

UNIVERSITY OF CALIFORNIA

Los Angeles

NO TITLE!?!?

A dissertation submitted in partial satisfaction  
of the requirements for the degree  
Doctor of Philosophy in Physics

by

Tyler Christopher Lam

2023

© Copyright by  
Tyler Christopher Lam  
2023

The dissertation of Tyler Christopher Lam is approved.

Michalis Bachtis, Committee Chair

University of California, Los Angeles

2023

## TABLE OF CONTENTS

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>The Standard Model and Searches for Displaced Photons</b>	<b>2</b>
2.1	The Standard Model	2
2.2	Motivations for Long Lived Scalar Bosons	2
<b>3</b>	<b>Experimental Apparatus</b>	<b>3</b>
3.1	The Large Hadron Collider	3
3.1.1	The High Luminosity LHC	6
3.2	The Compact Muon Solenoid	7
3.2.1	CMS Coordinate System	8
3.2.2	Charged Track Momentum Resolution	9
3.2.3	Inner Tracker	10
3.2.4	Electromagnetic Calorimeter	12
3.2.5	Hadronic Calorimeter	14
3.2.6	Muon Detectors	16
3.2.7	Trigger System	21
3.2.8	Event Reconstruction	25
<b>4</b>	<b>Real Time Muon Reconstruction at the Compact Muon Solenoid</b>	<b>29</b>
4.1	Introduction	29
4.2	The Kalman Filter Algorithm	30
4.3	The Kalman Barrel Muon Track Finder	31

4.3.1	Muon Trajectory Propagation . . . . .	32
4.3.2	Covariant Matrices and the Kalman Gain . . . . .	36
4.3.3	Initial Curvature Assignment . . . . .	38
4.3.4	Quality Cuts on Reconstructed Tracks . . . . .	40
4.3.5	Firmware Implementation of the KBMTF . . . . .	41
4.3.6	Architecture of the KBMTF . . . . .	42
4.4	Performance of the Kalman Barrel Muon Track Finder . . . . .	44
4.4.1	Trigger Efficiencies and Rate . . . . .	44
4.4.2	Firmware and Emulator Agreement . . . . .	48
4.5	The Phase-2 KBMTF . . . . .	50
<b>5</b>	<b>Search for a Long Lived Scalar Boson . . . . .</b>	<b>52</b>
5.1	Introduction . . . . .	52
5.2	Data and Simulated Samples . . . . .	53
5.2.1	Data Samples . . . . .	53
5.2.2	Simulated Samples . . . . .	55
5.3	Vertex Reconstruction of Displaced Photon Pairs . . . . .	65
5.3.1	Diphoton Kinematics . . . . .	65
5.3.2	Performance of the Diphoton Vertex Calculation . . . . .	68
5.4	Event Selection . . . . .	71
5.4.1	Triggers . . . . .	72
5.4.2	Preselection Criteria . . . . .	73
5.4.3	Signal Region Definitions . . . . .	80
5.5	Background Estimation . . . . .	89

5.5.1	Sources of Standard Model Background . . . . .	90
5.5.2	Method of Background Estimation . . . . .	91
5.5.3	Closure Tests . . . . .	92
5.6	Statistical Analysis . . . . .	93
5.6.1	Systematic Uncertainties . . . . .	100
5.6.2	Limit Setting Procedure . . . . .	103
5.6.3	Nuisance Parameter Impacts and Pull Distributions . . . . .	108
5.7	Results . . . . .	110
5.7.1	Observed Data . . . . .	110
5.7.2	Model Dependent Upper Limits . . . . .	110
5.7.3	Model Independent Upper Limits . . . . .	118
5.8	Summary . . . . .	118
<b>References</b>	. . . . .	<b>121</b>

## LIST OF FIGURES

3.1 LHC luminosity report. Left: breakdown of the CMS integrated luminosity by year from 2010-2022. Right: peak luminosity from 2016-2018 data taking [1]. . . . .	5
3.2 Diagram of the CERN accelerator complex during 2018 data taking. Protons begin as hydrogen atoms at Linac2 and collide at various detectors along the LHC at $\sqrt{s} = 13$ TeV . . . . .	5
3.3 Timeline of runs and shutdowns leading to the HL-LHC, showing the improvements to $\sqrt{s}$ and $\mathcal{L}$ [2]. The extended year-end technical stops (EYETS) are shorter shutdowns generally for routine maintenance, although upgrades such as the CMS Pixel Detector upgrade can occur during this time as well. . . . .	6
3.4 Cutaway diagram of CMS showcasing major detector components [3]. . . . .	7
3.5 Diagram showing the sagitta for a charged particle track . . . . .	9
3.6 $r$ - $z$ slice of the CMS pixel detector [4] . . . . .	12
3.7 $r$ - $z$ slice of the CMS silicon detector, detailing the four main components of the strip detector [5] . . . . .	13
3.8 Geometry of the CMS ECAL showing the barrel, endcap, and preshower, adapted from [6] . . . . .	15
3.9 $r$ - $z$ slice of the CMS HCAL showing the HB, HE, HF, and HO [7] . . . . .	17
3.10 Diagram of the CMS Muon system [8] . . . . .	18
3.11 Diagram of the CMS Drift Tubes [9] . . . . .	19
3.12 Cut-away diagram of a single CSC module [10] . . . . .	20
3.13 Diagram of the double RPC at CMS [11] . . . . .	21
3.14 CMS Level-1 Trigger architecture during Run 2 [12] . . . . .	23

3.15	Designed trigger architecture of the HL-LHC. The addition of the track primitives allows the computation of more advanced reconstruction algorithms at the L1 level [13]. . . . .	24
4.1	The iterative process of propagation and updating a muon track through the KBMTF algorithm. The track properties at the innermost station are stored in order to trigger on muons not originating from the beam axis [14] . . . . .	33
4.2	Diagram showing a muon trajectory between two stations with the origin set at the outer station. The x-axis is oriented using the detector vertex and the outer station. . . . .	35
4.3	Calculation of phi propagation coefficients. Units are converted to their digitized form used in firmware calculations. . . . .	36
4.4	$\sigma$ of vertical slice fits of the 2-dimensional histogram in figure 4.3a used to calculate uncertainty coefficients. . . . .	37
4.5	Coordinate system for muon propagation beginning at the vertex and going outward.	39
4.6	Diagram of the CMS muon barrel system [15]. (Left) Each wheel is segmented in 12 sectors in $\phi$ and 4 DT and 3 RPC stations in $r$ . (Right) The 5 wheels in $z$ . .	43
4.7	Diagram showing the inputs of one MP7 board consisting, consisting of the Twin-Mux inputs from one central wedge (green) and the two adjacent wedges (blue). The FPGA reconstructs all tracks with the outermost stub in the central wedge.	43
4.8	Comparison of efficiencies between the KBMTF and phase-1 BMTF algorithm using a Z tag-and-probe sample collected during 2017 data taking. (a) Efficiency versus track $p_T$ , requiring L1 muon $p_T > 20$ GeV. (b) Efficiency versus track $\eta$ , requiring L1 muon $p_T > 20$ GeV and track $p_T > 25$ GeV in order to show the efficiency on the plateau. The drops in efficiency at $\eta = \pm 0.2$ are due to gaps between wheels in the muon system. . . . .	46

4.9	Comparison of efficiency vs cosmic track $d_{xy}$ using cosmic data from the 2018 data taking run. The prompt KBMTF takes the reconstructed $p_T$ using the vertex constrained measurement, while the displaced KBMTF uses the vertex unconstrained measurement. The displaced KBMTF shows substantial improvements to the previous BMTF [16] . . . . .	47
4.10	Rate comparison calculated using 2017 zero bias data of the Phase I BMTF, vertex constrained (prompt) KBMTF, vertex unconstrained (displaced) KBMTF, and displaced KBMTF for displaced tracks. . . . .	48
4.11	Comparison of the track $p_T$ between the firmware and emulator using data from a 2017 Z tag-and-probe sample. hwPt represents the integer value of $p_T$ used in hardware. The emulator and firmware output show 100% agreement across 500000 events. . . . .	49
4.12	Performance of the L1 Barrel Muon algorithms using simulated samples with HL-LHC run conditions. The Track + KBMTF algorithm matches tracks to loose KBMTF muons, while the L1 Track + Muon Stubs algorithm matches tracks to muon stubs. . . . .	51
5.1	Feynman diagrams for both quark induced and gluon induced Z+H production. . . . .	56
5.2	Feynman diagram showing the proposed decay channel of the Higgs boson to $\Phi\Phi$ , with one $\Phi$ decaying to a pair of photons and the other decaying to $d\bar{d}$ . The loop in the diphoton decay can contain any particle that could theoretically couple to the $\Phi$ . It is also possible for the $\Phi$ to decay directly to photons. . . . .	56
5.3	Pileup distributions and mean PU from 2011-2023 datataking [1]. . . . .	58
5.4	Trigger efficiency as a function of $\eta_{sc}$ for several ranges of electron $p_T$ for all data taking eras. . . . .	62
5.5	2018 electron HLT scale factors and uncertainties. . . . .	63

5.6	2017 electron HLT scale factors and uncertainties. . . . .	63
5.7	2016 post-VFP electron HLT scale factors and uncertainties. . . . .	64
5.8	2016 pre-VFP electron HLT scale factors and uncertainties. . . . .	64
5.9	Geometrical configuration used to reconstruct the displaced vertex of the di-photon pair . . . . .	66
5.10	Geometrical configuration showing the reconstruction of a negative impact parameter . . . . .	68
5.11	Performance of the vertex calculation on signal samples. The fit shows good agreement across all values of generator level $L_{xy}$ and a resolution of $< 3\text{ cm}$ . . . . .	69
5.12	The difference between the true and reconstructed vertex grows as the assumed mass departs from the true one. Gen-reconstructed $L_{xy}$ for a Z+H samples under several assumed mass points. . . . .	70
5.13	The $\rho$ for barrel photons and $ z $ for endcap photons for all reconstructed photons using <code>miniAODv2</code> signal samples . . . . .	71
5.14	Efficiency for a photon pair to pass preselection versus generator level $L_{xy}$ on ZH signal samples for all lifetimes in the $Z \rightarrow \mu\mu$ category. . . . .	80
5.15	$m(\ell^+\ell^-)$ data and simulated background after preselection criteria across all years in the anti-ID control region. . . . .	82
5.16	$m(\ell^+\ell^-)$ data and simulated background after preselection criteria across all years in the ID signal region. . . . .	83
5.17	Comparison of invariant mass distributions for FSR rejection using simulated $m_\Phi = 20\text{ GeV}$ , $c\tau = 0\text{ mm}$ signal and background. Background uses normalized anti-ID control region events for improved statistics while signal shows ID signal region events. . . . .	85

5.18 $m(\ell^+\ell^-)$ data and simulated background after rejecting FSR events across all years in the anti-ID control region. . . . .	86
5.19 $m(\ell^+\ell^-)$ data and simulated background after rejecting FSR events across all years in the ID signal region. . . . .	87
5.20 Comparison of leading and subleading photon $p_T$ in the anti-ID control region and signal region for the $Z \rightarrow \mu\mu$ category under the assumptions $m_\Phi = 20$ GeV (top) and $m_\Phi = 50$ GeV (bottom). . . . .	88
5.21 Final $m(\ell^+\ell^-)$ distributions for data and simulated samples in the anti-ID control region. Top row: $Z \rightarrow \mu\mu$ category. Bottom row: $Z \rightarrow ee$ category. . . . .	89
5.22 Final $m(\ell^+\ell^-)$ distributions for data and simulated samples in the ID signal region. Top row: $Z \rightarrow \mu\mu$ category. Bottom row: $Z \rightarrow ee$ category. . . . .	90
5.23 Cutflow efficiency for signal events with exactly 2 photons (top) and 4 photons (bottom). . . . .	91
5.24 Feynman diagram showing the possible background contributions from Drell-Yan events. Photons can originate from neutral pions produced in jets, initial state radiation from quarks, or final state radiation from leptons. . . . .	92
5.25 Shape comparison of ID signal region and anti-ID control region data in the negative $L_{xy}$ sideband for $m_\Phi = 15$ GeV in the $Z \rightarrow \mu\mu$ category (left) and $Z \rightarrow ee$ category (bottom) across all three years. . . . .	94
5.26 Shape comparison of ID signal region and anti-ID control region data in the negative $L_{xy}$ sideband for $m_\Phi = 20$ GeV in the $Z \rightarrow \mu\mu$ category (left) and $Z \rightarrow ee$ category (bottom) across all three years. . . . .	95
5.27 Shape comparison of ID signal region and anti-ID control region data in the negative $L_{xy}$ sideband for $m_\Phi = 30$ GeV in the $Z \rightarrow \mu\mu$ category (left) and $Z \rightarrow ee$ category (bottom) across all three years. . . . .	96

5.28 Shape comparison of ID signal region and anti-ID control region data in the negative $L_{xy}$ sideband for $m_\Phi = 40\text{ GeV}$ in the $Z \rightarrow \mu\mu$ category (left) and $Z \rightarrow ee$ category (bottom) across all three years. . . . .	97
5.29 Shape comparison of ID signal region and anti-ID control region data in the negative $L_{xy}$ sideband for $m_\Phi = 50\text{ GeV}$ in the $Z \rightarrow \mu\mu$ category (left) and $Z \rightarrow ee$ category (bottom) across all three years. . . . .	98
5.30 Shape comparison of ID signal region and anti-ID control region data in the negative $L_{xy}$ sideband for $m_\Phi = 55\text{ GeV}$ in the $Z \rightarrow \mu\mu$ category (left) and $Z \rightarrow ee$ category (bottom) across all three years. . . . .	99
5.31 Dependence of the normalization of limits on $\mathcal{BR}(\Phi \rightarrow \gamma\gamma)$ for the $Z \rightarrow \mu\mu$ categories across all masses and values of $c\tau$ . . . . .	105
5.32 Dependence of the normalization of limits on $\mathcal{BR}(\Phi \rightarrow \gamma\gamma)$ for the $Z \rightarrow ee$ categories across all masses and values of $c\tau$ . . . . .	106
5.33 Post-fit pulls and impacts of the 30 highest impact nuisance parameters for a $m_\Phi = 30\text{ GeV}$ , $c\tau = 100\text{ mm}$ signal hypothesis. Leftmost column shows the nuisance parameter names. Parameters starting with <code>CMS_bkg*</code> represent nuisance parameters for background normalization and non-closure uncertainty; all others correspond to systematic uncertainties of simulated signal. Middle column shows post-fit pull distributions, which lie within $1\sigma$ of the pre-fit uncertainty. Right column shows the impacts on the signal strength parameter $\mu = 20 \times \mathcal{BR}(H \rightarrow \Phi\Phi) \times \mathcal{BR}(\Phi \rightarrow \gamma\gamma)$ . . . . .	109
5.34 Observed number of $m_\Phi = 15\text{ GeV}$ events in the signal region of the $Z \rightarrow \mu\mu$ category (left) and the $Z \rightarrow ee$ category (right) across all three years. . . . .	111
5.35 Observed number of $m_\Phi = 20\text{ GeV}$ events in the signal region of the $Z \rightarrow \mu\mu$ category (left) and the $Z \rightarrow ee$ category (right) across all three years. . . . .	112

5.36 Observed number of $m_\Phi = 30$ GeV events in the signal region of the $Z \rightarrow \mu\mu$ category (left) and the $Z \rightarrow ee$ category (right) across all three years. . . . .	113
5.37 Observed number of $m_\Phi = 40$ GeV events in the signal region of the $Z \rightarrow \mu\mu$ category (left) and the $Z \rightarrow ee$ category (right) across all three years. . . . .	114
5.38 Observed number of $m_\Phi = 50$ GeV events in the signal region of the $Z \rightarrow \mu\mu$ category (left) and the $Z \rightarrow ee$ category (right) across all three years. . . . .	115
5.39 Observed number of $m_\Phi = 55$ GeV events in the signal region of the $Z \rightarrow \mu\mu$ category (left) and the $Z \rightarrow ee$ category (right) across all three years. . . . .	116
5.40 Model dependent upper limits for $\mathcal{BR}(H \rightarrow \Phi\Phi) \times \mathcal{BR}(\Phi \rightarrow \gamma\gamma)$ for different values of $m_\Phi$ and $c\tau$ under the assumption that $\mathcal{BR}(\Phi \rightarrow \gamma\gamma) = 0.5$ . . . . .	117
5.41 Observed limit of $\mathcal{BR}(H \rightarrow \Phi\Phi) \times \mathcal{BR}(\Phi \rightarrow \gamma\gamma)$ as a function of $m_\Phi$ and $c\tau$ under the assumption that $\mathcal{BR}(\Phi \rightarrow \gamma\gamma) = 0.5$ . . . . .	118
5.42 Model independent upper limits for $\sigma(Z + \Phi\Phi) \times \mathcal{BR}(\Phi \rightarrow \gamma\gamma)$ as a function of $c\tau$ for various values of $m_\Phi$ under the assumption that $\mathcal{BR}(\Phi \rightarrow \gamma\gamma) = 0.5$ . . . .	119
5.43 Observed model independent limits on $\sigma(Z + \Phi\Phi) \times \mathcal{BR}(\Phi \rightarrow \gamma\gamma)$ as a function of $m_\Phi$ and $c\tau$ under the assumption that $\mathcal{BR}(\Phi \rightarrow \gamma\gamma) = 0.5$ . . . . .	120

## LIST OF TABLES

4.1 Comparison of FPGA resource utilization [14]. The KBMTF utilizes DSP cores for arithmetic operations to propagate and update tracks and LUTs for the propagation coefficients and approximate Kalman Gain, while the BMTF uses LUTs to assign momentum. . . . .	42
5.1 CMS data sets used for each year in this analysis. Asterisks represents wildcards that cover the multiple data taking eras per year. . . . .	54
5.2 Table of integrated luminosity per year [17–19]. . . . .	54
5.3 Mass points, lifetimes, and number of simulated $H \rightarrow \Phi\Phi$ events for the $2\gamma 2q$ categories. Number of events generated is halved for the $4\gamma$ categories . . . . .	57
5.4 List of simulated background processes, their cross sections calculated using <code>GenXSecAnalyzer</code> [20], and CMS data set names. . . . .	59
5.5 Tag and probe criteria used to derive HLT scale factors . . . . .	61
5.6 HLT trigger paths used for the Run 2 datasets . . . . .	72
5.7 Cut based ID selection criteria for electrons in Run-2 [21]. . . . .	76
5.8 Cut based ID selection criteria for photons in Run-2 [22]. . . . .	79
5.9 Photon $p_T$ cut efficiency on control region background . . . . .	86
5.10 Cutflow table and expected yields on background Monte Carlo in the ID+ID signal region . . . . .	88
5.11 Description of multiplicative corrections for systematic uncertainties as a function of the nuisance parameter $\nu$ [23]. . . . .	100
5.12 Summary of the systematic uncertainties and corresponding correlations . . . .	103
5.13 Higgs production cross sections used to normalize simulated signal samples [24].	108

# **CHAPTER 1**

## **Introduction**

## **CHAPTER 2**

# **The Standard Model and Searches for Displaced Photons**

### **2.1 The Standard Model**

### **2.2 Motivations for Long Lived Scalar Bosons**

# CHAPTER 3

## Experimental Apparatus

### 3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) is a circular collider spanning the border between France and Switzerland, based at the European Organization for Nuclear Research (CERN). The central features of the LHC are the superconducting rings, located about 100 m underground with a circumference of 27 km, designed to collide counter-rotating beams of protons or heavy ions at highly relativistic energies. Along the rings lie four major experiments: ATLAS, CMS, ALICE, and LHCb. Both ATLAS and CMS are general purpose detectors, designed to probe a wide range of physics including the Higgs boson, precision measurements of fundamental constants, and physics beyond the standard model (BSM). The remaining two experiments are more specialized; ALICE measures quark gluon plasma produced in heavy ion collisions and LHCb focuses on  $b$ -quark physics and CP violation.

Two prominent aspects of the LHC are the high center of mass energy of the proton beams, referred to using the Mandelstam variable  $\sqrt{s}$ , and high instantaneous luminosity  $\mathcal{L}$ , often referred to as just luminosity. High  $\sqrt{s}$  allows for the production of more massive particles, giving more access to possible BSM physics, while high luminosity is essential for measuring rare processes and precision measurements. A process with cross section  $\sigma$  will have a rate  $R$  given by

$$R = \mathcal{L}\sigma \tag{3.1}$$

The cross section  $\sigma$  is a measure of how probable a process is to occur, and is measured in

units of area. A frequent used unit for cross sections is a barn (b), corresponding to  $100 \text{ fm}^2$ . Conversely, luminosity uses units of  $\text{Hz}/\text{b}$ . In cases where the relevant quantity is the total number of events, the integrated luminosity can be defined as  $\mathcal{L}_{\text{int}} = \int \mathcal{L} dt$  to give

$$N_{\text{events}} = \mathcal{L}_{\text{int}} \sigma \quad (3.2)$$

The luminosity depends on the characteristics of the proton beam and can be written in terms of the operational parameters of the detector given by

$$\mathcal{L} = \frac{N_b^2 n_b f_{\text{rev}} \gamma_r}{4\pi \epsilon_n \beta^*} F \quad (3.3)$$

where  $N_b$  is the number of particles per bunch,  $n_b$  is the number of bunches per ring,  $f_{\text{rev}}$  is the LHC revolution frequency,  $\gamma_r$  is the Lorentz factor for the proton,  $\epsilon_n$  is the transverse normalized beam emittance,  $\beta^*$  is the amplitude function at the collision point, and  $F$  is a geometric reduction factor based on the crossing angle of the two beams. The nominal design parameters of the LHC were intended to support a peak luminosity of  $12 \text{ Hz/nb}$  [25], but were exceeded by nearly twice that value of  $20.7 \text{ Hz/nb}$  during 2017 data taking and again with  $21.4 \text{ Hz/nb}$  in 2018. The LHC produced an integrated luminosity of  $41.6 \text{ fb}^{-1}$  in 2016,  $49.8 \text{ fb}^{-1}$  in 2017, and  $67.9 \text{ fb}^{-1}$  in 2018 for a total of  $164 \text{ fb}^{-1}$  during run 2 data taking. A more detailed breakdown of the LHC luminosity records can be seen in Fig. 3.1.

The protons used in collisions are extracted from hydrogen atoms by using ionizing electric fields to strip them of their electrons. Protons are first accelerated to an energy of  $50 \text{ MeV}$  through a linear accelerator Linac2 before entering the Proton Synchrotron Booster (PSB), where they reach a kinetic energy of  $1.4 \text{ GeV}$ . Next, the protons are accelerated by the Proton Synchrotron (PS) and Super Proton Synchrotron (SPS), where they are accelerated to  $26 \text{ GeV}$  and  $450 \text{ GeV}$  respectively. Finally, the beams are injected into the LHC where they undergo acceleration to  $6.5 \text{ TeV}$ , producing the desired center of mass energy of  $\sqrt{s} = 13 \text{ TeV}$ .

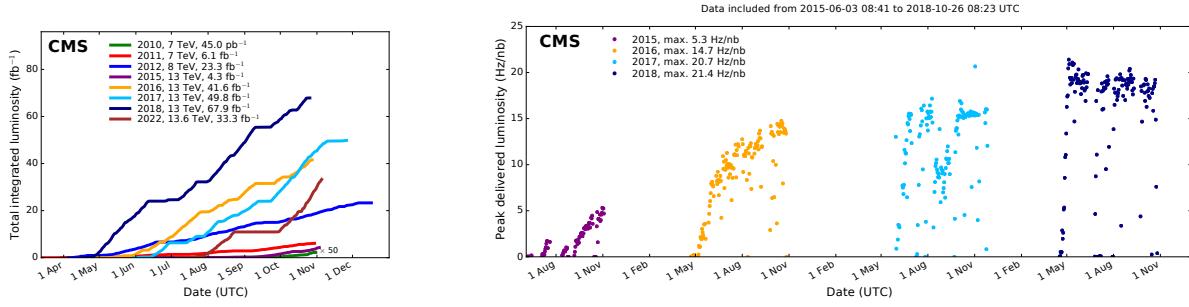


Figure 3.1: LHC luminosity report. Left: breakdown of the CMS integrated luminosity by year from 2010-2022. Right: peak instantaneous luminosity from 2016-2018 data taking [1].

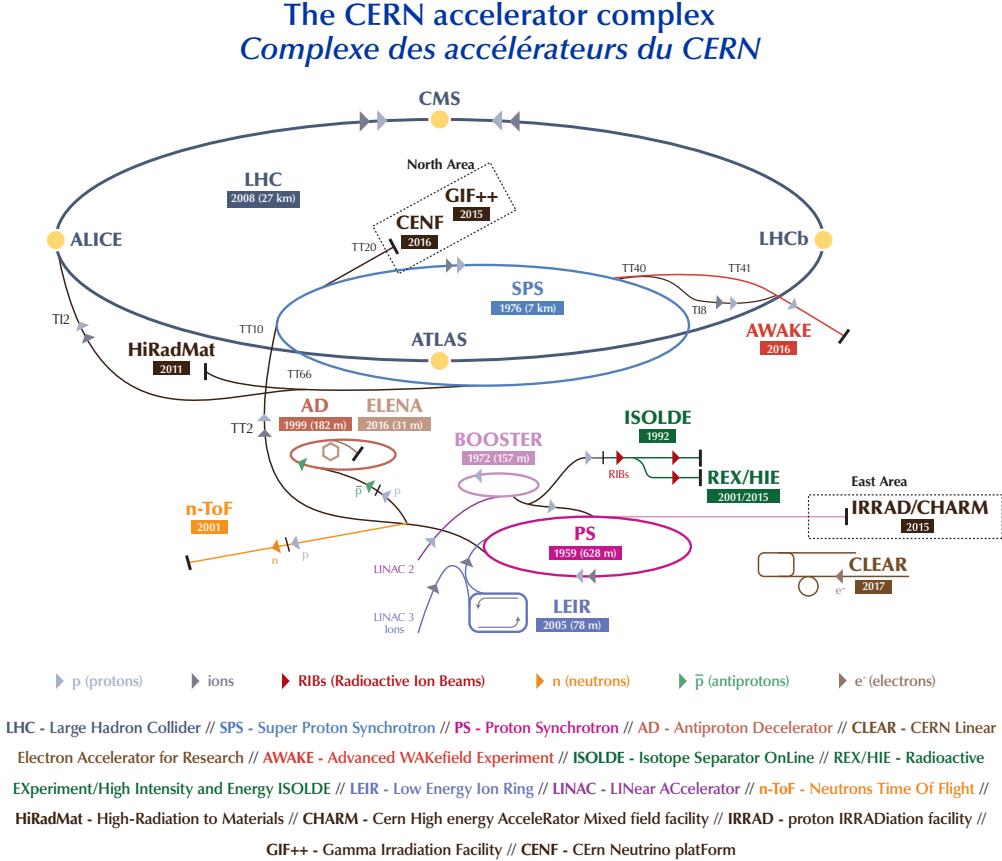


Figure 3.2: Diagram of the CERN accelerator complex during 2018 data taking [26]. Protons begin as hydrogen atoms at Linac2 and are accelerated in several stages to reach 6.5 TeV.

### 3.1.1 The High Luminosity LHC

The LHC is designed to go through periods of near continuous data taking, followed by long shut downs during which the detectors undergo various upgrades. As of the time of this writing, the LHC is currently in its third data taking run (Run 3), which is planned to operate from 2022-2025. The end of Run 3 marks the end of Phase-1 and the start of the third long shutdown (LS3), during which the LHC will receive major upgrades to increase the peak instantaneous luminosity up to 75 Hz/nb [14], over six times the designed peak luminosity and three times the peak luminosity achieved as of Run 2. The upgraded high luminosity LHC (HL-LHC) will gather substantially more data, which better allows experiments to probe rare processes and measure precise values predicted from the standard model.

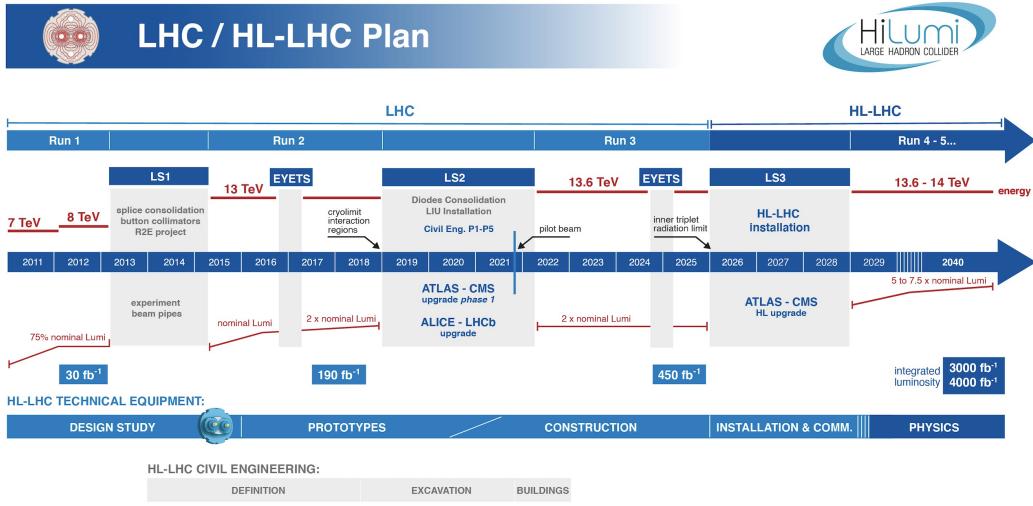


Figure 3.3: Timeline of runs and shutdowns leading to the HL-LHC, showing the improvements to  $\sqrt{s}$  and  $\mathcal{L}$  [2]. The extended year-end technical stops (EYETS) are shorter shutdowns generally for routine maintenance, although upgrades such as the CMS Pixel Detector upgrade can occur during this time as well.

## 3.2 The Compact Muon Solenoid

The Compact Muon Solenoid (CMS) is one of the two general purpose detectors at the LHC, designed to reconstruct physics events from proton-proton collisions. It is a cylindrical apparatus with its major axis aligned with the proton beams from the LHC, consisting of several concentric layers of specialized detectors. The innermost layer consists of a silicon tracker, followed by an electromagnetic and hadronic calorimeter (ECAL and HCAL, respectively). Surrounding those is the superconducting solenoid, which provides a 3.8 T magnetic field to the trajectory of charged particles. Lastly, layers of muon chambers are interspersed with an iron return yolk, designed to contain the magnetic field lines near the muon chambers.

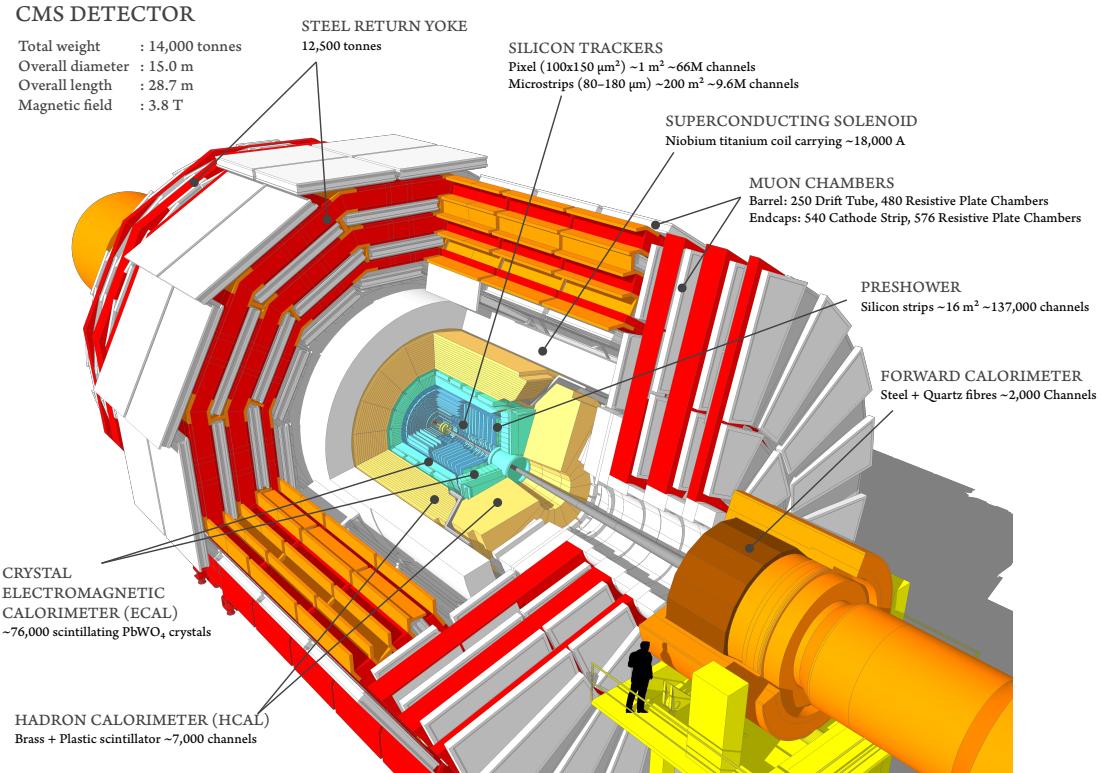


Figure 3.4: Cutaway diagram of CMS showcasing major detector components [3].

### 3.2.1 CMS Coordinate System

The origin of the CMS coordinate system is located at the center of the detector, at the nominal collision point of the proton beams. The  $+\hat{x}$  axis points radially towards the center of the LHC, while the  $+\hat{y}$  axis points vertically upward. This sets the  $+\hat{z}$  direction along the beamline, counterclockwise along the LHC. Due to the cylindrical symmetry of CMS, coordinates in the transverse  $\hat{x}\text{-}\hat{y}$  plane are commonly replaced with the radius  $r$  and the azimuthal angle  $\phi$ , which is measured from the positive  $\hat{x}$ -axis. The polar angle  $\theta$  is measured from the  $+\hat{z}$ -axis, though in practice this variable is rarely used. It is more common to define the polar angle in terms of the pseudorapidity  $\eta$ , which is defined as

$$\eta = \frac{1}{2} \ln \left( \frac{|\mathbf{p}| + p_z}{|\mathbf{p}| - p_z} \right) = -\ln \tan \left( \frac{\theta}{2} \right) \quad (3.4)$$

The motivation for using  $\eta$  instead of  $\theta$  stems from a fundamental property of hadron colliders: the center of mass frame for particle production rarely coincides with the lab frame. Thus, when measuring the separation between particles, it is useful to define quantities that remain invariant under Lorentz transformations in the  $\hat{z}$  direction. The rapidity  $y$  (not to be confused with the Cartesian coordinate  $\hat{y}$ ) is defined as

$$y = \frac{1}{2} \ln \left( \frac{E + p_z}{E - p_z} \right) \quad (3.5)$$

It is trivial to show that differences in  $y$  remain invariant under Lorentz boosts along the  $\hat{z}$ -axis. In the highly relativistic limit, which is valid for most particles produced at the LHC, rapidity approaches pseudorapidity as  $E \approx |\mathbf{p}|$ . The angular separation between two particles can then be expressed as an invariant quantity under Lorentz boosts as  $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ . One advantage of pseudorapidity is that  $\eta$  can be calculated using only geometric quantities of the detector, whereas  $y$  requires calculating both the energy and momenta of a particle, making pseudorapidity the natural choice for defining the polar

angle.  $\eta$  can range from  $(-\infty, \infty)$ , where  $\eta = \pm\infty$  points directly along the  $\pm\hat{z}$  axis. Higher values of  $|\eta|$  are commonly referred to as "forward".

### 3.2.2 Charged Track Momentum Resolution

The 3.8 T magnet curves the trajectory of charged particles, which allows us to precisely determine their momentum. Using the geometry described in section 3.2.1, a particle with charge  $q$  and transverse momentum  $p_T$  travels in a circular trajectory with radius  $R$  when viewed in the  $r\text{-}\phi$  plane. Per the Lorentz Force Law, these can be related by  $p_T = qBR$ . In particle physics, when working with an object of elementary charge, this is commonly rewritten as

$$p_T = 0.3BR \text{ [GeV/c]} \quad (3.6)$$

The track followed by a charged particle traveling over a length  $L$  can be described by the sagitta  $s$  shown in figure 3.5, defined by

$$s = R - \sqrt{R^2 - \frac{L^2}{4}} \quad (3.7)$$

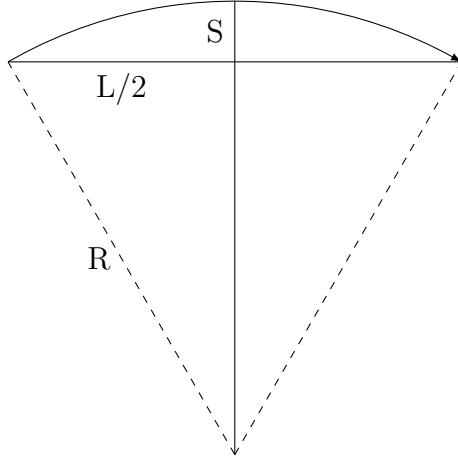


Figure 3.5: Diagram showing the sagitta for a charged particle track

For highly energetic particles, the length  $L$  is substantially smaller than the radius  $R$ , so

the sagitta can be approximated as

$$s \approx \frac{L^2}{8R} = \frac{0.3BL^2}{8p_T} \quad (3.8)$$

The relative uncertainty of the momentum is proportional to the uncertainty of the sagitta, the track length L, and the magnetic field B as

$$\frac{\delta p_T}{p_T} \propto \frac{p_T}{BL^2} \delta_s \quad (3.9)$$

The uncertainty on the sagitta  $\delta s$  is analogous to the hit resolution of a given detector. Therefore, a long lever arm and high magnetic field strength are crucial to precisely determine the momentum of energetic particles.

### 3.2.3 Inner Tracker

The CMS inner tracker is the first detector surrounding the primary interaction point (IP). Its purpose is to measure the tracks from charged particles as they curve in the strong magnetic field within CMS in order to calculate their momentum and reconstruct secondary vertices. As the closest detector to the primary IP, the tracker experiences the highest particle flux within CMS, and must have high enough granularity to distinguish the multitude of tracks. This high granularity also lends the inner tracker the best momentum resolution for charged tracks out of all the subdetectors comprising CMS, with a momentum resolution of 2.8% for a 100 GeV muon [27]. The high flux also makes the detector susceptible to radiation damage, so the tracker must be robust in order to maintain efficiency over the long operational period of the LHC.

The inner tracker utilizes silicon tracking modules composed of P-N type junctions to detect the location of charged particles. When an external voltage is applied to a module, particles passing through will deposit energy and create electron-hole pairs, which drift to

their respective electrodes and generate an electrical signal. The applied voltage is tuned such that only a small amount of energy is required to produce electron-hole pairs, which reduces the total amount of material required in order to minimize the energy loss of the charged particles. The inner layer of the tracker is composed of higher granularity pixel detectors, while the outer layer is composed of coarser silicon strips.

### 3.2.3.1 Silicon Pixel Detector

The pixel detector is comprised of 124 million silicon pixel sensors, each with an area of  $100 \times 150 \mu\text{m}^2$  and a thickness of  $300 \mu\text{m}$ . The barrel (BPIX) consists of four cylindrical layers of pixel sensors spanning from  $z = -54 \text{ cm}$  to  $z = 54 \text{ cm}$  at radii of  $r = 2.9, 6.8, 10.9,$  and  $16 \text{ cm}$ . The endcap (FPIX) consists of three layers located at  $z = \pm 29.1, \pm 39.6,$  and  $\pm 51.6 \text{ cm}$ , which when combined with the barrel nets a total sensitive area of  $1.85 \text{ m}^2$  with coverage up to  $|\eta| < 2.5$ . With regards to detector performance, the pixel detector provides spatial resolution of  $9.5 \mu\text{m}$  in the  $r\text{-}\phi$  direction and  $22.2 \mu\text{m}$  in the  $z$  direction, as well as a hit efficiency of  $> 99\%$  in each layer at nominal LHC luminosity, with the innermost BPIX layer dropping to  $97.5\%$  at peak run 2 efficiency [4].

### 3.2.3.2 Silicon Strip Detector

The silicon strip detector surrounds the pixel detector, and is segmented into a tracker inner barrel (TIB), tracker outer barrel (TOB), tracker inner disks (TID), and tracker endcap (TEC). The TIB consists of four layers covering  $|z| < 65 \text{ cm}$  and radius  $25.5 \text{ cm} < r < 49.8 \text{ cm}$ . The endcaps of the TIB are covered by the TID, consisting of three disks covering  $90 < |z| < 90 \text{ cm}$ . Surrounding the TIB is the TOB, with six layers spanning  $|z| < 188 \text{ cm}$  and  $60.8 < |r| < 108 \text{ cm}$ . Lastly, both the TOB and TID are closed by the TEC, which covers radii ranging from  $22 < r < 113.5 \text{ cm}$  and stretches from  $124 < |z| < 280 \text{ cm}$ .

The strip detector modules function similarly to the pixel modules, but utilize much

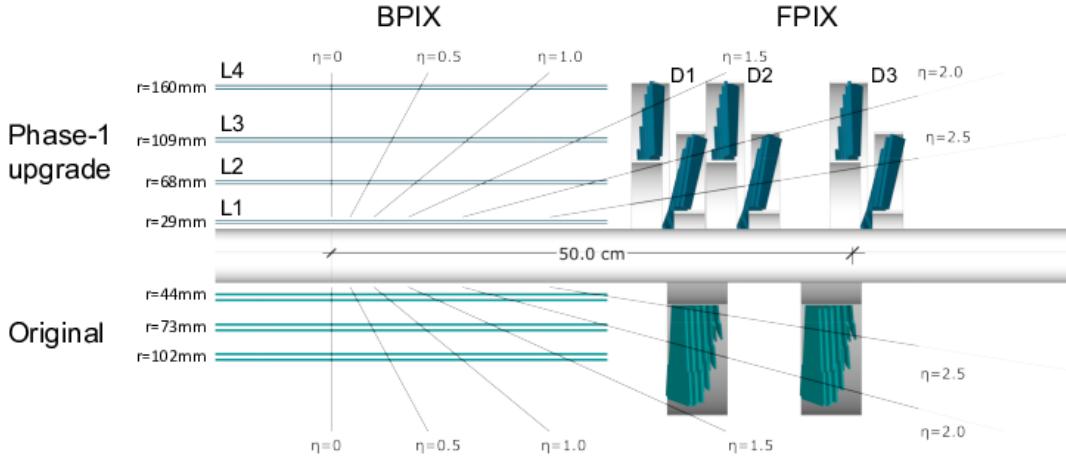


Figure 3.6:  $r$ - $z$  slice of the CMS pixel detector comparing original design (bottom, teal) with the Phase-I upgrade implemented in 2016/17, which added one layer to both the FPIX and BPIX (top, blue) [4].

larger silicon strips due to the expected reduced particle flux compared to the pixel detector. Each strip detector is composed of  $300\ \mu\text{m}$  thick micro-strip sensors with pitch varying from  $80 - 205\ \mu\text{m}$  to form  $7 - 12.5\ \text{cm}$  long strip modules. Overall, the strip detector uses 15,148 modules covering an area of  $200\ \text{m}^2$  up to  $|\eta| < 2.5$ .

### 3.2.4 Electromagnetic Calorimeter

The CMS Electromagnetic Calorimeter (ECAL) is designed to stop and capture the energy of photons and electrons. When a high energy photon or electron enters the ECAL, it primarily interacts with the material in the detector through  $e^+e^-$  pair production or bremsstrahlung, respectively. These processes create more photons and electrons that result in an electromagnetic cascade until the particles lose enough energy that ionization and compton scattering processes begin to dominate. This process is known as an electromagnetic shower. The light from these showers is measured using photodetectors to determine the initial energy of the incident particle with resolution given by

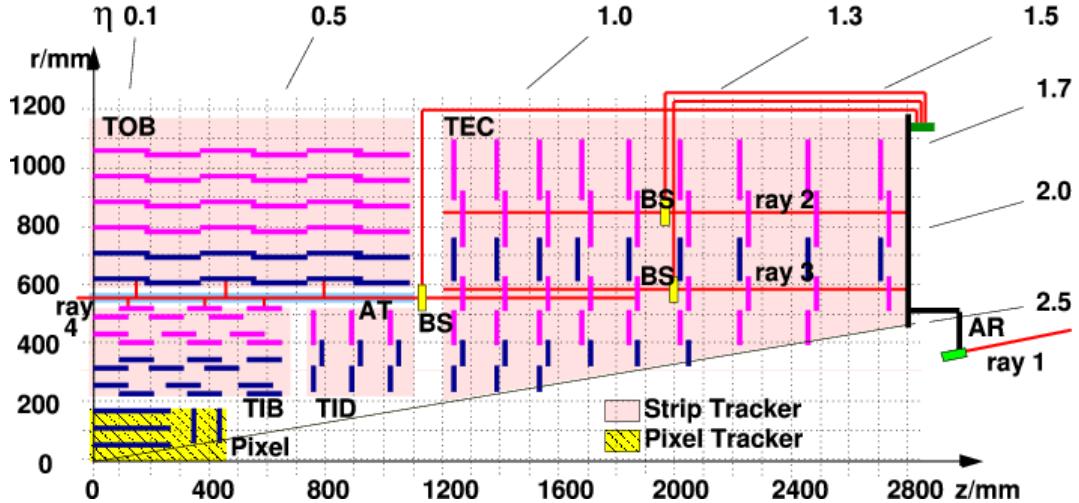


Figure 3.7:  $r$ - $z$  slice of the CMS silicon detector, detailing the four main components (TIB, TOB, TID, and TEC) of the strip detector. Pixel detector is shown in its Run-I configuration [5].

$$\left(\frac{\sigma_E}{E}\right)^2 = \left(\frac{2.8\%}{\sqrt{E/\text{GeV}}}\right)^2 + \left(\frac{12\%}{E/\text{GeV}}\right)^2 + (0.3\%)^2 \quad (3.10)$$

The first term arises from statistical fluctuations in the electromagnetic shower and light production, which is a poisson process and therefore proportional to  $1/\sqrt{E}$ . The second term is due to the noise in the detector electronics. The last constant term is due to imperfections in the detector such as non-uniformity, radiation damage, and calibration uncertainty. The material used for an electromagnetic calorimeter can be characterized by the radiation length ( $X_0$ ) and molière radius ( $R_M$ ). The radiation length is the average distance an electron will travel before its energy is reduced by a factor of  $1/e$ , and is roughly proportional to  $A/Z^2$ , where  $A$  is the atomic mass and  $Z$  is the atomic number. The molière radius is directly proportional to the radiation length and measures the spread of an electromagnetic shower in the transverse direction. Both are generally measured in units of  $[\text{g cm}^{-2}]$ , but can be divided by the density of the material to obtain a distance in [cm]. A low radiation length and Molière radius ensure all of the energy is captured in the detector and is contained to a

small area to precisely determine the position of an incident particle. It is for this reason that the ECAL is composed of lead-tungstate ( $\text{PbWO}_4$ ): a dense, high Z scintillator crystal with  $X_0 = 7.39 \text{ g/cm}^2$  (or  $0.89 \text{ cm}$  after dividing by the density of  $\text{PbWO}_4$ ) and  $R_M = 2.2 \text{ cm}$  [28].

The ECAL is constructed from  $61200 + 15000$  lead-tungstate crystals in the barrel + endcap region. Barrel crystals have a cross sectional area of  $2.2 \times 2.2 \text{ cm}^2$  to match the molière radius and a depth of  $23 \text{ cm}$  (or  $25.8 X_0$ ), and are arranged in rings along the  $\phi$  direction, tilted to align with lines of constant  $\eta$ . The barrel has an inner radius of  $129 \text{ cm}$  and covers an eta region up to  $|\eta| < 1.479$ . Endcap crystals have a slightly higher cross sectional area of  $2.6 \times 2.6 \text{ cm}^2$  and a slightly lower depth of  $22 \text{ cm}$ , and are arranged in an  $x$ - $y$  grid. The endcap provides the remaining coverage from  $1.479 < |\eta| < 3.0$ . A diagram showing the geometry of the ECAL can be seen in figure 3.8. The size and alignment of the crystals is designed to contain showers from incident photons/electrons within a  $2 \times 2$  grid of crystals.

One common process in CMS is the production of  $\pi^0$ s, which decay to two photons. When the two photons deposit their energy into the ECAL, they can fake a signal from a single high energy photon. In order to identify these background events, a higher granularity detector known as a preshower is placed before the ECAL endcap. It consists of a layer of lead absorber to initiate the shower, followed by silicon strip sensors to measure the tracks within the shower. The lead has a thickness of  $1.57 \text{ cm}$  (or  $2.8 X_0$ ), which is thick enough to reliably cause a shower while only causing energy loss of a few percent before the shower can reach the ECAL. The silicon strips are similar to those in the inner tracker, with a pitch of  $1.9 \text{ mm}$  [29].

### 3.2.5 Hadronic Calorimeter

The CMS Hadronic Calorimeter (HCAL) is a sampling calorimeter surrounding the ECAL. Unlike the ECAL, which is a homogeneous calorimeter where the crystals both induce showers and scintillate, the HCAL is composed of alternating layers of dense absorber and scintilla-

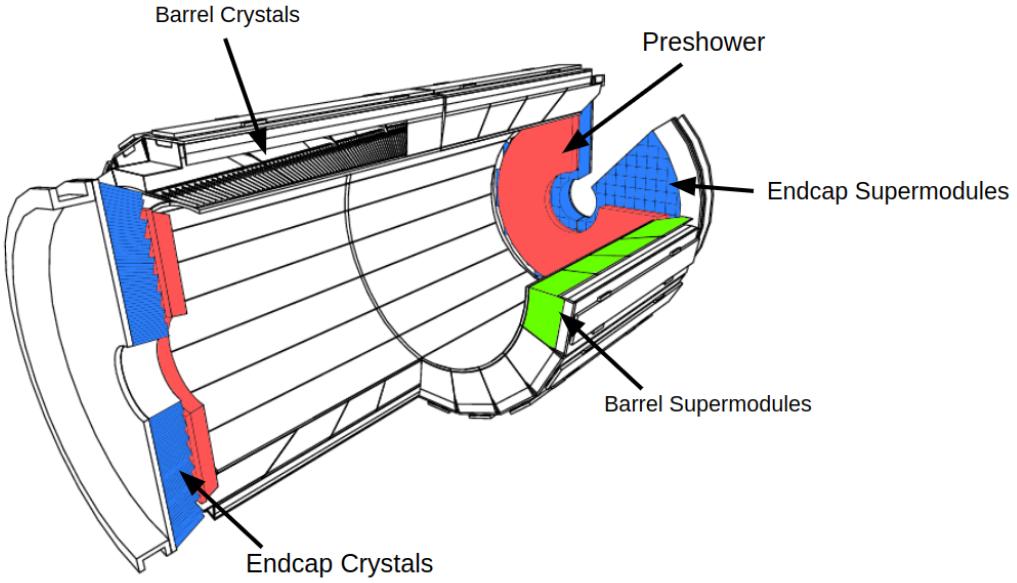


Figure 3.8: Geometry of the CMS ECAL showing the barrel, endcap, and preshower, adapted from [6].

tor. The absorber induces hadronic showers - cascades of hadrons resulting from inelastic scattering off a target nuclei. Particles in hadronic showers then pass through the scintillator, which produces light that gets transmitted through wavelength shifting fibers and read out by photodetectors, before repeating this process with the next layer of absorber/scintillator. The measured energy can be summed in sequential layers of detectors (referred to as "towers") to calculate the total energy of the incident hadron. As an additional complication, several hadrons decay to electrons or photons, which will create electromagnetic showers, requiring the scintillator to have good response to electromagnetic interactions. The HCAL must be hermetic, as a complete picture of the energy from proton collisions is required to infer the production of particles like neutrinos, which are otherwise invisible to the detector.

The relevant property for a hadronic absorber is the nuclear interaction length  $\lambda_I$ , which is the mean distance a hadron will travel before undergoing an inelastic nuclear interaction. Like the radiation length,  $\lambda_I$  is measured in  $[g\text{ cm}^{-2}]$  and can be divided by the density to

obtain a distance in [cm]. It is proportional to the density of nuclear matter, which goes as  $A^{1/3}$ . The HCAL uses a combination of brass and stainless steel, which have an interaction length of  $\sim 16.5$  cm [30].

The HCAL consists of four subsystems: an inner barrel (HB), an outer barrel (HO), an endcap (HE), and a forward calorimeter (HF). The HB covers  $|\eta| < 1.3$  with 36 azimuthal wedges, each containing 17 layers of alternating brass absorber and scintillator. These wedges begin at a radius of 1.806 m and are 89 cm thick, resulting in a minimum of  $5.3 \lambda_I$  at  $\eta = 0$ , increasing to  $10 \lambda_I$  and  $\eta = 1.2$ . The HE consists of two disks, also formed of 36 wedges with 19 layers of brass absorber and scintillator, totaling  $10 \lambda_I$  in order to contain highly boosted, forward showers resulting from quarks and gluons.. The HE extends the coverage of the HCAL up to  $|\eta| < 3.0$ . The HO is located outside the magnetic coil to catch high energy ( $> 100$  GeV) hadrons that escape the HB. It consists of five rings in eta, covering up to  $|\eta| < 1.2$ , and relies on the iron solenoid to induce hadronic showers. The HF covers the most forward region of  $3 < |\eta| < 5$ . This zone has the highest exposure of radiation, and so the detector must be sensitive only to the highest energy particles and robust against damage from heavy radiation. It uses steel as the absorber and quartz fibers to scintillate through Cherenkov radiation. The four subsystems provides a minimum of  $10 \lambda$  to capture incident hadrons.

### 3.2.6 Muon Detectors

The muon chambers lie outside the magnetic solenoid as the outermost subsystem in CMS. Due to the distance and amount of material between them and the primary IP, they are the lowest occupancy detectors, with few particles other than muons able to reach them before decaying or stopping in the inner subsystems. Although they lie outside the solenoid, the iron yolk pulls the magnetic field back to curve the trajectory of muons, which is required calculate their momenta. The detector subsystem is composed of three types of gaseous detection chambers: drift tubes (DTs), resistive plate chambers (RPCs), and cathode strip

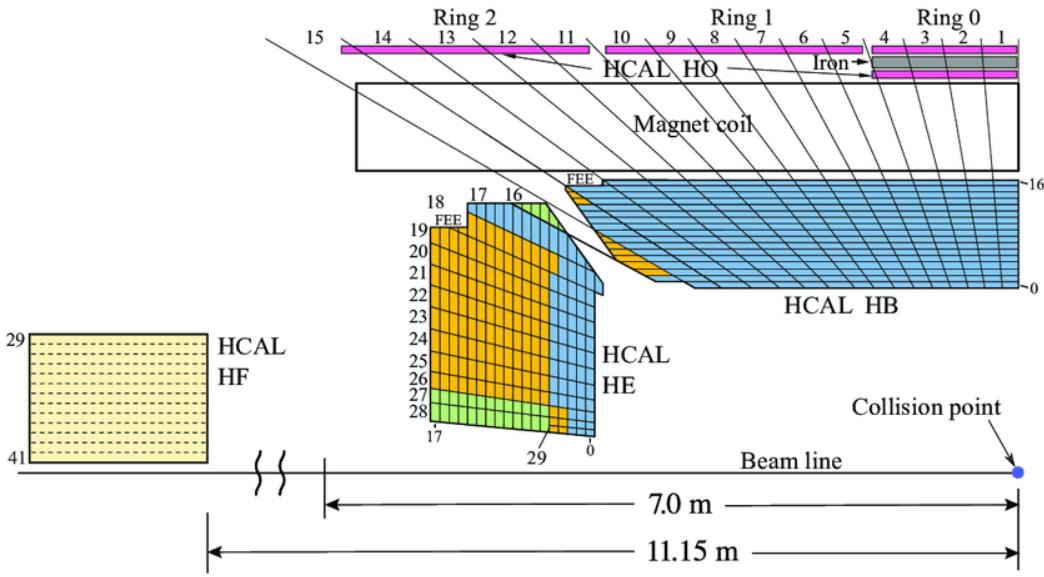


Figure 3.9:  $r$ - $z$  slice of the CMS HCAL showing the HB, HE, HF, and HO [7].

chambers (CSCs). A diagram showing the configuration of these detection chambers can be seen in figure 3.10.

The operating physics principles of all three chambers are similar: muons pass through a gas mixture and ionize the gas, knocking loose electrons. Anodes and cathodes create an electric field inside the gas chamber, causing the electrons to drift to the anode where they can be measured and read out by electronics. Electrons that are excited but not ionized from the incident muon will emit photons as they transition back to a lower energy level, thus a quenching gas is used to absorb these photons to prevent further cascades. The specifics of the three detectors are described in sections 3.2.6.1-3.2.6.3.

### 3.2.6.1 Drift Tubes

Drift tubes are used in the barrel of the muon system, covering  $|\eta| < 1.2$ . They consist of long chambers filled with an mixture of Ar/CO<sub>2</sub>, with an anode wire running lengthwise through the center and cathode strips along the edges. Electrode strips along the outer edge of the chamber shape the electric field lines to linearize the drift velocity of electrons

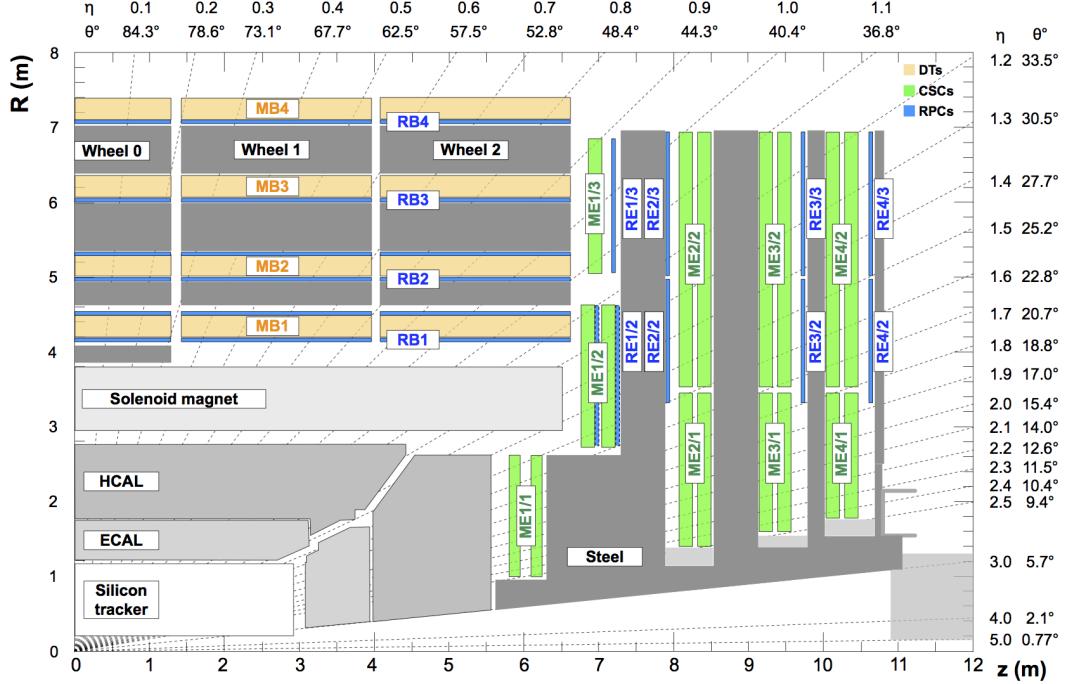


Figure 3.10:  $r$ - $z$  cross section of the CMS muon detector subsystem, configured for the 2016–2018 data taking run. The barrel consists of five wheels of DTs and RPCs in  $\eta$ , each wheel consisting of four layers in  $r$  and 12 sectors in  $\phi$ , covering up to  $|\eta| < 1.2$ . The endcap consists of four layers of DTs and RPCs in  $z$  [8].

throughout the chamber. When a muon passes through the chamber, ionized electrons follow the electric field lines towards the anode wire. Due to the known field inside the chamber, the drift time can be used to calculate the transverse distance from the wire to the incident muon with a resolution of  $250\ \mu\text{m}$ . The DTs are arranged in superlayers, which consist of 4 layers of DTs, with subsequent superlayers alternating in the  $r$ - $z$  and  $r$ - $\phi$  direction in order to obtain the muon trajectory/curvature in both the  $\eta$  and  $\phi$  direction. Figure 3.11 shows a diagram of a DT superlayer and a single DT cell.

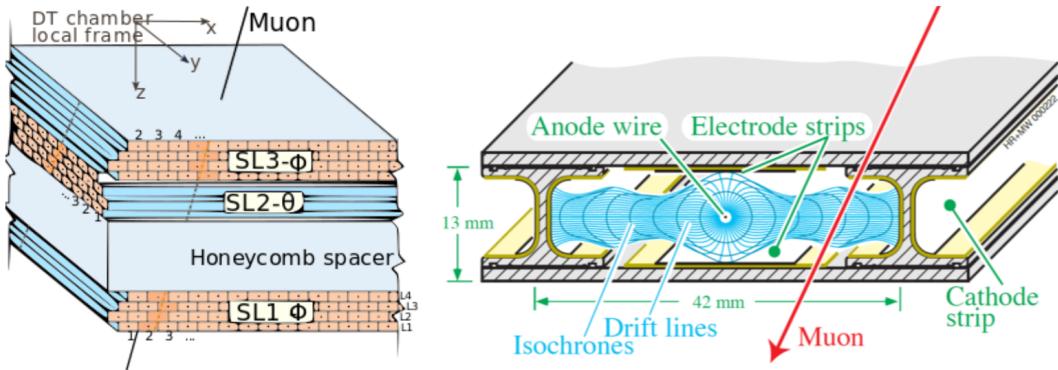


Figure 3.11: Drift tube superlayers are aligned in alternating orientations in order to provide a complete picture of muon trajectories (left). A cross section of a single drift tube chamber (right). Electrons ionized from the muon drift to the anode wire to produce a measurable current [9].

### 3.2.6.2 Cathode Strip Chambers

Cathode strip chambers are used in the endcaps of the muon system to provide coverage from  $0.9 < |\eta| < 2.4$ . CSC modules are trapezoidal chambers filled with a mixture of Ar/CO<sub>2</sub>/CF<sub>4</sub> and arranged in disks to cover the muon endcap. Each module is composed of six layers of radial cathode strips, which have high granularity in  $\phi$ , and transverse anode wires which provide granularity in  $r$ . Similar to the DTs, muons passing through CSC chambers ionize the gas molecules, knocking loose electrons which drift to the anode wires. The charge from these electrons can be measured and provides the position of the muon in  $r$ . Simultaneously,

the electrons induce a charge in the cathode strips, which provides the position of the muon in  $\phi$ . One chamber provides a spatial resolution of between  $40 - 50 \mu\text{m}$  with a time resolution of 3 ns.

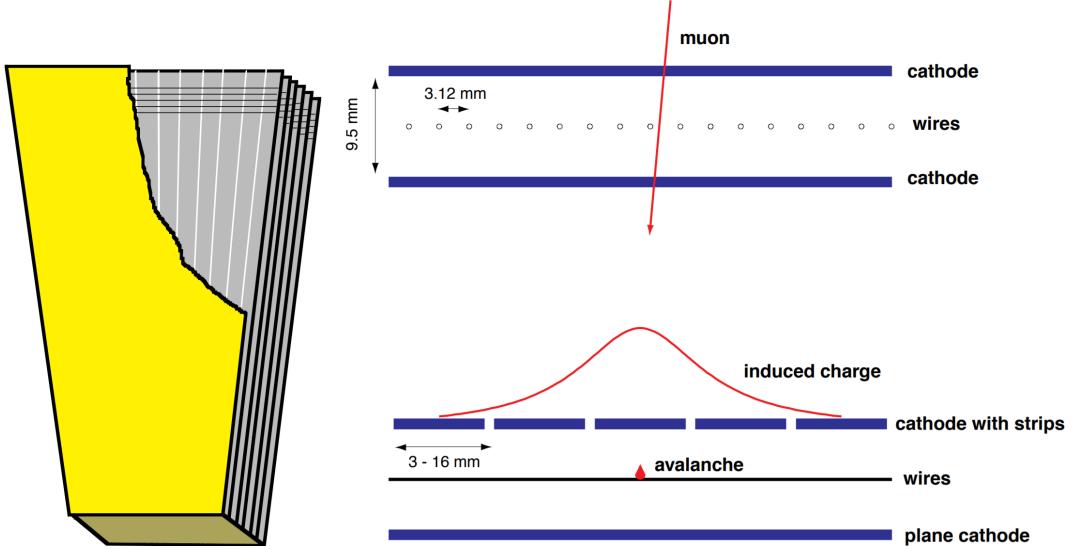


Figure 3.12: Cut-away diagram of a single CSC module showing the six layers of radial cathode strips and transverse anode wires (left). Muons ionize the gas chamber, causing an avalanche of electrons on the anode wires and inducing a charge on the cathode strips (right) [10].

### 3.2.6.3 Resistive Plate Chambers

RPC detectors are used in conjunction with the DTs and CSCs in both the barrel and endcap. One RPC consists of two high resistivity plastic plates: one positively charged anode and a negatively charged cathode which create a strong, uniform electric field between the two. The space between the two plates is filled with a mixture of 95%  $\text{C}_2\text{H}_2\text{F}_4$  and 5% iso- $\text{C}_4\text{H}_{10}$ . Incident muons ionize the gas between the two plates, which cause a cascade of electrons towards the anode plate due to the strong electric field. External metal strips lie outside the anode in order to measure the avalanche of electrons. The coarse nature of the strips means the RPC spatial resolution is lower than that of the CSCs and DTs at around

1 cm, but the time resolution is much faster on the order of 1 ns. For this reason they are used to supplement the CSC and DTs to aid with triggering and accurate timing.

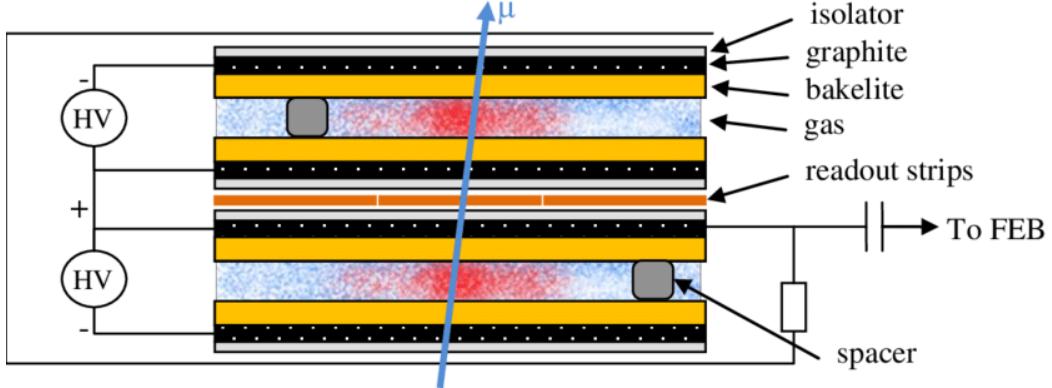


Figure 3.13: Diagram of the double RPC at CMS. Incident muons ionize the gas, causing a cascade of electrons towards the readout strips [11].

### 3.2.7 Trigger System

The high luminosity of the LHC yields an incredibly high rate of information to store from the various detectors. Each event contains approximately 1 MB of data to store, which would take bandwidth and storage space beyond current capacities to save given the 40 MHz collision rate. Additionally, many events are soft collisions that do not contain interesting physics worth storing for further analysis. In order to reduce the rate of events and select only the most interesting collisions, CMS employs a two level trigger system. The first level is known as the Level-1 (L1) trigger and uses hardware to identify physics objects in real time. The second is known as the High-Level trigger (HLT), consisting of a conventional CPU farm which performs more advanced calculations.

### 3.2.7.1 Level-1 Trigger

The L1 trigger utilizes special processors such as Field Programmable Gate Arrays (FPGAs) in order to process coarse information from the calorimeters and muon chambers in real time. The coarse detector information, known as trigger primitives, is subjected to several algorithms to create physics objects (e.g. muons, electrons/photons, etc). Objects passing a predetermined set of criteria are called trigger "seeds", with events that satisfy at least one seed getting passed to the HLT. A diagram of the L1 trigger architecture can be seen in figure 3.14. This entire decision making process must take place in  $3.2\ \mu\text{s}$  after a bunch crossing. As this is greater than the time between bunch crossings, the L1 trigger must be pipelined in order to operate continuously, while the full data is kept in a rolling buffer and read out only if the event passes a trigger seed. The L1 trigger was designed to accept events at a maximum rate of about 100 kHz, meaning only one in every four hundred events is passed to the HLT.

The L1 trigger will undergo several changes during LS3 as a result of the HL-LHC upgrade discussed in section 3.1.1. The main consequence of higher luminosity is an increase in additional proton collisions along the beamline per bunch crossing, known as pileup. Interesting physics events generally stem from a single hard scattering interaction, known as the primary vertex (PV), with lower energy pileup collisions creating background that can obscure the physics from the PV. The L1 trigger is well equipped for the average 40 pileup for Run 3, but will require upgrades in order to accommodate the expected 200 pileup for the HL-LHC. As part of the L1 trigger update, the maximum bandwidth will be increased from 100 kHz to 750 kHz and the trigger decision latency will be increased from  $3.2\ \mu\text{s}$  to  $12.5\ \mu\text{s}$ . Additionally, new tracker primitives will be available as inputs to the calorimeters and muon triggers. These primitives allow the precise  $p_T$ ,  $\eta$ , and  $\phi$  information from tracks to be associated with calorimeter clusters and muon primitives, greatly increasing the precision and efficiency of L1 reconstruction. The proposed trigger architecture for the HL-LHC can be seen in figure 3.15.

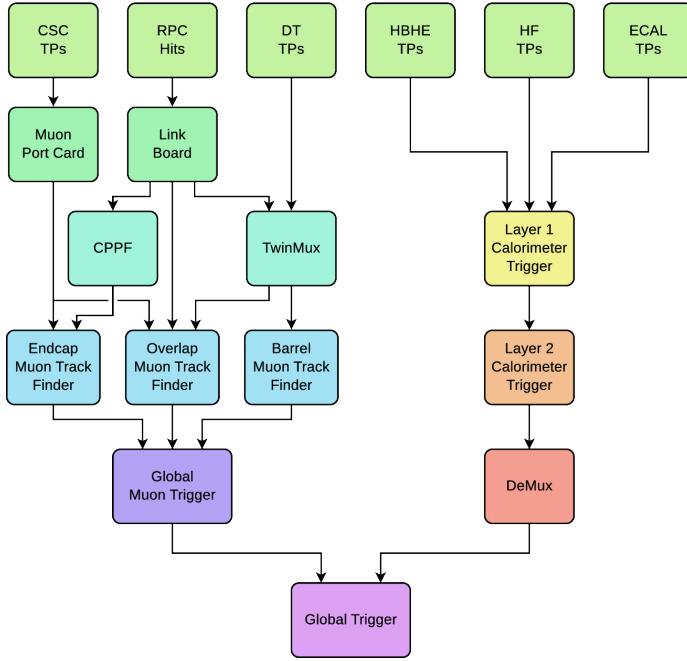


Figure 3.14: Block diagram of the CMS Level-1 Trigger architecture during Run 2 [12]. Muon objects found in the barrel, overlap, and endcap regions are ranked and sorted using their momentum and quality criteria before being passed to the Global Muon Trigger (GMT), which cleans duplicate tracks and selects the top eight muons. The muons are then passed to the global trigger, which combines the muons with objects created from the calorimeters and makes the final decision to pass the event to the HLT.

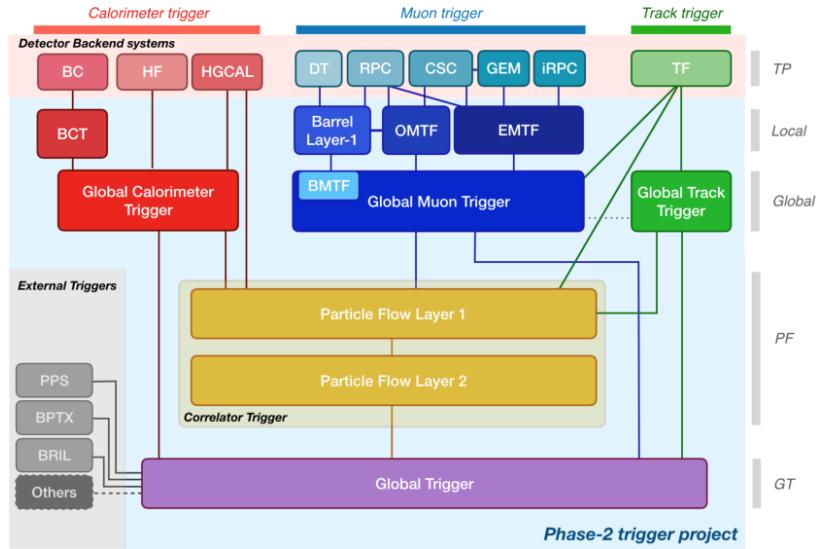


Figure 3.15: Designed trigger architecture of the HL-LHC. The addition of the track primitives allows the computation of more advanced reconstruction algorithms at the L1 level [13].

### 3.2.7.2 High Level Trigger

The HLT receives the full precision detector data for events passing the L1 trigger, and is designed to reduce the final event rate to approximately 100 Hz, five orders of magnitude less than the LHC bunch crossing rate. Information from the tracker allows the HLT algorithms to reconstruct a more complete picture of an event, as tracker information can be matched to both calorimeter hits and muon tracks to provide more detail to reconstructed physics objects. As with the L1 trigger, the HLT processes data through a series of trigger paths, which sequentially reconstruct and filter physics objects. In order to optimize computation time, the HLT uses L1 trigger seed to inform which trigger paths to use, and structures the paths in order of increasing complexity. For example, electrons are first created using only calorimeter data, then matched using coarse hits in the tracker, before being reconstructed using full tracker information. Events passing the HLT are then recorded permanently for offline reconstruction and analysis.

### 3.2.8 Event Reconstruction

Events that pass the HLT then undergo a complete event reconstruction, where data from all sub-detectors is combined to identify each physics object created in a collision. This section provides an description of the relevant objects for this dissertation, namely muons, electrons, and photons, as well as a brief description of the particle flow (PF) algorithm used for global event reconstruction.

#### 3.2.8.1 Tracks and Calorimeter Clusters

Raw detector information taken from the inner tracker and calorimeter are used to create tracks and clusters, which are the building blocks used to create a global event reconstruction. Tracks are built using a Kalman Filter algorithm to fit particle trajectories to hits in the pixel and strip detectors. This process begins with a track seed which is built from hits in the pixel detector compatible with a charged particle track. The initial seed track is propagated outward to collect nearby hits in successive layers of the tracker. Once the seed and hits are collected, the track is iteratively fit to determine the final properties of the charged particle, namely the origin,  $p_T$ , and direction. Various quality cuts are applied based on the number of hits in order to reject fake or misidentified tracks.

Clusters are seeded from individual cells in the calorimeters with an energy above a preset threshold. Adjacent cells are added to the cluster if they have an energy above twice the expected noise level to form superclusters. The average noise level per cell in the electromagnetic calorimeter is roughly 40 and 150 MeV in the barrel and endcap respectively, while the noise per tower in the hadronic calorimeter is around 200 MeV [31]. Lastly, various algorithms calculate the substructure of these topological clusters to separate contributions from individual sources into smaller clusters.

### 3.2.8.2 Muon Reconstruction

Muon reconstruction relies on tracks reconstructed in the inner tracker (tracker tracks) and tracks reconstructed in the muon systems. There are three categories of muon objects based on the method used to reconstruct tracks.

*Standalone-muon tracks* utilize a Kalman-filter algorithm to construct tracks using only information from the muon system. Tracks seeds are formed from CSC or DT track segments and propagated to nearby hits in the muon chambers.

*Global muon tracks* begin with a standalone track, which is propagated to the silicon tracker and matched to compatible tracker tracks. Two tracks are matched if the standalone track and tracker track can be propagated to a common surface. The track is then refit using a Kalman filter by combining the information from the tracker and standalone track.

*Tracker muon tracks* recapture muons in detector gaps and lower  $p_T$  muons that leave hits in the muon chambers but do not fully penetrate the muon system. In these cases, it is common for a muon to leave a segment of hits in consecutive DT/CSC layers that are compatible with a track but not produce a standalone track. Tracker muons are formed by propagating all tracker tracks with  $p_T > 0.5 \text{ GeV}$  and  $p > 2.5 \text{ GeV}$  into the muon chambers, and matching to nearby muon segments.

This reconstruction has an efficiency of 99% for muons produced within the geometric acceptance of the muon system [8]. Muon objects are then fed into the PF algorithm, which performs selection criteria based on the quality of the track. As the purity of each track varies based on type (e.g. hadrons that "punch through" the HCAL are more probable to be reconstructed as track muon tracks), the cuts vary for each category.

### 3.2.8.3 Electron and Photon Reconstruction

In theory, electrons should leave a smoothly curving track in the silicon tracker and deposit nearly all of their energy within a few crystals in the ECAL. However in practice, electrons

frequently interact with the material in the tracker, which induces electromagnetic showers before the particle can reach the calorimeter, resulting in tracks which can have sudden changes in curvature and ECAL clusters that are smeared in  $\phi$ .

The  $e/\gamma$  reconstruction algorithm accounts for these factors when track fitting and forming clusters for electrons. Electron track seeds are formed when the momentum of a track and the energy of an associated ECAL cluster are compatible with unity [31]. These tracks are refit using a Gaussian Sum Filter (GSF), which uses a sum of gaussian distributions to approximate the probability density function for energy loss due to bremsstrahlung [32].

Bremsstrahlung photons are emitted tangent to the trajectory of the electron, which bends in  $\phi$ . To ensure this energy is accounted for in the reconstruction, clusters located tangent to the GSF track are associated to the track and supercluster. Additionally, the total cluster energy is calibrated as a function of energy and  $\eta$  to account for energy loss. Photons, which are seeded by ECAL clusters with no linked GSF track, can undergo the same process of pair production to induce electromagnetic showers. As such, photon candidate clusters are subject to the same energy calibration.

Several cuts are applied to the candidate particles to reject background from hadronic processes. Electrons and photons are expected to have a low ratio of linked HCAL cluster energy to their ECAL cluster energy. Electron tracks have several variables that are fed to a boosted decision tree, which decides whether or not to accept an electron candidate. All tracks, GSR tracks, clusters, and superclusters associated with reconstructed electrons or photons are passed to the PF algorithm.

#### 3.2.8.4 Particle Flow

Particle flow (PF) is a global reconstruction algorithm that takes advantage of the unique signatures of various physics objects as they pass through each detector. It takes information passed from the muon, electron, and photon reconstruction as well as the remaining tracks

and clusters as input to form "pf blocks" in a process known as linking, which topologically associates objects from one or more detector. The particle flow algorithm then takes these blocks and reconstructs the individual particles. In order, the algorithm reconstructs muons, electrons and isolated photons, charged hadrons, and neutral particles. After each step the corresponding tracks and clusters are removed from the pf blocks to prevent double counting. Once all objects are reconstructed, a final post processing step is done to further reduce misreconstructed particles. The PF candidates are used to reconstruct higher level objects such as hadronized quarks and gluons, decay products from  $\tau$  mesons, or  $b$ -quark decays, and calculate the imbalance of transverse momentum created from particles that are invisible to the detector (called missing  $E_T$ , or MET).

## CHAPTER 4

# Real Time Muon Reconstruction at the Compact Muon Solenoid

### 4.1 Introduction

Section 3.2.7.1 provides an overview of the L1 trigger and its importance to the CMS trigger system. This section will detail the design and performance of a Kalman Filter algorithm used in the L1 trigger to identify muon tracks in the barrel region ( $|\eta| < 0.83$ ) of the CMS detector. This algorithm, known as the Kalman Barrel Muon Track Finder (KBMTF), was fully implemented for 2018 data taking and received improvements for continued use in Run-3.

The L1 barrel muon trigger receives inputs from the TwinMux system, which creates trigger primitives by combining information from the DT and RPC detectors to determine the position, trajectory, and timing of hits from incident muons at each station [33]. These hits, referred to as "stubs", are fed to the L1 muon trigger electronics where they are used to reconstruct muon tracks. The previous track finding algorithm, known as the Phase-I Barrel Muon Track Finder (BMTF), used a maximum of two stubs to reconstruct muon trajectories and was designed under the assumption that muons originate only from the beam line. As a part of upgrading the L1 Trigger, the KBMTF algorithm was developed to utilize stubs from all four muon stations as well as reconstruct displaced muon tracks resulting from decays of exotic long lived particles.

## 4.2 The Kalman Filter Algorithm

A Kalman Filter is a tracking algorithm that performs a recursive, iterative chi-square-like fit. Qualitatively, it propagates a system from its current state to the next and combines the predicted value with measurements to "update" the state of the system. This updated system is then iteratively propagated and updated with new measured values. This section focuses on a discrete linear Kalman Filter, which is applicable when a system can be described with a vector of variables whose evolution can be modeled with a linear transformation plus random uncertainties [34].

Abstractly, the state of a system at a given step  $n - 1$  can be described by a vector labeled  $x_{n-1}$  with covariance matrix  $P_{n-1}$ . Let the matrix  $F_n$  represent the linear transformation that propagates  $x_{n-1}$  to the next state  $x_n$ , and the matrix  $Q_n$  be the covariance matrix due to additional uncertainty. The predicted state of the system at step  $n$  can be given by

$$x_n = F_n x_{n-1} \quad \text{and} \quad P_n = F_n P_{n-1} F_n^T + Q_n \quad (4.1)$$

where  $F_n P_{n-1} F_n^T$  represents the propagation of the initial covariance matrix.

Now define a set of measurements taken at state  $n$  as  $z_n$  with covariance matrix  $R_n$ . The measured variables are not restricted to the same set of variables defining  $x_n$  as long as there exist a "change of basis" matrix  $H$  that relates the sets. It should be noted that mathematically  $H$  is not a strict change of basis matrix, as the measured values can have smaller dimensionality than the propagated ones. The predicted state and covariance matrix of the system written in the same variables as the measured quantities are defined as

$$\mu_n = H x_n \quad \text{and} \quad \Sigma_n = H P_n H^T \quad (4.2)$$

Updating the system relies on a matrix known as the Kalman Gain, which acts as a weight based on  $\Sigma_n$  and  $R_n$  and determines if the updated system should skew more towards

the measured or predicted values. The Kalman gain is defined as

$$K := \Sigma_n (\Sigma_n + R_n)^{-1} \quad (4.3)$$

which is then used to calculate the updated system as follows

$$\mu'_n = \mu_n + K(z_n - \mu_n) \quad \text{and} \quad \Sigma'_n = \Sigma_n - K\Sigma_n \quad (4.4)$$

The term  $z_n - \mu_n$  is frequently referred to as the residual of the prediction and measured value. A Kalman gain equal to the identity  $I$  would set the updated coordinates to the measured values, while a Kalman gain of 0 would effectively ignore the measured values. Finally, we substitute  $x_n$  and  $P_n$  into equation 4.4 using equation 4.2 and simplify to give

$$x'_n = x_n + K'(z_n - Hx_n) \quad \text{and} \quad P'_n = P_n - K'HP_n \quad (4.5)$$

where the Kalman Gain  $K$  has been redefined to

$$K' = P_n H^T (H P_n H^T + R_n)^{-1} \quad (4.6)$$

The system  $x'_n$  and  $P'_n$  can now be propagated to step  $n + 1$  where the Kalman Algorithm can be iterated.

### 4.3 The Kalman Barrel Muon Track Finder

A rough outline KBMTF algorithm for track finding is as follows:

1. A track seed is chosen from a stub in the muon station, from which a preliminary track is built. This seed cannot be chosen from the innermost muon station.
2. The track is propagated inward to the next station and matched to the closest stub.

If there is a matching stub, update the track with the stub information. Repeat until the track is at the innermost station.

3. The track is propagated from the innermost station to the beam axis. The track properties at this point are stored as the "vertex unconstrained" measurement. These properties include the  $p_T$ ,  $\phi$ ,  $\eta$ , and  $d_{xy}$ , which is defined as the closest distance from the propagated track to the beam axis.
4. The vertex propagated track is updated with the constraint that the track originated from the beam axis, and the track properties are stored as the "vertex constrained" measurement.

Muon track finding begins at the outer stations in order to get both the vertex unconstrained measurement and the vertex constrained measurement with only one iteration of the KBMTF algorithm. Starting from the inner station and propagating outward would require an additional propagation back inward in order to get the vertex constrained measurement, which would cause the algorithmic latency to exceed timing restrictions. The outer stations are also the lowest occupancy, meaning stubs are less likely to be from background. Tracks are then overlap cleaned, which ensures that stubs are not shared among multiple tracks, and cut and selected based on various goodness of fit criteria. A diagram showing the KBMTF propagation and update procedure can be shown in figure 4.1.

#### 4.3.1 Muon Trajectory Propagation

In order to implement a Kalman Filter, the propagation of muon tracks between stations must be expressed as a linear function of the track variables. From equation 3.6, a charged particle will travel in a circular orbit in the  $\hat{r} - \hat{\phi}$  plane. The high center of mass energy of the LHC results in highly energetic particles, whose trajectories have radii substantially larger than the size of the CMS detector. The large radius of curvature allows us to approximate these trajectories as parabolas. Assume a Cartesian coordinate system with the origin placed

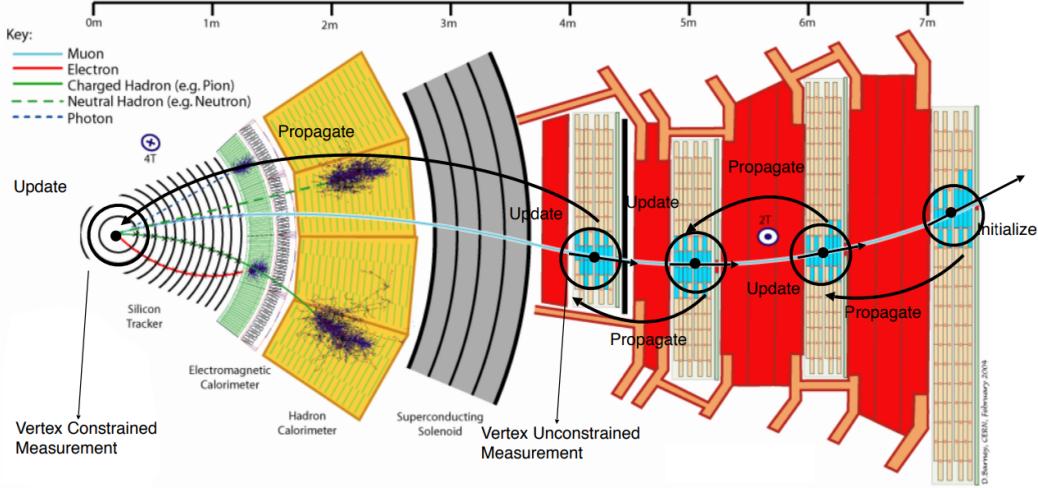


Figure 4.1: The iterative process of propagating and updating a muon track through the KBMTF algorithm. The track properties at the innermost station are stored in order to trigger on muons not originating from the beam axis [14].

along a particle's trajectory that has radius  $R$ . This trajectory can then be approximated as

$$y(x) = \frac{x^2}{2R} + bx \quad (4.7)$$

where  $b$  is a coefficient depending on the orientation of the coordinate system. If the initial trajectory is defined as  $\phi_{b,0}$ , taking the derivative and evaluating at the origin yields

$$y'(0) = \tan(\phi_{b,0}) = b \quad (4.8)$$

The curvature  $k = q/p_T$ , where the muon charge  $q = \pm 1$ , is the preferred variable to work with when propagating the trajectory of a muon, as the propagation is linear in  $k$ . Substituting values from equations 3.6 and 4.8 into equation 4.7 yields

$$y(x) = akx^2 + \tan(\phi_{b,0})x \quad (4.9)$$

where  $a = \frac{0.3B}{2}$ . Muon hits, or "stubs", provide information on the position  $\phi$  and bending angle  $\phi_b$ . Let the radius of two sequential muon stations be  $r_1$  and  $r_2$ , with the outer radius given by  $r_2$  and  $\Delta r = r_2 - r_1$ . Assuming a muon has curvature  $k$  at the outer station and ignoring energy loss, equation 4.9 can be used to propagate stubs from the outer station towards the inner station.

$$y(-\Delta r) = (a\Delta r^2) k - \Delta r \tan(\phi_{b,0}) \quad (4.10)$$

$$y'(-\Delta r) = -(2a\Delta r) k + \tan(\phi_{b,0}) \quad (4.11)$$

Converting these to the quantities to angles measured by the detector yields

$$\tan(\Delta\phi) = \frac{y(-\Delta r)}{r_1} \quad (4.12)$$

$$\phi_b = \Delta\phi + \tan^{-1}[y'(-\Delta r)] \quad (4.13)$$

The  $\Delta\phi$  term is added to equation 4.13 because the bending angle at each station is measured relative to an axis system oriented towards the center of the detector. Due to the large radius of curvature of the muon trajectories, the small angle approximation  $\tan(x) \sim \tan^{-1}(x) \sim x$  can be applied, giving the equations

$$\Delta\phi = \frac{a\Delta r^2}{r_1} k - \frac{\Delta r}{r_1} \phi_{b,0} \quad (4.14)$$

$$\phi_b = a\Delta r \left( \frac{\Delta r}{r_1} - 2 \right) k + \left( \frac{r_2}{r_1} \right) \phi_{b,0} \quad (4.15)$$

A diagram showing this propagation can be seen in figure 4.2.

Equations 4.14 - 4.15 show that the track propagation fits the criteria for a discrete linear Kalman Filter discussed in section 4.2. The matrices for propagation can now be constructed

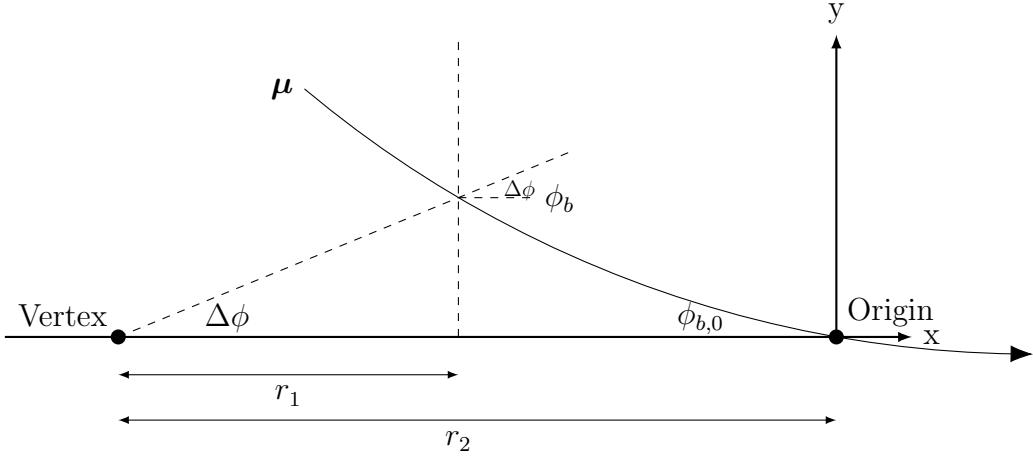


Figure 4.2: Diagram showing a muon trajectory between two stations with the origin set at the outer station. The x-axis is oriented using the detector vertex and the outer station.

as follows. The transfer matrix  $F_n$  from equation 4.1 can be expressed as

$$x_n = \begin{pmatrix} k \\ \phi \\ \phi_b \end{pmatrix}_n = \begin{pmatrix} 1 & 0 & 0 \\ \alpha_n & 1 & -\frac{\Delta r}{r_n} \\ \beta_n & 0 & \frac{r_{n-1}}{r_n} \end{pmatrix} \begin{pmatrix} k \\ \phi \\ \phi_b \end{pmatrix}_{n-1} = F_n x_{n-1} \quad (4.16)$$

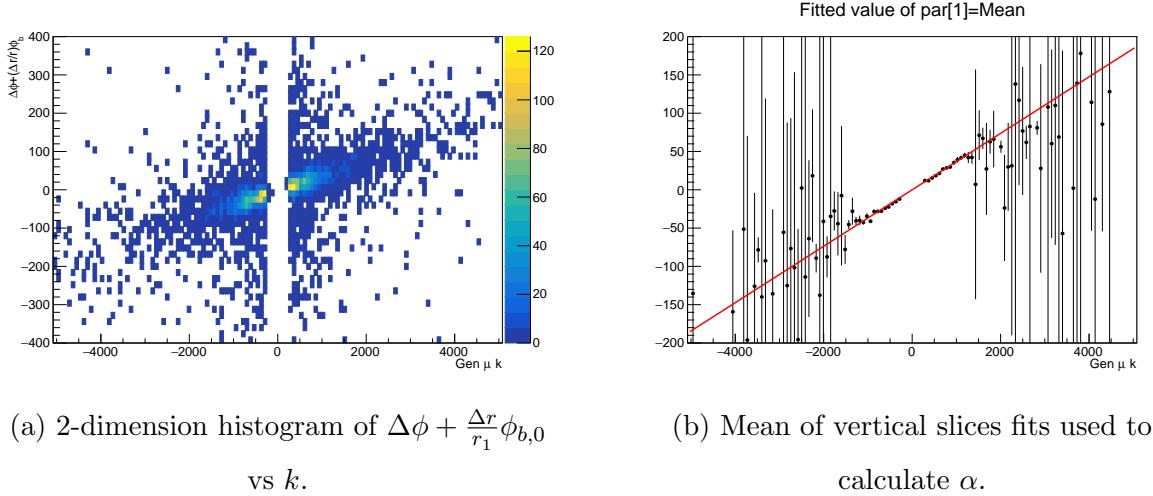
where the coefficients

$$\alpha_n = a \frac{\Delta r^2}{r_n} \quad \text{and} \quad \beta_n = a \Delta r \left( \frac{\Delta r}{r_n} - 2 \right) \quad (4.17)$$

with  $r_{n-1}$  being the radius of the previous (outer) muon station and  $r_n$  the radius of the sequential (inner) muon station.

While the radii of the muon stations are determined by detector geometry, the constants  $\alpha_n$  and  $\beta_n$  are measured using single muon monte-carlo samples where stubs are matched to the incident muons using generator level information. To determine  $\alpha_n$ , stubs in stations  $n$  and  $n - 1$  are first associated if they resulted from the same incident muon. For matching stubs, the quantity  $\Delta\phi + \frac{\Delta r}{r_1}\phi_{b,0}$  is plotted versus the muon  $k$  in a 2-dimensional histogram.

Slices along the y-axis are fit using a Gaussian distribution, and the mean value for each slice is used to calculate a linear fit as a function of  $k$ .  $\beta_n$  is calculated similarly by plotting  $\phi_b - (r_{n-1}/r_n) \phi_{b,0}$  vs  $k$ . This process can be seen in figure 4.3.



(a) 2-dimension histogram of  $\Delta\phi + \frac{\Delta r}{r_1} \phi_{b,0}$  vs  $k$ .  
(b) Mean of vertical slices fits used to calculate  $\alpha$ .

Figure 4.3: Calculation of phi propagation coefficients. Units are converted to their digitized form used in firmware calculations.

### 4.3.2 Covariant Matrices and the Kalman Gain

There are two sources of uncertainty that affect the covariance matrices. The first is multiple coulomb scattering, which occurs when muons scatter elastically with material in the detector. This can be thought of as a small deflection in momentum that alters the trajectory the muon. Lower momentum muons will be deflected at greater angles, so this uncertainty is proportional to  $k$ . Additionally, the probability for multiple scattering is proportional to the amount of material traversed by the muon, so this uncertainty must be calculated for each step in propagation. This is an uncertainty resulting from the propagation, and therefore plays the role of  $Q_n$  from equation 4.1. The second source is from the intrinsic detector resolution, which is a fixed value for each detector. However, the angular resolution is dependent on the radial distance from the center of the detector, so these must be calculated for each

station as well. Since this is an uncertainty on the measurements, the resolution plays the role of  $R_n$  in equation 4.3. The total uncertainty resulting from both of these factors is given by

$$\sigma(k) = \sqrt{\sigma_{ms}^2 k^2 + \sigma_{res}^2} \quad (4.18)$$

The coefficients can be calculated by using the Gaussian slice fits described when deriving  $\alpha$  and  $\beta$ . For each propagation, the sigma of the Gaussian fits is plotted versus  $k$  and fit to the form in equation 4.18. This fitting process for  $\phi$  propagation coefficients is shown in figure 4.4, and is functionally the same for  $\phi_b$ .

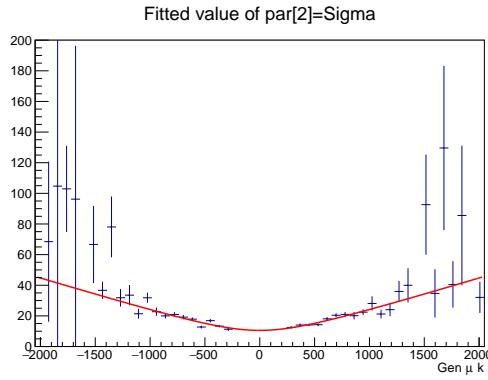


Figure 4.4:  $\sigma$  of vertical slice fits of the 2-dimensional histogram in figure 4.3a used to calculate uncertainty coefficients.

Each stub has a measured value representing the "quality" of the hits, with higher values representing a more precise reconstruction of the bending angle. For this reason, the  $\phi_b$  position resolutions are calculated separately for high and low quality stubs, as the uncertainty can differ by over a factor of 10 between the two.

Since the muon detectors only measure  $\phi$  and  $\phi_b$ , the matrix  $H$  described in equation 4.2 must be a  $2 \times 3$  matrix that takes the form

$$H = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (4.19)$$

From this, the Kalman Gain  $K'$  defined in equation 4.6 must be a 3x2 matrix with indices  $K_{ij}$ , and the update step can be explicitly written as

$$\begin{aligned} k_n^{\text{upd}} &= k_{n-1} + K_{00}\Delta\phi + K_{01}\Delta\phi_b \\ \phi_n^{\text{upd}} &= \phi_n^{\text{prop}} + K_{10}\Delta\phi + K_{11}\Delta\phi_b \\ \phi_{b,n}^{\text{upd}} &= \phi_{b,n}^{\text{prop}} + K_{20}\Delta\phi + K_{21}\Delta\phi_b \end{aligned} \quad (4.20)$$

where  $\Delta\phi$  and  $\Delta\phi_b$  are the residuals for  $\phi$  and  $\phi_b$ .

### 4.3.3 Initial Curvature Assignment

The propagation done by the KBMTF presupposes the curvature of the muon is known in order to calculate the  $\phi$  and  $\phi_b$  at the next station. In order to assign an initial value of the curvature to the track, we use the  $\phi_b$  of the seed. Take a muon track beginning at the vertex, and align the axis so the initial bending angle  $\phi_{b,0} = 0$  as seen in figure 4.5. Equation 4.9 then becomes

$$y(x) = akx^2 \quad (4.21)$$

From this, we can calculate the  $\Delta\phi$  of the muon stub located a distance  $R$  from the vertex as

$$\Delta\phi \approx \frac{y(R)}{R} = akR \quad (4.22)$$

The bending angle can then be calculated to be

$$\phi_b \approx y'(R) - \Delta\phi = akR \quad (4.23)$$

From this, we can see that  $\phi_b = \Delta\phi \propto k$ . When propagating from the vertex, the muon undergoes small but non-negligible energy loss. Assume the energy loss can be modeled by

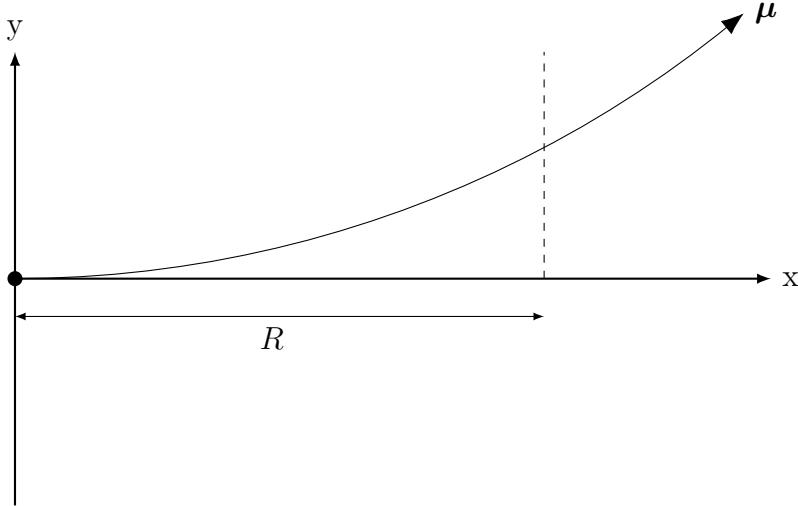


Figure 4.5: Coordinate system for muon propagation beginning at the vertex and going outward.

a constant change in momentum  $p \rightarrow p - \varepsilon$ . Then the curvature  $k$  goes as

$$k = \frac{q}{p} \rightarrow \frac{q}{p - \varepsilon} = \frac{q/p}{1 - \varepsilon/p} = \frac{k}{1 - \varepsilon|k|} \quad (4.24)$$

In a similar process to the propagation coefficient calculation, we use simulated samples to plot  $k$  versus  $\phi_b$  for each station and fit it to the form

$$k = \frac{a\phi_b}{1 + b|\phi_b|} \quad (4.25)$$

Each station then has coefficients  $a$  and  $b$  that are used to calculate the initial curvature from  $\phi_b$ .

Although the KBMTF is intended to be agnostic to the origin of the muon when reconstructing tracks, this derivation was done under the assumption that the muon originated from the vertex. We assign the initial curvature high uncertainty when setting the initial covariance matrix in order to avoid biasing the fit.

#### 4.3.4 Quality Cuts on Reconstructed Tracks

The KBMTF requires reconstructed tracks to pass a minimum quality selection in order to reject tracks from background such as punch through hadrons. The first measures the goodness of fit of the vertex unconstrained track by propagating the reconstructed track outward and measuring the difference in propagated and measured stub  $\phi$  and  $\phi_b$ . Let the track at the innermost station have curvature, position, and bending angle given by  $k$ ,  $\phi^{\text{track}}$ , and  $\phi_b^{\text{track}}$ . Applying equations 4.14 and 4.15 gives

$$(\phi^{\text{prop}} - \phi^{\text{track}}) + (\phi_b^{\text{prop}} - \phi_b^{\text{track}}) = ck \quad (4.26)$$

where  $c$  is a constant depending on the propagation coefficients. For each station, we derive  $c$  using simulation by plotting  $\Delta\phi + \Delta\phi_b$  versus  $k$ . The sum of residuals between stub and propagated coordinates is then

$$\sum_{i \in \text{stubs}} (\phi^{\text{prop},i} - \phi^i) + (\phi_b^{\text{prop},i} - \phi_b^i) = \sum_{i \in \text{stubs}} (\phi^{\text{prop},i} - \phi^{\text{track}}) + (\phi_b^{\text{prop},i} - \phi_b^{\text{track}}) + (\phi^{\text{track}} - \phi^i) + (\phi_b^{\text{track}} - \phi_b^i) \quad (4.27)$$

$$= \sum_{i \in \text{stubs}} c_i k + (\phi^{\text{track}} - \phi^i) + (\phi_b^{\text{track}} - \phi_b^i) \quad (4.28)$$

where  $(\phi^i, \phi_b^i)$  is the position and bending angle of the  $i$ -th stub and  $(\phi^{\text{prop},i}, \phi_b^{\text{prop},i})$  are the position and bending angle of the propagated track. The total residual sum is colloquially referred to as the approximate chi-square of the track. Tracks with three or more stubs must have an approximate chi-square below a preset threshold to be considered a candidate muon. The approximate chi-square is also used in the ranking of the reconstructed tracks.

#### 4.3.5 Firmware Implementation of the KBMTF

Two minor approximations are made in order to reduce algorithmic latency and memory usage for firmware implementation. The position resolution of  $\phi$  is substantially smaller than the multiple scattering resolution, so the track  $\phi$  is set automatically to the measured stub  $\phi$ . This is equivalent to fixing  $K_{10} = 1$  and  $K_{11} = 0$  in equation 4.20. Secondly, due to the high uncertainty associated with both the measured and propagated  $\phi_b$ , tracks using three or more stubs set  $K_{i1}$  equal to zero and only update using the residual from  $\phi$ .

Although the Kalman Gain can be exactly calculated exactly given the covariance matrices using equation 4.3, matrix inversion is a resource and latency expensive process for the hardware based trigger. Implementing division into hardware is significantly more resource intensive than multiplication, so the calculation of the determinant for matrix inversions would cause the algorithm to exceed the latency requirements for real time track finding. Instead, the values of  $K_{ij}$  are stored in look-up tables (LUTs) as a function of curvature and the combination of stations that were used to update the track, referred to as the track pattern. For tracks with only two stubs, the LUT contains the exact values of the gain that would have been calculated using equation 4.6. For tracks with more than two stubs, this is an approximation that loses some information from updates at previous stations but produces nearly identical performance.

As outlined in section 4.3.2, the  $\phi_b$  resolution coefficients are different for high and low quality stubs, thus the values for  $K_{ij}$  would have to be provided for every possible combination of high and low quality stubs for each track pattern. This would exponentially increase the number of gains stored on LUTs beyond the memory capacity of the hardware. However, from previous approximations, residuals for  $\phi_b$  are not used for tracks containing three or more stubs, which fixes  $K_{01}$  and  $K_{21}$  at 0 for these track patterns. Only gains for track patterns containing two stubs must account for the quality of the stubs. With four stations, this yields six track patterns with four combinations of stub quality that must be stored in

LUTs, which is within the capacity of the current hardware.

The firmware is written using Vivado High Level Synthesis (HLS) which compiles firmware written in C to HDL, optimizing it for a given FPGA and clock frequency [35]. When synthesized in a Virtex 7 690T FPGA with a clock frequency of 160 MHz, the KBMTF algorithm can process a single event in 9 bunch crossings, and is pipelined to allow for continuous processing of events. Table 4.1 shows the resource utilization of the KBMTF compared to the previous BMTF algorithm, which was implemented using an identical FPGA and clock frequency.

Algorithm	LUT	FF	BRAM	DSP
Phase I BMTF	43%	23%	35%	0%
KBMTF	16%	11%	15%	25%

Table 4.1: Comparison of FPGA resource utilization [14]. The KBMTF utilizes DSP cores for arithmetic operations to propagate and update tracks and LUTs for the propagation coefficients and approximate Kalman Gain, while the BMTF uses LUTs to assign momentum.

#### 4.3.6 Architecture of the KBMTF

The muon barrel detectors are segmented into 5 wheels along  $\hat{z}$ , which each wheel composed of 12 sectors in  $\phi$  and each sector composed of 4 DT and 3 RPC detectors in  $\hat{r}$  as seen in figure 4.6. A wedge is defined as the 5 adjacent sectors along all five wheels, shown in figure 4.7.

The hardware for the KBMTF consists of 12 Master Processor — Virtex-7 (MP7) boards, each hosting a Xilinx FPGA Virtex-7 chip, which are distributed among two  $\mu$ TCA crates [36]. The FPGA is responsible for receiving inputs from the TwinMux system through high speed optical links, reconstructing muon tracks, and delivering the muon candidates to the global muon trigger. Each MP7 corresponds to a central wedge and receives TwinMux stubs from all stations in both the central wedge and the two adjacent wedges. For a given MP7, the

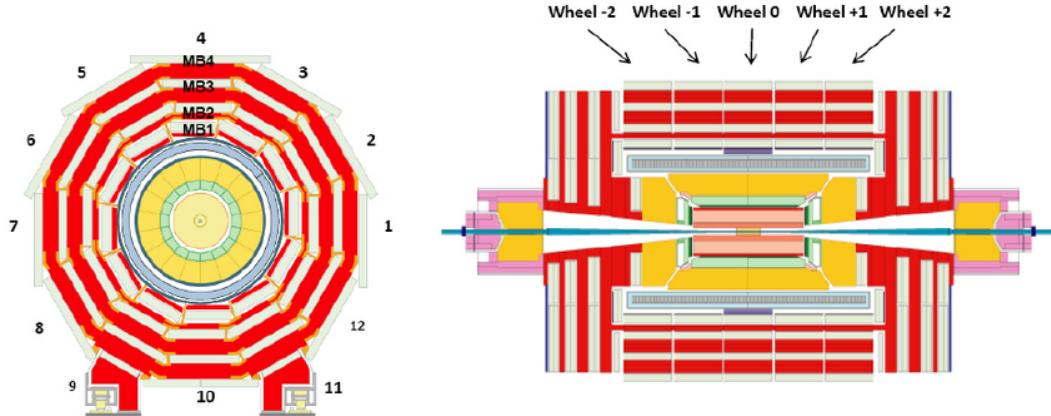


Figure 4.6: Diagram of the CMS muon barrel system [15]. (Left) Each wheel is segmented in 12 sectors in  $\phi$  and 4 DT and 3 RPC stations in  $r$ . (Right) The 5 wheels in  $z$ .

KBMTF algorithm reconstructs all tracks that have an outermost stub in the central wedge before overlap cleaning tracks that share stubs. It then picks the three best muon tracks based on track fit quality and  $p_T$  to deliver to the global muon trigger.

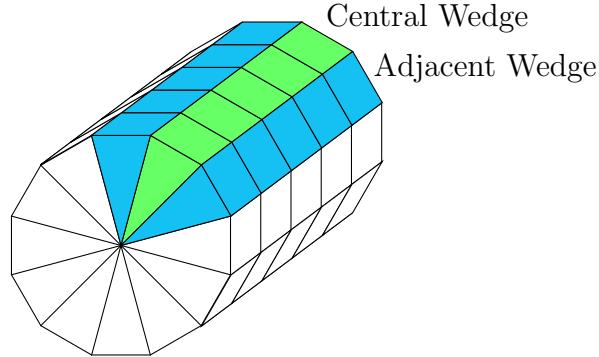


Figure 4.7: Diagram showing the inputs of one MP7 board consisting of the TwinMux inputs from one central wedge (green) and the two adjacent wedges (blue). The FPGA reconstructs all tracks with the outermost stub in the central wedge.

## 4.4 Performance of the Kalman Barrel Muon Track Finder

The performance of an trigger algorithm is primarily determined by the rate and efficiency. The efficiency is the fraction of signal events that successfully have a track reconstructed by the algorithm, while the rate is the frequency that the algorithm will trigger. A high efficiency is desirable to ensure any interesting events are stored for future analysis, while a low rate is crucial to reject background events and provide a reasonable bandwidth to store events.

### 4.4.1 Trigger Efficiencies and Rate

The previous L1 barrel muon trigger (BMTF) functioned by using pairs of stubs to form tracks, then using a LUT to assign track parameters based on the  $\phi$  and  $\phi_b$  of the stubs. This LUT takes only the  $\phi$  and  $\phi_b$  of the two stubs and was calculated with the assumption that the muon originated from the beam axis. In the case of displaced muons, which arise from secondary decays of long lived, beyond standard model particles, the assumption is no longer valid and results in large inefficiencies as the muon becomes more displaced from the beam axis.

The general efficiency of a trigger is defined as the fraction of signal muons that have a matching reconstructed track, and can be expressed abstractly as

$$\epsilon = \frac{N_{\text{signal}}(\text{Passing trigger})}{N_{\text{signal}}} \quad (4.29)$$

For muon triggers, the denominator of signal muons consist of events that have known information about the muons. This can be from generator information in simulated monte-carlo samples, tracker information using cosmic muon data, or probe muons using the tag-and-probe method on data. These signal muons are cut using the known muon information to produce a reasonable sample for track reconstruction. The numerator consists of signal

muons with a matching reconstructed track that may have addition cuts on the reconstructed track properties.

Figure 4.8 shows the efficiencies calculated on probe muons from a Z boson tag-and-probe sample collected during 2017 data taking. The associated tracker tracks for the probe muons are used as the known information, while the associated stubs are used to run the L1 muon reconstruction algorithms. Figure 4.8a shows the efficiency versus track  $p_T$  comparing the KBMTF and phase-1 BMTF algorithm, with the requirement that the matching L1 tracks have  $p_T > 20$  GeV. This requirement creates the turn-on curve effect, which has a sharpness proportional to the muon momentum resolution, followed by the plateau at which the algorithms reach maximum efficiency. Figure 4.8b shows the same efficiencies versus track  $\eta$ , with the additional requirement that the tracker tracks have  $p_T > 25$  GeV in order to show the efficiency on the plateau. Quality cuts in the KBMTF algorithm were tuned to match the efficiency of the BMTF.

Figure 4.9 shows the efficiency versus muon  $d_{xy}$  for three track finding algorithms. The denominator consists of cosmic ray muons passing through the barrel muon detectors with a  $p_T > 20$  GeV. The numerator requires a matching L1 track with reconstructed  $p_T > 10$  GeV. The lower threshold for the L1 track is set to prevent the momentum resolution of the reconstruction algorithms from affecting the efficiency. The first algorithm is the BMTF, which shows inefficiencies due to the  $p_T$  LUTs assuming the muon originated at the beam axis. The second is the prompt KBMTF, which uses the vertex constrained measurement. This also shows inefficiencies due to the vertex constraint, but is better than the BMTF as the vertex constraint isn't applied until the last step of track reconstruction. Lastly, the displaced KBMTF uses the vertex unconstrained measurement and shows substantial improvements compared to both prompt algorithms.

A reconstructed track must have a  $p_T$  above a given threshold to trigger on an event. Thresholds are set to optimize the analysis potential of the data while also minimizing the additional bandwidth to store events. The rate, which is the frequency at which a trigger

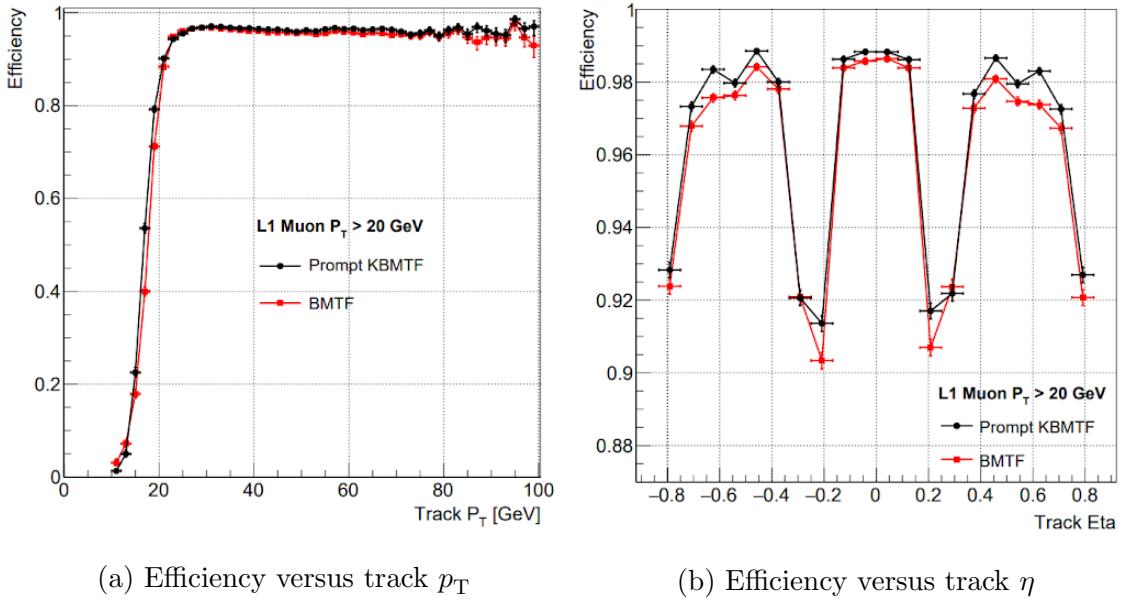


Figure 4.8: Comparison of efficiencies between the KBMTF and phase-1 BMTF algorithm using a Z tag-and-probe sample collected during 2017 data taking. (a) Efficiency versus track  $p_T$ , requiring L1 muon  $p_T > 20$  GeV. (b) Efficiency versus track  $\eta$ , requiring level-1 muon  $p_T > 20$  GeV and track  $p_T > 25$  GeV in order to show the efficiency on the plateau. The drops in efficiency at  $\eta = \pm 0.2$  are due to gaps between wheels in the muon system.

decides to store an event, can be calculated as a function of threshold using zero bias data, which is collected by storing random events regardless of trigger acceptance. This creates a representative data set of all  $pp$  collisions, hence the name "zero bias". The rate can be calculated from the zero bias data as follows

$$R = f_{\text{rev}} \times \langle N_{\text{bunches}} \rangle \times \frac{N_{\text{events}}(\text{L1 Track } p_T > \text{Threshold})}{N_{\text{events}}} \quad (4.30)$$

Where  $f_{\text{rev}}$  is the LHC revolution frequency of 11.245 kHz and  $\langle N_{\text{bunches}} \rangle$  is the average number of proton bunches that will collide in one revolution. For 2017 data taking these yield a scale factor of 20984 kHz. Figure 4.10 shows the rate versus threshold for the BMTF, vertex constrained KBMTF, and vertex unconstrained KBMTF. The KBMTF was tuned for

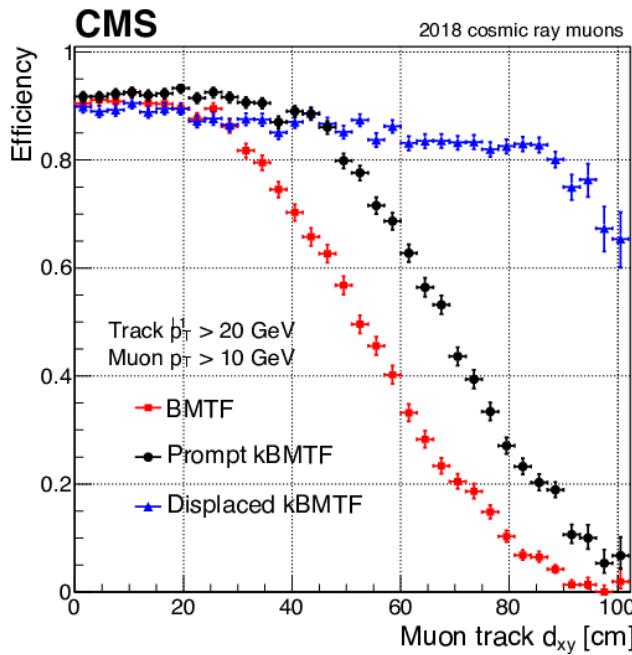


Figure 4.9: Comparison of efficiency vs cosmic track  $d_{xy}$  using cosmic data from the 2018 data taking run. The prompt KBMTF takes the reconstructed  $p_T$  using the vertex constrained measurement, while the displaced KBMTF uses the vertex unconstrained measurement. The displaced KBMTF shows substantial improvements to the previous BMTF [16].

similar efficiency but lower rate than the BMTF at 20 GeV. The displaced KBMTF initially shows lower rate at low  $p_T$  thresholds due to scale corrections to the reconstructed muon momentum.

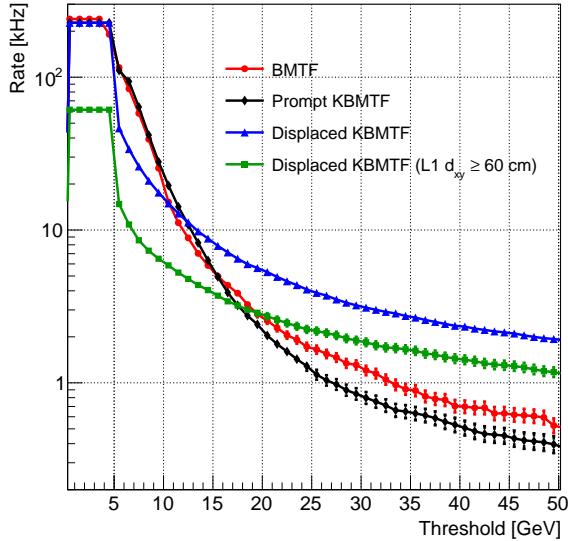


Figure 4.10: Rate comparison calculated using 2017 zero bias data of the Phase I BMTF, vertex constrained (prompt) KBMTF, vertex unconstrained (displaced) KBMTF, and displaced KBMTF for displaced tracks.

#### 4.4.2 Firmware and Emulator Agreement

The performance of the KBMTF is tested using a dedicated C++ software emulator which is integrated into the CMS analysis framework. This emulator follows similar logic to the HLS code, but is not a one-to-one replica. Several modifications are made to the HLS code in order to make it compatible with the synthesis into HDL and pipelining. As such, a test bench is used to ensure that the output of the firmware and emulator are identical. The emulator runs the algorithm over several events and outputs a text file containing information on the input stubs and the reconstructed tracks. This text file is used as input for the test bench, which runs the stub information through the HLS code and compares the properties of the

reconstructed tracks to those found in the emulator.

Disagreements between the emulator and firmware are used for a variety of debugging purposes. They frequently point to logic errors in the simplification of the algorithm for HLS synthesis, or in other cases show flaws in the emulator. During Run 2, the disagreements were almost entirely due to differences resulting from fixed point calculations in the firmware compared to floating point calculations in the emulator. These differences have been drastically reduced due to the addition of C++ classes that mimic fixed point calculations in the emulator, which now has a >99.9% agreement to the firmware. Figure 4.10 shows a comparison of the outputs of the firmware and emulator using events from a 2017 Z tag-and-probe sample.

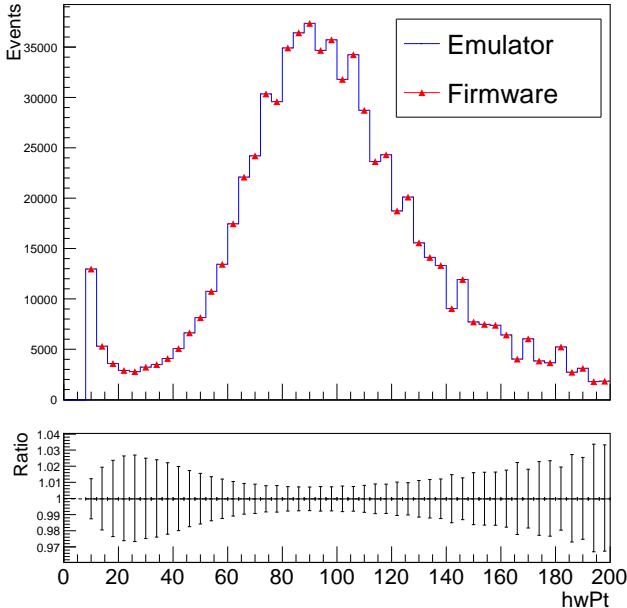


Figure 4.11: Comparison of the track  $p_T$  between the firmware and emulator using data from a 2017 Z tag-and-probe sample. hwPt represents the integer value of  $p_T$  used in hardware. The emulator and firmware output show 100% agreement across 500000 events.

## 4.5 The Phase-2 KBMTF

The modifications to the KBMTF algorithm for the HL-LHC result from increased precision from the input trigger primitives and modifications to the trigger architecture. These changes mainly affect the firmware implementation and do not require conceptual changes to the logic used to propagate and update tracks. The propagation coefficients and Kalman Gain must be recalculated as the number of bits allocated to  $k$ ,  $\phi$ , and  $\phi_b$  change. Additionally, the firmware will receive primitives from the entire barrel muon system simultaneously as opposed to processing each sector in  $\phi$  independently. The firmware will be time-multiplexed with a factor of 18, meaning data from each event will be sent to one of 18 processors running the KBMTF algorithm in parallel. Additionally, the FPGA will be changed to a ZYNQ Ultrascale+ and synthesized at a clock frequency of 200 MHz. Simulations show a resource utilization of 17% LUT, 3% FF, 11% DSP, and 7% BRAM [14].

The KBMTF algorithm performance is robust to the high pileup conditions of the HL-LHC. The efficiencies remain the same as Phase-1, and the rate scales approximately linearly with pileup as expected. Figure 4.12 shows the efficiency versus generated muon  $p_T$ , single muon rates versus threshold, and single muon rates versus pileup for the Phase-2 KBMTF and two proposed track finding algorithms. As discussed in section 3.2.7.1, the HL-LHC track trigger introduces new track primitives that will be used in novel muon triggers. These new algorithms have higher efficiency and precision for prompt muons due to the momentum resolution of the tracker, but will be unable to trigger on displaced muons as the track primitives do not include displaced tracks.

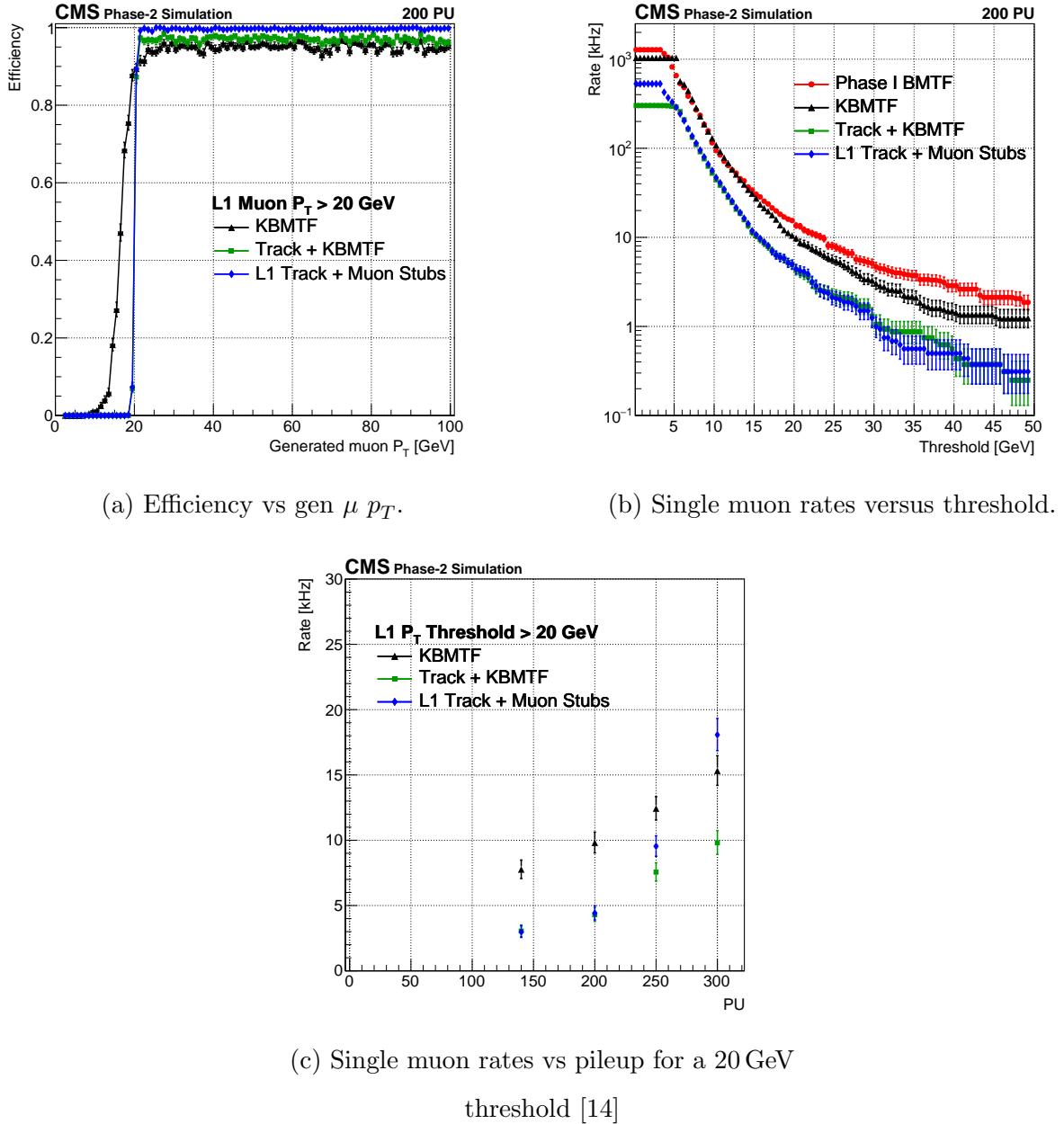


Figure 4.12: Performance of the L1 Barrel Muon algorithms using simulated samples with HL-LHC run conditions. The Track + KBMTF algorithm matches tracks to loose KBMTF muons, while the L1 Track + Muon Stubs algorithm matches tracks to muon stubs.

# CHAPTER 5

## Search for a Long Lived Scalar Boson

### 5.1 Introduction

This chapter describes the search for rare Higgs boson decays to long lived scalar bosons ( $\Phi$ ) decaying to a pair of photons using data taken by the CMS detector from 2016-2018. The following sections outline the choice of data sets and simulated samples, event selection criteria, methods of background estimation and statistical analysis for signal extraction, and results. This analysis is the first search for long lived photon pairs using a novel vertex reconstruction method to calculate the displaced vertex of the diphotons using kinematics.

The analysis is performed by searching for displaced diphoton decays of the Higgs bosons produced via associated production with a Z boson. Events are triggered and selected using the leptonic decays of the Z boson by requiring two high quality electrons or muons and two well reconstructed photons, with detailed criteria outlined in section 5.4. We utilize a data-driven background estimation shown in section 5.5 to model the background from Drell-Yan events containing photons from initial/final state radiation as well as neutral pions in jets, then extract the branching fraction of the Higgs boson to  $\Phi\Phi$ .

As described in section 3.2.4, the CMS ECAL is a homogeneous calorimeter that measures the energy deposits from electromagnetically interacting particles. This lends the ECAL precise energy resolution but low capabilities to determine the direction particles were traveling when they reached the calorimeter, referred to as pointing. Photons are electrically neutral and do not leave tracks in the inner tracker, meaning that without pointing it is impossible

to establish from where they originated. This places major restrictions on searches for displaced photons, as the transverse distance traveled by the  $\Phi$  before decaying to photons is the main variable used to extract the signal from data. We will show in section 5.3 that it is possible to recalculate the diphoton vertex using a kinematic constraint, which is performed for each hypothetical mass of the  $\Phi$ .

Results are calculated using a maximum likelihood based approach described in section 5.6 to extract upper limits on the signal strength parameter. In section 5.7 we set limits on the product of the Higgs boson branching ratio to  $\Phi\Phi$  times the LLP branching ratio to photons, as well as model independent limits on the generic  $Z + \Phi\Phi$  cross section.

## 5.2 Data and Simulated Samples

### 5.2.1 Data Samples

As discussed in chapter 3, events that pass at least 1 HLT path are stored for future analysis. These events are collected and sorted into data sets based on the number and type of physics objects used in the HLT path. It can be noted that these data sets are not disjoint; one event can be sorted into multiple data sets if it satisfies more than a single HLT path, so care must be taken to avoid double counting events. The data are maintained and updated by the Physics Data and Monte Carlo Validation (PdmV) group, which also provides recommendations on optimal usage for physics analysis [37]. This search utilizes the PdmV recommended data sets for Run-2 analyses, referred to as Ultra-Legacy (UL), which has been reprocessed using the most current calibration for detector responses and alignment. Although the characteristic signature of this analysis is two displaced photons, the HLT double photon triggers present in Run-2 are unsuitable for the phase space of signal photons. Therefore we utilize the leptonic decay of the associated  $Z$  which provides a robust trigger. Since we are interested in  $Z$  decays to either electrons or muons, we use the data sets labeled as SingleMuon or SingleElectron (renamed to EGamma in 2018). The complete

list of data sets is listed in table 5.1.

Table 5.1: CMS data sets used for each year in this analysis. Asterisks represents wildcards that cover the multiple data taking eras per year.

Year	CMS Data Set
2018	/EGamma/Run2018*-UL2018_MiniAODv2_NanoAODv9* /SingleMuon/Run2018*-UL2018_MiniAODv2_NanoAODv9*
2017	/SingleElectron/Run2017*-UL2017_MiniAODv2_NanoAODv9* /SingleMuon/Run2017-UL2017_MiniAODv2_NanoAODv9*
2016	/SingleElectron/Run2016*-UL2016_MiniAODv2_NanoAODv9* /SingleElectron/Run2016*_HIPM_UL2016_MiniAODv2_NanoAODv9* /SingleMuon/Run2016*-UL2016_MiniAODv2_NanoAODv9* /SingleMuon/Run2016*_HIPM_UL2016_MiniAODv2_NanoAODv9*

The reprocessed datasets must be cross referenced with the "Golden JSON": a certificate created by the Data Quality Monitoring (DQM) group within CMS by monitoring each sub-detector system in order to veto collisions that occur during sub-optimal conditions. The total integrated luminosity recorded by CMS during Run-2 was  $164\text{ fb}^{-1}$ ; however, only  $138\text{ fb}^{-1}$  of that was certified for analysis. The total integrated luminosity and uncertainty per year for collisions recommended by DQM can be seen in table 5.2.

Table 5.2: Table of integrated luminosity per year [17–19].

Year	2016	2017	2018	2016-2018
Luminosity [ $\text{fb}^{-1}$ ]	36.31	41.48	59.83	137.62
Uncertainty [%]	1.2	2.3	2.5	1.6

The 2016 data and simulation are divided into two distinct eras due to substantial changes to the silicon strip tracker. Saturation effects in the pre-amplifier of the APV25 readout chip resulted in a decreased signal-to-noise ratio and required a decrease to the feedback pre-amplifier bias voltage (referred to as VFP) [38]. As a result, simulated samples were produced separately for pre and post-VFP run conditions. The pre-VFP and post-VFP data corresponds to  $19.65$  and  $16.67\text{ fb}^{-1}$  respectively.

## 5.2.2 Simulated Samples

Simulated samples use Monte Carlo methods to generate events in several steps, from the matrix elements that determine cross sections for various physics processes, showering, and hadronization to detector responses and event reconstruction. These samples are generally created to model specific physics processes such as Drell-Yan, QCD,  $t\bar{t}$ , signal processes, etc. All samples used in this analysis use parton distribution functions given by NNPDF3.1 [39]. Simulated events are not created one-to-one to match data events; instead they are used to create weighted events that are normalized using the cross section of the simulated process and scaled to match the luminosity of the data. For this analysis we use both simulated background events as well as simulated signal samples.

### 5.2.2.1 Monte Carlo Signal

Standard Model Higgs production through quark induced and gluon induced ZH processes are simulated at next-to-next-to leading order (NNLO) for several candidate mass points  $m_\Phi$  and lifetimes  $c\tau$ . The primary hard scattering interaction producing the Higgs Boson and associated Z boson is simulated with matrix elements calculated at NNLO using POWHEG BOX v2 [40], while the subsequent decays of the Higgs Boson and showering processes are simulated using PYTHIA8 [41]. Leading order Feynman diagrams for the ZH production are shown in figure 5.1, while the decays of the  $\Phi$  boson can be seen in figure 5.2. The parameters for Pythia8 simulations were set to TuneCP5, which were the optimal values for agreement with measured data at NNLO at the time of producing the samples [42]. The simulated signal samples were produced centrally by the Monte Carlo and Interpretation subgroup of CMS.

The Higgs Boson is forced to decay to two long lived scalar bosons ( $\Phi$ ). Each  $\Phi$  has a 50% branching ratio to  $\gamma\gamma$  and a 50% branching ratio to  $d\bar{d}$  as shown in figure 5.2. The decay to  $d\bar{d}$  was chosen to create a generic jet with no characteristic signature that would affect

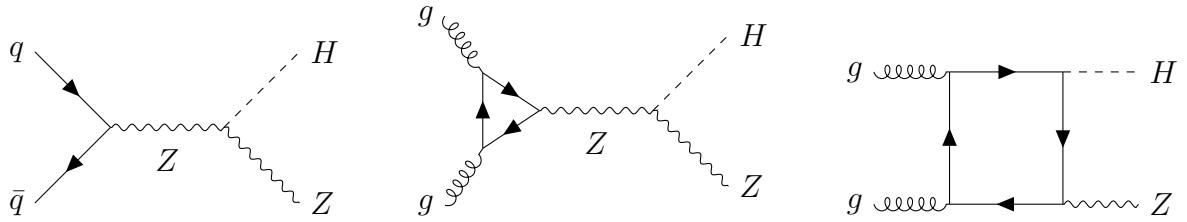


Figure 5.1: Feynman diagrams for both quark induced and gluon induced  $Z+H$  production.

the event topology. Additional samples with a 100% branching ratio to  $\gamma\gamma$  were generated to provide better statistical power for four photon decays. The  $Z$  is forced to decay leptonically, which is accounted for when normalizing samples for the limit setting procedure. A complete list of masses, lifetimes, and number of generated events is shown in table 5.3.

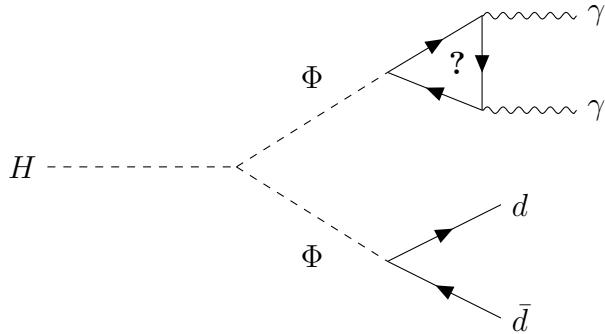


Figure 5.2: Feynman diagram showing the proposed decay channel of the Higgs boson to  $\Phi\Phi$ , with one  $\Phi$  decaying to a pair of photons and the other decaying to  $d\bar{d}$ . The loop in the diphoton decay can contain any particle that could theoretically couple to the  $\Phi$ . It is also possible for the  $\Phi$  to decay directly to photons.

### 5.2.2.2 Signal Lifetime Reweighting

One can interpolate between signal sample lifetimes using a process known as  $c\tau$  reweighting. As each generated event contains two  $\Phi$  with exponentially distributed lifetimes, an event weight can be applied to generate a sample of arbitrary lifetime. Assume an event generated with mean lifetime  $\tau_0$ , with each  $\Phi$  lifetime given by  $t_1$  and  $t_2$ . To reweight this sample

Table 5.3: Mass points, lifetimes, and number of simulated  $H \rightarrow \Phi\Phi$  events for the  $2\gamma 2q$  categories. Number of events generated is halved for the  $4\gamma$  categories

$m_\Phi$ [GeV]	$c\tau$ [mm]	Events
15	0, 10, 20, 50, 100, 1000	100000
20, 30, 40, 50, 55	0, 10, 20, 50, 100	50000
	1000	100000

to a new mean lifetime  $\tau_{\text{new}}$ , we apply a weight given by the ratio of poisson probability distributions

$$w(\tau_0, \tau_{\text{new}}, t_1, t_2) = \frac{\mathcal{P}(t_1|\tau_{\text{new}})\mathcal{P}(t_2|\tau_{\text{new}})}{\mathcal{P}(t_1|\tau_0)\mathcal{P}(t_2|\tau_0)} = \left(\frac{\tau_0}{\tau_{\text{new}}}\right)^2 \frac{e^{\frac{-(t_2+t_1)}{\tau_{\text{new}}}}}{e^{\frac{-(t_2+t_1)}{\tau_0}}} \quad (5.1)$$

The  $\Phi$  lifetimes  $t_i$  can be calculated from the generator level information as follows

$$ct = \frac{\beta}{\gamma} \cdot \text{ip}_{3D} = \frac{m|\mathbf{p}|}{E^2} \cdot \text{ip}_{3D} \quad (5.2)$$

Where  $c$  is the speed of light,  $m$ ,  $\mathbf{p}$ , and  $E$  are the mass, 3-momentum, and energy of the  $\Phi$  and the three-dimensional impact parameter  $\text{ip}_{3D}$  is the distance in the lab frame traveled by the  $\Phi$  from the PV before decaying.

Lifetimes are reweighted to 3, 14, 32, 70, and 316 mm to be uniformly distributed between generated signal samples on a logarithmic scale. The signal sample with the nominal lifetime closest to the target lifetime is chosen to be reweighted, excluding the 3 mm sample which is produced using the 10 mm sample.

### 5.2.2.3 Pileup Reweighting

Events passing the HLT generally consist of one interesting hard-scatter occurring at the primary vertex (PV). However, as discussed in chapter 3.2.7, colliding proton bunches often

produces secondary collisions along the beamline known as pileup (PU). Particles resulting from PU are generally uninteresting on their own but produce tracks and energy deposits that affect event reconstruction. During Run-2 data taking the average pileup ranged from 27 in 2016 to 37 in 2018. PU distributions in data can be seen in figure 5.3.

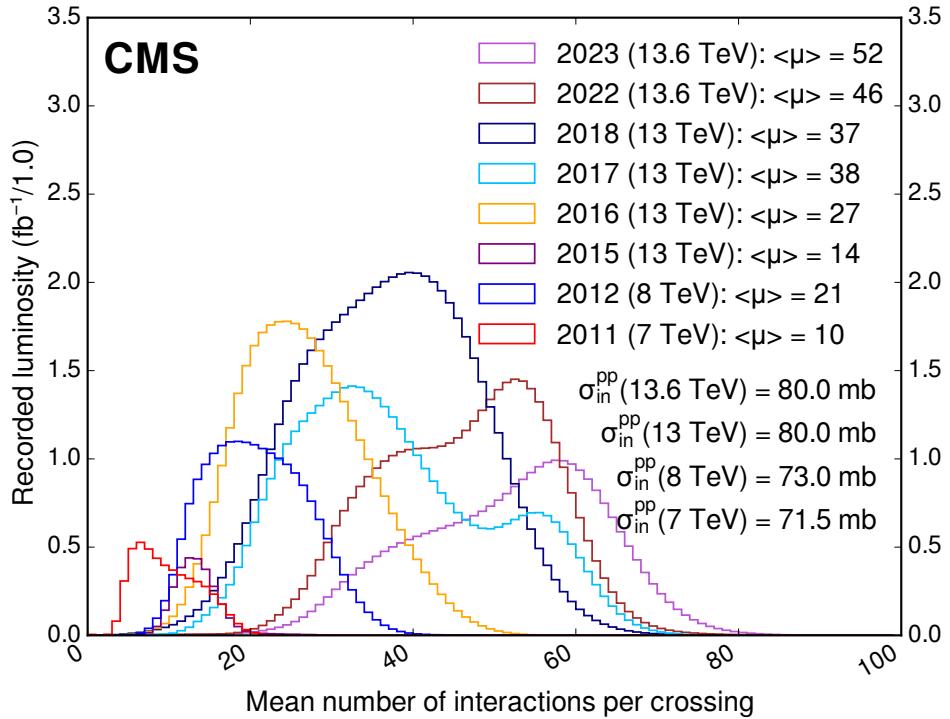


Figure 5.3: Pileup distributions and mean PU from 2011-2023 datataking [1].

Monte Carlo samples account for secondary collisions by generating PU interactions according to a predetermined distribution. However, the distribution of the number of interactions does not exactly match what is observed in data. To account for this, we apply a weight to simulated samples as a function of the number of simulated PU interactions by dividing a histogram of data PU distributions by the Monte Carlo PU distribution.

#### 5.2.2.4 Monte Carlo Background

Although this analysis employs a data driven method of background estimation, simulated samples are used to inform on properties of background events and tune cuts to discriminate between signal and background. For this analysis, we use simulated Drell-Yan (DY) to  $\ell\ell$  events with 0, 1 and 2 additional jets, simulated using matrix elements calculated at NNLO. Studies were done with inclusive (any number of additional jets) samples but found these samples to have worse agreement in control regions and statistical power compared to the jet categorized samples. The cross sections used for normalization were calculated using the `GenXSecAnalyzer`, a tool provided by CMS to calculate the cross sections for simulated samples [20]. A list of cross sections for each simulated process can be found in table 5.4.

Table 5.4: List of simulated background processes, their cross sections calculated using `GenXSecAnalyzer` [20], and CMS data set names. Asterisks stand for `18NanoAODv9-106X-upgrade2018_realistic_v16` for 2018, `17NanoAODv9-106X_mc2017_realistic` for 2017, `6NanoAODAPVv9-106X_mcRun2_asymptotic_preVFP` for 2016 pre-VFP, and `16NanoAODv9-106X_mcRun2_asymptotic` for 2016 post-VFP.

Process	Cross Section [pb]	Data set
DY + 0 jets	5090	<code>DYJetsToLL_0J_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL*</code>
DY + 1 jet	983.5	<code>DYJetsToLL_1J_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL*</code>
DY + 2 jets	353.5	<code>DYJetsToLL_2J_TuneCP5_13TeV-amcatnloFXFX-pythia8/RunIISummer20UL*</code>

#### 5.2.2.5 Scale Factor Corrections

Scale factors are corrective weights applied to simulated events in order to improve the agreement to data. The total weight applied is defined by the data-to-simulation efficiency ratio, given by

$$S = \frac{\epsilon(\text{data})}{\epsilon(\text{MC})} \quad (5.3)$$

The total efficiency  $\epsilon$  in data and Monte Carlo is a product of the efficiencies for each object used in the event selection. The efficiency for a given object is itself a product of several individual efficiencies. For muons, the efficiency  $\epsilon_\mu = \epsilon_{\text{reco}} \times \epsilon_{\text{trig}} \times \epsilon_{\text{ID}} \times \epsilon_{\text{iso}}$  is a product of the muon track reconstruction efficiency, trigger efficiency, ID efficiency, and isolation efficiency. For electrons and photons,  $\epsilon_{\text{iso}}$  is absorbed into the  $\epsilon_{\text{ID}}$  efficiency because the cut based ID includes the isolation. The efficiency for electrons is then  $\epsilon_e = \epsilon_{\text{reco}} \times \epsilon_{\text{trig}} \times \epsilon_{\text{ID}}$ . Photons are not used for any HLT triggers nor do they have associated tracks, so  $\epsilon_{\text{trig}}$  and  $\epsilon_{\text{reco}}$  are omitted, making the total efficiency  $\epsilon_\gamma = \epsilon_{\text{pix}} \times \epsilon_{\text{ID}}$ , where  $\epsilon_{\text{pix}}$  is the efficiency for the pixel seed veto. The total scale factor is then given by

$$S = \prod_{i \in \mu, e, \gamma} \frac{\epsilon_i(\text{data})}{\epsilon_i(\text{MC})} \quad (5.4)$$

All scale factors are centrally measured and provided by the respective POGs with the exception of the electron trigger efficiencies, which were privately produced. Almost all efficiencies are measured with the tag-and-probe technique, which uses events containing one lepton passing strict selection criteria (the "tag") and one oppositely charged same flavor lepton passing loose selection criteria (the "probe"), where the dilepton invariant mass lies within the Z resonance centered at 91 GeV. The strict tag and invariant mass criteria give a set of real leptons that are unbiased due to the loose probe criteria which are used to calculate various efficiencies. This method was adapted for photon ID efficiency calculations by ignoring the tracker information for  $Z \rightarrow ee$  events and treating the calorimeter clusters as photons. The photon pixel seed veto efficiencies require tracker information and cannot be calculated using  $Z \rightarrow ee$  tag-and-probe techniques. These were instead measured using  $Z \rightarrow \mu\mu + \gamma$  events, where the photon is from final state radiation. Invariant mass and topological cuts give a clean sample of unbiased photons which are used to calculate  $\epsilon_{\text{pix}}$ .

Electron HLT scale factors were not centrally provided and were calculated privately using the tag-and-probe technique described above. The tag+probe were required to have

a combined invariant mass of  $60 < m_{\text{tag-probe}} < 125 \text{ GeV}$ . Baseline criteria for the tag and probe can be seen in table 5.5.

Table 5.5: Tag and probe criteria used to derive HLT scale factors

Tag Criteria	Probe Criteria
Matched to Single Electron Trigger	$p_T > 15 \text{ GeV}$
Passing tight cut-based ID	$ \eta  < 2.5$
$ \eta  < 2.17$	$E_T > 5 \text{ GeV}$
$ \eta  < 1.4442$ or $ \eta  > 1.566$	$ d_z  < 0.2 \text{ cm}$
$p_T > 35 \text{ GeV}$	$ d_{xy}  < 0.2 \text{ cm}$
	Passes tight cut-based ID

Events that pass the baseline tag/probe criteria are binned according to the  $p_T$  and  $\eta_{sc}$  of the probe, and split into pass/fail histograms if the probe passed/failed to trigger the HLT path. For data, the histograms are fit analytically using the Z peak plus background to extract the number of signal events and calculate the efficiency. In Monte Carlo, the events are first matched to generator level electrons originating from the Z to ensure zero background events and fit to only the Z signal. The signal is modeled using a piecewise function with the upper and lower sides modeled using gaussian cores and exponential tails, which is then convolved with the world-average Z mass and width [43]. Background is modeled by multiplying an error function with a falling exponential. To calculate the systematic uncertainty of the fit, the efficiencies are calculated for data under alternative signal and background models. For the alternate signal, the data models the signal shape as a convolution of a gaussian and the Monte Carlo signal from the equivalent bin. For the alternate background, the shape is modeled as just an exponential, which can either grow or decay depending on the  $p_T$  and  $\eta$  range of the bin. This process is repeated for each era, as the trigger and run conditions can vary between years. The trigger efficiency as a function of  $\eta$  can be seen in figure 5.4. A full set of scale factors is shown in figures 5.5 - 5.8.

Muon reconstruction, isolation, and ID scale factors are applied for each muon used to reconstruct the Z as a function of  $p_T$  and  $|\eta|$ . Muon trigger scale factors are as a function of

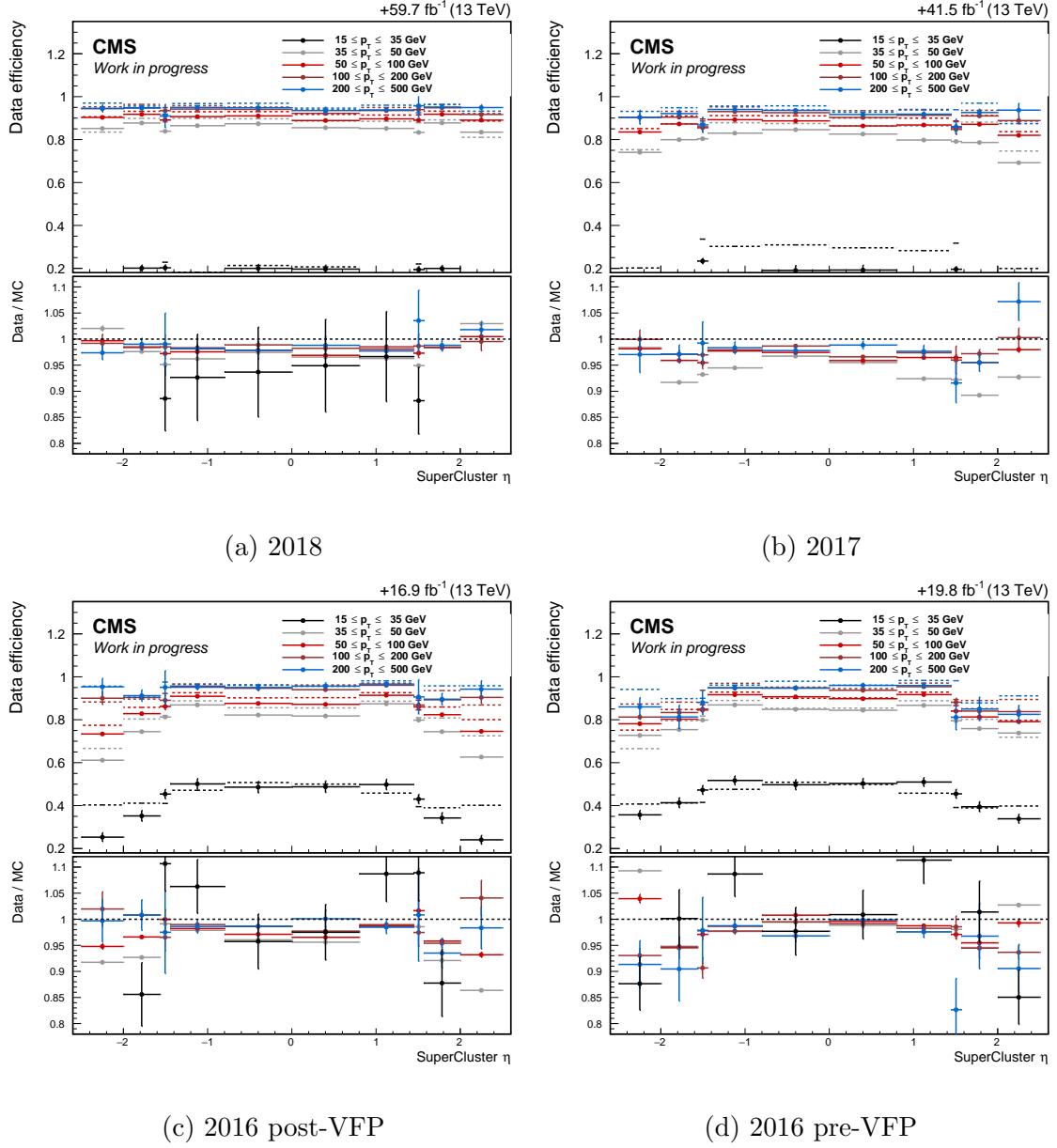


Figure 5.4: Trigger efficiency as a function of  $\eta_{sc}$  for several ranges of electron  $p_T$  for all data taking eras.

$p_T$ ,  $\eta$  and charge of the triggering muon, or the higher  $p_T$  muon if both pass the single muon trigger. The photon ID and all electron scale factors are applied as a function of electron  $p_T$  and  $\eta$ , with the electron HLT scale factor application following the same logic as the muon

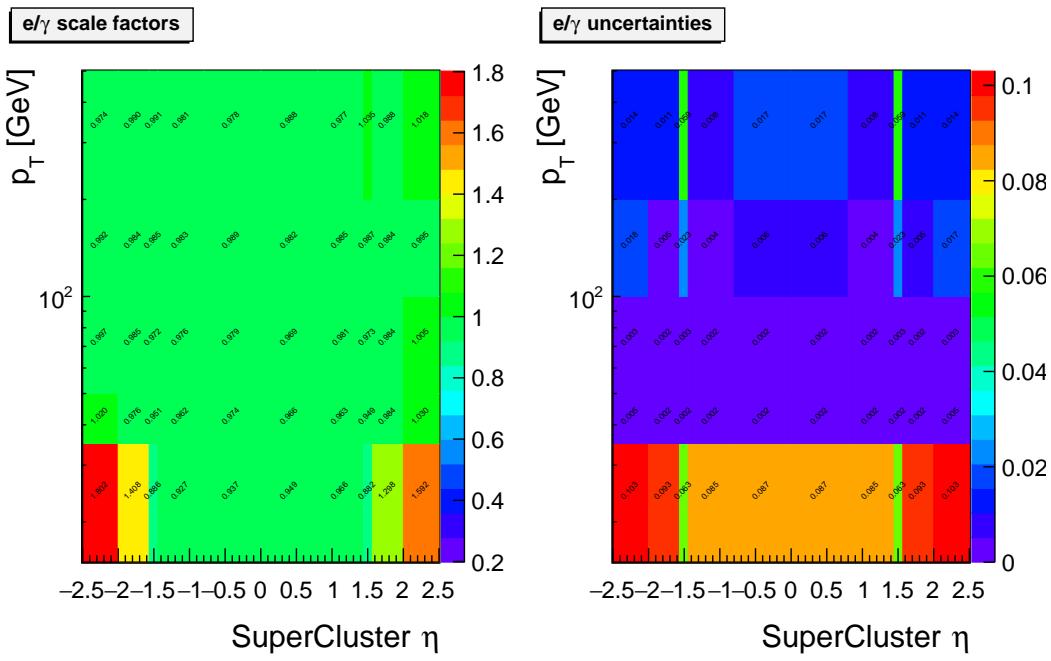


Figure 5.5: 2018 electron HLT scale factors and uncertainties.

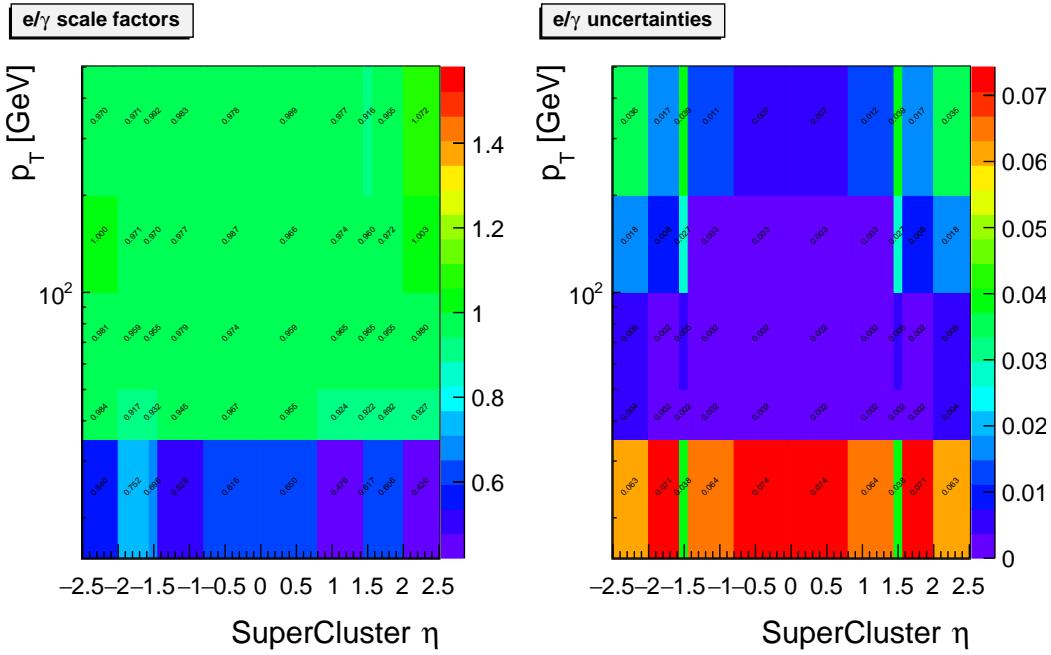


Figure 5.6: 2017 electron HLT scale factors and uncertainties.

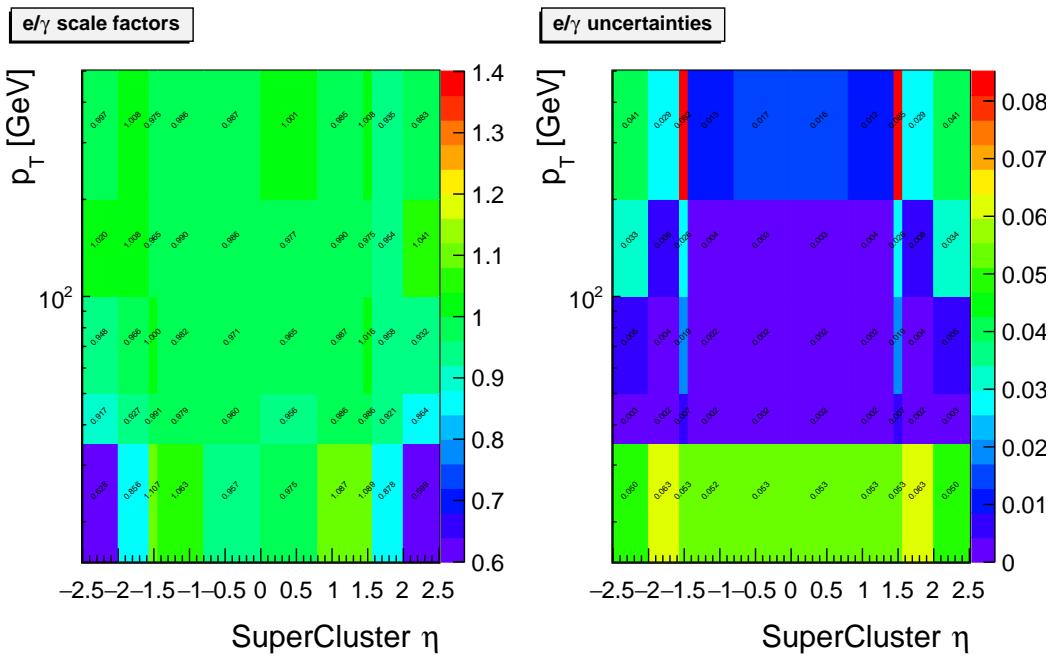


Figure 5.7: 2016 post-VFP electron HLT scale factors and uncertainties.

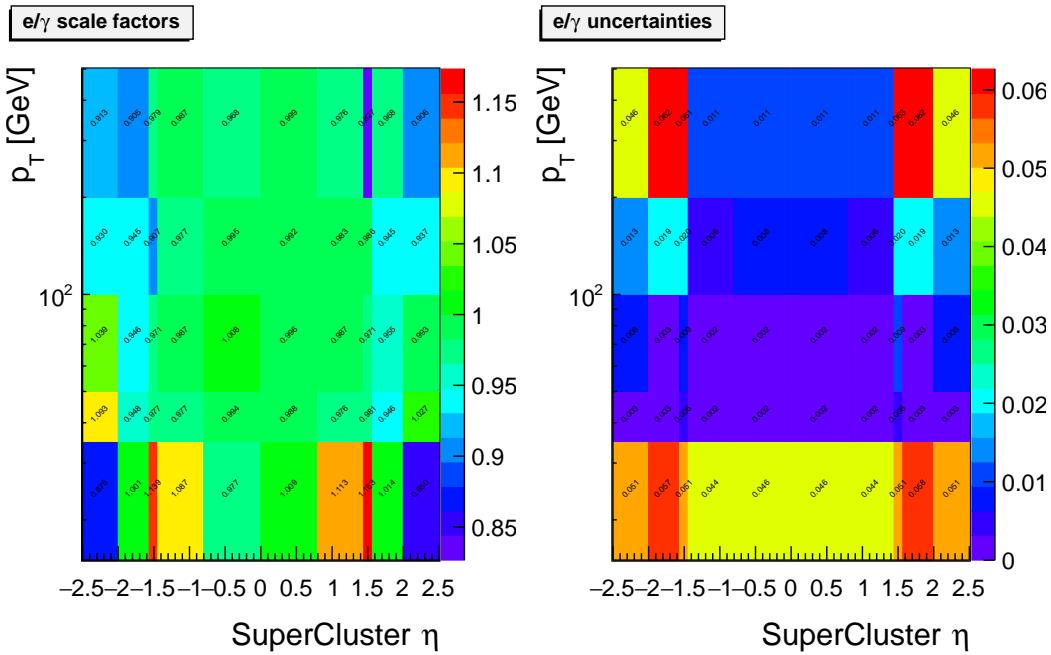


Figure 5.8: 2016 pre-VFP electron HLT scale factors and uncertainties.

HLT. Lastly, the photon pixel seed veto scale factor is applied as a function of the  $\eta$  region (barrel or endcap) of the photons. Most scale factors are on the order of a few percent with uncertainties less than 0.1%.

## 5.3 Vertex Reconstruction of Displaced Photon Pairs

### 5.3.1 Diphoton Kinematics

The kinematics of a long lived, massive scalar boson  $\Phi$  decaying to two photons can be precisely described by the following

$$m_\Phi^2 = 2E_1 E_2 (1 - \cos \theta) \quad (5.5)$$

where the  $m_\Phi$  is the mass of the  $\Phi$ ,  $E_1$  and  $E_2$  are the energies of the two photons, and  $\theta$  is the angle between the two photons. The energy of the two photons is measured by the CMS ECAL, thus  $\theta$  can be calculated as a function of  $m_\Phi$ . Due to the primary vertex, decay vertex, and ECAL clusters being coplanar, this calculation is treated as a 2 dimensional problem. It is a trivial procedure to rotate the system to calculate the vertex, then rotate back to detector coordinates once the decay vertex is calculated. Let the position of the two ECAL clusters be  $\mathbf{r}_1$  and  $\mathbf{r}_2$ . The system can be rotated by an angle

$$\varphi = \arccos \left( \frac{\mathbf{r}_1 \times \mathbf{r}_2}{\|\mathbf{r}_1 \times \mathbf{r}_2\|} \cdot \hat{z} \right) \quad (5.6)$$

around an axis given by

$$\hat{n} = \frac{\mathbf{r}_1 \times \mathbf{r}_2}{\|\mathbf{r}_1 \times \mathbf{r}_2\|} \times \hat{z} \quad (5.7)$$

to determine the decay vertex, then rotated by  $-\varphi$  around the same axis to give the final value.

The geometry of the vertex fit is described in figure 5.9. Assume we have two ECAL

clusters and a given value of  $m_\Phi$ . This fixes the value of  $\theta$  from equation 5.5. It will be shown that the locus of points satisfying these constraints corresponds to the arcs of two circles that overlap the ECAL clusters.

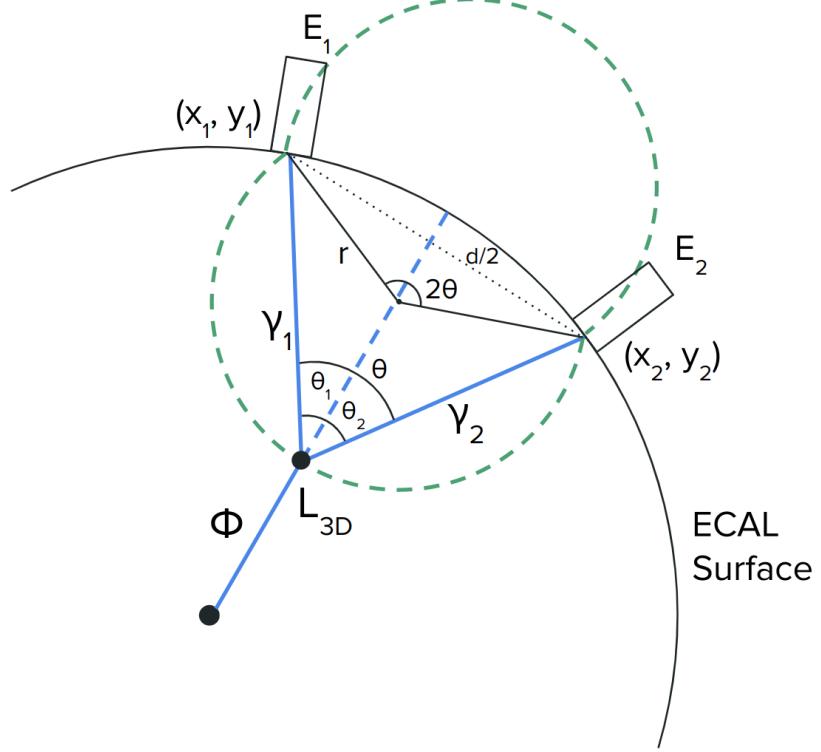


Figure 5.9: Geometrical configuration used to reconstruct the displaced vertex of the di-photon pair

Because three non-collinear points uniquely define a circle,  $\theta$  corresponds to an inscribed angle on a circle which overlaps the decay vertex and the two ECAL clusters located at  $(x_1, y_1)$  and  $(x_2, y_2)$ . The central angle of the arc between the two clusters has measure  $2\theta$  by the inscribed angle theorem. From this, the radius of the circle can be defined by

$$r = \frac{d}{2 \sin \theta} \quad (5.8)$$

where  $d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$  is the distance between the two ECAL clusters. The

center of the circle  $(x_0, y_0)$  is constrained by the two ECAL cluster coordinates and the equation of a circle  $(x - x_0)^2 + (y - y_0)^2 = r^2$ , which can be solved to give

$$x_0 = \frac{x_1 + x_2}{2} \pm \frac{y_1 - y_2}{q} \cdot \sqrt{r^2 - \frac{q^2}{4}} \quad (5.9)$$

$$y_0 = \frac{y_1 + y_2}{2} \pm \frac{x_2 - x_1}{q} \cdot \sqrt{r^2 - \frac{q^2}{4}} \quad (5.10)$$

$$\text{where } q = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (5.11)$$

The two solutions represent circles centered inside and outside of the ECAL. If  $\theta < \pi/2$ , then only the major arcs of the two circle satisfy equation 5.5. Alternatively, if  $\theta > \pi/2$  only the minor arcs are valid solutions. For either solution, an arc from one of two circles will lie outside the ECAL. This arc represents a nonphysical solution and can be discarded, as it would imply the  $\Phi$  both decayed outside of the ECAL and left ECAL clusters. The remaining arc represents the locus of possible decay vertices.

The angle between the two photons can be divided into the deflection of each photon ( $\theta_1$  and  $\theta_2$  respectively), and conservation of momentum requires  $E_1 \sin \theta_1 - E_2 \sin \theta_2 = 0$ . From here, the vertex can be calculated numerically by sampling points along the remaining arc and choosing the coordinates that satisfy this equation.

It should be noted that this method can yield a locus of points that pass behind the beamline resulting in a negative impact parameter. The vertex is still calculated in these cases in order to provide a disjoint region with respect to the signal region that will be used to validate analysis methods and estimate the background yield. A diagram showing a negative reconstructed vertex is shown in figure 5.10.

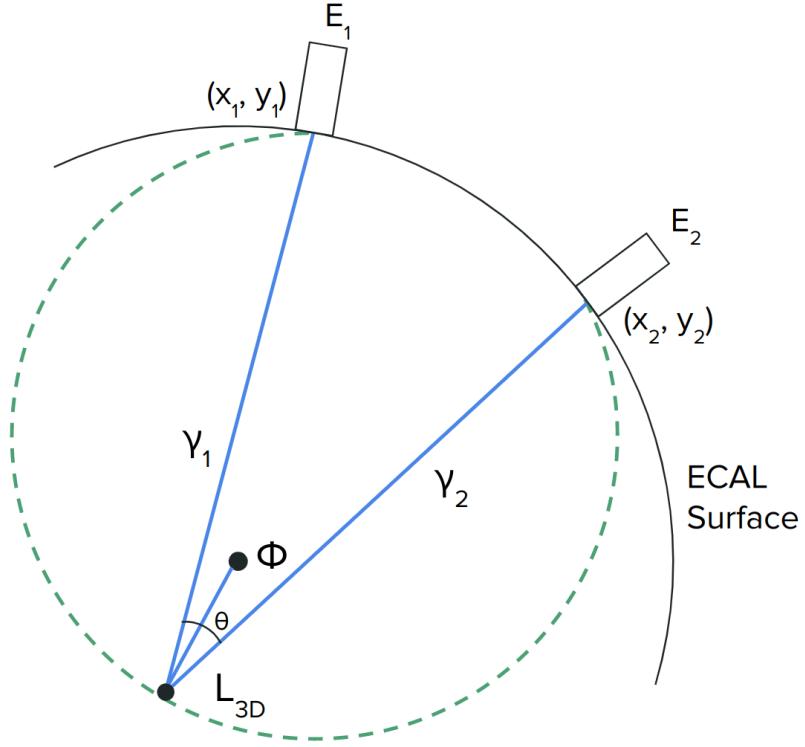


Figure 5.10: Geometrical configuration showing the reconstruction of a negative impact parameter

### 5.3.2 Performance of the Diphoton Vertex Calculation

It is evident that since the kinematic fit uses the hypothetical mass of the  $\Phi$ , the fit must be rerun for every hypothetical mass used in this analysis. As a result, the events in the signal region used for each value of  $m_\Phi$  are different. Figure 5.11a shows the reconstructed  $L_{xy}$  behaves as expected as the  $\Phi$  lifetime increases. To demonstrate the resolution of the kinematic fit, we plot the difference between the calculated  $L_{xy}$  and the generator level  $L_{xy}$  as a function of the generator level  $L_{xy}$ , performed using a  $Z \rightarrow \ell\ell$  samples with  $m_\Phi = 30 \text{ GeV}$ , using both the  $c\tau = 100 \text{ mm}$  and  $1000 \text{ mm}$  samples. These distributions can be seen in figure 5.11b. The vertical slices of this histogram were fit to Gaussian distributions so that the mean and sigma could be plotted as a function of generator level  $L_{xy}$ . The means shown in figure 5.11c show the differences in calculated and generator  $L_{xy}$  are compatible with zero,

and the sigma shown in figure 5.11d show a resolution of  $< 3$  cm, slightly decreasing with generator level  $L_{xy}$ . Both the mean and sigma remain consistent across all ranges of  $m_\Phi$ . The mean has a slight bias towards underestimating the  $L_{xy}$  due to the non-linear relation between the ECAL position/energy resolution and the vertex calculation. The sigma tends to decrease slightly as  $L_{xy}$  increases, as the  $\Phi$  decaying closer to the ECAL means the photons traverse less material before reaching the ECAL. This gives a lower probability to interact with the tracker which improves the position and energy resolution.

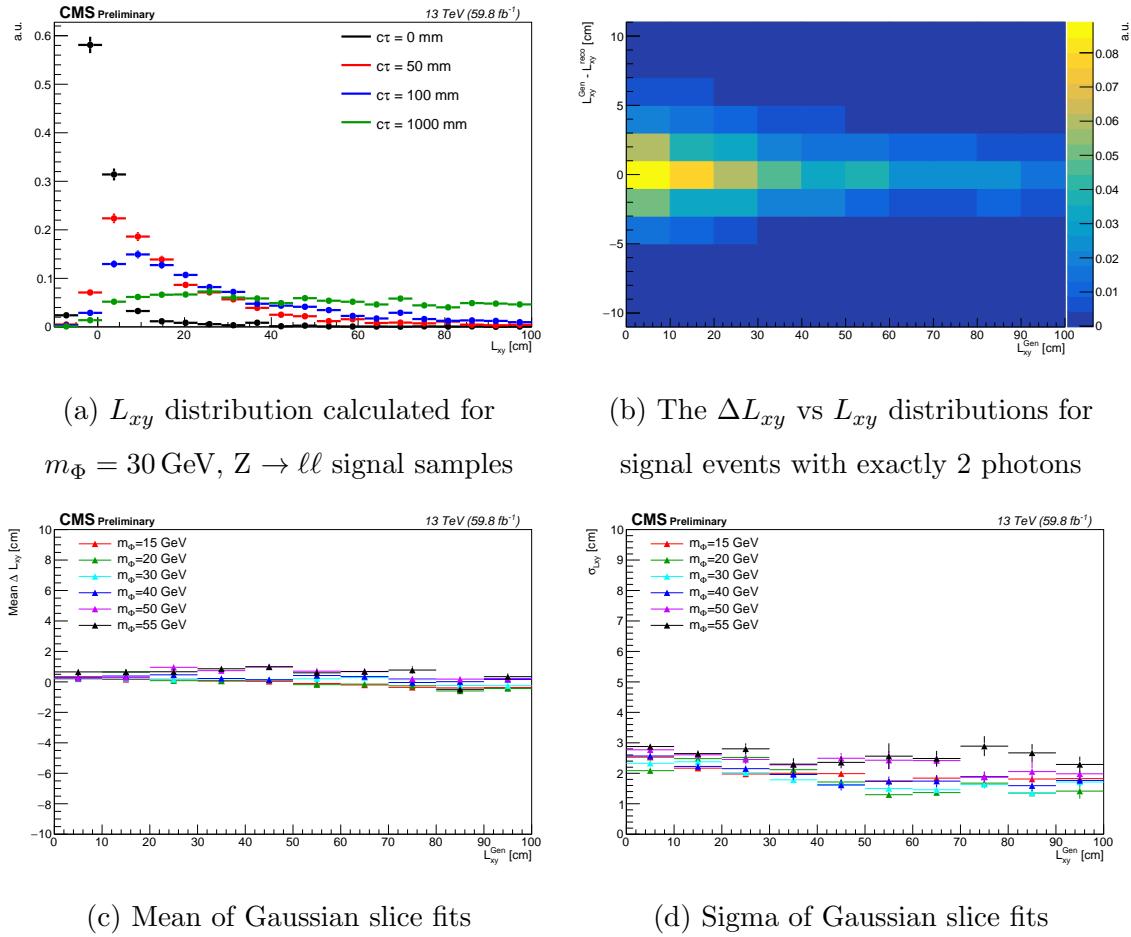


Figure 5.11: Performance of the vertex calculation on signal samples. The fit shows good agreement across all values of generator level  $L_{xy}$  and a resolution of  $< 3$  cm.

It can be shown that the calculated  $L_{xy}$  for a fixed value of prompt invariant photon mass

$m_{\gamma\gamma}$  is monotonically decreasing with the assumed value of  $m_\Phi$ . Assume that two photons have a prompt (four momenta calculated assuming the two clusters originated from the PV) invariant mass  $m_{\gamma\gamma}$ . Calculating the vertex assuming  $m_\Phi = m_{\gamma\gamma}$  would yield a vertex with  $L_{xy} = 0$  by construction. Calculating the vertex assuming  $m_\Phi > m_{\gamma\gamma}$  would cause the angle between the two photons given by equation 5.5 to decrease. As the ECAL cluster positions are fixed, this pushes the vertex backwards yielding  $L_{xy} < 0$ . By similar methods, assuming a value of  $m_\Phi < m_{\gamma\gamma}$  yields a vertex with  $L_{xy} > 0$ . The dependence makes the  $L_{xy}$  very sensitive to small changes in the assumed value of  $m_\Phi$ . If the assumed value of  $m_\Phi$  differs from the true mass by 5 GeV, the calculated  $L_{xy}$  can skew upwards or downwards by > 20 cm, depending on the value of the true mass. Figure 5.12 shows the differences in generated and reconstructed  $L_{xy}$ , combining samples for all lifetimes for a given true mass.

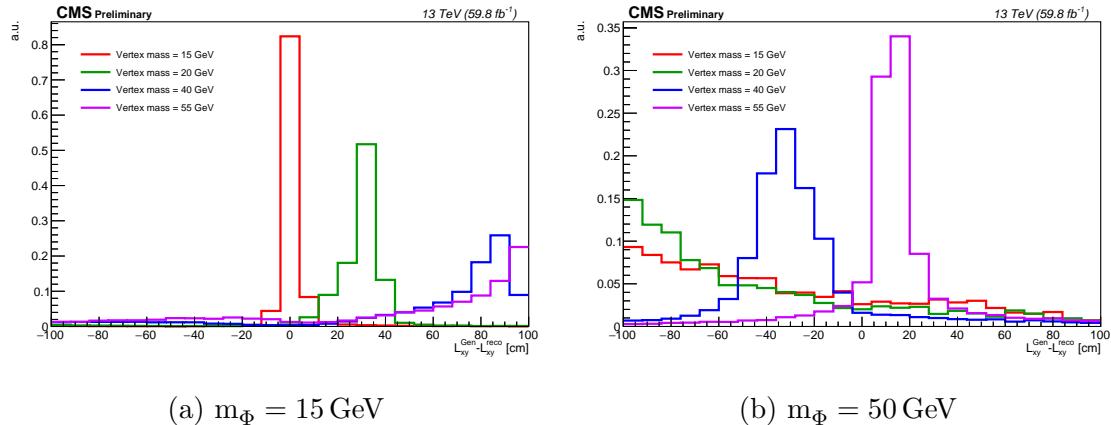


Figure 5.12: The difference between the true and reconstructed vertex grows as the assumed mass departs from the true one. Gen-reconstructed  $L_{xy}$  for a  $Z + H$  samples under several assumed mass points.

This analysis uses the nanoAODv9 data format - the most compact data tier available for analysis use in CMS. As opposed to lower tier formats which contain detailed collections of classes to store event properties, nanoAOD provides only the most important information for high level physics objects [44]. This information includes several properties of photons but omits the calorimeter cluster position, which is required to calculate the diphoton vertex. To

circumvent this, we propagate the photons from the PV to the ECAL using their  $\eta$  and  $\phi$ . The ECAL is approximated using a cylinder with radius  $\rho = 137$  cm and  $z = 324$  cm. These values were obtained using `miniAODv2` samples and are shown in figure 5.13. Diphoton vertices calculated using the propagated coordinates yield a  $L_{xy}$  within 3% of the  $L_{xy}$  calculations using the exact cluster coordinates, which is well within the intrinsic resolution of the vertex calculation.

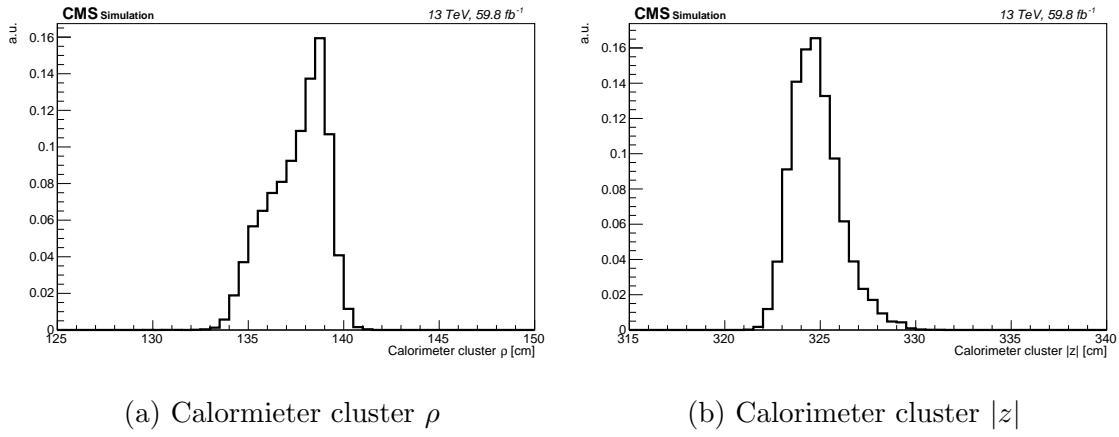


Figure 5.13: The  $\rho$  for barrel photons and  $|z|$  for endcap photons for all reconstructed photons using `miniAODv2` signal samples

## 5.4 Event Selection

The first step in collecting events for the analysis is to select events passing HLT lepton triggers. Events passing these triggers are then chosen based on the presence of two same flavor, opposite charge leptons and two photons. Various quality cuts, referred to as pre-selection criteria, are applied to these objects to reject background. Lastly, event level cuts are applied to reject photons from final state radiation (FSR) and low  $p_T$  photons that get promoted to displaced  $\Phi$  candidates by the kinematic vertex fit.

### 5.4.1 Triggers

Events are triggered through leptonic decays of the Z boson produced in association with the SM Higgs Boson. The Z candidates are selected from dilepton events passing the HLT isolated single lepton triggers. These triggers are the standard lepton triggers recommended both the EGamma and Muon physics object groups (POGs) in CMS. Double lepton and high  $p_T$  lepton triggers were considered but provide only negligible improvement when applied in addition to the single lepton triggers. To ensure events are not double counted between data sets, muon events must pass the single muon triggers and are vetoed by the single electron triggers, while electron events must pass single electron triggers and are vetoed by the single muon triggers. A complete list of HLT trigger paths used for each year can be seen in table 5.6.

Table 5.6: HLT trigger paths used for the Run 2 datasets

Year	Single Electron	Single Muon
2018	HLT_Ele32_WPTight_Gsf	HLT_IsoMu24
2017	HLT_Ele32_WPTight_Gsf	HLT_IsoMu27
2016	HLT_Ele27_WPTight_Gsf	HLT_IsoMu24    HLT_IsoTkMu24

Each HLT path name corresponds to the different criteria required to pass the trigger. The electron triggers require the trigger object have  $p_T > 32 \text{ GeV}$  ( $27 \text{ GeV}$  for 2016), pass tight selection criteria, and be reconstructed using the Gaussian sum filter (GSF) algorithm discussed in section 3.2.8.3. Muons are required to pass isolation and have  $p_T > 24 \text{ GeV}$  ( $27 \text{ GeV}$  for 2017). In order to improve the efficiency of the muon trigger in 2016, we also include the tracker muon trigger, which requires an isolated tracker muon with  $p_T > 24 \text{ GeV}$ .

## 5.4.2 Preselection Criteria

For both electrons and muons we define a set of tight and loose leptons based on cut based ID set by the EGamma and Muon POGs. To reconstruct a Z candidate, we require exactly two same flavor opposite charge tight leptons with invariant mass  $70 < m_{\ell\ell} < 110$  GeV and zero additional loose leptons. For photons we require at least two photons passing a set of preselection criteria, which are subject to additional classification based on cut based ID.

### 5.4.2.1 Particle Flow Isolation

One key application of the particle flow (PF) algorithm is the calculation of a particle's isolation. This quantity measures the  $p_T$  and energy contributions from PF candidates in a fixed  $\Delta R$  cone centered on the original particle. A lower value implies the particle is more isolated (i.e. there are fewer additional particles nearby), and vice versa. Isolation is a key variable to reject particles produced in jets from heavy-flavor hadronic decays or decays of pions and kaons [31]. A particle's isolation is defined as follows:

$$I_{\text{pf}} = \sum_{i \in \text{charged hadrons}} p_T^i + \max \left( 0, \sum_{i \in \text{neutral hadrons}} E_T^i + \sum_{\gamma \in \text{photons}} E_T^\gamma - \text{PU Corrections} \right) \quad (5.12)$$

The three contributions that are summed over in equation 5.12 are referred to as charged hadron isolation ( $I_{\text{ch}}$ ), neutral hadron isolation ( $I_{\text{n}}$ ), and photon isolation ( $I_\gamma$ ). It is also common to refer to a particle's isolation normalized by its  $p_T$  (or  $E_T$  for neutral particles), known as relative isolation which is defined as

$$I_{\text{rel}} = \frac{I_{\text{pf}}}{p_T} \quad (5.13)$$

Particles created from pileup can create calorimeter deposits within the  $\Delta R$  cone used to calculate isolation. These are unrelated to the physics process for the central particle and

can artificially inflate the isolation. Only charged hadrons with tracks originating from the PV are used when calculating the charged hadron isolation, which rejects charged hadrons originating from PU. Neutral particles, however, do not produce tracks, meaning calorimeter deposits from the PV and from pileup are indistinguishable. Because of this, there are additional corrections to remove the PU contributions from neutral hadron and photon isolation. Two common methods are  $\rho$  corrections, which calculate the average energy density per unit area from pileup ( $\rho$ ) multiplied by an effective area as a function of  $\eta$ , and  $\Delta\beta$  corrections, which calculate the energy from photons and neutral hadrons in pileup as a fraction of the charged hadron contribution from PU.

#### 5.4.2.2 Muon Criteria

The NanoAOD data format takes reconstructed muons and applies loose cuts before storing the muon properties. These branches contain all muons with  $p_T > 3 \text{ GeV}$  with minimal cuts to reconstruction quality. For the purpose of this analysis, muons are classified as loose if they pass all of the following criteria:

- $p_T > 10 \text{ GeV}$
- $|\eta| < 2.4$
- $D_{xy} < 0.2 \text{ cm}$ : the transverse distance from the PV to the beamline must be less than 0.2 cm
- $D_z < 0.5 \text{ cm}$ : the longitudinal distance from the PV to the beamline must be less than 0.5 cm
- $I_{\text{rel}}(\Delta R < 0.4, \Delta\beta\text{-corrected}) < 0.25$
- Passes loose quality criteria set by the Muon POG:
  - Muon must be reconstructed by the PF algorithm

- Muon must be either a global and/or tracker muon

We define tight muons as a subset of loose muons if they pass the following criteria:

- $I_{\text{rel}}(\Delta R < 0.4, \Delta\beta\text{-corrected}) < 0.15$
- Passes tight quality criteria set by the Muon POG:
  - Muon must be a global muon
  - The global muon track fit must have  $\chi^2/\text{ndof} < 10$
  - Global track must have at least one hit in a muon chamber
  - Muon track must have segments in at least 2 different muon stations
  - Tracker track  $D_{xy} < 2 \text{ mm}$
  - Tracker track  $D_z < 5 \text{ mm}$
  - Tracker track must have at least 1 hit in the pixel detector
  - Tracker track must have hits in more than 5 tracker layers

We also require that one of the two tight muons used to reconstruct the Z be the triggering muon, and that the triggering muon have  $p_T > 25 \text{ GeV}$  ( $28 \text{ GeV}$  in 2017) in order to operate above the single muon trigger  $p_T$  thresholds.

#### 5.4.2.3 Electron Criteria

Analogously to the muon criteria, loose electrons are defined from reconstructed electrons if they pass a set of preselection criteria.

- $p_T > 15 \text{ GeV}$
- $|\eta| < 2.4$

- $|\eta| > 1.57$  or  $|\eta| < 1.44$  to rejects clusters in the overlap region between the ECAL barrel and endcap
- $D_{xy} < 0.2 \text{ cm}$ ,  $D_z < 0.2 \text{ cm}$
- Passes cut based ID provided by the EGamma POG at the `veto` working point

Tight electrons are selected from this collection of they pass cut based ID at the `tight` working point set by the EGamma POG. A full set of criteria is shown in table 5.7 for barrel and endcap electrons. Similar to the muon criteria, we require one of the tight electrons be the triggering electron, and that the triggering electron have  $p_T > 35 \text{ GeV}$ .

Table 5.7: Cut based ID selection criteria for electrons in Run-2 [21].

(a) Barrel Electrons

Barrel Criteria	Veto	Tight
$\sigma_{inj\eta} <$	0.0126	0.0104
$ \Delta\eta_{\text{seed}}  <$	0.0463	0.00255
$ \Delta\phi_{\text{In}}  <$	0.148	0.022
$H/E <$	$0.05 + 1.16/E + 0.0324\rho/E$	$0.026 + 1.15/E + 0.0324\rho/E$
$I_{\text{rel}}(\Delta R < 0.3, \rho \text{ corrected}) <$	$0.198 + 0.506/p_T$	$0.0287 + 0.506/p_T$
$ 1/E - 1/p  <$	0.209	0.159
Missing inner hits $\leq$	2	1
Passes conversion veto	True	True

(b) Endcap Electrons

Endcap Criteria	Veto	Tight
$\sigma_{inj\eta} <$	0.0457	0.0353
$ \Delta\eta_{\text{seed}}  <$	0.00814	0.00501
$ \Delta\phi_{\text{In}}  <$	0.19	0.0236
$H/E <$	$0.05 + 2.54/E + 0.183\rho/E$	$0.0188 + 2.06/E + 0.183\rho/E$
$I_{\text{rel}}(\Delta R < 0.3, \rho \text{ corrected}) <$	$0.203 + 0.963/p_T$	$0.0445 + 0.963/p_T$
$ 1/E - 1/p  <$	0.132	0.0197
Missing inner hits $\leq$	3	1
Passes conversion veto	True	True

The variables used in the cut based ID are optimized to reject background such as photons in hadronic showers or photons that convert to  $e^+e^-$  pairs in the tracker. The spread of the

calorimeter cluster in  $\eta$  is measured using  $\sigma_{inj\eta}$ , which is the weighted second moment of the cluster shape in  $\eta$  measured in a  $5 \times 5$  grid of calorimeter crystals, centered on the most energetic crystal in the SC. This variable can be used to discriminate two photon showers produced from neutral meson decays, which have higher spread in  $\eta$ , from showers produced by prompt electrons or photons. The  $H/E$  ratio takes the energy of HCAL clusters located in a  $\Delta R < 0.15$  cone behind the ECAL cluster and divides it by the ECAL cluster energy. Genuine EM showers have three main contributions to hadronic energy: leakage through gaps into the HCAL, calorimeter noise, and pileup. The constant term in these cuts accounts for energy leakage from real electrons or photons, the term proportional to  $1/E$  is to account for noise in the HCAL, and the  $\rho/E$  term is the correction due to pileup. Tracker information is also used to separate signal from background. The  $|\Delta\eta_{seed}|$  and  $|\Delta\phi_{in}|$  are the difference in SC and track  $\eta/\phi$ , measured from the innermost tracker layer, which is expected to be small for prompt electrons. The difference between tracker momentum and calorimeter energy is measured using the quantity  $|1/E - 1/p|$ , which should be small as a good electron should have similar energy and momentum. The number of missing hits refers to number of gaps in the electron's trajectory through the tracker. Converted photons will leave tracks after the conversion vertex but leave no hits in the tracker before they convert, while prompt electrons should leave hits in the tracker starting from the PV. An additional conversion veto is applied by rejecting electrons with tracks matched to conversion vertices in order to further reject background from converted photons. The tight cut based ID is expected to have a 70% signal efficiency for real electrons.

#### 5.4.2.4 Photon Criteria

Reconstructed photons must pass baseline selection to be included in the nanoAOD data format. All nanoAOD photons must have  $p_T > 5 \text{ GeV}$ ,  $R_9 > 0.8$ , and pass cuts on total and relative charged hadron isolation given by  $I_{ch} < 20 \text{ GeV}$  and  $I_{ch}/p_T < 0.3$ . The  $R_9$  of a supercluster is defined as the energy the most energetic  $3 \times 3$  crystal grid in the supercluster

divided by the total supercluster energy. For an event to be considered for this analysis, we require at least two nanoAOD photons passing the following preselection criteria:

- $p_T > 20 \text{ GeV}$
- Passes pixel seed veto
- $|\eta| < 2.4$
- $|\eta| > 1.57$  or  $|\eta| < 1.44$
- $\Delta\eta(\gamma, \text{lepton})^2/0.4^2 + \Delta\varphi(\gamma, \text{lepton})^2/0.5^2 > 1$  for all loose leptons

Due to the production mechanism of the Higgs boson, the  $\Phi$  and subsequent photon pair are expected to be strongly correlated and have high energy. The following kinematic requirements based on the four-momentum of the diphoton system reject background due to uncorrelated or lower energy photon pairs while maintaining 99% signal efficiency:

- $p_T(\gamma, \gamma) > 20 \text{ GeV}$
- Invariant diphoton mass  $m_{\gamma\gamma} > 4 \text{ GeV}$

The pixel seed veto reject photons if there are at least two pixel seed hits that point to the supercluster, which is used to reject background from electrons that are misidentified as photons. Photons passing this preselection criteria are subject to further classification as ID'd or anti-ID'd if they pass or fail loose cut based ID outlined in table 5.8. The loose ID is measured to have approximately 90% efficiency for prompt photons, but was observed to have additional inefficiencies due to photon isolation and shower shape variables resulting from the displaced signature of the two photons. As the photon vertex becomes more displaced, the shower shape becomes more skewed and the two photons become closer together, which affects the cuts to  $\sigma_{i\eta i\eta}$  and  $I_\gamma$ . At lower masses, the photon isolation is the main source of inefficiency because the photons are more collimated, while at higher masses the  $\sigma_{i\eta i\eta}$

is the main inefficiency as the larger angle between the photons causes a larger spread in shower shape. The effect of each criteria in the cut based ID on the photon reconstruction efficiency is shown as a function of generator level  $L_{xy}$  in figure 5.14 for all values of  $m_\Phi$ . The efficiency denominator is all generator level (gen) photon pairs, where each gen photon has  $p_T > 20 \text{ GeV}$  and a decay vertex within the CMS ECAL. The numerator requires each gen photon to have a matching reconstructed photon that passes given selection criteria. Matching requires the reconstructed photon have  $\Delta R < 0.3$  and energy within 50% of the gen photon.

Given the large inefficiency of the  $I_\gamma$  cuts at low values of  $m_\Phi$ , studies were performed using modified photon isolation by removing the footprint of the partner photon from  $I_\gamma$ . This modified ID showed up to 50% improvement in signal efficiency for low  $m_\Phi$ , high  $L_{xy}$  events. However, this was shown to introduce substantial irreducible background with only marginal improvements to expected limits, so it was not used in the final analysis.

Table 5.8: Cut based ID selection criteria for photons in Run-2 [22].

Barrel Criteria	Loose ID
$H/E <$	0.04596
$\sigma_{in\eta} <$	0.0106
$I_{ch}(\Delta R < 0.3, \rho \text{ corrected}) <$	1.694
$I_n(\Delta R < 0.3, \rho \text{ corrected}) <$	$24.032 + 0.01512 * p_T + 2.259 * 10^{-5} * p_T^2$
$I_\gamma(\Delta R < 0.3, \rho \text{ corrected}) <$	$2.876 + 0.004017 * p_T$

Endcap Criteria	Loose ID
$H/E <$	0.0590
$\sigma_{in\eta} <$	0.0272
$I_{ch}(\Delta R < 0.3, \rho \text{ corrected}) <$	2.089
$I_n(\Delta R < 0.3, \rho \text{ corrected}) <$	$19.722 + 0.01512 * p_T + 2.3 * 10^{-5} * p_T^2$
$I_\gamma(\Delta R < 0.3, \rho \text{ corrected}) <$	$4.162 + 0.0037 * p_T$

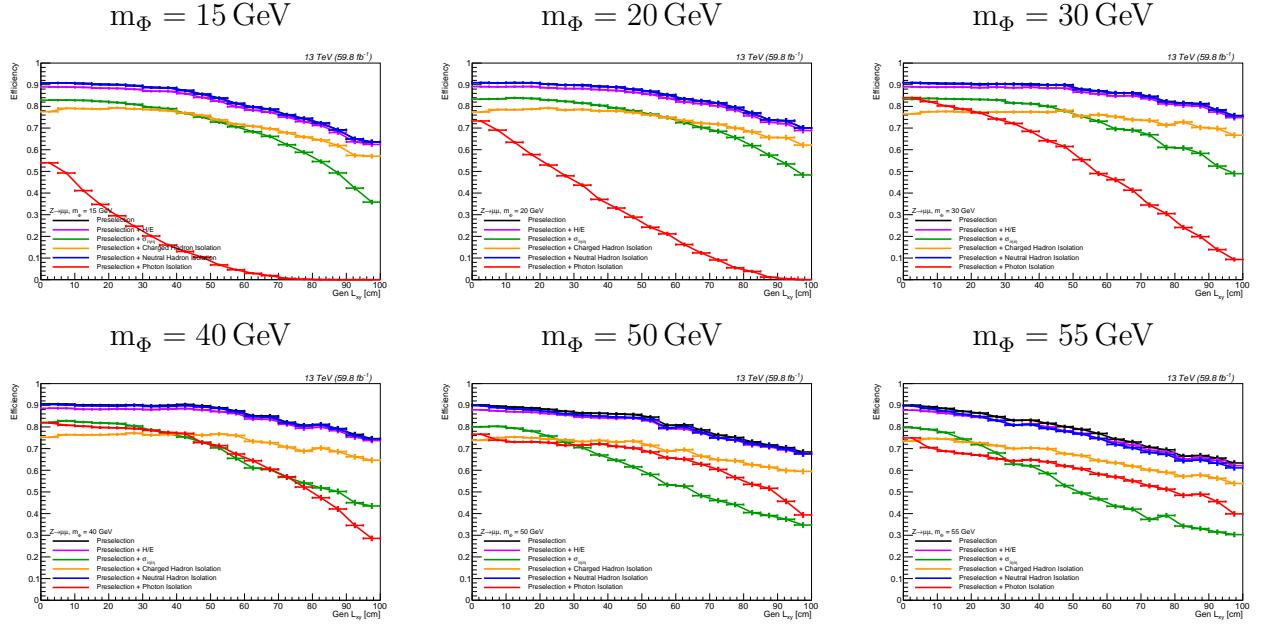


Figure 5.14: Efficiency for a photon pair to pass preselection versus generator level  $L_{xy}$  on ZH signal samples for all lifetimes in the  $Z \rightarrow \mu\mu$  category. Denominator is all generator level photon pairs with each photon having  $p_T > 20$  GeV that decay within the geometric acceptance of the ECAL. The numerator requires two matching ( $\Delta r < 0.3$ , energy within 50% of gen photon) reconstructed photons passing preselection criteria. Cut based ID is applied on top of preselection, showing that the dominant inefficiencies at high  $L_{xy}$  are due to  $I_\gamma$  and  $\sigma_{inj\eta}$ .

### 5.4.3 Signal Region Definitions

$\Phi$  candidates are reconstructed from events with at least two photons passing preselection criteria. If there are only two photons that pass preselection, the diphotons are chosen as a possible  $\Phi$  candidate. If more than two photons pass this criteria, the photons will be subject to pairing criteria for every candidate mass of the  $\Phi$ . This pairing process is designed to maximize pairing efficiency, as without pointing capabilities there is no consistent, robust way to determine whether signal photons originated from the same  $\Phi$ , while also minimizing the bias introduced through the pairing criteria. For each candidate mass, photon pairs are

chosen as follows:

- The pair with an energetically allowed decay is chosen (i.e. there is a physical solution to equation 5.5)
- If multiple pairs have allowed decays, the pair with more photons passing cut based ID is chosen
- If multiple pairs use the same number of photons passing ID, the pair with the largest prompt invariant mass  $m_{\gamma\gamma}$  is chosen
- If multiple pairs have the same  $m_{\gamma\gamma}$ , the pair with the highest  $p_T(\gamma\gamma)$  is chosen

For convention, the leading higher  $p_T$  photon is referred to as  $\gamma_1$  while the sub-leading lower  $p_T$  photon is referred to as  $\gamma_2$ .

The signal region is defined by the presence of one Z candidate and at least one  $\Phi$  candidate composed of two ID'd photons. Events are categorized based on the lepton flavor of the Z decay to either muons or electrons. For each signal region category, an anti-ID control region is utilized for background estimation. These categories are defined identically to the signal region but require the  $\Phi$  candidate to use one photon passing the loose cut based ID and one photon failing loose cut based ID defined in table 5.8.

To compare data to simulated background, the dilepton mass  $m_{\ell\ell}$  is plotted in all categories for the signal region and anti-ID control region in figures 5.15 - 5.16. These distributions are made agnostic to the vertex calculation and have no cuts on diphoton  $L_{xy}$ . The results show moderate agreement in the anti-ID control region and limited statistical power for simulated events in the ID signal region. Comparisons show an excess of events below 90 GeV due to the presence of final state radiation.

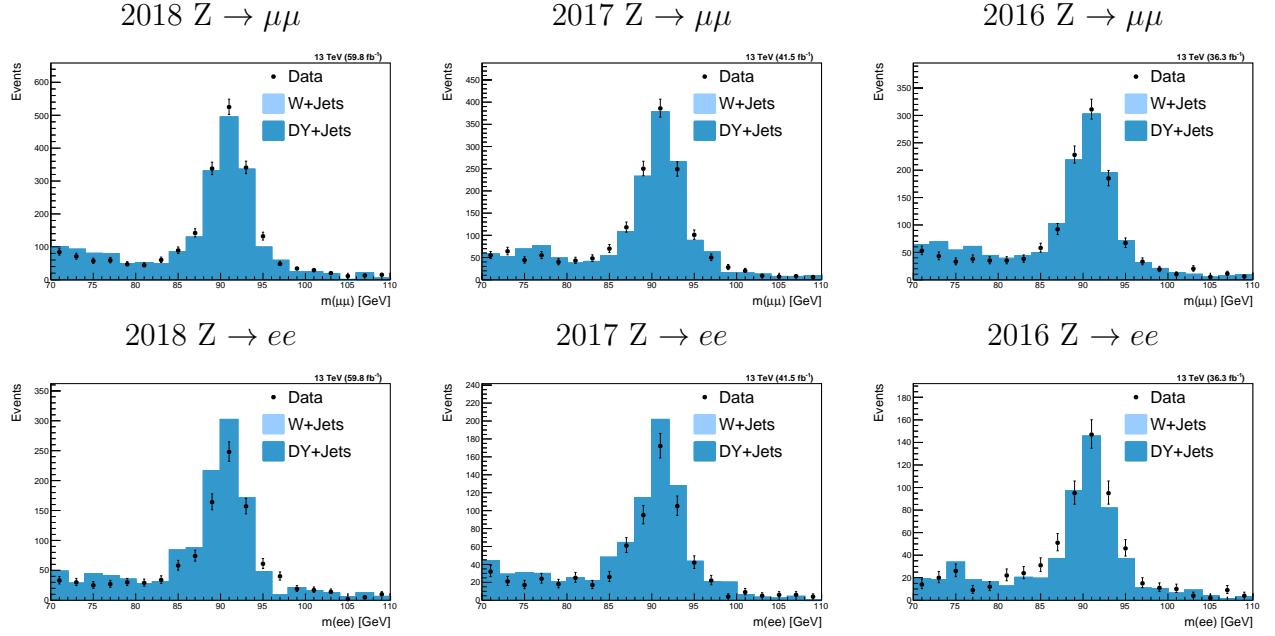


Figure 5.15:  $m(\ell^+\ell^-)$  data and simulated background after preselection criteria across all years in the anti-ID control region. Top row:  $Z \rightarrow \mu\mu$  category. Bottom row:  $Z \rightarrow ee$  category

#### 5.4.3.1 Positive and Negative $L_{xy}$ Regions

As discussed in section 5.3, it is possible for the vertex calculating to yield a negative value of  $L_{xy}$ , indicating that the  $\Phi$  would have decayed from behind the beamline relative to the ECAL clusters. The negative  $L_{xy}$  region is used for background estimation, while the positive  $L_{xy}$  region is used for the final signal extraction. To avoid possible signal contamination in the negative  $L_{xy}$  region, we define the two regions using the prompt invariant diphoton mass  $m_{\gamma\gamma}$ . This provides a robust cut that is not dependent on the vertex calculation.

By construction of the vertex constraint, if two photons have a prompt invariant mass  $m_{\gamma\gamma}$  then the vertex calculated assuming  $m_\Phi = m_{\gamma\gamma}$  would yield an  $L_{xy} = 0$ . As a corollary, the prompt invariant mass of a pair of displaced photons decaying from a particle of mass  $m_\Phi$  will always have a prompt invariant mass  $m_{\gamma\gamma} \leq m_\Phi$ . Since the mass of the  $\Phi$  is constrained by a maximum of  $m_H/2$ , the highest prompt invariant mass for potential signal events is

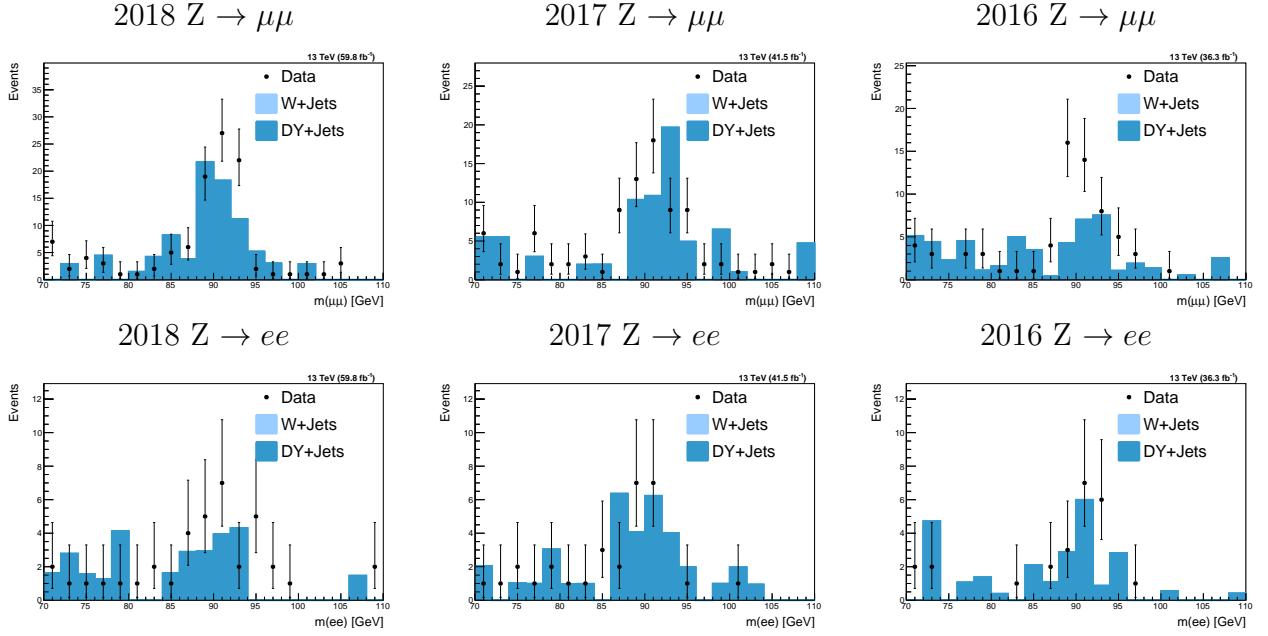


Figure 5.16:  $m(\ell^+\ell^-)$  data and simulated background after preselection criteria across all years in the ID signal region. Top row:  $Z \rightarrow \mu\mu$  category. Bottom row:  $Z \rightarrow ee$  category

$m_{\gamma\gamma} = m_H/2$ . Thus cutting on  $m_{\gamma\gamma} > m_H/2$  gives a set of events that yield a negative  $L_{xy}$  for all hypothetical  $m_\Phi$  that is disjoint from any possible signal events.

The positive  $L_{xy}$  signal region is defined by  $m_{\gamma\gamma} < m_\Phi + 5 \text{ GeV}$ . This allows slightly negative  $L_{xy}$  events into the signal region in order to account for the vertex calculation resolution for low displacement signal events.

#### 5.4.3.2 Final State Radiation Rejection

One source of background occurs when a lepton in the final state radiates a photon, known as final state radiation (FSR). This results in the band of excess events in  $m_{\ell\ell}$  seen in figures 5.15 - 5.16. FSR events are identified by calculating the invariant mass of the dileptons plus one or both of the photons used to reconstruct the  $\Phi$  candidate. If the combined invariant mass is within  $75 - 105 \text{ GeV}$  and closer to the  $Z$  boson mass of  $91 \text{ GeV}$  than the original dilepton mass, the event is flagged as having FSR and removed. FSR events are

classified based on which combination of photons yields a mass within the Z window, with the closest mass to 91 GeV being chosen if multiple satisfy the criteria. The categories are defined as follows:

- Tag 1:  $\ell\ell + \gamma_1$
- Tag 2:  $\ell\ell + \gamma_2$
- Tag 3:  $\ell\ell + \gamma_1 + \gamma_2$

Figure 5.17 show the comparison of invariant mass distributions for all three tags in the  $Z \rightarrow \mu\mu$  category. Simulated background uses the anti-ID control region due to poor statistical power in the ID signal region. Simulated signal uses an  $m_\Phi = 20$  GeV,  $c\tau = 0$  mm signal sample, but the distributions remain stable across all masses and lifetimes. It is shown that the largest background contribution comes from tag 2 followed by tag 1. Tag 3 shows minimal contribution, as the energy loss from double FSR photons usually causes the dilepton invariant mass to fall below 70 GeV.

In the anti-ID control region, this method rejects 16% of simulated background in the  $Z \rightarrow \mu\mu$  category and 19% in the  $Z \rightarrow ee$  category. In the ID signal region, this method rejects 9% of simulated background in the  $Z \rightarrow \mu\mu$  category and 35% in the  $Z \rightarrow ee$  category. Across all signal masses and lifetimes this method maintains  $> 99.5\%$  signal efficiency. The data and simulated background distributions after rejecting FSR events are shown in figures 5.18 - 5.19.

#### 5.4.3.3 Low $p_T$ Photon Background

Equation 5.5 shows that, for energetically allowed decays, the angle between two photons increases monotonically as the assumed value of  $m_\Phi$  increases. The position of the ECAL clusters is fixed, so the vertex becomes more displaced as the angle increases, leading to low  $p_T$  photon background with large  $L_{xy}$ . This makes the photon  $p_T$  a useful handle to

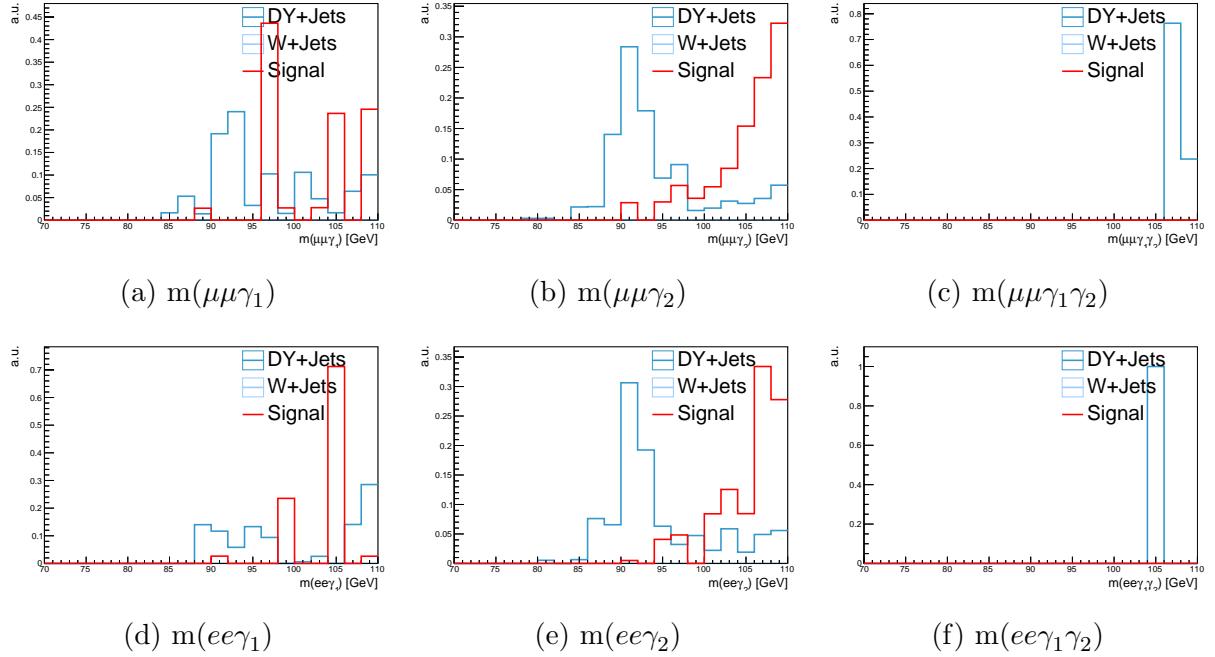


Figure 5.17: Comparison of invariant mass distributions for FSR rejection using simulated  $m_\Phi = 20 \text{ GeV}$ ,  $c\tau = 0 \text{ mm}$  signal and background. Background uses normalized anti-ID control region events for improved statistics while signal shows ID signal region events.

discriminate signal and background. We compare the  $p_T$  spectrum of leading and sub-leading photons in the anti-ID control region and signal region for events having a positive  $L_{xy}$  in order to determine optimal cut values. As one additional requirement, when comparing the leading  $p_T$  spectrum we select background events where  $\gamma_1$  is ID'd and  $\gamma_2$  is anti-ID'd (and vice versa when plotting the subleading photon). Plots comparing signal and background  $p_T$  distributions for leading and subleading photons can be seen in figure 5.20.

The final cuts were chosen to be  $p_T(\gamma_1) > 35 \text{ GeV}$  and  $p_T(\gamma_2) > 25 \text{ GeV}$ . These cuts reject between 70-80% of background and maintain about 70% signal efficiency. The exact values for background rejection can be seen in table 5.9, while the cut efficiency for signal can be seen in figure 5.23.

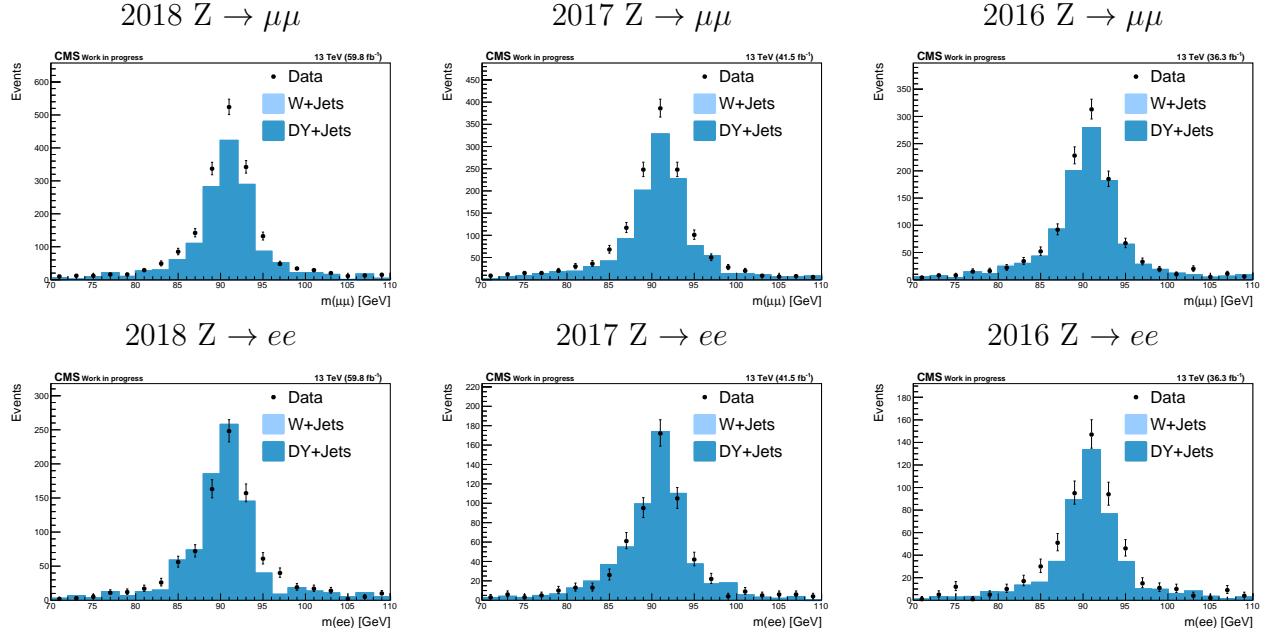


Figure 5.18:  $m(\ell^+\ell^-)$  data and simulated background after rejecting FSR events across all years in the anti-ID control region. Top row:  $Z \rightarrow \mu\mu$  category. Bottom row:  $Z \rightarrow ee$  category

Table 5.9: Photon  $p_T$  cut efficiency on control region background

$m_\Phi$ [GeV]	$Z \rightarrow \mu\mu$	$Z \rightarrow ee$
15	.24	.23
20	.27	.33
30	.27	.26
40	.29	.28
50	.29	.28
55	.30	.30

#### 5.4.3.4 Final Selection Requirements

The final  $L_{xy}$  agnostic data and Monte Carlo comparison plots after applying the FSR and tight photon  $p_T$  cuts can be seen in figures 5.21 - 5.22. Signal Monte Carlo distributions from the  $m_\Phi = 30$  GeV,  $c\tau = 0$  mm samples are overlaid at arbitrary normalization for shape

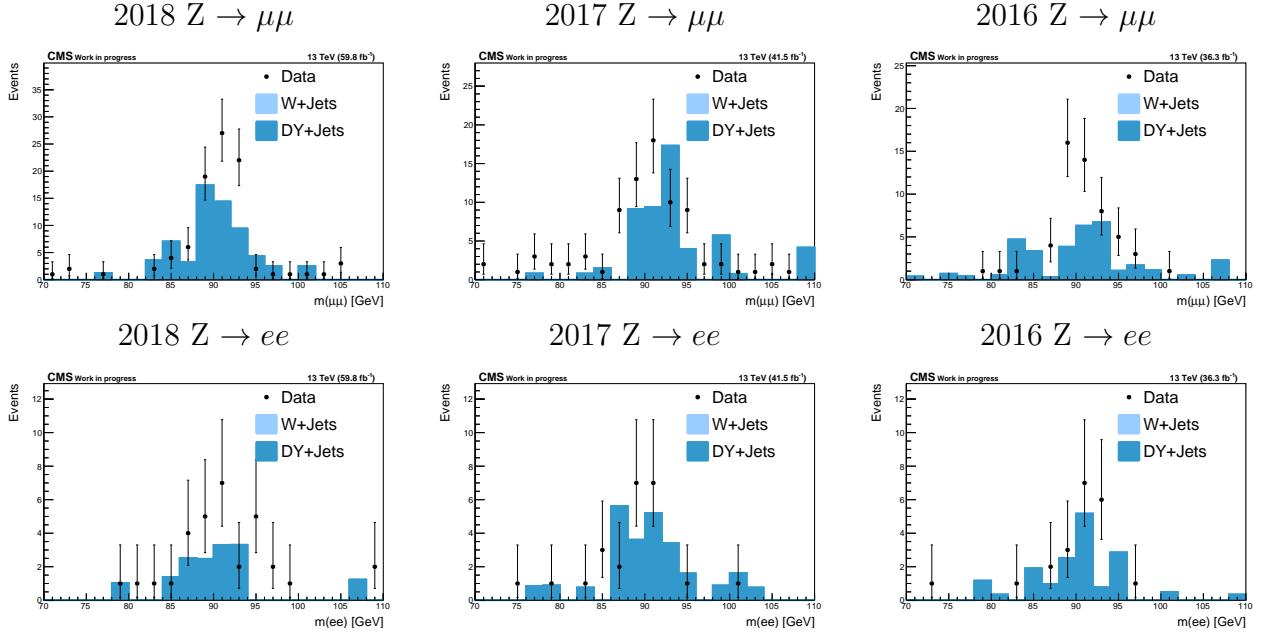


Figure 5.19:  $m(\ell^+\ell^-)$  data and simulated background after rejecting FSR events across all years in the ID signal region. Top row:  $Z \rightarrow \mu\mu$  category. Bottom row:  $Z \rightarrow ee$  category

comparison. In the ID signal region, the Monte Carlo background shows poor agreement with data and consists of very few weighted events with limited statistical power. For this reason we opt for a data driven method for background estimation.

The efficiencies and expected yields for simulated background are shown in table 5.10. For signal samples we plot the acceptances as an efficiency relative to the total number of generated events for all masses and lifetimes, separating two photon events where only one  $\Phi$  decays to photons and four final state photon events where both  $\Phi$  decay to photons. The HLT and object preselection cuts are applied as described in sections 5.4.1 and 5.4.2. The FSR rejection cuts are applied, but due to the high signal efficiency the points overlap with the previous cuts on loose cut based ID. The positive  $L_{xy}$  cuts result in a decrease in efficiency in the four photon categories due to mispaired photons, where one photon from each  $\Phi$  is chosen to calculate the vertex. The full set of signal efficiency plots are shown in figure 5.23.

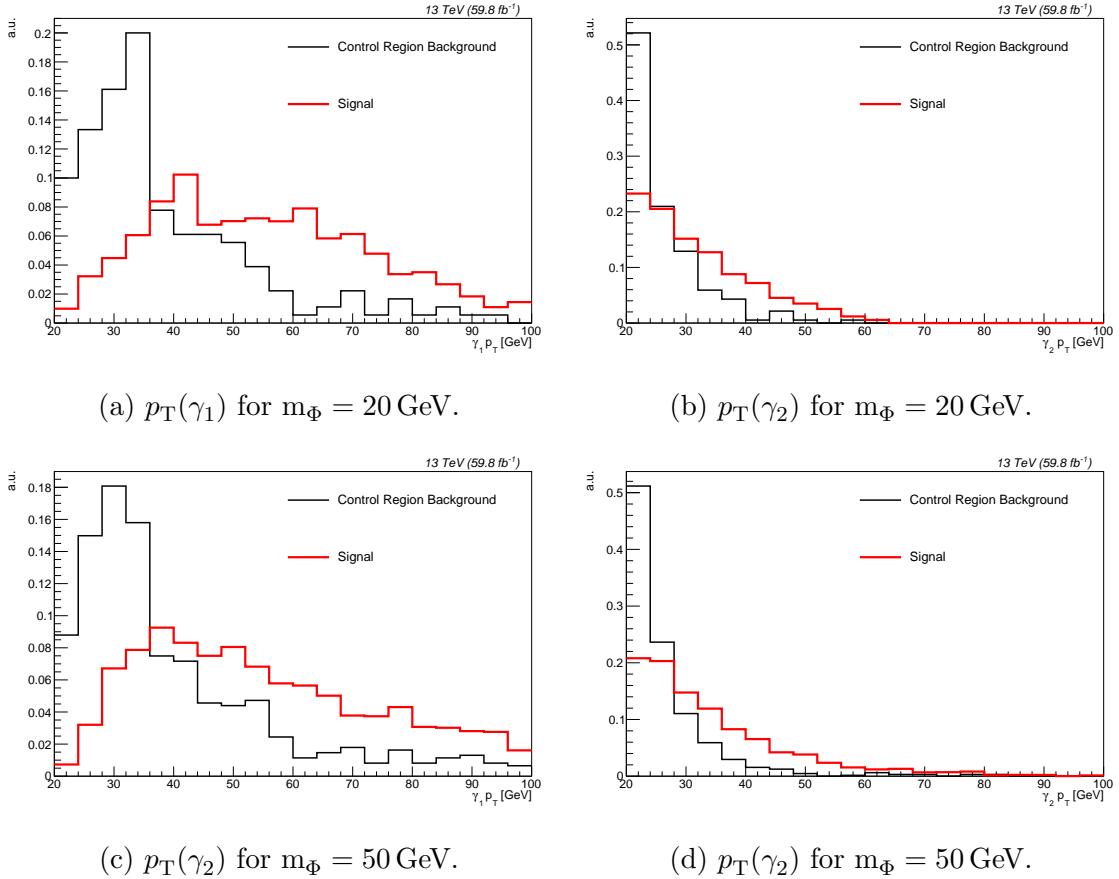


Figure 5.20: Comparison of leading and subleading photon  $p_T$  in the anti-ID control region and signal region for the  $Z \rightarrow \mu\mu$  category under the assumptions  $m_\Phi = 20 \text{ GeV}$  (top) and  $m_\Phi = 50 \text{ GeV}$  (bottom).

Table 5.10: Cutflow table and expected yields on background Monte Carlo in the ID+ID signal region

Selection	Simulated Events	
	$Z \rightarrow \mu\mu$	$Z \rightarrow ee$
Initial (from preselection)	67.4	22.1
Final State Radiation	64.7	13.6
$p_T(\gamma_1) > 35 \text{ GeV}, p_T(\gamma_2) > 25 \text{ GeV}$	28.0	5.5
Total Background Rejection	59%	75%

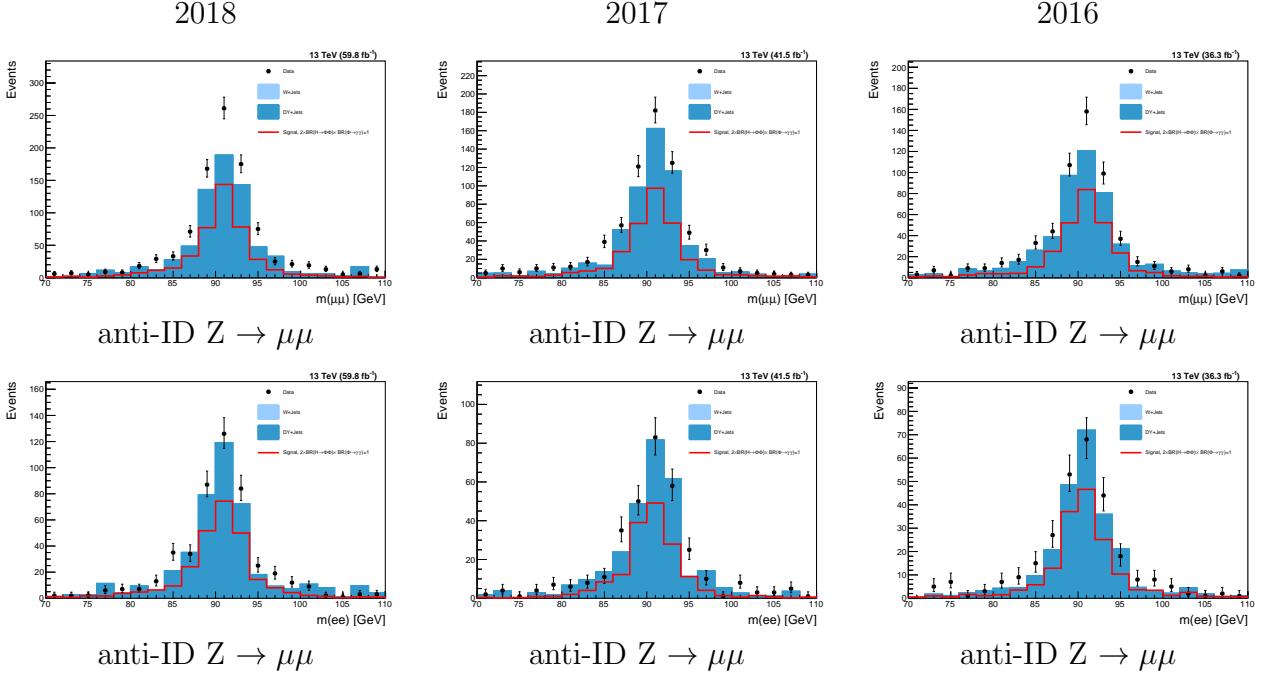


Figure 5.21: Final  $m(\ell^+\ell^-)$  distributions for data and simulated samples in the anti-ID control region. Top row:  $Z \rightarrow \mu\mu$  category. Bottom row:  $Z \rightarrow ee$  category.

## 5.5 Background Estimation

The limited statistics in the simulated background samples demand a data-driven method of background estimation. As the statistical estimator used to extract signal involves the  $L_{xy}$  distribution in the positive  $L_{xy}$  signal region, it is crucial to get an accurate estimate of both the yield and shape of the expected background. This section outlines the expected sources of background and the process used to estimate the background using the anti-ID control region and negative  $L_{xy}$  sidebands. We demonstrate closure tests which are used to derive the systematic uncertainties on the background estimation.

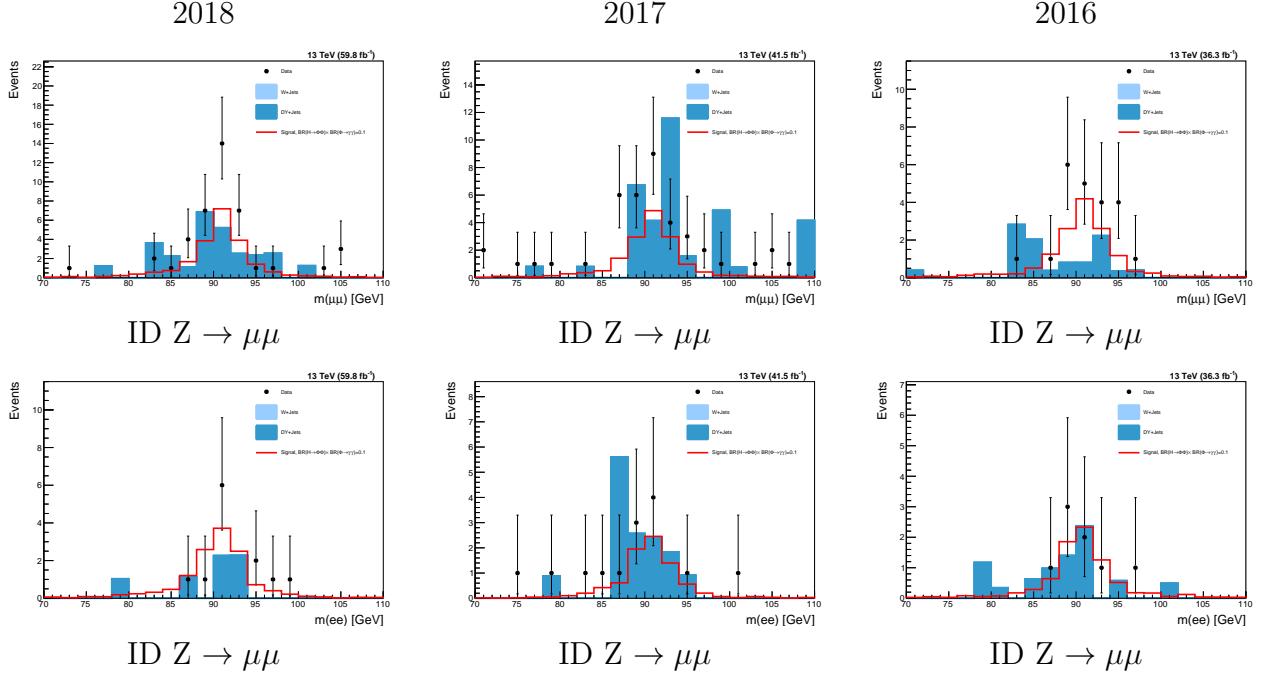


Figure 5.22: Final  $m(\ell^+\ell^-)$  distributions for data and simulated samples in the ID signal region. Top row:  $Z \rightarrow \mu\mu$  category. Bottom row:  $Z \rightarrow ee$  category.

### 5.5.1 Sources of Standard Model Background

There are two main expected SM background contributions originating from the Drell-Yan process, which correspond to combinations of "fake" photons from electromagnetic fluctuations of neutral pions produced in jets or PU, and "real" photons from initial or final state radiation (ISR/FSR). The FSR photons are strongly reducible using invariant mass cuts outlined in section 5.4.3.2. From studies using generator level information it was determined that the subleading photons primarily originate from jets. A Feynman diagram showing possible sources of background is shown in figure 5.24.

There are zero sources of resonant diphoton background expected in this analysis. While the standard model process  $H \rightarrow \gamma\gamma$  produces two final state photons, we expect very few events due to the low branching fraction of 0.0025 [28]. For the few events that pass full selection criteria, the invariant diphoton mass of  $m_{\gamma\gamma} = m_H = 125$  GeV would classify events

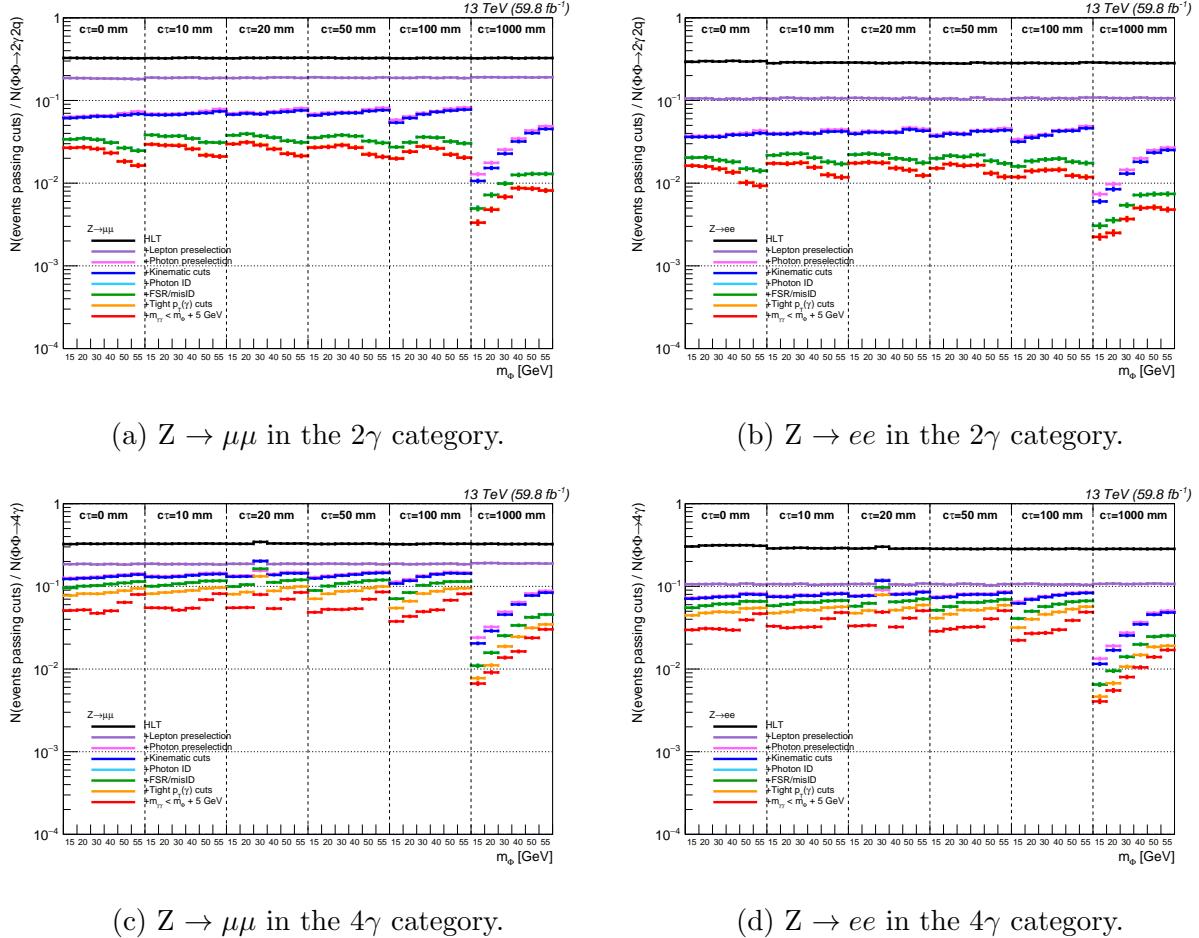


Figure 5.23: Cutflow efficiency for signal events with exactly 2 photons (top) and 4 photons (bottom).

in the negative  $L_{xy}$  sideband by section 5.3. This region is used to determine the yield of the expected background and would result in a negligible overestimation of background events.

### 5.5.2 Method of Background Estimation

The largest obstacle towards implementing a reliable background estimation for this analysis was the large statistical fluctuations of the estimated contributions using weights of events in both adjoint regions (e.g. fake-rate based method) and simulation. A mixed-event method

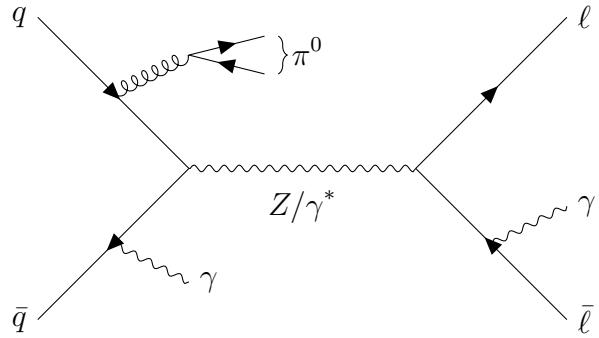


Figure 5.24: Feynman diagram showing the possible background contributions from Drell-Yan events. Photons can originate from neutral pions produced in jets, initial state radiation from quarks, or final state radiation from leptons.

which constructs diphotons using leading and sub-leading photons from different events was considered in order to create a large sample of events from uncorrelated combinations of real and fake photons but did not provide an adequate level of closure. In order to account for the correlation between photon kinematics, we predict the shape of the  $L_{xy}$  and  $m_{\gamma\gamma}$  distributions using the anti-ID control region, and normalize the expected number of background events by scaling the anti-ID control region to match the yield of the ID signal region in the negative  $L_{xy}$  sideband. The estimated background in bin  $i$  is given by

$$N_{\text{pkg}}^i = \frac{N(\gamma_{\text{ID}} + \gamma_{\text{ID}}, m_{\gamma\gamma} > m_H/2)}{N(\gamma_{\text{ID}} + \gamma_{\text{antiID}}, m_{\gamma\gamma} > m_H/2)} \times N(\gamma_{\text{ID}} + \gamma_{\text{antiID}}, i) \quad (5.14)$$

Where  $N(\gamma_{\text{ID}} + \gamma_{\text{(anti)ID}}, m_{\gamma\gamma} > m_H/2)$  is the number of events in the  $L_{xy}$  sideband of the (anti)ID region and  $N(\gamma_{\text{ID}} + \gamma_{\text{antiID}}, i)$  is the number of events in bin  $i$  of the anti-ID control region.

### 5.5.3 Closure Tests

The method of background estimation is validated by comparing the shape of the negative  $L_{xy}$  sidebands in both the ID signal region and anti-ID control region. The behavior of the

fit as  $L_{xy}$  approaches zero is used as an indicator for the agreement in the positive  $L_{xy}$  region, and is used to determine the bin-by-bin shape uncertainty. Error bars for the signal and control region data represent the 68% Poisson confidence interval given by

$$\text{C.I.} = [0.5 * \chi^2(2N, \alpha/2), 0.5 * \chi^2(2(N + 1), 1 - \alpha/2)] \quad (5.15)$$

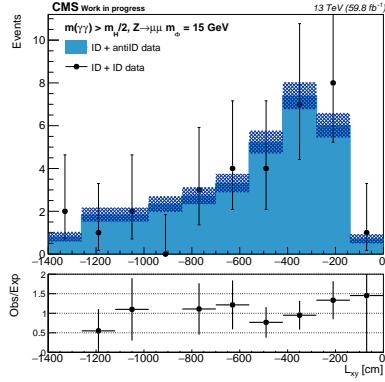
where  $\chi^2$  is the chi-square critical value,  $N$  is the number of observed events, and  $\alpha$  is the significance level of  $(1 - 0.68)$ . Since the anti-ID bin contents are scaled in order to match the yield of the ID signal region, the uncertainty is scaled by the same amount.

The anti-ID control region shows compatible values with the ID signal region across all ranges of negative  $L_{xy}$ . Based on the disagreement as  $L_{xy}$  approaches 0, we apply a conservative 40% uncertainty on each bin in the positive  $L_{xy}$  signal region. The full set of closure plots across all categories for each year can be seen in figures 5.25-5.30.

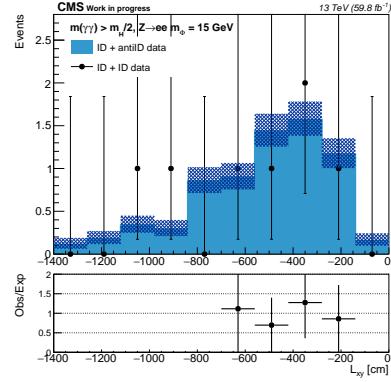
## 5.6 Statistical Analysis

The signal extraction is performed in two dimensional histograms of  $L_{xy}$  and  $m_{\gamma\gamma}$ . The data are binned in four bins of  $m_{\gamma\gamma}$  spanning 4 GeV to  $m_\Phi + 5$  GeV and rows in  $L_{xy}$  spanning from -10 cm to 100 cm. Each row represents the  $L_{xy}$  distribution for a given range of  $m_{\gamma\gamma}$ , and is initially binned using a bin width of 1 cm. Rows are then rebinned using the expected background distribution from anti-ID events derived in section 5.5.2. For each range of  $m_{\gamma\gamma}$ , the  $L_{xy}$  is rebinned into quantiles of 25% in order to ensure that there are no bins with zero estimated background. Events in the lowest range of  $m_{\gamma\gamma}$  are binned in quantiles of 50% due to the limited number of expected background events.

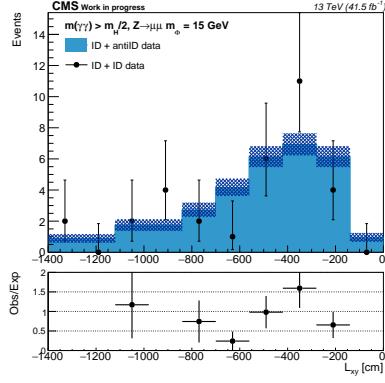
2018  $Z \rightarrow \mu\mu$



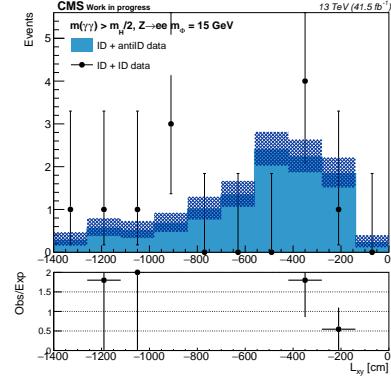
2018  $Z \rightarrow ee$



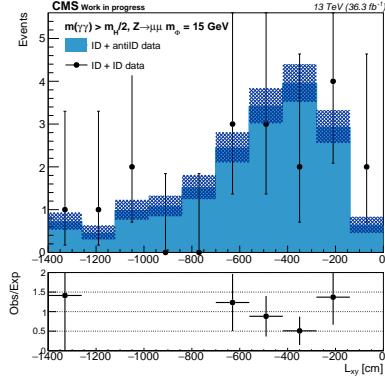
2017  $Z \rightarrow \mu\mu$



2017  $Z \rightarrow ee$



2016  $Z \rightarrow \mu\mu$



2016  $Z \rightarrow ee$

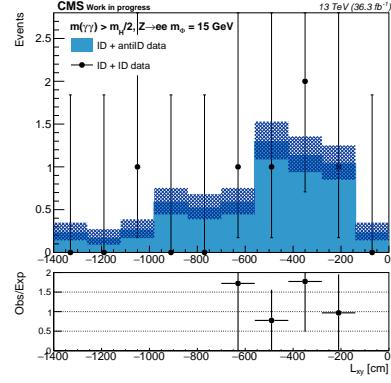
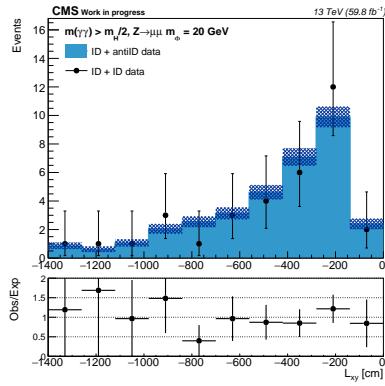
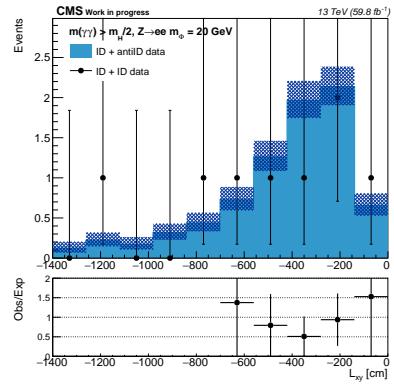


Figure 5.25: Shape comparison of ID signal region and anti-ID control region data in the negative  $L_{xy}$  sideband for  $m_\Phi = 15 \text{ GeV}$  in the  $Z \rightarrow \mu\mu$  category (left) and  $Z \rightarrow ee$  category (bottom) across all three years.

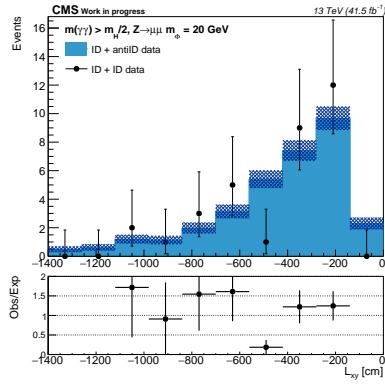
2018  $Z \rightarrow \mu\mu$



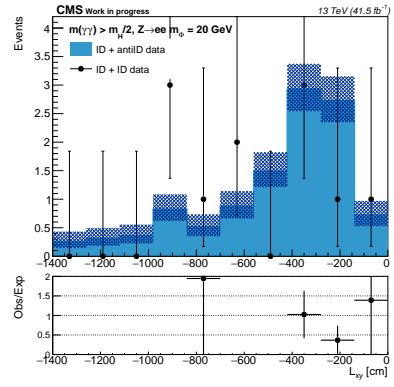
2018  $Z \rightarrow ee$



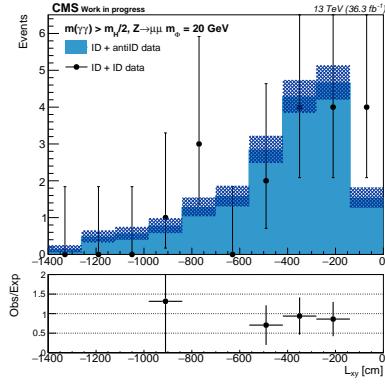
2017  $Z \rightarrow \mu\mu$



2017  $Z \rightarrow ee$



2016  $Z \rightarrow \mu\mu$



2016  $Z \rightarrow ee$

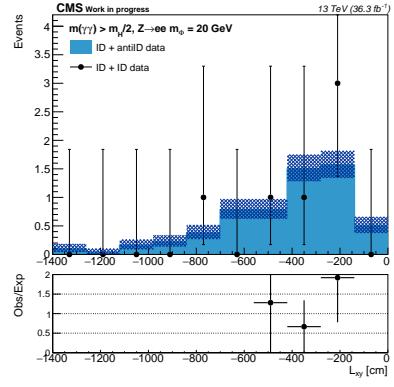
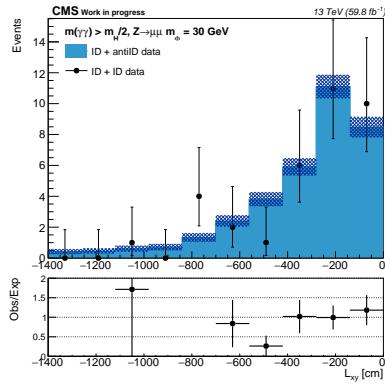
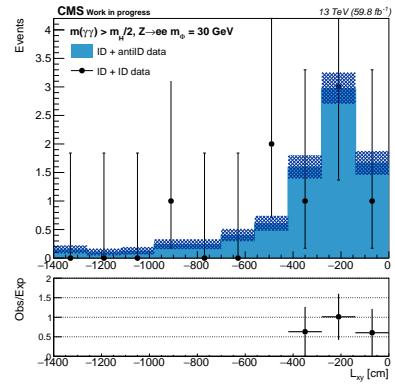


Figure 5.26: Shape comparison of ID signal region and anti-ID control region data in the negative  $L_{xy}$  sideband for  $m_\Phi = 20 \text{ GeV}$  in the  $Z \rightarrow \mu\mu$  category (left) and  $Z \rightarrow ee$  category (bottom) across all three years.

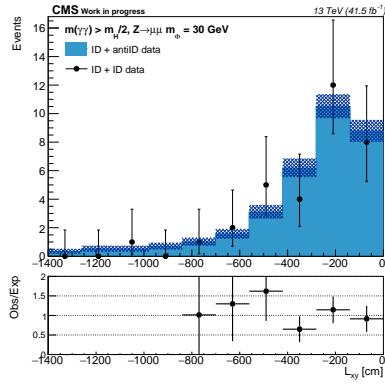
2018  $Z \rightarrow \mu\mu$



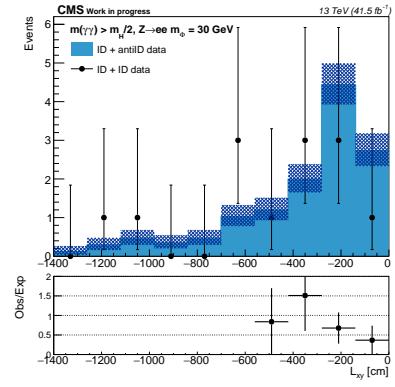
2018  $Z \rightarrow ee$



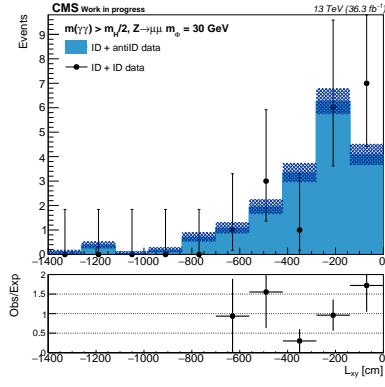
2017  $Z \rightarrow \mu\mu$



2017  $Z \rightarrow ee$



2016  $Z \rightarrow \mu\mu$



2016  $Z \rightarrow ee$

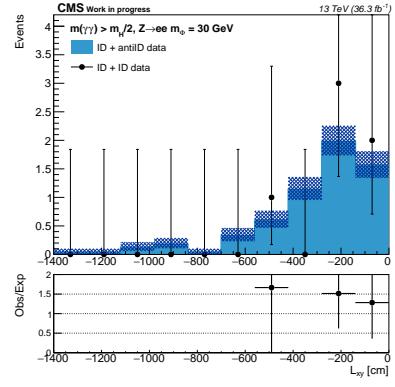


Figure 5.27: Shape comparison of ID signal region and anti-ID control region data in the negative  $L_{xy}$  sideband for  $m_\Phi = 30 \text{ GeV}$  in the  $Z \rightarrow \mu\mu$  category (left) and  $Z \rightarrow ee$  category (bottom) across all three years.

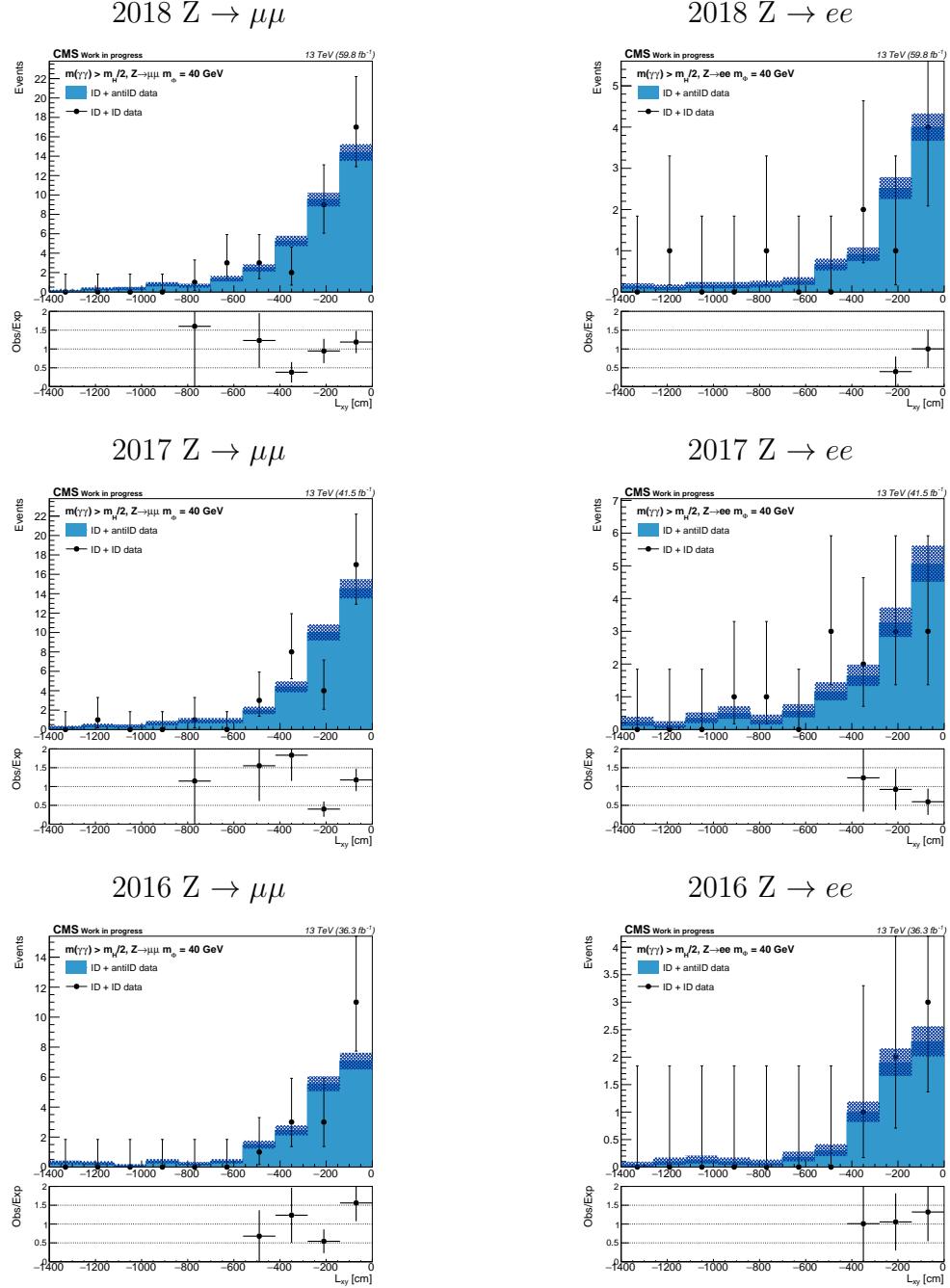
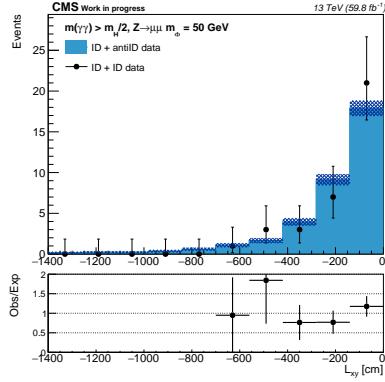
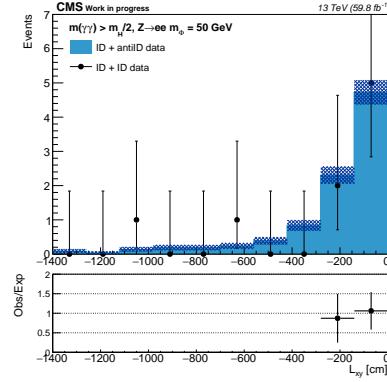


Figure 5.28: Shape comparison of ID signal region and anti-ID control region data in the negative  $L_{xy}$  sideband for  $m_\Phi = 40 \text{ GeV}$  in the  $Z \rightarrow \mu\mu$  category (left) and  $Z \rightarrow ee$  category (bottom) across all three years.

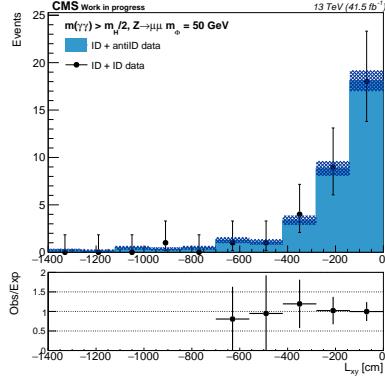
2018  $Z \rightarrow \mu\mu$



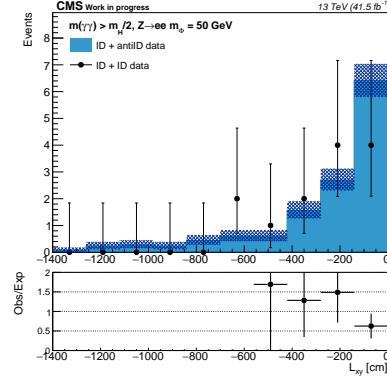
2018  $Z \rightarrow ee$



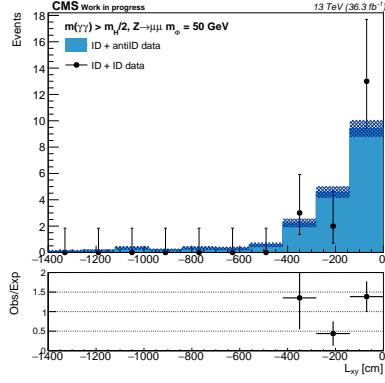
2017  $Z \rightarrow \mu\mu$



2017  $Z \rightarrow ee$



2016  $Z \rightarrow \mu\mu$



2016  $Z \rightarrow ee$

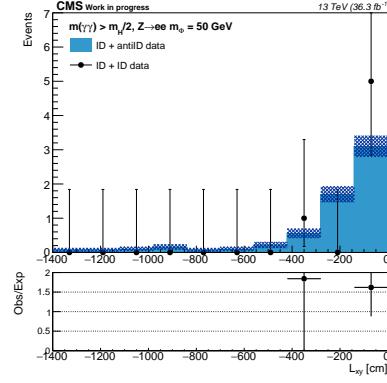


Figure 5.29: Shape comparison of ID signal region and anti-ID control region data in the negative  $L_{xy}$  sideband for  $m_\Phi = 50 \text{ GeV}$  in the  $Z \rightarrow \mu\mu$  category (left) and  $Z \rightarrow ee$  category (bottom) across all three years.

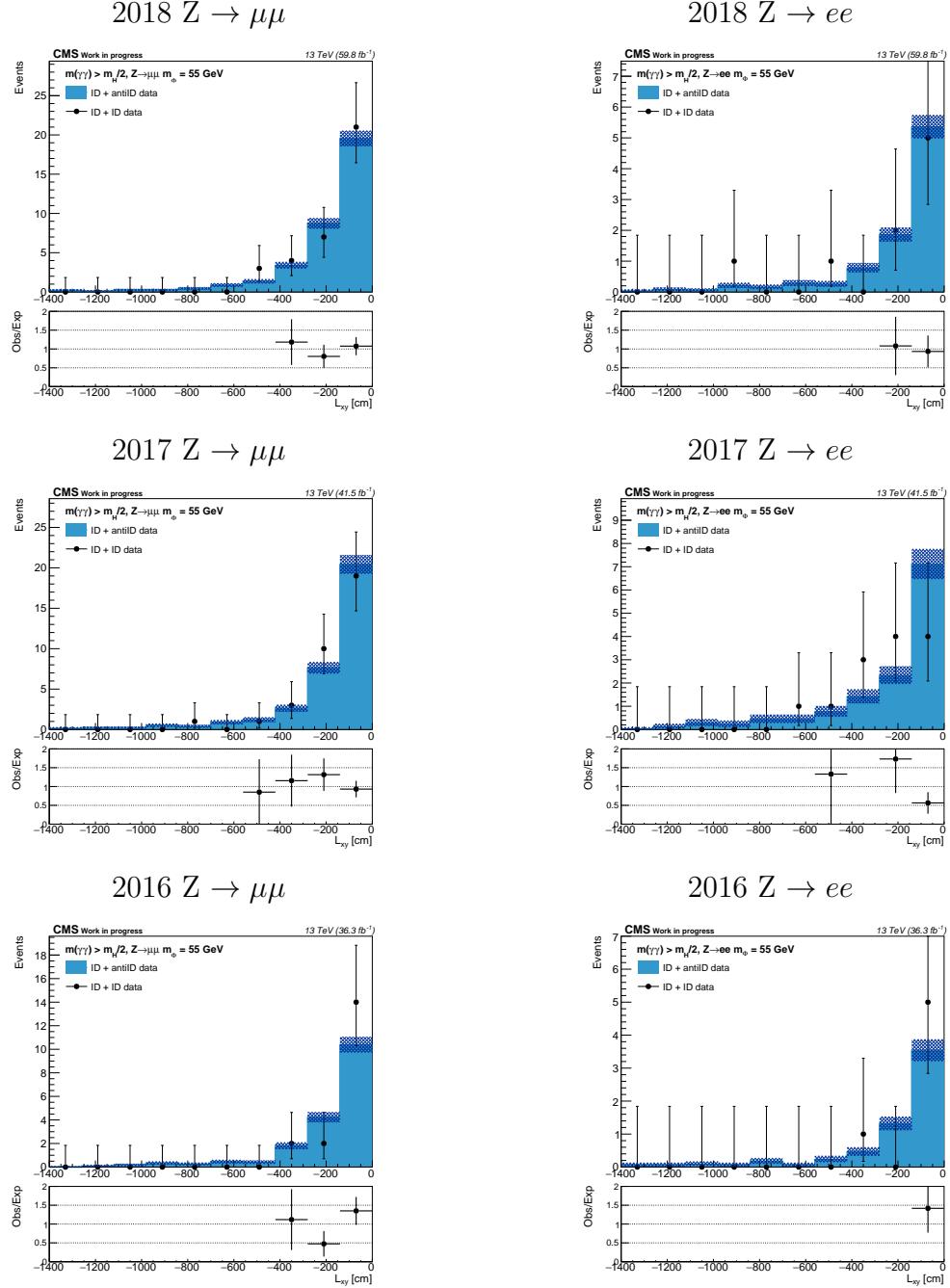


Figure 5.30: Shape comparison of ID signal region and anti-ID control region data in the negative  $L_{xy}$  sideband for  $m_\Phi = 55$  GeV in the  $Z \rightarrow \mu\mu$  category (left) and  $Z \rightarrow ee$  category (bottom) across all three years.

### 5.6.1 Systematic Uncertainties

Systematic uncertainties are implemented through the use of nuisance parameters which vary the yield of expected signal and background in each bin. For systematic uncertainties that act as multiplicative corrections to the expected yield, the nuisance parameters follow a log-normal (lnN) probability density function (PDF). If the uncertainty for a multiplicative correction is asymmetric, the uncertainty is implemented using asymmetric multiplicative factors scaling with log-normally distributed nuisance parameters. Lastly, the statistical uncertainty from normalization using a control region with  $N$  events and a scale factor  $\alpha$  is implemented as a systematic uncertainty using a Gamma distribution. The relation between the nuisance parameters and multiplicative corrections for each distribution can be seen in table 5.11. Systematic uncertainties can be correlated between bins, meaning that the multiplicative corrections for each uncertainty use the same nuisance parameter. Alternatively, uncorrelated uncertainties will each have their own independent nuisance parameter.

Table 5.11: Description of multiplicative corrections for systematic uncertainties as a function of the nuisance parameter  $\nu$  [23].

Uncertainty Type	Multiplicative Factor $f(\nu)$	Default Value
log-normal <sup>1</sup>	$f(\nu) = \kappa^\nu$	$\nu = 0$
asymmetric log-normal <sup>2</sup>	$f(\nu) = \begin{cases} \left(\kappa^{\text{Down}}\right)^{-\nu} & \nu < -0.5 \\ \left(\kappa^{\text{Up}}\right)^\nu & \nu > 0.5 \\ e^{\nu * g(\kappa^{\text{Down}}, \kappa^{\text{Up}}, \nu)} & -0.5 < \nu < 0.5 \end{cases}$	$\nu = 0$
Gamma	$f(\nu) = \nu/N$	$\nu = N + 1$

$$^*g(\kappa^{\text{Down}}, \kappa^{\text{Up}}, \nu) = 1/8 \left[ 4 \ln \left( \kappa^{\text{Up}} / \kappa^{\text{Down}} \right) + \ln \left( \kappa^{\text{Up}} \kappa^{\text{Down}} \right) (48\nu^5 - 40\nu^3 + 15\nu) \right]$$

<sup>1</sup> $\kappa = 1 + \Delta x/x$ , where  $\Delta x/x$  is the relative uncertainty of multiplicative correction

<sup>2</sup> $\kappa^{\text{Up/Down}}$  is 1 plus the relative change in  $x$  when shifting the uncertainty up/down by  $1\sigma$

These systematic uncertainties show small effects on the expected limits as the analysis is dominated by statistical Poisson uncertainties due to the small number of predicted background events. This section outlines the application of systematic uncertainties on the

predicted signal and background. Table 5.12 contains a summary of all systematic uncertainties and the corresponding correlations.

### 5.6.1.1 Signal Uncertainties

For normalization of signal samples relying on the SM Higgs boson cross section, we apply a theoretical uncertainty of  $(-3.1, +3.8)\%$  to the QCD scale and a  $\pm 1.6\%$  uncertainty in the parton distribution function (PDF) and  $\alpha_s$  [24]. These scale factors are applied directly to the normalization of the simulated events. The theoretical uncertainties are 100% correlated in all categories.

Each scale factor also has a measured uncertainty, which is applied by shifting the nominal weights up or down and reapplying the scale factor. The systematic uncertainty is then taken as the percent change in each bin due to the variation in the scale factors. These are multiplicative uncertainties and applied using asymmetric log-normal constraints. All scale factor uncertainties for the same physics object are 100% correlated between categories using that object. Therefore photon scale factor uncertainties are correlated between all categories, while the electron/muon uncertainties are only correlated within to the  $Z \rightarrow ee$  and  $Z \rightarrow \mu\mu$  channels respectively. These adjustments generally change the yield on the order of 0.1%.

Photons in simulated samples have energy corrections to match the mean and uncertainty of their energy resolution to the data. These are applied as a scale correction, which adjusts the central value, and a random smear to match the resolution. Both the scale and smearing corrections have an associated measured uncertainty. Unlike the scale factor corrections described in section 5.2.2.5, which are event level weights, the scale and resolution corrections directly modify the photon energy in each event. In order to adjust the yields based on the scale and resolution uncertainty, the energy of all photons is shifted up or down by the given uncertainty and the relevant distributions for signal extraction are calculated using the shifted values. The residuals when uncertainties are shifted up or down are then taken as the uncertainty. Although this acts as a shape uncertainty, it is implemented as an

asymmetric log-normal uncertainty on the yield in each bin. As the corrections are applied to all photons, the nuisance parameters are 100% correlated between all events.

Luminosity and pileup act as a multiplicative corrections for normalization and are treated as a lnN uncertainty. The corresponding luminosity uncertainties from table 5.2 are applied for each year. Since the method used to measure the luminosity is consistent for each year, the luminosity uncertainty is correlated for all categories across all three years of data taking. For the same reason, the 5% pileup uncertainty is correlated for all years as well.

### 5.6.1.2 Background Uncertainties

The predicted background has two systematic uncertainties for the shape and normalization. The shape uncertainty was determined from the non-closure uncertainty from section 5.5.3 and is implemented using 40% lnN in each bin, which is fully uncorrelated between each bin, year, and category. The normalization uncertainty is implemented using a Gamma function and is fully correlated between all bins for the same category and uncorrelated between different categories. This function has two parameters per bin given by  $N$  and  $\alpha_i$ .  $N$  symbolizes the number of events in the negative  $L_{xy}$  sideband of the ID signal region and  $\alpha_i$  represents the scale factor multiplied by  $N$  in order to obtain the expected number of events in bin  $i$ . Equation 5.14 can be rewritten in these terms as

$$N_{\text{bkg}}^i = \alpha_i \times N \quad (5.16)$$

where

$$\alpha_i = \frac{N(\gamma_{\text{ID}} + \gamma_{\text{antiID}}, i)}{N(\gamma_{\text{ID}} + \gamma_{\text{antiID}}, m_{\gamma\gamma} > m_{\text{H}}/2)} \quad (5.17)$$

and

$$N = N(\gamma_{\text{ID}} + \gamma_{\text{ID}}, m_{\gamma\gamma} > m_{\text{H}}/2) \quad (5.18)$$

Under this equivalent interpretation of equation 5.14, the anti-ID control region is used to derive a transfer factor that scales negative  $L_{xy}$  ID region events to positive  $L_{xy}$  signal region events.

Table 5.12: Summary of the systematic uncertainties and corresponding correlations

Uncertainty	Type	Value(%)	Signal	Background	Correlations
QCD scale (ZH)	lnN	-3.1, +3.8	✓	-	All categories
PDF ( $\sigma$ ) + $\alpha_S$	lnN	$\pm 1.6$	✓	-	All categories
Photon Energy Scale	lnN	$\pm \sigma$	✓	-	All categories
Photon Energy Resolution	lnN	$\pm \sigma$	✓	-	All categories
Photon ID	lnN	$\pm \sigma$	✓	-	All categories
Photon Pixel Veto	lnN	$\pm \sigma$	✓	-	All categories
Lepton ID	lnN	$\pm \sigma$	✓	-	Same flavor categories
Lepton RECO	lnN	$\pm \sigma$	✓	-	Same flavor categories
Muon Isolation	lnN	$\pm \sigma$	✓	-	Muon categories
Muon HLT	lnN	$\pm \sigma$	✓	-	Muon categories
Electron HLT	lnN	$\pm \sigma$	✓	-	Electron categories
Luminosity	lnN	1.2 – 2.5	✓	-	All categories
Pileup	lnN	$\pm 5.0$	✓	-	All categories
bkg normalization	gamma	$N \times \alpha$	-	✓	Bins in same category
bkg non-closure	lnN	40	-	✓	Fully uncorrelated

### 5.6.2 Limit Setting Procedure

As discussed in section 5.2.2.1, the signal samples have been generated using the process  $H \rightarrow \Phi\Phi$  with the  $\Phi$  fixed to decay either to a pair of photons or jets with a branching ratio of 50% for each decay channel. The final number of expected signal events is given by

$$N = \mathcal{L} \sigma_{ZH} \varepsilon_Z \mathcal{BR}(H \rightarrow \Phi\Phi) (2\mathcal{BR}(\Phi \rightarrow \gamma\gamma) (1 - \mathcal{BR}(\Phi \rightarrow \gamma\gamma)) \varepsilon_1 + \mathcal{BR}(\Phi \rightarrow \gamma\gamma)^2 \varepsilon_2) \quad (5.19)$$

where  $\varepsilon_Z$  is the  $\mathcal{BR}(Z \rightarrow \ell\ell)$  times the efficiency and acceptance to reconstruct leptonic decays of the  $Z$ ,  $\varepsilon_1$  is the efficiency to reconstruct a photon pair when only one  $\Phi$  decays to photons, and  $\varepsilon_2$  is the efficiency to reconstruct a photon pair when both  $\Phi$  decay to photons. Let  $\beta = \mathcal{BR}(\Phi \rightarrow \gamma\gamma)$ . Since we intend to extract the model independent product of the

branching fraction of the Higgs boson to two  $\Phi$  candidates with the branching fraction of  $\Phi \rightarrow \gamma\gamma$ , we can rewrite equation 5.19 as follows

$$N = \mathcal{L}\sigma_{ZH}\varepsilon_Z\mathcal{BR}(H \rightarrow \Phi\Phi)\mathcal{BR}(\Phi \rightarrow \gamma\gamma)(2\varepsilon_1(1 - \beta) + \varepsilon_2\beta) \quad (5.20)$$

It is evident that the limit on  $\mathcal{BR}(H \rightarrow \Phi\Phi)\mathcal{BR}(\Phi \rightarrow \gamma\gamma)$  will have a model dependence on  $\beta$  by a factor of  $1/[2\varepsilon_1(1 - \beta) + \varepsilon_2\beta]$ . Isolating the dependence of  $\beta$  yields

$$f(\beta) = \frac{1}{1 - (1 - \frac{\varepsilon_2}{2\varepsilon_1})\beta} \quad (5.21)$$

This factor ranges from  $[1, \frac{2\varepsilon_1}{\varepsilon_2}]$  as  $\beta$  goes from 0 to 1. If the reconstruction of each  $\Phi$  was independent of the other, then we would expect  $\varepsilon_2 = 1 - (1 - \varepsilon_1)^2$ . In this case, evaluating  $f$  for  $\beta = 1$  would give  $f(1) = 1/(1 - \varepsilon_1/2) > 1$ . However, it is common to mispair photons from events where only 1 photon from each  $\Phi$  passes reconstruction, which inflates  $\varepsilon_2$  and leads to  $\varepsilon_2 > 1 - (1 - \varepsilon_1)^2$ , causing  $f(1) < 1$ . This effect is larger for higher values of  $m_\Phi$  as the mispaired  $\Phi$  candidate is more likely to pass selection criteria and have a positive  $L_{xy}$ . Results from this analysis are performed using  $\beta = 0.5$  but can be reinterpreted for other branching ratios by scaling them by equation 5.21. Figures 5.31 and 5.32 show the dependence of the limits on  $\beta$  for the  $Z \rightarrow \mu\mu$  and  $Z \rightarrow ee$  categories across all hypothetical values of  $m_\Phi$  and  $c\tau$ .

Upper limits are set using the `combine` toolkit provided by CMS, which was created to combine results from multiple decay channels in the Higgs boson discovery [23]. Each bin is treated as a Poisson based likelihood with two contributions: signal and background. For each bin, the number of expected events is given by

$$E[n_i] = \mu s_i(\boldsymbol{\theta}_s) + b_i(\boldsymbol{\theta}_b) \quad (5.22)$$

where  $s_i(\boldsymbol{\theta}_s)$  and  $b_i(\boldsymbol{\theta}_b)$  are the expected signal and background events given a set of nuisance

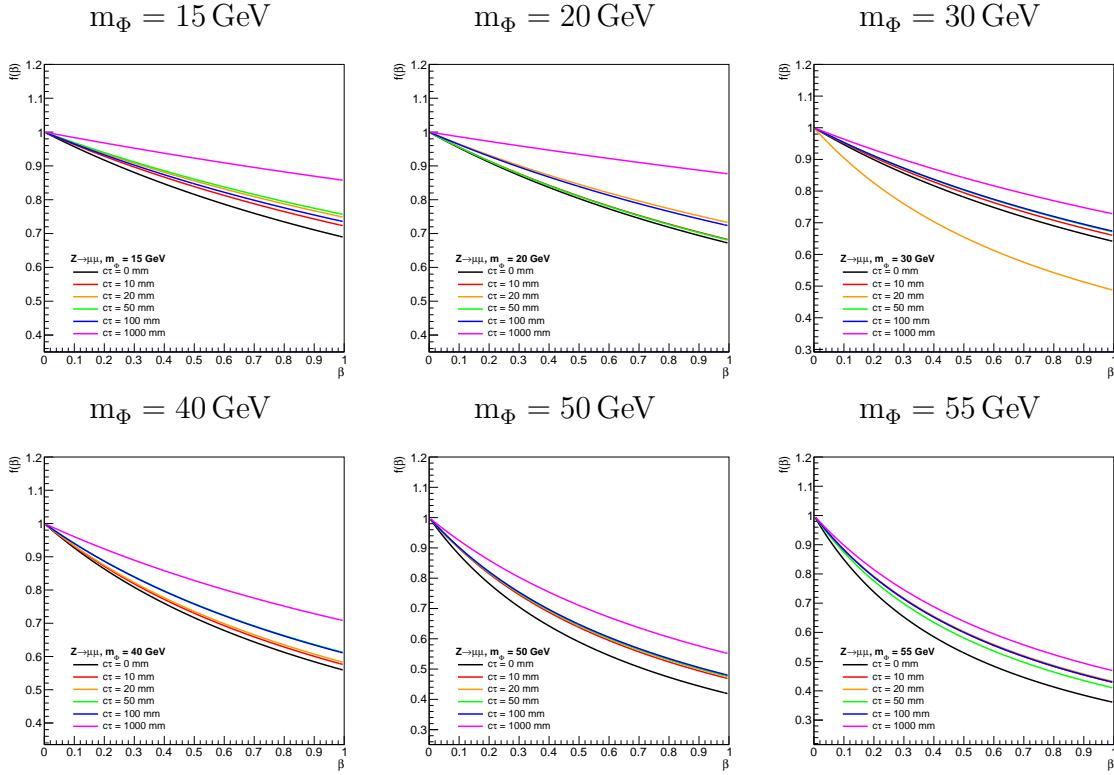


Figure 5.31: Dependence of the normalization of limits on  $\mathcal{BR}(\Phi \rightarrow \gamma\gamma)$  for the  $Z \rightarrow \mu\mu$  categories across all masses and values of  $c\tau$ .

parameters  $\boldsymbol{\theta}$ . Each nuisance parameter is associated with one or more systematic uncertainties outlined in section 5.6.1. The total expected background also treated as a nuisance parameter. The variable  $\mu$  is the signal strength parameter, with  $\mu = 0$  corresponding to the background-only, or "null", hypothesis and  $\mu = 1$  corresponding to the nominal signal hypothesis. In order to extract the signal strength, we allow the nuisance parameters and number of background events to vary in order to define a test statistic  $t_\mu$  as

$$t_\mu = -2 \ln \frac{\mathcal{L}(\mu, \hat{\boldsymbol{\theta}}(\mu) | \text{data})}{\mathcal{L}(\hat{\mu}, \hat{\boldsymbol{\theta}}(\hat{\mu}) | \text{data})} \quad (5.23)$$

The term  $\hat{\boldsymbol{\theta}}(\mu)$  represents the set of nuisance parameters that maximizes the likelihood  $\mathcal{L}$  for a given value of  $\mu$ . In the denominator, the term  $\hat{\mu}$  represents the signal strength parameter

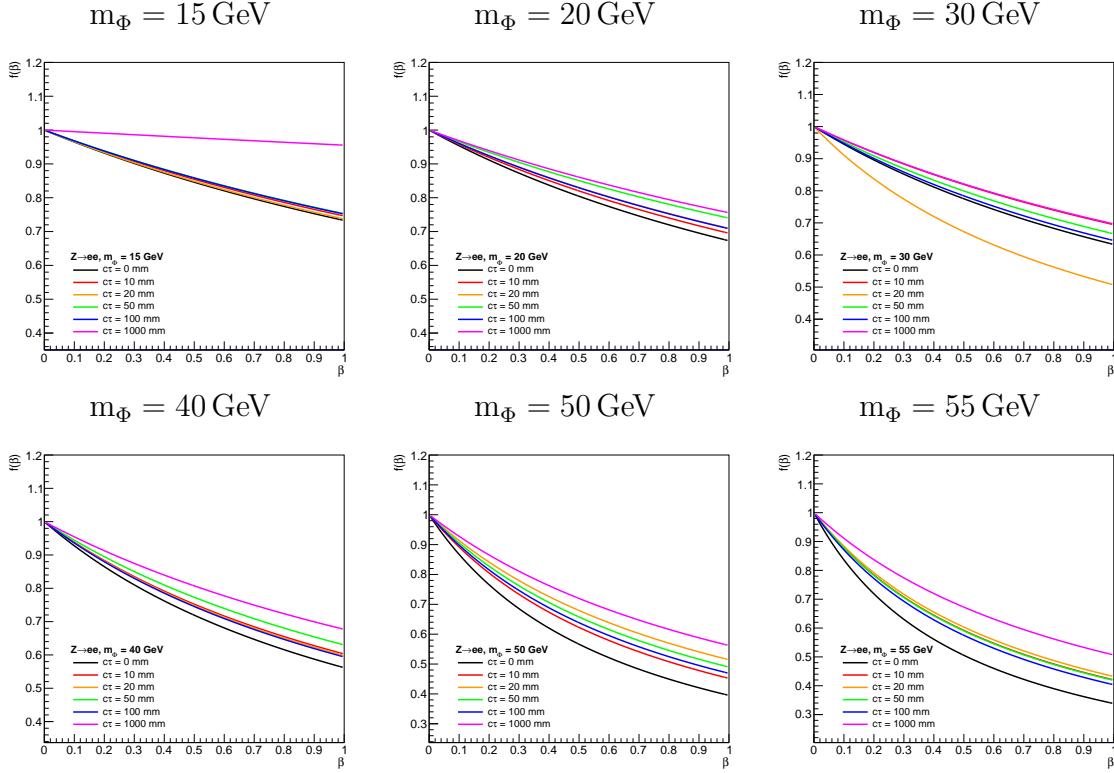


Figure 5.32: Dependence of the normalization of limits on  $\mathcal{BR}(\Phi \rightarrow \gamma\gamma)$  for the  $Z \rightarrow \mu\mu$  categories across all masses and values of  $c\tau$ .

that gives the global maximum of the likelihood. By construction, a lower test statistic indicates higher compatibility between data and a signal strength of  $\mu$ , while a higher test statistic indicates lower compatibility. For cases where we wish to set an upper limit on a strictly positive signal strength parameter, such as this analysis where a negative signal strength would represent a non-physical solution, the modified test statistic  $\tilde{q}_\mu$  is defined as

$$\tilde{q}_\mu = \begin{cases} 0 & \hat{\mu} > \mu \\ -2 \ln \frac{\mathcal{L}(\mu, \hat{\theta}(\mu) | \text{data})}{\mathcal{L}(\hat{\mu}, \hat{\theta}(\hat{\mu}) | \text{data})} & \mu \geq \hat{\mu} \geq 0 \\ -2 \ln \frac{\mathcal{L}(\mu, \hat{\theta}(\mu) | \text{data})}{\mathcal{L}(0, \hat{\theta}(0) | \text{data})} & \hat{\mu} < 0 \end{cases} \quad (5.24)$$

The test statistic follows a probability density function (PDF) denoted as  $f(\tilde{q}_\mu | \mu)$ , which

has an approximate analytic form by applying the Wald approximation [45]. It is also useful to define the specific test statistic for the case where  $\mu = 0$  as

$$q_0 = \begin{cases} 0 & \hat{\mu} < 0 \\ -2 \ln \frac{\mathcal{L}(0, \hat{\theta}(0) | \text{data})}{\mathcal{L}(\hat{\mu}, \hat{\theta}(\hat{\mu}) | \text{data})} & \hat{\mu} \geq 0 \end{cases} \quad (5.25)$$

This background only test statistic has its own corresponding PDF denoted by  $f(q_0|0)$ , once again having an analytic form from the Wald approximation [45]. It should be noted that while the best fit value  $q_{0,\text{obs}}$  is equivalent to the best fit value from  $\tilde{q}_\mu$  when  $\mu$  is set equal to 0, the PDF is a distinct distribution from  $f(\tilde{q}_\mu|\mu)$ . From these two PDFs, we define two p-values: the probability of measuring at least this many events given the null hypothesis ( $p_b$ ) and the probability to measure more events given the best fit signal hypothesis ( $p_\mu$ ) as the following:

$$1 - p_b = \int_{q_{0,\text{obs}}}^{\infty} f(q_0|0) dq_0 \quad (5.26)$$

$$p_\mu = \int_{\tilde{q}_{\mu,\text{obs}}}^{\infty} f(\tilde{q}_\mu|\mu) d\tilde{q}_\mu \quad (5.27)$$

From these two p-values, the modified frequentist confidence limits  $\text{CL}_s$  are defined as

$$\text{CL}_s(\mu) = \frac{p_\mu}{1 - p_b} \quad (5.28)$$

A signal strength of  $\mu$  is said to be excluded at a confidence level (CL) of  $\alpha$  for  $\text{CL}_s(\mu) < 1 - \alpha$ . For example, a value of  $\text{CL}_s(\mu) < 0.05$  would exclude  $\mu$  at 95% CL.

In order to set a limit on the desired branching ratios, the expected number of signal events described in equation 5.20 are scaled such that  $\mu$  represents  $\mathcal{BR}(\text{H} \rightarrow \Phi\Phi) \times \mathcal{BR}(\Phi \rightarrow \gamma\gamma)$ . The simulated signal is normalized using the measured cross section for quark and gluon induced Z+H production found in table 5.13. The Z boson in all signal samples is forced to decay leptonically, so the samples must be normalized to the total leptonic branching ratio to  $e, \mu$ , and  $\tau$  of 0.100986 [28].

Table 5.13: Higgs production cross sections used to normalize simulated signal samples [24].

Production Model	Cross Section
ZH	0.8696
ggZH	0.1057

### 5.6.3 Nuisance Parameter Impacts and Pull Distributions

The fitting procedure done by `combine` has several checks to validate that the nuisance parameter are properly constrained and the fits are well behaved. We provide checks on the pull distributions and impacts of the nuisance parameters. The pull distribution refers to the difference in initial pre-fit and final post-fit values of the nuisance parameters, divided by the initial pre-fit uncertainty. The impact refers to the correlation between a nuisance parameter and the signal strength  $\mu$ . To calculate the impact of a nuisance parameter, we vary the parameter to its post fit value  $\pm 1\sigma$  of the post-fit uncertainty, then refit the remaining nuisance parameters and extract the new signal strength. The impact is defined as the difference between the signal strength parameter obtained from this fitting procedure and the signal strength parameter obtained from the nominal fit. Broadly speaking, a larger impact signifies the analysis is more sensitive to a given nuisance parameter. Although performing the fit to data involves calculating the signal strength  $\mu$ , the exact value is kept unknown in order to remain blinded.

A set of pulls and impacts of the highest impact nuisance parameters for a signal hypothesis of  $m_\Phi = 30 \text{ GeV}$  and  $c\tau = 100 \text{ mm}$  is shown in figure 5.33. Central values of the post fit nuisance parameters lie within  $1\sigma$  of the pre-fit values. The post-fit uncertainties are similar to the pre-fit uncertainties, signifying that they are not tightly constrained by the fit process. Impacts are measured for a signal strength parameter corresponding to  $\mu = 20 \times \mathcal{BR}(H \rightarrow \Phi\Phi) \times \mathcal{BR}(\Phi \rightarrow \gamma\gamma)$ . The highest impact nuisance parameters shift  $\mu$  on the order of 0.5%, corresponding to a shift in the product of branching ratios of 0.025%.

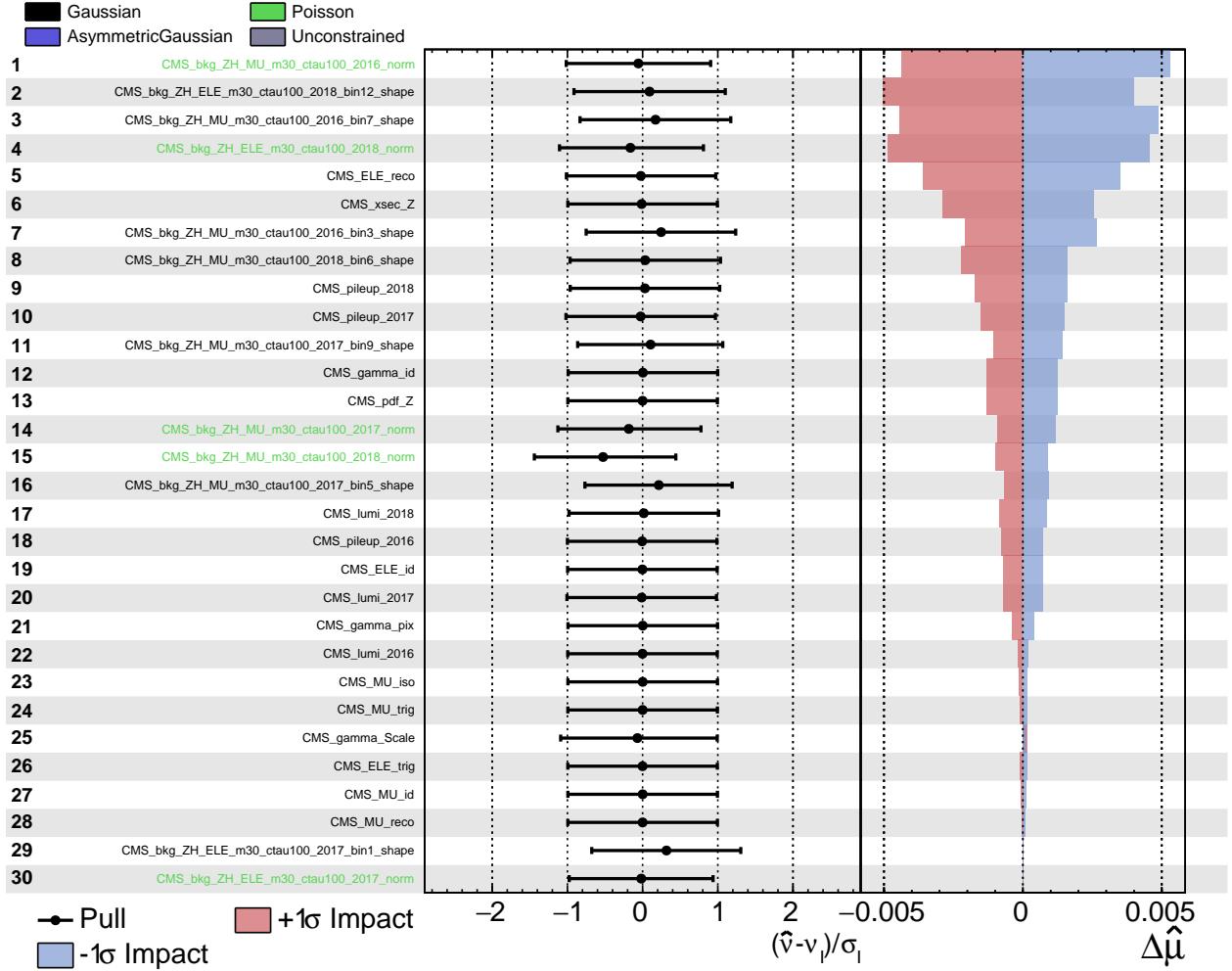


Figure 5.33: Post-fit pulls and impacts of the 30 highest impact nuisance parameters for a  $m_\Phi = 30 \text{ GeV}$ ,  $c\tau = 100 \text{ mm}$  signal hypothesis. Leftmost column shows the nuisance parameter names. Parameters starting with CMS\_bkg\* represent nuisance parameters for background normalization and non-closure uncertainty; all others correspond to systematic uncertainties of simulated signal. Middle column shows post-fit pull distributions, which lie within  $1\sigma$  of the pre-fit uncertainty. Right column shows the impacts on the signal strength parameter  $\mu = 20 \times \mathcal{BR}(H \rightarrow \Phi\Phi) \times \mathcal{BR}(\Phi \rightarrow \gamma\gamma)$ .

## 5.7 Results

The following section covers the observed data and measured limits for model dependent SM Z+H production as well as limits on model independent Z+ΦΦ production.

### 5.7.1 Observed Data

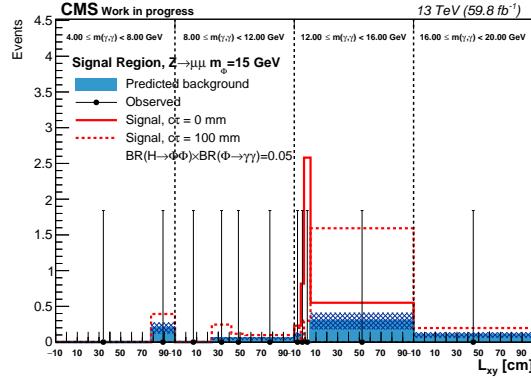
The observed number of events in the positive  $L_{xy}$  signal region as well as the predicted background are shown for all categories in figures 5.34 - 5.39. Shaded bands represent the 40% uncertainty applied to each bin determined by the closure tests. Signal samples with lifetimes  $c\tau = 0 \text{ mm}$  and  $100 \text{ mm}$  are normalized to  $\mathcal{BR}(\text{H} \rightarrow \Phi\Phi) \times \mathcal{BR}(\Phi \rightarrow \gamma\gamma) = 0.05$  and stacked on the predicted background.

### 5.7.2 Model Dependent Upper Limits

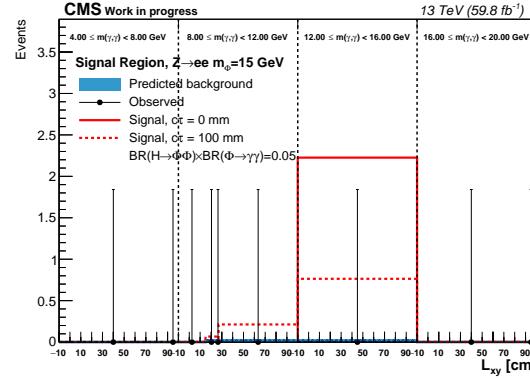
The expected and observed limits are interpreted using the asymptotic  $\text{CL}_s$  method described in section 5.6.2. Upper 95% CL limits are set on a signal strength parameter corresponding to  $\mathcal{BR}(\text{H} \rightarrow \Phi\Phi) \times \mathcal{BR}(\Phi \rightarrow \gamma\gamma)$  assuming the associated production of a Z boson and SM Higgs boson. Results are presented as a function of  $c\tau$  for all candidate signal masses listed in section 5.2.2.1. Limits are interpolated between the centrally provided and reweighted lifetimes. The one-dimensional limits versus  $c\tau$  for all candidate mass points are shown in figure 5.40. Two-dimensional limit plots of  $m_\Phi$  versus  $c\tau$  shown in figure 5.41 also interpolate between the  $m_\Phi$  points given from centrally produced signal samples.

The peak sensitivity occurs at intermediate mass points around  $m_\Phi = 30 \text{ GeV}$  and low to intermediate lifetimes below  $50 \text{ mm}$ . As mass increases, the diphoton  $I_\gamma$  cut efficiency increases because the  $\Delta R$  between the two photons increases with  $m_\Phi$ , while the  $\sigma_{inj}$  efficiency decreases as the shower shape becomes more skewed. Additionally, the estimated background also increases, as there are more diphoton events that can reconstruct a vertex in the positive  $L_{xy}$  signal region. These two effects cause peak sensitivity at intermediate values

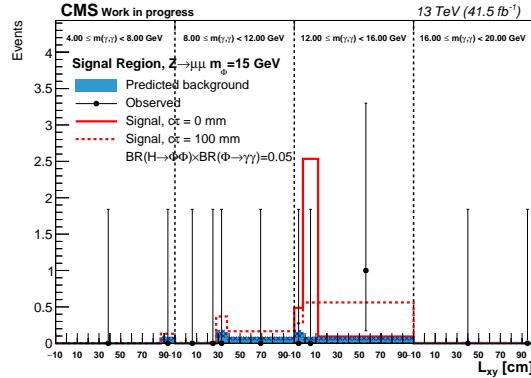
2018  $Z \rightarrow \mu\mu$



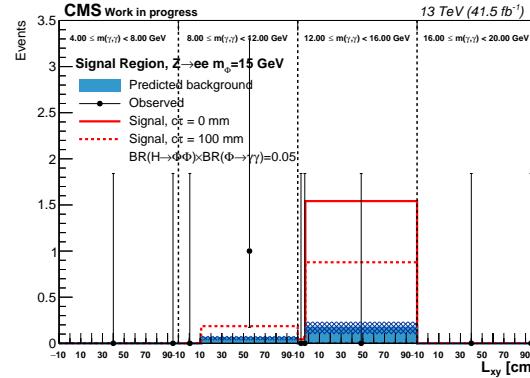
2018  $Z \rightarrow ee$



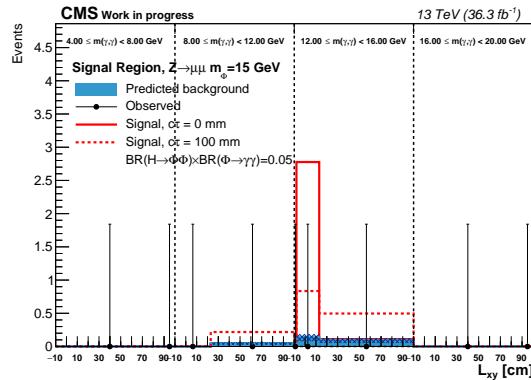
2017  $Z \rightarrow \mu\mu$



2017  $Z \rightarrow ee$



2016  $Z \rightarrow \mu\mu$



2016  $Z \rightarrow ee$

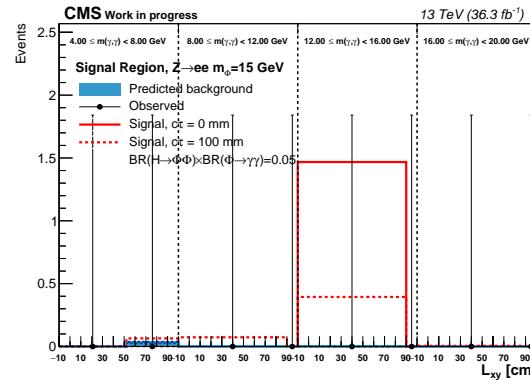
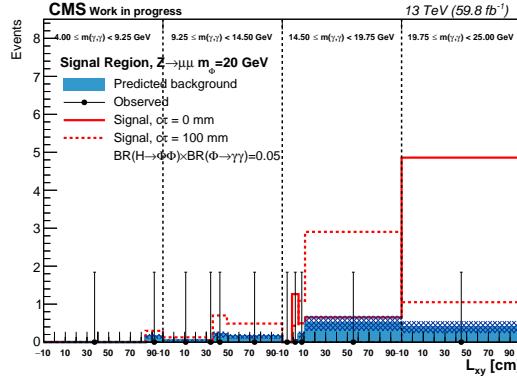
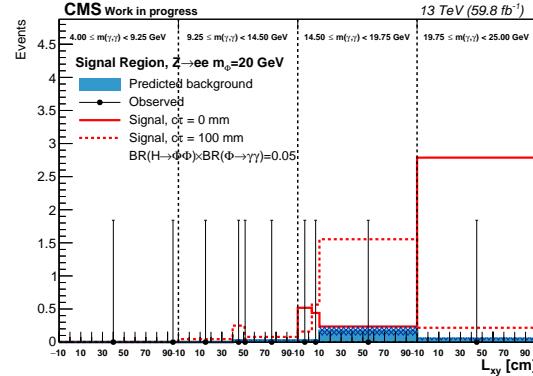


Figure 5.34: Observed number of  $m_\Phi = 15$  GeV events in the signal region of the  $Z \rightarrow \mu\mu$  category (left) and the  $Z \rightarrow ee$  category (right) across all three years.

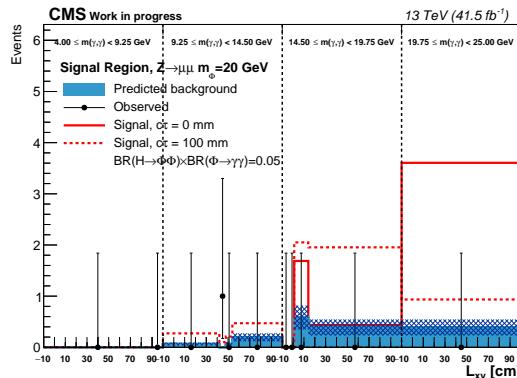
2018  $Z \rightarrow \mu\mu$



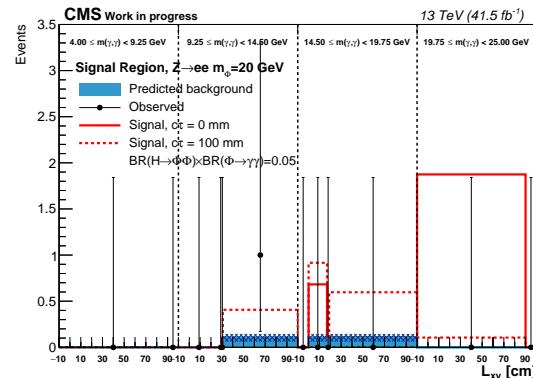
2018  $Z \rightarrow ee$



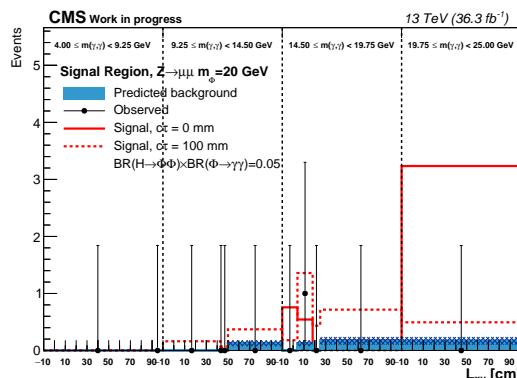
2017  $Z \rightarrow \mu\mu$



2017  $Z \rightarrow ee$



2016  $Z \rightarrow \mu\mu$



2016  $Z \rightarrow ee$

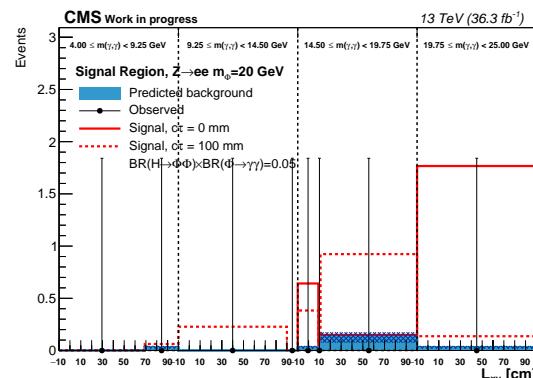
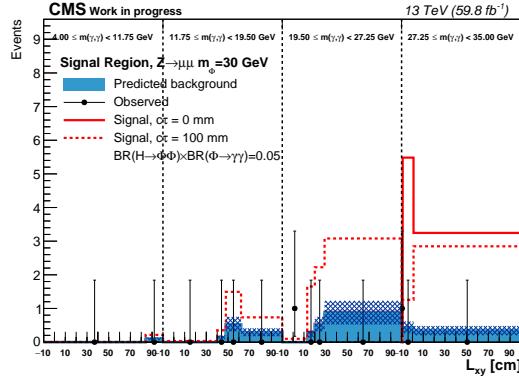
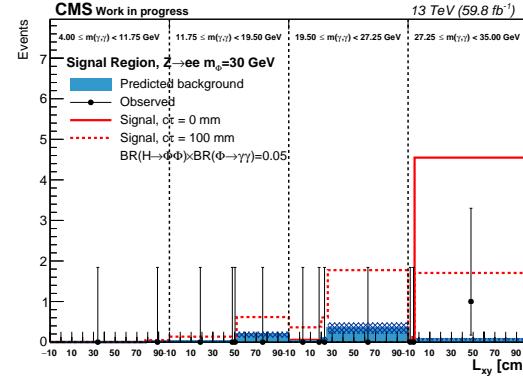


Figure 5.35: Observed number of  $m_\Phi = 20$  GeV events in the signal region of the  $Z \rightarrow \mu\mu$  category (left) and the  $Z \rightarrow ee$  category (right) across all three years.

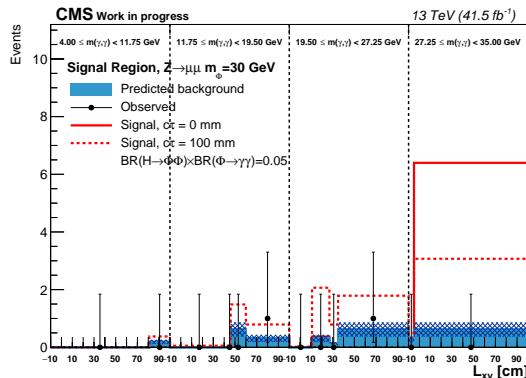
2018  $Z \rightarrow \mu\mu$



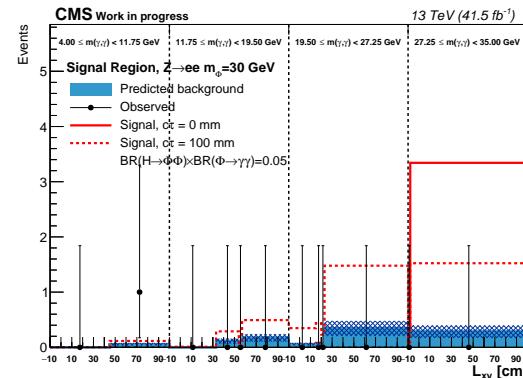
2018  $Z \rightarrow ee$



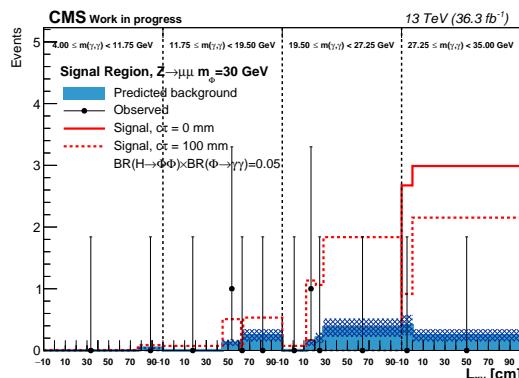
2017  $Z \rightarrow \mu\mu$



2017  $Z \rightarrow ee$



2016  $Z \rightarrow \mu\mu$



2016  $Z \rightarrow ee$

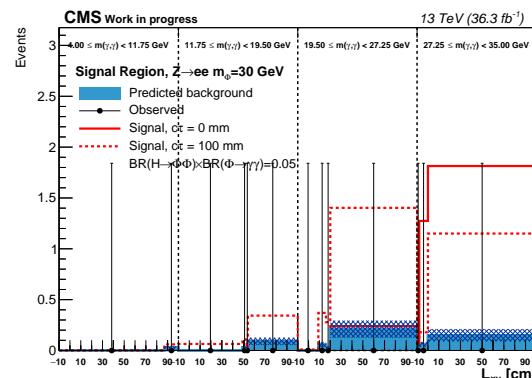
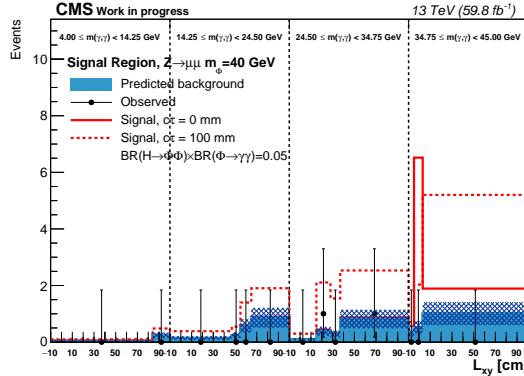
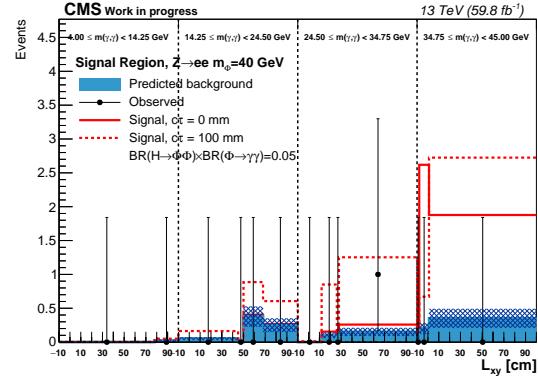


Figure 5.36: Observed number of  $m_{\Phi} = 30$  GeV events in the signal region of the  $Z \rightarrow \mu\mu$  category (left) and the  $Z \rightarrow ee$  category (right) across all three years.

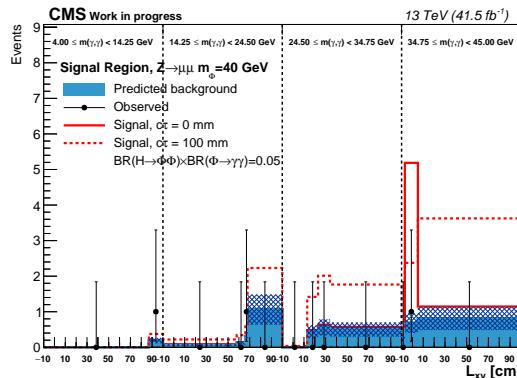
2018  $Z \rightarrow \mu\mu$



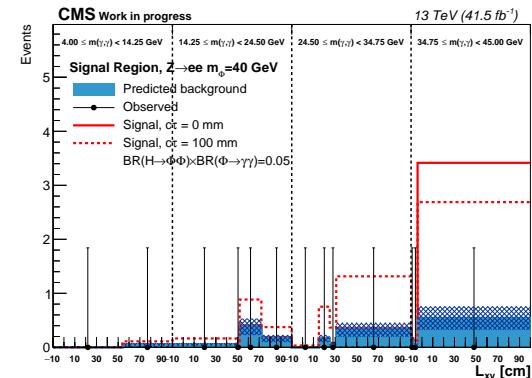
2018  $Z \rightarrow ee$



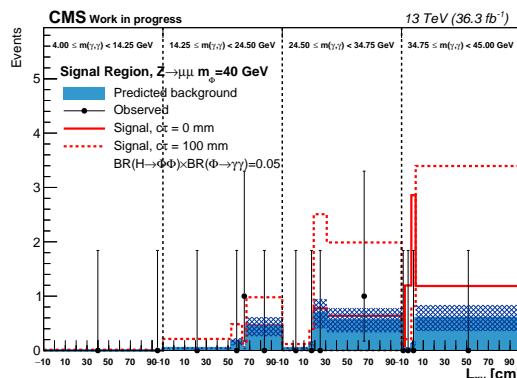
2017  $Z \rightarrow \mu\mu$



2017  $Z \rightarrow ee$



2016  $Z \rightarrow \mu\mu$



2016  $Z \rightarrow ee$

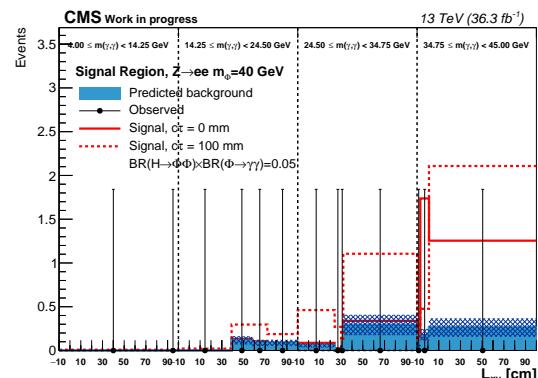
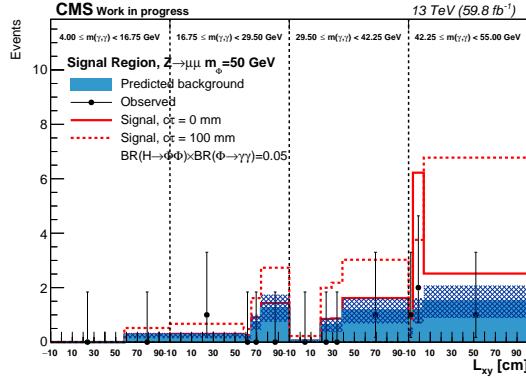
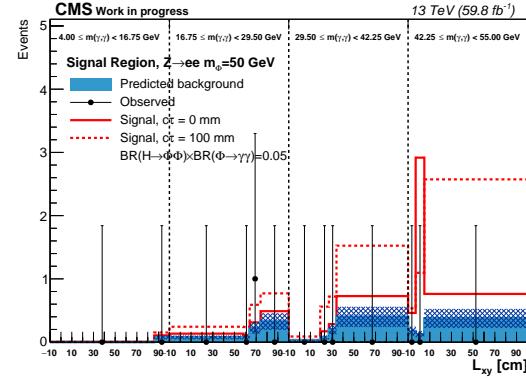


Figure 5.37: Observed number of  $m_\Phi = 40$  GeV events in the signal region of the  $Z \rightarrow \mu\mu$  category (left) and the  $Z \rightarrow ee$  category (right) across all three years.

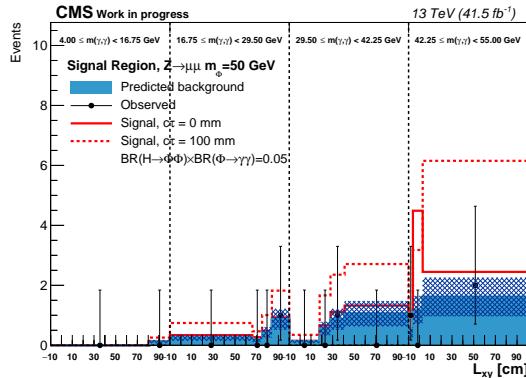
2018  $Z \rightarrow \mu\mu$



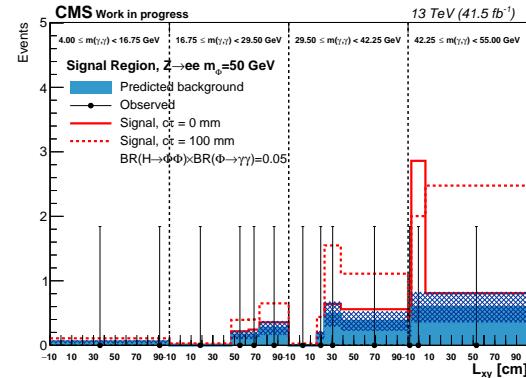
2018  $Z \rightarrow ee$



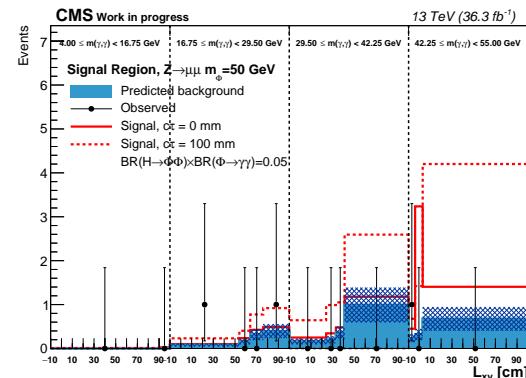
2017  $Z \rightarrow \mu\mu$



2017  $Z \rightarrow ee$



2016  $Z \rightarrow \mu\mu$



2016  $Z \rightarrow ee$

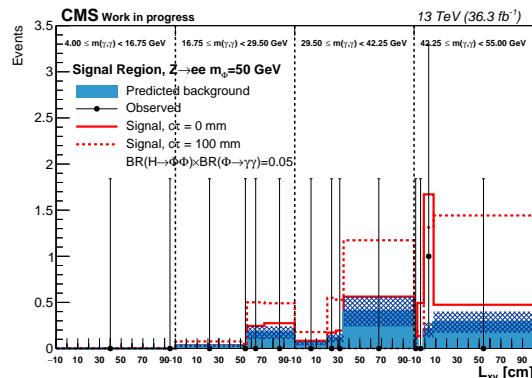
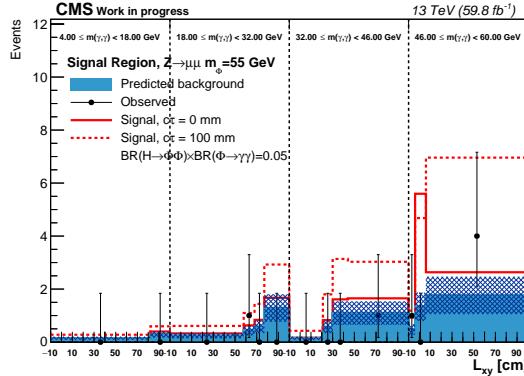
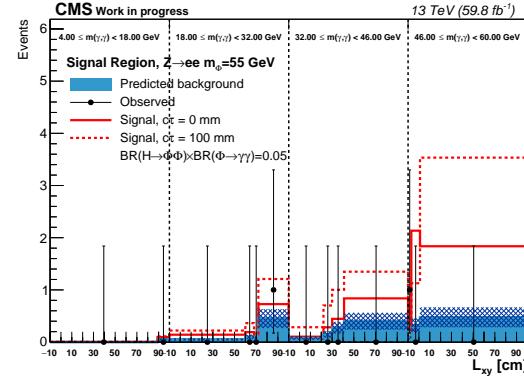


Figure 5.38: Observed number of  $m_\Phi = 50$  GeV events in the signal region of the  $Z \rightarrow \mu\mu$  category (left) and the  $Z \rightarrow ee$  category (right) across all three years.

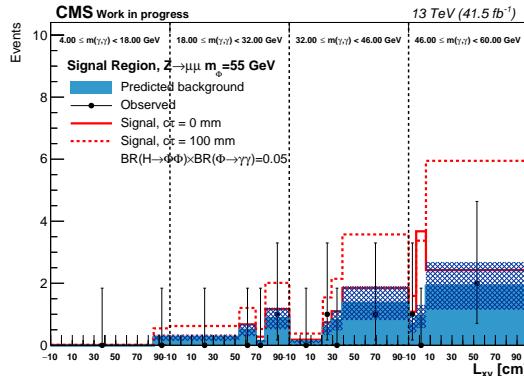
2018  $Z \rightarrow \mu\mu$



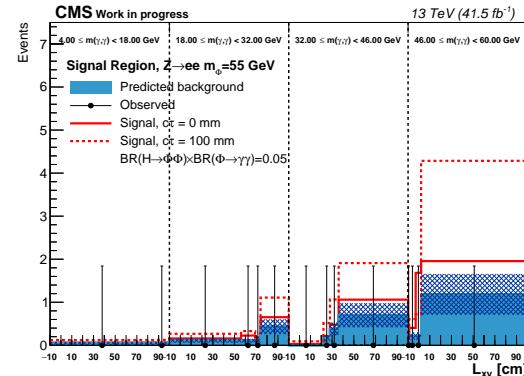
2018  $Z \rightarrow ee$



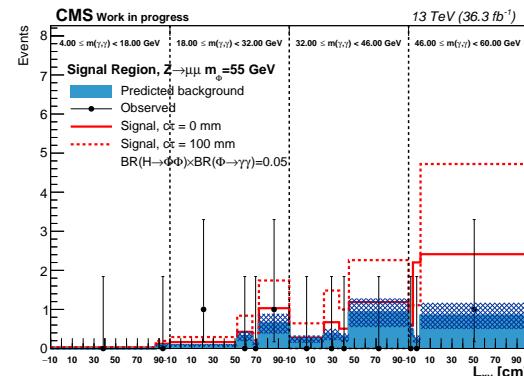
2017  $Z \rightarrow \mu\mu$



2017  $Z \rightarrow ee$



2016  $Z \rightarrow \mu\mu$



2016  $Z \rightarrow ee$

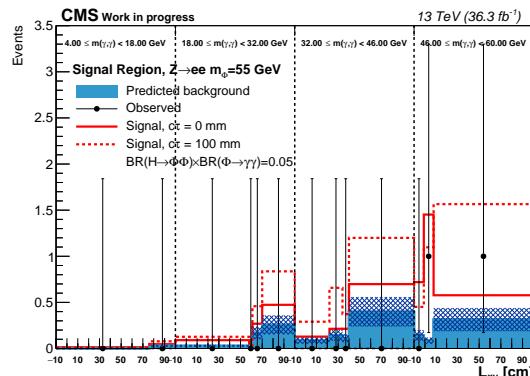


Figure 5.39: Observed number of  $m_\Phi = 55$  GeV events in the signal region of the  $Z \rightarrow \mu\mu$  category (left) and the  $Z \rightarrow ee$  category (right) across all three years.

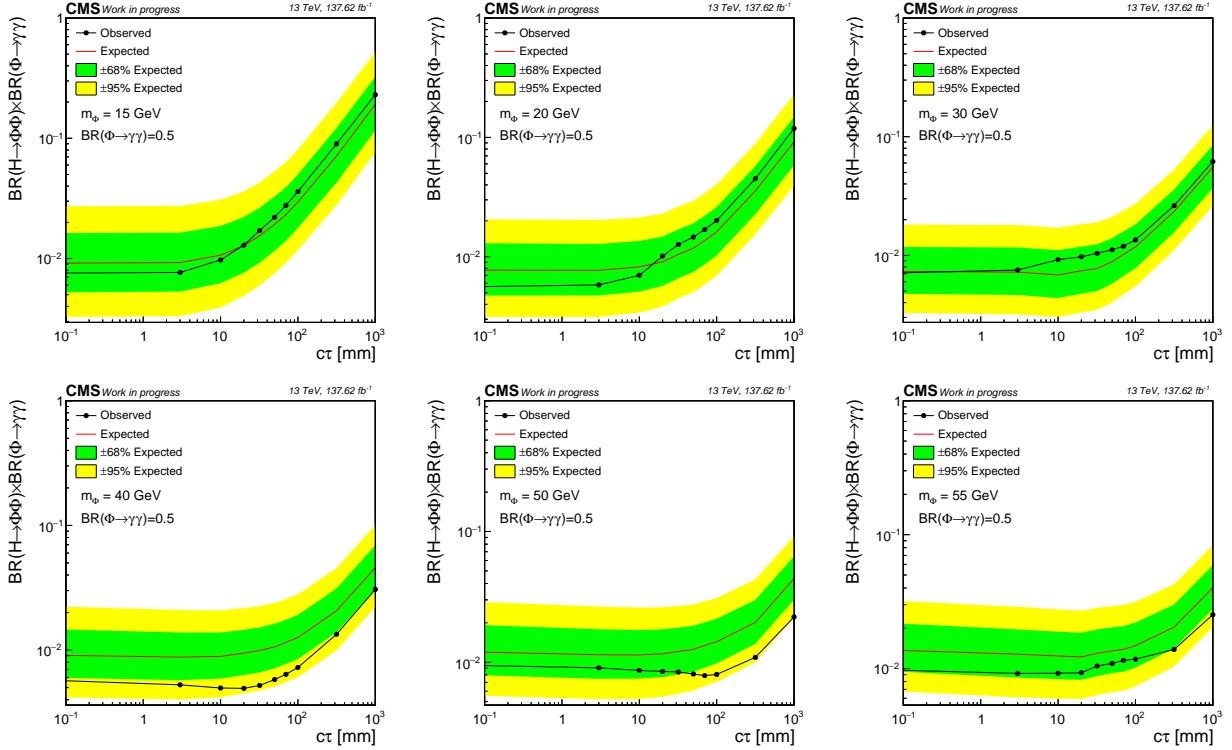


Figure 5.40: Model dependent upper limits for  $\mathcal{BR}(H \rightarrow \Phi\Phi) \times \mathcal{BR}(\Phi \rightarrow \gamma\gamma)$  for different values of  $m_\Phi$  and  $c\tau$  under the assumption that  $\mathcal{BR}(\Phi \rightarrow \gamma\gamma) = 0.5$ .

of  $m_\Phi$ . The behavior of the limits versus  $c\tau$  at low lifetimes is dominated by the intrinsic resolution of the vertex fit. Since the signal distributions vary negligibly below  $c\tau = 3$  mm, the limits also remain constant. Due to this, the range of lifetimes is only shown to a minimum of 0.1 mm but show a constant value below this. At intermediate lifetimes around 10 mm the signal distributions become distinct, but do not produce significant improvements to the upper limits. This is because the background distribution  $L_{xy}$  is not determined from the tracker information or from detector pointing capabilities but from calculations based on the energy and separation of the photons. Because of this, the expected background does not steeply fall with  $L_{xy}$ .

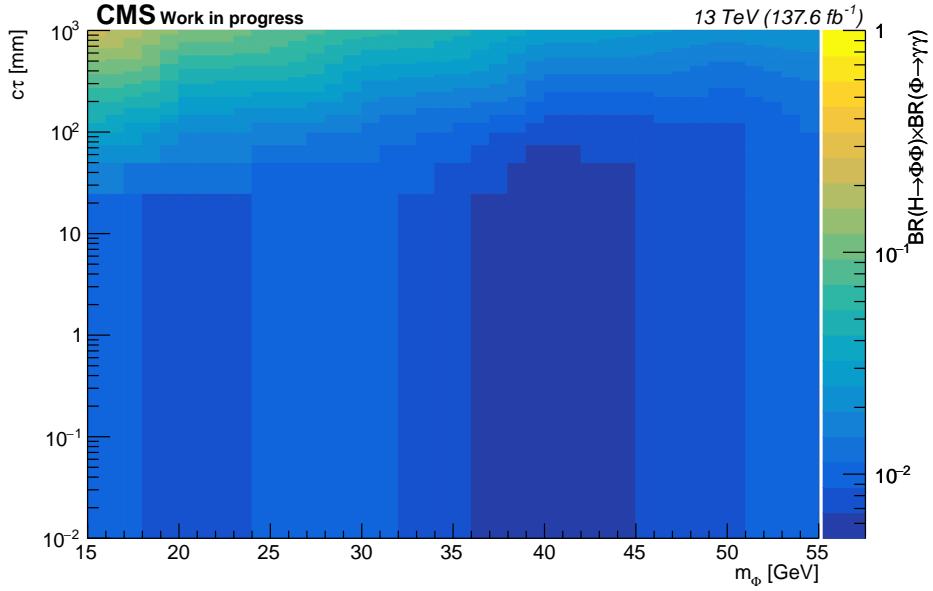


Figure 5.41: Observed limit of  $\mathcal{BR}(H \rightarrow \Phi\Phi) \times \mathcal{BR}(\Phi \rightarrow \gamma\gamma)$  as a function of  $m_\Phi$  and  $c\tau$  under the assumption that  $\mathcal{BR}(\Phi \rightarrow \gamma\gamma) = 0.5$ .

### 5.7.3 Model Independent Upper Limits

The previous model dependent limits assume that a SM Higgs boson was created and decayed directly to  $\Phi\Phi$  by normalizing the expected events using the SM cross section  $\sigma(ZH)$ . However, some models such as those discussed in chapter 2 predict the existence of BSM Higgs bosons, which would have different couplings than the SM Higgs. We present reinterpreted upper limits on a signal strength parameter  $\mu = \sigma(Z + \Phi\Phi) \times \mathcal{BR}(\Phi \rightarrow \gamma\gamma)$  by normalizing the number of expected signal events to a cross section to  $\sigma(Z + \Phi\Phi) = 1 \text{ pb}$ .

## 5.8 Summary

This analysis presents a search for rare Higgs boson decays to two long-lived scalar bosons decaying to two photons. We use data collected from 2016 to 2018 from center of mass energy  $\sqrt{s} = 13 \text{ TeV}$  proton-proton collisions corresponding to  $138 \text{ fb}^{-1}$ . A novel vertex

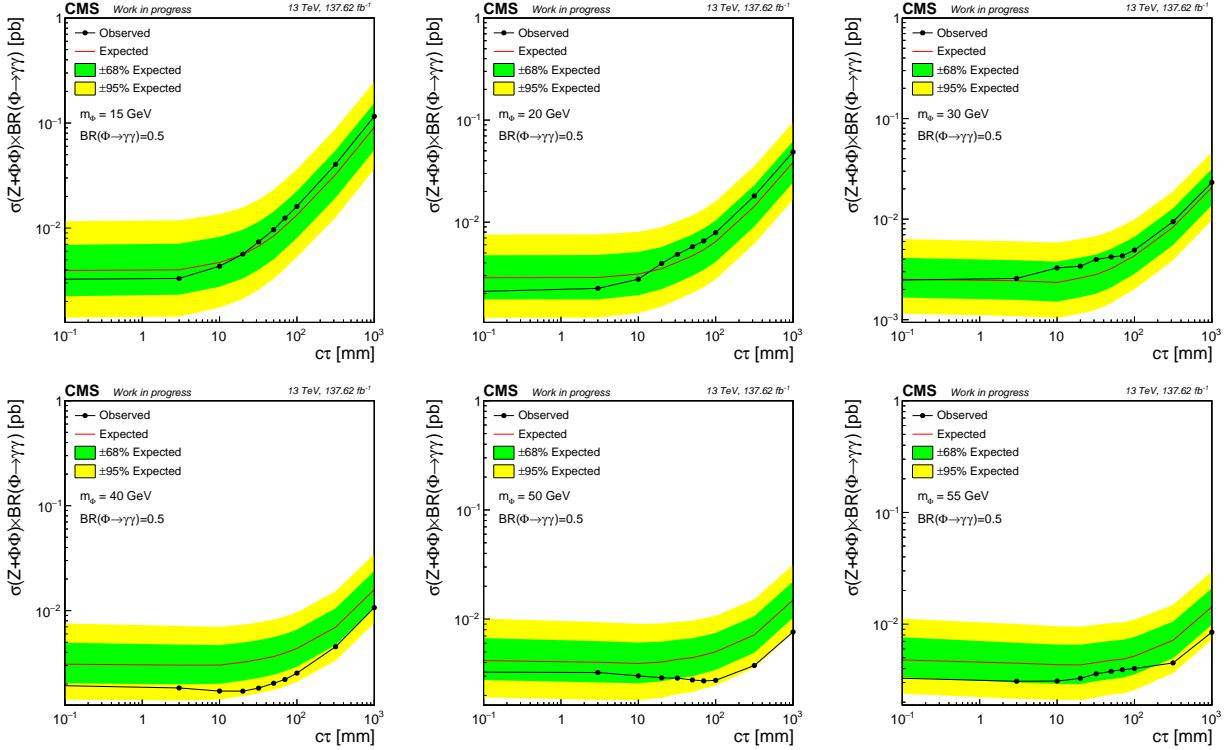


Figure 5.42: Model independent upper limits for  $\sigma(Z + \Phi\Phi) \times \mathcal{B}\mathcal{R}(\Phi \rightarrow \gamma\gamma)$  as a function of  $c\tau$  for various values of  $m_\Phi$  under the assumption that  $\mathcal{B}\mathcal{R}(\Phi \rightarrow \gamma\gamma) = 0.5$ .

reconstruction technique described in section 5.3 is used to calculate the vertex of displaced photon pairs using kinematic constraints with a resolution of  $< 3 \text{ cm}$ . The reconstructed vertices are used to extract upper limits on the signal strength using a maximum likelihood approach outlined in section 5.6. We set limits on the product of the SM Higgs boson branching ratio to long-lived scalar bosons  $\Phi$  times the branching ratio of the  $\Phi$  boson to two photons for a wide range of hypothetical  $\Phi$  masses and lifetimes. No significant excesses are observed in the data and upper limits are calculated for the product of branching ratios below 1%.

Results from this analysis validate the diphoton vertex reconstruction as a viable method to set limits on long lived photon decays. At the time of writing this dissertation, there is work in progress to extend this analysis for additional Higgs boson production mechanisms.

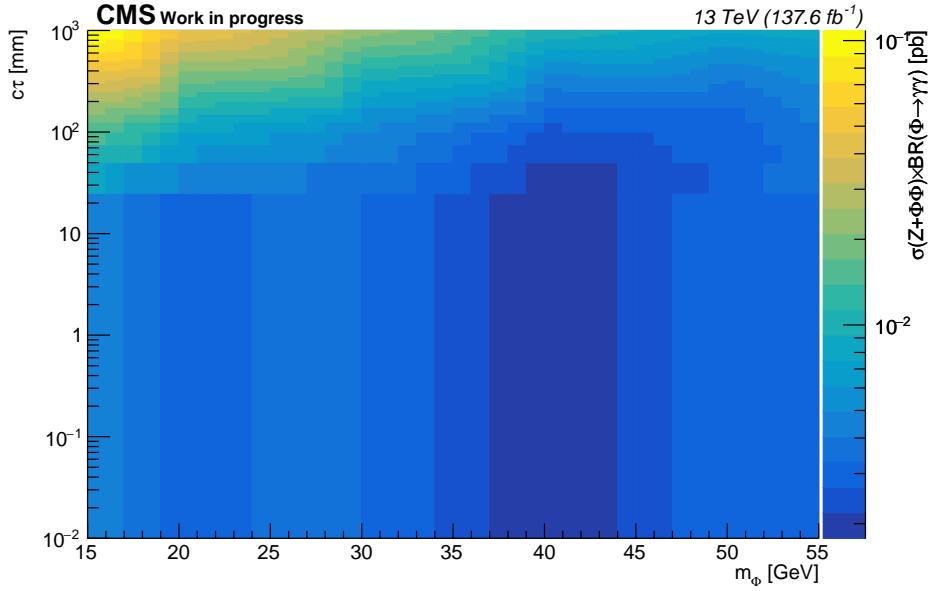


Figure 5.43: Observed model independent limits on  $\sigma(Z + \Phi\Phi) \times \mathcal{BR}(\Phi \rightarrow \gamma\gamma)$  as a function of  $m_\Phi$  and  $c\tau$  under the assumption that  $\mathcal{BR}(\Phi \rightarrow \gamma\gamma) = 0.5$ .

Higgs boson production with an associated W boson is a natural extension that can follow the event selection and signal extraction procedure outlined above. In parallel, an analysis searching for Higgs production through gluon fusion will extract the signal strength using the invariant mass of the four photons to reconstruct the Higgs boson mass. The four photon mass resolution shows improvements by correcting the photon four-momentum using the vertex constraint, which is expected to improve the expected limits. Preliminary plans to use Higgs bosons produced through vector boson fusion are ongoing. These analyses are expected to show improved results by incorporating data taken from 2022 to 2024, which currently corresponds to a recorded  $91.49 \text{ fb}^{-1}$  [1].

## REFERENCES

- [1] CMS Collaboration. CMS luminosity - public results, 2018. URL: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults>.
- [2] CMS Collaboration. The high-luminosity lhc (hl-lhc). URL: <https://voisins.web.cern.ch/en/high-luminosity-lhc-hl-lhc>.
- [3] Tai Sakuma. Cutaway diagrams of CMS detector. 2019. URL: <http://cds.cern.ch/record/2665537>.
- [4] The Tracker Group Of The CMS Collaboration. The cms phase-1 pixel detector upgrade, 2020. URL: <https://arxiv.org/abs/2012.14304>, doi:10.48550/ARXIV.2012.14304.
- [5] S Chatrchyan et al. Alignment of the CMS silicon tracker during commissioning with cosmic rays. *JINST*, 5:T03009, 2010. URL: <https://cds.cern.ch/record/1211825>, arXiv:0910.2505, doi:10.1088/1748-0221/5/03/T03009.
- [6] Badder Marzocchi. Simulation of the cms electromagnetic calorimeter response at the energy and intensity frontier. *Journal of Physics: Conference Series*, 1162:012007, 01 2019. doi:10.1088/1742-6596/1162/1/012007.
- [7] Albert M Sirunyan et al. Calibration of the CMS hadron calorimeters using proton-proton collision data at  $\sqrt{s} = 13$  TeV. *JINST*, 15:P05002, 2020. Replaced with the published version. Added the journal reference and the DOI. All the figures and tables can be found at <http://cms-results.web.cern.ch/cms-results/public-results/publications/PRF-18-001> (CMS Public Pages). URL: <https://cds.cern.ch/record/2691403>, arXiv:1910.00079, doi:10.1088/1748-0221/15/05/P05002.
- [8] A.M. Sirunyan et al. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at  $\sqrt{s} = 13$  TeV. *JINST*, 13(06):P06015, 2018. Replaced with the published version. Added the journal reference and the DOI. All the figures and tables can be found at <http://cms-results.web.cern.ch/cms-results/public-results/publications/MUO-16-001> (CMS Public Pages). URL: <http://cds.cern.ch/record/2313130>, arXiv:1804.04528, doi:10.1088/1748-0221/13/06/P06015.
- [9] CMS Collaboration. Muon drift tubes. URL: <https://cms.cern/detector/detecting-muons/muon-drift-tubes>.
- [10] Isabelle Helena J De Bruyn. Electronics upgrade for the CMS CSC muon system at the High Luminosity LHC. Technical report, CERN, Geneva, 2019. URL: <http://cds.cern.ch/record/2797803>.

- [11] Priyanka Kumari. Improved-rpc for the cms muon system upgrade for the hl-lhc, 2020. URL: <https://arxiv.org/abs/2005.11396>, doi:10.48550/ARXIV.2005.11396.
- [12] Albert M Sirunyan et al. Performance of the CMS Level-1 trigger in proton-proton collisions at  $\sqrt{s} = 13$  TeV. *JINST*, 15(10):P10017, 2020. Replaced with the published version. Added the journal reference and the DOI. All the figures and tables can be found at <http://cms-results.web.cern.ch/cms-results/public-results/publications/TRG-17-001> (CMS Public Pages). URL: <https://cds.cern.ch/record/2721198>, arXiv: 2006.10165, doi:10.1088/1748-0221/15/10/P10017.
- [13] Benjamin Radburn-smith. Overview of the hl-lhc upgrade for the cms level-1 trigger. page 639, 12 2022. doi:10.22323/1.414.0639.
- [14] The Phase-2 Upgrade of the CMS Level-1 Trigger. Technical report, CERN, Geneva, 2020. Final version. URL: <https://cds.cern.ch/record/2714892>.
- [15] Sergey Chatrchyan, Vardan Khachatryan, Albert Sirunyan, W Adam, Beth Arnold, H Bergauer, Thomas Bergauer, Marko Dragicevic, M Eichberger, J. Erö, M. Friedl, Rudolf Frühwirth, V Ghete, Josef Hammer, S. Hänsel, M Hoch, N. Hörmann, J Hrubec, M. Jeitler, and A. White. Performance of the cms drift-tube local trigger with cosmic rays. *Journal of Instrumentation*, 5:T03003, 03 2010.
- [16] Aram Hayrapetyan et al. Development of the CMS detector for the CERN LHC Run 3. Technical report, CERN, Geneva, 2023. URL: <https://cds.cern.ch/record/2870088>, arXiv:2309.05466.
- [17] Albert M Sirunyan and other. Precision luminosity measurement in proton-proton collisions at  $\sqrt{s} = 13$  TeV in 2015 and 2016 at CMS. *Eur. Phys. J. C*, 81(9):800, 2021. Replaced with the published version. Added the journal reference and the DOI. All the figures and tables can be found at <http://cms-results.web.cern.ch/cms-results/public-results/publications/LUM-17-003> (CMS Public Pages). URL: <https://cds.cern.ch/record/2759951>, arXiv:2104.01927, doi:10.1140/epjc/s10052-021-09538-2.
- [18] CMS luminosity measurement for the 2017 data-taking period at  $\sqrt{s} = 13$  TeV. Technical report, CERN, Geneva, 2018. URL: <https://cds.cern.ch/record/2621960>.
- [19] CMS luminosity measurement for the 2018 data-taking period at  $\sqrt{s} = 13$  TeV. Technical report, CERN, Geneva, 2019. URL: <https://cds.cern.ch/record/2676164>.
- [20] CMS Collaboration. How to compute cross sections with the genxsec analyzer. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/HowToGenXSecAnalyzer>.
- [21] CMS Collaboration. Cut based electron id for run 2. URL: <https://twiki.cern.ch/twiki/bin/view/CMS/CutBasedElectronIdentificationRun2>.

- [22] CMS Collaboration. Cut based photon id for run 2. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/CutBasedPhotonIdentificationRun2>.
- [23] Aram Hayrapetyan et al. The CMS statistical analysis and combination tool: COMBINE. Submitted to *Comput. Softw. Big Sci.*, 2024. arXiv:2404.06614.
- [24] CERN. Cern yellow reports: Monographs, vol 2 (2017): Handbook of lhc higgs cross sections: 4. deciphering the nature of the higgs sector, 2017. URL: <https://e-publishing.cern.ch/index.php/CYRM/issue/view/32>, doi:10.23731/CYRM-2017-002.
- [25] Oliver Sim Brüning, Paul Collier, P Lebrun, Stephen Myers, Ranko Ostojic, John Poole, and Paul Proudlock. *LHC Design Report*. CERN Yellow Reports: Monographs. CERN, Geneva, 2004. URL: <https://cds.cern.ch/record/782076>, doi: 10.5170/CERN-2004-003-V-1.
- [26] Esma Mobs. The CERN accelerator complex - august 2018. complexe des accélérateurs du CERN - août 2018. Aug 2018. General Photo. URL: <https://cds.cern.ch/record/2636343>.
- [27] The CMS Collaboration. Description and performance of track and primary-vertex reconstruction with the cms tracker. *Journal of Instrumentation*, 9(10):P10009, oct 2014. URL: <https://dx.doi.org/10.1088/1748-0221/9/10/P10009>, doi:10.1088/1748-0221/9/10/P10009.
- [28] R. L. Workman and Others. Review of Particle Physics. *PTEP*, 2022:083C01, 2022. doi:10.1093/ptep/ptac097.
- [29] E. Tournefier. The preshower detector of cms at lhc. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 461(1):355–360, 2001. 8th Pisa Meeting on Advanced Detectors. URL: <https://www.sciencedirect.com/science/article/pii/S0168900200012432>, doi: 10.1016/S0168-9002(00)01243-2.
- [30] G Baiatian et al. Design, Performance, and Calibration of CMS Hadron-Barrel Calorimeter Wedges. Technical Report 1, CERN, Geneva, 2008. URL: <http://cds.cern.ch/record/1049915>, doi:10.1140/epjc/s10052-008-0573-y.
- [31] A.M. Sirunyan et al. Particle-flow reconstruction and global event description with the cms detector. *Journal of Instrumentation*, 12(10), October 2017. URL: <http://dx.doi.org/10.1088/1748-0221/12/10/P10003>, doi:10.1088/1748-0221/12/10/p10003.
- [32] W Adam, R Frühwirth, A Strandlie, and T Todorov. Reconstruction of electrons with the gaussian-sum filter in the cms tracker at the lhc. *Journal of Physics G: Nuclear*

- and Particle Physics*, 31(9):N9–N20, July 2005. URL: <http://dx.doi.org/10.1088/0954-3899/31/9/N01>, doi:10.1088/0954-3899/31/9/n01.
- [33] A. Triossi, M. Bellato, J.M. Cela Ruiz, J. Ero, C. Fernandez Bedoya, G. Flouris, C. Foudas, L. Guiducci, N. Loukas, A. Navarro Tobar, S. Mallios, G. Masetti, F. Montecassiano, E. Paradas, I. Redondo Fernandez, D. Redondo Ferrero, J. Sastre, P. Sphicas, S. Ventura, and C. Wulz. The cms barrel muon trigger upgrade. *Journal of Instrumentation*, 12(01):C01095, jan 2017. URL: <https://dx.doi.org/10.1088/1748-0221/12/01/C01095>, doi:10.1088/1748-0221/12/01/C01095.
  - [34] R. Frühwirth. Application of kalman filtering to track and vertex fitting. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 262(2):444–450, 1987. URL: <https://www.sciencedirect.com/science/article/pii/0168900287908874>, doi:10.1016/0168-9002(87)90887-4.
  - [35] Michail Bachtis. Upgrade of the CMS Barrel Muon Track Finder for HL-LHC featuring a Kalman Filter algorithm and an ATCA Host Processor with Ultrascale+ FPGAs. Technical report, CERN, Geneva, 2019. URL: <https://cds.cern.ch/record/2648953>, doi:10.22323/1.343.0139.
  - [36] J. Ero, I. Evangelou, G. Flouris, C. Foudas, L. Guiducci, N. Loukas, N. Manthos, I. Papadopoulos, E. Paradas, S. Sotiropoulos, P. Sphicas, A. Triossi, and C. Wulz. The cms level-1 trigger barrel track finder. *Journal of Instrumentation*, 11(03):C03038, mar 2016. URL: <https://dx.doi.org/10.1088/1748-0221/11/03/C03038>, doi:10.1088/1748-0221/11/03/C03038.
  - [37] CMS Collaboration. Pdmvgroupdescription. URL: <https://twiki.cern.ch/twiki/bin/viewauth/CMS/PdmVGroupDescription>.
  - [38] Simulation of the Silicon Strip Tracker pre-amplifier in early 2016 data. 2020. URL: <https://cds.cern.ch/record/2740688>.
  - [39] Richard D. Ball, Valerio Bertone, Stefano Carrazza, Luigi Del Debbio, Stefano Forte, Patrick Groth-Merrild, Alberto Guffanti, Nathan P. Hartland, Zahari Kassabov, José I. Latorre, Emanuele R. Nocera, Juan Rojo, Luca Rottoli, Emma Slade, and Maria Ubiali. Parton distributions from high-precision collider data: Nnpdf collaboration. *The European Physical Journal C*, 77(10), October 2017. URL: <http://dx.doi.org/10.1140/epjc/s10052-017-5199-5>, doi:10.1140/epjc/s10052-017-5199-5.
  - [40] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re. A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX. *JHEP*, 06:043, 2010. arXiv:1002.2581, doi:10.1007/JHEP06(2010)043.

- [41] Christian Bierlich, Smita Chakraborty, Nishita Desai, Leif Gellersen, Ilkka Helenius, Philip Ilten, Leif Lönnblad, Stephen Mrenna, Stefan Prestel, Christian T. Preuss, Torbjörn Sjöstrand, Peter Skands, Marius Utheim, and Rob Verheyen. A comprehensive guide to the physics and usage of pythia 8.3, 2022. [arXiv:2203.11601](https://arxiv.org/abs/2203.11601).
- [42] A. M. Sirunyan, A. Tumasyan, W. Adam, et al. Extraction and validation of a new set of cms pythia8 tunes from underlying-event measurements. *The European Physical Journal C*, 80(1), Jan 2020. doi:10.1140/epjc/s10052-019-7499-4.
- [43] A.M. Sirunyan et al. Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC. *Journal of Instrumentation*, 16(05):P05014, may 2021. URL: <https://doi.org/10.1088/2F1748-0221%2F16%2F05%2Fp05014>, doi:10.1088/1748-0221/16/05/p05014.
- [44] Marco Peruzzi, Giovanni Petrucciani, Andrea Rizzi, and for the CMS Collaboration. The nanoaod event data format in cms. *Journal of Physics: Conference Series*, 1525(1):012038, apr 2020. URL: <https://dx.doi.org/10.1088/1742-6596/1525/1/012038>, doi:10.1088/1742-6596/1525/1/012038.
- [45] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells. Asymptotic formulae for likelihood-based tests of new physics. *The European Physical Journal C*, 71(2), February 2011. URL: <http://dx.doi.org/10.1140/epjc/s10052-011-1554-0>, doi:10.1140/epjc/s10052-011-1554-0.