Springer Theses
Recognizing Outstanding Ph.D. Research

Michael Andrews

Search for Exotic Higgs Boson Decays to Merged Diphotons

A Novel CMS Analysis Using End-to-End Deep Learning

Springer

Springer Theses

Recognizing Outstanding Ph.D. Research

Aims and Scope

The series "Springer Theses" brings together a selection of the very best Ph.D. theses from around the world and across the physical sciences. Nominated and endorsed by two recognized specialists, each published volume has been selected for its scientific excellence and the high impact of its contents for the pertinent field of research. For greater accessibility to non-specialists, the published versions include an extended introduction, as well as a foreword by the student's supervisor explaining the special relevance of the work for the field. As a whole, the series will provide a valuable resource both for newcomers to the research fields described, and for other scientists seeking detailed background information on special questions. Finally, it provides an accredited documentation of the valuable contributions made by today's younger generation of scientists.

Theses may be nominated for publication in this series by heads of department at internationally leading universities or institutes and should fulfill all of the following criteria

- They must be written in good English.
- The topic should fall within the confines of Chemistry, Physics, Earth Sciences, Engineering and related interdisciplinary fields such as Materials, Nanoscience, Chemical Engineering, Complex Systems and Biophysics.
- The work reported in the thesis must represent a significant scientific advance.
- If the thesis includes previously published material, permission to reproduce this must be gained from the respective copyright holder (a maximum 30% of the thesis should be a verbatim reproduction from the author's previous publications).
- They must have been examined and passed during the 12 months prior to nomination.
- Each thesis should include a foreword by the supervisor outlining the significance of its content.
- The theses should have a clearly defined structure including an introduction accessible to new PhD students and scientists not expert in the relevant field.

Indexed by zbMATH.

Michael Andrews

Search for Exotic Higgs Boson Decays to Merged Diphotons

A Novel CMS Analysis Using End-to-End Deep Learning

Doctoral Thesis accepted by
Carnegie Mellon University, Pittsburgh, USA

Springer

Author
Dr. Michael Andrews
Department of Physics
Carnegie Mellon University
Pittsburgh, PA, USA

Supervisor
Dr. Manfred Paulini
Professor of Physics
Carnegie Mellon University
Pittsburgh, PA, USA

ISSN 2190-5053 ISSN 2190-5061 (electronic)
Springer Theses
ISBN 978-3-031-25090-3 ISBN 978-3-031-25091-0 (eBook)
https://doi.org/10.1007/978-3-031-25091-0

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Supervisor's Foreword

The theory describing the smallest building blocks of matter and the forces acting between them is called the standard model of particle physics. It is an enormously successful theory describing the interactions of all known elementary particles, the quarks, leptons and gauge bosons acting as force mediators. Developed over the past 60 years, starting with the quark model in the 1960s, the discovery of the charm quark in 1974, the τ lepton seen in experiments from 1974 to 1977, the bottom quark in 1977, the W and Z bosons in 1983, the top quark in 1995, and culminating in the discovery of the Higgs boson in 2012, there is to date no experimental evidence contradicting the predictions of the standard model. Although it is successful in describing all phenomena at the subatomic scales, it is not a complete "theory of everything" that can explain all known observations. For example, no particle exists in the standard model that constitutes a possible candidate for the dark matter that makes up about one quarter of the energy–mass content of the universe. The quest for finding phenomena that are not described by the standard model is one reason why physicists at the CERN Large Hadron Collider (LHC) are searching for yet-unknown particles, which can pave the way to postulate theories beyond the standard model.

The Ph.D. research conducted by Dr. Michael Andrews under my supervision in the Department of Physics at Carnegie Mellon University using proton–proton collision data collected with the Compact Muon Solenoid (CMS) experiment at the LHC is not just another search for phenomena beyond the standard model. What sets the data analysis in Dr. Andrews' thesis apart from conventional CMS searches is the use of several innovative approaches and "firsts" for CMS. It is also a story about the beauty of being a professor that allows you to learn together with and from your students. Let's go back in time to better understand…

About five years I got interested in the application of modern machine learning (ML) techniques in particle physics. Somehow I had ignored ML for a long time given that we had been using neural networks in particle physics for over 30 years. Together with Dr. Andrews, I learned very quickly that recent ML advances, in particular in the field of computer vision, have led to breakthrough applications of convolutional neural networks to scientific challenges, if the data can be expressed as an image or series of images. In particular, we became interested in exploring whether ML can help to get beyond limitations of traditional analysis techniques. As a first project, Dr. Andrews' work demonstrated the application of image-based deep learning techniques to separate electron from photon showers in one of the CMS sub-detectors, the electromagnetic calorimeter, a task that is not achievable with conventional approaches. This brought us to establish the concept of an end-to-end event classification that directly leverages low-level detector data as input to classify event signatures such as using images from low-level detector data to go directly to classify an event signature without using data reconstruction.

Fueled by the initial success, Dr. Andrews became quite involved in ML and very quickly an expert in the usage of different deep learning networks and ML techniques. His thesis analysis follows the path of exploring what is the maximum information that can be extracted from detector data when modern ML approaches are unleashed. He studied the hypothetical decay of the Higgs boson into a pair of light particles $\mathrm{H} \rightarrow \mathcal{A}\mathcal{A}$, each of which may in turn decay into a pair of photons $\mathcal{A} \rightarrow \gamma\gamma$. The branching fraction for $\mathcal{A} \rightarrow \gamma\gamma$ is maximized at light $m_{\mathcal{A}}$ masses, but in this regime, each of the $\mathcal{A} \rightarrow \gamma\gamma$ decays is highly merged, and the diphotons are reconstructed as a single photon shower in the CMS electromagnetic calorimeter consisting of lead-tungstate crystals. Using end-to-end ML techniques, Dr. Andrews was able to develop a mass regression algorithm that maintains sensitivity even in the limit, where the two photons from the $\mathcal{A} \rightarrow \gamma\gamma$ system deposit their energy in the same calorimeter crystal. On the way to setting the first CMS limit for the theoretically highly interesting mass regime $m_{\mathcal{A}} < 200$ MeV, Dr. Andrews solved several issues with sensitivity toward the $m_{\mathcal{A}} \rightarrow 0$ mass endpoint that I leave for the interested reader to discover in his thesis entitled "Search for exotic Higgs boson decays to merged photons employing a novel deep learning technique at CMS".

This well-written and nicely organized Ph.D. thesis contains very accessible introductions for the novice to particle physics but also allows the expert to find useful new information. For example, Chap. 2 is an engaging introduction to the LHC and the CMS detector that should be accessible for a reader less familiar with particle physics, while Chaps. 7 and 8 detail the mass regression method and the data analysis for the experts. There is something for everyone in this thesis.

Finally, let me conclude by expressing my appreciation for the Ph.D. thesis work of Dr. Michael Andrews. I feel honored and grateful about the opportunity that I had working with him and learning from him.

Pittsburgh, USA Dr. Manfred Paulini
December 2022

Abstract

A search for exotic Higgs boson decays, of the form H$\rightarrow$ aa, with a $\rightarrow$ $\gamma\gamma$, is performed. The hypothetical particle a is a light, scalar or pseudoscalar particle decaying to two highly merged photons reconstructed as a single photon-like object in the CMS detector. A novel, end-to-end deep learning-based technique is developed to directly measure the invariant mass of merged a $\rightarrow$ $\gamma\gamma$ candidates for the first time at CMS. Analysis criteria similar to those used in the standard model H$\rightarrow$ $\gamma\gamma$ search are applied, to probe the possibility that existing measurements in this decay mode may conceal a contribution from a low-mass particle a. The search is performed using the full CMS Run II data set, corresponding to a total integrated luminosity of 136 fb^{-1}, at a proton–proton center-of-mass collision energy of $\sqrt{s} = 13$ TeV. No significant excess of events over standard model expectations is found. Branching fractions for this process of $\mathcal{B}(\mathrm{H} \rightarrow \mathrm{aa} \rightarrow 4\gamma) \gtrsim 0.9$–$3.3 \times 10^{-3}$ are excluded at 95% confidence level, for particle masses between $0.1 \leq m_a \leq 1.2$ GeV, assuming negligible lifetime.

Acknowledgments

I would like to express my profound gratitude to my adviser, Manfred Paulini, for his enduring and unwavering support over an admittedly longer-than-usual Ph.D. program. His faith and support have been instrumental in allowing me to explore riskier ideas that, while in retrospect have proven career-defining, could have easily been nipped at the bud during their earlier, more fragile days.

To the ECAL community for entrusting their cherished detector to my hands during Run II. Being ECAL run-coordinator has been the professional thrill of a lifetime. I can only hope that I have given back to them as much as they have given to me.

To Sergei Gleyzer, for his partnership in developing end-to-end ML. To my earlier advisers, Michael Wang, who gave me my big break into the world of high energy physics, and Neil Caranto, for prodding me to pursue physics when I was still a business graduate.

Finally, to my family for their love, moral (and financial) support, and to my friends at different stages of my Ph.D., for their company and day-to-day emotional support, wherever they may be today.

Contents

Chapter 1
Introduction

In 2012, a new boson was observed by the CMS and ATLAS experiments [1, 2] operating at the CERN Large Hadron Collider (LHC) with properties consistent with the standard model (SM) Higgs boson decaying to $\mathrm{H} \rightarrow \mathrm{ZZ}^{*} \rightarrow 4\ell$ and $\mathrm{H} \rightarrow \gamma\gamma$. Since then, additional decay modes have been observed, building confidence that the new boson is, in fact, the SM Higgs boson [3, 4], capping off a major puzzle in the origin of electroweak symmetry breaking and particle mass. Indeed, the results of the broader CMS search program of recent years suggest that the physics probed by the LHC is just as predicted by the SM. Yet, astronomical observations and theoretical inconsistencies [5] make it clear that the SM cannot be the final theory of particle physics. With the LHC now collecting unprecedented amounts of data, this has prompted a number of searches for physics beyond the standard model (BSM) that venture farther out into unexplored corners of phase space, where they may have been overlooked by more conventional search strategies.

Even in the light of current LHC constraints, the Higgs sector remains an important search space for BSM physics, due to its accessibility to SM-neutral hidden sectors. In such scenarios, because of the small decay width of the SM Higgs boson, even minute couplings to BSM physics can lead to sizeable branching fractions for exotic new states that may be accessible at the LHC. With current constraints on $\mathrm{H} \rightarrow \mathrm{BSM} \approx 20-60\%$ [6], depending on assumptions, this leaves much room for exploration in the exotic Higgs sector.

At the same time, recent advances in analysis tools, particularly those based on advanced machine learning (ML) or so-called deep learning, have empowered the pursuit of experimentally challenging topologies, which were theoretically well-motivated but simply not feasible to pursue previously. A prime example, which is the focus of this thesis, is the exotic decay of the Higgs boson to a pair of light scalars, each subsequently decaying to two photons, $\mathrm{H} \rightarrow \mathrm{aa}$ with $\mathrm{a} \rightarrow \gamma\gamma$ [7], or $\mathrm{H} \rightarrow \mathrm{aa} \rightarrow 4\gamma$ for short. Not all applications of ML, however, lead to breakthroughs. Historically, the LHC experiments used highly processed inputs representing physically meaningful quantities like particle 4-momenta to train ML algorithms. However, for many new physics searches, sensitivity is limited not by the ability to extract useful information from particle 4-momenta but by inefficiencies in the reconstruction of the 4-momenta

M. Andrews, *Search for Exotic Higgs Boson Decays to Merged Diphotons*, Springer Theses, https://doi.org/10.1007/978-3-031-25091-0_1

quantities themselves. In this thesis, we describe the first LHC physics analysis that trains advanced ML algorithms directly on low-level, "raw" detector data to address this bottleneck. As we will show, a direct probe of the $\mathrm{H} \rightarrow \mathrm{aa} \rightarrow 4\gamma$ signal at CMS is impossible without end-to-end deep learning. Thus, for suitable applications, end-to-end deep learning delivers breakthrough sensitivity and extends our physics reach to entirely new domains.

Decays like those of $\mathrm{H} \rightarrow \mathrm{aa} \rightarrow 4\gamma$ are candidates in various BSM models involving Minimal Composite Higgs Models (MCHM), two-Higgs-double-like models (2HDM), Next-to-Minimal Supersymmetric Standard Model (NMSSM), and any SM extension involving an additional hidden sector coupling to a new singlet [8, 9]. Moreover, such decays are of particular interest in searches for axion-like particle (ALP) production [10–13] because of their potential impact on our understanding of the early universe and stellar formation. ALPs are also an important potential candidate for dark matter [14–17]. In astrophysical and cosmological searches, the particle a is additionally identified as a spin-0 or CP-odd particle, known as a pseudoscalar. The experimental search we present, however, is insensitive to the CP quantum numbers of a, since its polarization is not measured.

While different model assumptions allow for varying $\mathrm{a} \rightarrow \gamma\gamma$ branching fractions, the $\mathrm{a} \rightarrow \gamma\gamma$ decay mode is generally enhanced when m_{a} is less than the pair production threshold for decays to the heavier SM states [8]. For masses below the charmonium production threshold ($m_{\mathrm{a}} \lesssim 3\,\mathrm{GeV}$), the particle a will be increasingly preferred to be long-lived [8]. If the a decays outside of the detector volume, it will not be reconstructed at all. Moreover, even if the a decays promptly, if it arise from $\mathrm{H} \rightarrow \mathrm{aa}$, the $\mathrm{a} \rightarrow \gamma\gamma$ photons will be highly collimated. Each $\mathrm{a} \rightarrow \gamma\gamma$ will thus be misreconstructed as a single photon-like object $\Gamma(\mathrm{a} \rightarrow \gamma\gamma)$, or Γ for short, by existing particle reconstruction algorithms. In this scenario, the $\mathrm{H} \rightarrow \mathrm{aa} \rightarrow 4\gamma$ decay will present an invariant mass resonance approximately degenerate with that of the SM $\mathrm{H} \rightarrow \gamma\gamma$ decay [18]. Therefore, if realized in nature, the low-m_{a} $\mathrm{H} \rightarrow \mathrm{aa} \rightarrow 4\gamma$ signal will be buried in existing events resembling SM $\mathrm{H} \rightarrow \gamma\gamma$ decays [6, 19].

Motivated by these challenges and opportunities, in this thesis, we present the first $\mathrm{H} \rightarrow \mathrm{aa} \rightarrow 4\gamma$ search that directly measures the invariant mass spectrum of merged photon candidates Γ reconstructed in events resembling a SM $\mathrm{H} \rightarrow \gamma\gamma$ final state. That is, the search is performed in the experimentally challenging regime where the $\mathrm{a} \rightarrow \gamma\gamma$ decays are merged, but where the branching fraction for this decay mode is most theoretically attractive. The analysis is made possible by the development of a novel particle reconstruction technique, which we likewise describe in this thesis. The technique utilizes an end-to-end deep learning strategy to reconstruct the invariant mass of merged photon candidates, m_{Γ}, directly from the energy deposits in the CMS electromagnetic calorimeter. The full CMS Run II data set is used, corresponding to an integrated luminosity of $136\ \mathrm{fb}^{-1}$. We probe $\mathrm{H} \rightarrow \mathrm{aa} \rightarrow 4\gamma$ decays with particle a masses in the range range $m_{\mathrm{a}} = 0.1$–1.2 GeV. In this first analysis, we assume that the as decay promptly and analyze only $\mathrm{a} \rightarrow \gamma\gamma$ candidates reconstructed in the barrel section of the detector, for simplicity.

While a number of ATLAS measurements have performed similar searches [18, 20], this analysis represents the first attempt at the LHC to directly probe the $\mathrm{a} \rightarrow \gamma\gamma$

invariant mass spectrum. A number of other CMS analyses have been published [21–25], or are underway, to either directly or indirectly search for particle a decays to other states $\mathrm{a} \to xx$, as well its possible production from yet another new state, $\mathrm{X} \to \mathrm{aa}$. Generic decays of the form $\mathrm{a} \to \gamma\gamma$ have also been studied outside of $\mathrm{H} \to \mathrm{aa}$ decays in collider experiments [26, 27], as well as in astrophysics and cosmology [5, 28, 29], although at much lighter masses $m_{\mathrm{a}} \sim \mathrm{eV}$.

This thesis is based on two CMS results. The first of these is a technique paper [30] describing the development and validation of the end-to-end deep learning technique in the context of $\mathrm{a} \to \gamma\gamma$ decays. The second of these is a physics analysis [31] focusing on the application of this technique to perform the first direct search for $\mathrm{H} \to \mathrm{aa} \to 4\gamma$ in its most experimentally challenging but theoretically attractive regime. It will take many years to bear out the ultimate value of the end-to-end deep learning to physics searches but this work represents an important first step in demonstrating its potential. Already, entirely new tools and searches at CMS are underway that push the limits of what can be probed with the CMS detector using end-to-end deep learning at their core. The lasting significance of this work, therefore, will arguably be its demonstration of the feasibility and breakthrough potential of the end-to-end deep learning technique for physics searches.

This thesis is arranged as follows. Following this chapter which describes the motivation for the $\mathrm{H} \to \mathrm{aa} \to 4\gamma$ search, a description of the CERN LHC experimental apparatus and the CMS detector collecting the data is provided in Chap. 2. The theoretical basis of the SM, the extended Higgs sector, and the phenomenology of the $\mathrm{H} \to \mathrm{aa} \to 4\gamma$ decay are then presented in Chap. 3. In Chap. 4, we outline the analysis strategy for discriminating $\mathrm{H} \to \mathrm{aa} \to 4\gamma$ signal events. The CMS data sets used for the analysis, and the criteria employed to select $\mathrm{H} \to \mathrm{aa} \to 4\gamma$ candidate events, are detailed in Chaps. 5 and 6, respectively. Chap. 7 is dedicated to describing the training and validation of the novel end-to-end ML-based m_{Γ} regression algorithm. The main physics analysis, detailing the signal and background models used to perform the $\mathrm{H} \to \mathrm{aa} \to 4\gamma$ signal search, is given in Chap. 8. The results of the analysis are presented in Chap. 9, and our conclusions are summarized in Chap. 10.

References

1. Chatrchyan S, Khachatryan V, Sirunyan AM, Tumasyan A, Adam W, Aguilo E, Bergauer T, Dragicevic M, Ero J, Fabjan C et al (2012) Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. Phys Lett B 716:30 Sep
2. Aad G, Abajyan T, Abbott B, Abdallah J, Abdel Khalek S, Abdelalim A, Abdinov O, Aben R, Abi B, Abolins M et al (2012) Observation of a new particle in the search for the standard model Higgs boson with the atlas detector at the LHC. Phys Lett B 716:1
3. Sirunyan AM, Tumasyan A, Adam W, Ambrogi F, Asilar E, Bergauer T, Brandstetter J, Dragicevic M, Ero J, Escalante Del Valle A et al (2018) Observation of Higgs boson decay to bottom quarks. Phys Rev Lett 121

4. Aaboud M, Aad G, Abbott B, Abdinov O, Abeloos B, Abhayasinghe D, Abidi S, AbouZeid O, Abraham N, Abramowicz H et al (2018) Observation of Higgs boson production in association with a top quark pair at the LHC with the atlas detector. Phys Lett B 784:173 Sep
5. Zyla P et al (2020) Review of particle physics. PTEP 2020(8):083C01
6. Sirunyan AM, Tumasyan A, Adam W, Ambrogi F, Asilar E, Bergauer T, Brandstetter J, Dragicevic M, Ero J, Del Valle AE et al (2019) Combined measurements of Higgs boson couplings in proton-proton collisions at $\sqrt{s} = 13$ TeV. Eur Phys J C 79:421
7. Dobrescu BA, Landsberg G, Matchev KT (2001) Higgs boson decays to CP-odd scalars at the Fermilab Tevatron and beyond. Phys Rev D 63:075003 Feb
8. Curtin D, Essig R, Gori S, Jaiswal P, Katz A, Liu T, Liu Z, McKeen D, Shelton J, Strassler M et al (2014) Exotic decays of the 125 GeV Higgs boson. Phys Rev D 90
9. Curtin D, Essig R, Gori S, Shelton J (2015) Illuminating dark photons with high-energy colliders. J High Energy Phys 157
10. Peccei RD, Quinn HR (1977) CP conservation in the presence of pseudo particles. Phys Rev Lett 38:1440–1443 Jun
11. Bauer M, Neubert M, Thamm A (2017) Collider probes of axion-like particles. JHEP 12:044
12. Peccei RD (2008) The strong CP problem and axions. Springer, Berlin Heidelberg, p 3
13. Essig R et al (2013) Dark sectors and new, light, weakly-coupled particles
14. Raffelt GG (2008) Astrophysical axion bounds. Springer, Berlin Heidelberg, p 51
15. Sikivie P (2008) Axion cosmology. Springer, Berlin Heidelberg, p 19
16. Marsh DJ (2016) Axion cosmology. Phys Rep 643:1
17. Chadha-Day F, Ellis J, Marsh DJE (2021) Axion dark matter: what is it and why now?
18. ATLAS Collaboration (2012) Search for a Higgs boson decaying to four photons through light CP-odd scalar coupling using 4.9 fb^{-1} of 7 TeV pp collision data taken with ATLAS detector at the LHC
19. Khachatryan V et al (2014) Observation of the diphoton decay of the Higgs boson and measurement of its properties. Eur Phys J C 74(10):3076
20. Aaboud M et al (2019) Search for pairs of highly collimated photon-jets in pp collisions at $\sqrt{s} = 13$TeV with the ATLAS detector. Phys Rev D 99:012008 Jan
21. Sirunyan AM, Tumasyan A, Adam W, Ambrogi F, Asilar E, Bergauer T, Brandstetter J, Dragicevic M, Ero J et al (2018) Search for an exotic decay of the Higgs boson to a pair of light pseudoscalars in the final state of two muons and two τ leptons in proton-proton collisions at $\sqrt{s} = 13$ TeV. J High Energy Phys 2018
22. Sirunyan AM, Tumasyan A, Adam W, Ambrogi F, Asilar E, Bergauer T, Brandstetter J, Dragicevic M, Ero J, Escalante Del Valle A et al (2019) A search for pair production of new light bosons decaying into muons in proton-proton collisions at 13 TeV. Phys Lett B 796:131 Sep
23. Sirunyan AM, Tumasyan A, Adam W, Ambrogi F, Asilar E, Bergauer T, Brandstetter J, Dragicevic M, Ero J, Escalante Del Valle A et al (2020) Search for light pseudoscalar boson pairs produced from decays of the 125 GeV Higgs boson in final states with two muons and two nearby tracks in pp collisions at $\sqrt{s} = 13$ TeV. Phys Lett B 800:135087 Jan
24. Sirunyan AM, Tumasyan A, Adam W, Ambrogi F, Asilar E, Bergauer T, Brandstetter J, Dragicevic M, Ero J, Escalante Del Valle A et al (2019) Search for an exotic decay of the Higgs boson to a pair of light pseudoscalars in the final state with two muons and two b quarks in pp collisions at 13 TeV. Phys Lett B 795:398 Aug
25. Sirunyan AM, Tumasyan A, Adam W, Ambrogi F, Bergauer T, Dragicevic M, Ero J, Escalante Del Valle A, Fruhwirth R et al (2020) Search for a light pseudoscalar Higgs boson in the boosted $\mu\mu\tau\tau$ final state in proton-proton collisions at $\sqrt{s} = 13$ TeV. J High Energy Phys 2020
26. Knapen S, Lin T, Lou HK, Melia T (2017) Searching for axionlike particles with ultraperipheral heavy-ion collisions. Phys Rev Lett 118:171801
27. Adhikari S et al (2022) Search for photoproduction of axionlike particles at GlueX. Phys Rev D 105:052007
28. Graham PW, Irastorza IG, Lamoreaux SK, Lindner A, van Bibber KA (2015) Experimental searches for the axion and axion-like particles. Annu Rev Nucl Part Sci 65:485

29. Irastorza IG, Redondo J (2018) New experimental approaches in the search for axion-like particles. Prog Part Nucl Phys 102:89
30. Reconstruction of decays to merged photons using end-to-end deep learning with domain continuation in the CMS detector (2022)
31. Search for exotic Higgs boson decays $\mathrm{H} \to \mathcal{A}\mathcal{A} \to 4\gamma$ with events containing two merged diphotons in proton-proton collisions at $\sqrt{s} = 13$ TeV (2022)

Chapter 2
The LHC and the CMS Detector

In this chapter, we describe the experimental apparatus involved in the production, collection, and reconstruction of the particle physics data used in this analysis. The basic unit of statistically independent physics data is the collisions event, or event for short. In Sect. 2.1, we begin with a description of the Large Hadron Collider (LHC), which is the primary apparatus responsible for the production of high energy collision events. This is followed in Sect. 2.3 by a description of the Compact Muon Solenoid (CMS) detector, which is responsible for the collection of data generated by the LHC, and the main data source for this analysis. A short primer on the interaction of particles with matter is presented in Sect. 2.2, prior to the description of the CMS detector, in order for the design of the CMS detector to be better appreciated. Following these, the steps involved in the filtering and reconstruction of the detector data are described. Due to the untenable volume of data generated by the LHC, a dedicated event filtering or *triggering* system is implemented in the CMS detector, to select only events of interest, described in Sect. 2.4. For events passing the trigger, the data collected from the CMS subdetectors are subsequently reconstructed into physics objects used for analysis, as described in Sect. 2.5. Note that the reconstruction here pertains to those of standard CMS physics objects, not those reconstructed by the end-to-end technique, which is instead described in Chap. 7. Finally, as particularly relevant for the end-to-end reconstruction technique, we conclude this chapter in Sect. 2.6 with an overview of the detector simulation process and its basic validation.

2.1 The LHC

The CERN LHC is presently the largest and most energetic man-made particle collider ever built. It straddles the border of France and Switzerland, between the foothills of the Jura mountain range and Lac Léman, some 100 km underground.

The LHC, while designed to be a general purpose collider, was conceived primarily to study the nature of electroweak symmetry breaking, for which the Higgs mechanism was thought to be responsible. Today, it is chiefly known for its discov-

M. Andrews, *Search for Exotic Higgs Boson Decays to Merged Diphotons*,
Springer Theses, https://doi.org/10.1007/978-3-031-25091-0_2

ery of the Higgs boson, jointly discovered by the CMS and ATLAS experiments in their Run I (2011–2012) data sets, for which the 2013 Nobel prize in physics was awarded to the duo of Francois Englert and Peter Higgs. The LHC remains the only operational collider able to probe the electroweak energy regime and thus continues to host a broad research program investigating both high-precision, high-energy SM physics, as well as searches for physics beyond the standard model.

In this section, we describe the design choices that motivated the LHC's construction, detail its basic operation, and highlight key features that drive its physics performance.

Collider Design. The LHC, at its most basic level, is a synchrotron accelerator that accelerates beams of charged particles in a circular orbit. In the case of the LHC, there are two counter-rotating beams of protons, which, at pre-determined points in the orbit, are steered into collision, from which particle collisions are generated.

As opposed to linear accelerators, colliders based on circular accelerators have the distinct advantage of having much higher collision rates. At an energy of 6.5 TeV per beam, each proton in the beam orbits the 27 km circumference of the LHC ring at a rate of more than 11 kHz, orders-of-magnitude higher than would be achievable with linear accelerators that would need to be time-consumingly refilled after every collision. As a result, the LHC has one of the highest nominal collision rates of any collider, 40 MHz, placing it in a unique position to probe the rarest of physics decays.

As opposed to striking fixed targets, by introducing two counter-rotating beams, the LHC is additionally able to maximize collision energy. For a given beam energy, the collision energy, parametrized by the Mandelstam variable $\sqrt{s}$, is maximized when the incident particles collide in their center-of-mass frame. By utilizing two counter-rotating beams of similar mass and energy, the physics potential of the LHC beams is therefore maximized in the lab frame. As a result, the LHC is also the most energetic collider ever built, with a $\sqrt{s} = 13\,\text{TeV}$, giving it the ability to probe the highest energy physical phenomena, or equivalently, the smallest length scales, in a laboratory setting.

A disadvantage of circular colliders, however, is that they require magnets with large bending power in order to deflect particles into a circular orbit. For an orbit radius R, a particle of charge q and momentum $|\mathbf{p}|$ requires a magnetic field of strength

$$|\mathbf{B}| = \frac{|\mathbf{p}|}{qcR}, \tag{2.1}$$

where c is the speed of light. Moreover, accelerating charged particles (as in a circular orbit) dissipate energy in the form synchrotron radiation. For a particle of mass m, this occurs at a rate proportional to

$$P \propto \frac{q^2}{m^4 R^2} \tag{2.2}$$

As the above equations suggest, these effects can be mitigated by constructing large radius accelerator rings and using heavy particles. It should come as no surprises