

MEASURING PROPERTIES OF THE HIGGS BOSON UTILIZING ITS OFF-SHELL PRODUCTION

by
Lucas Kang

A dissertation submitted to The Johns Hopkins University in conformity
with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
August 2025

© 2025 Lucas Kang
All rights reserved

Abstract

The Higgs boson is the last observed piece of the Standard Model of particle physics, and due to its couplings to other fundamental particles, provides a promising portal through which we can potentially probe new physics beyond the Standard Model. This thesis revolves around an analysis which incorporates the full Run 2 dataset from the CMS experiment at the LHC to make precision measurements of the Higgs boson, utilizing a combination of its on- and off-shell production. It also sets the stage for novel extensions to analysis of Higgs boson production in the off-shell region, so that future endeavors can better constrain coupling strengths between the Higgs boson and other particles in the search for Beyond-the-Standard-Model (BSM) physics.

Keywords: Higgs, boson, precision, measurement, width, off-shell, golden channel, HVV, kappa, anomalous, coupling, Standard Model, SM, BSM, CMS, LHC, CERN.

Primary Reader and Advisor: Andrei Gritsan

Secondary Readers: Toni Šćulac, Morris Swartz, Robert Leheny, Kobe Marshall-Stevens

*This thesis is dedicated to my mother
and all others who guided the way*

Acknowledgements

There are so many people who contributed to the work described in this thesis. But certainly, none of this could have been possible without a steady hand on the tiller and a steely vision of the research landscape.

Andrei, thank you so much for your relentless support and zeal for good science. You have been an inspiration many times during my PhD, and it must be made clear how appreciated you are in your role as a researcher, advisor, and guiding force.

Savvas, our dear postdoc, thank you for the effort you put into our analyses and group. You kept things moving, even when unexpected challenges presented themselves, and you reliably provided a voice of reason when the going got tough. You were a collaborator and friend the whole way, as we figured things out together.

Mohit, my off-shell heir apparent, I will forever appreciate your skills, camaraderie, and wit. Same goes for Nick, Zhiyuan, and Jeff, my CMS comrades in the trenches. Thank you all for making this group what it is now, and for being fantastic colleagues and companions these past few years.

Toni, how great is it that we got to spend so much time with you. Thank you for the chats, advice, and motivation over the years. You've always been a lovely collaborator but now you've added incredible liveliness and enthusiasm to our group and it's a blessing to end my PhD journey on such a note.

I might as well mention others in the Hopkins HEP family: Morris, Petar, Danielle(s),

Acknowledgements

thank you for the conversations and insights; even with our limited cross-section, it was very enjoyable and I cherish what I've learned.

Michalis, Lucas G., Valdis, Ruoxi, Annie, Sanjana, Tamas, Lucas C., Oz, it was a pleasure to overlap as much as we did. I will treasure the time we spent together and am grateful that we occasionally even got to work together.

Reid, Toby, Bob, thank you as well for the conversations we have had. I really valued having your perspectives, unconditional support, and continued advice.

And, of course, I need to thank my past academic advisors, who have inspired and molded my academic journey. Stephen, Meenakshi, Jim, you have taught and helped me greatly throughout my research career. Thank you, sincerely.

Finally, I wouldn't be here without my family. Mom, thank you for consistently putting in the work and enriching our lives. You always prioritized our education and interests above all, and I owe everything to you. Dad, I appreciate your constant passion and exploration. And Louis and Lawrence, I will never stop admiring your tenacity and brilliance—you both inspire me everyday.

Hanc marginis exiguitas non caperet.

– PIERRE DE FERMAT

Table of Contents

| | |
|---|------------|
| Abstract | ii |
| Dedication | iii |
| Acknowledgements | iv |
| Epigraph | vi |
| List of Tables | x |
| List of Figures | xii |
| Chapter 1 Introduction | 1 |
| Chapter 2 The Theory | 3 |
| 2.1 The Standard Model of particle physics | 3 |
| 2.1.1 Elementary particles | 3 |
| 2.1.2 Forces and symmetries | 4 |
| 2.2 The Higgs mechanism | 7 |
| 2.3 Phenomenology of the off-shell Higgs boson | 14 |
| 2.3.1 Production modes | 15 |
| 2.3.2 Decay modes | 17 |
| 2.4 Properties of the Higgs boson | 19 |
| 2.5 Off-shell technique | 21 |
| 2.6 Higgs boson couplings | 23 |
| 2.6.1 Amplitude EFT couplings | 24 |
| 2.6.2 Kappa framework couplings | 27 |
| Chapter 3 The Experiment | 29 |
| 3.1 The Large Hadron Collider (LHC) | 29 |

Table of Contents

| | | |
|------------------|--|-----------|
| 3.2 | The Compact Muon Solenoid (CMS) | 34 |
| 3.2.1 | Solenoid | 36 |
| 3.2.2 | Silicon tracker | 38 |
| 3.2.3 | Electromagnetic calorimeter (ECAL) | 43 |
| 3.2.4 | Hadron calorimeter (HCAL) | 45 |
| 3.2.5 | Muon system | 46 |
| 3.3 | Triggers and Data Acquisition (DAQ) system | 53 |
| 3.3.1 | Level-1 (L1) trigger | 54 |
| 3.3.2 | High-Level Trigger (HLT) | 55 |
| 3.4 | Computing | 56 |
| 3.5 | Particle reconstruction | 57 |
| 3.6 | Alignment of the CMS tracker | 58 |
| 3.6.1 | Parameterization | 58 |
| 3.6.2 | Algorithms | 61 |
| 3.6.3 | Validation | 66 |
| 3.6.4 | Alignment and validation with HipPy | 76 |
| Chapter 4 | Analysis of Higgs boson data at the LHC | 81 |
| 4.1 | Modeling of off-shell production | 83 |
| 4.1.1 | K factors | 85 |
| 4.1.2 | Corrections for jet modeling | 91 |
| 4.2 | Building a $H^{(*)} \rightarrow ZZ^{(*)} \rightarrow 4\ell$ analysis | 93 |
| 4.2.1 | Reconstructing and selecting events | 94 |
| 4.2.2 | Matrix Element Likelihood Approach | 98 |
| 4.2.3 | Defining categories and observables | 99 |
| 4.2.4 | Modeling of background | 110 |
| 4.2.5 | Parameterizing the on-shell Higgs boson | 111 |
| 4.2.6 | Parameterizing the off-shell Higgs boson | 112 |
| 4.2.7 | Systematic uncertainties | 116 |
| 4.3 | Measurement of Higgs boson properties | 118 |

Table of Contents

| | | |
|------------------|---|------------|
| 4.3.1 | Higgs boson width | 118 |
| 4.3.2 | Off-shell signal strength | 122 |
| 4.3.3 | Kappa framework | 124 |
| Chapter 5 | Conclusion | 126 |
| | Bibliographic references | 127 |

List of Tables

| | | |
|------------------|---|-----|
| Table 3.1 | Design parameters for the LHC [11] | 31 |
| Table 4.1 | The calculation for the NNLO to N ³ LO K factor for the gluon fusion MC samples. The factor calculated at 125 GeV is applied to all the gluon fusion processes in the full off-shell region. . . . | 86 |
| Table 4.2 | The calculation for the LO to NNLO K factor for the EW off-shell samples at the Higgs boson mass peak (125 GeV). The calculated factor is applied to all the off-shell EW MC samples. | 90 |
| Table 4.3 | Generator-level cuts on jets and leptons in all utilized samples. Note that Phantom provides EW simulation but MCFM is used for both EW and ggH events. | 92 |
| Table 4.4 | Efficiencies of generator-level requirements on leptons ($p_T(\ell_{1,2,3,4}) > 3$ GeV, $ \eta(\ell_{1,2,3,4}) < 2.7$) in the off-shell EW samples as a function of $m_{4\ell}$. The EW (had) includes VBF and hadronic VH production, while EW (full) also includes leptonic VH production. | 92 |
| Table 4.5 | Efficiencies of generator-level requirements on jets in the off-shell EW samples as a function of $m_{4\ell}$. The EW (had) includes VBF and hadronic VH production, while EW (full) also includes leptonic VH production. These efficiencies are calculated after generator-level requirements had been applied on leptons, as indicated in Table 4.4. The last two columns incorporate efficiency of lepton cuts together with jet cuts. | 93 |
| Table 4.6 | Summary of the three production categories in the off-shell $m_{4\ell}$ region. All discriminants are calculated with the JHUGen signal and MCFM background matrix elements. The VH interference discriminant in the VH-tagged category is the simple average of the ones corresponding to the ZH and WH processes [100]. . . . | 102 |

| | | |
|-------------------|--|-----|
| Table 4.7 | Observed and expected yields for the Higgs boson signal and background contributions in the off-shell region $m_{4\ell} > 220$ GeV, for each of the four-lepton categories and the total. The yields from interference of the signal and background and the ZH cross-feed are also shown. The expected yields are adjusted within their respective uncertainties from the fit to the data [100]. . . . | 111 |
| Table 4.8 | Summary of the ReReco-based total Higgs boson width Γ_H measurement, showing the 68% CL (central values with uncertainties) and 95% CL (in square brackets) intervals for the $H \rightarrow ZZ \rightarrow 4\ell$ channel in combination with the off-shell $H \rightarrow ZZ \rightarrow 2\ell 2\nu$ channel [100]. | 119 |
| Table 4.9 | Summary of all UL-based results. Total Higgs boson width Γ_H and signal strength μ_j measurements, with 68% CL (central values with uncertainties) and 95% CL (in square brackets) intervals for the $H \rightarrow ZZ \rightarrow 4\ell$ channel. | 122 |
| Table 4.10 | Summary of constraints at 68% CL (central values with uncertainties) and 95% CL (in square brackets) on the parameters μ_F and μ_V with and without including κ_Q unconstrained [100]. . . | 125 |

List of Figures

| | | |
|-------------------|---|----|
| Figure 2.1 | Particle content of the Standard Model. | 4 |
| Figure 2.2 | Visualization of the characteristic “Mexican hat” shaped Higgs potential, as parameterized by the two fields in the Higgs doublet. | 10 |
| Figure 2.3 | SM Higgs boson production cross sections as a function of m_H at the LHC at $E_{CM} = 13$ TeV [4]. | 16 |
| Figure 2.4 | Feynman diagrams for the dominant Higgs boson production modes at the LHC: (a) gluon fusion (ggH), (b) vector boson fusion (VBF), (c) associated (VH) production, and (d) $t\bar{t}H$ production. | 17 |
| Figure 2.5 | SM Higgs boson decay branching ratios (a) and total decay width (b) as a function of mass m_H [4]. In the off-shell region above 220 GeV, the Higgs boson decay to a pair of vector bosons dominates. We see an enhancement of the Higgs boson width at energies above $2m_V$ when the two vector bosons become on-shell. | 18 |
| Figure 2.6 | Feynman diagram for the decay $H^{(*)} \rightarrow ZZ^{(*)} \rightarrow 4\ell$ | 19 |
| Figure 2.7 | A published sketch of the Higgs boson mass distribution reconstructed from ZZ decays. The peak on the left marks on-shell production while the broad shoulder on the right represents the off-shell process. Beware that the corresponding rates indicated here are not to scale. | 21 |
| Figure 2.8 | Example Feynman diagrams for gluon fusion off-shell Higgs boson production (right) and corresponding background (left) with EFT insertions [8]. | 25 |
| Figure 2.9 | Example Feynman diagrams for electroweak off-shell Higgs boson production (left), vector boson scattering (VBS) (center), and additional background (right) with EFT insertions [8]. | 26 |

List of Figures

| | | |
|--------------------|--|----|
| Figure 3.1 | An aerial view of CERN and the LHC, located near Geneva on the border between France and Switzerland. | 30 |
| Figure 3.2 | Cumulative luminosity versus day delivered to CMS during stable beams for pp collisions at nominal center-of-mass energy. This is shown for data-taking in 2010 (green), 2011 (red), 2012 (blue), 2015 (purple), 2016 (orange), 2017 (light blue), 2018 (navy blue), 2022 (brown), and 2023 (light purple). These plots use the best available offline calibrations for each year. | 35 |
| Figure 3.3 | The CMS detector is shaped like a cylindrical onion, with several concentric layers of components. | 36 |
| Figure 3.4 | View of the CMS magnet during construction. | 37 |
| Figure 3.5 | Schematic view of the CMS detector. | 38 |
| Figure 3.6 | Illustration of the differences between the original design of the silicon tracker and the Phase-1 upgrade [17, 19, 20]. This thesis utilized data from LHC Run 2 after the Phase-1 upgrades to the tracker. | 39 |
| Figure 3.7 | Schematic view of one quarter of the tracker in the r-z plane. This details the full silicon detector with all modules in their original positions. | 39 |
| Figure 3.8 | Silicon pixel detector before installation. | 40 |
| Figure 3.9 | Illustration of the CMS silicon sensor modules. | 41 |
| Figure 3.10 | CMS Tracker layers shown in the plane perpendicular to the beam. | 42 |
| Figure 3.11 | Photograph of the silicon strip detectors in the barrel module. | 43 |
| Figure 3.12 | Photograph of the CMS ECAL during construction. | 44 |
| Figure 3.13 | Insertion of the HCAL barrel inside the magnet in 2008. | 45 |

List of Figures

| | | |
|--------------------|---|----|
| Figure 3.14 | A muon, in the plane perpendicular to the LHC beams, leaves a curved trajectory in four layers of muon “stations.” Note that the muon passes through the silicon tracker first, which measures momentum, charge, trajectory, and impact parameter, or how close the muon’s path is to the primary vertex. | 48 |
| Figure 3.15 | Hierarchical structures of the pixel detector and strip tracker [41]. | 59 |
| Figure 3.16 | Simple illustration of a track through misaligned layers [43]. . . | 59 |
| Figure 3.17 | Example local coordinates of a module. Global parameters are shown in parentheses for modules in the TIB and TOB [41]. . | 60 |
| Figure 3.18 | Process diagram of the Hits-and-Impact-Points (HIP) alignment algorithm [43]. | 63 |
| Figure 3.19 | Process diagram of the Hits-and-Impact-Points-Past-Year-1 (HipPy) alignment algorithm [43]. | 64 |
| Figure 3.20 | Example global tracker plot for GC validation, comparing analogous HipPy and Millepede alignments using beam data, both calculated from PromptGT. | 67 |
| Figure 3.21 | Example DMR validation plot for x' residual in the BPIX, comparing HipPy (blue) and Millepede (red) alignments using beam data, both calculated from the starting geometry of PromptGT (black). | 69 |
| Figure 3.22 | Example DMR validation plot for x' residual in the FPIX, comparing HipPy (blue) and Millepede (red) alignments using beam data, both calculated from the starting geometry of PromptGT (black). | 70 |
| Figure 3.23 | Example DMR validation plot for y' residual in the BPIX, comparing HipPy (blue) and Millepede (red) alignments using beam data, both calculated from the starting geometry of PromptGT (black). | 71 |

| | | |
|--------------------|--|----|
| Figure 3.24 | Example DMR validation plot for y' residual in the FPIX, comparing HipPy (blue) and Millepede (red) alignments using beam data, both calculated from the starting geometry of PromptGT (black). | 72 |
| Figure 3.25 | Example PV validation plot for PIX, comparing HipPy (blue) and Millepede (red) alignments using beam data, both calculated from the starting geometry of PromptGT (black). | 73 |
| Figure 3.26 | Example PV validation plot for BPIX Layer 1, comparing HipPy (blue) and Millepede (red) alignments using beam data, both calculated from the starting geometry of PromptGT (black). | 74 |
| Figure 3.27 | Muons generated by cosmic rays are extremely important for the commissioning and the alignment of CMS. Despite being 100 m underground, every second hundreds of muons generated on top of the atmosphere cross CMS. Presented are two views, transverse (left) and longitudinal (right), of a single cosmic ray event [49]. | 75 |
| Figure 3.28 | Example DMR validation plot for the BPIX in local x' coordinate, comparing HipPy (cyan) and Millepede (yellow) alignments using early 2023 CRUZET data, both calculated from the starting geometry of mp3628 (green). | 78 |
| Figure 4.1 | The four-lepton $m_{4\ell}$ invariant mass distributions for gluon fusion (left) and electroweak production in association with two jets (right) at the LHC with a 13 TeV pp collision energy. The total SM production (“H+bkg+int”) and background-only (“bkg”) components are shown in black. Three operators $c_{z\Box}$ (magenta), c_{zz} (blue), and \tilde{c}_{zz} (red) are shown in color, as analogs of the anomalous couplings described in Section 2.6.1, and they are introduced in place of the SM interaction with their strength constrained to reproduce the SM cross section of the on-shell Higgs boson signal production [8]. | 84 |
| Figure 4.2 | LO-to-NNLO and LO-to-NLO K factors plotted as functions of invariant mass for ggH production signal at $\sqrt{s} = 13$ TeV, cited from YR4 [4]. | 87 |

| | | |
|-------------------|---|-----|
| Figure 4.3 | Illustrations of the NLO-to-NNLO QCD K factor for the $q\bar{q} \rightarrow ZZ$ background (left) and LO-to-NLO EW K factor as a function of invariant mass (right). Both plots take $\sqrt{s} = 13$ TeV [89, 91, 92]. | 91 |
| Figure 4.4 | Diagrams with relevant kinematic observables labeled in the VBF (left), VH (center), and ggH (right) production modes of the Higgs boson. | 99 |
| Figure 4.5 | Distributions of the observed (points) and predicted (stacked histograms) $\mathcal{D}_{\text{bkg}}^{\text{kin}}$ of the four-lepton system, in the inclusive final state. The predictions for the Higgs boson signal and the three main backgrounds are given by the different colors. The vertical bars on the points show the statistical uncertainties in the data [100]. | 103 |
| Figure 4.6 | Corrections as a function of $m_{4\ell}$ for off-shell EW process selection efficiencies in the Untagged, VBF-tagged, and VH-tagged categories. A polynomial fit is performed extending up to 2.5 TeV. Events with $m_{4\ell} > 2.5$ TeV assume the correction value at 2.5 TeV. The corrections are shown for all of Run 2 combined. | 106 |
| Figure 4.7 | Comparison of $\mathcal{D}_{\text{bkg}}^{\text{VH+dec}}$ and $\mathcal{D}_{\text{bsi}}^{\text{VH+dec}}$ distributions of ggH events from JHUGen and POWHEG for $m_{4\ell} > 220$ GeV (left) and $250 < m_{4\ell} < 370$ GeV (right). | 107 |
| Figure 4.8 | Comparison of $\mathcal{D}_{\text{bkg}}^{\text{VBF+dec}}$ and $\mathcal{D}_{\text{bsi}}^{\text{VBF+dec}}$ distributions of ggH events from JHUGen and POWHEG for $m_{4\ell} > 220$ GeV (left) and $250 < m_{4\ell} < 370$ GeV (right). | 108 |

- Figure 4.9** Off-shell data (points) and pre-fit distributions (histograms) for the Untagged (left), VBF-tagged (middle), and VH-tagged (right) categories. The upper row shows $m_{4\ell}$ distributions with a requirement on $\mathcal{D}_{\text{bkg}}^{\text{kin}} > 0.6$ (left), $\mathcal{D}_{\text{bkg}}^{\text{VBF+dec}} > 0.6$ (middle), or $\mathcal{D}_{\text{bkg}}^{\text{VH+dec}} > 0.6$ (right) applied for illustration purposes to enhance signal over background contributions. The middle row shows $\mathcal{D}_{\text{bkg}}^{\text{kin}}$ (left), $\mathcal{D}_{\text{bkg}}^{\text{VBF+dec}}$ (middle), $\mathcal{D}_{\text{bkg}}^{\text{VH+dec}}$ (right) distributions, where an additional requirement $m_{4\ell} > 340$ GeV is applied to enhance signal-over-background contributions. The lower row plots the \mathcal{D}_{bsi} with both the $m_{4\ell}$ and $\mathcal{D}_{\text{bkg}}^{\text{kin}}$ requirements specified above. Contributions from the four processes are shown by the different colors, where “s”, “b”, and “i” refer to the signal, background, and interference contributions, respectively. The vertical bars on the points give the statistical uncertainties in the data, and the horizontal bars represent the bin widths. For the pre-fit distributions, the different cross sections are set to their SM values [100]. 109
- Figure 4.10** Left: Distributions over $m_{4\ell}$ of on-shell cross-feed events, off-shell signal, and the dominant $q\bar{q} \rightarrow 4\ell$ background process (scaled down by three orders of magnitude). Right: Distribution of $m_{2\ell 2q}$ events reconstructed in the off-shell region of the $H \rightarrow 4\ell$ analysis. Note the peak at 125 GeV with negligible yield elsewhere. 115
- Figure 4.11** Observed impacts of nuisance parameters on the Higgs boson width. 115
- Figure 4.12** ReReco-based observed (solid) and expected (dashed) profile likelihood projections from the Higgs boson width fit using on-shell [80] and off-shell $H \rightarrow ZZ \rightarrow 4\ell$ production, combined with off-shell $H \rightarrow ZZ \rightarrow 2\ell 2\nu$ [69]. The black horizontal dashed lines show the 68% and 95% CL values [100]. 120
- Figure 4.13** Observed (solid) and expected (dashed) likelihood projections from the Higgs boson width fit using on- and off-shell production, conducted with UL processed off-shell samples. The black horizontal dashed lines show the 68% and 95% CL values. . . . 121

- Figure 4.14** Observed and expected profile likelihood scans for $\mu_F^{\text{off-shell}}$ (top-left), $\mu_V^{\text{off-shell}}$ (top-right), and $\mu^{\text{off-shell}}$ (bottom-left) conducted using UL processed samples. Additionally plotted is the observed 2D profile likelihood projection (bottom-right) of the off-shell signal strength parameters ($\mu_F^{\text{off-shell}}$, $\mu_V^{\text{off-shell}}$) from the fit to the off-shell $H \rightarrow ZZ \rightarrow 4\ell$ channel. The best fit value is shown by the black cross and the SM prediction by the red x. The 68% and 95% CL contours are given by the dashed and solid curves, respectively. The color scale to the right of the plot relates the quantitative values. 123

Chapter 1

Introduction

In a sense, Democritus can be regarded as the intellectual father of what we now recognize as “particle physics.” The ancient Greek atomists, including Democritus and his mentor Leucippus, proposed that all substances observed in the natural world could be reduced to a set of indivisible, fundamental constituents they termed “atoms.” According to their philosophical framework, these atoms were eternal, indestructible, and varied only in shape, size, and motion, thereby giving rise to the macroscopic diversity of materials. This notion, while purely speculative, laid the groundwork for later scientific developments concerning the composition of matter [1].

By the 19th century, the concept of atoms had been refined and formalized within the framework of chemistry, where the term “atom” came to denote the smallest unit of a chemical element that retains its identity in a reaction. However, the atomic theory of chemistry, as it developed through the work of figures like John Dalton, Dmitri Mendeleev, and later Niels Bohr, eventually revealed that these so-called atoms were not, in fact, indivisible. Advances in experimental physics, particularly the discovery of the electron by J.J. Thomson in 1897 and the atomic nucleus by Ernest Rutherford in 1911, demonstrated that what chemists referred to as “atoms” were themselves composite structures, composed of subatomic particles—electrons, protons, and neutrons [1]. Thus, the original atomist vision of fundamental, structureless building blocks found a more fitting realization not at the atomic scale, but in what modern physics refers to as elementary particles.

To contextualize the significance of the analysis presented in this thesis, it is essential to first examine the Standard Model of particle physics. The Standard Model represents the most comprehensive theoretical framework for classifying and describing the fundamental particles and their interactions, as currently understood. Developed over the course of the mid-20th century, with its foundations laid in quantum field theory and gauge symmetries, the Standard Model reached its modern formulation in the 1970s. It successfully unifies three of the four fundamental forces—the electromagnetic force, the weak nuclear force, and the strong nuclear force—within a single theoretical structure, while classifying all known elementary particles into two broad categories: fermions, which constitute matter, and bosons, which mediate interactions.

Despite its tremendous success, both in terms of predictive power and its alignment with experimental data, the Standard Model is known to be incomplete. While it provides an extraordinarily accurate description of particle interactions at accessible energy scales, it does not incorporate gravity, nor does it account for dark matter, dark energy, or the observed mass hierarchy of elementary particles. Nevertheless, it remains the most rigorously tested and empirically validated theory of fundamental particles to date. The discovery of the Higgs boson at the Large Hadron Collider in 2012 provided the final experimental confirmation of the Standard Model’s mechanism for mass generation of elementary particles, reinforcing its status as a cornerstone of modern physics. Yet, ongoing research continues to probe the limitations of this framework, seeking insights that may lead to a deeper, more fundamental theory that extends beyond the Standard Model.

Chapter 2

The Theory

2.1 The Standard Model of particle physics

2.1.1 Elementary particles

Elementary particles are the most fundamental constituents of the universe, forming the basis of all known matter and mediating the fundamental forces that govern physical interactions. These particles, as described by the Standard Model (SM) of particle physics illustrated in Figure 2.1, are not composed of any smaller constituents that we know of and are broadly classified into fermions, which constitute matter, and bosons, which mediate interactions.

Fermions are divided into two distinct families: quarks and leptons. A fundamental property of quarks is color charge, which confines them within composite particles called hadrons, preventing their isolation in nature. They come in six flavors—up, down, charm, strange, top, and bottom—each possessing unique masses and charge properties. Quarks are susceptible to all fundamental forces, but interact predominantly via the strong force, which grows stronger as quarks attempt to separate, ensuring that they exist as elementary components of color-neutral bound states such as baryons (e.g., protons and neutrons) and mesons (e.g., pions).

Leptons, unlike quarks, are elementary particles that do not experience the strong force.

| | | | | | |
|----------------|---|---------------------------------------|--------------------------------------|------------------------|-------------------------------|
| mass → | $\approx 2.3 \text{ MeV}/c^2$ | $\approx 1.275 \text{ GeV}/c^2$ | $\approx 173.07 \text{ GeV}/c^2$ | 0 | $\approx 126 \text{ GeV}/c^2$ |
| charge → | $2/3$ | $2/3$ | $2/3$ | 0 | 0 |
| spin → | $1/2$ | $1/2$ | $1/2$ | 1 | 0 |
| | u up | c charm | t top | g gluon | H Higgs boson |
| QUARKS | $\approx 4.8 \text{ MeV}/c^2$ | $\approx 95 \text{ MeV}/c^2$ | $\approx 4.18 \text{ GeV}/c^2$ | 0 | |
| | $-1/3$ | $-1/3$ | $-1/3$ | 0 | |
| | $1/2$ | $1/2$ | $1/2$ | 1 | |
| | d down | s strange | b bottom | γ photon | |
| | $0.511 \text{ MeV}/c^2$ | $105.7 \text{ MeV}/c^2$ | $1.777 \text{ GeV}/c^2$ | $91.2 \text{ GeV}/c^2$ | |
| | -1 | -1 | -1 | 0 | |
| | $1/2$ | $1/2$ | $1/2$ | 1 | |
| | e electron | μ muon | τ tau | Z Z boson | |
| LEPTONS | $< 2.2 \text{ eV}/c^2$ | $< 0.17 \text{ MeV}/c^2$ | $< 15.5 \text{ MeV}/c^2$ | $80.4 \text{ GeV}/c^2$ | |
| | 0 | 0 | 0 | ± 1 | |
| | $1/2$ | $1/2$ | $1/2$ | 1 | |
| | ν_e electron neutrino | ν_μ muon neutrino | ν_τ tau neutrino | W W boson | |
| | | | | GAUGE BOSONS | |

Figure 2.1: Particle content of the Standard Model.

The three charged leptons—electron, muon, and tau—interact via the electromagnetic and weak forces, while their corresponding neutrinos (electron neutrino, muon neutrino, and tau neutrino) interact only through the weak force. Neutrinos are particularly intriguing due to their extremely small masses and their ability to oscillate between different flavors.

2.1.2 Forces and symmetries

Fundamental interactions in quantum field theory arise from the exchange of gauge bosons, which mediate three of the fundamental forces at work in the universe: the strong force, the weak force, and the electromagnetic force. They work over different

ranges and have different strengths: the electromagnetic force has infinite range and can predominantly drive interactions at intermediate scales; the weak and strong forces are effective only over a very short range and dominate at the level of subatomic particles. The weak force is so named because it is the weakest of the three forces at the scale of partons like quarks. The strong force, as the name suggests, is the strongest of all fundamental interactions between elementary particles at similar scale.

The strong nuclear force is mediated by gluons, which carry the force that binds quarks together to form protons, neutrons, and other hadrons. Unlike other force carriers, gluons themselves possess color charge, allowing them to interact with one another. At nuclear scales (approximately 1–3 femtometers), a residual effect of this fundamental force termed the “nuclear force,” carried by mesons such as pions, binds protons and neutrons together to form atomic nuclei.

The electromagnetic force, responsible for interactions between electrically charged particles, is mediated by photons. This force contributes to macroscopic interactions at various scales, from atomic bonding to everyday life (though significantly larger scales are dominated by gravity), and has an infinite range, diminishing in strength with the inverse square of the distance. The electromagnetic force is vastly stronger than gravity but is often neutralized at larger scales due to the balance of positive and negative charges in matter.

The weak interaction, mediated by the W and Z bosons, governs processes such as beta decay, or neutrino interactions. The W bosons facilitate charged-current interactions, enabling quarks to change flavor (e.g., converting a neutron into a proton in beta

decay) or transforming a lepton into its corresponding neutrino. The Z boson mediates neutral-current interactions, which do not change particle identities but influence their scattering behavior. Unlike the strong and electromagnetic forces, the weak force has a highly limited range (10^{-18} meters), due to the large masses of the W and Z bosons, which restrict their influence.

A fundamental question in early quantum field theory was why the strong and weak interactions are not long-range, and why there exists only a residual nuclear force outside the proton. Goldstone's theorem, a critical result in quantum field theory, originally suggested that any spontaneously broken continuous symmetry would necessarily yield massless scalar bosons [1]. This posed a major theoretical challenge, as no such massless particles were observed in experiments involving the weak force. It was thus believed that Yang-Mills gauge theories, which rely on local symmetries, could not correctly describe the weak interaction, since they would seemingly predict unobserved massless bosons.

However, this reasoning was incomplete. In the case of the strong interaction, the resolution lies in the non-abelian nature of quantum chromodynamics (QCD). Unlike electromagnetism, where photons interact with charged particles but remain neutral themselves, gluons possess color charge and therefore interact not only with quarks but also with each other. This self-interaction of gluons inside nuclei gives rise to “confinement:” quarks and gluons are permanently bound within hadrons such as protons and neutrons, and the strong force does not manifest significantly beyond nucleonic scales. While the strong force is immensely powerful at short distances, its effects are effectively screened at larger distances and from outside the proton, there's

almost no evidence of the strong force.

However, a completely different mechanism is responsible for the short range of the weak force. It turns out that the Higgs field, which permeates all of space, undergoes spontaneous symmetry breaking, providing a mass mechanism for the weak gauge bosons. In the SM, the electroweak symmetry group $SU(2)_L \times U(1)_Y$ is spontaneously broken to the electromagnetic subgroup $U(1)_{\text{EM}}$ by the Higgs vacuum expectation value, or VEV. This breaking mechanism gives mass to the W and Z bosons while leaving the photon massless, thereby allowing electromagnetism to remain long-range while restricting the weak force to subatomic scales. Despite the weak interaction's intrinsic coupling strength, the large masses of its mediators ensure that weak processes occur only over extremely short distances, making the weak force appear weak at macroscopic scales.

2.2 The Higgs mechanism

The origins of the Higgs mechanism can actually be traced to condensed matter physics, where Philip Anderson, in his 1963 paper on gauge invariance in superconductors, suggested that a field permeating space could confer mass to otherwise massless gauge bosons. This idea laid the groundwork for the development of a relativistic field-theoretic framework that could explain the mass of weak interaction bosons.

In 1964, several independent groups formalized this concept within the context of particle physics. François Englert and Robert Brout first published a paper describing spontaneous symmetry breaking in gauge theories, demonstrating how vector bosons

could acquire mass without violating gauge invariance. Shortly thereafter, Peter Higgs elaborated on this mechanism and explicitly predicted the existence of a new scalar boson, now known as the Higgs boson. Around the same time, Gerald Guralnik, Carl Hagen, and Tom Kibble presented an alternative but equivalent formulation of the same underlying physics. This collective work established what is now recognized as the Higgs mechanism.

In 1967, Steven Weinberg incorporated the Higgs mechanism into electroweak theory, building upon Sheldon Glashow's earlier work. Glashow had introduced a unified framework for the weak and electromagnetic interactions, postulating the existence of the W and Z bosons along with the photon. However, his model required a mechanism to break symmetry while preserving gauge invariance. Weinberg demonstrated that the Higgs field could accomplish this, allowing for the spontaneous breaking of electroweak symmetry and thereby endowing the W and Z bosons with mass while keeping the photon massless. Around the same time, Abdus Salam independently developed a similar theory. The electroweak theory, formulated by Glashow, Weinberg, and Salam, was instrumental in shaping the SM and earned the three physicists the 1979 Nobel Prize in Physics.

A major hurdle for early versions of the electroweak theory was the question of renormalizability—whether the theory could yield finite, predictive results at all energy scales. In 1971, Gerard 't Hooft, under the supervision of Martinus Veltman, demonstrated that gauge theories with spontaneously broken symmetry, including the electroweak theory, were renormalizable. This work provided the final theoretical validation of the SM framework and solidified the Higgs mechanism as a cornerstone

of modern particle physics.

Although widely referred to as the Higgs boson, some physicists have suggested that its name should reflect the collective contributions of multiple researchers. The term “ABEGHHK’tH boson” has been proposed, incorporating the initials of Anderson, Brout, Englert, Guralnik, Hagen, Higgs, Kibble, and ’t Hooft. However, historical convention has cemented the use of “Higgs boson” in both scientific and popular discourse.

In today’s SM, the Higgs field is introduced as a complex scalar doublet under the electroweak gauge group $SU(2)_L \times U(1)_Y$:

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}, \quad (2.1)$$

where ϕ^+ and ϕ^0 are complex scalar fields. This field transforms under $SU(2)_L \times U(1)_Y$ as $\Phi \rightarrow e^{i\alpha^a \tau^a} e^{i\beta Y} \Phi$, where τ^a are the generators of $SU(2)$, and Y is the hypercharge.

The dynamics of the Higgs field are governed by the Higgs potential:

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2. \quad (2.2)$$

The form of this potential depends on the sign of the real-valued parameter μ^2 . If

$\mu^2 > 0$, the minimum of $V(\Phi)$ occurs at $\Phi = 0$, preserving the full electroweak symmetry. If $\mu^2 < 0$, the potential acquires a “Mexican hat” shape, leading to spontaneous symmetry breaking (SSB) as illustrated in Figure 2.2.

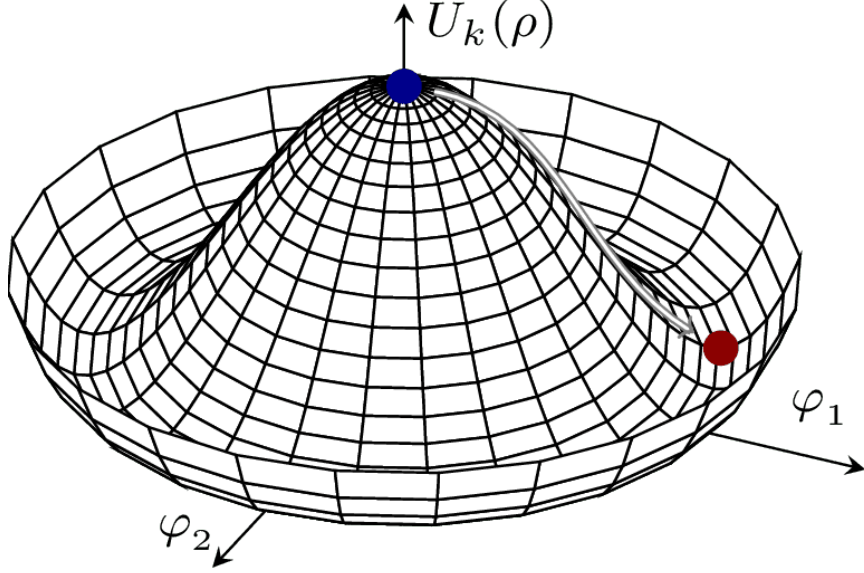


Figure 2.2: Visualization of the characteristic “Mexican hat” shaped Higgs potential, as parameterized by the two fields in the Higgs doublet.

For $\mu^2 < 0$, the Higgs field acquires a vacuum expectation value (VEV), breaking the electroweak symmetry:

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \quad (2.3)$$

Here, v is determined by minimizing the potential¹:

¹Experimentally, the VEV is found to be approximately $v \approx 246$ GeV.

$$v = \sqrt{\frac{-\mu^2}{\lambda}}. \quad (2.4)$$

The Higgs field interacts with the electroweak gauge bosons via the covariant derivative:

$$D_\mu \Phi = \left(\partial_\mu - igW_\mu^a \tau^a - ig'B_\mu Y \right) \Phi. \quad (2.5)$$

Expanding around the VEV and extracting the quadratic terms in the Lagrangian, the mass terms for the gauge bosons emerge. The kinetic term for the Higgs field,

$$(D_\mu \Phi)^\dagger (D^\mu \Phi), \quad (2.6)$$

yields mass terms for the weak bosons:

$$M_W = \frac{1}{2}gv, \quad M_Z = \frac{1}{2}\sqrt{g^2 + g'^2}v. \quad (2.7)$$

The photon remains massless, as expected, since the Higgs mechanism only breaks $SU(2)_L \times U(1)_Y$ down to $U(1)_{\text{EM}}$.

Fermion masses arise through interactions of the Higgs field with the fermion fields. The strengths of these interactions are dictated by so-called Yukawa couplings, which take the form:

$$\mathcal{L}_Y = -y_f \bar{\psi}_L \Phi \psi_R + \text{h.c.} \quad (2.8)$$

After the Higgs field acquires a VEV, this interaction produces mass terms proportional to each fermion's Yukawa coupling strength y_f :

$$M_f = \frac{y_f v}{\sqrt{2}}. \quad (2.9)$$

Thus, the Higgs mechanism provides a consistent and gauge-invariant way for the weak bosons and fermions to acquire mass while preserving renormalizability. The Higgs scalar field, through its spontaneous symmetry breaking, generates mass terms for the W and Z bosons, while fermions gain mass through Yukawa interactions.

For decades, the Higgs boson remained the final missing piece of the SM. Its discovery required a collider capable of reaching energy scales high enough to produce it. This challenge was met by the Large Hadron Collider (LHC) at CERN, where the ATLAS and CMS experiments, in 2012, independently observed a particle consistent with the predicted properties of the Higgs boson. This landmark discovery confirmed the Higgs mechanism and earned François Englert and Peter Higgs the 2013 Nobel Prize in Physics.

While the Higgs boson's discovery marked a triumph for the SM, it also highlighted unresolved mysteries in fundamental physics. Among the persistent shortcomings of the SM, we note that there is no inclusion of a force carrier for gravity, or so-called

“graviton.” Furthermore, the SM still does not contain any description of dark matter, which seems to outweigh visible matter roughly six to one. Or any explanation for the matter-antimatter asymmetry that we observe in our universe, despite the fact that, in principle, particles formed in the early universe should have had an equal chance of becoming matter or antimatter.

Answers to these questions, or at least hints to guide our way, hinge upon the search for evidence of physics Beyond the Standard Model (BSM). Given the extraordinary success of the SM in describing the known fundamental particles and their interactions, any necessary modifications or extensions to this framework are most effectively probed through the observation of subtle anomalies, discrepancies between theoretical expectations and precision measurements. The Higgs boson offers a natural point of connection to potential new physics, particularly in the context of dark matter². Since dark matter must possess mass to interact gravitationally, it is plausible that dark matter particles couple to the Higgs boson.

If dark matter (or any other anomalous particle) does indeed couple to the Higgs field, then precise measurements of Higgs boson production and decay processes could reveal subtle anomalies—shifts in branching ratios, invisible decays, or unexpected interaction strengths—that point toward the existence of new, BSM particles. Thus, studying the Higgs boson in unprecedented detail represents one of our most promising strategies for uncovering the true nature of dark matter and extending our understanding of fundamental physics beyond the SM.

²Dark matter remains elusive, with its non-gravitational interactions yet to be directly detected.

2.3 Phenomenology of the off-shell Higgs boson

Since 2012, both ATLAS and CMS have observed a Higgs boson with mass around 125 GeV [2] [3] which is consistent with our SM expectations. The Higgs boson’s mass, as the last free parameter of the SM, is necessarily determined experimentally and is instrumental in our understanding of the SM, along with the Higgs boson’s other properties such as its width and couplings to other particles. This is because the Higgs mass relies on the value of λ as seen in 2.2, an a priori unknown free parameter. It is also tied to the value of the Higgs boson width and μ^2 . Ergo, any precision measurements of the Higgs boson are invaluable in our search for deviations from our SM expectation.

Now, in the SM, “virtual” particles can have a mass that is different from the mass that the particle would have if it were “real,” or physical. Inside a Feynman diagram³, virtual particles can have any mass because they aren’t physical, stable particles; they are excitations in quantum fields. This lets us draw Feynman diagrams with intermediary particles of masses very different from their SM mass. An example of this is illustrated in Figure 2.6 in which a Higgs boson ($m_{H,SM} \approx 125$ GeV) can decay to two physical Z bosons ($2m_{Z,SM} \approx 182$ GeV). The SM mass, also referred to as the “pole mass” is the mass value which matches the mathematical pole of each particle’s renormalized propagator in momentum space. It can also be thought of as the physical mass of the particle as observed, for example if measured in scattering experiments.

³Feynman diagrams are pictorial representations of interactions between subatomic particles. They provide an intuitive way to illustrate and organize the many terms which arise in expansions of quantum field theory (QFT) descriptions of such interactions.

We call particles with masses away from their pole mass “off-shell.” This terminology arises from the notion that a particle is produced “on the mass shell” when its invariant mass⁴ matches the SM expectation, whereas an “off the mass shell” particle is produced with invariant mass higher or lower than its nominal pole mass. This thesis will focus on measurements of the off-shell Higgs boson, for reasons to be presented.

2.3.1 Production modes

At collision energies around 13 TeV, we certainly have enough energy to produce the Higgs boson at the LHC, as observed in the plot of proton-proton cross sections in Figure 2.3. To provide a sense of scale, assuming SM cross sections, ~ 1 Higgs bosons are produced every second. Further details are presented in Section 3.1.

Production occurs via four dominant modes which have the Feynman diagrams illustrated in Figure 2.4. As one can see, each of the dominant production diagrams begin with quarks or gluons from the proton-proton collisions. Our dominant production mode, gluon fusion, can be seen in Figure 2.4a and at leading order involves a quark loop since the Higgs boson does not couple directly to the massless gluons. The primary contribution is with a loop of top quarks (the heaviest quark), with a small contribution from the bottom quark (the second heaviest). At higher order, radiated gluons and other jets (created through QCD effects from the initial state gluons) can be produced along with the Higgs boson.

The second most productive mechanism is vector boson fusion (VBF), shown in

⁴Note that the invariant mass is a Lorentz-invariant quantity defined from a particle’s four-momentum, and includes both energy and momentum such that it remains constant regardless of boost or frame. For on-shell particles, this coincides with the rest mass.

Figure 2.4b, in which two vector bosons are radiated from our initial state quarks and then interact to “fuse” and generate a Higgs boson. Because the W and Z vector bosons have mass around 80-90 GeV, VBF can involve the radiation of hundreds of GeV especially at higher energies when the vector bosons come on-shell. This means that the two quark jets emitted from a VBF event characteristically possess large transverse energies and are often emitted opposite each other.

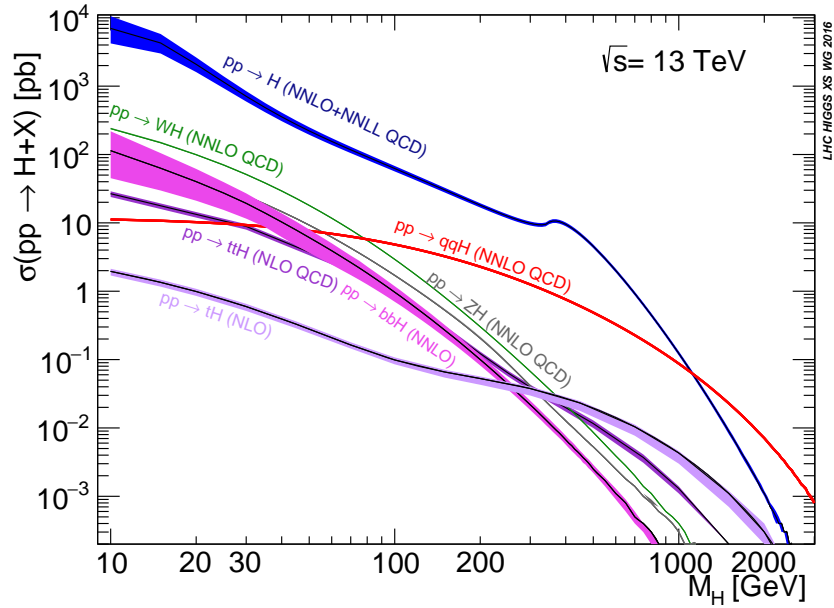


Figure 2.3: SM Higgs boson production cross sections as a function of m_H at the LHC at $E_{CM} = 13$ TeV [4].

The third most common production mode for both on-shell and off-shell Higgs boson production is associated production with a vector boson, or VH production. This is also known as Higgs-strahlung production, since it closely resembles how a photon can be radiated by an electron through Bremsstrahlung radiation. In VH production, as seen in Figure 2.4c, two initial state quarks will produce an energetic vector boson which radiates a Higgs boson. Such that a $q\bar{q}$ initial state results in a VH final state.

Finally, most of the remaining Higgs boson cross section comes from $t\bar{t}H$ production, Figure 2.4d. This production mode, along with $b\bar{b}H$ and other less common mechanisms, becomes less relevant in the production of the off-shell Higgs boson as one can see in Figure 2.3. In the off-shell region (chosen to include energies above 220 GeV), the relevant production modes to consider are ggH , VBF, and VH.

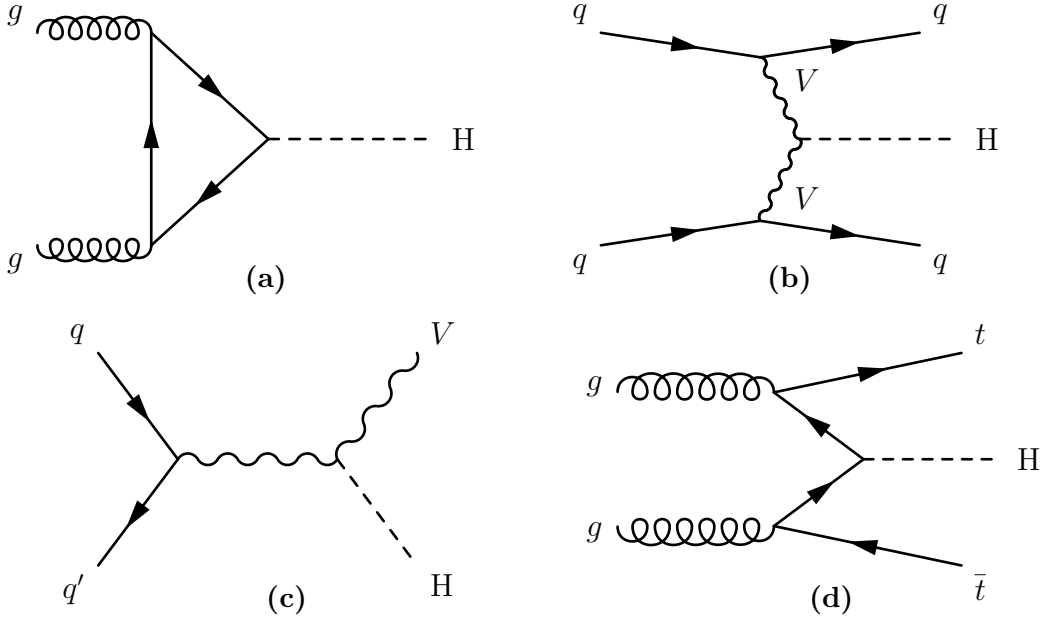


Figure 2.4: Feynman diagrams for the dominant Higgs boson production modes at the LHC: (a) gluon fusion (ggH), (b) vector boson fusion (VBF), (c) associated (VH) production, and (d) $t\bar{t}H$ production.

2.3.2 Decay modes

The interaction strength of the Higgs boson with its decay products is directly proportional to their masses. As illustrated in Figure 2.5a, at a Higgs boson mass around 125 GeV, the dominant decay mode is $H \rightarrow b\bar{b}$, since the bottom quark is the heaviest kinematically accessible fermion. However, as the Higgs boson's mass

increases beyond this threshold, decays into electroweak gauge bosons, specifically $H \rightarrow WW$ and $H \rightarrow ZZ$, become dominant due to their larger coupling strength and enhanced phase space availability as there is enough energy for the vector bosons to exist on-shell. The loop-induced processes $H \rightarrow \gamma\gamma$ and $H \rightarrow Z\gamma$ have significantly lower branching fractions, as the Higgs boson does not couple directly to massless photons; these decays proceed via higher-order quantum corrections involving virtual heavy particles in the loop, such as top quarks or W bosons.

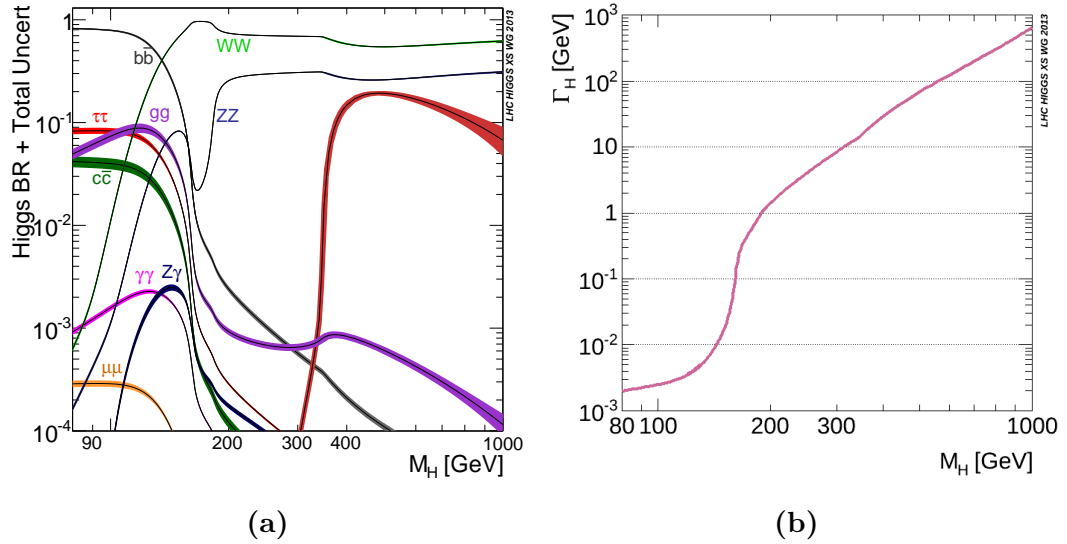


Figure 2.5: SM Higgs boson decay branching ratios (a) and total decay width (b) as a function of mass m_H [4]. In the off-shell region above 220 GeV, the Higgs boson decay to a pair of vector bosons dominates. We see an enhancement of the Higgs boson width at energies above $2m_V$ when the two vector bosons become on-shell.

The experimental sensitivity of a given decay channel is determined not only by the branching fraction but also by factors such as the mass resolution of reconstructed final states and the level of irreducible and reducible backgrounds. The $H \rightarrow b\bar{b}$ channel, despite its large branching ratio, suffers from overwhelming quantum chromodynamic (QCD) backgrounds and poor mass resolution, making precise measurements chal-

lenging. The $H \rightarrow WW$ decay, while more distinctive, involves final-state neutrinos that escape detection, leading to missing transverse energy and complications in fully reconstructing the Higgs boson's invariant mass.

The $H \rightarrow ZZ$ channel provides a particularly clean final state, especially in the fully leptonic decay mode $H \rightarrow ZZ \rightarrow 4\ell$ (where $\ell = e, \mu$). This process benefits from a low background contamination, fully reconstructible decay products, and excellent mass resolution, making it an optimal channel for precision measurements of the Higgs boson properties. As a result, this thesis will primarily focus on the $H \rightarrow ZZ$ decay mode for detailed studies of the Higgs boson, as visualized in Figure 2.6.

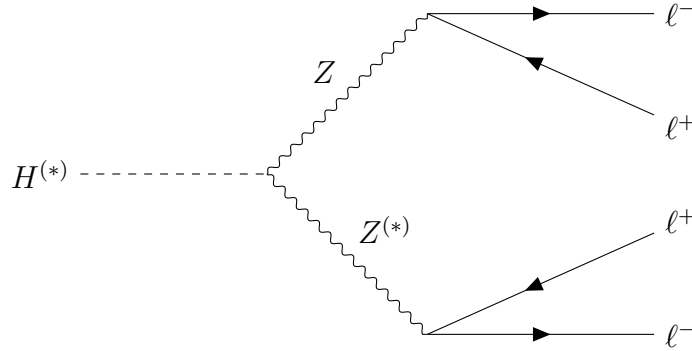


Figure 2.6: Feynman diagram for the decay $H^{(*)} \rightarrow ZZ^{(*)} \rightarrow 4\ell$.

2.4 Properties of the Higgs boson

One of the most effective ways to investigate potential deviations from SM predictions is through precise measurements of the properties of the Higgs boson. As an extremely short-lived particle, the Higgs boson rapidly decays into lighter particles, and its finite lifetime τ is inherently connected to the width Γ of its resonance via the uncertainty principle, $\Gamma = \hbar/\tau$. This width can be interpreted as the intrinsic uncertainty in the

particle’s mass. Since the decay rate of the Higgs boson is dictated by its couplings to other particles, any deviation from SM predictions in these couplings would manifest as an alteration in its total width. Consequently, high-precision studies of the Higgs sector, particularly measurements of Γ_H , provide an indirect but highly sensitive probe of BSM physics, including possible exotic decay channels or modifications to Higgs boson interactions.

For particles with broad resonances, the total width can be extracted directly from the Breit–Wigner line shape, which characterizes the probability distribution of invariant mass around the nominal mass of the particle. However, the Higgs boson presents a unique challenge in this regard. Due to its extremely short lifetime, the corresponding resonance width is remarkably narrow relative to its mass, making direct measurement from the line shape infeasible. Even in the so-called “golden channel,” where the Higgs boson decays via $H \rightarrow ZZ \rightarrow 4\ell$ (a channel renowned for its clean experimental signature and excellent mass resolution), the finite detector resolution limits the direct width measurement to an uncertainty of approximately 1 GeV. This is orders of magnitude larger than the predicted SM Higgs boson width of around 4 MeV, rendering a direct determination of Γ_H at the peak practically impossible.

Given these limitations, alternative methods must be employed to infer the Higgs boson’s width. One such approach involves off-shell Higgs boson production, where measurements in kinematic regions far from the Higgs boson pole mass can provide indirect sensitivity to its total width. This method separates Higgs boson-mediated and non-resonant background processes, and utilizes the observed yields of Higgs boson events to extract a measurement of Γ_H that is not constrained by detector resolution.

The ability to probe the Higgs boson width through such indirect techniques represents a significant contribution to current precision Higgs boson studies in the ongoing search for deviations from the SM and potential new physics.

2.5 Off-shell technique

In this thesis, we employ the “off-shell technique,” originally developed by Caola and Melnikov [5] and implemented by the CMS collaboration [6], which currently provides the most stringent constraints on the Higgs boson total width. This method exploits Higgs boson production in two distinct kinematic regimes: the “on-shell” region, defined in this case within the invariant mass window of 105 to 140 GeV, and the “off-shell” region, characterized by Higgs boson candidates exceeding 220 GeV.

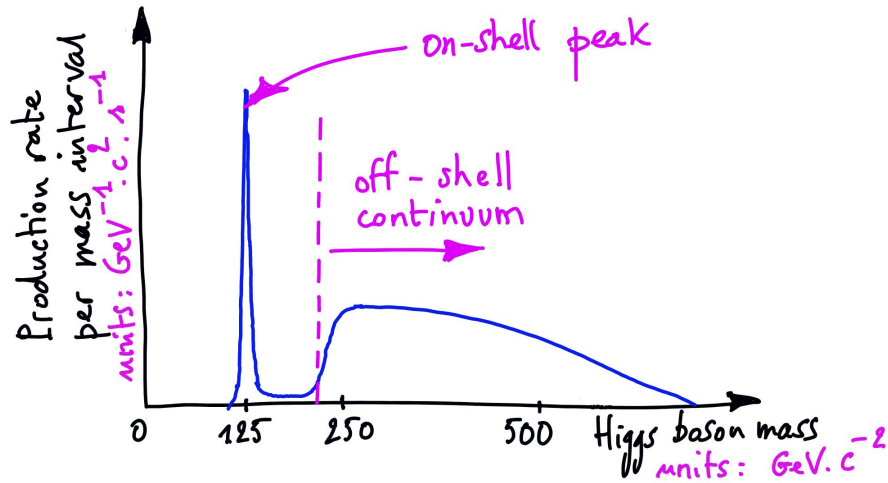


Figure 2.7: A published sketch of the Higgs boson mass distribution reconstructed from ZZ decays. The peak on the left marks on-shell production while the broad shoulder on the right represents the off-shell process. Beware that the corresponding rates indicated here are not to scale.

A seminal study conducted in 2013 by researchers at Johns Hopkins University [5] demonstrated that the total width of the Higgs boson, Γ_H , can be inferred by relating the production rates of on-shell and off-shell Higgs bosons. Specifically, by measuring the rate of Higgs boson production through gluon fusion and subsequent decay via the so-called “golden channel” ($H^* \rightarrow ZZ \rightarrow 4\ell$), a direct, resolution-constrained measurement of the Higgs boson width can be circumvented. Instead, a precise, statistically limited determination of Γ_H is achieved by taking the ratio of off-shell to on-shell Higgs boson event yields [7]. The proportionalities

$$\sigma_{VV \rightarrow H \rightarrow 4\ell}^{\text{on-shell}} \propto \mu_{VVH} \frac{1}{\Gamma_H} \quad \text{and} \quad \sigma_{VV \rightarrow H \rightarrow 4\ell}^{\text{off-shell}} \propto \mu_{VVH}, \quad (2.10)$$

where μ_{VVH} is the signal strength modifier (defined as the ratio of the observed to expected number of on-shell 4ℓ events), give us the relationship

$$\Gamma_H \propto \sigma_{VV \rightarrow H \rightarrow 4\ell}^{\text{off-shell}} / \sigma_{VV \rightarrow H \rightarrow 4\ell}^{\text{on-shell}}. \quad (2.11)$$

Since the off-shell to on-shell Higgs boson cross section ratio scales proportionally with Γ_H , this method provides a highly sensitive probe of the Higgs boson width.

This approach offers several advantages. In the decay of an on-shell Higgs boson to a pair of massive vector bosons (e.g., ZZ), at least one of the Z bosons must be produced off-shell due to the Higgs boson mass being below the kinematic threshold of $2m_Z \approx 180$ GeV. However, in the off-shell regime ($m_H > 220$ GeV), the Higgs boson can decay into two fully on-shell Z bosons, each with a mass of approximately 91 GeV.

This results in a significant enhancement⁵ of the off-shell Higgs boson production rate at high invariant masses, thereby increasing the available event statistics and improving the precision of the width determination.

Furthermore, the focus on Higgs boson decays via the $ZZ \rightarrow 4\ell$ channel ensures a clean experimental signature with excellent mass resolution, while also enhancing sensitivity to potential deviations in the Higgs’ couplings to vector bosons. As a result, this method not only facilitates a precise measurement of the Higgs boson width but also serves as a robust probe of anomalous contributions to the Higgs–vector boson interaction vertex (HVV). With the LHC operating at a center-of-mass energy of 13.6 TeV in Run 3, the anticipated increase in high-energy collisions will further enhance the statistical power of this technique, enabling even more stringent tests of potential BSM physics.

2.6 Higgs boson couplings

In the SM, the Higgs boson couples to gluons via quark loops with Yukawa interactions, where the dominant contribution arises from the top-quark loop, as illustrated in Figure 2.4a. But BSM physics could allow for contributions from new heavy particles in this loop. Yukawa couplings also govern direct interactions with fermion-antifermion pairs, such as in $t\bar{t}H$ and $b\bar{b}H$ production. These interactions are less relevant at high energies due to their suppression, shown in Figure 2.3, but they are included in the analysis of on-shell Higgs boson production under similar assumptions as for gluon

⁵This enhancement is “significant” relative to the dramatic suppression of production as the Higgs boson moves off-shell. Overall, the off-shell range only accounts for $\sim 14\%$ of the total number of Higgs boson events.

fusion.

The Higgs boson couples to the electroweak vector bosons in HVV vertices, without the need for quark mediators. Since we are targeting off-shell Higgs boson production, we have one such vertex on the production sides of two of our primary Feynman diagrams, VBF and VH. And because of our focus on the $H \rightarrow ZZ$ decay channel, we have additional HVV vertices in all of our Feynman diagrams on the decay side. This makes us particularly sensitive to the HVV amplitude overall, which we can use to our advantage.

Beyond the Higgs boson couplings to SM particles, we can also incorporate anomalous coupling strength terms which can be constrained to develop Effective Field Theory (EFT) descriptions. Some examples of familiar vertices in Feynman diagrams that can feature anomalous EFT contributions are shown in Figure 2.8. The anomalous contributions are treated as deviations in the Higgs boson's interactions with two gauge bosons. These anomalous couplings can manifest in both Higgs boson production and decay, regardless of the invariant mass at which the Higgs boson is produced.

2.6.1 Amplitude EFT couplings

We assume that the Higgs boson couples to two gauge bosons VV , such as WW , ZZ , $Z\gamma$, or $\gamma\gamma$, which in turn couple to fermions. These fermions appear either as four-lepton final states in the Higgs boson decay, or as quarks and leptons in its production and in the decays of associated electroweak bosons.

A general form of the signal scattering amplitude of the H boson with two vector

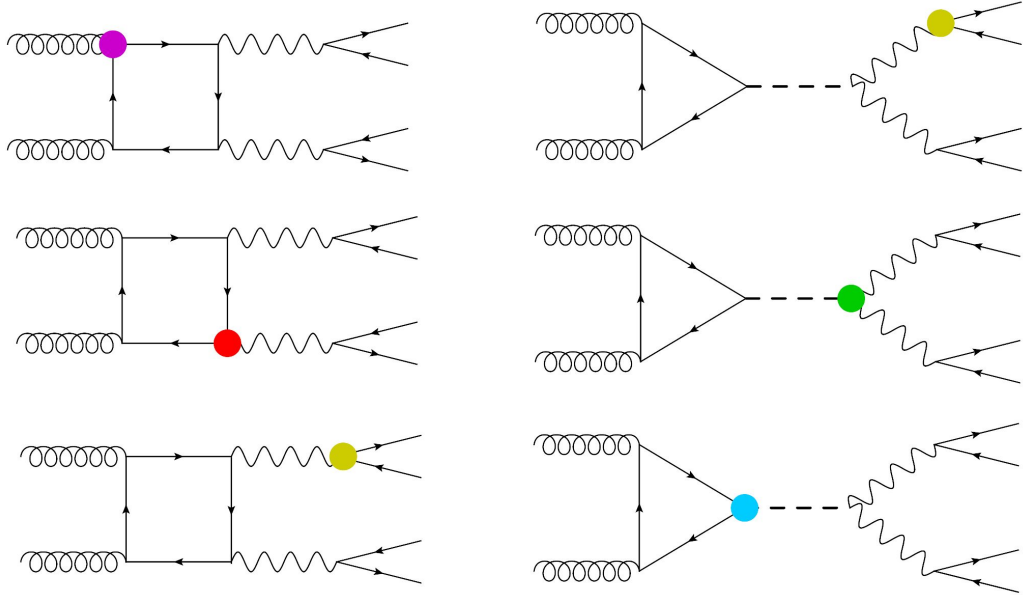


Figure 2.8: Example Feynman diagrams for gluon fusion off-shell Higgs boson production (right) and corresponding background (left) with EFT insertions [8].

bosons can be written in terms of polarization vectors ϵ_i and momenta q_i of bosons V_i , with scalar tensor $f^{(i)\mu\nu} = \epsilon_i^\mu q_i^\nu - \epsilon_i^\nu q_i^\mu$, as:

$$A(HVV) \propto \left[a_1^{VV} + \frac{\kappa_1^{VV} q_1^2 + \kappa_2^{VV} q_2^2}{(\Lambda_1^{VV})^2} + \frac{\kappa_3^{VV} (q_1 + q_2)^2}{(\Lambda_Q^{VV})^2} \right] m_{V1}^2 \epsilon_{V1}^* \epsilon_{V2}^* \quad (2.12)$$

$$+ a_2^{VV} f_{\mu\nu}^{*(1)} f^{*(2),\mu\nu} + a_3^{VV} f_{\mu\nu}^{*(1)} \tilde{f}^{*(2),\mu\nu},$$

such that the scales of BSM physics are set by Λ_1 and Λ_Q , and the a_2^{VV} and a_3^{VV} scaled terms provide CP-even and CP-odd contributions, which may be sensitive to evidence of CP violation. Figure 2.9 highlights the double HVV vertex present in our electroweak Feynman diagrams.

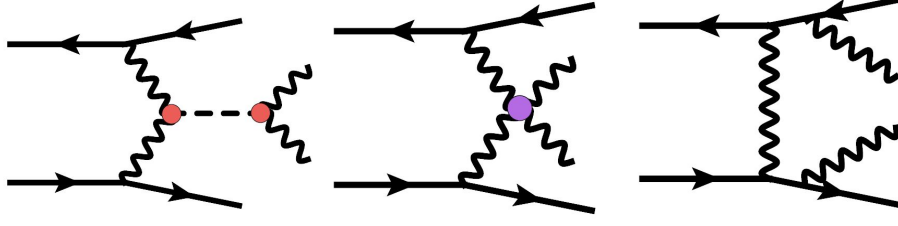


Figure 2.9: Example Feynman diagrams for electroweak off-shell Higgs boson production (left), vector boson scattering (VBS) (center), and additional background (right) with EFT insertions [8].

When at least one of the gauge bosons V is massive, m_{V1} is the pole mass of that gauge boson. All the a_i^{VV} (which denotes a_1^{VV} , $1/\Lambda_1$, $1/\Lambda_Q$, a_2^{VV} , and a_3^{VV}) terms act as coupling-strength modifiers on their corresponding contributions to the amplitude.

Instead of constraining any a_i directly, we can constrain the fractional contribution of its corresponding term to the signal scattering amplitude of the Higgs boson, as a ratio of observable cross sections f_{ai} . In Equation 2.12, pure SM behavior is described by a_1^{VV} alone. This would correspond to a measurement of $f_{a_1} = 1$. If, for example, we took $f_{a_1} = 0.5$, then we would have a 50% contribution to our overall cross section from the SM and 50% from BSM physics.

This parameterization via f_{ai} has the benefit of being invariant to the a_n coupling scale, and having values bounded by 0 and 1. Also, in experimental measurements of f_{ai} most systematic uncertainties cancel in the ratios to give us a cleaner measurement [9]. We define ratio f_{ai} and phase ϕ_{ai} as:

$$f_{ai} = \frac{|a_i|^2 \sigma_i}{\sum_{j=1,2,3\dots} |a_j|^2 \sigma_j}, \phi_{ai} = \arg\left(\frac{a_i}{a_1}\right). \quad (2.13)$$

2.6.2 Kappa framework couplings

Another representation of the Higgs boson couplings utilizes the so-called κ framework, in which couplings of the Higgs boson to the SM particles are modified by a scale factor κ_i for each particle i .

As previously discussed, SM-like behavior implies the top quark's predominance in the gluon fusion loop during Higgs boson production. If there are additional contributions in this loop, such as from undiscovered heavy particles, the probability distribution for Higgs boson production in the off-shell region would change. To model this, one could introduce a heavy quark Q with coupling strength to the Higgs boson κ_Q [9, 10].

In the framework of effective field theories, the contribution from Q can be reasonably interpreted as a point-like interaction which integrates all effects from heavy particles present in the loop. One can include both top and bottom quarks in the loop, characterized by the coupling strength κ_q , where $\kappa_q = 1$ corresponds to the SM. Similarly, κ_V is the Higgs boson coupling strength to the Z and W bosons, which impact EW production and Higgs boson decay and where $\kappa_V = 1$ in the SM.

As described in Equation 2.10, we usually parameterize our model in terms of signal strength. In this analysis, we use μ_F for Higgs boson production mechanisms driven by fermion couplings, i.e. production via ggH or in association with a $t\bar{t}$ ($t\bar{t}H$) or $b\bar{b}$ pair ($b\bar{b}H$). For EW production, i.e., production via vector boson fusion (VBF) or in association with a W or Z boson (VH), the signal strength is denoted as μ_V .

The κ framework requires a re-parameterization in terms of κ_q , κ_Q , and κ_V [10]:

$$\begin{aligned}\mu_V &= \kappa_V^4 \times \frac{\Gamma_0}{\Gamma} \\ \mu_F &= \kappa_q^2 \kappa_V^2 \times \frac{\Gamma_0}{\Gamma}\end{aligned}\tag{2.14}$$

where Γ_0 is the reference value of the Higgs boson width used in simulation, and

$$\begin{aligned}\kappa_Z &= \kappa_W = \kappa_V = \text{sign}(\mu_V) \left(|\mu_V| \frac{\Gamma}{\Gamma_0} \right)^{\frac{1}{4}} \\ \kappa_t &= \kappa_b = \kappa_q = \text{sign}(\mu_F) |\mu_F|^{\frac{1}{2}} \left(\frac{1}{|\mu_V|} \frac{\Gamma}{\Gamma_0} \right)^{\frac{1}{4}}\end{aligned}\tag{2.15}$$

The bulk of this thesis focuses on measurements of the SM Higgs boson, but it will also include an extension of the analysis which sets constraints on these κ framework coupling strengths. It should become apparent that this analysis also sets the stage for upcoming measurements of anomalous coupling strengths in our amplitude basis, presented in Section 2.6.1.

Chapter 3

The Experiment

3.1 The Large Hadron Collider (LHC)

Since its inaugural operation on 10 September 2008, the LHC has remained the most powerful and expansive particle accelerator ever constructed. It is so named because it is very large, consisting of a 27 km ring of superconducting magnets; accelerates hadrons, usually protons but sometimes various ions such as the nuclei of lead, argon or xenon atoms; and collides these particles via two beams traveling in opposite directions, which intersect at four points along the circumference of the machine. By pushing energy frontiers, the LHC enables scientists to test fundamental theories and explore high-energy interactions where rare and undiscovered phenomena might occur.

The energy of a circular particle accelerator is constrained by its size and the strength of its guiding magnetic fields. For the LHC, the center-of-mass energy is proportional to the curvature radius of the magnets, which drives the size of the ring, and the field strength of the superconducting dipole magnets. The LHC reuses the 27 km underground tunnel of its predecessor, the Large Electron-Positron Collider (LEP), which was decommissioned in 2000. The decision to reuse this infrastructure was driven by cost-effectiveness, as tunneling underground was financially more viable than acquiring land and developing an equivalent surface facility. Additionally, the Earth's crust provides natural radiation shielding, reducing background noise in detectors and protecting personnel from radiation. The dimensions of the tunnel,

along with constraints from superconducting magnets, RF cavities, and cryogenic systems, establish the design energy of 7 TeV per proton beam [11].



Figure 3.1: An aerial view of CERN and the LHC, located near Geneva on the border between France and Switzerland.

Geological considerations dictated the placement of the tunnel at an average depth of 100 m, with variations ranging from 175 m beneath the Jura Mountains to 50 m near Lake Geneva. The tunnel follows a gentle gradient of 1.4% to accommodate geological stability and minimize excavation costs. The location was also chosen to ensure minimal disruption to surface infrastructure and avoid heavily populated areas while maintaining proximity to CERN’s primary research facilities.

The LHC relies on superconducting electromagnets to maintain and guide the circulating beams. Specifically, 1232 dipole magnets, each measuring 15 meters in length, generate a magnetic field of 8.3 T to bend the trajectory of the protons. These dipoles,

| Design Parameter | Value |
|---------------------------------|--------------------------|
| Circumference | 26659 m |
| Dipole operating temperature | 1.9 K (-271.3 °C) |
| Number of magnets | 9593 |
| Number of main dipoles | 1232 |
| Number of main quadrupoles | 392 |
| Number of RF cavities | 8 per direction |
| Energy, protons | 7 TeV |
| Energy, ions | 2.76 TeV/u |
| Peak magnetic dipole field | 8.3 T |
| Distance between bunches | ~ 7.5 m |
| Luminosity (protons) | $10^{34} cm^{-2} s^{-1}$ |
| No. of bunches per proton beam | 2808 |
| No. of protons per bunch | 1.1×10^{11} |
| Number of turns per second | 11245 |
| Number of collisions per second | 1 billion |

Table 3.1: Design parameters for the LHC [[11](#)]

along with 392 quadrupole magnets (each between 5 and 7 meters long) responsible for beam focusing, are composed of niobium-titanium (NbTi) superconducting cables. These materials exhibit superconductivity below 10 K (-263.2°C), though the LHC operates at an even lower temperature of 1.9 K (-271.3°C). The ultrahigh vacuum maintained within the beam pipes (10^{-10} to 10^{-11} mbar) is colder than the cosmic microwave background temperature (2.7 K) and approaches the surface pressure of the Moon.

This extreme cooling is necessary to ensure minimal resistance in the superconducting magnets and to reduce unwanted interactions that could degrade beam quality. To sustain these cryogenic conditions, the LHC employs 96 tonnes of superfluid helium, making it the world's largest cryogenic system. Beam control and stability are further enhanced by specialized multipole magnets and final focusing quadrupoles near interaction points. Prior to collision, the beams undergo final focusing, compressing their transverse size to maximize interaction probability. The precision required for this process is akin to aligning two needles 10 km apart such that they collide at their midpoint.

All accelerator operations, including beam injection, acceleration, steering, and diagnostics, are centrally coordinated from the CERN Control Center. The LHC facilitates collisions at four primary interaction regions housing major particle detectors: ATLAS (A Toroidal LHC ApparatuS), CMS (Compact Muon Solenoid), ALICE (A Large Ion Collider Experiment), and LHCb (Large Hadron Collider beauty). Of these, ATLAS and CMS serve as general-purpose detectors, while LHCb specializes in flavor physics, particularly the study of b-quarks, and ALICE investigates heavy-ion collisions. Each

of these experiments has dedicated research objectives, with ATLAS and CMS playing a central role in Higgs boson studies and searches for BSM physics, while ALICE focuses on recreating conditions similar to those moments after the Big Bang, such as with quark-gluon plasmas. This thesis presents results obtained using pp collision data recorded by the CMS experiment.

Proton beams at the LHC are structured into well-defined bunches due to the constraints imposed by RF acceleration. Under nominal conditions, each beam comprises 2808 bunches, with each bunch containing approximately 10^{11} protons. Given the minute transverse dimensions of the bunches, the probability of individual proton-proton collisions becomes very large through the high luminosity achieved through intense beam focusing and bunch structuring. At each interaction point, up to 60 simultaneous collisions can occur per bunch crossing. With bunch crossings occurring at a rate of ~ 40 MHz, the LHC generates approximately 2.5×10^9 proton-proton interactions per second. This vast amount of data necessitates sophisticated triggering systems to selectively record the most interesting events while discarding redundant or low-energy interactions. Advances in machine learning techniques have also been integrated into the LHC's data-processing pipeline to optimize event selection and improve detection of rare physics processes.

A key parameter quantifying collider performance is luminosity, defined as the number of interactions per unit area per unit time. While luminosity at the LHC is often reported integrated over a period of active beam, the instantaneous luminosity can be calculated in terms of the accelerator parameters as follows:

$$\mathcal{L} = \frac{\gamma f N_p^2 N_b}{4\pi\epsilon_n\beta^*} \mathcal{F} \quad (3.1)$$

where N_p is the number of particles per bunch, N_b is the number of bunches per beam, f is the revolution frequency of the bunches around the ring, ϵ_n is the transverse emittance, β^* is the amplitude function at the interaction or crossing point, \mathcal{F} is a geometric factor correlated to the crossing angle of the beams, and γ is the relativistic Lorentz factor $\frac{1}{\sqrt{1-\beta^2}}$. A plot of the integrated luminosities delivered and recorded by CMS is shown in Figure 3.2. This thesis includes data through 2018 (Run 1 + Run 2).

The number of events for a particular interaction also depends on its probability, expressed as a cross section in units of area. When investigating a specific physics process, the event rate is determined by the product of the luminosity (\mathcal{L}) and the process-specific cross-section (σ):

$$\frac{dN}{dt} = \mathcal{L} \times \sigma. \quad (3.2)$$

3.2 The Compact Muon Solenoid (CMS)

The CMS detector is built around a large superconducting solenoid, for which it is named. It spans 6 meters in its internal diameter and generates a magnetic field of 3.8 Tesla. This very large field bends the trajectories of charged particles which helps to identify their momenta along with charge information. Inside the solenoid we find

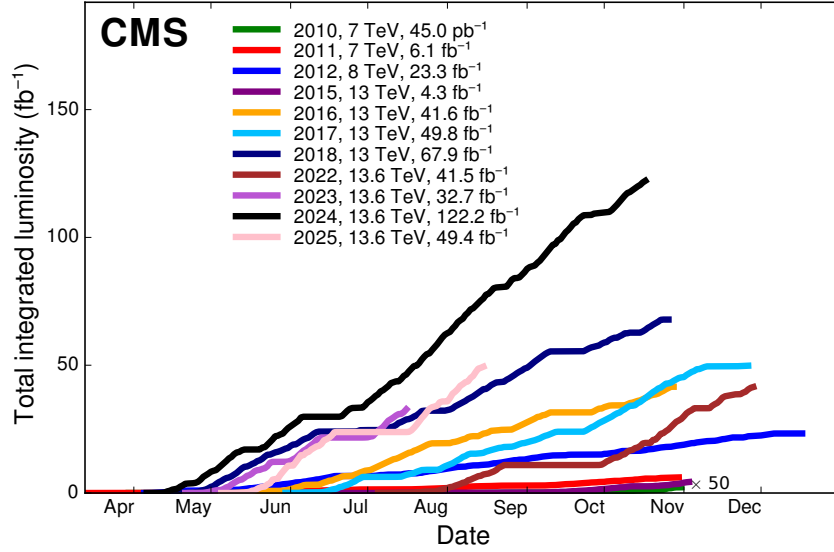


Figure 3.2: Cumulative luminosity versus day delivered to CMS during stable beams for pp collisions at nominal center-of-mass energy. This is shown for data-taking in 2010 (green), 2011 (red), 2012 (blue), 2015 (purple), 2016 (orange), 2017 (light blue), 2018 (navy blue), 2022 (brown), and 2023 (light purple). These plots use the best available offline calibrations for each year.

silicon trackers and calorimeters, while outside the solenoid we have layers of muon detection chambers [12–14].

A useful analogy is that CMS effectively acts like a high-speed camera, taking 3-dimensional “snapshots” of particle collision events in all directions up to 40 million times each second. Although most of the initial collision products decay rapidly, the stable products can be detected by CMS as they travel away from the beamline. By identifying the stable particles produced in each collision and then piecing them together, the detector can reconstruct an understanding of the collision which can

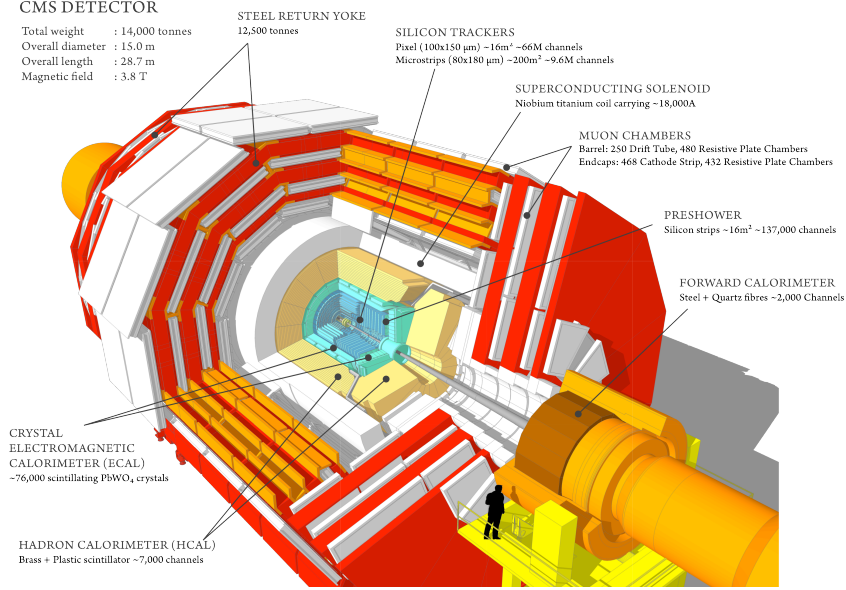


Figure 3.3: The CMS detector is shaped like a cylindrical onion, with several concentric layers of components.

then be analyzed to identify constituent physics phenomena [13, 15–18].

3.2.1 Solenoid

The eponymous superconducting solenoid is central to the construction and operation of CMS, and features an internal diameter of 6 m. As the largest superconducting magnet ever built, its coils and return yoke weigh in at 12,500 tonnes and it is cooled to 1.9 K (-268.5°C), approaching the temperature of deep space. It also generates a uniform magnetic field of 3.8 T, which is approximately 10^5 times stronger than the Earth’s magnetic field.

This intense field bends the trajectories of charged particles produced in high-energy



Figure 3.4: View of the CMS magnet during construction.

collisions. The degree of deflection is inversely proportional to a particle's momentum, allowing precise momentum measurements through high-resolution tracking in conjunction with the well-calibrated magnetic field. High-energy particles exhibit minimal curvature, necessitating exceptional spatial resolution to accurately reconstruct their momenta. This motivates the precise alignment of the tracker in CMS.

The inner tracking system and electromagnetic and hadronic calorimeters (ECAL and HCAL) are housed within the solenoid coil. The solenoid itself provides primary structural support for the CMS experiment, and is robustly engineered to withstand the immense forces generated by its own magnetic field. Surrounding the solenoid is a 12-sided iron return yoke, which integrates the muon detection system and serves to contain and guide the magnetic field. Comprising of three layers, the yoke extends to a diameter of 14 m and functions as a physical filter, allowing only muons and weakly interacting particles, such as neutrinos, to pass through.

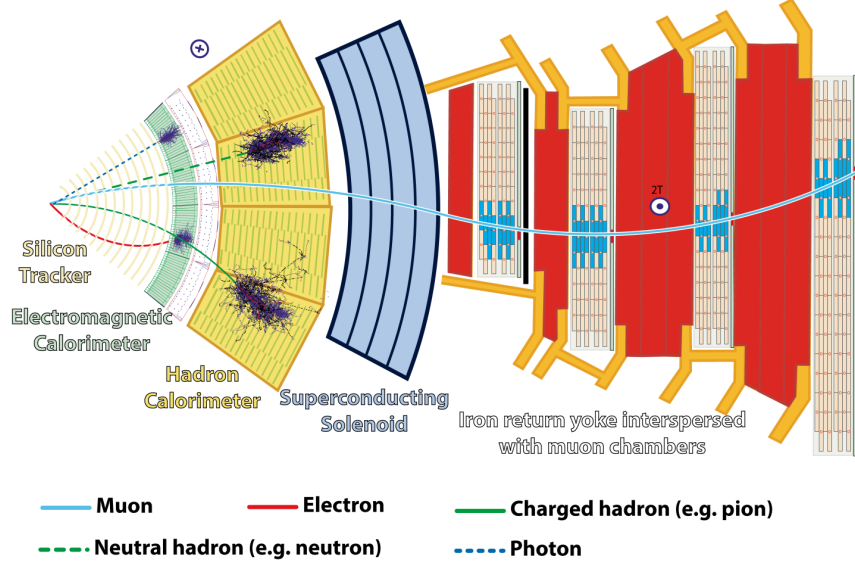


Figure 3.5: Schematic view of the CMS detector.

3.2.2 Silicon tracker

The CMS tracker is designed to precisely record charged particle trajectories while minimizing material interactions to reduce additional scattering and energy loss. This is achieved through high-precision spatial measurements, enabling accurate track reconstruction with a minimal number of data points. Each position measurement is recorded with a resolution of $\sim 10 \mu\text{m}$, an order of magnitude smaller than the width of a human hair, ensuring precise momentum determination.

As the innermost subsystem of CMS, the tracker operates in a highly irradiated environment due to its proximity to the interaction point. Consequently, its construction materials were meticulously selected for radiation hardness to maintain performance.

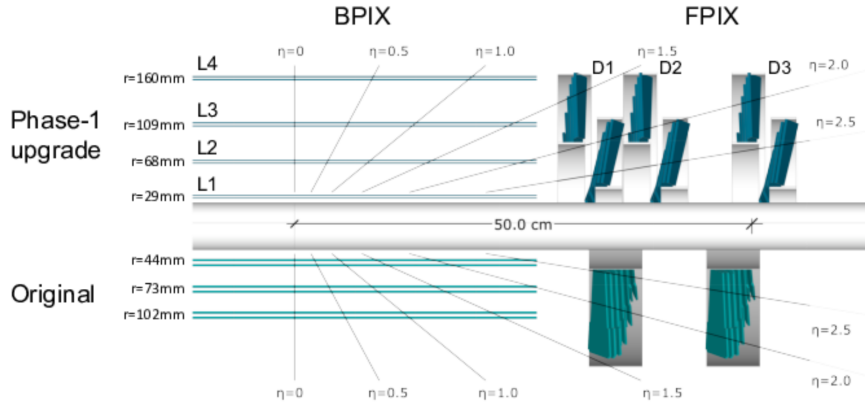


Figure 3.6: Illustration of the differences between the original design of the silicon tracker and the Phase-1 upgrade [17, 19, 20]. This thesis utilized data from LHC Run 2 after the Phase-1 upgrades to the tracker.

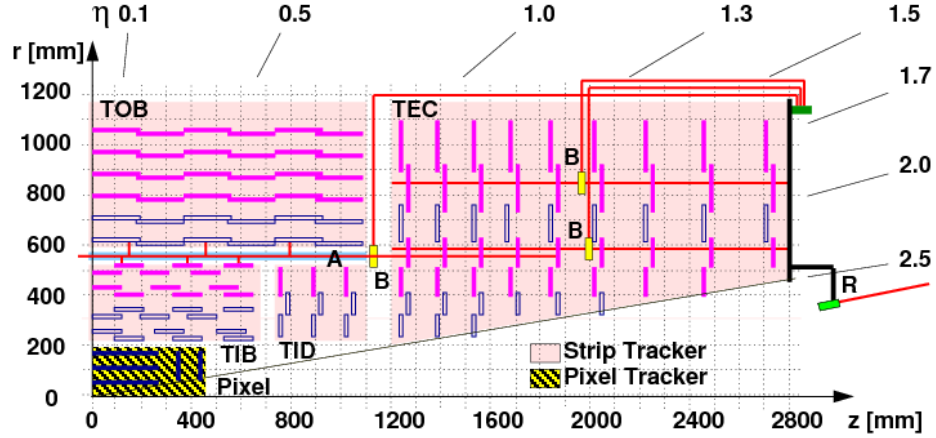


Figure 3.7: Schematic view of one quarter of the tracker in the r - z plane. This details the full silicon detector with all modules in their original positions.

A cross sectional view of the CMS tracker is shown in Figures 3.6 and 3.7. The tracker is composed of multiple silicon-based detector layers optimized for spatial resolution and radiation tolerance. It consists of two pixel subdetectors, the Tracker Pixel Barrel (TPB) and Tracker Pixel Endcap (TPE), which provide the highest granularity near the beamline. Beyond the pixel detectors, four silicon strip subdetectors extend tracking coverage: the Tracker Inner Barrel (TIB) and Tracker Inner Disks (TID)

provide intermediate precision measurements, while the Tracker Outer Barrel (TOB) and Tracker Endcap (TEC) contribute to tracking particles at larger radii from the collision point [21].

3.2.2.1 Pixel detector

The CMS pixel detector (a photograph of which is presented in Figure 3.8) contains an array of 124 million pixels and is positioned closest to the beam pipe, which makes it essential for reconstructing the short-lived decay products of heavy particles. In the Phase-1 geometry, the pixel detector consists of four cylindrical barrel (BPIX) layers located at radii of approximately 3 cm, 7 cm, 11 cm, and 16 cm, with additional forward and backward endcap disks (FPIX) extending coverage at angles along the beamline (regions with higher values of pseudorapidity, parameterized by η as seen in Figures 3.6 and 3.7).

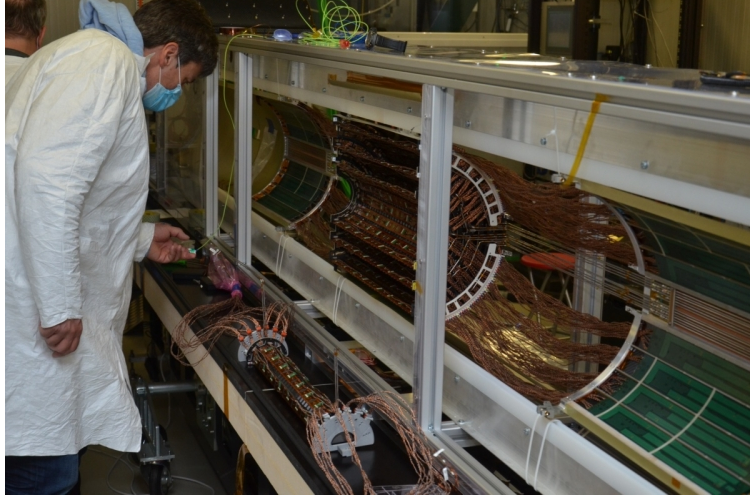


Figure 3.8: Silicon pixel detector before installation.

Each of the four layers is composed of silicon sensor modules, subdivided into individual

pixels as visualized in Figure 3.9. Each pixel measures $100\text{ }\mu\text{m} \times 150\text{ }\mu\text{m}$, comparable to the width of two human hairs. When a charged particle traverses a pixel sensor, it ionizes the silicon atoms along its path, liberating electron-hole pairs. A bias voltage applied across the sensor collects these charges, generating a small electric signal. This signal is subsequently amplified and processed by an electronic readout chip, with each module integrating 16 readout chips (ROCs) to handle data acquisition efficiently. Given the detector's close proximity to the beamline, the particle flux is extreme, reaching approximately 6×10^8 particles per cm^2 per second at a radius of 3 cm [13].

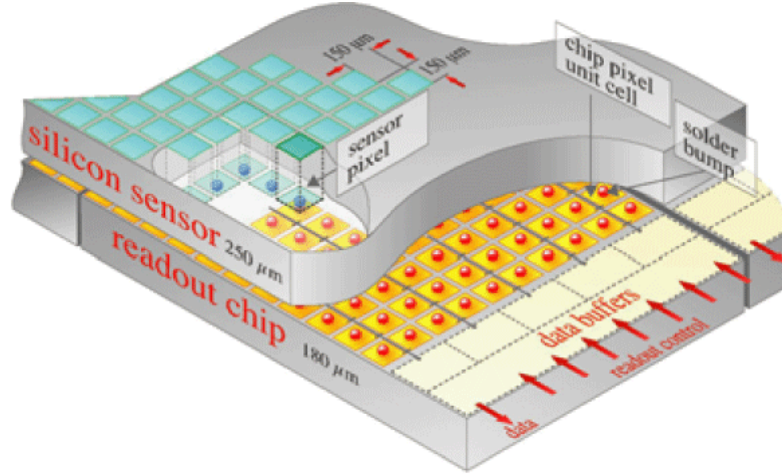


Figure 3.9: Illustration of the CMS silicon sensor modules.

3.2.2.2 Strip detector

After passing through the pixel detector, charged particles traverse ten successive layers of silicon strip detectors, extending the tracking coverage to a radius of 130 cm. These layers provide high-precision position measurements over a larger volume, complementing the pixel detector in track reconstruction, particularly for higher-momentum particles that travel greater distances before bending significantly in the

magnetic field [21]. A diagram and photograph of the strip layers are presented in Figures 3.10 and 3.11, respectively.

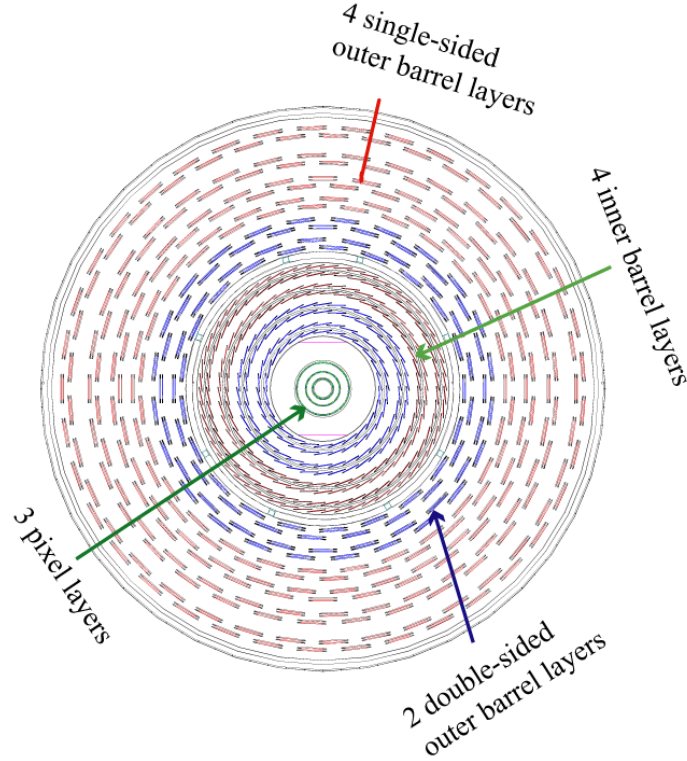


Figure 3.10: CMS Tracker layers shown in the plane perpendicular to the beam.

The silicon strip tracking system consists of multiple subdetectors, each optimized for its position within the detector. The inner tracking volume comprises the Tracker Inner Barrel (TIB), composed of four concentric layers, and the Tracker Inner Disks (TID), each consisting of three small disk-shaped structures. Encasing the TIB and TID is the Tracker Outer Barrel (TOB), which consists of six additional layers, providing extended coverage in the barrel region. At both ends of the tracker, the Tracker Endcaps (TEC) ensure full pseudorapidity coverage. The geometries of these subsystems are carefully optimized to balance spatial resolution, material budget, and mechanical stability [22].

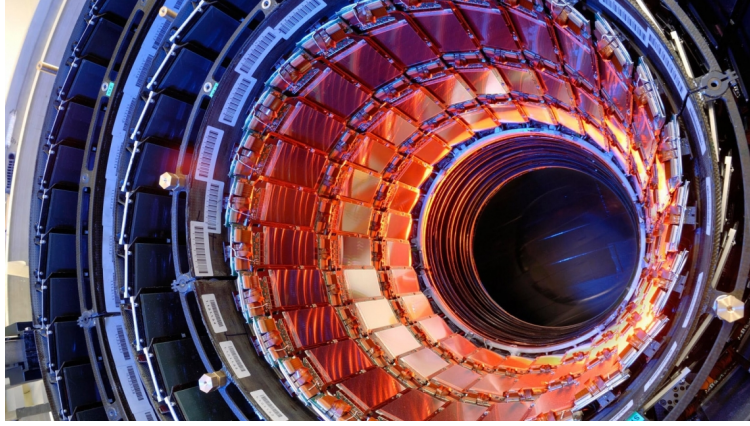


Figure 3.11: Photograph of the silicon strip detectors in the barrel module.

The silicon strip tracker contains a total of 15,200 highly sensitive modules, comprising approximately 10 million detector strips read out by 72,000 microelectronic chips. Each module consists of three primary components: one or two silicon strip sensors, a mechanical support structure ensuring precise alignment and stability, and readout electronics responsible for amplifying and processing signals. This extensive system enables precise tracking of charged particles throughout their passage within the CMS detector, ensuring robust vertex reconstruction and momentum measurements [23].

3.2.3 Electromagnetic calorimeter (ECAL)

The ECAL is the innermost of the two calorimeters in CMS and is designed to measure the energy of electrons and photons with high precision by fully absorbing them.

The ECAL houses approximately 80,000 lead tungstate (PbWO_4) crystals, a scintillating material selected for its high density ($\rho = 8.28 \text{ g/cm}^3$), short radiation length ($X_0 = 0.89 \text{ cm}$), and fast response time. Each crystal, primarily composed of lead bound with oxygen in a tungstate lattice structure, took two days to grow and has a

mass of 1.5 kg and is characterized by exceptional optical transparency despite its metallic composition. When traversed by high-energy electrons or photons, these crystals emit scintillation light in short, well-defined bursts, enabling precise energy measurements and particle identification. The ECAL barrel is shown in Figure 3.12.

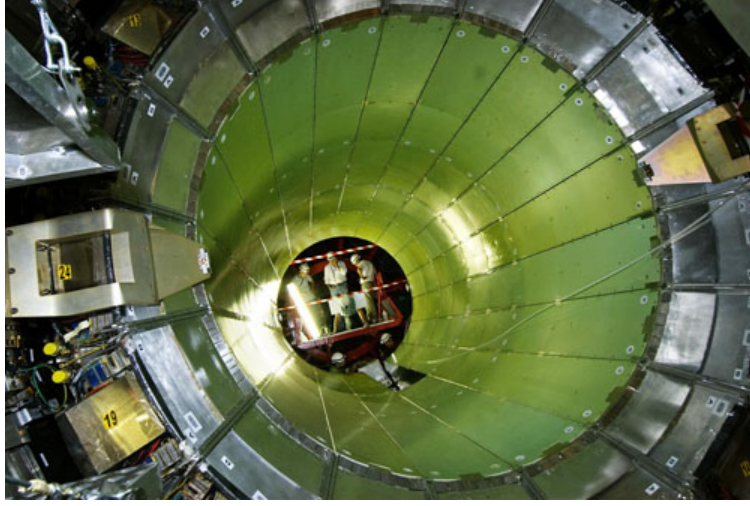


Figure 3.12: Photograph of the CMS ECAL during construction.

The emitted scintillation light is detected by photodetectors, specifically avalanche photodiodes (APDs) in the barrel region and vacuum phototriodes (VPTs) in the endcaps. These photodetectors convert the optical signal into an electrical signal, which is then amplified and digitized for precise energy reconstruction. The ECAL's excellent energy resolution ($\sim 1\%$ at 100 GeV) for high-energy electrons and photons makes it a critical component for the identification of physics phenomena such as Higgs boson decays and precision Standard Model measurements [13, 24, 25].

3.2.4 Hadron calorimeter (HCAL)

Hadrons, composite particles composed of quarks and gluons bound by the strong interaction, pass through the ECAL with minimal energy deposition before being absorbed in the HCAL [26, 27]. A photograph of the HCAL barrel is shown in Figure 3.13.

The HCAL is designed to measure the energy, position, and timing of hadronic showers. It consists of alternating layers of dense absorber material (brass and steel plates), which induces hadronic interactions, and fluorescent scintillator tiles, which produce optical signals proportional to the deposited energy. These scintillation signals are collected via wavelength-shifting fibers and transmitted to hybrid photodiodes (HPDs) for digitization [26, 27]. The absorber material is critical, as hadronic interactions necessitate a substantial thickness—material of ~ 5.82 nuclear interaction lengths in the barrel region and ~ 10 nuclear interaction lengths along the beam—to ensure full shower containment [26, 27].

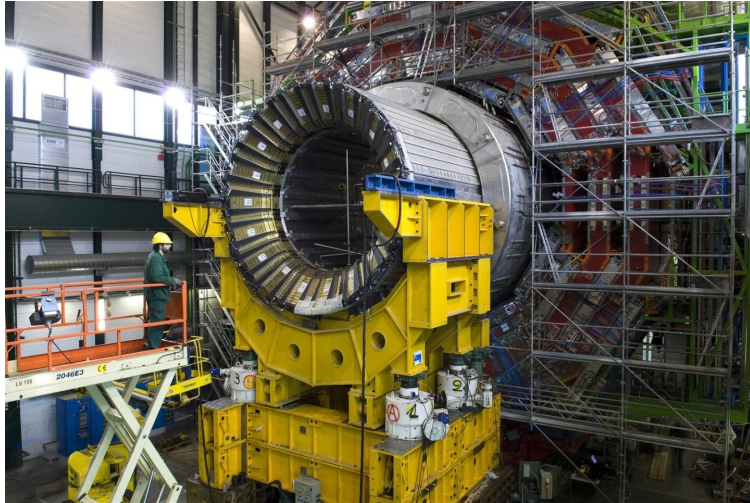


Figure 3.13: Insertion of the HCAL barrel inside the magnet in 2008.

The HCAL is structured into four primary components: the Hadron Barrel (HB), the Hadron Outer (HO), the Hadron Endcap (HE), and the Hadronic Forward (HF) calorimeters. The HB and HE employ brass absorber plates interleaved with plastic scintillator tiles, selected for their balance between density and cost-effectiveness. The HO, positioned outside the superconducting solenoid, captures late-developing showers that might otherwise escape detection. The HF, located at high pseudorapidity (η), utilizes quartz fibers embedded in a steel absorber, leveraging Cherenkov light production to make measurements in the forward region [28].

The HE calorimeters consist of 36 wedge-shaped modules, covering the pseudorapidity range $1.3 < |\eta| < 3.0$, extending hadronic energy measurements beyond the barrel region. The HF calorimeters, spanning $3.0 < |\eta| < 5.2$, play a crucial role in detecting forward jets and reconstructing missing transverse energy (E_T^{miss}), which is essential for identifying weakly interacting particles such as neutrinos or potential BSM candidates [29].

3.2.5 Muon system

The detection of muons is one of the primary objectives of the CMS experiment¹. Muons are charged leptons, similar to electrons and positrons, but approximately 200 times heavier. This increased mass grants them a significantly greater penetration depth because it reduces the energy lost by the traveling particles through various mechanisms, allowing them to traverse several meters of material with minimal energy loss. Unlike most particles produced in high-energy collisions, muons are not

¹While we depend on our detection of electrons too, muons are more difficult to capture and measure, so we will focus on them here.

fully absorbed by the electromagnetic (ECAL) or hadronic (HCAL) calorimeters. Consequently, dedicated muon detection chambers are positioned in the outermost regions of CMS, beyond the solenoid magnet, where muons are often the only particles producing a distinct and measurable signal.

The CMS muon system consists of four muon stations, which are interleaved with the iron plates of the return yoke. The system incorporates 1400 muon chambers, comprising 250 Drift Tubes (DTs) and 540 Cathode Strip Chambers (CSCs) responsible for tracking and triggering, along with 610 Resistive Plate Chambers (RPCs) and 72 Gas Electron Multiplier chambers (GEMs), which provide a redundant trigger to ensure rapid data acquisition and filtering.

Muon tracking is performed by reconstructing the curvature of muon trajectories within the 3.8 T magnetic field generated by the CMS solenoid, as illustrated in Figure 3.5 and Figure 3.14. The reconstruction of muon momentum is achieved through a combination of measurements from the muon chambers, which provide track trajectory information, and the silicon tracker, which offers high-precision tracking near the interaction point. The combined resolution for transverse momentum measurements (p_T) is approximately 3–5% for $p_T > 10$ GeV/ c .

At low transverse momentum ($p_T < 200$ GeV/ c), the resolution of the muon system is primarily limited by multiple scattering effects in the material preceding the muon stations. However, at high transverse momentum ($p_T > 200$ GeV/ c), the dominant limiting factor is the intrinsic spatial resolution of the muon chambers. The highest precision for low-momentum muons is obtained from the inner tracker alone, while for

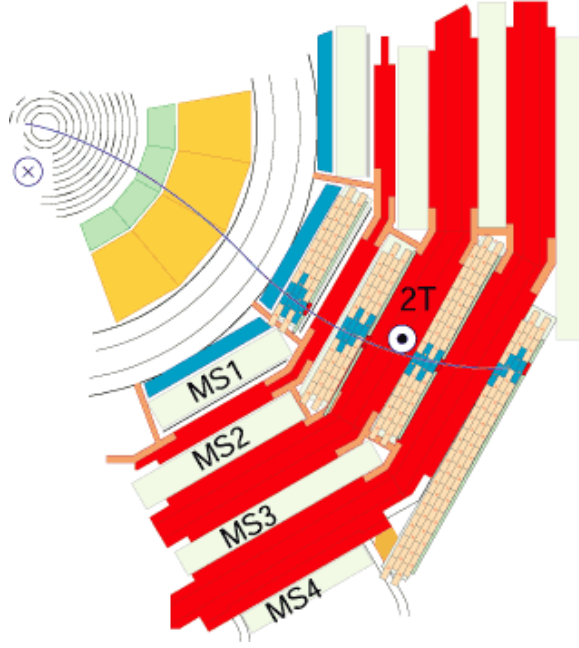


Figure 3.14: A muon, in the plane perpendicular to the LHC beams, leaves a curved trajectory in four layers of muon “stations.” Note that the muon passes through the silicon tracker first, which measures momentum, charge, trajectory, and impact parameter, or how close the muon’s path is to the primary vertex.

higher momentum muons, optimal resolution is achieved by combining measurements from both the muon system and the silicon tracker.

Accurate tracking requires precise alignment of the muon chambers and silicon tracker, which is achieved through cosmic ray data and dedicated calibration runs, ensuring sub-millimeter alignment precision. Each detector type undergoes rigorous calibration procedures to maintain accurate timing and spatial measurements. This includes gas gain calibration for the DT and CSC chambers, as well as timing calibrations for the RPC chambers. Overall, the CMS muon system is able to achieve an efficiency exceeding 95% for the detection of high-momentum muons.

3.2.5.1 Muon Drift Tubes (DTs)

The DT system is responsible for tracking and identifying muons within the barrel region of the CMS detector. It detects charged particles as they pass through an array of gas-filled DTs. Each tube is a cylindrical chamber, 4 cm in width, containing a thin, positively charged anode wire aligned along its axis. The tubes are filled with a gas mixture, typically argon and carbon dioxide (CO_2), chosen to optimize ionization efficiency and ensure stable electron drift.

When a muon or any charged particle traverses a DT, it ionizes the gas atoms, liberating electrons. These electrons, guided by an applied electric field, drift toward the central anode wire, where they are amplified and produce a measurable charge pulse. The electron drift velocity within the gas mixture remains relatively constant, enabling precise determination of the particle's trajectory. The time delay between ionization and signal detection provides an accurate measurement of the muon's crossing point within the drift cell.

Each DT chamber consists of multiple layers of staggered drift cells, grouped into units known as “superlayers.” Each superlayer contains four layers of DTs, offering redundant position measurements and improving spatial resolution. The DT system achieves a spatial resolution of approximately 1–2 mm, ensuring precise muon tracking.

The DT chambers are strategically positioned in the barrel region, covering a pseudo-rapidity range of $|\eta| < 1.2$. They are organized into four muon stations, which are embedded within the iron return yoke of the solenoid. The return yoke is structured into five wheels, each subdivided into twelve azimuthal sectors. Within this layout,

each DT chamber is paired with one or two RPCs, forming a complementary detection system that enhances muon identification and trigger efficiency.

A high-energy muon traversing the barrel region can interact with up to four DT chambers and six RPCs, resulting in as many as 44 recorded hits. This extensive tracking redundancy significantly improves momentum resolution and enhances robustness against inefficiencies in individual detectors. The DT system plays a crucial role in both tracking and triggering, delivering high-precision spatial and timing measurements that contribute to accurate muon momentum reconstruction [30–32].

3.2.5.2 Cathode Strip Chambers (CSCs)

CSCs are employed in the endcap disks, where the magnetic field is non-uniform, and particle fluxes are significantly higher than in the barrel region. These chambers are designed to provide precise muon tracking and triggering capabilities in this challenging environment. Each CSC consists of a multi-wire proportional chamber (MWPC) structure, comprising an array of positively charged anode wires intersecting with negatively charged copper cathode strips within a gas-filled chamber.

When a muon traverses a CSC, it ionizes the gas atoms, liberating electrons. These free electrons drift toward the anode wires under the influence of a strong electric field, initiating an avalanche multiplication process that results in a detectable charge pulse. Simultaneously, the positively charged ions drift toward the cathode strips, inducing a secondary charge signal. Since the anode wires are oriented perpendicularly to the cathode strips, the combined readout from both elements provides two-dimensional position information for the muon track.

CSCs are strategically positioned in the endcap regions, covering a pseudorapidity range of $1.2 < |\eta| < 2.4$. Each chamber consists of finely spaced cathode strips that enable high-resolution position measurements in the radial direction, achieving a spatial resolution of approximately 1 mm. The use of cathode strip readout, in conjunction with anode wire signals, allows for robust tracking and precise reconstruction of muon trajectories.

Given their location in the high-radiation forward region, CSCs play a crucial role in both muon tracking and triggering. They contribute significantly to the Level-1 (L1) trigger system, providing rapid and accurate muon identification to facilitate real-time event selection. Their ability to operate efficiently in regions with high background rates and strong magnetic field gradients makes them an essential component of the CMS muon detection system [32–34].

3.2.5.3 Resistive Plate Chambers (RPCs)

RPCs are fast gaseous detectors that play a crucial role in the CMS muon trigger system. They consist of two parallel resistive plates acting as the anode and cathode, separated by a thin gas-filled volume. These plates are constructed from a high-resistivity plastic material, which helps localize the induced signal and ensures stable operation at high particle rates.

CMS employs double-gap RPCs, where two parallel high-resistivity Bakelite electrodes enclose a 2 mm-wide gas gap. This configuration is repeated with another identical gap, separated by a layer of copper readout strips. Operating in avalanche mode, these chambers achieve excellent time resolution while maintaining efficiency under

high-radiation conditions. The RPCs identify muon candidates by analyzing the pattern of fired strips across multiple layers, comparing them against expected Monte Carlo predictions. The trigger system utilizes this information to determine whether the event data should be retained.

The RPCs are strategically positioned within both the barrel ($|\eta| < 1.6$) and endcap ($1.6 < |\eta| < 2.4$) regions, ensuring comprehensive muon detection coverage. Their construction consists of resistive plates separated by a carefully optimized gas mixture, typically composed of tetrafluoroethane ($\text{C}_2\text{H}_2\text{F}_4$) and sulfur hexafluoride (SF_6), which enhances ionization properties and detector longevity. Charged particles traversing the RPC volume ionize the gas, triggering localized electron avalanches that induce a fast signal across the resistive plates. This mechanism enables a timing resolution on the order of 1 ns, making RPCs particularly well-suited for real-time event selection.

Due to their fast response time and high granularity, RPCs serve primarily as a trigger detector, efficiently identifying high-energy muons and significantly contributing to the CMS L1 trigger system. Their integration with the DT and CSC subsystems allows for robust muon tracking and precise momentum reconstruction, enhancing the overall performance of the CMS muon system [32, 35].

3.2.5.4 Gas Electron Multipliers (GEMs)

GEMs represent a recent enhancement to the CMS muon detection system. They have been introduced to complement the existing detectors in the forward region near the beampipe, an area subject to increasingly high radiation doses and event rates during Phase-2 of the LHC. Their incorporation aims to improve tracking resolution

and trigger performance in this challenging environment.

Similar to other CMS muon subsystems, GEMs are gaseous detectors. Their distinguishing feature is the GEM foil, a thin ($50\text{ }\mu\text{m}$) insulating polyimide layer coated on both sides with copper conductors. This foil is perforated with microscopic holes arranged in a regular hexagonal pattern. The CMS GEM chambers are structured with two printed circuit boards (PCBs) enclosing a gas volume, within which a stack of three GEM foils is positioned.

The operational principle of GEMs relies on the ionization of an argon-carbon dioxide (Ar/CO_2) gas mixture by incoming charged particles. A strong potential difference applied across the GEM foils generates intense electric fields within the microscopic holes. Electrons liberated during the ionization process drift toward the foils and undergo successive amplification within the holes, producing an electron avalanche. The resulting charge multiplication induces a readout signal on finely spaced readout strips, enabling precise spatial resolution and robust tracking capabilities.

By integrating GEMs into the CMS muon system, the experiment enhances its ability to detect and reconstruct muons in the high-occupancy forward region, thereby contributing to improved efficiency and performance in the demanding conditions of LHC Phase-2 [\[36\]](#).

3.3 Triggers and Data Acquisition (DAQ) system

At the LHC design luminosity, the proton-proton (pp) collision rate exceeds 1 GHz, but only a small fraction of these events are relevant to CMS physics research. Due

to storage and computational constraints, only a limited subset can be retained for offline analysis. Efficient event reconstruction is crucial for identifying key physics objects, such as muons, electrons, and jets, within the most significant events.

Muon detection is central to the CMS trigger system, as high-energy muons often signal rare processes, including potential BSM physics. With a bunch crossing rate of 40 MHz, the LHC generates approximately 10^9 pp interactions per second. Given that each recorded event averages 1.5 MB and only $\mathcal{O}(1000)$ events per second can be stored, the trigger system must achieve an event rejection factor of $\mathcal{O}(10^6)$ while maintaining high efficiency for key signals.

The CMS trigger system reduces the event rate in real time while preserving sensitivity to rare physics processes. It consists of two stages: the L1 trigger, implemented in custom hardware, and the High-Level Trigger (HLT), which runs software algorithms on a computing farm. Together, these systems ensure that only the most relevant events are stored for analysis [37].

3.3.1 Level-1 (L1) trigger

The L1 trigger rapidly processes detector signals to identify candidate physics objects, such as electromagnetic energy deposits, hadronic jets, and ionization patterns from muons. To manage the 100 kHz output limit of CMS readout electronics, the trigger dynamically adjusts selection thresholds based on the LHC’s instantaneous luminosity.

Using coarse-grained detector data, the L1 trigger selects events with high transverse momentum particles and anomalous energy deposits. Muon chambers—including DTs,

CSCs, and RPCs—provide initial p_T estimates for candidate muons. Operating within a strict latency of a few microseconds, the L1 trigger ensures rapid event selection in the LHC’s high-rate environment. The LHC delivers bunch crossings at 40 MHz, and the L1 trigger reduces the incoming event rate to ~ 100 kHz.

3.3.2 High-Level Trigger (HLT)

The DAQ system assembles L1 data, and passes it to the HLT which refines L1 selections by incorporating full-granularity detector data and advanced reconstruction algorithms. For muon identification, it integrates muon chamber signals with precise silicon tracker measurements, improving momentum resolution and suppressing background. By applying more stringent selection criteria, the HLT ensures that only the most promising events are retained for physics analyses.

Events passing the L1 trigger are transferred to front-end readout buffers before being processed by the HLT on a dedicated computing farm. The HLT software further reduces the event rate from 100 kHz to 1 kHz for mass storage. Event selection follows a staged approach: initial decisions rely on calorimeter and muon system data, followed by inner tracker pixel information, and ultimately, full event reconstruction when necessary.

To optimize storage efficiency, events are categorized into non-exclusive data streams based on HLT decisions. Most data undergo immediate processing, while lower-priority datasets can be deferred using a technique known as “parking.” This capability allows CMS to record additional data on tape at the expense of increased processing latency.

3.4 Computing

Even with the triggers reducing the amount of unnecessary data coming out of our detector, CMS still produces a huge amount of data that must be analyzed—more than 5 petabytes (Pb) per year when running at peak performance. The storage and computational demands of processing such data far exceed the capabilities of a single facility, necessitating a globally distributed computing infrastructure.

To manage and process CMS data efficiently, the Worldwide LHC Computing Grid (WLCG) provides a hierarchical computing architecture, integrating resources from 170 computing centers across 42 countries. This infrastructure consists of four distinct tiers, each with specialized functions [37–39].

Following event selection by the trigger system, accepted events are transferred to the CMS Storage Manager for archival storage and subsequent processing. Initially, data are stored on local disks before being transmitted to the Tier-0 center at CERN. This facility is responsible for the initial reconstruction of raw collision data (the first step toward physics analysis) and creates a primary backup before distributing the data to Tier-1 centers for further processing.

Tier-1 centers, located in thirteen national laboratories and research institutes across France, Germany, Russia, Italy, Spain, Taiwan, the United Kingdom, and the United States, receive data from Tier-0. These centers perform a second round of event reconstruction, incorporating refined calibration constants derived from ongoing experimental analysis. They also provide long-term storage and distribute datasets to Tier-2 centers.

Around 160 Tier-2 centers, primarily based at universities and research institutions, receive subsets of data from Tier-1. These facilities are dedicated to specific analyses and large-scale Monte Carlo simulations required for physics studies.

Tier-3 consists of smaller computing clusters, including individual research group resources, which provide additional flexibility for end-user analysis. Unlike the other tiers, Tier-3 centers do not have a formal commitment to WLCG but are crucial for local data processing.

3.5 Particle reconstruction

The Particle Flow (PF) algorithm is designed to achieve a comprehensive reconstruction and identification of all stable particles in an event by leveraging information from all CMS sub-detectors. By optimally combining measurements from the tracker, calorimeters, and muon systems, the PF algorithm determines the direction, energy, and particle type with high precision. I will introduce the concept here, but elaborate on the details of particle reconstruction and selection as they apply to this thesis in [Section 4.2.1](#).

The reconstructed list of individual particles serves as the foundation for higher-level physics objects. It is treated similarly to final-state particles generated by Monte Carlo event generators, enabling the construction of hadronic jets, from which the kinematic properties of originating quarks and gluons can be inferred. Additionally, the PF algorithm plays a crucial role in determining missing transverse energy (MET), providing an estimate of the momentum carried away by neutrinos and other undetected

particles. Furthermore, it facilitates the reconstruction and identification of particles through their decay products, enhancing the sensitivity of CMS analyses to a wide range of physics processes [40].

3.6 Alignment of the CMS tracker

As aforementioned, the precision of the silicon tracker is critical to CMS, especially in measuring the trajectories of particles within our solenoid. The alignment of the tracker was also crucial to our research group, which contributes to the experimental service effort and is the sole maintainer of one of primary tracker alignment algorithms. Over the course of my thesis, I ran alignment campaigns and validated alignment candidates, as well as provided some updates to our Hits-and-Impact-Points alignment algorithm also known as “HipPy”—the algorithm and my efforts are described in Sections 3.6.2 and 3.6.4, respectively.

3.6.1 Parameterization

Various parts of the CMS detector have been detailed in Section 3.2. The pixel detector and strip tracker, in particular, can be further divided into their subdetector structures or modules. The hierarchical structure of the CMS tracker is illustrated in Figure 3.15.

The accuracy of track reconstruction is significantly affected by uncertainties in module positioning. If the module position is not precisely known, the recorded hit positions and the expected impact points (extrapolated from other measured hits) of a track will exhibit systematic displacements. The discrepancy between these two quantities,

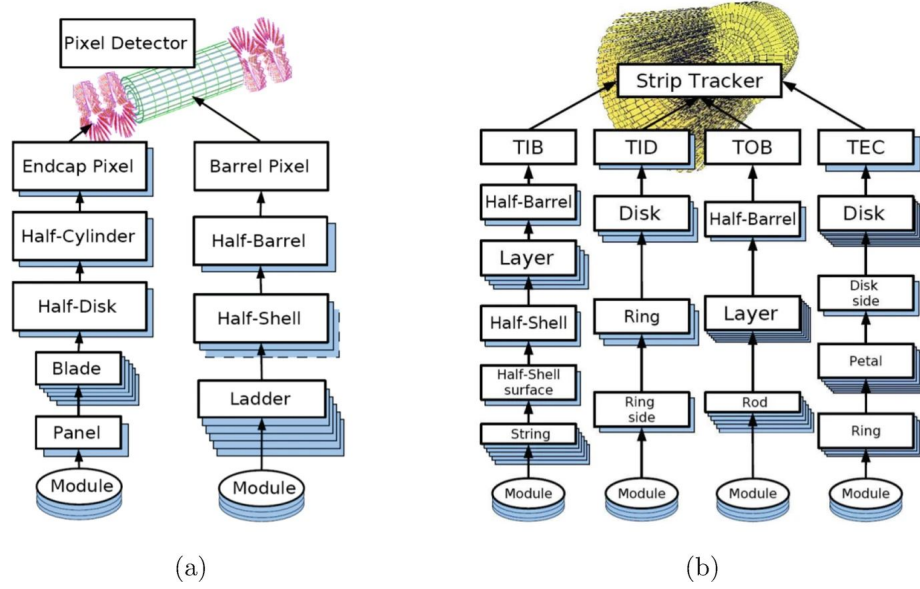


Figure 3.15: Hierarchical structures of the pixel detector and strip tracker [41].

expressed in the local coordinate system of the module, defines the *track-hit residuals* illustrated in Figure 3.16 [42].

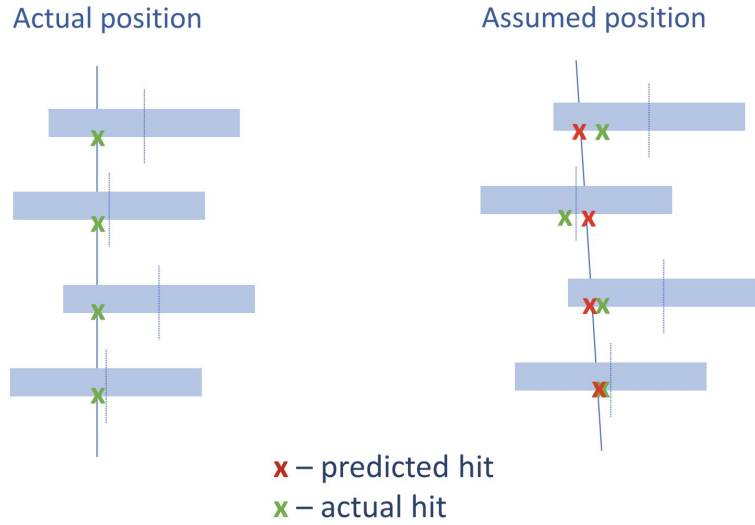


Figure 3.16: Simple illustration of a track through misaligned layers [43].

Each detector module is parameterized within a local coordinate system (u, v, w) , as

illustrated in Figure 3.17. The coordinate axes are defined such that the w -axis is normal to the module plane, while the u - and v -axes lie within the plane. The u -axis is oriented along the more sensitive measurement direction. Rotational misalignments of the module are described by three angles: α , β , and γ , which correspond to rotations about the u -, v -, and w -axes, respectively. Precise calibration of these parameters is essential for minimizing systematic biases in track reconstruction [42, 44].

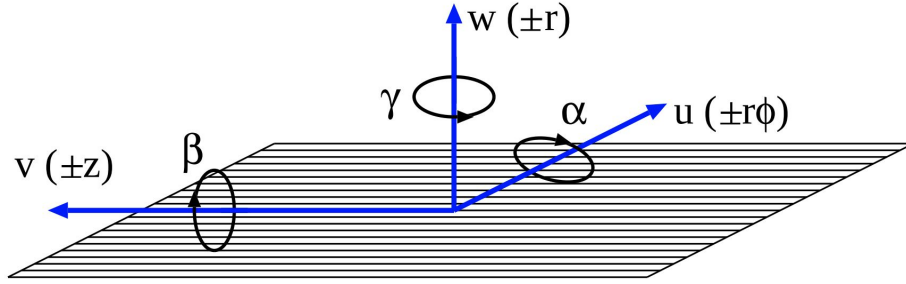


Figure 3.17: Example local coordinates of a module. Global parameters are shown in parentheses for modules in the TIB and TOB [41].

Each module is also parameterized in a global coordinate system (x, y, z) . The transformation from the global coordinate system to the local module coordinate system is expressed as

$$\mathbf{q} = \mathbf{R}(\mathbf{r} - \mathbf{r}_0), \quad (3.3)$$

where $\mathbf{r} = (x, y, z)$ represents the position in the global coordinate system, $\mathbf{q} = (u, v, w)$ denotes the corresponding position in the local module system, $\mathbf{r}_0 = (x_0, y_0, z_0)$ is the position of the detector module center in global coordinates, and \mathbf{R} is a rotation matrix that accounts for the module's orientation. This transformation ensures that local hit positions can be consistently related to the global detector geometry.

Due to the hierarchal nature of our modules, it is sometimes useful to describe the

position of a group of sensors belonging to a composite tracker structure. In this case, we use a simple change of variable to represent the same definitions as above, but for the larger structures. \mathbf{g} represents the coordinates in the composite basis, \mathbf{g}_0 gives the composite origin in global coordinates, and \mathbf{G} rotates from the global to the composite system [42].

Initially, the rotation matrix \mathbf{R} and the detector positions \mathbf{r}_0 can be determined through detector assembly and survey measurements. Then, after assembly, these parameters will be corrected incrementally through alignment procedures to account for mechanical shifts, thermal expansion, and other systematic effects that may alter the module positions over time. It should be noted that tracker alignment is not a physical procedure, but rather a correction of our interpretation of the detector data in software.

3.6.2 Algorithms

Two primary algorithms are used by CMS to minimize track-hit residuals and improve alignment: Millepede and HipPy.

3.6.2.1 *Millepede*

Millepede-II (colloquially referred to as “Millepede” or “MP”) employs a global matrix approach, eliminating the need for iterative procedures. Millepede-II [45] is an upgraded version of the original Millepede [46] program, designed to perform a global fit by minimizing the χ^2 function while simultaneously accounting for both track and alignment parameters. Since angular corrections remain small, the problem can be

effectively linearized while remaining a suitable approximation for alignment studies.

The optimization MP performs is structured to focus exclusively on the n alignment parameters, reducing the problem to solving a matrix equation of size n [41]. The χ^2 function, expressed in terms of both local track parameters \mathbf{q} and global alignment parameters \mathbf{p} , is given by

$$\chi^2(\mathbf{p}, \mathbf{q}) = \sum_j \sum_i \frac{(y_{ji} - f_{ji}(\mathbf{p}, \mathbf{q}_j))^2}{\sigma_{ji}^2} \quad (3.4)$$

where y_{ji} represents uncorrelated hit measurements of track j with associated uncertainties σ_{ji} , while $f_{ji}(\mathbf{p}, \mathbf{q}_j)$ denotes the predicted measurement based on the alignment and track parameters. Given reasonable initial estimates \mathbf{p}_0 and \mathbf{q}_{j0} , the function f_{ji} can be linearized. Applying the least-squares method to minimize χ^2 results in a system of linear equations, with one equation per alignment parameter and all the track parameters of each track [41].

The system of equations can be reduced, ultimately leading to a matrix equation of the form $\mathbf{C}\mathbf{a} = \mathbf{b}$ where \mathbf{a} corresponds to small corrections to the initial alignment parameters \mathbf{p}_0 . This formulation allows for efficient computation, making Millepede-II particularly well-suited for large-scale alignment tasks.

3.6.2.2 HipPy

The Hits-and-Impact-Points (HIP) algorithm [42, 44] served as the predecessor to what later became HipPy. Initially implemented during the commissioning of the CMS

tracker [41] and extensively used throughout Run 1 [21, 47], the algorithm underwent significant modifications over time [48]. These refinements led to the development of an improved version, designated as Hits-and-Impact-Points-Past-Year-1 (HipPy), which was formally introduced at the start of Run 2.

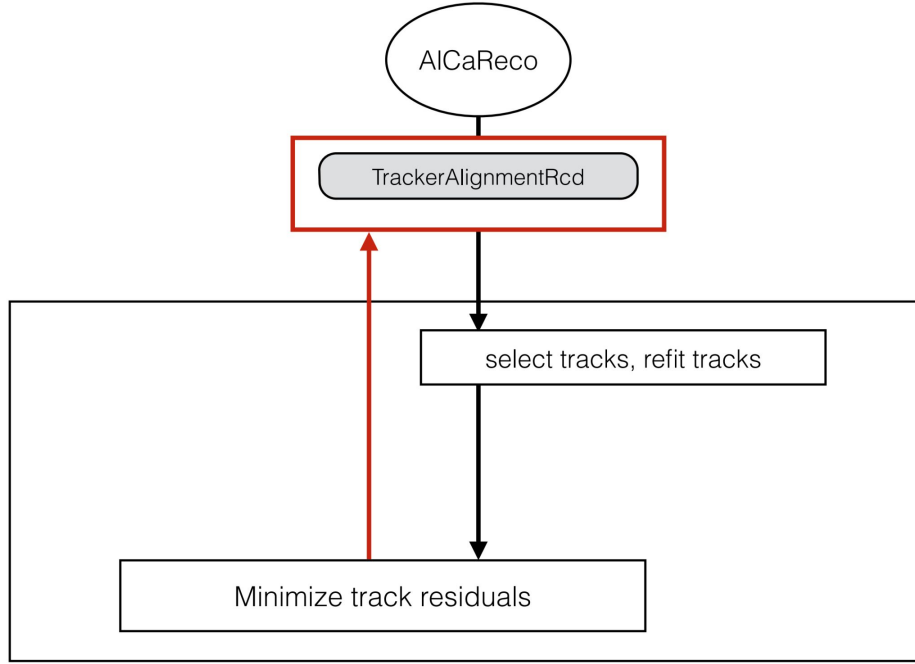


Figure 3.18: Process diagram of the Hits-and-Impact-Points (HIP) alignment algorithm [43].

Key advancements in the alignment methodology after CMS' first year online included the addition of three alignment parameters to account for sensor curvature, extending beyond the conventional six parameters that describe sensor position and orientation. Enhancements were also made to allow for the weighting of specific input types, the introduction of a sequential and hierarchical alignment strategy across multiple time periods (multi-IOV), and the incorporation of optional mass and vertex constraints in events with well-characterized underlying physics processes. These improvements

significantly enhanced the alignment precision and stability, contributing to the success of Run 2. The HipPy-based alignment parameters were adopted for CMS data reconstruction and were used in both prompt and end-of-year (EOY) alignments alongside Millepede-II (MP) constants [48].

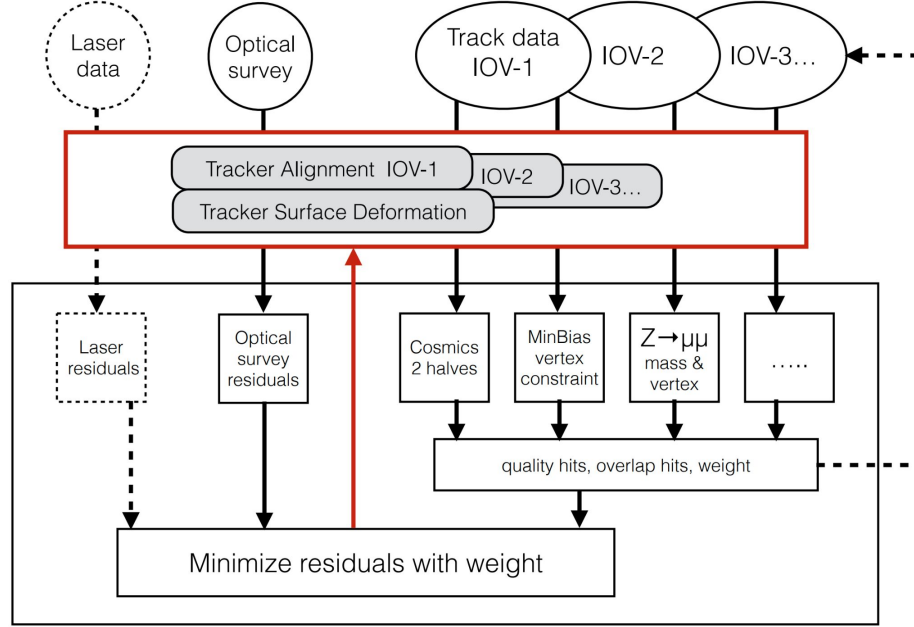


Figure 3.19: Process diagram of the Hits-and-Impact-Points-Past-Year-1 (HipPy) alignment algorithm [43].

The fundamental principle of HipPy involves iteratively refining the local alignment of individual detector sensors while keeping all other parameters fixed. The procedure begins with loading track data and hit measurements, followed by track fitting using the current estimate of alignment parameters. A χ^2 function, defined in Eq. (3.5), is computed for the selected hits for a given sensor (module). For each sensor, the alignment parameters are updated by minimizing χ^2 with respect to local shifts while holding all other sensor parameters constant. The process iterates until convergence is achieved.

The χ^2 function for HipPy can be written using a covariance matrix \mathbf{V} of the measurement uncertainties

$$\chi^2 = \sum_i^{\text{hits}} \mathbf{r}_i^T(\mathbf{p}, \mathbf{q}) \mathbf{V}_i^{-1} \mathbf{r}_i(\mathbf{p}, \mathbf{q}) \quad (3.5)$$

where $\mathbf{r}_i(\mathbf{p}, \mathbf{q})$ are residuals dependent on both track parameters \mathbf{q} and alignment parameters \mathbf{p} . The update step for alignment parameters is obtained from

$$\mathbf{p}_m = \left[\sum_i^{\text{hits}} \mathbf{J}_i^T \mathbf{V}_i^{-1} \mathbf{J}_i \right]^{-1} \left[\sum_i^{\text{hits}} \mathbf{J}_i^T \mathbf{V}_i^{-1} \mathbf{r}_i \right] \quad (3.6)$$

where \mathbf{J}_i denotes the Jacobian matrix of residuals with respect to the alignment parameters, and can be solved analytically via the small angle approximation. Correlations between different modules and effects on the track parameters are accounted for by iterating the minimization process and by refitting the tracks with new alignment constants after each iteration [41].

A key advantage of HipPy is its seamless integration with the CMS software framework, particularly in the context of track reconstruction. This integration enables the use of kinematic constraints such as mass and vertex constraints, which improve alignment precision. Additionally, HipPy simplifies development by leveraging Kalman filter-based algorithms. The algorithm treats the position and orientation of each sensor independently in every iteration. While each step consists of a straightforward matrix inversion [41, 42], multiple iterations are required to account for correlations between

sensor parameters. This necessitates repeated track fitting which can become time- and CPU-intensive.

3.6.3 Validation

Several validation methods are used within CMS to quantify the quality of calculated alignment parameters. Some common validation methods include Geometry Comparisons (GC), Distribution of the Median of Residuals (DMR), Primary Vertex (PV), cosmic track splitting, and $Z \rightarrow \mu\mu$ or “dimuon” validation [43].

3.6.3.1 Geometry Comparison (GC)

GC validations provide a direct and effective means of visualizing differences between tracker alignments. Once an alignment fit is performed, the resulting geometry is compared against a reference geometry, which may correspond to the design geometry or a previously aligned dataset. Systematic differences observed in such comparisons can reveal potential distortions in the tracker structure. While geometry comparisons alone do not establish the validity of systematic shifts in module positions, they serve as an essential diagnostic tool, offering insights into possible alignment effects.

Unphysical distortions, identifiable through detector design constraints, can signal biases in the alignment procedure. The geometry comparison framework, adapted from HipPy’s optical survey constraint tools, aligns two geometries by applying translations and rotations before computing their differences.

To facilitate meaningful comparisons, global shifts and rotations of large structures are

removed, allowing for precise measurement of module displacements, denoted as Δz , Δr , and $\Delta\phi$, as functions of z , r , and ϕ , respectively. Additionally, the remaining three spatial coordinates (x , y , and z) and three angular rotations can also be visualized in a similar manner. Since geometry comparison validation does not rely on reconstructed track data, it can be applied universally with respect to any reference geometry, including design, survey, or previously aligned track-based geometries. An example of such a validation is shown in Figure 3.20.

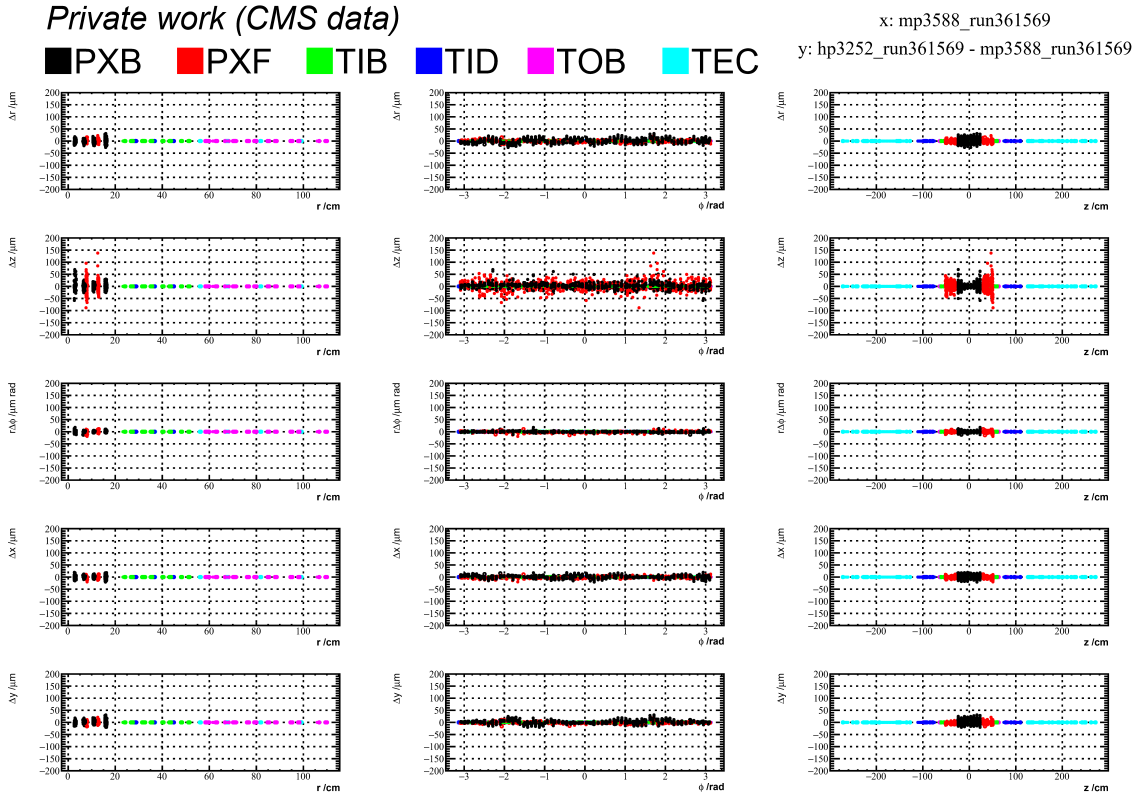


Figure 3.20: Example global tracker plot for GC validation, comparing analogous HipPy and Millepede alignments using beam data, both calculated from PromptGT.

3.6.3.2 *Distribution of the Median of Residuals (DMR)*

By analyzing each module individually, DMR validations provide a measure of alignment improvement across subdetectors. This method involves refitting each track using all hits except for the one associated with the module under study. The residual, defined as the difference between the predicted and measured hit positions, is computed for each track, and the median of this distribution serves as a representative statistic for the module. These median residuals are then plotted across all modules, allowing for a comprehensive alignment evaluation.

To study various sources of bias, DMR validation is performed separately for the local x and y coordinates. A well-aligned detector should produce a DMR distribution centered around zero, with its width reflecting the statistical precision of the alignment (assuming a large enough sample of hits is used for validation). Deviations from zero suggest systematic biases, while a broader width indicates limitations in alignment accuracy. Since the width also depends on the sample size, valid comparisons require that all datasets have identical track statistics, as ensured in the validation process. Examples of DMR validation are shown in Figures [3.21](#), [3.22](#), [3.23](#), [3.24](#), and [3.28](#).

3.6.3.3 *Primary Vertex (PV)*

The quality of tracker alignment is further assessed through its impact on the performance of physics objects, particularly the reconstruction of the PV. The PV, typically defined as the vertex associated with the track of highest transverse momentum (p_T), is primarily determined by the pixel detector, as it is closest to the interaction point and offers the best intrinsic hit position resolution.

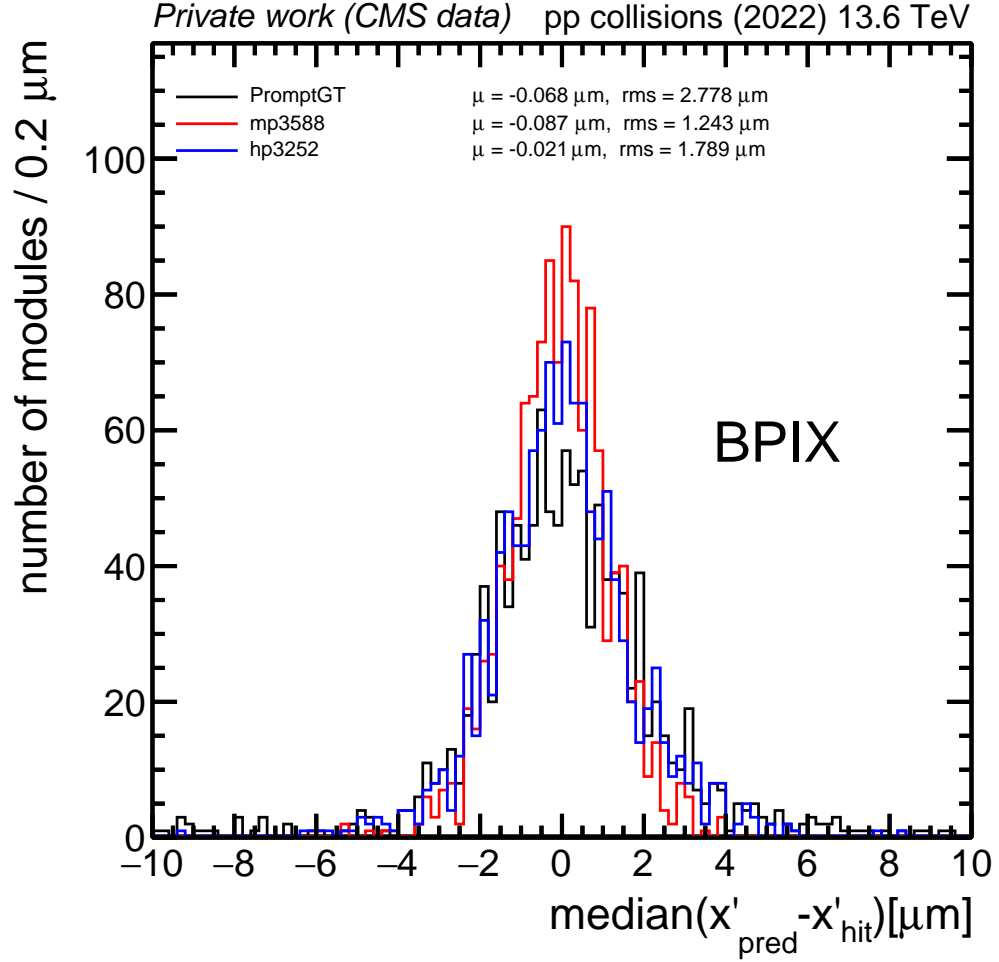


Figure 3.21: Example DMR validation plot for x' residual in the BPIX, comparing HipPy (blue) and Millepede (red) alignments using beam data, both calculated from the starting geometry of PromptGT (black).

The precision of vertex reconstruction is evaluated using unbiased track-vertex residuals, also known as track impact parameters, which provide a direct measurement based on data. For each track, the PV position is reconstructed while excluding the track under consideration. To ensure robustness against pileup, a deterministic annealing clustering algorithm is utilized. Examples of PV validation are shown in Figures 3.25 and 3.26.

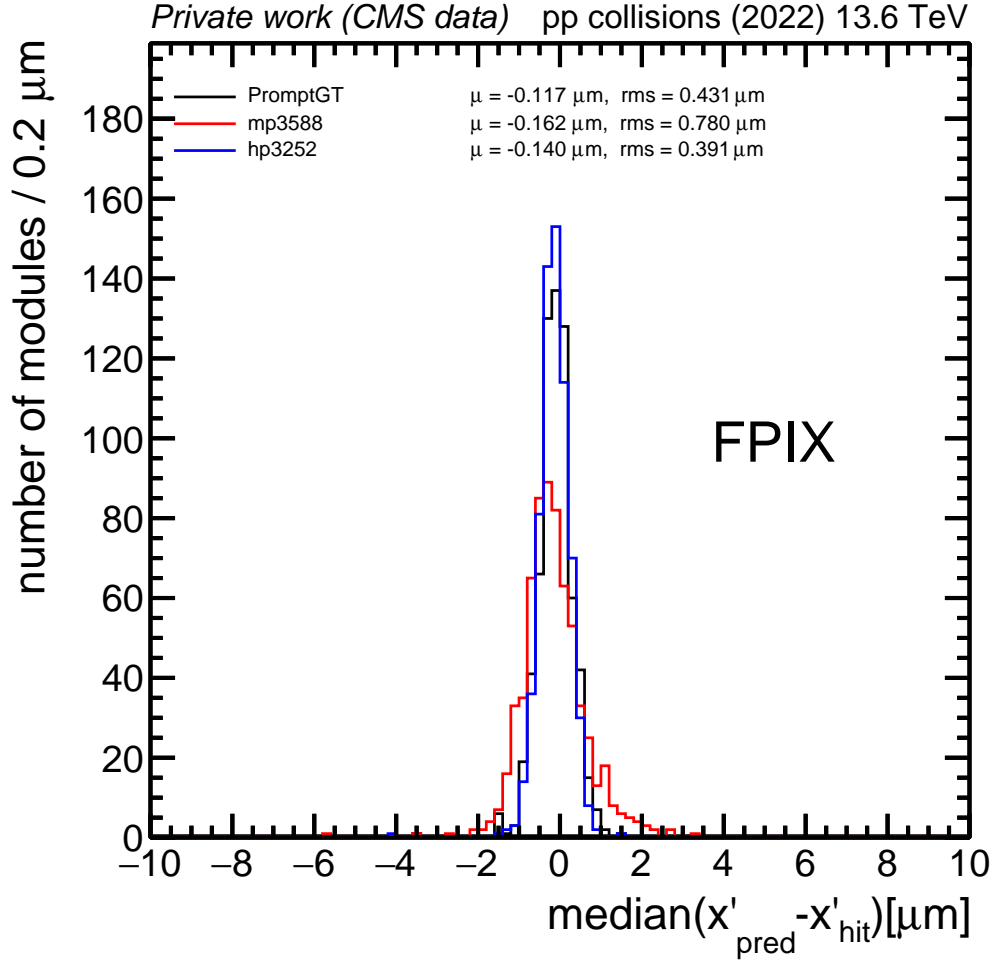


Figure 3.22: Example DMR validation plot for x' residual in the FPIX, comparing HipPy (blue) and Millepede (red) alignments using beam data, both calculated from the starting geometry of PromptGT (black).

3.6.3.4 Cosmic track splitting

Cosmic ray muons, often referred to as cosmics or cosmic events, recorded by the CMS detector are used for detector commissioning and calibration. Before the magnetic field is turned on, events are recorded during the Cosmic RUNs at ZERo Tesla (CRUZET). Cosmic ray muon tracks are also recorded in the 3.8 T magnetic field provided by the

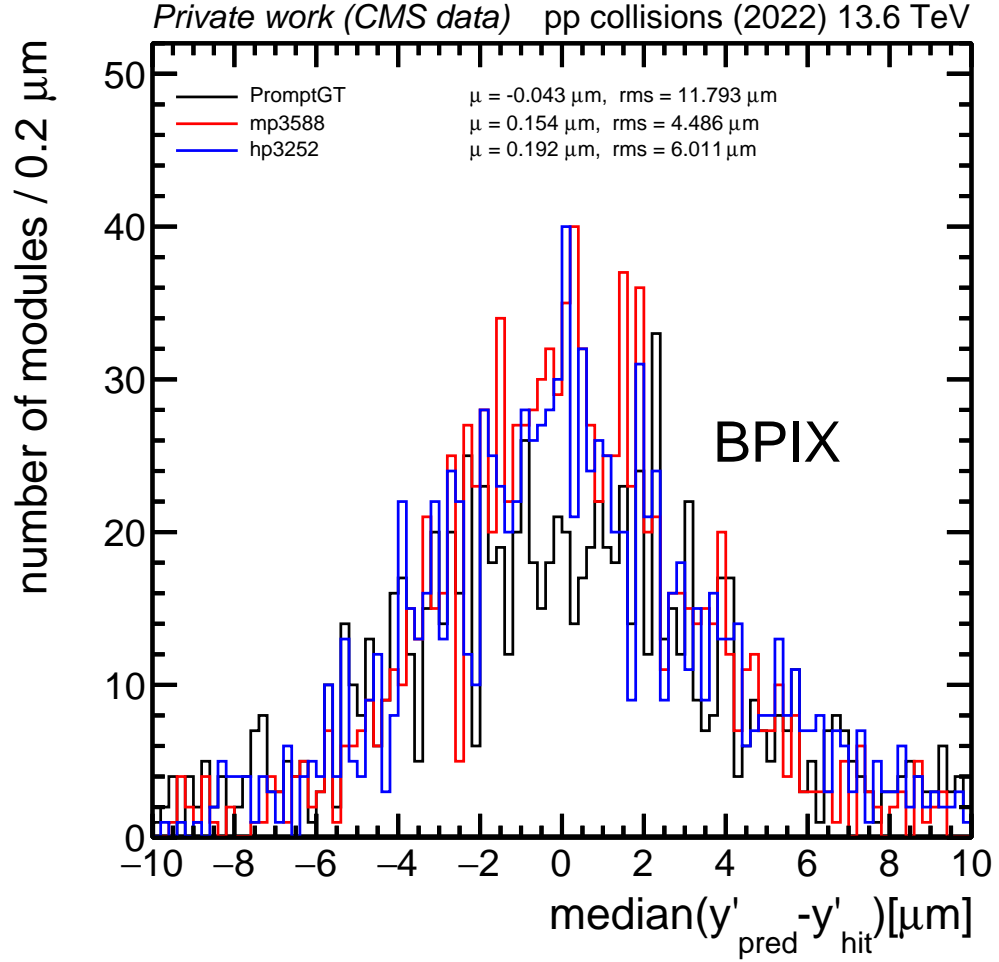


Figure 3.23: Example DMR validation plot for y' residual in the BPIX, comparing HipPy (blue) and Millepede (red) alignments using beam data, both calculated from the starting geometry of PromptGT (black).

CMS solenoid during the Cosmic Runs At Four Tesla (CRAFT).

Tracks from cosmic ray muons are crucial for deriving alignment constants for two main reasons. First, they can be recorded before the start of LHC collisions and are therefore used to derive the initial alignment corrections after a shutdown period. Second, they have a very different topology compared to collision tracks. Unlike

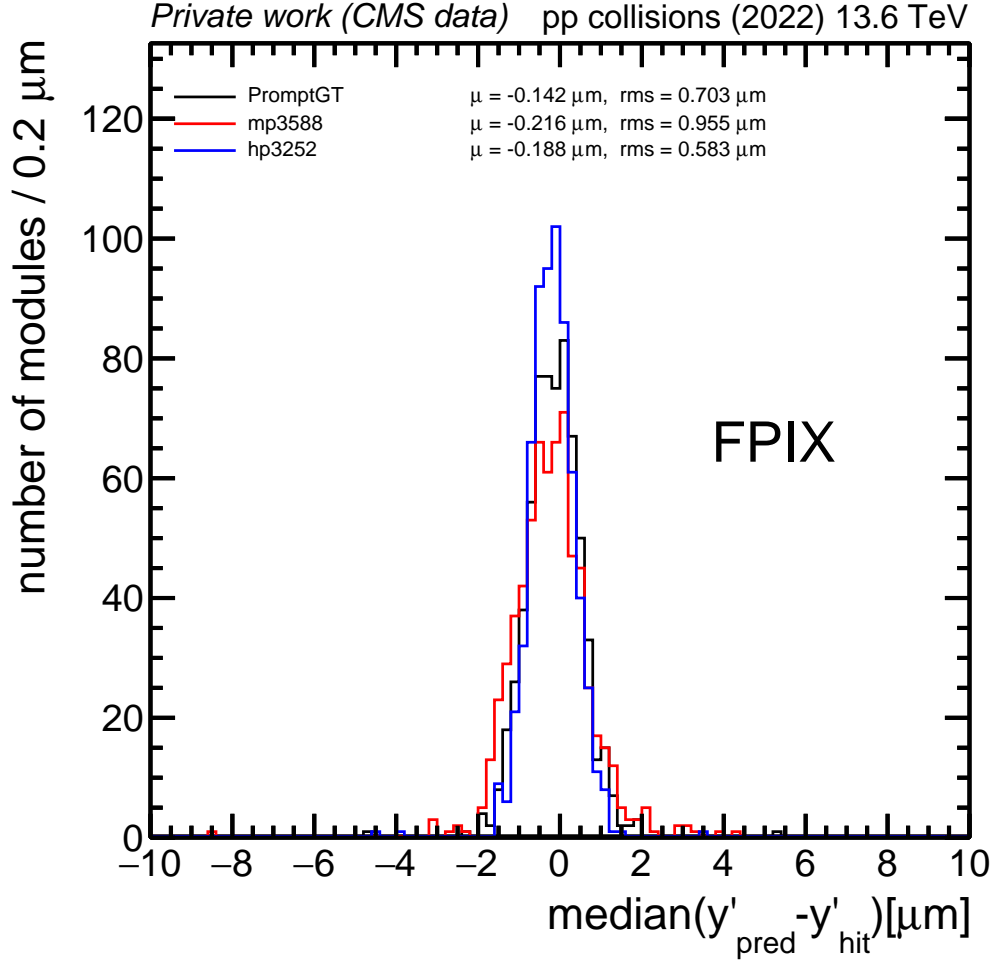


Figure 3.24: Example DMR validation plot for y' residual in the FPIX, comparing HipPy (blue) and Millepede (red) alignments using beam data, both calculated from the starting geometry of PromptGT (black).

collision tracks, cosmic ray muon tracks traverse the entire detector and connect modules in the top and bottom halves of the tracker. This breaks the cylindrical symmetry typical of collision tracks and helps to constrain several types of distortions.

Throughout Run 2, cosmic ray muon events were recorded before the start of LHC collisions during dedicated commissioning runs and in the time intervals between two

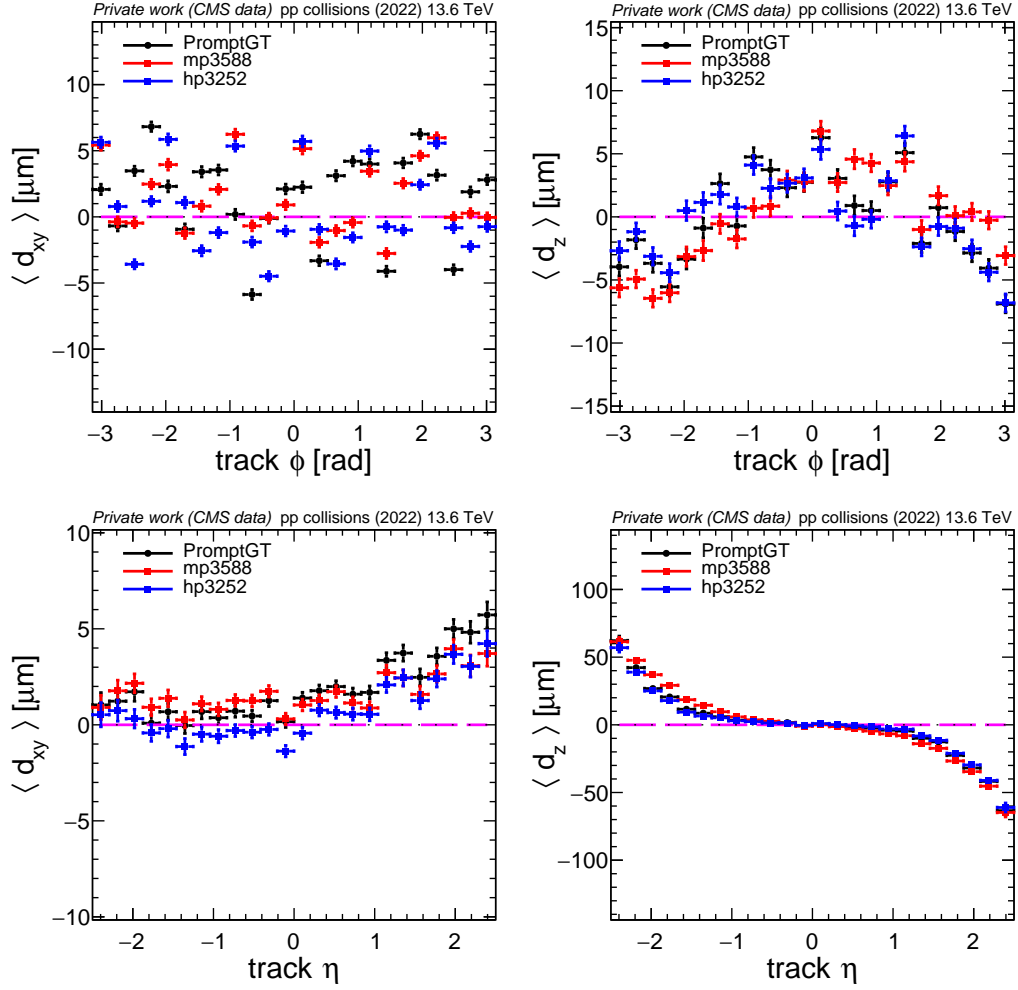


Figure 3.25: Example PV validation plot for PIX, comparing HipPy (blue) and Millepede (red) alignments using beam data, both calculated from the starting geometry of PromptGT (black).

LHC fills (interfill runs). Since 2018, cosmic events have also been recorded during collision data-taking.

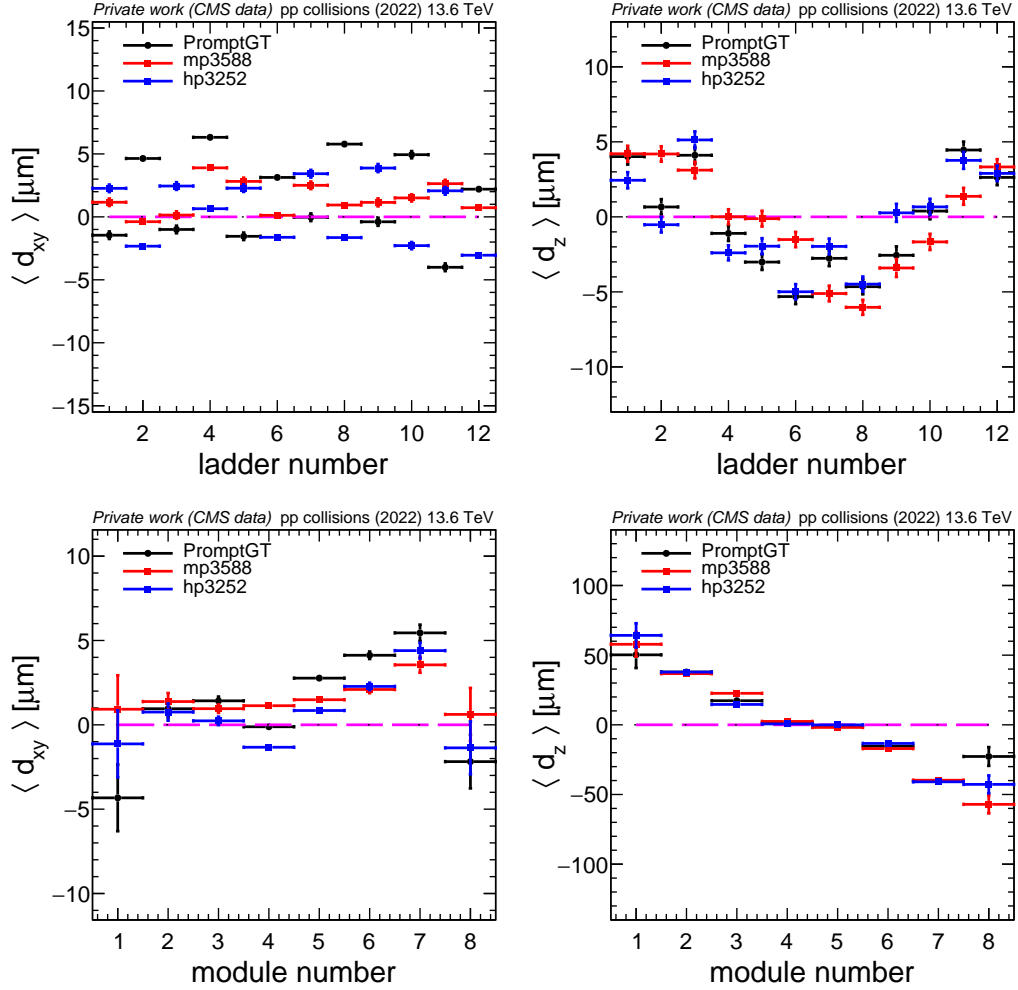


Figure 3.26: Example PV validation plot for BPIX Layer 1, comparing HipPy (blue) and Millepede (red) alignments using beam data, both calculated from the starting geometry of PromptGT (black).

3.6.3.5 Dimuon validation

In an ideally aligned tracker, the reconstructed invariant mass of a resonance $X \rightarrow \mu^+ \mu^-$ should exhibit minimal dependence on the trajectory of the muons within the detector. Consequently, the quality of a given set of alignment constants can be evaluated by examining potential biases in the reconstructed mass of a well-known

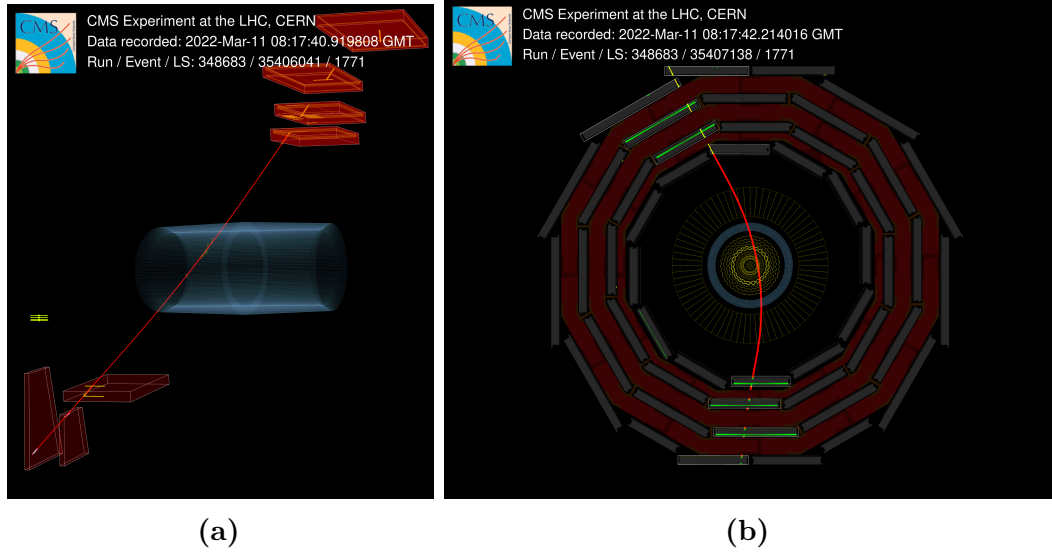


Figure 3.27: Muons generated by cosmic rays are extremely important for the commissioning and the alignment of CMS. Despite being 100 m underground, every second hundreds of muons generated on top of the atmosphere cross CMS. Presented are two views, transverse (left) and longitudinal (right), of a single cosmic ray event [49].

resonance. While any resonance could, in principle, be used for this purpose, the Z -boson decay into muons is the primary choice. The Z -boson is frequently produced with a relatively small boost, leading to a topology in which the two muons traverse opposite ends of the tracker, making alignment effects more pronounced.

Tracker misalignment may be indicated if the mean reconstructed mass deviates from the expected value of 91.2 GeV, either uniformly or as a function of pseudorapidity η and azimuthal angle ϕ . This validation is performed using LHC collision data after accumulating a sufficiently large sample of $Z \rightarrow \mu^+ \mu^-$ events. Since this method relies on well-understood physics processes in stable operating conditions, it serves as a powerful tool for assessing alignment precision in the detector.

3.6.4 Alignment and validation with HipPy

Aside from assisting with occasional Beam Spot (BS) calculations, most of my service work for the CMS experiment has involved generating tracker alignment parameters with HipPy and validation against analogous Millepede alignments. I was also responsible for documenting the HipPy alignment procedure and initiating updates to our scripts.

3.6.4.1 *Preparing HipPy for Run 3*

Alignment work began with the resurrection of the HipPy workflow and running of the first CRUZET alignments in 2021. We also began using CRAFT and beam data to align the pixel and strip detectors with HipPy. I initiated documentation of the HipPy software and alignment validation. I also performed validation of our results against the Millepede analogs for alignment of high-level structures (HLS) in the pixel detector as well as alignment of the pixel ladders (inside the BPIX), pixel panels (inside the FPIX), and strip HLS (TIB, TOB, TID, and TEC).

HipPy achieved performance comparable to that of Millepede for the alignment of pixel HLS. Its performance lagged behind Millepede in the alignment of additional modules, but this was consistent with expectations at the time due to the limited number of iterations HipPy was able to run (later improved as described in Section 3.6.4.5). Thus, HipPy continued to be utilized for alignment campaigns and acted as a cross-check for alignment parameters calculated with Millepede.

3.6.4.2 HLT alignment at 400 V

In late 2022, CMS introduced an updated operational condition in the Barrel Pixel detector (BPIX), specifically increasing the High Voltage (HV) applied to the sensors in Layer 1 from 350 V to 400 V, motivated by the cumulative radiation damage experienced by the silicon sensors.

However, this HV increase also induced subtle mechanical shifts and modified charge distributions within the pixels, impacting local hit reconstruction. Consequently, the alignment constants utilized by the HLT required updating to ensure continued accurate track reconstruction at the trigger level.

My contribution involved assisting in update of HLT alignment parameters, focusing on the verification of track residual distributions and shifts in the tracker geometry. Using beam data, I prepared a HipPy alternative to the Millepede alignment of the pixel detector at module level while keeping the strips fixed, and ran validation of the new alignment parameters which were then incorporated into the HLT global conditions database (GT), to optimize the performance of trigger selections during Run 3 Era G data-taking.

3.6.4.3 2023 commissioning and startup

I also contributed to the alignment of the tracker during 2023 commissioning and collisions startup. Using early 2023 CRUZET data, I ran a HipPy alignment of the pixel ladder and half-cylinder, as well as the strip HLS (TIB, TOB, TEC). We were able to match the performance of Millepede for alignment of the pixel detector and

validated the DMRs for the new alignment parameters, as visualized in Figure 3.28.

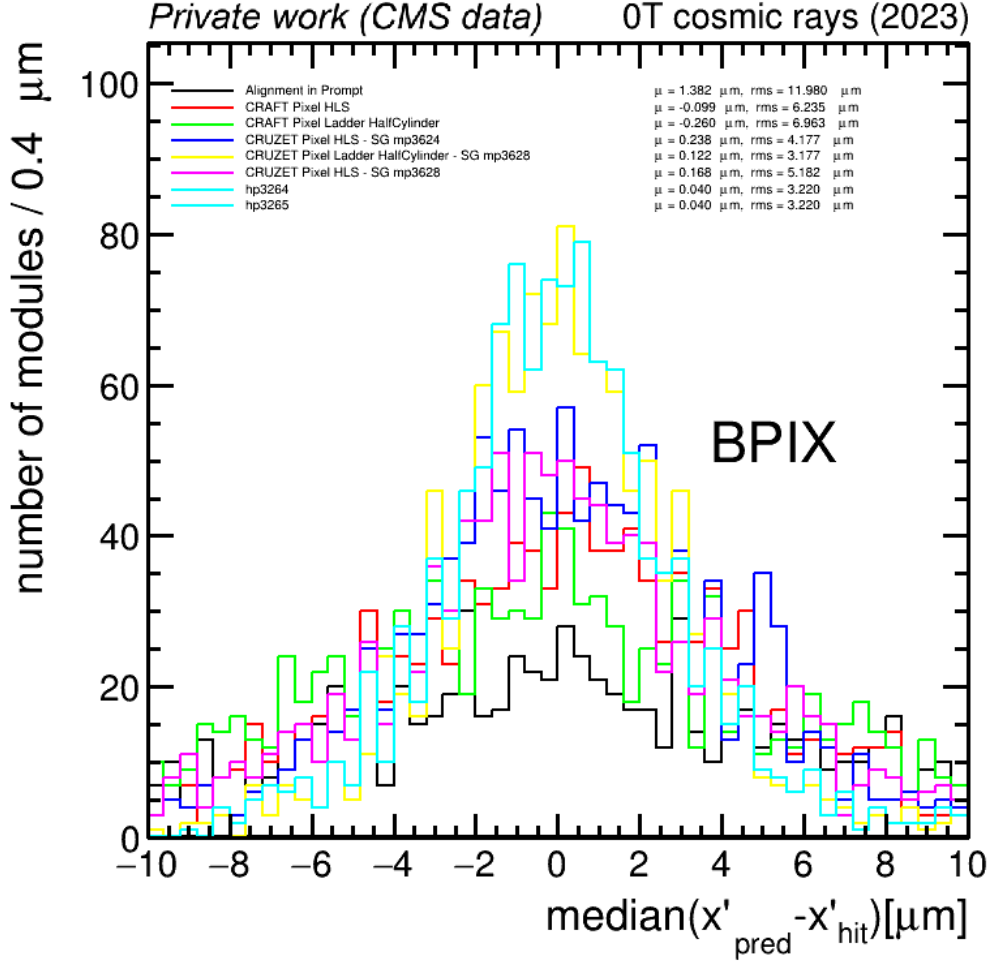


Figure 3.28: Example DMR validation plot for the BPIX in local x' coordinate, comparing HipPy (cyan) and Millepede (yellow) alignments using early 2023 CRUZET data, both calculated from the starting geometry of mp3628 (green).

3.6.4.4 Deriving Beam Spot (BS) conditions

The BS represents the average interaction region of proton-proton collisions in CMS and serves as a critical reference point for tracking and vertexing. Precise knowledge of the BS position is essential for computing impact parameters, reconstructing primary

vertices, and validating alignment parameters. During alignment updates (particularly those affecting the HLT) the most up-to-date BS conditions are needed to ensure consistency in the Full Track Validation (FTV), to evaluate the performance of track reconstruction. As part of my contributions to CMS tracker alignment, I derived and plotted BS payloads for several alignment candidates, in collaboration with the Millepede team.

3.6.4.5 HipPy updates

While working with the HipPy algorithm, we have been making upgrades and investigating how to improve its performance for tracker alignment (particularly when using multiple datasets). While the latter is ongoing, I have modernized parts of the HipPy code to improve functionality and ease of use. Through an update to the official CMS Software (CMSSW) repository, I improved the usability of the codebase by implementing bug fixes and updating the HipPy algorithm to python3. I also modified the execution scripts to log additional information about submitted alignment jobs, track ongoing processes without requiring persistent sessions, and avoid interruptions of the algorithm to aid in the calculation of more iterations and consequently achieve better alignment performance.

I had previously presented the HipPy algorithm at the Second CMS Tracker Alignment Workshop, in preparation for our contributions to tracker alignment in Run 3. After implementing my HipPy updates, I gave a follow-up presentation with a relevant and interactive example of aligning the pixel half-barrel and half-cylinder at the CMS Tracker Training Days. This alignment campaign calculated new positions for the pixel

modules while keeping strips fixed, using a “MinBias” dataset—as close as possible to an unbiased sample of pp collisions—comprising of 23 million events in this case.

Analogous alignment campaigns were undertaken by HipPy and Millepede using the same MinBias dataset and starting from the same initial starting geometry taken from the Prompt Global Tag (GT), or “PromptGT.” A GT being a named snapshot of CMS conditions (calibrations, alignments, various detector parameters), and the PromptGT being the GT used to reconstruct newly-taken data at Tier-0. (Using the PromptGT ensures a starting geometry which matches the exact conditions used to reconstruct the data being used for alignment.) Validation plots for the resulting HipPy alignment compared against the analogous Millepede alignment and their shared PromptGT starting geometry are shown in Figures [3.20](#), [3.21](#), [3.22](#), [3.23](#), [3.24](#), [3.25](#), and [3.26](#).

Chapter 4

Analysis of Higgs boson data at the LHC

As previously discussed in Section 2.2, the SM postulates the existence of a scalar field—the Higgs field—responsible for the generation of mass for fundamental particles through the mechanism of electroweak symmetry breaking. The excitation of this field is known as the Higgs boson (H) [50–56]. The properties of the Higgs boson, observed with a mass of approximately 125 GeV by the ATLAS and CMS Collaborations [2, 57, 58] at the LHC, are found to be in agreement with SM predictions [59, 60].

The mass of the Higgs boson (m_H) is a free parameter in the SM and governs all other Higgs boson properties, necessitating its measurement with the highest possible precision. For instance, the Higgs boson couplings to vector bosons depend strongly on m_H and are predicted with high precision within the SM framework. Another key parameter is the Higgs boson’s total decay width (Γ_H), which determines its lifetime. The SM predicts a width of $\Gamma_H = 4.1$ MeV, corresponding to a lifetime of 1.6×10^{-22} s [4] for $m_H = 125$ GeV. Any deviation from this prediction could signal anomalous Higgs boson couplings or decays to invisible or non-SM particles.

Using $\sqrt{s} = 7$ and 8 TeV proton-proton (pp) collision data collected during Run 1 (2011–2012), the ATLAS and CMS Collaborations jointly measured the Higgs boson mass as 125.09 ± 0.24 GeV [61], based on a combined integrated luminosity of 25 fb^{-1} per experiment. This result has since been superseded. The ATLAS Collaboration, combining the $H \rightarrow \gamma\gamma$ and $H \rightarrow 4\ell$ ($\ell = e, \mu$) decay channels using both Run 1

and Run 2 ($\sqrt{s} = 13$ TeV, 2015–2018) data, reported a measurement of $m_H = 125.11 \pm 0.11$ GeV, with a statistical uncertainty of 0.09 GeV [62]. The most recent CMS result, using the same decay channels and including Run 1 and 36 fb^{-1} of 2016 data, yields $m_H = 125.38 \pm 0.14$ GeV, with a statistical uncertainty of 0.11 GeV. Independent measurements from ATLAS and CMS based solely on $H \rightarrow 4\ell$ decay and 2016 data give $m_H = 124.94 \pm 0.18(\pm 0.17)$ GeV and $m_H = 125.26 \pm 0.21(\pm 0.19)$ GeV, respectively, where the value in parentheses is the statistical uncertainty only [61, 62].

Using only on-shell production, CMS placed an upper limit on the Higgs boson width of $\Gamma_H < 1.10$ GeV at 95% confidence level (CL), constrained by the four-lepton invariant mass resolution [63, 64]. Both ATLAS and CMS have also set more stringent constraints using the off-shell production method [6, 7, 65–69] which, as discussed in Section 2.5, exploits the population of Higgs boson events at higher energies [5, 70, 71]. This method yields recent measurements of $\Gamma_H = 3.2^{+2.4}_{-1.7}$ MeV by CMS [69], and $\Gamma_H = 4.3^{+3.3}_{-2.5}$ MeV by ATLAS [72]. Additionally, from limits on the Higgs boson’s flight distance in the detector, CMS set a lower bound of $\Gamma_H > 3.5 \times 10^{-9}$ MeV at 95% CL [66].

This chapter presents an updated CMS measurement of the Higgs boson width using off-shell production in the $H^* \rightarrow 4\ell$ decay channel. Additionally, results from an extension of the off-shell analysis are summarized, as well as a Higgs boson mass measurement conducted by collaborators using the shared on-shell data sample. The analysis includes 138 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV, collected from Run 2 (2016–2018), combined with Run 1 data.

4.1 Modeling of off-shell production

This analysis takes advantage of new tools for off-shell simulation that are now available through JHUGen [9, 73–77]. With the release of JHUGen v7.2.4, options were added to generate off-shell EW events, including hadronic VH production, the continuum background, and up to two scalar resonances. By release v7.3.0, the generation of interference in 4ℓ final states in off-shell EW production was improved and it became possible to simulate off-shell EW Higgs boson production with the inclusion of the anomalous couplings described in Section 2.6.1. In conjunction with the previously available simulations for off-shell gluon fusion production, illustrated in Figure 4.1, this functionality enables future off-shell analysis of these anomalous couplings with all primary Higgs boson production modes simulated through JHUGen. When anomalous couplings are incorporated into future analyses, reweighting techniques also enable the creation of distributions for hypotheses with mixed anomalous contributions to the HVV amplitude, without the need to uniquely generate such samples.

Notably, as is done in this analysis, one can reweight various MC samples to the same SM-like hypothesis, providing additional statistics to make a more precise measurement of the SM-like Higgs boson. The matrix element likelihood approach (MELA), which is discussed further in Section 4.2.2, is used to reweight all simulated off-shell events to the SM hypothesis through modified MCFM matrix elements [57, 78]. This eliminates the need for out-of-the-box simulation of additional SM events by utilizing statistics from pre-existing BSM simulations instead. Furthermore, this analysis improves on past papers [7] by including the entire Run 2 dataset (spanning years 2016, 2017, and 2018) and using Ultra-Legacy (UL) processing, as explained in Section 4.2.

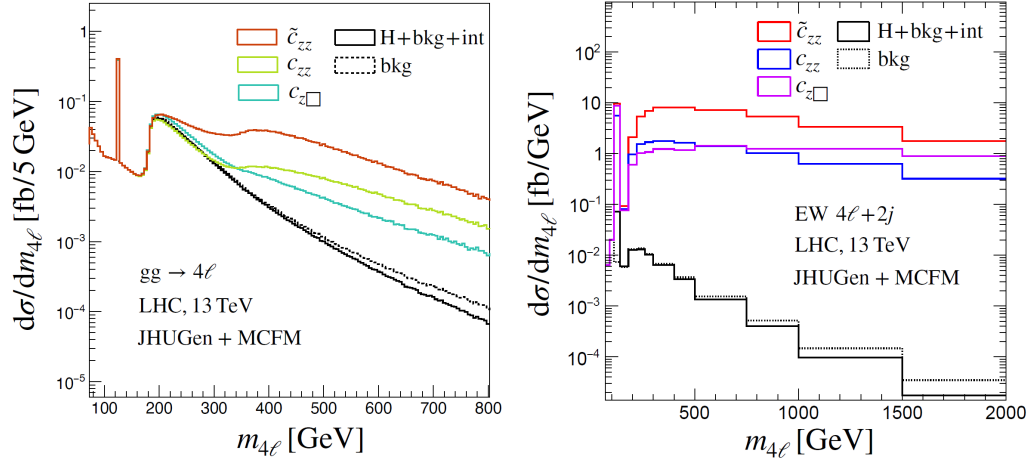


Figure 4.1: The four-lepton $m_{4\ell}$ invariant mass distributions for gluon fusion (left) and electroweak production in association with two jets (right) at the LHC with a 13 TeV pp collision energy. The total SM production (“H+bkg+int”) and background-only (“bkg”) components are shown in black. Three operators $c_{z\Box}$ (magenta), c_{zz} (blue), and \tilde{c}_{zz} (red) are shown in color, as analogs of the anomalous couplings described in Section 2.6.1, and they are introduced in place of the SM interaction with their strength constrained to reproduce the SM cross section of the on-shell Higgs boson signal production [8].

There already exist on-shell samples (centered around the 125 GeV Higgs boson peak) for the dominant production modes generated with POWHEG v2 and JHUGen v7.0.11, which have been validated and utilized in CMS analyses [7, 79, 80]. These samples are inherited for the on-shell Higgs boson cross-section, and doubly used as a reliable point of comparison for newly generated MC events. To provide additional statistics for electroweak production, on-shell samples are generated with JHUGen+JHUGenMELA and off-shell samples of the combined BSI (background+signal+interference) events are generated via Phantom. These samples are valuable in the validation of the recently produced off-shell samples for gluon fusion and electroweak production, which were generated using JHUGen v7.3.0 interfaced to MCFM v7.0.1 and MCFM v7.0.6.

4.1.1 K factors

MC simulations are improved by applying K factors, which correct leading-order (LO) calculations with higher-order effects (taking them to NLO, NNLO, N³LO, etc.). A K factor is defined as the ratio of a cross section at higher order to the LO prediction, $K = \sigma_{\text{N}^n\text{LO}}/\sigma_{\text{LO}}$. For gluon-fusion Higgs boson production (ggH), QCD effects dominate the K factors. Electroweak (EW) corrections are generally smaller but their relative importance grows at high energies due to off-shell vector bosons. Backgrounds such as $q\bar{q} \rightarrow ZZ$ and $gg \rightarrow ZZ$ backgrounds rely primarily on QCD and loop-induced corrections. In practice, all these K factors are applied as multiplicative weights to MC samples, so that simulated yields and shapes match the best available theoretical predictions.

It is worth noting that K factors are not constant and vary with kinematics. In the off-shell Higgs boson analysis, their dependence on the four-lepton invariant mass $m_{4\ell}$ and on jet multiplicity is especially important. Using differential K factors ensures that both the overall rate and the distribution shapes of signal, background, and interference processes are modeled more accurately across phase space. Corrections relevant to jet modeling as a function of invariant mass are discussed in Section 4.1.2.

4.1.1.1 ggH processes

The cross section for the SM Higgs boson produced via gluon fusion in the on-shell region is reported in Yellow Report 4 (YR4) [4] at N³LO as 48.58 pb. Because the MC cross section already includes the LO to NNLO K factor, matching the YR4 value requires the calculation of an additional NNLO-to-N³LO K factor, which is

independent of the Higgs boson mass, to rescale the cross section of the on-shell gluon fusion signal.

The N³LO cross section, reported in YR4, is corrected for the generator cut efficiencies and the Higgs boson branching fraction, and calculated for the three lepton flavor combinations separately. The K factor is expected to be independent of the lepton flavor, therefore the result is averaged in the end. The calculation is performed at 125 GeV mass point, for which the N³LO cross section is reported, but the correction is assumed to be independent of the mass. The calculation and results are summarized in Table 4.1. The ratio between the YR4 and MC cross sections is treated as an additional NNLO to N³LO K factor of 1.154. This K factor is applied to all the gluon fusion samples across all mass values (pure signal, interference, and background).

| | YR4 σ [pb] | BR | GEN cut efficiency | σ [pb] |
|---|-------------------|-----------------------|--------------------|---------------|
| ggH \rightarrow 2e2 μ | 48.58 | 5.897×10^{-5} | 65.285% | 0.001870 |
| ggH \rightarrow 4e | 48.58 | 3.254×10^{-5} | 64.413% | 0.001018 |
| ggH \rightarrow 4 μ | 48.58 | 3.254×10^{-5} | 64.413% | 0.001018 |
| MC ggH \rightarrow 2e2 μ | | | | 0.001628 |
| MC ggH \rightarrow 4e | | | | 0.000888 |
| MC ggH \rightarrow 4 μ | | | | 0.000888 |
| NNLO QCD to N ³ LO QCD and NLO EW K factor | | | | 1.154 |

Table 4.1: The calculation for the NNLO to N³LO K factor for the gluon fusion MC samples. The factor calculated at 125 GeV is applied to all the gluon fusion processes in the full off-shell region.

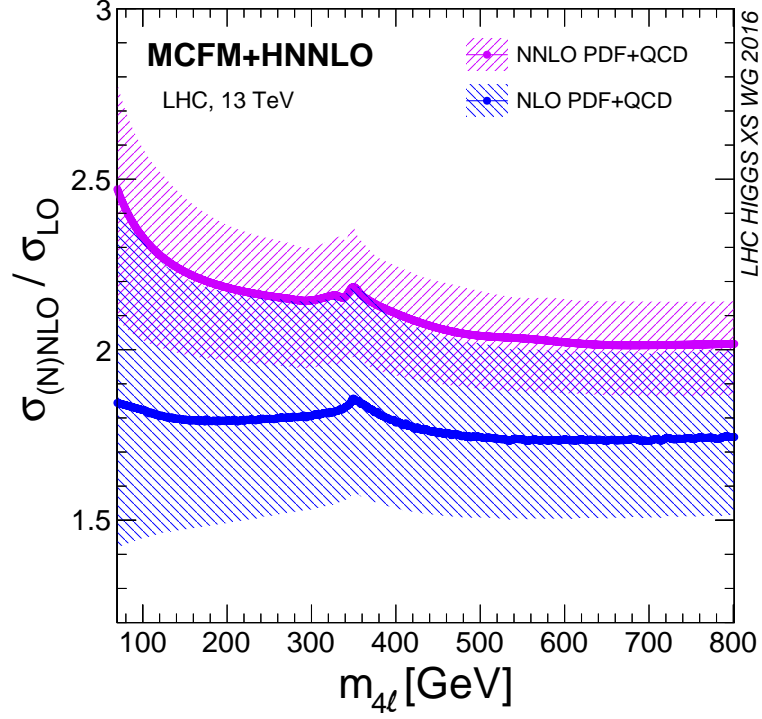


Figure 4.2: LO-to-NNLO and LO-to-NLO K factors plotted as functions of invariant mass for ggH production signal at $\sqrt{s} = 13$ TeV, cited from YR4 [4].

The gluon fusion simulation is performed at LO and a mass dependent LO to NNLO K factor is applied, as illustrated in Figure 4.2. All off-shell production simulations are conducted at LO in QCD, with higher-order corrections included through a K factor. In the gluon fusion process, the factorization and renormalization scales are set to run dynamically by equating them to $m_{4\ell}/2$. To incorporate higher-order QCD corrections, signal cross section calculations are performed at leading order (LO), next-to-leading order (NLO), and next-to-next-to-leading order (NNLO) using the MCFM and HNNLO 2 programs [81–83] under the narrow-width approximation, covering a broad range of masses spanning both the on-shell and off-shell regions.

The ratios of NNLO to LO cross sections, referred to as LO-to-NNLO K factors, are used to reweight [4] the $m_{4\ell}$ distributions obtained from LO QCD simulations produced with MCFM and JHUGen. A constant K factor of 1.154 (as seen in Table 4.1) is applied across the entire $m_{4\ell}$ spectrum to normalize the Higgs boson production cross section via gluon fusion to the next-to-next-to-next-to-leading order (N³LO) prediction at $m_{4\ell} \approx 125$ GeV [4]. The $m_{4\ell}$ distributions from the POWHEG+JHUGen simulation of the $gg \rightarrow H$ process are reweighted using NLO-to-NNLO K factors.

While the LO-to-NNLO K factor is directly applicable to the signal, it serves only as an approximation for the $gg \rightarrow 4\ell$ background and its interference with the signal. An approximate NLO calculation [84–87] is available for both the background and the interference. The resulting LO-to-NLO K factors for these components agree with the signal K factor within approximately 10% in the $m_{4\ell} > 220$ GeV range relevant to this analysis. Consequently, the same LO-to-NNLO K factor and N³LO normalization, computed for the signal, are applied to the background and interference, along with their associated uncertainties. An additional 10% uncertainty is assigned to the background component, and a 5% uncertainty—corresponding to the square root of the variation—is applied to the interference.

4.1.1.2 *EW processes*

Using the JHUGen+MCFM programs, we simulate EW Higgs boson production via VBF and VH processes, with inclusive decays of the Higgs boson to all lepton flavors (including τ leptons). In the VH process, the vector boson decays are restricted to hadronic modes, as the EW simulation is designed to model the $4\ell + 2\text{jets}$ final state.

These LO samples are generated with generator-level lepton and jet cuts. The lepton cuts are identical to those applied in the gluon fusion simulation described earlier and do not affect the final event selection after reconstruction. The jet cuts, however, reduce the number of selected events by approximately 15%, impacting only topologies that are not explicitly VBF- or VH-tagged.

YR4 provides cross-section predictions for individual EW processes (VBF, ZH, WH) in the on-shell region, calculated up to NNLO in QCD and NLO in EW corrections [4]. As in the gluon fusion case, we compare the on-shell cross section of the simulated EW samples to the YR4 predictions and extract a universal correction factor (K factor) to be applied consistently across all EW simulated samples, including pure signal, interference, and background contributions.

For this comparison, we apply the generator-level (GEN) selection efficiency to each YR4 cross-section value. Additionally, for WH and ZH processes, we multiply by the hadronic branching fraction of the vector boson [88]. Finally, the $H \rightarrow 4\ell$ branching fraction (including decays to τ leptons) is applied to all processes. The resulting three contributions are summed and compared to the simulated pure signal cross section at the Higgs boson mass peak, as summarized in Table 4.2.

We adopt the uncertainties reported in YR4 for the NNLO cross section and estimate an inclusive uncertainty for each source contributing to the K factor. The resulting LO-to-NNLO QCD and NLO EW correction factor is:

$$K = \mathbf{1.039} \quad \left\{ {}^{+1.33\%}_{-0.8\%} (\text{QCD scale}) \pm 1.89\% (\text{PDF}) \pm 0.65\% (\alpha_{\text{strong}}) \right\}.$$

| | YR4 [pb] | GEN cut efficiency | V hadronic | $H \rightarrow 4\ell$ | σ [pb] |
|---|----------|--------------------|------------|-----------------------|---------------|
| VBF | 3.782 | 62.073% | | 0.0002745 | 0.000644 |
| WH | 1.373 | 52.611% | 67.41% | 0.0002745 | 0.000134 |
| ZH | 0.8839 | 55.867% | 69.911% | 0.0002745 | 0.000095 |
| Total | | | | | 0.000873 |
| MC | | | | | 0.000840 |
| LO QCD+EW to NNLO QCD and NLO EW K factor | | | | | 1.039 |

Table 4.2: The calculation for the LO to NNLO K factor for the EW off-shell samples at the Higgs boson mass peak (125 GeV). The calculated factor is applied to all the off-shell EW MC samples.

The $q\bar{q} \rightarrow 4\ell$ background is simulated at NLO in QCD and LO in electroweak (EW) theory using POWHEG. The fully differential cross section for this process has been computed at NNLO in QCD [89], and a differential NLO-to-NNLO K factor as a function of $m_{4\ell}$ is applied to the POWHEG sample. This K factor is approximately 1.1 at $m_{4\ell} = 125$ GeV and varies between 1.0 and 1.2 in the region $m_{4\ell} < 500$ GeV. The uncertainty from missing EW corrections in the low-mass region ($m_{4\ell} < 2m_Z$) is small compared to QCD-related uncertainties.

To account for EW corrections, an LO-to-NLO K factor [90] is applied to events containing two on-shell Z bosons. This K factor decreases with increasing $m_{4\ell}$, from approximately 1.0 near 125 GeV to around 0.9 in the high-mass region ($m_{4\ell} > 500$ GeV). The associated uncertainty is the dominant systematic for the off-shell

width measurement and is included as a function of $m_{4\ell}$. The differential QCD and EW K factors are illustrated for in Figure 4.3.

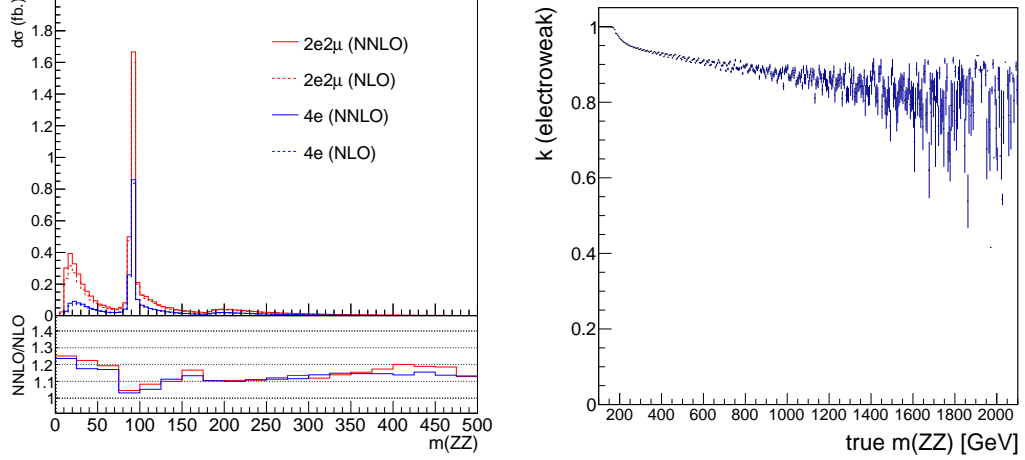


Figure 4.3: Illustrations of the NLO-to-NNLO QCD K factor for the $q\bar{q} \rightarrow ZZ$ background (left) and LO-to-NLO EW K factor as a function of invariant mass (right). Both plots take $\sqrt{s} = 13$ TeV [89, 91, 92].

4.1.2 Corrections for jet modeling

It is important to note, as on-shell events are compared between all these samples using the cross-section and branching ratio values from YR4 [4], that the off-shell samples share common generator inputs which are not applied to the on-shell samples [4]. These cuts and their associated efficiencies are documented in Table 4.3.

Generator-level cuts are applied on leptons during the generation of the samples. The efficiency of these cuts is evaluated to ensure proper normalization of the cross sections through independent Monte Carlo (MC) studies, where various processes are generated both with and without the aforementioned lepton selection.

Chapter 4. Analysis of Higgs boson data at the LHC

| Generator | Sample / Analysis | Leptons | Jets |
|---------------|-------------------------------|---|--|
| JHUGen+MCFM | EW off-shell | $pT_{l_{1,2,3,4}} > 3\text{GeV}, \eta_{l_{1,2,3,4}} < 2.7$ | $pT_{j_{1,2}} > 15\text{ GeV}, m_{jj} > 30\text{ GeV}, \eta_{j_{1,2}} < 6.5, \Delta R > 0.3, \Delta\eta_{j_{1,2}} > 0$ |
| | ggH off-shell | $pT_{l_{1,2,3,4}} > 3\text{GeV}, \eta_{l_{1,2,3,4}} < 2.7$ | $pT_{j_{1,2}} > 15\text{ GeV}, \Delta R > 0.4$ |
| Phantom | EW off-shell (BSI, $4e$) | $pT_{l_{1,2,3,4}} > 3\text{ GeV}, \eta_{l_{1,2,3,4}} < 2.7$ | $m_{jj} > 30\text{ GeV}, \eta_{j_{1,2}} < 6.5$ |
| | EW off-shell (BSI, 4μ) | $pT_{l_{1,2,3,4}} > 3\text{ GeV}, \eta_{l_{1,2,3,4}} < 2.7$ | $m_{jj} > 30\text{ GeV}, \eta_{j_{1,2}} < 6.5$ |
| | EW off-shell (BSI, $2e2\mu$) | $pT_{l_{1,2,3,4}} > 3\text{ GeV}, \eta_{l_{1,2,3,4}} < 2.7$ | $m_{jj} > 30\text{ GeV}, \eta_{j_{1,2}} < 6.5$ |
| POWHEG+JHUGen | VBF on-shell | No specified cuts | No specified cuts |
| | W^-H on-shell | No specified cuts | No specified cuts |
| | W^+H on-shell | No specified cuts | No specified cuts |
| | ZH on-shell | No specified cuts | No specified cuts |
| | ggH on-shell | No specified cuts | No specified cuts |
| JHUGen+JHUGen | VBF on-shell | No specified cuts | No specified cuts |
| | WH on-shell | No specified cuts | No specified cuts |
| | ZH on-shell | No specified cuts | No specified cuts |

Table 4.3: Generator-level cuts on jets and leptons in all utilized samples. Note that Phantom provides EW simulation but MCFM is used for both EW and ggH events.

In the case of gluons fusion off-shell samples, the following lepton selection requirements are applied: $p_T(\ell_{1,2,3,4}) > 3\text{ GeV}, |\eta(\ell_{1,2,3,4})| < 2.7$. It has been validated that no leptons which would otherwise be reconstructed in the analysis are lost due to these cuts, since these selection cuts remove leptons that fall outside the detector acceptance or reconstruction capabilities.

| m_H | VBF | WH | ZH | EW (had) | EW (full) |
|---------|--------|--------|--------|----------|-----------|
| 125 GeV | 71.35% | 61.64% | 63.09% | 69.15% | 68.45% |
| 200 GeV | 78.90% | 71.45% | 72.18% | 78.13% | 77.83% |
| 300 GeV | 81.70% | 76.78% | 77.39% | 81.46% | 81.36% |
| 400 GeV | 84.57% | 82.06% | 82.46% | 84.50% | 84.47% |
| 500 GeV | 86.95% | 85.77% | 85.91% | 86.93% | 86.92% |
| 600 GeV | 88.72% | 88.19% | 88.28% | 88.71% | 88.71% |

Table 4.4: Efficiencies of generator-level requirements on leptons ($p_T(\ell_{1,2,3,4}) > 3\text{ GeV}, |\eta(\ell_{1,2,3,4})| < 2.7$) in the off-shell EW samples as a function of $m_{4\ell}$. The EW (had) includes VBF and hadronic VH production, while EW (full) also includes leptonic VH production.

In the case of EW off-shell samples, we evaluate efficiency of the generator-level cuts separately for the VBF, WH, and ZH samples, and then combine them according to their relative cross sections, used as weight. First we evaluate effect of lepton selection requirements are applied: $p_T(\ell_{1,2,3,4}) > 3$ GeV, $|\eta(\ell_{1,2,3,4})| < 2.7$. The effect of these cuts on EW samples, generated with JHUGen signal only, is shown in Table 4.4. The effect of jet cuts on EW samples, generated with JHUGen signal only, is shown in Table 4.5.

| | efficiency of jet cuts | | | | | jet + lepton cuts | |
|---------|------------------------|--------|--------|----------|-----------|-------------------|-----------|
| m_H | VBF | WH | ZH | EW (had) | EW (full) | EW (had) | EW (full) |
| 125 GeV | 86.92% | 85.49% | 88.38% | 86.84% | 86.81% | 60.05% | 59.42% |
| 200 GeV | 87.09% | 88.59% | 90.34% | 87.33% | 87.42% | 68.23% | 68.03% |
| 300 GeV | 87.28% | 91.50% | 92.48% | 87.51% | 87.61% | 71.29% | 71.28% |
| 400 GeV | 87.20% | 93.38% | 94.00% | 87.39% | 87.47% | 73.84% | 73.88% |
| 500 GeV | 86.84% | 94.71% | 94.98% | 87.00% | 87.07% | 75.62% | 75.67% |
| 600 GeV | 86.56% | 95.61% | 95.83% | 86.69% | 86.92% | 76.90% | 76.95% |

Table 4.5: Efficiencies of generator-level requirements on jets in the off-shell EW samples as a function of $m_{4\ell}$. The EW (had) includes VBF and hadronic VH production, while EW (full) also includes leptonic VH production. These efficiencies are calculated after generator-level requirements had been applied on leptons, as indicated in Table 4.4. The last two columns incorporate efficiency of lepton cuts together with jet cuts.

4.2 Building a $H^{(*)} \rightarrow ZZ^{(*)} \rightarrow 4\ell$ analysis

For measurements of the Higgs boson width and production signal strength, this analysis uses UL processed CMS data and simulated samples, which provide the most up-to-date and homogeneous reconstruction of the full Run 2 dataset (2016–2018). The UL campaign incorporates final detector calibrations, alignments, and

reconstruction algorithms, thereby improving the overall detector performance and reducing systematic uncertainties compared to earlier re-reconstruction campaigns (ReReco). While ReReco samples were produced year by year, each using the best available knowledge of the detector at that time, they do not guarantee uniform conditions across the full Run 2 dataset. UL samples on the other hand, were generated in a single comprehensive reprocessing after Run 2 data-taking concluded, ensuring a consistent treatment across all years.

4.2.1 Reconstructing and selecting events

Event reconstruction is based on the particle-flow (PF) algorithm [93], which endeavors to reconstruct and identify each individual particle in an event by combining information from the various subdetectors of the CMS experiment. PF candidates are classified as photons, electrons, or muons, and they are then used to build higher-level objects, such as jets.

Photon energies are determined directly from ECAL measurements, and are not expected to present a signal in the HCAL or tracker. Electrons are identified from a multivariate discriminant involving their momentum at the PV as measured by the tracker, a geometrically corresponding ECAL cluster of energy deposition, and the integrated energy of all bremsstrahlung (radiated) photons determined to be spatially consistent with emission from the electron’s trajectory [79]. Muons, like electrons, also start in the tracker, but have trajectories and momenta that can be extrapolated to the muon system while depositing little to no energy in the ECAL and HCAL. Their energies are inferred from the curvature of their tracks in the magnetic field. After the

isolated muons, electrons, and photons are classified, the remaining particles are used to identify hadrons. For charged hadrons, energies are determined by combining their momenta measured in the tracker with the corresponding ECAL and HCAL energy deposits. The energies of neutral hadrons are obtained from their corrected ECAL and HCAL energy deposits.

Electrons are reconstructed within the geometrical acceptance defined by $|\eta| < 2.5$ and transverse momentum $p_T > 7 \text{ GeV}$ [94]. While muons with transverse momentum $p_T > 5 \text{ GeV}$ are reconstructed within the geometrical acceptance of $|\eta| < 2.4$. For $H \rightarrow 4\ell$ candidates, track parameters relevant to momenta are reconstructed using information from the tracker [95]. Selection of leptons from among the reconstructed track candidates is subject to stringent quality criteria applied to tracks in both the muon chambers and the silicon tracker, complemented by requirements of minimal energy deposition in the calorimeter systems, ensuring high purity in identification.

Additionally, to distinguish prompt lepton products of high- p_T Z boson decays from those originating in hadronic jets—such as leptons produced in electroweak decays of hadrons—an isolation requirement is applied. For muons, an isolation variable is computed to quantify the level of surrounding activity and suppress backgrounds from non-prompt sources [79]. The isolation variable is defined relative to the muon’s transverse momentum p_T as

$$\mathcal{I}^\mu \equiv \left(\sum p_T^{\text{charged}} + \max \left[0, \sum p_T^{\text{neutral}} + \sum p_T^\gamma - p_T^{\text{PU}} \right] \right) / p_T, \quad (4.1)$$

where p_T^{charged} , p_T^{neutral} , and p_T^γ are summed over the charged and neutral hadrons and

photons within a cone radius of $\Delta R \equiv \sqrt{(\eta^i - \eta^j)^2 + (\phi^i - \phi^j)^2} = 0.3$ around the lepton's direction at the interaction vertex. Here, ϕ denotes the azimuthal angle and indices i and j refer to the hadron or photon and muon, respectively. This isolation variable exhibits particular sensitivity to energy deposits from pileup interactions. Therefore, a contribution p_T^{PU} from pileup is subtracted, as seen in Equation (4.1) [96]. Finally, each muon in a selected event must satisfy $\mathcal{I}^\mu < 0.35$.

Instead of applying a separate isolation requirement—as in the case for muons—the electron multivariate discriminant incorporates the isolation sums ($\sum p_T^{\text{charged}}$, $\sum p_T^{\text{neutral}}$, and $\sum p_T^\gamma$) directly [79]. The inclusion of these isolation components offers better performance for electrons compared to a simple threshold on the relative isolation variable. The XGBOOST (eXtreme Gradient Boosting) library is used on simulated events to train and optimize the multivariate discriminant employed for electron identification and isolation. Events are separated into six regions, defined by two transverse momentum ranges ($7 < p_T < 10 \text{ GeV}$ and $p_T > 10 \text{ GeV}$) and three pseudorapidity intervals: the central barrel ($|\eta_e| < 0.8$), outer barrel ($0.8 < |\eta_e| < 1.479$), and endcaps ($1.479 < |\eta_e| < 2.5$). Separate discriminant training is performed for each of the three data-taking periods in Run 2, and the selection requirements are optimized such that the signal efficiency is uniform across all three periods.

The effect of the final-state radiation (FSR) from leptons are also recovered [93]. (Bremsstrahlung photons already associated to electrons in the reconstruction step are not considered.) Isolated photons satisfying $p_T^\gamma > 2 \text{ GeV}$, $|\eta^\gamma| < 2.4$, and $\mathcal{I}^\gamma < 1.8$ are associated with the nearest selected lepton (muon or electron) in the event. Photons must also satisfy the criteria $\Delta R(\gamma, \ell)/(p_T^\gamma)^2 < 0.012 \text{ GeV}^{-2}$ and $\Delta R(\gamma, \ell) < 0.5$. When

multiple photon candidates fulfill these conditions, the one minimizing $\Delta R(\gamma, \ell)/(p_T^\gamma)^2$ with respect to the given lepton is retained. FSR photons are excluded from the computation of the relative isolation parameter.

A “tag-and-probe” technique [97], based on samples of Z boson events in both data and simulation, is employed to measure the reconstruction and selection efficiencies for prompt electrons and muons across bins of transverse momentum (p_T) and pseudorapidity (η). The differences between the efficiencies measured in data and those obtained from simulation are used to derive correction factors, which are applied to rescale the event yields in the simulated samples [79]. These simulated events are also utilized to calibrate the momentum scale and resolution of prompt leptons, again in bins of p_T and η [98, 99].

For this analysis, selection of Higgs boson candidates requires four prompt and isolated leptons following the aforementioned criteria. One lepton must satisfy $p_T > 20$ GeV and at least one additional lepton must have $p_T > 10$ GeV. Then, Z boson candidates are constructed from e^+e^- or $\mu^+\mu^-$ pairs with invariant masses in the range 12–120 GeV. These dilepton pairs are combined to form a Higgs boson candidate. We choose Z_1 to label the pair with invariant mass closest to the nominal Z boson mass, and call the remaining pair Z_2 .

Four possible combinations are considered and treated independently: 4μ , $4e$, $2e2\mu$, and $2\mu2e$, where the mixed-flavor final states are distinguished based on the flavor composition of Z_1 . These four combinations exhibit different four-lepton mass resolutions (predominantly determined by the flavor composition of Z_1) and contain different

amounts of reducible background (largely dependent on the flavor composition of Z_2). Since none of these flavor-dependent characteristics impact the off-shell Higgs boson analysis, all flavor channels are combined to make our measurement [100].

Signal candidates must satisfy $m_{4\ell} > 70$ GeV. When multiple Higgs boson candidates can be formed in an event, the one with the highest value of the kinematic discriminant $\mathcal{D}_{\text{bkg}}^{\text{kin}}$ (defined in Section 4.2.3) is chosen. However, if these Higgs boson candidates are comprised of the same four leptons, the candidate whose Z_1 invariant mass lies closest to the nominal Z boson mass is selected.

In this off-shell analysis, events undergo further categorization based on jets associated with the Higgs boson candidate [100]. Jets are clustered using the anti- k_t jet finding algorithm [101, 102] with a distance parameter of 0.4. The jet momentum is determined as the vector sum of all constituent particle momenta. Jets must satisfy $p_T > 30$ GeV and $|\eta| < 4.7$ and must be spatially separated from all selected lepton candidates and any selected FSR photons by requiring $\Delta R(\ell/\gamma, \text{jet}) > 0.4$. Jets originating from b quark hadronization are identified using the DeepCSV algorithm [103], which integrates information on impact parameter significance, secondary vertex characteristics, and jet kinematic variables. These b-tagged jets are then used as part of our criteria for categorizing events, as detailed in Section 4.2.3.

4.2.2 Matrix Element Likelihood Approach

With up to 13 observables (Ω) describing the Higgs boson kinematic distributions in the $2 \rightarrow 6$ process, as visualized in Figure 4.4, it can be a challenging task to perform an optimal analysis. MELA lets us calculate the matrix elements (ME) for each Higgs

boson interaction as a function of its couplings and interacting particles. We can utilize these ME in ratios to change our hypotheses and reweight our Monte-Carlo samples, or treat them as probabilities and construct discriminants from the ME to use as observables (as discussed further in Section 4.2.3).

The MELA approach is designed to reduce the number of observables to the minimum while retaining all essential information. This is a very powerful tool which ensures that our analysis observables are still based in first principles, while affording us the flexibility to reweight distributions to new hypotheses or combine differently simulated statistics for greater precision in a single measurement.

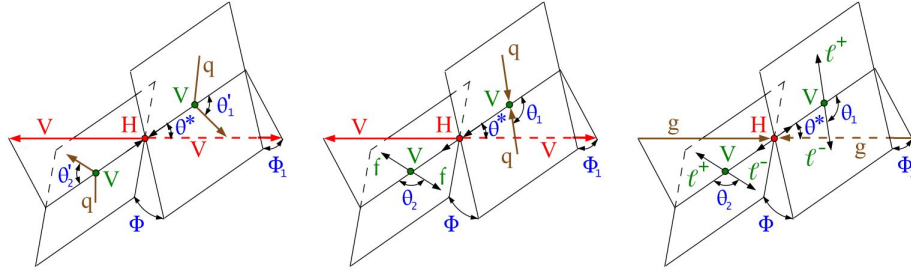


Figure 4.4: Diagrams with relevant kinematic observables labeled in the VBF (left), VH (center), and ggH (right) production modes of the Higgs boson.

4.2.3 Defining categories and observables

In order to perform a dedicated study of the particular kinematic topology of HVV, visualized in Figure 4.4, events are split into several mutually exclusive categories based on the presence of other particles produced in association with the Higgs boson candidate [79]. Two types of discriminants are defined for either the production or decay process, which can be combined to describe the full $2 \rightarrow 6$ process [7, 64, 79, 104]. We use the values of these kinematic discriminants and other selection requirements

to perform the categorization. They are calculated using the MELA approach while employing the matrix elements at leading order (LO) in quantum chromodynamics (QCD):

$$\mathcal{D}_{\text{alt}}(\boldsymbol{\Omega}) = \frac{\mathcal{P}_{\text{sig}}(\boldsymbol{\Omega})}{\mathcal{P}_{\text{sig}}(\boldsymbol{\Omega}) + \mathcal{P}_{\text{alt}}(\boldsymbol{\Omega})} \quad (4.2)$$

$$\mathcal{D}_{\text{int}}(\boldsymbol{\Omega}) = \frac{\mathcal{P}_{\text{int}}(\boldsymbol{\Omega})}{2 \sqrt{\mathcal{P}_{\text{sig}}(\boldsymbol{\Omega}) \mathcal{P}_{\text{alt}}(\boldsymbol{\Omega})}}, \quad (4.3)$$

where the probability of a certain process \mathcal{P} is calculated using the full kinematics characterized by $\boldsymbol{\Omega}$ for the hypotheses denoted as “sig” for a signal model and “alt” for an alternative model, which could be an alternative Higgs boson production mechanism (used to categorize events), background (used to isolate signal), or an alternative Higgs boson coupling model (used to measure coupling parameters). The “int” label represents the interference between the two model contributions. The probabilities \mathcal{P} are calculated from the matrix elements provided by the MELA package and are normalized to give the same integrated cross sections in the relevant phase space of each process.

A set of discriminants $\mathcal{D}_{2\text{jet}}$ is constructed, following Equation 4.2, where \mathcal{P}_{sig} corresponds to the signal probability for the VBF (WH or ZH) production hypothesis in the VBF-tagged (VH-tagged) category, and \mathcal{P}_{alt} corresponds to that of Higgs boson production in association with two jets via gluon fusion. These discriminants use full kinematic information from the Higgs boson with its associated jets and are labeled to indicate a specific topology (2jet) and production mechanism (VBF, WH, ZH), which is discriminated against the dominant gluon fusion process: $\mathcal{D}_{2\text{jet}}^{VBF}$, $\mathcal{D}_{2\text{jet}}^{ZH}$, and $\mathcal{D}_{2\text{jet}}^{WH}$. When more than two jets pass the selection criteria, the two jets with the highest p_T

are chosen for the matrix element calculations. Thus, the $\mathcal{D}_{2\text{jet}}$ discriminants separate the target production mode of each category from gluon fusion production, in all cases using only the kinematics of the Higgs boson and two associated jets [100].

Sequential selection criteria are used to define the event categories:

- The VBF-2jet category requires exactly four leptons. In addition, events must contain either two or three jets with at most one identified as originating from a b quark (b-tagged jet) [79], or at least four jets with no b-tagged jets. Furthermore, events must satisfy $\mathcal{D}_{2\text{jet}}^{VBF} > 0.5$ [79].
- The VH-hadronic category requires exactly four leptons. Additionally, events must have either two or three jets, or at least four jets with no b-tagged jets. Events are required to satisfy $\max(\mathcal{D}_{2\text{jet}}^{WH}, \mathcal{D}_{2\text{jet}}^{ZH}) > 0.5$ [79].
- The Untagged category comprises all remaining events.

The selected events are split into the three categories summarized with their observables in Table 4.6: VBF-tagged, VH-tagged, and Untagged. Because the EW off-shell SM signal samples are dominated by VBF production, the fraction of VH events is predictably small, as reflected in Table 4.7.

In each event category, the observables $\vec{x} = \{m_{4\ell}, \mathcal{D}_{\text{bkg}}, \mathcal{D}_{\text{int}}\}$ are defined following Equations 4.2 and 4.3, and as summarized in Table 4.6. The \mathcal{D}_{bkg} observable is calculated with category-dependent methodology. In the Untagged category, \mathcal{P}_{bkg} is evaluated for the dominant $q\bar{q} \rightarrow 4\ell$ background process. The signal and background

| Category | VBF-tagged | VH -tagged | Untagged |
|-------------|---|---|--|
| Selection | $\mathcal{D}_{2\text{jet}}^{\text{VBF}} > 0.5$ | $\mathcal{D}_{2\text{jet}}^{ZH}$ or $\mathcal{D}_{2\text{jet}}^{WH} > 0.5$ | Rest of events |
| Observables | $m_{4\ell}, \mathcal{D}_{\text{bkg}}^{\text{VBF+dec}}, \mathcal{D}_{\text{bsi}}^{\text{VBF+dec}}$ | $m_{4\ell}, \mathcal{D}_{\text{bkg}}^{VH+\text{dec}}, \mathcal{D}_{\text{bsi}}^{VH+\text{dec}}$ | $m_{4\ell}, \mathcal{D}_{\text{bkg}}^{\text{kin}}, \mathcal{D}_{\text{bsi}}^{\text{gg,dec}}$ |

Table 4.6: Summary of the three production categories in the off-shell $m_{4\ell}$ region. All discriminants are calculated with the JHUGen signal and MCFM background matrix elements. The VH interference discriminant in the VH-tagged category is the simple average of the ones corresponding to the ZH and WH processes [100].

probabilities incorporate both the matrix element probability, derived from the four-lepton kinematics, and a probability density function for the four-lepton invariant mass ($m_{4\ell}$), which is parameterized using simulated events to account for detector effects. The Higgs boson mass is constrained to $m_H = 125.38 \text{ GeV}$ [105]. An illustrative example of the differentiating value of $\mathcal{D}_{\text{bkg}}^{\text{kin}}$ is shown in Figure 4.5.

In the VBF-tagged and VH-tagged categories, \mathcal{P}_{bkg} and \mathcal{P}_{sig} incorporate not only four-lepton kinematics but also the kinematic properties of the two associated jets. The \mathcal{P}_{bkg} probability density characterizes the combined EW and QCD background processes $4\ell + 2\text{jets}$, while \mathcal{P}_{sig} represents the EW processes VBF and VH. Inclusion of jet kinematics in the \mathcal{D}_{bkg} calculation enhances discrimination of the targeted signal production against both background and Higgs boson production through gluon fusion.

The third observable, \mathcal{D}_{int} defined in Equation 4.3, discriminates between the interfer-

ence of SM-like Higgs boson coupling and background as an alternative model, and is designated \mathcal{D}_{bsi} for the signal-background interference in the off-shell region. In the Untagged category, decay kinematic information is utilized in the calculation of \mathcal{D}_{int} . In the VBF-tagged and VH-tagged categories, production information incorporating the two associated jets is employed.

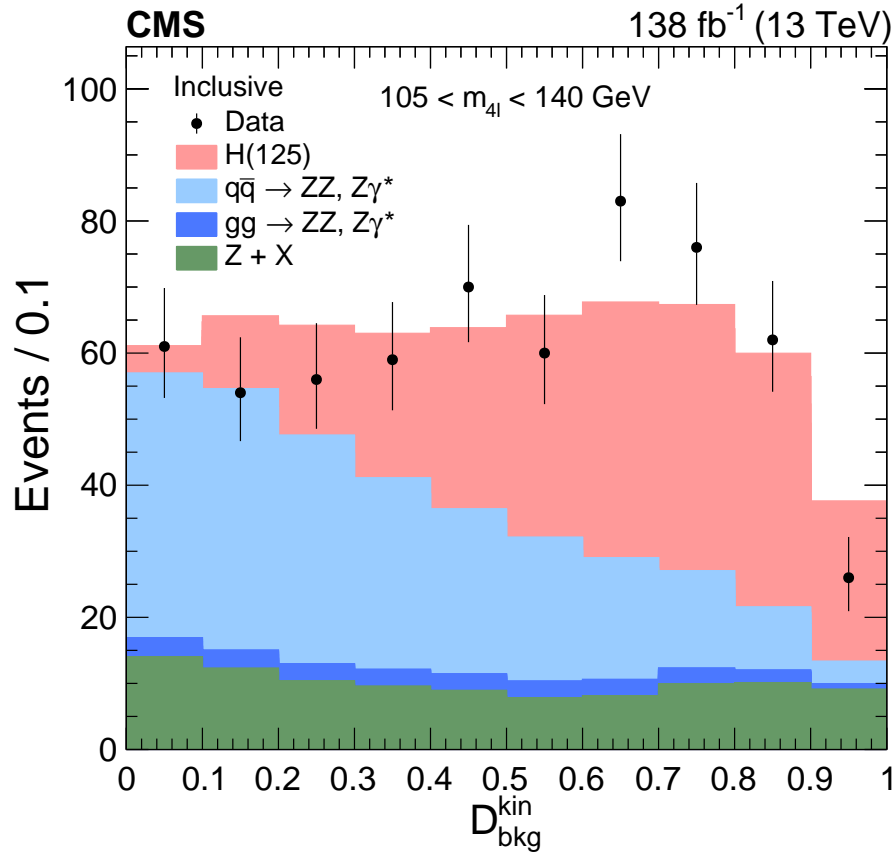


Figure 4.5: Distributions of the observed (points) and predicted (stacked histograms) $\mathcal{D}_{\text{bkg}}^{\text{kin}}$ of the four-lepton system, in the inclusive final state. The predictions for the Higgs boson signal and the three main backgrounds are given by the different colors. The vertical bars on the points show the statistical uncertainties in the data [100].

The gluon fusion cross section is computed utilizing the highest-order QCD and

electroweak (EW) expansions available for inclusive ZZ production [4]. However, event categorization relies on the modeling of associated jets through PYTHIA’s parton showering and hadronization algorithms, which necessitates proper matching to the underlying hard-scattering process. Off-shell gluon fusion production is generated at leading order (LO) without associated jets at the matrix element level, implying that all reconstructed jets originate from PYTHIA’s modeling. The parton showering and hadronization processes require specification of the hadronization scale, which is typically dependent on the characteristic energy scale of the process—set to $m_{4\ell}$ in the case of $gg \rightarrow 4\ell$.

For EW production mechanisms, the inclusive production cross section of ZZ with two associated jets is similarly calculated using the highest-order QCD and EW expansions available [4]. In contrast to gluon fusion, two hard jets—typically the leading jets in the event—are generated intrinsically at the matrix element level in the LO simulation. These jets correspond to either the forward-backward jets characteristic of vector boson fusion (VBF), or to the hadronic decay products of an associated W or Z boson. Consequently, the dependence on PYTHIA’s parton shower and hadronization algorithms, and their matching to the hard-scattering production, is substantially reduced for these EW processes.

The categorization efficiency for simulated ggH and EW Higgs boson production can be validated through comparison with alternative POWHEG and MINLO simulations. The POWHEG samples are generated across a broad range of off-shell Higgs boson masses at next-to-leading order (NLO) in QCD, with one jet generated at the matrix element level and additional jets simulated through PYTHIA matching. The MINLO

simulation of ggH production is performed for Higgs boson masses of 125 and 300 GeV at NLO in QCD, with at least two jets generated at the matrix element level and additional jets modeled via PYTHIA matching.

While JHUGen categorization efficiencies are consistent with those from POWHEG and MINLO within the uncertainties associated with the QCD scale implementation in PYTHIA, the ggH process exhibits deviations in central values and corresponding uncertainties of up to 20% in the signal-dominated $m_{4\ell}$ range of 300–500 GeV, with category-dependent variations. For EW processes, categorization efficiencies typically agree within 5–10% between methodologies, for the primary categories of Untagged and VBF-tagged events. We implement a correction procedure whereby the categorization efficiency as a function of $m_{4\ell}$ is adjusted to match that of the SM signal derived from POWHEG samples, with the assumption that this correction applies equally to background and interference contributions. This procedure ensures conservation of the total event yield across all three categories at each $m_{4\ell}$ value. An illustration of these categorization corrections for the EW production modes can be found in Figure 4.6.

The simulation of the \vec{x} observables enumerated in Table 4.6 remains unaffected by jet modeling in the Untagged category. Furthermore, the modeling of observables in jet-tagged categories for EW processes demonstrates remarkable consistency between direct MCFM+JHUGen samples and reweighted POWHEG+JHUGen implementations. However, jet modeling in jet-tagged categories for the ggH process significantly impacts the resulting probability distributions, as illustrated in Figures 4.7 and 4.8. Therefore, within these two jet-tagged categories, the gluon fusion process is implemented through reweighted POWHEG+JHUGen simulation, which provides more

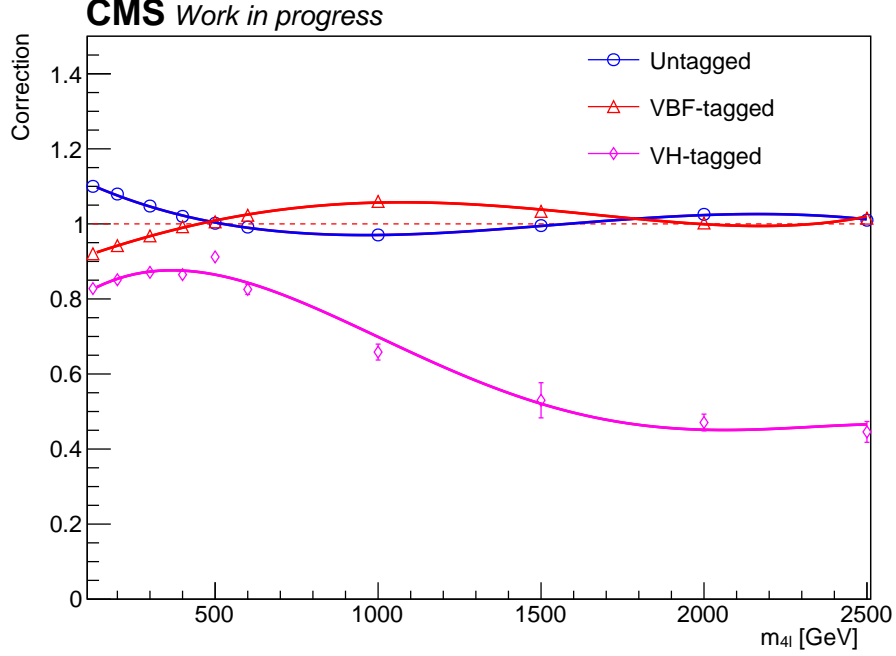


Figure 4.6: Corrections as a function of $m_{4\ell}$ for off-shell EW process selection efficiencies in the Untagged, VBF-tagged, and VH-tagged categories. A polynomial fit is performed extending up to 2.5 TeV. Events with $m_{4\ell} > 2.5$ TeV assume the correction value at 2.5 TeV. The corrections are shown for all of Run 2 combined.

precise parton shower matching and consequently superior modeling of associated jet activity. In this framework, samples are reweighted according to the appropriate theoretical model using the MELA package.

The observed and expected distributions of observables \vec{x} for events in the off-shell region are illustrated in Figure 4.9 for each of the three categories. The expected distributions are found using the SM predicted signal and background cross sections.

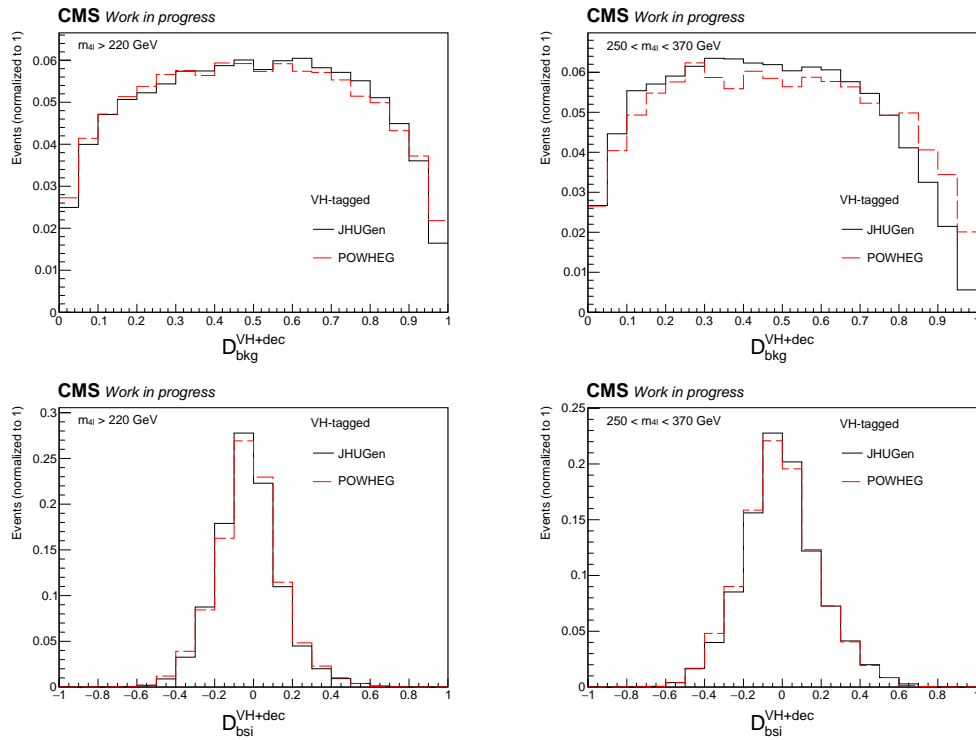


Figure 4.7: Comparison of $\mathcal{D}_{\text{bkg}}^{\text{VH+dec}}$ and $\mathcal{D}_{\text{bsi}}^{\text{VH+dec}}$ distributions of ggH events from JHUGen and POWHEG for $m_{4\ell} > 220$ GeV (left) and $250 < m_{4\ell} < 370$ GeV (right).

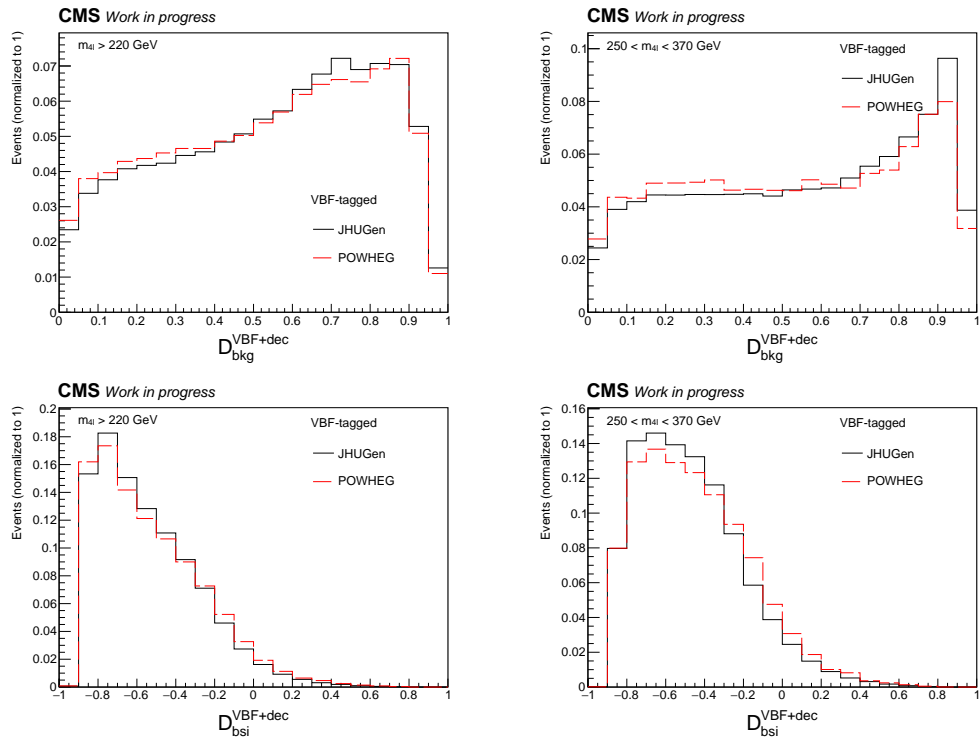


Figure 4.8: Comparison of $D_{bkg}^{VBF+dec}$ and $D_{bsi}^{VBF+dec}$ distributions of ggH events from JHUGen and POWHEG for $m_{4\ell} > 220$ GeV (left) and $250 < m_{4\ell} < 370$ GeV (right).

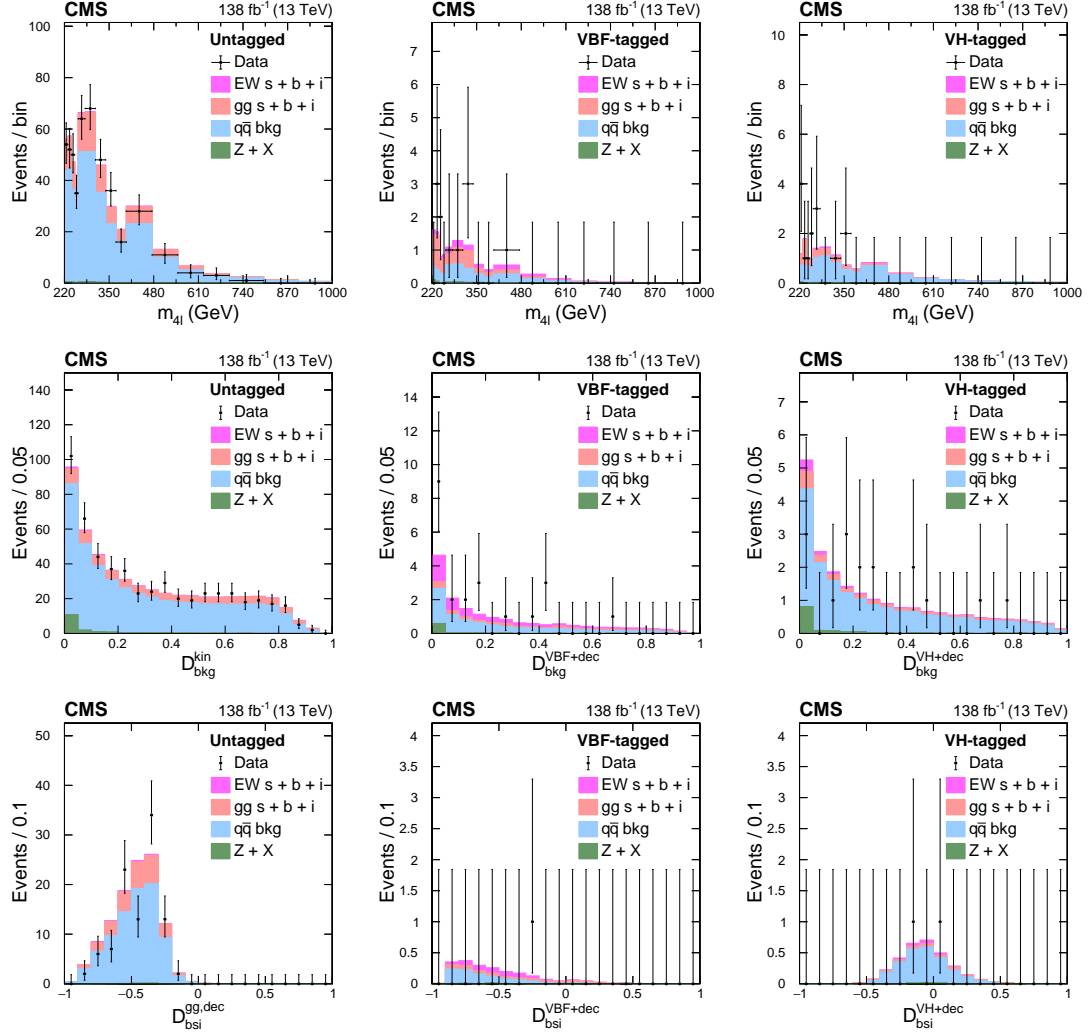


Figure 4.9: Off-shell data (points) and pre-fit distributions (histograms) for the Untagged (left), VBF-tagged (middle), and VH-tagged (right) categories. The upper row shows $m_{4\ell}$ distributions with a requirement on $\mathcal{D}_{\text{bkg}}^{\text{kin}} > 0.6$ (left), $\mathcal{D}_{\text{bkg}}^{\text{VBF+dec}} > 0.6$ (middle), or $\mathcal{D}_{\text{bkg}}^{\text{VH+dec}} > 0.6$ (right) applied for illustration purposes to enhance signal over background contributions. The middle row shows $\mathcal{D}_{\text{bkg}}^{\text{kin}}$ (left), $\mathcal{D}_{\text{bkg}}^{\text{VBF+dec}}$ (middle), $\mathcal{D}_{\text{bkg}}^{\text{VH+dec}}$ (right) distributions, where an additional requirement $m_{4\ell} > 340$ GeV is applied to enhance signal-over-background contributions. The lower row plots the \mathcal{D}_{bsi} with both the $m_{4\ell}$ and $\mathcal{D}_{\text{bkg}}^{\text{kin}}$ requirements specified above. Contributions from the four processes are shown by the different colors, where “s”, “b”, and “i” refer to the signal, background, and interference contributions, respectively. The vertical bars on the points give the statistical uncertainties in the data, and the horizontal bars represent the bin widths. For the pre-fit distributions, the different cross sections are set to their SM values [100].

4.2.4 Modeling of background

The largest background in the $H \rightarrow 4\ell$ channel originates from the process $q\bar{q} \rightarrow ZZ/Z\gamma^*/\gamma^*\gamma^* \rightarrow 4\ell$. Additionally, the processes $gg \rightarrow ZZ/Z\gamma^*/\gamma^*\gamma^* \rightarrow 4\ell$ and electroweak (EW) production—including vector boson scattering and VZZ processes—also contribute to the background and interfere with off-shell Higgs boson production. In the on-shell region, this interference is negligible; however, the EW background additionally includes other processes such as VVV , $t\bar{t}VV$, and $t\bar{t}V$ [100].

Another important background arises from processes collectively referred to as $Z + X$, where leptons are misidentified from decays of heavy-flavor hadrons, in-flight decays of light mesons within jets, or charged hadrons overlapping with π^0 decays. The dominant contribution to this background is from Z +jets events, estimated using data-driven methods in dedicated control regions. These control regions are defined by selecting events containing one lepton pair meeting the Z_1 candidate requirements, along with an additional pair of opposite-sign leptons passing looser identification criteria compared to the main analysis selection. These four leptons must then satisfy the Z_1, Z_2 selection criteria. The background yield in the signal region is obtained by weighting the control region events by the lepton misidentification probability (f_ℓ), defined as the probability for a non-prompt lepton to pass the analysis selection. A detailed description of this method can be found in Reference [64]. The $m_{4\ell}$ distribution for the $Z + X$ background is modeled by a Landau function, determined from data over an extended invariant mass range of 70–770 GeV due to limited event counts.

The observed numbers of data events, expected backgrounds, and signal yields in the off-shell region are summarized in Table 4.7. The on-shell events are inherited

| | VBF-tagged | VH-tagged | Untagged | Total |
|---|------------|-----------|----------|---------|
| $gg \rightarrow 4\ell$ signal | 1.70 | 1.44 | 40.78 | 43.92 |
| $gg \rightarrow 4\ell$ background | 18.65 | 15.28 | 495.76 | 529.69 |
| $gg \rightarrow 4\ell$ interference | -2.86 | -2.40 | -69.57 | -74.83 |
| EW signal | 1.78 | 0.12 | 2.47 | 4.37 |
| EW background | 15.41 | 1.77 | 19.25 | 36.43 |
| EW interference | -3.84 | -0.06 | -4.63 | -8.53 |
| ZH cross-feed | 0.24 | 0.44 | 7.30 | 7.98 |
| $q\bar{q} \rightarrow 4\ell$ background | 25.90 | 46.00 | 1895.20 | 1967.11 |
| $Z + X$ background | 4.32 | 4.50 | 75.47 | 84.29 |
| Total expected | 61.09 | 66.72 | 2455.20 | 2583.01 |
| Total observed | 63 | 71 | 2420 | 2554 |

Table 4.7: Observed and expected yields for the Higgs boson signal and background contributions in the off-shell region $m_{4\ell} > 220$ GeV, for each of the four-lepton categories and the total. The yields from interference of the signal and background and the ZH cross-feed are also shown. The expected yields are adjusted within their respective uncertainties from the fit to the data [100].

from the latest CMS analysis which considers anomalous Higgs boson couplings in the on-shell region [80]. The signal and ZZ background yields are derived from simulation, whereas the $Z + X$ background yield is obtained directly from data. Further details on the modeling of the Higgs boson signal, interference effects with background processes, and ZH cross-feed in the off-shell analysis are provided in Section 4.2.6.

4.2.5 Parameterizing the on-shell Higgs boson

The probability density for the on-shell region includes both signal and background contributions. It is normalized to the total event yield for each process j and category

k , and can be written as

$$\mathcal{P}_{jk}(\vec{x}; \vec{\xi}_{jk}, \vec{\zeta}) = \mu_j \mathcal{P}_{jk}^{\text{sig}}(\vec{x}; \vec{\xi}_{jk}, \vec{\zeta}) + \mathcal{P}_{jk}^{\text{bkg}}(\vec{x}; \vec{\xi}_{jk}), \quad (4.4)$$

where signal strength μ_j is defined as the ratio of the signal yield to the SM expectation, $\vec{\zeta}$ are unconstrained parameters of interest, $\vec{\xi}_{jk}$ are constrained nuisance parameters for a particular parametrization, and \vec{x} are the observables.

In the on-shell Higgs boson mass and width analysis, six signal production processes are considered: ggH , VBF , VH , $t\bar{t}H$, $b\bar{b}H$, and tqH . The primary background contributions arise from three processes: $q\bar{q}/gg \rightarrow ZZ/Z\gamma^*/\gamma^*\gamma^* \rightarrow 4\ell$ and $Z + X$.

When constraints on the total Higgs boson width (Γ_H) are extracted through a simultaneous fit to both the on-shell and off-shell regions, the on-shell component follows Scheme 2 from the previous on-shell analysis [80] which also provides further information on event categorization as well as signal and background modeling.

4.2.6 Parameterizing the off-shell Higgs boson

In the off-shell region, the probability density follows Equations (2.10) and (4.4) closely, with the additional contribution of the interference (“int”) between the signal and background amplitudes, as well as a cross-feed (“cross”) component discussed below. It is parametrized as

$$\mathcal{P}_{jk}(\vec{x}; \vec{\xi}_{jk}, \vec{\zeta}) = \frac{\mu_j \Gamma_H}{\Gamma_0} \mathcal{P}_{jk}^{\text{sig}}(\vec{x}; \vec{\xi}_{jk}) + \sqrt{\frac{\mu_j \Gamma_H}{\Gamma_0}} \mathcal{P}_{jk}^{\text{int}}(\vec{x}; \vec{\xi}_{jk}) + \mu_j \mathcal{P}_{jk}^{\text{cross}}(\vec{x}; \vec{\xi}_{jk}) + \mathcal{P}_{jk}^{\text{bkg}}(\vec{x}; \vec{\xi}_{jk}), \quad (4.5)$$

where Γ_0 is the reference value of the Higgs boson width (not necessarily its SM value) used in simulation. Otherwise, the definition of the terms is the same as in Equation (4.4). (Notably, μ_j is the signal strength on-shell.) In the off-shell width analysis, there are two production modes, j (ggH and EW, the latter of which includes both VBF and VH), and three jet-tagged categories, k . All lepton flavors and data periods are combined in this analysis. The contributions from $t\bar{t}H$, $b\bar{b}H$, and tqH are expected to be negligible at high energies.

The $\mathcal{P}_{jk}^{\text{sig}}$, $\mathcal{P}_{jk}^{\text{int}}$, $\mathcal{P}_{jk}^{\text{cross}}$, and $\mathcal{P}_{jk}^{\text{bkg}}$ probability densities are normalized to the expected number of events, and are implemented as binned histograms (templates) of the observables \vec{x} listed in Table 4.6. These templates are obtained as weighted linear combinations of existing simulated signal or background samples [100].

4.2.6.1 Cross-feed effect

The off-shell region includes all events with $m_{4\ell} > 220$ GeV. In this region, other processes can mimic off-shell Higgs boson production and decay to four leptons. In particular, we consider on-shell Higgs boson events where the Higgs boson decays to $2\ell + X$, where the X can be replaced by mislabeled associated particles such that the event satisfies the four-lepton selection criteria. The dominant on-shell Higgs boson process that can contribute to the off-shell region is $Z(\ell\ell)H(2\ell + X)$.

This ZH cross-feed contribution is estimated using on-shell ZH simulated samples with $H \rightarrow WW$ and ZZ decays, allowing for both hadronic and leptonic decays of the Z boson. The most significant contributions come from $H \rightarrow 2\ell 2\nu$ and $H \rightarrow 2\ell 2q$ final states, produced in association with $Z \rightarrow 2\ell$. Other possible contributions include the

on-shell $t\bar{t}H$ production (0.57 events) and the ZH production of the Higgs boson with subsequent decay to two muons (< 0.1 events).

Templates for the statistical analysis are prepared using the on-shell ZH samples to model cross-feed as a separate on-shell process. At the same time, to avoid double counting, on-shell cross-feed events are explicitly removed from the off-shell simulated samples. Any simulated $H \rightarrow 2\ell 2q$ events in the off-shell EW JHUGen sample with mass close to the Higgs boson pole mass within 100 MeV are removed. This procedure has close to 100% efficiency of selecting on-shell cross-feed events, while having a negligible effect on the other events. In addition to the ZH templates, the $t\bar{t}H$ contribution is modeled and included in the statistical analysis.

In Figure 4.10 (left), the $m_{4\ell}$ distributions of three simulated samples are shown: the on-shell cross-feed process ZH(125), the genuine off-shell EW signal process with cross-feed contributions removed, and the dominant $q\bar{q} \rightarrow 4\ell$ background, which is scaled down by three orders of magnitude to facilitate comparison. The cross-feed and background processes exhibit similar kinematic distributions, while the genuine off-shell signal displays a distinct shape with a dominant contribution at higher energies.

The right panel of Figure 4.10 shows the $m_{2\ell 2q}$ invariant mass distribution for events reconstructed in the off-shell signal region (plotted at MC-truth level). A narrow peak corresponding to the on-shell cross-feed component is clearly visible at 125 GeV. A study of the event categorization reveals that the majority of these cross-feed events ($> 91\%$) are classified into the Untagged category, with the remainder distributed between the VBF- and VH-tagged categories used in the analysis.

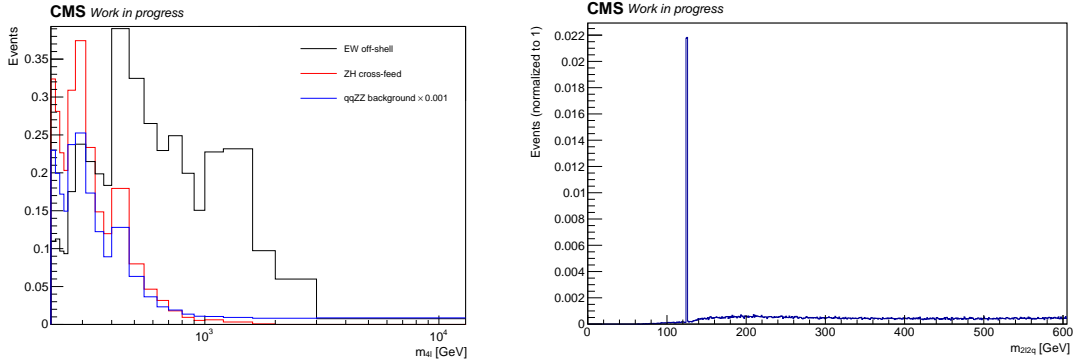


Figure 4.10: Left: Distributions over $m_{4\ell}$ of on-shell cross-feed events, off-shell signal, and the dominant $q\bar{q} \rightarrow 4\ell$ background process (scaled down by three orders of magnitude). Right: Distribution of $m_{2\ell 2q}$ events reconstructed in the off-shell region of the $H \rightarrow 4\ell$ analysis. Note the peak at 125 GeV with negligible yield elsewhere.

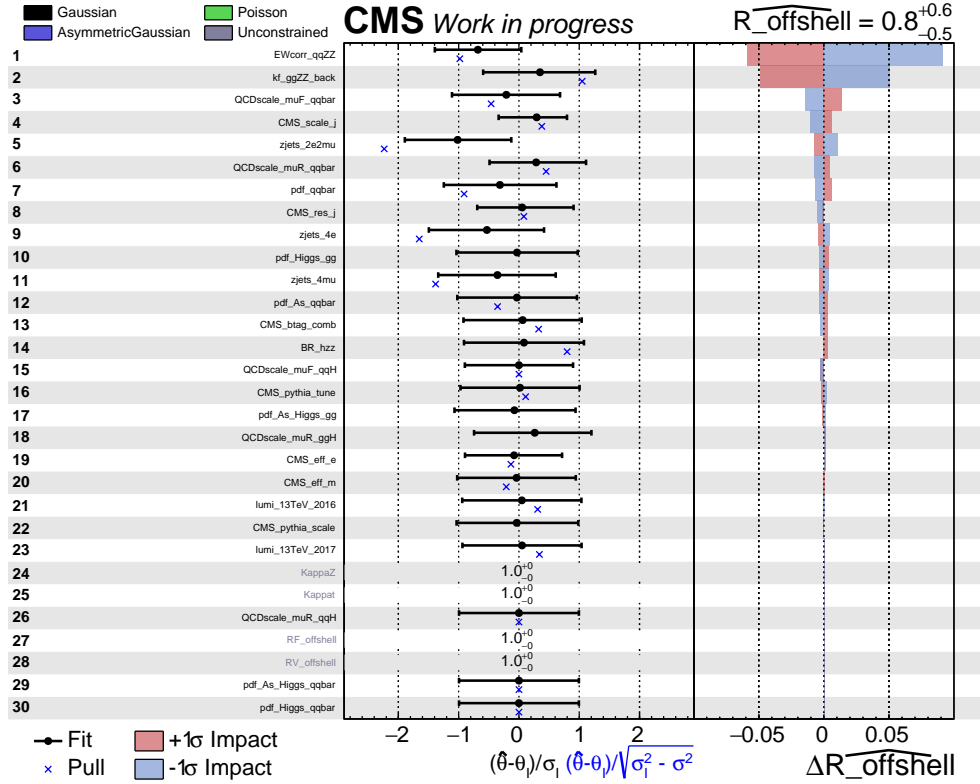


Figure 4.11: Observed impacts of nuisance parameters on the Higgs boson width.

4.2.7 Systematic uncertainties

Several systematic uncertainties are evaluated in the measurement of the constrained parameters $\vec{\xi}_{jk}$. The template shapes used to describe the probability distributions in Equations (4.4) and (4.5) are independently varied within their theoretical and experimental uncertainties. The resulting changes in the constrained parameters are taken as the systematic uncertainties associated with each source.

In the off-shell width measurement, the largest systematic uncertainty originates from the modeling of the dominant background process, $q\bar{q}/gg \rightarrow ZZ/Z\gamma^*/\gamma^*\gamma^* \rightarrow 4\ell$. Experimental uncertainties specific to the off-shell analysis include the jet energy calibration, which primarily affects the VBF- and VH-tagged categories.

Three sources of jet reconstruction-related uncertainties are studied and incorporated into the off-shell analysis. These include uncertainties in the jet energy scale (JES), jet energy resolution (JER), and the b-tagging scale factor. While the event categorization depends on the presence and properties of associated jets, the total event yield integrated over all three categories is not expected to vary significantly under these uncertainties. Instead, the dominant effect is a redistribution of events among the categories and corresponding changes in the shapes of their discriminant variables.

To account for these effects, alternative calculations of the categorization discriminants and the jet-related logic (including the number of b-tagged jets) are performed for all variations of each uncertainty source. In addition, variations of the final observable discriminants are also computed. From these, up- and down-shifted templates are constructed and incorporated in the analysis as systematic shape uncertainties.

Theoretical uncertainties impacting both the signal and background predictions include those associated with the renormalization and factorization scales, as well as the choice of parton distribution functions (PDFs). The uncertainty from the renormalization and factorization scales is estimated by varying both scales by factors of 0.5 and 2 relative to their nominal values, while maintaining the ratio between them within the range $[0.5, 2]$. PDF uncertainties are evaluated by computing the root mean square of variations across replicas of the default NNPDF set. An additional 10% uncertainty is applied to the K factor used in the ggZZ background prediction.

The integrated luminosities for the 2016, 2017, and 2018 data-taking periods carry individual uncertainties of 1.2–2.5% [106–108], with a combined uncertainty of 1.6% for the full 2016–2018 dataset. The uncertainties in lepton identification, reconstruction, and selection efficiencies range from 2% (for the 4μ final state) to 14% (for the $4e$ final state), affecting both signal and background yields.

For the estimation of the $Z + X$ background, differences in the flavor composition of hadronic jets misidentified as leptons between the $Z + 1\ell$ and $Z + 2\ell$ control regions introduce an uncertainty of approximately 30% in the background yield. Additionally, uncertainties in the modeling of the misidentification rate as a function of p_T and η , combined with the statistical uncertainty in the $Z + 1\ell$ control region, lead to total uncertainties in the background yields ranging from 32% in the $4e$ final state to 39% in the 4μ final state.

An impact plot of the nuisance parameters on the latest observed Higgs boson width result is presented in Figure 4.11.

4.3 Measurement of Higgs boson properties

4.3.1 Higgs boson width

An extended binned maximum likelihood fit is performed on the combined on-shell and off-shell event samples, divided into multiple analysis categories. The final measurements of m_H , Γ_H , and the signal strength parameters μ_j are carried out using the CMS statistical analysis tool COMBINE [109].

The extended likelihood function is constructed from the probability densities defined in Equations (4.4) and (4.5), where each event is characterized by a discrete category index k and typically three observables \vec{x} . The likelihood \mathcal{L} is maximized with respect to the nuisance parameters $\vec{\xi}_{jk}$, which encode systematic uncertainties, and the signal strength parameter μ (representing the total signal yield), or alternatively μ_F and μ_V , corresponding to the gluon-fusion and electroweak production processes, respectively. Confidence intervals at 68% and 95% CL are determined using the profile likelihood method, with thresholds of $-2\Delta \ln \mathcal{L} = 1.00$ and 3.84 [110].

Constraints on Γ_H are extracted from a simultaneous fit to $H \rightarrow ZZ \rightarrow 4\ell$ events in both the on-shell and off-shell regions. Treatment of the on-shell region is inherited as previously discussed in Section 4.2.5, and uses six mutually exclusive categories with all anomalous Higgs boson interactions set to zero. This fit yields two signal strength parameters, $\mu_F^{\text{on-shell}}$ and $\mu_V^{\text{on-shell}}$, as defined in Equations (4.4) and (4.5), corresponding to production via fermionic and vector boson couplings of the Higgs boson, respectively. The Higgs boson mass is fixed to $m_H = 125.38 \text{ GeV}$ [105].

| Channel | Observed Γ_H (MeV) | Expected Γ_H (MeV) |
|--|--------------------------------|---------------------------|
| 4ℓ on- and off-shell + $2\ell 2\nu$ off-shell | $3.0^{+2.0}_{-1.5}$ [0.6, 7.3] | 4.1 ± 3.5 [0.1, 10.5] |

Table 4.8: Summary of the ReReco-based total Higgs boson width Γ_H measurement, showing the 68% CL (central values with uncertainties) and 95% CL (in square brackets) intervals for the $H \rightarrow ZZ \rightarrow 4\ell$ channel in combination with the off-shell $H \rightarrow ZZ \rightarrow 2\ell 2\nu$ channel [100].

As described in Section 4.2, the measurements of Higgs boson width and signal strengths were performed using UL processed samples. However, before the UL processing became available, the analysis was first conducted and published using ReReco processing [100].

In the ReReco case, the results of this analysis with SM-like couplings are combined with the prior CMS off-shell $H \rightarrow ZZ \rightarrow 2\ell 2\nu$ analysis [69], giving the first CMS measurement of Γ_H using the full 4ℓ and $2\ell 2\nu$ data sample collected during Run 2. These combined fit results supersede the previous CMS results [69] under the SM-like coupling assumption.

The ReReco-based observed and expected constraints on the Higgs boson width are presented in Table 4.8, while the profile likelihood scan of Γ_H , obtained using the asymptotic approximation, is shown in Figure 4.12. These results rule out the scenario of no off-shell Higgs boson production with a CL corresponding to 3.8 standard deviations (average expected 2.4).

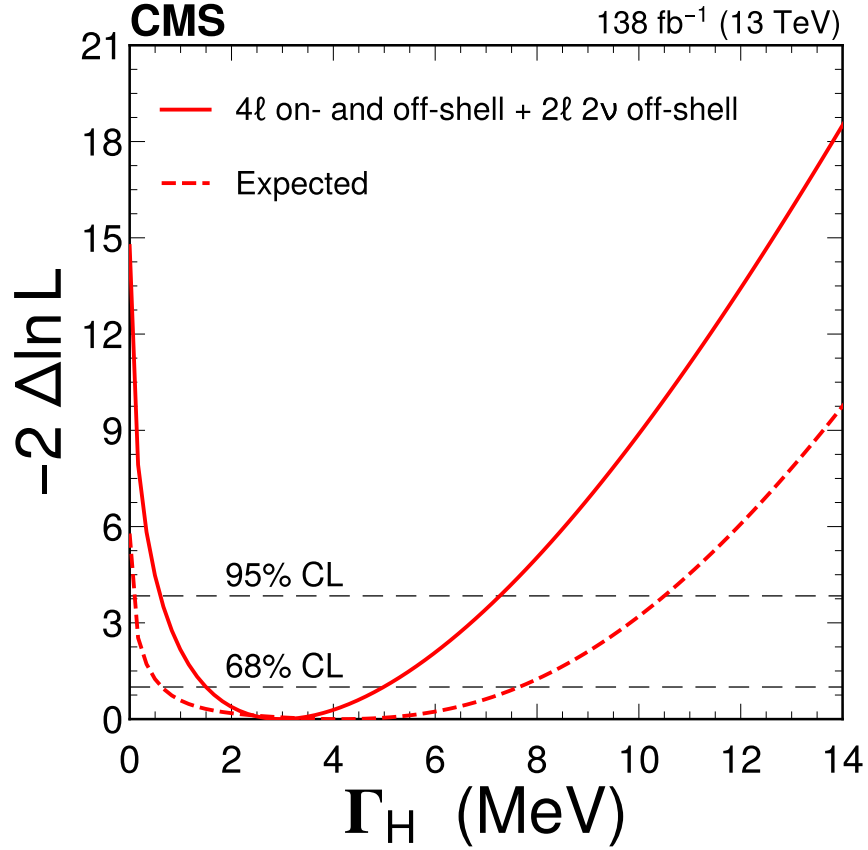


Figure 4.12: ReReco-based observed (solid) and expected (dashed) profile likelihood projections from the Higgs boson width fit using on-shell [80] and off-shell $H \rightarrow ZZ \rightarrow 4\ell$ production, combined with off-shell $H \rightarrow ZZ \rightarrow 2\ell 2\nu$ [69]. The black horizontal dashed lines show the 68% and 95% CL values [100].

The observed limits on Γ_H are more stringent than the average expectations derived from simulation. This behavior is supported by the template distributions, in particular the upper-left panel of Figure 4.9, where the number of observed events in the region defined by $m_{4\ell} > 340$ GeV and $\mathcal{D}_{\text{bkg}} > 0.6$ in the Untagged category is lower than expected, though still statistically consistent.

A reduced number of events in this sensitive region favors the hypothesis of greater negative interference between the signal and background contributions, which becomes dominant over the pure signal contributions for Γ_H values near the SM value. As a result, both large and very small values of Γ_H are disfavored by the data.

Conducting the width measurement using newly available UL processed samples gives us a slightly modified result, with our observed Higgs boson width moving closer to our new expected value but with observed bounds still remaining more stringent than expected, as shown in Figure 4.13 and summarized in Table 4.9.

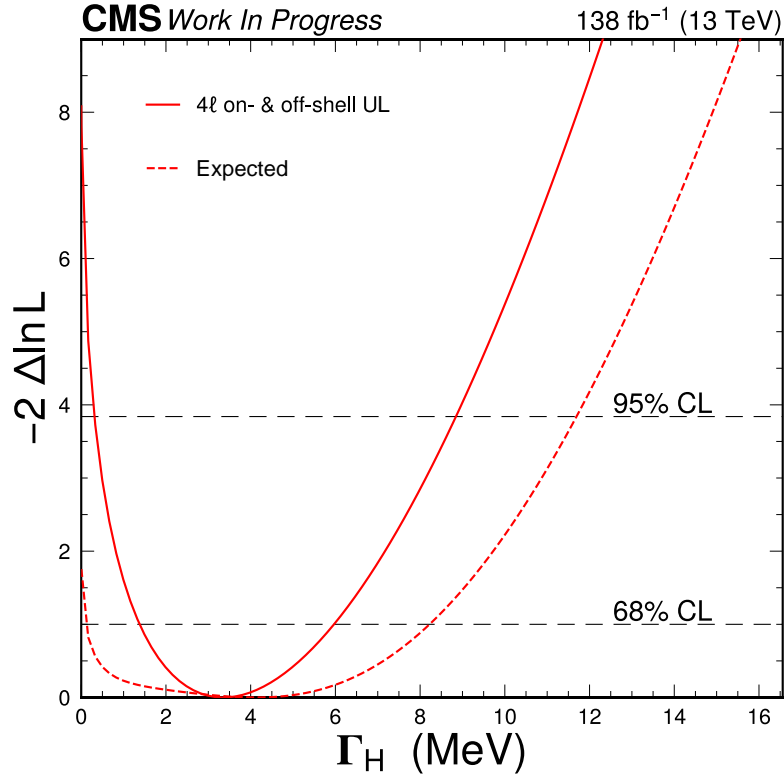


Figure 4.13: Observed (solid) and expected (dashed) likelihood projections from the Higgs boson width fit using on- and off-shell production, conducted with UL processed off-shell samples. The black horizontal dashed lines show the 68% and 95% CL values.

4.3.2 Off-shell signal strength

An alternative approach to the off-shell region fit involves decoupling its signal strength from that of the on-shell region. In this configuration, the signal yield is scaled by a single parameter, $\mu^{\text{off-shell}}$, which is common to all production mechanisms. The total Higgs boson width is fixed to its SM value, $\Gamma_H = \Gamma_H^{\text{SM}}$, in Equation (4.5), and the SM expectation corresponds to $\mu^{\text{off-shell}} = 1$. The likelihood scans for this parameter are presented in Figure 4.14, conducted using UL processed samples. Note that these new UL-based results now supersede the previously published ReReco-based measurements [100].

In addition, a more general fit is performed with two unconstrained signal strength parameters: $\mu_F^{\text{off-shell}}$ and $\mu_V^{\text{off-shell}}$, corresponding to gluon-fusion and electroweak production mechanisms, respectively. The UL-based measured values of these parameters are reported in Table 4.9 and illustrated in Figure 4.14, with their joint CL regions shown in the two-dimensional likelihood scan to the bottom-right of Figure 4.14.

| Parameter | Observed | Expected |
|----------------------------|------------------------------------|-------------------------------|
| Γ_H | $3.4^{+2.6}_{-2.0}[0.3, 8.9]$ | $4.1^{+4.1}_{-4.0}[<11.7]$ |
| $\mu^{\text{off-shell}}$ | $0.75^{+0.57}_{-0.44}[0.07, 1.94]$ | $1.00^{+1.02}_{-0.96}[<2.83]$ |
| $\mu_F^{\text{off-shell}}$ | $0.66^{+0.63}_{-0.46}[0.02, 1.97]$ | $1.00^{+1.04}_{-1.00}[<2.89]$ |
| $\mu_V^{\text{off-shell}}$ | $0.92^{+1.46}_{-0.78}[<4.40]$ | $0.96^{+3.29}_{-0.96}$ |

Table 4.9: Summary of all UL-based results. Total Higgs boson width Γ_H and signal strength μ_j measurements, with 68% CL (central values with uncertainties) and 95% CL (in square brackets) intervals for the $H \rightarrow ZZ \rightarrow 4\ell$ channel.

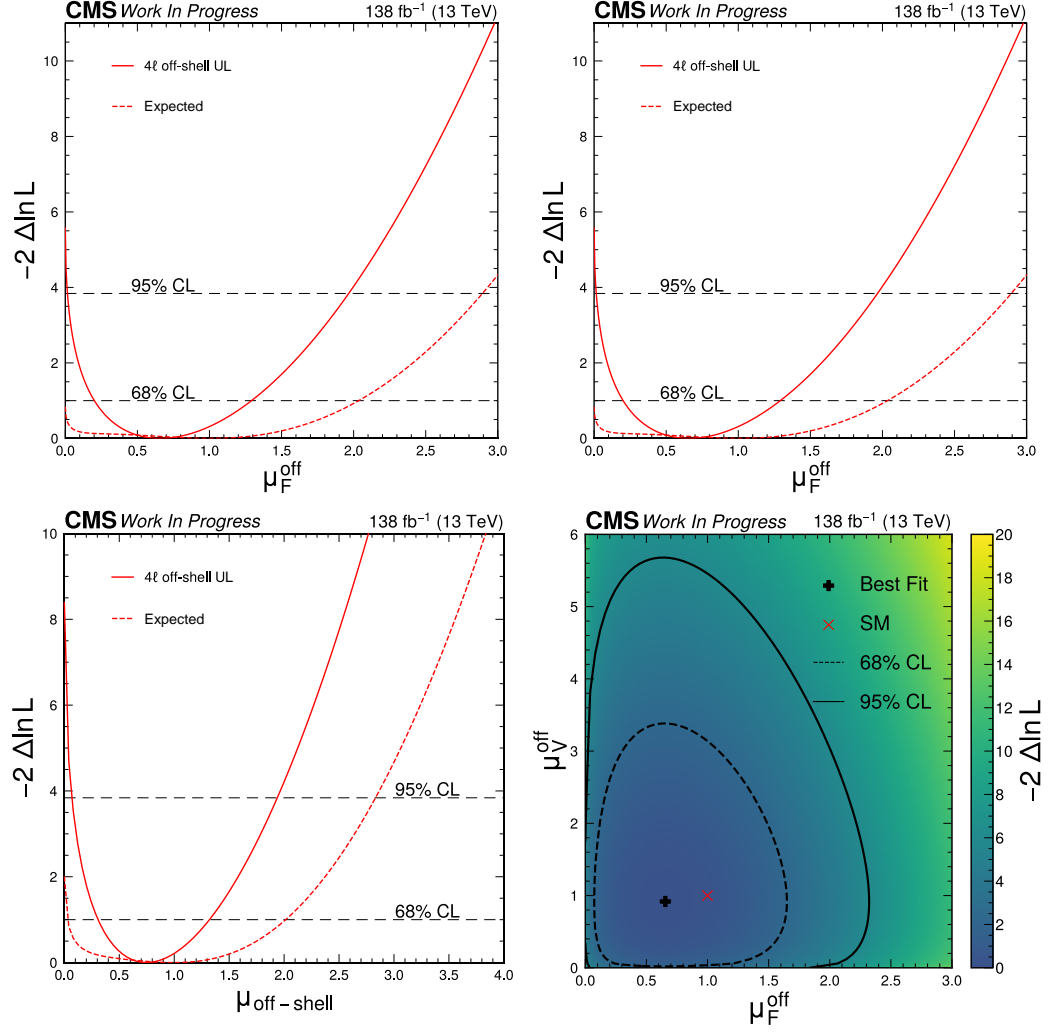


Figure 4.14: Observed and expected profile likelihood scans for $\mu_F^{\text{off-shell}}$ (top-left), $\mu_V^{\text{off-shell}}$ (top-right), and $\mu_{\text{off-shell}}$ (bottom-left) conducted using UL processed samples. Additionally plotted is the observed 2D profile likelihood projection (bottom-right) of the off-shell signal strength parameters $(\mu_F^{\text{off-shell}}, \mu_V^{\text{off-shell}})$ from the fit to the off-shell $H \rightarrow ZZ \rightarrow 4\ell$ channel. The best fit value is shown by the black cross and the SM prediction by the red x. The 68% and 95% CL contours are given by the dashed and solid curves, respectively. The color scale to the right of the plot relates the quantitative values.

4.3.3 Kappa framework

The Γ_H constraints presented in Section 4.3.1 assume SM-like evolution of the Higgs boson couplings across a wide $m_{4\ell}$ range. Anomalous contributions to the HVV vertex in EW production and Higgs boson decay were investigated in earlier analyses using smaller datasets [7, 69], and the resulting Γ_H constraints remained consistent. However, both the current and previous analyses assume that the gluon-fusion production mechanism is dominated by the top quark loop. If additional heavy particles contribute to the loop, this assumption would no longer hold, and the $m_{4\ell}$ dependence in the off-shell region could be significantly modified. This possibility is explored using ReReco processed samples [100].

To explore the impact of potential contributions from new heavy particles in the $gg \rightarrow H$ loop, a hypothetical heavy quark Q is introduced with an unconstrained coupling strength κ_Q , such as is described in Section 2.6.2 [9, 10]. In the framework of Effective Field Theory (EFT), the influence of such a particle is modeled as a point-like interaction that encapsulates its loop effects. The parametrization in Equation (4.5) is extended to include template components proportional to κ_Q and κ_Q^2 , obtained from simulations reweighted using the MELA package in the limit of infinite Q mass.

The $m_{4\ell}$ distribution in the off-shell region exhibits different behaviors for SM-like gluon fusion—dominated by the top quark with the $2m_t$ threshold—and for scenarios involving the additional heavy quark Q [9]. Allowing κ_Q to vary freely introduces additional uncertainty in the $m_{4\ell}$ spectrum, which weakens the sensitivity to Γ_H . Nonetheless, both the on-shell and off-shell $H \rightarrow 4\ell$ data constrain the allowed values of κ_Q , and the Γ_H constraints remain broadly consistent with the $\kappa_Q = 0$ case.

From a simultaneous fit to the on-shell and off-shell regions, based on ReReco processing, the observed width is measured to be $\Gamma_H = 2.7^{+2.7}_{-1.8}$ MeV, with an expected value of $4.1^{+5.5}_{-4.1}$ MeV [100]. The observed 95% confidence interval is [0.1, 8.8] MeV, with an expected range of [<14.4] MeV [100]. Future combinations with measurements from other on-shell Higgs boson production and decay channels are expected to significantly improve constraints on κ_Q , thereby reducing the allowed deviations from SM-like coupling evolution across the high- $m_{4\ell}$ regime.

Utilizing Equations 2.14 and 2.15, we can also consider likelihood scans of the parameters μ_F and μ_V with and without including κ_Q unconstrained. The numerical results are outlined in Table 4.10.

| Parameter | Configuration | Observed | Expected |
|-----------|--------------------|-------------------------------------|--------------------------------------|
| μ_F | κ_Q floated | $1.01^{+1.40}_{-0.41}$ [0.23, 9.28] | $1.00^{+5.49}_{-0.91}$ [0.80, 1.21] |
| μ_F | $\kappa_Q = 0$ | $0.87^{+0.09}_{-0.09}$ [0.69, 1.08] | $1.00^{+0.11}_{-0.10}$ [< 15.88] |
| μ_V | κ_Q floated | $1.05^{+0.44}_{-0.37}$ [0.35, 1.98] | $1.00^{+0.46}_{-0.40}$ [0.27, 1.96] |
| μ_V | $\kappa_Q = 0$ | $1.04^{+0.43}_{-0.37}$ [0.35, 1.96] | $1.00^{+0.46}_{-0.40}$ [0.27, 1.96] |

Table 4.10: Summary of constraints at 68% CL (central values with uncertainties) and 95% CL (in square brackets) on the parameters μ_F and μ_V with and without including κ_Q unconstrained [100].

It is worth noting that the constraints on μ_V remains largely unaffected by κ_Q being floated or fixed, whereas the constraints on μ_F exhibit significant variation. This is because the EW production mechanism, characterized by μ_V , is independent of κ_Q , whereas the ggH process is strongly influenced by both κ_Q and κ_q —the latter being related to μ_F through Equation (2.15).

Chapter 5

Conclusion

This thesis is centered around precision measurements of the Higgs boson properties (width) using the $H^{(*)} \rightarrow ZZ^{(*)} \rightarrow 4\ell$ decay channel in proton-proton collisions at $\sqrt{s} = 13 \text{ TeV}$ collected by the CMS experiment, corresponding to an integrated luminosity of 138 fb^{-1} . The strength of off-shell Higgs boson production is also measured, excluding the no off-shell production hypothesis at 3.8 standard deviations, after ReReco-based combination with $H \rightarrow ZZ \rightarrow 2\ell 2\nu$. The published ReReco-based and latest UL-based results are summarized in Tables 4.8 and 4.9, respectively.

A simultaneous fit to on- and off-shell production in the 4ℓ final state, combined with results from the $2\ell 2\nu$ final state, gives us our most data-inclusive measurement to date and yields $\Gamma_H = 3.0^{+2.0}_{-1.5} \text{ MeV}$, consistent with the SM prediction of 4.1 MeV. A 4ℓ -only measurement made with more recent UL processing returns $\Gamma_H = 3.4^{+2.6}_{-2.0} \text{ MeV}$.

These results provide rigorous tests of the Standard Model and establish an experimental baseline for characterizing Higgs boson properties. Looking ahead, this analysis provides a robust foundation for future studies targeting deviations from SM expectations. The methods and statistical techniques developed here are directly extensible to searches for anomalous Higgs boson couplings and BSM effects in the HVV vertex. This framework will support increasingly precise constraints as we move toward a Standard Model Effective Field Theory (SMEFT) description of the Higgs boson and probe Electroweak Symmetry Breaking (EWSB) to potentially uncover new physics.

Bibliographic references

1. Griffiths, D. *Introduction to Elementary Particles* (John Wiley & Sons, New York, USA, 1987).
2. The ATLAS Collaboration. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Physics Letters B* **716**, 1–29. doi:<https://doi.org/10.1016/j.physletb.2012.08.020> (2012).
3. The CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B* **716**, 30–61. doi:<https://doi.org/10.1016/j.physletb.2012.08.021> (2012).
4. LHC Higgs Cross Section Working Group. *CERN Yellow Reports: Monographs, Vol 2 (2017): Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector* 2016. doi:[10.23731/CYRM-2017-002](https://doi.org/10.23731/CYRM-2017-002). arXiv: [1610.07922 \[hep-ph\]](https://arxiv.org/abs/1610.07922).
5. Caola, F. & Melnikov, K. Constraining the Higgs boson width with ZZ production at the LHC. *Physical Review D* **88**, 054024. doi:[10.1103/PhysRevD.88.054024](https://doi.org/10.1103/PhysRevD.88.054024) (5 2013).
6. The CMS Collaboration. Constraints on the Higgs boson width from off-shell production and decay to Z-boson pairs. *Physics Letters B* **736**, 64–85. doi:[10.1016/j.physletb.2014.06.077](https://doi.org/10.1016/j.physletb.2014.06.077) (2014).
7. The CMS Collaboration. Measurements of the Higgs boson width and anomalous HVV couplings from on-shell and off-shell production in the four-lepton final state. *Physical Review D* **99**. doi:[10.1103/physrevd.99.112003](https://doi.org/10.1103/physrevd.99.112003) (2019).
8. Azatov, A., Falkowski, A., Gritsan, A. V., Grojean, C., Kauer, N., Salvioni, E., Thomas, M., Vryonidou, E., Kang, L., De Blas, J. & Sarica, U. *Off-shell Higgs Interpretations Task Force : Models and Effective Field Theories Subgroup Report* en. 2022. doi:[10.17181/LHCHWG-2022-001](https://doi.org/10.17181/LHCHWG-2022-001).
9. Gritsan, A. V., Roskes, J., Sarica, U., Schulze, M., Xiao, M. & Zhou, Y. New features in the JHU generator framework: Constraining Higgs boson properties

- from on-shell and off-shell production. *Physical Review D* **102**. doi:[10.1103/PhysRevD.102.056022](https://doi.org/10.1103/PhysRevD.102.056022) (2020).
10. Davis, J., Gritsan, A. V., Guerra, L. S. M., Kyriacou, S., Roskes, J. & Schulze, M. Constraining anomalous Higgs boson couplings to virtual photons. *Physical Review D* **105**, 096027. doi:[10.1103/PhysRevD.105.096027](https://doi.org/10.1103/PhysRevD.105.096027) (9 2022).
11. CERN Press Office. *CERN Brochure - LHC: The Guide* 2010.
12. The CMS Collaboration. *CMS, the Compact Muon Solenoid: Technical proposal* (CERN, Geneva, 1994).
13. The CMS Collaboration. The CMS experiment at the CERN LHC. *Journal of Instrumentation* **3**, S08004. doi:[10.1088/1748-0221/3/08/S08004](https://doi.org/10.1088/1748-0221/3/08/S08004) (2008).
14. Contardo, D., Klute, M., Mans, J., Silvestris, L. & Butler, J. *Technical Proposal for the Phase-II Upgrade of the CMS Detector* tech. rep. (CERN, Geneva, 2015). doi:[10.17181/CERN.VU8I.D59J](https://doi.org/10.17181/CERN.VU8I.D59J).
15. Karimäki, V., Mannelli, M., Siegrist, P., Breuker, H., Caner, A., Castaldi, R., Freudenreich, K., Hall, G., Horisberger, R., Huhtinen, M. & Cattai, A. *The CMS tracker system project: Technical Design Report* (CERN, Geneva, 1997).
16. The CMS Collaboration. *The CMS tracker: addendum to the Technical Design Report* (CERN, Geneva, 2000).
17. The Tracker Group of the CMS Collaboration. The CMS Phase-1 Pixel Detector Upgrade. *Journal of Instrumentation* **16**, P02027. doi:[10.1088/1748-0221/16/02/P02027](https://doi.org/10.1088/1748-0221/16/02/P02027). arXiv: [2012.14304](https://arxiv.org/abs/2012.14304) (2021).
18. The CMS Collaboration. *The Phase-2 Upgrade of the CMS Tracker* tech. rep. (CERN, Geneva, 2017). doi:[10.17181/CERN.QZ28.FLHW](https://doi.org/10.17181/CERN.QZ28.FLHW).
19. The CMS Collaboration. *Technical proposal for the upgrade of the CMS detector through 2020* tech. rep. (2011).
20. CMS Technical Design Report for the Pixel Detector Upgrade (eds Matzner Dominguez, D. A. *et al.*) doi:[10.2172/1151650](https://doi.org/10.2172/1151650) (2012).

Bibliographic references

21. The CMS Collaboration. Alignment of the CMS tracker with LHC and cosmic ray data. *Journal of Instrumentation* **9**, P06009. doi:[10.1088/1748-0221/9/06/P06009](https://doi.org/10.1088/1748-0221/9/06/P06009) (2014).
22. The CMS Collaboration. *The CMS tracking system: Technical Design Report* tech. rep. CERN-LHCC-98-006 (CERN, 1998).
23. The CMS Collaboration. Performance of the CMS tracking system with proton-proton collisions at $\sqrt{s} = 13$ TeV. *Journal of Instrumentation* **15**, P10003. doi:[10.1088/1748-0221/15/10/P10003](https://doi.org/10.1088/1748-0221/15/10/P10003) (2020).
24. The CMS Collaboration. Energy response and resolutions of the CMS electromagnetic calorimeter in LHC Run 2. *Journal of Instrumentation* **15**, P01020. doi:[10.1088/1748-0221/15/01/P01020](https://doi.org/10.1088/1748-0221/15/01/P01020) (2020).
25. The CMS Collaboration. *The CMS electromagnetic calorimeter: Technical Design Report* tech. rep. CERN-LHCC-97-33 (CERN, 1997).
26. The CMS Collaboration. *The CMS hadron calorimeter project: Technical Design Report* (CERN, Geneva, 1997).
27. Mans, J., Anderson, J., Dahmes, B., de Barbaro, P., Freeman, J., Grassi, T., Hazen, E., Mans, J., Ruchti, R., Schimdt, I., Shaw, T., Tully, C., Whitmore, J. & Yetkin, T. *CMS Technical Design Report for the Phase 1 Upgrade of the Hadron Calorimeter* tech. rep. (CERN, 2012).
28. The CMS Collaboration. Performance of the CMS Hadron Calorimeter with Cosmic Ray Muons and LHC Beam Data. *Journal of Instrumentation* **5**, T03013. doi:[10.1088/1748-0221/5/03/T03013](https://doi.org/10.1088/1748-0221/5/03/T03013) (2010).
29. The CMS Collaboration. Performance of the CMS Level-1 Trigger during Run-II at the LHC. *Journal of Instrumentation* **15**, P10017. doi:[10.1088/1748-0221/15/10/P10017](https://doi.org/10.1088/1748-0221/15/10/P10017) (2020).
30. The CMS Collaboration. Performance of the CMS Drift Tube Chambers in Proton-Proton Collisions at $\sqrt{s} = 13$ TeV. *Journal of Instrumentation* **16**, P11022. doi:[10.1088/1748-0221/16/11/P11022](https://doi.org/10.1088/1748-0221/16/11/P11022) (2021).

Bibliographic references

31. The CMS Collaboration. Alignment of the CMS Muon Drift Tube Chambers Using Cosmic-ray Muons. *Journal of Instrumentation* **9**, P06005. doi:[10.1088/1748-0221/9/06/P06005](https://doi.org/10.1088/1748-0221/9/06/P06005) (2014).
32. The CMS Collaboration. The Muon Project: Technical Design Report. *CERN-LHCC* **97-032** (1997).
33. The CMS Collaboration. Performance of the CMS cathode strip chambers with cosmic rays. *Journal of Instrumentation* **5**, T03018. doi:[10.1088/1748-0221/5/03/T03018](https://doi.org/10.1088/1748-0221/5/03/T03018) (2010).
34. Pozzobon, N. *The CMS Muon System performance during the LHC Run-2* tech. rep. (CERN, Geneva, 2019). doi:[10.1088/1748-0221/14/11/C11031](https://doi.org/10.1088/1748-0221/14/11/C11031).
35. Eysermans, J., Morales, M. I. P. & the CMS Collaboration. Operation and performance of the CMS Resistive Plate Chambers during LHC run II. *Journal of Physics: Conference Series* **912**, 012015. doi:[10.1088/1742-6596/912/1/012015](https://doi.org/10.1088/1742-6596/912/1/012015) (2017).
36. The CMS Collaboration. *CMS Technical Design Report for the Muon Endcap GEM Upgrade* tech. rep. (CERN, 2015).
37. Cittolin, S., Rácz, A. & Sphicas, P. *CMS The TriDAS Project: Technical Design Report, Volume 2: Data Acquisition and High-Level Trigger. CMS trigger and data-acquisition project* (CERN, Geneva, 2002).
38. Shiers, J. The Worldwide LHC Computing Grid (worldwide LCG). *Computer Physics Communications* **177**, 219–223. doi:<https://doi.org/10.1016/j.cpc.2007.02.021> (2007).
39. Bayatyan, G. L., Della Negra, M., Foà, Hervé, A. & Petrilli, A. *CMS computing: Technical Design Report* (CERN, Geneva, 2005).
40. Beaudette, F. The CMS Particle Flow Algorithm. *International Conference on Calorimetry for the High Energy Frontier*. arXiv: [1401.8155](https://arxiv.org/abs/1401.8155) (2013).
41. The CMS Collaboration. Alignment of the CMS silicon strip tracker during stand-alone commissioning. *Journal of Instrumentation* **4**, T07001. doi:[10.1088/1748-0221/4/07/T07001](https://doi.org/10.1088/1748-0221/4/07/T07001) (2009).

Bibliographic references

42. Karimaki, V., Lampen, T. & Schilling, F.-P. *The HIP Algorithm for Track Based Alignment and its Application to the CMS Pixel Detector* tech. rep. (CERN, Geneva, 2006).
43. The CMS Collaboration. Strategies and performance of the CMS silicon tracker alignment during LHC Run 2. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **1037**, 166795. doi:<https://doi.org/10.1016/j.nima.2022.166795> (2022).
44. Karimaki, V., Heikkinen, A., Lampen, T. & Linden, T. Sensor alignment by tracks. *eConf* **C0303241**, TULT008. arXiv: [physics/0306034](https://arxiv.org/abs/physics/0306034) (2003).
45. Blobel, V. Software alignment for tracking detectors. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **566**, 5–13. doi:<https://doi.org/10.1016/j.nima.2006.05.157> (2006).
46. Blobel, V. & Kleinwort, C. *A New Method for the High-Precision Alignment of Track Detectors* 2002. arXiv: [hep-ex/0208021](https://arxiv.org/abs/hep-ex/0208021) [[hep-ex](#)].
47. The CMS Collaboration. Alignment of the CMS silicon tracker during commissioning with cosmic rays. *Journal of Instrumentation* **5**, T03009. doi:[10.1088/1748-0221/5/03/T03009](https://doi.org/10.1088/1748-0221/5/03/T03009) (2010).
48. Brown, D., Gritsan, A., Guo, Z. & Roberts, D. Local alignment of the BaBar Silicon Vertex Tracking detector. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **603**, 467–484. doi:<https://doi.org/10.1016/j.nima.2009.02.001> (2009).
49. The CMS Collaboration. *Displays of cosmic ray events as seen in the CMS detector in March 2022* 2022.
50. Glashow, S. L. Partial-symmetries of weak interactions. *Nuclear Physics* **22**, 579. doi:[10.1016/0029-5582\(61\)90469-2](https://doi.org/10.1016/0029-5582(61)90469-2) (1961).
51. Englert, F. & Brout, R. Broken symmetry and the mass of gauge vector mesons. *Physical Review Letters* **13**, 321. doi:[10.1103/PhysRevLett.13.321](https://doi.org/10.1103/PhysRevLett.13.321) (1964).

Bibliographic references

- 52. Higgs, P. W. Broken symmetries, massless particles and gauge fields. *Physics Letters* **12**, 132. doi:[10.1016/0031-9163\(64\)91136-9](https://doi.org/10.1016/0031-9163(64)91136-9) (1964).
- 53. Higgs, P. W. Broken symmetries and the masses of gauge bosons. *Physical Review Letters* **13**, 508. doi:[10.1103/PhysRevLett.13.508](https://doi.org/10.1103/PhysRevLett.13.508) (1964).
- 54. Guralnik, G. S., Hagen, C. R. & Kibble, T. W. B. Global conservation laws and massless particles. *Physical Review Letters* **13**, 585. doi:[10.1103/PhysRevLett.13.585](https://doi.org/10.1103/PhysRevLett.13.585) (1964).
- 55. Weinberg, S. A model of leptons. *Physical Review Letters* **19**, 1264. doi:[10.1103/PhysRevLett.19.1264](https://doi.org/10.1103/PhysRevLett.19.1264) (1967).
- 56. Salam, A. *Weak and electromagnetic interactions in Elementary particle physics: relativistic groups and analyticity* (ed Svartholm, N.) (Almqvist & Wiksell, Stockholm, 1968), 367.
- 57. The CMS Collaboration. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Physics Letters B* **716**, 30–61. doi:[10.1016/j.physletb.2012.08.021](https://doi.org/10.1016/j.physletb.2012.08.021) (2012).
- 58. The CMS Collaboration. Observation of a new boson with mass near 125 GeV in pp collisions at $\sqrt{s} = 7$ and 8 TeV. *Journal of High Energy Physics* **06**, 081. doi:[10.1007/JHEP06\(2013\)081](https://doi.org/10.1007/JHEP06(2013)081). arXiv: [1303.4571](https://arxiv.org/abs/1303.4571) [[hep-ex](#)] (2013).
- 59. The ATLAS Collaboration. A detailed map of Higgs boson interactions by the ATLAS experiment ten years after the discovery. *Nature* **607**, 52. doi:[10.1038/s41586-022-04893-w](https://doi.org/10.1038/s41586-022-04893-w) (2022).
- 60. The CMS Collaboration. A portrait of the Higgs boson by the CMS experiment ten years after the discovery. *Nature* **607**, 60. doi:[10.1038/s41586-022-04892-x](https://doi.org/10.1038/s41586-022-04892-x) (2022).
- 61. The ATLAS and CMS Collaborations. Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments. *Physical Review Letters* **114**, 191803. doi:[10.1103/PhysRevLett.114.191803](https://doi.org/10.1103/PhysRevLett.114.191803). arXiv: [1503.07589](https://arxiv.org/abs/1503.07589) [[hep-ex](#)] (2015).
- 62. The ATLAS Collaboration. Combined measurement of the Higgs boson mass from the $H \rightarrow \gamma\gamma$ and $H \rightarrow ZZ^* \rightarrow 4\ell$ decay channels with the ATLAS detector

- using $\sqrt{s} = 7, 8$ and 13 TeV pp collision data. *Physical Review Letters* **131**, 251802. doi:[10.1103/PhysRevLett.131.251802](https://doi.org/10.1103/PhysRevLett.131.251802). arXiv: [2308.04775](https://arxiv.org/abs/2308.04775) [[hep-ex](#)] (2023).
63. The CMS Collaboration. Precise determination of the mass of the Higgs boson and tests of compatibility of its couplings with the standard model predictions using proton collisions at 7 and 8 TeV. *European Physical Journal C* **75**, 212. doi:[10.1140/epjc/s10052-015-3351-7](https://doi.org/10.1140/epjc/s10052-015-3351-7). arXiv: [1412.8662](https://arxiv.org/abs/1412.8662) [[hep-ex](#)] (2015).
64. The CMS Collaboration. Measurements of properties of the Higgs boson decaying into the four-lepton final state in pp collisions at $\sqrt{s} = 13$ TeV. *Journal of High Energy Physics* **11**, 047. doi:[10.1007/JHEP11\(2017\)047](https://doi.org/10.1007/JHEP11(2017)047). arXiv: [1706.09936](https://arxiv.org/abs/1706.09936) [[hep-ex](#)] (2017).
65. The ATLAS Collaboration. Constraints on the off-shell Higgs boson signal strength in the high-mass ZZ and WW final states with the ATLAS detector. *European Physical Journal C* **75**, 335. doi:[10.1140/epjc/s10052-015-3542-2](https://doi.org/10.1140/epjc/s10052-015-3542-2). arXiv: [1503.01060](https://arxiv.org/abs/1503.01060) [[hep-ex](#)] (2015).
66. The CMS Collaboration. Limits on the Higgs boson lifetime and width from its decay to four charged leptons. *Physical Review D* **92**, 072010. doi:[10.1103/PhysRevD.92.072010](https://doi.org/10.1103/PhysRevD.92.072010). arXiv: [1507.06656](https://arxiv.org/abs/1507.06656) [[hep-ex](#)] (2015).
67. The CMS Collaboration. Search for Higgs boson off-shell production in proton-proton collisions at 7 and 8 TeV and derivation of constraints on its total decay width. *Journal of High Energy Physics* **09**, 051. doi:[10.1007/JHEP09\(2016\)051](https://doi.org/10.1007/JHEP09(2016)051). arXiv: [1605.02329](https://arxiv.org/abs/1605.02329) [[hep-ex](#)] (2016).
68. The ATLAS Collaboration. Constraints on off-shell Higgs boson production and the Higgs boson total width in $ZZ \rightarrow 4\ell$ and $ZZ \rightarrow 2\ell 2\nu$ final states with the ATLAS detector. *Physics Letters B* **786**, 223. doi:[10.1016/j.physletb.2018.09.048](https://doi.org/10.1016/j.physletb.2018.09.048). arXiv: [1808.01191](https://arxiv.org/abs/1808.01191) [[hep-ex](#)] (2018).
69. The CMS Collaboration. Measurement of the Higgs boson width and evidence of its off-shell contributions to ZZ production. *Nature Physics* **18**, 1329. doi:[10.1038/s41567-022-01682-0](https://doi.org/10.1038/s41567-022-01682-0). arXiv: [2202.06923](https://arxiv.org/abs/2202.06923) [[hep-ex](#)] (2022).
70. Kauer, N. & Passarino, G. Inadequacy of zero-width approximation for a light Higgs boson signal. *Journal of High Energy Physics* **08**, 116. doi:[10.1007/JHEP08\(2012\)116](https://doi.org/10.1007/JHEP08(2012)116). arXiv: [1206.4803](https://arxiv.org/abs/1206.4803) [[hep-ph](#)] (2012).

71. Campbell, J. M., Ellis, R. K. & Williams, C. Bounding the Higgs width at the LHC using full analytic results for $gg \rightarrow e^-e^+\mu^-\mu^+$. *Journal of High Energy Physics* **04**, 060. doi:[10.1007/JHEP04\(2014\)060](https://doi.org/10.1007/JHEP04(2014)060). arXiv: [1311.3589](https://arxiv.org/abs/1311.3589) [[hep-ph](#)] (2014).
72. The ATLAS Collaboration. Evidence of off-shell Higgs boson production from ZZ leptonic decay channels and constraints on its total width with the ATLAS detector. *Physics Letters B* **846**, 138223. doi:[10.1016/j.physletb.2023.138223](https://doi.org/10.1016/j.physletb.2023.138223). arXiv: [2304.01532](https://arxiv.org/abs/2304.01532) [[hep-ex](#)] (2023).
73. Gao, Y., Gritsan, A. V., Guo, Z., Melnikov, K., Schulze, M. & Tran, N. V. Spin determination of single-produced resonances at hadron colliders. *Physical Review D* **81**. doi:[10.1103/physrevd.81.075022](https://doi.org/10.1103/physrevd.81.075022) (2010).
74. Bolognesi, S., Gao, Y., Gritsan, A. V., Melnikov, K., Schulze, M., Tran, N. V. & Whitbeck, A. Spin and parity of a single-produced resonance at the LHC. *Physical Review D* **86**. doi:[10.1103/physrevd.86.095031](https://doi.org/10.1103/physrevd.86.095031) (2012).
75. Anderson, I., Bolognesi, S., Caola, F., Gao, Y., Gritsan, A. V., Martin, C. B., Melnikov, K., Schulze, M., Tran, N. V., Whitbeck, A. & Zhou, Y. Constraining anomalous HVV interactions at proton and lepton colliders. *Physical Review D* **89**. doi:[10.1103/physrevd.89.035007](https://doi.org/10.1103/physrevd.89.035007) (2014).
76. Gritsan, A. V., Röntsch, R., Schulze, M. & Xiao, M. Constraining anomalous Higgs boson couplings to the heavy-flavor fermions using matrix element techniques. *Physical Review D* **94**. doi:[10.1103/physrevd.94.055023](https://doi.org/10.1103/physrevd.94.055023) (2016).
77. Martini, T., Pan, R.-Q., Schulze, M. & Xiao, M. Probing the CP structure of the top quark Yukawa coupling: Loop sensitivity versus on-shell sensitivity. *Physical Review D* **104**. doi:[10.1103/physrevd.104.055045](https://doi.org/10.1103/physrevd.104.055045) (2021).
78. Campbell, J. M. & Ellis, R. MCFM for the Tevatron and the LHC. *Nuclear Physics B - Proceedings Supplements* **205–206**, 10–15. doi:[10.1016/j.nuclphysbps.2010.08.011](https://doi.org/10.1016/j.nuclphysbps.2010.08.011) (2010).
79. The CMS Collaboration. Measurements of production cross sections of the Higgs boson in the four-lepton final state in proton–proton collisions at $\sqrt{s} = 13$ TeV. *European Physical Journal C* **81**, 488. doi:[10.1140/epjc/s10052-021-09200-x](https://doi.org/10.1140/epjc/s10052-021-09200-x). arXiv: [2103.04956](https://arxiv.org/abs/2103.04956) [[hep-ex](#)] (2021).

- 80. The CMS Collaboration. Constraints on anomalous Higgs boson couplings to vector bosons and fermions in its production and decay using the four-lepton final state. *Physical Review D* **104**. doi:[10.1103/physrevd.104.052004](https://doi.org/10.1103/physrevd.104.052004) (2021).
- 81. Catani, S. & Grazzini, M. An NNLO subtraction formalism in hadron collisions and its application to Higgs boson production at the LHC. *Physical Review Letters* **98**, 222002. doi:[10.1103/PhysRevLett.98.222002](https://doi.org/10.1103/PhysRevLett.98.222002). arXiv: [hep-ph/0703012](https://arxiv.org/abs/hep-ph/0703012) [[hep-ph](#)] (2007).
- 82. Grazzini, M. NNLO predictions for the Higgs boson signal in the $H \rightarrow WW \rightarrow \ell\nu\ell\nu$ and $H \rightarrow ZZ \rightarrow 4\ell$ decay channels. *Journal of High Energy Physics* **02**, 043. doi:[10.1088/1126-6708/2008/02/043](https://doi.org/10.1088/1126-6708/2008/02/043). arXiv: [0801.3232](https://arxiv.org/abs/0801.3232) [[hep-ph](#)] (2008).
- 83. Grazzini, M. & Sargsyan, H. Heavy-quark mass effects in Higgs boson production at the LHC. *Journal of High Energy Physics* **09**, 129. doi:[10.1007/JHEP09\(2013\)129](https://doi.org/10.1007/JHEP09(2013)129). arXiv: [1306.4581](https://arxiv.org/abs/1306.4581) [[hep-ph](#)] (2013).
- 84. Caola, F., Melnikov, K., Röntsch, R. & Tancredi, L. QCD corrections to ZZ production in gluon fusion at the LHC. *Physical Review D* **92**, 094028. doi:[10.1103/PhysRevD.92.094028](https://doi.org/10.1103/PhysRevD.92.094028). arXiv: [1509.06734](https://arxiv.org/abs/1509.06734) [[hep-ph](#)] (2015).
- 85. Melnikov, K. & Dowling, M. Production of two Z-bosons in gluon fusion in the heavy top quark approximation. *Physics Letters B* **744**, 43. doi:[10.1016/j.physletb.2015.03.030](https://doi.org/10.1016/j.physletb.2015.03.030). arXiv: [1503.01274](https://arxiv.org/abs/1503.01274) [[hep-ph](#)] (2015).
- 86. Campbell, J. M., Ellis, R. K., Czakon, M. & Kirchner, S. Two loop correction to interference in $gg \rightarrow ZZ$. *Journal of High Energy Physics* **08**, 011. doi:[10.1007/JHEP08\(2016\)011](https://doi.org/10.1007/JHEP08(2016)011). arXiv: [1605.01380](https://arxiv.org/abs/1605.01380) [[hep-ph](#)] (2016).
- 87. Caola, F., Dowling, M., Melnikov, K., Röntsch, R. & Tancredi, L. QCD corrections to vector boson pair production in gluon fusion including interference effects with off-shell Higgs at the LHC. *Journal of High Energy Physics* **07**, 087. doi:[10.1007/JHEP07\(2016\)087](https://doi.org/10.1007/JHEP07(2016)087). arXiv: [1605.04610](https://arxiv.org/abs/1605.04610) [[hep-ph](#)] (2016).
- 88. Particle Data Group, Zyla, P. A., *et al.* Review of particle physics. *Progress of Theoretical and Experimental Physics* **2020**, 083C01. doi:[10.1093/ptep/ptaa104](https://doi.org/10.1093/ptep/ptaa104) (2020).

89. Grazzini, M., Kallweit, S. & Rathlev, D. ZZ production at the LHC: Fiducial cross sections and distributions in NNLO QCD. *Physics Letters B* **750**, 407–410. doi:[10.1016/j.physletb.2015.09.055](https://doi.org/10.1016/j.physletb.2015.09.055) (2015).
90. Bierweiler, A., Kasprzik, T. & Kühn, J. H. Vector-boson pair production at the LHC to $\mathcal{O}(\alpha^3)$ accuracy. *Journal of High Energy Physics* **12**, 071. doi:[10.1007/JHEP12\(2013\)071](https://doi.org/10.1007/JHEP12(2013)071). arXiv: [1305.5402](https://arxiv.org/abs/1305.5402) [[hep-ph](#)] (2013).
91. Baglio, J., Ninh, L. D. & Weber, M. M. Massive gauge boson pair production at the LHC: a next-to-leading order story. *Physical Review D* **88**, 113005. doi:[10.1103/PhysRevD.88.113005](https://doi.org/10.1103/PhysRevD.88.113005). arXiv: [1307.4331](https://arxiv.org/abs/1307.4331) [[hep-ph](#)] (2013).
92. Gieseke, S., Kasprzik, T. & Kühn, J. H. *Vector-boson pair production and electroweak corrections in HERWIG++* 2014. arXiv: [1401.3964](https://arxiv.org/abs/1401.3964) [[hep-ph](#)].
93. The CMS Collaboration. Particle-flow reconstruction and global event description with the CMS detector. *Journal of Instrumentation* **12**, P10003. doi:[10.1088/1748-0221/12/10/P10003](https://doi.org/10.1088/1748-0221/12/10/P10003). arXiv: [1706.04965](https://arxiv.org/abs/1706.04965) [[physics.ins-det](#)] (2017).
94. The CMS Collaboration. Electron and photon reconstruction and identification with the CMS experiment at the CERN LHC. *Journal of Instrumentation* **16**, P05014. doi:[10.1088/1748-0221/16/05/p05014](https://doi.org/10.1088/1748-0221/16/05/p05014). eprint: [2012.06888](https://arxiv.org/abs/2012.06888) (2021).
95. The CMS Collaboration. Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}=13$ TeV. *Journal of Instrumentation* **13**, P06015. doi:[10.1088/1748-0221/13/06/p06015](https://doi.org/10.1088/1748-0221/13/06/p06015). eprint: [1804.04528](https://arxiv.org/abs/1804.04528) (2018).
96. The CMS Collaboration. Pileup mitigation at CMS in 13 TeV data. *Journal of Instrumentation* **15**, P09018. doi:[10.1088/1748-0221/15/09/P09018](https://doi.org/10.1088/1748-0221/15/09/P09018). arXiv: [2003.00503](https://arxiv.org/abs/2003.00503) [[hep-ex](#)] (2020).
97. The CMS Collaboration. Measurement of the Inclusive W and Z Production Cross Sections in pp Collisions at $\sqrt{s}=7$ TeV. *Journal of High Energy Physics* **10**, 132. doi:[10.1007/JHEP10\(2011\)132](https://doi.org/10.1007/JHEP10(2011)132). arXiv: [1107.4789](https://arxiv.org/abs/1107.4789) [[hep-ex](#)] (2011).
98. The CMS Collaboration. *ECAL 2016 refined calibration and Run 2 summary plots* CMS Detector Performance Summary CMS-DP-2020-021 (CERN, 2020).

Bibliographic references

- 99. Bodek, A., van Dyne, A., Han, J. Y., Sakumoto, W. & Strelnikov, A. Extracting muon momentum scale corrections for hadron collider experiments. *European Physical Journal C* **72**. doi:[10.1140/epjc/s10052-012-2194-8](https://doi.org/10.1140/epjc/s10052-012-2194-8) (2012).
- 100. The CMS Collaboration. Measurement of the Higgs boson mass and width using the four-lepton final state in proton-proton collisions at $\sqrt{s} = 13$ TeV. *Physical Review D* **111**, 092014. doi:[10.1103/PhysRevD.111.092014](https://doi.org/10.1103/PhysRevD.111.092014) (9 2025).
- 101. Cacciari, M., Salam, G. P. & Soyez, G. The anti- k_t jet clustering algorithm. *Journal of High Energy Physics* **04**, 063. doi:[10.1088/1126-6708/2008/04/063](https://doi.org/10.1088/1126-6708/2008/04/063). arXiv: [0802.1189](https://arxiv.org/abs/0802.1189) [[hep-ex](#)] (2008).
- 102. Cacciari, M., Salam, G. P. & Soyez, G. FastJet user manual. *European Physical Journal C* **72**, 1896. doi:[10.1140/epjc/s10052-012-1896-2](https://doi.org/10.1140/epjc/s10052-012-1896-2). arXiv: [1111.6097](https://arxiv.org/abs/1111.6097) [[hep-ph](#)] (2012).
- 103. The CMS Collaboration. Identification of heavy-flavour jets with the CMS detector in pp collisions at 13 TeV. *Journal of Instrumentation* **13**, P05011. doi:[10.1088/1748-0221/13/05/P05011](https://doi.org/10.1088/1748-0221/13/05/P05011). arXiv: [1712.07158](https://arxiv.org/abs/1712.07158) [[physics.ins-det](#)] (2018).
- 104. The CMS Collaboration. Constraints on anomalous Higgs boson couplings using production and decay information in the four-lepton final state. *Physics Letters B* **775**, 1. doi:[10.1016/j.physletb.2017.10.021](https://doi.org/10.1016/j.physletb.2017.10.021). arXiv: [1707.00541](https://arxiv.org/abs/1707.00541) [[hep-ex](#)] (2017).
- 105. The CMS Collaboration. A measurement of the Higgs boson mass in the diphoton decay channel. *Physics Letters B* **805**, 135425. doi:[10.1016/j.physletb.2020.135425](https://doi.org/10.1016/j.physletb.2020.135425). arXiv: [2002.06398](https://arxiv.org/abs/2002.06398) [[hep-ex](#)] (2020).
- 106. The CMS Collaboration. Precision luminosity measurement in proton-proton collisions at $\sqrt{s} = 13$ TeV in 2015 and 2016 at CMS. *European Physical Journal C* **81**, 800. doi:[10.1140/epjc/s10052-021-09538-2](https://doi.org/10.1140/epjc/s10052-021-09538-2). arXiv: [2104.01927](https://arxiv.org/abs/2104.01927) [[hep-ex](#)] (2021).
- 107. The CMS Collaboration. *CMS luminosity measurement for the 2017 data-taking period at $\sqrt{s} = 13$ TeV* CMS Physics Analysis Summary CMS-PAS-LUM-17-004 (CERN, 2018).

Bibliographic references

108. The CMS Collaboration. *CMS luminosity measurement for the 2018 data-taking period at $\sqrt{s} = 13$ TeV* CMS Physics Analysis Summary CMS-PAS-LUM-18-002 (CERN, 2019).
109. The CMS Collaboration. The CMS Statistical Analysis and Combination Tool: COMBINE. *Computing and Software for Big Science* **8**, 19. doi:[10.1007/s41781-024-00121-4](https://doi.org/10.1007/s41781-024-00121-4) (2024).
110. Wilks, S. S. The large-sample distribution of the likelihood ratio for testing composite hypotheses. *Annals Math. Statist.* **9**, 60. doi:[10.1214/aoms/1177732360](https://doi.org/10.1214/aoms/1177732360) (1938).