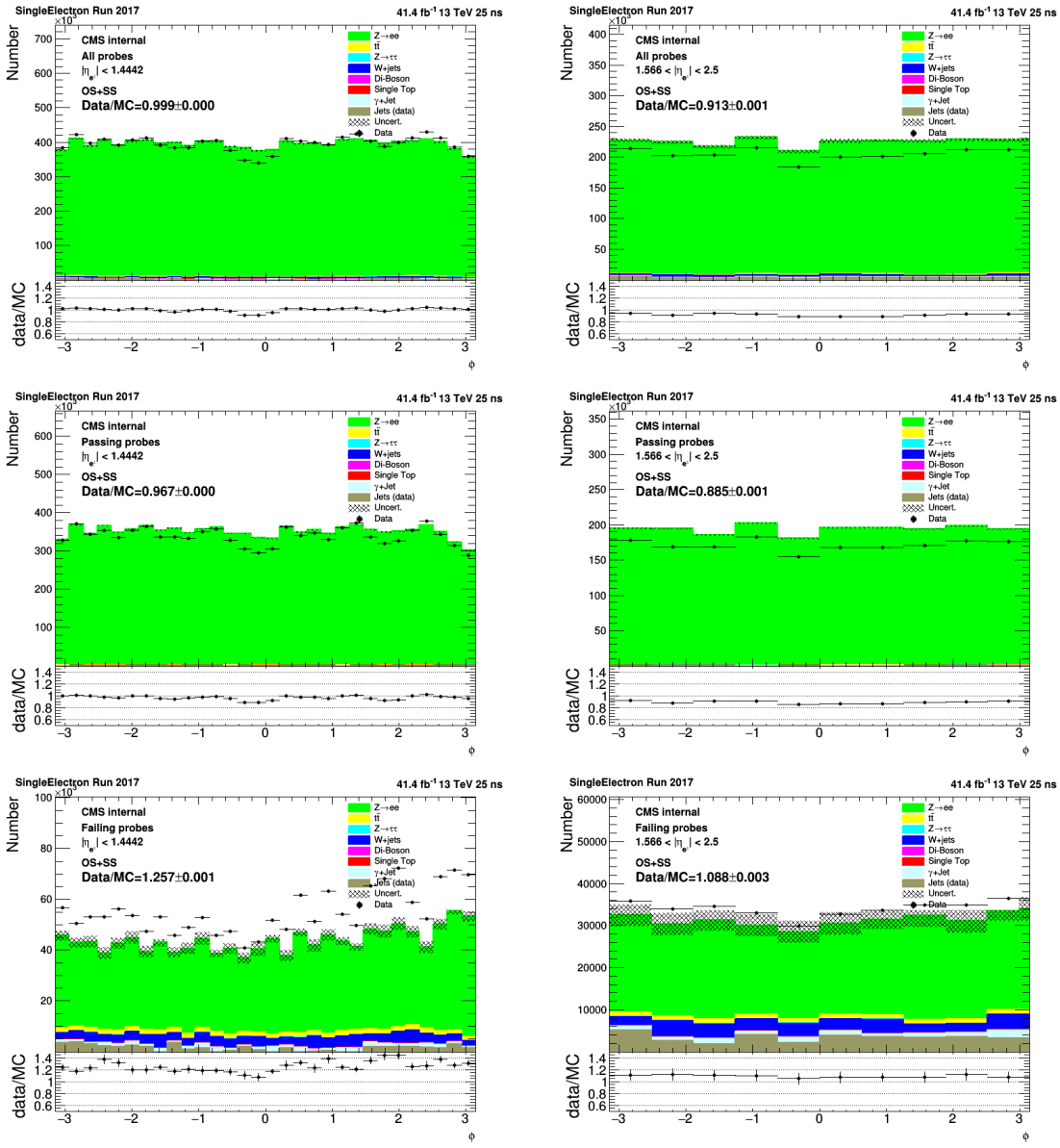

Figure 5.26: ϕ of probe in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle), and only failed probes are included (bottom) for 2017.

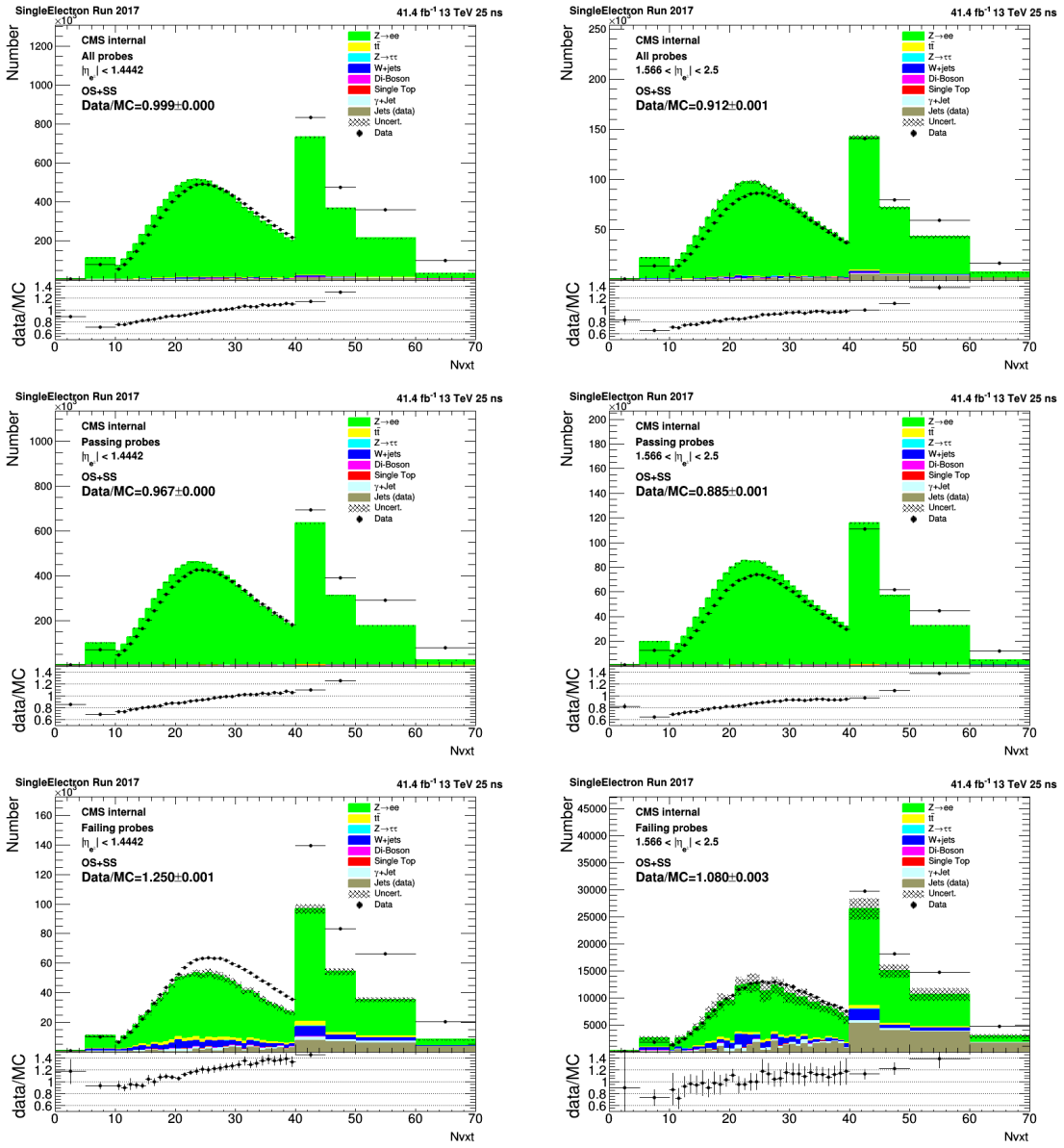

Figure 5.27: Number of primary vertex of tag and probe pair event in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle), and only failed probes are included (bottom) for 2017.

5.5.2 HEEP ID efficiencies and scale factors

In 2016, the HEEP ID efficiencies for data and MC are shown as a functions of E_T of the probe, η of the probe, ϕ of the probe, and N_{vtx} in Figures 5.28-5.31 (left), here data includes both DY and non-DY events. In order to check if the contributions of non-DY backgrounds are estimated correctly, the DY efficiency is compared to data from which non-DY contributions are subtracted in Figures 5.28-5.31 (right). The scale factors obtained with these two approaches are in a good agreement in all variables. In addition, the HEEP ID efficiency scale factors between data and MC are shown in the bottom pad of the same plot. Same plots in 2017 are shown in Figures 5.32-5.35.

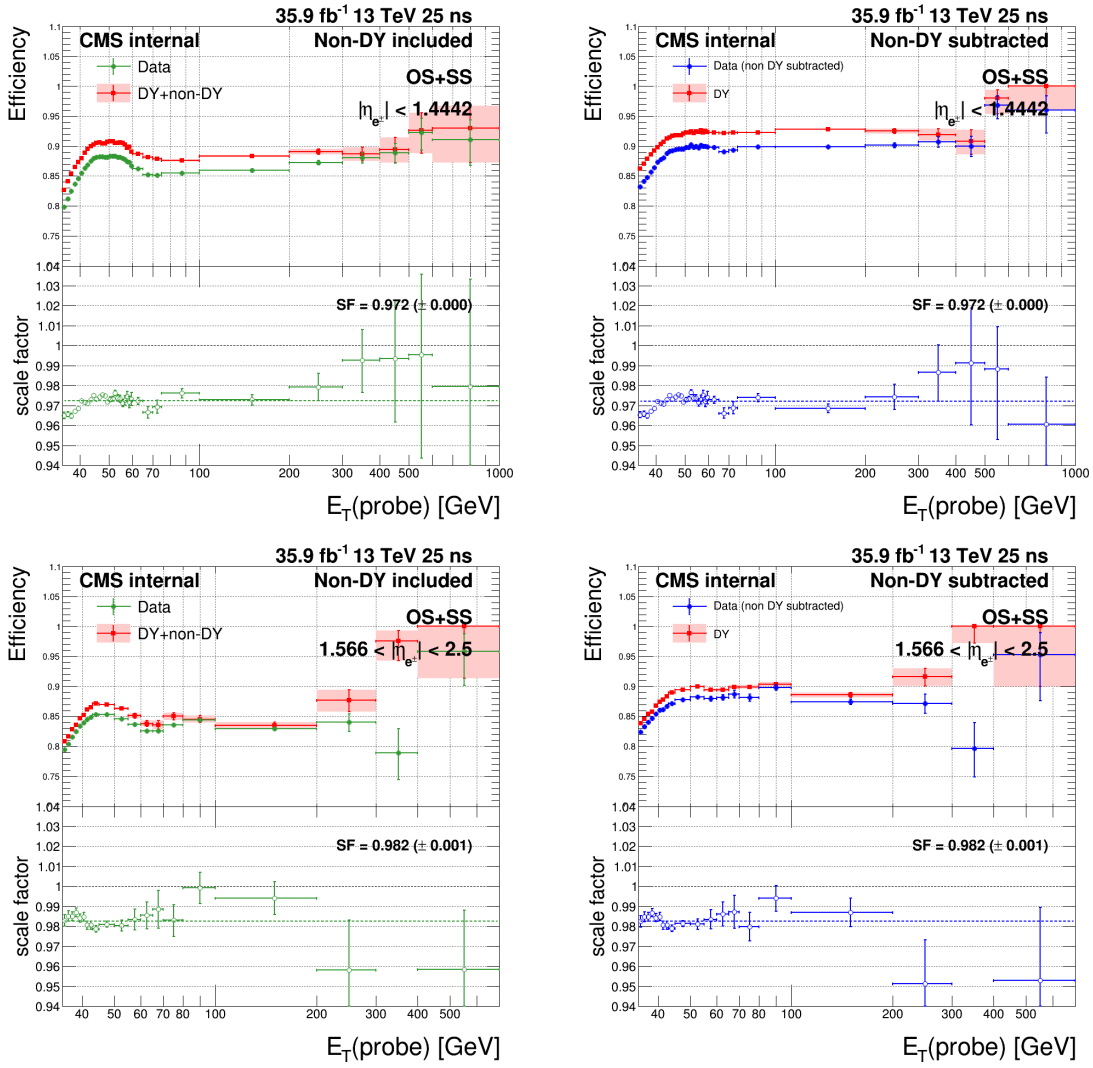

Figure 5.28: The HEEP efficiencies in data and MC, as well as the scale factors when the probe is in the barrel (top) and endcap (bottom) where the non-DY processes are included (left) and subtracted (right) as functions of probe E_T in 2016.

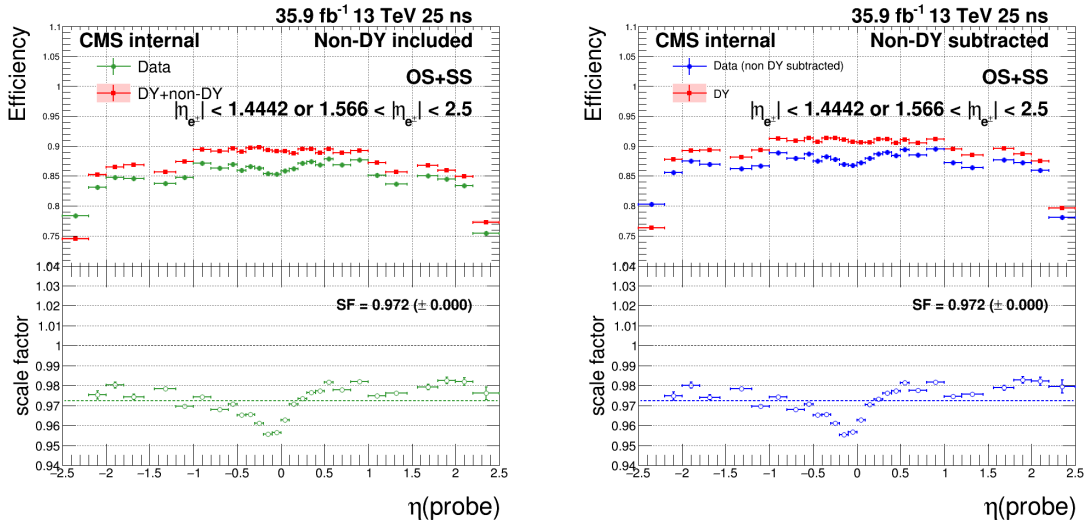

Figure 5.29: The HEEP efficiencies in data and MC, as well as the scale factors where the non-DY processes are included (left) and subtracted (right) as functions of probe η in 2016.

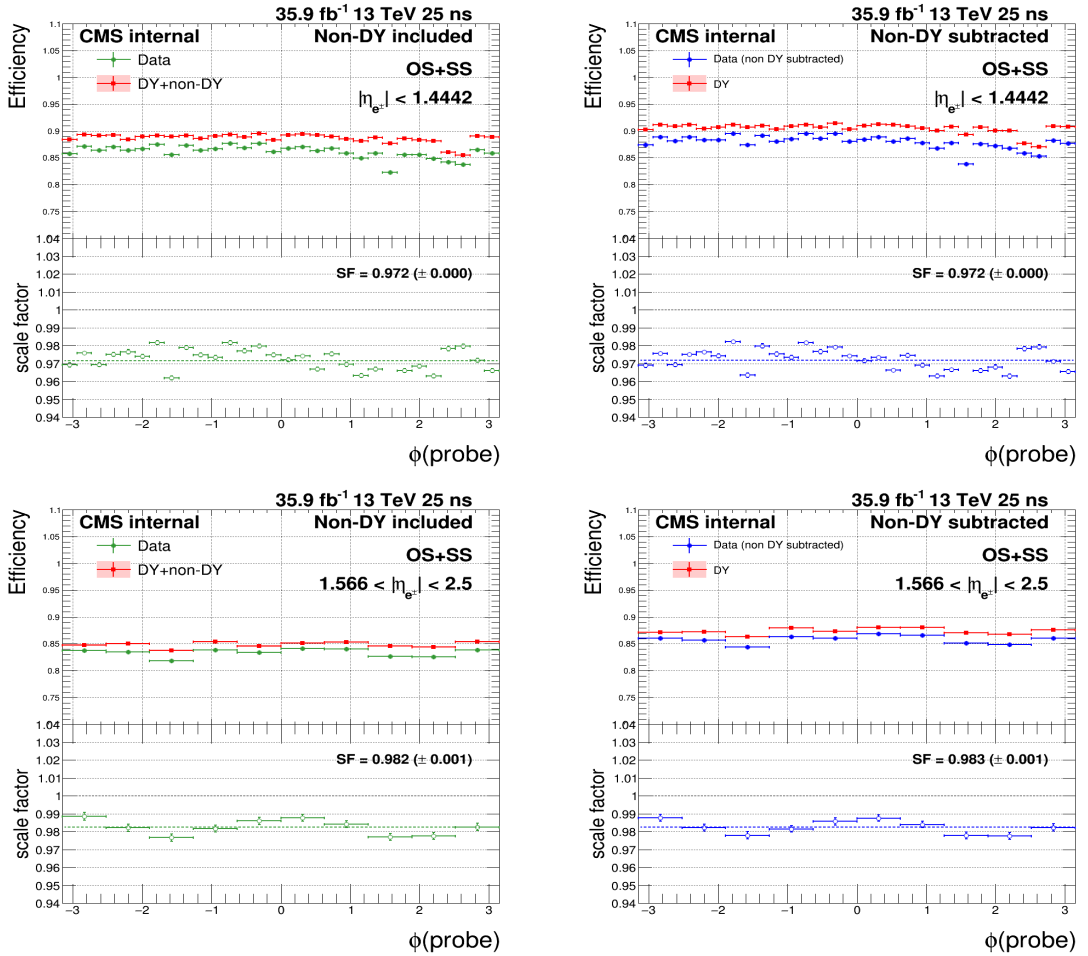

Figure 5.30: The HEEP efficiencies in data and MC, as well as the scale factors when the probe is in the barrel (top) and endcap (bottom) where the non-DY processes are included (left) and subtracted (right) as functions of probe ϕ in 2016.

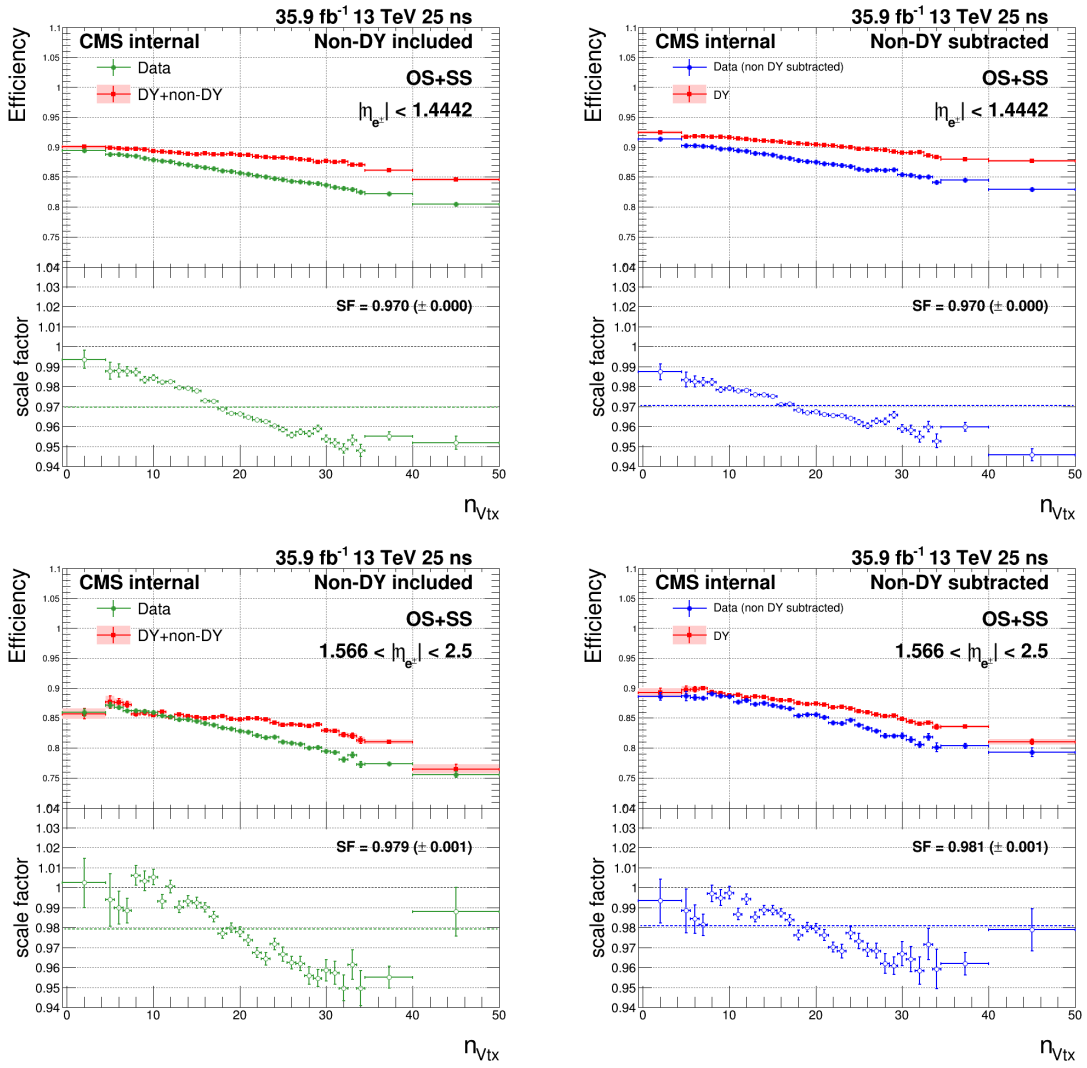

Figure 5.31: The HEEP efficiencies in data and MC, as well as the scale factors when the probe is in the barrel (top) and endcap (bottom) where the non-DY processes are included (left) and subtracted (right) as functions of the number of primary vertices, n_{Vtx} in 2016.

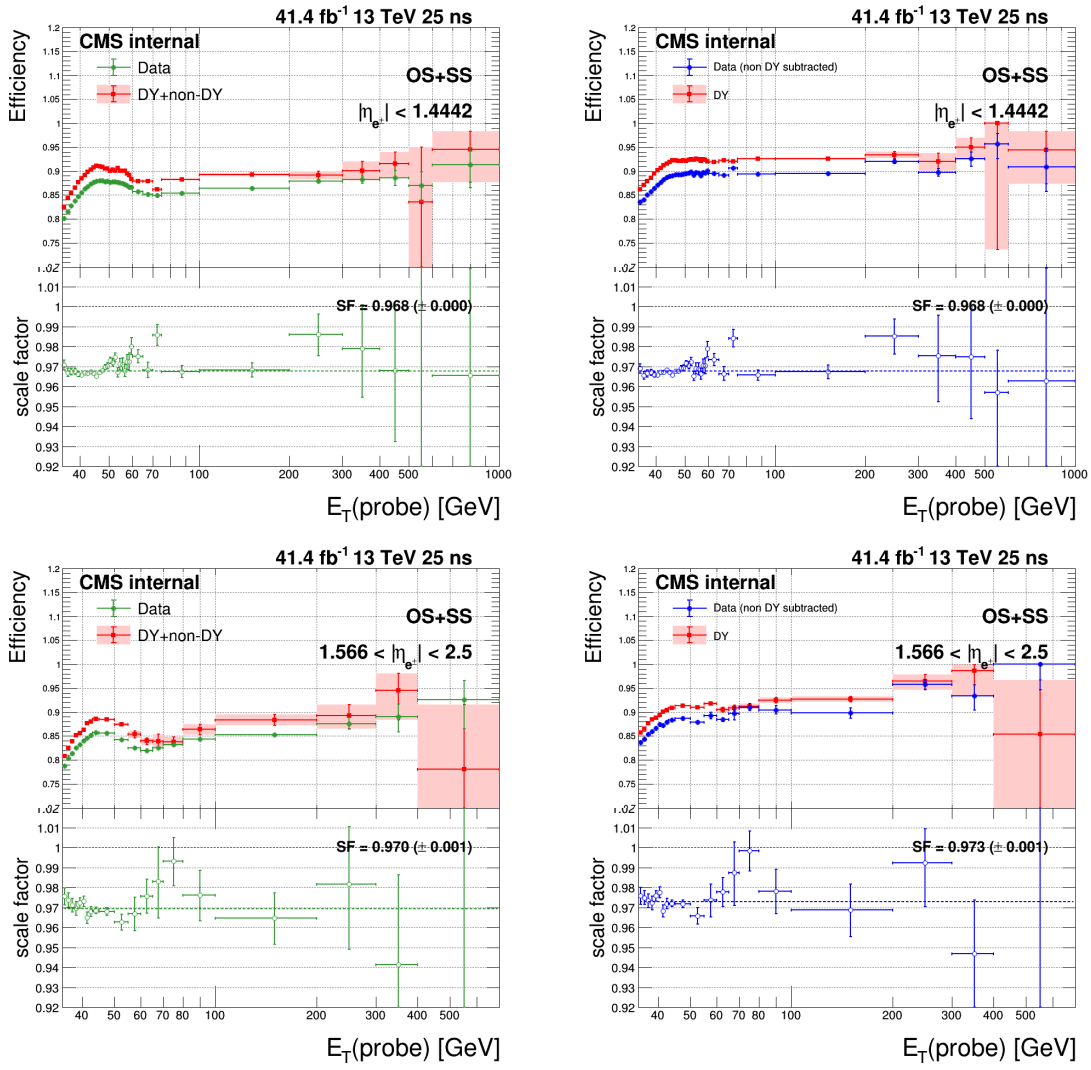

Figure 5.32: The HEEP efficiencies in data and MC, as well as the scale factors when the probe is in the barrel (top) and endcap (bottom) where the non-DY processes are included (left) and subtracted (right) as functions of probe E_T in 2017.

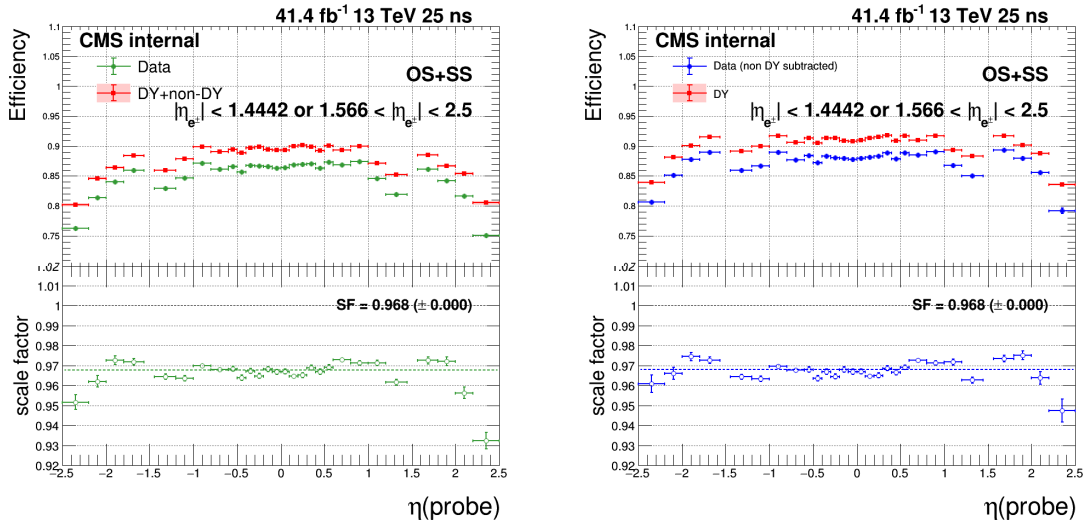

Figure 5.33: The HEEP efficiencies in data and MC, as well as the scale factors where the non-DY processes are included (left) and subtracted (right) as functions of probe η in 2017.

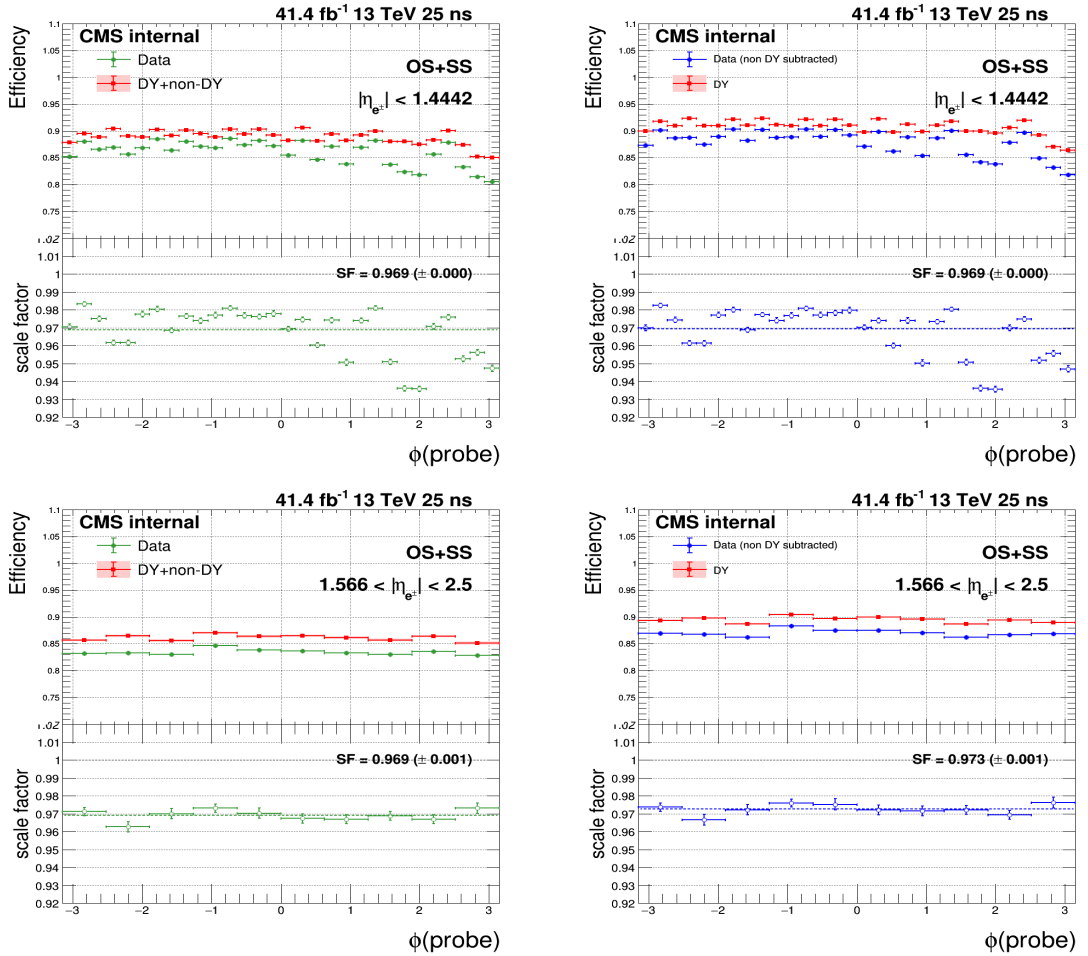

Figure 5.34: The HEEP efficiencies in data and MC, as well as the scale factors when the probe is in the barrel (top) and endcap (bottom) where the non-DY processes are included (left) and subtracted (right) as functions of probe ϕ in 2017.

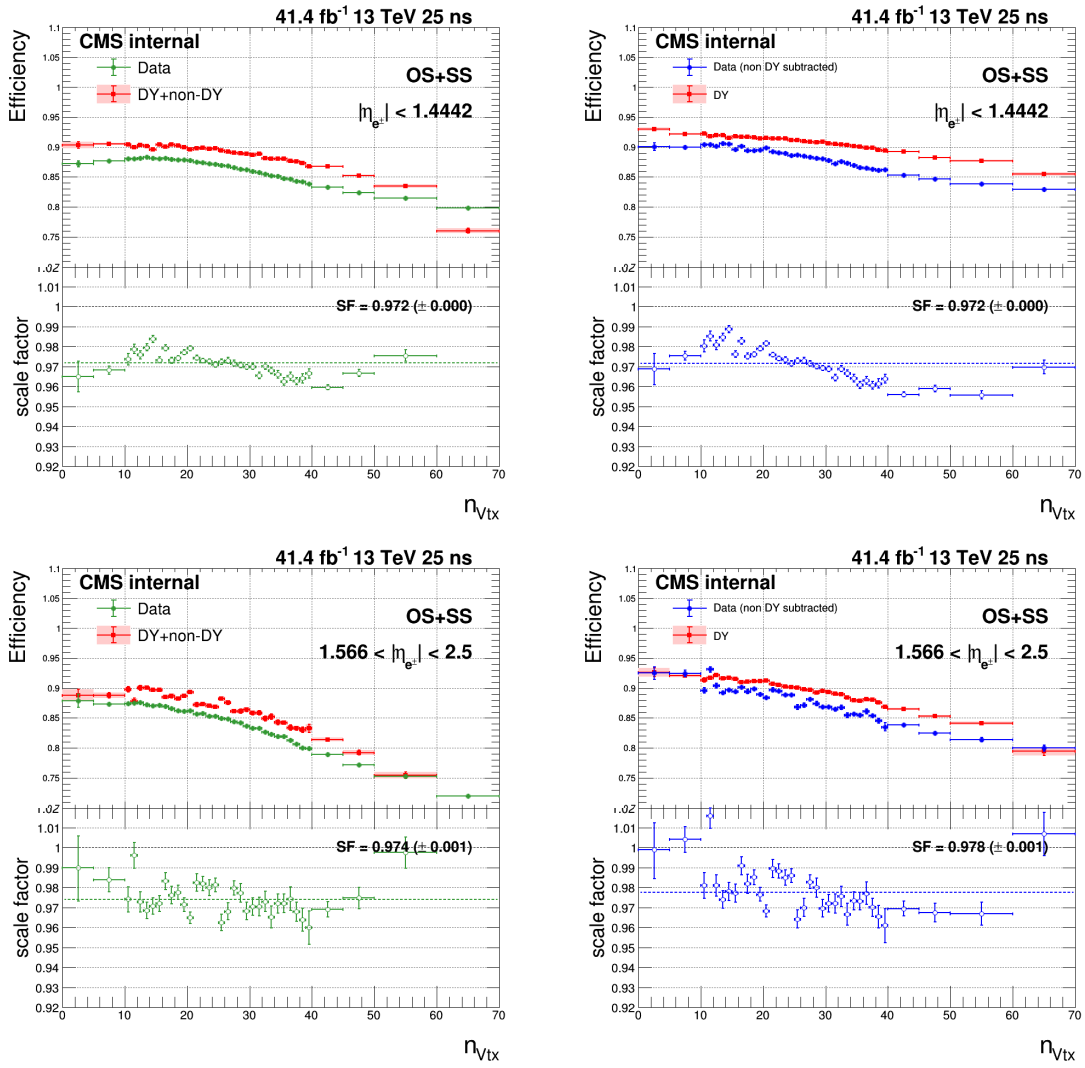

Figure 5.35: The HEEP efficiencies in data and MC, as well as the scale factors when the probe is in the barrel (top) and endcap (bottom) where the non-DY processes are included (left) and subtracted (right) as functions of the number of primary vertices, n_{Vtx} in 2017.

The main systematic uncertainties on the scale factor are from the normalization of non-DY processes. The cross sections of $t\bar{t}$ and W+jets processes are varied by 10% and 50%, respectively. For QCD estimation, it is considered with 50% uncertainty. Summary of the HEEP ID efficiencies and the scale factors are shown in Table 5.6.

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 5.6: The HEEP efficiencies in data and MC, as well as the scale factors when the probe is 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 B.2 and B.3.

- 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 Figure B.19 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 fluctuation or something in Moriond17 MC DY sample. In Appendix B.2.3, we cross checked the HEEP efficiency and scale factor using DYJetsToLL amcatnlo sample which shows the DY efficiencies in endcap in last two bins are normal. From Figure B.19, it can be seen that the efficiency in the MC for 500 GeV electrons is normal if you do not 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 B.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 B.2.1 Figure B.12). This effect is small (<1%) so it is ignored.

- In 2017
- In run F the pixel detector has a lower efficiency in some regions which can be seen

in Figure B.33 in Appendix B.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 5.36. 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 because the pile up is significant increased after run E which is shown in Figure 5.1. 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 is shown in Appendix B.3.2. Comparing the nominal "cut count method" and "fit method" they are consistent within 1% for barrel and 2% for endcap.

More details about the HEEP ID scale factor study can be found in Ref. [114] ([115]) for 2016 (2017).

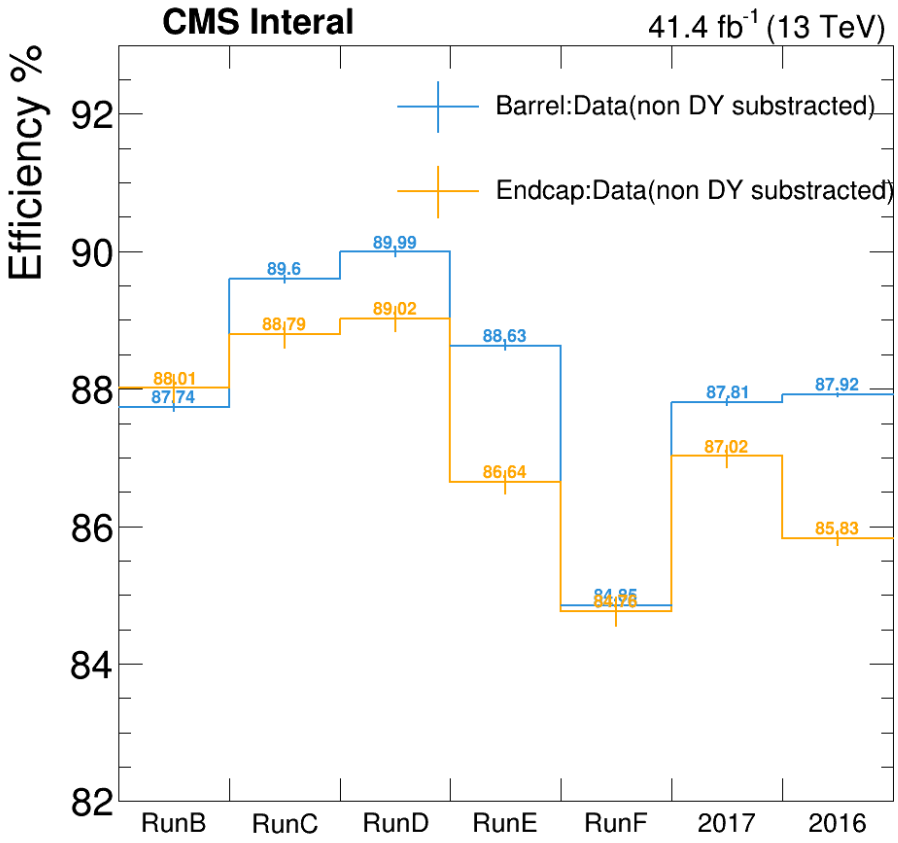

Figure 5.36: HEEP ID efficiency in data after non-DY contributions are subtracted for different runs for barrel and endcap in 2017.

5.6 Standard Model Backgrounds

There are three main kinds of backgrounds for this analysis. The most important and irreducible one is the SM Drell-Yan process.

The second most important background type includes $t\bar{t}$ and $t\bar{t}$ like processes, which contain WW, WZ, ZZ, tW, and $Z/\gamma^* \rightarrow \tau^+\tau^-$. The $t\bar{t}$ is the dominant background comparing to the others. However, WW events become increasingly important at high mass. Indeed at high mass, the top and antitop are boosted. The top decay products are then collimated, the b jet is close to the electron from the W decay. In this case, the electron usually fails the isolation requirements, reducing the contribution of $t\bar{t}$ background at high mass.

The third type of background is the jet background, where jets are misidentified as electrons. The jet background mainly comes from W+jets and multijets processes.

5.6.1 SM Drell-Yan background

The SM Drell-Yan background is estimated by MC simulation and the MC samples are normalized to the data in the Z boson peak region of 60-120 GeV. They are also corrected with the trigger turn on curve.

Figure 5.37 shows the data and MC at the Z peak for the barrel-barrel and barrel-endcap regions. In this case, in order to see the normalisation agreement between data and MC, MC events are normalized to the luminosity. Besides, the trigger turn on curve is applied and the scale factors between data and MC for gsf electron reconstruction efficiency and HEEP ID efficiency are applied. Further to this, the official EGamma corrections for electron energy scale and smearing are applied to data and MC in order to get better data and MC agreement.

A measured DY cross sections including the trigger efficiencies and data/MC efficiency scale factors are shown in Table 5.7. The SM DY cross section in mass range 60-120 GeV at NNLO is 1928±72 pb. The measurement is in good agreement with the theory value. There is a ∼2% difference for barrel-barrel region and a ∼0.3% difference for barrel-endcap region in 2016. For 2017, it is ∼2% for barrel-barrel region and ∼4% for barrel-endcap region.

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×eff	0.0880±0.001(stat.)	0.0315±0.001(stat.)	0.0807±0.001(stat.)	0.0289±0.001(stat.)
Data/MC gsf RECO SF	0.979	0.985	-	-
Data/MC HEEP ID Eff SF	0.943±0.001(stat.)	0.953±0.002(stat.)	0.935±0.002(stat.)	0.947±0.004(stat.)
Luminosity (pb^{-1})	35867	35867	41368	41368
DY cross-section (pb)	1967±3(stat)±51(lumi)	1922±3(stat)±50(lumi)	1974±3(stat)	1854±3(stat)
Ratio to theory (1928 pb)	1.02	0.997	1.024	0.962

Table 5.7: Measurements of the DY cross section in the range of $60 < M_{ee} < 120$ GeV. The scale factors for HEEP ID efficiency between data and MC are taken from Table 5.6. It should be noticed that the scale factors in Table 5.6 are for individual electrons, while here they are for electron pairs. For 2017, the gsf electron reconstruction efficiency scale factor between data and MC is already included in MC acc×eff [110, 111].

5.6.1.1 DY Background Correction and Uncertainty

The main uncertainties on the Drell-Yan background originate from the parton density function (PDF) (which is the probability of finding a parton with a fraction x of the

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 5.8: The PDF uncertainties on the DY cross section relative to the Z boson peak region (60-120 GeV) as a function of mass [110].

proton momentum in the proton) and higher-order effects. The cross-section has been evaluated using FEWZ 3.1.b2 with next-next-leading order (NNLO) accuracy in QCD and next-leading order (NLO) in electroweak (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 (60-120 GeV) 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 Ref. [116]. 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 5.38.

It is immediately noted that the POWHEG NNPDF3.0 prediction is increasingly higher than the FEWZ prediction as the mass increases. In the analysis a FEWZ to POWHEG k-factor is applied. This slightly improves data/MC agreement at high mass. The functional form for this k-factor (accounting for the fact we normalize in the 60-120 GeV region) is shown in Figure 5.38. 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 5.8. The uncertainties quoted are on the DY cross section ratio of the invariant mass range considered to the Z peak region of 60 to 120 GeV. The uncertainties are fitted by polynomial which is shown in Figure 5.39.

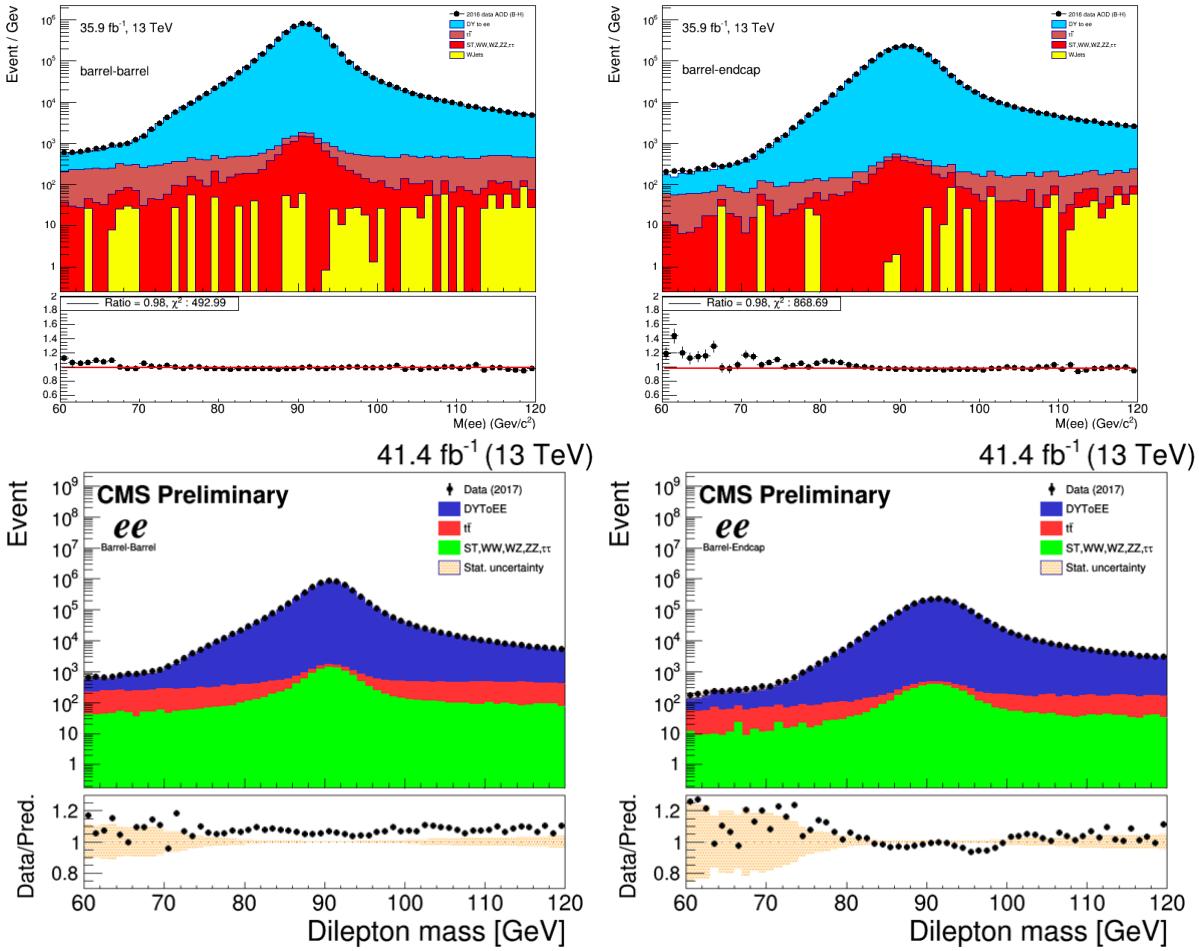

Figure 5.37: Data - MC agreement in the Z peak region with the MC normalized to the luminosity of data for the barrel-barrel (left) and barrel-endcap (right) regions. The trigger turn on curve, gsf reconstruction and HEEP ID scale factor are applied for MC in 2016 (top) [110] and 2017 (bottom) [111].

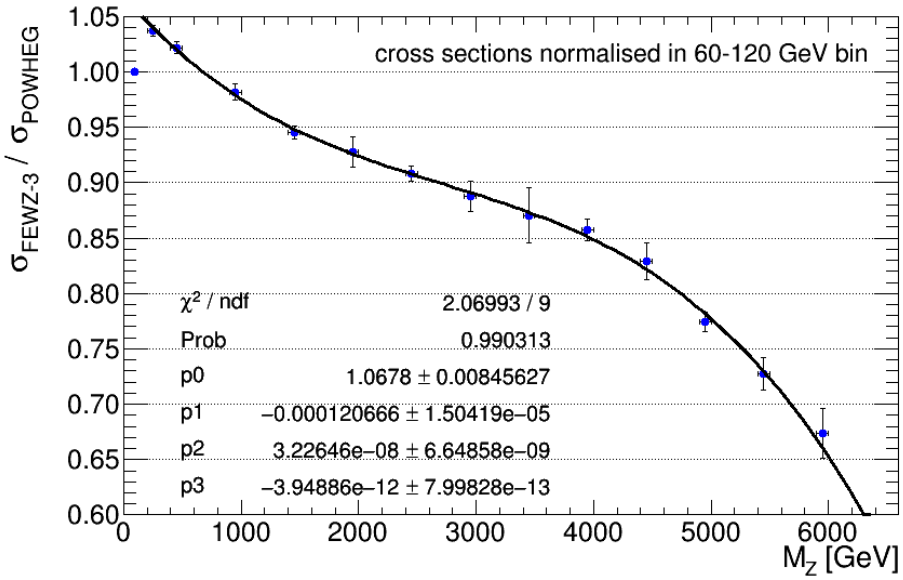

Figure 5.38: The ratio of the $Z/\gamma^* \to e^+e^-$ cross-section as predicted by FEWZ 3.1.b2 using the LUXqed_plus_PDF4LHC15_nnlo_100 PDF set and that predicted by POWHEG at NLO using the NNPDF3.0 PDF set in various mass bins. The cross-sections are normalized to each other in the 60 to 120 GeV bin [110].

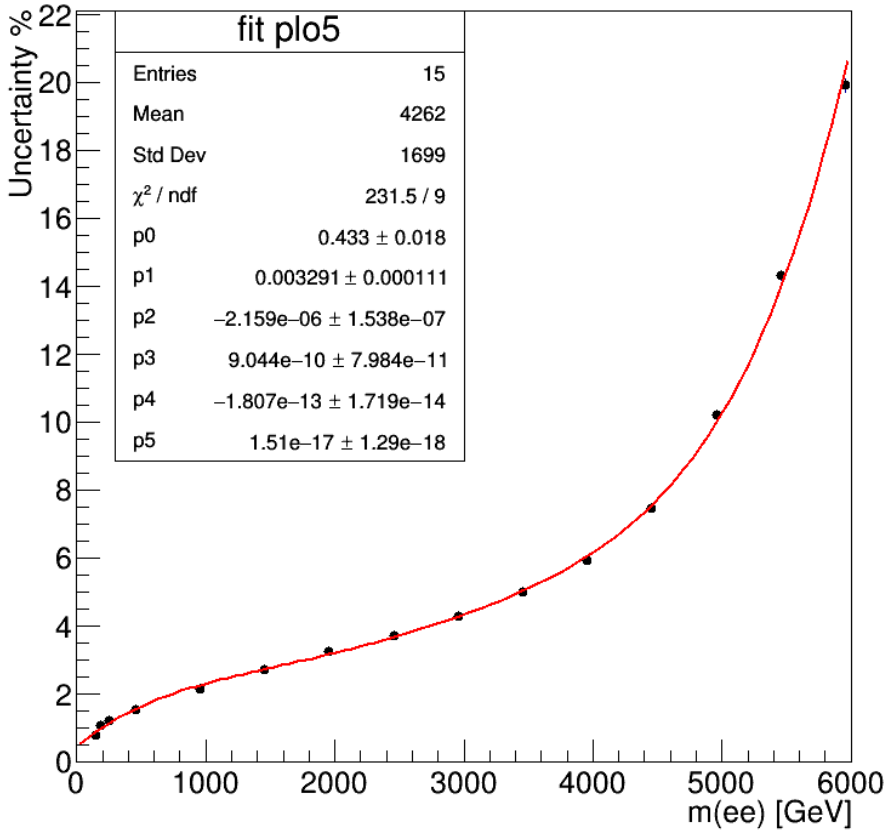

Figure 5.39: The PDF uncertainties on the DY cross section relative to the Z peak region as a function of dielectron invariant mass.

5.6.2 $t\bar{t}$ and $t\bar{t}$-like backgrounds

The background for electrons coming from $t\bar{t}$ and $t\bar{t}$-like processes is estimated directly from MC simulation.

As we know, all these processes are flavour-symmetric and the branching ratio to eμ is twice larger than the branching ratio to ee for all these processes. Because of that the e-μ invariant mass spectrum provides an excellent control region to validate the MC predictions for these processes. This validation is described below. In order to study the e-μ mass spectrum, we use the SingleMuon datasets and the "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 [117], 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 with 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 5.2, all scale factors have been applied to MC events.

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 eμ 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 eμ 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 eμ pair should be the same.

The invariant mass spectra of eμ events in data and MC are shown for both datasets in Figure 5.40.

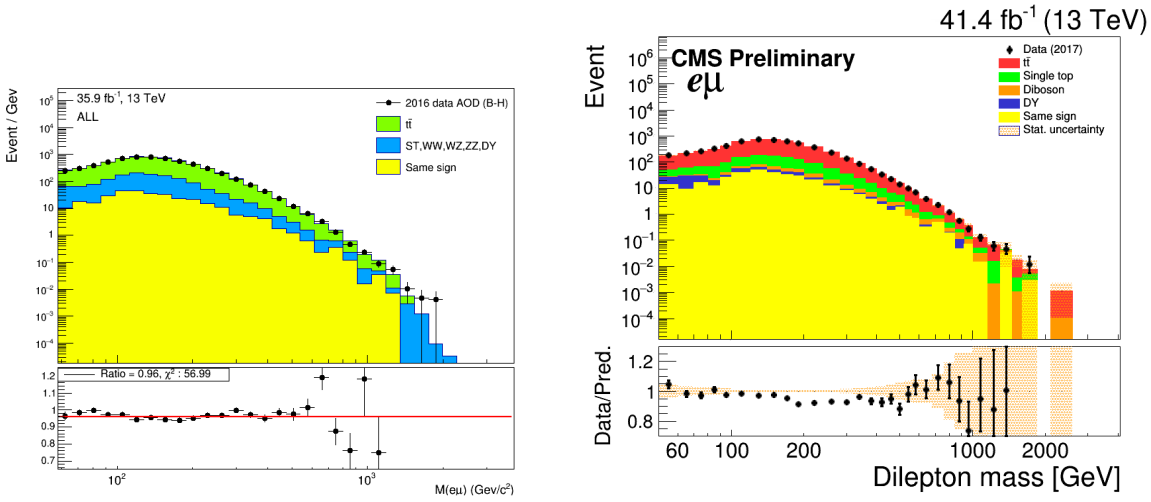

Figure 5.40: Invariant mass spectra of the eμ events in data and MC for 2016 (left) [110] and 2017 (right) [111].

After having checked that the simulations describe the sample of eμ events well, the contribution of these backgrounds to the ee spectrum is estimated from Monte Carlo. The tendency of the data/MC ratio to decrease at high mass come from a known effect from a misdescription of the top p_T in the $t\bar{t}$ simulation samples (from higher order corrections),

which is not taken into account in the above plots.

5.6.3 Jet background

The smallest of the backgrounds is coming from the jet background. The primary components of the jet background contains W + jets process (where the W decays to an eν pair and a jet passes the electron ID criteria) and QCD process (where two jets have passed electron ID criteria). The jet background is estimated by the "fake rate method" which is explained below.

5.6.3.1 Jet to electron misidentification probability measurement

The "fake rate" is defined as the probability for a jet reconstructed as an electron to pass the HEEP ID selection. The "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 5.9 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 $Z/\gamma^* \to e^+e^-$ 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 normalized 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 normalized to the observed N-1 track isolation distribution in the range of 10 $<$Isol p_T 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.

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

To summarize 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 normalized 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 5.41 (5.42), together with simple fitted functions to allow the fake rate to be easier to apply in the analysis. The fit parameters are summarized in Table 5.10. A 50% uncertainty on the method is assumed

Variable	Barrel	Endcap
$\sigma_{i\eta i\eta}$	<0.013	<0.034
H/E	<0.15	<0.10
nr. missing hits	<= 1	<= 1
|dxy|	< 0.02	< 0.05

Table 5.9: 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 5.10: The functional approximation of the measured fake rate for HEEP electrons in the barrel and endcap vs E_T [110, 111].

and this covers deviations from the arbitrarily chosen fits. More details can be found in [110].

5.6.3.2 Jet mass spectrum estimates

The jet background contribution is estimated through selecting electron pairs which passing the primary trigger of the analysis and one of the electron passing the HEEP selection and another passing the fake rate (FR) pre-selection (see in Table 5.9) but failing the HEEP selection. This is called "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. 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 $Z/\gamma^* \to e^+e^-$ events which is corrected for by directly subtracting off the MC estimate.

The 1FR estimate includes the background from W + jets and QCD processes, while because of combinatorial effects, the 1FR estimate overestimates the QCD contribution by a factor 2. The QCD contribution can be estimated by selecting electron pairs passing primary analysis trigger and both electrons pass the FR pre-selection but fail the HEEP selection. This is called "2FR estimate". The selected events are then 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. Finally, the 1FR estimate is subtracted by the 2FR estimate to estimate the total jet background contribution without 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 being in the endcap region. The endcap-endcap region is used because there are more jets in endcap comparing with the barrel region. The results are shown in Figure 5.43 (the uncertainty band is explained in Section 5.7) and one can see the data and MC agree within uncertainty.

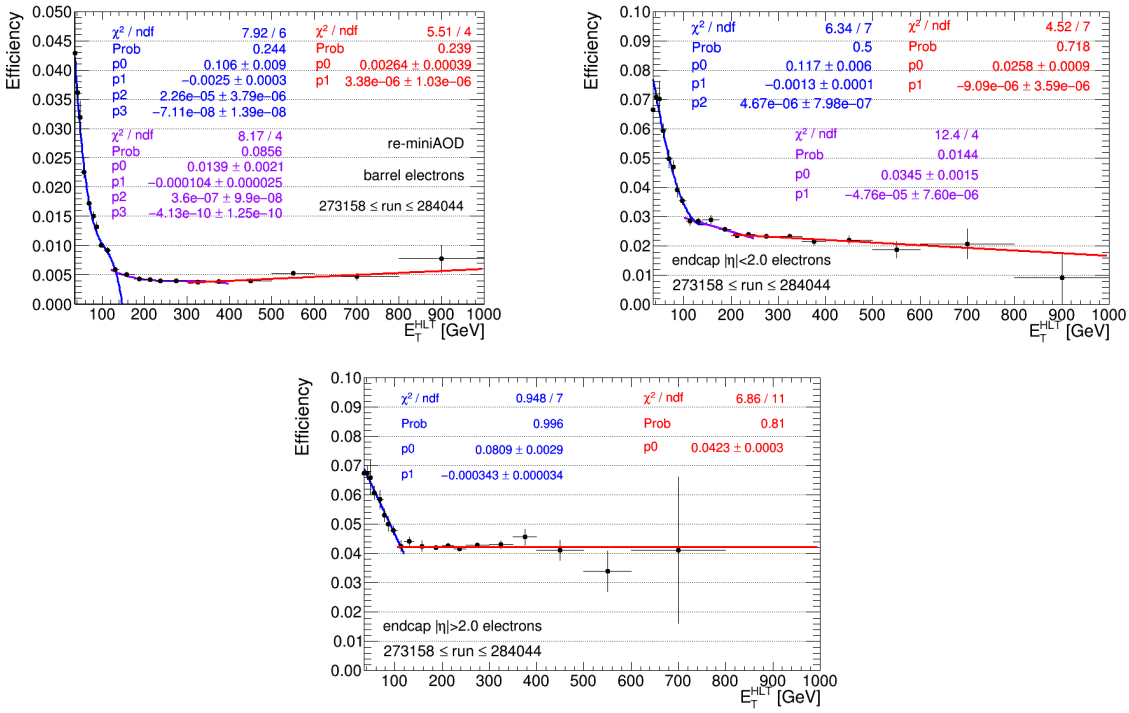

Figure 5.41: The measured HEEP ID fake rate vs E_T for the barrel region (top left), the endcap low $|\eta|$ region (top right) and the endcap high $|\eta|$ region (bottom) in 2016 [110].

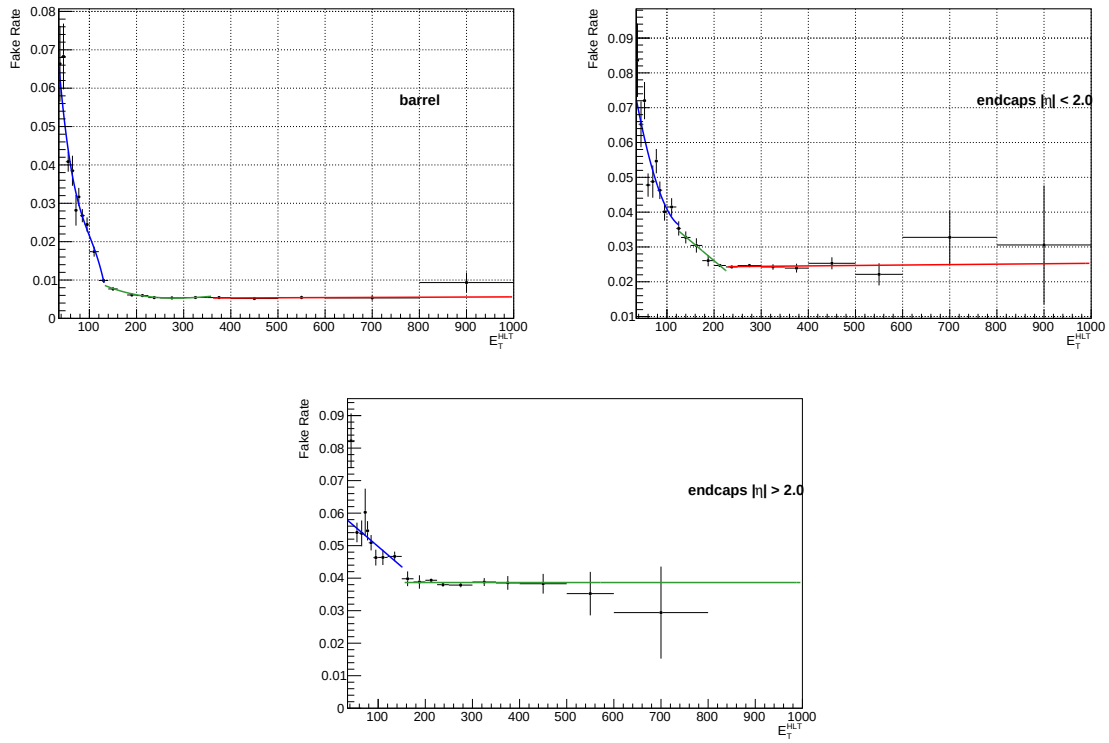

Figure 5.42: The measured HEEP ID fake rate vs E_T for the barrel region (top left), the endcap low $|\eta|$ region (top right) and the endcap high $|\eta|$ region (bottom) in 2017 [111].

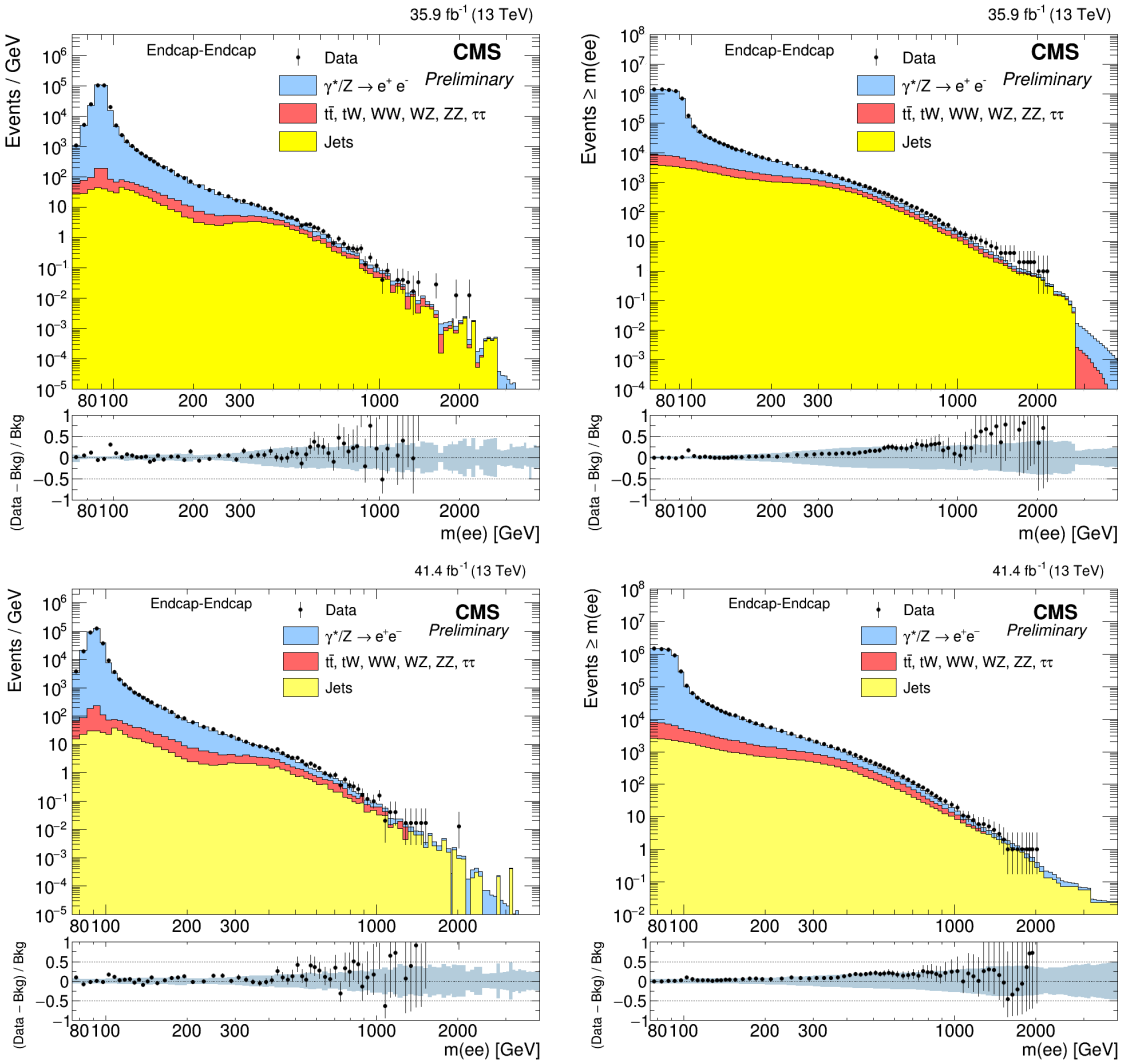

Figure 5.43: The dielectron mass spectra (left) and the cumulated distributions (right) for both electrons in the endcaps in 2016 (top) and 2017 (bottom).

5.7 Invariant Mass Spectra

The final dielectron mass spectra from data and standard model prediction are shown in Figure 5.44 (5.45) together with the integral of the mass spectra in 2016 (2017). The ratios of dielectron mass spectra between data and standard model prediction in signal region are shown in Figure 5.46 (5.47) in 2016 (2017).

The following systematic uncertainty sources are considered.

- Pile up reweighting: due to the HEEP ID scale factor (see Figures 5.22 and 5.27) is not flat versus pile up. Therefore, different pile up reweighting condition will give different total number of MC events. This uncertainty is considered by scaling the minimum bias cross section (which is used for producing the pile up distribution of data) up and down by 1 sigma (4.6%).
- DY PDF: due to the DY mass spectrum may dependent on the choosing of PDF, the effect of using different PDFs is estimated (see Section 5.6.1.1). This uncertainty is considered by applying a mass dependent uncertainty on the DY cross section (for mass $>$ 120 GeV) (see Figure 5.39).
- Cross section of background processes: 7% (around 1 sigma for $t\bar{t}$) uncertainty is considered on the cross section of the non-DY background estimated from MC for different processes.
- Jet background: 50% uncertainty is considered on fake jet contribution due to the uncertainty on the fake rate estimation.
- Electron energy scale: From the energy scale study using boosted Z events (see Figures 5.13 and 5.14), the considered uncertainty on electron energy scale for high mass region is 2% for barrel-barrel and 1% in barrel-endcap (for mass $>$ 120 GeV).
- HEEP ID scale factor: From the HEEP ID scale factor study (see Section 5.5), one can see in low E_T region the scale factor can be measured precisely, while for high E_T region, due to the lack of statistics, the uncertainty on the scale factor is increased. Therefore, a E_T dependent uncertainty on HEEP ID scale factor is considered. For barrel it is 1% below 90 GeV and 1-3% linearly increase for 90 GeV - 1 TeV range and 3% for higher than 1 TeV. For endcap it is 1% below 90 GeV and 1-4% linearly increase for 90 GeV - 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 90 GeV - 300 GeV and 5% for higher than 300 GeV.
- Normalization to Z peak: As we normalized MC to data in Z peak region, there are other uncertainties (which is not mentioned above) could affect Z peak normalization like the uncertainty on trigger efficiency measurement, gsf reconstruction efficiency measurement, and DY cross section in Z peak. All these uncertainties can be merged into one uncertainty which is the Z peak normalization uncertainty. Based on the DY cross section measurement in Z peak region study (see Section 5.6.1), the Z peak normalization uncertainty is considered as 1% in 2016 and 2% (4%) for barrel-barrel (barrel-endcap) in 2017.

In Table 5.11, predicted SM background and observed data yields are shown as a function of dielectron invariant mass.

In Table 5.12, the relative effect of each systematic uncertainty is shown as a function of dielectron invariant mass.

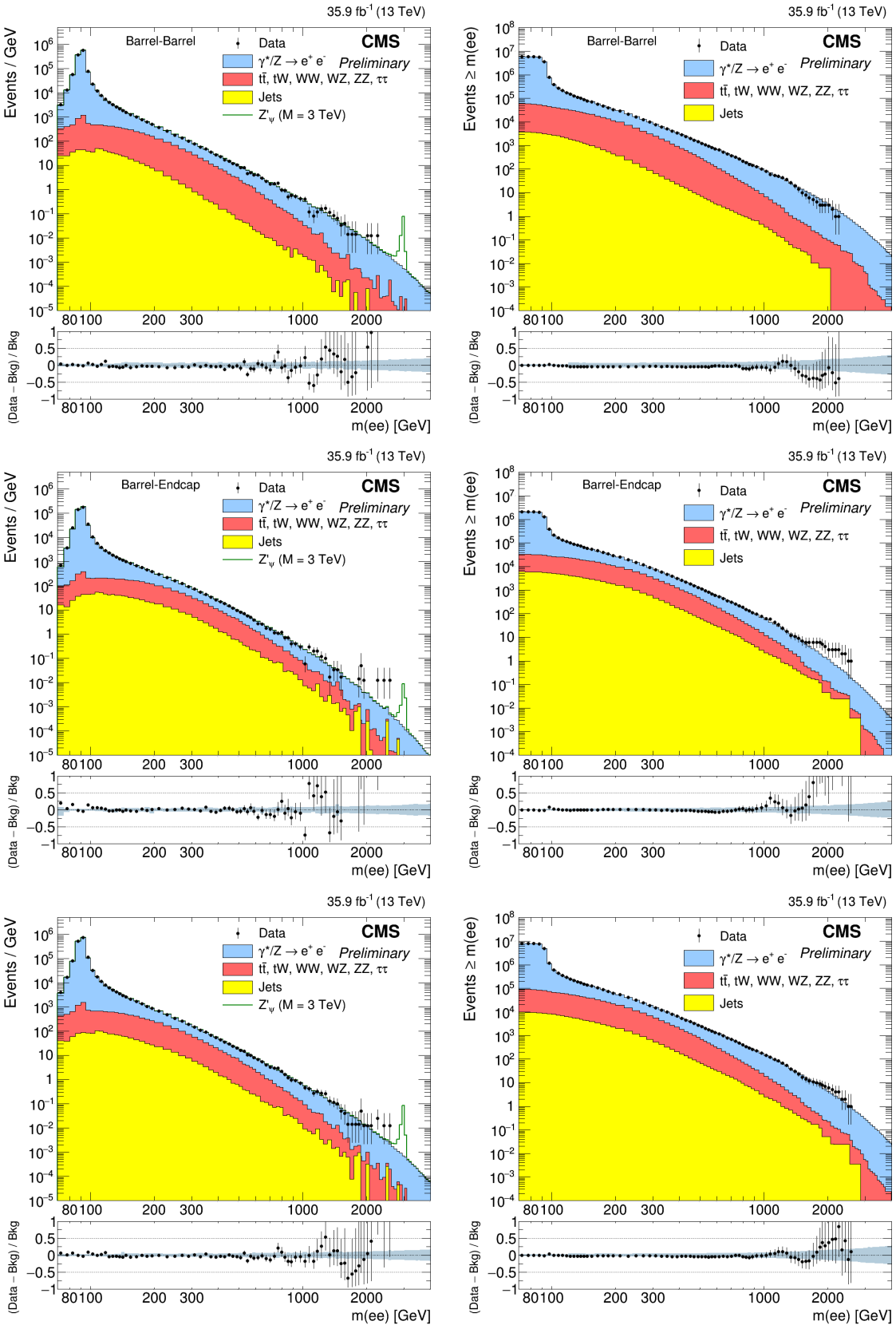

Figure 5.44: The observed dielectron mass spectrum (left) and the integral of the measured dielectron mass spectrum (right) for barrel-barrel (top), barrel-endcap (middle) and sum of the barrel-barrel and the barrel-endcap (bottom) together with the predicted standard model backgrounds in 2016.

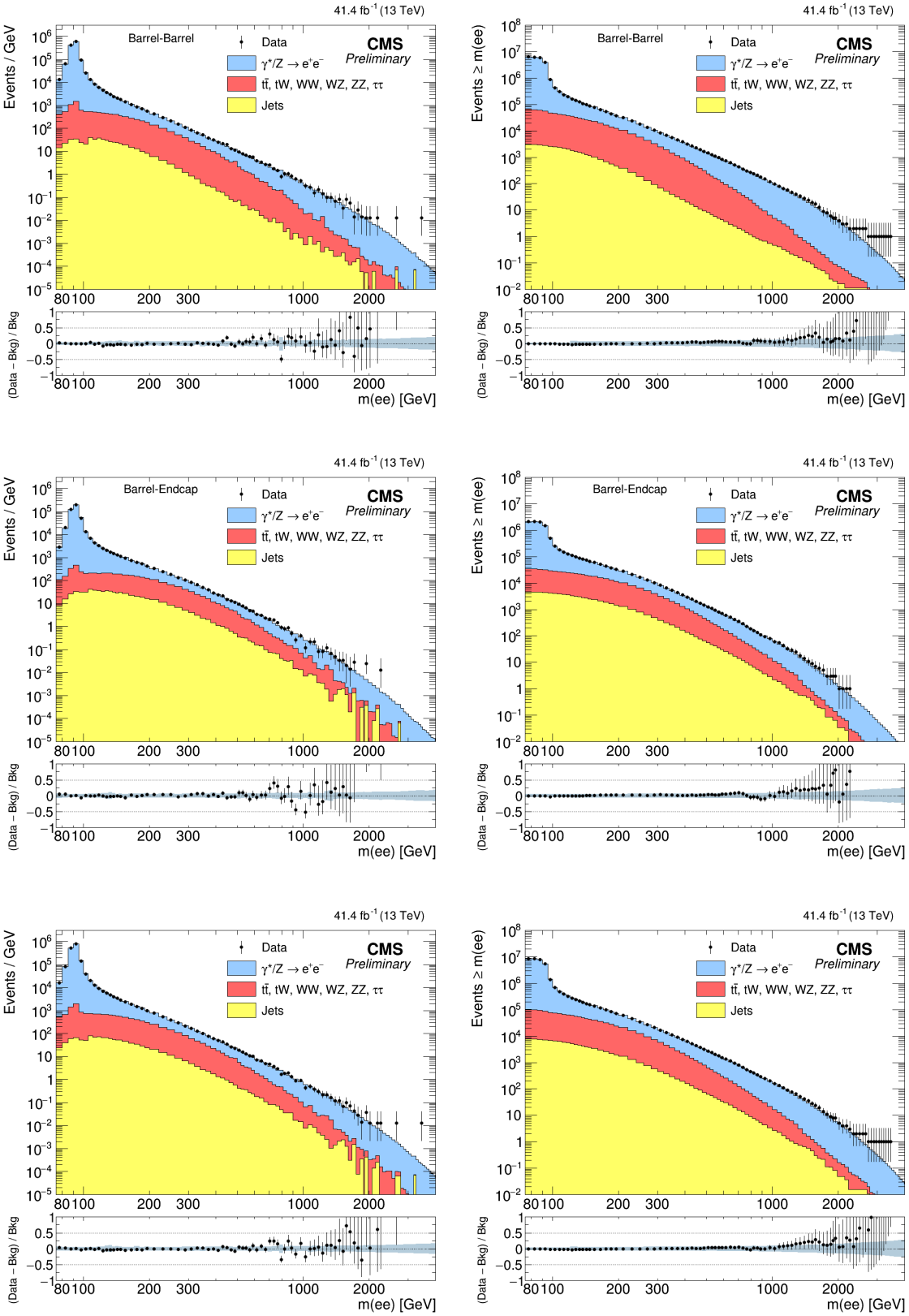

Figure 5.45: The observed dielectron mass spectrum (left) and the integral of the measured dielectron mass spectrum (right) for barrel-barrel (top), barrel-endcap (middle) and sum of the barrel-barrel and the barrel-endcap (bottom) together with the predicted standard model backgrounds in 2017.

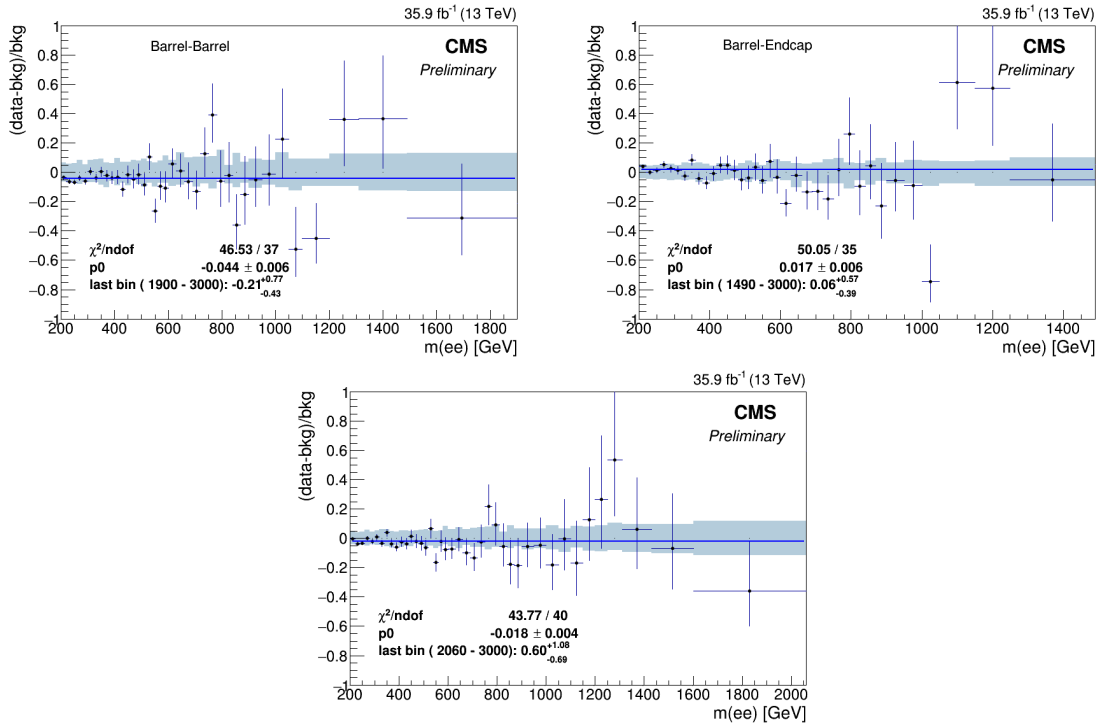

Figure 5.46: The ratio of observed dielectron mass spectrum in signal region for barrel-barrel (top left), barrel-endcap (top right) and sum of the barrel-barrel and the barrel-endcap together (bottom) in 2016.

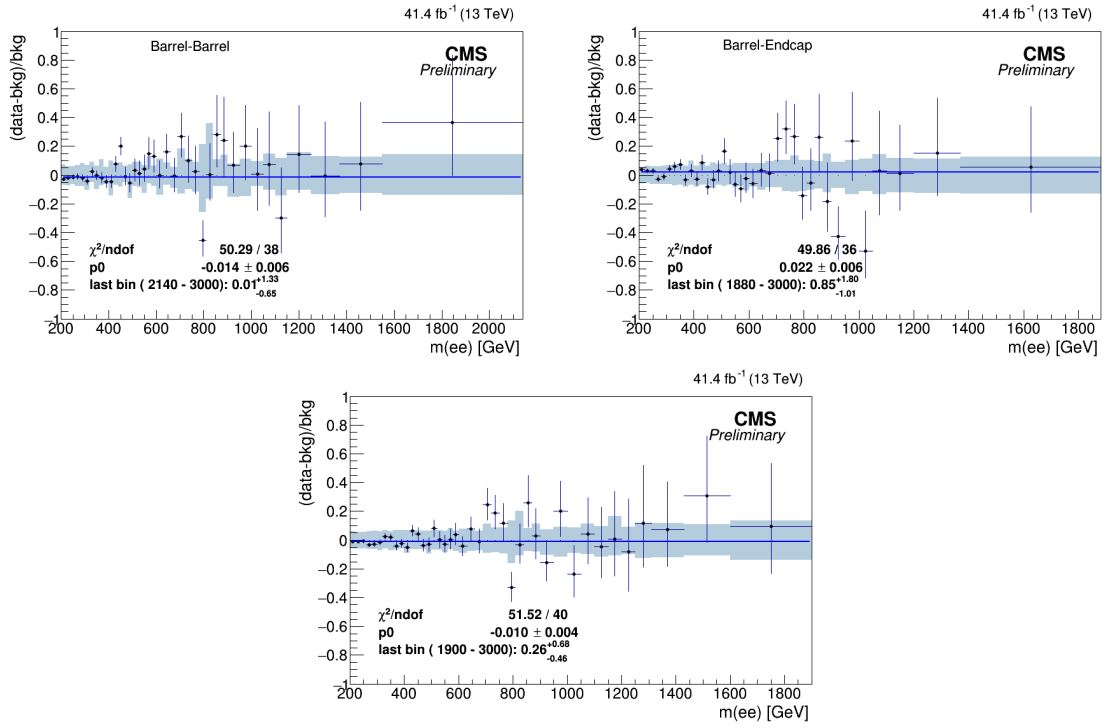

Figure 5.47: The ratio of observed dielectron mass spectrum in signal region for barrel-barrel (top left), barrel-endcap (top right) and sum of the barrel-barrel and the barrel-endcap together (bottom) in 2017.

Year	Mass range (GeV)	Data	Total bkg	$\gamma^*/Z \to ee$	$t\bar{t}$ and $t\bar{t}$-like bkg	Jets
2016	Barrel-Barrel					
	60 – 120	5760346	5762889.7 ± 133911.3	5730973.8 ± 133096.7	29369.6 ± 1277.6	2546.3 ± 1273.2
	120 – 400	146598	152496.1 ± 11452.7	120819.5 ± 10364.5	29824.2 ± 1572.3	1852.5 ± 926.2
	400 – 600	2163	2295.3 ± 183.6	1636.5 ± 124.6	643.5 ± 63.2	15.4 ± 7.7
	600 – 900	523	520.1 ± 51.1	425.9 ± 39.2	91.8 ± 13.3	2.4 ± 1.2
	900 – 1300	100	107.9 ± 10.6	96.5 ± 9.4	10.9 ± 1.8	0.6 ± 0.3
	1300 – 1800	24	21.5 ± 2.8	20.5 ± 2.6	0.9 ± 0.2	0.1 ± 0.0
	1800 – 6000	3	5.4 ± 0.9	5.2 ± 0.9	0.2 ± 0.0	0.0 ± 0.0
	Barrel-Endcap					
	60 – 120	2051759	2054401.1 ± 40405.8	2042472.6 ± 40296.8	9270.3 ± 345.6	2658.2 ± 1329.1
	120 – 400	98503	99151.9 ± 4158.3	77350.1 ± 3306.5	17813.2 ± 728.6	3988.6 ± 1994.3
	400 – 600	2134	2117.0 ± 112.6	1243.8 ± 58.0	751.3 ± 44.6	121.9 ± 61.0
	600 – 900	420	463.3 ± 27.9	311.7 ± 18.5	128.3 ± 9.9	23.2 ± 11.6
	900 – 1300	82	78.4 ± 5.9	59.2 ± 4.6	15.9 ± 1.6	3.3 ± 1.7
	1300 – 1800	9	12.8 ± 1.2	10.3 ± 1.0	1.9 ± 0.4	0.6 ± 0.3
	1800 – 6000	6	2.1 ± 0.3	1.8 ± 0.2	0.1 ± 0.0	0.1 ± 0.1
	Barrel-Barrel + Barrel-Endcap					
	60 – 120	7812105	7817300.8 ± 166656.9	7773476.7 ± 165766.3	38627.6 ± 1524.1	5196.5 ± 2598.2
	120 – 400	245101	252071.4 ± 12933.9	198526.1 ± 11326.7	47709.5 ± 2148.8	5835.8 ± 2917.9
	400 – 600	4297	4424.5 ± 232.4	2887.1 ± 147.9	1400.2 ± 87.6	137.2 ± 68.6
	600 – 900	943	985.9 ± 63.9	739.3 ± 48.7	221.0 ± 17.2	25.6 ± 12.8
	900 – 1300	182	186.6 ± 13.5	155.9 ± 11.9	26.8 ± 2.3	3.9 ± 1.9
	1300 – 1800	33	34.3 ± 3.4	30.9 ± 3.2	2.8 ± 0.5	0.6 ± 0.3
	1800 – 6000	9	7.5 ± 1.1	7.0 ± 1.1	0.3 ± 0.0	0.1 ± 0.1
2017	Barrel-Barrel					
	60 – 120	6190697	6194808.2 ± 178177.2	6156571.2 ± 177324.0	34116.3 ± 1921.1	4120.7
	120 – 400	162005	167925.2 ± 13932.9	132981.8 ± 12618.8	33128.5 ± 2365.8	1815.0
	400 – 600	2503	2404.7 ± 215.9	1782.7 ± 144.5	605.2 ± 79.6	16.8
	600 – 900	588	560.2 ± 51.5	478.6 ± 44.9	78.7 ± 15.6	3.0
	900 – 1300	118	113.4 ± 13.1	105.1 ± 11.4	7.8 ± 2.8	0.5
	1300 – 1800	28	23.1 ± 3.0	23.0 ± 3.0	0.0 ± 0.0	0.1
	1800 – 6000	7	5.7 ± 1.0	5.7 ± 1.0	0.0 ± 0.0	0.0
	Barrel-Endcap					
	60 – 120	2096490	2098260.5 ± 96902.6	2086010.3 ± 96566.8	10473.3 ± 675.3	1777.0
	120 – 400	109771	110357.8 ± 6227.6	87277.3 ± 5107.0	19860.0 ± 1450.6	3220.6
	400 – 600	2365	2364.5 ± 164.6	1442.8 ± 99.9	810.5 ± 68.3	111.3
	600 – 900	518	488.3 ± 37.9	341.8 ± 26.5	124.1 ± 13.9	22.4
	900 – 1300	75	86.5 ± 7.7	69.5 ± 6.3	14.0 ± 2.7	3.0
	1300 – 1800	16	14.2 ± 1.7	11.7 ± 1.3	2.0 ± 1.0	0.6
	1800 – 6000	3	2.2 ± 0.3	2.1 ± 0.3	0.0 ± 0.0	0.1
	Barrel-Barrel + Barrel-Endcap					
	60 – 120	8287187	8290233.1 ± 369104.3	8242650.8 ± 364680.2	44535.5 ± 2834.7	3046.7
	120 – 400	271776	280802.3 ± 18298.7	222376.6 ± 15710.7	53411.9 ± 3976.3	5013.9
	400 – 600	4868	4841.2 ± 329.4	3267.6 ± 216.8	1446.0 ± 128.0	127.6
	600 – 900	1106	1062.0 ± 77.4	829.4 ± 63.0	207.4 ± 22.6	25.3
	900 – 1300	193	201.3 ± 18.0	176.2 ± 16.0	21.6 ± 3.9	3.5
	1300 – 1800	44	37.6 ± 4.2	34.8 ± 4.0	2.1 ± 1.1	0.7
	1800 – 6000	10	7.9 ± 1.2	7.8 ± 1.2	0.0 ± 0.0	0.1

Table 5.11: 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 regions. The uncertainty contains statistic uncertainty and systematic uncertainty.

Year	Uncertainty	60-120 GeV	120-400 GeV	400-600 GeV	600-900 GeV	900-1300 GeV	1300-1800 GeV	1800-6000 GeV
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%
	Bgk_scale_up	0.036%	1.334%	1.870%	1.460%	1.092%	0.958%	0.888%
	Bgk_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%

Table 5.12: The relative effect of each systematic uncertainty as a function of dielectron invariant mass.

5.7.1 Complementary plot

In addition to the invariant mass plots, the distributions of the following variables are shown in Figure 5.48 (5.49) for 2016 (2017).

– invariant mass of the selected electron pair in Z peak region for different catagories
– 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

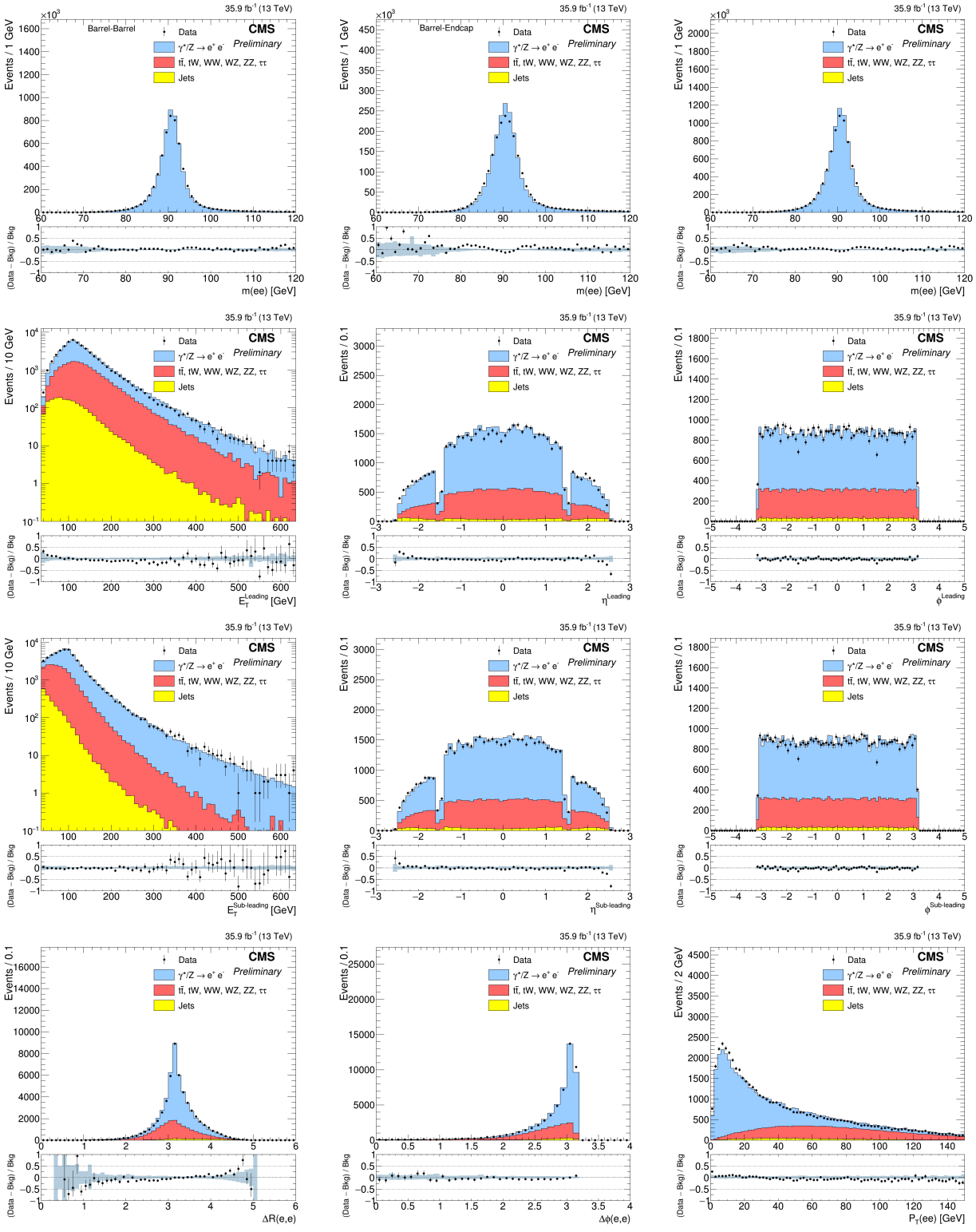

Figure 5.48: The distributions of invariant mass of two electrons in barrel-barrel, barrel-endcap and barrel-barrel + barrel-endcap (first row), E_T, η and ϕ of leading electron (second row) and E_T, η and ϕ of sub-leading electron (third row),ΔR, $\Delta\phi$ between two electrons and p_T of Z (fourth row) in 2016.

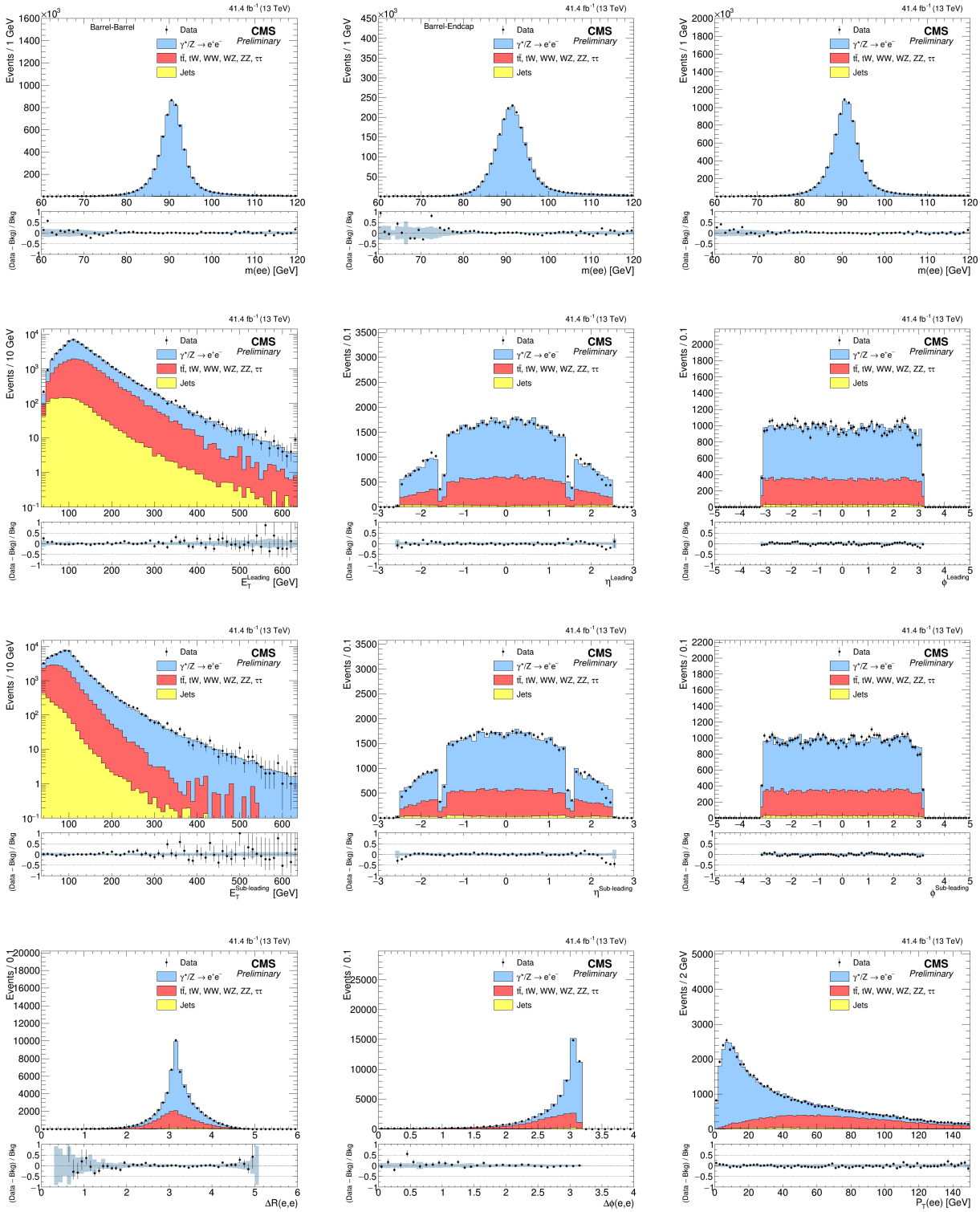

Figure 5.49: The distributions of invariant mass of two electrons in barrel-barrel, barrel-endcap and barrel-barrel + barrel-endcap (first row), E_T, η and ϕ of leading electron (second row) and E_T, η and ϕ of sub-leading electron (third row),ΔR, $\Delta\phi$ between two electrons and p_T of Z (fourth row) in 2017.

5.8 Statistical Interpretation

Given that no 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 [118, 119, 120]. 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:

$$f(m|\boldsymbol{\theta},\boldsymbol{\nu}) = q_1 \cdot f_{\mathrm{S}}(m|\boldsymbol{\theta},\boldsymbol{\nu}) + (1-q_1)\cdot f_{\mathrm{B}}(m|\boldsymbol{\theta},\boldsymbol{\nu}) \tag{5.5}$$

where m is the dilectron 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 $\mathrm{Res}(m|\sigma,\boldsymbol{\theta})$:

$$f_{\mathrm{S}}(m|\Gamma,\sigma,\boldsymbol{\theta},\boldsymbol{\nu}) = \mathrm{BW}(m|\Gamma)\otimes \mathrm{Res}(m|\sigma,\boldsymbol{\theta},\boldsymbol{\nu}) \tag{5.6}$$

where Γ is the intrinsic width of the signal and σ is the mass resolution. As described in Section 5.4 the resolution function $\mathrm{Res}(m|\sigma,\boldsymbol{\theta},\boldsymbol{\nu})$ is described by a Cruijff function (which is defined as a Gaussian core connected with an exponential tail on each side) in 2016, and in 2017 the double-sided crystal ball function is used.

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 dielectron invariant mass and it can be expressed by:

$$\begin{aligned} &m^{\kappa}\exp\Big(\sum_{i=0}^{3}\boldsymbol{\alpha}_i m^i\Big), \text{if } m \le 600 \text{ GeV} \\ &m^{\lambda}\exp\Big(\sum_{i=0}^{3}\beta_i m^i\Big), \text{if } m > 600 \text{ GeV}, \end{aligned} \tag{5.7}$$

The parameters are determined by a fit to the simulated dielectron mass spectrum. The background spectra together with the fitted function are shown in Figure 5.50 for the barrel-barrel and barrel-endcap categories.

The unbinned likelihood is defined as:

$$\mathcal{L}\left(\boldsymbol{m}|R_\sigma,\boldsymbol{\nu}\right) \;=\; \prod_{i=1}^{N} f\left(m_i|R_\sigma,\boldsymbol{\nu}\right) \tag{5.8}$$

where the product is over the events in the dataset and $\boldsymbol{m}$ is the vector of corresponding dielectron 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:

$$R_\sigma = \frac{\sigma_{\mathrm{Z}'}\cdot \mathrm{BR}(\mathrm{Z}'\to \mathrm{ee})}{\sigma_{\mathrm{Z}}\cdot \mathrm{BR}(\mathrm{Z}\to \mathrm{ee})} \tag{5.9}$$

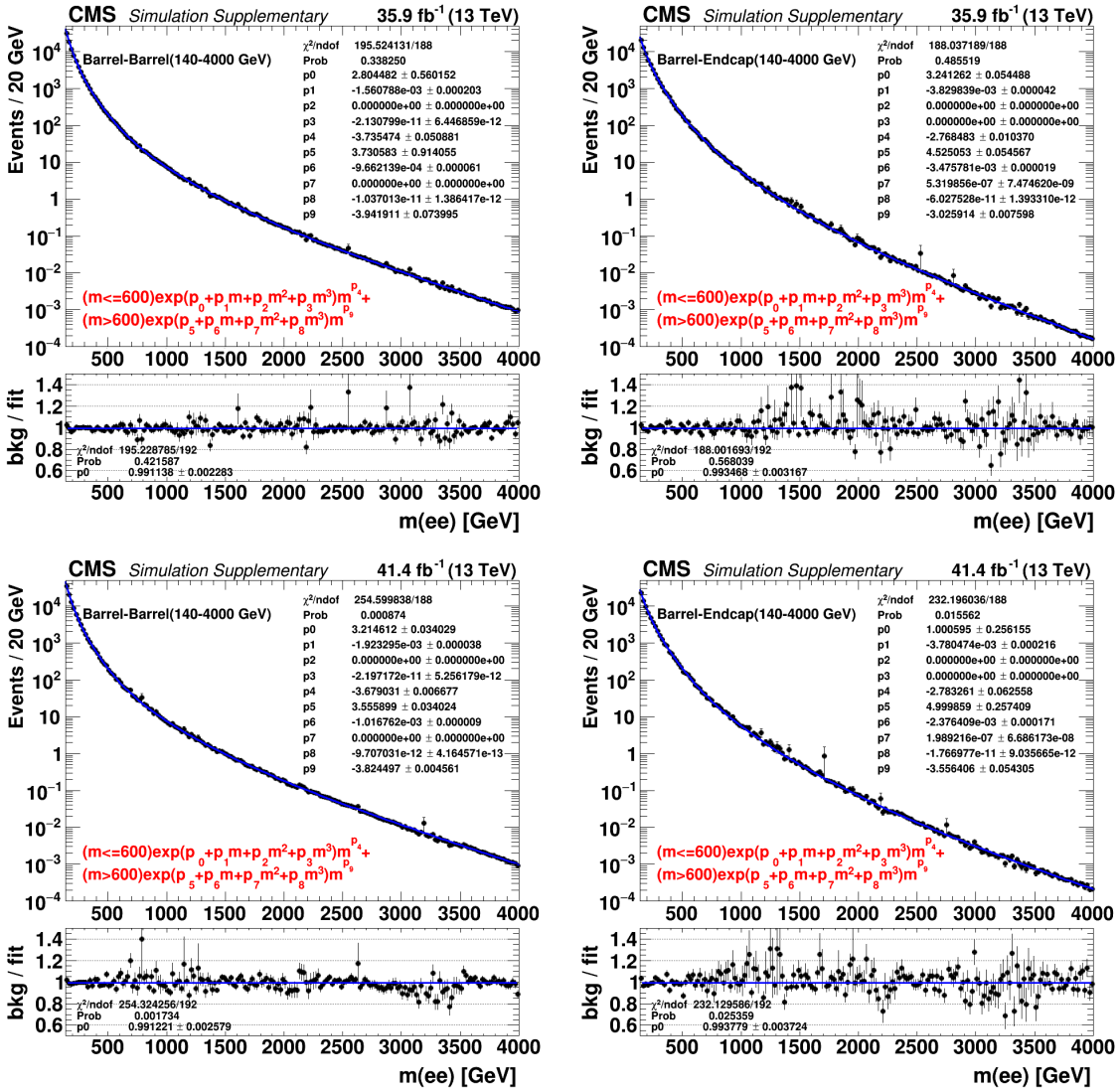

Figure 5.50: The total SM background together with the fitted functional form used to enter it into the limit setting tools for the barrel-barrel (left) and barrel-endcap (right) channels in 2016 (top) and 2017 (bottom).

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 (5.8) and inserting the equation for the signal and background pdfs detailed above, one obtains:

$$\mathcal{L}(\boldsymbol{m}|R_\sigma,\boldsymbol{\nu}) = \frac{\mu^N e^{-\mu}}{N!}\cdot\prod_{i=1}^{N}\left(\frac{\mu_S(R_\sigma,\boldsymbol{\nu})}{\mu}f_S(m_i|R_\sigma,\boldsymbol{\nu}) + \frac{\mu_B(R_\sigma,\boldsymbol{\nu})}{\mu}f_B(m_i|R_\sigma,\boldsymbol{\nu})\right) \quad (5.10)$$

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:

$$\mu_S = R_\sigma\frac{(\mathrm{Acc}\times\epsilon)_{Z'}}{(\mathrm{Acc}\times\epsilon)_Z}N_Z \quad (5.11)$$

where $(\mathrm{Acc}\times\epsilon)_{Z'}$ and $(\mathrm{Acc}\times\epsilon)_Z$ are the acceptance times efficiency of the Z' (shown in Figure 5.51) and the Z bosons 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

$$\boldsymbol{\nu} = \hat{\boldsymbol{\nu}}\cdot(1+\delta\boldsymbol{\nu})^\beta \quad (5.12)$$

where $\hat{\boldsymbol{\nu}}$ is the estimate of $\boldsymbol{\nu}$, $\delta\boldsymbol{\nu}$ 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

$$\mathcal{L}(\boldsymbol{m}|R_\sigma,\boldsymbol{\nu}) = \mathcal{L}(\boldsymbol{m}|R_\sigma,\boldsymbol{\nu}\ )\cdot\prod_j Gauss(\beta_j|0,1) \quad (5.13)$$

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:

$$f(R_\sigma,\boldsymbol{\nu}|\boldsymbol{m})\cdot p(\boldsymbol{m}) = \mathcal{L}(\boldsymbol{m}|R_\sigma,\boldsymbol{\nu})\cdot p(R_\sigma,\boldsymbol{\nu}) \quad (5.14)$$

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

$$p(R_\sigma|\boldsymbol{m})\cdot p(\boldsymbol{m}) = \mathcal{L}(\boldsymbol{m}|R_\sigma)\cdot p(R_\sigma) \quad (5.15)$$

The expression for the posterior pdf immediately follows as

$$p(R_\sigma|\boldsymbol{m}) = \frac{\mathcal{L}(\boldsymbol{m}|R_\sigma)\cdot p(R_\sigma)}{p(\boldsymbol{m})} = \frac{\mathcal{L}(\boldsymbol{m}|R_\sigma)\cdot p(R_\sigma)}{\int\mathcal{L}(\boldsymbol{m}|R_\sigma)\cdot p(R_\sigma)\,\mathrm{d}R_\sigma} \quad (5.16)$$

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:

$$\int_0^{R_\sigma^{95}} p(R_\sigma|\boldsymbol{m})\,\mathrm{d}R_\sigma = 0.95 \quad (5.17)$$

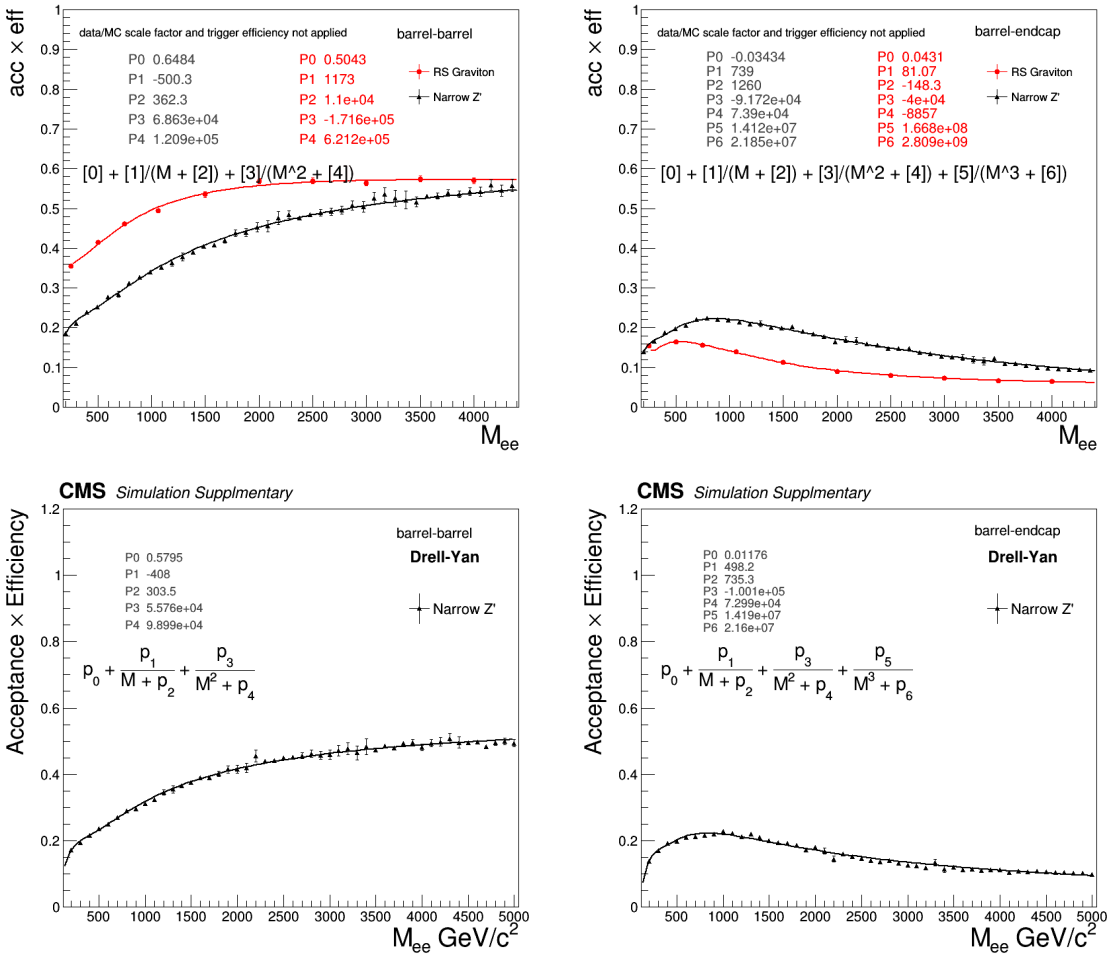

Figure 5.51: The acceptance times efficiency for a spin-1 or spin-2 particle to be selected by the analysis in the barrel-barrel region (left) and barrel-endcap region (right) together with the fitted functional form used to enter it into the limit setting tools in 2016 (top) [110] and 2017 (bottom) [111]. 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-Hasting algorithm [121, 122].

Finally, it is worth to mention that (see Equation (5.11)) the parametrization of the acceptance times efficiency $(\text{Acc} \times \epsilon)_{Z'}$ is shown in Figure 5.51. Other inputs required for the limit setting tool are listed in Table 5.13, as well as the number of data events and acceptance at the Z peak region. Since the limits are normalized 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 times 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 5.13: 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 [110, 111].

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

$$\mu_B = \hat{\mu}_B \cdot (1 + \delta\mu_B)^{\beta_B} \tag{5.18}$$

where β_B is again a random number extracted from a normal distribution as the case of Equation (5.12). 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.

5.8.1 Upper limits

Using the method described above, the observed and expected 95% upper limits on R_σ for a resonance width of 0.6% are shown in Figure 5.52. The signal R_σ curves are shown on the plot in order to obtain a mass limit on two specific Z′ signal model.

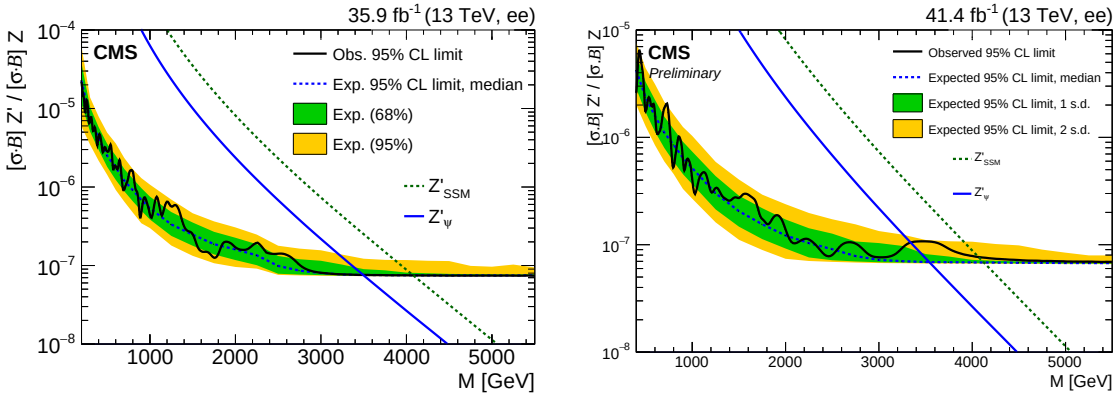

Figure 5.52: The 95% CL upper limits on R_σ for a spin 1 resonance with a width equal to 0.6% of the resonance mass for 2016 (left) [31] and 2017 (right) [32]. The shaded bands correspond to the 68% and 95% quantities for the expected limits. Theoretical predictions for the spin 1 Z'_{SSM} and Z'_{ψ} 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 [117]. 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 are shown in Figure 5.53.

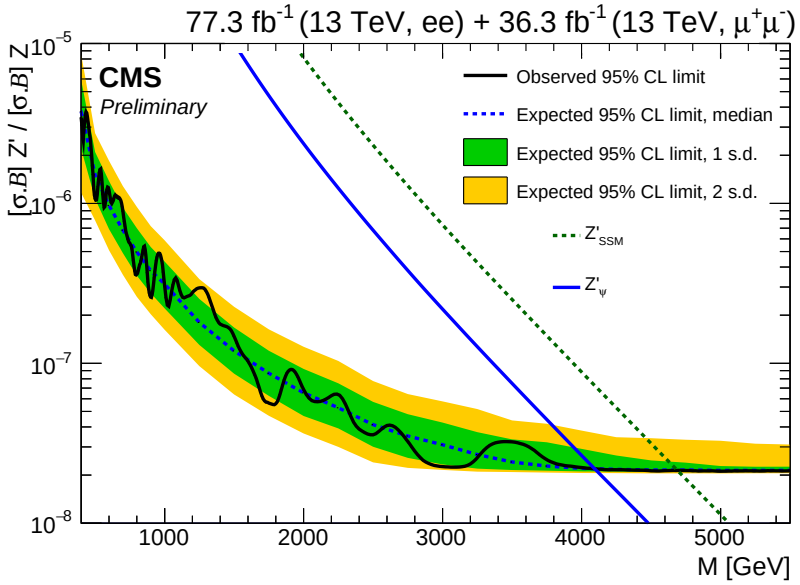

Figure 5.53: The 95% CL upper limits on R_σ for a spin 1 resonance with a width equal to 0.6% of the resonance mass for the dielectron (using 2016 and 2017 data) and dimuon (using 2016 data) final states combined. The shaded bands correspond to the 68% and 95% quantities for the expected limits. Theoretical predictions for the spin 1 Z'_{SSM} and Z'_{ψ} resonances are also shown [32].

In addition, in 2016 the results for widths equal to 0.6%, 3%, 5%, and 10% of the resonance mass are shown in Figure 5.54 and one can see at high masses the limits do not exhibit any dependence on the assumed resonance width.

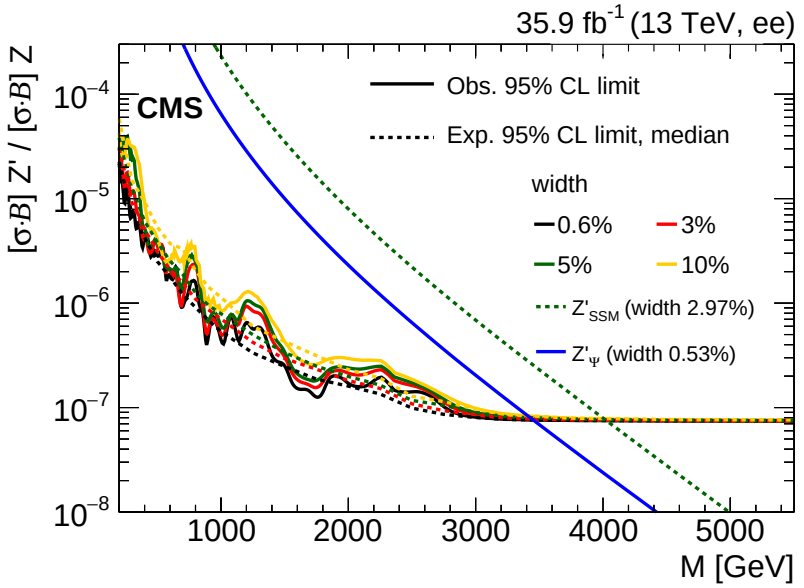

Figure 5.54: The 95% CL upper limits on R_σ for a spin 1 resonance with a width equal to 0.6%, 3%, 5%, and 10% of the resonance mass in 2016 [31].

Table 5.14 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.

	Z'_{SSM}		Z'_{ψ}	
Channel	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) + μμ (2016)	4.70	4.70	4.10	4.10

Table 5.14: The observed and expected 95% CL lower limits on the masses of spin 1 Z'_{ψ} and Z'_{SSM} bosons, assuming a signal width of 0.6% (3%) of the resonance mass for Z'_{ψ} (Z'_{SSM}) [31, 32].

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 $k/\overline{M}_{Pl}$ of 0.01, 0.05, and 0.10 are shown in Figure 5.55 for the dielectron channel and dielectron dimuon combined. Table 5.15 presents the values of the observed and expected 95% CL lower limits of the aforementioned models.

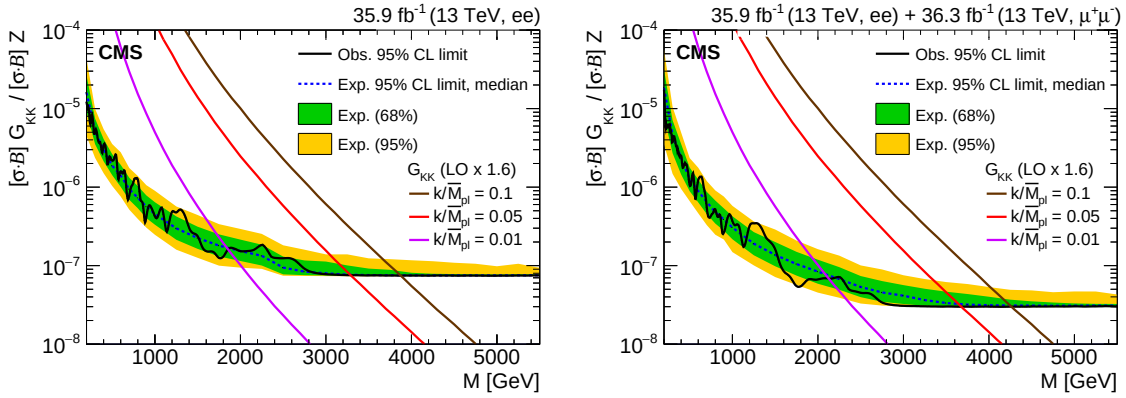

Figure 5.55: The 95% CL upper limits on R_σ for a spin 2 resonance for the dielectron (left) and its combination with dimuon (right) final states in 2016. The shaded bands correspond to the 68% and 95% quantities for the expected limits. Theoretical predictions for the spin 2 resonances with width equal to 0.01, 0.36, and 1.42 GeV corresponding to coupling parameters $k/\overline{M}_{Pl}$ of 0.01, 0.005, and 0.10 are shown for comparison [31].

Channel	$k/\overline{M}_{Pl}$=0.01		$k/\overline{M}_{Pl}$=0.05		$k/\overline{M}_{Pl}$=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) + μμ (2016)	2.10	2.05	3.65	3.60	4.25	4.25

Table 5.15: 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 $k/\overline{M}_{Pl}$ of 0.01, 0.005, and 0.10 [31].

5.9 Summary

A search for narrow resonances in the 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 $\overline{M}_{Pl}$ 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 $\overline{M}_{Pl}$ value of 0.01 (0.10).

My personal contributions to the analysis include performing weekly 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 understood it. The HEEP ID is also used in other CMS analyses, so providing HEEP ID scale factor becoming one of my EPR (Experimental Physics Responsibilities) 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 do not have saturated events in our final mass distribution in data. In addition, I 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 [123] and the other corresponding to the analysis of 41.4 fb^{-1} of integrated luminosity collected in full 2017 [32]. There is one paper [31] where the results are obtained by 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. For instance, ATLAS [124] puts a 95% confidence level lower mass limits 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. Recently ATLAS published their full Run2 analysis (see Ref. [37]), and found the following 95% confidence level lower mass limits: 5.1 TeV for the Z'_{SSM} model and 4.5 TeV for the Z'_ψ.

Chapter 6

Search for New Physics via Top Quark Production in Dilepton Final State

This chapter introduces the search for new physics via top quark production in ee and μμ final states (the motivation is described in Section 2.3). 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. The data and MC samples are introduced in Section 6.1. In Section 6.2 the triggers used are given. The object identification are presented in Section 6.3. The event selection is given in Section 6.4. In Section 6.5 the SM backgrounds are described. The data and MC comparisons for different variables are presented in Section 6.6. The signal extraction is expressed in Section 6.7. The systematic uncertainties are described in Section 6.8. Finally, the limits for different effective couplings are given in Section 6.9.

6.1 Data-sets and MC Samples

6.1.1 Data samples

The primary data sets used in this analysis are summarized in Table 6.1. The data from the Moriond17 rereco campaign (Run2016 03Feb2017 Re-miniAOD) are 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 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 6.1: Data sets (X) used in this analysis. X = "DoubleEG" and "SingleElectron" for the ee channel. X = "DoubleMuon" and "SingleMuon" for μμ channel analysis.

6.1.2 MC samples

The analysis uses centrally produced MC samples from "RunIISummer16 MiniAODv2" campaign as are recommended for Moriond17. All the MC samples are listed in Table 6.2.

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 6.2: MC samples.

The pile up distributions for MC and data, which are calculated by using 69.2 mb as

the minimum bias cross section, are shown in the left plot of Figure 5.1. MC events are re-weighted to account for the pile up difference between data and MC.

6.2 Triggers

In this analysis we use various sets of triggers in order to achieve an optimal selection efficiency. For each dilepton channel (ee, μμ) we take the logical OR of the dilepton triggers listed in Table 6.3. The double electron trigger "HLT_Ele23_Ele12_caloIdL_TrackIdL_IsoVL_DZ" (which selects two electrons and one of them with $p_T > 23$ GeV and another with $p_T > 12$ GeV, besides both electrons pass loose calo ID, loose track ID, very loose isolation cut, and DZ cut which is the z distance between the track and primary vertex) and the single electron trigger "HLT_Ele27_WPTight_Gsf" (which selects one electron with $p_T > 27$ GeV and passing tight working point) are used in ee channel. For μμ channel, the double muon trigger "HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL" (which selects two global muons (see definition in Section 4.2 and one of them with $p_T > 17$ GeV and another with $p_T > 8$ GeV, besides both muons pass very very loose track isolation cut) is used in MC and in data from run B to run G. For data run H the DZ cut is added for this trigger, due to the increase of the instantaneous luminosity delivered by the LHC in run H. A similar trigger "HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL" (replace "Mu8" with "TkMu8" which is tracker muon (see definition in Section 4.2) with $p_T > 8$ GeV) is used. Moreover, the single muon triggers "HLT_IsoMu24" (isolated global muon with $p_T > 24$ GeV) and "HLT_IsoTkMu24" (isolated tracker muon with $p_T > 24$ GeV) are also used.

We complement the partially inefficient dilepton triggers by single lepton triggers (by using $t\bar{t}$ and tW MC samples, it is found that by adding single lepton triggers, the trigger efficiency is increased by around 5%.) and remove the overlap between two primary data sets by vetoing events that fired a dilepton trigger in the single lepton primary data sets.

Channel	Path	Dataset
ee	HLT_Ele23_Ele12_caloIdL_TrackIdL_IsoVL_DZ	data & MC
	HLT_Ele27_WPTight_Gsf	data & MC
μμ	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 6.3: Summary of the signal triggers.

Since the trigger information is available in simulated samples, we require MC samples to fire the trigger and finally apply corrections for any data/MC disagreement using scale factors. The TOP group recommended trigger scale factors (described in [125]) are used. We also have estimated the trigger scale factors by ourself as a cross check, and it gave similar results (although they are not used in this analysis).

6.3 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 and one b-jet. The identifications for different object are presented in the following.

6.3.1 Lepton selection

6.3.1.1 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 Physics Object Group (POG). Furthermore, they are required to be particle-flow (see Section 4.5) 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 summarized in the following and correspond to the so-called Tight muon identification [126].

- $p_T > 20$ 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 constrains 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 (PU) 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 (PV). The muon isolation is therefore defined as:

$$I^{\mu}_{rel} = \frac{1}{{p_T}^{\mu}} \quad (\quad \Sigma p_T(\text{ch-had from PV}) + max(0, \Sigma E_T(\text{neut-had}) + \Sigma E_T(\text{photon}) - 0.5\Sigma p_T(\text{ch-had from PU}))), \tag{6.1}$$

where "ch-had" means charged hadron and "neut-had" means neutral hadron. The factor 0.5 corresponds to an approximate average of neutral to charged particles and has been measured in jets in Ref. [127]. In our analysis the muon candidates must have $I^{\mu}_{rel} < 0.15$ to be considered as isolated [126]. Scale factors are used to correct for differences in the reconstruction, ID and Isolation efficiencies in data and MC. They are evaluated using the tag and probe technique, and both the scale factors and their uncertainty prescriptions are provided by the Muon POG [128]. In addition, muon energy scale and smearing is applied based on Rochester group recommendations [129].

6.3.1.2 Electron

Electron candidates are selected from the reconstructed gsf electrons with $p_T > 20$ GeV and $|\eta| < 2.4$ while the ECAL barrel-endcap gap is removed ($1.4442 < |\eta_{SC}| < 1.566$, the η_{SC} is the η of super-cluster (see Section 4.1) corresponding to the electron). Electrons need to pass the tight cut based POG recommended working point [130] which includes

	Barrel($\|\eta_{SC}\|$) <= 1.479	Endcap($\|\eta_{SC}\|$) > 1.479
full5×5 $\sigma_{I_\eta I_\eta}$ <	0.00998	0.0292
$\|\Delta\eta_{\text{Inseed}}\|$ <	0.00308	0.00605
$\|\Delta\Phi_{\text{In}}\|$ <	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<=	1	1
pass conversion veto	yes	yes
d0 <	0.05	0.1
dZ <	0.1	0.2

Table 6.4: The cuts for the electron identification in the barrel and endcap.

the requirement to pass the conversion veto. The Table 6.4 contains the cuts used for the tight electron identification and more details can be found in [130].

An Multivariate Analysis (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 [131]. In addition, electron energy scale and smearing is applied based on Egamma POG recommendations [132].

6.3.2 Jet selection

Particle candidates found by the PF algorithm are clustered into jets using the anti-k_T algorithm (see Section 4.3) with distance parameter R =0.4. 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. Level 1 corrections (which removes the energy coming from pile-up events), Level 2 and Level 3 MC-truth jet energy corrections as well as Level 2 and Level 3 residual corrections (which corrects the remaining small differences for jet response in data and MC, it is applied only for data) are applied using the latest set of Jet Energy Corrections (JECs). Jets in MC are smeared using the latest set of Jet Energy Resolutions (JERs). Corrected jets with $p_T > 30$ GeV and $|\eta| < 2.4$ are selected if they pass the loose jet identification criteria [133], 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 version 2 algorithm (CSVv2). The CSVv2 is based on secondary vertex and track-based lifetime information, it is an updated version of the CSV algorithm (see Section 4.3.1) by combining the variables with a neural network instead of a likelihood ratio. In this analysis, a jet is b-tagged when its CSVv2 value passes the medium working point (i.e. CSVv2 > 0.8484) [134].

6.3.3 Missing Transverse Energy

Missing transverse energy (MET, see the definition in Section 4.4) 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 [135].

In order to reduce the instrumental noise in the detector, MET filters are applied as are recommended by Jet-MET POG [136]. These filters are summarized 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)

6.3.4 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/2016dataReReco Efficiencies/tracking/Tracking_EfficienciesAndSF_BCDEFGH.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/2016dataReReco Efficiencies/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

6.3.5 Top p_T reweighting

In order to better describe the p_T distribution of the top quark in $t\bar{t}$ events, the top quark p_T spectrum simulated with POWHEG is reweighted to match the differential top quark p_T distribution at NNLO QCD accuracy and includes EW corrections calculated in Ref. [137].

6.4 Event Selection

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

6.4.1 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 6.3. At least 1 pair of opposite charge leptons with invariant mass > 20 GeV is required while the leading lepton should have $p_T > 25$ GeV. The first two selected leptons which are sorted due to the p_T should have the same flavor and opposite sign. If the two highest p_T leptons have different flavors (e.g. $e\mu$ event) or same sign, the 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 ($\geq$2-jets,2-tags).

6.4.2 Event selection (step 2)

In order to suppress the contribution of Drell-Yan events, we reject events with a dilepton invariant mass around the Z peak [76, 106]. In addition, a MET cut > 60 GeV is applied. In Figure 6.1, one can see that the Data/MC agreements for MET distributions in ee and $\mu\mu$ channels are not good. We have investigated this problem (see Appendix B.5). Because of that the normalization of the DY background simulation is estimated from data which is described in Section 6.5.1. The final event selection criteria are shown in Table 6.5.

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

Table 6.5: Event selection criteria for ee and $\mu\mu$ channels

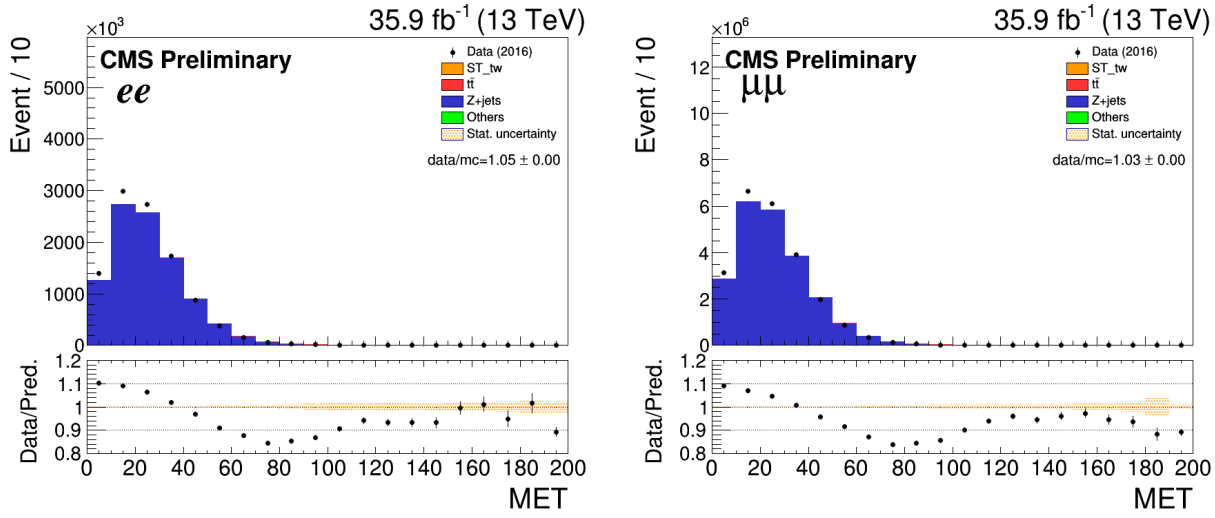

Figure 6.1: The distributions of MET for ee (left) and μμ (right) channels after step 1 selection. All backgrounds are estimated from MC simulations.

6.5 Background Predictions

6.5.1 Prompt Background

The background from processes giving two prompt leptons is taken from MC samples and normalized to the luminosity. It consists mostly of events from $t\bar{t}$ production, Drell-Yan, and WW productions. 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 [138, 139, 140, 141], 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 eμ 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:

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

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:

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

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 $N^{ee_{in,loose}}$ 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:

$$k_{ee} = \sqrt{\frac{N^{ee_{in,loose}}}{N^{\mu\mu_{in,loose}}}}$$

$$k_{\mu\mu} = \sqrt{\frac{N^{\mu\mu_{in,loose}}}{N^{ee_{in,loose}}}}$$

The values of k_{ll} for different njet-nbjet regions are shown in Table 6.6. We use explicitly k_{ll} for different njet-nbjet regions.

Channel	all	1jet,1tag	2jet,1tag	>=2jet,2tag
$N^{ee_{in,loose}}$ (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

Table 6.6: The values of k_{ll} for different njet-nbjet regions. Errors are statistical uncertainties only.

The global scaling factors

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

are determined. The results and scaling factors are summarized in Table 6.7.

	All		1jet,1tag		2jet,1tag		$>= 2jet, 2tag$	
	ee	μμ	ee	μμ	ee	μμ	ee	μμ
$N_{in}^{ll,Z+jets\ MC}$	243878.3	562506.9	2712.3	6490.7	1550.6	3734.2	306.8	712.7
$N_{out}^{ll,Z+jets\ MC}$	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
$N_{in}^{ll,data}$	220435	501781	3805	8291	3735	7550	2493	5180
$N_{in}^{e\mu,data}$	34322	34322	4453	4453	6230	6230	6259	6259
$N_{out}^{ll,Z+jets\ data}$	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 6.7: Data-driven Z+jets background estimation in the μμ and ee channels after the "$Step1 + Step2$" selection requirements. Errors are statistical uncertainties only.

6.5.2 Fake Background

Another source of events with misidentified leptons is the Wγ process, where the W decays to an electron and a neutrino 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, coming mostly from W+jets process and QCD. The method which is used to estimate the jet background 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 the same sign control region. The contributions of the prompt backgrounds are subtracted from data in the same sign region using MC samples to find the jets contribution in the opposite sign region (signal region).

6.6 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 6.8 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 6.2-6.7. The definitions of H_T^{syst}, p_T^{syst}, and MT^{syst} in Figure 6.6 are $H_T^{syst} = \sum {p_T}^{jet} + \sum {p_T}^{lepton}$, ${p_T}^{syst} = \sum \overrightarrow{{p_T}^{jet}} + \sum \overrightarrow{{p_T}^{lepton}}$, $MT^{syst} = \sqrt{(H_T^{syst})^2 - ({p_T}^{syst})^2}$. In Figure 6.8, 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 6.8 and the plots of Figures 6.2-6.8 are explained in Section 6.8.

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

Table 6.8: Numbers of expected events from tW, $t\bar{t}$, and DY productions, 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.

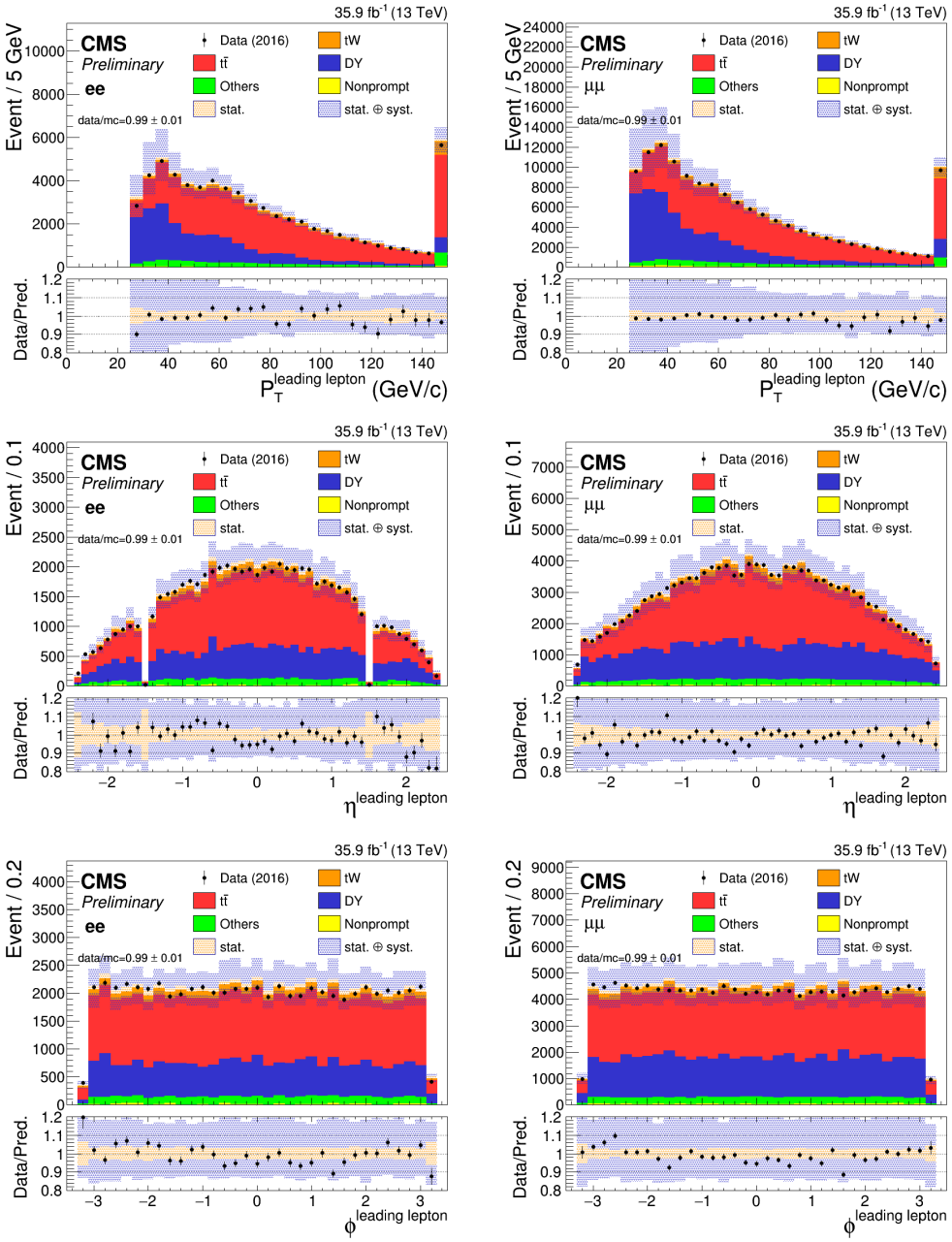

Figure 6.2: The distributions of p_T (top), η (middle) and ϕ (bottom) of leading lepton for ee (left) and $\mu\mu$ (right) channels after step 2 (full selection).

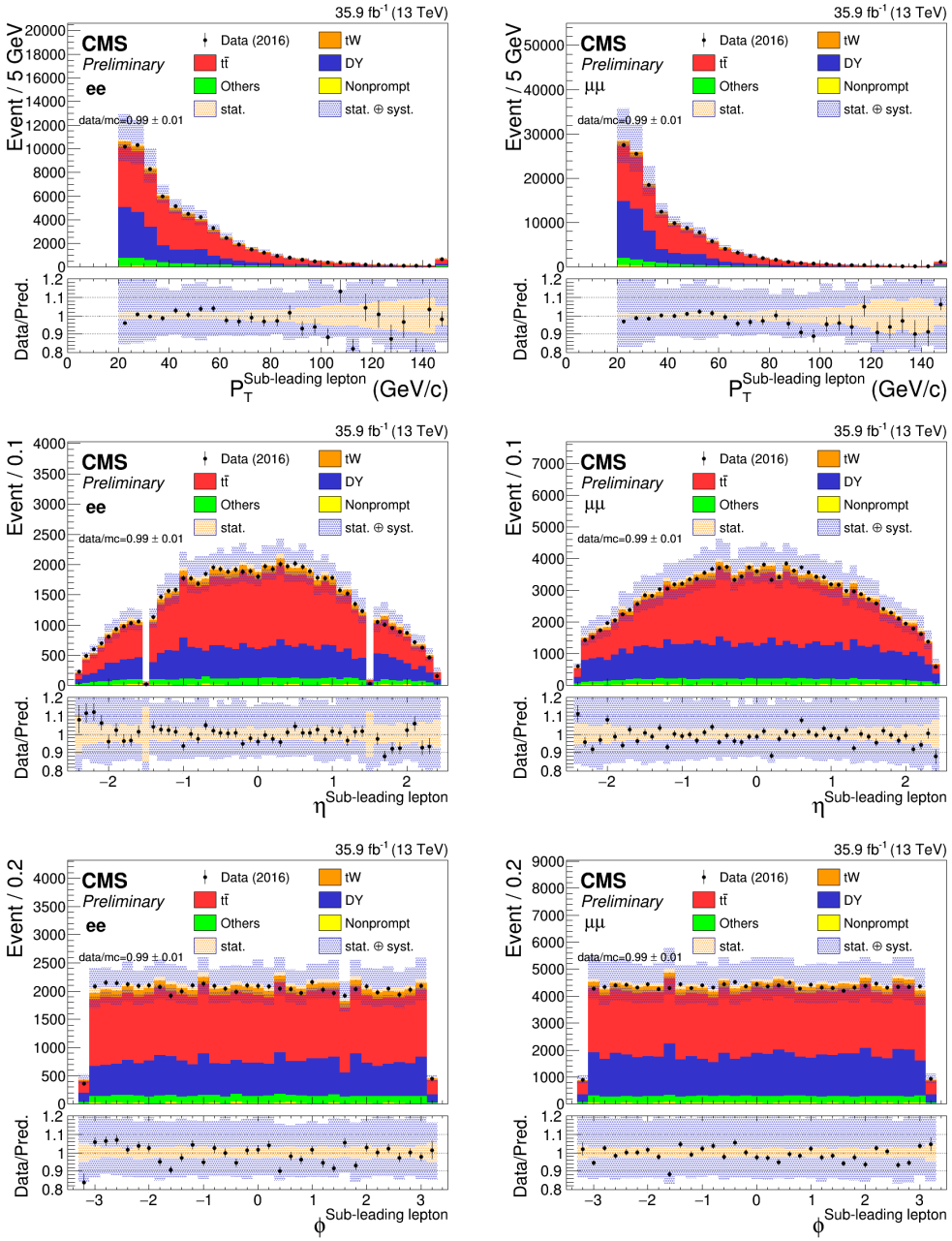

Figure 6.3: The distributions of p_T (top), η (middle) and ϕ (bottom) of sub-leading lepton for ee (left) and $\mu\mu$ (right) channels after step 2 (full selection).

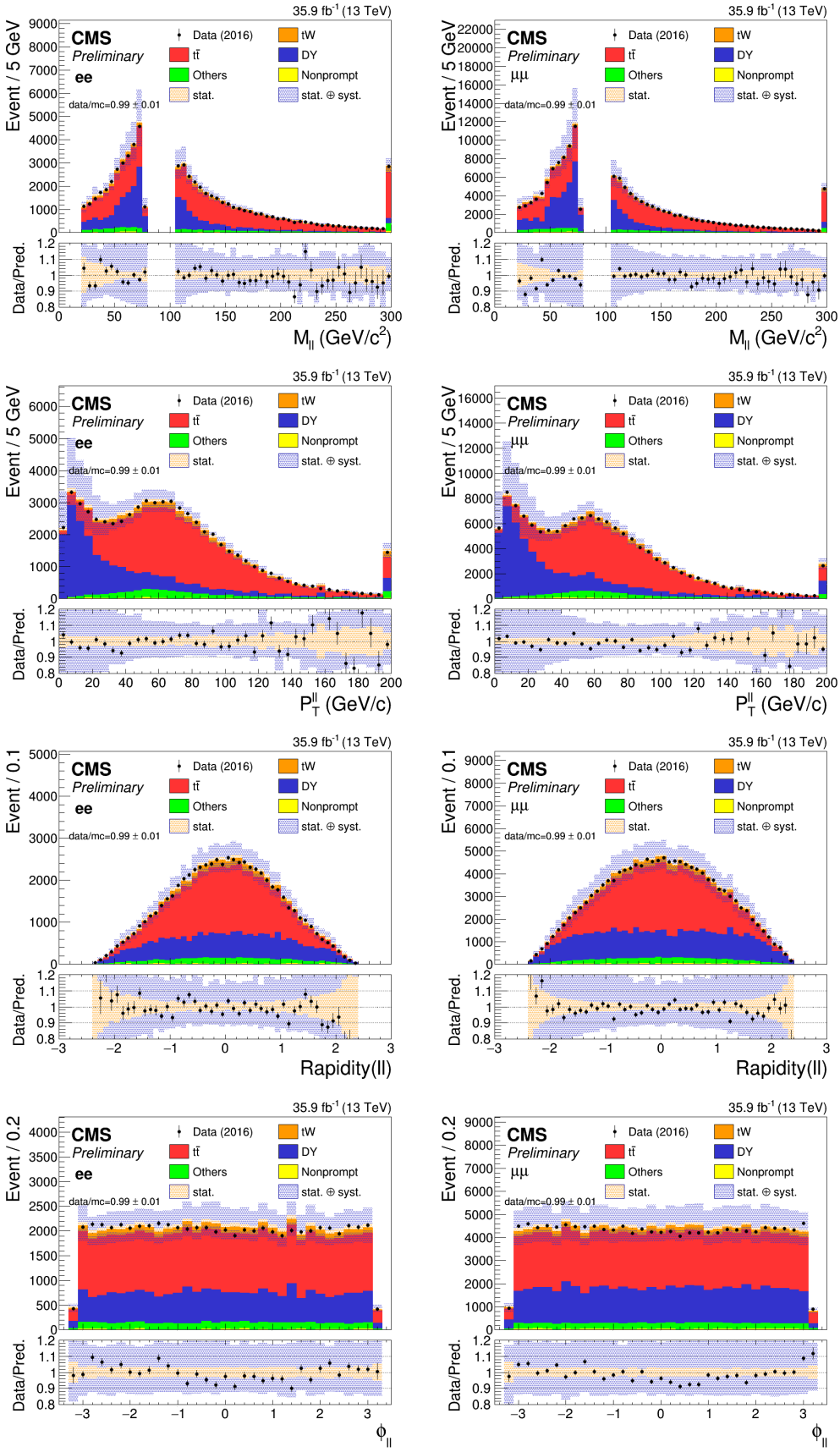

Figure 6.4: The distributions of invariant mass of two leptons (first row), p_T of two leptons (second row), Rapidity of two leptons (third row) and ϕ of two leptons (last row) for ee (left) and $\mu\mu$ (right) channels after step 2 (full selection).

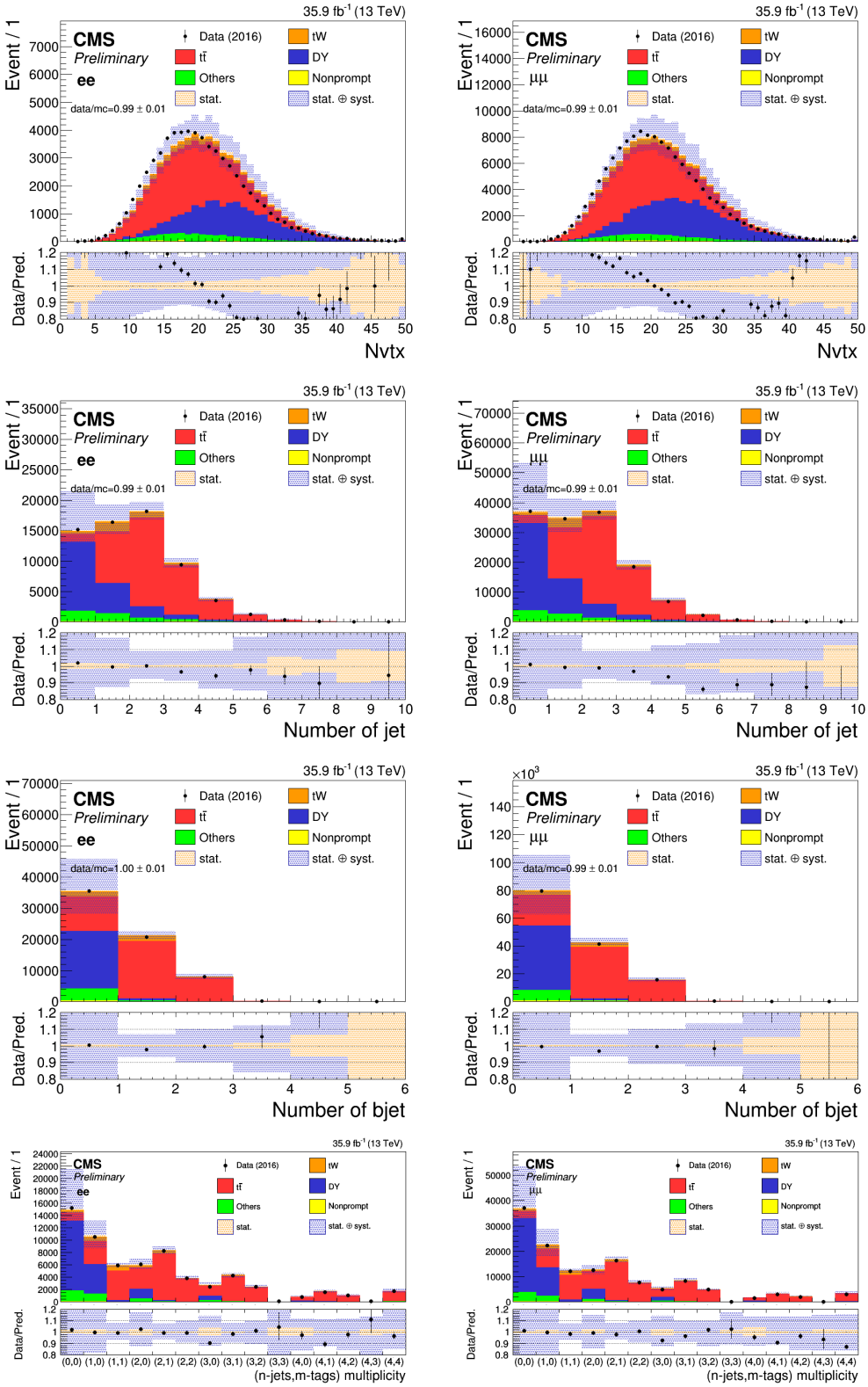

Figure 6.5: The distributions of number of vertices (first row), number of jet (second row), number of b jets (third row) and number of jet-bjets (last row) for ee (left) and μμ (right) channels after step 2 (full selection).

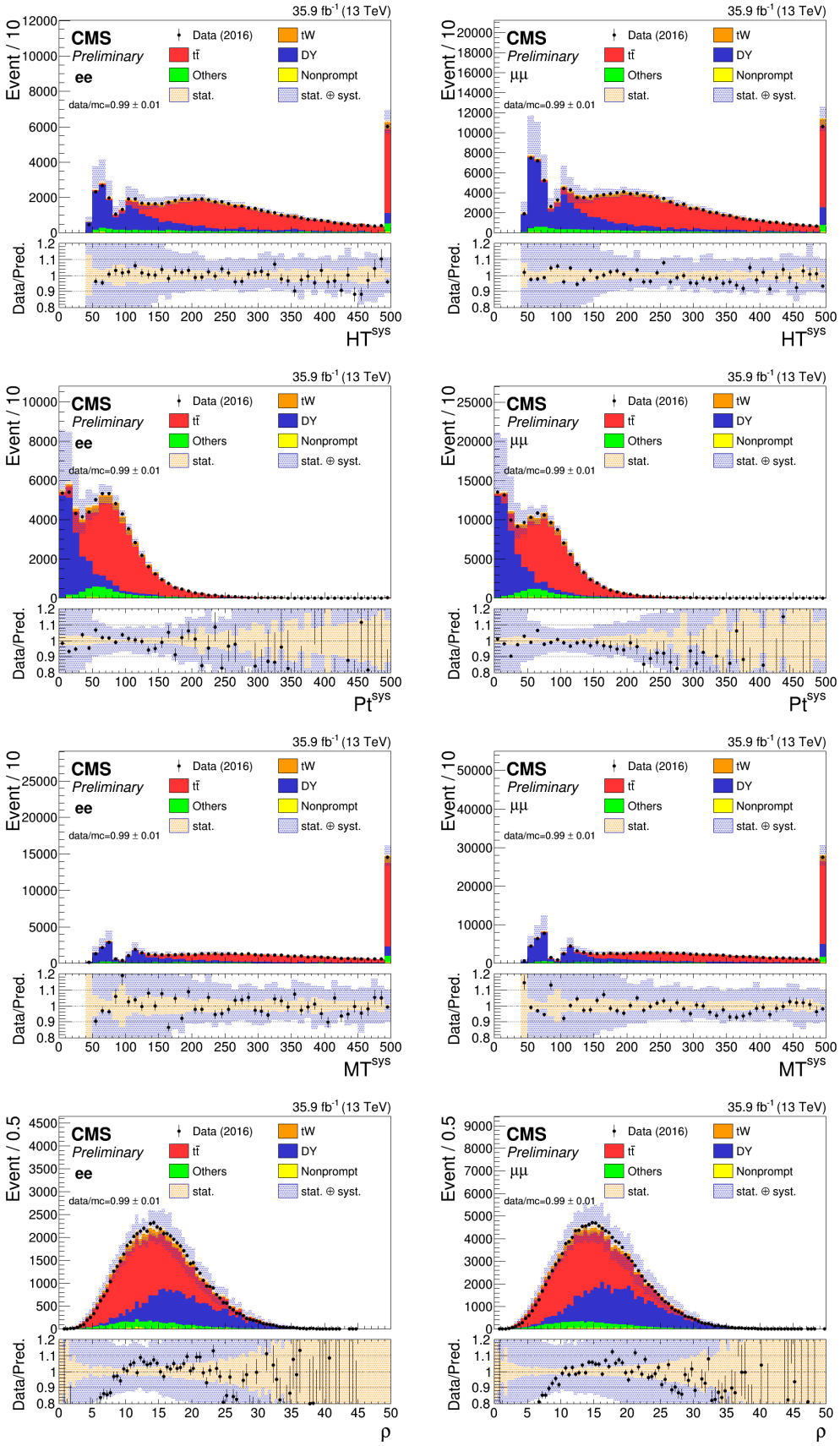

Figure 6.6: The distributions of HT^{syst} (first row), $p_T{}^{syst}$ (second row), MT^{syst} (third row) and ρ for ee (left) and $\mu\mu$ (right) channels after step 2 (full selection).

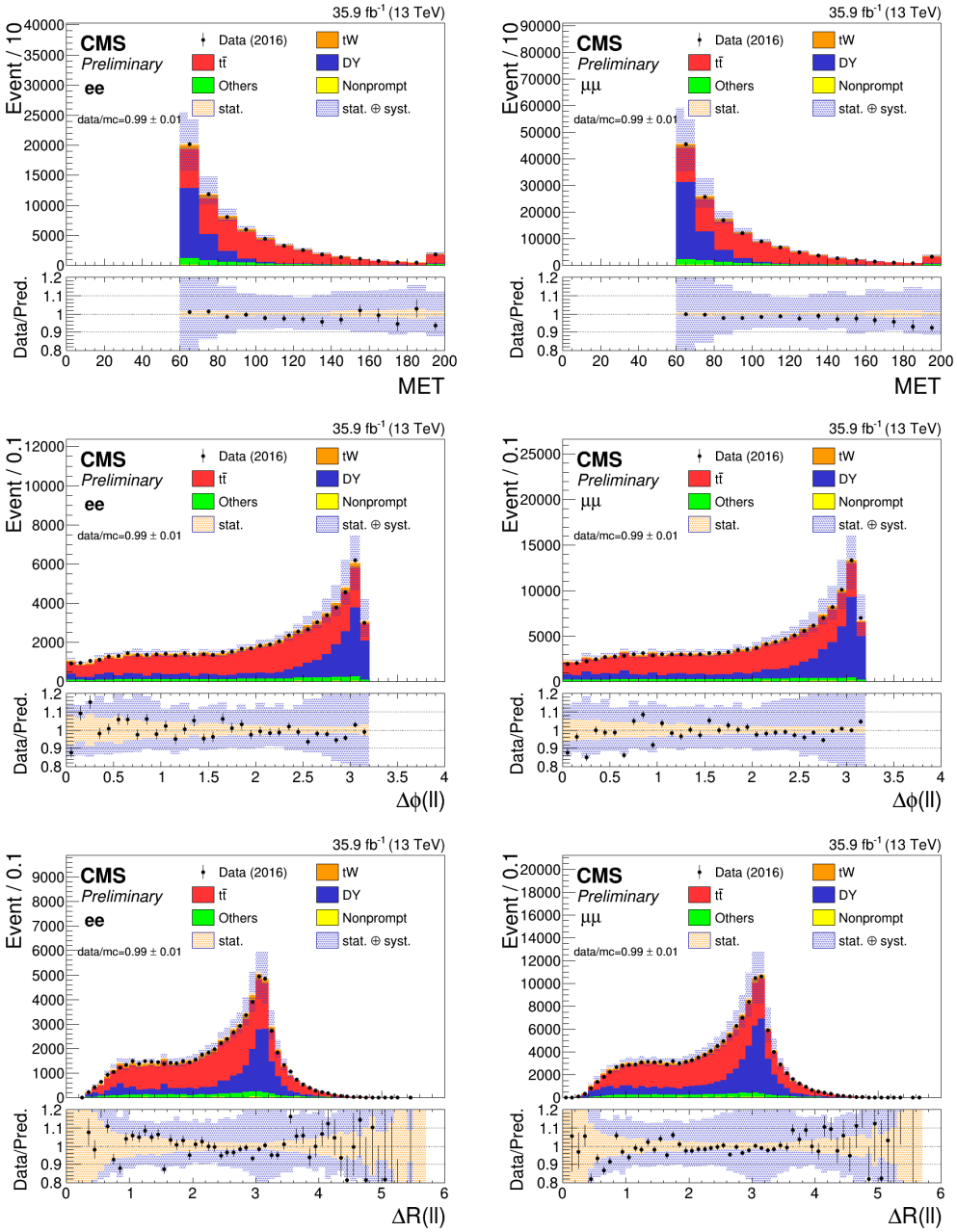

Figure 6.7: The distributions of MET (top row), $\Delta\phi$ between two leptons (middle row) and ΔR between two leptons (bottom row) for ee (left) and $\mu\mu$ (right) channels after step 2 (full selection).

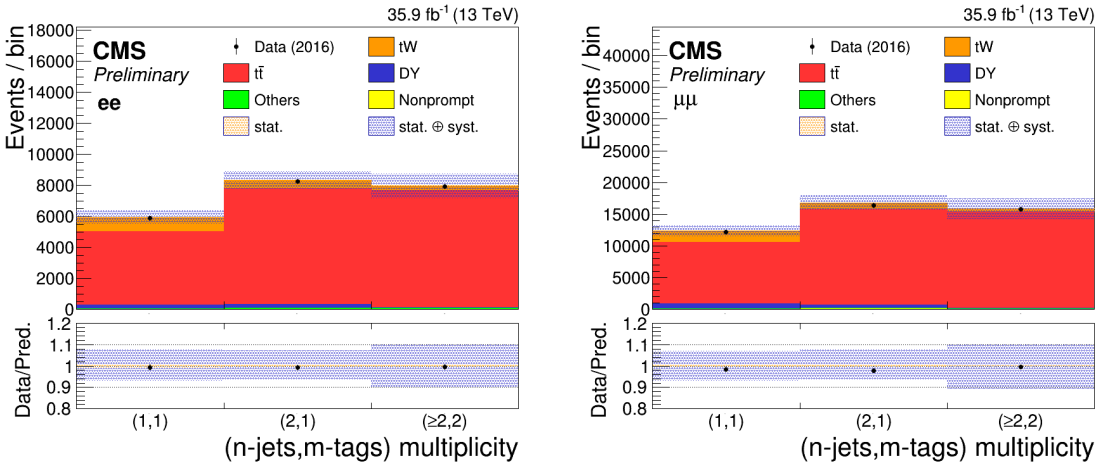

Figure 6.8: The observed numbers of events and SM background predictions in the search regions of the analysis for the ee (left), μμ (right) channels. The hatched band corresponds to the quadratic sum of statistical and systematic uncertainties in the event yield for the SM background predictions. The 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.

6.7 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 [142].

All the effective couplings introduced in Section 2.3 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. [85] 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 6.4, 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 [143].

The presence of the O_{uG} and O_{cG} operators changes the initial-state particle (see Figure 2.3), 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 summarized in Table 6.9. The various input variables for the NN introduced above are shown in Table 6.10.

The distributions of Multiple Layer Perceptron (MLP) input variables for signal and background events are shown from Figures 6.9 to 6.11. The MLP output for test and train samples are shown in Figure 6.12.

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
	μμ	Yield	Yield	-	Yield
$C_{\phi q}^{(3)}$, C_{tW}, C_{tG}	ee	NN_{11}	NN_{21}	-	Yield
	μμ	NN_{11}	NN_{21}	-	Yield
C_{uG}, C_{cG}	ee	-	-	NN_{FCNC}	-
	μμ	-	-	NN_{FCNC}	-

Table 6.9: Summary of the observables used for probing effective couplings in various (n-jets,m-tags) categories in the ee and μμ channels.

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(ℓ^{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			√
$\Delta\phi(\ell\ell,\text{jet}^{\text{leading}})$	Δφ 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			√

Table 6.10: 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.

6.7.1 Data/MC comparison for MVA input variables

The data and MC comparisons for MVA input variables are shown from Figure 6.13 to 6.14 for NN_{11}, and from Figure 6.15 to 6.16 for NN_{21} and from Figure 6.17 to 6.19 for NN_{FCNC}. Sources of systematic uncertainties which are considered in the plots are explained in Section 6.8. Considering the systematic uncertainties, data/MC are in good agreements and no obvious large mis-modelling is observed.

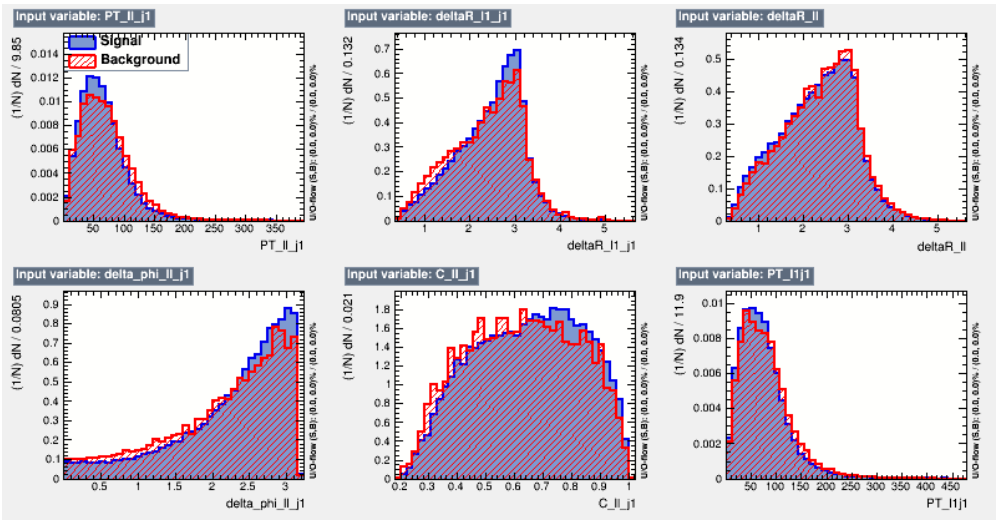

Figure 6.9: MLP input variable distributions for NN_{11}. Signal distributions are drawn in blue and the background distributions in red [144].

.

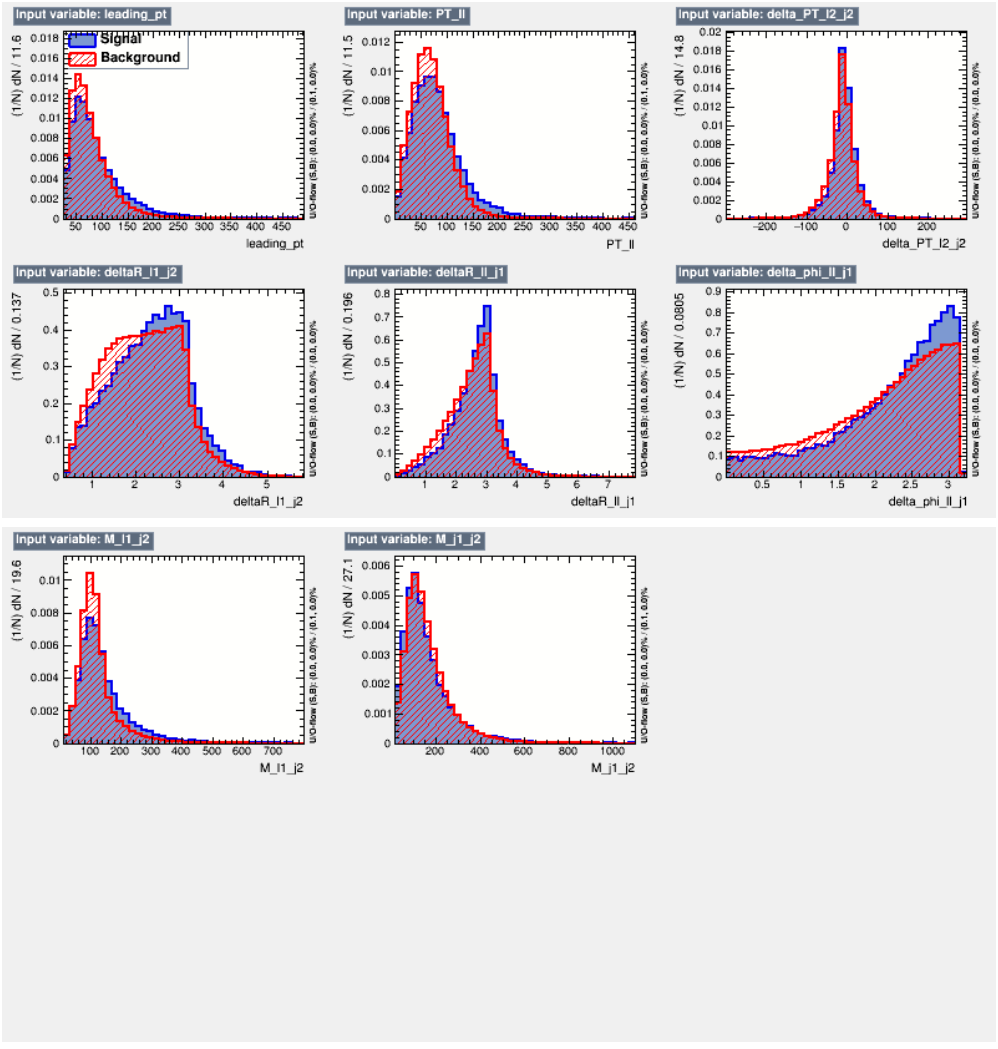

Figure 6.10: MLP input variable distributions for NN_{21}. Signal distributions are drawn in blue and the background distributions in red [144].

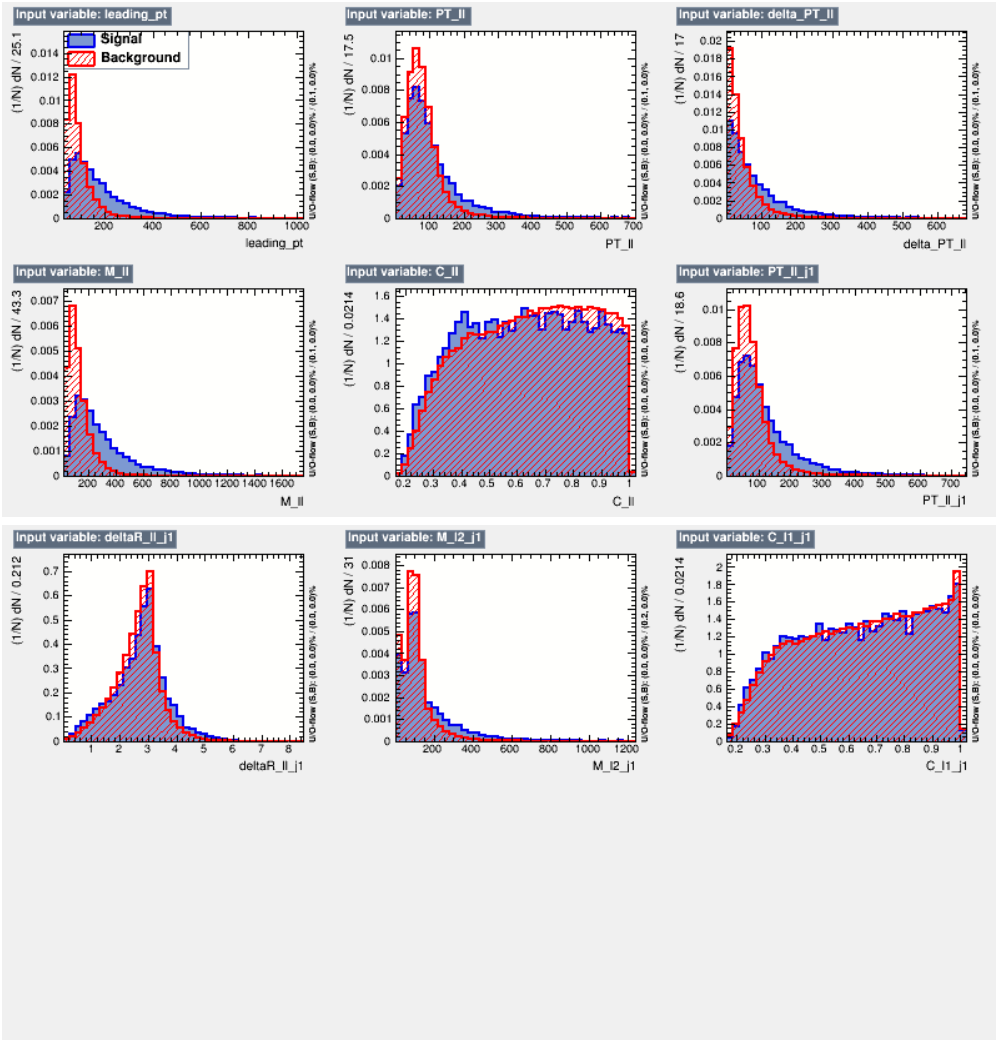

Figure 6.11: MLP input variable distributions for NN_{FCNC}. Signal distributions are drawn in blue and the background distributions in red [144].

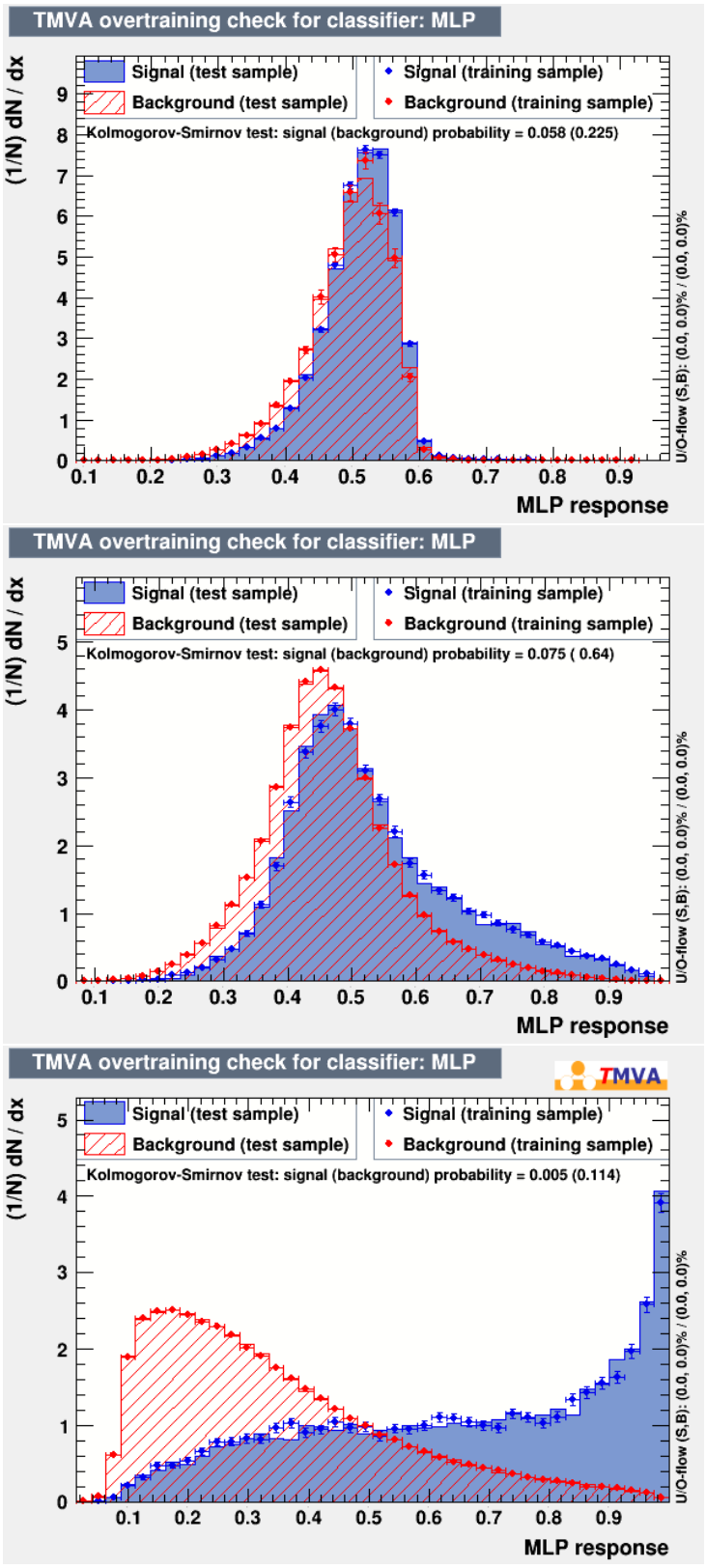

Figure 6.12: MLP output for test and train samples for NN_{11} (top), NN_{21} (middle) and NN_{FCNC} (bottom) [144].

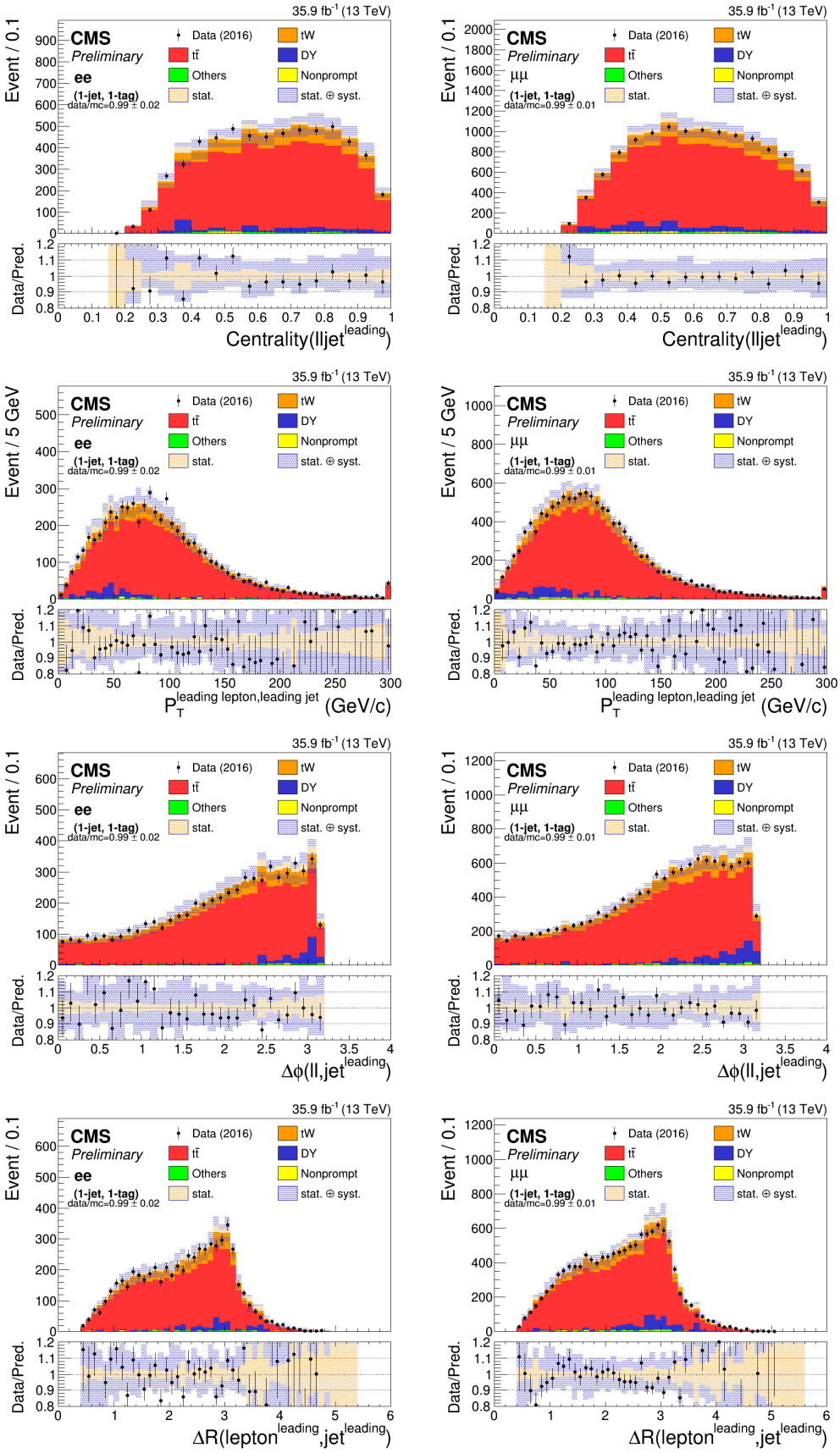

Figure 6.13: The distributions of variables used for MVA input for ee (left) and μμ (right) channels for NN_{11}.

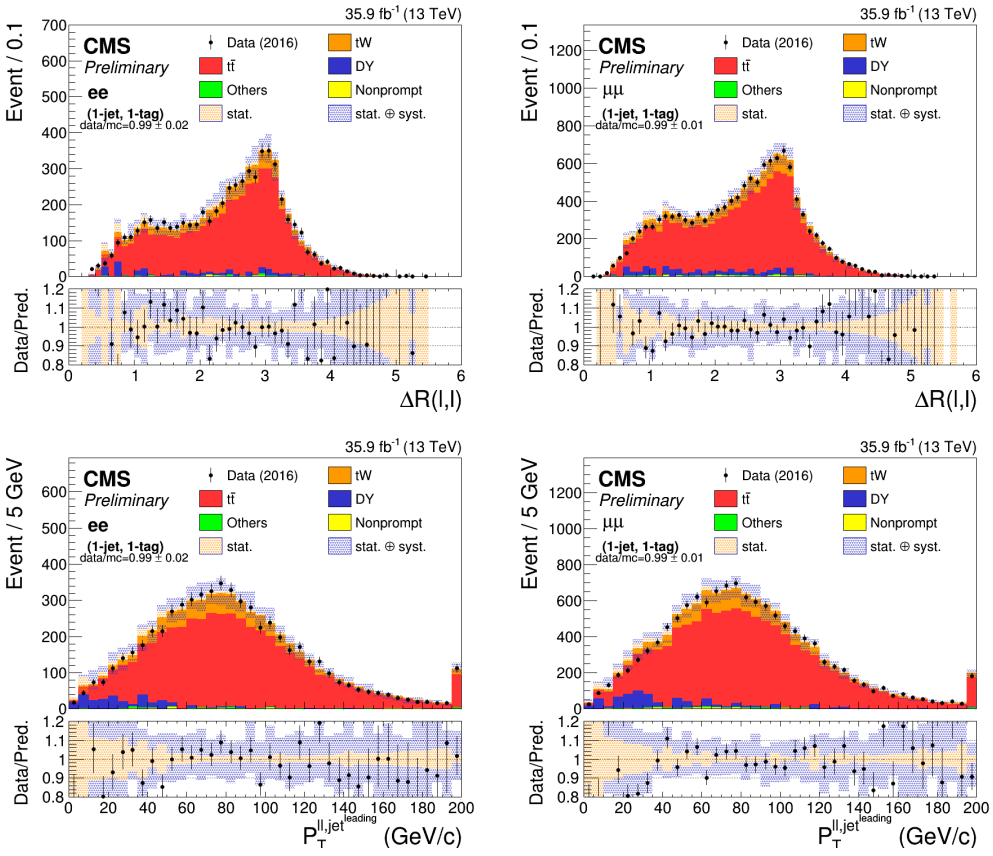

Figure 6.14: The distributions of variables used for MVA input for ee (left) and μμ (right) channels for NN_{11}.

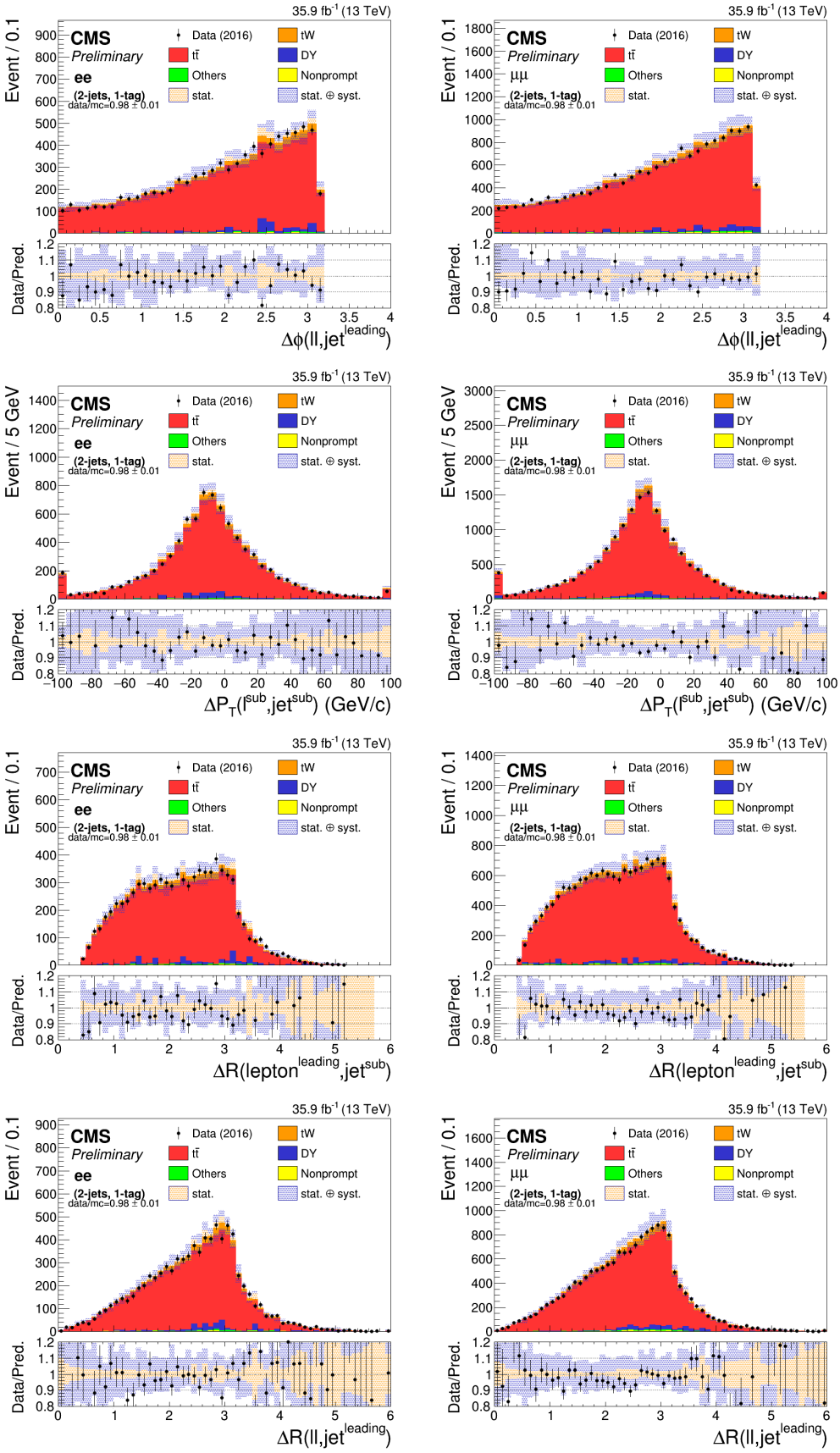

Figure 6.15: The distributions of variables used for MVA input for ee (left) and μμ (right) channels for NN_{21}.

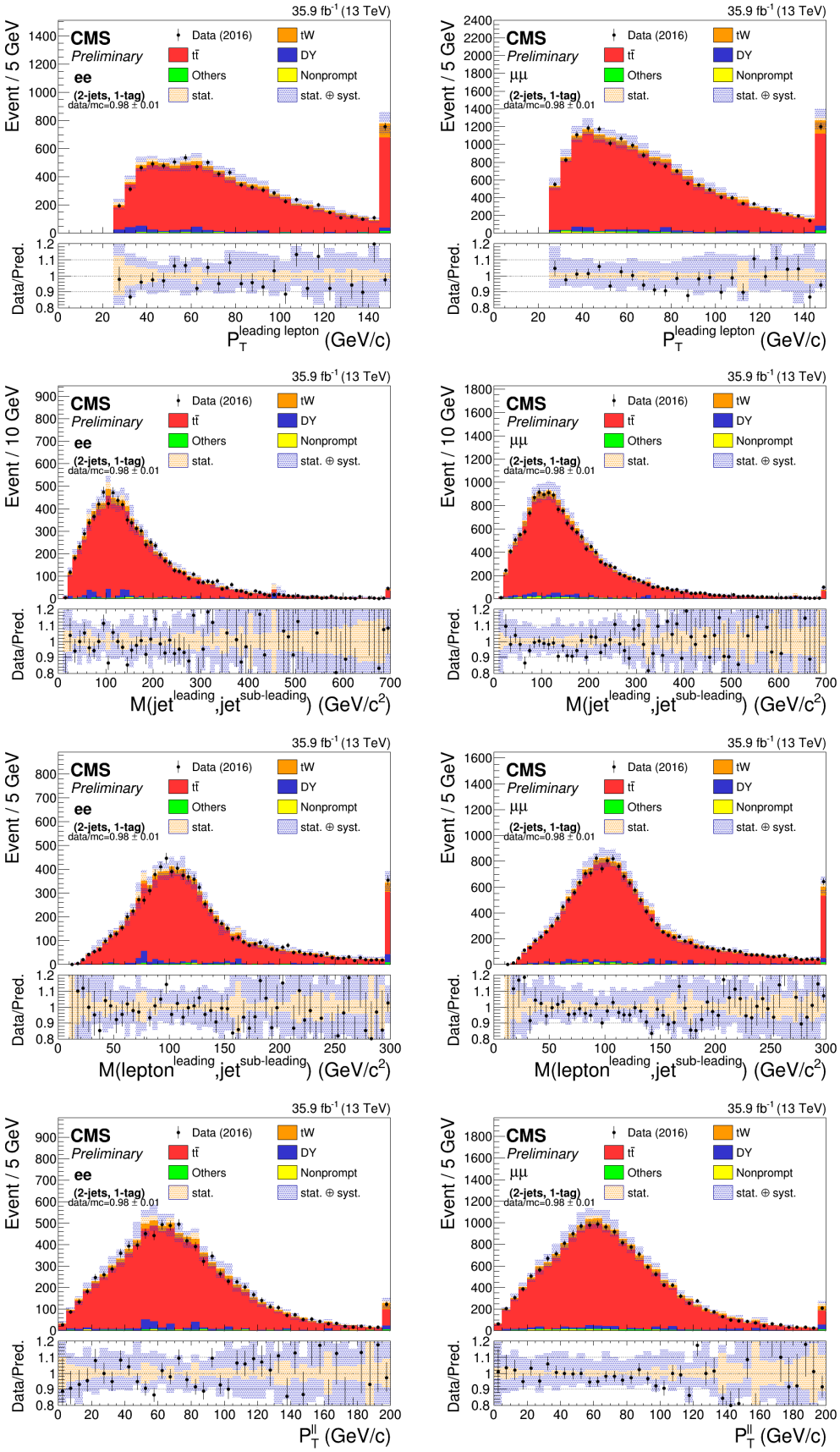

Figure 6.16: The distributions of variables used for MVA input for ee (left) and μμ (right) channels for NN_{21}.

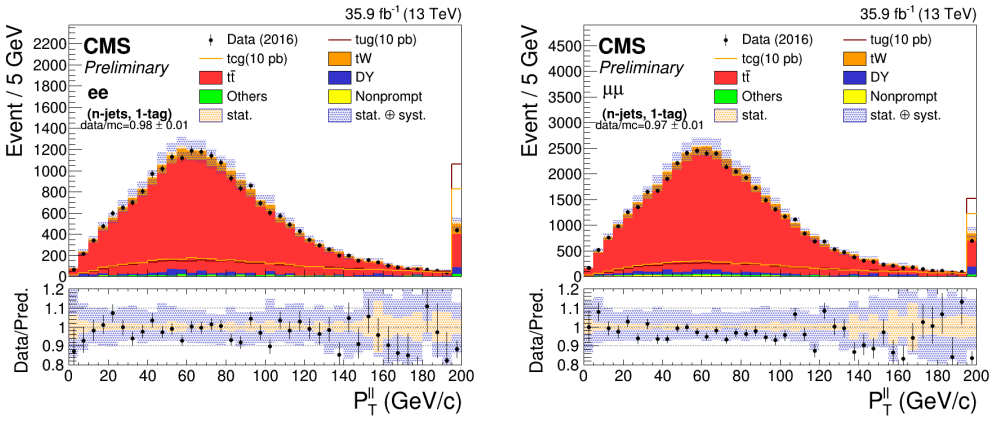

Figure 6.17: The distributions of variables used for MVA input for FCNC study for ee (left) and μμ (right) channels for $\mathrm{NN_{FCNC}}$.

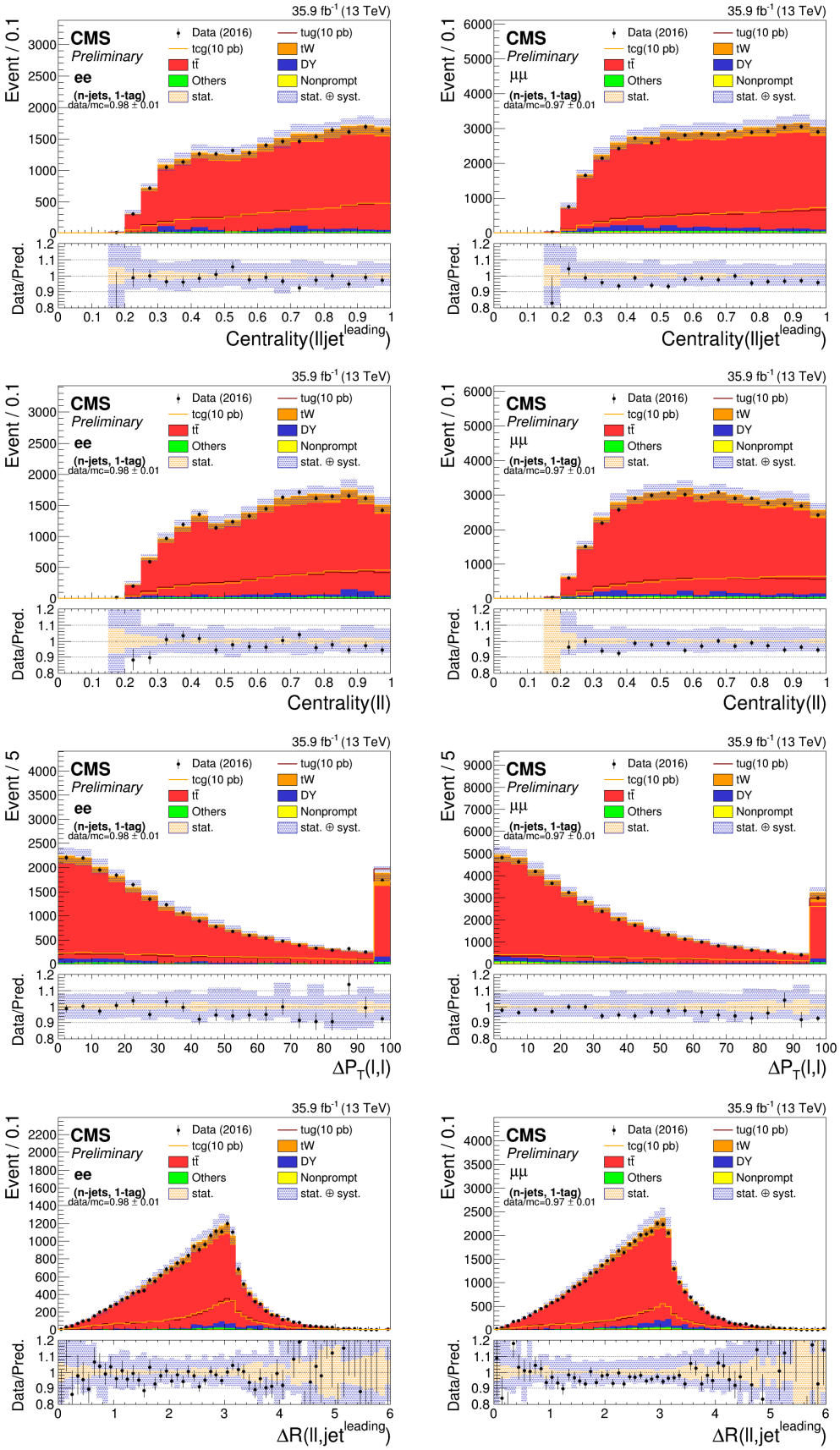

Figure 6.18: The distributions of variables used for MVA input for FCNC study for ee (left) and μμ (right) channels for $\mathrm{NN_{FCNC}}$.

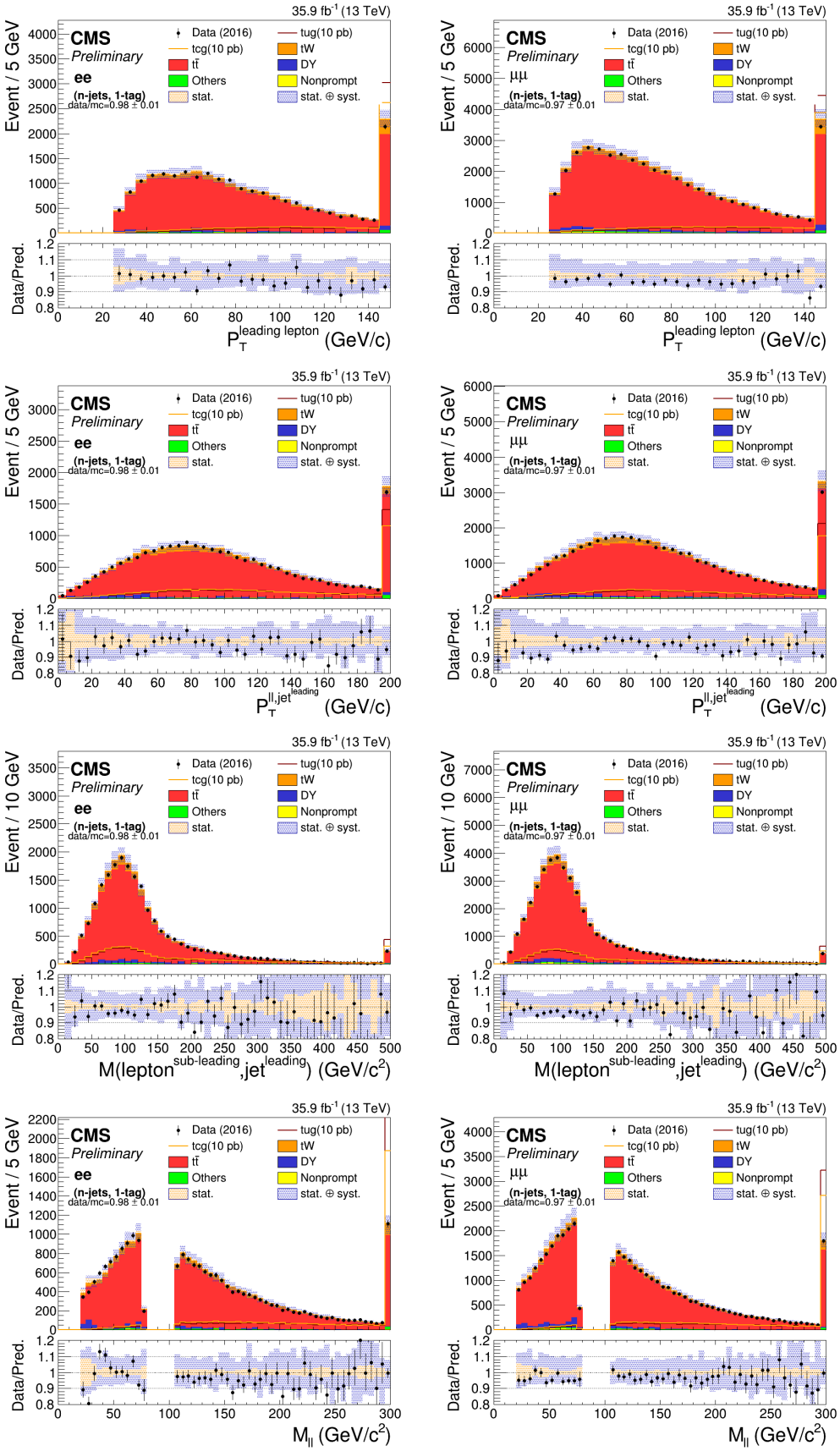

Figure 6.19: The distributions of variables used for MVA input for FCNC study for ee (left) and μμ (right) channels for $\mathrm{NN_{FCNC}}$.

6.8 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 inputs for generating and simulating $t\bar{t}$ and tW processes. In Table 6.11, 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 fluctuates up while the other fluctuates 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 6.11).
- **Parton Distribution Functions (PDF) uncertainty:** The magnitude of the uncertainties related to the parton distribution functions and the variation of the strong coupling constant for each simulated signal process is obtained using the replicas of the NNPDF 3.0 set. This source of uncertainty is only considered for $t\bar{t}$.
- **Top mass:** The most recent measurement of the top quark mass by CMS yields a total uncertainty of $\pm$ 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 [145]. 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 (UE):** It is known when two protons have a collision, the quarks or gluons in the two protons give hard interactions and the remnants of the two protons goes away. The particles from the remnants of the protons are form a UE. However, the QCD process in the remnants of the protons can not be derived from perturbation theory, in order to simulate UE, the PYTHIA parameters required to be tuned according to experiment results. The uncertainty from UE simulation is derived from varying the default parameters in the tune "CUETP8M2T4" according to their uncertainties. This source of uncertainty is only considered for $t\bar{t}$.
- **Color reconnection (CR):** As we know, the proton is colorless. After the collision of two protons, the full event is also color. The color reconnection in the remnants of the two collided protons needs to be modeled. The uncertainty from CR is considered

by varying the CR model with respect to the default using alternatives including the resonant decay products in possible reconnections to the UE. The default simulation (Multiple Parton Interaction-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 6.11: MC samples used for systematic study.

In Figures 6.20-6.21, the relative effect of each source of $t\bar{t}$ and tW modeling uncertainties for the MLP output distributions are shown. Relative uncertainty 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.

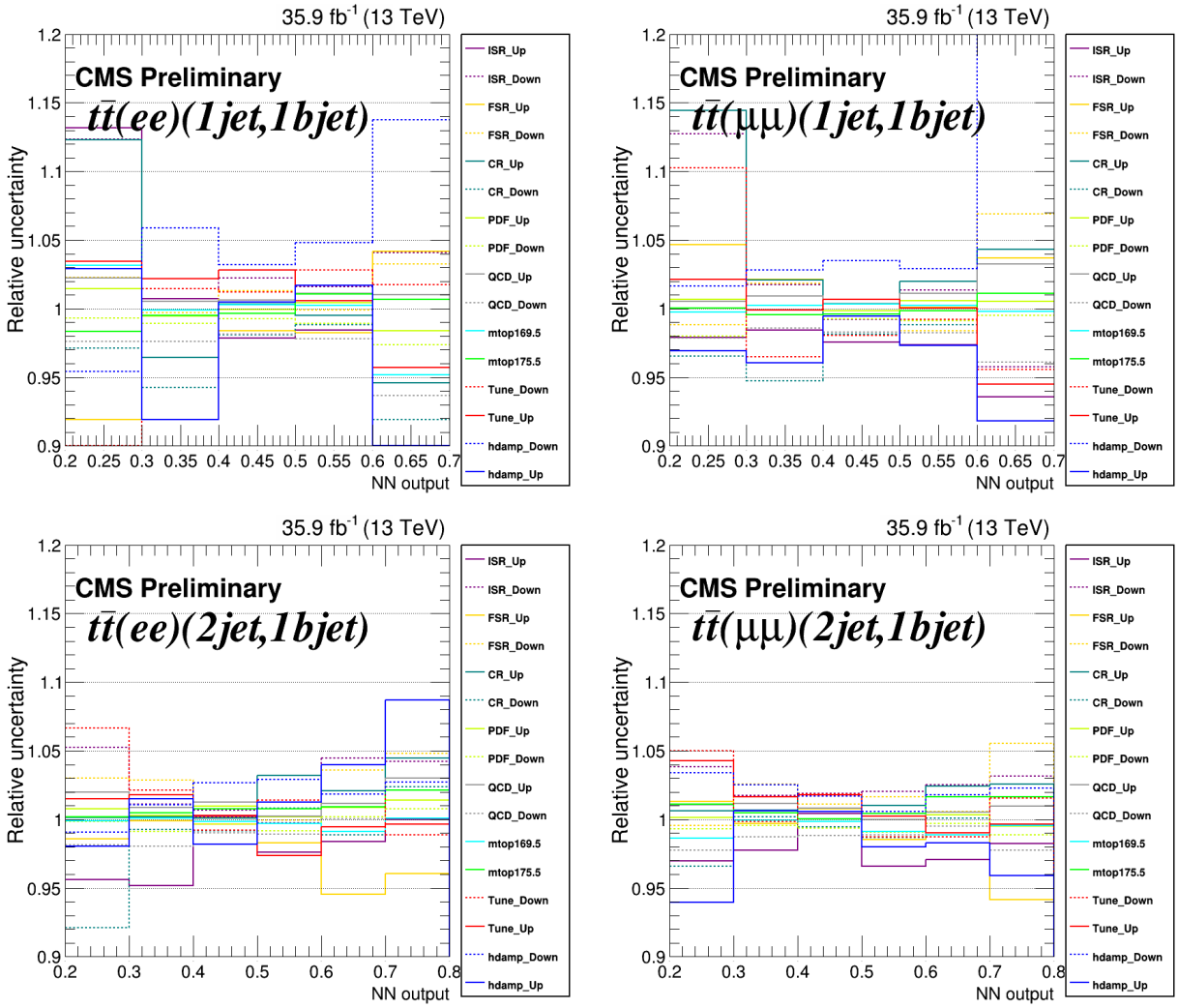

Figure 6.20: The relative effect of the $t\bar{t}$ modeling uncertainties in MLP distribution for different (1jet,1b-jet) region (top row), (2jet,1b-jet) region (bottom row) for ee channel (left column) and μμ channel (right column).

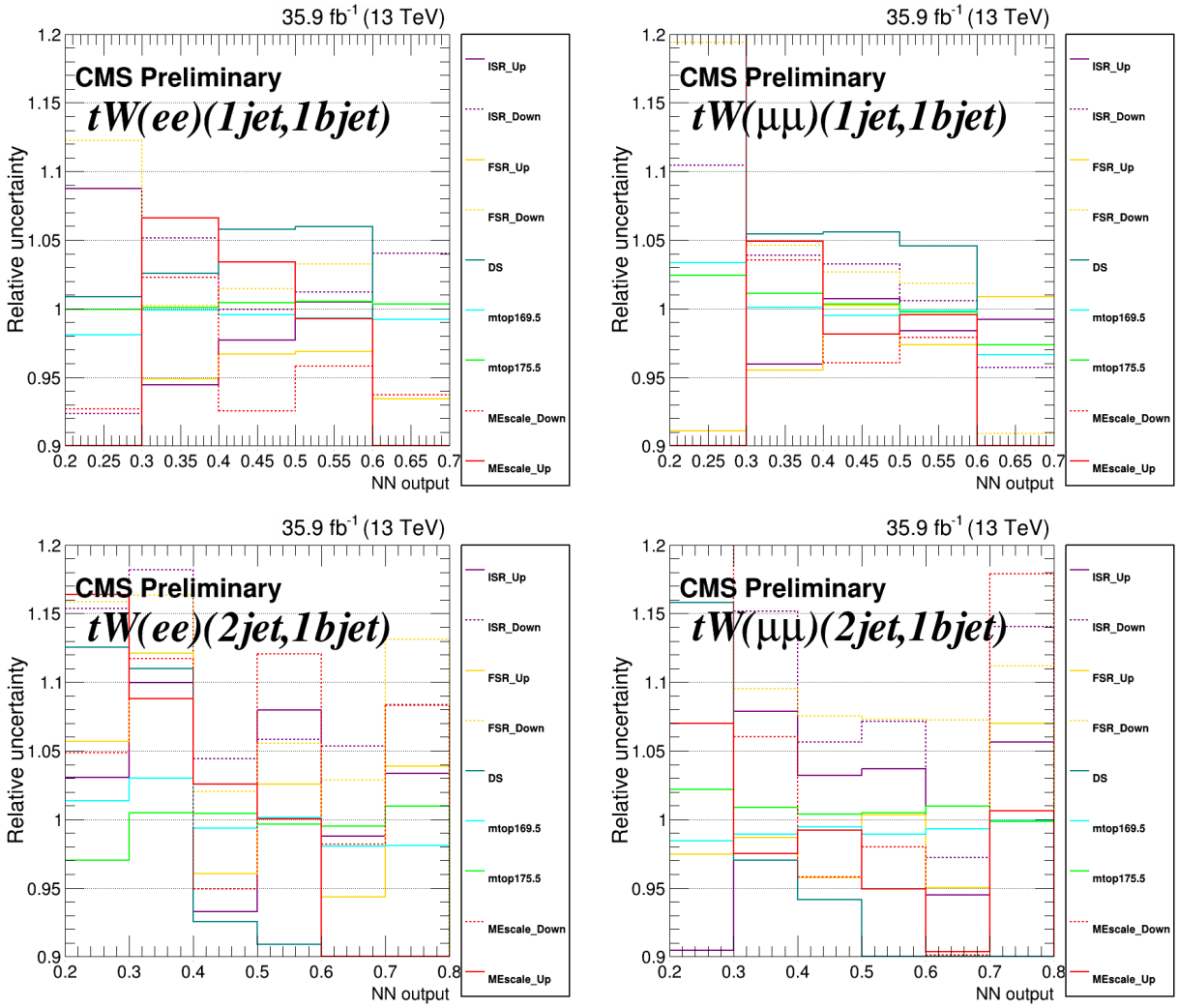

Figure 6.21: The relative effect of the tW modeling uncertainties in MLP distribution for different (1jet,1b-jet) region (top row), (2jet,1b-jet) region (bottom row) for ee channel (left column) and μμ channel (right column).

• **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 are provided officially by JET/MET POG in the recommended global tag [146]. 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:** As the MET is made from PF candidates but not all energies make PF candidates. So the MET will be varied by adding unclustered energy which is considered as an uncertainty from unclustered energy.

• **Trigger scale factor:** Uncertainties 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 [147].

• **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 [148].

• $t\bar{t}$ **normalization:** Uncertainty on $t\bar{t}$ normalization is considered to be 5% [149] for $O^{(3)}_{\phi q}$, 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 μμ channels (data-driven normalization) is considred 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 is considered to be 50%.

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

• **PDF uncertainty:** The magnitude of the uncertainties related to the PDF 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 renomalization and factorization scale variation. The largest deviation from the nominal value is taken as QCD scale error.

• **Parton shower QCD scale uncertainty:** The scales of the initial (ISR) and final (FSR) state showers are varied up and down by a factor of two with respect to the FCNC tW simulated signal samples. MC samples used to estimate the parton shower QCD scale uncertainties are summarized in Table 6.12.

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 6.12: Systematic samples for FCNC signal study.

In Figure 6.22, the effects of shape dependent uncertainties except $t\bar{t}$ and tW modeling are shown.

The relative effects of the all uncertainties on the total normalization are summarized in Tables 6.13 and 6.14 for ee and $\mu\mu$ channel respectively.

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 6.13: All systematic uncertainty effects for ee channel.

All	All	1jet, 1bjet	2jet, 1bjet	>= 2jet, 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%

Table 6.14: All systematic uncertainty effects for μμ channel.

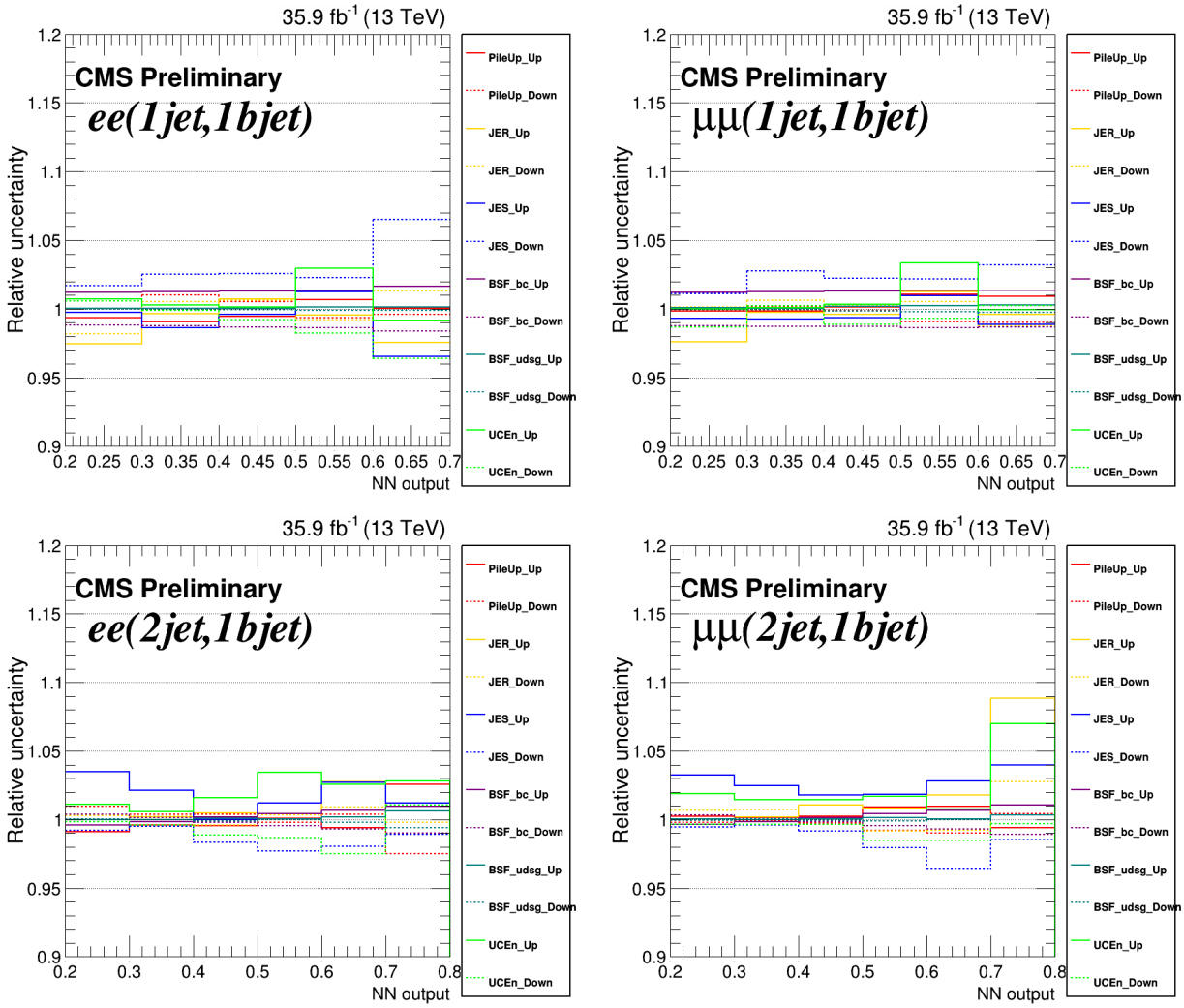

Figure 6.22: The relative effect of the shape dependent uncertainties in MLP distribution for different (1jet,1b-jet) region (top row), (2jet,1b-jet) region (bottom row) for ee channel (left column) and μμ channel (right column).

6.9 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 B.6) which makes sure we have understood 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 2.3 in the universal FeynRules output format [150] through the Feynrules package [151] and use the MadGraph5_aMC@NLO event generator [152, 153] 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

$$\sigma = \sigma_{SM} + C_i\sigma_i^{(1)} + C_i^2\sigma_i^{(2)}, \tag{6.2}$$

where the C_i's are effective couplings introduced in Equation 2.3. 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 $\sigma_{SM}^{t\bar{t}} = 832^{+20}_{-29}(scales) \pm 35(PDF + \alpha_s)$ 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 productions, respectively [154, 155]. The scales reflect 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 6.15.

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 [156]	-	-
	$\sigma_i^{(2)-LO}$	102.3	-	-	16.4	-	-
	$\sigma_i^{(2)-NLO}/\sigma_i^{(2)-LO}$	-	-	-	1.44 [156]	-	-
tW	$\sigma_i^{(1)-LO}$	-	6.7	-4.5	3.3	0	0
	$\sigma_i^{(1)-NLO}/\sigma_i^{(1)-LO}$	-	1.32 [157]	1.27 [157]	1.27 [157]	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 [157]	1.18 [157]	1.06 [157]	1.27 [83]	1.27 [83]

Table 6.15: 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.

6.9.1 Limit setting procedure

For those operators which interfere with the SM, C_G-C_{tG}-$C_{\phi q}^{(3)}$-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 6.2 in which σ_{SM}, $\sigma_i^{(1)}$ and $\sigma_i^{(2)}$ are fixed parameters (see Table 6.15) 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, uncertainty due to PDF is considered. Results of the mentioned variations are only shown for observed limits for comparison to the nominal results. In Table 6.16, nominal values for $\sigma_i^{(1)}$ and $\sigma_i^{(2)}$ are shown together with errors.

Channel	σ_{SM} (scale unc.) (PDF+α_s unc.)
$t\bar{t}$	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.)
$t\bar{t}$	202.83 C_{tG} (+24.54,-26.98) , 31.9 C_G (+8.1,-6.9)
tW	8.844 $C_{\phi q}^{(3)}$ (+0,-0) ,-5.65 C_{tW} (+0.08317,-0.061846) , 4.223 C_{tG} (+0.0294,-0.0398)

Channel	$\sigma_i^{(2)}$ (scale unc.)
$t\bar{t}$	23.545 ${C_{tG}}^2$ (+0,-0) , 102.3 ${C_G}^2$ (+22.7,-15.3)
tW	0.275 $C_{\phi q}^{(3)^2}$ (+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)

Table 6.16: 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.

Note: Before going the limit results it is worth to mention that all these studies are also performed for eμ final state. Therefore results from eμ channel are shown for comparison, and the results for ee, eμ and μμ channels combined are also presented.

6.9.2 Exclusion limits on C_G effective coupling

It was discussed in Section 2.3 that the operator O_G 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 yields in (1-jet,1-tag), (2-jet,1-tag) and (≥2-jet,2-tag) categories for ee and μμ channels. The results for ee, μμ and eμ (a analogous analysis but using eμ final state) individual channels and all three channels combined are listed in Table 6.17. The result of the combined likelihood scan of the C_G coupling is shown in Figure 6.23.

	Regions	Best fit exp./obs.	68% exp./obs.limit	95% exp./obs.limit
C_G	ee yield (1j1t), (2j1t), (≥2j,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]
	μμ yield (1j1t), (2j1t), (≥2j,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]
	eμ yield (1j0t), (1j1t), (2j1t), (≥2j,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 6.17: Summary of allowed 68% CL and 95% CL intervals on C_G effective coupling obtained in ee, eμ, μμ and combined channels ($\Lambda = 1$ TeV).

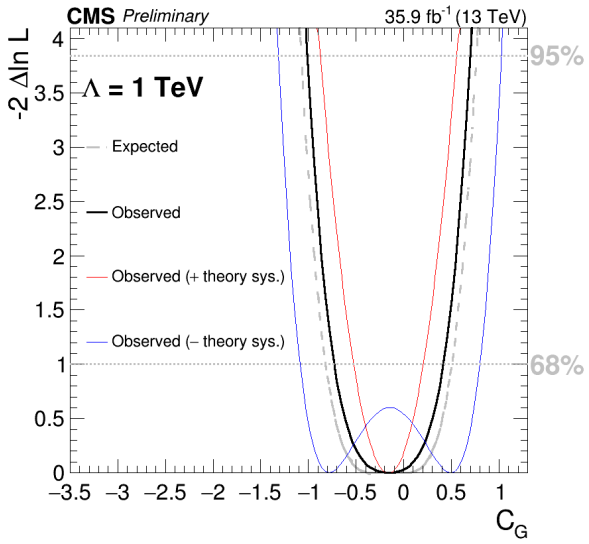

Figure 6.23: Likelihood scan of C_G effective coupling for combined channels.

6.9.3 Exclusion limits on C_{tG}, $C_{\phi q}^{(3)}$ and C_{tW} effective couplings

The deviation from the SM tW production from the interference terms between the SM and the O_{tG}, $O_{\phi q}^{(3)}$ 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 6.9.1, 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 are 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}, $C_{\phi q}^{(3)}$ 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 ($\geq$2jet,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 6.7. The inclusion of the ($\geq$2jet,2b-jet) and (2jet,1b-jet) regions helps to constrain the normalization and systematic uncertainties of the $t\bar{t}$ background. Comparisons between observed data and the SM background prediction for the MLP output shape in various jet-bjet regions are shown in Figure 6.24.

Three Wilson coefficients sensitive to new physics contribution in top quark interactions, as defined in Equation 2.3 are tested in observed data. The results for individual ee, $\mu\mu$ and eμ (a analogous analysis but using eμ final state) channels and all combined channels are listed in Table 6.18. The results of the likelihood scans of the Wilson coefficients on the full 13 TeV dataset are shown in Figure 6.25 for all combined channels.

		Best fit exp./obs.	68% exp./obs.limit	95% exp./obs.limit
$C_{\phi q}^{(3)}$	ee NN output for (1j1t+2j1t) + yields($\geq$2j,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($\geq$2j,2t)	0.00 / 1.13	[-2.20 to 1.92]/[-0.87 to 2.86]	[-4.68 to 3.66]/[-3.58 to 4.46]
	eμ NN output for (1j0t+1j1t+2j1t) + yields($\geq$2j,2t)	0.00 / -0.70	[-1.34 to 1.12]/[-2.16 to 0.59]	[-2.57 to 2.15]/[-3.74 to 1.61]
	Combined	0.00 / -1.52	[-1.05 to 0.88]/[-2.71 to -0.33]	[-2.04 to 1.63]/[-3.82 to 0.63]
C_{tW}	ee NN output for (1j1t+2j1t)+yields($\geq$2j,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($\geq$2j,2t)	0.00 / -1.40	[-2.18 to 6.97]/[-3.00 to 7.79]	[-3.63 to 8.42]/[-4.23 to 9.01]
	eμ NN output for (1j0t+1j1t+2j1t)+yields($\geq$2j,2t)	0.00 / 1.64	[-1.40 to 6.19]/[-0.80 to 5.59]	[-2.39 to 7.18]/[-1.89 to 6.68]
	Combined	0.00 / 2.38	[-1.14 to 5.93]/[0.22 to 4.57]	[-1.91 to 6.70]/[-0.96 to 5.74]
C_{tG}	ee NN output for (1j1t+2j1t)+yields($\geq$2j,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($\geq$2j,2t)	0.00 / -0.15	[-0.19 to 0.18]/[-0.34 to 0.02]	[-0.40 to 0.35]/[-0.53 to 0.19]
	eμ NN output for (1j0t+1j1t+2j1t)+yields($\geq$2j,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 6.18: Summary of allowed 68% CL and 95% CL intervals on C_{tG}, $C_{\phi q}^{(3)}$ and C_{tW} effective couplings obtained in ee, eμ, $\mu\mu$ and combined channels ($\Lambda = 1$ TeV).

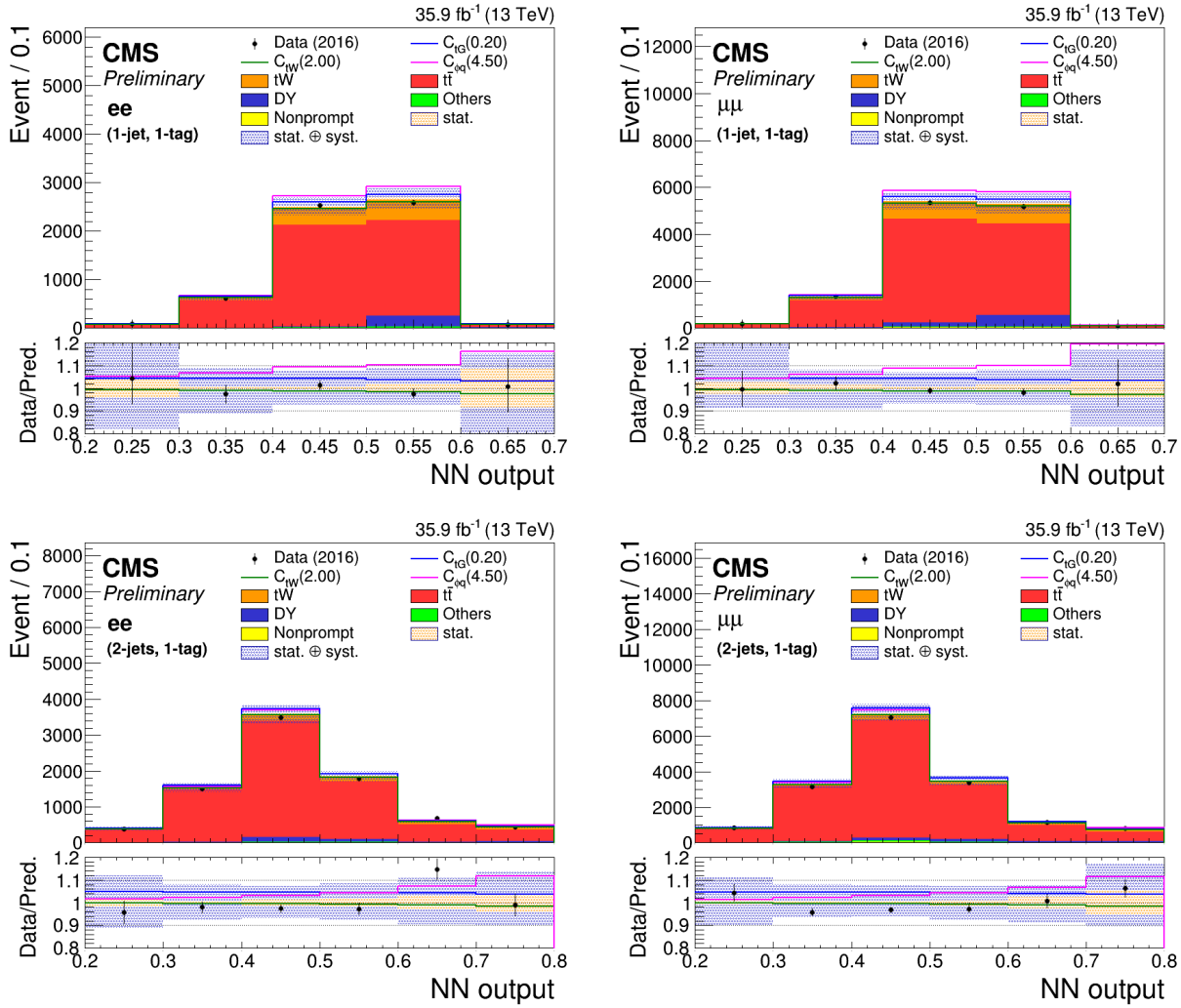

Figure 6.24: The MLP distributions of data and MC in different regions: (1jet,1b-jet) (top row), (2jet,1b-jet) (bottom row) used in limit setting for ee channel (left column) and μμ channel (right column). The blue hatched bands correspond to the sum of statistical and systematic uncertainties in the 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.

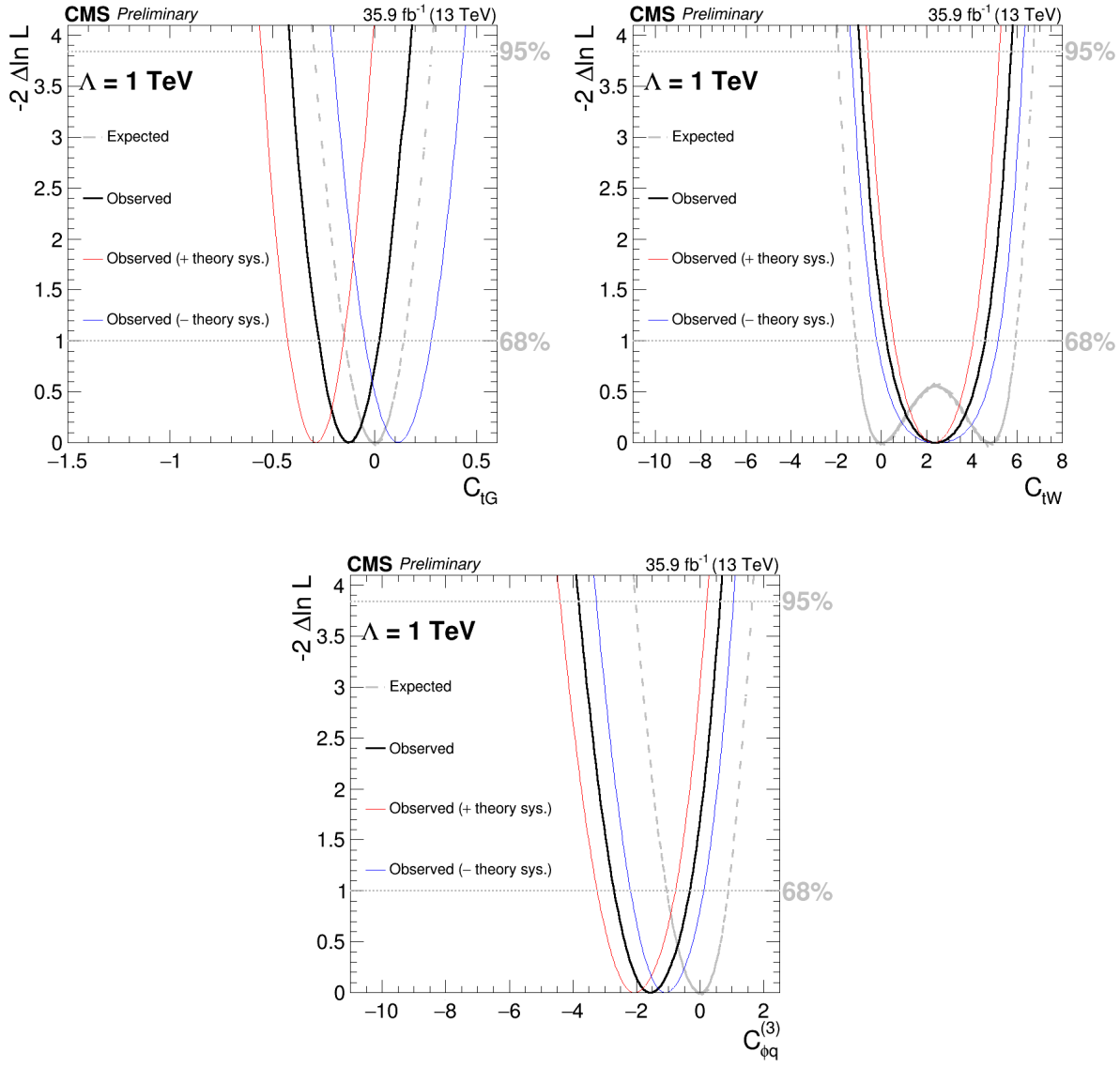

Figure 6.25: Likelihood scans of C_{tG}, $C_{\phi q}^{(3)}$ and C_{tW} effective couplings for combined channels.

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

Since the tW production via FCNC interactions does not interference 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 comparisons of the MLP output for the data, SM background and signal (tW events via FCNC interactions) in 1b-jet region are shown in Figure 6.26. Here the MLP is trained to separate FCNC tW events from SM tW and $t\bar{t}$ events as discussed in Section 6.7.

The limit results for individual ee, μμ and eμ (a analogous analysis but using eμ final state) channels and all combined channels are listed in Table 6.19. The results of the likelihood scans of the Wilson coefficients on the full 2016 dataset are shown in Figure 6.27 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 the combined channel in Table 6.20.

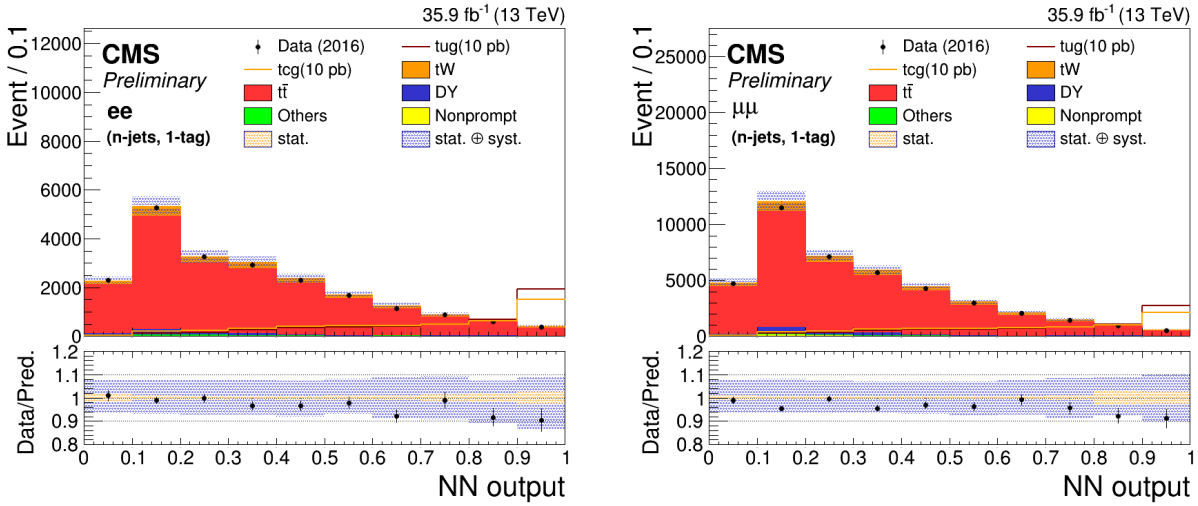

Figure 6.26: The MLP distributions of data, MC and FCNC signals in 1b-jet region used in limit setting for ee channel (left column) and μμ channel (right column). The blue hatched bands correspond to the sum of statistical and systematic uncertainties in the event yield for the sum of signal and background 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]
	μμ 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]
	eμ 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]
	μμ 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]
	eμ 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 6.19: Summary of allowed 68% CL and 95% CL intervals on C_{uG} and C_{cG} effective coupling obtained in ee, eμ, μμ and combined channels ($\Lambda = 1$ TeV).

The expected and observed 95% CL intervals on the Wilson coefficients obtained from the combination of all channels and signal regions are visualized in Figure 6.28.

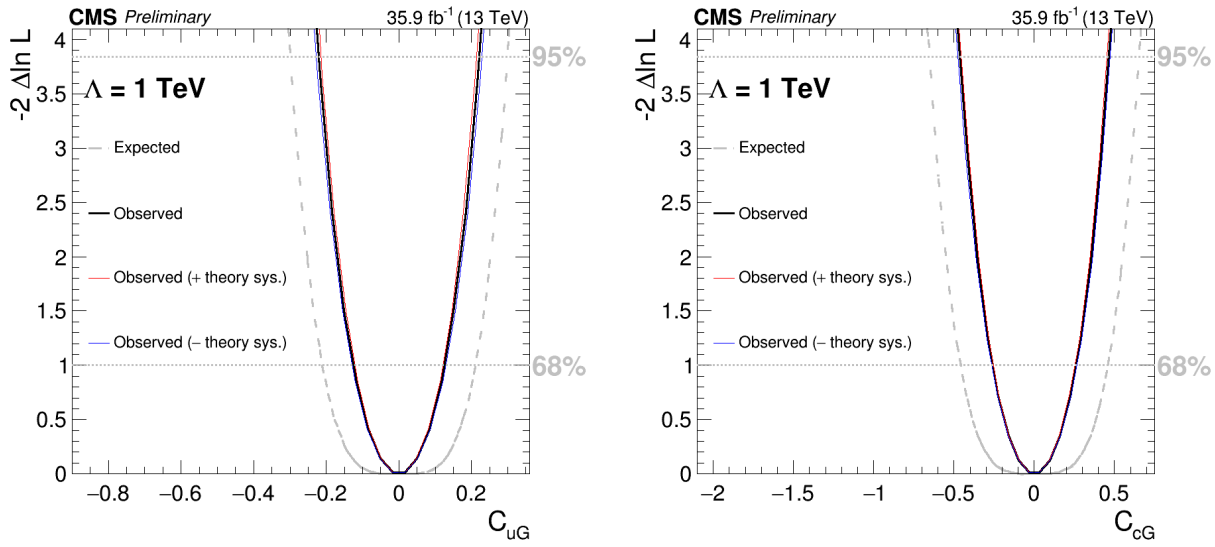

Figure 6.27: Likelihood scans of C_{uG} and C_{cG} effective couplings for combined channels.

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

Table 6.20: The expected and observed 95% CL intervals on the cross section of tW production via C_{uG} and C_{cG} effective couplings times square of the branching fraction $B(W\to \ell\nu)$, the effective couplings C_{uG} and C_{cG}, and the corresponding branching fractions $B(t \to ug)$ and $B(t \to cg)$.

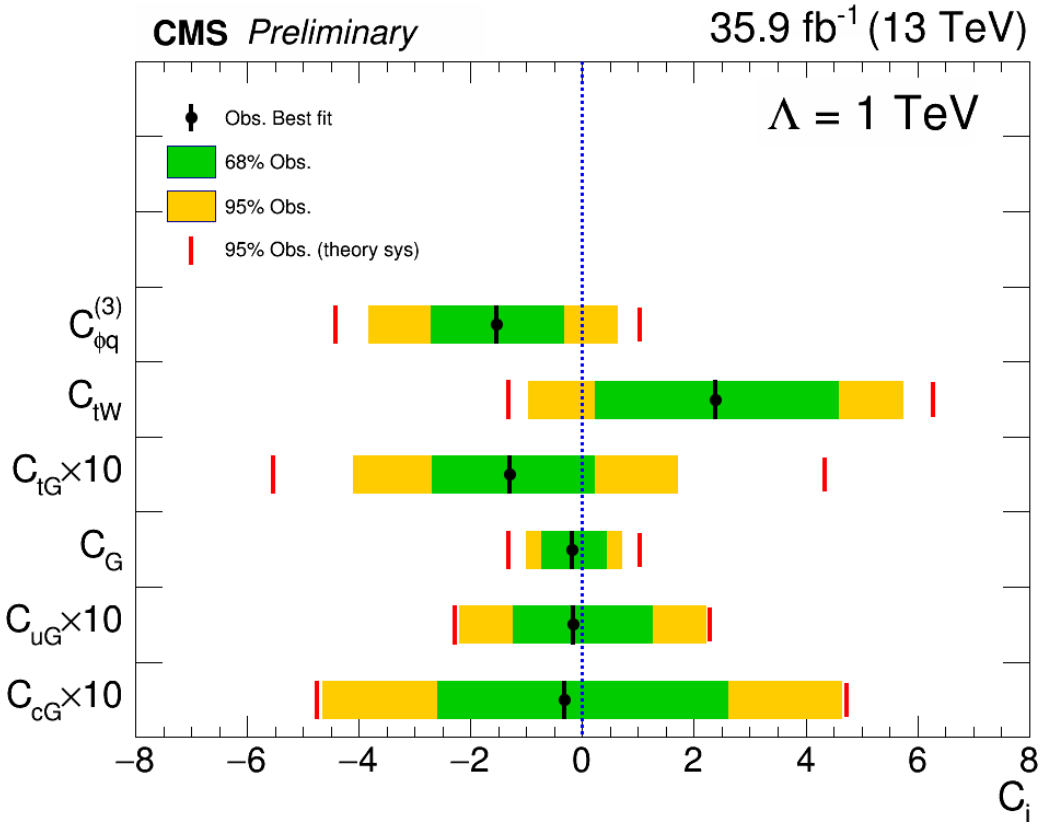

Figure 6.28: Observed and expected 95% CL intervals on the top quark effective couplings for combined channel ($\Lambda = 1$ TeV) [158].

6.10 Summary

A search for new physics via top quark production in ee and μμ 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 bands 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 limits 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 in this analysis: performing checks of distributions between data and simulations for different variables; working on the trigger scale factor study; estimating the various systematics uncertainties of the analysis; measuring the cross section of tW process for cross check; providing the final neural network spectra and the limits for different effective couplings. There is one publication for this analysis which is CMS-PAS-TOP-020.

This analysis is in publication by the CMS collaboration with one PAS document [52] and it has also been submitted to journal.

Chapter 7

Conclusions and perspectives

This thesis presented the latest available results from two separate searches for new physics beyond the standard model with the CMS detector at the LHC. One is searching for heavy resonances in dielectron final state, another is searching new physics in top quark production in dielectron or dimuon with some jets/bjets final states. The strategies of these two analyses are different. One is directly searching for the localized excess in the dielectron mass spectrum using a as simple and robust as possible way. The another one is indirectly searching for new physics and using a dedicated multivariate analysis to separate tW and $t\bar{t}$ processes to make the analysis more sensitive to new physics.

The search for new heavy resonances decaying in the dielectron final state was described in details in Chapter 5. The data used are the ones collected by the CMS experiment in 2016 with 35.9 fb^{-1} and the ones in 2017 with 41.4 fb^{-1}. The event selection is optimized in order to be highly efficient for high-energy electrons/positrons and to avoid loosing potential signal events. The main source of background, the Drell-Yan process, is estimated from simulation. Data-driven approach is used for validating the subleading background processes which are estimated from simulations, also it is used for the determination of the background coming from quantum chromo-dynamics process. After having inspected the dielectron invariant mass, no significant excess over the standard model background is observed, and upper limits at 95% confidence level are set on the ratio of production cross-section times branching ratio of a new resonance to the one at the Z boson peak, using a Bayesian approach. With the measured upper limits on the cross-section ratio, lower limits on the resonance masses have been set for particles predicted by various models. In particular, for spin 1 resonances, masses below 4.7 TeV, for the Z'_{SSM} particle from the sequential standard model, and below 4.1 TeV for the superstring inspired Z'_{ψ} particle could be excluded with the combination of dielectron (using 35.9 fb^{-1} from 2016 + 41.4 fb^{-1} from 2017) and dimuon (using 36.3 fb^{-1} from 2016) channels. This is the CMS most stringent limits to-date on the topic. The ATLAS collaboration has also provided several publications on the same topic. For instance, ATLAS [124] puts a 95% confidence level lower mass limits 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 thesis. Recently ATLAS published their full Run2 analysis (see Ref. [37]), and found the following 95% confidence level lower mass limits: 5.1 TeV for the Z'_{SSM} model and 4.5 TeV for the Z'_{ψ}.

The second analysis presented in this thesis is the search for new physics in top quark production and was described in details in Chapter 6. The single top quark production in association with a W boson is probed together with the top quark pair production to find the new physics signatures. Due to the similarity of the final states for tW and $t\bar{t}$ processes, a dedicated multivariate analysis is used to separate these two processes.

The data used is collected by the CMS experiment in 2016 with 35.9 fb^{-1}. Using the ee and $\mu\mu$ final states and combining with eμ final state, 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}, [−3.82,0.63] for $C_{\phi q}^{(3)}$, [−0.22,0.22] for C_{uG} and [−0.46,0.46] for C_{cG}. 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.

Although we haven't find the new physics until now, we may find it in the future with the help of much more data from LHC Run 3 and "high luminosity LHC" (HL-HLC). The foreseen operating scenario for Run 3 is to reach 13 - 14 TeV pp collision energy and keep the instantaneous luminosity at the level of 10^{34} $cm^{-2}s^{-1}$. At the end of Run 3 (the end of 2023), the LHC will deliver $\sim$ 300 fb^{-1} data which is 2 times larger than what we have now. From 2026 on, the HL-LHC starts and it is expected to operate at an enhanced luminosity of $5 \times 10^{35} cm^{-2}s^{-1}$. The total integrated luminosity delivered at the end of HL-LHC will be $\sim$ 3000 fb^{-1} which increases the statistics by around one order of magnitude comparing to the total delivered luminosity at the end of Run 3. Form a detector point of view, there will be many challenges for a detector to work very well at the future HL-LHC. For instance, due to the much higher radiation from the increased instantaneous luminosity, there is a plan to replace the full tracker and the endcap calorimeter of the ECAL subdetector with a high granularity calorimeter made of silicon sensors and tungsten absorber. Besides, the average number of pileup interactions will be much increased which is around 4 times higher what it is now. Therefore, the trigger system should be carefully optimized to save the interesting events. From physics analysis point of view, the results about SM parameter measurements can be updated precisely, for example the Higgs properties measurement. In addition, some rare SM processes (e.g. four top production, H $\rightarrow$ $\mu\mu$) could be observed thanks to the much increased statistics. Last but not least, using the massive cumulated data the new physics can be thorough searched at TeV energy scale, and a basic question about whether or not any new physics exists at the TeV energy scale can be addressed by the LHC experiments.

Publications, talks, posters, and internal notes

Publications

Journals

1. CMS Collaboration, "Search for high-mass resonances in dilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV", JHEP 2018.6 (2018). DOI: 10.1007/JHEP06(2018)120
2. CMS Collaboration, "Search for narrow resonances in dilepton mass spectra in proton-proton collisions at $\sqrt{s} = 13$ TeV and combination with 8 TeV data", Phys. Lett. B 768 (2017) 57. DOI:10.1016/j.physletb.2017.02.010
3. Barbara Clerbaux et al. "Model-independent constraints on the CKM matrix elements $|V_{tb}|$, $|V_{ts}|$ and $|V_{td}|$", JHEP (2019). DOI:10.1007/JHEP03(2019)022

Physics analysis summary

1. CMS Collaboration, "Search for new physics via top quark production in dilepton final state at at $\sqrt{s} = 13$ TeV", CMS-PAS-TOP-17-020
2. CMS Collaboration, "Search for high mass resonances in dielectron final state", CMS-PAS-EXO-18-006
3. CMS Collaboration, "Search for high-mass resonances in the dilepton final state using proton-proton collisions at $\sqrt{s} = 13$ TeV", CMS-PAS-EXO-16-047
4. CMS Collaboration, "Search for a high-mass resonance decaying into a dileptonfinal state in 13 fb^{-1} of pp collisions at $\sqrt{s} = 13$ TeV", CMS-PAS-EXO-16-031
5. CMS Collaboration, "Search for a narrow resonance produced in 13 TeV pp collisions decaying to electron pair or muon pair final states", CMS-PAS-EXO-15-005

Conference proceedings

1. W. Fang for the CMS Collaboration, "Search for heavy resonances in the dilepton final state with the CMS detector", proceedings for LHCP, CMS-CR-2016-189

Talks and posters

Talks at conferences

1. Parallel talk, "Search for new physics via top quark production in dilepton final state at 13 TeV", Chinese LHCP, 19 December 2018
2. Parallel talk, "Latest CMS results on heavy resonance search in the dilepton final state", General Meeting of the Belgian Physical Society, 11 April 2018

3. Parallel talk, "Search for high mass ee/eμ resonances in CMS", Chinese LHCP, 22 December 2017
4. Workshop talk, "Identification of high energy e/γ", CMS EGM workshop on 2017 performance and plans for 2018, 21 November 2017

CMS approval (pre-approve) talks

1. Approval talk of TOP-17-020, "Search for new physics in top quark production in di-lepton final state", June 2018
2. Pre-Approval talk of EXO-18-006, "High mass resonance search in ee channel with full 2017 data", February 2018
3. Approval talk of EXO-16-031, "Search for high mass dielectron resonances with 2016 data", July 2016

Posters at international conferences

1. "Search for heavy resonances in the dilepton final state with the CMS detector at the Large Hadron Collider", General Meeting of the Belgian Physical Society, UMons, Belgium, May 2017
2. "Search for heavy resonances in the dilepton final state with the CMS detector", LHCP conference, June 2016

Contribution to internal notes of the CMS collaboration

1. "Search for high-mass resonances in the di-electron final state with 2017 data", CMS-AN-18-021, 13 authors, 94 pages
2. "HEEP selection efficiency scale factor study for CMS with the full 2017 data sample", CMS-AN-18-143, 3 authors, 64 pages
3. "Search for new physics via top quark production in dilepton final state at 13 TeV", CMS-AN-17-134, 4 authors, 34 pages
4. "Measurement of the cross-section for tW production in dilepton final states at 13 TeV using 2016 data", CMS-AN-17-132, 4 authors, 72 pages
5. "HEEP selection efficiency scale factor study for CMS with the full 2016 data sample", CMS-AN-17-077, 3 authors, 54 pages
6. "Search for high mass di-electron resonances with the full 2016 data", CMS-AN-16-404, 14 authors, 104 pages
7. "Search for High Mass Di-Electron Resonances with 2016 Data", CMS-AN-16-190, 14 authors, 65 pages
8. "HEEP Selection Efficiency Scale Factor Study for CMS in Run 2", CMS-AN-15-320, 5 authors, 65 pages
9. "Dielectron resonance search in Run 2 at $\sqrt{s} = 13$ TeV pp collisions", CMS-AN-15-222, 17 authors, 60 pages
10. "Studies of two Egamma tag and probe HLT trigger paths for CMS in Run 2", CMS-AN-15-134, 4 authors, 17 pages

Acronym table

ALICE	A Large Ion Collider Experiment
ATLAS	A Toroidal LHC ApparatuS
CL	Confidence Level
CMS	Compact Muon Solenoid
CP	Charge Parity
CR	Color Reconnection
CSV	Combined Secondary Vertex
CSVv2	Combined Secondary Vertex version 2
DY	Drell-Yan
ECAL	Electromagnetic Calorimeter
EFT	Effective Field Theory
FCNC	Flavor Changing Neutral Current
FSR	Final State Radiation
GUT	Grand Unified Theory
HCAL	Hadronic Calorimeter
ISR	Initial State Radiation
JEC	Jet Energy Correction
JER	Jet Energy Resolution
LHC	Large Hadron Collider
LHCb	Large Hadron Collider beauty
MC	Monte Carlo
ME	Matrix Elements
MET	Missing Transverse Energy
MLP	Multiple Layer Perceptron
MVA	Multivariate Analysis
NLO	Next-Leading Order
NN	Neural Network
NNLO	Next-Next-Leading Order
PAG	Physics Analysis Group
PDF	Parton Distribution Function
PF	Particle-Flow
POG	Physics Object Group
PS	Parton Shower
PU	Pile Up
SM	Standard Model
SSM	Sequential Standard Model
UE	Underlying Event

Appendices

Appendix A

The Appendices for Theory

A.1 The Feynman Calculus

A.1.1 Lifetimes

As we know all particles can be separated by two categories, one is stable particle another is unstable particle. Such as the electron, photon and proton which are stable, while the muon, tau and neutron which are unstable. Actually most of particles are unstable and we are interested in what is the mass, lifetime, decay channel, spin and so on of these particles. For the unstable particles the number of particle versus time can be expressed by A.1 which comes from the integration of A.2, here the Γ means the decay rate of the particle and $\tau = \frac{1}{\Gamma}$ is called the mean lifetime of the particle.

$$N(t) = N(0)e^{-\Gamma t} \tag{A.1}$$

$$dN = -\Gamma\ N\ dt \tag{A.2}$$

Moreover if the particle can decay by different channels then each channel has different decay rate so the total decay rate is the sum of decay rate of each channel which is shown in A.3.

$$\Gamma_{tot} = \sum_{i=1}^{n} \Gamma_i \tag{A.3}$$

In addition to τ, we want to calculate the branching ratios which is the fraction of total particles decay to each channel. It is determined by the decay rates showed in A.4.

$$Br_i = \frac{\Gamma_i}{\Gamma_{tot}} \tag{A.4}$$

Therefore the most important physical quantity to describe the decay is the Γ and the way to calculate it will be described in section A.1.3.

Actually using the Schrödinger equation (the A.5) the wave function of the decay particle can be expressed by A.6, the E_R is the energy of the particle. And the existence possibility of the particle versus time is $\Psi\Psi^* \propto e^{\Gamma t}$ which is consistent with A.1.

$$(\frac{-\hbar^2}{2m}\nabla^2 + V)\Psi = i\hbar\frac{\partial}{\partial t}\Psi \quad \text{(Schrödinger equation)} \tag{A.5}$$

$$\Psi(t) = \Psi(0)e^{-iE_Rt}e^{\frac{-t}{2\tau}}, \text{ here the } \hbar = c = 1 \tag{A.6}$$

Using Fourier transform we can represent the wave function as the function of energy which is shown in A.7.

$$\chi(E) = \int_0^{+\infty} \Psi(t)e^{iEt}dt = \Psi(0)\int_0^{+\infty} e^{-t[\frac{\Gamma}{2}+i(E_R-E)]}dt = \frac{\Psi(0)}{(E-E_R)-\frac{i\Gamma}{2}} \tag{A.7}$$

Finally the existence possibility of the particle versus energy is $\chi\chi^*$ which is shown in A.8 and it gives the form of non-relativistic Breit Wigner (BW) distribution. Using the particle itself as the reference frame we can replace the E_R by M_R which is the mass of the particle. The full wave at half maximum (FWHM) of BW distribution is the Γ.

$$\chi(E)\chi(E)^* = \frac{\Psi(0)\Psi(0)^*}{(E-E_R)^2+\frac{\Gamma^2}{4}} \tag{A.8}$$

A.1.2 Cross-section

One very important physical quantity in high energy physics is "cross-section". The cross-section characterize the possibility of one physical process will be happen and its unit is cm^{-2}. For example for electron elastic scattering process $e + e \rightarrow e + e$ if we know the cross-section σ and the luminosity $\mathcal{L}$ which means the number of e e scattering (include elastic and inelastic) events in one cm^2 then the number of e e elastic scattering events can be expressed by A.9.

$$N_{elastic} = \sigma_{elastic} * \mathcal{L} \tag{A.9}$$

If we want to measure the cross-section of e e scattering for all possible process like $e + e \rightarrow$ anything, then it equal sum of the cross-section for each process.

$$\sigma_{tot} = \sum_{i=1}^{n} \sigma_i \tag{A.10}$$

The way to calculate the cross-section is explained in A.1.3.

A.1.3 The Golden Rule

The transition rate for a given process is described by the Fermi's "Golden Rule" shown in A.11. The $\mathcal{M}$ is called the amplitude for the process which contains all the dynamical information and it can be calculated by evaluating the relevant Feynman diagrams using the "Feynman rules". The phase space factor contains only the kinematical information, it depends on the masses, energies and momenta of the participants. For example a heavy particle decay into many light particles will involves a large phase space factor because there are many ways to apportion the available energy. By contrast, the decay of neutron ($n \rightarrow p + e + \bar{\nu}_e$) has very small phase space for the final state particles because the proton's mass just a bit smaller than neutron's.

$$\text{transition rate} = \frac{2\pi}{\hbar}|\mathcal{M}|^2 * (\text{phase space}) \tag{A.11}$$

- The golden rule for decay. Suppose one particle decays into several other particles:

$$1 \rightarrow 2 + 3 + ... + n \tag{A.12}$$

The decay rate is given by the formula A.13. Here the $p_i = (E_i/c, \mathbf{p_i})$ is the four-momentum of the ith particle with $m_i^2c^4 = E_i^2 - \mathbf{p_i^2c^2}$. The δ function enforces conservation of energy and momentum between decaying particle and final state particles. The decaying particle is assumed to be at rest so the $p_1 = (m_1c^2, \mathbf{0})$. the S is equal $\frac{1}{j!}$ for each group of j identical particles in the final state.

$$d\Gamma = |\mathcal{M}|^2 \frac{S}{2\hbar m_1}[(\frac{c\ d^3\mathbf{p_2}}{(2\pi)^3 2E_2})(\frac{c\ d^3\mathbf{p_3}}{(2\pi)^3 2E_3}) ... (\frac{c\ d^3\mathbf{p_n}}{(2\pi)^3 2E_n})]*(2\pi)^4\delta^4(p_1 - p_2 - p_3 ... - p_n) \tag{A.13}$$

The equation A.13 is the differential rate for a decay in which the momentum of particle 2 is within the range $d^3\mathbf{p_2}$ about the value $\mathbf{p_2}$, similar for other final state particles. Usually we are more interested in the decay rate for full phase space of final state particles, so we can integrate the equation A.13 over all momenta of final state particles to get the total decay rate Γ. For example for two body decay the total Γ is expressed in A.14, besides after some algebraic calculations the equation A.14 can be simplified with equation A.15 and without knowing $\mathcal{M}$ which is the function of final state particles momenta. Here the $\mathbf{p}$ is the magnitude of momentum for either final state particle. However there are no such simplified equation for more than two body decay until we know the specific functional from of $\mathcal{M}$. In such case we can only go back to the equation A.13 and work it out from scratch.

$$\Gamma = \frac{S}{\hbar m_1}(\frac{c}{4\pi})^2 \frac{1}{2}\int \frac{|\mathcal{M}|^2}{E_2 E_3}\delta^4(p_1 - p_2 - p_3)\, d^3\mathbf{p_2} d^3\mathbf{p_3} \tag{A.14}$$

$$\Gamma = \frac{S|\mathcal{M}|^2|\mathbf{p}|}{8\pi\hbar m_1^2 c},\ |\mathbf{p}| = \frac{c}{2m_1}\sqrt{m_1^4 + m_2^4 + m_3^4 - 2m_1^2m_2^2 - 2m_1^2m_3^2 - 2m_2^2m_3^2} \tag{A.15}$$

- The golden rule for scattering. Suppose two particles collide and producing several particles:

$$1 + 2 \rightarrow 3 + 4 + ... + n \tag{A.16}$$

The cross-section is given by the formula A.17. Here, as before, the $p_i = (E_i/c, \mathbf{p_i})$ is the four-momentum of the ith particle with $m_i^2c^4 = E_i^2 - \mathbf{p_i^2c^2}$. The δ function enforces conservation of energy and momentum. the S is equal $\frac{1}{j!}$ for each group of j identical particles in the final state.

$$\begin{aligned} d\sigma &= |\mathcal{M}|^2 \frac{\hbar^2 S}{4\sqrt{(p_1\cdot p_2)^2 - (m_1m_2c^2)^2}}[(\frac{c\, d^3\mathbf{p_3}}{(2\pi)^3 2E_3})(\frac{c\, d^3\mathbf{p_4}}{(2\pi)^3 2E_4}) ... (\frac{c\, d^3\mathbf{p_n}}{(2\pi)^3 2E_n})] \\ &* (2\pi)^4\delta^4(p_1 + p_2 - p_3 - p_4 ... - p_n) \end{aligned} \tag{A.17}$$

Equation A.17 gives the cross-section when the momentum of particle 3 within the range $d^3\mathbf{p_3}$ about the value $\mathbf{p_3}$, similar for other final state particles. If we want to know the total cross-section for full phase space, we can integrate the equation over the momenta of all particles. Similar with decay, for two body scattering $(1 + 2 \rightarrow 3 + 4)$ in the central mass (CM) frame we can get simplified equation which is shown in A.18. Here the $|\mathbf{p_i}|$ is the magnitude of either incoming momentum and $|\mathbf{p_f}|$ is the magnitude of either outgoing momentum, the $d\Omega = sin\theta\ d\theta\ d\phi$ is the differential solid angle of either outgoing particle. For more than two body scattering we need work it out from scratch using equation A.17.

$$\frac{d\sigma}{d\Omega} = (\frac{\hbar c}{8\pi})^2 \frac{S|\mathcal{M}^2|}{(E_1 + E_2)^2}\frac{|\mathbf{p_f}|}{|\mathbf{p_i}|} \tag{A.18}$$

Now we know the form of golden rule for calculating the decay rate and scattering cross-section, while in order to get the exact value we still need to know the $\mathcal{M}$ in the equations and it can be calculated by evaluating the relevant Feynman diagrams using the "Feynman rules" which is explained in section A.1.4.

A.1.4 The Feynman Rules for A Toy Theory

This section will gives the Feynman rules for a toy theory which means the spin of the particles are not considered, so it will not present the real world which is more complicated

and will be addressed in section, but still it gives the basic idea how Feynman rules works. Below is the ritual:

1. **Draw the Feynman diagram** : The figure A.1 shows an example of the lowest-order Feynman digram for particle A decay into particle B and C in the left plot and two particles scattering ($A + A \rightarrow B + B$) in the right. The p_i means the four-momenta of incoming or outgoing particles, the q_i means the four-momenta of intermediate particles like the particle C in the right plot of A.1. The arrow on each line represent the direction of the particle's movement (for internal lines it is arbitrary).

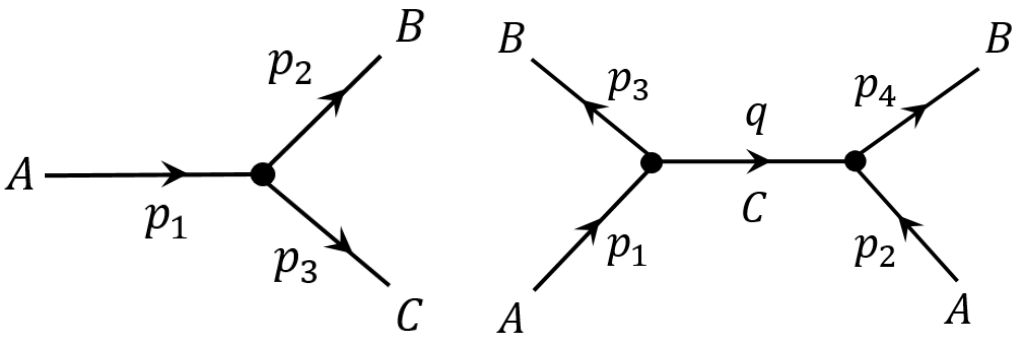

Figure A.1: The lowest-order Feynman diagram for $A \rightarrow B + C$ (left) and $A + A \rightarrow B + B$ (right).

2. **Coupling Constant** : For each vertex write down a factor of $-ig$. The g is called coupling constant which represent the strength of the interaction between different particles.
3. **Propagator** : For each internal line write down a factor of $\frac{i}{q_i^2 - m_i^2 c^2}$, here the q_i is the four-momentum of the line ($q_i^2 = (\frac{E_i}{c})^2 - \mathbf{p_i}^2$) and the m_i is the rest mass of the particle the internal line describes. Note that $q_i^2 \neq m_i^2 c^2$, because the propagator is virtual particle and it does not lie on its mass shell.
4. **Conservation of energy and momentum** : For each vertex write down a δ function with the form of $(2\pi)^4 \delta^4(k_1 + k_2 + k_3)$. Here the k_i are the four-momentum of the particle which involved in the vertex, if the arrow of line k_i is toward to the vertex then put a positive sign in front of k_i otherwise put a negative sign. This factor imposes the conservation of energy and momentum at each vertex, since the δ function is zero unless the sum of incoming momenta equals the sum of outgoing momenta.
5. **Integration over internal momenta** : For each internal line write down a factor $\frac{1}{(2\pi)^4 d^4 q_i}$ and integrate over all internal momenta.
6. **Cancel the δ function** : Erase $(2\pi)^4 \delta^4(p_1 + p_2 + ... + p_n)$ in the result which enforcing overall conservation of energy and momentum. Finally what remains is $-i\mathcal{M}$.

Now we can use Feynman rules to get the amplitude for lowest-order of particle A decay into particle B and C (see figure A.1 left). Through rule 1 to 5 it gives $(-ig)(2\pi)^4(p_1 - p_2 - p_3)$ and using rule 6 it remains $-ig$ only which equal $-i\mathcal{M}$. Finally the $\mathcal{M}$ equal g. Therefore the decay rate of particle A is found by plugging $\mathcal{M}$ into equation A.15 which is shown in A.19. The lifetime of the particle A is shown in

A.20.

$$\Gamma = \frac{g^2|\mathbf{p}|}{8\pi\hbar m_A^2 c}, \quad |\mathbf{p}| = \frac{c}{2m_A}\sqrt{m_A^4 + m_B^4 + m_C^4 - 2m_A^2m_B^2 - 2m_A^2m_C^2 - 2m_B^2m_C^2} \tag{A.19}$$

$$\tau = \frac{1}{\Gamma} = \frac{8\pi\hbar m_A^2 c}{g^2|\mathbf{p}|} \tag{A.20}$$

Similarly using Feynman rules we can get the amplitude for lowest-order of two particles A scattering process $A + A \to B + B$ (shown in figure A.1 right). The result of rule 1 to rule 5 is shown in A.21. After some algebraic calculations and using rule 6, the final result for $\mathcal{M}$ is shown in A.22.

$$\int (-ig)^2 \frac{i}{q^2 - m_C^2c^2} (2\pi)^4\delta^4(p_1 - p_3 - q) (2\pi)^4\delta^4(p_2 + q - p_4) \frac{1}{(2\pi)^4} d^4q \tag{A.21}$$

$$\mathcal{M} = \frac{g^2}{(p_1 - p_3)^2 - m_C^2c^2} = \frac{g^2}{(p_4 - p_2)^2 - m_C^2c^2} \tag{A.22}$$

But there is another diagram contribute to the lowest-order of the process $(A + A \to B + B)$ which is shown in figure A.2. This is obtained by twisting the outgoing lines of the right plot of figure A.1 and making p_1 connect to p_4, p_2 connect to p_3. Similar we can using Feynman rules to get the amplitude for this process. Finally sum up the $\mathcal{M}$ form these two lower-order diagram, we can get the total $\mathcal{M}$ for $A + A \to B + B$ processes which is shown in A.23.

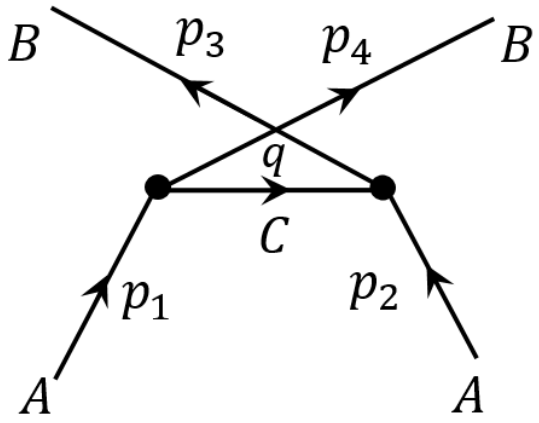

Figure A.2: Another Feynman diagram contributing in lowest-order to $A + A \to B + B$.

$$\mathcal{M} = \frac{g^2}{(p_1 - p_3)^2 - m_C^2c^2} + \frac{g^2}{(p_3 - p_2)^2 - m_C^2c^2} \tag{A.23}$$

A.2 Quantum Electrodynamics

A.2.1 Dirac Equation

As we know in nonrelativistic quantum mechanics particles are described by the Schrödinger equation (the A.5). In relativistic quantum mechanics particles of spin $\frac{1}{2}$ are described by Dirac equation which is shown in equation A.24 with the definition of ∂_μ A.25 and γ^μ A.26 (the standard "Bjorken and Drell" convention). For particles with other spins such as spin 0 particles they are described by the Klein-Gordon equation A.27 and spin 1 particles are described by Proca equation.

$$i\hbar\gamma^\mu\partial_\mu\Psi - mc\Psi \ = 0 \ \ \text{(Dirac equation)} \tag{A.24}$$

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu},\ \partial_0 = \frac{1}{c}\frac{\partial}{\partial t},\ \partial_1 = \frac{\partial}{\partial x},\ \partial_2 = \frac{\partial}{\partial y},\ \partial_3 = \frac{\partial}{\partial z} \tag{A.25}$$

$$\begin{aligned} \gamma^0 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},\ \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \\ \sigma^1 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},\ \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned} \tag{A.26}$$

$$-\frac{1}{c^2}\frac{\partial^2\Psi}{\partial t^2} + \nabla^2\Psi \ = (\frac{mc}{\hbar})^2\Psi \ \ \text{(Klein} - \text{Gordon equation)} \tag{A.27}$$

We can using Dirac equation to solve the plane wave function A.28 of particles with spin $\frac{1}{2}$. Here the a is normalization constant and $u(p)$ represent the spin of particles.

$$\Psi(x) \ = \ ae^{-(i/\hbar)x\cdot p}u(p),\ \text{with x} = (\text{ct}, \mathbf{x}),\ \text{p} = (\frac{\text{E}}{\text{c}}, \mathbf{p}) \tag{A.28}$$

Putting this into the Dirac equation A.24 it gives equation A.29 and using equation A.30 the we get A.31.

$$(\gamma^\mu p_\mu - mc)u = 0 \tag{A.29}$$

$$\gamma^\mu p_\mu = \gamma^0 p_0 - \boldsymbol{\gamma}\cdot\mathbf{p} = \frac{E}{c}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} - \mathbf{p}\cdot\begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix} = \begin{pmatrix} E/c & -\mathbf{p}\cdot\boldsymbol{\sigma} \\ \mathbf{p}\cdot\boldsymbol{\sigma} & -E/c \end{pmatrix} \tag{A.30}$$

$$(\gamma^\mu p_\mu - mc)u = \begin{pmatrix} \frac{E}{c} - mc & -\mathbf{p}\cdot\boldsymbol{\sigma} \\ \mathbf{p}\cdot\boldsymbol{\sigma} & -\frac{E}{c} - mc \end{pmatrix}\begin{pmatrix} u_A \\ u_B \end{pmatrix} = \begin{pmatrix} (\frac{E}{c} - mc)u_A - \mathbf{p}\cdot\boldsymbol{\sigma}u_B \\ \mathbf{p}\cdot\boldsymbol{\sigma}u_A - (\frac{E}{c} + mc)u_B \end{pmatrix} \tag{A.31}$$

Finally after some algebraic calculations we get the two solutions for electron shown in A.32 and two solutions for positron shown in A.33 and the normalization constant $N = \sqrt{(|E| + mc^2)/c}$ in order to satisfy the normalization requirement $u^\dagger u = 2|E|/c$.

$$u^{(1)} = N\begin{pmatrix} 1 \\ 0 \\ \frac{cp_z}{E+mc^2} \\ \frac{c(p_x+ip_y)}{E+mc^2} \end{pmatrix},\ u^{(2)} = N\begin{pmatrix} 0 \\ 1 \\ \frac{c(p_x-ip_y)}{E+mc^2} \\ \frac{-cp_z}{E+mc^2} \end{pmatrix}, \text{with } E = \sqrt{m^2c^4 + \mathbf{p}^2c^2} \tag{A.32}$$

$$u^{(3)} = N\begin{pmatrix} \frac{cp_z}{E-mc^2} \\ \frac{c(p_x+ip_y)}{E-mc^2} \\ 1 \\ 0 \end{pmatrix},\ u^{(4)} = N\begin{pmatrix} \frac{c(p_x-ip_y)}{E-mc^2} \\ \frac{-cp_z}{E-mc^2} \\ 0 \\ 1 \end{pmatrix}, \text{with } E = -\sqrt{m^2c^4 + \mathbf{p}^2c^2} \tag{A.33}$$

Moreover if we choose the direction of motion as Z axis then the $u^{(i)}$ will be the eigenstate of S_z which is defined in A.34, the $u^{(1)}$ and $u^{(3)}$ are spin up, $u^{(2)}$ and $u^{(4)}$ are spin down.

$$\mathbf{S} = \frac{\hbar}{2}\begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} \tag{A.34}$$

A.2.2 The Photon

As we know in classical electrodynamic theory the electric field ($\mathbf{E}$) and magnetic field ($\mathbf{B}$) is described by Maxwell's equations shown in A.35 with charge density ρ and current density $\mathbf{J}$.

$$\begin{aligned} &\textbf{(i)}\ \ \nabla\cdot\mathbf{E} = 4\pi\rho \qquad &&\textbf{(iii)}\ \ \nabla\cdot\mathbf{B} = 0 \\ &\textbf{(ii)}\ \ \nabla\times\mathbf{E} + \frac{1}{c}\frac{\partial\mathbf{B}}{\partial t} = 0 \quad &&\textbf{(iv)}\ \ \nabla\times\mathbf{B} - \frac{1}{c}\frac{\partial\mathbf{E}}{\partial t} = \frac{4\pi}{c}\mathbf{J} \end{aligned} \tag{A.35}$$

If we using the "field strength tensor" $F^{\mu\nu}$ shown in A.36 (e.g. $F^{01} = -E_x$, $F^{12} = -B_z$, etc.) and J^μ show in A.37 then the equation **i** and **iv** can be expressed by equation A.38.

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \tag{A.36}$$

$$J^\mu = (c\rho, \mathbf{J}) \tag{A.37}$$

$$\partial_\mu F^{\mu\nu} = \frac{4\pi}{c} J^\nu \tag{A.38}$$

Because of the antisymmetry of $F^{\mu\nu}$ (that is $F^{\mu\nu} = -F^{\nu\mu}$), using equation A.38 we can get the "continuity equation" which expressing the conservation of charge shown in A.39.

$$\partial_\mu J^\mu = 0 \ \text{ or } \ \nabla\cdot\mathbf{J} = -\frac{\partial\rho}{\partial t} \tag{A.39}$$

For (**iii**) in Maxwell's equations it can be rewritten as $\mathbf{B}$ equal the curl of vector potential $\mathbf{A}$ shown in A.40, then the (**ii**) in Maxwell's equations becomes equations A.41 which is equivalent to say the $\mathbf{E} + (1/c)(\partial\mathbf{A}/\partial t)$ can be written as the gradient of scale potential V shown in A.42.

$$\mathbf{B} = \nabla\times\mathbf{A} \tag{A.40}$$

$$\nabla\times(\mathbf{E} + \frac{1}{c}\frac{\partial\mathbf{A}}{\partial t}) = 0 \tag{A.41}$$

$$\mathbf{E} = -\nabla V - \frac{1}{c}\frac{\partial\mathbf{A}}{\partial t} \tag{A.42}$$

In relativistic notation, the equations A.40 and A.42 can be written as equation A.43.

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu, \ \text{ with } A^\mu = (V, \mathbf{A}),\ \partial^\mu = (\frac{1}{c}\frac{\partial}{\partial t}, -\nabla) \tag{A.43}$$

Then the equation A.38 becomes A.44.

$$\partial_\mu\partial^\mu A^\nu - \partial^\nu(\partial_\mu A^\mu) = \frac{4\pi}{c} J^\nu \tag{A.44}$$

Until now we do not know the value of $\mathbf{A}$ and V, actually if we put $A^{\mu'}$ which is A^μ plus $\partial\lambda$ (λ is a scaler) shown in A.45 into equation A.43 there will no change to $F^{\mu\nu}$. This is called "gauge transformation" and we can exploit it to impose an extra constraint on the potential which is called "Lorentz condition" shown in A.46.

$$A^{\mu'} = A^\mu + \partial\lambda \tag{A.45}$$

$$\partial_\mu A^\mu = 0 \tag{A.46}$$

With "Lorentz condition" the equation A.44 is simplified to A.47.

$$\Box = \frac{4\pi}{c} J^\nu, \quad \text{with } \Box \equiv \partial_\mu \partial^\mu = \frac{1}{c^2}\frac{\partial^2}{\partial t^2} - \nabla^2 \tag{A.47}$$

In empty space where $J^\mu = 0$ we can pick $A^0 = 0$ (called "Coulomb gauge") then the equation A.46 becomes A.48.

$$\nabla \cdot \mathbf{A} = 0 \tag{A.48}$$

In quantum electrodynamics A^μ becomes the wave function of photon. For free photon it should satisfies equation A.47 with $J^\mu = 0$. Similar with electron case, we can make a plane wave for photon which is equation A.49, here the ϵ is polarization vector which represent the spin of photon, a is normalization factor and $p = (E/c, \mathbf{p})$. Putting equation A.49 into A.47 it gives A.50, equation A.50 just tell us the photon is massless particle which is the same as expected.

$$A^\mu(x) = a e^{-(i/\hbar)p\cdot x}\epsilon^\mu(p) \tag{A.49}$$

$$p_\mu p^\mu = 0 \quad \text{or} \quad E = |\mathbf{p}|c \tag{A.50}$$

Using Lorentz gauge A.46 and Coulomb gauge ($A^0 = 0$) we can get the solution for ϵ^μ which is shown in A.51.

$$\epsilon^0 = 0, \quad \epsilon \cdot \mathbf{p} = 0 \tag{A.51}$$

A.51 tell us ϵ is perpendicular to the $\mathbf{p}$ and if we choose the direction of photon's motion as Z axis, then only ϵ^1 and ϵ^2 are non-zero which can make the photon has spin 1 or -1.

A.2.3 The Feynman Rules for QED

Different with section A.1 which is a toy theory of Feynman rules, here we are going to describe the Feynman rules for QED in real world. The process is similar with section A.1 which is shown in the following:

1. **Draw the Feynman diagram** : Draw the corresponding Feynman diagrams for the process. Label four-momenta p_1, p_2, ... for incoming and outgoing particles together with the corresponding spins s_1, s_2, ... Label four-momenta q_1, q_2, ... for internal particles. The arrow in the line represent the direction of motion of electron or photon, while for positron it is opposite with its motion.
2. **External line** : For each external line write down u ($\bar{u}$) for incoming (outgoing) electron or v ($\bar{v}$) for incoming (outgoing) positron or ϵ^μ ($\epsilon^{\mu*}$) for incoming (outgoing) photon.
3. **Vertex factor** : For each vertex write down a factor of $ig_e\gamma^\mu$. The coupling constant g_e represents the strength of the electromagnetic interaction with value $g_e = e\sqrt{4\pi/\hbar c}$.

4. **Propagator** : For each internal line write down a factor which is shown in A.70.

$$\begin{aligned} &\text{For electron and positron} \ : \ \frac{i(\gamma^{\mu}q_{\mu}+mc)}{q^2-m^2c^2} \\ &\text{For photon} \ : \ \frac{-ig_{\mu\nu}}{q^2} \end{aligned} \tag{A.52}$$

5. **Conservation of energy and momentum** : For each vertex write down a δ function with the form of $(2\pi)^4\delta^4(k_1 + k_2 + k_3)$. Here the k_i are the four-momentum of the particle which involved in the vertex, if the arrow of line k_i is toward to the vertex then put a positive sign in front of k_i otherwise put a negative sign (opposite for positron). This factor imposes the conservation of energy and momentum at each vertex, since the δ function is zero unless the sum of incoming momenta equals the sum of outgoing momenta.
6. **Integration over internal momenta** : For each internal line write down a factor $\frac{1}{(2\pi)^4 d^4q_i}$ and integrate over all internal momenta.
7. **Cancel the δ function** : Erase $(2\pi)^4\delta^4(p_1 + p_2 + ... + p_n)$ in the result which enforcing overall conservation of energy and momentum. Finally what remains is $-i\mathcal{M}$.
8. **Antisymmetrization** : Including a minus sign between diagrams which differ only in interchange of two incoming (or outgoing) electrons or positrons, or of an incoming electron with an outgoing positron (or vice versa).

Using the Faynman rules for QED we can calculate the amplitude of the QED process. For example for electron electron scattering process which is shown in figure A.3 using Feynman rules from 1 to 6 for left plot it gives the result A.53.

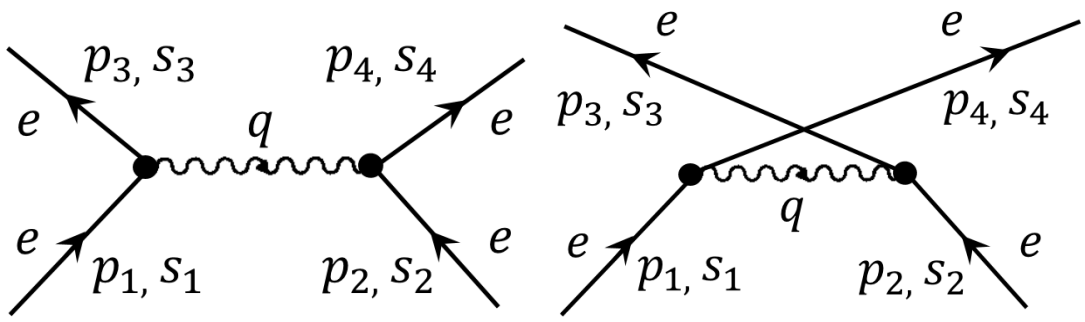

Figure A.3: The two lowest-order Feynman diagrams for electron electron scattering process.

$$\begin{aligned} &(2\pi)^4 \int [\bar{u}^{(s_3)}(p_3)(ig_e\gamma^{\mu})u^{(s_1)}(p_1)]\frac{-ig_{\mu\nu}}{q^2}[\bar{u}^{(s_4)}(p_4)(ig_e\gamma^{\nu})u^{(s_2)}(p_2)] \\ &* \delta^4(p_1-p_3-q)\delta^4(p_2+q-p_4)d^4q \end{aligned} \tag{A.53}$$

After do the integration and using Feynman rule 7 it gives the $\mathcal{M}$ for left plot in figure A.3 which is shown in A.54.

$$\mathcal{M} = -\frac{g_e^2}{(p_1-p_3)^2}[\bar{u}^{(s_3)}(p_3)\gamma^{\mu}u^{(s_1)}(p_1)][\bar{u}^{(s_4)}(p_4)\gamma_{\mu}u^{(s_2)}(p_2)] \tag{A.54}$$

Similar we can get the $\mathcal{M}$ for the right plot in figure A.3 which is shown in A.55

$$\mathcal{M} = -\frac{g_e^2}{(p_1-p_4)^2}[\bar{u}^{(s_4)}(p_4)\gamma^{\mu}u^{(s_1)}(p_1)][\bar{u}^{(s_3)}(p_3)\gamma_{\mu}u^{(s_2)}(p_2)] \tag{A.55}$$

According Feynman rule 8 these two diagrams need to be subtracted because the only difference between these two diagrams is the interchange of two outgoing electrons, it does not matter which diagram subtract which diagram because $|\mathcal{M}|^2$ is used ultimately. Finally we get the $\mathcal{M}$ for lowest-order (or tree level) of electron electron scattering process which is shown in A.56.

$$\mathcal{M} = -\frac{g_e^2}{(p_1 - p_3)^2}[\bar{u}(3)\gamma^\mu u(1)][\bar{u}(4)\gamma_\mu u(2)] + \frac{g_e^2}{(p_1 - p_4)^2}[\bar{u}(4)\gamma^\mu u(1)][\bar{u}(3)\gamma_\mu u(2)] \quad \text{(A.56)}$$

Until now we only consider to tree level of the electron electron scattering process, but if we add the high order contributions like one loop in the photon propagator (left plot of figure A.4), the result will be infinite. In order to remove this divergence, a method called "renormalization" is used which absorb the infinities into "renormalized" masses and coupling constant. For example transfer the $g_e = e_{bare}\sqrt{4\pi/\hbar c}$ to $g_R = e_{eff}\sqrt{4\pi/\hbar c}$ which means when consider the high order effect we should use the effective charge of an electron not the original charge (or the bare charge), this is because of vacuum polarization effect which behaviours like an electron virtually emits photons and the photons produce electron-positron pairs and those pairs emit further photons and so on which is shown in right plot of figure A.4. Hence, an electron turns out to be surrounded by many virtual electrons and positrons and due to Coulomb attractive force, positrons become closer to the original electron and thus the vacuum is polarized. Besides the effective charge will be changed by the momentum transferred q in the collision, higher momentum transfer means more closer between two collided particles and higher effective charge for each particle. So the coupling g_R is changing with q and is called "running" coupling constant. However, the change of g_R or fine structure constant ($\alpha = \frac{g_R^2}{4\pi} = \frac{e_{eff}^2}{\hbar c}$) versus q is very small which is shown in right plot of figure A.5.

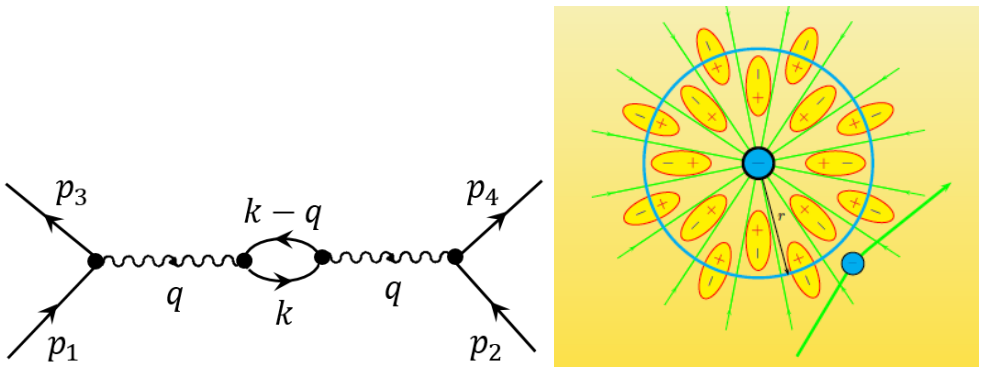

Figure A.4: The representation of the Feynman diagram with one loop in the photon propagator (left) and the vacuum polarization phenomenon causing charge screening by virtual pairs (right) [159].

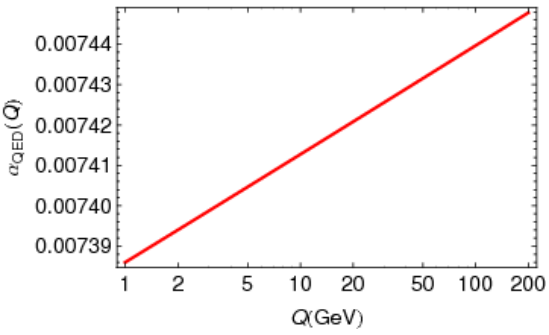

Figure A.5: The momentum transfer evolution of fine structure constant α_{QED} [160].

A.3 Quantum Chromodynamics

A.3.1 Quark Color

The fact of quarks have colors is supported by many experiments and a direct evidence for the species of color be to 3 comes from the e^+e^- annihilations experimental result on the R value which defined in A.57. The first step of $e^+ + e^- \rightarrow hadrons$ is $e^+ + e^- \rightarrow \gamma \rightarrow q + \bar{q}$ and the Feynaman diagram is shown in left plot of figure A.6. This is ordinary QED process and the same as $e^+ + e^- \rightarrow \gamma \rightarrow \mu^+ + \mu^-$ when we replace the mass and charge of the quark by muon's mass and charge. Using the Feynman rules in A.2.3 we can finally get the cross-section of $e^+ + e^- \rightarrow \gamma \rightarrow q + \bar{q}$ or $e^+ + e^- \rightarrow \gamma \rightarrow \mu^+ + \mu^-$ which is shown in A.58 when the energy of electron is substantially above the threshold ($E_e > Mc^2 \gg m_e c^2$, M is the mass of quark or muon). The second step of $e^+ + e^- \rightarrow hadrons$ is called "hadronization" which means when the two produced quarks fly away and reach a separation distance of around 10^{-15} m (the diameter of a hadron) the strong interaction becomes so great that new quark-antiquark pairs are produced which mainly from gluons and the phenomenon is shown in right plot of figure A.6.

$$R = \frac{\sigma(e^+e^- \rightarrow hadrons)}{\sigma(e^+e^- \rightarrow \mu^+\mu^-)} \tag{A.57}$$

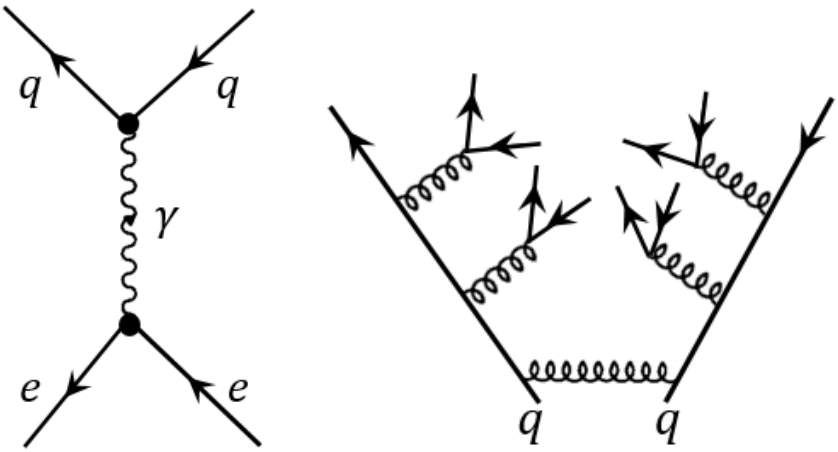

Figure A.6: The Feynman diagram for $e^+ + e^- \rightarrow \gamma \rightarrow q + \bar{q}$ process (left). The representation of the hadronization phenomenon for quarks (right).

$$\sigma = \frac{\pi}{3}(\frac{\hbar Q c \alpha}{E})^2 \tag{A.58}$$

Putting equation A.58 into A.57, it gives $R = \sum q e_q^2$, where the e_q is the electric charge of the quark q in quarks pairs produced in the e^+e^- annihilation. Beyond the s quark pair production threshold but lower then c quark pair production threshold, only u, d, s quarks contribute to this ratio R and yield A.59.

$$\begin{aligned} R &= e_u^2 + e_d^2 + e_s^2 = \frac{2}{3}, \text{ (without color)} \\ R &= 2, \text{ (with 3 color)} \end{aligned} \tag{A.59}$$

Similarly, for higher energies beyond the c quark pair production threshold but lower then

b quark pair production threshold, the R is shown in A.60.

$$\begin{aligned} R &= e_u^2 + e_d^2 + e_s^2 + e_c^2 = \frac{10}{9}, \text{ (without color)} \\ R &= \frac{10}{3}, \text{ (with 3 color)} \end{aligned} \tag{A.60}$$

And for more higher energies beyond the b quark pair production threshold but lower then t quark pair production threshold, the R is shown in A.61.

$$\begin{aligned} R &= e_u^2 + e_d^2 + e_s^2 + e_c^2 + e_b^2 = \frac{11}{9}, \text{ (without color)} \\ R &= \frac{11}{3}, \text{ (with 3 color)} \end{aligned} \tag{A.61}$$

The experimental result prefer the 3 colors for any case.

A.3.2 Feynman Rules for Chromodynamics

For strong interaction the coupling is $g_s = \sqrt{4\pi\alpha_s}$ which is similar with the coupling of electromagnetic interaction $g_e = \sqrt{4\pi\alpha}$. In order to describe quarks, we need not only Dirac spinor $u^{(s)}(p)$ which gives its momentum and spin but also a three-element column vector c_i which gives its color shown in A.62.

$$c_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \text{ for red, } \quad c_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \text{ for blue, } \quad c_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ for green} \tag{A.62}$$

Usually, the quark changes color at quark gluon interaction vertex and gluon takes the difference of the color between incoming quark and outgoing quark like figure A.7. Each gluon carries one color and one anticolor, then there are nine kinds of gluons:

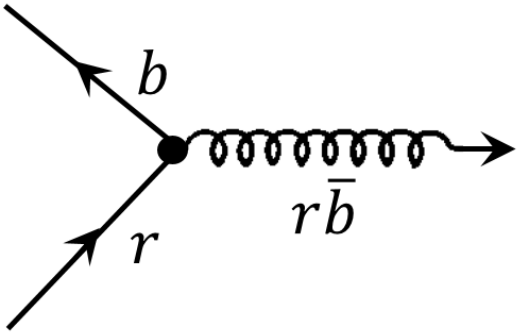

Figure A.7: The representation of the quark color changes at a quark-gluon vertex.

$r\bar{r}$, $r\bar{b}$, $r\bar{g}$, $b\bar{r}$, $b\bar{b}$, $b\bar{g}$, $g\bar{r}$, $g\bar{b}$, $g\bar{g}$. In terms of color $SU(3)$ symmetry, these nine states make up a "color octet" A.63 and a "color singlet" A.64. If the color singlet A.64 exist then it could be a mediator between two color singlets (e.g. a proton and a neutron) and giving rise to a long-range force with strong coupling, whereas we know the strong force is very short range. Therefore, there are evidently only eight kinds of gluons in our world.

$$\begin{aligned} |1\rangle &= (r\bar{b} + b\bar{r})/\sqrt{2} & |5\rangle &= -i(r\bar{g} - g\bar{r})/\sqrt{2} \\ |2\rangle &= -i(r\bar{b} - b\bar{r})/\sqrt{2} & |6\rangle &= (b\bar{g} + g\bar{b})/\sqrt{2} \\ |3\rangle &= (r\bar{r} - b\bar{b})/\sqrt{2} & |7\rangle &= -i(b\bar{g} - g\bar{b})/\sqrt{2} \\ |4\rangle &= (r\bar{g} + g\bar{r})/\sqrt{2} & |8\rangle &= (r\bar{r} + b\bar{b} - 2g\bar{g})/\sqrt{2} \end{aligned} \tag{A.63}$$

$$|9\rangle = (r\bar{r} + b\bar{b} + g\bar{g})/\sqrt{3} \tag{A.64}$$

Gluons are massless particles of spin 1, liking the photon A.2.2 we use polarization vector ϵ^{μ} to represent it which is orthogonal to the gluon momentum (see A.51). Besides we need an eight-element column vector in order to describe the color state of the gluon shown in A.65.

$$a^{\alpha} = \begin{pmatrix}1\\0\\0\\0\\0\\0\\0\end{pmatrix} \text{ for } |1\rangle\ , \quad a^{\beta} = \begin{pmatrix}0\\1\\0\\0\\0\\0\\0\end{pmatrix} \text{ for } |2\rangle\ , \quad a^{\gamma} = \begin{pmatrix}0\\0\\1\\0\\0\\0\\0\end{pmatrix} \text{ for } |3\rangle\ , \quad \text{and so on} \tag{A.65}$$

Before the statement of the Feynman rules for QCD, we need know the Gell-Mann "λ-matrices" which are shown in A.66 and the commutators of the λ matrices A.67 which define the "structure constants" $f^{\alpha\beta\gamma}$ of the $SU(3)$ group. The non-zero value of $f^{\alpha\beta\gamma}$ is shown in A.68.

$$\begin{aligned}
\lambda^1 &= \begin{pmatrix}0&1&0\\1&0&0\\0&0&0\end{pmatrix} \quad \lambda^2 = \begin{pmatrix}0&-i&0\\i&0&0\\0&0&0\end{pmatrix} \quad \lambda^3 = \begin{pmatrix}1&0&0\\0&-1&0\\0&0&0\end{pmatrix}\\
\lambda^4 &= \begin{pmatrix}0&0&1\\0&0&0\\1&0&0\end{pmatrix} \quad \lambda^5 = \begin{pmatrix}0&0&-i\\0&0&0\\i&0&0\end{pmatrix} \quad \lambda^6 = \begin{pmatrix}0&0&0\\0&0&1\\0&1&0\end{pmatrix}\\
\lambda^7 &= \begin{pmatrix}0&0&0\\0&0&-i\\0&i&0\end{pmatrix} \quad \lambda^8 = \begin{pmatrix}1&0&0\\0&1&0\\0&0&-2\end{pmatrix}
\end{aligned} \tag{A.66}$$

$$[\lambda^{\alpha}, \lambda^{\beta}] = 2if^{\alpha\beta\gamma}\lambda^{\gamma}, \quad \text{here summation over } \gamma \text{ from 1 to 8 is implied} \tag{A.67}$$

$$\begin{aligned}
&f^{123} = 1,\ f^{147} = f^{246} = f^{257} = f^{345} = f^{516} = f^{637} = \frac{1}{2},\\
&f^{458} = f^{678} = \frac{\sqrt{3}}{2}
\end{aligned} \tag{A.68}$$

Similar with QED A.2.3 the Feynaman rules for QCD tree level (for loop diagrams it require special rules) diagrams are explained in following:

1. **Draw the Feynman diagram** : Draw the corresponding Feynman diagrams for the process. Label four-momenta $p_1,\ p_2,\ ...$ for incoming and outgoing particles together with the corresponding spins $s_1,\ s_2,\ ....$ Label four-momenta $q_1,\ q_2,\ ...$ for internal particles. The arrow in the line represent the direction of motion of quark or gluon, while for antiquark it is opposite with its motion.
2. **External line** : For each external line write down $u^{(s)}(p)c$ ($\bar{u}^{(s)}(p)c^{\dagger}$) for incoming (outgoing) quark or $\bar{v}^{(s)}(p)c^{\dagger}$ ($v^{(s)}(p)c$) for incoming (outgoing) antiquark where c represents the color of the correspond quark or antiquark. Besides write down $\epsilon_{\mu}(p)a^{\alpha}$ ($\epsilon^{*}_{\mu}(p)a^{\alpha *}$) for incoming (outgoing) gluon.

3. **Vertex factor** : Each vertex introduces a factor and there are vertex for quark-gluon or gluon-gluon interaction. Because the gluons carry color (in contrast to the photon, which is electrically neutral), they can couple with each other and there is a three-gluon vertex and a four-gluon vertex shown in figure A.8. The value of each vertex is shown in A.69.

$$\begin{aligned}
&\text{Quark}-\text{gluon (left plot of A.8)}: \frac{-ig_s}{2}\lambda^{\alpha}\gamma^{\mu}\\
&\text{Three gluon (middle plot of A.8)}: -g_s f^{\alpha\beta\gamma}[g_{\mu\nu}(k_1-k_2)_{\lambda}+g_{\nu\lambda}(k_2-k_3)_{\mu}+g_{\lambda\mu}(k_3-k_1)_{\nu}]\\
&\text{, here the gluon momenta } (k_1,k_2,k_3) \text{ are assumed to point into the vertex,}\\
&\text{if any point outward, need change the sign.}\\
&\text{Four gluon (right plot of A.8)}: -ig_s^2[f^{\alpha\beta\eta}f^{\gamma\delta\eta}(g_{\mu\lambda}g_{\nu\rho}-g_{\mu\rho}g_{\nu\lambda})\\
&+f^{\alpha\delta\eta}f^{\beta\gamma\eta}(g_{\mu\nu}g_{\lambda\rho}-g_{\mu\lambda}g_{\nu\rho})+f^{\alpha\gamma\eta}f^{\delta\beta\eta}(g_{\mu\rho}g_{\nu\lambda}-g_{\mu\nu}g_{\lambda\rho})] \quad (\text{summation over } \eta \text{ implied.})
\end{aligned} \tag{A.69}$$

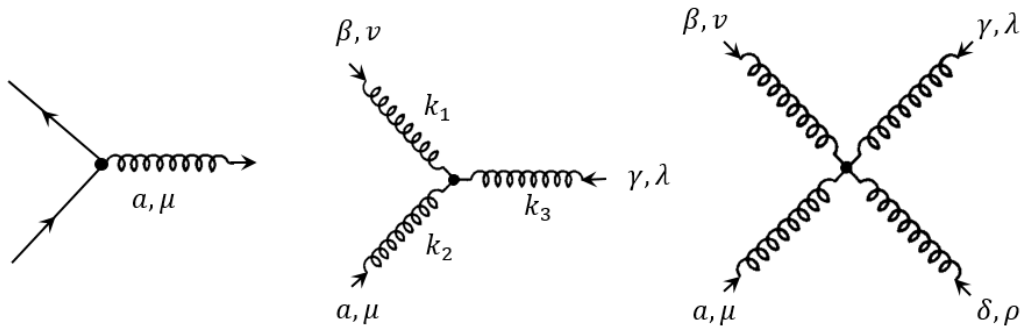

Figure A.8: The representation of the coupling between gluons.

4. **Propagator** : For each internal line write down a factor which is shown in A.70.

$$\begin{aligned}
&\text{For quark and antiquark}: \frac{i(\not{q}+mc)}{q^2-m^2c^2},\ \not{q}\equiv q^{\mu}\gamma_{\mu}\\
&\text{For gluon}: \frac{-ig_{\mu\nu}\delta^{\alpha\beta}}{q^2}
\end{aligned} \tag{A.70}$$

5. **Conservation of energy and momentum** : For each vertex write down a δ function with the form of $(2\pi)^4\delta^4(k_1+k_2+k_3)$. Here the k_i are the four-momentum of the particle which involved in the vertex, if the arrow of line k_i is toward to the vertex then put a positive sign in front of k_i otherwise put a negative sign (opposite for antiquark). This factor imposes the conservation of energy and momentum at each vertex, since the δ function is zero unless the sum of incoming momenta equals the sum of outgoing momenta.
6. **Integration over internal momenta** : For each internal line write down a factor $\frac{1}{(2\pi)^4 d^4 q_i}$ and integrate over all internal momenta.
7. **Cancel the δ function** : Erase $(2\pi)^4\delta^4(p_1+p_2+...+p_n)$ in the result which enforcing overall conservation of energy and momentum. Finally what remains is $-i\mathcal{M}$.

8. **Antisymmetrization** : Including a minus sign between diagrams which differ only in interchange of two incoming (or outgoing) quarks or antiquarks, or of an incoming quark with an outgoing antiquark (or vice versa).

Using this Feynman rules we can calculate the $\mathcal{M}$ of the tree level QCD process. For example for $u + \bar{d} \to u\bar{d}$ process which is shown in figure A.9 it gives A.71 and finally the $\mathcal{M}$ is given in A.72.

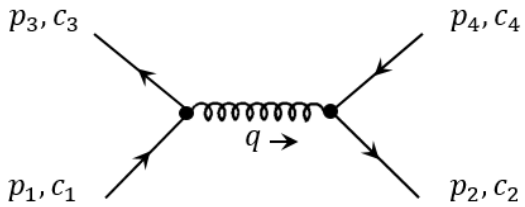

Figure A.9: The representation of the quark-antiquark interaction.

$$
\begin{aligned}
-i\mathcal{M} &= [\bar{u}(3)c_3^\dagger][-i\frac{g_s}{2}\lambda^\alpha\gamma^\mu][u(1)c_1][\frac{-ig_{\mu\nu}\delta^{\alpha\beta}}{q^2}] \\
&* [\bar{v}(2)c_2^\dagger][-i\frac{g_s}{2}\lambda^\beta\gamma^\nu][v(4)c_4]
\end{aligned}
\tag{A.71}
$$

$$
\mathcal{M} = \frac{-g_s^2}{4}\frac{1}{q^2}[\bar{u}(3)\gamma^\mu u(1)][\bar{v}(2)\gamma_\mu v(4)](c_3^\dagger\lambda^\alpha c_1)(c_2^\dagger\lambda^\alpha c_4), \text{(summation over } \alpha \text{ is implied.)} \tag{A.72}
$$

A.3.3 Asymptotic Freedom

As we know for QED we using renormalization method to remove the divergence from higher level loop diagram, similar we can use it for QCD loop diagram like left plot of figure A.10 just thinking color as electrical charge. However, for QCD we also have virtual gluon loop which is shown in right plot of figure A.10, this changes the story. It turns out that the gluon contribution works in the opposite direction which behaviours like "antiscreening" and the formula for the running coupling constant in QCD is A.73, where n is the number of colors (3 for the Standard Model) and f is the number of flavors (6 in the Standard Model). When $11n > 2f$ the antiscreening effect will be dominate and the coupling will decrease with increasing $|q^2|$ which means at short distances the strong force becomes relatively weak and is call "asymptotic freedom". Because of the asymptotic freedom we can use perturbation theory and Feynman calculus in QCD and it is a basic ingredient in the theory of quarkonium, otherwise the Chromodynamics would have been abandoned.

$$
\alpha_s(|q^2|) = \frac{\alpha_s(\mu^2)}{1 + (\alpha_s(\mu^2)/12\pi)(11n - 2f)\ln(|q^2|/\mu^2)} \quad (|q^2| \gg \mu^2) \tag{A.73}
$$

In order to make formula A.73 more compact, the new variable Λ is defined in A.74. Then the running coupling constant can be express in terms of a single parameter shown in A.75.

$$
\ln\Lambda^2 = \ln\mu^2 - 12\pi/[(11n - 2f)\alpha_s(\mu^2)] \tag{A.74}
$$

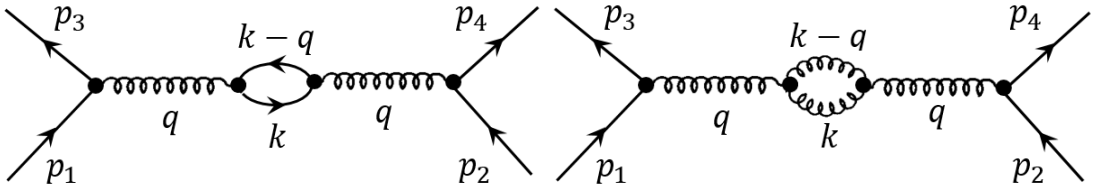

Figure A.10: The representation of the Feynman diagram with one quark-antiquark loop in the gluon propagator (left) and the Feynman diagram with one gluon-gluon loop in the gluon propagator (right).

$$\alpha_s(|q^2|) = \frac{12\pi}{(11n-2f)\ln(|q^2|/\Lambda^2)} \quad (|q^2| \gg \Lambda^2) \tag{A.75}$$

Unfortunately, it is hard to determine Λ precisely from experimental data, but Λc appears to lie somewhere in the range 100 MeV $< \Lambda c <$ 500 MeV.

Difference with the QED coupling which varies only minutely over the accessible energy range A.5, the variation in the QCD coupling is substantial shown in figure A.11. When q^2 gets close to Λ the $\alpha_s(|q^2|)$ will increase to infinite, So for the process which q^2 close to or below Λ we have difficulty to calculate $\mathcal{M}$. Therefor for low q^2 or long-range QCD process there are a lot of work need to be done.

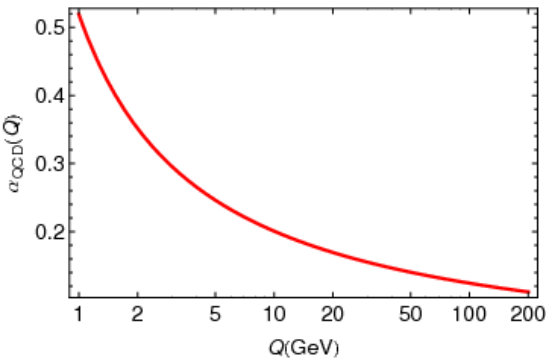

Figure A.11: The momentum transfer evolution of QCD coupling constant α_{QCD} [160]

A.4 Groups: a brief introduction

The formulas in this section are from book [54].

Group theory is a branch of mathematics that underlines the treatment of symmetry. It plays an very important role particle physics. Some basic concepts and terminologies for the group theory will be given in this section. Here we use operators of 3D rotation ($R(\theta_1, \theta_2, \theta_3)$), the θ_i means the rotated angle respect to each axis) as an example. Suppose all the rotation form a rotation group and the "multiplication" of two rotations R_2R_1 means first do rotation R_1 then do R_2, then this group must full fill following four requirements: 1, the R_2R_1 should equal to an other group element R_3. This is correct for the rotation because two rotations can be expressed by one rotation; 2, the existence of identity element. This is satisfied by the existence of a rotation $R(0,0,0)$; 3, the existence of an inverse element for each element. This is satisfied by the existence of $R(-\theta_1, -\theta_2, -\theta_3)$ with respect to $R(\theta_1, \theta_2, \theta_3)$; 4, the equality of $(R_1R_2)R_3$ and $R_1(R_2R_3)$. This is also cor-

rect obviously. It should be noticed that it is not necessary to have $R_1R_2 = R_2R_1$ for a group.

As we know the experiment result or the physics is not dependent with the orientation of the laboratory. Therefore the rotation group is a symmetry group. Suppose the transformation of system state by the rotation R is

$$|\psi\rangle \to |\psi'\rangle = U|\psi\rangle$$

and the probability of a system described by $|\psi\rangle$ will be found in state $|\varphi\rangle$ must unchanged by R,

$$|\langle\varphi|\psi\rangle|^2 = |\langle\varphi'|\psi'\rangle|^2 = |\langle\varphi|U^\dagger U|\psi\rangle|^2.$$

So the $U^\dagger U = 1$ and U is unitary operator. Besides, the rotation R can be represented by U. Because the rotation group is Lie group which has a crucial property that is each rotation can by expressed by a product of succession of infinitesimal rotations. For instance, the rotation on z axis with infinitesimal angle ϵ, to first order ϵ, can by expressed by

$$U = 1 - i\epsilon J_z$$

where the J_z is called the generator of the rotations about the z axis. Now

$$1 = U^\dagger U = (1 + i\epsilon J_z^\dagger)(1 - i\epsilon J_z) = 1 + i\epsilon(J_z^\dagger - J_z) + 0(\epsilon^2),$$

it gives $J_z^\dagger = J_z$ which means J_z is hermitian and hence is a observable.

A.4.1 SU(2) groups

In the lowest-dimension nontrivial representation of the rotation group, the generators can be written as

$$J_i = \frac{1}{2}\sigma_i, \text{ with } i = 1, 2, 3,$$

where the σ_i is Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{A.76}$$

and the base states for this representation is conventionally chosen to be the eigenvectors of σ_3, that is

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{A.77}$$

which describes spin $\frac{1}{2}$ particle with spin up spin and spin down in z direction or it can describe the isospin state of proton $(+\frac{1}{2})$ and neutron $(-\frac{1}{2})$ because they are indistinguishable in nuclear interaction. The unitary transformation matrices can by written as

$$U(\theta_i) = e^{-i\theta_i\sigma_i/2}.$$

The set of all 2×2 unitary matrices is known as the group $U(2)$. However, the $U(2)$ is larger then the $U(\theta_i)$, because all elements in $U(\theta_i)$ has zero trace. For any zero trace hermitian matrix σ, it can be proved that

$$\det(\sigma) = e^{iTr(\sigma)} = 1.$$

Therefore, all traceless 2×2 unitary matrices $U(\theta_i)$ form a subgroup of $U(2)$, which is call $SU(2)$ group, the "S" means special.

A.4.2 SU(3) groups

All 3×3 unitary matrices with det$U = 1$ form a $SU(3)$ group. There are $3^2 - 1 = 8$ linearly independent hermitian generators.

The fundamental representation of $SU(3)$ group is a triplet. The three colors (R, G, B) of a quark form a fundamental representation of a $SU(3)$ group. In this representation, the 8 generators are 3×3 traceless unitary matrices and are traditionally called λ_i with i from 1 to 8. There are only two diagonal generator matrices which are

$$\lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \lambda_8 = \sqrt{\frac{1}{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 0 \end{pmatrix} \tag{A.78}$$

with simultaneous eigenvectors

$$R = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad G = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{A.79}$$

These states are plotted in Figure A.12 in term of their λ_3 and λ_8 eigenvalues. The figure also shows how the base states are transformed using the λ_i generators. The λ_i are known as the Gell-Mann matrices (see A.3.2), the λ_1, λ_2, and λ_3 correspond to the three Pauli matrices, which means they exhibit explicitly one $SU(2)$ subgroup of $SU(3)$.

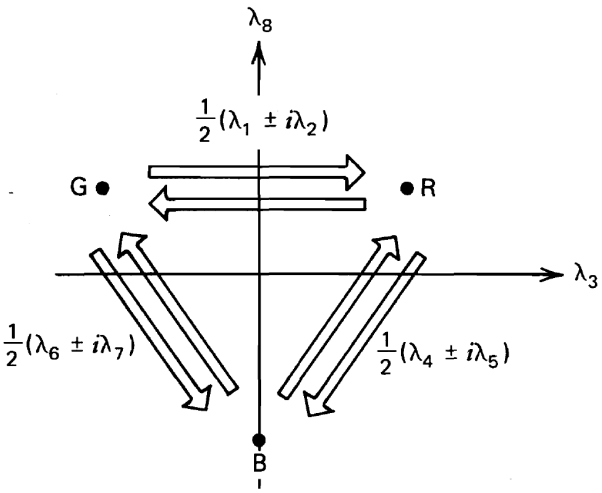

Figure A.12: The action of the generators on fundamental representations of $SU(3)$ [54].

Appendix B

The Appendices for Experiment

B.1 Mass resolution fit results

As described in 5.4, 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. B.1 for the Barrel-Barrel category (BB) and in fig. B.2 for the Barrel-Endcap (BE) category.

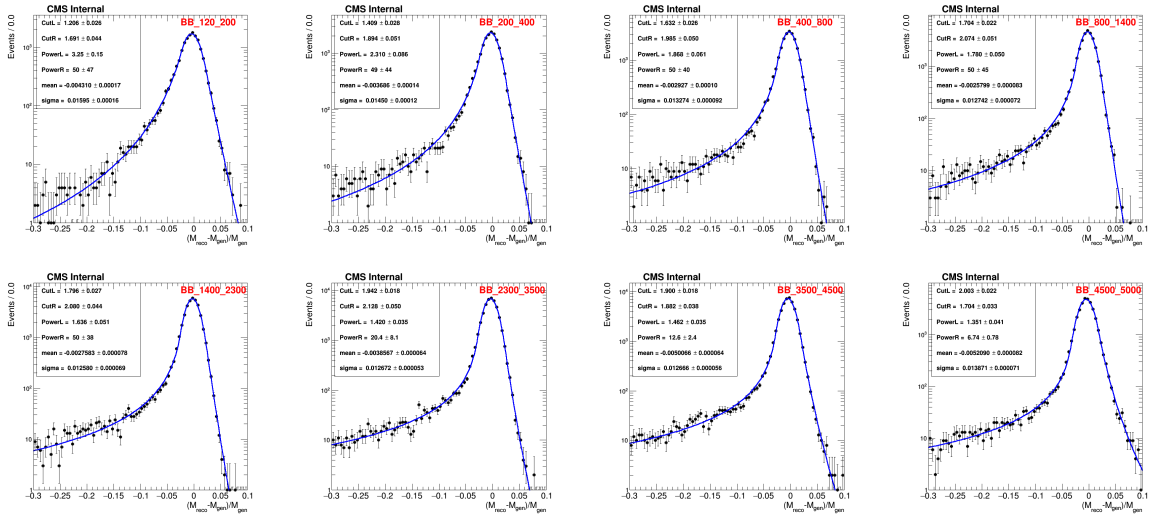

Figure B.1: Fit results of the $\frac{(M_{reco}-M_{gen})}{M_{gen}}$ histograms in the BB category per M_{gen} bins.

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. B.3 and B.4).

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 B.3 and B.4.

B.2 For 2016 HEEP ID scale factor

B.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

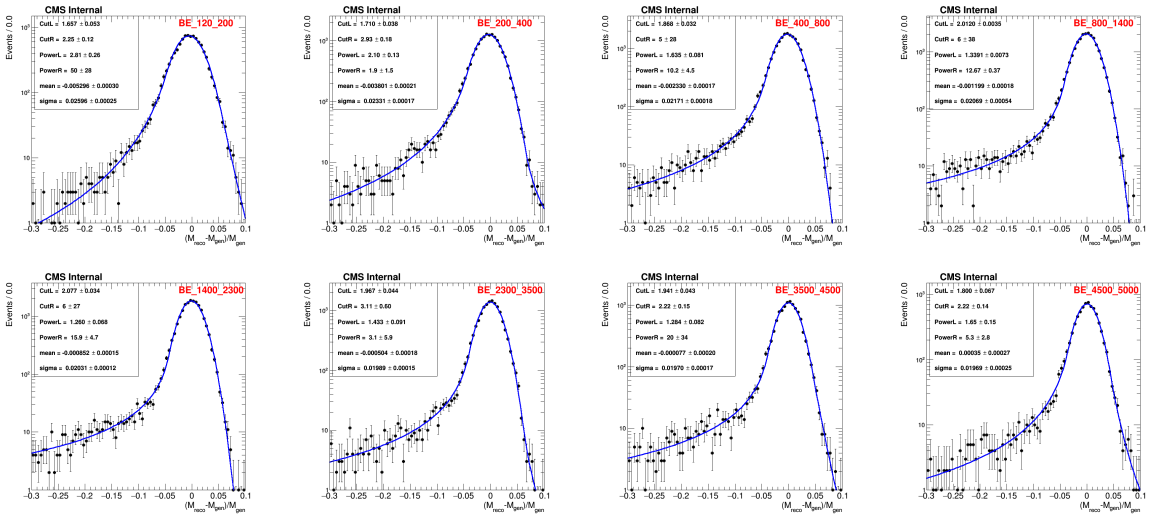

Figure B.2: Fit results of the $\frac{(M_{reco}-M_{gen})}{M_{gen}}$ histograms in the BE category per M_{gen} bins.

N-3 efficiencies and scale factors are shown as functions of E_T, η of the probe and of N_{vtx} in figures B.5-B.15.

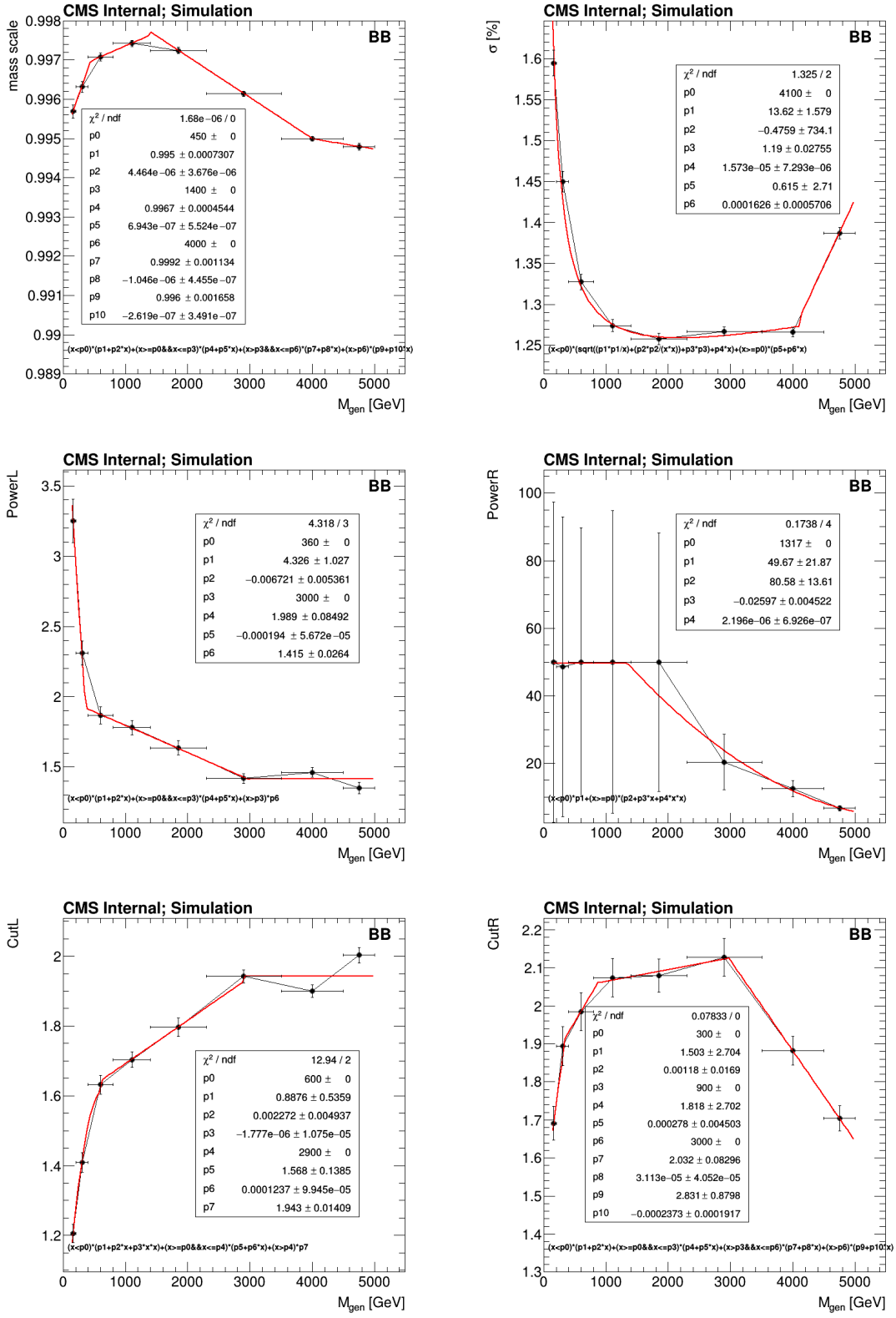

Figure B.3: Parameters of the dCB as a function of M_{gen} for BB category.

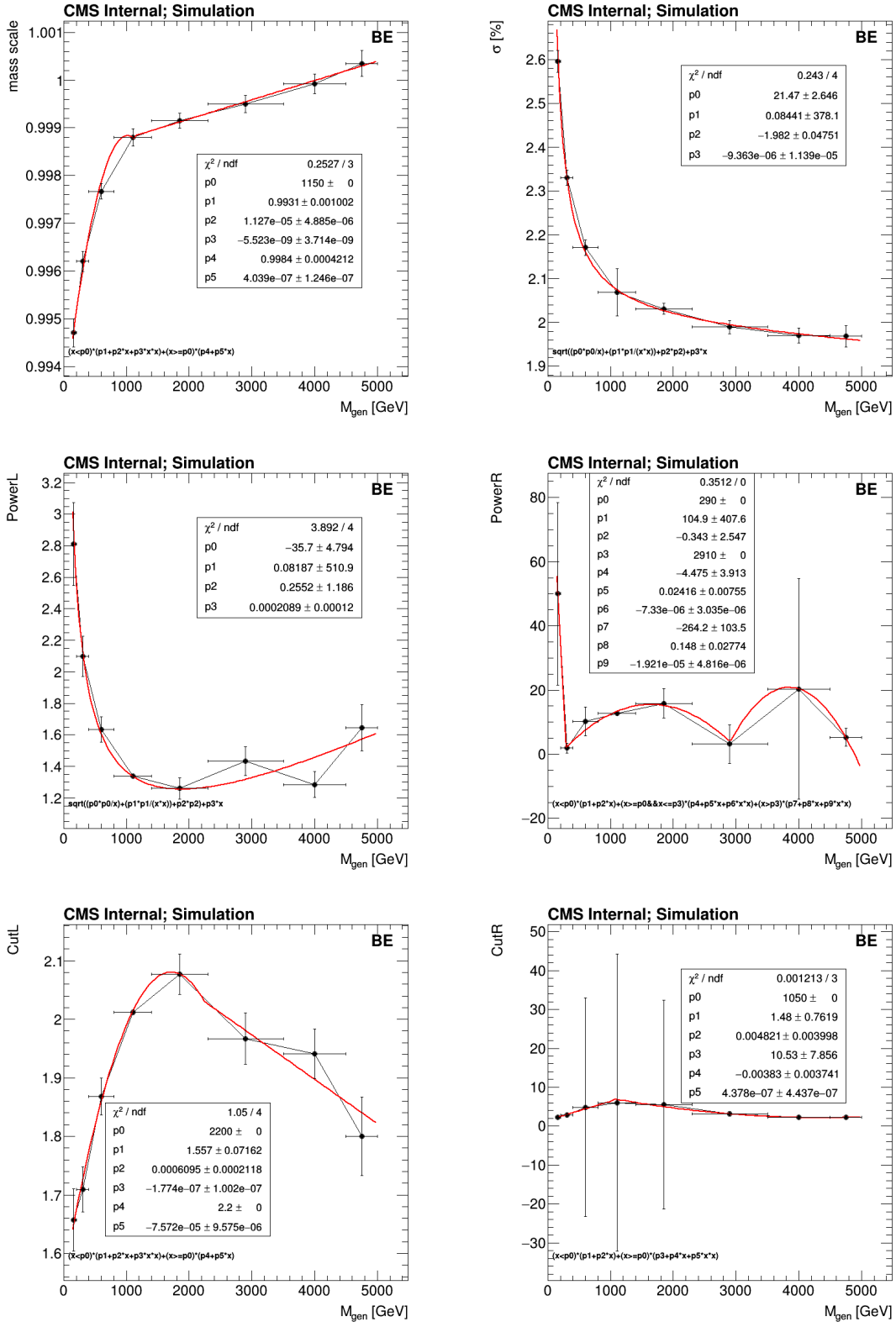

Figure B.4: Parameters of the dCB as a function of M_{gen} for BE category.

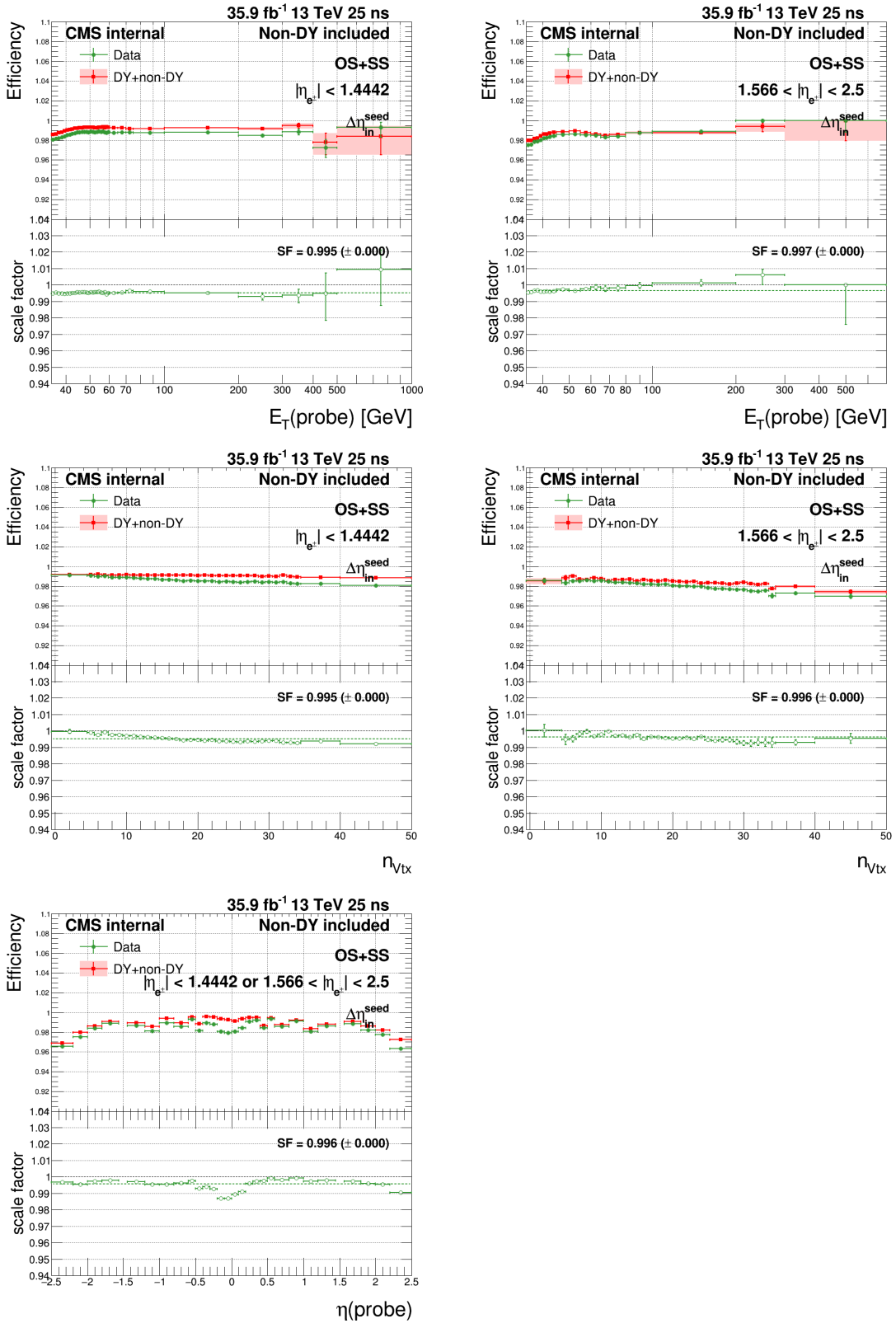

Figure B.5: $\Delta\eta_{in}^{seed}$ N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

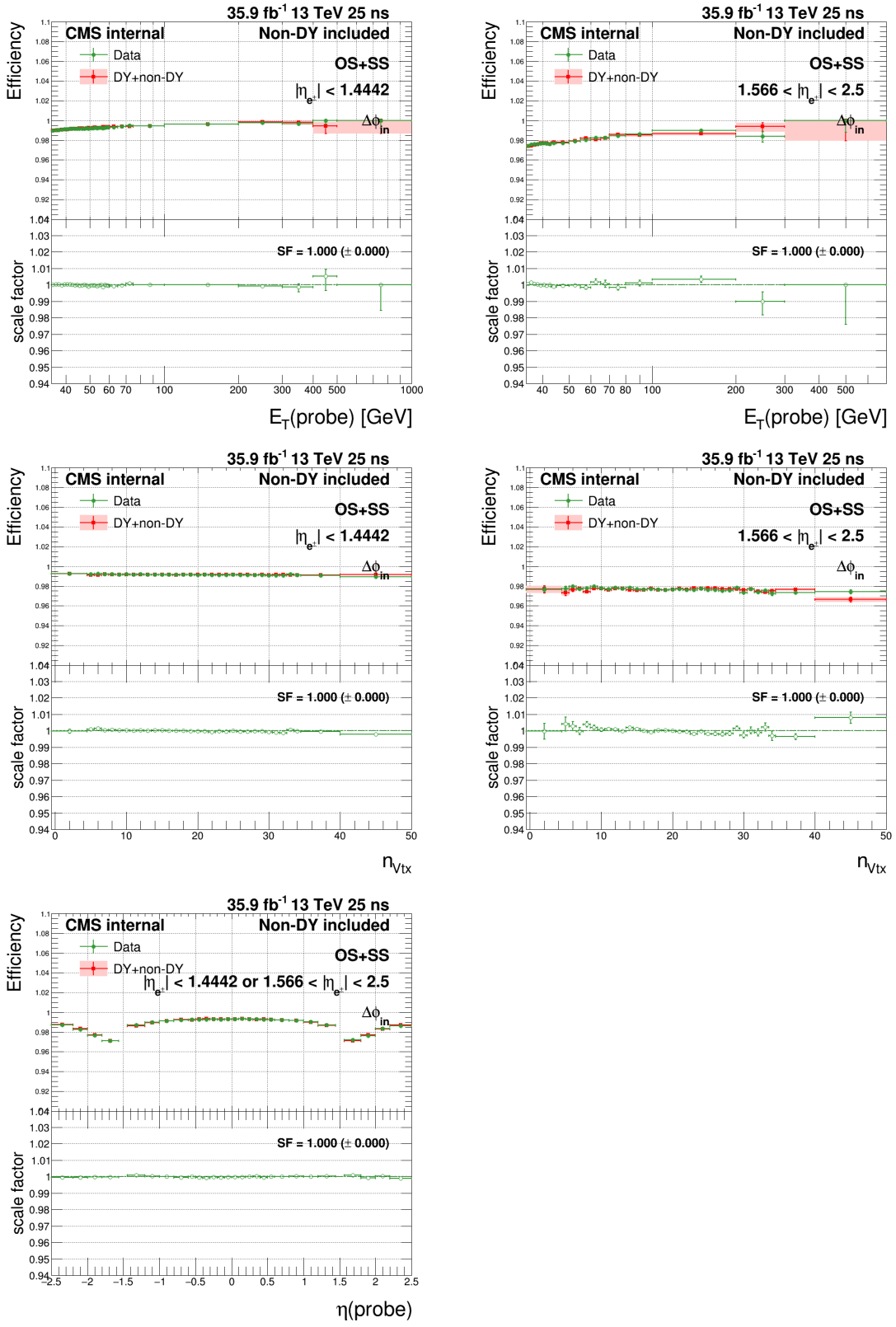

Figure B.6: $\Delta\phi_{in}$ N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

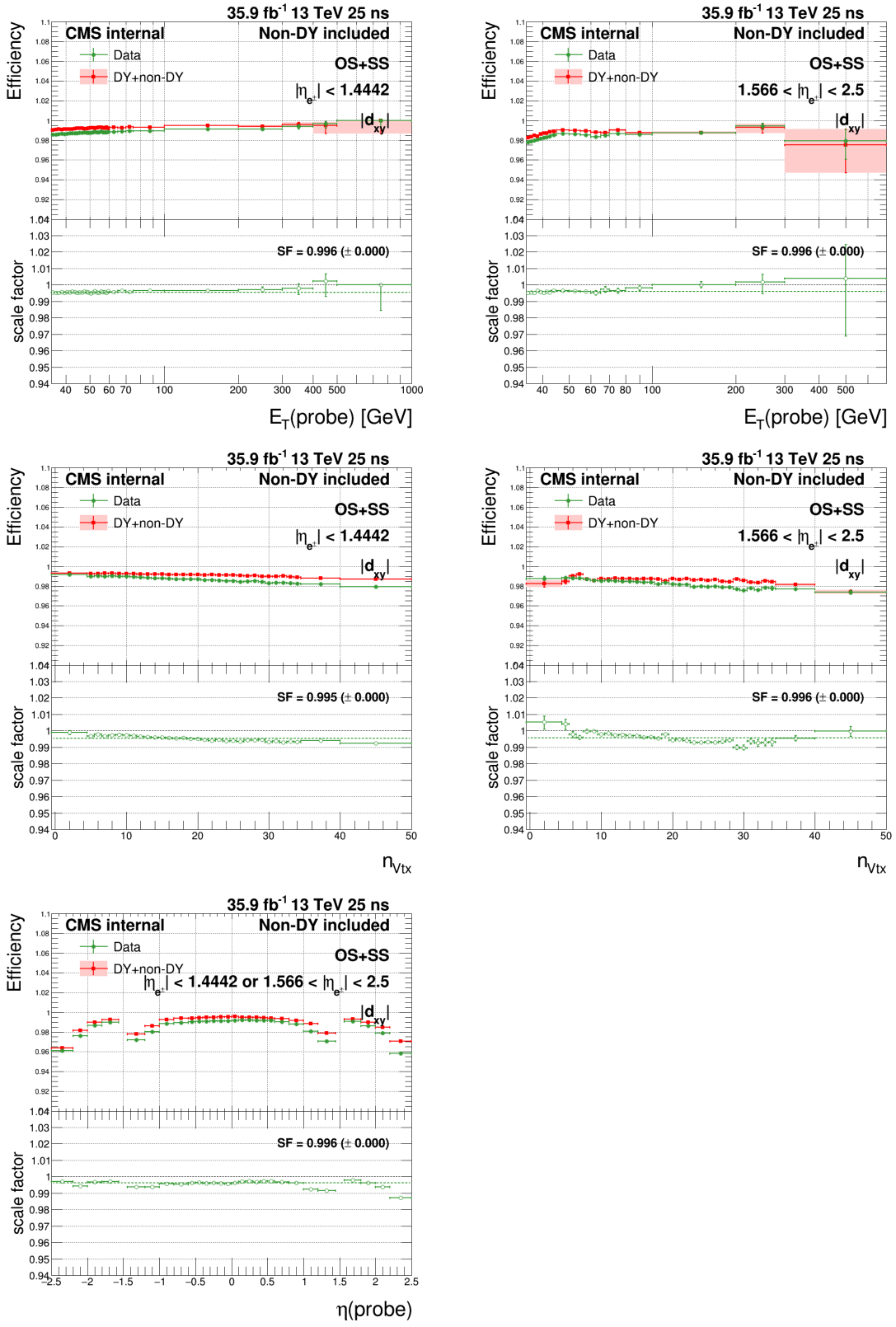

Figure B.7: $|d_{xy}|$ N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

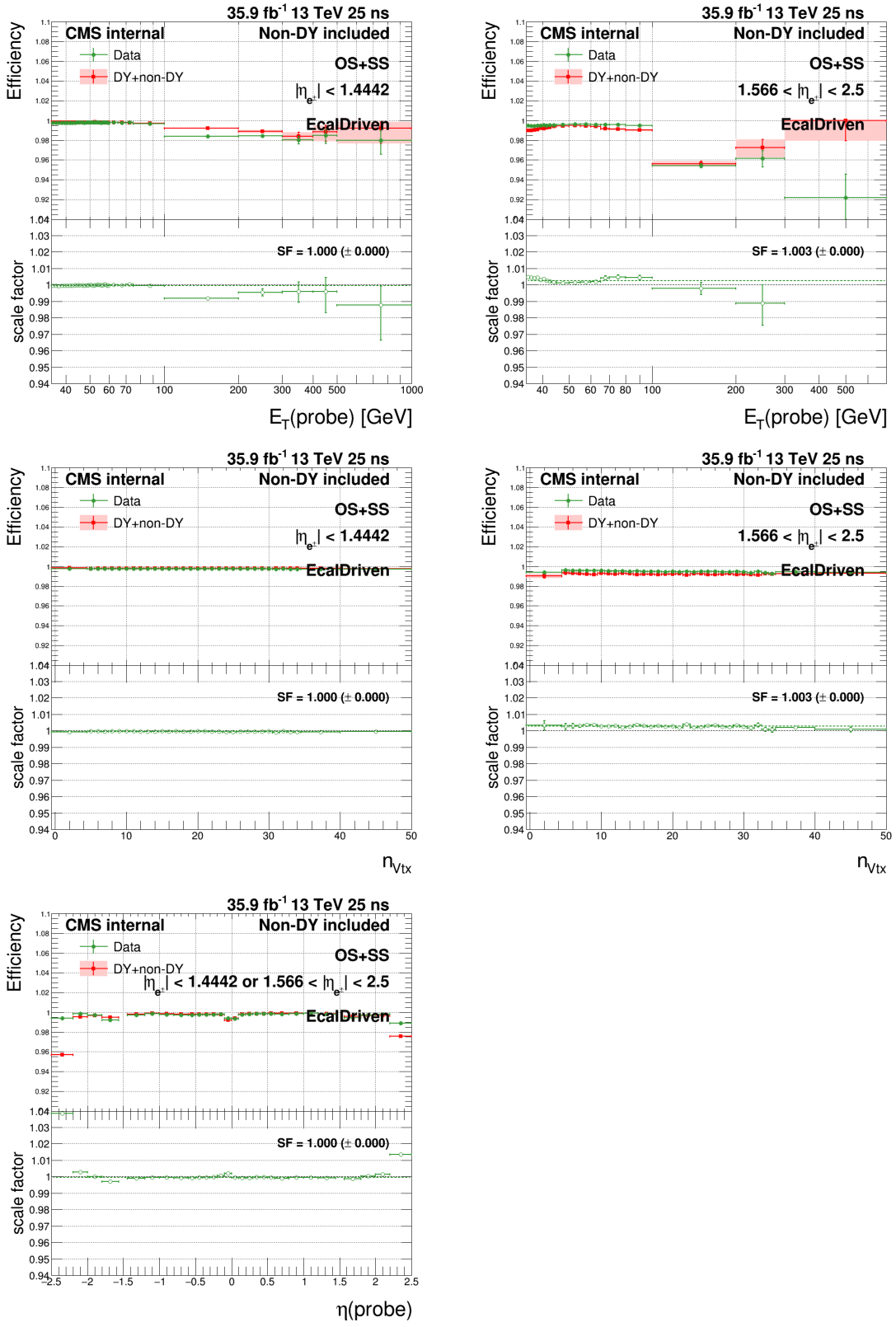

Figure B.8: *EcalDriven* N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

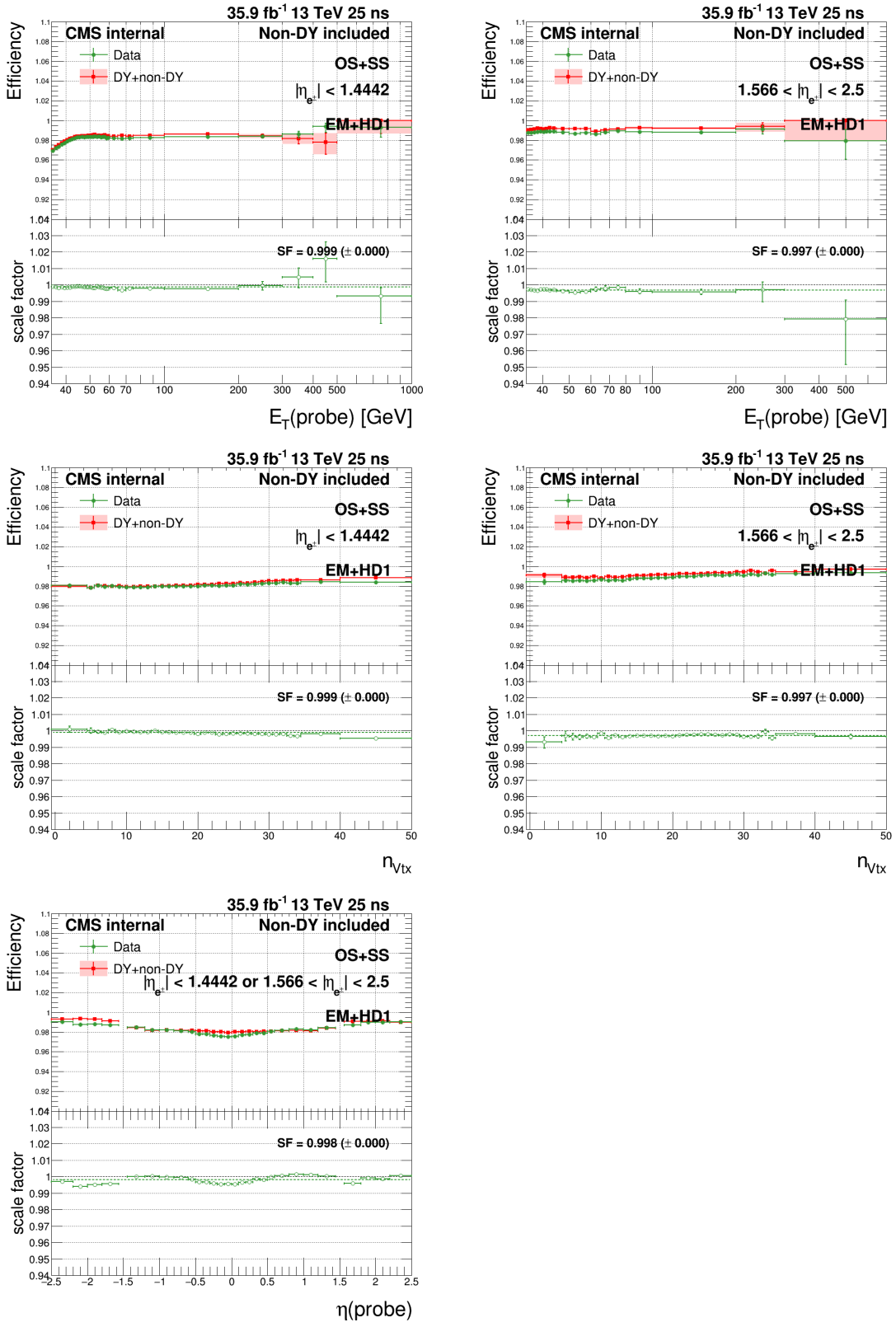

Figure B.9: $EM + HD1$ N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

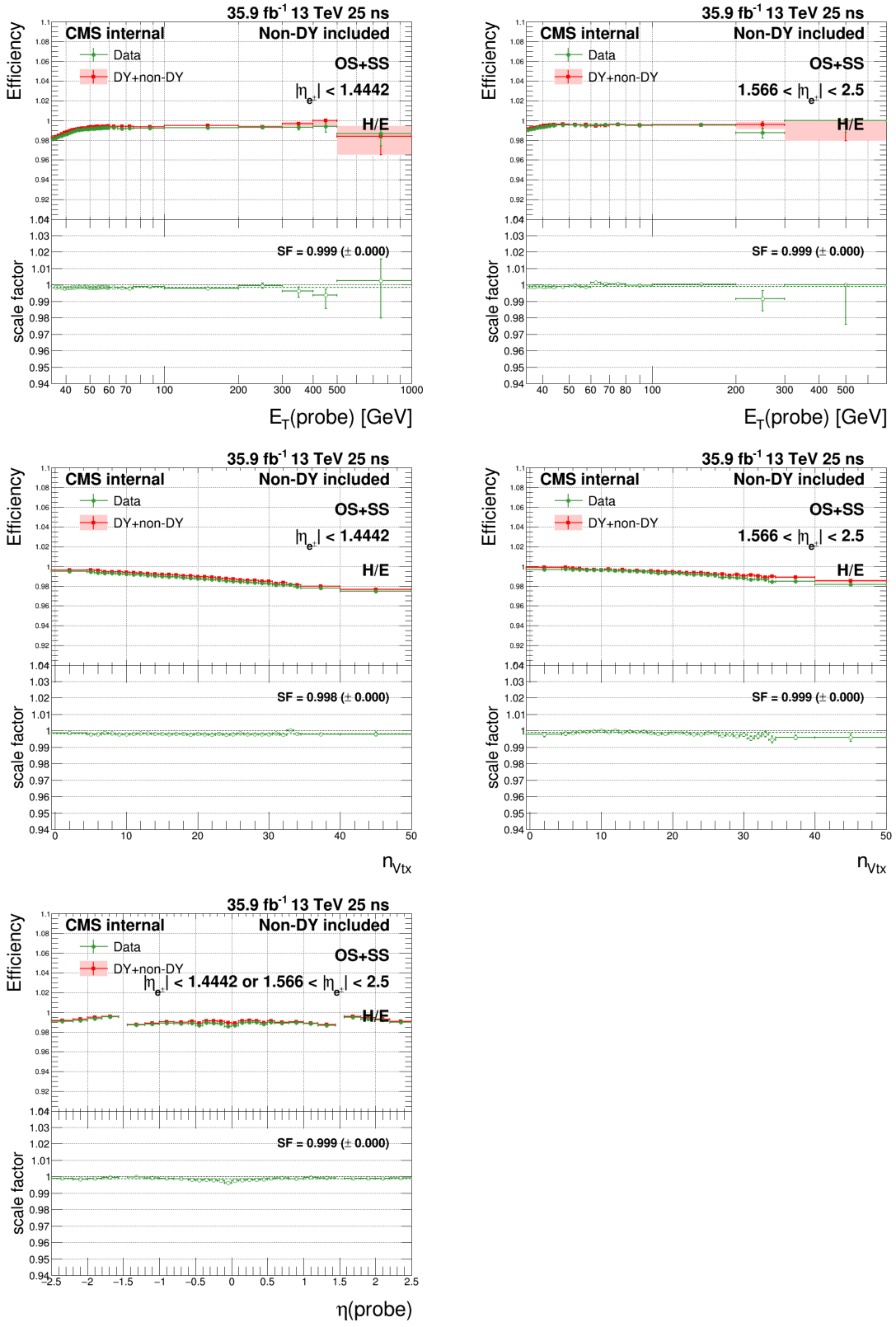

Figure B.10: H/E N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

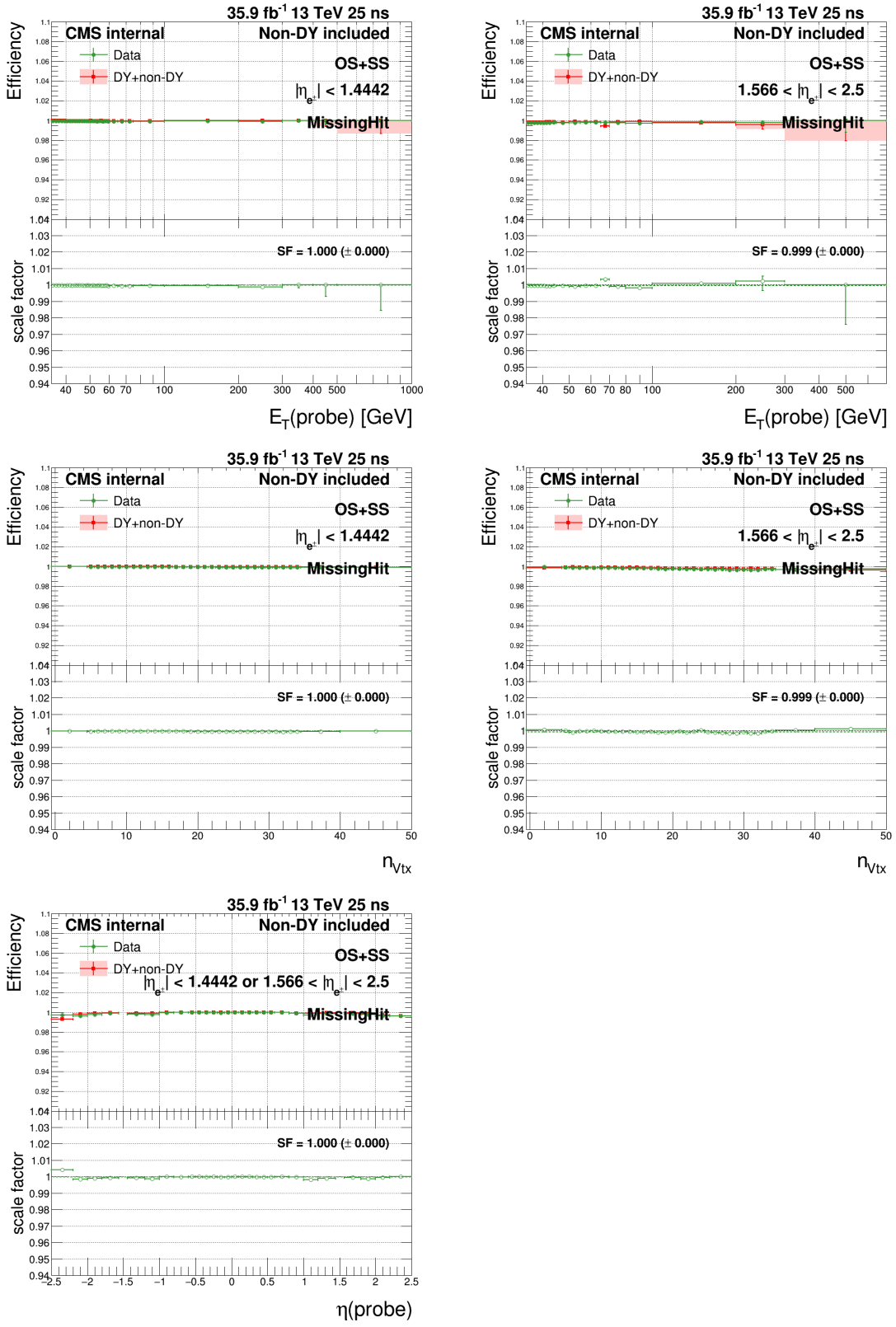

Figure B.11: *MissingHit* N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

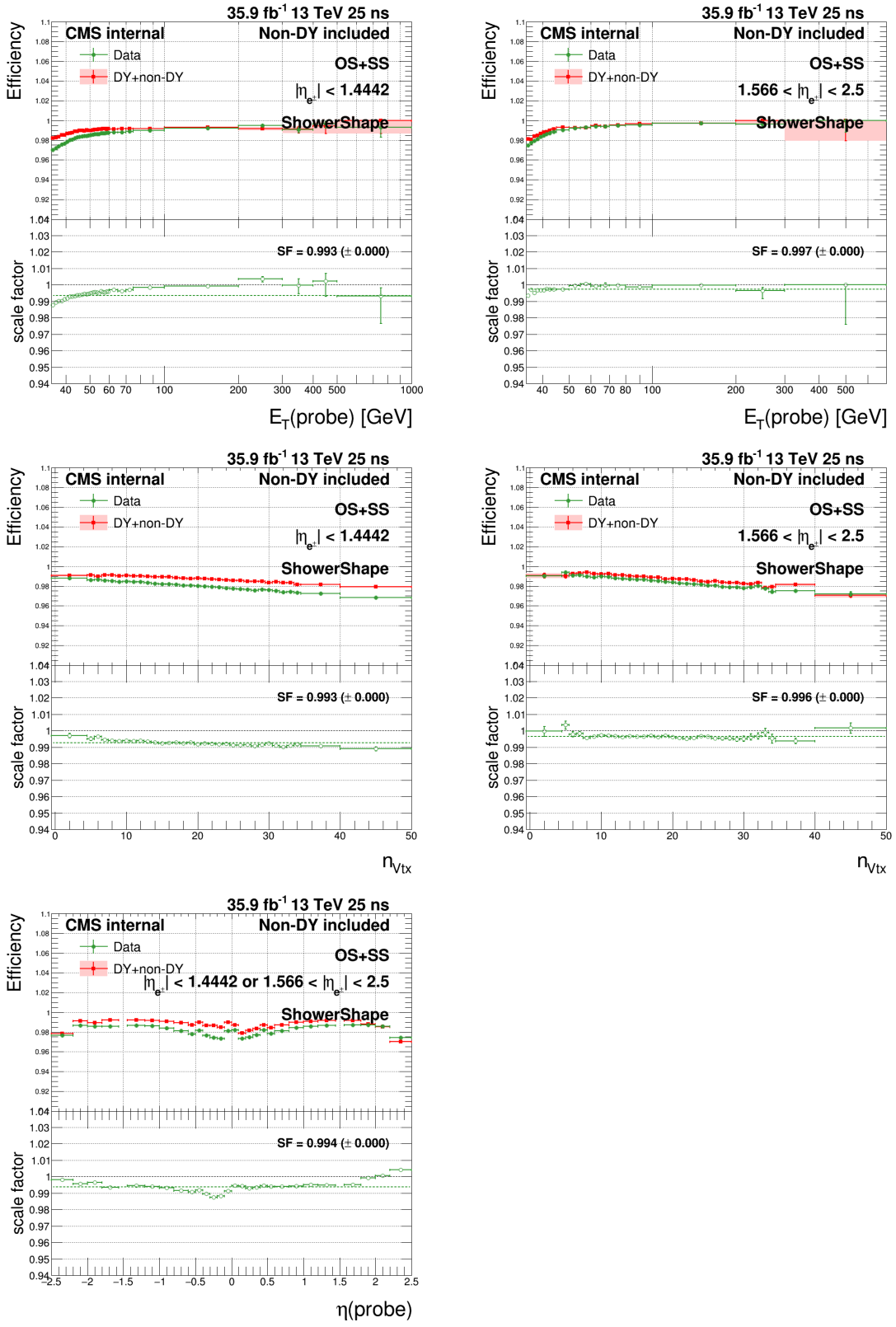

Figure B.12: *ShowerShape* N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

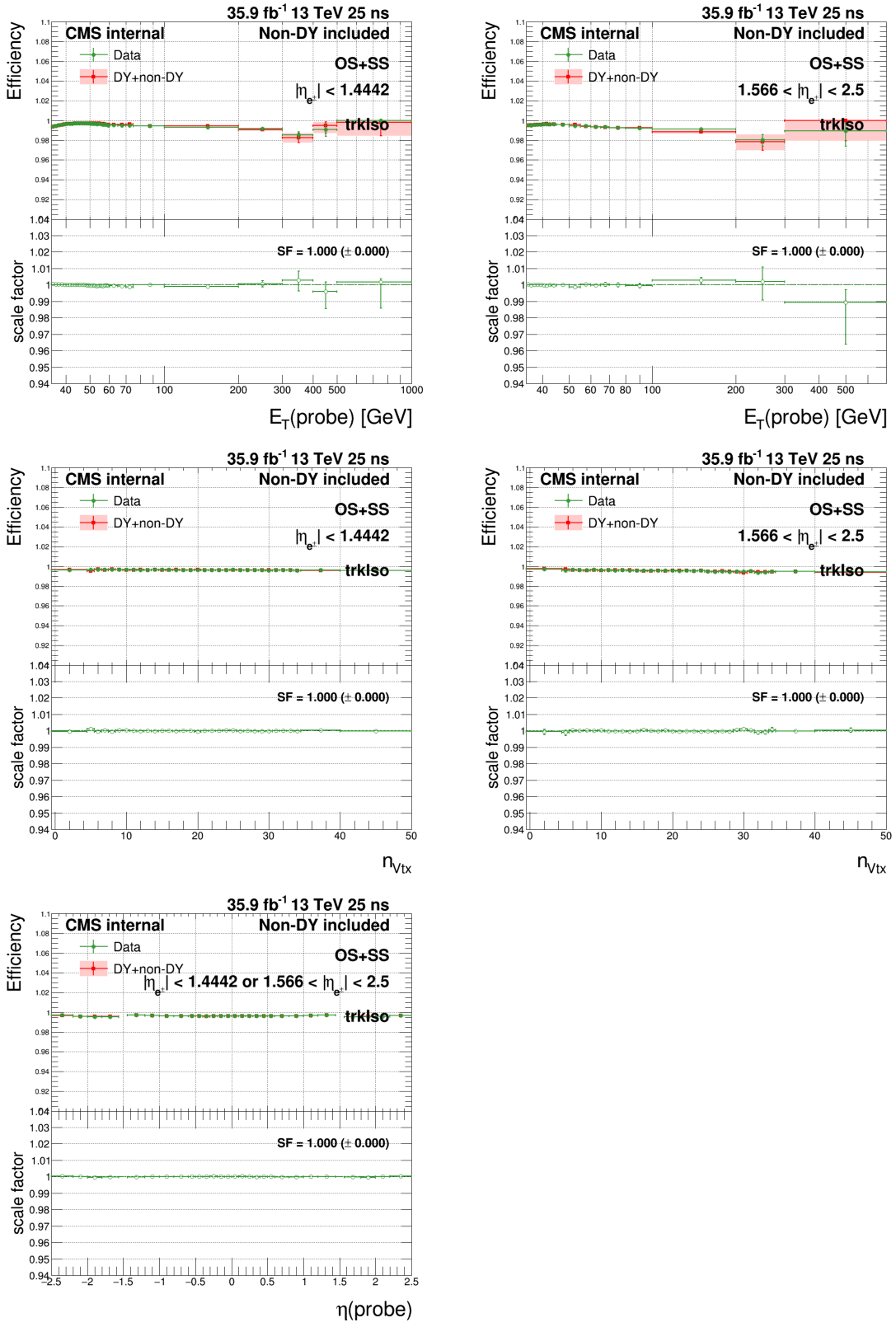

Figure B.13: *trackisolation* N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

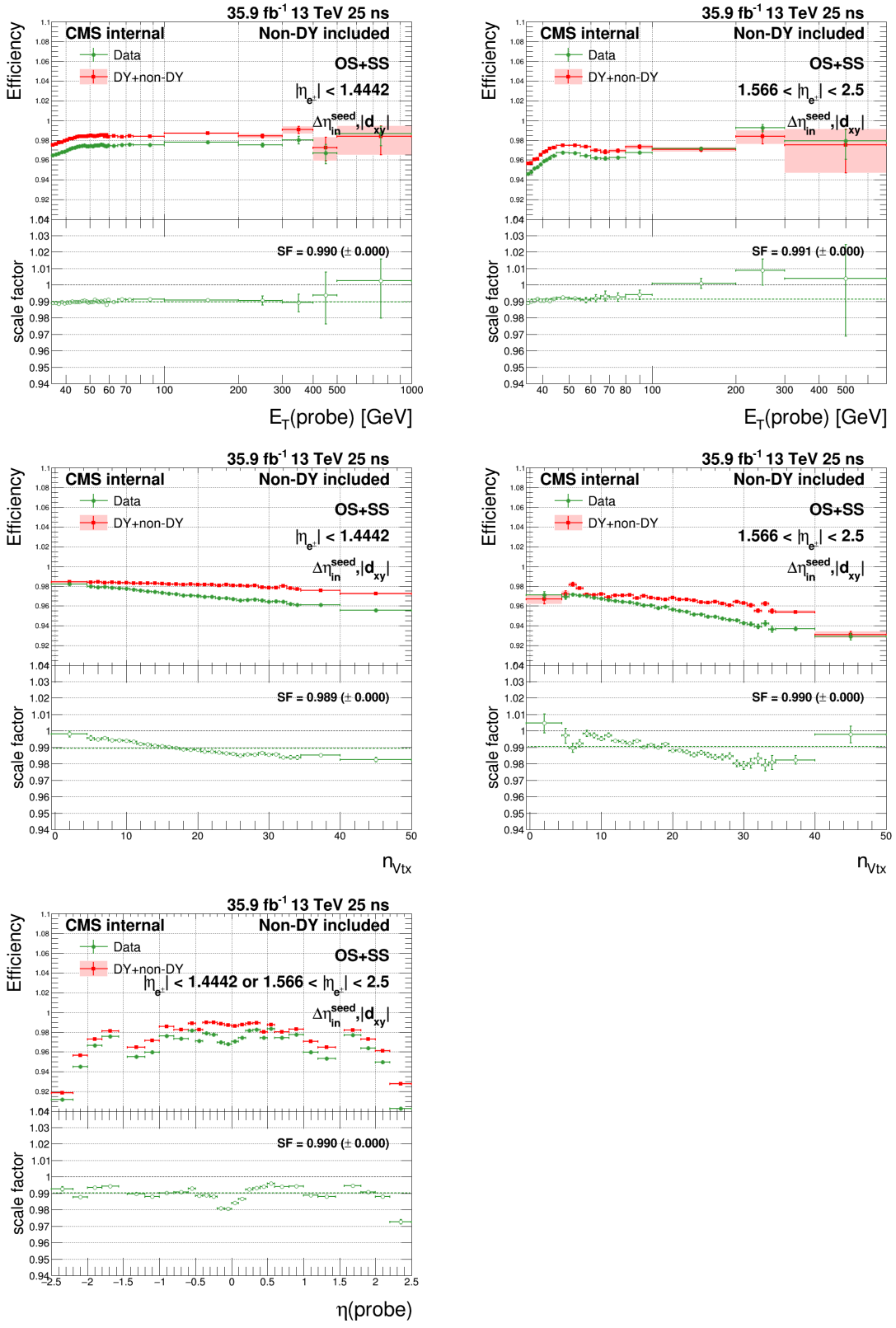

Figure B.14: $\Delta\eta_{in}^{seed}$ and $|d_{xy}|$ N-2 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

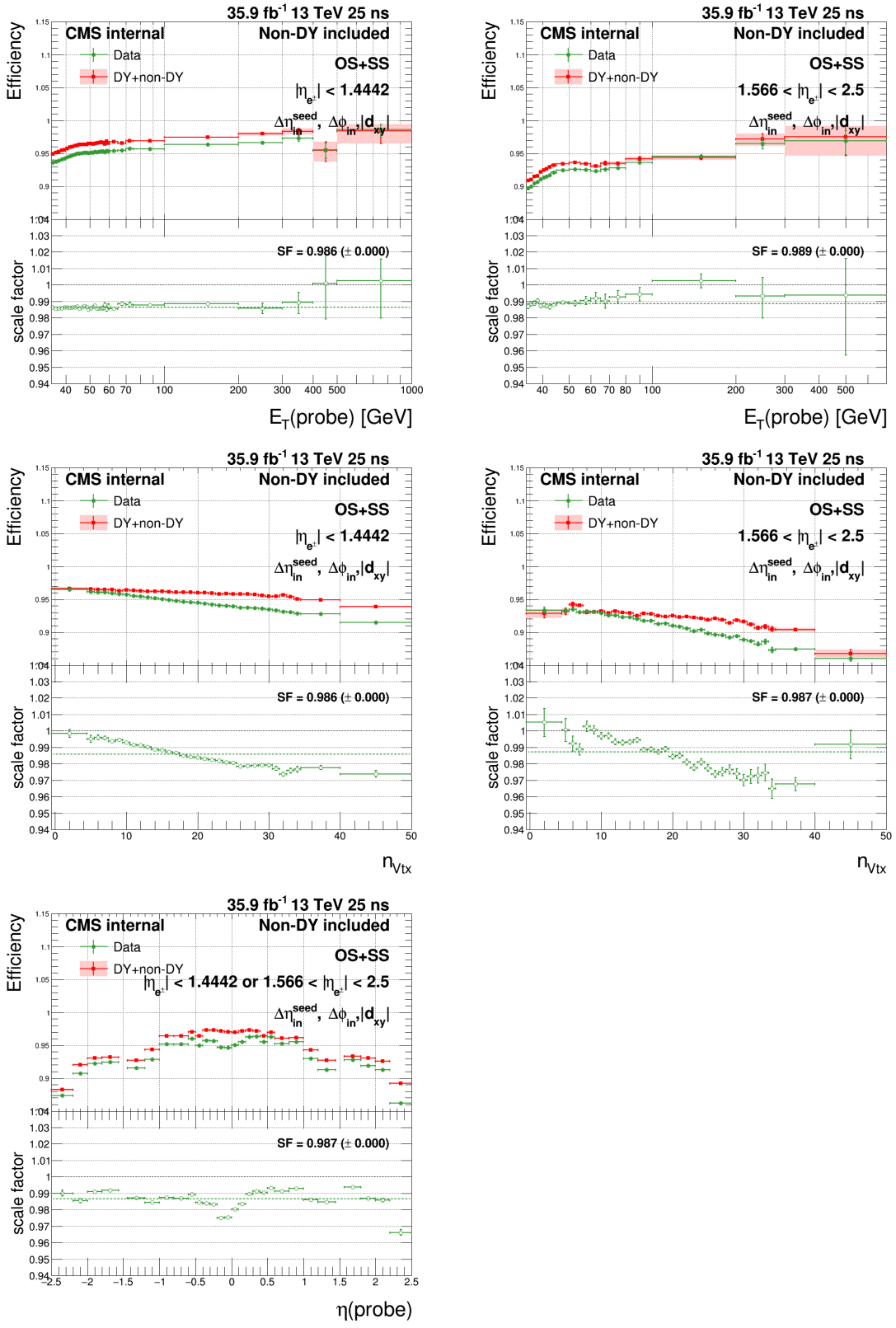

Figure B.15: $\Delta\eta_{in}^{seed}$, $|d_{xy}|$ and $\Delta\phi_{in}$ N-3 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), N_{vtx} (middle) and probe η (bottom).

B.2.2 HEEP efficiency versus η for different E_T bins

As can be seen in Figure 5.29, 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 B.16. 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 B.1

Table B.1: HEEP scale factors for different η bins.

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

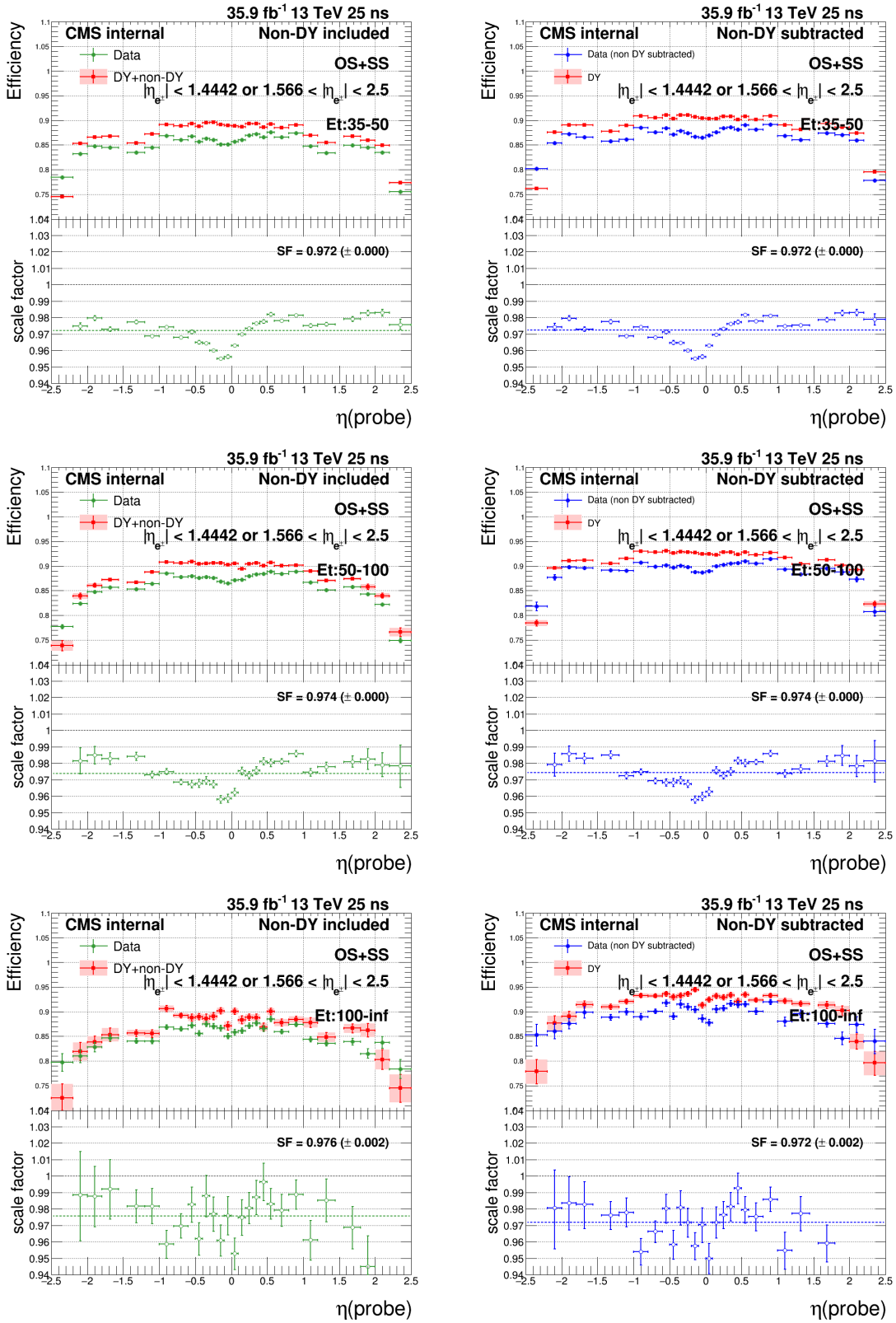

Figure B.16: Efficiencies and scale factors in MC and data for E_T of probe bin $35-50$(GeV) (top), $50-100$(GeV) (middle) and >100(GeV) (bottom) where the non-DY processes are included (left) and subtracted (right) as functions of probe η.

B.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 B.17. The HEEP efficiency and scale factor for different E_T are shown in figures B.18.

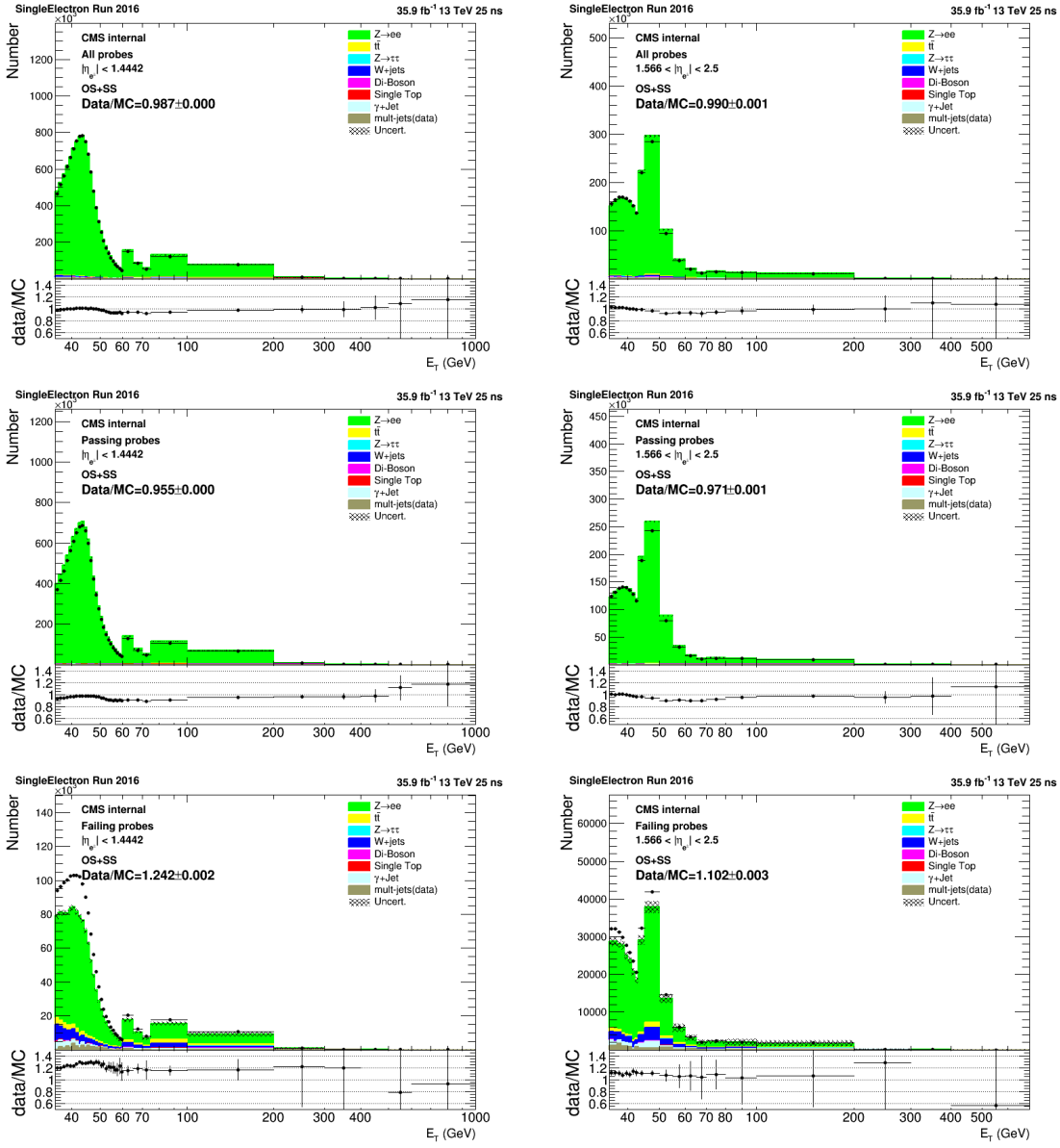

Figure B.17: E_T of probe in the barrel (left) and endcap (right) where all the probes are included (top), only passing probes are included (middle) and only failed probes are included (bottom).

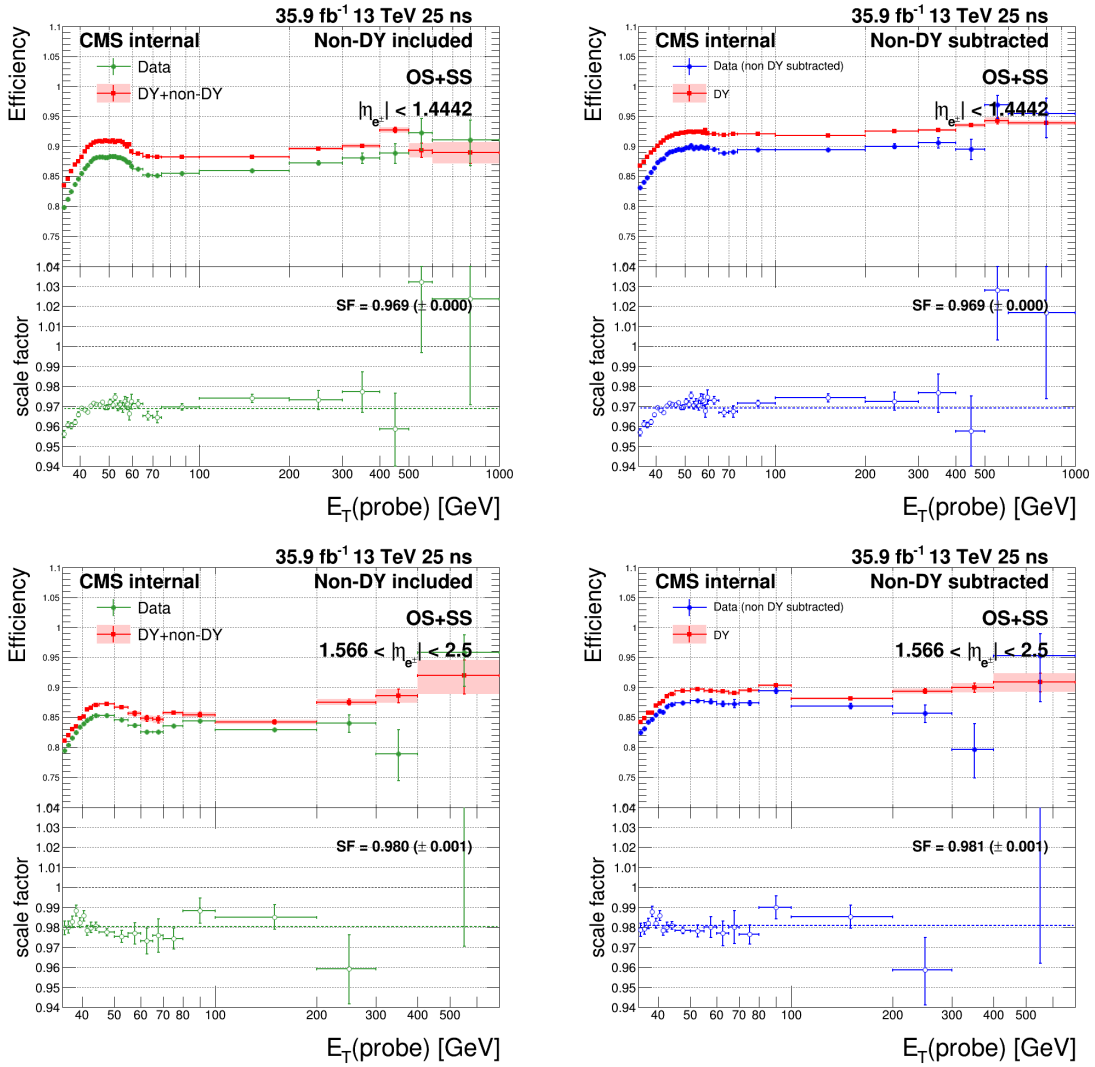

Figure B.18: Efficiencies and scale factors in MC and data in the barrel (top) and endcap (bottom) where the non-DY processes are included (left) and subtracted (right) as functions of probe E_T.

B.2.4 HEEP efficiecny 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 B.19-B.21. From Figure B.19 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 B.20 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 B.21 one can see ZToEE mass bin powheg and DYJetsToLL amcatnlo agree well.

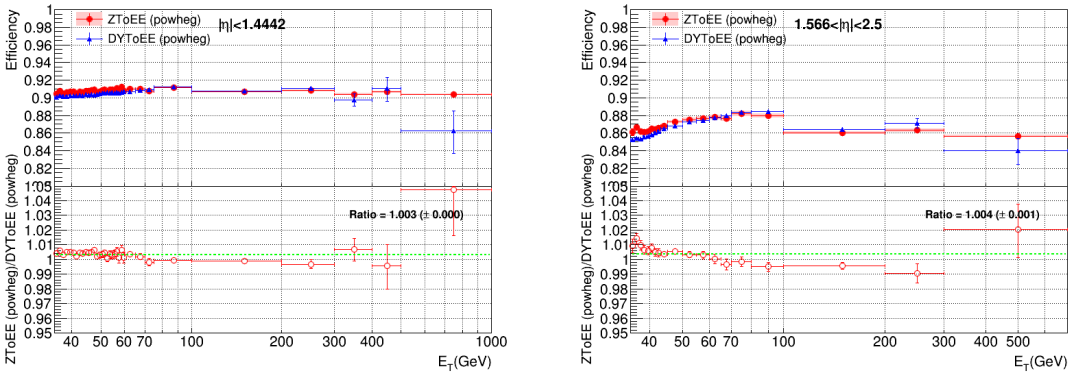

Figure B.19: The HEEP efficiencies for mc matched electron for DYToEE powheg sample and ZToEE powheg sample for barrel (left) and endcap (right) for different generated E_T of electron.

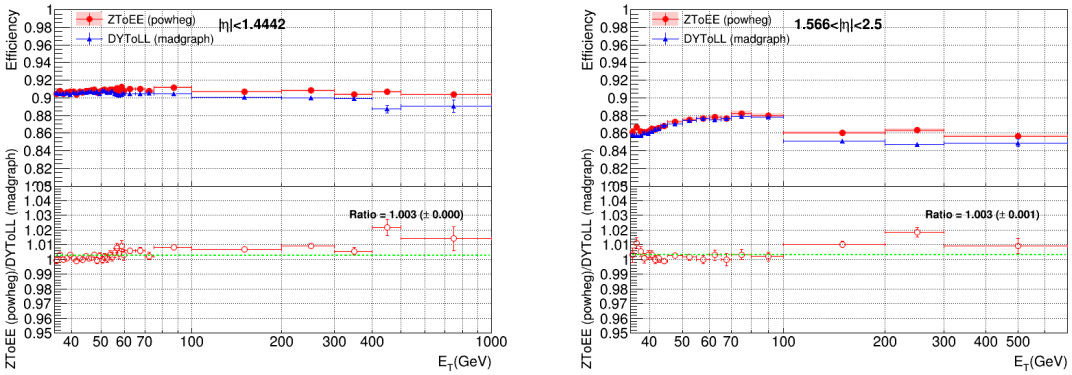

Figure B.20: The HEEP efficiencies for mc matched electron for DYToLL madgraph sample and ZToEE powheg sample for barrel (left) and endcap (right) for different generated E_T of electron.

B.3 For 2017 HEEP ID scale factor

B.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 B.22-B.32.

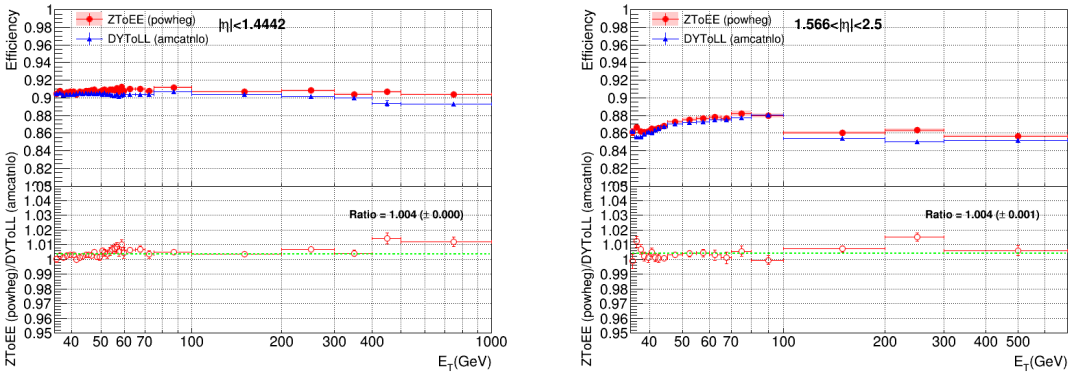

Figure B.21: The HEEP efficiencies for mc matched electron for DYToLL amcatnlo sample and ZToEE powheg sample for barrel (left) and endcap (right) for different generated E_T of electron.

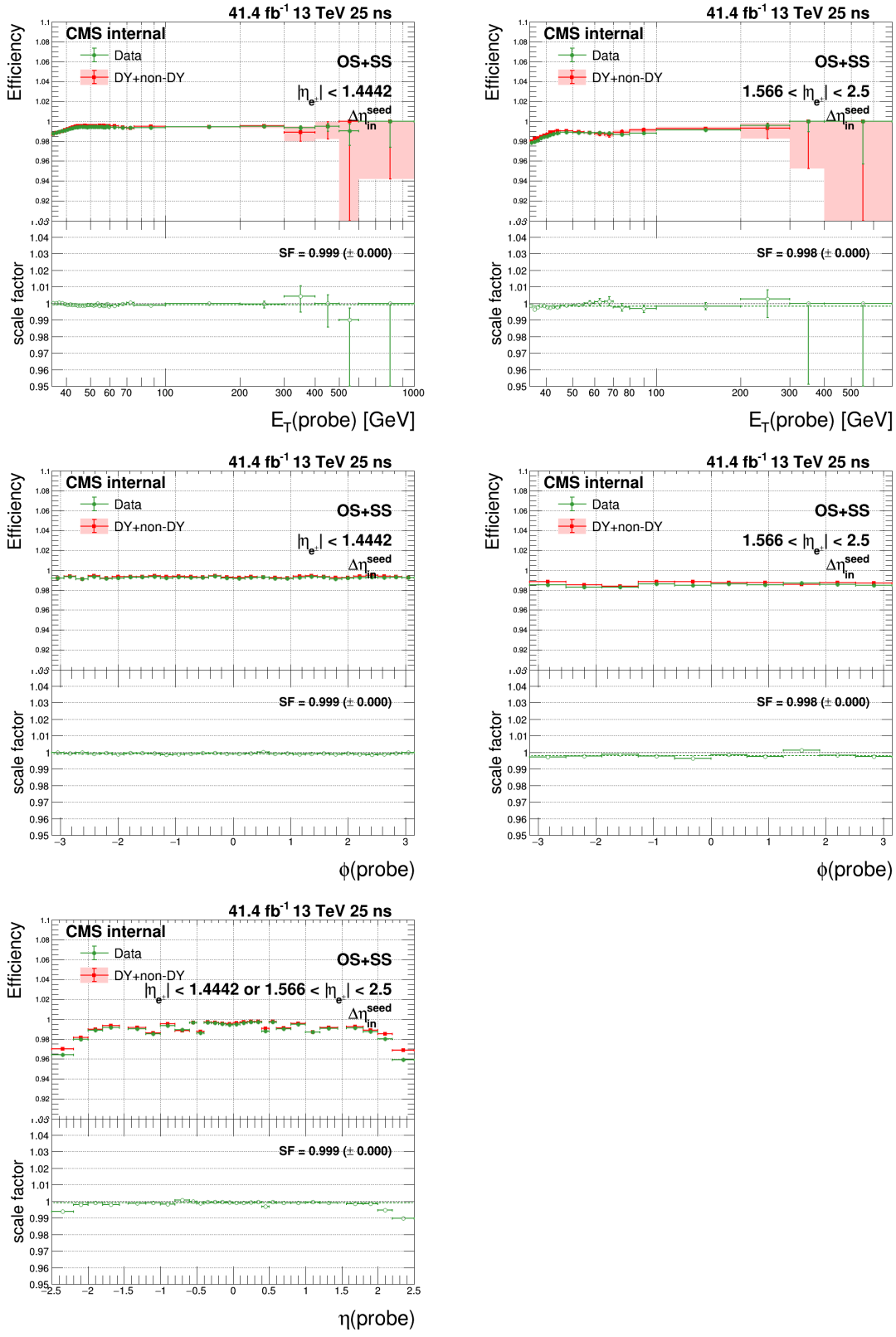

Figure B.22: $\Delta\eta_{in}^{seed}$ N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

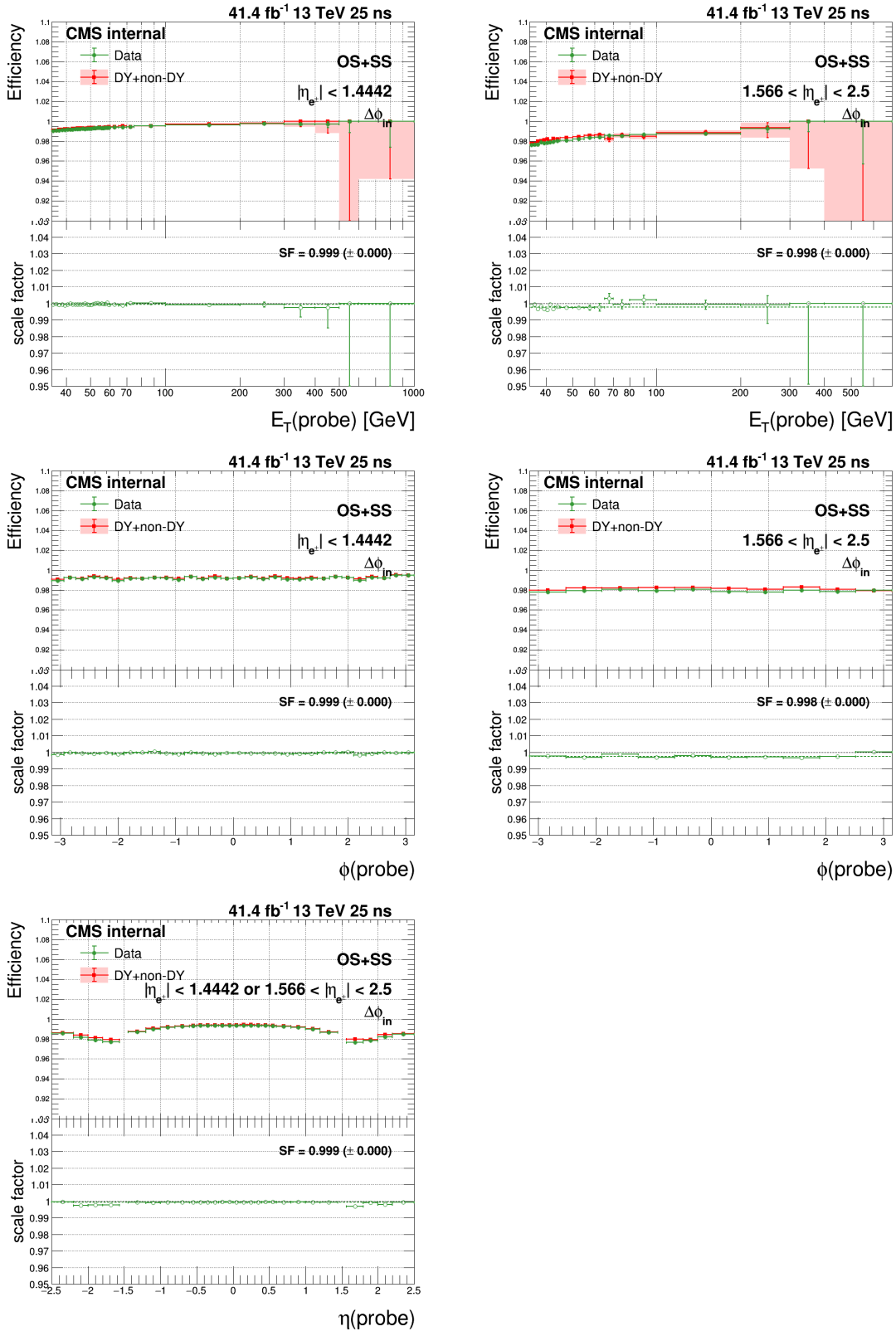

Figure B.23: $\Delta\phi_{in}$ N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

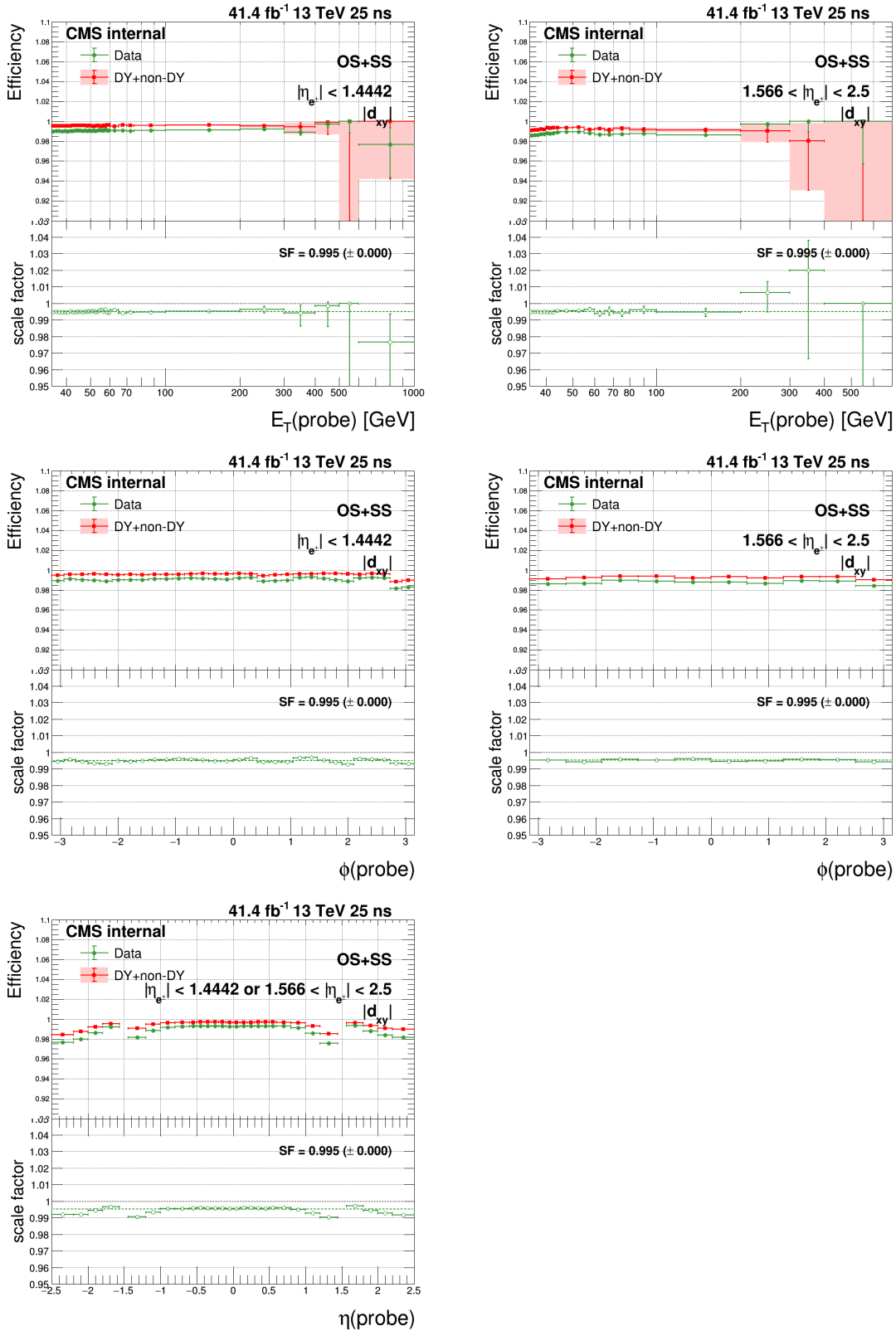

Figure B.24: $|d_{xy}|$ N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

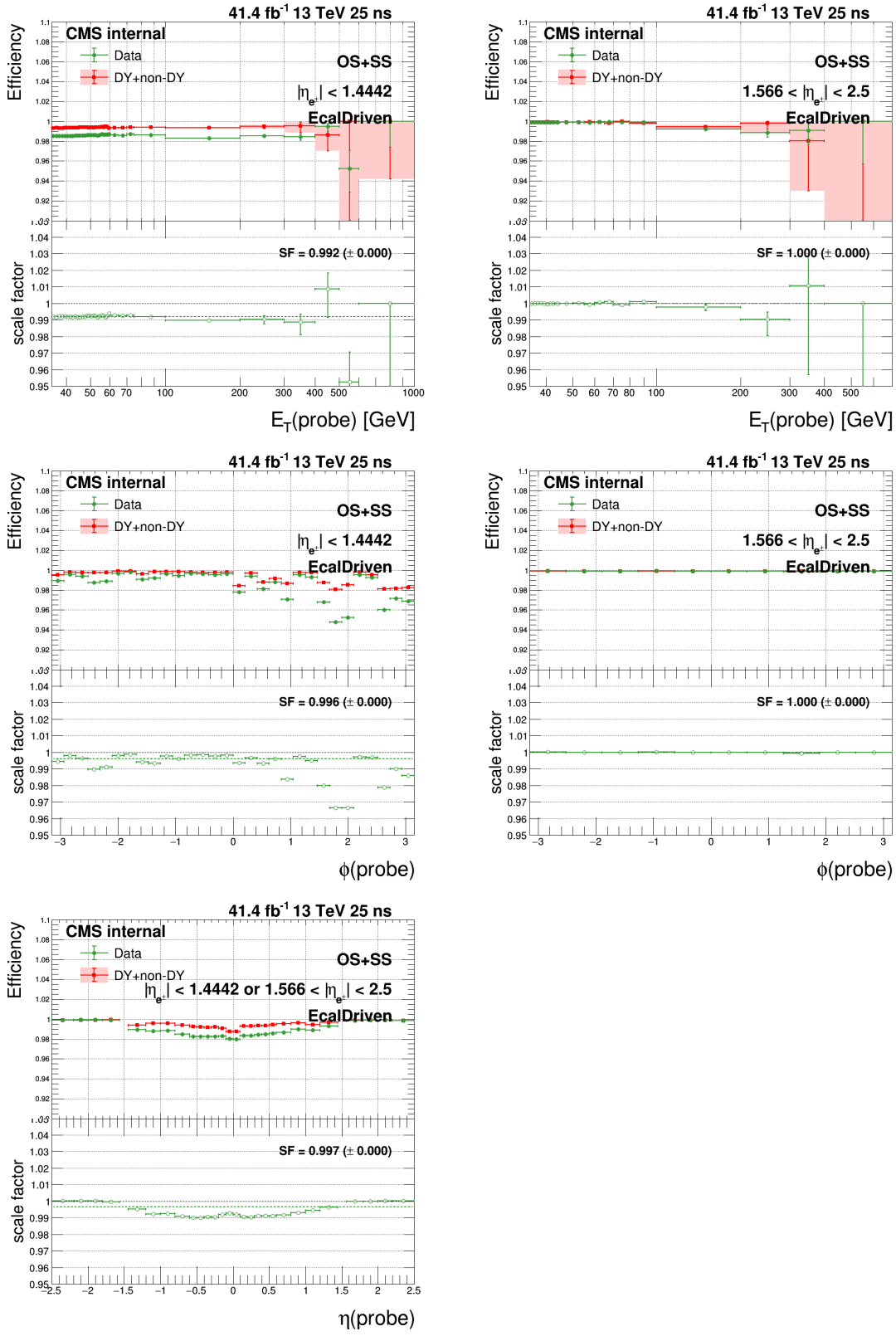

Figure B.25: *EcalDriven* N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

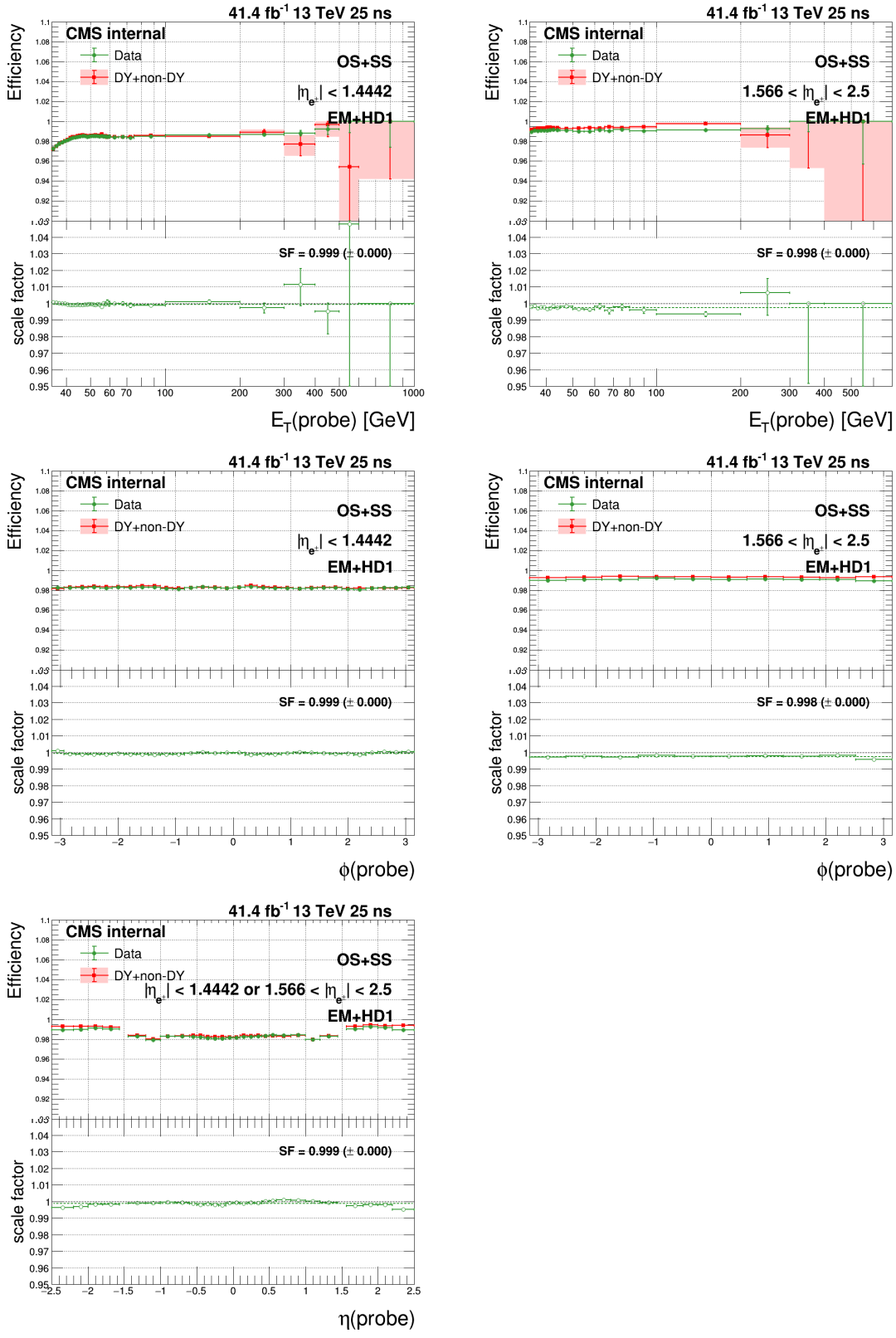

Figure B.26: $EM+HD1$ N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

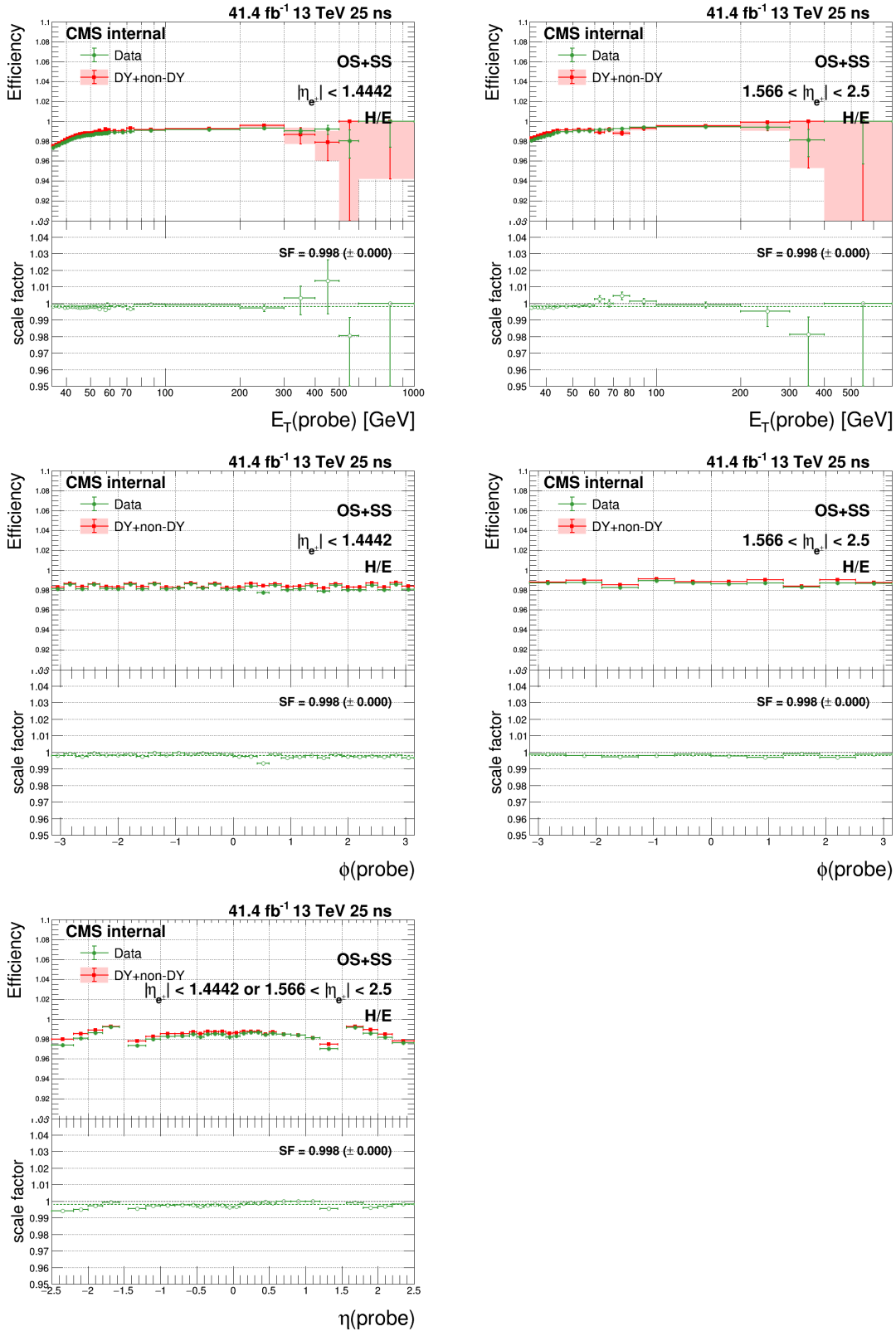

Figure B.27: H/E N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

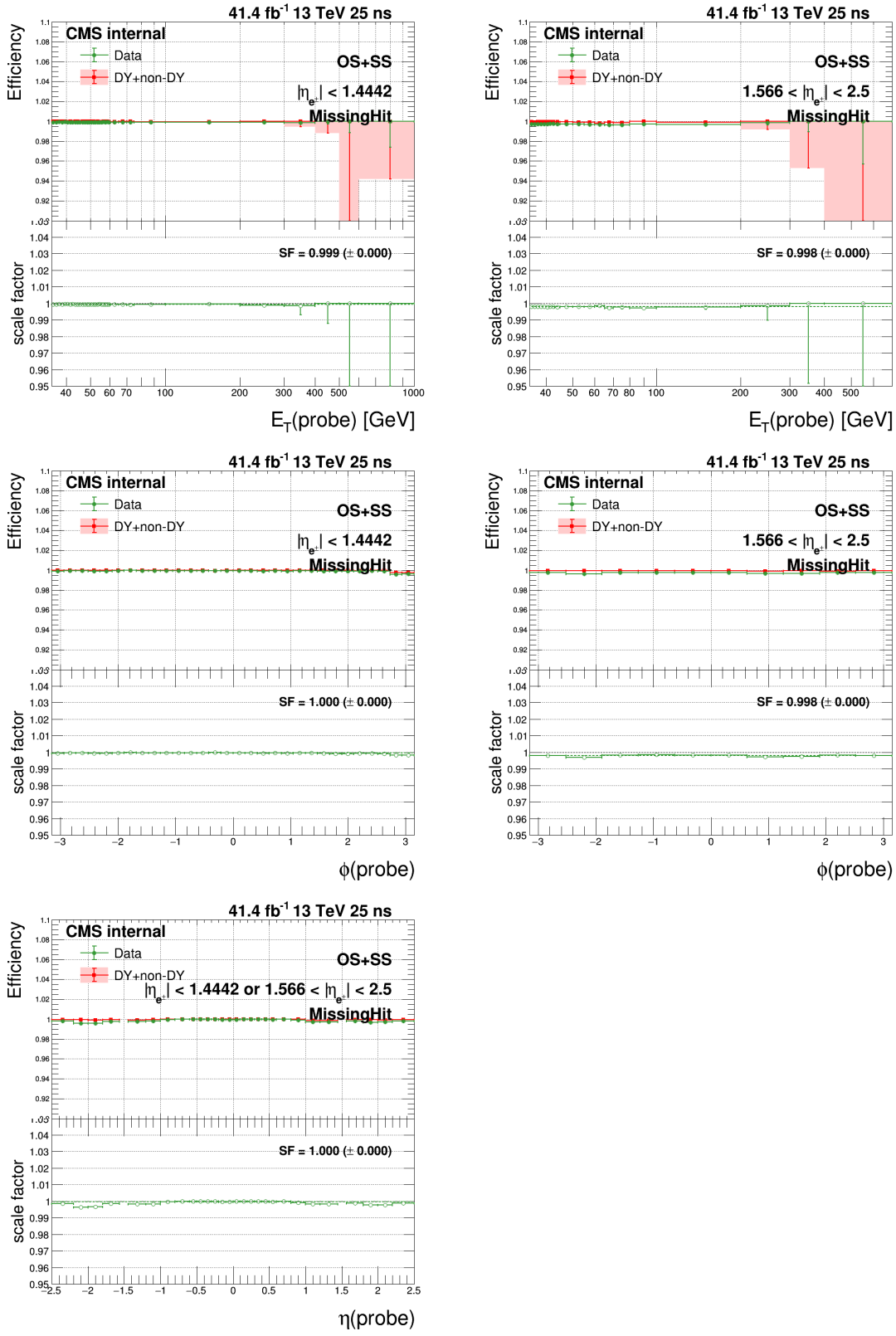

Figure B.28: *MissingHit* N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

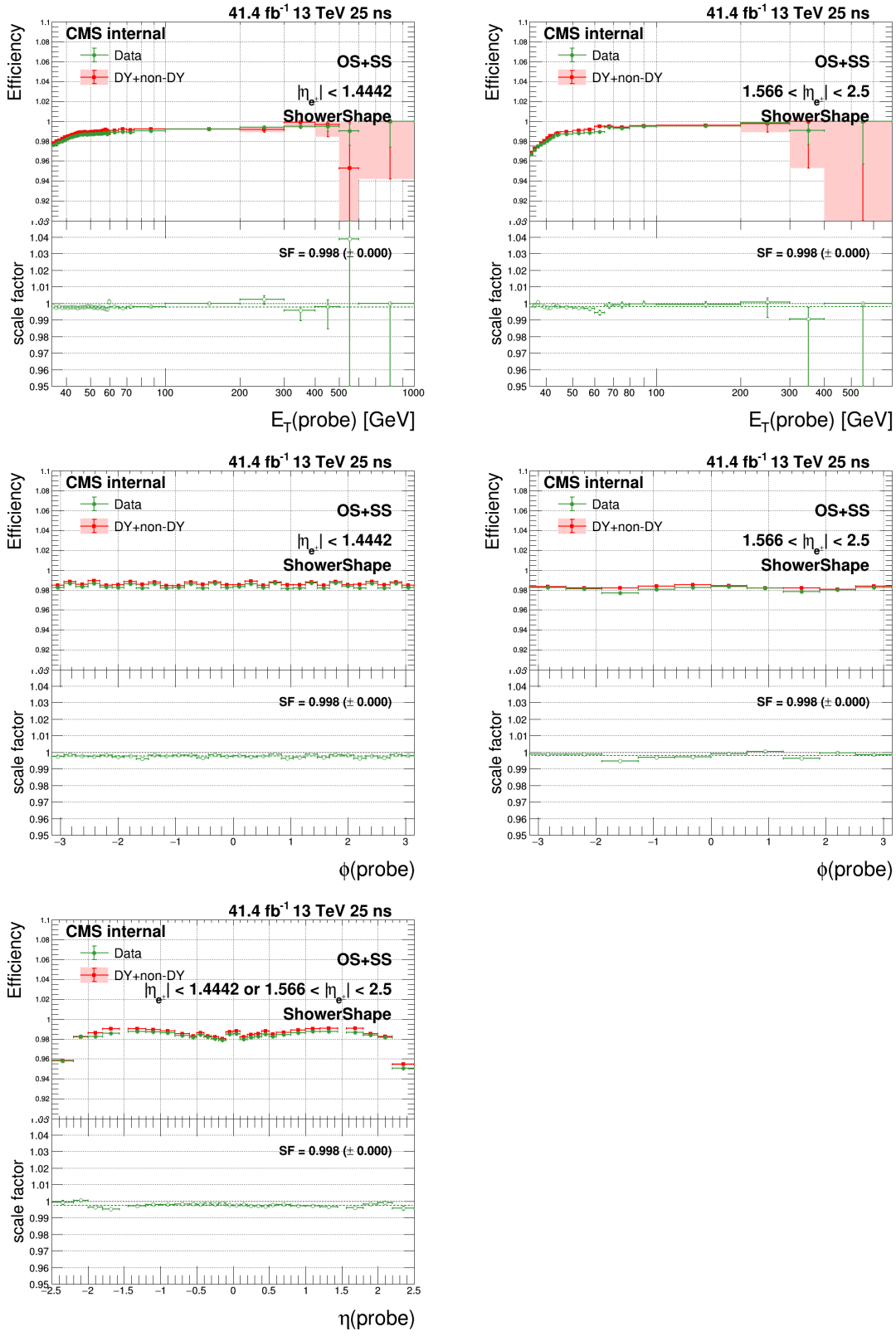

Figure B.29: *ShowerShape* N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

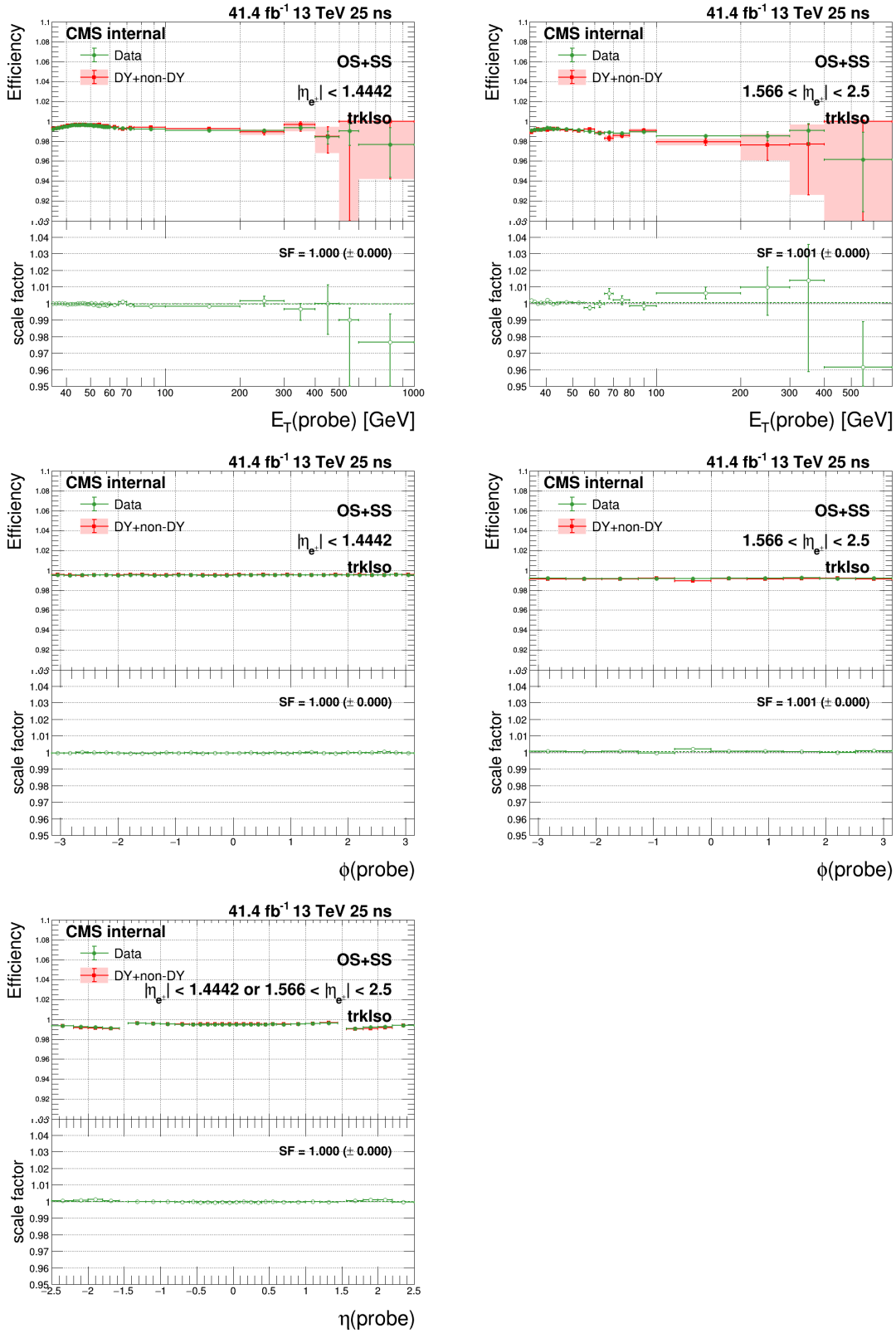

Figure B.30: *trackisolation* N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

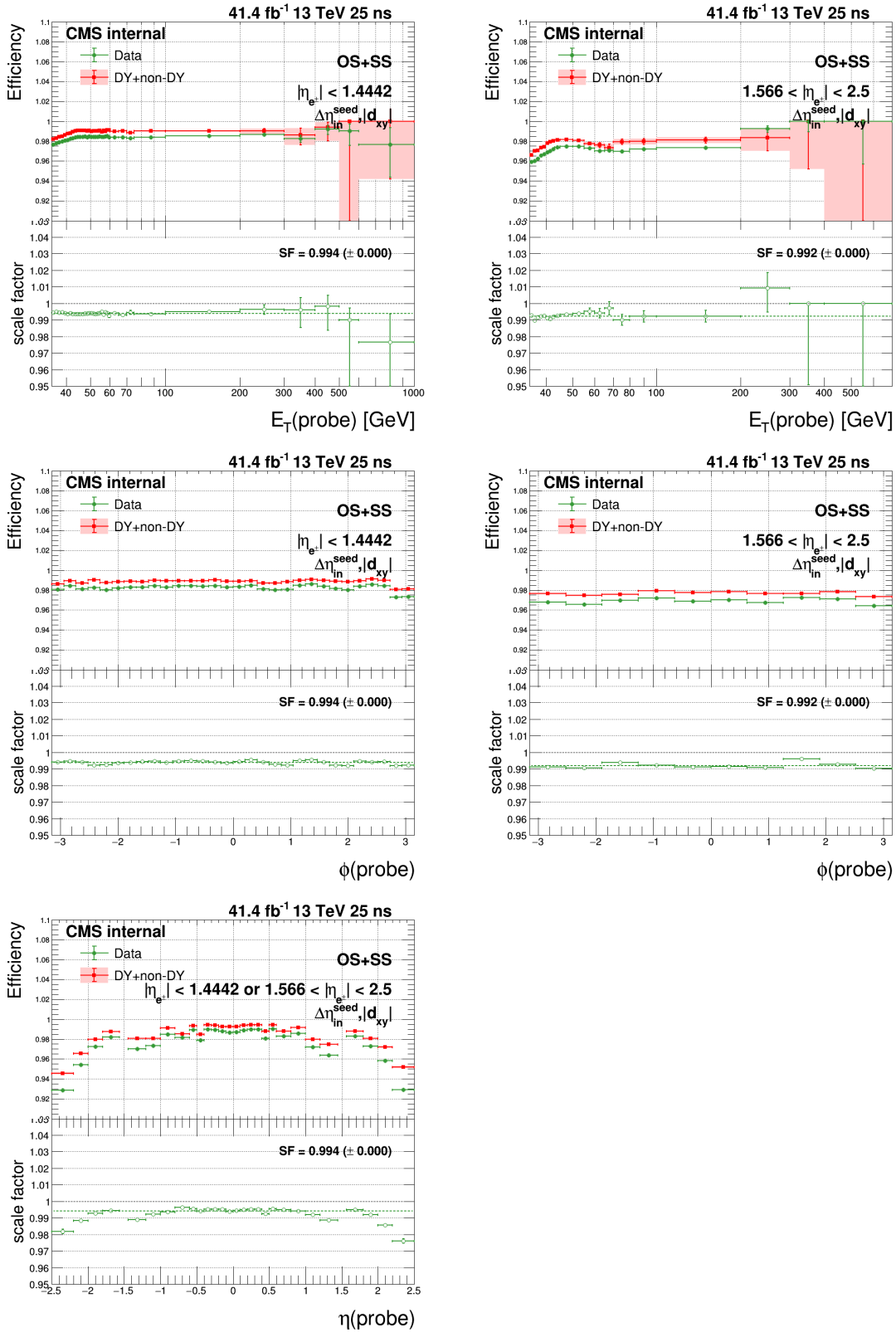

Figure B.31: $\Delta\eta_{in}^{seed}$ and $|d_{xy}|$ N-2 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

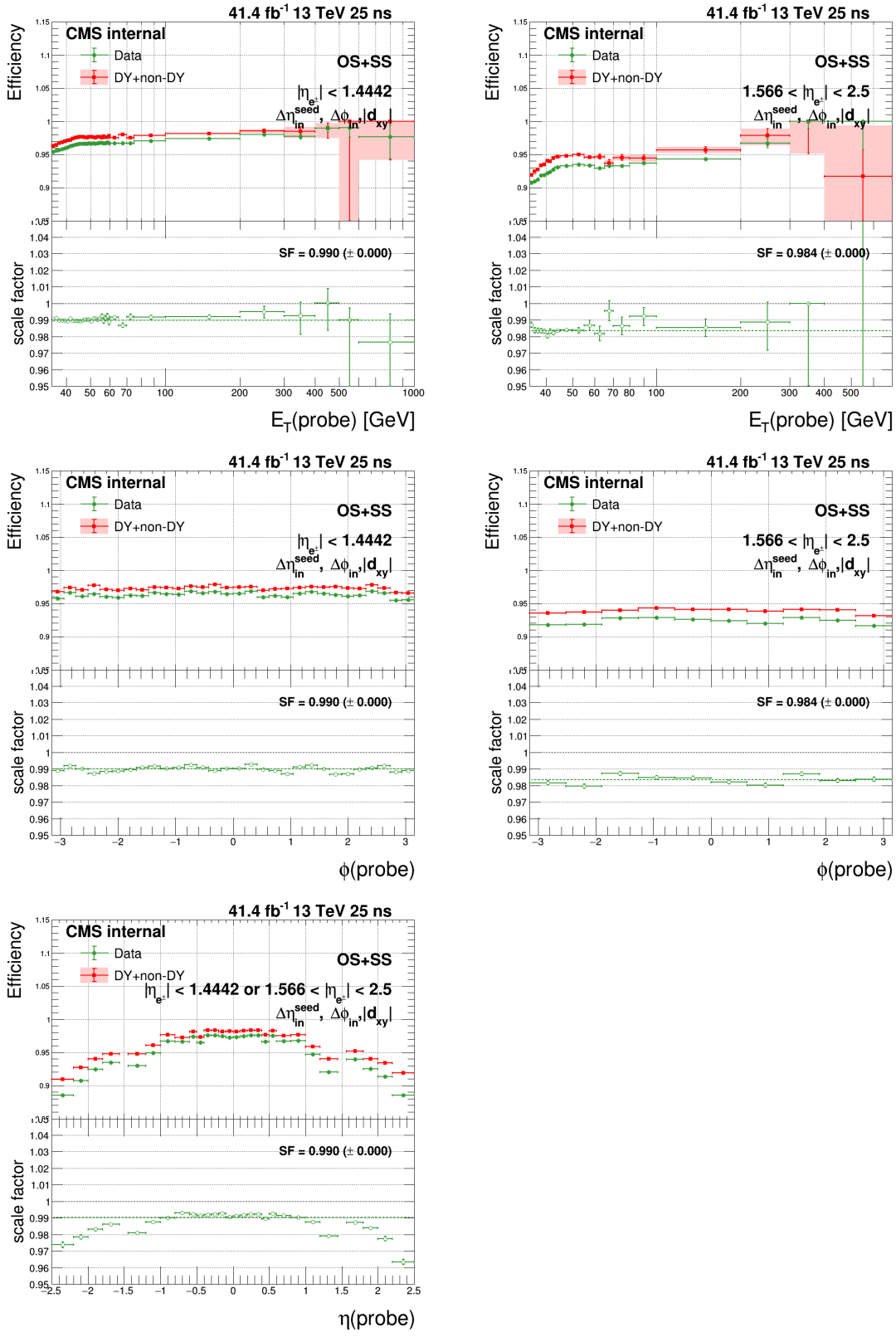

Figure B.32: $\Delta\eta_{in}^{seed}$, $|d_{xy}|$ and $\Delta\phi_{in}$ N-3 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom).

B.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 B.34 and B.35. The HEEP ID efficiency and scale factor is shown in Figure B.36. The comparation of the HEEP ID efficiency and scale factor from cut method and fit method are shown in Figure B.37.

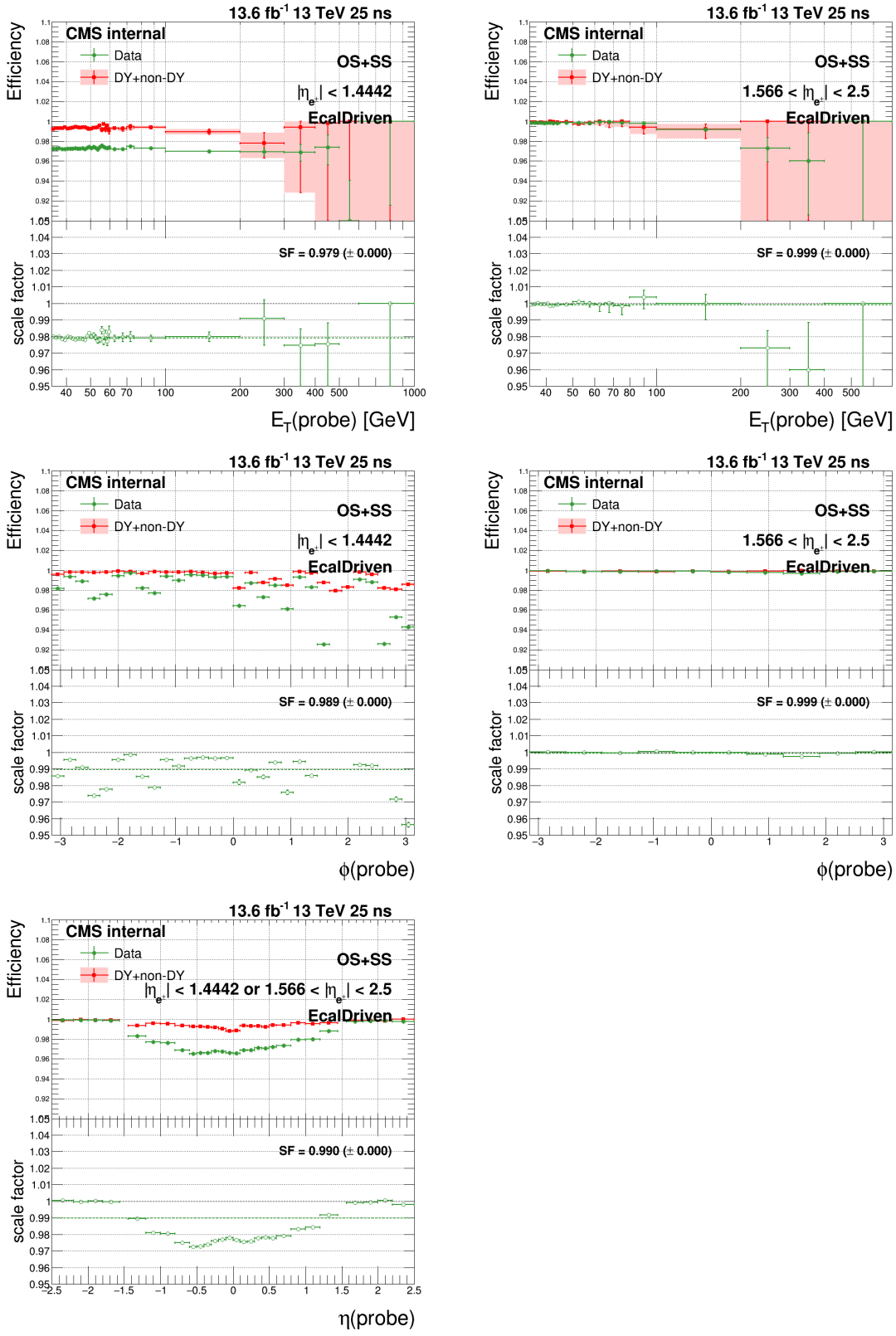

Figure B.33: *EcalDriven* N-1 efficiencies and scale factors in MC and data in the barrel (left) and endcap (right) as functions of probe E_T (top), ϕ (middle) and probe η (bottom) for run F.

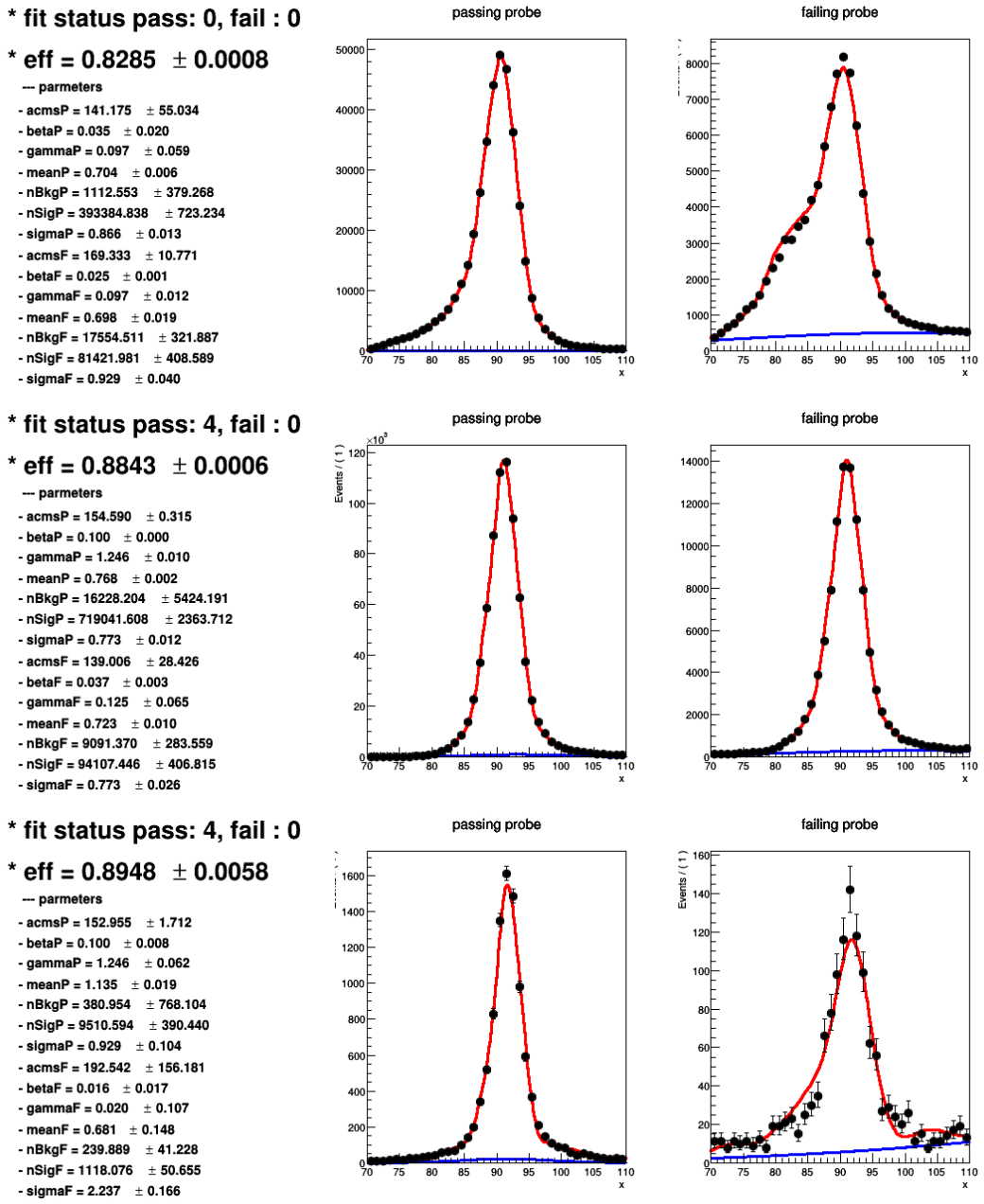

Figure B.34: The fit plot of passing probes (left) and failing probes (right) for E_T of probe between 35-36 GeV (top), 44-45 GeV (middle) and 200-1000 GeV (bottom) in the barrel.

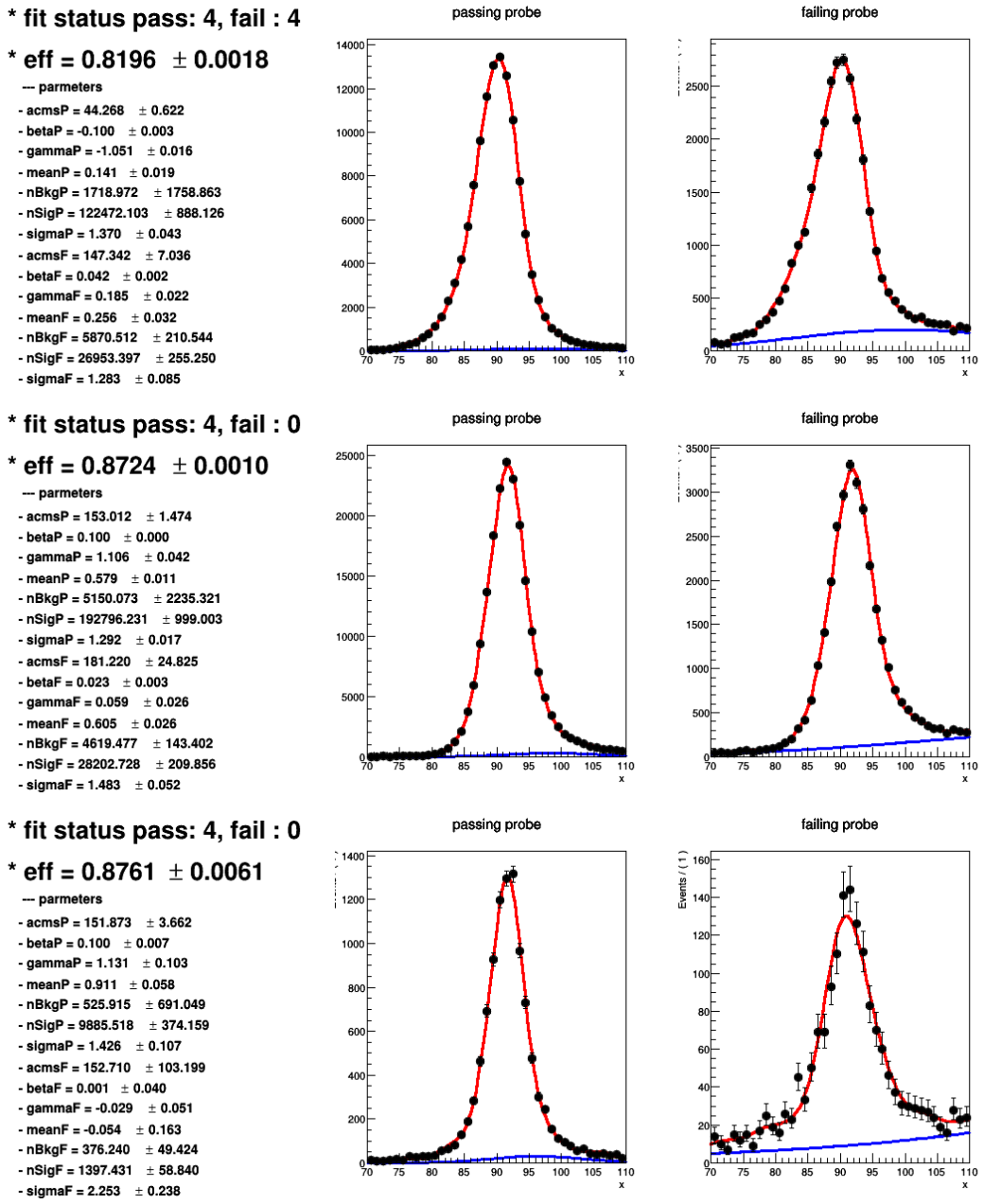

Figure B.35: The fit plot of passing probes (left) and failing probes (right) for E_T of probe between 35-36 GeV (top), 43-45 GeV (middle) and 100-1000 GeV (bottom) in the endcap.

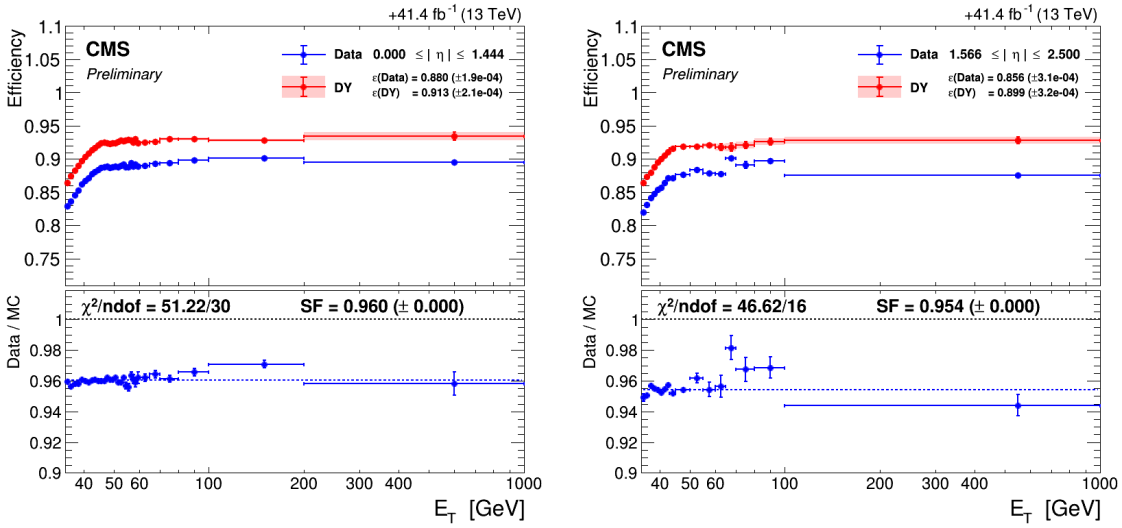

Figure B.36: Efficiencies and scale factors in MC and data as functions of probe E_T.

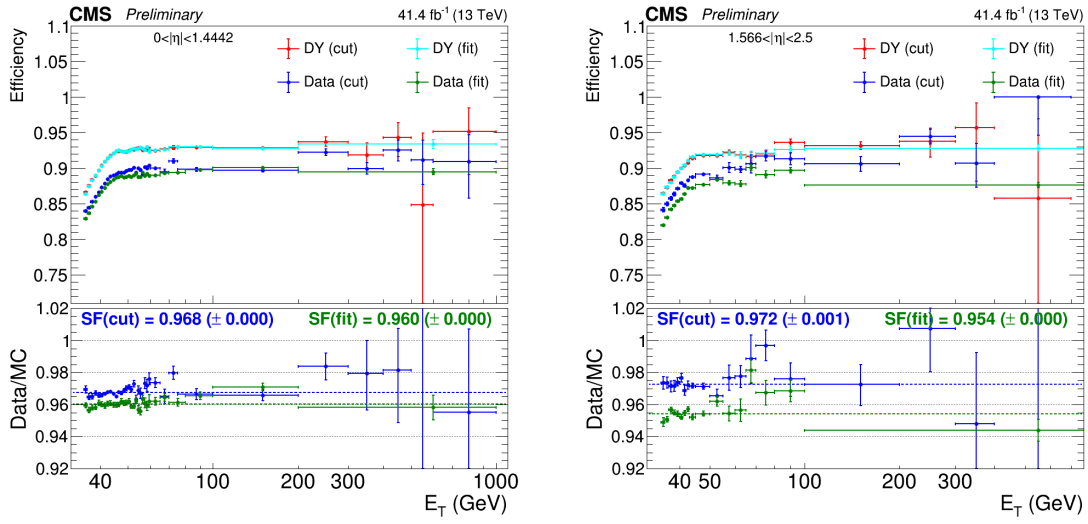

Figure B.37: The comparation of efficiencies and scale factors in MC and data between cut method and fit method as functions of probe E_T.

B.4 Electron Saturation Study

When the deposite energy in single ECAL crystal is very high, the electric readout will be saturated because of limited dynamic range. Therefore the energy of saturated electron measured by detector with be incorrect. Here we adopt MVA method to get the correct energy of saturated electron which can be seen in chapter B.4.1. Besides, we also check the ECAL linearity response in chapter B.4.2. Finally we study the saturate effection to the HEEP ID efficiency in chapter B.4.3.

B.4.1 Get true energy of saturated electron

B.4.1.1 Method

Here we use the MVA BDTG (Gradient Boost) method which is the variants of BDT (Boosted Decision Tree) to get the correct energy of saturated electron. The training target defined as:

$$T = log(\frac{E_{MC}}{E_{rawSC}}) \tag{B.1}$$

where the E_{MC} is the true energy of electron and E_{rawSC} is the raw SuperCluster energy. The training input variables are listed below:

- E_{rawSC} : raw energy of the SuperCluster
- $\frac{E_{max}}{E_{rawSC}}$: energy of the most energetic crystal in the SC normalized to E_{rawSC}
- $\frac{E_{3\times3}}{E_{rawSC}}$: energy of the 3×3 matrix centered around the seed normalized to E_{rawSC}
- $\frac{E_{5\times5}}{E_{rawSC}}$: energy of the 5×5 matrix centered around the seed normalized to E_{rawSC}
- $\frac{E_{left}}{E_{rawSC}}, \frac{E_{right}}{E_{rawSC}}, \frac{E_{top}}{E_{rawSC}}, \frac{E_{bottom}}{E_{rawSC}}$: energy of the four crystals around the seed normalized to E_{rawSC}
- $\frac{E_{2\times5\ left}}{E_{rawSC}}, \frac{E_{2\times5\ right}}{E_{rawSC}}, \frac{E_{2\times5\ top}}{E_{rawSC}}, \frac{E_{2\times5\ bottom}}{E_{rawSC}}$: energy of the four 2×5 crystal dominoes around the seed belonging to the 5×5 matrix normalized to E_{rawSC}
- $\frac{E_{1\times5\ max}}{E_{rawSC}}$: energy of the most energetic 1×5 domino belonging to the 5×5 matrix normalized to E_{rawSC}
- $\frac{E_{preShower}}{E_{rawSC}}$: energy measured in the PreShower normalized to E_{rawSC} (only for endcap electrons)
- η and ϕ of the SC
- η and ϕ width of the SC
- i_η and i_ϕ of the SC seed
- $\frac{H}{E}$
- ρ

The configuration of the training and testing sample is `factory->PrepareTrainingAndTestTree ("","nTrain_Regression=0:nTest_Regression=0:SplitMode=Random:NormMode= NumEvents:!V")` which means the the sample is split in half for training and testing, the events are selected randomly and the weight is one. The configuration of the BDTG method is `factory->BookMethod(MVA::Types::kBDT, "BDTG","!H:!V:NTrees=1000:: BoostType=Grad:Shrinkage=0.1:UseBaggedBoost:BaggedSampleFraction=0.5: nCuts=5:MaxDepth=3")`

The training target and input variables are the same with diphoton analysis in ref. [161]. The electrons for the training are mc matched ($\Delta R < 0.1$) saturated electrons with mc energy less than 10 TeV and the η ranges for the training are barrel and endcap separately.

B.4.1.2 Sample

The training and testing sample used is `/DoubleElectron_FlatPt-300To6500/RunIISpring16DR80-PUFlat0to50_80X_mcRun2_asymptotic_2016_v3-v1/AODSIM` which is flat in pseudorapidity η and transverse momentum P_T in range 0.3 to 6.5 TeV. The distributions of P_T and energy of generated electron versus η are shown in Figure B.38, one can see it is almost flat in η and P_T as expected. Besides, in Figure B.38 one can see the threshold energy for electron being saturated in barrel is around 2 TeV, while in endcap the threshold is not a consant which increases with the η. More details can be seen in Figure B.39 which is the distribution of E_{max} versus η for unsaturated and saturated electron, one can see the E_{max} for saturated electron in barrel is almost flat while for endcap it is increase with η, the reason why E_{max} will increase with η in endcap for saturated electron maybe because of the "darkness" of crystal increase with η in endcap. Moreover from Figure B.40 we know $E_{sc} = E_{max} + E_{24}$ is correct also for saturated electron. From Figure B.41 which gives the distributions of E_{24}, E_{max}, E_{sc} versus generated energy E_{mc} for saturated electron and the plots of E_{sc} versus E_{mc} should be equal to plots of E_{24} versus E_{mc} add the plots of E_{max} versus E_{mc}. In addition, the saturated fraction of electron for different energy is shown in Figure B.42, one can see there are a sharp turnon curve for barrel while for endcap it is gentle. Finally the number (or its fraction) of saturated crystal in 3×3 matrix with seed crystal in the center for different mc energy is shown in Figure B.43, one can see the fraction of having two saturated crystals is increase with the mc energy especially in barrel and the strange behaviour around the highest mc energy in barrel is because of the electrons which are very close to the gap.

B.4.1.3 Result

The distribution of $(recoE - trueE)/trueE$ are shown in Figure B.44 and one can see the energy of saturated electron from MVA is very close to true (or mc) energy. A fit to the MVA result using double-side CB is performed, the peak position and the standard deviation of the Gaussian core of the distributions are estimated through the fitted values of mean and sigma, respectively. The $'effective'$ standard deviation σ_{eff}, defined as half of the smallest interval around the peak position containing 68.3% of the events, is used to assess the resolution, while taking into account possible non-Gaussian tails. We also checked the result by using MLP method which gives worse resolution shown in Figure B.59 in section B.4.5, therefore our basic method is BDTG method. Then we check the results for different η which are shown in Figure B.45 and B.46. From figures B.45 and B.46 we know the MVA works well for different η. Besides we check the MVA regressive results for different true energy of saturated electron which are shown in Figure B.47. From Figure B.47 we see for barrel the results are stable for different energy, while for endcap the results are not very well in low energy, because for low electron energy the possibility of crystal to be saturated in endcap is very small and the training statistics for low energy and saturated electron is very less which can be seen in the bottom right plot of Figure B.38.

B.4.1.4 Check the result in ZToEE samples

Here we check the MVA result using ZToEE powheg sample with M_Z great than 4500 GeV samples. The results are shown in figures B.48 and B.49. The result looks good.

B.4.1.5 Apply to data

Using DoubleEG dataset from 2016 in 5.1, we find three events in which there are and only one saturated HEEP (High Energy Electron Pairs) electron. The energy of SC and MVA regressived energy are shown in Table B.2. The events displays are shown in Figure B.50.

	event number	η	SC energy (GeV)	Regressived energy (GeV)
electron1	1076867675	1.21759	2370.34	2279.53
electron2	897834686	1.56931	2954.7	3167.63
electron3	400840829	1.1476	2048.81	3151.99

Table B.2: Detail of saturated HEEP electrons.

B.4.1.6 Prediction of saturated events

The number of saturated DY events which contain saturated electron for 30 fb^{-1} luminosity is shown in Figure B.51 and one can see the possibility to have one saturated DY event is $\sim 6\%$.

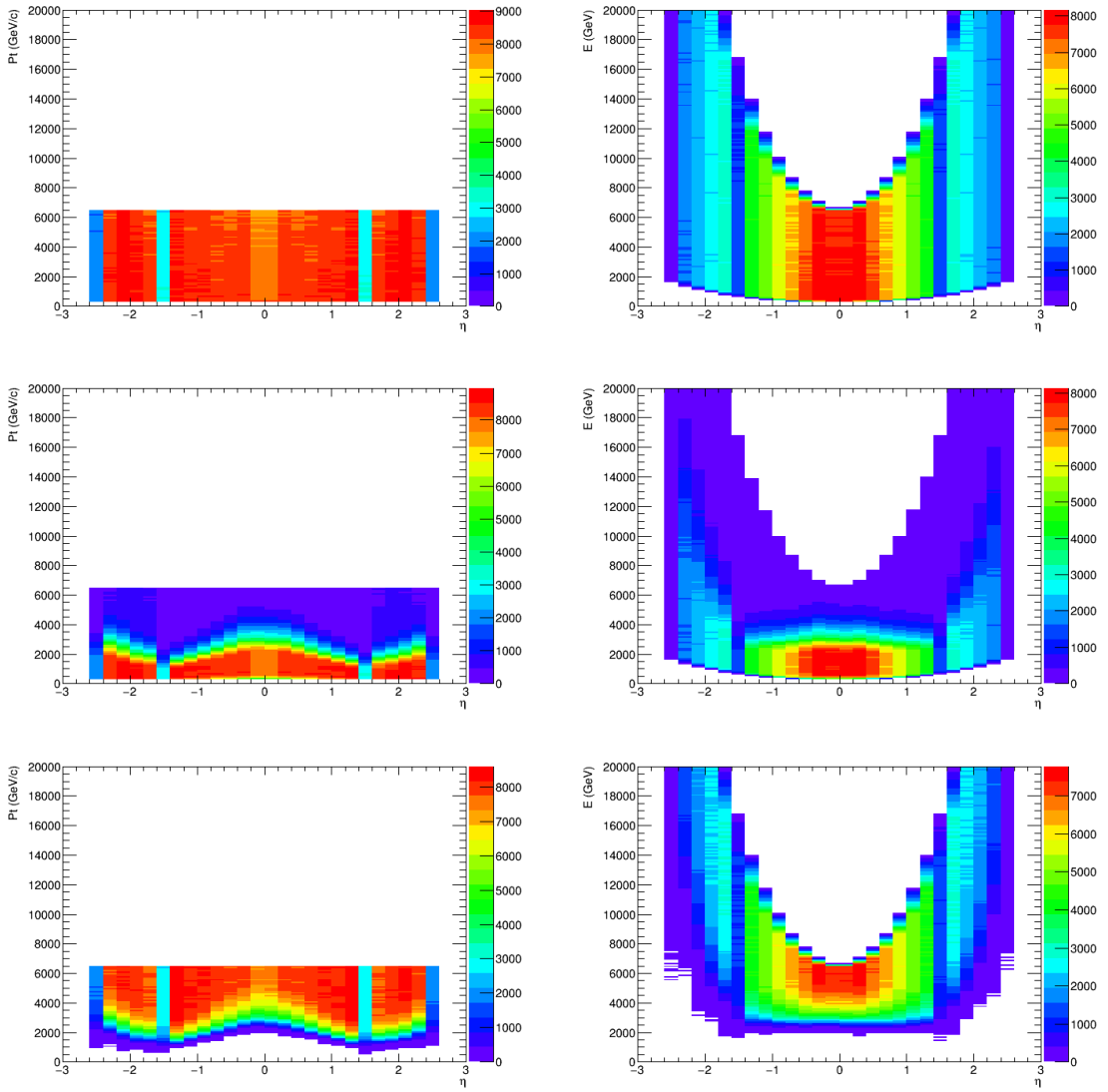

Figure B.38: The distributions of P_T for generated electrons versus η (left) and energy of generated electrons versus η (right) for all electrons (top), unsaturated electrons (middle) and saturated electrons (bottom).

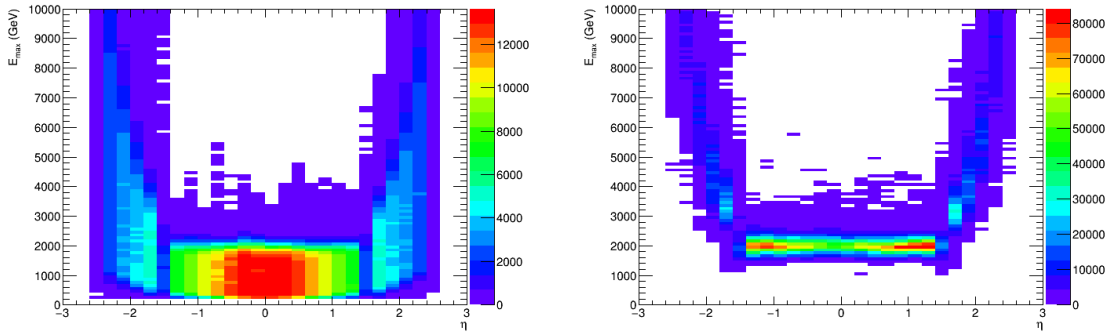

Figure B.39: The distributions of E_{max} versus η for unsaturated electrons (left) and saturated electrons (right).

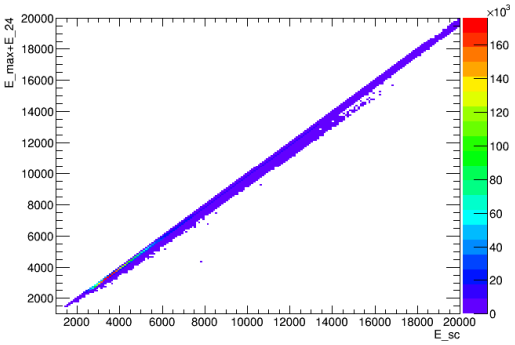

Figure B.40: The distribution of $E_{max} + E_{24}$ versus E_{sc} for saturated electrons.

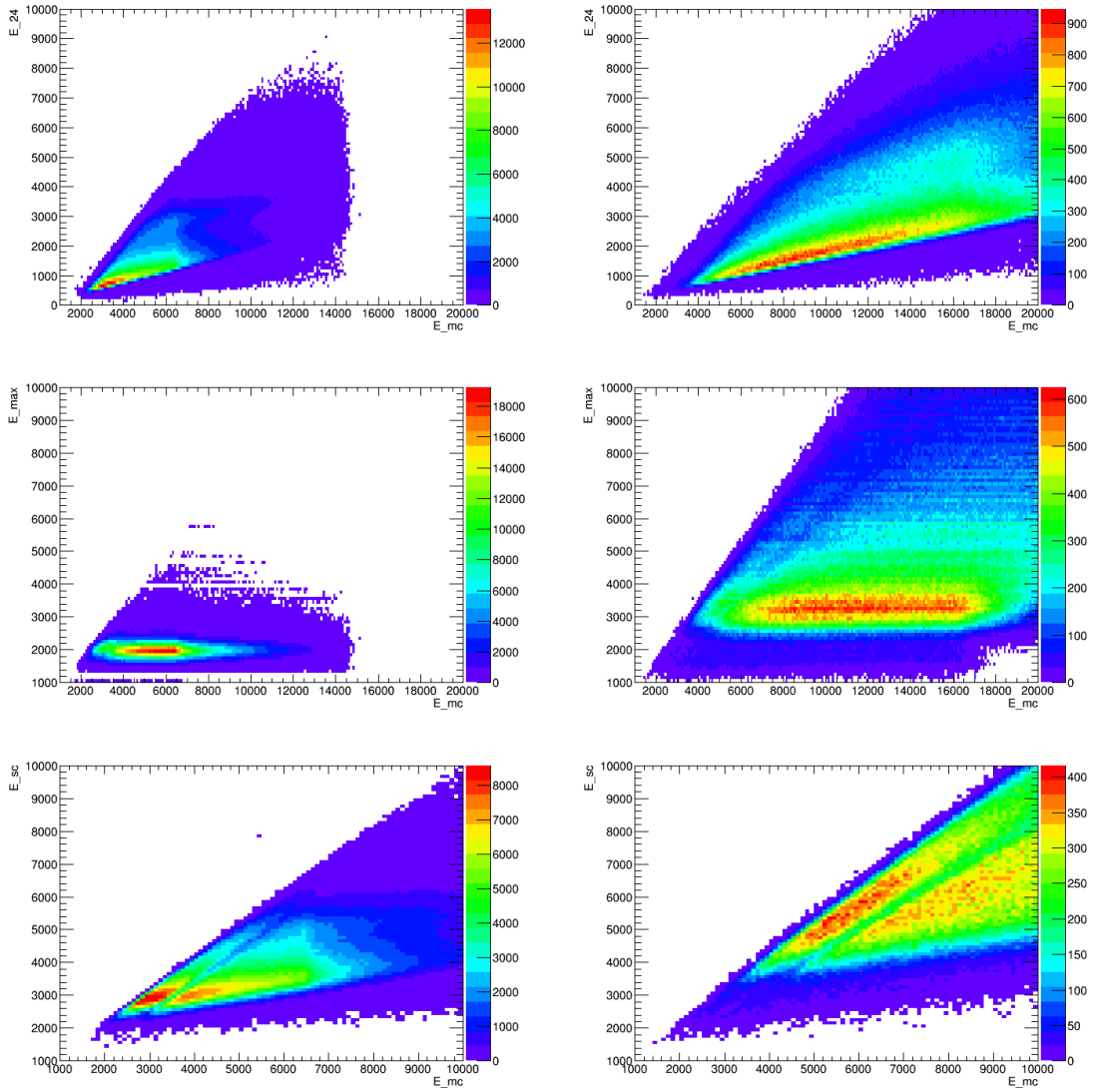

Figure B.41: The distributions of E_{24} versus E_{mc} (top), E_{max} versus E_{mc} (middle) and E_{sc} versus E_{mc} (bottom) for saturated electrons for barrel (left) and endcap (right).

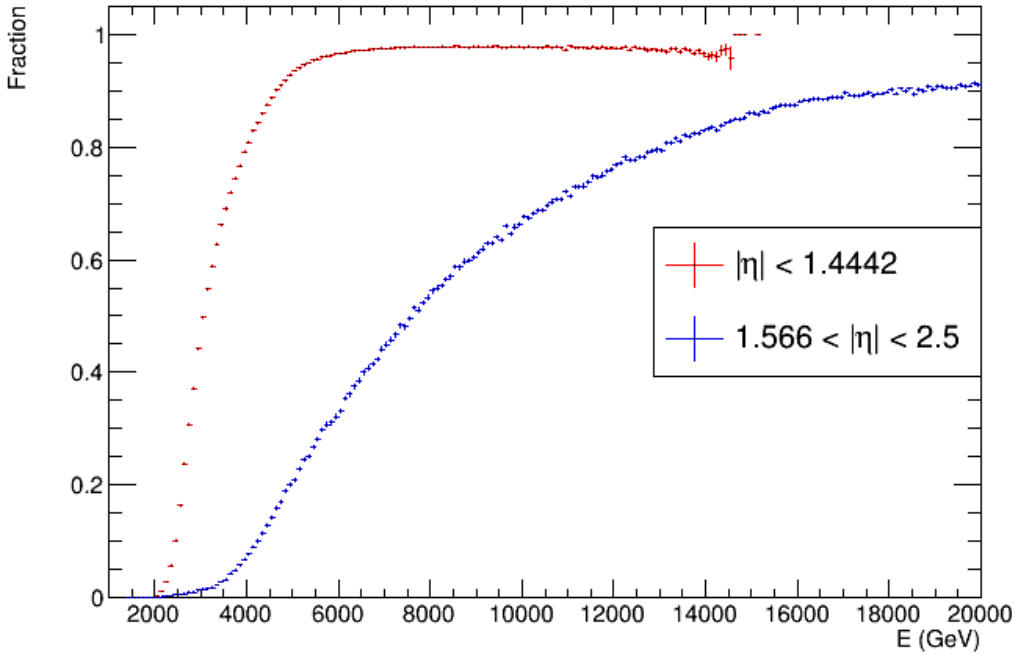

Figure B.42: The saturated fraction versus true (or mc) energy of electron for barrel (red points) and endcap (blue points).

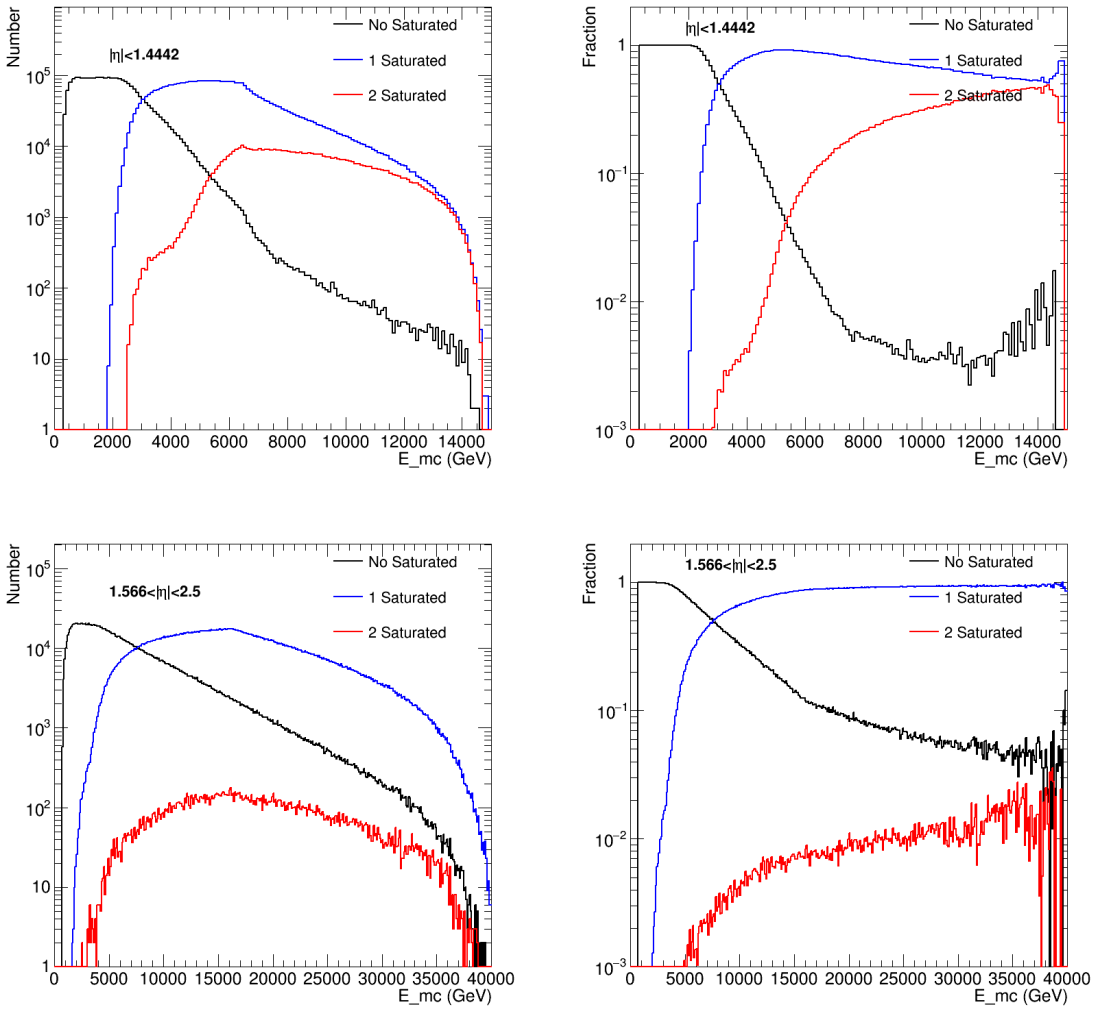

Figure B.43: The number of saturated crystals in 3×3 matrix with seed crystal in the center for different mc energy (left) and the fraction of number of saturated crystals in 3×3 matrix with seed crystal in the center for different mc energy (right) for barrel (top) and endcap (bottom).

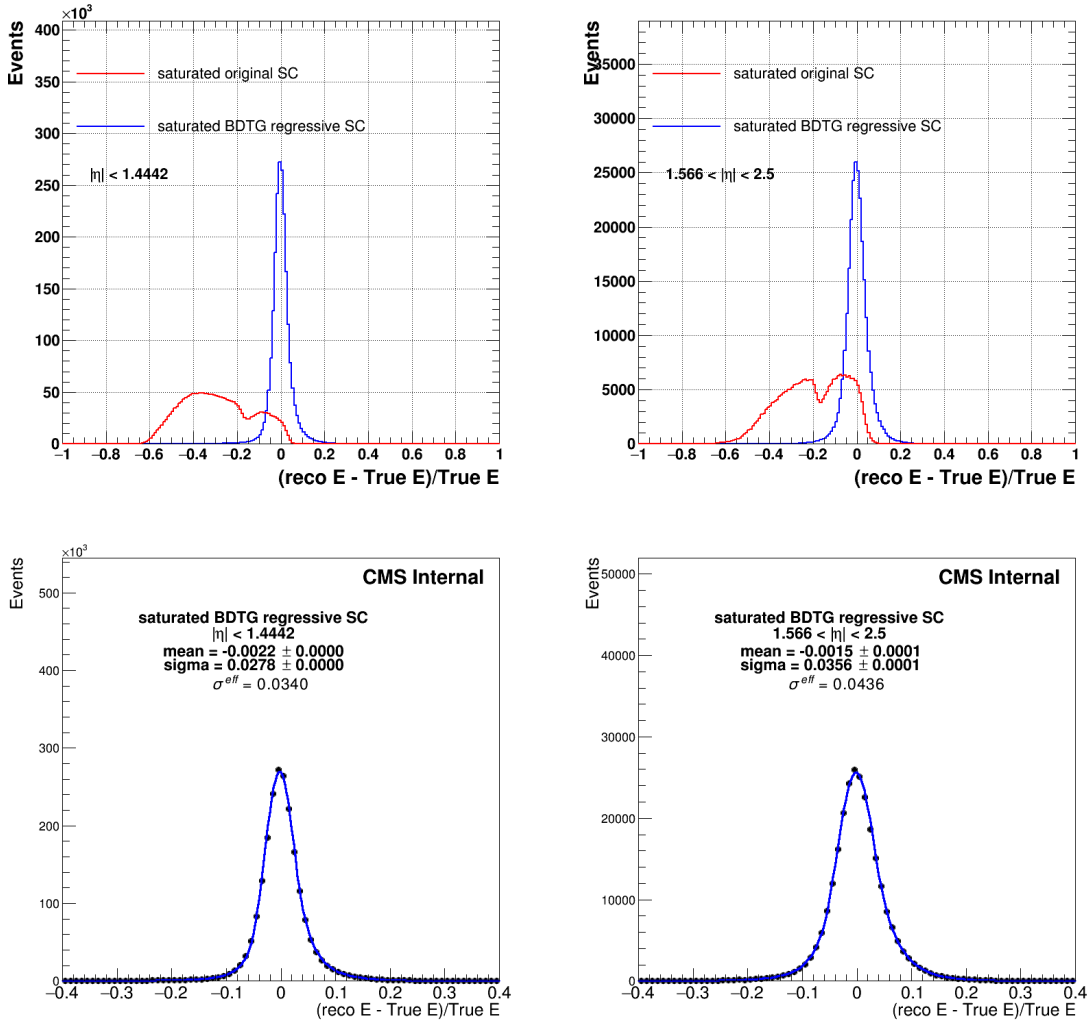

Figure B.44: The top plots are the distribution of supercluster energy minus ture energy divided by true energy for saturated electron for barrel (left) and endcap (right), the red histogram is for reconstructed supercluster enery, the blue histogram is for MVA regressive energy. The bottom plots are the fit of the blue histogram for barrel (left) and endcap (right).

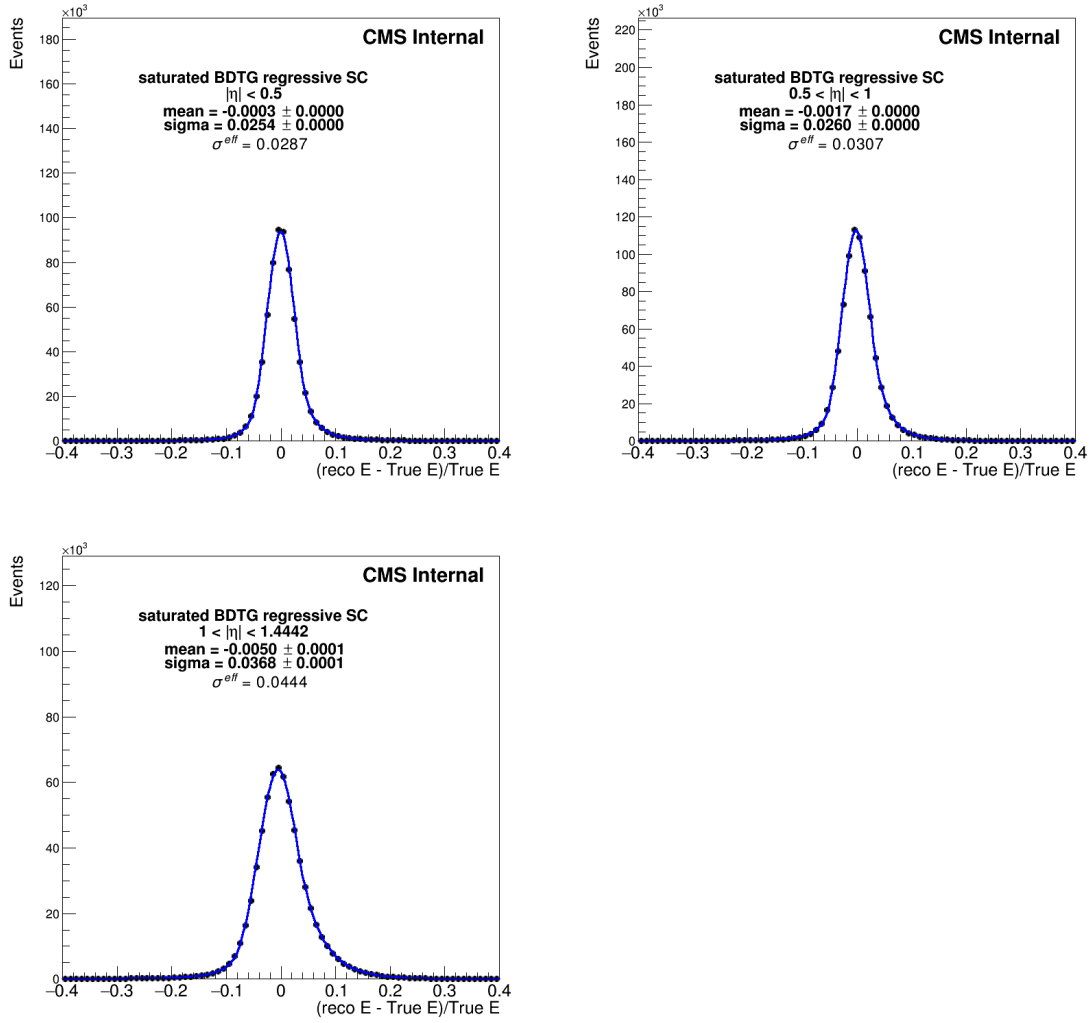

Figure B.45: The distributions of MVA regressive energy minus ture energy divided by true energy for saturated electrons in barrel for different η ranges.

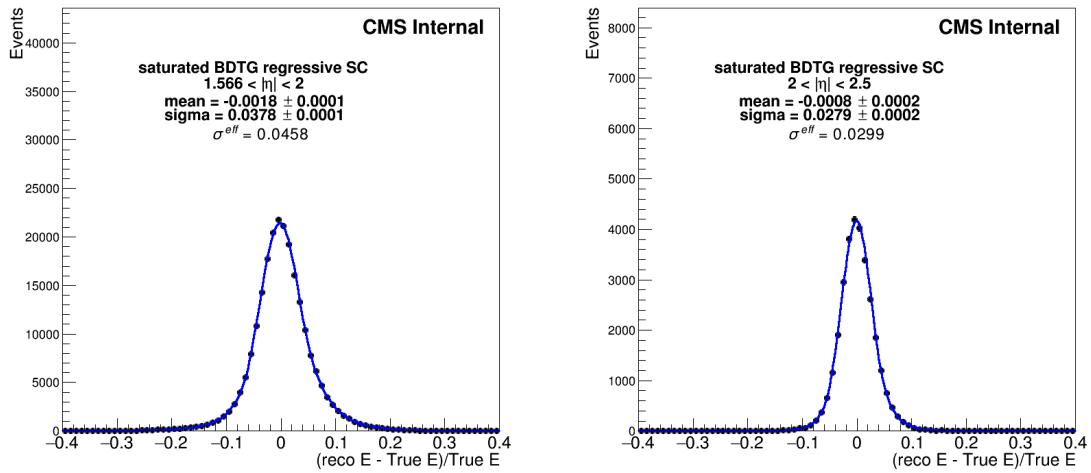

Figure B.46: The distributions of MVA regressive energy minus ture energy divided by true energy for saturated electron in endcap for different η ranges.

	Energy	2-3TeV	3-4TeV	4-5TeV	5-6TeV	6-7TeV	7-8TeV	8-9TeV	9-10TeV	<10TeV
Barrel	Mean	0.0036	0.0001	-0.0002	-0.0011	-0.0032	-0.0028	-0.0056	-0.0142	-0.0022
	Sigma	0.0254	0.0265	0.0265	0.0268	0.0275	0.0289	0.0312	0.0306	0.0278
Endcap	Mean	0.0424	0.0172	0.0108	0.0064	0.0069	0.0071	0.0014	-0.0143	-0.0015
	Sigma	0.0775	0.0525	0.0431	0.0402	0.0396	0.0380	0.0336	0.0308	0.0356

Figure B.47: The fit results for the distributions of MVA regressive energy minus ture energy divided by true energy for saturated electron for different true energy in barrel and endcap.

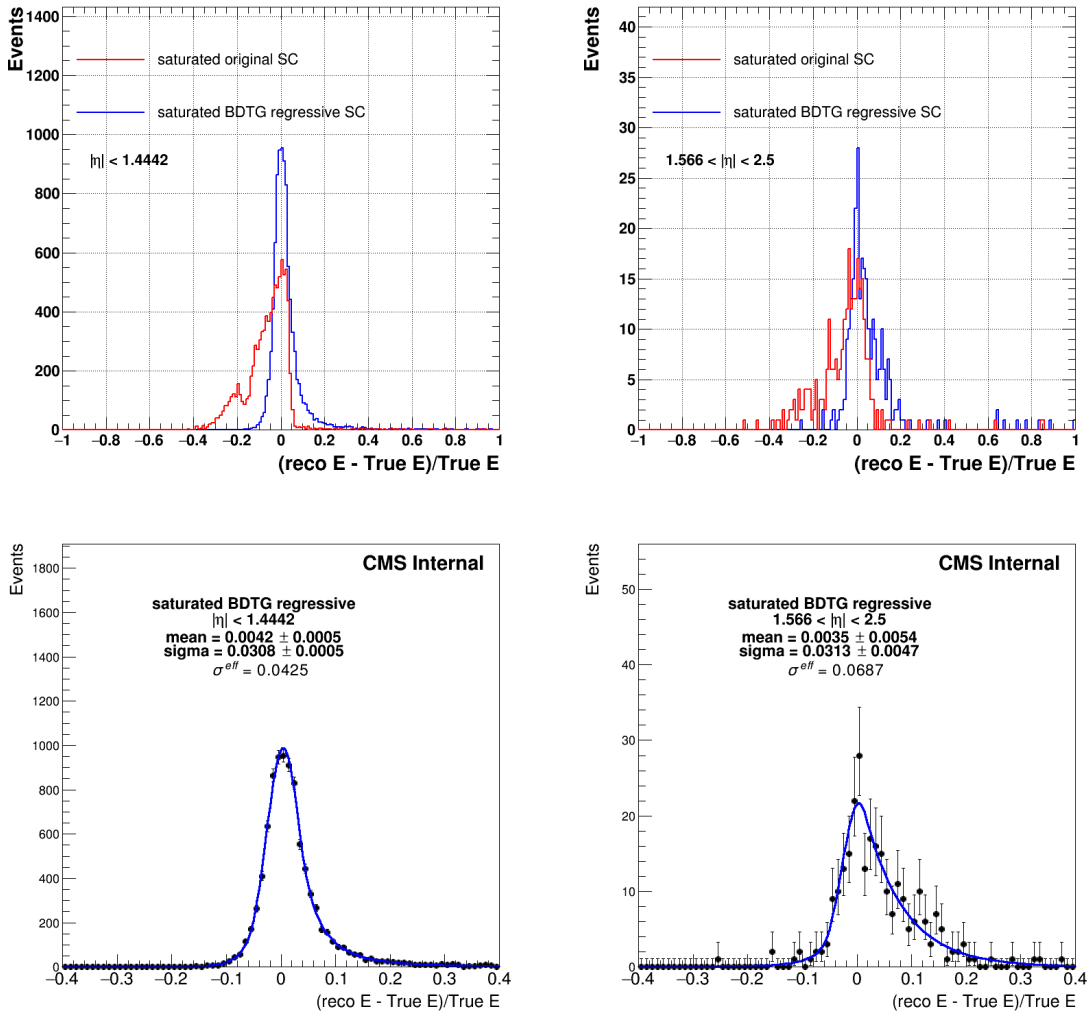

Figure B.48: The top plots are the distribution of supercluster energy minus ture energy divided by true energy for saturated electron for barrel (left) and endcap (right) for ZToEE_4500_6000 sample, the red histogram is for reconstructed supercluster enery, the blue histogram is for MVA regressive energy. The bottom plots are the fit of the blue histogram for barrel (left) and endcap (right).

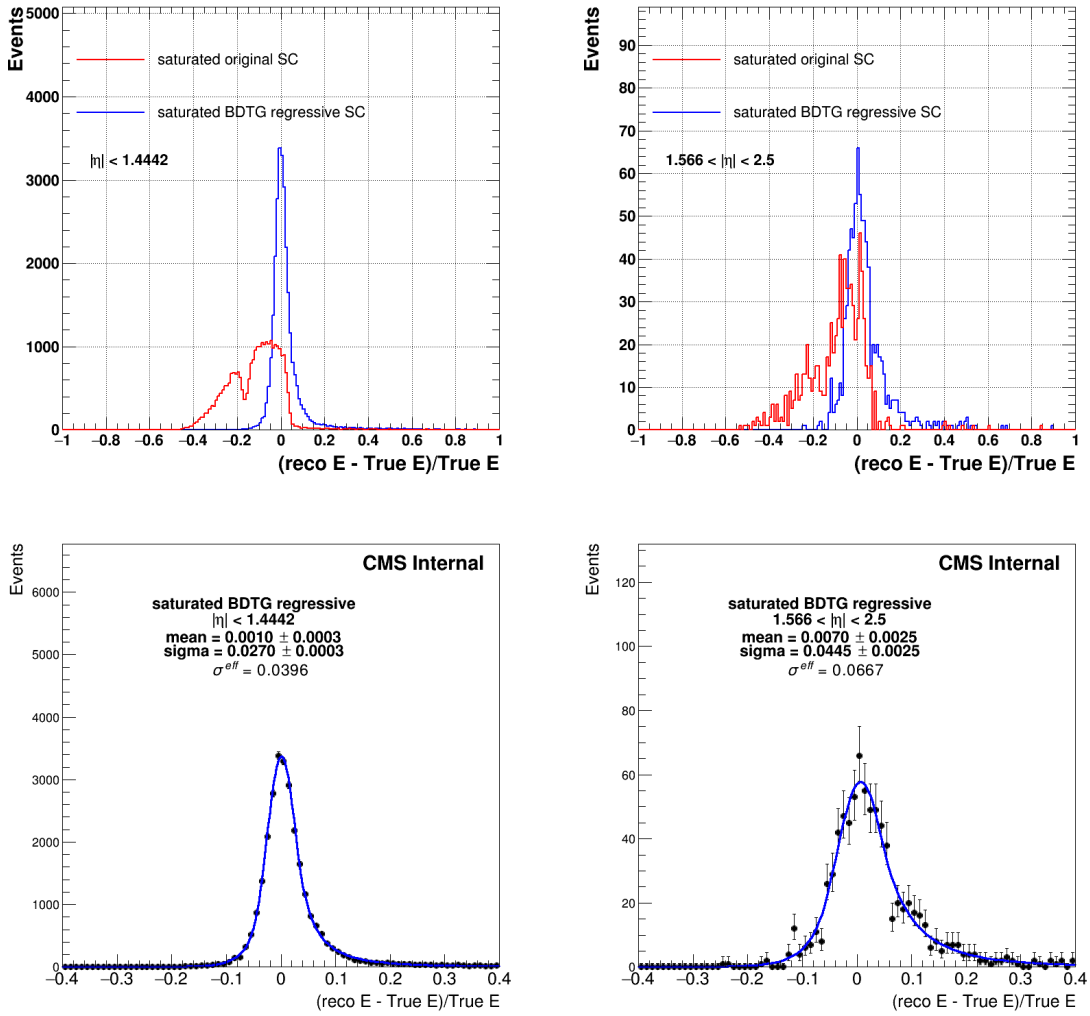

Figure B.49: The top plots are the distribution of supercluster energy minus ture energy divided by true energy for saturated electron for barrel (left) and endcap (right) for ZToEE_6000_Inf sample, the red histogram is for reconstructed supercluster enery, the blue histogram is for MVA regressive energy. The bottom plots are the fit of the blue histogram for barrel (left) and endcap (right).

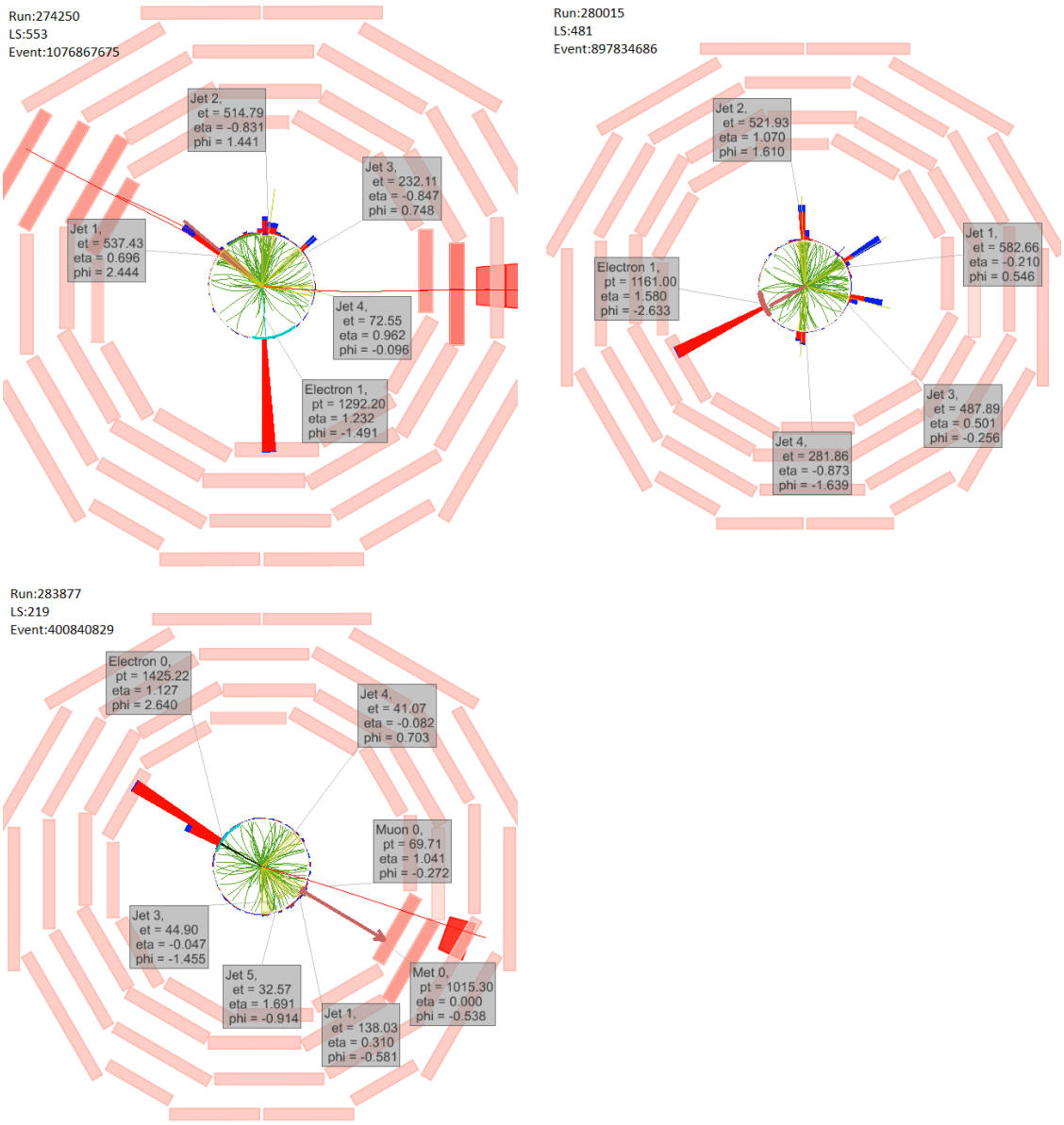

Figure B.50: The events displays of saturated HEEP electrons in data for $\rho\phi$ visual angle.

Mass range	Corss section (pb)	Total events	Saturated events	Saturated fraction	Expected saturated events(30fb^{-1})
800_1400	0.0168	100000	0	0	0
1400_2300	0.00139	100000	57	0.000570	0.023769
2300_3500	8.948E-05	100000	909	0.009090	0.024401
3500_4500	4.135E-06	100000	7803	0.078030	0.009680
4500_6000	4.56E-07	99000	28653	0.289424	0.003959
6000_Inf	2.06E-08	99800	67210	0.673447	0.000416
Total					0.062225

Figure B.51: The number of saturated DY events for 30 fb^{-1} luminosity.

B.4.2 Check ECAL linearity response

Here we use MVA BDTG method training on unsaturated electron in MC and apply it to unsaturated electron in data to check ECAL linearity response for high energy electrons.

The training target defined as:

$$T = \frac{E_{calo}}{E_{24}} \tag{B.2}$$

where the E_{calo} is the calorimeter energy of electron and E_{24} is the sum energy of crystals in 5×5 matrix without seeded crystal. The training input variables are listed below:

- $\frac{E_{left}}{E_{24}}, \frac{E_{right}}{E_{24}}, \frac{E_{top}}{E_{24}}, \frac{E_{bottom}}{E_{24}}$: energy of the four crystals around the seed normalized to E_{24}
- $\frac{E_{2\times5\ left}}{E_{24}}, \frac{E_{2\times5\ right}}{E_{24}}, \frac{E_{2\times5\ top}}{E_{24}}, \frac{E_{2\times5\ bottom}}{E_{24}}$: energy of the four 2×5 crystal dominoes around the seed belonging to the 5×5 matrix normalized to E_{24}
- $\frac{E_{preShower}}{E_{24}}$: energy measured in the PreShower normalized to E_{24} (only for endcap electrons)
- η of the SC

The MC samples for training and testing are `ZToEE_NNPDF30_13TeV-powheg_M_400_800 (also 800_1400)` Moriond sample. The electrons used for training are unsaturated HEEP electrons with calorimeter energy higher than 500 GeV in barrel and 600 GeV in endcap separately. The data used for applying the MVA method are `DoubleEG_Run2016B(C,D,E,F,G,H) -03Feb2017_v*_MINIAOD` datasets ($L \sim 35.9\ fb^{-1}$).

B.4.2.1 Result

The result from MC shown in Figure B.52 is for electron energy from 500 to 700 GeV in barrel and 600 to 1000 GeV in endcap, the result in Figure B.53 is for electron energy more than 700 GeV in barrel and more than 1000 GeV in endcap. Simliarly the result from data shown in Figure B.54 is for electron energy from 500 to 700 GeV in barrel and 600 to 1000 GeV in endcap, the result in Figure B.55 is for electron energy more than 700 GeV in barrel and more than 1000 GeV in endcap. In order to reduce the fake electrons in data we require there are two HEEP electrons and at least one in barrel for the data event. So the result from MC in barrel is very good, for endcap the peak has around 1-2% shift. The result from data is seems fine, the value from MAV has 1% lower than the reconstructed energy (calo energy).

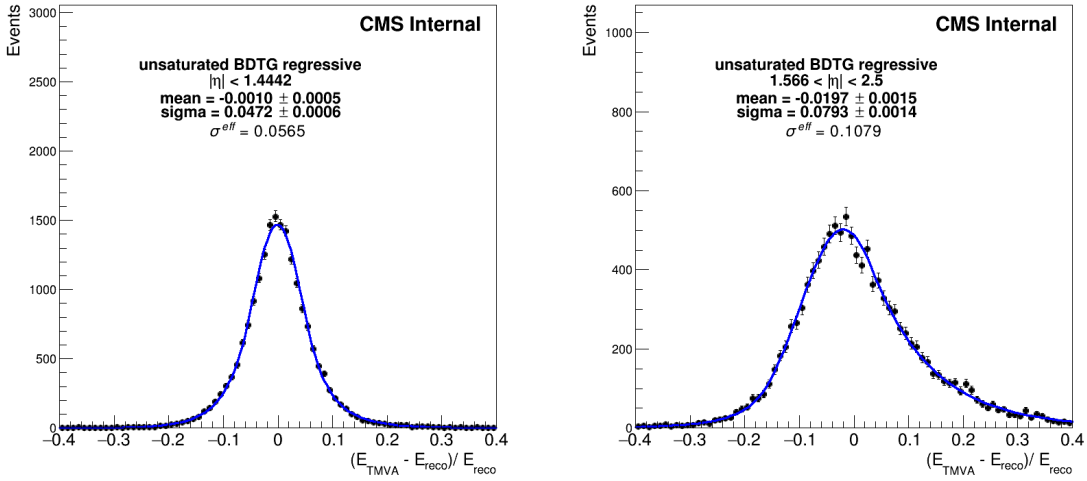

Figure B.52: The distributions of unsaturated HEEP electrons energy from MVA minus reconstructed energy devide reconstructed energy for barrel for the reconstructed energy between 500 to 700 GeV (left) and for endcap for the reconstructed energy between 600 to 1000 GeV (right) from MC.

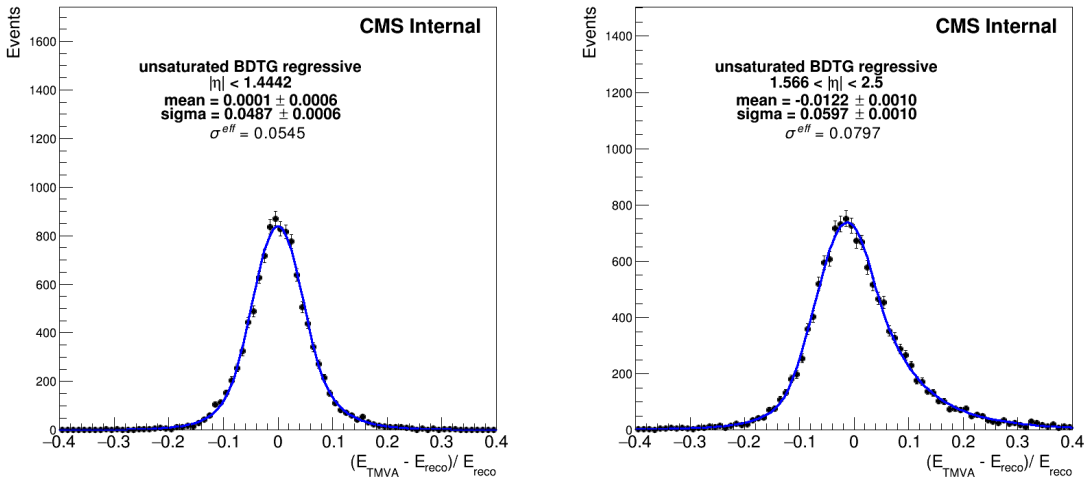

Figure B.53: The distributions of unsaturated HEEP electrons energy from MVA minus reconstructed energy devide reconstructed energy for barrel for the reconstructed energy more than 700 GeV (left) and for endcap for the reconstructed energy more than 1000 GeV (right) from MC.

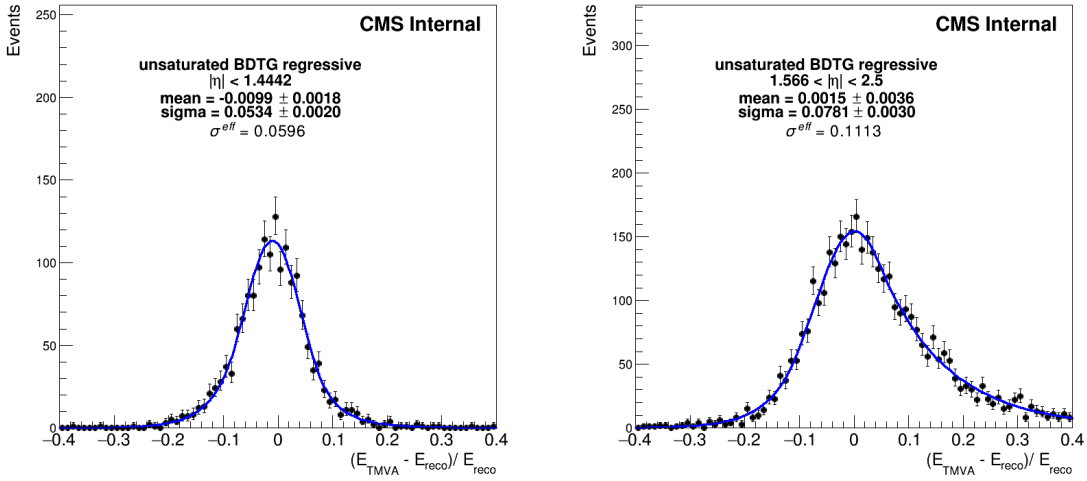

Figure B.54: The distributions of unsaturated HEEP electrons energy from MVA minus reconstructed energy devide reconstructed energy for barrel for the reconstructed energy between 500 to 700 GeV (left) and for endcap for the reconstructed energy between 600 to 1000 GeV (right) from data.

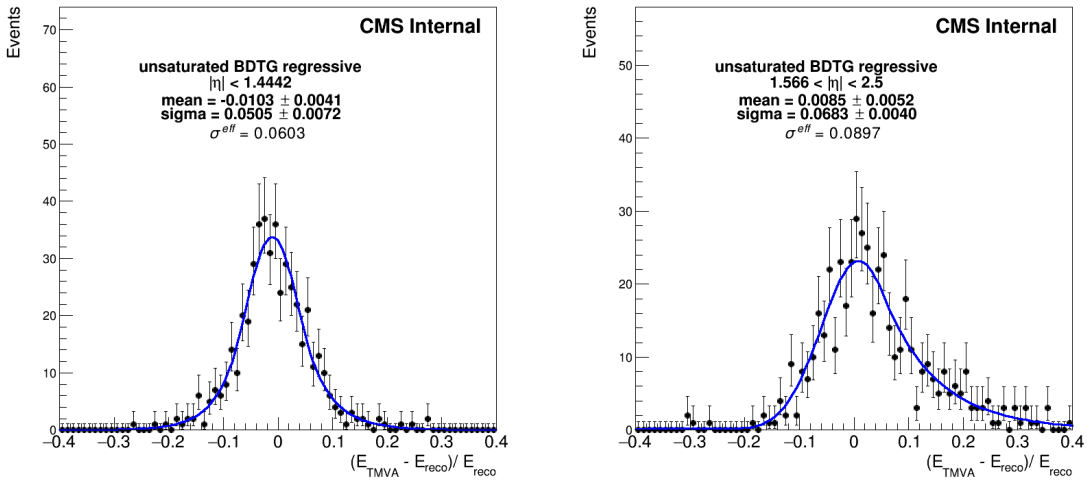

Figure B.55: The distributions of unsaturated HEEP electrons energy from MVA minus reconstructed energy devide reconstructed energy for barrel for the reconstructed energy more than 700 GeV (left) and for endcap for the reconstructed energy more than 1000 GeV (right) from data.

B.4.3 Saturation effect to HEEP ID efficiency

Here we use sample `/DoubleElectron_FlatPt-300To6500/RunIISpring16DR80-PUFlat0to50_80X_mcRun2_asymptotic_2016_v3-v1/AODSIM` to study the effect of saturation to HEEP ID efficiency. The HEEP ID efficiency for saturated and unsaturated electron is shown in Figure B.56 and one can see the HEEP ID is safe for saturated electron in endcap while for barrel it works well when the energy is lower than around 3.2 TeV. The reason for the HEEP ID efficiency decrease quickly with energy for saturated electron in barrel is because of the showershape cut which can be seen in Figure B.57 and the efficiency of HEEP ID without showershape is shown in Figure B.58. The difference of the efficiency between saturated electron and unsaturated electron in Figure B.58 in barrel is mainly from $\frac{H}{E}$ cut and in endcap it is mainly from EcalDriven cut.

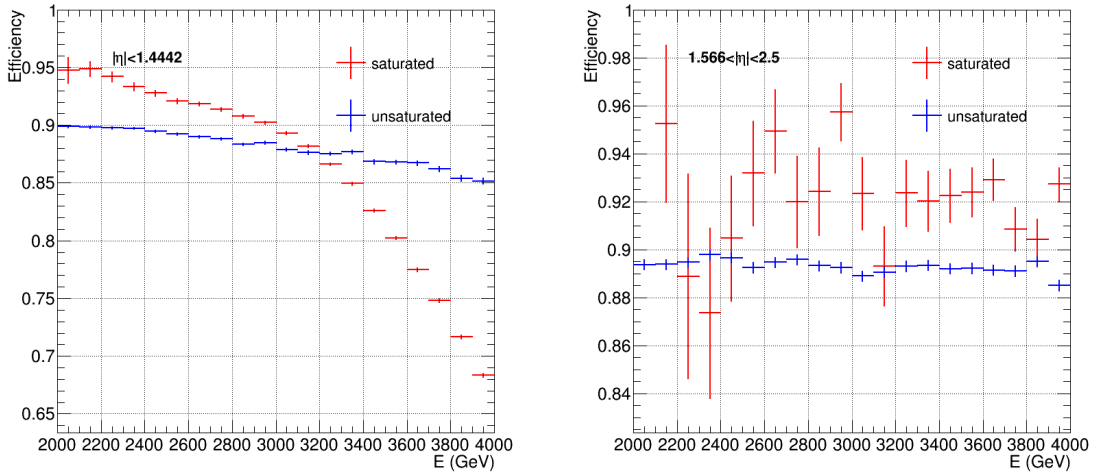

Figure B.56: The HEEP ID efficiency for saturated electron (red histogram) and unsaturated electron (blue histogram) for barrel (left) and endcap (right).

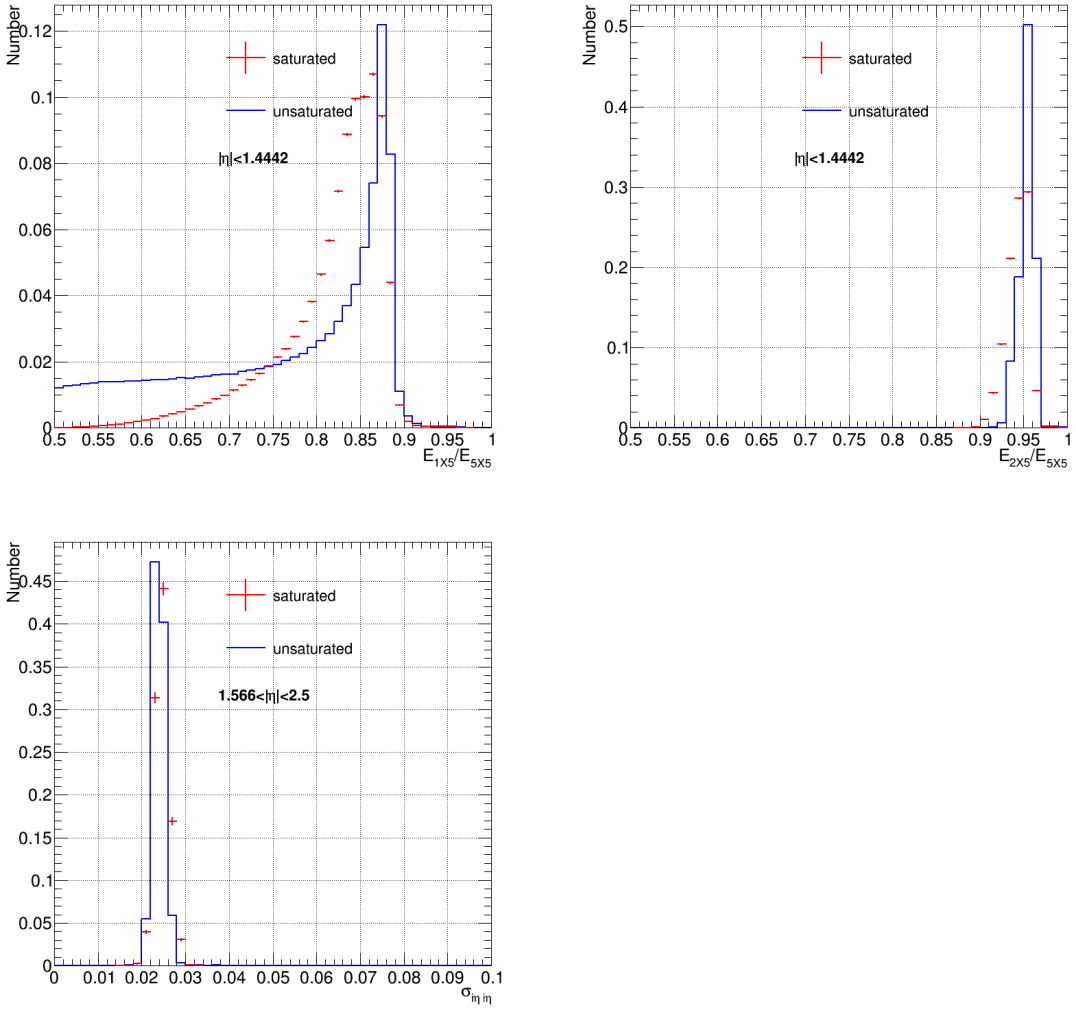

Figure B.57: The distributions of the value of HEEP ID showershape variables for barrel (top) and endcap (bottom) for saturated (red histogram) and unsaturated (blue histogram) electrons.

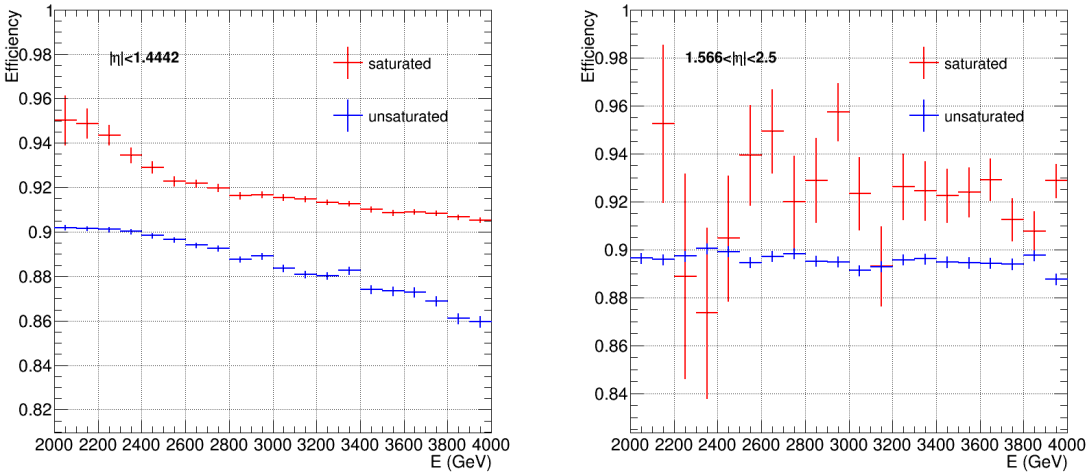

Figure B.58: The HEEP ID without showershape criteria efficiency for saturated (red histogram) and unsaturated electrons (blue histogram) for barrel (left) and endcap (right).

B.4.4 Conclusions

In order to get the correct energy for saturated electron which has very high energy, we tried MVA method and it works well in barrel in wide energy range, for endcap there are a small bias in low erengy because of low statistics for training. Besides we also checked the ECAL linearity response in data which seems good and there are 1% difference between the reconstructed energy and energy from MVA. Finally we studied the effection of saturation to the HEEP ID efficiency which shows the HEEP ID still works well for saturated electron.

B.4.5 Checking with MLP method

Here we replace BDTG method with MLP method to replay the result in Figure B.44. The configuration of the MLP method is `factory->BookMethod( MVA::Types::kMLP, "MLP", "!H:!V:VarTransform=Norm:NeuronType=tanh:NCycles=500:HiddenLayers=N+20: TestRate=6:TrainingMethod=BFGS:Sampling=0.3:SamplingEpoch=0.8: ConvergenceImprove=1e-6:ConvergenceTests=15:!UseRegulator" )`. From the result in Figure B.59 one can see the MLP method also works but comparing with BDTG method its resolution is worse than BDTG method.

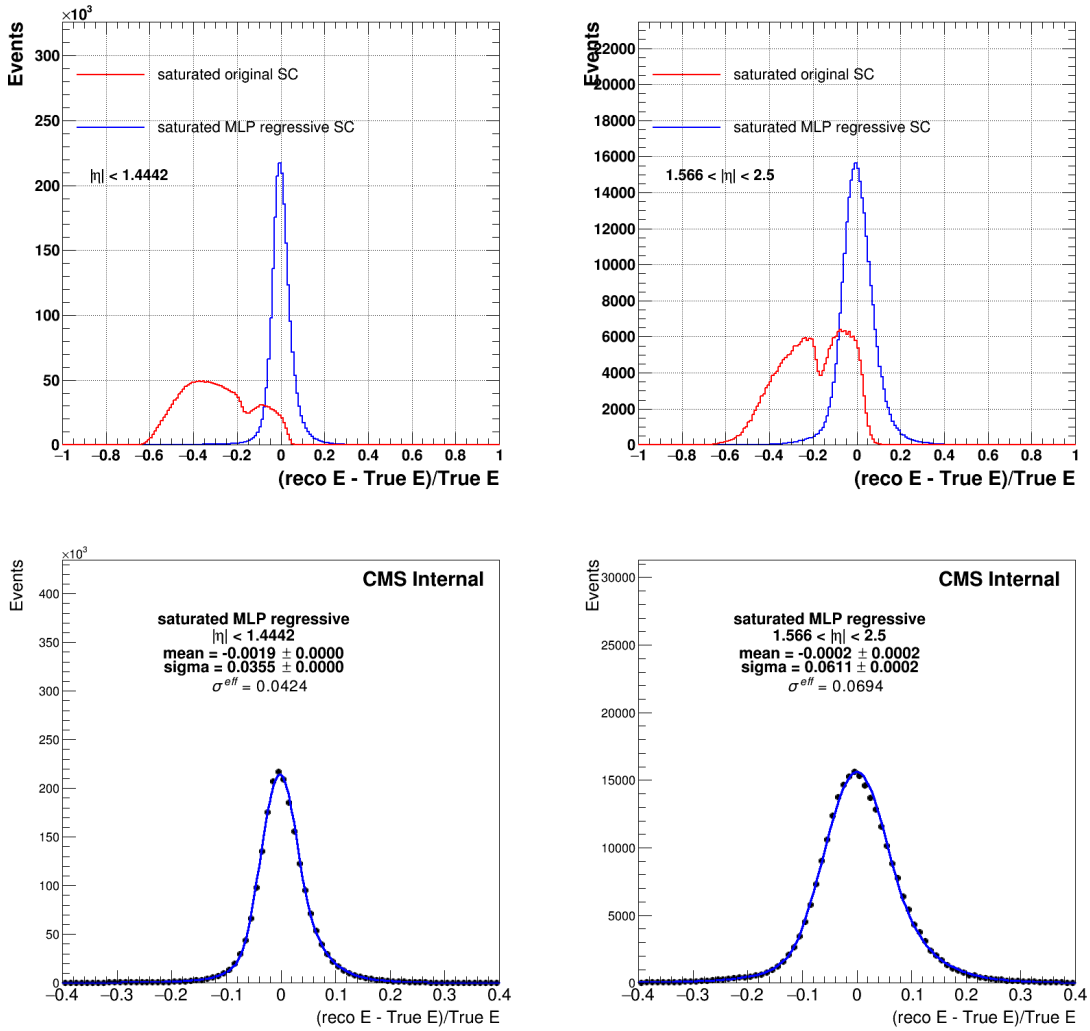

Figure B.59: The top plots are the distribution of supercluster energy minus ture energy divided by true energy for saturated electron for barrel (left) and endcap (right), the red histogram is for reconstructed supercluster enery, the blue histogram is for MVA regressive energy. The bottom plots are the fit of the blue histogram for barrel (left) and endcap (right).

B.5 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 6.1. The problem is present in same flavor channel mostly from DY process which has not real MET.

B.5.0.1 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 [162]

• 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 B.60, the distribution of $|\Delta\phi|$(MET,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 $|\Delta\phi|$(MET,ll)<2 and looked at MET distribution which can be seen in Figure B.60. 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 B.61, 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 B.62, 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 6.1) although the ratio of data MC normalization is not 1 because of the fact that pile-up reweighting does not 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 B.63. The data MC agreement has improved clearly.

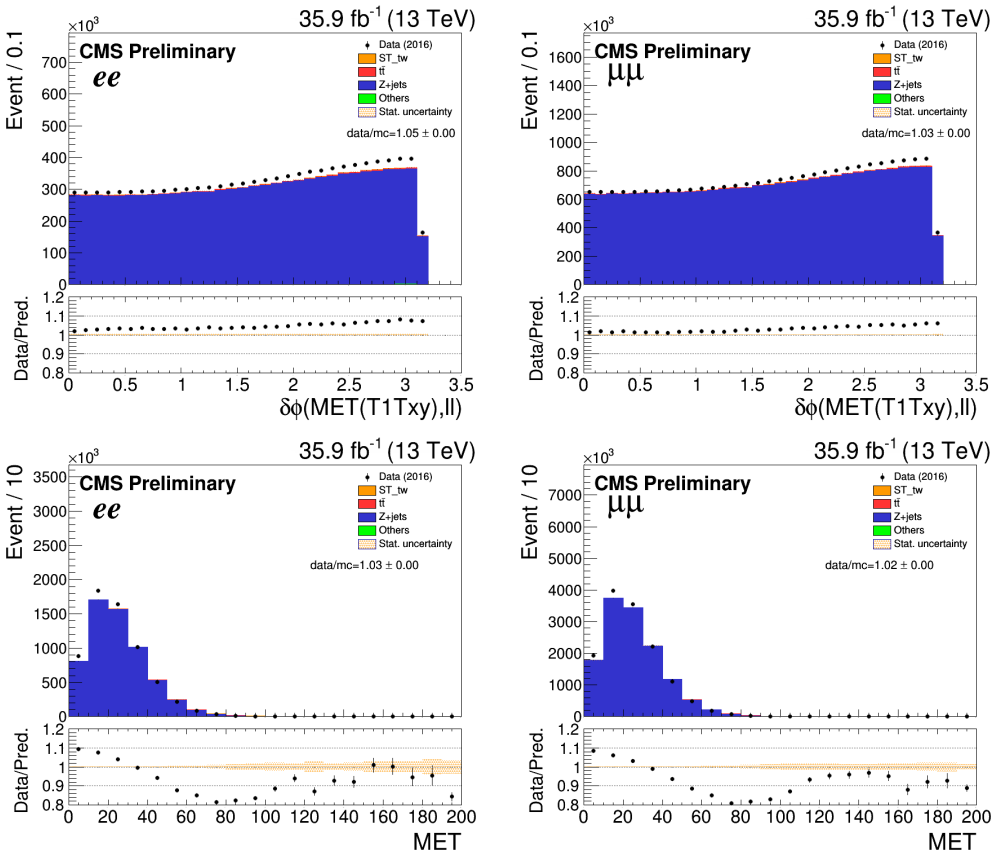

Figure B.60: The distributions of the $\Delta\phi$ (MET,ll) (first row), MET (second row) for events with $\Delta\phi$ (MET,ll)<2, for ee (left) and $\mu\mu$ (right) channels after step 1.

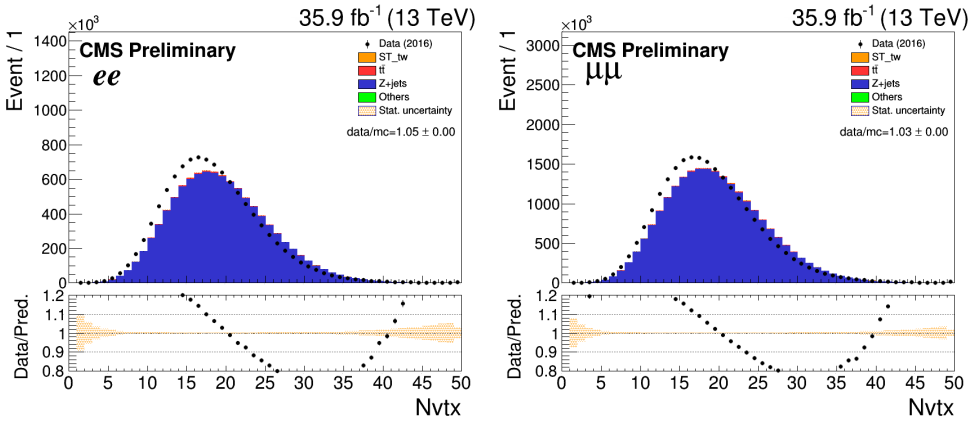

Figure B.61: The distributions of number of reconstructed vertices for ee (left) and $\mu\mu$ (right) channels after step 1.

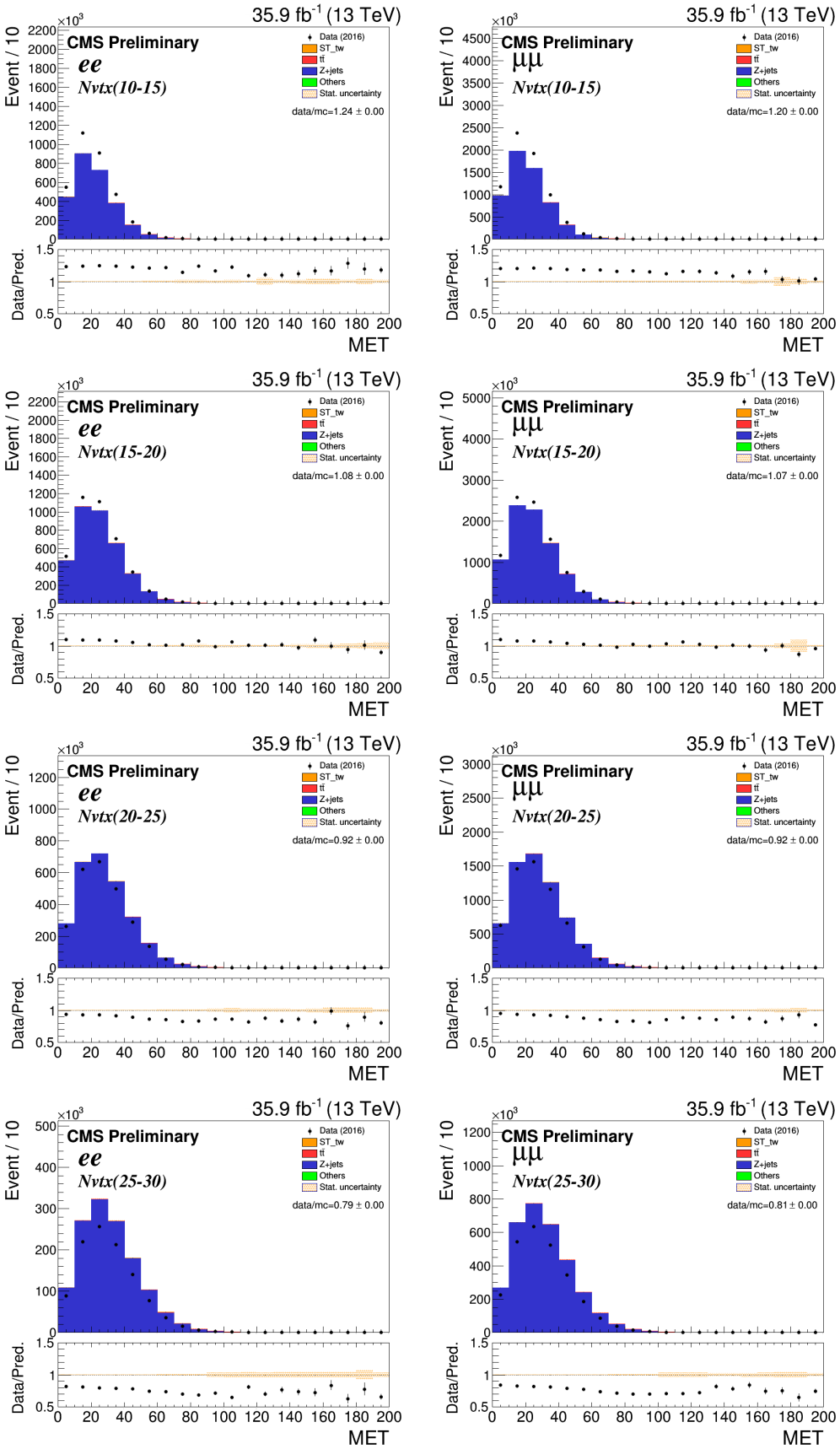

Figure B.62: The distributions of MET for events with number of vertices between 10-15 (first row), 15-20 (second row), 20-25 (third row) and 25-30 (last row) for ee (left) and μμ (right) channels.

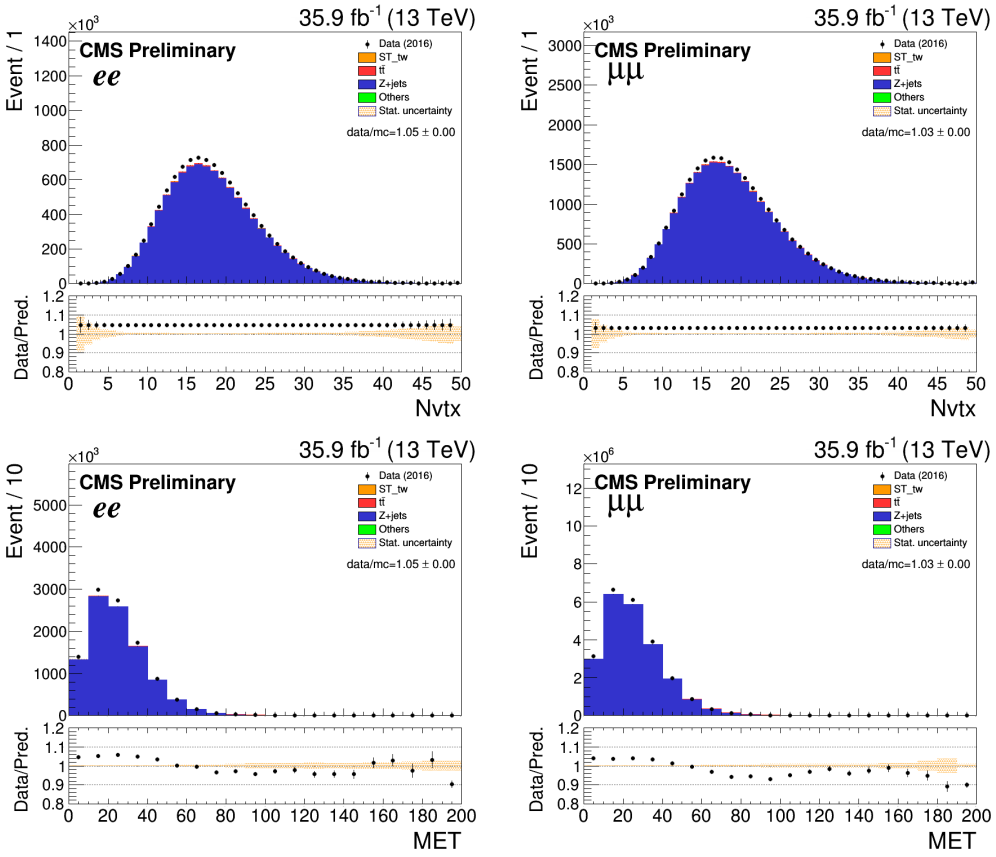

Figure B.63: The distributions of number of vertices (top) and MET (bottom) after number of vertices re-weighting for ee (left) and μμ (right) channels.

B.6 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 6.7) output distributions for both data and MC expectation in the (1jet,1b-jet) and (2jet,1b-jet) regions and event yield in the (≥2jet,2b-jet) region for ee and μμ channels. The inclusion of the (≥2jet,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 B.64. All sources of systematic uncertainties discussed in Section 6.8 are included in our results.

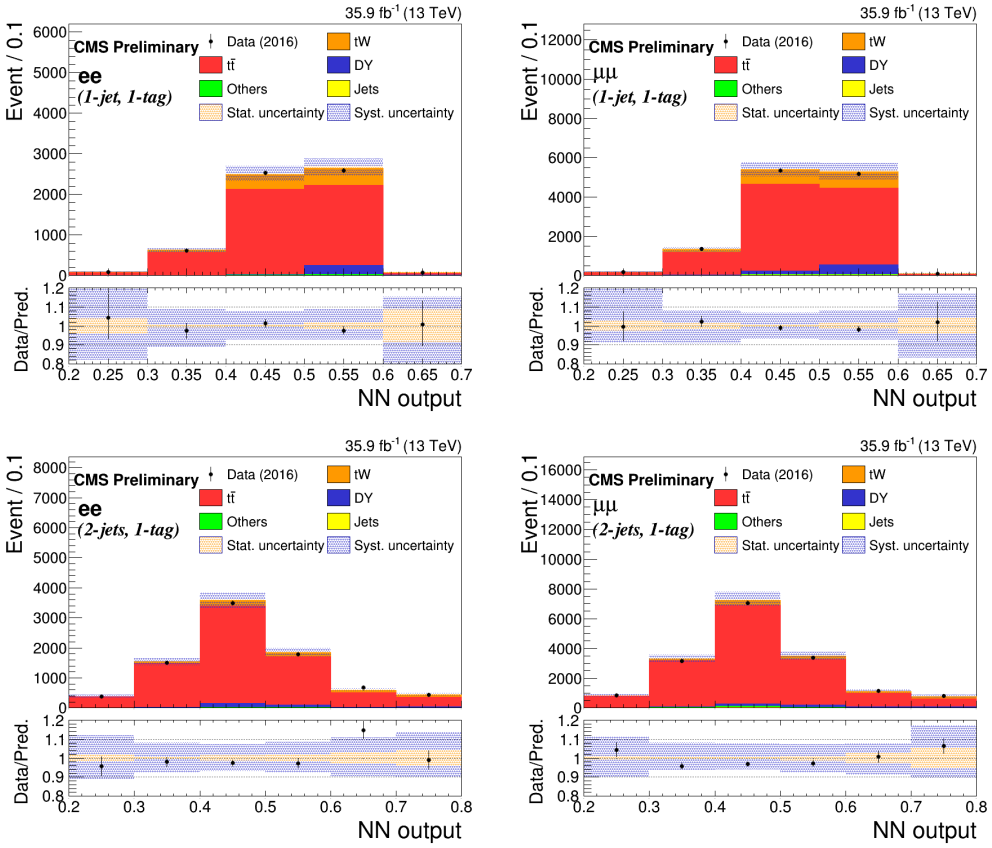

Figure B.64: The MLP distributions for different (1jet,1b-jet) region (top row), (2jet,1b-jet) region (bottom row) for ee channel (left column) and μμ channel (right column).

In Figure B.65, likelihood scan for the signal strength are shown for different region for ee and μμ channels. In Table B.3, tW measured cross section is reported for ee and μμ channels compared with the result from $e\mu$ channel and combined channels. In figures B.66 and B.67, impact of individual systematic sources for ee, μμ, $e\mu$ and combined channels are shown for expected and observed. In Table B.4, the effect of each systematic uncertainty source to the combined fit is shown.

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

Table B.3: 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 B.4: 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%

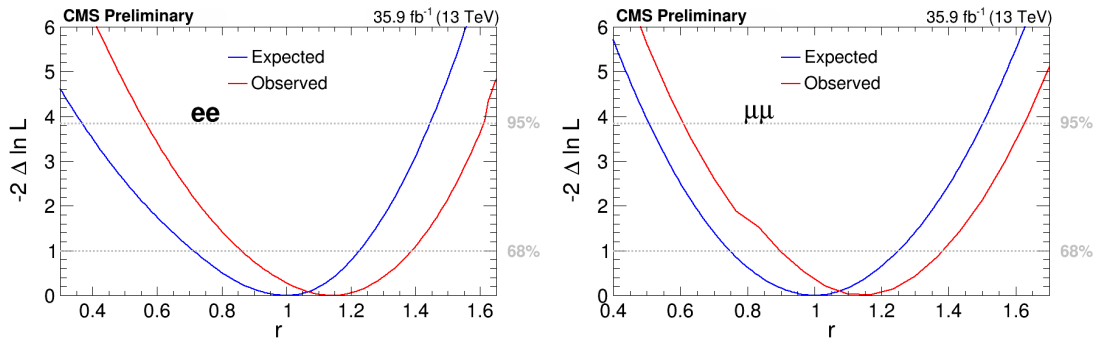

Figure B.65: The continue likelihood scan for various regions for ee channel (left) and μμ (right) channels.

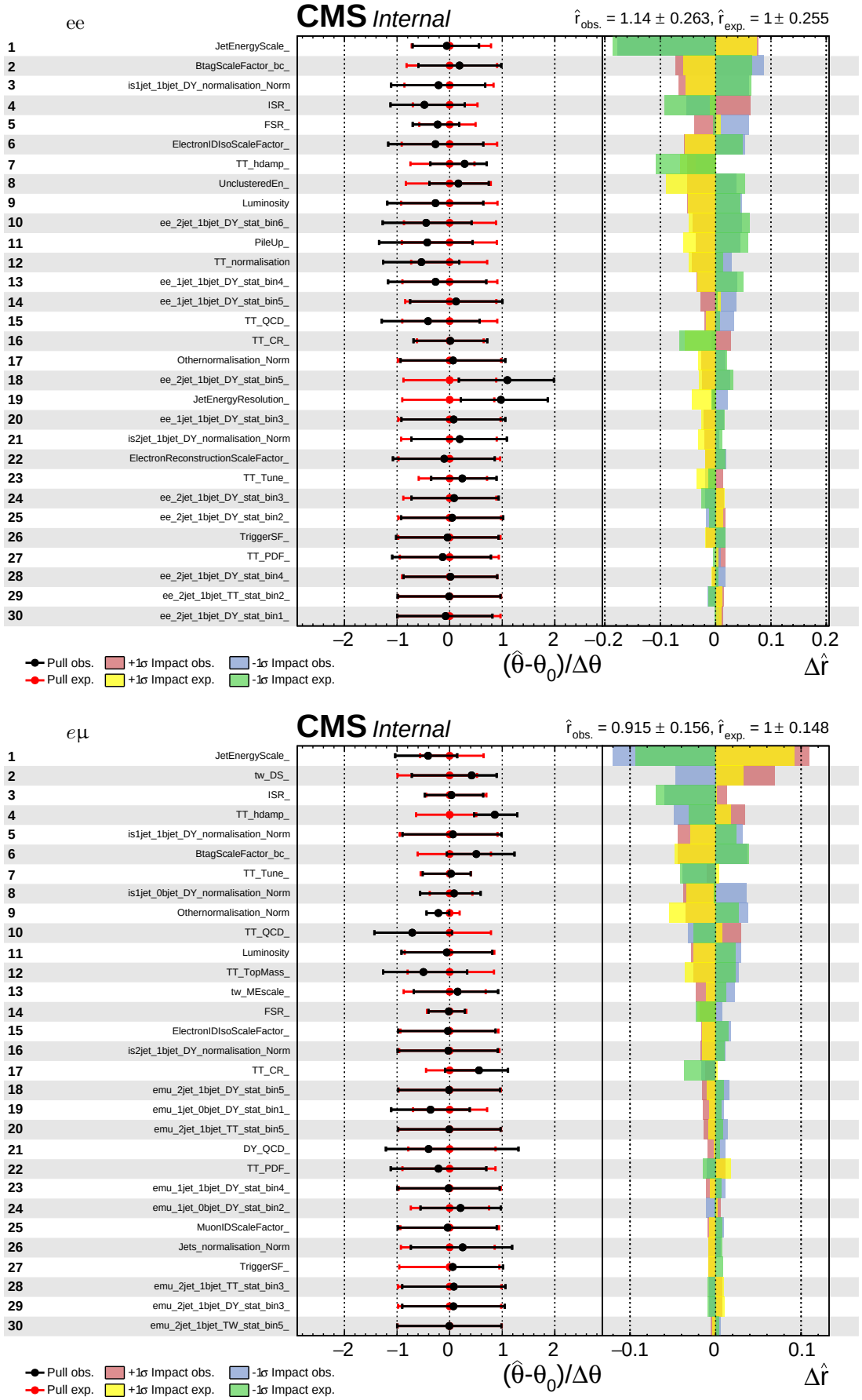

Figure B.66: The expected and observed impacts of the most important uncertainty sources on the measurement of tW cross section in ee, eμ channels.

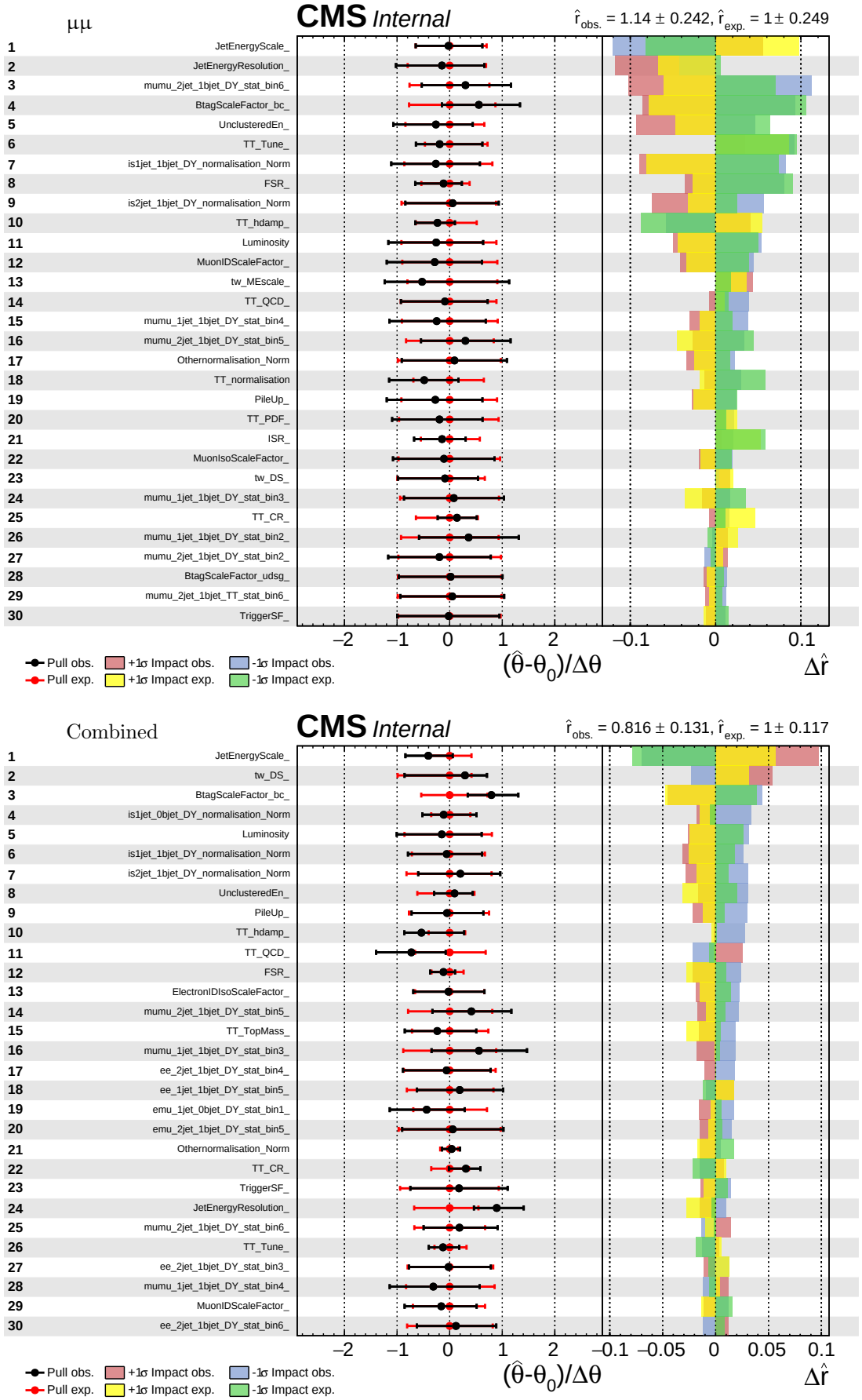

Figure B.67: The expected and observed impacts of the most important uncertainty sources on the measurement of tW cross section in μμ channel and combined.

Bibliography

[1] David Griffiths. “Introduction to elementary particles”. In: (1987).

[2] Alan D. Martin Francis Halzen. “QUARKS AND LEPTONS: An Introductory Course in Modern Particle Physics”. In: (1984).

[3] S.N. Mukherjee T. Morii C.S. Lim. “The Physics of the Standard Model and Beyond”. In: (2004).

[4] The ATLAS Collaboration. “Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC”. In: *Phys. Lett.* B716 (2012), pp. 1–29. DOI: 10.1016/j.physletb.2012.08.020. arXiv:1207.7214 [hep-ex].

[5] The CMS Collaboration. “Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC”. In: *Phys. Lett.* B716 (2012), pp. 30–61. DOI: 10.1016/j.physletb.2012.08.021. arXiv:1207.7235 [hep-ex].

[6] Sidney D Drell and Tung-Mow Yan. “Partons and their applications at high energies”. In: *Annals of Physics* 66.2 (1971), pp. 578 –623. ISSN: 0003-4916. DOI: https://doi.org/10.1016/0003-4916(71)90071-6. URL: http://www.sciencedirect.com/science/article/pii/0003491671900716.

[7] Vera C. Rubin and W. Kent Ford Jr. “Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions”. In: *Astrophys. J.* 159 (1970), pp. 379–403. DOI: 10.1086/150317.

[8] Katherine Freese. “Review of Observational Evidence for Dark Matter in the Universe and in upcoming searches for Dark Stars”. In: *EAS Publ. Ser.* 36 (2009), pp. 113–126. DOI: 10.1051/eas/0936016. arXiv:0812.4005 [astro-ph].

[9] Y. Fukuda et al. “Evidence for oscillation of atmospheric neutrinos”. In: *Phys. Rev. Lett.* 81 (1998), pp. 1562–1567. DOI: 10.1103/PhysRevLett.81.1562. arXiv:9807003 [hep-ex].

[10] Y. Abe et al. “Indication for the disappearance of reactor electron antineutrinos in the Double Chooz experiment”. In: *Phys. Rev. Lett.* 108 (2012), p. 131801. DOI: 10.1103/PhysRevLett.108.131801. arXiv:1112.6353 [hep-ex].

[11] R. Barbier et al. “R-parity violating supersymmetry”. In: *Phys.Rept.* 420 (2005). DOI: 10.1016/j.physrep.2005.08.006. arXiv:0406039 [hep-ex].

[12] G. Altarelli, B. Mele, and M. Ruiz-Altaba. “Searching for new heavy vector bosons in $p\bar{p}$ colliders”. In: *Z. Phys. C* 45 (1989), p. 109. DOI: 10.1007/BF01556677.

[13] A. Leike. “The Phenomenology of extra neutral gauge bosons”. In: *Phys.Rept.* 317 (1999), p. 143. DOI: 10.1016/S0370-1573(98)00133-1.

[14] J. L. Hewett and T. G. Rizzo. “Low-Energy Phenomenology of Superstring Inspired E(6) Models”. In: *Phys.Rept.* 183 (1989), p. 193. DOI: 10.1016/0370-1573(89)90071-9.

[15] N. Arkani-Hamed, S. Dimopoulos, and G. Dvali. “The hierarchy problem and new dimensions at a millimeter”. In: *Phys. Lett. B* 429 (1998), p. 263. DOI: 10.1016/S0370-2693(98)00466-3. arXiv:9803315 [hep-ph].

[16] L. Randall and R. Sundrum. “A Large mass hierarchy from a small extra dimension”. In: *Phys. Rev. Lett.* 83 (1999), p. 3370. DOI: 10.1103/PhysRevLett.83.3370.

[17] Battistoni et al. “The Application of the Monte Carlo Code FLUKA in Radiation Protection Studies for the Large Hadron Collider”. In: *Progress in Nuclear Science and Technology* 2 (Oct. 2011), pp. 358–364. DOI: 10.15669/pnst.2.358.

[18] CMS. “Public CMS Luminosity Information”. In: *CMS Twiki* https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults ().

[19] Antonella Del Rosso. "HL-LHC updates in Japan. Projet HL-LHC : une r茅union fait le point au Japon". In: BUL-NA-2014-272. 51/2014 (2014), p. 4. URL: https://cds.cern.ch/record/1975962.

[20] S. Chatrchyan et al. "The CMS Experiment at the CERN LHC". In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.

[21] CERN. "CERN website". In: http://cms.web.cern.ch/news/cms-detector-design (2011).

[22] CERN. "CERN website". In: http://cms.web.cern.ch/org/cms-presentations-public (2011).

[23] CMS Collboration. "CMS Physics: Technical Design Report Volume 1: Detector Performance and Software". In: *Technical Design Report CMS. CERN, Geneva* http://cds.cern.ch/record/922757 (2016).

[24] CMS Collboration. "The Electromagnetic Calorimeter Technical Design Report". In: *Technical Design Report CMS* http://cds.cern.ch/record/349375/files/ (1997).

[25] *The CMS hadron calorimeter project: Technical Design Report*. Technical Design Report CMS. The following files are from <a href=. Geneva: CERN, 1997. URL: http://cds.cern.ch/record/357153.

[26] Serguei Chatrchyan et al. "Search for resonances in the dilepton mass distribution in pp collisions at $\sqrt{s} = 7$ TeV". In: *JHEP* 05 (2011), p. 093. DOI: 10.1007/JHEP05(2011)093. arXiv:1103.0981 [hep-ex].

[27] Serguei Chatrchyan et al. "Search for narrow resonances in dilepton mass spectra in pp collisions at $\sqrt{s} = 7$ TeV". In: *Phys. Lett. B* 714 (2012), p. 158. DOI: 10.1016/j.physletb.2012.06.051. arXiv:1206.1849 [hep-ex].

[28] Serguei Chatrchyan et al. "Search for heavy narrow dilepton resonances in pp collisions at $\sqrt{s} = 7$ TeV and $\sqrt{s} = 8$ TeV". In: *Phys. Lett. B* 720 (2013), p. 63. DOI: 10.1016/j.physletb.2013.02.003. arXiv:1212.6175 [hep-ex].

[29] Vardan Khachatryan et al. "Search for physics beyond the standard model in dilepton mass spectra in proton-proton collisions at $\sqrt{s} = 8$ TeV". In: *JHEP* 04 (2015), p. 025. DOI: 10.1007/JHEP04(2015)025. arXiv:1412.6302 [hep-ex].

[30] The CMS Collaboration. "Search for narrow resonances in dilepton mass spectra in proton-proton collisions at $\sqrt{s} = 13$ TeV and combination with 8 TeV data". In: *Phys. Lett.* B768 (2017), pp. 57–80. DOI: 10.1016/j.physletb.2017.02.010. arXiv:1609.05391 [hep-ex].

[31] The CMS Collaboration. "Search for high-mass resonances in dilepton final states in proton-proton collisions at $\sqrt{s} = 13$ 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.

[32] The CMS Collaboration. "Search for high mass resonances in dielectron final state". In: *CMS Physics Analysis Summaries* CMS-PAS-EXO-18-006 (2018).

[33] Georges Aad et al. "Search for high mass dilepton resonances in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS experiment". In: *Phys. Lett. B* 700 (2011), p. 163. DOI: 10.1016/j.physletb.2011.04.044. arXiv:1103.6218 [hep-ex].

[34] Georges Aad et al. "Search for high-mass resonances decaying to dilepton final states in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector". In: *JHEP* 11 (2012), p. 138. DOI: 10.1007/JHEP11(2012)138. arXiv:1209.2535 [hep-ex].

[35] Georges Aad et al. "Search for high-mass dilepton resonances in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector". In: *Phys. Rev. D* 90 (2014), p. 052005. DOI: 10.1103/PhysRevD.90.052005. arXiv:1405.4123 [hep-ex].

[36] Morad Aaboud et al. "Search for new high-mass phenomena in the dilepton final state using 36.1 fb^{-1} of proton-proton collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector". In: *JHEP* 10 (2017), p. 182. DOI: 10.1007/JHEP10(2017)182. arXiv:1707.02424 [hep-ex].

[37] Georges Aad et al. "Search for high-mass dilepton resonances using 139 fb^{-1} of pp collision data collected at $\sqrt{s}$ =13 TeV with the ATLAS detector". In: (2019). arXiv:1903.06248 [hep-ex].

[38] T. Aaltonen et al. "Search for high-mass e^+e^- resonances in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV". In: *Phys. Rev. Lett.* 102 (2009), p. 031801. DOI: 10.1103/PhysRevLett.102.031801. arXiv:0810.2059 [hep-ex].

[39] T. Aaltonen et al. "A search for high-mass resonances decaying to dimuons at CDF". In: *Phys. Rev. Lett.* 102 (2009), p. 091805. DOI: 10.1103/PhysRevLett.102.091805. arXiv:0811.0053 [hep-ex].

[40] Victor Mukhamedovich Abazov et al. "Search for Randall–Sundrum gravitons in the dielectron and diphoton final states with 5.4 fb^{-1} of data from $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV". In: *Phys. Rev. Lett.* 104 (2010), p. 241802. DOI: 10.1103/PhysRevLett.104.241802. arXiv:1004.1826 [hep-ex].

[41] Victor Mukhamedovich Abazov et al. "Search for a heavy neutral gauge boson in the dielectron channel with 5.4 fb^{-1} of $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV". In: *Phys. Lett. B* 695 (2011), p. 88. ISSN: 0370-2693. DOI: 10.1016/j.physletb.2010.10.059. arXiv:1008.2023 [hep-ex].

[42] T. Aaltonen et al. "Search for high mass resonances decaying to muon pairs in $\sqrt{s} = 1.96$ TeV $p\bar{p}$ collisions". In: *Phys. Rev. Lett.* 106 (2011), p. 121801. DOI: 10.1103/PhysRevLett.106.121801. arXiv:1101.4578 [hep-ex].

[43] T. Aaltonen et al. "Search for new dielectron resonances and Randall–Sundrum gravitons at the Collider Detector at Fermilab". In: *Phys. Rev. Lett.* 107 (2011), p. 051801. DOI: 10.1103/PhysRevLett.107.051801. arXiv:1103.4650 [hep-ex].

[44] 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].

[45] 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].

[46] 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].

[47] 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].

[48] 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].

[49] 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].

[50] 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].

[51] 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].

[52] 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).

[53] Particle Data Group. "Review of Particle Physics". In: *Chin. Phys.* C40.10 (2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001.

[54] Alan D. Martin Francis Halzen. "QUARKS AND LEPTONS: An introductory Course in Modern Particle Physics". In: JOHN WILEY and SONS (1984).

[55] Emmy Noether. "Invariant variation problems". In: *Transport Theory and Statistical Physics* 1.3 (1971), pp. 186–207. DOI: 10.1080/00411457108231446. eprint: https://doi.org/10.1080/00411457108231446. URL: https://doi.org/10.1080/00411457108231446.

[56] F. Englert and R. Brout. "Broken Symmetry and the Mass of Gauge Vector Mesons". In: *Phys. Rev. Lett.* 13 (9 1964), pp. 321–323. DOI: 10.1103/PhysRevLett.13.321. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.321.

[57] P.W. Higgs. "Broken symmetries, massless particles and gauge fields". In: *Physics Letters* 12.2 (1964), pp. 132 –133. ISSN: 0031-9163. DOI: https://doi.org/10.1016/0031-9163(64)91136-9. URL: http://www.sciencedirect.com/science/article/pii/0031916364911369.

[58] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (16 1964), pp. 508–509. DOI: 10.1103/PhysRevLett.13.508. URL: https://link.aps.org/doi/10.1103/PhysRevLett.13.508.

[59] Steven Weinberg. "A Model of Leptons". In: *Phys. Rev. Lett.* 19 (21 1967), pp. 1264–1266. DOI: 10.1103/PhysRevLett.19.1264. URL: https://link.aps.org/doi/10.1103/PhysRevLett.19.1264.

[60] Abdus Salam and J. C. Ward. "Weak and electromagnetic interactions". In: *Il Nuovo Cimento (1955-1965)* 11.4 (1959), pp. 568–577. ISSN: 1827-6121. DOI: 10.1007/BF02726525. URL: https://doi.org/10.1007/BF02726525.

[61] Sheldon L. Glashow. "The renormalizability of vector meson interactions". In: *Nuclear Physics* 10 (1959), pp. 107 –117. ISSN: 0029-5582. DOI: https://doi.org/10.1016/0029-5582(59)90196-8. URL: http://www.sciencedirect.com/science/article/pii/0029558259901968.

[62] B.R. WEBBER R.K. ELLIS W.J. STIRLING. "QCD and Collider Physics". In: (1996).

[63] CODATA Value: Fermi coupling constant. "The NIST Reference on Constants, Units, and Uncertainty". In: *US National Institute of Standards and Technology* (June 2015. Retrieved 2016-10-31).

[64] The CMS collaboration. "Measurement of the Drell-Yan cross section in pp collisions at $\sqrt{s} = 7$ TeV". In: *Journal of High Energy Physics* 2011.10 (2011), p. 7. ISSN: 1029-8479. DOI: 10.1007/JHEP10(2011)007. URL: https://doi.org/10.1007/JHEP10(2011)007.

[65] Dimitri Bourilkov. "Exploring the LHC Landscape with Dileptons". In: *arXiv* 1609.08994 (2017).

[66] M. Peskin. "Beyond the Standard Model". In: *Lectures notes of the 1996 European School of High-Energy Physics* (Carry-le-Rouet, France, 1996).

[67] K.A. Olive and Particle Data Group. "Review of Particle Physics". In: *Chinese Physics C* 38.9 (2014), p. 090001. URL: http://stacks.iop.org/1674-1137/38/i=9/a=090001.

[68] Adam G. Riess et al. "Observational evidence from supernovae for an accelerating universe and a cosmological constant". In: *Astron. J.* 116 (1998), pp. 1009–1038. DOI: 10.1086/300499. arXiv:astro-ph/9805201 [astro-ph].

[69] S. Perlmutter et al. "Measurements of Omega and Lambda from 42 high redshift supernovae". In: *Astrophys. J.* 517 (1999), pp. 565–586. DOI: 10.1086/307221. arXiv:astro-ph/9812133 [astro-ph].

[70] A D. SAKHAROV. "Violation of CP invariance, 小asymmetry, and baryon asymmetry of the universe". In: vol. 5. Nov. 1998, pp. 84–87. ISBN: 978-981-02-3606-9. DOI: 10.1142/9789812815941_0013.

[71] Nicola Cabibbo. "Unitary Symmetry and Leptonic Decays". In: *Phys. Rev. Lett.* 10 (12 1963), pp. 531–533. DOI: 10.1103/PhysRevLett.10.531. URL: https://link.aps.org/doi/10.1103/PhysRevLett.10.531.

[72] Makoto Kobayashi and Toshihide Maskawa. "CP-Violation in the Renormalizable Theory of Weak Interaction". In: *Progress of Theoretical Physics* 49.2 (Feb. 1973), pp. 652–657. ISSN: 0033-068X. DOI: 10.1143/PTP.49.652. eprint: http://oup.prod.sis.lan/ptp/article-pdf/49/2/652/5257692/49-2-652.pdf. URL: https://doi.org/10.1143/PTP.49.652.

[73] T. Schäfer and E. V. Shuryak. "Instantons in QCD". In: *Rev. Mod. Phys.* 70 (2 1998), pp. 323–425. DOI: 10.1103/RevModPhys.70.323. URL: https://link.aps.org/doi/10.1103/RevModPhys.70.323.

[74] The CMS and ATLAS Collobarations. "Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV". In: *JHEP* 08 (2016), p. 045. DOI: 10.1007/JHEP08(2016)045. arXiv:1606.02266 [hep-ex].

[75] Stephen P. Martin. "A Supersymmetry primer". In: *Adv. Ser. Direct. High Energy Phys.18,1* (1997). DOI: 10.1142/9789812839657_0001, 10.1142/9789814307505_0001. arXiv:hep-ph/9709356 [hep-ph].

[76] R. Barbier et al. "R-Parity-violating supersymmetry". In: *Physics Reports* 420.1 (2005), pp. 1 – 195. ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2005.08.006. URL: http://www.sciencedirect.com/science/article/pii/S0370157305003327.

[77] H. Georgi and S. Glashow. "Unity of All Elementary Particle Forces". In: *Phys. Rev. Lett.* 32 (1974), p. 438. DOI: 10.1103/PhysRevLett.32.438.

[78] R. C. King, L. Dehuai, and B. G. Wybourne. "Symmetrized Powers of Rotation Group Representations". In: *J. Phys.* A14 (1981), pp. 2509–2538. DOI: 10.1088/0305-4470/14/10/009.

[79] Edward Witten. "Symmetry Breaking Patterns in Superstring Models". In: *Nucl. Phys.* B258 (1985), p. 75. DOI: 10.1016/0550-3213(85)90603-0.

[80] Edward Witten. "Symmetry breaking patterns in superstring models". In: *Nuclear Physics B* 258 (1985), pp. 75 –100. ISSN: 0550-3213. DOI: https://doi.org/10.1016/0550-3213(85)90603-0. URL: http://www.sciencedirect.com/science/article/pii/0550321385906030.

[81] John H. Schwarz. "Introduction to superstring theory". In: *NATO Sci. Ser. C* 566 (2001), pp. 143–187. DOI: 10.1007/978-94-010-0522-7_4. arXiv:0008017 [hep-ex].

[82] Cen Zhang and Scott Willenbrock. "Effective-Field-Theory Approach to Top-Quark Production and Decay". In: *Phys. Rev.* D83 (2011), p. 034006. DOI: 10.1103/PhysRevD.83.034006. arXiv:1008.3869 [hep-ph].

[83] 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].

[84] 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].

[85] 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].

[86] 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).

[87] CERN. "CERN website". In: http://home.cern/fr (2015).

[88] 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.

[89] E. Meschi at al. "Electron Reconstruction in the CMS Electromagnetic Calorimeter". In: *CMS Notes* CMS Note-2001-034 (2001).

[90] 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].

[91] 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.

[92] 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.

[93] 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.

[94] The CMS Collaboration. *The CMS Physics Technical Design Report*. CERN/LHCC 2006-001, CMS TDR 8.1. 2006.

[95] Sanmay Ganguly. *Jet Measurements In CMS*. Tech. rep. arXiv:1303.6038. 2013. URL: https://cds.cern.ch/record/1532824.

[96] 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.

[97] 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).

[98] 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).

[99] 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.

[100] 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].

[101] 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].

[102] 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].

[103] 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].

[104] 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].

[105] 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].

[106] 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].

[107] 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.

[108] 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.

[109] The CMS Collaboration. "Generic Tag and Probe Tool for Measuring Efficiency at CMS with Early Data". In: *CMS Notes* CMS AN-2009-111 (2009).

[110] B. Clerbaux at al. "Search for high mass dielectron resonances with the full 2016 data". In: *CMS Notes* CMS AN-2016-404 (2017).

[111] B. Clerbaux at al. "Search for high-mass resonances in the di-electron final state with 2017 data". In: *CMS Notes* CMS AN-2018-021 (2018).

[112] . "HEEP ID". In: *Twiki* https://twiki.cern.ch/twiki/bin/view/CMS/HEEPElectronIdentificationRun2?rev=r23 ().

[113] 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).

[114] 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.

[115] 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.

[116] 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.

[117] 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).

[118] S. Folgueras, N. Neumeister, and J.-F. Schulte. *Statistical Analysis for a Search for a Narrow Resonance*. CMS Note 2016/307. CERN, 2016.

[119] S. Schmitz et al. "Statistical Inference in a Search for a Narrow Resonance". In: *CMS Notes* CMS-AN-2012-185 (2012).

[120] V. Giakoumopoulou et al. "Z′ stats meeting". In: *CMS Indico Pages* (2015). URL: https://indico.cern.ch/event/387187/.

[121] 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.

[122] 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.

[123] The CMS Collaboration. "Search for high-mass resonances in dilepton final states in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *CMS Physics Analysis Summaries* CMS-PAS-EXO-16-047 (2016).

[124] 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).

[125] T Arndt, C.D Pardos, and A Meyer. "Measurement of the 2016 Trigger Efficiencies for a dilepton selection for a ttbar analysis". In: *AN-2016-392* (2016).

[126] https://twiki.cern.ch/twiki/bin/viewauth/CMS/SWGuideMuonIdRun2.

[127] "Commissioning of the Particle-Flow reconstruction in Minimum-Bias and Jet Events from pp Collisions at 7 TeV". In: (2010).

[128] https://twiki.cern.ch/twiki/bin/view/CMS/MuonWorkInProgressAndPagResults.

[129] https://twiki.cern.ch/twiki/bin/viewauth/CMS/RochcorMuon.

[130] https://twiki.cern.ch/twiki/bin/view/CMS/EgammaIDRecipesRun2.

[131] 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.

[132] https://twiki.cern.ch/twiki/bin/view/CMS/EGMSmearer.

[133] https://twiki.cern.ch/twiki/bin/viewauth/CMS/JetID.

[134] https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation80XReReco.

[135] https://twiki.cern.ch/twiki/bin/viewauth/CMS/MissingETRun2Corrections.

[136] https://twiki.cern.ch/twiki/bin/view/CMS/MissingETOptionalFiltersRun2.

[137] 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].

[138] "Measurement of the tt 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].

[139] 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].

[140] 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].

[141] 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.

[142] 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.

[143] 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].

[144] B. Clerbaux at al. "Measurement of the cross-section for tW production in dilepton final state at 13 TeV using 2016 data". In: *CMS Notes* CMS AN-2017-132 (2017).

[145] 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.

[146] https://cms-conddb.cern.ch/cmsDbBrowser/diff/Prod/gts/80X_mcRun2_asymptotic_2016_TrancheIV_v6/80X_mcRun2_asymptotic_end2016_forEGM_v0.

[147] https://twiki.cern.ch/twiki/bin/viewauth/CMS/BtagRecommendation80XReReco.

[148] *CMS Luminosity Measurements for the 2016 Data Taking Period.* Tech. rep. CMS-PAS-LUM-17-001. Geneva: CERN, 2017.

[149] Vardan Khachatryan et al. "Measurement of the $t\bar{t}$ production cross section using events in the eμ 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].

[150] 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].

[151] 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].

[152] 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].

[153] 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].

[154] 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].

[155] Nikolaos Kidonakis. "Theoretical results for electroweak-boson and single-top production". In: *PoS* DIS2015 (2015), p. 170. arXiv:1506.04072 [hep-ph].

[156] 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.

[157] 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].

[158] B. Clerbaux at al. "Search for new physics via top quark production and decay in dilepton final state at 13 TeV". In: *CMS Notes* CMS AN-2017-134 (2017).

[159] http://w3.lnf.infn.it/the-variable constant/?lang=en. "THE 鈥椄ARIABLE鈥機ONSTANT". In: (2017).

[160] Craig D. Roberts. "Strong QCD and Dyson-Schwinger Equations". In: (2012). arXiv:1203.5341 [nucl-th].

[161] T. Tabarelli De Fatis et al. "Early search for new physics in the diphoton spectrum at high mass at 13 TeV". In: *CMS Notes* CMS AN-2015-241 (2015).

[162] https://twiki.cern.ch/twiki/bin/view/CMS/MissingETRun2RecoilCorrection.