Measurement of $H \to b\bar{b}$ in Associated Production with the CMS Detector

by

Daniel Robert Abercrombie

B.S., Pennsylvania State University (2014)

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2020

Author . . . .
Department of Physics
June 3, 2020

Certified by . . . .
Christoph M. E. Paus
Professor of Physics
Thesis Supervisor

Accepted by . . . .
Nergis Mavalvala
Associate Department Head, Physics

Measurement of $H \to b\bar{b}$ in Associated Production with the CMS Detector

by

Daniel Robert Abercrombie

Submitted to the Department of Physics
on June 3, 2020, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

We measured $VH \to b\bar{b}$ with the CMS Detector.

Thesis Supervisor: Christoph M. E. Paus
Title: Professor of Physics

Acknowledgments

Thanks.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

One of the most curious features of physics at small scales, which will likely frustrate students for the rest of time, is that certain sequences of events only have a probability of happening. There is no guarantee that an electron and a positron approaching each other at high energies will annihilate and produce an muon and an anti-muon. However, this event might still occur at a later time at with the exact same initial conditions. Furthermore, the observation of resonances where this is more likely to happen when the electron and positron approach each other at particular speeds does not mean that a Z boson was present in a given interaction. It just means that the weak component of the electroweak force significantly increases the probability of the muonic final state, given the total energy of the initial state. The sum of probabilities from different possible field interactions with particular initial conditions is the only thing we can measure. This is also only possible when observing many events with the same initial conditions.

This point is difficult to convey concisely, so many laypeople, as well as some practicing physicists, are confused by the terminology adopted by the field. But this distinction is relevant to the topic of this work. This document presents a measurement of a cross section. Cross section is the name given to the probability of an interaction occurring. Reported cross sections can be split up to describe different contributions to final states, and they can be collated into what are called “production cross sections” which describe the probabilities of particular intermediate states "occurring" (even though intermediate states never exist in reality).

The main point is that if there exists some interesting particle, and it interacts with other particles, you can see an increased probability of certain initial states resulting in certain final states. This can teach the observer about the role of the interesting particle, without ever directly seeing it.

1.1 Measurement of the Higgs Cross Section

The purpose of the following document is to present the methods and results of measuring the strength of the coupling between the Higgs Boson and bottom quarks. In this context, the Higgs Boson makes up one of the previously mentioned intermediate states that cannot be shown as present in a given event. The cross section measurement relies on a number of physics processes that will be accounted for in this document.

To measure this coupling, the Higgs Boson must first be "produced" before measuring its coupling strength to bottom quarks. Since we are not technologically advanced enough to achieve this generation using bottom quarks directly, we use measured Higgs generation rates from normal constituents of protons. We constrain ourselves further by requiring that the Higgs is generated by associated production.

After the Higgs is generated, it can decay into a number of different particles. This work is only concerned with one kind of decay.

The math that this all relies on is presented in Chapter 2.

1.2 Motivation for the Measurement

This is Thomas Kuhn's "normal science".

Precision measurements are needed to be certain of what we think.

Precision measurements often lead to discrepancies that are explained by a fundamental shift in the model.

The Standard Model is a good one. It will not be fully replaced, but at worst expanded upon. Just like Newton's Laws are still a reasonable approximation for General Relativity, The Standard Model is a good approximation for most things we have been able to interact with so far.

The only lingering questions are Dark Matter and Dark Energy, but there is no reason to assume that precise measurements of known phenomena will not lead to an explanation.

1.3 Historic Context

First, we have The Standard Model.

Parts were proven correct by the observation of the weak bosons.

The Higgs was observed in 2013.

The Higgs decaying to $b\bar{b}$was observed in 2018. [2]

1.4 Using the CMS Detector

The CMS detector is a general purpose detector used to make many observations of conditions unattainable on Earth outside of the LHC.

It has many stationary parts, and a couple of moving ones too.

This device is described in detail in Chapter 3.

Chapter 2

Theory

Before diving into the description of the experimental apparatus, an explanation of why it is expected to work is needed. There are many textbooks that cover the Standard Model, as there are many students who study it. Much of what follows is taken from the book by Mark Thompson [7].

The Standard Model Lagrangian can be defined as a sum of Lagrangians that each describe the interactions between different fermions and bosons. Equations of motion can be extracted from a Lagrangian $\mathcal{L}$ for a particle field ϕ_i using the Euler-Lagrange equations.

$$\delta_\mu \left(\frac{\delta \mathcal{L}}{\delta(\delta_\mu \phi_i)} \right) - \frac{\delta \mathcal{L}}{\delta \phi_i} = 0 \tag{2.1}$$

In the measurement of $H \to b\bar{b}$ in associated production, many components of the Standard Model are of interest. These will be introduced as needed. First, I will give a brief explanation of Higgs field's non-zero vacuum energy, a trait that makes the Higgs one of the central keystones to Standard Model. After that, the electroweak Lagrangian will be described since the cross section of associated production depends on the coupling of the Higgs Boson to the W and Z vector bosons. The coupling of the electroweak force to fermions is also important to understand both the generation of these intermediate states and the resulting final state that the CMS detector records.

Another important factor for this work is a the decay of the Higgs boson itself into bottom quarks. This depends on the Higgs directly coupling to fermions. Finally, we will briefly consider the part of the Lagrangian describing the strong force. Since the LHC is a hadron collider, understanding of the strong force is required to extract data from LHC collisions.

2.1 The Higgs Mechanism

In both of these components of the Standard Model Lagrangian, the Higgs coupling actually gives the vector bosons and massive fermions their mass. (In this work, neutrinos can be treated as massless.) The granting of mass happens for two reasons: the Higgs field has a non-zero vacuum expectation value, and the Higgs field couples to vector boson and massive fermion fields.

The Higgs can be described as two complex scalar fields in a weak isospin doublet with a quartic potential. The Lagrangian for a free Higgs is then

$$\mathcal{L} = (\delta_\mu \phi)^\dagger (\delta^\mu \phi) - (\mu^2 (\phi^\dagger \phi) + \lambda (\phi^\dagger \phi)^2) \tag{2.2}$$

Through the virial theorem, the potential has a minimum value when

$$\phi^\dagger \phi = \frac{-\mu^2}{2\lambda} = \frac{v^2}{2} \tag{2.3}$$

This potential of the Higgs field breaks the SU(2) $\times$ U(1) symmetry of the Standard Model Lagrangian. Through this non-zero vacuum expectation value, the Higgs then has a constant influence in other parts of the Standard Model Lagrangian. For this measurement, three interactions that the Higgs makes with this influence need to be considered: the Higgs interacting with itself, the Higgs interacting with the electroweak vector bosons, and the Higgs interacting with quarks.

The first two interactions manifest in the Lagrangian when we force the SU(2) $\times$

U(1) symmetry on the Lagrangian in Equation 2.2. The derivatives must be replaced.

$$\delta_\mu \rightarrow D_\mu = \delta_\mu + i\frac{g_W}{2}\boldsymbol{\sigma} \cdot \mathbf{W}_\mu + ig'\frac{Y}{2}B_\mu \tag{2.4}$$

ϕ can be rewritten to satisfy the vacuum expectation value in the gauge that will give us the massless neutral boson known as a photon.

$$\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{2.5}$$

This leads to the following expansion for the kinetic term of the Lagrangian.

$$\begin{aligned}(D_\mu\phi)^\dagger(D^\mu\phi) =& \frac{1}{2}(\delta_\mu h)(\delta^\mu h) + \frac{1}{8}g_W^2(W_\mu^{(1)} + iW_\mu^{(2)})(W^{(1)\mu} - iW^{(2)\mu})(v+h)^2 \\ &+ \frac{1}{8}(g_W W_\mu^{(3)} - g'B_\mu)(g_W W^{((3)\mu} - g'B^\mu)(v+h)^2\end{aligned} \tag{2.6}$$

Terms that are quadratic in terms of the gauge boson fields reveal the mass of the fields. Taking $h(x) \rightarrow 0$, the terms for $W^{(1)}$ and $W^{(2)}$ are the just

$$\frac{1}{4}g_W^2 v^2 W_\mu^{(1)} W^{(1)\mu} \quad \text{and} \quad \frac{1}{4}g_W^2 v^2 W_\mu^{(2)} W^{(2)\mu},$$

giving the mass.

$$m_W = \frac{1}{2}g_W v \tag{2.7}$$

The quadratic terms for $W^{(3)}$ and B mix to give a non-diagonal mass matrix $\mathbf{M}$.

$$\frac{v^2}{8}\begin{pmatrix} W_\mu^{(3)} & B_\mu \end{pmatrix}\mathbf{M}\begin{pmatrix} W^{(3)\mu} \\ B^\mu \end{pmatrix} = \frac{v^2}{8}\begin{pmatrix} W_\mu^{(3)} & B_\mu \end{pmatrix}\begin{pmatrix} g_W^2 & -g_W g' \\ -g_W g' & g'^2 \end{pmatrix}\begin{pmatrix} W^{(3)\mu} \\ B^\mu \end{pmatrix} \tag{2.8}$$

The non-diagonal matrix allow $W^{(3)}$ and B to mix. Physical states must be represented by a diagonal Hamiltonian. Diagonalizing the term above gives masses of the physical states.

$$\frac{1}{8}v^2 \begin{pmatrix} A_\mu & Z_\mu \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & g_W^2 + g'^2 \end{pmatrix} \begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix} = \frac{1}{2} \begin{pmatrix} A_\mu & Z_\mu \end{pmatrix} \begin{pmatrix} m_A^2 & 0 \\ 0 & m_Z^2 \end{pmatrix} \begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix} \tag{2.9}$$

This gives us the masses of the neutral gauge bosons.

$$m_A = 0 \quad \text{and} \quad m_Z = \frac{1}{2}v\sqrt{g_W^2 + g'^2} \tag{2.10}$$

From the simple act of requiring SU(2) $\times$ U(1) symmetry on the Lagrangian of a scalar doublet with non-zero vacuum expectation value, the masses of all the electroweak gauge bosons have been produced. The next thing to consider is the couplings also produced by this process. The couplings will allow us to determine more precisely the parameters above by measuring cross sections.

2.2 Associated Production

$$\mathcal{L}_{EWK} \tag{2.11}$$

2.2.1 Production Mechanisms

2.2.2 Decay Channels of Vector Bosons

2.3 Characteristics of the Higgs

2.3.1 Energy Spectrum

2.3.2 Decay to $b\bar{b}$

$$\mathcal{L}_f \tag{2.12}$$

2.4 Other Relevant Standard Model Processes

Chapter 3

The CMS Detector

The Compact Muon Solenoid (CMS) detector is a complex device located at the LHC. Ideally, a lot of time could be saved for the reader by only detailing the parts of the detector that are relevant for the analysis presented in this work. Unfortunately, a characteristic of contemporary collider measurements is that almost all sub-systems of the detector are used to make the final measurement.

A brief overview of all the detector subsystems are therefore presented in this chapter. A curious reader can learn more about the design and motivations for said design in the TDR [1].

3.1 Associated Production at the LHC

Before describing the devices that are used to observe and record events, the method of generating interesting events must be described. The CMS detector is located at the Large Hadron Collider (LHC). Described in detail in multiple publications [4], a brief description is given here.

The LHC, with a circumference of 26.7 km, is large enough to be considered located in multiple towns and countries, but it will suffice to say it is near Geneva, Switzerland at the European Organization for Nuclear Research (CERN), the main campus of which is addressed in Meyrin, Switzerland. This campus itself also spans the border between Switzerland and France.

Quark interactions can lead to Vector boson production, since the protons are at high energy.

Virtual vector bosons can radiate a Higgs.

This happens along with additional event stuff, caused by breaking apart the protons and pileup.

3.2 Observables of Associated Production

We can count charged leptons or infer neutrinos to tag vector bosons.

We are interested in Higgs to $b\bar{b}$.

3.3 Detector Requirements

We need to identify interesting particles, be able to reconstruct missing transverse momentum, and remove pileup.

3.3.1 Particle Types and Energies

3.3.2 Pileup Conditions at the LHC

3.4 Detector Design

3.4.1 Silicon Pixel Detector

3.4.2 ECAL

3.4.3 HCAL

3.4.4 Muon Chambers

3.5 Detector Performance

3.5.1 Test Beam Performance

3.5.2 Trigger

3.5.3 Online Calibration

3.6 Data Format

3.6.1 Event Reconstruction

3.6.2 Offline Calibration

3.7 Accessing Data

Chapter 4

Simulation

4.1 Backgrounds to the Analysis

In order to effectively measure Higgs production, we need to be able to accurately estimate other events that end up in our selection.

4.2 Event Generation

We use different generators.

Details in Appendix D.

4.2.1 Tree Level Simulation

4.2.2 Parton Showers

4.3 Detector Simulation

4.4 Corrections to Simulation

4.4.1 Smearing

Muons

Electrons

Jets

4.4.2 Selection Efficiencies

4.4.3 Control Regions

Light Flavor Jets

Heavy Flavor Jets

$t\bar{t}$

Chapter 5

Event Selection

5.1 Object Definitions

Section 3.6.1 describes how detector responses are linked to possible physical particles. We want to remove false positives, so here are some tighter requirements for counting for event selection.

Once we have our objects defined, we can count them.

5.1.1 Muons

Muons can show up in weakly decaying jets, so we're only interested in isolated ones here.

5.1.2 Electrons

Electrons do the same thing as muons, but messier because the ECAL isn't as clean as the muon chambers.

5.1.3 Jets

Jets are messier still.

Pileup removal is a big deal here.

5.1.4 MET

MET is corrected.

5.1.5 Undesirable Particles

There are certain particles that we do not want present. We make very loose selections for those and veto on them.

Photons

Tau Leptons

5.2 Removal of QCD

We have some cuts across the board on our objects in order to remove events that are just QCD.

5.3 Categories of Vector Boson Decay

Now that we are ready to count, we can count leptons in order to characterise potential vector bosons.

5.3.1 0 Leptons

5.3.2 1 Lepton

5.3.3 2 Leptons

5.4 Topology of Higgs Decay

5.4.1 Resolved Jets

We reconstruct two b jets.

5.4.2 Boosted Jet

When the Higgs has very high p_T, the jet clustering algorithms can find both daughter particles as being part of a single jet.

Chapter 6

Analysis Results

6.1 Systematic Uncertainties

6.2 Combination Fit Method

6.3 Results

Chapter 7

Conclusions

Appendix A

Detector Projects

Each collaborator must contribute to the operation of the CMS detector before his or her name is added to the author list. The operation of the detector is distinct from analyzing the data generated by the detector, so all collaborators must adopt some role outside of being a physicist.

This appendix details projects I completed in order to contribute to the operation of the CMS detector. The first project presented is the Dynamo Consistency project. It is a plugin for the dynamic data management system Dynamo [5] that compares the inventory of files Dynamo expects at a site with the files that are actually at a site. The other project described is known as Workflow Web Tools. This is a dynamic web server that displays errors reported by the CMS computing infrastructure to operators, and allows those operators to perform corrective actions through the web page. Workflow Web Tools also tracks operator actions for future use in training various machine learning models. Both projects produced software packages written in Python [8, 9] and available through the Python Package Index (PyPI) as `dynamo-consistency` and `workflowwebtools`.

A.1 Dynamo Consistency

[3] [6]

A.2 Workflow Web Tools

Appendix B

Physics Calculations

Appendix C

Data Format

Appendix D

Generator Parameters

Appendix E

Data Card

Bibliography

[1] The CMS Collaboration. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software.* Technical Design Report CMS. CERN, Geneva, 2006.

[2] The CMS Collaboration. Observation of higgs boson decay to bottom quarks. *Physical Review Letters*, 121(12), Sep 2018.

[3] Alvise Dorigo, Peter Elmer, Fabrizio Furano, and Andrew Hanushevsky. Xrootd-a highly scalable architecture for data access. *WSEAS Transactions on Computers*, 1(4.3):348–353, 2005.

[4] Lyndon Evans and Philip Bryant. LHC machine. *Journal of Instrumentation*, 3(08):S08001–S08001, aug 2008.

[5] Yutaro Iiyama, Benedikt Maier, Daniel Abercrombie, Maxim Goncharov, and Christoph Paus. Dynamo – handling scientific data across sites and storage media, 2020.

[6] Erwin Laure, A Edlund, F Pacini, P Buncic, M Barroso, A Di Meglio, F Prelz, A Frohner, O Mulmo, A Krenek, et al. Programming the grid with glite. Technical report, CERN, Geneva, Switzerland, 2006.

[7] Mark Thomson. *Modern particle physics.* Cambridge University Press, Cambridge, 2013.

[8] Guido Van Rossum and Fred L. Drake. *Python 3 Reference Manual.* CreateSpace, Scotts Valley, CA, 2009.

[9] Guido Van Rossum and Fred L Drake Jr. *Python reference manual.* Centrum voor Wiskunde en Informatica Amsterdam, 1995.